Book Indexing: A Reflection On The Contrasting Complexities And Ease Of Conceptualization And How We Can Deepen Our Understanding*

Ina Fourie
Department of Information Science, University of Pretoria, Pretoria, South Africa
Ina.Fourie@up.ac.za

Abstract
The complexities, as well as seemingly simplicities of book indexing are considered against conceptualisation and determining what an entity is about. This is done by considering an academic, theoretical perspective as well as the perspective of indexers in praxis. Various viewpoints on conceptionalisation are briefly considered, namely the simplistic approach, content-oriented approach, requirements-oriented approach and socio-cognitive view. Suggestions on bridging the gap between pragmatic indexing and theoretical indexing discourse are offered.

Keywords
Aboutness, book indexing, conceptualisation, indexing theory, indexing praxis

1 Introduction
In the subject literature there is an abundance of evidence for the importance of tools that provide access to information (e.g. back-of-book indexes, indexing and abstracting journals, databases and more lately also subject portals). These are based on conceptualizing what an entity such as a book, journal article, conference paper or website is all about. It seems as if there are contrasting opinions on the complexities (or ease) of conceptualisation and book indexing. This may also reflect on the quality of book indexes, the quality of training, and the comprehensiveness of our understanding of

conceptualization from both theoretical and practical perspectives. This can also impact on the scope of our research efforts – or lack thereof.

- This article will reflect on the complexities book indexers face in conceptualisation, by addressing the following aspects:
  - Ease or complexity of book indexing
  - Book indexing and conceptualisation from an academic, theoretical perspective
  - Book indexing and conceptualisation from an indexer perspective
  - Suggestions on bridging the gap between pragmatic indexing and theoretical indexing discourse
  - Need for research.

1.1 Importance Of Adequate Access To Subject Content

The importance of high quality book indexes is clear from the statements of experts in indexing and information organisation and retrieval: “Several years ago a British politician introduced a bill that would have denied copyright and imposed a fine on a publisher who issued a book without an index, but the bill failed” (Cleveland & Cleveland 2001:142).

Book indexes are compiled to help readers to find information in a book (or monographs such as reports, yearbooks, encyclopaedias) when they need it, and preferably in the way they need it (that is using the words and combination of words they might have in mind). The purpose is to support readers in finding information, to save time and effort, and to prevent frustration. Wellisch (1992:69) describes it as follows: “Indexing is the pinpointing of salient bits and pieces of information potentially relevant for future searches, as well as the indication of their physical space in the storage medium. Without indexing retrieval is either very spotty and haphazard or totally impractical”.

Book indexers are expected to identify all the important concepts discussed without missing anything. They should save readers time and effort in finding information. They should also not mislead readers by focusing on aspects hardly discussed. The process of identifying these important concepts is called conceptualisation, content analysis or subject analysis. The main purpose is to determine what a book or other entity is about. What are the topics, subjects, issues, concepts or main points discussed, and what do you need to consider when deciding on this? When thinking about topics, subjects, etc, it all seems so easy. It feels as if it is something you can intuitively explain, until asked to. When explaining conceptualisation to novice indexers, it sometimes seems a very simple process, as will be pointed out. I once even heard the remark “but surely this is something a secretary can do”.

112
With first practice (application), it however, often is experienced as a very complex, uncertain and difficult process. After gaining some experience, the process becomes easier for some. The question is whether it is really easier, or whether indexers are over-simplifying conceptualisation and missing the point in creating high quality indexes? Or are academics complaining about the lack of progress in conceptualisation, unrealistic about the realities book indexers are dealing with in practice? Book indexers are paid to get a job done, often within very tight timeframes, and to the best of their abilities: how much time do they then have to reflect on things like conceptualisation and its complexities? Will reflection make a difference?

In spite of all the standards and guidelines for book indexing and all the hints on layout and structure, the bottom-line is that a book index is only as good as the conceptualisation on which it is based. From the research literature it seems as if conceptualisation is still a major concern to academics and researchers. It is therefore necessary to reconsider the realities between different perceptions of the complexity of conceptualisation, and how this can be dealt with from both a theoretical and a practical perspective.

1.2 Concerns From Indexing Research Literature

When studying the research literature there still seem to be so many concerns about the seemingly “easy” process of indexing. Milstead (1994:582) for example remarks that although we have been making indexes for hundreds of years, there has been relatively little research on indexing as such. In fact, according to Anderson (1994) indexing standards are based on expert opinion rather than research results, and it is up to individual indexers to decide how to apply indexing standards.

When looking at the following examples of doubtful conceptualisations from recipe books, one can understand the concern:

- For butternut soup, the index entry appears under soup, sunshine; under butternut there is only an entry for butternuts, with bacon.
- Rum and brandy mousse (no entry under mousse)
- Quick brandied honey tart (there is no entry under tarts, honey or brandied)
- Cakes, large and quick breads, recipes (this is the index entry).

There are of course also many excellent examples of indexes receiving rewards from associations such as the American Society of Indexers (http://www.asindexing.org) and ASAIB (Association of Southern African Indexers and Bibliographers) (http://www.asaib.org.za/).

The question arises: will it make a difference if book indexers take note of these concerns, and if they reflect more on the complexities of conceptualisation? Will book
indexes improve, or will it be a frustrating waste of time? What can indexers and academics do to bridge the gap between theory and praxis, and to stop the ongoing complaints of lack of progress and lack of understanding of practical realities?

2 Ease Or Complexity Of Book Indexing

2.1 Seemingly Easy Steps To Book Indexing

If things seem to be easy, everybody wants to try their hand at it: nothing wrong with that, as long as we also realise that things are not always what they seem to be. In the case of book indexing it is certainly more complex than it may seem at first impression.

Basic indexing guidelines as portrayed in textbooks such as Borko and Bernier (1978), Cleveland and Cleveland (2001) and Lancaster (2003), as well as in indexing standards Recommendations for examining ... (1984), approach indexing with seemingly easy step-by-step directions: the indexer scans the book, and makes notes on the key concepts discussed. The indexer then decides what the book is all about and chooses index entry terms for these. This should not be too difficult, especially when you are advised to look at chapter headings, main headings, subheadings, the introduction, the closing paragraph or summary, words in italics and bold, figures and diagrams, etc. The things to consider when deciding what an entity is about seem to be very simple and straightforward.

When we train book indexers we often also simplify the principles and processes. This is good according to teaching theory: complex activities are broken down into smaller more manageable steps, so that trainees will believe that they can master the skill, and that practice will make perfect. Indexing is portrayed as an art (Wellisch 1992). Indexers only need to practice and build confidence in order to become good indexers: it is proclaimed that practice makes perfect. Cleveland and Cleveland (2001:97) for example explain: “A person who examines and uses indexes critically will gradually learn what an ideal index should be like, although indexes vary widely in their characteristics and quality”. But is it really as simple as this to become a good indexer?

Guidelines from theoretical discourse often portray a more complex picture. Hjørland (2001:774) for example adapted the guidelines offered by Wilson on how to determine what a document is about:

• Determine the author’s intentions and purpose in writing the document;
• Read the document and weigh the relative dominance and subordination of different elements covered in the document;
• Group or count the concepts mentioned in the document; and
• Develop a set of rules to determine what is essential in the document.
Novice book indexers are thus often faced with a job considerably trickier than expected. The question arises whether they realise how complex conceptualisation really is, and why it is so difficult? Do they link their indexing practice to indexing theory? Do they think about indexing and conceptualisation and its complexities? Do they think about the questions they should ask and the reasons why indexing theorists are so concerned about the lack of progress in indexing? To what extent are they aware of these concerns? And also, to what extent should they give this some thought?

2.2 Knowing The Indexing Basics: Easy So Easy, Or Doubting Whether Your Are Getting It Right

When indexing in real life, novice indexers soon realise that they have to deal with a number of things at the same time. Farrow (1991) explains this as different kinds of conceptualisation, namely conceptualisation of the:

- Subject
- Index structure (layout, etc.)
- Specific requirements set by for example the publisher
- Users and their needs
- Background or general world knowledge.

Novice indexers also soon realise that:

- Indexers will differ on what a document is about, and on the indexing terms to use. In indexing theory this is called inter indexer consistency. Ingwersen (1992) considers this to be a good thing, since different approaches in assigning indexing terms will also be consistent with the different approaches readers will use in finding information. In practice it can, however, cause a lot of anxiety among novice indexers to decide who is right, and who is wrong.

- An indexer will assign different indexing terms at different points in time. In indexing theory this is called intra indexer consistency, and once again Ingwersen (1992) argues that this can have a positive effect on information retrieval, because readers will also use different terms when approaching the indexing at different points in time.

Consistency in indexing, does however, not necessarily imply good indexing: indexers have shown that they can also be consistently wrong in their conceptualisation.

Chu and O’Brien (1993:440) further adds to the complexities of conceptualisation: “There are several approaches to this large question of determining aboutness in text. The theoretical approach offers behavioural and/or linguistic interpretations and the practical approach is to be found in textbooks and manuals on indexing and other historical background and guidance on indexing with an emphasis on individual systems and using them to assign indexing terms”.

115
Before assessing whether it is necessary to give more thought to conceptualisation, one, however, has to acknowledge the realities with which book indexers have to deal, namely:
- Limited amount of time allowed to complete an index
- Limited fees being paid
- Their own understanding of the impact of indexing on information retrieval, which might be either rather limited or daunted by the complexities of conceptualisation.

Farrow (1991:157) remarks: “An indexer’s conceptual knowledge will clearly take time to develop”. It therefore may seem as if time and practice may be the solution to quality conceptualisation and indexes. Is this assumption, however, true, or should one add the condition that indexers should first of all be aware of the need to develop their conceptual knowledge in order to improve indexing as well as information retrieval?

2.3 No Need To Understand “Aboutness” Or “Topics” Or “Concepts”, But Pick Them Correctly

Book indexing is about conceptualisation. That means deciding what a book is about, selecting the concepts that are worth mentioning in the index and translating these concepts into words (i.e. concepts that reflect what the book is about). Some academics and researchers portray this in very complex terms. According to Lancaster (2003:15) on the other hand conceptual analysis means “nothing more than identifying the topics that are discussed or otherwise represented in a document”. “Conceptual analysis, first and foremost, involves deciding what a document is about – that is, what it covers” (Lancaster 2003:9). Questions that should come to mind include:
- What in this book would readers like to know about?
- What is the intellectual content?
- Which topics are covered?
- Which themes, things or aspects are discussed?
- What are the units of thought?
- What is the philosophical content?
- What is the focus of the work?

There are, however, also more complex interpretations of aboutness (e.g. Bruza, Song & Wong 2000; Hjørland 2001, 1992; Hjørland and Albrechtesen 1995). Weinberg (1988:3) also explains: “It is contended in this paper that indexing which is limited to the representation of aboutness serves the novice in a discipline adequately, but does not serve the scholar or researcher, who is concerned with specific aspects of or points-of-view on a subject”. According to Hutchins (1977:17): “One of the most crucial problem areas of information science concerns the identification of what documents are ‘about’.”
In contrast to highly theoretical discourse on the meaning of “aboutness” and “conceptualisation” Lancaster (1991:11) takes a more pragmatic approach: “However, if one must reach agreement on the precise definition of terms before pursuing any task one is unlikely to accomplish much – in indexing or any other activity”. According to Lancaster (2003:13) it is sufficient to recognise that a document is of interest to a particular user group, and to translate these into concepts that will support information retrieval.

From the preceding discussion it is clear that indexing and conceptualisation often appears to be relatively simple processes when considering the guidelines offered during training and in textbooks. In practice indexers may initially, however, experience it as confusing and complex. Is it then simple or complex? Often indexers become more confident about their skills with practice and experience; thus easy becoming easier, which also seems in line with Lancaster’s reasoning. Although it is good for indexers to gain self-confidence, this might mean that the complexity of indexing might be underestimated, which would explain the concern expressed by theoreticians that there is no progress in indexing practice and that indexing research has too little impact. If we are very confident about what we are doing, there is no need for us to reconsider our practices and to learn from research findings and theoretical discourse.

A next question to ask is then: how complex can conceptualisation be, and do we need to look at it from different points of view?

### 3 Need For Different Ways Of Looking At Conceptualization

Apart from the complexity of the meaning of conceptualisation, there are also different views of looking at conceptualisation. Albrechtsen (1993) distinguishes the following: 

#### 3.1 Simplistic Approach

The topics mentioned in books are regarded as absolutely objective, and they can be used as indexing terms by merely picking the topics form the words in the text (e.g. the title, main headings, etc.). It is a very pragmatic approach, not too time consuming and much easier to apply than the other approaches. It, however, means that no value is added (e.g. in linking the content to specific problems that it might solve) and that users’ needs and the potential usefulness of a book for the intended user group are not considered. The simplistic as well as the content-orientated approach can be compared to Soergel’s (1985:227) entity-orientated indexing: all the indexing terms are directly selected from the entity (e.g. book). It seems as if this can be linked to rule-based indexing as explored by Quinn (1994):140): “The goal of indexing theory thus becomes not
one of discovering subliminal processes but of constructing explicit, well formulated rules that can be used to yield indexing phrases from texts”.

3.2 Content-Orientated Approach

The content-orientated approach includes an interpretation of a book’s contents that involves the identification of topics that are not explicitly stated in the text. The indexer also needs to pick up on implied concepts. An article on the University of South Africa, for example, also implies distance education. Identification of implied concepts requires that the indexer should know something about the subject matter. Although the content-orientated approach takes a deeper look at the content of a book, it still does not recon with the users’ needs, or the potential use of the information contained in the book. This can be supplemented by Adams and Baker’s (1968) views on mission and discipline in indexing.

3.3 Requirements-Orientated Approach

The requirements-orientated approach takes into account possible requests from users and also considers contextual frameworks. Albrechtsen (1993:222) explains that the indexer should ask: “… how should I make this document, or this particular part of it, visible to potential users? What terms should I use to convey its knowledge to those interested?” With this approach a book is analysed to predict its potential usefulness. This can be compared to Soergel’s (1985:230) request-orientated approach.

The advantage of the requirements-oriented approach is that it can help in solving problems experienced by users, and in the transfer of knowledge. It is more time consuming than the other approaches and requires a very good knowledge of the users, their needs and the problems they need to solve.

The socio-cognitive view can be linked to the requirements-orientated approach. Hjørland and Albrechtsen (1995) initially referred to this as the “domain analytical approach”, while Fidel (1994) writes about user-centred indexing.

3.4 Socio-Cognitive View

The question arises whether we should consider the users as individuals, or as people belonging to a larger group, sharing similar information needs and vocabulary. The socio-cognitive approach seems to stress the latter (Hjørland 2002a, 2002b, Hjørland and Christensen 2002). According to Jacob and Shaw (1998:142) the socio-cognitive approach shifts attention from individual knowledge structures to discourse domains and thus to knowledge-producing, knowledge-sharing, and knowledge-consuming communities.
A simplification of the socio-cognitive approach will read as follows: people live, work, study and communicate in a social environment. They form part of a group of people who talk to one another (a discourse community). They are influenced by their culture, historical developments, political ideologies, economic situation, educational and work environments, and so on. This can be referred to as a socio-cultural environment that applies to children, students, academics from different disciplines, professionals from different sectors, et cetera. These people communicate with one another. They share information, ideas, differences, and motivation. They learn from the “history” of the particular group and the inherent culture. They share a common language (to some extent) of concepts, theories, expressions, methods, et cetera. They use this language (vocabulary) when

- searching for information
- communicating information (e.g. informal discussions, conference papers, articles, books)
- producing or describing information (e.g. news reports, lectures, adverts, building plans, software design)
- using information (e.g. interpreting published documents in terms and concepts that they can use for their productions).

Wilson (as cited by Jacob & Shaw 1998:150) for example found that social workers’ document representations should focus on problems they actually face. In his research Swift (as cited by Jacob & Shaw 1998:150) found that social scientists need multiple document representations for all documents (e.g. theoretical focus, research methods, research paradigms). Tibbo (1992, 1994) again found that specialized guidelines are necessary for indexing and abstracting in disciplines from the Humanities such as History. Although the before-mentioned distinctions may seem complex and confusing, it is really necessary for indexers to consider the view they are taking on conceptualization and how this may affect information retrieval – the ultimate purpose of indexing.

4 Book Indexers And Conceptualisation: The Academic And Theoretical Point Of View

Quinn (1994:146) makes the following statement about indexing theory and practice: “Everyone is a “theorist” in that he/she holds beliefs about something. The value of theory is that it makes explicit what we all do implicitly. By making one’s beliefs explicit, one becomes more aware of one’s ideology and its strengths and weaknesses. Becoming conscious of the assumptions and principles that underlie one’s techniques, operations, and services is the first step to improving them”. This should also apply to views on conceptualisation. Bridging the gap between indexing theory and practice is,
however, somewhat more complicated. It sometimes seems as if the issue is viewed from two completely different worlds.

The following are some of the reasons for the gap:

- Indexing theory is often described in highly academic language, and published in journals not normally read by indexers. Travis and Fidel (1982:43) for example remark that “subject analysis is a mature field with a large body of theory and practice gained over more than a century of systematic study. The very existence of this body of knowledge is largely unknown to the general public”.

- Indexing practice is influenced by the knowledge, background, experiences, et cetera of the indexer. This is referred to as his/her cognitive framework and includes knowledge of the document’s subject, experience in conceptualisation, knowledge of the information retrieval system (in this case book indexes) and knowledge of the potential users. To adopt theoretical views on indexing, indexers need to open their minds and develop their theoretical cognitive frameworks. They also need to learn the language of academic discourse. Although very noble, I think this might be asking too much if one considers the fees indexers can expect in contrast to the time and effort required to understand academic research and discourse, and furthermore, to figure out how to apply this in practice.

- Theoreticians consider indexing standards and textbooks to be too pragmatic with too little explicit application of research findings. A lot of criticism is offered, which might be experienced in a negative way by indexers.

- Academics and researchers specializing in indexing often publish in journals that are not read by indexers (e.g. Journal of the American Society for Information Science and Technology), or deliver their papers at more academic conferences not attended by indexers.

Some suggestions will be offered in section 6 on how this gap can be bridged.

5 Book Indexers And Conceptualisation: A View Point From Praxis

When considering indexers in real life, we should especially look at the

- misconception that there is no acknowledgement in practice for indexing theory
- harsh, practical realities indexers often have to face.

Many examples of theoretical underpinnings offered by practising indexers can be found in the more “practically orientated” literature. A good example is Kendrick and Zafran’s (2001:ix) citation of Thomas: “It is not only about competency with the language of the law or the ability to execute the job. Legal indexing is also about values.
There is a generosity that great indexers bring to their work. There is a deep-seated passion to help a faceless future user who will be working an important case, perhaps late into the night. And above all, there is a desire to get it right”.

The following examples offer further illustration on how indexers reflect on theoretical complexities of indexing without always explicitly acknowledging it:

<table>
<thead>
<tr>
<th>Consider the reader’s background and vocabulary.</th>
<th>“The sophistication of those who will be using the index influences what detail is needed. Targeting patenting attorneys or bankruptcy attorneys is different from indexing for a more general audience that may include law students or the general public” (Knapp, as cited by Kendrick and Zafran 2001:viii).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendrick and Zafran (2001:viii) also cites Baggiano who offers advice on indexing for law school casebooks: “emphasize concrete terms for the law student, and abstract ones for the professor, and terms of art for both, e.g. the mailbox rule, the mirror image rule, etc.”.</td>
<td></td>
</tr>
<tr>
<td>Consider the usability of conceptualisation and indexing terms in existing indexes.</td>
<td>Learn from the readers and ask their advice on existing indexes (Karnezi, as quoted by Kendrick and Zafran 2001:viii).</td>
</tr>
<tr>
<td>Consider the reader’s search strategies and search terms (also reflected in the socio-cognitive approach).</td>
<td>According to theoreticians there is more uniformity among readers than one might expect.</td>
</tr>
<tr>
<td>Acknowledge that conceptualisation is influenced/based on an understanding of the subject area, especially for concept-based indexing.</td>
<td>Indexers are advised to deepen their understanding of the subject area in which they are working.</td>
</tr>
</tbody>
</table>

From the before-mentioned examples it seems as if in spite of the concern about the lack of progress in implementing indexing research results and recommendations in practice, as well as the concerns about the inadequate progress made towards perfect indexing, there is sensitivity for indexing theory amongst some indexers. The question is now, how to spread this sensitivity to more indexers, and also to sensitise theoreticians to the practical realities indexers are facing. Such realities include: limited time to complete a book index, limited fees, limitations on the scope of the index, and restrictions placed by publishers who do not always understand the impact of book indexes on information retrieval.

When considering book indexing and conceptualisation one should realise that there are good and poor indexes. Good indexes, and the more “theoretical” advice offered by some practising indexers should be the benchmark for novice indexers. Although
it might be difficult for novice indexers to identify such “theoretical underpinnings” for indexing practice, it is worth pointing these out in training sessions.

As for the practical realities, theoreticians should perhaps take a closer look at indexing guidelines offered by publishers and database producers, and how to lobby for guidelines that will reflect theoretical concerns, before merely offering criticism on indexing practices.

6 A FEW SUGGESTIONS ON ACKNOWLEDGING THE COMPLEXITY OF CONCEPTUALIZATION, BUT AT THE SAME TIME MAKING IT EASIER

The following are a few suggestions that might help to bring the different interpretations of the complexity of conceptualisation closer and to bridge the gap between indexing theory and practice:

- If we expect indexers to take theory more seriously, we should make theory and research findings accessible in language that can be understood by novice indexers, and in publications that are more widely read (e.g. The Indexer).
- Academics should be prepared to offer presentations at conferences attended by practising indexers (e.g. ASAIB conference or conferences of the American Society of Indexers), and to explain how their research findings or ideas can be implemented in practice – with acknowledgement of the realities faced by indexers.
- Academics should also be realistic about the fact that it takes time and effort to create the “perfect index” – something which indexers do not have or are most often not paid for. If this is not acceptable, methods should be found for lobbying more reasonable time schedules and payment for indexers, or to investigate how much time users are prepared to spend searching for information in contrast to paying more. Academics should help indexers look for realistic ways to move to “perfect indexing” and “perfect conceptualisation”.

Practising indexers also need to do more “homework” by getting more involved in indexing theory and making a definitive effort of at least noting indexing research projects and findings, for example through monitoring conference programs and the tables of contents of academic journals dealing with indexing. They could also get involved in grassroots research, for example what goes on in their minds when indexing, how do they go about in conceptualization, and what are their perceptions of aboutness and conceptualization. They can also query the success of their indexing and the effect it may have on information retrieval. Other questions they might ask include:

- What am I trying to do with the index? (purpose)
- Can I try my index out with a few potential readers? Examples of usability studies are reported by Olason (2000) and Jörgensen and Liddy (1996).
- Can I ask their advice on some existing indexes (e.g. previous editions)?
• What do I know about the potential readers? Are there any guidelines in the subject literature? (e.g. indexing for law, medicine)? Are there perhaps different groups of readers?
• Is there any way that I can get to know how readers are actually using my index?
• Do I know enough of the subject? What can I do to learn more?
• How does my work compare with the work of other indexers?
• Is it time to query my own skills and spend time at a conceptualisation workshop with other indexers (getting to understand our cognitive processes)?
• Do I know enough about book indexing? Where can I learn more?
• How will I try to read more about indexing for specific target groups?
• Which factors are influencing my work, and is there anything I can do about it?

The bottom-line of the advice to indexers is: take time off to study theoretical discussions, and make it their own (e.g. Bates 1998, Bertrand & Celliers 1995, Blair 1990, Borko 1997, Borko & Bernier 1978 and Diodata & Grandt 1991). The bottom-line of advice to practitioners: write for fellow academics, but also put ideas into words that can be understood by “ordinary” people. Get in touch with practicing indexers, learn about their problems, and make it tempting for them to learn about indexing research.

Publishers can also contribute to bridging the gap between indexing theory and praxis by deciding what they need for their readers and how much they are prepared to pay for, and to note research findings.

7 Conclusion

There are many ways to look at indexing and especially conceptualisation. Two of the dangers are to over simplify the process or to make it so complex that it turns into a daunting task that no practising indexer would dare to attempt.

In answering the question set in the introduction, I would like to conclude that nothing in life is ever as simple as it seems. To lure us into attempting to try something, it is in order to simplify things. Once we feel comfortable with the new skill, we should ask ourselves: what more can there be to this skill? Indexing should, however, also not be so complicated that we do not want to attempt it.
References


Tibbo, HR. 1992. Abstracting across the disciplines: a content analysis of abstracts from the natural sciences, the social sciences, and the humanities with implications for abstracting standards and online information retrieval. *Library & Information Science Research* 14:31-56.


