A journal editor on indexing

At the first publisher I worked for, indexes were the responsibility of the journal editors, who probably delegated them to research assistants. It was always hard work chasing them, and the December issues were often delayed as a result. When I joined the Society for Endocrinology, the indexes were compiled by one of the desk editors who performed the task valiantly and very time-consuming. When this editor left and was replaced by someone with less experience, I felt she had enough to do in her first year without worrying about indexes, of which she had no experience. I therefore decided that, for one year only, I would accept the expense of having the indexes prepared outside. I was recommended to contact Michèle Clarke, of the Society of Indexers, who agreed on fairly short notice to index both journals. She supplied sample pages for our comment, so that once we received the final disk, there were not too many changes or corrections. Our printer was able to use the disk so there was no retyping to be done. Most interesting, though, was the cost, which was substantially less than I would have expected for indexing a scientific publication. This caused me to examine the previous cost of the desk editor’s time for this task and the end result is that Michèle is continuing to index the journals for us!

Sue Thorn
Manager, Journal of Endocrinology Ltd, Bristol

British/American thesaurus editions

I am delighted with Philip Bradley’s review of The Oxford thesaurus, certainly one of the more perceptive amongst the few that have appeared. (‘The Oxford thesaurus: British and American editions’, The Indexer 18(3), 192–4.)

In it, Mr Bradley asks what might well be metaphorical questions; but I think that they are all legitimate and merit a response.

[References are as in the review. For the sake of convenience, I repeat the questions.]

‘(1) sacrilege

UK The horrendous Buckinghamshire county office building is an example of architectural sacrilege.

US The horrendous office building is an example of architectural sacrilege.

... It may be pernickety to suggest this, but could not the name of the building in the UK edition be omitted or an equally horrendous building in the US be included in their version?’

I took the opportunity in coining the sample sentences for the Thesaurus to express some of my pet peeves. The county office building in Aylesbury is a singularly ugly reinforced concrete structure, an eyesore that dominates all others in the area, a paragon of bad taste. [Designed by an architect by the name of Pooley, it is widely known as ‘Pooley’s folly’; notwithstanding its truly execrable appearance, it won a design award—presumably conferred by some society for the blind—in the year of its completion.] In the driving instructions for reaching my house, I include, ‘... in the distance you will see the tallest building in Aylesbury and the ugliest building in the world’. When I was preparing the US edition, I tried to think of an equally offensive monstrosity but, while such undoubtedly exist, I could think of none that could match the low quality epitomized in the Bucks. County Office Building.

‘(2) sad

UK It was a sad day for all of us when the England team lost the semifinal.

US It was a sad day when the US team failed to win the gold medal.

... One might wonder why “for all of us” is included only in the UK edition.’

The simple answer is that a copy editor at Oxford University Press, in New York, was bent upon saving lines as best he could, here and there, to make the text fit the procrustean space allotted for it. Unless such changes distorted my examples, I allowed them to pass.

‘(3) sag

UK ... sagged under his weight.

US ... sagged under my weight.’

The NY editor changed his to my probably to avoid any accusations of sexism!

Essentially, the brief to the copy editor was, ‘If it ain’t broke, don’t fix it.’ Many of the changes—which go far beyond the few documented here—allow for changes in idiom between the two Englishes; some were made to save space; others are gratuitous and sometimes a bit weird. I reviewed the US text as changed, paying little attention to the UK text (for it did not occur to me that anyone would bother making a comparison), and, if I found the change satisfactory, I let it go.

As for the Indexes, they are a matter of preference: Americans are more accustomed to the kind found in the US edition, people in Britain to that in the UK. The UK Index was a computer-driven extract by the Oxford editorial staff, cleaving to the system preferred in the UK; the US Index reflects a form more com-
monly encountered in that country (as in Roget-type thesauruses, for instance). The differences are, in part, a reflection of the ways in which dictionaries display phrases and idioms in the two countries: in Britain, particularly in the Oxford dictionaries, idioms and verbal phrases are listed under their respective verbs, in no certain order; in the US, such lexical elements are listed in alphabetical order under what is perceived to be the 'key' word, that is, the most significant; in rare cases, a phrase may be entered under both of two key words. I suggest that those interested compare the treatments of the (many) idiomatic expressions listed under a word like take in typical American and British dictionaries.

The proof of the pudding is in the eating: although I am accustomed to both kinds of arrangement, when I look up certain things in a UK source I look for them in different places from where I would expect to find them in a US source. So it is with negative phrases (like no mean trick).

As for differences like nonaligned/non-aligned, they are mere spelling conventions, the British preference being for the hyphenated form, the American for that without a hyphen; that can be confirmed by looking down the lists of those words in any pair of dictionaries. Mr Bradley's personal preference might well make good sense, but the forms shown are those that reflect actual usage, neither personal preference nor logic.


definition of keywords.

Hanne Albrechtsen's paper on subject analysis and indexing (The Indexer 18 (4), 219–24) greatly assists in the establishment of a theory of indexing, although the author admits that some of the aims will remain unattainable as it is impossible to postulate user requirements. What should be more attainable, however, is to be able to elicit what the author of the work being indexed wishes to communicate. Furthermore, this process should be capable of being algorithmic as texts tend to follow predetermined patterns: anarchy in textual construction is not encouraged by a wide range of checks within the medium of publishing. Only very short texts, notably document surrogates such as abstracts, deliberately exclude themselves from broader contexts. On the other hand one must agree that it is necessary to broaden the horizons of mechanized indexing systems by incorporating as much lexical and encyclopaedic assistance as practicable.

The most important feature which tends to be overlooked in Albrechtsen's analysis (and in many other academic studies of indexing) is the failure to recognize that there is a primary requirement for indexes, especially those within books, to be able to recall what the reader knows to be present within the text. In part this aspect is noted with the reference to the UNIX command COMM where the index had failed to provide the key entry of compare. In broad economic terms, both the reader and the author of the text must feel cheated if an index fails to provide this service. In equally broad terms it would appear to be a pre-eminent requirement that subject indexes are at least as good as a concordance in this ability. As concordances are simple to prepare by computerized methods, the task of developing automatic indexing programs for books is perhaps not so daunting as many fear.

Like many indexers from a background of library science, Albrechtsen emphasises the importance of facet analysis. Whilst not restricting herself to Ranganathan's original absurd limit of five, she does quote nine which may, or may not, have been relevant for the index in question. The notion of some form of primary categories, themes, or facets is useful in determining indexing strategies and is broadly in accord with similar processes within authorship.

I recently spent many happy holiday hours reading Paul Theroux's The happy isles of Oceania (Penguin, 1992), which lacks an index. Amongst the themes detected were: the islands themselves (Trobriands, Easter Island, etc); famous people encountered (King of Tonga, Dame Cath Tizard and David Lange); people who infuriate the author (Thor Heyerdahl, Ronald Reagan and most Australians); food and drink (yams and former cannibals' taste for pork); religions: a total eclipse of the sun; the response of the islanders to the Gulf War, and the failure of the author's marriage. The first corresponds to Ranganathan's space facet, but the rest! And where does one place Dame Cath's neologism of fuckwit?

Kevin Jones
Welwyn Garden City, Herts

Vive la différence! The survival of the softest

Although I had read your latest paper, 'Vive la différence: The survival of the softest,' in the Proceedings of the American Society of Indexers' 1993 Annual Meeting,¹ I reread it in The Indexer (18 (4) October 1993) because I admire your writing and believe that you bring a new dimension to the indexing literature: an emphasis on sensitivity to language.

I share your assessment of the importance of this aspect of book indexing. When Richardson mocked your article on 'fine indexing' in the ASI Newsletter,² I wrote a letter to the editor defending your use of the term, while maintaining that the indexing of many publications that you might consider technical involves subtle interpretation of a text rather than mere extraction of keywords.³
Your latest article places indexing standards in opposition to ‘soft’ indexing. As Mary Piggott’s review of Cataloging Heresy (in the same issue) demonstrates, I do not advocate slavish adherence to standards, but rather careful thinking about when to follow them and when to deviate from them. I devote a large percentage of my time to reviewing and developing information standards, as well as teaching library and information science students about them. I am therefore writing to take exception to two statements about standards in your latest article:

The paper places thesauri in opposition to soft indexing, suggesting that thesauri list preferred terms that may be unsuitable for a particular book. Among the guidelines of the revised American National Standard for Thesaurus Construction, which is about to be published,4 is frequent reference to the term preferred by users in the domain of the thesaurus, rather than an insistence on either popular or technical terms, for example. This principle may be applied to book indexing: the terms most appropriate to the potential audience of the book should be selected as index headings.

Your paper suggests also that indexing standards have hard-and-fast rules for the preferred form of name headings. Some indexing manuals contain such rules, but the major principle of the most widely used code for establishing name headings is ‘the best known form of the best known name’.5 It could be real name or pseudonym; royal name or common name.

I cannot speak for the jury of the Wheatley medal, but having served on several juries for the Wilson award, I can state with assurance that comparison of the index with the text is a standard part of the evaluation process. True, it is easy to look first at the formal and structural aspects of indexes to eliminate those without continuation headings or with circular cross-references, for example. The ‘finalist’ indexes are evaluated in conjunction with the text, however.

In my published reviews of indexes, I focus not only on index structure but also on indexable matter. For example, in my review of Lancaster’s indexing text,6 I noted that a passage that discussed selective dissemination of information without naming the concept was not indexed under that term.

In reading, I constantly flip from the text to the index to check that the passages I would want to return to in the future are covered—by the terms I would seek. I annotate the indexes to my own books when they lack such terms.

I do not think that a separate society for ‘soft’ indexers is warranted. I believe that the work of technical and humanities indexers is more similar than it is different; that members of both groups can learn much from each other’s ideas; and that well designed indexing standards, cataloging codes, and thesauri can assist us rather than control us.

References

BELLA HASS WEINBERG
Division of Library and Information Science
St. John’s University, New York

This article (The Indexer 18 (4) 231–6) set me thinking (again) about the nature of indexing. Here are just a few observations.

I am a ‘genuine freelance book indexer’, but I don’t recognize the alleged divide between subject areas, nor the categories of ‘hard’ and ‘soft’ texts. You say that ‘a whole series of subject specialisms has been hived off’ (p.231) and the mainstream ‘regarded, with some condescension, as merely generalist’ (p.232). By whom, how and where? The ‘merely generalist . . . residue’ you describe is, as I see it, a collection of very important specialist subjects and forms.

Most experienced freelance indexers are probably prepared to index any subject, apart from those outside their knowledge and understanding. They aim to deal with each text in the most appropriate way. No matter what the subject is, a text needs analysis and interpretation.

All the books I deal with (many in the fields of economics, education, social sciences, and linguistics) require me to exercise my intellectual muscles and certainly make me use my imagination and ingenuity in the construction of the index. I find the same need whenever I index, whether the material is a collection of photographs or a file of committee minutes or anything else. Indexing is, after all, a creative activity, not just a semi-automatic process of word-spotting; it involves understanding each document and ‘feeling’ its message, identifying the significant concepts and representing them in appropriate words and phrases, separating out subthemes and devising suitable subheadings, and considering alternative approaches to topics.
LETTERS TO THE EDITOR

The author replies: with reference to paragraph two, subject specialisms are hived off predominantly in SI's annual publication, Indexers Available, whose layout gives more prominence to subject specialisms than to indexers themselves and indexing in general. After the single sentence in the 'Advice to publishers' on page 1, 'Most indexers will undertake popular subjects for the general reader, but subject specializations are listed under each entry', page 3 indexes the 'Subjects special skills and services'; pages 4–19 give these again in alphabetical order, with indexers undertaking each listed under them; indexers are listed individually only in the second half of the booklet, each one with their specialisms.

Hazel Bell

'Soft' indexing and the Wheatley Medal

Your article in the October Indexer was doubtless light relief at the end of the indexers' conference for which it was primarily written, but as an article in print it requires very rigorous examination. In particular, the assessment of how the Wheatley Panel operates seemed to us (as current members) insufficiently researched and rather partisan. Accordingly we must ask for the opportunity to provide your readers with a more accurate picture.

We find the implication that the Panel is dazzled by thick, difficult, specialist works and passes over the non-specialist or general works not only quite wrong but also damaging. We analyse the 'hard/dry' texts (using the article's terms) in the same way as the 'soft' texts, and for the most recent award we had the opportunity to call in specialist opinions. The picture painted of the Panel's operations strikes us as frankly insulting, not only to us as individual members but also, by implication, to the Library Association and the Society of Indexers. The Panel members carry out a difficult and invidious task as conscientiously as possible. We are looking for outstanding work, indexes

The Indexer Vol. 19 No. 1 April 1994
LETTERS TO THE EDITOR

that serve the reader particularly well, in whatever subject. Indeed, one recent submission which claimed very serious attention was a £9 paperback on a social science (ie non-hard) topic; in the end this very good and apt index had to give way to one that we felt could fairly be called outstanding.

You say your ‘impression’ is that the Panel pays ‘particular attention to conformity with rules and standards’ at the expense, apparently, of testing the index’s relation to the text itself. Nothing could be further from the truth; the seven Panel members have, between them, made extensive tests of each book’s text against its index entries. How else could we judge? As for ‘conformity with rules and standards’, that comes well down in our list of priorities in effect; suffice it to note that one of the ‘Commended’ awards in the last round had been noted as failing to conform to BS 3700!

We are astonished by your reckoning that only three ‘soft’ indexes have won the medal in 30 years. We cannot make it less than ten*, and that ignores the Commended and Highly Commended awards for books in the ‘lives and letters’ and similar ‘soft’ categories. Surely ten ‘soft’ medals in 30 years is a reasonable balance, especially when you consider that last year (for instance) the submissions of ‘hard’ texts outweighed the ‘soft’ ones by a factor of three if not four (depending on how you classify them). In other words a ‘hard’ text had three or four times more chance of winning. (Perhaps, to allay misgivings, consideration should be given to publishing the complete list of submissions each year, not just the winners? This would allow people to judge by facts rather than suppositions.)

It was also unfair to generalize, from the statement of just one (atypical) Wheatley Medal winner who had worked from abstracts and keywords rather than article texts, that this is a normal method of working for ‘hard/dry’ texts. It only takes a quick glance at the list to show that such a method is patently not typical of Wheatley submissions, among which journal indexes are pretty rare in any event.

The article seemed to us really a ‘soft’ piece in the sense of being overly dependent on assertion and impression as against hard fact. ‘Soft’ indexers certainly have their place, but they must argue their case in muscular fashion and not allow emotion—and divisiveness—to creep in. Good indexing is the choice of terms and their arrangement to get the user from the index to the right point in the text in the shortest possible time; what kind of text it is—polemical, scientific, biographical, anecdotal or whatever—and how that text is held—electronically or on paper—is irrelevant.

Cherry Lavell, Jill Ford, Nikki Coplestone and Olwen Terris
Wheatley Medal panel members

The author replies:
As Mary Piggott, quoted in my article, described recipients of the Wheatley Medal, these have gone mostly to ‘compilers of bibliographies or to indexers of a long sequence of periodicals or related documents’: dry texts and collections. Many such compilations are not intended for reading through, but for reference and selective consultation. They can be taken for indexing a unit at a time. Subsequent references in the index to the same topics treated in discrete units are repetition—structurally related—rather than continuous. Narrative texts require to be read through before indexing starts, to get a perspective of the continuous development of themes, and I would think their indexes should similarly be assessed from the perspective of familiarity with the whole text.

I have heard several indexers of periodicals state that articles in these are not read through and analysed for indexing in the same way as the text of books, and that keywords may be relied on instead: notably, several speakers at the ASI conference in May 1993, and Moira Greenhalgh at the SI’s, in York in September. Geoffrey Hamilton, then British Library Reference Division’s representative on the BSI committee revising the British Standard on the preparation of indexes, told an ALPSP seminar in June 1986 that the same topic treated in a book and in a periodical would be analysed and supplied with subheadings in the former case, but not the latter.

The nine criteria by which Wheatley submissions are judged, printed on the back of the nomination form, include ‘the avoidance of strings of undifferentiated page references’, which I have argued to be necessary to indicate repeated, minor appearances of characters in biographies, and heard many indexers of biographies agree; and ‘the clarity of the introductory note’—a mandatory inclusion, since there is no proviso, ‘when a note is necessary’, while for soft texts for general readers there may well be no need for a note. With two of the published criteria disincentives to nominate soft texts for the Wheatley Medal, the nomination of only few is unsurprising, especially with a former Chair of the ASI/Wilson award having written in this journal, ‘The case made for the information conveyed by undifferentiated references is very strong, but the principle that a string of locators without subheadings suggests minor references would have to be adopted by the members of indexing award juries, who

*ie: Clemency Canning, How to catch trout, Hakluyt Voyages, Winston Churchill, Dickens letters, History of King’s Works (definitely not ‘hard’!), John Wesley, King Lear texts, Pepys Diary, Wilson’s School.

The Indexer Vol. 19 No. 1 April 1994
automatically reject a book with six locators per entry’ (Bella Hass Weinberg, *The Indexer* 16 (4), 279).

Long, sustained continuity is one major difficulty of indexing narrative texts, resulting particularly in huge entries for main characters in biographies whose breakdown is a whole art in itself, not considered in the Wheatley criteria.

I do not question the panel’s industry and conscientiousness, but suggest that these are the reasons why, in 31 years of the medal’s history, with 27 medals awarded and 26 works commended (11 Highly; 15 Commended), totalling 53 works noticed, only five (9.4%) are sustained narratives concerning human lives (three medals—Canning biography, 1962; Churchill biography, 1967; Pepys diary, 1983; and two Commendeds—*Crisis and compromise: politics in the fourth Republic*, 1964; *Dickens biography*, 1990). The other 48 works noticed are all drily disciplinary academic specialisms, technical texts, or documentary collections, with prescribed terms and subheadings.

No soft text has won an award from the Wilson Panel.

**Hazel Bell**

**Dry texts not cut and dried**

In her address to the 25th annual American Society of Indexers Conference, Hazel K. Bell pled most eloquently for the art of ‘soft indexing’, the compendium of skills, intelligence and judgment needed to produce a back-of-the-book index for non-specialist, human-focused texts. In the process, however, she suggested of drier, technical, specialist material that ‘the sole aim is clarity; it is purely information-bearing text’. In fact, specialist texts may also be written by ‘imaginative writers with particular vision’, and the topics they deal with are far more nuanced than the non-specialist may imagine. I am currently running a department for the indexing of legal texts, and despite the highly technical nature of many of our publications, our indexing skills cannot be replaced by a thesaurus or standardized scheme. Even the most technical indexing has an art all its own.

In the legal materials I work with, three major obstacles stand in the way of both automation and standardization of the indexing process: the difficulties of textual interpretation; the variability of terminology; and different levels of readership.

Legal texts make extensive use of examples, usually from administrative and judicial case decisions, to clarify theoretical points of law and to illustrate precedence for particular arguments. The actual substance of the text, its gut, may not have anything at all to do with the specific point of the case used as an illustration. I may want to index a case annotation about an avocado tree plantation, for instance, or a chinchilla farm, under child labor laws or migrant workers—the avocados and chinchillas are beside the point. One of my favorite index entries, which I found in an old tax manual—‘Dogs, garage walls falling on’—was in fact totally mis-indexed; the gut of the text was about tax deductions for damages to personal residences.

Gauging the meaning of the text may not be a simple matter, however. Lawyers have subtle minds and the law is a complex thing; the detailed relationship between case and subject is not always obvious and it may be hard to find a way to conceptualize the idea at issue—which brings me to my second problem, variability of terminology. Many people, from non-specialist indexers to specialist text users, tend to assume that technical terminology is precise and unchanging, and that indexing it is correspondingly a simple matter. In fact technical terms lay on additional levels of meaning to the ordinary language of the text. Terms like actions and procedures have precise legal meaning but they will also be used in the text as ordinary English words. Some concepts have two or more terminological representatives; current fashion or usage may favor first one, then the other. Some terms are related but not quite exactly the same—a multi-employer pension plan and a multiple employer pension plan are defined as two separate terms in the US Internal Revenue Code, but nowadays they are generally referred to together as multi-employer plans, and many lawyers are unaware of or have forgotten the differences between them. In large multi-volume texts using several authors, differences in background, expertise, and simple human idiosyncrasy may result in a cacophony of related terms which have to be harmonized by the indexer. The terminology of ‘dry’ text is anything but cut and dried.

Of all the complex issues facing the specialist indexer, perhaps the most difficult is the problem of readership. While a publication may be theoretically aimed at readers with a particular level of expertise, or at those involved in a particular specialty, the actual readers may vary widely. Lawyer users may be specialists or generalists, corporate workers, legal clerks, students, new to the law or well-established practitioners; non-lawyers may be business people, accountants, personnel administrators or secretaries. The index has to work for as many of these people as possible.

Until editors and readers are automated, even the most technical of manuals will be written and read by human beings, with all their faults and foibles and fuzzy thinking, and it will take another human being to help them find what they want in the text.

**Kate Mertes**

*Alexandria, Virginia*