

What Works for Female Offenders: A Meta-Analytic Review

Craig Dowden
D. A. Andrews

Although the question of what works for general offender populations has received considerable attention within the rehabilitation literature, very little research has examined female offenders. The present investigation examined the principles of effective correctional treatment for female offenders through a meta-analytic review. The results indicated that the clinically relevant and psychologically informed principles of human service, risk, need, and responsivity identified in past meta-analytic reviews were associated with enhanced reductions in reoffending.

Considerable debate has occurred regarding the effectiveness of correctional interventions. Although the notion that nothing works (Martinson 1974) predominated in the 1970s, the advent of meta-analytic research has swung the pendulum to a what works perspective (McGuire 1995). Several meta-analytic reviews have been conducted on the rehabilitation literature and, for the most part, have suggested that some types of correctional interventions can effectively reduce recidivism (Andrews, Dowden, and Gendreau forthcoming; Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen 1990; Dowden 1998; Hill, Andrews, and Hoge 1991; Izzo and Ross 1990; Lipsey 1989, 1995; Losel 1996).

Although the effectiveness of rehabilitation for general offender populations has received widespread attention (Gibbons 1999), research dedicated to female offender populations has been quite limited (Koons, Burrow, Morash, and Bynum 1997). Morash, Bynum, and Koons (1995) identified 67 studies that reported promising intervention strategies for female offenders, but only 12 included an outcome measure, and none linked recidivism to program components.

A qualitative survey study conducted by Koons and her colleagues (1997) also attempted to identify promising intervention strategies for female offenders. In their report, correctional administrators identified treatment

CRAIG DOWDEN: Doctoral candidate, Carleton University, Ottawa, Ontario, Canada.
D. A. ANDREWS: Professor, Department of Psychology, Carleton University, Ottawa, Ontario, Canada.

CRIME & DELINQUENCY, Vol. 45 No. 4, October 1999 438-452
© 1999 Sage Publications, Inc.

needs that they believed were related to successful treatment outcome. These needs included substance abuse education and treatment and the development of parenting and life skills as well as interpersonal and basic education skills. Interestingly, these program targets only partially overlapped with the list of more promising targets for change (i.e., criminogenic needs) outlined by a number of scholars in the correctional treatment literature (Andrews and Bonta 1998; Andrews, Bonta, and Hoge 1990).

Koons and her colleagues (1997) argued that female offenders have several unique needs and concerns such as childcare, pregnancy, and sexual or physical abuse victimization. Nonetheless, their survey results and introductory overview of the what works literature suggested that the principles of effective correctional treatment (i.e., risk, need, and responsivity) may apply regardless of the gender of the treatment population. At the same time, the authors stated that "the question of whether or not these findings (i.e., principles of effective correctional treatment) can be generalized to the female offender population still is very much in need of an answer" (p. 517).

What are the principles of risk, need, and responsivity?

Essentially, the risk principle is concerned with identifying those clients who should receive the most intensive allocation of correctional treatment resources and those who require less attention. In other words, this principle states that the amount of intervention that an offender receives must be matched to his or her risk level to reoffend. The highest levels of service should be reserved for the higher risk cases, whereas the minimal levels of service and supervision should be provided for the lower risk cases (Andrews and Bonta 1998; Andrews, Bonta, et al. 1990).

The need principle, on the other hand, is concerned with the targets for change identified within the treatment program. More specifically, theory and some research have demonstrated that when certain risk factors have been altered, reductions in recidivism for offenders have occurred. These more promising targets or dynamic risk factors have been commonly referred to as criminogenic needs (Andrews and Bonta 1998; Andrews, Bonta, et al. 1990), and they are different from other less promising targets that are classified as noncriminogenic needs. Past research has suggested that programs that targeted these latter needs have not led to significant reductions in recidivism (Andrews and Bonta 1998; Andrews, Bonta, et al. 1990). Accordingly, the need principle states that if the primary goal of treatment is reduced recidivism, the criminogenic needs of offenders must be emphasized and targeted.

Although the need principle has received empirical support within a number of meta-analytic reviews (Cleland, Pearson, and Lipton 1996; Dowden 1998), research has been sorely lacking regarding the applicability of this principle to female offenders. More specifically, research has not focused on whether the criminogenic needs of female offenders are the same as those for general offender populations. For example, one of the promising treatment targets identified within the Koons, Burrow, Morash, and Bynum (1997) study involved focusing on issues of victimization and targeting an offender's level of self-esteem. Traditionally, self-esteem has been classified as a non-criminogenic need (Andrews and Bonta 1998). One of the main goals of the present meta-analytic investigation was to determine whether the list of criminogenic and noncriminogenic needs identified by Andrews and his colleagues (Andrews and Bonta 1998; Andrews, Bonta, et al. 1990) was valid when applied to female offenders.

A third principle of effective correctional treatment is responsivity. This principle is directly concerned with the characteristics of program delivery, and it states that the styles and modes of service used within a treatment program should be matched to the learning styles of offenders. Both general and specific responsivity considerations are encompassed within the responsivity principle. General responsivity states that the most effective types of service are based on cognitive-behavioral and social learning approaches. Specific responsivity focuses on offender characteristics such as interpersonal sensitivity, interpersonal anxiety, and verbal intelligence to name a few. For example, Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990) stated that "the success of highly verbal, evocative, and relationship-dependent services seems to be limited to clients with high levels of interpersonal, self-reflective, and verbal skill" (p. 376). However, the meta-analytic evidence to date has only examined the effectiveness of the general responsivity principle and has suggested that the most powerful treatment approaches are those that have used concrete social learning and behavioral strategies.

Support for the clinical effectiveness of these three principles was derived from the Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990) meta-analysis, in which a four-level type-of-treatment variable was used to identify effective correctional treatment programs. This four-level variable consisted of criminal sanctions, inappropriate service, unspecified service, and appropriate service. The results revealed that the appropriate service category (programs that adhered to each of the principles of risk, need, and responsivity) was associated with significantly larger reductions in reoffending when compared to each of the remaining categories. Based on these results, the authors concluded that the most effective correctional interventions were ones that

incorporated the principles of effective correctional treatment within their program framework.

METHODOLOGY

Sample of Studies

The present study combined two distinct samples of studies. The first was taken directly from Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990), whereas the second ($n = 220$) was composed of additional studies reported by Andrews (1996), Andrews and Bonta (1998: Resource Note 10.1), and Dowden (1998). Studies selected to be included in the present meta-analysis possessed the following characteristics:

- (1) The study was composed predominantly or entirely of female offenders.
- (2) The study included a follow-up period. If several follow-up periods were reported, data from the longest follow-up period were coded to ensure the maximum time at risk in the community.
- (3) The study compared a group of offenders who received some form of intervention to a control group who did not receive the primary intervention. Individual control groups could have received a diluted form of the treatment program and could even have received alternate services as long as these services could be differentiated from those received by the treatment group.
- (4) A measure of recidivism was included in the report. Recidivism was defined in several ways. Acceptable definitions included rearrest, reconviction, and parole failures or revocations. The preferred measure of recidivism was reconviction.

Variables Included in the Analysis

Risk. Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990) directly coded the risk principle only when the primary study reported outcome data for high- and low-risk groups separately. Consequently, an overall examination of the risk principle for their entire sample was not available. For the present meta-analysis, an aggregate approach to coding risk (Lipsey 1989) was introduced. In other words, each study was coded as involving higher or lower risk offenders, depending on whether the majority of those in the study had penetrated the justice system at the time of the study or had a previous criminal offense. The study was coded as involving high-risk cases if either of these characteristics were present. The advantage of this approach over the one originally used by Andrews, Zinger, Hoge, Bonta, Gendreau,

and Cullen (1990) was that it allowed a risk score to be entered for each study included in the meta-analysis.

General responsivity. Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990) only coded for general responsivity within their meta-analysis. To ensure consistency and facilitate comparison of the findings, this method of coding the responsivity principle was maintained. More specifically, social learning or cognitive-behavioral programs that used modeling, role-playing, reinforcement, and graduated practice were considered to be appropriately addressing general responsivity.

Need. Each individual treatment program was examined to determine whether any of the more or less promising targets for intervention outlined by Andrews and Bonta (1998) had been included within it. Once all of the needs within a particular treatment program were identified, an overall need variable was constructed by considering the difference score between the number of criminogenic and noncriminogenic needs targeted within the treatment program of interest. The variable was a binary measure (0 = the difference score was less than or equal to 0 and 1 = the difference score was greater than or equal to 1). This variable was used as the overall test of the need principle. In other words, programs targeting more criminogenic than noncriminogenic needs were considered to be appropriately addressing the need principle; conversely, programs that targeted an equal or greater number of noncriminogenic needs than criminogenic ones were considered to be inappropriately addressing the need principle.

Type of treatment. Because, in the present investigation, the principles of risk, need, and responsivity were examined separately, the next logical step involved the development of a type-of-treatment variable similar to the one used by Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990). This variable had to do with how well the principles of risk, need, and responsivity were incorporated within a particular treatment program, and it was an objective extension of the original coding procedures used by Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990). The scoring for the variable was based on the composite risk, need, and responsivity scores, with a simple count to determine the number of principles that the treatment program appropriately addressed. This score determined the type-of-treatment rating received by each particular treatment program.

The possible range of scores for this variable ranged from 0 to 3. Criminal sanctions without the provision of human service were automatically coded

0. The four categories, in order of appropriateness, were *inappropriate service*, *weak service*, *promising service*, and *most promising service*. For example, a human service program that incorporated only two of the three principles of risk, need, and responsivity was coded as a promising service.

Procedure

Because the meta-analysis was conducted on a subsample of studies taken from a larger meta-analytic review and the same set of variables was used, we assumed that the interrater agreement ratings would remain the same; therefore, the results for the interrater reliability reported below were based on those presented in the original meta-analysis (Dowden 1998).

The first author coded all of the studies that met the inclusion criteria for the present meta-analysis, using the previously mentioned coding procedures. An honors student in psychology was also trained to use the coding manuals and was given a preliminary sample of five studies to code. Once the student had finished coding these studies, any discrepancies in scoring were discussed with the first author. This procedure ensured that the other rater fully understood the underlying constructs presented within the coding manuals. Once the other coder felt comfortable with the coding manuals, the first author provided him with a random sample of 29 studies, equally drawn from the justice, inappropriate, unspecified, and appropriate categories identified by Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990).

The interrater reliability ratings were calculated by dividing the total number of correct classifications by the total number of classified variables. The rates of agreement for the core variables were 100 percent (any treatment, $r = 1.00$) and 90 percent for behavioral ($r = .79$), risk level ($r = .79$), and criminogenic need ($r = .79$). The interrater agreement rate was 76 percent ($r = .88$) on the four-level type-of-treatment variable.

Calculation of effect sizes. The effect-size measure used in the current study was the phi coefficient. Phi was used as the measure of treatment effect because it provides the magnitude and direction of the association between two binary variables (treatment participation and recidivism) and it is equivalent to the Pearson product-moment correlation coefficient. A valuable characteristic of phi is that it can be translated into the Binomial Effect Size Display (BESD; Rosenthal 1991). The BESD converts this statistic into a value that reflects the difference between the recidivism rates of the treatment group and the control group (assuming a base rate of recidivism of 50 percent and an equal number of cases in each group). For example, using the BESD, a

mean correlation coefficient of .20 translates into a recidivism rate of 40 percent for the treatment group (i.e., 50 percent $-20/2$) and a corresponding recidivism rate of 60 percent for the control group (50 percent $+20/2$).

It should be noted that the analyses were conducted on the unweighted effect-size estimates. Both weighted and unweighted effect-size estimates have been reported in the literature, and we decided to use the unweighted estimates for several reasons. Most importantly, the least-square approaches that can be conducted on the unadjusted estimates allow for a more sophisticated and effective exploration of the hypotheses. For example, several potential moderating variables can be identified, and their independent and joint contributions to effect size can be determined.

Analyses. Two sets of analyses were conducted for each of the major variables. The first set focused on the entire sample of treatment outcome studies in which female offenders predominated to ensure that the maximum number of effect sizes contributed to the analyses. The second set of analyses was conducted on only those studies that were composed entirely of female offender populations. This set of analyses provided an opportunity to examine whether any serious discrepancies existed between the predominantly and entirely female offender studies.

RESULTS AND DISCUSSION

What Works: A Focus on Risk, Need, Responsivity, and the Most Promising Service

The present meta-analysis consisted of 45 effect sizes extracted from 26 unique studies that examined the effectiveness of correctional treatment programs for female offenders. Sixteen of the studies were composed entirely of female offenders and contributed 24 effect sizes to the analysis. The small number of studies contributing to the meta-analysis highlighted the relative lack of research that has been conducted on the effectiveness of correctional treatment for female offenders.

A wide range of effect sizes was found within the entire sample of studies. More specifically, these varied from -0.43 to $+0.82$. The overall, mean effect size for the sample was $+0.14$ ($SD = .24$) with a 95 percent confidence interval of $+0.07$ to $+0.21$. Using the BESD introduced earlier, this value represented a recidivism rate of 43 percent for the treatment group and 57 percent for the control group. Interestingly, when the analysis focused exclusively on female offenders, the mean effect size was $+0.17$ ($SD = .24$).

TABLE 1: Mean Effect Size for Each Level of Human Service, Risk, Need, and Responsivity

Variable Label	Adheres to Principle		
	No (k^a)	Yes (k)	η
Human service			
Predominantly female	.01 (10)	.18 (35)	.31*
Solely female	.02 (4)	.20(20)	.29
Risk			
Predominantly female	-.04 (9)	.19 (36)	.40**
Solely female	-.04 (6)	.24 (18)	.51**
Need			
Predominantly female	.04 (24)	.26 (21)	.49**
Solely female	.09 (11)	.23 (13)	.32
Responsivity			
Predominantly female	.08 (30)	.27 (15)	.38**
Solely female	.12 (16)	.25 (8)	.26

a. k refers to the number of studies that contributed to the mean effect size of interest.
* $p < .05$. ** $p < .01$.

We hypothesized that human service in a justice context would yield greater reductions in recidivism than criminal sanctioning. Although the mean effect size for criminal sanctions was mildly positive (+0.01; $k = 10$; $SD = .07$), analysis of variance revealed that the mean effect size for human service interventions (+0.18; $k = 35$; $SD = .25$) was associated with a significantly greater mean reduction in recidivism ($\eta = .31$, $p < .04$). Using the BESD, the mean correlation coefficient for human service studies translated into a recidivism rate of 41 percent for human service programs and 59 percent for the control/comparison group. The magnitude of this trend was maintained when the analyses focused exclusively on female offenders (see Table 1).

These results suggested that human service programs played an important role in determining the therapeutic potential of a particular intervention. Further analyses were conducted to examine whether the principles of risk, need, and responsivity provided additional information concerning what works for female offenders.

Risk, need, and responsivity principles. Strong support was found for each of the principles of risk, need, and responsivity in the meta-analysis (see Table 1). More specifically, stronger treatment effects were revealed in programs that targeted higher versus lower risk cases ($\eta = .31$), predominantly focused upon criminogenic versus noncriminogenic needs ($\eta = .49$), and used behavioral-social learning versus nonbehavioral treatment strategies ($\eta = .38$).

TABLE 2: Mean Effect Sizes for Each Level of Type of Treatment

<i>Level of Type of Treatment</i>	<i>Predominantly Female (k^a)</i>	<i>Solely Female (k)</i>
Inappropriate service	.02 (14)	.03 (6)
Weak service	.03 (10)	.10 (5)
Promising service	.17 (9)	.18 (7)
Most promising service	.36 (12)	.34 (6)

a. *k* refers to the number of studies that contributed to the mean effect size of interest.

Type of treatment. The mean effect sizes for each level of the type-of-treatment variable are presented in Table 2. Using the Scheffe correction, the appropriate treatment category yielded a significantly higher mean effect size compared to the weak and inappropriate programs ($p < .05$). These findings replicated the results of Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen (1990) and suggested that programs that appropriately implemented the principles of risk, need, and responsivity within their framework were associated with reductions in reoffending.

Supplementary analyses of need. The percentage distributions for each of the more and less promising targets for intervention are listed in Tables 3 and 4, respectively. The mean effect size for each need, when it was and was not targeted, and the magnitude of the association with effect size is presented in these tables.

The categorizations used for each of the criminogenic and noncriminogenic need targets were derived directly from a recent meta-analysis of the need principle conducted by Andrews, Dowden, and Gendreau (forthcoming). They combined similar criminogenic and noncriminogenic need factors to create composite need categories. These categories and their subcomponents are presented in Tables 3 and 4.

Regarding criminogenic needs, Table 3 reveals that 31 percent of the treatment programs focused on interpersonal criminogenic need targets, defined as family process or antisocial associate variables. These need targets yielded the strongest positive association with reduced reoffending ($\eta = .45$). The other category of criminogenic needs, personal criminogenic need targets, included focusing on either antisocial cognition or self-control deficits. This category also yielded a significant positive correlation with effect size ($\eta = .37$).

An interesting pattern of results emerged when the analyses shifted to the school/work and substance abuse variables and their corresponding relationships with recidivism. More specifically, both school/work ($-.08$) and substance abuse ($-.01$) had slightly negative correlations with reduced reoffending.

TABLE 3: Criminogenic Needs and Their Magnitude of Correlation With Effect Size: Percentage of Tests With Need Targeted, Mean Effect Size When and When Not Targeted and Correlation With Effect Size

Need Area Targeted	Percentage	Mean Phi (k) ^a		Correlation With Phi
		Not Targeted	Targeted	
Personal criminogenic targets				
Antisocial cognition and skill deficits	18	.11 (37)	.31 (8)	.32*
Antisocial cognition	11	.11 (40)	.38 (5)	.36*
Self-control deficits	11	.13 (40)	.22 (5)	.12
Interpersonal criminogenic targets				
Family and peers	31	.07 (31)	.30 (14)	.45**
Family process	20	.08 (36)	.38 (9)	.51***
Antisocial associates	14	.09 (36)	.35 (9)	.45**
School/work	18	.15 (38)	.10 (7)	-.08
Substance abuse	11	.14 (40)	.14 (5)	-.01

NOTE: Two targets occurred in less than 5 percent of the studies: Relapse prevention ($k=1$, $r=N/A$) and barriers to treatment ($k=3$, $r=.52^{***}$). Components of antisocial cognition include antisocial attitudes ($k=1$, $r=N/A$) and anger ($k=4$, $r=.34^*$). Components of family process include affection ($k=9$, $r=.51^{***}$) and supervision ($k=4$, $r=.62^{***}$). Components of antisocial associates include increase contact with prosocial ($k=8$, $r=.39^{**}$) and decrease contact with antisocial ($k=3$, $r=.08$ n.s.). Components of school/work include school ($k=6$, $r=-.04$ n.s.) and vocational skills ($k=1$, $r=N/A$). Components of substance abuse include treatment ($k=4$, $r=.03$ n.s.) and information ($k=3$, $r=.08$ n.s.).

a. k refers to the number of studies that contributed to the mean effect size of interest.

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

Personal and interpersonal noncriminogenic need targets were also examined. The results revealed that almost one quarter of the effect sizes targeted personal noncriminogenic needs. Not surprisingly, personal noncriminogenic needs had a slightly negative correlation with effect size ($-.03$). On the other hand, interpersonal noncriminogenic needs were targeted in only 13 percent of the cases and yielded a moderate negative association with reduced recidivism ($-.23$). An even more interesting finding appeared when analyses focused on other noncriminogenic forms of family interventions. Although the family-functioning variables classified as criminogenic needs were associated with enhanced reductions in reoffending, vague forms of family intervention (i.e., providing family counseling without specifically identifying the targets for intervention) had a significant negative relationship with effect size ($-.32$). Clearly, the specific targets of family intervention were important indicators of the therapeutic potential of these particular programs.

TABLE 4: Noncriminogenic Needs Rank Ordered by Magnitude of Correlation With Effect Size: Percentage of Tests With Need Targeted, Mean Effect Size When and When Not Targeted, and Correlation With Effect Size

Need Area Targeted	Percentage	Mean Phi (k) ^a		Correlation With Phi
		Not Targeted	Targeted	
Personal noncriminogenic targets	24	.15 (34)	.13 (11)	-.03
Vague emotional/personal problems	20	.14 (36)	.15 (9)	.02
Interpersonal noncriminogenic targets	13	.16 (39)	.01 (6)	-.23
Family (other)	11	.17 (40)	-.06 (5)	-.32*

NOTE: Low frequency targets include targeting of respect for criminal thinking ($k=1$, $r=N/A$) and cohesive peers ($k=1$, $r=N/A$). Components of personal noncriminogenic targets include fear of official punishment ($k=2$, $r=-.10$) and respect criminal thinking ($k=1$, $r=-.12$). Components of interpersonal noncriminogenic targets include increase cohesiveness of antisocial peer groups ($k=1$, $r=N/A$).

a. k refers to the number of studies that contributed to the mean effect size of interest.

* $p < .05$.

Methodological considerations. Several important methodological factors were also considered to determine their relationship with effect size. The majority of these were taken directly from the strongest methodological variables identified by Andrews, Dowden, and Gendreau (forthcoming) in their examination of the strongest potential threats to validity (see Table 5). The remaining two variables, age and randomness of design, were included due to their importance within both meta-analytic and treatment outcome research. The simple and partial correlation coefficients (with type of controlled treatment) for each of these variables are presented in Table 5.

Only the involved evaluator variable was significantly correlated with effect size. Although some individuals have claimed that an involved evaluator is a biasing factor in rehabilitation research, we share Lipsey's (1995) view that an involved evaluator may not be a sign of bias but a sign of increased therapeutic integrity. In other words, having an involved evaluator in the implementation and evaluation of the program increases the likelihood that the program will be correctly implemented. Evaluators who are not so involved may not take the same time and consideration in the preliminary stages of program implementation.

CONCLUSION

The meta-analysis reported here indicated that the principles of risk, need, and general responsivity were important contributors to treatment outcome

TABLE 5: The Most Important Methodological and Control Variables and Their Relationships With Effect Size

<i>Variable</i>	<i>Frequency (percentage)</i>	<i>Simple r</i>	<i>Partial r</i>
CJ ^a sponsor	84	-.29	-.08
CJ ^a referral	87	.34	.21
Nonresidential	71	-.08	.25
Small sample	58	.26	.12
New program	38	.26	-.17
Involved evaluator	18	.64***	.53***
Older (age 18+)	22	.03	.17
Random assignment	31	-.02	-.10

a. CJ stands for criminal justice.

*** $p < .001$.

for female offenders. These results support previous theoretically derived viewpoints (Andrews and Bonta 1998; Andrews, Bonta, et al. 1990; Andrews, Zinger, et al. 1990).

However, some more specific findings should be emphasized. One of these was that the most promising targets for intervention (i.e., substance abuse and basic education skills) that were identified by a national sample of correctional administrators (Koons et al. 1997) did not emerge as important in our study, rather, the strongest predictors of treatment success were interpersonal criminogenic need targets and, in particular, family process variables. Furthermore, it is clear that personal and interpersonal noncriminogenic needs were not related to treatment outcome; in fact, they were associated with recidivism increases within the treatment group.

Although only a moderate number of effect sizes contributed to the present meta-analysis, it is significant that the principles of risk, need, and responsivity survived all of the strongest threats to validity documented within previous meta-analytic research (Andrews, Dowden, et al. forthcoming). In addition, treatment programs that appropriately addressed these principles were associated with enhanced reductions in recidivism.

The Koons, Burrow, Morash, and Bynum (1997) study suggested that programs that focused on dealing with past victimization issues and targeted self-esteem are promising targets for change for female offenders. Because none of the studies discussed here focused on these treatment targets, it remains unclear as to whether these are criminogenic or noncriminogenic needs for female offenders. These are matters for future study.

Despite these overall findings, one final point should be made. Although the present investigation explored the effectiveness of the principles of risk, need, and responsivity for female offender populations, it did not look at gen-

der as a specific responsivity consideration. More specifically, we did not examine whether making the treatment program more responsive to the specific learning styles of women offenders (i.e., relationship-oriented treatment) had any impact on recidivism. Exploring the effects of gender as a specific responsivity consideration will be the focus of a future meta-analysis.

In conclusion, although promising evidence was uncovered for each of the principles of risk, need, and responsivity, more work needs to be done.

APPENDIX

The following studies were included in the meta-analysis. Studies that were composed predominantly or entirely of female offenders met our inclusion criteria. This standard was adopted to ensure that the maximum number of effect sizes contributed to the meta-analysis. Please note the similarities between the trends for the entire set of studies and those that solely involved female offenders in Table 1 and Table 2.

- Alexander, James F., Cole Barton, R. Steven Schiavo, and Bruce V. Parsons. 1976. "Systems-Behavioral Intervention With Delinquents: Therapist Characteristics, Family Behavior, and Outcome." *Journal of Consulting and Clinical Psychology* 44:556-664.
- Borduin, Charles M., Barton J. Mann, Lynn T. Cone, Scott W. Henggeler, Bethany R. Fucci, David M. Blaske, and Robert A. Williams. 1995. "Multisystemic Treatment of Serious Juvenile Offenders. Long-Term Prevention of Criminality and Violence." *Journal of Consulting and Clinical Psychology* 63:569-78.
- Buckner, John C. and Meda Chesney-Lind. 1983. "Dramatic Cures for Juvenile Crime: An Evaluation of a Prisoner-Run Delinquency Prevention Program." *Criminal Justice and Behavior* 10:227-47.
- Davidson, William S. and Timothy R. Wolfred. 1977. "Evaluation of a Community-Based Behavior Modification Program for Prevention of Delinquency." *Community Mental Health Journal* 13:296-306.
- Davies, Jean and Nancy Goodman. 1972. *Girl Offenders Aged 17 to 20 Years*. London, UK: H.M. Stationary Office.
- Deng, Xiaogang. 1997. "The Deterrent Effects of Initial Sanction on First-Time Apprehended Shoplifters." *International Journal of Offender Therapy and Comparative Criminology* 41:284-97.
- Druckman, Joan M. 1979. "A Family-Oriented Policy and Treatment Program for Female Juvenile Status Offenders." *Journal of Marriage and the Family* 41:627-36.
- Gordon, Donald A., Jack Arbutnot, Kathryn E. Gustafson, and Peter McGreen. 1988. "Home-Based Behavioral Systems Family Therapy With Disadvantaged Juvenile Delinquents." *The American Journal of Family Therapy* 16:243-55.
- Gruher, Martin. 1979. "Family Counseling and the Status Offender." *Juvenile and Family Court Journal* 30:23-7.
- Johnson, David C., Ronald W. Shearon, and George M. Britton. 1974. "Correctional Education and Recidivism in a Women's Correctional Center." *Adult Education* 24:121-9.
- Klein, Nancy C., James F. Alexander, and Bruce V. Parsons. 1977. "Impact of Family Systems Intervention on Recidivism and Sibling Delinquency: A Model of Primary Prevention and Program Evaluation." *Journal of Consulting and Clinical Psychology* 3:469-74.