advanced draft of
Allan Franklin *What Makes a Good Experiment? Reasons and Roles in Science*.
University of Pittsburg Press 2016. 372+viii pp. hbk.$55.
essentially as it will appear in *Metascience* **102**, 2016

What *does* make a good experiment? This is an important question, understood both as asking what we should want from experiments and as asking how we can get what we want from them. For a central part of trust in science is thinking that well-conducted experiment is essential to understanding the world. Franklin describes and comments on eighteen important experiments, from 1865 to the present, which reveal ways in which an experiment can be valuable. (See the list below.) All except two of them are in physics and all except two of the remaining sixteen are in particle physics, broadly conceived. (The two that are not physics are biology, both in genetics.) I do not think there is an implicit claim that the best experiments tend to be found here. The point is surely instead that one should analyse what one knows best, since the details matter. But there is a price to pay for this focus, since many of the experiments that have the greatest impact on our lives are not in physics, and there are important methodological questions about many of these. The book is a rich source of examples and stimulating observations. It does not give a systematic account of what makes a good experiment, though it does classify experimental virtue under several headings and does justify the attribution of them to particular experiments with a wealth of detail. There are more systematic treatments in other works by Franklin, but one would like to know whether his views have changed, and how they stand up to the details in these case studies

In a preface and a conclusion Franklin addresses more general issues. He states that he does not think there is "an algorithm for a simple set of criteria that would do justice to the wide variety of good experiments." If this means that ingenuity, imagination, and a deep knowledge of one's subject are irreplaceable in designing and carrying out experiments, and these are not governed by any rules that we know, then few will disagree. If it means that there is little to say about why and how particular features of experiment are essential to science, then it is more controversial. In the introduction Franklin divides good experiments into the conceptually important, the technically good, and the pedagogically important. Of course an experiment can have two or more of these virtues. To this he adds that an experiment can be methodologically good, by which he means that it gives good reasons to believe its results. He gives a list of ten strategies which help an experiment be methodologically good, which presumably take some non-mechanical good sense to apply in detail. It is unclear to me how methodological value and conceptual importance are related, since his criterion for conceptual importance is that it supports or undermines a major theory. (This criterion for conceptual importance is explicit in the chapter on the search for magnetic monopoles.) In any case, there is an important distinction between two aspects of experimental value. On the one hand we have aspects that can be at least roughly ascertained at the time, such as adherence to basic methodological norms and provision of clearly telling evidence for a widely accepted or widely discussed theory. On the other hand we have aspects that can only be attributed with hindsight, such as impact on the development of science, or attention to a crucial or significant phenomenon, or proving to be well suited to later expositions of a topic in the light of later developments.

The first work Franklin discusses is Mendel's experiments on hybridization. Franklin describes these as "an almost perfect set of experiments". His reason is that the logic of the experiments is clear and they produce powerful evidence for their conclusions. One should mention also, though Franklin does not emphasize this point, that Mendel simultaneously presents the theory of dominant and recessive characteristics and provides evidence for it. So it is an achievement of theorizing as well as of confirmation. This gives another reason for considering Mendel's experiments among those that are conceptually important, because they introduce new concepts as well as supporting an important theory. (An experiment could in principle introduce valuable new concepts in the course of supporting an unimportant theory or one that did not achieve long-term acceptance.) Franklin does not deduct any points for Mendel's failure to use statistical safeguards that would now be compulsory. The inclusion of experiments which with hindsight might be taken to anticipate the nonconservation of parity under conceptually important experiments is interesting in this connection since they did not have a major impact on physics.

Franklin describes some experiments as crucial, in the sense that they establish a theory definitively. He says this of Mendel's experiments, the experiment by Wu and others showing that parity is not conserved in weak interactions, and, with reservations, of Meselson and Stahl's experiment establishing the mechanism for the replication of DNA. He defends this claim against the old Quine/Duhem point, that enough ingenuity with ancillary hypotheses can always make any evidence consistent with any theory, by relying on a sociological conception of theory acceptance. If all the authorities take an experiment to establish a theory, it is written into textbooks, and so on, then the experiment has definitively established the theory. It would be interesting to consider experiments which were once taken to put beyond doubt theories which later were taken to be seriously flawed. An example might be experiments on the interference of light and their supposed demonstration that light has a wave rather than a corpuscular nature.

Franklin does discuss Millikan's measurement of Planck's constant and in the course of this points out that Millikan did not take himself to have confirmed the corpuscular theory even though he made a central use of the photoelectric effect whose only available theoretical underpinning was corpuscular. And he cites Millikan and others as taking this line because interference so strongly supported the wave theory. A relativized "general acceptance" conception might take into account theoretical resources that are available at one time and not at another, resulting in a conception of crucial experiments given the explanations that could be given and would be taken seriously at a point in scientific development.

The organization of the book suggests a different way of classifying experiments. Part I is about conceptually important experiments (Mendel on genes, Wu on parity, Meselson and Stahl on DNA, Christenson et al on parity, the non-discovery of parity nonconservation in the 1920s), part II about measuring important quantities (Millikan on Planck's constant, Millikan on electron charge), part III about evidence for the existence of entities (neutrinos, η mesons, the Higgs boson), part IV about solving open problems (the absorption and spectrum of β rays, solar neutrinos), and part V about demonstrating that entities or effects do not exist (the 17-keV neutrino, aether drift, a possible fifth force, magnetic monopoles). Franklin may be suggesting that these are the things that experiments can do well, and one might wonder how these interact with the general virtues he describes. He might also be suggesting that all good experiments fit one of these classifications. The obvious topic to discuss, then, though he does not, would be the nature of exploratory experiments, intended to produce data that might stimulate theorizing and suggest other experiments rather than to affect particular existing theories. Experimentalists informally admit to doing more of this than one would gather from their publications.

Franklin describes the experiments that have convinced physicists of the existence of the Higgs boson. He emphasizes the meticulous attempts to exclude extraneous factors. He also emphasizes the sophisticated statistics, though he does not go deeply into the details, in particular the slightly unusual combination of Monte Carlo methods and significance testing. There is a thread here from the physics experiments that are his main focus to experiments in medicine and social science that he does not discuss. He also does not mention the sense of disappointment among many physicists that the results were in accordance with the standard model rather than hinting at dramatic new possibilities. There is a virtue of experiment here that is often at odds with others, of issuing a call for thinking outside the box.

This is a very useful sourcebook of classic experiments, giving enough detail to show what is going on in each of them but discussing enough separate experiments that one can see a variety of experimental virtues. Franklin's attention to detail and his epistemological caution inhibit him from tackling some more adventurous questions. On what range of topics can we hope for evidence that is as convincing as this? Do essential aspects of experiment vary from one discipline to another? How relative to its theoretical context is the evidence provided by even the best experiment? These are important questions, but they require an element of speculation. Perhaps a responsible and well-informed book such as this can provide material for some such further projects.

Adam Morton