Pergamon

Neuropsychologia, Vol. 34, No. 4, pp. 263–273, 1996

0028–3932(95)00104–2

Printed in Great Britain
0028–3932/96 $15.00 + 0.00

Response suppression, initiation and strategy use following frontal lobe lesions

PAUL W. BURGESS*† and TIM SHALLICE†‡

†Department of Psychology, University College London, Gower Street, London WC1E 6BT, U.K.; and
‡Department of Clinical Neuropsychology, The National Hospital for Neurology and Neurosurgery, London WC1, U.K.

(*Received* 30 *July* 1993; *accepted* 30 *May* 1995)

Abstract—Ninety-one patients with cerebral lesions were tested on a task involving two conditions. In the first condition (response initiation) subjects were read a sentence from which the last word was omitted and were required to give a word which completed the sentence reasonably. In the second condition (response suppression) subjects were asked to produce a word unrelated to the sentence. Patients with frontal lobe involvement showed longer response latencies in the first condition and produced more words which were related to the sentence in the second, in comparison to patients with lesions elsewhere. Moreover, in the second condition patients with frontal lobe lesions produced fewer words which showed the use of a strategy during response preparation. Performance on the initiation and suppression conditions was unrelated at the group or single case level. The relationships between response initiation, suppression and strategy use are discussed.

Key Words: frontal lobes; executive function; strategy; response suppression; inhibition; initiation.

Introduction

It has long been known that frontal lobe lesions can cause deficits in both response initiation and suppression [7, 8]. For instance, on the Verbal Fluency test (VFT) patients with frontal lobe lesions produce fewer words than patients with lesions elsewhere [e.g. 2, 10, 12; but see 18]. One view is that this reduced word fluency is linked to a general aspontaneity of thought [e.g. 7, 17] and so is a feature of frontal initiation problems. A more specific interpretation of reduced verbal fluency in certain patients has been given by De Lacey Costello and Warrington [4], namely that reduced word fluency is one characteristic of the syndrome of dynamic aphasia, in which patients present with reduced spontaneous use of language in the context of normal confrontation naming, language comprehension and speech production. They further suggest that dynamic aphasia represents a deficit in verbal planning at a stage prior to sentence construction.

Response suppression, too, is well known to present problems for frontal lobe patients. Thus, Drewe [6] found that patients with frontal lobe lesions were significantly poorer than patients with lesions elsewhere at a go–no go task which required the subjects to withhold responses to one or two stimuli. More specifically, Verfaellie and Heilman [21] gave a patient with medial frontal lobe lesions the task of responding to a touch of one hand by raising the contralateral hand. The patient consistently raised the ipsilateral hand in response to the touch. The authors interpreted this as an inability to inhibit an incorrect response since the patient showed self-corrections and was able to describe verbally the correct response when asked to do so.

The issue of whether the inability to suppress the most salient response is related to any response initiation difficulty is raised by the work of Perret [16] who used a form of the Stroop test with groups of patients with cortical lesions, and found that the patients whose lesion invaded the left frontal lobe were significantly poorer at suppressing the habitual response than patients with lesions elsewhere. In addition, those patients who

*To whom all correspondence should be addressed.

performed poorly on the Stroop test were also those who were most likely to achieve a poor Verbal Fluency score.

Perret's favoured explanation of this finding was that the Stroop task and the Verbal Fluency task make similar cognitive processing demands, despite their apparent differences in format. He argued, in reference to the VFT, that the "usual criterion in the search for words" (p. 323) is word meaning. By asking the patients to perform their search according to initial letter, one is asking them "to suppress the habit of using words according to their meaning" (p. 324). Thus for Perret, both the Stroop test and the VFT require the subjects to suppress the habitual response. A possibility which Perret does not consider, but which is a variant of this view, is that an initial step that is required before the generation of a new word is the suppression of the currently active representation.

However, Perret also considered an alternative explanation, namely that the two processes might be functionally independent, but that the cerebral structures which subserve them could be in close proximity in the brain so that both processes might tend to be affected together in patients with cerebral damage. Yet a third explanation is possible for his result, namely that a high-level "supervisory system" [18] is damaged in frontal lobe patients, thus both the initiation of novel responses and response suppression abilities are impaired. The question of the relationship between response initiation and inhibition therefore remains open.

A difficulty in the interpretation of the performance of frontal patients across different frontal tests such as the Stroop and word fluency arises from the differing characteristics of the specific tests. The present study, therefore, utilises a task in which the differing components of initiation and inhibition can be examined with minimal changes in the background characteristics of the task. Sentences in which the last word is omitted are presented to the patient, the sentences being chosen such that there is a high probability of a particular response occurring. The patient's task is to complete the sentence, in one condition with any word and in the other with a word which makes no sense in the context; in this second condition the rational response has to be inhibited.

Method

Subjects

A consecutive series of 108 patients referred for routine neuropsychological assessment to the Department of Clinical Neuropsychology at the National Hospital for Neurology and Neurosurgery were seen as part of the present study. Patients were then excluded if they did not pass all the following criteria:

Right-hand dominant (as assessed by the admitting neurologist) and aged between 18 and 75, with a unilateral lesion restricted to one or two lobes (with the exception of bilateral frontal lesions with no posterior involvement, who were considered as a separate group; no history of psychiatric problems, alcohol or substance abuse, or any previous neurological conditions; no hydrocephalus or long-standing epilepsy; first language as English; having a CT brain scan available for their current condition; and of course being capable of, and willing to undertake the experimental procedure in addition to a full neuropsychological assessment.

Seventeen of the original 108 patients failed to meet these criteria and were excluded, leaving 91 patients. The remaining patients were classified according to site of lesion based on the radiographer's report of the CT scan. Any patient who had involvement of the frontal lobes was classified as 'anterior', and those patients who had involvement elsewhere in the cortex, but not involving the frontal lobes, were classified as 'posterior'. Patients who had bilateral frontal lobe lesions with no posterior (i.e. non-frontal) involvement were considered separately as a 'bifrontal' group. The ages of these patients, together with their pre-morbid IQ estimates based on the NART [14], and current levels of intellectual functioning [22] are shown in Table 1. Of the 91 cases, 44 were confirmed as suffering from primary tumours, the majority of which were forms of glioma with 15 cases of menigioma and six with cystic lesions. Sixteen cases had space-occupying lesions, presumed neoplastic, the precise nature of which could not be confirmed at the time of testing. Twenty-one cases were suffering from lesions which were vascular (infarct and haemorrhage) in origin and six cases had a subdural haematoma. Additionally there were two cases with a metastatic lesion and two cases with a cerebral abscess. Overall 60.9% of the anterior group and 77.8% of the posterior group were confirmed as having, or were presumed (on the basis of CT scan) to be suffering from, a tumour. Of the anterior cases 28.1% had vascular lesions, compared with 11.1% of the posteriors. Of the posterior group 57.9% had lesions confined to one lobe, compared with 61.8% of the anterior group. Details of the patients along with a group of 20 normal subjects matched for age and IQ are given in Table 1. The controls were mainly relatives accompanying patients. None had experience of psychology or related disciplines.

Materials

The Hayling test consists of 30 sentences in which the final word was omitted. They were selected from those given by Bloom and Fischler [3] who present norms for completion of the sentences; the 30 chosen were those for which there was a particularly high probability of one specific response. The American terms in the sentences were translated into their British equivalents (see Appendix 1 for examples). The 30 sentences were pseudo-randomly assigned to one of two groups of 15 sentences to be used in either the first or second sections of the test. In addition four more sentences from Bloom and Fischler were selected to be used as examples for the subjects.

Method and task instructions

The task consists of two sections (A and B), each containing 15 sentences. Subects were tested with only the examiner present, in a quiet standard hospital office.

Table 1. Subject characteristics (S.D. in brackets)

Group	n	Age	WAIS-IQ*‡	NART†
L anterior	21	40.8 (13.2)	100.1 (13.8)	106.3 (10.5)
R anterior	26	48.5 (13.3)	109.6 (13.4)	114.3 (8.8)
L posterior	12	37.8 (14.3)	102.1 (15.8)	111.2 (11.6)
R posterior	15	48.1 (14.0)	108.6 (9.5)	111.5 (9.8)
Bifrontal	17	51.4 (13.0)	100.6 (11.0)	107.2 (9.6)
Collapsed to give:				
Anteriors	47	45.1 (13.7)	105.3 (14.3)	110.7 (10.3)
Posteriors	27	43.5 (14.8)	105.7 (12.8)	111.4 (10.4)
Bifrontal	17	51.4 (13.0)	100.6 (11.0)	107.2 (9.6)
Controls	20	49.7 (13.7)	—§	112.0 (12.9)

*Certain subjects were given the WAIS—R; a correction factor of + 8 points was then used [5].
†National Adult Reading Test [16].
‡Wechsler Adult Intelligence Scale [22].
§Not administered.

Section A: Initiation

The subject was given spoken instructions of which the critical part is "I want you to listen carefully to each sentence and when I have finished reading it, your job is to give me a word you think could fit at the end of the sentence as quickly as possible". Two practice sentences are read to the subject and s/he was required to give a reasonable response. If the subject was still unsure about the task, further explanation was given (in practice this was extremely rare). Sentences were then read to the subject at a normal reading pace. Timing of response latency was by stop-watch, started as soon as it was judged the last word of the sentence had been read, with timing being stopped as soon as the subjects began their response.

Section B: Response suppression

Section B was conducted immediately after section A and used the same type of sentence stem. Between the two tasks the subjects were given instructions with the following critical section "this time I want you to give me a word which makes *no sense at all* in the context of the sentence. I want the word you give me to be *unrelated* to the sentence in every way". The two examples were then given to the subject. If the subjects was unable to think of an unrelated word, the example word 'banana' (which is unrelated to either of the two practice items) was suggested to the subject. If at any time during this stage of the test, the subject gave a sentence completion rather than an unrelated word, s/he was told that the word was too related to the sentence, and the task instructions were repeated. In practice, with a few of the subjects who were very poor at the task, this resulted in their being reminded about the task instruction after each of the 15 trials. If the patient had not produced a word within 60 sec, that trial was terminated and a response latency of 60 sec was recorded.

Results

No significant laterality effects appeared in any of the measures (see Table 2), so the unilateral patients were collapsed to give four groups, unilateral anterior or posterior, a bi-frontal group and controls.

In the statistical analysis we will principally consider the contrast between the control group and the two unilateral lesion groups since the control group and particularly the unilateral posterior group are the critical contrast groups for drawing theoretical conclusions about any putative unilateral anterior group deficit. To include the bifrontal group would be inappropriate since any selective deficit the group shows is potentially explicable in terms of lesion size effects and moreover their inclusion would contaminate analyses of lesion extent effects within the unilateral lesion patients. The bifrontal anterior patients are therefore contrasted with the unilateral anterior patients to see if any selective unilateral anterior deficit is exacerbated in the bilateral anterior group.

Table 2. Response latencies across groups (sec; S.D. in brackets)

Group	Section A	Section B	B–A
Controls	12.6 (6.5)	26.4 (19.5)	13.8 (16.3)
Posteriors	17.4 (14.1)	40.6 (38.9)	23.2 (28.9)
Anteriors	28.6 (34.5)	83.1 (106.8)	54.4 (79.3)
Bifrontals	35.8 (23.1)	135.4 (119.3)	99.6 (118.9)
LH (n=33)	26.2 (22.3)	59.6 (68.2)	33.3 (56.0)
RH (n=41)	23.2 (33.9)	74.1 (105.1)	50.9 (74.5)

LH = all patients with (unilateral) left hemisphere lesions considered together.
RH = all patients with (unliateral) right hemisphere lesions.

The study used neuropsychological group study methodology in which selection of unilateral patients was determined as far as the lesion was concerned only by no more than two lobes being involved. Thus patients with a number of different aetiologies and differing lesion sizes are included. As a control check we therefore followed significant findings by a multiple regression analysis which included aetiology and number of lobes involved.

Section A of the task (straightforward completion)

The quality of the answers was satisfactory for all groups since less than 1% of responses in this section were not one of those predictable from the results of Bloom and Fischler [3] on the quality of the replies. The response latencies (sum of latencies across 15 trials) for all groups are shown in Table 2. There was a significant difference between the unilateral anterior, posterior and control groups [ANOVA following log transformation to eliminate skewness d.f. 2, 91; $F=7.80$, $P<0.002$]. Tukey–Kramer *post-hoc* comparisons showed that the anterior group were significantly slower than the controls at the $P<0.01$ level. None of the other comparisons were significant for $P<0.05$.

Analysis of the patients' times by multiple linear regression indicated both age [t-ratio 3.27, $P<0.005$] and FSIQ [t-ratio -3.12, $P<0.005$] to be significant predictors of performance on this measure. An analysis of covariance (following log transformation) was therefore performed using age and FSIQ as covariates. Under these conditions the anterior–posterior difference in the section A times was significant at the 0.02 level. The bifrontal group were poorer still than the unilateral anteriors [$t=-2.12$, d.f. $=36$, $P<0.05$].

In order to investigate the effect of lesion aetiology and number of lobes involved the patient's lesions were classified into three broad categories: vascular, tumour (or presumed tumour) or 'other'. Aetiology of lesion, together with number of lobes affected and whether the lesion was anterior or posterior in location were used as predictors of response latencies using multiple linear regression (log transformed data). When all predictors were considered together, only the anterior–posterior classification were significantly predictive of performance [t-ratio 2.14, $P<0.05$]. Neither the unilateral posterior or anterior single-lobe involvement cases were significantly different in their pattern of performance from their unilateral two-lobe counterparts [posteriors: one lobe = mean 17.8 (16.1), two lobes = 17.2 (8.9); anteriors: one lobe = mean 22.0 (22.0), two lobes = 20.6 (11.0)].

Section B of the task (anomalous completion condition)

Error classification. In order to discover whether the groups differed in their ability to produce words unrelated to the stimulus sentence in section B, three raters, blind to the purpose of the study, rated each of the 1665 subject reponses (111 subjects × 15 sentences). They were asked to classify each of the responses into one of the eight possible categories shown in Fig. 1, following the instructions presented in Appendix 1.

These eight categories can be subdivided into three main groups. Firstly, there are responses which are sensible completions of the sentence, thus clearly violating the task instructions (category C in Fig. 1). Secondly, there are responses that are semantically connected to the sentence in some way (categories SO, SA, SB and SC in Fig. 1). This second set of responses may be semantically related to either the expected completion of the sentence (including opposites of what might constitute a sentence completion), or the subject of the sentence, or it may be a word, which whilst not entirely completing the sentence in a reasonable fashion, nevertheless gives the sentence some meaning (socially inappropriate responses were included in this category). Thirdly, there are those responses which are unrelated to the sentence, as required by the task instructions (these subcategories will be differentiated later).

Two or more raters agreed on 94.1% of occasions, which is most satisfactory given that each response can be assigned to one of eight classifications. In the remaining 5.9% of responses, a fourth rater, blind to the purpose of the study, was asked to decide which one of the three raters was correct.

Error scoring. An error score was devised which is a measure of the overall semantic relatedness of each response to its stimulus sentence (see Fig. 1). Three points were given if the word was a straightforward completion of the sentence (category C), and 1 point for a word that was semantically related to the sentence in some way (categories SO, SA, SB and SC). Categories U, UR, UL and URL represent successful completion of the task requirements and so were given no score. A summary of the error score results is given in Table 3. Under these conditions the difference in performance between the control group, unilateral anterior group and the posterior group was significant at the 0.002 level [ANOVA, d.f. $=2$, 91; $F=7.34$]. A Tukey–Kramer *post-hoc* comparison showed the difference between the anterior group and the control group is significant at the 0.05 level [$Q=4.1$, $r=3$, d.f. $=91$], with the difference between the anterior group and the posterior group significant at the 0.01 level [$Q=4.4$, $r=3$, d.f. $=91$] (see Table 3). Multiple linear regression showed that age and IQ were significant predictors of performance as measured by the error score. In order to check that the significant anterior–posterior effect in this analysis was not explicable in terms of these two factors, an analysis of covariance was performed on the patients' scores using age and FSIQ as covariates. Under these conditions, the anterior–posterior difference was significant at the 0.005 level.

Investigation of the effects of lesion involvement

Step		Classification	Error score
1.	Does the word reasonably complete the sentence (i.e. it's a word you might give yourself if asked to provide a word that would fit at the end).	C	3
2.	Is the word an opposite of what you might expect as an answer?	SO	1
3.	Is the word obviously semantically connected to the subject of the sentence?	SA	
4.	Is the word obviously semantically connected to the expected response?	SB	1
5.	Does the word vaguely fit at the end of the sentence, but in a way that makes the sense of the sentence ludicrous – or is the word a slang semi-obsenity?	SC	1
	If the word completely unconnected to the sentence:		
6.	Is the word an item you might find in a hospital office?	UR	0
7.	Is the word semantically connected to the subject's last response?	UL	0
8.	Is the word both an item you might find in an office, and is also related to the last answer?	URL	0
9.	Are none of the above true?	U	0

Fig. 1. Rater's classification guide.

[anterior or posterior; one or two lobes affected; laterality; lesion type (tumour, vascular, other)] by multiple regression revealed that only the anterior/posterior classification was a significant predictor of error score [t-ratio 2.72, $P<0.01$].

Analysis of variance of each subjects' mean latencies on section B (log transformed) according to the three main groups of error type (e.g. completions, semantically connected or unconnected) described above revealed no significant differences, despite some tendency towards longer response latencies in patients with frontal lobe involvement. In other words, none of the groups showed significantly different response latencies on those trials where they produced a word which was related to the sentence, thus breaking the task requirements. The group × error type latencies interaction, similarly, did not approach significance.

Responses revealing the use of strategies. Further analyses were carried out on the third category of responses (i.e. production of words unrelated to the sentence, as requested). Observation of the performance of controls suggested that two strategies were used frequently when the subject was performing the task. The first method for producing unrelated words was to look round the room, choosing objects within sight and name these in response to the stimulus sentences. The second method was for the subject to generate a semantic category and choose members of that category as responses. (In both cases the responses could be checked before being produced that they satisfied the test conditions.) Raters were therefore asked to classify responses unrelated to the sentence into one of four additional categories. First they had to judge whether the completion responses were objects normally found in a standard hospital office/room (UR). Second they had to judge whether the words were related to the *last* answer they gave (UL). Obviously on occasions responses might satisfy both conditions, and so there was a category (URL) for such responses. If the word was unrelated to the sentence, was not an object normally found in an office and was not semantically related to the last response, it was classified as U (see Fig. 1).

The unilateral anterior group, the posterior group and the control group differed in the proportion of their responses that fell in one or other of these categories,

Table 3. Error scores across groups (S.D. in brackets)

Group	Error score (max 45)
Controls	4.2 (5.7)
Posteriors	4.8 (4.8)
Anteriors	11.5 (11.3)
Bifrontals	19.4 (11.1)

Table 4. Response latencies by response type across groups (mean sec; S.D. in brackets)

Group	Completions	Semantically related	Unrelated	Response failures*
Controls	2.1 (2.0)	3.3 (4.6)	1.6 (2.4)	0
Posteriors	5.0 (8.2)	3.4 (3.5)	2.5 (3.6)	0
Anteriors	7.4 (9.0)	6.6 (9.0)	3.4 (5.2)	9
Bifrontals	6.9 (9.3)	7.7 (9.8)	8.7 (10.6)	6

*This column gives the numbers of occasions that a member of the group was unable to produce any response within 60 sec.

namely in being either semantically related to their last response, or being objects normally found in an office, or both [Kruskal–Wallis $H=6.53$, d.f. $=2$, $P<0.05$]. Multiple comparison procedures showed that the unilateral anterior group produced a significantly smaller proportion of these types of response (at the <0.05 level) than either the posterior group or the control group, which did not differ significantly from each other (see Table 5). The proportion of the responses that came from the strategy categories was, on average, less than half for the bifrontal group than for the unilateral anterior group.

Clearly however, as the anterior patients in section B showed a greater tendency to complete the sentence, whilst at the same time showing decreased evidence of strategy use, these characteristics may not be independent; there may be some characteristic of the task which leads to a trade-off between them. This possibility was explored using factor analysis. The two types of task failure (either straightforward completion, or the production of one of the four types of semantically connected response—SO, SA, SB, SC) were considered alongside the two forms of task success (either producing a 'non-strategy' unconnected response (category U) or one of the types of 'strategy' unconnected responses (categories UL, UR, URL). The numbers of instances of responses falling into these categories were considered across the unilateral patients and controls, using principal components analysis.

The results indicated an inverse relationship between occurrence of strategy-type answers and the two levels of task failure. Thus number of strategy responses loaded positively on the first principal component (eigenvalue 2.05, explaining 51% of the variance), and both number of completions and number of semantically connected responses loaded negatively. Number of non-strategy but satisfactory (category U) responses was comparatively unrelated to this component. However they were positively loaded on the second principal component (eigenvalue 1.3, explaining a further 32% of the variance), while strategy answers (UR, UL, URL), by contrast, had a negative loading. The two types of error responses were not loaded on this factor (see Table 6).

The factor loadings suggest a complex relationship between strategy use and completions of the sentence (i.e. failure of response suppression) in the Hayling task. Removing the variance explicable by group from the raw data by regression techniques before the principal component analysis is carried out did not alter the factor structure, suggesting that this complex causal relationship is a characteristic of the specific demands made by the Hayling task rather than being a characteristic of patients who are unable to perform it well. There was, however, a significant relationship between the *first* principal component factor scores (derived for each individual) and group [$F=5.74$, d.f. $=2, 91$, $P<0.005$] with this effect largely attributable to the anterior/other group comparison [t-ratio -3.15, $P<0.005$]. In other words, the performance of the anteriors was characterised by a combination of decreased strategy use together with simultaneous increase in responses

Table 5. Percentage responses in each category across groups

	C	SO	SA	SB	SC	UR	UL	URL	U*
Cont	4.2	2.7	2.5	1.5	5.8	31.5	5.0	1.7	45.2
Posts	4.7	3.7	2.7	3.7	9.6	27.2	8.9	1.7	37.5
Ants	16.9	8.2	3.4	3.1	10.9	24.8	2.7	1.1	27.7
BiFr	32.2	11.0	6.7	3.1	12.5	8.2	3.9	0.4	22.7

Section totals

	Completions	Semantically related	Unrelated
Cont	4.2	12.5	83.3
Posts	4.7	19.8	75.3
Ants	16.9	25.7	56.2
BiFr	32.2	32.5	35.3

*See Fig. 1 for explanation of categories.
Cont = controls; Posts = unilateral posteriors; Ants = unilateral anteriors; BiFr = bifrontals.

Table 6. Summary of principal component loadings across response types

		PC 1	PC 2
Failure type	Completions	−0.54	NS*
(Error score 1 or 3)	Semantically connected	−0.57	NS
Success type	Strategy responses	0.57	−0.82
(Error score = 0)	Non-strategy response	NS	0.50

*Indicates factor loading <0.3.
NS = not significant.

semantically connected to the sentence, thus breaking the task rules (see Tables 5 and 6).

Response latencies in both sections of the task

The response latencies of all groups were significantly longer on section B, compared with section A [paired t-tests, all groups $P<0.005$]. A one-way ANOVA of the unilateral patients and controls revealed differences in response latencies significant at the $P<0.01$ level [$F=6.39$, d.f. = 2, 91], with Tukey–Kramer *post-hoc* comparisons showing the anterior-control comparison to be significant at the $P<0.01$ level; none of the other comparisons were significant.

The effects in the unilateral patients, of lesion size (affecting one or two lobes), laterality (left or right hemisphere), location (anterior or posterior) and type (tumour, vascular or other) were investigated using multiple linear regression (backward elimination). Of these factors, only the anterior/posterior classification was a significant predictor of response latencies in section B [$F=4.73$, d.f. = 1, 72, $P<0.05$]. As with the performance on section A, the bifrontal group were significantly slower than the unilateral anterior group on this section [following log transformation: $t=2.92$, d.f. = 36, $P<0.01$] (see Table 2).

One patient (with a right frontal lesion) was unable to produce a word within 60 sec in eight of the 15 trials. A number of other patients with frontal lesions failed to produce words within the time limit on single trials (such a failure to respond, which was never seen in section A, was scored as a 'completion' for the error score analyses). No patient with a posterior lesion or a control ever showed this pattern.

B–A latencies. As a measure of the difference that the experimental manipulation in section B made, the section B latencies minus the section A latencies were considered for each subject (see Table 2). This presumably represents the additional 'thinking time' that is required in having to produce a novel word rather than a straightforward sentence completion. Obviously this serves to remove the possible confounding factor of initiation problems when considering the latencies in the response suppression condition.

The B–A times for the unilateral anterior, posterior and control groups were significantly different [$F=4.35$, d.f. = 2, 91, $P<0.01$], with Tukey–Kramer *post-hoc* comparisons showing the anterior-control differences to be significant at the $P<0.01$ level. None of the other comparions was significant. Moreover, the correlation between the B–A latencies in the unilateral patients (i.e. suppression condition minus initiation condition latencies) and the proportion of responses revealing the use of strategies was -0.45, which is significantly different from zero [$P<0.01$, 2-tailed]. In other words, those patients that showed longer thinking times tended to produce fewer strategy responses.

Overall relation between performance on section A and section B

Correlations between performance on section A (response latency) and section B (error score) were considered separately for each group. The controls and posteriors showed correlations between the two sections of the test which were significantly different from zero [controls: $r=0.593$, $P<0.01$; posteriors: $r=0.389$, $P<0.05$]. This was not the case for the unilateral anteriors, where performance on section A appeared unrelated to performance on section B [$r=0.191$, $P=0.2$]. This pattern was confirmed by considering the scores of individual patients. Using a strict criterion for a dissociation between tasks as being where the subject performs with 1 S.D. of the mean of the control group on one measure, but worse than 3 S.D. below their mean on another [see 19], four patients with unilateral anterior lesions, one bifrontal and one with a posterior lesion showed this pattern for good performance on section A with poor section B performance, and one patient with a unilateral anterior lesion showed the opposite pattern (poor A with good B performance). Such low correlations in performance might occur if the reliability of the measure were low. However when the Spearman–Brown split-half reliabilities for section A and section B (0.63 and 0.78, respectively) were taken into account, there were individual cases (in the anterior group only) who exhibited individual difference scores (following the procedure outlined in [1] pp. 136–137, which calculates the likelihood of occurrence of a given magnitude of discrepancy between a subject's performance on two different tests) that were significant at the 0.01 level in favour of each section of the task. In addition, one bifrontal showed a significant individual difference score in favour of section A. Thus it seems that performance on section A may be independent of performance on section B.

On the correlation analysis the bifrontal group showed a different pattern from the unilateral anteriors, with performance on the two sections of the test being highly related [$r = 0.634$, $P < 0.01$]. This difference in the correlations between the unilateral anteriors and bifrontals is significant at the 0.05 level. The bifrontal–unilateral anterior difference may suggest differences in laterality of function between the two sections of the test. However, within the unilateral anteriors there was no effect at all of laterality of lesion on either parts of the Hayling test (laterality of lesion explaining 0% of the variance in a regression analysis for each section). It would appear instead that the bifrontals tended to show more uniform impairments across the test (10/17 bifrontals were more than 2 S.D.s below the controls on section A, and 11/17 were at least 2 S.D.s below the controls on section B).

Discussion

This study has two main findings. The first is that, compared with a group of age- and IQ-matched patients with lesions to other parts of the brain (the posterior group), patients with lesions that included the frontal lobes (the anterior group) showed deficits on a measure of response initiation. The second finding is that the anterior group were also poorer than the posterior group on a task which involved suppression of the habitual response. In subsidiary analyses it was shown that a bilateral anterior group had significantly greater deficits than the unilateral anterior group.

In comparision with the patients with posterior lesions, the patients with anterior lesions performed in a qualitatively poorer fashion. They made significantly more straightforward completions of the sentence, and those answers they gave which were not completions nevertheless had a higher number judged to be semantically related to the sentence.

Turning to the latency of responding, the anterior group were significantly slower than the posterior group in section A. When the latencies in the two sections were compared, all groups took longer in section B than section A. Since all characteristics of the task are the same in the two sections except that the subject must produce a word unconnected to the meaning of the sentence in section B, latency B minus latency A presumably gives a measure of the thinking time in section B. The two unilateral lesion groups did not differ significantly on this measure.

In the context of the anterior group's qualitatively worse performance this means that the anterior group's thinking time was used less effectively in fulfilling the task requirements. However, although the unilateral anterior group's performance was poorer than that of the posterior group on two different aspects of the task—the overall error score on part B and the speed of responding on part A—the relationship between performance on the two sections was not the same for all groups. Thus whilst performance on section A was related to section B performance for the controls and posteriors, this was not the case for the unilateral anteriors. The bifrontal group, by contrast, was generally poor on both sections of the test.

One possible explanation of the anterior patients' deficits on the Hayling test is that they might be suffering from a greater degree of generalised cognitive impairment than the posterior group. However, this can be rejected as the two groups were matched for performance on a standard neuropsychological measure of general intellectual functioning (the WAIS).

A second possible explanation of the results might be that the anterior group performed the tasks badly due to distractibility or to motivational problems. However, this seems implausible since the anterior group did not perform other neuropsychological tests poorly (e.g. WAIS). An explanation of the results therefore needs to be based upon the way the tests employ particular processing resources or abilities that are damaged in patients with lesions primarily in the frontal lobes.

Perret [16] suggested that the high correlation in his study between performance on the Verbal Fluency test (VFT) and the Stroop task suggests that both VFT and the Stroop task tap the same process—suppression of the habitual response. The alternative hypothesis was that VFT requires good initiation or verbal planning abilities and the Stroop requires a different process, the suppressing of a prepotent response, but that initiation and suppression abilities tend to be damaged together because they are part of the same system or that they tend to be impaired together due to the anatomical proximity of the cortical areas subserving these processes.

However, the findings of the present study suggest that initiation or verbal planning abilities and response suppression may be impaired singly in unilateral anteriorly lesioned patients. These findings seem unlikely to be due to any overall lack of reliability in either a response time measure *per se* or in the error score in part B, as the error score correlated with (B–A) response times at the 0.01 level. In addition, when reliability was taken into account statistically, individual anterior patients still showed significant individual difference scores [1] on the two measures. In any case, low task reliability is unlikely to be a good explanation of the double dissociations noted within single cases [see 9, p. 158]. Instead, the most likley explanation of the present results is that the two processes are separable and that where both processes are impaired, they are so because the cerebral areas which subserve such processes are damaged together.

Why then should there be a positive correlation in the bifrontal group? One possibility is that medial frontal lesions, which would of course be almost inevitably produced in this group, give rise to general initiation problems which affected performance on both sections of the task. An alternative explanation could be given in terms of lesion size effects. Neither of these accounts are likely to be satisfactory in explaining the positive correlation in the posterior group, however. Here it is more plausible that the correlation is a result of deficits in the processing requirements of the task which are incidental to the main focus of this study. For instance, mild sub-clinical aphasic problems would be expected to affect sections of the task roughly equally. No doubt many other forms of disorder would act similarly.

However, there is a complicating factor for the preceding discussion. More detailed examination of performance on section B shows that an important factor in the way the different groups performed in section B is how well they employed strategies appropriate to the task. A failure to utilise an appropriate strategy when dealing with a novel task has often been treated as one aspect of the patient's difficulties [e.g. 13, 20]. However, it has been rare for the utilisation of appropriate strategies to be actually measured in frontal lobe group studies, the only case known to us being the study of Owen *et al.* [15].

In the present situation control subjects frequently employed some heuristic which enabled them to generate a word likely to be unrelated to the sentence stem. Among the many possible heuristics available two strategies were commonly used by control subjects. One was to pick objects from the surroundings, and the other was to generate a semantic category, and then choose exemplars of this category as responses, maybe after checking that it produced an appropriate result. A strategy measure was therefore employed of the frequency with which subjects produced the names of items typically found in hospital offices or words related to the previous response. The anterior group demonstrated a greater tendency towards completion of the sentence, whilst simultaneously showing a decrease in responses showing evidence of strategy use.

One possibility is that the frontal group might be unable to inhibit inappropriate responses and thus differed on the strategy measure because they did not have the time to operate the strategy. In fact there was a significant negative correlation between B–A latencies and the production of strategy-indicating responses. The patients who did not employ strategies effectively took relatively longer on task B than task A. These results therefore do not support the lack of time hypothesis. Instead they raise the possibility that the cause of the anterior group's poor task performance was their inability to acquire or realise an appropriate strategy and not being merely due to a deficit in response suppression abilities. In performing the Hayling task there would then be a necessary reciprocal causal link between strategy generation and response suppression abilities. Indeed it is plausible that the first factor derived in the principal components analysis reflects this reciprocal link.

The extent to which the factor of strategy use is critical in other situations where a frontal deficit in prepotent response suppression has been held to be important e.g. Stroop [16] or Wisconsin Card-Sorting [11] remains to be investigated. In the present situation, however, the critical unilateral anterior group show no significant relation between their performance on section A and that on section B. There appears to be little necessary relation other than the possible anatomical contiguity of their material substrates, between processes that concern strategy generation and realisation and those concerned with initiation or verbal planning that are involved in straightforward sentence completion.

Acknowledgements—We would like to thank Professor E. K. Warrington for providing the facilities for this research to be carried out, and two anonymous referees for their helpful comments on an earlier draft. This research was supported by a grant from the U.K. Medical Research Council.

References

1. Anastasi, A. *Psychological Testing*. Macmillan, New York, 1990.
2. Benton, A. L. Differential behavioural effects in frontal lobe disease. *Neuropsychologia* **6,** 53–60, 1968.
3. Bloom, P. A. and Fischler, I. Completion norms for 329 sentence contexts. *Mem. Cognit.* **38,** 631–642, 1980.
4. Costello, A. de L. and Warrington, E. K. Dynamic aphasia: The selective impairment of verbal planning. *Cortex* **25,** 103–114, 1989.
5. Crawford, J. R., Allan, K. M., Besson, J. A. O., Cochrane, R. H. B. and Stewart, L. E. A comparison of the WAIS and WAIS—R in matched UK samples. *Br. J. clin. Psychol.* **29,** 105–109, 1990.
6. Drew, E. A. Go–no go learning after frontal lobe lesions in humans. *Cortex* **11,** 8–16, 1975.
7. Kleist, K. Gehirnpathologie. Barth, Leipzig, 1934.
8. Luria, A. R. *Traumatic Aphasia.* Mouton, The Hague, 1970.
9. McNemar, Q. *Psychological Statistics*. John Wiley, London, 1964.
10. Miller, E. Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *Br. J. clin. Psychol.* **23,** 53–57, 1984.
11. Milner, B. Effects of different brain lesions on card sorting. *Arch. Neurol.* (Chicago) **9,** 90–100, 1963.
12. Milner, B. Brain mechanisms suggested by studies of the temporal lobes. In *Brain Mechanisms Underlying Speech and Language*, C. W. M. Whitty and O. L. Zangwill (Editors), pp. 122–145. Grune and Stratton, New York, 1967.

13. Milner, B. Petrides, M. and Smith, M. L. Frontal lobes and the temporal organization of memory. *Hum. Neurobiol.* **4,** 137–142, 1985.
14. Nelson, H. E. and O'Connell, A. Dementia: The estimation of pre-morbid intelligence levels using the new adult reading test. *Cortex* **14,** 234–244, 1975.
15. Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E. and Robbins, T. W. Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia* **28,** 1021–1034, 1990.
16. Perret, E. The left frontal lobe in man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia* **12,** 323–330, 1974.
17. Ramier, A. M. and Hecaen, H. Role respectif des atteintes frontales et de laterallisation lesionelle dans les deficits de la "fluence verbale". *Rev. Neurol. (Paris)* **123,** 17–22, 1970.
18. Shallice, T. Specific impairments of planning. *Phil. Trans. R. Soc. Lond.* **B298,** 199–209, 1982.
19. Shallice, T. *From Neuropsychology to Mental Structure.* Cambridge University Press, New York, 1988.
20. Shallice, T. and Evans, M. E. The involvement of the frontal lobes in cognitive estimation. *Cortex* **14,** 294–303, 1978.
21. Verfaellie, M. and Heilman, K. M. Response preparation and response inhibition after lesions of the medial frontal lobe. *Arch. Neurol.* **44,** 1265–1271, 1987.
22. Wechsler, D. *Wechsler Adult Intelligence Scale.* The Psychological Corporation, New York, 1955.

Appendix

Hayling sentence completion test—raters' instructions and examples of test items

[Raters were also given the 'decision tree" shown in Fig. 1.]

The sheets you have been given have two sets of 15 sentences on them. We are only interested in scoring the SECOND SET of sentences.

Each subject had each sentence read out to them in turn, and the response the subject made is written at the end of that sentence. We would like you to rate each of these words according to the "decision tree" below.

Examples:

C rating: (a straightforward completion of the sentence)

The captain wanted to stay with the sinking	BOAT or SHIP
He bought them in the sweet	SHOP or STORE

SO rating: (an opposite of what might be expected)

They went as far as they	COULDN'T
The whole town came to hear the major	REMAIN SILENT

SA rating: (a word obviously semantically connected to the subject of the sentence)

Most sharks attack very close to	FISHES or SEA or BITE
She called the husband at this	PHONE or WIFE or DIAL

SB rating: (a word showing an obvious semantic connection to the word(s) that you would expect to appear at the end of the sentence)

None of the books made any	UNDERSTANDING
Most cats see very well at	DAWN or DUSK or WINTER

SC rating: (a word which makes vague sense at the end of the sentence but which makes the sentence ludicrous. Also swear-words, obscenities or another inappropriate word.)

The whole town came to hear the major	FART
The dog chased our cat up the	TOWER OF PISA

UR rating: (a word that is completely unconnected to the sentence *and* which might reasonably be expected to be found in a normal office room—including office fixtures and fittings and clothes that someone in an office might wear)

The dough was put in the hot	COATHANGER/PAPERCLIP
Jean was glad the affair was	DESK/CHAIR/WINDOW/DOOR

UL rating: (a word that is unconnected to the sentence *and* which is semantically connected to their response to the last sentence. NB—cannot apply to sentence 1)

Most cats see well at	TALK (where previous answer was CONVERSATION)
The whole town came to hear the major	SANDCASTLE (where previous response was SEA)

URL rating: this is where both UL and UR are true.

U rating: a word that is unconnected to the sentence and which is not connected to the subject's response to the previous sentence and you would not generally expect to find in an office.