



**HAL**  
open science

## Extrapolation from Animal Model of Depressive Disorders: What's Lost in Translation?

Mael Lemoine

► **To cite this version:**

Mael Lemoine. Extrapolation from Animal Model of Depressive Disorders: What's Lost in Translation?. Wakefield, Jerome C; Demazeux, Steeves. Sadness or Depression? International Perspectives on the Depression Epidemic and Its Meaning, Springer, pp.157-172, 2015, 10.1007/978-94-017-7423-9\_11 . halshs-01270260

**HAL Id: halshs-01270260**

**<https://shs.hal.science/halshs-01270260>**

Submitted on 6 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Extrapolation from animal model of depressive disorders:  
what's lost in translation?**

**Maël Lemoine**

University of Tours, INSERM U930, IHPST (Paris)

**Abstract** Animal models of depression are problematic and results drawn from them is moderately convincing. The main problem, it is often argued, is that it is impossible to model a mental disorder, i.e. specifically human, in animals like rodents: it is a matter of resemblance of symptoms. Yet in this field it is generally assumed that animal models of depression are more or less 'valid' according to three criteria: predictive, construct, and face validity, with only the latter concerned with the resemblance of symptoms. It is argued here that the problem is actually not with resemblance to the clinical features or to the factors of depression: it is not their being *mental* parameters. It lies, rather, in the fuzziness of the definition of a human entity and in the difficulty of linking together supposedly involved biological mechanisms into a consistent picture of the underlying process of the disease. It is therefore not that we cannot model what we know to be depression, it is rather that we do not know *what* to model.

**Keywords:** translational medicine, depression, validity, model organism, mental disorder.

**Acknowledgments** Professors C. Belzung, neurobiologist, and V. Camus, psychiatrist, have read and discussed the paper. Thanks to them for having always been an invaluable help since the beginning of my philosophical work in their team.

### **Introduction: translational research and extrapolation in psychiatry**

Philosophers tend to import their own problems into foreign domains, not always for the sake of the greater good. An alternative strategy in the philosophy of science consists in trying to illuminate problems as scientists encounter them. Modeling mechanisms of diseases in organisms belongs to what is called ‘translational research’. As a matter of fact, in contemporary medicine, ‘translational research’ is summarized through the rhetorical motto ‘from the bench to the bedside’. As defined more specifically by the Translational Research Working Group regarding cancer research,

“Translational research transforms scientific discoveries arising from laboratory, clinical, or population studies into clinical applications to reduce cancer incidence, morbidity, and mortality”. (Translational Research Working Group 2007, 99)<sup>1</sup>

---

<sup>1</sup> Other definitions have been proposed (see McArthur and Borsini 2008, xix).

Yet this simple definition contains multiple meanings. In an editorial for the first issue of *Science Translational Medicine*, Elias A. Zerhouni, a former director of the National Institutes of Health and formerly a strong advocate of this approach, states that the term ‘translational’ can be understood in at least three senses: the rendering in clinical terms of what is *understood* at the basic level; the therapeutic *application* of basic biological knowledge; and the *extrapolation* made possible by the “profound unity of biology” resulting from “shared evolutionary pathways” (Zerhouni 2009). Indeed, application does not automatically follow understanding. This seems to be all the more true in psychiatry, where translational research has recently become a motto (the first issue of *Translational Psychiatry*, a publication by Nature Publishing Group was released in April 2011). Many potential treatments have resulted in disappointment and many exciting *in vivo* and *in vitro* models have failed to tackle the issue of human mental disorders.

As regards *in vivo* or animal models, experimental as well as more theoretical issues have been raised. As a matter of fact, concerns about the rationale of extrapolation from animal models have been both objected to and dismissed by philosophers (LaFollette and Shanks 1995; Schaffner 1998a; Schaffner 1998b; Schaffner 2000; Ankeny 2001; Weber 2005), so the question may need further consideration in the specific case of animal models of *mental disorders*. More specifically still, scientists spontaneously distinguish disorders that seem to hit potentially anyone, such as anxiety and depressive disorders, and those disorders that seem to threaten only a clinical

subpopulation, such as autistic disorders and schizophrenia; the latter seem to be even harder to model in animals than the former. The main reason seems to be the impossibility of modeling mental features that we find hard to understand ‘intimately’ in humans, a problem scientists apparently consider to be less stringent in the case of mood and anxiety disorders, where cognitive traits look less mysterious and behavioral traits far more recognizable.

Is modeling depression along with its mental processes, factors, and symptoms in animals truly unproblematic? My contention in this chapter is that the main problem scientists encounter in the field of mental health is not the fact that depression is difficult to model because its symptoms are ‘mental’, that is, personal, experienced, and contextual (1); rather, it is the fact that human depression is a fuzzy target of modeling, and that “piecemeal theorizing” (Murphy 2006) is required, which is a challenge to causal reasoning in medicine in general (2).

### **1 Translation and extrapolation about depression: the mind-body problem?**

If rodents cannot conceive of guilt, worthlessness, despondency and dejection, or if they cannot worry about what the future may bring and ‘consider’ suicide, is there any causal pathway left for them to develop genuine depression? Broadly speaking, there are two series of objections here:

1. the target condition, i.e. *mental* disorders, cannot exist in animals (*dissimilarity of animal models to their human target*);
2. the causal network relevant to this condition, including *mental factors*, is not relevant for animals (*impossibility of bypassing mental causality*).

In this section, I consider the way translational psychiatry deals with both problems and conclude that problems raised by modeling mental disorders in animals are not relevant to the mind-body problem.

### ***1.1 How translational psychiatry deals with the problem of similarity to depressive symptoms in animal models***

#### **1.1.1 Feature-to-feature resemblance**

In this section, I attempt to give a brief presentation of a field largely unknown to philosophers of psychiatry. When submitted to environmental factors similar to some of those precipitating depression in humans, like moderate chronic stress, animal models are expected to produce some behavioral symptoms and biological changes similar to those found in humans. In an experimental test of a pharmacological treatment, the animal should not only resemble the human, but the whole experimental situation should resemble the whole human situation too. This ‘situation’ is generally construed in the following way:

1. A disease entity instantiates in a human population marked by a genetic vulnerability through the occurrence of a pathogenic sequence of events enticing a neurobiological dysfunction.
2. This dysfunction can be assessed on the basis of clinical symptoms and biological markers thanks to diagnostic tests.
3. Any possible chemical treatment is a molecule with a pharmacological target called the endpoint.

The relevant features of resemblance in this situation are reduced to abstracted and idealized parameters (see table 1).

<b>Parameters (human condition)</b>	<b>Corresponding parameters (animal experiment)</b>
Human depressive disorder	Animal equivalent of the depressive disorder
Genetic and environmental vulnerability factors	Animal species and strand
Depressogenic sequence of events	Stress protocol
Dysfunctional neurobiological mechanism	Dysfunctional neurobiological mechanism
Clinical symptoms	Behavioral and cognitive changes
Biomarkers	Biomarkers
Diagnostic tests	Biological, behavioral and cognitive tests
Endpoint	Endpoint
Treatment (molecule, vehicle and	Treatment (molecule, vehicle and

---

 posology)
 

---



---

 posology)
 

---

**Table 1. Animal models of depression: Parameters of feature-to-feature resemblance.**

Each parameter in itself is a matter of concern:

- Disease entities* do not necessarily cross the boundaries of species.
- Animal species* display specific natural and artificial properties with known and unknown advantages and drawbacks for modeling a specific disease.
- Animal strands* are known for specific genetic vulnerabilities, some being an exaggeration of what can be encountered “in the wild”, that is in a natural human population.
- Pathogenic sequences of events* in humans have only partial correspondents in mice.
- Too many *neurobiological mechanisms* are potentially dysfunctional in depression.
- Some *clinical symptoms* of human depression have reasonably convincing equivalents in mice like psychomotor agitation or retardation, insomnia, weight loss, and even anhedonia, while others do not, such as feeling of guilt or worthlessness, irritability, or suicidal ideation.
- Biomarkers*, that is, evaluated indicators of the intrinsic causes of an illness, its clinical course, and its modification by treatment (Frank and Hargreaves 2003), not pathognomonic or cutting-off signs, cannot generally be measured *in situ*, but only indirectly, and



a thorough knowledge of the specifics of human and animal physiology is required to translate.

-Usual *clinical tests* differ in humans (questionnaires) and in animals (measurement and observation of activity).

-*Endpoint* of a candidate treatment, i.e. is the locus of action (receptor, behavior), is ideally the same in both humans and mice, but species may have different potential acceptors of the molecule leading to different potential side effects.

-A potential *treatment* itself has to be adapted in many ways, because of differences in required dosage (Lin 1998), or transposition of places to be stimulated in the case of transmagnetic stimulation (TMS).

Two concluding remarks are noteworthy. The first is that there are many more features to compare between species than it seems at first sight. Some are deeply problematic, others, not at all. The second remark is that scientists generally do not focus on the possibility of modeling, but rather on the strategic choices to make in order to design the best model possible. There are obviously good choices and bad choices, given what biologists generally assume about the inner working or main symptoms of depression. Nevertheless, all this indeed does not prove skepticism wrong about the potential results of animal research on depression. My main claim in this chapter is the reverse, i.e. that skepticism about resemblance of animal models to psychiatric conditions, if justified, does not entail that no significant result can come from this field.

### 1.1.2 Face validity, predictive validity, and construct validity

It is crucial to understand that translational research does not assess an animal model by *similarity* to the target, but by the *validity* of the inference. The similarity of the animal model to its human counterpart is neither a sufficient nor even a necessary condition of the validity of the extrapolation of a result based on animal experiments to a human population.

Many conditions of validity have been proposed and considered, for instance, in the field of the study of mood disorders, where this reflection happens to be most developed in translational psychiatry (Van der Staay 2006; van der Staay et al. 2009; Belzung and Lemoine 2011). Paul Willner proposed the most often cited conception of validity in this field (Willner 1984; Willner 1994; Willner and Mitchell 2002); he distinguishes:

- *Face validity*: “the extent of similarity between the model and the disorder (...) on as wide as possible a range of symptoms and signs” (Willner and Mitchell 2002);
- *Predictive validity*: “similar response to treatment” (Willner 1984);
- *Construct validity*: specific similarity of the animal experiment to the theoretical entity referred to as the disease and supposed to explain its symptoms, and to this theoretical entity only.

These three aspects of validity are not only different, but also independent of one another, so that when one is fulfilled, the others are not necessarily satisfied.

Face validity is what is commonly understood by the ‘resemblance’ of the animal model to its human target, especially in critical assessments of translational psychiatry. Yet face validity is not considered as equally important as construct and predictive validity. About the relations between construct validity and face validity, for instance, Willner says:

“Face validity only requires the demonstration of similarity between the model and symptoms of the disorder being modelled. *Construct validity does not require superficial similarity which may, indeed, be absent.* It does, however, require the demonstration of homology – the same theoretical constructs must be applicable in the two cases – and an empirically supported rationale for believing that the construct in question is fundamental to the disorder, rather than an epiphenomenon.” (Willner 1986, 684, my emphasis)

For instance, when examining two standard protocols, the Tail Suspension Test (TST) and the Unpredictable Chronic Mild Stress (UCMS), Willner notes that while the first has poor face validity and construct validity but strong predictive validity, the second has fair face validity and less convincing construct validity (depending on one’s hypothesis on the relation between stress and depression). A mouse that stops moving when suspended by its tail bears little resemblance to either the observable features or any received explanatory model of depression. Yet it is strongly predictive of the action of a drug in depression, and this is considered sufficient to assess

whether a treatment should be tested on humans. A mouse submitted for a protracted time to mild stressors such as nocturnal light, humidified soil, predator sounds, etc., shows signs very much alike to some of those depressed human subjects display, but it can be, and actually is, discussed on a theoretical level whether what stress entails is indeed an equivalent of depression.

It might be objected that predictive and construct validity also are specific forms of resemblance. As a matter of fact, predictive validity is more readily interpreted as a correlation of results of experiments than as a degree of resemblance between, say, effects of a treatment on mice and on humans. It says, roughly, that when the experiment is successful on mice, it will also be on humans. As to construct validity, it is not exactly the *similarity* of a model to its target, but rather the *conformation* of both the model and its target to a *theoretical* construct. Both what is observed in the model and in its target must be explained by the same underlying theoretical disease entity, all other theoretical disease entities excluded. It therefore depends on the nature of the theoretical construct, that is, whether animal modeling of depression is possible or not: as a highly sophisticated mental process, depression is hardly what rodents undergo. Nevertheless, the problem is that there is no consensual theoretical model of what depression consists of in translational research. Experimenting on animal models does not beg the question, but indeed excludes possible explanations – like highly elaborated psychodynamic models. This should not be considered a reductive claim, but a biological bet. What scientists really expect is not to make a point,

but depends on the results of the experiments: it is a strategy, good or bad, not dogma. Of course, some scientists may try to make reductionist points. Yet objecting that experimenting on animal models cannot achieve any knowledge of the allegedly corresponding human condition is indeed both dogmatic and bad strategy.

Some may ask: what if the success of antidepressant medication was precisely defined through the very hypothesis that animal models implement? For instance, if ‘depression’ was defined as a ‘low level of serotonin in the brain’, then surely a certain animal model with a low level of serotonin in the brain could provide a wonderfully predictive model... at a small price. This objection of circularity is to be carefully considered. So far, the efficiency of antidepressant medication has not been assessed through biological markers, but rather as a result of scales, like the MADRS or the HAMD described above, that are not semantically, but empirically connected (or not) with drug intake.

For all these reasons, the resemblance of the animal model to its human target is but one series of problems among the more general question of the validity of the extrapolation, probably the less important, because what matters most is both the power of prediction and the theoretical interpretability of the model.

## ***1.2 How psychiatry deals with the problem of heterogeneous factors***

An additional problem comes from modeling causal factors of depression. This section presents how it is addressed.

### **1.2.1 Multifactorial determinants of psychiatric disorders vs. animal research**

By famously urging the use of multifactorial models in psychiatry as well as in somatic medicine, Engel (1977) clearly opposed two attitudes he labeled reductionism and exclusionism. Whereas ‘reductionism’ was the view that biomolecular models should suffice to account for diseases both somatic and mental, ‘exclusionism’ was the view that conditions not amenable to biomolecular models were simply not diseases, which to some is the case for mental disorders (Szasz 1960). He proposed instead that the interaction of all factors, biological, psychological, and social, be studied in a system-theoretic approach. Since then, epidemiological studies of depression have repeatedly shown the importance of sociological and economical factors in its pathogenesis.

It requires a body and a nervous system to produce a sadness reaction, but it does not necessarily require stressful or demoralizing environmental conditions to deplete serotonin, stop hippocampal neurogenesis, or produce chronically high levels of cortisol. Psychological and social factors obviously supervene on some bio-

logical factors, whereas other biological factors do not make any psychological or social sense. Systemic approaches should therefore take great care to avoid considering the same factor twice, that is, as a biological as well as a sociological or a psychological factor. That is obviously not easily operationalized.

On the other hand, it is not necessary for someone working on animal models of depression to deny or even neglect the causal power of meaningfulness on an intelligent system (Bolton and Hill 2004). It is natural, on the contrary, to assume that if they exist at all in animals, mental causal factors, i.e. meanings, are already taken into account in their biological form and should not be ‘added’ somehow. In any case, cognitive bias, personality types, early-life events such as maternal care deprivation, neuroticism, and so on, all have proposed animal equivalents.

The problem is therefore not that some causal factors are not taken into account, but rather that in animals, the underlying biological phenomena of meanings might be absent or might underlie something else, so that in the best of cases, only an incomplete part of the biological mechanisms of depression can be studied. I assume that most scientists in the field would acknowledge that. I also assume that here again, what they are doing is a methodological bet rather than an ontological claim: it is possible to study some essential aspects of depression in a system that does not display other essential aspects of this mental disorder.

### **1.2.2 Knowledge approach vs. Treatment approach**

This methodological bet is an essential thing to understand. Let us take the example of drug discovery research. The point in modeling is not to create a homolog on which testing drugs is acceptable. It is not *knowledge-based* in the sense that scientists would think:

- If we knew how mental disorders worked, we could devise efficient treatments;
- We must experiment to know how they work;
- We cannot safely experiment directly on humans, we must therefore experiment on animals;
- What are the best animals to experiment on in order to know how the human disease works?
- Once we have good models, test of treatments on them will be trustworthy.

This line of thought can be encountered as a rhetorical justification of the method. However, as a matter of fact, the actual reasoning is rather *treatment-based*, and consists in the following:

- We know that some chemical agents have dramatic effects on mental disorders;
- These agents have effects on animals too;
- Which effects, on which animals, will guide us toward selecting the right chemical agents, i.e. with efficient and specific action, to test on humans?
- In order to restrain the possibilities of further agent candidates, we want to know how the right treatments work both on animals and on humans, i.e. what the targets could be.



The crucial point is that the analogy of causal effects is not assessed as a result of a previously existing theory, but rather that whatever the theory, any effective treatment on humans will have effects on some animals too; the question remains, though, which molecules affect which animals. Theories are not a starting point, but only a heuristic tool to restrain the domain of testable molecules.

### **1.2.3 Relevant differences: a philosophical account of how the problem of heterogeneous factors is successfully dealt with**

A final concern is the possibility of modeling when significant factors differ – such as guilt, despair, and feelings of worthlessness. Modeling mental disorders such as depression is but a case of what philosopher Daniel Steel addresses as “the problem of extrapolation in heterogeneous populations” (Steel 2008). This refers to the possibility of inferring from one experiment on a test population to properties of a target population, when both populations differ in causal aspects *relevant* to the inferred properties. What is needed, according to Steel, is knowledge that relevant dissimilarities at corresponding stages of the mechanism of interest in the model and target have no significant effect on the point of comparison, ‘significant’ meaning affecting the extrapolation. He calls this approach “comparative process tracing” (Steel 2008, 89).

A facilitating condition is that it is possible to abstract some relevant factors, namely, upstream factors (resulting in the same starting point of the mechanism of interest), and downstream factors (inter-

vening after the last stage of the mechanism of interest). The former can be abstracted away provided that there is no bypassing mechanism linking the upstream factor to a given stage in the mechanism of interest (Steel 2008, 90). The latter can also be abstracted away, provided that there is no feedback loop. Mental factors in depression are considered upstream factors, whereas mental reactions such as rumination are considered downstream factors. What is at stake is the mechanism that takes place in-between.<sup>2</sup>

Moreover, what is sought for is not a deterministic, but rather a probabilistic claim. The existence of any mechanism from X (cause) to Y (effect) in any individual in the animal population increases the probability of Y in the animal population when X is the case (Steel 2008, 109). At least, this is true if and only if there is at least one undisrupted mechanism from the cause to the effect, which in turn has the experimental consequence that any intervention on the cause makes a difference to the probability of the effect (*sublata causa, tollitur effectus*). This makes the case of mental phenomena a particularly difficult experimental problem. The reason is that they are very sensitive, or causally central, that is, possibly influenced by many more causal factors than most physical phenomena. Anything, from deprivation of a particular kind of food (say, chocolate) to living in a particular place, having undergone such or such experience in the past or having to in the future, can affect the mood in the most

---

<sup>2</sup> Steel also refers to another facilitating condition, namely, that the inference is about negative or positive causal relevance only, not on the effect size. Thereby, a certain degree of dissimilarity in corresponding stages of the mechanism does not affect the soundness of the extrapolation. This obviously applies to the problem of metabolism referred to above in the case of modeling depression.

radical way. But this does not make it specific in nature, and it does not preclude a carefully designed animal model that can teach us something about causal mechanisms in depression.

All in all, in the case of the study of depression in animal models, the treatment approach I described above entails the preliminary assumption that what is directly at stake are the molecular interactions within the brain, and only indirectly the mental or social factors that supervene on them. This, however, does not completely abstract away mental factors, for it does not abstract away the molecular causal effects that underlie these factors in the brain. The problem is therefore, to identify

- the modular neuronal mechanism where some dysfunction is implied in the mental disorder in human subjects as well as in animal models;
- the modular neuronal mechanism realizing this particular kind of cognitive process that presumably does not take place in rodents (say, despondency after losing a job, despair at the perspective of the void of a life to come without a loved one, and so on);
- the causal interactions between both modules.

That done, animal modeling applies to the case of mental disorders without any specificity, provided that what is looked for is a picture of crucial parts of the mechanism, not the entirety of the mechanism, and that the approach is treatment-driven rather than knowledge-driven.

The conclusion is that modeling depression in animals does not raise specific problems because it is a mental disorder. The reasons why are: 1) most relevant features can be reasonably considered similar in both animals and humans; 2) face validity (i.e. resemblance) is but a minor point in extrapolation; 3) mental factors can be considered as supervening on neurobiological mechanisms; 4) the approach is generally treatment-based, that is, does not presuppose a known mechanism to be similar, but rather, an interesting effect to occur in animals too; and 5) relevant factors can be abstracted away consistently even in the case of mental disorders. However, this does not mean that mental disorders in general, and depression in particular, do not pose specific problems to animal modeling. The problem with animal models of depression is therefore not an instance of the mind-body problem. What is it, if any?

## **2 Something is modeled in translational psychiatry: what, exactly?**

The problem of inferring causal relations from animal models of depression to human targets is not the similarity of the model, but the indeterminacies of the target. Depressed human subjects are a fuzzy target. The definition of depression as a syndrome contributes to this fuzziness, and that must have consequences on the relevance of likely animal models: should they display all the symptoms animal models can, or a significant subset of them? So does the fact

that the various mechanisms known to be possibly involved in depression do not fall neatly in place. In this situation, the scope of the extrapolation cannot be precisely determined.

### ***2.1 Fuzziness***

The target population is fuzzily determined, first, through the uncertainty of the disease entity due to its polythetic semiology, second, through the width of the spectrum of mood disorders, and third, through the dimensionality of the criteria of depressive disorders. Consequently, the poor results of factor analysis and principal component analysis have repercussions on the evidential power of animal models of such conditions.

In the *Diagnostic and Statistical Manual of Mental Disorders* (DSM), a major depressive episode (MDE) is defined through alternative combinations of symptoms. This kind of diagnostic definition, called *polythetic*, leads to much heterogeneity in the population of people suffering from a MDE. Østergaard et al. (2011) calculated that for 5 items out of 9, the first or the second being necessary, the possibilities number up to 227. Moreover, considering that 3 criteria contain alternate but incompatible subcases (*either* weight gain *or* weight loss, etc.), the number of possible forms of MDE could amount to 1497. The result is that many configurations of symptoms have not even one symptom in common, and many more share only non-specific symptoms (such as insomnia and weight loss). The problem is that no common underlying dysfunctional system has yet

been discovered from which such heterogeneous patterns could be causally derived, and thereby the unity of the syndrome, justified. What is, then, the possible configuration, or the core symptom, that animal models should display? Anhedonia and ‘helplessness’ behaviors are most often proposed; this is a strong theoretical claim, which clinicians may question.

MDEs occur within the course of many mental disorders, including major depressive disorder (MDD), but also bipolar disorders or schizophrenia. It is a question whether the nature of the MDE is the same under the surface of symptoms in these different cases. Moreover, major depressive disorder is different from but next to disruptive mood dysregulation disorder, chronic depressive disorder (formerly known as dysthymia), premenstrual dysphoric disorder, the controversial and abandoned ‘mixed anxiety/depression’, substance-induced depressive disorder, depressive disorder associated with a known general medical condition, and a few other specified depressive disorder and unspecified depressive disorder. It is not so clear whether there are biological differences underlying these distinctions and above all how, if biological, these differences could be modeled in animals. Moreover, one can wonder whether what is induced or observed in animals is specific to MDE. The Novelty Suppressed Feeding test (i.e., testing whether a rodent will or will not eat in an unknown environment), for instance, is sometimes considered as a test of anxiety, and sometimes as a test of depression<sup>3</sup>. This uncer-

---

<sup>3</sup> As neurobiologist Catherine Belzung explained to me (personal communication).

tainty unintentionally reflects many hesitations in the clinic about the distinction to be made between these two disorders. The upshot is that whereas animals are supposed to model one mental disorder, this specificity is highly questionable and there is a genuine problem in knowing what, exactly, is modeled in a ‘depressed’ mouse.

Two further problems are worth mentioning. The first is the dimensionality of criteria. Most features of depression can, and maybe should, be assessed quantitatively: disturbed sleep and weight change, for example; but things such as loss of interest and feelings of worthlessness can be quantified as moderate, mild or severe (see the Beck Depression Inventory (BDI), the Hamilton Rating Scale for Depression (HDRS/HAMD) and the Montgomery-Åsberg Depression Rating Scale (MADRS)). It is often noted that the main outcome of such continuity is fuzzy boundaries between normal and pathological sadness and pessimism. This, too, is a difficulty for modeling: should only severe cases be modeled? However, these scales are generally used as a measure of severity rather than presence of a MDE.

Second, there is the problem of the poor results of factor and cluster analyses. For instance, the key result of Paykel’s study in the 1970s was that the main distinction was between old subjects with severe depression and young subjects with mild depression, the former subdividing between psychotic and anxious types, the latter, between hostile types and those with personality disorders (Paykel 1971). A classical factor analysis of depression found three profiles, “anxious-tense depression”, “hostile depression” and “retarded de-

pression” (Overall et al. 1966); yet another found that the main factors were the neurotic-psychotic axis and the depressive vs. paranoid factor (Kay et al. 1969). As philosopher Rachel Cooper rightly emphasized, the problem with all studies of this kind is theoretical and comes either with the choice or with the interpretation of variables and results (clusters or profile) (Cooper 2010). In other terms, the problem is the elaboration of the relevant construct. Animal models should conform to that construct, but a prior decision is to be made: should psychiatrists include observable variables that have clear equivalents in mice? The second question is: can the result of a cluster or factor analysis in humans be back-translated into mice?

## ***2.2 Mosaicism and chimerism of models***

Despite the treatment approach I emphasized in the first section, some knowledge of the whole biological picture is necessary to model the disease. We have some indeed, but it is very patchy: areas of known mechanisms upon which used treatments act are surrounded by largely unknown causal chains. This situation leads to what philosopher Dominic Murphy calls, in *Psychiatry in the Scientific Image*, “piecemeal theorizing”, that is, the fact that psychiatry about mental disorders in general employs “generalizations or causal models of limited scope at different levels” rather than “one big elegant theory” (Murphy 2006, 240-1). In the more specific case of one given mental disorder in turn, Murphy distinguishes global models, de-



signed to explain all symptoms by one general dysfunction, and modular models, designed to explain all symptoms by a series of part-dysfunction, each explaining one symptom or one part of the symptoms.

The requirements of animal experiments are partly responsible for the prevalence of piecemeal theorizing and modular models in biological approaches to depression. One has to choose *one* putative mechanism to model it. In translational psychiatry, scientists therefore consider their animal models to instantiate *one* component mechanism of the disease (hopefully *the key* mechanism). But they do not exactly know where, in the global plan of the disease, this mechanism stands – that is, what are its causal relations with the rest. It may be central or peripheral, it may precede, follow, or add up to other mechanisms in the course of the mental disorder.

The main biological hypothesis on depression until recently, the monoamine hypothesis, is an illustration of piecemeal theorizing. The serendipitous discovery of seemingly effective treatments of depression, tricyclic antidepressant (imipramine), and monoamine oxidase inhibitors (iproniazid), led to research on their target, mainly, the serotonergic system (Maxwell and Echhardt 1990; McArthur and Borsini 2008). Further exploration of this biological mechanism in animal models led in turn to better animal models on the one hand, and better molecules on the other hand (mostly, with less side effects): chiefly, fluoxetine. Many have thought that an imbalance in the serotonergic system, grossly in the form of a depletion of synaptic serotonin, expressed the main mechanism of depression (Hirsch-

feld 2000). Yet this position hides a paradox, that of the delay of action of mood regulators: how can depression consist mainly in a serotonin deficit if normal levels are restored almost immediately under medication, but improvement occurs only 3 weeks after the treatment begins? This question has more recently lead scientists to question the centrality of serotonergic system in depression, and look for complementary, alternative, or competing hypotheses. Among others, a possible disruption in hippocampal neurogenesis, which has in turn been linked to a mechanism known to be crucial in anxiety disorders, the so-called hypothalamo-pituitary-adrenergic axis of stress, and so on to a dozen other mechanisms possibly involved in depression (Licino and Wong 2005).

This situation, i.e. many different animal models of various mechanisms of mostly unknown importance and causal relations with one another, though referring to the same mental disorder, I propose to call *mosaicism* of animal models. I distinguish it from *chimerism* of animal models, i.e. the fact that several different animals are usually required for the instantiation of one hypothesis, each instantiating one part of the component mechanism of the disease, none instantiating the whole mechanism. In other terms, there is not one animal model similar to the mental disorder according to hypothesis *X* on its inner mechanisms, but rather several animal models implementing parts of the whole mechanism supposed to take place in the disease in humans. For instance, the hypothesis that the s/s polymorphism of the serotonin transporter gene is a vulnerability factor to depression, and requires that a whole chain of events from this polymorphism to

symptoms of depression be established. As a matter of fact, it has been, but not the whole chain into a single model – it required at least three different rodent models.

Both mosaicism and chimerism of animal models are consequences of modeling in the dark about the ins and outs of mental disorders. The main outcome of that necessity to model piecemeal is on the scope of the extrapolation. First, one does not know the strategic place of the part of the human disease modeled in the animal and, second, one does not extrapolate from one type of animal to the human target population on the basis of a one-to-one resemblance, but from several types of animals to the human target population on the basis of a many-to-one resemblance.

## **Conclusion**

What is lost in the translation of animal models to human targets? It was not a well-formed meaning, but rather the conviction that there was one in the first place. The question is not the impossibility of translating mental properties and mental factors into ‘animal language’. The real problem with animal models of depression is depression itself. First, the human population to be modeled is not itself strictly determined, and second, animal models cannot be thought of as experiments about mechanisms precisely placed and fitting nicely into the sound template of a biological theory of this mental disorder, for there is no such theory. They indicate fixed di-

rections and sound natural facts about those mechanisms, but we do not know which directions and which facts.

In the face of this problem, there have been theoretical attempts at gathering all we know about the biology of depression (Kendler et al. 2002; Kendler et al. 2006; Willner et al. 2012). The problem is always about searching either for consistent, if alternative, underpinning mechanisms of a same clinical condition, or for a relevant definition of the clinical presentation of the same biological dysfunction. Should we redefine depression after what we think we know of its underlying biology or should we search for a biological rationale of what we think we observe of its clinical presentation? Translational psychiatry of depression is at the heart of this problem. It tends to use concepts, such as that of ‘endophenotypes’, to resolve the question. As opposed to exophenotypes, that is, clinical presentations of depression affected by many environmental and developmental factors, endophenotypes are hypothetical underlying presentations of genetic factors of depression. In the present field, they consist in behavioral patterns affected by genetics only, that is, hereditary, possibly hidden or disguised by cultural or biographical events (John and Lewis 1966; McArthur and Borsini 2008). Their function is obviously to deconstruct the clinically defined syndromes into (yet) undetermined alternative phenotypes, of which traditional clinical entities might have been an approximate picture. As endophenotypes would not necessarily be species-bound, and could be defined after whatever genetic determinants of depression can be discovered, they would totally redefine the human condition that de-

pression is into a biological, trans-species condition. If the unity and reality of the condition were specifically human and relied on meanings, then this attempt would be doomed to failure.

This ultimate consequence of animal modeling – to rewrite the very notion of what depression is about – is neither to be fought nor favored by philosophers. As a matter of fact, scientific progress is about the naturalization of prescientific notions such as movement, heat, reproduction, species, but also diabetes, epilepsy, and, possibly, schizophrenia and depression. Some of these attempts at naturalization seem to be bound to fail, some, to succeed: how could a philosopher know which ones can successfully be naturalized before they are? For a successful naturalization, i.e., roughly, an explication of a profane notion through the terms of natural sciences, is not about the faithfulness of the scientific concept to the prescientific concept (Murphy 2006; Lemoine forthcoming). It is about capturing, in a consistent picture, how nature works.

#### References

- Ankeny RA (2001) Model Organisms as Models: Understanding the “Lingua Franca” of the Human Genome Project. *Proceedings of the Philosophy of Science Association* 2001:S251–.
- Belzung C, Lemoine M (2011) Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression. *Biol Mood Anxiety Disord* 1:9. doi: 10.1186/2045-5380-1-9

- Bolton D, Hill J (2004) *Mind, Meaning and Mental Disorder: The nature of causal explanation in psychology and psychiatry*, 2nd ed. OUP Oxford
- Campbell DT, Stanley J (1963) *Experimental and Quasi-Experimental Designs for Research*, 1st ed. Wadsworth Publishing
- Cooper R (2010) *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*, 1st ed. Softcover of orig. ed. 2005. Springer
- Frank R, Hargreaves R (2003) Clinical biomarkers in drug discovery and development. *Nat Rev Drug Discov* 2:566–580. doi: 10.1038/nrd1130
- Hirschfeld RMA (2000) History and Evolution of the Monoamine Hypothesis of Depression. *J Clin Psychiatry* 61 (suppl. 6): 4–6.
- John B, Lewis KR (1966) Chromosome Variability and Geographic Distribution in Insects. *Science* 152:711–721. doi: 10.1126/science.152.3723.711
- Kay DW, Garside RF, Beamish P, Roy JR (1969) Endogenous and neurotic syndromes of depression: a factor analytic study of 104 cases. Clinical features. *Br J Psychiatry* 115:377–388.

Kendler KS, Gardner CO, Prescott CA (2002) Toward a comprehensive developmental model for major depression in women. *Am J Psychiatry* 159:1133–1145.

Kendler KS, Gardner CO, Prescott CA (2006) Toward a comprehensive developmental model for major depression in men. *Am J Psychiatry* 163:115–124. doi: 10.1176/appi.ajp.163.1.115

LaFollette H, Shanks N (1995) Two Models of Models in Biomedical Research. *Philosophical Quarterly* 45:141–160.

Lemoine M (forthcoming) The naturalization of the concept of disease. G. Lambert, M. Silberstein & P. Huneman (eds), *Classification, Disease and Evidence. New Essays in the philosophy of medicine*, Springer.

Licino J, Wong ML eds (2005) *Biology of Depression. From Novel Insights to Therapeutic Strategies*, Wiley.

Lin JH (1998) Applications and limitations of interspecies scaling and in vitro extrapolation in pharmacokinetics. *Drug Metab Dispos* 26:1202–1212.

Maxwell RA, Echhardt SB (1990) *Drug Discovery: A Casebook and Analysis*. Humana Press Inc.

McArthur RA, Borsini F (2008) *Animal and Translational Models for CNS Drug Discovery: Psychiatric Disorders*. Academic Press.

- Murphy D (2006) *Psychiatry in the Scientific Image*. MIT Press.
- Overall JE, Hollister LE, Johnson M, Pennington V (1966) Nosology of depression and differential response to drugs. *JAMA* 195:946–948.
- Paykel ES (1971) Classification of depressed patients: a cluster analysis derived grouping. *Br J Psychiatry* 118:275–288.
- Schaffner KF (1998a) Genes, Behavior, and Developmental Emergentism: One Process, Indivisible? *Philosophy of Science* 65:209–252.
- Schaffner KF (1998b) Model Organisms and Behavioral Genetics: A Rejoinder. *Philosophy of Science* 65:276–288.
- Schaffner KF (2000) Behavior at the Organismal and Molecular Levels: The Case of *C. Elegans*. *Philosophy of Science* 67:288.
- Van der Staay FJ (2006) Animal models of behavioral dysfunctions: basic concepts and classifications, and an evaluation strategy. *Brain Res Rev* 52:131–159. doi: 10.1016/j.brainresrev.2006.01.006
- Van der Staay FJ, Arndt SS, Nordquist RE (2009) Evaluation of animal models of neurobehavioral disorders. *Behav Brain Funct* 5:11. doi: 10.1186/1744-9081-5-11



Steel D (2008) *Across the Boundaries: Extrapolation in Biology and Social Science*. Oxford University Press

Szasz TS (1960) The myth of mental illness. *American Psychologist* 15:113–118. doi: 10.1037/h0046535

Translational Research Working Group (2007) *Transformation Translation – Harnessing Discovery for Patient and Public Benefit*. National Cancer Institution, National Institutes of Health, U.S. Department of health and human services., Washington, D.C.

Weber M (2005) *Philosophy of Experimental Biology*. Cambridge University Press

Willner P (1984) The validity of animal models of depression. *Psychopharmacology (Berl)* 83:1–16.

Willner P (1986) Validation criteria for animal models of human mental disorders: learned helplessness as a paradigm case. *Prog Neuropsychopharmacol Biol Psychiatry* 10:677–690.

Willner P (1994) Animal models of depression. In: den Boer J, Ad Sitsen J (eds) *Handbook of depression and anxiety. A biological approach.*, Marcel Dekker. New York - Basel - Hong Kong, pp 291–316

Willner P, Mitchell PJ (2002) The validity of animal models of pre-disposition to depression. *Behav Pharmacol* 13:169–188.

Willner P, Scheel-Krüger J, Belzung C (2012) The neurobiology of depression and antidepressant action. *Neurosci Biobehav Rev*. doi: 10.1016/j.neubiorev.2012.12.007

Zerhouni EA (2009) Space for the cures: science launches a new journal dedicated to translational research in biomedicine. *Sci Transl Med* 1:1ed1. doi: 10.1126/scitranslmed.3000341