Symposium on Selective Indexing

While it cannot be denied that a book index should be comprehensive if it is to meet fully the needs of all who may consult it, many indexers must have been faced with the task of compiling a selective index—putting a quart into a pint pot—because not enough space had been allowed by the publisher or for some other reason. The problem then arises as to how to decide which items to leave out, or what kind of entry to omit, or, preferably, how to telescope the content of the entries so as to include as many as possible in the given space.

It has been found that if omissions from indexes are systematic, being confined to one or two types of reference, retrieval can be obtained through subject entries. For example, in one case all personal names (which are legion) mentioned incidentally and once only were omitted. These could be retrieved by using a dual set of subject entries which provided a crossbeam. The nature of the material to be indexed dictates the selection and the method, more than do general principles.

The following example is supplied by Mrs. A. M. Quinn, who has suggested the subject of the symposium and the way contributors might deal with it. It is her latest attempt to make a selective index and is chosen because selection was not rigorous and the method also self-evident; it is therefore easier to follow than one involving more than one category of omission. The decision to be selective was her own. The book is a large one, the text (without the illustrations) running to 714pp. This selection was decided upon to ensure that there would be space for the valuable and professional bibliography which had been compiled. The bibliography occupies 20pp. and the index 25.


INDEX

For the voyages of discovery the activities of Cook, ships' officers, masters and professional supernumeraries are not indexed in detail. They are covered by the movements and visits of the ships in which they served, under the following abbreviations:

- (1768-71) *Endeavour* End.
- (1776-80) *Resolution and Discovery* Res. (2) Res.: Dis.

*Adv.*, *End.*, *Dis.*, *Res.* (1) and *Res.* (2) are used with names of persons to denote voyage and ship; with place names they indicate where the ship is sailing alone. *Res.*: *Adv.*, and *Res.*: *Dis.* are used when the ships are sailing in company.

Names of ships are under the heading Ships; the heading *Ships (Cook's)* lists the ships in which Cook served and includes the *Adventure* and *Discovery*.

Examples of entries:

Clerke, Charles, master's mate *End.*, 138-9, 141, 269, 271; lieut. *Res.* (1), 281, 294, 297-8; commander, 444, *Dis.*, 495-6, 506; *Res.* (2), 675; illness, 568-9, 679; death, 682; burial, 683

Queen Charlotte Sound, N.Z., 279, 319, 512; longitude, 336, 423


*see also* Ship Cove

The Hon. Editor asked a number of members to contribute to the symposium. The following contributions have been received.
from M. D. Anderson

All indexing involves a process of selection. For the purposes of this symposium, selective indexing may be taken to mean concentrating the entries as far as possible while maintaining effective coverage. In an article in Vol. 7 of The Indexer (pp. 121-2) on how to make an index to a specified length, I suggested the method of over-indexing to some extent, followed by pruning and condensation, since it is easier to decide after working through a book which entries of lesser importance can best be spared, and how entries may best be combined. Selective indexing is thus achieved by a double process of selection.

I recently carried out a rather extensive operation of this kind, and subjoin the details. I had to index a book of 157 pages, called Social sense and practice. The publisher asked for either a 2-column index with 300 lines of 40 characters, or preferably a 3-column index with 475 lines of 25 characters. I set out to make the latter, and collected about 450 lines, but so many of them were longer than 25 characters that the 2-column index was obviously preferable, and a third of the lines had to go.

I began the shortening process by throwing out several classes of entry. The book had a section on how to eat special types of food, with a little drawing for each, and I had accordingly indexed "asparagus, how to eat", "corn-on-the-cob, how to eat", etc. I decided that a sub-heading under "foods", "special, how to eat", could cover the lot. Again, there was a section on how to address peers of the realm, with directions for each rank. I omitted "earls, how to address", "viscounts, how to address", etc., considering that they were covered by "peers, how to address", but I did leave "dukes and duchesses, how to address", and also "knights, how to address", for knights are not peers, and also this entry had another sub-heading, "precedence of different orders of", that I wanted to preserve. Other sections on how to address dignitaries of the churches, the law, the universities, and civic life, were similarly treated; entries for the Pope, and for Archbishops and Bishops, were retained. Advice on how to behave in different European countries was originally indexed under the countries named, but these were replaced by adapting the section heading from the book to give "Continent, manners and customs on", as a collective entry. Entries for "pages at weddings", "photographers at weddings", etc., were thought to be covered by "weddings . . . arrangements for". For all the above groups, it appeared that the reader ought to be able to recover the desired item from the collective entry; the relevant length of text was 3 pages for the foods, 4 for the peers, and 2 for the continental countries.

Another kind of shortening was achieved by making do with single entries instead of double or triple ones. For example, "passports" was left, but additional entries for "British Visitor's Passport" and "United Kingdom Passport" were deleted. "General Register Offices" was removed to leave only "Register Offices, General".

And finally, lines were saved by abbreviating individual entries. Thus "architects, formalities in dealing with", was reduced to just "architects", and "cremation", which had had sub-headings "expression of wish for", and "formalities for", was also reduced to the single word.

By these methods, I was able to produce an index of the length required, and I hoped that the result would prove to give reasonable coverage of the text. If I had originally been making the shorter index, I should have used from the start many of the expedients described above, and should probably have been able to estimate the length so that not more than 20 or 30 lines had finally to be cut. This method of double selection does seem to involve unnecessary work, but with a fairly short and simple book the time involved is surprisingly little; the total time spent on the above index was 10 hours. It would not be a sensible way to tackle the index for a longer and more serious book, such as that described by Mrs. Quinn, for which preliminary planning would obviously be essential.

One method of lightening an index load is that adopted by the indexer of Steven Runciman's book The fall of Constantinople, 1453 (Cambridge University Press, 1965). At the head of the index is this note:

Names that recur continually in the text, such as Constantinople, Greeks, Byzantines, Turks, Ottoman, Italians, are not listed.

The number of lines in the index is about seven per cent of the number of lines in the book, i.e., about the expected length for a book of this kind. The only long entries are those for the Emperor Constantine XI (11 lines), the Sultan Mehmet II (21 lines), and Venice (10 lines), and in addition (in spite of the preliminary note), a very long entry for Constantinople (2 columns),
subdivided into Churches and monasteries, Fortifications and gates, Harbours, Public buildings and monuments, and Quarters (Byzantine and modern).

"Names that recur continually in the text" are those that give indexers the most trouble. We have been advised by G. V. Carey (No Room at the Top, The Indexer, Vol. 2, p. 120) not to make a long elaborate entry for the subject of a biography, but to distribute most of the references appropriately among other entries. In the index to Runciman's book, there is an extension of this principle, and it would seem worth while, when planning an index, to consider if it could usefully be applied.

from Roger F. Pemberton

Selective indexing involves the exclusion of some details so that others, presumably more important, may be included. The first difficulty is in deciding which are the more important, for the indexer cannot know who may be using the book or what the user may wish to find by means of the index. The indexer must therefore rely on his imagination, his knowledge of the subject and his experience, when making his decisions on what to include and what to exclude.

The next problem is to condense what has to be indexed into as small a space as possible; so that the maximum guidance may be given in the minimum space, thereby ensuring the greatest possible extent of retrieval. Obviously the maximum coverage in any given space can be assured by using 'blind references' only. (By 'blind references' I mean main headings and page numbers only, without distinguishing subheadings). This certainly extends the coverage greatly, but it also gives the user a lot of trouble. Yet it may, in some circumstances, be desirable if the probable readership is extensive and varied; and it can be argued that an indexer should not be expected to do all the user's work, particularly if the user is on research.

Since, however, indexers are generally expected to provide, by means of subheadings, some indication of what the more important references are about (the less important being either 'blind references' immediately after the main heading, or grouped under the subheading 'other mentions' at the end of the entry, or omitted altogether), some consistent method of selection, condensation and grouping must be adopted. A brief Note, simply and clearly explaining the method adopted and any codes used, must be put at the head of the index, for the guidance of users. This note must be really brief, for if it is not brief users will not read it. Mrs. Quinn's illustration of such a note and code is a good example of what is required.

Apart from such ingenious and necessary devices, much condensation may be achieved: (a) by assembling similar references, such as 'foreign relations' or 'foreign policy', under one of those subheadings or a similar subheading; and (b) by combining convenient subheadings which have the same or adjacent page numbers. Two examples are given below. They are from the Cambridge ancient history, Volume II, Part 2 (not yet published in its revised form). In the first example, the main series page references are given in bold type immediately after the main heading, for convenience; and these are followed by detailed references for each country and people named.

Example (a)
The entry in full

Egypt, external relations, 81-6, 218-21, 226-30, 232-4, 241-4; during XXI Dynasty, 656; with African dependencies, 85-6; Alashiya, 202-3, 204; Amqa, 84; Amurru, 85, 140, 261; 'Asiatics', 82, 84, 85, 519; Asy, 203, 204; Babylonia, 23-6, 29, 81, 283; Cyprus, 197-9 passim, 201 (see also Alashiya); Damascus, 5 (see also Syria); Eastern Mediterranean area, 198; Edom, 656; Ethiopia (after Piankh, q.v.), 657; Great Powers in Near East, 81; Greece, Mycenaean, 201; Hittites (Khatti), 7-9 passim, 17, 18 passim, 49, 81, 83, 85, 120, 121, 125, 128, 129, 140-1 passim, 221, 226-9, 234, 252, 257-8, 258-9, 265, 277 (treaties, 9, 141, 182, 229, 234, 236, 258-9, 260, 279); Hurrians, 10; Israel (under David and Solomon), 320, 656 bis; Kinza, 8, 17; the Lebanon, 656; the Levant, (generally), 242-3; Libya q.v., 230, 233-4, 242-3, 339; Mitanni, 5, 69, 81-3 passim; the 'Nine Nations', 82; Nubia, 224, 230, 244, 643; Nukhash, 84; Palestine, 5, 82, 85, 98-116, 218-20, 221, 228, 234, 235 bis, 588, 656; 'Peoples of the Sea', q.v., 147 (see also Ramesses III); Qadesh, 84, 140, 220, 221; Sumura, 12; Syria, 1-5 passim, 8, 10, 12, 16, 19, 21, 81, 83, 84, 85, 99, 123, 128, 198, 220, 221 bis, 252, 259, 642-3, 656; Troy, 163; Ugarit, 4, 10, 85, 133-4, 137-41 passim, 144, 147.
First alternative

Egypt, external relations, 81-6, 218-21, 226-30, 232-4, 241-4, 656; see also: African dependencies, Alashiya, Amqa, Amurru, 'Asiatics', Assy, Babylonia, Cyprus, Damascus, Eastern Mediterranean area, Edom, Ethiopia, Great Powers in Near East, Greece (Myceanaean), Hittites, Hurrians, Israel, Kinza, the Lebanon, the Levant, Libya, Mitanni, the 'Nine Nations', Nubia, Nukhash, Palestine, 'Peoples of the Sea', Qadesh, Sumura, Syria, Troy and Ugarit.

Second alternative

Egypt, external 81-6, 218-21, 226-30, 232-4, 241-4, 656; see also under names of countries and peoples concerned.

N.B. If either of the two above alternatives relations with Egypt are indexed in the entries for the countries and peoples named.

Example (b)

The entry in full

Tutankhamun of Egypt (died aet xix): his reign, 66-71; his tomb-filling, 64; artefacts from El Amarna, 64-5; his canopied coffinette, 65; his burial furniture, 65, 654; his age at accession, and marriage, 65; his residence 65-6 & n(8); wine-jars from his tomb, 66; his burial place, findings, 66, 69, 70; monuments to his reign (few), 66; his efforts to restore prosperity to Egypt, 66-7; restores worship of Amun, 67; and monuments to Amenophis III, 67, 73; prepares his tomb, 67; his Restoration Stela at Karnak, 66, 73, 75, 84; the last male heir of XVIII Dynasty, 69; succeeded by Vizier Ay, 69 bis; his sarcophagus, 71, 95; his Deputy, 71; and Horemheb 72-3; his tomb robbed, 75; lawlessness during his reign, 75; his relationship to ruling house (of Egypt), 79-80; his parentage and birth-date, 80; and Syria, 84 bis; his District Commissioners, 84 and Central Palestine, 85; visited by African prince and princess, 86; his tomb-paintings, 97; applied arts under, 97; his golden chair, 597; his death (childless) and burial (c. 1352 B.C.), 69, 70, 73, 85, 217; his widow and Shuppiluliumash of Khatti (messages quoted), 18 bis, 69, 84.

Alternative entry (condensed)

Tutankhamun of Egypt (died aet xix): his reign, 66-71; artefacts from El-Amarna 64-5; his tomb and its contents, 64-71 passim, 95, 97, 597, 654 (robbed, 75); accession, marriage, and residence, 65-6 & n(8); monuments to and by, 66, 67, 73; internal affairs (in Egypt), 66-7, 75, 97; Restoration Stela, 66, 73, 75, 84; relationship to XVIII Dynasty, 69, 79-80; his successor, 69 bis; his Deputy, 71; and Horemheb, 72-3; external affairs, 64 bis, 85, 86; his District Commissioners, 84; his death and burial, 69, 70, 73, 85, 217; his widow and the Hittite king, 18 bis, 69, 84.

When some entries, presumably less important but possibly wanted by some users, are 'lost' (i.e. not directly traceable in the index), they may be retrieved by some indirect method, such as through subject entries or entries under the names of persons or places. In these cases, some hint as to how the user may find them should be put in the Note (above mentioned) at the head of the index, e.g. 'economic and political details may be found under the names of the countries concerned'.

from Robert Collison

The choice of this subject will remind some of the older librarians of a shortlived experiment in the 'thirties in the field of "selective cataloguing". The idea was received with great enthusiasm and much thought was given to the problem of eliminating "unnecessary" catalogue entries and references without harming the services required of the catalogue. The topic, however, died out as rapidly as it had been introduced, for it was soon clear that in any selective system some harm is bound to occur. Nevertheless, any efforts toward selective indexing are well worth pursuing, for we are all aware that there are many entries in indexes that are never or rarely used and therefore do not justify the space they occupy.

The Indexer Vol. 9 No. 2 October 1974
My own belief is that no selective system should be introduced until all the typographical possibilities have been exhausted. Thus, in the great majority of books the index is printed in two columns. In most cases this results in a mass of white empty space, and I suggest that a three-column index—with some important adjustments in layout—should be considered. An analysis of the average index will show that the great majority of entries will fit very well into a three-column index, but certain lengthy entries will fare badly. Here I think we should take a leaf out of the indexes in some American encyclopaedias and works of reference where outstanding lengthy entries are printed in a two-column-width box at the correct position in the alphabetical sequence, thus giving prominence to so important an entry and allowing enough space to give it special layout attention. The current evolution of this method can be seen in many telephone directories; for example, in Britain it will be noticed that long lists of telephone numbers relating to various public utilities are often featured separately—at the appropriate alphabetical position—in the white pages, while the yellow pages (admittedly in the form of an advertisement) follow a similar system where a number of telephone numbers need to be differentiated. The possibilities of this method are easily appreciated, and indexers—with their knowledge of effective typography and layout—are in a good position to improve on these ways of conveying important information economically.

Unfortunately far too little notice has so far been taken of the typographical ingenuity that, in isolated cases, has been applied to indexing problems, and it is especially disappointing that the new edition of the Britannica—so revolutionary in other respects—did not break new ground in this area. America is not however the only source of new ideas in indexing: an experiment that passed almost unnoticed was carried out in Mr. D. H. Borchardt’s *How to find out in philosophy and psychology* (London: Pergamon Press, 1968). As can be seen from the illustration, the layout of the index has been very carefully designed to cope with a special problem—in this

INDEX

Numbers on the left of the subject term refer to the running number of the LIST OF WORKS REFERRED TO; page references to the text are given on the right.

Aesthetics 58
Analytical philosophy 36
Ancient philosophy 26
Aristotle 67
Associations see Societies and Associations
Augustine, St. 67

Bibliographic Index 3
Bibliography of philosophical bibliography 32
Bibliography of psychological bibliography 36
Biobibliography 15-21, 27, 63-70, 72

British philosophy 68
Broad, C. D. 68
Buddhism 52

Carnap, R. 68
Cassirer, E. 68
Child psychology 30, 62
Chinese philosophy 26, 44-46
Chronology 19

Descartes, R. 66
Dewey, J. 68
Dialectical materialism 59-60
Dictionaries and Encyclopaedias (Philosophy) 4-12, 14-15, 56
Dictionaries and Encyclopaedias (Psychology) 12-14, 15
case, the necessity for indexing both the subject-matter of the book and a large number of bibliographical references to the same subjects. This has been very skilfully solved by Mr. Borchardt—and it also gives us new ideas of how such a layout can be adapted to different kinds of requirements.

In the case of selective indexing, the Borchardt layout could be used to good effect. Imagine that the column on the right occupies only a third of the width of the page: then we have two-thirds in which to display lengthy entries to their full extent. It will be noticed that by limiting the right-hand column to a third width, only eight entries need more than one line—and four of these will probably be the ones that are candidates for the more extensive “display” treatment proposed. Before we go any further with selective indexing that eliminates entries, I suggest we try to get some good typographers interested in what can be done to get more entries into the limited space available to us.

from F. T. Dunn

Every good index, of course, must be selective—in that it should include only those entries likely to be helpful to the reader. There is no point in slavishly pursuing comprehensiveness if this means including every single reference regardless of its importance or relevance; and this I would say is the distinction between cataloguing—that is, a fully factual listing of contents—and indexing, which is more in the nature of a guide and signpost. To waste time looking up a reference which turns out to be a mere incidental mention (where, for instance, a sentence like ‘This was the age of Bennett, Shaw and Wells’ can yield three useless references) can be more exasperating than finding no reference at all.

Selection then, like ripeness, is all; but most of us know cases where limited space forces us to engineer an even more selective index than usual. In such cases it seems useful to omit certain types of entry altogether and to treat the remaining entries in some detail, rather than provide a fuller range of subjects more scantily treated with faceless strings of undifferentiated page references. Obviously the nature of the text will indicate the categories to be selected. In a military biography, for instance, the battles, campaigns and commanders demand detailed treatment, and other personal and place names could well be omitted. In a philosophical or technical study, on the other hand, it may be necessary to concentrate on concepts or processes to the complete exclusion of personal names—in which case the limited nature of the index can be clearly indicated by calling it a ‘Subject Index’.

There is great scope for selection in that feature usually considered obligatory in indexes to biographies—the individual entry for the main subject. Too often this becomes a wasteful duplication of entries already in the index, whereas it should be confined to personal attributes, abstract items or events in the subject’s career for which it is difficult to find obvious keywords of their own—such as ancestry, childhood, ill-health, first stage appearance, marriage, financial worries and so on. The fact that Our Hero attended Summerfields, Eton and Oxford will have yielded at least three entries in the main index, all of which can be telescoped into one reference to ‘education’ in the individual entry.

These are straightforward cases, but at the moment I am engaged on a project which seems to blur the earlier mentioned distinction between cataloguing and indexing. This is a cumulative index to the seventy volumes (1905-74) of The Dickensian, the magazine of the Dickens Fellowship. Ideally it should list everything—the thousands of references to the individual works and characters, the countless members of the various Fellowship branches; in practice some degree of selection becomes essential. So many of the references are incidental, so many people are mentioned only as attending a branch meeting (and not even speaking), so many of the references to Mr. Pickwick merely repeat that he was a wonderful creation, that I cannot believe a user of the index will thank me for leading him to them through seventy volumes. Therefore the decision was taken to index works, characters and Fellowship members only if the reference is of some value—in the knowledge that readers’ ideas of value may well differ from mine. In the same way, advance notices of branch meetings are considered to be by now useless pieces of information and therefore omitted, but subsequent reports of those meetings are of course indexed. Letters are indexed if from important people, or from unimportant people if on a subject of interest or concern,

64 The Indexer Vol. 9 No. 2 October 1974
and also if giving information in reply to readers' queries. And perhaps this is the most dangerous criterion of all—to decide who is important and who is not (or who, unimportant when he wrote his letter, later became important). How much easier indexing would be if judgement and selection were not required, if one could accept the Gradgrind dictum that facts are alone wanted—if, indeed, one were content to catalogue and not to guide.

from Philip Wright

The indexer faced with the need for compression can perhaps take solace from the thought that all indexes are, presumably, selective in some degree.

I am unable to propound any scheme for the retrieval of omitted items. As to what types of entry can be dispensed with, this must surely depend on the nature and function of the work being indexed. For instance, recently in making a cumulative index for 20 volumes of the Transactions of the Jewish Historical Society of England, I assumed that items not directly affecting Jews, individually or collectively, could be omitted, e.g. when one of several Bristol magnates involved in denouncing a certain Jew was described as being interested in the colonization of Newfoundland, I omitted Newfoundland.

A type of reference which might often appear to be a strong candidate for omission is the indexing of authors and titles of works cited simply as authorities for statements in the text. To revert to the Jewish Historical index, the work contained numerous citations of the newspaper, the Jewish Chronicle. As I could not conceive what use anyone could make of a complete string of such references, I indexed only those references which gave information about the history and editorial policy of the paper. The remainder could be 'retrieved', if need be, through the relevant subject entry in the index.

Apart from the omission of entries from the index, some economies can obviously be obtained in the presentation of the entries which do appear (as in the example provided by Mrs. Quinn). The well-known commandment against having more than three undifferentiated page references after a heading or sub-heading is a useful ideal, but need it be regarded as inviolate? There may even be cases where a reader would actually find a straight run of half-a-dozen or more page references after a name more serviceable than a re-arrangement under sub-headings. I would also like to put in a word for the old-fashioned (?) style of reference such as '73f'; in some cases, just as adequate as its fussy alternative, '73, 74, 76, 78-9, 81, 82'.

Readers are invited to submit further contributions (with examples if possible) or correspondence relating to this Symposium.

ISSN 0019 - 4131

Readers will have noticed the above on the cover of this issue, and some may have wondered what it means.

It is an International Standard Serial Number and is allocated by the International Centre (IC) of the International Serials Data System (ISDS) which is located in Paris, and was established by Unesco in 1973. Numbers have been allocated to all existing serials, and new serials will be given numbers by the IC unless a national centre of ISDS has been set up. This has happened in many countries. For example, in Australia it is at the National Library of Australia, Canberra; in Canada at the National Library, Ottawa; in the U.K. at the Bibliographical Services Division of the British Library; and in the U.S.A. at the Library of Congress.

The purpose of the number is to communicate briefly, and with a minimum of error, the identity of a particular serial; it may be used by subscription agencies and publishers for such processes as ordering, invoicing and inventory control, and authors may use it for copyright purposes.

Indexing Techniques

In our next issue we hope to have a review of Title derivative indexing techniques: a comparative study by Dr. Hilda Feinberg. It deals with the effectiveness of various techniques used for permuted title indexing.