

Representation and Practical Accomplishment in the Laboratory: When is an Animal Model *Good-Enough*?

Lewis JL, Atkinson PA, Harrington J, Featherstone K.
Sociology (Forthcoming).

Abstract

Science is dependent on standardised rodents. These lab-rats and lab-mice are different from their wild cousins. Interest in the genetic basis of a variety of phenotypes has meant that lab-rodents have been bred over many decades as resources for experimental science. We focus on the use of lab-rodents as animal models in the study of single-gene human medical conditions. The animals' status as models is not a given. Scientists calibrate animals against the medical phenomena which they are intended to represent. In turn, human medical conditions and the patients who manifest them have to be calibrated against the rodent models. The creation of animal models and their interpretation is, therefore, part of the practical work of biomedical scientists and their adjustment is a key aspect in determining when the model is *good-enough*.

Introduction

We document the work of biomedical scientists in using rodent models of human diseases for single-gene disorders. We illustrate the work of choosing animal models to represent human medical phenomena, and of making those medical conditions, in turn, fit the models. The choice of species and the calibration of models are, therefore, fundamental aspects of the practical work conducted by biomedical scientists, as is the determination that any given model is good-enough. The notion of 'good-enough' in the social sciences derives, in part, from the work of Winnicott (1953) and has subsequently been applied to a wide number of phenomena, in management for instance being related to the Pareto Principle: diminishing returns mean that striving for perfection is self-defeating. It is also used in systems engineering and software development to describe a level of representation adequate for functionality, and is widely used in discussions of scientific measurement. Our own use here derives more from the phenomenological tradition, especially the work of Alfred Schütz (1967), to capture the practical work of deciding how and when a representation is deemed adequate for the users' interests-at-hand.

All forms of research involve a type of simplification and representation. Delamont, Atkinson and Parry (2000) show how research fields - from anthropology to

biochemistry, to artificial intelligence - are 'predicated on particular ways of establishing conventional representation and simplifications' (p114). Modelling, in particular, is a type of simplification and representation intrinsic to experimentation, and animal modelling is a particular form that plays a fundamental role in the biomedical disciplines.

Mice, rats and zebrafish account for over 90% of the animals used as models in biomedical research. Laboratory-rodents include transgenic mice (where DNA from an organism has been inserted into the embryo), knock-out and knock-in rodents (where a specific gene has been removed from or inserted into the animal) and animals with surgical lesions. These lab-rodents are among a wide range of contemporary *bio-objects*: organisms and assemblages that transform boundaries between species, between the human and the non-human, and between the organic and non-organic (Vermeulen, Tamminen and Webster 2012). Transgenic rodents are part natural, part artificially created (Holmberg and Ideland 2009, 2012; Shapiro 2002). They are 'progressively transformed from holistic naturalistic creatures into analytical *objects* of technical investigation' (Lynch 1988: 266). They are, as Michael (2001) suggests, 'bespoke' organisms. Such transformation requires work and interpretation on the part of laboratory scientists.

There have been animal 'models' of the human and the social in many cultural and historical contexts. Fable and folklore have used exemplary animal species to 'model' human social behaviour: for example the anthropomorphising of the industrious bee or ant, the wise owl and the cunning fox. They are all illustrative of the ways in which we take animals to be essentially *like* humans despite differences (Daston and Mitman 2005). The pervasive phenomenon of totemic species represents another way in which natural kinds can be used as metaphors whereby social and cultural categories have their homologies in the natural world (Lévi-Strauss 1963). Likewise, Ritvo (1987) documents culturally-defined homologies between social status and animal pedigrees.

Contemporary animal models are used in medical research to create a distinctive form of correspondence. They are used to represent distinctive features of human medical conditions. As we shall see, they cannot do so in every possible detail, and there are often trade-offs between functional value and representational detail. Laboratory rodents are, therefore, simultaneously like and unlike their human counterparts, as are

the manifestations of illness, produced by genetic manipulation or surgically-induced lesions. They are part of a constellation of biomedical phenomena that bring human and animal domains into alignment (cf Agamben 2004; Haraway 1997; Brown and Michael 2004; Birke, Arluke and Michael 2007).

Methods

The paper derives from three projects examining the scientific and medical work within stem cell laboratories and clinical genetics services. Animal modelling is fundamental to the practical work at all these sites. First, we draw on observation and interviews at a laboratory we call Cell-lab, conducting experimental foetal tissue research into Huntington's and Parkinson's disease (Author 1). Second, the paper draws on observations and interviews at four research centres focussing on stem cell treatments, which we call TransLab (Author 3). Third, the paper draws on an ethnographic study of Rett Syndrome documenting multiple sites where it is enacted and represented (Author 4). Fieldwork at the research sites was conducted between 2006 and 2010. Author 1 spent six months with research technicians, and cell and behavioural scientists in an animal house observing the performance of behavioural tasks on laboratory-rodents. Author 3 spent varying periods of up to three weeks over four years observing and interviewing stem cell scientists, clinicians and animal welfare officers. Author 4 spent eighteen months carrying out ethnographic fieldwork following a local clinical and scientific team and the wider scientific field at international conferences. Interview data from the three sites were transcribed and analysed, while fieldnotes were written-up immediately following observations of laboratories, clinics, and conferences. All data-sets have been analysed thematically, and are the basis of publications elsewhere (Harrington and Stephens 2010; Featherstone and Atkinson 2011; Lewis and Atkinson 2011; Stephens, Lewis and Atkinson in press). Our analytic aim here is to draw out a *generic* account of themes that have been identified across the various research sites dealing with the good-enough model.

Models

A model, by definition, in order to function *as* a model, must preserve some essential features of the original, while discarding or simplifying others. If not, it will expand to encompass the same degree of detail as the original. It would be like the map described

by Borges (1988). The map-maker progressively incorporated features until the map grew to the same size as the original territory rendering it functionally useless. Modelling implies a number of activities or functions: models simplify, they standardise, and they stand proxy for other objects. Animal models do all of these things – substituting for human actors, embodying selected biological and medical traits in standardised and manageable ways, and being among several versions or representations of disease entities. Models can vary depending on the number of features they mimic and what they omit, each being a gloss on the original (Garfinkel and Sacks 1970). Simplified representations are often preferred over more detailed versions for the purposes of scientific presentation (Mulkay and Gilbert 1984) and complex data are routinely broken down in order to render them tractable (Star 1983). Furthermore, models embody the practices of framing and answering scientific questions (Fox-Keller 2000).

In biomedicine, ‘model’ has two connotations. On the one hand, the biomedical communities we study refer to laboratory animals as ‘models’. This is what might be called the *bio-object*. Analytically speaking, on the other hand, we can identify a much broader, generic process whereby representations are constructed and circulated. These can take many concrete forms – from laboratory animals, to photographs and video-recordings, to statistical representations, to graphic representations, and so on. The latter are all ‘models’ in the broader sense.

Our general contention is that the power of any given model - in whatever material embodiment - is not a given. There is no natural set of representations or equivalences that confers the status of model. Models need to be achieved. Their representational capacities have to be worked at continually. The degree of correspondence needs calibrating and adjusting in the course of everyday laboratory work. Human medical conditions and versions of them in rodents are no different in this respect. They too require alignment through such work and through professional judgments of adequacy including deciding when the model is good-enough.

Choosing and Calibrating Animals

First, scientists need to consider what species to use. When considering the ‘appropriate material’, the modeler must ask: what is the model’s function – for

example to map a particular biological pathway or to mirror a human disease? Which animal is then best suited to achieve this and is the animal practically and morally viable for use in experimentation? Indeed, the decision over which animals should be used in experimentation has had an awkward and often controversial history, encouraging ethical as well as biological debates (French 1975; Knight 2011). For example, when deciding what constitutes an animal in the laboratory there may be distinctions drawn between vertebrate and invertebrate species; such demarcation lines are entangled with emotional, ethical and cultural values as well as scientific ones (Paton 1984).

In the laboratory, a stable hierarchy and menagerie of species has emerged, including yeast, worm, fly, mouse, rat, through to monkey and chimpanzee. This scale reflects the complexity of the organism, the amount of DNA shared with humans, as well as more elementary characteristics such as the size of organs and speed of reproduction. Once made, the choice of species, in turn, implies investment in infrastructure, and commitment to an array of specific experimental techniques and expertise. Paton (1984) reports, it was towards the beginning of the twentieth century that rodents emerged as the most common subject for experimentation. The appearance of the mouse, for example, has been traced back to C.C. Little's research into cancer and genetics (Radar 1998; 2004); while Clause (1993) tells a similar story about the standardization of the Wistar rat. Endersby (2007) discusses the significance of the two rodent species, which have become standard animal-models for biomedical research. They are used to display the molecular basis of many human diseases and enable proof-of-principle studies for novel diagnostic and treatment strategies (cf. Wanga *et al.* 2005; Davis *et al.* 1998). Although animal models are artefacts designed to be shared across scientific communities and are often referred to as boundary objects (Star and Griesemer 1989), boundary walkers (Haraway 1997), or boundary crawlers (Holmberg and Ideland 2009), the specific features of each model have become associated with distinct fields of biomedical science. Research groups have therefore been known to align and identify themselves with particular animal models (Kohler 1994). Accordingly, each species has large-scale, international, scientific communities attached to them.

Next, a research group must decide whether an animal-model achieves a good enough equivalence to their target medical condition: whether it will develop a symptom or disease or respond adequately to treatment. There is always, in principle, the practical problem for the model user: How ‘good’ is a good-enough animal model? How many of the original features of the human disease must be preserved? How many functions need to be implemented in the design? How reliable does the model need to be to work? In other words, the adequacy and fidelity of any model must be judged.

Two dimensions are therefore involved in animal modelling. First, there is the *vertical* dimension: that is, the *choice* of animal species for the model. This reflects degrees of general similarity to humans, from the simplest organisms such as yeasts to more complex organisms such as primates. For example, Brown and Michael (2001) discuss scientists’ discursive work in establishing the position of pigs within such a calculus of comparison with humans. This vertical axis also reflects pragmatic issues, such as the ease of breeding laboratory populations and the infrastructure they require. Indeed, the vertical axis represents a series of similarities and differences between species that corresponds to the process of replacement or transposition (Friese and Clarke 2012). Second, there is the ‘*horizontal*’ dimension; that is, the axis of similarity to a *specific* medical condition or symptom that can be achieved. The horizontal axis represents the processes whereby particular populations or even individual animals are matched to selected features of the target medical condition. We argue that the vertical axis represents relationships of a *metaphorical* nature (based on difference, separation and substitution), while the horizontal axis represents relationships of a *metonymic* kind, (based on selected affinities). The latter, we maintain, depends on processes of *calibration*ⁱ: the iterative adjustment of the rodent model and the disease-model to bring them into alignment for the purposes of practical experimentation. In the rest of the paper we use empirical data to example how the metaphorical and metonymic are intertwined in the process of making the good-enough model.

Rodents as Standard

As already indicated, the choice of species is the first decision to be made. Mouse and rat currently sit in that good enough space to satisfy a Darwinian inspired evolutionary resemblance to humans (warm blooded mammals with a close enough genetic affinity), whilst are also arguably ethically less problematic than other possible

candidates (e.g. ape, or the domesticated cat or dog). That is, mouse and rat are both *similar to* and *different enough* from humans to have emerged as the marquee animals for use in experimentation. These similarities and differences combined with their experimental amenability make rodents functional experimental animals, although they cannot replicate various kinds of human functions, such as higher cognitive processes, language, or bipedal posture and gait.

Rats and mice – rodents that embody the instrumentation of modern genetics - are both used in the biomedical research reported here. The two standard species co-exist, in part, because different scientific communities (molecular biology and behavioural studies) currently work together under the domain of translational research. Two clear examples of this are research into Huntington's disease (HD) and Rett syndrome (RTT).

HD is a dominantly inherited, degenerative neurological condition characterised by complex phenotypic behavioral and neurological features, which cause involuntary twitching and chronic jerky uncontrollable movements, leading to complete disability, cognitive impairment and dementia. In the world of molecular and histological research on HD, the mouse is the material of choice as it can be manipulated to be representative of the human HD gene. In the following quote Dr Greenⁱⁱ explains how the mouse is the superior model for the sequence of nucleotides CAG:

‘The whole genetics behind it is [different]. They could only get, I think, 50 CAG repeats into the rat... In Huntington's there's up to 120 CAG repeats or something but they couldn't get anything high. Whereas with mice you can get various lengths of CAG repeats' [Dr Green: Cell-lab].

Dr Green's comments are representative of the wider field of transgenic animals. The genetic engineering of animals – with the exception of mice - has been a rather slow process requiring large financial investment. As she explains, within HD research they have not been able to create a transgenic rat with high enough CAG repeats and this renders it a poor genetic model compared to the mouse.

Further exploration of HD research, however, reveals that if the behavioral features of the condition are the focus, then the rat becomes the experimenter's preferred rodent.

When asked therefore to justify why the laboratory uses both rats and mice, the response was:

‘Because you’ve got different flexibilities in each. So in mice it is much easier to make transgenics [for Huntington’s]... The reason for using rats is that they are much better at doing animal behavior - really substantially better - so if you’re looking at complex behaviors that are maybe more akin to human behavior, then you’re much better off with a rat. It’s also got a bigger brain, so you know if you’re doing transplantations... it’s much easier, the bigger the brain the easier it is to do those sorts of experiments’ [Prof Stephens: Cell-lab].

Professor Stephens’ comment provides us with confirmation that the mouse is easier to manipulate at the genetic level and, likewise, that the rat is the preferred model to study complex behavior: rats perform experimental tasks more willingly and demonstrate a range of behaviors that are perceived to be more ‘akin’ to human behavior. Rats also have larger brains, which mean surgical modifications such as inserting therapeutic grafts are easier to perform. She uses the term ‘flexibilities’ to describe how each animal (even those animals from the same order of species) and therefore each model has its advantages and disadvantages.

A colleague elaborates further on the relationships between mouse and rat. Corroborating the previous two extracts she explains that mice are ‘easier’ to manipulate at a molecular level. She goes further, however, explaining that although this delicacy and intricacy is valued within molecular biology, these same ‘traits’ mean that they are not always robust enough for behavioural experimentation, particularly in comparison to the more gregarious rat.

‘Mice were generated first because it's easier to generate a [transgenic] model in a mouse... But mice are not as easy to test with because they're a bit temperamental and they have a habit of just dying... Rats are a bit more robust. I'm a bit mouseist because I'm allergic to them, but rats are just generally more robust animals and they have less tendency to being nervous about being handled and less tendency to just keel over and die’ [Dr Dunbar: Cell-lab].

Dr Dunbar aligns herself with rats, claiming that they are more sturdy and comfortable around humans than mice. Nevertheless she recognises the history of the mouse-model and it is clear that the laboratory use both animals. Therefore when it comes to

calibrating the choice of animal to the human disease, different scientific groups value different animals. Species alignment is based on local preferences.

A particular species may become associated with a disease, or a specific feature of that disease. For example, it would appear that the transgenic mouse has a closer *genetic* affinity to human HD, while the rat with a lesion provides a better imitation of human HD *behaviour*. The global choice of species used to stand in place for the human (metaphorical relations) therefore has to be matched with local laboratory expertise. Scientists endeavour to calibrate particular features of the animal against different features of the medical condition (metonymic relations).

There are, of course, other practical reasons for using rats as models over and above any behavioral equivalence to humans. Although significantly larger than the mouse, the rat is still a smaller and cheaper option compared to other mammalian candidates. Larger mammals tend to live longer and, as such, might mirror the human lifecourse or disease trajectory more faithfully. However smaller mammals with a shorter life-cycle such as the rat are considered to be more in harmony with the temporal rhythms of experimentation and research, an exemplar of how models are situated in particular worlds and judged through different lenses and temporal frames. Therefore what makes a model good enough is dependent on which attributes are prioritised and valued, and by whom. In terms of biological resemblance it might be argued that a primate would be a better model, but in terms of experimental practicalities (which include regulatory barriers, facilities, methods, resources and infrastructures) the rat may be considered to be a better option as the following quote indicates:

‘You’ve got to look at the whole disease and then you’ve also got to look at how we can *represent* that disease to perhaps try to find strategies to alter the disease progress. In terms of what's available for testing for example, arguably, a primate would be much better... to use to model human disease. But, of course, primates are very expensive, you can't use many of them, there's lots of legal implications, there's lots of social implications...so we have to find other models and so we use the rat. Rats are very cheap, you can produce lots of them, they're in common use. Socially there's less of a problem than with primates’ [Dr Grant: Cell-lab].

The claim from Dr Grant is not that the mouse and rat are [necessarily] the best candidates for biological imitation, but that any decision is also bound up in ethical, social, cultural and institutional judgments. The choice of mouse or rat may reflect

their ready availability just as much as any special suitability. To state that ‘there we are, so we have to find other models and so we use rat’ is to suggest that rodents are good enough because they are made to be.

Models for Particular Purposes

As we have already suggested, rodents need to be aligned, or calibrated, with the selected features of the human medical condition. In the case of HD there is no single rodent-model but a proliferation of models within models, each representing a smaller or a larger array of symptoms. Models can be matched to a particular feature of a disease or biological pathway. In Professor Stephens’ work examining the effectiveness of inserting therapeutic grafts into the brains of rats, the rat with a surgical lesion provides a stable model adapted to experimental requirements.

‘If you have more models and if you understand fully the complexities of those models then it gives you more options as to match the therapeutics up to that model. That’s the big advantage. For example lesion-models are probably not very good for therapeutics but they’re ideal for testing very quickly whether your grafts are going to work rather than waiting for an animal to get to two years old before you sacrifice it and see whether your grafts have survived or not, which is an issue that we have’ [Professor Stephens: Cell-lab].

Rather than having to keep models alive for extended periods, with all the associated staff and resources costs, the lesion-rat model provides reliable and speedy results. The global choices of which animal to use are, however, influenced heavily by local practices, local traditions and local expertise. The following example describes the close alignment of Cell-lab with the mouse-model. The team has been using mice-models for a long time, and they exude a sense of comfort with this locally understood model.

‘The mouse-model has been around for a long time and so people just use what’s available. The rat-model is quite new and certainly new to our group. We’ve had it for three years now... I mean each model has got its own characteristics. Like for example one model will have lots of histology and others will have lots of behavioural problems. So, if you want to test, if you want to do lots of histology we’ll do one model, if you want to do lots of behavioural stuff you would try another model’ [Mr Edwards: Cell-lab].

The good-enough animal is therefore the most practical animal: the one that is available, known and understood by experts within the particular laboratory. The practical reality is that at the bench there is not an abundance of models to choose

from, and researchers are often guided but sometimes constrainedⁱⁱⁱ by the historical and institutional tendencies of their laboratory.

Anthropomorphising and Evidential Visibility

As described, the model animal is calibrated against the target phenomenon. In other words, scientists are always faced with the issue of how good a fit the model is. In research on human medical conditions, this depends on adjudicating how well the animal model replicates the desired human characteristics. Here, we see a process of anthropomorphising taking place.

‘You can train them very well. They have paws which when you look closely at them, they’re not too dissimilar to hands you know?’ [Dr Grant: Cell-lab].

This relationship and equivalence between paw and hand is used to demonstrate the jerky, uncontrollable movements associated with HD. The same process of equivalence can be demonstrated in mouse-models of Rett Syndrome (RTT); a severe neuro-developmental disorder causing profound intellectual and physical disability, particularly in females. Linked with mutations in the MECP2 gene (Amir *et al.* 1999), a range of mouse-models of RTT have been built without functional MeCP2 (Chen *et al.* 2001; Guy *et al.* 2001; Shahbazian *et al.* 2002; Pelka *et al.* 2006). The MeCP2 deficient mice typically have a stunted body and small brain, develop tremors, and suffer from hind limb clasping, and irregular breathing, usually dying 6-12 weeks after birth (Chen *et al.* 2001; Guy *et al.* 2001). Phenotypically it is described as having similar features to children with RTT and therefore demonstration of the mouse-model is central to many conference papers delivered on the subject.

A particularly strident version of this relationship was given by one of the leading scientists in the field who had created one of the first and most widely used mouse-models. During her paper, she pronounced that the phenotypic presentation of her mouse-model and the human ‘are the same’. To demonstrate this point she played a video of a tiny mouse in a cage; its whole body shaking uncontrollably. She described this mouse as displaying some of the key features of RTT in a child; the ‘paw clasping’ represented hand stereotypies (repetitive, purposeless actions) and the uncontrollable shaking a sign of epilepsy:

‘The mouse and human phenotypes are the same, they are normal at 2-3 months, then they display fore paw clasping, and bite when handled and they can be more severe and have seizures’ [Fieldnotes].

However, what constitutes the key clinical features of RTT appears to be modified in this context. Whilst ‘hand stereotypies’ is a central clinical feature of the syndrome, breathing irregularities, epilepsy or EEG abnormalities are supportive features and are not central to a diagnosis of classic RTT (Rett Syndrome Working Group, 1988). In the mouse-model described above though the pronounced physical features displayed *become* the syndrome. The mouse’s paw-clasping and uncontrollable shaking is said to reveal the inner molecular change and the key physical signs of the syndrome:

‘Why Retts? Well we have mice, it has symptoms which can be expressed in the mouse e.g. seizures, abnormal gate, breathing. So not only at a cellular level in the brain but also physically- seizures, altered gait- it’s got the phenotype as well’ [Fieldnotes].

Furthermore, returning to the example of the HD mouse, the relationship between rodent-paw and human-hand changes in the context of touch. Contrary to the close association made between paw and hand, in an earlier discussion, a research technician at Cell-lab explained the function of the rat’s tongue is to feel and touch, doing some of the jobs of human hands.

‘Rats can feel and touch and search for food with their tongues. They have remarkably sensitive tongues’ [Fieldnotes].

It would therefore appear that the anthropomorphising of human traits can be selective depending on what scientists want to see, with some features selected whilst others are ignored. Nevertheless, a model appears to be deemed good-enough when it provides evidential visibility of a condition or disorder. As Mr. Edwards states, the model needs to ‘mirror’ the human disease within a particular social world to become good-enough.

‘Unless it mirrors human disease then your strategies [for] checking it aren't very good. I mean that’s from a behavioral point of view, you may look at the histology and if the histology mirrors human disease you may want to use strategies to alter the histology. But, from my point of view my projects are primarily behavioural. So it becomes good-enough when the rat starts looking like somebody with Huntington’s’ [Mr Edwards: Cell-lab].

Fundamentally, biomedical scientists need to decide how faithfully the mouse replicates standard features found in the human, and these can occur on multiple levels:

‘Well you have to look at your model on different levels. You have to look at whether it mimics the pathology of the disease, whether it mimics the behavioral and the cognitive aspects of the disease and whether when you try and treat it, it reduces the deficits that you had before. If it produces the right deficits...if it has the correct pathology and all that kind of thing and they all add in together, then you say “oh, that’s a good model”’ [Ms Whaling: Cell-lab].

As is explained in the extract, the model ‘needs to be viewed from various levels’ or, as we maintain, viewed both vertically (metaphorically) and horizontally (metonymic). Taken together the model is therefore made to work through evidential visibility.

Making the Model Work: Creating Real-Life Scenarios

In the laboratory, animals are also linked to the human by connecting the ways in which the model responds to stimuli and the ways in which the disease responds to experimental interventions. Returning to the field of HD, the quote below captures how lesion-rats mirror some phenotypic aspects of the human disease. It suggests experimental findings indicate that ‘stimulated’ rats have better functional recovery than those who have little or no stimulation, and likens this to the work of a physiotherapist treating and ‘stimulating’ patients with HD. In making this connection, the translation from the impact of stimuli on models to clinical interventions is forcefully made:

‘One of the post-docs here, who recently left, did a lot of work on environmental enrichment... He basically put...one set of rats into a cage where they didn’t have much stimulation and a set of rats where there are lots of things to play with, you see that there’s sort of best functional recovery in the ones where they’re sort of stimulated. That could be translated to the real world, like for example a physio could have some intervention where she’s sort of stimulating patients. So you sort of see parallels there’ [Dr Harris: Cell-lab].

Here the laboratory and the clinic are drawn closer together as behavioral scientists endeavour to replicate clinical interventions. On occasions equivalence is therefore not just found in the *bio-objects* created in the laboratory but also in the surrounding

milieu, since in targeted experiments the good-enough model also includes an adjustment of the animal's surroundings.

As described above, the use of a model depends on practical judgments about which features of the disease can be expressed and manipulated in the model and which aspects can be left out. It also depends on the practical work of simply making the experiment work (cf. Fujimura 1988). This is always about practical local decision making on the part of the scientists. Personal choices, local compromises and imperfections are therefore necessary trade-offs when choosing which models to work with. As Dr Grant suggests, research communities have to make do with the best model or good-enough model currently available. As an example, he compares the use of pig, which is said to more closely resemble the human in size, physiology and anatomy, with rats and mice:

‘If you want to take pigs to human disease you’ve got to test your strategies. What sounds like a really good idea initially doesn’t turn out to be. So the advantage of using the rat is, as we said really, [they are] cheap to use and [there’s] lots of them. And they work quite well... It is just an indicator that the models [we have], perhaps, don’t represent the human disease as well as you would want. But, you know it’s the best that we’ve got really’ [Dr Grant: Cell-lab].

Dr Grant explains how the HD model does not represent the disease as well as he would like. Consequently, scientists continue to model different features of HD, suggesting there is unlikely to be a perfect model. It is therefore clear that it is not just the fidelity of the mouse in representing the human ‘in general’ that is at issue (the metaphorical relationship) but also the extent to which the simulation of ‘the disease’ or symptom is a good enough one (metonymic relationship).

The training of rodents by laboratory technicians to perform behavioral tasks for experimental research is paramount to this process. In addition to the practical problems of description and measurement already discussed, deviant cases and outliers must also be managed. The removal of outliers can be troublesome for technicians and scientists. For example, at what point should they remove what they consider to be an under or over performing mouse from an experiment? Furthermore, to what extent can the under/over performance be explained by the manipulation of the gene (the disease)

or by an uncooperative animal? Here the rodent is reintroduced into the situation as an animal as well as some abstraction of a disease.

‘How are they trained? Actually they're quite good in that they do learn very quickly. For a situation such as the balance beam where it's on a slope, you put them at the bottom of the slope, it seems mice naturally go upwards. So they'll just turn round and they'll walk upwards and then, to encourage them, you put a cage mate at the top so they've got something to go to... Rotor rod you just do - gradually build up the speed. They get it very quickly. Whereas water maze, I don't actually teach them to swim but you kind of dip them in the water... You sit them on a platform which is out of the water and with a big flag on it and you kind of dip them at the side of the platform so they go straight to the platform because they see the flag. You just encourage them to swim further and further away very slowly, and if they do panic, take them out, give them a bit of a dry and then try again, slowly’ [Ms Johnson: Cell-lab].

To model is also to control and make uniform. There is however a long-standing tension between the desire to mechanise and standardise and the individual specificities of animal care. Standardisation is an attempt to remove local variations from modern science (Birke, Arluke and Michael 2007). Nevertheless, during the pilot stage of each test, local decisions *are* made on the aptitude and ability of particular animals to perform the tasks; if classified as unfit, animals may be rested from an experimental practice or taken out of the study altogether.

‘If I started testing a mouse in water maze and it wasn't doing very well compared to what it normally does, I may stop it after four trials in that day's data, feed it up and then start it again the next day. So I guess, in the end, I'll probably just delete all the mouse's data because it obviously wasn't right in the first place or things like that. But I don't think I've ever been in a situation where I've actually needed to do that really, just like a couple of times. You obviously get the odd mouse that's died during the evening or something... I don't think it ever really makes a difference to the results. I've never known a mouse that's freaked out enough to not give good results, if that makes sense’ [Ms Johnson].

Although uncommon, animals that cannot produce ‘good data’ despite ‘encouragement’ are removed from study samples. As a research technician explains, if the animals are not able to perform prior to the intervention they are not put through surgery and do not become a disease model.

‘Either you try and encourage them a bit to do the tests or if they really won't do anything you just have to leave them out at the end of the day. What you normally do is test them prior to having a lesion. If they won't do anything

then you leave them out at that stage and you don't put them through the surgery and that kind of thing' [Ms Johnson: Cell-lab].

Nevertheless, this does not completely solve the problem of whether a poor experimental performance by a rodent following the insertion of a lesion is the result of the surgical intervention or simply an "off" day for the rodent.

To establish the continuity of experimentation it is therefore essential to have a stock of mice with the same genetic pattern. However, successive breeding can cause deterioration in the familial line, with progressive litters of mice becoming 'sicker and sicker'. The result is that the stock of mice may be too sick or die before an experiment can be run. A post-doc at TransLab explains the difficult balance between having a stock of mice with the 'correct' amount of features that approximate disease, and being able to run a successful experiment:

'Experiment-wise my main restriction is my animals. You know about my sick animals, that we have our knock-down mice that are very small and have no hair, and they die very quickly. So my work is kind of restricted as to how many of them are alive at the time of the experiment, or I have to plan my weeks around when they are born and when they will be used... So that's a big influence... They were too young last week [to carry out the experiment], they would be ideal this week but they're dead. We never used to lose them this early, we were losing them at just before four weeks, but now they are dying at three weeks and they're not back crossed anymore' [Dr Harold: TransLab].

Here we have an example of how scientists need to construct a good enough, robust model. There is a danger in going too far: although increased molecular manipulation may be a closer representation of the disease, the mice can become too sick to function as a working model.

A workable model does not, therefore, retain all potentially relevant features. Although visual evidence can be matched in the model organisms, other aspects do not necessarily correspond. Induced disease in rodents, for instance, may not retain the developmental nature of medical conditions in humans:

'I suppose the only downside of some of the models we use is quite often we use them, start off with a healthy animal and we give it a lesion and then it immediately becomes unhealthy, immediately has a disease, whereas it generally progresses in humans and patients' [Ms Whaling: Cell-lab].

When modelling neurodegenerative disorders, there is limited success in replicating the degenerative nature of human disease or how the diseased human interacts with their social environment. The dependence on experimentation means that the environment and temporal development may be stripped out.

Finally, the adequacy of a disease model depends on the description of the disease. It is not just the animal that is the model – it is also the disease category, which is itself an ideal-type. Researchers therefore need to determine what counts as the essential features of the disease itself. The process of calibration therefore is not a one-way process. In order to generate adequate, standardised models, the human medical condition itself requires calibration. If the model is intended to replicate the disease, then the disease also requires standardisation. This is from Cell-lab:

‘Professor Endacott presents a split slideshow. On the left are some rodent behavioural tests being performed in the laboratory. While on the right of the screen are HD patients performing similar tests in the clinic. He explains how the behavioural tests performed in the laboratory have influenced the behavioural tests in the clinic’ [Fieldnotes].

The switch from human HD to rodent HD is apparently as smooth as the switch between rodent HD and human HD. The same is true for RTT: the representativeness of the mouse-models used within this field was never called into question at conferences and it was common for speakers to seamlessly switch between the description of their mouse-models and the child with RTT. An example is given below where throughout her presentation, a Japanese clinical expert described by a delegate as ‘the Japanese Rett queen’ switched fluidly between the animal and child. She repeatedly emphasised that the knock-out mice she used had clinical and behavioral features of the syndrome that were ‘similar to the girls’; the mouse ‘replicates’ the syndrome.

“‘We are looking at the hippocampus and neuron developments in the brain of *mecp2*-null mice”. She describes the clinical features of the mice without *MECP2* as displaying key clinical features to girls with the syndrome “these mice are similar to the girls, at 3-5 weeks they are pre-symptomatic, at 6-9 weeks they are symptomatic, claspings, ataxia, seizure like behavior. So the *mecp2* mouse behaviorally replicates the delayed onset and symptoms observed in human patients”” [Fieldnotes].

The reflexive relationship between rodent and human therefore not only redefines the disease it can also become the disease.

Discussion

The problem for all modellers is equivalent to the phenomenon of experimental regress (Collins 1985). We cannot know if an experiment is appropriately designed and conducted until it yields ‘the right’ results; yet we simultaneously use those experiments to derive the results in the first place. In the same way, scientists cannot know how good a model is until it predicts the ‘right’ outcome, but cannot know the right outcome in the absence of accurate models. Like all problems of regress (Collins 1985; Stephens, Lewis and Atkinson in press), it reflects the fact that in practice there cannot be a gold-standard that is independent of the procedures used to make observations, measurements, predictions, or estimates. Consequently, when we turn to the use of animal models in medical research, we encounter precisely the same issues. Scientists are confronted with making decisions concerning the adequacy of the model itself, the interpretation of their results, and their extrapolation to the target medical condition.

In practice, regress is not a matter of anxiety or uncertainty on the part of scientists. Practical scientific work rests on the good-enough principle. Animal models do not have to be perfect. They are selected because they furnish more-or-less adequate representations of certain key features. Judgments of adequacy depend on disciplinary preferences and local cultures, as well as the specific research problem at hand. Therefore, if animal models are pervasive in contemporary biomedical research, so too are the judgments that inform their choice and use. We have stressed how the choice of animal species reflects both practical and representational issues. Mice and rats provide complementary experimental opportunities, as well as furnishing different practical affordances. But their relevance for the specific medical conditions under investigation is not a given. It rests on further processes of judgment and practical action in the laboratory. The scientists need to adjudicate whether and how the lab-animal adequately represents specific features of the target disease in humans, and therefore whether the model is good-enough for all practical purposes. As we have also shown, this is not a simple, one-way process. The key features of the disease itself have to be aligned with symptomatic and behavioural features of the rodent model.

References

- Agamben G (2004) *The Open: Man and Animal*. Stanford CA: Stanford University Press.
- Amir RE (1999) Rett Syndrome is caused by Mutations in X-linked MECP2, encoding Methyl-CpG-binding protein 2. *Nature Genetics* 23: 185-188.
- Birke L, Arluke A and Michael M (2007) *The Sacrifice: How Scientific Experiments Transform Animals and People*. West Lafayette: Purdue University Press.
- Borges JL (1988) (trans A Hurley) *Collected Fictions*. New York: Penguin Books.
- Brown N and Michael M (2001) Switching between science and culture in transpecies transplantation, *Science, Technology and Human Values*, 26, 1: 3-22.
- Brown N and Michael M (2004) Risky creatures: Institutional species boundary change in biotechnology regulation, *Health, Risk and Society*, 6: 207-22.
- Chen RZ, Akbarian S, Tudor M and Jaenisch M. (2001) Deficiency of Methyl-CpG-Binding Protein-2 in CNS Neurons Results in a Rett-like Phenotype in Mice. *Nature Genetics* 27: 327-331.
- Collins HM (1985) *Changing Order: Replication and Induction in Scientific Practice*. London: Sage.
- Clause BT (1993) The Wistar Rat as a Right Choice: Establishing Mammalian Standards and the Ideal of a Standardized Mammal. *Journal of the History of Biology* 26: 329-349.
- Daston L and Mitman B (2005) *Thinking with Animals: New Perspectives on Anthropomorphism*. NY: Columbia University Press.
- Davis JA, Naruse S, Chen H, Eckman C, Younkin S, Price DL et al. (1998) An Alzheimer's Disease-Linked PS1 Variant Rescues the Development Abnormalities of PS1-Deficient Embryos. *Neuron* 20(3): 603-609.
- Delamont S, Atkinson P and Parry O (2000) *The Doctoral Experience*. London: Falmer Press.
- Endersby J (2007) *A Guinea Pig's History of Biology: The Plants and Animals who Taught us the Facts of Life*. London: William Heinemann.
- Featherstone K and Atkinson P (2011) *Creating Conditions: The Making and Remaking of a Genetic condition*. London: Routledge.
- Fox-Keller E (2000) Models of and Models for: Theory and Practice in Contemporary Biology. *Philosophy of Science* 67 (Supplement): S72-S86.

- French RD (1975) *Antivivisection and Medical Science in Victorian Society*. London: Princeton University Press
- Friese C and Clarke A (2012) Transposing Bodies of Knowledge and Technique: Animal Models at Work in Reproductive Science. *Social Studies of Science* 42(1): 31-52.
- Fujimura J (1988) The Molecular Biological Bandwagon in Cancer Research: Where Social Worlds Meet. *Social Problems* 35: 261-283.
- Garfinkel H and Sacks H (1970) The Formal Properties of Practical Action. In McKinney JC and Tiryakian EA (eds) *Theoretical Sociology*. New York: Appleton-Century-Crofts.
- Guy J, Hendrich B, Holmes M, Martin JE, Bird A (2001) A Mouse Mecp2-null Mutation causes Neurological Symptoms that Mimic Rett Syndrome. *Nature Genetics* 27: 322-326.
- Haraway D (1997) *Modest_Witness@Second Millenium: Female Man_Meets_OncoMouse*. London: Routledge.
- Harrington J and Stephens N (2010) A Social Science View on the FRAME Symposium: Identities and Networks. *ATLA: Alternatives To Laboratory Animals* 38(Supplement): 101-104
- Holmberg T and Ideland M (2009) Transgenic Silences: The Rhetoric of Comparisons and Transgenic Mice as Ordinary Treasures. *BioSocieties* 4: 165-181.
- Knight A (2011) *The Cost and Benefits of Animal Experimentation*. Basingstoke: Palgrave Macmillan.
- Kohler R (1994) *Lords of the Fly: Drosophila Genetics and the Experimental Life*. Chicago: University of Chicago Press.
- Lewis J and Atkinson P (2011) The Surveillance of Cellular Scientists' Practice. *BioSocieties* 6(4): 381-400.
- Lévi-Strauss C (1963) (trans. R Needham) *Totemism*, London: Merline Press.
- Lynch M (1988) Sacrifice and the Transformation of the Animal Body into a Scientific Object: Laboratory Culture and Ritual Practice in the Neurosciences. *Social Studies of Science* 18(2): 265-289.
- Michael M (2001) Technoscientific bespoking: Animals, publics and the new genetics, *New Genetics and Society*, 20, 3: 205-24.
- Mulkay MJ and Gilbert N (1984) *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*. Cambridge: Cambridge University Press.

- Paton W (1984) *Man and Mouse: Animals in Medical Research*. Oxford: Oxford University Press.
- Pelka GJ, Watson, CM, Radziewicz T, Hayward M, Lahooti H, Christodoulou J et al. (2006) Mecp2 Deficiency is Associated with Learning and Cognitive Deficits and Altered Gene Activity in the Hippocampal Region of Mice. *Brain* 129(4): 887-898.
- Radar K (1998) The Mouse People': Murine Genetics Work at the Bussey Institution, 1909-1936. *Journal of the History of Biology* 31: 327-354.
- Radar K (2004) *Making Mice: Standardizing animals for American Biomedical Research, 1900-1955*. Princeton: Princeton University Press.
- Rett Syndrome Diagnostic Criteria Working Group (1988) Diagnostic Criteria for Rett Syndrome. *Ann Neurol* 23: 425-428.
- Ritvo H (1987) *The Animal Estate*. Harvard MA: Harvard University Press.
- Schütz A (1967) *The Phenomenology of the Social World*. Evanston Ill: Northwestern University Press.
- Shahbazian M, Young J, Yuva-Paylor L, Spencer C, Antalffy B, Noebels J et al. (2002) Mice with truncated MeCP2 Recapitulate Many Rett Syndrome Features and Display Hyperacetylation of Histone H3. *Neuron* 35(2): 243-254.
- Shapiro K (2002) A Rodent for your Thoughts: The Social Construction of Animal Models. In Henninger-Voss M. (ed) *Animals in Human Histories*. Rochester, University of Rochester Press, 439-469.
- Star SL (1983) Simplification in Scientific Work: An Example from Neuroscience Research. *Social Studies of Science* 13: 205-228.
- Star SL and Griesemer JR (1989) Institutional Ecology, Translations and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology 1907-39. *Social Studies of Science* 19: 387-420.
- Stephens N, Lewis J and Atkinson P (In Press) Closing the Regulatory Regress: GMP Accreditation in Stem Cell Laboratories. *Sociology of Health and Illness*.
- Vermeulen N, Tamminen S and Webster A (eds) (2012) *Bio-Objects: Life in the 21st Century*. Farnham: Ashgate.
- Wanga Y, Lui W, Wada E, Murata M, Wada K, Kanazawa I (2005) "Clinico-Pathological Rescue of a Model Mouse of Huntington's Disease by siRNA". *Neuroscience Research* 53(3): 241-249.

Winnicott, D. (1953). Transitional objects and transitional phenomena, *International Journal of Psychoanalysis*, 34: 89-97.

ⁱ *Calibration* is a type of comparison between measurements, one of known correctness and another made in as similar a way as possible to the first.

ⁱⁱ All scientist's names are pseudonyms.

ⁱⁱⁱ Bench scientists could be said to be *bannistered* (both guided and constrained) by the traditional preferences of their laboratory.