Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety

Neil McNaughton\textsuperscript{a,}\textsuperscript{*}, Jeffrey A. Gray\textsuperscript{b}

\textsuperscript{a}Department of Psychology and Centre for Neuroscience, University of Otago, P.O. Box 56, Dunedin, New Zealand
\textsuperscript{b}Department of Psychology, Institute of Psychiatry, Maudsley Hospital, London, UK

Abstract

According to “The Neuropsychology of Anxiety” [Gray, J.A., 1982. The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-hippocampal System, Oxford University Press, Oxford; Gray, J.A., McNaughton, N., 2000, The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-hippocampal System, 2nd ed., Oxford University Press, Oxford], anxiolytic drugs of all types act on a behavioural inhibition system, the most important neural component of which is the septo-hippocampal system. Anxiolytics affect septo-hippocampal function by impairing the subcortical control of hippocampal ‘theta’ activity — the principle response of the septo-hippocampal system to arousal. Our recent experiments show that there are multiple systems controlling theta activity and that anxiolytics act on several, but not all, of these systems. This pattern of results implies that there are many different types of arousal, only some of which appear to contribute to the generation of anxiety in normal subjects and to the etiology of pathological anxiety. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Anxiety; Arousal; Generalized anxiety disorder; Behavioural inhibition system; Anxiolytic; Septo-hippocampal system; Hippocampus; Theta activity; Theta rhythm

1. Introduction

Gray (1982) published a “Neuropsychology of Anxiety” which had as its key suppositions that:

1. the neural and behavioural actions of anxiolytic drugs in other animals can provide us with keys to the nature of anxiety in people;

2. the behavioural actions of the anxiolytics drugs are best described as impairment of the ‘Behavioural Inhibition System’ (BIS) depicted in Fig. 1;

3. the most important common neural actions of the anxiolytic drugs are to impair the control of ‘theta activity’ in the septo-hippocampal system (SHS);

4. changes in septo-hippocampal function, and especially theta activity, can underlie both normal and pathological changes in anxiety.

An updated “Neuropsychology of Anxiety” (Gray and McNaughton, 2000) retains all of the above
suppositions and summarizes the extensive confirmatory evidence for them which has accrued since 1982. It also describes modifications of the specific mechanisms of the theory needed to bring it into line with the details of these recent data. In this paper we will summarise the 1982 theory and briefly describe the aspects of its updating which are directly relevant to the role of theta activity (omitting many other aspects). We then consider in more detail the implications for anxiety of what we now know about the mechanisms of control of theta activity and of the actions of anxiolytics on that control.

2. The Neuropsychology of Anxiety (1982)

The foundations of the theory (both in 1982 and in its updated form) are the data showing that:

1. anxiolytic drugs produce behavioural effects in animals that are qualitatively (although not always quantitatively) the same as lesions of the SHS;
2. anxiolytic drugs impair the control of theta activity (rhythmic burst firing of cells in the SHS at 5–12 Hz in the free moving rat), the occurrence of which depends on a pacemaker input from the medial septal area.

The theory then attributes the behavioural, including clinical, effects of anxiolytics to the disturbance which they produce in the control of theta.

The details of the theory are more elaborate. In particular, they take into account the fact that the SHS, while central to the theory, is only one node in a complex set of neural networks which contribute to different aspects of anxiety — and it is the interaction of the SHS with other structures that determines changes in behaviour. Thus, the SHS is not the only structure which we postulate is involved in the control of anxiety. Nor is the proposed function of the SHS solely to control anxiety. Indeed, in both the 1982 theory and its updated version, the proposed role of the SHS is to act as a simple comparator of inputs. It is the nature of the mismatches detected by the comparator, and the nature of the action taken as a result of significant output from the comparator to other brain systems, which give rise to the strong relation we postulate between activity in the SHS (and hence BIS) and anxiolytic action.

The comparator function envisaged in the 1982 version of the theory has a number of key features of its operation, for our present purposes:

1. much of the time the comparator is receiving inputs about the state of the world and the expected state of the world but produces no output as it is ‘just checking’;
2. these inputs code not only for simple events (and particularly for the occurrence of reinforcers) and for predictions about expected events but also for the upcoming step in the currently executing motor program and for future steps (i.e. plans);
3. when there is a mismatch between the expected and actual events the comparator produces significant output that not only tends to halt the current motor programme but also to cause it (and hence
future plans) to be executed more slowly and carefully in the future or to be replaced by attempts to resolve the problem which gave rise to the mismatch (e.g. increases in exploratory behaviour).

In the context of the BIS of Fig. 1, this comparator will of course produce an output to each of signals of punishment, signals of reward omission and novel stimuli. In each of these cases a prepotent response, which previously led to appetitive stimuli (or to the avoidance of aversive stimuli), has resulted in an unexpected event.

While these proposed operations of the SHS comparator, and of the BIS in which it is embedded, can be described fairly simply in words, the neural machinery attributed to them is not simple. Words such as ‘plan’ are held to imply extensive neural machinery (involving, in the case of ‘plan’ itself, the bulk of the frontal and prefrontal cortex — a substantial part of the cortical mantle). The key neural elements which are proposed to instantiate the comparator within the SHS are detailed in Gray (1982).

There is also one part of the neural machinery of the 1982 theory, the septal area, which might seem superfluous in terms of the model of Fig. 1 and our view of the comparator since all the required functions are attributed to other specific areas. The septal area was (and is) included in the theory for quite different, neurophysiologically-driven, reasons than the psychologically-driven approach which has characterized our description so far. The septal area is included (as are the noradrenergic input from the locus coeruleus and the serotonergic input from the raphe) because it has been shown to control theta activity and the latter has been shown to be affected by anxiolytic drugs. The medial septal area is, therefore, a key node in the circuits through which the anxiolytic drugs can alter the functioning of the hippocampal formation and hence, in our theory, alter the comparator function.

We will discuss theta activity and the effects on it of anxiolytic drugs in more detail in Section 4. Here we need only note that in the 1982 theory theta activity was held to be required for the hippocampal comparator to operate correctly and, in particular, was necessary for the quantization of information passing round recursive loops between the hippocampus and neocortex (Parmeggiani et al., 1971; see also Miller, 1991).

3. The Neuropsychology of Anxiety (second edition)

The 1982 theory attempted to encompass all the information available at that time on the neural and behavioural actions of the anxiolytic drugs, the nature of theta activity and the neural and behavioural functions of the SHS. The explosion of data since its publication prompted us to produce a second edition of “The Neuropsychology of Anxiety” to assess the predictive validity of the 1982 theory and to update its formulation. Some of the updating is cosmetic — altering only the way the theory is expressed; but some involves theoretically significant alterations of the detailed machinery of the theory. A particular concern at the cosmetic level was to make clearer the relationship between the proposed functions of the hippocampus in the control of anxiety and the known involvement of the hippocampal formation in temporal lobe amnesia. (Briefly, our view is that so-called ‘amnesia’ is better described as a ‘catastrophic hypermnesia’. That is, the loss of behavioural inhibition resulting from impairment of SHS function causes a failure to suppress storage or recall of incorrect alternatives rather than, as is normally assumed, a failure to store or recall correct alternatives.)

The most important development of the last 15 years or so has been the appearance of the novel anxiolytics: drugs such as buspirone which act at or through 5HT1A receptors and which share the primary anxiolytic action, but none of the side effects, of the classical anxiolytics (e.g. benzodiazepines, barbiturates). These drugs provide a stringent set of predictive tests of the core tenets of the theory. It passes these tests with flying colours: novel anxiolytics are like classical anxiolytics in (a) affecting the control of theta activity (see Section 4) and (b) producing behavioural effects like those of hippocampal lesions (Table 1). The hippocampal-like action of classical and novel anxiolytics extends even to tests with a maximally hippocampal and minimally anxiolytic flavour such as the Morris water maze.
Table 1

A comparison of the behavioural effects of hippocampal formation lesions and anxiolytic drugs. The data are taken from Gray and McNaughton (2000). In most cases the symbols summarise the effect generally found in a large number of experiments. In the case of the symbols surrounded by squares the indicated result has been obtained not only with classical anxiolytics but with the novel anxiolytic buspirone. In the case of buspirone the results are typically from only one experiment and are usually obtained only at low doses. The pattern of effects suggests that hippocampal lesions and anxiolytics have very similar effects; that they do not impair motivation, perception or memory in any general sense; and that the bulk of their effects reflect a simple loss of behavioural inhibition (see Fig. 1).

<table>
<thead>
<tr>
<th>Task</th>
<th>HIP</th>
<th>ANX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rewarded Bar Pressing, CRF</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rewarded Running, CRF</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Simultaneous Discrimination</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shock, Skilled Escape, Active Avoidance</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Resistance to Extinction</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Two Way Active Avoidance</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Non-spatial Active Avoidance</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Rewarded Bar Pressing, Intermittent Reinforcement</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Successive Discrimination</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Spontaneous Alternation</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rearing</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mazes (not Radial Arm)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Passive Avoidance</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Differential Reinforcement of Low Rates</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Spatial Discrimination Reversal</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Water Maze</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

= Tested with Buspirone

Another positive development for our theory is the change that has occurred in theories of hippocampal amnesia. In 1982 there was an apparent chasm between our theory (which identified the hippocampus with anxiety) and other theories (which identified it with the storage of memories, particularly spatial memories; O‘Keefe and Nadel, 1978). More recent theories of hippocampal function (e.g. Cohen and Eichenbaum, 1993) do not see the hippocampus as the site of long-term memory storage. Rather they see the hippocampal formation as being critically involved in particular kinds of cognitive processing (e.g. of relations between stimuli in the Cohen and Eichenbaum case). This emphasis on processing as opposed to storage brings these recent theories very close to the updated version of our theory (see below). But they universally see the hippocampal formation as aiding the production of certain types of true memories; whereas, we see it as aiding the suppression of certain types of false memories. This is more than a semantic distinction as it leads us to see the presence or absence of interference within a paradigm as more important than the nominal type of memory being tested in determining sensitivity to hippocampal damage. It also allows us to account for the presence of intrusion errors in the responses of amnesics. (It is difficult to see why previously correct items should be produced incorrectly on subsequent trials if the main result of hippocampal damage was to prevent their initial storage.)

Many data have, thus, fulfilled the predictions of the prime tenets of Gray’s (1982) theory: anxiolytic drugs of all chemical types and modes of action impair the control of theta activity; anxiolytics of all chemical types and modes of action have behavioural effects similar to hippocampal lesions; and the SHS is still best viewed as the most crucial neural substrate of the conceptual BIS. Other data, however, have caused us to rethink either the justification for elements of the original theory or the details of its specific mechanisms.

The most substantial rethinking has been a simplification of the definition of the inputs to the BIS and a simplification of the information processing operations attributed to the BIS itself. There has also been a major addition: the allocation of a role for the amygdala in the processing of anxiety. We will deal with these in order below. However, while they all involve changes in the precise manner in which the SHS is held to discharge the functions of the BIS, they do not involve changes in the fundamental operation of the BIS, nor in our predictions about the effects of hippocampal dysfunction nor in any of the other gross features of the 1982 theory.

Our simplification of the definition of the inputs to the BIS is derived directly from the ethopharmacological work of Caroline and Robert Blanchard (Blanchard and Blanchard, 1989, 1990a, b; Blan-
Using a semi-natural environment in which rats could live continuously, and using natural stimuli (e.g. a cat) to elicit defensive behaviour, they have made a strong distinction between fear-related behaviours (essentially those appropriate when a predator is immediately present, e.g. fight/flight) and anxiety-related behaviours (essentially those appropriate when a predator may or may not be present, e.g. risk assessment). They also showed that ‘fear’ and ‘anxiety’ (as they define them) each produces somewhat different types of behaviour depending on what they term the ‘defensive distance’ or degree of threat. Conveniently for us, anxiolytic drugs as a class were found to consistently affect the key anxiety-related measures of risk assessment and not to affect the key fear-related measures of avoidance, flight and unconditioned freezing. Also, importantly, the effect of the drugs on rearing (a measure of risk assessment) was to increase this if it was low because of a small defensive distance, but to decrease it if it was high because of an intermediate defensive distance. Anxiolytics did not, therefore, so much suppress rearing as shift the animal’s choice of behaviour to that characteristic of a greater defensive distance. That is, the drugs reduced anxiety itself rather than simply suppressing anxiety-related behaviours.

While the Blanchards themselves distinguish fear and anxiety by the presence or absence of a predator, we see this as a special case of a more general principle. In learning experiments it is not the delivery of shock within a test situation which renders behaviour sensitive to anxiolytic drugs but rather whether the animal must suppress responding to avoid shock (Gray, 1977). To the concept of defensive distance (which, as the Blanchards use it, determines the precise defensive behaviour emitted) we add the notion of defensive direction (which determines not only the class of defensive behaviours emitted but also their sensitivity to anxiolytics). Pure fear, then, is held by us to control behaviour which allows an animal to leave a dangerous situation; and anxiety to control behaviour which allows an animal to enter a dangerous situation. On this view, anxiety also necessarily involves a conflict between two incompatible goals — in the simple defensive case, a conflict between, say, obtaining food for yourself and not becoming food for someone else. Learning theoretic, ethological and pharmacological experiments (Gray, 1982) also led us to equate the effects of shock with the effects of the omission of an expected reward (frustrative non-reward) and to see the latter as engendering anxiety when it results in an approach–avoidance conflict. With the accent, then, not on approach or avoidance as such but on the conflict between them, we also added to the list of anxiety-provoking stimuli those novel stimuli which have both threatening and curiosity-inducing aspects. With these the same stimulus tends to elicit both approach (curiosity driven) and avoidance (potential threat driven). The Blanchards’ work, then, allows us to derive from ethological first principles the inputs to the BIS (Fig. 1) which were each separately postulated ad hoc in the 1982 version of the theory.

Our second simplification of the information processing operations attributed to the BIS was driven by two issues. First, data driven, was the remarkably simple and repetitive neural organisation of the hippocampus. This was well known in 1982 but was not fully taken into account in our attribution of a relatively large number of complex information processing operations to the SHS. Second, theory driven, was the redefinition of the BIS described above. This emphasizes approach–avoidance conflict as a common feature of the inputs to the BIS more than it emphasizes the differences between the stimuli themselves. These two issues, taken together, led us to a modified view of the SHS, and particularly of its comparator function (represented in Fig. 2). Fig. 2 shows the general way in which comparisons are made between goals in our theory. It could be taken to imply (as we did in 1982) that the SHS contains only one comparator. However, we now see more stimulus-based goal differences (see Fig. 2) as usually being processed in entorhinal cortex and more response-based goal differences as usually being processed in hippocampus proper. Fig. 2 is also simplified in that it does not include the fact that the way the circuitry operates within each structure differs depending on modulatory inputs, particularly monoaminergic — and this known variation is important for the specific operations we attribute to the comparators.

As can be seen from Fig. 2, we see the SHS as receiving information from all areas which have the capacity to activate goal directed behaviour. These
areas send to the SHS what are, in effect, efference copies of the activity which, within the goal processing system, defines the goal. A key point about the use of the word goal here is that it conflates stimulus and response features. A goal is neither a stimulus nor a response *per se* but is a stimulus (or stimulus complex) with an associated action tendency. Thus, in the diagram, Sb (a specific environmental feature) can activate both G2 and G3 — goals which are distinguished by their differing action tendencies. Similarly, G3 and G4 represent the same action tendency but are different goals because they are directed to different aspects of the environment.

R2 and R3 could represent approach and avoidance respectively, in which case the concurrent activation of G2 and G3 would represent an ap-
proach–avoidance conflict. In the model this is detected very simply by the comparator within the SHS. If the SHS receives input from only one highly activated goal it monitors this fact but produces no functional out. (This represents the ‘just checking’ condition in the 1982 theory.) As soon as a second goal becomes similarly activated, summation of these activities will pass a threshold and produce output from the SHS. (This represents the ‘mismatch’ condition in the 1982 version of the theory.)

Note that, in the model, each stimulus (environmental feature) has the capacity, in principle, to elicit both affectively positive (+) or affectively negative (−) associations of any particular goal depending on the prior reinforcement history (or innate consequences) of that particular RiSj combination. (If R2 and R3 represent approach and avoidance respective- ly, the effect of conditioning with Sb should be opposite on G2 and G3. For example, pairing Sb with a shock should tend to suppress G2 and activate G3. Note that in terms of the diagram this means that increase in the input marked ‘−’ will tend to increase the activation of G3.)

Output from the comparator is held to have two effects. First, it tends to suppress both of the currently conflicting responses. This effect is likely to be selective to the two (or more) equally and highly active goals — since conflict can result in displacement activities. We would see displacement activity in this model as resulting from the release of, e.g., a moderately active G4 when highly active G2 and G3 conflict with each other and are then suppressed by the comparator. Second, it increases the valence of the affectively negative associations of each of the conflicting goals. A particularly important feature of this latter output for explanations of anxiety is that this increase in negative affective value will, through normal associative mechanisms, cause a reduction in activation or even total suppression of affectively positive goals on subsequent occasions and will tend to cause an increase in activation of affectively negative goals. This is an affectively biased alteration of memory representations and is the point at which our theory makes direct contact with the hippocampal memory literature. Insufficient activity in the SHS will result in a failure to suppress competing incorrect responses. It is also the basis for our view of generalized anxiety.

Excessive activity in the SHS will, on our view, lead to excessive perception of environmental threat and, as an indirect consequence, the creation of increased numbers of threatening associations of previously negative stimuli.

We now come to the major addition to our theory: the role of the amygdala. Fig. 2 is silent as to the location in the brain of G1–Gn and we have already stated that these can be located in any area which processes goals. Where conflict results from, e.g. the concurrent activation of G3 and G4 (which are distinguished by addressing Sb and Sc respectively), this reflects a difficulty in deciding where to make a response such as a lever press not a difficulty in deciding which response to make. This kind of conflict will be most prevalent in hippocampal-sensitive ‘memory’ paradigms and G3 and G4 will usually be encoded in frontal or temporal cortex. (It is important to note that, in our model, the SHS deals only with conflicts between concurrent goals, including subgoals. Conflict in the ordering of subgoals within a sequence is dealt with by frontal cortex and conflict in the motor programmes required to achieve a single clear goal, as in mirror drawing, is dealt with by the motor system.) Where conflict results from, e.g., the activation of G2 and G3 and where R2 and R3 represent approach and avoidance tendencies, respectively, defence systems in the brain, and particularly the amygdala, will often be involved.

Since 1982 there has been an extensive assault on the organisation of defensive systems. Results, especially from the laboratories of LeDoux and Davis (e.g. LeDoux, 1992, 1994; Rogan et al., 1997; Davis, 1992a, Davis et al., 1997) have shown that the amygdala is a key site at which the plasticity underlying fear conditioning occurs and this fits in with a view of the amygdala as one of the higher centres of a hierarchical defence system (e.g. Graeff, 1994). In our view (Fig. 3) this system is best seen as being divided into multiple levels (which essentially reflect the Blanchards’ concept of defensive distance) and into two parallel streams (which essentially reflect the two possible defensive directions; i.e. the Blanchards’ fear/anxiety distinction).

We agree, then, with those who believe that the amygdala (and related areas such as the bed nucleus of the stria terminalis) is engaged by defensive
Fig. 3. The Hierarchical Defense System. This diagram represents an integration of the views of Davis (1992b, Davis et al., 1997), LeDoux (1992, 1994) and, particularly, Graeff (1994) on the neural organisation of defense and is organised in terms of the concepts of defensive distance and direction derived from the work of Blanchard and Blanchard (1989, 1990a). The lowest neural levels of the system (particularly the periaqueductal gray) control responses in a very simple and immediate fashion when the defensive distance is very small. As defensive distance increases, more complex defensive strategies appear, controlled by ever higher neural levels of the system with the cingulate cortex representing the highest levels. Defensive avoidance (‘fear’) is controlled by the amygdala and anterior cingulate. Defensive approach (‘anxiety’) occurs when a strong appetitive tendency conflicts with avoidance; is characterized by high levels of risk assessment behaviour; and is controlled by the septo-hippocampal system and the posterior cingulate.

anxiety. But there are tasks (e.g. fixed interval schedule) which should, in theory, engage the BIS and which are, in practice, sensitive to anxiolytic drugs but which are not sensitive to amygdala lesions and are sensitive to SHS lesions. So, we see the SHS as coding critical aspects of anxiety and, on occasion, engaging the amygdala to produce key (particularly fear-related) outputs. We do not see the SHS (in contrast to, e.g., LeDoux, 1994) as an area which encodes complex (e.g. contextual) stimuli and so is required as an input to the amygdala only for a subset of cases where the latter requires these to support anxiety. In our view, then, the SHS encodes anxiety of all types and, in a subset of cases involving fear, generates its increase in negative bias by increasing avoidance tendencies controlled by the amygdala. The amygdala, then, controls behavioural avoidance consequent on anxiety because it mediates the fear which underlies both active and passive avoidance tendencies and because anxiety generates passive avoidance.

There is one exception to the otherwise hippocam- po-centric view we present of the control of the outputs of the BIS — the arousal output. Unlike anxiolytic drugs, SHS lesions do not affect arousal in anxiety-provoking situations and, in the first edition, we attributed the arousal output to the noradrenergic input from the locus coeruleus to the hypothalamus. Since then, however, the specific case of the arousal output demonstrated in the paradigm of ‘fear potentiated startle’ has been shown to be unaffected by hippocampal lesions and to be mediated by the amygdala (McNish et al., 1997). We, now, therefore, attribute to the amygdala rather than the hypo-thalamus the ‘increase arousal’ output of the BIS.

Why should this one output of the BIS be me-
mediated solely by the amygdala and not at all by the SHS? This is most easily understood in threatening situations. Where there is conflict, the animal must often switch very quickly between fear-related and anxiety-related behaviours. These are in general opposite to each other and so would be expected to require a system, such as the SHS, to resolve the conflict between them. However, there should be relatively constant autonomic output not only because this affects bodily organs only with considerable delay (inconsistent with fast switching) but because this will be as advantageous with fear as with anxiety.

In summary, in our view, activation of the hippocampus but not the amygdala will constitute non-anxious rumination; activation of the amygdala (or rather the fear-related parts of it) by itself will constitute ‘pure’ fear; and activation of both together constitutes anxiety. Given this parallel activation, some outputs of the BIS are controlled largely by the circuitry of the SHS (e.g. behavioural inhibition which depends on frustration rather than fear — which is not affected by amygdala lesions), some are controlled largely by the amygdala (e.g. fear potentiated startle) and some depend on the combination of the two (e.g. behavioural inhibition which depends on fear). In some cases we expect the amygdala and hippocampus will interact via their direct reciprocal connections. However, we have postulated, above, that the amygdala can affect SHS processing indirectly by its capacity to increase arousal.

We now turn to consideration of the theta input to the hippocampus which can be thought of as reflecting just such arousal.

4. Anxiolytic action on the control of theta activity

In Gray (1982), as in O'Keefe and Nadel (1978) and in Miller (1991), theta activity was held to be essential for hippocampal function. In all three cases, too, the precise frequency of theta was held to be important, reflecting some form or other of indexing function. However, the reformulation of the comparator shown in Fig. 2 suggests a less absolute, but nonetheless fundamental, role for theta activity in the SHS.

In Fig. 2, the comparator is held to receive inputs from the various goal areas, compare them, and produce an output when there is no single goal which is clearly more highly activated than any other. The output increases negative bias by modification of the power of existing negatively affective inputs to the conflicting goals. This will, in turn, alter the strength of activation of those goals which have such an input (some goals, of course, may have no history of negative consequences) and hence, often, alter the balance of the inputs to the comparator. If this model is, in fact, instantiated in the known hippocampal-cortical loops, then the passage of information, from hippocampal comparator, to Gn and then back to comparator should generally take of the order of 50–200 ms (Miller, 1991).

It should be noted that the alteration in bias produced by a single passage of information round the proposed loops should often be insufficient to resolve the conflict. In this case, we propose that the hippocampal comparator increases its output and essentially continues to do so, recursively, until one or another goal becomes predominant (meanwhile eliciting, e.g. exploration which can obtain new information to help resolve the conflict). A potential problem, here, is that pathways with short (50 ms) conduction times will then be more heavily biased than pathways with long (200 ms) conduction times if this recursive process is allowed to continue unfettered. In cases where conflict resolution is simple this may not impair performance and, in these cases, we would hold theta activity is not necessary for adequate operation of the SHS. However, where conflict resolution is complex, requiring many iterations of recursive processing, the presence of theta activity (which results from a phasic inhibitory input from the medial septal/diagonal band complex) will allow significant processing only briefly and at intervals of between 100 and 200 ms. This will keep the computations being performed by different circuits in register and so improve the accuracy or acuity of the system.

The medial septal area has been shown to maintain a direct pacemaking control over theta activity in the entorhinal cortex, subiculum and posterior cingulate cortex as well as in the hippocampus proper. (Indeed, we use this input to define the elements of what we refer to as the SHS.) In our view, then,
elimination of theta activity will produce behavioural effects unlike those of lesions of the hippocampus proper in two ways. First, as implied in the previous paragraph, the effects will be quantitatively less — a reduction in the acuity of, not an elimination of, processing. Second, given the multiple components of the SHS, just listed, the effects will be qualitatively more — involving impairments in functions dependent as much on, say, entorhinal or posterior cingulate cortex as on hippocampus proper.

Note that theta represents only an input to the SHS and cannot be used as a marker of SHS functional output. Theta is known to vary in frequency in relation to movement and, in a stationary animal, to the level of arousal (Sainsbury, 1985). For example, high frequency theta can be obtained by presenting a rat with a predator. Thus, theta is known to occur and to show a large number of tight relationships to behaviour in situations which are totally insensitive to SHS lesions.

In understanding the effects of the anxiolytic drugs on behaviour, then, we must take into account the fact that mediation of these effects by changes in the control of theta activity implies that they will act like mild or diffuse dysfunction to the entire SHS. We must also take into account that their effects on the control of theta activity are themselves subtotal. They do not eliminate theta activity in any general way. Interestingly, they impair the control of theta activity in two quite separate and distinct ways.

The effect of anxiolytic drugs on the control of theta activity which received most attention in Gray (1982) was their effect on septal driving of theta. As we have noted already, theta activity in the SHS depends on a pacemaker input from the medial septal area. As shown in Fig. 4A and 4B, electrical

![Fig. 4. The effects of anxiolytic drugs on septal driving of theta activity. Trains of single pulses delivered at frequencies in the region of 5–10 Hz to the medial septal area (B) cause rhythmic discharge of septal cells and then hippocampal cells with each burst of single cell firing phase-locked to a stimulus. Concurrent phase locking of large numbers of hippocampal cells (both with septal stimulation and under normal physiological conditions) gives rise to a gross extracellular theta rhythm (Ai, ii). With incorrect electrode placement (Aiv) evoked potentials rather than theta rhythm are recorded. Ai and Aiii were obtained from the same animal with higher intensity stimulation producing the evoked potentials of Aiii. The threshold stimulation intensity required to drive theta rhythm is lowest in male rats in the region of 7.7 Hz (130 ms) as shown by the pre-drug curves (○) in panel C. Anxiolytic drugs (●: BUS = buspirone; IMI = imipramine; EtOH = ethanol) flatten the curve by raising the threshold in the region of 130 ms. Other drugs such as methysergide (METH), haloperidol (HAL) and scopolamine (SCOP, see text) are either totally inactive or have quite different effects on the threshold-frequency function.](image-url)
stimulation of the medial septum at frequencies in the normal range of theta activity can ‘drive’ theta — each electrical pulse produces a theta wave, phase-locked to it. In freely moving male rats the threshold voltage (or current) at which this driving can be produced is a U shaped function of the stimulation frequency with a minimum in the centre of the normal range (7.7 Hz, 130 ms inter-pulse interval). Clinically effective anxiolytic drugs all increase the driving threshold in the region of 7.7 Hz, flattening the threshold-frequency function (Fig. 4C). An important support for the 1982 theory is that novel anxiolytics, such as buspirone and the anxiolytic antidepressant imipramine, have now been shown to share this effect of the classical anxiolytics.

Drugs which are not clinically effective as anxiolytics generally do not produce this effect. Of particular interest, here, is the effect of the anticholinergic drug scopolamine. This appears (Fig. 4C) to produce an increase in the threshold for driving — but this result is obtained in a peculiar way. Septal stimulation at the pre-drug threshold sometimes produces theta but sometimes has no effect at all. Thus the threshold itself has not been increased as it would be by an anxiolytic. Rather the presence or absence of theta has become dependent on whether the animal is moving or not, respectively — a phenomenon previously noted with spontaneous theta activity and forming the basis for Vanderwolf’s (1969, 1988) distinction between two types of theta. We will discuss the role of ‘cholinergic theta’ further below.

The effect of the anxiolytic drugs on septal driving of theta can be reproduced by selective neurotoxic lesions of the dorsal ascending noradrenergic bundle which arises in the locus coeruleus and supplies all the noradrenergic input to the hippocampus (Gray et al., 1975). An important prediction of the 1982 theory, then, was that such lesions should reproduce the common effects of anxiolytic drugs and SHS lesions. A review of the literature showed that this prediction was confirmed, but only for about half of the relevant behavioural tasks (McNaughton and Mason, 1980).

The anxiolytic drugs have a second effect on theta activity which has received most study since 1982. High frequency stimulation of the reticular activating system elicits theta activity at a frequency which is proportional to the strength of stimulation. Interestingly, there is no sign of non-linearity of the intensity–frequency function in the region of 7.7 Hz and no effect on it of depletion of noradrenaline. It is, therefore, neurophysiologically and pharmacologically distinct from the septal driving function discussed above. All classes of anxiolytic drug, including the novel anxiolytics buspirone and imipramine, which have only been tested since 1982, reduce the frequency of reticular-elicited theta (Fig. 5) and drugs which are not clinically anxiolytic generally do not reproduce this effect. (Taking the septal driving and reticular elicitation tests together there are no anxiolytic drugs which do not affect these tests appropriately and no non-anxiolytic drugs which do.) As with the septal driving case, administration of scopolamine did not affect the basic neurophysiological function but did make the presence of theta depend on whether the animal was moving.

This effect of the anxiolytic drugs can be reproduced by dysfunction of the supramammillary nucleus (see next section) which, in a fixed interval schedule, reproduces quantitatively as well as qualitatively not only the effect on theta frequency but also the effect on behaviour of anxiolytic drugs (Woodnorth and McNaughton, in preparation). Since this task is insensitive to dorsal bundle lesions, this result shows that the common behavioural profile of anxiolytic drugs and SHS lesions results from at least two quite distinct effects of the anxiolytics on the control of theta activity.

5. The subcortical control of theta activity . . .

We have already indicated that the locus coeruleus provides a noradrenergic input to the hippocampal formation that alters the threshold at which input from the medial septum can entrain theta activity. Thus, under some circumstances, reductions in this noradrenergic input could prevent the occurrence of theta in the middle of the frequency range when otherwise it would have occurred. This effect is best viewed as a ‘gating’ of theta since loss of noradrenaline does not affect the frequency or amplitude at which theta rhythm occurs, but only whether it occurs or not. It turns out that there are a number of
quite diverse subcortical inputs of this type. The most obvious in our discussion so far has been the cholinergic system. Theta activity can occur in a non-moving animal, but only if cholinergic systems are intact (Vanderwolf, 1988). We have recently (McNaughton et al., 1997) used stimulation of the pedunculopontine tegmental nucleus to elicit theta in urethane-anaesthetized animals (Vertes, 1981; see Vertes and Kocsis, 1997). We then mapped the ascending theta control system with injections of local anaesthetic and found (Fig. 6A) that the output from the pedunculopontine tegmental nucleus made cholinergic relays in an area within the midbrain reticular formation, in the substantia nigra, in the superior colliculus and in the amygdala. Interestingly, blockade of any one of the branches of this system was sufficient to abolish theta activity. Since Dringenberg and Vanderwolf (1996) had shown that theta elicited from the amygdala was relayed by cholinergic synapses in the medial septal area, our results suggest a diffuse co-operative cholinergic network (shaded in Fig. 6A) which has a permissive (gating) effect on theta activity. There is also evidence that the permissive gating of theta activity which occurs during movement, and which is insensitive to cholinergic blockade, is serotonergic in character and, presumably, arises in the raphe nuclei (Vanderwolf, 1988; Fig. 6A).

The subcortical control of theta frequency appears no simpler than that of cholinergic gating. In rats anaesthetized with urethane, stimulation of nucleus reticularis pontis oralis can elicit theta activity (Vertes, 1981) the frequency of which is encoded in the supramammillary nucleus (Kirk and McNaughton, 1991, 1993). However, in un-anaesthetized animals, nucleus reticularis pontis oralis stimulation produces theta not only via the supramammillary nucleus but also via some other, as yet unidentified, nucleus (McNaughton et al., 1995; marked as DMH? in Fig. 6A). Whether the supramammillary nucleus controls theta frequency or not depends on the behaviour of the animal. During exploration, the supramammillary nucleus appears to be involved neither in the control of theta frequency nor behaviour (Thinschmidt et al., 1995; Woodnorth and McNaughton, in preparation); in the Morris water maze, the supramammillary nucleus makes only a modest contribution to theta frequency and to behaviour (Pan and McNaughton, 1997); whereas in a fixed interval schedule, chlordiazepoxide injected into the supramammillary nucleus has as big an effect on theta frequency and on behaviour as it does

Fig. 5. The effects of anxiolytic drugs on reticular-elicited (100 Hz stimulation) theta activity. Increasing strength of reticular stimulation produces a linear increase in theta frequency. Compared to pre-drug (○), anxiolytic drugs (●: CDP = chlordiazepoxide; DZP = diazepam; ALP = alprazolam; AMY = amylobarbitone) reduced the frequency of theta. Drugs which act to block acetylcholine (ACh), serotonin (5HT), dopamine (DA) or noradrenaline (NA) do not have this effect. For further comment on both ACh and NA, see text.
Fig. 6. The subcortical control of theta activity. Solid lines represent known direct anatomical connections. Dotted lines represent inferred indirect connections. A. Some of the neural systems known to control distinct aspects of theta activity. (i) serotonergic (5HT) inputs from the raphe can permit theta to occur, without contributing to frequency or amplitude as such. They can best be thought of as gating theta activity, and do so in parallel with (ii) cholinergic (Ach) gating input which is the net result from co-operative activity in a network (shaded boxes) of interlinked structures. The structures currently known to be elements of this network are: PPT = pedunculopontine tegmental nucleus; MRF = midbrain reticular formation; SC = superior colliculus; SN = substantia nigra; AMYG = amygdala; MS = medial septum/diagonal band complex. (iii) GABAergic inhibitory input from the medial septum (MS, unshaded) provides a phasic 'pacemaker' for theta activity. The frequency of this phasic input is determined in the supramammillary nucleus of the hypothalamus (SUM) acting in concert with other nuclei (possibly including the dorso medial hypothalamus, DMH?). This system has been shown to produce frequency as a linear function of 'arousal', i.e., the intensity of input from areas such as nucleus reticularis pontis oralis (RPO) and possibly the periaqueductal gray (PAG?). (iv) Noradrenergic (NA) input from the locus coeruleus (LC) controls the threshold for theta elicitation (see Fig. 4). The change in threshold can be thought of as a weak, frequency-specific, form of the gating seen with 5HT and Ach. B. A speculative model of the operation of the multiple parallel components of the theta control system shown in A. In different behavioural conditions both gating and frequency may depend more or less selectively on one or another subset of nuclei. Given their known inputs, for example, it seems quite likely that during threat, superior colliculus will be significantly involved in cholinergic gating and the non-supramammillary pathway (hypothetically PAG–DMH–MS) will be significantly involved in frequency control. Likewise, given that PPT and RPO are known to be involved in the control of sleep, it seems likely that the theta activity which occurs in REM sleep is cholinergically gated by PPT and depends for its frequency on the RPO–SUM–MS pathway.
6. ... and its implications for the role of arousal in anxiety

There has always been a clear link between arousal and the occurrence of theta activity. Green and Arduini (1954) stated that arousal led to theta activity and Brucke et al. (1959) went so far as to describe theta as ‘der “Hippocampus-arousal-Reaktion”’. More recently, Sainsbury (1985) has linked the occurrence of theta in a non-moving animal to the presence of arousal — and has obtained particularly high frequencies of theta in response to the arousal produced by predatory threat. However, the organisation of the subcortical systems controlling theta activity means that we cannot simply identify arousal with theta activity nor theta activity with anxiety.

Neurally speaking, arousal is often identified with activity in the ascending reticular activating system. The level of this activity is transduced, continuously, into ‘theta frequency’ (by e.g. the supramammillary nucleus) but this does not mean that the SHS shows theta activity continuously. In addition to the ‘arousal’, which determines theta frequency, there must also be permissive gating input from either the cholinergic or serotonergic systems — which can each be thought of as coding ‘arousal’ in other senses. With a natural stimulus such as a predator we know (Sainsbury, 1985) that the cholinergic gating system will be activated concurrently with the frequency control system. But this unity of psychological stimulus does not mean there is a unity of the nature of arousal. In the absence of the cholinergic gating input, the theta activity that occurs during movement is known to be of the same frequency as it would have been if it had been preceded by a burst of cholinergically-gated theta. This means that the frequency control mechanism is integrating ‘arousal’ quite independently of whether the cholinergic system is active or not.

Perhaps the most important implication of the subcortical control of theta activity is that arousal (in some sense or other) is a necessary but not sufficient condition for the generation of anxiety. As we have already noted the presence of arousal (encoded as an increase in theta frequency within, say, the supramammillary nucleus) does not guarantee that theta activity will occur in the SHS. But, further, theta activity in the SHS does not guarantee that there will be functional ‘anxiogenic’ output from the SHS. There are many tasks (e.g. running down a runway) in which behaviour is accompanied by theta activity but is not dependent on an intact SHS. In our theory this is because the hippocampus is in its ‘just checking’ mode and, only when the comparator detects conflict, will there be significant functional output.

The circuitry also shows there are different types of arousal. Theta frequency, which we take as the index of the level of hippocampal arousal, is calculated by a dedicated circuit and has been shown, in at least some tasks, to reflect much more, e.g., planned acceleration of movement through space than it does, e.g. level of autonomic activation.

Lesion or pharmacological manipulation also indicates different psychological contributions of the different types of arousal. Deficient activity in the cholinergic gating system appears to produce amnestic effects — presumably as a result of the fact that it totally eliminates theta (at least when the animal is not moving). By contrast, anxiolytic drugs have much milder effects on memory while affecting the perception of threat. On our theory, this lesser effect of anxiolytics would be because they reduce the frequency of theta (which would be directly related to the level of threat) and do so even-handedly whether the animal is moving or not.

It also appears likely that different types of ‘arousal’ fulfill different functions even in relation to anxiety. Thus both the gating and frequency of theta could be controlled by one set of structures (speculatively superior colliculus and dorsomedial hypothalamus; Fig. 6B) in the presence of threat and by a second set (the pedunculopontine tegmental nucleus and nucleus reticularis pontis oralis; Fig. 6B) in tasks such as the fixed interval schedule.

7. Conclusions

While this overview of our theory and of the subcortical control of theta activity has been both
brief and selective (see Gray, 1982; Gray and McNaughton, 2000, for more extended treatment), we believe it gives good reason to see amnesia and generalized anxiety disorder as opposite poles of a continuum defined by the reactivity of the SHS. While it may seem peculiar that both extremes of output from the SHS should result in psychopathology, it should be noted that the complex circuitry controlling the nature and occurrence of theta activity would not be necessary if theta and hippocampal output were universally beneficial.

Arousal has always been a slippery term within psychology and we have made no attempt to define it here. Rather, we argue that the known architecture of the theta control system and the face validity of theta activity as indexing hippocampal arousal means that there must be multiple, neurally quite distinct, types of arousal (which may need very careful definition); that only some of these are affected by anxiolytic drugs; and that anxiolytic drugs act to affect multiple types of arousal — each of which can be involved in the control of anxiety under specific limited behavioural conditions.

It also follows from our analysis that arousal and anxiety may be linked, at least to some extent, in a positive feedback loop. Arousal, in the most general sense, increases theta frequency and permissive theta gating. This, in turn, increases the acuity of hippocampal processing. This, in turn, could lead to an increase in negative bias, i.e. the perception of threat, as a result of output from the SHS to the amygdala and so to an increase in arousal which would close the positive feedback loop.

References


