Does MST work?

Comments on a systematic review and meta-analysis of MST

TERJE OGDEN, Director of research, terje.ogden@atferdssenteret.no

KRISTINE AMLUND HAGEN, Assistant director of research

k.a.hagen@atferdssenteret.no

Norwegian Center for Child Behavioral Development, Unirand, University of Oslo

In this article, we comment on a systematic review and meta-analysis of Multisystemic Therapy (MST) presented by Littell, Popa, and Forsythe (2005) for the Nordic Campbell Collaboration. The review attracted a fair amount of attention and has, in consequence, contributed to renewed discussions about the use of MST in particular and evidence-based methods in general. According to the Nordic Campbell Center, there is "no evidence to suggest that MST is better for young people than alternative forms of treatment. Nor is there, however, any evidence that MST is worse." The authors of the systematic review concluded, based on 8 evaluations of MST, that it is premature to determine the effectiveness of MST compared to other services. In this article, we argue that the foundation on which this review is based is insufficient and assert that it is the meta-analysis that is premature. In consequence, the review seems unfit for evaluating the effectiveness of MST on adolescent behavioral problems, particularly for the Nordic countries.

Note: This article has been translated from Norwegian with permission of the authors and has been reviewed by the authors for accuracy. The full citation for the published version of this article in its original form is as follows:

Ogden,T., & Hagen,K.A. (2006). Virker MST? Kommentarer til en systematisk forskningsoversikt og meta-analyse av MST. Nordisk Sosialt Arbeid, 26,3, 222-233.

Multisystemic Therapy (MST) is a family- and community based treatment method for families with youth exhibiting serious behavioral problems. Since the introduction of MST to Norway in 1999, MST-teams have now been established in every county across the country and approximately 1600 adolescents and their families have received treatment. MST has been evaluated with Norwegian families in a post treatment evaluation, about 6 months after intake, (Ogden & Halliday-Boykins, 2004) and again in a follow-up study two years later (Ogden & Amlund-Hagen, 2006). Results showed that MST was effective in preventing out-of-home placement and in reducing youth problem behavior compared to regular child welfare services. MST has also been evaluated in a series of earlier studies conducted by the program developer, Scott Henggeler and his colleagues, and by and large these studies, too, demonstrate significantly better outcomes for families receiving MST (for an overview, see Curtis, Ronan, & Borduin, 2004). Less encouraging results come from one study conducted in Canada (Lescheid & Cunningham, 2002).

In a protocol from the Campbell Collaboration, Littell, Popa, and Forsythe (2005) present a systematic review and a meta-analysis of the results from eight evaluations of MST in the USA, Canada and Norway[i]. The conclusions of the research review are contrary to those reached in the primary studies on which the meta-analysis is based; namely, that the effects of MST are non-significant. As a consequence, researchers and practitioners who are familiar with the program are left baffled and doubts about MST as an effective treatment method have surfaced. We feel, however, that the meta-analysis is premature and thereby poorly suited as a basis for assessing how effective MST is in the treatment of serious behavioral problems, particularly in the Nordic countries.

In this article we present some of the main themes from the systematic review (Littell et al. 2005) and the subsequent discussion between Littell and Henggeler et al. These are published in the Children and Youth Services Review with the following titles: "Lessons from a systematic review of effects of Multi-systemic Therapy" (Littell 2005), "The Littell paper: Methodological critique and meta-analysis as a Trojan horse" (Henggeler Schoenwald, Borduin & Swensson 2006) and "The case for Multi-systemic Therapy: Evidence or orthodoxy?" (Littell 2006).

Which studies are included in the meta-analysis?

Littell et al. (2005) place certain methodological requirements on the studies to be included in the meta-analysis. The studies must have used a quantitative group design with random assignment of participants to the treatment alternatives (randomized controlled trials, or RCT), and standardized instruments for pre-post comparison of results across studies. Littell et al. included MST-studies with adolescents with serious social and emotional and/or behavioral problems. They did not take into consideration the fact that results often vary in relation to the characteristics of the target group (e.g. results have been less positive when the participants also had serious psychiatric problems). The selection procedure resulted in the inclusion of eight studies—six from the USA, one from Canada, and one from Norway. Littell et al. (2005) did not, however, require that they had been published in peer-reviewed publications, as is often recommended in guidelines for research on evidence-based methods (Flay et al. 2005) and which provides studies with a professional stamp of quality. The selection process meant that this meta-analysis, in contrast to others, included the unpublished Canadian MST study (Lescheid & Cunningham, 2002). Littell (2006) acknowledges that this may have resulted in her conclusions' deviations from those drawn in other meta-analyses.

What does the meta-analysis tell us about MST?

While seven of the eight primary studies found that there were significant differences between treatment conditions in favor of MST, Littell et al. (2005) concluded in their meta-analysis that even though the data had a tendency to favor MST, the differences were not statistically significant. The explanation given by Littell et al. (2005) was that the effects across the studies were inconsistent. In the meta-analysis, Littell et al. re-analyzed the results variable by variable across the eight studies; a total of 21 individual analyses. But while 8 studies were included in the overall review, each individual variable analysis consisted only of between 2 and 5 studies. In other words, none of the analyses included more than five studies, and in the majority of cases only two or three. Within this very limited selection, the studies were also quite heterogeneous, a methodological situation that is expected with studies that differ from each other in important areas such as

choice of study group, observation period, implementation quality, and design. With such a small number of primary studies, considerable homogeneity in design and results is required in order to be able to demonstrate statistically significant differences.

The Canadian MST study (Lescheid & Cunningham, 2002) weighs heavily in the meta-analysis because it had a large sample (N=409) and constituted 21-45 percent of the study population in the meta-analysis. At the same time, it clearly reports less encouraging results compared to the other evaluations. Sundell (2005) speculates that if the unpublished Canadian study had been excluded from the meta-analysis, several significant differences between MST and the comparison groups would have emerged. This assumption is supported by results from the meta-analysis by Curtis, Ronan and Borduin (2004), which included six of the eight studies that are part of Littell et al.'s (2005) review. Despite the fact that there are several similarities in the choice of studies between the reviews by Curtis et al. (2004) and Littell et al. (2005), they arrive at very different conclusions. Curtis et al. (2005) found that MST significantly reduced problem behavior, with an effect size of Cohen's d = .87 in demonstration studies (efficacy) and Cohen's d = .26 in studies from regular practice (effectiveness).

The Canadian study deviates from other MST trials on additional key characteristics. For one, it included adolescents with less severe problematic profiles, second, the intervention in the control group was of higher quality than in the USA, and third, the quality of implementation was worse (Lescheid & Cunningham 2002; Sundell 2005). In another comment regarding the meta-analysis, Sundell (2005) writes that it is difficult to determine whether the Canadian results are attributable to MST as an intervention, to the characteristics of the target group, to contextual factors, or to the degree of program integrity. He indicates that it is important to resolve whether lack of effectiveness is due to weaknesses in the treatment program or weaknesses in the implementation (e.g. poor preparation, or lack of treatment integrity). Littell et al. make no attempt at resolving this issue.

Who should be included in the analyses?

Littell et al. (2005) suggest a quality ranking of the MST studies, and designate "intention to treat" (ITT) analyses to be of the highest quality. In their meta-analysis, only

the Canadian MST study used such a design. An ITT analysis includes all subjects who meet the inclusion criteria and who have been asked to participate in the study, regardless of treatment completion. Adolescents who do not participate (refusers) or who drop out of treatment (drop-outs) may have more serious problems than those who complete the program, and consequently the efficacy of a program may be exaggerated if only the completers are counted. An alternative explanation may be that drop-outs or refusers are families who resolve their problems or feel that their difficulties are not serious enough to warrant MST. Analyses of "treatment of the treated" (TOT, or as described by others "totally treated"), which includes participants who have actually undergone the treatment, are ranked lower by Littell et al. But for those wishing to know how effective MST is, it may be of greater interest to know how those who were actually treated function at post treatment. In such instances, TOT analyses may be more appropriate.

An additional problem accompanying ITT analyses is that they often underestimate the effectiveness of a program by overlooking how such a design can influence the so-called "never-takers" (those who refuse or drop out). "Never takers" can easily become demoralized as a result of their inability to complete the treatment and can, therefore, also negatively skew the results in the experimental group. Consequently, the two treatment conditions are compared on a different basis. As Jo (2003) describes: "This negative psychological effect [of demoralization or discouragement] would not occur for never-takers assigned to the control condition, since the treatment is never offered" (p. 4). Our belief, that ITT analyses can be a poor alternative in RCT analyses, is shared by statistical experts (Jo 2003; Jo & Muthen 2003).

Criticism of the primary studies in MST

On a more detailed level, Littell et al. (2005) highlight what they consider to be methodological weaknesses in the MST studies. These include 1) inconsistent reporting of the number of participants in cases of multiple publications from the same project, 2) inconsistent coupling of pairs of MST and comparison cases, 3) unclear randomization procedures, 4) non-standardized observation periods in follow-up studies, and 5) subjective definitions of treatment termination. Littell et al. claim that selective drop-out in the MST studies has lead to more positive results than would have been the case had drop-

outs been dealt with differently. Littell et al. implies that such choices were made in order to achieve better treatment results and that they therefore undermine the experimental design, the precision of the results and the reliability of the studies. This criticism put forth by Littell et al. (2005) is rejected by Henggeler et al. (2006) who explain and rationalize their procedures. We find that some of these claims are difficult to prove one way or the other, but chose to believe that the researchers responsible for the primary studies are in the best position to report what is accurate, precise and relevant. At the same time, Littell et al. seem more concerned about deviations from what they perceive to be correct procedures for conducting and reporting efficacy studies, than they are about what the consequences of the criticism might be for the treatment results.

Littell et al. (2005) thus maintain that there is a general failure of MST evaluations to adhere to the strict norms of conducting randomized controlled trials (RCTs). Our position is that while these demands are useful in the sense that they function as ideals for which to strive, complete adherence to such standards is unrealistic, especially in effectiveness trails and hence does not serve the very people for whom treatment is intended. Additionally, a study's design, participant characteristics and choice of outcome variables will often be guided by changing needs, interests and opportunities for publication. And we question whether the results from the MST studies that are included in the review would have differed substantially as a result of such conformity. Moreover, we see a logical slip in Littell et al.'s criticism. When Littell et al. disapprove of the methods, analyses, and procedures of the investigations included in their meta-analysis they simultaneously undermine their own study: A meta-analysis can never be better the investigations on which the analysis is based.

Factors that modify the effects of MST

The disagreement between Henggeler et al. (2006) and Littell et al. (2005) also concerns the extent to which one should take into consideration factors that likely have a moderating effect on the results. We hold that such factors are important to consider not only in the execution of evidence-based treatment methods, but also in meta-analyses. These are: 1) Differentiating between efficacy (university-based) and effectiveness

(community-based) studies, 2) Examination of treatment fidelity, 3) Consideration of site differences, 4) Program maturity.

1. Lack of differentiation between research-based and practice-based studies.

RCTs of treatment effect can be catalogued as either "efficacy" or "effectiveness" studies. *Efficacy* studies demonstrate the results a program can furnish under optimal conditions (usually with important contributions from the program developer or researcher). *Effectiveness* studies, on the other hand, evaluate how the program works in regular practice (the program developer having a peripheral role). Henggeler (2004) notes that there is still significant work to be done before we know how to implement research-based methods in regular practice and achieve results that come close to those reported in demonstration studies (or efficacy studies). Littell et al. (2005) do not, however, differentiate between "efficacy" and "effectiveness" studies in their meta-analysis, and believe that the difference between these is unclear in the MST research. Consequently, any attempt to explain variations based on such a division will be, in their opinion, speculative.

Petrosino and Soydan (2005) found that when the evaluator, or the program developer, played a significant role in the development of design and implementation of a treatment program, higher effect sizes were more easily obtained than in other evaluations. Lipsey (1995) questions how such results should be interpreted: "A cynical view might attribute this (i.e. better results in cases where the researcher determines design and implementation) to some biasing or 'wish fulfilling' influence researchers have on the outcomes of studies they control. I see another interpretation as plausible… When a researcher is closely involved in treatment design… there is likely to be a high level of treatment integrity…" (p. 75). One can choose to interpret this as either that the expectations of the program developer influence the results or that the difference is due to an extra focus on achieving a high degree of program and treatment fidelity. In our opinion, Littell and her colleagues (2005) chose the "cynical view" and discounted the fact that the differences may be a matter of implementation quality (Wilson 2004).

2. The significance of treatment integrity

Treatment integrity in MST is assessed using the "treatment adherence measure" (TAMS, Henggeler & Borduin 1992). The TAM has been shown to correlate significantly with

treatment results (Henggeler 2004). That is, the greater the treatment adherence, the better the outcomes for the participating families. Littell et al. (2005) feel, however, that the TAM does not discriminate between MST and other interventions, arguing that the elements being measured are not unique to MST (e.g. commitment, treatment participation and therapeutic alliance). We believe that Littell et al. are criticizing the TAM scale for lack of discriminating properties based on inaccurate assumptions and without empirical foundation. In our opinion, all programs consist of a combination of both general components important to any therapeutic method and unique program components. MST, for example, requires that the therapists master general processing abilities which are common to most treatment programs, but also that they are knowledgeable about and proficient with the use of specific MST components (working towards increased contact and fit between families and surroundings, and the targeting of defined problems).

3. Differences in treatment effect between test sites (site differences)

Henggeler et al. (2006) argue that variations in treatment integrity between treatment sites should be taken into consideration. Littell et al. (2005) claim that the data do not support the contention that MST is more effective at some sites than others. They refer to the fact that in two of three MST studies involving multiple testing sites, results were not reported from the individual treatment sites. They also suggest that it is wrong to claim differences in treatment effects between sites in hindsight (post hoc) and call for multilevel studies which can explain differences in treatment effect between both therapists and sites (so-called nesting effects). This request, however, is somewhat unrealistic given that multi-level studies require a high number of testing units (i.e. the study should include 20 to 30 testing sites with several therapists operating in each site). Littell et al. (2005) also claim that results from individual testing sites ("site effects") are of no interest in a meta-analysis, the purpose of which is to aggregate the results from all available research. Littell's call for a more advanced analysis of variations between testing sites seems to us contradictory, when she simultaneously rejects its relevance in meta-analyses.

4. Effects of duration of program use ("evaluating programs before they mature")

Results from MST studies show, according to Henggeler et al. (2006), that the first year of program operation is particularly demanding for many programs when they are implemented in new sites and without direct contribution from the program director.

Consequently, evaluations of programs immediately after their initiation result in an underestimation of their potential. Littell's (2006) response is based predominantly on the Canadian study in which she, in contrast to Henggeler et al., claims that the data do not support the conclusion that the results of the program were influenced by its maturity.

In summary, Henggeler and colleagues (2006) assert that it is important to differentiate between treatments operating in various phases of the program development. Best results are achieved in studies in which there is a high degree of control of treatment integrity, a high degree of program integrity at the testing site, and where there is a certain degree of experience with the use of the program. Littell, on her part, rejects the assertion that MST research supports this.

Other issues

Meta-analyses in this field should contain far more studies than the number of investigations included in the Littell et al. review. Considering the limited number of studies and the method by which these were selected, weighed and ranked, the resulting conclusion appears somewhat oversimplified and certainly incongruent with the primary studies upon which the analysis is based. Littell et al. (2005) acknowledges the limitation of their small sample by stating "Since statistical power is low, we cannot conclude that MST is not more effective than other services" (p. 22). The review's inadequate sample size could have been compensated for by including evaluations of other programs focusing on the same target group. As a result, the statistical strength of the analyses would have been increased, along with a broadened focus on the study of family-based treatment programs.

An even more serious shortcoming of the meta-analysis, however, is the failure to acknowledge the heterogeneity of the selected studies. We refer here to a recommendation forwarded by The Cochrane Collaboration presented on their web-site (2002) which is supportive of our view. The Cochrane Collaboration is an international and independent information-providing organization and is referenced by Littell et al. (2005) on several occasions. It has developed guidelines for conducting meta-analyses and gives the following advice: "If you identify and suspect that important diversity or heterogeneity is present in your review, there are several options open to you. [...] one option is that of not

performing a meta-analysis. An unwise meta-analysis can lead to highly misleading conclusions." We are surprised by the fact that this advice is ignored in the present Campbell and Cochrane review of MST.

Finally, we have difficulties distinguishing between criticism directed toward the research procedures of the MST evaluations and criticism directed toward MST as a treatment method. We assume that the former is the official purpose of this review, but are struck by several statements indicating the latter. For example, Littell et al. (2005) write that they are doubtful as to whether a short-term treatment program delivered to individual families can improve the lives of youths with serious behavioral problems. We feel that it is important to differentiate between research criticism, and criticism of MST as a treatment method, and hold that criticism of the quality of the research publications does not on its own provide grounds for expressing an opinion about the effectiveness of the treatment program. They also write that more robust, long-term interventions, combined with consistent financial, educational, medical and therapeutic support for adolescents and their families may be necessary in order to bring about improvements in the lives of these adolescents. This assumption is obviously not grounded in any empirical work, but may rather signify the authors' personal convictions about MST as a treatment method. We have been unable to find empirical support for these viewpoints in the systematic research review—and doubt that this prescription is suitable for Norwegian and Nordic conditions. This is because the services that Littell et al. request have been part of the existing clinical work with serious behavioral problems in the Nordic countries, without it appearing to be producing the desired result (Storvoll, 1997). And this is perhaps the reason why authorities in the Nordic countries have called for a treatment alternative organized on the basis of the scientific principles represented by MST (i.e., research-based, multi-systemic, home-based services, client participation etc.)

Summary

According to Littell et al. (2005), the main intention of the systematic research review was to focus on the gap which can exist between research and practice in working with meta-analyses. Furthermore, their purpose was to synthesize results from MST evaluations in order to make general statements about its effectiveness. In doing so, they

request stringent adherence to ideal standards from method books in the execution of real world investigations and in the reporting of findings. One can not, however, argue the inadequacies of a body of research and at the same time defend conclusions based on results from that very research.

Much of the discussion and disagreement between Henggeler and Littell is concerned with the weight given to (or the inclusion of) the Canadian study. Henggeler et al. (2006) stress that the study is unpublished (and its quality thus questionable), and that the results are not as unilaterally negative as Littell claims. Littell, on her part, states that the Canadian study is the first large independent replication study of MST and that it therefore does not have the same degree of "bias" as the studies conducted by Henggeler and his colleagues. Littell et al. do acknowledge, however, that this study weighs heavily in the meta-analyses by virtue of the sample size (N=409). No other MST study involves more than 200 participants.

Littell and colleagues (2005) assert the need for more independent studies in order to confirm or disaffirm MST's effectiveness compared to other child welfare services. And they further claim that conclusions regarding the effectiveness of MST are premature. We believe that more studies of MST need to be conducted before a proper and thorough meta-analysis of MST can be undertaken and that, as such, it is the meta-analysis conducted by Littell et al. that is premature.

The research review comes at a time when the discussion about the value of evidence-based methods in general, and MST in particular, in the Nordic countries is at a peak. At the same time, the Nordic Campbell Center has been established, the intention of which is to present research reviews of current treatment programs and methods. In this context, the review and meta-analysis of MST must appear confusing. Most people have gained the understanding that MST has produced good results in controlled evaluation studies, and they probably expected that the research review to some degree would reflect this. The Nordic Campbell Center, however, presents the research review under the heading "Doubt about the effects of MST," and unreservedly states that the method is neither better nor different than other methods (www.nc2.net). Littell and colleagues (2005) are somewhat more restrained and despite their disapproving conclusions, they assert that there is no evidence suggesting that any known intervention is more effective

than MST. Based on the Norwegian evaluation results (Ogden & Halliday-Boykins 2004; Ogden & Amlund Hagen 2006), we will assert that MST appears satisfactory in a Norwegian context—compared with other child welfare services for adolescents with serious behavioral problems, providing that the program's core components are implemented in line with the guidelines from the program developer.

[1] p-value of the result variable, "rearrests" was: 0.0504, which many people would have judged as being a significant difference in favor of MST at the 5% level.
[2] According to Lipsey (2000), the test (Q-test) used in such analyses has weak "power" and at least 20 separate studies are required in order for one to be able to draw any conclusions about mean results.
[3] Low "statistical power" increases the risk of so-called type II error, i.e. that real differences are not identified by statistical hypothesis testing (Sundell 2005).

REFERENCES

Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004) Multisystemic treatment: A meta-analysis of outcome studies. *Journal of Family Psychology,* 18, 411-419.

Flay, Brian. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., Moscicki, E. K., Schinke, S., Valentine, J.C., & Ji, P. (2005) Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science,* 6, 151-171.

Henggeler, Scott.W. (2004) Decreasing effect sizes for effectiveness studies - Implications for the transport of evidence-based treatments: Comment on Curtis, Ronan and Borduin .*Journal of family psychology*, 18, 420-423.

Henggler, Scott. W. & Borduin, C. M. (1992) *Multisystemic therapy adherence scales.* Unpublished instrument. Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina.

Henggeler, Scott W., Schoenwald, S. K., Swenson, C. C., & Borduin, C. M. (2006). Methodological critique and meta-analysis as a Trojan horse. *Children and Youth Services, 28,* 447-457.

Jo, Booil (2003) Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics, 27*, 385-409.

Jo, Booil & Muthen, B. O. (2003) Longitudinal studies with intervention and noncompliance: Estimation of causal effects in growth mixture modeling. In S. Travel & N. Duan (Eds). *Multilevel modeling: Methodological advances, issues, and applications.* (pp. 112-139). Mahwah, NJ, US: Lawrence Erlbaum Associates.

Leschied, Allan. & Cunningham, A. (2002) *Seeking effective interventions for serious young offenders: Interim results of a four-year randomized study of Multisystemic Therapy in Ontario, Canada.* Centre for Children and Families in the Justice System, London, Canada.

Lipsey, Mark (2000) *Practical meta-analysis.* Seminar at CUS, Socialstyrelsen, September 2000, ref. from Sundell,K. (2005).

Lipsey, Mark. (1995) Cited here from A. Petrosino (2004). *Presentation prepared for Social methods seminar, 9 December, 2004, Centre for evaluation of social services, National Board of health and welfare, Stockholm.*

Littell, Julia. H. (2005) Lessons learned from a systematic review of multisystemic therapy. *Children and youth services review,* 27, 445-463.

Littell, Julia. H. (2006) The case for Multisystemic Therapy: Evidence or orthodoxy? *Children and youth services review*, 28, 458-472.

Littell, Julia. H., Popa, M.& Forsythe, B. (2005) *Multisystemic therapy for social, emotional and behavioral problems in youth aged 10-17.* Rapport 05:01, Nordic Campbell Center www.nc2.net

Ogden, Terje. & Amlund-Hagen, K. (2006) Multisystemic Treatment of Serious Behaviour Problems in Youth: Sustainability of Effectiveness Two Years after Intake, *Child and adolescent mental health,* in print.

Ogden, Terje. & Halliday-Boykins, C. A. (2004) Multisystemic treatment of antisocial adolescents in Norway: Replication of clinical outcomes outside of the US. *Child and adolescent mental health,* 9, 76-82.

Petrosino, Antonio., & Soydan, H. (2005) The impact of program developers as evaluators of criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of experimental criminology*, 1, 435-450.

Storvoll, Elisabeth. (1997) Children and young people with serious behavioral problems: Who are they, and what help are they being offered? Report 21. Oslo: NOVA.

Sundell, Knut. (2005) *Scientific scrutiny of Systematic and nonsystematic reviews of effects of Multisystemic Therapy (MST) by Julia Littell.* Report to the Institute for Methods in Social Work (IMS), Socialstyrelsen, Stockholm.

The Cochrane Collaboration (2002) The Cochrane Collaboration Open Learning Material. Retrieved from http://www.cochrane-net.org/openlearning/HTML

Wilson, Sandra. J. (2004) *Do program developers influence program outcomes? An example from meta-analysis.* Social methods seminar, 9 December, 2004, Centre for evaluation of social services, National Board of health and welfare, Stockholm.

Summaries

In a systematic research overview from Campbell Collaboration, Littell, Popa and Forsythe (2005) present a meta-analysis of the results from eight evaluations of Multisystemic Therapy (MST). The overview concludes in that it is premature to draw

conclusion about the efficiency of MST compared with other services. This article argues that it is the meta-analysis which is premature, and that consequently it is poorly suited as a basis for assessing how effective MST is in working serious behavioral problems, particularly in the Nordic countries.

We comment on a systematic review and meta-analysis of Multisystemic Therapy (MST) presented by Littell, Popa, and Forsythe (2005) for the Campbell Collaboration. The authors concluded, based on eight evaluations of MST, that it is premature to determine the effectiveness of MST as compared to other services. In this article, we argue that the foundation on which this review is based is insufficient and assert that it is the meta-analysis that is premature. In consequence, the review seems unfit for evaluating the effectiveness of MST on adolescent behavioral problems, particularly for the Nordic countries.

---

i i