Systematic Review Workshop

Systematic Review

Workshop 2011

Zoe Handley
(zoe.handley@education.ox.ac.uk)

Plan

• Background
• Further information
• How to carry out a systematic review
  1. Setting the scope and methods for the review
  2. Gathering and describing research
  3. Appraising and synthesising data
  4. Making use of the review
• Systematic review checklist
• Administrative systems

What is a Systematic Review?

• Systematic reviews attempt to reduce the subjective bias characteristic of many traditional literature reviews through the use of a transparent and explicit protocol, exhaustive database searches, explicit inclusion/exclusion criteria, and quality assurance measures, i.e. hand searches to validate database searches and double blind reviews of individual studies.

Zoe Handley
Applied Linguistics, University of Oxford
History of Systematic Review

- The evidence movement
  - Search for evidence to support public policy decisions
  - Critical appraisal and presentation of research in a systematic manner

- The Cochrane Collaboration (Evidence-based medicine)
  - Systematic management of knowledge in the domain of healthcare

- The Campbell Collaboration
  - Adapting the Cochrane methodology to broader public policy
  - Examining the effects of social interventions
  - Social welfare, crime and justice, education

- The EPPI Centre
  - Evidence and Policy in Practice Information Co-ordinating Centre
  - Adaptive methodology in education and social welfare

FURTHER INFORMATION

Methods

- EPPI-Centre: [http://eppi.ioe.ac.uk/cms](http://eppi.ioe.ac.uk/cms)

Options for further information:
- [Examples of reviews](http://eppi.ioe.ac.uk/cms/Home.aspx?TabId=81)
- [Overview of methods](http://eppi.ioe.ac.uk/cms/Home.aspx?TabId=89)
- [Methods references](http://eppi.ioe.ac.uk/cms/Home.aspx?TabId=1919)

Citing the EPPI-Centre:
- EPPI-Centre (March 2007) EPPI-Centre methods for conducting systematic reviews. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
CONDUCTING
A SYSTEMATIC REVIEW

Overview of the Systematic Review Process

1. Setting the scope and methods for the review
   - Define review questions
   - Define inclusion/exclusion criteria

2. Gathering and describing research
   - Key word map
   - Searching, screening, keywording, mapping

3. Appraising and synthesising data
   - In-depth review
   - Qualify and relevance appraisal, synthesis, conclusions

4. Making use of the review
   - Communication

1. SETTING THE SCOPE AND METHODS FOR THE REVIEW

Setting the scope and methods for the review

1. Define review questions
2. Define inclusion/exclusion criteria
3. Define protocol
   - Statement of review question
   - Conceptual framework and method for review

Note:
- Systematic reviews are intended to be policy and practice oriented. It is therefore good practice to involve users in the definition of research questions and inclusion/exclusion criteria.
- If your review is initially broad, the definition of review questions and inclusion/exclusion criteria might involve an iterative process.

Zoe Handley
Applied Linguistics, University of Oxford
Setting the scope and methods for the review

- **Users**
  - Oxford University Press and English language teaching professionals

- **Review questions**
  - What empirical research could be found on the use of new technologies in language learning and teaching with learners in primary and secondary schools since 1990?

- **Inclusion/exclusion criteria**
  1. Report on the use of technology
  2. In foreign/second language learning
  3. With school age (primary and secondary) learners
  4. Focus on learners
  5. Describe or include an empirical study carried out by the author(s)
  6. Has been reported between 1990 and 2009
  7. Have been published peer-reviewed journal articles, and
  8. Have been published in English.

---

2. GATHERING AND DESCRIBING RESEARCH

Gathering and describing research

1. **Searching for studies**
   a. Database searches
   b. Hand searches
   c. Literature review searches
   d. General search engines, e.g., Google Scholar
   e. Specialist websites
   f. Experts in the field

2. **Screening studies (against inclusion/exclusion criteria)**
   a. Title and abstract
   b. Full document

3. **Describing studies (keywording)**
   a. Generic
   b. Review specific

4. **Mapping studies (produce keyword map)**

Searching for studies using Databases

- **Aim is to identify a comprehensive and unbiased set of papers**
  - Including published and unpublished papers
  - Easily accessible and harder to find papers
- **Use a range of databases (available on OXIP+)**
  - ERIC (Education Resources and Information Center)
  - LILACS (Latin America and the Caribbean's Health Literature in Social and Public Health Sciences)
  - PsychINFO (Psychology)
  - PESTEC (Computer Science)
  - Index to Theses
  - Dissertations and Theses (ProQuest) / Dissertation Abstracts International

---

Zoe Handley
Applied Linguistics, University of Oxford
Searching for Studies using Databases

- Developing your search strategy
  - Based on inclusion/exclusion criteria
  - Combination of:
    - Free text searches (keyword searches)
    - Thesaurus searches

Free text searches

- Techniques: Regular expressions
  - Search for all variations of a word
    - Truncation (e.g., * or $)
      - *: "educ*" will find education, educators, educational, educate, educated...,
      - $: "educ$" will find educational, education, etc.
    - Wild character (\?):
      - ??: "educ??r" will also find research, researchers, etc.
    - Search for phrases:
      - "language learning"
- Techniques
  - Combining search terms
    - Boolean operators: OR, AND, NOT
      - OR: broadens your search
      - AND: narrows your search
      - NOT: excludes terms
    - OR should be used before AND
      - bullying AND secondary school OR secondary education

Zoe Handley
Applied Linguistics, University of Oxford
Thesaurus searches

- Thesaurus searches
  - Thesaurus terms are attached by trained indexers
    (in the ERIC database they are referred to as descriptors - an example is presented on the next slide)
  - Developing a list of thesaurus terms
    - Peer grouping
    - Use the thesaurus terms that have been used in closely papers which you know well the inclusion criteria.
    - Thesaurus mapping
    - Map your own search terms to subject headings.
    - Do not include methodological terms as these are usually not well indexed.

Piloting Searches

- Searches should aim to identify all and only studies which meet the inclusion/exclusion criteria for the review.
- Tha results should aim to achieve a balance between precision and recall.
- Tha results should only studies relevant to the review, (even exactly).
- Developing a search strategy is therefore an iterative process.
- Screen titles and abstracts against inclusion exclusions criteria.
  - Include abstracts. Note potential additional search terms (i.e. keywords) and also terms that provided virtual results.
  - Check thesaurus to add additional descriptors (thesaurus terms) and also description that you may wish to execute.
- Repeat the process.
- Keep a log of pilot searches.

Example Database: ERIC

- What empirical research has been undertaken on the use of technologies in language learning and teaching with learners in primary and secondary schools since 1990?
- Free text / Keyword search
  - initial search
  - Language AND (learning AND (computer OR technology))
  - using regular expressions
  - Measurement of term, Pronouns
- Developing and refining the search
  - Exploring
  - English only
  - Education level
- Combine searches from search history using OR/AND.
Saving your Search Strategies (ERIC)

1. Language learning or A253
2. Language learning and A253
3. Language learning
4. Language learning or A253
5. Language learning
6. Language learning
7. Language learning
8. Language learning
9. Language learning
10. Language learning
11. Language learning
12. Language learning
13. Language learning
14. Language learning
15. Language learning
16. Language learning
17. Language learning
18. Language learning
19. Language learning
20. Language learning

**Erwin 3.0 query**

- [language learning] or
- [language learning] or
- [language learning] or
- [language learning] or
- [language learning] or
- [language learning] or
- [language learning]

For an extended example see:

http://www.eecs.ucsb.edu/erwin/erwin3x/3x_english/

Saving Your References

- Save your references to a reference management system:
  - EndNote
  - RefWorks

- RefWorks guides:
  - Quick start:
    http://www.refworks.com/content/quick_start_guide.wsp
  - Webinars:
    http://www.refworks.com/learn/help/Refworks.htm

- RefWorks course:
  - OUCS, Wednesday 20th March:
    http://www.oucs.ox.ac.uk/file/journey/detail/57029

Logging your Searches

<table>
<thead>
<tr>
<th>Database searched</th>
<th>Check &amp; present searching</th>
<th>Screening</th>
<th>First period of search</th>
<th>No. of titles examined</th>
<th>RefWorks imported?</th>
<th>Imported into which reference?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERIC</td>
<td>23.08.01</td>
<td>Yes</td>
<td>01/01/2005</td>
<td>2043</td>
<td>Yes</td>
<td>ERIC</td>
</tr>
<tr>
<td>PsycINFO</td>
<td>23.08.01</td>
<td>Yes</td>
<td>01/01/2005</td>
<td>297</td>
<td>Yes</td>
<td>PsycINFO</td>
</tr>
<tr>
<td>Social science citation notes</td>
<td>23.08.01</td>
<td>Yes</td>
<td>01/01/2005</td>
<td>24</td>
<td>Yes</td>
<td>Social science citation notes</td>
</tr>
</tbody>
</table>

Example Database: ERIC

- What empirical research has been undertaken on the use of technologies in language learning and teaching with learners in primary and secondary schools since 1990?

- Thesaurus search
  - Identifying descriptors
    - Use keyword from articles that match your inclusion/exclusion criteria
    - Explore the thesaurus
    - Developing and refining the search
      - Data range
      - English title
      - Education level
    - Combine searches from search history using AND/OR

Zoe Handley
Applied Linguistics, University of Oxford
Example Database: PsycInfo

- What empirical research has been undertaken on the use of technologies in language learning and teaching with learners in primary and secondary schools since 1990?

- Thesaurus search
  - Identifying descriptors
  - Limited search: topic to subject heading (contextual knowledge and Thesaurus search)
  - Search tools: single tool or several

- Developing and refining the search
  - Location: publication year
  - Contains keywords from search history using AND/OR
  - Field string articles

Example Database: PsycInfo

- Saving your Search Strategies (PsycInfo)

Zoe Handley
Applied Linguistics, University of Oxford
Google Scholar

- A useful tool for identifying initial articles and search terms
- Advantages:
  - Allows full-text searches
- Limitations
  - Scanty about its coverage — some publishers do not allow it to crawl their journals
  - Ranks articles on a combined measure of relevance, citation counts, publication date, etc.
  - Puts high weight on citation counts — so, the rich get richer in terms of citation counts
  - Puts high weight on words in titles

Database Searches (Two-Stage Screening)

- For each database
  1. Pilot searches
  2. Run final search
  3. Screen titles and abstracts, applying inclusion/exclusion criteria
  4. Keyword studies which meet inclusion criteria
  5. Screen full texts, applying inclusion/exclusion criteria
  6. Keyword studies which meet inclusion criteria

Practical

- Start developing the search strategy for your own study using ERIC or PsycInfo
  1. Go to OXLIPE
  2. Select database
  3. Follow steps for developing a search strategy
- Discuss observations and problems encountered

Hand Searches (One-Stage Screening)

- Searches of electronic databases will not show up every article
- Hand searching the principal journals on the topic allows researchers to validate database searches
- For each journal:
  1. Screen titles, abstracts and full texts, applying inclusion/exclusion criteria
  2. Keyword studies which meet inclusion criteria

Zoe Handley
Applied Linguistics, University of Oxford
Literature Review Searches

- Searching the reference list of literature reviews on the topic also allows researchers to validate database searches
- For each literature review:
  1. Screen titles, abstracts and full texts, applying inclusion/exclusion criteria
  2. Keyword studies which meet inclusion criteria

Keywording Studies

- Generic keywording
  - [Link](http://www.ioe.ac.uk/EPP/WidContent/downloads/EPP_P1_Keyword_strategy_0.9.7.pdf)
  - Language, Country, Topic, Curriculum, Population (teachers vs. learners), Age of learners, Sex of learners, Type of study (descriptive, exploratory, relationships, naturally-occurring experiment, researcher-manipulated experiment)
- Review specific keywording
  - Technologies, SLA theories, SLA pedagogy

Zoe Handley
Applied Linguistics, University of Oxford
Mapping Studies: An Overall Map

- See handout

Mapping Studies: Cross-tabulations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>N</td>
<td>N</td>
<td>P</td>
<td>N</td>
<td>P</td>
</tr>
<tr>
<td>Vocab</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Grammar</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reading</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Writing</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Speaking</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Listening</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Zoe Handley
Applied Linguistics, University of Oxford
3. APPRAISING AND SYNTHESISING DATA

Appraising and Synthesising Data

- In-Depth Review
  1. Define research questions
  2. Define inclusion/exclusion criteria
  3. Screen studies in keyword map
  4. Double blind reviewers
     a) Complete evaluation grid for each study
     b) Evaluate weight of evidence of each study
     c) Compare and come to an agreement with second reviewer
  5. Synthesise the findings

Evaluation Grid

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Criteria for quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Effective abstract provide sufficient information for initial overview</td>
</tr>
<tr>
<td>Introduction/Background</td>
<td>When did it take place?</td>
</tr>
<tr>
<td>Methodology/Inference theory</td>
<td>Why was it carried out?</td>
</tr>
<tr>
<td>Purpose</td>
<td>For what purpose was it carried out?</td>
</tr>
<tr>
<td>Study design</td>
<td>Was it carried out in a particular group of people?</td>
</tr>
<tr>
<td>Statistical power</td>
<td>Which statistical test was used?</td>
</tr>
<tr>
<td>Limitations</td>
<td>Were the limitations clear?</td>
</tr>
<tr>
<td>Methodology</td>
<td>Were the methods of the study reliable?</td>
</tr>
<tr>
<td>Results</td>
<td>Were the results reproducible?</td>
</tr>
<tr>
<td>Generalizability</td>
<td>Are the findings generalizable?</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Were the conclusions supported?</td>
</tr>
</tbody>
</table>

Evaluation Grid

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Clearly stated or implied?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>Does it use sophisticated language?</td>
</tr>
<tr>
<td>Methodology</td>
<td>What is the methodology adopted?</td>
</tr>
<tr>
<td>Inference theory</td>
<td>What is the inference theory stated?</td>
</tr>
<tr>
<td>Variables</td>
<td>What are the independent/dependent variables?</td>
</tr>
<tr>
<td>Sampling</td>
<td>What was the sampling method?</td>
</tr>
<tr>
<td>Population</td>
<td>What was the population studied?</td>
</tr>
<tr>
<td>Sample size</td>
<td>Was the sample adequate?</td>
</tr>
<tr>
<td>Bias</td>
<td>Were there any biases?</td>
</tr>
<tr>
<td>Effect sizes</td>
<td>Were the effect sizes calculated?</td>
</tr>
<tr>
<td>Generalizability</td>
<td>Are the findings generalizable?</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Were the conclusions supported?</td>
</tr>
<tr>
<td>Limitations</td>
<td>Were the limitations clearly stated?</td>
</tr>
</tbody>
</table>

Zoe Handley
Applied Linguistics, University of Oxford
Evaluation Grid

<table>
<thead>
<tr>
<th>Grouping</th>
<th>How many groups?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>How were the samples divided into groups?</td>
</tr>
<tr>
<td></td>
<td>Did the groups have the same statistical variances?</td>
</tr>
<tr>
<td></td>
<td>What treatment was given to each group?</td>
</tr>
<tr>
<td></td>
<td>Data collection</td>
</tr>
<tr>
<td></td>
<td>Were the data collected through the use of a reliable method?</td>
</tr>
<tr>
<td></td>
<td>Were there sufficient numbers of data collected?</td>
</tr>
</tbody>
</table>


Data analysis

<table>
<thead>
<tr>
<th>Data analysis</th>
<th>Is the data analysis appropriate for the purpose?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Does the analysis seem like a valid way of analyzing the data?</td>
</tr>
<tr>
<td></td>
<td>Does the analysis seem like a valid way of approaching the data?</td>
</tr>
<tr>
<td></td>
<td>Are the results in the paper consistent with the research question?</td>
</tr>
<tr>
<td>Result/findings</td>
<td>Are there any limitations in the reporting of the results?</td>
</tr>
<tr>
<td></td>
<td>Do the results match your assessment of the findings/results?</td>
</tr>
<tr>
<td></td>
<td>Are limitations of the study discussed (e.g. confounding variables)?</td>
</tr>
<tr>
<td></td>
<td>Are there implications for teaching and training?</td>
</tr>
<tr>
<td></td>
<td>Do the implications match the study findings?</td>
</tr>
</tbody>
</table>


Weight of Evidence

- Rate each of the following on a 3 point scale (High, Medium, Low) for each research question:
  - Relevance of particular focus of the study for addressing the research question of this systematic review
  - Appropriateness of research design and analysis for addressing the research question of this systematic review
  - Testworthiness – can the study findings be trusted in answering the research question of this systematic review
  - Contribution of the study to answer the research question of this systematic review

Methods for Synthesis

- Statistical meta-analysis
  - "A statistical meta-analysis is a statistical procedure designed to combine the numerical results of primary research studies addressing similar research questions.
- Narrative synthesis
  - "A narrative synthesis brings together the results of primary research that go beyond a narrative form in providing an evaluative discussion of results from individual studies in abstract form in summary tables."
- Conceptual synthesis
  - "A conceptual synthesis is a way of understanding or conceptualizing the framework brought together to create new concepts or concepts."
- The selected method will depend on the type of study (quantitative, qualitative) and the quality of the reporting (is there enough information to calculate effect size?)

Zoe Handley
Applied Linguistics, University of Oxford

26/01/2011
4. MAKING USE OF THE REVIEW

Final Report

- EPPI Centre produces reports in 3 formats
  - One page summary – key findings
  - Short report – concentrates on findings and provides some details of method
  - Technical report – includes all details, including search strategies

- Provide enough detail for someone to update the review, i.e. re-run searches

SYSTEMATIC REVIEW PROCESS CHECKLIST

MARKING

Zoe Handley
Applied Linguistics, University of Oxford
Marking Conventions for MSc ALSLA

- For candidates submitting a Systematic Review instead of an empirical study, examiners will, in addition to many of the criteria above, be looking for:
  - Evidence of understanding the principles underlying systematic review
  - Evidence of completeness and coverage of review
  - Evidence of bias-avoidance and objectivity
  - Transparency of the study selection process
  - Contribution to knowledge building in the specific review field
  - Contribution to teaching and learning

NB: Systematic Reviews should contain a non-technical summary accessible to readers who are not necessarily experts in the field.

(UGC ALSLA handbook, 2011: 24)

Producing the Flow Chart

- Health Warning!!

- You will need to produce a flow chart which shows how you filtered the papers from the output of the database searches to the studies included in the in-depth reviews.

- It is important to set up good administrative systems from the start because you will require a lot of information in order to produce the flow chart for the review.

- The information that you will need to produce at each stage is summarised on the following slides.

Producing the Flow Chart (2)

- Database searching (Two-stage screening)
  - Stage 1: Screening titles and abstracts
    - Number of articles which the database searches threw up
    - Number of articles included
    - Number of articles excluded
    - For each inclusion/exclusion criterion, the number of articles that were included for that reason
    - Number of articles unsure
  - Stage 2: Screening full documents
    - Number of articles screened (unseen from previous stage)
    - Number of articles included
    - Number of articles excluded
    - For each inclusion/exclusion criterion, the number of articles that were excluded for that reason
    - Number of articles that you could not get hold of

Zoe Handley
Applied Linguistics, University of Oxford
### Producing the Flow Chart (3)

- **Hand searching (One-stage screening)**
  - Stage 1: Screening titles, abstracts, and full documents
    - Number of articles screened
    - Number of articles included
    - Number of articles excluded
    - For each inclusion/exclusion criterion, the number of articles that were excluded for that reason.

- **In depth review**
  - Number of articles screened
  - Number of articles included
  - Number of articles excluded
  - For each inclusion/exclusion criterion, the number of articles that were excluded for that reason.

### Administrative Systems

- Save all searches
- Reference management system in which all references are saved (RefWorks recommended; see slide 26)
- Spreadsheet recording all searches (Excel recommended; see slide 27)
- Spreadsheet recording filtering, keywording, and weight of evidence (SPSS recommended; see slide 29)
- Save all evaluation grids

### On the up side ...

- All the effort systematically coding studies according to different themes etc. is worth it in the end.
- One problem I often encounter when reviewing literature is determining how to structure the synthesis.
- Systematically coding the studies really helps you to see the different themes in the research and identify possible structures for your synthesis.

---

**Zoe Handley**  
**Applied Linguistics, University of Oxford**
Further help

- EPPI Centre website: http://eppi.ioe.ac.uk/cms

- Library:
  - Training sessions: http://www.bodleian.ox.ac.uk/education/services/training
    [Mon 14th Nov, Mon 31st Jan]
  - WebLearn: http://weblearn.ox.ac.uk/portal/hierarchy/secsci/education/library

- Contact me: zoe.handley@education.ox.ac.uk

How can I be systematic?

- What quality assurance measures can I use in my review?
  - Validate the application of inclusion/exclusion criteria to database/searcher with hand searches
  - Validate keywording with evaluation grids
  - Note reasons for Weight of Evidence ratings
  - Check intra-rater reliability for Weight of Evidence ratings

Zoe Handley
Applied Linguistics, University of Oxford