

Observation Versus Experiment: An Adequate Framework for Analysing Scientific Experimentation?

Saira Malik¹

Published online: 7 May 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Observation and experiment as categories for analysing scientific practice have a long pedigree in writings on science. There has, however, been little attempt to delineate observation and experiment with respect to analysing scientific practice; in particular, scientific experimentation, in a systematic manner. Someone who has presented a systematic account of observation and experiment as categories for analysing scientific experimentation is Ian Hacking. In this paper, I present a detailed analysis of Hacking’s observation versus experiment account. Using a range of cases from various fields of scientific enquiry, I argue that the observation versus experiment account is not an adequate framework for delineating scientific experimentation in a systematic manner.

Keywords Philosophy of experiment · Ian Hacking · Observation · Experiment · Scientific experimentation · Scientific practice

1 Introduction

“They [the Greeks] observed but did not experiment”.¹

This quote from Desmond Lee, the famous translator of Aristotle’s scientific works, identifies the two categories that form the bedrock of modern scientific practice.²

This quote also identifies well one of the principal markers used to delineate modernity. It is now a well-worn axiom that what distinguishes Western modernity—and implicit in

¹ See Desmond Lee’s *Introduction* to his translation of Aristotle’s *Meteorologica*.

² See ‘Introduction’ by Lorraine Daston and Elizabeth Lunbeck in *Histories of Scientific Observation*; in particular page 3.

✉ Saira Malik
MalikS7@cardiff.ac.uk

¹ Cardiff University, Cardiff CF10 3EU, UK

this, Western hegemony—is the phenomenon of the Scientific Revolution in the West. What marks out the scientific practices of the Scientific Revolution and thereafter in the West from other scientific enterprises in the past—in the popular imagination—is supposed to be ‘experiment’. It is posited that *this* is the hallmark of the Scientific Revolution—what went before is ‘observation’.³ What Desmond Lee appears to be doing here is setting up a binary of ‘observation versus experiment’ rather than ‘observation *and* experiment’. Many decades later, the distinguished philosopher of science Ian Hacking echoes Lee by positing, “Observation and experiment are not one thing [,] nor even opposite poles of a smooth continuum”.⁴ The casting of experiment in opposition to observation as Hacking and Lee do, rather than in addition to it, is a very modern turn.

Experiment—as *experimentum* (and its cognates) in Latin and *empeiria* or *peira* in Greek—has a continuity of usage as a category for scientific learning finding its genesis in the works of Hippocrates, Aristotle and Pliny (Pomata 2011, 45–46).⁵ Observation, as a scientific category, does not enjoy the same continuity as the essays in *Histories of Scientific Observation*, edited by Lorraine Daston and Elizabeth Lunbeck, show.⁶

In fact, what may be thought of as scientific observational practices were subsumed under a myriad of terms in Latin: *experientia*, *experimentum*, *contemplatio*, *consideratio*—and the least used—*observatio*. Where Greek is concerned, there is no equivalent for observation (*observatio*)—*teresis*—in the scientific canon of Hippocrates and Aristotle (Pomata 2011, 45). It is only in the seventeenth century that observation and experiment as scientific categories—respectively as *observatio* and *experimentum*—become established as well as conjoined (Daston 2011, 81–113). Despite this distinguishment, the terms remained conjoined: “Observation, by the curiosity it inspires and the gaps that it leaves, leads to experiment; experiment returns to observation by the same curiosity that seeks to fill and close the gaps still more; thus one can regard experiment and observation as in some fashion the consequence and complement of one another” (Daston 2011, 86). The difference implied between the two then, on the eve of the nineteenth century, was that experiment implied intervention and manipulation whereas observation did not (Daston 2011, 86). In many cases even this implied distinction was subsumed for others, including notables such as Robert Boyle and Robert Hooke, who appear to make no distinction between observation and experiment as long as both were dedicated to the cause of knowledge acquisition of the natural world (Anstey 2014, 105).

It is in the nineteenth century that one sees observation being cast in opposition to experiment rather than in addition to it (Daston and Lunbeck 2011, 3). During this period one can see the reconfiguration of vision insofar as it becomes detached from a referent and thus abstracted, leading to the inevitable subjectivity of the observer (Crary 1992).⁷ This leads, as Jonathan Crary explains, to the ‘social remaking of the observer’ (Crary 2001, 4). The observer goes from a passive receiver of the external world to an active producer of it

³ This, of course, belies the considerable academic scholarship by historians of science that exists on the nature and characteristics of scientific practice, in particular, scientific experimentation, in pre-modern cultures that has shown the significant limitations of this position; Greek, Latin, Arabic and Chinese to name just a few. See Lloyd (2004, 2006) on Greek and Chinese science and references therein. For Arabic science, see Sabra (1996). For Latin, see Lindberg (2007). For an example from the exact sciences, see the case of geometrical optics; for Greek, see Smith (1996), and for Arabic, see Sabra (2003). For the case of medicine, see Pormann and Savage-Smith (2007).

⁴ See Hacking (1983, 173).

⁵ In particular, see footnote 12 in Pomata (2011).

⁶ See Park (2011, 15–44), Pomata (2011, 45–80) and Daston (2011, 81–113).

⁷ See pp. 148–149 in particular.

(Crary 2001, 95–97). This nineteenth century reconfiguration of vision and the observer has been made clear not just in Jonathan Crary’s work on the *camera obscura* and works of art, but also in Christoph Hoffmann’s work on scientific practices, the senses and instrumentation (astronomical, in particular) during the same period (Hoffmann 2006) where the author shows that any qualitative distinction between the observer and instrumentation fades away.⁸

In light of the importance of observation and experiment as categories in scientific practice, particularly scientific experimentation, it is surprising that relatively little attention has been paid to them as a binary within the modern academy of philosophy of science⁹—in spite of philosophers of experiment such as Hans Radder and David Gooding calling for such.¹⁰ An exception is Ian Hacking—in *Representing and Intervening* (Hacking 1983).¹¹ In this essay, I scrutinise Hacking’s account of observation and experiment in order to assess its efficacy as an adequate account for delineating scientific experimentation. I show that there are significant weaknesses in Hacking’s account when used to analyse a range of cases from different fields of scientific enquiry.

2 Hacking: Observation Versus Experiment

In *Representing and Intervening* (Hacking 1983), Ian Hacking makes a category distinction between experiment and observation. He states, ‘Observation and experiment are not one thing nor even opposite poles of a smooth continuum’ (Hacking 1983, 173). According to Hacking, ‘Much of the discussion about observation, observation statements and observability is due to our positivist heritage’ (Hacking 1983, 168). He thinks the need to make these distinctions at all, and to take them seriously, is a task very much confined to professional philosophers. According to Hacking, these distinctions do not worry scientists. He gives the example of Francis Bacon to show what he means (Hacking 1983, 168–169).

Francis Bacon does not mention the term once in his discussion of the inductive sciences despite the term being in circulation during his time. Observation at this time was restricted in its use—used mainly for observations made of the heavenly bodies via telescopes. That is, the use of the term observation in the natural sciences was associated with the use of instrumentation. Instead of observation, Bacon uses the term ‘prerogative instances’. In his *Novum Organum* of 1620, he lists 27 different ‘prerogative instances’: these include a range of activities which today one may refer to as scientific practices:

⁸ See also Schickore (2007) for the case of the microscope. Also see Daston and Galison (2007).

⁹ Those who have been interested in the detailed historical accounts of particular experiments include Galison (1982, 1983), Pickering (1981), Gooding (1982), Worral (1982), Wheaton (1983), Stuewer (1975) and Franklin (1986). Others have been concerned with the role of experiment in knowledge acquisition such as Gooding (2000), Kuhn (1976), Dear (1995) and Tiles (1993). Some have been interested in the philosophy of scientific experimentation (Radder 2003a, b) which takes into account the nexus that experimentation provides for the meeting of theory, technology and modelling amongst others. Others have been concerned with the relationship between theory, observing and experimentation such as Latour and Woolgar (1986), Collins (1985), Galison (1987), Bogen and Woodward (1988) and Rheinberger (1997). Philosophers of science interested in observation include Shapere (1982) and Fodor (1983).

¹⁰ See Radder (2003b, 15) and Gooding (1992, 68).

¹¹ Hacking’s primary aim in *Representing and Intervening* (Hacking 1983), however, lies in the juxtaposition of experiment to theory rather than an analysis of experiment relative to observation per se. Although Hacking takes up the subject of experiment again in some of his later work, there he is more concerned with other matters. He deals with the anti-realist position (see Hacking 1989, for a response, see Shapere 1993) or with trying to defend the stability of laboratory practice (see Hacking 1992).

experiments, tests to distinguish between hypotheses, notable observations, some are made with devices that ‘aid the immediate actions of the senses’. The latter includes microscopes as well as telescopes, rods, astrolabes and similar devices. He calls devices that aid the senses ‘evoking devices’, devices that ‘reduce the non-sensible to the sensible; that is, make manifest, things not directly perceptible, by means of others which are’ (Hacking 1983, 168–169).

Bacon recognises the difference between what is directly perceptible and that which is hidden from the senses and needs to be ‘evoked’. He recognises it and does not give it much significance—for Bacon, the difference is not important. For Bacon, there is no difference between directly seeing the sun overhead at noon and seeing a planet via a telescope at night.

Hacking states it is only later, in the nineteenth century, that the difference between things that are directly perceptible and those that are hidden from the senses and have to be ‘evoked’ becomes important. It becomes important because the very notion of ‘seeing’ undergoes a transformation. In the nineteenth century, ‘to see’ is to see the surface—and only the surface—and all knowledge must be derived via this way. This marks the beginnings of positivism and phenomenology. Positivism needs to distinguish between inference and seeing with the unaided eye (Hacking 1983, 169). Thus, unlike Bacon, there is a difference between seeing the sun overhead at noon and seeing a planet at night via a telescope. For the positivist, the planet seen via a telescope can only be inferred—it is not an observation. According to Hacking, this marks the start of the distinction made between observation and theory in the philosophy of science and is articulated by someone like Bas van Fraassen (1980). This view has come to be contested in two ways: one that emphasises the scope of observation and the other of theory. Grover Maxwell is a good exemplar of the former view (Maxwell 1962) while Paul Feyerabend is an example of the latter (Hacking 1983, 172–173).

Hacking deals with Maxwell thus (Hacking 1983, 170). Maxwell makes a historically contingent argument. He suggests that what may be unobservable at some particular time may subsequently become observable—or in Bacon’s language be ‘evoked’—with the development of adequate instrumentation and/or the expansion of the capacity of existing instrumentation. For example, in the case of visual perception, there is a continuum that starts with seeing through a vacuum, through the atmosphere, through a simple microscope and, at present, finishes with seeing through the current batch of advanced microscopes. In this way, what in previous generations would have been unobservable—and according to positivists only inferred, and thus theoretical—becomes observable with the development of appropriate instrumentation. For example, prior to Louis Pasteur, the notion of microbial entities responsible for disease was considered theoretical. However, with the advent of microscopy these entities became observable. Other examples include genes on chromosomes, cell bodies in cells and the fine structure of metals. In all these cases the entities were regarded as theoretical until the development of adequate instruments rendered them observable. For Maxwell there is no significant difference between knowledge gained directly through the senses and that gained indirectly with the aid of instrumentation—Bacon’s ‘evoking’ devices.

The second type of critique of the positivist stance is based on the notion that the distinction between observation and theory is redundant. That is, all observations—whether made directly via the senses or not—are theoretical—that is, there are no pure observations. All observations are ‘theory laden’ to coin Norwood Hanson’s term from his *Patterns of Discovery* (Hanson 1958, 19). Hanson states, ‘seeing is a “theory laden” undertaking’.

Paul Feyerabend agrees with Hanson but goes even further (Hacking 1983, 172–173). For Feyerabend, there is no difference between observation and theory. In fact, he has rejected the term ‘theory laden’ on the grounds that there can be no observation without theory. He states, ‘Nobody will deny that such distinctions [between observation and theory] *can* be made, but nobody will put great weight on them, for they do not play any decisive role in the business of science’ (Hacking 1983, 173). He comments on the everyday practices of science, ‘observational reports, experimental results, “factual statements”, either *contain* theoretical assumptions or assert them by the manner in which they are used (Hacking 1983, 173).

Hacking chooses to align himself with Grover Maxwell rather than Feyerabend; and is particularly scathing of Feyerabend’s [lack of] understanding of scientific practice exemplified by the statement, ‘observational reports, experimental results, “factual statements”, either *contain* theoretical assumptions or assert them by the manner in which they are used’. Hacking explains why using two historical examples: the work of Albert Michelson and Edward Morley along with that of William Herschel.

The work of Michelson and Morley is well known to historians of the physical sciences (Hacking 1983, 174). It is famous because, on reflection, it refuted the existence of ‘electromagnetic aether’ and led to the establishment of the special theory of relativity. Hacking focuses on the scientific practices of Michelson and Morley and what these mean with respect to Feyerabend’s comment, ‘observational reports, experimental results, “factual statements”, either *contain* theoretical assumptions or assert them by the manner in which they are used’. The published ‘report’ of the experiment of 1887 was 12 pages long. The ‘observations’ made were done so for a total of a couple of hours over 4 days in July. The ‘results’ of the experiment remain controversial: Michelson believed that this work showed that the earth’s motion was independent relative to the [presumed] aether. Hacking goes on to identify the components which (in his view) contributed towards the impact of this work—in its own time and up to the 1920s. These components include, *inter alia*, the making and re-making of apparatus, getting the apparatus to actually work and, most importantly—knowing when the apparatus was working. Interestingly, the most important result of this work, according to Hacking, had less to do with aether and more to do with the transformation of measurement. Hacking concludes, “In short, ‘Feyerabend’s factual statements, observation reports, and experimental results’ are not even the same kind of thing. To lump them together is to make it impossible to notice anything about what goes on in experimental science” (Hacking 1983, 174).

Hacking then proceeds to show that Feyerabend’s notion that *all* observations carry theory is false. Hacking uses the historical case study of William Herschel (d. 1822) as a rebuttal to Feyerabend’s, ‘all observations carry theory’.

William Herschel was an astronomer, who, in the year 1800, is attributed to have discovered radiant heat whilst conducting his astronomical work with his telescope (Hacking 1983, 176). On using different coloured filters in his telescope, Herschel realised that different colour filters gave off different amounts of heat. Herschel, in the reporting of his work in the *Philosophical Transactions of the Royal Society* for the year 1800 states, “When I used some of them I felt a sensation of heat, though I had but little light, while others gave me much light with scarce any sensation of heat”. It was this incidental observation to his principal work on the sun which led Herschel in a new experimental direction and the discovery of radiant heat: that the sun emits both visible and invisible rays and that human sight is sensitive to only the visible rays. This incidental observation led him to conduct a whole series of experiments investigating the transmission, reflection and refraction of these rays (Hacking 1983, 177). Hacking concludes, “Feyerabend says

that observations reports, etc., always contain or assert theoretical assumptions. This assertion is hardly worth debating because it is obviously false” (Hacking 1983, 174).

Thus, Hacking’s notion of observation appears to be very much aligned with Grover Maxwell, and with Francis Bacon’s ‘evoking devices’. Hacking’s anti-positivist stance on observation becomes even clearer when considering his position on observation of sub-atomic particles via indirect methods, and finally, with his view of observation of entities via a microscope.

On observation of sub-atomic particles using indirect methods, Hacking is in agreement with Dudley Shapere (Shapere 1982). Shapere uses the discourse of ‘observing the interior of the sun or another star’ as his starting point in his argument to show what is meant by ‘to see’ in modern science—and in the process shows how far we have journeyed along the path of Bacon’s ‘evoking devices’. Shapere analyses the solar neutrino experiment in which physicists claim that the core of the sun (or any star) can be directly observed via the detection of neutrinos.¹² Shapere shows the various layers of detection—‘seeing’—involved in the ‘direct observation’ of neutrinos emitted from the core of the sun. Shapere argues that despite what appears to be a complicated series of events entailed in the detection of neutrinos, it is justifiable to term this process as ‘direct observation’ (Shapere 1982, 492).

Hacking suggests that it is the fact that the theories underlying the detection mechanism are *not* entwined with the subject matter under investigation is what gives credence to the claim in the solar neutrino experiment that the “stellar core of the sun can be directly observed” (Hacking 1983, 185).¹³ For Hacking therefore what would count as an observation would include the detection of electrons in a bubble-chamber, as the theory used in the manufacture and operation of the bubble-chamber does not directly use theory about electrons.¹⁴ For Hacking, this also holds for the use of microscopes—that is, that the theories, assumptions and norms on which microscopes are built and used (from simple light microscopes to electron-scanning and X-ray diffraction ones) are independent of the subject matter being studied (Hacking 1983, 186–209) and for him therefore what is seen using these instruments counts as an observation.

Hacking is very much wedded to an anti-positivist stance on what constitutes observation—very much in the tradition of Francis Bacon and his “evoking devices making manifest, things which are not directly perceptible, by means of others which are”. Hacking, thus, aligns himself with working scientists for whom ‘to see’ includes detection methods ranging from the simple microscope to its X-ray diffraction and electron scanning versions—things Bacon could not have imagined when he composed his *Novum Organum*. What Hacking thus means by observation is detection.

Where experiment is concerned—again—Hacking appears to be in support of Bacon (Hacking 1983, 246–250) as the following citations from Bacon show, “The secrets of nature reveal themselves more readily under the vexation of art than when they go their own way” (Hacking 1983, 246), “shake out the folds of nature” and to “twist the lion’s tail” (Hacking 1983, 246). Hacking says this alludes to “Bacon’s good sense” (Hacking 1983, 250). However, for Hacking an experiment is not just a case of intervention, or ‘to twist the lion’s tail’, as it is for Bacon as we see below.

¹² Also see Pinch (1985).

¹³ Brigitte Falkenburg has proposed that this position has limited value as theories of entities such as neutrinos, their detectors and the way information is transmitted from the source are all inextricably linked (see Falkenburg 2000).

¹⁴ For a detailed explanation see Galison (1985).

What is Hacking saying an experiment is? He stipulates very clearly—It is the “creation of phenomena” (Hacking 1983, 220). This is made more emphatic in “Experiment is the creation of phenomena” (Hacking 1983, 229). What does Hacking mean? First—phenomena: Hacking says that he agrees with scientists as to what is meant by phenomena,

A phenomenon is *noteworthy*. A phenomenon is *discernable*. A phenomenon is commonly an event or process of certain type that occurs regularly under defined circumstances. When we know the regularity exhibited in a phenomenon we express it in a law-like generalization. The very fact of such a regularity is sometimes called the phenomenon. (Hacking 1983, 221)

This description fits very closely with the etymology of the (Greek) term phenomenon: a thing, an event or process that can be seen. However, phenomenon, Hacking points out, has quite a different sense in philosophy.

Phenomenon has a long history in its philosophical usage (Hacking 1983, 220–221) quite different to its etymological roots. Phenomenon, for philosophers—both ancient and modern—has come to indicate something related to the senses. For many ancients, phenomena were in opposition to reality insofar as phenomena—perceived via the senses—were subject to change (Hacking 1983, 221). The fact of phenomena being the subject of change led to the juxtaposition of phenomena to noumena: phenomena were only appearances of things whereas noumena were things as they actually were. Kant took up this distinction and proposed that only phenomena could be known—the noumena could not. With the advent of positivism, phenomena came to indicate sense-data—things that are “private, personal sensations” (Hacking 1983, 221)—rendered as ‘phenomenalism’, and according to one of its principal proponents, J. S. Mill, “things are only the permanent possibilities of sensation, and that the external world is constructed out of actual and possible sense-data” (Hacking 1983, 221).

Hacking breaks from the way philosophers have come to use the term phenomenon, and aligns himself with the scientists. He says,

My use of the word ‘phenomenon’ is like that of the physicists. It must be kept as separate as possible from the philosophers’ phenomenalism, phenomenology and private, fleeting sense-data. A phenomenon, for me, is something public, regular, possibly law-like, but perhaps exceptional. I pattern my use of the word [phenomenon] after physics and astronomy. (Hacking 1983, 222)

Hacking illustrates what he means by using the ‘Hall effect’ from the field of modern physics as an exemplar.

Edwin (E. J.) Hall’s work on the relationship between a magnetic field and electric potential is referred to as the Hall effect (Hacking 1983, 224–225). In the late 1870s, Hall, under the supervision of Henry Rowland at John Hopkins University, had been expanding on some of James Clerk Maxwell’s ideas from his *Treatise on Electricity and Magnetism*. In the *Treatise* Maxwell had proposed that, where a conductor carrying an electric current was under the influence of a magnetic field, the magnetic field acts on the conductor rather than the current. Hall proposed that if this were the case then there should be two possible outcomes: either the resistance of the conductor would be affected by the magnetic field or that an electric potential across the field would be produced. Hall discarded the first possibility as he failed to observe any effect by the magnetic field on the resistance of the conductor. However, the second possibility bore fruition: he was successful at measuring an electric potential across the magnetic field. He obtained an electric potential when he placed a gold leaf at right angles to the magnetic field and electric current. Hall says,

It seemed hardly safe, even then, to believe that a new phenomenon has been discovered, but now after nearly a fortnight has elapsed, and the experiment has been many times and under various circumstances successfully repeated ... it is perhaps not too early to declare that the magnet does have an effect on the electric current or at least an effect on the circuit never before expressly observed or proved. (Hacking 1983, 225)

Hacking tells us that by the 1880s it was common for physicists to call a phenomenon an effect: as in the Compton effect,¹⁵ the Zeeman effect¹⁶ and the photoelectric effect¹⁷ (Hacking 1983, 224). He states,

Phenomena and effects are in the same line of business: noteworthy discernable regularities. The words ‘phenomena’ and ‘effects’ can often serve as synonyms, yet they point in different directions. Phenomena remind us, in that semiconscious repository of language, of events that can be recorded by the gifted observer who does not intervene in the world but who watches the stars. Effects remind us of the great experiments after whom, in general, we name the effects: the men and women, the Compton and Curie, who intervened in the course of nature, to create a regularity which, at least at first, can be seen as regular (or anomalous) only against the further background of theory. (Hacking 1983, 225)

Here Hacking starts by telling us phenomena and effects have similar aims—they yield “noteworthy discernable regularities”—albeit he draws a difference between them in so far as the kinds of activities they are: phenomena as ‘events’ noted by those who do not ‘intervene in the world’ while effects are things which “remind us of the great experiments” done by those “who intervened in the course of nature, to create a regularity...”.

If we now turn to consider what Hacking means by ‘creation’ in his stipulation of experiment—‘creation of phenomena’, we find that Hacking is conferring a very constricted meaning to creation.

Hacking (again) uses Hall’s work to illustrate what he means by creation. He says, “Hall’s effect does not exist outside of certain kinds of apparatus” (Hacking 1983, 226). This is made more emphatic in, “Hall’s effect did not exist until, with great ingenuity, he [Hall] had discovered how to isolate, purify, create it in the laboratory” (Hacking 1983, 226). To give even more emphasis, Hacking cites another example: the ‘Josephson effect’ referring to the work of Brian Josephson in the 1960s. Again, the example Hacking chooses is from modern physics and, in this case, concerns the subject of electrical conduction by super-conductors.¹⁸ He says, “The Josephson effect did not exist in nature until people created the apparatus” (Hacking 1983, 229). For Hacking, it is these effects—bounded by the apparatus in which they can be demonstrated in laboratory conditions that appear to fulfil his criteria for what qualifies as experiment.

¹⁵ The Compton effect refers to the scattering of X-rays by electrons in work done by Arthur Compton in the 1920s.

¹⁶ The Zeeman effect refers to the splitting of the energy levels of an atom when it is placed in a magnetic field. Pieter Zeeman and Hendrik Lorentz did this work in the 1890s.

¹⁷ The photoelectric effect refers to the detection of a current when light is shone on some metals and is taken as an indication of the emission of electrons.

¹⁸ Many substances act as superconductors at temperatures near to absolute zero. Brian Josephson (in 1962) predicted that a weak current (subsequently named a ‘super-current’) would flow between two superconductors that were separated by a thin sheet of electrical insulation. Philip Anderson and John Rowell confirmed Josephson’s prediction a year later in 1963.

Hacking explains what he means by his statement, “the Hall effect does not exist outside of certain kinds of apparatus” (Hacking 1983, 226). He asks rhetorically, “Does not a current passing through a conductor, at right angles [*sic*] to a magnetic field, produce a potential, anywhere in nature?”, answering ambivalently, “Yes and no”. According to Hacking, if there were such an event in nature, which occurred in isolation of any other processes, then it could be said that the Hall effect occurs in nature. However, it is only in laboratory conditions that the Hall effect can be produced independent of any other processes. It is with this explanation that it becomes clear that, for Hacking, in order for a phenomenon to be created—it needs to be produced in isolation, or what he calls “in a pure state” (Hacking 1983, 226).

Hacking’s commentary on the work of Edwin Hall tells us what Hacking means when he stipulates “experiment is the creation of phenomena”. What we see is that Hacking’s stipulation of experiment as ‘creation of phenomena’ becomes highly constricted because of his insistence on the fact that the phenomena under consideration needs to be produced in a ‘pure state’ or in isolation. His repeated emphasis on ‘creation’ gives emphasis on the importance of this aspect in his stipulation of experiment. The emphasis on ‘pure state’ is underlined from the highly selective way he chooses his case studies in support of his position. He cites many examples but chooses either not to deal with them in any sustained way or dismisses them—on occasion, flippantly (Hacking 1983, 227–228). Amongst the many examples Hacking cites, he chooses to focus only on cases from modern physics, such as the Hall and Josephson effects. This appears to be a deliberate strategy on his part as the following illustrates.

Hacking introduces the medical work of Claude Bernard (published as *Introduction to the Study of Experimental Medicine* in 1865) as a potential case study to show the distinction between experiment and observation (Hacking 1983, 173). Hacking states,

Consider Dr Beauchamp [*sic*] who, in the Anglo-American war of 1812 [*sic*], had the fortune to observe, over an extended period of time, the workings of the digestive tract of a man with a dreadful stomach wound. Was that an experiment or just a sequence of fateful observations in almost unique circumstances? (Hacking 1983, 173)

In this example, Hacking not only makes a couple of errors in transposing historical details from Bernard,¹⁹ but more importantly, considerably truncates the details of William Beaumont’s study on the digestive physiology of the human stomach.²⁰ Hacking finishes by choosing not to engage with this case from medical physiology, saying, “I do not want to pursue such points” (Hacking 1983, 174).

In contrast, I think it worth pursuing this case from medical physiology, as well as some others from different fields of scientific enquiry, in order to assess how Hacking’s observation versus experiment account maps onto cases from a range of scientific experimentation.

3 Cases

First, returning to Beaumont’s story. William Beaumont himself believed that the work he was doing was an experimental investigation of human digestion (Beaumont 1833, 5–6). However, more important is how this case fits with Hacking’s account as he chooses not to

¹⁹ Dr. Beauchamp should read Dr. Beaumont (see Bernard 1957, 8). The work was conducted during the 1820s, not a decade earlier as stated (see Bernard 1957, 8).

²⁰ See Beaumont (1833).

address this question himself. First—a very brief overview of Beaumont’s work on digestion.

Beaumont, in his capacity as a doctor, treated a patient suffering from a gunshot wound, which had caused damage to his left lung and stomach (Beaumont 1833, 10). The patient recovered but with a very unusual outcome: the stomach lining did not heal in a uniform way. Instead it formed a fistula with an exterior valve. Beaumont used this valve as the access point in conducting a series of investigations on digestion (Beaumont 1833, 11–23). He used a pipetting technique to both put substances into the stomach, as well as to draw them out. In this way he examined the digestion of various substances in the stomach.

If one were to use the Hacking’s criteria for experiment, the ‘creation of phenomena’—Beaumont’s work within the stomach falls short of qualifying as experiment. That is, although the digestion process qualifies as a phenomenon (a discernable change) as well as an effect (requires activity and intervention on the part of the investigator); it does not fulfil Hacking’s criteria for creation—that is, the effect is not occurring in isolation of other processes.²¹ However, Beaumont goes on to perform a series of investigations looking at the digestive action of the ‘gastric juice’ in isolation (Beaumont 1833, 73–101).²² This series of investigations would appear to fulfil Hacking’s criteria for experiment as Beaumont sets up apparatus (however rudimentary) which gives rise to an effect—in isolation of others.

Thus, if we use Hacking’s criteria for experiment in respect to William Beaumont’s work, the outcome would be that only part of Beaumont’s work qualifies as experiment—the in vitro part. That is, the part done within the apparatus that Beaumont sets up to investigate digestion outside the stomach. The in vivo part of Beaumont’s work, that is, the work done on the stomach directly, fails to qualify as experiment as the effects occurring are not in isolation of other processes. This ignores the crucial point in William Beaumont’s investigations: the in vitro part (experiment) is contingent on the work that William Beaumont has previously done on the stomach. Beaumont would never have set up the apparatus part of his work if he had not already done the work on the stomach.

According to Hacking, thus, if Beaumont’s in vivo work is not experiment then is it a series of observations? If so, Hacking has told us that, “[o]bservation and experiment are not one thing nor even opposite poles of a smooth continuum” (Hacking 1983, 173). This statement would imply that, in Beaumont’s case, the work done in vivo and that done in vitro are not part of ‘a smooth continuum’. This is not reasonable given that the work done in vitro is continuous with the work done in the stomach. The limitations of Hacking’s framework are not only confined to this case from physiology.

In evolutionary biology, the name of Henry Kettlewell is well known. Kettlewell’s work on moths in the 1950s was important in understanding the process of natural selection.²³ Kettlewell used three kinds of moths: typical (*Biston betularia*), intermediate (*Biston insularia*) and dark (*Biston carbonaria*). In Britain, the typical moth had been prevalent in

²¹ Such as neurological processes which control the mechanical and nerve impulse activities of the stomach.

²² In particular, see p. 82.

²³ For a synopsis of Henry Kettlewell’s study on moths, see Franklin (2012). Kettlewell’s work was published in *Heredity* (1955, 1956, 1958). David Rudge has worked extensively on the history of Kettlewell’s work, see Rudge (2005a, b, 2006, 2009, 2010). He has also dealt extensively with the issue of statistical error in Kettlewell’s numerical analysis (Rudge 2001, 2005a, b) and the issue of validity of control experiments (1999) in which he deals in particular with Joel Hagen’s critique of Kettlewell’s use of controls (Hagen 1999); for an overview of the issue of the use of controls on Kettlewell’s experiments, see Brandon (1999). The validity of the controls Kettlewell used relate to the geographical areas in which he performed the experiments (Birmingham, UK and Dorset, UK).

most areas prior to industrialization. However, the proportion of the typical variety in relation to the other two types changed during the twentieth century. Kettlewell showed that this change was due to the change in colour of the landscape. Kettlewell first showed that the different kinds of moths were more or less conspicuous depending on the colour of the background on which they were settled. He did this by using volunteers to rank the degree of conspicuousness of each type of moth on different colour backgrounds. In the next stage, he put all three types of moths in a cage with different colour bark on which they could settle. He then introduced birds (predator to moths) into the cage. He found that the rate at which the moths were eaten depended on the colour of the bark on which they were settled. As three different kinds of moths were used along with different colour barks, the data analysis was very complex in this part of the study. The third part of his study was done in native conditions. Kettlewell released all three kinds of moths in both polluted (dark background) and unpolluted (lighter background) areas and tracked how many survived. This last part of the investigation depended on previously marked moths that had survived being recaptured in traps. Kettlewell showed that the dark species of moths survived better in a polluted (dark) environment than the lighter colour varieties whereas the lighter typical species survived better in the less polluted (light) environment compared with the darker varieties. He showed this was due to the colour of the landscape.

What do we see when we map Hacking's observation versus experiment account onto this case? The observation part of the account can be done in a straightforward manner. Hacking has told us that the observation part is a source of detection—in this case the numerical values related to what kind of moth species is conspicuous on which colour bark, the numerical values related to different species surviving predation in the cage, the kind of bait used for re-capture in native conditions. However, what, according to Hacking, is the experiment part? The phenomenon under study here—natural selection—is not being produced in isolation of other processes and therefore, in Hacking's account, cannot—or should not—be included in his category of experiment.

We see this anomalous consequence of Hacking's account in fields of scientific enquiry other than the two (physiology, evolutionary biology) mentioned already.

In the field of study of animal behavior and psychology, the work of Harry Harlow is well known amongst those working on attachment theory.²⁴ Harlow did his experimental work on rhesus monkeys (macaques) during the 1950s and 60s.²⁵ His work on isolated infant monkeys had shown that the infants formed a close attachment to the soft materials in their cages (diapers, bedding) whereas those infants who had their mothers in the cage did not form this attachment. Harlow conducted a series of experiments to measure degrees of attachment of an infant monkey to the quality of a carer.

Eight new-born monkeys were separated from their mothers immediately after birth. Each was placed in a cage with two 'surrogate mothers'—one surrogate was made of wire with a box face while the other surrogate was made of soft cloth with a quasi-monkey face. Milk was dispensed from each surrogate. Harlow measured the time that each infant monkey spent with each surrogate over a period of some months. He found that the infant monkeys spent more time with the cloth-covered surrogate than with the wire one. He then withdrew milk dispensation from the cloth surrogate. He found that the total time that the infants spent with the cloth surrogate was still much greater than that time spent with the

²⁴ 'Attachment theory' relates to the notion that non-material provision from a (primary) carer is significant in the cognitive formation and development of higher mammals.

²⁵ See Prior and Glaser (2006), Ainsworth (1991), Blum (2002); see a review of the latter at: <http://primate.uchicago.edu/2004PC.pdf> (accessed 8 Mar 2015).

wire surrogate—the infants would only go to the wire surrogate to feed when hungry—as soon as their hunger abated, they returned to the cloth covered surrogate. Harlow concluded from these particular experiments that infant monkeys had requirements (social, cognitive, emotional) beyond those of (just) nutrition (milk) in their early years.

In the field of geology, the work of Nevil Maskelyne and colleagues gave an initial indication of the density of the earth (Danson 2009).²⁶ Their work was based on the notion that a pendulum, placed near a mountain in a uniform gravitational field, would shift from the true vertical. This shift could be measured against a reference such as a fixed star and—given Newton’s proposal that force is proportional to the mass of an object—the density of the earth could be calculated. Isaac Newton himself, in the *Principia* had indicated that this should be possible but had discarded the idea as he believed the instrumentation of the day would not be able to detect the small changes in the shift of the pendulum.²⁷

Just over a century later in 1772, Maskelyne, the Royal Astronomer to George III, believed that the instrumentation at the Royal Observatory in Greenwich was up to the task that Newton had set. The French astronomer Pierre Bouguer had carried out Newton’s proposal of using a mountain in South America some decades earlier—but had not met with much success on account of numerous technical obstacles (Danson 2009, 40–42; 97–98).

Maskelyne met with greater success at a mountain in central Scotland, Schiehallion (chosen for its symmetry). The investigation was divided into two stages. The first entailed measurement of the deflection of the pendulum with respect to positions of fixed stars for which two observatories were built—one on the north side and one on the south. The measurements taken were in the astronomical measure of arc minutes. The other stage of the investigation involved the survey of the mountain in order to measure its volume. These measurements were expressed in terms of height (feet/inches). The work took until 1778 to complete and the final density of the earth was computed to within 20 % of that calculated by Henry Cavendish some twenty odd years later using a torsion balance to measure the attraction between two lead spheres.

Staying within geology and in Scotland, James Hutton’s extensive investigations on soil erosion helped significantly shape understanding of landscape formation (Dean 1992).²⁸ Over a span of decades, Hutton made extensive surveys and measurements of various areas of Britain as well as France, Belgium and Holland. Much of this work was subsequently the starting point for Charles Lyell and Charles Darwin in their work on geology (Rudwick 2005a).²⁹ The first outline of Hutton’s work was circulated as *Abstract of a Dissertation Concerning the System of Earth, its Duration and Stability* in 1786. His work consisted of analysis of rock strata and analysis (chemical, thermogenic) of different kinds of rock formations (granite and gneiss, sediment[ary] and volcanic [igneous] as well as the identification and recording of the frequency of the occurrence of fossils in these different rock strata.³⁰ The records of his results consisted of temperature readings at which different kinds of rocks changed appearance, the recording of what these changes entailed, the

²⁶ See Chapters 11–15. See also Smallwood (2009).

²⁷ See the edition of Andrew Motte’s translation of *Newton’s Principia: The Mathematical Principles of Natural Philosophy*, pp. 527–528.

²⁸ Also see Repcheck (2004). For reception of Hutton’s work amongst his contemporaries, see Dean (1973). For a synopsis of Hutton’s biography see his entry in the *Dictionary of Scientific Biography*.

²⁹ See also Rudwick (1985, 2004, 2005b).

³⁰ A visitor to Hutton’s home in Edinburgh remarked, “his study is so full of fossils and chemical apparatus of various kinds that there is barely room to sit down”.

recording of which (if any) rock kinds reacted with different kinds of chemicals, extensive classification and frequency tabulations of fossil finds, numerous drawings of fossil finds and rock strata.

I now want to turn to physics—the principal focus of study for Hacking. In the early part of the twentieth century, Robert Millikan conducted a series of investigations to establish that the charge of the electron was quantized (had a discrete fundamental value) and occurred in situ as multiples of this value rather than a continuum as had been previously proposed by Thomas Edison, amongst others (Holton 1978).³¹

The received narrative of Millikan's investigations is presented as an ingenious use of the cloud chamber developed by Charles Wilson (Franklin 1986, 216). In Wilson's original, within a sealed container, ions act as loci around which water droplets can form. Wilson used a sealed container filled with air and water vapour at the point of condensing—a supersaturated environment—which he produced using a vacuum pump for first compressing and then expanding air inside the sealed container ('chamber'). Any charged particle in the container containing this supersaturated mixture causes ionization as it moves. This ionization acts as loci around which vapour ('cloud') forms as a consequence of condensation. The movement (or fall) of this 'cloud' in this 'chamber' under gravity can be detected via a viewer (short focal distance telescope) and the visible ionization path measured (by calibrating the eyepiece of the telescope). If an electric field is applied (vertically) across the chamber (in the form of two charged plates—positive at the top and negative at the bottom with a DC voltage applied to each plate via a battery)—then the change in the rate at which the cloud moves/falls under gravity can be detected. Measurement of the velocities of the fall of the cloud under just gravity and then with a known voltage should determine the charge on the electron.

J. J. Thompson had attempted to measure the charge on the electron in this way but had tried to measure the charge of the whole cloud and had met with little success—owing in the main to practical obstacles (Goodstein 2001, 54).

Millikan, in attempting the same as Thompson, found that applying a much greater electric field across the charged plates resulted not in the cloud being suspended, as had been predicted, but most of the cloud dispersing, leaving only a few drops suspended between the plates. Millikan deduced that working with individual droplets would overcome many of the logistical and numerical obstacles that Thompson had faced in working with a whole cloud (*ibid.*).

Millikan (and his graduate students) set about repeating Thompson's work with single droplets of water but found no success as the single water drops tended to evaporate quickly, making reliable measurements impossible. They thus set about adapting Wilson's cloud chamber, as well as Thompson's method, over a period of some years. The appearance of simplicity in Millikan's final investigative set-up belies the various stages it took for the investigation to mature.

The first issue they had to overcome was that of evaporation. They did this by replacing water drops with substances whose evaporation rate would have a negligible effect on their measurements. The first substance they used was oil with a low vapour pressure that would easily form a spray (they produced the oil drops as a spray with a perfume atomizer using watch oil bought at minimal cost at a local market). Although Millikan's published work dealt with the results from work done with the oil drops, Millikan and his group had done the same investigations with glycerine and mercury. Evaporation issues were only the first

³¹ See also Franklin (1981), Barnes et al. (1996), and Goodstein (2001). Also see Niaz (2005) for an appraisal of the studies of Holton, Franklin, Barnes et al. and Goodstein.

of many obstacles they had to overcome to arrive at a working system, including *inter alia*: temperature within the chamber affecting viscosity of the air, allowing for the evaporation of the oil (as well as glycerine and mercury)—however minimal, the motion of the air inside the chamber, the fluctuation of the charge applied by the battery source (Franklin 1981).³²

Their final set up (which led to Millikan's published work on the quantization of the charge on the electron in 1910 and 1913) ran as follows. Within a sealed container Millikan et al. placed two charged plates 16 mm apart which were connected to a DC supply (battery). Above the top plate was an aperture through which the atomizer could spray droplets into the container. The top charged plate had a small aperture through which oil (glycerine, mercury) droplets could drop under gravity. In the space between the two plates were three apertures: one for the short focal telescope to view the drops, one for a light source in order to be able to see the drops and the other for an X-ray source to induce ionization of the air. The actual measurements were made in units of time—in seconds (range 11–19 s)—taken for an oil drop to move across a known distance of 10.21 mm (Millikan 1913).³³ The voltage (when used) was set at 5 kV. Differences in the time measured for an oil drop to move across the given distance (10.21 mm) under (just) gravity and then under the given current (5 kV) allowed Millikan to calculate the charge and, with repeated measurements under varying conditions, deduce that the charge was quantized.³⁴

As with Beaumont and Kettlewell, how does Hacking's observation versus experiment account map onto the cases outlined above from various fields of scientific enquiry?

As with Beaumont and Kettlewell, the observation part of Hacking's account—a means of detection—can be identified easily in the mentioned cases:

In Harlow's work with infant monkeys, his measurements of time spent with each surrogate;

In Maskelyne's investigations on the density of the earth, measurements of arc minutes to measure the shift in the pendulum from the true vertical with respect to fixed star positions and those of height in terms of feet and inches in order to measure volume;

In Hutton's case, the results of chemical and heat changes of different kinds of rock formations, drawings of fossils and rock strata as well as frequency tabulation in the form of integer numbers related to recording of numbers of fossils found in various rock formations;

In Millikan's case, the measurements of the time taken for an oil drop to traverse the distance of 10.21 mm.

Again as with the cases of Beaumont and Kettlewell, it is much more difficult to see how Hacking's category of experiment fits with these cases. In none of them is it clear to see where the 'creation of phenomena' with its emphasis on 'pure state', as we saw with the Hall or Josephson Effects, lies:—emotional attachment for Harlow, density for Maskelyne, landscape erosion for Hutton—even in physics—Millikan's quantization of charge.

Hacking's observation versus experiment account—as a means of delineating scientific experimentation as part of *practice*—thus appears not very helpful when faced with cases

³² Also see Franklin (1986, 215–224).

³³ See in particular pp. 124–125.

³⁴ Millikan's conclusions were contested amongst specialists in the field for more than a decade after publication of this work; see Holton (1978) in particular; for a defence of Millikan, see Goodstein (2001).

from a range of fields of scientific enquiry as those described. Even within modern physics—as Millikan’s work shows—Hacking’s account has limited use.³⁵

4 Experiment and Observation as Processes

David Gooding notes that it is in facing real accounts of scientific experimentation that what he calls “the familiar distinction between observation and experiment” collapses; calling the distinction an “artifact of the disembodied, reconstructed character of retrospective accounts” (Gooding 1992, 68). We should then perhaps not be surprised that Hacking’s observation versus experiment framework does not survive intact when put to the test in a range of cases of scientific experimentation³⁶ such as those described.

Hacking’s account, in its attempt to reify and stipulate the notion of experiment, fails to capture the range and complexity of actions (mental and physical) entailed in what is indicated by experiment in scientific practice. If we return to the examples of scientific experimentation described above, in all cases—some undertaken over decades—the investigations consisted of the accumulation of parts: in Beaumont’s case, his *in vivo* as well as his *in vitro* work; in Kettlewell’s case, his struggles in finding the appropriate control landscapes; in Harlow’s case, trials with different kinds of ‘soft material’ used as a surrogate with which the infant monkeys could identify; Maskelyne’s case consisted of two distinct parts—the astronomical measurements made in the two observatories and the land survey of Schiehallion which followed the astronomical part and took nearly 2 years to complete (due to weather conditions); Hutton’s work on investigating rock strata and formations, and their relationship with the age of the earth, took decades and consisted of two distinct parts—analysis of the rock strata and work with fossils; Millikan’s experiment—which appears straightforward—went through a number of stages as it was optimized for substances (different kinds of oils, glycerine and mercury) and conditions (such as temperature, air viscosity) and calibration (different scales used to measure distance).

Looking at these examples of scientific work, should we refer to them as experiments or a series of experiments? Gooding proposes a potentially helpful way of thinking about this question. Gooding asks us not to talk about experiment but experimentation and think of it

³⁵ In defence of Hacking, his principal aim in *Representing and Intervening* in making his observation experiment distinction is in service of other philosophical ends such as entity realism. Further, within its own time, Hacking’s drawing of a polarity between observation and experiment served the purpose of challenging the hitherto identification of experiment with observation (as a perceptual rather than a detection form). One may therefore reasonably posit that the criteria Hacking puts forward as his description of experiment should not be applied rigidly. However, I believe he appears quite committed to his stipulation of experiment as ‘creation of phenomena’ in a formalistic way—he emphasizes the ‘creation’ part of ‘creation of phenomena’ in his discussion at length explicitly and reiterates this commitment by the examples with which he chooses to engage at length—certain kinds of cases from physics (Hall effect, Josephson effect) while consciously stepping away from others such as those such as the work of William Beaumont which, as we see above, are not easily receptive to the observation versus experiment account. In addition, Hacking underlines his commitment to ‘creation’ in his description of experiment as ‘creation of phenomena’ as well as ‘purity’ of said in his later work (Hacking 1992, 37; here Hacking uses the photoelectric effect as an exemplar). I think therefore it is not unreasonable to take Hacking at his own (repeated) word. If one does do that, then it appears from our discussion that Hacking’s stipulation of experiment as ‘creation of phenomena’, and his emphasis on ‘pure state’, gives rise to anomalies in a range of cases of scientific practices as shown.

³⁶ Hans Radder uses ‘scientific experimentation’ in a way which reflects the importance of the processual nature of scientific work and practices, (Radder 2003b, 15).

as a process³⁷ (1992, 65–67). Hasok Chang too, using a different lexicon, asks us to think of experiment as a series of activities which themselves are composed of processes (2011, 208–210).³⁸ The idea of process fits well with the range of examples described above. However, does viewing experimentation as a process help us in delineating experiment from observation as categories distinct from each other as Hacking does in his account? We have seen that observation, for Hacking, is a means of detection. However, this too, more often than not, tends to be a process. If we take just one case from amongst those that Hacking categorizes as observation this becomes apparent. One example (cited earlier) Hacking uses as an example of an observation is of the detection of solar neutrinos (1983, 182). The detection of solar neutrinos runs thus.³⁹

Solar neutrinos are produced as a by-product of nuclear fusion in the core of the sun (Pinch 1985, 5). As they are highly unreactive, they can pass through the outer layers of the sun and pass through the earth's atmosphere (predominantly) in the state in which they were produced in the sun's interior. The fact that they are highly unreactive of course makes them very difficult to detect. In the 1960s, Raymond Davis Jr. developed the methodology for detection of solar neutrinos of a particular kind (pp or proton–proton). A 100–400 k gallon container of dry cleaning fluid (perchloroethylene) was buried over a mile underground (in a disused mineshaft). The chlorine in perchloroethylene contains traces of a radioactive chlorine isotope (37) with which the solar neutrinos are able to react. The reaction between the chlorine isotope and solar neutrinos gives rise to the production of a radioactive argon isotope (37). This argon isotope is allowed to accumulate over typically a month (not longer as the half-life of the argon isotope (37) is 35 days). Other isotopes of argon (36 or 38) are added which aids the argon isotope (37) to bind to the helium gas, which is flushed through the container to remove the argon isotope (37). The helium containing the argon isotope (37) is then passed through pre-cooled charcoal, which collects the argon isotope (37). It is the decay of the argon isotope (37), which can be detected via a pre-calibrated Geiger counter.

It is apparent that this process of detection of solar neutrinos is exactly that—a process—with a multitude of different manipulations, practices and interpretations. In fact, very similar to the practices and processes of experimentation described in the cases above, and act for Gooding's claim that in the face of real cases of scientific practice, to (try to) distinguish between observation and experiment is futile. In Hacking's account, experiment is defined in a verbal phrase—'creation of phenomena'—based on activity and 'endless different tasks' (Hacking 1983, 230). However, observation too in this account entails the same as a means of detection and in most fields of scientific enquiry requires 'endless different tasks'—one could replace the case of the solar neutrinos described above by numerous others including *inter alia*: other sub-atomic particle decay experiments in physics, chain reactions in chemistry (organic and inorganic), cascade and chain reactions in biochemistry. Both observation and experiment in practice involve undertaking various activities, manipulations, interventions and interpretations.

³⁷ Hacking too has used the term 'experimentation', 'Experimentation has many lives of its own' (1983, 165). However, Hacking uses 'experimentation' in contrast to theory, saying "..., let us not pretend that the various phenomenological laws of solid state physics required a theory—any theory—before they were known. Experimentation has many lives of its own" (*ibid.*). In contrast, David Gooding uses 'experimentation' as a process *qua* process.

³⁸ Also see Rouse (1996).

³⁹ This case has been analyzed in detail by Shapere (1982) and Pinch (1985) as well as dealt with in summary by Bogen and Woodward (1988, 316).

Hasok Chang has proposed that the pursuit of a systematic analysis of activities entailed in scientific practice is a worthy goal (Chang 2011). He proposes a “philosophical grammar of scientific practice” (ibid, 206) where he tentatively draws a taxonomy of what he says are only some of the “epistemic activities” entailed in scientific practice including, *inter alia*, Describing, Explaining, Hypothesizing, Testing, Observing, Measuring, Classifying, Representing, Modelling, Simulating, Synthesizing, Analyzing, Causing, Abstracting, Idealizing. David Gooding too has made an attempt to describe scientific practice (Gooding 1990, 1992) albeit in diagrammatic form—in what he calls “experimental maps” (Gooding 1992, 67), rather than discursively as Chang does. However, both are interested in considering the nature (in Chang’s case) and ordering (in Gooding’s case) of the multitude of epistemic activities entailed in scientific practice; rather than in differentiating between them as Hacking appears to be doing in his experiment versus observation account, as a means of categorization in an ‘either/or’ way. It is therefore not surprising that Hacking’s account, based as it is on casting observation and experiment as polarities, rather than seeing them as parts of a continuum within the process of experimentation, is not able to adequately account for cases of scientific experimentation other than from the very narrow area of physics on which he chooses to focus, such as high-energy lasers and such like.

The categorical distinction Hacking makes between observation and experiment would seem to rely in the main on his very particular definition of experiment—‘creation of phenomena’ and the many issues arising out of his stipulation of ‘creation’ as demonstrated earlier.⁴⁰ If one disregards Hacking’s stipulation of ‘creation’ in his definition for experiment, then it is difficult to see how a category distinction can be maintained between observation and experiment. As we saw earlier, both encompass generation of data so this could not act as an adequate marker.

If one were to broaden Hacking’s notion of experiment to consider another candidate as a marker for making a distinction between observation and experiment, then one of the more obvious ones is intervention. Lorraine Daston, in her account of practices of observation in the period 1600–1800, gives a glimpse of the various views circulating around the projected distinction between observation and experiment during this period (Daston 2011, 85–87). Amongst these views, many gave importance to intervention (or its synonyms) as an important marker for distinguishing between observation and experiment. However, even then (that is, before the use of increasingly complex instrumentation became ubiquitous in scientific experimentation and practice in the modern age) some could see ambiguities arising. Gottfried Wilhelm Leibniz notes, “there are certain experiments that would be better called observations, in which one considers rather than produces the work” (ibid. 86).

This attempt to cast ‘intervention’ as a potential marker to distinguish between observation and experiment as categories, of course, pre-dates the crucial nineteenth century shift towards the dissipation of any qualitative difference between ‘seeing’ with help such as with instrumentation with its associated range of interventions, and that without—as the works of Clary, Hoffmann and Schikore amongst others have shown.⁴¹ Looking back to the case of William Beaumont and his work on human digestion we see that: both his *in vivo* and *in vitro* work needs intervention (of some kind) to be satisfactorily completed making it impossible to distinguish (in any consistent and coherent way) between what should be observation and what experiment. The case of observation of solar neutrinos also makes

⁴⁰ Others too have noted the ambiguities arising out of the very particular way Hacking stipulates his category of experiment, See Feest (2011, 63–64).

⁴¹ See earlier references to each of these authors.

very clear, with its numerous and complex manipulations, that intervention is not a reasonable candidate for acting as a category distinguisher between observation and experiment. The category distinction Hacking makes between observation and experiment thus rests very much on his narrow definition of experiment—‘creation of phenomena’—with the anomalous consequences that arise when this definition is used across various instances of scientific experimentation as the cases earlier demonstrate.

Other accounts than those of Chang and Gooding have also been advanced to analyse scientific experimentation; although interestingly—but perhaps unsurprisingly in light of our discussion thus far, very few use Hacking’s nomenclature of observation/experiment. Like Gooding and Chang, most believe that scientific experimentation should be viewed as a continuous process rather than one entailing discrete parts—and the terminology used underlines this sense of continuousness. Friedrich Steinle and Richard Burian have coined the term ‘exploratory experimentation’ which gives the same sense of the continuousness of the experimental process as do Chang and Gooding in their work: Steinle working on the early history of electromagnetism (Steinle 1997, 2002), and Burian in his work on molecular biology (Burian 1997, 2007).⁴²

Steinle, in analysing the experimental work of Oersted, Ampere and Faraday draws a distinction between two kinds of experiments: those designed with the specific aim of tracing particular effects which were expected because of the field of knowledge within electromagnetism at the time, and those experiments set up where the investigators had, what Steinle calls, “no theory—or—even more fundamentally—no conceptual framework” (Steinle 1997, S65). Richard Burian, too, has used the term ‘exploratory experimentation’. Burian first used the term in his analysis of the work of Jean Brachet’s experiments on the localization and functioning of nucleic acids (Burian 1997). Burian examines the research of Brachet’s on the distribution of nucleic acids across cell life cycles. Burian shows that Brachet was not guided by theoretical considerations about how the nucleic acids may be distributed across the lifetime of cells in various organisms. This was very much in contrast to Brachet’s peer, Francis Crick, working on the same subject matter, who was much more theoretically inclined which greatly influenced the kinds of experiments he chose to undertake (*ibid.*, 40–41). Burian therefore also uses the term in the same sense as Steinle insofar as to distinguish a particular kind of experimentation from theory. Since this inception, ‘exploratory experimentation’ has gradually gained more definitive structure: for example, it is clear that it is not the case that exploratory experimentation is free from theory—rather, the question is how theory influences the experimental process thus leading to a distinction between ‘theory-directed’ and ‘theory-informed’ (Waters 2007, 277); leading to calls for the creation of its own sub-structure that can account for historical cases more adequately than it does in its present form (O’Malley 2007).

Another term, ‘experimental system’, has been used within writing on epistemology of experimentation, which conveys the same sense of ‘continuousness’ as does exploratory experimentation. The term ‘experimental system’ was first used by Hans-Jörg Rheinberger to refer to the experimental research on protein synthesis (1997). Rheinberger describes experimental systems as “systems of manipulation designed to give unknown answers to questions that the experimenters themselves are not yet clearly to ask” (*ibid.*, 28). As with the term ‘exploratory experimentation’, Rheinberger uses ‘experimental system’ in order to distinguish it from a theory-dominated approach, arguing that experimental work in biology always “begins with the choice of a system rather than with the choice of

⁴² Rose-Mary Sargent too uses the term but for descriptive rather than analytical purposes (Steinle 1997, S71).

theoretical framework” (ibid., 25). Other similar terms to experimental systems have been used to indicate scientific experimentation as a process with the sense of continuousness embedded at their centre such as ‘manipulable systems’ (Turnball and Stokes 1990) and ‘production systems’ (Kohler 1991).

All these terms (exploratory experimentation, experimental system, manipulable system, production system) and their respective accounts emerge with the aim of distinguishing them from theory-dominated accounts such as hypothesis testing. None of these accounts seek to do what Hacking does with his stipulation of experiment: distinguish between different kinds of activities and interventions within the process of scientific experimentation. Hasok Chang’s aim in delineating the various kinds of activities and interventions involved in scientific practice uses a descriptive rather than a stipulative approach (Chang 2011). James Woodward uses the terms observation and experiment as distinct categories but with the aim of defending what he calls a “manipulationist account of causation” rather than in an attempt to delineate scientific experimentation as process (Woodward 2003, 88).⁴³

However, elsewhere, James Woodward, together with James Bogen, has sought to put forward an account which seeks to specifically delineate the process of scientific experimentation. It completely abandons the vocabulary of observation and experiment and uses data and phenomena instead.⁴⁴ Bogen and Woodward initially put forward their data phenomena account in 1988 (Bogen and Woodward 1988).⁴⁵

Bogen and Woodward tell us that data should be thought of as that which provides evidence for the existence of phenomena (Bogen and Woodward 1988, 305). Data can (usually) be detected. However, data (usually) cannot be predicted. Phenomena, on the other hand, can only be detected through the use of data (Bogen and Woodward 1988, 306). Examples of data include bubble chamber photographs, patterns of discharge in electronic particle detectors and records of reaction times and error rates in psychological experiments. These instances of data provide evidence for the following phenomena respectively: weak neutral currents, decay of the proton and chunking effects in human short-term memory.

Bogen and Woodward analyse a number of examples to illustrate the distinction between data and phenomena (Bogen and Woodward 1988, 308–322). Examples they use

⁴³ See also Woodward (2013) for a review of the topic where he deals with the different positions on the subject matter, including his own.

⁴⁴ In so doing, they avoid linguistic oddities such as, ‘What has been shown as well is that, in actual practice, making scientific observations often includes doing genuine experiments’ (Radder 2003b, 15).

⁴⁵ Since then it has been re-stated by Woodward on a number of occasions (Woodward 1989, 2000, 2011). In these revised versions, however, Woodward has been more concerned with dealing with the relationship between this account and its relationship with scientific theory. The data phenomena account has been contested on various grounds. These contestations have tended to focus on two areas. First, whether it is reasonable to draw a distinction between data and phenomena at all whereas both should be viewed as patterns within data sets (Glymour 2000). Further, even if one were to draw a distinction between them, how one does this—in particular, the role of assumptions in this process (McAllister 1997). Woodward (2011, 175–176) amongst others (Apel 2011, 27–31), have responded to these points in recent years. The other area of focus has been the relationship of data and phenomena within Bogen and Woodward’s account to theory (Schindler 2007, 2011); in particular, as it relates to the influence of theory on observation and its implications for reliability. Woodward counters this view in detail (Woodward 2011, 172–174) by suggesting that the charge that data in the data phenomena account is assumed to be independent of ‘additional assumptions’ or ‘theory free’ is unfounded. However, where used to delineate scientific practice qua practice, it appears reasonably robust—as even its detractors concede (Schindler 2011, 54).

to show what they mean include the melting point of lead (from chemistry) and weak neutral currents (from physics).

Bogen and Woodward analyse the following statement about the melting point of lead to show what they mean by their data phenomenon distinction: 'lead melts at 327.5 ± 0.1 degrees centigrade'. However, this is not what actually happens. It is not possible to determine the melting point of lead by taking a single thermometer reading.⁴⁶ It is necessary to take a series of measurements. Even if systematic errors are reduced, there will be variations in the thermometer readings such as to give a scatter of results that all differ from each other, even if potential sources of error are minimized. The figure 327.5 represents the mean of the scatter of thermometer readings while the figure 0.1 represents the standard deviation.

Within Bogen and Woodward's account, the thermometer readings fall within the category data while the calculated melting point, 327.5 degree centigrade fall within the category phenomena. It is the latter, phenomena, which becomes the object of systematic scientific explanation. Thus, in the case of the melting point of lead, the figure 327.5 degree centigrade becomes the object of explanation in terms of the molecular structure of metals. This would be expressed in terms such as metallic bonding mechanisms and type of co-ordination.

The data too can become the object of scientific explanation. However, the terms in which explanations regarding data would be made would be different from those made for phenomena. Explanations related to data would include discussion of the accuracy of the thermometer, the purity of the lead sample used, the point at which the thermometer is taken (when the sample of lead starts to melt, at mid-way, when the sample has all melted), the reliability of the heating mechanism and such like. These terms and considerations are very different to those related to discussions in terms of molecular structure. In Bogen and Woodward's account, thus, data are distinguished from phenomena by the fact that the terms in which phenomena are explained are distinct from the terms in which data are explained.

Another difference between data and phenomena in Bogen and Woodward's account relates to phenomena possessing regular characteristics, which are detectable from very different kinds of data (or evidence). Bogen and Woodward use the following example to show what they mean.

The evidence for the existence of the phenomenon of weak neutral currents came from two different kinds of investigations. One was at CERN in Switzerland and the other at the NAL (National Accelerator Laboratory) in the US. The data from CERN comprised of bubble chamber photographs (where the detection method depended on the formation of bubbles) while that from NAL consisted of patterns of discharge in particle detectors (where the detection method registered the passage of charged tracks by electronic means). These two very different kinds of data—from very different kinds of apparatus—provided the evidence for the same phenomenon: the weak neutral current.

The terms of explanation for the phenomenon, the weak neutral current, comprise the interaction of the Z particle with the weak force—this is common in both cases: from the data from CERN as well as the very different data from the NAL. However, the terms of explanation of the two different data sets have very little in common: the data set from CERN comprises of terms consisting of, inter alia, the nature of the neutron beam, the shielding chamber, the size of the chamber as well as the type of liquid used in the chamber. The data from the NAL, however, required explanation in terms of, inter alia, the

⁴⁶ For details of how the melting point of lead is determined under laboratory conditions, see Bogen and Woodward (1988, 309–310).

strength of the magnetic field, the characteristics of the calorimeter used to stop, absorb and measure a particle's energy, and the nature of the tracking device.

For Bogen and Woodward, phenomena are "in the world, as belonging to the natural order itself and not just to the way we talk about or conceptualize that order" (1988, 321). This may include, "particular objects, objects with features, events, processes, and states" (Bogen and Woodward 1988, 321). To Bogen and Woodward, the key feature of phenomena is that they be the objects of general scientific explanation, rather than the particular explanations, which are the characteristic feature of data, and from which they are distinct (1988, 322). Data are highly localized and idiosyncratic and demand explanations that are framed in very different terms to that of phenomena for which they act as evidence (Bogen and Woodward 1988, 319).

Mapping the data phenomena account onto the cases cited earlier would thus yield the following outcomes. For Beaumont, the data relates to the results of digestion from both the *in vivo* and *in vitro* parts of his investigation and the explanatory terms in which they are framed relate to degrees of acidity, temperature readings and measurements of time, while the terms in which the explanatory terms for the phenomena are framed include peristaltic movement, the anatomy and composition of gastric cell types and the physical topography of the stomach with respect to the rest of the gastrointestinal tract. For Kettlewell, knowledge about data would relate to what kind of moth is conspicuous on which colour bark, the numbers of different kinds of moths surviving exposure to predation in the cage, what kind of bait is used to trap surviving moths in native conditions. The phenomenon is accounted for by discussion in terms of the changing colour of the landscape owing to pollution and degrees of conspicuousness to predators. For Harlow, the data would be framed in terms such as time spent with each kind of surrogate while discussion of phenomena would be framed in terms of emotional bonding, cognitive support and imitation along with the need for physical contact with an animate like material. For Maskelyne, the data would be framed in terms of arc minutes for the astronomical measurements and feet/inches for the physical survey of the mountain while the phenomena concerned (density of the earth) was expressed in terms of a numerical value (4500 kg/m^3). For Hutton, the data was framed in terms of chemical, temperature and field measurements and anatomical differentiation in fossil records while the phenomena was framed in terms of soil erosion and the influence of the physical elements (wind, water) on this erosion as indicative of changing climate and its correlation with fossil deposits. For Millikan, the data would be framed in terms of time taken for an oil drop to travel a distance of 10.21 mm, the viscosity of the oil, the temperature of the cloud chamber while the phenomena was framed in terms of the discrete nature of the charge carried by an electron, its interactions with other parts of the atom, the nature of these interactions and the value of the charge itself ($1.5924 \times 10^{-19} \text{ C}$).

Although both Hacking and Bogen and Woodward aim, in each of their accounts of delineating scientific experimentation, to use a criteria based approach, the criteria they use are very different.

It is worth noting some points of conceptual overlap, as well as departure, between the two accounts—notwithstanding the different lexicon of each. There is considerable congruence between Hacking's category of observation (or results(s) of observation) and Bogen and Woodward's category of data—the outcomes when mapping each account onto the mentioned cases makes this clear: Beaumont's time measurements for digestion, Kettlewell's survival rates, Harlow's record of time spent with each kind of surrogate, Maskelyne's astronomical and survey measurements in arc minutes and units of height, Hutton's complex and varied array of temperature, chemical, maps and diagrammatic records and Millikan's measurements of time taken for an oil drop to traverse 10.21 mm. It is when we turn to

consider the second part of each account—Hacking’s experiment (creation of phenomena) and Bogen and Woodward’s phenomena—that the conceptual overlap starts to dissipate. At first glance, both use phenomena in a similar way. For Hacking, it is a discernable regularity, “[a] phenomenon is *noteworthy*. A phenomenon is *discernable*. A phenomenon is commonly an event or process of certain type that occurs regularly under defined circumstances” (Hacking 1983, 221, emphasis in original). For Bogen and Woodward, a phenomenon has “stable, repeatable characteristics which will be detectable by means of a variety of different procedures, which may yield quite different kinds of data” (Bogen and Woodward 1988, 317). However, the common use of vocabulary—both in terms of naming and description—should not prevent us from noting the different ways each is conceptualized in each different account—as Bogen and Woodward themselves note (*ibid*, 306), suggesting that although there are certain similarities between their notion of phenomena and that of Hacking’s, they find Hacking’s notion limited insofar as Hacking’s notion “is not correct as a general characterization of phenomena” and continue on to say that the “features which Hacking ascribes to phenomena are more characteristic of data”. Others too have noted the ambiguity in Hacking’s description of the relationship between experiment and phenomena.⁴⁷

Bogen and Woodward here identify the principal limitation in Hacking’s observation versus experiment account as a means for (systematically) delineating scientific experimentation as practice: in the observation versus experiment account (the results of) observation and experiment *both* are ways of generating data. It is therefore perhaps not surprising that earlier we saw quite anomalous and ambiguous outcomes on mapping Hacking’s account on to cases of scientific experimentation from a range of fields of enquiry.

Bogen and Woodward use an explanation,—or what Rheinberger has called ‘epistemic object’—based,⁴⁸ criteria. Hacking, however, uses a narrowly construed criteria, centred on (kinds of) action/activity/intervention in his stipulation of experiment as ‘creation of phenomena’. This stipulative approach, as we have seen, has limited value when used in practice across a whole range of fields of enquiry.

5 Concluding Remarks

We have seen from our discussion that the observation versus experiment account has significant weaknesses as a means of delineating scientific experimentation within scientific practice—across a range of cases from various fields of scientific enquiry. This would suggest that the experiment versus observation framework—where observation and experiment are cast as polarities, rather than as complements of each other—as Hooke and Boyle did—is not a sound basis on which to make value judgments.

Acknowledgments I would like to thank Emilie Savage-Smith and Nick Jardine for their comments during the early gestation of this work. I am also very grateful to the anonymous referees, as well as the Editors, for their very helpful comments during review.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

⁴⁷ See Feest (2011, 63–64).

⁴⁸ See Rheinberger (1997, 28).

References

- Ainsworth, M. D. S. (1991). Attachments and other affectional bonds across the life-cycle. In B. Caldwell & H. Ricciuti (Eds.), *Review of child development research* (pp. 1–94). Chicago: The University of Chicago Press.
- Anstey, P. (2014). Philosophy of experiment in early modern England: The case of Bacon, Boyle and Hooke. *Early Science and Medicine*, 19, 103–132.
- Apel, J. (2011). On the meaning and the epistemological relevance of the notion of scientific phenomena. *Synthese*, 182, 23–38.
- Aristotle. (1962). *Meteorologica*. Introduction and Translation by H. D. P. Lee. London: Heinemann, Loeb Classical Library Series.
- Barnes, B., Bloor, D., & Henry, J. (1996). *Scientific knowledge: A sociological analysis*. Chicago: The University of Chicago Press.
- Beaumont, W. (1833). *Experiments and observations on the gastric juice and the physiology of digestion*. Boston: F. P. Allen.
- Bernard, C. (1957). *An introduction to the study of experimental medicine*. New York: Dover Publications Inc.
- Blum, D. (2002). *Love at Goon Park: Harry Harlow and the science of affection*. Cambridge, MA: Perseus.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, 97, 303–352.
- Brandon, R. N. (1999). Introduction. *Biology and Philosophy*, 14, 1–7.
- Burian, R. (1997). Exploratory experimentation and the role of histochemical techniques in the work of Jean Brachet, 1938–1952. *History and Philosophy of the Life Sciences*, 19, 27–45.
- Burian, R. (2007). On micro RNA and the need for exploratory experimentation in post-genomic molecular biology. *History and Philosophy of the Life Sciences*, 29, 285–312.
- Chang, H. (2011). The philosophical grammar of scientific practice. *International Studies in the Philosophy of Science*, 25, 205–221.
- Collins, H. (1985). *Changing order*. Chicago: The University of Chicago Press.
- Crary, J. (1992). *Techniques of the observer: On vision and modernity in the nineteenth century*. Cambridge, MA: MIT Press.
- Crary, J. (2001). *Suspensions of perception: Attention spectacle and modern culture*. Cambridge, MA: MIT Press.
- Danson, E. (2009). *Weighing the world: The quest to measure the Earth*. Oxford: Oxford University Press.
- Daston, L. (2011). The empire of observation. In L. Daston & E. Lunbeck (Eds.), *Histories of scientific observation* (pp. 81–113). Chicago: The University of Chicago Press.
- Daston, L., & Galison, P. (2007). *Objectivity*. New York: Zone Books.
- Daston, L., & Lunbeck, E. (2011). Introduction. In L. Daston & E. Lunbeck (Eds.), *Histories of scientific observation* (pp. 1–9). Chicago: The University of Chicago Press.
- Dean, D. R. (1973). James Hutton and his public, 1785–1802. *Annals of Science*, 30, 89–105.
- Dean, D. R. (1992). *James Hutton and the history of geology*. Ithaca: Cornell University Press.
- Dear, P. (1995). *Discipline and experience: The mathematical way in the scientific revolution*. Chicago: The University of Chicago Press.
- Falkenburg, B. (2000). How to observe quarks. In E. Agazzi & M. Pauri (Eds.), *The reality of the unobservable* (pp. 329–341). Dordrecht: Kluwer.
- Feest, U. (2011). What exactly is stabilized when phenomena are stabilized? *Synthese*, 182, 57–71.
- Fodor, J. (1983). Observation reconsidered. *Philosophy of Science*, 51, 23–43.
- Franklin, A. D. (1981). Millikan's published and unpublished data on oil drops. *Historical Studies in the Physical Sciences*, 11, 185–201.
- Franklin, A. (1986). *The neglect of experiment*. Cambridge: Cambridge University Press.
- Franklin, A. (2012). Experiment in physics. In *Stanford encyclopedia of philosophy*. <http://plato.stanford.edu>. accessed on 14 May 2014.
- Galison, P. (1982). Theoretical predispositions in experimental physics: Einstein and the gyromagnetic experiments, 1915–1925. *Historical Studies in the Physical Sciences*, 12, 285–323.
- Galison, P. (1983). How the first neutral current experiments ended. *Reviews of Modern Physics*, 55, 477–509.
- Galison, P. (1985). Bubble chambers and the experimental workplace. In P. Achinstein & O. Hannaway (Eds.), *Observation, experiment and physical science* (pp. 309–373). Cambridge, MA: MIT Press.
- Galison, P. (1987). *How experiments end*. Chicago: The University of Chicago Press.
- Glymour, B. (2000). Data and phenomena: A distinction reconsidered. *Erkenntnis*, 52, 29–37.
- Gooding, D. (1982). Empiricism in practice: Teleology, economy, and observation in Faraday's physics. *Isis*, 73, 46–67.

- Gooding, D. (1990). *Experiment and the making of meaning*. Boston: Kluwer Academic.
- Gooding, D. (1992). Putting agency back into experiment. In A. Pickering (Ed.), *Science as practice and culture* (pp. 65–112). Chicago: The University of Chicago Press.
- Gooding, D. (2000). Experiment. In W. H. Newton-Smith (Ed.), *A companion to the philosophy of science* (pp. 117–126). Oxford: Oxford University Press.
- Goodstein, D. (2001). In defence of Robert Andrews Millikan. *American Scientist*, 89, 54–60.
- Hacking, I. (1983). *Representing and intervening*. Cambridge: Cambridge University Press.
- Hacking, I. (1989). Extragalactic reality: The case of gravitational lensing. *Philosophy of Science*, 56, 555–581.
- Hacking, I. (1992). The self-vindication of the laboratory sciences. In A. Pickering (Ed.), *Science as practice and culture* (pp. 29–64). Chicago: The University of Chicago Press.
- Hagen, J. B. (1999). Retelling experiments: H. B. D. Kettlewell's studies of industrial melanism in peppered moths. *Biology and Philosophy*, 14, 39–54.
- Hanson, N. R. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Hoffmann, C. (2006). *Unter Beobachtung: Naturforschung in der Zeit der Sinnesapparate*. Göttingen: Wallstein.
- Holton, G. (1978). Subelectrons, presuppositions, and the Millikan–Ehrenhaft dispute. *Historical Studies in the Physical Sciences*, 9, 161–224.
- Kettlewell, H. (1955). Selection experiments on industrial melanism in the Lepidoptera. *Heredity*, 9, 323–342.
- Kettlewell, H. (1956). Further selection experiments on industrial melanism in the Lepidoptera. *Heredity*, 10, 287–301.
- Kettlewell, H. (1958). A survey of the frequencies of *Biston betularia* (L.) and its melanic forms in Great Britain. *Heredity*, 12, 51–72.
- Kohler, R. (1991). Systems of production: Drosophila, neurospora and biochemical genetics. *Historical Studies in the Physical and Biological Sciences*, 22, 87–130.
- Kuhn, T. S. (1976). Mathematical vs. experimental traditions in the development of physical sciences. *Journal of Interdisciplinary History*, VII(I), 1–31.
- Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton: Princeton University Press.
- Lindberg, D. C. (2007). *The beginnings of western science*. Chicago: The University of Chicago Press.
- Lloyd, G. E. R. (2004). *Ancient worlds, modern reflections: Philosophical perspectives on Greek and Chinese science and culture*. Oxford: Oxford University Press.
- Lloyd, G. E. R. (2006). *Principles and practices in ancient Greek and Chinese science*. Aldershot: Variorum.
- Maxwell, G. (1962). The ontological status of theoretical entities. In H. Feigl & G. Maxwell (Eds.), *Scientific explanation, space and time: Minnesota studies in the philosophy of science* (pp. 181–192). Minnesota: University of Minnesota Press.
- McAllister, J. (1997). Phenomena and patterns in data sets. *Erkenntnis*, 47, 217–228.
- Millikan, R. A. (1913). On the elementary charge and the Avogadro constant. *The Physical Review*, II, 109–143.
- Newton, I. (1846). *Newton's principia. The mathematical principles of natural philosophy*. Trans. A. Motte. New York: Daniel Adee.
- Niaz, M. (2005). An appraisal of the controversial nature of the oil drop experiment: Is closure possible? *British Journal for the Philosophy of Science*, 56, 681–702.
- O'Malley, M. A. (2007). Exploratory experimentation and scientific practice: metagenomics and the proterorhodopsin case. *History and Philosophy of the Life Sciences*, 29, 337–360.
- Park, K. (2011). Observation in the margins, 500–1500. In L. Daston & E. Lunbeck (Eds.), *Histories of scientific observation* (pp. 15–44). Chicago: The University of Chicago Press.
- Pickering, A. (1981). The hunting of the quark. *Isis*, 72, 216–236.
- Pinch, T. (1985). Towards an analysis of scientific observation: The externality and evidential significance of observational reports in physics. *Social Studies of Science*, 15, 3–36.
- Pomata, G. (2011). Observation rising: Birth of an epistemic genre, 1500–1650. In L. Daston & E. Lunbeck (Eds.), *Histories of scientific observation* (pp. 45–80). Chicago: The University of Chicago Press.
- Pormann, P., & Savage-Smith, E. (2007). *Medieval Islamic medicine*. Washington D.C.: Georgetown University Press.
- Prior, V., & Glaser, D. (2006). *Understanding attachment and attachment disorders: Theory, evidence and practice*. London: Jessica Kingsley.
- Radder, H. (Ed.). (2003a). *The philosophy of scientific experimentation*. Pittsburgh: University of Pittsburgh Press.

- Radder, H. (2003b). Towards a more developed philosophy of scientific experimentation. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 1–18). Pittsburgh: University of Pittsburgh Press.
- Repcheck, J. (2004). *The man who found time*. London: Perseus.
- Rheinberger, H.-J. (1997). *Towards a history of epistemic things: Synthesizing proteins in the test tube*. Stanford: Stanford University Press.
- Rudge, D. W. (1999). Taking the peppered moth with a grain of salt. *Biology and Philosophy*, 14, 9–37.
- Rudge, D. W. (2001). Kettlewell from an Error Statistician's Point of View. *Perspectives on Science*, 9, 59–77.
- Rudge, D. W. (2005a). Did Kettlewell commit Fraud? Re-examining the evidence. *Public Understanding of Science*, 14, 249–268.
- Rudge, D. W. (2005b). The beauty of Kettlewell's classic experimental demonstration of natural selection. *Bio-Science*, 55, 369–375.
- Rudge, D. W. (2006). H. D. P. Kettlewell's research 1937–1953: The influence of E. B. Ford, E. A. Cockayne and P. M. Sheppard. *History and Philosophy of the Life Sciences*, 28, 359–388.
- Rudge, D. W. (2009). H. B. D. Kettlewell's research 1934–1961: The influence of J. W. Heslop Harrison. *Transactions of the American Philosophical Society*, 99, 243–270.
- Rudge, D. W. (2010). Tut-tut tutt, not so fast: Did Kettlewell really test Tutt's explanation of industrial melanism? *History and Philosophy of the Life Sciences*, 28, 359–388.
- Rudwick, M. J. S. (1985). *The Great Devonian controversy: The shaping of scientific knowledge among gentlemanly specialists*. Chicago: The University of Chicago Press.
- Rudwick, M. J. S. (2004). *The new science of geology: Studies in the earth sciences in the age of revolution*. London: Ashgate.
- Rudwick, M. J. S. (2005a). *Lyell and Darwin, geologists: Studies in the earth sciences in the age of reform*. London: Ashgate.
- Rudwick, M. J. S. (2005b). *Bursting the limits of time: the reconstruction of geohistory in the age of revolution*. Chicago: The University of Chicago Press.
- Sabra, A. I. (1996). Situating Arabic science: Locality versus essence. *Isis*, 87, 654–670.
- Sabra, A. I. (2003). Ibn al-Haytham's revolutionary project in optics: The achievement and the obstacle. In J. P. Hogendijk & A. I. Sabra (Eds.), *The enterprise of science in Islam* (pp. 85–118). Boston, MA: The MIT Press.
- Schickore, J. (2007). *The microscope and the eye: A history of reflections, 1740–1870*. Chicago: The University of Chicago Press.
- Schindler, S. (2007). Rehabilitating theory: Refusal of the 'bottom-up' construction of scientific phenomena. *Studies in the History and Philosophy of Science*, 38, 160–184.
- Schindler, S. (2011). Bogen and Woodward's data-phenomena distinction, forms of theory-ladenness, and the reliability of data. *Synthese*, 182, 39–55.
- Shapere, D. (1982). The concept of observation in science and philosophy. *Philosophy of Science*, 49, 485–525.
- Shapere, D. (1993). Astronomy and antirealism. *Philosophy of Science*, 60, 134–150.
- Smallwood, J. R. (2009). John Playfair on Schiehallion, 1801–1811. In C. L. E. Lewis & S. J. Knell (Eds.), *The making of the Geological Society of London* (pp. 279–298). London: The Geological Society.
- Smith, A. M. (1996). *Ptolemy's theory of visual perception*. Philadelphia: The American Philosophical Society.
- Steinle, F. (1997). Entering new fields: Exploratory uses of experimentation. *Philosophy of Science (Proceedings)*, 64, S65–S74.
- Steinle, F. (2002). Experiments in history and philosophy of science. *Perspectives on Science*, 10, 408–432.
- Stuewer, R. (1975). *The Compton effect*. New York: Science History Publications.
- Tiles, J. E. (1993). Experiment as intervention. *British Journal for the Philosophy of Science*, 44, 463–475.
- Turnball, D., & Stokes, T. (1990). Manipulable systems and laboratory strategies in a biomedical institute. In H. E. Le Grand (Ed.), *Experimental enquiries* (pp. 167–192). Dordrecht: Kluwer.
- van Fraassen, B. C. (1980). *The scientific image*. Oxford: Oxford University Press.
- Waters, C. K. (2007). The nature and context of exploratory experimentation: An introduction to three case studies of exploratory research. *History and Philosophy of the Life Sciences*, 29, 275–284.
- Wheaton, B. (1983). *The tiger and the shark*. Cambridge: Cambridge University Press.
- Woodward, J. (1989). Data and phenomena. *Synthese*, 79, 393–472.
- Woodward, J. (2000). Data, phenomena, and reliability. *Philosophy of Science*, 67, S163–S179.
- Woodward, J. (2003). Experimentation, causal inference, and instrumental realism. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 87–118). Pittsburgh: University of Pittsburgh Press.
- Woodward, J. (2011). Data and phenomena: A restatement and defense. *Synthese*, 182, 165–179.
- Woodward, J. (2013). Causation and manipulability. In E. N. Zalta (Ed) *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/win2013/entries/causation-mani/>.
- Worrall, J. (1982). The pressure of light: The strange case of the vacillating crucial experiment. *Studies in History and Philosophy of Science*, 13, 133–171.