CHAPTER 1

Introduction

> If two scientists disagree on any issue, and the issue is within the ambit of science, then it must be possible for them to agree on a procedure which they can both accept as a critical test of their points of difference. For reasons of personality they may not be able to get together to work out such a test procedure, but it must exist as a possibility. If such a critical test cannot be imagined as possible, then the issue between them is not a scientific issue. The scientific method does not vary with the subject-matter, but is the same irrespective of its results and basically the same in all the sciences.
>
> L.T. Wilkins: Social Deviance, page 4.

In reviewing the final report on Cranfield I, N.D. Stevens (Ref. 1) described it as being 'extremely complex'; even after 'several careful readings' he found parts of it 'still bewildering', and said that 'there are so many side issues that the author neglects the clear and detailed presentation of the main headings; the reader finds himself sidetracked by these, or other interesting diversions'. Since this reviewer was by no means the only person who made such comments, it has been our particular endeavour in this report to make quite clear what has been done, how it has been done and what has been the outcome, even though at times this has lead to what some people may consider undue verbosity and repetition.

In one respect, this project is easier to report, for, being in a more concentrated field, it does not raise many of the side-issues - such as indexing times, indexer qualifications, etc. - which came up in Cranfield I and which were sufficiently interesting to sidetrack the reader. On the other hand, to those who have been involved in Cranfield II, the earlier project seems to have been child's play to what has now been attempted, and the complexity of the present work is inevitably reflected in what has to be reported. To those readers who, like ourselves, tend to view with dismay the many papers on information retrieval which consist substantially of some twenty pages of mathematics, we can only apologise that it has become necessary to introduce a number of equations into this volume. However, it is certain that there is no mathematics

which should be beyond the comprehension of a schoolboy of average intelligence, but, even so, it is suggested that Chapter 3 might be skipped by those who are not closely concerned with the particular problem of performance measurement. Important though the work in this Chapter is felt to be, yet the arguments may well be of marginal interest to many readers. At the beginning of Chapter 4, which presents the main set of test results, full information is given concerning the performance measures which are actually used; Chapter 3 explains in some detail why those measures were selected in preference to other possible measures.

To a lesser extent, the same is true of Chapter 2 which discusses at length the variables which were being investigated and the environment in which the test was carried out. Again, we have tried to make Chapter 4 complete in itself in that such matters are briefly recorded therein. Only if the reader is puzzled as to why such seemingly unnecessarily tortuous actions have been taken, need he refer to Chapter 2 to find the possible justification.

The test results presented in Chapter 4 make up the main bulk of the report. Some may cavil at the way in which, at the slightest provocation, we include plots of the results. Undoubtedly these add to the bulk, but we can only hope that they will allow readers more quickly to get a general idea of what has been happening. The following chapter presents substantially the same set of results in a simpler but probably more controversial manner. In Chapter 6, extracts have been taken from the main test results and presented in such a way as to illustrate different aspects of the investigation.

Subsidiary to the main test was an attempt to make a comparative evaluation of citation indexing and bibliographic coupling. While there should be no serious problems in making such an evaluation under operational conditions, the value of testing this form of index in an artificial environment appears dubious. However, with considerable reservations the results are given in Chapter 7.

Up to this stage the results have been presented without any attempt being made to draw conclusions. All such have been relegated to the final chapter of this volume, in which an attempt is also made to relate the results to other investigations in this field.

There is one general apology that should be made and that is for the introduction into this report of yet more jargon. Many terms first used in reporting Cranfield I now appear to have gained general acceptance, but it is unlikely that such phrases as 'maximum starting term coordination level method' or 'proportional coordination level method' will crop up very frequently in the literature - and we certainly hope they won't - but it has been necessary to find terms to describe certain procedures so that, in reports of other tests, one has a chance of knowing which of several

possible methods has been used. We can only plead that we have not - as some apparently delight to do - concocted new terms to describe measures or methods when existing terminology has already appeared in the literature. If we have offended in this way, it was unintentional and we hope that our attention will be drawn to any lapse. The only case of which we know is where a term has been changed from that used in Volume I. The term 'generality ratio' has been dropped in favour of 'generality number'. It is hoped that the argument in Chapter 3 will provide the reasons for this change.