How Do We Know What Works?

The Campbell Collaboration: International Efforts to Synthesize Evidence for Practice

National Association of Deans and Directors (NADD)
Of Schools of Social Work
San Antonio, TX
17 September 2005

Julia H. Littell
jlittell@brynmawr.edu

Portions of this work were funded by the Smith Richardson Foundation, the Swedish Institute for Evidence-Based Social Work Practice (IMS), and the Nordic Campbell Center (Danish National Institute of Social Research)

Where's the Evidence for Evidence-Based Practice (EBP)?

- EBP models: many sources and types of evidence are relevant for practice, including
  - Qualitative, quantitative and anecdotal evidence
  - About consumer needs, values, preferences, and effects of interventions
- This presentation focuses on empirical evidence on *effects* of interventions
  - Not because this is "better" or more important than other evidence
  - Because if we are going to review and summarize empirical evidence of intervention effects, we should do it well.
  - This knowledge is cumulative, changing, incomplete
  - Where is this evidence? How is it synthesized? What do we know? With what certainty? What *don't* we know?
  - To what extent is knowledge of intervention effects based on science vs tradition, authority, and other sources?

Practice of Research Synthesis

Traditional research reviews use

- Convenience samples of published studies
  - Vulnerable to publication bias (Begg, 1994; Rothstein, Sutton & Weinstein, in press)
- Narrative analysis
- Cognitive algebra or “vote counting” to synthesize results
  - Relies on statistical significance in primary outcome studies (may be underpowered)
  - Vulnerable to selection bias, confirmation bias

Practice of Research Synthesis (cont'd)

- Criteria for evaluating treatment effects have been developed by
  - government and professional organizations
  - meta-analysts
- Diverse criteria have been applied to bodies of evidence to determine "what works"
- Results have been used to create lists of "effective" or "model" programs
- These categorizations affect funding decisions

Science of Research Synthesis

- Handbook of Research Synthesis (Cooper & Hedges, 1994)
- Advances in
  - Information retrieval (e.g., Rothstein, Turner, & Lavenberg, 2003)
  - Research designs for causal inference (e.g., Shadish, Cook, & Campbell, 2002; Shadish & Myers, 2003)
  - Meta-analysis (e.g., Becker, Hedges & Pigott, 2003; Lipsey & Wilson, 2001)

Science of Research Synthesis (cont'd)

Systematic Reviews (SRs) treat review process as a form of research

- follow basic steps in research process
- use transparent procedures to minimize bias, including
  - Explicit inclusion/exclusion criteria
  - Systematic strategies for locating all potentially-relevant studies
  - Inter-rater agreement on decisions about text retrieval, study eligibility, and coding
  - Systematic coding and analysis of included studies methods, treatments, samples, outcomes
  - Meta-analysis (when possible) to estimate pooled effect sizes (ES) and moderators of ES

Issues

- Science and practice of research synthesis are not well-connected
  - Lists based on traditional reviews
  - Meta-analyses not based on systematic reviews
  - “Systematic” reviews without meta-analysis
- “Science is supposed to be cumulative, but scientists only rarely cumulate evidence scientifically” (Chalmers, Hedges & Cooper, 2002, p. 12)
- Practitioners are urged to pay attention to “scientific” evidence
  - Shouldn’t scientists do the same?
  - Shouldn’t this evidence be cumulated scientifically?

Two Collaborations

- Bridge science and practice of research synthesis
- International
- Interdisciplinary
- Networks of scholars, policy makers, practitioners, and consumers
- Nonprofit organizations
- Commitment to producing, updating, and disseminating SRs

Cochrane and Campbell Collaborations

- The Cochrane Collaboration (CC) is devoted to cumulating evidence in medical and health sciences. Includes 50 review groups and 10 methods groups (www.cochrane.org).
- The Campbell Collaboration (C2) is devoted to synthesizing evidence about effects of social and behavioral interventions (3 substantive coordinating groups, 6 methods groups, users group, communications group) (www.campbellcollaboration.org).
- CC and C2 relate to each other via overlap in Steering Groups and some subgroups (e.g., Methods, Social Welfare).
- Prominent social work scholars have been involved in C2 since its inception in 1999 (Gambrill, Mullen, Schuerman)

CC and C2 Reviews

- Follow procedures and standards adopted by international, interdisciplinary Steering Groups
- Title registration (declares review team's intent)
- Protocol (plan) for SR is developed in advance
- Protocol and completed SR are vetted by international experts in the substantive area and SR methods (information retrieval, research design, meta-analysis)
- Conflict of interest statements required
- *Not* limited to RCTs, but RCTs are treated separately
  - Glazerman, Levy, and Myers (2002)
- SRs updated every 2-3 years
- Products and commentaries posted on web

An Example: SR of effects of Multisystemic Therapy

- Title registration
  - with joint C2 and CC Developmental, Psychosocial, and Learning Problems Group (Bristol, UK)
- Protocol development
  - vetted by C2 and CC substantive and methodological experts (editors, trial search coordinators, and statisticians)
  - published in Cochrane Library (Issue 2, 2004) and available on C2 website
- Completed review
  - Critiqued by 10 anonymous readers and C2 and CC experts
  - Published in the Cochrane Library (Issue 4, 2005) and available on C2 website

(Related article in April 2005 issue of Children and Youth Services Review, with debate in press)

What is Multisystemic Therapy (MST)?

- Intensive, short-term, family- and community-based intervention for youth and families
  - Originally developed with juvenile delinquents and offenders
  - Extended to youth with other social, emotional, and behavioral problems
- Aims to reduce out-of-home placements, crime and delinquency, youth and family problems
- Intervention in multiple social systems (e.g., family, peers, school, neighborhood)
- Staffed by Master's level therapists (psychologists and social workers)
- Emphasis on
  - adherence to 9 MST "principles" (vs specific techniques)
  - staff training and support

Previous Reviews of MST Outcome Studies

- More reviews than primary outcome studies
  - 82 reviews published after 1996 (not in reports on MST studies)
  - Most are “lite” reviews (based on other reviews)
  - 34 reviews analyzed (the “best” reviews)
- Most reviews looked at MST, as one of several treatments for
  - Conduct disorder and delinquency
  - Child abuse and neglect
  - Serious emotional disorders in youth
- Criteria and methods of 34 reviews vary
  - Most were narrative reviews of convenience samples of published studies
  - Some used keyword searches OR sought unpublished data OR used meta-analysis
- Most conclude that MST “works”

Another Review

- Fully systematic
- Different methods, different results
- Review questions
  - What are the impacts of MST on out-of-home placements? indicators of youth and family well-being?
  - Are results consistent across studies? If not, what factors might account for inconsistencies?

MST Review: Inclusion Criteria

- Randomized controlled trials (RCTs) only
- Licensed MST intervention
- Youth with social, emotional, and/or behavioral problems (not medical conditions)
- Any comparison condition (usual services, alternative treatment, no treatment)
- Studies reported before 2003
- No language or geographic restrictions

Search Strategy

- Available reference lists
- Personal contacts
  - with program developers, PIs, other experts
- Keyword searches of electronic databases and websites (listed in published protocol and SR) using:
  - (multisystemic OR multi-systemic) AND
  - (treat* OR therap*) AND
  - (evaluat* OR research OR outcome*)
- Results:
  - 5290 hits
  - 266 unique citations

Retrieval and Inclusion Decisions

- 2 independent reviews of titles and abstracts (of 266 citations)
- 95 full-text reports retrieved
- 35 unique studies of MST outcomes
  - 13 excluded (no randomization, wrong population, etc.)
  - 14 ongoing (incomplete)
  - 8 included
- Additional information from primary investigators
  - Sought *all* reports on included studies and additional data

Coding of Included Studies

Independent, double-coding of all:

- Studies
  - Research methods
  - Intervention characteristics
  - Sample characteristics
- Reports (multiple reports per study)
  - Bibliographic information
  - Sample and subsamples
- Outcomes (multiple outcomes per report)
  - Instrumentation
  - Data collection processes
  - Timing
  - Valid N of cases in each group
  - Results

Problems Encountered in Included Studies (not mentioned in previous reviews)

- Unclear randomization procedures in most studies
  - Methods not reported or not fool-proof
  - Not clear whether all cases were randomly assigned in some studies
- Unclear sample sizes (conflicting reports) in 4 studies
  - Number of cases in experiment drops in successive reports (e.g., 210, 200, 176)
- Unyoked designs
- Unstandardized observation periods within studies
  - Follow-up period ranges from 16 to 97 weeks in one study, described as a 57 week follow-up
  - Fixed-interval data (e.g., one-year follow-up) not available for some studies
- Systematic omission of those who
  - Refused treatment, did not complete MST, or did not complete MST “successfully”

Levels of Confidence in Findings

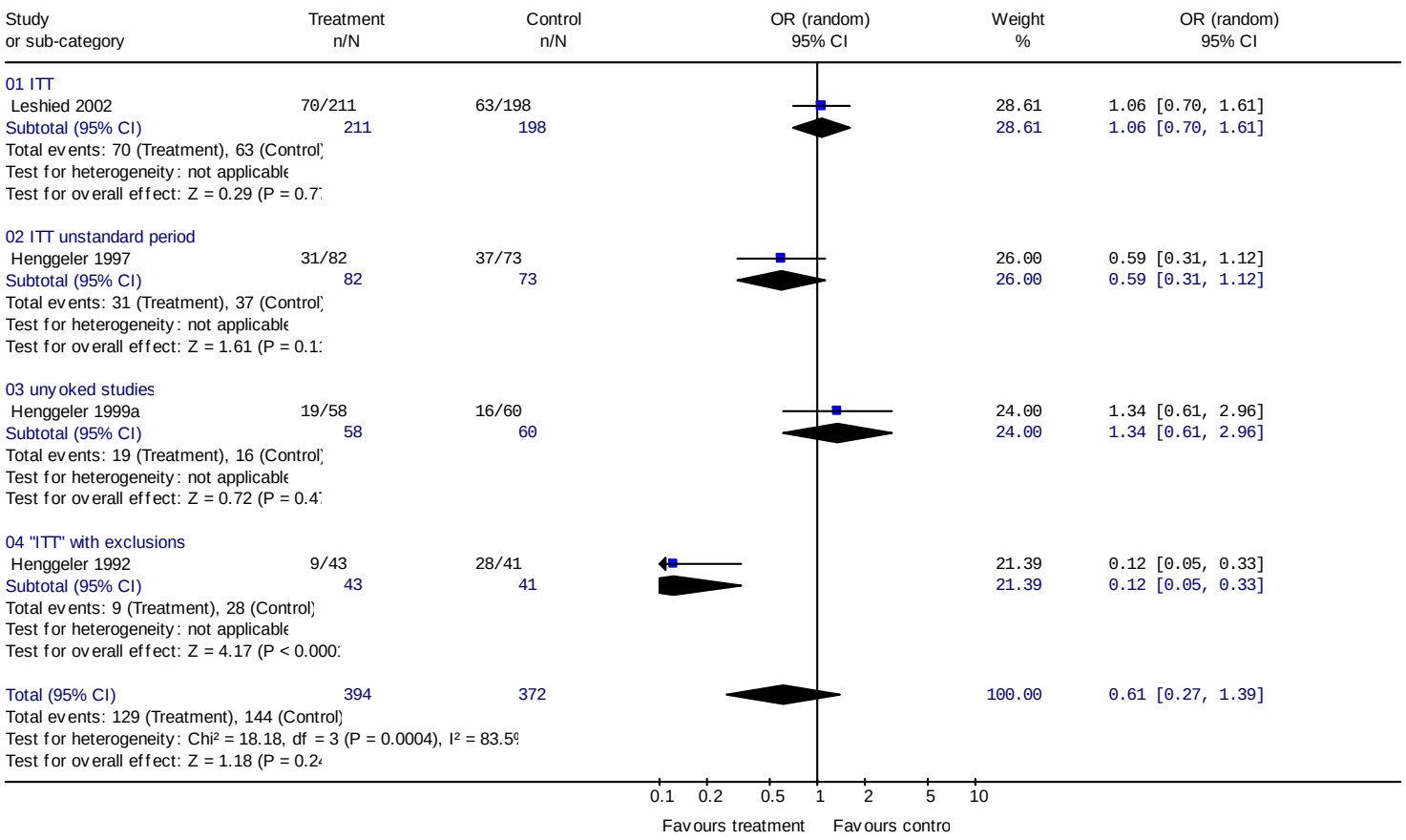
Ranked studies in terms of

- Ability to support intent-to-treat (ITT) analysis
  - No exclusion of MST drop-outs
- Quality of follow-up data
  - One year follow-up vs variable observation periods

- 5 levels of confidence
- Sorted findings by level of confidence
- Pooled results weighted using inverse variance methods

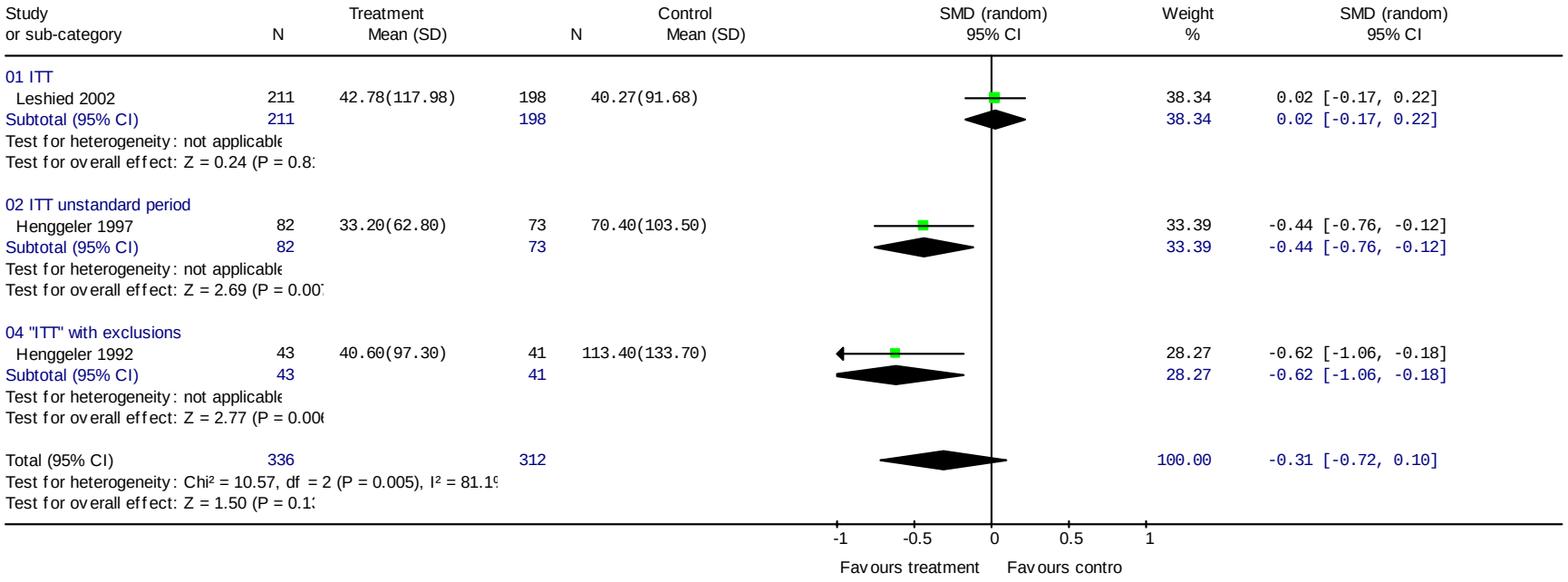
Incarceration (dichotomous)

Review: Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-1
Comparison: 01 Out-of-home placement
Outcome: 01 Incarceration



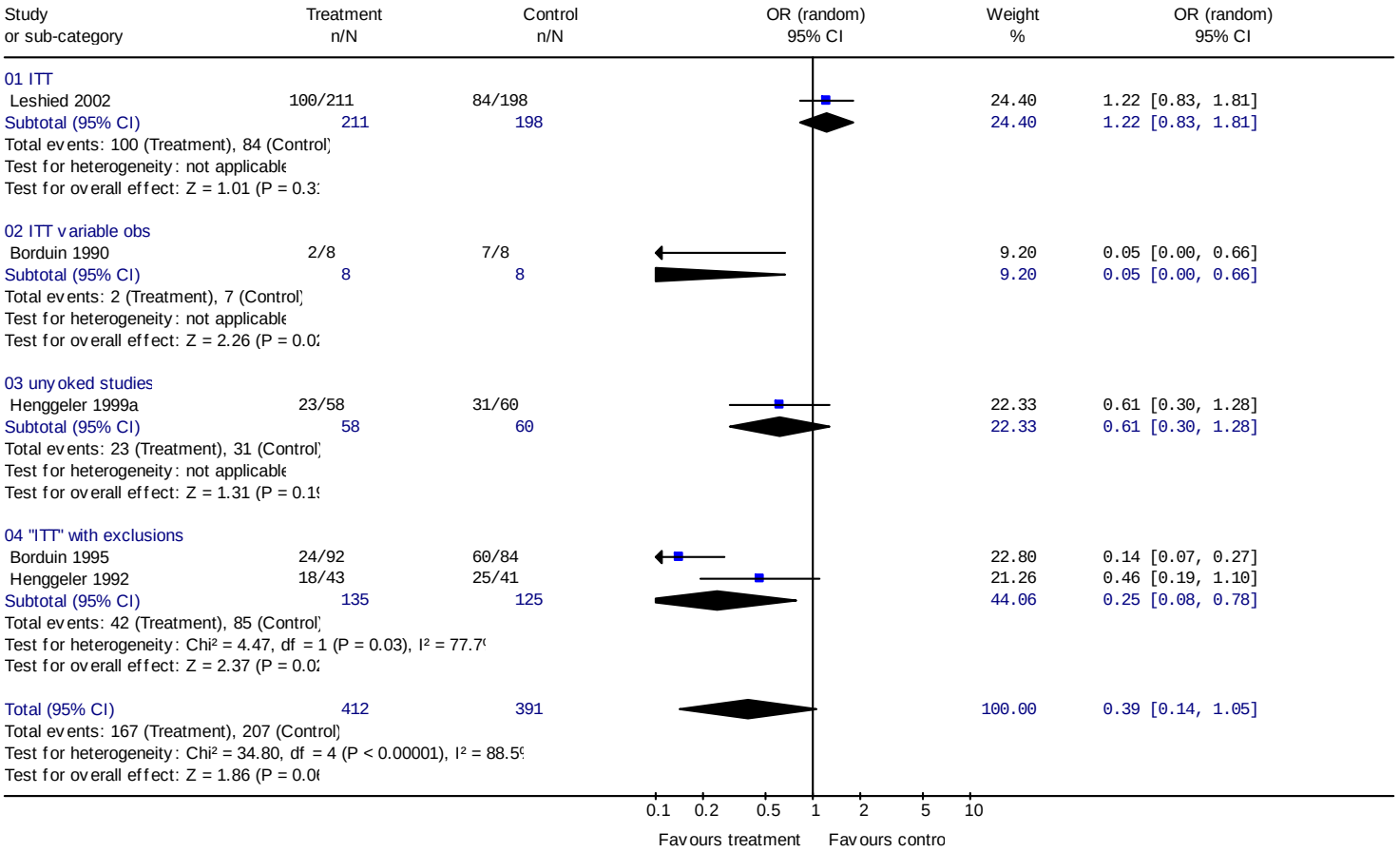
Days incarcerated (continuous)

Review: Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-1
Comparison: 01 Out-of-home placement
Outcome: 02 Days incarcerated



Recidivism (arrested/convicted)

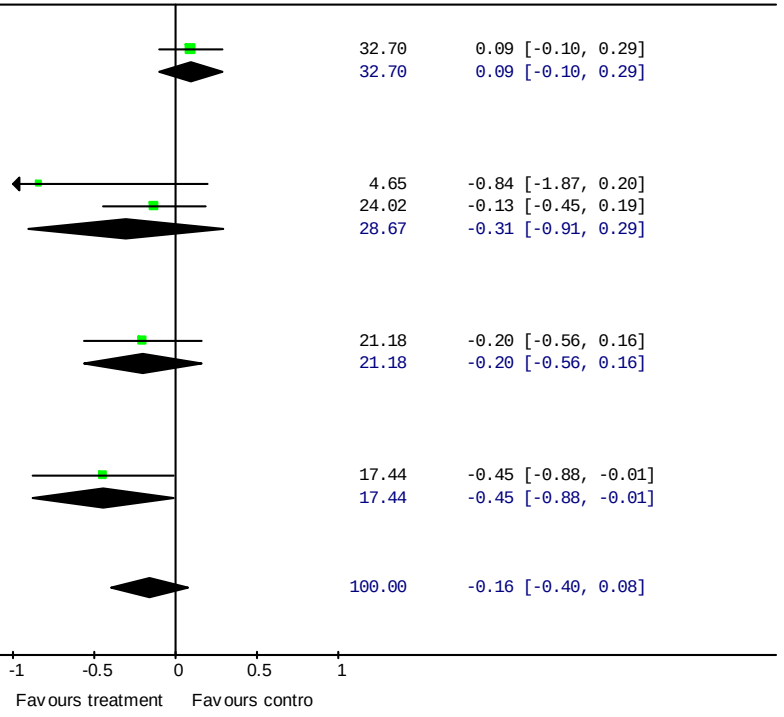
Review: Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-1
Comparison: 02 Arrest or conviction
Outcome: 01 Arrest or conviction



Number of arrests/convictions

Review: Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-1
Comparison: 02 Arrest or conviction
Outcome: 02 Number of arrests or convictions

Study or sub-category	N	Treatment Mean (SD)	N	Control Mean (SD)	SMD (random) 95% CI	Weight %	SMD (random) 95% CI
01 ITT							
Leshied 2002	211	0.74(0.98)	198	0.65(0.93)		32.70	0.09 [-0.10, 0.29]
Subtotal (95% CI)	211		198			32.70	0.09 [-0.10, 0.29]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.95 (P = 0.3							
02 ITT variable obs							
Borduin 1990	8	0.75(1.49)	8	3.88(4.76)		4.65	-0.84 [-1.87, 0.20]
Henggeler 1997	82	0.89(1.39)	73	1.20(3.11)		24.02	-0.13 [-0.45, 0.19]
Subtotal (95% CI)	90		81			28.67	-0.31 [-0.91, 0.29]
Test for heterogeneity: Chi² = 1.64, df = 1 (P = 0.20), I² = 39.2							
Test for overall effect: Z = 1.00 (P = 0.3							
03 unyoked							
Henggeler 1999a	58	0.40(0.61)	60	0.53(0.67)		21.18	-0.20 [-0.56, 0.16]
Subtotal (95% CI)	58		60			21.18	-0.20 [-0.56, 0.16]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.09 (P = 0.2							
04 "ITT" with exclusions							
Henggeler 1992	43	0.87(1.34)	41	1.52(1.55)		17.44	-0.45 [-0.88, -0.01]
Subtotal (95% CI)	43		41			17.44	-0.45 [-0.88, -0.01]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.01 (P = 0.0							
Total (95% CI)	402		380			100.00	-0.16 [-0.40, 0.08]
Test for heterogeneity: Chi² = 8.39, df = 4 (P = 0.08), I² = 52.3							
Test for overall effect: Z = 1.32 (P = 0.1							



Post-treatment effects for program completers (TOT analysis)

No significant average effects on:

- Self-reported delinquency (SRD scale)
- Peer relationships (MPRI scale)
- Behavior problems (RBPC)
- Youth psychiatric symptoms (SCL-90-R, GSI, BSI)
  - Internalizing and externalizing problems (CBCL)
- Parent psychiatric symptoms (SCL-90-R, GSI, BSI)
- Family functioning (FACES Cohesion, Adaptability scales)

Summary: Impacts of MST

- Inconsistent across studies
- No significant effects in ITT analysis
- Few effects in weaker analyses (single studies), not significant on average (across studies)
- Suggests that MST is not consistently better or worse than other services

- Contrary to conclusions of other reviews
  - Which suggest that the effectiveness of MST is well established

Why are these results different from those of prior reviews?

- Traditional reviews appear to prefer:
  - Recent reports (vs. all study reports), don't examine study implementation problems
  - Published reports (publication bias, confirmation bias)
- Uncritical acceptance of RCTs is common
  - Not all RCTs are created equal
  - Some RCTs produce quasi-experimental results
- Different review methods yield different results
  - Narrative summaries of convenience samples of published reports vs.
  - Clear inclusion criteria, systematic search, include unpublished studies, analysis of study quality, and quantitative synthesis

Implications

- Encourage more rigorous primary research on intervention effects
  - Better reporting, using the 2001 CONSORT (CONolidated Standards Of Reporting Trials)
- Encourage more rigorous, systematic reviews of research
  - Use CC and C2 guidelines and standards to minimize bias
  - Better reporting using the QUORUM (QUality Of Reporting of Meta-analysis) standards
- To get better estimates of effects of social programs

Recent Developments

- C2 Social Welfare Initiatives in North America
  - Initial organizational meeting Jan. 2005
  - Work teams
    - Consumer input
    - Communications
    - Funding
- Future C2 Colloquia
  - Feb. 2006 in Los Angeles
  - Feb. 2007 in London

What you can do

- Encourage social work faculty and students to
  - Learn about and use SR methods
  - Identify SR topics, potential reviewers
  - Conduct a SR (lead/join a review team)
  - Join a C2 editorial board or work group
  - Attend C2 colloquia, workshops, interest group at SSWR
  - Read SRs and use results (in developing curricula, programs, research proposals)
- Contact jlittell@brynmawr.edu with questions and suggestions