The making of an index

F. Howard Collins

Edited by Michael Robertson

Introduction

Michael Robertson

Frederick Howard Collins (1857-1910), the author of the Authors' and printers' dictionary, the standard guide to British editorial practice (now published as the Oxford dictionary for writers and editors), was also the author of two early texts about indexing, which are presented here. The first is a chapter on 'The Making of an Index', from the second edition of Charles T. Jacobi's Some notes on books and printing: a guide for authors and others (1902); and the second is a paper on Subject indexes: some principles which underlie them, prepared for the Royal Society in 1896 for participants in a conference on an 'International Catalogue of Science'.

The first of these texts is exactly contemporary with Henry Wheatley's indexing classic, How to make an index,2 and the second precedes it by six years. (Although Wheatley mentions several late nineteenth-century writers of articles about indexing, Collins's name does not appear in his book.) Collins's indexing guide gives us a glimpse into the indexer's study of a century ago. Prior to the manufacturing of ready-made index cards that could be arranged in boxes, the practice of cutting and pasting paper entry slips for indexes, described by Jacobi, Collins, and Wheatley, seems, on their evidence, to have been general. Today's computerized professional indexers may gain some idea of the level of technological difficulty involved from Wheatley's advice:

There will always be anxiety to the indexer while his work is being cut up and sorted. A breeze from a window when a door is opened may blow some of his slips away... A good cook will make good paste, but if you are especially particular you can make it yourself... If the paste is made at home, it need not be lumpy; and lumps, when you are pasting, are irritating to the last degree.3

By contrast, the second of the texts presented here, Subject indexes: some principles which underlie them, gives an insight into the theoretical difficulties and controversies which indexers of the time faced.

Collins wrote in a variety of fields, including such items as an article in the Fortnightly Review (1888) on 'The ways of orthodox critics': a pamphlet on The diminution of the jaw in the civilized races, an effect of disease (1891); and charts of the tidal streams of the North Sea (1894), the west coast of Scotland (1894), and the Channel Islands (1897). In addition to the Authors' and printers' dictionary, his other major success as an author was a 600-page Epitome of the synthetic philosophy of Herbert Spencer (1889), with a preface by Spencer himself, which went through five editions and was translated into French, Hungarian, and Russian. In the wake of a recent revival of interest in Spencer's philosophy, the fifth edition of Collins's Epitome (1901) received the accolade of being reprinted by University Microfilms in 1975; the Dictionary of national biography commented on the book that, although it was an 'excellent' one-volume summary, its 'formality and necessary brevity... render it unsuitable for reading and its chief use is as an elaborate index to the Philosophy'.4 As Collins himself mentions, in the preface to the second of the papers presented here, it was largely through his work on Spencer's philosophy that he gained his experience in indexing: his conclusions, he states, 'are the result of compiling various scientific subject indexes during the past twelve years—including all those to Mr. Herbert Spencer's "System of Synthetic Philosophy" and his other works—and, when compiled, having to use them'. His debt to Spencer is also acknowledged in the preface to the first edition of the Authors' and printers' dictionary: 'One of the chief inducements which led me to undertake [this book] was the encouragement of my late dear friend Herbert Spencer; and, when I had commenced it, his willing help with its difficulties, while his health lasted.5

Jacobi (1853-1933) had published the first edition of Some notes on books and printing in 1892. It was a revised reprint, in an edition of 1000 copies, of his 1891 booklet On the making and issuing of books, of which only 50 copies were printed, and it contained the following two-paragraph introduction to indexing, by Jacobi himself:

The best and quickest way to make an index is to write each item on a separate slip of paper. Each slip should contain the head to be indexed, with the page reference attached to it. Assuming you have chosen what subjects to index, for instance, say, all names of persons and places, let every one of these be written out as often as they...
occur in the text, commencing with the first page and taking the whole in sequence. As they are written, throw them into a box or basket, and, when finished, gather them together and place all in alphabetical order. The next step is to eliminate all duplicate headings, but before doing this the page reference of the one thrown out must be transferred to the slip retained, and in its numerical order. This considerably reduces the bulk of the index.

This plan is the only royal road to making a correct index, without the chance of duplication or omission. Care must be taken a second time in checking the strictly alphabetical arrangement of the slips. When you have assured yourself of this, they may be pasted up in sheet form; this reduces the risk of losing any of the slips, and is a more convenient form for handling.

Perhaps on the strength of the paper Collins had presented to the Royal Society in 1896 (the second of the texts presented here), Jacobi invited him to contribute a new introduction to indexing for the second, enlarged, edition of Some notes on books and printing, published in 1902 in an edition of 500 copies. Jacobi acknowledges Collins in his Preface: 'I am greatly indebted to Mr. F. Howard Collins for the chapter of the Index, and for many kind suggestions and much help in the literary portion of the work.' A third edition of the book appeared in 1903, and a fourth edition in 1912.

The text of the chapter on indexing in the fourth edition contains some minor revisions and additions, which may have been introduced by Jacobi. The text presented below uses Collins's 1902 text as its base, but includes matter added in the 1912 text within angle brackets, and matter deleted in the 1912 edition within square brackets. To read the 1902 edition, therefore, the reader should ignore everything within angle brackets and include everything in square brackets; to read the 1912 edition, the reader should, conversely, include everything within angle brackets and ignore everything in square brackets. Two minor alterations from italics into roman in the later edition are silently ignored; and the misprint 'genius' for 'genus' in the text of Subject indexes has been silently corrected. The word 'catchword' in both texts is spelt inconsistently, with or without a hyphen, and where necessary the hyphen has been removed to conform with the usage of the Authors' and printers' dictionary.

I

The <Making of an> Index (1902)

-Nearly- all works are -undoubtedly- the better for an Index, and its absence in those dealing with scientific matters is frequently noticed as a great fault by reviewers. The 'Athenaeum' [nearly always] gives credit to a volume which contains a good index, and -rightly- points out the omission should it be a work requiring one.

In making an index, it is not enough for the indexer merely to understand the subject of the book, but he must understand the wants of the reader: two quite distinct things. The art of finding the correct title—or, as it is called, the 'catchword'—under which to describe the thing indexed, must be combined with the knowledge that under that head, and no other, will the reader most probably look for it. Hence a specialist in the particular subject of which the book treats may not -always- compile so valuable an index as one who, while knowing less of the subject, has had more experience in the ways of readers, and the catchwords they use.

A <mere> [good] index <of proper names, i.e. persons and places, of course, is a simple affair, but a subject-index> must be exhaustive; <It> must include the various 'points' of a book; must gather under one heading the same subjects; must be concise, yet explicit; must contain the subjects of chapters, and the smaller divisions, such as sections, not merely single words as they occur; and must define each entry that occurs under the same title, so that the reader may turn at once to the desired page.

-[An index cannot be commenced until the sheets as finally printed are at hand, otherwise errors in pagination are sure to occur.] <If, however, the volume is one not calling for a very special subject-index, the author is the best one to make it, but failing this the printer's reader who has had charge of the work through the press is suggested. This indexing should be attacked as fast as each sheet is finally printed, for it will save that interval of time which is sometimes damaging in getting the work out to a certain date.> Supposing the sheets are at hand, the best way to commence is to have an ordinary ruled copy-book, and to make each entry as it is decided upon, writing the catchword somewhat more plainly than the subsequent definition, so as to catch the eye when turning back for reference, as is often necessary. [Between] <After> each entry leave one line blank for the subsequent cutting up and sorting which will be necessary. <An alternative is to write the references at once direct on to separate slips.
which can be sorted out afterwards as recommended later on.>

As each entry is made, underline the word indexed in the text, if it is there, or write it in the margin if not, so that when the index proof comes to be checked with the sheets, the eye may readily catch the sought-for word.

In front of the desk, and visible without much effort, should be kept the list of cross-references. This should be written up as each heading is decided upon, and selected from one or more possible synonyms. For instance, if ‘Printing’ is the title selected and entered on the cross-reference sheet, there will be no subsequent re-arranging required from say such headings as ‘Book,’ ‘Caxton,’ ‘Fount,’ ‘Typography,’ for these will all be referred on the cross-reference sheet to ‘Printing.’

Do not imagine that it is sufficient to select mere words as they occur in the text. Entries must be made for the contents of whole chapters, sections, and paragraphs for the indexer must give the author credit for wanting to say some special thing in each subdivision of his work, and must similarly credit the intellectual reader with—having once found that particular subject dealt with—a desire to refer to it again, although pagination and all else concerning it has vanished from his mind.

When all the sheets have been extracted, and all the entries made in the copy-book—if this method of entry be adopted—it then remains to separate the book into loose leaves, and to cut with scissors along all the intermediate blank lines; then to sort the slips thus freed into alphabetical order; and then all slips with the same title in page order, commencing with the lowest. When this is done, the slips should be pasted or gummed on sheets of paper in proper order, and for the press, putting in cross-references from the sheet where a record of them has been kept, properly punctuating, eliminating duplicate headings by transferring all page numbers, references in numerical order, and otherwise giving the compositor as little trouble as possible when it comes into his hands.

The punctuation of an index may seem a simple matter, but it may be such as to make or mar it for a bibliographic eye. To make the catchword or heading of each entry prominent, it may be printed in a special thick-faced fount, or even in italic or small capitals, or, if this be not desired—as it may involve extra expense—a good plan, which still throws up the heading a little, is to put a colon (:) after the first word. A comma should be put between the last word of an entry and the first figure of the page, and a semicolon between each entry under the same heading. Where there are several pages given to one reference, a comma after each.

When the index proof comes from the printer it must be checked with the pages of the book itself, for who has not known the irritation produced by an incorrect page number in an index? Many hours of labour will be saved in this dreary task if the word indexed has been—as recommended—underlined on the page itself at the time it was selected, and marginal notes made of anything not in the text.

<Usually the index is printed in double columns, but this is a detail best left to the printers who will arrange it according to the length of the titles and the average number of references to each item indexed.>

II

Subject Indexes: Some Principles Which Underlie Them (1896)

Preface

Some former communications of mine to the Catalogue Committee of the Royal Society lead me to think that the following notes may be of interest to some at least of those attending the Conference on ‘The International Catalogue of Science’. They are printed in their present form, as the Royal Society do not propose to have any papers read by those who are not delegates.

It is needless to say that the conclusions drawn are not those universally held. It is perhaps not so needless to say that they are the result of compiling various scientific subject indexes during the past twelve years—including all those to Mr. Herbert Spencer’s ‘System of Synthetic Philosophy’ and his other works—and, when compiled, having to use them.

My best thanks are due to Dr. Foster, Secretary R.S., for kindly undertaking to distribute this paper to the distinguished visitors at Burlington House.

F. Howard Collins.

Churchfield.

Edgbaston,

Birmingham.

Subject Indexes: Some Principles Which Underlie Them

The discussion which has been taking place with reference to the proposed Subject Index to the Royal
THE MAKING OF AN INDEX

Society Catalogue of Scientific Papers,' and still more recently with reference to the proposed 'International Catalogue of Science,' may cause an account of the general principles involved to be of interest, more especially as, so far as I am aware, they have not been previously stated.

In all subject indexing the first step is of course to select the most important word of the Title under which the subject matter of the paper appears to be best generalized; the selection, that is, of the primary 'catchword' or 'key-word'. Then, if, as the work proceeds, many papers come under this one heading, to sub-divide this first group by the selection of a second catchword, and so on, until only groups of moderate compass are left. For instance, taking the first few titles from Vol. VII. of the 'Royal Society Catalogue of Scientific Papers:'—'The development of fungus spores in the alimentary canal' would come under 'Fungus spores,' and, were there many of these headings, 'development' would form the second key-word, and so on. 'On the approximate solution of numerical equations,' 'Equations: numerical' with 'approximate solution' if necessary. 'Electrolysis of Itaconic Acid,' under 'Itaconic Acid,' if necessary. Electrolysis; a cross reference from Electrolysis also being given. 'The conditions which influence the amplitude of the parallactic displacement of the ophthalmoscopic Image' under 'Ophthalmoscope: image of,' etc., etc.

Whether the titles given by the various authors are those best suited to represent the subject matters of the papers, and if they are not, how better titles should be selected, is not necessary here to discuss, as the 'catchword' method to be adopted would be the same in either case. Having now obtained the material from which the subject index is to be compiled—the bricks with which the house is to be built—the next step to consider is in what order these catchwords shall be arranged. They may be put in simple alphabetical sequence, as are the words in all ordinary dictionaries and encyclopaedias: Fungus under F., Equations under E., and so on; or they may be classified under the sciences and their sub-divisions to which they belong. Fungus in this method would be placed under the general heading Botany; then Botany Cryptogamic, then Thallophyta, then Schizomycetes, and so on. Equations primarily under Mathematics, and so forth. The index generally. A feeling which is never dispelled, leading eventually to absolute distrust of the index. This brings us to:

2. That any and all classifications are altered by advancing knowledge. Going no further back than the last half century it is exemplified on all sides. In Botany, the classifications of the cryptogams; in Zoology, the varied positions that the Sponges and the Vermes have taken; in Chemistry, there is the fresh title, 'The New Chemistry:' in Physics, Kinematics, Kinetics, and ether waves, are headings for many matters which formerly would have been differently...
named; possibly the ‘X-rays’ may some years hence be treated of as the phenomenon dealt with by an ‘X-ology;’ the old term. Political Economy, seems to have given way to Economics; in Biology, we find a part of the science of heredity now termed Weismannism, and adaptation more or less included in Ecology. That the theory of harmony, counterpoint, fugue, and orchestration may become a branch of physics Helmholtz has shown, while Physiography tends to cloud the distinction between geography and geology. And the Aurora Borealis, from belonging to Meteorology, hovered for some time over electricity and magnetism, finally coming to anchor under the former. In Astronomy, we have the long journey of the nebulae to find a home which will be satisfactory to spectroscopic and the other Sciences, and ‘The whole domain of optics is now annexed to electricity.’

3. If, as we are bound to assume from the foregoing few instances, classifications alter with the progress of time, it becomes a mathematical certainty that any classified arrangement of scientific papers made at one period will, in the future, become obsolete; will become historically interesting as a record of what was known at the time of compilation, and a lasting monument of contemporary ignorance; a point d’appui for future moralists on the scientific intelligence! Now, as a scientific subject index to be of use, must be kept up to date by publishing new volumes at periods of a few years, the question is certain to arise after the publication of the first volume, whether the same classification is to remain in a subsequent volume, or whether it is to be altered into accordance with the then present knowledge. Only two alternatives are possible, let us consider the first—that it should. Very well! look a few years ahead, and what do we see? That everyone who refers to the index for one special subject, must learn the state of knowledge of that subject, at the time each volume was published. To suppose say even ten volumes had been published is a sufficient answer to the question shall it be altered? If we say no, the classification adopted in the first volume shall remain, and no notice shall be taken of recent advances of knowledge; we not only perpetuate what we know to be wrong, and put titles where everyone knows they ought not to be, but put a positive hindrance in the way of scientific progress. In addition to this, we shall not be able to place correctly things which were not thought of when the first volume was compiled.

4. One of the most favourite reasons for classifying is that it brings titles together which treat of similar and allied subjects, and is thus a help to the memory of one who refers to the index. This is incontrovertible if it be granted that subjects can be classified in only one way, which is not true. Every group is susceptible of diverse forms of arrangement. A classification can only fulfil one general purpose at at time, and the better it fulfils this one function the worse it fulfils any other; the more difficult it is to use it for any other purpose; the more it acts as a positive obstruction to aid in other directions. The morphologist, the anatomist, the physiologist, and the chemist, each desire their own special arrangement, and a perfect classification for one is so much the less perfect for any of the others. The geologist may look at his science from the point of view of Geognesy, Dynamic, Geotectonic, Palaeontologic, Stratigraphic, or Physiographic Geology, and each one of these may require a special treatment. The engineer may want the working of metals and their physical properties, or the forms and kinds of engines—whether steam, air, oil, or gas.

5. A subject index to all sciences must not be of use only to specialists in each particular science, but must be of use to those specialists who wish to get information upon their own subject by referring to other sciences, which they think may throw light upon their own. and these will rarely be inclined to read up and master the classifications of these other sciences, for the mere purpose of being able to refer to such a key as a subject index. The geologist will not want to learn the system of classification adopted by the mathematician, merely that he may know where to look up references upon the age of the earth from the mathematical side. Nor will the mathematician and the physicist the geological classification when they are working at the same problem.

It seems appropriate here to mention a reason for a classified arrangement made by men who, having evolved such a system for their official, professional, or private letters, think it would answer for science generally. Assuming, for the sake of argument, that letters arranged alphabetically under authors or subjects, is not a better plan—usually it is, though it has never been tried—the fact staring us in the face that these men are specialists in their own letters: know everything about them and their subject matters, and hence, their own arrangement would be of little use, and possibly a hindrance to men of other professions who are not so intimately acquainted with the contents of the letters. No two men would treat similar letters in the same way without they had a specialist’s knowledge of them.

6. Assuming, for the sake of hypothesis, that a fairly good scheme of classification has been decided upon, however perfect it may appear to be, there will always crop up titles which cannot be fairly placed within it, even if the initial difficulty mentioned, of an entry being equally well placed in two divisions, be passed over. A few difficulties may be taken from Vol. VII. of the Royal Society Catalogue: Battledore scales, brewery deposit and brewers’ grains, cement, damage by caterpillars, distillery dregs, parabolic rifling, oil cake, spurious diamonds, dyeing materials, and many others from a very few consecutive pages.
7. Every classified subject index must have an alphabetical key. This not being generally disputed, what does it amount to? That everyone who uses the index will first refer to the key, to see where the subject ought to be found in the body of the volume or volumes. And that in every case, instead of one reference, two will be entailed. But why two? Why not have one only, i.e., the references printed at once in the alphabetical key? But this is to beg the whole question and adopt the method which practical experience on every side has proved to be the best.

Let not the reader start when he learns that we are now going to discuss in a similar manner the objections to the alphabetical arrangement. Having given most of the objections to the classified system, we will give all those which have ever been brought against the other method. They are only two, and would more correctly be styled quasi- or pseudo-objections. The more important is a variant upon the foregoing one number 4—that it fails to bring together similar subject matter. That the Psychologist looking up ‘Ant’ at the beginning of the volume is not reminded of ‘Wasps’ occurring at the end; although, as he is dealing with instinct, the two ought to be together, and contain, in addition, ‘Bees.’ But this is to suppose that the index is badly made or compiled; for if it were not, there would be a reference each from Ants, Wasps, and Bees to Hymenoptera, which latter heading, after dealing with all titles referring to this one division of the Insecta in its ensemble, would conclude with references to all entries of greater and less taxonomic value to which titles had been found—whether they be sub-kingdoms, classes, orders, families, tribes, groups, genera, or species, and similarly references from these smaller groups back to the larger ones. In other words, so long as anyone using the index knows the name of any one single object which he wants—this is not assuming much—he will not only find this one object at once in its proper alphabetical place, but will also then find references to all other allied objects, whether of greater generality or whether of less generality, and each of these can be found instantly, assuming, as we must, that a user of the index will have learnt his alphabet! I have already answered the criticism that classifying all such things together would save these references being made when stating that a classification can only serve one purpose at a time.

The second objection has nothing to do with the scientific value of the index, but merely its economic value to the proprietors. As references to similar matters are scattered all over the volume, an astronomer, say, cannot purchase that part only which refers to the heavenly bodies, or a botanist that part only which refers to plants. They will each be compelled to purchase the entire volume, and thus the sale of many parts will be lost. Not so, however, for the index is a mere finger-post to the ‘Transactions, Proceedings, Journals,’ etc., of scientific societies which are not likely all to be in the private library of the specialist. Hence, he will delay referring to it until he is surrounded by these volumes. He will delay using it until he visits the Library of one of the Societies which he knows contains not only the index, but the volumes to which it refers.

To sum up: a classified subject index is marred by the fact that there will always be titles as much belonging to one science as to another; there will always be doubt where to put the titles, and a fortiori always be doubt where to find them. All classifications are altered by advancing knowledge, and hence consecutive volumes must be classified differently, or things purposely put where it is well known they ought not to be. This implies that the arrangement of each volume will have first to be learnt. The more perfect a classification is for one object, the greater hindrance it becomes for every other; it cannot fulfil two functions equally well at the same time. It must be of use to specialists who wish references to other Sciences than their own; each one must accordingly acquire a knowledge of the system of the other Sciences. There will always be titles occurring which cannot be forced without undue strain into any of the divisions, which implies that they will not easily be found, if at all. An alphabetical key is a necessity.

There are thus seven good objections to a classified form of subject index, and two pseudo-objections of slight importance to the alphabetical subject index. Need anything further be said than that, having endeavoured as far as possible to state the problem fairly for both sides, the alphabetical system is incomparably the better one?

Notes and references

1. I should like to express my thanks to Hazel Bell, editor of The Indexer, for her encouragement in the preparation of this article; and, for their assistance with research: Mr Peter S. Fodden, Archivist of Oxford University Press; the staff and Syndics of Cambridge University Library; and the Reference Department of the New York Public Library. Some of the information in the introduction previously appeared in my review of the Oxford dictionary for writers and editors in the Society of Freelance Editors and Proofreaders Newsletter 23 (April 1991), pp. 7–10.
2. London: Elliot Stock, 1902.
THE MAKING OF AN INDEX


9. Cf. the note by Collins in the Preface to the *Authors’ and printers’ dictionary* (10th ed., Oxford: Oxford University Press, 1956): ‘A sketch of the way in which this work was compiled may be of interest. All suitable words, phrases, &c., were copied on to separate slips. These were then arranged alphabetically, duplicates eliminated, and the manuscript copy made.’

10. London: Chiswick Press, 1896. Collins is designated on the title page as the ‘Author of “An Epitome of the Synthetic Philosophy,” “The Tides of the West Coast of Scotland,” etc.’ This item appears to be a bibliographical rarity, as it is not listed by the British Library Catalogue, and the National Union Catalogue lists the New York Public Library’s copy as unique.

---

THE INDEXER

Thirty years ago

Volume 3, No. 4 of *The Indexer*, the autumn 1963 issue, was the last edited by John Thornton. In his farewell editorial (after five years in the office), he welcomed his successor, L. M. Harrod, and looked forward to perusing issues that we have not read in manuscript, edited, proof-read, pasted-up, re-read in page form, and despatched.

Three highly distinguished indexers came to the fore in this issue. A. R. Hewitt wrote nine expert pages on legal indexing, on which he subsequently published classic volumes. E. J. Coates described the ‘aims and methods of the *British Technology Index*’. M. D. Anderson made two appearances, writing of ‘Indexers at play’ (humorous indexes) and ‘The indexer as proof corrector’. Three pages of correspondence continued to debate the favourite topics: whether novels should have indexes; indexing from unpaged galley proofs; word-by-word versus letter-by-letter; one index, or more than one? Eleven books were reviewed, from *Machine indexing: progress and problems* to *The Bee World Index*.

This issue, completing volume 3 (164 pages), included the title-page and three-page index, compiled by the Society’s then President, Gordon V. Carey.

---

Author–indexer approved

Wholly correct authorial attitudes are shown by Homer, hero of Jane Langton’s *The memorial hall murder* (Gollancz, 1990):

‘He was supposed to be finishing the last chapter of the textbook, *The Great Cloud Darkening the Land*, which was growing out of the course of lectures. But he was bored with the last chapter. It was the index that really captured his interest. The index was going to be the best part. It was going to be the most informative, garrulous, cross-indexed index there ever was. A magnificent index. At the moment the index was only a crawling swarm of three-by-five cards, proliferating all over the table.’

This author’s wife, too, seems to be involved on an enthusiastic, rather than *faute de mieux* level:

‘Homer and Mary took a walk along the Charles after lunch, and then Mary turned around to go home to work on her half of the index, with which she was as infatuated as Homer. “Now listen, Homer, don’t forget to stop at the grocery store when you’re through at Widener. Have you got the list?” “Right here,” said Homer. “The trouble is, I’ve got two lists. One of them is the references I’ve got to check, and what I’m afraid I’ll do is march up to the call desk in Widener and pound on the counter and demand a dozen tortillas and a can of enchilada sauce.” “Well, just be sure you don’t hand me twelve volumes of the *Proceedings of the Massachusetts Historical Society* and expect me to turn them into a Mexican dinner.”’

Our thanks to Messrs. Victor Gollancz for permission to quote these passages.