Searching the Literature

“The typical review involves tasks that many scientists find irksome—an ambitious literature search, obtaining of documents, extensive reading, reconciliation of conflicting claims, preparation of citations—and the bulk of effort is centered on the work of others, as if one had to write Chapter 2 of one’s dissertation all over again... The present movement in research synthesis... encourages consideration of all empirical studies on a subject—not only the published but the unpublished ones—so as to capture... the full range of reported statistical effects.”

White (1994)

Searching the Literature

- Goal of a good literature search:
  - Uncover all relevant studies
    (i.e., studies that meet inclusion criteria)

- More realistic goal:
  - Minimize differences between retrieved and unretrieved studies
The Population of Studies

- Population
  - All relevant studies
- Accessible population
  - All retrievable studies
- Sample
  - All retrieved studies

Generic vs. Specialized Search Engines

- Generic search engines (e.g., Google) can sometimes be a good resource for unpublished material
  - Far less efficient though

<table>
<thead>
<tr>
<th>Search Type</th>
<th>Primary Benefit</th>
<th>Primary Drawback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specialized db</td>
<td>Retrieval of unpublished work</td>
<td>False hits</td>
</tr>
<tr>
<td>Generic search engine</td>
<td>Retrieval of unpublished work</td>
<td>False hits (dramatically)</td>
</tr>
</tbody>
</table>

Electronic Search Strategies: Potential Problems

- Publication bias
  - Known difference in statistical significance of published vs. unpublished studies
  - Hence, don’t rely on electronic searches alone
- English language bias
  - Tendency for studies published in English language journals to have stronger findings than studies published in other languages
  - Hence, don’t rely on works published in English alone
- Inconsistent usage of terms across disciplines
- Inconsistent labeling of methodology
Inconsistent Labeling of Methodology

• Reviewers only interested in experiments will sometimes use search strategies aimed at capturing design
  (After-school or afterschool) AND
  (random or evaluation or RCT)
• Several studies have shown that this strategy will miss relevant studies (25-67%)!

Hope for the Future

• CONSORT statement
  – Specific elements of studies that ought to be reported
    • Standardizes terms
    • Adopted by major medical journals and the APA
• Structured abstracts
  – Specific effort aimed at standardizing the information contained in abstracts

Searching the Literature: Finding the “File Drawers” by Snowballing

• Contacts & Solicitations
  – Email correspondence to active researchers
  – Postings on relevant listservs
  – Email to organizations that might conduct relevant research
    • Reviewer must be aware of these (again, subject matter expertise is critical)
• Reference Treeing
  – Search the reference sections of all relevant studies and reviews for citations to works not found
• Databases of unpublished and ongoing research
  – Usually topic specific
• Citation Searches
  – SSCI & Google Scholar
Social Science Citation Index & Google Scholar

- Can find …
  - works that cite a particular reference
  - works by a particular author
  - Note that for SSCI, the journal must be indexed by Thompson Scientific (about 4000, or roughly 20% of all journals in all academic disciplines)
  - Example on next slide looks up
Example of Non-electronic Search

- Cooper et al. (2002) on Modified School Calendars
  - Wrote to the 100 largest school districts in the U.S.
  - Contacted all authors of included studies
  - Reference tree for all included studies

Information Retrieval: A Continuous Process

- Preliminary Searches
  - Supports beginning steps: Definition of key concepts & research question
  - Use of standard reference tools and broad searches for review articles and key primary studies

- Main Searches
  - Identification of primary studies through searches of online databases, printed indices, Internet, branching, hand-searches
  - Most difficult given a number of challenges

- Final Searches
  - Occurs towards the end of the Review Process
  - Refine search terms and update original searches
Main Searches: Decisions

- Selection of Information Retrieval Tools
  - Scope of search: Which disciplines or subject fields should be searched (including all related fields)?
  - Availability of indexing tools & expertise: Which tools do we have access to at our institution? Are there others who can perform searches for us? (more about this later)
  - Format of indexing tools: What format are they in (e.g. online, print, web-based)?
  - Dates: How far back does the indexing go for each tool?
  - Language: What is the language of the material that is indexed? How can we locate non-English material?
  - Unpublished work: How can we access dissertations, reports, & other grey literature?

Selection of Databases

- Consult your academic library’s website to learn what databases are accessible from your institution.
  - Example: Social Work Resources at LUC

Examples of Databases

- Education: ERIC, British Education Index, Australian Education Index, CBICA Education, Education: A SAGE Full-text Collection, Education Abstracts.
- Psychology: PsycINFO, PubMed (Medline), ApaLine, Psychology: A SAGE Full-Text Collection, Criminology: A SAGE Full-Text Collection
- Sociology: Sociological Abstracts, Contemporary Women’s Issues, Sociology: A SAGE Full-text Collection
- Multidisciplinary: Academic Search Premier, ProQuest Dissertations and Theses, FRANCIS, Social Sciences Index, SCOPUS, Web of Science
Main Searches: Decisions

- Preparation of Search Strategies
  - What are the key concepts to be searched?
  - How are these represented in each discipline?
  - What are their related terms?
  - How are these key concepts represented in the controlled vocabulary within each database to be searched?

Sample Research Question

The purpose of this review is to assess the impacts of MST on out-of-home living arrangements, crime and delinquency, and behavioral and psychosocial outcomes for youth and families

Using a Thesaurus

1. From the research question, determine the main concepts to be searched (usually there are three):
   - Multisystemic therapy
   - Multisystemic treatment
   - Research or evaluation or outcome
2. Consult the main database to be searched.
3. Look up each concept in the thesaurus for this database.
   - A thesaurus is an alphabetical listing of the controlled vocabulary (or descriptors) used within a subject database
   - A hierarchical arrangement is used so that Broader, Narrower and Related headings may be discovered
   - The user will be sent from invalid headings to valid headings
Main Searches: Decisions

- Construction of the Search Statements
  - What terms should be searched as descriptors or as "free text"?
  - What Boolean operators should be used?
  - Where should truncation characters be used?
    - (e.g. parent* will retrieve parent, parents, parental)
  - What limiting features are available to narrow results?
    - (e.g. use of Publication Type codes)?
  - What time period should be searched?

Boolean Operators

AND: Both terms must be present in order for a record to be retrieved. Used to combine different concepts.
* e.g. parent participation AND achievement

OR: Either term may be present in order for a record to be retrieved. Used to search for related terms or synonyms.
* e.g. parent OR family

NOT: Used between two terms to ensure that the second term will not appear in any of the results.
* e.g. literacy NOT adult

(Parental involvement OR parent participation) AND academic achievement AND (elementary OR primary education)
Next Steps

Repeat these steps for each database to be searched.

Additional Databases

- Consult with the C2 Information Retrieval Methods Group
  - Learn availability of Trials Search Coordinator or IRMG liaison

Managing Your Results

- Export the results
  - Save as a Text file
- Import into a bibliographic management software:
  - Reference Manager,
  - EndNotes
- Edit your inhouse database
  - Add Source code for each database searched (e.g. ERIC1, PsycINFO1,...)
  - Add notes to the records (e.g. includes vs excludes)
- Compile a Search History document listing the original search strategies
  - Use of IRMGECG worksheet (to come)
Key Decisions in Literature Retrieval

- Important to have inter-rater agreement on key decisions
  - Does the study look like it might be relevant?
    - Based on reading titles and abstracts
    - If yes, retrieve the full text
  - Is the study in fact eligible for inclusion?
    - Based on information in the full text of study
    - If yes, code study

- When in doubt, double code
  - At least two trained raters working independently
  - Generally I double code everything

Information Retrieval: Wrap Up

“Shoestring-budget information retrieval is likely to introduce bias, and should be avoided.”
(IR Policy Brief, 2004)

- Importance of information retrieval process
  - Not a “one-shot” deal
  - Requires expertise in the planning and implementation of searches
  - Consulting with the IRMG or an Information Specialist is highly recommended (Some coordinating groups have Trial Search Coordinators)

- Use of bibliographic management software
  - Store, manage and organize results

- Ability to replicate review
  - Documentation of entire process, including search strategies used for each database, decisions taken, etc.