

# Hard Drugs and Soft Terms: the Challenge of Description

Anne Bailey

DrugScope

23 October 2004

# A History of the ISDD Thesaurus

- **1968**
  - ISDD Library derives a set of 200 subject headings from those used by the Library of the UN Division of Narcotic Drugs
- **1973**
  - US Drug Abuse Council awards 3-year grant to extend these headings
  - UK Department of Health and Social Security extends ISDD grant so professionally-qualified staff can be employed.

# A History of the ISDD Thesaurus (2)

- **1977**
  - King Edward VII Fund grant awarded
- **1978**
  - Preliminary edition issued for in-house use and testing
- **1980**
  - First edition of the ISDD Thesaurus published
- **1989**
  - Revised edition incorporating changes and feedback

# A History of the ISDD Thesaurus (3)

- **1989**
  - Computerization of ISDD catalogue begins
- **1993**
  - 2<sup>nd</sup> edition Thesaurus published, incorporating changes to aid in computerized retrieval
- **2003**
  - Creation of Information Officer – Cataloguing and Indexing post, dedicated to cataloguing and thesaurus development

# Words or Concepts

- In creating a thesaurus we are mapping concepts – ideas that form the important topics in our subject area
- The words used to describe these concepts are labels – short-hand to signify the underlying concepts
- *This is particularly important when creating a thesaurus that covers more than one language or more than one version of a language: what needs to be “translated” is the concept – not just the label*
  - e.g. Prevention in the UK is a wider concept than in the US, and includes harm reduction activities, including safer injection.

# Coordination

- The ISDD / DrugScope Thesaurus was designed to use largely post-coordinate terms – i.e. broad terms are combined to specify narrow concepts
  - *cf.* chain indexing, where pre-coordinated terms are used
  - Only “*largely* post-coordinate” – it has been necessary, for pragmatic reasons, to include some quite specific, *pre*-coordinated terms – e.g. motor abilities-performance
  - Main advantage of post-coordination is flexibility – the searcher can combine terms to reflect different nuances to meet their needs.
  - Main disadvantage – false drops, which are catered for by precoordinating narrower terms – e.g. Treatment Outcomes

# Currency - General

- Currency is a big issue for most indexing systems
  - Controlled vocabularies take time to develop
  - It is important to maintain controlled vocabularies – ideally on a monthly or weekly basis
  - Accuracy and currency can sometimes appear at odds with each other as goals for a thesaurus
    - New vocabulary appears every day and it is a challenge to identify what is here to stay and what is just a buzz-word
    - Academic study can focus on loci of interest for brief moments: it is a challenge to identify new developments in the subject as opposed to brief obsessions.

# Currency - DrugScope

- At DrugScope we cater for this by inputting abstracts
  - Abstracts are free-text and introduce current terminology instantly
  - In-house, we can search keywords field and the abstracts field separately or concurrently
  - Online we provide a simplified subject search that allows keyword, title and abstracts fields to be searched together
  - Searching keywords alone retrieves fewer documents with high specificity
  - Searching keywords & abstract together retrieves more documents with less specificity



# Hard Terms

- e.g. scientific terms (drug names, medical terms)
- Always easier to handle
  - Process of scientific nomenclature is designed around specificity
  - To a certain degree, science is an International language
  - Less likely to change over time
- At DrugScope we prefer scientific drug terms, with slang / street names as cross-references.
  - Argot varies from country to country and region to region (e.g. in UK English, “dope” usually means cannabis. In US English it means cocaine.

# Soft Terms

- “Soft terms” are more flexible and can be more wordy – they describe concepts that may be more culturally-specific or given to change over time.
  - It can be difficult, or even arbitrary, to decide which soft term should be preferred and which should be simple cross-references
- Such terms can appear as “woolly” or “fuzzy” or “difficult to define” but, used in combination with other terms, they often allow retrieval of material that would be impossible to find using “hard terms” alone.

# Examples of Hard and Soft Terms

- Hard Terms

- Cannabis
- Body
- Extraction
- Humans

- Soft Terms

- Identification of drugs
- Screening
- Diagnosis of dependence
- Social Interaction
- Educational level

# Intuitive Language

- Less-intuitive indexing systems have the advantage of “flagging up” the existence of terminology that the user and the system designer may be using in different ways.
- There is a constant challenge to balance the desire to welcome the user to the system with the need to provide an indexing system that is accurate.

# Towards a European Thesaurus

- Multilingual thesauri have issues around translating concepts, cultures and labels when creating terms.
- The level of pre-coordination and post-coordination can be difficult to set
- Hierarchy can be a real issue – within different cultures there can be a mapping to and from different levels of a hierarchy
- It becomes harder to make a thesaurus structure and terms intuitive across many cultures
- *How* the thesaurus is used is vital – will these terms be the users' only entry-points?

# Indexing ELISAD

- A good thesaurus “maps” the key areas in a subject area. After decades of indexing, it is time to ask:

How different is substance use “cartography”  
across our shared continent?