Using Robots to Model Mental Disorders*

Matthew Lewis and Lola Cañamero

Embodied Emotion, Cognition and (Inter-)Action Lab,
University of Hertfordshire, UK
{M.Lewis4|L.Canamero}@herts.ac.uk

Abstract

We are currently at a point where the use of robots to model human mental disorders is possible, and this capability will only increase. By considering the lessons learned from animal models, we argue that robot models of human mental disorders can complement existing approaches in mental health research.

1 Introduction

We are currently at a point where the use of robots to model human mental disorders is possible, and this capability will only increase. It is true that current robots are a long way from reproducing the capabilities of human beings, and it might be seen as insulting that complex human disorders could be modelled with such distant approximations. However, for many years, animal models of mental disorders have been used for human mental disorders. We therefore take stock, and examine how we should progress in the use of robot models.

2 Models for Mental Disorders

In his section, we define and briefly discuss four types of models for mental disorders.

A *conceptual model of a mental disorder* is a theoretical construct that links underlying causes (etiology), either proposed or observed, with observed symptoms and correlates. A conceptual model serves as a framework for understanding, and should have explanatory and predictive power with respect to the condition being modelled. For example, Shafran [2] gives five models for Obsessive-Compulsive Disorder based variously on: a faulty appraisal of normal intrusive thoughts, an excessive emphasis on control of one's own thoughts, and a self-perpetuating mechanism of checking behaviour. There is not necessarily a need for one "true" model, and different models may be complementary, having different emphases, levels of abstraction or uses

An *animal model of a mental disorder* is a non-human animal used to study brain–behaviour relations with the goal of gaining insight into, and to enable predictions about, these relations in humans [3]. Animal models may be induced by genetic manipulation, drugs, or by environmental manipulation. Alternatively, they may be naturally occurring. They have the advantage that they model a complete system (organism and environment), and they use a real animal, hence a real nervous system. However, there are limits to how closely a non-human animal can be used to model human mental disorders. There are also ethical issues associated with animal experimentation.

A *computational model of a mental disorder* is a realisation, or partial realisation, of a theoretical model in a computer. The emerging field of computational psychiatry includes within its scope the development of computational models of psychiatric disorders [4]. These models have the advantage that, by their nature, they are highly specified and so any results should be replicable and can be analysed in detail. However, due to the complexity of implementing such a model, they are typically only partial implementations (e.g. of a neurological subsystem) or they work at a relatively high level of abstraction. In addition, they do not necessarily include any behavioural element, a true closed-loop interaction with the environment, or the effects of contextual and environmental elements.

A *robot model of a mental disorder* include an embedded realisation of a conceptual model in an embodied, interacting robot and its environment. This introduces elements that are present in animal models, but which purely computational models lack.

Thus far, there have been relatively few explicit robot models of mental disorders, with work Yamashita and Tani [5] being one of the rare examples. However, work in biologically-inspired autonomous robots is linked, since models of behaviour can also potentially serve as models of pathological behaviour. For example, our previous work on pleasure and behaviour [6] has links to addiction.

*Arguments presented in this paper are summarised from [1].

3 Learning from Animal Models

In order to maximise the potential of robot models, it is instructive to learn from what has been learned from many years of using animal models.

Animal models can be evaluated and validated along four criteria [7]: *face validity* (phenomenological similarity), *construct validity* (validity of the underlying mechanism), *predictive validity* (whether, for example, the model can predict effective interventions) and *reliability* (whether results are robust and reproducible).

Thinking about how these criteria relate to robot models, we make the following observations.

Face validity can be easy to achieve in robot models: a robot can simply be programmed to behave in a pathological fashion. However, without construct validity, in this case face validity becomes meaningless. Therefore, we should not work from the direction of face validity, but use it as a validation criterion.

If robot models are based on theoretical models, then construct validity either comes from this development process, or, if the model is hypothetical, then the robot model serves are a test of the model itself.

Predictive validity is important in the context of developing clinical interventions, it is therefore something that should be targeted by robot models in order to maximise their contribution to translational research.

Reliability is something that to some extent comes naturally from using robot models: it is possible to replicate experiments. However, since it may be the theoretical model that is of interest, not the specific implementation, efforts should also be made to produce alternative implementations of the same theoretical model, to demonstrate that the behaviour of a model is not due to a detail of implementation (you may then ask if that implementation detail should be included in the theoretical model).

Finally, with robot models we can easily "look inside" our model to examine cognitive processes as they happen. While this can be done to some extent with animals, it is limited, and the techniques may be invasive. By doing this we may gain insights into how mental illness is experienced from the inside (symptoms such as confusion, alienation from one's own actions, paranoia), going beyond the behavioural aspects that are the most readily examined in animal models.

4 Advantages of Robot Models

One significant advantage of computational models, including robot models, over animal models is that it allows precise operationalisation and explicit implementation of an underlying theoretical model. While models may be implemented in animals, experimenters may not always have enough control over the biology to implement it a theoretical model as precisely as desired.

In addition to this, robot models have vastly reduced ethical issues compared to animal models.

It is also easier to control environmental confounding factors that are not part of the model with robots than with animals. For example, experimental results in animals have been unexpectedly affected by experimenter smells [8]. Such effects can be limited in robots since we have more knowledge of, and control over, their sensory systems.

Bearing in mind the criterion of predictive validity, above, robot models may also allow us to test interventions. For example, they may test simulated drugs that have a targeted effect on one element of the model, such as a receptor, when no such chemical is yet known, or when known chemicals have undesirable side effects. However, in order to do this, potential targets (pharmaceutical or otherwise) needs to be part of the underlying model.

Robot models can also take advantage of their embodied aspect. Mental disorders frequently have have embodied aspects, such as a distorted sense of the body, and some therapeutic interventions are also embodied (e.g. exercise, art therapy). Symptoms of mental disorders may also be partly due to dysfunctions in the perception-action loop. By taking this into account, robot models can be used as tools where purely computational models are not suitable.

5 Conclusions

Robot models for mental disorders are a promising direction for research, to be used in conjunction with existing animal and purely computational models. However, in order to achieve their potential thought needs to be given to how they are used. Implementation of existing theoretical models has promise, but these models need to be assessed in terms of face validity (of both phenotypes and "hidden" endo-phenotypes), predictive validity and reliability. With predictive power, robot models can then contribute to translational research of treatments.

References

[1] M. Lewis and L. Cañamero, "Robot models of mental disorders," in *Proc. 7th International Conference on Affective Computing and Intelligent Interaction, Workshops and Demos (ACIIW 2017)*, (San Antonio, TX), pp. 193–200, IEEE, 2017.

[2] R. Shafran, "Cognitive-behavioral models of OCD," in *Concepts and Controversies in Obsessive-Compulsive Disorder* (J. S. Abramowitz and A. C. Houts, eds.), pp. 229–260, Boston, MA: Springer, 2005.

[3] F. J. van der Staay, "Animal models of behavioral dysfunctions: Basic concepts and classifications, and an evaluation strategy," *Brain Research Reviews*, vol. 52, no. 1, pp. 131–159, 2006.

[4] Q. J. M. Huys, T. V. Maia, and M. J. Frank, "Computational psychiatry as a bridge from neuroscience to clinical applications," *Nature Neuroscience*, vol. 19, no. 3, pp. 404–413, 2016.

[5] Y. Yamashita and J. Tani, "Spontaneous prediction error generation in schizophrenia," *PLoS ONE*, vol. 7, no. 5, pp. 1–8, 2012.

[6] M. Lewis and L. Cañamero, "Hedonic quality or reward? A study of basic pleasure in homeostasis and decision making of a motivated autonomous robot," *Adaptive Behavior*, vol. 24, pp. 267–291, 2016.

[7] P. Willner, "Validation criteria for animal models of human mental disorders: Learned helplessness as a paradigm case," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 10, no. 6, pp. 677–690, 1986.

[8] R. E. Sorge, L. J. Martin, K. A. Isbester, S. G. Sotocinal, S. Rosen, A. H. Tuttle, J. S. Wieskopf, E. L. Acland, A. Dokova, B. Kadoura, P. Leger, J. C. S. Mapplebeck, M. McPhail, A. Delaney, G. Wigerblad, A. P. Schumann, T. Quinn, J. Frasnelli, C. I. Svensson, W. F. Sternberg, and J. S. Mogil, "Olfactory exposure to males, including men, causes stress and related analgesia in rodents," *Nature Methods*, vol. 11, no. 6, pp. 629–632, 2014.