EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

EIPPEE Training Workshop

Bologna meeting room
15th Floor of the Ministry of Education, Culture and Science (OCW),
Rijnstraat 50, 2515 XP, The Hague, Netherlands

www.eippee.eu

Twitter @EIPPEEnet
#EIPPEE2012

Systematic review methods

David Gough and Janice Tripney

8 May 2012

Education and Culture DG

Social Science Research Unit

Five streams of work:

- Childhood Studies
- Evaluations of Social Interventions
- Sexual Health, Reproduction and Social Exclusion
- Evidence for Policy and Practice Information and Co-ordinating Centre
- Perspectives, Participation and Research

http://eppi.ioe.ac.uk/

EPPI-Centre

Evidence informed policy and practice

Methods development in EIPP

Software tools (EPPI-Reviewer)

Online and traditional courses and a MSc

Undertake systematic reviews

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Outline

1. Research and decision making
2. Accessing research
3. Stages and types of review
4. Review questions
5. Searching and screening
6. Coding, quality appraisal
7. Synthesis
8. Using research

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

1. Research and decision making

Is research relevant to decision making?

Research: ‘the systematic study of materials and sources in order to establish facts and reach new conclusions’

Oxford Concise Dictionary

YES!

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Modernizing Government, UK 1999

“What must change…… This Government expects more of policy makers. More new ideas, more willingness to question inherited ways of doing things, better use of evidence and research in policy making and better focus on policies that will deliver long-term goals.” (page 11)

Demand for research

Genuine interest from policy makers and practitioners for research to inform policy - though normal policy dynamics apply

Irritation from contradictory advice (without clarity as to why) which devalues academic input (perceived as opinion)

Long term gains in credibility and use of research in policy and practice

2. Accessing research

How would you know if relevant research is available?

- Quantity – even academics can not keep up with research
- Location – distributed in many journals
- Language - technical
- Skill in quality and relevance appraisal

Activity:
Whole group discussion

You're in a Parent Teachers Association Meeting.

The head teacher is speaking about the school's homework policy. In the room are teachers, parents (one of whom is also a researcher) and school governors.

Someone says "I've read a literature review that says there is no point in trying to get children to do homework with their parents: it just doesn't work".

What might it be useful to know about this review?

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Trust individual studies?

Studies may be excellent but even if of high quality, dangers from:

- Relevance of study / assumptions
- Chance findings
- Publication bias
- Fashion and the ‘seems sensible bias’
- Vested interests

Trust traditional literature reviews

Traditional reviews may be excellent but dangers from explicit:

- Theoretical and ideological assumptions (perspectives driving review)
- Boundaries of knowledge (sampling bias)
- Methods of synthesis (interpretative bias)

Need to know what we know from *all relevant and trustworthy* research?

So need some methodology of review

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Trust experts and expert panels?

Experts may be skilled but dangers from lack of explicit:

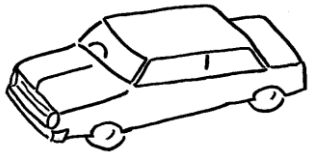
- Methods (boundaries of knowledge)
- Theoretical and ideological assumptions ("single topic pressure groups")
- Boundaries of knowledge (sampling bias)
- Up to dateness (e.g. BSE advice to government)
- Methods of synthesis (interpretive bias)
- Distinction between research and practice knowledge
- Opinion: claim that Himalayan glaciers could melt away by 2035 was a scientist's 1999 media interview.

What is a systematic review?

- Formal accountable method for bringing together what we know – accessible and understandable and explicit about how framed and how executed:
- Systematic: ‘done or acting according to a fixed plan or system; methodical’
- Review: ‘a critical appraisal of a book, play, or other work’ (OED)

EIPP: basing decisions on systematic evaluations of the literature (not individual studies, traditional reviews and expert opinion)

Need for reviews? an example



The UK government has proposed to introduce driver education programmes in schools and colleges. A recent review suggests it may lead to:

- Early licensing
- A modest but potentially important increase in the proportion of teenagers involved in traffic crashes

Roberts, Kwan and the Cochrane Injuries Group Driver Education Reviewers. School based driver education for the prevention of traffic crashes (Cochrane Review). In: The Cochrane Library, Issue 4, 2002. Oxford: Update Software.

3. Stages and types of review

What do we want to know from research *(question)*?

What studies would help (*inclusion criteria*)?

How do we locate these (*search strategy and screening*)?

What do we need to know from each study?

How to we assess them (*qual & relevance*)?

How can we integrate them to answer the question (*synthesis*)?

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Stages of map and synthesis

Form review team (involve 'users')

↓

Formulate review question, conceptual framework and inclusion criteria

↓

Search strategy + screening to identify relevant studies

↓

Describe studies

↓

Assess study quality (and relevance)

↓

Synthesise findings

↓

Communicate and engage

↓

Map

Synthesis

May be narrower than map

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Map: what has been studied

The nature of the research field. What we know about what has been studied – and what has not been studied!

Synthesis: what it tells us

What we know about the overall messages

- inform decision making by policy makers, professionals, the public
- inform the focus of new primary research

But which research paradigm? Quantitative, qualitative, both?

• Reviews well known for ‘what works’ and cost effectiveness questions using randomized controlled trials

• But we can apply to all research questions (including process) and thus all primary research designs

• From quantitative experimental (statistical empirical meta analysis) to meta ethnography (conceptual synthesis including meta ethnography)

Examples

- What research has been undertaken on school size? – systematic *map* of various aspects of studies to date

- What is the relative effect of school size on various educational and social outcomes? – *statistical meta synthesis* of RCTs

- What are the processes by which school size has these effects– *conceptual synthesis* of qualitative process studies

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Aggregative reviews

Aggregative reviews predominately add up (aggregate) findings of primary studies to answer a review question…

… to indicate the direction or size of effect

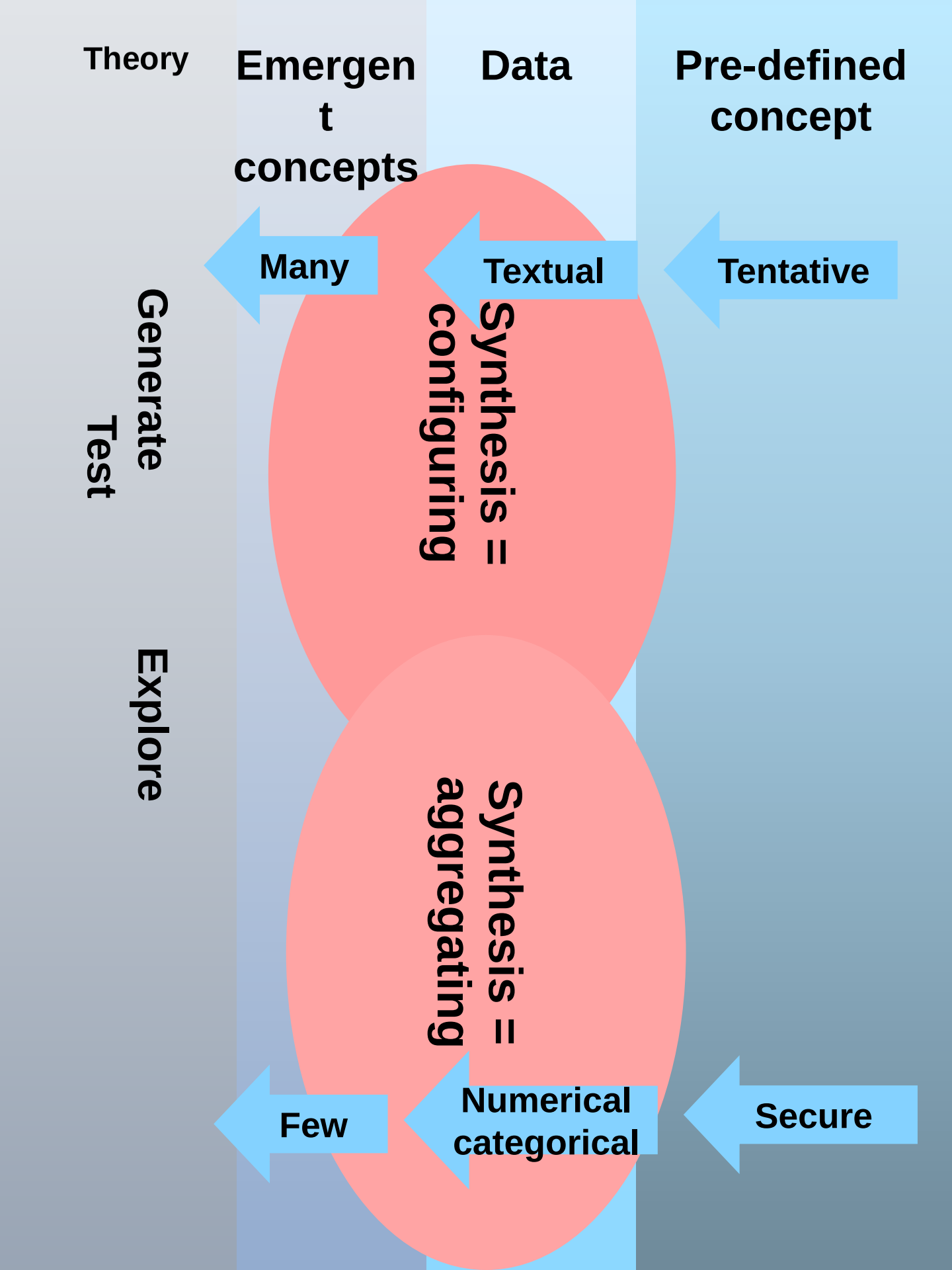
EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Configurative reviews

Configurative reviews predominately arrange (configure) the findings of primary studies to answer the review question….

… to offer a meaningful picture of what research is telling us

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE


Theory
Emergent concepts
Data
Pre-defined concept
Many
Textual
Tentative
Generate
Test
Synthesis = configuring
Explore
Synthesis = aggregating
Few
Numerical categorical
Secure

How broad a review?

- Research is often narrowly framed
- Narrow on topic
- Narrow on method
- Two strategies:
    i. Mapping
    ii. Mixed methods reviews

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Breadth and depth and time

- Reviews (maps and syntheses) are not of a set breadth and depth

- Time, other resources, and type and extent of review need to be fit for purpose

- Reviews can be Rapid Evidence Assessments

- Easier if part of a strategic approach to knowledge production

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Why only research knowledge?

Knowledge source	**Knowledge gained by/ from…**
Research	Doing research and producing research findings
Practitioners	‘Doing’
Policy community	The wider policy and political context
Service users	Experience of, and reflection on, service provision
Organisations	Regulations and procedures

Adapted from Pawson R, Boaz A, Grayson L, Long A, Barnes C (2003) *Types and quality of knowledge in social care.* London: Social Care Institute for Excellence.

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

and whose research knowledge?

• All knowledge is question, theory, and ideology dependent - so whose questions?

• Different perspectives lead to different questions and thus different answers – biased thus relevant

• How to consider which perspectives are driving (or could drive) reviews of research

• Policy, practice, user of service, driven and owned research
- school student led review example

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Knowledge synthesis and use as a driver for fit for purpose research:

What decisions could be assisted by research findings?

What do we know already and how do we know it?
(systematic maps and synthesis)

What more do we need to know and how can we know it?
(new primary research)

So user driven relevant research questions and fit for purpose research methods

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Review limitations

- Limited perspectives of many reviews
- Limited questions of many reviews
- Limited study methods considered
- Limited data reported to inform applicability
- Narrow/shallow rather than broad/deep
- Convergent rather than divergent
- Rapid reviews? (narrow or expert reliant)

4. Review questions

Form review team (involve 'users')

↓

Formulate review question, conceptual framework and inclusion criteria

↓

Search strategy + screening to identify relevant studies

↓

Describe studies

↓

Assess study quality (and relevance)

↓

Synthesise findings

↓

Communicate and engage

↓

Map

Synthesis

May be narrower than map

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

The review question

- Is the driver for all review processes
- Is an investigative statement rather than a topic of interest
- Is in dynamic interplay with theory and inclusion/exclusion criteria
- Should be clear and answerable
- Should need an answer and be worth answering

Types of research question

What are the questions that users want answered?

- What do people want? (Needs)
- What's the balance of benefit and harm of a given intervention? (Impact/effectiveness)
- Why/how does it it work? (Process/explanation)
- What is happening? (Implementation)
- What relationships are seen between phenomena? (Correlation)
- What are people's experiences?
- ….

Different research designs best for addressing different questions (fit for purpose)

Moving from a research topic to a research question

e.g. “How can UK develop educational leadership courses for heads in the UK?”

- To what extent have leadership courses been implemented in UK? (Implementation levels)
- What do head teachers in UK say they need from leadership courses? (Levels of need)
- What are the benefits and costs of on-line learning for leadership development of head teachers? (Impact)
- How do these strategies produce their effects on professional development of head teachers? (Process or explanation)

Formulating the review question

What type of question?

Specify/clarify main concepts

Useful acronyms: ‘PICOC’ or ‘SPICE’

- Population?
- Intervention and Comparison (or other phenomenon, processes, perspectives…)?
- Outcome or other Evaluative measure?
- Context or Setting?

……..Also Time?

Conceptual framework

Theoretical assumptions that underpin the review question

Influences all stages of the review

- criteria for including or excluding reports
- terms for searching for relevant literature
- terms for sorting/describing literature
- detailed questions to ask of each research report
- framework for the synthesis stage of review

Example review question

How do parents influence the impact of homework on attainment in primary school children?

What might the conceptual framework look like for this question?

This is a broad question – sub-questions could include:

- What are the effects on student attainment of initiatives that teach parents how to help with homework?
- How do different kinds of parental involvement in homework influence student attainment?
- Through what mechanisms does parental involvement in students' homework affect student attainment?

Nature of homework, attainment, parental involvement, students and theories about relationships between them

'Measures' of all of these concepts

Study types that can help address these questions

Dimensions for inclusion criteria, e.g.

Nature of what's being studied

- (specific practices/approaches/outcomes/ other phenomena?)

Setting and population

- (e.g. school type, student age, adults, countries)

Date of research

- (ever, since 1920, since 1990, Why?)

Research methods

- (all methods, only empirical, only certain designs?)

Language of report

- (e.g. English only, Norwegian only, both?)

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Draft inclusion criteria for a review

Review question: How do parents influence the impact of homework on attainment in primary school children?

1. Participants - Must be students, parents, teachers of primary school age pupils (age 5-11)
2. Settings - State funded elementary schools in an OECD country
3. Intervention - Homework (= Programme of work set by a teacher to be completed out of school hours in addition to work carried out within the classroom)
4. Study design
    a. Must be empirical examination of the relationship between parents' participation in homework and attainment; that also collects own data (not a review)
    b. Must have a variable reporting amount / type homework given
    c. Must have a variable reporting attainment
5. Published after 1980
6. Published in English language

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Workshop Activity:

Developing review questions and inclusion criteria

Individually (20 minutes)

- *using flip-chart paper provided*
- *write down a review question of interest to you*
- *draft 5 or so inclusion criteria*

Work in pairs and whole class discussion (10 minutes)

5. Searching and screening

Form review team (involve ‘users’)

↓

Formulate review question, conceptual framework and inclusion criteria

↓

Search strategy + screening to identify relevant studies

↓

Describe studies

↓

Assess study quality (and relevance)

↓

Synthesise findings

↓

Communicate and engage

↓

Map

Synthesis

May be narrower than map

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Searching: common characteristics

Searching is the act of seeking studies that might be relevant for the review's question

Searches are derived from the review's question

Practically constrained (by database limitations, restricted time and resources available to reviewers)

As in other stages, reviewers:

- Have a rationale for their searching methods:
- Draft, test and implement a structured search plan ('search strategy');
- Report fully on their methods and findings.

Seeking specific types of study

Beware variation in the ways study types are understood and labelled. Individual study types, e.g. ‘survey’ ‘controlled trial’

- may not identify themselves as such in publications
- are often indexed poorly by commercial databases

Systematic reviews are worth seeking out, but challenging

- sources of studies, as well as being studies in their own right
- may also be labelled as meta-analysis, research synthesis….
- tend to be clustered on websites of review producing organisations

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Searching is a skilled and time consuming activity

Tap into the skills of information, subject and systematic review specialists, to:

- Identify sources to search
- Deal with technicalities (e.g. of databases)

Plan time for developing and testing searches

Approaches to searching

‘Purposive’ searching

- E.g. to identify main themes in the literature (sometimes uses saturation sampling),
- Searching plans may develop as evidence comes to light.

Comprehensive searching

- Considers ideal as having access to all studies that answer the review question. An unbiased sample is next best.

Challenges of comprehensive searching (1)

Findings **evade** us:

- Profusion of published and unpublished material
- Much hidden: only 50% abstracts presented at conferences are later published in full
- Different databases use different terms to classify studies
- Different databases cover different, largely discrete areas of literature
- As individuals we are constrained by our own disciplinary and policy area knowledge (which journals, which books, which databases?)

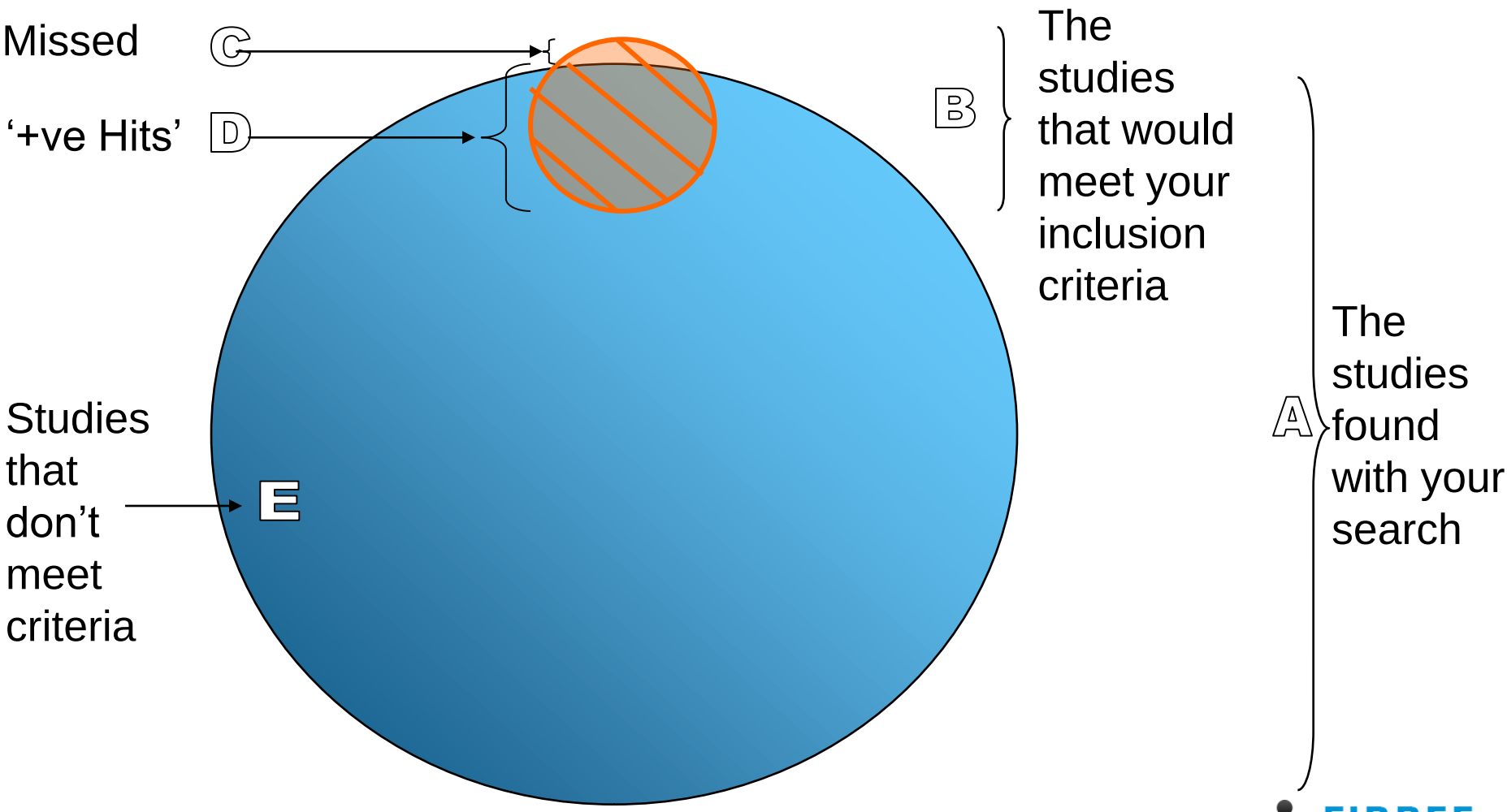
Challenges of comprehensive searching (2)

Some evasion is systematic*

E.g. statistically significant, ‘positive’ results more likely to be:
- Published
- Published rapidly
- Published in English
- Published more than once
- Cited by others

*work cited in Egger M et al (2003) . How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health Technology Assessment 2003; Vol. 7(1)

Comprehensive searching – a picture of a sensitive search



EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Sensitivity and specificity

Sensitivity

- Ability to locate all studies of interest
- = the number of relevant studies located as a proportion of all that exist (D/B)
- Low sensitivity means that many relevant studies missed by search

Specificity

- Accuracy of the search strategy in identifying studies of interest
- = the proportion of the total number of studies identified by search which are deemed ‘relevant’ (D/A)
- Low specificity means that the search identified many studies that were not relevant to the review.

There is always a trade off between sensitivity and specificity

The importance of searching a range of sources

How are decisions made about the entry of people aged 65+ to care services?

	SSCI	Medline	CINAHL	Caredata
Unique articles retrieved	237	182	27	16
Unique relevant articles	116	73	24	15

Taylor B, Dempster M and Donnelly M (2003) Hidden Gems: systematically Searching Electronic Databases for Research Publications for Social Work and Social Care. C J Social Work, 33:423-429.

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Types of sources of research

- Bibliographic databases
  - ‘General’ (e.g. ERIC, Medline)
  - ‘Specialised’ (e.g. CERUK register of ongoing educational research, SIGLE for grey literature)
- Internet search engines and gateways

(e.g. Google, Google Scholar, Intute, PolicyHub)

- Hand-searching journals
- Scanning reference lists
- Professional/personal contacts

The basics of bibliographic databases

Indexers use standardized lists of ‘controlled terms’ to describe key features of papers.

- E.g. ‘Descriptors’ in ERIC
- A ‘thesaurus’ or index and ‘scope notes’ explains to indexers (and you) the meaning of each controlled term

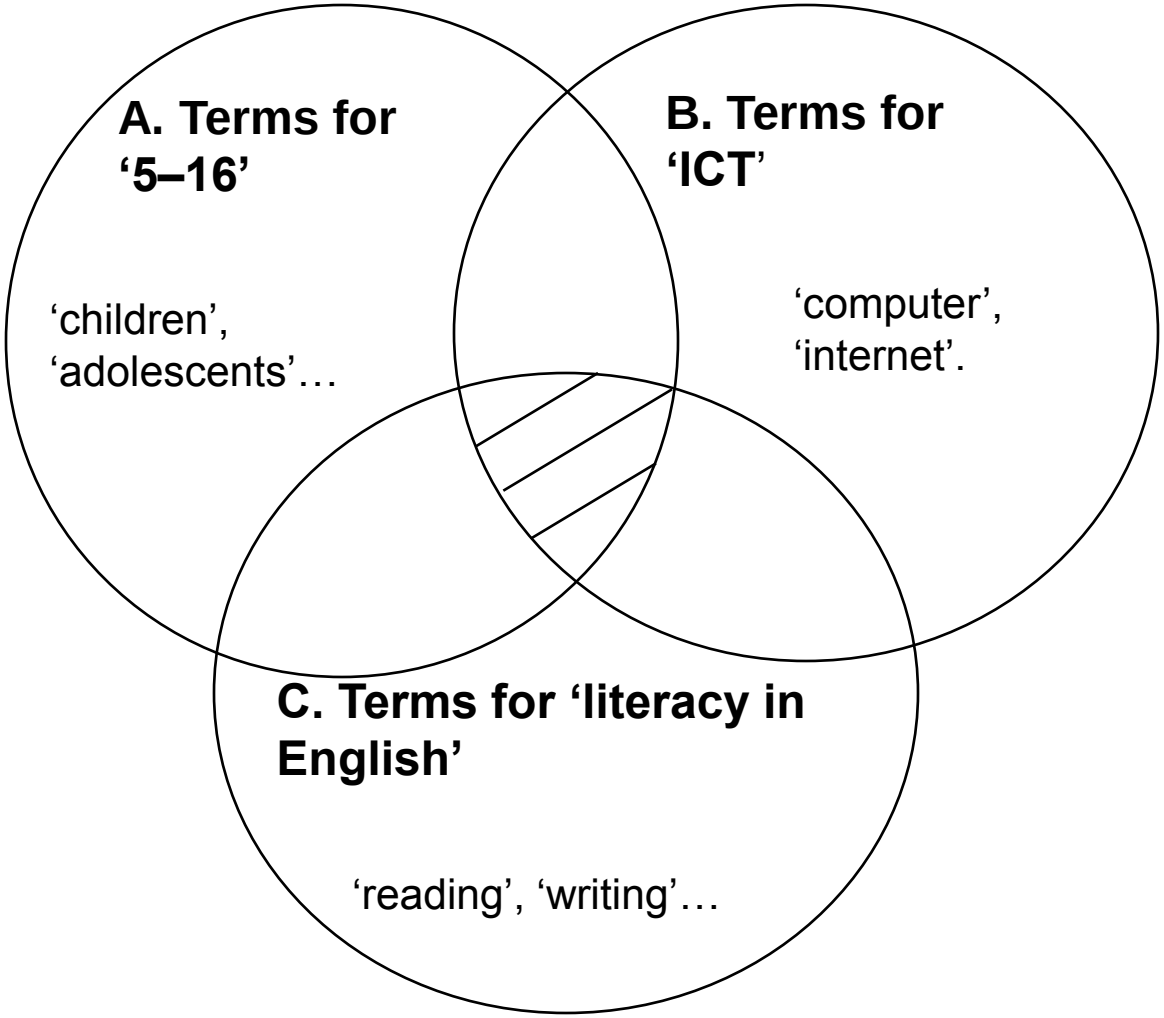
Identify controlled terms for each of the concepts in your review

Build searches using these controlled terms

- Check search terms against controlled terms for known relevant studies

Supplement searches for controlled terms with searches for terms in titles and abstracts (‘free-text’)

‘What is the impact of information communication technology (ICT) on 5-16 year olds’ literacy in English?’*- Concepts and search terms



*Andrews R et al (2002) A systematic review of the impact of networked ICT on 5–16 year olds’ literacy in English. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Do the reviewers want these?

Cleary M, *English* G. The small schools movement: implications for health education. Journal of School Health; v75 n7 p243 Sep 2005

Cerezo, Fuensanta; Ato, Manuel. Bullying in Spanish and *English* Pupils: A Sociometric Perspective Using the BULL-S Questionnaire. Educational Psychology; v25 n4 p353-367 Aug 2005

What is screening?

The process of deciding which studies found in your search meet the review's inclusion criteria

Can be one-stage or two-stage:

- One-stage – full papers are used
- Two-stage

  i. titles and abstracts used first, then full paper only if study still looks relevant (sometimes authors contacted also)

  ii. Important to be over-inclusive initially, then apply criteria more strictly to full papers

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

6. Coding, quality appraisal

Form review team (involve ‘users’)

↓

Formulate review question, conceptual framework and inclusion criteria

↓

Search strategy + screening to identify relevant studies

↓

Describe studies

↓

Assess study quality (and relevance)

↓

Synthesise findings

↓

Communicate and engage

↓

Map

Synthesis

May be narrower than map

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Describing studies

How is this done? What aspects of studies need to be described?

Describing studies for systematic maps

Describing studies for systematic research synthesis

Demo of EPPI-Reviewer

The process of describing studies

Individual research reports are studied systematically in order to provide data for a systematic map and/or synthesis

Labels for the process differ:

- ‘coding’ used here as generic term for describing studies (with both codes and text);
- ‘keywording’ – coding for systematic maps describing what research has been undertaken;
- ‘data extraction’ – coding for syntheses of study findings.

Procedures to promote coding quality, e.g. coding guidance (guidelines), piloting of guidelines, double coding with resolution of differences (standard in coding for synthesis).

Different products require different data

Basic data to describe what research has been done and how

⇒ **produce a map**

Detailed data on methods and results

⇒ **for aggregative synthesis**

⇒ **for configurative synthesis**

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

What is critical appraisal?

An approach used (in systematic reviews and elsewhere) to identify variation in quality and relevance between studies

‘is the process of systematically examining research evidence to assess its validity, results and relevance before using it to inform a decisions’ *

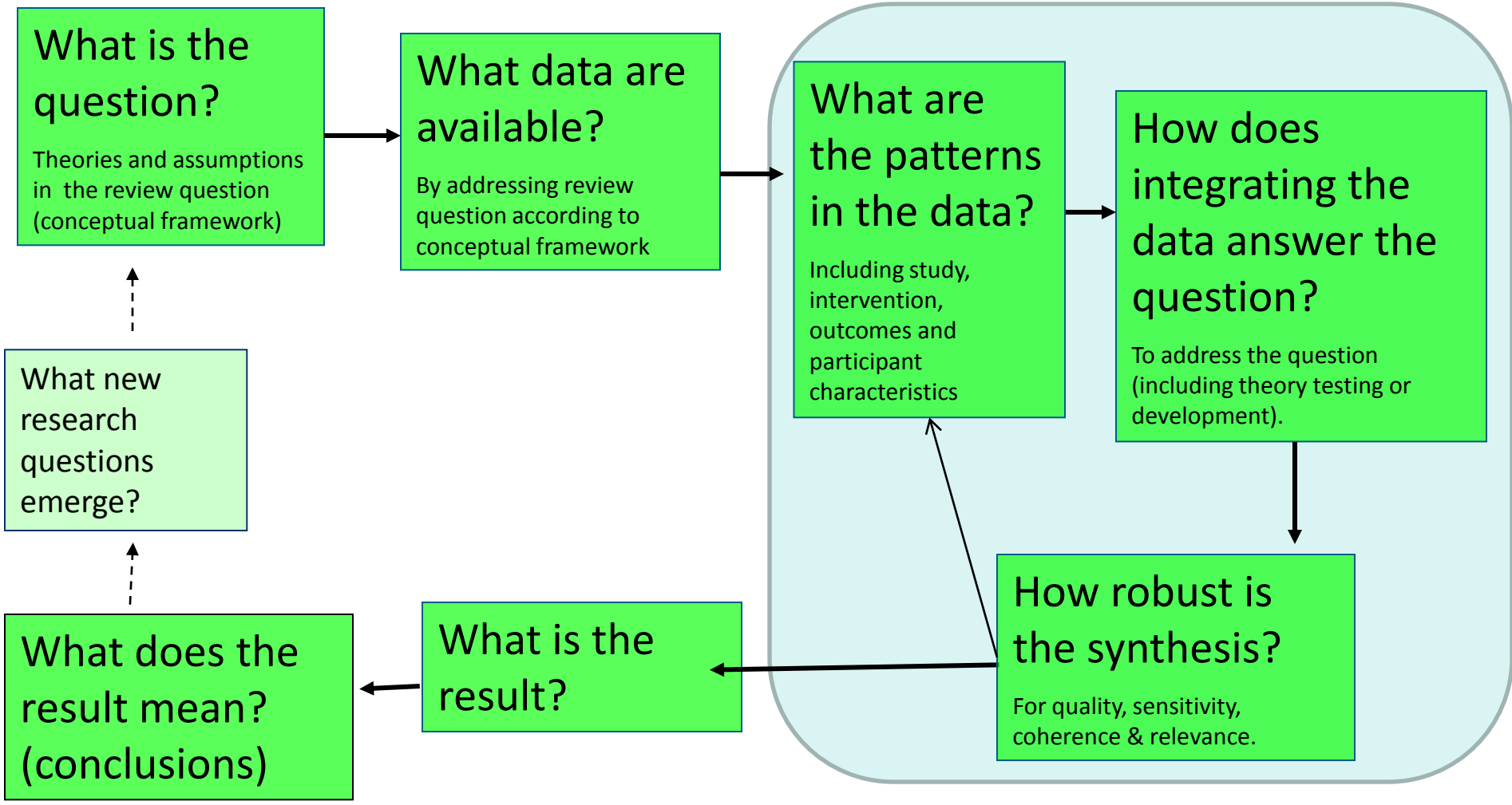
**Hill A, Spittlehouse C. What is critical appraisal?*
London: Hayward Medical Communications, 2006.

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

The three dimensions of critical appraisal for systematic reviews

A. The methodological quality of the study being considered
B. The relevance of that research design for answering the review question
C. The relevance of the study focus for answering the review question

7. Synthesis



Cooper, H.M. (1982) Scientific Guidelines for Conducting Integrative Research Reviews *Review Of Educational Research* 52; 291
See also: Popay et al. (2006) *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews.* Lancaster: Institute for Health Research, Lancaster University.
http://www.lancs.ac.uk/fass/projects/nssr/research.htm

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Different questions, different research, data and synthesis

Aggregative synthesis

- ‘Adding or averaging’ empirical data
- Often ‘quantitative’ primary research
- Methods of synthesis reflect the primary research
- For example, a priori methods
- Conceptual work in advance; e.g. Theory of change

Aggregative example: Is this intervention effective?

- Inclusion criteria: experimental trials
- Search strategy: exhaustive
- Data: empirical (quantitative experimental data)
- Quality and relevance appraisal: avoid hidden bias
- Synthesis: aggregative (e.g. statistical empirical meta analysis)

Two challenges in meta-analysis

Calculating the ‘effect size’ for each study

Combining the individual effect sizes from each study in an appropriate way

Selecting outcomes (1)

As always, try to pre-specify

Sometimes papers present the same outcome in different ways

Data extraction needs to be:

- Consistent
- Unbiased (which outcome to choose?)

Selecting outcomes (2)

Sometimes papers will present outcomes of sub-groups – e.g. age

The correct way of combining these outcomes, so that you have one outcome per study, is to do a ‘mini-meta-analysis’

Important to check that the same people are not appearing more than once in any one analysis

Different ways of presenting an intervention's effect

Often called 'effect sizes'

Binary (or dichotomous) outcomes

- Odds ratios, risk ratios, etc
- Calculated from a 2 x 2 table

Continuous outcomes

- Mean, and standardised mean differences
- Calculated from, e.g. mean and standard deviation

Time to event

- Odds / hazard ratio

Basics and challenges of meta-analysis

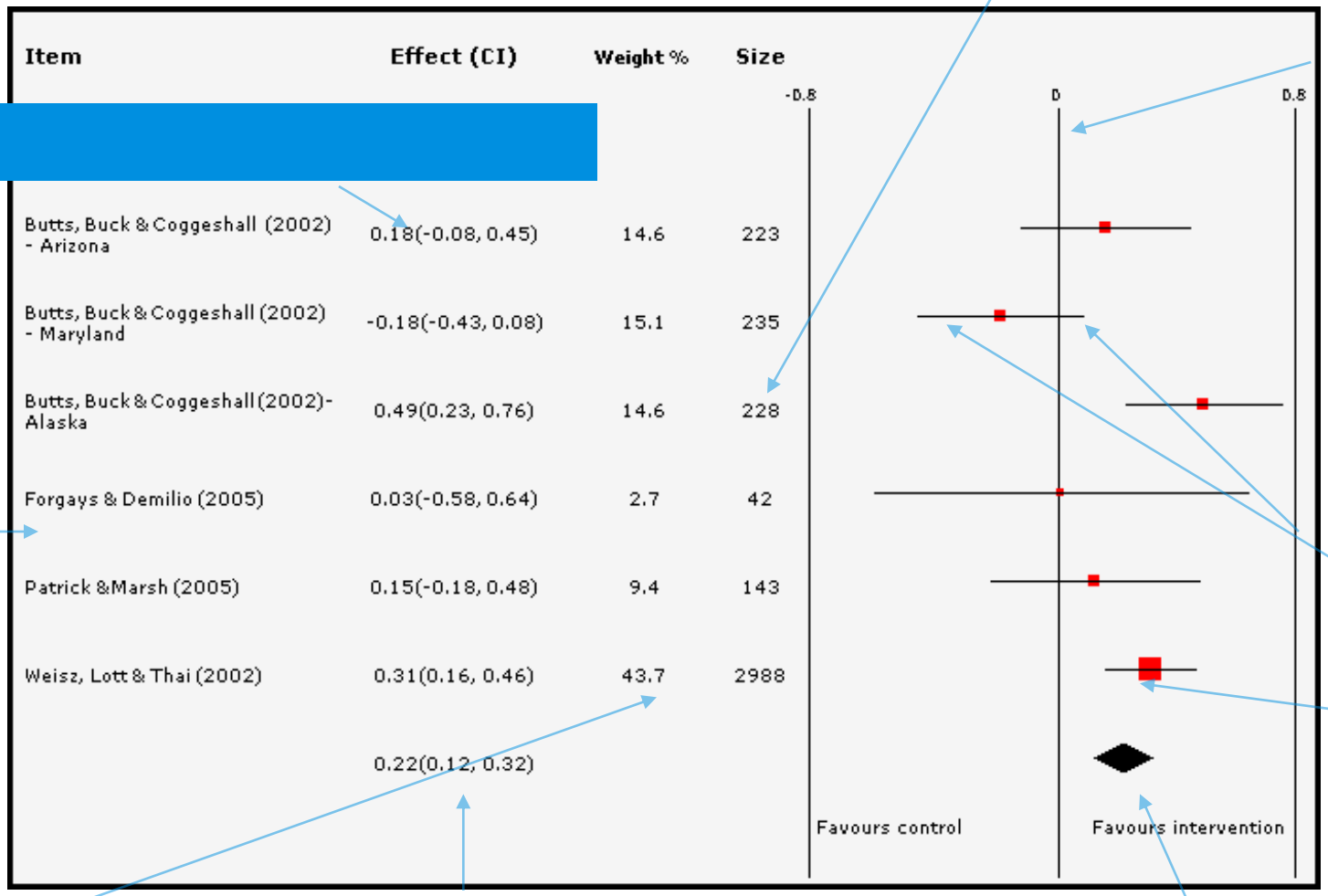
Need to specify methods in protocol

- Measures / outcomes
- Imputation
- Models
- Subgroup analyses
- Sensitivity analyses
- Further analyses (e.g. meta-regression)
- Avoid data dredging

When is it appropriate?

How to deal with heterogeneity

Anatomy of a forest plot



N in study

Line of no effect

Studies

C.I

Study effect size

Weighting of study in meta-analysis

Pooled effect size

Pooled effect size

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Heterogeneity

Studies are usually different to one another

Key question to address is “how different”?

- Different types of people, methods of recruitment, settings, interventions, comparisons, measurement of outcomes
- Chance

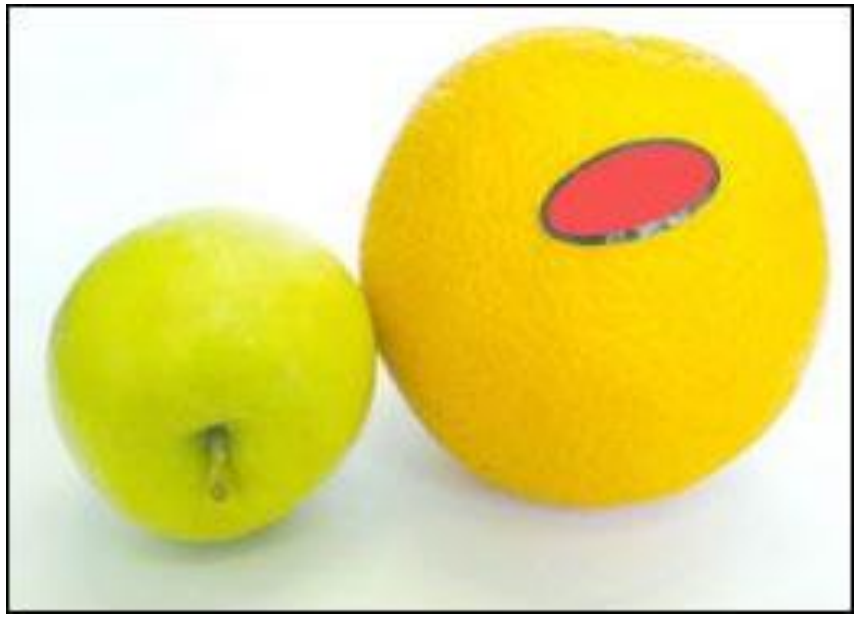
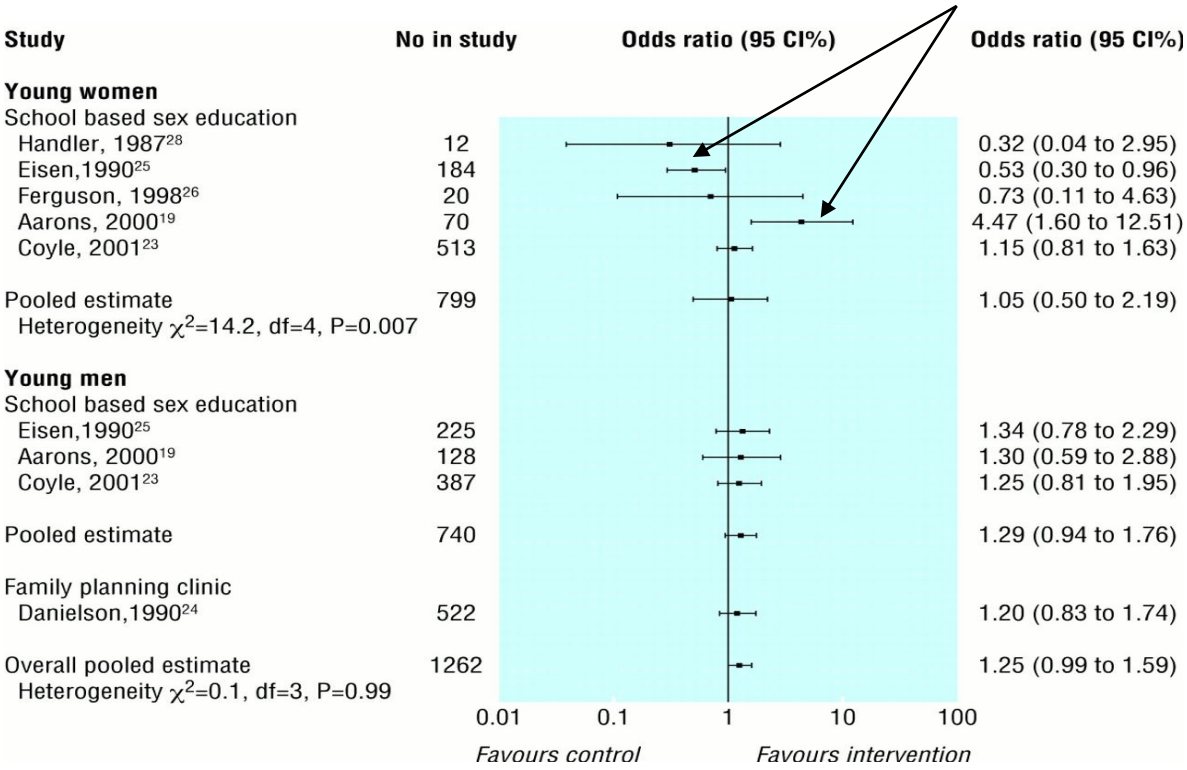


Image source: www.sxc.hu

Heterogeneity (eyeball)

Confidence intervals do not overlap



EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Heterogeneity (statistical tests)

Q statistic tells us if we have statistical heterogeneity

- The null hypothesis is that the underlying effect is the same in each study and the observed effects differ because of sampling variation only
- It has low power with small numbers of studies
- It can detect insignificant heterogeneity with large numbers of studies

Proportion of total variability attributed to between-study heterogeneity (I^2)

- Quantifies the amount of heterogeneity
  - How much is too much ?

How much heterogeneity is too much?

First consideration (reviewer judgement): whether the samples, interventions and endpoints or outcomes measured in the individual studies are sufficiently similar that combining them will provide a valid estimate of the benefit or harm of intervention X

Second consideration (statistical test): whether the results of the individual studies are sufficiently similar (i.e. are not too statistically heterogeneous) for the valid application of the statistical process of meta-analysis

The degree to which a given amount of heterogeneity is acceptable will depend on the type of question being addressed by the review

- Only very general questions can be answered if there is a large amount of heterogeneity

Two main statistical models for meta analysis

Fixed effect model

Random effects models

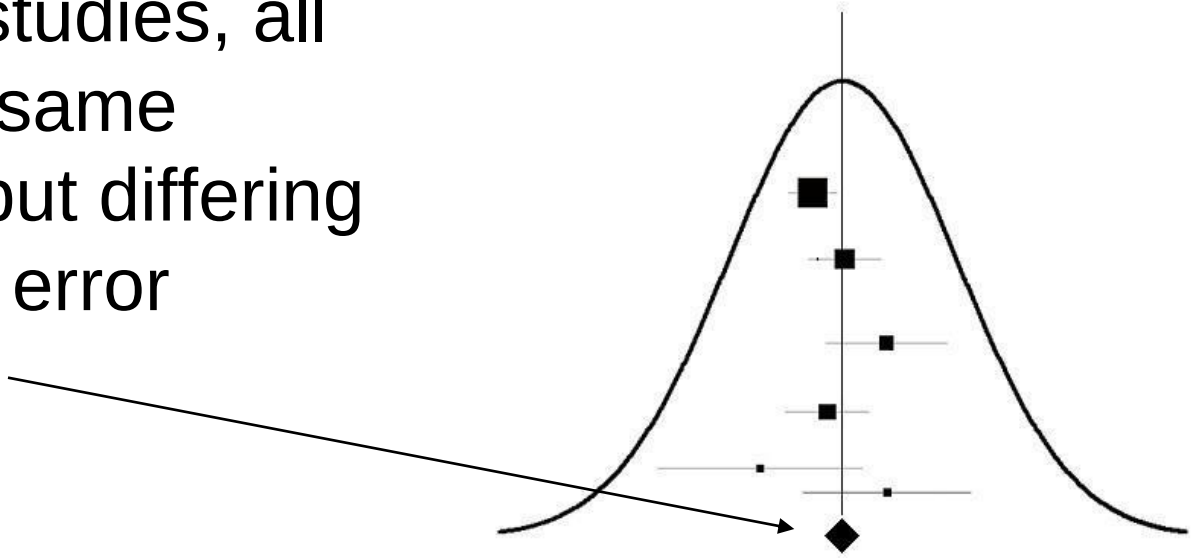
Fixed effect model

Fixed effect meta-analysis assumes that each study is trying to estimate the same *true* effect size. Differences are due to chance and larger studies given more weight.

- N.B. Fixed effect: one effect

Fixed effect model

Each study is seen as being a sample from a distribution of studies, all estimating the same overall effect, but differing due to random error



EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Random effects model

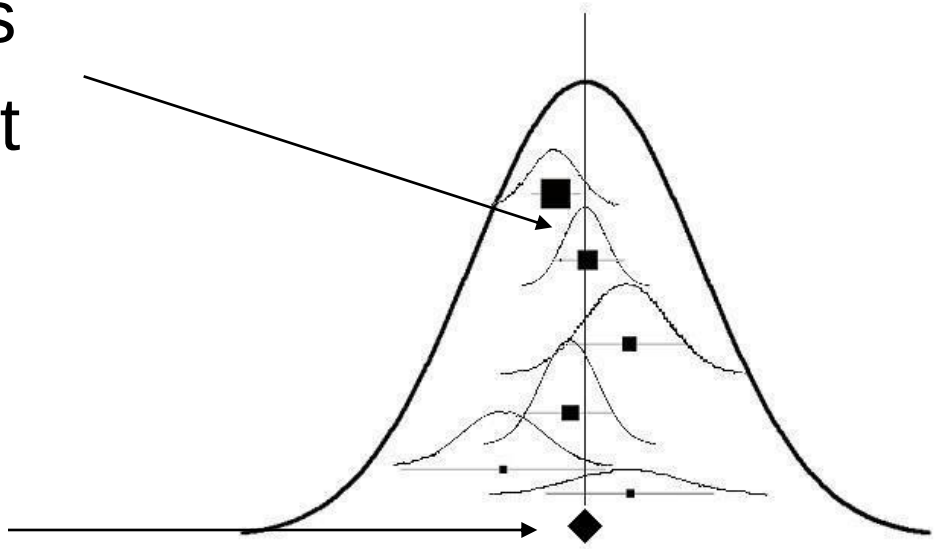
Random effects meta-analysis assumes that each study has a *different* true effect that each is trying to estimate. Differences are expected and incorporated into the model giving more weight to smaller studies than in fixed effect meta-analysis.

- N.B Random effects: plural, many effects

Random effects model

Each study is seen as representing the mean of a distribution of studies

There is still a resultant effect size



EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Publication bias

What it is

Why it is important

- statistical model

How to detect it

What to do about it

What is publication bias?

The *differential* inclusion of studies in a synthesis

- (i.e. the missing studies are not a random sample of all possible studies, but a particular type or sub-set)

Not all studies:

- Are written up by busy researchers
- Are submitted
- Are accepted for publication
- Are identified for inclusion in reviews

Sources of publication bias

Decisions by authors regarding the desirability and chances of publication

You're more likely to be published if you:

- Write in English
- Are famous
- Have exciting results to report

Similar biases operate regarding the chances of a study being found for inclusion in a systematic review

How to detect publication bias

Funnel plots

Statistical tests

- Begg’s
- Egger’s

Sub-group analyses

What to do about publication bias

Trim and fill

- estimates missing studies and their effect sizes

Sensitivity analysis

Do not combine studies statistically

Sensitivity analysis

Checks for robustness in the analysis

- (Something that happens throughout the review)

In a meta-analysis, checks are usually made for:

- Establishing the degree to which the analysis depends on small numbers of studies
- Examining differential impact of other factors – e.g. study quality

Meta-regression

Estimates the extent to which covariates (e.g. age) can explain between study heterogeneity by extending random effects meta-analysis

If covariate is not associated with heterogeneity then it will not be significant in the regression

Different questions, different research, data and synthesis

Configurative synthesis

- Arranging data
- Often conceptual
- Often ‘qualitative’ primary research
- Methods of synthesis reflect the primary research; for example, iterative methods
- Interpretative/conceptual work within the review

Configurative example: How can we understand this?

- Inclusion criteria: ethnography
- Search strategy: ‘purposive’
- Data: conceptual
- Quality and relevance appraisal:
- Synthesis: configure (e.g. conceptual synthesis such as thematic synthesis or meta ethnography)

Meta ethnography example

Britten *et al* (2002) Using meta ethnography to synthesise qualitative research: a worked example. J Health Serv Res Policy Vol 7 No 4: 209-215

Products: new interpretative constructions, ‘translating the studies into one another’

- Transferring ideas, concepts and metaphors across studies;
- Data: interpretations and explanations of original authors;
- Looking for reciprocal and refutational studies and lines of argument;
- Role of quality and sampling varies.

Exact methods vary amongst the small number of studies reported so far

This review was about…

How do the perceived meanings of medicines affect patients' medicine taking behaviour and communication with health professionals?

Was not concerned with searching or assessing quality – it is an example of a method of synthesis.

Steps taken in the synthesis

The researchers looked across the papers for common and recurring concepts.

Used Schutz's notion of 'first- and second-order constructs.

- First order: everyday understandings of ordinary people
- Second order: constructs of the social sciences

'Third order interpretations' were derived from a 'line of argument' based on 1st and 2nd order concepts.

Translating studies into one another

Table 1 Completed grid: lay meanings of medicines

Methods and concepts	Donovan and Blake[14]	Morgan[16]	Britten[15]
Sample	54 patients with suspected inflammatory arthropathy	60 white and Afro-Caribbean patients treated for hypertension for at least one year	30 patients, attenders and non-attenders
Data collection	Home interviews pre and post consultation; observation of consultations	Home interviews	Home interviews
Setting	Three rheumatology units	15 general practices	Two general practices
Type(s) of medicine	NSAIDs and second-line drugs	Antihypertensive drugs	Unselected
Adherence/compliance	(Patients do not perceive compliance to be an issue)	Stable adherence and problematic adherence	Correct behaviour and routine medicine-taking
Self-regulation	Levels of non-compliance	Leaving off drugs	Preference for not taking drugs
Aversion	Dislike of taking drugs, fear of side-effects, weakness, dependence	Fear of side-effects, addiction, harmful effects of drugs	Aversion to medicines, medicines as harmful
Alternative coping strategies	Range of alternative remedies	Traditional (herbal) remedies	Use of alternative medicine
Sanctions	–	Patients warned by doctors and told severely about the need to take the tablets regularly	–
Selective disclosure	Patients did not tell doctors of altered doses	–	–
Explanation/theory (second-order interpretation)	'Patients carry out a "cost–benefit" analysis of each treatment, weighing up the costs/risks of each treatment against the benefits as they perceive them'	Medicine-taking is influenced by cultural meanings and cultural resources	Patients may not articulate views that they do not perceive to be medically legitimated

Example of conceptual synthesis:

How do the perceived meanings of medicines affect patients' medicine taking behaviour and communication with health professionals?

CONCEPTS described in primary studies	**Second order interpretations**	**Third order interpretations**
Adherence/compliance:correct and routine medicine taking **Self regulation:** problematic adherence; leaving off drugs **Aversion**: eg dislike of taking drugs; harmful side effects **Alternative coping strategies**:eg Traditional remedies	a)Patients conduct cost-benefit analyses: weigh up risks vs benefits b) Medicine taking influenced by cultural meanings and resources	c)Self regulation includes the use of alternative coping strategies
Sanctions: Drs talk severely about need to take tablets regularly; coercion from significant others, fear of coercion	d) Self regulation is… inhibited by… the threat of social and professional sanctions	e) Self regulation flourishes if sanctions are not severe
Selective disclosure: patients don't tell Dr of altered doses; manage info to psychiatrists **Adapted from Britten et al., 2003**	f) Patients may not articulate views which they do not perceive as medically legitimated	g) Alternative coping strategies are not seen by patients as medically legitimate h) Fear of sanctions and guilt produce selective disclosure

Conclusions about the method

Methods developed for synthesising ethnographic research are applicable to the synthesis of other kinds of (non-ethnographic) study.

By developing ‘third order interpretations’, this method goes beyond the primary studies and develops new concepts.

The hypotheses it produces could be tested by other researchers.

- Third example shows how hypotheses generated from a synthesis of qualitative studies can be tested within the same systematic review.

Example: Children and healthy eating: a systematic review of barriers and facilitators

What is known about the barriers to, and facilitators of, healthy eating amongst children aged 4 to 10 years?

Full review available online:
http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/hp/reports/healthy_eating02/healthy_eating02.htm

Thomas J, Harden A (2008) Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8:45 doi:10.1186/1471-2288-8-45 (http://www.biomedcentral.com/1471-2288/8/45)

Thomas J, Harden A, Oakley A, Oliver S, Sutcliffe K, Rees R, Brunton G, Kavanagh J (2004) Integrating qualitative research with trials in systematic reviews: an example from public health. *British Medical Journal* **328**: 1010-1012. (http://www.bmj.com/cgi/content/full/328/7446/1010)

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Underlying principles of our approach

Policy and practice concerns often precede, or go beyond, questions of effectiveness.

Different types of questions require different combinations of study types to be included.

Different combinations of study types demand different types of synthesis methods.

The review question should drive the review methods.

One review with three syntheses

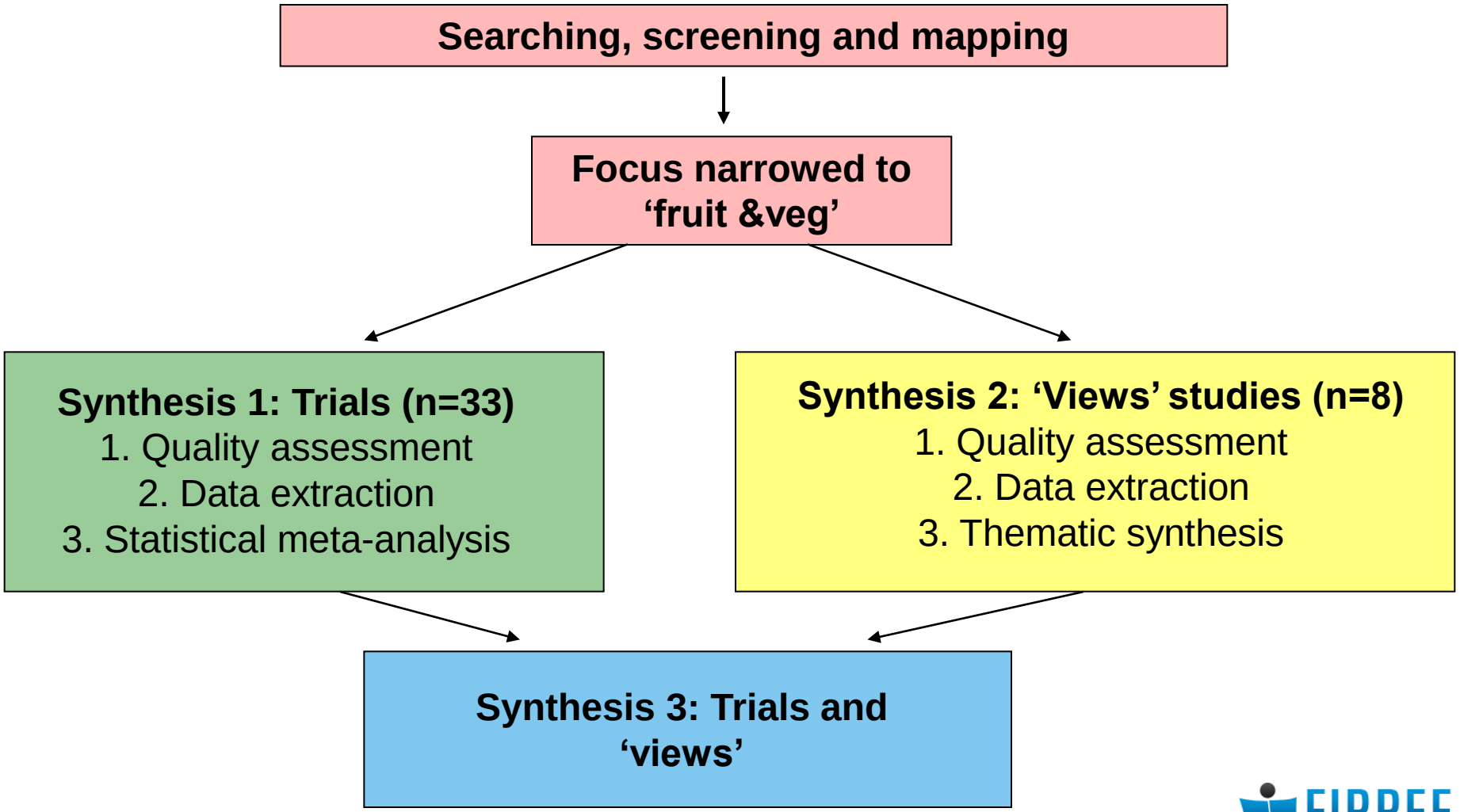
1. Quantitative methods used to conduct a meta-analysis of data from trials.

2. Qualitative methods used to synthesise textual data from ‘views’ studies (aided by NVivo).

3. Both qualitative and quantitative methods used to combine experimental studies of effectiveness (trials) with studies of people’s views (descriptive).

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

REVIEW PROCESS



EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Thematic analysis: 1

Data extraction: results from primary studies

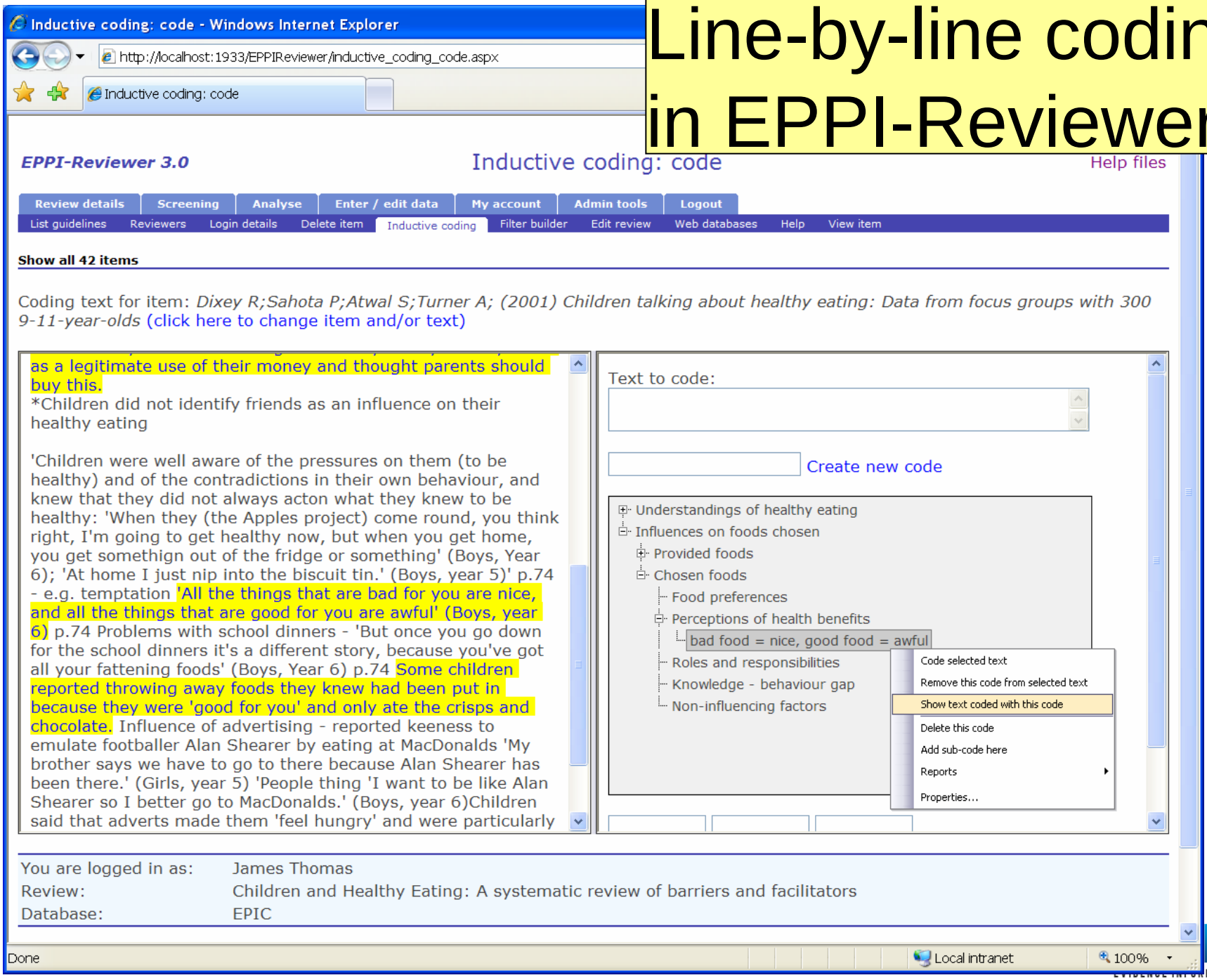
Coded the themes described in our data extraction (e.g. ‘bad food = nice, good food = awful’)

- 36 initial *descriptive codes*

Looked for similarities and differences among *descriptive* codes in order to group them

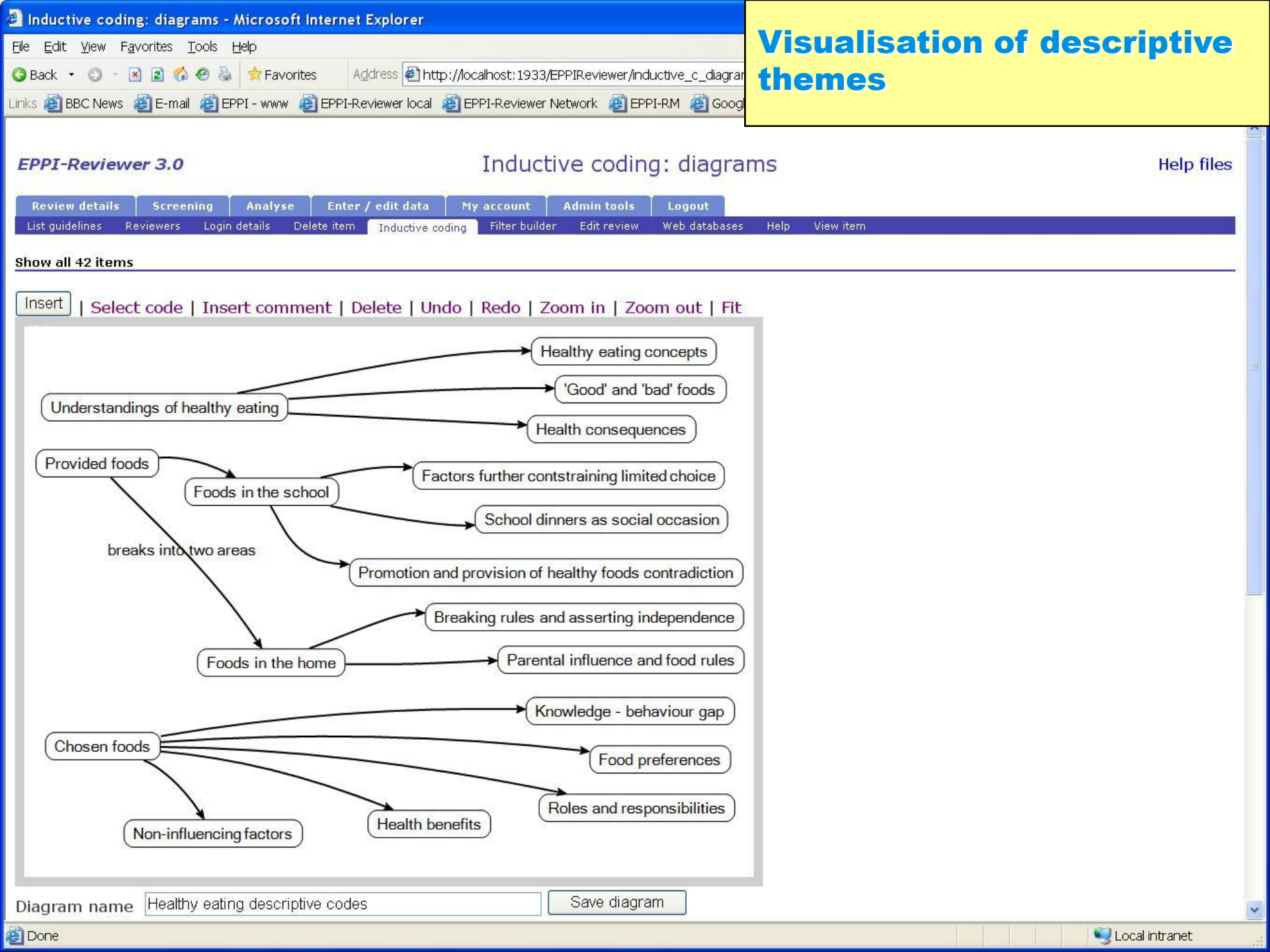
- 13 *descriptive themes* (e.g. ‘Perceptions of health benefits’)

Line-by-line coding in EPPI-Reviewer



EPPI
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

Visualisation of descriptive themes



Sub-questions for Synthesis 2: driven by main review question

What are children's perceptions of and attitudes towards healthy eating? What does healthy eating mean to children?

What do children think stops them from eating healthily?

What do children think helps them to eat healthily?

What ideas do children have for what could or should be done to promote their healthy eating?

Analytical themes

1) Children don't see it as their role to be interested in health.

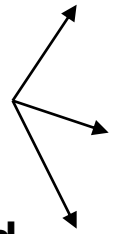
2) Children do not see future health consequences as personally relevant or credible.

3) Fruit, vegetables and confectionary have very different meanings for children.

4) Children actively seek ways to exercise their own choices with regard to foods.

5) Children value eating as a social occasion.

6) Children recognise contradiction between what is promoted and what is provided



Brand fruit and vegetables as 'tasty' rather than 'healthy'.

Reduce health emphasis of messages

Do not promote fruit and vegetables in the same way within the same intervention.

Create situations for children to have ownership over their food choices.

Ensure messages promoting fruit and vegetables are supported by appropriate access to fruit and vegetables

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

This method of synthesis across study types:

preserves the integrity of the findings of the different types of studies

allows us to integrate ‘quantitative’ estimates of benefit and harm with ‘qualitative’ understanding from people’s lives.

allows the exploration of heterogeneity in ways in which it would be difficult to imagine in advance.

Different questions, different research, data and synthesis

Example 3: Realist synthesis

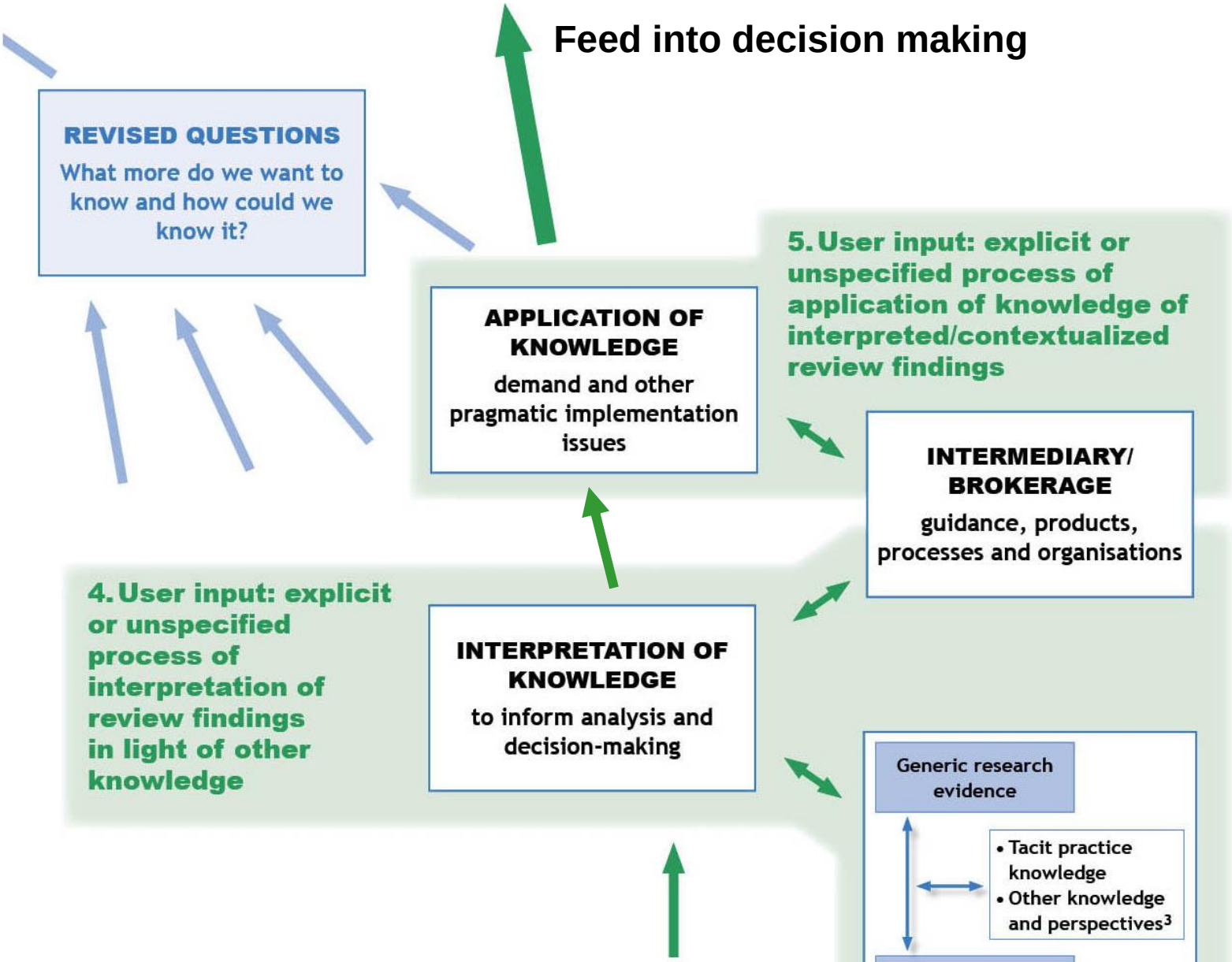
- Study evidence in support of middle level theories across social policy areas
- Unpack model (configure)
- Test evidence for model assumptions/ causal pathways
- Inclusion criteria: empirical studies
- Search strategy: purposive /iterative
- Data: empirical
- Quality and relevance appraisal:
- Synthesis: aggregate (e.g. Realist Synthesis)

8. Using research

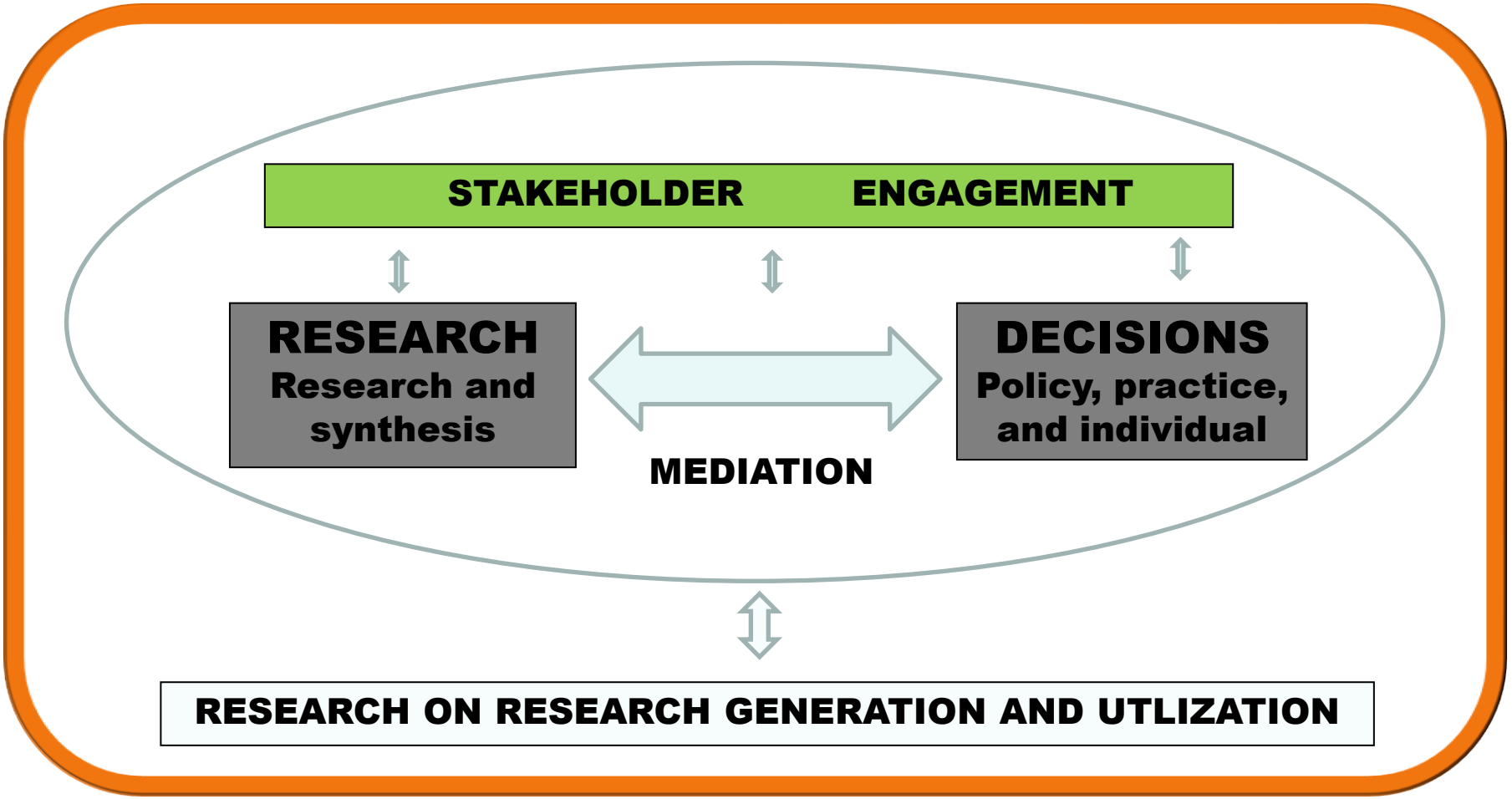
Interpretation and implementation

A system approach

EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE


Feed into decision making
REVISED QUESTIONS
What more do we want to know and how could we know it?
APPLICATION OF KNOWLEDGE
demand and other pragmatic implementation issues
5. User input: explicit or unspecified process of application of knowledge of interpreted/contextualized review findings
INTERMEDIARY/ BROKERAGE
guidance, products, processes and organisations
4. User input: explicit or unspecified process of interpretation of review findings in light of other knowledge
INTERPRETATION OF KNOWLEDGE
to inform analysis and decision-making
Generic research evidence
• Tacit practice knowledge
• Other knowledge and perspectives3
Context-sensitive research evidence
Communication of review findings

Evidence informed policy and practice systems



EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

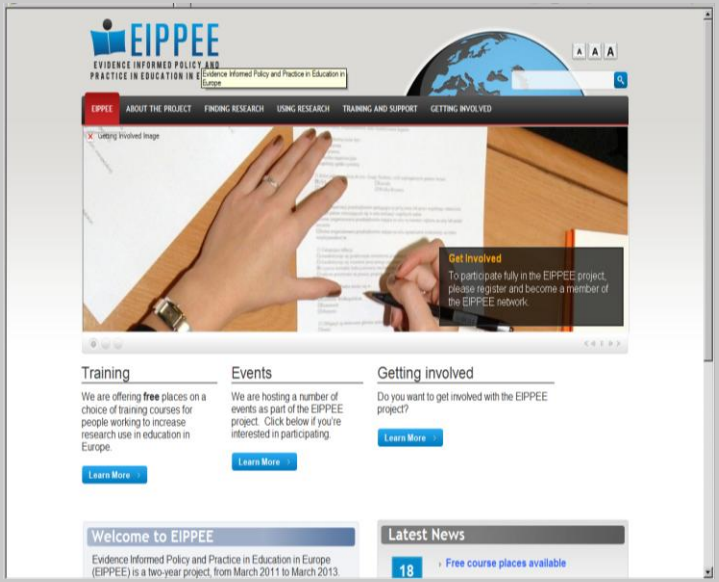
Review of whole course

Whole group activity and discussion

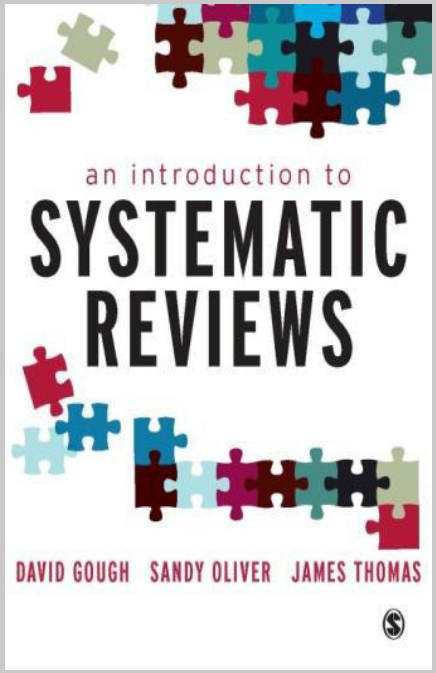
EIPPEE
EVIDENCE INFORMED POLICY AND PRACTICE IN EDUCATION IN EUROPE

EIPPEE

http://www.eippee.eu



An introduction to systematic reviews: Sage Publications Ltd



Evidence & Policy

EPPI-Centre, Social Science Research Unit
Institute of Education, University of London
Email: d.gough@ioe.ac.uk
Web: http://eppi.ioe.ac.uk/

Gough D, Thomas J, Oliver S (In press). Clarifying differences between review designs and methods. *Systematic Reviews Journal*. http://www.systematicreviewsjournal.com