

Supporting policy and practice: Why systematic reviews matter

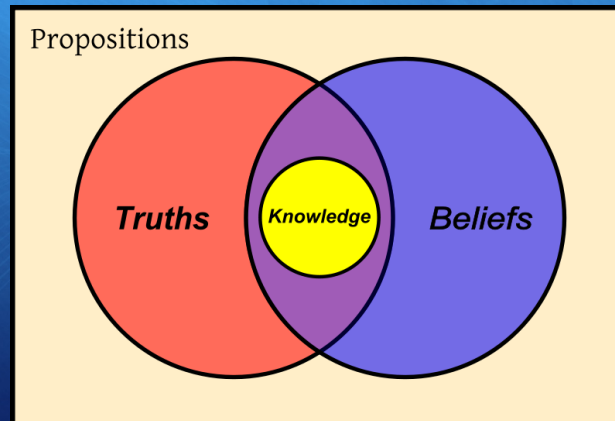
Julia H. Littell
Graduate School of Social Work and Social Research
Bryn Mawr College

9th annual Campbell Collaboration Colloquium
Oslo, Norway
18 May 2009

Why we care about evidence

- Good intentions are not enough
 - Can have unintended, even harmful effects
- Building knowledge to inform policy and practice

A Classical Definition of Knowledge

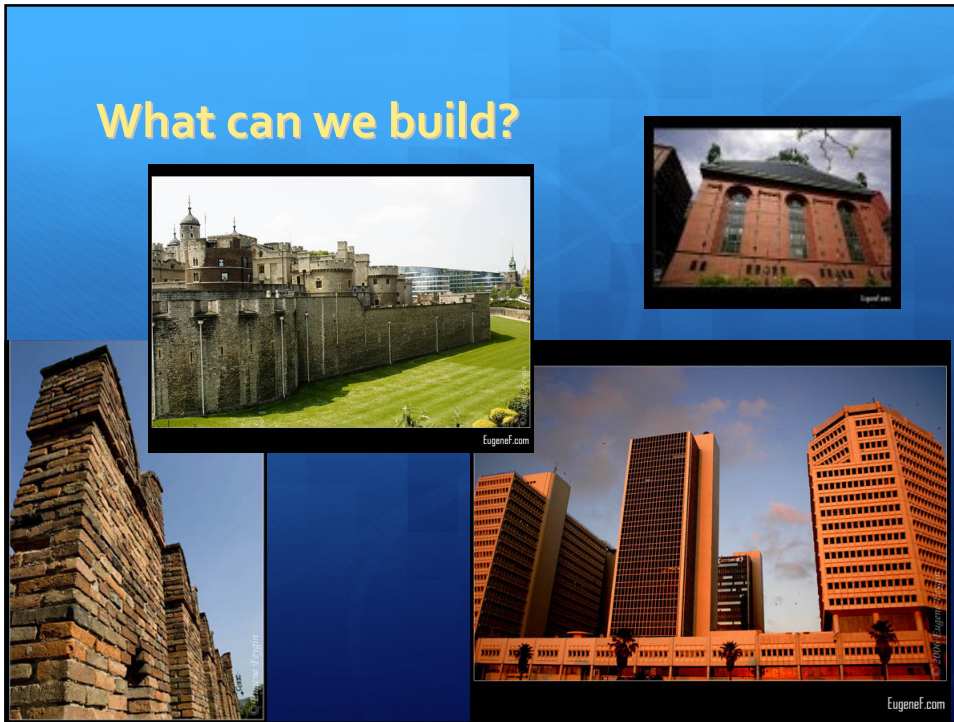


Szczepan, 1990

Building knowledge

- In the information age there is great need for syntheses/summaries of relevant evidence
- “What can you build with thousands of bricks?” (Lipsey, 1997)
 - Combining results of studies “can lead to the building of better intervention theories needed to guide effective program development and evaluation design”

What can we build?



How do we build evidence?

- What are our blueprints?
- What are the raw materials?
- What methods are used to combine results across studies?

Blueprints

- Plans for review
- Reviews vary in amount of planning, transparency, rigor
 - Traditional, narrative reviews
 - Systematic reviews
 - Cochrane/Campbell systematic reviews

Traditional reviews

- Convenience samples of published studies
- Narrative description of studies
- Cognitive algebra or “vote counting” to synthesize results
 - Relies on statistical significance in primary studies, which may be “underpowered” (too small or too weak to detect effects)
- Vulnerable to bias

Bias

- Bias is a distortion of information
- Bias can lead to the wrong conclusions
- Several known sources of bias affect research reviews

Publication bias

- Studies with statistically significant results are more likely to be published than studies with non-significant (null) results
 - (Begg, 1994; Dickersin, 2005; Scherer et al., 2004)
- Sources of publication bias are complex

Outcome reporting bias

- Within studies with mixed results, significant results are more likely to be
 - reported (mentioned at all)
 - fully reported (i.e., data provided)
 - (Chan et al., 2004, 2005; Williamson et al., 2006)

Confirmation bias

- Tendency to seek and accept information that confirms prior expectations (hypotheses) and ignore evidence to the contrary
 - (Bacon 1621/1960, Watson 1960, 1968; Mahoney, 1977; Fugelsang et al., 1994; Nickerson, 1998; Schrag, 1999)
- Murphy's Law of Research
 - If you look hard enough, you will find evidence that supports your theory

Summary: Bias

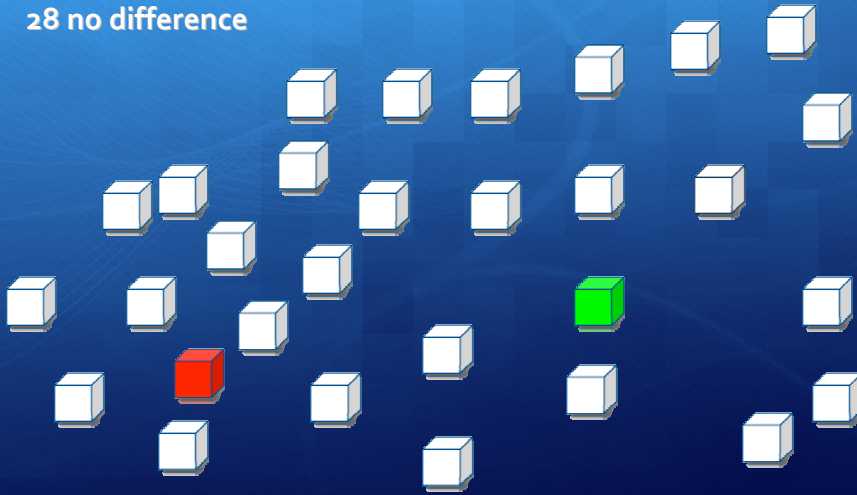
- Publication and reporting biases appear to be cumulative (Altman, 2006)
 - Tend to inflate estimates of effects
- Problems are ubiquitous, but often ignored

How do we minimize bias?

- Systematic reviews are a form of research
 - Aim to minimize bias at each step
 - Follow pre-determined plan (protocol)
 - Use transparent (well-documented, replicable) procedures to locate, analyze, and synthesize results of previous studies

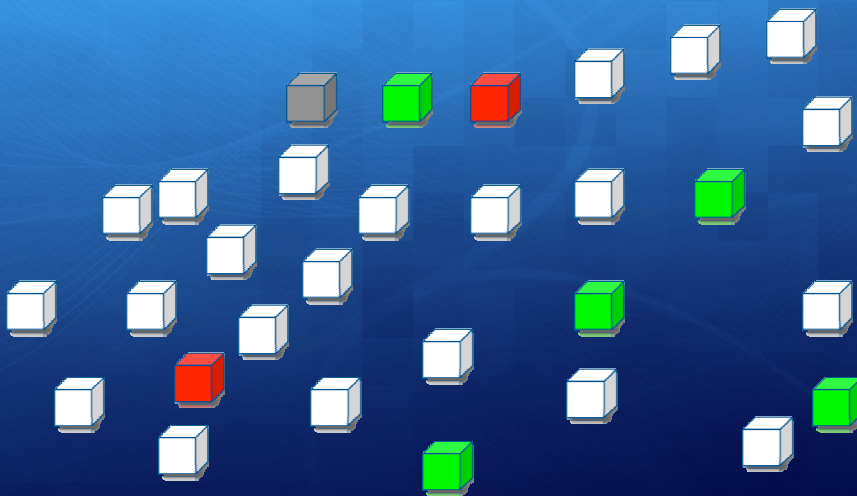
Results expected by chance (30 outcomes)

1 favor PT
1 favor MST
28 no difference



Results obtained (Brunk et al., 1987)

Parent Training vs Multisystemic Therapy



Results obtained (Brunk et al., 1987)

Parent Training vs Multisystemic Therapy

Outcome data



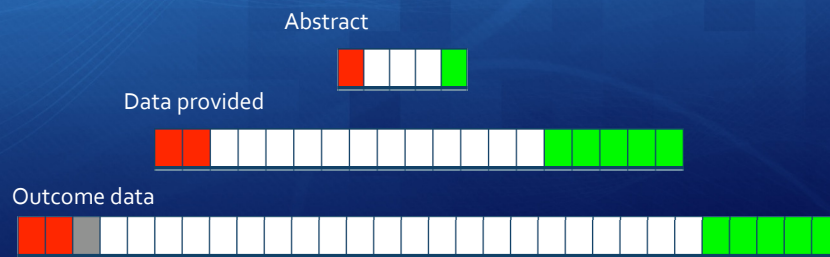
What did the investigators make of these results?

Outcome data



What did the investigators build?

A balanced report
(with some missing data)



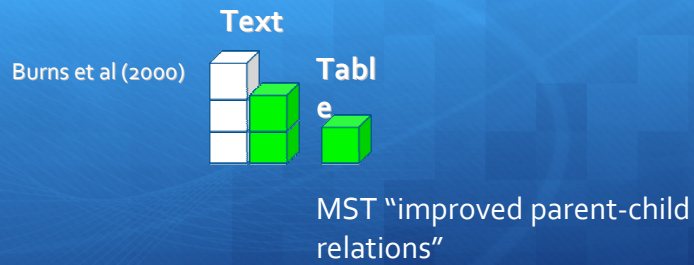
What did published reviews build?



"Parents in both groups reported decreases in psychiatric [symptoms] and reduced overall stress....both groups demonstrated decreases in the severity of the identified problems....[MST] improved [parent-child] interactions, implying a decreased risk for maltreatment of children in the MST condition" (p. 293).



What did published reviews build?



Outcome data



What did published reviews build?



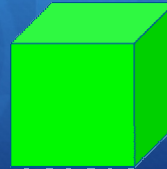
Outcome data



What did published reviews build?

Curtis et al (2004)

average ES across all outcomes
 $d = 1.32$ (sd=.65)



















Average *MST case* outperformed
 90% of *PT cases* on average
 across all outcomes ??

Outcome data



Summary: Published reviews describing Brunk et al.

	Text	Table	
Burns et al (2000)	X		
		X	
Corcoran (2000)	X		
Curtis et al (2004)		X	
Henggeler, Lee (2003)		X	
Henggeler et al (1998)	X		
Henggeler et al (2002)	X		
		X	
Henggeler, Sh (2003)	X		
Hoagwood et al (2001)	X		
Kazdin (1998)	X		
Kazdin & Weisz (1998)	X		
Pushak (2002)	X		
Shownwald, Ro (2002)	X		
Swenson, Heng (2003)	X		
Swenson et al (1998)	X		

Outcome data



Summary: Published reviews describing Brunk et al.

- Most reviews used a single phrase to characterize results of this study, highlighting advantages of one group (MST)
- Losing information on
 - relative advantages and disadvantages of different approaches
 - equivalence of different interventions

Outcome data



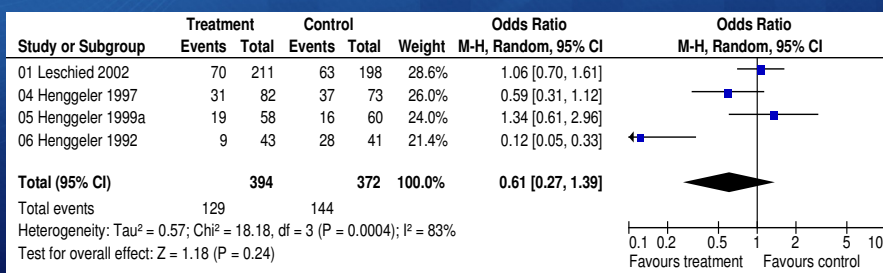
Research reviews include multiple studies

- 86+ reviews of research on effects of MST published after 1996
 - more reviews than studies!
 - Assessed 66 reviews
 - Most were "lite" reviews (relied on other reviews)
- 37 reviews cited one or more primary studies
 - Analysis of review methods (Littell, 2008)
 - Most were traditional narrative summaries of convenience samples of published reports
 - Most conclude that MST "works" (is more effective than alternatives)

- "The efficacy of MST...was established through three randomized clinical trials with delinquents, and effectiveness through the transfer of MST to other clinical populations ... and to multiple organizational settings" (Burns et al., 2000, 309).
- "MST is unique... providing multiple replications across problems, therapists, and settings.... This shows that the treatment and methods of decision making can be extended and that the treatment effects are reliable.... MST is quite promising given the quality of evidence and consistency of the outcomes" (Kazdin & Weisz, 1998, 28).

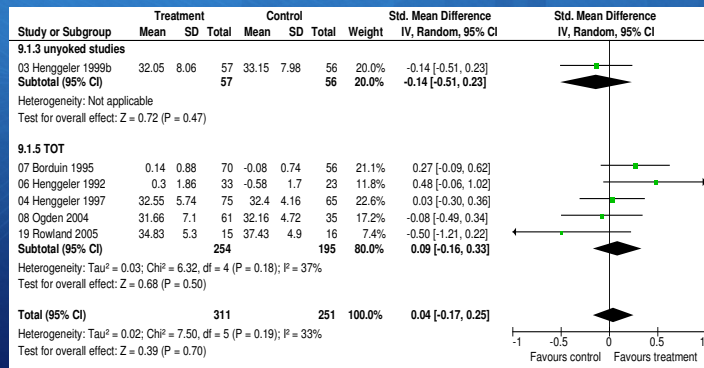
Campbell/Cochrane review

- Effects of Multisystemic Therapy (Littell, Popa, Forsythe, 2005)
- Are not consistent within or across studies
 - Few studies
 - All studies have mixed results across outcomes (except those that have null results)



Systematic review

- Family cohesion



Issues

- The science and practice of research synthesis are not well-connected (Littell, 2005, 2008)
 - “Science is supposed to be cumulative, but scientists only rarely cumulate evidence scientifically” (Chalmers, Hedges & Cooper, 2002)
- Scientists urge policy makers and practitioners to pay attention to “scientific” evidence... shouldn’t scientists do the same?
 - Shouldn’t evidence be cumulated scientifically?

The Scope of Systematic Reviews

- Specific, narrow questions
 - Few studies, little variation
 - Restricted domain for generalization
- Broad, global questions
 - More studies, more variation
 - Global, wide domain of generalization
- **Lipsey (2008)**

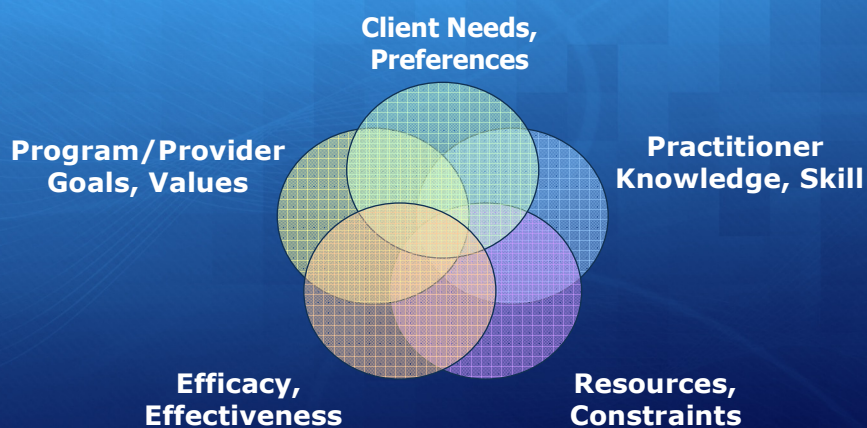
The Scope of Systematic Reviews

- Not limited to questions about effects
 - Can address trends, epidemiology, accuracy of diagnostic and prognostic tests
- Not limited to randomized controlled trials or quantitative data

Why systematic reviews matter

- Systematic reviews can provide more accurate, reliable syntheses of empirical evidence than traditional reviews
 - Ought to be used to inform policy and practice
 - Can be used to synthesis data on many questions (not just questions about effects)
- Better to have few, rigorous systematic reviews than many (86+) nonsystematic reviews

Evidence for practice and policy: Effectiveness in context



Adapted from: Gibbs (2003), Davies (2004)

Thank you!
Enjoy the Colloquium

Contact information:
jlittell@brynmawr.edu