SEPTEMBER 2012

Training working memory and fluid intelligence in older adults:

Developing measures and exploring outcomes

Sinéad Hynes

This thesis is submitted for the degree of Doctor of Philosophy at the University of Cambridge.

Acknowledgements

Firstly I would like to thank my principal supervisor Dr. Tom Manly and my second supervisor Dr. Jessica Fish for their guidance and support during the research process. I am thankful for their patience and help throughout the course of my PhD. My thanks also go to staff and researchers at the MRC Cognition and Brain Sciences Unit. I feel privileged to have had the opportunity to complete my studies at the CBU. Thanks especially go to a few people in particular. To Dr. Adam Hampshire for his generous help with developing, programming and providing the training tasks used in chapters 4-7. To Prof. Barbara Wilson who allowed me to work with her client S.M. – her passion for her work is extremely motivating for me to see as a student. To Dr. Russell Thompson for providing the Feature Match task used in chapter two and for his encouraging advice over cups of tea in the kitchen. And my thanks go to two visiting students to the unit Veronica Montani and Sarah Griffiths who helped me with the video ratings in chapter three. Without their help this chapter would not have been possible.

Financial support for my PhD was provided by the National University of Ireland and the Cambridge European Trust. My sincere gratitude to these committees for funding me.

My thanks to Prof. Agnes Shiel in NUIG– “for the use of the parish hall” over the summer months and for all the chats and encouraging words late into the evenings.

To my roastin’ woolly jumper-wearing mucksavage lab-mates Theresa, Davy and Oriel who made my time at the unit sick, in the best possible way! I never expected to find such good friends in such a small room. Without the support that I had from you guys as well as the baking, singing, emails and advice sessions I can confidently say that this thesis would not be completed. I feel truly blessed to have made such good friends in you and a gigantic thanks and loads of love for making my time at the unit such a fun, bright and happy time.

My final year in Cambridge was one of the best for one big reason – 149 Catherine St. I was so lucky to find the best housemates ever in Dee, Laura and Maeve. All the cups of tea, bouncing castles, punch, BBQs and craic will be with me forever as some of the best memories of my years in Cambridge. Thanks girlies. Thanks to all my friends and to Gear and Eoghan; Theo (for all the match reports); and Adam.

And finally thanks millions to my mama and dada. Ye’ve been mighty, as I hope ye know. All of your help has gotten me where I am and I love ye big-bits.

So if I may, I would like to dedicate this thesis to my Bu-Ti-Full Mama and Da.

“Briseann an dúchas trí shúile an chait”, mar a deir siad. Tá súil agam gur fíor seo!

Preface

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text.

Dissertation summary

Name: *Sinéad Hynes* Title: *Training Working Memory and Fluid Intelligence in Older Adults: Developing measures and exploring outcomes*

This thesis investigates computerised cognitive training in older adults, with a focus on training working memory and fluid intelligence. A series of studies is reported, with two broad aims. The first was to develop and validate outcome measures appropriate for use in this population, and the second was to examine whether established gains in cognitive functioning generalised to everyday life.

In relation to the first aim, two studies were conducted which concerned the development of a sensitive measure of organisational abilities within a computerised paradigm, the Games Evaluation Task (GET). A further study made use of an existing naturalistic measure, the Multiple Errands Task (MET, Shallice & Burgess, 1991), and investigated whether it was possible to obtain reliable ratings of performance on the basis of video footage taken from the participant's perspective by means of a body-worn camera. Both the GET and MET were used as outcome measures in the subsequent training studies.

In relation to the second aim, three studies of cognitive training are reported. The first is a case study of a man with problems in working memory and time perception following a stroke. He underwent training on an intensive working memory package within a single-case experimental design that incorporated an active control condition. This approach was then extended in a larger sample of healthy older adults, who trained intensively on tasks that focussed either on working memory, or fluid intelligence and problem solving. In the final study participants trained on a combination of both working memory and fluid intelligence tasks. In addition, they watched training videos that focused on teaching various cognitive strategies. The aim of this video supplement was to help participants draw links between the computer training and real-life situations, and hence to foster generalisation of any benefits to everyday life.

The thesis concludes with a general discussion which examines the major findings of the studies presented, their clinical applications, the limitations of the research and possible future directions.

Table of Contents

List of Tables

List of Figures

Chapter One:

Introduction

This thesis investigates whether computerised cognitive training - specifically targeting working memory (WM) and fluid intelligence – can improve cognitive capacity and daily functioning. As part of this investigation chapters two and three report the development and validation of outcome measures appropriate for use with older adults. This introductory chapter focuses on working memory training (WMT) as well as strategy-based training and training targeting fluid intelligence.

Introduction

Limited capacity and working memory

The human brain is capable of processing vast amounts of sensory information (e.g. colours, textures, surfaces, edges, movement) in parallel. In other respects it behaves like a serial, capacity limited system, at a first approximation, only able to focus on one thought or object, or produce responses to one task at a time. Frameworks reflecting this differences include contrasting 'automatic' and 'controlled' processes (Schneider & Shiffrin, 1977) and concepts of attentional selection (Broadbent, 1971). From studies examining the effects of tasks occurring in different modalities Baddeley and Hitch (1974) developed a simple and hugely influential model of these capacity limitations in human thought. Adopting the term Working Memory (WM) to differentiate capacity-limited on-line processing from the vast capacity of memory in general, this was characterised as "a brain system that provides temporary storage and manipulation of the information necessary for such complex tasks as language comprehension, learning and reasoning…[and where] Working memory has been found to require the simultaneous storage and processing of information." (Baddeley, 1992; pp556). There are two distinct modality specific stores are dedicated to phonological ('phonological loop') and visuo-spatial ('visuo-spatial scratchpad') information respectively. These slave systems however required some common processes (i.e. that could lead to interference between phonological and visuo-spatial tasks) which they termed the Central Executive.

Whilst there is often emphasis on WM as a short-term store, reflecting the serial recall tasks that are often used to assess it, Baddeley and Hitch's conceptualisation went well beyond this with WM being essentially synonymous with the capacity for current conscious thought. In this it has common ground with other views of 'executive' function such as Norman and Shallice's Supervisory Attentional System (Shallice 1982). The original Baddeley and Hitch (1974) WM model has remained remarkably resilient to almost three decades of research (in 2000 Baddeley updated this to include an "episodic buffer", a temporary interface between the slave systems and long term memory, but this does not materially affect the general point here). It is clear from this description that WM capacity is likely to be relevant to real-life activities and skills (see Cohen & Conway, 2008) and that any problems with WM may have wide-ranging implications. These could include, for example, making it difficult for a person to keep active the various components of long sentences, problems keeping relevant information 'in mind' to make informed decisions, and difficulty keeping track of a plan. In turn it would be expected that this would have a detrimental effect on education, occupation and social function. Evidence of this has been reported in developmental and neurological populations with compromised WM (Gathercole et al., 2005; Robertson & Murre, 1999). WM impairments have been associated with multiple sclerosis, Alzheimer's disease and related dementias (Baddeley, Logie, Bressie, Della Sala & Spinnler, 1986 ; Kensinger, Shearer & Locascio, 2003; Pelosi, Geesken, Holly, Hayward & Blumhardt, 1997; Robertson & Murre, 1999), as well as being a feature of emotional disorders including depression (Castaneda, Tuulio-Henriksson, Marttunen, Lonngvist & Suvisaari, 2008) and acute Post-traumatic Stress Disorder (Lagarde, Doyon & Brunet, 2010). As is addressed further in subsequent sections, if it were possible to increase WM capacity via training, benefits across many functional activities could follow.

Limited capacity and general intelligence

Just as WM can be seen to be a central capacity relevant to many different sorts of cognitive tasks, so too can general intelligence. In its empirical form, evidence for this concept emerged from observed positive correlations between a wide variety of tests; people who tended to do well on one sort of test also tended to do well on another even if its demands were ostensibly very different. It was therefore possible to separate variance due to a common factor (general intelligence or *g*) from demands specific to the individual tests (Spearman, 1904). Much debate has followed on the process or processes that may underpin *g*. As with WM, if it were possible to increase this capacity then, by definition, benefits could potentially generalise to many domains of function. It is important to note however that training WM or *g* are neither entirely independent nor synonymous. One method of measuring general intelligence ("intelligence quotient" IQ) that follows from Spearman's observations is to administer a wide range of different measures and take the individual's average level as reflecting *g*. Such batteries often include measures with a high WM demand (e.g. Wechsler, 1981). Another method is to administer a single test that has a particularly high loading on *g* (Duncan, Burgess & Emslie, 1995). As discussed further below, these novel problem solving tests such as the Raven's Progressive Matrices, Cattell Culture Fair and Matrix Reasoning have a high, but not exclusive, demand on WM.

Multiple demand regions: An overlap between WM and *g*?

Research at the neural level also suggests scope for overlap between WM and *g* and, by inference, the effects of WM or g training. A meta-analysis of functional imaging studies showed that increased demands across a diverse range of tasks was associated with common increased recruitment of a fronto-parietal network, termed descriptively the "Multiple Demand" (MD) network (Duncan and Owen, 2000). This echoed single cell recording studies that have showed a remarkable capacity for a large proportion of dorsolateral frontal neurons to adaptively respond to whatever is relevant to the task at hand, and to change this tuning as the task changes (Roa, Rainer, & Miller, 1997; Sigala, Kusunoki, Nimmo-Smith, Gaffan & Duncan, 2008). Increased difficulty in WM tasks *and* novel problem solving tasks are associated with increased MD activity, providing a potentially common neural mechanism for either WM or g training to produce generalised benefits. An interesting question is whether WM or other form of training may be the best way to achieve these effects.

It is important at this point to distinguish between “crystallised” and “fluid” intelligence and state that the latter is likely to be of greatest importance to WM or *g* training, including in neuropsychological rehabilitation. Put simply, crystallised intelligence broadly refers to the fruits of previous intelligent activity and learning. People’s general knowledge or knowledge of how to pronounce irregularly spelled words, for example, are highly correlated with g and novel problem solving high g measures in the young, healthy population. With aging or neurological insult, however, these relationships can decline. People may remember that Neil Armstrong was the first man on the Moon or that “gaol” is pronounced like “jail” but show marked reductions in (new) novel problem solving. For this reason, general knowledge and irregular word pronunciation are used as ‘hold’ tests to estimate people’s likely level of general intelligence before a brain injury (e.g. Nelson, 1991; Wechsler, 1981).

In support of a particular role of the MD prefrontal cortex (PFC) in fluid intelligence Duncan et al. (1995), tested three well-educated patients with frontal lobe lesions and reported a significant discrepancy in fluid compared with crystallised intelligence scores that was absent in five patients with damage to the posterior areas of the cortex. Duncan, Emslie, Williams, Johnson and Freer (1996) also showed how, in the normal population, low g scores were associated with a heightened tendency to ignore a stated intention – a feature that is also prevalent in those with frontal damage (see chapter 2 for more detail). The work of Duncan, Johnson, Swales and Freer (1997) and Roca et al. (2010) indeed suggests that much of the variance in conventional ‘executive’ tests long held to be sensitive to frontal lobe damage can be fully explained by changes in fluid intelligence measures.

In summary there are capacity limitations in normal human cognition. Individual differences in these limitations, including as a result of brain injury, are likely to affect how well people cope with a range of everyday demands. The key issue for this thesis is whether and how these capacities can be trained and whether this produces generalised gains in everyday function.

Cognitive training

People who undertake particular sports training tend to get better at that sport and increase their general levels of fitness; the training has a specific and generalised effect. In a similar way it can be argued that schoolwork teaches specific knowledge but also more generalised study skills, stamina, concentration and so on. There are grounds to hope therefore that if particular cognitive capacities can be trained that this would be to the benefit of a range of everyday activities that require the capacity. Such training could be of benefit to many in improving scholastic and occupational achievement, efficiency and so on. The greatest benefit could however be for people who are disadvantaged with developmental or acquired cognitive deficits.

In neuropsychological rehabilitation the benefits of repetitive training have, however, generally proved elusive. Repeatedly learning list of words improves memory for those words but is of minimal benefit when faced with a new list (Glisky, Schacter, & Tulving, 1986). In the realm of attention (such as sustaining or dividing attention) there is some evidence that gains on the trained tasks can generalise to similar but untrained tasks (‘near transfer’) but evidence on ‘far transfer’ to everyday life is scant (Gray, Robertson, Pentland, & Anderson, 1992; Sturm, Willmes, Orgass, & Hartje, 1997). A similar picture is apparent in training people to remember to do something at a future point (prospective memory; Raskin & Sohlberg, 2009). The lack of generalisation from a particular training approach does not, of course, mean that no training could be effective. In addition, measuring transfer from simple controlled experimental conditions to noisy, complex everyday situations is difficult. Often this relies on self- or other report questionnaires or very broad indices of improvement such as occupational status (i.e. generalisation may have occurred but not have been detected). Finally, the last 20 years have seen remarkable changes in the availability of personal computers and connection to the internet that mean qualitatively different forms of training are now possible.

Computers are extremely good at providing repetitive training over long durations that can remember adapt to the user’s skill level. It is only relatively recently that programmers who are affordable by (or who are) cognitive researchers could produce attractive game-like adaptive interfaces. The main capacity reported to have been the subject of this kind of training over recent years is WM training (WMT). Before turning to that literature, alternative approaches to remediating the effects of ‘executive’ impairments are briefly reviewed to place the WMT literature in context. The term “Strategy based rehabilitation” is used to distinguish these approaches from WMT. However, as is discussed further, it is neither clear that WMT acts via increased capacity (rather than the development of strategy) nor that ‘strategy’ based approaches do not lead to changes in capacity and plastic reorganisation. This review also gives emphasis to studies related to the maintenance of cognitive function in the healthy aging population as work with this group forms a substantial component of this thesis.

Strategy-based rehabilitation for executive impairment

Historically damage to the frontal cortex has been linked with a range of impairments including marked changes in personality, loss of creative problem solving skills (with preserved ability to function along habitual lines), mental inflexibility, problems forming/following realistic plans, distractibility and impulse control (e.g. Eslinger, Gratten & Geder, 1995; Fuster, 1989; Goldstein, 1942; Luria, 1966; Shallice & Burgess, 1991; Stuss & Benson, 1983). Unified accounts of these diverse problems have generally been in terms of damage to high-level processes required to produce coherent, goal directed action (e.g. Shallice, 1986) akin to the Central Executive within the Baddeley and Hitch (1974) WM model. In line with most neuropsychological impairments being named according to their functional consequences (amnesia, prospopagnosia) rather than their putative ‘location’ within the brain (‘temporal lobe syndrome’, ‘fusiform syndrome’) there has been a migration from discussing “frontal” symptoms to the language of executive impairment.

Whilst the nature of executive impairments have been known for over a century, systematic reports of attempts to remediate the consequences of these difficulties are a relatively recent development. von-Cramon et al. (1991) addressed the question of whether rash, impulsive and inflexible problem-solving styles acquired through brain injury could be replaced by a more measured step-by-step approach. Participants in their Problem Solving Training group were asked to solve various sorts of novel problems (e.g. what information was most relevant to a word-limited for-sale advert, how to solve a whodunnit novel) and trained in steps. These included first defining and encoding the problem (e.g. re-read and summarise instructions in your own words), thinking about what was relevant and irrelevant information, generating as many solutions as possible, evaluating the pros and cons of each, forming a plan, executing the plan, checking back to see if it is working and so on.

This lengthy training was associated with improved novel problem solving compared with a ‘memory training’ placebo control although the extent to which these techniques generalised to patients’ everyday lives remained unknown.

Influenced by the von Cramon study, Levine et al. (2000) developed a psycho-educational approach designed to offset the negative consequences of executive impairment called Goal Management Training (GMT). In their study of 30 patients with traumatic brain injury Levine et al. (2000) compared goal management training to the placebo control of motor skills training. In the motor skills training sessions participants completed activities such as reading and tracing motor-reversed words, while the GMT focused on the five stages of GMT training – “Stop; Define main task; List steps; Learn steps; Check”. In this initial proof-of-concept study, both groups were trained over two 4-6 hour sessions. Only the GMT group showed gains on untrained problem solving tests, slowing down their performance to offset errors. No examination of the generalisation of these strategies was undertaken in this study but was addressed to some extent in a case study described in the same paper (Levine et al., 2000). GMT principles were used to teach a participant with organisational problems following encephalitis meal preparation. Performance was improved as measured by observation and self-report measures.

As part of a series of five papers on cognitive rehabilitation in the elderly, using a multiple baseline cross-over design as introduced by Stuss et al. (2007), Levine et al. (2007) describe a modified version of GMT that they used with 49 adults aged between 71 and 87 years. They used interactive small group sessions with participants that were related to real-life events and addressed difficulties that participants may be experiencing that could be addressed using the principles of GMT such as stopping, “present mindness”, prioritising, breaking down complex tasks, and using a mental blackboard. They used paper and pencil and desktop exercises to apply the constructs that were taught during sessions. Following four weeks of training and at follow-up gains in performance were reported and maintained on the table-top simulated real-life tasks that served as outcome measures. The gains were in line with those reported by Levine et al. (2000) and, as with that study, did not investigate generalisation to everyday life.

Winocur, Palmer, Dawson, Binns, Bridges and Stuss (2007a) looked at the psychological functioning of this group of older participants with the aim of improving overall well-being and coping. They integrated aspects of the two previous training modules by including information on aspects of GMT such as organising and breaking tasks down and also emphasising the relationship between psychosocial wellbeing and memory. The authors made an effort to relate all that was discussed in the sessions, such as overstating problems and society expectations, to home life. Participants were encouraged to keep logs and set goals for themselves. Improvements were reported on self-reported quality of life, happiness, well-being and coping, which were observed at follow-up but there was a failure to generalise to the real-life difficulties reported by participants.

In a larger scale study using an intervention based on GMT and von Cramon et al.'s Problem Solving Therapy, Spikman, Boelen, Lamberts, Brouwer and Fasotti (2010) conducted a trial in which 75 ABI patients were randomized to either GMT/PST or a control condition involving computerized training of reaction time, attention and planning. The study included 6-month follow-up, blind assessment and a primary outcome measure defined by everyday function rather than test performance. Despite the active control condition, GMT/PST was associated with greater gains in areas like work and leisure involvement and was the only condition linked with post-training continued improvement. Similar benefits were observed on a measure of participants' ability to set and accomplish realistic goals, and the planning of real-life tasks.

Positive results were also reported by Miotto, Evans, de Lucia and Scaff (2009) from an intervention also informed by Problem Solving Therapy and GMT, developed at the Oliver Zangwill Centre as the Attention and Problem Solving (APS) group approach. Thirty participants, including people with focal frontal injuries and TBI, took part and were allocated to three groups; Attention and Problem-solving Group (APS), a similar duration of receiving information and education on brain injury, or treatment as usual. Initial assessment included a version of (Shallice & Burgess, 1991) Multiple Errands Task. As is discussed more extensively in subsequent chapters of this thesis, the gave participants a series of goals that needed to be completed within a particular time, and according to certain rules, in a shopping centre. Raters recorded which of the tasks were achieved, the efficiency of the strategy and instances of rule breaks. The APS groups were run in a weekly 90 minute session for 10 weeks. The results showed that APS was associated with significantly greater improvements on the Multiple Errands measure and a decline in reported difficulty compared with the control groups. In the design, both control groups then crossed-over into receiving APS and broadly showed similar gains, strengthening the attribution of the improvement specifically to APS. No significant further improvements were observed at 6-months follow-up but the gains relative to baseline were well maintained.

Jobe et al. (2001) describe a study of 2832 older adults aged between 65 and 94 years - the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE). The results of this multi-site training which consisted of ten 60 - 75 minute sessions given over 5- 6 weeks was reported by Ball et al. (2002). Small-group teaching (~5 people) or individual sessions were used. There were three groups in the study – memory, reasoning, and speed of processing. Following training there was improvement reported in everyday performance from the three groups although the training effects were not as large as in other studies of its like and were not maintained at two years follow-up (Ball et al., 2002). The measures of everyday life included financial management, food preparation, medication use and driving. Near measures in all groups showed improvements. At five-year follow-up (Willis et al., 2006) found that the positive effects of the three trainings were maintained.

A group of papers on the SIMA project, first introduced by Oswald, Rupprecht, Gunzelmann and Tritt (1996) report cognitive and psychomotor training in older adults. The aim of this multidisciplinary project was to maintain independent living in older adults. In the most recent study Oswald, Gunzelmann, Rupprecht and Hagen (2006) examined this training in 375 older adults and gave a five year perspective of the SIMA study. They examined cognitive training, physical training and psychoeducational training. Cognitive training focused on fluid intelligence, attention and memory and involved visual search tasks and speed of information training, as well as training on strategies such as verbal elaboration and visual imagery to remember names and were shown various different memory aids to help with everyday memory slips. Practice at home was encouraged throughout. Physical training was a mixture of various gymnastic exercises and games involving balance training, perceptual and motor co-ordination. Psychoeducational training consisted of information sessions on compensatory aids and appliances that might be helpful, information on nutrition, falls prevention and financial assistance, among other topics. It aimed at strengthening coping responses in participants in relation to everyday demands. A battery of various cognitive, physical, mood, and everyday life measures were used. As with the earlier study (Oswald et al., 1996) the combination of cognitive training and physical training had the most positive effect on cognitive, physical and emotional domains and the gains were best maintained in this group at follow-up, although not all measures reached statistical significance. These effects are similar to that in the ACTIVE study where long term training effects were also maintained, although improvements were not reported in the same number as domains for example health status and physical function. The principles from this study were applied to a nursing home dwelling community of 294 participants. They applied a programme of cognitive activity and physical activity and they found improvements in activities of daily living such as mobility, eating and dressing. Significant improvements were seen also in cognitive domains and the training was well-accepted by staff members and participants in the nursing home.

This section has reviewed evidence of a range of educational/strategy application interventions designed to offset the consequences of executive impairment in neurological populations or to help preserve cognitive function, including organisation, in the context of normal aging. As would be expected from such complex interventions and diverse groups, the results are mixed. However, overall there is reasonable proof-of-concept that these approaches can be associated with immediate gains on untrained outcome measures (although the degree to which these truly differ from the trained activities remains an issue). The generalisation of the training gains to everyday life, which is crucial to clinical benefit but also extremely difficult to assess, has been less frequently addressed but there are some positive indications in this respect (particularly from Spikman et al., 2011 and Miotto et al., 2010). There are important questions about which components within these approaches give rise to benefits and the nature of the changes that occur. For example, if the step-by-step, measured stance towards problem solving that is initially effortful and conscientiously applied comes, through use, to be a more automatic, ‘habitual’ part of an individual’s disposition, does it remain ‘compensatory’?

An interesting question, to which this thesis is partly applied, is whether training that is more ostensibly about developing cognitive capacity could form a useful component within these broader, educational approaches. It is timely at this stage, therefore, to turn to the best evaluated of these techniques, working memory training.

Working Memory Training

Studies of working memory training in neurologically healthy adults

One of the best known instances of ‘working memory training’ was described by Chase and Ericsson (1978, reported in Chase, Lyon & Ericsson, 1981). They trained a young male student, SF, on digit span tasks (i.e. where the participant hears a series of digits and is tasked with repeating them back in the same order), for a period of 20 months, with an hour’s practice 3-5 days per week (i.e up to 400 hours). Remarkably, S.F.’s digit span increased from the normal level of 7 to the decidedly abnormal level of almost 80 digits repeated in their exact sequence! Despite this his performance on the very similar letter-span task, on which he had not trained, was unchanged. In accounting for this highly specific effect, SF, a long distance runner had developed strategies only applicable to digits; he chunked the numbers into groups of 3-5, associating each with a running time, memorable ages and dates. With further practice he was able to build up these chunked sequences using a subgroup-supergroup structure. Presumably, had he spent a similar amount of time training on letter span, he could have developed similar highly specific strategies

Much of the literature on WMT has focused on effects in young undergraduates and there have been several reports of beneficial effects of computerised cognitive training packages, including the all-important evidence of generalisation to untrained tasks. Dahlin, Stigsdotter Neely, Larsson, Bäckman and Nyberg (2008) for example, used functional Magnetic Resonance Imagining (fMRI) to examine the patterns of neural activation in a group of neurologically healthy young adults performing three types of cognitive task before and after training. The first type of task was a so-called “updating” task, in which a series of letters was presented, with the participant having to recall the four most recently presented letters. The task progressed in difficulty by increasing the length of the letter series. This task was also used for training, although the training also made use of parallel versions of the task (where letters were substituted for digits, spatial locations or shapes). An untrained n-back task formed a near-transfer measure. Here participants had to determine whether a given stimulus was the same as the stimulus presented n trials earlier in the series. Finally, far transfer was assessed on a version of the Stroop task. The Stroop task could be said to engage attentional/central executive processes, though ones of a qualitatively distinct nature than those engaged by the updating and n-back tasks. Fifteen participants underwent 15 training sessions each lasting 45 minutes, over a period of five weeks. Seven additional control participants underwent fMRI scanning at baseline and five-weeks later, without any training. Analysis of the behavioural data revealed a significant interaction between time and training group, such that the improvement in task performance was superior in the training group compared to the control group, on both the letter-updating and n-back tasks. No effect was found for the Stroop. The authors report that, relative to baseline, the n-back and updating tasks showed common areas of fMRI increased blood oxygen level dependent (BOLD) signal whilst the Stroop had a different signature. Training was associated with increased striatal and decreased frontal activation during the Letter Memory task, and increased frontal and striatal activation in the n-back task. The authors suggest therefore that training-related behavioural improvements were mediated by the striatum. A subsequent study in healthy older adults, a participant group more relevant for examining potential strategies for rehabilitation, however, showed comparable effects of training on trained measures, but no transfer to the n-back task.

Jaeggi, Buschkuehl, Jonidas and Perrig (2008) used a similar training procedure in a study where 35 participants were trained on a dual n-back task for a period of either 8, 12, 17, or 19 days, for 25 minutes per day. A separate group of 34 participants were tested at the same intervals without training. The dual n-back task involved two series of stimuli being presented simultaneously, one auditory and visuo-spatial. The participant had to confirm whether or not each stimulus was the same as that presented n trials previously, with n increasing as participants' accuracy increased. The striking finding in this study was that those who had undergone training on the dual n-back task showed significant improvements on tests of fluid intelligence (Raven's Advanced Progressive Matrices and another similar test) in comparison with untrained controls. There was a significant linear relationship between the duration of training and the size of the gain in measures of fluid intelligence, even though these were relatively short training durations (between 8 and 19 days). This finding is of particular interest firstly because the tests of fluid intelligence were not trivially similar to the tasks used in training, and secondly fluid intelligence has been thought to be a capacity relatively impervious to change via training (Sternberg,1999). A criticism of these findings, discussed in Redick et al. (2012), was that the authors collapsed the data across all four studies – the groups that received 8, 12, 17 or 19 sessions – despite differences in procedure including in the choice and time-limits of outcome measures and the different patterns of transfer (see Redick et al., 2012 for full discussion).

Later studies were conducted by Jaeggi, Studer-Luethi, Buschkuehl, Yi-Fen, Jonides and Perrig (2010). In two different experiments with 104 and 99 university students, they showed little if any difference between training on a dual n-back tasks and training on a single n-back task over four weeks. Compared with untrained controls, both types of WMT were again linked with gains on fluid IQ.

A dual n-back paradigm was used by Schweizer, Hampshire and Dagleish (2011). Their interest was particularly in examining the potential effects of WMT on cognitive control of emotional material/responses. Forty-five post-graduates trained on WM tasks that were either emotional or neutral, or a control non-WM procedure. The WMT was based on Jaeggi et al. (2008). In the emotional condition the stimuli included emotive faces and words with strong valence (e.g. rape). The authors found gains from emotional and neutral WMT that generalised to an untrained WM measure and to the speed of completion of a fluid IQ task. Only the emotional WMT was linked with gains on an emotional Stroop outcome task. It would be interesting to see, given the results of Jaeggi et al (2010), if the same results would be seen if a single n-back condition had been used. This study used a small, highly-educated sample but the findings are promising in demonstrating the added benefits in the clinically important area of affective control.

Owen et al. (2010) recently conducted a very large-scale study of “brain training” which included a significant WMT component. More than 11000 volunteers opted in to the online study that was a tie-in with a popular science TV programme (“Bang

Goes the Theory", BBC, 2009). Participants aged 18-60, who were randomised to 3 groups, were first directed to a website to perform a series of outcome measures. These included measures of reasoning, verbal short-term memory, spatial working memory and paired-associates learning. The training conditions were a general 'brain training' group, a reasoning training group and a control group. The control group was asked to use the internet to find answers to obscure general knowledge questions. The reasoning training group were trained on 6 adaptive planning, reasoning and problem solving tasks. The general brain training group trained on 6 tests of memory, attention, visuo-spatial processing and arithmetic calculations designed to replicate those commonly used in commercial 'brain training' packages. Data were only analysed from participants who trained for at least 10 minutes, three times per week and who completed the second set of outcome measures. The sessions were shorter in duration than those in many of the previous studies reported and take-up differed across groups (the reasoning group completed an average 28 sessions compared with the general group's 24 sessions and control group's 18 sessions). Effect size analyses showed no detectable effect of either training condition over and above re-test effects present in the control group.

This study raises interesting issues. On the one hand a series of small-scale intervention studies are not equivalent to a single well-powered study. Studies with fewer participants are more likely to produce extreme (unreliable) results. When combined with positive publication biases (from authors and journals) the picture is almost inevitably rosier than from a single high-powered (hence more reliable) study. Against this, a single study can only ask one question – does this *particular* training, at this *particular* duration, and in this *particular* group produce gains? A series of smaller (and ideally independent) studies can provide a more nuanced, cumulative picture of whether a certain *type* of intervention is effective and the influence of factors such as intensity, feedback and so forth. Nevertheless, Owen et al.'s study raises important cautionary points, particularly about the likely efficacy of the multibillion dollar "brain training" software industry's products in the healthy population.

Shipstead, Redick & Engle (2010; 2012) reviewed evidence on WMT efficacy, primarily in the healthy population. Among the key concerns they raise are the use of no-treatment test retest control groups (whereby the general effects of experimenter attention, participant expectation and any sort of repetitive cognitive activity are not controlled) and the use of relatively few outcome measures. To address these issues Redick et al. (2012) conducted a new study. Seventy-five young adults were placed randomly into one of three groups – dual n-back (n=24), adaptive visual search (n=29) and no contact control (n=20). The two training groups completed 20 sessions. Assessment took place at baseline, mid-way through training and post-training. Reportedly in contrast to previous WMT studies, the experimenters took care not to give participants any expectation that the training would change performance on the outcome measures. Although, as would be expected, participants improved on training tasks in both training groups providing the basis for generalised effects, at post-training there was no difference between the WMT group and either active or no-contact control groups on any cognitive outcome measures. From the mid-point assessments there was also no evidence of any dose-dependent effect reported by other authors. The study had high statistical power and a strong methodology. One limitation of this study was that although the authors set out to have young participants of various educational backgrounds, only three people used in the study were not college students. The finding of this well controlled study lend support to the findings reported by Owen et al. (2010) and as suggested by Redick et al. (2012) it could be that negative findings of WMT are under-reported.

In summary, the results from WMT studies in the healthy, predominantly young and college-educated population are decidedly mixed. Arguably these participants are likely to have the strongest pre-training WM capacity against which it may be difficult to show significant gains. This review now turns to clinical groups in whom WM is likely to be deficient.

WMT in children with developmental disorders and poor working memory

Perhaps the most encouraging findings regarding generalised benefits of computerised WMT come from a series of studies by Klingberg and colleagues. They developed an attractive game-like WMT program with an age appropriate versions for children (RoboMemo) and adults (CogMed http://www.cogmed.com/). The typical intervention involves five-week, 20-40 minute daily programme of adaptive WMT. The authors only recommend the training be undertaken with a

CogMed-trained practitioner and claim benefits in many different groups. Two studies have examined the effectiveness of a WMT program in remediating the cognitive deficits associated with Attention Deficit Hyperactivity Disorder (ADHD). These studies were one small-scale RCT where n=7 in both the control and experimental conditions (Klingberg, Forssberg & Westerberg, 2002), and one larger multi-site RCT (Klingberg et al., 2005). The findings from both studies were consistent, so focus here is on the better powered 2005 study. Fifty-six children aged between 7 and 12 years with a diagnosis of ADHD (any subtype) were screened for suitability, with 53 children being randomized to condition (exclusion criteria included being on medication for ADHD symptoms, IQ < 80, comorbid operational defiant disorder, autism/Asperger's syndrome or depression). Further exclusions meant that the final analysis included data from 20 participants in the treatment group and 24 participants in the control group. The WMT involved performance of WM tasks such as spatial span tasks, with the difficulty level titrated to individual performance. This would mean that if a child could accurately recall a 6-item series, the subsequent trial would comprise 7-items, whereas if the 6-item trial was incorrect, the subsequent trial would include 5 items. The comparison group underwent training on a software package identical to the WMT condition, with the exception that task difficulty was fixed and relatively easy. Both groups completed approximately 25 home-based sessions, each lasting 40 minutes, over a period of five weeks. All participants were assessed at baseline, post-training and 3-month follow-up on various standardised measures of WM (digit span, spatial span, both forwards and backwards), response inhibition (Stroop test), and reasoning (Raven's Coloured Progressive Matrices), along with parent and teacher-rated scales of ADHD symptoms (from the DSM-IV items, and the short version of the revised Conners Rating Scale). Significant interactions between training group and time in the direction of increased benefits with WM training in comparison with 'placebo' training were observed on all cognitive tests, including Ravens Matrices. The only non-significant difference was on the number of head movements made during performance of a 15-minute computer task (an index of motor activity). Furthermore, those who had undergone progressive WMT were rated by parents as exhibiting fewer symptoms on inattention, and hyperactivity/impulsivity (teacher-reported symptoms were unchanged). Most effects were unchanged at follow-up, and effect sizes were comparable with those reported in medication studies. For example, the effect size on the span board task, which was maintained at follow-up, was .93; and the effect size on parent-rated symptoms was 1.21 immediately post-training, reducing to .67 at follow-up (where effect sizes >.8 considered to be strongly clinically significant, and over .5 moderately clinically significant).

One factor limiting possible generalisation of these results is that, unlike in many other countries, the children in this sample were generally not on stimulant medication. However, this issue was addressed by Holmes, Gathercole, Place, Dunning, Hilton and Elliott (2009a). In this study children (8-11) were tested on a battery designed to assess different components of WM functioning (broadly corresponding to the storage versus manipulation of verbal and visuospatial information) and the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) at four time points: pre-training off medication, pre-training on medication, post-training on medication, and 6-month follow-up on medication. The training package was identical to that used by Klingberg et al. (2005), at the same intensity and for the same duration. The only difference was that training was conducted in small groups, in a school setting, rather than alone in the child's home. The authors found that medication significantly improved visuospatial WM (including manipulation of information) but no other component. In contrast, Holmes et al. (2009a) found significant improvements in performance on the untrained WM tasks were found on all WM components after WMT, and all improvements were maintained at follow-up, with the exception of those for verbal storage. This finding was interpreted in light of differing reliance of the outcome measures upon central executive resources. Unlike the Klingberg et al. (2005) study, there was no observed transfer of benefit to IQ (although this was assessed by the more crystallised IQ measure Wechsler Abbreviated Scale of Intelligence).

Another study conducted by Holmes and colleagues (Holmes, Gathercole & Dunning, 2009b) involved recruiting students in mainstream schools whose WM performance was categorised as "poor" (<10th percentile), between the ages of 8 and 11 years. Such children were randomised to receive either standard CogMed training (n=22) or non-progressive CogMed training (n=20), again in groups of 4-8 at school, and assessed using the same standardised measure of WM, along with Wechsler tests of Reading and Mathematical ability. Significant interactions between time and training groups, indicating benefits specific to adaptive training, were found for three out of four WM domains, again on those components with demonstrated Central Executive loading, and also on an ecologically valid test of WM. After training, 68% of children had WM performance of an age-appropriate level, when only 25% had been classified as such prior to training. Furthermore, no difference in IQ was observed. An interesting supplementary result of this study was that those children who received adaptive WM training were significantly better on the Wechsler Objective Number Dimensions (WOND; Wechsler, 1996) test at 6-month follow up than those who had undergone static WM training. This suggests that the increased WM capacity allowed children to engage in standard classroom teaching to a greater degree.

In a group of younger students, Thorell et al. (2008) conducted a study to examine the effects of WMT in comparison with Inhibition Training (IT) in typically developing Swedish children aged 4-5 years. The classes, which each included 14-18 children, were randomly assigned to one of four conditions: WMT, IT, active control (computer games with low WM load), or passive control. The two types of training were both adaptive to performance, had a similar interface (inhibition training was based on standard go-no go, flanker and stop-signal paradigms). The children performed training tasks in school, for 15 minutes per day over a 5-week period, with pre- and post- training assessment conducted blind to condition. WMT, but not IT, was found to produce significant benefits (with large effect sizes) on untrained WM tasks and, with smaller effect sizes, on the continuous performance test. There was no effect of either training programme on other untrained tests of inhibition.

In another study of young children (101 four year olds), Bergman Nutley, Soderqvist, Bryde, Thorell, Humphreys and Klingberg (2011) used a double-blinded randomised controlled trial to examine if non-verbal reasoning training, WMT, a combination of both or a placebo version of the combined training would lead to improvements in fluid intelligence. Five weeks of training was undertaken by the children. WMT again used CogMed. The non-verbal reasoning training used geometrical figures and was based on tests that have a high fluid intelligence loading. WMT was associated with gains on WM outcome measures but not on IQ. IQ improvements occurred only after the verbal reasoning training. Improvements were positively correlated with the amount of training. This ‘dose effect, consistent with the findings of Jaeggi et al. (2008), increases confidence that the effect is related to the training.

There are a series of studies reporting positive results from WMT primarily on the basis of teacher reports on ADHD questionnaire scales. Beck, Hanson, Puffenberger, Benninger and Benninger (2010) report results from 42 children aged between 7 and 17 years with ADHD who trained for five weeks using the CogMed program. These children were compared with a no-contact control group with the same diagnosis and had met the initial same criteria for participation. The results from teacher or parent report measures/interviews were consistent with improvements in inattention, planning, working memory and symptoms of ADHD. No objective measures or self-reports by the participants themselves were used in the study. Using a different training package, "Memory Booster", St Clair-Thompson, Stevens, Hunt and Bolder (2010) looked at improving working memory in two hundred and fifty-four children (117 in training group) aged 5-8 years over a period of 6-8 weeks. The difference between control and training group was significant on WM tasks that looked at the phonological loop and auditory recall. The results showed no transfer to untrained reading and mathematics measures. However, whilst these tests doubtless have a current WM demand it would be expected that, if improved WM enhanced classroom attention and learning, gains on measures of attainment would occur later in time.

Despite the promising findings reported above, a recent meta-analysis by Melby-Lervåg and Hulme (2012) (of twenty-three studies of neurologically healthy children) was critical of a lack of follow-up effects are argued that, when benefits occurred, these were primarily to near rather than far transfer measures. In a somewhat contradictory manner, the authors expressed concern about the plausibility of far transfer in the absence of near benefits (e.g. Jaeggi et al., 2008; 2010's IQ gains in the absence of clear WM improvement). This presupposes however that all measures are equally sensitive to change and (understandably) that the mechanism of any benefit of WMT is increased WM. If however WMT was associated with increased confidence in test situations, or the application of further effort, these effects may well show-up on IQ test that would benefit from greater persistence but not on a WM task that was already at capacity. Whilst this might mean that WMT was not 'training' WM, it may just happen to be a good way of achieving these other gains. Taken together with Shipstead et al. (2012) this review nevertheless adds caution to the idea that WMT can produce lasting generalised benefits in neurologically healthy adults and children.

Studies of WM training in people with intellectual disabilities

Reduced WM performance is likely in children and adults with intellectual disability (Alloway & Alloway, 2010; Gathercole & Alloway, 2006). A number of authors have looked at training strategies in this population (e.g. Comblain, 1994) but the use of computerised training remains relatively under-researched. Ninety-five adolescents with mild to borderline intellectual disabilities were assigned to either an adaptive WMT, non-adaptive WMT or control condition of identifying the odd-one out in stimuli (Van der Molen, Van Luit, Van der Molen, Klugkist & Jongmans, 2010). Participants completed five weeks of training and were assessed immediately afterwards and at a ten-week follow-up. The training package was developed by the authors and involved three-times per week training sessions with the sessions lasting six minutes. In the training condition, task difficulty was titrated according to performance. The non-adaptive training had a low demand. Adaptive training was associated significantly greater improvements in verbal short-term memory, maintained at follow-up.

Dahlin (2011) looked at the effects of the CogMed/RoboMemo programme on the reading performance of 57 children with special needs. It was unclear what specific educational or learning needs the children had or if they had any formal diagnosis, e.g. intellectual disabilities. An uneven number of participants were allocated to the training (42) and control (15) groups and no information is provided on how allocation took place. In addition, the control group only completed the outcome measures at baseline with inferences being drawn on their likely post-training abilities by comparison with data from Klingberg et al's. (2005) control participants (raising the question of the need for any control group at all in this study). In that context it is difficult to fully interpret the finding that RoboMemo WMT disproportionately improved reading ability.

Alloway and Alloway (2009) carried out a small-scale study with fifteen participants with learning disabilities (eight of whom were allocated to a training condition). The training was challenging and adaptive and lasted 8 weeks, 3 times a week for 30-minutes. The control group received the same duration of targeted educational support. The authors reported improvements in measures of crystallised intelligence following WMT but not in the control group. As this paper was a short report from conference proceedings the results need to be viewed cautiously until they are subjected to peer-review and more information on methodological aspects of the study are available.

As with other clinical groups, therefore, there is some promising preliminary evidence on adaptive WMT in people with intellectual disabilities that warrants further research.

Computerised working memory training with older adults

There is great interest in the value of cognitive and other activities in protecting older people against age-related declines including, presumably, in WM capacity. Various methods have been investigated. Bugos et al. (2007), for example, taught sixteen older adults to play the piano over a six-month period (at a recommended rate of 3 hours per week). Compared to an untreated control group they improved on two of the assessment measures – trail making and digit symbol (although these benefits were not maintained once rehearsal discontinued). Trail making and digit symbol are not traditionally considered WM tasks in that they do not specifically target mental retention or manipulation. However, it is perhaps illustrative of the ubiquity of WM demands in almost any activity that the authors claim a WM benefit from these results. Similarly, learning the piano (or any new skill) undoubtedly has a higher WM demand that highly practiced ‘automatic’ play. The study is therefore perhaps a useful reminder that one does not have to engage in neuropsychological test style WM activities to exercise WM. However, the lack of an active control condition means that the benefits cannot be attributed with any certainty to piano practice per se.

Basak et al. (2008) trained 19 older adults on real-time strategy computer games. Fifteen hour and a half training sessions were completed. They reported improvements in game playing and, relative to a control group, on untrained tests of task-switching, reasoning and working memory. The computer games placed large demands on participants in terms of WM and problem-solving and so these gains are arguably near-transfer effects. The lack of an active control condition however again limits direct attribution of the benefits to the particular characteristics of the games.

Ackerman and colleagues (2010) investigated a commercially available cognitive training package with a group of 78 older adults. Unlike Owen at al. (2010) the training was only carried out on older adults and was compared with the same number of reading sessions on current affair topics – i.e. technology, medical drugs. A cross-over design was used as each participant completed both the “Nintendo Wii Big Brain Academy” software and the reading sessions (20 one hour sessions of each). Unlike Basak et al. (2008) but consistent with Owen et al. (2010), there was within task gain but no transfer of training reported from either the reading or the Wii activity.

McDougall and House (2012) investigated generalised gains from hand-held Nintendo DS brain training software. Six-weeks of training programme was associated with greater improvements on the WM heavy backward digit span test than in the control group who had no contact with experimenters between assessment sessions. Additionally, the authors reported that baseline quality of life ratings interacted with treatment gains, such that those with higher QoL displayed greater training-related improvement. The authors propose that the improvements seen in backward digit span may be due to the similar demands of the training task where information had to be remembered, manipulated and re-produced also to certain extent - again consistent with near-transfer gains. Taken together the studies show beneficial effects of practicing *something*, at least on measures that are similar to that which has been practiced. The lack of an active control however makes it difficult to attribute the gains specifically to one or other form of training. It is always possible that any kind of intervention (e.g. exercise programme, art, music, break-dancing) would lead to gains. Nevertheless, the results are at least consistent with a benefit from repetitive cognitive training including on tasks with a strong WM component and, in that, are encouraging.

The lack of transfer to far measures seems to be a common theme for many authors who use cognitive training with an older population (e.g. Buschkuehl et al., 2008;

Zinke, Zeintl, Eschen, Herzog & Kliegel, 2012). Brehmer, Westerberg & Backman (2012) looked at the effects and the differences of WM training on younger (n=55 aged 20-30 years) and older (n=45, aged 60-70 years) adults. Using the CogMed program, participants trained for 5 weeks. An active control group (half of the sample) was used in which the task difficulty level was kept at a low level. Differences between younger and older participants in the degree of generalisation were expected, in part on the basis of the failure of other studies in the older population to show such effects. In fact, there were negligible differences between the groups in this respect. The adaptive WMT group showed greater gains on the untrained tests of Stroop, PASAT and CFQ relatively to the control group, consistent with generalisation of the WMT. Such effects were not however detectable on tasks Raven's and RAVLT. Whether this reflects the underlying processes necessary to complete these tasks or psychometric sensitivity is less clear.

Richmond, Morrison, Chein and Olsen (2011) in an earlier study using an active control group design, showed generalisation of WMT in an older adult population (n=21) to Reading Span, the number of repetitions on the California Verbal Learning Test and improvement in the number of self-reported lapses of attention as measured by the Test of Everyday Attention, relative to controls. As with Jaeggi et al.'s (2008; 2010) digit span showed no improvement. Although this is an iconic WM task, its demands can be argued to fall primarily on the phonological loop rather than the central executive requirement. This is brought more to the fore when the stimuli require manipulation as in the backward digit span.

In summary, studies with older adults have produced mixed findings. Gains on the trained tasks are often reported and these quite frequently transfer to similar, measures. Such findings would be consistent with a WM capacity training effect but also on the development of strategy, comfort with this sort of task and so on. Far transfer to measures ostensibly very different to the trained tasks has been less commonly observed and transfer to everyday function has rarely been examined. There are a number of issues in the study design that are important. As discussed, the absence of an active control makes interpretation of specific benefits of WMT and related interventions difficult. The use of the arguably conservative non-progressive WMT as a control also carries risks. The first is that repetitive practice and focused cognitive activity, regardless of whether it gets more challenging, is itself beneficial and real gains are missed in these comparisons. The second is that participants infer from the easy tasks that they are in the control group which affects motivation and effort. The plausibility of these control conditions has rarely been addressed. If health policy makers and potential users are to be convinced of the value of cognitive training a range of different study designs are required – firstly doing something is better than doing nothing (non-active control required), secondly, of the things you could do, discovering which is associated with the greatest gains (active control groups required).

Studies of WMT in people with brain injury

Problems with working memory are common a common consequence of brain injury and can have a negative effect on recovery (Lundqvist, Grunstrom, Samuelsson & Ronnberg, 2010). With the surge in WMT research in the healthy and developmental populations, unsurprisingly investigations into its potential efficacy following neurological damage have begun.

There is preliminary evidence that WM capacity in tasks with a high executive component may be malleable in this group. For example, McDowell, Whyte and D'Esposito (1998) conducted a preliminary study on the effects of dopaminergic medication following traumatic brain injury (TBI). They report beneficial effects on tasks involving the manipulation of information within WM, dual-tasks (such as part B of the Trail Making Test), Stroop colour-word interference, letter fluency and reduced perseverative responses on the Wisconsin card sorting test. In contrast, no benefits were found in tasks requiring only simple maintenance on information in WM.

Westerberg et al. (2007) reported a study using the CogMed WMT with a group of 18 adults (35-65 years, 12-36 months post-stroke) who had self-reported problems with attention. The structure of the training was identical to that described in Klingberg et al (2005) except that comparison group did not undergo the non-progressive training but were re-tested at the same time interval. Statistically significant WMT effects were found on the Span Board task, Digit Span task, the Ruff 2 & 7 task (another attentional/WM task), and the Paced Auditory Serial Addition Test (PASAT – see below), along with a reduction in the number of cognitive lapses self-reported on the

Cognitive Failures Questionnaire, arguably evidence of some far transfer and functional effects in everyday life although general effects of active training cannot be ruled out. No statistically significant effects were found however on Raven's Matrices, the Stroop task, or a declarative memory test (learning word lists). There was no follow-up assessment.

Kim, Chun, Kim and Park (2011) examined a computerised cognitive training programme in a group of patients following stroke that focused on memory and attention. A second group combined this training with virtual reality training that looked at moving objects, juggling etc. Following 30 minutes of training 3 times a week over 5 weeks the only improvements seen were in visual continuous performance and backward digit span. There was no improvement seen in measures of ADL and there was no no-contact control group in the study. Participants were not trained on strategies to improve their performance.

Serino et al (2007) reported findings consistent with those of Westerberg et al.(2007) using a non-proprietary WM training program in 9 adults with TBI (at least 6 months post-injury). They used a within-subjects design, incorporating a 'placebo training' control condition which always preceded the 'active training' condition. The placebo training involved listening to series of items belonging to a particular category (e.g. digits, months of the year), and making a judgements regarding the most recently presented item (e.g. classifying odd/even, or winter month/non-winter month). All participants underwent four training sessions per week for a period of 4 weeks. Participants then underwent WMT, again for four sessions per week for four weeks. WM training involved repeated practice on the PASAT. In this test participants hear a sequence of single digits read aloud. The task is to add the current item to the previously heard item and report the answer aloud whilst at the same time listening out for the next digit to add to the now previously heard one (1 7 "9" 1 "8" 3 "4" and so on). It makes a number of demands including on sustained attention, numerical knowledge as well as in holding on to the last digit in WM whist simultaneously performing calculations. The months task required that participants said aloud whichever of the two previously presented months was either earlier in the year, and the words task involved participants listening to a series of words, and for each one, generating a word beginning with a particular letter, that letter being the third one in the *previously* presented word. Participants were tested on a small battery of neuropsychological tests at baseline, following placebo training and following WMT. There were significantly greater improvements following the high-WMT relative to the low-WMT control in performance on the n-back and divided attention subtasks of the Test of Attentional Performance (TAP), as well as on letter fluency, Tower of London, delayed memory recall, but no effect was observed on Speed of Processing or Sustained Attention subtask of the TAP. Additional improvements were observed on the two self-reported measures of everyday function – the Rivermead Head Injury follow-up questionnaire, and the Patient Competency Rating Scale.

So far there has only been limited study of the effects of WMT in children with head injury. Madsen Sjo, Spellerberg, Weidner and Kihlgren (2010) report an uncontrolled pilot study. Seven children with WM problems trained on an individual basis for 18-20 weeks on various aspects of attention, memory and repetition. Improvements were reported on subtests of memory and learning which is encouraging although further work is clearly required to rule out simple practice effects.

Summary and aims of thesis

It is encouraging that humans will get progressively better at almost any non-trivial task that they practice. This is evident in the WMT literature as elsewhere. Numerous factors are likely to be relevant to these improvements including familiarity with what is required and the time-course of particular tasks, increased confidence through success, the development of strategies that minimise task demands and so forth. Specific training effects are very useful if the task is something that is needed in everyday life, or which, like computer games, Sudoku etc. provide inherent pleasure in their mastery. However, if tasks are practiced in the hope of providing spreading benefits to everyday skills this needs to be demonstrated rather than assumed.

In terms of WMT, the evidence is currently mixed. Well-powered and conducted studies using randomisation and active, conservative control conditions have reported generalization of WMT gains to measures of fluid IQ and questionnaire measures of everyday function. Benefits have been reported by independent groups working with different populations including children with developmental conditions, adults with neurological injury and in the healthy population. Adaptive training based on n-back, reverse span and PASAT style tasks have all returned positive results.

Different interpretive issues affect the different studies but overall there are grounds for optimism that adaptive computerised WMT can be effective. Against this, systematic reviews and large, well-powered studies provide grounds for caution, particular with respect to positive publication bias.

Whilst the efficacy of WM training is open, it is clearly not the only potential route to improving capacities required across many tasks. It is legitimate to ask, if measures of fluid intelligence have the highest relationship to *g,* which, by definition, is a factor influencing performance across a range of tasks, can repetitive progressive training on fluid intelligence tasks produce generalised gains? It is important to stress that this is a qualitatively different question to ‘training the test’ (e.g. tests of pronunciation of irregular words like the NART are correlated with IQ, it would not be expected that teaching someone this particular set of words would improve IQ on any other measure!).

Training fluid intelligence?

In a series of related papers (Adult Development and Enrichment Project -ADEPT; Plemons, Willis & Baltes, 1978; Willis, Blieszner & Baltes, 1981; Willis & Nesselroade, 1990) interventions aimed at improving, maintaining or slowing loss of general intellectual function was carried out with small groups of older adults. The training was based on task analysis of the Culture Fair Test (Scale 2; Cattell & Cattell, 1960), and involved identifying different figural relation rules and determining logical patterns. Plemons et al. (1978) trained fifteen people in groups of three, with a mean age of 69, twice weekly for four weeks. Compared with the control group there were advantages on the untrained Cattell-Horn Figural Relations Battery that were maintained at follow-up. Few benefits were seen however on other measures of Induction and Verbal Comprehension.

Willis et al. (1981) used a similar training regime with 58 older adults (mean age 69.8years) which was again associated with well-maintained gains on similar but untrained tasks of figural relations – the ADEPT Figural Relations Test (developred for the study), the Culture Fair Tests and the Raven’s Matrices. There was again no evidence of change on far-transfer measures – Letter Series, Number Series and

Letter Sets, as well as the Identical Pictures set. A version of this training was conducted with 38 older adults (mean age 69 years) over three phases (sample size reduced slightly with each phase of training – final sample was of twenty-five participants) which involved five training sessions (Willis & Nesselroade, 1990). They showed positive training effects to their own measure – ADEPT Figural Relations Test at all three phases and at follow-up but no generalisation of benefits was reported. There was no report of any other measures given. The primary outcome measure in all three studies (The ADEPT Figural Relations Test; Plemons et al., 1978) was specifically developed for these studies and, of course, had notable *conceptual* overlap with the trained tasks even if the precise rules and stimuli differed.

Stein-Morrow, Parisi, Morrow & Park (2008) invited 87 older adults (mean age 73) to take on a long-term problem to solve over the 20 weeks of their study. Participants chose a problem in the area of literature, history or science. They worked individually and in groups (plus coach) of 5-7 people engaged in the same problem. Significant improvements were noted on measures of fluid intelligence relative to controls at pre and post- training.

In another study from that year, forty-four adults with a mean age of 67.82 years took part in a series of stimulating cognitive activities for 10-12 weeks (Tranter & Koutstaal, 2008) at home and in groups. A small but significant improvement in the primary outcome measure, the Cattell, was observed when compared to the no-contact control group. There was an upper-age limit on the participants who took part which could influence the application of the research to older adults as a whole.

Summary and rationale for the empirical sections of this thesis

It is now possible for people, including those at risk from cognitive impairment through age, neurological illness or developmental difficulty, to engage in intensive computerised cognitive training at low financial cost. WMT is the best developed and investigated specific capacity training of this sort. Early studies in the healthy population showed that training produced gains that generalised to untrained fluid IQ measures, previously thought rather immutable to such effects. Subsequent work has challenged the methodology and failed to replicate such effects in large samples. WMT studies of people with low baseline WM performance has produced a more positive but still mixed picture. Appropriate controls in some cases mean that these cannot all be dismissed as regression to the mean or general effects of 'intervention'. It would be premature however to conclude that WMT is more effective where WM capacity is low – too few studies have been conducted. Very few studies have examined the logical corollary to fluid IQ benefit generalisation, namely improved performance in a range of everyday situations. There are some positive results from self- and teacher- report questionnaires of such gains but it remains possible that, even when there is apparent IQ transfer, it is still encapsulated in a 'I am doing a test' context that resembles that of training. Against this it is much more difficult to quantify everyday function than test performance and real gains may have been missed or may have occurred at a longer latency than studied. Demonstrating such generalised benefit will be crucial however if WMT is to be recommended, including to clinical groups. Without that there is a real risk of people spending long periods in ultimately fruitless activity. Premature dismissal similarly runs the risk of missing one of the most potentially important consequences of the digital age.

Whilst questions remain about WMT, it seems relatively clear that versions which increase difficulty with performance gains produce better results than those that remain at a stable, easily achievable level. Adaptive training could be applied to a range of other specific cognitive operations such as inhibition, selective, sustained and divided attention, prospective memory and so forth. In each case the same questions of generalisation would apply. As discussed, an area of interest with some preliminary suggestions of efficacy is to train the sort of logical pattern identification and mental flexibility required by tests with a high *g* loading, such as the Raven's and Cattell.

Both the WM and g training literature raise important questions on the underlying components of intelligent behaviour – what is it that contributes to so many different sorts of tasks and which is particularly taxed on the abstract logical puzzles of the Cattell? One approach has been informed by 3 key observations; the effects of lesions to the prefrontal cortex on intelligent, flexible goal-directed behaviour; the common activation in healthy fMRI of a fronto-parietal "multiple demand" network by many tasks *as they become more challenging;* and by the strong activation of these MD areas by tests most sensitive to *g*. Whilst intelligence may not be synonymous with the activity of this network a good place to start is by asking what are these regions contributing to task performance?

The first observation, highlighted by Duncan (1995) is that in single cell recording work, a remarkably high proportion of studied neurons in dorsolateral prefrontal cortex respond to *whatever* is currently relevant to the task at hand. Unless the researchers have been particularly lucky, the implication is that a large proportion of the total cell population in these regions are able to tune to what is relevant. This is in contrast with the general pattern in more posterior, sensory areas where a cell will tend to respond when its preferred stimulus is encountered, usually more so when it is task relevant or attended, but will not switch to coding a different sort of event. This work suggests that an arbitrary intention, given to the participant by the experimenter with respect to displayed stimuli (which is a target and so on) is very quickly able to transform this bit of cortex into a 'task solver' (Roa et al., 1997). The implication is that, in a limited capacity system, if these cells were not coding elements relevant to the current instructed task, the participant would either have completely 'misunderstood' the task, be doing another task or be doing nothing.

The next remarkable property of this region is that patterns of activity across neurons are sensitive to the part of the task in which the participant currently finds himself. For example, in a task where there is a cue, the presentation of a target or distractor and then the response, patterns of firing across cells reliably demarcate these epochs (Sigala et al., 2008). The region not only codes what is relevant but updates information about which objects and actions are relevant as the task unfolds. The implication is that, if these distinctions were not made, the participant may have a very lose collection of targets, rules and possible responses that may be enacted at the wrong time and to the wrong stimulus or repeatedly produced without obvious reason. Performance is unlikely to be efficient.

If a participant performs an arbitrary instructed cue – target – response task accurately on the first trial, the components of that task must have been preassembled at the time of instructions. In line with this, MD regions show patterns of discrete activity during the receipt of instructions (Woolgar et al., 2010). Other characteristics of neural responses in this area include the capacity to fire across a delay between a cue and the ability to execute the cued behaviour (Petrides, 1994).

These characteristics are not ‘intelligent’ in one sense, in that the participant is just doing what he or she is told in a simple task. This is a consequence of the sort of simple, repeated task structure necessary to reveal reliable underlying patterns of activity. However, it provides the basis of mechanisms that would be needed to enact coherent goal-directed behaviour free from simple response to environmental contingency.

Another characteristic of this region closer to an idea of inductive intelligence is chunking. As discussed in relation to strategy in WM, this can refer to the breaking down of an undifferentiated stream of information into meaningful chunks that will facilitate recall (by reducing WM demand). Bor, Duncan and Owen (2001) looked at fMRI responses as participants viewed spatial span tasks that either implicitly encouraged chunking (e.g. the sequence traced out a geometric shape) or in which chunking was difficult/ futile. Much greater MD activity was observed in the high chunking situation despite this never being instructed. Potentially, this is an example of spontaneous breaking down of a difficult task into manageable sections. Extrapolations to complex goals (such as planning a holiday) into meaningful sub-goals (selecting time, destination, booking, packing etc.) remain just that but it is relatively easy to see how, as with coding relevance and sequence, failure to do so would lead to a lack of ‘intelligent’ goal-directed behaviour.

The first empirical chapters of this thesis focus on the assessment of outcome from WMT. Chapter 2 focuses on a new computerised task (The Games Evaluation Test) that is designed to capture key elements of behavioural organisation in complex unstructured life-like settings. Emphasis then turns to an existing measure of such organisation, the Multiple Errands Test. An issue in this measure, which involves monitoring participants on a complex real-world shopping task, is whether the assessment is reliable and/or open to bias when the status of the participant (e.g. patient/control, pre/post intervention) is known. To this end, the practicalities and outcome of first-person video recording and off-line blind assessment are explored as well as the relationships with the Games Evaluation Task and other measures. The thesis then takes a brief detour to examine the effects of intensive adaptive WMT in a brain injured patient. Here the focus was on whether effects generalised to another specific clinical problem that he experienced, that of time perception. A series of studies are then described that then apply multi-tasking outcome measures first to WMT, then to a form of adaptive “g” problem solving training. Finally, the potential benefits of embedding adaptive capacity training within a broader Goal Management Training approach designed to push facilitation is explored.

Chapter Two:

Developing a new computerised multi-tasking computer game and examining its relationship with existing organisational measures

Abstract

Previous studies indicate that aspects of executive function may be best measured in loosely structured, multi-component tasks that reflect complex real-life demands. In particular, the phenomenon of ‘goal neglect’ may be more easily captured in such measures. Goal neglect refers to situations in which people are able to report what it is that they should do at a given time or circumstance but, when this occurs, fail to do it. A downside to such assessments is that they are often difficult or time-consuming to contrive. This chapter reports an evaluation of a new; easy-to-administer computer-based multiple component test, the Games-Evaluation Task (GET).

Recent research has suggested that goal neglect may be increased by the presentation of irrelevant instructions for a task. A second aim of the study was therefore to examine whether the likely sensitivity of the task could be increased by the addition of such redundant instructions. To this end, 20 adults from the older healthy population (55-70 years) completed the task under different conditions. No performance difference related to instructions was apparent. The task was, however, shown to be a useful measure of multi-tasking in itself.

Study two set out to examine the relationships between the GET and performance on measures of related constructs. The results show that indeed poor GET performance was correlated with self-reported frequency of everyday cognitive lapses and measured by the Cognitive Failures Questionnaire. Despite expectations, no significant correlations between GET and the Multiple Errands Task (MET) performance was observed. In conclusion, there are reasonable grounds for using the GET as a quick, practical and potentially sensitive measure of organisational skills.

Introduction

Difficulties in planning and multitasking

Difficulties in a planning and multitasking are commonly reported following brain injuries (Wilson, Alderman, Burgess, Emslie, & Evans, 1996), particularly following frontal lesions (Luria, 1966; T. Shallice, 1988). This can often lead to negative occupational outcomes (Crepeau & Scherzer, 1993) and influences performance and functioning in everyday tasks. An accurate assessment of cognition is crucial in many clinical settings. Its use in diagnosis, impairment screening and measuring change can provide information that is valuable to an individual's ability to return to work and in setting realistic rehabilitation goals (Makatura et al., 1999). It has been pointed out that traditional neuropsychological measures may not adequately capture such problems (Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Shallice & Burgess, 1991). Specifically, that such tests generally provide structure, clear instructions on when to stop and start, emphasise a single rather than multiple tasks and minimise the effects of strategy. In contrast, in everyday life patients may need to develop and act on strategies to manage multiple, competing demands. Accordingly, Shallice and Burgess set out to standardise complex multitask situations that might better assess these difficulties. In the first, the Multiple Errand Task (MET; Shallice & Burgess, 1991), participants were asked to perform a series of tasks in a set of shops in London's Southampton Row (e.g. buy a birthday card, check which was the warmest location in the UK yesterday). Various constraints were imposed, including how much money could be spent and that no shop may be entered more than once. The task therefore emphasised planning, initiating and following that plan as well as a degree of reasoning and knowledge. Shallice and Burgess (1991) reported that a group of 3 patients with traumatic brain injury and a high level of reported disorganisation in everyday life showed considerably more disorganisation on the task despite adequate or even good performance on a range of neuropsychological tests. The second task developed by Shallice and Burgess was the 6-Elements test, a desk-top measure designed also to stress planning, plan enactment and keeping in mind the overall goal. In the task patients were asked to complete something from each of 6 tasks within a limited time. Crucially, participants were told that they would be unable to complete all of the tasks within the total time available and that they would therefore have to decide when to switch from one to another. Three different types of task were used (writing down the names of a series of pictures, performing and writing down the answers to a series of simple calculations, and describing events into a Dictaphone), each with two parts (making up the 6-elements). This allowed the additional rule that part 2 of a given task could not be attempted immediately after part 1. Again, the authors report that the performance of patients, who performed well on standard neuropsychological assessments despite profound disorganisation in everyday life, was quantitatively and qualitatively worse than that of control participants. Despite participants being able to state the aims of the task and their plan both before and after completing it, they showed a strongly increased tendency to disregard this during their actual performance, a phenomenon that has been termed ‘goal neglect’ (see below).

Existing measures

Subsequent variations on the MET and 6-Elements test have been reported. MET and 6-Elements style tasks can be complex to administer and difficult to score (for example, if crucial behaviours are missed by the administrator). Accordingly, Miotto and Morris (1998) developed a computerised virtual reality version of a MET, offering the potential advantages in terms of automated scoring and easy administration. They tested 25 patients with frontal lesions and 25 controls on a virtual board game like task that involves planning and completing a number of set items (booking and collecting flight tickets or shopping and preparing a meal). They report a significant difference between performance of patients and controls in tasks that required organisation. Virtual Reality (VR) tasks have a number of advantages including the potential to be used with people with mobility problems, hemiparesis, and in imaging studies. However, even if the task is similar to something that participants might have themselves done or would readily understand (such as stacking boxes), the mechanisms for interacting with the task via the mouse or joystick may, unless there is considerable opportunity to practice, form a barrier for some participants. In addition, achievable VR implementations (i.e. without the vast budgets/expertise of the games industry), could be argued currently to lack immersion (the sense that one is actually in the environment, therefore undermining the ‘everyday’ nature of the tasks) and to produce activities that possibly lack sustained interest. A problem for 6-element type tests is, if the component tasks are not particularly interesting, participants may be more likely to switch between the tasks, reducing the sensitivity of the measure.

The primary aim of this study was to examine the usability of the task, specifically whether participants could understand and remember what was required, make the necessary actions and crucially, whether through task omissions or timing discrepancies, the task was sensitive to goal neglect. Volunteers from the older healthy population were assessed primarily because of their relevance to subsequent studies in this thesis. In that such people may have less experience of computers/computer games than younger participants, however, this also provided a potentially greater test of the GET's usability.

Study one

Study aim

The primary aim in the current study was to investigate conditions that may produce the most sensitive version of the Games-Evaluation Task. The GET had been developed before the current study by Prof. Jonathan Evans and Dr. Tom Manly working with programmers at the University of Aberdeen. A second and original aim of the current study was to investigate conditions that may produce the most sensitive version of the Games-Evaluation Task. Recent studies conducted by Duncan and colleagues are interesting in this respect. Duncan et al. (1996) originally coined the term “Goal Neglect” in relation to a phenomena highlighted above in which people may be able to state their intention (e.g. in the 6-Elements Test to switch between tasks) and yet, in practice, fail to act upon it. Whilst this occurs with increased frequency following brain injury, it can also be observed in the general population. Duncan and colleagues have used variants on a basic task to probe this phenomenon. In a typical task participants watch a series of letter or number pairs appearing on a screen and are asked to monitor just one side of the display and respond if a pre-specified target appears. At some stage in the task, a central cue is presented that could either be a plus or minus sign. If the plus appears, participants are asked to monitor the right hand side of the screen. If it is a minus, they must monitor the left hand side of the screen. Typically, participants are able to perform the first part of the task well but a proportion will fail to switch on the cue despite knowing and being able to state what it means. Commonly, after a certain number of trials, participants will accurately switch and then tend to maintain this accuracy for the remainder of the task. A series of studies by Duncan et al. (2008) examined factors that made such goal neglect more likely and found that the key factor was the complexity of the whole body of knowledge specified in the task instructions. Put simply, the more people had to remember, the greater their chance of neglecting a given goal despite being able to state what the goal is.

Accordingly it was hypothesised that errors and disorganisation in the GET would be increased by adding to participants' instructional load. To this end, a series of instructions were developed that were true of the tasks (the aim was not to increase errors by misleading participants) but which were unnecessary for performance. For example, one of the GET games involves moving a platform to bounce a ball against a brick wall. The key instruction in the task is to move the platform to the left and right in order to intercept the ball. The additional instruction "each brick is worth a different number of points between 5 and 30" is true but irrelevant because it was not possible to selectively aim the ball at high value bricks, only to keep the ball in play. Similarly, another game required a funnel to be moved to catch balls that were descending from the sky. Proving information about the range of colours in which the balls come does not add in any way to the strategy required but serves to increase the overall instruction load.

A final hypothesis was that errors on the task would be related to the order in which instructions were given. Based on Duncan and colleagues work it was hypothesised that initial instructions are processed more actively and tend to out-compete (cause neglect of) later instructions. The participants were therefore randomly allocated to three instruction conditions: *Relevant* – in which only the relevant instructions were given; *Relevant first* – in which the relevant instructions were given followed by irrelevant details about the task and; *Relevant last* – in which the same instruction set was given but with the information necessary to perform the task and, in particular, the requirement to switch, was given last. It was hypothesised that errors and time-deviation in the GET would increase in a gradient across the three conditions.

Methodology

Participants

Twenty volunteers (fifteen female) between the ages of 55 and 70 years (mean 66.75 years, SD 4.41) took part. Participant's IQ, as measured by the Cattell ranged from 76 to 120 (mean 108.91, SD 13.26), NART mean 10.11 (SD 7.06). The three instruction groups, described below, did not significantly differ in terms of Cattell Score [$F(2,21) = 0.96$, $p = .40$], NART score [$F(2,21) = 1.77$, $p = .20$] or age [$F(2,21) = 2.0$, $p = .16$].

All participants gave written, informed consent for their participation from the Medical Research Council Cognition and Brain Sciences Unit (MRC-CBU) database of healthy volunteers. These volunteers are recruited via local advertisement, through the MRC-CBU website, on open days, science fairs and through word-of-mouth and paid an honorarium for their contribution to studies. Participants had no history of significant mental health problems, head injury, or neurological impairment. Potential participants were screened and recruited over the phone.

Members of the CBU volunteer panel are asked to give consent for scores, such as on the NART, to be made available for subsequent studies (reducing unnecessary retesting etc.). Accordingly, to attempt to balance the groups on likely variables of importance, participants were first ranked/matched on the variables of age, gender and NART score (where available) and then randomly allocated in groups of the 3 closest match to the three GET experimental conditions. Seven completed *Relevant first* and *Relevant last* respectively. Six completed the *Relevant* condition. Testing took place on an individual basis in a room relatively shielded from visual and auditory distraction. All participants completed the computerised GET and the brief goal neglect "Feature Match Task", each presented at a comfortable viewing distance on a DELL 17" monitor. They then completed the Cattell Culture Fair and the Cognitive Failures Questionnaire paper and pencil measures. No feedback was provided on performance. The other measures will be discussed in the study 2.

<u>The Games Evaluation Task (GET)</u>

The Games Evaluation Task (GET) was programmed in Flash and presented using a Dell OptiPlex 755 computer. The GET consisted of 4 games. The first was a car racing game in which participants had to steer a moving car left and right to refuel by running over petrol pumps distributed across a three-lane road, their scores increasing with each successful contact. The second involved moving a funnel left and right to catch balls falling from the top of the screen, with each successful catch increasing the score. The third was a squash game in which a paddle was moved left and right to bounce a ball against 3 walls, all with different point values, and prevent it falling off the bottom of the screen (which resulted in a loss of points). The fourth game also involved moving a paddle left and right to bounce a ball to knock out bricks representing different point values with again points being lost if the ball was allowed to drop off the screen. In this last game, were a participant to successfully remove all the bricks, they would be instantly replaced with the original configuration, allowing the game to continue. Screen shots can be seen in figure 2.1 below:

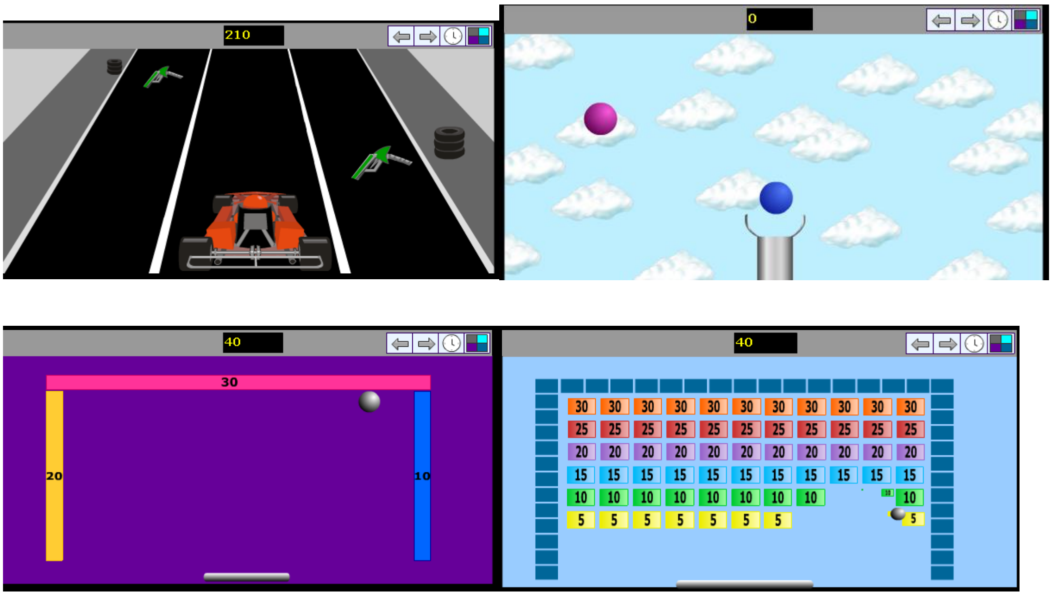


Figure 2.1: Screen shots of GET computer games

In addition to the stimuli relating to each game the screen also showed the updating points total centrally and instructions relating to the keys at the top right-hand side of the screen. These corresponded to the G H J & K keys on a standard QWERTY keyboard and represented “move left”, “move right”, “check time” and “change game” respectively. If the check time button was pressed a digital clock display would appear for 6 seconds at the top left-hand side of the screen indicating how many minutes were left. Pressing the change game key would scroll to the next game and repeated presses would cycle through all 4 games.

The GET was programmed to allow it to be run at different durations. Here the time was set to five minutes. The program automatically logged which games were played, their start and finish times, the number of points scored in each instance of play and the number and times of any clock checks. Where participants scrolled through more than one game without playing it, this was recorded by the game duration being set at 0 allowing subsequent exclusion from analysis.

Participants across all groups were asked to play the games to get a ‘feel’ for how engaging they were.

Questionnaire:

A difficult issue concerned the difference between participants *forgetting* and *neglecting* the critical instructions. If, due to the addition of irrelevant instructions, participants forgot or did not register the critical instructions, the conditions for goal neglect would not be met. Reminders, on the other hand, risked rendering the irrelevance and order manipulations obsolete. Accordingly an instruction questionnaire was introduced. The questionnaire contained a question relating to each instruction that was given to participants, regardless of the instruction condition they were in. The questionnaire was given to check retention but contained questions about the irrelevant instruction and kept the relevant first/last order. The questionnaire was given before the GET was administered (after the task instructions were given) and also after participants completed the GET to ensure retention of instructions. If participants answered any question incorrectly prior to beginning the task they were given the correct answer and understanding was again checked. All three questionnaires are available in Appendix A. The instructions given to participants in each group are described below.

Instruction conditions

In each condition the instructions were given in a written form (using PowerPoint on the computer screen and screenshots from the games) and explained by the experimenter. In each condition participants completed the questionnaire about the instructions before beginning the task. Any errors were corrected and it was a prerequisite for beginning that participants were able to correctly report all relevant instructions.

Relevant condition:

In the *Relevant only* condition participants were given the following instructions:

"In this task you are being asked to rate your experience in playing this computer game. I would like you to let me know at the end of the game how engaging you found it.

Have a look at this PowerPoint Presentation where I will give you the necessary instructions [PowerPoint presentation is then turned on on the computer that is to be used for the game – verbal instructions are also displayed on screen with visual prompts].

There are four games in this task- driving, catching, squash and bricks.

At the end of the task I will ask you to tell me how engaging you found the games on a scale between 1 and 10, 1 being not at all engaging and 10 being very engaging.

You need to play all the four games twice.

You will have 5 minutes overall to complete this task.

Please try to spend an equal amount of time on all the games.

You use this button [point to "clock" button] to check the time, this button [point to left and right buttons] to control the direction in each game, and this button [point to square button] to change games.

Do you have any questions about that? [Clarify/answer any question].

Ok – before we start I would like you just to answer a few questions on this sheet to make sure that you know the instructions.

[Give sheet to participants to fill in. Check sheet for answers – if correct then continue but if incorrect clarify and correct any wrong answers]

I will now give you a quick practice on the task. [Give participants 30 seconds to get a feel for the controls and to see how the games work].

Now we will start the task. Please do not speak to me during the task unless there is problems with the computer.

[At the end of the task ask participants to fill in the questionnaire again]."

Relevant first condition:

Participants in the *Relevant first* condition received the above instructions and then a further 18 instructions that were accurate but not crucial for performance of the GET. The information given is chunked under the headings of the four games – both on the PowerPoint presentation of the rules and in the questionnaire that is given to participants. The most relevant information to the task is given at the beginning under each game's heading. For example in the "Driving" section the instruction relating to switching tasks (relevant) and controlling direction (relevant) are given before the next four instructions which are irrelevant. This is the case for all four games - following each game's heading a relevant (or two relevant) instructions was given before irrelevant instructions (those that were not needed for successful completion of the task) were given. The layout of the questionnaire was the same for all three conditions as can be seen in Appendix A.

In the *Relevant first* condition participants were given the following instructions:

"In this task you are being asked to rate your experience in playing this computer game. I would like you to let me know at the end of the game how engaging you found it.

Have a look at this PowerPoint Presentation where I will give you the necessary instructions [PowerPoint presentation is then turned on on the computer that is to be used for the game – verbal instructions are also displayed on screen with visual prompts].

[Picture of countdown screen]

This task begins with a countdown from 12 to 0. You start the tasks after it reaches 0.

[Picture of driving game]

You switch to the next game by pressing the square button [point at square button].

You press this button [point to left and right buttons] to control the direction in each game.

In the driving task you are trying to drive over the petrol pumps- you score points this way.

You can see the tyres here [point to tyres on screen] – they're not relevant to the task at all.

The car will continue to travel at the same speed- you can't slow down or speed up.

[Changes to picture of Catching game]

You will have 5 minutes overall to complete this task.

You use this button [point to "clock" button] to check the time

In this catching game you are trying to catch as many falling balls as you can.

If you check the time then the time disappears again after a few seconds.

The main thing I would like you to do in this task as I said is to evaluate how engaging you find all the games.

At the end of the task I will ask you to tell me how engaging you found the games on a scale between 1 and 10, 1 being not at all engaging and 10 being very engaging.

[Changes to picture of Squash Game]

Please try to spend an equal amount of time on all the games.

There are four games in this task- driving, catching, squash and bricks.

The aim of the squash game is to score as many points as possible by bouncing the ball off the coloured walls and you can score different points depending on which wall the ball hits.

If the ball falls off the screen you lose points.

I will ask you to evaluate how you found the games when the time has elapsed.

[Changes to Picture of Bricks Game]

You need to play all the four games twice.

The aim of the bricks game is to score as many points as possible by bouncing the ball off the bricks, without letting the ball fall off the screen.

You get different points depending on which bricks are removed – different bricks are worth different amount of points.

You don't get any points for hitting the outside walls and if the ball falls off the screen then you lose points.

Do you have any questions about that? [Clarify/answer any question].

Ok – before we start I would like you just to answer a few questions on this sheet to make sure that you know the instructions.

[Give sheet to participants to fill in. Check sheet for answers – if correct then continue but if incorrect clarify and correct any wrong answers]

I will now give you a quick practice on the task. [Give participants 30 seconds to get a feel for the controls and to see how the games work].

Now we will start the task. Please do not speak to me during the task unless there is problems with the computer.

[At the end of the task ask participants to fill in the questionnaire again]."

<u>Relevant last condition:</u>

Participants were given the same 24 instructions as above but with the relevant instructions appearing last in each game's section.

In the *Relevant last* condition participants were given the following instructions:

“In this task you are being asked to rate your experience in playing this computer game. I would like you to let me know at the end of the game how engaging you found it.

Have a look at this PowerPoint Presentation where I will give you the necessary instructions [PowerPoint presentation is then turned on on the computer that is to be used for the game – verbal instructions are also displayed on screen with visual prompts].

[Picture of countdown screen]

This task begins with a countdown from 12 to 0. You start the tasks after it reaches 0.

[Picture of driving game]

In the driving task you are trying to drive over the petrol pumps- you score points this way.

You can see the tyres here [point to tyres on screen] – they’re not relevant to the task at all.

The car will continue to travel at the same speed- you can’t slow down or speed up.

You switch to the next game by pressing the square button [point at square button].

You press this button [point to left and right buttons] to control the direction in each game.

[Changes to picture of Catching game]

In this catching game you are trying to catch as many falling balls as you can.

If you check the time then the time disappears again after a few seconds.

The main thing I would like you to do in this task as I said is to evaluate how engaging you find all the games.

At the end of the task I will ask you to tell me how engaging you found the games on a scale between 1 and 10, 1 being not at all engaging and 10 being very engaging.

You will have 5 minutes overall to complete this task.

You use this button [point to “clock” button] to check the time

[Changes to picture of Squash Game]

The aim of the squash game is to score as many points as possible by bouncing the ball off the coloured walls and you can score different points depending on which wall the ball hits.

If the ball falls off the screen you lose points.

I will ask you to evaluate how you found the games when the time has elapsed.

Please try to spend an equal amount of time on all the games.

There are four games in this task- driving, catching, squash and bricks.

[Changes to Picture of Bricks Game]

The aim of the bricks game is to score as many points as possible by bouncing the ball off the bricks, without letting the ball fall off the screen.

You get different points depending on which bricks are removed – different bricks are worth different amount of points.

You don't get any points for hitting the outside walls and if the ball falls off the screen then you lose points.

You need to play all the four games twice.

Do you have any questions about that? [Clarify/answer any question].

Ok – before we start I would like you just to answer a few questions on this sheet to make sure that you know the instructions.

[Give sheet to participants to fill in. Check sheet for answers – if correct then continue but if incorrect clarify and correct any wrong answers]

I will now give you a quick practice on the task. [Give participants 30 seconds to get a feel for the controls and to see how the games work].

Now we will start the task. Please do not speak to me during the task unless there is problems with the computer.

[At the end of the task ask participants to fill in the questionnaire again]."

Main GET variables

Participants received 1 point for each game instance that they played. If participants played more than 8 games, this was capped at 8 but not penalised in any way, the maximum score was therefore 8. The deviation scores were calculated from the absolute difference between duration of each time a game was played and the optimal of 37.5 seconds (i.e. had all games been played twice in the 5 minutes available). The direction of the deviation was not taken into account such that underplaying or overplaying were equally penalised.

Results

General performance and score distributions on the Games Evaluation Test (GET)

Overall, for Games Tried, of the 8 points that would have been an optimal performance of the task, participants managed to score a mean of 6.32 (i.e. started 6.32 games; SD 1.64). Examining the distributions, 31.8% of the sample (n = 7) scored the maximum 8 points with approximately 60% scoring 7 or 8 points. The score was skewed towards ceiling (Skewness = -0.56, Kurtosis (-3) = -1.08).

The mean summed deviation from optimal time (see methods) was 135.25 seconds (SD 48.17). As would be expected, a more normal distribution was observed (Skewness = 0.70, Kurtosis = 0.142). The positive values here for skewness and kurtosis indicate too many low scores in the distribution and a pointy and heavy-tailed distribution.

The effect of instructions on Games Evaluation Test performance

Although one-way ANOVA can be robust to violations of its assumptions, the skew on the Games Tried Score was such that that the non-parametric equivalent Kruskal-Wallis Test was applied. This was seen as appropriate too given the small sample involved.The Kruskal-Wallis test ranks for the *Relevant*, *Relevant first* and *Relevant last* conditions were 8.14, 9.50 and 17.14 respectively. There was a significant effect of instruction condition on the number of games attempted H (2)= 7.785, P= 0.02 with participants in the *Relevant* condition scoring a mean of 5.50 (SD 1.64), *Relevant first* participants scoring a mean of 5.88/8 points (SD 1.73) and *Relevant last* participants a mean of 7.71 (SD 0.49), see figure 2.3.

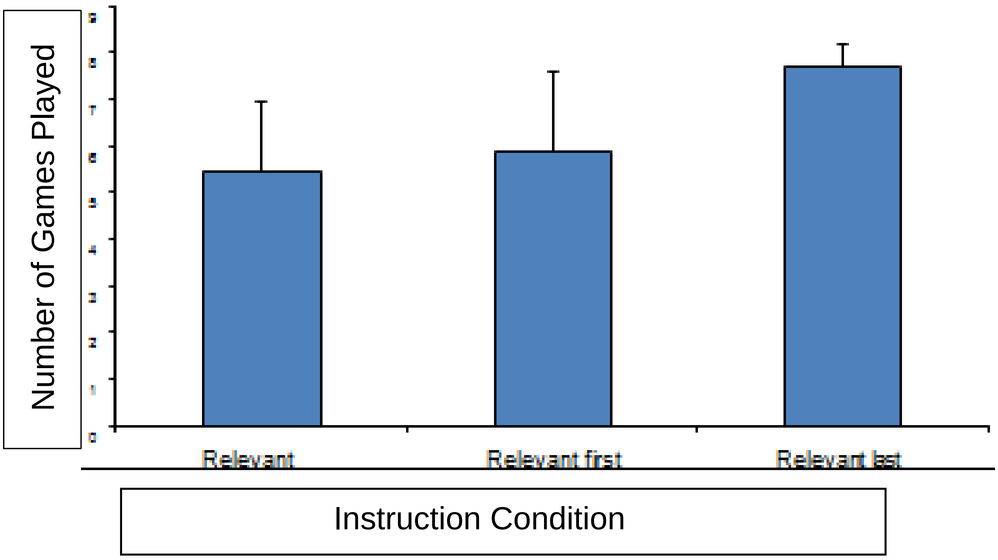


Figure 2.2 The number of points (games played) scored by participants under the three instruction conditions (error bars = 1 standard deviation).

As might be expected from these mean values, post-hoc analysis using Mann-Whitney two-group tests showed no statistical difference between the performance of the three groups (*Relevant last* vs. *Relevant first*, P = 0.14; *Relevant last* vs. *Relevant* P = 0.22, *Relevant firs*t vs. *Relevant* P = 0.76; two-tailed P values).

This result was therefore contrary to the hypothesis that irrelevant instructions would worsen performance on the test and that relevant aspects of the task would be neglected by being presented later in the instructions. Whilst a recency effect might explain why participants who received the relevant instructions just before performing the task might slightly outperform those for whom there was an intervening period of irrelevant instructions it is difficult to see how this could account for the *Relevant last* group outperforming, although not to a significant level, those who only had the relevant instructions (i.e. that preceding the relevant with irrelevant instructions in some way helped!)

Turning to the time-deviation score, the mean values for the *Relevant, Relevant first and Relevant last* conditions were 138.08 (SD 38.58), 138.81 (SD 36.39) 131.57 (SD 72.33) respectively. One-way ANOVA with the dependent variable of time-deviation and the factor of condition revealed no statistically significant difference on this potentially more sensitive measure (F (2,21) = 0.042, P= 0.96). However, the Levene test for homogeneity of variance (implemented in SPSS version 16) returned a positive value (p =.047). As can be seen in figure 2.3, the standard deviation associated with the relevant-instructions-last condition was approximately double that of the other conditions. The non-parametric Kruskall-Wallis Test was therefore performed which also was consistent with no statistically significant differences between the conditions in terms of time deviation H(2) = 0.68, p>0.71.

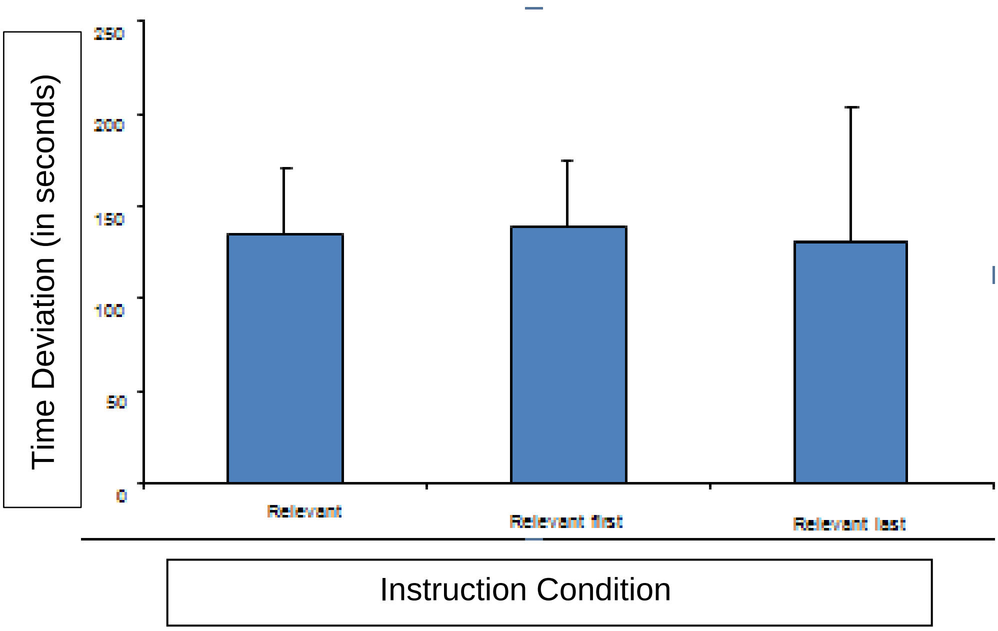


Figure 2.3: Time deviation scores of the participants under the different instructions

The results so far suggest that the hypothesised negative effect of irrelevant instructions was not detected either in terms of the number of games played or the time-deviation score.

Selection of the principal GET outcome measure

To help derive the best score from the GET the correlation between the games completed score and the time-deviation score was calculated. If one score is highly predictable from another, there is a redundancy in using two scores. For this purpose, data from all participants across the instruction conditions were used. It seems likely that there would be a significant negative relationship between these two variables; if a participant only played, say, 3 games, their time deviation from the 8-game/37.5 second ideal would inevitably be large. To take into account the skewed distribution of the score a non-parametric Spearman's correlation coefficient was carried out, which ranks the (ordinal) data and then applies Pearson's equation. There was, however, no statistically significant relationship between the number of games played and the participants' time deviation scores, *r*(19) = -0.26, p(two tailed)

=0.25). To examine whether, as seems likely, this is due to the high frequency of ceiling scores for the points total, the correlation was repeated with these scores removed. When participants who were performing at ceiling were removed from analysis then the number of games played was significantly correlated with participant's time deviation scores $r(14) = -0.67$, p(two-tailed), $p = 0.009$) in the hypothesised direction.

The correlation results above suggest that, aside from the effect of ceiling scores, the number of games played was strongly related to the time-deviation score. For that reason, and given the preferable distribution of the time-deviation score this was taken as the primary outcome measure of the GET.

Checking the clock in the GET

During the GET, participants had the option of checking how much time was left. Across all participants clock checks occurred an average of 7.53 times (SD 4.62; Range 1- 21) during the 5 minute task. Clock checking was not significantly modulated by instruction condition ($F(2, 29) = .20$, $P = .82$) and was not significantly correlated with performance as indexed by the time-deviation score (Pearson's $r(32) = -0.21$, $P = 0.25$)

Discussion

Taken together the results from Games Tried and Time Deviation provide no evidence for a detectable effect of instruction load in the hypothesised direction. Before turning to the instruction manipulation, it is worth examining what the basic performance of participants reveals about the GET in terms of its likely validity and psychometric properties. There are at least two simple possibilities that would challenge its validity as a measure of "goal neglect". The first is that participants simply did not remember what they were supposed to be doing and just played the first computer game for 5 minutes (i.e. they forgot rather than neglected the goal). The second would be that participants did remember the main goal but, due to unfamiliarity with the keyboard, menus etc., were not able to switch from the first game. The results show, however, that every single participant did switch tasks (the minimum number was 3 the most common was 8) and did check the clock at least once, generally more. This, in itself, strongly suggests that the goal of switching was remembered. None of the participants had difficulty remembering instructions before or after playing the GET.

A test that does not discriminate between individuals from the relevant population is not useful. As might be expected, in this healthy sample, there were strong ceiling effects on the number of instances of game play by participants. The sensitivity of this measure could be potentially improved in a number of ways including running the task over a longer duration or imposing a dual-task demand, for example. However, the timing-deviation score showed a useful distribution, even with a 5-minute version of the task. Whilst it would be expected that people are not perfect at estimating 37.5 intervals, the point is that some are better than others and that this may be related to their ability to keep the main goal of switching actively in mind. The question is whether the GET difference tells us anything useful about performance in everyday settings and tasks.

Recent work using another measure of Goal Neglect had shown that this phenomenon could be observed more frequently when the totality of the “task set” that participants needed to encode was increased (Duncan et al., 2008). It is not difficult to imagine how having 10 things to remember would make any one of those things more prone to being overlooked. Here, therefore, the effect of adding accurate but irrelevant instructions to the GET was examined. This produced one statistically significant and entirely contrary to hypothesis difference, in which participants first exposed to irrelevant instructions and then given the key instructions outperformed the two other groups in terms of the number of games attempted.

Caution needs to be applied in interpreting the Games Tried score effect due to the small sample size, the possibility that the groups were unbalanced in terms of the relevant ability despite broad NART matching, the presence of the ceiling effect, and the lack of a clear hypothesis as to why irrelevant instructions should improve performance. The results cannot be taken as evidence that the hypothesised effect of instruction load reported in Duncan et al. (2008) would not be detectable on the GET with a larger and better-matched sample. However, the primary interest here was whether it would have a *substantial* impact on the likely sensitivity of the task that would inform its use as an outcome measure. The best interpretation of the manipulation in this respect is that the GET seems rather robust to differences in instructions. Given the complexities that might arise from giving complex irrelevant instructions, this lack of difference suggests that the simplest, *Relevant* condition instructions are employed in future work described in this thesis.

Here too, the relationships between GET performance, the Cattell, CFQ and other measures will be examined within larger samples.

Study two

Methodology

Participants

Given no effect of instruction, the groups from study 1 (n = 20) were collapsed to increase power for examining correlations. An additional 11 participants were also tested using only the relevant instructions. The combined sample was made up of 31 people (twenty female) with a mean age of 64.83years (SD 4.51). Procedures were the same as in study one.

Measures

Measures administered to participants in this study were the Cattell Culture Fair (Cattell & Cattell, 1960), A Feature Match Test (developed and kindly shared by Russell Thompson and John Duncan), the Cognitive Failures Questionnaire (CFQ) self-report version (Broadbent, Cooper, Fitzgerald & Parkes. 1982) and the National Adult Reading Test-Revised (NART: Nelson, 1991). These additional tasks were included to help characterise the GET and ultimately to be used together with the GET as potential outcome measures for the training studies described in this thesis. The Multiple Errands Task (Shallice & Burgess, 1991) was also used but as it was with a separate sample will be described at a later point separately. The procedures were as in study one.

<u>The Games Evaluation Task (GET)</u> was the same as described above.

<u>The Cattell Culture Fair</u> (Cattell & Cattell, 1960) is a measure of fluid intelligence. It consists of four different sections, which are time-limited: series completion, odd elements, matrix completion and dot task. The test does not require general

knowledge or depend on linguistic ability. Participants are not told how long they have to complete the tasks. Such measures form strong correlates of *g*, an estimate of general intellectual ability. Research suggests that the Cattell may be particularly dependent on the operation of a fronto-parietal brain network that is active in many attentionally demanding tasks (Duncan and Owen 2000; Duncan, Seitz et al. 2000). Work also suggests that Cattell performance is negatively correlated with the frequency of goal neglect observed in some experimental tasks (people with lower scores show more goal neglect) (Duncan 1995) and that it can explain much of the variance seen in the performance of traditional ‘executive’ tests in patients with frontal lesions (Roca, Parr et al. 2010). However, there is also evidence that another measure of goal neglect, the Hotel Test (Manly, Hawkins, Evans, Woldt & Robertson. 2002), appears to reflect a rather independent factor not fully explained by differences in fluid intelligence (Roca et al. 2010).

The Feature Match task is a computerised measure, developed by Thompson and colleagues (Duncan et al., 2008) designed to elicit goal neglect. In each trial, participants see a pair of numbers appear on the screen, each number being within a coloured shape outline (circle, square or triangle each potentially appearing in red, blue or green). Participants must select the appropriate response according to a series of rules (e.g. if both numbers are surrounded by the same shape but these are of a different colour, a key corresponding to the higher number must be pressed; if the shape differed but the colour was the same, again, the key corresponding to the higher value should be selected but; if both shapes and colour were the same, no key should be pressed). A low score indicates good performance – the score given in this chapter reflects the number of incorrect answers a participant makes during the task.

The Cognitive Failures Questionnaire (Broadbent et al., 1982) was developed to examine the frequency of everyday cognitive slips in the healthy population. It asks people to rate how frequently events like walking into a room and failing to remember what you have come for occur. A high score on this measure means that there are more cognitive slips reported.

The National Adult Reading Test (Nelson, 1991) requires participants to read a series of words with irregular letter to sound correspondence (e.g. gaol, i.e. words that you have to know, you cannot just work out). In the general population this test correlates well with general intelligence but, unlike other capacities required in IQ test, this knowledge is relatively resistant to the effects of brain injury. Hence the NART is most commonly used to estimate pre-morbid IQ and to look at discrepancies with current intellectual function. In this study it is therefore difficult to make a specific prediction and its use is primarily to characterise the sample. Here the number of errors is reported.

Results

Examining the distribution 45.16% of participants, (n=14) scored the maximum 8 points on the GET with approximately 67.74% (n=21) scoring 7 or 8 points. Scores were skewed towards ceiling (Skewness = -.741, Kurtosis = -1.032) in the number of games played but with the time deviation score, the main and most reliable variable of the GET, as would be expected, a more normal distribution was observed (Skewness =.747, Kurtosis =.176). Table 2.1 below shows the scores for the participants in the variables that were assessed and used to correlate with the GET.

	Min	**Max**	**Mean**	**Std. Dev.**	**Skew**	**Kurtosis**
Feature Match (number of errors)	11	56	24.4	10.56	1.46	3.08
CFQ (number of self-reported cognitive slips)	18	63	35.64	9.59	0.84	2.03
Cattell Culture Fair (IQ)	76	118	101.05	8.97	-0.5	1.95
NART (number of errors)	1	24	10.05	6.85	0.63	-0.77
Age	55	70	66.86	4.22	-2.22	4.55

Table 2.1 Distributions of Feature match, CFQ, Cattell, NART and age

Two distributions (Feature Matching and age) gave rise to concern about the use of parametric statistics. As shown in table 2.2 the GET timing score, as would be anticipated, was negatively and significantly correlated with fluid intelligence, as measured by the Cattell (r=-.35, p(one-tailed =.02) - participants with lower IQ performed worse on the GET. A trend can also be seen in participants who report more everyday cognitive slips (CFQ scores) and less than optimal performance on the GET, (r=.29, p(one-tailed)=.06). There was little obvious relationship between GET performance and performance on the goal neglect measure, however the latter was also significantly related to IQ scores (r=.-35, p(one-tailed) =.03).

		Time Dev.	Feature Match	CFQ	NART	Cattell IQ	Age
Time Dev. (GET)	R	1	0.15	0.29	0.07	-0.35	0.13
	Sig.		0.23	0.06	0.36	0.02	0.25
Feature Match	R	0.15	1	0.29	-0.03	-0.35	0.06
	Sig.	0.23		0.06	0.44	0.03	0.39
CFQ	R	0.29	0.29	1	0.16	-0.22	0.09
	Sig.	0.06	0.06		0.22	0.12	0.32
NART	R	0.07	-0.03	0.16	1	-0.23	-0.07
	Sig.	0.36	0.44	0.22		0.13	0.37
Cattell IQ	R	-0.35	-0.35	-0.22	-0.23	1	-0.27
	Sig.	0.02	0.03	0.12	0.13		0.07
Age	R	0.13	0.06	0.09	-0.07	-0.27	1
	Sig.	0.25	0.39	0.32	0.37	0.07	

Table 2.2: Correlations of GET variables with other measures
Correlations significant at p(one tailed) <.05

None of the other measures reached significance in terms of correlation with the GET time deviation scores.

Discussion

The results of this study showed that, in a group of 31 participants from the older healthy population the novel GET produced a normal range of scores, particularly when timing was taken into account.

The results were consistent with a high GET time deviation score (i.e. less optimal performance) being associated with a higher degree of self-reported cognitive slips on the CFQ, and increased error rates on the Cattell measures. There was no obvious relationship between GET scores and the NART. It may at first glance appear anomalous that the GET was related to Cattell IQ but not to another measure of IQ based on the National Adult Reading Scale. The first point here is that NART scores were not available for all participants. Secondly, speeded fluid intelligence as measured by the Cattell is known to show significant declines in normal aging. In contrast, the NART is often used to estimate pre-morbid function in the context of head injury etc. because it is resistant to even quite severe deteriorations in intellectual function. Given that the sample here was from the older population, this lack of a relationship is not so unexpected.

The most encouraging findings from the current study in terms of the GET's potential validity come from the relationship with other measures. Most importantly, in terms of predicting everyday difficulties, the time-deviation GET score showed a relationship in the expected direction with participants' self-reported tendency to attentional lapses on the Cognitive Failures Questionnaire. GET scores also correlated in the expected direction with Cattell IQ scores.

The results are encouraging in suggesting some overlap of process between the different measures – that something that contributes to everyday cognitive slips is being taxed in the GET, that something required to solve complex novel pattern sequences in the Cattell is required for efficient GET performance. As outlined in the introduction, this was by no means a given. The nature, duration and characteristics of the tests differ markedly and it was entirely possible that, for example, the fine visual discrimination required in the Cattell but not in the GET could have eroded the strengths of correlations. Similarly the CFQ is not a direct measure of everyday slip but of people's recall of their everyday slip frequency. In that people may differ in their propensity to notice, remember and report such slips again it would not have been surprising were correlations more fragile than were actually observed.

Correction for multiple comparisons was not applied. It has been argued (Fan, Yao & Tong, 1996) that such corrections may be appropriate if there is a completely a-theoretical examination of a series of random variables but may be too conservative when, as here, the same unidirectional hypothesis is being addressed in the different comparisons. The general convergence in results here can be argued to strengthen confidence rather than require correction. However, of course, correlations based on relatively small samples should be considered provisional.

Study Three

Introduction

In this study the relationship between the GET and the multiple errands task (MET, Shallice & Burgess, 1991) was examined. As discussed, the MET is carried out in a real-life setting – a shopping centre and participants have to gather a number of pieces of information and buy a certain number of items whilst following a set of rules within a certain time. To date, performance has been scored by a rater who follows participants during the task. Like the GET and 6-Elements Test, participants must plan how to organise their time, keep this plan in mind and keep aware of the time. It cannot be assumed however that simply because of these common surface features that the tests will prove equally sensitive and/or correlate. It is possible that the cluster of different materials needed for the 6-Elements or the natural hurly burly of a shopping centre contribute to those tasks demands. Stripping these elements down to a simple computer interface may miss key aspects of the tasks. Similarly, whilst the GET games were designed to be engaging, for example, by adopting elements of popular commercial games (such as “Pong”), success is by no means guaranteed. Commercial game companies often spend considerable time and money adjusting the look, feel difficulty level and built in reward of games. Accordingly, this study continued to examine the psychometric properties of the GET and in particular, the degree to which GET performance correlated with existing measures of executive function. If a correlation was found between the two measures of organisation then it could be possible to carry out the brief GET task in place of the more time consuming MET in appropriate situations.

Methodology

Participants

A separate group of participants were recruited from the MRC- CBU panel of volunteers and were all older adults over the age of 65years. Participants were recruited over the phone and following this sent information on the study by post. The sample consisted of 19 older adults (mean age 69.04 years, SD 5.22). The participants described here are the same as those described in chapter three. It was of interest here, as mentioned, to examine any relationship between the GET and the modified MET, described below. Participants in the two studies described previously above did not complete the MET and so were not included in this study.

Participants were assessed in the testing labs of the MRC-Cognition and Brain Sciences Unit on the GET prior to carrying out the MET at a separate time in the Grafton Shopping Centre. On all occasions informed consent was sought and participants were told that they were free to withdraw at any point. Participants were given and honorium for their participation. Participants were debriefed following each assessment session and were given a summary of the tasks that they completed and a brief reasoning behind them. Participants were also at a later date posted feedback forms on their experience of taking part.

Measures

The MET

The MET was carried out in the Cambridge Grafton shopping centre, which was familiar to many participants. Informed consent was given by each participant prior to carrying out the MET

Before being given the instructions, participants were asked two questions (*"How efficient would you say you were with tasks like shopping?"* "*How well do you know this shopping centre?*") For the first there was a 10-point response scale with end points labelled ("1" – "hopeless", "10"- "excellent"). The second had a four-point scale ("1" - "never visited", "2" –"visited once or twice", "3" – "visit occasionally", "4" – "visit regularly"). Participants were read and given the instructions on a piece of paper (with a clip board if wanted) and given a pen, a plastic bag and a ten pound note. Participants were all asked to wear a wrist watch if they had one. If they did not then they were given a phone to carry with a clear clock display on the screen without having to press any buttons.

Participants were first told the geographical limits of the shopping centre. The task instructions were as follows:

> "In this exercise I want you to complete three tasks. The tasks are: to buy the five items listed on this sheet (*indicate and describe items on the sheet*); to obtain and write down five pieces of information (*indicate and describe items on sheet*); and to meet me here in 25minutes after I have said "…begin the exercise" and tell me the time. However, whilst completing this exercise you must obey the rules listed on your instruction sheet (*indicate and describe rules on sheet*).
>
> You must carry out all these tasks but you may do so in any order. You should spend no more than £6; although I've given you £10 you should spend no more than six. You should stay within the limits of this shopping centre. You are free to go upstairs if you like but you must not go outside any of the outside doors. No shop should be entered other than to buy something, so if you go into a shop it should be with the intention of buying something. You should not go back into a shop you have already been in, so if you've been into a particular shop you should not go back into it again. You should only buy items from shops, not stalls. You should buy no more than two items from Poundland. Take as little time as possible to complete this exercise without rushing excessively.
>
> Finally, approach me and tell me when you have completed the exercise.
>
> Is that clear, have you any questions?" (*Clarify any questions the participant has*)
>
> "Now tell me what you must do." (*Ensure participant is clear about what they must do*)
>
> "*Begin the exercise"* (*Start timing at this point*)

Participants were given an instruction sheet to keep with them, which listed the rules and tasks. At the end of the test participants were asked to rate two more questions: "*How easy did you find the task*?" using a five-point scale with weighted end points ("1" – "very difficult", "5"- "very easy") and "*How well do you think you did with the shopping task?*" using a ten-point scale with weighted end points ("1" – "hopeless", "10"- "excellent"). Instruction and scoring sheets are available in Appendix B. Participants were paid an honorarium for their time. There were two versions of the task to allow for re-assessment and participants were randomly given either version

A or version B. The instructions for both versions were the same. The only difference was in the information that was to be gathered and the items to be bought. A pilot study showed that the difference between the two versions was negligible.

Scoring

Scoring of the task was in line with the categorisation of errors specified by Shallice and Burgess (1991, examples given in Table 2.3 below): inefficiencies, rule breaks, interpretation failures and task failures. An error was marked as an "*inefficiency*" when the participant could have used a different method to achieve the task more efficiently. "*Rule breaks*" apply to both the rules of the task and also social rules, for example shouting at a shop assistant. If a subtask was misunderstood it was deemed an "*interpretation failure*" and a "*task failure*" was when subtasks – buying items or collecting information- are not finished satisfactorily.

Task Failure	Inefficiency	Rule Break	Interpretation failure
Does not tell the time at the end. Does not purchase an item or gather information necessary	Does not use a time-efficient route during task Does not ask for help from shop assistant when it would be more efficient to do so	Speaks to the instructor during the task Leaves the outside limits of the shopping centre. Social rule break	Purchases sponges instead of dishcloths Gets the phone-number of the wrong phone box

Table 2.3: Examples of errors made across these categories

There was no maximum number of errors that participants could make. If participants performed well it was possible for them to make no errors. As well as the number of errors made, a record was kept by the rater of the number of items bought (and after how many minutes), the number of pieces of information gathered, the order in which tasks were performed, the amount of money spent. It was also attempted to keep a record of the time participants spent planning.

Variables used for this study were "Tasks" – the number of tasks completed (max 10), "Mistakes" – number of rule breaks and errors (as per examples in Table 2.3), and "Time" which is the time taken to complete the task. The two variables of the GET - "number of tasks" and "time deviation" are reported here

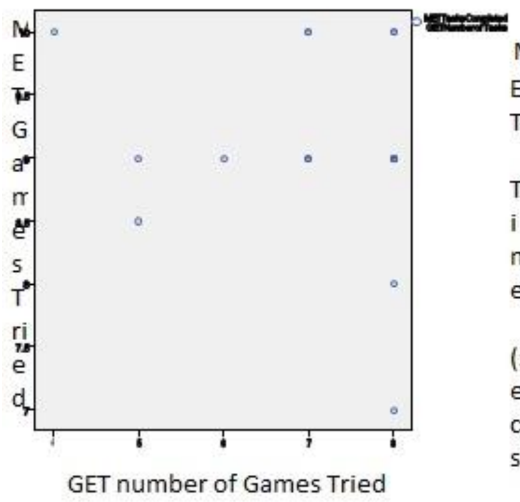
Results

A summary of scores in MET and GET can be seen in Table 2.4.

	Minimum	Maximum	Mean	Standard Deviation
MET tasks Completed	7	10	9.08	.75
MET Time Deviation (secs)	998	2295	1513.89	314.12
MET Mistakes	0	6	3.29	1.73
GET Tasks	4	8	7.11	1.24
GET Time Deviation (secs)	92	363	158.47	70.68

Table 2.4 Descriptive statistics of GET and MET variables

Scatter plots were initially used to examine the relationship between variables. The number of tasks completed and the time taken/deviation in both the MET and GET can be seen in Figures 2.5 and 2.6.



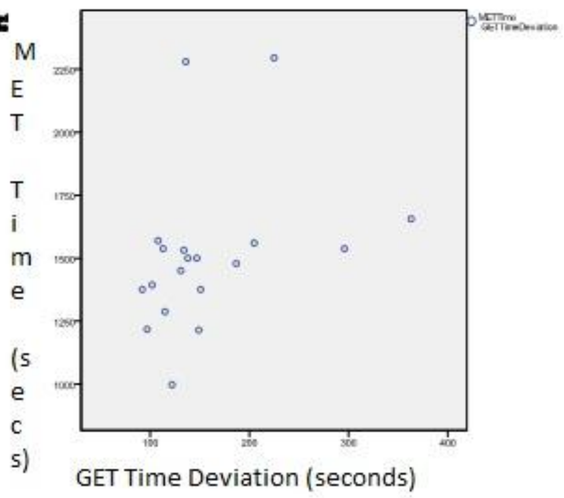


Figure 2.4: The number of tasks completed on MET and GET
Figure 2.5: The time deviation/time taken on GET and MET

Correlations between the MET and GET were examined and can be seen in Table 2.5. Spearman's correlation coefficient was used with the non-parametric data. All participants completed both the MET and the GET.

	MET Tasks	MET Time	MET Mistakes	GET Tasks	GET Time Dev.
MET Tasks					
MET Time	-.103				
MET Mistakes	-.37	-.12			
GET Tasks	-.15	-.43	.46*		
GET Time Dev.	-.24	.45	-.11	-.33	

Table 2.5 Correlations between MET and GET
*. Correlation is significant at the 0.05 level (2-tailed).

The only relationship that was seen in the measures was between the number of mistakes made in the MET and the number of games played in the GET, r=.46, p (two-tailed) <.05. No other relationships were seen between any of the other variables.

Discussion

The performance of this participant group was strong across both MET and GET, as shown above. There was a narrow range of tasks completed for both (MET 7-10, GET 4-8) with scores heavily skewed towards ceiling. From examining the relationships between the MET and the GET it can be seen that there is not a strong correlation between the measures. The only relationship was seen between the mistakes in the MET and the number of tasks in the GET. This indication of poorer performance on the MET being related to performance on the GET gives an idea that there may be similarities in what is being measured if the sample was bigger to examine this fully. It could be useful to have a single variable as a measure of success on the MET which is not the case at present, as discussed further in chapter three. As participants were rather good at both measures and a sample of 19 is a relatively small number with which to examine correlations it is still premature to conclude that they are measuring different things. Both desk-top and real-life measures give varying information and so complement each other, and so it might always be useful, when possible, to carry out both measures.

General Discussion

The main aim of this study was to establish the likely value of the GET as an outcome measure for a training study. The results in that respect are positive in terms of the internal and external validity checks and the sensitivity of the time-deviation score. It is unfortunate that the instruction manipulation did not return a clear answer on what form of instruction might produce the *most* sensitive test. Given the numerically greater ‘disorganisation’ shown by participants who received irrelevant as well as relevant instructions, however, it seems most useful to retain this feature of the task, perhaps distributing the relevant instructions not just at the beginning and end, but interspersed (and perhaps therefore more ‘hidden’) among the irrelevant instructions.

The work of Duncan and colleagues (i.e. Duncan et al., 1995) has consistently shown that goal neglect on their fluid intelligence tasks is more prevalent among people who perform poorly on the Cattell Culture Fair. This was shown here on the increase of cognitive slips on the CFQ and errors on the Cattell with poorer performance on the GET in study two. Unfortunately, no relevant relationship was seen between the MET and the GET and so there is a possibility that the two measures may be looking at differing constructs, although this is yet unclear. A difficulty that is seen with the MET is whether in fact the rater can capture everything that is happening, and also if it could be done blind. The next chapter will provide the methods for doing this.

At this early stage of development of the instrument it is difficult to make too many predictions about the future clinical use of this measure. Results to date, when viewed prudently, show that the GET could be a promising new measure in reflecting everyday cognitive slips in an older population. The possible use of the GET in a clinical population needs further testing before it can be confidently used as an executive measure. If the measure was shown to be an accurate estimate it would be a clinically useful tool as it is easy to administer, brief, requires little set-up and should be useful in a variety of age-groups and conditions. Further testing of this instrument is justified in order to have a more complete picture of the psychometric properties of the GET as well as age-related norms before being confident about the instrument's predictive validity as a measure of executive functioning. Correlations seen to date need to be replicated in a larger sample and may be a reflection of the sensitivity rather than the specificity of the measure. Further work with a larger sample would allow more intricate analysis of GET scores using, for example, regression analyses. Similarly, with due regard to practice effects in particular tasks, the effect of increasing irrelevant instructions could be considered more systematically using within-subject designs or much more powerful between-subject designs. Overall it has been shown that the GET could to be used as a fast, easily accepted and potentially sensitive measure of disorganisation.

Chapter Three:

Using video ratings to assess multitasking performance in a naturalistic paradigm.

Abstract

Multitasking measures, in which a series of tasks must be completed within a naturalistic setting not fully under the experimenter’s control, have been shown to be more sensitive than traditional measures in detecting organisational problems in people with subtle difficulties in executive functioning. There are a number of drawbacks to such tasks however. They can take a considerable time to administer and are demanding in terms of examiners noting and recording all relevant aspects of performance. They require sometimes subtle judgments (e.g. whether a rule has been infringed). This potentially leaves them more open to subtle bias due to examiner leniency, prior knowledge of the participant’s status (e.g. patient vs. control), and experimental condition (e.g. post-rehabilitation vs. control). One method that could offset these limitations is to video record performance such that it can be accurately scored off-line by one or more raters who are naïve as to participant/condition status. Accordingly, the use of a video ratings is investigated in this chapter with nineteen participants (mean age 69.04 years) who completed a Multiple Errands Task (MET) while wearing a body-worn camera. Their performance was scored “live” as well as by an independent rater who had only access to the combined video and audio footage of the task. Significant relationships were seen on all variables of the MET between the live and video ratings. The inter-rater reliability of the measure appears strong. This chapter provides initial support for the use of a video rater for participants who have carried out an MET.

Introduction

Psychologists often try to develop highly controlled tasks to isolate a particular cognitive capacity and to minimize the effect of different prior experience by using novel, abstract materials -the Wisconsin Card Sorting Test is a good example. Here people are asked to sort a pack of cards showing shapes of different colours in different groupings according to logical rules (by colour, shape or number of items; Heaton, 1981). Every so often the participant is asked to switch the rule. The task is conducted under quiet one-to-one conditions with the examiner indicating when to start and stop the task. From this, inferences are drawn about people's ability to hold on to and switch mental set with presumed predictive validity for their abilities in everyday settings that require these skills. It has been noted by a number of authors (e.g. Shallice & Burgess, 1991) that such traditional desk-top measures may in fact be rather *insensitive* to executive problems that are manifest in normal situations. They argue that when you reduce a task to assess a specific capacity you throw out many features with which such people may struggle. A key feature of everyday situations is that we generally have *multiple* goals which, due to our capacity to only complete one at a time, are in competition with each other (planning tomorrow's packed lunch is in competition with doing the washing up which is in competition with watching the TV). This delay in being able to act on a goal can lead to it being forgotten. Similarly, everyday life contains many habitual triggers for actions that we may not even consciously intend to complete (I may go into a shop to buy eggs but return with some washing powder that was on special offer and forget all about the eggs). Particularly where there is insufficient time to complete all of one's goals it is necessary to prioritise and review 'on the hoof' in a way that takes into account the opportunities and barriers that one encounters. Accordingly researchers have attempted to develop tasks that build in rather than exclude these features.

The development of the MET

In 1991 Shallice and Burgess developed a measure, the Multiple Errands Test (MET), with this aim in mind. Participants were asked to complete a series of goals within a given section of a London shopping street. These included buying specific items and finding out information. Some general knowledge and inference was required (e.g. how you could find which part of the UK had been hottest on the previous day). Participants also had to comply by task-specific rules (such as not re-entering a shop) as well as socially normative and legal rules (e.g. not insulting shop staff or stealing items). Planning was required to develop a strategy likely to complete the tasks within a given time (e.g. ordering the tasks to minimize the distance that needed to be walked) and this plan had to be held in mind and, if necessary, updated over the period of performance. The participant's behaviour was carefully recorded by an examiner who followed them at a distance, who was also ready to intervene if rule breaking became problematic! Shallice and Burgess (1991) showed that, in three patients with frontal lobe deficits who performed well on tests of IQ, perception, language and cognition, the MET elicited the types of errors that were apparent in their everyday lives. They committed a large number of rule breaks and developed inefficient strategies. They had problems with the task because it focused on areas of organisation and managing multiple sub-goal tasks that had been shown to be difficult for them. The task also requires motivation and memory, and controls do not often score at ceiling. From this paper it is clear, despite the small sample involved, that an MET, multiple sub-goal-type task, might be an appropriate measure to use with people who score within the normal level on other psychometric tests.

Although it does not make reference to the MET, Boyd and Sautter (1993) developed an unstructured task that could be used with a brain-injured population. Participants' (n=31) route-finding ability was assessed in a hospital setting where the raters looked at task formulation, strategy application, their dependence on cueing and the detection and correction of mistakes. In this task, as with the MET, participants were free to complete the task in any number of ways. Unlike the point-by-point recording of the MET, however, a Likert scale was used to rate overall performance. Although the authors cite good inter-rater reliability, this single score raises interesting questions about the precise criteria used by different raters. Burgess et al. (2000) tackled this issue in a revised MET that attempted to break down performance into theoretically separable categories. For example, asking people about their plan before and after performance allows planning to be separated from memory for, and tendency to follow, the plan.

Since the initial development of the MET a number of versions have been reported for use with clients with more severe difficulties than those reported by Burgess and Shallice (1991). A hospital-based version of the MET (MET-HV) was designed by Knight, Alderman and Burgess (2002) for clients with behavioural problems (where there were safety issues in a public environment as well as such settings proving potentially too distracting). They assessed 20 healthy controls and 20 patients on this simplified version and found that patients showed most problems with subtle planning, prospective memory and when a task was “ill-structured”. Patients broke more rules, made more mistakes, achieved fewer tasks and were more reliant on others to help them. They found that this version of the MET correctly classified 85% of patients. Knight et al. (2002) reported that in the MET-HV, particular failure to achieve tasks, combined with responses on the Dysexecutive Questionnaire (DEX; Burgess, 1996) gave an overall indication of the presence and severity of behavioural difficulties. They raise the important issue that failure on MET style measures can occur due to neglect of the plan but also frank amnesia for the tasks. In this respect they noted that memory impairment on the RBMT was associated with memory errors in the MET-HV. Measures of general intellectual function were, however, rather unrelated to MET performance in this study.

In a further study, Rand et al. (2009) assessed nine patients post-stroke, twenty healthy young controls and twenty healthy older adults on a virtual MET (VMET). The participants carried out a real and virtual MET. They found that the virtual version of the task differentiated between older and younger participants and between healthy participants of all ages and the clinical group. They promote it as a safe way in which to experiment, make mistakes and learn but it is not clear how much generalisation of gains there may be from learning in a virtual environment and transference of gains to everyday life. There has been an increase in the potential use of virtual multitasking measures in recent years but the research to date has mostly focused on healthy adults (e.g. Logie, Trawley & Law, 2011; Jovanovski, Zakzanis, Campbell, Erb & Nussbaum, 2012). Logie and colleges investigated the use of a multitasking tool with 153 undergraduate students (mean age 19.59, SD 2.43 years). They used a simulated shopping mall – the Edinburgh Virtual Errands Test (EVERT). The task was relatively short (8 minutes) but required a substantial amount of pre-planning, navigation and memory as all task requirements had to be remembered. Participants were required to carry out eight tasks, some of which had two parts and some, which were time dependent. The authors indicate the importance of pre-planning in success of the test. They suggest the possible use of the task as a test of multitasking but as the sample used was very high functioning it is unclear if this would be an appropriate measure for a population that was different to that measured. Jovanovski et al. (2012) also looked at using a virtual environment but they used a virtual city (Multitasking in the City Test; MCT) which was different to the shopping centre used previously. They had a smaller sample (n=30) of undergraduates that were the same age as those in Logie et al (2011). Their task was longer, at fifteen minutes, but gave participants very few explicit rules, unlike Burgess and Shallice (1991) who had a set of rules. Participants were also required to make a plan before beginning the task and were given unlimited time for this. The environment itself resembled a real town with a post-office, shops, doctor, restaurant, participant's own home etc. the authors found that although fifteen minutes was given, the majority of participants finished the task in five minutes. There were fewer tasks to complete than in the original study by Shallice and Burgess (1991). They found that only three participants followed the plan that they had made prior to beginning the task but that those who planned well performed better. The authors suggest using this measure (MCT) as an estimate of executive functioning. They report significant correlations between it and the Modified Six Elements Test (Burgess et al.,1998) and the Trail Making Test (Reitan, 1979) but the sample used is small and not representative of the target group. These three studies give support for the potential use of virtual measures in a clinical setting, although further research is essential before they can be confidently used in place of existing reliable and valid measures especially with regards its appropriateness with older adults.

For patients with lower IQ but who can be safely assessed in a public setting Alderman, Burgess, Knight and Henman (2003) looked into a simplified version of the original MET (MET-SV). Patients had an IQ post-injury, as measured by the WAIS-R FSIQ, of 84.1 (SD12.7) and the majority (75%) were categorised as very severe traumatic brain injury, as determined by duration of post-traumatic amnesia and duration or depth of coma when first admitted to hospital. This version of the MET had simplified task demands, and more concrete rules, and had more time available for task completion. They reported on the basis of data from 46 controls and 50 patients that the key MET-SV variables discriminating the groups were the number of rule breaks and task failures. Patients made approximately three times more errors and significantly (19 times) more social rule breaks that healthy participants. Age was again negatively associated with performance. As with previous studies it was noted that some of the patients who struggled with the MET had performed relatively well on traditional desktop measures of executive function.

Empirical validation of METs

As discussed, it has been argued that the unstructured nature of METs and the multiple competing goals make them more likely to be sensitive to everyday dysexecutive problems. Examples have also been presented of patients who perform well on traditional executive tasks but perform poorly on MET and everyday life situations. Such examples do not necessarily mean that this is generally the case. In an empirical test, Burgess et al. (1998) assessed 92 mixed aetiology patients and 216 controls on ten different measure of executive functioning in order to test their relative ecological validity. Responses from relatives/carers of participants on the DEX were used as the measure of 'real-life' function against which the predictive validity of various executive tests was judged. From the DEX scores five potentially distinct factors of executive function were derived. Performance on the traditional executive measures Cognitive Estimates, Verbal Fluency and the Wisconsin Card Sorting Test (WCST) were shown to be rather weak predictors of any factor and not to differentiate well between patients and controls, which replicated an earlier finding from Burgess et al. (1998). The authors point towards using a multiple errands task to observe behaviour, when possible. They find it to be the most revealing of measures, although indicate that it may not be appropriate for use in every situation, and they suggest that the Zoo Map, which involves planning a route around a zoo according to a set of rules, might be an alternative in these situations. Manchester, Priestly and Jackson (2004) also reported poor sensitivity for Word Association and the Cognitive Estimates. In the Burgess et al. study (1998), the 6-Elements test, which as discussed is a form of desktop based MET, was most closely related to intentionality factors. A later study in 2000 by Burgess, Veitch, de Lacy and Shallice support the idea of fractionation in 60 controls and 60 patients. They showed support for the use of certain executive measures in measuring different cognitive functions, dependant on which of the five cognitive factors they wish to target – for example using MET to target intentionality.

The METs outlined so far have been based around shopping tasks. Performance in other ecological everyday situations has also been examined. Chevignard et al (2008) examined meal cooking performance in 45 patients with dysexecutive syndrome after a head injury. They found that patients made significantly more mistakes and took longer to complete the task than controls. Over half of patients did not complete the task and half displayed dangerous behaviour during performance. Their results showed that performance was related to injury severity and age- with older patients performing worse- and that more than half of participants displayed dangerous behaviour while undertaking the task. Although the authors linked performance with ‘executive function’, in this and other METs factors such as memory for instructions, pre-existing knowledge and skills, perception and motor function are important issues in interpretation. It may be a useful tool in situations when it might not be possible to leave the clinical or home environment with clients. Clinicians are increasingly looking for more interventions and assessments that are transparent to patients and that are representative of their needs and generalise to different settings and areas of a patient’s life. The MET lends to that possibility and has the potential to be accepted by both patients and clinicians.

Study aims

In summary therefore, there is good evidence that the complex, unstructured, multiple competing goals nature of MET, in that these mimic real life situations, can make the test better predictors of dysexecutive everyday errors than highly reduced/abstract desktop tests. A drawback to these tests is that this very complexity can make it difficult to clearly interpret errors (was a task omitted because of goal neglect, frank amnesia for the instruction or for motivational reasons?). In addition, the tasks are lengthy to administer. Finally, important limitations that this chapter seeks to address are that MET are hugely reliant on the attention of the administrator in noting what occurs and when and that, unlike many paper and pencil or computerized tests, the examiner’s report provides the only available record. This is important because mistakes cannot be corrected, differences in scoring propensity and the possibility of unconscious bias. Here therefore the practicality and outcome of using first-person video recordings from a device worn by the participants was examined. Specifically, contemporary ratings from a “live” MET examiner were compared with those from an independent rater who scored during off-line viewing of video footage. The key questions were:

1. To what extent did the independent raters' scores accord with those of the live examiner?
2. Did the video recordings allow greater accuracy in some respect (e.g. events were noted that were missed by the live examiner, timings were more accurate etc.)?
3. Did the video reduce accuracy in some respects compared with the live rater (e.g. viewing perspective was non-optimal, technical glitches occurred etc.)?
4. Was the accuracy of the video recordings such that future examiners using this technology would be able to keep a watchful eye on participants' safety etc. but not concern themselves with live scoring?

To these ends, a variant of Shallice and Burgess' (1991) MET was developed to fit the layout and shops of Cambridge's Grafton Shopping Centre. Participants drawn from the older healthy population (the key groups investigated in the other chapters) were given instructions by an examiner who then followed them during the task performance recording their behaviour on a score-sheet. The participants wore a hidden video camera designed to continuously record sound and video during performance from the participants 1^{st} person perspective (i.e. the participants themselves were not seen but their voices, location, arm actions etc. should be visible). The choice of using a hidden video camera was to prevent attention being drawn to the participants, others behaving in unusual ways and/or having concerns about why filming was taking place. This required close consultation with the ethics committee and MRC Regulatory Statutory Support Unit about the legality of this type of filming and how the recordings were used.

Methodology

Four participants (one male; mean age 28.5 years, SD 10.47) took part in the pilot. For the main study 26 participants from the older healthy population were recruited from the MRC Cognition and Brain Sciences Unit Volunteer Panel and carried out the video MET. Technical problems including poor video quality/angle, obscured camera angle, muffled audio and/or problems with the battery meant that seven of the original 26 videos could not be used. This left a sample size of 19 participants (13 male) with a mean age of 69.04 years (SD 5.22).

Ethical approval for this study was granted by Cambridge Psychology Research Ethics Committee (CPREC 2009.53). Permission to carry out the task in the Grafton Shopping Centre was granted in writing by its management. All participants gave informed consent both before and after taking part, were fully aware of the video camera and were reimbursed for their time. Consent was taken again following completion of the task in order to confirm that they were still happy for their data to be used as the task may have been different to what they had anticipated. CCTV footage is routinely recorded throughout the shops and public areas of the Grafton Centre and there are prominent signs that this is the case. Consultation with the Ethics Committee and the MRC Regulatory Support Centre suggested that people are aware that they are quite frequently videod either by CCTV or incidentally in the background of personal video/phone recordings. Crucially, shoppers and shop staff who appeared in our video recordings were incidental to the aims of the study.

Pilot

The two versions of the MET developed here (see below) were first piloted on four younger participants Some difficulties were identified, for example one shop that was to be used was closed down and there were difficulties with positioning the camera. The timings of the two versions of the task were compared and seen to be of equal length and difficulty. Feedback from participants indicated that they found the task to be a reasonable challenge within the time and budget allowed. They also stated that the task would prove more difficult if it was to happen at a busier time as there would be longer queues and more distractions. As a result when possible the task was carried out before noon. Pilot data also indicated that the two versions were of equal difficulty and length.

The MET

The MET was carried out in the Cambridge Grafton shopping centre. It is a centre which was familiar to many participants. Informed consent was given by each participant prior to carrying out the MET. Before starting the participants donned the body-worn camera (CCD Button Camera). Initially a button camera was used attached to the body strap of a sports bag but camera angle and battery life were unreliable. A second camera (Swann Pen Cam™) was sourced and used. This widely available commercial product had wide-angle lens built into the lid of a pen, such that when the pen was clipped into a breast pocket it collected a stable, first person perspective view (see figure 3.1). Participants were asked to wear clothing with a breast pocket if possible. If not, solutions were improvised (such as clipping the pen to a cross-body bag strap).

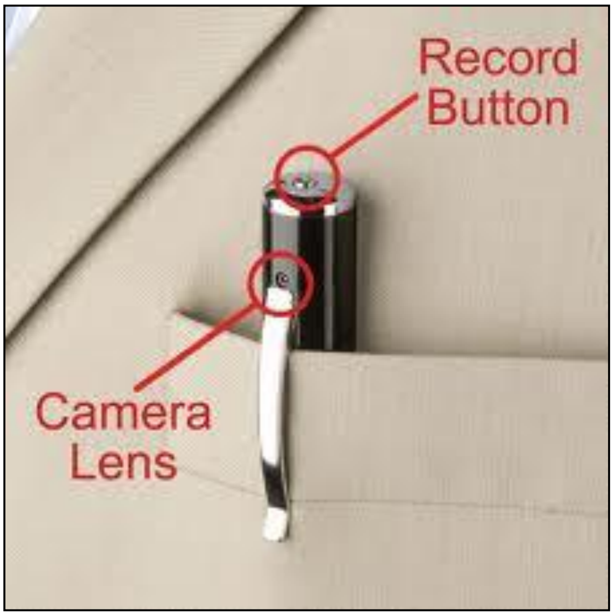


Figure 3.1. The covert video recorder pen used in the MET.

Participants were asked prior to beginning the task about their self-rated efficiency at shopping generally and about their familiarity with the shopping centre. The task instructions were read and then given on a piece of paper to participants to carry with them during the task. They were also given a pen, a plastic bag and a ten pound note. They either wore a wristwatch or carried a mobile phone with a clear time display on it. The limits of the shopping centre and the task instructions were then read and after this participants were free to ask any questions they had prior to starting the task. The instructions were the same as those reported in Study 3 of Chapter Two and can also be found in Appendix B1 and B5. Once the task began - when the participant is told that they may begin- participants were not to speak to the rater until the end of the task when they told the rater the time, indicating that they had finished the task. At this point the participants were asked to rate their own performance on the task and to state how easy they found the task. Participants were paid an honorarium for their time. Two versions of the task, of equal difficulty were available for use. As mentioned above the pilot study showed that the difference between the two versions was negligible. This task was to be used as an outcome measure for later studies and having available a parallel version would help reduce the effects of practice and make the task more challenging and interesting for participants on re-assessment as they would be less likely to be able to use strategies used previously.

Scoring

The task was scored in the same way as described in Chapter Two – using the same categorisation of errors described by Shallice and Burgess (1991): inefficiencies, rule breaks, interpretation failures and task failures. An "*inefficiency*" happened when a participant did not use the best way of achieving a task. A "*rule break*" was when any of the rules listed were broken by a participant. An "*interpretation failure*" was when a task was not completed properly because the participant did not understand correctly what needed to be done. A "*task failure*" was when a subtask was not adequately finished or when it was not attempted at all. A participant received one point for every error they made – i.e. a lower score is a more efficient completion of the overall task. A record was kept of the order and timings of when items were bought and information was gathered as well as amount of money spent and the time taken to complete the task. An attempt was made to also keep a record of the time participants spent planning. This proved difficult, as will be discussed later.

Raters

The live ratings of the MET were scored as the task was taking place and directly on completion of the task. Two raters scored the videos at a later time. The two video raters had no contact with each other and were not aware of either the live ratings or the ratings of the other video rater. Video ratings were carried out at the rater's convenience in a quiet environment. The video raters were trained in the categories of errors etc. by the live examiner and used assessment sheets identical to those used in live scoring. Video raters were able to rewind, pause and replay any sections of the videos and to take their timings from the video clock.

Results

Timing

The participants were instructed to tell the examiner the time (from his or her watch) at which they had completed the test. The live examiner and video-raters recorded whether this had been achieved and the time that it occurred relative to the start of the test according to the stopwatch/video clock. Where participants did not remember to report the time the 'finish' point was set as 1500 (25mins) and it was marked as a task failure. On three occasions, the video recorder was switched off by participants before this point. In this case the video-raters recorded the time the video ended.

As shown in table 3.1, even when a margin of +/- 10 seconds was used to take into account small differences in when timing started, rounding up etc., the raters only all agreed in 3/19 participants cases. This was also where the greatest variance was in performance between participants. As might be expected agreement was higher between the video raters but still only occurred in about half of cases. This level of agreement suggests that the scoring criteria were understood by the raters and applied consistently where the events seen/filmed were unambiguous.

Partic'	Live	Video 1	Video 2	All agree	Video raters agree	Overall discrep	Video rater discrep	L-V1 bias	L-V2 bias	V1-V2 bias
1	1500	1552	1500			52	52	-52	0	52
2	2280	2290	2340			60	50	-10	-60	-50
6	1560	1560	1500			60	60	0	60	60
7	1538	1600	1560			62	40	-62	-22	40
9	1570	1560	1500			70	60	10	70	60
11	1376	1380	1380	Y	Y	4	0	-4	-4	0
12	998	1020	1020		Y	22	0	-22	-22	0
13	1215	1380	1224			165	156	-165	-9	156
14	1218	1260	1260		Y	42	0	-42	-42	0
17	1288	1300	1260			40	40	-12	28	40
20	1656	1700	1630			70	70	-44	26	70
24	1500	1500	1500	Y	Y	0	0	0	0	0
25	1394	1440	1440		Y	46	0	-46	-46	0
26	1376	1370	1380	Y	Y	10	10	6	-4	-10
27	2295	2280	2220			75	60	15	75	60

28	1532	1500	1500		Y	32	0	32	32	0
30	1538	1530	1500			38	30	8	38	30
31	1451	1440	1440		Y	11	0	11	11	0
32	1479	1480	1490		Y	11	10	-1	-11	-10
Mean (SD)	1513.89 (314.12)	1533.79 (303.03)	1507.58 (307.58)	0.00	0.00	45.79 (37.53)	33.58 (39.52)	-19.89 (43.71)	6.32 (38.08)	26.21 (45.02)
Agree				15.70%	47.3%					

Table 3.1. Time to completion (seconds) scores from three raters (live, video rater 1 and video rater 2) for each of the participants MET performance. Agreement rates are taken as being within +/- 10 seconds of the other rater's score. Overall discrepancy represents the difference between the highest and lowest value reported by any rater. Video Discrepancy represents the difference between the highest and lowest values returned by the video raters. The bias scores take into account the direction of a discrepancy: L-V1 bias = Live rater – Video rater 1; L-V2 bias = Live rater – Video rater 2; V1-V2 bias = Video rater 1 – Video rater 2.

Tasks

The variables that were used in the MET task were "Tasks" which was made up of both the number of items bought and the pieces of information gathered (max. 10); "Mistakes" which was the number of errors made during the task including inefficiencies, rules breaks, interpretation failures and task failures; and "Time" which is the time it took to complete the MET in seconds.

Rater	Measure	Minimum	Maximum	Mean	Std. Deviation
Live	Efficiency	2	10	7.21	1.619
	Tasks	7	10	9.08	0.75
	Time	998	2295	1513.89	314.115
	Mistakes	0	6	3.29	1.727
Video 1	Efficiency	2	10	7.21	1.619
	Tasks	6	10	8.63	1.065
	Time	1020	2290	1533.79	303.032
	Mistakes	0	6	3.21	1.475
Video 2	Efficiency	2	10	7.21	1.619
	Tasks	5	10	8.32	1.204
	Time	1020	2340	1507.58	307.434
	Mistakes	1	7	3.26	2.156

Table 3.2 MET variables and participant performance (n=19)

It should also be noted that the "efficiency" reported in Table 3.2 are subjective scores given by the participants and not therefore subject to examiner variability. As the focus of this study was not to look at relationships for example between efficiency and actual task performance it was seen as appropriate to exclude it in this case. Table 3.2 shows the range in performance of the five MET variables as scored by the live rater, first and second video raters.

Scores for the number of tasks completed (max 10) between the 3 raters for the 19 participants were examined. In terms of how best to address these data, measures of inter-rater reliability such as Cohen's Kappa are generally based on categorical judgments and take into account baseline probabilities of agreement (e.g. if two people judge whether a series of stimuli are "pink" or "red" each does so at a certain frequency that would determine the level of chance agreement: if both said "pink" in all cases the probability of 'chance' agreement would be very high). It has been pointed out that this can be a very conservative test that precludes the possibility that all the stimuli were, in fact, "pink". With a notional 10-point scale here very few exact correspondences would occur were the numbers generated at random *from across the scale.* However, as can be seen in table 3.3 below, the raters did not produce values from across the scale instead using between 3 and 5 points, all clustered around the 8-9/10 score level. This increases the probability of chance convergence. However, much as the possibility that all stimuli were pink in the above example, this analysis should be grounded in the reality that the ratings fall within this tight region because this is broadly the level reached by these healthy participants.

Given this, looking at the scores presented in table 3.3, there is perhaps a surprising lack of agreement in the number of tasks achieved by each participant. The raters' totals agreed in only 6/19 (32%) of cases between all three raters. The level was slightly higher between the two video raters (9/19, 47%). Examining these discrepancies, 10/13 (77%) was of 1 point or fewer but in two instances were 3 and 4 points. A likely case of non-random discrepancy would be if the achievement of some task was apparent to the live rater but obscured on the video. This would be consistent with the live rater's generally higher scores (see bias scores in table 3.1). However, in one of these larger discrepancies, the live-rater (9 tasks complete) was in better agreement with the first video rater (8 tasks) than the second (5 tasks) suggesting either a period of inattention in the final rater or something rather ambiguous about this participant's performance. The reverse pattern was apparent in the second substantial discrepancy, with the live rater and video rater 2 being in greater agreement and video rater 1 noting markedly less task achievement. Unlike the purchasing of items it is often less clear to the videos raters if certain items of information have been collected. Some information, such as the number of the public phone box are more obvious on camera than other pieces of information such as the number of mobile phone shops in the centre or the number of shops beginning with a certain letter. It is less obvious through the video if participants are taking note of the number of shops or not – this is a lot more clear live. Some participants also use different strategies such as asking for help or counting the number of shops using the centre floor plan. Some strategies are easier than others to detect. The correlation between the live and video rater 1 and video rater 2 were (Spearman's rho) 0.78 and 0.69 (P = 0.001) respectively. Video raters 1 and 2 correlated 0.64 (P = 0.001), with these values being likely reduced by the narrow spread of the scores in this healthy group. Overall, in terms of tasks achieved, therefore there was reasonable agreement between the video and live rating methods but the results suggest that, even when two people are looking at the *same* video clips, discrepancies do occur. This suggests that the reliability of the conventionally used 'live' method may be similarly noisy – a factor that has not been taken into account in previous studies.

Partic'	Live	Video 1	Video 2	All agree	Video raters agree	Overall discrep	Video rater discrep	L-V1 bias	L-V2 bias	V1-V2 bias
1	9	9	8			1	1	0	1	1
2	10	10	10	Y	Y	0	0	0	0	0
6	8	8	7			1	1	0	1	1
7	9	9	8			1	1	0	1	1
9	9	9	9	Y	Y	0	0	0	0	0
11	9	9	9	Y	Y	0	0	0	0	0
12	10	10	9			1	1	0	1	1
13	9	8	5			4	3	1	4	3
14	9	9	9	Y	Y	0	0	0	0	0
17	9	8	9			1	1	1	0	-1

20	10	10	8			2	2	0	2	2
24	7	7	7	Y	Y	0	0	0	0	0
25	10	9	10			1	1	1	0	-1
26	9	8	8		Y	1	0	1	1	0
27	9	9	8			1	1	0	1	1
28	9	8	8		Y	1	0	1	1	0
30	8.5	8	8		Y	0.5	0	0.5	0.5	0
31	10	10	10	Y	Y	0	0	0	0	0
32	9	6	8			3	2	3	1	-2
Mean (SD)	9.08 (0.75)	8.63 (1.07)	8.32 (1.20)			0.97 (1.06)	0.74 (0.87)	0.45 (0.76)	0.76 (0.98)	0.32 (1.11)
agree				31.58%	47.37%					
Max	10	10	10							
Min	7	6	5							

Table 3.3. Tasks achieved scores from three raters (live, video rater 1 and video rater 2) for each of participants MET performance. “Bias” scores calculated by subtracting each of video rater 1’s scores from each of the live rater’s scores, video rater 2 scores from the live rater scores and then video rater 2’s scores from video rater 1. If there is no particular tendency for a rater to score high or low, the mean of these values will tend towards 0. If, however, there is a tendency in one direction positive or negative values will be returned.

Difference between live v video 1 and video 2

Ratings of individual variables – “Tasks”, “Time” and “Mistakes” – between the three raters were compared to see if there was a correlation between their scorings. There was a significant relationship between all the ratings on all the variables of the MET p (two-tailed) <.05, as seen in table 3.3. All variables apart from Mistakes Live and Mistakes Video 2 were significant to p < .01. This suggests a strong inter-rater reliability of the task, even when scoring of the task is not live and immediate.

	Tasks Live	Time Live	Mistakes Live	Tasks Video 1	Time Video 1	Mistakes Video 1	Tasks Video 2	Time Video 2	Mistakes Video 2
Tasks Live									
Time Live	-.103								
Mistakes Live	-.367	-.117							
Tasks Video 1	.783**	.147	-.130						
Time Video 1	-.062	.966**	-.003	.182					
Mistakes Video 1	-.424	.191	.719**	-.337	.302				
Tasks Video 2	.688**	-.143	-.237	.635**	-.203	-.447			
Time Video 2	-.073	.967**	-.037	.160	.970**	.295	-.163		
Mistakes Video 2	-.098	.222	.503*	.169	.283	.576**	.124	.354	

Table 3.4 Correlations between raters and variables – Live rater, video rater 1 and video rater 2
**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

It was also shown that ratings between video rater 1 and video rater 2 were significantly correlated in all variables. As can be seen in table 3.4 there was no relationship between any of the other individual variables – scores were not predictive of each other.

Difference between live and video ratings

Video 1 and Video 2 scores were then collapsed to see if there was a difference overall between live ratings and one single video score – “Video (mean)”. Again, from this it was shown that live ratings were significantly correlated with video ratings p (two tailed) < .01. As demonstrated in table 3.5 below, all variables showed a positive correlation in live and video ratings, providing further evidence for the reliability of the method.

	Tasks Live	Time Live	Mistakes Live	Tasks Video (mean)	Time Video (mean)	Mistakes Video (mean)
Tasks Live						
Time Live	-.103					
Mistakes Live	-.367	-.117				
Tasks Video (mean)	.773**	.029	-.200			
Time Video (mean)	-.062	.975**	-.012	.026		
Mistakes Video (mean)	-.248	.202	.683**	-.059	.305	

Table 3.5 Correlations between live and video ratings
**. Correlation is significant at the 0.01 level (2-tailed).

Intra-class correlations were used to investigate the inter-rater reliability of the three main variables of the MET – tasks, time and mistakes to see if there was absolute consistency between raters. The correlation between the raters was as follows: r=.825, p<.001 for tasks, r=.997, p<.001 for time and r=.861, p<.001 for mistakes. This shows that there is little disagreement between the raters on the three main variables.

Discussion

From this reasonably small sample significant correlation was found on all three key variables by all three raters in live and video scoring of the task and there appears to be strong inter-rater reliability. Discussions between raters on what is meant by each of the "errors" listed were important and proved valuable, as was piloting the video scoring. An example of this was that Video rater 2 was scoring "asking for help" (by a shop assistant) as an "inefficiency" while Video rater 1 noted it as an efficient strategy. From this it was decided to keep a record of the number of times help was asked for but not to score it as an error. Positive results from this study indicate the possibility of using this method of scoring and assessment of the MET in subsequent studies.

The most common error made by participants was the final task that they had to do – telling the rater the time when they had finished- indicating a common difficulty with prospective memory for this group of participants. This the most difficult subtask as it was failed most often. Participants seemed to be so relieved to be finished the task that they forgot this final subtask. It would be interesting to see if this error would be as common if the subtask was to be moved to a different part of the MET – “After ten minutes tell me what time it is”. This subtask was at times difficult to detect from the video recordings as at times participants turned off the video before they told the time and so there was discrepancy between live and video ratings on this subtask.

Planning time was difficult to accurately measure in this task as some participants stood at the beginning of the task and made a plan as to how they would achieve their subtasks, while others jumped straight into the task and planned as they went along. Although it might have been of interest to get a measure of this it was not possible due to the differing manner of planning that participants used. Miotto and Morris (1998) showed that patients were not found to be any slower than controls indicating that having a score of “planning time” may not contribute over and above other types of score. Other authors have used planning times in their virtual versions of the task and have shown it to make a difference to performance so it could be an issue worth addressing in the future. It would be useful to see if participants were required to plan their route and their time and budget prior to beginning the task as in certain other studies (e.g. Logie et al., 2011; Jovanovski et al., 2012) if an improvement in performance would be seen

Certain difficulties came from undertaking such a task. The first was in recruiting adequate number of participants. As this was to be used as an outcome measure for the training study to be discussed in Chapters 5-7 it was offered to participants as an optional measure to undertake. As a result a large number of participants opted out of taking part due to time constraints or the difficulty of getting to the shopping centre in question. For those who did take part there was a large number who were not interested in coming back following their training to carry out the MET a second time in order to be able to use it as an outcome measure (discussed in a later chapter). A second difficulty came in carrying out the task in public. Because the task took place in a shopping centre (with permission from the centre's manager) the suspicions of the security staff were raised on many occasions. There were many different staff members working over the testing days. As participants generally followed a similar route, entered many of the same shops, were walking around carrying a clipboard and taking notes over many days and many sessions while being followed around the centre there was often a security guard in turn following the live rater! In order to minimise the potential disruption this might cause – stopping the task to explain to staff what was happening- it was decided not to use the clipboard and to do as much of the preparation and consent giving and payment outside the centre. Staff in shops seemed uneasy with customers carrying clipboards but had no problems with customers carrying around pages which potentially had shopping lists on. There were problems too, as mentioned, with the quality of some videos. Two different cameras were used and the second proved much better quality and easier to use. This pen stayed in place better, had better quality audio recording, was less conspicuous, and had better battery life. Most of the previous difficulties that were seen in the first camera were resolved with the second camera. Only one video from the second camera needed to be discarded and this was because the participant was carrying the instruction sheet in front of the camera lens.

With further investigation and validation it would be of interest to see if this assessment measure could also have the potential to be as a clinical intervention strategy. It could be scored in conjunction with the client in order to point out potential areas of difficulty and point to mistakes made and help to devise strategies to make performance of the task more efficient. Knight et al. (2002) suggested using the MET with clients to facilitate the process of goal-setting. Using the video may help clients who might have memory problems and may not be able to accurately recall their performance. It also could be used in a multi-disciplinary team to demonstrate to team members areas to work on with clients that may not have shown on traditional measures. It could also be used to clearly demonstrate examples of difficulties clients are having when handing a client over to another clinician (with client's consent). Clinicians would then potentially have a clear measure of progress over time.

As it is a task that can be used in different settings it would be useful to see if repetition in various settings would lead to improvements in other everyday settings. This would give clients opportunities to practice dealing with unexpected situations, devising efficient strategies etc. in order to achieve small goals set in a real-life but safe situation where they will be able to retrospectively review their own performance. A task such as this also allows clinicians the potential to taper the difficulty and demands of the task by gradually introducing settings that for example have more distracters and more potential for rule breaks or social interaction depending on the level of functioning and goals of the client. Various versions of the task already exist and have been validated, as previously discussed.

As well as being used as an intervention for people with organisational difficulties this device has the potential to be used with people with autobiographical memory problems- people who have difficulty remembering things that they have done during the day. It could be worn when patients have to attend important meetings or appointments and replayed later in the day. It would not be practical to wear the camera all through the day and it is often not necessary to remember everything one does during the day. Studies using the SensCam ™ device (now Vicon Motion Systems) have looked at this to some degree but as it does not have audio recording using this video and audio recording device (with consent) might be more appropriate in important situations that do not last for a long period of time, as in the aforementioned “important meeting”.

Participants in this study were a relatively homogenous group of high functioning older adults. None of the group had any significant everyday organisational problems and although none scored at ceiling for this task it would be of interest to see if scores between live and video ratings still had as strong a relationship when participants experienced more difficulty with the task. In summary, inter-rater reliability of this task appears strong and lends itself to the potential to be used in further investigations with different populations. It provides initial validation for the use of remote scoring of assessments, in this case the MET, in reducing bias and promoting better use of clinician’s time as they do not necessarily need to concern themselves with live scoring.

Chapter Four:

Intensive working memory training: A single case study of a patient following a cerebrovascular accident

Abstract

Previous research on computerised cognitive training with patients with a brain injury or cerebrovascular accident has shown mixed results in terms of generalisation of gains to everyday situations and untrained tasks. This chapter describes a single-case study rehabilitation strategy used with a patient, S.M. S.M., following hypoxic brain damage, reported difficulties with working memory and time-perception. This was further confirmed through a battery of assessment and clinician reports. Previous studies have compared adaptive working memory training (AWNT) with non-adaptive versions of the tasks or against no active control. Here the specificity of any AWMT effects were examined by contrasting its impact with an adaptive, challenging computerised training that emphasised novel problem solving (PS). Following initial assessment S.M. trained for four weeks (20 days), 20-30 minutes a day on the PS tasks before repeating key outcome measures, including in time-perception. He then trained for an identical period on AWMT. Interestingly, an index of WM, albeit a low executive load task, digit span showed no disproportionate gains following AWMT. This condition was however specifically associated on transfer executive measures of planning and organisation and with significantly better time-perception performance, a capacity that had previously been resistant to rehabilitation. The results provide encouraging evidence in the context of low WM capacity following brain injury that AWMT may have benefits. Limitations of the study, including in terms of WM assessment are discussed.

Introduction

Single-case experimental design in rehabilitation studies

One of the biggest criticisms of single case experimental design studies is the extent to which findings generalise to other cases (Perdices & Tate, 2009). An equal problem is a typical lack of experimental controls that make it difficult to establish whether beneficial effects are due to the treatment applied or other uncontrolled variables. Against this, group studies are often obliged to lump together individuals who share certain characteristic (aetiology, functional difficulty) but who may differ in many ways relevant to the outcome. In contrast, appropriately controlled single-case experimental design studies allow highly targeted interventions, may be the only option in rare conditions, and can provide quick, therapeutically relevant feedback. Generalisation to others is not crucial but the accumulation of successful interventions certainly provides useful guidance for other clinicians (Tate, Mc Donald, Perdices et al., 2008; Mateer, 2009).

Tate and colleagues (2008) developed a single-case experimental design (SCED) quality scale. It includes criteria such as target behaviours definition, design, baseline, availability of raw data, inter-rater reliability, independence of assessors, statistical analysis, replication and generalisation. If all the criteria are clearly addressed and the study has been shown to be of high quality then generalisation of the observed effects to other cases is possible, and hence one of the major difficulties in interpreting results from single-case design studies is addressed. Although a criterion of the SCED, Mateer (2009) also recommends that the clinical history of the participant is made available.

Working memory training with clinical populations

Damage to working memory can have negative effects on a person's social, vocational and everyday functioning and is often seen as a predictive factor of recovery post brain-injury (Lundqvist et al., 2010). As discussed, cognitive training in people with cognitive impairment post-ABI has historically had rather disappointingly impact (e.g. Stigsdotter Neely & Bäckman, 1995) and resultantly a more practical clinical focus on compensating for deficits has been adopted. However, interest in cognitive training within neuropsychological rehabilitation has increased in recent years, following the publication of impressive and seemingly generalisable effects of WMT in children with ADHD (e.g. Klingberg et al., 2002; Klingberg et al., 2005) and there has been a focus of computerised WMT in adults post-brain injury in more recent years. As outlined in chapter 1, these have been followed by preliminary studies in people with brain injury (Serino, Ciaramelli, Santantonio, Malagu, Servadei

& Ladavas, 2007; Westerberg, Jacobaeus, Hirvikoski, Clevberger, Ostensson, Bartfai & Klingberg, 2007).

This investigation focused on an individual, S.M., a bright graduate who had acquired rather specific WM problems following a brain injury. As outlined below, he felt this difficulty in keeping things in mind was one of his primary impairments, undermining his confidence and ability to return to work, manage at home and so on. Unlike many of the WMT studies in which poor WM function has either been inferred due to a clinical condition (e.g. ADHD), has occurred within a broader spectrum of cognitive impairment, or has been absent (e.g. in healthy undergraduate studies) this was a case in which far transfer of WMT gains was less critical than the issue of whether WM capacity could be increased. To examine this specificity AWMT was contrasted with the effects of a challenging and adaptive computerised training of the same duration that emphasised a range of novel problem solving skills (PS). The titration of difficulty in both sorts of training makes this an extremely conservative comparison and meets high quality standards in single-case methods outlined above.

In single-case designs the effect of a variable that is hypothesised to influence a behaviour but not produce a permanent change in the individual can be investigated using a variety of ABA designs, in which the factor is absent during A and present during B. In this manner order effects, the effect of time, spontaneous recovery and so on can be controlled. Where lasting change is hypothesised however, inferences can be drawn from varieties of ABA designs where the change with the onset of the factor can be interpreted with reference to the variance observed over multiple baseline observations (multiple assessments points at A). Where those observations are of a behaviour, such designs are excellent but when they require specific assessment they can be limited by practice effects and low acceptability to participants. In addition, these would generally fail to control for non-specific effects of the factor (e.g. beginning *any* form of intervention). Accordingly, this study adopted an ABC design in which test performance at baseline (A) was contrasted with that after a carefully matched and plausible control intervention (B) and then the hypothesised to be effective intervention (C). There are always limitations and this design cannot deal with other performance-relevant events coinciding with one of the intervals, is clearly vulnerable to cumulative practice effects and is problematic if the effects of B take time to accumulate.

The development of the PS training used here is described in more detail in subsequent chapters. It consisted of a variety of tasks including deducing the next pattern in a logical sequence, comparing two patterns to see if they are identical, deciding whether one pattern is a canonical rotation of another, and deducing a rule by which an odd-one-out can be defined (see below for details). The claim here is not that these do not require WM, but rather that the WM demands are very different to standard AWMT. In AWMT the ‘rule’ linking events in the task is invariably simple and sequential (span, n-back) and the major challenge lies in retaining the information and, in some tasks, manipulating it (e.g. reversing it). In many of the PS tasks the rule linking elements is not obvious and must be derived but the information need not be maintained as it remains on the screen.

The key issue in this study was therefore whether a marked improvement in S.M.’s WM function would follow AWMT, suggesting a *specific* effect of AWMT, or PS, suggesting a more general, potentially *g* related, effect, or neither. WM gains can be further subdivided into those that may benefit capacity within a particular ‘slave’ system (such as the phonological loop) and those that exert a putatively wider influence via the central executive. To this end two different sorts of outcome measure were employed – the low executive demand of Digit Span, the relatively high executive load of backward digit span, and the requirements to plan, monitor and switch in the Games Evaluation and 6 Elements tasks. The digit span tests were far from conservative as both were part of AWMT training. In contrast, the GET tasks represent a fairer and more conservative test of executive transfer. The case also raised an interesting issue of potential far transfer. In addition to his poor WM, S.M. reported problems with time-perception. In contrast with his pre-brain injury state, he found it very difficult to estimate, for example, the amount of time he had spent in a waiting room before an appointment (see case description for more details). This difficulty had proved resistant to previous rehabilitation attempts (see below).

There has been considerable interest in human and other species ability to track time even in the absence of obvious external cues. Models positing dedicated time-perception systems have been contrasted with time-perception arising as an intrinsic part of other systems (such as memory decay; Ivry & Schlerf, 2008). There are clearly processes that are relevant to time-perception that are relatively automatic and, by inference, free from interference from concurrent tasks. Rats, for example, are reportedly able to accurately time periods of 40 seconds after the complete removal of cortex (Jaldow, Oakley & Davey, 1989). At the same time, contextual factors including concurrent cognitive load have a significant influence on time-perception during that interval (Khan, Sharma & Dixit, 2008). A recent study (Woehrle & Magliano, 2012) contrasted the ability of healthy students with relatively high and low WM capacity to estimate a time-interval during which they performed a series of mental arithmetic calculations. The high capacity students were *less* aware of the time, attributed by the authors (perhaps rather circularly) to their greater engagement in the maths test and them consequently having less available WM capacity to monitor time. In keeping with this sort of argument, children with ADHD – in which low WM is common - were reported to be less accurate in replicating intervals when there was no secondary task (Meaux & Chelonis, 2003) Whilst it is far from certain that S.M.'s time-perception difficulties were secondary to his WM loss, there are reasonable grounds to suppose that WM can influence time perception. This could be related to the availability of capacity to monitor internal and external time cues and the different experience of the "*specious present"* due to the capacity to maintain its elements in mind. A secondary hypothesis in this study was therefore that improvements in S.M.'s WM capacity – whether following PS or AWMT – would be accompanied by gains in his previously resistant time-perception abilities (see case description for previous interventions).

Case Description

S.M. was a 45 year old man, who sustained bilateral hypoxic brain damage following a bicycle accident 3 years previously, in which the pressure from the straps of the bike helmet he was wearing lead to bilateral dissection of the carotid arteries. He was left with left-sided hemiparesis and multiple cognitive impairments.

S.M. was married and lived at home with his wife and two young children. He previously completed a PhD in biology and was working as a researcher in a biology lab at the time of his accident. When assessed, S.M. was not in voluntary or paid employment. He had left-sided weakness with reduced function in his upper limb but had developed compensatory strategies and was fully independent in self-care activities. He required no assistance walking but had reduced mobility due to his hemiparesis. He was computer literate prior to his injury and continued to experience no difficulties in this respect (the training would be computer delivered). He attributed his problems primarily to cognitive rather than physical problems. These were principally in keeping things in mind, time-perception (not realising how much time had passed) and perception. He had stopped driving.

Previous problems and rehabilitation

At 12 months post-injury S.M. had been through intense assessment and 18 weeks of neuropsychological rehabilitation at the Oliver Zangwill Centre. This rehabilitation included a mixture of individual and group activities on areas such as psychological support, cognition, understanding brain injury, vocational rehabilitation and mood. At the 6 month follow-up stage S.M. continued to have difficulties in two main areas. He was having problems with tasks that required working memory and he also was having problems in estimating durations of time – both time that had passed and in estimating how long a set period of time might be – i.e. if he was told to “wait 5 minutes” he would find it difficult to have a sense of how long that would be. Previous work had been done with S.M. over 4 sessions on time perception where he was provided with feedback on time spent on daily tasks. He would estimate how long he spent on a task such as brushing his teeth and he was then given the correct time after. There was some improvement reported during sessions. This intervention was based on a study by Roy, Mitten and Christenfeld (2008) where participants without a head injury were shown to have provided a more accurate estimation of the duration of a task when supplied with feedback. They reduced their errors by half. With this group of university students following feedback they had a tendency to go from underestimating tasks to overestimating the duration of tasks. S.M. used to overestimate time – waiting for a few minutes in the waiting room felt like a very long time. The improvement that was seen in his intervention did not carry from session-to-session or when treatment was withdrawn. There was no generalisation in any improvements and so the intervention was withdrawn. He did, however, find it useful to time himself on certain tasks, such as shaving, to have a note of how long these tasks take in order to plan his time more effectively.

Methodology

Initial contact and summary of baseline screening assessment

An informal discussion took place at an initial meeting with S.M. The demands of the training were explained, along with the rationale as to why it might be of benefit to him. He was given the opportunity to ask any questions he wanted, as well as being provided with an email address for future queries. He was then asked to sign a consent form that explained his right to withdraw from the training program without giving a reason, and arrangements for data protection.

As S.M. had previously had extensive neuropsychological assessment as an inpatient and as a client of the Oliver Zangwill Centre and these tests were not repeated (the effect of prior exposure is not always known).

The Cortical Vision Screening Test (CORVIST; James, Plant & Warrington, 2001) is an assessment that is used to identify visual impairments in people with normal or corrected vision. In order for S.M. to participate effectively in the computer based training we needed to ensure that he had no extreme visuoperceptual impairment that may have resulted from his head injury. The test is made up of ten subtests that assess different areas of visual processing. The tests look at acuity, shape discrimination, size discrimination, shape detection, hue discrimination, scattered dot counting, fragmented numbers, word reading, crowding and face perception. A variety of visuoperceptual areas are targeted such as spatial scanning, perceptual identification, reading, impairments of acuity and face perception amongst others. S.M.'s results were all within normal range as can be seen by Table 4.1 below:

CORVIST (Cortical Visual Screening Test)	Snellen Equivalent/ correct?	Within normal range?
Symbol Acuity	Three incorrect	Yes
Shape Discrimination Test	Three incorrect	Yes
Size Discrimination Test	One correct, One incorrect	Yes
Hue Discrimination Test	All correct	Yes
Scattered Dot Counting	All correct	Yes
Fragmented Numbers Test	One incorrect	Yes
Word Reading Test	All correct	Yes
Face Perception Test	All correct	Yes
Crowding Test	All correct	Yes

Table 4.1: CORVIST results screening measure

Outcome measures

Time Perception Task (personal communication Fish, Wilson & Manly 2010)

In the absence of standardised time-perception tasks one was developed for this study. Programmed in E-Prime 1 (Psychology Software Tools), the task required S.M. view a shape appearing on the screen for a varying duration and then recreate each interval by holding down a key (during which the shape also appeared). The task was under single- and dual-task conditions (interleaved trials). The dual task was designed to interfere with any vocal or sub-vocal counting strategy was conducted on both 'watch' and 'reproduce' sections of each trial. S.M. was asked to generate and speak aloud a random sequence of numbers. Five time intervals (5, 10, 15, 20 and 30 seconds) were used, each twice under each condition. Scores are reported as the mean percentage deviation from the correct time on each trial – lower scores indicate better performance.

Story Recall

Story Recall subtest of the Rivermead Behavioural Memory Test II (RBMT II; Wilson, Cockburn & Baddeley, 2003) – immediate and delayed versions.

In this subtest a prose passage of about 7 or 8 sentences is read aloud with participants then being asked to recall as much as possible. Scoring is based on the number of "ideas" recalled. The delayed version requires a second free recall of the same material after a 10 minute delay. There are four parallel stories that we used to minimise practice effects over the repeated assessments. The Story Recall subtask was included as a measure of immediate and delayed memory capacity and was of interest to see if AWMT would improve performance on this memory test.

The Games Evaluation Task (GET)

As described in Chapter 2.

The Raven's Standard Progressive Matrices (Raven's SPM)

The Raven's SPM, a non-verbal problem solving measure, was used here. It contains 60-items (multiple choice) and was untimed here. The items are presented in 5 sets (A-E) with 12 items in each set. Each item depicts a logical series and participants are asked to select from an array the item that best completes the sequence or pattern. Many patterns are presented in a matrix form. Here the single raw scores are reported.

The Goal Management Questionnaire (GMQ)

The GMQ (Manly et al. personal communication) was developed to closely track the key topics of Goal Management Training. It comprises 34 questions requiring responses from 0 “not a problem at all” to 10 “a very major problem”. Items include “Worrying too much about things that you need to achieve?”, “Losing track of the time?”, Having to go back for something that you had forgotten to take with you?”, or “Avoided thinking about a problem because it just seems too complicated?”. The lower the score on this scale the better a person’s performance.

The Hospital Anxiety and Depression Scale (HADS)

The HADS (Zigmond & Snaith, 1983) is a brief self-report tool widely used in many clinical groups. Its reduced focus on somatic and motor characteristics makes it less prone to false positives in populations with likely physical injuries than general anxiety and depression scales (Dawkins, Cloherty, Gracey & Evans, 2006; Bjelland, Dahl, Haug & Neckelmann, 2002). It presents 14 statements (half relating to anxious symptoms, half to depression) and respondents are asked to rate on a scale of 0-3 how often something might happen or how strongly they agree with the statements.

The Behavioural Assessment of Dysexecutive Syndrome (BADS)

The Modified 6-Elements subtest of the BADS (Wilson, Aderman, Burgess, Emslie & Evans,1996).

This is a simplified version of Shallice and Burgess’ (1991) measure. Participants have six tasks to attempt within 10-minutes. The 6 tasks actually comprise 3 basic tasks, each with two parts. These are writing answers to a series of simple mental arithmetic problems presented in two separate booklets, writing the names of common objects depicted in two separate booklets, and dictating information about two recent personal events. The 10 minutes is not nearly enough to complete all of the tasks. Accordingly participants must plan and remember to switch tasks whilst adhering to the rule that part 2 of a given task cannot be started immediately after part 1. The number of tasks attempted and rule breaks are scored.

The Zoo Map subtest of the BADS

This task comprises two parts. In the first the participant has to plan a route around a zoo map that meets various criteria given in rule sheet (listing the places that must be visited and constraints, for example about whether you can return to the same path). The second version has the same requirement but now the participants receive step-by-step instructions, reducing the requirement for planning and retaining a plan. Accuracy and rule breaks are scored. As a measure of planning and following rules this task was included as a generalisation measure.

Procedure

An ABC design was used. Following baseline assessment, S.M. completed 4 weeks (20-days) of PS for 20-30 minutes a day. Reassessment was then followed by a further 4 weeks of adaptive challenging training of the same duration and final assessment session.

Problem Solving Training (PS)

PS was comprised of the following tasks, each of which provided instant feedback on performance and were adaptive, increasing in difficulty following correct responses and reducing following incorrect responses.

Spatial Slider – Numbered tiles appeared on the screen arranged in numerical order on a grid. The tiles were then shuffled so that some were no longer in the correct position. The task was to rearrange the tiles to the correct order in as few moves as possible by sliding them in and out of the blank spaces.

Feature Match – Two boxes appeared on the screen, each containing a complex array of abstract shapes. The task was to decide if the two boxes were identical or different.

Odd-one-out – Nine patterns appeared in a grid on the screen. Each pattern comprised of a set of features, for example, colour, shape and number. The task was to detect the odd-one-out that could be defined by difference in a single feature or a combination of features.

Rotations- Two boxes appeared on the screen, each filled with red and green squares. The task was to decide whether one was a canonical rotation of the other.

Polygons – Two panels appeared, one containing two overlapping shapes and the other containing just one shape. The task was to decide whether the single shape was identical to one of the overlapping shapes or subtly different.

Simple choice – A simple, spatial reaction time task. Five boxes appeared on the screen. The task was to mouse-click the box in which a red square target appeared as quickly as possible.

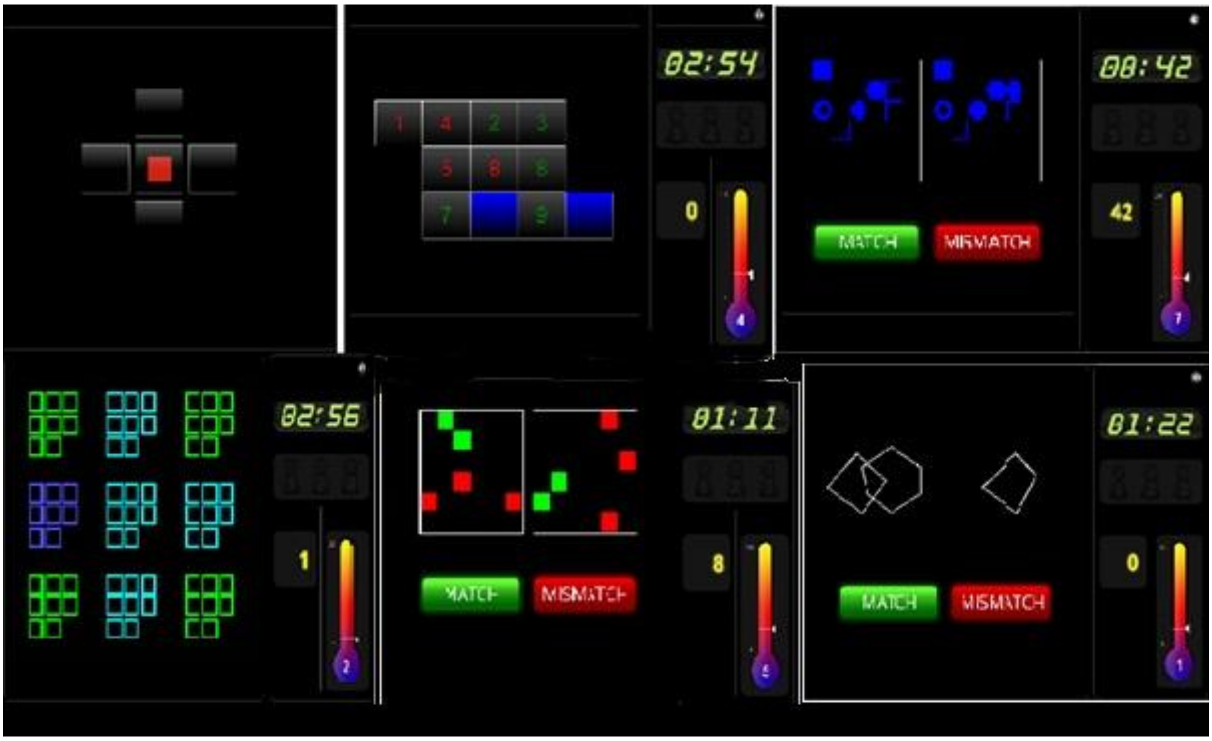


Figure 4.1: Screenshots of placebo training tasks: (From top left) Simple choice, Spatial Slider, Feature Match, Odd-One-Out, Rotations, Polygons.

<u>Adaptive Working Memory Training (AWMT) tasks</u>

Digit Span – Participants had to remember a sequence of numbers that appear on the screen one after the other. Sequence is increased or decreased depending on performance.

Spatial Span Ladder – A 4x4 grid appears on the screen. A number of boxes light up one after the other. Participants have to remember the sequence and click on the boxes as they lit up. The number of boxes that light up increases or decreases depending on performance.

Paired Associates – Boxes appear at different locations on the screen, each box appears a different image. Images are shown one by one. Participants need to remember which image appears in which box. An additional box appears which will present an image that has already been shown. Participants need to click on the box which contained the image that was presented.

Monkey Ladder – A number of boxes appear on the screen. Each box will be numbered. The numbers will disappear and participants have to click on the boxes in numerical order as they appeared on the screen.

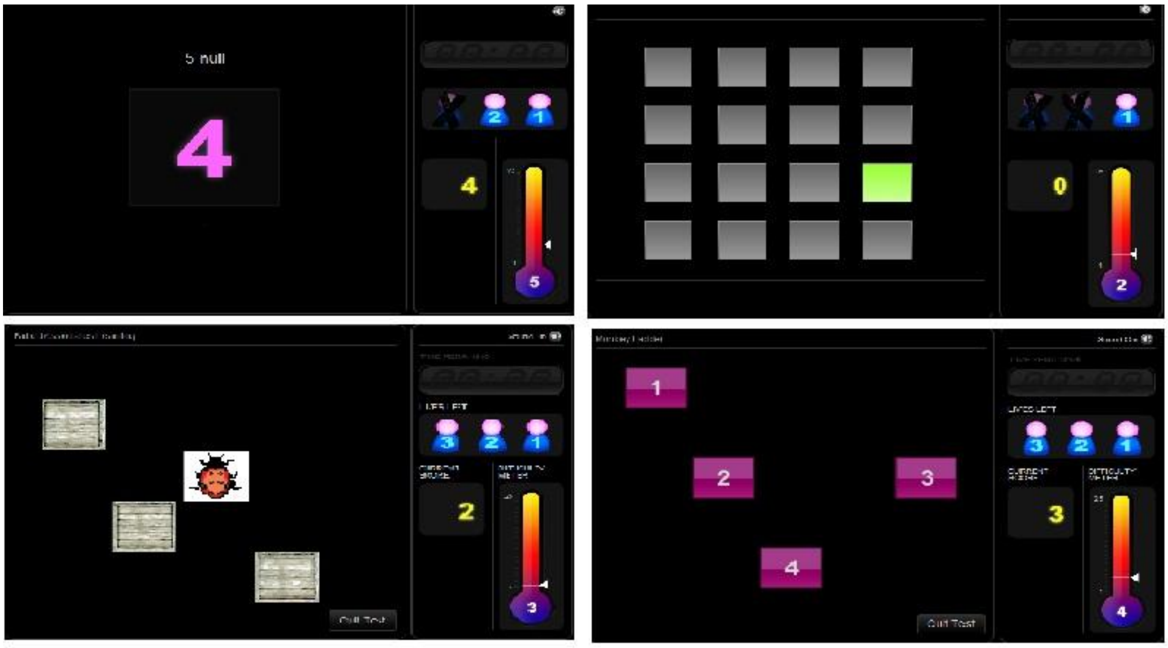


Figure 4.2: Screen shots of working memory training tasks. (From top left) Digit Span, Spatial Span Ladder, Paired Associates, Monkey Ladder

Training schedule

The training schedule for the PS and AWMT were the same. S.M. was assessed on the Monday morning of the week in which he was due to start the training. He began the training at home later that day. The training lasted for four weeks with five days

(M-F) of training per week with the recommended amount being 20-30 minutes in each session. The training was carried out at S.M.'s convenience. He tended to train during the week and leave his weekends free. The times of the day in which he carried out the training varied depending on his demands that day. He was assessed after the four weeks of both training conditions on the Monday following training cessation. In the 30 minutes of training S.M. completed on average 12 different tasks with some repetition of tasks within each session given the limited number. Although the researcher was available to contact at any time, S.M. was also contacted weekly to ensure that he was still happy to continue with the training and also to see if there were any problems/questions he wanted to ask.

Results

Acceptability of intervention.

S.M. reported being happy with both training programmes and with their intensity. He complied with the recommended schedule and completed each day of training (40days) as required.

Direct training effects.

Unfortunately, S.M. employed the sensible strategy of writing down digit sequences in order to perform the digits backwards training task (which he found very difficult). This measure is therefore unavailable. He did not apply this strategy to forward spans or other training tasks, according to his report.

With due caution regarding the psychometric sensitivities of the tasks, all results from the PS are suggestive of marked improvements across the training period. In contrast, very little change occurred on the trained WM tasks.

At baseline, and in line with his previous assessments, S.M. had the very low digit span score of 4. This was unchanged after PS training. Following AWMT training *which involved multiple trials of digit span,* his score was 5. Whilst this increase is technically in line with the rationale behind the ABC design, the low level of change and the absence of extensive baseline data (showing that 5 was very unlikely without the training) urges caution. His spatial span performance was unchanged after AWMT.

Story recall.

S.M.'s ability to remember key points from 3 well-matched stories was assessed at each time point. Here the pattern (see figure 4.3 below) was consistent with AWMT being associated with a lower levels of recall.

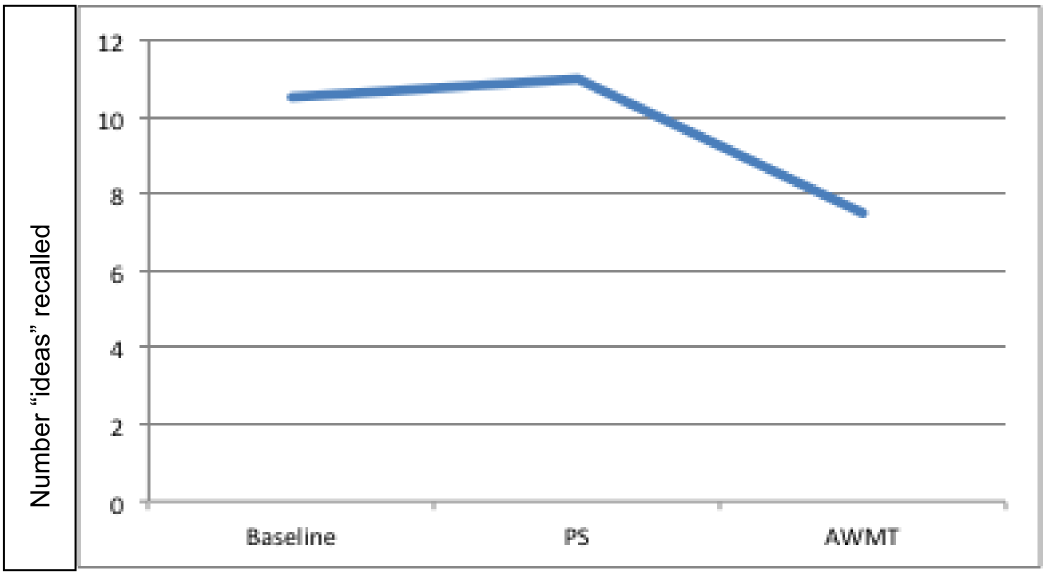


Figure 4.3: S.M.'s story recall on 3 parallel stories at baseline

<u>Executive transfer</u>

S.M.'s performance on the GET task (time deviation score) showed consistent improvement across the 3 measurement points of the study. Although no feedback is given and time-perception (see below) has a clear role in the task, without control it is not possible to distinguish gains from PS and AWMT from simple practice effects. The pattern was rather different on the conceptually similar 6-elements timing score that flattened in its improvement following PS. S.M.'s scores on this task are not optimal on any three data points.

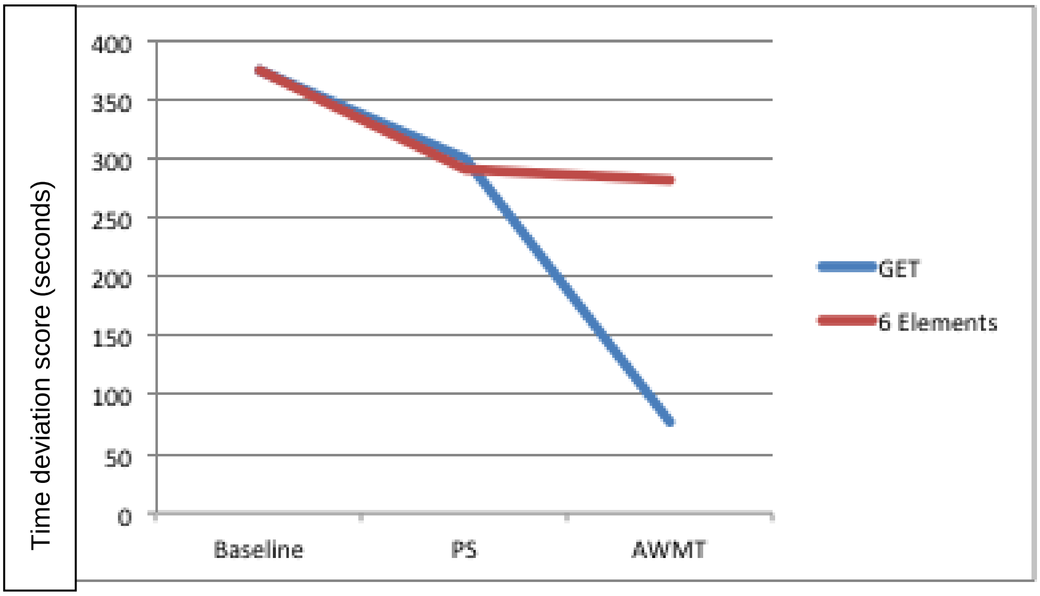


Figure 4.4: S.M.'s time deviation scores on the GET and 6-Elements tasks at baseline, following PS and AWMT (6-elements scores scaled by 4.84 to equate initial levels).

Perhaps unexpectedly, despite exposure to conceptually similar tasks, Raven's accuracy scores showed a dip following PS. However, although the Raven's is not time-limited a record was kept of how long it took and many of the PS measures were timed and post- PS performance is conducted rather faster than at other points. Taking this into account with a time/accuracy score, the pattern for this test is consistent with a benefit from AWMT that did not occur through simple practice after PS.

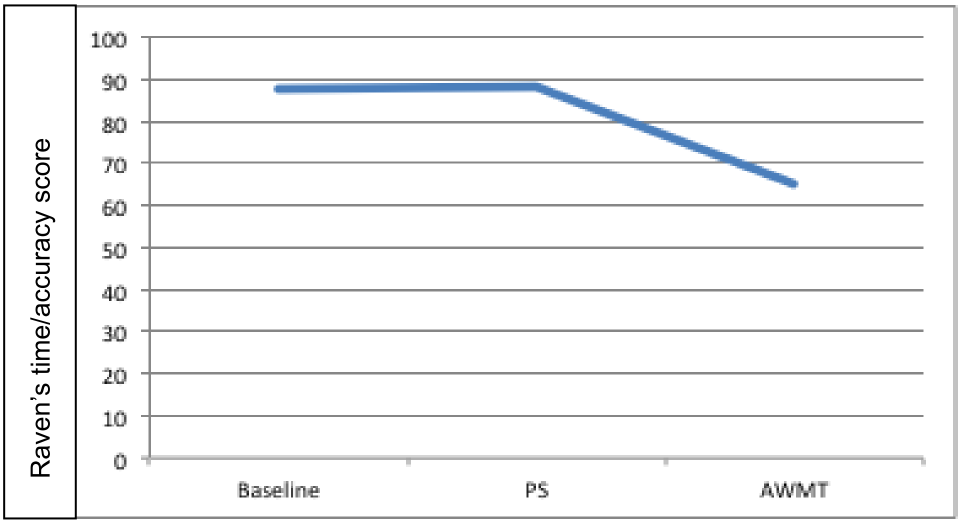


Figure 4.5: S.M.'s performance on Raven's Matrices (time taken/accuracy) after Problem Solving (PS) and adaptive working memory training (AWMT).

Unfortunately it was not possible to reassess S.M. on the Zoo Map test in the post-PS assessment due to time limits on the intervention session – S.M. had transport booked for a certain time.

Time perception task.

As shown in figure 4.6 below, the results of the single-task time-perception measure, in which WM resources can arguably be deployed, show minimal gains following PS but then a rather substantial drop following AWMT. In contrast, dual task performance shows a drop after PS that is well maintained, more consistent with a simple effect of practice.

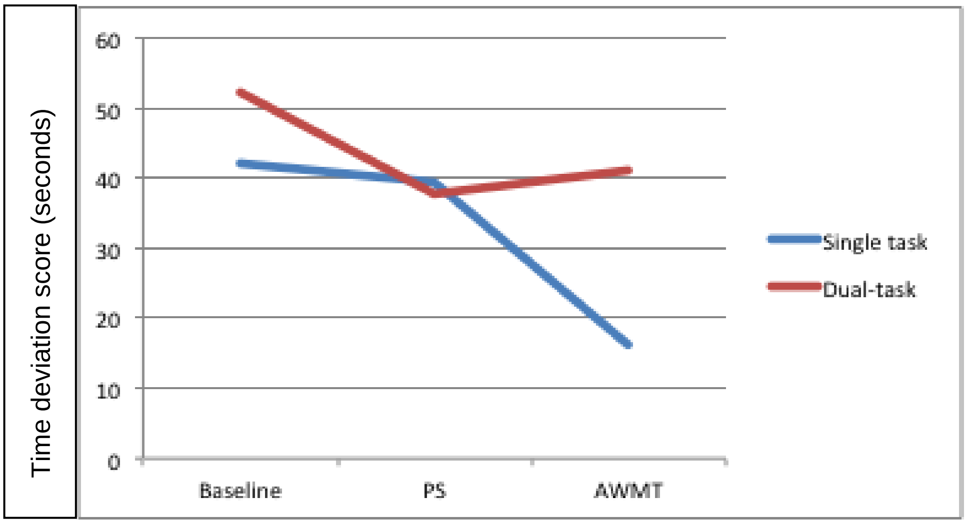


Figure 4.6: S.M performance on single- and dual-task time-perception measures at baseline and after PS and AWMT

Mood.

S.M.'s self-reported mood was relatively stable from the beginning to the end of study period although his depression score showed something of a spike following PS. It is unclear whether this was related to training or other events. S.M. did not mention anything that may have caused this spike. Low scores indicate fewer symptoms of anxiety/depression.

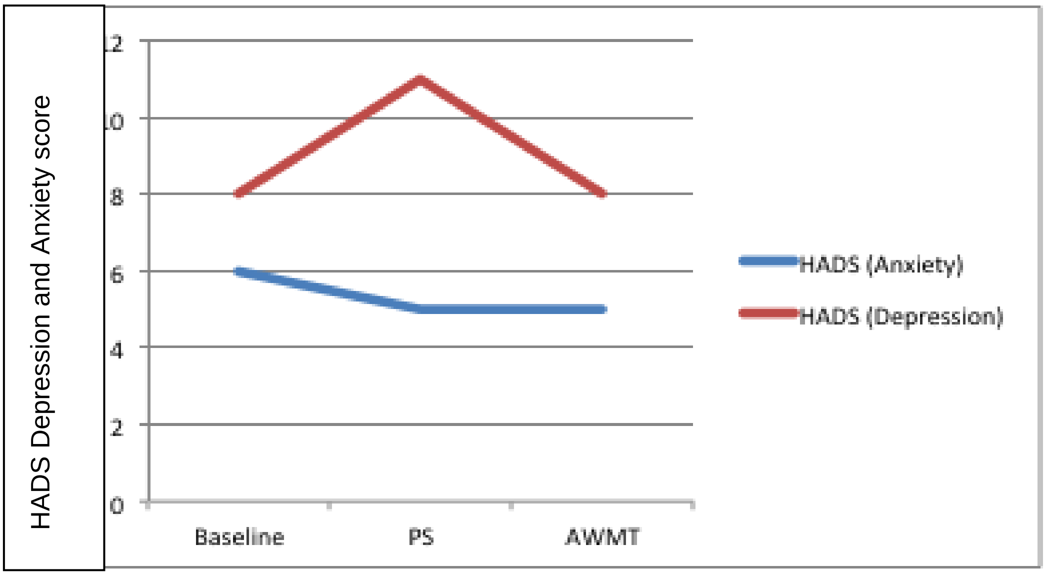


Figure 4.7. Measures of S.M.'s mood at baseline and following PS and AWMT.

Below are the summarised scores following both training interventions for S.M.

Measure	**Baseline**	**After PS**	**After AWMT**
Time Perception-Single Task Mean % Deviation	42.13	39.29	16.22
Time Perception-Dual Task Mean % Deviation	52.36	37.75	40.92
Six Elements – n tasks	3	3	3
Six Elements - time deviation	774	600	582
Story Recall-immediate raw	10.50	11	7.50
Story Recall-delayed raw	8	9	6
Games Task - n tasks	3	4	10
Games Task - time deviation	375	300	77
Zoo Map - time (sec)	414	No parallel version	174
Zoo Map - overall score	7	Did not want to administer so frequently	3
HADS (Anxiety)	6	5	5
HADS (Depression)	8	11	8
GMQ (average score)	2.18	2.60	1.90

Raven's score	30	24	30
Raven’s time (seconds)	2628	2123	1948
Digit Span	4	4	5

Table 4.2 Summarised results before and after low and high demand WM training

Discussion

Taking into account numerous caveats related to single-case design and single assessment points the results of comparing two plausible and challenging cognitive training programmes were consistent with:

- S.M. showed good evidence of improvement with training on a range of initially novel problem solving tasks.
- In contrast, whilst his digit span improved slightly from baseline and post PS levels, in general his performance was very stable on WM measures used in the training.
- S.M showed improvements in strategy, planning and monitoring as measured by the GET and 6 Elements test over the study period but it is hard to rule out simple practice.
- On the Raven’s matrices, the pattern was consistent with a specific effect of AWMT in line with some previously reported results.
- S.M. had specific problems with time-perception that had previously proved resistant to rehabilitation based on feedback. When this was assessed with a dual task (during which strategy application to the time-perception task is more difficult) the results were consistent with a general practice effect on this particular task – although this was apparently in contrast with previous interventions. On the single-task measure, in which arguable WM resources could be deployed to support performance, the result was consistent with a specific benefit of WMT.
- Other results included this intensity of training being acceptable to the participant and a relatively stable mood profile over the course of the study.

This study raises interesting questions about clinical research. A patient had WM and time-perception problems. We do not know how common that combination is (time-perception is rarely assessed and was addressed in this case only because the patient was aware of it). Even were a group to be assembled who show WM and time-perception problems, they may differ in many relevant respects. The key clinical issue here was whether a given intervention is plausibly linked with a reduction in a presenting symptom that had previously been resistant to intervention. The result at that level is positive, with the magnitude of the effect (a 25%improvement in accuracy over intervals from 5-30 seconds) suggestive of clinically significant benefit. The result would certainly suggest that further work examining relationships between WMT and time-perception (for example in ADHD) is warranted.

Taking at face value, the single- vs. dual-task time perception results are consistent with S.M. being better able to deploy some strategy, such as sub-vocal counting, during the single-task version. It is plausible that increased WM, management of distraction etc. from the training helped that effect. If this is the case it raises an interesting question: If the improvement in time-perception is linked to AWMT, can WM capacity occur despite no clear change on the WM measures that are used in the training? Conceptually at least the answer is yes and relates to the coarseness of the measures: If I score 4 points by running 100 m in 20 seconds but need to run it in under 7 seconds to score 3 (by some arbitrary scoring criteria), my score will remain at 4 for a very long period despite hard training which will nevertheless be of no benefit to me in catching the bus! It is not clear that we yet have an adequate rubric for equating a change in span or n-back of one with WM requirements in other tests.

In other respects that are of course strict limitations on what can actually be inferred from the results. People vary considerably in their performance of a task from one time to another and it is all too easy to take credit for the time-perception gain whilst ignoring the ostensibly deleterious effects of the training on story recall as irrelevant. Given more time it would be useful to develop parallel outcome measures that allow repeated assessment within the epochs allowing statistical analysis to sort random from consistent variation. With that in mind, the results from this chapter provide a cautious pointer to effects that need to be explored in more rigorous single-case and group designs.

Chapter Five:

Working memory training in a healthy older adult population

Abstract

This chapter reports the effects of a twenty-five session intensive computer-based working memory training regime on a group of older adults. Participants trained five days a week for five weeks on a variety of adaptive WM tasks. The control group received no training. Twelve participants were in the working memory training group (mean age 66.75 years) and twelve participants were in the control group (mean age 70.36 years). Participants in both groups showed improvements on untrained outcome measures of working memory, recognition memory and fluid intelligence attributable to practice but a disproportionate benefit of WMT was undetectable. Important caveats include the relatively small group sizes and the generally good abilities of this sample.

Introduction

As discussed in the Chapter 1, repetitive training of cognitive skills has generally proved disappointing in neuropsychological rehabilitation. Improvements on the trained tasks are common but generalisation to other tests, let alone everyday activities, has rarely been reported. However, a series of recent studies have suggested that adaptive training on working memory (WM) tasks can produce generalised benefits including to untrained measures of fluid intelligence (Klingberg et al., 2005; Jaeggi et al., 2008). Not all results have been positive however (Owen et al., 2010, Redick et al., 2012) and replication is essential.

Although the aims of the work described in this thesis were to inform neuropsychological rehabilitation, claims have been made about the efficacy of WM training in the healthy population (Mindsparke, 2011; Morrison & Chein, 2011; Jaeggi, Studer-Luethi, et al., 2010; Klingberg, 2010; Sternberg, 2008). Accordingly a series of proof-of-concept studies are presented in this and the following chapters with healthy elderly participants. There were practical advantages to this in terms of access to the populations, availability of internet connected computers in the home etc. In addition it is likely to be a population in which WM skills are somewhat reduced compared with the generally young participants reported by Owen et al. (2010) and Redick et al. (2012). If lower WM capacity is a factor in the generalised efficacy of the training, it is more likely to be seen in this group.

Previous work reporting positive effects of adaptive Working Memory Training (AWMT) have generally used variants of Klingberg and colleagues program (Cogmed Working Memory Training) although positive results have been reported via training on the Paced Auditory Serial Addition task (PASAT; Serino et al., 2007). A further advantage of the current studies was to examine a new on-line training program developed at the CBU. In the subsequent chapter an adaptive novel problem solving training program is described. Here a novel set of AWMT, developed by Adam Hampshire (Hampshire, personal communication, Owen et al., 2010) were evaluated. It should be noted that the current study was run before the negative results of Owen et al. (2010) were known.

Because of the early developmental stage of the training program only five different WM tasks were designed, tested and available for use. These five tasks were the same as those reported in chapter four, but with the addition of backward digit span. The tasks were paired associate learning, spatial span ladder, monkey ladder, digit span forward and digit span backward and were accessed by participants through their own personal links to an online site. All tasks were adaptive, becoming more challenging or easier depending on participants' performance.

Previous AWMT studies have used objective, performance-based measures and self- or other- report of cognitive function, behaviour and coping in everyday situations (e.g. Klingberg et al, 2005; Westerberg et al, 2007). The objective measures have included near and far transfer tasks. Near transfer refers to a test which has not been directly trained but which is thought to require very similar processes to those which have been trained. An untrained WM task would be a clear example. Far transfer refers to tests which ostensibly require quite different processes but which may share a common demand with the training measures, fluid IQ tests of novel problem solving being the best example (i.e. these may be seen to require WM but much else besides).

The current study had similarities and differences with this approach. Outcome was assessed on self- or other-rated questionnaire reports on coping, behaviour and cognitive function (Goal Management Questionnaire, Cognitive Failures Questionnaire). Far transfer tests of fluid IQ (Raven’s Standard Progressive Matrices, Cattell Culture Fair) were also included. As discussed, a particular feature of the studies reported here was to investigate whether benefits of AWMT transferred to measures of executive function that emphasised planning and prospective memory. Accordingly outcome measures included, for the first time in AWMT studies, the novel Games Evaluation Task (see chapter two) and The modified Multiple Errands Test (see chapter three). The Multiple Errands Task (Shallice & Burgess, 1991) was used here as both a measure of far transfer and also as an indication of everyday life performance. This measure has not been reported to date as an outcome measure in WM training. As shown in chapter three, the MET can often be more indicative of difficulty (Shallice & Burgess, 1991) than other measures in a clinical population so it is worth investigating if training can lead to any improvement in function in this everyday life task.

Because of the time demands of these measures it was decided to drop near-transfer assessment and instead examine performance change on a WM measure used as part of the training, Digit Span.

The Hospital Anxiety and Depression Scale (HADS) was also included as an outcome measure. It was of interest to see if participation in organised training would have any influence over self-reported symptoms of anxiety or depression in this older adult population. Participants were, however, screened for any significant problems with mood (see below) as it was thought that any clinical levels of depression or anxiety could influence participant’s participation in the study.

Finally a non-WM memory measure, the Doors and People test was included in a more exploratory fashion. This task allows the assessment of visual and verbal memory over a short interval, retention over a delay and separate estimates of recall and recognition. There is considerable interest in the effects of aging on these non-WM capacities and evidence that AWMT was effective in bolstering these functions would be important. It is certainly plausible that practice, increased capacity and/or strategy development in taking in and manipulating information in WM could produce benefits in the initial processing of these materials that would foster retention.

In summary, twenty-four participants were recruited and randomised into an experimental and control group. Both groups completed an initial series of measures. The AWMT group were then asked to engage in daily (weekday) AWMT sessions of 20-30 minutes over a 5 week period. At the end of the training, the initial measures were repeated for both groups.

Methodology

Participants

Healthy volunteers who were members of the CBU participant panel were contacted by phone and email and offered a place in the study. Inclusion criteria were age of 55 or over, no history of neurological impairment, access to a home computer with broadband connection, availability for the duration of the study and scores within the normal range on screening measures (see below). Participants were paid for their contribution and were free to withdraw at any stage without giving a reason. The final sample consisted of twelve participants in both the training and control groups. The working memory training group consisted of twelve participants, eight men and four women, with a mean age of 66.75 years (SD 5.74). There were twelve participants (4 male) in the control group with a mean age of 70.36 years (SD 6.14).

Participants were told that they could be placed in either control, working memory training or problems solving training (described in chapter six) groups. They were informed that whoever took part in control group would be offered training on the programme (if any) that proved most effective in the initial comparisons. Allocation to groups was at random. Participants were consecutively numbered as they were recruited to the study. Randomisation was via http://www.randonization.com (Dallal, 2008). From this a randomisation plan was generated, identifying which participants were to be allocated to each of the three groups.

Screening measures

Whilst mood and its potential interaction with cognitive training is an important issue in brain injured and other populations it is a potential confound in relatively small training studies. To increase the likely homogeneity of the sample in this respect two screening measures were employed, The Geriatric Depression Scale (GDS) and Beck Depression Inventory (BDI) on which participants had to score within the normal range to participate. In practice, all did.

The Geriatric Depression Scale (Yesavage, Brink, Rose, Lum, Huang, Adey & Leirer, 1982-1983) is a 30-item self-report screening measure used with adults over the age of 60 (Yesavage et al., 1982-1983). It is valid and has good internal consistency including in older adults (Hamilton, 1960; Zung, 1965; Parmelee, Powell Lawton & Katz, 1989; Debruyne, van Buggenhout, Le Bastard, Aries, Audenaert, DeDeyn and Engelborgh, 2009; Alexopoulos, Abrams, Young & Shamoian, 1988). A lower score indicates more depressive symptoms.

Beck Depression Inventory (Beck, Ward & Mendelson, 1961) is a very widely used measure consisting of 21 statement groups of statements. Numerous studies attest to its validity, psychometric properties and consistency (Beck, Steer & Garbin, 1988; Richter, Werner, Heerlein et al., 1998). The lower the score here the less depressive symptoms reported.

Independent Samples t-tests showed no differences between the groups on all screening measures and scores were within normal range so no participants were excluded from the study.

	Group	Mean	Std. Dev.
Geriatric Depression Scale	Control	4.00	3.693
	AWMT	4.00	4.068
Becks Depression Inventory	Control	5.67	4.334
	AWMT	5.50	3.989

Table 5.1: Screening measures of participants

Outcome measures

The outcome measure used were a mixture of computer-based, table-top, questionnaire and every-day measures and were completed out in the following order - The Goal Management Questionnaire (GMQ), the Cognitive Failures

Questionnaire (CFQ), Raven's Standard Progressive Matrices, The Games Evaluation Task (GET), The Doors and People Test, the Cattell Culture Fair, and Digit Span, the Multiple Errands Tasks (MET; where completed) and the Hospital Anxiety and Depression Scale (HADS).

The Cognitive Failures Questionnaire (Broadbent et al., 1982) is a 25-item questionnaire which gives an approximation on how often participants might have slips of memory and cognition in daily life or make everyday mistakes.

Cattell Culture Fair Scale 2 Form A (Cattell & Cattell 1960; 1973) was developed as a measure of reasoning that was relatively independent of education and culture by using abstract patterns that were notionally novel for everyone. It consists of four different sections and contains forty-six items in which participants need to identify the relationship and rules that exist between shapes and figures. This measure is time limited.

The Raven's Standard Progressive Matrices (Raven's SPM; Raven, 1979) is an untimed novel problem-solving measure of fluid intelligence that participants completed at the beginning of the testing session. It contains 60 items (5 sets of 12 items) that increase in difficulty and is described in more detail in chapter four.

The Games Evaluation Task is a measure of planning and self-organisation, described in chapter two. Participants are required to play each of four tasks, switching between them as necessary. The number of games played and the time-allocation are scored.

The Goal Management Questionnaire is an self-report questionnaire developed to closely map the content of Goal Management Training (Manly et al. personal communication). It is described in further detail in chapter four. It comprises 34 items addressing everyday coping including memory and organisation.

The Multiple Errands Task (Shallice & Burgess, 1991). As detailed in chapter three, in this test participants were asked to carry out a set number of tasks – collecting information and purchasing items – within a time limit and following rules in a natural (and therefore somewhat unpredictable) setting. Performance was assessed by a rater who followed participants during the task and took note of rule breaks, inefficient strategies, pieces of information collected etc. This is a lengthy assessment that took place on a separate testing occasion and was optional. Consequently not all participants completed this assessment.

The Doors and People Memory Test (Baddeley, Emslie & Nimmo-Smith, 1994) is a neuropsychological test of memory that looks at recall, recognition and forgetting. It consists of four sections and uses faces, names, shapes and doors as stimuli. It takes 35-45 minutes to administer. Updated cut-off scores have been recommended by Davis, Bradshaw and Szabadi (1999).

Digit Span – Digit span was selected as a notional outcome measure from the AWMT training set. The highest two scores obtained were averaged to give the "digit span" score for each participant.

The Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983) is a self-assessment measure of depression and anxiety. It is a brief measure and consists of 14 statements rated on a scale as to how frequently they occur.

Development of training tasks

The training tasks were developed by, and generously loaned to this study by, Adam Hampshire at the CBU in part to support a large-scale on-line study run in collaboration with the BBC (Owen et al., 2010). To maximise inclusion and international replication the emphasis was on non-verbal and numeric materials/tasks. Prior to the tasks and the website being used by participants it was piloted with 5 people (aged 26-56 years). Feedback from this group was useful in identifying and correcting technical issues before the study began.

The training tasks, as described in more detail in the previous chapter, were as follows:

Digit Span (forward and backward): This task required participants to re-enter seen sequences of single digits either in the order presented or the reverse order. The task started with 3 digits and increased or decreases in difficulty by one digit depending on performance. If 3 consecutive errors occurred, the task terminated.

Spatial span: Participants were required to remember the sequence in which boxes lit up on a 4x4 grid.

Monkey ladder: Participants were required to remember the sequence of a set of numbered boxes after the numbers disappear from the boxes and reproduce the order.

Paired associate: Participants were required to remember images which appear on boxes and then disappear. Another box appeared with an image on it and participants clicked on the box which earlier contained that image. Figure 5.1 below shows an example of this task.

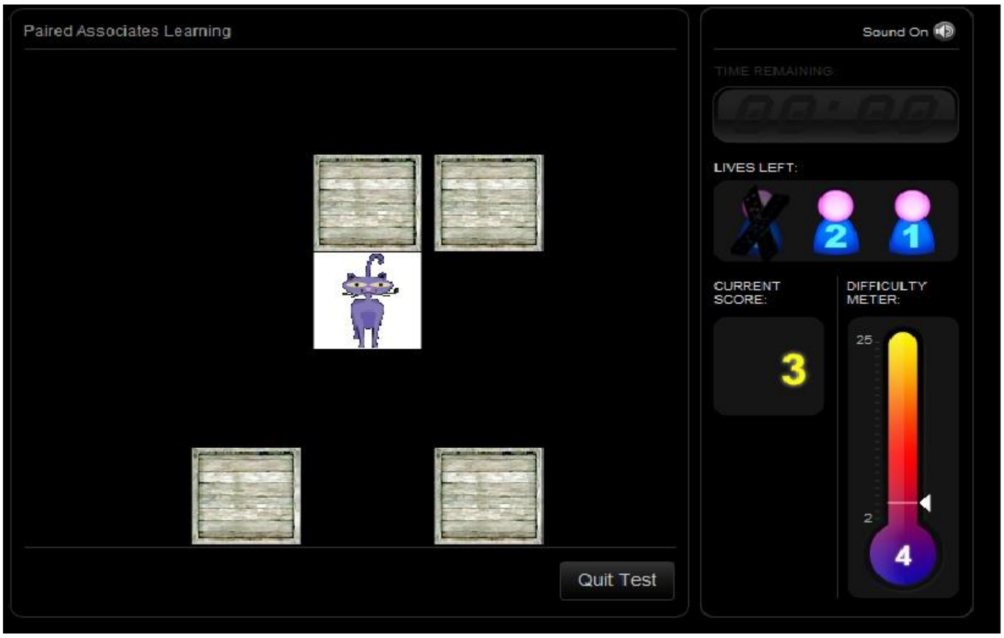


Figure 5.1: Screenshot of paired associate learning task

Training regime

Following assessment participants were emailed details of how to begin the training. They were encouraged to train five days a week but allowances were made if participants had to miss a day of the training. Participants were emailed links to each of their daily training sessions. The training consisted of twelve tasks each day and took between 25 and 35 minutes. Because of the limited number of tasks participants carried out the five tasks described above a number of times – this was paired associates, monkey ladder and spatial span ladder three times each and digit span three times - backward and/or forward. The order of presentation varied daily. The training consisted of 25 sessions. Control participants came in at baseline and re-assessment and had no contact between the two sessions except to arrange next appointment.

Results

Compliance

Compliance in all participants was high. Occasional sessions were missed or deferred but no participant's adherence gave rise to concern about their inclusion. Only one participant did not complete all sessions. This participant missed only one session. The average duration of the training was 28minutes 15seconds.

Differences following training

Independent samples t-test and One-way ANOVAs showed no differences between the two groups on age or test variables at baseline. Training effects were examined using repeated measures ANOVA with the key dependent variables from each measure at time 1 and time 2 as the within-subject factors and group (AWMT vs. Control) as the between-subject factor. Whilst the issue of multiple comparisons is important, in the following section uncorrected P values are reported. The key measure of differential change in one or other group is the presence of a statistically significant time x group interaction term.

Outcome Measure	**Group**	**Time 1**	**Time 2**	**Main Effect of Time**	**Group x Time Interaction**
		Mean (SD)	**Mean (SD)**		
Cognitive Failures Questionnaire	Control	31.91 (10.26)	32.83 (12.22)	F(1,22) = 0.46, P = 0.504	F(1,22) = 0.0, P = 1.0
	AWMT	34.83 (12.55)	35.75 (14.34)		
Goal Management Questionnaire	Control	87.66 (51.24)	83.83 (57.44)	F(1,22) = 1.68, P = 0.21	F(1,22) =0.54, P = 0.47
	AWMT	88.25 (56.18)	74.33 (49.91)		
Cattell	Control	32.17 (3.90)	31.92 (4.78)	F(1,22) = 2.10, P = 0.65	F(1,22) = 0.53, P = 0.47

	AWMT	33.75 (4.29)	34.83 (4.76)		
Raven's Score	Control	49.42 (4.80)	50.25 (4.51)	F(1,22) = 0.55, P = 0.47	F(1,22) = 0.06, P = 0.81
	AWMT	52.75 (4.81)	53.17 (4.17)		
Raven's Time (seconds)	Control	2073.83 (509.03)	1719.58 (414.83)	F(1,22) = 16.21, P = 0.001	F(1,22) =0.03, P = 0.87
	AWMT	2069.42 (754.82)	1684.00 (525.40)		
GET Tasks	Control	7.85 (1.62)	7.50 (1.31)	F(1,22) = 0.038, P = 0.85	F(1,22) = 0.15, P = 0.70
	AWMT	7.33 (1.15)	7.58 (1.17)		
GET Time Dev. (seconds)	Control	151.67 (81.90)	132.79 (52.42)	F(1,22) = 2.74, P = 1.112	F(1,22) = 0.27, P = 0.61
	AWMT	143.33 (56.82)	107.17 (68.43)		
HADS (Depression)	Control	2.58 (1.51)	2.33 (1.43)	F(1,22) = 0.39, P = 0.54	F(1,22) = 0.99, P = 0.76
	AWMT	3.00 (2.08)	2.91 (1.83)		
HADS (Anxiety)	Control	3.17 (2.66)	3.16 (2.37)	F(1,22) = 0.40, P = 0.54	F(1,22) = 0.99, P = 0.76
	AWMT	4.50 (3.48)	4.92 (3.40)		
Doors and People	Control	13.83 (3.27)	15.17 (2.21)	F(1,22) = 15.36, P = 0.001	F(1,22) = 1.19, P = 0.29
	AWMT	12.72 (3.13)	15.09 (2.30)		
Digit Span	Control	5.28 (0.56)	6.32 (0.90)	F(1,22) =16.77, P = 0.001	F(1,22) = 1.41, P = 0.26
	AWMT	5.56 (0.61)	6.22 (0.82)		

MET Time (seconds)	Control	1390 (239.81)	1382.5 (291.53)	F(1,22) = 0.78, P = 0.41	F(1,22) = 0.66, P = 0.44
	AWMT	1475.33 (72.32)	1296 (303.39)		
MET Tasks	Control	9.50 (0.55)	9.17 (0.75)	F(1,22) = 1.56, P = 0.25	F(1,22) = 0.00, P = 1.00
	AWMT	9.33 (0.58)	9.00 (1.00)		
MET Mistakes	Control	2.50 (2.43)	1.50 (0.84)	F(1,22) = 0.60, P = 0.463	F(1,22) = 0.02, P = 0.88
	AWMT	3.33 (1.16)	2.67 (2.31)		

Table 5.2 Repeated Measures ANOVA of key variables

As can be seen in the table above there were no statistically significant differences following training on any variables when compared with the control group. It is perhaps worth noting, given the relatively small sample, that AWMT was associated with numerically greater gains in 7/9 cognitive measures in which change was likely due to the absence of ceiling effects (GMQ, Cattell accuracy, Raven's time, GET time allocation, Doors and People and MET time). Only on the CFQ and Raven's accuracy did the control group show numerically greater gains over the training interval.

Practice Effects

Although not a primary aim of the study, it was of interest to see if participants improved on tasks from beginning to the end of training. The practice effects of tasks were especially of interest here when there was no transfer effect seen on outcome measures. The final two scores (usually session 24 and 25) that participants obtained were averaged for each participant at the beginning and end of training. All data reported are two-tailed.

There were differences in the AWMT group before and after training on digit span t(11)= -3.75, p=.003. Significant differences were also seen on spatial span t(11)= -4.10, p=.002, and backward digit span t(11)= -2.68, p=.02. Following training on monkey ladder task the difference was nearing significance t(11), -2.11, p=.059 but not significant. Difference on paired associates was not significant following training t(11), = -1.27, .23. Differences pre and post training reached significance in most of the training tasks.

Power and Effect Size

A post-hoc power analysis was conducted and was calculated by using the sample size, effect size and the probability level. Observed power/retrospective power analysis was not used as it is based on the assumption that the sample effect size is the same as the effect size in the population it is taken from (Field, 2009) and as the sample here is relatively small this may be false. A priori power analysis would have been preferable but as it was difficult to control the sample size due to time and resources available a post-hoc power analysis was used. The small, medium and large effect sizes described by Cohen (1988) were used to estimate effect sizes.

The effect size was calculated using the values from the contrasts that were obtained from the repeated factors ANOVA – value for F for Group x Time interactions. Because the analysis looked at focused comparisons between the two groups the following equation was used to calculate the effect size (Field, 2009):

$$r = \sqrt{\frac{F(1, df_R)}{F(1, df_R) + df_R}}$$

Taking results from Doors and People as an example (all the other variables were calculated in the same way to obtain effect size and power)

$$r = \frac{1.19}{1.19 + 22}$$

$$r = \sqrt{0.051315}$$

$$= .23$$

Small effect sizes were seen in all variables – CFQ= 0.00; GMQ = 0.13;Cattell = 0.13; Raven's Score = 0.19; Raven's Time= 0.04; GET tasks = 0.08; GET time deviation = 0.11; HADS Depression = 0.21; HADS Anxiety = 0.21; Digit Span 0.25; MET Task= 0.00;MET Time = 0.17 and MET Mistakes = 0.03.

A post hoc power analysis was conducted using the software package, GPower computer program (Faul, Erdfelder, Buchner & Lang, 2009). The alpha level used for this analysis was $p < .05$ and the sample size was 24. The post hoc analyses revealed the statistical power for the Doorsand People with this sample was .30. Output from G* is as follows:

Analysis:	Post hoc: Compute achieved power		
Input:	Tail(s)	=	One
	Effect size |ρ|	=	0.23
	α err prob	=	0.05
	Total sample size	=	24
Output:	Noncentrality parameter δ	=	1.1578053
	Critical t	=	1.7171444
	Df	=	22
	Power (1−β err prob)	=	0.3006219

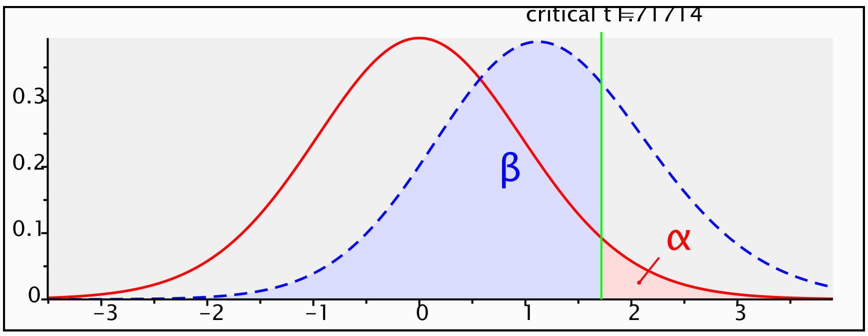


Figure 5.2: Output from G* power of CFQ.

The other outcome measures were calculated in the same manner and the power was as follows:

CFQ= 0.05; GMQ = .15; Cattell = .15; Raven's Score = .23; Raven's Time = .07; GET Tasks = .10; GET Time Deviation = .13; HADS (Depression) = .27; HADS (Anxiety) = .27; ; Digit Span = .34; MET Time = .20; MET Tasks = .05; and MET Mistakes = .07

Discussion

Following twenty-five sessions of training on AWMT, numerically greater benefits were observed on 7/9 key cognitive variables relative to those seen in an untrained control group. Statistical analysis revealed that none of these benefits approached significance.

At the time this study was conducted the result with respect to Cattell and particularly Raven's PM was surprising in that they contrasted with results from adaptive WMT in the healthy population of apparently similar intensity (e.g. Jaeggi et al., 2008; 2010; Schweizer et al., 2011). The much larger sample sizes in those studies (35-104) was one possible account of this difference. However, the results here are entirely consistent with those subsequently reported by Owen et al. (2010) using the same training tasks. This suggests that the current sample could have been multiplied by 1000 and still not produced anything beyond minimal effect sizes.

The contrast between the results of Jaeggi et al (2008; 2010) and Owen et al. (2010) and here are unlikely to stem from the duration of training (Jaeggi et al's sample completed between just 8-19 days). The difference could be related to the training tasks which for Jaeggi et al. (2008; 2010) and Schweizer and colleagues (2011) were n-back measures. There is no doubt that n-back measures do become extremely challenging as n increases but so too do span and other tasks used in the current study. It is hard to argue that one task that exceeds capacity is 'more difficult' than another. It is possible however that there are important difference in the dynamics of adaptive training between the different sorts of tasks and adaptations that are relevant and this required further evaluation. Against this, Redick et al (2011) applied very similar n-back tasks to Jaeggi et al. to a large healthy young sample and failed to replicate Jaeggi et al.'s fluid IQ transfer effect.

Redick et al. advance several reasons for this difference including the use (in their case) of an active control and, interestingly, their efforts to reduce participant expectation that general cognitive gains might occur. In the current study this was left more ambiguous as participants were told that we were trying a number of different training packages with the possibility that they were participating in a control condition.

The issue of expectation is potentially important. It is possible to imagine two rather different styles of training. In one, as a paid participant with no expectation of generalised gains, it would be possible to go through the motions of training and doubtless show performance gains on the task but without real engagement, without thinking about the implications of what you were doing for other aspects of your life. In contrast, a person who expects change and for whom this is the primary motivation in taking part may have a quite different experience and may, as a consequence, be better able to 'take' capacity, stamina, persistence, confidence etc. from the trained to untrained situations. Whether such effects are important in moderating training remains to be seen.

In summary and contrary to original hypothesis the current study found no evidence of transferable benefits of 25 sessions of adaptive computerised WMT in the healthy older population.

Chapter Six:

Problem solving training in a healthy older adult population

Abstract

This chapter describes an intensive computerised training programme that was undertaken by ten older adults (mean age 67.5years). The training was based on novel problem solving and consisted of twenty-five days of training for 20-35minutes over five weeks. Following training participants showed improvements relative to the untrained control group in only one of the variables, self-reported anxiety. Although caution is required due to the sample size, the results provide little support for the use of this particular training in isolation.

Introduction

The emphasis so far in this thesis has been on WMT and the development and evaluation of new outcome measures to examine potential transfer of training effects to performance in complex everyday situations. The rationale for training WM has been the ubiquity of situations that require the online retention and manipulation of information and/or increasing general intellectual capacity via this route. Span and n-back tasks lend themselves well to this kind of adaptive training because manipulating difficult amounts to +/- 1. Other approaches are possible however and apriori it seems likely that the broader the base of cognitive tasks on which people train, the more probable transfer to a set of untrained tasks would be.

An obvious candidate set of tasks are those that are most closely related to g and which strongly activate multiple demand regions. The work of Plemons et al. (1978) and Willis et al. (1981) provides one guide. Here the operations required to perform one such test, the Cattell, were analysed and participants were trained in identifying these underlying regularities. As a consequence, their performance on untrained though similar problems was enhanced.

An interesting issue, not least in terms of the underlying processes necessary to solve them, is whether Cattell-like tasks can be *automatically* generated at different level of difficulty to support the large number of trials at each level required for computerised training. An early prototype, developed for this study was as follows:

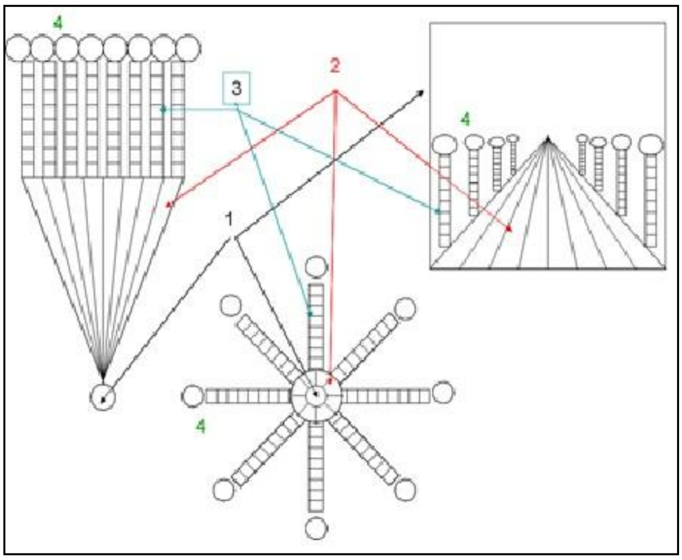


Figure 6.1. Three analogous pattern structures to which similar rules could be applied to generate problem solving items of progressive difficulty.

As shown in figure 6.1, one model is to develop analogous but ostensibly different problems to which the same rule transformations can be applied. Here 3 or more objects could be generated which have a common underlying structure. This comprised a single object (see 1 in figure 6,1), an array with 8 segments (2), a further array (3) each unit of which is subdivided and so on. Easy logical sequence items could be generated by applying one rule to one element and keeping all others constant (as shown in figure 6.2 below). In this case a highlighted element in array 4 moves systematically around the figure. Alternative simple rules would include a single element alternating on and off, the highlighted element flipping from one side of the figure to the other and so on. The 3 incorrect response options can be generated by applying the same rule but starting at a different point.

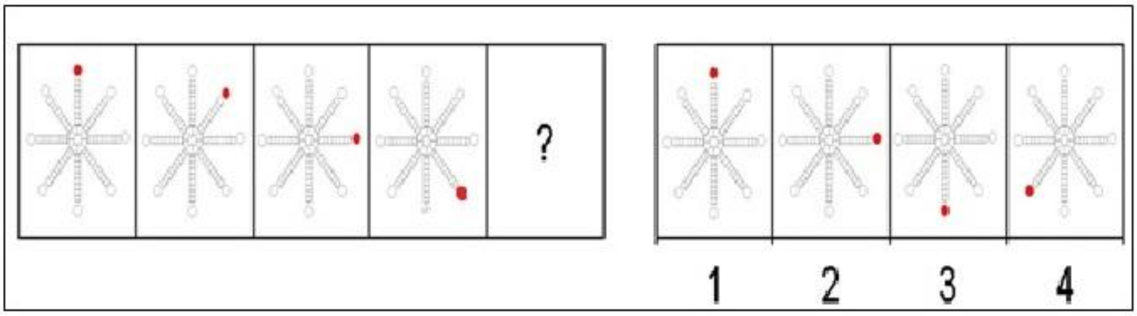


Figure 6.2. An easy item in a logical series.

The problem could be made more challenging by bringing two simple rules into play. (e.g. Object 1 has an alternation rule applied, object 2 is constant (off) object 3 has 1 element selected segment rule applied and object 4 is off) – again with wrong answers being generated by different randomly selected start points.

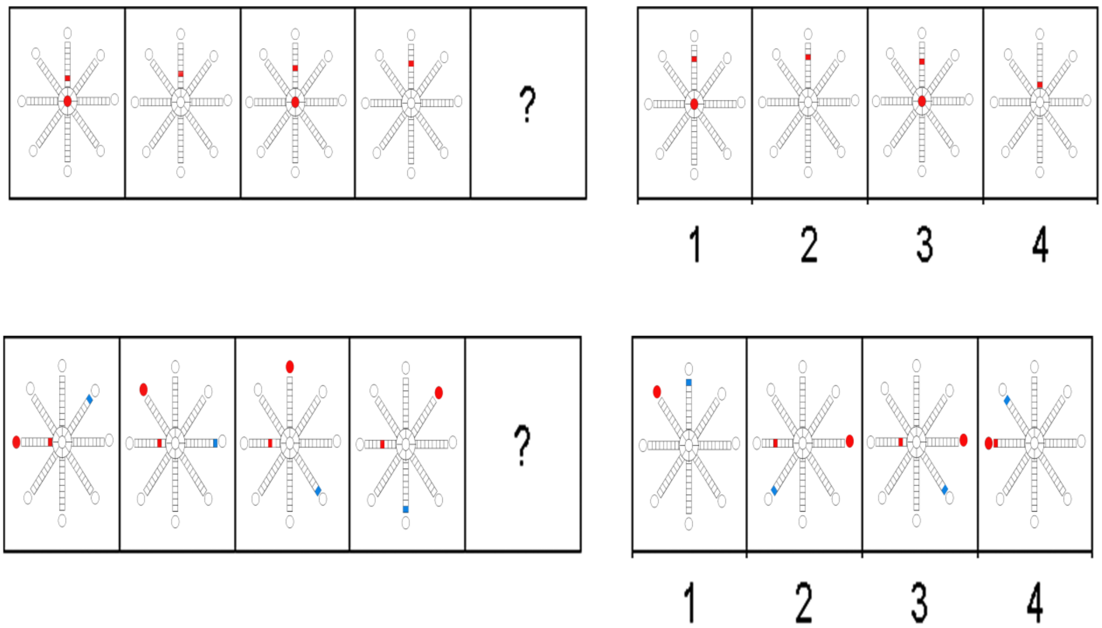


Figure 6.3. Increasing the complexity of a logical series by applying two simple rules

Difficulty could then be increased by applying simple independent rules to all elements. Introducing conjunction rules (if object 1 is in state 1, then object 2 will behave in one way; if object 1 is in state 2, object 2 will behave in another).

It would be expected that participants would gain experience of tracking each object array for simple rules. However, this could be offset by introducing conjunction rules within the constraints of being able to pick up the series within 4 items. For example, a simple conjunction rule would be if the central item is highlighted, one other element is also always highlighted.

A key challenge in high g tasks may be in casting aside the rule determining the response and even the type of task presented in the previous item in order to reset for the current item. Accordingly, in planning the tasks for this study the aim was for participants to have to switch between figural logical sequences (such as above), making fine perceptual discriminations to discern whether two figures were identical, detecting the odd-one-out based on relationships between objects, solving letter-number codes and the pattern in mathematical sequences.

Discussions of these topics, the pragmatics of programming for on-line training and issues of what was easily visible within a small screen area contributed to final set of Problem Solving Training (PST) tasks, most particularly the Object Reasoning, Odd One out and Letter Codes tasks. In addition Adam Hampshire generously allowed the use of his Hampshire Tree Task, Spatial Spider and adaptation of a grammatical relations task (see below).

The aim of this study was therefore to examine whether an adaptive computerised training based on a variety of novel problem solving tasks would fare better than the adaptive WMT described in the previous chapter.

Methodology

Participants

All participants in this study were over the age of 55 years and were recruited from the MRC-CBU database of healthy volunteers. All participants volunteered to take part in the study after being contacted by email or phone. All who took part were paid an honorarium for their contribution to the study. Participants had no history of significant mental health problems, head injury, or neurological impairment. Potential participants were screened and recruited over the phone and through the questionnaires sent to them by post following initial screening over the phone.

Participants were required to have access to a computer with broadband connection. They also had to be available for the entire length of the study as the training needed to take place in one full block with no long interruptions. It was taken into consideration that participants might need to miss a day or two of training for various reasons but in order to be included in the final sample participants had to undertake 20 of the 25 training sessions. Following recruitment and screening participants were randomly assigned to groups, as described in chapter five through randomization.com. Participants in groups were not matched.

The control group used was the same as that described in chapter five.

Four participants dropped out following initial assessment in the training group – they were unable to commit to the time required of in the study. Following this the sample consisted of eleven participants. The panel for healthy older subjects had at that point almost been exhausted and as different recruitment method than in the previous chapters may influence results, as the sample might be different, it was decided to leave the sample at eleven. A total of 10 of the 11 participants tested at baseline completed the full training schedule. Participants had to complete 20 out of the 25 training sessions in order to be included in the analysis. All 12 control participants assessed at baseline came to the follow-up re-assessment session.

There were 10 participants (6 female)in the training group – mean age 67.5 years (range 62 – 77) - and 12 participants (8 female) in the control group- mean age 70.36 years (SD 6.14). Overall mean of the group (n=22) was 69 years (SD 5.59). Although the control group was older the two groups did not differ in terms of age $t(19) =.251$. All except one participant in the control group was educated to at least degree level, while in the training group only four participants had a degree or higher.

Participants were informed prior to taking part that if one of the training programmes (working memory or problem solving training) was clearly superior participants who were in the other group will be offered this training package.

The first and final assessment sessions were conducted, at a convenient time, in the CBU – this allowed participants to get to know the training set-up, ask questions, complete assessments and also to report back on their experiences. The remainder – the training sessions- were conducted in the participants’ own homes using their own computers. It was hoped is that any significant difficulty in participating in the training would emerge in the initial meeting with participants. If however, people found the procedure onerous or inconvenient it was stress throughout that they could withdraw from the study at any stage without having to give a reason.

Outcome measures

Outcome measures that were used were the same as those described in the working memory training study, chapter five. They consisted of Cognitive Failures Questionnaire, Goal Management Questionnaire, The Cattell Culture Fair, Raven's Standard Progressive Matrices, The Games Evaluation Task, Hospital Anxiety and Depression Scale and the Doors and People Test. Participants were posted the self-rated questionnaires before and after training and the remainder of the assessments were carried out at assessment sessions in the MRC- CBU. The same procedure was carried out as described in the previous study. There were not enough MET scores in this training group as in the previous study so this was not included in analysis. It is thought that as this was an optional part of the study and it may have been a busy time of year for participants (Christmas) more decided against taking part, or did not complete the re-assessment of the MET. Digit Span was not assessed in training group as it was measured at the beginning and end of training with the AWMT group and was not part of training schedule for this group.

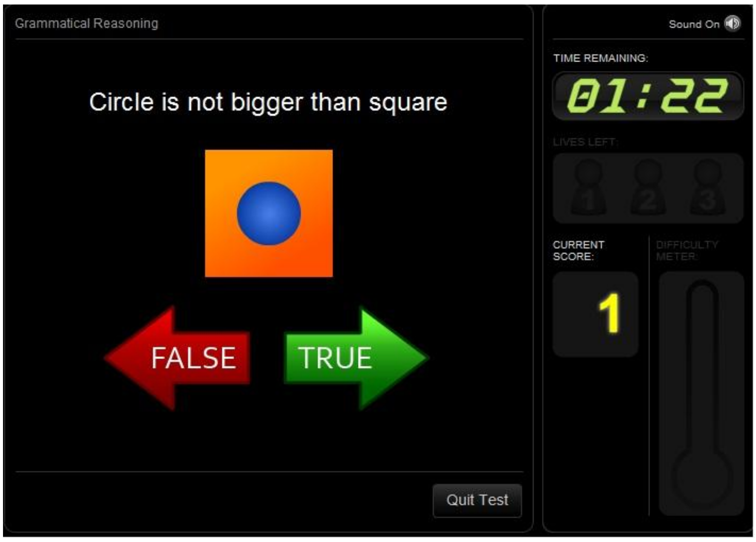
Beck's Depression Inventory and the Geriatric Depression Scale were used as screening measures for participation in the study. No participants were over the cut-off mark of what was considered normal functioning on each of the measures, except for one participant whose sister had died prior to assessment session. Assessments were returned by participants prior to the initial session and so this participant was contacted and advised against taking part in the study at that time. This participant joined the study at a later date (scores at this point were within normal range).

Problem Solving Training tasks

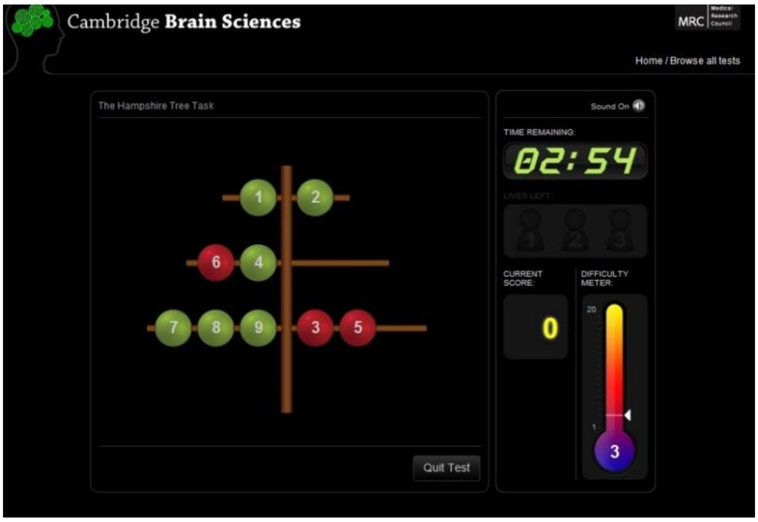
The tasks that participants trained on were as follows (task order changed daily).

Spatial Slider – In this task numbered tiles appeared on the screen arranged in numerical order on a grid. The tiles were then shuffled so that some were no longer in the correct positions. The tiles had to be rearranged into the correct order in as few moves as possible by sliding them in and out of the blank spaces.

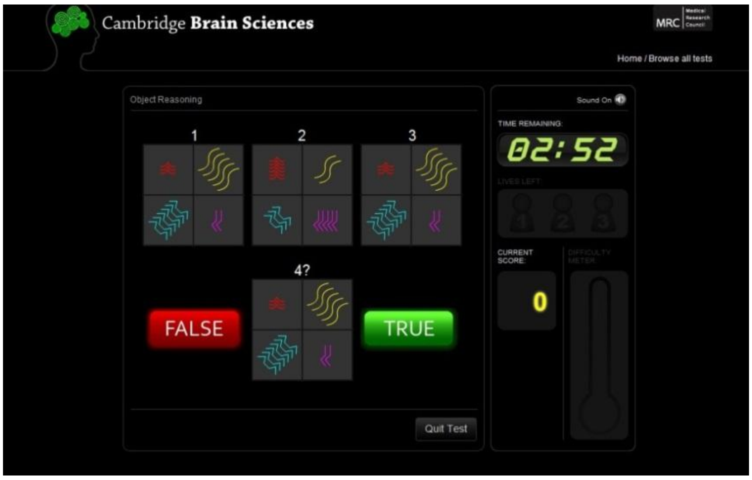
Grammatical Reasoning (1 and 2) – In this task a statement appeared at the top of the screen with two shapes underneath. The statement described a relationship between the shapes. Participants had to decide whether the statement was true of false according to the, completing as many as possible correctly within each 90-second trial.



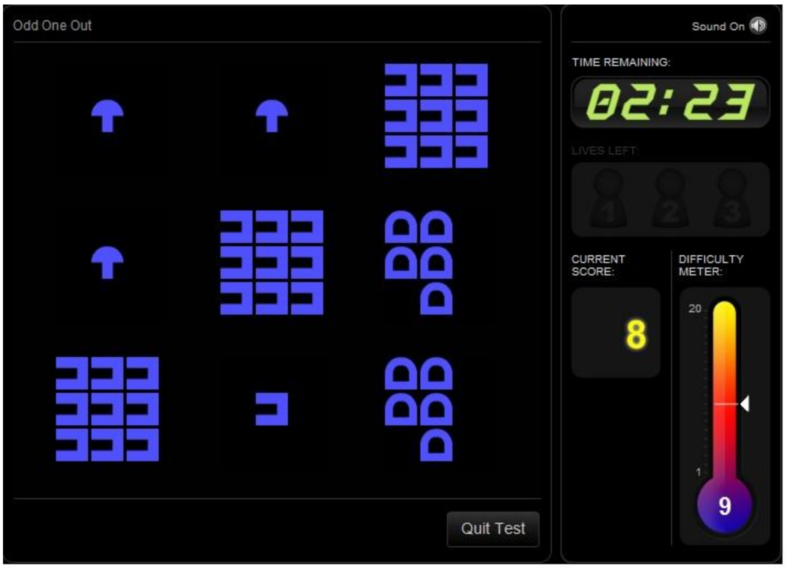
Hampshire Tree task – In this task a tree-shaped frame appeared, as shown below. Nine numbered balls were on the branches of the tree. Participants were required to re-arrange the balls in numerical order with as few moves as possible. Only balls that were not blocked could be used and only one ball could be moved at a time. Participants have three minutes to solve as many problems as possible



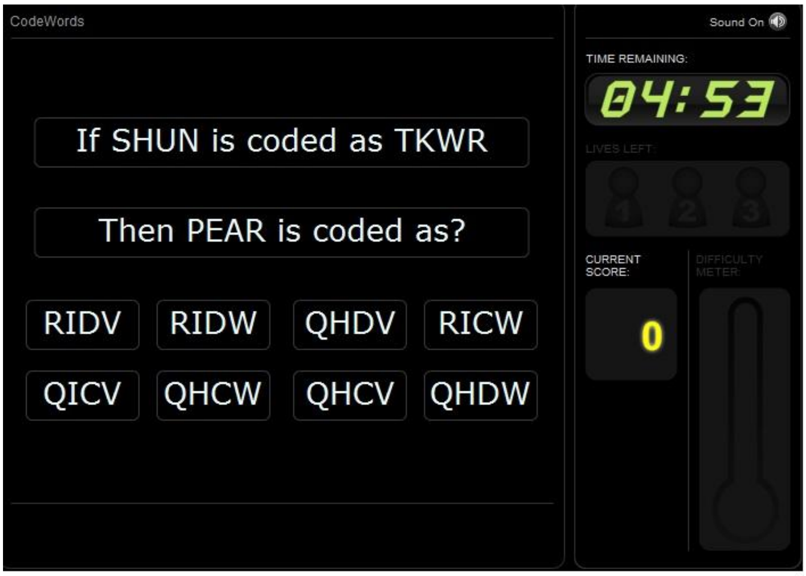
Object reasoning – In this task four panels appeared on the screen – three on top and one on the bottom. The panels on the top differed from each other according to a rule or a number of rules. Participants had to deduce the rule(s) and decide if the final panel logically completed the sequence or not. Participants had three minutes to solve as many problems as possible.



Odd One Out - Nine patterns appear in a grid on the screen, as below. Each pattern is comprised of a set of features, for example, colour, shape and number. One of the patterns is different from the others because of one or more of its features. Participants need to work out which pattern is the odd-one-out and click on it.



Code Words- Participants are required to figure out the rules that code a certain word, as shown below and they then have to apply the same rule to a new word in order to come up with the code word for that new word. They can choose from eight options.



Procedures and training regime

Training was conducted at home, over the internet, and took between 20 minutes and 35 minutes a day (depending on performance on tasks) for 5 weeks. All training sessions were designed to be roughly of equal length every day. Each assessment session took between 1hour 15minutes and 2 hours.

Following initial assessment participant were sent an email that provided links to their first day's training. They were provided with a username (their email) and a password. When they opened the link provided the first page requested log in details. From this they had to consent to take part (daily) and for their data to be used. They were then taken to the start page. Once they clicked on "Start" button the training began. Tasks took different times lengths but generally were 3minutes in length. Once one task was finished the next task was automatically queued. Participants at this point could take a brief rest, if required, before starting on the next task. When all the day's tasks were completed they were informed that the training was finished and they could close their browser. Results were stored on the training website. The investigator could access them but participants could not. Participants were sent ten links at each time and told to email the investigator when they had completed the links sent and required more. Through this it was hoped to increase compliance with the training regime and also it was an opportunity for participants to ask any questions they might have had or to let the investigator know if there was any difficulties with the website (which was the case for some participants) as it was a newly-developed website.

Results

Direct effects of training

Within the training tasks, as would be expected, significant initial-final differences were seen on all of the tasks (Spatial Slider $t(9)=-4.20$, $p=.002$; Grammatical Reasoning $t(9)= -10.51$, $p=.000$; Hampshire Tree Task $t(9)= 5.95$, $p=.000$; Object Reasoning $t(9) = -3.41$, $p=.008$; Odd-one-out $t(9)= -2.47$, $p=.036$; and Code Words $t(9) = -4.58$, $p=.001$.

Generalisation

Following training, compared with the untrained control group, participants showed a numerically greater gain in Cattell, GET and Doors and People Scores as well as on the HAD mood measures. These differences did not however reach statistical significance in terms of a group x time interaction other than the HADS anxiety score ($p<0.05$).

Outcome Measure	Group	Time 1	Time 2	Main Effect of Time	Group x Time Interaction
		Mean (SD)	Mean (SD)		
Cognitive Failures Questionnaire	Control	31.92 (10.26)	32.83 (12.22)	F(1,20) = 0.51, P = 0.48	F(1,20) = 0.02, P = 0.88
	PS	38.00 (14.25)	38.60 (13.76)		
Goal Management Questionnaire	Control	88.67 (51.24)	83.83 (57.44)	F(1,20) = 0.11, P = 0.75	F(1,20) = 0.08, P = 0.78
	PS	77.10 (40.86)	76.80 (31.00)		
Cattell	Control	32.17 (3.90)	31.92 (4.78)	F(1,20) = 1.17, P = 0.29	F(1,20) = 0.59, P = 0.45
	PS	30.60 (4.97)	29.10 (6.21)		
Raven's Score	Control	49.42 (4.80)	50.25 (4.52)	F(1,20) = 0.31, P = 0.58	F(1,20) = 0.83, P = 0.38
	PS	48.20 (5.25)	48.00 (5.50)		
Raven's Time (seconds)	Control	2073.83 (509.03)	1719.58 (414.84)	F(1,20) = 22.10, P = 0.00	F(1,20) = 2.13, P = 0.16
	PS	2123.70 (478.56)	1937.40 (342.25)		
GET Tasks	Control	7.58 (1.62)	7.50 (1.31)	F(1,20) = 1.57, P = 0.23	F(1,20) = 2.13, P = 0.16
	PS	5.90 (1.66)	7.00 (1.56)		
GET Time Dev. (seconds)	Control	151.67 (81.90)	132.79 (52.42)	F(1,20) = 6.79, P = 0.02	F(1,20) = 2.48, P = 0.13
	PS	244.90 (75.70)	168.40 (93.54)		
HADS (Depression)	Control	2.58 (1.51)	2.33 (1.44)	F(1,20) = 2.71, P = 0.12	F(1,20) = 0.61, P = 0.45
	PS	2.70	2.00		

		(1.77)	(1.56)		
HADS (Anxiety)	Control	3.17 (2.66)	3.17 (2.37)	F(1,20) = 4.82, P = 0.04	F(1,20) = 4.82, P = 0.04
	PS	6.60 (4.53)	5.20 (3.01)		
Doors and People	Control	13.83 (3.27)	15.67 (2.21)	F(1,20) = 14.78, P = 0.001	F(1,20) = 2.89, P = 0.11
	PS	12.67 (2.92)	16.11 (1.62)		

Table 6.1: Repeated Measures ANOVA on key variables

Power and Effect Size

Cohen's (1994;1988) levels of effect size were used here: small= 0.1; medium= 0.3; large =0.5. The alpha level used for this analysis was *p* < .05 and the sample size was 22 and the effect size was calculated using contrasts from repeated measures ANOVA and the following equation:

$$r = \sqrt{\frac{F(1, df_R)}{F(1, df_R) + df_R}}$$

Effect sizes were small for CFQ, GMQ, Cattell, Raven's Score and HADS Depression 0.03, 0.06, 0.03, 0.20, 0.17 respectively.

Medium effect sizes were seen in Raven's Time = 0.31; GET Time 0=.33 and GET Tasks 0.31; HADS Anxiety = 0.44 and Doors and People 0.36.

Power was calculated as described in chapter 5 using GPower computer programme (Faul, Erdfelder, Buchner & Lang, 2009). From this the power was calculated and was as follows for each of the outcome measures: CFQ = .07; GMQ = .08; Cattell = .07; Raven's Score = .24; Raven's Time = .43; GET Tasks = .43; GET Time deviation = .48; HADS Depression = .07; HADS Anxiety = .72; Doors and People = .54.

Discussion

The results show that, as with S.M., substantial and statistically significant gains occurred across a range of challenging and adaptive problem-solving training tasks. Comparisons of untrained outcome measures before and after the training, however, revealed little in the way of disproportionate effect beyond that seen with just retest in the control group. The significant change in anxiety score was not predicted and would require further replication before being considered reliable.

Why are gains not transferring to what seem to be conceptually similar measures? The first point is that they might not be as similar at the relevant level as was imagined. The Cattell and Raven's tasks were developed by people applying their intellect and creativity to each item. In switching from one item to the next there is often need to adopt a quite different stand-point to uncover the problem. In the automated tasks used here, blocks were conducted in which rather similar challenges were repeatedly at play; that may simply not have prepared participants enough for the clinical tests. It would be relatively simple to change the block structure of the training. If there is something inherently problematic with automatic production however, the future of computerised training of this sort seems in doubt.

A second and possibly not trivial issue is that even intelligent participants were not able to see past the surface structure to the problems. Training and gains occurred on computerised tasks in a particular and constant screen structure (e.g. the tasks being central and to the same scale, a clock to the right, tasks timing out on you etc.). If improved performance is only possible or likely within that tightly constrained framework, the potential of such tasks for rehabilitation is limited.

A third point is whether the sensitivity of the outcome measure is too coarse to pick up changes over the training duration. This is possible but seems unlikely given the sensitivity of the Cattell and Raven's to small individual differences. The 'dose' of training delivered here was equivalent to and in some cases much greater than some studies reporting positive results from WMT. Examining whether increasing the dose would improve the situation would benefit from stronger preliminary evidence. A 'sleeper effect' in which initially small and under detectable effects of training increase over time, e.g. though greater application in everyday life, is possible. Detecting it would require randomised trials with long follow-up.

A clear possibility is that the current study was not sufficiently powered to detect real differences. Medium effect sizes were detected on a number of measures which is suggestive that there may be some evidence of training effects if the sample was larger. Most of the WMT studies reviewed in the first chapter featured samples 3 times the size available here. However, in light of the Owen et al.'s (2010) negative result from more than 11,000 healthy participants completing training tasks including those used here it was pragmatic, in the time-scales of a PhD, to discontinue this data collection and focus on the final study described in the next chapter.

Chapter Seven:

Combined strategy training with a healthy older adult population

Abstract

As has been discussed, in neuropsychological rehabilitation the evidence that one can enhance capacity via specific training is scant. In contrast, the evidence that patients can adopt a range of strategies that minimise the functional consequences of the impairment is ubiquitous. In the domain of executive impairment a range of such interventions, including von Cramon et al.'s Problem Solving Training (von Cramon et al., 1991) and Goal Management Training (Spikman et al., 2011) have been linked with generalised changes (although less frequently have changes to everyday function been examined).

In light of the failure to detect spontaneous generalised changes in cognitive function in healthy older participants in the preceding chapters, this chapter examines the idea of combining computerised 'capacity' training with explicit strategies designed to promote facilitation.

Thirteen older adults (mean age 71.62 years) took part in this 25 day training programme. Participants trained for 5 weeks on progressively increasing duration cognitive training. Nine videos relating to the training and its application to everyday life were developed as part of the package. Unlike the previous chapters, participants were also encouraged to keep a logbook and set regular goals for themselves. Despite these features, significant detectable transfer effects remained absent in this group.

Introduction

The context giving rise to the current study has been well rehearsed in previous chapters. Within neuropsychological rehabilitation attempts to improve function by repetitive training of specific cognitive functions has generally proved disappointing. In contrast there is good evidence that many types of intervention that help patients to strategically offset the consequences of their impairments can lead to major gains. The simplest example is using a diary to compensate for poor episodic memory or an electronic device that helps keep one oriented in time and place. In the domain of remembering to do things in the future (prospective memory) there is compelling evidence that direct electronic reminders delivered via a pager ("take your medication") can dramatically increase goal attainment and independence (Neuropage; Hersh & Treadgold, 1994). In contrast there is minimal evidence that practicing Prospective Memory tasks over increasing intervals has any generalised effect (Sohlberg, White, Evans & Mateer, 1992).

Within executive function rehabilitation there are a number of positive indications from such work. Von Cramon et al. (1991) showed that, if someone is spontaneously unable to generate and follow realistic plans, they can be helped by a step-by-step strategy that can be applied across many types of problems. Positive results have also been reported from the strategies of Goal Management Training, including gains to everyday life (e.g. Spikman et al,, 2011).

The published results from WMT, at least initially, stood in contrast to this general pattern. At last it appeared there was a progressive cognitive training regime that could, with sufficient duration, produce generalised gains. Criticisms and failures to replicate some of these key claims have been amply discussed and the results from the last two chapters are certainly in keeping with this more negative picture.

Given the importance accorded to far transfer in the WMT literature, there is surprisingly little discussion about steps that might facilitate such generalisation; the emphasis is entirely on spontaneous transfer. In a recent study, for example, Redick et al (2011) go to some lengths to prevent their WMT participants from having *any* idea that what they are learning in training tasks could generalise to other measures. There is a certain theoretical purity to this approach but it raises interesting questions of what constitutes 'legitimate' transfer. If people during training were encouraged to think about other situations which required similar processes, or in everyday life situations to think back to how they managed see patterns in spatial span that allowed them to improve performance, could far transfer be supported?

The aims of this study were therefore to combine intensive, adaptive computerised cognitive training with material that could encourage generalisation and personal goal setting. One of the clear advantages of computerised cognitive training is that it can be conducted over practically unlimited durations, in many settings and at little cost. In developing the facilitation materials, the aim was that this too could be delivered via the internet at a time convenient to the trainee, rather than in face-to-face meetings. To this end a series of scripts for short video sequences were developed to be viewed by trainees at specified stages of the overall training programme. The themes, given in more detail below, included thinking about everyday situations relevant to the training materials during training, the value of setting and monitoring goals in training and elsewhere, breaking large problems down into sub-goals (in training and everyday life), managing distraction, and avoiding perfectionist “if I can’t do it perfectly I won’t engage at all” thinking. In conjunction with the training, participants from the healthy older population were also given log books to record their training and to set and monitor personal goals.

Methodology

Participants

The recruitment criteria, as with the previous two studies were age over 55, no significant mental health or medical problems that would influence their participation in the study. There was no upper age limit. Participants were recruited by email and phone and screened. All participants had to have access to a computer with broadband connection for the entire duration of the study in order to take part. The test re-test control group used was the same as that described in chapters five and six. One participant dropped out because he was not familiar enough with using computers to carry out the training on his own at home. The final sample consisted of thirteen participants in the training group and twelve participants in the control group.

The thirteen participants in the training group (2 male) had a mean age of 71.62 (SD 5.87) years. The control group consisted of twelve participants (4 male) with a mean age of 70.36 years (SD 6.14). All participants in both the training and control groups were within the normal range on the screening measures, the GDS and the BDI.

Assessment sessions took place in the testing labs in the CBU and the training took place in participants' own homes. Participants were in contact with the researcher throughout through phone-calls and emails. Any difficulties were dealt with and explained immediately when encountered. Participants could withdraw from the study at any stage without having to give a reason.

The training videos

As previously stated the content of the videos was based on content from previous research in which education and information-giving formed part of the training. There was a combination of teaching, using strategies and homework and participants were also encouraged to keep logs and set goals as was the case in Winocur et al. (2007). Participants were given log books where they could track their progress, set every-day and training goals and there was also a summary of each of the training videos. A copy of the logbook can be seen in the Appendix D. There were nine training videos in total. The first video was to be watched prior to beginning the training. The final video is watched before the last training day and the other videos are placed at intervals of 4-5days throughout the training. The videos vary in length from 5'15" to 14'30" as can be seen below and consisted of a voice-over plus text and graphics to emphasise key points. The topics of the training videos were as follows. An example of one of the scripts and training links can be found in the Appendix E:

1. Intro to training – who is the programme for, why is it important, how do I start (6mins, 02secs)
2. Goal setting – setting SMART (Specific, Measurable, Agreed, Realistic and Time-limited) goals for training and everyday life, how to set goals, examples (5mins, 36secs)
3. Rewarding yourself- challenge perfectionist thinking, reward effort and time as well as progress, expect variability (5mins, 15secs)
4. Stating your intentions – how writing down what you plan on doing or saying it aloud can increase the likelihood of you following through, examples given. (5 mins,15secs)
5. Bullet Points- ways of separating information into relevant chunks, useful in breaking down hard logical problems into easier to manage steps (8mins, 40secs)
6. Why is working memory important? – what is WM, what tasks train WM, what everyday situation require WM, why are we interested in WM (5mins, 47secs)

7. Managing Distraction – internal and external distracters, managing your environment, coping with and managing distractions, deliberately introducing distracters (6mins, 17secs)
8. Memory Strategies – using mental imagery, grouping items (14mins, 30sec)
9. Making the most of training – challenging common misconceptions about memory and aging, remembering strengths, seeking challenges and setting goals for the future, keeping active.(10mins, 21secs)

The videos were accessed by participants through YouTube links. The links were not publically accessible. Screenshots below are given as examples.

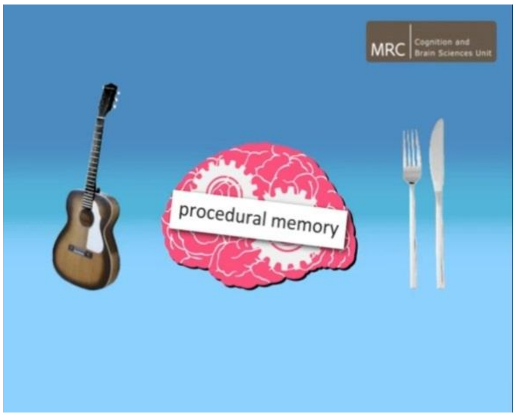


Figure 7.1: Screenshots of training videos

Two narrators, a man and woman, read every second script for variety. The importance of watching the videos and their relevance was explained to participants.

Adaptive Cognitive Training

The on-line adaptive cognitive training package used in the current study combined tasks from the Problem Solving and WMT reported in the previous two chapters. The tasks were Forward Digit Span, Backward Digit Span, Spatial Slider, Grammatical Reasoning, Hampshire Tree Task, Object Reasoning, Odd One Out, Code Words, Spatial Span, Monkey Ladder and Paired Associates.

Outcome Measures

The outcome measures used were mostly the same as those reported in previous chapters. In addition for this study The Beck Depression Inventory (BDI) and the Geriatric Depression Scale (GDS) were applied pre and post-training.

Two assessment measures were excluded from this chapter (The Doors and People and the MET) for reasons of time and practicality and to allow the inclusion of three new measures. The other measures (Digit Span, Cattell Culture Fair, Raven's Standard Progressive Matrices, Games Evaluation Task, Cognitive Failures Questionnaire, Goal Management Questionnaire) were as described in the previous chapters. The three new measures are described below. It is worth noting at this stage that due to practicalities and the use of a common untreated control group over this and the two previous studies, pre-post comparisons on these new measures were uncontrolled and their use was exploratory to examine potential change.

The Hotel Test (Manly, Hawkins, Evans et al., 2002),

This measure of planning and monitoring was a variant of Shallice and Burgess' (1991) 6-Elements Test. During the task participants were asked to imagine that they were working in a hotel and have been asked by a manager to sample 5 tasks 'to get a feel for how long they would take to complete'. The five tasks were then described. These were, sorting a randomly shuffled set of conference name badges into alphabetical order, sorting a coin collection into piles of each denomination and excluding foreign coins, proof-reading a draft pamphlet, compiling individual bills from a common till roll and looking up business telephone numbers in the Yellow Pages directory. Participants were told, accurately, that it would take longer than an hour to complete all of the tasks and that they had only 10 minutes to try some of each of the tasks. They were asked to try and spend as much time on each of the tasks as possible. Comprehension of these main goals was checked prior to the test beginning and instructions repeated if necessary. A covered clock was available for participants to check as often as they wished (the cover allowed the examiner to note when this occurred) and they were asked to remove watches etc. for the test duration. The examiner recorded when participants began and ended each task and clock checks. Subsequently the cumulative time allocated to each task was calculated. The key measures were the number of tasks attempted and the cumulative deviation from the optimal time allocation (for example, if task 1 was completed for 1 minute and 30 seconds, this would represent a deviation of 30 seconds form the optimal 2 minutes. Similarly, if task two was completed for 2 minutes and 30 seconds, this too would represent a 30 second deviation to be added to the total). The tasks looks at a participant's ability to monitor time, switch between tasks and remember what they are required to do (Manly et al., 2002).

The Paced Auditory Serial Addition Task (PASAT; Gronwall, 1977) is a measure of cognitive function and was designed to assess processing speed and flexibility to auditory information. Participants listen to a series of numbers presented at 2 second intervals and each new digit must be added to the previous digit heard. One mark is given for each correct addition and the maximum score is 60.

The Rivermead Behavioural Memory Test-II Story Recall subtest (Wilson, Cockburn & Baddley, 2003)-, as described in chapter four, is a measure of explicit verbal memory and was useful as it has four different versions which lend themselves well to re-assessment.

Procedure

Informed consent was collected and the outcome measures were administered in a 1.5-2 hour session before training. These were repeated within two days of the final training session. Participants in this study were asked to initially train for 5 -10 minutes a day during the first week and then systematically increase training duration to 20-25 minutes per day. By the third week the participants were training at full capacity. This was done in order to familiarise participants with the training and to ease them into the training regime, as was suggested by participant's feedback from the earlier two studies. One session of a full length training sessions consisted of twelve blocks (twelve different training tasks), as was the case with the previous two training packages –AWMT and PS.

The training took place at home at a time that suited participants. Participants were encouraged to complete the 25 days of training and to watch all the videos that were provided. As with the WM and PS training, success at an initially relatively easy level of difficulty would lead to progressively more difficult items being presented.

Results

Compliance

Compliance was high throughout the training. All participants completed the minimum requirement of the training. One participant did not complete the final two days of training due to prior holiday commitments that were planned. One participant carried out two training sessions in one day twice in order to complete the all training days. During these days one session was carried out in the morning and one in the evening. When participants were training at full capacity (in the final three weeks of training) the daily training sessions took on average 29.53minutes to complete.

Improvement during training

Initial levels of performance on the training tasks were compared with final levels of performance training to see if there was a training effect. The last two sessions that each participant completed was obtained and then averaged to give a final score. Paired samples t-tests were used to look for change and all data reported are two-tailed. From Table 7.1 it can be seen that all training tasks showed an upward significant change over time between the beginning and end of training.

Training Task	*t* value	*p* value
Digit Span	-2.42	.030
Spatial Slider	-2.15	.050
Grammatical Reasoning	-7.91	.000
Hampshire Tree Task	-2.41	.030
Object Reasoning	-4.69	.001
Odd One Out	-2.45	.031
Code Words	-4.35	.001
Spatial Span	-6.11	.000
Monkey Ladder	-3.19	.008
Paired Associates	-5.10	.000
Backward Digit Span	-2.68	.020

Table7.1 Practice effects on training tasks

Outcome measures

At baseline there was no significant difference between the two groups on any of the outcome measures. There was, however, an almost significant difference between control and training group performance on the GET with the training group performing less well .055 (two-tailed).

Table 7.2 below summarises the results from both the control and strategy training group on the variables of interest at Time 1 (pre-training) and Time 2 (post-training/re-assessment). As stated above the new variables were not included in the repeated measures ANOVA as results were not available for the control participants. These results are reported separately below.

Outcome Measure	Group	Time 1	Time 2	Main Effect of Time	Group x Time Interaction
		Mean (SD)	Mean (SD)		
Cognitive Failures Questionnaire	Control	31.92 (10.26)	32.83 (12.22)	F(1,23) = 0.01, P = 0.93	F(1,23) = 0.44, P = 0.52
	Strategy	39.15 (14.19)	38.46 (15.50)		
Goal Management Questionnaire	Control	87.67 (51.24)	83.83 (57.44)	F(1,23) = 0.001, P = 0.98	F(1,23) = 0.25, P = 0.62
	Strategy	93.62 (65.75)	97.08 (70.16)		
Cattell	Control	32.17 (3.90)	31.92 (4.78)	F(1,23) = 1.31, P = 0.26	F(1,23) = 0.61, P = 0.44
	Strategy	32.15 (6.54)	30.85 (4.49)		
Raven's Score	Control	49.42 (4.80)	50.25 (4.52)	F(1,23) = 1.65, P = 0.21	F(1,23) = 0.004, P = 0.95
	Strategy	49.00 (9.17)	49.92 (7.26)		
Raven's Time (seconds)	Control	2073.83 (509.03)	1719.58 (414.84)	F(1,23) = 2.86, P = 0.10	F(1,23) =0.14, P = 0.71
	Strategy	2585.46 (833.35)	2359.85 (810.37)		
GET Tasks	Control	7.58 (1.62)	7.50 (1.31)	F(1,23) = 0.20, P = 0.66	F(1,23) = 0.41, P = 0.53
	Strategy	6.69 (2.06)	7.15 (2.03)		
GET Time Dev. (seconds)	Control	151.67 (81.90)	132.79 (52.42)	F(1,23) = 0.20, P = 0.66	F(1,23) = 0.43, P = 0.52
	Strategy	218.65 (83.86)	222.31 (106.51)		

HADS (Depression)	Control	2.58 (1.51)	2.33 (1.44)	F(1,23) = 0.75, P = 0.40	F(1,23) = 1.81, P = 0.19
	Strategy	2.15 (1.77)	3.31 (4.07)		
HADS (Anxiety)	Control	3.17 (2.66)	3.17 (2.37)	F(1,23) = 0.59, P = 0. 45	F(1,23) = 0.59, P = 0.45
	Strategy	3.69 (3.12)	4.46 (5.41)		
Digit Span	Control	5.28 (0.56)	6.32 (0.90)	F(1,23) =3.54, P = 0.08	F(1,23) = 0.002, P = 0.97
	Strategy	5.61 (0.74)	6.60 (1.21)		
Geriatric Depression Scale	Control	4.00 (3.69)	3.92 (3.50)	F(1,23) = 0.57, P = 0.46	F(1,23) = 0.72, P = 0.41
	Strategy	4.23 (3.68)	5.69 (6.56)		
Beck's Depression Inventory	Control	6.00 (4.38)	5.82 (5.76)	F(1,22) = 0.30, P = 0.59	F(1,22) = 0.49, P =0.49
	Strategy	5.31 (4.82)	6.77 (8.80)		

Table 7.2: Repeated measures ANOVA between control and strategy participants

As can be seen in table 7.2 above, as with chapters five and six, no significant improvement was seen in performance in this group following strategy training when compared with the control group.

Mean and standard deviation scores for the variables that were not used on the control group – the Hotel task, PASAT and the RBMT II Story Recall- can be seen in table 7.3 below. Improvements on re-assessment can be seen on some of the variables, especially the PASAT and Story recall measure. The working memory measure of PASAT t(11)= -3.588, p=.002 (one-tailed) reached significance following training. The story recall subtest of the RBMT II, a measure of explicit verbal memory, also improved significantly after immediate recall t(12)=-1.968, p=.037 (one-tailed) and was approaching significance when retold after a delay t(12) = 1.671, p= .06 (one-tailed). In the Hotel Task there were no other significant differences on re-assessment but the number of subtasks completed was nearing significance (p=.069).

Outcome Measure	Time 1 Mean (SD)	Time 2 Mean (SD)
Hotel Task - Number of Subtasks	4.20 (1.55)	4.70 (1.16)
Hotel Task - Number of Clock Checks	6.2 (4.69)	5.10 (2.47)
Hotel Task - Time Deviation (seconds)	294.60 (245.21)	316.20 (278.39)
PASAT	28.25 (9.60)	33.42 (7.97)
RBMT II Story Recall Immediate	7.19 (2.16)	8 (2.84)
RBMT II Story Recall Delayed	6.85 (2.38)	7.54 (3.09)

Table 7.3 Strategy training variables – mean and standard deviations

Power and Effect Size

Effect sizes were calculated as in Chapter 5 and Chapter 6 and were small in all variables reported CFQ (0.14), GMQ (0.10, Cattell (0.16), Raven's Time (0.08), Raven's Score (0.01), GET Tasks (0.13), GET Time (0.14), HADS Depression (0.27), HADS Anxiety (0.16), Digit Span (0.01), Geriatric Depression Scale (0.17) and BDI (0.15).

A post-hoc power analysis was carried out on the data here to see if the non-significant results seen were as a result of a lack of statistical power. Again the analysis was as in chapter five and six and used GPower computer programme (Faul et al., 2009).Power for each of the outcome measures was as follows: CFQ=.17; GMQ = .12; Cattell = .20; Raven's Score = .06; Raven's Time = .10; GET Tasks = .16; GET Time Deviation = .17; HADS (Depression) = .39; HADS (Anxiety) = .20; Digit Span = .06; GDS = .13.

As can be seen, statistical power was low which suggests that there may not have been sufficient power in the study to detect a significant effect, given our sample size.

Discussion

Participants in the training group in this chapter completed 25 days of training and watched nine videos that were related to the training and the application of the training to everyday life. Following this period of training improvements, when compared to the control group, were not evident in the training group. Disappointingly, even when participants were trained in strategies to help improve performance on cognitive tasks and apply this training to everyday life they showed no disproportionate gains. These findings are in line with those already reported in chapters five and six but it was thought that with more of a variety of tasks available and the clear application of the training to daily life that participants would benefit more from this holistic training.

Improvements in measures of memory were seen on the PASAT and story-recall subtest of the RBMT II. One of the limitations that is evident from this chapter is the unavailability of the control data for these new measures. Participants improved significantly on re-assessment but due to missing data it is unclear if the same improvement would be seen in the control group through practice of the items. It was hoped to form a new control group for this chapter but as the panel of volunteers was near exhausted for this age group there were not sufficient numbers available. It was decided not to change the recruitment method for this chapter when there was to be no change of control group so that comparisons could easily be made, if necessary between the training conditions. Participants from the control group were contacted and encouraged to return for two assessment sessions in which they were measured on the new outcome measures. As participants were at this stage aware that they were in the control condition and it had been some time since they first participated they were reluctant to return. Only four of the original twelve group members agreed to it and so it was decided that this would not be sufficient. It could be that participant had become de-motivated and disinterested in the study because they had to wait to be offered the most superior training package. This de-motivation in later group was a potential difficulty for Ball et al. (2002). It cannot be said as a result if the training benefits are due to training or just to re-assessment but there may be some influence of training. The RBMT II has different versions, which were used to reduce the likelihood of improvement being due to a practice effect but this possibility cannot be ruled out fully.

Unlike a lot of research described previously, this chapter reports on a training programme that is entirely computer-based. The training and the teaching/education elements of the programme are online. This was decided upon due to resources that are available to clinicians and researchers. If a computer-based programme was seen to be effective in improving everyday functioning the number of contact hours that patients had with clinicians could be used on other aspects of functioning. A programme such as this may be effective with patients who wish to take part in a cognitive rehabilitation programme but who may not wish to learn or may not be able to effectively learn in a group situation. This allows patients to train in their own time. This does eliminate the social and psychological aspects of rehabilitation that comes with being a member of a group. It also is more difficult to monitor than traditional therapy – there is no way of knowing if the patient did the training themselves or if they allowed their children to do it. There is also no way of knowing if they watched the videos or if they pressed the play button and walked away. It is hoped that if participants are interested in recovering that they will be motivated to participate in therapy but this is not necessarily always the case. On inspection of the number of "hits" the strategy training videos got on YouTube it is clear that not every participant watched every video. It was suggested by one participant that perhaps our real-life examples and some of the training material and strategies were too simplistic and common-sense-like for our high functioning group.

An element that was not included in this training package but that might be worth including in future studies is exercise, as it was shown when used in conjunction with a cognitive rehabilitation programme to be more beneficial than the cognitive programme alone (SIMA; Oswald et al., 1996). Participants in this study were encouraged to participate in exercise in the videos but it was not an essential element, only a suggestion. If this was an essential compulsory element of the training it may have improved the results seen by cognitive training alone.

It was disappointing to see that the measures that were designed to show improvements in everyday life – the CFQ, GMQ and Hotel Task -did not show any differences following training. Improvements in psychosocial well-being were not also not evident. It would be interesting to see if any improvements would be seen at follow-up in this group. The final video that was shown to participants gave tips on how to continue their cognitive training, encouraged participants to keep active and to get involved in activities. It would be worth noting if participants decided to take on any activities etc. that they were not previously a part of and if at follow-up this improved or maintained their scores on the variables used. As participants were encouraged to set goals during the training programme and after the final training session it would have been a good everyday and clinical measure to have Goal Planning as a way of evaluating the training and not by neuropsychological tests alone, as suggested by Wilson (2002). Although participants were encouraged to set goals for themselves throughout training it is unclear what number of participants set goals and worked towards achieving these. Participants kept track of their goals in their logbooks and were permitted to keep their logbooks following training. In order for participants to set personal goals that they may not have wished to share it was decided not to collect logbooks and so no accurate record of this is available.

The results of this chapter echo those of Owen et al. (2010) and Redick et al. (2012) as well as the previous two chapters where no gains can be reported from an intensive period of training on a mixture of cognitive tasks – focusing especially on working memory and problem solving. The training in this chapter was of a slightly lower intensity to the previous two chapters as the training began at 5mins per day and increased weekly but yet ran for the same amount of time as in the previous chapters. It is difficult to say for some of the measures if the improvements seen are as a result of the training or that just by being in a study that there is a higher chance of improvements being seen. From the three studies of working memory, problem solving and strategy training no clear conclusions can be made about the benefits of cognitive training to older adults. One cannot endorse their use as a way of enhancing cognition in this population group, although it can be said that further investigation of the topic might be worthwhile.

Chapter Eight:

General Discussion

Summary of findings

This thesis explored a number of different areas with the broad aim of adding to the research and knowledge base of assessment and training of cognition in older adults. The first of the two main aims was to develop and validate outcome measures that could be used with this same population group. The second main aim was to investigate computerised cognitive training in older adults. The focus of the training was on both working memory and fluid intelligence which were trained separately and then as a combined training package. In keeping with the applied clinical nature of the research aims, it was of particular interest to examine training-related gains in cognitive functioning, and the generalisation of any such gains to everyday life.

The first empirical chapter concerned the development of a new computerised measure of organisational abilities, the Games Evaluation Task (GET). This was based on existing measures such as the 6-Elements and Hotel test but had the putative advantages of automated scoring, minimal materials and, in that only finger actions were required, compatibility with fMRI and other functional imaging methods. The first of three studies reported in this chapter focused on whether the sensitivity of the measure to goal neglect could be enhanced by the addition of accurate but irrelevant instructions. A further manipulation was to the order such that irrelevant instructions were heard first or last. In fact, instruction condition exerted no detectable between-group effect. The second and third studies examined correlations between GET and other measures. Organisation of performance on the GET was significantly negatively correlated with a well-established measure of cognitive lapse frequency in everyday life, the Cognitive Failures Questionnaire and positively correlated with the Cattell Culture fair. It was however not significantly related to performance on another measure of Goal Neglect. This is interesting because a study (Rocas et al. 2010) found that Cattell performance accounted for much of the variance in traditional 'frontal' tests but not the Hotel Test, which is conceptually similar to the GET. Nevertheless, the results suggest that the GET may have ecological validity as a brief and practical measure of executive function.

Chapter three examined the psychometric properties of an adapted version of an existing measure, the Multiple Errands Task. The focus was on whether performance could be practically scored from a 1st-person perspective video recording. The results highlighted a number of issues critical to timing the tests. Inter-rater reliability, between a live rater and two separate raters using only the video and audio footage, was adequate. One value of this approach is that, in the context of intervention studies etc., blind and independent assessment can be performed off-line.

In chapter four, a case-study is reported of a man with difficulties in of time perception and working memory post-stroke. In the first of a number of subsequent studies investigating the use of cognitive training, a programme was implemented that, following a period of problem solving training, focused on intensive training on working memory tasks. Improvements in performance were seen on a measure of time perception and also on the newly developed organisational measure, the GET following AWM training, with no consistent changes after PS training. Whilst interpretation is limited in single cases the results were certainly consistent with a generalised benefit from AWMT.

In chapter five a group of older adults undertook an intensive working memory training programme that lasted five weeks and consisted of twenty-five sessions. The training was similar to that in chapter four but a no-contact control group was included, rather than a within-subject PS training. The results were inconclusive with no significant improvements seen in any of the measures used and consequently no indication of generalisation of training, even when practice effects were seen on training tasks. These results may have been influenced by a number of factors for example the sample size of the lack of variety in the training programme itself. It was somewhat surprising that although positive results were seen using the same training programme in chapter four, chapter five failed to give the same indication. However, given the differences in participants, the experimental design, and the subsequent analysis, it is difficult to make direct comparisons.

In chapter six the theme of cognitive training was further developed. Here there was an attempt to distil the sort of reasoning/novel problem solving required for high g tasks into a training regime. As with the other studies compliance was high and strong gains were observed on the trained materials. However, no evidence of generalisation to the key outcome measures including of fluid IQ was observed relative to the untrained control group.

Finally, chapter seven saw the two training packages – working memory and problem solving – combined. This training programme therefore had a broader cognitive focus and a greater variety of tasks. Another addition to this training package was the inclusion of strategy training in conjunction with the computerised aspect. In this, participants watched a number of videos throughout the five weeks of training that focused on a number of different themes from relating to the cognitive training to everyday life to teaching participants specific strategies that could help improve their performance on some of the training tasks. These included the value of setting goals, monitoring progress and rewarding oneself for goal attainment, thinking about the relevance of training to everyday activities (and vice versa), avoiding perfectionist thinking, and managing distraction. Disappointingly, despite the changes made to the training package no disproportionate gains were seen in any of the outcome measures relative to the control group.

In summary, from the two broad aims of this thesis, in terms of assessment and training of cognition, mixed results have been found. The two assessments that have been under investigation – the MET and the GET – show the potential to be used as outcome measures but are still at an early stage of development. In terms of the cognitive training package the overall results are disappointing. None of the three training packages succeeded in showing any significant difference in older adult's daily lives following intensive training. The results with this population failed in almost all outcome measures, including measures of near and far transfer. There is still a potential for this training package to be of use as positive results were seen in the case-study of a man with an acquired brain injury. The training package might be more appropriate for use with a group that was not at the high functioning level of the group here. This can be no recommendation that this cognitive training package should be used with healthy older adults.

Theoretical implications

Having available reliable measures that can be scored independently, either by a separate blind rater or a computer programme is an important factor when considering the strength of a study. If outcome measures can be free from the influence of bias and are themselves valid, reliable measures then it would be beneficial when assessing the results of a given piece of research, in this case the effectiveness of cognitive training. If the outcome measures used are not of a high standard then the results are often questioned. The positive results that are indicated in the video scoring of the MET and in the use of the GET as a quick measure of organisation give initial support for their use in these circumstances.

Recent imaging studies (e.g. Takeuchi et al., 2010; Olesen, Westerberg & Klingberg, 2003) have shown plasticity in the brain after a period of training focused on working memory. Takeuchi and colleagues (2010) showed that following two months of training over 25minutes a day, changes consistent with increased myelination was found in frontoparietal regions using voxel-based analysis, while in an fMRI study Olesen et al. (2003) found that after five weeks of intensive working memory training increased activation was found in the cortical areas that are associated with working memory – the prefrontal and parietal regions. This suggests that not only does working memory have the potential to be trained, as was discussed in chapter one, it can also have an influence at a cortical level. The results from this thesis are in contrast to this and suggest that training-related benefits are hard to harness. The indication from this series of studies is that this training did not elicit change in working memory capacity. In terms of problem-solving training, this thesis either failed in isolating the particular novel problems that are found to be the strongest loading on *g,* failed to measure generalised improvement successfully or contradicts the potential for this to even be possible. Redick et al. (2012) questioned the theoretical grounding of working memory training. They point out that even when positive transfer is seen there is no reporting on the mechanisms that are responsible for this transfer. If transfer can be seen then why is it so difficult to identify the underlying reasons for this? The lack of consistent findings, consistent transfer and consistent reporting and methodology may well be the basis for some of these reasons. Dismantling designs and conducting mediation analysis could shed some light on this in future studies.

The findings are not, however, fully unexpected despite the fact that recent theory might give that impression. As previously noted, two well-controlled, large-scale studies (Owen et al., 2010; Redick et al., 2012) also did not show support for a cognitive training regime. Their lack of any meaningful transfer of gains from intensive cognitive training is in line with what has been reported here. What further emphasises this was the fact that many of the training tasks used in the study by Owen and colleagues were used in this thesis also. As mentioned by Shipstead et al. (2012) and Redick et al. (2012), there may be more studies out there that support this lack of transfer that never see the light of day due to positive publication bias in authors and journals. As stated in chapter one, many of the studies that report positive finding have a large number of methodological flaws, from small sample sizes, to collapsing data when the reasons are not particularly strong (see chapter one), to lack of control group and so even the positive results may need to be viewed prudently. It remains possible that cognitive training may be beneficial in improving performance on certain measures, in certain groups but this thesis provides little or no evidence for use of either of the three cognitive training programmes reported on a healthy high-functioning older adult group.

Clinical implications

Assessment

The importance of having accurate and reliable outcome measures in both research and clinical settings cannot be over stated. Measures need to discriminate between groups, be reliable, be validated against existing measures and be reasonably easy to administer. In research claims are made based on results from outcome measures, as was the case in this thesis. In clinical settings appropriateness of an outcome measure is important for setting goals and having an indication of rehabilitation potential of a patient. The GET showed itself to have at the very least the potential to be new, quick measure which would give an indication of a person's ability not only to organise themselves and to multitask but also suggest if a patient might be experiencing cognitive slips in their everyday lives. The measure, no doubt needs further testing before it is used as a clinical test of executive difficulties but it has been shown to have promise in this field.

Instruments need to produce the same results not only at repeated intervals but there must also be stability across different raters (Corr & Siddons, 2005). The video MET study showed that when three different raters, in two different circumstances, rated participant performance there was consistency in scoring in the tool. As already discussed using a video rater in a research setting could eliminate potential bias, and in a clinical setting a similar result could be seen by having an objective measure of a patient’s functioning. The task, although time consuming, loads on a variety of different cognitive and organisational factors and has the added advantage of being set in an everyday situation. The video does not add any further difficulty to the task on the part of the participant and does not necessitate changing the measure in any way. It is encouraging to see that the video ratings for the MET are as reliable as in live ratings but as with the GET the task demands further rigorous psychometric testing at this early stage.

Treatment/rehabilitation

From a treatment and rehabilitation point of view the results, as stated, are not so encouraging. Training needs to show that it will extend to improvements in daily life in order for it to be of any benefit to patients in terms of rehabilitation. Not only did this not happen in the series of studies reported here, there was also a lack of consistent far transfer and everyday transfer reported in many other studies to date, as previously mentioned. The idea that cognitive training would be a useful, reasonably enjoyable, and relatively quick (5 weeks) method of slowing down is a tempting one and from that viewpoint it is easy to see the reasons why there is a recent surge in interest and research in the area. Although results here have not been encouraging, it may be that this training is not suitable for such a high functioning group but that the training works well with a clinical population who might be functioning less well. As the participants appeared to be at a high level of cognitive functioning there was little room for improvement on tasks and outcome measures used. In a group of participants functioning less well cognitively there may be more potential for change following training and the training and/or strategies may be more beneficial.

The two main positives findings that have come from this training data are in relation to the results from the single-case experimental design and the compliance and enjoyment of the training for the participants who completed it. Positive results were seen in chapter four in areas that were of difficulty to the man under study. His performance improved in time perception, organisation and working memory, compared to PS training. Replication of this with a series of single case experimental designs and with a larger sample would give more information about the possible rehabilitation potential of this training package with a clinical population. Until this time there is not enough evidence on which to base an endorsement of the training.

Despite the absence of training effects, compliance with the three types of training was high. Not only did participants finish the entire training schedule the majority of the participants who filled in anonymous feedback forms reported having enjoyed the training. Of the thirty-five participants who completed the training (feedback on the training was not requested from control participants), twenty-one participants completed a brief questionnaire that was sent to them requesting feedback on their experiences of the training after they had finished the final assessment session (see Appendix F). Only two participants said that they would not recommend the training to others – one from the working memory group and one from the strategy training group. Below are some quotes taken from the feedback forms that show the positive and negative experiences of those who took part.

> *“I found the whole exercise fascinating as well as interesting…All in all a marvellous experience” P1*

> *“It went on too long…it was very repetitive-the tests didn’t sustain my interest very much.” P2*

Participants were asked to rate on a Likert scale from 1 to 10 (not at all – a lot) how much they enjoyed the training. The mean score for enjoyment was 7.67 (SD 1.93) which is encouraging given the relative repetitive nature of the training.

> *“Overall it was a fun experience and I’m glad I participated” P3*

> *“The training was excellent, I am impressed with that”P4*

When asked how much they thought they benefited from the training participants from 1 (not at all) to 4 (a lot) participants gave a mean score of 2.62 (SD 0.67).

> *“I think my memory was faster after training” P5*

“Although toward the end of the exercise I found the tasks quite boring, I certainly felt that I had improved memory for some of the numerical problems, though whether this was actually the case only you will know!”P6

“I felt they were mental exercises which had little or no influence for everyday activities. So the skills acquired by doing the exercises were wasted on me.” P2

It may have been that the tasks, as mentioned, were not challenging enough for the participants. This was shown in the feedback form when participants were asked how difficult they found the training from 1(very difficult) to 5 (very easy) and the mean score (3.1, SD 0.77) indicated that the training was only moderately difficult for participants. One participant mentioned that the training may have benefitted from being longer in duration:

“I got very frustrated that I didn’t seem to get any better at computer tasks. I suspect that it takes more than five weeks to make a change for a ‘well’ person”P7

When the group that received the strategy training was asked about how useful they found the training videos on a scale of 1 (not useful) to 10 (very useful) the results were also moderate (mean 5.33, SD 1.87) with some participants finding the videos more useful than others.

“The training made me think more deeply about all of the above (confidence, speed of thinking, concentration, memory organisation, everyday life), to realise that I could work on them … I think I have made some of the strategies part of my everyday life. Definitely a worthwhile experience for me.” P8

It is clear that one of the benefits that can be salvaged from the negative results of the three studies is that there is potential there, if in the future there is further development on the training, for it to be accepted well and for participants to be compliant in the training regime.

Limitations and implications for future research

A clear difficulty in interpreting the results of all of the studies reported in this thesis is the fact that the samples in each of the studies were relatively small and the group that made up this sample were relatively homogenous. As can be seen from analysis in chapters five-seven there was not adequate power in most cases to make any definite conclusions. The effect sizes reported in chapters 5 and 6 with the Raven's Time, GET Time Deviation and Doors and People indicate some potential benefits of AWMT and PS training, which could be further explored in a larger group. As older adults were the focus of the research the lack of variety in age was not an issue but many of the participants were at a very high level of functioning and may not be representative of older people in general. They may, however, be relatively representative of an older Cambridge population and in that case any generalisation might be limited to just this specific group. Again, further testing would be required for this to be confidently stated.

There is much discussion about when corrections for multiple comparisons should be employed. When a very large number of correlations are examined without specific hypotheses, for example, such correction is clearly warranted. However, in this case there were a relatively limited number of comparisons with clear hypotheses (e.g. that GET performance would be positively related to fluid intelligence and negatively related to everyday cognitive errors). Correction in that context appears too conservative and risks missing important issues.

In the training studies, chapters five to seven, a repeated measures design was used but it also has its limitations. The likelihood of having sampling bias and inflated amounts of Type 1 errors might be increased in using this model if there are missing values, which although not a big problem was the case for a number of measures (Gueorguieva, 2004; Muller, 1989). The repeated measures ANOVA is very vulnerable to this.

As discussed by Shipstead et al (2010; 2012) and Redick et al. (2012) there are issues surrounding the use of a no-contact control group in relation to motivation, "Hawthorn effect" (French, 1953) etc. and the likelihood of getting positive results. As it turned out, this is not an issue in interpreting negative training results presented here nor for S.M.'s results, where PS training was used as a comparison. However it is worth noting that non-active control groups are often the most plausible in terms of actual clinical practice – many patients do not face a choice between different rehabilitation programmes and examining general as well as specific effects of intervention gives a less conservative guide to likely gains. In practice, multiple types of control group provide the best overall picture.

With respect to the tasks that were used in the training studies there was an obvious lack of variability and variety in what the participants trained on:

> *"Would have liked a bit more variety" P9*

This was especially an issue in the problem solving group. In previous studies (e.g. Redick et al, 2012) that use dual n-back tasks participants use the same task throughout their training. In comparison with this, the current studies actually included quite a varied array of tasks. Indeed, many of the participants found the amount of variety satisfactory.

> "*The only change would be the test involving the 'tree'. That is a horrid one, not for the faint hearted*."*P10*

In the problem solving group, however, the theoretical reasoning behind the training was that the tasks had to be novel so that participants are constantly being challenged to problem solve on new tasks. The idea was that participants would not be able to come up with strategies to tackle problems with practice. Certain task, e.g. The Hampshire Tree Task, made it more difficult to create strategies but there was still an element of familiarity and practice that came with training on a limited number of stimuli over 25 sessions.

As discussed in chapter one, many studies that report gains following training used dual n-back tasks. These tasks were not used here. It was thought for this group of older adults that the intensity and difficulty level of the WM tasks was at an adequate level to challenge participants. On reflection having seen the results and the high functioning level of the group it might have been appropriate to use dual n-back training with this group. Some individual participants reported gains from training nonetheless.

> "*I can now reproduce a 9 digit number which I would have previously divided into two to remember".P11*

Goal setting, as was addressed to a certain extent in the strategy training study, should be included in future research and in a clinical setting as both a measure of meaningful improvement and as a way of setting targets for rehabilitation (Wilson, 2003). Goal setting would work as a very good compliment to other functional or cognitive assessments and gives an easy measure of everyday performance and improvement. Neuropsychological assessments give an estimate of the likelihood of everyday problems but they are not always informative regarding a person's daily functioning. It is essential to have available naturalistic measures as well as setting meaningful goals to see if these problems identified on measures actually impact a person's functioning or not (Wilson, 2003). Shallice and Burgess' (1991) study lends further evidence to this.

An area that was not full addressed in these studies was that of quality of life (QoL) and older adults. Quality of life and cognitive functioning might be worth addressing if treating patients and so would be worth including in future research. Mc Dougall and House (2012) found that participants who reported better QoL benefited more from cognitive training. In the training studies reported here basic mood disorders were screened for such as anxiety and depression but QoL was not specifically examined. Loneliness and social isolation according to Mc Dougall & House (2012) should be considerations both in terms of treatment and research and would be something that should be addressed in future studies.

Whether future research proves or disproves the effectiveness of cognitive training it is clear that many different methodological issues need to be addressed. Many issues mentioned above should be considered as well as using a more varied, representative sample of the population. It is clear that as mixed and contradictory results continue to be found in terms of differing cognitive training regimes a full identification of the methodological issues needs to be identified and targeted fully. The gains appear to be more consistently strong with children (Melby-Lervag & Hulme, 2012). Redick et al (2012) make a good start at this but further evidence is required in order to make any certain conclusions, and as with other studies the sample used here is a young educated group. Following this series of studies it is clear that this area requires more research in order for any clear deductions to be made from findings.

Conclusions

The cognitive training described in this thesis failed to show any transfer to the measures of mood, fluid intelligence, working memory, everyday memory and cognitive slips, and multitasking that was anticipated. The absence of transfer effects in the light of practice effects seen on individual items is especially interesting as it shows that the participants are improving in performance but this is solely restricted to the items they trained on over the 25 sessions. Participants were compliant with the training demands and claim to have enjoyed the experience of taking part in the investigation. The positive results seen in the single case experimental design lend some hope to the potential of this training regime being used with a clinical population. There is a need for further well-controlled studies in order to add to the relatively small base of cognitive interventions that have a sound evidence base. Taken together the results from this thesis support the provisional use of the measures that were under investigation and further investigation of the training reported in a clinical population and with older adults is justified.

References

Ackerman, P.L., Kanfer, R. & Calderwood, C. (2010). Use it or lose it? Wii brain exercise practice and reading for domain knowledge. *Psychology and Aging*, 25(4), 753-766.

Alderman, N., Burgess, P.W., Knight, C. & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*,9, 31-44.

Alexopoulos, G.A., Abrams, R.C., Young, R.C. & Shamoian, C.A. (1988) Cornell scale for depression in dementia. *Biological Psychiatry*, 23:271-284.

Alloway, T.P. & Alloway R.G. (2009). The efficacy of working memory training in improving crystallized intelligence. *Nature Precedings*. HDL:10101/npre.2009.3697.1

Alloway, T.P. & Alloway, R.G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *J Exp Child Psychol.* 106(1):20-9.

Atkinson, R. C. & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, Aug, 82-90.

Awh, E. & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci*., 5, 119-126.

Baddeley, A.D. (1992). Working memory. *Science*, 255, 556-559.

Baddeley, A. D. & Hitch, G. J. (1974). Working Memory. In G.A.Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (pp. 47-90). New York: Academic Press.

Baddeley, A. D. (1988). But what the hell is it for? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1: Memory in everyday life* (pp. 3–18). Chichester, England: Wiley.

Baddeley, A. D., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology, 38A,* 603-618.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science* 4 (11), 417–423.

Baddeley, A.D., & Hitch, G. (1974). *Working memory.* In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A.D., Emslie, H. & Nimmo-Smith, I. (1994). Doors and People: A Test of Visual and Verbal Recall and Recognition. Flempton, Bury St. Edmunds: Thames Valley Test Company.

Ball, K., Berch, D.B., Helmers, K.F., Jobe, J.B., Leveck, M.D., Marsiske, M., Morris, J.N., Rebok, G.W., Smith, D.M., Tennstedt, S.L., Unverzagt, F.W.& Willis, S.L. (2002). Effects of cognitive training interventions with older adults. A randomised controlled trial. *JAMA*, 288(18), 2271-2281.

Basak, C., Boot, W.R., Voss, M.W. & Kramer, A.F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychology and Aging*, 23(4), 765-777.

Beck, A. T., Steer, R.A., & Garbin, M.G. (1988) Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review,* 8(1), 77-100.

Beck, A.T., Ward, C.& Mendelson, M. (1961). Beck Depression Inventory (BDI). *Arch Gen Psychiatry,* 4 (6): 561–571.

Beck, S.J., Hanson, C.A., Puffenberger, S.S., Benninger, K.L. & Benninger, W.B. (2010). A controlled trial of working memory training for children and adolescents with ADHD. *Journal of Clinical Child and Adolescent Psychology*, 39(6), 825-836.

Bedard, A. C., Martinussen, R., Ickowicz, A., & Tannock, R. (2004). Methylphenidate improves visual-spatial memory in children with attention-deficit/hyperactivity disorder. *J Am.Acad.Child Adolesc.Psychiatry*, 43, 260-268.

Bergman Nutley, S., Soderqvist, S., Bryde, S., Thorell, L.B., Humphreys, K. & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: a controlled, randomised study. *Developmental Science*, 14(3), 591-601.

Bjelland, I., Dahl, A.A., Haug, T.T. & Neckelmann D. (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. 52(2):69-77.

Bor, D., Duncan, J., & Owen, A. M. (2001). The role of spatial configuration in tests of working memory explored with functional neuroimaging. *Scandinavian Journal of Psychology, 42*, 217-224.

Boyd, T.M. & Sautter, S.W. (1993). Route finding: a measure of everyday executive functioning in the head-injured adult. *Applied Cognitive Psychology*, 7, 171-181.

Brehmer, Y., Westerberg & Bäckman, L. (2012). Working-memory training in younger and older adults: training gains, transfer and maintenance. *Frontiers in Human Neuroscience*, 6(63), 1-7, DOI:10.3399/fnhum.2012.00063.

Broadbent, D. B. (1971). *Decision and Stress*. London: Academic Press.

Broadbent, D. B., Cooper, P. F., Fitzgerald, P. & Parkes, K.R.. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology* 21,1-16.

Brossart, D. F., Parker, R. I., Olson, E. A. & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification*, 30, 531-563.

Bugos, J.A., Perlstein, W.M., McCrae, C.S., Brophy, T.S. & Bedenbaugh, P.H. (2007). Individualised piano instruction enhances executive functioning and working memory in older adults. *Aging & Mental Health*, 11(4), 464-471.

Burgess P. (1996). The Dysexecutive Questionnaire. In B. A. Wilson, N. Alderman, P. W. Burgess, H. Emsley & J. Evans (Eds.), *The Behavioural Assessment of the Dysexecutive Syndrome*. Bury St Edmunds: Thames Valley Test Company.

Burgess, P. W. (1997). Theory and Methodology in Executive Function Research. In P. Rabbitt (Ed.), *Methodology of Frontal and Executive Function* (pp. 81-111). Hove: Psychology Press.

Burgess, P. W., Alderman, N., Evans, J., Emslie, H. & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, 4, 547-558.

Burgess, P.W., Alderman, N., Forbes, C., Costello, A., Coates, L, M-A., Dawson, D.R., Anderson, N.D., Gilbert, S.J., Dumontheil, I. & Channon, S. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*,12, 194-209.

Burgess, P.W., Veitch, E., de Lacy, A. & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38, 848-863.

Buschkuehl, M.,Jaeggi, S.M., Hutchison,S., Perrig-Chiello, P., Dapp, C., Muller, M., Breil,F., Hoppeler, H. & Perrig, W.J. (2008). Impact of working memory training on memory performance in old-old adults. *Psychol.Aging,* 23, 743–753.

Castaneda, A., Tuulio-Henriksson, A., Marttunen, M., Lönnqvist, J.& Suvisaari, J. (2008). A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *J. Affect. Disord.* 106,1–27.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1–22.

Cattell, R.B. & Cattell, A.K.S. (1960).*Culture Fair Intelligence Test (CFIT), Scale 2, Form A*.Champaign, IL: Institute for Personality and Ability Testing.

Cattell, R.B. & Cattell, A.K.S. (1973). *Measuring Intelligence with the Culture Fair Tests*. Champaign, IL: The Institute for Personality and Ability Testing.

Chase, W.G., Lyon, D.R. & Ericsson, K.A. (1981). Individual differences in memory span. In M.P. Friedman, J.P. Das & N. O' Connor (Eds.), *Intelligence and learning*. New York: Plenum.

Chevignard, M.P., Taillefer, C., Picq, C., Poncet, F., Noulhiane & Pradat-Diehl, P. (2008). Ecological assessment of the dysexecutive syndrome using execution of a shopping task. *Neuropsychological Rehabilitation*, 18(4), 461-485.

Clark, C. R., Paul, R. H., Williams, L. M., Arns, M., Fallahpour, K., Handmer, C. & Gordon, E. (2006). Standardized assessment of cognitive functioning during development and aging using an automated touchscreen battery. *Arch.Clin.Neuropsychol*., 21, 449-467.

Clark, F., Azen, S.P.; Zemke, R., Jackson, J., Carlson, M., Mandel, D., Hay, J., Josephson, K., Cherry, B., Hessel, C., Palmer, J.& Lipson, L. (1997). Occupational Therapy for Independent-Living Older Adults. A Randomized Controlled Trial. *JAMA*. 278(16):1321-1326.

Cogmed Working Memory Training. Available:http://www.cogmed.com/cogmed-united-kingdom-uk-republic-ireland-roi [28/09/2011]

Cohen, G. & Conway, M.A. (2008). *Memory in the real world* (3rd Ed.) Hove, England: Psychology Press.

Cohen, J. (1994). The earth is round (*P*<0.05). *American Psychologist*, 49, 997–1003.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Academic Press, New York

Cohen, J. D., Pearstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J. & Smith, E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386, 604-608.

Comblain, A. (1994). Working memory in Down's syndrome: Training the rehearsal strategy. *Down's Syndrome: Research and Practice*, 2 (3), 123-126.

Conway, A. R., Kane, M. J. & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends Cogn Sci*., 7, 547-552.

Corr, S. & Siddons, L. (2005). An introduction to the selection of outcome measures. *British Journal of Occupational Therapy*. 68(5), 202-206.

Courtney, S. M., Ungerleider, L. G., Keil, K. & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, 386, 608-611.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-114.

Crépeau, F. & Scherzer, P. 1993. Predictors and indicators of work status after traumatic brain injury: A meta-analysis*. Neuropsychological Rehabilitation*, 3(1), 5-35.

Curtis, C. E. & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci*., 7, 415-423.

Curtis, C. E., Rao, V. Y. & D'Esposito, M. (2004). Maintenance of spatial and motor codes during oculomotor delayed response tasks. *Journal of Neuroscience*, 24, 3944-3952.

Dahlin, E., Stigsdotter Neely, A., Larsson, A., Bäckman, L. & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, 320, 1510-1512.

Dahlin, K.I.E. (2011). Effects of working memory training on reading in children with special needs. *Read Writ*, 24, 479-491.

Dallal, G.E. (2008) Randomization.com. Available: http://www.randomization.com [Accessed: 5/10/'09]

Dawkins, N., Cloherty, M.E., Gracey, F.& Evans, J.J. (2006). The factor structure of the Hospital Anxiety and Depression Scale in acquired brain injury. *Brain Injury,* 20(12):1235-9.

Debruyne H., Van Buggenhout M., Le Bastard N., Aries M., Audenaert K., De Deyn P.P.& Engelborghs S. (2009). Is the geriatric depression scale a reliable screening tool for depressive symptoms in elderly patients with cognitive impairment? *International Journal of Geriatric Psychiatry* 24(6):556-62.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 13494-13499.

Duncan J., Parr A., Woolgar A., Thompson R., Bright P., Cox S., Bishop S. & Nimmo-Smith I. (2008) Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology*: General, 137(1), 131-148.

Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 3, 271-290.

Duncan, J. (2006). EPS mid-career award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology*, 59, 2-27.

Duncan, J. & Owen, A.M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences* 23(10): 475-483.

Duncan, J. D. (1995). Attention, intelligence and the frontal lobes. *The Cognitive Neurosciences*. M. S. Gazzaniga. Cambridge, MIT Press**:** 721-733.

Duncan, J., Burgess, P. & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions*. Neuropsychologia*, 33(3), 261-268.

Duncan, J., Emslie, H., Williams, P., Johnson, R. & Freer, C. (1996). Intelligence and the frontal lobe: The organisation of goal-directed behaviour*. Cognitive Psychology*, 30, 257-303.

Duncan, J., Johnson, R., Swales, M. & Freer, C. (1997). Frontal lobe deficits after head injury: unity and diversity of function. *Cognitive Neuropsychology*. 14(5), 713-741.

Duncan, J., R. J. Seitz, Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F.N. & Emslie, H. (2000). A Neural Basis for General Intelligence. *Science* 289: 457-460.

Duncan, J. (1986) Disorganisation of behaviour after frontal lobe damage*. Cognitive Neuropsychology,* 3, 271-290*.*

Eslinger, P.J. Grattan, I.M. & Geder, L. (1995). Impact of frontal lobe lesions on rehabilitation and recovery from acute brain injury. *Neurorehabilitation*, 5(2), 161-182.

Fan, J., Yao, Q. & Tong, H. (1996). Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika*, 83, 189-206.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160

Field, A. (2009)*. Discovering statistics using SPSS (3rd Edition)*. SAGE: London, pp. 481.

Freeman, K. A. 2003. Single subject designs. In Thomas, J.C. & Hersen, H. (Eds.) *Understanding research in clinical and counselling psychology*, 181–208. Mahwah, NJ: Lawrence

French, J. R. P. (1953), Experiments in Field Settings. In Festinger, L. & Katz, D.(Eds.). *Research Methods in the Behavorial Sciences*, Dryden Press, p.101.

Fuster, J.M. (1989). *The prefrontal cortex (2nd Ed*). New York: Raven Press

Gathercole, S. E., & Alloway, T. P. (2006). Short-term and working memory impairments in neurodevelopmental disorders: Diagnosis and remedial support. *Journal of Child Psychology & Psychiatry*. *47*, 4-15.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2003). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psycholog*y, 18, 1-16.

Gathercole, S. E., Tiffany, C., Briscoe, J., Thorn, A. S. C., & ALSPAC Team (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry*. *46*, 598-611.

Glisky, E.L., Schacter, D.L. & Tulving, E. (1986). Computer learning by memory-impaired patients: acquisition and retention of complex knowledge. *Nia*, 24, 313-328.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14, 477-485.

Goldstein, K. (1942). *After effects of brain injuries in war*. New York: Grune and Stratton.

Gray, J.M., Robertson, I.H., Pentland, B. & Anderson, S.I. (1992). Microcomputer based cognitive rehabilitation for brain damage: a randomised group controlled trial*. Neuropsycological Rehabilitation*, 2, 97-116.

Gronwall, D.M. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Percept Mot Skills,*.44(2), 367–373

Gueorguieva, K. (2004). More Over ANOVA. *Arch Gen Psychiatry* 61: 310-317.

Hamilton, M (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry.* 23: 56-62.

Heaton, R.K. (1981) *Wisconsin Card Sorting Test Manual*. Psychological Assessment Resources, Odessa, Florida

Hersh, N. & Treadgold, L. (1994). The rehabilitation of memory dysfunction by prosthetic memory cueing. *Neuorehabilitation*, 4, 187-197.

Holmes, J. Gathercole, S.E., Place, M., Dunning, D.L., Hilton, K.A. & Elliott, J.G. (2009a). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology*. DOI: 10.1002/acp.1589.

Holmes, J., Gathercole, S.E. & Dunning, D.L. (2009b). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*. 12, F1-F7.

Ivry, R.B. & Schlerf, J.E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Science*, 12(7), 273-2780. PMID:18539519

Jaeggi, S.M., Buschkuehl, M., Jonidas, J. & Perrig, W.J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America.* 195, 6829-6833.

Jaeggi, S.M., Studer-Luethi, B., Buschkuehl, M., Yi-Fen, S., Jonides, J. & Perrig. (2010). The relationship between n-back performance and matrix reasoning – implications for training and transfer. *Intelligence*, 38, 625-635.

Jaldow, E.J., Oakley, D.A. & Davey, G.C.L. (1989). Performance of decorticated rats on fixed interval and fixed time schedules. *European Journal of Neuroscience*, 1(5): 461

James, W. (1890). *Principles of Psychology*. New York: Holt.

James,E.K., Plant, G.T. & Warrington, M. (2001). Cortical Vision Screening Test. Pearson Assessment: Harlow.

Jobe, J.B., Smith, D.M., Ball, K., Tennstedt, S.L., Marsiske, M., Willis, S.L., Rebok, G.W., Morris, J.N., Helmers, K.F., Leveck, M.D. & Kleinman, K. (2001). ACTIVE: A cognitive intervention trial to promote independence in older adults. *Controlled Clinical Trials*, 22, 453-479.

Jovanovski, D., Zakzanis, K., Campbell, Z., Erb, S. & Nussbaum, D. (2012). Development of a novel, ecologically orientated virtual reality measure of executive function: the multitasking in the city test. *Applied Neuropsychology:Adult*, 0, 1-12.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99, 122-149.

Kensinger, E.A., Shearer, D.K., Locascio, J.J., Growdon, J.H. & Corkin, S. (2003). Working memory in mild Alzheimer's disease and early Parkinson's disease. *Neuropsychology*, 17(2), 230-9.

Khan, A., Sharma, N.K. & Dixit, S. (2008). Cognitive load and task condition in event- and time-based prospective memory: an experimental investigation. *J Psychol*, 142(5), 517-531.

Kim, B.R., Chun, M.H., Kim, L.S. & Park, J.Y. (2011). Effects of virtual reality on cognition in stroke patients. *Ann Rehabil Med*, 35, 450-459.

Klingberg, T. (2010). Training and plasticity of working memory. Trends in Cognitive Science, 14(7), 317-324.

Klingberg, T. (1998). Concurrent performance of two working memory tasks: potential mechanisms of interference. *Cerebral Cortex*, 8, 593-601.

Klingberg, T., Fernell, E., Olesen, P., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C.G., Forssberg, H. & Westerberg, H. (2005). Computerized Training of Working Memory in Children with ADHD – a Randomized, Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44, 177-186.

Klingberg, T., Fernell, E., Olesen, P., Johnson, M., Gustafsson, P., Dahlstrom, K., Gillberg, C.G., Forssberg, H. & Westerberg, H. (2008). Computerised training of working memory in children with ADHD – a randomised controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. 44, 177-186.

Klingberg, T., Forssberg, H. & Westerberg, H. (2002). Training of working memory with children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, 24, 781-791.

Knight,C., Alderman, N. & Burgess, P.W. (2002). Development of simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation*, 12(3), 231-255.

Kyllonen, P. C. & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, 14, 389-433.

Lagarde,G., Doyon, J. & Brunet, A. (2010).Memory and executive dysfunctions associated with acute posttraumatic stress disorder. *Psychiatry Research*, 177(1–2), 144–149.

Levine, B., Robertson, I.H., Clare, L., Carter, G., Hong, J., Wilson, B.A., Duncan, J. & Stuss, D.T. (2000). Rehabilitation of executive functioning: an experimental-clinical validation of goal management training. *Journal of the International Neuropsychological Society*, 6, 299-312.

Levine, B., Stuss, D.T., Winocur, G., Binns, M.A., Fahy, L., Mandic, M., Bridges, K. & Robertson, I.H. (2007). Cognitive rehabilitation in the elderly: effects on strategic behaviour in relation to goal management. *Journal of the International Neuropsychological Society*, 13, 143-152.

Logie, R.H., Trawley, S. & Law, A. (2011). Multitasking: multiple, domain-specific cognitive functions in a virtual environment. *Mem Cogn*, 39, 1561-1574.

Lundqvist, A., Grunstrom, K., Samuelsson, K. & Ronnberg, J. (2010). Computerised training of working memory in a group of patients suffering from acquired brain injury. *Brain Injury*, 24(10), 1173-1183.

Luria, A. R. (1966). *Higher cortical functions in man*. London: Tavistock.

Madsen Sjo, N., Spellerberg, S., Weidner, S. & Kihlgren, M. (2010). Training of attention and memory deficits in children with acquired brain injury. *Acta Paediatrica*. DOI:10.1111/j.651-2227.2009.01587.x

Makatura, T.J., Lam, C., Leahy, B.J. Castillo, M. & Kalpakjian, C.Z. (1999). Standardized memory tests and the appraisal of everyday memory. *Brain Injury*, 13(5), 355 -367.

Manchester, D., Priestly, N. & Jackson, H. (2004). The assessment of executive functions: coming out of the office. *Brain Injury*, 18(11), 1067-1081.

Manly, T., Hawkins, K., Evans, J., Woldt, K., & Robertson, I. H. (2002). Rehabilitation of executive function: facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia*, 40(3), 271-281.

Mateer, C.A. (2009). Neuropsychological interventions for memory impairment and the role of single-case design methodologies. *Journal of the International Neuropsychological Society*,15, 623-628.

Mc Dougall, S. & House, B. (2012). Brain training in older adults: Evidence of transfer to memory span performance and pseudo-Matthew effects. *Aging, Neuropsychology, and Cognition*, 19(1-2), 195-221.

Mc Dowell, S., Whyte, J.& D'Esposito, M. (1998). Differential effects of a dopaminergic agonist on prefrontal function in head injury patient, *Brain*, 121, 1155-1164.

Mc Grath, K.O.& Wong, S.O. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*, 1(1),30-46.

Meaux, J.B. & Chelonis, J.J. (2003). Time perception differences in children with and without ADHD. *Journal of Pediatric Health Care*, 17(2), 64-71.

Mehta, M. A., Owen, A. M., Sahakian, B. J., Mavaddat, N., Pickard, J. D., & Robbins, T. W. (2000). Methylphenidate enhances working memory by modulating discrete frontal and parietal lobe regions in the human brain. *Journal of Neuroscience*, 20, RC65.

Melby-Lervag & Hume, C. (2012). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, DOI:10.1037/a0028228.

Miller, G. A. (1956). The magical number seven, plus-or-minus two or some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

Mindspark (2011). Mindsparke- think.work.test.better. Available: http://www.mindsparke.com/ [12/01/2012]

Miotto, E.C.& Morris, R.G. (1998). Virtual planning in patients with frontal lobe lesions. *Cortex*, 34, 639-657.

Miotto, E.C., Evans, J.J., de Lucia, M.C.S. & Scaff, M. (2009). Rehabilitation of executive dysfunction: a controlled trial of an attention and problem solving treatment group. *Neuropsychological Rehabilitation*, 19(4), 517-540.

Morrison, A. & Chein, J. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory in the following paginated issue of *Psychonomic Bulletin & Review*, 18(1), 46-60, DOI: 10.3758/s13423-010-0034-0.

Muller, B. (1989). Approximate Power for Repeated -Measures ANOVA lacking sphericity. *Journal of the American Statistical Association* 84 (406): 549-555.

Nelson, H. E. (1991). *National adult reading test (NART): test manual (revised)*. Windsor, NFER-Nelson.

Olesen, P., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal brain activity after training of working memory. *Nature Neuroscience*, 7, 75-79.

Oswald, W.D., Gunzelmann, T., Rupprecht, R. & Hagen, B. (2006). Differential effects of single versus combined cognitive and physical training with older adults: the SimA study in a 5-year perspective. *European Journal of Ageing*, 3, 179-192.

Oswald, W.D., Rupprecht, R., Gunzelmann, T. & Tritt, K.(1996). The SIMA project: effects of 1 year cognitive and psychomotor training on cognitive abilities of the elderly. *Behavioural Brain Research*. 78, 67-72.

Owen, A. M. (1997). The functional organization of working memory processes within human lateral frontal cortex: the contribution of functional neuroimaging. *European Journal of Neuroscience*, 9, 1329-1339.

Owen, A.M., Hampshire, A., Grahn, J.A., Stenton, R., Dajani, S., Burns, A.S., Howard, R.J. & Ballard, C.G. (2010). Putting brain training to the test. *Nature*. DOI:10.1038/nature09042

Parmelee, P.A., Powell Lawton, M. & Katz, I.R., (1989). Psychometric properties of the Geriatric Depression Scale among the institutionalized aged. *Psychol. Assess*. 1, 331–338.

Pelosi, L., Geesken, J.M., Holly, M., Hayward, M. & Blumhardt, L.D. (1997). Working memory impairment in early multiple sclerosis: evidence from an event-related potential study of patients with clinically isolated myelopathy. *Brain*, 120, 2039–2058.

Perdices, M. & Tate, R. (2009). Single-subject designs as a tool for evidence-based clinical practice: are they unrecognised and undervalued? *Neuropsychological Rehabilitation*, 19(6), 904-927.

Petrides, M. (1994). Frontal lobes and working memory: evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (pp. 59-81): Elsevier Science.

Plemons, J.K., Willis, M.S. & Baltes, P.B. (1978). Modifiability of fluid intelligence in aging: a short-term longitudinal training approach. *Journal of Gerontology*, 33(2), 224-231.

Poulton, E. C. (1977). Arousing stresses increase vigilance. In R. R. Mackie (Ed.), *Vigilance: Theory, Operational Performance and Physiological Correlates* (pp. 423-460). New York: Plenum Press.

Rand, D., Basha-Abu Rukan, S.,Weiss, P.L. & Katz, N. (2009). Validation of the virtual MET as an assessment tool for executive functions. *Neuropsychological Rehabilitation*, 19(4), 583, 602.

Raskin, S. & Sohlberg, M. (2009). Prospective memory intervention: A review and evaluation of a pilot restorative intervention. *Special Issue of Brain Impairment*, 10 (1), 76-86.

Raven, J. (1981). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No.1: The 1979 British Standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales, Together With*

Comparative Data From Earlier Studies in the UK, US, Canada, Germany and Ireland. San Antonio, TX: Harcourt Assessment

Redick T.S., Shipstead, Z., Harrison, T. L., Hicks, K.L., Fried, D.E., Hambrick, D.Z., Kane, M.J. & Engle, R.W. (2012). No Evidence of Intelligence Improvement After Working Memory Training: A Randomized, Placebo-Controlled Study. *Journal of Experimental Psychology*: *General*, Jun 18.

Reitan, R. M. (1979). *Manual of administration of neuropsychological test batteries for adults and children* . Tucson , AZ : Neuropsychological Laboratory

Richter, P., Werner, J., Heerlein, A., Kraus, A. & Sauer, H. (1998). On the validity of the Beck Depression Inventory, *Psychopathology* 31,160-168.

Richmond, L. L., Morrison, A.B., Chein, J.M. & Olson, I.R. (2011). Working memory training and transfer in older adults. *Psychology and Aging*, DOI: 10.1037/a0023631.

Roa, S.C., Ranier, G. & Miller, E.K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821-824.

Robertson, I.H. & Murre, J.M.J. (1999). Rehabilitation of brain damage. Brain plasticity and principles of guided recovery. *Psychological Bulletin*, 125(5), 544- 575.

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F. & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions*. Brain* 133(1): 234-247.

Roy, M.M., Mitten, S.T. & Christenfeld, N.J.S. (2008). Correcting memory improves accuracy of predicted task duration. *Journal of Experimental Psychology: Applied*, 14(3), 266-275.

Sawaguchi, T. (1991). D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science*, 251, 947-50.

Schweizer, S., Hampshire, A. & Dagleish, T. (2011). Extending brain-training to the affective domain: increasing cognitive and affective executive control through emotional working memory training. *PLoSONE*, 6,e24372.

Serino, A., Ciaramelli,E., Di Santantonio, A., Malagu, S., Servadei, F. & Ladavas, E. (2007). A pilot study of the rehabilitation of central executive deficits after traumatic brain injury. *Brain Injury*. 21(1), 11-19.

Shallice, T. (1982). Specific impairments of planning.. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 298 (1089), 199–209.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.

Shallice, T. & Burgess, P. (1991). Deficit in strategy application following frontal lobe damage in man*. Brain*, 114, 727-741.

Shallice, T. & Burgess, P. (1996). The domain of supervisory processes and temporal organisation of behaviour. *Philosophical Transactions of the Royal Society of London*, 351, 1405-1412.

Shipstead, Z., Redick, T.S. & Engle, R.W. (2010). Does working memory training generalize? *Psychologica Belgica*, 50(3&4), 245-276.

Shipstead, Z., Redick, T.S. & Engle, R.W. (2012). Is working memory effective? *Psychological Bulletin.* Advanced online publication. DOI: 10.1037/a0027473.

Sigala, N., Kusunoki, M., Nimmo-Smith, I., Gaffan, D. & Duncan, J. (2008). Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proceedings of Natural Academy of Science USA*, 105(33), 11969-11974.

Sohlberg, M.M., White, O., Evans, E. & Mateer. (1992). Background and initial case studies into the effects of prospective memory training. *Brain Injury*, 6(2), 19-138. DOI: 10.3109/02699059209029651.

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology* 15 (2), 201–292.

Spikman, J. M., Boelen, D. H., Lamberts, K. F., Brouwer, W. H., & Fasotti, L. (2010). Effects of a multifaceted treatment program for executive dysfunction after acquired brain injury on indications of executive functioning in daily life. *Journal of the International Neuropsychological Society : JINS, 16*(1), 118-129.

St Clair-Thompson, H., Stevens, R., Hunt, A. & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology*, 30(2), 203-219.

Sternberg, R. J. (1999): The theory of successful intelligence. *Review of General Psychology, 3*, 292-316.

Sternberg, R.J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences of the United States of America.* 105(19),6791-9792.

Stigsdotter Neely, A. & Bäckman, L.(1995). Effects of multifactorial memory training in old age: Generalizability across tasks and individuals*. Journal of Gerontology Psychological Sciences*,3, 134–140.

Stine-Morrow, E. A. L., Parisi, J. M., Morrow, D. G. & Park, D. C., (2008). The Effects of an Engaged Lifestyle on Cognitive Vitality: A Field Experiment. *Psychology and Aging*, 23(4),778-786.

Sturm, W., Wilmes, K., Orgass, B. & Hartje, W. (1997). Do specific attention deficits need specific training? *Neuropsychological Rehabilitation*, 7(2), 81-103.

Stuss, D.T. & Benson, D.F. (1983). Neuropsychological studies of the frontal lobes. *Psychological Bulletin*, 95, 3-28.

Stuss, D.T., Robertson, I.H., Craik, F.I.M., Levine, B., Alexander, M.P., Black, S., Dawson, D., Binns, M.A., Palmer, H., Downey-Lamb, M. & Winocur, G. (2007). Cognitive rehabilitation in the elderly: A randomized trial to evaluate a new protocol. *Journal of the International Neuropsychological Society*, 13, 120-131.

SüB, H. M., Oberauer, K., Wittmann, W. W., Wilhelm, O. & Schulze, R. (2002). Working-memory capacity explains reasoning ability - and a little bit more. *Intelligence*, 20, 261-288.

Swanson, H. L. (1999). What develops in working memory? A life span perspective. *Dev.Psychol*, 35, 986-1000.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, Yamanouchi, T., Suzuki, S. & Kawashima, R. (2010). Training of working memory impacts structural connectivity. *The Journal of Neuroscience*, 30(9), 3297-3303.

Tate, R., McDonald, S., Perdices, M., Togher, L., Schultz, R. & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation*, 18(4), 385-401.

Thorell, L.B., Lindqvist, S., Bergman, S., Bohlin, G. & Klingberg, T. (2008). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 11, 969-976.

Tranter, L.J. & Koutstaal, W. (2008). Age and flexible thinking: an experimental demonstration of the beneficial effects of increased cognitively stimulating activity on fluid intelligence in healthy older adults. *Aging, Neuropsychology and Cognition*, 15, 184-207.

van Balen, H.G.G., Westzaan, P.S.H. & Mulder, T. (1996). Stratified norms for the Rivermead Behavioural Memory Test. *Neuropsychological Rehabilitation*. 6(3), 203-217.

Van der Molen, M.J., Van Luit, J.E.H., Van der Molen, W., Klugkist, I. & Jongmans, M.J. (2010). Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research*. 54(5), 433-447.

von Cramon, Matthes-von-Cramon, G. & Mai, N. (1991). Problem-solving deficits in brain-injured patients. *Neuropsychological Rehabilitation*, 1(1),45-64. DOI: 10.1080/09602019108401379.

Wechsler, D. (1996). *Wechsler Objective Number Dimensions (WOND).* New York: Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale- Revised*. New York: Psychological Corporation. Brace Jovanovich: Harcourt.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX.: The Psychological Corporation.

Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Ostensson, M.L., Bartfai, A. & Klingberg, T. (2007). Computerised working memory training after stroke – a pilot study. *Brain Injury*, 21(1), 21029.

Willis, S.L. & Nesselroade, C.S. (1990). Long-term effects of fluid ability training in old-old age. *Developmental Psychology*, 26(6), 905-910.

Willis, S.L., Blieszner, R. & Baltes, P.B. (1981). Intellectual training research in aging: modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology*, 73(1), 41-50.

Willis, S.L., Tennstedt, S.L., Marsiske, M., Ball, K., Elias, J., Mann Koepke, K., Morris, J.N., Rebok, G.W., Unverzagt, F.W., Stoddard, A.M. & Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA*, 296(23), 2805-2814.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H. & Evans, J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome*. Bury St Edmunds: Thames Valley Test Company.

Wilson, B., Cockburn, J. & Baddeley, A. (2003). *The Rivermead Behavioural Memory Test Second Edition. Manual*. Thames Valley Test Company: Bury St Edmunds.

Wilson, B., Cockburn, J., Baddeley, A. & Hiorns, R. (2003*). The Rivermead Behavioural Memory Test- Second Edition. Supplement Two*. Thames Valley Test Company: Bury St Edmunds.

Wilson, B.A. (2002). Towards a comprehensive model of cognitive rehabilitation. *Neuropsychological Rehabilitation*, 12(2), 97-110.

Wilson, B.A., Greenfield, E., Clare, L., Baddeley, A., Cockburn, J., Watson, P., Tate, R., Sopena, S., Nannery. R. & Crawford, J. (2008). *Rivermead Behavioural Memory Test 3*. Harlow: Pearson Assessment.

Wilson, B.A., Watson, P., Baddeley, A.D., Emslie, H. & Evans, J.J.(2000). Improvement or simply practice? The effects of twenty repeated assessments on people with and without brain injury. *Journal of the International Neuropsychological Society*, 6,469-479.

Winocur, G., Craik, F.I.M., Levine, B., Robertson, I.H., Binns, M.A., Alexander, M., Black, S., Dawson, D., Palmer, H., McHugh, T. & Stuss, D.T. (2007a). Cognitive rehabilitation in the elderly: overview and future directions. *Journal of the International Neuropsychological Society*, 13, 166-171.

Winocur, G., Palmer, H., Dawson, D., Binns, M.A., Bridges, K. & Stuss, D.T. (2007b). Cognitive rehabilitation in the elderly: an evaluation of psychosocial factors. *Journal of the International Neuropsychological Society*, 13, 153-165.

Woehrle, J. & Magliano, J. (2012). Time flies faster if a person has a high working-memory capacity. *Acta Psychologica*, 139(2), 314-319.

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., Roca, M., Antoun, N., Manes, F. & Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America.*107(33), 14899-14902.

Yesavage, J., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M. & Leirer, V.O. (1982-1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research***,** 17(1), 37–49.

Zigmond, A.S. & Snaith, R.P.(1983). The hospital anxiety and depression scale. *Acta Psychiatr Scand*., 67(6),361-70.

Zinke, K., Zeintl, M., Eschen, A., Herzog, C. & Kliegel, M. (2012). Potentials and limits of plasticity induced by working memory training in old-old age. *Gerontology.* 58:79-87.

Zung, W.W. (1965). A self-rating depression scale. *Archives of General Psychiatry* 12, 63–70.

Appendix A: Questionnaires for Study One Chapter Two GET

Appendix A.1: Relevant Only Condition

1. How many games are there in the task?

4	*10*	*1*	*2*

2. How will you record how engaging each task is?

By getting high points on the most engaging tasks	*I don't need to record how engaging the tasks are*	*By playing the best games again at the end*	*On a scale between 1 (not at all) and 10 (very)*

3. How many times must you play each game?

You must attempt to play only half of the games	*You must attempt to play all four games twice*	*You must attempt to play games as many times as possible*	*You must attempt to play all four games once*

4. What is the overall time for all tasks?

5 minutes	*There is no time limit. Tasks finish when all games are completed*	*15 minutes*	*20 minutes*

5. How much time must you aim to spend on each task?

You must spend the most amount of time on the task that you like the most	*You must spend as much time as is needed on tasks to get the highest score possible*	*You must spend roughly equal amounts of time on each task*	*It does not make any difference how much time is spent on each task*

6. How do you check the time?

You cannot check the time	*Using the clock button*	*The time is always displayed*	*Using the H button*

7. How do you switch to the next task?

Using the square key	*By asking the experimenter*	*The tasks switch automatically*	*By clicking on the mouse*

8. How do you control the direction?

By moving the mouse left and right	*By using the left and right arrow keys*	*Using the 6 and 8 keys.*	*You can't control the direction*

Appendix A.2: Relevant instructions first condition

1. There is a countdown from 12 to 0, what should you do?

Start before it reaches 5	*Start before it reaches 0*	*Start after it reaches 0*	*Start before it reaches 7*

Driving

2. How do you switch to the next game?

Using the square key	*By asking the experimenter*	*The tasks switch automatically*	*By clicking on the mouse*

3. How do you control the direction in each game?

By moving the mouse left and right	*By using the left and right arrow keys*	*Using the 6 and 8 keys.*	*You can't control the direction*

4. What is the aim of the driving task?

To stay in the middle of the road as much as possible	*To drive over the petrol pump symbols*	*To avoid the tyres on the side of the road*	*To move the car as much as possible*

5. How do you increase your score?

By collecting the tyres on the side of the road	*By spending as much time as possible in the middle of the road*	*By avoiding the petrol pumps*	*By driving over the petrol pumps*

6. What is the purpose of the tyres in the task?

They are obstacles to be avoided	*To increase your score by collecting them*	*The tyres are used as replacements when the car gets a puncture*	*The tyres are not relevant to the task*

7. How do you speed up or slow down the car?

By pressing the up and down arrow buttons	*It is not possible to speed up or slow down the car*	*By driving over the petrol pump symbols*	*By pressing the numbers between 1 and 9*

Catching

8. What is the overall time for all tasks?

5 minutes	*There is no time limit. Tasks finish when all games are completed*	*15 minutes*	*20 minutes*

9. How do you check the time?

You cannot check the time	*Using the clock button*	*The time is always displayed*	*Using the H button*

10. What is the aim of the catching task?

To catch as many of the falling balls as possible	*To avoid the falling balls*	*To catch as many of the blue balls as possible*	*To avoid the blue balls*

11. What happens after you check the time?

The time disappears after a few seconds	*The time is always displayed*	*The time disappears only when you change tasks*	*The time appears when you change tasks*

12. What is your main task?

To get very good at each game	*To evaluate how engaging I found each game*	*To fill in time before I do something else*	*To check that the games work*

13. How will you record how engaging each task is?

By getting high points on the most engaging tasks	*I don't need to record how engaging the tasks are*	*By playing the best games again at the end*	*On a scale between 1 (not at all) and 10 (very)*

Squash

14. How much time must you aim to spend on each task?

You must spend the most amount of time on the task that you like the most	*You must spend as much time as is needed on tasks to get the highest score possible*	*You must spend roughly equal amounts of time on each task*	*It does not make any difference how much time is spent on each task*

15. How many games are there in the task?

4	*10*	*1*	*2*

16. What is the aim of the squash task?

To score as many points as possible by avoiding the ball and allowing it to drop off the screen	*To score as many points as possible by bouncing the ball off the coloured walls*	*To score as many points as possible by hitting only the top pink wall with the ball*	*To keep your score over 200 points at all times*

17. Does it make a difference which wall the ball bounces off?

You can score different numbers of points depending on which wall the ball bounces off	*All the walls are worth the same number of points so it makes no difference*	*You need to only hit the top pink wall with the ball and avoid the side walls*	*There are no walls in this game*

18. What happens if the ball drops off the screen?

The aim of the game is to avoid the ball	*You gain points when the ball drops off the screen*	*You lose points if you allow the ball to drop out of the screen*	*Nothing happens when the ball drops off the screen*

19. When will you be asked to evaluate each task?

At the end of each task	*After having tried all tasks once*	*When the time has elapsed*	*You won't be asked to evaluate tasks*

Bricks

20. How many times must you play each game?

You must attempt to play only half of the games	*You must attempt to play all four games twice*	*You must attempt to play games as many times as possible*	*You must attempt to play all four games once*

21. What is the aim of the bricks task?

To score as many points as possible by avoiding the ball and allowing it to drop off the screen	*To keep your score over 200 points at all times*	*To bounce the ball off the bricks with the highest scores only*	*To score as many points as possible by bouncing the ball off the bricks and preventing it dropping off the screen*

22. How are points scored?

By hitting the ball off the paddle as much as possible	*By avoiding the ball and allowing it to drop off the screen*	*By playing the game for as long as possible while trying not to bounce the ball off the bricks*	*You can score different numbers of points depending on which bricks are removed*

23. Are bricks worth the same number of points?

No, they are worth different points	*There are no bricks in this game*	*Each brick is worth the same number of points*	*Removing bricks does not increase scores*

24. How many points are scored for hitting the outside wall?

You don't get any points for hitting the outside wall	*You score 100 points by hitting the outside wall*	*You must avoid hitting the outside wall*	*You loose points every time you hit the outside wall*

25. What happens if the ball drops off the screen?

You don't gain or lose any points for dropping the ball off the screen	*You score points by dropping the ball off the screen*	*It is not possible for the ball to drop off the screen*	*If the ball drops off the screen you lose points*

Appendix A3:. Relevant instruction last condition

1. There is a countdown from 12 to 0, what should you do?

Start before it reaches 5	*Start before it reaches 0*	*Start after it reaches 0*	*Start before it reaches 7*

Driving

2. What is the aim of the driving task?

To stay in the middle of the road as much as possible	*To drive over the petrol pump symbols*	*To avoid the tyres on the side of the road*	*To move the car as much as possible*

3. How do you increase your score?

By collecting the tyres on the side of the road	*By spending as much time as possible in the middle of the road*	*By avoiding the petrol pumps*	*By driving over the petrol pumps*

4. What is the purpose of the tyres in the task?

They are obstacles to be avoided	*To increase your score by collecting them*	*The tyres are used as replacements when the car gets a puncture*	*The tyres are not relevant to the task*

5. How do you speed up or slow down the car?

By pressing the up and down arrow buttons	*It is not possible to speed up or slow down the car*	*By driving over the petrol pump symbols*	*By pressing the numbers between 1 and 9*

6. How do you switch to the next task?

Using the square key	*By asking the experimenter*	*The tasks switch automatically*	*By clicking on the mouse*

7. How do you control the directions in the games?

By moving the mouse left and right	*By using the left and right arrow keys*	*Using the 6 and 8 keys.*	*You can't control the direction*

Catching

8. What is the aim of the catching task?

To catch as many of the falling balls as possible	*To avoid the falling balls*	*To catch as many of the blue balls as possible*	*To avoid the blue balls*

9. What happens after you check the time?

The time disappears after a few seconds	*The time is always displayed*	*The time disappears only when you change tasks*	*The time appears when you change tasks*

10. What is your main task?

To get very good at each game	*To evaluate how engaging I found each game*	*To fill in time before I do something else*	*To check that the games work*

11. How will you record how engaging each task is?

By getting high points on the most engaging tasks	*I don't need to record how engaging the tasks are*	*By playing the best games again at the end*	*On a scale between 1 (not at all) and 10 (very)*

12. What is the overall time for all tasks?

5 minutes	*There is no time limit. Tasks finish when all games are completed*	*15 minutes*	*20 minutes*

13. How do you check the time?

You cannot check the time	*Using the clock button*	*The time is always displayed*	*Using the H button*

Squash

14. What is the aim of the squash task?

To score as many points as possible by avoiding the ball and allowing it to drop off the screen	*To score as many points as possible by bouncing the ball off the coloured walls*	*To score as many points as possible by hitting only the top pink wall with the ball*	*To keep your score over 200 points at all times*

15. Does it make a difference which wall the ball bounces off?

You can score different numbers of points	*All the walls are worth the same*	*You need to only hit the top pink wall with the*	*There are no walls in this game*

depending on which wall the ball bounces off	*number of points so it makes no difference*	*ball and avoid the side walls*	

16. What happens if the ball drops off the screen?

The aim of the game is to avoid the ball	*You gain points when the ball drops off the screen*	*You lose points if you allow the ball to drop out of the screen*	*Nothing happens when the ball drops off the screen*

17. When will you be asked to evaluate each task?

At the end of each task	*After having tried all tasks once*	*When the time has elapsed*	*You won't be asked to evaluate tasks*

18. How much time must you aim to spend on each task?

You must spend the most amount of time on the task that you like the most	*You must spend as much time as is needed on tasks to get the highest score possible*	*You must spend roughly equal amounts of time on each task*	*It does not make any difference how much time is spent on each task*

19. How many games are there in the task?

4	*10*	*1*	*2*

Bricks

20. What is the aim of the bricks task?

To score as many points as possible by avoiding the ball and allowing it to drop off the screen	*To keep your score over 200 points at all times*	*To bounce the ball off the bricks with the highest scores only*	*To score as many points as possible by bouncing the ball off the bricks and preventing it dropping off the screen*

21. How are points scored?

By hitting the ball off the paddle as much as possible	*By avoiding the ball and allowing it to drop off the screen*	*By playing the game for as long as possible while trying not to bounce the ball off the bricks*	*You can score different numbers of points depending on which bricks are removed*

22. Are bricks worth the same number of points?

No, they are worth different points	*There are no bricks in this game*	*Each brick is worth the same number of points*	*Removing bricks does not increase scores*

23. How many points are scored for hitting the outside wall?

You don't get any points for hitting the outside wall	*You score 100 points by hitting the outside wall*	*You must avoid hitting the outside wall*	*You loose points every time you hit the outside wall*

24. What happens if the ball drops off the screen?

You don't gain or lose any points for dropping the ball off the screen	*You score points by dropping the ball off the screen*	*It is not possible for the ball to drop off the screen*	*If the ball drops off the screen you lose points*

25. How many times must you play each game?

You must attempt to play only half of the games	*You must attempt to play all four games twice*	*You must attempt to play games as many times as possible*	*You must attempt to play all four games once*

Appendix B: Multiple Errands Task Instruction and Performance Sheets – Versions A and B

Appendix B.1: Instructions (Version A)

Ensure participant has the following items:

- pen/pencil
- instructions on a clipboard for the participant
- carrier bag
- £10 note
- a watch and set it to 11:00 at the start and tell participants end time

Before leaving, obtain ratings for shopping efficiency and familiarity with shopping area (see separate sheet)

Once in shopping area walk participants through shopping centre, if not familiar with it, BEFORE giving instructions. Proceed to meeting place (next to front door of the shopping centre). Give the subject the clipboard, pen/pencil, carrier bag and £10 note.

Read the following instructions to the participant:

"**In this exercise I want you to complete three tasks. The tasks are: to buy the five items listed on this sheet** (*indicate and describe items on the sheet*); **to obtain and write down five pieces of information** (*indicate and describe items on sheet*); **and to meet me here in 25minutes after I have said "…begin the exercise" and tell me the time.**

However, whilst completing this exercise you must obey the rules listed on your instruction sheet (*indicate and describe rules on sheet*).

You must carry out all these tasks but you may do so in any order. You should spend no more than £6; although I've given you £10 you should spend no more than six. You should stay within the limits of this shopping centre. You are free to go upstairs if you like but you must not go outside any of the outside doors. No shop should be entered other than to buy something, so if you go into a shop it should be with the intention of buying something. You should not go back into a shop you have already been in, so if you've been into a particular shop you should not go back into it again. You should only buy items from shops, not stalls. You should buy no more than two items from Poundland. Take as little time as possible to complete this exercise without rushing excessively.

Finally, approach me and tell me when you have completed the exercise.

Is that clear, have you any questions? (*clarify any questions the participant has*)

Now tell me what you must do. (*ensure participant is clear about what they must do*)

"***Begin the exercise"*** (*Start timing at this point*)

On return, obtain the rating for how well the subject felt they executed the task (see separate sheet).

Appendix B.2: Task Instructions given to participants (Version A)

In this exercise you should complete the following three tasks:

1) You should buy the following items:

- packet of plasters ____
- birthday card ____
- small bottle of water ____
- one packet of dishcloths ____
- a single pen ____

2) You should obtain the following the information and write it down in the spaces below:

1. What is the phone number of the public phone closest to Mothercare?

____

2. What is the closing time of the shopping centre on Saturday?

____

3. What is the price of 2 student cinema tickets on weekend evenings?

____

4. How many mobile phone shops are there in the shopping centre?

____

5. What shop is between Evan and Barratts?

3) You must meet me at the front doors 25 minutes after you have started this task and tell me the time.

____

Whilst carrying out this exercise you must obey the following rules:

- You must carry out all these tasks but may do so in any order
- You should spend no more than £6
- You should stay within the limits of the shopping centre. Do not leave the outside doors.
- No shop should be entered other than to buy something
- You should not go back into a shop you have already been in
- You should not buy any item from the stalls
- You should buy no more than 2 items in Poundland
- Take as little time to complete this exercise without rushing excessively

Appendix B.3: Scoring sheet (version A)

<u>Errors:</u>

- *inefficiencies* – where a more efficient strategy could be applied

- *rule breaks* – where a specific rule (social or task rule) was broken

- *interpretation failure*- requirements of task were misunderstood

- *task failure* - a task not completed satisfactorily

Tasks:

Items to be bought:

- packet of plasters _____
- birthday card _____
- small bottle of water _____
- One packet of dishcloths _____
- A single pen _____

Observations:

Order in which the task was carried out:

Number of items bought after:

3 minutes:

6 minutes:

9 minutes:

12 minutes:

15 minutes:

18 minutes:

21 minutes:

24 minutes:

End:

What is the phone number of the public phone closest to Mothercare?

What is the closing time of the shopping centre on Saturday?

What is the price of 2 student cinema tickets on weekend evenings?

How many mobile phone shops are there in the shopping centre?

What shop is between Evan and Barratts?

Time taken to complete exercise:

Number of times help was asked for:

Money spent:

Appendix B.4.: Participant self-rating scales (Used in versions A and B)

Before beginning the task:

1. How efficient would you say you were with tasks like shopping?

hopeless excellent

1 2 3 4 5 6 7 8 9 10

2. How well do you this shopping centre?

not at all visited once or twice visit occasionally visit regularly

1 2 3 4

On completion of the task:

3. How easy did you find the task?

very difficult difficult moderate easy very easy

1 2 3 4 5

4. How well do you think you did with the shopping task?

hopeless excellent

1 2 3 4 5 6 7 8 9 10

Appendix B.5 Instructions (Version B)

Ensure participant has the following items:

- pen/pencil
- instructions on a clipboard for the participant
- carrier bag
- £10 note
- a watch and set it to 11:00 at the start and tell participants end time

Before leaving, obtain ratings for shopping efficiency and familiarity with shopping area (see separate sheet)

Once in shopping area walk participants through shopping centre, if not familiar with it, BEFORE giving instructions. Proceed to meeting place (next to front door of the shopping centre). Give the subject the clipboard, pen/pencil, carrier bag and £10 note.

Read the following instructions to the participant:

“**In this exercise I want you to complete three tasks. The tasks are: to buy the five items listed on this sheet** (*indicate and describe items on the sheet*); **to obtain and write down five pieces of information** (*indicate and describe items on sheet*); **and to meet me here in 25minutes after I have said “…begin the exercise” and tell me the time.**

However, whilst completing this exercise you must obey the rules listed on your instruction sheet (*indicate and describe rules on sheet*).

You must carry out all these tasks but you may do so in any order. You should spend no more than £6; although I’ve given you £10 you should spend no more than six. You should stay within the limits of this shopping centre. You are free to go upstairs if you like but you must not go outside any of the outside doors. No shop should be entered other than to buy something, so if you go into a shop it should be with the intention of buying something. You should not go back into a shop you have already been in, so if

you've been into a particular shop you should not go back into it again. You should only buy items from shops, not stalls. You should buy no more than two items from Poundland. Take as little time as possible to complete this exercise without rushing excessively.

Finally, approach me and tell me when you have completed the exercise.

Is that clear, have you any questions? (*clarify any questions the participant has*)

Now tell me what you must do. (*ensure participant is clear about what they must do*)

"***Begin the exercise"*** (*Start timing at this point*)

On return, obtain the rating for how well the subject felt they executed the task (see separate sheet).

Appendix B.6: Task instructions given to participants (Version B)

In this exercise you should complete the following three tasks:

1) You should buy the following items:

- packet of safety pins ____
- Thank You card ____
- a packet of chewing gum ____
- a bar of soap ____
- a pair of black shoe laces ____

2) You should obtain the following the information and write it down in the spaces below:

6. What is the phone number of the public phone closest to the front entrance (near Next)?

____

7. What is the closing time of the shopping centre on Wednesday?

____

8. What is the price of 2 student cinema tickets on weekday afternoon?

____

3) You must meet me at the front doors 25 minutes after you have started this task and tell me the time.

____

Whilst carrying out this exercise you must obey the following rules:

- You must carry out all these tasks but may do so in any order
- You should spend no more than £6
- You should stay within the limits of the shopping centre. Do not leave the outside doors.
- No shop should be entered other than to buy something
- You should not go back into a shop you have already been in
- You should not buy any item from the stalls
- You should buy no more than 2 items in Poundland
- Take as little time to complete this exercise without rushing excessively

Appendix B.7: Scoring sheet (Version B)

Errors:

- *inefficiencies* – where a more efficient strategy could be applied

- *rule breaks* – where a specific rule (social or task rule) was broken

- *interpretation failure*- requirements of task were misunderstood

- *task failure* - a task not completed satisfactorily

Tasks:

Items to be bought:

- packet of plasters _____
- birthday card _____
- small bottle of water _____
- One packet of dishcloths _____
- A single pen _____

Observations:

Order in which the task was carried out:

Number of items bought after:

3 minutes:

6 minutes:

9 minutes:

12 minutes:

15 minutes:

18 minutes:

21 minutes:

24 minutes:

End:

Other:

What is the phone number of the public phone closest to the front entrance (near Next)?

What is the closing time of the shopping centre on Wednesday?

What is the price of 2 student cinema tickets on weekday afternoon?

How many shops are there in the shopping centre beginning with “H”?

What shop is between “Vodafone” store and “3” store?

Time taken to complete exercise:

Number of times help was asked for:

Money spent:

Appendix C: Goal Management Questionnaire

This questionnaire is about problems that most people experience from time-to-time. Please circle the number that best describes how much of a problem this has been for *you* in the *last two weeks.* The scale goes from 0 (not a problem at all) through to 10 – a really major problem.

	<<no problem at all	a very>> major problem
1. Finding that you don't achieve many of the things that you want to get done in a day?	0 1 2 3 4 5 6 7 8 9 10	
2. Walking into a room and forgetting what it was that you had come for?	0 1 2 3 4 5 6 7 8 9 10	
3. Finding that you don't have time to stop and think?	0 1 2 3 4 5 6 7 8 9 10	
4. Something that you needed to do just "slipped your mind" (e.g. forgetting to post a letter whilst you were out)?	0 1 2 3 4 5 6 7 8 9 10	
5. Not actually having a very clear idea of what you are trying to achieve?	0 1 2 3 4 5 6 7 8 9 10	
6. Having to go back and re-read a paragraph because you didn't take the information in the first time?	0 1 2 3 4 5 6 7 8 9 10	
	<<no problem at all	a very>> major problem
7. Not being realistic about how long something will take to complete?	0 1 2 3 4 5 6 7 8 9 10	
8. Forgetting something that needed to be done at a certain time (e.g. an appointment, a TV programme that you wanted to watch, calling someone)?	0 1 2 3 4 5 6 7 8 9 10	
9. Feeling too busy, hassled, like you aren't in control?	0 1 2 3 4 5 6 7 8 9 10	
10. Trying to please everybody?	0 1 2 3 4 5 6 7 8 9 10	
11. Find that you haven't been listening to important information that someone is telling you?	0 1 2 3 4 5 6 7 8 9 10	
12. Making a mistake because you weren't thinking about what you were doing at the time?	0 1 2 3 4 5 6 7 8 9 10	
13. Not remembering where you had got to in a task?	0 1 2 3 4 5 6 7 8 9 10	
14. Worrying too much about things that you need to achieve?	0 1 2 3 4 5 6 7 8 9 10	

15. Find that you have done things in the wrong order (e.g. checking that you have something you need when you arrive at a destination rather than before you leave, cooking something that takes a short time before something that needs longer)?	0 1 2 3 4 5 6 7 8 9 10
16. Trying to do or think about too many things at once?	0 1 2 3 4 5 6 7 8 9 10
17. Spent too long searching for things (e.g. spectacles, keys, a scrap of paper with an important phone number on it)?	0 1 2 3 4 5 6 7 8 9 10
18. Not having what you need with you at the right time (e.g. going to the shops without a payment card or enough cash, turning up at a meeting without the relevant papers, going out without your keys)?	0 1 2 3 4 5 6 7 8 9 10
19. Not remembering whether you had done an everyday activity or not (e.g. not remembering whether you had locked the door or turned the light off)?	0 1 2 3 4 5 6 7 8 9 10
20. Feeling that others expect too much from you?	0 1 2 3 4 5 6 7 8 9 10
21. Losing track of the time?	0 1 2 3 4 5 6 7 8 9 10
22. Getting distracted from an important activity by something that is less important?	0 1 2 3 4 5 6 7 8 9 10
23. Getting “carried away” with something, not stopping to think about it?	0 1 2 3 4 5 6 7 8 9 10
24. Avoided thinking about a problem because it just seems too complicated?	0 1 2 3 4 5 6 7 8 9 10
25. Feeling worried about how well you are coping?	0 1 2 3 4 5 6 7 8 9 10
26. Feeling overwhelmed by the things that you need to do?	0 1 2 3 4 5 6 7 8 9 10
27. Not thinking something through before acting (e.g. making a “rash” purchase, accidentally offending someone)?	0 1 2 3 4 5 6 7 8 9 10
28. Running out of time because you got too caught up in something that you were doing?	0 1 2 3 4 5 6 7 8 9 10
29. Daydreaming rather than thinking about what you were doing?	0 1 2 3 4 5 6 7 8 9 10
30. Finding your mind “wandering off” a task you were doing?	0 1 2 3 4 5 6 7 8 9 10
31. Having difficulty making decisions?	0 1 2 3 4 5 6 7 8 9 10
32. Had problems in organizing your time (e.g. arranging appointments that clash, not giving yourself enough time to get somewhere, being late).	0 1 2 3 4 5 6 7 8 9 10

33. That you keep making the same mistake (e.g. when using a computer or dialling a number)?	0 1 2 3 4 5 6 7 8 9 10
34. Having to go back for something that you had forgotten to take with you?	0 1 2 3 4 5 6 7 8 9 10

Appendix D: Logbook given to participants in chapter seven

<u>Hello!</u>

Welcome to this cognitive training programme.

Many thanks for agreeing to take part in this study. Your participation is greatly appreciated.

We are asking all participants in this study to use this logbook. This logbook serves a number of uses – to keep track of progress, to have somewhere to make notes, to have a record of notes from video sessions and to make note of any difficulties you may come across during training. Hope you find it a useful part of the training programme.

There are nine videos in this training package. The first one should be watched prior to beginning the first training session and the final video is designed to be watched at the end of the final training session. The other videos should be watched as part of the training every 4 days or so. They will be listed in the links for the training that is sent to your email inbox and should work by clicking them.

The videos are in the following order:

1) Introductory video
2) Goal setting video
3) Rewarding yourself video
4) Stating your intentions video
5) Bullet points in everyday life video
6) Why is working memory important? video
7) Managing distraction video
8) Memory strategies video
9) Final video- making the most of your cognitive training

If you have any difficulties then do not hesitate to get in contact with any questions. The researchers running this study are Sinead Hynes (sinead.hynes@mrc-cbu.cam.ac.uk, 01223 273635 (direct line) or 01223 355294 ext 599, and Tom Manly (tom.manly@mrc-cbu.cam.ac.uk, 01223 760683.

I hope you enjoy the training.

Contents:

Consent:

I understand that all information collected in this training is treated with the strictest confidentiality within the research team. No information from which a person could be identified is ever released.

I understand that I am completely free to withdraw from the training at any stage without having to provide a reason.

I am willing for the contents of this logbook to be included in our analysis of the research.

I certify that I have read and understood this form, and agree to take part in the project of my own will:

Name: ______________________________ (please print)

Signed: _______________________________ Date: ___________________

<u>Video One – Introductory video</u>

1. Who is the programme for?

The programme is designed for anyone who is interested in tuning their cognitive skills. We will ask you to set your **own goals** about how far and how fast you want to progress.

2. What happens in the programme?

In this programme you will be able to practice highly focused tasks that emphasise different sorts of cognitive skill.

You may find some of the tasks more challenging than others but it is not about which you can do best in – they are all designed to tackle different areas.

3. How much should I do? How long will it take?

Practice for 5-days a week and then give yourself a couple of days off for the five-weeks of the programme. It is likely that the intensity of the training is important for the eventual gains and leaving too long a gap between training sessions may take you back a few stages.

4. How do I start?

Click on the links sent to your email. The link will then take you to the first task. Click on instructions to find out what you need to do. We strongly recommend that you wait until the next day to tackle the next training session because the gains may be consolidated during this period. Aim for daily training.

5. When to do the training?

Try and clear you mind of the other things that you need to do.

Don't try the training when you are feeling drowsy or sleepy or feel unwell.

Don't try the training when there are many distractions going on around you

<u>Video Two – Goal setting</u>

- To help make the most of your training in everyday life, to motivate you and to monitor your progress it is useful to set some targets or goals for yourself.
- This video is all about setting SMART goals.
- What is a SMART goal?
    - Specific - try to keep each goal only about one thing
    - Measurable - it has to be clear whether or not the goals were reached
    - Agreed - the goals are important to you
    - Realistic - attainable in relation to what you are working on and the time you have to do it
    - Time Limited - they should all have a time-scale within which they can be achieved
- Set 2 or 3 goals for yourself each week relating to the training tasks and also at least one everyday goal that relates to your own life. There are examples in the goal setting section of this logbook.
- Set short, medium and long-term goals for yourself relating to training and everyday tasks.
- In setting goals for yourself you are making the training more individualised. Having realistic, achievable targets set should make the training more interesting to undertake.
- It is hoped that setting and reviewing different types of goals will make you more aware of reviewing your own everyday challenges – i.e. cooking a meal.

<u>Video Three – Rewarding yourself</u>

- Most things in life work better if we are rewarded and rewards have been shown to help in learning.
- For each goal set try and build in a reward for when you have achieved that goal.
- Don't forget that cognitive training is not about getting the answers correct, it's all about ***trying*** to get the answers correct – the effort *is* the training not the number of items that you get correct.
- As part of the training, we would encourage you to take some of these challenges and just see how you get on!
- Feeling more confident in our cognitive and other abilities can help us feel better and engage more in activities that we might otherwise avoid.

Video Four – Stating your intentions

- Telling ourselves aloud what we will do can have some interesting effects and it may be interesting to incorporate this into your cognitive training programme and try it out in everyday life.
- Try saying to yourself: "When I encounter a difficult problem in everyday life I will break it down into manageable steps, like the problems in the training"

  or

  "When I have to take in a lot of information, I will try and get into the same mode that I developed when repeating back digits etc in the training."
- Try writing some of these intentions down and then checking back to see whether these made any difference.

Video Five – Bullet points in everyday life

- The aim of this video is to encourage you to break problems down into more readily solvable steps and apply this in real-life tasks as well as the training.
- It is a way of separating information into relevant chunks that can make it easier to think about problems, to remember your intentions and to act on them.
- Breaking big problems down into smaller manageable steps leaves us with an immediate set of tasks that is a lot less overwhelming than one big problem – essentially forming a simple "to-do" list.
- Try to think of a challenge that you have faced that you either did break down in this way or that might have benefitted from this approach. It is a useful exercise to sit down with a piece of paper breaking down each main goal into different subgoals and thinking about the best order to do things in.
- Using bullet points can be helpful for both breaking down logical problems in the training and everyday life problems.

Video Six- Why is working memory important?

- The brain has quite a lot of distinct memory systems.
- Working memory is a very temporary store for holding information- like a mental blackboard that is constantly being written on and wiped out.
- Working memory has a very limited capacity and this limited capacity is the main reason that people are only really good at doing one thing at a time and means that the current contents of conscious thought has a limited capacity.
- Working memory is needed for pretty much every single activity in which we are involved.
- Scientists have shown in recent studies with people of different ages and population groups that repetitive training on working memory, such as the training you are undertaking, have led to ***general*** improvements in performance. They have shown that people improve on tasks that they have trained on and tasks and assessments that were different and unrelated to the training.

<u>Video Seven – Managing distraction</u>

- It would be strange to spend so long increasing your working memory capacity but **not** take steps to maximise your available capacity by cutting down distractions where possible.
- Practice trying to clear your mind of thoughts about how well you are doing or how much time is left and try to engage fully in the task.
- Try and put aside concerns or worries for the duration of the training session.
- Take concrete steps where possible to manage your environment to maximise your focus on the tasks. If you find good ways of clearing your mind of clutter and getting into the zone during training, try and use these strategies when needed in everyday situations.
- Distractions of one sort or another consume your working memory capacity.
- Managing distractions – both internal and external - are an important part of training. Learn to focus available resources on the task at hand.
- In general, you will probably get the most out of training by being fully engaged and focused.
- As an experiment, however, you can explore resisting deliberate distraction to see if you can develop these skills - practicing maintaining your focus on the tasks ***despite*** the distraction.

<u>Video Eight – Memory strategies</u>

- Scientists have long known that the ***way*** in which people learn information changes their ability to remember it.
- The more deeply we process information, the more we think about it, the more we will remember.
- Processing information deeply by creating images and stories can make it easier to remember, particularly if they are striking or unusual.
- Using images can help us to remember information – the more bizarre the better!
- We can reduce that amount of information that we need to remember …
    - By using images to group things together
    - By grouping numbers together to form bigger units
    - By grouping locations together into shapes.
- As with all cognitive training, the *use* of all of these strategies is made easier by practice.

Video Nine – Making the most of your cognitive training

- Many older people with extremely good memories, people can ***think*** that they have poor memories when, in reality, they do not. Older people begin to ***notice*** their memory mistakes more and interpret this as evidence that they have a poor memory.
- Retirement can lead people to feel differently about their abilities. They may lose confidence in their abilities and begin a cycle in which they become less involved.
- Low mood, perceived loss of role and events such as bereavement can have a very real impact on memory and other cognitive function.
- Even where there is some real decline in cognitive abilities, factors such as loss of confidence, inactivity and lower mood can make the situation yet worse.
- Take time to think about everyday situations where stopping and thinking, or focusing your attention might be useful.
- Try to find an ideal balance between seeking stimulating challenges and taking things in a step-by-step fashion.
- Seek out opportunities to test your wits in whatever time you have available - it's about trying, persevering, thinking about the problems in different ways and maintaining your focus.
- Seek out opportunities to remain engaged.
- Use a PQRST technique if tackling a challenging reading assignment:
    - P –preview
    - Q- questions
    - R- read
    - S – summarise
    - T- test
- Help to consolidate and continue the training process by seeking out cognitive challenges in everyday life, practicing and gradually increasing what you can achieve.
- Set yourself manageable, realistic progressive cognitive challenges using SMART goals.

- Notice your successes and your progress.
- Try the same thing with physical activity – set manageable, realistic, progressive goals.

Checklist of training completed

This page is a checklist of each day of training completed. You can use it to tick off the days of training completed. There is also a space beside each day to fill in how long the training took. The training will more than likely take different people different lengths of time and may vary by day so it would be helpful if we had a rough estimation of how long the training is taking. If you forget to fill this section in or miss a few days don't worry about it.

Day one	________	Time taken	____________
Day two	________	Time taken	____________
Day three	________	Time taken	____________
Day four	________	Time taken	____________
Day five	________	Time taken	____________
Day six	________	Time taken	____________
Day seven	________	Time taken	____________
Day eight	________	Time taken	____________
Day nine	________	Time taken	____________
Day ten	________	Time taken	____________
Day eleven	________	Time taken	____________
Day twelve	________	Time taken	____________
Day thirteen	________	Time taken	____________
Day fourteen	________	Time taken	____________
Day fifteen	________	Time taken	____________

Day sixteen	________	Time taken	____________
Day seventeen	________	Time taken	____________
Day eighteen	________	Time taken	____________
Day nineteen	________	Time taken	____________
Day twenty	________	Time taken	____________
Day twenty-one	________	Time taken	____________
Day twenty-two	________	Time taken	____________
Day twenty-three	________	Time taken	____________
Day twenty-four	________	Time taken	____________
Day twenty-five	________	Time taken	____________

Goal Setting

These pages relate to the second video about goal setting. It is a space for you to make note of your goals set during the training. These goals are to be filled in week by week and focus on both everyday challenges and goals that are related to the online training. There is space for short, medium and long-term goals.

If you take the digit span test as an example – remembering a sequence of numbers in a row- then a short-term goal for that, something that could be achieved in a week, might be something like: "By Wednesday of next week I will have increased my digit span by one, making it 5". Obviously if your digit span to start with was already 5 then there would be very little reason in having this as a goal. If your digit span was something like 2 or 3 digits then depending on how difficult you find the task it could be achievable or it might be something that you need to allow yourself more time to achieve – "A week from next Wednesday (STATE DATE) I will…"

An everyday equivalent of remembering a sequence of numbers could be something like: "By Wednesday of next week I will have successfully shopped for 5 items in the supermarket without using a shopping list".

Goals should try to relate to different areas of life such as: communication- remembering names, productivity – remembering a shopping list, or leisure – remembering the scores of a football match. Other goals may relate to planning, prioritising etc.

This is space for your weekly (short-term) goals to be written – in SMART form.

Training goals:

____
____
____
____
____
____
____
____
____

Every-day life goals:

-

____
____
____
____
____
____
____
____
____

This is space for your weekly (short-term) goals to be written – in SMART form.

Training goals:

____

____

____

____

____

____

____

____

____

Every-day life goals:

-

____

____

____

____

____

____

____

____

____

This is space for your medium-term goals to be written – in SMART form.

Training goals:

____

____

____

____

____

____

____

____

____

Every-day life goals:

-

____

____

____

____

____

____

____

____

____

This is space for your longer-term goals set- these may be by the end of training programme or half-way through the training.

Training goals:

____

____

____

____

____

____

____

____

____

Every-day life goals:

-

____

____

____

____

____

____

____

____

____

This is a space for you to reflect on whether there are any other goals that you wish to work on when you have finished the training. There may be goals that you have not achieved yet that you want to continue to work on or it may be that you wish to set yourself new challenges now that the training has finished.

Training goals:

____

____

____

____

____

____

____

____

____

Every-day life goals:

-

____

____

____

____

____

____

____

____

____

<u>Notes</u>

Appendix E: Video scripts and training links

Appendix E.1: Final video script

Video 9 - Making the most of your cognitive training

In previous videos we talked about setting SMART goals for yourself – goals that are

Specific

Measureable

Agreed

Realistic and

Time-limited.

We talked about how SMART goals could be applied to making the most of your training and how they can be very useful for achieving things in everyday life.

We talked about how reward can increase training effects and how linking the achievement of your SMART goals with some pleasurable reward can be useful.

We talked about the dangers of over-perfectionist thinking – not having a go at problems because you judge early on that it will be too difficult.

In the next video we described some interesting effects of stating aloud your intentions to yourself– for example, saying to yourself before you begin a cognitive training session "If I reach a problem that I find particularly difficult I will make sure to give it my full effort"

We then talked about bullet points – ways of separating information into relevant chunks that can make it easier to think about problems, to remember your intentions and to act on them.

We then looked at how practicing breaking down logical problems into bullet points of things to check could help and how, although this slows you down, can help you get into good mental habits that then allows you to speed up again – but possibly speed up better!

We examined how everyday challenges can be broken down into bullet points that allow you to focus on the most relevant information and help you to remember key stages in your plans.

We then talked about working memory – how this describes the capacity of your moment-to-moment conscious thought – and how the training was directed at increasing this capacity and helping you to manage your working memory better.

We talked about managing internal and external distraction as a means of increasing the working memory capacity available to the task at hand.

In the last video, we talked about useful memory strategies – how the more you process or think about some piece of information the more likely you are to remember it and how forming to-be-remembered information into a single image can make it much easier to remember. We applied this to remembering strings of numbers or locations by grouping the numbers into double digit numbers or 4-digit dates like 1946 or by remembering shapes made by different locations lighting up.

This last video is about using the gains you have made in your cognitive training programme in everyday life and maintaining a cognitive training stance to your everyday activities.

Before we do that, we want to challenge some commonly held ideas about the aging process.

Many people think that getting older is inevitably associated with a loss of cognitive ability, in particular memory.

This is not true; there are many, many older people with extremely good memories.

In addition, people can ***think*** that they have poor memories when, in reality, they do not.

One reason for this is that older people begin to ***notice*** their memory mistakes more and interpret this as evidence that they have a poor memory. Exactly the same lapse could have occurred to a 20-year-old, who would probably think nothing of it.

Retirement can lead people to feel differently about their abilities. Suddenly, for example, they may feel that they have lost their useful work role and drop their levels of physical and cognitive activity markedly. They may lose confidence in their abilities and begin a cycle in which they become less involved.

Low mood, which can sometimes accompany retirement, perceived loss of role and events such as bereavement can have a very real impact on memory and other cognitive function. If people think this is part of a real, irreversible decline this can further lower their mood and set up a vicious circle – in effect a self-fulfilling prophesy.

People can ignore strengths that come with age – such as a wide experience and a lifetime of using strategies for coping.

Even where there is some real decline in cognitive abilities, factors such as loss of confidence, inactivity and lower mood can make the situation yet worse.

So, returning to the cognitive training programme - as we have said in previous videos, the point about the programme is to improve your abilities in everyday situations.

To help this happen, try and “take your cognitive training head” out and about with you.

Think about problems that you come across in everyday life like you thought about the problems during the training.

Think about how well you focused on remembering digits and so on during the training and try and ***use*** that focus in everyday activities, at work or when you are shopping.

When you are doing the training, take time to think about everyday situations where stopping and thinking, or focusing your attention might be useful.

Make clear intentions to use these strategies in those situations.

_____

In this last part of the video we offer some advice that might at first seem contradictory. That is because we are talking about an ideal balance between seeking stimulating challenges and taking things in a step-by-step fashion. Where the ideal balance lies will depend on your particular circumstances – where you are starting from and how you respond to challenges and disappointments. We will leave that balance up to you.

So to start with - *seeking challenges*

Seek out opportunities to test your wits in whatever time you have available.

If you don’t feel confident about your abilities, take on private quizzes and challenges in the newspaper, buy or borrow puzzle books from the library. Remember that, just like with this cognitive training, it’s not always about getting the answers right (although that is always nice) it’s about trying, persevering, thinking about the problems in different ways and maintaining your focus – that is a good way to take your cognitive training forwards.

Use SMART goals and statements to guide how you take on challenges. If you currently don’t take on these challenges set realistic, achievable goals like “I will try and complete one page of puzzles or seduko by the end of this week.” Increase your activity as you achieve each goal. Remember, achieving realistic goals can be motivating. Not achieving unrealistic goals can reduce your motivation.

If you are the type of person who is usually very happy to leave organisational issues to others, try volunteering for a manageable task. Treat it like a project taking the time to think about what is required, making bullet-point lists of what you need to do when in order to get the project done.

If you have retired, seek out opportunities to remain engaged. For example, many people use some of their time to make their skills available to others by volunteering for charities in various ways.

Keep active; there is good evidence that regular physical activity can help people in many ways including supporting their cognitive skills – their memory, thinking and attention. What active means will depend a bit on your age, health and current levels of activity and we are not suggesting that you go out and run a marathon!

Again, a SMART goal approach is useful. Where ever you are starting from set goals which are a small step from there. If you can walk for 5 minutes, set a first goal to walk for 6.

Set yourself realistic challenges in cognitive activity. For example, you might decide to read and persevere with a book that you previously might have feared that you would not understand.

A good way to increase your understanding and memory for what you have read is to use a PQRST technique. This stands for

P - Preview the book or information that you need to understand – have a quick flick though to see what it might be about.

Q – Set yourself some questions that you will be able to answer after reading the information – this will give your reading greater purpose.

R – Read the information with these questions in mind.

S – Summarise what you have read in your own words

T – Test yourself. See how well you can remember what you have read without looking back at the text.

Set yourself realistic goals in your reading and stick to them.

There was a nice research project with a man who had suffered a brain injury. Like many of us he found that his mind wandered off when he was reading difficult information, in his case an accountancy text book.

He found the amount of time that he could read for without his mind wandering off. It wasn’t very long.

But, he set himself the goal of reading for 10% longer. For example if he started out being able to read for about a minute without his mind drifting he set himself the goal of reading for 1 minute and 6 seconds. **Then he stopped.**

The next time he read he would aim for 10% longer again – and then stop.

By stopping at his time limits he made it extremely unlikely that he would experience a slip.

Gradually, he increased the duration for which he could read and after not too long, achieved his goal.

You can probably see that a general theme is emerging. The theme is breaking down big steps into lots of manageable small steps as a way of achieving extraordinary things. This technique has the benefits of offering you lots of positive reinforcement as your manageable steps are achieved. This can have beneficial effects on people's mood – giving them multiple experiences of success rather than failure.

So, in summary

People can sometimes feel dispirited because they feel that a loss of cognitive skills is inevitable as we age. Think about how things like ***worry*** about losing abilities, inactivity and loss of confidence can influence your cognitive abilities.

Try and make the gains from the cognitive training programme pay dividends in your everyday life by

- Thinking about the training in everyday life situations and applying some of the strategies that you may have developed.
- Thinking about everyday challenges during your training

Help to consolidate and continue the process by seeking out cognitive challenges in everyday life, practicing and gradually increasing what you can achieve.

Set yourself manageable, realistic progressive cognitive challenges using SMART goals.

Notice your successes and your progress.

Try the same thing with physical activity – set manageable, realistic, progressive goals.

Remember, you can watch any of these videos again by clicking on the links at the bottom of the email.

Finally, thanks very much for taking part in this Medical Research Council Cognitive Training programme. We welcome any feedback that you may have!

Appendix E.2: Strategy training video links

Training Video 1-Windows Media.wmv
<http://www.youtube.com/watch?v=eSApFLOIMN8> Video 1 Intro to training

Training Video 2-Windows Media.wmv <http://www.youtube.com/watch?v=znHz-Si3CaU> Video 2 Goal setting

Training Video 3-Windows Media.wmv
<http://www.youtube.com/watch?v=jETFVDfxf74> Video 3 Rewarding yourself

Training Video 4-Windows Media.wmv
<http://www.youtube.com/watch?v=i9K0OKNgsGk> Video 4 Stating your intentions

Training Video 5-Windows Media.wmv
<http://www.youtube.com/watch?v=RzxlqWdamiI> Video 5 Bullet Points

Training Video 6-Windows Media.wmv
<http://www.youtube.com/watch?v=KafkDMk27qk> Video 6 Why is WM important?

Training Video 7-Windows Media.wmv <http://www.youtube.com/watch?v=bxgn-qpxH14> Video 7 Managing Distraction

Training Video 8-Windows Media.wmv
<http://www.youtube.com/watch?v=D3jDMvNylpk> Video 8 Memory Strategies

Training Video 9-Windows Media.wmv
<http://www.youtube.com/watch?v=7Bw38iQOQtA> Video 9 Making the most of training

Appendix F: Participant feedback form

It would be useful for us in the design and implementation of future studies to get your reflections on the brain training study you participated in. If you could fill in these few questions it would be great.

5. How much did you enjoy the training?

Not at all a lot

1 2 3 4 5 6 7 8 9 10

6. How do you think you benefited from the training?

not at all	a small bit	a fair bit	a lot
1	2	3	4

7. How easy did you find the training?

very difficult	difficult	moderate	easy	very easy
1	2	3	4	5

8. How useful/informative did you find the videos?

Not useful at all Very Useful

1 2 3 4 5 6 7 8 9 10

5. Would you recommend the training?

__________ Yes

__________ No

6. Did you notice any effects on your confidence, speed of thinking, concentration, memory, organisation or everyday life? If so, then how?

____

____

____

____

____

____

7. Reflections on the process. What would you change about the training, if anything?

____

____

____

____

____

____

____

____

____

____

Many thanks again for your help with this study,

All the best,

____

Sinéad Hynes