Pergamon

0028-3932(95)00165-4

Neuropsychologia, Vol. 34, No. 8, pp. 793–808, 1996
Copyright © 1996 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0028-3932/96 $15.00+0.00

False recognition and the right frontal lobe: A case study

DANIEL L. SCHACTER,*† TIM CURRAN,† LISSA GALLUCCIO,†
WILLIAM P. MILBERG‡ and JULIANNA F. BATES§

†Department of Psychology, Harvard University, Cambridge, Mass., U.S.A.; ‡Geriatric Research Education and Clinical Center and West Roxbury Department of Veterans Affairs Medical Center and §Center for Morphometric Analysis, Department of Neurology, Harvard Medical School and Massachusetts General Hospital, Cambridge, Mass., U.S.A.

(*Received* 15 *January* 1995; *accepted* 1 *October* 1995)

Abstract—We described a patient, BG, who exhibited a striking pattern of false recognition after an infarction of the right frontal lobe. Seven experiments document the existence of the phenomenon, explore its characteristics, and demonstrate how it can be eliminated. BG showed pathologically high false alarm rates when stimuli were visual words (experiments 1 and 4), auditory words (experiment 2), environmental sounds (experiment 3), pseudowords (experiment 5), and pictures (experiment 7). His false alarms were not merely attributable to the semantic or physical similarity of studied and non-studied items (experiments 4 and 5). However, BG's false recognitions were virtually eliminated by presenting him with categorized stimuli and testing him with new stimuli from non-studied categories (experiments 6 and 7). The results suggest that BG's false alarms may be attributable to an over-reliance on memory for general characteristics of the study episode, along with impaired memory for specific items. The damaged right frontal lobe mechanisms may normally support the monitoring and/or retrieval processes that are necessary for item-specific recognition. Copyright © 1996 Elsevier Science Ltd.

Key Words: false recognition; frontal lobes; episodic memory.

Introduction

Research concerning the role of the frontal lobes in memory has occupied an increasingly prominent position in neuropsychological research. Studies of patients with frontal-lobe damage have implicated these regions in memory for temporal order information [5, 26–28, 47], source memory [7, 19, 42, 48], and various other aspects of encoding conditions and retrieval [10, 23, 49, 53; for reviews see Refs. 38, 46, 52]. Recent neuro-imaging studies, too, have underscored the importance of frontal regions in human memory functions [33, 40, 44, 53].

Important insights into the role of frontal-lobe regions in memory have also been provided by observations concerning certain kinds of memory distortions, or false memories, in patients with frontal-lobe damage. For example, a number of investigators have argued that confabulation is closely associated with frontal-lobe damage [20, 30, 51], although additional damage to nearby basal forebrain structures may be necessary to produce extensive confabulation [11]. In addition, recent reports indicate that frontal-lobe damage is associated with the phenomenon of false recognition, where patients incorrectly claim to recognize distractor or lure items that had not appeared previously in an experiment. Delbrug-Derousne *et al.* [9] and Parkin *et al.* [32] have each reported case studies of patients with ruptured anterior communicating artery aneurysms and associated frontal-lobe damage, who produced an unusually high number of false alarm responses on recognition tests. Moreover, these false recognitions were accompanied by high confidence.

False recognition is theoretically important because it may provide clues concerning the nature of encoding and retrieval processes in episode memory [4, 15–17, 37, 41, 56]. Moreover, phenomena of false memory have recently assumed great practical importance in controversies concerning the accuracy of recovered memories of childhood [24]. Scientific knowledge concerning the neural basis of false memory is meager [39, 41], and neuropsychological investigations can further illuminate this theoretically and practically important issue.

---

* Address for correspondence: Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138, U.S.A.; Fax: 617-496-3122.

In this paper we describe a patient, BG, who exhibited a striking pattern of false recognitions after an infarction of the right frontal lobe. We report seven experiments that document the existence of the phenomenon, explore its characteristics and demonstrate how it can be eliminated. We propose an account of BG's false memories that appeals to differing levels of representation in episodic memory, and to the impairment of retrieval and monitoring processes that may be specifically associated with right frontal-lobe regions.

Experiment 1

To examine recognition memory in BG, we initially used a simple levels-of-processing paradigm. BG and matched control subjects studied a series of target words under either semantic encoding conditions (subjects rated how much they like each word) or non-semantic encoding (subjects counted the number of T-junctions—the places where two lines meet—in each word). After a delay of several minutes, old and new words were presented in a recognition test. To examine the qualitative nature of recollective experience, we used the remember/know procedure developed by Tulving [54] and Gardiner *et al.* [13]. When BG and control subjects responded that a word had appeared on the study list, they were asked to indicate (a) whether they possessed a specific recollection of having encountered the word previously (a "remember" response), or (b) whether they just "knew" that the item appeared on the list, even though they did not have a specific recollection of having encountered it (a "know" response).

Although remember and know responses probably overlap considerably with high- and low-confidence judgments, a substantial experimental literature indicates that remember and know responses are not merely substitutes for high- and low-confidence responses, but instead appear to tap qualitatively different aspects of recollective experience [reviewed by Ref. 13]. By examining remember and know responses, we hoped to shed light on the qualitative nature of both true and false recognition responses.

Method

Case report. BG is a 66-year-old right-handed man who was admitted to the hospital in December 1993 complaining of diarrhea, cough and lethargy, and pleuritic chest pain. His neurological examination was normal. He was diagnosed with pneumococcal pneumonia and placed on antibiotics. On the third day after admission BG's behavior became "inappropriate", agitated, and confused. He continued to deteriorate medically, becoming somnolent, and was transferred to the medical intensive care unit where he was given additional antibiotic therapy. Approximately 2 weeks later, as his pneumonia began to improve and he became more alert, his neurological examination was notable for left arm and leg weakness. A CT scan was ordered, which revealed a right frontal-lobe infarction in the territory of the middle cerebral artery, possibly of embolic origin, that was assumed to have occurred sometime during his admission. In October 1994, a magnetic resonance imaging (MRI) scan was performed (Fig. 1). BG's lesion extended back to the center sulcus, primarily involving motor and premotor cortex (Brodmann's areas 4 and 6). The lesion involves the rostral two-thirds to three-quarters of the precentral gyrus, as well as a bit of the postcentral gyrus, particularly at the rostral (lateral, inferior) tip of the central sulcus. The inferior frontal gyrus, pars opercularis, is also affected. The upper bank of the Sylvian fissure is completely destroyed, over ~20 mm length (anterior–posterior) starting at about the level at which the temporal lobe is first attached, as seen in coronal section. The insula appears to be intact. Subcortically, the ventricles are enlarged, and the neostriatum is affected, particularly the body and tail of the caudate, which are severely reduced in the right hemisphere. Portions of the thalamus were also involved, particularly the anterodorsal, anteroventral and ventrolateral nuclei, with lesser involvement of the midline mediodorsal nucleus. In the left hemisphere, a much smaller infarct was observed in the caudal and ventral portion of the putamen (in the region of anterior choroidal artery distribution). Basal forebrain structures such as the nucleus accumbens and substantia innominata appear to be intact in both hemispheres.

BG possesses a masters degree (18 years of education), and he was employed for most of his working life in financial managerial positions. He was last employed in this capacity in 1988, when he retired after his company was sold. His medical history includes hypertension, coronary artery disease and possibly heavy alcohol use.

The present experiments were conducted between April and July of 1994. A summary of BG's neurophsychological assessment is presented in Table 1. On the Wechsler Adult Intelligence Scale-Revised (WAIS-R), BG received a verbal IQ of 99, a performance IQ of 93, and a full scale IQ of 96. Although these scores fall within the average range, they are somewhat below what we would estimate to be his high average level of premorbid intellectual functioning, based on his education, his occupation and his reading score of 112.96 on the American National Adult Reading Test (ANART). This discrepancy between estimated premorbid intelligence and current IQ suggests that BG has suffered at least a mild degree of general intellectual decline.

BG's MQ (memory quotient) of 100 is consistent with his present level of general intellectual functioning, with evidence of some slight loss of information across a retention interval (delay score = 93). However, his attention score of 84 suggests mild to moderate impairment. In addition, a mild to moderate impairment was evident in his initial and delayed recall of the list of words from the California Verbal Learning Test (CVLT; initial recall 8–10/16 items), though delayed recognition memory was normal (15/16 items with two false positives). The finding that BG made relatively few false positive responses on this test, in contrast to the experiments that we report, probably reflects the fact that target lists are presented repeatedly on the CVLT, thereby creating ceiling effects on subsequent recognition performance for all but the most severely amnesic patients. BG's copy of the Rey–Osterreith Complex figure was unimpaired relative to his age group (33/36). Although he showed a slight decline in recall of this figure with short (20/36) and long (16/36) delay, his performance on these measures was also unimpaired relative to his age group.

On the Warrington Recognition Test, which assesses two alternative forced-choice recognition memory for words and faces, BG's performance on the word subtest was normal (63%ile), but he showed severe impairments on the face subtest, scoring at the chance level (<5%ile). These results contrast with BG's relatively normal performance on a difficult test of facial perception, the Benton Face Recognition Task, where he was correct on 43/54 items. However, BG did show moderate

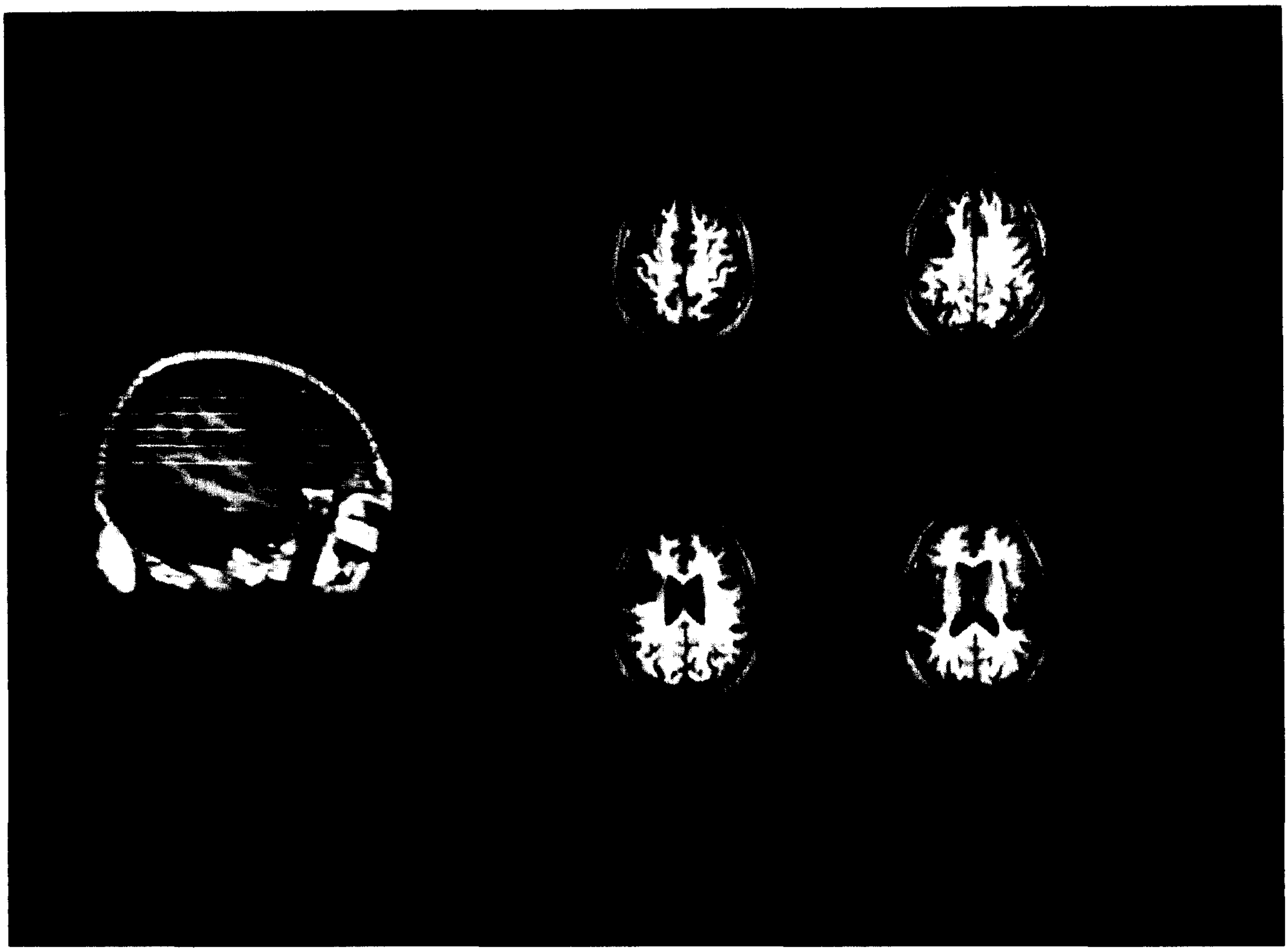

Fig. 1. Lesion localization for BG in parasagittal and axial (horizontal) planes as revealed by MRI. The entire precentral gyrus is affected, as seen in the parasagittal view. The superior–inferior and lateral–medial extents of this lesion can be appreciated in a series of axial slices which proceed from superior to inferior (levels A–D). These images are derived from a 3-D-SPGR T1-weighted spoiled gradient echo pulse sequence performed on a GE 1.5 Tesla Signa MR system (TR = 40, TE = 5, flip angle = 400, FOV = 24 cm, slice thickness = 3.0 mm). Slices A, B, C and D are spaced equally, 10 mm apart in each case. See text for further details of lesion analysis.

to severe impairments on the Benton Line Orientation Test (7/30 correct).

Clinically, BG exhibited little difficulty remembering his recent experiences and did not engage in extensive spontaneous confabulation. We administered confabulation batteries that have been developed by Dalla Barba [8] and Moscovitch [30], and he exhibited no evidence of confabulation on either battery.

BG's adjusted score on the phonemic word list generation task (FAS) placed him in the 25–29%ile, suggesting a moderate impairment. In addition, he was also able to generate only 2/6 categories, on the Wisconsin Card Sorting Test (77 total errors and 38 perseverative errors), which also suggests at least a moderate impairment. Finally there was no clinical evidence of language impairment. BG's spontaneous speech was fluent and well articulated with no evidence of word finding difficulties. Indeed, his performance on the Boston Naming Test performance was entirely normal (57/60 correct).

Control subjects. Eight control subjects who were closely matched to BG according to age (65.6 years) and years of education (17.0) participated in the experiment. Eight controls with similar characteristics were also matched to BG in each of the subsequent six experiments. Four of the control subjects participated in all seven experiments, and others participated in varying numbers of experiments; a total of 11 control subjects were needed to yield eight controls/experiment. Because experiments 2 and 3 involved auditory presentation and/or testing, control subjects were given audiometric assessment and were matched to BG, who exhibited a normal loss of sensitivity to high frequencies that is often seen in elderly adults.

Materials and design. The experimental stimuli consisted of 144 concrete English nouns. The words were divided into six subsets of 24 which were roughly equated for word length (Mean = 5.1, S.D. = 1.4, range = 3–8) and frequency of usage (Mean = 33.3, S.D. = 22.8, range = 10–99; [22]). Word subsets were randomly assigned to a particular experimental condition (liking, T-junction, non-studied). Another 36 words with similar characteristics were used as buffer and practice items.

Subjects completed two study test blocks, with study task manipulated on a within-subjects basis. Each study list was divided into two 30-item sublists. In the first block subjects provided liking ratings for the first sublist and counted T-junctions for the second sublist. The study task order was reversed for the second study-test block. Study sublists each contained 24 experimental words surrounded by three-word primacy and recency buffers. Test lists contained 24 words from the liking task, 24 from the T-junction task and 24 non-studied words. Order of words on the test list was random with the constraint that no more than three words appeared consecutively from the same condition.

Words were presented on a MacIntosh Powerbook in uppercase, 24-point Geneva font. Responses were written on response forms by the subject.

Table 1. Neuropsychological assessment of patient BG

Test	Score	Percentile (where available)
Wechsler Adult Intelligence Scale-Revised		
Verbal IQ	99	47.4
Performance IQ	93	32.2
Full scale IQ	96	39.6
Wechsler Memory Scale-Revised		
General memory	100	50
Attention	84	14.6
Delay	93	32.2
Warrington Recognition Test		
Faces	25/50	<5
Words	44/50	63
California Verbal Learning Test		
Learning trials 1–5 of Monday list	8, 9, 8, 10, 8/16	●
Immediate free recall of Tuesday list	6/16	●
Short delay free recall Monday list	7/16	●
Short delay cued recall Monday list	9/16	●
Long delay free recall Monday list	7/16	●
Long delay cued recall Monday list	7/16	●
Long delay recognition Monday list		●
Hits	15/16	
False positives	2/16	
Benton Line Orientation Test	7/30	●
Benton Face Recognition Test	43/54	33–59 range
Rey Osterreith Complex Figure		
Copy	33/36	60
Immediate recall	16/36	●
Delay recall	20/36	●
Word List Generation (items per minute)		
FAS	12, 7, 10	25–29 range
Grocery list	20	●
Wisconsin Card Sort Test		●
Sorts	2/6	
Total errors	77	
Perseverative errors	38	
Boston Naming Test	57/60	●
American National Adult Reading Test	112.96	●

Procedure. Subjects first completed a short practice block to introduce the study and test procedures. The practice block included two four-word study lists (liking encoding task, then T-junction encoding task), followed by a 12-word test list consisting of eight studied and four non-studied words. After this practice segment, subjects completed the two experimental study-test blocks.

During the encoding tasks, words were presented for 4 sec each with a 1-sec interstimulus interval. Subjects were encouraged to respond to each word before the next appeared. In the liking task, subjects rated each word on a 5-point scale according to how much they liked its meaning—a rating of "1" was given to words that subjects strongly disliked, "3" to neutral words, and "5" to words that subjects strongly liked. In the T-junction task, subjects counted instances in which two lines within a letter intersected to form a T-shaped formation.

Prior to each test list, subjects performed an unrelated serial reaction time task for 2 min. After completion of the serial reaction time task, the recognition test was administered. Subjects were instructed in the use of the "remember", "know", and "new" responses using instructions adapted from Rajaram [36], and were required to summarize the instructions for the experimenter to ensure adequate understanding of them. Test lists were self-paced such that words appeared on the screen until subjects responded and initiated presentation of the next word by pressing the space bar on the keyboard. The three response options were displayed below the test word vertically: (R)emember, (K)now, (N)ew.

Results and discussion

Figure 2 presents the proportion of "remember" (R), "know" (K) and "old" (R+K) responses to new words, old words that appeared during the T-junction encoding task, and old words that appeared during the liking encoding task. For control subjects, the mean proportions are displayed for each kind of response in each of the three main conditions. In addition, we indicate the maximum and minimum proportion of responses observed in any individual subject in each condition in order to define the range of normal performance for purposes of comparison with BG.

The most striking feature of the data in Fig. 2 is that BG made many more false alarm responses (R+K) to new words than did controls: he made false alarms to 0.50 of non-studied words, whereas control subjects, on average, made false alarms to 0.17 of non-studied words. BG's false alarm rate fell well outside the range of control performance; the highest proportion of false alarms made by any one of the control subjects was 0.29. To provide a statistical comparison between BG's responses and those of control subjects, we used the non-parametric comparison of counts test described by Bennet and Franklin [2]. BG made significantly more false alarm responses than did control subjects, $P < 0.01$. Further inspection of subjects' responses to new items indicates that the difference in false alarm rates between BG and controls was entirely attributable to the R responses: BG claimed to "remember" 0.38 of new words, whereas controls provided "remember" responses to only 0.05 of the new words, $P < 0.01$. By contrast, BG and control subjects provided a virtually identical number of K responses to non-studied words.

It is possible that BG's elevated false alarm rate is simply a reflection of generally degraded memory for study list items: BG may have been unable to remember the study list items and instead offered haphazard guesses. However, inspection of responses to studied items (Fig. 2) indicates that this is not the case. In both tasks, BG's hit rate was substantially higher than his false alarm rate, thus indicating that he was not simply guessing randomly. After the liking task, the hit rates of both BG and controls approached ceiling levels (0.96), whereas after the T-junction task, BG's hit rate for old words (R+K) was higher than that of control subjects (0.81 versus 0.60, $P < 0.05$). We further assessed BG's recognition accuracy by applying the standard correction procedure in which the false alarm rate is subtracted from the hit rate. BG's corrected recognition score was much lower than that of controls after both the T-junction task (0.31 versus 0.43) and the liking task (0.46 versus 0.79), but this is entirely attributable to his elevated false alarm rate.

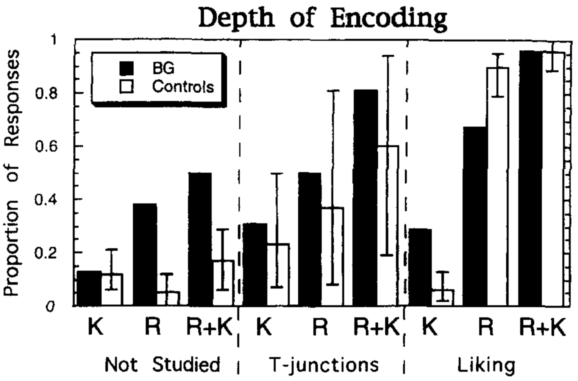

Fig. 2. Mean proportion of "remember" (R), "know" (K) words, and "old" (R+K) responses in each condition for BG and controls in experiment 1. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject.

In previous research examining the effects of levels of processing manipulations on R and K responses, it was found that relative to non-semantic encoding, semantic encoding selectively enhances the proportion of R responses to old words [12]. Both BG and control subjects exhibited this general pattern: the semantic encoding task produced more R responses than the non-semantic encoding task (Fig. 2).

BG's pattern of responses to studied items are potentially relevant to the high proportion of R responses that he made to non-studied words. It is possible that this false recollection phenomenon reflects a general bias on the part of BG to use the R response more frequently than control subjects. If so, then BG should also consistently make more R responses to studied words than control subjects do. However, this outcome was not observed: BG made non-significantly more R responses to studied words than did controls after the t-junction task, and significantly fewer ($P < 0.01$) remember responses to studied words than did control subjects after the liking task. Collapsed across the two study tasks, BG made slightly fewer R responses to studied words (0.59) than did control subjects (0.64). These results suggest that the high proportion of R responses that BG provided to non-studied words does not reflect a general bias on BG's part to use the R responses more often than control subjects. However, subsequent experiments indicate that BG's proportion of R responses to studied items (relative to control subjects) fluctuates considerably across conditions, and we will postpone until the General Discussion further consideration of whether his R responses to non-studied items reflect a generalized bias to use the R response.

In summary, experiment 1 has revealed a striking false recognition effect in patient BG. He made approximately three times as many false alarm responses as did control subjects, and he made over seven times as many R responses to new words as did the controls. The effect is not attributable to random guessing, nor does it appear to reflect a generalized bias to use the R responses more often than control subjects.

Experiment 2

The main purpose of experiment 2 was to determine whether we could replicate the false recognition phenomenon observed in experiment 1 under different conditions. For example, it is conceivable that the false recognition exhibited by BG reflects some sort of idiosyncratic response to the particular words that were assigned to the non-studied condition in experiment 1. Alternatively, BG's false recognitions may be tied in some way to the fact that both study and test were conducted in the visual modality. To address these possibilities and explore the generalizability of the results from experiment 1, we used a new set of stimulus materials and presented and tested all target words in the auditory modality.

Method

Materials, design and procedure. The basic design of experiment 2 was similar to that of experiment 1, except that only the semantic encoding task (liking rating) was used. The target materials were 96 common English words that were studied and tested in spoken form (length: Mean = 6.91, S.D. = 1.10, range = 5–10; frequency [22]: Mean = 17.14, S.D. = 28.83, range = 0–185). Two subsets of 48 words were randomly assigned to the studied or non-studied conditions. Each word was digitally recorded on a MacRecorder (Soundedit program) in one of six different voices (four female and two male); words were spoken at normal conversational levels. Words were presented by a Macintosh PowerBook through headphones. Subjects rated the 48 studied words with the same liking-rating task used in experiment 1, with 6 sec allowed for rating each word. After completing the study list, subjects completed the same 2-min serial reaction time task used in experiment 1. The 48 studied words and 48 non-studied words were randomly intermixed for the recognition task, and subjects were given remember–know test instructions along the lines described in experiment 1. Subjects were able to replay the test words as often as required in order to ensure that they heard each test word.

Results and discussion

As displayed in Fig. 3, the key outcome of experiment 2 is that BG made significantly more false alarms to non-studied items than did control subjects, $P < 0.05$. His overall proportion of false alarms (R + K), 0.27, was not as high as in experiment 1, but again fell outside the range of the eight control subjects. As in experiment 1, most of BG's false alarms were R responses, whereas most of control subjects' false alarms were K responses. BG provided R responses to 0.25 of non-studied words, a value that was significantly higher than the corresponding mean for control subjects (0.04; $P < 0.01$), and that fell outside the range of control responses.

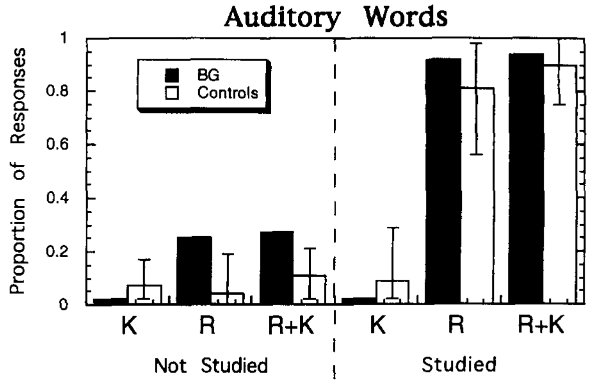

Fig. 3. Mean proportion of "remember" (R), "know" (K), and "old" (R + K) responses in each condition for BG and controls in experiment 2. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject.

BG's hit rate to studied items (0.94) was comparable to that of control subjects (0.90) and much higher than his false alarm rate. These observations indicate that BG possessed a good deal of accurate memory for study list items, and that his responses on the recognition test were not simply random guesses. As in experiment 1, BG's corrected recognition score was considerably lower than that of control subjects (0.67 versus 0.89), but this was entirely attributable to his elevated false alarm rate. In addition, BG exhibited a nonsignificant trend to provide more R responses to studied words than did control subjects (0.92 versus 0.81).

Experiment 2, then, has replicated the major false recognition findings from experiment 1 with a different set of stimulus materials and in a different sensory modality. Thus, BG's excessive false recognitions are not restricted to the visual modality nor to any particular materials. In experiment 3, we examined further the generality of the phenomenon.

Experiment 3

In both experiments 1 and 2, the target materials were familiar words. It is possible that BG's pathological false recognition is restricted to verbal materials. To examine whether BG would also provide excessive numbers of false alarms to non-verbal materials, we tested his memory for everyday environmental sounds.

Method

Materials, design and procedure. The target stimuli were 60 digitally recorded environmental sounds, such as a door closing, telephone ringing, baby crying, and so forth. Each recorded sound had a 5-sec duration. The set of 60 sounds was randomly divided into two subsets, one that was exposed both during the study and test phases of the experiment, and another that was exposed only during the test phase. During the study task, subjects rated how much they liked each sound. All other aspects of the design and procedure were identical to experiment 2.

Results and discussion

As shown in Fig. 4, BG once again exhibited an excessively high false alarm rate. Overall, he made false alarm responses to 0.47 of non-studied sounds, which was significantly higher than the mean false alarm rate of control subjects (0.20; $P < 0.01$), and also fell well outside the range of control performance. As in previous experiments, many more of BG's false alarms were R responses than K responses, whereas control subjects exhibited the opposite pattern. BG made R responses to 0.37 of non-studied sounds, whereas control subjects made R responses to 0.08 of non-studied sounds, $P < 0.01$; again, BG's performance fell outside the range of control subjects.

BG's overall hit rate to studied sounds (1.0) was higher than that of control subjects (0.85), $P < 0.05$. Nevertheless, because of his elevated false alarm rate, BG's corrected recognition score was still somewhat lower than that of control subjects (0.53 versus 0.65). Unlike in previous experiments, BG provided R responses to every studied item. Control subjects provided significantly ($P < 0.05$) fewer R responses (0.73) to previously studied sounds than did BG.

The results of experiment 3, then, indicate clearly that BG makes an abnormally high number of false recognitions to nonverbal sounds, and thereby demonstrate that his elevated false alarm rate is not tied specifically to verbal materials. Having observed the phenomenon in three separate experiments, we can now begin to address the critical question: why does BG exhibit pathological false recognition? One possibility is suggested by a theoretical account of false recognition in normal subjects that was advanced by Underwood [56] three decades ago. Underwood demonstrated that subjects could be induced to commit false alarms to words that are associatively related to previously studied items. He suggested that at the time of study, subjects make "implicit associative responses" to target words, and that these associations are confused with studied items on the recognition test. With respect to BG, he might be either more likely than control subjects to make implicit associative responses, or less able than controls to distinguish, at the time of test, between associations that were elicited during the study task and items that were actually encountered during the study task. Given previously noted evidence that patients with frontal-lobe damage exhibit source memory deficits, this latter possibility seems quite plausible.

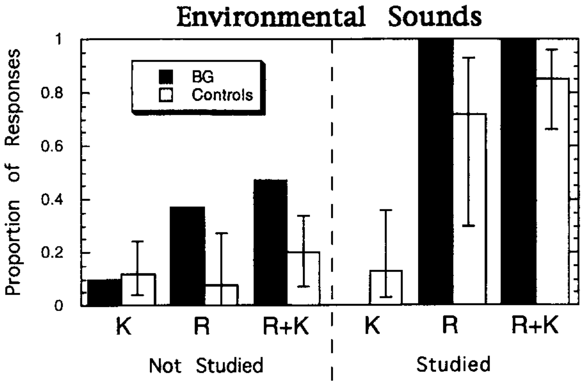

Fig. 4. Mean proportion of "remember" (R), "know" (K), and "old" (R + K) responses in each condition for BG and controls in experiment 3. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject.

Analysis of BG's false alarms in previous experiments provides some evidence consistent with this suggestion. For example, we examined the 18 non-studied items in experiment 1 that BG claimed to "remember". We were able to find an associatively related word in the prior study list for nearly all of these items. Thus, for instance, the non-studied word "cellar" was preceded on the study list by "basement", "chair" was preceded by "seat", and "cotton" was preceded by "wool". BG's false recognition responses may have been triggered by the associative overlap between these items, perhaps because he was less able than controls to distinguish between prior associative responses and prior study list items. In experiment 2, the words that BG falsely "remembered" were not related in any obvious way to any of the studied words. Although BG produced many more false alarms than did control subjects in this experiment, his false alarm rate was only about half of what it was in experiment 1. In experiment 3, many of the non-studied sounds were perceptually similar to, and thus highly confusable with, previously studied sounds, and BG's false alarm rate was nearly as high as in experiment 1.

These considerations raise the possibility that BG's elevated false alarm rate is largely or entirely attributable to his inability to distinguish between non-studied items that bear some associative or perceptual similarity to previously studied items. BG may claim to "remember" many of these items because he has in fact encountered associatively or perceptually similar items on the study list. On the other hand, the fact that BG still exhibited some pathological false recognition in experiment 2, where there was little relationship between studied and falsely recognized items, suggests that inter-item similarity may not account for all of BG's false alarm responses. To explore these issues, we specifically manipulated the associative (experiment 4) and perceptual (experiment 5) relationship between studied and non-studied words.

Experiment 4

In experiment 4 non-studied items were either associatively related or unrelated to studied words. Studied words were presented either once or three times, in order to determine whether frequency of exposure to an associate systematically affected false alarm rates. The major question addressed in experiment 4 was whether BG would exhibit pathological false recognition only in the conditions in which non-studied words bore an associative relationship to previously studied words.

Method

Materials, design and procedure. Presentation frequency (0, 1, or 3) and associative relationship of non-studied words to studied words (associated versus non-associated) were manipulated on a within-subjects basis. The experimental stimuli were 192 familiar words, consisting largely of concrete nouns. Ninety-six words were taken from Postman's [34] word association norms, consisting of 48 pairs of highly related associates (e.g., sofa/couch). Another 96 non-associated control words were selected to match these associates on a item-by-item basis for word length (associates: Mean = 5.56, S.D. = 1.28, range = 3–9; controls: Mean = 5.44, S.D. = 1.45, range = 3–10) and frequency of usage (associates: Mean = 91.71, S.D. = 133.33, range = 0–760; controls: Mean = 89.14, S.D. = 131.86, range = 0–715; [22]). Twelve words with similar characteristics were used as buffers. Buffers and non-associated words were selected such that none were obviously related to each other nor to any of the associated words. Each associate and its non-associated control was randomly assigned to one presentation condition (non-studied, one presentation, three presentations) and one of two study-test blocks.

Twenty-four words were presented once and another 24 were presented three times, yielding a list length of 96. Within each presentation frequency condition, 12 words were selected from among the associates and 12 were the corresponding (i.e., matched) non-associated controls. The study lists were divided into two 48-item sublists with a subject-paced rest break intervening. Each sublist was surrounded by three-word primacy and recency buffers, and words from each condition were equally distributed between the two sublists. Repetitions always occurred within the same sublist and were separated by at least five other items. Presentation duration was the same as in experiment 1 and liking ratings were given to all studied words.

Test lists included words from five different conditions: (1) 12 non-associated words that were studied once; (2) 12 non-associated words that were studied three times; (3) 12 non-studied words that were associated with words studied once; (4) 12 non-studied words that were associated with words studied three times; and (5) 24 non-studied words that were not associated to any of the studied words. The associated words that were presented during the study phase of the experiment never appeared on the recognition test. Order of items on the recognition test was random with the constraint that no more than three items appeared consecutively from the same condition. The test instructions and procedure were identical to those described for experiment 1. As in previous experiments, each study-test retention interval included the 2-min serial reaction time task.

Results and discussion

Figure 5 displays the results for non-studied words. Consider first the non-studied words that had no associative relation to previously studied words. As in previous experiments, BG's overall false alarm rate (0.17) was more than double the overall false alarm rate of the control subjects (0.08). However, unlike in previous experiments, one of the control subjects made more overall false alarms (0.21) than did BG, so the difference between BG and controls did not attain statistical significance. Nevertheless, as in previous experiments, BG made more R than K false alarm responses, whereas controls exhibited the opposite pattern. The proportion of R responses that BG gave to non-studied words (0.17) was again significantly ($P < 0.01$) higher than the proportion of R false alarms (0.01) made by control subjects, and was well outside the range of control responses. Thus, although the differences are not quite as large as in experiments 1–3, BG still exhibits evidence of pathological false recognition even when test items have no associative relationship to previously studied items.

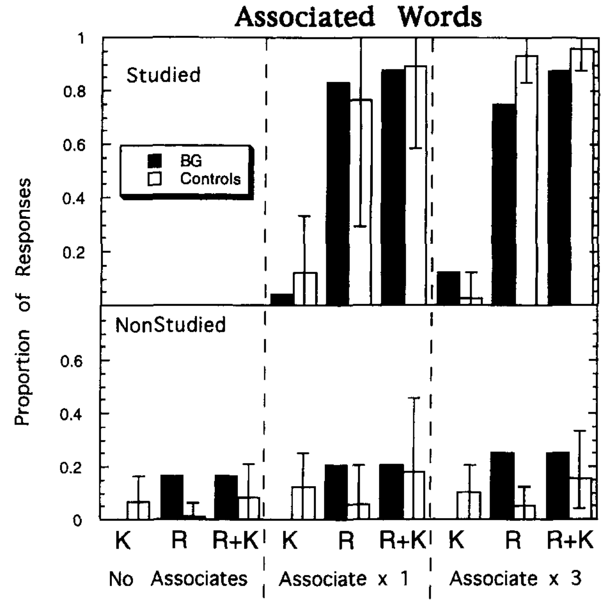

Fig. 5. Mean proportion of "remember" (R), "know" (K), and "old" (R + K) responses in each condition for BG and controls in experiment 4. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject. Studied words were presented once (× 1) or three times (× 3). Non-studied words were associated with no studied words (no associates), associated with a once-presented studied word (associate × 1), or associated with a thrice-presented studied word (associate × 1), or associated with a thrice-presented studied word (associate × 3).

Consider next false alarms to words that are strongly associated to study-list words. Focussing first on the condition in which associates were presented once during the study list (Fig. 5: "Associate × 1"), BG's performance is quite similar to the no-associate condition: He made false alarm responses to 0.21 of the non-studied associates, and all of these were R responses. However, control subjects now made nearly as many false alarms as BG (0.18), and BG's performance fell within the control range. BG still exhibited more R responses to non-studied words than did controls (0.21 versus 0.06), but one control subjects exhibited as many R false alarms as did BG, so the difference between BG and control subject did not attain statistical significance. Results from the condition in which associates of non-studied words were exposed three times during study list presentation yielded a generally similar pattern of results (Fig. 5: "Associate × 3"). BG made more overall false alarm responses than control subjects (0.25 versus 0.16), but two controls made more false alarms than BG, and the difference between BG and controls did not approach significance. However, BG made significantly ($P < 0.05$) more R false alarm responses than did any of the control subjects (0.25 versus 0.05).

BG's responses to previously studied (non-associated) words that were presented either once or three times were generally similar to those of control subjects (Fig. 5). For words presented once, the hit rates of BG and control subjects were nearly identical (0.88 versus 0.89); for words presented three times, BG's hit rate was non-significantly lower than that of controls (0.88 versus 0.96). When false alarm rates from the non-studied no-association condition were subtracted from the hit rates, BG's corrected recognition scores were again lower than those of control subjects (0.71 versus 0.81 for words presented once, and 0.71 versus 0.88 for words presented three times). BG made non-significantly more R responses to words presented once than did control subjects (0.83 versus 0.77), and made significantly fewer R responses to words presented three times than did controls (0.75 versus 0.93, $P < 0.05$).

Overall, the results of experiment 4 do not support the view that BG's pathological false recognition is attributable to his inability to distinguish between previously studied items on the one hand and associative responses to those items on the other. If BG's false recognitions were especially sensitive to associative overlap between study and test items, he should have exhibited much more false recognition than control subjects in the conditions in which associatively related words were studied, and little or no pathological false recognition in the no-associates condition. However, BG did exhibit evidence of pathological false recognition in the no-associates condition, although the absolute magnitude of the false recognition effect was rather modest. More importantly, there was little evidence that BG's false recognition was disproportionately inflated in the two conditions in which associatively related words were present. BG, like control subjects, made more false alarms to associatively related words than to unrelated words, and our analysis of experiments 1–3 also indicated a role for associative and possibly perceptual relatedness between studied and non-studied items. But the fact that BG exhibited some pathological false recognition of unrelated items in experiment 2 and in the no-associates condition of experiment 4, suggests that other factors are relevant to BG's overall pattern of false alarm responses.

Some insight into what one other such factor might be emerges from considering a salient property of the materials used in all previous experiments: the words used in experiments 1, 2, and 4 and the sounds used in experiment 3 were all in some sense familiar to BG prior to the experiment. The target words were all common items of English vocabulary that were presumably part of BG's pre-experimental lexicon. Similarly, the environmental sounds used in experiment 3 were examples of everyday sounds that people encounter frequently in everyday life. Perhaps BG's high false alarm rate is attributable to the fact that he has special difficulty distinguishing between the occurrence of a word or a sound in the laboratory and previous encounters with these words or sounds outside the laboratory. As noted in the Introduction, patients with frontal lobe damage often exhibit deficits in source and temporal memory, and the existence of such a deficit could have produced BG's elevated false alarm rates in experiments 1–4. We tested this hypothesis in experiment 5.

Experiment 5

If BG's pathological false recognition is attributable to an inability to distinguish between pre-experimental and experimental encounters with a target item, then it should be possible to eliminate the phenomenon by exposing BG to materials that he has never encountered previously. To accomplish this objective, we exposed BG to pronounceable non-words (pseudowords) and then tested his recognition for studied and non-studied pseudowords. Because BG has presumably never seen any of the pseudowords prior to the experiment, he should not experience confusion between pre-experimental and experimental encounters with these items. We also examined further BG's sensitivity to the relatedness of study and test items. Some of the non-studied pseudowords bore no physical resemblance to nonwords in the study list. In two other conditions, each non-studied pseudoword was physically similar to a study list item. In one of these conditions the study list was presented once and in the other condition it was presented three times.

Method

Materials, design and procedure. The design was identical to experiment 4 except that pronounceable pseudowords served as stimuli. Pseudowords were created by randomly replacing vowels from four- and five-letter words. Physically similar pseudowords were created by varying a single letter (e.g., spafe and spake). There were five test conditions: (1) 12 pseudowords that were studied once; (2) 12 pseudowords that were studied three times; (3) 12 non-studied pseudowords that were similar to pseudowords studied once; (4) 12 non-studied pseudowords that were similar to pseudowords studied three times; (5) 24 non-studied pseudowords that were dissimilar to the studied pseudowords.

The procedure was virtually identical to experiment 4 except for the encoding task. Subjects studied the pseudowords by rating their ease of pronunciation on a 5-point scale (1 = very difficult to pronounce and 5 = very easy to pronounce). Subjects were given a brief practice on the encoding task (10 words) and the serial reaction time test (30 sec) at the beginning of the session.

Results and discussion

Figure 6 displays the proportion of false alarms to non-studied pseudowords in the various experimental con-

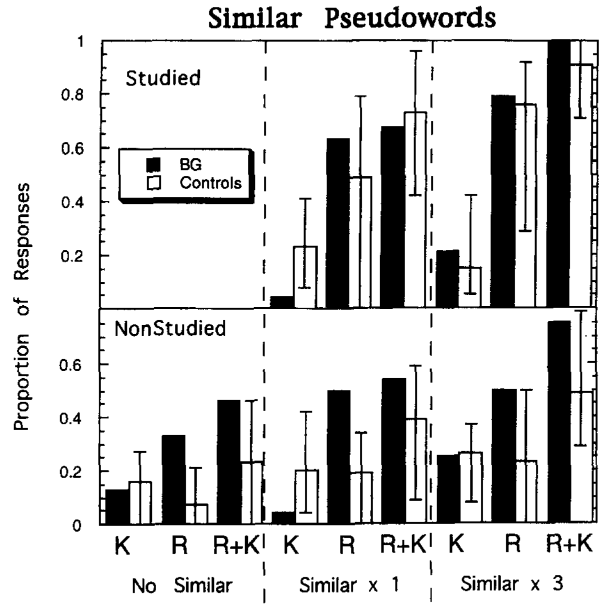

Fig. 6. Mean proportion of "remember" (R), "know" (K), and "old" (R + K) responses in each condition for BG and controls in experiment 5. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject. Studied pseudowords were presented once (× 1) or three times (× 3). Non-studied pseudowords were similar to no studied pseudowords (no similar), similar to a once-presented studied pseudowords (similar × 1), or similar to a thrice-presented studied pseudoword (similar × 3).

ditions. First, consider non-studied pseudowords that were physically dissimilar to study list pseudowords. BG's overall false alarm rate (0.46) was significantly ($P < 0.01$) higher than that of control subjects (0.23), although a single control made as many false alarms as did BG. As usual, BG made more R than K false alarms, and controls showed the opposite pattern. BG made many more R false alarms (0.33) than did controls (0.07), $P < 0.01$, and his proportion of R false alarms fell well outside the range of control subjects. Thus, BG exhibits pathological false recognition to pseudowords that he has never encountered prior to the experiment.

The same general pattern was observed when non-studied pseudowords were preceded in the list by physically similar pseudowords presented either one time or three times (Fig. 6: "Similar × 1" and "Similar × 3"), except that both BG and control subjects made more false alarms to physically similar words than to dissimilar words. For non-studied items preceded by a single presentation of a similar pseudoword, BG made more false alarms than did controls (0.54 versus 0.39), but one control subject made more false alarms than did BG (the same one who had a high false alarm rate in the no-similarity condition), so the difference did not attain statistical significance. Considering only R false alarms, however, BG still scored well outside the range of control subjects (means: 0.50 versus 0.19, $P < 0.01$). For non-studied items preceded by three presentations of a similar pseudoword, BG made significantly more false alarms overall than did controls (0.75 versus 0.49, $P < 0.05$), although the same control subject made more false alarms than did BG. BG also made significantly more R false alarms than did controls (0.50 versus 0.23, $P < 0.05$).

BG's overall hit rates were roughly comparable to those of controls subjects, with no significant differences. He showed a non-significant trend to make more R responses to studied pseudowords than did controls in the one-presentation condition (0.63 versus 0.49), and made only slightly more R responses to studied pseudowords than did controls in the three-presentation condition (0.79 versus 0.76).

Experiment 5 indicates clearly that BG's false recognitions cannot be attributed to a confusion between experimental and pre-experimental exposures to target items. Relative to control subjects, he consistently made excessive numbers of overall false alarms and R false alarms to pseudowords that he had presumably never encountered prior to the experimental session. In addition, the results fail to support the hypothesis that BG exhibits false recognition only when there is a strong degree of similarity between study and test items. Although BG, like control subjects, made more false alarms when non-studied pseudowords were preceded by physically similar pseudowords than when they were not, he made more false recognitions than did controls in both conditions. Thus, as in experiment 4, BG's pathological false recognition is evident even when there is little or no similarity between individual study and test items.

Experiment 6

The evidence that we have collected thus far still leaves open the question of why BG exhibits pathological false recognition. One possibility emerges from considering the properties of study and test items in experiments 1–5. In each of these experiments, all test items were drawn from the same general class as were the study items. Thus, for instance, in experiments 1, 2 and 4, BG studied familiar words and was tested on familiar words; in experiment 3, BG studied environmental sounds and was tested on environmental sounds; and in experiment 5, BG studied pseudowords and was tested on pseudowords. Even though BG exhibited high false alarm rates when *individual* items were not associatively or physically related to one another, the *sets* of items always had a similar description (e.g., familiar words, nonsense words, etc.). Perhaps BG's high false alarm rate arises because he responds inappropriately to the general or class-level similarity between study and test items. That is, BG may sometimes claim that a test item appeared on the list as long as it resembles the class or category of item that had appeared on the study list, even though the *specific* test item had not been presented. Parkin *et al.* [32] have offered a similar hypothesis concerning false recognition in their patient, JB, and have supported it by showing that JB did not make excessive false alarms to non-studied items that had no class resemblance to previously studied items.

To examine this hypothesis, we exposed BG to a categorized list of words, and then tested him with previously studied items and two types of non-studied items: words drawn from studied categories and words drawn from non-studied categories. False alarms normally increase when lures are taken from studied categories, and this increase depends on the number of studied exemplars [e.g. Refs. 14, 45]. We expected that BG would make more false alarms than control subjects to non-studied words from studied categories. The key question was whether BG would also exhibit a high false alarm rate to words from non-studied categories. According to the hypothesis sketched above, BG should make few if any false alarms to words from non-studied categories because they do not resemble the class of words that he was shown during the study episode.

Method

Materials, design and procedure. The target stimuli consisted of 138 words. Categorized words were taken from six of the Batting and Montague [1] categories: flowers (13), human dwellings (13), weather phenomena (13), colors (19), crimes (19), and food flavorings (19). The remaining 42 words were concrete nouns that were not members of the six critical categories.

Subjects studied a list consisting of 54 category members—six studied exemplars from three of the categories (flowers, dwellings, weather) and 12 studied exemplars from the other three categories (colors, crimes, flavorings). Non-studied items on the recognition test consisted of 42 exemplars from the six studied categories (seven per category) and the 42 non-category words. The procedure was similar to previous experiments: liking ratings (5 sec/word), serial reaction time for two min, and self-paced remember–know recognition test.

Results and discussion

Presenting non-studied words from non-studied categories effectively eliminated false recognition in patient BG: he did not make a single false alarm to a word from a non-studied category. Control subjects, too, almost never made a false alarm to an item from a non-studied category (one subject made one "K" false alarm). However, the significance of this finding is tempered by the fact that BG exhibited little evidence of pathological false recognition to words from studied categories. Figure 7 shows the results hits and false alarms for category words. For non-studied words preceded by six category members, BG's overall false alarm rate at (0.19) was non-significantly higher than the average for control subjects (0.13). For non-studied words preceded by 12 category members, BG's overall false alarm rate (0.20) was virtually identical to the average for control subjects (0.19). BG exhibited non-significant trends for more R false alarms than control subjects in both conditions (0.14 versus 0.7, and 0.15 versus 0.08, respectively). BG's hit rate was generally quite similar to that of control subjects (Fig. 7), except for a non-significant tendency to make more R responses when 12 category exemplars were studied (0.91 versus 0.81).

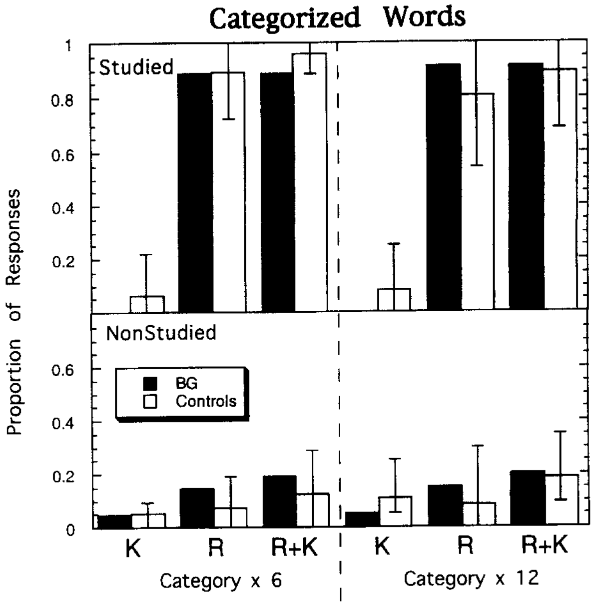

Fig. 7. Mean proportion of "remember" (R), "know" (K), and "old" (R + K) responses in each condition for BG and controls in experiment 6. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject. Studied words were taken from categories with either six or 12 studied category exemplars. Non-studied words were either not members of studied categories (not shown, see text), members of categories with six studied exemplars (category × 6), or members of categories with 12 studied exemplars (category × 12).

The fact that we were able to eliminate BG's false alarms entirely by presenting lure words from non-studied categories supports the idea that his pathological false recognition arises from inappropriate responding to non-studied items that resemble the class or category of studied items. However, this conclusion receives only limited support because, unlike in previous experiments, BG did not exhibit strong evidence of excessive false alarms to words that matched the category of studied items. One possible reason for this finding is related to the fact that BG's lesion is in the right hemisphere. The categorized verbal materials used in experiment 6 may have promoted greater reliance on his intact left hemisphere and, hence, minimized his problems on the recognition test. This consideration raises the possibility that BG might exhibit stronger evidence of false recognition to items from studied categories with non-verbal materials, thereby permitting a stronger test of the hypothesis that pathological false recognition can be eliminated by testing him with novel items from non-studied categories. Experiment 7 examined this possibility.

Experiment 7

In experiment 7, we presented BG with pictures of everyday objects drawn from a variety of categories, and then tested him with previously studied pictures, new pictures from studied categories, and new pictures from non-studied categories.

Method

Materials and design. Target stimuli were 130 pictures from Snodgrass and Vanderwart [50]. Eighty-eight of the pictures represented inanimate objects from six distinct categories: furniture (10), musical instruments (10), tools (14), toys (14), articles of clothing (20), things in a kitchen (20). We also used 21 pictures of animals and 21 pictures of miscellaneous inanimate objects that were not members of the other categories. Subjects studied pictures only from the six categories of inanimate objects, and the number of studied exemplars/category was varied: 3 (furniture, instruments), 7 (tools, toys), and 13 (clothing, kitchen). The test list included these 46 studied pictures, and also included three types of non-studied pictures: seven non-studied exemplars from each of the six studied categories (42), 21 animals, and 21 miscellaneous (non-category) inanimate objects.

Procedure. Each studied picture was presented for 4 sec with a 1-sec interstimulus interval. Subjects were asked to name each picture that appeared. Subjects then performed the previously described serial reaction time task for 2 min. Finally, a self-paced, remember–know recognition test was given for all experimental pictures. The order of study and test lists was randomly determined with the constraint that no more than three consecutive items were from the same condition.

Results

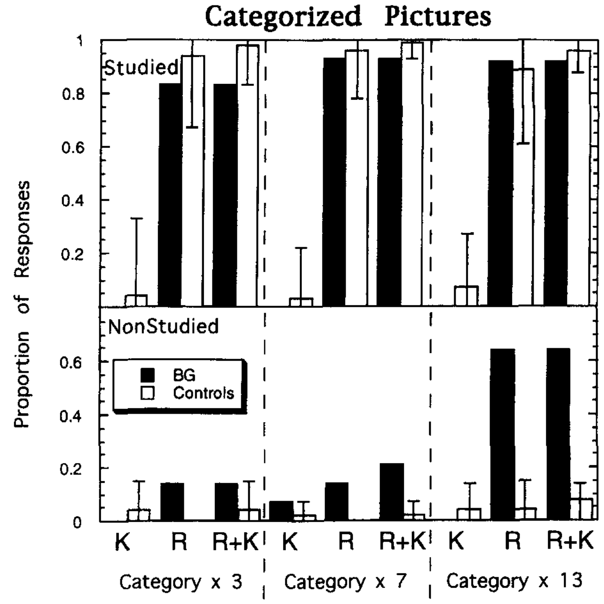

Fig. 8. Mean proportion of "remember" (R), "know" (K), and "old" (R + K) responses in each condition for BG and controls in experiment 7. Error bars denote the maximum and minimum proportion of responses observed in any individual control subject. Studied pictures were taken from categories with either three, seven or 13 studied category exemplars. Non-studied pictures were either not members of studied categories (not shown, see text), members of categories with three studied exemplars (category × 3), members of categories with seven studied exemplars (category × 7) or members of categories with 13 studied exemplars (category × 13).

BG's overall false alarm rate (Fig. 8) to pictures from studied categories (0.33) was significantly ($P < 0.01$) higher than, and outside the range of, control subjects' overall false alarm rate (0.04). Nearly all of BG's false alarms were R responses, whereas control subjects almost never made R false alarm responses. His overall proportion of R false alarms (0.31) was significantly ($P < 0.01$) higher than the average for control subjects (0.01) and fell well outside the control range. As the number of studied category exemplars increased, BG's false alarm rate fell increasingly beyond the range of controls—these differences were significant after studying 13 ($P < 0.01$) and seven ($P < 0.05$) exemplars but not after studying only three exemplars. Indeed, after studying 13 exemplars BG made false alarms (all of them "Rs") to 0.64 of non-studied pictures, which was approximately seven times the control rate of 0.09.

In contrast to his high false alarm rate for pictures from studied categories, BG made only a single false alarm to pictures of animals and also to pictures of inanimate objects from non-studied categories. Thus, the category manipulation had a strong impact on his false recognition responses. Control subjects never made a false alarm to either of these types of items.

BG's hit rates were generally high and quite similar to those of control subjects, although comparisons are hindered by the presence of ceiling effects. He made many more hits than false alarms, indicating a good deal of accurate memory. BG made similar numbers of R responses to studied pictures as did control subjects, although he exhibited a nonsignificant trend for fewer R responses to studied pictures in the three-exemplar condition.

In summary, experiment 7 indicates that with pictorial stimuli, BG exhibits considerable pathological false recognition when tested with novel pictures from studied categories, and almost no pathological false recognition when tested with novel pictures from non-studied categories.

General discussion

In this article we have established the existence of pathological false recognition in patient BG (experiments 1–3), shown that the phenomenon occurs even with novel stimuli and when non-studied items are associatively or physically unrelated to study items (experiments 4 and 5), and demonstrated that BG's false recognition can be virtually eliminated by presenting him with a categorized list and testing him with new words or pictures from non-studied categories (experiments 6 and 7).

We noted in the discussion of experiment 1 that BG's tendency to make "remember" false alarms—that is, to claim that he recollects specific details of the prior presentation of non-studied items—might be the result of a general bias to use the R response for both studied and non-studied items. In experiments 2, 3, 5 and 6 BG made more R responses to studied items than did control subjects, suggesting the existence of such a bias. But in experiments 1, 4 and 7 he made equal or fewer numbers of R responses to studied items than did controls. Thus, a general bias to use the R response may have played some role in BG's "remember" false alarms, but it probably cannot explain why above-normal R responses were more prevalent for non-studied than studied items.

Parkin *et al.*'s [32] patient, JB, whose pattern of false recognitions is roughly similar to BG's, suffered a ruptured anterior communicating artery aneurysm that resulted in an enlarged left frontal horn together with an adjacent area of low density. One clear difference between JB and BG is that when the remember/know procedure was used, JB made exclusively "know" false recognitions [32]—that is, he said that non-studied items seemed familiar, but did not claim to recollect specific details about these items. BG, by contrast, made predominantly "remember" false recognitions. It is conceivable that these differences are related in some way to the left-versus-right-sided damage that characterizes the two patients. Indeed, Blaxton and Theodore [3] have reported that patients with right-temporal lobe epilepsy make more "remember" than "know" responses, whereas normal controls and patients with left-temporal lobe epilepsy make more "know" than "remember" responses. Nonetheless, we must be cautious in drawing direct comparisons between JB and BG, because the use of R versus K responses may simply reflect different criteria for R and K responses used by the two patients, or differences in the way that they understood the task instructions.

The right lateralization of BG's lesion may also explain the different results with categorized words (experiments 6) versus categorized pictures (experiment 7). BG's within-category false alarms were similar to control subjects for words but extremely higher than controls for pictures. Previous research has shown that memory deficits arising from right frontal lesions are more severe for pictorial than verbal materials and vice versa for left frontal lesions [27]. Further research will be necessary to determine whether patterns of false recognition in left- and right-frontal damaged patients are attributable to disruption in the same processes or in different processes.

To understand false recognition in patient BG, it is necessary to account for two characteristics of his performance, which at first may appear paradoxical—i.e., BG exhibits false recognition even for non-studied words and pseudowords that have no apparent relationship to studied items, yet he does not exhibit false recognition for new words or pictures from non-studied categories. To resolve this puzzle, we suggest that BG sometimes commits false alarms when a test item is generally consistent with the class, category, or characteristics of the study list, and he fails to recollect the identity of particular items. Other research has suggested a distinction between memories that are based on general similarities between studied items and test items versus other sorts of memory that allow the retrieval of more specific information.

For example, within the domain of autobiographical memory, Conway and Rubin [6] referred to different forms of memory based on *general event knowledge* and *event-specific knowledge*. General events refer to high-level episodes, such as *going to the movies*, whereas event-specific knowledge refers to particular episodes that are nested within the general event, such as *spilling my popcorn* or *being surprised by the end of the film*. With respect to a memory experiment, the general event might be represented as *seeing a list of words* or *seeing some pictures of clothing and furnitures*, whereas event-specific knowledge would refer to memory for the specific items that were presented during the experiment. When a person is trying to remember whether a specific item was shown earlier in an experiment (event-specific knowledge), test items can activate a general event description. For normal subjects, such activation does not usually constitute sufficient evidence for claiming that an item appeared previously on the study list. Within this framework BG, by contrast, may be more likely to accept activation of a general event description as evidence for event-specific memory.

Within the domain of recognition memory, Hintzman and Curran [15–17] obtained evidence supporting a distinction between separate memory processes of familiarity and recall or recollection. Familiarity is a unidimensional measure of the overall similarity between a test item and the study list [e.g., Ref. 14] whereas recall allows retrieving the content of a specific experience (see Brainerd *et al.* [4], for a similar distinction between gist and verbatim memory). Using the response–signal method to examine the time course of retrieval, Hintzman and Curran [15] found that subjects initially base recognition on similarity between study and test items, so false alarm rates are especially high to similar lures. About 90 msec later, subjects are able to recall specific information that allows them to reject similar lures. This relatively slow access to item-specific information suggests that recollection may depend on effortful, intentional retrieval processes (also see Jacoby [18]).

Previous research on frontal lobe contributions to memory has suggested functions that are consistent with such an effortful recollection process. Milner *et al.* [49], among others [29, 43, 52], have argued that a variety of memory deficits in frontal patients may be attributable to deficient search and retrieval processes. Recent positron emission tomography studies have revealed that right frontal-lobe regions are consistently activated during episodic retrieval of recently studied items [55], and also suggest that such activations may reflect intentional or effortful retrieval processes [21, 40]. It is conceivable that the activation of right frontal regions during episodic retrieval in normal subjects reflects the intentional retrieval processes that we hypothesize are defective in patient BG.

The idea that BG's high false alarm rate is attributable to over-reliance on familiarity [14–17] or gist [14] seems consistent with his pattern of "yes" versus "no" responses. However, insofar as R responses indicate the retrieval of specific content from memory, BG's high rate of R false alarms cannot be attributable to a unidimensional signal such as familiarity.

Norman and Schacter [31] have argued that BG's false alarms are the product of retrieving actual content from memory. Norman and Schacter [31] suggest that the effortful retrieval processes supported by the frontal lobes play a role in generating focused descriptions of the study context that guide the search for specific items. In the Conway and Rubin's [6] terms, BG's recognition decisions may be based on a fuzzy general event description (e.g., "saw a bunch of words in the list") that is not focused enough to exclude many of the lures on a recognition test. Thus, BG's false alarms would not be entirely based on a familiarity signal that is devoid of content, but rather on retrieved content with impoverished details. BG is unable to set an appropriate criterion for deciding what was on the list because he attempts to match test items to a vague, unfocused description of what transpired at study. Normal subjects, by contrast, may be able to generate more focused descriptions of what went on at study that help them to reject new items. These ideas are consistent with the suggestions of Shimamura [46] that frontal regions play a role in filtering irrelevant information. A poorly focused retrieval description may serve as an ineffective filter, with the result that patients make their recognition decisions based on inappropriate or extraneous information [see Refs. 25, 35]. However, not all patients with frontal lobe lesions exhibit recognition memory deficits and elevated false alarm rates, as BG does, so we must be cautious about extrapolating from his case to others (for more detailed discussion, see Norman and Schacter [31]). Further investigation of intentional retrieval processes, focused descriptions, and the role played by frontal regions in supporting them, constitutes an important next step in developing a cognitive neuroscience of memory distortion.

Acknowledgements—This research was supported by National Institute of Neurological Disorders and Stroke Grants PO1 NS27950 and NS26895, National Institute on Aging Grant RO1 AG08441, and Department of Veterans Affairs Merit Review 097-44-3765-001. We thank Kimberly Nelson for help with preparation of the manuscript. Tim Curran is now at the Department of Psychology, Case Western Reserve University.

References

1. Battig, W. F. and Montague, W. E. Category norms for verbal items in 56 categories: a replication and extension of the Connecticut category norms. *Journal of Experimental Psychology* **80,** 1–46, 1969.
2. Bennett, C. A. and Franklin, N. L. *Statistical Analysis in Chemistry and the Chemical Industry.* John Wiley, New York, 1954.
3. Blaxton, T. A. and Theodore, W. H. The role of the temporal lobes in recognizing visuospatial materials: remembering versus knowing. *Brain Cognition,* in press.
4. Brainerd, C. J., Reyna, V. F. and Kneer, R. False-recognition reversal: when similarity is distinctive. *Journal of Memory and Language* **34,** 157–185.
5. Butters, M. A., Kasniak, A. W., Glisky, E. L., Eslinger, P. J. and Schacter, D. L. Recency discrimination deficits in frontal lobe patients. *Neuropsychology* **8,** 343–353, 1994.
6. Conway, M. A. and Rubin, D. C. The structure of autobiographical memory. In *Theories of Memory,* A. F. Collins, S. E. Gathercole, M. A. Conway and P. E. Morris (Editors). Lawrence Erlbaum Associates, Hillsdale, NJ, 1993.
7. Craik, F. I. M., Morris, L. W., Morris, R. G. and Loewen, E. R. relations between source amnesia and frontal lobe functioning in older adults. *Psychology and Aging* **5,** 148–151, 1990.
8. Dalla Barba, G. Confabulation: knowledge and recollective experience. *Cognitive Neuropsychology* **10,** 1–20, 1993.
9. Delbecq-Derouesné, J., Beauvois, M. F. and Shallice, T. Preserved recall versus impaired recognition. *Brain* **113,** 1045–1074, 1990.
10. della Rocchetta, A. I. Classification and recall of pictures after unilateral frontal or temporal lobectomy. *Cortex* **22,** 189–211, 1986.
11. DeLuca, J. Predicting neurobehavioral patterns following anterior communicating artery aneurysm. *Cortex* **29,** 639–647, 1993.
12. Gardiner, J. M. Functional aspects of recollective experience. *Memory and Cognition* **16,** 309–313, 1988.
13. Gardiner, J. M. and Java, R. I., Recognizing and remembering. In *Theories of Memory,* A. F. Collins, S. E. Gathercole, M. A. Conway and P. E. Morris (Editors). Lawrence Erlbaum Associates, Hove, UK, 1993.
14. Hintzman, D. L. Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review* **95,** 528–551, 1988.
15. Hintzman, D. L. and Curran, T. Retrieval dynamics of recognition and frequency judgments: evidence for separate processes of familiarity and recall. *Journal of Memory and Language* **33,** 1–18, 1994.
16. Hintzman, D. L. and Curran, T. When encoding fails: instructions, feedback, and registration without learning. *Memory and Cognition* **23,** 213–226, 1995.
17. Hintzman, D. L., Curran, T. and Oppy, B. Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning Memory Cognition* **18,** 667–680, 1992.
18. Jacoby, L. L. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory Language* **30,** 513–541, 1991.
19. Janowsky, J. S., Shimamura, A. P. and Squire, L. R. Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia* **27,** 1043–1056, 1989.

20. Kapur, N. and Coughlan, A. K. Confabulation and frontal lobe dysfunction. *Journal of Neurology, Neurosurgery and Psychiatry* **43,** 461–463, 1980.
21. Kapur, S., Craik, F. I. M., Jones, C., Brown, G. M., Houle, S. and Tulving, E. Functional role of the prefrontal cortex in retrieval of memories: a PET study. *NeuroReport* **6,** 1880–1884, 1995.
22. Kucera, H. and Francis, W. N. *Computational Analysis of Present-day American English.* Brown University Press, Providence, RI, 1967.
23. Leonard, G. and Milner, B. Contribution of the right frontal lobe to the encoding and recall of kinesthetic distance information. *Neuropsychologia* **29,** 47–58, 1991.
24. Loftus, E. F. The reality of repressed memories. *American Psychologist* **48,** 518–537, 1993.
25. Luria, A. R. *The Neuropsychology of Memory.* Winston, Washington, DC, 1976.
26. McAndrews, M. P. and Milner, B. The frontal cortex and memory for temporal order. *Neuropsychologia* **29,** 849–859, 1991.
27. Milner, B., Corsi, P. and Leonard, G. Frontal-lobe contribution to recency judgments. *Neuropsychologia* **29,** 601–618, 1991.
28. Milner, B., Petrides, M. and Smith, M. L. Frontal lobes and the temporal organization of memory. *Human Neurobiology* **4,** 137–142, 1985.
29. Moscovitch, M. Memory and working-with-memory: Evaluation of a component process model and comparisons with other models. In *Memory Systems* 1994, D. L. Schacter and E. Tulving (Editors). MIT Press, Cambridge, MA, 1994.
30. Moscovitch, M. Confabulation. In *Memory Distortion*, D. L. Schacter, J. T. Coyle, G. D. Fischbach, M. M. Mesulam and L. E. Sullivan (Editors). Harvard University Press, Cambridge MA, 1995.
31. Norman, K. A. and Schacter, D. L. Implicit memory, explicit memory, and false recollection: a cognitive neuroscience perspective. In *Implicit Memory and Metacognition*, L. M. Reder (Editor). Lawrence Erlbaum Associates, Hillsdale, NJ, in press, 1996.
32. Parkin, A. J., Bindschaedler, C., Harsent, L. and Metzler, C. Pathological false alarm rates following damage of the frontal lobes. Submitted for publication, 1996.
33. Petrides, M., Alivisatos, B., Evans, A. C. and Meyer, E. Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proceedings of the National Academy of Sciences* **90,** 873–877, 1993.
34. Postman, L. The California norms: association as a function of word frequency. In *Norms of Word Association*, L. Postman and G. Keppel (Editors). Academic Press, San Diego, 1970.
35. Pribham, K. H., Ahumada, A., Hartog, J. and Roos, I. A progress report on the neurological processes disturbed by frontal lesions in primates. In *Frontal Granular Cortex and Behavior*, J. M. Warren and K. Akert (Editors). McGraw-Hill, New York, 1964.
36. Rajaram, S. Remembering and knowing: two means of access to the personal past. *Memory and Cognition* **21,** 89–102, 1993.
37. Roediger, H. L. III and McDermott, K. B. Creating false memories: remembering words not presented in lists. *Journal of experimental Psychology: Learning, Memory and Cognition* **21,** 803–814, 1995.
38. Schacter, D. L. Memory, amnesia, and frontal lobe dysfunction. *Psychobiology* **15,** 21–36, 1987.
39. Schacter, D. L. Memory distortion: history and current status. In *Memory Distortion*, D. L. Schacter, J. T. Coyle, G. D. Fischbach, M. M. Mesulam and L. E. Sullivan (Editors). Harvard University Press, Cambridge, MA, 1995.
40. Schacter, D. L., Alpert, N. M., Savage, C. R., Rauch, S. L. and Albert, M. L. Conscious recollection and the human hippocampal formation: Evidence from positron emission tomography. *Proceedings of the National Academy of Sciences* **93,** 321–325, 1996.
41. Schacter, D. L. and Curran, T. The cognitive neuroscience of false memories. *Psychiatric Annals* **25,** 726–730, 1995.
42. Schacter, D. L., Harbluk, J. L. and McLachlan, D. R. Retrieval without recollection: an experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior* **23,** 593–611, 1984.
43. Shallice, T. *From Neuropsychology to Mental Structure.* Cambridge University Press, Cambridge, 1988.
44. Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J. and Dolan, R. J. Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* **368,** 633–635, 1994.
45. Shiffrin, R. M., Huber, D. and Marinelli, K. Effects of category length and strength on familiarity in recognition. *Journal of experimental Psychology: Learning, Memory and Cognition* **21,** 267–287, 1995.
46. Shimamura, A. P. Memory and frontal lobe function. In *The Cognitive Neurosciences*, M. Gazzaniga (Editor). MIT Press, Cambridge MA, 1995.
47. Shimamura, A. P., Janowsky, J. S. and Squire, L. R. Memory for the temporal order or events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia* **28,** 803–813, 1990.
48. Shimamura, A. P. and Squire, L. R. A neuropsychological study of fact memory and source amnesia. *Journal of experimental Psychology: Learning, Memory and Cognition* **13,** 464–473, 1987.
49. Smith, M. L. and Milner, B. Estimation of frequency of occurrence of abstract designs after frontal or temporal lobectomy. *Neuropsychologia* **26,** 297–306, 1988.
50. Snodgrass, J. G. and Vanderwart, M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental Psychology: Human Learning and Memory* **6,** 174–215.
51. Stuss, D. T. Disturbance of self-awareness after frontal system damage. In *Awareness of Deficit After Brain Injury: Clinical and Theoretical Issues*, G. P. Prigatano and D. L. Schacter (Editors). Oxford University Press, New York, 1991.
52. Stuss, D. T., Eskes, G. A. and Foster, J. K. Experimental neuropsychological studies of frontal lobe functions. In *Handbook of Neuropsychology*, F. Boller and J. Grafman (Editors). Elsevier, Amsterdam, 1994.
53. Stuss, D. T., Pogue, J., Buckle, L. and Bondar, J.

Characterization of stability of performance in patients with traumatic brain injury: variability and consistency on reaction time tests. *Neuropsychology* **8,** 316–324, 1994.

54. Tulving, E. Memory and consciousness. *Canadian Psychologist* **26,** 1–12, 1985.
55. Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M. and Houle, S. Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences* **9,** 2016–2020, 1994.
56. Underwood, B. J. False recognition produced by implicit verbal responses. *Journal of experimental Psychology* **70,** 122–129, 1965.