Allan Franklin, ***Selectivity and Discord: Two Problems of Experiment***. Pittsburgh: University of Pittsburgh Press (2002), 288pp., $38.50 (cloth).

This book can perhaps best be characterized as an anthology of papers dealing with the two issues of selectivity and discord in experimental physics during the twentieth century. (Many of the chapters have been previously published as articles.) The first issue refers to selectivity in the choice of data or in the methods of their analysis, governed either by an experimenter's bias, hidden systematic errors or by the need to introduce some kind of cut-off in order to arrive at any data (illustrated here in chapter 1 on the measurement of the K^{+}_{e2} branching ratio, a mid-1960s high-energy physics experiment in which Franklin himself had taken part). This selectivity among other factors such as the use of different instruments, data analysis procedures, experimental skills, etc., often leads to discordant results between two or more experiments, in one case (chapter 9) even to two different results from the same experiment. Such discord inevitably poses the problem of how to decide and how to resolve the issue. Franklin's agenda briefly put is the following: "Although a consensus is usually achieved within a reasonable time, I believe that one must demonstrate that the methods by which such resolution is achieved proved grounds for scientific knowledge—in other word, that they are based on epistemological and methodological criteria" (239).

Franklin attains this goal by identifying various experimental strategies, foremost experimental checks and calibration, artifact reproduction to confirm the instrument's smooth functioning, elimination of plausible sources of error and alternative explanations of the obtained result, using characteristic features of the results themselves or an independent well-corroborated theory, a well-understood apparatus, or statistical arguments to argue their validity. Aside from these seven strategies (commented on 3ff.) he also mentions elsewhere independent confirmation and blind analysis (34 and 132ff.)—familiar to us from experimental psychology as effectively eliminating experimenter bias. This is discussed further in one of the few chapters containing material the reviewer hadn't already seen in one of Franklin's earlier publications. Neither successful replication nor any of the other strategies, taken in isolation, are a guarantee for the correctness of an experimental result: it is their combination in scientific practice that leads to the robustness of experimental results, often even over deep changes in theoretical guiding assumptions. The strength of the

Philosophy of Science, 71 (October 2004) pp. 607–610. 0031-8248/2004/7104-0013$10.00

book is its discussion of all of these experimental strategies in ten historical case studies, ranging from the absorption of β-rays to atomic parity violation or low-mass electron positron states (which after a decade of experimentation were ultimately refuted as an artifact of the cut-off procedures used).

Rather than superficially glancing at all the examples, many of them highly technical and probably difficult for anyone without a good knowledge of 20th-century physics to understand, let me discuss some issues in more depth, picking one of Franklin's oldest (1981), easiest, and best-known examples: Millikan's oil-drop experiment, which confirmed the quantization of charge. Franklin disagrees with other historians of physics who had analyzed Millikan's notebooks in the archives at Caltech, such as Gerald Holton and Daniel Siegel, who had concluded that Millikan went about "choosing data according to his presuppositions, and then using those data to support his presuppositions" (Dan Siegel, quoted on 256). Against what seems to be a classic example of the experimenters' regress, Franklin argues that in the overwhelming majority of such omissions, there were intrinsic reasons for suppressing some of the data taken because of identified disturbing factors such as changes in temperature, fluctuations in the voltage, or the size of the drops being too large so as to require second-order correction to Stokes's law. Thus Millikan's touching up of 30 of the 58 published events in order to slightly improve his statistical uncertainty is dubbed "cosmetic surgery" (74), even though Franklin concedes that "the exclusion of drops for which [Millikan] calculated a value of *e* and could thus select the value he wanted as well as his choice of calculational method are not justified." But Franklin's detailed statistical reanalysis of Millikan's data shows that his omissions only very slightly reduced the statistical uncertainty of his final results but did not have any significant effect on the final value of *e* (257, fn. 14).

Given all this, I am at a loss to understand how Franklin can uphold the claim that "there is, however, no evidence that the public and private arguments are different" (246). Such exaggerated claims don't help Franklin's case in convincing constructivists: He should rather have said that science remains a rational enterprise *despite* these occasional differences between public and private data, between what is published and what is entered in the lab notebook. Millikan's case shows more than that: There is no way around admitting that Millikan transgressed the boundaries of proper conduct by lying outright to his readers by stating in 1913 that "this is not a selected group of drops but represents all of the drops experimented upon during 60 consecutive days" (72; Franklin's statement "This is not correct" sounds unnecessarily euphemistic to me). There can also be no doubt that ultimately social reasons were responsible in particular for the silent omission of one measurement orginally singled out as particulary good and reliable: Millikan's impulse not to play into the hands of Ehrenhaft by publishing a single value which his opponent could have quoted as supporting evidence for his case, i.e., the existence of fractional charges.[1] But *despite* this violation of the code by tuning the experimental cuts and concealing this selectivity in the presentation of his findings, Millikan's claims on charge quantization were *de facto* firmly based on far more data than he eventually published. He himself (and many of his successors who later repeated this experiment) implemented many of the experimental strategies listed above, so well-suited for safeguarding against fraud and self-deception.

The core of Franklin's claims as I understand them is not that misconduct doesn't happen in science, but that it will be discovered sooner or later because of the amazing efficiency of these experimental strategies. Any local experimenters' regress as might be found and temporarily stabilized (be it in Weber's gravitational wave detectors, or during the search for the 5th force—here covered in chapters 2 and 7 respectively) will ultimately be broken up and corrected. It is not a global and indefinite vicious circle only decidable by social pressure—as Harry Collins would have us believe—but a local and temporary one. Contingency, clever rhetorics, concealed evidence, and self-deception only have a short lifespan in science: that's the message, not the mistaken claim that these issues are totally irrelevant. Yes, "there is no instant rationality in science. Problems of selectivity and discord may take some time to resolve" (247). Any philosopher of science wishing to understand how modern science achieves this amazing degree of in-built self-correction and resolving capacity in cases of discord will have to start from the experimental strategies described so clearly in Franklin's case studies. What is missing is a more structural account of their epistemological effect, and the way they are intertwined with each other, as is beautifully exemplified in many of these case studies, but somehow buried in the details.

Philosophers of science may be most interested in the introduction where Franklin also situates his own 'conjectural realism' as compared with Collins, Galison, Hacking, Pickering, and a few others (none from outside the US and the UK, though). Newcomers to the field will welcome

1. Actually, the omitted drop of 16 April 1912, which is "among Millikan's most consistent measurements" and initially commented upon by him with "Publish. Fine for showing two methods of getting *v*," leads to 0.6*e* as we learn from Franklin's highly interesting footnote 12 on 256f., which really belongs in the main text along with further discussion. In general, such historical material apparently playing into the hands of the constructivists must be addressed with particular care up front rather than be tucked away in the endnotes.

the book as a good synopsis of his central claims, focusing on the two key issues of selectivity and discord which crop up time and again in the debates between realists and/or 'rationalists' vs. constructivists. But I doubt whether the latter will be induced to change their opinion on the basis of this book, mostly because Franklin does not make an effort to 'translate' his findings into their language, and occasionally makes unnecessarily overdrawn claims which will breed distrust in his other—in my opinion quite justified—central theses.

KLAUS HENTSCHEL, UNIVERSITÄT BERN