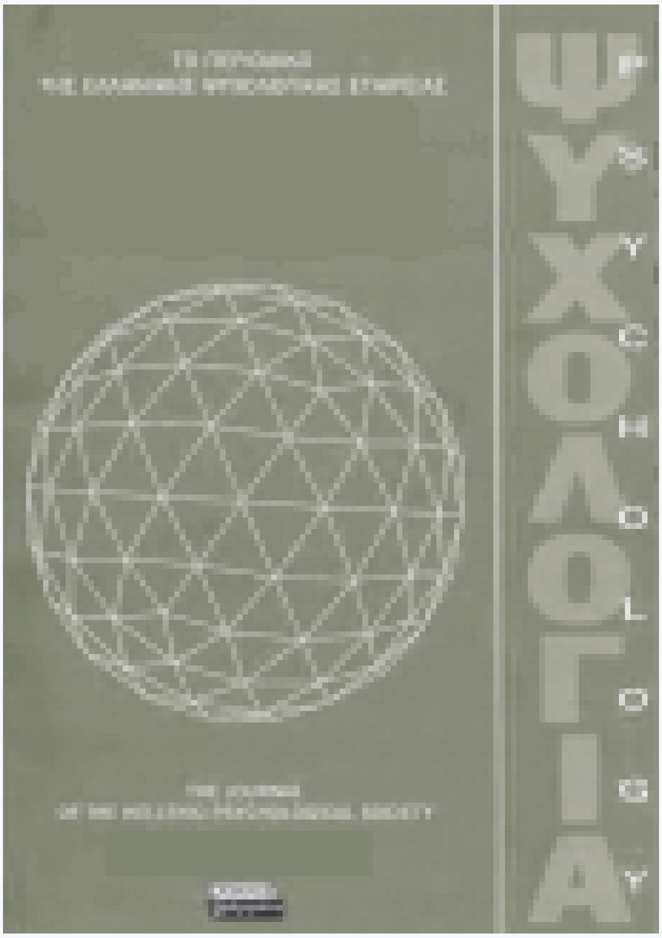
ePublishing

Psychology: the Journal of the Hellenic Psychological Society

Vol 14, No 2 (2007)



The investigation of planning on neuropsychology: from the Tower of Hanoi and Tower of London paradigms to Virtual Reality

Maria Kotitsa

doi: [10.12681/psy_hps.23859](https://doi.org/10.12681/psy_hps.23859)

Copyright © 2020, Maria Kotitsa

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0](https://creativecommons.org/licenses/by-sa/4.0/).

To cite this article:

Kotitsa, M. (2020). The investigation of planning on neuropsychology: from the Tower of Hanoi and Tower of London paradigms to Virtual Reality. *Psychology: The Journal of the Hellenic Psychological Society*, *14*(2), 188–208. https://doi.org/10.12681/psy_hps.23859

https://epublishing.ekt.gr | **e-Publisher:** EKT | **Downloaded at:** 24/12/2024 04:04:17

The investigation of planning in neuropsychology: from the Tower of Hanoi and Tower of London paradigms to Virtual Reality

MARIA KOTITSA[1]

ABSTRACT The aim of this work is to review selected tests employed in the investigation of high-level cognitive/executive functioning of planning and problem-solving abilities in patients with frontal lobe damage. Following a brief presentation of conventional planning tests, the review will focus on two strands of research, namely, the Tower of Hanoi/London tasks and more modern tests, such as the Six Elements Test (Wilson et al., 1996), as these two types of procedures formed the basis of two novel computer-based tasks reported herein: *the Bungalow Task and* the Warehouse Six Elements Test. For this set of computerised tasks, the patients are presented with complex and open-ended everyday life situations (a 'house removal' and a 'factory' scenario), as opposed to simple and well-defined problems that are characteristic of traditional tests. Importantly, these computerised tasks appear to be successful in capturing the rich diversity of impairments in patients following frontal lobe neurosurgery. The review concludes with the notion that computer-based environments -by taking experimentation out of the laboratory and closer to real life- may offer the best opportunity we currently have for investigating planning abilities in a controlled, yet ecologically valid fashion.

Key words: Planning, Frontal lobes, Virtual reality.

Introduction

Planning is an activity that individuals engage in regularly, since it is used in all facets of everyday life, from doing ones shopping to starting a business. The ubiquity of planning has impeded investigation, in that teasing apart the common components of the vast array of situations where planning is required has not been readily amenable to research. Hence, the precise operationalisation of planning may have been problematic due to overlapping constructs.

From a historical perspective, Luria (1966) viewed planning as the ability to organise behaviour in relation to a desired goal and was seen as part of problem solving ability, for which he identified three phases. The first is strategy selection, a preliminary stage where the hypotheses

1. *Address*: Dr Maria Kotitsa, Lecturer, Regent's College, Inner Circle, Regent's Park, London, NW1 4NS, UK. Tel./Fax. +44(0)20 8809 7378, E-mail: kotitsam@regents.ac.uk.

are created, and the strategies are selected in order to produce a solution to the problem. The second and third phases refer to the application of the operations and the evaluation of the outcome, respectively; these two phases are seen to be executive in nature. Friedman and Scholnick (1997) note that planning was initially seen as a distinct function and a general psychological process with various authors taking it to mean for example representation, or selection of a strategy, or strategy application. Subsequently, as these authors note, planning begun to be seen as the *orchestration* of a number of *inter*-dependent and diverse processes of a mechanism aimed at reaching a desired goal. Having as a starting point the notion that planning is not a unitary process or skill, experimental psychologists have divided planning into a series of processes, associating planning with problem solving (Anzai & Simon, 1979; Klahr, 1994; Simon, 1975; Zhang & Norman, 1994).

Several information-processing theories have sought to outline the set of activities that must be performed during problem solving (e.g. Miller, Galanter & Pribram, 1960; Pea & Hawkins, 1987). Friedman and Scholnick (1997) have identified a number of elements that information-processing theories share: problem representation, goal definition, strategy anticipation and execution (with a strategy monitoring mechanism taking place at the same time). Planning then involves the selection of a goal or a desired end state, where a desired state is more than a simple change in the current scheme of affairs. Furthermore, Scholnick and Friedman (1993) and Baker *et al.* (1996) defined planning as the accomplishment of a desired goal through a series of intermediate steps, incorporating the following aspects of executive functioning: (a) *problem representation*, where the problem is defined by comparing the current state with the desired goal state, (b) *goal selection*, (c) *decision* to plan, (d) *strategy choice*, (e) *strategy execution* and (f) *monitoring of effectiveness* of prior actions (*ibid.*); and also 'the use of knowledge for a purpose, the construction of an effective way to meet some future goal' (Scholnick & Friedman, 1993, p. 145). Essential requirements for an individual to engage in planning activity includes having both sufficient working memory and processing capacity, attention, sequencing ability and self-regulation (meaning inhibitory control) (Overton & Newman, 1982). In particular, working memory serves as the basis for the ability to perceive, integrate, and represent critical features of the environment (in particular in novel situations). By systematically varying the level of task complexity and hence manipulating the degree to which the various planning sub-processes are required, Kotovsky *et al.* (1985) and Zhang and Norman (1994) have demonstrated the facilitatory role of external aids in problem solving tasks; they concluded that planning places considerable demands on working memory capacity. Since task complexity impacts the demands on working memory, it was deemed necessary in the present experiments to keep working memory demands to a minimum. One way to do this is by cueing the participants to the requirements of the task (*i.e.* by making available to them a cue card with a summary of the task instructions).

Certain researchers have sought to broaden the planning literature by incorporating planning in familiar situations (e.g. Hayes-Roth & Hayes-Roth, 1979; Pea & Hawkins, 1987); to this end, domains such as errand planning was explored drawing upon theories of cognition that apart from processing skills were taking into consideration the person's knowledge base as well (Chi *et al.*, 1982; Hammond, 1990; Scholnick, 1995). Hence, a new dimension was added, in that instead of viewing planning as merely the result of processing skills it was taken to indicate a person's knowledge about a related domain, and moreover, about a *prior planning* ability (Nurmi, 1991). This may partly explain the interesting dissociation that seems to exist between planning in unfamiliar or novel and familiar situations: in encountering a novel problem solving setting, people often find that information and ideas contained in the long-term memory store is of little benefit, due to the lack of correspondence between the to-be-solved problem and past problems or successful attempts. In contrast, when faced with a familiar situation, individuals are most likely to rely on their knowledge base to retrieve material to produce a solution, or even already existing solutions. Consequently, attempting to solve a novel problem will require the development of a new solution; poor planning is expected in *unfamiliar* situations in cases where new solutions fail to develop, and in *familiar*, where peoples' knowledge store may not be sufficiently large or maybe faulty.

The current paper aims to review a series of tasks used in the neuropsychological study of planning and problem solving, starting with simple, classical laboratory procedures and followed by more modern tests (naturalistic, laboratory-based and some also making use of computerized technology). In so doing this paper seeks to demonstrate that –with the incorporation of real-life mimicking tasks in its repertoire– the investigation of complex planning is possible, also in a more 'ecologically' valid fashion through the use of virtual reality (VR). To this end, two VR procedures have been developed and are reported herein; but before we turn to those, it is worth reviewing some classical tests of planning and problem solving and in particular the *Tower of Hanoi or London* (ToH/ToL) paradigm. The reason for the emphasis placed on this paradigm is because it served as the main laboratory predecessor of the VR tasks. In the first VR task for example, the Bungalow, the planning component, the inclusion of task constraints, as well as the idea of goal-subgoal conflict and certain measures have all been inspired by the Tower of Hanoi/London paradigm. More recent influences on the VR tasks include the *Multiple Errands Task* by Shallice and Burgess (1991) and the *BADS* battery (or the Behavioural Assessment of the Dysexecutive Syndrome) (Wilson *et al.*, 1996), procedures designed to possess higher levels of ecological validity than earlier (laboratory) tests, and in so doing, to more readily capture the complexities of real life planning and problem solving. Indeed, the VR tasks have been developed in the same spirit.

Laboratory-based planning neuropsychological procedures

Much of the research on planning to date has focused on the cognitive functions that allow plan formulation. From a historic perspective, the work of Penfield and Evans in the '30s offers a striking clinical observation pinpointing impairments in what they called 'planned initiative' as the most important symptom of frontal lobe excisions. These researchers described the condition of a young woman a few months following a right frontal lobe resection. The patient, who was in fact Penfield's sister, was unable to plan and cook a family meal but despite this, was perfectly able of preparing the individual dishes (Penfield & Evans, 1935). However, maybe the first experimental demonstrations of a planning deficit in patients after prefrontal excisions came from a study by Porteus and Kepner (1944) and found that patients with prefrontal lobotomies showed a defective problem solving ability. Although this finding was later replicated (e.g. Mettler, 1952), it was also shown that a decrement in maze-learning ability was not restricted to frontal damage (Walsh, 1978). Measures of fluency are amongst those tests characteristic of planning and executive function. In the *verbal* fluency test (Benton, 1968) the participants are required to name as many words as they can think of beginning with the letter F, then A, and lastly with the letter S, allowing a predetermined time period for each letter (sixty seconds). Benton (1968) and Borkowski *et al.* (1967) found that the patients with left, and more so bilateral frontal damage, showed poor performance on the verbal fluency task, and this was in line with results by Ramier and Hécaen (1970), confirming a left frontal deficit. In the *design* versions of the FAS test, Jones-Gotman and Milner (1977) and Ruff and his colleagues (1987) asked their participants to create as many different, unnamable designs, as possible in a given time. Whilst, patients with left frontal lobe excisions have been frequently shown to be impaired on the *verbal* fluency test, impaired performance of patients with right frontal or fronto-central damage has been reported on the non-verbal, i.e., the design fluency tests (e.g. Milner, 1964; Jones-Gotman & Milner, 1977; Benton, 1968; Perret, 1974).

Another early procedure was the Wisconsin Card Sorting Test (WCST), a problem-solving task devised by Grant and Berg in 1948. In this task the patient is presented with four different key cards: the first showing a red triangle, the second two green stars, the third three yellow crosses, and the last one four blue circles. The patient is then given two stacks containing 128 cards with the aforementioned colours, shapes and symbols, the task being to match the cards into piles of the same colour, shape and symbol. The sorting of the cards to one of the four key cards is done according to a rule, which is not known to the patient. After doing this for a while the rule is changed by the experimenter, and the patient has to stop sorting, for example by colour, and sort by symbol. The experimenter changes the rule after ten consecutive responses by the patient, with each response being followed by feedback as to whether this response is correct or not (*i.e.* matches the rule or not). Patients with frontal lobe lesions appear to have problems both adapting to the new rules and modifying their responses according to the feedback given to them. Hence, they make significantly more perseverative errors, ignoring the feedback that their response was incorrect (Milner, 1964).

With regards to more recent procedures Bechara and his associates (1994) designed the Iowa Gambling Task, which was the first attempt to investigate in the laboratory and quantify the decision-making impairments in patients with frontal lesions. The test -presented in the form of a card game- was intended to mimic decision-making in situations featuring uncertainty of premises and outcomes, in addition to reward and punishment. Bechara and colleagues (1994) administered the Iowa Gambling Task to patients with medial orbitofrontal damage and to healthy controls. Participants were shown four decks of cards (A-D), and were instructed to select a card from one deck at a time. Only after turning over each card, are they informed of the amount of money they win or lose: the task's objective is to maximise the profit. Specifically, certain decks not known to participants, decks A and B, contain high paying cards ($100 gain each card). However, accompanying these large, immediate monetary rewards is the probability of large (e.g. $1250), future monetary penalties. Decks C and D are characterised as low paying, because they give small ($50 each card) immediate monetary rewards. They are, however, more advantageous compared to the first ones, as they also give small penalties, hence, leading to overall profit in the long run. The task finishes after one hundred selections. According to the probabilistic contingencies involved in the task, the correct responses cannot be easily computed; participants must rely on their 'instinct' to guide behaviour, and this is an important aspect of the task. The normal controls were found to develop a preference for the advantageous, low reward and low risk decks after about forty selections, with this preference being the result of experience from the task and also from taking into account the feedback from the tester. In sharp contrast, the patients with damage to the medial orbitofrontal region seemed to be unable to learn that selecting from decks A and B is less likely to maximise overall winnings, and so switch to the less risky decisions. The results suggested a diminished ability to assess future consequences, or as the authors put it patients '*are oblivious to the future consequences of their actions, and seem to be guided by immediate prospects only*' (1994, p. 7).

The Tower of Hanoi or London paradigms

The Tower of Hanoi (ToH) is a classic planning task that requires a strategic approach, in that planners must construct a strategy (a sequence of actions) in order to get from the initial to the final goal state. The apparatus consists of a board with three identical pegs, where a series of rings or discs are arranged. These discs differ in size; the task requires rearranging the discs in order to reach a certain goal state. However, in transferring the rings certain rules must not be violated, for example, participants can move only one ring at a time and a ring can only be placed on a peg on which it will be the smallest. The *means-end analysis* heuristic (Simon, 1975; Zhang & Norman, 1994) is particularly useful in solving the ToH, and similar problems: it requires the generation of a number of subgoals aimed at removing any obstacles that prevent the final goal from being reached. For example, one strategy, termed '*hill-climbing*', refers to the progressive nearing of the final goal by attaining a number of subgoals, each of which closes the distance between final goal and subgoals (Simon, 1975) describes additional strategies, that differ in their complexity and the demands they place on working memory). To achieve the goal, it is necessary to generate a *sequence of subgoals* and follow them through in a recursive, not linear fashion. However, as was pointed out by Shallice, the level of problem difficulty could not be easily manipulated in this task, rendering the ToH a rather 'unsatisfactory test for psychometric purposes' (Shallice, 1988, p. 347). Consequently, Shallice and McCarthy (see Shallice, 1982) developed another test, the Tower of London (ToL), in which difficulty could be graded more easily, and planning and 'look ahead' skills can be assessed in a more valid manner.

Due to the prominence of the Tower paradigm in the study of problem solving in patients with brain damage, but more importantly due to its relevance to the development of the VR tasks (to be described later), what follows is an in-depth review of experiments using the ToH/ToL. Also, as a guide for the identification of links between planning and brain areas, we will draw upon two approaches: lesion studies with patients with frontal lobe damage and neuroimaging data from healthy volunteers (even though this should not be mistaken for the notion that these two methodologies are interchangeable, as no direct concordance can be assumed to exist with certainty between neuropsychological and brain imaging findings).

I. Neuropsychological Studies on the ToH/ToL

Shallice (1982) tested patients with anterior and posterior brain lesions on the ToL: the patient is presented with three different coloured beads (threaded into sticks) and is required to use a stated minimum number of moves in order to arrange the beads so as to match a model arrangement. The number of moves required to solve the puzzle ranges from two to five. The results showed that patients with left anterior lesions were more impaired than patients with right anterior damage and left or right posterior damage, with impairment being reflected in the number of moves needed to solve the problems (Shallice, 1982) –even though these findings were not replicated in a later study by the same investigator, a possible reason for this relating to sample characteristics (Shallice, 1988). Studies by Owen and his colleagues sought to follow up the original Shallice (1982) results (Owen *et al.*, 1990). These researchers used the computerised adaptation of the ToL (see Morris *et al.*, 1988), which comes in two formats; the first involves a simple matching-to-sample strategy, concentrating on the generation of the correct motor sequence, and thus necessitating only minimal cognitive planning (or 'forward thinking' according to Owen, 2005). This version requires two moves to solve the puzzle, *i.e.* to produce the pre-specified arrangement. However, in the more complex version more moves are needed; thinking and mental imagery operations are required, and not infrequently, participants need to engage in visually *counter-intuitive* moves (see later).

In the Owen (1990) study, 'planning time', the time between the presentation of the task and the first move, was found to be unaffected in a sample of twenty six patients with frontal lesions, while the 'subsequent thinking time' (after they had started the task) was significantly prolonged compared to healthy controls; the patients were less efficient in generating solutions, although no laterality effect was found (Owen *et al.*, 1990). Owen and his colleagues also compared the frontal lobe patients with a group of amygdalohippocampectomy patients as well as patients with medial temporal lobe damage (Owen *et al.*, 1995). Again, the findings indicated a selective impairment in the patients with frontal lesions in terms of spending significantly more time completing the ToL task, once started on the task. This finding was replicated in a study by Carlin *et al.* (2000), who also tested patients with frontal lobe lesions on the ToL and found that, although they showed similar planning times to their matched controls, their solution times were significantly slowed; additionally, patients also required an increased number of moves. Furthermore, these researchers also administered the ToH task and found strong correlations between the two tasks, this indicating that similar planning abilities were being tapped (Carlin *et al.* 2000). Morris's studies also consistently report poor performance in their focal frontal patients on the Tower tasks (e.g. Morris *et al.*, 1997a; b – see next two sections for these studies). Finally, Colvin *et al.* (2001) tested patients with circumscribed frontal lesions on the ToH and also on the procedurally similar Water Jug task, a computerised task involving three jars with large, medium and small capacity, and an initial and a goal state. In the initial state the large jar is full of water; participants must discover a series of moves that would lead to the large and medium jar having specific amounts of water (goal state). No constraints are applied in pouring between jars and, like the Tower problems, this task also involved a hill-climbing strategy and counter-intuitive moves. Sixteen of the twenty-seven patients were also given the ToH in order to assess the relationship between performance on the two tasks. Colvin *et al.* (2001) found no significant correlations between measures of the tasks, but a selective impairment emerged on the Water Jug task, with patients with left-sided or bilateral lesions being more impaired than patients with right-sided frontal lobe damage.

II. Neuroimaging Studies on the ToH/ToL

With respect to neuroimaging data associating the frontal lobes with planning, two early studies using Single Photon Emission Computed Tomography, or SPECT, reported increased frontal brain metabolism during performance on the ToL task (Morris *et al.*, 1993; Rezai *et al.*, 1993). In particular, in the Morris *et al.* study the activation increased in participants that found the task difficult and took longer to complete the task. However, the SPECT technique has poor spatial resolution. Thus, in a later study using PET Owen *et al.* (1996) sought to explore specific frontal areas associated with tackling easy or difficult ToL problems. Participants were given a control task and the comparison of the difficult problem to the control condition revealed activation in the mid-dorsolateral frontal region of the left hemisphere and part of the caudate nucleus; when activations observed during the simple planning condition were subtracted from the activation observed in more difficult condition the increases in regional cerebral blood flow referred to the caudate nucleus and the thalamus only. The authors also reported activity in the right hemisphere, although, as they note, it was not significant by standard criteria (Owen *et al.*, 1996). The Owen results were more or less replicated in another PET study conducted by Baker *et al.* (1996). Task administration differed from the previous study, in that it did not require the execution of the solution itself, as participants are to perform this action mentally, *i.e.*, find the minimum number of moves required to solve the problem and indicate this by pressing the key that corresponds to this number, *without* carrying out the moves. Clear dorsolateral activations were observed bilaterally and activity in BA 10 in the right hemisphere, as well as to premotor, cingulate, parietal and occipital areas. Moreover, it was noted that activity was more robust in participants that found the task demanding and took longer to complete it (Baker *et al.*, 1996).

Further studies investigated brain areas associated with task complexity, identifying a large network, including prefrontal areas, the anterior cingulate, posterior parietal areas and parts of the caudate nucleus (Dagher *et al.*, 1999 [PET]; Asloun *et al.*, 2001; Van der Heuvel *et al.*, 2003 [fMRI]). Schall *et al.* (2003) combined PET and fMRI data and also administered the task using the Baker *et al.* (1996) technique. They reported task-difficulty-dependent increases in the cerebellum, left dorsolateral cortex, premotor cortex, cingulate, precuneus and structures in the basal ganglia. In addition, using fMRI and paradigm similar to Baker *et al.* (1996), Cazalis *et al.* (2003) tested participants on problems of varying solution difficulty, 'control', 'easy', and 'difficult', and categorised participants into superior and standard performers; they found that superior performers showed larger activation in the left dorsolateral cortex than standard performers, who conversely, showed a tendency to have greater activations in the anterior cingulate than superior performers. The authors suggest that successful problem solving invokes specific patterns of brain activity implicating the dorsolateral cortex and the anterior cingulate (Cazalis *et al.*, 2003). Another PET study sought to pinpoint brain areas subserving processes related to specific components of the ToL task (Rowe *et al.*, 2001). In this study, participants in the 'plan' conditions were required to plan the best solution, while in the 'control' conditions they only had to either produce a given number of moves (without having a goal state), or imagine executing these moves. Under both these conditions the same pattern of activations was observed, involving the dorsolateral convexity, the premotor and parietal cortex, and also the cerebellum. Contrasting the activity associated with the plan task to that of the control tasks, showed no residual pre-frontal activity, leading the authors to suggest that dorsolateral activity can be attributed to generating, selecting and/or remembering mental moves (Rowe *et al.*, 2001).

Further confirming evidence on the critical role particularly of the prefrontal cortex is provided by Lazeron *et al.* (2000), who in an fMRI study found activation in the dorso-lateral prefrontal cortex, in addition to non-dorsolateral areas, such as the cingulate, insula, cuneus and precuneus, and parietal regions. Prefrontal specificity has been suggested in a study that sought to investigate the planning and visuospatial components of ToL under three conditions, featuring 'easy', 'moderate', and 'difficult' problems (Newmann *et al.*, 2003). The results showed that the prefrontal cortex in both hemispheres were equally engaged when solving moderate and difficult problems, also that right prefrontal activity was correlated with individual differences in working memory and additionally, the left and right prefrontal cortices were found to have different functional connectivity. On the basis of these results, the authors proposed a differential involvement of right and left prefrontal areas, with the former playing a larger role in plan generation, and the latter in plan execution; in terms of parietal regions, they suggested that right parietal areas are more implicated in attention and the left in visuospatial processes (Newmann *et al.*, 2003).

Taken together, neuroimaging studies consistently report increased activation in specific frontal regions whilst solving the ToL puzzle. Whether the task requires participants to mentally solve the problem (e.g. Baker *et al.*, 1996; Schall *et al.*, 2003; Newmann *et al.*, 2003) or execute the moves (e.g. Owen *et al.*, 1996; Dagher *et al.*, 1999; Rowe *et al.*, 2001), the region that seems to be consistently associated with solving the more difficult problems in particular is the mid-dorsolateral frontal cortex (BA 9/46), although by no means are these dorsolateral areas exclusively engaged. Bearing in mind that patients with frontal damage can show marked planning deficits on this task, it could be suggested that the prefrontal activity observed in healthy individuals while performing on the ToL is related to planning operations taking place.

III. The Goal-Subgoal Conflict

Reference was made in the previous section to the need to engage in counter-intuitive moves; this applies to situations whereby adopting a 'clear', gradual goal reduction approach does not lead to the solution. The Tower problems appear to be particularly suitable for the study of goal-subgoal conflict situations, which typically arise from having to carry out actions that seem to take one *away* from the desired situation/solution. Morris and colleagues (1997a) investigated the effects of a goal-subgoal conflict on planning ability following brain damage. In particular, they assessed two groups of patients, one with frontal and another with temporal lobe lesions, using the ToH problem. In this task, a goal-subgoal conflict is evident when participants have to move the disc away from the final position. The researchers predicted that patients with frontal lobe lesions would have difficulties in showing the correct, but counterintuitive moves, compared to the healthy and also brain damaged control participants. To measure accurate performance, the number of moves above the minimum necessary to solve the problem was recorded. The results demonstrated a significant deficit in patients with left frontal and right temporal damage in terms of inhibiting a pre-potent response; this impairment applied to the four-move problems. However, this impairment in the left-lesioned frontal group dissipated (whilst it generalised in the patients with right temporal lesion) for the five-move version of the problem. Morris *et al.* (1997a) went on to administer a visuospatial memory test to examine the hypothesis that memory deficits were responsible for the observed patterns of deficits. This was found to be true of the temporal lobe group. Regarding the left frontal lobe patients, it was speculated that the novelty implicated in having to solve a goal-subgoal conflict may account for the impairment occurring in these patients (*ibid.*). Hence, solving a goal-subgoal conflict appears to present different brain damaged individuals with different problems.

Another related study on goal-subgoal conflict is that by Goel and Grafman (1995), who called into question the ability of ToH task to evaluate planning ability on the grounds that its requirement for inhibition of pre-potent responses may contaminate the outcome measure. Goel and Grafman argue that the difficulties experienced by the patients on the task are weakly, if at all, related to the planning or 'look ahead' impairments, and instead they propose that the deficits may be due to goal-subgoal conflict resolution difficulties. The authors concluded that 'the point is not that frontal patients do not have planning deficits, but simply that, the Tower of Hanoi tasks does not warrant any conclusions about planning abilities' (1995, p. 640). (Although this is an interesting point that the researchers make, this may not be entirely true, as they did not specifically instruct participants to plan prior to commencing the task.) The conceptualization of the task and the interpretation of the results aside, (whether they are interpreted in terms of planning problems or difficulties resolving goal-subgoal conflicts) the fact of the matter is that patients with frontal lesions have been consistently found to perform poorly on Tower type tasks, thus supporting the prefrontal involvement in the task and helping establish the task's validity.

The need for real-life planning tasks

Concerns were voiced over two decades ago by Lezak (1982) regarding the nature of the traditional 'frontal' tests, which invariably feature well-defined goals and provide clear initiation prompts, all this resulting in extremely well structured situations, but which bear no resemblance to the world that the test-taker lives, functions and has problems in, and subsequently, the world that the test ought to represent. That is, *if* the test aims to provide a reliable index of a patient's frontal difficulties. Indeed, a landmark study in this regard is that of Eslinger and Damasio (1985) reporting the patient EVR. Apart from demonstrating of course the debilitating impact of planning difficulties, this is an exemplar study calling for more ecological validity in testing procedures. This patient had undergone neurosurgery for the removal of a bilateral orbitofrontal tumour; he had a premorbid IQ of over 130 and when tested on a long standard neuropsychological battery, including measures typically sensitive to frontal lobe damage, his performance was above average. The fact that this patient was highly competent in the cognitive domain came in sharp contrast to his gross impairments in organizing daily life activities (such as going out to dinner, or to the cinema). He had also enormous difficulties in maintaining a balanced personal and working life, all this culminating in his job loss and inability to remain employed or keep his marriage. Following this case study there were other similar descriptions of patients with apparent daily life organisational problems but who showed normal performance on certain frontal tests (Anderson *et al.*, 1991; Shallice & Burgess, 1991). Shallice and Burgess (1991) give an account of three individuals with frontal lobe damage who were not able to function well in everyday situations, requiring for example, shopping or working within time constraints, but interestingly they performed normally on tests considered to be sensitive to frontal dysfunction.

Hence, almost ten years later, the same problems as pointed out by Lezak in 1982 continued to plague both the research and assessment of frontal lobe functions; Shallice and Burgess (1991) noted that standard tests required that task trials were very short, one single problem was tapped at one time, successful completion was well defined and also that task initiation was heavily prompted by the tester –this latter point was in agreement with what Burgess and Alderman had also pointed out, that patients were deprived of the opportunity to initiate activity, monitor their performance and/or modify their behaviour accordingly (Burgess & Alderman, 1990). In addition, evidence suggested that such tests may also be sensitive to damage to areas outside the frontal cortex (Teuber, 1964; Anderson *et al.*, 1991; Reitan and Wolfson, 1994), and so the sensitivity and specificity of a range of standard 'frontal' tests (such as the ToL, the Stroop, the WCST) have been at times called into question (Reitan & Wolfson, 1994, 1995; Phillips, 1997). The widely used Tower of Hanoi/London task and variants increasingly receives criticisms concerning its suitability for assessing frontal lobe syndromes and dysfunction, the criticisms converging on the same issue, *i.e.* that it does not represent daily life situations: see Shallice and Burgess, (1991), Burgess *et al.* (1998); more recently Morris *et al.* (2005) observed that the 'mental activity involved is somewhat removed from planning in the real world' and to this effect, Phillips and her colleagues (2005) also noted that one should be particularly cautious in making inferences concerning real-life planning skills on the basis of performance on the ToL, precisely because the realistic context in absent. Therefore, the development of ecologically valid tests became central in planning research. Ecological validity is a psychometric property that tells us how good a test is in providing accurate information concerning the participant's ability to function in everyday situations (Pramuka & McCue, 2000).

The beginning of a new era in the design of frontal lobe and planning tests is marked by the development of certain 'frontal' tests, which are as creative as they are promising: the Multiple Errands Test (Shallice & Burgess 1991) and the

Six Elements Test (the latter contained in the Behavioural Assessment of the Dysexecutive Syndrome, or BADS, battery [Wilson *et al.*, 1996]). A much-cited case study is the Goldstein *et al.* (1993) study, which describes a patient with marked planning difficulties following left frontal lobectomy. Not only because of this important finding, but this study is also notable in terms of the test used to assess the patient. These researchers implemented one of the tests that emerged at that period, involving testing in a natural environment, with the results demonstrating the capacity of such procedures to invoke everyday planning and organisational deficits; hence, the Multiple Errands is discussed next.

'Multiple Errands' and BADS

The Multiple Errands Test (Shallice & Burgess, 1991) was designed to simulate real-life activities. Conducted in an unfamiliar pedestrian area, it involved completing eight tasks, whilst following a number of pre-learned rules. In terms of the tasks, six of them made simple demands (for example, to buy a brown loaf, a packet of throat pastilles). The seventh task was a little more demanding, in that it required the participant to stay at a specific place for 15 minutes after starting. The last task was even more complex, as it required participants to write four pieces of information on a postcard; these were: (a) to give the name of the shop in the street expected to sell the most expensive item, (b) to write the price of a pound of tomatoes, (c) to write the name of the coldest place in Britain the day before, and (d) to give the rate of the exchange of the French franc the day before. The postcard with this information was then to be sent to one of the experimenters. In terms of the task constraints, these involved: (a) spending as little money and taking as little time as possible; (b) not entering a shop unless to buy something; (c) when leaving a shop telling one of the experimenters accompanying the participant what was bought there; and (d) not using anything not bought on the street to assist with the conduction of the experiment.

Although the patients performed well on standard frontal, the ToL, and intelligence tests they failed, however, to develop and follow an efficient plan in order to carry out the list of errands. The patients made more errors and their behaviour was also qualitatively different (for example, walking out of a shop without paying) compared to the controls. Shallice and Burgess classified the patients' behaviour into: rule breaks, inefficient strategy, task interpretation errors, and task failures, with patients performing worse than the controls. The advantages of the Multiple Errands test include high ecological validity and the tackling of non-routine settings, imitating real, day-to-day aspects of planning, organisation, and initiative. Although the test has proved successful in detecting frontal lobe impairment (e.g. Shallice & Burgess, 1991; Goldstein *et al.*, 1993; Worthington *et al.*, 1999), the disadvantages underlying its use is mainly its lack of practicability, as it entails substantial amounts of time and planning on the part of the researcher as well as the presence of two research assistants to conduct the experiment and record the patient's movements; so its obvious disadvantage of being difficult to standardise makes it rather unlikely to succeed as a widely used paradigm, and this has been the case. It has been validated, however, for use with individuals of below average intellect (Aitken *et al.*, 1993), while recent work also explored the utility of a simplified version of the Multiple Errands Test within a hospital setting with very promising results (see Knight *et al.*, 2002).

With regard to the BADS battery, it consists of a questionnaire and six tasks, namely: (a) the 'Rule Shift Cards Test' (a test consisting of 21 ordinary playing cards and examining the patient's ability to respond correctly to a rule and shift from one rule to another); (b) the 'Action Program Test' (a novel problem-solving task, involving a number of apparatuses and requiring the participant to remove the cork from a tall tube with certain restraints); (c) the 'Key Search Test' (which participants are presented with a large square drawn on a piece of paper, the task being to imagine this is a field and draw the path they would take in search of some lost keys); (d) the 'Temporal Judgment Test' (consisting of four questions and assessing the patients' ability to make judgments about time); (e) the 'The Zoo Map Test', where participants are presented with a map of a zoo, the task being to plan a route around the zoo, by drawing a line on the map, in order to visit a number of pre-specified locations. However, apart from having to show the order of each visit, participants are requested to follow certain rules, namely to start at the entrance and finish at the picnic area; additionally, rules apply relating to using certain 'paths' in the zoo more than once. The test has been devised so as only four different routes can be produced without breaking the rules; and (f) the 'Modified Six Elements Test'. This requires participants to plan their work in order to carry out a number of tasks in a limited period of ten minutes, whilst following a specific rule. Participants are told to attempt at least something from each task. The tasks are organised in three groups, 1, 2, and 3, each of which is further subdivided in two parts, A and B, producing six tasks in total. The rule is that part A of one task should not be attempted immediately after part B of the same task, and likewise, part B should not come after part A of a particular task. The SET reflects a multiple subgoals scenario, taxing the ability to plan, organise and monitor behaviour (Burgess, 1997).

An early pilot study, comparing patients with brain injury to controls (Alderman *et al.*, 1993), indicated that the BADS was a better predictor of daily life problems than the modified WSCT, with the Modified SET and the Zoo Map Test being the two most sensitive subtests in the battery. Further validation, using a larger sample of 216 controls and 90 people with brain injury revealed a good inter-rater and test-retest reliability (Wilson *et al.*, 1996). It appears that this battery successfully addresses ecological validity issues and also the question of patients underestimating their functional difficulties due to reduced or lack of insight (e.g. Stuss, 1991; Prigatano, 1991; Alderman *et al.*, 2001), though there seems to be also some evidence of the BADS not being sensitive. For example, a recent study by McGeorge *et al.* (2001) employed the BADS battery to assess executive difficulties in patients with head injury. The authors found that the BADS did not detect impairment in these patients, although they had documented daily life planning problems and performed significantly worse than controls on a novel planning task designed by McGeorge and colleagues.

Additional 'ecological' approaches to planning

Another early study attempting to bridge the gap between laboratory assessment and real-life planning skills was carried out Boyd and Sautter (1985; 1993). They looked at executive aspects of route-finding, such as, task formulation, strategy, or error identification and correction, investigated in a natural environment. In particular, thirty-one patients with traumatic brain injury were required to start at a specified location and find the route, which would get them to a specified office on an unfamiliar campus. Two members of staff were with the patient while performing the task, the examiner and an independent rater to enable the investigation of the test reliability aspect of the test, which was found to be very high (Boyd & Sautter, 1993). However, the number of trained assistants needed to carry out such a test, poses obstacles with regards to wider use and/or standardization.

Miotto and Morris (1998) tested twenty-five patients with focal unilateral or bilateral frontal lesions on a procedure designed to imitate the planning and organisation of naturalistic and relatively simple and familiar daily activities. The 'Virtual Planning' test had the form of a verbal board game; participants were instructed to carry out a number of actions within a week by selecting a series of cards, which represented the different actions. The cards were placed on a board in front of the participant, and a summary card was also supplied to reduce demands on memory. This task sought to address the following aspects of planning associated with frontal lobe disorder: (a) whether frontal patients tend to fail to do activities if the purpose of the activities are made more remote. For this purpose two types of irrelevant activities were inserted (distractor items related to a 'trip' and to the 'week'): the introduction of irrelevant actions is an element which in the main was absent from the classical 'frontal' tests, thus taking the test one step closer to real-life situations; (b) the second dimension of the test tapped into prospective memory ability following frontal lobe damage, hypothesizing that patients will show impairment in remembering to do certain activities at a certain day or time; (c) finally, the VIP test addressed the question of whether the complexity of an activity affected performance, the prediction being that patients would show a more fragmented pattern of behaviour. A significant deficit was revealed in the patients as a heterogeneous group. With regard to the specific measures, the control group were prone to selecting distractor items related to the 'trip' as opposed to 'week' related ones, while the patients were equally prone to both types of irrelevant actions. The patients did not differ from the controls in their ability to perform prospective memory tasks, contrary to previous findings linking prospective memory deficits with the prefrontal cortex. The authors attributed the lack of difference between patients and controls to the fact that the participant is strongly cued by the cue card as to when the activities have to be performed. In terms of the third dimension, it was found that the more complex actions generated more errors than the less complex ones.

Channon and Crawford (1999) investigated social problem-solving using a story-or video-based task, for which they presented their participants with stories and videotapes depicting a range of interpersonal awkward situations –for example, in one scenario the protagonist had problems with his newly moved neighbours, who happened to own dogs barking during the night and making it impossible to have a proper night's sleep. The participants were tested on their ability to generate different solutions and solve the problem by selecting the one that was effective and socially appropriate. The results demonstrated that the patients with frontal lesions were impaired in their ability to produce possible solutions compared to patients with lesions in the posterior area, suggesting a sensitivity of this procedure to frontal lobe damage. Finally, classical are the ecological approaches to planning by Goel and his associates (Goel *et al.*, 1997; Goel & Grafman, 2000): using the realistic scenario of household finance they assessed planning abilities of ten frontal lobe patients by requiring them to 'help' a young couple achieve a number of immediate and long-term goals, such as, buy a house or send their children to college (Goel *et al.*, 1997). The task was designed so that in order to achieve these goals, participants had to use strategies pertaining to management of income and expenses, and/or reallocation of assets. Patients were found to be unable to discover and apply these strategies, showing more organisation difficulties than the controls. Moreover, in a case study by Goel and Grafman (2000), a patient with frontal lobe damage also exhibited more difficulties than the matched control, when tested on a real life planning and organisation task. In particular, this patient (a professional architect) was assigned the task of re-arranging a real office area so as to meet certain criteria, e.g. be functional, accommodate a number of staff and use minimal resources, and as noted was significantly impaired.

Virtual Reality: bridging the gap between laboratory and real world tests

It appears from this review[2] that researchers have faced a dilemma as regards measuring complex abilities, such as planning, in that conventional tests though high in experimental control, tend to lack ecological validity. A solution to this was to test participants in real life (e.g. Multiple Errands), but the main advantage of this approach seems to be in terms of triggering the development of new tests, presented in the form of board games (e.g. BADS), rather than contributing to the systematic investigation of planning. Yet these new tests -though they may fare better compared to ToH/ToL- precisely because of the game like quality, may be less ecologically valid than their real world precursors. The answer to the issue of ecological validity then may come from the field of Virtual Reality (VR). VR is a rapidly growing new field, which plays an increasingly important role in psychology, both as an assessment and rehabilitation tool (see Kotitsa [submitted] for a review). Indeed, Zalla *et al.* (2001) made use of a computerised task to test seven patients with frontal lobe lesions. Participants were required to form and execute all the actions needed in a familiar daily life situation (from getting up in the morning to leaving the house), with the results indicating impairments in the frontal lobe patients in plan execution.

More recently, Kotitsa (2005) tested thirty-five neurosurgical patients on a set of experimental VR procedures modeling naturalistic types of behaviour in order to explore the effects of discrete frontal lesions on executive functioning. These procedures (introduced in detail below) were developed by Morris and his colleagues (2005) using existing software -previously used by the Virtual Reality Group at the University of East London, UK, to investigate aspects of memory and training people with learning difficulties or following brain damage (Brooks *et al.*, 1999; Rose *et al.*, 1999)- in order to provide the environment for the presentation of the paradigms. These novel procedures were designed so as to be in keeping with preserving the realistic context of planning (recently Phillips *et al.* [2005] also emphasized this issue); to incorporate the element of novelty in tapping executive functioning (as pointed out by theorists, e.g. Stuss & Benson, 1986); and to be in agreement with the information-processing approach favouring novel and open-ended procedures as more likely to capture frontal lobe deficits. More specifically:

(a) *Maintenance of the realistic context of planning*: the requirements of the tasks (to be described next) are infused a number of contingencies similar to those found in everyday life. For example, in order to more closely mimic real life planning activity, participants are to work together with other two (imaginary) persons and generally, in their imagination, their planning behaviour will have an impact on other people.

(b) *Encompassing appropriate levels of novelty and familiarity*: whilst participants are required to tackle a novel situation, which will necessitate the generation and application of a non-obvious strategy plan encompassing the various task requirements, at the same time, however, the familiarity factor was also considered important to highlight: this is reflected in the general nature of the tasks (the tasks involving entering and spending time in an average house or a warehouse, and engaging in familiar activities: moving from room to room to find certain furniture and placing objects onto to specified locations). A task with relative familiarity, apart from being inherently ecologically valid, is also very likely to be interesting; this in turn may elicit higher levels of co-operation and

2. For a comprehensive review on the cognitive psychology of planning, see Morris and Ward (2005).

motivation, hence optimum performance is more likely to be attained.

(c) *Featuring open-ended multitasking scenarios:* it is in these situations that established routine operations for successful performance may not be readily available, this demanding the function of the hypothesised SAS processes (Norman & Shallice, 1986; Shallice & Burgess, 1991).

(d) *Allowing the investigation of different facets of executive functioning:* executive functions feature a large and diverse set of abilities, and the tasks herein described seeking to probe *aspects of executive functioning*, such as rule following, strategy formation and various types of prospective memory [or memory for future actions, Einstein & McDaniel (1990)].

While the above comprise the general principles that *conceptually* characterize these tasks, in terms of the procedures, the Bungalow Task features a virtual house, a bungalow, which consists of four rooms and a hallway. A number of items and furniture are distributed throughout the rooms in a semi-random fashion. The participants are informed that they are the 'removal person', their task being to collect all furniture and items (with the exception of specific, glass items) for the arrival of the removal van (in practice, participants would touch the screen upon which furniture would disappear, indicating successful removal). The owners of the bungalow are moving to an eight-room house and have left *specific instructions* as to how the removal is to be carried out. These instructions serve to generate a number of rules or requirements that must be followed (a cue card summarises these rules and is available to the participant throughout the task). The first rule refers to the fact that the furniture must be collected according to which room in the new house they are going. Specifically, the furniture that could go to the new lounge must be collected first, followed by furniture for the dining room, then the nursery, the kitchen, the study room, the music room, the bedroom and finally the new hallway. This first requirement gives rise measures of rule following ability and it also allows the investigation of spatial strategy formation, since –to be able to select the different categories of furniture in order– participants must develop certain strategies of going around the four rooms. The second requirement asks the participant to place 'fragile' labels on the glass items, which should be left behind for the owners to collect. This requirement enables the measurement of event-based prospective memory (forgetting to place the label) and behavioural inhibition (in the case where a glass item was erroneously collected). According to the third requirement, which measures activity-based prospective memory, participants have to close the door of the bedroom (so that the cat does not go there) every time they go out of this room and in addition they must close the front door every time they open it (again, so that the cat does not run out). According to the fourth requirement, looking at time-based prospective memory, since the doorbell is out of order, participants have to open the front door every five minutes and check whether the removal van has arrived to load the furniture. A clock is provided for them to monitor the time, as well as the time at which they ought to be checking for the van: should they fail to visit the front door within a two minutes of the set time, they are reminded to do so.

The second task, termed the Warehouse Six Elements Task (WSET), features an industrial environment, consisting of one room, in which a number of objects, cylinders and boxes, are stored on four stacks. Prior to the task, participants undergo a familiarization phase, whereby they learn to perform three main tasks: task 1 asks them to place objects onto a trolley that is in the room; task 2 requires them to place objects on a conveyor belt which takes them to a different room; and for task 3 they learn to weigh objects and then put them on the trolley (these activities are achieved by pointing to the item on the screen and then pointing to the desired location). In line with the Six Elements Test (SET) by Shallice and Burgess (1991) described earlier, each of these three tasks is further divided into two parts –one referring to the collection of boxes, and the other to the collection of cylinders– and the participants should not attempt, for any task, one part immediately after the other. A time limit of ten minutes applies, while the participants cannot stop before the time is up. These aspects of the task allow the exploration of rule-following ability in a multi-tasking situation. Imbedded in the task is also the requirement to collect the same amount of objects for each task, and the participants' performance in respect to this was quantified to yield a measure of their ability to allocate effort evenly across tasks.

Perhaps one particular building block of the Bungalow and WSET tasks, namely the all so important planning element is, to some extent, reminiscent of the ToH/ToL paradigm, the latter having proved very influential, precisely because it highlighted a particular method for investigating planning, *i.e.* by incorporating specific constraints which would make planning the key to solving it successfully. The ToH/ToL test contributed to the development of the above tasks, more so of the Bungalow Task, by lending the notion of a planning or 'look-ahead' component; in the Bungalow this is induced by the predetermined order in which categories of furniture need to be selected. Additionally, in an analogous manner that planning competency on the Tower problems is inferred by the number of (extra) moves required to achieve the goal state (with the fewer moves reflecting better planning), in the Bungalow too one general measure of planning refers to the (total) number of room visits required to complete the task. Finally, in terms of the goal-subgoal conflict proposed to account for the planning impairments on the Tower tasks, here too such a conflict may arise between selecting one specific category of furniture, e.g., for the lounge (main goal) and selecting other, non-lounge items or furniture that the participants encounter. For example, selecting for the dining room, or the kitchen could be said to represent (sub)goals, from which the participants should refrain whilst in the lounge category (refraining involves inhibiting the pre-potent response of collecting items as one goes along, irrespective of the room order). Clearer and more intuitive similarities exist between the Bungalow on the one hand, and the Zoo Map and Multiple Errands, on the other: the latter tasks, it will be remembered, require forming plans and setting goals, and generating and applying organized sequences of going to various places either in a board game format (Zoo Map) or in real-time shopping (Multiple Errands). In terms of the WSET, as its name also implies, it is primarily founded on the Modified SET, described in the preceding section and thought to reflect the prototypical multitasking situation (e.g. Burgess, 2000).

The results from the Bungalow Task measures showed that the combined patient group (including patients with right, left and bilateral frontal excisions) was significantly more impaired than controls: the patients showed significantly more *violations* of the first task requirement (selecting categories of furniture in turn, starting with the lounge furniture and finishing with the hallway); in addition, whilst they showed the same task completion time as the controls, however they visited the bungalow rooms significantly fewer times, suggesting a *poor strategy* in going around the rooms. In order to comply with the first task requirement the participants must devise a pattern or strategy of visiting the four rooms, and repeat this until they finish collecting the furniture. Performance data from the control group indicated the existence of two main strategies: the first is to visit room 1, then room 2 and then rooms 3 and 4, while the second strategy would be to go to room 1, then to room 4, and then to rooms 3 and 2. The absence of an organized strategy for moving from room to room characterized the performance of the patients here. With regards to prospective memory, the two groups did not differ on the activity-based measure (remembering to close specific doors), however there were deficits on the other two measures: in particular, the patients showed a deficit at trend level on event-based prospective memory (remembering to label fragile items) and a highly significant impairment on the *time-based* task (visiting the front door at specified intervals). Furthermore, the examination of possible contribution of memory (as measured by the Logical Memory subtest of the Wechsler Memory Scale Revised [WMS-R] [Wechsler, 1987]) indicated a particular pattern of results: co-varying for memory did not alter the results in terms of the rooms visited, strategy formation and time-based prospective memory, though it did remove a trend for impairment on the event-based task, pointing to the contribution of retrospective memory to remembering future intentions; in terms of variables that could not use parametric tests, correlations occurred in the patients only, in that poor memory was associated with poor performance on the activity-based task, and also with increased rule breaking behaviour. Similarly, the Warehouse Task also elicited significant impairments in the patients: here, the patients with right frontal lesions committed significantly more *rule violations* than the controls, while the patients' worse performance on the *evenness* measure (suggesting to allocate effort or resources evenly across a number of tasks) was not related to hemispheric side of the lesion. Again, in order to explore the potential effect of memory, following analysis of covariance the group difference on evenness remained robust, while non-parametric correlations revealed that poor memory was associated with increased number of rule violations in the patients only. It is also interesting to note that the patients were tested on specifically designed paper-and-pen prospective memory tasks, but also on well-known conventional, and ecologically valid procedures, namely the SET and Zoo Map Test. The performance of the patients with frontal lobe lesions on these tasks, however, did not differ significantly from the normal controls.

Conclusion

The experimental tasks used for the investigation of planning and problem-solving skills in patients with frontal lobe damage vary from simple tasks performed in the laboratory to more multi-tasking, real-life simulating procedures. Although the ToL/ToH paradigm remains a good way of furthering understanding, however, there has been a strong move in recent years towards more ecologically valid tests. Indeed, the studies reviewed above support the idea that facets of planning ability may be successfully explored through the use of VR tests, designed to invoke processes analogous to the ones required in everyday life situations. Clearly, more empirical work is needed so that VR-based techniques can be rigorously evaluated, but – by combining experimental control and ecological validity – the potential exists for these techniques to at least complement standard research testing and become a viable means for investigating higher order abilities.

References

Aitken, S., Chase, S., McCue, M., & Ratcliff, G. (1993). An American adaptation of the Multiple Errands test: Assessment of executive abilities in everyday life. *Archives of Clinical Neuropsychology, 8*(3), 212.

Alderman, N., Evans, J. J., Burgess, P., & Wilson, B. A. (1993). Behavioral-Assessment of the Dysexecutive Syndrome. *Journal of Clinical and Experimental Neuropsychology, 15*(1), 69-70.

Alderman, N., Dawson, K., Rutterford, N. A., & Reynolds, P. J. (2001). A comparison of the validity of self-report measures amongst people with acquired brain injury: A preliminary study of the usefulness of EuroQol-5D. *Neuropsychological Rehabilitation, 11*(5), 529-537.

Anderson, S. W., Damasio, H., Jones, R. D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology, 13*, 909-922.

Anzai, Y. & Simon, H. (1979). The theory of learning by doing. *Psychological Review, 94*, 192-210.

Asloun, S., Cazalis, F., Granon, S., Bittoun, J., Torot, R., Robbins, T., Rogers, R., Azouvi, P., & Burnod, Y. (2001). Functional MR Imaging study of Tower of London performance. *NeuroImage, 13*(6), 638.

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S. J., & Robbins, T. W. (1996). Neural systems engaged by planning: A PET study of the Tower of London task. *Neuropsychologia, 34*(6), 515-526.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 5*, 7-15.

Benton, A. L. (1968). Differential behavioural effects in frontal lobe disease. *Neuropsychologia, 6*, 53-60.

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5*, 135-140.

Boyd, T. M. & Sautter, S. W. (1985). Route-finding by the head-injured: qualitative analysis of executive control in an everyday task. Paper Presented at the Ninth Annual Postgraduate Course on Rehabilitation of the Brain Injured Adult and Child, Williamsburg, VA.

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.

Brooks, B. M., McNeil, J. E., Rose, F. D., Greenwood, R. J., Attree, E. A., & Leadbetter, A. G. (1999). Route learning in a case of amnesia: The efficacy of training in a virtual environment. *Neuropsychological Rehabilitation, 9*(1), 63-76.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function*, (pp. 81-116). England: Psychology Press.

Burgess, P. W. (2000). Strategy application disorder: the role of the frontal lobes in human multitasking. *Psychological Research, 63*, 279-288.

Burgess, P. W., & Alderman, N. (1990). Rehabilitation of dyscontrol syndromes following frontal lobe damage in a cognitive neuropsychological approach. In R. L. Wood & I. Fussey (Eds), *Cognitive rehabilitation in perspective*. Hove, UK: 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*(6), 547-558.

Carlin, D., Bonerba, J., Phipps, M., Alexander, G., Shapiro, M., & Grafman, J. (2000). Planning impairments in frontal lobe dementia and frontal lobe lesion patients. *Neuropsychologia, 38*(5), 655-665.

Cazalis, F., Valabregue, R., Pelegrini-Issac, M., Asloun, S., Robbins, T. W., & Granon, S. (2003). Individual differences in prefrontal cortical activation on the Tower of London planning task: implication for effortful processing. *European Journal of Neuroscience, 17*(10), 2219-2225.

Channon, S. & Crawford, S. (1999). Problem-solving in real-life-type situations: The effects of anterior and posterior lesions on performance. *Neuropsychologia, 37*, 757-770.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). *Expertise in problem solving. Advances in the psychology of intelligence* (pp. 7-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colvin, M. K., Dunbar, K., & Grafman, J. (2001). The effects of frontal lobe lesions on goal achievement in the water jug task. *Journal of Cognitive Neuroscience, 13*(8), 1129-1147.

Dagher, A., Owen, A. M., Boecker, H. & Brooks, D. J. (1999). Mapping the network for planning: A correlational PET activation study with the Tower of London task. *Brain, 122*, 1973-1987.

Einstein, G. O. & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 16*(4), 717-726.

Eslinger, P. J. & Damasio, A. R. (1985). Severe disturbances of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology, 35*, 1731-1741.

Friedman, S. L. & Scholnick, E. K. (1997). An evolving 'blueprint' for planning: Psychological requirements, task characteristics, and social-cultural influences. In S. L. Friedman & E. K. Scholnick (Eds), *The Developmental Psychology of Planning: Why, How, and When Do We Plan?* (pp. 3-22). NJ, London: Lawrence Erlbaum Associates.

Goel, V. & Grafman, J. (1995). Are the frontal lobes implicated in 'planning' functions: Interpreting the data from the Tower of Hanoi. *Neuropsychologia, 33*, 623-642.

Goel, V., Grafman, J., Tajik, J., Gana, S., & Danto, D. (1997). A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain, 120*, 1805-1822.

Goel, V. & Grafman, J. (2000). Role of the right prefrontal cortex in ill-structured planning. *Cognitive Neuropsychology, 17*(5), 415-436.

Goldstein, L. H., Berbard, S., Fenwick, P. B. C., Burgess, P. W., & McNeil, J. (1993). Unilateral frontal lobectomy can produce strategy application disorder. *Journal of Neurology Neurosurgery and Psychiatry, 56*(3), 274-276.

Grant, A. D. & Berg, E. A. (1948). A behavioural analysis of reinforcement and use of shifting to new responses in a Weigl-type card sorting. *Journal of Experimental Psychology, 38*, 404-411.

Hammond, K. J. (1990). Case-based planning – A framework for planning from experience. *Cognitive Science, 14*(3), 385-444.

Hayes-Roth, B. & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science, 3*, 275-310.

Jones-Gotman, M. & Milner, B. (1977). Design fluency: The invention of non-sense drawings after focal cortical lesions. *Neuropsychologia, 15*, 653-674.

Klahr, D. (1994). *Discovering the present by predicting the future. The development of future oriented processes* (pp. 177-218). Chicago: University of Chicago Press.

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

Kotitsa, M. (submitted). The use of Virtual Reality technologies in neuropsychological studies. *Vima ton Koinonikon Epistimon.*

Kotitsa, M. (2005). *Planning and organisational ability in patients with focal frontal lobe neurosurgery, investigated using Virtual Reality*, PhD thesis. Institute of Psychiatry, Kings College London, UK.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard – evidence from Tower of Hanoi. *Cognitive Psychology, 17*(2), 248-294.

Lazeron, R. H. C., Rombouts, S. A. R. B., Machielsen, W. C. M., Scheltens, P., Witter, M. P., Uylings, H. B. M., & Barkhof, F. (2000). Visualizing brain activation during planning: The Tower of London test adapted for functional MR imaging. *American Journal of Neuroradiology, 21*(8), 1407-1414.

Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology, 17*, 281-297.

Luria, A. R. (1966). *Human brain and psychological processes.* New York: Harper and Row.

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Della Sala, S., Milne, A. B., Hamilton, S., & Callender, J. S. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence-Teleoperators and Virtual Environments, 10*(4), 375-383.

Mettler, F. A. (1952). *Psychosurgical problems.* New York: Blakiston.

Miller, G. A., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behaviour.* New York: Holt, Rinehart and Winston.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren and K. Akert (Eds), *The Frontal Granular Cortex and Behaviour* (pp. 313-335). New York: NcGraw Hill.

Miotto, E. C., Bullock, P., Polkey, C. E., & Morris, R. G. (1996). Spatial working memory in patients with frontal lesions. *Cortex, 32*, 613-630.

Miotto, E. C. & Morris, R. G. (1998). Virtual planning in patients with frontal lobe lesions. *Cortex, 34*(5), 639-657.

Morris, R. G., & Ward, G. (2005). *The cognitive psychology of planning*. England: Psychology Press.

Morris, R. G., Ahmed, S., Syed, G. M., & Toone, B. K. (1993). Neural correlates of planning ability: frontal lone activation during the Tower of London test. *Neuropsychologia, 31*, 1367-1378.

Morris, R. G., Downes, J. J., Sahakian, B. J. Evenden, J. L., Heald, A., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 51*, 757-766.

Morris, R. G., Miotto, E. C., Feigenbaum, J. D., Bullock, P., & Polkey, C. E. (1997a). The effect of goal-subgoal conflict on planning ability after frontal and temporal lobe lesions in humans. *Neuropsychologia, 35*(3), 1147-1157.

Morris, R. G., Miotto, E. C., Feigenbaum, J. D., Bullock, P., & Polkey, C. E. (1997b). Planning ability after frontal and temporal lobe lesions in humans: The effects of selection equivocation and working memory load. *Cognitive Neuropsychology, 14*(7), 1007-1027.

Morris, R. G., Kotitsa, M., & Bramham, J. (2005). Planning in patients with focal brain damage: from simple to complex task performance. In R. G. Morris & G. Ward (Eds), *The cognitive psychology of planning* (pp. 153-180). England: Psychology Press.

Newman, S. D., Carpenter, P. A., Varma, S., & Just, M. A. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia, 41*, 1668-1682.

Norman, D. A. & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds) Cottsciousness and self-regulation New York: Plenum Press.

Nurmi, J. E. (1991). How do adolescents see their future - a review of the development of future orientation and planning. *Developmental Review, 11*(1), 1-59.

Overton, W. F. & Newman, J. L. (1982). Cognitive development: A competence activation/ utilization approach. In T. M. Field, A. Huston, H. C. Quay, L. Troll, and G. E. Finley (Eds), *Review of Human Development* (pp. 217-241). New York: Wiley.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28*(10), 1021-1034.

Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuospatial short-term recognition memory and learning after temporal-lobe excisions, frontal lobe excisions or amygdalohippocampectomy in man. *Neuropsychologia, 33*(1), 1-24.

Owen A. M., Doyon, J. Petrides, M. & Evans, A. C. (1996). Planning and spatial working memory examined with positron emission tomography. *European Journal of Neuroscience, 8*, 353-364.

Owen, A. M. (2005). Cognitive planning in humans: new insights from the Tower of London task. In R. G. Morris & G. Ward (Eds), *The cognitive psychology of planning* (pp. 135-152). England: Psychology Press.

Pea, R. D. & Hawkins, J. (1987). Planning in a chore-scheduling task. In S. L. Friedman, E. K. Scholnick, & R. R. Cocking (Eds), *Blueprints for thinking: The role of planning in cognitive development* (pp. 273-302). Cambridge: Cambridge University Press.

Penfield, W. & Evans, J. (1935). The frontal lobe in man: a clinical case study of maximum removals. *Brain, 58*, 115-123.

Perret, E. (1974). The left frontal lobe of man and the supression of habitual responses in verbal categorical behaviour. *Neuropsychologia, 12*, 323-330.

Petrides, M. & Milner, B. (1982). Deficits on subject-ordered tasks after frontal-and temporal lobe lesions in man. *Neuropsychologia, 20*(3), 249-262.

Phillips, L. H., MacLeod, M. S., & Kliegel, M. (2005). Age, the frontal lobes and executive functioning. In R. G. Morris & G. Ward (Eds), *The cognitive* *psychology of planning* (pp. 111-134). England: Psychology Press.

Phillips, L. H. (1997). Do 'Frontal Tests' measure executive function?: Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.). *Methodology of frontal and executive function* (pp. 191-213). England: Psychology Press.

Porteus, S. D. & Kepner, R. (1944). Mental changes after bilateral prefrontal lobotomy. *Genetic Psychology Monographs, 29*(4), 115.

Pramuka, M., & McCue, M. (2000). Assessment to rehabilitation: Communicating across the gap. In R. D. Vanderploeg (Ed.), *Clinician's guide to neuropsychological assessment* (pp. 337-355). New Jersey: Lawrence Erlbaum Associates.

Prigatano, G. P. (1991). Disturbances of self-awareness of deficit after traumatic brain injury. In G. P. Prigatano & D. L. Schacter (Eds), *Awareness of deficit after brain injury: Clinical and theoretical issues* (pp. 111-126). New York: Oxford University Press.

Ramier, A. M. & Hecaen, H. (1970). Role respectif des attentes frontales et de la lateralisation lesionnelle dans les deficits de la 'fluence verbale'. *Revue Neurologie, 123,* 17-22.

Reitan, R. M., & Wolfson, D. (1994). Practical approaches to puzzling problems in interpretation using the Halstead-Reitan Battery. *Workshop presented at the Annual Meeting of the National Academy of Neuropsychology,* Orlando.

Reitan, R. M., & Wolfson, D. (1995). Consistency of responses on retesting among head injured subjects in litigation versus head-injured subjects not in litigation. *Applied Neuropsychology, 2,* 67-71.

Rezai, K., Andreasen, N. C., Allinger, R., Cohen, G., Swayze, V., & O'Leary, D. S. (1993). The neuropsychology of the prefrontal cortex. *Archives of Neurology, 50,* 636-642.

Rose, F. D., Brooks, B. M., Attree, E. A., Parslow, D. M., Leadbetter, A. G., McNeil, J., Jayawardena, S., Greewood, R., & Potter, J. (1999). A preliminary investigation into the use of virtual environments in memory retraining of patients with vascular brain injury: Indications for future strategy? *Disability and Rehabilitation, 21,* 548-554.

Rowe, J. B., Owen, A. M., Johnsrude, I. S., & Passingham, R. E. (2001). Imaging the mental components of a planning task. *Neuropsychologia, 39*(3), 315-327.

Ruff, R. M., Light, R. H., & Evans, R. W. (1987). The Ruff Figural Fluency Test: A normative study with adults, *Developmental Neuropsychology, 3,* 37-52.

Schall, U., Johnston, P., Lagopoulos, J., Juptner, M., Jentzen, W., Thienel, R., Dittmann Balcar, A., Bender, S., & Ward, P. B. (2003). Functional brain maps of Tower of London performance: a positron emission tomography and functional magnetic resonance imaging study. *NeuroImage, 20*(2), 1154-1161.

Scholnick, E. K. & Friedman, S. L. (1993). Planning in context-developmental and situational considerations. *International Journal of Behavioral Development, 16*(2), 145-167.

Scholnick, E. K. (1995). Knowing and constructing plans. *SRCD Newsletter, 17,* 1-2.

Shallice, T. (1982). Specific impairments of planning. In D. E. Broadbent & L. Weiskrantz (Eds). *The neuropsychology of cognitive functions* (pp. 199-209). London: The Royal Society.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.

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

Simon, H. A. (1975). The functional equivalence of problem-solving skills. *Cognitive Psychology, 7,* 268-288.

Stuss, D. T. (1991). Disturbances of self-awareness after frontal system damage. In G. P. Prigatano & D. L. Schacter (Eds), *Awareness of deficits after brain injury: Clinical and theoretical issues* (pp. 63-83). Oxford: Oxford University Press.

Stuss D. T. & Benson, D. F. (1986). *The frontal lobes.* New York: Raven Press.

Teuber, H. L. (1964). The riddle of the frontal lobes in man. *The frontal granular cortex and behaviour.* New York: McGraw-Hill.

Van den Heuvel, O. A., Groenewegen, H. J., Barkhof, F., Lazeron, R. H. C., Van Dyck, R., & Veltman, D. J. (2003). Frontostriatal system in planning complexity: a parametric functional magnetic resonance version of Tower of London task. *NeuroImage, 18*(2), 367-374.

Walsh, K. W. (1978). *Neuropsychology: A clinical approach.* Edinburgh: Churchill Livingstone.

Wechsler, D. (1987). *Wechsler Memory Scale Revised.* New York: Harcourt Brace Janovich.

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

Worthington, A. (1999). Dysexecutive paramnesia: strategic retrieval deficits in retrospective and prospective remembering. *Neurocase, 5*, 47-57.

Worthington, R. A., Arumugam, T. V., Hansen, M. A., Balcar, V. J., Barden, J. A. (1999). *Identification and localisation of ATP P2X receptors in rat midbrain. Electrophoresis 20*, 2077-2080.

Zalla, T., Plassiart, C., Pillon B., Grafman, J., & Sirigu, A. (2001). Action planning in a virtual context after prefrontal cortex damage. *Neuropsychologia, 39*, 759-770.

Zhang, J. J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*(1), 87-122.