

A randomized controlled trial of the effectiveness of multisystemic therapy in the Netherlands: post-treatment changes and moderator effects

Jessica J. Asscher · Maja Deković ·
Willeke A. Manders · Peter H. van der Laan ·
Pier J. M. Prins ·
Dutch MST Cost-Effectiveness Study Group⁴

Published online: 7 November 2012
© Springer Science+Business Media Dordrecht 2012

Abstract

Objective In the present randomized controlled trial, the effectiveness of multisystemic therapy (MST) in The Netherlands was examined. Moderator tests were conducted for ethnicity, age and gender.

Methods The sample consisted of $N=256$ adolescents, referred because of conduct problems, and randomized to MST or treatment as usual (TAU). Assessments (questionnaires and observational ratings) took place before and immediately after the treatment.

Results MST was more effective than TAU in decreasing externalizing behavior, ODD, CD and property offences, but not for violence. Findings were mixed for adolescents' and parental cognitions: the MST group, compared to TAU, showed an improvement in parental sense of competence, and a decrease in adolescents' hostility, but no change in self-esteem and an increase in personal failure. MST was effective for positive dimensions of parenting and associations with prosocial peers, but not for relationships with deviant peers. MST was equally effective for adolescents of different ages and with different ethnicities. However, MST showed larger (and more positive) effects for adolescent cognitions for boys than for girls.

Members are: D.E.M.C. Jansen, K.M. Vermeulen, A.H. Schuurman-Luinge, E. Buskens, E. J. Knorth, and S.A. Reijneveld

J. J. Asscher (✉)

Forensic Child and Youth Care Sciences, University of Amsterdam, Nieuwe Prinsengracht 130,
1018 VZ Amsterdam, Netherlands
e-mail: J.J.Asscher@uva.nl

M. Deković · W. A. Manders
Utrecht University, Utrecht, Netherlands

P. H. van der Laan
Free University Amsterdam/NSCR, Amsterdam, Netherlands

P. J. M. Prins
Child Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands

Conclusions Effects of MST in The Netherlands are generally comparable to the positive findings reported in American and Norwegian trials. MST seems equally effective across age and ethnic minority groups, but some gender moderator effects were found for adolescent cognitions.

Keywords Effectiveness · Multisystemic therapy (MST) · Externalizing behavior problems · Parenting · Cognitions · Peers

Introduction

It is nowadays generally recognized that, given the multi-determined nature of antisocial behavior, interventions that aim to reduce antisocial behavior must address multiple causes (“broad focus” criterion) and should be delivered in the natural environment, i.e. in the context in which this behavior occurs (“ecological validity” criterion). Multisystemic therapy (MST) is based on these criteria. It is a multi-faceted, intensive home- and community-based treatment for youths who show serious, violent and chronic antisocial behavior (Henggeler and Lee 2003; Henggeler and Schaeffer 2010).

Previous short-term effectiveness studies of MST conducted in the USA (Henggeler et al. 1992, 1997; Rowland et al. 2005; Timmons-Mitchell et al. 2006), Norway (Ogden and Halliday-Boykins 2004) and the UK (Butler et al. 2011) showed positive effects: MST-treated juveniles showed less delinquent behavior and recidivism than juveniles from the control groups (see, for a recent overview, Henggeler 2011). Also, long-term effectiveness trials on MST exist, which suggest that the effects are lasting well into adulthood (Borduin et al. 2009; Henggeler et al. 2002; Schaeffer and Borduin 2005; Welsh et al. 2012). However, there have also been studies finding null results for MST or more positive outcomes for the comparison groups (e.g., Brunk et al. 1997; Leschied and Cunningham 2002). Moreover, Littell et al. (2005) criticized the methodology of some of the previous studies showing positive results by questioning the random assignment procedures, the unexplained attrition, the absence of intent-to-treat analyses, the use of unstandardized observation periods, and the undisclosed conflicts of interest in some MST trials.

There are several reasons why replication of the effectiveness of MST is needed. First, the trials conducted in Canada (Cunningham 2002) and Sweden (Sundell et al. 2008) did not replicate the positive findings for MST. Second, there are large differences in composition of the populations between, for example, the USA and European countries, but also among European countries. In the Netherlands, for example, there is a high chance that ethnic minorities will participate in MST, as they are overrepresented in the juvenile justice system (Blokland et al. 2010). The composition of the ethnic minority groups in The Netherlands differs substantially from ethnic minority groups in the USA in terms of religion, duration of time since migration, language difficulties, and culture. Third, there are differences between countries in the quality of youth services, and consequently in the quality of treatment offered to the control group and thus that of participants in the comparison/control condition. In The Netherlands, as in Sweden, in-home services are used more frequently than in the USA, where juvenile offenders are most likely to be treated within the juvenile justice system (Sundell et al. 2008; Lipsey 1999). These

differences in intensity and type and quality of treatment offered to the control group might affect the results of effectiveness trials. Additionally, differences in treatment adherence between countries might influence the results with respect to effectiveness (Henggeler et al. 1997; Schoenwald et al. 2004). A meta-analysis by Curtis et al. (2004) indicated that effect sizes were larger in efficacy trials, in which treatment was carried out by highly supervised university students, than in effectiveness studies, in which treatment was conducted by community professionals.

Besides replicating existing studies of MST in different circumstances with different participant groups, the present study enhances the previous work in two ways. First, the present study includes outcomes that were previously not addressed that may be affected by MST and have been found to be related to delinquent behavior. In their meta-analysis of MST effectiveness studies, Curtis et al. (2004) indicated that researchers should “broaden their assessment” (Curtis et al. 2004, p. 417) to focus more on the aims of MST. They suggest that both parental and adolescent changes in cognitions should be additional outcome measures. Parents’ cognitions, such as parental competence, were shown to be important for parenting behavior in previous research (Coleman and Karraker 1997; Donovan et al. 1990). Potier and Day (2007) suggested that parental cognitions may be an important link in the treatment of conduct problems. Deković et al. (2010) showed that participation in the Home-Start parenting support program was related to changes in maternal sense of competence, which in turn predicted changes in parenting. Furthermore, there is ample evidence that adolescent’s cognitive distortions (van Vugt et al. 2011), self-esteem (Mason 2001; Donnellan et al. 2005), and perceptions of failure (Schneider and Leitenberg 1989) are related to aggressive behavior and delinquency. Moreover, despite MST’s aims to increase relationships with prosocial peers, research to date has solely focused on the question whether associations with deviant peers have decreased. In this study, the following broad outcomes are included: different types of problem behaviors (general measure of externalizing problems, ODD and CD symptoms, and different types of offenses, such as property offence and violence), several dimensions of parenting (positive discipline, inept discipline, quality of relationship between parent and adolescent), and outcomes not previously assessed such as both parents’ (sense of competence) and adolescent’s cognitions (hostility, self-esteem, and personal failure), and adolescent’s relationship with not only deviant but also prosocial peers. Furthermore, multi-informant assessment is used: adolescent self-report, parent report, and observational ratings.

Second, moderator effects were tested to examine differential effects of MST for different subgroups. Ethnicity, gender and age were examined as possible moderators of MST effectiveness. Given the above-mentioned differences between American and European ethnic minority groups, it is important to study potential differential effectiveness of MST in Dutch ethnic minorities. Next, little is known about the gender specificity of juvenile justice programs (Zahn et al. 2009). This is remarkable as research consistently shows juvenile female delinquents to have specific problems, more frequently being sexual abuse victims and experiencing mental health problems (Emeka and Sorensen 2009). As a result, female delinquents are more often traumatized, which warrants specific treatment. Finally, MST is offered to adolescents aged between 12 and 18. Recently, research has shown that the influence of dynamic risk factors decreases during the course of adolescence (Van der Put et al. 2011). It is important to study differential effectiveness of MST for different ethnic minorities, gender and age groups.

Methods

Recruitment and randomization

All referring agencies (Child Protection Council, juvenile judges, Bureaus Youth Care, local referral institutions) were informed beforehand about the study and gave their consent. Adolescents were referred by primary health care (GP) or child social workers in 39 % of cases. Fifty-one percent of the adolescents was court-ordered and 11 % were self-referred. The referrers informed the juveniles and their families that were considered to be eligible for MST that a study was being conducted to investigate the effectiveness of youth care. If the families met the inclusion criteria for MST according to the MST supervisors of the participating institutions, research procedures were explained to the juveniles and their families and their informed consent to participate in the study was obtained by researchers.

Immediately after referral, participants were randomized with the use of a computerized randomization program. This program was executed separately for each site. The randomization ratio was adjusted in a 1:2 ratio in favor of MST for a 6-month period, due to a low number of referrals. Once an adolescent was randomized to TAU, the staff involved in the referral (from Bureau Youth Care or Child Protection Council) together with the MST supervisor, referred the adolescent to an alternative treatment. For a more elaborate description of the randomization process, see Asscher et al. (2007).

A total of $N=256$ juveniles were assigned to MST ($n=147$) and a 'treatment as usual' (TAU) control group condition ($n=109$). According to Cohen (1992), a sample size of 64 per treatment condition is sufficient to test the stated hypotheses assuming .80 power, an alpha of .05, and a medium effect size. Data were collected by research assistants before the start of the treatment and immediately after termination of the treatment ($M=5.72$, $SD=1.90$, months after pretest) in the homes of the participants. The same assessments took place before and after intervention, with the exception of demographic characteristics, such as ethnicity, gender and date of birth, which were asked only once. The majority of the research assistants who visited the families at home were not informed of the family's randomly assigned condition. Research assistants were blind to the study hypotheses.

During the home visits, parents and adolescents filled out the questionnaires and their behavior was observed and rated after the home visit. The complete research protocol can be obtained from the first author.

The design of the study was approved by the institutional review board and the medical ethic committee of Utrecht University. The trial was registered in the Dutch Trial register (number: 1390). Each family member received 10 euros for completing each assessment.

Participants

Participants were juveniles who were referred to MST between 2006 and 2010 and who met the inclusion criteria for MST (MST services 2011). Inclusion stopped when sufficient participants were participating. Data collection took place in three MST institutions spread over the Netherlands. Inclusion criteria for the study were identical to those used by the MST offering agencies: severe and violent antisocial behavior at

home, school or community, sufficiently serious to require treatment, age between 12 and 18 years. Exclusion criteria were: ongoing treatment by another agency, substance abuse without antisocial behavior, sexual offending, autism, acute psychosis, or imminent risk of suicide, and presence of the youth in the home posed a serious risk to the youth or to the family.

Judicial records of offending behavior were used to calculate the number of arrests before treatment. According to official justice data, 71 % of the participants had been arrested at least once before treatment. According to the self-reports, 64 % of the adolescents had contact with the police at some point during the year before the baseline assessment. Using the clinical cut-off score of the 90th percentile for the CBCL, 75 % of the adolescents were reported to have clinically elevated scores on externalizing behavior problems by their parents at baseline. Figure 1 charts the flow of participants from referral to data analyses. Despite extensive tracing efforts, 33 participants were lost to post-intervention assessment. Participants lost to post-intervention assessment did not differ significantly from those retained on any assessed variable. Little's MCAR test indicated that data were missing completely at random, $\chi^2(3,097)=32,00.556, p=.095$. Thus, all 256 participants were included in

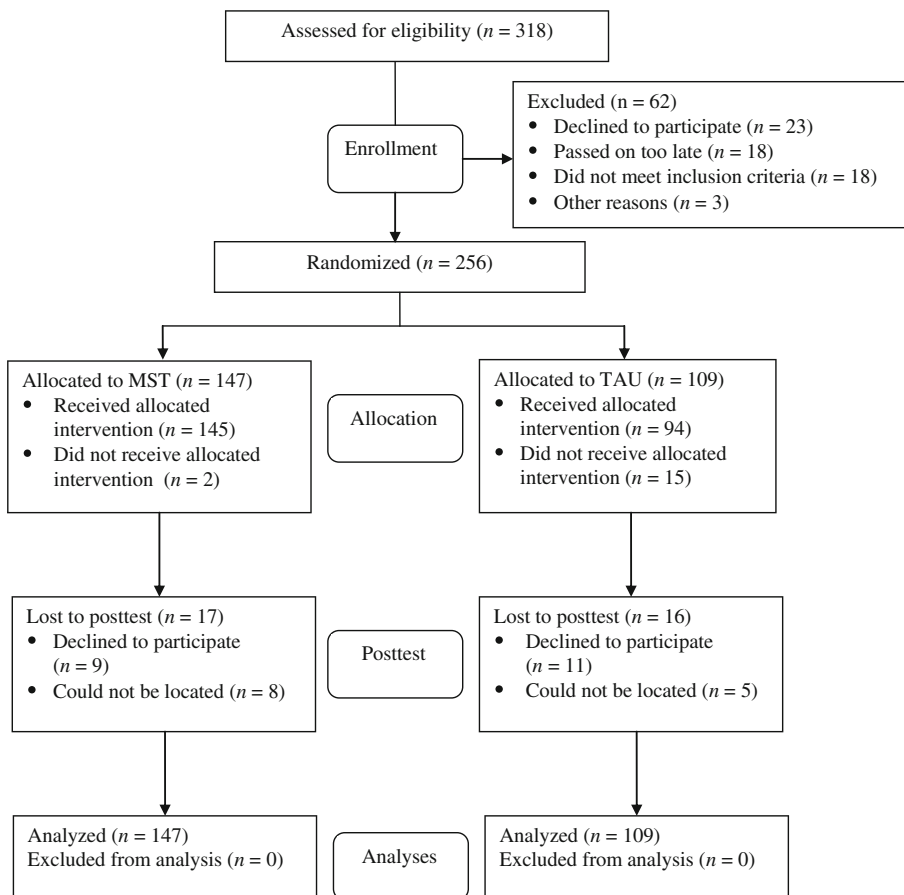


Fig. 1 Flow diagram of the participants

the analyses, and using LISREL 8.8, the multiple imputation was carried out by the expectation maximization algorithm (Graham 2009).

The sample consisted of $n=188$ boys and $n=68$ girls, with an average age of $M=16.02$, $SD=1.31$, years. Fifty-five percent of the adolescents had a Dutch ethnicity. Of the adolescents belonging to ethnic minority groups, most had a Moroccan (34 %) or a Surinamese (32 %) background. Half of the adolescents came from a single-parent family. Fifty percent of the mothers and 36 % of the fathers were unemployed. Forty-five percent of the families experienced financial strains and more than half of the families (56 %) lived below minimum income levels. Independent samples *t*-tests for continuous variables and chi-square analyses for categorical variables were used to examine difference between treatment conditions at T1 on demographic and the outcome variables. No significant differences were found on any of these variables, suggesting that randomization was successful.

Conditions

MST MST is based on social ecological and family systems theories, and on research on the causes and correlates of serious antisocial behavior (Henggeler 2011; Henggeler and Schaeffer 2010). It addresses several key systems in which the adolescent is embedded: family, school, peer group, and neighborhood. MST services are often provided in homes at times that are convenient for the families, but meetings are also held in schools, neighborhood settings, or social service agencies. In consultation with family members, the therapist identifies a well-defined set of treatment goals, assigns the tasks required to accomplish these goals, and monitors the progress in regular family sessions at least once a week. The MST therapist training protocol is quite comprehensive, yet the treatment itself is highly individualized to address specific needs of clients.

Treatment as usual (TAU) Participants in the control condition received an alternative treatment that would have been offered had MST not been available. Mostly, these services included individual treatment (individual counseling or supervision by probation officer or case manager, 21 %), and family-based interventions (family therapy, parent counseling, parent groups, or home-based social services, 53 %). Seven percent received a combination of care (e.g., individual treatment and family counseling), and 4 % was placed in a juvenile detention facility. Fourteen percent eventually received no treatment due to various reasons such as moving house or repeated no show at treatment sessions.

Measures

Primary outcomes

Externalizing problems and delinquent behavior In order to assess *externalizing behavior*, several instruments and informants were used. Parents were asked to report on the externalizing behavior of their child with the Child Behavior Checklist (Verhulst et al. 1990). For externalizing behavior problems (aggression and delinquent behavior, 33 items) items had to be answered on a 3-point scale, ranging from 0

(*never*) to 2 (*often*). Cronbach's alpha's for externalizing behavior were α T1=.82; α T2=.81.

Parents filled out the DSM symptom scales for behavioral problems assessed with the Disruptive Behaviors Disorder rating scales (Oosterlaan et al. 2000). The subscales *Oppositional Defiant Disorder* (9 items) and *Conduct disorder* (18 items) had to be answered on a 4-point scale, ranging from 1=*not at all*, to 4=*a lot*. The Oppositional Defiant Disorder (ODD) problems scale (e.g., argues with adults) reached an α of .92 (both T1 and T2). The Conduct Disorder (CD) subscale (e.g., broke into and entered someone's home or car) had α s of .71 (T1) and .83 (T2).

Juveniles were asked to report on their own *externalizing behavior* problems by means of the externalizing behavior problems subscale of the Youth Self Report (Achenbach 1991; Verhulst and Van der Ende 1992), which consists of the aggression and delinquency subscale, in total consisting of 30 items, to be answered on a 3-point scale, ranging from 0 (*never*) to 2 (*often*). Cronbach's alphas were α T1=.89; α T2=.90.

Additionally, two subscales of the Self-Report Delinquency scale (SRD) (Blom and van der Laan 2006) were used to assess self-report delinquency. Participants were asked to indicate on a list of potential delinquent behaviors whether they engaged in the described behaviors during the past 6 months ("yes" or "no"). The SRD *Violent offending* (5 items, α T1=.74; α T2=.73) and *Property offences* (10 items, α T1=.76; α T2=.82) scales were used.

Secondary outcomes

Parent and adolescent cognitions

Parental competence Parental competence was assessed with the competence subscale of the Parenting Stress Index (PSI; Abidin 1983), which consists of 15 items, to be answered on a 6-point scale, ranging from 1=*completely disagree* to 6=*completely agree*. Cronbach's alphas were .87 (T1) and .90 (T2).

Adolescent cognitions Adolescent cognitions were assessed with adolescent report on two subscales of the Children's Automatic Thought Questionnaire (Schniering and Rapee 2004). Juveniles were asked to indicate on a 5-point scale, ranging from 0=*not at all*, to 5 *all the time*, how many times they had a specific thought the past week. The *Personal failure*, 10 items (e.g., I am worthless), α =.81 (T1) and α =.88 (T2), and *Hostility*, 10 items (e.g., when someone hurts me, I have the right to hurt that person), α =.73 (T1) and α =.83 (T2) scales were included in the present study.

Self-esteem was assessed with the self-perception profile for children (Harter 1985). The general self-esteem subscale was used, consisting of 5 items to be answered on a 4-point scale, α =.70 (T1), and α =.73 (T2).

Parenting: parent and adolescent reports

Three theoretically relevant dimensions of parenting: positive discipline, inept discipline and quality of the parent-adolescent relationship, were assessed by both parents and adolescents using multiple indicators. For each dimension, we first selected the

indicators (scales from established measures), and tested the internal consistency of each indicator. To assess the feasibility of combining different indicators of the same construct in a composite score, for each dimension we conducted confirmatory factor analysis (CFA), separately for parent and adolescent. If a single-factor model produced an adequate fit with significant factor loadings of indicators, and the internal consistency of the composite score was satisfactory, the composite score representing that dimension was computed by averaging the scores of the indicators. All items were standardized before computing the composite. Standardization was performed across the full sample and across both time points, so that the relative differences in variability across time were preserved.

Positive discipline The same four indicators of the construct positive discipline were used for both parental and adolescent reports. The first indicator, *consistency*, was assessed with a subscale of the Parenting Dimensions Inventory (PDI) (Deković et al. 2003; Slater and Power 1987), consisting of 8 items to be rated on a 6-point scale (1 = *I totally disagree* to 6 = *I totally agree*). The second indicator, *behavioral control*, was assessed with six items of Parenting Practices questionnaire (Kerr and Stattin 2000), ranging from 1 = *never*, to 5 = *always*. Third, parents and adolescents were presented with three hypothetical situations from the PDI, describing adolescent misbehavior followed by several possible parental reactions. They were asked to indicate how probable (1 = *very improbable* to 6 = *very probable*) it was that parents would use each reaction. Two items describe parental use of *inductive discipline* (i.e. explaining the reasons for rules and pointing out consequences of misbehaviors for adolescent or for others). A mean score across three situations was calculated. Finally, parental *monitoring* (i.e. knowledge regarding adolescent whereabouts) behavior was assessed with 6 items, to be answered on a 4-point scale, ranging from 1 = *I do not know anything about this*, to 4 = *I know all about this* (Brown et al. 1993).

A confirmatory factor analysis (CFA, LISREL 8.80) was used to evaluate whether the indicators measure the same construct. The CFA, based on the covariance matrix and using maximum likelihood estimation, showed an adequate fit to the data for both parent [$\chi^2(6)=7.85$, $p=.250$, RMSEA=.035, CFI=.967], and adolescent reports [$\chi^2(6)=3.10$, $p=.796$, RMSEA=.000, CFI=1.000], with all estimated factor loadings being significant. The composite alphas for the parent were .85 (T1) and .86 (T2), and for the adolescent .82 (T1) and .78 (T2).

Inept discipline Two indicators of inept discipline were the same for both parents and adolescents, and both were based on the responses on three hypothetical situations from the PDI. First, *harsh discipline*, was the mean score across situations for the following reactions: yelling and physical punishment. The second indicator, *other punishments*, included ignoring, love withdrawal, and exercise of power. Third indicator was adolescent report on the level of *psychological control* used by the parents. The Psychological Control Scale, Youth Report (Barber 2002) consists of 8 items, to be answered on a 6-point scale, ranging from 1 = *completely disagree* to 6 = *completely agree*.

The CFA on adolescent data, with one error covariance being constrained to provide one degree of freedom to allow fit indexes to be estimated, showed an adequate fit [$\chi^2(4)=0.91$, $p=.916$, RMSEA=.000, CFI=1.00]. The fit of the parent CFA model could not be meaningfully estimated (i.e. it was perfect), because there

were only two indicators of the latent variable. All of the estimated factor loadings were significant in both models. The composite alpha's were .77 (T1) and .84 (T2) for the parent, and .86 (T1) and .87 (T2) for the adolescent.

Quality of parent-adolescent relationship The parent's perception of the quality of parent-adolescent relationship was assessed by four indicators. The first indicator, *responsiveness*, is an 8 item scale of the Nijmegen Parenting Questionnaire (Gerris et al. 1993), rated on a 6-point scale (1=*I totally disagree* to 6=*I totally agree*). The second indicator, *acceptance of the child*, was measured with a 12 item scale from the Parenting Stress Index (Abidin 1983), rated on a 6-point scale (1=*I totally disagree* to 6=*I totally agree*). Third, the degree of *conflict and antagonism* in the parent-adolescent relationship, was assessed with 6 items from the Network of the Relationship Inventory (NRI) (Furman 1996) rated on a 5-point scale (1=*little or none* to 5=*very much*). Fourth, the *problems in communication* were assessed with 5 items to be answered on a 6-point scale (1=*completely disagree* to 6=*completely agree*), from the Parent-adolescent Communication Scale (Barnes and Olson 1985). For adolescent perception of the quality of parent-adolescent relationship, the last two indicators were also used. In addition, *attachment* of the adolescent to the parent was assessed with the Inventory of Parent and Peer Attachment (IPPA) (Nada Raja et al. 1992), consisting of 12 items, to be answered on a 4-point scale (1=*hardly ever* to 4=*almost always*).

The CFA on four indicators in parent data yielded an adequate fit, [$\chi^2(6)=4.08$, $p=.666$, RMSEA=.000, CFI=1.000]. This was also the case for the adolescent model with three indicators [$\chi^2(4)=8.10$, $p=.088$, RMSEA=.003, CFI=.976]. All estimated factor loadings were significant. The alpha for the composite score *Positive quality of adolescent relationship* for the parent were .79 (T1) and .85 (T2), and for the adolescent .69 (T1) and .63 (T2).

Parenting: observational data

The observational assessment of parenting was based on the Coder Impressions Inventory (CII – 81 items) (Webster-Stratton 1998). The CII has been used for home observations (e.g., Webster-Stratton 1998) and is based on the observers' overall impressions of the parent, child and their interactions during unstructured home observation. It is designed to collect information that may be missed in structured questionnaires or strictly defined observational procedures. The coders were trained extensively until they achieved agreement of >80% during training. The research assistants observed parent and adolescent couples during home visits which lasted for about 1½ h. Immediately after a home visit, the research staff coded 14 items assessing *positive parenting behavior* (alpha's: .93 at T1 and .78 at T2) on a 3-point scale (1=*did not occur* to 3=*four or more examples*). The observers' ratings of 26 items from the CII, tapping harsh/critical parenting and lax/permissive parenting, were used as an observed measure of *inept discipline*. Alpha were .94 (T1) and .83 (T2). Finally, the observers' ratings of 20 items from the CII, tapping parental responsiveness/nurturance and child bonding with parent were used as an observed measure of *positive quality of the parent-adolescent relationship*. Alpha's were .88 (T1) and .78 (T2).

Peer relationships

Involvement with deviant peers Involvement with deviant peers was assessed with a combination of the trouble subscale of the peers scale of the Family, Friends, and Self Scale (Simpson and McBride 1992) (4 items) and with the deviant peers subscale of the Basic Peer Questionnaire by Weerman and Smeenk (2005, 7 items). These 11 items assess the amount of deviancy in the peer network of the adolescent (e.g., “How many of your friends have stolen something?”), had to be answered on a 5-point scale, ranging from (1=*none* to 5=*almost all*). Cronbach’s alpha was .91 (T1) and .90 (T2)

Involvement with prosocial peers Prosocial peer involvement was assessed with the Family, Friends, and Self Scale (Simpson and McBride 1992) with a scale consisting of 7 items (e.g., “How many of your friends get good grades in school?”), to be answered on a 5-point scale, ranging from (1=*none* to 5=*almost all*). Cronbach’s alpha were .84 (T1) and .85 (T2).

Treatment adherence Treatment adherence was assessed with the 15 item Treatment Adherence Measure (TAM) (Henggeler and Borduin 1992; Huey et al. 2000). This version was also used in the Swedish MST study (Sundell et al. 2008). This instrument assesses the adherence to the nine MST principles according to the parents, who are phoned monthly to assess treatment adherence. Items were rated on a 5-point scale (1=*not at all* to 5=*very much*). The scores were averaged per family to indicate the mean level of therapist adherence experienced by a family. The mean adherence score ranged from 4.14 to 4.58 (M=4.36, SD=.51) indicating that treatment adherence was satisfactory and comparable to the TAM scores in the US (M=4.41, SD=.49 ; Letourneau et al. 2002) and higher than reported in the Swedish study (M=4.00; SD=.61) (Sundell et al. 2008).

Analytic strategy

First, overall effectiveness was examined for all outcome measures by conducting an ANCOVA, with the outcome measures at post-test as dependent variables, treatment condition as factor and pre-intervention scores of the outcome variables as co-variates.

For the moderator analyses, the same ANCOVA’s were conducted, with the moderators as factor. Posthoc analyses for moderator effects were conducted by splitting the file according to the moderator and again conducting an ANCOVA and calculating effect sizes separately for each group. Effect sizes were computed as *d*, based on adjusted means and standard errors, with a positive sign indicating improvement in the MST group relative to the control group.

Results

Intervention effects

Externalizing problems and delinquent behavior Table 1 shows that for both parent and adolescent reported externalizing behavior problems, a significant effect was

Table 1 Means and standard deviations for MST and TAU groups pre- and post-intervention

	MST				TAU				<i>F</i> for group	<i>d</i> (95% CI)
	T1		T2		T1		T2			
	M	SD	M	SD	M	SD	M	SD		
Externalizing/delinquency										
Externalizing problems (P)	23.32	12.60	17.64	11.57	22.55	12.95	19.25	10.56	4.21*	.26 (.01, .51)
ODD (P)	2.03	.84	1.78	.71	1.93	.75	1.91	.68	7.86**	.36 (.11, .61)
Conduct Disorder (P)	1.36	.38	1.21	.26	1.34	.29	1.28	.29	5.54*	.30 (.05, .55)
Externalizing problems (A)	12.40	9.25	10.39	7.92	12.36	8.32	11.95	7.56	4.17*	.26 (.01, .51)
Violence (A)	.38	.58	.33	.52	.36	.57	.31	.53	.06	.03 (.22, .28)
Property offences (A)	.31	.43	.15	.19	.29	.45	.22	.39	4.12*	.25 (.01, .50)
Cognitions										
Sense of competence (P)	4.02	.87	4.32	.94	4.24	.91	4.20	1.05	8.17**	.36 (.11, .61)
Self-esteem (A)	3.09	.64	3.01	.63	3.00	.62	2.87	.80	1.90	.18 (-.07, .42)
Personal failure (A)	.28	.41	.36	.55	.34	.46	.24	.47	4.18*	-.26 (-.51, -.01)
Hostility (A)	.91	.69	.60	.57	.89	.67	.73	.62	4.13*	.26 (.01, .51)
Parenting										
Positive discipline (P) ^a	.03	.64	.11	.62	-.03	.64	-.15	.64	13.13***	.47 (.21, .71)
Inept discipline (P) ^a	-.03	.76	-.07	.81	.11	.81	.02	.85	.06	.03 (-.22, .28)
Quality of relationship (P) ^a	.01	.26	.02	.24	-.00	.30	-.04	.24	6.03*	.31 (.07, .56)
Positive discipline (A) ^a	.07	.63	.00	.56	.05	.58	-.15	.67	5.07*	.28 (.04, .53)
Inept discipline (A) ^a	.02	.69	-.01	.75	.08	.74	-.09	.71	1.71	-.17 (-.40, .08)
Quality of relationship (A) ^a	-.00	.38	.01	.35	.01	.34	-.02	.33	.60	.10 (-.15, .35)
Positive discipline (O)	1.82	1.01	2.03	.95	2.01	.92	1.89	.83	6.90**	.33 (.08, .58)
Inept discipline (O)	1.16	.58	1.10	.64	1.30	.57	1.32	.55	4.58*	.27 (.02, .52)
Quality of relationship (O)	1.82	.95	1.92	.91	1.94	.84	1.81	.71	8.72**	.37 (.12, .62)
Peers										
Deviant peers (A)	1.82	.76	1.62	.63	1.91	.85	1.65	.66	.01	.05 (-.24, .26)
Prosocial peers (A)	3.14	.84	3.20	.77	3.00	.90	2.87	.83	9.09**	.41 (.13, .63)

P Parent report, *A* adolescent report, *O* observations

* $p < .05$, ** $p < .01$, *** $p < .001$

^a Composite score based on standardized indicators

found, indicating that MST was more effective than TAU in decreasing externalizing behavior problems. Additionally, MST was more effective than TAU in decreasing Oppositional Defiant Disorder, Conduct Disorder, and property offences. For violence, no intervention effect was found.

Parent and adolescent cognitions A significant intervention effect was found for three, out of four, measures of parent and adolescent cognitions. Parents in MST group reported a larger increase in sense of competence than the parents from TAU group, where sense of competence slightly decreased. Adolescents in the MST group reported a larger decrease in hostility than those in the TAU group. In the MST group, significant increases of personal failure were reported by the adolescents, whereas, in the TAU group, personal failure decreased. There were no significant changes in adolescent self-esteem in both groups.

Parenting behavior For positive discipline, MST participants showed more improvements than TAU participants, according to parents, adolescents and raters. For relationship quality, MST was effective according to parents and observers, but not according to adolescents, for whom no differences between MST and TAU were found. For inept discipline, no intervention effects were found for parent and adolescent reports, but an effect was found for observed inept discipline MST was more effective than TAU.

Peer relationships Finally, the effects of MST on peer relationships were assessed. These analyses indicated that MST was not more effective in changing the affiliation with deviant peers than TAU: in both groups there was a significant decrease in involvement with deviant peers (MST: $F=11.78$, $p<.001$; TAU: $F=12.66$, $p<.001$). MST did, however, affect the association with prosocial peers: in the MST group, there was an increase in contacts with prosocial peers, whereas there was a decrease in the TAU group.

Moderators of effectiveness

In order to examine whether demographic characteristics affected the effectiveness of MST, three moderator effects were tested.

Ethnicity Moderator analyses of ethnicity revealed no significant interactions between ethnicity and condition, with one exception: there was a significant interaction between ethnicity and condition for parent reported externalizing behavior problems ($F(4,251)=6.01$, $p<.05$). Post hoc analyses indicated that the intervention effects were stronger for the native Dutch juveniles ($F(1,138)=9.48$, $p<.01$, $d=.53$) than for ethnic minority groups ($F(1,112)=.26$, $p=.61$, $d=.10$). As this was the only moderator effect for ethnicity, it seems that, apart from the parent reported behavior problems score, MST was equally effective for different ethnic groups.

Gender Moderator analyses revealed several moderator effects for gender. A significant interaction between gender and condition was found for self-esteem ($F(4, 251)=11.33$, $p<.001$). Post-hoc analyses showed significant intervention effects for boys (F

(1, 185)=7.43, $p < .01$, $d = .40$): boys in TAU group showed decreased self-esteem, whereas boys in MST group showed no significant changes in self-esteem between T1 and T2. The intervention effect was also significant for girls ($F(1, 65)=7.05$, $p < .05$, $d = -.34$), but the direction of this effect was different than for boys: between T1 and T2, self-esteem in girls increased in the TAU group, whereas it decreased in the MST group. Additionally, a significant gender \times condition interaction was found for personal failure ($F(4, 251)=4.05$, $p < .05$). Posthoc analyses revealed that the effect size for girls ($F(1, 65)=8.21$, $p < .01$, $d = -.74$) was larger than for boys ($F(1, 185)=.36$, $p = .55$, $d = .09$), indicating that boys in both the MST and TAU group showed hardly any change in personal failure, whereas personal failure of girls in the MST group increased and personal failure in the TAU group decreased. Finally, a significant gender \times condition effect was found for hostility ($F(4, 251)=4.48$, $p < .05$). Post hoc analyses showed that the effectiveness of MST was larger for boys ($F(1, 185)=8.02$, $p < .01$, $d = .42$) than for girls ($F(1, 65)=.65$, $p = .42$, $d = .21$). Boys in the MST group showed a larger decrease in hostility than boys in the comparison group, whereas girls in MST and TAU showed a similar decrease in hostility.

Age In order to find out whether younger or older adolescents benefited most from MST, the group was subdivided into a group of juveniles younger than 16 years of age ($n=54$) and in a group of juveniles of 16 and older ($n=189$). No moderating effects were found for age.

Discussion

The present study examined the effects of MST in The Netherlands on problem behavior, adolescent's cognitions, parenting, and relations with peers. It showed that MST was more effective than TAU in changing externalizing behavior problems, ODD and CD symptoms, and involvement in property offences. No significant effects were found for adolescent reported violent offences.

Additionally, MST was more effective than TAU in increasing parental competence and decreasing adolescent hostility. However, no intervention effect was found for adolescent self-esteem, and an unexpected negative effect was found for personal failure: MST was less effective than TAU in decreasing personal failure, which increased in the MST group and decreased in the TAU group. Effects for parenting showed that MST was more effective than TAU in the sense that positive dimensions of parenting (positive discipline and relationship quality) increased in the MST group, whereas they decreased in the TAU group. Adolescent reported positive discipline decreased in both the TAU and MST group; however, the decrease was larger in the TAU group. MST seems effective in decreasing observed inept discipline, but no differences were found between MST and TAU in parent and adolescent reported inept discipline. Finally, MST was more effective than TAU in increasing associations with prosocial peers, but no differences between MST and TAU were found in the decrease of the relationships with deviant peers.

When comparing the results of the present study to previous research on MST, it seems that, despite some variation in effect sizes, the Dutch MST results are

comparable to the American and Norwegian studies (Curtis, et al. 2004; Ogden and Halliday-Boykins 2004), although, generally, the effect sizes are somewhat smaller than in those studies. In the present study, as in the American studies, there were no indications that the effectiveness of MST varied between participants of different ethnic origins (with the exception of parent reported externalizing behavior), despite the fact that in The Netherlands ethnic minority groups differ from the ethnic minority groups in the USA. So differences in study population cannot be an explanation for the somewhat smaller effect sizes. In the present study, treatment adherence was comparable to the American studies (Curtis et al. 2004), suggesting that the implementation of MST in The Netherlands has been successful. Thus lower treatment adherence cannot explain the smaller effect sizes either. In contrast, differences between countries in the quality of youth services, and consequently, the control condition may explain the differences in magnitude of the effect sizes. In The Netherlands, as in Sweden, in-home services are used more frequently than in the USA, where juvenile offenders are most likely to be treated within the juvenile justice system (Sundell et al. 2008; Lipsey 1999). It therefore seems that the quality of TAU may be crucial for establishing effectiveness of MST, in the sense that when, more comprehensive in-home services are offered as TAU, the effects of MST are likely to be smaller than when MST is being compared to residential care within the juvenile justice system. In sum, differences between countries in the quality of youth services, and consequently in the quality of treatment offered to the control group, can affect the results of effectiveness studies. Failure to draw attention to this point (at least from a policy perspective) can lead to misinterpretation of the results of studies and, consequently, undermine an effective program.

The present study extended previous research by including additional outcome variables which may have been influenced by the multisystemic approach of MST, such as parenting variables, cognitions and contacts with prosocial peers. For the parenting variables, effects were found for parental competence, supportive parenting and relationship quality. The positive findings for parental competence and parenting behaviors are in line with the focus of MST on family interventions. Explanations for the positive results on parenting may be the intensity of the contacts between therapists and parents, and the 24/7 therapist availability, which may have empowered parents to improve their parenting skills. No parent or adolescent reported effects were found for inept discipline. Possibly, MST therapists focus more on changing positive behaviors than on negative ones. Perhaps, changing inept discipline behaviors is something MST therapists should focus on more, as adolescents and parents reported it as related to externalizing behavior problems at pre-test.

In the present study, adolescent cognitions were included as outcome measures, given their association with the development of delinquent behavior (Barriga et al. 2001; Gibbs 2003). The present study revealed an effect of MST for hostility. Changing hostile attributions may thus be an important first step in decreasing aggressive, violent delinquent behavior (Schniering and Rapee 2004; Beck 1999). Noteworthy is the finding that MST was less effective than TAU in decreasing personal failure; the sense of inability to solve one's problems. Personal failure has been associated with feelings of depression (Schniering and Rapee 2004). This could mean that adolescents developed some awareness of the seriousness of their situation as a consequence of MST, although apparently there was no concurrent effect on the

adolescents' self-esteem. Acknowledging the problems and attributing those to themselves may be a first step in behavior change.

Finally, it was remarkable that no intervention effect was found for deviant peers, while MST did have more positive effects on involvement with prosocial peers than TAU. Apparently, the MST adolescents were open to more positive peer contact but did not show a decrease in involvement with deviant peers more than TAU adolescents. Possibly, MST treatment could be improved by focusing more explicitly on diminishing deviant peer contacts. Apparently TAU is also effective in diminishing contacts with deviant peers.

Several moderator effects were found; however, these effects were not consistent over the various outcome variables. The only consistent pattern that appeared were gender differences regarding adolescent cognitions, with MST showing larger (and more positive) effects for boys than for girls. Girls in MST group showed decreases in self-esteem after treatment, increases in personal failure, and a similar degree of decrease in hostility as girls from the control group. It thus seems that, for girls, MST is less effective in changing cognitions than for boys, and it might even produce some undesirable effects. This may be due to the fact that girls are less often referred to programs. Therefore, therapists may have been less equipped to change cognitions in girls. Although these findings need to be replicated, it might be worthwhile to give specific attention to symptoms of depression and personal failure in girls in treatment programs.

As only a few moderator effects have been found, these results should be interpreted with caution. The non-significant moderator effects that were found in the present study are in line with previous studies of MST (e.g., Henggeler et al. 2002) who did not find any or few (Sundell et al. 2008) moderator effects and concluded that the effects of MST did not vary as a consequence of demographic characteristics.

A first limitation of the present study is that the focus was on the short-term effectiveness of MST immediately after treatment. A follow-up assessment is needed, which should include adolescent and parent reports, observational data and official judicial records to assess recidivism. As a decrease in recidivism is the ultimate goal of MST, it is important to include this outcome measure in future studies.

A second limitation of the present study is that the power in some of the subgroups that were examined in the moderator analyses may have been too low to detect subgroup effects (e.g., the sample size for females was $n=68$). This may have resulted in an underestimation of the gender effects. However, those moderator effects of gender that have been found, seem quite robust.

Finally, to assess treatment integrity we relied on a sole measure (i.e., TAM). Although several studies provided evidence for the validity of this measure (Henggeler et al. 2002; Huey et al. 2000; Schoenwald et al. 2000, 2003, 2005), it must be recognized that TAM assesses therapist adherence to the nine principles of MST, covering aspects that might not be unique to MST (such as client satisfaction with services, bonding with the therapist and agreement on the goals and tasks of treatment), rather than assessing whether a specific component or a specific intervention technique is implemented with a particular level of quality. In future studies, more effort should be made to assess the treatment integrity with a more specific instrument.

In summary, the present study is the first RCT of MST in a Dutch sample. The sample size is larger than the most of the previous MST trials, which increases the

power to detect treatment effects. Moreover, results based on youth and parent reports often converged, showing that the results were fairly consistent across informants. This study showed positive outcomes for MST in the European context with a mixed ethnicity sample. Positive changes in problem behavior of the juveniles have been found as well as changes in parent and adolescent cognitions, parenting behavior and involvement with prosocial peers. It thus seems that MST indeed is able to empower parents to make changes in their parenting behavior and consequently diminish a number of risk factors for delinquent behavior in the juvenile's environment. However, additional long-term cost-effectiveness or cost-benefit studies, and studies including recidivism data are needed to make a definite statement on the value of MST in the Dutch context. Aos et al. (2004) and Welsh et al. (2012) showed that in the US, MST is cost-beneficial. These findings should be replicated in the Dutch context.

Acknowledgments This work was supported by grants from the Netherlands Organisation for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW). We thank the therapists and staff of De Waag and Yorneo for their cooperation, and the adolescents and families who participated in the project.

References

- Abidin, R. R. (Ed.). (1983). *Parenting stress index manual*. Charlottesville: Pediatric Psychology Press.
- Achenbach, T. M. (1991). *Manual for the youth self-report and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia: Washington State Institute for Public Policy.
- Asscher, J. J., Deković, M., Van der Laan, P. H., Prins, P. J. M., & Van Arum, S. (2007). Implementing randomized experiments in criminal justice settings: an evaluation of multisystemic therapy (MST) in The Netherlands. *Journal of Experimental Criminology*, 3, 113–129.
- Barber, B. K. (2002). *Intrusive parenting. How psychological control affects children and adolescents*. Washington, DC: APAO.
- Barnes, H., & Olson, D. H. (1985). Parent-adolescent communication and the circumplex model. *Child Development*, 56, 438–447.
- Barriga, A. Q., Gibbs, J. C., Potter, G. B., & Liau, A. K. (2001). *Test manual for the how I think questionnaire*. Champaign: Research Press.
- Beck, A. T. (1999). *Prisoners of hate. The cognitive basis of anger, hostility, and violence*. New York: Harper Collins.
- Blokland, A., Grimbergen, K., Bernasco, W., & Nieuwbeerta, P. (2010). Criminaliteit en etniciteit. Criminele carrières van autochtone en allochtone jongeren uit het geboortecohort 1984. *Tijdschrift voor Criminologie*, 52, 122–152 [Criminality and ethnicity: criminal careers of Dutch and migrant juveniles from birth cohort 1984].
- Blom, M., & van der Laan, A.M. (2006). *Monitor Jeugd Terecht*. Retrieved November, 24, 2009 from <http://www.wodc.nl/onderzoeksdatabase/monitor-jeugd-terecht-2006.aspx>.
- Borduin, C. M., Schaefer, C. M., & Heblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile sex offenders: effects on youth social ecology and criminal activity. *Journal of Consulting and Clinical Psychology*, 77, 26–37.
- Brown, B. B., Mounts, N., Lamborn, S. D., & Steinberg, L. (1993). Parenting practices and peer group affiliation in adolescence. *Child Development*, 64, 467–482.
- Brunk, M., Henggeler, S. W., & Whelan, J. P. (1987). Comparison of multisystemic therapy and parent training in the brief treatment of child abuse and neglect. *Journal of Consulting and Clinical Psychology*, 55, 171–178.

- Butler, S., Baruch, G., Hickey, N., & Fonagy, P. (2011). A randomized controlled trial of multisystemic therapy and a statutory therapeutic intervention for young offenders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*, 1220–1235.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.
- Coleman, P. K., & Karraker, K. H. (1997). Self-efficacy and parenting quality: findings and future applications. *Developmental Review*, *18*, 47–85.
- Cunningham, A. J. (2002). *One step forward: lessons learned from a randomized study of multisystemic therapy in Canada*. Praxis: Research from the Centre for Children & Families in the Justice System. June 2002.
- 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.
- Deković, M., Janssens, J. M. A. M., & van As, N. M. C. (2003). Family predictors of antisocial behavior in adolescence. *Family Process*, *42*, 223–236.
- Deković, M., Asscher, A. A., Hermanns, J., Reitz, E., Prinzie, P., van den Akker, A. L. (2010). Tracing changes in families who participated in the Home-Start Parenting Program: parental sense of competence as mechanism of change. *Prevention Science*, *11*, 263–274
- Donnellan, M. B., Trzeniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science*, *16*, 328–335.
- Donovan, W. L., Leavitt, L. A., & Walsh, R. O. (1990). Maternal self-efficacy, Illusory control and its effect on susceptibility to learned helplessness. *Child Development*, *61*, 1638–1647.
- Emeka, T. Q., & Sorensen, J. R. (2009). Female juvenile risk: is there a need for gendered assessment instruments? *Youth Violence and Juvenile Justice*, *7*, 313–330.
- Furman, W. (1996). The measurement of friendship perceptions: Conceptual and methodological issues. In W. M. Bukowski & A. F. Newcomb (Eds.), *The company they keep: friendship in childhood and adolescence*. Cambridge studies in social and emotional development (pp. 41–65). New York: Cambridge University Press.
- Gerris, J. R. M., et al. (1993). *Parents, adolescents, and young adults in dutch families: a longitudinal study*. Nijmegen: University of Nijmegen.
- Gibbs, J. C. (2003). *Moral development and reality*. Thousand Oaks: Sage.
- Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annual Review of Psychology*, *60*, 549–576.
- Harter, S. (1985). *Manual for the self-perception profile for children*. Denver: University of Denver.
- Henggeler, S. W. (2011). Efficacy studies to large-scale transport: the development and validation of multisystemic therapy programs. *Annual Review of Clinical Psychology*, *7*, 351–381.
- Henggeler, S. W., & Borduin, C. M. (1992). *Multisystemic therapy adherence measures*. Unpublished Instrument. Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina.
- Henggeler, S. W., & Lee, T. G. (2003). Multisystemic treatment of serious clinical problems. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 301–322). New York: Guilford Press.
- Henggeler, S. W., & Schaeffer, C. M. (2010). Treating serious emotional and behavioral problems using multisystemic therapy. *Australian and New Zealand Journal of Family Therapy*, *31*, 149–164.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: an effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, *60*, 953–961.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent chronic juvenile offenders and their families: the role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology*, *65*, 821–833.
- Henggeler, S. W., Clingempeel, W. G., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance abusing and substance dependent juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 868–874.
- Huey, S. J., Henggeler, S. W., Brondino, M. J., & Pickrel, S. G. (2000). *Treatment Adherence Measure (TAM)*. Unpublished manuscript.
- Kerr, M., & Stattin, H. (2000). What parents know, how they know it, and several forms of adolescent adjustment: further support for a reinterpretation of monitoring. *Developmental Psychology*, *36*, 366–380.
- Leschied, A. W., & Cunningham, A. (2002). *Seeking effective interventions for serious young offenders. Interim results of a four-year randomized study of multisystemic therapy in Ontario, Canada*. London (CA): Centre for children and families in the justice system of the London Family Court Clinic.

- Letourneau, E. J., Sheidow, A. J., & Schoenwald, S. K. (2002). *Structure and reliability of the MST therapist adherence scale in a large community sample*. Family Services Research Center, Medical University of South Carolina.
- Lipsey, M. W. (1999). Can intervention rehabilitate serious delinquents? *Annals of the American Academy of Political and Social Research*, *564*, 142–166.
- Littell, J. H., Popa, M., & Forsythe, B. (Eds.). (2005). *Multisystemic therapy for social, emotional, and behavior problems in youth age 10–17*. *Cochrane Library*, issue 3. Chichester: John Wiley & Sons Ltd.
- Mason, W. A. (2001). Self-esteem and delinquency revisited (again): a test of Kaplan's self-derogation theory of delinquency using latent growth curve modeling. *Journal of Youth and Adolescence*, *30*, 83–102.
- MST services (2011). Multisystemic therapy (MST). Brief description, retrieved from http://www.mstservices.com/pa_mst_medicaid_standard.doc, January 2011.
- Nada Raja, S., McGee, R., & Stanton, W. R. (1992). Perceived attachments to parents and peers and psychological well-being in adolescence. *Journal of Youth and Adolescence*, *21*, 471–485.
- Ogden, T., & 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*, 77–83.
- Oosterlaan, J., Scheres, A., Antrop, I., Roeyers, H., & Sergeant, J. A. (2000). *Vragenlijst voor Gedragsproblemen bij Kinderen (VvGK)*. Handleiding. Lisse: Swets & Zeitlinger.
- Potier, J., & Day, C. (2007). Childhood onset conduct problems: a preliminary investigation into the role of mothers' interpersonal schemas and their relationship to parenting behaviour. *Behavioural and Cognitive Psychotherapy*, *35*, 457–472.
- Rowland, M. D., Halliday-Boykins, C. A., Henggeler, S. W., Cunningham, P. B., Lee, T. G., Kruesi, M. J. P., et al. (2005). A randomized trial of multisystemic therapy with Hawaii's Felix Class Youths. *Journal of Emotional and Behavioral Disorders*, *13*, 13–23.
- Schaeffer, C. M., & Borduin, C. M. (2005). Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology*, *73*, 445–453.
- Schneider, M. J., & Leitenberg, H. (1989). A comparison of aggressive and withdrawn children's self-esteem, optimism and pessimism, and causal attributions for success and failure. *Journal of Abnormal Child Psychology*, *17*, 133–144.
- Schniering, C. A., & Rapee, R. M. (2004). The structure of negative self-statements in children and adolescents: a confirmatory factor-analytic approach. *Journal of Abnormal Child Psychology*, *32*, 95–109.
- Schoenwald, S. K., Henggeler, S. W., Brondino, M. J., & Rowland, M. D. (2000). Multisystemic therapy: monitoring treatment fidelity. *Family Process*, *39*, 83–103.
- Schoenwald, S. K., Sheidow, A. J., Letourneau, E. J., & Liao, J. G. (2003). Transportability of multisystemic therapy: evidence for multi-level influences. *Mental Health Services Research*, *5*, 223–239.
- Schoenwald, S. K., Sheidow, A. J., & Letourneau, E. J. (2004). Toward effective quality assurance in evidence-based practice: links between expert consultation, therapist fidelity, and child outcomes. *Journal of Clinical Child and Adolescent Psychology*, *33*, 94–104.
- Schoenwald, S. K., Letourneau, E. J., & Halliday-Boykins, C. A. (2005). Predicting adherence to a transported family-based treatment for youth. *Journal of Clinical Child and Adolescent Psychology*, *34*, 658–670.
- Simpson, D. D., & McBride A. A. (1992). Family, friends, and self (FFS) assessment scales for Mexican American youth. *Hispanic Journal of Behavioral Sciences*, *14*, 327–340.
- Slater, M. A., & Power, T. G. (1987). Multidimensional assessment of parenting in single-parent families. *Advances in Family Intervention, Assessment, and Theory*, *4*, 197–228.
- Sundell, K., Hansson, K., Löfholm, C. A., Olsson, T., Gustle, L. H., & Kadesjö, C. (2008). The transportability of multisystemic therapy to Sweden: short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, *22*, 550–560.
- Timmons-Mitchell, J., Bender, M. B., Kishna, M. A., & Mitchell, C. C. (2006). An independent effectiveness trial of multisystemic therapy with juvenile justice youth. *Journal of Clinical Child and Adolescent Psychology*, *35*, 227–236.
- Van der Put, C. E., Deković, M., Stams, G. J. J. M., Van der Laan, P. H., Hoeve, M., & Van Amelsfort, L. (2011). Changes in risk factors during adolescence: implications for risk assessment. *Criminal Justice and Behavior*, *38*, 248–262.
- Van Vugt, E. S., Hendriks, J., Stams, G. J. J. M., Van Exter, F., Bijleveld, C., van der Laan, P., & Asscher, J. J. (2011). Moral judgment, cognitive distortions and moral reasoning in juvenile sex offenders. *Forensic Psychiatry, & Psychology*, *22*, 603–619.

- Verhulst, F., & Van der Ende, J. (1992). Agreement between parents' and adolescents' self-reports of problem behavior. *Journal of Child Psychology and Psychiatry*, 33, 1011–1023.
- Verhulst, F., Koot, J. M., Akkerhuis, G. W., & Veerman, J. W. (1990). *Praktische handleiding voor de CBCL. [CBCL manual]* Van Gorcum, Assen/Maastricht.
- Webster-Stratton, C. (1998). Preventing conduct problems in head start children: strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, 66, 715–730.
- Weerman, F. M., & Smeenk, W. H. (2005). Peer similarity in delinquency of different types of friends: a comparison using two measurement methods. *Criminology*, 43, 499–524.
- Welsh, B. C., Lipsey, M. W., Rivara, F. P., Hawkins, J. D., Aos, S., & Hollis-Peel, M. E. (2012). Promoting change, changing lives: Effective prevention and intervention to reduce serious offending. In R. Loeber & D. P. Farrington (Eds.), *From juvenile delinquency to adult crime: criminal careers, justice policy, and prevention* (pp. 245–277). New York: Oxford University Press.
- Zahn, M. A., Day, J. C., Mihalic, S. F., & Tichavsky, L. (2009). Determining what works for girls in the Juvenile justice system a summary of evaluation evidence. *Crime & Delinquency*, 55, 266–293.

Jessica J. Asscher, developmental psychologist, is an assistant professor Forensic Child and Youth Care Sciences at the University of Amsterdam. Her research interests concern the effectiveness of (juvenile justice) interventions and the development of behavior problems.

Maja Deković developmental psychologist, is a full professor at Department of Child and Adolescent Studies, Utrecht University and leader of the research programme “Development and treatment of psychosocial problemst”. Her research interests include the development of problem behaviour, parent–adolescent relationships, family interaction and effects of family-based interventions.

Willeke A. Manders studied Psychology at Tilburg University. She obtained her PhD at the Radboud University for a study on families and adolescent problem behaviour. Currently she is conducting a study on the implementation and effectiveness of Multisystemic Therapy in the Netherlands as a postdoctoral fellow at Utrecht University. Her research interests are family assessment, the development of problem behavior, and the implementation and effectiveness of interventions.

Peter H. van der Laan, Ph.D., is Professor of Probation and Parole at the VU University Amsterdam, The Netherlands, and is Senior Researcher at the Netherlands Institute for the Study of Crime and Law enforcement (NSCR), Amsterdam. His research interests regard juvenile delinquency, criminal justice issues and the effectiveness of criminal justice interventions.

Prof Dr. Pier J. M. Prins is a full professor of child clinical psychology at the Department of Developmental Psychology at the University of Amsterdam. Prof. Prins is director of the research program “Mechanisms of Change and Psychosocial Treatments for Behaviorally and Emotionally Disordered Youth” at the University of Amsterdam. Currently, he is involved in the multisite Amsterdam-Groningen Study evaluating moderators and mediators of CBT with anxiety disordered children and adolescents. He is also involved in the development and evaluation of a computer game to remediate executive functions in children with ADHD. Pier Prins is a licensed GZ-psychologist and also a licensed behavior therapist.