Brain (1997), **120,** 1805–1822

A study of the performance of patients with frontal lobe lesions in a financial planning task

Vinod Goel,[1] Jordan Grafman,[2] Jinous Tajik,[1] Sheldon Gana[1] and David Danto[1]

[1]*Department of Psychology, York University, North York, Ontario, Canada and* [2]*Cognitive Neuroscience Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland, USA*

Correspondence to: Vinod Goel, Department of Psychology, York University, North York, Ontario, Canada M3J 1P3 or Jordan Grafman, Cognitive Neuroscience Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892-1440, USA

Summary

It has long been argued that patients with lesions in the prefrontal cortex have difficulties in decision making and problem solving in real-world, ill-structured situations, particularly problem types involving planning and look-ahead components. Recently, several researchers have questioned our ability to capture and characterize these deficits adequately using just the standard neuropsychological test batteries, and have called for tests that reflect real-world task requirements more accurately. We present data from 10 patients with focal lesions to the prefrontal cortex and 10 normal control subjects engaged in a real-world financial planning task. We also introduce a theoretical framework and methodology developed in the cognitive science literature for quantifying and analysing the complex data generated by problem-solving tasks. Our findings indicate that patient performance is impoverished at a global level but not at the local level. Patients have difficulty in organizing and structuring their problem space. Once they begin problem solving, they have difficulty in allocating adequate effort to each problem-solving phase. Patients also have difficulty dealing with the fact that there are no right or wrong answers nor official termination points in real-world planning problems. They also find it problematic to generate their own feedback. They invariably terminate the session before the details are fleshed out and all the goals satisfied. Finally, patients do not take full advantage of the fact that constraints on real-world problems are negotiable. However, it is not necessary to postulate a 'planning' deficit. It is possible to understand the patients' difficulties in real world planning tasks in terms of the following four accepted deficits: inadequate access to 'structured event complexes', difficulty in generalizing from particulars, failure to shift between 'mental sets', and poor judgment regarding adequacy and completeness of a plan.

Keywords: planning; problem solving; prefrontal cortex; information processing theory

Abbreviations: WAIS = Wechsler Adult Intelligence Scale; WAIS-R = WAIS—revised; WCST = Wisconsin Card Sorting Test; WMS = Wechsler Memory Scale

Introduction

The neuropsychology literature contains many compelling observations by physicians and clinicians about social, emotional and cognitive consequences of lesions in the prefrontal cortex (Harlow, 1868; Penfield and Evans, 1935; Rylander, 1939). The following remarks made by Eslinger and Damasio (1985) about their patient E.V.R., a successful 35-year-old accountant who underwent an operation for the removal of a large orbitofrontal meningioma, are typical of many patients with frontal lobe lesions. After the operation E.V.R. tested in the 'above average' range on the Wechsler Adult Intelligence Scale (WAIS) and Wechsler Memory Scale (WMS). However, 'After a 3-month recovery period, he returned to accounting and bookkeeping . . . He soon became involved in a home-building partnership with a former coworker, a man of questionable reputation who had been fired from the company . . . The business failed and he had to declare bankruptcy . . . Thereafter, he drifted through several jobs. He worked as a warehouse laborer, as a building manager, and as an accountant . . . but was fired from each. Employers complained about tardiness and disorganization, although basic skills, manners, and temper were appropriate. Similar difficulties led to a deterioration of his marital life . . . Unable to hold a job and separated from his family, E.V.R. moved in with his parents . . . Employment problems

© *Oxford University Press 1997*

continued . . . He needed 2 hours to get ready for work in the morning, and some days were consumed entirely by shaving and hair-washing. Deciding where to dine might take hours . . . He would drive to each restaurant to see how busy it was . . . Purchasing small items required in-depth consideration of brands, prices, and the best method of payment. He clung to outdated and useless possessions, refusing to part with dead houseplants, old phone books, six broken fans . . .' (pp. 1731–2).

These observations speak of difficulties in judgement, decision-making and problem solving in real-world, open-ended situations. Often the problem solving involves planning and look-ahead components.

Over the decades, many laboratory tasks have been developed and administered to patients to understand the nature of the cognitive deficits that follow lesions to the prefrontal cortex more fully. Some studies have directly targeted complex cognitive tasks. For example, card sorting tasks (Milner, 1963), word similarity tasks, proverbs tasks (Rylander, 1939), and word definition tasks have been used to measure abstraction and generalization ability. Nonsense drawing tasks (Smith and Milner, 1988) and word generation tasks have been used to measure nonverbal and verbal fluency, respectively. Shell games have been used to measure rule/pattern induction (McCarthy and Warrington, 1990). Choice reaction time studies have been used to measure use of advance information (Alivisatos and Milner, 1989). The 'Tower of London' has been used to measure looking ahead/anticipatory abilities (Shallice, 1982) and cognitive estimation has been used to measure judgement (Shallice and Evans, 1978).

Other studies have been designed to evaluate processes that may modulate complex cognitive processes. For example, Stroop-type tasks have been used to measure selective attention (Perret, 1974), while boring/monotonous tasks have been used to measure sustained attention (Wilkins *et al*., 1987). Maze tracing has been used to measure instruction following (Corkin, 1965). Drawing tasks have been used to measure perseveration (Goldberg and Bilder, 1987). The 'A-not-B' task (Diamond, 1990) and the 'Antisaccade' task (Roberts *et al*., 1994) have been used to study inhibitory mechanisms.

While much has been learned from these neuropsychological tests, an increasing number of researchers are beginning to realize that there is a striking discontinuity in the level at which these tests engage the patient and the level at which the above compelling clinical observations emerge. We would like to raise three specific concerns.

First, it is not clear, at least superficially, that the types of tasks that have been administered to patients with frontal lobe lesions (card sorting, visual search, proverbs, estimation, etc.) entitle one to any conclusions about problem-solving abilities, let alone to differentiate between planning and nonplanning capacities, or between ill-defined and well-defined problem-solving capabilities. We are struck by the time scale and simplicity of the neuropsychology tasks. It is clear that they all engage high level cognitive processes, but it is not clear that they engage problem-solving capacities as generally understood in the cognitive literature (Newell and Simon, 1972).

Problem solving, by most definitions, requires at least the following conditions: (i) there be two distinct states of affairs; (ii) the agent is in one state and wants to be in the other state; (iii) it is not apparent to the agent how the gap is to be bridged; (iv) bridging the gap is a consciously guided (at least at the top executive level), multi-step process. Most of the tasks seem to fail condition (iii). For example, the object sorting tasks meet conditions (i), (ii), and (iv). There are two distinct states: a start state, the set of objects in front of the patient; and a goal state, the same set of objects grouped in some other way. The agent is in one state and wants to be in the other, and bridging the gap is a consciously guided multi-step process. However, it is not clear that the task satisfies condition (iii). It is only when the agent does not have at hand a single operator to bridge the gap that a problem space is instantiated to construct the sequence of operators that will affect the transformation.

Exceptions to this criticism are the 'Tower of Hanoi' and 'Tower of London' tasks (Shallice, 1982; Goel and Grafman, 1995), which do meet each of these criteria. But even here, in so far as these tasks are supposed to be paradigmatic cases of planning, the following difficulty presents itself. Planning problems require an agent to chart a path from A to B in some space, without 'bumping' into the world. All the 'bumping' must be done in some modelling space, and some satisfactory path extracted. That is, the whole idea of planning is that we want to know the consequences of an action before the action is executed. The only way to do this is to execute the plan in some modelling space and observe the consequences. If the results are satisfactory, we are prepared to execute the plan in the real-world. If the results are not satisfactory, we revise the plan until we believe it to be satisfactory. Goel and Grafman (1995) have argued that planning or 'looking ahead' is neither necessary nor sufficient to solve the Tower of Hanoi puzzle. It is not necessary because it is not required by any of the strategies used by the subjects. These strategies can be implemented in computer programs and such programs do not need to search several levels deep into the state space to determine the next move. Planning ability is not sufficient to solve the Tower of Hanoi task because you can look ahead all you like, but unless you see the 'trick', the counter-intuitive backward move, you will not solve the problem. Goel and Grafman (1995) have offered an alternative interpretation of the performance of patients with frontal lobe lesions in the Tower of Hanoi that appeals to working memory deficits and a failure to inhibit the prepotent global goal in favour of a conflicting local subgoal. They did not question whether the prefrontal cortex is implicated in planning functions (the clinical evidence is very compelling) but only claim that the Tower of Hanoi is not an ideal planning test.

A third concern is that most of the laboratory tasks are artificial, well-structured tasks, i.e. not the type of task that one is confronted with in the world at large. Real-world planning problems are invariably ill-structured problems. The ill-structured/well-structured distinction stems from an analysis of problem types by Reitman (1964). Reitman classified problems based on the distribution of information within a problem vector. [A problem vector is a tuple of the form (A, B, ⇒), where components A and B are the start and end states, respectively, and the component ⇒ is the transformation function.] Problems where the information content of the vector components is absent or incomplete are said to be ill-structured. For example, consider the following problem: design a toy airplane. The start state A is unspecified (e.g. what should it be made of? wood? cardboard, steel?, etc.). The goal state B is under specified (e.g. how large should it be? what colour? should it fly?, etc.). The transformation function ⇒ is also unspecified (e.g. how should it be made? by folding paper? cutting cardboard? stealing the design from a competitor?, etc.). Well-structured problems, on the other hand, are characterized by the presence of information in each of the components of the problem vector. The Tower of Hanoi problem provides a convenient example (Simon, 1975; Goel and Grafman, 1995). The start state is completely specified. There is a clearly defined test for the goal state. The transformation function is specified in advance and restricted to one of the following operations (depending on the strategy used): (i) move a disk, or (ii) move a pyramid.

Of course there are deep reasons (e.g. having to do with standardization, measurement, time constraints, etc.) for using well-structured, standardized tasks. But given that frontal lobe deficits are most apparent in social, real world situations, it seems at best odd to limit investigations to standardized tasks only. Furthermore, there are arguments in the cognitive science literature that suggest different computational mechanisms are required to deal with well-structured and ill-structured problems (Reitman, 1964; Chandrasekaran, 1987; Goel, 1995).

Our concerns are not unlike those recently voiced by others (Shallice and Burgess, 1991*b*; Bechara *et al*., 1994). Shallice and Burgess (1991) write that it is ‘necessary to develop quantifiable analogues of the open-ended multiple subgoal situations where this subset of frontal patients would theoretically have problems’ (p. 728). They go on to propose and administer a ‘Six Element Task’ that involves the pursuit of multiple goals over a 15-min period. Bechara *et al.* (1994) also complain about the lack of ‘a laboratory probe to detect and measure an impairment that is so obvious in its ecological niche’ and propose ‘an experimental neuropsychological task which simulates, in real time, personal real-life decision-making relative to the way it factors uncertainty of premises and outcomes, as well as reward and punishment’ (p. 8). They go on to describe a card game that has some of these elements.

We believe that both of these tasks take us in the right direction. The present study is conducted in a similar spirit and is designed to take us further along this road. We present data from 10 patients and 10 normal control subjects engaged in a real-world financial planning task. We also introduce a theoretical framework and methodology, developed in the cognitive science literature, for analysing the complex data generated by problem-solving tasks.

Method

Subjects

Ten male patients, ranging in age from 45 to 53 years, participated in the study. Eight of the patients were drawn from a Vietnam Head Injury population. Of the other two, one patient suffered a subarachnoid haemorrhage secondary to a right anterior communicating artery aneurysm, and the other patient had neurosurgical intervention to relieve pressure due to a right frontal intracerebral haemorrhage. The former patient was tested 1 year after surgery, the latter was tested 3 years after surgery. The eight Veterans came from similar socio-economic and educational backgrounds. [This determination was made on the basis that these patients were all drafted into the armed forces as enlisted men (not officers) during the Vietnam War. Many of these men came from middle-class or working-class families from the Southern United States.] They all received penetrating head injuries during their service in Vietnam in the late 1960s, and had been tested most recently between 1992 and 1994. Thus their etiology, injury dates, and recovery periods are similar. The sensory, motor and language functions of all patients were relatively intact, as determined by previous neurological and neuropsychological testing (*see* Table 1 for some language ability scores), and all patients seemed quite functional to casual observation. The experimental protocol was approved by the National Institute of Neurological Disorders and Stroke review board and all patients and control subjects gave informed consent.

The age, education and cognitive profiles of patients, along with the size and laterality of lesions (as determined by MRI) are noted in Table 1. The involvement of specific structures for eight of the patients, also determined from MRI (Damasio and Damasio, 1989), are specified in Fig. 1. These patients were matched for age and education with 10 normal volunteers.

Task

The task is taken from the domain of household finance. The rationale for choosing this domain is that (i) it is a real-life domain, (ii) it provides good examples of planning, (iii) and, while like all real-world tasks, it requires knowledge of the world, it does not require specialized knowledge. It is a domain with which we all have some familiarity, certainly by the time we are in our forties, as all of the patients in this study are. It is simply not possible to succeed in modern society without some such knowledge.

The problem involves helping a young couple balance

Table 1 *Characteristics of patients with frontal lobe lesions and normal control subjects*

	Normal subjects (n = 10)	Lesion in R hemisphere (n = 5)	Lesion in L hemisphere (n = 1)	Bilateral lesion(s) (n = 4)	All patients (n = 10)
Age (years)	43.50	49.67	44.00	45.00	47.23
Education (years)	15.21	15.40	15.00	12.50	14.10
WAIS-R (IQ)					
General	–	106.80	110.00	90.75	100.70
Verbal	–	110.80	109.00	91.00	102.70
Performance	–	99.60	111.00	90.25	97.00
WMS-R					
General	–	111.80	93.00	97.50	104.20
Verbal	–	120.75	96.00	97.75	107.78
Visual	–	105.00	92.00	98.25	100.56
WCST					
Categories	–	3.40	6.00	5.50	4.50
Perseveration	–	27.20	14.00	12.75	20.10
Picture arrangement	–	9.40	11.00	7.50	8.80
Word fluency	–	51.00	42.00	29.00	40.22
Tower of Hanoi	1323.00	787.69	1185.33	855.72	854.67
Volume loss (cc) (eight patients)	–	32.14	62.3	47.29	43.49

WAIS-R = Wechsler Adult Intelligence Scale—revised; WMS-R = Wechsler Memory Scale—revised; WCST = Wisconsin Card Sorting Test.

their budget; purchase a home within the next 2 years; send their children to college in 15–20 years; and have sufficient funds to retire in 35 years. The task is to help the couple achieve these four specific goals by various manipulations of income and expenditures and/or reallocation of assets. The couple's financial information was conveyed by way of an income statement and balance sheet. The financial information was such that subjects were required to manipulate income and expenses, and restructure assets and liabilities to achieve the four goals. The actual problem scenario is reproduced in Appendix A. It meets all of the criteria (*see* above) of an ill-structured problem.

All of our patients have worked for a living, saved money, bought or rented a house, raised children, and sent them to school, and they are approaching retirement. They understand and identify with the task.

Data collection

Subjects were brought into a testing room and presented with the task. They were given written instructions that explained both the experimental procedure and the task. Both of these sets of instructions are reproduced in Appendix A. The task was presented as a 'problem scenario'. Subjects were informed that they were to help a young couple get their household finances in order and achieve some goals. Subjects were asked to talk aloud as they proceed through the task. The sessions were videotaped. Subjects were warned not to try to explain what they were thinking but, rather just to vocalize the fragments of thoughts and ideas they might be attending to at that time. They were told that while time was not a critical factor, it was important to fully engage the task. They were also told that the information contained in the problem scenario was incomplete and they were encouraged to ask questions as necessary. The experimenter was present in the room and answered any questions the subject asked, but he/she did not initiate questions or discussion. Subjects had access to pen and paper, and financial calculation aids. The experimenter offered to do all financial calculations for subjects.

The task was administered in two parts. The first part required subjects to answer specific questions designed to familiarize them with the given information. If they were unable to find some piece of information, the experimenter pointed it out to them. This insured that when subjects began the second (planning) part, they were on an equal footing with respect to understanding the task and the given information.

Data analysis

The video recordings along with anything the subjects wrote constitute the raw data. They are clearly very complex data. In analysing them, we are interested in not only how well each group performs the task but, more importantly, how they perform the task (i.e. their cognitive strategies and processes).

The accepted way of analysing such data is with a methodology called 'protocol analysis' (Ericsson and Simon, 1984). While unfamiliar to many neuropsychologists, this is a quite common methodology in the cognitive science literature. It is tailored to such complex data and allows for both qualitative and quantitative analysis. It is not an exaggeration to say that most of our substantive results concerning complex human problem solving stem from a

Fig. 1 Location of lesions in eight patients, based on MRI.

combination of this methodology and information processing theory.

Information processing theory (Simon, 1961; Newell and Simon, 1972; Simon, 1978) offers an account of human problem solving in terms of an information processing system, a task environment and a problem space (Fig. 2). The information processing system is a computational characterization of the cognitive agent who has the problem. The task environment is the external environment, inclusive of the problem, of the information processing system. The problem space is an abstract logical construct defined by the characteristics of the information processing system, and more importantly, the task environment.

An information processing system is a physical symbol system (Newell, 1980) with a memory, a processor, sensory receptors and motor effectors. There are actually three separate memories: a long-term memory, a short-term memory and an external memory. Each is characterized by its organization and read/write times. The processor performs some basic elementary processes, e.g. read, write, test, compare, discriminate and replace symbols, but there is no necessary or sufficient set of processes. These elementary processes operate on one or two symbols at a time and are strictly sequential. They can be combined to carry out any arbitrarily complex computation. The main function of the receptors and effectors is to access external memory. Today there are many, more sophisticated, accounts of the structure of the information processing system (Anderson, 1983; Newell, 1990). However, the original Newell and Simon (1972) account provides a good first order approximation which is consistent with many of the recent, more specific characterizations.

The task environment consists of (i) a goal (the ‘desire’ to solve the problem), (ii) a problem and (iii) any other relevant factors. The motivation of the information processing system is assumed, and little is known about how general environmental factors affect cognitive processes. The emphasis has been on how the structure and content of a problem situation gets mapped onto the problem space. The assertion is that people are ‘severely stimulus-bound with respect to the particular representation they construct’ (Hayes and Simon, 1974, p. 197). They construct naive/transparent models based on surface features of the environment, unless they have some specialized knowledge which allows them to construct more sophisticated models. The models are assumed to be invariant across time and subjects.

The problem space is a modelling space where problem solving occurs as a computational process. It is defined by states, operators, evaluation functions and control strategies. States are symbolic representations of a problem at a point in time. Operators are the procedures which transform one state into the next state. Evaluation functions measure the

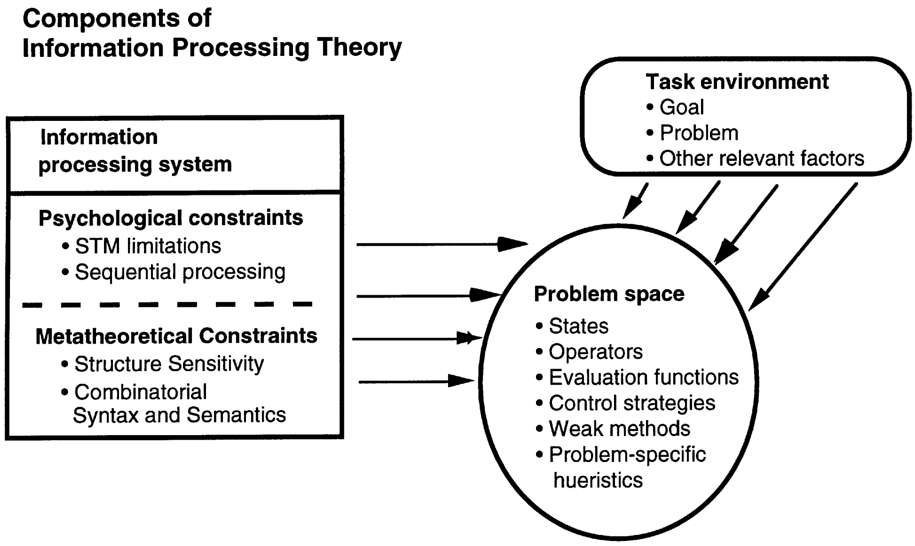

Fig. 2 Components of an information-processing system and their relationships, adapted from Goel (1995) with permission from MIT Press.

'goodness' of any current state and guide the search. Weak method search strategies (e.g. means-ends analysis, breadth-first search, depth-first search) are hardwired into the information processing system. But the strategy employed in any situation depends on the knowledge the system has explicitly available.

Information processing theory provides both a general framework and a specific vocabulary for discussing complex cognitive processes, and allows an interpretation of 'talk-aloud' verbalization data (Newell and Simon, 1972; Ericsson and Simon, 1984). The verbalizations (and written output) constitute the database and are interpreted against the model provided by the theory. Very crudely, the model indicates that the verbalization stream is an incomplete dump of the contents of short-term memory. On this assumption, the verbalization constitutes the state-space of the subject's problem space, and allows for the inference of the transformation functions or operators. Tracing the connections of sequences of states allows for the inference of control strategies and heuristic knowledge. Many more higher-level analyses and inferences are possible (Goel and Pirolli, 1992). The end result of the various analyses is an explication of the cognitive processes (states, operators, control strategies, etc.) engaged in by the subject during that problem-solving session.

Protocol analysis treats verbal data like any other behavioural data. At the top level subjects are generating noise/sound waves. We interpret these 'phonetic acts' as 'phatic acts', that is, as sentences of a natural language. Furthermore, we freely assign meaning to these sentences. All this is prior to any explicit analysis. We can do this freely because we believe the utterances to be meaningful and understand their meaning by virtue of belonging to the same linguistic community as the subjects. The explicit analysis and interpretation of the data begins with the transcription of the verbalizations. At this point there is some pretheoretical preprocessing that involves the filtering of facial gestures, hand waving, intonation, etc. The preprocessed transcribed text is correlated with the written material, and coded. The coding can be with *a priori* categories, or the understanding of the verbalization and formation of categories can occur in parallel. In either case the vocabulary/categories are given by a theory (Newell and Simon, 1972; Newell, 1990).

A common misconception of the methodology is to regard it as a form of introspection. Note, however, that in instructing subjects, one is not asking them to tell what they are doing or thinking. It is the theorist's job to figure this out. What one is asking subjects to do is to verbalize whatever thought contents they are attending to as they do the task. The accepted interpretation of these verbalizations, within information processing theory, is that they give us a trace of the contents of the subject's activated memory structures. From this trace the theorist infers the operations the subject applied to these contents, and the control strategies that guided the subjects through the problem space. A number of studies have indicated that the extra demands of such requested verbalizations may affect the speed of problem solving but otherwise do not affect performance (Ericsson and Simon, 1984).

Coding scheme

A three-level scheme developed by Goel and Pirolli (Goel and Pirolli, 1992; Goel, 1994, 1995) was used to code the data. Each level of the scheme is associated with a different vocabulary and granularity, and provides insight into a different aspect of the subject's cognitive process. A general overview appears in Fig. 3. A more detailed account follows in Fig. 4. The first step involves breaking the protocols into individual statements

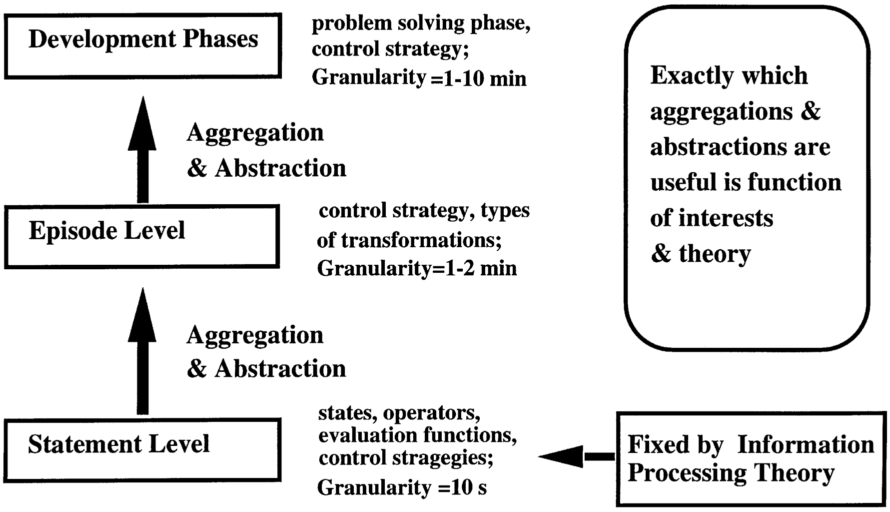

Fig. 3 Overview of coding scheme.

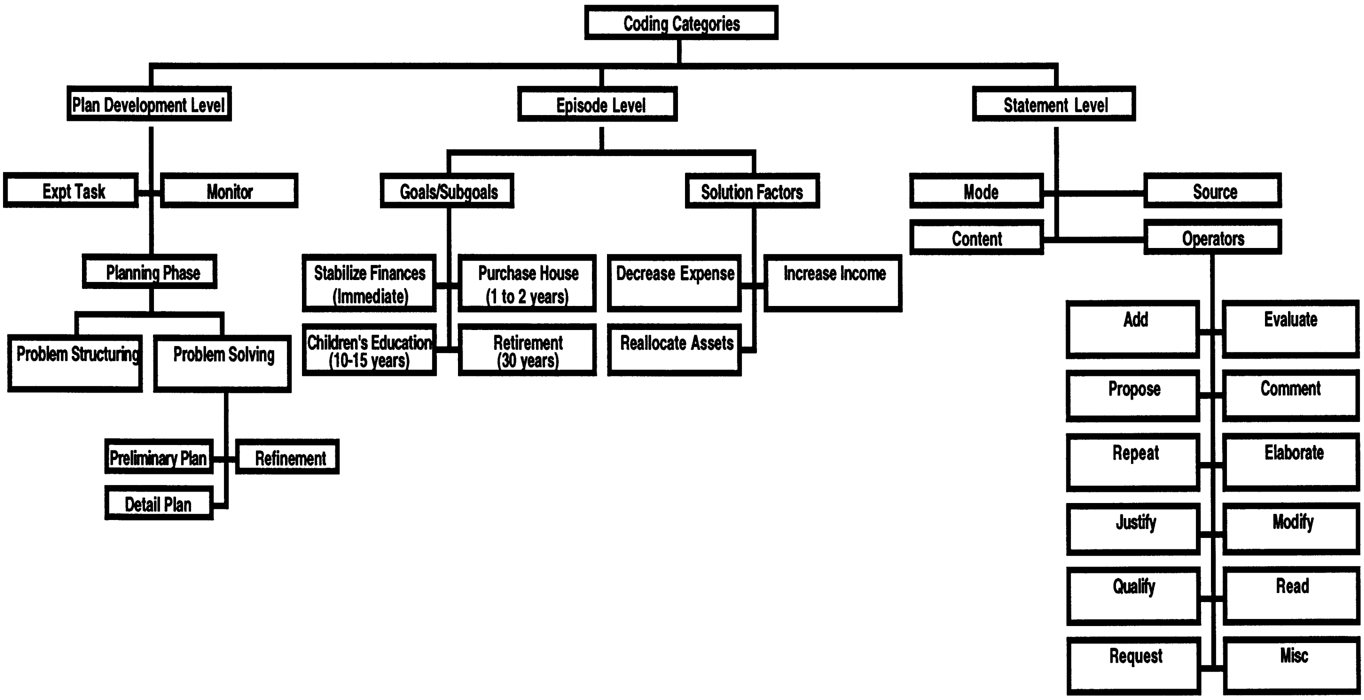

Fig. 4 Specific coding categories.

representing single 'thoughts' or ideas. Content cues, syntactic cues and pauses are used to effect this individuation. The vocabulary at this level consists of states, operators and evaluation functions, and comes straight from information processing theory. We take each statement as constituting a state in the subject's problem space and infer the operator applied to it. We found the following 12 operators adequate for our purposes: add, evaluate, propose, comment, repeat, elaborate, justify, modify, qualify, read, request and miscellaneous. This level of analysis gives a picture of cognitive processes at the granularity of a few seconds.

These statements are then aggregated into episodes, which are connected sequences of statements in the service of a common subgoal. In our financial planning task, episodes are organized around goals/subgoals and strategies. Four goals are given in the problem scenario: stem negative cash flow/ stabilize situation, purchase a house, send kids to college and save for retirement. Subjects generate a number of subgoals as they traverse the problem space. Typical subgoals are things such as 'reducing shelter expenses', 'repaying the car loan' and 'qualifying for a mortgage.' There are three types of strategies that subjects can utilize to achieve the goals/ subgoals: increasing income, decreasing expenditures and reallocating assests. Episodes typically have a duration of 1–2 min. The vocabulary at this level is one of goals, subgoals and strategies.

Table 2 *Sample of coded protocol from a normal control subject*[*]

Verbalization segmented into statements	Coding categories applied to statements
SUBJECT: They are gonna buy this house and pay $840 a month.	L1; plan-development: refine L2; strategy: general: house L3; write: self: [home and payments]: add
I'm gonna assume all these other things can't go down, . . .	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [assumption]: add
. . . because they don't seem outrageous.	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [justification]: justify
Food, $350 a month for food, . . .	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: d-bri: [food]: add
. . . for four people, . . .	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [number of people]: elaborate
. . . that doesn't sound like a lot.	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [opinion]: evaluate
. . .	
Clothing, $175 a month?	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [question for clothing]: repeat
That seems like a lot.	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [opinion]: evaluate
EXPERIMENTER: For three people?	E
SUBJECT: Oh come on now!	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [answer to exp]: evaluate
I haven't spent that much money per month.	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: [answer to exp]: comment
Well that's gonna have to change.	L1; plan-development: preliminary L2; strategy: decrease: clothing L3; verbal: self: [reduce funds for clothing]: add
I'm sorry.	L1; plan-development: preliminary L2; strategy: decrease: clothing L3; verbal: self: [apology]: comment
She should have a lot of baby stuff left from the first kid, . . .	L1; plan-development: preliminary L2; strategy: decrease: clothing L3; verbal: self: [sources of clothing]: elaborate

[*]The subject has just determined a monthly expenditure for a house purchase and is going back to consider some of the expenditures associated with food and clothing.

Episodes are further aggregated into a plan-development level. The plan-development level consists of experimental task statements, monitor statements and planning phase statements. The planning phase is further divided into problem-structuring and problem-solving phases. Problem structuring is a necessary prerequisite for the solution of ill-structured problems. It involves generating information missing from the problem scenario so that the problem space can be constructed. Once the problem space is specified, problem solving can begin. Problem solving is further differentiated into preliminary planning, refinement and detailing. Preliminary plan statements result in the initial generation and exploration of ideas. Refinement statements serve to elaborate and develop an idea. Detailing statements specify the final form of an idea. These phases typically have durations of 1–10 min, and are generally engaged sequentially, starting with preliminary planning, passing through refinement, and ending with detailing. However, it is common for subjects to return to an earlier phase as previously unnoticed aspects emerge.

The specifics of our coding categories are discussed more fully in Appendix B. Samples of coded protocols are presented in Tables 2 and 3. Table 2 presents a brief excerpt from a normal control subject and Table 3 presents an excerpt from a patient with a frontal lobe lesion. The verbalizations, broken into statements, appear in the left hand column. The three level, hierarchical coding scheme appears in the right hand column. While it is difficult to understand an arbitrary piece of verbal protocol extracted from its context, the samples give the reader an idea of

Table 3 *Sample of coded protocol from a patient with a frontal lobe lesion, who had difficulty with the task*[*]

Verbalization segmented into statements	Coding categories applied to statements
PATIENT: Shelter was the biggest one.	L1; plan-development: problem-structuring L2; strategy: general: housing L3; verbal: self: elaborate
EXPERIMENTER: Okay.	E
PATIENT: Now, if they eliminate that, the . . . that and . . . 10,800 they save that a year.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: propose
But yet . . .	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: evaluate
EXPERIMENTER: But you need a place to live.	E
PATIENT: Yes.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: experimenter: add
Course I know a place that sells tents cheap.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: add
You can buy one of those.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: add
EXPERIMENTER: Uh-huh.	E
PATIENT: But, no.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: modify
EXPERIMENTER: The . . .	E
PATIENT: Yeah. Or you move into a . . .	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: propose
No.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: modify
If I eliminate . . . (inaudible).	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: propose
That's clothing.	L1; plan-development: preliminary L2; strategy: decrease: housing L3; verbal: self: add
Hum. What?	L1; plan-development: preliminary L2; strategy: general L3; verbal: self: request
I really don't have any ideas.	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: comment
I don't have any ideas.	L1; plan-development: problem-structuring L2; strategy: general L3; verbal: self: repeat
EXPERIMENTER: Okay.	E

[*]The patient is trying to eliminate the expense for shelter.

the granularity of the statement level, and a sense of how the coding scheme was applied.

In the first statement of Table 2, the subject has just made some decisions regarding a house purchase and determined that it is going to cost $840 a month. This gets categorized as a problem-solving (refinement) statement; the goal being pursued is buying a house; the information is being added into the problem space; the source of the information is the subject; and the format of the information is verbal. Then the subject steps back and considers information given with respect to food and clothing expenditures. These are problem-structuring statements because the subject is just fleshing out given information. During this phase, the subject is not explicitly pursuing any of the four goals nor utilizing one of the three strategies, so these get coded as 'general'. Some of the subject's statements are adding information to the problem space, others are evaluating and justifying it. With the statement 'Well that's gonna have to change' the subject switches from problem structuring to problem solving. She begins to make some preliminary decisions or generate preliminary ideas. The strategy she is following now is one of decreasing expenses. She adds information to the problem space with the first statement, makes a comment with the second, and goes on to elaborate it with the third statement. Comparing the samples in Tables 2 and 3 gives a qualitative sense of the data which we briefly discuss in the Results section below.

The objectivity of the methodology lies in the fact that, once the categories have been developed and their recognition criteria explicated, different individuals can apply the categories to the data with similar results. The data were coded by three research assistants. The coders did not know the identity of patients and control subjects. A recoding of 10% of the data by the first author resulted in a 92% rate of agreement.

Results

Verbal protocols can be analysed at both qualitative and quantitative levels. Both analyses have merit. A qualitative analysis captures something of the contents of the protocols and is often used in case studies. A quantitative analysis captures the structure of the problem-solving process as revealed by the coding scheme. Because we have 10 patients and 10 control subjects we have chosen to do a quantitative analysis. But it is worth noting that there were some qualitative differences between patients and control subjects that are lost in a quantitative analysis.

For example, the normal control subject in Table 2 proceeds with confidence and brings her personal experience to bear on the task (e.g. ‘Clothing, $175 a month? That seems like a lot. Oh come on now, I haven’t spent that much money per month. Well that’s going to have to change . . .’). The patient, with a lesion in the prefrontal cortex, in Table 3, is much less confident and at a bit of a loss. He is trying to eliminate (as opposed to reduce) shelter expenses because that will save $10 800 per year. He makes the suggestion ‘I know a place that sells tents cheap. You can buy one of those.’ He does not seem to realize the oddity of the suggestion and has not thought out all of its implications. He repeats several times that he is not sure what to do. This admittedly is a case of a patient that had extreme difficulty with the task. It may be worthwhile to capture some of these differences in subsequent qualitative case studies.

We begin our analysis of the results by examining the duration of the protocols and the rate of verbalization. The mean problem-solving time for normal control subjects was 57.9 min versus 47.0 min for patients [$t(18) = 1.5$, $P = 0.14$]. The control subjects generated significantly more statements (mean 702.4) during their session than did patients (mean 437.6) [$t(18) = 2.5$, $P < 0.05$]. The patients generated 9.5 statements per minute while the control subjects generated 12.7 statements per minute [$t(18) = 1.6$, $P = 0.12$]. Furthermore, there were no significant changes in the rate of generation of statements throughout the problem-solving session in either group. To determine this we divided the

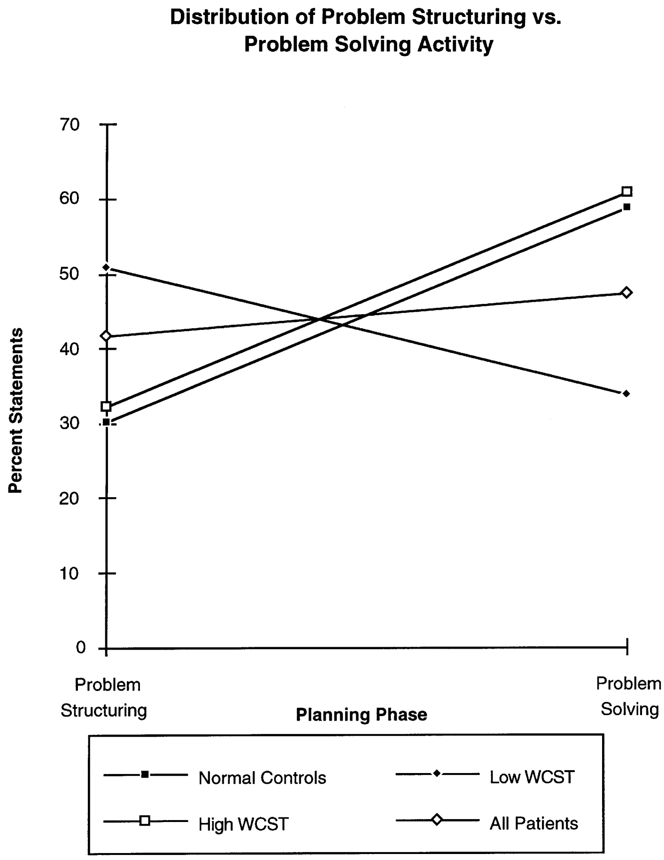

Fig. 5 Percentage of statements which control subjects and patients devoted to problem structuring and solving.

protocols of four randomly chosen patients and four control subjects into 5-min intervals and totalled the number of statements in each interval. Both patients and control subjects generated an equal number of statements during the first 5 min, the middle 5 min and the last 5 min of their protocol [$F(2,12) = 0.17$, $P = 0.84$].

The coding scheme analyses the data at three different levels of granularity: (i) plan-development level, (ii) episode level and (iii) statement level. There were significant differences in the problem-solving behaviour of patients and control subjects at the plan-development and episode levels. The behaviour of patients and control subjects was identical at the statement level.

At the plan-development level we found that control subjects and patients spent, respectively, 89.11% and 89.04% of the protocol statements on the planning phase (problem structuring and problem solving) as opposed to monitoring and experimental task issues. However, the distribution of statements between problem structuring and problem solving was quite different for patients and control subjects (Fig. 5). The patients used a significantly larger proportion of their statements (mean 41.6%) than control subjects (mean 30.2%) on problem structuring [$t(18) = 2.05$, $P = 0.05$]. The control subjects used a significantly larger proportion of their statements (mean 58.9%) than patients (mean 47.4%) on

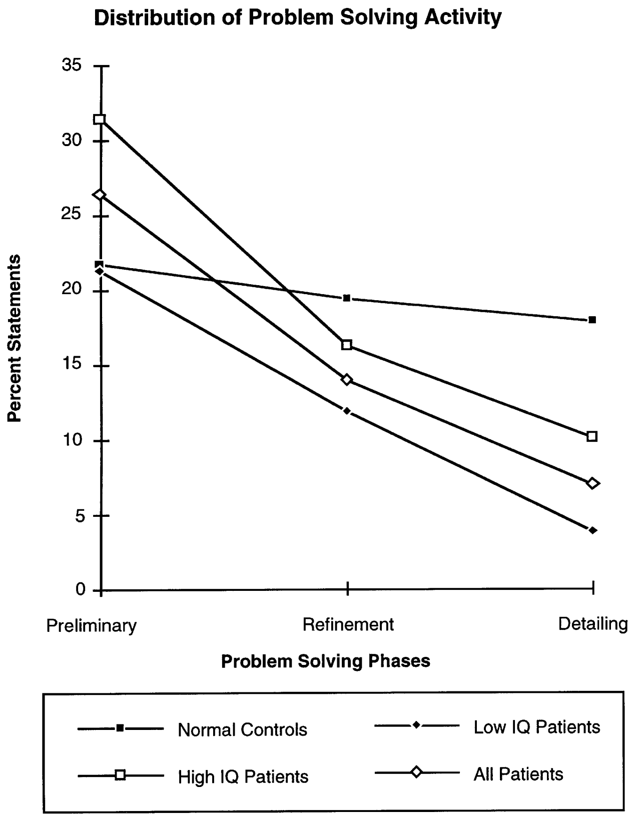

Fig. 6 Percentage of statements which control subjects and patients devote to each phase of problem solving.

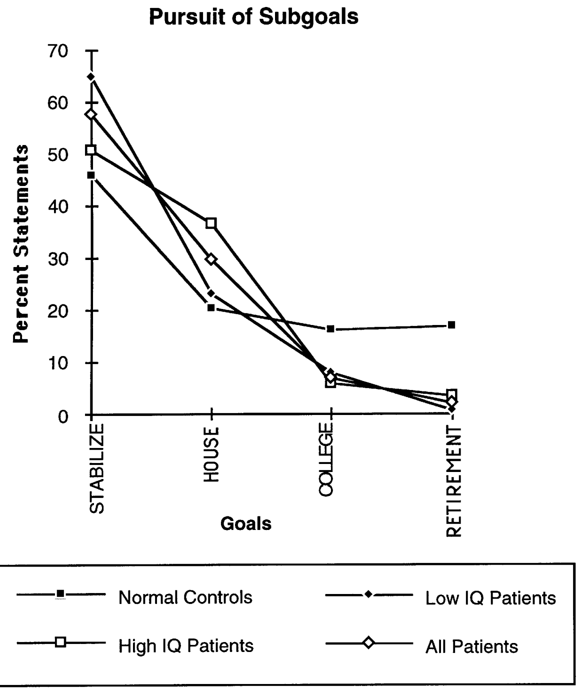

Fig. 7 Percentage of statements which control subjects and patients devote to each goal.

Table 4 *Goal satisfaction and strategy utilization on the financial planning task*

	Controls	Patients
Goals		
Acknowledgment (%)	100	70
Active pursual (%)	95	58
Strategies		
Decrease expenses (%)	90	80
Increase income (%)	90	50
Reallocate assets (%)	100	80

problem solving [$t(18) = 2.2$, $P < 0.05$]. The interaction between subjects and the planning phase was significant [$F(1,18) = 5.1$, $P < 0.05$].

There were also significant differences between control subjects and patients within the problem-solving phase (Fig. 6). Control subjects distributed their statements approximately equally between preliminary planning (mean 21.7%), refinement (mean 19.4%) and detailing (mean 17.8%). Patients, on the other hand, spent most of their time on the preliminary plan (mean 26.4%) and had less time left for refinement (mean 14.1%) and detailing (mean 7.0%). The resulting interaction approached significance [$F(2,36) = 2.4$, $P = 0.10$].

The episode-level analysis provides a measure of goals and strategies pursued (*see* Table 4). The task was designed to have four explicit goals and three solution factors or strategies that can be manipulated in various combinations to achieve the goals. Control subjects acknowledged all four goals and actively pursued a mean of 98%. Patients on the other hand acknowledged and actively pursued fewer goals (70% and 58%, respectively). Satisfaction of the goals requires the 'discovery' and manipulation of three solution factors: (i) decreasing expenses, (ii) increasing income and (iii) restructuring assets and liabilities. The control subjects were more successful than patients in utilizing each of these strategies. A number of patients focused exclusively on the decreasing expenses strategy. Only half the patients considered the 'increasing income' strategy and when patients did utilize the asset reallocation strategy, it tended to be at a superficial level.

Examination of Fig. 7 shows a significant interaction in the number of statements that control subjects and patients spent on the four goals [$F(3,54) = 2.5$, $P = 0.06$]. Normal control subjects devoted an equivalent number of statements to the three future goals (house, college and retirement), whereas patients with frontal lobe lesions devoted more statements to the immediate goal (stabilize finances) and a decreasing number of statements to future goals.

Examination of results at the operator level reveals that, overall, the distribution of operators was almost identical in the two groups (*see* Table 5); there was no difference in the overall distribution of operators by subject [$F(1,180) = 0.169$, $P = 0.69$] nor any interaction between subjects and operators [$F(10, 180) = 0.39$, $P = 0.95$]. Both patients and control subjects spent a similar amount of time adding, elaborating, evaluating and qualifying their statements in the

Table 5 *Distribution of categories at the statement level*

	Normal subjects	High IQ patients	Low IQ patients	All patients
Output mode (%)				
Verbal	87.27	86.22	95.53	90.88
Written	12.63	12.07	4.47	8.27
Source of information (%)				
Experimenter	9.33	7.23	23.42	15.33
Problem scenario	6.96	4.97	3.28	4.12
Self	80.14	82.11	71.20	76.65
Inference	3.45	3.94	2.10	3.02
Operators (%)				
Add	25.33	24.12	23.09	23.61
Evaluate	17.28	19.14	16.91	18.02
Propose	6.96	9.81	6.87	8.34
Comment	5.84	5.14	10.23	7.68
Repeat	8.09	5.88	7.17	6.53
Elaborate	16.18	13.98	15.23	14.60
Justify	1.84	2.62	1.59	2.10
Modify	2.21	3.06	2.84	2.95
Qualify	0.60	0.95	0.50	0.73
Read	3.94	2.56	1.43	1.99
Request	8.32	9.07	9.82	9.45

problem space. Nor were there any significant differences between subjects' use of individual operators.

On visual inspection there are some interesting differences between patients and subjects in their mode of output (verbal versus written) and the source of information utilized (*see* Table 5). Specifically, the patients relied more heavily on the experimenter for information than did control subjects. Control subjects utilized the problem scenario more extensively or generated information from their background knowledge. However, neither of these overall differences was significant.

To see if these results were a function of reduced IQ scores on the part of patients we divided the patients into two groups of five. The high IQ group had a mean WAIS—revised (WAIS-R) score of 112 while the low IQ group had a mean WAIS-R score of 89. We compared the performances of the two groups with those of the normal control subjects. In most cases, the above noted differences remained (*see* Figs 6 and 7, and Table 5), but they were no longer statistically significant due to the reduction in power. In contrast, differences between problem structuring and problem solving, the source of information and output mode parameters were influenced by IQ.

High IQ patients devoted the same number of statements to problem structuring and problem solving as did normal control subjects (Fig. 5). With respect to output mode, the main effect due to subjects became significant [$F(2, 17) = 3.6$, $P = 0.05$] and the interaction between subjects and the mode of output approached significance ($F(7,17) = 2.1$, $P = 0.15$). This is due to the fact that low IQ patients generated fewer written statements (and more verbal statements) than normal control subjects and high IQ patients. These differences approached significant levels: [$t(13) = 2.0$, $P = 0.06$] and [$t(8) = 1.9$, $P = 0.09$], respectively.

Similarly, for the source of information, the main effect due to subjects [$F(2,17) = 3.7$, $P < 0.05$] and the interaction between subjects and the source of information [$F(6,51) = 22.5$, $P < 0.05$] became significant. Analytical comparisons show that this is due to the fact that the low IQ patients solicited information more frequently from the experimenter and less frequently from the problem scenario than normal control subjects [$t(13) = 2.3$, $P < 0.05$ and $t(13) = 2.25$, $P < 0.05$, respectively].

We also divided up the patient group into two equal halves on the basis of their performance on two established frontal lobe tasks, the Tower of Hanoi task and the Wisconsin Card Sorting Test (WCST) (Table 1). However, four of our five high performers on the WCST and the Tower of Hanoi were also high IQ patients so the results look very much like the high and low IQ results. We are not reporting hemisphere comparisons because we only had one patient with only the left hemisphere lesioned.

Discussion

The main reason for administering this task was to test the long accepted claim that patients with frontal lobe lesions suffer from 'planning' deficits (Harlow, 1868; Penfield and Evans, 1935; Rylander, 1939). Before we discuss our results with respect to planning, it is worth considering how patients cope with the verbal protocol methodology.

On a per minute basis, control subjects generate ~25% more statements than patients, though the difference is not statistically significant. By viewing verbalization as a self monitoring task requiring dual or alternating attention, and recalling that patients with frontal lobe lesions often have attention deficits, one may well expect patients to have difficulties verbalizing while engaged in a complex problem-solving task. It is then possible to interpret the lack of statistical significance as just an artifact of the relatively low power of our study, and conclude that patients are indeed being hampered by the verbalization requirement and that the results are not indicative of their real planning or problem-solving ability. However, there is considerable evidence in the cognitive science literature indicating that the type of verbalizations solicited from subjects does not require self-monitoring, and, in fact, has no measurable overhead costs associated with it in normal populations (Ericsson and Simon, 1984). Given that (i) the difference is not significant, (ii) no subject complained about the verbalization requirement, or needed excessive reminding, and (iii) there were no differences between patients and control subjects in the rate of verbalization (as measured in 5-min intervals) across the duration of the problem-solving session, we favour the latter interpretation and conclude that patients with frontal lobe lesions can give an interpretable verbal protocol while engaged in a complex problem-solving task.

With respect to the planning task itself, our notion of planning comes from the design and cognitive science literature, and is much richer than the notion of 'sequencing' used in much of the neuropsychology literature. Goel (1995) characterizes design/planning situations in terms of a dozen constraints on the task environment. A planning situation exists only if a majority of these constraints are present. We discuss our results with respect to the following five constraints: problem structuring; phases of problem solving; no right or wrong answer or official termination point; lack of feedback from the world; and negotiable constraints.

Real-world planning problems are always ill-structured. The ill-structured/well-structured distinction has already been described in the Introduction. Before an ill-structured problem can be solved, it must be structured. That is, the information needed to specify the problem space must be (i) retrieved from background knowledge, the problem scenario, the experimenter, etc. and (ii) mapped onto the problem space. Our results indicate that normal control subjects structure the problem space relatively quickly and then spend the majority of their time solving the problem. On the other hand, patients with frontal lobe lesions take much longer on the structuring phase and spend less time on the problem-solving phase. Not only do they take longer, but based on the fact that they rely more heavily on the experimenter and the problem scenario as a source of information, we suggest that they have some difficulty with information retrieval. This difficulty is predicted by theories such as Grafman's 'structured event complex' (Grafman, 1995). Structured event complexes are large-scale knowledge structures that guide much of our routine behaviour. Certainly patients can retrieve individual events, even small sets of events. However, it does appear that some events are retrieved out of order, and others demonstrate inappropriate duration of activation (subjects staying too long in an activity, for example). A situation where the entire structured event complex may be difficult to retrieve, but fragments of the structured event complex are potentially accessible, would explain difficulties in problem structuring.

Once the information has been retrieved, mapping the information onto the problem space may be another source of difficulty for patients. The information that is retrieved by a subject will invariably be from a specific situation in their past experience. The mapping process will require that the information be abstracted from its specific context and then be instantiated in the current context (problem space). The determination of the relevance of information is the key to such mapping processes (Rescher, 1980; Carbonell, 1983). A recent PET neuroimaging study (Goel *et al*., 1997) indicates that these mapping processes implicate the left medial prefrontal cortex.

Difficulty in inferring abstract principles from particular instances is a classic frontal lobe deficit as demonstrated by sorting/classification tasks requiring similarity judgments (Halstead, 1940; Goldstein and Scheerer, 1941; Milner, 1963; Drewe, 1974; Malmo, 1974; Nelson, 1976) and nonsorting tests like the similarities and vocabulary subset of the WAIS scale and the proverb-definition task (Rylander, 1939).

The problem-solving phase in planning problems can be differentiated into preliminary planning, refinement and detailing (Goel, 1995). While normal control subjects spend roughly the same amount of time in each phase, patients with frontal lobe lesions spend progressively less time in each subsequent phase. One reason for this may be that patients simply become exhausted and overwhelmed by the task. However, the fact that the pace of statement generation does not decrease during the task strongly suggests that patients are not becoming tired. A second possibility is that patients run out of time. But patients actually complete the task 11 min before normal control subjects and they know they can have additional time if they wish. We think the real issue is one of patient judgment.

It is the case that in most ill-structured problems, there are no right or wrong answers, though there are certainly better and worse answers (Rittel and Webber, 1974). Another important characteristic of planning problems is that a plan represents a blueprint for achieving some future state of affairs. The 'goodness' of the plan cannot be determined until it is actually executed and its consequences allowed to unfold. However, the problem solver needs some measure of goodness (feedback) as the plan is being developed. This feedback must be self-generated, i.e. the problem solver must construct a mental model of the world and execute the plan in this model and 'observe' the consequences. This information can then be used to improve and modify the plan.

Given the fact that our task has no right or wrong answer, and thus no official termination point, subjects must make a judgment as to when they have satisfied the task requirements. Furthermore, the 'goodness' of their solution is also self-determined and will be a function of how well they model the relevant parts of the world necessary to test the plan. Patients are simply quicker to determine that they have satisfied the problem requirements. It is really a judgment issue as to what constitutes a satisfactory solution and when it has been reached. Patients genuinely believe they have specified a complete plan when they stop.

Another possibility is that, because they cannot successfully structure the problem space, there is not enough information there for them to successfully complete each of the phases of problem solving. But again, while this may be true, the judgment issue remains because, in the end, patients judge that they have successfully solved the problem. Similar errors in judgment are reported by Eslinger and Damasio (1985) for their patient E.V.R. and by Shallice and Evans (1978) in their cognitive estimation task.

Patients also make poor judgments about subgoal satisfaction. Both normal control subjects and patients spend more time on the first or immediate goal than the subsequent future goals. But the normal control subjects spend an equal amount of time on each of the three future goals (buy a house, college education for kids and retirement). Frontal lobe patients on the other hand spend a decreasing amount of time on future goals.

There are some similarities between this result and Shallice and Burgess' (1991*b*) findings that given a multiple goals task, patients organized their time poorly and spent too much time on individual tasks. They explained their results in terms of patient inability to reactivate previously formed intentions in the absence of a current trigger. While this is one possible explanation for our own results, there are some important differences in our experimental design that suggest that this may not be the most appropriate explanation for our patients' performance. Our subjects were given four related goals, and these stayed in front of them during the session. In fact the experimenter prompted subjects about each goal. Furthermore, there was no time restriction on our task. So given that patients were not under time pressure, and that the prompting by the experimenter would (should) serve as a current trigger, we need some other explanation of our patients' performance.

On the basis that subjects were required to write things and had access to the problem scenario documents at all times, we can discount memory explanations for patient failure to address future goals. Furthermore, if poor memory was a factor in patient performance, patients would add information into the problem space, and then forget that they had already done so, and add it again at a later time. This would show up in our coding scheme as an increase in the number of 'repeat' operators. Our results indicate no significant difference in the number of repeat operators between patients and control subjects.

We favour an alternative explanation for patients inadequately addressing future goals. The goals are temporally spaced into the future (stabilize situation now, buy a house in 1–2 years, send kids to collage in 15 years, and retire in 30 years). The greater the projection into the future, the less the projected situation will resemble the current situation. Grafman's (1995) structured event complex theory predicts that patient difficulty in retrieving information would increase as the task domain becomes more unfamiliar, because the strength of representation of large-scale knowledge structures such as a structured event complex is dependent on the frequency of exposure to these structures and the frequency of their activation. Thus, just like an aphasic who may have a relatively easier time processing a high frequency, high imagery word compared with a low frequency abstract word, the patient with a prefrontal cortical lesion would have more difficulty in processing a low frequency (unfamiliar) structured event complex compared with a high frequency (familiar) one.

The Norman and Shallice model (Shallice and Burgess, 1991*a*) makes a similar prediction. Given the unfamiliarity of future situations the routine knowledge scripts would no longer be applicable, and would result in an impasse or the triggering of the contention scheduler. Normally the supervisory system would intervene and call upon 'special-purpose cognitive subsystems' to carry on. In patients with frontal lobe lesions, the intervention by the supervisory system is not successful.

On the Grafman (1995) account it is the long-term memory traces that are damaged or lost. On the Shallice account the problem is with the control mechanism. While the distinction between data structures and control structures is well-defined in classical computational systems, it is notoriously difficult to make on the basis of behavioural data without some strong assumptions about the functional architecture (Pylyshyn, 1984).

Another differentiating feature of real-world planning problems is the nature of the constraints associated with them. Constraints typically fall into three categories: (i) logical, definitional or constitutive, (ii) nomological, and (iii) social, economic, political, cultural, etc. In puzzles and games, the constraints are logical or constitutive of the task, i.e. if one violates a constraint or rule, one is simply not playing that game [e.g. if we are playing chess, and I move my rook (castle) diagonally across the board, I am simply not playing chess]. The constraints we encounter in real-world planning problems are of the latter two kinds. Nomological constraints are constraints dictated by natural law. So, for example, if a beam is to support a downward thrust of *x* pounds per square inch, it must exert an upward thrust of equal or greater amount. These constraints, while never negotiable, are also not definitional or constitutive of the task. The social, economic, political and cultural constraints are negotiable (e.g. if you go to an architect and ask him to build a new house for you, and he convinces you to renovate your existing house instead, or move in with him, it seems odd to say that he is not playing the game of design).

Our results indicate that patients with frontal lobe lesions did not fully appreciate or utilize the negotiability of constraints. For example, only half of the patients utilized the 'increasing income' strategy. To utilize this strategy subjects had to consider situations that were not given in the problem scenario (e.g. Ted getting a second job or starting a part-time home business, Carol going to work, etc.). Control subjects were much more likely to consider these possibilities than patients. This is consistent with existing observations that frontal lobe patient's ability to shift quickly between concepts and perspectives or 'mental sets' is impaired.

This result has been obtained with both the sorting tasks and the body orientation task. In the sorting tasks, the patients may get the first category right, but are unable to switch rapidly to another category (Rylander, 1939; Luria, 1966). The body-orientation task requires subjects to match body position (on their own body) with that of a line drawing of a human figure. The line drawing was presented in one of four different orientations (facing toward subject, facing away from subject, upside down and upright). Patients displayed selective impairments which were attributed to a deficit in switching between perspectives (Semmes *et al*., 1963).

Conclusion

Our results indicate that patients with frontal lobe lesions do have certain difficulties with real-world planning problems. In particular, patient performance is impoverished at a global level but not at the local level, i.e. at the statement level, a time-slice on the order of seconds, their performance is indistinguishable from that of control subjects. They instantiate the same set of operators, in the same sequence, with the same frequency. But when we examine their performance over a scale of minutes to hours, differences begin to emerge. Patients have difficulty in organizing and structuring their problem space. Once they begin problem solving, they have difficulty in allocating adequate effort to each problem-solving phase. Patients also have difficulty dealing with the fact that there are no right or wrong answers, nor an official termination point in real-world planning problems. They also find it problematic to generate their own feedback. They invariably terminate the session before the details are fleshed out and all the goals satisfied. Finally, patients do not take full advantage of the fact that constraints on real-world problems are negotiable.

It is apparent from this list that there is no single unifying difficulty that patients with frontal lobe lesions encounter that can be termed a 'planning' deficit. It is possible to understand patients' difficulties in coping with ill-structured real-world planning problems in terms of accepted frontal lobe deficits. We have implicated four accepted deficits to account for the poor performance of patients with frontal lobe lesions in real-world planning tasks: inadequate access to 'structured event complexes' (due to damage to control mechanisms or the memories themselves); generalization from particulars; failure to shift between 'mental sets'; and poor judgment regarding the adequacy and completeness of a plan.

Furthermore, there is no reason to believe that these difficulties are intrinsically related to planning problems. They should arise in many reasonably complex, ill-structured, real-world problem-solving situations. So, while the literature has focused on planning problems, the dimension of ill-structured versus well-structured may be the more relevant one. It just so happens that all real-world planning problems are ill-structured, but many nonplanning problems are also ill-structured. One of us has made the argument that the ill-structured/well-structured distinction is a fundamental one, and that the computational mechanisms that have proved so useful in dealing with well-structured problems are, in principle, inadequate to deal with ill-structured problems (Goel, 1995). Ill-structured and well-structured problems may well require fundamentally different types of computational mechanisms.

Finally, since we do not have a patient control group with lesions in areas other than the prefrontal cortex, we cannot conclude that these performance characteristics are unique to patients with frontal lobe lesions. Our results are consistent, however, with the clinical observations and experimental findings of other researchers (Shallice and Evans, 1978; Eslinger and Damasio, 1985; Shallice and Burgess, 1991*b*; Sirigu *et al*., 1995; Sirigu *et al*., 1996).

Acknowledgements

We wish to thank two reviewers of this journal for valuable comments and advice that greatly improved the quality of this paper. Portions of this work were supported in part by grants from York University's President's NSERC Fund, Faculty of Arts funds, and the National Sciences and Engineering Research Council of Canada.

References

Alivisatos B, Milner B. Effects of frontal or temporal lobectomy on the use of advance information in a choice reaction time task. Neuropsychologia 1989; 27: 495–503.

Anderson JR. The architecture of cognition. Cambridge (MA): Harvard University Press, 1983.

Bechara A, Damasio AR, Damasio H, Anderson SW. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 1994; 50: 7–15.

Carbonell JG. Learning by analogy: formulating and generalizing plans from past experience. In: Michalski RS, Carbonell JG, Mitchell TM, editors. Machine learning: an artificial intelligence approach. Palo Alto (CA): Tioga Press, 1983: 137–62.

Chandrasekaran B. What kind of information processing is intelligence?: a perspective on AI paradigms and a proposal. In: Partridge D, Wilks Y, editors. Source book on the foundations of AI. Cambridge: Cambridge University Press, 1987: 14–46.

Corkin S. Tactually-guided maze learning in man: effects of unilateral cortical excisions and bilateral hippocampal lesions. Neuropsychologia 1965; 3: 339–51.

Damasio H, Damasio AR. Lesion analysis in neuropsychology. New York: Oxford University Press, 1989.

Diamond A. Developmental time course in human infants and infant monkeys and, the neural bases of inhibitory control in reaching. [Review]. Ann N Y Acad Sci 1990; 608: 637–76.

Drewe EA. The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. Cortex 1974; 10: 159–70.

Ericsson KA, Simon HA. Protocol analysis: verbal reports as data. Cambridge (MA): MIT Press, 1984.

Eslinger PJ, Damasio AR. Severe disturbance of higher cognition after frontal lobe ablation: patient EVR. Neurology 1985; 35: 1731–41.

Goel V. A comparison of design and nondesign problem spaces. Art Intell Eng 1994; 9: 53–72.

Goel V. Sketches of thought. Cambridge (MA): MIT Press, 1995.

Goel V, Gold. B, Kapur S, Houle S. The seats of reason? An imaging study of deductive and inductive reasoning. Neuroreport 1997; 8: 1305–10.

Goel V, Grafman J. Are the frontal lobes implicated in 'planning' functions: interpreting data from the Tower of Hanoi. Neuropsychologia 1995; 33: 623–42.

Goel V, Pirolli P. The structure of design problem spaces. Cogn Sci 1992; 16: 395–429.

Goldberg E, Bilder RM Jr. The frontal lobes and hierarchical organization of cognitive control. In: Perecman E, editor. The frontal lobes revisited. New York: IRBN Press, 1987: 159–87.

Goldstein K, Scheerer M. Abstract and concrete behavior: an experimental study with special tests. Psychol Monographs 1941; 53 (Whole No. 239).

Grafman J. Similarities and distinctions among current models of prefrontal cortical functions. [Review]. Ann N Y Acad Sci 1995; 769: 337–68.

Halstead WC. Preliminary analysis of grouping behavior in patients with cerebral injury by the method of equivalent and non-equivalent stimuli. Am J Psychiat 1940; 96: 1263–94.

Harlow JM. Recovery after severe injury to the head. Pub Mass Med Soc 1868; 2: 327–46.

Hayes JR, Simon HA. Understanding Written problem instructions. In: Gregg LW, editor. Knowledge and cognition. Potomac (MD): Lawrence Erlbaum, 1974: 167–200.

Luria AR. Higher cortical functions in man. London: Tavistock, 1966.

Malmo HP. On frontal lobe functions: psychiatric patient controls. Cortex 1974; 10: 231–7.

McCarthy RA, Warrington EK. Cognitive neuropsychology: a clinical introduction. San Diego: Academic Press, 1990.

Milner B. Effects of different brain lesions on card sorting: the role of the frontal lobes. Arch Neurol 1963; 9: 90–100.

Nelson HE. A modified card sorting test sensitive to frontal lobe defects. Cortex 1976; 12: 313–24.

Newell A. Physical symbol systems. Cognit Sci 1980; 4: 135–83.

Newell A. Unified theories of cognition. Cambridge (MA): Harvard University Press, 1990.

Newell A, Simon HA. Human problem solving. Englewood Cliffs (NJ): Prentice-Hall, 1972.

Penfield W, Evans J. The frontal lobe in man: a clinical study of maximum removals. Brain 1935; 58: 115–33.

Perret E. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia 1974; 12: 323–30.

Pylyshyn ZW. Computation and cognition: toward a foundation for cognitive science. Cambridge (MA): MIT Press, 1984.

Reitman WR. Heuristic decision procedures, open constraints, and the structure of ill-defined problems. In: Shelly MW, Bryan GL, editors. Human judgments and optimality. New York: John Wiley, 1964: 282–315.

Rescher N. Induction: an essay on the justification of inductive reasoning. Pittsburgh: University of Pittsburgh Press, 1980.

Rittel HWJ, Webber MM. Dilemmas in a general theory of planning. DMG-DRS J 1974; 8: 31–9.

Roberts RJ, Hager LD, Heron C. Prefrontal cognitive processes: working memory and inhibition in the Antisaccade task. J Exp Psychol Gen 1994; 123: 374–93.

Rylander G. Personality changes after operations on the frontal lobes. Acta Psychiat Neurol Scandi 1939; Suppl 20.

Semmes J, Weinstein S, Ghent L, Teuber H-L. Correlates of impaired orientation in personal and extra-personal space. Brain 1963; 86: 747–72.

Shallice T. Specific impairments of planning. Philos Trans R Soc Lond B Biol Sci 1982; 298: 199–209.

Shallice T, Burgess P. Higher-order cognitive impairments and frontal lobe lesions in man. In: Levin HS, Eisenberg HM, Benton AL, editors. Frontal lobe function and dysfunction. New York: Oxford University Press, 1991a: 125–38.

Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. Brain 1991b; 114: 727–41.

Shallice T, Evans ME. The involvement of the frontal lobes in cognitive estimation. Cortex 1978; 14: 294–303.

Simon HA. Human cognition and problem solving. In: University of California Medical School Symposium on 'Control of the mind'. 1961: 1–8.

Simon HA. The functional equivalence of problem solving skills. Cognit Psychol 1975; 7: 268–88.

Simon HA. Information-Processing theory of human problem solving. In: Estes WK, editor. Handbook of learning and cognitive processes, Vol. 5. Hillsdale (NJ): Lawrence Erlbaum, 1978: 271–95.

Sirigu A, Zalla T, Pillon B, Grafman J, Agid Y, Dubois B. Selective impairments in managerial knowledge following pre-frontal cortex damage. Cortex 1995; 31: 301–16.

Sirigu A, Zalla T, Pillon B, Grafman J, Agid Y, Dubois B. Encoding of sequence and boundaries of scripts following prefrontal lesions. Cortex 1996; 32: 297–310.

Smith ML, Milner B. Estimation of frequency of Occurrence of abstract designs after frontal or temporal lobectomy. Neuropsychologia 1988; 26: 297–306.

Wilkins AJ, Shallice T, McCarthy RA. frontal lesions and sustained attention. Neuropsychologia 1987; 25: 359–65.

Received March 6, 1997. Accepted May 23, 1997

Appendix A
Experimental task instructions

Thank you for participating in our study. Very generally, we are interested in people's problem-solving and reasoning processes, particularly those associated with planning tasks. Therefore, during the next hour we are going to ask you to engage in a planning exercise as called for in the accompanying problem scenario.

As you can see, the session will be taped on a video recorder. The tapes provide us with a trace of your problem-solving activity and allow us to engage in an in depth analysis at a later date.

For the recordings to be of maximal benefit to us, we are going to ask you to *talk aloud* as you proceed through the task. By this we do not mean that you should explain what you are thinking. Rather, you should just try to vocalize the fragments of thoughts and ideas that you might be attending to at the time. We would like to get a continuous stream of such vocalizations from you.

It is not easy (or even normal) to attend to a complex task and verbalize at the same time. Therefore, you will undoubtedly lapse into periods of silence. This is to be expected. During such periods we will prompt you to speak. This is a routine part of our experimental methodology.

Due to practical considerations we have decided to limit the session to approximately one hour. However, time is not an important factor. (We have two hours of tape and will not mind if you choose to go over one hour.) What is important is that you address the problem as fully as possible and outline a reasonably detailed plan of action.

Also, the enclosed problem scenario is rather sparse. It is intended that you converse with the experimenter to iron out any difficulties and shortcomings.

Please begin.

Problem scenario for financial planning task

Ted and Carol are members of the post baby-boom generation. Both are reasonably well educated. Ted graduated from university in 1987 with a Ph.D. in computer science. He received several job offers upon graduation. Favouring a teaching career, he accepted a position as assistant professor of computer science at a small regional college in the Washington DC area. Carol is a homemaker who currently has her hands full with the couple's 3-year-old daughter.

This year Ted is up for review and promotion and he and Carol are expecting their second child in 3 months. They are using the occasion to think about the course of their lives and to articulate some long-term plans and goals.

Ted enjoys his job and Carol derives a certain satisfaction from raising their daughter. Both are looking forward to the arrival of the second child. However, a number of concerns about the long-term economic future have recently been on their minds. How will they purchase their first house, pay for a college education for their children, and save for retirement? The economic prosperity and security which seemed to come so effortlessly for their parent's generation seems beyond their grasp.

Enclosed are documents relating to Ted and Carol's current life style and financial situation. You may ask for any other information that you feel is relevant. The experimenter will respond on behalf of Ted and Carol.

You are required to do two tasks. (1) The first is to analyse Ted and Carol's financial situation and answer the attached questions. (2) The second task is to help Ted and Carol to plan and structure their savings, expenditures, and life style such that they are able to achieve the following goals:

(i) Stem their negative cash flow.
(ii) Purchase a home (preferably during the next 2 years).
(iii) Send their two children to college.
(iv) Have sufficient funds to retire at the age of 65 years.

You should provide Ted and Carol with a plan of action. The plan needs to contain the following three things:

(A) A list of things they should do,
(B) When they should do them, and
(C) What they can expect the financial consequences to be. The consequences should be specified in terms of budget projections for each goal.

Some financial tables and a calculator are provided to assist you with financial calculations. The experimenter is also prepared to do any financial calculations for you at your request.

Appendix B
Coding scheme, adapted from Goel (1995) with permission from MIT Press

The *statement level* code (Fig. 4) has four independent fields; the *operator* applied, the *content* to which it is applied, the *mode* of the output and the *source* of knowledge used.

Operators

Operators are a labelling of statements by the function they serve in the problem space. While no theoretical commitment is made to any specific set, the eleven noted below are adequate for current purposes:

Add: The basic operation of putting something into the problem space with some degree of commitment.

Propose: Indicates that an idea is being entertained but is not yet committed to the problem space.

Evaluate: Means that the statement is an explicit evaluation of a previous statement or plan component in the problem space.

Comment: It is by and large the report of an activity rather than the execution of it. Comments generally occur with monitoring statements. They often involve the subject explaining what he has just done, or just making some remarks, which, while not directly related to his progress, are none the less illuminating.

Modify: a statement which deletes or alters an existing idea or element which is already a part of the problem space. It is sometimes difficult to distinguish between *add* and *modify,* i.e. to distinguish between an old idea being modified and a new idea being added.

Elaborate: Expands an existing idea or element.

Justify: Offers a rationale for the addition, modification or elaboration of ideas or elements in the problem space.

Read: Any time the subject reads from the experimental task instructions, problem scenario, or any other documents supplied with the task.

Qualify: A statement used to hedge or further qualify the previous statement.

Request: Statements used to ask questions of or make suggestions to the experimenter.

Repeat : The application of the same operator to the same content again. While any operator can be repeated, it is usually only *add, modify,* and *elaborate* operations which actually are repeated.

Miscellaneous: Any statement which can't be coded with one of the above operators.

Content

The content to which the operator is applied is also noted.

Mode of output

The mode of output of a statement is encoded as either verbal or written: Hand and facial gestures are not encoded.

Verbal: Statements which are only uttered verbally, with no accompanying mark-on-paper.

Written: Statements accompanied by marks-on-paper. These statements may or may not have an associated verbalization.

Source of knowledge

Each statement is also encoded for the source of knowledge for the statement. The four categories used are the *experimenter,* the *problem -scenario, self* (retrieved from long-term memory), and *inferred* (deductively) from the information existent in the problem space.

Experimenter: This is information which is either given to the subject by the experimenter, or actively solicited by the subject from the experimenter.

Problem scenario: This is information which the subject has obtained directly from the problem scenario statement and any accompanying documents.

Self: This is information which the subject either generates or retrieves from his long-term memory.

Infer: This is information which the subject infers (in the strong deductive sense) from the information existent in the problem space.

Plan-development level

The plan-development level (Fig. 4) codes the statement as either an experimental task (expt-task), monitoring phase (monitor), planning phase (planning-phase), or miscellaneous statement type, where each is defined as follows.

Expt-task: Any statement having to do with the experimental design and setup.

Monitor: Any statement used to take stock, further, review, or comment on the problem-solving process itself. Most of such statements correspond to what in the literature have been called meta-cognitive statements.

Plan-development: Statements that advance the state of the plan/design.

Miscellaneous: Any statement that does not fall into one of the above categories.

Planning phase statements

The planning phase statements are further categorized into the following four subcategories.

Problem-structuring: Statements that serve to solicit or generate information to structure the problem.

Preliminary-plan: Statements that result in the initial generation and exploration of some aspect of the plan.

Refine: Statements that serve to elaborate and further the commitment to an already generated plan idea or element.

Detail: Statements that serve to detail and give the final form to some aspect of the plan.