



Fallacies of Mice Experiments

Erik De Schutter¹

Published online: 2 April 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Recently, one of my PhD students complained that while presenting her poster, the scientific relevance of her modeling work was questioned aggressively by an experimentalist. So even at the end of the second decade of this millennium, theoreticians still have to justify the relevance of their work towards understanding the brain.¹ In this editorial I want to demonstrate that, unfortunately, such high standards are not always applied to experimental work, in particular in mice.

A first example is the recent news that two major Phase II trials for Alzheimer drugs have been canceled.² These trials of humanized anti-amyloid- β monoclonal antibodies were based on the convergence of two sets of data: genetic risk for Alzheimer disease in humans indicating the importance of amyloid metabolism and extensive studies in transgenic mice.³ Many studies have shown that transgenic mice expressing gene mutations associated with human familial Alzheimer disease progressively develop brain amyloid plaques and memory deficits.⁴ Immunization against amyloid- β peptide rapidly reversed memory defects in some transgenic models^{3,5} leading to the subsequent clinical trials. In hindsight, is the failure of such treatments in patients surprising? There were

already papers suggesting that the antibodies did not work in all mouse models of Alzheimer disease; noticeably these papers were published in lower impact journals.⁶ But I want to argue that, in general, mouse models are not very predictive of human disease.

This conjecture is based on a preceding sequence of failures with drugs derived from mouse and other animal studies: those used for the treatment of septic shock. Of the 69 clinical studies performed in 1982–2013 that were analyzed in,⁷ only 8 resulted in some benefits and 4 actually harmed the patients, all others showed no effect. All of these studies used compounds that were beneficial in mice, baboons or rabbits. A simple reason of these differences may be that humans are much more sensitive to bacterial lipopolysaccharide than mice or baboons, but this has not stopped the use of these animals in drug testing. At least there were no fatal incidents in Phase I trials of septic shock drugs, as was recently the case for a new compound that was supposed to work both for Parkinson's disease and chronic pain.⁸

A rigorous reason why murine models are not predictive for septic shock was provided by a genomic study showing that changes in gene expression in mouse models to inflammatory stress have zero correlation with the corresponding gene expression changes in humans.⁹ This study was widely advertised in the popular and scientific press and generated a lot of reactions. But, unfortunately, it did not lead to a fundamental change in our approach to murine models of human diseases and, in particular, there has been little interest in applying the lessons learned from septic shock to other categories of disease models.

¹ De Schutter, E. (2008). Reviewing multi-disciplinary papers: a challenge in neuroscience? *Neuroinformatics*, 6(4), 253–255. <https://doi.org/10.1007/s12021-008-9034-x>.

² <https://www.biospace.com/article/a-long-line-of-failures-roche-drops-alzheimer-s-drug-trials/>

³ Panza, F., Lozupone, M., Logroscino, G., & Imbimbo, B. P. (2019). A critical appraisal of amyloid- β -targeting therapies for Alzheimer disease. *Nature Reviews Neurology*, 1–16. <https://doi.org/10.1038/s41582-018-0116-6>.

⁴ Karran, E., Mercken, M., & De Strooper, B. (2011). The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. 1–15. <https://doi.org/10.1038/nrd3505>.

⁵ Dodart, J.-C., Bales, K. R., Gannon, K. S., Greene, S. J., DeMattos, R. B., Mathis, C., et al. (2002). Immunization reverses memory deficits without reducing brain A β burden in Alzheimer's disease model. *Nature Neuroscience*, 5(5), 452–457. <https://doi.org/10.1073/pnas.94.4.1550>.

⁶ Mably, A. J., Liu, W., Donald, J. M. M., Dodart, J.-C., Bard, F., Lemere, C. A., et al. (2015). Neurobiology of Disease. *Neurobiology of Disease*, 82(C), 372–384. <https://doi.org/10.1016/j.nbd.2015.07.008>.

⁷ Fink, M. P. (2013). Animal models of sepsis. *Virulence*, 5(1), 143–153. <https://doi.org/10.1172/JCI116529>.

⁸ Butler D, & Callaway E. (2016) Scientists in the dark after French clinical trial proves fatal. *Nature*, 529, 263–264.

⁹ Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., et al. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences*, 110(9), 3507–3512. <https://doi.org/10.1073/pnas.1222878110>.

✉ Erik De Schutter
erik@oist.jp

¹ Computational Neuroscience Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

Some will argue that there are no alternatives to mice, though for some diseases human organoids¹⁰ may soon become the primary model system. But we should not underestimate sheer inertia, the availability of easy to get grants for translational research and the vested interests of the industry and university centers that support murine research: all of these factors converge to ensure a bright future of the mouse model irrespective of its variable usefulness.

Returning to neuroscience, the relevance of murine models of psychiatric disease should especially be questioned. Although these mouse models are often named for the psychiatric syndrome, e.g. schizophrenia or autism, in reality they are models of a single endophenotype.¹¹ Endophenotypes are discrete behavioral traits that in combination form the whole syndrome, but of course there is no guarantee that the disease can really be decomposed in such a way. To compensate for this, it is now established practice to study several transgenic lines simultaneously based on the, somewhat naive, expectation that the relevance of an observed effect correlates with the number of transgenic lines in which it is observed. Resolving the mechanistic cause of human psychiatric diseases, which for most syndromes remains a mystery, should really be a more pressing challenge than investigating mouse endophenotypes in detail.

If mice have limited use in studying human disease, are they at least useful in understanding brain function? The technical revolution caused by optogenetic methods and imaging of genetically expressed calcium dyes has led to a rapid shift from rats to mice as the preferred experimental animal.¹² But again, a mouse is not a human and therefore, it is not the best animal to study every interesting neuroscience topic. An example is visual cortex, which is the main target of the Allen Institute's Project MindScope.¹³ It is well known that mice have low visual acuity¹⁴ and use olfaction and whisking as their main sensory input. Although it was recently shown that mice do use vision in specific behaviors,¹⁵ they are

- like rats - not binocular animals¹⁶ and - lacking pinwheels¹⁷ - their visual cortex is organized quite differently compared to primate visual cortex. Based on these differences, studying mouse visual cortex is comparative neuroscience, likely as relevant as studying the *Drosophila* visual system towards understanding human vision.

Finally, even when relevant brain functions are studied in mice, the standards used to design the behavioral component are much lower than what is common in human imaging studies.¹⁸ A basic - though inaccurate¹⁹ - neuroimaging technique is to subtract images obtained when performing a control condition (e.g. pushing a button or seeing an image) from images acquired when performing the condition of interest (e.g. pushing a specific button for a specific image). But even this is absent in mouse behavioral design. In many papers reporting on in vivo murine experiments all neural activity is implicitly assumed to be caused by the cognitive task, usually there is no attempt to decompose the behavior and attribute activity to specific subcomponents.

An example of how this lack of sophistication in study design and analysis leads to confusing results can be found in a recent study of the cortico-cerebellar loop.²⁰ At one level this study is groundbreaking because it demonstrates that activity in cerebellar nuclei is required for motor planning, specifically using sensory discrimination to plan a future directional licking movement. The mice have to use their whiskers to locate a pole relative to their fixed head, wait, and then perform either left or right licking. It was known that such a task requires persistent activity in the frontal cortex during the waiting period - akin to working

¹⁰ Paşca, S. P. (2019). Assembling human brain organoids. *Science (New York, NY)*, 363(6423), 126–127. <https://doi.org/10.1038/d41586-018-04813-x>.

¹¹ Salgado, J. V., & Sandner, G. (2013). A critical overview of animal models of psychiatric disorders: challenges and perspectives. *Revista Brasileira De Psiquiatria*, 35(suppl 2), S77–S81. <https://doi.org/10.1093/schbul/sbn176>.

¹² Kim, C. K., Adhikari, A., & Deisseroth, K. (2017). Integration of optogenetics with complementary methodologies in systems neuroscience. *Nature Reviews Neuroscience*, 18(4), 222–235. <https://doi.org/10.1038/nrn.2017.15>.

¹³ Hawrylycz, M., Anastassiou, C., Arkhipov, A., Berg, J., Buice, M., Cain, N., et al. (2016). Inferring cortical function in the mouse visual system through large-scale systems neuroscience. *Proceedings of the National Academy of Sciences*, 113(27), 7337–7344. <https://doi.org/10.1109/TPAMI.2007.56>.

¹⁴ Huberman, A. D., & Niell, C. M. (2011). What can mice tell us about how vision works? *Trends in Neurosciences*, 34(9), 464–473. <https://doi.org/10.1016/j.tins.2011.07.002>.

¹⁵ Hoy, J. L., Yavorska, I., Wehr, M., & Niell, C. M. (2016). Vision Drives Accurate Approach Behavior during Prey Capture in Laboratory Mice. *Current Biology*, 26(22), 3046–3052. <https://doi.org/10.1016/j.cub.2016.09.009>.

¹⁶ Wallace, D. J., Greenberg, D. S., Sawinski, J., Rulla, S., Notaro, G., & Kerr, J. N. D. (2013). Rats maintain an overhead binocular field at the expense of constant fusion. *Nature*, 498(7452), 65–69. <https://doi.org/10.1038/nature12153>; Payne, H. L., & Raymond, J. L. (2017). Magnetic eye tracking in mice. *eLife*, 6. <https://doi.org/10.7554/eLife.29222>.

¹⁷ Ohki, K., Chung, S., Kara, P., Hübener, M., Bonhoeffer, T., & Reid, R. C. (2006). Highly ordered arrangement of single neurons in orientation pinwheels. *Nature Publishing Group*, 442(7105), 925–928. <https://doi.org/10.1093/cercor/13.3.225>; Bonin, V., Histed, M. H., Yurgenson, S., & Reid, R. C. (2011). Local Diversity and Fine-Scale Organization of Receptive Fields in Mouse Visual Cortex. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 31(50), 18,506–18,521. <https://doi.org/10.1523/JNEUROSCI.2974-11.2011>.

¹⁸ Amaro, E., Jr., & Barker, G. J. (2006). Study design in fMRI: Basic principles. *Brain and Cognition*, 60(3), 220–232. <https://doi.org/10.1016/j.bandc.2005.11.009>.

¹⁹ Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., & Dolan, R. J. (1996). The trouble with cognitive subtraction. *NeuroImage*, 4(2), 97–104. <https://doi.org/10.1006/nimg.1996.0033>.

²⁰ Gao, Z., Davis, C., Thomas, A. M., Economo, M. N., Abrego, A. M., Svoboda, K., et al. (2018). A cortico-cerebellar loop for motor planning. *Nature Publishing Group*, 1–27. <https://doi.org/10.1038/s41586-018-0633-x>.

memory²¹ - and the authors showed recently that the thalamus is required for this persistent activity^{21,22} It is through the thalamus that the cerebellum interacts with frontal cortex. Up till now, this summary of²⁰ describes an interesting mice experiment that is consistent with the increasing evidence for a cognitive function of the cerebellum.²³

However, the surprise of the study is that specifically the fastigial nucleus is required for this task. At first view this is completely unexpected, because the fastigial nucleus is the phylogenetically oldest cerebellar nucleus and is highly conserved in mammalian evolution.²⁴ Conversely, the dentate nucleus is known to project extensively to non-motor areas in cortex,²⁵ is greatly expanded in humans²⁶ and is generally assumed to be the structure involved in cognitive tasks of the cerebellum. The authors do not discuss why the fastigial nucleus is activated in their task and, because it was so unexpected, several colleagues in the cerebellar field do not believe the results reported in.²⁰ However, if one decomposes the behavior into its different components it becomes much easier to explain. Discriminating the position of a structure close to the head of the animal to decide about a movement is a task for which activation of the fastigial nucleus is not surprising, because one of its main functions is axial and proximal motor control.²³ In other words, it is probably the sensory component of the task, not the short-term memory component, that causes the fastigial nucleus to

be involved and changes to the sensory component may cause other cerebellar nuclei to become necessary for the frontal cortex activation, as suggested by an unpublished study.²⁷ But this message is not conveyed by the paper.

In this editorial I discussed only a few examples of the lack of introspection and quality control that unfortunately affects much neuroscience research and I focused on the experimental side. Obviously more examples can be found, also in computational neuroscience,²⁸ and many of the challenges are not specific to neuroscience. As mentioned, these problems are exacerbated by the inertia of established scientific organizations and the group-thinking that guide many of the choices scientists make. To combat this, one has to look both inward and outward. Question yourself, are you trying to solve big questions about the healthy or diseased brain, or just collecting more data that will not contribute much to better understanding or effective treatment? Look beyond the boundaries of your field - I extensively did so in this editorial - and use this knowledge to improve your scientific planning. To return to the initial point, the experimentalist who intimidated my student could instead have tried to understand what modeling can contribute to his science.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

²¹ Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science (New York, NY)*, 173(3997), 652–654.

²² Guo, Z. V., Inagaki, H. K., Daie, K., Druckmann, S., Gerfen, C. R., & Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. *Nature Publishing Group*, 545(7653), 181–186. <https://doi.org/10.1038/nature22324>.

²³ Schmähmann, J. D. (2019). Neuroscience Letters. *Neuroscience Letters*, 688, 62–75. <https://doi.org/10.1016/j.neulet.2018.07.005>.

²⁴ Zhang, X.-Y., Wang, J.-J., & Zhu, J.-N. (2016). Cerebellar fastigial nucleus: from anatomic construction to physiological functions. *Cerebellum & Ataxias*, 1–10. <https://doi.org/10.1186/s40673-016-0047-1>.

²⁵ Strick, P. L., Dum, R. P., & Fiez, J. A. (2009). Cerebellum and Nonmotor Function. *Annual Review of Neuroscience*, 32(1), 413–434. <https://doi.org/10.1146/annurev.neuro.31.060407.125606>.

²⁶ Sultan, F., Hamodeh, S., & Baizer, J. S. (2010). The human dentate nucleus: a complex shape untangled. *Neuroscience*, 167(4), 965–968. <https://doi.org/10.1016/j.neuroscience.2010.03.007>.

²⁷ <https://www.biorxiv.org/content/10.1101/335703v2.full>

²⁸ Chen, W., & De Schutter, E. (2017). Time to Bring Single Neuron Modeling into 3D. *Neuroinformatics*, 1–3. <https://doi.org/10.1007/s12021-016-9321-x>.