

Implementing randomized experiments in criminal justice settings: An evaluation of multi-systemic therapy in the Netherlands

Jessica J. Asscher · Maja Deković ·
Peter H. van der Laan · Pier J. M. Prins ·
Sander van Arum

Published online: 3 May 2007
© Springer Science + Business Media B.V. 2007

Abstract This paper addresses several issues that we encountered when implementing a randomized experiment to evaluate a promising intervention for serious juvenile delinquents, multi-systemic therapy (MST). The issues overlap with issues others experienced when trying to conduct randomized experiments in clinical practice, but are complicated because MST takes place in a legal context. The complex juvenile justice system makes it difficult to come up with a simple system of randomization. Problems encountered are the complexity of the referral process in this type of intervention, legal issues and the ethical concerns raised by clinicians. Our experiences might be useful to others planning to carry out a randomized efficacy study within the juvenile justice system. Firstly, other researchers might feel supported if they realize that others experience similar problems, and, secondly, we provide some concrete suggestions concerning the following: selecting the procedures and determining the moment of randomization, dealing with and overcoming the resistance of institutions involved, maintaining the cooperation of both the institution and the sample, and keeping turnover of research staff to a minimum.

Keywords Criminal justice · Delinquency · Evaluation · Multi-systemic therapy · Randomized experiments

J. J. Asscher (✉) · M. Deković
Faculty of Social Sciences, Research Centre Psychosocial Development in Context,
Utrecht University, P. O. Box 80140, 3508 TC Utrecht, Netherlands
e-mail: J.J.Asscher@fss.uu.nl

P. H. van der Laan · P. J. M. Prins
University of Amsterdam, Amsterdam, Netherlands

P. H. van der Laan
NSCR, Leiden, Netherlands

S. van Arum
De Waag, Utrecht, Netherlands

In a recent review of experimental studies in the field of criminology, Farrington and Welsh (2005) pointed out that relatively few randomized experiments in criminology are conducted outside the USA. Although there is a large body of research in Dutch criminology, very few studies, if any, which aim to examine efficacy of interventions to reduce delinquency and recidivism, can be characterized as randomized experiments (Deković et al. 2007; van der Laan 2007). We can only guess at the reasons for the infrequent use of experiments, but we believe that Dutch researchers do not lack awareness of the importance and advantages of this type of design (Bijleveld 2003). The main reason is probably the anticipation of the difficulties the researchers are likely to encounter when implementing such a design in the criminal justice setting, which may lead to the conclusion that it is not worth the trouble. This is consistent with the findings of a study that examined the reasons that researchers of criminal justice evaluations in the USA offer for their choice of methodology (Lum & Yang 2005). There was general consensus among researchers that randomized experimentation is the best approach, but practical reasons (e.g. difficulties in ensuring cooperation from institutions and individuals) impeded researchers from choosing this design.

In the past two years, however, there have been some positive developments in The Netherlands. The need to work in an evidence-based manner has been increasingly emphasized by state health and criminal justice agencies (Ministry of Justice 2003). In 2005 a Best Practices initiative (Ministry of Justice 2005), focusing on identifying the most promising approaches to the prevention of youth delinquency and violence, was launched. These developments are expected to have a positive effect on the willingness of stakeholders to cooperate with researchers in conducting evaluation studies.

In the present paper we describe the implementation process of a randomized experiment that tests the efficacy of multi-systemic therapy (MST), an intervention that aims to prevent recidivism and to decrease antisocial behaviour in delinquent youths. The problems and obstacles encountered during the implementation and possible ways to overcome them are discussed. Although we describe one specific example, we believe that the problems we encountered are relevant for many researchers who plan to carry out a randomized experiment in criminal justice settings.

Background of the evaluation study of MST in the Netherlands

Every year in The Netherlands, some tens of thousands of adolescents between 12 and 17 years of age come into contact with the police because they are suspected of committing a crime. The number of serious delinquents among these youths has increased (van der Laan 2007). The need to develop effective interventions for adolescents who exhibit serious and frequent antisocial behaviour is well established. Unfortunately, many interventions developed for this group have had only limited success, probably due to the narrow focus of these interventions (Kazdin 1997). There are, however, interventions that seem promising. One of these promising interventions, which has recently (in 2004) been introduced in The Netherlands, is MST, developed originally in the USA by Henggeler and colleagues (e.g. Henggeler et al. 1992, 1996, 1997).

Multi-systemic therapy: A description of the intervention

MST uses a home-based model of service delivery, which helps to overcome barriers to services access and increases the chances that families will adhere to the treatment. It addresses several key systems in which the adolescent is embedded: family, school, peer group, neighbourhood. 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 goals of the treatment are family driven, rather than therapist driven, and the treatment itself is highly individualized. Intervention strategies include strategic family therapy, structural family therapy, behavioural parent training, and cognitive therapies. Given the complexity of interventions involving MST, the behaviour of the therapists is governed by a set of nine treatment principles (Henggeler et al. 1997), rather than being delineated explicitly in a treatment protocol. Treatment is typically delivered for 4 to 6 months. Therapists, who have low caseloads (four to six families maximum), are available 24 hours a day, seven days per week.

The treatment quality is monitored closely. The MST therapists receive extensive 5-day training; there is weekly consultation with an MST expert who is external to the programme site (MST consultant) and there are quarterly on-site booster sessions. An additional important feature of MST is the continuous monitoring of treatment integrity, defined as the adherence of the therapist to a set of nine treatment principles specified in a treatment manual (Schoenwald et al. 2000, 2003).

MST seems to be a promising intervention for the following reasons. First, the problems targeted in the intervention are selected on the basis of theories and empirical findings regarding the development and maintenance of antisocial behaviour. Second, MST has several important characteristics, which have been shown in previous research to increase effectiveness. It addresses multiple causes of antisocial behaviour (“broad focus” criterion). The intervention is delivered in the natural environment, i.e. in the context in which problem behaviour occurs (“ecological validity” criterion). Furthermore, MST is intensive, with the therapist being available 24 hours per day (Henggeler et al. 1997, 1999). In short, MST includes all the characteristics (multifaceted, intensive, home- and community-based, with carefully monitored treatment adherence) that are nowadays generally recognized as features of effective, clinically relevant interventions (McGuire and Priestly 1996; Andrews 1996).

Research into the effectiveness of MST

The studies examining the effectiveness of MST indeed show promising results in terms of reducing antisocial behaviour and preventing recidivism in serious adolescent offenders (e.g. Ogden and Halliday-Boykins 2004; Huey et al. 2000; Borduin et al. 1995). Recently conducted meta-analyses of the effectiveness of MST (Curtis et al. 2004), including 11 studies and 708 participants, showed a moderate effect size ($d=0.50$) on indicators of criminal activity.

These positive indications notwithstanding, there are also reasons for caution. A recently published review by Littell (2005) has questioned the conclusion that MST is effective by pointing out that inconsistent and incomplete reports have been published, which sometimes contained unclear randomization procedures and inadequate

analyses. Also, with a few exceptions (Leschied and Cunningham 2002; Ogden and Halliday-Boykins 2004; Timmons-Mitchell et al. 2006), most of the empirical support, including the meta-analysis by Curtis et al. (2004), comes from the studies conducted by the same group of researchers who also developed and implemented MST. A meta-analysis conducted by Littell et al. (2005) showed much less promising results. There is a clear need for confirmation of these results by an independent group of researchers. Even more importantly, MST has been developed and tested in the USA. Although there are no reasons to assume that theories on which MST is based do not adequately explain antisocial behaviour in Dutch adolescents (Deković et al. 2004), due to the differences between the two countries in social and political climate, organization of mental health services, availability of different interventions, type and ethnic background of clients, etc., the question remains whether the same favourable results can be obtained here.

Thus, although MST has fast become *the* solution for certain young people, based on claims in the US literature, it has not yet been demonstrated how well the programme translates to the Dutch culture and Dutch criminal setting. Further evaluations using a randomized controlled trial are indispensable in order to examine the effectiveness of MST compared to available treatments, especially considering the fact that MST is an intensive, and therefore expensive, programme in terms of time and money.

Dutch study of the effectiveness of MST

Soon after the introduction of MST in two large cities, Utrecht and Amsterdam, by De Waag, a clinic for forensic psychiatry in The Netherlands, a team of researchers from Utrecht University, the University of Amsterdam and the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) obtained a grant from the Dutch Organisation for Scientific Research (NWO) to conduct a study of the efficacy of MST in the Netherlands. The first aim was to examine whether MST produces outcomes that are superior to the comprehensive treatment already available (“treatment as usual”). The primary outcomes are defined as a decrease in the rate and seriousness of antisocial behaviour and prevention of recidivism. The second aim is to understand the processes through which MST works by testing the hypothesized mediators of beneficial intervention: increase in the competence of both the adolescent and his/her parents, improvements in the parent–adolescent relationship, decrease in the adolescent’s involvement with deviant peers and increased involvement with conventional, pro-social peers). The third aim is to determine the conditions that may affect the effectiveness of MST: characteristics of clients (ethnicity, psychopathology, marital dysfunctioning) or characteristics of the intervention (adherence, length, therapist experience, etc.)

The study involves random assignment into experimental (MST) condition and control (“treatment as usual”) condition, a pre-test, a post-test immediately after intervention (about 5 months later) and a follow-up a year after intervention. The reasons for the choice of randomized experiments were the following. In order to carry out a systematic evaluation of any intervention, an experimental design is the condition sine qua non. It is the gold standard or highest level in the classification of

the Scientific Methods Scale (Farrington 2003) or the Maryland Scientific Methods Scale (Sherman, in: Lum & Yang 2005). Randomized experiments reduce a number of threats to validity (Clingempeel and Henggeler 2002; Farrington 2003). The problem with quasi-experiments is that one can never be sure that the comparison groups are alike for every relevant factor, or that all possible confounders are identified (Weisburd 2003). In contrast, the experiments can rule out (or, at least, make highly implausible) alternative explanations such as passage of time, the effects of assessment, or the presence of different types of clients in the experimental and control conditions (Cook 2003).

In the present study the families in the control condition received treatment as usual (existing services). This enabled us to determine whether the programme works better than treatment as usual. Therefore, this study provides explicit information regarding the relative benefits of a competing intervention. An additional advantage of the 'treatment as usual' control condition is that this design is more likely to be tolerated in clinical and legal settings and can overcome clinical staff resistance to effectiveness studies.

Although the choice in favour of experimental design was not difficult, there were many issues to consider when implementing this design, as the reality was far more complicated than initially anticipated. In the following sections we elaborate on the choices we made and discuss several problems experienced when implementing this design.

Challenges with implementing an experiment

Referral paths for MST

The first issue is the high complexity of the Dutch juvenile justice system, which makes it extremely challenging to identify at what stage and by whom young people are referred for MST, and consequently to determine an optimal moment for randomization. The justice system through which juveniles can be referred for MST is not unambiguous. Broadly speaking, there are two trajectories through which young people can be referred: either following criminal trials, or through civil protection procedures (van der Laan 2007).

In criminal cases, when a young person is arrested and taken into custody, a counsellor from the Youth Protection Council examines what would be best for this individual, writes an advisory report for the prosecutor and contacts the MST-providing agency. If there is agreement that MST is the best option, then a request for suspension will be made at the arraignment. The examining magistrate, suspect (defending council), and youth prosecutor are heard, and suspension of custody with MST and a probation measure as a special condition are requested by De Waag and by the Bureau of Youth Care, respectively. Then, a hearing in chambers will take place, and, if three judges in the juvenile court agree, the verdict will be suspension of custody on condition of MST and juvenile probation. This procedure is summarized in Fig. 1.

In civil protection orders, on the other hand, the procedure is as follows. When the young person shows risk behaviour, the family guardian notifies the child

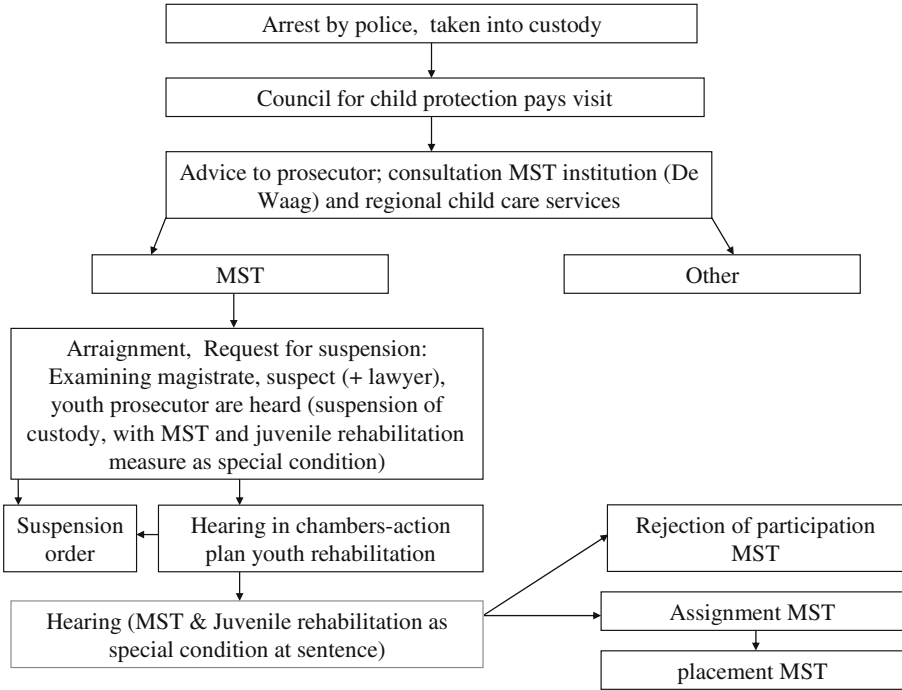


Fig. 1 Criminal law procedure

protection council, which examines the child and may advise a supervision order. The case is taken to court, and the judge will pronounce a supervision order. Following the order, a family guardian is assigned who designs an intervention plan. In the meantime the parents and the young person are supported by the child protection worker. This plan might be MST or placement in a secure or open youth institution. The civil protection procedure is shown schematically in Fig. 2. As shown in Fig. 2, the child protection worker can also refer the individual directly to MST.

Thus, numerous people and institutions are involved in the decision whether or not a young person is to be enrolled in MST: the Child Protection Council, Bureau of Youth Care, the prosecutors, the judges in the juvenile court, and the child and parents themselves. A complicating factor is also that the procedures differ quite a bit in the different regions of the Netherlands. In addition, other groups can be involved in the decisions, such as the formal cooperation between the Child Protection Council, the prosecutor, the police, and Bureau of Youth Care in Amsterdam (Ketenunits), although the final referral is always checked and confirmed by the Bureau of Youth Care, broadly speaking.

When to randomize

A consequence of this complex Dutch system is the difficulty in deciding at what moment the randomization should be carried out. There are several possibilities, each having its own advantages and disadvantages.

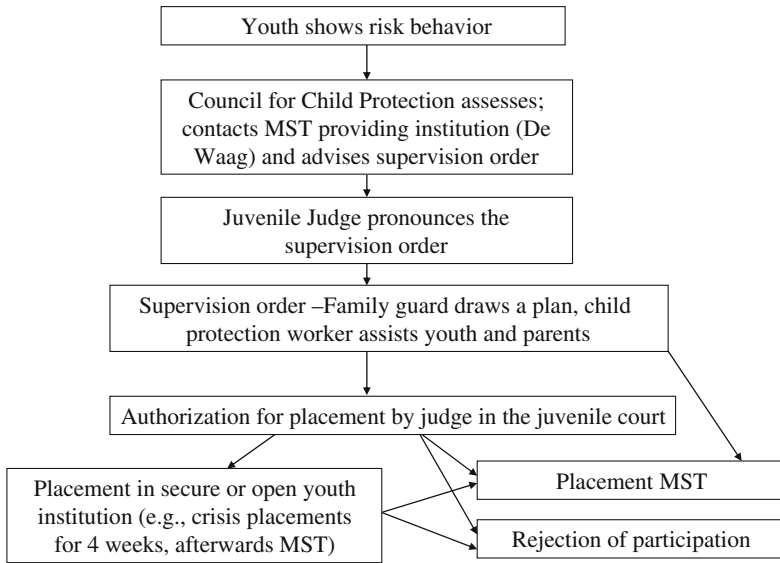


Fig. 2 Civil law procedure

The first possibility, and the initial preference in the present study, would be to randomize “as late as possible”, at the end of the referral process (see Figs. 1 and 2), when it is absolutely certain that the child will enter the MST programme. The young person would undergo random selection at De Waag itself. Consequently, the risk of drop-out during the phases of the referral process is minimal. Also, the randomization takes place at one central point, which makes it easier to carry out.

However, there are several problems with randomization at this stage. When all the parties concerned have come to a careful judgement and have decided that the individual must undergo MST, combined with a verdict of the judge, it is very hard to exclude the control group from the hoped-for intervention. Second, even though the randomization takes place centrally, all referrers involved in the earlier stages of the referral process still need to be informed and must be willing to collaborate, because they are involved in determining the treatment as usual.

An alternative possibility is to randomize “as early as possible”, whenever referrers consider MST as an option for the first time (see Figs. 1 and 2). When the randomization is carried out this early, the practitioners who first advise MST, i.e. the counsellor from the Child Protection Council or the family guardian, are the ones who write their advice according to the outcome of the randomization. This works as follows. Whenever a child protection worker or family guardian thinks that an individual is suitable for MST, they contact the MST-providing agency (De Waag). De Waag checks whether the family is indeed suitable for MST (i.e. the young person shows antisocial behaviour, and at least one parent/ caregiver at home will participate in MST). If the family meets both criteria, randomization takes place. This moment of randomization has the advantage that clinicians are actively involved in the randomization process. Further procedures remain as usual, and the advice can be confirmed by the judge in the juvenile court. Another advantage of

this moment of randomization is that the judge is able to state explicitly under what condition—i.e. MST—the case is to be suspended or a conditional sentence is to be imposed, which is legally required in certain regions.

However, there are disadvantages to this moment of randomization. There is a high risk of drop-out, because there are quite a few steps still to be taken after randomization, in which the families might turn out not to qualify for MST and drop out of the experimental condition. Moreover, it is quite an investment of time to inform all referrers and convince them of the use and importance of randomization.

In the present study the decisions about when randomization should take place were mainly driven practically. Even though both options (“as early as possible” and “as late as possible”) are acceptable, in the present study we decided to randomize as early as possible. The most important reason was that the judges in the juvenile court wanted to know what therapy a specific individual would eventually receive. To summarize, although we are aware of the advantages and disadvantages of the different randomization moments, we had to adjust our procedures to the wishes and demands of the referrers. Both in Utrecht and Amsterdam, the randomization at the Child Protection Council and at the branches of the Bureau of Youth Care takes place early in the referral process.

Legal issues

During the implementation of our study various legal dilemmas arose. A quite general problem with MST, related to the principle of legality, is the question of whether judges can sentence ‘whole’ families to MST when it is the child who commits the crime. Strictly speaking, in the Dutch justice system, one cannot sentence parents for the actions of their children. Therefore, MST can be imposed only on young people whose parents are sufficiently motivated to help their children. Actually, this first legal issue is not related to research specifically, but it might cause different approaches to MST in different regions. For example, the judges in the juvenile court in Amsterdam use this line of reasoning as an argument for why they never refer young people to MST in their verdicts. In Utrecht, however, judges include MST in their verdict whenever they think it is necessary. Another legal issue encountered is the question of whether, in some cases, MST might be considered a condition too severe for a relatively minor offence, because MST is a very intensive intervention (principle of proportionality). A third issue encountered is related to the principle of equal treatment (Graebisch 2000). Criminal sanctions usually need to be provided in a way such that people in similar circumstances must be treated equally. A randomization procedure tends to do the opposite. If, however, the ‘treatment as usual’ is comparable to the treatment, as in our study, randomization can be accepted as an option. In addition, as one of the referrers told us during the process: referral processes in juvenile justice in the Netherlands are often quite random already, with decisions based on the personal preference of the referrer concerned, waiting lists, and gut feelings. This means that, in fact, we often do experiments on people already. However, as the circumstances are not controlled for, we are unable to learn from them.

Resistance to random allocation

It goes without saying that the cooperation of agencies during the implementation and running of the experiment is of pivotal importance. However, despite the political push to evaluate the effectiveness of sentences imposed by the criminal justice system and the awareness of researchers that experimental designs are the optimal way to conduct the evaluation (Feder and Boruch 2000; Feder et al. 2000; Weisburd 2000), such designs are not popular, to put it mildly, among professionals in criminal justice agencies. It should be noted that the resistance of clinical staff is a problem encountered in any field evaluation study, regardless of its design, as agency personnel may be suspicious of the researchers' intent. Thus, resistance is also likely to be encountered in non-experimental studies. One of the reasons for this resistance is the perceived threat of negative research findings, i.e. the fear of negative consequences for the practitioners involved if the results show that interventions do not work (Feder et al. 2000). In our case, however, we met no resistance at all from the agency whose work we intended to evaluate (De Waag). On the contrary, the chief administrator actually initiated the study by approaching the research staff. He also helped to organize the meeting to inform the MST therapists, signifying the agency's general endorsement of the experiment and legitimizing interactions with agency personnel. Those directly affected by experiments, the MST therapists, whose work was to be evaluated and who needed to provide (part of) the data, were eager to find out whether their efforts do indeed work, and they indicated their full support.

The problems were encountered with referring agencies, those on which we counted in order to gain access to participants. These problems were almost exclusively focused on concerns regarding the random allocation of clients to two conditions. That is, these problems seem to be specific to experiments rather than to non-experimental studies. There are several ethical concerns. First, practitioners often believe that a particular intervention is better, despite the fact that the efficacy has not yet been proven. When asked to take part in randomized experiments, they feel that they are being asked to set aside their professional judgment and responsibility, and they consider it unethical to withhold the, in their eyes, superior, intervention from participants on the basis of chance alone. However, referrers tend to forget that we cannot continue to assume that our good intentions might not be harmful (Dishion et al. 1999). The argument that we take large risks by mandating an intervention that has not been rigorously tested (Boruch et al. 2000; Metcalf and Thornton 1992), is often convincing only after long and intense discussions. In the present study, likewise, such long discussions with all involved parties were necessary, as has been illustrated by our description of the decision about when randomization ought to take place.

Other issues included questions concerning the possibility of withholding experimental intervention from those who refuse to participate in the research. In the present study, we decided not to do so. So, when a family started MST, but did not want to participate in the study, they continued to receive treatment. However, we tried to keep the number of families refusing to participate as small as possible by asking MST therapists and the practitioners of the control group to explain to parents how important participation would be for them.

These ethical problems are serious points, and careless reactions of researchers to these points might violate ethical norms held by many. On the other hand, steps taken to protect conflicting ethical standards may severely weaken the methodological power or relevance of the effectiveness study (Metcalf and Thornton 1992). In the case of the present study, such an unlucky compromise is the agreement that referrers can assign families the exceptional status of receiving MST without randomization. In order to obtain cooperation with the institutions involved, we had to agree to their demand to continue to be able to make exceptions. Exceptions were allowed if MST was the *only* alternative to out-of-home placement. However, we took measures to ensure that not every young person would be an exception. Those measures were that a worker could never decide on his/her own that an exception had to be made. This always had to be confirmed by someone from the management. In practice it turned out that no exceptions were made at all, even though some individuals in the control condition were eventually placed out of home.

Apart from the ethical concerns, the staff feared the inconveniences that accompany a randomized controlled trial, the increase in workload and interferences with their everyday work. Since the research staff carries out the randomization at the university, and since the researchers are very proactive in keeping in contact about the families, this does not function as badly as was initially feared. However, of course, it still generates some extra work for workers at De Waag who need to communicate with the referrers about the research and for the therapists, who sometimes have to encourage families to participate.

Further threats to the progress of this study

In the further conduct of the study, too, there are several issues that might threaten the validity of the conclusions or the further continuation of the study, which we will briefly discuss below. Again, we would like to stress that many of these problems might also occur in other designs.

Maintaining the cooperation of institutions involved

The first issue that might threaten the further conduct of the study is that maintaining the cooperation of the agencies and institutions involved in the experiment can be difficult (Feder et al. 2000). Owing to misunderstandings, there is a risk that parties agreeing to participate withdraw later on. Even when willingness to participate is pronounced, there will always be people who disagree, and, sometimes, they have the power to influence those who have already agreed to participate. We also experienced this in the present study. Initially, for example, we offered to introduce the research to the judges in the juvenile court in Amsterdam, while, at the time, they did not consider the introduction necessary. After about 6 months, however, they were having doubts about the randomization process. We were then invited there and explained the procedures again.

Maintaining the cooperation of the individuals in the sample

The second group of threats concern the sample. It may turn out to be difficult to recruit sufficient numbers of participants to obtain an adequate statistical power. In the first place it might be difficult to recruit this group of participants and maintain their involvement in the research. A second problem is over-estimating the size of the eligible population (Feder et al. 2000). People tend to overestimate the group targeted for the service (Boruch 1997). A third concern in recruiting the sample is that it is common for the number of referrals to drop as soon as an evaluation study is carried out (Cunningham 2002). In the present study this did not seem to be the case. A reason for this might be that, from the beginning, there was close contact between the referrers, De Waag and the researchers. The referrers were guided towards an acceptable alternative, so it was not as if the referrer had problems if MST was not going to take place.

Once an appropriate sample for research is recruited, there are other threats. There is a risk of selective drop-out, both from treatment and from research. Despite a neat randomization procedure, there is a risk of selective drop-out of *treatment*. In the MST intervention, for example, therapists visit the families at home. Because of that, it probably requires less effort for families to have the therapist at their homes than to go to the clinic to receive an intervention, resulting in a higher drop-out rate within the control group. Another issue is the selective drop-out of the *research*. Families that participate in MST are possibly more motivated to participate also in the efficacy trial, whereas control group families might be less motivated. Alternatively, there might be a smaller drop-out in the MST condition because families are used to the procedures that accompany the study, for example, home visits and video-recordings. Comparison families are not used to these procedures and might experience them as intrusive. Apart from this, a threat concerning the sample is the fact that researchers are unable to control fully the conditions of the control group. Families may participate in different youth services, for example.

High staff turnover

A final threat to the continuation of experimental clinical field studies in criminology is the high staff turnover (Feder et al. 2000). This staff turnover is due to the often hostile attitude of the agencies and institutions who are asked to participate. Researchers frequently have to explain the basics of research and ethical issues over and over again. There also often exists an unfriendly attitude outside the agencies who have to participate, emanating, for example, from researchers of institutions who see the intervention to be evaluated or the researchers as a threat to their own number of referrals. As a consequence of the unfriendliness encountered outside or inside the agencies participating in the study, there is often high staff turnover in these studies (Cunningham 2002). In the present study there was also a high staff turnover, mainly among the research assistants. The assistants became unmotivated because they had to put a lot of effort into reaching the families, who were often not that interested in the research.

What can researchers do?

Overcoming resistance and maintaining the cooperation of the institutions involved

What can researchers do to motivate staff within the participating agencies to participate and to maintain their cooperation? Of foremost importance is investing time in communication, as we did in the present study, informing all involved parties, paying attention to their objections, and explaining why a randomized experiment is the only possible design to carry out a sound evaluation. It takes an initial investment to acquire support for the research from the management, but it also needs time during the actual carrying out of the study to acquire and maintain cooperation from the clinicians. It can be helpful to involve the clinicians in the study from the beginning and to involve them actively in the decisions in the research, such as the decision about what kind of ‘treatment as usual’ would be an acceptable and workable intervention for those families within the control condition. In our case this helped overcome ethical issues concerning unequal treatment of participants. However, this also gave us the problem that, during these meetings, they insisted on being able to make exceptions.

Time also had to be invested in communication, in the present study. As described above, it took some time to explain to the institutions concerned why this research was necessary and, mainly, why it had to be done with a randomized controlled trial. After we obtained formal cooperation, we held information sessions at each of the institutions concerned, depending on the time and needs of the institutions. In both Child Protection Councils, we presented the research and procedures twice: once as an introduction, and a second time just before the study actually started. A meeting was held with the judges in the juvenile courts in Utrecht and Amsterdam. At both branches of the Bureau of Youth Care all workers were informed by e-mail. Furthermore, all institutions and people concerned were updated regularly by newsletters, e-mails and visits of the researchers in order to make sure that, if there were problems, we would soon hear.

Secondly, it is important to keep the time the clinicians have to spend on the research at a minimum and to allow the experiment to disrupt normal procedures as little as possible (Metcalf and Thornton 1992; Borkovec and Echemendia 2001). We therefore try to incorporate the randomization procedure into their regular procedures. Referrers usually telephone De Waag to discuss whether or not a client might be a case for MST and to talk about the length of the waiting list, in order to decide whether or not MST is an option. In our case, whenever an individual was referred for MST and randomization took place, it was explained that the study was going on, which meant that whether or not the child would actually receive MST would depend on the randomization, and that, if MST would not be started, the person who answered the telephone would proactively be involved in finding the best alternative for this specific child. That way, all referrers were reminded of the procedure whenever they thought of MST. After this telephone conversation, the research was explained, and, immediately after that, the randomization took place. Usually, the referrer was called back within 5 minutes and would be told whether or not MST was going ahead. That way, the randomization procedure disrupted the

normal procedures as little as possible. The randomization was done by the principal researcher, who was always available to carry out randomization. Thus, De Waag had as little inconvenience as possible, as did the referrers. In order to facilitate matters, a written explanation concerning ethical aspects of randomization and describing what exactly the referrers need to be told was given to De Waag, and all referrers were provided with a written explanation of the research, ethical matters and the procedure as well. A simple research protocol was used.

Thirdly, in maintaining the cooperation of clinical staff, it might help if clinicians were to be reimbursed for the time devoted to the study (Borkovec and Echemendia 2001). If it is not possible to reward clinicians for their efforts with money, it is advisable to try to find other ways to make the research project of value to the clinicians. Research questions can be of relevance to clinicians, and so the research outcomes might help the participating referrers and agencies to improve their work (Feder et al. 2000). Unfortunately, in the present study, there was no budget to reimburse clinicians for their time devoted to the study. However, instead, the principal researcher came to the Waag each week to discuss possible problems and to check on the new families, so that the ones responsible for the research there did not have to spend time trying to reach the researcher. Finally, more time should be devoted to the education and training of clinicians from various agencies involved, in order to habituate and motivate them to evaluate their work.

Maintaining the cooperation of the individuals in the sample

In order to maintain cooperation of the participating families, there are several things that can be done. First, it is crucial that participants understand why participation is important and why certain questions are posed: the relevance of the study must be communicated clearly to the client, and procedures must be explained carefully. It takes investment to convince clients to cooperate. Second, attrition will be lower when participation in the research is not too much effort for the participants. It might help to use brief, clear and relevant questionnaires. It is also useful to send research assistants that speak the language of the families (in the Netherlands, for instance sometimes Arabic or Berber) to the homes of the participants to help them to read and to fill out the questionnaires (and to make sure they are returned). Finally, payment per completed assessment battery often motivates clients to participate (Borkovec and Echemendia 2001).

In the present study, it indeed turned out to be difficult to recruit individuals for the study. We tried to tackle this problem by asking the therapists to mention the research at the first intake and to ask them to stress the importance of this study for De Waag and for MST. However, they were not able to help to motivate the control group. In order to reach and motivate this group we needed contact with the those who were to provide the alternative treatments. Whenever they were carried out in the region where the research took place, it was viable; however, whenever a young person was taken into custody far out of the region, it was more difficult to reach them, because the institutions involved often did not know of the study. In most cases we were able to track the families via their referrers, who were very willing to help. However, it was not possible to reach all families.

High staff turnover

The issue of high staff turnover among our research assistants has been addressed by paying the assistant per attempt to make contact, rather than per successful result. Furthermore, time and attention is being devoted to preparing and training the assistants to make the telephone calls and visits to the families and by supporting them by telephone after each visit.

Discussion and conclusions

At the moment in The Netherlands, non-evaluated interventions continue to dominate the field of evaluations of interventions, especially juvenile justice interventions. In the absence of any form of accountability or quality control to ensure that evidence-based interventions are promoted, a diverse range of untested, perhaps even harmful, interventions are offered to the public. There is a large need for methodologically sound evaluation studies, but conducting a randomized experiment in a criminal justice setting is presently like skating on thin ice, under constant threat of Murphy's law. Given the large number of people who need to support the implementation of such a study, the chance that something (someone) will go "wrong" is indeed great.

The awareness that decisions in clinical practice should be evidence-based, rather than "good intention"- or intuition-based, is increasing, but the resistance to randomized experiments is still large, based mostly on misunderstandings of the principles of research. Although evaluation research in criminal justice settings will never be easy, due to a large number of factors that are beyond the control of the researchers, it could be made easier if professionals would understand and fully support such research. In order to get that far, much more attention should be given to the education and training of professionals working within criminal justice settings regarding the need for evidence-based work and understanding the logics and the basic principles of different types of research designs. This would make communication easier and, hopefully, lead to a realization that researchers and practitioners are on the same side and have the same aims: to improve interventions and help those who need it. Furthermore, it has been noticed that funding agencies in criminal justice play an important role in defining the methodology used (Weisburd 2003; Farrington and Welsh 2005). The clear encouragement of experimental studies by funding agencies is likely to lead to a more general use of experiments in criminal justice. Finally, it is important to stress that it is morally imperative (Weisburd 2003) for professional agencies to examine the efficacy of the treatments and interventions they offer. Since carrying out randomized experiments is the best method to examine efficacy, it is necessary to incorporate evaluation studies consisting of randomized experiments into agency policy.

However, implementing a randomized controlled trial in youth justice settings is complicated by several factors. First of all, there are many people and institutions involved in the decision of whether or not a young person will be enrolled for MST. Researchers, therefore, must be aware of the fact that it will cost time and energy to convince those involved and incorporate this time into research proposal planning. Second, there is still often a lot of resistance in the clinical field to this type of

experiment. The only way to overcome this resistance is to cooperate with the clinicians, listening carefully and investing time in explaining why things need to be done the way they are done. Generally, many more experiments should take place. Only when that happens will the resistance gradually decrease as clinicians become used to that way of working. And finally, there are the issues of keeping the cooperation of those involved and making sure that the groups are equal in the end, because everyone who is included once ought to stay within the study until the end. Keeping the cooperation of the institutions involved can be achieved by being in close contact, being available and by being proactive whenever problems occur. Making sure all families participate in the research and stay in the research requires close contact with the treating therapists and with the families themselves.

The focus of this paper is MST and juvenile justice in the Netherlands; however, many of the issues are encountered by anyone who plans to carry out a randomized controlled trial to examine the effectiveness of any treatment. We hope that this discussion about the issues encountered and the solutions we report will encourage and motivate others planning to examine the effectiveness of any treatment also to carry out a randomized controlled trial (RCT) as well. Even though we emphasized in the present paper the problems and obstacles that we (and many others) encounter when conducting randomized experiments, it is not our intention to discourage researchers to carry out such studies. On the contrary, the increased use of this design would probably lead to more acceptance by the agencies involved. As Feder et al. (2000) aptly put it, "...there is one thing that will help get experiments more widely accepted and used: doing more experiments." (p. 398). Such experimental studies provide not only valuable information for clinical practice but are also ways of testing models devised from fundamental research and can be used to increase our understanding of the mechanisms involved in the development of antisocial behaviour and delinquency.

References

- Andrews, D. (1996). The psychology of criminal conduct and effective treatment. In J. McGuire (Ed.), *What works; reducing reoffending* (pp. 3–34). Chichester: Wiley.
- Bijleveld, C. C. J. H. (2003). *Mens duf te meten! Over methoden en technieken van criminologisch onderzoek. [Human, dare to measure! On methods in criminology research]*. Den Haag: Boom Juridische Uitgevers.
- Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., et al. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, *63*, 569–578.
- Borkovec, T. D., & Echemendia, R. J. (2001). The Pennsylvania practice research network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice*, *8*, 155–167.
- Boruch, R. F., Victor, T., & Cecil, J. S. (2000). Resolving ethical and legal problems in randomized experiments. *Crime & Delinquency*, *46*, 330–354.
- Boruch, R. F. (1997). *Randomized experiments for planning and evaluation: A practical guide*. Thousand Oaks, CA: Sage.
- Clingempeel, W. G., & Henggeler, S. W. (2002). Randomized clinical trials, developmental theory, and antisocial youth: Guidelines for research. *Development and Psychopathology*, *14*, 659–711.
- Cook, T. D. (2003). Why have educational evaluators chosen not to do randomized experiments? *Annals of the American Academy of Political and Social Sciences*, *589*, 114–149.
- 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.
- Đeković, M., Janssens, J. M. A. M., & As, N. M. C. van (2007). Family factors and adolescent violent antisocial behavior. In R. Loeber, & M. Tonry (Eds). *Crime and Justice Series* (in press).
- Đeković, M., Wissink, I. B., & Meijer, A. M. (2004). The role of family and peer relations in adolescent antisocial behavior: Comparison of four ethnic groups. *Journal of Adolescence, 27*, 497–514.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm. Peer group and problem behavior. *American Psychologist, 9*, 755–764.
- Farrington, D. P. (2003). Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Sciences, 587*, 49–69.
- Farrington, D. P., & Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology, 1*, 9–38.
- Feder, L., & Boruch, R. F. (2000). The need for experiments in criminal justice settings. *Crime & Delinquency, 46*, 291–295.
- Feder, L., Jolin, A., & Feyerherm, W. (2000). Lessons from two randomized experiments in criminal justice settings. *Crime & Delinquency, 46*, 380–400.
- Graebisch, C. (2000). Legal issues of randomized experiments on sanctioning. *Crime & Delinquency, 46*, 271–282.
- Henggeler, S. W., Cunningham, P. B., Pickrel, S. G., Schoenwald, S. K., & Brondino, M. J. (1996). Multisystemic therapy: An effective violence prevention approach for serious juvenile offenders. *Journal of Adolescence, 19*, 47–61.
- 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., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance-abusing and -dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research, 1*, 171–184.
- Huey, S. J., Henggeler, S. W., Brondino, M. J., Pickrel, S. G. (2000). Mechanisms of change in multisystemic therapy: Reducing delinquent behavior through therapist adherence and improved family and peer functioning. *Journal of Consulting and Clinical Psychology, 68*, 451–467.
- Kazdin, A. E. (1997). A model for developing effective treatments: Progression and interplay of theory, research and practice. *Journal of Clinical Child Psychology, 26*, 114–129.
- Leschied, A., & Cunningham, A. J. (2002). *One step forward; Lessons learned from a randomized study of multiSystemic therapy in Canada*. London: ON.
- Littell, J. H. (2005). Lessons from a systematic review of effects of multisystemic therapy: *Children and Youth Services Review, 27*, 445–463.
- Littell, J. H., Popa, M., & Forsythe, B. (2005). Multisystemic therapy for social, emotional and behavioural problems in youth aged 10–17. (Cochrane Review).
- Lum, C., & Yang, S. M. (2005). Why do evaluation researchers in crime and justice choose non-experimental methods? *Journal of Experimental Criminology, 1*, 191–213.
- McGuire, J., & Priestley, P. (1996). Reviewing ‘What Works’: Past, present and future. In J. McGuire (Ed.), *What works; reducing reoffending* (pp. 3–34). Chichester: Wiley.
- Metcalf, C. E., & Thornton, C. (1992). Random assignment. *Children and Youth Services Review, 14*, 145–156.
- Ministry of Justice (2003). *Jeugd terecht. Actieprogramma aanpak jeugdcriminaliteit [Action program for prevention of youth criminality]*. Den Haag: Ministerie van Justitie.
- Ministry of Justice (2005). *Installatie erkenningscommissie gedragsinterventies Justitie. [Best practices initiative behavioural interventions in justice settings]*. Den Haag: Ministerie van Justitie. Web address: <http://www.justitie.nl/actueel/persberichten/archief2005/installatie-erkenningscommissie-Gedragsinterventies-Justitie.aspx>.
- 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* (2), 77–83.
- Schoenwald, S. K., Halliday-Boykins, C. A., & Henggeler, S. W. (2003). Client-level predictors of adherence to MST in community service settings. *Family Process, 42*, 345–359.

- Schoenwald, S. K., Henggeler, S. W., Brodino, M. J., & Rowland, M. D. (2000). Multisystemic therapy: Monitoring treatment fidelity. *Family Process, 39*, 83–103.
- 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 Laan, P. H. (2007). Police and justice interventions. In R. Loeber, & M. Tonry (Eds). *Crime and Justice Series* (in press).
- Weisburd, D. (2000). Randomized experiments in criminal justice policy: Prospects and problems. *Crime & Delinquency, 46*, 181–194.
- Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice. The moral imperative for randomized trials. *Evaluation Review, 27*, 336–354.

Jessica Asscher developmental psychologist, is a postdoctoral fellow at Utrecht University, currently carrying out a Dutch randomized clinical trial in order to evaluate the effectiveness of multi-systemic therapy in The Netherlands. In a previous project she evaluated the effectiveness of the Home-start programme. Her research interests include the development of behaviour problems in children and the evaluation of interventions.

Maja Deković developmental psychologist, is a full professor at Utrecht University and leader of the research programme “Development and treatment of psychosocial problems in context”. Her research interests include the development of problem behaviour, parent–adolescent relationships, family interaction and effects of family-based interventions. She is the co-author of several books and more than 130 articles in these areas.

Peter van der Laan is professor of Social and Educational Care at the University of Amsterdam, senior researcher at the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) in Leiden and a member of the steering committee of the Campbell Crime and Justice Group. He has a long-lasting interest and expertise in the field of juvenile delinquency and juvenile justice research. His publications concentrate increasingly on the effectiveness of judicial and semi-judicial interventions and policies.

Pier Prins is director of the research programme “Mechanisms of Change and Psychosocial Treatments for Behaviourally and Emotionally Disordered Youth” at the University of Amsterdam. He has a long record of international peer-reviewed publications in the area of evaluating childhood behavioural interventions and over the past decade has been project leader in various large-scale, multi-site treatment–outcome projects dealing with serious social problems, attention deficit hyperactivity disorder (ADHD) and conduct disorder in youth.

Sander van Arum MSc. is chief clinical director of youth and heads the multi-systemic therapy (MST) programme at De Waag—a centre for forensic ambulatory psychiatry with seven locations in the Netherlands. He is a registered psychotherapist, an approved family therapy supervisor, and he chairs the forensic systems therapy section of the Dutch Association for Relational, and Family, Therapy. In his clinical work in child and adolescent psychiatry he focuses on disruptive behaviour disorders, post-traumatic stress disorder, relational violence and family resilience.