

A Knockout Experiment: Disciplinary Divides and Experimental Skill in Animal Behaviour Genetics

NICOLE C. NELSON*

Department of the History of Science, University of Wisconsin – Madison,
1225 Linden Drive, Madison, WI 53706, USA

Abstract: In the early 1990s, a set of new techniques for manipulating mouse DNA allowed researchers to ‘knock out’ specific genes and observe the effects of removing them on a live mouse. In animal behaviour genetics, questions about how to deploy these techniques to study the molecular basis of behaviour became quite controversial, with a number of key methodological issues dissecting the interdisciplinary research field along disciplinary lines. This paper examines debates that took place during the 1990s between a predominately North American group of molecular biologists and animal behaviourists around how to design, conduct, and interpret behavioural knockout experiments. Drawing from and extending Harry Collins’s work on how research communities negotiate what counts as a ‘well-done experiment,’ I argue that the positions practitioners took on questions of experimental skill reflected not only the experimental traditions they were trained in but also their differing ontological and epistemological commitments. Different assumptions about the nature of gene action, eg., were tied to different positions in the knockout mouse debates on how to implement experimental controls. I conclude by showing that examining representations of skill in the context of a community’s knowledge commitments sheds light on some of the contradictory ways in which contemporary animal behaviour geneticists talk about their own laboratory work as a highly skilled endeavour that also could be mechanised, as easy to perform and yet difficult to perform well.

Keywords: Behaviour Genetics, Molecular Biology, Mouse Models, Transgenic Animals, Skill, Replication

In the summer of 1992, two research groups presented a series of experiments on the genetics of learning and memory that caused ‘quite a stir’ at the annual Cold Spring Harbor Symposium on Quantitative Biology.¹ The topic of that year’s annual symposium

* Email address for correspondence: nicole.nelson@wisc.edu

Thanks to Harald Kliems for his research assistance and to Thomas Schlich and Nicholas Whitfield for creating a welcoming space for cross-disciplinary discussions on skill and for providing helpful comments throughout the revision process. I would also especially like to thank Thomas Broman for reading multiple versions of this paper and offering thoughtful suggestions on how best to frame it. This project was funded by a Doctoral Dissertation Improvement Grant from the National Science Foundation (Award No. SES 0749635) and a Doctoral Fellowship by the Social Sciences and Humanities Research Council of Canada.

¹ Marcia Barinaga, ‘Knockouts Shed Light on Learning’, *Science*, 257, 5067 (1992), 162.

was 'The Cell Surface', but the papers on memory had generated such excitement that the conference organisers created a separate session, unrelated to the cell surface, so that two papers could be presented back-to-back.² The researchers had used a suite of techniques for manipulating DNA to produce 'knockout' mice: by taking advantage of homologous recombination (where sections of genetic material with similar sequences are exchanged between two strands of DNA), researchers could 'knock out' specific genes in mice by replacing functional genes with non-functional copies in mouse stem cells and inserting the modified cells into a developing mouse embryo. Susumu Tonegawa's laboratory at the Massachusetts Institute of Technology (MIT) targeted a gene encoding calcium-calmodulin-dependent kinase II alpha (α -CaMKII), an enzyme that previous studies had implicated in memory formation.³ Eric Kandel's group at Columbia University took a similar approach, studying mice lacking four different tyrosine kinase receptors that were highly expressed in the hippocampus, an area of the brain thought to be important in long-term memory formation.⁴

Knockout experiments created an opening for claims about genes and behaviour with an unprecedented degree of specificity. While prior methods might have provided evidence for general claims about the molecular mechanisms of memory formation or the heritability of psychiatric disorders, researchers argued that knockout techniques would allow them to drive those claims all the way down to the level of the individual gene. Kandel's laboratory, eg., had previously experimented with drugs that inhibited tyrosine kinase activity to implicate these enzymes in memory formation but noted these drugs often only partially inhibited the action of the kinases under study and disrupted the activity of other kinases as well.⁵ For Kandel's group, knockout techniques provided a way to distinguish between the activities of different kinases and make claims about individual enzymes.⁶ Tonegawa and colleagues similarly remarked that studies comparing different inbred strains of mice had been able to generate correlations between brain levels of particular kinases and performance on learning and memory tasks, but that the differences identified through those experiments were 'clearly not the result of differences in a single gene'.⁷ They argued that their experiments, in contrast, demonstrated the 'selective but drastic' impact that a single genetic change could have.⁸ Other researchers shared their enthusiasm about the technique's potential for making convincing links between specific genes and complex behavioural deficits. A neuroscientist interviewed for a news article on the Cold Spring Harbor Laboratory papers commented that the most exciting part for him was that Tonegawa and Kandel's results demonstrated that it was indeed possible to 'analyse a phenomenon as complex as [long term potentiation] via this knockout approach'.⁹

² Eric Kandel, *Search of Memory* (New York: WW Norton, 2006), 291.

³ Tonegawa's paper was published shortly after the Cold Spring Harbor symposium as Alcino J. Silva, Richard Paylor, Jeanne M. Wehner and Susumu Tonegawa, 'Impaired Spatial Learning in Alpha-Calcium-Calmodulin Kinase II Mutant Mice', *Science*, 257, 5067 (1992), 206–11.

⁴ Kandel and Soriano's paper was published several months later as Seth G.N. Grant, Thomas J. O'Dell, Kevin A. Karl, Paul L. Stein, Philippe Soriano and Eric R. Kandel, 'Impaired Long-Term Potentiation, Spatial Learning, and Hippocampal Development in *fyn* Mutant Mice', *Science*, 258, 5090 (1992), 1903–10.

⁵ Thomas J. O'Dell, Eric R. Kandel and Seth G.N. Grant, 'Long-Term Potentiation in the Hippocampus is Blocked by Tyrosine Kinase Inhibitors', *Nature*, 353 (1991), 558–60.

⁶ Grant *et al.*, *op. cit.* (note 4), 1903.

⁷ Silva *et al.*, *op. cit.* (note 3), 210.

⁸ *Ibid.*

⁹ Barinaga, *op. cit.* (note 1), 162.

Experiments taking this form – where researchers created a targeted mutation and then tested the resulting mouse for differences in its behaviour – proliferated in the months and years after. These first papers were followed in a quick succession by more studies from the Tonegawa group and others, published in high-profile journals such as *Nature*, *Science*, and *Cell*. Several papers reported on the results of knocking out other kinases on the mouse's performance in learning and memory tests,¹⁰ as well as tests for other behavioural traits such as aggression, anxiety, and activity levels.¹¹ Not everyone, however, was convinced that knockout experiments clinched the debate on the role of specific enzymes such as α -CaMKII in behavioural traits. Robert Morris, who developed the eponymous Morris water maze test for learning and memory that the 1992 papers employed, pointed out that there were a number of alternative explanations that could account for the results these papers reported, and warned that 'a little learning is a dangerous thing'.¹² Another letter to the editor published in *Science* from a psychologist at the University of California, San Diego similarly took issue with the strong claim that these experiments had linked a mutation to a learning deficit and pointed out other possible interpretations.¹³ Throughout the 1990s, as more researchers began using knockouts in their laboratories, more commentaries appeared on the merits of knockout studies as a way to make claims about the molecular basis of behaviour, many of them questioning the value of this supposedly revolutionary technique. In one particularly strident letter to the editor in a 1995 issue of *Nature*, entitled 'Knockout mouse fault lines', another psychologist suggested that the entire field was being 'led into a technological cul-de-sac' by the increasing adoption of knockout techniques that were 'wholly inappropriate for resolving the issues for which [they were] intended'.¹⁴

This paper takes advantage of this intersection of different knowledge communities around a shared set of new techniques to investigate the relationship between scientific practitioners' understandings of experimental skill and their ontological and epistemological commitments, drawing from and extending on Harry Collins's work on how research communities negotiate what counts as a 'well-done experiment'.¹⁵ Through an examination of debates between molecular biologists and animal behaviour geneticists about how best to conduct and interpret knockout experiments, I aim to show that the

¹⁰ Asa Abeliovich, Richard Paylor, Chong Chen, Jeansok J. Kim, Jeanne M. Wehner and Susumu Tonegawa, 'PKC Gamma Mutant Mice Exhibit Mild Deficits in Spatial and Contextual Learning', *Cell*, 75, 7 (1993), 1263–71; Asa Abeliovich, Chong Chen, Yukiko Goda, Alcino J. Silva, Charles F. Stevens and Susumu Tonegawa, 'Modified Hippocampal Long-Term Potentiation in PKC γ -Mutant Mice', *Cell*, 75, 7 (1993), 1253–62; Atsu Aiba, Chong Chen, Karl Herrup, Christian Rosenmund, Charles F. Stevens and Susumu Tonegawa, 'Reduced Hippocampal Long-Term Potentiation and Context-Specific Deficit in Associative Learning in mGluR1 Mutant Mice', *Cell*, 79, 2 (1994), 365–75; Atsu Alba, Masanobu Kano, Chong Chen, Mark E. Stanton, Gregory D. Fox, Karl Herrup, Theresa A. Zwingman and Susumu Tonegawa, 'Deficient Cerebellar Long-Term Depression and Impaired Motor Learning in mGluR1 Mutant Mice', *Cell*, 79, 2 (1994), 377–88.

¹¹ Chong Chen, Donald G. Rainnie, Robert W. Greene and Susumu Tonegawa, 'Abnormal Fear Response and Aggressive Behavior in Mutant Mice Deficient for Alpha-Calcium-Calmodulin Kinase II', *Science*, 266, 5183 (1994), 291–4; Ming Xu, Rosario Moratalla, Lisa H. Gold, Noboru Hiroi, George F. Koob, Ann M. Graybiel and Susumu Tonegawa, 'Dopamine D1 Receptor Mutant Mice Are Deficient in Striatal Expression of Dynorphin and in Dopamine-Mediated Behavioral Responses', *Cell*, 79, 4 (1994), 729–42.

¹² Richard G.M. Morris and Mary B. Kennedy, 'The Pierian Spring', *Current Biology*, 2, 10 (1992), 511–4.

¹³ J. Anthony Deutsch, 'Spatial Learning in Mice', *Science*, 262, 5134 (1993), 760–3.

¹⁴ Aryeh Routtenberg, 'Knockout Mouse Fault Lines', *Nature*, 374, 6520 (1995), 314.

¹⁵ Harry M. Collins, *Changing Order: Replication and Induction in Scientific Practice* (London: Sage Publications, 1985); *idem*, 'The Seven Sexes: A Study in the Sociology of a Phenomenon, or the Replication of Experiments in Physics', *Sociology*, 9, 2 (1975), 205–24; *idem*, 'Son of Seven Sexes: The Social Destruction of a Physical Phenomenon', *Social Studies of Science*, 11, 1 (1981), 33–62.

positions practitioners took on these issues reflected not only the different experimental traditions they were trained in but also their understandings of the nature of gene action and of how to generate knowledge about the relationship between genes and behaviour. In the first section of the paper, I outline some of the differences between molecular biologists' and animal behaviour geneticists' approaches to studying the genetics of behaviour and trace out some of the pathways through which practitioners from these fields converged on similar experimental systems. The second section of the paper examines the positions that molecular biologists and animal behaviour geneticists took throughout the 1990s on methodological questions about how to conduct knockout experiments and relate these positions to their experimental traditions and epistemological commitments. Molecular biologists' suggestions for standardising the way that knockout mice were constructed, eg., reflected an experimental tradition in which knowledge was produced by contrasting mutants with standardised, 'wild type' organisms as well as an investment in the idea that the mechanisms of complex phenomena could be unravelled by studying them one gene at a time.

The final third of the paper uses interviews and ethnographic material to explore the relationships between ontology, epistemology, and skill in greater detail.¹⁶ Animal behaviour geneticists' retrospective accounts of the knockout debates, while not necessarily accurate in representing the positions of molecular biologists or the historical events as they unfolded, are nevertheless useful for examining the relationship between knowledge commitments and conceptions of skilful practice. Recalling the influx of new practitioners from molecular biology into the study of behaviour serves as a foil for animal behaviour geneticists to articulate what is centrally important to them in their laboratory practice. They emphasise, eg., their assumptions about the nature of gene action and the stability of behavioural tests, which, they argue, informed their stances on whether inconsistent results between laboratories were to be expected or whether they were evidence of badly conducted experiments. These different ontological and epistemological commitments, I argue, prevented researchers from being able to form a consensus on the meaning of knockout experiment data even as they came to agreement on some questions about experimental skill. The article concludes with an examination of ethnographic data on how researchers talk about skill in an animal behaviour genetics laboratory at 'Western University'.¹⁷ I argue that examining these statements in the context of the community's ontological and epistemological commitments sheds light on some of the contradictory ways in which animal behaviour geneticists talk about their own laboratory work as a highly skilled endeavour that also could be mechanised, as easy to perform and yet difficult to perform well.

¹⁶ Interviews were conducted between 2006 and 2010, and the majority of these interviews were conducted under the condition that the material would be used anonymously. This anonymity, while potentially lessening the value of these interviews for historical inquiry, was a necessary precondition for being able to gather this information because of the fears that many practitioners held about the public representation of behaviour genetics and the potential for controversy. All of the excerpts used for this paper come from interviews with researchers in North America who identified as 'behaviourists'; that is, they had spent the majority of their careers working on questions about behaviour using either genetic or neurobiological approaches.

¹⁷ The ethnographic fieldwork that this paper draws on was conducted between 2006 and 2009 in animal behaviour genetics laboratories and scientific meetings, and in particular in laboratories in the department of neuroscience on the west coast of the United States that I refer to here as 'Western University'. This research was also conducted under the condition that it would be used anonymously, and the name 'Western University' is a pseudonym.

Knockout Experiments and Intersecting Knowledge Communities

Behavioural knockout studies generated some strange scientific bedfellows in their early days, since conducting these studies required researchers to assemble skill sets that were not typically found together in scientific knowledge communities in the early 1990s. Although most of the scientists involved in the first behavioural knockout studies could ostensibly be called ‘geneticists’, in practice their work drew on very different conceptual and technical histories.

Studying the Genetics of Behaviour

Constructing a mouse with a targeted gene deletion required a very specific set of skills in manipulating DNA at the molecular level and working with cell cultures – skills central to molecular biology at the time. Hans-Jörg Rheinberger argues that there have been two important ‘shifts of assemblage’ in the history of molecular biology.¹⁸ He locates the first in the post-WWII period, with the characterisation of the structure of the double helix and the introduction of a series of biophysical techniques such as electron microscopy, X-ray structure analysis, and radioactive tracing, and the second in the 1970s, with a turn towards genetic engineering and the introduction of cellular ‘tools’ such as restriction enzymes, plasmids, and polymerases. Each of these shifts brought molecular biologists into contact with new knowledge communities – the first shift, eg., brought molecular biologists into conflict with neighbouring biological fields such as cell biology and biochemistry,¹⁹ while the second brought them into close contact with the emerging biotechnology industry in the United States.²⁰ Although the central projects of molecular biology in both of these periods were focused on understanding the structure and function of genetic material, molecular biologists did take an interest in neuroscience and behaviour. Sydney Brenner, who collaborated with Francis Crick on studies of messenger RNA, famously declared in the 1960s that all of the classical problems of molecular biology had been solved and turned his attentions instead to the more exciting problem of the development of the nervous system.²¹ Brenner’s approach to studying these questions with the worm *C. elegans*, informed by his prior work in molecular biology, involved first establishing a map of each cell in the normal, ‘wild type’ worm and then comparing this standard organism to genetic mutants created through exposure to mutagenic chemicals.²²

In contrast to molecular biologists’ use of manipulations at the level of the genetic material, research in animal behaviour genetics in the post-WWII period focused largely on experimental interventions at the level of whole organisms and their

¹⁸ Hans-Jörg Rheinberger, ‘What Happened to Molecular Biology?’, *BioSocieties*, 3 (2008), 303–10, here: 304.

¹⁹ Phina G. Abir-Am, ‘The Politics of Macromolecules: Molecular Biologists, Biochemists, and Rhetoric’, *Osiris*, 7 (1992), 164–91.

²⁰ Nicolas Rasmussen, *Gene Jockeys: Life Science and the Rise of Biotech Enterprise* (Baltimore: Johns Hopkins University Press, 2014).

²¹ S. Brenner to M. Perutz, 5 June 1963, cited in Soraya de Chadarevian, ‘Of Worms and Programmes: *Caenorhabditis Elegans* and the Study of Development’, *Studies in History and Philosophy of Biology and Biomedical Sciences*, 29, 1 (1998), 81–105.

²² Rachel A. Ankeny, ‘The Natural History of *Caenorhabditis elegans* Research’, *Nature Reviews Genetics*, 2 (2001), 474–9; *idem*, ‘Fashioning Descriptive Models in Biology: Of Worms and Wiring Diagrams’, *Philosophy of Science*, 67 (Supp., 2000), S260–72. Soraya de Chadarevian has also shown how Brenner’s approach to this new area of study was influenced by his prior work in gene function and protein synthesis, from his choice of a ‘simple’ experimental organism such as *C. elegans* to his invocation of the notion of a ‘programme’ by which genes controlled development; see de Chadarevian, *ibid.*

environments. Selective breeding experiments, comparisons of inbred strains of mice, behavioural tests, and statistical analysis were the primary tools used to demonstrate the heritability of characteristics such as intelligence and temperament.²³ Researchers in human behaviour genetics were likewise interested in parsing the contributions of genetics and environment in forming behavioural traits but pursued these questions using methods such as comparisons of identical twins reared apart.²⁴ The methodological and conceptual history of animal behaviour genetics has important connections to the field of psychology, both because practitioners defined themselves against the then prevailing behaviourist theories²⁵ and because many of the behavioural tests were either developed by psychologists for work in comparative psychology or drew on psychological theories for their rationale.²⁶ These animal behavioural tests were particularly important for assembling knockout studies because they offered readily available ‘theory–methods packages’ for evaluating the differences between live, behaving, genetically engineered mice and their unaltered littermates.²⁷ Despite what might seem like obvious overlaps between the research agendas of behaviour genetics and other fields of genetics research, behaviour genetics remained relatively isolated from mainstream genetics in the 1970s and 1980s. Aaron Panofsky argues that this disconnect was the result of public controversies in the late 1960s around race and the heritability of intelligence, which made mainstream geneticists worry that working on behaviour would revive their field’s association with eugenics.²⁸

Thus, while animal behaviour geneticists and some molecular biologists could both be said to be doing ‘the genetics of behaviour’ in the early 1990s, what practitioners from each of these fields meant by ‘genetics’ would have likely been quite different. One senior mouse geneticist recalled clashes between her understanding of ‘genetics’ and those of molecular biologists she taught in training courses at the Jackson Laboratory:²⁹

²³ John Paul Scott’s experimental work at the Jackson laboratory is a good example of what research in animal behaviour genetics looked like in the post-WWII period. On Scott’s dog breeding experiments, see Diane Paul, ‘The Rockefeller Foundation and the origin of behavior genetics’, in K.R. Benson, J. Maienschein and R. Rainger (eds), *The Expansion of American Biology*, 231–61. On Scott’s experiments with mice and aggression, see Gregg Mitman, ‘Dominance, Leadership, and Aggression: Animal Behavior Studies During the Second World War’, *Journal of the History of the Behavioral Sciences*, 26, 1 (1990), 3–16.

²⁴ On the history of twin studies in the United States, see M. Susan Lindee, *Moments of Truth in Genetic Medicine* (Baltimore: Johns Hopkins University Press, 2008), especially ch. 5. On the different methodological approaches that constitute the heterogeneous field of behaviour genetics, see Aaron Panofsky, *Misbehaving Science: Controversy and the Development of Behavior Genetics* (Chicago: University of Chicago Press, 2014).

²⁵ Panofsky, *ibid.*, see especially ch. 2.

²⁶ For example, one commonly used test for anxiety, the open field test, was originally developed in the 1930s by psychologist Calvin S. Hall as a test of ‘emotionality’ in rodents. Another common test for anxiety, the light/dark emergence task, was developed in the early 1980s by a behavioural neuroscientist but draws strongly on the social psychologist Kurt Lewin’s theory of approach – avoidance conflicts. For more on how the validity of these tests as models for human anxiety is constructed and maintained, see Nicole C. Nelson, ‘Modeling Mouse, Human, and Discipline: Epistemic Scaffolds in Animal Behavior Genetics’, *Social Studies of Science*, 43, 1 (2013), 3–29.

²⁷ Joan Fujimura, ‘Crafting Science: Standardized Packages, Boundary Objects, and “Translation”’, in A. Pickering (ed.), *Science as Culture and Practice* (Chicago: University of Chicago Press, 1992), 168–211.

²⁸ Panofsky, *op. cit.* (note 24).

²⁹ The Jackson Laboratory is a non-profit organisation that conducts genetic research using mouse models and provides scientific services to the mouse community (such as providing a centralised repository for mouse strains). For a history of the Jackson Laboratory and its role in the standardisation of the mouse as a model organism, see Karen Rader, *Making Mice: Standardizing Animals for American Biomedical Research, 1900–1955* (Princeton: Princeton University Press, 2004).

In a lot of schools during the 80s and 90s there were molecular genetics programs, and that's what they thought geneticists were. And all they knew how to do was run PCR assays and sequence DNA and you know [*laughter*]. And so actually a big thing that we do at the lab is to help scientists entering mouse genetics from those angles to learn how to handle mice, learn how to work with them, learn things about genetic backgrounds, and so on.³⁰

Rheinberger has argued that the fuzzy and ill-defined nature of the gene concept is one of the reasons that it has been so productive in so many different experimental systems and research programs.³¹ And as this researcher's comments demonstrate, the differences in the nature of these experimental systems and research programs in the late 1980s and early 1990s could be quite pronounced: while one geneticist's research practice might have relied heavily on breeding and handling animals, another geneticist might have built a successful career having examined hundreds of gels but never touching a live, behaving mouse.

Molecular Biology Meets Animal Behaviour Genetics

Given the very different material and conceptual histories of molecular biology and animal behaviour genetics, how did practitioners from these fields converge on similar experiments in the early 1990s? Historical examinations of the molecularisation of the life sciences, such as Angela Creager's work on the tobacco mosaic virus and Soraya de Chadarevian's work on the Laboratory of Molecular Biology at Cambridge, have focused on the importance of *in vitro* techniques and simple model organisms such as viruses and worms, but have not explored the circumstances that led molecular biologists to turn towards higher organisms such as the mouse.³² Histories of animal behaviour genetics, such as Diane Paul's work on the dog-breeding program at the Jackson Laboratory, locate this line of research within pre-WWII research programs on eugenics but do not explore its contacts with contemporaneous scientific fields.³³ Given the centrality of molecular techniques, mice as model organisms, and neuroscientific questions in contemporary biology, the convergence of these fields is worth exploring in much greater detail.³⁴ Here I offer a brief sketch of how some of these intersections emerged in the specific case of knockout experiments.

Some behavioural knockout projects came about because researchers with existing research programs on behaviour and neuroscience saw potential in the new genetic engineering techniques. Kandel, eg., had devoted his career to studying the molecular basis of memory storage using the sea slug *Aplysia* as a model system, work for which he received the Nobel Prize in Physiology or Medicine in 2000. He recalls in his autobiography that he was intrigued by knockout techniques because he thought they would make it feasible to study memory in higher organisms in the same molecular detail as he had previously in lower organisms.³⁵ Kandel hired Seth Grant as a postdoc in his

³⁰ Researcher C, interview, 23 January 2008.

³¹ Hans-Jörg Rheinberger, 'Gene concepts: fragments from the perspective of molecular biology', in P. Buerton *et al.* (eds), *The Concept of the Gene in Development and Evolution* (Cambridge: Cambridge University Press, 2000), 219–39.

³² Angela N.H. Creager, *The Life of a Virus* (Chicago: University of Chicago Press, 2001); Soraya de Chadarevian, *Designs for Life: Molecular Biology After World War II* (Cambridge: Cambridge University Press, 2002). See also Hans-Jörg Rheinberger, *Toward a History of Epistemic Things* (Stanford: Stanford University Press, 1997).

³³ Paul, *op. cit.* (note 23). Panofsky, *op. cit.* (note 24) does take up the question of contacts between behaviour genetics and other fields, but from a sociological perspective.

³⁴ On the prevalence of the 'neuro' in contemporary biology, see Nikolas Rose and Joelle M. Abi-Rached, *Neuro: The New Brain Sciences and the Management of the Mind* (Princeton: Princeton University Press, 2013).

³⁵ Kandel, *op. cit.* (note 2), see especially chs 20 and 21.

laboratory in 1989: an MD who had previously held a postdoctoral fellowship at Cold Spring Harbor Laboratory and whose familiarity with mouse genetics Kandel recalls as key to undertaking this new line of research with transgenic mice.³⁶ He also formed collaborations to obtain knockout mouse stocks from other researchers for his experiments, such as Philippe Soriano at Baylor University and Stephen Goff at Columbia University.³⁷

Other instances represented efforts by molecular biologists to move into new areas of study. In 1987, Susumu Tonegawa had also won the Nobel Prize in Physiology or Medicine, but for his work on the genetics of antibody diversity rather than on neuroscience. Kandel recalls that Tonegawa was looking for new scientific challenges after these accomplishments, and the two met in 1987 to discuss Tonegawa's emerging interest in studying consciousness.³⁸ Around the same time, Alcino Silva, who became the first author on the Tonegawa lab's 1992 paper, recalls that another graduate student told him of Tonegawa's interest in the brain, which prompted Silva to email Tonegawa about the possibility of doing a project in his laboratory on memory.³⁹ Tonegawa had prior experience in using knockout techniques for studying the immune system and had collaborated on a paper about the molecular basis of nerve cell differentiation⁴⁰ but was still quite new to neuroscience when his group published their knockout studies in 1992. The lab's early success in applying knockout techniques to the study of learning and memory, however, secured Tonegawa's position in this new field. The MIT news office highlighted his research as one of the major breakthroughs of the year in their 'Reports to the President' in 1993, and a year later he was named the inaugural head for the newly established Center for Learning and Memory at MIT.⁴¹

The polysemy of the genes themselves also resulted in unlikely collaborations being formed. Researchers who created knockout mice found themselves drawn into new areas of research because of the unexpected results of creating such mutations, or fielding requests for collaborations from researchers in distantly related fields who were interested in the same gene but for different reasons. One of the first knockout experiments, eg., targeted a gene thought to be important in cancer, but the researchers discovered that knocking out this gene also resulted in severe brain defects.⁴² The knockout mouse lines developed by Philippe Soriano and Stephen Goff similarly targeted suspected oncogenes but were useful for Kandel's experiments because these kinases were some of the same

³⁶ *Ibid.*, 289.

³⁷ Two of the four knockout mouse lines used in Kandel's 1992 paper were developed in Soriano's laboratory, which Soriano likely provided to Kandel in exchange for co-authorship on the paper. For more on exchange practices within the mouse community, see Fiona Murray, 'The Oncomouse That Roared: Hybrid Exchange Strategies as a Source of Distinction at the Boundary of Overlapping Institutions', *American Journal of Sociology*, 116, 2 (2010), 341–88.

³⁸ Kandel, *op. cit.* (note 2), 291.

³⁹ See Silva's autobiographical note on his personal website, http://www.silvalab.org/alcino_silva.html (accessed 26 January 2015).

⁴⁰ Peter Mombaerts, Alan R. Clarke, Martin L. Hooper and Susumu Tonegawa, 'Creation of a Large Genomic Deletion at the T-Cell Antigen Receptor Beta-Subunit Locus in Mouse Embryonic Stem Cells by Gene Targeting,' *PNAS*, 88, 8 (1991), 3084–7; Asa Abeliovich, David Gerber, Osamu Tanaka, Motoya Katsuki, Ann M. Graybiel and Susumu Tonegawa, 'On Somatic Recombination in the Central Nervous System of Transgenic Mice', *Science*, 257, 5068 (1992), 404–10.

⁴¹ MIT Communications Office (ed.), *Reports to the President, 1993* (Cambridge: MIT, 1993), 488; *idem*, *Reports to the President, 1994* (Cambridge: MIT, 1994), 41, 380, 385.

⁴² This research was conducted in the laboratory of Mario Capecchi, who was one of the recipients of the 2007 Nobel Prize in Physiology or Medicine for this work. See Kirk R. Thomas and Mario R. Capecchi, 'Targeted Disruption of the Murine Int-1 Proto-Oncogene Resulting in Severe Abnormalities in Midbrain and Cerebellar Development', *Nature*, 346, 6287 (1990), 847–50.

ones under study in his group's learning and memory work.⁴³ Molecular biologists who had generated knockouts also took it upon themselves to find out if their mutant mice might have interesting properties other than the ones they anticipated. Jacqueline Crawley, an expert on rodent behavioural analysis, recalls that in the early days of knockout experiments she received many calls from researchers who had little prior experience in behaviour or neuroscience but became interested in it after producing mice with mutations in genes expressed in the brain. In an article authored with Richard Paylor, one of the behaviour geneticists who collaborated with the Tonegawa laboratory on testing their α -CaMKII knockouts, they estimated that by 1997 they had tested over a dozen different transgenics and knockouts created by other research groups,⁴⁴ and Crawley estimates that her laboratory alone had as many as forty such collaborations by the mid-2000s.⁴⁵

As might be expected, such collaborations between researchers with very different disciplinary backgrounds were not always easy. To some behaviour geneticists and behavioural neuroscientists, knockout studies seemed to be yet another in a long line of colonising moves made by molecular biologists. As Pnina Abir-Am has shown in her analysis of molecular biology's relations with neighbouring disciplines in the 1960s, molecular biology was perceived as a threat by organismal biologists and biochemists because it was 'redefining, and hence appropriating, many concepts, both central and peripheral, around which the "classical" disciplinary monopolies were constituted'.⁴⁶ The same could be said of the field of behaviour genetics in the 1990s, where molecular biologists appeared to be redefining what it meant to do 'the genetics of behaviour' in terms that worked to their advantage.⁴⁷ The research agenda that molecular biologists proposed resonated with the dominance of genetics in medical thought and practice and the cultural cachet of the gene in the 1990s,⁴⁸ and their proposals attracted substantial amounts of funding. Steven Hyman, director of the National Institute of Mental Health 1996–2001, recalls that funding for genetic studies in animal models 'markedly increased' during his tenure, but he notes that most of this investment went into knockout mouse models and large-scale mouse mutagenesis projects.⁴⁹

Asymmetries in the resources available to molecular biologists were further compounded by what some saw as unsportsmanlike behaviour from these new entrants to the behavioural field. Tonegawa was accused of being particularly unwilling to exchange

⁴³ Pamela L. Schwartzberg, Alan M. Stall, Jeff D. Hardin, Katherine S. Bowdish, Teresa Humaran, Sharon Boast, Margaret L. Harbison, Elizabeth J. Robertson and Stephen P. Goff, 'Mice Homozygous for the *abl*^{m1} Mutation Show Poor Viability and Depletion of Selected B and T Cell Populations', *Cell*, 65, 7 (1991), 1165–75; Philippe Soriano, Charles Montgomery, Robert Geske and Allan Bradley, 'Targeted Disruption of the C-*Src* Proto-Oncogene Leads to Osteopetrosis in Mice', *Cell*, 64, 4 (1991), 693–702.

⁴⁴ Jacqueline N. Crawley and Richard Paylor, 'A Proposed Test Battery and Constellations of Specific Behavioral Paradigms to Investigate the Behavioral Phenotypes of Transgenic and Knockout Mice', *Hormones and Behavior*, 31 (1997), 197–211.

⁴⁵ Jacqueline N. Crawley, interview, 21 November 2007.

⁴⁶ Abir-Am, *op. cit.* (note 19).

⁴⁷ Panofsky, *op. cit.* (note 24) describes the late 1980s and early 1990s as a time of both excitement and anxiety in behaviour genetics because of the rapid advances in the molecular techniques. He argues that techniques brought new money into the field at a time when NIH grants for behaviour genetics were in decline but also generated tensions between veteran behaviour geneticists and the new entrants to the field.

⁴⁸ Dorothy Nelkin and M. Susan Lindee, *The DNA Mystique: The Gene as a Cultural Icon* (New York: Freeman, 1995); Lindee, *op. cit.* (note 24).

⁴⁹ Steven Hyman, 'Using genetics to understand human behavior: promises and risks', in Erik Parens, Audrey R. Chapman and Nancy Press (eds), *Wrestling with Behavioral Genetics: Science, Ethics and Public Conversation* (Baltimore: Johns Hopkins University Press, 2006), 109–30, here: 109–10.

mouse strains, even though this was considered common practice in the mouse community and his laboratory was large and well funded.⁵⁰ To established animal behaviour researchers who had limited experience in molecular techniques or limited resources to develop new research programs, it seemed that their options for responding to the molecular turn were equally unappealing: either collaborate or be conquered. In their article on behavioural testing of transgenic mice, Crawley and Paylor acknowledged the reluctance of many of their colleagues to accept entreaties to collaboration from molecular biologists, humorously suggesting that behaviourists' first reaction to such calls might be best captured by the words of poet Ogden Nash: 'If called by a panther/Don't anther.'⁵¹

Crawley and Paylor encouraged their colleagues to resist this first impulse, calling the difficulties of cross-disciplinary collaboration 'trivial' in comparison to the potential knowledge gained and stressing the opportunities that such collaborations opened up for behavioural neuroscientists to enter this 'fascinating, rapidly expanding field'.⁵² At the same time, however, they had a very specific vision of how behavioural experiments with knockout mice should be conducted. Crawley and Paylor outlined a lengthy sequence of experiments that they believed were necessary to fully evaluate the characteristics of a new knockout, and Crawley went on to publish an entire textbook on behavioural testing of transgenic mice aimed at newcomers to the field.⁵³ The aspects of behavioural experimentation emphasised by Crawley and her colleagues – proper controls, handling techniques, and breeding schemes – were the types of details that became contentious in debates about how to conduct and interpret knockout experiments throughout the 1990s. While researchers from many different experimental backgrounds agreed in principle that knockouts offered an exciting opportunity for generating new knowledge about behaviour, they differed on how exactly to do such experiments and how do them well. These interdisciplinary negotiations over what exactly it meant to conduct a 'well-done' knockout experiment are the focus of the next section.

A Well-done Knockout Experiment

In his seminal work on replication, Harry Collins argues that notions of experimental skill and experimental phenomena are co-constituted in the practice of scientific research, especially at the evolving frontiers of science.⁵⁴ In contrast to experiments dealing with established phenomena where the expected results are well known and it is possible to judge a practitioner's competency or experimental design based on the results she produces, when investigating new phenomena, it is impossible to assess whether the experiment has been executed well based on the results alone, since the outcome is unknown. Conducting further experiments only adds to the uncertainty, as the underlying problem of how to assess the validity of the experimental data remains unresolved. Collins terms this quandary 'the experimenter's regress' and argues that scientists overcome this problem of establishing a basis for judging novel experimental

⁵⁰ Jon Cohen, 'Share and Share Alike Isn't Always the Rule in Science', *Science*, 268, 5218 (1995), 1715–8.

⁵¹ Crawley and Paylor, *op. cit.* (note 44), 205.

⁵² *Ibid.*, 206, 197.

⁵³ The textbook contains an introductory note welcoming molecular geneticists to the world of behavioural neuroscience while at the same time warning them that behavioural testing is too complex to be undertaken without an experienced collaborator. See Jacqueline N. Crawley, *What's Wrong with My Mouse? Behavioral Phenotyping of Transgenic and Knockout Mice* (New York: Wiley-Liss, 2000), 3.

⁵⁴ Collins, *Changing Order*, *op. cit.* (note 15).

findings through a process of social negotiation, where they form a consensus about who counts as a skilled practitioner and what counts as a 'well-done experiment'.⁵⁵ Interpretations of new experimental phenomena are thus formed alongside estimations of experimental skill.

The Genetic Background Effect

As Collins has shown, negotiations over the interpretation of novel experimental results can be a protracted process, even amongst a relatively small set of highly specialised practitioners. In the case of behavioural knockout experiments, the substantial disciplinary differences between the practitioners involved amplified divergent opinions on how to conduct and interpret these experiments. One of the most widely discussed issues with knockout experiments was a problem that came to be known as the genetic 'background effect'. When knockout techniques were first developed, the embryonic stem cells that were used for producing the targeted molecular alterations came from a strain of mice known as the 129, since their cells were particularly amenable to being raised in culture. Ironically, while the cells were easy to work with, adult 129 mice were not; they were susceptible to disease and difficult to breed in a laboratory setting. To circumvent these difficulties, researchers often mated their mutant 129 mice with more robust mice (such as the widely used C57) to create a more stable knockout mouse line. The process of 'transferring' the knockout mutation into a new mouse strain through crossbreeding, however, resulted in chimeric mice who had a mixture of genetic material from both strains. The question at issue in the research community was whether this variation in the genetic 'background' of the mouse mattered or had any impact on interpreting the results of the specific genetic alteration.

Evidence of the importance of the strain used to create knockouts was present even from the first studies – the Kandel laboratory's 1992 paper noted in a footnote that the spatial learning deficits they observed varied depending on the genetic background of the mice – but attention to the issue was amplified through a series of opinion pieces on the issue. In a highly cited debate article in *Trends in the Neurosciences*, Genentech researcher Robert Gerlai summarised the problem as one of 'the confounding effects of background genes'.⁵⁶ He reviewed fourteen behavioural knockout studies and found that the majority of these studies used mice that were hybrids of two different strains, and in a few cases the published study did not even report on whether they had crossbred their mice. Particularly disconcerting to him was the fact that the 129 mice, the source of the stem cells used to make nearly all of the knockouts he reviewed, had pronounced behavioural anomalies, such as severe deficits in learning and memory. He pointed out that, in several knockout studies, such as the Tonegawa lab's papers, the deficits in learning and memory that were attributed to knocking out a particular gene bore a striking resemblance to the deficits seen in unaltered 129 mice. 'It is therefore possible', he argued, 'that the differences observed between mutant and control mice were, in fact, due to the genetic differences (in the linked background genes) between the inbred strains used in the generation of the null-mutant animals and not to the null mutation.'⁵⁷

⁵⁵ Collins, 'Son of Seven Sexes', *op. cit.* (note 15), 34.

⁵⁶ Robert Gerlai, 'Gene-targeting Studies of Mammalian Behavior: Is It the Mutation or the Background Genotype?', *Trends in Neurosciences*, 19, 5 (1996), 177–81, here: 181.

⁵⁷ *Ibid.*, 179. On the confounds complicating the assessment of learning and memory performance in knockout mice see also Hans-Peter Lipp and David P. Wolfer, 'Genetically Modified Mice and Cognition', *Current Opinion in Neurobiology*, 8, 2 (1998), 272–80.

The influence of practitioners' disciplinary backgrounds on their opinions on how to construct appropriate genetic backgrounds is evident in two reports published in 1997 offering recommendations on the issue. The first emerged from a series of roundtable discussions that took place at a Banbury conference on genetic approaches to learning and memory, held at the Cold Spring Harbor Laboratory in December 1996 and attended largely by molecular biologists and neuroscientists.⁵⁸ The report, authored by Silva and colleagues, presented recommendations aimed at standardisation and facilitating replication between laboratories.⁵⁹ They suggested that all publications should include a detailed description of the genetic background of the mice, that readily available mouse strains should be used so that others could replicate knockout experiments, and that the use of a common genetic background across experiments would be desirable.⁶⁰ The second paper emerged from a workshop on behavioural phenotypes of inbred strains of mice, hosted by behaviourists Crawley and Paylor at the National Institutes of Health (NIH) campus.⁶¹ In contrast to the Banbury conference report's emphasis on standardisation, Crawley and colleagues argued that there was 'no one best strain that [could] be recommended' for all knockout experiments and that the specific hypothesis being tested would determine the most appropriate strain to use.⁶² If the gene under study was suspected to impair learning, for example, then a strain with strong learning skills would be the best mouse to use for a knockout study because it would increase the experimenters' chances of seeing a behavioural change. Their report went on to list in extensive detail the behavioural characteristics of many commonly used mouse strains so that future researchers could make such informed choices.

These contrasting ideas on the appropriate approach to addressing the background effect reveal several differences between behaviourists' and molecular biologists' ideas about what constitutes a well-done knockout experiment. First, it reveals differing opinions on whether the problem merited an immediate change in methodology or a wait-and-see approach. While the NIH report took it as a given that changes in experimental practice were needed, the Banbury report's recommendations for facilitating future replications suggested a less pressing need for reform. Opinion articles on the background effect indicate that there was, at least initially, a good deal of debate about how real the genetic background effect even was. In a commentary on a 1995 study that systematically demonstrated the impact of genetic background on the results of knocking out the epidermal growth factor (EGF) gene, behavioural researcher Wim Crusio commented that 'this result apparently surprised many, which in itself is quite amazing and telling because [the experiment] should not have been necessary.'⁶³ These discussions were intensified

⁵⁸ *Cold Spring Harbor Laboratory Annual Report 1996* (Cold Spring Harbor: CSHL Archives, 1996), 304–5.

⁵⁹ Alcino J. Silva, Elizabeth M. Simpson, Joseph S. Takahashi, Hans-Peter Lipp, Shigetada Nakanishi, Jeanne M. Wehner, Karl P. Giese, Tim Tully, Ted Abel, Paul F. Chapman, Kevin Fox, Seth Grant, Shigeyoshi Itohara, Richard Lathe, Mark Mayford, James O. McNamara, Roger J. Morris, Marina Picciotto, John Roder, Hee-Sup Shin, Paul A. Slesinger, Daniel R. Storm, Michael P. Stryker, Susumu Tonegawa, Yanyan Wang and David P. Wolfner, 'Mutant Mice and Neuroscience: Recommendations Concerning Genetic Background', *Neuron*, 19, 4 (1997), 755–9.

⁶⁰ *Ibid.*

⁶¹ Jacqueline N. Crawley, John K. Belknap, Allan Collins, John C. Crabbe, Wayne Frankel, Norman Henderson, Robert J. Hitzemann, Stephen C. Maxson, Lucinda L. Miner, Alcino J. Silva, Jeanne M. Wehner, Anthony Wynshaw-Boris and Richard Paylor, 'Behavioral Phenotypes of Inbred Mouse Strains: Implications and Recommendations for Molecular Studies', *Psychopharmacology*, 132 (1997), 107–24.

⁶² *Ibid.*, 120.

⁶³ Wim E. Crusio, 'Gene Targeting Studies: New Methods, Old Problems', *Trends in the Neurosciences*, 19, 5 (1996), 186–7, quote: 187. The original EGF study is Maria Sibilio and Erwin F. Wagner, 'Strain-Dependent Epithelial Defects in Mice Lacking the EGF Receptor', *Science*, 269, 5221 (1995), 234–8.

by the fact that the available techniques for controlling the genetic background were expensive and time consuming. Gerlai pointed out that the methods proposed for avoiding background effect confounds could potentially at least double the effort required to carry out a gene-targeting study.⁶⁴ Researchers who doubted the existence of the background effect were thus reluctant to adopt these controls simply as a precautionary measure.

Secondly, the contrasting consensus reports suggest that there were differing opinions on whether the background effect was a burden or an opportunity – differences which reflected the experimental traditions of each field. The emphasis on standardisation in the Banbury report aligns closely with approaches in molecular biology that relied on comparisons between ‘wild type’ and mutant organisms. As Rachel Ankeny has shown, constructing a ‘standard’ organism to serve as the basis for future comparisons was a critical part of Sydney Brenner’s early work with *C. elegans* on the nervous system,⁶⁵ and molecular biologists’ desire to arrive at a consensus on a single mouse strain for constructing knockouts reflects a similar logic of creating a standard object to serve as a reference point for targeted mutations. Behaviourists reacted differently to the idea of results that varied along with the genetic background. Crusio noted that while ‘many researchers will probably consider such interactive effects a nuisance’ and would choose solutions circumventing these issues, he thought that ‘for a number of reasons, this would be a pity’.⁶⁶ The longstanding interest in individual differences and comparisons between strains in animal behaviour genetics made the background effect seem not only expected, but potentially as a profitable area for studying questions about genetic redundancy or interactions between genes. Crusio argued that studying the effects of knocking out genes on several different backgrounds would also offer a more realistic representation of how genes functioned in nature, where individuals with different genotypes might react to a mutation or treatment differently.⁶⁷

Conflicting Experimental Reports

Similar disciplinary divides can be seen in discussions about other interpretational issues in knockout experiments, such as how to assess conflicting experimental reports. With an increasing number of knockout mice being constructed in the mid-1990s, in more than one instance different research groups found themselves developing and testing mice with similar targeted mutations in similar tests and often producing different results. In one case, two independent research groups (one of which was Tonegawa’s) created knockouts of the dopamine D1 receptor, and while one group found that the mutation produced an increase in the locomotor activity of the mice, the other group found no such behavioural difference.⁶⁸ In another instance, a research group at Columbia University reported two different behavioural profiles for their serotonin 1B receptor knockout mice: in early studies they found that mice lacking the receptor showed no difference in the open field test (a test of anxiety), but in later work they found that these same knockouts showed more anxiety behaviour in this same test.⁶⁹

⁶⁴ Gerlai, *op. cit.* (note 56), 180.

⁶⁵ Ankeny, ‘Fashioning Descriptive Models’, *op. cit.* (note 22).

⁶⁶ Crusio, *op. cit.* (note 63), 187.

⁶⁷ *Ibid.*

⁶⁸ Xu *et al.*, *op. cit.* (note 11); John Drago, Charles R. Gerfen, Jean E. Lachowicz, Heinz Steiner, Tom R. Hollon, Paul E. Love, Guck T. Ooi, Alexander Grinberg, Eric J. Lee, Sing Ping Huang, Perry F. Bartlett, Pedro A. Jose, David R. Sibley and Heiner Westphal, ‘Altered Striatal Function in a Mutant Mouse Lacking D1A Dopamine Receptors’, *PNAS*, 91, 26 (1994), 12564–8.

⁶⁹ Sylvie Ramboz, Frédéric Saudou, Djamel Ait Amara, Catherine Belzung, Louis Segu, René Misslin, Marie-Christine Buhot and René Hen, ‘5-HT_{1B} Receptor Knock Out – Behavioral Consequences’, *Behavioural*

Behaviourists pointed out that such differences might be due to the background effect or other genetic differences between the animals but could also be due to differences in the way that practitioners conducted their experiments.⁷⁰ In one controversial and highly cited study published in *Science* in 1999, three behavioural research laboratories attempted to make the importance of even minute procedural differences visible by conducting the same set of mouse behavioural experiments simultaneously in their laboratories, attempting to control for as many variables in the testing equipment, protocols, and environment as they could.⁷¹ They found that despite this close attention to standardisation, in several cases there were statistically significant (and in a few cases large) differences between the test results from each laboratory, which they attributed to as-yet-unidentified sources of variation in the laboratory environment. They concluded by urging caution in interpreting the results of knockout studies: ‘For behaviors with smaller genetic effects (such as those likely to characterize most effects of a gene knockout)’, they wrote, ‘there can be important influences of environmental conditions specific to each laboratory, and specific behavioral effects should not be uncritically attributed to genetic manipulations such as targeted gene deletions.’⁷²

In interpreting conflicting results of knockout experiments, molecular biologists once again tended to focus on standardisation as the solution to experimental problems, assuming that a properly standardised experiment would produce a consistent experimental outcome. In cases such as the comparison study in *Science*, where standardisation did not seem to resolve the issue, they assumed there must be a problem either with the tests themselves or with the practitioners executing them. Behaviourists, in contrast, saw these supposedly conflicting findings as a degree of variability that was expected and even acceptable given the interactive nature of behaviour and the complexity of the experimental setups used to study it. Such methodological problems, they argued, had long been known in the field and were simply part of the price of admission to studying behavioural phenomena. In a follow-up article to the 1999 *Science* study, the authors noted that reactions to the study ‘covered the spectrum of opinion from an indignant sigh that we all knew this already to hysterical outbursts that our findings invalidate the entire field of behavioural genetics’.⁷³ Moreover, they noted that these reactions tended to follow disciplinary lines, with researchers outside of the field of mouse behaviour genetics adopting the pessimistic view that behavioural tests were simply unreliable tools.⁷⁴

At the same time, some behaviourists began to argue that it was knockout techniques that were inherently flawed, and that removing entire gene products from a complex, interactive

Brain Research 73, 1–2 (1995), 305–12; Xiaoxi Zhuang, Cornelius Gross, Luca Santarelli, Valerie Compan, Anne-Cécile Trillat and René Hen, ‘Altered Emotional States in Knockout Mice Lacking 5-HT1A or 5-HT1B Receptors’, *Neuropsychopharmacology*, 21 (1999), 52S–60S. The conflict between these two results is noted in Deborah A. Finn, Mark T. Rutledge-Gorman and John C. Crabbe, ‘Genetic Animal Models of Anxiety’, *Neurogenetics*, 4, 3 (2003), 109–35.

⁷⁰ Finn *et al.*, *ibid.*

⁷¹ John C. Crabbe, Douglas Wahlsten and Bruce C. Dudek, ‘Genetics of Mouse Behavior: Interactions with Laboratory Environment’, *Science*, 284, 5420 (1999), 1672.

⁷² *Ibid.*

⁷³ Douglas Wahlsten, Pamela Metten, Tamara J. Phillips, Stephen L. Boehm II, Sue Burkhart-Kasch, Janet Dorow, Sharon Doerksen, Chris Downing, Jennifer Fogarty, Kristina Rodd-Henricks, René Hen, Carrie S. McKinnon, Catherine M. Merrill, Cedar Nolte, Melike Schalomon, Jason P. Schlumbohm, Jason R. Sibert, Charlotte D. Wenger, Bruce C. Dudek and John C. Crabbe, ‘Different Data from Different Labs: Lessons from Studies of Gene–Environment Interaction’, *Journal of Neurobiology*, 54, 1 (2003), 283–311, here: 305.

⁷⁴ *Ibid.*, 306.

biological system generated more knowledge production problems than it solved. In a follow-up to his original debate article on the background effect, Gerlai noted that while several practical solutions had been put forward to assess the contribution of background genes versus targeted mutations, other, less tractable problems with the knockout approach remained.⁷⁵ His major concern was that even well-designed experiments still raised questions about changes induced by the targeted deletion at the system level. The problem, as he saw it, turned on the fact that the knockout method took the individual gene as the primary unit of biological organisation, rather than systems of genes working in concert.⁷⁶ Knocking out a gene could result in a number of compensations from other biological systems, which would present serious interpretational issues – making a mutation that impacted a mouse’s sense of smell, eg., might result in a mouse with more acute vision, even though the gene targeted had nothing to do with the visual system. In another opinion article, a behavioural neuroscientist argued that the knockout approach was fundamentally flawed because it was not representative of how the molecular mechanisms of behaviour functioned in nature, where behavioural differences were produced by natural genetic variation rather than deletions of whole genes.⁷⁷ Rather than making drastic modifications that bore little resemblance to naturally occurring populations, he argued that researchers would be better off studying milder mutations in the interest of pursuing a ‘kinder, gentler genetic analysis of behavior’.⁷⁸

These commentaries questioning the basic value of knockout techniques and behavioural tests as experimental tools suggest that at stake were not just questions of methodology but also epistemological questions about how best to generate knowledge about the relationship between genes and behaviour. Not only did behaviourists and molecular biologists disagree on issues such as whether the best experimental design involved using a standard mouse strain or choosing a strain based on its specific behavioural features, they disagreed more fundamentally on whether knocking out genes was an appropriate means for investigating the relationship between genes and behaviour at all. Recognising both the methodological and epistemological dimensions of these discussions is important because resolutions to disagreements of one kind did not necessarily bring about closure on other disagreements. Researchers might have agreed on appropriate breeding schemes for controlling the background effect while still disagreeing on the value of the data produced through such a well-done experiment – what for one researcher was an elegant technique for precisely demonstrating the importance of a particular enzyme remained for another a fairly crude method for probing a complex network of genes. It is this relationship between skilfully designed experiments and practitioners’ knowledge commitments that I take up in the next section.

Articulating Ontologies, Epistemologies, and Skill in Animal Behaviour Genetics

So far, I have laid out some of the differences in the conceptual and experimental histories of molecular biology and animal behaviour genetics, and shown how these experimental traditions informed the positions that practitioners took when the two fields

⁷⁵ Robert Gerlai, ‘Gene Targeting: Technical Confounds and Potential Solutions in Behavioral Brain Research’, *Behavioural Brain Research*, 125, 1–2 (2001), 13–21.

⁷⁶ *Ibid.*, 15.

⁷⁷ Ralph J. Greenspan, ‘A Kinder, Gentler Genetic Analysis of Behavior: Dissection Gives Way to Modulation’, *Current Opinion in Neurobiology*, 7 (1995), 805–11.

⁷⁸ *Ibid.*, 811.

began to intersect in the early 1990s and disagreements arose around behavioural knockout experiments. In this section, I will switch gears and draw from a very different set of source material to further the discussion of skill in animal behaviour genetics – interviews with senior animal behaviour geneticists in which they reflected on their participation in the knockout debates, and ethnographic observations from contemporary animal behaviour genetics laboratories. These retrospective accounts should not be taken as faithful historical narratives of the events that transpired in the 1990s, but they offer nonetheless a productive site for further developing the relationship visible in published opinion and debate pieces between practitioners’ epistemologies and their views on how to design, conduct, and interpret knockout experiments. Indeed, the ways in which actors exaggerate some events and de-emphasise others in their retrospective retellings can be useful analytical tools for discerning the perspectives and commitments of different actors.⁷⁹

The Nature of Gene Action and of Behavioural Tests

In developing the concept of the experimenter’s regress, Collins draws on a case study of a controversy that took place in the 1970s around the detection of gravitational radiation.⁸⁰ Einstein’s theory of general relativity predicted the existence of such gravitational waves, but finding evidence of these waves presented a difficult experimental challenge because of the extraordinarily small signal that they were expected to produce in a detector and the large amounts of signal noise from other sources. The controversy that Collins follows turned on a straightforward question – did the experiments under examination detect gravity waves or not? – and, as he shows, a conclusion to the controversy remained out of reach so long as there was no ‘universally agreed criterion of experiment quality’ available.⁸¹ Collins thus characterises the problem confronted by the experimenters and its solution as one of breaking the loop between interpretations of experimental data and evaluations of experimental skill.

The crux of the controversy in Collins’s case can be reduced to a single, binary question because it took place within a ‘core set’ of specialists who shared similar assumptions.⁸² Those involved in the debate agreed on the theoretical possibility of gravity waves and how they could be expected to behave, and agreed that constructing a device that vibrated at the predicted frequency of the waves would be a good way of detecting them. What remained at issue were relatively narrowly defined questions about experimental skill, such as whether the signal-to-noise ratio suggested high quality data or the computer programs used to analyse the data contained errors. In the controversy about knockout experiments, the heterogeneous nature of the experimental community meant that much less could be taken for granted. Those involved in the debates disagreed about how genes produced behavioural effects and how best to detect those relationships, as well as about more detailed methodological questions. The lack of consensus on these fundamental ontological and epistemological issues, I argue, means that a consensus on the meaning of knockout mouse experiments could not wholly be resolved through negotiations around experimental skill.

One of the key disagreements between behaviourists and molecular biologists concerned an ontological question about the nature of genes and how they functioned in producing

⁷⁹ Nicole C. Nelson, ‘Shooting Genes, Distributing Credit: Narrating the Development of the Biologic Gene Gun’, *Science as Culture*, 21, 2 (2012), 205–32.

⁸⁰ Collins, ‘Son of Seven Sexes’, *op. cit.* (note 15); Collins, *Changing Order*, *op. cit.* (note 15).

⁸¹ Collins, *Changing Order*, *op. cit.* (note 15), 84.

⁸² For a definition of the ‘core set’, see *ibid.*, 142–5.

diseases or behavioural traits. In recalling discussions about knockout experiments, behaviourists frequently emphasised this difference between themselves and the molecular biologists. They thought that the ‘gene jocks’ who were new to the study of behaviour brought with them an overly simplistic set of assumptions about the molecular basis of behaviour. One senior American behaviour geneticist commented on this difference:

The perspective from the point of view of molecular biology is sort of ‘one gene at a time’, and you’re not faced at the outset with the issue of control of particular phenotypes by multiple genes, multiple sets of genes. And so that aligns itself with the medical model that historically had tried to find a gene for a disorder, right? That contrasted very much with the field of behaviour genetics as it developed, which was immediately acutely tuned in to this complex control system likelihood. Anything that we were likely to study, if it had a genetic component, would likely be controlled by multiple genes, perhaps interacting, perhaps not, who knows?⁸³

This researcher portrayed behaviour genetics and molecular biology at two opposite ends of a spectrum of understanding about gene effects: On the one end, he aligned molecular biology with the search for discrete characters that could be linked to single genes, such as white-eyed mutations in flies or inborn errors of metabolism resulting from a single mutated protein. On the other end, he placed behaviour genetics and other fields such as agricultural genetics that dealt with continuously varying traits such as degree of intelligence or crop yield that were assumed to be the aggregate product of many different gene effects. The problem with the initial set of knockout experiments, he argued, was that the molecular biologists who executed them were interpreting the data based on a ‘medical model’ paradigm where an individual gene could be assigned a discrete function, while animal behaviour geneticists such as himself saw these experiments as perturbations of a multigenic system.

A senior behaviour geneticist in Canada made a similar argument about the different assumptions made by researchers from different backgrounds and pointed out that these assumptions had specific methodological consequences. He argued, eg., that the molecular biologists’ original position that the background effect was a non-issue was reasonable if they were operating under the assumption that the effect of knocking out a gene would be pronounced. He explained:

For [molecular biologists], a lot of these questions were less problematic. For example, if you knock out a gene that influences, let’s say, limb development and you end up with a mouse that doesn’t have any legs, okay? That’s such a major developmental alternation that these background genes and modifying effects or compensatory changes or the so-called flanking allele problem is basically irrelevant. It’s such a robust mutation that it doesn’t really matter. But for us, behavioural neuroscience and behavioural geneticists, we cannot think like that. . . We have to be a bit more worried about these seemingly minuscule, negligible genetic effects. For us, these effects are real and they’re not negligible.⁸⁴

He argues that doing behaviour genetics warrants a different set of assumptions about how genes act in producing the phenomena in question and, consequently, a different set of practices to investigate and control for small genetic effects.

Other differences that animal behaviour geneticists emphasised were the assumptions they made about the complex nature of behaviour and how to measure it. Behaviour, as they saw it, was the product of multiple inputs and interactions between those inputs, and therefore challenging to measure. Even when operating in a controlled laboratory setting and even when dealing with routine tests, behavioural researchers stressed that some degree of variability in the data was to be expected. One graduate student at Western University explained the emergence of unexpected baseline results in a common behavioural test to me in the following way:

⁸³ Researcher A, interview, 25 August 2006.

⁸⁴ Researcher D, interview, 6 January 2009.

The more complex behaviour is, the harder it is to get it to work. And so for whatever unknown reason, things don't always turn out as you expect them to . . . And so there are just difficulties with that, to get it set up and running. It's not so much a technical problem, it's just something that since it's so complicated, you can't figure out what's going on that you're not controlling for.⁸⁵

Even though the behaviour of unaltered mice in this particular test had already been extensively characterised, the student still allowed for the possibility that unexpected results could happen. This was especially true, he argued, because the test depended on multiple interlocking behaviours and capacities, such as learning, memory, and spatial orientation. He therefore saw variability in the results not as a 'technical problem' with a malfunctioning experimental setup or his own lack of skill as a researcher but as an outcome of working with a complex system where not all of the relevant variables to be controlled for were known.

Once again, researchers pointed to ways in which these assumptions had important methodological consequences. A senior American researcher argued that researchers' interpretations of conflicting results from knockout experiments depended on how they viewed the nature of behaviour and behavioural tests. He explains:

Unreliability in the biochemists' world relates to repeatability of an assay and so on, and that just means that you probably weren't good enough with your pipetting. Or your weighing out of your reagent crystal, you weren't paying enough attention. And so unreliability there is inherently associated with doing poor methodology, whereas in psychology, unreliability comes not necessarily from that kind of thing but also heavily from inherent differences in measurements of things at one time and another. Brains change, right? It's not the same thing when I test this mouse and this mouse. This mouse is different! Even if they're genetically identical, they're different! This mouse may have come from a cage where it grew up with two siblings, and this one with four siblings, and there's a disinterest in that nuanced level of complexity that psychologists would naturally be interested in.⁸⁶

He argues that while molecular biologists saw conflicting experimental results as evidence of 'bad bench work', behaviourists attributed these differences to the nature of behaviour and the tests used to measure it, and didn't necessarily view consistency as a marker of experimental skill. Another animal behaviour geneticist at Western concurred. To molecular biologists, he argued, '[behavioural tests were] assays, no different than a Lowry protein assay. Except that it kind of is! If you do it this way, you get one answer, if you do it a little differently, you're going to get a different answer'.⁸⁷ Using Hans-Jörg Rheinberger's terminology, it could be said that animal behaviour geneticists tended to regard behaviours and behavioural tests as 'epistemic things' whose responses to experimental manipulation had not yet been fully characterised. Unravelling the many genetic and environmental inputs into behaviour and establishing the controls needed to measure it consistently, then, appeared as intellectually rewarding tasks.⁸⁸ They portrayed molecular biologists, in contrast, as treating behavioural tests as 'technical things' that would produce stable and consistent findings and were not interesting in and of themselves. In one view, behaviours were fixed quantities to be measured; in the other, they were loosely defined characters that had not yet been fully explored.

In emphasising the differences between themselves and molecular biologists, animal behaviour geneticists articulate some of the ways in which their understandings of skilled practice were dependent on particular knowledge commitments. If behaviour was seen as a complex phenomenon that was still not fully understood in the experimental setting, then a skilled practitioner was one that could manage these uncertainties and distinguish

⁸⁵ Graduate Student A, interview, 8 April 2008.

⁸⁶ Researcher A, *op. cit.* (note 83).

⁸⁷ Researcher E, interview, 14 August 2006.

⁸⁸ Rheinberger, *op. cit.* (note 32).

promising results from the experimental background noise. In contrast, if behaviour was viewed as the product of a few genes with large effects and behavioural experiments as stabilised tools, then a skilled practitioner was one who could employ these techniques to produce repeatable gene–behaviour correlations. Skill, in one view, is epitomised by accuracy of technique that generates consistent results, while in the other, skill is evidenced by the degree to which the practitioner pays attention to ‘nuanced levels of complexity’ in designing experiments and interpreting their outcomes. It is little wonder, then, that practitioners often seemed to be talking past each other in the knockout mouse debates, preventing closure from forming around the specific claims about genes and behaviour that experimenters put forward.

Interpreting Contradictory Portrayals of Skill

Understanding animal behaviour geneticists’ knowledge commitments also helps to make sense of some of the seemingly paradoxical ways in which contemporary animal behaviour geneticists talk about the place of skill in their experimental work. Given the emphasis I have outlined above on a practitioner’s ability to discern and control for multiple inputs into a complex, interactive system, I was surprised to find that many animal behaviour geneticists expressed a desire to mechanise aspects of their laboratory practice or talked about themselves as machines executing protocols. Sharon, a senior researcher at Western, described to me in great detail how she tried to act like a machine when she performed behavioural experiments. In one experiment where she had to test nearly a hundred mice over the course of a day, she found that her arm was getting tired when she lifted the mice off the table and began to worry that the last mice she tested were not getting the same testing experience as the first. In the next experiment, she took care to lift the mice a shorter distance at the beginning so she could maintain the same movement throughout the day. These two representations of skill are at odds with the way that skill has been discussed in the social science literature, where the mere mechanical reproduction of an action has been contrasted with the contextual understanding of the significance of those actions.⁸⁹

At first glance, then, animal behaviour geneticists’ descriptions of the degree of skill required to do behavioural work with mice seem contradictory, in that they describe it as highly skilled work that could nonetheless be performed at least as well (or perhaps even better) by machines. This apparent paradox becomes more intelligible when practitioners’ ontological and epistemological commitments are taken into account. In a community where the assumption is that the genetic effects being sought are small and easily lost amongst the noise of variation in other genes or the laboratory environment, eliminating some of these variables through mechanisation can be seen as a tactic for increasing the quality of experimental data rather than de-skilling this work. When I asked a graduate student which experimental task he would most like to mechanise, he replied:

It would definitely be something behavioural, because a robot is so good at doing things repetitively, exactly the same way. It would definitely have to be behavioural. Probably some sort of handling robot, something that would pick up the mouse from the cage and put it in the apparatus. Something like that would be ideal, because that’s where you get a lot of variability with behaviour, is the experimenter picking it up and putting it in the cage. . . there’s definitely variability [with machines], but I would say that it’s more reliable than a human. Because humans have a bad day, and you know they put an extra couple Newtons of force when they grab a mouse that makes it more stressed or something like that.⁹⁰

⁸⁹ See for example Harry M. Collins, G. H. de Vries and Wiebe E. Bijker, ‘Ways of Going On: An Analysis of Skill Applied to Medical Practice’, *Science, Technology & Human Values*, 22, 3 (1997), 267–85.

⁹⁰ Graduate Student A, *op. cit.* (note 85).

The desire to automate animal handling, which animal behaviour geneticists regard as a centrally important component of their experimental practice, stems not from an assumption that the skill required for performing these tasks is minimal, but because the skill required for a human practitioner to behave in a highly consistent manner is so great. It is worth noting that molecular biologists have made similar arguments about the capacity of mechanisation to increase the quality of experimental data, but with reference to a very different suite of variables more salient to their epistemic culture. Rather than expressing concern about the demeanour, body odour, or handling skills of a human practitioner, molecular biologists have called upon mechanisation to solve problems with imprecise measurements, mixing up reagents, or mislabelling of samples.⁹¹

Likewise, researchers at Western described the skill required to take up mouse behavioural testing in a paradoxical manner. While running mouse after mouse through a maze and spending hours scoring videotapes of the mice's movements, researchers would complain that 'a trained monkey could do [their] jobs' and that the laboratory might be better off investing in a computer that could do the scoring automatically. At the same time, they would complain about how difficult it was to train new technicians or graduate students in those same techniques and recount the troubles that they themselves had when trying to get experiments working. When discussing the touchy subject of non-behaviourists taking up behavioural testing, they complained both that those not trained in the field had the mistaken impression that behaviour is 'finicky' or a 'shaky thing that is difficult to measure' and also that they didn't 'take behaviour seriously' and acted as though 'any trained idiot' could perform behavioural tests.

These contradictory positions on how much skill is required to take up behavioural testing are again more intelligible when considered in relation to practitioners' assumptions about the nature of behaviour and behavioural tests. Behaviourists' continued attention to procedural details in running behavioural experiments is motivated by an assumption that the behaviours being tested are not stable objects, and that changing the experimental parameters could in essence change the behaviour. Behavioural tests may be easy to perform in the sense that – barring a mouse that jumps out of the maze – the test will always generate some kind of data about the animal's behaviour. But behaviourists see them as difficult to perform in a way that will allow them to distinguish the subtle influence of the genetic signal they seek from the procedural and environmental noise. For outside practitioners who want to employ mouse tests as standardised tools that can be incorporated into their experimental systems, behaviourists' level of attention to procedural details might seem to be either an overabundance of caution rather than an exercise of experimental skill or, conversely, an indication that the test is simply too complicated or unreliable to be adopted for routine use.

Conclusion

When knockout techniques were first developed in the late 1980s and early 1990s, researchers from many disciplinary backgrounds were excited by the new opportunities for experimental intervention that these techniques offered. In the study of behaviour, researchers trained in very different traditions – molecular biology and animal behaviour – quickly converged on very similar experimental approaches, using existing behavioural

⁹¹ For an excellent description of how mechanisation has been employed in genomics, see Hallam Stevens, 'On the Means of Bio-Production: Bioinformatics and How to Make Knowledge in a High-Throughput Genomics Laboratory', *BioSocieties*, 6 (2011), 217–42.

tests in combination with knockout animals to make new types of claims about the molecular basis of behaviour. The ensuing debates about how to do behavioural knockout studies and the heterogeneity of the practitioners involved offer a productive site for examining the relationship between disciplinary backgrounds and conceptions of experimental skill. The positions that researchers took on how to conduct knockout experiments not only reflected their experimental training but also revealed different commitments about the nature of gene action and how to study it. I have argued that understanding these ontological and epistemological commitments is key to understanding the role of skill in these experimental communities – what it means to be a skilled practitioner, why agreements on skill do not necessarily resolve disagreements on interpreting natural phenomena, and how skill can be treated as something highly specialised yet also capable of being mechanised are all more comprehensible in light of researchers' foundational knowledge commitments.

In the years following the debates that I have chronicled, there is some evidence that researchers' knowledge commitments have begun to transform as well. Contemporary animal behaviour geneticists argue that they need to be even more sensitive to the complexities of behaviour in their experimental practice than they once assumed, and acknowledge that even molecular biologists have begun to talk about the complexity of behaviour in grant proposals – although some suggested to me this was a tactical move taken because they realised that 'simple knee jerk determinism wasn't going to cut it anymore'.⁹² Other animal behaviour geneticists noted that there was no longer talk of finding a single gene for a particular disorder amongst researchers and that current estimations of the number of genes involved in behavioural traits ranged in the hundreds, although once again some suspected their colleagues of secretly hoping that the number of important genes for their behaviour might be closer to a dozen. These commentaries point to the importance of understanding the continually shifting landscape of ontological and epistemological positions when attempting to understand why scientific practitioners make the methodological choices they do.

Finally, examining how different knowledge communities interact around a similar suite of materials and techniques highlights the possibilities for different configurations of experimental practice and knowledge. Recent portraits of scientific practice have tended to argue for a close coupling between experimental practice and theory, between instruments and scientific findings. Hans-Jörg Rheinberger's work on experimental systems, eg., draws a tight relationship between the materials and technical objects available to researchers, the traces these devices leave and the new avenues they suggest, and the knowledge that experimentalists produce.⁹³ While it is valuable to show the ways in which theories are linked to specific experimental systems, it is also worth keeping open for inquiry a space between practice and knowledge. The convergence of many researchers around very similar designs for behavioural knockout studies and the very different interpretations of what those experiments could produce suggest a degree of flexibility in the way that knowledge and practice are linked. Examining intersecting knowledge communities engaged in debates about experimental skill offers a way of probing the interpretive flexibility of experimental systems.

⁹² Researcher F, interview, 19 July 2006.

⁹³ Rheinberger, *Toward a History of Epistemic Things*, *op. cit.* (note 32).