

Behavioural models in psychopharmacology

PAUL WILLNER

Most psychopharmacologists who work with animals do so in the belief that their research will ultimately be relevant to people. The behavioural paradigms and procedures that they use therefore serve, in some sense, as models of human behaviour. However, the precise nature of this relationship is often left implicit. If we attempt to define what, in this context, is meant by the term 'model', it rapidly becomes apparent that its meaning is as varied as the interests of psychopharmacologists, who come from a diversity of academic backgrounds and work in a variety of institutional settings. Inevitably, attempts to define the term 'model' succeed in describing only some of its uses. Take a recent example:

Animal models represent experimental preparations developed in one species for the purpose of studying phenomena occurring in another species. In the case of animal models of human psychopathology one seeks to develop syndromes in animals which resemble those in humans in certain ways in order to study selected aspects of human psychopathology. (McKinney, 1984)

This seems a perfectly reasonable definition. But is it? Consider, for example, one of the older animal models of depression, the muricide test (Horovitz, 1965). In this model, potential antidepressant drugs are tested for their ability to prevent rats from killing mice. The model has been used quite widely, despite being inefficient as a way of detecting antidepressant activity and ethically objectionable (Willner, 1984a). But does it fall within the scope of our definition of 'animal models'? If animal models depend upon a resemblance between the model and the condition being modelled, it clearly does not: depressed people do

not typically kill mice, and it is difficult to see any other respect in which muricide and depression are similar.

One way to preserve the integrity of the definition is to argue that the muricide test is not really an animal model of depression, but rather, a screening test for antidepressant drugs. We would then wish to distinguish between 'animal models' and 'screening tests'. However, this has not been the position taken in reviews of the literature on 'Animal Models of Depression', for they invariably include the muricide test (see, e.g. Katz, 1981; McKinney, 1984; Willner, 1984a; Jesberger & Richardson, 1985). Should we then accept the implicit view of these reviewers that antidepressant screening tests are a type of animal model of depression? If we do, we are tempted to extrapolate models like muricide beyond their proper domain, and all manner of fallacious claims are likely to arise as a result. The dopamine hypothesis of schizophrenia provides a very obvious example of the conceptual confusion that arises when a theory of the mechanism of drug action (antischizophrenic drugs are antagonists at dopamine receptors ...) becomes a theory of the nature of the disorder being treated (... therefore schizophrenics suffer from excessive transmission at dopamine synapses).

It is clearly necessary to consider carefully, in every case, the domain in which the use of a particular model is justified. However, this prospect is not so daunting as it may appear, since some general principles may be elaborated that apply to broad classes of behavioural model. This chapter will describe three such classes, and consider what

the models in each class are designed to achieve and how they may be assessed. The term 'behavioural model', rather than 'animal model', is used deliberately, to allow an initial consideration of as wide a range of models as possible.

1.1 A TAXONOMY OF BEHAVIOURAL MODELS

A straightforward taxonomy of behavioural models emerges from simply specifying the field in which the models are to be deployed. It is clear from the most cursory glance at any of the relevant journals or textbooks, that the name 'psychopharmacology' is actually a misnomer. The subject matter of psychopharmacology is not simply the effects of drugs on mental processes, but also the underlying physiological mechanisms. The name 'neuropsychopharmacology', which is clumsier but more accurate, reminds us that we are dealing with the meeting point of three disciplines: neuroscience, psychology, and pharmacology. Correspondingly, there are three classes of behavioural model: behaviour may be used to model brain function, psychological processes, or drug actions. These three classes of model will be referred to as behavioural bioassays, simulations, and screening tests. Screening tests are 'drug centred': they are designed solely to expedite the discovery of new drugs, and therefore form a branch of pharmacology. Behavioural bioassays fall within the realm of neuroscience, in that they are 'brain centred': their purpose is to use behaviour to measure activity in a specified brain system. Simulations attempt to use animals to further our understanding of human mental processes. Thus, they are 'mind centred', and fall within psychology.

1.1.1 Screening tests

A screening test models a drug action: the search for novel psychotropic agents is based upon the actions of known drugs, which serve as reference points against which to compare the performance of new candidates. Two types of screening procedures may be distinguished, which correspond to different strategies for drug development. Traditionally, the object of screening tests has been to identify agents likely to have a specific type of clinical action: neuroleptic, antidepressant, anxiolytic, and so on.

Tests of this kind may be capable of identifying clinically effective drugs that vary widely in their chemical structure. However, they carry no guarantee, and may sometimes fail to identify structurally novel compounds. This problem is perhaps best recognized in relation to the so-called atypical antidepressants, several of which were discovered accidentally during testing in humans after failing all the standard traditional screening tests (see Danysz *et al.*, this vol., Ch. 6).

A second strategy is to identify specific biochemical actions as targets for drug development. This strategy becomes appropriate once we understand the mode of action of existing treatments. Anxiolytic drug development, for example, now proceeds primarily via the initial identification of compounds that interact with the GABA/benzodiazepine receptor complex (Squires & Braestrup, 1977). Screening for specific biochemical actions is carried out primarily by biochemical tests, but behavioural techniques are also used for this purpose. The advantage of screening for a specific neurochemical action is that the usefulness of the resulting compounds may transcend traditional diagnostic boundaries. Serotonin (5-HT) uptake inhibitors are probably the best example: developed primarily as antidepressants, these drugs may also be of value as anorectics (see Sepinwall & Sullivan, this vol., Ch. 9), as cognitive enhancers in dementia (Bergman *et al.*, 1983), as anxiolytics (Westenberg & den Boer, 1988), and as analgesics (Ogren & Holm, 1980).

The biochemical strategy has the major disadvantage of precluding the discovery of chemically novel modes of treatment. One of the most popular ways of screening for neuroleptic activity, for example, is to test the ability of drugs to antagonize apomorphine-induced stereotyped behaviour. The potency of drugs in this test correlates almost perfectly with their clinical potency in schizophrenia, and also with their ability to bind to dopamine receptors (Creese & Snyder, 1978). But why is it that virtually all drugs effective in the treatment of schizophrenia are dopamine receptor antagonists? It may be that the dopamine synapse is indeed a crucial site for intervention in schizophrenia. However, it seems equally likely that the apparent specificity of neuroleptics for the dopamine receptor simply reflects the degree to which neuroleptic screening tests succeed in identifying

dopamine receptor antagonism (see Ahlenius, this vol., Ch. 12).

If the objective is to discover drugs having a specific clinical indication, then screening tests may attempt to simulate the condition for which drugs are being sought. Many of the screening tests for anxiolytic drugs, for example, embody procedures involving some degree of conflict and a presumption that the animal is anxious (see Stephens & Andrews, this vol., Ch. 3). However, simulation is in no way a requirement in a screening test, and many screening tests have an exclusively empirical foundation; their rationale is simply that agents known to have the desired property are found to alter behaviour in the test. The literature abounds with empirically based screening tests: the muricide test for antidepressant activity has already been mentioned; suppression of apomorphine-induced emesis, one of the classic tests for neuroleptic activity (Niemegeers & Janssen, 1979), is another. The increasingly popular 'drug discrimination' procedures fall within this category. In drug discrimination experiments, animals are rewarded for making one response if they have been injected with a reference drug, and for a different response if they have not; the similarity between a novel agent and the reference drug is then assessed by measuring the degree to which the animals respond to the novel agent as though it were the reference drug (Colpaert & Slangen, 1982). Thus these procedures require the animal to discriminate its internal state. However, at the behavioural level, the consequence of that judgement is always the same: the rat presses one of two levers.

Irrespective of the manner in which they are constructed, screening tests are subject primarily to one very simple requirement: the test should predict accurately the desired activity; it should accept drugs that are effective and reject those that are ineffective. However, if a screen is less than perfect (as most of them are) then the two types of error are not of equal importance. If a screening test accepts an ineffective compound (false positive) the error will eventually come to light in further testing, and no permanent damage will have been done. However, if the test wrongly rejects an effective compound (false negative) a potentially beneficial drug will be lost irretrievably.

When assessing the ability of a screening test to discriminate active from inactive compounds, it

should be borne in mind that failure to discriminate may not necessarily be the fault of the screening test: it is always possible that the clinical effects of a drug have been incorrectly classified. Many antidepressant screens, for example, admit anticholinergic drugs as 'false positives'. But are they really false? Anticholinergics have never been subjected to properly controlled clinical trials, but there are numerous reports from the 1940s and 1950s that they have antidepressant activity (Janowsky & Risch, 1984). Conversely, a 'false negative' could reflect the fact that a drug considered to be clinically effective, actually, is not. A false impression of clinical effectiveness can easily arise in early, uncontrolled clinical trials. What is less frequently recognized is that this impression can continue through into controlled studies, if they compare the new drug against a reference drug that is known to be active. In such a trial, it is possible that the new agent might perform worse than the reference drug, but for a variety of reasons, the difference might not be statistically significant. The new drug is then said to be 'as good as drug X'. However, if a placebo condition had been included, we might have observed that in addition to being 'no worse than drug X', the new drug was also 'no better than no drug at all'. This seems to be what happened in the case of the antidepressant iprindole, which has caused havoc by challenging antidepressant researchers to explain how it works, but actually appears not to be clinically effective in properly designed, placebo-controlled tests (Zis & Goodwin, 1979).

In view of the large number of candidate drugs entered into screening programmes, a screening test should ideally respond to a single administration of an active compound (though it will also be necessary in later testing to confirm that the drug retains its activity following chronic administration). There appears to be a degree of conflict between the logistical requirement that a test respond to acute treatment and the clinical observation that many psychotherapeutic agents act slowly over a period of days, weeks, or even months. However, while this is an important consideration in assessing the validity of a simulation (see Section 1.2.2), it plays no part in the assessment of a screening procedure, which is concerned solely with the ability of the test to make accurate predictions.

Part I: Introduction

In addition to fulfilling the scientific imperative of efficient prediction, and the logistic imperative of speed, a good screening test will also meet a number of other, essentially economic, criteria. It should be cheap, simple, and reliable; it should allow adequate experimental design as well as efficient data processing and statistical analysis; it should not be labour intensive; and it should not rely unduly on the goodwill or sustained concentration of the operator. For several of these reasons, automated procedures have much to recommend them.

1.1.2 Behavioural bioassays

Behavioural bioassays model a physiological action: the whole animal is used as a measuring device to assess the functional state of an underlying physiological system, in exactly the same way that such measurements might be made in an isolated tissue or in a test tube. Examples are the increases in locomotor activity and stereotyped behaviour induced by psychomotor stimulants, which may be used to measure the responsiveness of dopamine receptors in nucleus accumbens and striatum, respectively (Kelly *et al.*, 1975). One use of behavioural assays is as screening tests to identify a specified biochemical action. More commonly, behavioural assays are used to study the mechanisms responsible for changes in brain function, typically those resulting from chronic drug administration, brain lesions, or other similar experimental manipulations.

If the object is to measure a physiological action, why use behavioural methods? There at least are four reasons why indirect behavioural measurement might at times be preferable to a direct biochemical assay. First, behavioural methods are non-destructive. They do not require removal of the brain, which has a number of advantages both for the experimenter and for the experimental subjects. Second, the use of a behavioural measure guarantees that the agents being tested are actually reaching the brain: the poor bioavailability of many novel drugs is only apparent in the whole animal. Third, behavioural measures are functional measures. Many biochemical indices are notoriously unreliable as guides to functional changes. This applies particularly to the parameters derived from receptor binding studies: to

take just one example, the well established super-sensitivity of dopamine receptors in the nucleus accumbens that follows chronic antidepressant administration has proved extremely difficult to confirm using receptor binding methods (Martin-Iverson *et al.*, 1983). Fourth, behaviour integrates the activity of the whole brain. Behavioural measures therefore take into account any corrective changes occurring 'downstream' from the point of measurement. As a result, a behavioural assay may measure a change in brain function less accurately than a biochemical assay, but nevertheless, may give a clearer indication of its functional significance.

Like screening tests, behavioural assays are subject primarily to a single, simple requirement: in this case, that they should measure what they claim to measure. In practice, this may actually be rather difficult to demonstrate, and demonstration is essential, since 'obvious' assumptions have sometimes been shown to be false. It tends to be assumed, for example, that behavioural effects of the 5-HT precursor 5-HTP provide a measure of activity at central 5-HT synapses. It is true that most effects of 5-HTP are indeed mediated by 5-HT, but it can often be demonstrated that the 5-HT systems responsible are actually outside the brain (Carter *et al.*, 1978).

Additionally, if behaviour is being used as a measuring instrument, it is important that some attention be paid to basic principles of measurement (Martin & Bateson, 1986). One issue of obvious importance is the need to ascertain that the behaviour is sensitive to changes in the underlying physiological variable, and is not subject to 'floor' or 'ceiling' effects. A second pertinent issue is the level of measurement achieved: quantitative comparisons require that measurements be at least at the level of ordinal scaling (i.e. rank ordering). For many purposes, interval scaling (i.e. arithmetically linear relationships) would be preferable, and this desideratum is frequently translated into an unjustified assumption that interval scaling has been achieved. Interval scaling is implicit in the use of parametric statistical tests, such as *t*-tests or analysis of variance; if parametric analyses are applied to non-linear data the 'obvious' interpretation of the results may well be incorrect. A monotonic dose-response function, though not essential, is a third desirable feature in a behavioural

assay. The head-twitch response to 5-HT agonists provides a useful measure of (spinal) 5-HT receptor function. However, the response wanes at high doses, with the emergence of other forms of abnormal behaviour (Drust *et al.*, 1979). There is no problem in interpreting an increase in head-twitching as an increase in responsiveness, but a decrease in head-twitching could result either from a decrease in responsiveness or from a large increase, and in published studies the direction of change is not always obvious.

1.1.3 Simulations

Having discussed models of drug action (screening tests) and models of brain function (behavioural assays), it is clear that the term 'animal model of anxiety/mania/obesity/. . .' properly applies to the third class of model, the simulation of human behaviour. In principle, any facet of behaviour is open to simulation. However, simulations of 'normal' behaviour usually take the form of a demonstration that effects first characterized in animals may also be observed in human subjects: examples include classical conditioning (Prokasy, 1965), the control of operant behaviour by schedules of reinforcement (Bradshaw *et al.*, 1976; Davison & Morley, 1987), or the application of animal learning theory to educational practice (Gagne, 1970). It would be more accurate to view this literature as consisting of studies of human models of animal behaviour, rather than the reverse.

While simulations of normal human behaviour are far from rare, simulations of abnormal behaviours predominate, for a variety of social, financial, ethical and legal reasons. An animal model of an abnormal behaviour attempts to simulate a symptom of the disorder, a group of symptoms, or exceptionally, a complete syndrome. Methods of constructing the simulation vary greatly; they include brain damage, selective breeding, selection of extreme individuals, and the application of a variety of factors assumed to be implicated in the etiology of mental disorders, such as stressors, social isolation, or ageing. The object of these manipulations is to produce a behavioural state that can be used as a tool to study aspects of the disorder being modelled. Broadly, four facets of the disorder may be addressed in a model: its

etiology, its treatment, its physiological basis, and the physiological mechanisms underlying its successful treatment. A given model may be appropriate to all or to only some of these problems.

In using an animal model to study a human disorder, the overriding consideration is the validity of the simulation. While it is relatively straightforward to assess the validity of a screening test or a behavioural assay, assessing the validity of a simulation is far more complex. However, the value of data derived from simulations depends crucially on their validity as models of human behaviour. For this reason, the problem of validating simulations will be discussed in some detail.

1.2 ASSESSING THE VALIDITY OF SIMULATIONS OF HUMAN BEHAVIOUR

Models are tools. As such, they have no intrinsic value; the value of a tool derives entirely from the work one can do with it. Conclusions arising from the use of a simulation of abnormal behaviour are essentially hypotheses, that must eventually be tested against the clinical state. An assessment of the validity of a simulation gives no more than an indication of the degree of confidence that we can place in the hypotheses arising from its use. However, the fact that a simulation appears to be valid carries no guarantee that such predictions will be fulfilled. We may feel that a model has a high degree of validity and yet see a prediction falling flat when tested in the real world; on the other hand, a model of apparently low validity may occasionally hit the jackpot. The validity of a simulation is a matter of judgement, rather than measurement.

Against this background, there are a number of yardsticks on which a judgement may be based. They address three broad aspects of a model, from which a picture may be built up of its overall validity. By analogy with the procedures used for validating psychological tests, I have called these three facets predictive validity, face validity and construct validity. Predictive validity means that performance in the test predicts performance in the condition being modelled; face validity means that there are phenomenological similarities between the two; and construct validity means that the model has a sound theoretical rationale (Willner, 1984a, 1986).

Part I: Introduction

Earlier attempts to develop criteria for validating animal models of human behaviour have tended to concentrate largely on the assessment of face validity (McKinney & Bunney, 1969; Abramson & Seligman, 1977). The identification of two further categories reflects two ways in which the literature has developed in recent years. First, there has been a considerable expansion in the literature dealing with the pharmacological exploitation of animal models, much of which contributes to the assessment of predictive validity. Second, there has been significant growth in our understanding of the psychological mechanisms underlying psychopathological states, and examination of construct validity provides a convenient way of bringing animal models into contact with this very relevant literature. The exercise of distinguishing different types of validity has practical value, in that it allows ready identification of areas in which information about a particular model is weak or missing, and ensures that comparisons between different models are made on the basis of comparable data.

1.2.1 Predictive validity

In principle, questions of predictive validity can be addressed to a number of features of simulations, including their etiology and physiological basis. In practice, the predictive validity of the behavioural models used in psychopharmacology is determined primarily by their response to therapeutic drugs. Assessing the predictive validity of a simulation is therefore similar to assessing a screening test: the first question is whether the model discriminates efficiently between those agents that are clinically effective and those that are not. However, there are a number of important differences; in some areas we would ask more of a simulation than of a screening test, while in others we would ask less.

There are three major areas in which a simulation might be expected to go further than a screening test. One is that whereas a screening test is designed to identify drugs that improve the clinical condition, a simulation should respond also to treatments that make the clinical condition worse. It should be possible to show, for example, that the behavioural abnormalities apparent in a simulation of schizophrenia are increased by acute administration of amphetamine or l-dopa, drugs

that exacerbate schizophrenic symptoms (Angrist *et al.*, 1973); this demonstration would have no relevance in a screening test for anti-schizophrenic drugs. A second difference is that a simulation should be responsive to all of the classes of drugs that are useful in treating the disorder, whereas a screening test might embody a strategic decision to look for drugs with particular chemical properties; this would then be recognized as an explicit limitation of the test. At the present time, for example, it makes good strategic sense to search for anxiolytic agents among the benzodiazepines and related compounds, and to design screening tests to minimize their undesirable side-effects; however, a simulation of anxiety might be expected to respond not only to benzodiazepines, but also to beta-blockers and various drugs that interact with the 5-HT system (see Lader, this vol., Ch. 4).

A third effect that should be demonstrable in a simulation, but is superfluous in a screening test, is that the relative potencies of different agents in the model correlate positively with their potencies in clinical use. This is potentially a very powerful test: witness, for example, the almost perfect correlations between the clinical potencies of neuroleptic drugs in schizophrenia, their ability to inhibit apomorphine-induced stereotyped behaviour, and their affinity for dopamine D2 receptors (Creese & Snyder, 1978). However, this test should be applied cautiously, having regard to species differences in drug absorption and metabolism. Furthermore, it can only be applied if drugs are chosen that vary widely in their clinical potency. This condition is satisfied for the neuroleptics, which vary over four orders of magnitude in their clinical potency. It is much less satisfactory for antidepressants, for example, most of which are used clinically within a very narrow dosage range; nevertheless, even in this case, there are sufficient outliers to allow a correlation of potencies to provide useful information (Willner, 1984b).

The major point of divergence between a screening test and a simulation is that while the value of a screening test is totally undermined by a failure to predict efficiently, the presence of false positive or false negative responses would not automatically invalidate a simulation. In discussing screening tests, we have already observed that a lack of agreement between a model and the clinical condition might arise from the incorrect classification

of drugs as clinically active or inactive. False positives or negatives can also arise from species differences in drug kinetics. In particular, species differences in the rate or the route of drug metabolism can have profound effects on the concentration of the drug reaching its site of action; discrepancies arising in this way are especially likely in the case of a drug which has active metabolites. A third source of variability is the regime of drug administration. Clinical trials almost always assess the efficacy of drugs after a period of chronic treatment, and experiments carried out using acute drug administration may not accurately reproduce the appropriate conditions. The development of tolerance on chronic treatment can obliterate a 'false positive' response; conversely, the development of sensitization may overcome a 'false negative'.

In addition to these factors that can apply equally to false positives and false negatives, there are two further considerations. In the case of a false positive, it is important to consider whether a drug that appears to be clinically ineffective might reveal a positive effect if higher doses had been used. In practice, clinical trials at adequate dosage might prove impossible to carry out, owing to the emergence of unacceptable side-effects, which in a simulation would be far less of a deterrent. It is possible, for example, that the very poor performance of cholinergic agonists in dementia, in contrast to their positive effects in many animal models (see Dunnett & Barth, this vol., Ch. 14), might reflect a difficulty in achieving an appropriate dose level in the clinical studies (Summers *et al.*, 1986). False negatives, on the other hand, raise the interesting possibility of clinical subgroups, distinguishable by their differential drug response. Unlike other antidepressants, for example, monoamine oxidase inhibitors fail to normalize behaviour in olfactory bulbectomized rats; some workers have suggested that this may be related to the fact that these drugs are useful clinically only in an atypical group of patients (Jesberger & Richardson, 1985).

It should by now be clear that the process of validating an animal model cannot proceed mechanistically to a yes/no answer. A model may have predictive validity despite failing to discriminate efficiently between agents that are clinically active and those that are not. It is a matter of judgement to decide on the significance

of discrepancies between drug effects in the model and in the clinical state. Inefficient prediction may not be incompatible with predictive validity, provided that we can understand the origin of the discrepancy.

1.2.2 Face validity

Face validity refers to a phenomenological similarity between the model and the disorder it simulates: on one hand, the model should resemble the disorder, while on the other, there should be no major dissimilarities. In a seminal paper that for the first time endowed the topic of modelling psychopathology in animals with a degree of scientific respectability, McKinney & Bunney (1969) suggested that to be valid, a model should resemble the condition it models in four respects: etiology, symptomatology, treatment, and physiological basis. Though admirable in principle, this requirement is in fact too stringent. In practice, we cannot require a simulation to correspond in every particular with the disorder it simulates, for the very good reason that there are major gaps in our understanding of the disorders themselves. It is not reasonable, for example, to require similarity of etiology in a simulation of schizophrenia, when the etiology of schizophrenia is virtually a closed book. Indeed, the main objective in setting up a simulation is precisely to fill out the missing pages.

We can, however, reasonably ask that a simulation be compared with the disorder it models in those areas where we do understand something of the disorder. A corollary to this position is that in areas where more is known about the disorder, we should examine the model more carefully for the degree of phenomenological similarity. A second corollary is that as the understanding of a disorder develops, criteria for evaluating the face validity of an animal model will automatically change. Where facts have been established about the etiology of a disorder or its physiological basis, they should obviously be included in an assessment of face validity. In general, however, the etiology and physiological basis of psychiatric disorders are poorly understood; assessment of face validity will usually be based primarily on symptomatology, about which most is known, and to some extent, on treatment.

A number of relevant questions may be asked

about treatment, additional to those issues raised under the heading of 'predictive validity'. Do abnormalities in the model respond to effective behavioural modes of treatment, where they exist? Are drug effects in the model achieved at reasonable doses, that produce tissue concentrations of the drug comparable to those found clinically? However, the most important of these questions concerns the treatment regime. Most prescriptions of psychotropic drugs envisage that the drug will be taken regularly over an extended period of weeks or months, and consequently, the drugs must continue to exert their effects after chronic treatment. It is therefore essential to establish that drug effects in the model are also demonstrable after a period of chronic administration. It is also necessary to establish that therapeutic effects are actually present while the drug is being administered: it is often expedient, following chronic drug treatment, to precede testing by a period of drug 'wash-out', and in these circumstances confusion between drug effects and withdrawal symptoms can easily arise (Noreika *et al.*, 1981; see also Section 1.3 below).

A related issue is less straightforward. In clinical use, not only must drugs continue to act after chronic treatment, but also, their onset of action is frequently delayed. This fact has been used to argue against the validity of models in which therapeutic effects can be demonstrated using acute drug treatment. The problem with accepting this argument is that delayed onsets of clinical action are poorly understood, and may in part reflect a delay in delivering an adequate concentration of drug to the brain (see, e.g., Willner, 1989). As with so many of the guidelines for validating animal models, an acute drug effect, in a model of a disorder that appears to require chronic treatment, may or may not indicate a lack of validity.

The central activity in an assessment of face validity is likely to be a comparison of symptoms displayed in the disorder with behavioural and other abnormalities apparent in the model. Since most disorders are syndromes that include a number of different symptoms, it should usually be possible to make multiple comparisons between the disorder and the model; the symptom check-list approach adopted in the DSM-III system of diagnosis provides a useful starting point for this exercise (American Psychiatric Association, 1987). In

practice, models often tend to focus on a single behaviour. In this case, the significant question is that of its centrality in the disorder; again, the DSM-III distinction between essential symptoms and optional extras may be helpful. However, the question of centrality cannot be answered fully at the level of phenomenology. Even though a reduction in food intake is the essential feature of anorexia nervosa, it remains a moot point whether anorexia is properly described as a disorder of feeding or a disorder of personality (see Montgomery, this vol., Ch. 8); if the latter, then simulations based on a reduction in food intake may be missing the target. A related question is the specificity of symptoms for a particular disorder (i.e. their value in differential diagnosis); clearly, a model will be less valid if the symptoms it simulates are common to a variety of disorders (Abramson & Seligman, 1977).

Where a number of points of similarity may be adduced between a model and the disorder it simulates, it becomes necessary to ask whether the identified cluster of symptoms forms a coherent grouping that might realistically be seen in a single patient, or whether they are drawn from a variety of diagnostic subgroupings. It is important in this context not to rely too heavily upon the diagnostic categories refined in DSM-III and related systems, since these categories were derived consensually rather than empirically, and themselves stand in need of validation (see, e.g., Carroll, 1983). Simulations offer a way of relating clusters of symptoms to specific etiological factors, specific treatment modalities, and specific physiological abnormalities. It remains an optimistic possibility that through this process, valid simulations of psychiatric disorders might themselves contribute to the clearer definition of diagnostic boundaries.

1.2.3 Construct validity

There is one important natural limitation to the attempt to establish face validity by mapping a point-to-point correspondence between a disorder and an animal model: there is no good reason to suppose that a given condition will manifest itself in identical ways in different species (Hinde, 1976) and, in fact, there are many obvious instances of divergence. For example, rearing on the hind legs is a prominent component of stimulant-induced

stereotyped behaviour in rats but not in primates, while the reverse is true of scratching (Randrup & Munkvad, 1970). The physical topography of these behaviours is quite different; nevertheless, we are able to say that they are homologous across species; the judgement of homology arises from an understanding that the two behaviours have the same physiological substrate. Patterns of maternal behaviour, which vary widely across species, provide another example of homology. In the case of a human mother cradling her baby and a female rat retrieving her pups, the judgement of homology arises primarily from an understanding of the psychosocial environment in which the behaviours take place. In both instances, the decision that different behaviours in different species reflect different manifestations of a similar underlying process is based upon a theoretical rationale, which derives from consideration of factors other than the behaviours themselves.

The theoretical rationale behind a model requires evaluation, and a satisfactory outcome endows the model with construct validity. In advancing this concept it is assumed implicitly that it is possible to construct theories of psychopathology that have some application to species other than people. As this has in the past been a source of some confusion (exemplified by the extreme difficulty of defining what might be meant by a schizophrenic rat) it may be helpful to consider briefly the general shape of psychopharmacological theories.

I have argued elsewhere (Willner, 1984c, 1985) that it is inadequate for a psychopharmacological theory simply to describe a link between a set of biochemical events (e.g. antagonism of benzodiazepine receptors) and a change in human experience (relief of anxiety). An adequate theory must also interpose an account at two intermediate levels. At the first, it must explain the consequences of the identified biochemical changes for functional activity within the brain (which pathways are affected, and how), while at the second, it must explain how these changes in functional activity affect the brain's ability to process information (which requires some understanding of the role of the affected systems in the cognitive activity of the normal brain). It then becomes necessary to consider, in addition, the manner in which those underlying cognitive changes are incorporated into

subjective experience, and so to arrive at a proper understanding of the experiential effects of a drug.

Having distinguished between these various levels of analysis we are now in a position to consider which aspects of a theory of psychopathology may be addressed in animals, and which may not. For most practical purposes, we do not possess experimental tools that will allow us to address queries to animals about their subjective state; the questions we can ask them are strictly limited by the necessity that their replies must be behavioural rather than verbal. Even in the case of the drug discrimination procedure, the data from human experiments caution against equating internal states with subjective states (Chait *et al.*, 1986; see also Goudie, this vol., Ch. 17). We do not, therefore, attempt to simulate the experiential aspects of drug action or of psychopathology. What we are attempting in a simulation is to model the constraints on information processing and behaviour that underly the experiential changes.

From this perspective, the assessment of construct validity is a three-stage process (Willner, 1986). The first step is to identify the behavioural variable that is being modelled; the second is to assess the degree of homology between the identified variable and behaviour in the simulation; the third is to assess the significance of the identified variable in the clinical picture.

Each of these stages is fraught with difficulty. It might, for example, seem a straightforward matter to decide what is being modelled, and this may indeed be so in some cases, where the behavioural objectives of the model are specified in advance. Many models of dementia, for example, embody an explicit choice to simulate specific disorders of learning and memory (see Dunnett & Barth, this vol., Ch. 14). However, there are many other starting points for a simulation, such as the application of a supposed etiological factor, or the use of drugs to induce a supposedly relevant physiological state; the behavioural state simulated may then be shrouded in obscurity. It remains a total mystery what aspect of depression was being addressed by experiments in which monkeys were isolated in the dark, in a vertical cylinder nicknamed 'the well of despair' (Harlow & Suomi, 1971).

Having established what is being modelled, the next step is to assess whether the two behaviours are homologous. However, this can only be done if

Part I: Introduction

both behaviours are well understood, and frequently they are not. The learned helplessness model of depression provides an excellent example of this problem. In this model, prior exposure to uncontrollable aversive events causes a subsequent learning deficit; this effect may be demonstrated very readily in both rats and people (Miller *et al.*, 1977). However, despite a very substantial research effort, the nature of the underlying behavioural mechanisms remains unclear. The explanation originally proposed was that subjects exposed to uncontrollable events learn that events are uncontrollable: a homologous process in rats and people (Seligman, 1975). However, in rats there is evidence that motor disabilities and stress-induced analgesia contribute to the behavioural deficits (Jackson *et al.*, 1978, 1979), while in people helplessness effects are sensitive to very minor changes in the experimental procedure (Buchwald *et al.*, 1978). Returning to the original examples of maternal behaviour and drug-induced stereotypies, we are able to see that for two behaviours to be homologous, they should share a similar physiological basis and occur in a similar behavioural context. In learned helplessness, the growth of information about the two behaviours has, if anything, reduced the likelihood of homology.

The final step in the assessment of construct validity is an evaluation of the significance in the overall clinical picture of the behaviour modelled by the simulation. Many animal models of drug dependence, for example, focus on the rewarding properties of drugs. However, while rewarding effects are of undoubted importance in the early stages of drug use, some theoretical formulations emphasize the fear of withdrawal as a major factor in dependence (see Goudie and Hartnoll, this volume, Chapters 17 and 19). If this is correct, then models that simulate rewarding effects may be of only limited relevance. A similar problem concerning the significance of low food intake in anorexia nervosa has already been remarked: the anorectic's low food intake should not be mistaken for a reduced appetite for food (see Montgomery, this vol., Ch. 8). Because debates of this kind take place entirely within the human literature, these problems do not figure prominently in discussions of animal models. However, it is obviously of some considerable importance to know whether a simulation is asking the correct question.

Of the three roads to validity, construct validity is the most fundamental: while a model may, for a variety of reasons, fail to meet criteria for predictive or face validity and still survive, it would be difficult to retain confidence in a model whose theoretical rationale had been exploded. Construct validity is the most difficult aspect of a model to establish, but also, the most challenging.

1.3 USING MODELS

Of course, while it is important to distinguish clearly between screening tests, behavioural assays and simulations, in practice the three types of model interact, in a variety of ways. One obvious point at which models of different types meet is through the cross-talk in their scientific development. Screening tests, for example, are frequently developed as offshoots of attempts to simulate a disorder, or as a result of insight into the physiological mechanisms revealed by a simulation. This process can also work in the opposite direction: the early neuroleptic screening tests threw up a range of active agents that were later discovered to be dopamine receptor antagonists; as a result, later screening tests were based on bioassays for dopamine antagonism, and chronic overstimulation of dopamine receptors forms the basis of some current attempts to simulate schizophrenia (see Lyon, this vol., Ch. 11).

A second significant point of intersection is that the same procedure may be used for two quite different purposes. In such cases, assessing the model's suitability for its two uses might involve radically different sets of criteria. The use of the same procedure either as a screening test or as a bioassay has already been discussed. If the procedure is being used as an assay then it must have good metric properties; however, if it is being used as a screening test all we ask is that it should discriminate accurately between agents that are active and those that are inactive. Similarly, many procedures may serve either as screening tests or as simulations. The suppression of lever pressing by electric shock, for example (Geller & Seifter, 1960), may be used either as a simulation of anxiety, or as a screening test for anxiolytic drugs (see Green & Hodges, and Stephens & Andrews, this vol., Ch. 2, 3). Again, the important point is that the criteria employed to decide whether or not this should be

considered a 'good' model depend upon the purpose for which it is being used. If the Geller-Seifter conflict procedure is being used as a screening test, then its validity as a simulation is irrelevant: the only requirement is that it should efficiently predict anxiolytic activity. However, if the test is being used to investigate the brain mechanisms of anxiety, then the overriding concern is with its validity as a simulation; its predictive power is relatively unimportant since, as discussed above, there are a number of factors that might turn a valid simulation into an inefficient predictor of clinical activity.

A third point of contact is that two different types of model may be used within a single experiment. This applies particularly to the simultaneous use of a simulation and a behavioural assay. A case study will serve to illustrate both the development of a behavioural assay and its use in conjunction with a simulation procedure. The problem addressed in these experiments was whether antidepressant drugs, chronically administered, reduce the sensitivity of presynaptic dopamine autoreceptors, and whether this effect contributes to their clinical action; the literature was equivocal with respect to the first question and silent on the second. The model used to assay the responsiveness of dopamine autoreceptors was the suppression of food intake by a low dose of the dopamine agonist apomorphine. To validate this model we demonstrated that the response was blocked by centrally acting dopamine receptor antagonists, but not by a dopamine antagonist that does not enter the brain, or by a variety of other types of drug; that the response could be elicited by administering apomorphine directly to the ventral tegmental area of the midbrain, where the cells of origin of the mesolimbic dopamine system are located; and that administration of a dopamine antagonist to the ventral tegmental area blocked the effect of systemically administered apomorphine (Willner *et al.*, 1985; Muscat *et al.*, 1986; Towell *et al.*, 1986a). Then we examined the effect of apomorphine on food intake during chronic treatment with antidepressant drugs, and found that there was indeed a reduced response, indicative of dopamine autoreceptor subsensitivity. However, the effect could only be demonstrated following withdrawal from the antidepressants (Towell *et al.*, 1986b). Could it be relevant to the clinical action? To

answer this, we turned to a simulation of depression. In this model, which simulates the cardinal symptom of endogenous depression, the inability to experience pleasure, rats are exposed chronically to a variety of mild stressors; over a period of weeks, they gradually lose their preference for a highly rewarding sweet solution (Willner *et al.*, 1987; see also Willner, this vol., Ch. 5). On administering apomorphine, to measure dopamine autoreceptor sensitivity, we found that the attenuated response typical of withdrawal from antidepressants was also present in the stressed animals. Antidepressant administration restored normal preference behaviour in the stressed animals, but caused no further desensitization of dopamine autoreceptors. We therefore concluded that this effect is not, in fact, one of the mechanisms by which antidepressants exert their clinical action (Muscat *et al.*, 1988).

I have presented these experiments in some detail as relatively few studies have exploited the possibilities of using behavioural assays within the context of a simulation procedure. However, studies of this kind must inevitably increase in the future. Numerous effects of psychotherapeutic drugs on brain function have been described, but almost invariably in 'normal' animals. Before accepting that such effects are responsible for clinical improvement, it will be necessary in every case to demonstrate that the effect is also present in simulations of the disorder (or in clinical trials). As the above example shows, there is no guarantee of success in the attempt to demonstrate the clinical relevance of drug effects described in 'normal' animals.

1.3.1 Models and their users: three caricatures

Before moving from general principles to a more detailed examination of the behavioural models used in specific areas of psychopharmacology, we need to ask two further questions. Who uses behavioural models, and why? Unlike many disciplines, psychopharmacology contains not one, but three scientific communities: academics, industrial pharmacologists, and clinicians. These three groups relate to the discipline in different ways; they have different objectives, which give rise to different scientific priorities, and lead them to different uses in the available behavioural tools.

Part I: Introduction

The major concern of an academic psychopharmacologist is (or should be) the construction, development and evaluation of theories. I have already discussed what is required of a theory in psychopharmacology. To reiterate briefly, an adequate theory of psychotropic drug action must provide a coherent account of how the drug interacts with the brain, the specific anatomical sites of action, the resulting physiological changes, the consequent changes in cognitive and behavioural abilities, and the relationship between cognitive change and subjective experience. The job of the academic psychopharmacologist is to provide explanations in some or all of these domains: to elucidate the biochemical, physiological and psychological mechanisms that underly psychopathological behaviour and its treatment by psychotherapeutic drugs. The major behavioural tool deployed to these ends is the simulation, and assessment of the validity of simulations therefore falls within this area of work. Behavioural assays are also used extensively, in the investigation of physiological mechanisms. However, the academic psychopharmacologist has very little scientific interest in screening tests, whose only contribution to the development of explanatory theory is that new drugs often prove to be of value as experimental tools.

The industrial psychopharmacologist, on the other hand, is concerned primarily to discover new drugs. The screening test is therefore the major behavioural tool for this group of workers. They also make heavy use of behavioural assays, both as screening tests and as tools to investigate the physiological actions of clinically active drugs. However, industrial psychopharmacologists are interested in simulations (and their validity) only to the extent that a valid simulation might serve as the starting point from which to develop a superior screening test.

The major concern of clinical psychopharmacologists is to improve the care of their patients. Their involvement with animal models is usually less direct than that of the other two groups; they do not themselves use the models, but rather, apply and test clinically the fruits of others' work. Clinicians are obviously interested in simulations, since the main object of a simulation is to throw light on the nature of the disorder. However, because clinical psychopharmacologists are pri-

marily clinicians and only secondarily pharmacologists, unlike the other two groups they have no commitment to drugs, if other forms of treatment or prevention should prove more effective. Their interest is primarily in the behavioural aspects of a model rather than its physiological mechanisms: many of the behavioural techniques in psychotherapy derived initially from simulations (Eysenck & Martin, 1987). Even though their contact with simulations is somewhat distant, the interests of clinicians extend more widely than that of the other groups, to include areas in which there are behavioural simulations of disorders that are not treated pharmacologically (Keehn, 1986). Clinical psychopharmacologists are also interested in screening tests, and look to them to provide compounds that will be more efficacious than existing drugs, and/or have fewer side-effects. However, as clinicians, they have little interest in using behavioural models as assay systems. Information of this kind becomes important only when it leads to new treatments or to diagnostic improvements.

Obviously, these sketches of the academic, the industrial, and the clinical psychopharmacologist are grossly oversimplified caricatures, and many psychopharmacologists do not fit neatly into these stereotypes. Many industrial pharmacologists are academics at heart, whose wish to contribute to the development of scientific theory is at least the equal of their desire to develop new products; some pharmaceutical companies are sufficiently enlightened to recognize this as a legitimate aspiration, and make appropriate provision for 'pure' research. Equally, many academics are attracted by the possibility of making an immediate practical impact in the real world, by developing screening tests for the pharmaceutical industry. Scientific collaboration between the academic and industrial sectors is highly developed, to the benefit of both: industry takes theoretical advances and uses them to develop both new products and new routes to product development; conversely, novel agents produced by the industry are among the most powerful analytical tools available to conduct 'pure' research. Clinical psychopharmacology plays a central role in guiding preclinical research, through the essential function of testing the hypotheses arising from the use of animal models; at the same time, the agenda for clinical psychopharma-

cology is largely determined by preclinical research, which identifies the physiological questions to be asked in the clinic and provides the novel agents that form the subject matter of clinical trials. Furthermore, there is considerable movement of personnel, at all levels, between the academic world and industry; the adoption of a new persona appropriate to the new institutional context can be a long drawn-out process. In addition, many clinicians take time out from their clinical responsibilities to do basic research, and in so doing slip into the other two roles.

1.3.2 Theoretical, industrial and clinical perspectives

Despite the blurring of distinctions between the three sections of the psychopharmacological community, it is important to recognize that their interests are not identical, and that these differences are reflected in the direction of research and in the choice of experimental methodologies. In this book, we have tried for the first time to identify the particular perspectives of academic, industrial and clinical psychopharmacologists working in different areas of psychopharmacology. In the six areas covered, we have attempted to examine the ways in which the three perspectives interact, and to clarify the contribution of behavioural models to progress in each of the three domains.

Of the three caricatures outlined above, that of the academic psychopharmacologist is probably the most accurate: the chapters that follow demonstrate a remarkable creativity in generating new ideas for modelling psychiatric disorders in animals. However, in most of the areas surveyed the emphasis has been largely on the development of models, rather than their use to probe the mechanisms underlying the disorders modelled. In large part, this reflects the sad fact that while great vigour has been applied to the development of a wide variety of models, many have little to recommend them. Nevertheless, it is clear that the framework outlined in this chapter (Section 1.2) provides a straightforward means of assessing the strengths and weaknesses of a model; it is also clear that in some instances, converging guidance from a number of models can be informative despite the weakness of each model individually. Thus, while the actual contribution of simulations to our

understanding of psychiatric disorder has been limited, their potential is considerable.

The least accurate of the three sketches is that of the industrial psychopharmacologist. It is apparent in reading accounts of the screening process that industrial psychopharmacologists do in fact share the academic preoccupation with the validity of models as simulations of psychiatric disorders. It is instructive to examine why this should be, given that the theoretical perspective is not a necessary feature of a screening programme. One reason is more political than scientific: the development of a new drug represents a major investment of resources, which can more easily be justified to the company decision makers by evidence derived from a test that has face validity, and so makes intuitive sense (Broekkamp, personal communication). However, there are also important considerations within the scientific context of drug development, which broadly reflect the belief that a better understanding of the disorder may lead to new avenues for drug development. For some disorders, such as dementia or obesity, there are no well established treatments, and therefore there is no basis for an empirical predictive test. In these cases the only alternative to seeking valid simulations is to take promising compounds directly into the clinic; this is a high-risk strategy on which few companies would wish to rely exclusively. At the other extreme, the established antidepressants have a great variety of biochemical actions, which have formed the basis for the development of a bewildering variety of novel agents that may have antidepressant properties (see Danysz *et al.*, this vol., Ch. 6). The number of potential antidepressants undergoing clinical trials has therefore seen rapid expansion. An unforeseen consequence is that as we do not at present know which of the new agents work and which do not, it has become impossible to assess with any degree of accuracy how well the available screening tests succeed in predicting clinical activity. For other disorders, most notably schizophrenia, a particular treatment modality is so successful that drug development is based primarily around particular chemical structures with specified biochemical actions. However, this strategy by its nature precludes the discovery of compounds acting through different mechanisms. The example of anxiolytic drugs illustrates perfectly the dangers of this preemptive reliance on medicinal

Part I: Introduction

chemistry: no sooner had screening programmes settled down to searching for compounds active at the benzodiazepine receptor, than several new families of potential non-benzodiazepine anxiolytics were discovered, that interact in various ways with 5-HT receptors (see Part 2). The fact that these new agents differ significantly from benzodiazepines, and from one another, in traditional anxiolytic-sensitive behavioural tests, highlights an urgent need to clarify which aspects of anxiety the different tests model.

It is clear from this discussion that the quest for valid simulations is actually central to industrial psychopharmacology, since such models may be used to identify desirable biochemical properties. However, as discussed earlier (Section 1.1), screening tests and simulations have very different requirements. This essential tension at the heart of industrial psychopharmacology is most commonly resolved by an uneasy compromise: a layered screening programme in which rough and ready primary screens are followed by tests of greater complexity that approximate more closely to valid simulations.

The clinical perspective on behavioural models, in reality, is twofold. On the one hand, clinicians view animal models of psychiatric disorders with a healthy scepticism, and provide a valuable critical service in pointing out their many shortcomings, when measured against the complexities of human behaviour. On the other hand, the literature generated from animal models encourages clinicians to confront their own limitations, and these are many. Psychiatric diagnosis is to a large extent descriptive and consensual, and there are few indications that this situation is likely to improve in the near future. Ideas of the etiology of psychiatric disorders are still in their infancy (note, for example, the pervasive role of 'stress'). Even treatment, the jewel in the crown of biological psychiatry, remains largely empirical. The absence of a sound theoretical basis is illustrated by the way in which treatments find application outside their original indications: for example, antidepressant drugs as anxiolytics or anorectics, and neuroleptic drugs as anxiolytics or antidepressants. (Consider also that mild dementias may be cured by treating an underlying depression, but depression is exacerbated by cholinergic agonists, the current front-runners in the race to find an anti-dementia drug.)

As one of the 'clinical' contributors to this volume notes, because behavioural models are explicitly related to a broader body of theory, they fulfill a valuable function in forcing clinicians to examine critically their own assumptions. It is all too easy, from a clinical standpoint, to denigrate animal models of psychiatric disorders. However, it could equally be seen as an area of strength that the models lend themselves so readily to critical analysis. Ultimately, the major contribution of animal models may well be that they encourage clinicians to apply a similar rigour to the analysis of disordered human behaviour.

REFERENCES

- Abramson, L.Y. & Seligman, M.E.P. (1970). Modelling psychopathology in the laboratory. History and rationale. In: J.D. Maser & M.E.P. Seligman (eds.), *Psychopathology: Animal Models*, pp. 1-26. Freeman: San Francisco.
- American Psychiatric Association (1987). *Diagnostic Criteria from DSM-III-R*. APA, Washington D.C.
- Angrist, B., Sathananthan, G.S. & Gershon, S. (1973). Behavioural effects of L-dopa in schizophrenic patients. *Psychopharmacologia* **31**, 1-12.
- Bergman, I., Brane, G., Gottfries, C.G., Jostell, K.-G., Karlsson, I. & Svennerholm, L. (1983). Alaproclate: A pharmacokinetic and biochemical study in patients with dementia of the Alzheimer type. *Psychopharmacology* **80**, 279-83.
- Bradshaw, C.M., Szabadi, E. & Bevan, P. (1976). Behaviour of humans in variable interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior* **27**, 275-9.
- Buchwald, A.M., Coyne, J.C. & Cole, C.S. (1978). A critical evaluation of the learned helplessness model of depression. *Journal of Abnormal Psychology* **87**, 180-93.
- Carroll, B.J. (1983). Neurobiologic dimensions of depression and mania. In: J. Angst (ed.), *The Origins of Depression: Current Concepts and Approaches*, pp. 163-86. Springer: Berlin.
- Carter, R.B., Dykstra, L.A., Leander, J.D. & Appel, J.B. (1978). Role of peripheral mechanisms in the behavioral effects of 5-hydroxytryptophan. *Pharmacology, Biochemistry and Behavior* **8**, 249-53.
- Chait, L.D., Uhlenluth, E.H. & Johanson, C.E. (1986). The discriminative stimulus and subjective effects of phenylpropanolamine, mazindol and d-amphetamine in humans. *Pharmacology, Biochemistry and Behavior* **24**, 1665-72.
- Colpaert, F. & Slangen, J.L. (1982). *Drug*

1. Behavioural models in psychopharmacology

- Discrimination: Applications in CNS Pharmacology*. Elsevier: Amsterdam.
- Creese, I. & Snyder, S.H. (1978). Behavioral and biochemical properties of the dopamine receptor. In: M.A. Lipton, A. DiMascio & K.F. Killam (eds.), *Psychopharmacology: A Generation of Progress*, pp. 377–88. Raven Press: New York.
- Davison, M. & Morley, D. (1987). *The Matching Law: A Research Review* Lawrence Erlbaum: Hillsdale, N.J.
- Drust, E.G., Sloviter, R.S. & Connor, J.D. (1979). Effect of morphine on 'wet dog' shakes caused by cerebroventricular injections of serotonin. *Pharmacology* **18**, 299–305.
- Eysenck, H.J. & Martin, I. (1987). *Theoretical Foundations of Behaviour Therapy*. Plenum Press: New York.
- Gagne, R.M. (1970). *Conditions of Learning*. Holt, Reinhart & Winston: New York.
- Geller, I. & Seifter, J. (1960). The effects of meprobamate, barbiturates, D-amphetamine and promazine on experimentally induced conflict in the rat. *Psychopharmacologia* **1**, 482–92.
- Harlow, H.F. & Suomi, S.J. (1971). Production of depressive behaviours in young monkeys. *Journal of Autism and Child Schizophrenia* **1**, 246–55.
- Hinde, R.A. (1976). The uses of similarities and differences in comparative psychopathology. In: G. Serban & A. Kling (eds.), *Animal Models in Human Psychobiology*, pp. 187–202. Plenum Press: New York.
- Horovitz, Z.P. (1965). Selective block of mouse killing by antidepressants. *Life Science* **4**, 1909–12.
- Jackson, R.L., Maier, S.F. & Rapoport, P.M. (1978). Exposure to inescapable shock produces both activity and associative deficits in rats. *Learning and Motivation* **9**, 69–98.
- Jackson, R.L., Maier, S.F. & Coon, D.J. (1979). Long-term analgesic effect of inescapable shock and learned helplessness. *Science* **206**, 91–3.
- Janowsky, D.S. & Risch, S.C. (1984). Cholinomimetic and anticholinergic drugs used to investigate an acetylcholine hypothesis of affective disorders and stress. *Drug Development Research* **4**, 125–42.
- Jesberger, J.A. & Richardson, J.S. (1985). Animal models of depression: Parallels and correlates to severe depression in humans. *Biological Psychiatry* **20**, 764–86.
- Katz, R.J. (1981). Animal models and human depressive disorders. *Neuroscience and Biobehavioral Reviews* **5**, 231–46.
- Keehn, J.D. (1986). *Animal Models for Psychiatry*. Routledge: London.
- Kelly, P.H., Seviour, P.W. & Iversen, S.D. (1975). Amphetamine and apomorphine responses in the rat following 6-OHDA lesions of the nucleus accumbens septi and corpus striatum. *Brain Research* **94**, 507–22.
- Martin, P. & Bateson, P. (1986). *Measuring Behaviour: An Introductory Guide*. Cambridge University Press: Cambridge.
- Martin-Iverson, M., Leclere, J.F. & Fibiger, H.C. (1983). Cholinergic-dopaminergic interactions and the mechanisms of action of antidepressants. *European Journal of Pharmacology* **94**, 193–201.
- McKinney, W.T. (1984). Animal models of depression: An overview. *Psychiatric Developments* **2**, 77–96.
- McKinney, W.T. & Bunney, W.E. (1969). Animal model of depression: Review of evidence and implications for research. *Archives of General Psychiatry* **21**, 240–8.
- Miller, W.R., Rosellini, R.A. & Seligman, M.E.P. (1977). Learned helplessness and depression. In: J.D. Maser & M.E.P. Seligman (eds.), *Psychopathology: Animal Models*, pp. 104–30. Freeman: San Francisco.
- Muscat, R., Willner, P. & Towell, A. (1986). Apomorphine anorexia: A further pharmacological characterization. *European Journal of Pharmacology* **123**, 123–31.
- Muscat, R., Towell, A. & Willner, P. (1988). Changes in dopamine autoreceptor sensitivity in an animal model of depression. *Psychopharmacology* **94**, 545–50.
- Niemegeers, C.J.E. & Janssen, P.A.J. (1979). A systematic study of the pharmacological activities of dopamine antagonists. *Life Science* **24**, 2201–16.
- Noreika, L., Pastor, G. & Liebman, J. (1981). Delayed emergence of antidepressant efficacy following withdrawal in olfactory bulbectomized rats. *Pharmacology, Biochemistry and Behavior* **15**, 393–8.
- Ogren, S.-O. & Holm, A.-C. (1980). Test-specific effects of the 5-HT uptake inhibitors alaproclate and zimelidine on pain sensitivity and morphine analgesia. *Journal of Neural Transmission* **47**, 253–71.
- Prokasy, W.F. (ed.) (1965). *Classical Conditioning: A Symposium*. Appleton, Century, Crofts: New York.
- Randrup, A. & Munkvad, I. (1970). Biochemical and psychological investigations of stereotyped behaviour. In: E. Costa & S. Garattini (eds.), *Amphetamines and Related Compounds*, pp. 695–713. Raven Press: New York.
- Seligman, M.E.P. (1975). *Helplessness: On Depression, Development and Death*. Freeman: San Francisco.
- Squires, R.F. & Braestrup, C. (1977). Benzodiazepine receptors in the brain. *Nature* **266**, 732–4.
- Summers, W.K., Majovski, L.V., Marsh, G.M., Tachiki, K. & Kling, V. (1986). Oral tetrahydroaminoacridine in long-term treatment of

Part I: Introduction

- senile dementia. *New England Journal of Medicine* **315**, 1241-5.
- Towell, A., Muscat, R. & Willner, P. (1986a). Apomorphine anorexia: The role of dopamine cell body autoreceptors. *Psychopharmacology* **89**, 65-8.
- Towell, A., Willner, P. & Muscat, R. (1986b). Behavioural evidence for autoreceptor subsensitivity in the mesolimbic dopamine system during withdrawal from antidepressant drugs. *Psychopharmacology* **90**, 64-71.
- Westenberg, H.G.M. & den Boer, J.A. (1988). Serotonin uptake inhibitors and agonists in the treatment of panic disorder. *Psychopharmacology* **96** (Suppl.), 56.
- Willner, P. (1984a). The validity of animal models of depression. *Psychopharmacology* **83**, 1-16.
- Willner, P. (1984b). The ability of antidepressant drugs to desensitize beta-adrenergic receptors is not correlated with their clinical potency. *Journal of Affective Disorders* **7**, 53-8.
- Willner, P. (1984c). Drugs, biochemistry and subjective experience: Towards a theory of psychopharmacology. *Perspectives in Biology and Medicine* **28**, 49-64.
- Willner, P. (1985). *Depression: A Psychobiological Synthesis*. Wiley: New York.
- Willner, P. (1986). Animal models of human mental disorders: Learned helplessness as a paradigm case. *Progress in Neuropsychopharmacology and Biological Psychiatry* **10**, 677-90.
- Willner, P. (1989). Sensitization to the actions of antidepressant drugs. In: M.W. Emmett-Oglesby & A.J. Goudie (eds.), *Psychoactive Drugs: Tolerance and Sensitization*, pp. 407-59. Humana: Clifton, N.J.
- Willner, P., Towell, A. & Muscat, R. (1985). Apomorphine anorexia: A behavioural and neuropharmacological analysis. *Psychopharmacology* **87**, 351-6.
- Willner, P., Towell, A., Sampson, D., Sophokleous, S. & Muscat, R. (1987). Reduction of sucrose preference by chronic mild unpredictable stress, and its restoration by a tricyclic antidepressant. *Psychopharmacology* **93**, 358-64.
- Zis, A.P. & Goodwin, F.K. (1979). Novel antidepressants and the biogenic amine hypothesis of depression. The case for iprindole and mianserin. *Archives of General Psychiatry* **36**, 1097-107.