I. Introduction

In this paper I wish to propose a modification to Eysenck's (1957, 1967) well-known theory of introversion-extraversion. This theory is concerned both with the psychological and with the physiological nature of introversion, and the modification proposed here has a similarly dual aspect. In brief, it consists of the following two alterations to Eysenck's theory: (1) It replaces the Ascending Reticular Activating System (ARAS: Moruzzi and Magoun, 1949; Magoun, 1963), regarded by Eysenck as the physiological substrate of introversion, with a more extensive system, consisting of the ARAS together with the medial septal area, the hippocampus, and the orbital frontal cortex, and the interconnections between these structures, to perform the same theoretical function. (2) It replaces the notion of "conditionability", thought by
Eysenck to be greater in the introvert than the extravert, with the notion of "susceptibility to punishment", which, I shall suggest, is related in a similarly positive manner to degree of introversion.

Just as the proposed modification to Eysenck's theory has both a psychological and a physiological aspect, so there are both psychological and physiological considerations which argue for the necessity of the modification. The physiological considerations will be dealt with first.

Consider the significance of two of the major experimental facts known about the physiological basis of introversion: (1) that the barbiturate drugs (in particular, sodium amobarbital) and alcohol have an extraverting effect on behaviour (Eysenck, 1967), and (2) that lesions to the frontal cortex have a similar effect (Willett, 1960). Consider in particular what is known about the effects of these treatments on the behaviour of animals.

II. The Behavioural Effects of Sodium Amobarbital

It is clear from a large number of experimental studies that the behavioural effects of both amobarbital and alcohol are specific with respect to the reinforcement contingencies governing the animal's behaviour. The effects of amobarbital have been studied in more detail, so we shall confine the discussion to this drug, although it should be added that, where data on alcohol are available, its effects on behaviour do not appear to differ in any important respects from those of the barbiturates. The extensive work of Miller (1959) has shown that small doses of amobarbital (typically 15-20 mg/kg intraperitoneally in the rat) are able to reduce the behavioural effects of punishment (i.e., passive avoidance) in approach–avoidance conflict situations, while not altering the behavioural effects of reward. In more recent years it has become clear, both from Miller's (1964) experiments and from those of other workers (Gray, 1967a; Ison, 1968; Wagner, 1966) that this drug also reduces the behavioural effects of frustrative nonreward (Amsel, 1962). Thus extinction is retarded by injections of amobarbital (Barry, Wagner and Miller, 1962), the partial reinforcement acquisition and extinction effects (Lewis, 1960) are attenuated or blocked completely (Wagner, 1963; Stretch, Houston and Jenkins, 1964; Gray, 1969; Ison and Pennes, 1969), discrimination learning is impaired owing to increased responding to the negative cue (Ison and Rosen, 1967), and the Crespi depression effect is reduced (Rosen, Glass and Ison, 1967). In contrast to these important effects of amobarbital on passive avoidance and on behaviour resulting from frustrative nonreward, simple approach learning (to a reward) is unimpaired by the drug and active avoidance learning, as in
the Miller-Mowrer shuttle-box, may even be enhanced by it (Kamano, Martin and Powell, 1966).

The nature of active avoidance learning requires some further discussion. It has been proposed elsewhere (Gray, 1971a, in press, b) that

![Block diagram of the "arousal-decision model" proposed by Gray and Smith (1969) for conflict situations.](image)

**Fig. 1.** Block diagram of the "arousal-decision model" proposed by Gray and Smith (1969) for conflict situations. R₁ and P₁: inputs to the reward and punishment mechanisms, Rew and Pun. D. M.: the decision mechanism. A: the arousal mechanism. B. Com.: behaviour command to "approach", on the reward side, or to "stop" (passively avoid), on the punishment side. Beh.: the observed motor behaviour. B. Cons.: the consequences (rewarding or punishing) of the behaviour that occurs. Comp.: comparator mechanisms which compare the actual consequences of behaviour with the expected consequences and make appropriate reward or punishment inputs. If the actual reward is less than the expected reward, there is an input to the punishment mechanism (the "fear = frustration hypothesis"). Conversely, if the actual punishment is less than the expected punishment, there is an input to the reward mechanism (the "hope = relief hypothesis"). For further details, see Gray and Smith (1969).

this kind of learning is in fact reinforced by activity in the same positively reinforcing brain system which is involved in ordinary reward learning, a view which has also been expressed by Olds and Olds (1964). Furthermore, the two-way active avoidance involved in the shuttle-box is in fact a conflict between approach to the now-safe side of the apparatus and
III. THE PHYSIOLOGICAL SITE OF ACTION OF AMOBARBITAL

The behavioural effects of small doses of sodium amobarbital may be summarized then by the statement that the animal's sensitivity to both punishment and frustrative nonreward is reduced by this drug, whereas its sensitivity to reward or to relieving nonpunishment is unimpaired. Now the assumption is commonly found in the psychopharmacological literature (Killam, 1962) that barbiturates act primarily by depressing activity in the ARAS. No doubt this is correct in so far as the sedative and anaesthetic properties of these drugs are concerned. However, after the administration of 20 mg/kg sodium amobarbital, a rat is not normally sedated to any marked degree. It is capable of running an alley in a fast and well-motivated manner to a food or water reward and, as in Kamano, Martin and Powell's (1966) experiment, it may even out-perform control animals in a shuttle-box. In the work on human personality summarized by Eysenck (1967), the extraverting effects of the barbiturates and of alcohol which have interested experimental psychologists have usually been obtained with doses well short of those needed for sedation or anaesthesia. Furthermore, it seems implausible that the ARAS, usually regarded as the controlling system for general level of arousal or alertness (Samuels, 1959; Gray, 1964a), should be the chief site of action of a drug whose effects are specific to the type of reinforcement (punishment or passive avoidance of this same side, since the animal has just been shocked there. Thus it is to be expected that amobarbital will enhance active avoidance in the shuttle-box (as found by Kamano, Martin and Powell, 1966), in exactly the same way as it increases the likelihood that an animal will take a reward in spite of punishment in an ordinary approach-avoidance conflict (Miller, 1959).

These various facts concerning the action of amobarbital may all be understood in terms of a model for conflict situations proposed by Gray and Smith (1969), presented in Fig. 1. As it can be seen, this model incorporates the "fear = frustration hypothesis" (Gray, 1967a), according to which both punishment and frustrative nonreward act on the same physiological system, here labelled the "punishment mechanism". It also incorporates the "hope = relief" hypothesis (Gray, 1971a, in press, b), according to which reward and the omission of anticipated punishment ("relieving nonpunishment") both act by way of the "reward mechanism". The simplest assumption concerning the effects of amobarbital is then that this drug antagonizes the activity of the punishment mechanism, while leaving the reward mechanism unaffected.
reward) involved in the situation. It would therefore seem worth considering other possible sites of action for the behaviourally effective doses of amobarbital which have occupied our attention above.

I have proposed elsewhere (Gray, 1970a) that an alternative site of action is to be found in the hippocampus and in the cluster of cells in the medial septal area which acts as the pace-maker for the hippocampal theta rhythm (Stumpf, 1965). Several lines of evidence argue in favour of this suggestion. In the first place, there is a remarkable similarity between the effects of lesions to the septal area, lesions to the hippocampus and injections of amobarbital. Like amobarbital, lesions to both areas impair passive avoidance and extinction of once-rewarded behaviour (Douglas, 1967; McCleary, 1966), but enhance two-way active avoidance (Isaacson, Douglas and Moore, 1961; Olton and Isaacson, 1968; Rabe and Haddad, 1969; Green, Beatty and Schwartzbaum, 1967; McCleary, 1966). More direct evidence of antagonism between the barbiturate drugs and these portions of the limbic system is not lacking. Harvey et al. (1964), in an extensive investigation of the effects of brain lesions on barbiturate-induced sleep, found that only lesions in the septal area, the hippocampus, and the dorsomedial tegmentum (which has extensive interconnections with the septal area and the hippocampus: Nauta, 1958) affected this, both increasing the speed of induction of sleep and prolonging its duration. Furthermore, in a direct investigation of single neuron activity in the ARAS, Adey, Segundo and Livingston (1957) and Livingston (1959) reported that there is a hippocampal inhibition of upstream conduction from the midbrain to the thalamic portion of the ARAS, and that this inhibition is reduced by sodium pentobarbitone in doses which are too small to act directly on the ARAS.

Given that data exist which make it plausible that low doses of the barbiturate drugs act on the hippocampus and the septal area, one is naturally inclined to seek for a single mechanism which could account for the action of these drugs on both these parts of the limbic system. Such a mechanism is not difficult to find. It is known that the medial septal area (specifically, the medial septal nucleus and the nucleus of the diagonal band of Broca) contains the pacemaker cells for the hippocampal theta rhythm (Stumpf, 1965). It is possible, therefore, that the communality between the effects of small doses of barbiturates and those of both septal and hippocampal lesions arises because these drugs depress the septal pacemaker for the hippocampal theta rhythm.

This possibility is strengthened by the fact that recent data from experiments using small lesions within the septal area suggest that it is lesions in the medial septal area, where the pacemaker cells are located, which are able to impair extinction of rewarded behaviour (Butters and
IV. THE BEHAVIOURAL EFFECTS OF DRIVING AND BLOCKING THE THETA RHYTHM

The writer (in collaboration with Dr. G. G. Ball) has directly tested the hypothesis that the behavioural effects of small doses of sodium amobarbital are due to an action on the septal mechanism for production of the hippocampal theta rhythm (Gray, 1970a). These experiments have been conducted with free-moving rats chronically implanted with bipolar recording electrodes in the dorsal hippocampus and with bipolar stimulating electrodes in the medial septal area. As reported by Stumpf (1965) and his co-workers for the curarized rabbit, it is possible in the free-moving rat to drive the hippocampal theta rhythm by low frequency (6–10 Hz) septal stimulation, to disrupt the theta rhythm temporarily by high-frequency (200 Hz) stimulation of the same region, or to disrupt permanently the theta rhythm by electrolytic lesions of the medial septal area, a finding also reported by Donovick (1968). We have also observed that the threshold for septal driving of the hippocampal theta rhythm is raised very considerably after an intraperitoneal injection of 20 mg/kg sodium amobarbital, the dose which is commonly employed in the behavioural experiments discussed earlier in this paper. Furthermore, the rise in the theta-driving threshold is particularly great at precisely that frequency (7–8 Hz) which is spontaneously apparent in the rat’s hippocampal record when the animal is exposed to frustrative nonreward (nondelivery of a water reward for which the animal has previously been trained to run down an alley). The effects of 20 mg/kg sodium amobarbital on the theta-driving threshold at 6 Hz (displayed spontaneously in the hippocampal record when the animal is consuming the reward) and at 9–10 Hz (seen when the animal is running down the
alley towards the reward) are very much smaller (Gray and Ball, 1970). Thus, these results fit very well with the hypothesis that the behavioural effects of sodium amobarbital (impairment of extinction and other behaviour resulting from frustrative nonreward, but no alteration in rewarded behaviour) arise from a depression of the septal pacemaker for the hippocampal theta rhythm.

Further support for this hypothesis was obtained in a series of experiments (Gray, in press, a) in which we investigated the effects of both theta-driving and theta-blocking (by electrical stimulation of the medial septal area at 7-7 and 200 Hz respectively) on behaviour in a learning situation. If amobarbital works by blocking the theta rhythm in the 7-8 Hz frequency band, then artificially producing a theta rhythm in this band should have effects opposite to those produced by injections of the drug. In accordance with this deduction it was found (a) that theta-driving applied to the animal during extinction increased the rate of extinction; and (b) theta-driving applied in the goalbox on a random 50% of rewarded trials (water being available on every trial) during acquisition retarded subsequent extinction. These effects may be regarded as the opposites, respectively, of the reduced rate of extinction caused by injections of amobarbital during extinction (Barry, Wagner and Miller, 1962; Miller, 1964; Gray, 1969) and of the blocking of the partial reinforcement extinction effect produced by injections of amobarbital during acquisition (Gray, 1969; Ison and Pennes, 1969).

Conversely, blocking the theta rhythm by means of high-frequency stimulation of the septal area should act like an injection of sodium amobarbital. In agreement with this hypothesis it was found (Gray, 1970a) that the application of theta-blocking stimulation in the goalbox on the nonrewarded trials of a partial reinforcement schedule (water delivered for running the alley on a random 50% of trials) reduced resistance to extinction compared to control animals given the partial reinforcement schedule without stimulation. In a further experiment in this series (Gray, Quintão and Araujo-Silva, in press), rats were prepared with lesions of the medial septal area which permanently disrupted the hippocampal theta rhythm. The lesions were carried out either before animals were trained on continuous or partial reinforcement schedules, or after training was over but before extinction. When the speeds in the goal section of the runway were examined, controls, subjected to sham-operations, displayed the usual partial reinforcement extinction effect (Lewis, 1960), but this effect was reduced in both sets of lesioned animals. Furthermore, the reduction in the partial reinforcement extinction effect could be attributed about equally to a rise in the running speed during extinction of animals trained on continuous reinforcement,
as reported by McCleary (1966), and to a fall in extinction running speed in rats trained on partial reinforcement.

In summary, then, these experiments offer evidence that: (1) theta-driving during extinction speeds up extinction; (2) theta-driving on a random 50% of rewarded trials creates a "pseudo partial reinforcement extinction effect", that is, an increase in subsequent resistance to extinction; (3) theta-blocking by electrical stimulation of the septal area on the nonrewarded trials of a partial reinforcement schedule reduces resistance to extinction relative to normal controls trained on partial reinforcement; and (4) lesions to the medial septal area which disrupt the theta rhythm markedly attenuate the partial reinforcement extinction effect. All these results are readily understood on the hypothesis that amobarbital affects behaviour by antagonizing the normal theta-rhythm response to frustrative nonreward.

V. The Frontal Cortex

The similarities between the effects of injections of amobarbital, lesions to the medial septal area, and lesions to the hippocampus are repeated once again, at a higher level, when we consider the effects of lesions to the frontal cortex. These too have been shown to impair passive avoidance of punishment and extinction of once-rewarded behaviour (see Grossman, 1967, for review) and to enhance two-way active avoidance (Albert and Bignami, 1968). Furthermore, in a specific comparison between the effects of septal and frontal cortex lesions in the rhesus monkey, Butters and Rosvold (1968) showed that lesions to the medial septal area had similar effects to those resulting from lesions to the orbital frontal cortex (Butter, Mishkin and Rosvold, 1963; Butter, 1969), from which the medial septal nuclei receive an important projection (Johnson, Rosvold and Mishkin, 1968). Thus it seems that in primates (no doubt including Man) the cortical representation of the septo-hippocampal system whose functions we have been delineating lies in the orbital frontal cortex.

Now this conclusion, of course, takes on particular significance in the light of the facts with which we began this discussion: namely, that both amobarbital (and alcohol) and lesions to the frontal cortex have an extraverting effect on behaviour in Man (Eysenck, 1967; Willett, 1960). If the orbital frontal cortex is the highest level of the septo-hippocampal system whose functioning is indicated by the occurrence of a hippocampal theta rhythm, and if the behavioural effects of amobarbital are due to an antagonistic action on this system, then this similarity
between frontal cortical lesions and amobarbital administration in Man is exactly what we would expect.

It seems, then, that, by considering as we have the physiological locus of action of the barbiturates, we have also arrived at an understanding of the communality of effects on human personality between frontal lesions and administration of these drugs. But from these two important conclusions flow: (1) the physiological basis of introversion–extraversion must consist of that system whose functions are depressed by the barbiturates in small, behaviourally effective doses; (2) the psychological nature of introversion–extraversion must involve individual differences in the behavioural functions exercised by this same physiological system. To each of these points we now turn.

VI. THE PHYSIOLOGICAL BASIS OF INTROVERSION

The proposal which is being put forward, then, is that *it is activity in this frontal cortex-medial septal area-hippocampal system which determines the degree of introversion*: the more sensitive or active this system is, the more introverted will the individual be. Now the equivalent assumption made by Eysenck (e.g., 1967) in his most recent statements of his theory of introversion–extraversion is that degree of introversion is determined by level of activity in the ARAS. But in fact these two hypotheses do not differ as much as first appears. Indeed, I hope to show that the system I have delineated, *in conjunction with the ARAS*, is able to perform exactly those functions which Eysenck, on psychological grounds, has always supposed to lie at the root of individual differences along the introversion–extraversion dimension, even before he made any attempt to locate these functions in a particular physiological structure.

In outlining the role played by the septal area and the hippocampus in the exercise of those behavioural functions which appear to be antagonized by sodium amobarbital, I have stressed that these functions appear to be exercised at times when the hippocampus, as a result of neural input from the medial septal area, is displaying a theta rhythm. But there is good evidence (Stumpf, 1965) that the pacemaker cells in the medial septal area act to produce a theta rhythm as a result of neural input received from the midbrain reticular formation. As the voltage of a stimulating current applied to the midbrain reticular formation is gradually increased, a theta rhythm first appears in the hippocampus and then increases in frequency, until, at some limiting stimulus voltage, the theta rhythm is replaced by a fast, low-voltage, desynchronized hippocampal EEG (Stumpf, 1965). As well as this reticulo-septal-
hippocampal circuit, there is a hippocampal-recticular formation connection, one which is evidently of an inhibitory nature. As shown by Adey, Segundo and Livingston (1957) and Livingston (1959), stimulation of the hippocampus causes an inhibition of upstream conduction from the midbrain reticular formation to the thalamic portion of the ARAS; and I have already commented on the fact that this inhibition is disrupted by small doses of barbiturates, indicating the likelihood that it is normally exercised at times when the hippocampus displays a theta rhythm. Thus there is every reason to suppose that the ARAS and the septo-hippocampal system whose operation is indicated by a theta rhythm are coupled together to form a negative feedback loop. That is to say, increased activity in the ARAS results in an increase in hippocampal activity, and one result of this increased hippocampal activity is to inhibit any further increase in the activity of the ARAS. In this way, for a given sensory input and under given conditions, a temporary equilibrium between reticular activation and hippocampal inhibition is reached.

At the same time, the evidence suggests that the effects of hippocampal inhibition are not confined to the ARAS.

As we have seen, lesions to the medial septal area, which controls the theta rhythm, and to the hippocampus itself, as well as to the orbital frontal cortex, impair passive avoidance and extinction of appetitive behaviour. We have argued above that the most parsimonious interpretation of these facts is that these kinds of behavioural inhibition are carried out by the septo-hippocampal system at a time when the hippocampus is displaying a theta rhythm—i.e., at a time when there is an important input from the ARAS to the medial septal area.

In addition to what might be called the "act" inhibition which is involved in passive avoidance and extinction of appetitive behaviour, there is good evidence that the hippocampus and the septal area exert an inhibitory influence on sensory input. Feldman (1962) demonstrated that electrical stimulation of the hippocampus inhibits the potentials evoked in the hypothalamus by stimulation of the sciatic nerve, and Redding (1967) showed a similarly inhibitory effect of hippocampal stimulation on both visual and auditory cortical evoked potentials. This hippocampal inhibition of sensory input seems to be exerted both in the midbrain reticular formation (Redding, 1967) and in the septal area (Feldman, 1962). Inhibitory effects from stimulation of the septal area itself have been demonstrated by Lorens and Brown (1967) using evoked potentials in the visual cortex. The results of experiments in which the septal area has been lesioned also support this view of septal function. A chief symptom of the "hyper-emotionality" syndrome caused by such lesions (see Grossman, 1967, for review) is an increased reactivity to tactile
stimuli; and Lints (cited by Lorens and Brown, 1967) has reported a lowered threshold to electric shock in septal-lesioned rats. Other experiments have shown that septal-lesioned rats display increased reactivity to light (Schwartzbaum et al., 1967) and to gustatory stimuli (Beatty and Schwartzbaum, 1967).

We see, then, that an increase in the activity of the ARAS is likely to increase the probability that the septo-hippocampal system described will initiate inhibitory influences of three kinds: (1) on the ARAS itself; (2) on behavioural acts leading to punishment or frustrative nonreward; and (3) on sensory inputs.* Thus, the hypothesis put forward here that activity in the septo-hippocampal system underlies degree of introversion and Eysenck's (1967) hypothesis that it is activity in the ARAS which performs this function are not rivals, but rather complementary to each other. We may therefore unite them in a more complete statement of the view proposed in this paper: the level of introversion is determined by the degree of activity in a negative feedback loop consisting of the ARAS together with the orbital frontal cortex, the medial septal area and the hippocampus (Fig. 2).

In discussing elsewhere (Gray, 1967b) the simple arousal hypothesis for the nature of the dimension of introversion-extraversion, I have pointed out that this hypothesis, if correct, would raise the possibility that introversion and the dimension of personality studied in the Soviet Union under the name "strength of the nervous system" (Teplov, 1956-67; Nebylitsyn, 1966; Gray, 1964b) are in fact identical, for there is a considerable amount of support for the view that this dimension too is one of arousability (Gray, 1964a). The rather more complex set of hypotheses for the physiological basis of introversion advanced in the present paper still leaves this hypothesis open. Indeed, it may be the case that the phenomenon of transmarginal inhibition (response decrement at high stimulus intensities), which is of central importance in the theory of strength of the nervous system (Gray, 1964b), is due precisely to the operation of the kind of feedback loop discussed here as the physiological basis of introversion. In discussing this phenomenon in an earlier paper (Gray, 1964a), I suggested that the cortico-reticular negative feedback loop described by Dell, Bonvallet and Hugelin (1961) might constitute the physiological mechanism involved. It is possible that the system made

* It seems likely that the inhibition of sensory input which results from septo-hippocampal activity is functionally part of a selective attention mechanism. It is in agreement with this view that impairments in selective attention have been shown to arise as a result of hippocampal lesions (Hendrickson, Kimble, and Kimble, 1969) and amobarbital injections (McGonigle, McFarland and Collier, 1967).
up of the orbital frontal cortex, the medial septal area and the hippocampus forms the descending limb of this feedback loop. At any rate, the exciting possibility that the Western work on introversion and the Soviet

![Diagram](image)

**FIG. 2.** Negative feedback loop whose activity is presumed to underlie the dimension of introversion–extraversion. See text for further explanation.

work on strength of the nervous system have in fact been concerned with the same fundamental aspect of personality is still very much with us and deserves the closest attention.

**VII. THE PSYCHOLOGICAL NATURE OF INTROVERSION**

It will be obvious to anyone familiar with Eysenck’s theoretical writings that the general architecture of the set of interactions between the ARAS and the frontal cortex-septo-hippocampal system described above is remarkably similar to the psychological relationship which he has postulated (e.g., Eysenck, 1957) to lie at the root of the dimension of introversion–extraversion. He has consistently presented a picture of increased cortical arousal in the introvert leading to increased behavioural inhibition, though he has sometimes used the vocabulary of “arousal”
or "excitation" (higher in the introvert) and sometimes that of "cortical inhibition" (higher in the extravert). If we now substitute for "cortical arousal", the degree of activity in the feedback loop just described, we see that there is indeed a considerable amount of evidence that activity in this system leads to the kinds of inhibitory effects with which Eysenck and his group have been concerned.

Even though the ARAS and the septo-hippocampal inhibitory system function so closely together, and even though the view at which we have just arrived has such close similarities to Eysenck's basic theoretical position, it should be emphasized that this view does carry with it different predictions from those which may be derived from the simpler arousal/ARAS hypothesis as to the psychophysiological basis of introversion. These differences stand out clearly if we return to the barbiturate effects with which this paper began. On a simple arousal view of introversion, we would expect that the relationship between this dimension of personality and performance in learning situations would be independent of the nature of the reinforcement used in the situation. But, as we have seen, the effects of amobarbital on behaviour are very closely dependent on reinforcement contingencies, as are the effects of lesions to the frontal cortex, the medial septal area, and the hippocampus. Thus, in the light of the view of introversion advanced here, we would expect that differences along this dimension of personality would also interact with the nature of the reinforcement involved in a particular situation. Indeed, since we may describe the effects of amobarbital as a reduction in sensitivity to punishment and frustrative nonreward (see above), it follows that we may regard the dimension of introversion—extraversion as a dimension of susceptibility to punishment and nonreward: the greater the degree of introversion, the greater is this susceptibility.

Such a view has a great deal of face validity. Psychopathic behaviour in the extraverted neurotic is easily regarded as a tendency to take a reward (by, say, stealing, lying or sexual gratification) without thought for the consequences, i.e., with no fear of punishment. The recidivism which is such a feature of psychopathic behaviour (Eysenck, 1964) is also most simply seen as a relative insensitivity to punishment. Conversely, the symptoms of the introverted, or "dysthymic" (Eysenck, 1957), neuroses are in many cases perfectly clear expressions of fear, i.e., of sensitivity to punishment and the threat of punishment. This is obvious enough in the phobias or the anxiety state. It may easily be observed in obsessive-compulsive neuroses, when any attempt to prevent the patient from carrying out the obsessional ritual—which is best understood as a form of active avoidance response (Gray, 1971a)—produces frank signs of fear. Indeed, the only syndrome among the dysthymic neuroses requiring
more complex analysis in the light of the hypothesis advanced here is the neurotic or "reactive" depression (Mayer-Gross et al., 1960; Eysenck, 1960). However, one of the main strengths of the present position is that it can make very good sense out of the fact that individuals suffering from reactive depression have personality profiles which are identical to those of individuals suffering from the other major dysthymic syndromes (Cattell and Scheier, 1961). This strength arises from the "fear = frustration" hypothesis (Gray, 1967a; 1971a), according to which the behavioural effects of punishment and of the omission of expected reward are functionally and physiologically very similar, and perhaps identical. This hypothesis, the evidence for which is rather strong (Gray, 1967a; 1971a; Wagner, 1966; Miller, 1964), is incorporated into Gray and Smith's (1969) mathematical model for conflict behaviour, shown in Fig. 1. Now Mayer-Gross et al. (1960) note that reactive depressions usually result from "a sudden critical change, the death of a wife, the breaking of a love-affair, and are much less frequently precipitated by a gradual accumulation of miseries". These conditions are naturally described as the removal of accustomed sources of reward. Thus, taken together, the fear = frustration hypothesis and the hypothesis that introverts are relatively highly susceptible to punishment would predict that introverts will suffer more than extraverts from reactive depression; as is indeed the case (Eysenck, 1960; Cattell and Scheier, 1961).

VIII. The Role of Socialization in the Development of Introversion

I have discussed elsewhere some of the other reasons which may be used to support the adoption of the hypothesis that introverts are relatively highly susceptible to punishment (Gray, 1970b). Chief among them is the fact that this hypothesis can help resolve an important difficulty in Eysenck's approach to the process of socialization. This process is of central importance in his general theory of introversion-extraversion.

It is easier to understand Eysenck's theory of introversion if we first consider the problems to which this theory is an attempted solution. The starting-point for his theory lies in the relation of the dimensions of introversion and neuroticism to each other, as well as to a cluster of psychiatric disorders called by Eysenck (1957) "dysthymic" and to a second cluster of offences against society called "psychopathic". These relations are set out in Fig. 3. The problems they pose for the student of personality are these: (1) What are the psychological and/or physiological variables which result in the high susceptibility of individuals with high
degrees of neuroticism both to dysthymic disorders and to the commission of psychopathic offences? (2) What are the psychological and/or physiological variables which differentiate neurotic individuals along the dimension of introversion-extraversion (which is quite independent of neuroticism) and which cause the introverted neurotic to be susceptible to the dysthymic disorders, but the extraverted neurotic to display antisocial behaviour of a psychopathic kind? Although there have been many changes in detail, the broad lines of Eysenck’s solutions to these problems have remained fairly constant over the years; they are set out in a highly simplified manner in Fig. 4.
In brief commentary on Fig. 4, the following points must be made. (We shall ignore neuroticism in the present paper; the interested reader will find a brief treatment of this dimension in Gray, 1970b.)

(1) The first link in Eysenck's chain of explanation is to describe the dysthymic neuroses as disorders of over-socialization and psychopathic behaviour as a disorder of under-socialization (e.g., Eysenck and Rachman, 1965). In the introverted neurotic the establishment of the conscience (treated by Eysenck as a cluster of classically conditioned fear reactions) has proceeded so effectively that the individual concerned is disabled in adulthood by a variety of manifestations of his conditioned fears (in the form of phobias, obsessions and compulsions, frank anxiety states, reactive depression, etc.). In the extraverted neurotic, by contrast, there has been a relative failure of socialization, resulting in the lack of a sense of responsibility towards society and the various forms of anti-social behaviour (juvenile delinquency, sexual delinquency, lying, careless or drunken driving, or more serious breaches of the law) displayed by the psychopath.

(2) Introverts form the conditioned reflexes comprising the conscience with greater ease than extraverts because they are in general more highly conditionable.
(3) The greater conditionability of the introvert is attributed by Eysenck (e.g., 1967) to a relatively higher level of arousal at this pole of the introversion–extraversion dimension.

(4) The higher level of arousal in the introvert is due to a higher level of activity in the ARAS.

We have already suggested that this theoretical structure must be modified at point (4), in so far as the evidence suggests that the physiological basis of introversion must consist of a more complex system comprising the ARAS together with the inhibitory circuits present in the orbital frontal cortex, the medial septal area and the hippocampus. The second—psychological—modification is necessary at point (2).

As I have discussed elsewhere (Gray, 1970b), the evidence, mainly derived from experiments on eyeblink conditioning (Eysenck, 1965, 1967; Spence, 1964; Levey, 1967; Ominsky and Kimble, 1966; Eysenck and Levey, this volume), gives support to the following conclusions:

(1) Introverts are better than extraverts at eyeblink conditioning under some conditions, but not all;

(2) These conditions can be predicted on the hypothesis that introverts are relatively more highly aroused than extraverts;

(3) The superior conditioning of the introvert appears in experiments which involve at least some degree of threat.*

It is the first of these conclusions which has most importance for Eysenck's general theory of introversion. The idea that introverts are more highly conditionable than extraverts has to bear a heavy burden within this theory (Fig. 4). It is not at all clear that this burden can be carried by the weaker assumption (demanded by the experimental facts) that introverts condition better than extraverts only under some conditions. It appears that these conditions are, in general, relatively under-arousing ones (Eysenck, 1967; Gray, 1970b). But there is no reason to suppose that the parental conditioning techniques presumably involved in the formation of the conscience are more often under-arousing than over-arousing, so that there is no longer any reason to predict the over-socialization of the introvert which is critical to Eysenck's theoretical edifice.

Faced with this impasse, let us retreat one step. If we accept Eysenck's description of introvert behaviour as relatively over-socialized, and also his view that the process of socialization consists in the formation of a set of conditioned fear reactions, then we must agree that Eysenck has asked the right question: Why do introverts form conditioned fear

* Hobson (1969) has, however, recently reported results in conflict with this conclusion.
reactions more strongly than extraverts? We have just rejected the answer: "because they are better at conditioning". The only alternative answer is: "because they are more susceptible to fear". That this is the right answer is strongly suggested in any case by the fact that the introvert superiority in eyeblink conditioning depends on the degree of threat contained in the experimental situation (Ominsky and Kimble, 1966; Gray, 1970b). Furthermore, it is the same answer as the one we have already arrived at by considering the nature of the behavioural alterations produced by the barbiturates and by lesions to the frontal cortex: namely, the introvert is more susceptible to punishment.

IX. Susceptibility to Punishment and Level of Arousal

We see, then, that there are good grounds, of both a psychological and a physiological nature, for concluding that the psychological nature of introversion involves a relatively great susceptibility to punishment. Now, Eysenck derives his notion of greater conditionability in the introvert from the hypothesis that the introvert is relatively more highly aroused than the extravert (Fig. 4). There is, in fact, good evidence in support of the view that introverts and extraverts do differ in level of arousal* (Eysenck, 1967; Gray, 1967b); so that it would be in the interests of parsimony if we could now relate differences in susceptibility to punishment to differences in arousability. One way of doing this is to start from the fact that any stimulus, if it is sufficiently intense, may act as a punishment; and then to recall that differences in arousability, may be regarded as differences in the degree to which individuals amplify (at the high arousal pole of the continuum) or dampen (at the low arousal pole) stimulation (Gray, 1964a). It must follow that, as any physical stimulus is increased in intensity, the point at which it becomes punishing will be reached sooner, the more highly aroused the individual, i.e. (ex hypothesi) the more introverted he is (direct evidence for the introvert's tendency to avoid intense stimulation is reviewed by Eysenck, 1967). In this way, the greater susceptibility to punishment of the introvert may be derived from the same fundamental substrate of introversion–extraversion postulated by Eysenck: the introvert is more highly aroused than the extravert and is therefore more susceptible to punishment.

This conclusion is, in fact, psychologically isomorphic with the physiological conclusions we reached earlier, according to which the ARAS (arousal level) and the behaviour-inhibition system consisting of the frontal cortex, the medial septal area and the hippocampus (suscepti-

* See Corcoran, this volume.
bility to punishment) function as a single feedback loop: the higher the activity in the ARAS (the greater the arousal level), the stronger is the activity in the inhibitory part of this loop, i.e., the greater the inhibition of behaviour (susceptibility to punishment). Furthermore, although we have found it necessary to modify some of the details of Eysenck’s theory, on both psychological and physiological grounds, the resulting theoretical edifice (Fig. 5) not only retains the same basic structure as

<table>
<thead>
<tr>
<th>Level</th>
<th>Introverts</th>
<th>Extraverts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Socio-Psychiatric</td>
<td>dysthymic disorders</td>
<td>psychopathic behaviour</td>
</tr>
<tr>
<td>2. Socialization</td>
<td>good</td>
<td>poor</td>
</tr>
<tr>
<td>3. Conditioning of fear</td>
<td>good</td>
<td>poor</td>
</tr>
<tr>
<td>4. Susceptibility to punishment</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>5. Arousalability</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>6. Physiological</td>
<td>Feedback loop comprising ARAS, frontal cortex, septal area and hippocampus</td>
<td>high activity</td>
</tr>
</tbody>
</table>

Fig. 5. Proposed modification of Eysenck’s theory of introversion–extraversion. Compare with Fig. 4.

Eysenck’s (Fig. 4), but is in very close conformity with his basic notion of high arousal in the central nervous system giving rise to inhibited behaviour (e.g., Eysenck, 1957).

X. SUMMARY

Two areas of research with critical implications for Eysenck’s theory of introversion–extraversion have been reviewed. These concern: (1) the physiological locus of action of small, but behaviourally effective, doses of sodium amobarbital, and (2) the nature of the psychological differences between introverts and extraverts in the conditioning process.
These sets of data suggest the need for modifications to Eysenck's theory both at a physiological level and at a psychological level. The proposed modifications are: (1) that the physiological system underlying degree of introversion consists, not of the Ascending Reticular Activating System alone, but of this in conjunction with an inhibitory system consisting of the orbital frontal cortex, the medial septal area, and the hippocampus; and (2) that there are differences between introverts and extraverts, not in conditionability in general, but in susceptibility to punishment and frustrative nonreward, the introvert having higher susceptibility. It is shown that these modifications are isomorphic to each other; and also that the modified theory of introversion proposed has the same overall structure as Eysenck's original theory, although there are important differences in the detailed predictions which can be derived.

ACKNOWLEDGEMENTS

The experimental work of my own which is discussed in this paper was supported by grants from the Medical Research Council of Great Britain and from the Foundations' Fund for Research in Psychiatry, and also by a grant to Dr. N. E. Miller from the National Institute of Mental Health. Much of it was done while I held an M.R.C. Travelling Fellowship in Dr. N. E. Miller's laboratory at The Rockefeller University, New York. I am grateful to Dr. Miller for his constant encouragement, advice, and support. Several of the experiments on the hippocampal theta rhythm were done in collaboration with Dr. G. G. Ball, to whom I am particularly grateful, both for guidance in the mysteries of the brain and for many hours of fruitful exchange of ideas.

REFERENCES


