Journal of Consulting and Clinical Psychology
2007, Vol. 75, No. 1, 187–193


0022-006X/07/$12.00 DOI: 10.1037/0022-006X.75.1.187

# Multidimensional Treatment Foster Care for Girls in the Juvenile Justice System: 2-Year Follow-Up of a Randomized Clinical Trial

Patricia Chamberlain, Leslie D. Leve, and David S. DeGarmo
Oregon Social Learning Center

This study is a 2-year follow-up of girls with serious and chronic delinquency who were enrolled in a randomized clinical trial conducted from 1997 to 2002 comparing multidimensional treatment foster care (MTFC) and group care ($N = 81$). Girls were referred by juvenile court judges and had an average of over 11 criminal referrals when they entered the study. A latent variable analysis of covariance model controlling for initial status demonstrated maintenance of effects for MTFC in preventing delinquency at the 2-year assessment, as measured by days in locked settings, number of criminal referrals, and self-reported delinquency. A latent variable growth model focusing on variance in individual trajectories across the course of the study also demonstrated the efficacy of MTFC. Older girls exhibited less delinquency over time relative to younger girls in both conditions. Implications for gender-sensitive programming for youths referred from juvenile justice are discussed.

*Keywords:* gender, delinquency, girl, intervention, foster care

Although girls have multiple factors in common with their male counterparts that put them at risk for delinquency (e.g., low parental monitoring, association with delinquent peers, substance use, family criminality and instability), other factors identified from developmental psychopathology appear to play a role in how delinquency unfolds for girls. In particular, longitudinal and developmental research suggests that girls' involvement in juvenile justice often follows from exposure to trauma and abuse and often co-occurs with anxiety and mood problems (Teplin, Abram, McLelland, Dulcan, & Mericle, 2002), negative interpersonal relationships (Ehrensaft, 2005), and social aggression (Underwood, 2003). The increasing awareness of the developmental pathways to delinquency for girls, plus concerns about serving girls in existing male-oriented programs and institutions, speaks to the need for gender-sensitive services. This is underscored by data showing that female delinquency is increasing relative to male delinquency, resulting in problems for existing public child-service systems that have few community-based treatment alternatives for girls (Siegel & Senna, 2000).

Patricia Chamberlain, Leslie D. Leve, and David S. DeGarmo, Oregon Social Learning Center, Eugene, Oregon.

Patricia Chamberlain is one of the owners of TFCC, Inc., which disseminates MTFC to community agencies in the United States and Europe.

Support for this research was provided by the Oregon Youth Authority and by the following grants: National Institute on Drug and Alcohol Abuse Grants DA15208 and DA17592 and National Institute of Mental Health Grant MH54257.

We thank J. P. Davis, Peter Sprengelmeyer, and Dana Smith for implementing the intervention; Courtenay Padgett and Sara Webb for data collection and management; the Oregon Youth Authority directors (Rick Hill, Karen Brazeau, and Robert Jester) and the Lane County Department of Youth Services for their assistance and support; Judy Boler for editorial assistance; and all the youths, parents, and foster parents who volunteered to participate in this research study.

Correspondence concerning this article should be addressed to Patricia Chamberlain, Oregon Social Learning Center, 10 Shelton McMurphey Boulevard, Eugene, OR 97401-4928. E-mail: pattic@oslc.org

In this report, we focus on results of a randomized clinical trial of girls with chronic delinquency who participated in a study comparing multidimensional treatment foster care (MTFC; Chamberlain, 2003) with group care (GC). Past studies have indicated that MTFC is an effective treatment approach for delinquent boys, producing outcomes superior to GC in terms of arrest and incarceration rates (Chamberlain & Reid, 1998; Eddy, Whaley, & Chamberlain, 2004). In addition, the cost effectiveness of MTFC relative to GC for boys was evaluated in an independent economic evaluation in which long-term cost savings to taxpayers were estimated to range from $21,836 to $87,622 per youth (Aos, Phipps, Barnoski, & Leib, 2001). Four variables have accounted for significant variance in delinquency outcomes: relationship with a mentoring adult, close supervision, clear limit setting, and low association with delinquent peers (Eddy & Chamberlain, 2000). The current application of MTFC retained these factors and was modified to map onto recent studies of girls (e.g., Putallaz & Bierman, 2004). Enhanced components include strategies for increasing emotional regulation and coping skills, recognizing anxiety and symptoms related to abuse and trauma, and decreasing social aggression, areas identified as potentially malleable antecedents of girls' conduct problems.

At 1-year assessments, girls in MTFC had a significantly greater reduction in days in locked settings and caregiver-reported delinquency than girls in GC (Leve, Chamberlain, & Reid, 2005). Therefore, follow-up efficacy models are evaluated in this article with directional hypotheses (Aron & Aron, 1994). We focus on (a) maintenance of intervention effects on delinquency at 2 years and (b) intervention effects on growth trajectories using repeated assessment data.

## Method

### *Participants*

Juvenile court judges in Oregon referred 103 girls between 1997 and 2002; girls had been mandated to out-of-home care because of

problems with chronic delinquency. The project manager enrolled all referred girls who were 13–17 years old and not currently pregnant. Girls were randomly assigned to the experimental condition (MTFC; $n = 37$) or the control condition (GC; $n = 44$). The flow of participants through the study is presented in Figure 1. Caseworker and youth written consent to participate was attained prior to participation and was accompanied by project staff's verbal explanations of the study and its risks and benefits. All study youths and caregivers knew that they were participating in a research study and receiving treatment services. Analyses included the full intent-to-treat randomized sample, although treatment length varied. The mean length of stay in the randomized placement was 174 days ($SD = 144$ days), and the average time between baseline and intervention entry was 47 days (neither of which differed significantly by group). The study was conducted in compliance with our institution's internal review board.

Girls were 15–19 years old at the 24-month assessment ($M = 17.3$, $SD = 1.0$); the sample was 74% Caucasian, 2% African American, 9% Hispanic, 12% Native American, 1% Asian, and 2% other or of mixed ethnic heritage. In comparison, 93% of girls ages 13–19 living in the region were Caucasian (U.S. Department of Commerce, 1992). Additional demographic information is given in Leve et al. (2005). There were no group differences on the rates or types of prebaseline offenses or on other demographic characteristics. No adverse events occurred during the study.

### *Procedure*

Prior to entering their out-of-home placements, each girl and her parent or other primary preplacement caregiver participated in a 2-hr baseline (BL) assessment at the research center. Staff members responsible for data collection and data entry were blind to

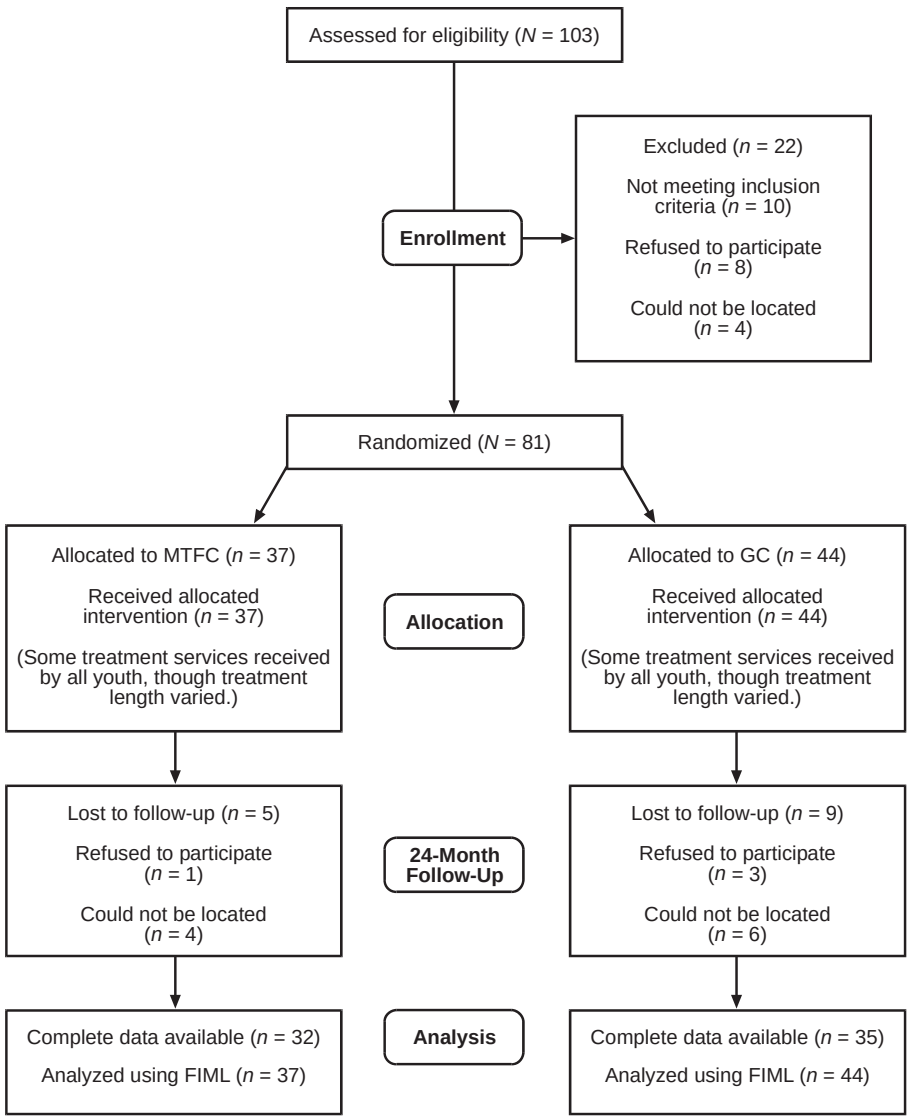


*Figure 1.* CONSORT statement flow diagram of participants through each stage of the randomized clinical trial. MTFC = multidimensional treatment foster care; GC = group care; FIML = full-information maximum likelihood.

participants' group assignment and were not involved in delivering the intervention. At 12 and 24 months post-BL, girls completed a follow-up assessment, and juvenile court records were collected.

*MTFC intervention.* Girls were placed in highly trained and supervised foster homes with state-certified parents. Although the intervention was individualized for each girl, it included standardized components: (a) Daily (Monday through Friday) telephone contact with the foster parents, via the Parent Daily Report Checklist (Chamberlain & Reid, 1987), monitored treatment fidelity and case progress. Weekday fidelity data were collected on parent implementation of an individualized, in-home, daily point-and-level program to reinforce girls' strengths and provide consequences for problem behaviors. Case progress data tracked girls' performance on the point-and-level system, amount of unsupervised time, school attendance and performance, and foster parent stress level. These data were used to adjust treatment in the foster home during weekly foster parent supervision and support sessions. (b) Weekly foster parent group training, supervision, and support meetings were led by an experienced program supervisor, who also coordinated all intervention components across settings and supervised individual and family therapy. (c) Individual therapy for each girl emphasized here-and-now problems. (d) Family therapy (for the family of origin) focused on improving parent management strategies. (e) School functioning was closely monitored with a daily school card signed by teachers. (f) Program staff were on call for the youths and foster and biological parents. (g) Psychiatric consultation was conducted.

For the implementation of gender-related components, foster parents and therapists were trained and supervised to teach and reinforce girls to avoid social–relational aggression and to develop alternative strategies for dealing with perceived rejection and other stress. Foster parents and girls were supervised to focus on learning strategies for emotional regulation (e.g., recognizing feelings of distress and coping strategies), generating options for dealing with problems, and planning for the future. Program supervisors viewed videotaped family and individual sessions and supervised therapists and foster parents to correct fidelity issues (see Chamberlain, 2003, for a program description).

*Control condition.* Girls went to 1 of 19 community-based GC care programs located throughout the state of Oregon that represented typical services for girls in out-of-home care. The programs had 2 to 51 youths in residence ($M = 21$), 1 to 50 staff members ($Mdn = 2$), and on-site schooling. The majority of programs (86%) reported endorsing a specific treatment model, with the primary philosophy being behavioral (70%), eclectic (26%), or family style (4%). Seventy percent of the programs reported delivering therapeutic services at least weekly.

## Measure

### *Delinquency Construct*

A multiple-method delinquency construct was computed from three indicators assessing behavior during the prior 12 months: number of criminal referrals, number of days in locked settings, and self-reported delinquency. Criminal referrals were collected from state police records and circuit court data, which have been found to be reliable indicators of externalizing behavior (Capaldi & Stoolmiller, 1999). The number of days spent in locked settings was measured by girls' report of total days spent in detention, correctional facilities, jail, or prison. Self-reported delinquency was measured with the Elliott General Delinquency Scale (Elliott, Huizinga, & Ageton, 1985). The 21-item subscale records the number of times girls report violating laws during the preceding 12 months.

### *Analytic Approach*

Two-year maintenance effects were tested with two methods. The first focused on the 2-year end point to examine whether differences between groups were still evident at 24 months. Therefore, we specified an autoregressive latent variable construct using structural equation modeling, specifying the factor measurement loadings as equal across time. This model provided the ability to partial measurement error, evaluate measurement invariance across time, and handle missing data, thus providing enhanced validity for the criterion outcome. Second, we examined effects across the course of the study, focusing on individual variation in delinquency trajectories using latent growth curve models for experimental designs (Curran & Muthén, 1999). Growth models take into account individual differences in change trajectories. To examine trajectories over time, we rescaled each continuous indicator 0 to 1 and then averaged them. For example, the minimum and maximum ranges of the Elliott General Delinquency Scale were rescaled from 0 to 1 at each wave. For count data, the data must be bounded by the same minimum and maximum value across time. Therefore, the count data were bounded by the maximum value across time before rescaling. Delinquency indicators were significantly correlated within each wave, and a principal-components factor analysis produced a single-factor solution at each wave, with eigenvalues ranging from 1.5 to 1.6 across time (loadings ranged from .44 to .82). Criminal referrals and days in locked settings were log transformed before rescaling to correct for distributional skewness and kurtosis.

A key advantage for both analytic approaches over repeated measures analysis of variance is the ability to control for initial status and the ability to model missing data using full-information maximum likelihood. Full-information maximum likelihood uses all available information from observed data. Compared with mean imputation, listwise, and pairwise models, the method provides more statistically reliable standard errors (Wothke, 2000). A missing-values analysis determined that all indicators and predictors were missing completely at random, Little's missing completely at random $\chi^2(44, N = 81) = 44.82$, $p = .43$. Similarly, there was no differential rate of attrition for those participants lost to follow up by group, as shown in Figure 1.

## Results

Means, standard deviations, and comparisons for the delinquency construct and indicators are provided in Table 1. There were no baseline differences for any of the variables. We also note that although individual indicators were differentially affected, the criterion outcome and focus was the delinquency construct. A latent variable composed of convergent communality across indicators provides a more continuous multisource measure and therefore a more reliable index than any one source alone (Patterson, 1996). Results of the structural equation model testing the main-

Table 1
*Means, Standard Deviations, and Comparisons for Rescaled Delinquency Indicators and Construct Score*

| Variable | GC | | MTFC | | | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *t* | *df* |
| Delinquency growth construct | | | | | | |
| Time 1: baseline | 0.48 | 0.18 | 0.47 | 0.18 | 0.46 | 79 |
| Time 2: 12 mo. | 0.30 | 0.20 | 0.22 | 0.17 | 1.98* | 74 |
| Time 3: 24 mo. | 0.25 | 0.21 | 0.12 | 0.16 | 2.66** | 65 |
| Elliott General Delinquency | | | | | | |
| Time 1: baseline | 0.36 | 0.26 | 0.33 | 0.21 | 0.51 | 79 |
| Time 2: 12 mo. | 0.15 | 0.19 | 0.18 | 0.19 | 0.82 | 70 |
| Time 3: 24 mo. | 0.12 | 0.16 | 0.11 | 0.18 | 0.15 | 60 |
| Log no. criminal referrals | | | | | | |
| Time 1: baseline | 0.53 | 0.24 | 0.53 | 0.27 | 0.06 | 79 |
| Time 2: 12 mo. | 0.25 | 0.24 | 0.15 | 0.21 | 1.93* | 74 |
| Time 3: 24 mo. | 0.22 | 0.26 | 0.13 | 0.18 | 1.63† | 61 |
| Log days in locked settings | | | | | | |
| Time 1: baseline | 0.60 | 0.27 | 0.56 | 0.33 | 0.59 | 74 |
| Time 2: 12 mo. | 0.51 | 0.38 | 0.31 | 0.34 | 2.28* | 74 |
| Time 3: 24 mo. | 0.42 | 0.40 | 0.14 | 0.28 | 3.10** | 61 |

*Note.* GC = group care; MTFC = multidimensional treatment foster care; mo. = months.
† $p < .06$. * $p < .05$. ** $p < .01$.

tenance effects are shown in Figure 2 by means of standardized path coefficients.

When age was controlled, MTFC was associated with greater reductions in delinquency compared with GC ($\beta = -.36$, $p < .01$). Older girls exhibited lower levels of 24-month delinquency ($\beta = -.36$, $p < .01$). Factor loadings were strong, and the model provided an excellent fit to the data, $\chi^2(23, N = 81) = 18.44$, $p = .73$ (comparative fit index = 1.00). In addition, tests of equally constrained factor loadings across time versus freely estimated loadings for each respective indicator demonstrated measurement invariance across time. The MTFC effect size for the 24-month follow-up was estimated as Cohen's *d*, a mean comparison of the delinquency construct shown in Table 1, and as eta-squared, with controls for baseline and age. Cohen's *d* was .65, and eta-squared

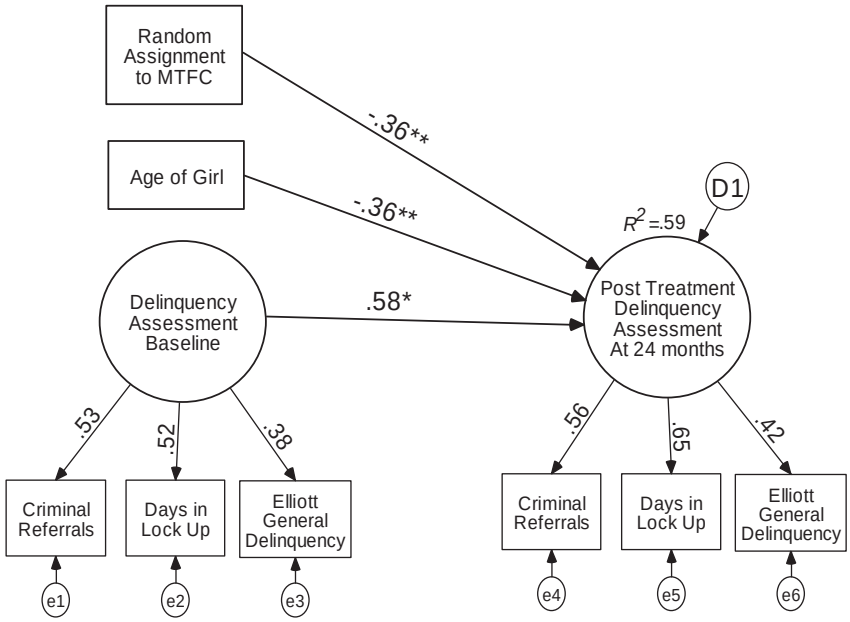


*Figure 2.* Latent variable structural equation path model testing 2-year maintenance effects of multidimensional treatment foster care (MTFC) on girls' delinquency compared with group care (GC). Paths are standardized coefficients, $\chi^2(23, N = 81) = 18.44$, $p = .73$ (comparative fit index = 1.00). D = latent variable error term; e = manifest variable error term. * $p < .05$. ** $p < .01$.

was .08; both of these effect size coefficients were considered a medium effect (Cohen, 1988).

We next specified a linear growth curve model with the initial status factor, or random intercept, fixed at 1.0 for each time point and the linear growth factor fixed at 0.0, 1.0, and 2.0 for each respective time period. Inspection of the means in Table 1 indicates that relatively more change occurred from BL to 12 months than occurred from 12 to 24 months. Indeed, a strictly linear model provided a marginally adequate fit to the data, $\chi^2(6, N = 81) = 14.77$, $p = .02$ (comparative fit index = .73). One approach to nonlinear data is to specify an additional quadratic factor. Because there were only three waves of data, however, this would require fixing the error terms at, for instance, 20% of the variance (e.g., 1 − reliability). A more appropriate alternative to specifying a quadratic is to specify a linear spline model (Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004; Stoolmiller, 1995). A spline factor combines the growth rate information and shape of any nonlinear growth by fixing the growth parameters at 0.0 and 1.0 for the first two waves and freeing the third time point. The spline model provided a significant improvement in fit for a nested comparison, $\Delta\chi^2(1, N = 81) = 12.82$, $p < .001$. Results of the spline model are provided in Figure 3, $\chi^2(5, N = 81) = 1.95$, $p = .86$ (comparative fit index = 1.00). When initial status and age were controlled, MTFC girls obtained a greater rate of decrease in delinquency over the course of the study relative to GC girls ($\beta = -.42$, $p < .01$).

Taken together, theses analyses provide unique information regarding evaluation of the 2-year follow-up assessment. First, the effects were maintained on delinquency at the end of the study, and, second, the variation in trajectories of repeated assessments showed a significantly greater rate in reductions of delinquency for MTFC. The group mean trajectories are displayed in Figure 4 for the observed means and their 95% confidence intervals. Any mean of one group outside the confidence intervals of another group represents a significant effect of the MTFC treatment.

## Discussion

Participation in MTFC resulted in better outcomes than placement in GC at 12- and 24-month follow-ups. Findings show that effects found at 1 year (Leve et al., 2005) were maintained at the 2-year assessment with a slightly larger effect size and that trajectories of reductions across the course of the study were significantly larger for MTFC. Age predicted delinquency outcomes as well, with younger girls showing more vulnerability than older girls. These results do not address potential mediators of intervention effectiveness; this work is planned for subsequent analyses. Because we consider it important to first address the intervention's impact on the primary problem for which girls were referred and removed from their family home by the juvenile court authorities (i.e., delinquent behavior), we do not report on a broad array of outcome domains in this article, as suggested by Hoagwood, Jensen, Petti, and Burns (1996). Some are reported elsewhere (see Leve & Chamberlain, in press). Even the somewhat narrow findings reported here have potential policy and cost implications. It appears that girls, like boys, can be more effectively treated for delinquency in well-trained and supervised community foster homes and that recruitment and retention of those foster homes is feasible, which speaks to the potential clinical utility of these findings.

In addition, the number of days spent in locked facilities displayed the largest effect size as an indicator and has cost implications. In raw scale number of days, MTFC girls spent over 100 fewer days in locked settings during the 2 years post-BL than GC girls (mean difference = 104.82 days). Previous analyses have found that placement in MTFC costs from one third to one half less on a daily basis than placement in GC (Aos et al., 2001). Participation in MTFC includes attending public school, whereas GC participants most often attend on-site schools, which contributes to additional costs not captured in program rates.

Limitations of this study include the small sample size, the fact that the majority of participating girls were Caucasian (represen-

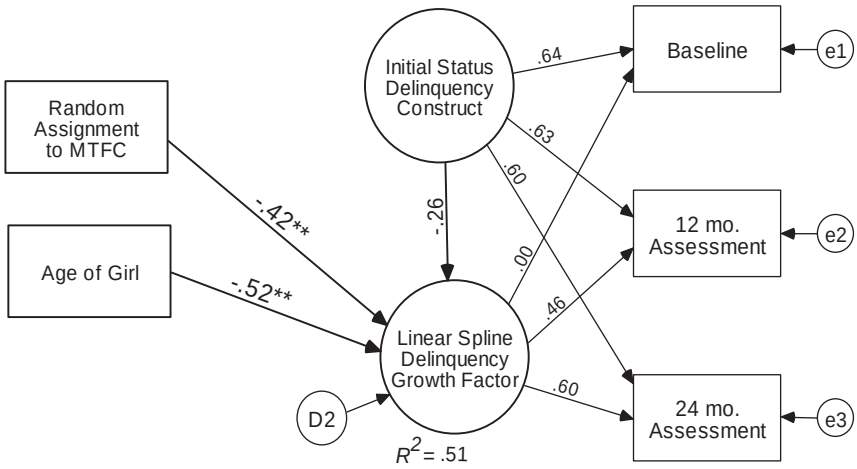


*Figure 3.* Latent variable growth model testing multidimensional treatment foster care (MTFC) 2-year effects on individual differences in trajectories of delinquency across the course of the study. Paths are standardized coefficients, $\chi^2(5, N = 81) = 1.95$, $p = .86$ (comparative fit index = 1.00). D = latent variable error term; e = manifest variable error term. $^{**}p < .01$.

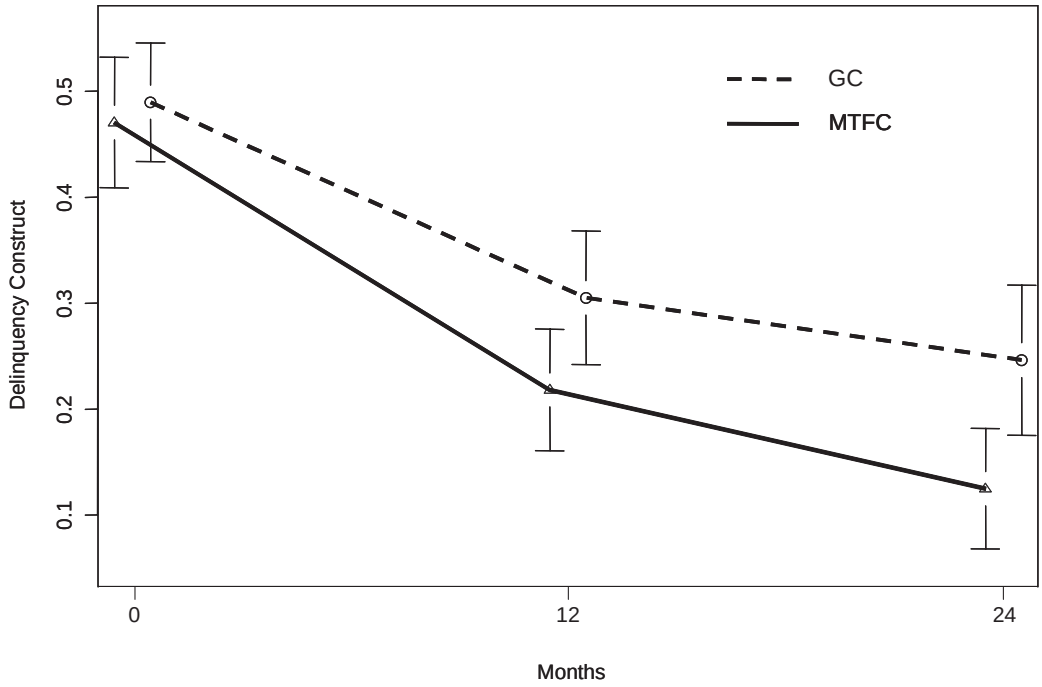


*Figure 4.* Group × Time observed mean trajectories of delinquency and their 95% confidence intervals. By default, the SPLUS graphing program "jitters," or slightly offsets, the confidence interval bands for ease of interpretation. MTFC = multidimensional treatment foster care; GC = group care.

tative of the region but not of the female juvenile justice population at large), and the fact that the findings are the first of their kind and need to be verified by replication. Although we used the literature on developmental psychopathology in girls to guide modifications to the MTFC model, the two-condition random assignment design did not allow us to disaggregate the intervention to test the efficacy of the added gender-specific components. Future studies will test whether gender-specific treatment components add to intervention effectiveness above factors relevant for both genders. The current study adds to the sparse empirical literature on intervention efficacy for girls and represents a modest step toward the development of research-based approaches for this understudied, vulnerable population.

## References

Aos, S., Phipps, P., Barnoski, R., & Leib, R. (2001). *The comparative costs and benefits of programs to reduce crime* (No. 01–05-1201). Olympia, WA: Washington State Institute for Public Policy.

Aron, A., & Aron, E. N. (1994). *Statistics for psychology.* Englewood Cliffs, NJ: Prentice Hall.

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9,* 30–52.

Capaldi, D. M., & Stoolmiller, M. (1999). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: III. Prediction to young-adult adjustment. *Development and Psychopathology, 11,* 59–84.

Chamberlain, P. (2003). *Treating chronic juvenile offenders: Advances made through the Oregon Multidimensional Treatment Foster Care model.* Washington, DC: American Psychological Association.

Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. *Behavioral Assessment, 9,* 97–109.

Chamberlain, P., & Reid, J. (1998). Comparison of two community alternatives to incarceration for chronic juvenile offenders. *Journal of Consulting and Clinical Psychology, 6,* 624–633.

Cohen, J. (1988). *Statistical power analysis for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Curran, P. J., & Muthén, B. O. (1999). The application of latent curve analysis to testing developmental theories in intervention research. *American Journal of Community Psychology, 27,* 567–595.

Eddy, J. M., & Chamberlain, P. (2000). Family management and deviant peer association as mediators of the impact of treatment condition on youth antisocial behavior. *Journal of Consulting and Clinical Psychology,* 68, 857–863.

Eddy, J. M., Whaley, R. B., & Chamberlain, P. (2004). The prevention of violent behavior by chronic and serious male juvenile offenders: A 2-year follow-up of a randomized clinical trial. *Journal of Emotional and Behavioral Disorders, 12,* 2–8.

Ehrensaft, M. K. (2005). Interpersonal relationships and sex differences in the development of conduct problems. *Clinical Child and Family Psychology Review, 8*(1), 39–63.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining delinquency and drug use.* Beverly Hills, CA: Sage.

Hoagwood, K., Jensen, P. S., Petti, T., & Burns, B. J. (1996). Outcome of mental health care for children and adolescents: I. A comprehensive conceptual model. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1055–1063.

Leve, L. D., & Chamberlain, P. C. (in press). A randomized evaluation of multidimensional treatment foster care: Effects on school attendance and homework completion in juvenile justice girls. *Research on Social Work Practice.*

Leve, L. D., Chamberlain, P., & Reid, J. B. (2005). Intervention outcomes

for girls referred from juvenile justice: Effects on delinquency. *Journal of Consulting and Clinical Psychology, 73,* 1181–1185.

Patterson, G. R. (1996). Some characteristics of a developmental theory for early onset delinquency. In M. F. Lenzenweger & J. J. Haugaard (Eds.), *Frontiers of developmental psychopathology* (pp. 81–124). New York: Oxford University Press.

Putallaz, M., & Bierman, K. L. (Eds.). (2004). *Aggression, antisocial behavior, and violence among girls: A developmental perspective.* New York: Guilford Press.

Siegel, L. J., & Senna, J. J. (2000). *Juvenile delinquency: Theory, practice, and law* (7th ed.). Belmont, CA: Wadsworth.

Stoolmiller, M. (1995). Using latent growth curve models to study developmental processes. In J. M. Gottman & G. Sackett (Eds.), *The analysis of change* (pp. 105–138). Hillsdale, NJ: Erlbaum.

Teplin, L. A., Abram, K. M., McClelland, G. M., Dulcan, M. K., & Mericle, A. A. (2002). Psychiatric disorders in youth in juvenile detention. *Archives of General Psychiatry, 59,* 1133–1143.

U.S. Department of Commerce. (1992). *1990 census of population: General population characteristics, Oregon* (USDC Publication No. 1990 CP-1–39). Washington, DC: U.S. Government Printing Office.

Underwood, M. K. (2003). *Social aggression among girls.* New York: Guilford Press.

Wothke, W. (2000). Longitudinal and multi-group modeling with missing data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multiple group data: Practical issues, applied approaches and specific examples* (pp. 269–281). Mahwah, NJ: Erlbaum.

Received January 25, 2006
Revision received October 25, 2006
Accepted October 31, 2006 ■