QUALITY OF STIMULI AND PREFRONTAL LESIONS EFFECTS ON REVERSAL LEARNING IN GO–NO GO AVOIDANCE REFLEX DIFFERENTIATION IN CATS

Kazimierz ZIELIŃSKI and Julitta CZARKOWSKA

Department of Neurophysiology, Nencki Institute of Experimental Biology
Warszawa, Poland

Abstract. Go-no go avoidance reflex differentiation of two acoustic stimuli has been previously established in cats. Then the signalling properties of the conditioned stimuli were reversed and the course of acquisition of a new go-no go differentiation was studied in normal cats and in cats with prefrontal lesions. The hypothesis based on stimulus intensity dynamism theory was: (i) in normal cats acquisition of the new differentiation would be easier if the more effective stimulus of the pair were used as the positive stimulus, and (ii) in prefrontal cats this effect of stimulus quality would be less or absent. Results confirmed both predictions. The differences in rapidity of reversal learning were almost exclusively due to differences in responding to the new positive stimulus. Prefrontal cats were not deficient in comparison with normal cats in their inhibitory abilities in spite of the fact that extinction of bar-pressing to the new negative conditioned stimulus was much slower than transfer of the avoidance response to the new positive stimulus. At the beginning of reversal learning two opposite changes in responding on no-go trials were observed: (i) a decrease in the number of long-latency responses, which reflected the changed signalling value of the previously positive stimulus, and (ii) an increase in the number of short-latency responses, which was positively correlated with the increase in rate of intertrial responding. The increases in number of short-latency responses to the new negative stimulus and in rate of intertrial responding observed at the beginning of reversal learning were smaller in prefrontal than in normal cats. Results of the experiment indicate that the "drive disinhibition hypothesis" does not account for the effects of prefrontal lesions on avoidance behavior.

INTRODUCTION

In a long series of experiments performed chiefly in this Department on dogs and in the Section on Neurophysiology of the National Institute
of Mental Health in Bethesda on monkeys it was firmly established that prefrontal animals are deficient in retention of a go-no go differentiation task (7, 29, 34, 42). The symptoms were labeled as the “disinhibitory prefrontal syndrome”, since only small changes in execution of the conditioned response (CR) on positive trials 1 were observed; whereas the ability of the prefrontal animals to inhibit performance of the same CR on negative trials and during the intertrial interval was poor.

In earlier papers a similarity between the symptoms observed after prefrontal lesions in the go-no go differentiation task and in the conditioned inhibitor task were described (5), and it was shown that both classically conditioned and instrumental reflexes were disinhibited when food was used as an incentive (6). In further experiments the loci responsible for the syndrome (11, 13, 40, 41) and effects of some experimental variables on the severity of impairment (12, 17) were determined.

In contrast to experiments with alimentary reflexes, the disinhibitory prefrontal syndrome has rarely been found in experiments with defensive reflexes. Although Auleytner and Brutkowski (4) obtained the disinhibitory syndrome in prefrontal dogs, in other studies with classically conditioned defensive reflexes only minute or inconsistent changes in performance were observed after prefrontal lesions in dogs (1, 38, 39). When the avoidance procedure was used, the performance on negative trials was not impaired after prefrontal lesions, either in dogs (2) or in cats (47, 50, 51).

This discrepancy between results obtained in experiments with alimentary and with defensive reflexes raises some problems for hypotheses concerning the mechanisms by which the frontal lobes may influence the ability to inhibit the execution of the CR on negative trials. The “drive disinhibition” theory proposed by Brutkowski (7, 8, 12) postulates that after removal of the medial aspect of the prefrontal cortex in dogs or after lesions in the orbitofrontal cortex in monkeys the drive functions are released from cortical inhibitory control and that this causes impairment of performance on negative trials. In the case of defensive reflexes

---

1 We follow the terminology in which the conditioned stimulus (CS) is called positive or CS+ if it is associated with the presentation of the unconditioned stimulus (US), independently of whether the “reinforcing” agent is attractive or aversive, under probability p higher than zero (0 < p ≤ 1). The CS associated with presentation of the US under probability p = 0 is called negative or CS−. In the go-no go differentiation procedure reinforced and unreinforced CSs are presented in random succession or in predetermined order. The positive trials started with the CS+ onset and the negative trials started with the CS− onset. During the reversal learning the same pair of CSs is used but their signalling properties are reversed, thus the previous CS+ becomes the CS− and, consequently, the previous CS− becomes the CS+. 

and, particularly, the avoidance reflex, release of drive brings into action more primitive modes of defensive behavior which interfere with skilled avoidance performance, thus preventing its appearance on disinhibited negative trials and leading to some decrease of performance on positive trials. Similar reasoning has been proposed to explain changes in performance observed after prefrontal lesions in cats and rabbits trained with the avoidance method (8, 9, 14).

It has been argued before that changes in the avoidance reflex (44) or in the go–no go avoidance reflex differentiation (50) observed after prefrontal lesions cannot be covered by the "drive-disinhibition" theory. However, the strength of these arguments may be weakened by the fact that in the studies cited above the retention of the preoperatively acquired avoidance reflexes was tested. Thus nociceptive shock stimulation was used rather infrequently both before and after the lesions. A suggestion was raised, essentially within the framework of the "drive-disinhibition" theory, that strength of the reinforcement may be the main parameter influencing the presence of the disinhibition prefrontal syndrome in defensive reflexes so that the syndrome would be observed when the nociceptive stimulation is used rather regularly (27). A counter-argument was provided by experiments with the bar-pressing escape response which terminated shock applied on each trial both before and after prefrontal lesions. Such an escape reflex showed perfect retention after prefrontal lesions in cats (45, 47).

It seems that even stronger evidence against the usefulness of the "drive-disinhibition" theory for the postoperative changes in the instrumental defensive reflexes may be obtained if the signalling properties of the conditioned stimuli (CS) used in the go–no go differentiation are reversed. In effect, the acquired tendencies to respond to one stimulus and not to the other must be reversed. To solve this task the CR has to be transferred from the previously positive conditioned stimulus (CS+) to a new positive stimulus (CS'+), and to be inhibited to a new negative stimulus CS'−, which was previously the CS−. In the case of reversal learning after the go–no go avoidance reflex differentiation, the frequency of nociceptive stimulation drastically increases and the stronger the perseveration of non-responding to the CS'+, which was previously the CS−, the higher is the frequency of shocks received. Thus, the inhibitory abilities of the animals would be tested against a strong excitatory background in this task.

The aim of the present study was to compare the course of reversal learning in normal cats and in cats with prefrontal lesions. Since it is widely accepted (see 8, 34) that changes in performance in prefrontal animals cannot be attributed to a sensory defect, we were not interested
in their discriminatory abilities. Hence, two easily discriminable CSs, tone and click, were used. These stimuli differed in quality and also in their inherent effectiveness, or, following Pavlovian terminology, in reflexogenic strength. Acquisition of the avoidance reflex was more rapid when the click was used as the CS+, and some differences in responding on negative trials were observed depending on whether tone or click was used as the CS− (50). These differences in the course of the go–no go differentiation were assumed to depend on the stimulus intensity dynamism effect, since it was found in different species that differentiation training is easier when the stronger stimulus from the pair is used as the CS+ than when the weaker is used as the CS+ (3, 23, 24, 32, 35, 36, 43).

It may be expected that this rule operating during training in differentiation will also have an effect on the course of reversal learning. Thus, reversal learning would be more difficult in animals which were previously trained to differentiate the stronger CS+ and the weaker CS− than in animals which had been trained to differentiate the weaker CS+ and the stronger CS−. In other words, if the same pair of stimuli is used, in the first group of animals the intensity relations between the CS+ and CS− will facilitate original differentiation and make the reversal learning more difficult; whereas in the second group after learning a more difficult differentiation the reversal learning task ought to be easier.

This problem was not considered in earlier studies on reversal learning, although there is some information showing dependence of the rapidity of reversal learning on the properties of the CS (19, 21).

It is of interest whether or not this asymmetry in acquisition of the differentiation (and, presumably, in reversal learning) depending on intensity relations between a pair of stimuli will operate not only in normal but also in prefrontal animals. It was found in cats that the higher the intensity of the white noise used as the CS+, the greater was the proportion of avoidance responses executed with short latencies (46–49). The stimulus intensity effect was more easily detected by this index than by the commonly used measure as length of training necessary to reach criterion of CR acquisition (46, 48, 50). After removal of the proreal and orbital gyri in cats, the ability to perform short-latency avoidance responses was markedly deteriorated and did not recover during the postoperative training (45, 47, 48). It seems that the stimulus intensity effect was attenuated after the prefrontal lesion in cats although it was shown that the group differences related to the CS+ intensity were preserved after such a lesion (48).

In light of these considerations it is probable that in normal animals during reversal learning differences due to the intensity relations bet-
ween the CSs would be similar to those described for differentiation training; whereas in frontal animals such differences might be less or absent. This study was designed to answer these questions.

MATERIAL AND METHODS

Experiments were carried out on 16 adult male cats. They were trained in a cage, $55 \times 55 \times 40$ cm, having a bar, $10 \times 2$ cm, located 10 cm above the grid-floor in the middle of an oblong wall of a cage. The cage was placed in a sound-proof CR chamber. The sources of conditioned stimuli were located near the center of the ceiling of the box, a loudspeaker through which a mild tone of about 65 db and 100 cycle/sec from the tone generator was presented and an electromagnetic relay, which produced a rhythmic 5/sec click. The grid-floor was activated with 50 cycle/sec a-c from a circuit in which 100,000 ohm resistance was connected in series with the cat to apply shock to the paws of the animal. For most cats 2.5 ma shock was used.

The experiment consisted of several stages: (i) acquisition of avoidance CR, (ii) differentiation training, (iii) surgery and/or rest period, (iv) testing of postoperative retention, and (v) the reversal learning. Only the final stage of the experiment is presented here, the remainder have been previously reported (50).

A $2 \times 2$ experimental design was used. Four experimental groups, four cats each, differed in CS used (tone vs. click) and in treatment after reaching criterion in differentiation training (surgery and rest period vs. rest period alone). The conditioned stimuli used for each group are indicated in Table I. The positive trial started with the CS$^+$ onset and was terminated with the bar-pressing response. A bar press executed within 5 sec after CS onset terminated the CS$^+$, prevented the activation of the grid floor, and was scored as an avoidance response. If an avoi-
dance response did not occur within 5 sec after the CS\textsuperscript{+} onset, the grid floor was activated and both stimuli lasted until the S performed the bar-press; this was labelled as an escape response. The negative trials consisted of a 5 sec presentation of the CS\textsuperscript{−}; no shock was applied in these trials and the duration of the CS\textsuperscript{−} was independent of the S's behavior.

Cats from the Groups 1F and 2F received prefrontal lesions (bilateral removal of the proreal and orbital gyri by suction under Nembutal anesthesia in one operation) the day after they reached the criterion of differentiation training. As a criterion of the differentiation at least 90 avoidance responses to the CS\textsuperscript{+} and no more than 10 bar-presses to the CS\textsuperscript{−} presentations during 10 consecutive daily sessions were required. Reconstructions of the lesions have been published before (50) and an example of a typical lesion is shown in Fig. 1. Postoperative retention testing began 10 days after surgery. The Ss from Groups 1N and 2N were left intact, but they rested in their home-cages for 10 days before retention testing. Retention testing was carried out until the same 90\% criterion was met again. The next day reversal learning was begun and was continued for 30 sessions.

During the acquisition of avoidance reflex each daily session consisted of 10 positive trials. In subsequent stages of the experiment 10 positive and 10 negative trials were presented daily. The order of positive and inhibitory trials was changed each day according to the Gellerman series (22). The length of intertrial intervals was 40, 60, and 80 sec, randomly distributed.

The main indices of behavior were: (i) number of avoidance responses executed to the CS\textsuperscript{+}, (ii) number of disinhibited negative trials, (iii) latencies of the bar-pressing responses executed in presence of the CS,
and (iv) number of intertrial responses (ITR). Latencies of responses were measured to the nearest 0.2 sec. To analyse the course of reversal learning, the 30 sessions were divided into six blocks of 5 sessions each and compared with the last 5 retention sessions (Block 0).

RESULTS

Transfer of the avoidance response to the CS'⁺

The course of reversal learning is indicated in Fig. 2, for each group independently, by the mean percentages of bar-pressing responses performed on positive and on negative trials in consecutive blocks of sessions. As seen from Fig. 2, transfer of the CR from CS⁺ to the new positive stimulus CS'⁺ was very rapid; even in Block 1 the level of performance in each group exceeded the 50% level. In most cats the first avoidance response was observed on the 3rd–6th presentation of the CS'⁺, except for Ss from Group 1N in which the first avoidance response to the tone

![Fig. 2. Group data showing the course of reversal learning indicated by the mean percentages of avoidance responses to positive CS (black triangles and circles) and bar-pressing responses to negative CS (open triangles and circles) in consecutive blocks of sessions.](image-url)
CS$^+$ did not appear until the second session of the reversal learning. This group difference was significant statistically as shown by $2 \times 2$ analysis of variance in which both main effects were at a chance level, but the CS quality and treatment (surgery vs. rest period) interaction was significant at $p < 0.05$ level. This result cannot be explained by differences in the performance to the CS$^-$ before reversal learning since during Block 0 no one animal pressed the bar more than three times in 50 negative trials, with group means ranging from 0.25 to 1.0 presses.

The retardation in transfer of the avoidance response to the tone CS$^+$ shown by normal animals at the very beginning of the reversal did not compensate in further learning, as seen from the level of performance in consecutive blocks of sessions (Fig. 2). During 30 sessions of reversal learning only one cat from Group 1N reached the criterion of 90 avoidance responses in 100 consecutive positive trials. All Ss from Group 2N, two Ss from Group 1F and two Ss from Group 2F mastered the same criterion.

Fig. 3. Rate of responses during intertrial intervals in consecutive blocks of sessions, before (Block 0) and in the course of reversal learning (Blocks 1–6).
Group differences were also observed in general emotionality indicated by the intertrial responses (ITR). At the beginning of reversal learning the frequency of shock application was much higher than before, and due to this an increase in the ITR rate was observed in all groups, more pronounced in normal Ss than in prefrontal cats (Fig. 3). A $2 \times 2$ analysis of variance, in which raw data were ITR rates in Block 1, showed that effect of treatment was significant at the $p < 0.05$ level; effect of CS quality was not significant statistically and interaction of the two main effect was significant at $p < 0.025$.

It seems that the retardation of Group IN in transfer of the avoidance response to the tone CS' cannot be related to group differences in the ITR rate. As seen from Fig. 3, the highest ITR rate at the beginning of the reversal learning was observed in Group 2N, and the lowest in Group 1F, both of them showed a higher level of performance on positive trials than Group 1N. Analysis of individual data indicate that the level of performance of avoidance responses to the CS' during Block 1 was only slightly correlated with the increase in ITR rate in Block 1 over the ITR rate in Block 0 ($r_s = +0.50$, $p < 0.10$, $n = 16$, Spearman rank order correlation coefficient).

In the course of reversal learning the ITR rate gradually decreased and the blocks effect was significant at $p < 0.01$ (analysis of variance; mixed design, Type III) (see 31). In this analysis, data from all six blocks of the reversal learning were taken into account.

Figure 3 shows several differences between normal cats and those with lesions: lower ITR rate in normal cats before reversal, more pronounced increase at the beginning, and more rapid decrease of the ITR rate in the course of reversal learning in Groups 1N and 2N.

Some differences were also observed between normal and frontal cats in distribution of response latencies to the CS' (Fig. 4). Prefrontal cats performed a smaller proportion of short-latency avoidance responses than normal animals trained to the same CS', both at the beginning and toward the end of the reversal learning. These differences were tested statistically using the Smirnov two-tailed test (16), and results of analyses are presented in Table II. At the beginning of the reversal learning the largest vertical distance ($D_{\text{max}}$) between cumulative distributions of response latencies to the same CS' estimated for normal and prefrontal cats was observed close to the moment of the shock application; thus analyses for Block 1 correspond roughly to differences in the level of performance in positive trials. As seen from Fig. 4 and Table II, normal cats trained with the tone CS' were worse than prefrontal cats, and normal cats trained with the click CS' were better than prefrontal cats trained to the same stimuli.
In the course of the reversal learning the proportion of short-latency avoidance responses increased in all groups, and this change was significant statistically (blocks effect $p < 0.001$, analysis of variance, mixed design, Type III) (see 31$^2$). This increase in proportion of short-latency avoidance responses was greater in normal animals (however the treat-

$^2$ In this analysis the raw data were numbers of avoidance responses performed in consecutive blocks of sessions with latencies shorter than 2.3 sec since it was found before (50) that at this point the largest vertical distance ($D_{max}$) between cumulative distributions of response latencies to the tone and the click CS on the avoidance acquisition stage was observed.

Fig. 4. Cumulative frequency distributions of the latencies of bar-pressing responses to the CS' at the beginning (Block 1) and toward the end (Block 6) of the reversal learning.
TABLE II

Effect of the quality of CS' + on latencies of the instrumental response in positive trials at the beginning (Block 1) and toward the end (Block 6) of reversal learning. \( S_{1N} \) denote the empirical distribution function of response latencies for Group 1N.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Direction of difference</th>
<th>Point of maximum discrepancy (sec)</th>
<th>( D_{\text{max}} )</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1N vs. 1F Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block 1</td>
<td>( S_{1N} &lt; S_{1F} )</td>
<td>5.3</td>
<td>0.155</td>
<td>( p &lt; 0.025 )</td>
</tr>
<tr>
<td>Block 6</td>
<td>( S_{1N} &gt; S_{1F} )</td>
<td>1.1</td>
<td>0.175</td>
<td>( p &lt; 0.005 )</td>
</tr>
<tr>
<td>2N vs. 2F Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block 1</td>
<td>( S_{2N} &gt; S_{1F} )</td>
<td>4.7</td>
<td>0.270</td>
<td>( p &lt; 0.001 )</td>
</tr>
<tr>
<td>Block 6</td>
<td>( S_{2N} &gt; S_{2F} )</td>
<td>2.5</td>
<td>0.370</td>
<td>( p &lt; 0.001 )</td>
</tr>
</tbody>
</table>

Inhibition of the avoidance response to the CS'–

Extinction of the CR to the previously positive stimulus was very slow. As seen from Fig. 2, during the Block 1 of reversal learning more than 80% of negative trials were disinhibited in each experimental group. In the last 10 sessions of the reversal learning only 6 cats (1 or 2 \( S_{5} \) from each group) performed less than 10 bar-presses to the CS'– and in 3 cats (2 from the Group 1N and 1 from the Group 1F) more than 50% of negative trials were disinhibited. In normal cats extinction of the avoidance response to the click CS'– was more difficult than to the tone CS'–; but in cats with frontal lesions differences between the two CS'– were much smaller and in the opposite direction. However, these differences were not significant statistically (analysis of variance, mixed design, Type III) (31).

At the beginning of the reversal learning the bar-pressing responses to the CS'– were performed even more vigorously then before when
onset of the same stimulus initiated positive trials. This was due to the increase in general emotionality of the cats caused by much higher frequency of shock application since they did not perform avoidance responses to the previously negative CS. In effect, if during the first sessions of the reversal learning the bar-pressing response to the CS'− occurred, its latency in most cases was shorter than to the same stimulus

Fig. 5. Cumulative frequency distributions of the latencies of avoidance responses to the CS+ before reversal learning (Block 0) and of the bar-pressing responses to the same stimulus after reversal of its signalling properties.
during Block 0. This may be seen in Fig. 5 in which distributions of latencies to the CS+ during Block 0 and to the CS'- during Blocks 1 and 6 are presented for each experimental group independently. In both groups of normal cats the proportion of short-latency responses to the CS' during Block 1 was greater than to the CS+ during Block 0, and group differences due to stimulus intensity dynamism were preserved even in the last sessions of the reversal learning. In prefrontal cats such a clear increase in the proportion of short-latency responses was observed only in Group 1F trained to the click CS', and group differences related to stimulus intensity dynamism disappeared toward the end of the reversal learning.

It has to be pointed out that by Block 1 some extinction of bar-pressing responses on negative trials was observed, as indicated by numbers of responses performed during the 5 sec of the CS'- which were in each group smaller than numbers of the avoidance responses to the same stimuli used as CS+ in Block 0. Comparison of the latency distributions for Block 0 and Block 1 reveal that changes in long-latency responses were in accordance with the reversed signalling properties of stimuli, but the increase of short-latency responses observed at the beginning of the reversal learning was opposite to that expected with the new signalling properties of the conditioned stimuli.

When data presented in Fig. 5 were compared with changes in the ITR rate (Fig. 3), a hypothesis emerged that the increase in proportion of short-latency responses to the CS' observed at the beginning of the reversal learning was related to the increase of the emotionality indicated by the ITR rate. To test this hypothesis for each S the $D_{max}$ between relative cumulative distributions of latencies of avoidance responses performed in Block 0 and of bar presses to the CS' in Block 1 were estimated. It was found that in all cats the proportion of short-latency responses was greater in Block 1 than in Block 0, and the Smirnov two-tailed test (16) showed that this increase was statistically significant at the $p < 0.05$ level or better in seven normal cats but only in two prefrontal cats. The points of the largest vertical distance varied in individual cats from 0.7 to 3.1 sec. Then the values of the $D_{max}$ were correlated with the increase in ITR rate (difference between the ITR rates in Block 1 and in Block 0) observed in a given cat. It was found that the greater the increase in the ITR rate, the higher was the value of $D_{max}$ ($r_s = +0.52$, $p < 0.05$, $n = 16$, Spearman rank correlation coefficient).

It is important to add that the number of bar presses performed during the 5 sec of the CS' in Block 1 was not related to the changes in the ITR rate ($r_s = +0.14$, $p > 0.10$, $n = 16$, Spearman rank correlation coefficient). This indicates that short-latency instrumental defensive
responses in cats tested in the described experimental situation are controlled by independent variables other than the over-all level of performance to a given CS.

Several measures of the speed of inhibitory training, i.e. the length of time required for reaching a session in which there was no bar-pressing response in five, seven or nine negative trials were tested statistically using $2 \times 2$ analysis of variance. In no one analysis did the main effects (CS quality and treatment) or their interaction reach significance.

Taking all data presented in this part of the paper into account, it can be inferred that ability to inhibit bar-pressing responses to the previously positive stimulus was equally good in prefrontal as in normal cats.

**Acquisition of the new click vs. tone differentiation**

During 30 sessions of reversal learning nine cats reached the criterion of 90 avoidance responses in 100 consecutive positive trials. Six cats mastered the criterion of 10 or less disinhibited negative trials in 100 consecutive presentations of the CS$^-$. However, only two (C-22/68 and C-38/67, both from Group 2N) showed the level of performance which was required as criterion for differentiation: 90% of correct responses both in positive and in negative trials in 10 consecutive daily sessions. Thus, this criterion measure was not sensitive enough for detailed analysis of group differences in the ability to acquire a new click vs. tone differentiation. In the original differentiation training most cats reached the same criterion during the first 10 sessions, and only one needed more than 30 sessions of training.

Before proceeding with further analysis of the data it was necessary first to consider the problem of measurement of the degree of differentiation accomplished by a given subject at each consecutive step of the reversal learning. Among many different methods which may be proposed, those measures will be presented which seem to be the most appropriate for the data collected.

1. Comparison of the response latencies on positive and negative trials. If it is possible to use the same measure of the strength of responses performed both to the CS$^+$ and to the CS$^-$ presentations, one may consider data collected on positive trials of a given session as one sample and data collected on negative trials of the same session as another sample. Different measures may be chosen depending on the nature of the response. In the reversal of the go-no go differentiation, when a high level of performance of the CR to both stimuli is observed, latencies of responses on positive and on negative trials may be compared and the degree of differentiation measured for each S and experimental session.
independently, using statistics appropriate for the two sample case. This method gives one a chance to discern differentiation of two stimuli in spite of 100% responding both on positive and on negative trials, it is a case in which response latencies to one CS are stochastically shorter than to the second CS. However, this measure will be biased when a CR does not occur on some positive and on some negative trials. According to the methods applied in this study the maximum duration of the negative trial was 5 sec; whereas the positive trial lasted until the cat performed the instrumental response. Thus, the response latencies to the CS were not longer than 5 sec, and on positive trials they had no limitations. When there was no response to the CS, we assumed that the response latency was 5.1 sec; and to minimize the bias related to theoretically much longer response latencies on positive trials, the non-parametric Mann-Whitney $U$ statistic was used for estimation of the difference between two sets of latencies.

The comparison of the response latencies on positive and negative trials was made for each experimental session of Block 0 and of reversal learning. Results of this analysis are presented in Fig. 6 in which levels of significance shown by the Mann–Whitney $U$ test are indicated by bars

---

**Fig. 6.** Individual data showing the course of the reversal learning as indicated by comparison of the response latencies on positive and on negative trials in each experimental session. Highest bars indicate difference in response latencies significant at $p < 0.01$; bars of intermediate height, significance at $p < 0.02$; lowest bars, significance at $p < 0.05$; and lack of a bar indicates no significant difference (Mann–Whitney $U$ test, two-tailed). The white bars denote sessions in which cats responded according to the signalling properties of CSs used in original differentiation, and the crossed bars denote sessions in which cats responded according to reversed signalling properties of CSs.
of different height. As seen from the Figure, at the beginning of the
reversal learning the differences in response latencies to the CS' + and
to the CS' - were in accordance with the old signalling properties of the
CS. Very soon, however, marked group differences appeared. Normal
cats trained to the click CS' + and tone CS' - (Group 1N) were as a rule
very persistent in responding according to the former signalling pro-
properties of the stimuli in spite of the fact that already during Block 1 50% of
avoidances occurred to the tone CS' + (see Fig. 2 and 4). In each cat from
this group the response latencies to the CS' + were only in isolated cases
stochastically shorter than response latencies to the click CS' -, even at
the end of the reversal learning. On the other hand, normal cats trained
to the tone CS' + and click CS' - (Group 2N) very quickly ceased to respond
to the old signalling values of these stimuli; and after some sessions of
"hesitation", they responded rather regularly with latencies stochastically
shorter to the CS' + than to the CS' -. Both groups with lesions shared
an intermediate position, they were neither so persistent in responding
to the old signalling properties of the stimuli as was Group 1N, nor so
regular in responding to the reversed signalling properties of the stimuli
as was Group 2N.

When the number of sessions, with significantly shorter response la-
tencies to the CS' + than response latencies to the CS' - were compared
using the $2 \times 2$ analysis of variance, it was found that the effect of
CS quality was significant at the $p < 0.005$ level and the CS quality and
treatment interaction — at the $p < 0.001$ level (Table III, Method A).

| TABLE III |
| Analysis of variance of number of sessions during reversal learning in which cats responded according to the new signalling properties of the conditioned stimuli as estimated by the three methods of measurement of differentiation |

<table>
<thead>
<tr>
<th>Groups</th>
<th>Method A</th>
<th>Method B</th>
<th>Method C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group means</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1N</td>
<td>2.00</td>
<td>7.00</td>
<td>0</td>
</tr>
<tr>
<td>2N</td>
<td>21.50</td>
<td>19.00</td>
<td>16.25</td>
</tr>
<tr>
<td>1F</td>
<td>19.25</td>
<td>14.75</td>
<td>16.25</td>
</tr>
<tr>
<td>2F</td>
<td>11.25</td>
<td>11.50</td>
<td>1.75</td>
</tr>
<tr>
<td>Source of variation</td>
<td>df</td>
<td>Values of the $F$ statistics</td>
<td></td>
</tr>
<tr>
<td>CS quality</td>
<td>1</td>
<td>17.70**</td>
<td>3.78</td>
</tr>
<tr>
<td>Treatment</td>
<td>1</td>
<td>0.32</td>
<td>0.01</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>43.69***</td>
<td>11.32*</td>
</tr>
<tr>
<td>Error (Mean Square)</td>
<td>12</td>
<td>(15.85)</td>
<td>(20.48)</td>
</tr>
</tbody>
</table>

* $p < 0.01$; ** $p < 0.005$; *** $p < 0.001$. 
2. Comparison of the level of performance on positive and on negative trials. Numbers of CRs performed during the CS-US interval on positive trials and numbers of disinhibited negative trials were considered as measures of the level of performance. The Fisher exact probability test (37) was used to answer whether or not proportions of responding and non-responding on positive and on negative trials were statistically different. This measure is not biased when the CS-US interval on positive trials equals the duration of the CS−, as it was in the present study. However, such a method is not sensitive in situations when high level of performance both on positive and negative trials is observed and/or when a small number of trials is applied in one experimental session.

It has been found that this method, when applied to data collected in this study, was less sensitive than the previous one for discovering persistence of responding to the old signalling properties of stimuli. Only in the Group 1N, which was retarded in emission of the first avoidance response to the tone CS'+, was there a significantly greater proportion of bar presses to the CS'− than to the CS'+ during the 1st and/or the 2nd day of the reversal learning.

The number of sessions in which there was a significantly greater proportion of bar presses to the CS'+ than to the CS'− was compared, using the 2 × 2 analysis of variance. It was found that only the interaction of the two main effects was significant at the $p < 0.01$ level (see Table III, Method B).

3. Comparison of the relative response latencies on positive and on negative trials. In the Russian literature three steps of reversal learning are distinguished (20): (i) initial reversal in which response strength to the CS'+ for the first time is greater than response strength to the CS'−; (ii) final reversal in which the differences in response strengths become permanent, and (iii) full or absolute (see 30) reversal in which response strength to the CS'+ is equal to or exceeds previous performance to the CS+, and response strength to the CS'− is equal to or smaller for the CS−. The methods of evaluation of the degree of reversal learning accomplished by a given subject described above are suitable for detection of the initial and final stages of reversal. In the present study no one animal attained the step of full (absolute) reversal since in each cat greater numbers of negative trials were disinhibited during Block 6 than during Block 0. However, there is reason to compare within each session of reversal learning the response latencies in positive and in negative trials using as the point of reference the median latency of responses performed by a given cat during positive trials of Block 0. This method exaggerates the influence of stimulus intensity dynamism on the course of reversal learning in comparison with the original differentiation train-
ing, and thus it is a very efficient way to determine such an influence. The proportions of trials in which response latency was greater (or smaller) than the point of reference was used as a measure because it eliminates the bias due to differences in theoretical length of latencies on positive and negative trials. The question of whether or not such proportions on positive and negative trials were different was answered by the Fisher exact probability test.

It was found that this method yielded very similar results to the method of comparison of absolute response latencies. The numbers of sessions in which cats persisted in responding to the old signalling properties of the stimuli, as indicated by comparison of absolute (Method A) and relative response latencies (Method C), was positively correlated \( (r_s = +0.65, p < 0.01, n = 16, \text{Spearman rank correlation coefficient}) \). The number of sessions in which they responded to the new signalling values of the conditioned stimuli showed an even stronger correlation, as seen from Table IV. The number of sessions in which the Fisher exact probability test discovered significantly shorter relative latencies on positive than on negative trials was tested statistically, and the \( 2 \times 2 \) analysis of variance showed a significant interaction of the two main effects (see Table III, Method C).

In fact, all three measures of the acquisition of the new discrimination proposed in this section of the paper showed similar results. As indicated in Table III, independently of the measure used, Group 1N showed the best scores and Group 2N, the worst. In each case the groups with lesions shared an intermediate position; however, better scores were observed in Group 1F than in Group 2F. A similar correspondence of the individual data was also found. The number of sessions in which significant differences in responding to the CS\(^+\) and to the CS\(^-\) in accordance to their new signalling properties were detected by the three methods were strongly and positively correlated (Table IV). However,
in spite of this correspondence of the individual data, only Method A showed a statistically significant effect of the CS quality on the acquisition of the new click vs. tone differentiation. The interaction of CS quality and treatment (lesion vs. control groups) effects was significant statistically in each case.

DISCUSSION

The data confirm results observed on the same animals at earlier stages of the experiment. In addition new data concerning the interaction of CS quality and prefrontal lesions were obtained. In the earlier paper effects of the quality of CSs on the avoidance reflex acquisition and differentiation learning had been demonstrated: more rapid learning in animals trained to the click CS+ than in those trained to the tone CS+; and smaller generalization of the CR to the tone CS− than to the click CS− (50). These results were obtained on normal cats before two experimental groups, 1F and 2F, received prefrontal lesions.

In the course of the reversal learning in normal cats similar effect of CS quality was observed. All cats learned very quickly to perform avoidance CRs on positive trials after reversal of the signalling properties of the conditioned stimuli; however, the normal cats trained to the tone CS'+ were retarded in emission of the first avoidance response. During all 30 sessions of the reversal learning the level of responding on positive trials in Group 1N trained to the tone CS'+ was lower than in Group 2N trained to the click CS'+. As far as responding on negative trials is concerned, the group differences were not statistically significant, but the number of bar-presses to the click CS'− in Group 1N was greater than to the tone CS'− in Group 2N. In effect, marked differences were observed between the two groups in rapidity of reversal learning. Most cats from Group 1N continued to respond according to the previous signalling properties of the stimuli, but those from Group 2N quickly changed their responses in accordance with the new signalling properties of the stimuli. Thus, in normal cats acquisition of the new go–no go differentiation was easier when the stronger stimulus from the pair was used as the CS'+ than when the weaker from the pair of stimuli was used as the CS'+. Results obtained on normal cats confirm the prediction that reversal learning will be more difficult in animals which were previously trained to differentiate the stronger CS+ and the weaker CS− than in animals which had been trained with the reversed intensity relations of stimuli. These differences in the course of original differentiation and in reversal learning may be predicted from the effect of stimulus intensity on generalization from the CS+ to the CS−. They were
demonstrated before only for the original differentiation training (3, 23, 24, 32, 35, 36, 43).

These considerations are not valid for cats with prefrontal lesions. Groups 1F and 2F showed similar levels of performance on positive trials and shared an intermediate position between the two groups of normal cats. There was even a tendency for prefrontal cats trained to tone, the weaker CS+ and click, the stronger CS− to show better response scores to the new signalling properties of the conditioned stimuli, than observed in those from the Group 2F, which were trained to differentiate the click CS+ and the tone CS−.

On the basis of the previous data showing attenuation of the stimulus intensity effect, indicated by changes in ability to perform short-latency avoidance responses after prefrontal lesions in cats (45, 47, 48), it was expected that differences in the course of the reversal learning between the two prefrontal groups would be smaller than between the groups of normal cats. Data presented in this paper confirm this prediction. Whereas in normal cats intensity relations between the two CSs exerted a definite influence on the course of the reversal learning, in prefrontal cats such differences in effectiveness of the stimuli had no effect either on extinction of the old habits or on acquisition of the new differentiation. This was indicated particularly by the interaction of stimulus quality and treatment effects, which was significant statistically in a number of analyses, showing the ability to change responding in accordance with the new signalling properties of conditioned stimuli. These data provide additional arguments in favor of the notion that behavior of prefrontal cats is not as strongly influenced by stimulus intensity dynamism as is the behavior of normal cats trained in the active avoidance reflex.

Group differences discussed above were almost exclusively due to differences in responding to the CS+. All measures of performance on negative trials yielded non-significant statistical differences between groups. Prefrontal cats were no worse than normal in their ability to inhibit bar-pressing to the CS−. Results presented in this paper confirm previous data regarding retention of preoperatively acquired avoidance reflex differentiation (50) in spite of the fact that the inhibitory abilities of the prefrontal cats were tested against a strong excitatory background. Similarly, in earlier experiments on cats it was found that the numerous shock applications necessary for retraining of the avoidance reflex after prefrontal lesion did not worsen retention of the difficult conditioned inhibition task (47, 51).

Perfect retention of the preoperatively acquired differentiation and conditioned inhibition tasks as well as lack of differences between the
normal and prefrontal cats in responding to the CS' during reversal learning indicate that the drive disinhibition hypothesis is not applicable to the active-avoidance bar-pressing situation. Analysis of data concerning the intertrial responding also confirms this conclusion.

In the course of learning in most situations CRs are performed not only during the CS presentations but also during intertrial intervals. Extinction of these intertrial responses (ITR) is considered as an effect of an active inhibitory process, which is involved in discrimination between the CS+ (i.e. positive CS and the experimental situation) and the experimental situation alone, which constitutes the negative stimulus (15, 26, 28, 33). Increase in the ITR rate, described in some studies concerning effects of prefrontal lesions on alimentary conditioned reflexes, was considered as one of the symptoms of the disinhibitory syndrome (6, 10, 12). In cats trained in the active bar-pressing avoidance reflex a clear increase of the ITR after prefrontal lesion was observed only in one experiment (48). In the present study the ITR rate was slightly higher in prefrontal animals prior to the reversal learning; however, this difference between normal cats and those with lesions was not significant statistically, as in most experiments in which the same method was used.

As a consequence of the high frequency of shock application on positive trials at the beginning of reversal learning, the animals became more excited and the ITR rate increased greatly, much more in normal cats than in those with lesions. This is in disagreement with the drive disinhibition hypothesis, which would predict a greater increase of the ITR rate in animal with prefrontal lesions.

However, the differences in the ITR rate observed at the beginning of the reversal learning do not necessarily indicate that fear drive conditioned to the experimental situation was lower in prefrontal animals than in normal cats. It was shown before that prefrontal lesions in cats selectively decrease the ability to perform short-latency avoidance responses, which seems to indicate deterioration of initiation and not of execution of the avoidance response (45, 47, 48). In light of the notion that measures of performance on positive and negative trials during the go–no go differentiation are not fully orthogonal, one would expect that after prefrontal lesions initiation of bar-pressing on negative trials and of intertrial responses would also be more difficult than in normal cats. This statement was supported also by the fact that the increase in proportion of short-latency responses to the CS'+, which will be discussed later, was smaller in prefrontal than in normal cats.

Although prefrontal cats emitted fewer ITR in the beginning of reversal learning, they showed a greater difficulty than normal cats in extinguishing intertrial responding in the course of further training. It
seems that in active bar-pressing avoidance situation prefrontal cats, which were not worse in comparison with normal cats in various inhibitory tasks, have difficulty improving performance with further training. It should be recalled that the only effect of prefrontal lesions on learning the original differentiation was lack of further improvement of performance on negative trials similar to that observed in normal cats after the 90% criterion was reached (50).

The last problem which has to be discussed is the relation between performance on negative trials and responding during intertrial intervals. At the beginning of reversal learning two opposite processes were observed with respect to responding on negative trials: a decrease of the over-all level of performance to the previously positive CS and an increase in proportion of short-latency bar-pressing responses executed in the presence of the same CS. This increase in proportion of short-latency responses to the CS' was correlated positively with the increase in the ITR rate, but the over-all level of performance on negative trials did not show such a correlation. Decrease of performance of long-latency responses to the CS' observed even at the beginning of reversal learning reflected the changed signalling value of this CS; whereas increase in the probability of short-latency responses was in opposition to the new signalling properties of the conditioned stimuli. This discrepancy indicates that different mechanisms were responsible for execution of responses with different latencies. It seems that short-latency responses were related to the level of the general arousal or of the fear drive exhibited to the experimental situation and indicated by numerous ITR, but performance of the long-latency responses more directly reflected the signalling value of a given CS.

In a recent study on differentiation learning in the classically conditioned defensive reflex situation it was found that the onset of the CS, independently of whether positive or inhibitory, produces a rapid increase in heart rate. Clear differences between positive and negative trials were observed only in the second part of the CS-US interval: high heart rate to the CS+ and low rate to inhibitory stimuli (25). Kakigi considered these results as indicating that the heart rate response with short latency was based mainly upon nonconditional factors, i.e., sensitization and/or orienting reflex and that the response observed in the second part of the CS-US interval was based on a conditional factor, or, speaking differently, reflected the signalling value of a given CS. It ought to be added that in these experiments increase of the heart rate to CS onset developed in the course of training, as the subjects were habituated to these stimuli and no changes in the heart rate were observed prior to the differentiation learning.
It seems that this parallelism in the theoretical account of data collected in the present study and in that by Kakigi is not accidental. More and more experimental results indicate that different mechanisms are responsible for execution of CRs with different latencies. As far as the avoidance bar-pressing situation is concerned, the probability of performance of short-latency responses in cats increases with intensity of the acoustic CS (46–48), decreases as the difference between CS and background intensity levels decreases (49), markedly decreases after prefrontal lesions and does not change in postoperative training (45, 47, 48), and in some situations is positively correlated with the ITR rate. This relation of the short-latency responses to the ITR rate was shown so far only on negative trials in those stages of learning when the CS− has partially excitatory properties, as at the beginning of reversal learning or differentiation training. During differentiation the rate of responding to the click CS− and during intertrial intervals are positively correlated, although similar correlation in cats trained to the tone CS− are not statistically significant (50). Responses to the click CS− were performed mainly with short latencies; whereas to the tone CS− they were uniformly distributed during the time of its presentation.

On the basis of behavioral and neurophysiological data it was hypothesized before that fear associated with a given CS (and not with the whole experimental situation) is more involved in execution of long-latency than in short-latency avoidance responses (47). The present results are consonant with this hypothesis; however, the data showing a correlation between the short-latency responses executed on negative trials and the ITR rate call for more careful investigation of relations between motivational aspects of the sporadic conditioned stimuli and of the experimental situation, or, more precisely, the stimulus situation existing during the intertrial intervals. Such studies would be especially important in light of the “elicitation theory” (see 18), which assumes that in the active avoidance reflex situation short-latency “relief” and long-latency “relaxation” responses occur during intertrial intervals and that development of these responses strongly influences both acquisition and extinction of the avoidance response.

This investigation was supported by Project 09.4.1 of the Polish Academy of Sciences and by Foreign Research Agreement 05.275.2 of the U.S. Department of Health, Education and Welfare under PL 480.

REFERENCES

15. CAPPOCK, H. and MOWRER O. H. 1947. Inter-trial responses as "rehearsal": a study of "overt thinking" in animals. Amer. J. Psychol. 60: 608-616.
21. FEDOROV, V. K. 1962. Some results of investigation of the typological cha-
racteristics of animals higher nervous activity (in Russian). Dokl. AN SSSR. 142: 1432-1435.

Received 8 January 1973

Kazimierz ZIELIŃSKI and Julitta CZARKOWSKA, Department of Neurophysiology, Nencki Institute of Experimental Biology, 3 Pasteur, 00-973 Warszawa, Poland.