

Research on Social Work Practice

<http://rsw.sagepub.com>

Are Mainstream Programs for Juvenile Delinquency Less Effective With Minority Youth Than Majority Youth? A Meta-Analysis of Outcomes Research

Sandra Jo Wilson, Mark W. Lipsey and Haluk Soydan

Research on Social Work Practice 2003; 13; 3

DOI: 10.1177/1049731502238754

The online version of this article can be found at:
<http://rsw.sagepub.com/cgi/content/abstract/13/1/3>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Research on Social Work Practice* can be found at:

Email Alerts: <http://rsw.sagepub.com/cgi/alerts>

Subscriptions: <http://rsw.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations <http://rsw.sagepub.com/cgi/content/refs/13/1/3>

Are Mainstream Programs for Juvenile Delinquency Less Effective With Minority Youth Than Majority Youth? A Meta-Analysis of Outcomes Research

Sandra Jo Wilson

Mark W. Lipsey
Vanderbilt University

Haluk Soydan

Center for Evaluation of Social Services, Stockholm, Sweden

Objective: A meta-analysis was undertaken to synthesize research results about the effectiveness of mainstream service programs for minority juvenile delinquents relative to White delinquents. The analysis addresses the question of whether mainstream interventions that are not culturally tailored for minority youth have positive outcomes on their subsequent antisocial behavior, academic performance, peer relations, behavior problems, and other outcomes. In addition, outcomes were compared with those for White samples receiving the same interventions to identify any differences in the responsiveness of minority and majority youth. Method: 305 studies were selected from a large meta-analytic database in which the participant samples were either predominantly (60% or more) minority or White youth. Effect sizes and more than 150 study descriptors were coded from these studies and analyzed using standard meta-analytic techniques. Results: The results showed positive overall intervention effects with ethnic minority respondents on their delinquent behavior, school participation, peer relations, academic achievement, behavior problems, psychological adjustment, and attitudes. Overall, service programs were equally effective for minority and White delinquents. Although there were slight differences in effectiveness for different service types between minority and majority youth, none of these differences was statistically significant. Conclusions: The use of mainstream service programs for ethnic minority juvenile delinquents without cultural tailoring is supported by these findings.

The mission of the social work profession includes work with marginalized, needy, and economically disadvantaged groups and individuals in society. This often entails services for immigrant and ethnically diverse populations.

Authors' Note: The research reported in this article was supported in part by grants from the Centre for Evaluation of Social Services of the National Board of Health and Welfare, Stockholm, Sweden, the National Institute of Mental Health (MH42694 and MH39958) and the Russell

Research on Social Work Practice, Vol. 13 No. 1, January 2003 3-26
DOI: 10.1177/1049731502238754
© 2003 Sage Publications

In the early decades of its development in the United States, the social work profession was called upon to assist newly arrived immigrants, especially in urban and densely populated areas (Addams, 1895; Soydan, 1999). Generations later, the clientele for social work continues to be characterized by ethnic diversity, as well as a multiplicity of languages, religions, and value systems. In other parts of the world, particularly in Europe, social work faces similar circumstances as migration has made populations more diverse (Williams, Soydan, & Johnson, 1998).

Ethnic diversity, in particular, has presented major challenges to social work practice. When people of different ethnic groups are assumed to have different needs or to respond differently to services, the methods for serving those groups become a matter of debate. At issue is whether clients of ethnic minority groups should be treated with the same methods, interventions, and programs as the majority population of a particular country. Especially in the United States, the necessity of tailoring social work practice to meet the special needs of ethnically diverse populations has been much discussed (Soydan, Jergeby, Olsson, & Harms-Ringdal, 1999). Such concerns have resulted in various social work practice models for serving ethnic populations, for example, Cultural Awareness (Green, 1995), the Process Stage Approach in Minority Treatment (Lum, 1996), and Ethnic-Sensitive Social Work Practice (Devore & Schlesinger, 1996). These approaches are usually presented as generic models for ethnic minorities that are appropriate to different types of social work services such as family counseling, group counseling, behavior contracting, interpersonal skills training, drug treatment programs, and the like.

However, the need for, and effectiveness of, ethnically tailored social work approaches have been vigorously debated in social work research and practice (Soydan et al. 1999; de Anda, 1997). The nature of the controversy is illustrated by the papers in the volume, *Controversial Issues in Multiculturalism*, edited by de Anda (1997). The questions raised there include: whether the emphasis on multicultural practice has resulted in more effective and appropriate services for ethnic minority clients; whether programs and service delivery systems should be culture-specific in their design; whether the therapeutic process is more effective if the client and the helping professional are of the same ethnic/cultural group; and whether ethnic agencies can more effectively serve ethnic communities than mainstream agencies. Debating

Sage Foundation. A bibliography of studies used in the meta-analysis is available from the first author or the following Web site: <http://www.vanderbilt.edu/CERM/>. Correspondence concerning this article should be addressed to Sandra Jo Wilson, Center for Evaluation Research and Methodology, Vanderbilt Institute for Public Policy Studies, 1207 18th Ave. South, Nashville, TN 37212; telephone: (615) 343-7215; e-mail: sandra.j.wilson@vanderbilt.edu.

whether the emphasis on multicultural practice has resulted in more effective and appropriate services for ethnic minority clients, for instance, John Longres wrote,

Dr. Brown argues that multicultural practice has brought about more effective and appropriate services. I argue that it has not. . . . In the first place, Dr. Brown is largely talking about appropriateness, not effectiveness. She offers anecdotal evidence to demonstrate that her students, her colleagues, and apparently their clients seem to be satisfied with the counseling they are receiving. This anecdotal evidence hardly stands up to rigorous evaluation and so has to be taken for what it is, the opinion of an educator. Even if her evidence were more rigorously represented, Dr. Brown supplies no evidence of effectiveness: the clients and their helping professionals may feel good, but do the clients behave differently, and have their lives been changed for the better? The evidence suggests that as a collective, people of color are treading water; their lives have not been improved by the growth of a new multicultural sensitivity, however more appropriate it may appear to be. (de Anda, 1997, pp. 18-19)

As this comment indicates, arguments about this issue are seldom based on well-founded empirical knowledge. When relevant evidence is consulted, the results are often surprising. For example, empirical research related to one specific social work intervention, foster care for children who are maltreated or whose parents are unable to care for them, demonstrates that differences between ethnic groups in the outcome of mainstream programs, if they occur at all, do not necessarily favor majority groups. Using data on death rates, Barth and Blackwell (1998) showed that White and Hispanic children in foster care have higher rates of death than their counterparts in the general population, but the death rates for African American foster children are no worse than those for the African American children in the general population. On the other hand, Jonson-Reid and Barth (2000) found that the risk of incarceration following foster care was greater for African American youth than for Hispanic or White youth, even when gender, age at first placement, and characteristics of placement history were controlled.

Although the effectiveness of services in producing positive outcomes for the specific problems to which they are addressed is not the only important consideration in social work practice, it is certainly a critical one. Where such outcomes are concerned, the question of whether culturally tailored or mainstream social work practice is more effective for minority youth is an empirical one. Unfortunately, at present we lack a solid body of research on this issue and where we do have some empirical findings, as in the foster care example above, the results do not always point in the same direction.

The present study was undertaken to assemble otherwise scattered research results about the effectiveness of service programs for minority

juvenile delinquents relative to White majority delinquents. It uses the techniques of meta-analysis to address the question of whether mainstream interventions that are not culturally tailored for minority youth have positive outcomes on subsequent antisocial behavior. In addition, those outcomes are compared with the intervention effects for White majority samples to find out if there are any differences in the responsiveness of minority and majority youth to mainstream juvenile delinquency services.

META-ANALYSIS

Meta-analysis is a technique for recording and analyzing the statistical results of a collection of empirical research studies. Central to meta-analysis is the effect size statistic, which represents the quantitative findings of each study in a standardized way that permits comparison across studies. For the intervention studies of interest here, the effect size statistic (ES) is the standardized mean difference (Cohen, 1988; Lipsey & Wilson, 2001), defined as the difference between the treatment and control group means on an outcome variable divided by their pooled standard deviation. This effect size statistic indexes the outcomes for the treatment group relative to the control group in standard deviation units. Thus, $ES = .50$ indicates that the outcome for the treatment group was more favorable than that for the control group by an amount equivalent to half a standard deviation on the respective outcome measure. For binary outcomes (e.g., arrested/not arrested), the mean difference effect size was derived using Cohen's arcsine transformation for proportion values (Lipsey & Wilson, 2001). If a study assesses the effects of an intervention on more than one outcome, say reoffending and social adjustment, that study would generate two effect sizes, one for each outcome.

The effect size is not the only piece of information recorded about each study that contributes to a meta-analysis. Studies of juvenile delinquency programs involve youth from a variety of different age groups, ethnicities, risk levels, and so forth and use different procedures and methods to evaluate program effectiveness. Of course, the interventions themselves also differ across studies. Thus, in addition to the effect sizes, a meta-analytic database includes detailed information about each study's methods, participants, treatments, and other such descriptive characteristics that may be relevant for understanding the results. Analysis of the coded information from a collection of studies can then investigate relationships between various characteristics of the studies and the effect sizes those studies produce. In addition, statistical techniques can be used to control for many of the sources of error and natural differences between studies so that better estimates of the effects of intervention can be derived (Cooper & Hedges, 1994; Hedges & Olkin, 1985).

The studies included in the meta-analysis presented here are drawn from a database of nearly 500 studies on the effects of intervention with juvenile delinquents (Lipsey, 1992, 1995; Lipsey & Wilson, 1998). This larger synthesis of research results has shown that intervention for juvenile offenders is generally effective at reducing recidivism (Lipsey, 1995), even among serious, violent, or chronic juvenile offenders (Lipsey & Wilson, 1998). However, the results of these analyses have also shown that there is considerable variability in effectiveness across different research studies and that this variability is related to the research methods used, participant characteristics, and format and nature of the treatment. The analysis reported below uses this extensive database to address the relative effects of service programs for minority versus majority juvenile offenders.

METHOD

The database from which the studies for this meta-analysis are drawn includes empirical research on the effects of juvenile delinquency programs conducted between 1950 and 1996, both published and unpublished (Lipsey, 1992; Lipsey & Wilson, 1998). Trained personnel coded over 150 items for each study that describe the methods and procedures, participant samples, treatment and program characteristics, effect sizes, and other important information. Interrater reliability for the coding of effect sizes was .93 and interrater agreement was generally above 80% for items describing study characteristics. For the present analysis, the studies that allowed comparison of the effectiveness of delinquency programs for minority versus majority youth were selected from this database. Specifically, the 141 studies with participant samples comprised at least 60% minority youth, and the 164 studies with at least 60% White participants were selected. The mean percentage of minority youth in the studies selected to represent intervention with ethnic minorities was 82% and the mean percentage of White youth in the comparison studies was about 80%. Among the minority samples, most were predominately African American but Hispanic youth and, in much smaller numbers, other minority youth were also represented (see Table 1).

As with all studies in the full database, the selected studies met the following criteria:

- The youth in the study sample were identified as delinquent or displaying antisocial behavior; or, if the youth were not explicitly described as delinquent or antisocial, the study included delinquency or antisocial behavior among its primary outcome variables.

(Text continued on pg. 10)

TABLE 1: Characteristics of the 305 Studies Used in the Meta-Analysis

| <i>Variable</i> | <i>N</i> | <i>%^a</i> | <i>Variable</i> | <i>N</i> | <i>%</i> |
|---|----------|----------------------|--|----------|----------|
| General study characteristics | | | Institutionalized | 79 | 26 |
| Publication year | | | Percentage of sample with prior offenses | | |
| 1950-1969 | 43 | 14 | None | 7 | 2 |
| 1970-1979 | 119 | 39 | Some (< 50%) | 56 | 19 |
| 1980-1989 | 116 | 38 | Most (≥ 50%) | 43 | 14 |
| 1990-1996 | 21 | 7 | All | 143 | 47 |
| Missing | 6 | 2 | Some, cannot estimate % | 38 | 12 |
| Type of publication | | | Missing | 18 | 6 |
| Technical report | 156 | 51 | Heterogeneity of treatment sample | | |
| Journal article, chapter | 93 | 31 | Low | 107 | 35 |
| Dissertation | 30 | 10 | Moderate | 132 | 43 |
| Book | 24 | 8 | High | 66 | 22 |
| Conference paper | 2 | < 1 | Source of participants for treatment | | |
| Country in which study was conducted | | | Volunteers | 18 | 6 |
| United States | 280 | 92 | Referred by parents/friends | 1 | < 1 |
| United Kingdom | 14 | 5 | Referred by non–juvenile justice (JJ) agency | 19 | 6 |
| Canada | 9 | 3 | JJ referral (voluntary) | 91 | 30 |
| Other English-speaking country | 2 | < 1 | JJ referral (mandatory) | 142 | 47 |
| Characteristics of juveniles in treatment | | | Multiple sources | 12 | 4 |
| Gender mix | | | Solicited by researcher | 18 | 6 |
| No males | 9 | 3 | Missing | 4 | 1 |
| Some males (< 50%) | 16 | 5 | Treatment characteristics | | |
| Mostly males (≥ 50%) | 144 | 47 | Program age | | |
| All males | 136 | 45 | Relatively new | 187 | 61 |
| Mean age at time of treatment | | | Established (2+ years) | 113 | 37 |
| < 12.9 | 14 | 5 | Defunct | 1 | < 1 |
| 13.0-13.9 | 27 | 9 | Missing | 4 | 1 |
| 14.0-14.9 | 54 | 18 | Program sponsorship | | |
| 15.0-15.9 | 56 | 18 | Research project | 85 | 28 |
| 16.0-16.9 | 67 | 22 | Private agency | 29 | 10 |
| 17.0-17.9 | 38 | 12 | Public agency, non-JJ | 54 | 18 |
| 18.0-18.9 | 11 | 4 | Public agency, JJ | 132 | 43 |
| 19.0 | 27 | 9 | Missing | 5 | 2 |
| Missing | 11 | 4 | Duration of treatment (weeks) | | |
| Predominant ethnicity (> 60%) | | | 1-10 | 48 | 16 |
| African American | 80 | 26 | 11-20 | 81 | 26 |
| Hispanic | 19 | 6 | 21-30 | 47 | 15 |
| Other (Asian, American Indian) | 8 | 3 | 31-40 | 66 | 22 |
| Mixed (minority > 60%) | 34 | 11 | 41-50 | 13 | 4 |
| White | 164 | 54 | 51 and up | 50 | 17 |
| Delinquency level | | | | | |
| Predelinquent | 89 | 29 | | | |
| Delinquent | 137 | 45 | | | |

TABLE 1: (continued)

| Variable | N | % ^a | Variable | N | % |
|------------------------------------|-----|----------------|----------------------------------|-----|----|
| Frequency of treatment event | | | Public, non-JJ site | 64 | 21 |
| Continuous | 57 | 19 | Private | 80 | 26 |
| Daily | 55 | 18 | Mixed | 43 | 14 |
| 2-4 times/week | 31 | 10 | Other | 20 | 7 |
| 1-2 times/week | 90 | 30 | Treatment type | | |
| Less than weekly | 24 | 8 | Multimodal programs | | |
| Missing | 48 | 16 | Institutional | 16 | 5 |
| Implementation problems | | | Noninstitutional | 22 | 7 |
| Yes | 124 | 41 | Casework, service brokerage | 31 | 10 |
| Possible | 67 | 22 | Counseling, noninstitutional | 41 | 13 |
| No | 114 | 38 | Counseling, institutional | 42 | 14 |
| Rated amount of meaningful contact | | | Behavioral, cognitive-behavioral | 14 | 5 |
| 1 (Trivial) | 4 | 1 | Employment related | 26 | 9 |
| 2 | 41 | 13 | Wilderness/challenge | 9 | 3 |
| 3 | 40 | 13 | Academic services | 25 | 8 |
| 4 (Moderate) | 67 | 22 | Interpersonal skills | 6 | 2 |
| 5 | 69 | 23 | Vocational | 9 | 3 |
| 6 | 61 | 20 | Probation and variations | 21 | 7 |
| 7 (Substantial) | 23 | 8 | Other | 43 | 14 |
| Who delivers treatment? | | | Method characteristics | | |
| JJ personnel | 91 | 30 | What control group receives | | |
| School personnel | 14 | 5 | Receives nothing | 51 | 17 |
| Public mental health personnel | 20 | 7 | Minimal contact | 19 | 6 |
| Private mental health personnel | 44 | 14 | School, treatment as usual | 18 | 6 |
| Non-mental health counselors | 66 | 22 | Usual probation services | 72 | 24 |
| Laypersons | 47 | 15 | Usual institutional treatment | 84 | 28 |
| Researcher | 6 | 2 | Other treatment as usual | 49 | 16 |
| Other | 6 | 2 | Placebo | 12 | 4 |
| Missing | 11 | 4 | Role of evaluator | | |
| Primary treatment format | | | Delivered treatment | 13 | 4 |
| Juvenile alone | 8 | 3 | Planned or supervised treatment | 88 | 29 |
| Juvenile and provider | 81 | 27 | Influential, no direct role | 52 | 17 |
| Juvenile group | 147 | 48 | Independent of treatment | 137 | 45 |
| Juvenile and family | 19 | 6 | Missing | 15 | 5 |
| Mixed | 41 | 13 | Method of group assignment | | |
| Other | 9 | 3 | Random | 141 | 46 |
| Treatment site | | | | | |
| Public, JJ | 98 | 32 | | | |

(continued)

TABLE 1: (continued)

| <i>Variable</i> | N | % ^a | <i>Variable</i> | N | % |
|-----------------------------|-----|----------------|-----------------|----|----|
| Matching | 80 | 26 | Sample size | | |
| Other nonrandom | 84 | 28 | Up to 49 | 57 | 19 |
| Blinding in data collection | | | 50-99 | 65 | 21 |
| No | 233 | 76 | 100-199 | 84 | 28 |
| Yes | 72 | 24 | 200-299 | 34 | 11 |
| | | | 300-399 | 16 | 5 |
| | | | 400 and up | 49 | 16 |

a. Percentages may not add up to 100 because of rounding.

- The youth in treatment were between the ages of 12 and 21 and resided in the United States, Canada, Great Britain, New Zealand, or Australia.
- The research design was an experimental or quasi-experimental comparison of at least one treatment and one control/comparison group.

The selected studies reported intervention effects on a variety of outcome variables. The primary outcome of interest was subsequent delinquent behavior, most commonly measured by police contacts or arrests. Many studies also assessed other outcome variables that we have grouped into broad conceptual domains. Each of these domains covers a variety of specific individual variables, thus results presented for effects in these domains should be interpreted only as broad summaries. The domains we identified include the following outcome constructs. The number of studies contributing outcomes in each category is shown in parentheses.

- Academic achievement: grades and other measures of achievement in various academic subjects ($n = 34$)
- Attitude change: attitudes about delinquency or toward school, work, or community ($n = 57$)
- Behavior problems: nonaggressive behavior problems such as acting out, disruptiveness, and hyperactivity ($n = 23$)
- Employment status: getting or keeping a job, number of jobs held, and the like ($n = 28$)
- Family functioning: measures of family relations, parental discipline, family stress, and the like ($n = 15$)
- Internalizing problems: withdrawal, shyness, anxiety, and other similar problems ($n = 23$)
- Peer relations: relations with peers, interpersonal skills, and social adjustment ($n = 31$)
- Psychological adjustment: locus of control, personality adjustment, and the like ($n = 36$)

- School participation: tardiness, truancy, and dropping out ($n = 56$)
- Self-esteem: measures of self-esteem or self-concept ($n = 27$)

CHARACTERISTICS OF THE DELINQUENCY TREATMENT STUDIES

Table 1 presents a summary of the 305 studies that comprise the database for this meta-analysis. The characteristics of these studies were generally similar to those for the entire database from which they were drawn. In addition, the characteristics of the studies with minority samples were not appreciably different from those with predominantly White samples. Some of the general features of the studies are as follows:

- The majority of studies was conducted in the United States. Most studies were published subsequent to 1970 with the most common form of publication being technical reports.
- The juvenile samples were largely male with most of the youth age 15 or older. The predominant ethnic classification among minority samples was African American with smaller proportions of mixed and predominantly Hispanic samples.
- Most of the studies involved youth who were delinquent or institutionalized and, for most of the samples, all or the majority of the juveniles had prior offense histories.
- In nearly half of the studies, the treatment was delivered to a group of juveniles (rather than individually) and about one third of the programs were delivered by juvenile justice personnel.
- The most frequent intervention strategies were institutional and noninstitutional counseling and casework or service brokerage-type services, although many other types of service programs were represented.

Nearly half of the studies used random assignment to place youth in treatment or comparison groups. In most of the studies, the youth in the comparison groups received usual or customary services, typically probation or institutionalization.

In addition to the descriptive information already coded on the selected studies, the original reports for those studies using minority samples were examined for statements relating to cultural tailoring of the interventions. We found one report indicating such tailoring (Wooldredge, Hartman, Latessa, & Holmes, 1994) and excluded it from our analyses. Thirteen of the 141 studies with minority youth mentioned using minority service providers with minority delinquents, but the services provided in these cases were not specifically

culturally tailored. By all available indications, therefore, the interventions provided to the minority youth, like those provided to the White youth in the comparison studies, were predominantly mainstream services without any special tailoring to the ethnic or cultural characteristics of the juveniles receiving those services.

EFFECTIVENESS OF JUVENILE DELINQUENCY SERVICES

Figure 1 reports the mean effect size with its 95% confidence interval for each of the major intervention outcome constructs for minority and White samples. Hedges's (1981) small sample correction was applied to each effect size and all computations with effect sizes were weighted by the inverse sampling variance to reflect the greater stability of estimates based on larger samples (Hedges & Olkin, 1985; Lipsey & Wilson, 2001; Shadish & Haddock, 1994).

For minority youth, the weighted mean effect size for delinquency outcomes across all treatment modalities was .11; for majority youth, the corresponding effect size value was .17. Both these values were statistically significant, as evidenced by confidence intervals that do not include zero. Though the mean effect size for White youth was somewhat larger than that for minority youth, this difference was not statistically significant (as shown by the overlapping confidence intervals for Whites and minorities). To illustrate in more intuitive form what these effect size values mean, Cohen's (1988) arcsine transformation was used to convert them to percentage differences in recidivism rates. If the approximate rate of recidivism for delinquents in control groups is set at 50% (which is very close to the actual rate in the database), the mean effect size of .11 for minority youth is then equivalent to about a 5 percentage point differential, that is, a 45% reoffense rate for treated juveniles compared to 50% for those in the control groups. The mean effect size of .17 for majority youth translates into a 42% reoffense rate for treated juveniles, that is, an 8 percentage point decrease from that for untreated juveniles.

For the other outcome constructs, mean effect sizes for both minority and White juveniles were greater than zero for all outcome categories except family functioning, although not all were statistically significant (confidence intervals for self-esteem and employment status overlapped zero for both groups, and those for internalizing problems and school participation overlapped for the minority samples). Thus, the mainstream interventions represented in these studies, on average, had positive effects on both subsequent delinquency and a number of other important outcomes. Of greatest

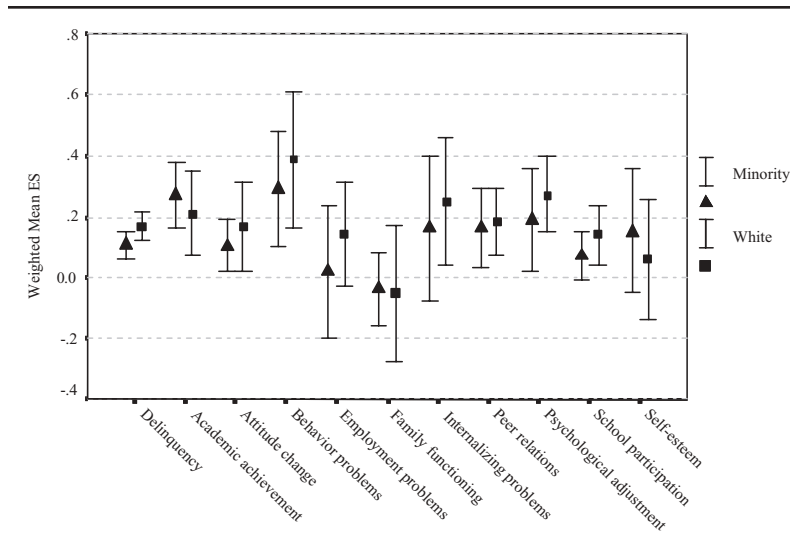


Figure 1: Weighted mean posttest effect sizes for minority and White youth for each outcome construct.

NOTE: ES = effect size.

relevance for present purposes, however, were the differences between the minority and White samples on the various outcome constructs. Although the mean effect sizes for White samples were greater than those for minority samples on 7 of the 10 nondelinquency outcome categories, none of the differences in either direction was statistically significant, as evidenced by the highly overlapping confidence intervals. Thus, without exception, across all the outcome domains represented in these intervention studies, there were no significant differences between the overall effects of mainstream intervention services on predominantly minority treatment groups and those on predominantly White treatment groups.

A CLOSER LOOK AT DIFFERENTIAL EFFECTS ON DELINQUENCY OUTCOMES

The overall mean effect size values in Figure 1 give a general affirmative answer to the question of whether mainstream delinquency interventions without special tailoring are as effective for ethnic minority youth as for White youth. However, there is great variability around the mean effects shown there. For the central delinquency outcomes, there was about three times as much effect size variability across studies as would be expected if all

the interventions produced the same effect with associated sampling error (the Q -statistic testing effect size heterogeneity was $Q_{305} = 991.5, p < .01$). Differences in delinquency outcomes observed across studies could be due to any of a number of factors, including between-study differences in method and procedure, participant characteristics, amount of treatment, and, of course, type of treatment. Indeed, the results of any one study are jointly determined by the nature of the treatments and participants in the study and the methods used to study them. It is, therefore, informative to investigate the study characteristics associated with larger or smaller effect sizes and, if necessary, to statistically control their influence on the comparison of intervention effects for minority and majority youth. The resulting comparison helps ensure that any difference, or lack of difference, found between the delinquency effect sizes for minority and majority youth represents comparable intervention circumstances and is not simply the result of a different mix of treatments, methods, and the like being used with the different ethnic groups.

The first possible source of effect size variability we examined for delinquency outcomes was differences across studies in methods and procedures. If two researchers use different methods to conduct their studies, and those differences influence the findings, it is difficult to tell whether those findings reflect the effectiveness of the interventions or the influence of the methods used to study them. Thus, methodological differences across studies are nuisance variables that could influence the observed outcomes of intervention differently for minority and majority youth. We therefore sought to identify the methodological characteristics that were related to the observed effect sizes in the studies and statistically control them. We also included several more general nuisance variables in this analysis; for instance, type of publication, which is associated with effect size because of the greater tendency for large effects to be formally published (Dickersin, 1997). For the purposes of this analysis, the studies using minority and majority samples were combined. Once influential between-study differences in method were identified and controlled, intervention effects for minority and majority samples were again compared to determine if any difference appeared.

To investigate the relationship between the methodological characteristics of the studies and the delinquency outcomes observed in those studies, we performed an inverse-variance weighted random effects multiple regression analysis using the delinquency effect sizes for all 305 studies as the dependent variable (Hedges & Olkin, 1985). Next, the regression model was used to predict the overall mean effect size for each study with the method-related independent variables held constant (i.e., assigned the mean value for all studies). This predicted value was added to the residual for each effect size to

generate a set of adjusted effect sizes that estimate the effects we would expect if all studies used similar methods.

This regression analysis showed that the delinquency effect sizes were significantly related to certain aspects of the study methods used. The model was significant ($Q_5 = 36.53, p < .01$) and about 10% of the variability in delinquency effect sizes was accounted for by the study methods used. The following method-related variables demonstrated the largest relationships with effect size in this analysis.

Pretreatment equivalence of experimental and control groups. Studies in which treatment and control groups were similar prior to treatment (indicated by random assignment to conditions, no pretest statistical differences between groups, and coder ratings of high similarity between treatment and control groups) produced smaller effect sizes than those in which treatment and control groups were not similar.

Type of publication. Unpublished technical reports tended to produce smaller effect sizes than published journal articles, books, and dissertations.

Role of evaluator. Studies in which the evaluator assumed only a research role tended to produce smaller effect sizes. That is, evaluators who were not involved in the design, planning, or delivery of treatment tended to be associated with smaller treatment effects.

Type of treatment received by control participants. Studies in which control participants received more services as part of "treatment as usual" control groups (e.g., institutionalization vs. no treatment or probation) resulted in smaller effect sizes.

Blinding in data collection. Studies in which those collecting outcome data were blind to the group status of participants produced larger effect sizes than those in which data collectors were not blind.

With the effect sizes adjusted for differences between studies on the variables identified as influential in this multiple regression analysis, we then sought to identify the treatment and participant characteristics most strongly related to observed effect sizes. This was accomplished with a second weighted random effects multiple regression analysis that used the method-adjusted effect sizes generated by the procedure described above as the dependent variable. The independent variables for this analysis were orga-

nized into four clusters of study characteristics that were entered into the regression hierarchically. The first cluster included such participant characteristics as age, gender, heterogeneity of the sample, delinquency level, and proportion with prior offenses. Note that ethnicity was not included in this analysis because we did not want to control it statistically but, rather, examine it directly at a later stage.

The remaining three clusters involved treatment characteristics. One cluster included general treatment characteristics, including the age of the program, source of participants, and program sponsorship. Another encompassed treatment dose, including implementation quality, treatment duration, treatment frequency per week, total hours of contact, and coder ratings of the intensity and meaningfulness of treatment. The final cluster represented treatment delivery personnel, format of treatment sessions, and treatment site.

The overall model was statistically significant, indicating that the array of study predictors in the model accounted for a significant proportion of the variance between effect sizes ($Q_{21} = 41.20, p < .01$). Nearly 15% of the variance among effect sizes was related to the clusters of study characteristics in the model. Of the four clusters of predictor variables, the participant cluster was the largest factor in the model, indicating that the treatment effects varied according to the characteristics of the juveniles in treatment. The individual variables in this cluster that were most influential are discussed below. The cluster relating to treatment dose was the next most important in the model, with more treatment and fewer implementation problems associated with larger effect sizes. The cluster including the variables describing treatment format and the one involving general program characteristics did not make statistically significant contributions to the overall model.

The next step was to identify the individual participant and treatment variables from the influential clusters that had the strongest relationships to effect size. This was done by dropping in a stepwise fashion the variables that made the smallest contributions to the overall model. The final reduced regression model included five variables (Table 2). As the results in Table 2 show, two variables relating to the characteristics of the juveniles in the study samples had strong, independent relationships with treatment effect: the participants' delinquency level and the proportion of the juvenile sample with prior delinquent offenses. Participant samples in which all of the juveniles had prior offense records tended to produce larger effect sizes than studies in which fewer participants (or none) had priors. In addition, samples with youth identified as delinquent but not institutionalized tended to show larger effects than those in which the youth were pre-delinquent (minor offenses but not adjudicated) or already institutionalized. This pattern is not surprising given that

TABLE 2: Weighted Mixed Effects Multiple Regression: Reduced Model

| <i>Variable</i> | β | B | p |
|--|---------|-------------------------|------|
| Participant characteristics | | | |
| Delinquents (vs. institutionalized and noninstitutionalized) | .1434 | .0946 | .012 |
| Prior offenses | .1262 | .0807 | .035 |
| Treatment characteristics | | | |
| Implementation quality | .1604 | .0578 | .004 |
| Amount of meaningful contact | .1330 | .0273 | .018 |
| Treatment delivered by juvenile justice personnel | -.1321 | -.0918 | .025 |
| Regression constant | | -.1322 | |
| Overall model | | $Q(5) = 31.01, p < .01$ | |
| Residual | | $Q(299) = 289.58, ns$ | |
| $R^2 = .10$ | | | |
| $N = 305$ | | | |

pre-delinquent samples have little delinquent involvement to begin with and thus less room to reduce their involvement as a result of effective treatment. Institutionalized youth, on the other hand, are the most serious offenders and may have problems that are less amenable to the effects of treatment.

Two variables related to treatment dose also showed strong, independent relationships with effect size. Treatments rated by coders as more meaningful in terms of their likelihood of engaging the juvenile were associated with larger effect sizes than those rated as less meaningful. Also, when difficulties in treatment implementation were reported (e.g., not all juveniles received treatment, or treatment delivery was not complete), smaller treatment effects were found, as would be expected. The final variable in the reduced model indicated that interventions delivered by juvenile justice personnel were not as effective as those delivered by mental health personnel and non-mental health counselors.

As in the previous regression analysis with methodological variables, the reduced model representing the key participant and treatment-related variables was used to statistically control for differences between studies on those variables. This was done by using the regression equation to "predict" the effect size value expected when the respective participant and treatment variables were given values equal to their means across all studies. This predicted value was then added to the residual for each effect size to create new effect size values that estimated the effect sizes that would result if all studies were equivalent on the independent variables represented in the regression model.

Minority-Majority Differences With Influential Study Characteristics Controlled

With the most influential study characteristics identified, the relative effects of intervention for minority versus majority samples could be examined while controlling for any differences associated with those study characteristics. The first regression analysis described above produced a set of *method-adjusted* effect sizes for which key methodological differences between studies were held constant. The reduced model from the second regression analysis (Table 2) then generated another set of effect sizes from the method-adjusted set for which the influential sample and treatment characteristics were also held constant. This second set of effect sizes, which we will call *equated* effect sizes, statistically controls for all the between-study differences identified by either regression analysis as having a significant relationship to observed effect size values. This level of statistical control permits comparison between studies of minority youth and those of majority youth with increased confidence that any differences in intervention effects, or lack thereof, reflect the role of ethnicity and not other influential study characteristics that happen to be unevenly distributed across the two sets of studies.

Of course, different types of intervention may have different effects, so as a further control, we compared method-adjusted and equated effect sizes separately within each major intervention category. The observed, method-adjusted, and equated effect size means and confidence intervals for minority and majority youth receiving each of the most common types of intervention are shown in Figures 2 and 3. Figure 2 shows counseling-type programs in institutional and noninstitutional settings. The services in these programs include group and individual counseling, milieu therapy, guided group interaction, family counseling, reality therapy, and the like. Figure 3 shows case-work and service brokerage-type programs, academic services such as tutoring and special classes, and probation with such juvenile justice enhancements as intensive supervision, restitution, and juvenile justice system education.

There are advantages and disadvantages associated with each of the three forms of effect size means shown in Figures 2 and 3. The observed means represent the findings of the individual studies as they were originally reported. However, their values are influenced by the various method, participant, and treatment characteristics discussed above, and those characteristics may not be uniform across studies with minority and majority youth. The method-adjusted effect sizes statistically control for between-study differences in method and procedure and thus give the best comparison of the effects

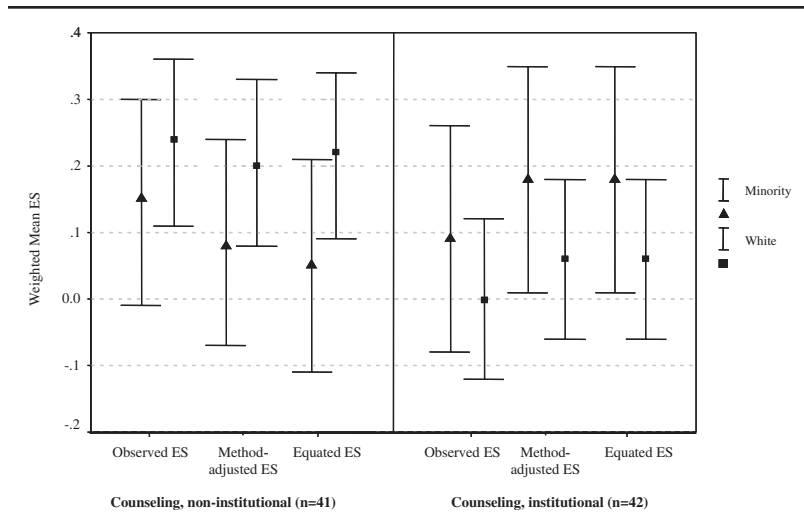


Figure 2: Mean observed and statistically adjusted effect sizes for minority and White youth for two types of counseling services.
NOTE: ES = effect size.

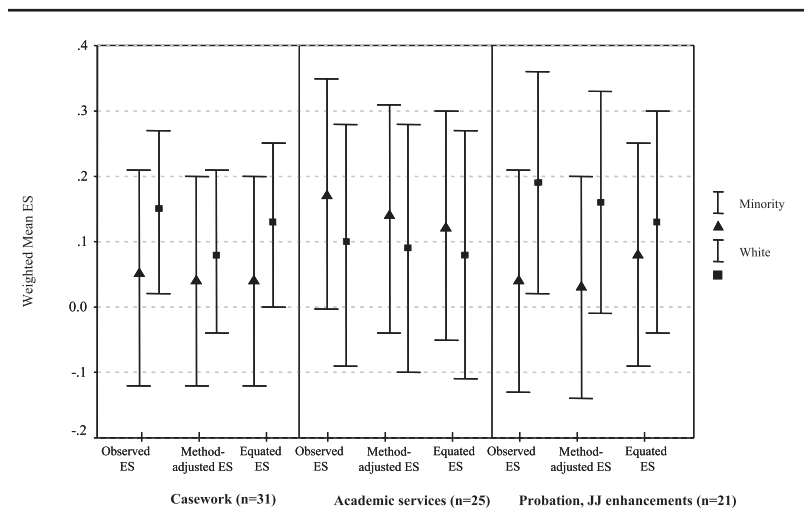


Figure 3: Mean observed and statistically adjusted effect sizes for minority and White youth for select program types.
NOTE: ES = effect size.

actually produced by the treatments that were delivered in the various studies. However, there may be important differences in the participant characteristics (e.g., delinquency level), treatment implementation and amount, personnel, and so forth between the interventions provided to the minority and majority youth in these studies that would yield different effects irrespective of ethnicity. The equated effect sizes simulate a situation in which key methodological, participant, and treatment characteristics are the same for minority and majority youth and thus permit direct comparison. But they represent configurations of method, participant, and treatment characteristics that, to some degree, were not actually provided to the youth in these studies.

With these considerations in mind, it is most appropriate to examine the three types of effect sizes together with attention to any patterns that may illuminate the differences or similarities between delinquency outcomes for minority and majority youth. The most strikingly consistent result in Figures 2 and 3 is the similarity of the delinquency effects for minority and majority youth across all the treatment types and all forms of the effect size. The overlapping confidence intervals for each pair of means show that the difference between the minority and majority means was not statistically significant for any comparison. Moreover, the pattern of nonsignificant differences does not strongly favor either group. The effect size means for minority youth are somewhat lower than those for majority youth when the interventions are noninstitutional counseling, casework, or enhanced probation services, but they are higher for academic services and counseling in institutional settings.

EFFECTS FOR MINORITY AND MAJORITY YOUTH RECEIVING THE SAME INTERVENTIONS

The results presented thus far indicate that mainstream treatments without cultural tailoring are as effective for minority youth as they are for majority youth. However, these results are all based on comparisons of different sets of studies, some using samples of minority youth and some using majority youth. Although the use of statistical controls reduces the influence of differences between the sets of studies on characteristics that are irrelevant to the issue of differential effects for minority and majority youth, the possibility remains that other differences that were not controlled still distort the comparisons. The most direct comparison of the effects of delinquency intervention programs for minority versus majority youth comes from studies with

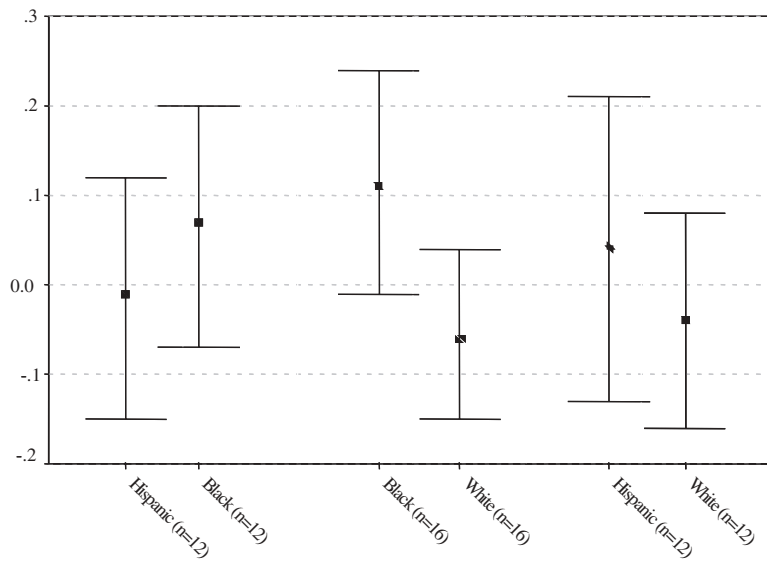


Figure 4: Differences in mean delinquency effect sizes for minority and majority groups in breakout samples.

both minority and majority participants that report effects separately for each group. Such studies would compare the outcomes for ethnic groups who received the same interventions that were evaluated under the same conditions with the same methods.

Relatively few studies in our database broke out and reported results separately for youth of different ethnicities. Figure 4 compares the mean effect sizes for each pair of ethnic groups from those studies that did compare outcomes for those groups. Thus, within each pair of effect size means, the youth in the respective ethnic groups participated side-by-side in the same treatment programs and research studies. Though the numbers are small, the results shown in Figure 4 are completely consistent with those from the other analyses reported above. None of the effect size differences between ethnic groups was statistically significant and the nonsignificant trends in those differences were in the direction of larger effects for minority youth than White youth. The limited data that permit direct comparison, therefore, also fail to support the view that the effects of mainstream programs for delinquency favor majority youth and are less effective with minorities.

DISCUSSION AND APPLICATIONS TO PRACTICE

The analyses reported above provide no evidence that mainstream delinquency intervention programs yield poorer outcomes for minority youth than for White youth despite their general lack of cultural tailoring for minority clientele. A large, representative selection of intervention studies showed no significant differences between minority and White samples in any outcome domain, including effects on delinquency, academic achievement, behavior problems, self-esteem, employment status, peer relations, internalizing problems, attitudes, school participation, family functioning and psychological adjustment. Additional analysis of delinquent reoffending, the major target outcome for these programs, further supported the initial finding of no difference in effects for minority versus White youth. In particular, introducing a range of statistical controls for methodological and substantive differences among the studies into the analysis did not appreciably alter the results. Moreover, direct comparison of outcomes for minority and White youth receiving the same treatments in the same studies using the same methods and measures also showed no differences.

A more interesting, and more definitive, analysis would involve a three-way comparison of the effects of culturally tailored programs for minority youth with the effects of mainstream programs on both minority and White youth. However, too few controlled studies of the outcomes of culturally tailored programs for minorities have been conducted and reported to permit such analysis. The database from which the studies for the meta-analysis presented here were drawn was developed through vigorous search for all qualifying published and unpublished studies conducted between 1950 and 1996, and we believe it provides very good coverage of the extant research for that period. When we examined each of these studies for indications of culturally tailored interventions, however, we found only one (Wooldredge et al., 1994) of the nearly 500 available that clearly involved such tailoring. The effect size for this study was .03, indicating that treatment group youth were not better off than control group youth after participating in culturally tailored treatment. Another 13 reported using minority personnel to provide services to minority youth but gave no indication that the nature of the service itself was otherwise adjusted on the basis of cultural considerations for those youth. The recidivism effect size for the programs with minority service providers was .13. Even if these latter cases are counted as minimal instances of cultural tailoring, there are too few studies to permit an adequate comparison of their outcomes with those of comparable mainstream programs for comparable minority youth. Furthermore, the results of these few studies do not suggest that tailoring by matching the cultural characteristics of juveniles and

providers produces better results than mainstream programming. Increased recent interest in culturally sensitive intervention programs may generate enough outcome research to support such comparison in the near future.

It should be noted that the mean effect sizes found for both minority and White youth in this meta-analysis are relatively modest. As Figure 1 indicates, there was no outcome domain for which the mean for either group exceeded .40, and most were in the range of .20 and below. The critical delinquency outcomes, in particular, showed a mean effect size of .11 for minority youth and .17 for White youth. By comparison with these values, the mean effect size for over 300 meta-analyses of the outcomes of psychological, educational, and behavioral interventions found by Lipsey and Wilson (1993) was .50.

One possible interpretation of the results presented in this study, therefore, is that the mainstream delinquency intervention programs reviewed are not generally successful in producing positive outcomes. Thus, the lack of any significant differences between the outcomes for minority and White youth demonstrates not that these programs are equally effective for minorities as Whites despite their lack of cultural tailoring but that they are equally ineffective for both groups of youth. The similarity in the outcomes of mainstream programs for minorities and Whites is only interesting if they have meaningful positive effects on both groups. If the programs do not work, no defensible case could be made for applying them to minority youth no matter how similar the results were to those for White youth.

For several reasons, we do not think the above interpretation of the results is correct. First, numerically small values of the standardized mean difference effect size statistic do not necessarily indicate that the practical significance of the effects is small. The mean effect sizes for the key delinquency reoffense outcomes, for instance, translate into 5-8 percentage point decreases from a 50% recidivism baseline among control groups. This difference, therefore, represents a 10%-16% reduction in the number of juveniles reoffending, which is far from trivial even though one could hope for more.

And, indeed, larger effects on delinquency are represented in the mean effect sizes found. The distributions of effect sizes for minority and White youth summarized by the respective mean values have relatively large variance. That is, the mean values average over a wide range of delinquency effects, some much smaller than the mean but many that are much larger. In other analyses, we have shown that the effects produced by the high-end delinquency interventions are considerable, ranging as high as 40% reductions in recidivism (Lipsey, 1995; Lipsey & Wilson, 1998). One implication of this state of affairs is that the mean effect size is not an especially good summary of the full effect size distribution. Another implication, however, is

that the statistically nonsignificant overall difference between the mean for minority youth and that for White youth encompasses the full range of effects, including those of unquestionable practical significance. This can be seen most clearly in Figures 2 and 3 where the confidence intervals around the mean effect sizes for different types of interventions are very broad, ranging upward to twice or three times the mean value, as well as downward an equivalent amount.

Overall, therefore, we believe the most defensible interpretation of the available research is that mainstream treatments for juvenile delinquents are generally effective and no less effective for ethnic minority youth than White youth. We must emphasize, however, that this does not mean that issues of cultural sensitivity are unimportant to such programs when minority youth are served. It could well be that the effects of programs with cultural tailoring would be larger than those of programs without even though those without do not have differential effects for minority and White youth. The evidence reviewed here only shows that cultural tailoring is not necessary for the programs to have positive outcomes and that the absence of such tailoring does not diminish the effects for minorities relative to Whites. As noted earlier, there are not yet sufficient outcome studies for programs with cultural tailoring to determine if they yield larger effects than comparable programs without such tailoring. In addition, all of the studies included here involved indigenous minority youth, rather than recent immigrants. Thus, our results do not speak to the particular needs of recent immigrant populations or the effectiveness of mainstream intervention for delinquent youth who are newly arrived.

Moreover, even if the major outcomes of mainstream programs for minority youth are comparable to those with White youth, there may be other benefits to culturally sensitive programming. It may well be that the likelihood of participation, the acceptance of the program plan, the ultimate satisfaction with the program experience, and other such factors not commonly measured in outcome studies are less positive for minority youth in mainstream programs than majority youth. This may be especially so for recent immigrant youth. Moreover, if such differences occur, programs especially tailored to specific ethnic groups may well alleviate them. What the evidence reviewed here indicates is only that such tailoring cannot be justified on the grounds that, without it, the programs are ineffective or not as effective as for majority youth. In addition, it is worth noting that in the course of developing culturally sensitive programs, some care must be taken to ensure that such tailoring does not reduce the effectiveness of the mainstream programs that are adapted. Ultimately, we must implement and evaluate a sufficient number and range of culturally tailored programs for delinquent youth to permit a

direct assessment of their outcomes and how they compare with those of mainstream programs.

REFERENCES

- Addams, J. (1895). *Hull House maps and papers. A presentation of nationalities and wages in a congested district of Chicago, together with comments and essays on problems growing out of the social conditions*. New York: Thomas Y. Crowell.
- Barth, R. P., & Blackwell, D. L. (1998). Death rates among California's foster care and former foster care populations. *Children and Youth Services Review, 20*, 577-604.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum.
- Cooper, H., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis*. New York: Russell Sage.
- de Anda, D. (Ed.). (1997). *Controversial issues in multiculturalism*. Boston: Allyn and Bacon.
- Devore, W., & Schlesinger, E. G. (1996). *Ethnic-sensitive social work practice*. Boston: Allyn and Bacon.
- Dickersin, K. (1997). How important is publication bias? A synthesis of available data. *AIDS Education and Prevention, 9*, 15-21.
- Green, J. W. (1995). *Cultural awareness in the human services. A multiethnic approach*. Boston: Allyn and Bacon.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Jonson-Reid, M., & Barth, R. P. (2000). From placement to prison: The path to adolescent incarceration from child welfare supervised foster or group care. *Children and Youth Services Review, 22*, 493-516.
- Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation: A casebook* (pp. 83-127). New York: Russell Sage.
- Lipsey, M. W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents. In J. McGuire (Ed.), *What works? Reducing reoffending—Guidelines from research and practice* (pp. 63-78). New York: John Wiley.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181-1209.
- Lipsey, M. W., & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Applied Social Research Methods Series, Vol. 49). Thousand Oaks, CA: Sage.
- Lum, D. (1996). *Social work practice and people of color. A process stage approach*. Monterey, CA: Brooks/Cole.
- Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261-281). New York: Russell Sage.

- Soydan, H. (1999). *The history of ideas in social work*. Birmingham, AL: Venture Press.
- Soydan, H., Jergeby, U., Olsson, E., & Harms-Ringdal, M. (1999). *Socialt arbete med etniska minoriteter: En litteratöversikt* [Social work with ethnic minorities. A literature review]. Stockholm: Liber.
- Williams, C., Soydan, H., & Johnson, M. R. D. (Eds.). (1998). *Social work and minorities. European perspectives*. London: Routledge.
- Wooldredge, J., Hartman, J., Latessa, E., & Holmes, S. (1994). Effectiveness of culturally specific community treatment for African American juvenile felons. *Crime & Delinquency*, 40, 589-598.