Failure to Respond Autonomically to Anticipated Future Outcomes Following Damage to Prefrontal Cortex

Antoine Bechara, Daniel Tranel, Hanna Damasio, and Antonio R. Damasio

Department of Neurology, Division of Behavioral Neurology and Cognitive Neuroscience, University of Iowa College of Medicine, Iowa City, Iowa 52242

Following damage to specific sectors of the prefrontal cortex, humans develop a defect in real-life decision making, in spite of otherwise normal intellectual performance. The patients so affected may even realize the consequences of their actions but fail to act accordingly, thus appearing oblivious to the future. The neural basis of this defect has resisted explanation. Here we identify a physiological correlate for the defect and discuss its possible significance. We measured the skin conductance responses (SCRs) of 7 patients with prefrontal damage, and 12 normal controls, during the performance of a novel task, a card game that simulates real-life decision making in the way it factors uncertainty, rewards, and penalties. Both patients and controls generated SCRs after selecting cards that were followed by penalties or by reward. However, after a number of trials, controls also began to generate SCRs prior to their selection of a card, while they pondered from which deck to choose, but no patients showed such anticipatory SCRs. The absence of anticipatory SCRs in patients with prefrontal damage is a correlate of their insensitivity to future outcomes. It is compatible with the idea that these patients fail to activate biasing signals that would serve as value markers in the distinction between choices with good or bad future outcomes; that these signals also participate in the enhancement of attention and working memory relative to representations pertinent to the decision process; and that the signals hail from the bioregulatory machinery that sustains somatic homeostasis and can be expressed in emotion and feeling.

Patients with damage to the prefrontal region, especially when the damage is centered in the ventral and medial aspects of this region, develop a severe impairment in personal and social decision making, in spite of otherwise largely preserved intellectual abilities (Damasio et al., 1991). Their condition has posed a double challenge: although the impairment is obvious in its ecological niche, there has been neither a laboratory probe to detect and measure it nor a satisfactory account of the neural and cognitive mechanisms underlying it. We have recently addressed the first challenge by showing that the impairment can be detected and measured using a laboratory task designed to simulate real-life decisions in the way it factors uncertainty, rewards, and penalties (Bechara et al., 1994). In the study presented below we seek to address the second challenge: the neural and cognitive basis of the condition.

Evidence from previous studies suggests that these patients possess and can access the knowledge necessary to conjure up options of actions and scenarios of future outcomes just as normal controls do (Saver and Damasio, 1991), and that their defect seems to be at the stage in the reasoning process at which the individual must act on such knowledge, for instance, the stage of reasoning at which a decision-making strategy must yield a choice. There are several possibilities to explain why they do not make a choice (or make a disadvantageous one). Here, we mention the most obvious ones. For instance, defects in basic processes such as attention (Posner, 1986) or temporal integration (Fuster, 1989) might preclude the effective display of knowledge representations required for the operations of reasoning. Another possibility is that knowledge representations, especially those of future outcomes, can be evoked by these patients but are unstable, for instance, representations would not be held in working memory long enough for reasoning strategies to be applied to them. This possibility invokes a defect in the process of working memory, a process whose existence Goldman-Rakic has demonstrated for the spatial domain relative to dorsolateral prefrontal cortex (Goldman-Rakic, 1987, 1992), and which she suggests also applies to other domains and to the prefrontal cortices in general. Yet another distinct possibility is that the representations of future outcomes would be properly attended, temporally integrated, and held in working memory, but they would not be marked with a negative or positive value. This possibility invokes the somatic marker hypothesis, which posits that overt or covert somatic states juxtapose a value mark in the processing of a cognitive scenario, and that the qualifying value mark helps endorse or reject an option for action. Because those biasing signals hail from the bioregulatory machine that sustains somatic homeostasis, we have called these signals somatic markers (Damasio et al., 1991). The term "somatic" is used here in its broadest sense to include both musculoskeletal and visceral structures of the soma, as well as their neural representations in the central nervous system. The term "somatic state" refers to changes in the soma, or changes in its neural representations, or both. The final possibility we would like to mention combines some features of the previous ones: somatic markers would help drive and maintain attention and working memory, and also mark and bias the representation of certain outcomes. We must make clear, at the outset, that given the current evidence, the last possibility provides the best fit to the condition of our patients.

The aim of the present study is the gathering of evidence with which one may eventually distinguish among these varied accounts, using the novel task described in a previous study (Bechara et al., 1994). In the task, the subject sits in front of four decks of cards, is given a loan of play money, and is engaged in a game in which the goal is to win as much money as possible (or lose as little). The game requires a series of card selections, one card at a time, from any of the four decks, until the subject is told to stop. After turning each card, the subject always receives some money, but the amount is consistently higher in two of the decks. After turning some cards, however, the subject is both given money and asked to pay a penalty. The penalty amounts are higher in the two high paying decks than in the low paying ones. The subject has no way of knowing when a penalty will arise, and no way of calculating precisely the net gain or loss for each deck. The only successful strategy is to develop a "hunch," for instance, to predict that in the long run the high paying decks are "bad," and the low paying are "good." Controls gradually formulate this prediction and select more from the advantageous decks (those that have low immediate gain, but larger future yield) and less from the disadvantageous decks (those that have

Cerebral Cortex Mar/Apr 1996;6:215-225; 1047-3211/96/$4.00

Table 1
Demographic information

	Age	Gender	Education	Handedness[a]	Chronicity[b]
Patients					
EVR-318	52	M	14	+100	17 years
VY-500	64	F	12	+100	10 years
PK-770	53	F	16	+100	9 years
DM-1336	85	M	9	+100	14 years
RS-1479	63	F	8	+100	4 years
JR-1584	51	M	8	+100	2 years
DV-1589	45	M	18	+100	12 years
Mean ± SD	59 ± 13*		12 ± 4†	+100	9.7 ± 5.3
Controls					
TC	30	F	19	+100	NA
KP	23	F	15	+100	NA
LM	30	F	16	+100	NA
PC	44	F	12	+100	NA
FW	50	M	16	+100	NA
TS	49	M	8	+100	NA
DA	66	M	16	+100	NA
JP	66	M	7	+100	NA
MD	69	F	12	+100	NA
DB	55	M	16	+100	NA
MB	52	F	16	+100	NA
ME	50	F	12	+100	NA
Mean ± SD	49 ± 15*		14 ± 4†	+100	NA

[a]Assessed with Geschwind-Oldfield Questionnaire, which ranges from +100 (full right-handedness) to −100 (full left-handedness).

[b]Time elapsed between onset of lesion and collection of data for the present study.

*No significant difference between the ages of patients and controls [$t(17) = 1.5$, $p > 0.1$].

†No significant difference in the education of patients and controls [$t(17) = 0.9$, $p > 0.1$]).

high immediate gain, but larger future loss). Ventromedial prefrontal patients do the exact opposite and select more from the disadvantageous decks than from the advantageous ones (Bechara et al., 1994). In this task, as in real life, the patients appear insensitive to the future consequences of their actions, and are thus guided by immediate prospects.

We hypothesized that if somatic state activation is necessary for the distinction between good and bad choices, then subjects performing the gambling task should show evidence of such an activation when they attempt to choose between the good and bad decks. Using the electrodermal skin conductance response (SCR) as a dependent measure of somatic state activation, we predicted that normal controls performing the gambling task would generate SCRs prior to their selection of a card from a given deck, that is, during the time they ponder from which deck to choose. By contrast, patients with prefrontal lesions would fail to generate anticipatory SCRs.

The use of skin conductance responses in these experiments was considered in the following historical perspective. Earlier studies had suggested that the frontal lobes might have a role in the central control of orienting reflexes, including skin conductance responses. For example, patients with large frontal lobe tumors showed defects in the vegetative components of the orienting reflex, including pathologically rapid habituation (Luria et al., 1964; Luria and Homskaya, 1970; Luria, 1973). Similar findings were obtained from studies of nonhuman primates (Grueninger et al., 1965; Kimble et al., 1965). Recent studies in humans, however, have revealed that damage centered in the ventromedial region of the prefrontal cortex does not cause a generalized SCR impairment (Damasio et al., 1990; Tranel and Damasio, 1994). Specifically, damage to the ventromedial frontal region alone is not sufficient to produce an SCR impairment to orienting stimuli (Damasio et al., 1990; Tranel and Damasio, 1994).

Table 2
Neuropsychological data

	EVR-318	VY-500	PK-770	DM-1336	RS-1479	JR-1584	DV-1589
VIQ	131	113	121	96	87	90	126
PIQ	135	88	89	85	88	111	98
WMS-MQ	143	90r	143	120	95r	67r	100r
BVRT							
C	9	6	9	4	5	7	7
E	1	6	1	14	7	6	4
Attn/Conc	138	109	135	110	75	108	112
Speech	Normal	Normal	Normal	Normal	Normal	Normal	Normal
COWA	49	31	65	21	51	25	41
WCST							
Cat	6	4	6	2	0	6	6
PE	4	36	5	36	84	10	5
FRT	43	45	34	41	44	39	43
JLO	30	29	21	25	22	30	23

VIQ = Verbal IQ; PIQ = performance IQ (from the WAIS-R). The WMS = MQ is a memory quotient prorated from the Wechsler Memory Scale, administered without the Visual Reproduction subtest; for instance, the score is derived from verbally based subtests only. MQs accompanied by an r are the Verbal Index from the Wechsler Memory Scale—Revised. BVRT = Benton Visual Retention Test: C = number correct (#/10); E = number errors. Attn/Conc = Attention/Concentration Index form the Wechsler Memory Scale—Revised. COWA = Controlled Oral Word Association test from the Multilingual Aphasia Exam. WCST = Wisconsin Card Sorting Test: Cat = categories completed; PE = perseverative errors. FRT = Facial Recognition Test. JLO = Judgment of Line Orientation test.

Materials and Methods

Characteristics of the Control and Target Subject Groups

The control group included six women and six men ($n = 12$) with an age range from 23 to 69 years, and 7 to 19 years of education (Table 1). Three of the control subjects were employees of The University of Iowa Hospitals & Clinics, and the rest were relatives of patients who visited the hospital from various communities in the midwest. Controls were screened with clinical interviews to determine that they were free from neurological and psychiatric disorders, and none had a history of learning disability or mental retardation. The patient (target) group included three women and four men ($n = 7$) with an age range from 45 to 85 years, and 8 to 18 years of education. The patients were selected from the Patient Registry of the University of Iowa's Division of Behavioral Neurology & Cognitive Neuroscience. The Registry contains an extensive database of neuroanatomical analyses based on neuroimaging data, as well as neuropsychological measurements, behavioral observations by clinicians, and factual data concerning patients' behavior provided by collateral sources. All subjects have been carefully screened by the investigators over a period of many years for signs of psychiatric disorder, drug or alcohol abuse, or current use of medications that may affect performance. At the time of the experiments reported here, none of these factors were present in any of the subjects. In addition, none of the subjects had a history of learning disability, psychiatric disorder, substance abuse, systemic disease that may affect the central nervous system, or mental retardation. All had at least an eighth grade education and were native English speakers. Screening for these factors was done according to standardized instruments, including the Wechsler Adult Intelligence Scale—Revised, Wide Range Achievement Test-3, and Minnesota Multiphasic Personality Inventory—2.

The criteria for patient selection for this study were (1) the presence of abnormal decision making, as evidenced by a consistent record of manifestly aberrant decisions in day-to-day life (see neuropsychological status of target subjects below) and (2) the presence of bilateral lesions involving the ventral and medial prefrontal region. Four of the target subjects in this study served as subjects in a previous study (Bechara et al., 1994). However, the SCR data presented here are obtained during a separate experiment, conducted at least 6 months after the behavioral experiment.

Neuropsychological Status

Table 1 provides demographic information on the seven patients. Table 2 summarizes their neuropsychological profiles (for references

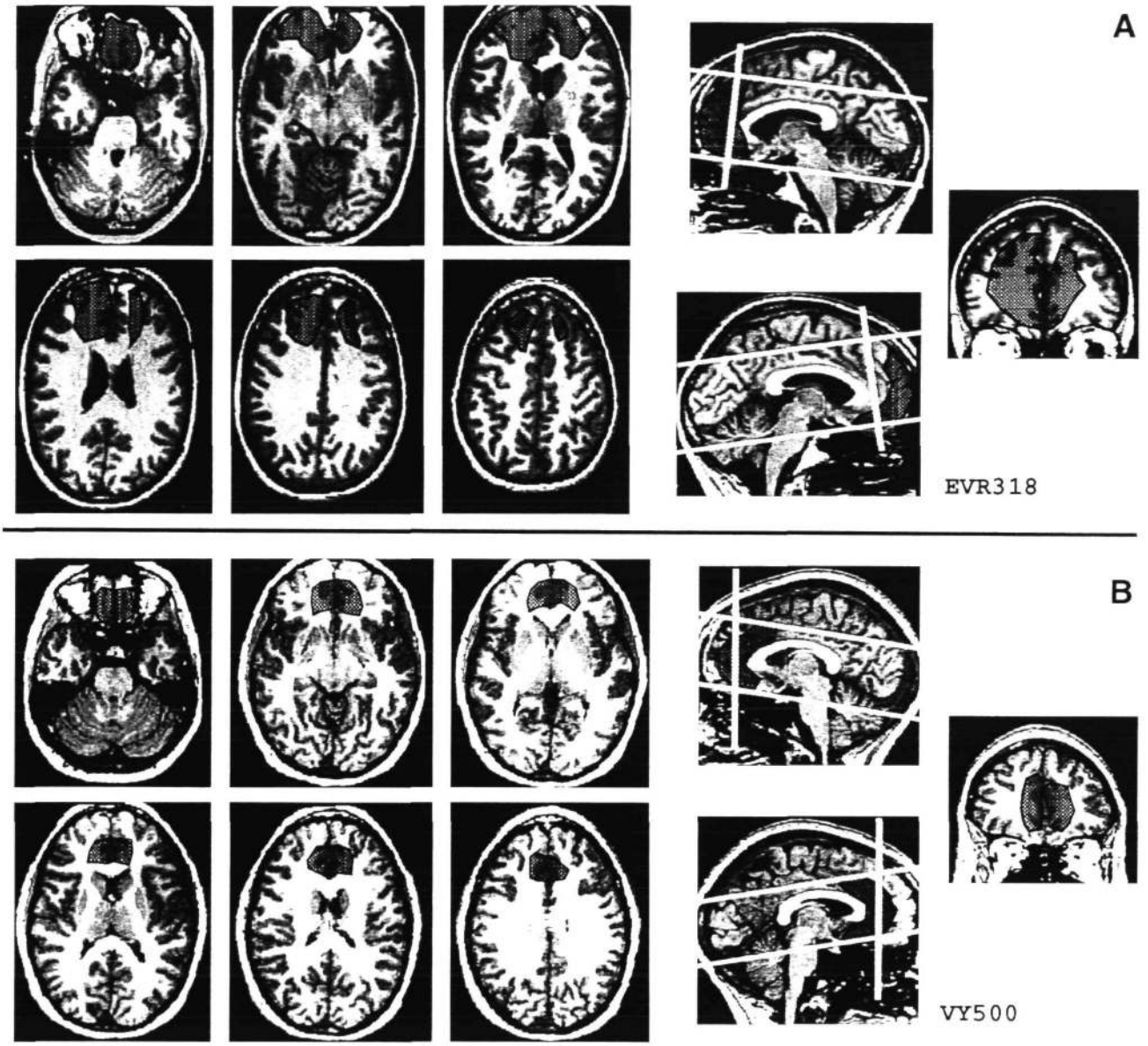

Figure 1. *A–G,* Transverse, coronal, and sagittal views of brain sections depicting the extent of the prefrontal lesions in the seven patients who participated in this study. Each panel depicts the lesion from one patient. The identification number of the patient is shown in the *lower right corner* of each panel.

to neuropsychological tests, see Benton et al., 1983; Lezak, 1995; Tranel, 1995). As Table 2 indicates, none of the target subjects suffered from a pervasive defect of intellect or memory, although some did have isolated impairments in some subtests. Working memory, indexed by the Mental Control subtest of the Wechsler Memory Scale and by the Benton Visual Retention Test, was normal or near normal in all seven subjects. All had normal speech and linguistic functioning. All performed normally on at least one of the visuoperceptual tests (Facial Recognition Test, Judgment of Line Orientation). So-called executive functions, as probed by the Wisconsin Card Sorting Test and the Controlled Oral Word Association test, were normal in four subjects, and mildly to moderately impaired in the other three.

The seven patients have been studied extensively with other clinical and experimental procedures aimed at measuring personality, social awareness, self-monitoring, and insight (Anderson et al., 1992; Saver and Damasio, 1991). Most of them manifest at least some degree of impairment in these domains, as would be expected on the basis of their bilateral ventromedial frontal lesions accompanied by the development of "acquired sociopathy" (Damasio et al., 1991; Tranel, 1994).

Anatomical Findings

The anatomical analyses were carried out on raw data from high-resolution magnetic resonance scans and x-ray computerized tomograms, using the standard procedures of the Division's Laboratory of Neuroimaging and Human Neuroanatomy. These include both template plotting (Damasio and Damasio, 1989), and three-dimensional volume reconstructions based on BRAINVOX (Damasio and Frank, 1992). All lesions were chronic, stable, and confined to the frontal region. The visual summary of the analyses of individual lesions is contained in Figure 1 (*A–G*). The description of each patient's lesion is as follows. EVR-318: bilateral damage in ventromedial sector. The cortex is severely damaged in the right hemisphere but less so in the left. The underlying white matter is equally damaged in both hemispheres. Bilateral damage is present in the frontopolar sector, more extensive on the right. Damage in the most anterior dorsolateral aspect of the sector in the right hemisphere. The left hemisphere is intact. VY-500: bilateral damage in ventromedial sector, largely confined to its posterior region. The frontopolar and dorsolateral sectors are intact. PK-770: bilateral damage in the posterior and inferior aspects of the ventromedial sector; bilateral damage in frontopolar sec-

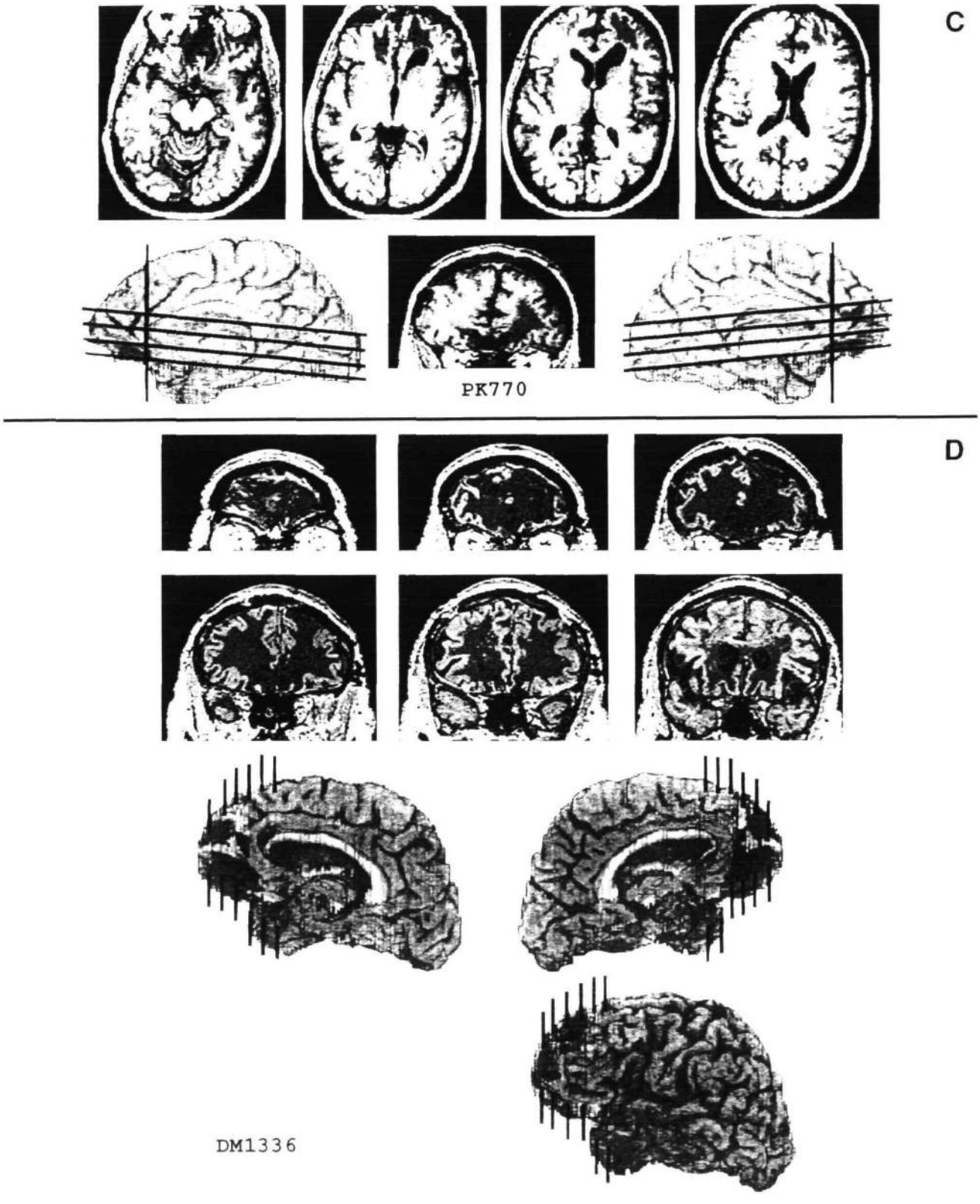

Figure 1. Continued.

tor, mostly involving the white matter and more extensive in the left hemisphere. The dorsolateral sector is intact. DM-1336: bilateral damage in ventromedial sector; bilateral damage in frontopolar sector confined to the white matter. The dorsolateral sector is intact. RS-1479: bilateral damage in the ventromedial sector, largely in white matter, sparing most of the cortex. The frontopolar sector is intact. Damage to the right anterior aspect of the dorsolateral sector is present. JR-1584: bilateral damage to the ventromedial sector, more extensive in the left hemisphere; minimal bilateral damage to the inferior aspect of the frontopolar sector. The dorsolateral sector is intact. Un-

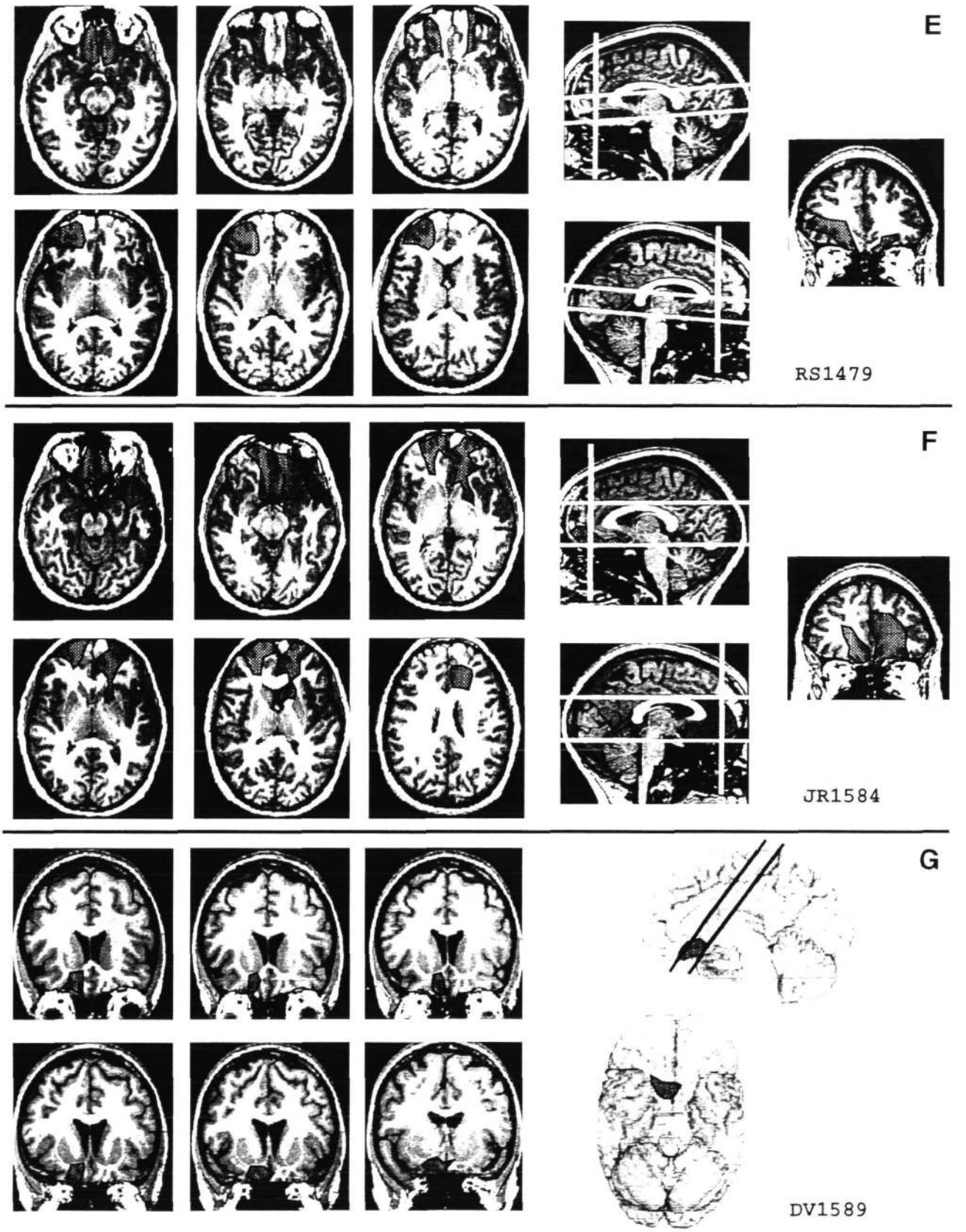

Figure 1. Continued.

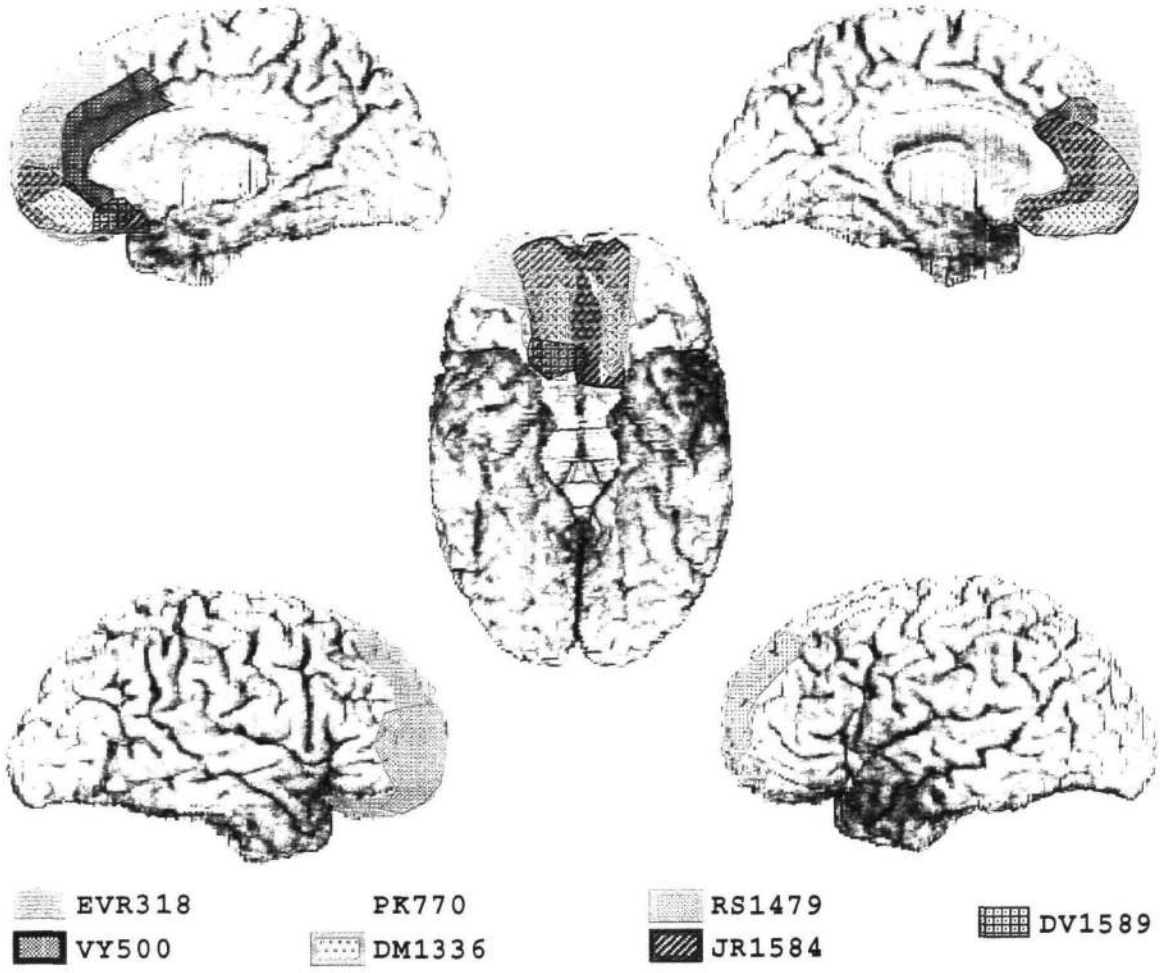

Figure 2. Right medial (*upper left*), right lateral (*lower left*), left medial (*upper right*), left lateral (*lower right*), and inferior (*center*) views of the brain, depicting a composite of the prefrontal lesions in the seven patients who participated in this study. The figure reveals by superposition the concentration of damage in medial and ventral aspects of frontal lobe. Each of the colors in the composite corresponds to one patient.

like other subjects, this patient also had damage in the head of the left caudate nucleus, and the left temporal lobe. DV-1589: damage to the posterior aspect of the ventromedial sector in the right hemisphere; minimal damage of the same territory in the left hemisphere. The frontopolar and dorsolateral sectors are intact.

A composite of the surface aspect of all lesions is shown in Figure 2, revealing the concentration of damage in the medial and ventral aspects of frontal lobe.

Characteristics of the Experimental Task

In the card game, subjects sit in front of four decks of cards equal in appearance and size, and are given a $2000 loan of play money (a set of facsimile U.S. bills), and the goal is to win as much money as possible. The subjects are told that the game requires a long series of card selections, one card at a time, from any of the four decks, until they are told to stop. After turning each card, the subjects receive some money (the amount is only announced after the turning, and varies with the deck). After turning some cards, the subjects are both given money and asked to pay a penalty (again, the amount is only announced after the card is turned and varies with the deck and the position in the deck according to a schedule unknown to the subjects). Turning any card from deck A or deck B yields $100; turning any card from deck C or deck D yields $50. However, the ultimate future yield of each deck varies because the penalty amounts are higher in the high paying decks (A and B), and lower in the low paying decks (C and D). For example, after turning 10 cards from deck A, the subjects have earned $1000, but they have also encountered five unpredicted punishments bringing their total cost to $1250 , thus incurring a net loss of $250. The same happens on deck B (except that instead of encountering five punishments, there is one large unpredicted punishment in the amount of $1250). On the other hand, after turning 10 cards from decks C or D, the subjects earn $500, but the total of their unpredicted punishments is only $250 (i.e., subject nets $250). In summary, decks A and B are equivalent in terms of overall net loss over the trials. The difference is that in deck A, the punishment is more frequent, but of smaller magnitude, whereas in deck B, the punishment is less frequent, but of higher magnitude. Decks C and D are also equivalent in terms of overall net loss. In deck C, the punishment is more frequent and of smaller magnitude, while in deck D the punishment is less frequent but of higher magnitude. Decks A and B are "disadvantageous" because they cost the most in the long run, while decks C and D are "advantageous" because they result in an overall gain in the long run. The preprogrammed schedules of reward and punishment are published elsewhere (Bechara et al., 1994).

Procedures and Measurement of the Psychophysiological Variable

We used the electrodermal skin conductance response (SCR) as a dependent measure of somatic state activation. The equipment, techniques, recording, and scoring have been described elsewhere in detail (Tranel and Damasio, 1988, 1989, 1994). After the electrodes were attached, the subject was seated in a comfortable chair, in front of a table on which the four decks of cards were placed. The decks were placed in a very close proximity to the subject, so that a minimal movement was required to select the card from the chosen deck. The subject was asked to be quiet and relaxed at all times, and not engage in activities such as exchanging money or adjusting the decks.

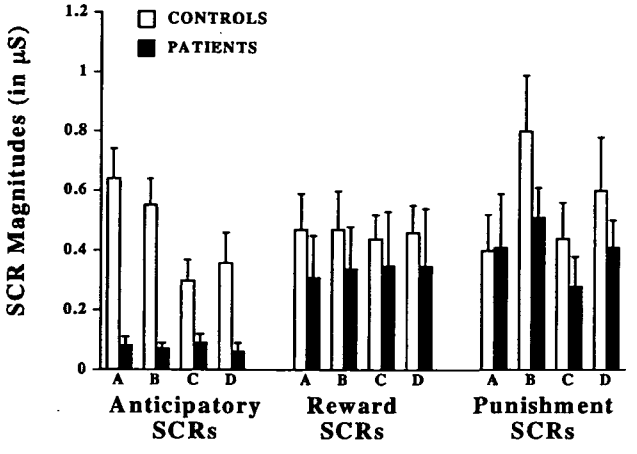

Figure 3. Means ± SEM of the magnitudes of anticipatory, reward, and punishment SCRs generated by normal controls (n = 12) and target patients (n = 7) averaged across all cards selected from a given deck.

All such activities were carried out by an experimenter who sat across the table from the subject, and delivered the card game task. Most importantly, the subject was instructed to select a card, only when a second experimenter (monitoring the SCR polygram) said the word "go." Otherwise, the subject must wait and consider which deck to choose next. This SCR recording ensured that every turn of a card from any deck (signalled by the word "go") coincided with a mark on the SCR polygram. Thus, SCRs generated in association with a specific card from a specific deck could be precisely identified on the polygram. The time interval between two card selections varied to some extent but conformed to two general rules: it was never shorter than 15 sec, and a "go" signal was not given if the subject was generating an SCR or was in the steep recovery limb of an SCR.

The SCRs generated during the task were divided into three categories: (1) reward SCRs were those generated after turning cards for which there was a reward and no penalty; (2) punishment SCRs were those generated after turning a card for which there was a reward followed immediately by a penalty; (3) anticipatory SCRs were those generated prior to turning a card from any given deck, during the time period the subject pondered from which deck to choose. Based on the results of pilot experiments, the time window for the rising onset of reward and punishment SCRs was within 4 sec after receipt of information (e.g., you have won X amount, or you have won X . . . but lost Y amount). SCRs generated during the time between the completion of money exchange and the next card selection were considered anticipatory SCRs.

Results

Both patients and controls generated SCRs in reaction to reward and punishment (Fig. 3). Controls, however, as they became experienced with the task, also began to generate SCRs prior to the selection of some cards (Fig. 4*A*). Patients entirely failed to generate such anticipatory SCRs (Fig. 4*B*). A two-way ANOVA on the means of anticipatory SCRs generated by controls and patients in association with decks A, B, C, or D revealed a significant main effect of group [$F(1,17) = 12.8, p < 0.01$], deck [$F(3,51) = 5.7, p < 0.01$], and interaction of group with deck [$F(3,51) = 6.1, p < 0.01$], suggesting that controls, but not patients, generated anticipatory SCRs in relation to decks A and B (disadvantageous decks), which were higher than those in relation to decks C and D (advantageous decks) (Fig. 3). For the control group, Newman-Keuls tests on the anticipatory responses from decks A and B as compared to decks C and D were highly significant ($ps < 0.001$). For the patient group, Newman-Keuls tests on the anticipatory SCRs were all nonsignificant. A similar two-way ANOVA on the means of reward SCRs did not reveal a significant main effect of group, or deck, or an interaction of group with deck. The reward SCRs of controls were higher than those of target patients (see Fig. 3), but these differences were not statistically significant. Finally, the two-way ANOVA on punishment SCRs also did not reveal a significant main effect of group, or an interaction of group with deck, but it did reveal a significant main effect of deck [$F(3,48) = 3.4, p < 0.05$], suggesting that the magnitudes of punishment SCRs generated in association with certain decks, specifically those in association with deck B (less frequent punishment but of high magnitude) were generally higher than those of decks A and C (more frequent punishment but of small magnitude), as revealed by Newman-Keuls tests ($ps < 0.05$) (Fig. 3). Figure 4*A* reveals that the anticipatory SCRs generated by controls (1) develop over time (i.e., after selecting several cards from each deck, and thus encountering several instances of reward and punishment), and (2) become more pronounced prior to the selection of cards from the disadvantageous decks (A and B). No such SCRs are present in target subjects (Fig. 4*B*).

Discussion

In our attempt to explain these results, we considered several possibilities. First, we excluded the possibility that patients fail to generate anticipatory SCRs because they can no longer activate SCRs, since the patients could generate reward and punishment SCRs appropriately. Although the magnitude of reward and punishment SCRs of patients was slightly smaller than that of controls, the differences were not statistically significant, and could not account for the complete failure of patients to generate any anticipatory SCRs.

Next, we excluded the possibility that the SCRs measured in the present study were confounded by motion artifact, for example, moving the hand to pick up a card from a given deck. To begin with, the anticipatory SCRs were usually generated before any movement was made toward a card. Second, although the subjects' motion was always the same, no anticipatory SCRs were observed for the early cards in the game. Third, the magnitude of SCRs was not equal for decks A and B compared to C and D. Fourth, the anticipatory SCRs generated in association with the earlier cards of decks A and B were far smaller than those of the later cards, suggesting that these SCRs were related to continued learning and experience with reward and punishment.

The possibility that the reward and punishment SCRs generated after the selection of a card were influenced by movement artifact can also be excluded. All the SCRs included in the results were recorded from the nonmoving hand, a procedure that has been shown to eliminate movement artifact (Raskin, 1973; Venables and Christie, 1973; Boucsein, 1992). Besides, movement artifact would not explain why punishment SCRs have their highest peak after cards followed by the highest magnitude of punishment (deck B). Moreover, SCRs elicited by movement normally habituate and disappear over repeated trials, unlike the reward and punishment SCRs observed in our experiment, which persisted throughout the experiment.

It is important to distinguish the nature of the SCR impairment observed in this study from that reported in early studies of patients with large frontal lobe tumors. Those patients showed defects in the vegetative components of the orienting reflex, namely pathologically rapid habituation to signal stimuli (Luria et al., 1964; Luria and Homskaya, 1970; Luria, 1973). Comparable results have also been reported in monkeys (Grueninger et al., 1965; Kimble et al., 1965). Unlike those early studies, however, which focused on arousal and the orienting response, we have identified a specific cognitive condition under which SCRs were abnormal. The lesion sites that correlate with these abnormalities (Damasio et al., 1990; Tranel and Damasio, 1994, the present study) are also distinct from those of the early studies.

Passive avoidance involves withholding a response to

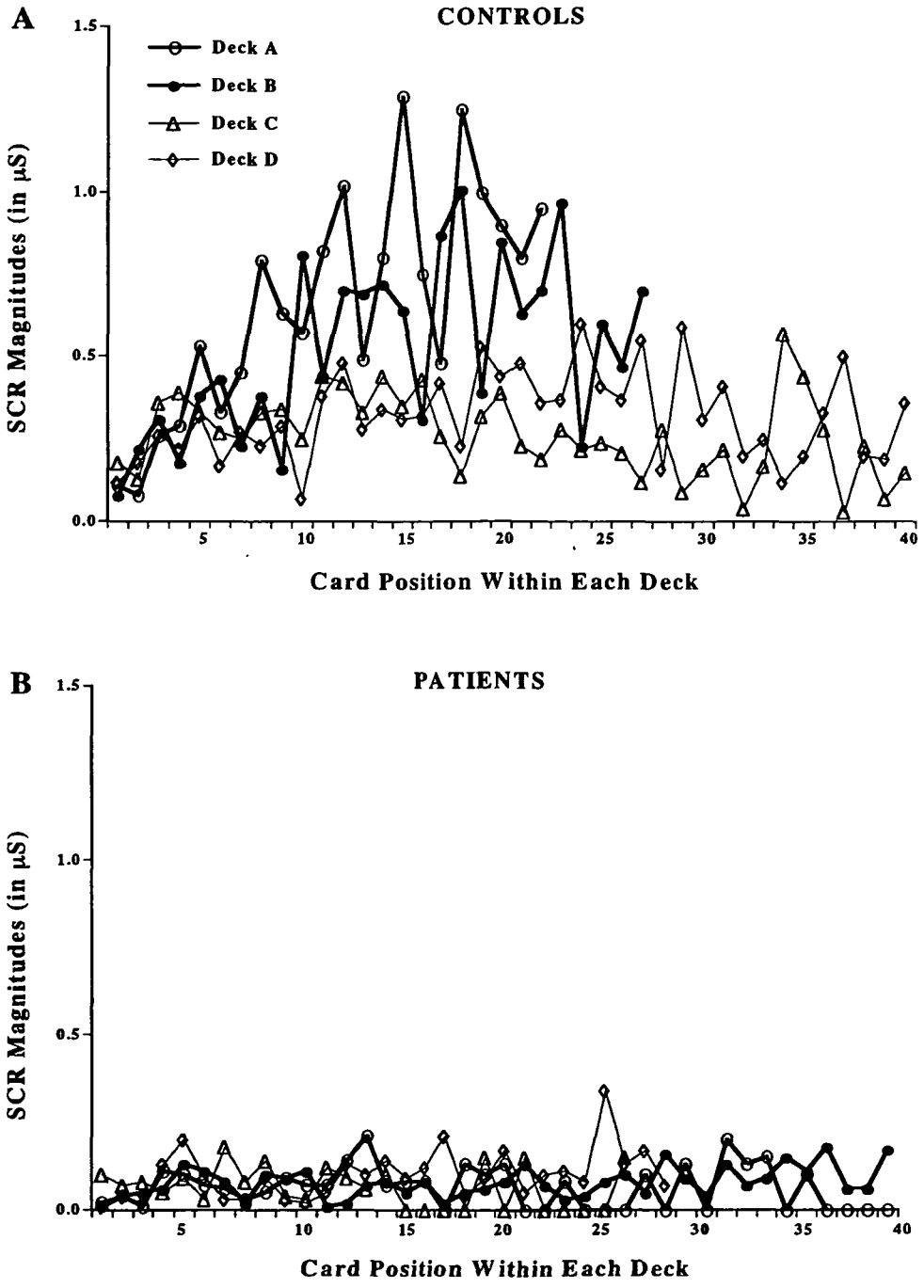

Figure 4. *A*, Profiles of the magnitudes of anticipatory SCRs generated by controls prior to their selection of the card indicated on the x-axis in a given deck. Each point represents the mean of the peak amplitudes of SCRs generated by (n = 2–12) subjects. *B*, Similar profiles of anticipatory SCRs generated by patients (n = 2–7). Two important points are revealed: (1) The performance profiles of controls and patients on the task are comparable to those obtained in a previous study (Bechara et al., 1994), in that normal controls make more selections from the good decks (*C* and *D*) as opposed to the bad decks (*A* and *B*). Patients do exactly the opposite. (2) Controls develop anticipatory SCRs that become very pronounced prior to the selection of cards from decks *A* and *B*, relative to *C* and *D*. Patients fail to do so.

avoid punishment, and an impairment in passive avoidance learning might be at the center of the patients' defect. Because patients with prefrontal lesions do not seem to learn from previous mistakes, and they frequently engage in behaviors that lead to negative consequences, the issue of passive avoidance learning must be considered. Yet we are not aware of any reports concerning human prefrontal damage and the relationship between punishment, autonomic arousal, and avoidance learning, although early investigations have addressed this relationship in sociopaths. The classic experiment on passive avoidance and autonomic arousal was reported by Lykken (1957), and the essential finding has been replicated by others (e.g., Schmauk, 1970). The experiments involved a mental maze task in which subjects were required to avoid making maze moves that led to punishment. Punishment consisted of an electric shock, or money loss, or being told you were wrong (social punishment). There were no differences between the control and sociopathic groups, in terms of their autonomic reactivity to punishment, as indexed by the SCR, although the magnitudes of SCRs in reaction to physical punishment were higher than those to money loss or social punishment. The sociopathic groups, however, were more likely to commit punished errors than controls, and the magnitudes of their SCRs in anticipation of a response previously paired with punishment were lower than those of controls (Schmauk, 1970). These results have intriguing parallels with the current findings, and are compatible with the notion that sociopathic behavior may depend on dysfunction in neural systems involving ventromedial frontal cortices among other regions (Damasio et al., 1990; Damasio, 1994).

It is important, nonetheless, to draw at least two distinctions between the behavioral task used in those early studies and our card task. First, unlike the immediate and consistent nature of punishment in the mental maze task, the delivery of punishment in our card task is delayed and inconsistent. There may be a potential advantage to the use of inconsistent punishment, since studies of psychopathic behavior suggest that psychopaths may respond to punishment when punishment is a virtual certainty, but fail to respond when punishment is uncertain (Siegel, 1978). Second, in addition to the "avoidance contingency" characteristic of the mental maze task, our task has a salient "approach contingency". That is, subjects are repeatedly rewarded for selecting cards from decks that also contain cards that lead to severe punishment. It is especially important to use conflicting reward/punishment contingencies in the patients under scrutiny, since individuals with disinhibited behavior demonstrate a passive avoidance deficit only when the avoidance requires them to inhibit a response that frequently led to reward (Newman et al., 1985).

There is a large body of literature on probability learning that considers the relationship between decision under certainty and thinking through disjunctions (see Shafir, 1994, for a review). It has been shown that people tend to make incorrect choices when thinking through disjunctions and deciding under uncertainty. It has been suggested that one reason for the difficulties that people have when deciding in these task situations is the sheer complexity that characterizes many questions on these tasks. However, difficulties in making correct decisions were observed in a number of simple contexts of decision and reasoning that do not seem readily attributable to complexity considerations (Shafir, 1994). Thus, we attribute these difficulties to the absence of an emotional experience with the context of the situation under which the subject must make a decision. For instance, in the Wason Selection Task (Wason and Johnson-Laird, 1972), one of the most popular probes of deductive reasoning, most subjects fail to select the logically correct answer, especially when the problem is abstract, unfamiliar, or has no social content. When the problem is familiar and has a social content, subjects yield a very high number of logically correct responses (Griggs and Cox, 1982). Although different theories have provided different explanations to this phenomenon (Cheng and Holyoak, 1989; Evans, 1989; Johnson-Laird and Byrne, 1991; Cosmides and Tooby, 1992), all the theories share in common the notion that familiarity and social context tend to promote a mode of reasoning in which prior experience with similar situations is used to decide in the task. Indeed, in experiments using modified problems of the Wason Selection Task, control subjects were presented with social problems, but some of the problems were likely to have been actually encountered by the subjects as opposed to some other problems (Adolphs et al., 1995). Control subjects tended to choose correctly on all the social problems, but they found the problems with familiar stories much easier than those with unfamiliar stories, and subjects did significantly better on the familiar than on the unfamiliar social problems (Adolphs et al., 1995). This facilitation was also observed when using problems with familiar, but nonsocial stories. On the other hand, while patients with ventromedial frontal lobe lesions performed similar to controls on most presented problems, they failed to benefit from the familiarity of the social and nonsocial stories, and thus failed to reason normally about familiar scenarios, both in the social and nonsocial domains (Adolphs et al., 1995). Together, these findings are consistent with the notion that an emotional/somatic signal derived from previous experiences with reward or punishment facilitates the implementation of an advantageous choice under conditions of uncertainty. Indeed, in experiments that used probabilistic classification learning tasks that provided feedback and an opportunity to learn from previous experience, both normal human subjects (Gluck and Bower, 1988) and amnesic patients (Knowlton et al., 1994) improved in their ability to select correct responses as information gradually accrued across trials. It would be interesting to see how patients with frontal lobe lesions perform on these probability learning tasks.

Thus, it appears that patients with prefrontal lesions do have a specific impairment in their ability to generate anticipatory SCRs in response to imagined scenarios. Since SCRs are physiologic indices of an autonomically controlled change in somatic state, it seems reasonable to conclude that the absence of anticipatory SCRs is an indication that these patients' ability to change somatic states in response to imagined scenarios is severely compromised. In this perspective, the failure to enact a somatic state appropriate to the consequences of a response would be a correlate of their inability to choose advantageously.

We suggest that somatic markers might assist the process of decision in a variety of ways. Covertly, somatic markers would (a) help inhibit the normal tendency to approach immediate reward, and (b) enhance and hold the representation of a future negative scenario in working memory, thus counterbalancing the automatic inclination to seek immediate gain. Overtly, the juxtaposition of a somatic marker to the representation of a future negative scenario would operate as an alarm signal by propitiating the inference that an option that causes immediate gain but future loss would be best avoided.

Why do patients fail to generate SCRs in anticipation of a response? The answer to this question, in neural terms, is that these patients have lost a critical system, centered in ventromedial prefrontal cortices, which normally (1) connects knowledge about the categorization of previous experiences, to different profiles of biological response including those that are part of an emotional response; and (2) has the ability to inhibit or activate the response appropriate to a given situation, by firing upon central bioregulatory structures such as those in the amygdalae and hypothalamus. The ventromedial frontal region is ideally suited to this role. It is the recipient, directly and indirectly, of signals pertaining to the soma and pertaining to nonsomatic sensory modalities; it is reciprocally interconnected with other prefrontal cortices and with subcortical structures involved in basic biological regulation, emotion, and social cognition and behavior, such as the amygdala and hypothalamus (see Goldman-Rakic, 1987; Barbas and Pandya, 1989; Pandya and Yeterian, 1990).

There is a different answer to the above question, in cognitive terms. An analysis of the learning contingencies required for the performance of the task, reveals the following: each deck contains both good cards and bad cards. Subjects never know whether the card that they are about to turn, from any deck, will be good or bad, nor can they calculate exactly how much is being gained or lost in each deck, at any time. How can the subjects develop, then, a sense of "goodness" or "badness" relative to each deck?

One possibility is that the subjects make a cognitive determination of the badness versus goodness of a given deck, based on a gross estimate of the number and magnitude of penalties for each deck. This estimate might influence behavior by direct conscious deliberation, or it might automatically result in a negative somatic state, which would promote the avoidance of the bad decks. In this formulation, reasoning towards a cognitive estimate precedes somatic signalling. However, somatic signalling would still be critical for the implementation of advantageous responses, since our observations reveal that ventromedial frontal lobe patients do not

choose advantageously, even when they eventually know which decks may be good and which ones may be bad.

But there is an alternate formulation. Since each deck has multiple occurrences of reward and punishment, these occurrences produce multiple good and bad somatic states. Later, when attention falls on a given deck, a somatic state would be triggered that best represents the balance of goodness and badness previously generated by that deck. For example, because in decks A and B the overall punishment is more severe than the reward, it is possible that negative somatic states overshadow positive ones, when these decks are being considered as a choice. Anticipatory SCRs generated prior to card selections would be an expression and an index of this process. This formulation would explain why the magnitude of the anticipatory SCRs increases with the experience of the player, and why it is especially marked for the disadvantageous decks. A negative somatic state as hallmarked by an anticipatory SCR, would nonconsciously "advise" the avoidance of the disadvantageous decks (A and B), while helping bring on line, cognitively the reasons for making the avoidance explicit. In brief, in this formulation, somatic signalling and covert estimation of goodness or badness precede conscious cognitive processing.

To conclude, ascribing a good or bad value to a given deck requires a mechanism for weighing the overall proportion of reward versus punishment within that deck, that is, an association between a stimulus (a given deck) and either its goodness (average reward value) and badness (average punishment value). We propose that in normal individuals, the ventromedial prefrontal cortices contain neural circuitry that links the stimulus configuration of a given deck (neutral stimulus), to the representations of both reward and punishment, of goodness and badness.

Notes

This work was supported by NINDS PO1 NS19632, James S. McDonnell Foundation, and MRC (Canada).

Address correspondence to Antoine Bechara, Department of Neurology, University of Iowa Hospitals and Clinics, Iowa City, IA 52242.

References

Adolphs R, Bechara A, Tranel D, Damasio H, Damasio AR (1995) Neuropsychological approaches to reasoning and decision-making. In: Neurobiology of decision-making. (Damasio AR, Damasio H, Christen Y, eds), pp 157–179. New York: Springer.

Anderson SW, Damasio H, Tranel D, Damasio AR (1992) Cognitive sequelae of focal lesions in ventromedial frontal lobe. J Clin Exp Neuropsychol 14:83.

Barbas H, Pandya DN (1989) Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. J Comp Neurol 286:353–375.

Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50:7–12.

Benton AL, Hamsher KD, Varney NR, Spreen O (1983) Contributions to neuropsychological assessment. New York: Oxford UP.

Boucsein W (1992) Electrodermal activity. New York: Plenum.

Cheng PW, Holyoak KJ (1989) On the natural selection of reasoning theories. Cognition 33:285–313.

Cosmides L, Tooby J (1992) Cognitive adaptations for social exchange. In: The adapted mind: evolutionary psychology and the generation of culture (Barkow JH, Cosmides L, Tooby J, eds), pp 163–228. New York: Oxford UP.

Damasio AR (1994) The somatic marker hypothesis. In: Descartes' error: emotion, reason, and the human brain. New York: Grosset/Putnam.

Damasio AR, Tranel D, Damasio H (1990) Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. Behav Brain Res 41:81–94.

Damasio AR, Tranel D, Damasio H (1991) Somatic markers and the guidance of behavior. In: Frontal lobe function and dysfunction (Levin H, Eisenberg H, Benton A, eds), pp 217–228. New York: Oxford UP.

Damasio H, Damasio A (1989) Lesion analysis in neuropsychology. New York: Oxford UP.

Damasio H, Frank R (1992) Three-dimensional *in vivo* mapping of brain lesions in humans. Arch Neurol 49:137–143.

Evans JStBT (1989) Bias in human reasoning: causes and consequences. Hove, UK: Erlbaum.

Fuster JM (1989) The prefrontal cortex, 2d ed. New York: Raven.

Gluck MA, Bower GH (1988) From conditioning to category learning: an adaptive network model. J Exp Psychol 117:227–247.

Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: Handbook of physiology, Vol 5, The nervous system (Plum F, ed), pp 373–401. Bethesda, MD: American Physiological Society.

Goldman-Rakic PS (1992) Working memory and the mind. Sci Am 267:111–117.

Griggs RA, Cox JR (1982) The elusive thematic-materials effect in Wason's selection task. Br J Psychol 73:407–420.

Grueninger WE, Kimble DP, Grueninger J, Levine S (1965) GSR and corticosteroid response in monkeys with frontal ablations. Neuropsychologia 3:205–216.

Johnson-Laird PN, Byrne RMJ (1991) Deduction. Hove, UK: Erlbaum.

Kimble DP, Bagshaw MH, Pribram KH (1965) The GSR of monkeys during orienting and habituation after selective partial ablations of the cingulate and frontal cortex. Neuropsychologia 3:121–128.

Knowlton BJ, Squire LR, Gluck MA (1994) Probabilistic classification learning in amnesia. Learn Memory 1:106–120.

Lezak MD (1995) Neuropsychological assessment, 3d ed. New York: Oxford UP.

Luria AR (1973) The frontal lobes and the regulation of behavior. In: Psychophysiology of the frontal lobe (Pribram KH, Luria AR, eds), pp 3–26. New York: Academic.

Luria AR, Homskaya ED (1970) Frontal lobes and the regulation of arousal processes. In: Attention: contemporary theory and analysis (Mostofsky DI, ed), pp 303–330. New York: Appleton-Century-Crofts.

Luria AR, Pribram KH, Homskaya ED (1964) An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). Neuropsychologia 19:65–71.

Lykken DT (1957) A study of anxiety in the sociopathic personality. J Abnormal Soc Psychol 55:6–10.

Newman JP, Widom CS, Nathan S (1985) Passive avoidance in syndromes of disinhibition: psychopathy and extraversion. J Personal Soc Psychol 48:1316–1327.

Pandya DN, Yeterian EH (1990) Prefrontal cortex in relation to other cortical areas in rhesus monkey: architecture and connections. Prog Brain Res 85:63–94.

Posner MI (1986) Chronometric explorations of the mind. New York: Oxford UP.

Raskin DC (1973) Attention and arousal. In: Electrodermal activity in psychological research (Prokasy WF, Raskin DC, eds), pp 125–155. New York: Academic.

Saver JL, Damasio AR (1991) Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. Neuropsychologia 29:1241–1249.

Schmauk FJ (1970) Punishment, arousal, and avoidance learning in sociopaths. J Abnorm Psychol 76:325–335.

Shafir E (1994) Uncertainty and the difficulty of thinking through disjunctions. Cognition 50:403–430.

Siegel RA (1978) Probability of punishment and suppression of behavior in psychopathic and nonpsychopathic offenders. J Abnorm Psychol 87:514–522.

Tranel D (1994) "Acquired sociopathy": the development of sociopathic behavior following focal brain damage. In: Progress in experimental personality and psychopathology research, Vol 17 (Fowles DC, Stucker P, Goodman SH, eds), pp 285–311. New York: Springer.

Tranel D (1996) The Iowa-Benton school of neuropsychological assessment. In: Neuropsychological assessment of neuropsychiatric disorders, 2d ed (Grant I, Adams KM, eds), in press. New York: Oxford UP.

Tranel D, Damasio AR (1988) Non-conscious face recognition in patients with face agnosia. Behav Brain Res 30:235–249.

Tranel D, Damasio H (1989) Intact electrodermal skin conductance responses in a patient with bilateral amygdala damage. Neuropsychologia 27:381–390.

Tranel D, Damasio H (1994) Neuroanatomical correlates of electrodermal skin conductance responses. Psychophysiology 31:427–438.

Venables PH, Christie MJ (1973) Mechanisms, instrumentation, recording techniques, & quantification of responses. In: Electrodermal activity in psychological research (Prokasy WF, Raskin DC, eds), pp 1–124. New York: Academic.

Wason PC, Johnson-Laird PN (1972) Psychology of reasoning: structure and content. London: Batsford.