

GENDER-RESPONSIVE SERVICES IN THE JUVENILE JUSTICE SYSTEM:
EFFECTIVENESS OF GIRLS' GROUP HOME INTERVENTIONS

By

Valerie Rose Anderson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Doctor of Philosophy

2015

ABSTRACT

GENDER-RESPONSIVE SERVICES IN THE JUVENILE JUSTICE SYSTEM: EFFECTIVENESS OF GIRLS' GROUP HOME INTERVENTIONS

By

Valerie Rose Anderson

The proportion of girls involved in the juvenile justice system has increased in recent years. This increase has prompted juvenile courts to develop more gender-specific services for young women entering the system to meet their unique needs as well as reduce recidivism. National statistics indicated that 21 percent of adjudicated females receive out-of-home placement as part of their treatment. The present study examined data from a mid-sized juvenile county court to determine the effects of the group home intervention for girls. The study compared group home girls to girls who did not receive group home treatment. Using two different propensity score matching (PSM) models, girls were compared on re-offense outcomes. The criminogenic risk level of girls was measured using the Youth Level of Service/Case Management Inventory (YLS/CMI) at intake and recidivism data were collected two years following release from the program. After employing the propensity score match models, mixed results were revealed. The first model revealed that the group home intervention significantly reduced recidivism rates for the treatment group two years following their program exit compared to similar girls who did not receive the intervention. This same effect was not found for the second propensity score model. The second model revealed that girls who received group home treatment did not significantly differ from the matched control in regard to recidivism within both one- and two-year follow-up periods. These findings have both substantive implications for gender-responsive programming and methodological implications for the use of propensity score matching for community-based program evaluation. Additional policy and

practice implications for gender-responsive services as well as future directions for research are discussed.

Copyright by
VALERIE ROSE ANDERSON
2015

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my dissertation committee. Dr. Bill Davidson, thank you so much for chairing all of my graduate school milestones, but especially my dissertation. You have been the best advisor and support over the last five years. I appreciate the many opportunities you have provided me during graduate school and I owe you a great deal of thanks for believing in my ideas. Many thanks to my committee members—Drs. Merry Morash, Deb Bybee, and Hi Fitzgerald—for helping me grow as a writer, researcher, and feminist scholar by providing insightful and valuable feedback throughout the dissertation process.

Special thanks to all of the past and present graduate and undergraduate student members of the Juvenile Risk Assessment Team and our community partners at the juvenile court. Without each of you this project would not have been possible.

My deepest gratitude goes to my very supportive family and friends, my wonderful partner Brendan and my adorable dog Cocoa Bean. Thank you all for your love and encouragement throughout this process.

This dissertation is dedicated in memory of the late Dr. Benjamin Hidalgo. Ben, thank you for introducing me to the field of community psychology at the University of Illinois Urbana-Champaign. I credit the beginning of this journey to you—thank you for everything you taught me, but most importantly, to always advocate for the disenfranchised and conduct research that promotes social justice. I miss you, my friend. I am eternally grateful for the inspiration.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	4
Purpose of the Juvenile Justice System	4
Gender-Responsive Policy Reform	5
Gender-Responsive Services	6
Theoretical Frameworks	6
Feminist Pathways Theory	8
Relational/Cultural Theory	9
Defining Gender-Responsive Treatment	11
Gender-Responsive Service Context	16
Gender-Responsive Service Content	16
Intervention Effectiveness by Gender	20
Gender-Responsive Intervention Effectiveness	22
Girls' Group Homes	26
Research Summary and Limitations	29
Current Study	32
CHAPTER 3: METHODS	34
Setting	34
Group Home Program Descriptions	34
Moderate Risk Group Home	34
High Risk Group Home	35
Gender-Responsive Elements of the Group Homes	36
Participants	39
Measures	42
Youth Level of Service/Case Management Inventory (YLS/CMI)	42
Additional Program Variables	42
Residential Placement	43
Intensive Probation Services (IPS)	44
Family Support Services	44
Court-Run Alternative School	45
Court-Run Evening Reporting Program	46
Recidivism	46
Constructing an Equivalent Time Metric: Model B	47
Propensity Score Matching	49
CHAPTER 4: RESULTS	51

Propensity Score Matching: Model A	51
Propensity Score Matching: Model B	56
CHAPTER 5: DISCUSSION	63
Summary and Discussion of Results	63
Model A Discussion	64
Model B Discussion	66
Time Frames.....	67
Over-Programming	68
Unobserved Variables	69
Methodological Implications of Propensity Score Matching	71
Limitations.....	72
PSM Model Specification and Sample Reduction	73
Defining Gender-Responsive Programming	74
Recommendations for Future Research	76
Conclusions	79
APPENDICES.....	81
Appendix A: Literature Review Table of Gender-Responsive Evaluation Studies	82
Appendix B: Youth Level of Service/Case Management Inventory (YLS/CMI).....	88
REFERENCES	89

LIST OF TABLES

Table 1. Key Elements of Gender-Responsive Programming.....	13
Table 2. Program Elements Matching Gender-Responsive Criteria	37
Table 3. Summary of Models	42
Table 4. Proportions of Girls Receiving Additional Court-Run Services	43
Table 5. Descriptive Statistics of Unmatched Samples of Girls: Model A	52
Table 6. Descriptive Statistics of Matched Samples of Girls: Model A	55
Table 7. One-Year Recidivism Rates: Model A.....	56
Table 8. Two-Year Recidivism Rates: Model A	56
Table 9. Descriptive Statistics of Unmatched Samples of Girls: Model B	57
Table 10. Descriptive Statistics of Matched and Unmatched Girls: Model B	60
Table 11. Descriptive Statistics of Matched Samples of Girls: Model B.....	61
Table 12. One-Year Recidivism Rates: Model B	62
Table 13. Two-Year Recidivism Rates: Model B	62
Table 14. Literature Review Table of Gender-Responsive Evaluation Studies	82

LIST OF FIGURES

Figure 1. Sampling Procedures Pre-Propensity Score Matching: Model A	40
Figure 2. Sampling Procedures Pre-Propensity Score Matching: Model B	41
Figure 3. Recidivism Time Frames: Model A	47
Figure 4. Constructing Equivalent Time Periods for Recidivism: Model B	49
Figure 5. Final Sampling Scheme Post-Propensity Score Matching: Model A	54
Figure 6. Final Sampling Scheme Post-Propensity Score Matching: Model B	59

CHAPTER 1: INTRODUCTION

In recent years, female juvenile offenders have comprised a growing proportion of juvenile court caseloads (Chesney-Lind & Shelden, 2004; Stevens, Morash, & Chesney-Lind, 2011). In particular, the greatest increase is for girls adjudicated for violent offenses (Puzzanchera Adams, & Sickmund, 2010). This increase in official female juvenile delinquency is largely seen as a reflection of the change in system-level policies and practices (Javdani, Sadeh, & Verona, 2011; Stevens et al., 2011) and changes in arrest patterns through the “upcriming” of girls’ offenses rather than an uptick in actual criminal behavior (Schwartz & Steffensmeier, 2012). Upcriming refers to the criminalization of minor or less serious charges (e.g., status offenses) (Schwartz & Steffensmeier, 2012). For example, the mandatory or pro-arrest policies for domestic violence cases appear to have a residual effect on girls being charged with violent offenses (Feld, 2009; Zahn et al., 2008). Strom and colleagues (2010) found that juveniles were more likely to be arrested than parents, particularly in mutually combative situations, and that these arrests policies had a stronger influence on the arrests of girls than boys. As a result of the increased proportion of juvenile cases involving girls, there is a growing interest and investment in gender-responsive services among juvenile justice practitioners and researchers (Chesney-Lind & Irwin, 2008; Chesney-Lind & Shelden, 2004). Further, there have been consistent calls for more rigorous evaluation studies on the effectiveness of gender-responsive programming (Chesney-Lind, Morash, & Stevens, 2008; Kerig & Schindler, 2013; Zahn et al., 2009).

The reauthorization of the Juvenile Justice and Delinquency Prevention Act in 1992 mandated states to include gender-responsive services (Chesney-Lind & Irwin, 2008). Much of the literature on female juvenile delinquency has focused on girls’ unique risk factors for entry

into the juvenile justice system (e.g., abuse/victimization, substance abuse, mental health needs, familial and peer contexts, etc.), girls' disparate treatment within the juvenile justice system, policies and practices that differentially impact girls in a negative way, and inadequate evaluation of programming for girls (e.g., Chesney-Lind, Morash, & Stevens, 2008). In particular, researchers, practitioners, and policymakers have drawn upon this literature when calling for more gender-responsive services for juvenile justice involved youth.

The literature on gender-responsive services broadly focuses on overarching themes and guidelines related to best practices with female youth (e.g., Ravoira, Graziano, & Patino-Lydia, 2012; Walker, Muno, & Sullivan-Colglazier, 2012). While a conceptual and theoretical basis suggests techniques to work effectively with female adolescents, there is limited existing research on the efficacy of the implied approach. Specifically there is a lack of evaluation of gender-responsive programming and the evaluation of programming disaggregated by gender (Kerig & Schindler, 2013; Matthews & Hubbard, 2008; Walker et al., 2012). Therefore, the purpose of this study is to evaluate the effectiveness of a community-based, gender-responsive intervention for female juvenile offenders. This study focused on re-offense outcomes for girls who received treatment in a gender-responsive group home placement versus girls in the same community who did not receive the group home treatment. This evaluation study adds to the growing—albeit limited—literature on gender-responsive interventions for juvenile justice involved girls.

The literature review covers two related areas regarding theory and previous research. First, the literature review provides an overview of the purpose and goals of the juvenile justice system. This section primarily focuses on the history of girls' involvement with the system and federal and state policy reform for the inclusion of gender-responsive treatment. Second, the

literature review discusses what is known to date about gender-responsive services for girls involved in the juvenile justice system. In particular, the second section reviews theoretical frameworks that inform the gender-responsive construct, how gender-responsive treatment is defined, the evidence for ineffectiveness of generic interventions for girls, and the effectiveness of gender-responsive interventions. The literature review concludes by discussing the limitations of and gaps within the current body of literature.

CHAPTER 2: LITERATURE REVIEW

Purpose of the Juvenile Justice System

The juvenile code of Michigan states that “each juvenile coming within the court's jurisdiction receives the care, guidance, and control, preferably in his or her own home, conducive to the juvenile's welfare and the best interest of the state. If a juvenile is removed from the control of his or her parents, the juvenile shall be placed in care as nearly as possible equivalent to the care that should have been given to the juvenile by his or her parents,” (Michigan Juvenile Code 712A.1-712A.32). The actions of the juvenile court and goals for intervention are guided by this treatment and rehabilitative framework rather than punishment or accountability. Thus, the purpose of the juvenile court is twofold: (1) to provide intervention-based treatment (rather than punishment), and (2) to separate youth from adult court. The system was originally designed to handle criminal misconduct among youth, but widened with the inclusion of status offenses—youth-specific noncriminal behavior (e.g., incorrigibility, running away, truancy, drinking alcohol) (Feld, 2009; Zimring, 2002).

Research on arrest and incarceration in the juvenile justice system has consistently shown that courts respond to girls primarily for these noncriminal status offenses and respond to boys mainly for criminal misconduct (Feld, 2009). Over the last three decades there has been an increasing number of girls coming into contact with the juvenile justice system (Puzzanchera, Adams, & Sickmund, 2010). Scholars largely see this as a reflection of policies and practices as opposed to an increase in actual delinquency among girls (Javdani, Sadeh, & Verona, 2011; Schwartz & Steffensmeier, 2012; Stevens et al., 2011). Finally, the juvenile court has historically been designed to handle and respond to the behavior of males. There is considerable evidence that the default approach in juvenile justice practice is built around the needs and experiences of

boys (see Chesney-Lind & Shelden, 2004; Kerig & Schindler, 2013). The observation that girls receive relatively harsh treatment for minor offenses has led scholars to call for a closer examination of girls' involvement with the juvenile justice system.

Gender-Responsive Policy Reform

Girls' historical involvement with the juvenile justice system can be seen as a residual impact of the child-saving movement on girls: the creation of a separate system for youth resulted in incarcerating more girls for largely "immoral behavior" (e.g., promiscuity, incorrigibility) (Chesney-Lind & Shelden, 2004). The reauthorization of the Juvenile Justice and Delinquency Prevention Act (JJDP) in 1992 specifically outlined the need for gender-responsive services and an examination of gender bias across levels of the system (Walker et al., 2012). The Office of Juvenile Justice and Delinquency Prevention (OJJDP) reissued these recommendations in 1998 in order to provide federal funding for separate, gender-responsive interventions (Kerig & Schindler, 2013). The amendment to the JJDP explicitly stated the need for physical and mental health services, education, and treatment for previous trauma/abuse. In addition, the American Bar Association and National Bar Association revealed gender bias across all levels of the system and called for gender equity in the juvenile justice system (2001).

Numerous task forces and committees were developed across states to implement gender-responsive services based on the principles put forth by the academic and governmental bodies of literature (Kerig & Schindler, 2013; Walker et al., 2012). Of particular note in gender-responsive reform was the development of the Girls' Study Group (GSG), a collaborative team of leading researchers on gender and crime, through the OJJDP in 2004. The GSG reviewed literature on girls' involvement with the juvenile justice system including risk and needs assessments, programs for girls, the causes and correlates of girls' involvement with the system,

and protective factors (Zahn et al., 2008). The GSG concluded that there was a systematic lack of attention to gender across all areas of juvenile justice research (Kerig & Schindler, 2013; Zahn et al., 2008). Additionally, based on analyses of a comprehensive database of juvenile justice program evaluations, Lipsey (2009) concluded that only 4 percent of programs served exclusively girls and 87 percent of program served all or mostly males. This is particularly problematic given that girls comprise roughly 30 percent of the juvenile court population (Puzzanchera, Adams, & Sickmund, 2010).

Gender-Responsive Services

The focus of gender-responsive policy and practice reforms in juvenile justice is due to the growing interest in, and literature on, female delinquency and the implementation of evidence-based practice in juvenile justice settings (Walker et al., 2012). The academic literature on gender-responsive services in the juvenile justice system has grown in recent years. From a gender-responsive perspective, girls' behaviors and the justice system's response are guided by girl-specific concerns (Matthews & Hubbard, 2008). Gender-responsive theories are especially important to expand upon for adolescent girls given that historically most theories of crime were developed based on the experiences and behaviors of boys (Chesney-Lind & Morash, 2011). In particular, research on juvenile delinquency and gender has illustrated the importance of gender-responsiveness in both theoretical and practice-based contexts to explain and respond to the unique needs of females (Mallicoat, 2007).

Theoretical Frameworks

Given that this study focuses on the outcomes of girls receiving gender-responsive treatment versus those who have not received gender-responsive treatment, this study will take an explicit feminist perspective. The influence of gender on the etiology of delinquency is well

supported in the criminological literature (see Belknap & Holsinger, 2006; Chesney-Lind, 2001; Daly & Chesney-Lind, 1988; Morash, 2010). In particular, feminist pathways theory and relational/cultural theory provide overarching frameworks and the rationale for the utility of gender-responsive services. Therefore, prior to reviewing the empirical evidence for the effectiveness of interventions for girls in the system, it is important to position the research within the context of relevant theoretical perspectives.

The feminist pathways perspective documents how a history of victimization and trauma is particularly salient in the childhood and adolescence of many women and girls involved in the justice system (Belknap et al., 1997; Wattanaporn & Holtfreter, 2014). Feminist criminological research has documented that many girls become involved in the juvenile justice system after running away from an abusive home and then committing “survival crimes” (Chesney-Lind, 2001). Research on girls and women’s pathways into crime has documented that girls have higher rates of substance abuse, victimization, depression, and anxiety than boys as well as social effects distinct from boys (e.g., family relationships, peers, social response to high-risk sexual behavior) (Belknap & Holsinger, 2006; Davis, 2007; Van Voorhis et al., 2010).

Relational/cultural theory emerged from the counseling psychological literature in which researchers and practitioners addressed the need to incorporate the experiences and voices of women into treatment (Jordan, 2013). In particular, relational/cultural theory posits that relationships—and life experiences within those relationships—are a central focus in the lives of girls and women and this knowledge must be integrated into interventions. This section will review and integrate two theoretical perspectives that inform the gender-responsivity framework and the need for gender-responsive programming in the juvenile justice system.

Feminist Pathways Theory

Daly (1992) introduced the feminist pathways framework by highlighting the differences between males and females as well as differences among females in regards to their pathways into crime. In particular, pathways theory provides a holistic, contextualized perspective that suggests that female crime is embedded within the broader conditions of their lives (Bloom et al., 2003; Van Voorhis et al., 2010). This includes violence against women and girls, intersectionality, and specific life events that act differentially by placing females at a greater risk for offending (e.g., frequency and severity of childhood abuse), and “multiple marginality” based on familial, educational, and work-related circumstances (Bloom et al., 2003; Owen & Bloom, 1998). Theories of violence are not necessarily equivalent across genders (e.g., between boys and girls) since female violence is more often rooted in female victimization (Chesney-Lind & Irwin, 2008). Thus, feminist pathways theory has contributed to the research on female offending and its connection to victimization. Wattanaporn and Holtfreter (2014) reviewed the salient contributions of the pathways framework to understanding female victimization, offending and its advancement of theory, policy, and practice. They concluded that the female pathways to crime perspective was an empirically rigorous and theoretically informed framework to study patterns of women’s and girls’ victimization and offending.

Pathways research has also documented girl-specific risks that are not usually seen with males such as dating significantly older partners and self-injury (Salisbury, Van Voorhis, & Spiropoulos, 2009). Childhood events (e.g., trauma, victimization) are distinct, gendered risk factors that create differential pathways for girls into the justice system (Foley, 2008). In particular, history of victimization, dysfunctional families, school troubles, mental health needs, substance abuse, and repeated status offenses (e.g., running away, truancy) are the primary

underlying issues that bring girls into the juvenile justice system (Bloom, Owen, & Covington, 2004; Foley, 2008). The strong link between history of victimization and system involvement informs the nature and extent of girls' delinquent behavior (Mallicoat, 2007; Owen & Bloom, 1998). Issues that differentially impact girls include physical abuse and sexual assault victimization that may influence them to run away from abusive homes and subsequent abuse within the juvenile justice system (Acoca, 1998; Goodkind, 2005; Hoyt & Scherer, 1998). Polyvictimization, or multiple forms of abuse, including caregiver violence, witnessing violence, gang/group attacks, dating violence, and sexual violence, is pervasive in the lives of girls in the juvenile justice system (Ford et al., 2010). Schaffner (2007) noted that trauma in the lives of girls and boys is qualitatively different, because for girls, it reaffirms their lower status in the gender hierarchy, while it contradicts boys' higher status.

The findings grounded in feminist pathways theory have brought about change in the way the justice system responds to females in the assessment process and programming decisions (Bloom et al., 2003; Bloom, Owen, & Covington, 2004; Wattanaporn & Holtfreter, 2014). As well, research using the pathways theoretical framework has substantially informed treatment approaches and programming in the justice system (Wattanaporn & Holtfreter, 2014).

Relational/Cultural Theory

Relational/cultural theory (RCT) emerged from dialogue about counseling and therapeutic relationships in which feminist scholars and practitioners addressed the need to integrate the voices and experiences of women into services and treatment (Jordan, 2013). Miller (1976) identified the centrality of relationships in the lives of women (and other historically marginalized groups), and noted that this centrality has been excluded from the discussion in traditional theories of human development. RCT assumes that isolation, oppression, and

exclusion are all relational traumas that women and girls routinely experience (Comstock et al., 2008). In turn, the lack of understanding of relationship experiences and context can lead to misunderstanding and marginalization. RCT focuses on the development of “growth fostering” relationships and explicitly addresses power dynamics in relationships and culture (Comstock et al., 2008; Jordan, 2013).

Developmental and relational theories focus on female development and emphasize the centrality of relationships in the lives of girls and the importance of a sense of connection to others (Bloom et al., 2003; Bloom et al., 2004). A fundamental principle of this theory rests on the notion that individuals develop within the context of relationships that are embedded in and influenced by broader sociocultural contexts (Jordan & Hartling, 2002). In turn, the psychopathology of female juvenile delinquents is related to disconnections from important interpersonal relationships (e.g., families, peer networks, acquaintances). In order to address these behaviors and needs, it is critical to focus on the meaningful relationships in the lives of girls (Covington, 2008; Foley, 2008). The RCT perspective also emphasizes restoring, healing, and building healthy relationships (Daly & Stubbs, 2006).

The concept of gender-responsivity is rooted in Gilligan’s (1982) relational approach to adolescent female development: the identity of females is shaped through relationships. While girls and boys may experience many of the same life stressors and risk factors for delinquency, the way in which these stressors impact girls’ juvenile justice trajectories differs from those of boys (Garcia & Lane, 2012). This is especially true in the context of girls’ socialization and emphasis on close relationships (Belknap & Holsinger, 2006). For example, Garcia and Lane examined data from focus groups (n = 27) with delinquent girls and found that nearly all focus group participants experienced some form of relationship strain (e.g., boyfriends, abuse/neglect,

fighting with other females, fighting with their mothers, etc.), which acted as a pathway to girls' involvement in the juvenile justice system.

Traumatic relationships, experiences such as physical and sexual abuse, relational aggression, domestic violence, and community violence are also more prevalent among female than male offenders (Covington, 2000; Covington et al., 2008; Garcia & Lane, 2012; Messina et al., 2010; Van Voorhis, 2012). These experiences of abuse and neglect negatively impact girls' emotional and behavioral well-being (Holsinger, 2000). Girls are more likely to have poor relationships with their parents and more likely than boys to report abandonment by a parent (Belknap & Holsinger, 2006; Mallicoat, 2007). While boys and girls in the juvenile justice system both experience trauma and relationship strain, the ways in which they manifest and are processed differs by gender (Garcia & Lane, 2012).

In sum, the rationale for gender-responsive practice is informed by both feminist pathways theory and relational/cultural theory. Feminist pathways theory explains how many girls become involved with the system due to trauma or a history of victimization (Belknap et al., 1997; Belknap, 2001). Relational/cultural theory emphasizes that there is a need to address the centrality of relationships and girls' life experiences within the context of those relationships in services (Jordan, 2013). Both of these theoretical perspectives provide a rationale for the need for an alternative treatment modality for girls in the juvenile justice system.

Defining Gender-Responsive Treatment

Consistent with feminist pathways theory and relational/cultural theory, gender-responsivity¹ can be seen as a paradigm for addressing girls' unique needs in the juvenile justice

¹*Important distinction in nomenclature:* The literature primarily discusses this topic by referring to services as “gender-responsive” or “gender-specific”—often times interchangeably. This review and study explicitly examines **gender-responsive** services as they differ from solely gender-specific

system. Gender-responsivity as an innovation seems to be congruent with the juvenile court's mission to rehabilitate youth. Gender-responsivity refers to a comprehensive systems response to female delinquency that emphasizes the importance of girls' experiences as well as addresses girls' unique psychological, developmental, social needs, and pathways into crime (Garcia & Lane, 2010, Green, Peters, & Associates, 1997). In essence, programming that is effective for girls involved in the juvenile justice system should be specifically designed and targeted for their unique, gender-based needs.

However, while this construct has strong conceptual and theoretical foundations in the feminist criminological literature (see Belknap, Holsinger, & Dunn, 1997; Bloom, Owen, & Covington, 2004), there is weaker evidence for, and operationalization of, gender-responsiveness or gender-specific programming in the empirical juvenile justice literature (Matthews & Hubbard, 2008). Kerig and Schindler (2013) outlined ten academic and practitioner-based publications that addressed the recommendations set forth by the OJJDP for gender-responsive intervention for girls involved in the juvenile justice system. In their review, Kerig and Schindler noted the importance of integrated treatment across levels of the system, intervention that addresses trauma, the promotion of resilience, and programming that has a relational and strengths-based focus. The review of these ten studies led the researchers to conclude that these were the critical components of effective intervention for juvenile justice involved girls. Table 1 provides an overview of key elements identified within all of these studies that constitute gender-responsive treatment. To date, the literature attempts to define the key elements and characteristics of gender-responsive services that focus on two broad areas of concern:

programming. That is, a program may be gender-specific (e.g., only targeted only targeted for girls in the juvenile justice system), but not be gender-responsive in the focus, scope, and/or content of the services.

(1) service context and (2) service content. One issue with this literature is that it consistently points out the broad areas that need to be addressed in the lives of girls, but not necessarily how those factors are to be addressed in programming. The table is organized by content and context related factors, which will be expanded upon in the following sections.

Table 1. Key Elements of Gender-Responsive Programming

Content-Related Factors	
Dangerous Neighborhoods	Addresses violence in the community, high levels of gang involvement, and unsafe schools (e.g., Laidler & Hunt, 2001; Leitz, 2003)
Delinquency History	Addresses reasons why girls get involved with the system and the extent to which girls' delinquency is less "serious" or less "chronic" than boys' involvement (e.g., Snyder & Sickmund, 1999)
Familial Relationships	Recognizes risk factors and needs related to family relationships including parental incarceration, death, homelessness, discord in the mother/daughter relationship, hostility/conflict/aggression, experiencing or witnessing violence/abuse in the home, lack of support/structure, abandonment or absence of the parents and/or involvement in the child welfare system. Treatment should include promoting healthy relationships with their families (e.g., trust, interdependence, especially with female family members), involving the family in treatment, and focus on positive relationship-building skills (e.g., Bloom et al. 2002; Cauffman et al., 2008; Kerig & Schindler, 2013; Ravora et al., 2012; Schaffner, 2007; Valentine Foundation, 1990)
Mental Health	Addresses girls' cognitive vulnerabilities such self-blame, rumination, excessive concern for others' opinions and strengths such as empathy and social intelligence, the presence of mental health issues (e.g., PTSD, depression, anxiety, suicide attempts) among girls, low self-esteem, and the extent to which there are higher rates of internalizing mental health problems among girls and less externalizing problems than boys (e.g., Watson & Edelson, 2012; Ravora et al., 2012; Matthews & Hubbard, 2008; Kerig & Schindler, 2013; Cauffman et al., 2008; Pasko, 2006; Teplin et al., 2002)
Peer Relationships	Addresses issues that arise in peer relationships and adversarial interpersonal relationships, specifically with other girls and teaches positive relationship-building skills by developing relationships of trust and interdependence with female friends (e.g., Cauffman et al., 2008; Kerig & Schindler, 2013; Valentine Foundation, 1990)
Physical Health	Addresses sexually transmitted infections, traumatic head injury or other injuries related to violence girls experience, general health education, and education around contraception, diseases and prevention, and healthy sexuality (e.g., Acoca & Dedel, 1998)

Table 1 (cont'd).

Pregnancy/Parenting	Recognizes girls' risk of pregnancy complications, socioeconomic disadvantage, relationship discord, and unresponsive parenting (e.g., Watson & Edelson, 2012; Raviora et al., 2012; Valentine Foundation, 1990; Cauffman et al., 2008)
Romantic Relationships	Promotes healthy relationships with significant others by teaching positive relationship-building skills and addresses the extent to which girls are involved with much older men, "bad boyfriends," and dating violence (e.g., Bloom et al., 2003; Kerig & Schindler, 2013; Cauffman et al., 2008; Schaffner 2006)
School Issues	Addresses missing school or truancy, doing poorly in school or academic failure, getting in trouble with school personnel, and need for special education services (e.g., Pasko, 2006)
Sexuality and Risky Sexual Behavior	Recognizes the juvenile justice system's sensitivity towards girls' sexuality (e.g., "moral crimes"/ "wayward girls") and benevolent patriarchy in the JJS (e.g., court as protecting girls from own sexuality). Addressing early puberty, prostitution/sex trafficking, risky sexual behavior (e.g., failure to use birth control, failure to use protection, sex with multiple partners), sexual minority status, and the "double standard" (e.g., Raviora et al., 2012; Kerig & Schindler, 2013)
Substance Abuse	Addresses girls' addiction, experience of withdrawal symptoms, substance use as a response to trauma (e.g., self-medicating through drug use) and the inclusion of services to address girls' substance abuse (e.g., Bloom et al. 2003; Kerig & Schindler, 2013; Acoca & Dedel, 1998; Gaarder et al., 2004; Pasko, 2006)
Trauma	Recognizes that girls' pathways may be different than boys, characterized by high rates of abuse and trauma [e.g., family violence, child abuse, polyvictimization (cumulative effects of multiple and chronic traumas), interpersonal traumas (e.g., abandonment by caregivers, multiple changes in caregivers), physical abuse, sexual abuse, observing violence]. Programming should address the after-effects of trauma, promote healing from trauma, recognize the dangers and risks that girls face because of gender, and juvenile justice staff should be trained in trauma-informed care (e.g., Acoca & Dedel, 1998; Bloom et al., 2004; Pasko, 2006; Raviora et al., 2012; Watson & Edelson, 2012)
Context-Related Factors	
Communication	Includes training of court workers for communicating and working with girls and adequate time for girls to talk and have nurturing conversations within ongoing relationships (e.g., Raviora et al., 2012; Valentine Foundation, 1990)
Community-Based	Establishes sustainable, community-based supervision and collaborative services for reentry (e.g., involvement with schools, community mental health, etc.) (e.g., Bloom et al., 2003; Valentine Foundation, 1990)

Table 1 (cont'd).

Comprehensive	Integrates family and community systems, utilizes a continuum-of-care, and provides collaborative services for community reentry. Comprehensive services deal with behavior in context by enabling girls to focus on their individual needs and how their risk factors may have impacted development (e.g., Bloom et al. 2003; Walker et al., 2012)
Culturally Responsive	Addresses risk and needs in context – context of culture and identity (gender, race, ethnicity, religion, socioeconomics, ability, sexual orientation) and recognizes heterogeneity among girls (e.g., class, sexual orientation, race/ethnicity, gender expression) in the provision of culturally-relevant services (e.g., Hubbard & Matthews, 2008; Bloom et al., 2003; Walker et al., 2012)
Gender-Informed Protocols	Includes assessments (e.g., risk, needs, assets), interviewing, strip searches, and program curriculum that are gender-informed (e.g., Ravora et al. 2012; Matthews & Hubbard, 2008)
Gender Matters	Acknowledges that gender matters and makes a difference with experiences in the juvenile justice system and celebrates and honors female development and girls' experiences (e.g., Bloom et al., 2003; Girls Inc., 1996)
Relationships	Recognizes that female development hinges on positive, mutual relationships and services promote healthy connections with others, specifically “prosocial females” by fostering supportive relationships with females, family, peers. Specifically programs that are relationship-based including someone to listen to them and female mentors, especially those who share experiences and who “exemplify survival and growth” (e.g., Bloom et al. 2003; Ravora et al. 2012; Valentine Foundation, 1990; Matthews & Hubbard, 2008; Walker et al., 2012)
Resources for Girls	Provides girls with opportunities to improve their socioeconomic status and recognize dangers and risks that girls face because of gender (e.g., sexism, racism, poverty). Provides adequate resources to ensure comprehensive program sustainability (e.g., Bloom et al., 2003; Valentine Foundation, 1990)
Safety	Addresses the development of trust, awareness of power-differentials, trauma-informed care (e.g., the promotion of healing from trauma) by creating physically and emotional safe spaces separate from males (e.g., Walker et al., 2012; Hubbard & Matthews, 2008; Valentine Foundation, 1990)
Strengths-Based	Focuses on girls' cultural strengths, girls' leadership, and the development of girls' strengths/assets, and include strengths-based assessments (e.g., Matthews & Hubbard, 2008; Valentine Foundation, 1990)
Voice	Integrates girls' voice into services – program design, implementation, and evaluation and values the female perspective (e.g., Ravora et al., 2012; Valentine Foundation, 1990; Girls Inc., 1996)

Gender-Responsive Service Context

Service context refers to the environmental aspects of programming whereas service content refers to the specific program elements (Bloom et al., 2003; Morgan & Patton, 2002). Gender-responsive service context includes a strengths-based orientation focusing on positive self-esteem, skill building, and self-efficacy (Bloom et al., 2003; Garcia & Lane, 2010; Green et al., 1997). Additionally, the program site or location should be a structured, safe, nurturing environment (Bloom et al., 2003; Smith & Smith, 2005; Valentine Foundation, 1990). Some research has indicated gender-responsive intervention should incorporate community-based treatment (Bond-Maupin et al., 2002). Service context is also inclusive of staff-related characteristics such as hiring professional, caring females with a similar background to the girls with whom they work (Bloom et al., 2003; Garcia & Lane, 2010; Green et al., 1998; Hubbard & Matthews, 2008; Ravoira et al., 2012, Walker et al., 2012). Service context focuses on using therapeutic intervention by matching staff and youth based on reciprocal interests and congruent personality characteristics (Bloom et al., 2003; Zahn et al., 2009) and utilizes cognitive-behavioral treatment strategies (Hubbard & Matthews, 2008).

Gender-Responsive Service Content

Service content includes a wide variety of topical areas that consider gender-responsive needs (Bloom et al., 2003; Morgan & Patton, 2002). The majority of the gender-responsive literature discusses services focusing on addressing the trauma and victimization histories of girls. In particular, researchers have used gendered theories (e.g., pathways) to understand girls' increased risk for offending due to traumatic events (Belknap et al., 2010). Gender-responsive treatment should address violence, sexual abuse, and other childhood traumas such as loss of a parent or experiencing other forms of abuse (Baines & Alder, 1996; Belknap et al., 2010; Bloom

et al., 2003; Bond-Maupin et al., 2002; Chesney-Lind et al., 2008; Covington, 2000; Gaarder et al., 2004; Garcia & Lane, 2010; Smith & Smith, 2005; Valentine Foundation, 1990).

Gender-responsive services also include focusing on building, restoring, and/or sustaining healthy relationships (e.g., romantic, peer, family) by fostering an atmosphere of respect, caring, one-on-one counseling, and active listening (Belknap et al., 1997; Bloom et al., 2003; Brubaker & Fox, 2010; Chesney-Lind et al., 2008; Chesney-Lind & Shelden, 2004; Gaarder et al., 2004; Garcia & Lane, 2010; Green et al., 1997; Hubbard & Matthews, 2008; Valentine Foundation, 1990). Gender-responsive services should also address the mental, physical, and sexual health related needs of girls, substance abuse treatment, and access to educational, recreational, and vocational resources (Bloom et al., 2003; Brubaker & Fox, 2010; Chesney-Lind et al., 2008; Covington, 2000; Gaarder et al., 2004; Garcia & Lane, 2010; Valentine Foundation, 1990).

Similarly, communication is a critical component of gender-responsivity (Baines & Alder, 1996). Research has indicated that girls respond best to court programming when trusting relationships are established between the girl and the adults with whom she is working (Baines & Alder, 1996). In order for a probation officer to effectively communicate with girls, he or she must first build a positive connection with the youth (Gaarder et al., 2004). One study on adult offenders noted gender differences in interpersonal skills and communication styles based on survey data from a sample of correctional administrators (Schram, Koons-Witt, & Morash, 2004). In particular, Schram and colleagues discovered that women were more talkative and verbal than men, correctional staff needed strong listening skills, women tended to ask more “why” questions, were more focused on problem-solving, and identified that an essential component to the needs of females is tied to their emotional need to speak and to be heard.

However, there is little research on communication patterns and needs among girls in the juvenile justice system. Advocates of gender-responsive reforms have noted that a core component to services is the inclusion of girls' voice in the treatment process (Raviera et al., 2012). Finally, important sociocultural factors such as poverty, race, language, and cultural background need to be addressed in gender-responsive programming (Bloom et al., 2003; Bond-Maupin et al., 2002; Brubaker & Fox, 2010; Chesney-Lind et al., 2008; Gaarder et al., 2004; Girls Inc., 1996; Hubbard & Matthews, 2008).

A final crucial factor from the literature on gender-responsive services is the role of integrative treatment, wraparound services, and continuums of care (Walker et al., 2012). Most girls in the juvenile justice system are also involved in, or in need of, other social services such as the mental health system, special education, and/or child protective services (American Bar Association & National Bar Association, 2001). Therefore developing linkages and fostering connection among the various social services intertwined in the lives of these girls is of critical importance for practitioners addressing this social problem.

Previous research has documented the lack of collaboration among key community stakeholders in the juvenile justice and child welfare systems (Bloom et al., 2003). In particular, Bloom and colleagues found that many probation staff felt that the juvenile justice system is often used in lieu of mental health services as well as in response to failure of the school system. Since there is no clear operationalization of what comprises gender-responsive services, developing a working definition with key terms/guiding principles may help guide practitioners attempting to provide a continuum-of-care. For example, the American Bar Association and National Bar Association (2001) found that developing and accessing gender-responsive services for girls in the juvenile justice system required a more concentrated effort at integrating social

services, building a collaborative approach between levels of government, and developing more comprehensive advocacy strategies for girls across these systems. In the context of building a collaborative infrastructure for juvenile justice involved girls, programmatic capacity would develop clearly focused objectives and goals and culturally competency across assessments and services. Programming would also identify intermediate goals or proximal outcomes of their initiative and work towards fulfilling the unmet needs of adolescent girls (e.g., identify gaps in service provision, linkages that can be strengthened among organizations, and/or uncovering indigenous services or interventions in the community already addressing gender-responsive areas of need)—all of which have been identified as critical elements in models of gender-responsivity (e.g., Ravora et al., 2012; Walker et al., 2012).

In sum, the literature on gender-responsive services focuses on the importance of comprehensive, wraparound services, use of a continuum-of-care, formal and informal service provision through collaboration, and the development of a therapeutic alliance between court officials and youth (Bloom et al., 2003; Brubaker & Fox, 2010; Matthews & Hubbard, 2008). One major issue that has been noted in the implementation of gender-responsive reforms is that the “critical elements” to achieve gender-responsiveness may not necessarily be gender-responsive (e.g., utility of a continuum-of-care), making it difficult to discern general best practices from gender-responsive practices (Kerig & Schindler, 2013; Matthews & Hubbard, 2008). Further research needs to identify best practices and gender-responsive practices separately by identifying those that are substantively different or if they are characteristics of quality juvenile justice practices regardless of gendered needs. For example, it should be noted that many of these key characteristics of services may be quality-driven, that is, programs that implement these strategies (e.g., interacting with empathic juvenile court staff) are more

effective than those that do not. The current state of the literature provides some preliminary evidence that (1) there is gender invariance in general juvenile court intervention, and (2) gender-responsive services are worthwhile and effective over and above generic court intervention. The extent to which this has been empirically tested and the quality of that evidence will be reviewed in the following two sections.

Intervention Effectiveness by Gender

Strong evidence for the effectiveness of “gender-neutral” interventions for both boys and girls is not easily discerned. Zahn and colleagues (2009) stated that analysis by gender is not consistently reported in studies evaluating juvenile justice programming. For example, Zahn and colleagues examined 392 programs from the Blueprints for Violence Prevention database and only 29 included analyses by gender, and only 6 of those 29 programs that included gender-based analyses used a sample with juvenile justice involved youth. In their review, they found that five of the six programs did not detect any gender differences in program outcome, while the other program detected a positive effect only for girls. In another review, Bell and colleagues (2012) reviewed 106 social interventions for youth that provided outcome data by gender, only 11 of which measured juvenile delinquency outcomes (e.g., self-report or official recidivism). Of those 11 programs, only one, Job Corps, had a positive impact on girls’ delinquency (reduction in re-offending).

Fagan and Lindsey (2014) reviewed experimental studies on community-based preventive interventions for girls’ and boys’ delinquency, substance use, and violence in adolescence. Fagan and Lindsey examined the Blueprints for Healthy Youth Development Website and found that 7 of 14 evaluation studies had outcomes that varied by gender (e.g., the program showing effectiveness more for males than females and vice versa)—about half of those

were more effective for males ($n = 4$) and the others were more effective for females ($n = 3$), but the authors found no discernable patterns in terms of the program setting, dosage, content, age of targeted population, etc. Outside of the Blueprints evaluations, Fagan and Lindsey identified 15 studies with gender-based analyses in which 7 of 15 showed stronger effects for females, 6 of 15 showed stronger effects for males, and 2 of 15 showed mixed effectiveness. Similar to their analysis of the Blueprints programs, there were no clear patterns in the components of what made programs more effective for males versus females. Fagan and Lindsey (2014) advocated for more experimental evaluation studies disaggregating program effectiveness by gender.

The juvenile justice system has long relied on intervention programs to reduce recidivism among juvenile offenders. However, there is a broad range of such programs, and although many have shown effectiveness, they are not necessarily equally successful (Hoge, 2001). Most programs utilized by the juvenile justice system offer specific services that a particular juvenile court might find suitable for a youth based on an assessment of their needs (Lipsey, 2009). In addition to these, there are also identified model programs such as Functional Family Therapy (FFT), Multisystemic Therapy (MST), cognitive-behavioral programs, and family counseling programs (Lipsey, 2009; Lipsey et al., 2010).

According to findings that have emerged from the reviews and meta-analyses conducted by Lipsey and colleagues, effective intervention programs tend to serve high-risk juvenile offenders, uphold a therapeutic rather than a control philosophy, and provide adequate quantity and quality in the service delivery (Lipsey, 2009; Lipsey et al., 2010; Lipsey & Howell, 2012). It is important to note that—within the “what works” paradigm—neither demographic characteristics of the youth, such as age, gender, or race/ethnicity, nor juvenile justice supervision appear to be directly related to program effectiveness (Lipsey et al., 2010; Lipsey & Howell, 2012).

Finally, it is important to note that many interventions that court systems implement are not evidence-based and/or there were no outcome studies conducted to examine program effectiveness (Bloom et al., 2002; Zahn et al., 2009). Thus, based on the current evidence, it is not possible to conclude that gender-neutral programs are ineffective or produce iatrogenic effects for girls. In sum, there has been a remarkable lack of consideration of gender differences in research on interventions for delinquent youth, including well-validated interventions (Kerig & Schindler, 2013); there are a number of effective interventions aimed at reducing recidivism. Specifically, various types of group homes have also shown effectiveness in reducing recidivism among juvenile offenders (Leeman et al., 1993; Lewis, 2005; Lazelere et al., 2004; Jones & Timbers, 2003; Kirigan et al., 1982; Rivard, 2005; McCurdy & McIntyre, 2004; Fields et al., 2006; Hooper et al., 2000). Therefore, it may be advantageous for juvenile justice systems to continue to utilize community-based intervention programs to not only reduce recidivism, but also to successfully re-integrate youth into the community. The models currently utilized have not properly examined the efficacy of juvenile delinquency prevention and intervention efforts with females nor have they based their intervention models on the specific needs of females.

Gender-Responsive Intervention Effectiveness

Even though the call for gender-responsive intervention dates back to the 1992 reauthorization of the JJDP, few interventions for juvenile justice involved girls have been rigorously evaluated (Chesney-Lind et al., 2008; Zahn et al., 2009). Most interventions only target one specific issue (e.g., teen pregnancy) and very few are comprehensive in addressing the interconnected factors associated with girls' system involvement (Kerig & Schindler, 2013). Chesney-Lind, Morash, and Stevens' (2008) review paper noted that currently all programs fall

short of addressing the multitude of girls' needs, but supported the need for a gender-based approach.

Chesney-Lind and colleagues (2008) reviewed nine formally evaluated gender-responsive programs that met criteria to prevent or respond to female delinquency (e.g., address history of abuse, health education, job training, housing assistance, etc.). These researchers operationalized gender-responsive needs and services as comprehensive and adaptable programming targeting younger girls (ages 9 to 14) by focusing on adolescent development, safety needs, abuses, relationships with older men, exploitation and prostitution, economic needs, and substance abuse treatment (Chesney-Lind et al., 2008). Chesney-Lind and colleagues found that few programs specifically designed for girls existed. They further suggested that the current state of the gender-responsive evaluation literature does not provide any insight into particular effective model programs, and encouraged the development of new gender-responsive services for girls.

Across the nine studies they reviewed, Chesney-Lind and colleagues (2008) described areas of need among girls that were addressed in the program: having someone to talk to, working on improving relationships with peers, education and support related to sex, sexuality, pregnancy, parenting, and romantic relationships, empowerment (via linking girls and their families with much needed resources), and programs that addressed multiple needs and utilized comprehensive services. However, the researchers highlighted the failure of the review's evaluation results to shed any light on effective gender-responsive programming. Based on the formal evaluation literature, there is limited evidence for the efficacy of services that address the multiple needs of girls that the feminist criminological and adolescent development literatures outlined.

In another review, Zahn and colleagues (2009) examined nine gender-responsive programs showing mixed support for their effectiveness in reducing recidivism. They found an overall pattern of improvement of relationships and school success, but less evidence for long-term effects on reduction in recidivism (more than six months). They did point to generally positive effects, and provided support for examining additional outcomes beyond reducing overall recidivism. These evaluation reviews of gender-responsive programming found that comprehensive programs that target multiple risk factors are the most effective in reducing delinquency regardless of gender (Chesney-Lind et al., 2008; Zahn et al., 2009). Similar to Chesney-Lind and colleagues' (2008) review findings, very few gender-responsive programs have been rigorously evaluated in this context. For example, Zahn and colleagues noted that only two gender-responsive programs had a rigorous evaluation design such as experimental randomization and use of a control group. One of those two programs, Working to Insure and Nurture Girls Success (WINGS), was an alternative probation service in which a case manager developed a comprehensive plan focusing on mother-daughter mediation, anger management, education, substance abuse, transportation issues, and career support. They found lower recidivism rates among the treatment group six months after the intervention compared to the control group, but not at 12-month and 18-month follow-ups. As well, while on probation girls in the program attended school twice as often as the girls in the control group, but this effect did not continue at follow-up (Chesney-Lind et al., 2008; Zahn et al., 2009). Finally, the evaluators found that girls demonstrated fewer risk factors and an increase in protective factors; however, the researchers did not provide follow-up data on the risk and protective factor outcomes.

The other experimentally evaluated program that was reviewed was Reaffirming Young Sisters' Excellence (RYSE). This program was characterized by home visits by bilingual

probation officers who provided (and linked girls to) services such as life skills courses, pregnancy and parenting services, therapy, and culturally tailored extracurricular activities (e.g., trips to local African-American plays and leadership conferences). The evaluation found that girls in RYSE were 50% more likely to complete probation than girls receiving standard probation services, had lower recidivism rates (18%) than the control group (39%), and had less severe re-offenses than the control group. Recidivism was measured after the services were completed, but the authors did not specify how that date was determined for the comparison group receiving traditional probation services. Finally, race/ethnicity moderated these findings in that African American and Hispanic girls showed lower recidivism rates than white and Asian girls (Zahn et al., 2009).

Appendix A includes an overview of 16 gender-responsive evaluation studies, two of which were the experimental studies described previously. This table provides a review of the programs evaluated in each of the studies including the research methods and outcomes and the specific core gender-responsive elements of the program. All non-experimental studies on gender-responsive programs reported some level of benefit to girls who received the treatment across a variety of outcomes (e.g., decrease in law violations, drug/alcohol usage, improving peer relationships, academic achievements, self-esteem, etc.). The programs varied in terms of scope and focus, but all of them identified some factors addressed as gender-responsive. For example, Davidson and colleagues (2011) evaluated Honolulu's Girls' Court and found that the core gender-responsive elements of intensive case management focused on building healthy relationships and individualized therapy to address trauma significantly decreased girls' law violations (e.g., new petitions/charges) and days spent on the run during the 1-5 year follow-up

periods for the five cohorts of girls who received the program versus the matched comparison group.

Girls' Group Homes

The most essential elements that need to be included in gender-responsive programming are core components of comprehensiveness, safety, empowerment, family and relationship support in the context of community-based services (Watson & Edelman, 2012). The juvenile justice system offers a variety of programming including foster care, detention centers, probation services, educational classes, community based diversion projects, residential treatment and other ranges of these types of services. While these services exist for female offenders, very little research on outcome measures based on gender differences have been conducted. As reviewed previously, Bloom and colleagues (2002) found that the two main aspects needed for program effectiveness and development were content- and context-related factors—meaning that programming for female offenders needs to not only be responsive to the multiple issues that impact their lives but also occur in a setting that is favorable to a therapeutic change process.

Group homes are out-of-home residential facilities in which juveniles receive long-term multidimensional treatment to better address problematic behavior and the circumstances that led to this behavior (Development Services Group, Inc., 2010). Group homes are also community-based allowing for continued integration with society in such ways as going to their main school, employment, and seeing social supports such as family members (Development Services Group, Inc., 2010). Group homes are small residential facilities that house fewer offenders in a community-based setting managed and secured by trained staff members (Ryan, Marshall, Herz, & Hernandez, 2008). A group home attempts to embody a safe and controlled setting yet is less restrictive than an in-patient or lock-down residential facility (Ryan et al., 2008). Girls' group

homes are often able to offer services that address the histories, backgrounds, and life experiences of female youth, which have been found to lead to such specific issues such as trauma, mental health, substance abuse, criminogenic behavior, and economic troubles (Bloom et al., 2002).

Much of the literature has focused on the evaluation of secure, residential placements for adolescent offenders. For example, the Harriet Tubman Residential Center in New York and the Touchstone Program in Connecticut are two types of residential facilities for girls. The Harriet Tubman Residential Center is a residential facility for female youth between the ages of 15 to 18 that were either first time offenders or committed minor crimes (Greene et al., 1997). The Harriet Tubman program aimed to create an environment that would help girls return home as productive members of society by administering educational and therapeutic services (Greene et al., 1997). The Touchstone program is a residential treatment program for female youth involved in the juvenile justice system, between the ages of 12-17 years of age (CT Touchstone Program, 2001).

Gender-responsive programming should address needs such as self-esteem, life skills, parenting classes, education, and substance abuse by providing an advocate or a counselor in which a one-on-one relationship can be achieved (CT Touchstone Program, 2001). While gender-responsive programming and group homes are growing in abundance, program success is difficult to evaluate for gender-responsive treatment because there has yet to be an established standard of assessing effectiveness due to the infancy of this area of research. Hubbard and Matthews (2008) noted that there are two different spheres of research on the subject that lead to different assessments of program goals and effectiveness. In one body of literature, “what works,” scholars and academics tend to look at effectiveness based on reduction in recidivism. In

the gender-responsive literature effectiveness is reviewed and evaluated on the basis of positive lifestyle outcomes in addition to recidivism (Hubbard & Matthews, 2008).

Chesney-Lind and colleagues (2008) compiled a review study on the evaluations of program effectiveness on multiple contemporary treatment models for female offenders. They found that some programs were reviewed on means of effectiveness in terms of lifestyle changes such as the evaluation of the Girls Incorporated program, based on levels of substance use and the evaluation of the Working to Insure and Nurture Girls' Success program, which included such strengths as improvement in school, involvement in community projects and organizations, and management of levels of self-control in difficult peer pressure situations (Chesney-Lind et al., 2008). A review of the data for female juvenile offenders in residential care facilities located in the state of New York showed that in the year 2008, 29% of the female juveniles reoffended 12 months after release from the program and 43% reoffended 24 months after program release (Recidivism Among Juvenile Delinquents and Offenders Released from Residential Care in 2011). Another study completed by the Department of Juvenile Services revealed that of the 427 female youth released from residential facility programming, only 18% of girls were reconvicted within the court system after being released for one-year while 58.1% were rearrested (Department of Juvenile Services Recidivism Report, 2012). Finally, a recent study that utilized propensity score matching in a quasi-experimental design, found positive effects of gender-responsive services for girls placed in residential treatment. Day, Zahn, and Tichavsky (2014) examined male and female youth in a detention facility and found that girls with gender-specific risk factors fared better in programs inclusive of gender-responsive elements (e.g., relationship-based, empowering, focus on mental health and self-esteem, trauma-informed, etc.) than programs that did not incorporate those factors. However, this was not the case for male youth in

treatment—these programs did not show any effect (positive or negative) for male youth with similar risk factors (Day et al., 2014). These findings suggest that girls are in need of differential approaches if they present with gender-specific risk factors. Though gender-specific programming is still a growing area of research in the juvenile justice system, there is evidence that shows positive outcomes from aforementioned programming in areas such as support systems, employment, self-esteem, and empowerment as well as decreases in recidivism (Day et al., 2014; Zahn et al., 2009). In sum, to date, the literature on gender-responsive services has grown; however, there has been a stark absence of evaluation evidence for the support of gender-responsive services to reduce recidivism (Day et al., 2014; Zahn et al., 2009).

Research Summary and Limitations

There are a number of shortcomings in the current body of literature on gender-specific interventions. First, there is a limited amount of strong empirical evidence, specifically restricted by small sample sizes and the lack of experimental/quasi-experimental research design to infer causal effects in evaluating program effectiveness by gender. For example, of the nine gender-responsive evaluations Zahn and colleagues (2009) examined only two using a randomized control design. Additionally, the outcome measures used were highly variable – some examined recidivism rates, drug/alcohol use, school performance, and psychological measures (e.g., self-esteem). Even the most common of outcomes (e.g., recidivism) was not necessarily measured in a consistent or comparable way (Zahn et al., 2009). For example, the WINGS program described measuring recidivism during and after the program, but did not specify how the follow-up dates were calculated or accounted for in the comparison sample.

Day and colleagues' quasi-experimental study had a small sample size and lost approximately 20% of the girls in their gender-responsive program sample because of the

propensity score matching procedures (e.g., there were not enough girls who did not receive residential placement to equivalently match). The subsample of girls that was dropped from the final analyses had higher average propensity scores, a larger proportion of Black girls/smaller proportion of White girls, had more prior detention incidences, and scored lower on most of the mental health assessment scales, limiting the generalizability of the study (Day et al., 2014).

A number of the other gender-responsive evaluation studies did not provide critical information about the studies. For example, Burke and colleagues (2001) did not provide statistical significance testing for the differences in recidivism rates. Similarly, the Wisconsin Department of Corrections (2005) did not provide comparison statistics for differential academic achievement outcomes. To date only two studies utilized randomized treatment and control conditions, whereas there has been a great deal of variation in outcome measures used, variation in program scope and focus, and sample sizes. Most of the evaluations focused on recidivism follow-up and academic success, but the operationalization of these outcomes were highly variable making it difficult to make comparisons or generalizations across programs.

Studies have not employed a longitudinal design to track girls during and after adolescence. As well, most rigorously evaluated programs to date have been on prevention efforts rather than intervention services for girls already involved with the system (Chesney-Lind et al., 2008; Davidson et al., 2011). Finally, Lipsey's studies show little relevant evaluation research information for girls given that most studies do not include gender-based analyses to test for equivalent effectiveness for both boys and girls (e.g., Lipsey, 2009). In its current state, the literature on gender-responsivity provides an expansive ideological argument for its purpose in the juvenile justice system, but little empirical work to support its use.

In sum, there is a lack of evaluation studies of gender-responsive programs and practices and those that do exist tend to lack methodological rigor (Chesney-Lind et al., 2008; Kerig & Schindler, 2013). The non-evaluative literature on gender-responsivity has especially focused on the needs of girls, giving girls voice in this process, and understanding girls' experiences within the system (e.g., Gaarder et al., 2004; Schaffner, 2007). Overall, gender-responsivity attempts to acknowledge gender differences and disparities in the system and develop a response (e.g., programming) to address these gaps and unique needs of girls (Chesney-Lind et al., 2008; Zahn et al., 2009). However, perhaps of greatest concern is the current formal evaluation literature's equivocal support for the use of gender-responsive intervention and a stark absence of rigorous program evaluation for juvenile justice involved samples of youth that disaggregate program effectiveness by gender.

Current Study

The court system is located in a midsized Midwestern county and sees approximately 300-350 new youth annually across their three main divisions (i.e., intake division, truancy court, and standard delinquency). The standard delinquency division is for youth on formal probation in which a juvenile court officer supervises them. In the standard delinquency division girls comprise approximately 25 percent of youth (Onifade et al., 2008). In truancy court, girls comprise approximately 50 percent of youth (Onifade et al., 2009). Over the course of the last decade this court has adopted and implemented the best practices/“what works” framework for juvenile justice interventions (Andrews et al., 1990, Andrews et al., 2011, Lipsey & Cullen, 2007). Specifically, the court system has implemented the risk-needs-responsivity (RNR) model for juvenile assessment, program design, and the tailoring of services based on the criminogenic risk and needs of youth in the court caseload. As well, this particular court system directly provides intensive case management and comprehensive services for youth and many of the core elements that Walker and colleagues (2012) described as fostering a gender-responsive context. Given the history of this particular court in implementing best practices and a stated concern with the needs of female delinquents, it represents a unique opportunity to observe how these principles play out in the field.

The current study investigates the effectiveness of two group homes for adjudicated females in a single county court. In particular, the study examined whether there are differences in re-offense outcomes for girls who received treatment in the group homes versus those who received standard probation services. The proposed study builds upon the gaps in the literature by evaluating a community-based intervention specifically designed for adjudicated girls with a longitudinal design tracking girls’ re-offenses two years following their program release. This

adds to the current literature, which is very limited in terms of the number of outcome studies using recidivism as the outcome of interest for gender-responsive intervention. Within the court setting, it was not feasible to conduct an experiment in which girls were randomly assigned to a treatment condition (group home) or a control group (standard probation services). Instead, this study employed a quasi-experimental design using propensity score matching in order to control for potential selection effects. The data were examined using two propensity score matching models. With this design, the study attempts to answer the following research questions:

Research Question 1: *Are there differences between girls who received group home treatment and girls who received standard probation services in terms of demographic and criminogenic risk factors?*

Research Question 2: *When matched based on propensity scores, are there any remaining differences between girls who received group home treatment and girls who received standard probation services in terms of their demographic and criminogenic risk factors?*

Research Question 3: *When matched based on propensity scores, are there differences in one- and two-year recidivism rates between girls receiving group home treatment and girls who received standard probation services?*

This is the first study to formally evaluate community-based gender-responsive girls' group homes. Given the limited experimental and quasi-experimental research on this topic, this study adds to the evaluation literature on gender-responsive programming.

CHAPTER 3: METHODS

Setting

The study was conducted in a mid-sized juvenile county court in the Midwest using archival data collected between 2005-2012. The juvenile court is comprised of multiple divisions including standard delinquency or formal probation, intake or informal probation, and truancy court. Girls involved with any division of the court, but typically formal probation or truancy court, may receive an out-of-home placement in a girls' group home for treatment. At the time of the study, this court system had two group homes for girls. All girls in truancy court or on formal probation are assigned to a juvenile court officer for supervision and treatment recommendations. Pertinent to the research questions, admission to a group home was determined by girls' initial risk assessment scores and other contextual factors deemed important by the assigned juvenile court officer and juvenile court staff.

Group Home Program Descriptions

Moderate Risk Group Home

The first group home is a facility for female adolescent offenders involved with the county court. Girls classified as moderate risk according to their Youth Level of Service/Case Management Inventory (YLS/CMI) score are eligible for this group home. This is a community-based program that seeks to improve the behaviors, social skills, self-esteem, and coping mechanisms of female juvenile offenders between the ages of 12 and 17. Program goals include reducing recidivism, improving of school performance, improving communication skills, developing individual self-worth, and building positive relationships and support systems. Personalized treatment plans are created for each individual youth by utilizing assessment tools

such as the Child and Adolescent Functional Assessment Scale (CAFAS), How I Think Questionnaire (HIT), and Juvenile Automated Substance Abuse Evaluation Survey (JASAE), among others. The group home aims to accomplish program goals by implementing Cognitive-Behavioral Therapy (CBT), Thinking for a Change (T4C) behavior curriculum, Girls Moving On (GMO) gender-responsive programming, and an aftercare service group. There is presently limited empirical evidence as to the effectiveness of T4C and GMO with juvenile offenders. Other structured program amenities include classes and activities such as scrapbooking, substance abuse education, tutoring, and community service projects. The average program duration is five to seven months of placement with approximately 100-200 hours of treatment services. The number of hours completed is dependent on the needs of each individual girl.

High Risk Group Home

The second facility was created through a collaborative effort between a community-based youth services initiative and the county court in March 2005 as an alternative option for moderate to high-risk female offenders in the juvenile division. Typically, this program serves high-risk girls as determined by the YLS/CMI. It is also a non-secure residential treatment center that provides services for female youth between the ages of 13 and 18 whom are currently on probation in either the truancy or delinquency division of the juvenile court. The group home has the following program goals: (1) to provide intensive and high fidelity evidence-based treatment services and programming for girls, (2) to reduce girls' recidivism as well as duration of stay at the residential treatment center, (3) to provide thorough and high quality aftercare services that reinforce community stability and the reconnecting of family unification and support systems.

Treatment models are based on criminogenic needs, family demands, substance abuse, and individual's needs based on assessment tools such as the Massachusetts Youth Screening

Inventory (MAYSI), Client Evaluations of Self and Treatment (CEST), Youth Level of Service/Case Management Inventory (YLS/CMI), and the Juvenile Automated Substance Abuse Evaluation (JASAE). Treatment models target these needs through the use of cognitive-behavioral therapy (CBT) in-group, individual and family settings, and seek to modify behavior patterns. These behavior patterns are approached by addressing such personality characteristics as antisocial cognitions, teaching non-criminal and alternative behaviors, building individuals skills such as problem-solving, anger management, communication, self-efficacy and means of coping, building positive relationships with peers and family members, enhancing school performance, reducing conflict by avoidance of negative social situations, and increasing pro-social leisure activities. Average program duration is four to six months of placement with approximately 100-200 hours of treatment services with the number of hours completed being dependent on the needs of individual girls. Aftercare services are also provided when girls transition out of the group home.

Gender-Responsive Elements of the Group Homes

Table 2 provides an overview of the elements of each of the group homes and the extent to which those program elements have been identified in the literature as meeting criteria for gender-responsive programming. Both group homes incorporate theoretically informed gender-responsive elements such as pathways (e.g., addressing trauma, abuse, and neglect) and RCT (e.g., focusing on the centrality of relationships, inclusion of girls' voice, and sense of connection to others). A variety of information was surveyed in order to determine the extent to which the group homes met the criteria outlined in the gender-responsivity literature. First, a systematic review of the group home program manuals and materials (e.g., curricula) were reviewed and contrasted to the core elements described in the gender-responsivity literature (see

Table 1 in literature review for detailed descriptions of each of the elements). For example, the group homes utilize Girls Moving On, which heavily focuses on improving girls' relationships—a core component of gender-responsive treatment.

Table 2. Program Elements Matching Gender-Responsive Criteria	
Community-Based	<ul style="list-style-type: none"> • Non-secure facility; • Girls continue attending their main school; • Girls can earn weekend passes to visit family;
Relational Focus	<ul style="list-style-type: none"> • Building positive relationships with other females in placement; • Building relationships with positive female role models; • Working on family relationships and reunification with the girls' families;
Trauma-Informed	<ul style="list-style-type: none"> • Utilizing the Girls Moving On Curriculum; • Trauma-Based Counseling;
Mental Health Needs	<ul style="list-style-type: none"> • Cognitive Behavioral Therapy; • Dialectical Behavior Therapy; • Substance abuse counseling; • Anger replacement training;
Communication and Relationships	<ul style="list-style-type: none"> • Small staff-to-youth ratio; • Extensive staff training on the needs of girls; • Individual therapy; • Individual goal planning;
Comprehensive	<ul style="list-style-type: none"> • Integrate family, schools, and other community-based agencies;
Use of Gender-Informed Protocols	<ul style="list-style-type: none"> • Girls Moving On Curriculum
Resources for Girls	<ul style="list-style-type: none"> • Rewards-based system; • Girls Moving On Curriculum

In addition to the manuals and program descriptions, interviews were conducted with juvenile court officers, a group home program manager, and court managers in which the practitioners described (1) considerations in placing girls into the program, and (2) details about

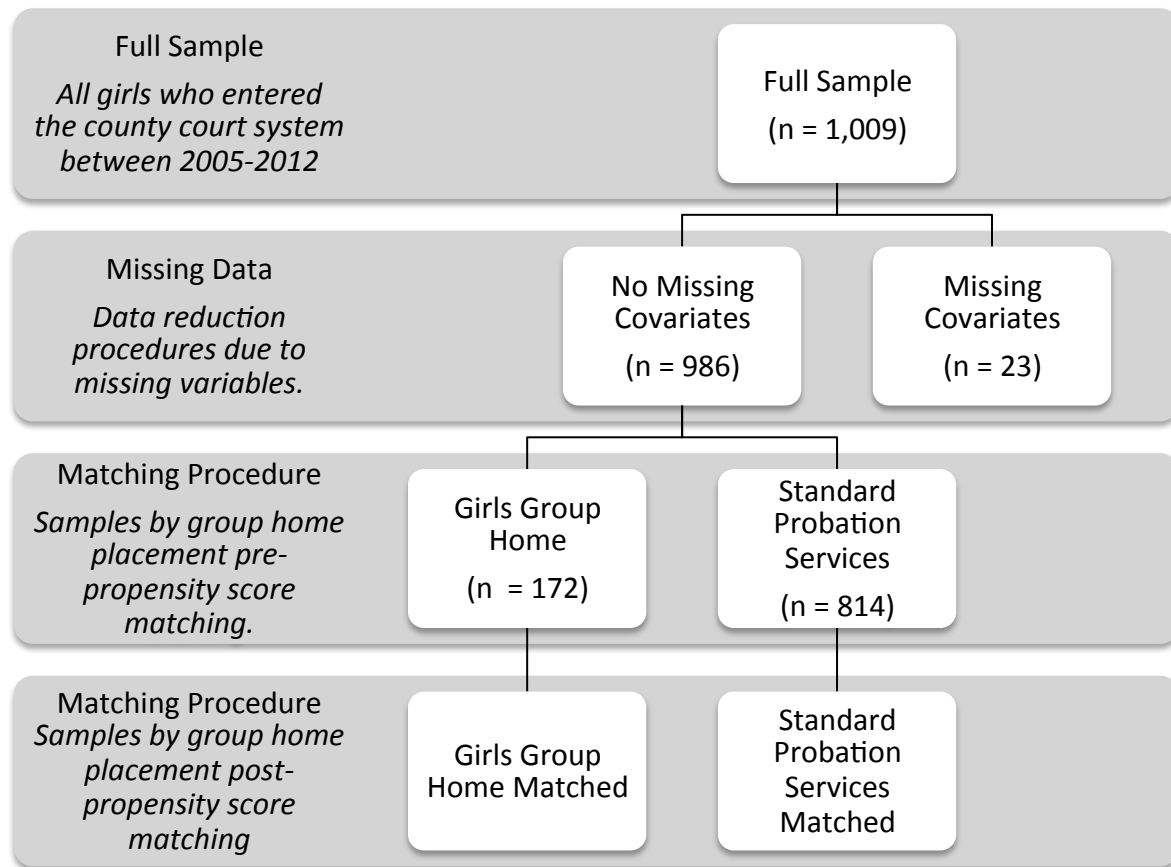
the core program elements and goals. In addition to the interviews, the weekly juvenile court staffing meetings were observed to better understand the placement process of girls into one of the group homes. These interviews and observations are part of a larger study (see Anderson & Davidson, 2015). Finally, the group home staff meets frequently to discuss girls' progress in the program (e.g., improvements in academics, progress in therapy), any issues (e.g., running away from the group home, fighting with other girls in the group home), and aftercare plans if the girl is nearing her exit time (e.g., discussing her home environment, alternative housing options, etc.). Each element that was identified in the program materials and/or through the interviews and observations were listed in Table 2.

Of particular interest in the group homes programs was their use of the Girls Moving On (GMO) gender-responsive curriculum. According to Orbis Partners (2014), GMO is a gender-responsive cognitive-behavior treatment program for at-risk girls between the ages of 12 to 21 years old. The program is informed by relational/cultural theory and incorporates both cognitive-behavior treatment approaches and motivational interviewing. The main goal of the program is to provide girls with skills, resources, an increased capacity for healthy relationships, and to reduce girls' risk for juvenile and criminal justice system involvement (Orbis Partners, 2014). The program includes seven modules delivered in both individual and group settings focusing on listening, healthy relationships, emotion expression, connections, and healthy decision-making (Orbis Partners, 2014). While there is some descriptive information available on this curriculum, to date, there are no reported outcome data on the curriculum (Kerig & Schindler, 2013; Zahn et al., 2008).

Participants

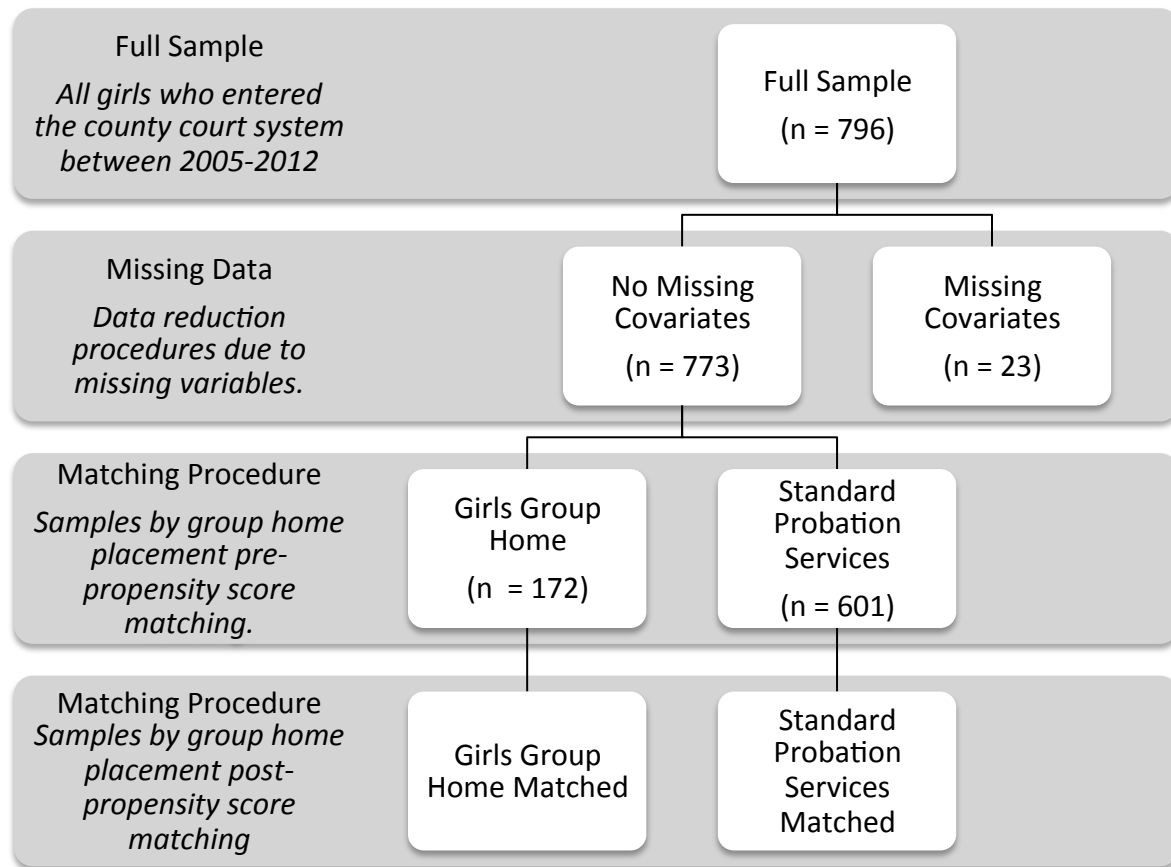
Data for this study were collected from the county court data management system for dates between 2005-2012. Girls in the sample came from both the delinquency (formal probation) and truancy divisions. All data were extracted for all girls in those two divisions during the relevant time period. Included in these databases are girls' race/ethnicity, age, division (e.g., whether the girl originally entered court via formal probation or via truancy court), criminogenic risk as determined by the Youth Level of Service/Case Management Inventory (YLS/CMI), length of time in the program, and other court-run or referred programming tracked in the county's data management system (e.g., secure residential placements). Two different models were analyzed in the study. Model A includes a larger sample of comparison girls than Model B, because it does not attempt to match on time at program exit. Therefore, there was a larger sample of girls that had two-year follow-up data. The sampling procedures prior to propensity score matching for Model A are outlined in Figure 1.

Figure 1: Sampling Procedures Pre-Propensity Score Matching: Model A



Model A includes 1,009 in the full sample for which there is two-year follow-up data and no missing covariates. Twenty-three girls were dropped from the full sample because they were missing key covariates (e.g., YLS/CMI scores). Within the final sample of 986 girls, 172 of the girls received the group home placement and 814 of the girls did not receive group home placement during the same time frame. The sampling procedures prior to propensity score matching for Model B are outlined in Figure 2.

Figure 2. Sampling Procedures Pre-Propensity Score Matching: Model B



Model B is a time control model in which comparison girls' recidivism is measured from a proximal exit date. The full sample for Model B includes 796 youth. There are 172 girls who received group home treatment and 601 non-group home girls involved in the court system during the same time period. Model B has a smaller sample due to creating an equivalent follow-up metric. Girls in the control group for Model B needed to have four-year follow-up data from their entry into the court because the model controls for time (see the following section 'Recidivism' for more detail). A summary of Model A and Model B, including similarities and differences between the models, are described in Table 3.

Table 3. Summary of Models

Elements	Model A	Model B
Total Covariates in Model	11 Covariates	16 Covariates
Program Variables	None	Five Programs
Time Frame – Control	YLS Date	Proxy Program Exit Date
Time Frame – Treatment	Program Exit Date	Program Exit Date
Original Sample Size	986 Girls	773 Girls
# of Matched Pairs	169	149
Matching Procedure	1:1 Nearest Neighbor, dl < .25	1:1 Nearest Neighbor, dl < .25

Measures**Youth Level of Service/Case Management Inventory (YLS/CMI)**

Juvenile court officers received extensive training on administering the YLS/CMI including coding cases, calculating inter-rater reliability across items, and discussing any disagreements. The YLS/CMI is comprised of 41 items across eight subscales. The subscales include prior offenses, education, leisure activities, peer relationships, substance abuse, family relationships, attitudes, and personality. All of the items and subscales of the measure are included in Appendix B. The eight YLS/CMI subscales were used as covariates in both Model A and Model B.

Additional Program Variables

There were five additional program variables used in Model B. Each of the five programs were coded as a binary variable (0 = no, 1 = yes) if girls received or were receiving the service prior to entry into the group home, as this is a pre-treatment covariate. If the girl did not receive the program until during treatment or after treatment it was coded as “no” for the matching procedure. Descriptions for each of the five programs and descriptive statistics are included in the following sections. Additionally, Table 4 includes a breakdown of the number of girls in the treatment group and in the control group who received each of these programs at any point in

time while under court jurisdiction, prior to group home placement, and after release from the group home.

Table 4. Proportions of Girls Receiving Additional Court-Run Services

Program	Group Homes (n = 172)	Standard Probation Services (n = 601)
Residential Placement	84 (48.8%)	59 (9.8%)
Residential Prior	53 (30.8%)	43 (7.2%)
Residential After	40 (23.3%)	14 (2.3%)
Intensive Probation Services (IPS)	61 (35.5%)	84 (14.0%)
IPS Prior	50 (29.1%)	67 (11.1%)
IPS During	18 (10.5%)	38 (6.3%)
IPS After	20 (11.6%)	34 (5.7%)
Family Support Services (FSS)	83 (48.3%)	50 (8.3%)
FSS Prior	38 (22.1%)	35 (5.8%)
FSS During	59 (34.3%)	21 (3.5%)
FSS After	64 (37.2%)	20 (3.3%)
Alternative School	49 (28.5%)	15 (2.5%)
Alternative School Prior	16 (9.3%)	7 (1.2%)
Alternative School During	34 (19.8%)	7 (1.2%)
Alternative School After	49 (28.5%)	11 (1.8%)
Evening Reporting	47 (27.3%)	31 (5.2%)
Evening Reporting Prior	30 (17.4%)	16 (2.7%)
Evening Reporting During	19 (11.0%)	12 (2.0%)
Evening Reporting After	25 (14.5%)	16 (2.7%)

Notes: It is not possible to receive residential at the same time as the group home so there is no 'Residential During' category. Also, these time categories are not mutually exclusive—for example, the same girl might be counted in both the 'prior' and 'after' categories if they were receiving the service for the full duration of the group home treatment period or were placed in the program more than once.

Residential Placement

Residential programming includes any out-of home placement in a secure facility as a result of their contact with the juvenile justice system. A youth might be sent to residential placement for temporary detention (e.g., awaiting disposition), long-term placement after being adjudicated for an offense, or while waiting for placement elsewhere. In the full sample of juveniles, 18.5% of girls received residential placement at any point in time while under court supervision and 48.8% of girls who received group home treatment also received residential placement, in which 30.8% was prior to treatment and 23.3% was after leaving the group home.

Intensive Probation Services (IPS)

Intensive Probation Services (IPS) was designed to provide counseling and intensive probationary supervision for adjudicated youth who meet either one of these criteria:

(1) adjudicated youth who would otherwise have been committed to the state's Department of Human Services or placed in a private residential facility, or (2) adjudicated youth who would have otherwise been placed in foster care, but were determined as being able to stay in their home given the aid of more intensive supervision and counseling services. IPS aims to reduce the number of out-of-home days youth are placed and therefore reduce costs of residential placement for the county and for the state. IPS is individualized to fit the needs of each youth and includes increased personal contact (e.g., daily) between the assigned juvenile court officer and the juvenile as well as the juvenile's family, school, and other agencies with which the youth is involved. In the full sample of juveniles, 18.8% of girls received IPS at any point in time while under court supervision and 35.5% of girls who received group home treatment also received IPS, in which 29.1% was prior to treatment, 10.5% was during treatment, and 11.6% received or were still receiving IPS after leaving the group home.

Family Support Services

Family Support Services (FSS) focuses on assisting parents of adjudicated youth. It provides guidance and support from the court through enhancing parenting strategies and utilizing a holistic strategy to help the family effectively function. The main goals of FSS include: (1) preventing the placement of the juvenile in an out-of-home facility, (2) providing support for earlier return and efficient re-entry of juveniles who have been placed out-of-home, (3) providing assistance to parents in need of stabilizing their home environment, (4) strengthening parenting skills, and (5) addressing barriers to effective parenting. FSS includes

weekly in-home meetings between parents and the assigned juvenile court officer, assessment of parental and family needs, individualized goal planning for the family developed in collaboration with the parent, parenting classes and support groups, court-run and community-based services for the juveniles, and assistance in accessing community resources. In the full sample of juveniles, 17.2% of girls received FSS at any point in time while under court supervision and 48.3% of girls who received group home treatment also received FSS, in which 22.1% was prior to treatment, 34.3% was during treatment, and 37.2% received or were still receiving FSS after leaving the group home.

Court-Run Alternative School

The alternative school is a highly structured day treatment program that is located outside of the court, but is operated and funded by the court. The purpose of the alternative school is to provide youth with an individualized learning environment that offers not only education, but also vocational and community-based services. The main goal of the alternative school is to provide youth the opportunity to receive a high school diploma and services to assist with that goal. The alternative school works with local school districts to provide an education and issue school credits. The school also collaborates with other community-based programming to assist with behavior management and obtaining employment. All youth who attend the alternative school also receive FSS for at least 90 days.

In the full sample of juveniles, 8.3% of girls went to the alternative school at any point in time while under court supervision and 28.5% of girls who received group home treatment also attended the alternative school, in which 9.3% was prior to treatment, 19.8% was while in the group home, and 28.5% attended or were attending the alternative school after leaving the group home.

Court-Run Evening Reporting Program

The court-run evening reporting program, aims to reduce the YLS/CMI risk level of youth by addressing criminogenic factors through the use of cognitive-behavioral intervention. The evening reporting program meets three times each week for high-risk youth and twice each week for moderate-risk youth (as determined by the YLS/CMI). Youth participate in a cognitive-behavioral program, Thinking for a Change (T4C), to modify patterns of criminogenic thinking. The program is set up with a five-phase system in which each phase takes four weeks to complete. In addition to the cognitive-behavioral treatment, the program staff also works with juveniles on incorporating other pro-social activities into their free time.

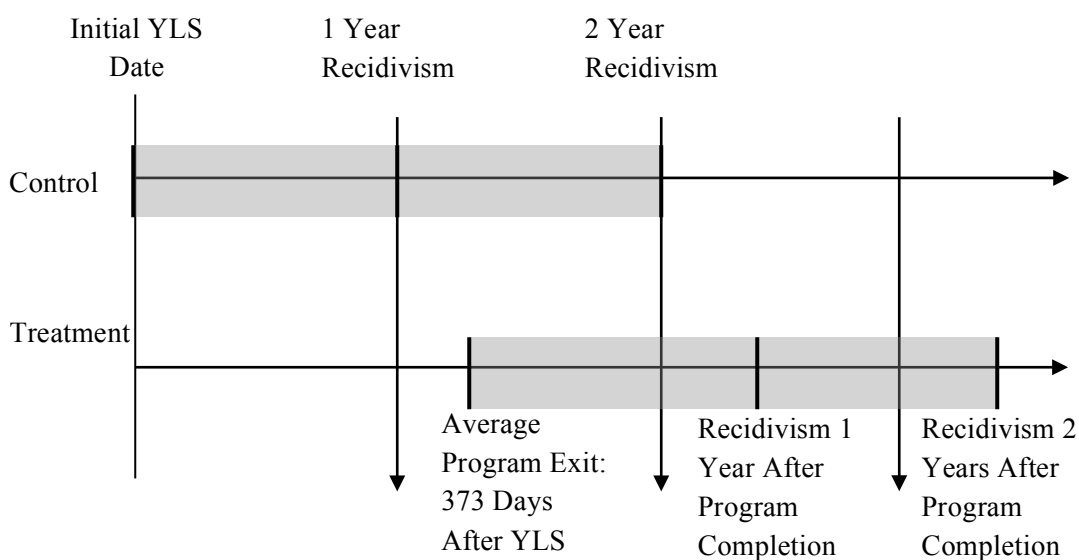
In the full sample of juveniles, 10.1% of girls received the evening reporting program at any point in time while under court supervision and 27.3% of girls who received group home treatment also received the evening reporting program, in which 17.4% was prior to treatment, 11.0% was while in the group home, and 14.5% received or were still receiving the evening reporting program after leaving the group home.

Recidivism

The dependent variable for the study was recidivism. Recidivism was collected from the court data management system. Recidivism, coded as a binary variable (0 = no, 1 = yes), was defined as any new petition to court two years (24 months) following their initial YLS/CMI assessment. Probation violations were not counted as re-offenses. If the girl aged out during the follow-up period, adult records were checked as well. A separate recidivism run was conducted for the girls who received group home treatment. In this separate run, instead of using the initial YLS/CMI assessment date, the girls' exit date from the program was used for the follow-up time interval to serve as a better proxy for long-term effectiveness of the program. Model A and

Model B use different time periods for the control group. Model A examined recidivism rates of girls in the control group from their YLS/CMI date. Model B examined recidivism rates of girls in the control group from a hypothetical exit-from-program date that mirrors the length of time between the treatment group’s entry into court and exit from the group homes. The time frames for the two models are presented in Figure 3 (Model A) and Figure 4 (Model B).

Figure 3. Recidivism Time Frames: Model A



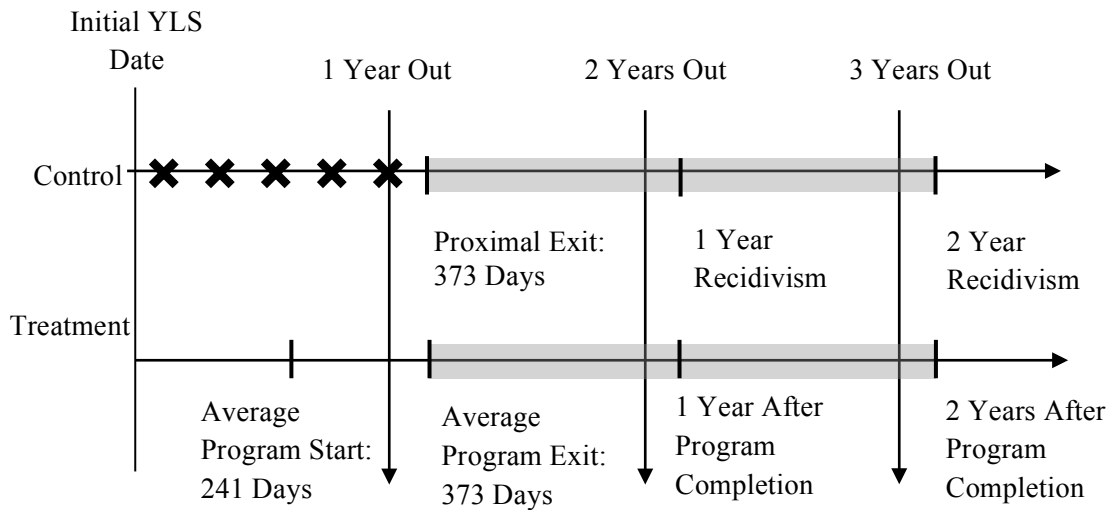
Constructing an Equivalent Time Metric: Model B

Since the control group does not have an exit from program date to measure, the recidivism follow-up period, a proxy for an exit date was constructed for the control sample using the mean of length of time between initial YLS/CMI assessment and exit from group home. It is important to create equivalent time frames since researchers have noted that time under court supervisions increases the likelihood of receiving future petitions (Barnes et al., 2015). Measuring recidivism from program release can mitigate the potential confounding effect of court supervision on recidivism.

The average time between YLS/CMI date and group home program start was 241.26 days ($SD = 223.98$). Days were normally distributed with a skewness of 0.76 ($SE = 0.19$). The average length of time girls were in the group was 132.00 days ($SD = 86.37$). The distribution was slightly skewed with skewness statistic of 1.08 ($SE = .185$). The median length of time girls spent in the group home was 118 days. Finally, the average length of time between the YLS/CMI date and exit from the group home was 373.26 days ($SD = 237.61$). Days were normally distributed with a skewness of 0.561 ($SE = 0.19$).

The program exit date is an absolute time for the treatment group, but for the control group a proximal exit date was used—373 days after the initial YLS/CMI assessment. Any petitions dated one and two years following that proximal exit date were counted as one- and two-year recidivism. That is, petitions prior to the actual exit date for the treatment group or the proximal exit date for the control group were not counted as re-offenses since these would have occurred prior to or during the treatment time period. Creating a proxy exit date for the control sample caused a reduction of 26.2% of the sample of control girls from Model A ($n = 814$) to Model B ($n = 601$). Figure 4 depicts how equivalent recidivism time frames were constructed for the control group and the treatment group in Model B.

Figure 4. Constructing Equivalent Time Periods for Recidivism: Model B



It was also necessary to create equivalent pre-treatment time frames for the other program variables (e.g., residential placement) for the control and treatment groups. For purposes of the matching procedure, only youth who received any of the five programs prior to starting the group home (for the treatment group) or prior to the average program start time (e.g., 241 days) for the control group were coded as “yes” for the program pre-treatment.

Propensity Score Matching

All data were entered in SPSS and analyzed using SPSS and the R plug-in psmatch. The primary data analytic strategy for the study was propensity score matching (PSM) in order to control for potential selection effects in a non-randomized design and produce a statistical balance on the observed covariates in the analysis (see Guo & Fraser, 2010; Stuart, 2010; Thoemmes & Kim, 2011 for reviews). In particular, PSM is a statistical matching technique that determines the estimated effect that predicted covariates have on particular intervention outcomes (Austin, 2011). The PSM analysis consisted of several steps. First, a binary logistic regression was calculated to create probability scores by regressing group home membership (0

= non-group home, 1 = group home) on the unique, theoretically salient variables on which full data exists for the final sample of girls. Model A utilized 11 propensity score covariates including age of the youth at which she received the YLS/CMI assessment, race/ethnicity (dichotomously coded as white or non-white), court division (dichotomously coded as delinquency or truancy), and the eight YLS/CMI subscale scores. Model B utilized 16 covariates including all Model A covariates plus five additional pre-treatment program variables (all dichotomously coded for the pre-treatment time period as yes or no). For each of the models, these covariates were combined into a single probability score to predict group home membership.

After the propensity scores were created, a nearest neighbor matching procedure was used to identify girls of similar risk and demographic profiles. In particular, the study utilized a 1:1 nearest neighbor caliper matching procedure without replacement based on the propensity score estimates produced by the logistic regression. Each girl receiving group home placement was matched to a girl who did not receive group home placement with the closest propensity score within 25 percent of the standard deviation of the propensity score estimate. This ensures that all matched pairs are similar in their propensity to receive treatment (Guo & Fraser, 2010; Rosenbaum & Rubin, 1985). This level ensures that nearly no imbalances remain between the two groups on the selected covariates. Utilizing this matching procedure allowed for developing a comparison group of girls with the closest propensity scores from the full sample of non-group home girls.

CHAPTER 4: RESULTS

This chapter begins by presenting the descriptive statistics for girls' demographic characteristics, risk assessment scores, and participation in other programming. This chapter is organized by model and by research question. The results for Model A are presented first and the results for Model B are presented second.

Propensity Score Matching: Model A

Research Question 1: *Are there differences between girls who received group home treatment and all girls who received standard probation services in terms of demographic and criminogenic risk factors?*

Descriptive statistics for the entire, unmatched, sample of girls are presented in Table 5 including X^2 and t statistics. The unmatched sample significantly varied on a number of key variables. Group home girls scored significantly higher than non-group home girls on all eight of the criminogenic risk/need domains on the YLS/CMI. Girls who received the group home treatment had a higher overall YLS/CMI score ($M = 18.91$, $SD = 5.88$) than standard probation girls ($M = 13.33$, $SD = 6.43$), $t(984) = -10.52$, $p < .001$. Girls did not significantly differ in age at which they entered the system, nor did they differ based on race/ethnicity or court division (delinquency or truancy). However, given the differences in criminogenic risk and need, it was important to compare girls who received group home treatment to other girls who did not receive group home treatment, but had similar risk and demographic profiles.

Table 5. Descriptive Statistics of Unmatched Samples of Girls: Model A

Covariates	<i>M (SD)</i>		Significance Test (<i>t</i> or X^2)
	Group Homes (<i>n</i> = 172)	Standard Probation Services (<i>n</i> = 814)	
% Non-White	64.0%	63.8%	0.02
% Delinquency	50.0%	43.6%	2.34
Age	14.15 (1.21)	14.36 (1.36)	1.91
Prior Offenses	0.38 (0.89)	0.25 (0.66)	-2.23*
Education	4.11 (1.54)	3.21 (1.70)	-6.45*
Leisure	2.02 (0.73)	1.75 (0.93)	-3.54*
Peers	2.70 (1.25)	1.91 (1.24)	-7.62*
Substance Abuse	1.40 (1.52)	0.77 (1.23)	-5.83*
Family	3.63 (1.59)	2.36 (1.71)	-8.99*
Attitudes	1.13 (1.23)	0.65 (1.07)	-5.16*
Personality	3.55 (1.62)	2.43 (1.78)	-7.59*

* $p < .001$

Research Question 2: *When matched based on propensity scores, are there any remaining differences between girls who received group home treatment and girls who received standard probation services in terms of their demographic and criminogenic risk factors?*

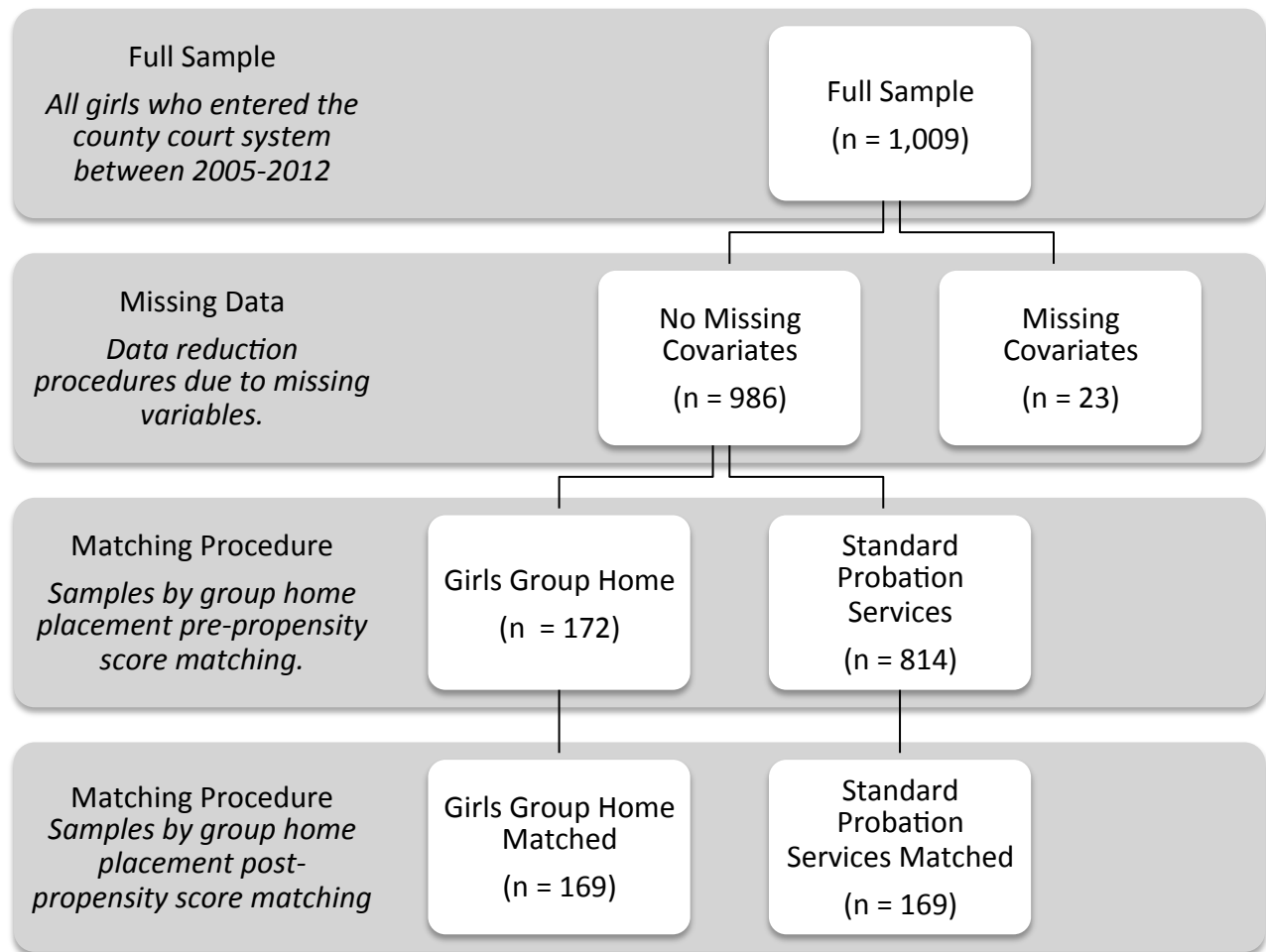
Propensity score matching was used based on entering each of the 11 specified covariates in Table 5 into a logistic regression model using group home membership as the binary outcome variable (yes = 1, no = 0). The conditional probabilities of being referred to the girls' group homes were used from the logistic regression model to match group home girls to girls on standard probation that did not receive group home treatment. Using SPSS and the R plug-in psmatch, a nearest-neighbor one-to-one matching procedure without replacement created the probability scores (Ho, Imai, King, & Stuart, 2007; Hansen, 2004; Hansen & Bowers, 2008; Thoemmes & Kim, 2011).

A successful matching procedure is defined by the reduction or removal of all previous significant differences between groups on any of the covariates. Additionally, the balance statistics provided by psmatch signified no covariates greater than $|\text{ld}| = .25$. Imposing a tolerance

level, or a caliper (e.g., $ldl = .25$), is important when utilizing nearest neighbor matching to improve the quality of the final matches (Caliendo & Kopeinig, 2005). Applying a caliper means that cases selected from the comparison group are matched to treatment cases that lie within this propensity range. One drawback of caliper matching is that it is difficult to know *a priori* what caliper level is reasonable, but most scholars suggest one quarter of a standard deviation or less is appropriate (Smith & Todd, 2005).

Figure 5 presents the final sampling scheme after the propensity score matching algorithm. The final matched sample includes 169 girls from the treatment group and 169 matched girls from standard probation services who never received group home placement. Only three (1.7%) of the girls in the treatment group were lost as a result of the caliper matching procedure because there was not an equivalent girl in the standard probation services comparison group who had a similar propensity score within $ldl = .25$.

Figure 5. Final Sampling Scheme Post-Propensity Score Matching: Model A



As seen in Table 6, the propensity score matching procedure created balanced mean scores across every observed covariate utilizing paired-sample analyses. The matching procedure was successful given that the mean probability score for the group home girls ($M = .283$, $SD = .15$) and the non-group home girls ($M = .276$, $SD = .15$) were equivalent for the two groups $t(168) = -.429$, $p = .668$. As well, girls' total YLS/CMI scores were not significantly different between the group home girls ($M = 18.75$, $SD = 5.79$) and the non-group home girls ($M = 18.64$, $SD = 6.10$), $t(168) = -.174$, $p = .862$. Therefore, there is reason to believe that any selection effects due to age, race/ethnicity, juvenile court unit, and criminogenic risk scores were controlled when matched based on propensity scores.

Table 6. Descriptive Statistics of Matched Samples of Girls: Model A

Covariates	<i>M (SD)</i>		Significance Test (<i>t or X²</i>)
	Group Homes (n = 169)	Standard Probation Services (n = 169)	
% Non-White	64.5%	59.2%	1.02
% Delinquency	50.3%	50.3%	0.00
Age	14.17 (1.20)	14.16 (1.40)	-0.04
Prior Offenses	0.37 (0.85)	0.33 (0.77)	-0.53
Education	4.08 (1.54)	4.07 (1.60)	-0.07
Leisure	2.01 (0.73)	2.12 (0.72)	1.43
Peers	2.69 (1.25)	2.63 (1.14)	-0.41
Substance Abuse	1.36 (1.49)	1.40 (1.59)	0.21
Family	3.62 (1.60)	3.66 (1.60)	0.27
Attitudes	1.10 (1.21)	1.10 (1.28)	0.00
Personality	3.53 (1.62)	3.33 (1.65)	-1.10

Next, one- and two-year recidivism rates were examined to see if there were differences between the two groups. In other words, was the girls' group home intervention effective in reducing recidivism over and above the standard probation services that similarly situated girls received?

Research Question 3: *When matched based on propensity scores, are there differences in one- and two-year recidivism rates between girls receiving group home treatment and girls who received standard probation services?*

Table 7 presents the one-year recidivism rates for the group home girls and the matched standard probation services. The sample of girls who received group home treatment had a one-year recidivism rate of 22.5% following program exit. The matched sample of girls who did not receive group home treatment had a one-year recidivism rate of 30.8%. A McNemar chi-square test did not reveal a statistically significant difference between the paired samples, $X^2 (1, N = 169), p = 0.10$ (exact *p* value).

Table 7. One-Year Recidivism Rates: Model A

	Recidivists n (%)	Non-Recidivists n (%)	Total
Group Homes	38 (22.5%)	131 (77.5%)	169
Standard Probation Services	52 (30.8%)	117 (69.2%)	169

Table 8 presents the two-year recidivism rates for the group home girls compared to standard probation services girls. The sample of girls who received group home treatment had a two-year recidivism rate of 28.4% following program exit. The matched sample of girls who did not receive group home treatment had a two-year recidivism rate of 42.0%. A McNemar chi-square test revealed a significant difference between the paired samples, $X^2(1, N = 169), p = 0.01$ (exact p value).

Table 8. Two-Year Recidivism Rates: Model A

	Recidivists n (%)	Non-Recidivists n (%)	Total
Group Homes	48 (28.4%)	121 (.716%)	169
Standard Probation Services	71 (42.0%)	98 (58.0%)	169

Propensity Score Matching: Model B

The second model differed from the first model because recidivism was calculated from a proximal exit date for the control group rather than from their entry into the court (refer back to Figure 4). The creation of this proxy exit date allowed for the inclusion of the five additional program variables as covariates in the model. The same three research questions were explored. Research Question 1: *Are there differences between girls who received group home treatment and all girls who received standard probation services in terms of demographic and criminogenic risk factors?*

Descriptive statistics for the entire, unmatched, sample of girls are presented in Table 9 including X^2 and t statistics. The unmatched sample significantly varied on a number of key

variables. Group home girls scored significantly higher than non-group home girls on seven of the eight criminogenic risk/need domains of the YLS/CMI. The prior offenses subscale was the only non-significant YLS/CMI subscale. The group home girls also received all five of the other court-run programs at a significantly higher rate than girls who did not receive group home treatment prior to entering the group home. In total, 12 of the 16 covariates were significantly different between the unmatched groups. Finally, girls who received group home treatment had a higher overall YLS/CMI score ($M = 18.91$, $SD = 5.88$) than standard probation girls ($M = 13.01$, $SD = 6.67$), $t(771) = -10.51$, $p < .001$. Given these differences, it was important to compare these girls to other girls who did not receive group home treatment, but had similar risk and demographic profiles.

Table 9. Descriptive Statistics of Unmatched Samples of Girls: Model B

Covariates	<i>M (SD)</i>		Significance Test (<i>t or X</i> ²)
	Group Homes (<i>n</i> = 172)	Standard Probation Services (<i>n</i> = 601)	
% Non-White	64.0%	62.9%	0.64
% Delinquency	50.0%	43.6%	2.22
% Residential	30.8%	7.2%	68.82*
% IPS	29.1%	11.1%	33.44*
% FSS	22.1%	5.8%	41.39*
% Alternative School	9.3%	1.2%	30.68*
% Evening Reporting	17.4%	2.7%	52.19*
Age	14.15 (1.21)	14.34 (1.38)	1.67
Prior Offenses	0.38 (0.89)	0.32 (0.75)	-0.92
Education	4.11 (1.54)	3.13 (1.72)	-6.73*
Leisure	2.02 (0.73)	1.70 (0.98)	-3.91*
Peers	2.70 (1.25)	1.82 (1.24)	-8.21*
Substance Abuse	1.40 (1.52)	0.76 (1.25)	-5.64*
Family	3.63 (1.59)	2.32 (1.74)	-8.87*
Attitudes	1.13 (1.23)	0.65 (1.08)	-4.94*
Personality	3.55 (1.62)	2.31 (1.83)	-8.59*

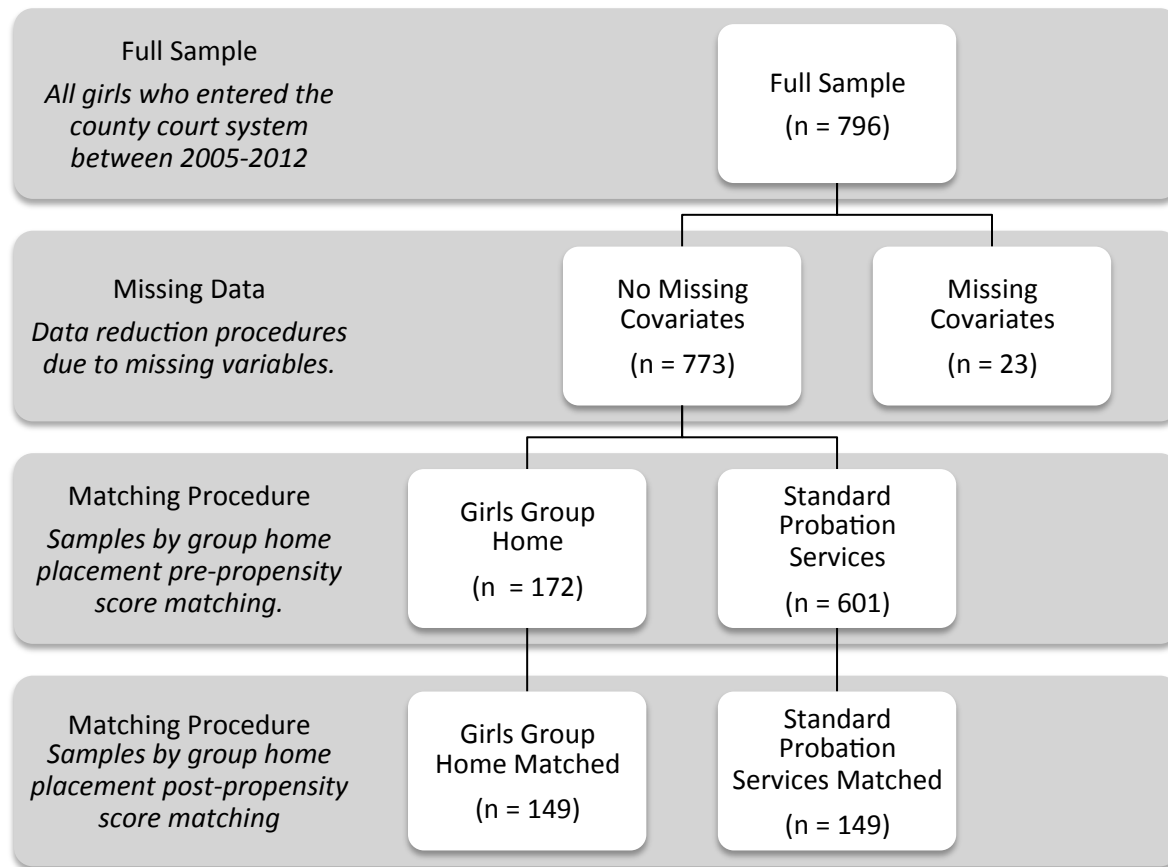
* $p < .001$

Research Question 2: *When matched based on propensity scores, are there any remaining differences between girls who received group home treatment and girls who received standard probation services in terms of their demographic and criminogenic risk factors?*

Propensity score matching was used based on entering each of the specified covariates in Table 9 into a logistic regression model using group home membership as the binary outcome variable (yes = 1, no = 0). The conditional probabilities of being referred to the girls' group homes were used from the logistic regression model to match group home girls to girls on standard probation that did not receive group home treatment. The same analytic procedure was conducted as in the first model. Using SPSS and the R plug-in psmatch, a nearest-neighbor one-to-one matching procedure without replacement created the probability scores (Ho, Imai, King, & Stuart, 2007; Hansen, 2004; Hansen & Bowers, 2008; Thoemmes & Kim, 2011).

Figure 6 presents the final sampling scheme after the propensity score matching algorithm. The final matched sample includes 149 girls from the treatment group and 149 matched girls from standard probation services who never received group home placement. Twenty-three (13.4%) of the girls in the treatment group were lost as a result of the caliper matching procedure because there was not an equivalent girl in the standard probation services comparison group who had a similar propensity score within $|\text{dl}| = .25$.

Figure 6. Final Sampling Scheme Post-Propensity Score Matching



This subsample of girls who were eliminated from the sample as a function of the matching procedure was different from those who were matched to a control girl in a number of ways. First, in terms of the overall YLS/CMI score, unmatched group home girls scored significantly higher ($M = 21.91, SD = 4.61$) than matched girls ($M = 18.46, SD = 5.93$), $t(170) = 2.67, p < .01$. There were a number of differences between the two groups on the YLS/CMI subscales including the substance abuse, peer relationships, and education domains. Finally, four of the five additional program variables also significantly differed between group home girls who were matched and the group home girls who were dropped from the sample. These results are summarized in Table 10.

Table 10. Descriptive Statistics of Matched and Unmatched Girls: Model B

Covariates	<i>M (SD)</i>		Significance Test (<i>t</i> or X^2)
	Matched (n = 149)	Unmatched (n = 23)	
% Non-White	63.1%	69.6%	0.36
% Delinquency	49.7%	52.2%	0.05
% Residential	28.2%	47.8%	3.60
% IPS	26.2%	47.8%	4.53*
% FSS	18.8%	43.5%	7.06*
% Alternative School	5.4%	34.8%	20.43*
% Evening Reporting	11.4%	56.5%	28.16*
Age	14.13 (1.23)	14.26 (1.10)	0.49
Prior Offenses	0.38 (0.90)	0.35 (0.78)	-0.17
Education	3.99 (1.57)	4.87 (1.06)	2.58*
Leisure	2.02 (0.74)	2.00 (0.67)	-0.12
Peers	2.58 (1.26)	3.48 (0.85)	3.28*
Substance Abuse	1.30 (1.49)	2.04 (1.55)	2.20*
Family	3.58 (1.62)	3.91 (1.47)	0.92
Attitudes	1.09 (1.24)	1.35 (1.15)	0.92
Personality	3.50 (1.61)	3.91 (1.68)	1.15

* $p < .05$

Overall, the unmatched sample of group home girls were higher risk as determined by the YLS/CMI and received more pre-treatment services than the matched sample of group home girls. Although dropping these girls from the sample limits the generalizability of the analyses, since the sample does not include the complete number of girls who received the program, a more rigorous test of program effectiveness can be conducted by guaranteeing the matched samples are as similar as possible on all of the observed covariates included in the model.

As seen in Table 11, the propensity score matching procedure created balanced mean scores across every observed covariate utilizing paired-sample analyses. The matching procedure was successful given that the mean probability score for the group home girls ($M = .37$, $SD = .22$) and the non-group home girls ($M = .36$, $SD = .21$) were not significantly different $t(148) = -.414$, $p = .679$. As well, girls' total YLS/CMI scores were equivalent for the group home girls (M

= 18.46, $SD = 5.93$) and the non-group home girls ($M = 18.87$, $SD = 6.33$), $t(148) = -.576$, $p = .565$. Therefore, there is reason to believe that any selection effects due to age, race/ethnicity, juvenile court unit, criminogenic risk scores, and pre-treatment programming were controlled.

Table 11. Descriptive Statistics of Matched Samples of Girls: Model B

Covariates	<i>M (SD)</i>		Significance Test (<i>t or X²</i>)
	Group Homes (n = 149)	Standard Probation Services (n = 149)	
% Non-White	63.1%	61.7%	0.06
% Delinquency	49.7%	53.0%	0.34
% Residential	28.2%	28.2%	0.00
% IPS	26.2%	30.2%	0.59
% FSS	18.8%	18.1%	0.02
% Alternative School	5.4%	4.7%	0.07
% Evening Reporting	11.4%	10.7%	0.03
Age	14.13 (1.23)	14.23 (1.24)	0.70
Prior Offenses	0.38 (0.90)	0.35 (0.85)	-0.33
Education	3.99 (1.57)	4.15 (1.58)	0.88
Leisure	2.02 (0.74)	2.01 (0.84)	-0.15
Peers	2.58 (1.26)	2.55 (1.16)	-0.24
Substance Abuse	1.30 (1.49)	1.32 (1.55)	0.08
Family	3.58 (1.62)	3.57 (1.67)	-0.07
Attitudes	1.09 (1.24)	1.21 (1.36)	0.80
Personality	3.50 (1.61)	3.70 (1.82)	1.05

Next, one- and two-year recidivism rates were examined if there were differences between the two groups.

Research Question 3: *When matched based on propensity scores, are there differences in one- and two-year recidivism rates between girls receiving group home treatment and girls who received standard probation services?*

Table 12 presents the one-year recidivism rates for the group home girls and the matched standard probation services. The sample of girls who received group home treatment had a one-year recidivism rate of 22.8% following program exit. The matched sample of girls who did not

receive group home treatment had a one-year recidivism rate of 24.2%. A McNemar chi-square test did not reveal a statistically significant difference between the paired samples, $X^2 (1, N = 149), p = 0.89$ (exact p value).

Table 12. One-Year Recidivism Rates: Model B

	Recidivists n (%)	Non-Recidivists n (%)	Total
Group Homes	34 (22.8%)	115 (77.2%)	149
Standard Probation Services	36 (24.2%)	113 (75.8%)	149

Table 13 presents the two-year recidivism rates for the group home girls compared to standard probation services girls. The sample of girls who received group home treatment had a two-year recidivism rate of 28.9% following program exit. The matched sample of girls who did not receive group home treatment had a two-year recidivism rate of 32.9%. A McNemar chi-square test did not reveal a difference between the paired samples, $X^2 (1, N = 149), p = 0.52$ (exact p value).

Table 13. Two-Year Recidivism Rates: Model B

	Recidivists n (%)	Non-Recidivists n (%)	Total
Group Homes	43 (28.9%)	106 (71.1%)	149
Standard Probation Services	49 (32.9%)	100 (67.1%)	149

In sum, the propensity score matching procedure created equivalent groups of girls based on theoretically salient pre-treatment covariates. The results from Model A indicated that the girls' group home significantly reduced recidivism at two-year follow compared to the control group. However, the results for Model B indicated that the girls' group homes do not significantly reduce recidivism at one- and two-year follow-up compared to the control group.

CHAPTER 5: DISCUSSION

This study used a quasi-experimental propensity score matching design to examine the effects of a group home intervention for adjudicated girls in the juvenile justice system. The study used two propensity score matching models to examine the effectiveness of a gender-responsive court intervention. These models differed in (1) the number of covariates included in the matching procedure and (2) the follow-up time periods for recidivism. The first research question asked if there were differences in demographic and criminogenic risk factors between girls who received group home treatment and girls who received standard probation services. The second research question asked if there were any remaining differences in demographic and criminogenic risk factors between the two groups after matching girls on propensity scores. The third research question asked if there were differences in the recidivism rates between girls who received group home treatment and girls who received standard probation services when matched on those same demographic and criminogenic risk factors using a propensity score matching procedure. These questions were critical in examining if group homes for girls were effective in reducing recidivism. The results from this study are important from both substantive and methodological perspectives.

Summary and Discussion of Results

The juvenile court uses the risk-needs-responsivity (RNR) model to identify areas of criminogenic risk and need and respond accordingly to those factors. Therefore, the higher the risk of the youth on their YLS/CMI domains, the more services the youth needs. The court system responded in a way that addressed this assumption in that the group home girls had higher risk scores and needs and subsequently received more intensive treatment than standard probation had to offer. It was to be expected that girls who did not receive a group home

placement would not present with the same level and type of needs than the group home. The results from both Model A and Model B signified that there were statistically significant differences between these two groups on a number of variables. Girls referred to the group homes had significantly higher YLS/CMI criminogenic risk scores. Model B revealed that group home girls received more court-supervised programming than girls not referred to the group home.

Model A Discussion

The results from Model A indicated that the group home intervention did not significantly reduce one-year recidivism rates group home girls compared to the control group (22.5% and 30.8%, respectively), but did significantly reduce the two-year recidivism rates of group home girls (28.4%) compared to the girls in the control group (42.0%). These findings provide some evidence for the effectiveness of the girls' group home intervention in decreasing future offending. In particular, by targeting girls' gender-specific needs in the group homes—such as offering intensive family and relational support services, a focus on mental health and self-esteem, substance abuse, all in community-based context—girls fared better in terms of reducing recidivism. A careful review of the group home program elements suggested that the girls' group home addressed the two broad areas of interest in gender responsive programming: girls' pathways into the system (e.g., implementing trauma-informed care) and girls' relationships (e.g., addressing and promoting healthy relationships with peers, romantic partners, and family members).

The theories that underpin gender-responsive intervention would suggest that girls who receive relationship-based and trauma-informed care would have better outcomes (e.g., lower recidivism) than girls who do not receive these types of services. More research is needed on

teasing apart the effectiveness of gender-responsive intervention and in what ways gender-responsivity is different from (or adds something in addition to) quality services for all youth in the system. However, this provides additional empirical evidence that girls presenting with gender-based needs and risk factors may need different types of resources over and above standard best practices (see Day et al., 2014).

However, there are appropriate cautions to interpreting the results from Model A. For example, there is a lack of data on which specific factors may have led to girls' placement in the group homes (e.g., if those were indeed gender driven based on girls' needs, if they were driven by offense type, by criminogenic risks and needs indicated by the YLS/CMI, or randomly by court decision-makers). However, there is likely good reason to believe that these decisions were informed by several of these processes. There is also a stark lack of gender-specific services both locally and nationally so it might be fair to say that program placement decisions in this context may be made by factors external to and/or in addition to factors identified by criminogenic risk/needs assessments.

There are also a number of limitations to this propensity score model. The two most notable concerns are related to the recidivism follow-up period for the comparison group and the heterogeneity of services girls may have received while involved with the court. The groups in Model A were compared on exit from program for the treatment group and entry into the court system for the control group. The treatment group, upon exit from the group home, was slightly older than the control girls and thus had less time at risk to re-offend as a juvenile (recall that status offenses are not offenses once the age of 17 is reached) which could partially account for the lower recidivism rates in the treatment group.

Additionally, this propensity score model did not incorporate any other programming that the girls received prior to group home placement. Without a program entry date for the control group it was not possible to match on group home girls to control girls on pre-treatment programming variables. It is likely that girls who received group home treatment also received other programming while involved with the court and the other programming could have impacted the reduction in recidivism among the treatment group. To test for these potential issues, a second propensity score matched model was developed that included pre-treatment program variables and a hypothetical equivalent exit date for the control group from which to measure recidivism.

Model B Discussion

The results of Model B indicated that the group home intervention did not significantly reduce the re-offense rates of girls at one- and two-year follow-up. After matching girls on demographic variables, criminogenic risk factors, and pre-treatment programming variables, girls who received group home treatment had a two-year recidivism rate of 28.9% following program exit. The control group had a similar recidivism rate of 32.9% at two-year follow-up. Contrary to the findings in Model A, these results suggested that the girls' group home intervention produced lower levels of recidivism, but not significantly so.

However, these results should also be interpreted with caution. The non-significant difference in recidivism rates does not necessarily provide evidence for the ineffectiveness of the girls' group homes. These results could be due to a number of effects including: (1) variation in time frames, (2) over-programming girls, and (3) unobserved variables. The implications of each of these effects are discussed in the following sections.

Time Frames

After calculating an equivalent exit time metric and not detecting differences in re-offense outcome, matching on age at exit was examined. That is, girls' age at program exit (calculated in days) was propensity score matched to control girls' age at entry, instead of using a mock program exit date for the control group. This matched analysis controlled for the girls' time at risk under court supervision. For example, younger girls have a higher probability of being on probation for longer periods of time by virtue of their age at entry into the court, which increases their risk for recidivism. The propensity score matching provided a similar loss of sample size to the original analysis resulting in only matching 149 pairs of girls. Matching on age at exit in addition to the other 15 covariates used in the Model B slightly altered the recidivism rates. For example, the age-matched sample of girls who received group home treatment had a two-year recidivism rate of 30.2% following program exit. The age-matched sample of girls who did not receive group home treatment had a two-year recidivism rate of 36.2%. A McNemar chi-square test did not reveal a difference between the paired samples, $\chi^2(1, N = 149), p = 0.321$.

Since the data available did not include a probation exit date for the control group, an approximate exit date had to be created based on a time frame similar to girls who received treatment. There are a number of issues with analyzing recidivism data without absolute exit dates from court supervision for both the control and treatment groups. First, the same time frame was applied to each girl in the control group based on her initial YLS/CMI date and thus does not capture any potential variation in non-group home girls' length of time under court supervision. For example, if a girl in the control group was dismissed from court three months after her initial YLS/CMI assessment, but her proximal measure of exit was over one year after

her initial assessment, there are nine months of potential new petitions to court that are not accounted for in the recidivism variable. Propensity score matching likely created two groups that had a similar length of time under court supervision, however, without absolute exit dates for both groups this cannot be measured.

Second, given that most group home girls continue to receive some type of court service, exit from the group home is not the same as exit from court supervision (see Table 4). For example, 64.5% of the group home girls received some type of court service after exit from the group home compared to only 10.8% of the non-group home girls from their approximated program exit date. While aftercare and continued support services are important post-release from an out-of-home placement, this additional post-intervention treatment and supervision could be inflating the recidivism rate of the treatment group due to additional supervision by court practitioners. Therefore, there is a need for future research to examine recidivism from absolute exit dates from court for all of the girls.

Over-Programming

There appears to be an overall pattern of over-programming girls, which could conflate the results. For example, 74.5% of the total sample of girls on probation did not receive any of the court run services, yet there was significant overlap in the number of services provided for girls who received any service. In particular, 22 out of the 172 (12.8%) of the girls receiving group home placement only received the group home. In other words, 87.2% of girls who were placed in the group home also received one additional court-run service. Furthermore, 58.7% of group home girls received two or more additional court-run programs. In contrast, the full sample of probation girls who were not placed at the group, only 25.5% received at least one court-run service, and 11.0% of the full sample of probation girls received two or more court-run

programs. This finding is important for a number of reasons. First, the amount of programs that girls received varied (and many times overlapped) across each measured time frame (i.e., pre-treatment, during treatment, post-treatment). This variation and overlap makes discerning any potential cause-and-effect relationships in a quasi-experimental design a challenge.

While group home girls received gender-responsive treatment, most of them also received other types of services while under court supervision that do not have gender-responsive elements (e.g., the court-run evening reporting program). By mixing girls into a variety of gender-neutral and gender-responsive treatment, detecting a true effect of either of the treatment modalities is difficult. This mixing could be influencing why girls in the propensity score matched groups did not differ on recidivism in the follow-up period. In addition, the effect of other programs girls received post-release from the group home is important to explore. Given that 64.5% of group home girls received or were receiving some type of service after the intervention may have influenced or lessened the effect of the treatment on re-offense.

Unobserved Variables

Many of the covariates in this study are correlated with recidivism (e.g., criminogenic risk factors measured by the YLS/CMI) and propensity score matching provides a method to balance any observable and measurable confounding characteristics. However, a caveat to using propensity score matching techniques is that balance can only be achieved on observed and measured characteristics (Rosenbaum & Rubin, 1985). Thus, propensity score matching cannot control for bias in unobserved covariates. The inability to control for bias in unobserved variables was relevant to the analyses and results for both Model A and Model B. Pertinent to this study, there are a number of potentially important pre-treatment variables that were not measured. For example, salient risk factors for girls' involvement with the juvenile justice

system such as previous traumatic experiences and running away were not tracked or measured. In addition to analyzing alternative pre-treatment covariates, some level of measurement of key variables in the intervention is needed. There were no variables to measure the group home services for the actual presence or absence of gender-responsive intervention at the individual level or the group level.

Further, studies on the validity of the YLS/CMI predicting recidivism for girls have produced varied results (Betchel et al., 2007; Flores et al., 2003; Onifade et al., 2008; Schmidt et al., 2011). They have also been critiqued for a lack of inclusion of gender-sensitive measures and over-predicting girls' risk for recidivism. Preliminary research has been conducted on the development and validation of gender-responsive assessments for women (see Salisbury et al., 2009; Van Voorhis et al., 2010). However, there are currently no validated gender-responsive assessments specifically for girls involved with the juvenile justice system.

Future research should focus on the development, implementation, and validation of gender-responsive assessments for female juvenile offenders. In the meantime, juvenile courts may consider assessing girls using mental health or trauma screening instruments. For example, the Massachusetts Youth Screening Instrument (MAYSI-2) is a mental health screening instrument commonly used in juvenile courts that includes a trauma subscale. Other researchers who have employed quasi-experimental designs to evaluate gender-responsive programming have examined these covariates for girls rather than the standard criminogenic risk factors (e.g., Day et al., 2014). The lack of measuring, and therefore not including, potential gender-responsive covariates in the model could contribute to the null effect of the intervention after matching the groups.

Methodological Implications of Propensity Score Matching

There is a growing body of literature on utilizing propensity score matching for quasi-experimental program evaluations. However, the literature is largely silent around the issue of equivalent time frames with matched comparison groups. Most program evaluation studies either have absolute dates for the two groups they are comparing or use propensity scores to compare subgroups of people based on gender, race/ethnicity, program type, etc. (see Baglivio et al., 2014 and Day et al., 2014 for examples). Therefore, there is currently no gold standard for how to measure equivalent periods of time between a group that received a service and has an end date to a group that did not receive the service and has no end date. In Model A it was important to consider recidivism from the initial YLS/CMI assessment because there is likely a large amount of variability in the length of time involved with the court for the control girls. For example, the majority of the full sample of control girls came from truancy court. Many of these girls may have only been under court supervision for a month or two while involved in truancy court and therefore creating a proxy edit date over a year later for them may misspecify their actual time at risk for reoffending by not capturing any new petitions to court during the first year after receiving their YLS/CMI assessment.

There is a need to build a stronger literature for addressing time frames in propensity score matching program evaluations. While these models used the same matching algorithm, they did vary in the assessed follow-up time frames. Arguments can be made for the use of both of these models as each one has its own set of strengths and weaknesses. For example, Model A retained most of the sample during the matching procedure, but did not match on the groups on an equivalent follow-up period. Model B controlled for more covariates and had a follow-up date for the control group recidivism, but the dates that were used could be an arbitrary cut off and

may not fully capture the extent of possible re-offense events among the girls in the control group. Furthermore, researchers need to state more explicitly their time points for follow-up. Most studies either do not mention the exact point in time recidivism was measured or use program completion dates (see Zahn et al., 2009). However, there are no set standards for how to address the lack of a program completion date for youth who did not receiving programming in a comparison group.

Finally, the issue of unobserved variables is present in both models. Propensity score estimates depend on meeting the ignorability assumption. That is, the available covariates are able to accurately describe the selection process and the same set of criteria is available for the comparison group. If relevant covariates are missing for either group, the propensity score models may yield inaccurate estimates (see Shadish, 2013 for a discussion on this topic in criminal justice program evaluation).

Limitations

There are a number of limitations to this study. First, given the context of the research setting, it was not feasible to randomize youth for a true experiment, so the only option was to utilize a quasi-experimental design. Since the study did not have the advantage of random assignment, over or underestimation of the outcome likely not systematic. While, PSM addresses a number of threats to internal validity (e.g., selection bias based on the measured covariates), this does not preclude other potential confounds such as regression to the mean.

Further, there is a lack of observed, systematic data on the specific factors and decision-making processes for placement in the group homes (e.g., if those were indeed gender driven based on girls' needs, if they were driven by offense type, by criminogenic risks and needs indicated by the YLS/CMI, or randomly by court decision-makers). However, there is likely

good reason to believe that these decisions were informed by multiple of these processes. There is also a stark lack of gender-responsive services both locally and nationally so it might be fair to say that program placement decisions in this context may be influenced by factors external or in addition to factors identified by criminogenic risk/needs assessments.

PSM Model Specification and Sample Reduction

There are a number of limitations to propensity score matching. First, propensity score matching techniques are only able to balance groups based on the observed, and thus measured, heterogeneity. As discussed there were a number of potential important variables (e.g., trauma history) that were not observed and thus not included in the models. Second, there are number of advantages and limitations to selecting specific parameters to estimate propensity scores and the appropriate matching algorithm. This study used a restrictive matching algorithm by employing a 1:1 pairwise nearest neighbor matching without replacement technique. That is, each comparison group girl could only be included as a matched case once. Dehejia and Wahba (2002) described that matching without replacement may not be adequate if there is little overlap of the propensity scores. While the current matching algorithm was able to match the majority of girls, a number of girls were dropped in the process of using a 1:1 method instead of a 1:k or sampling with replacement method.

Selecting a method of matching comes with needing to address trade-offs between variance and bias. For example, utilizing a nearest neighbor matching technique with a single neighbor and a caliper restriction (e.g., $ldl < .25$) decreases bias in the final match, but may increase variance. However, utilizing a without replacement method for the control group decreases the variance in the final match, but may increase bias (see Caliendo & Kopeinig, 2005 for a full discussion of trade-offs between bias and variance). In sum, there is no “right” method,

rather it is important to weigh the strengths and limitations to each method and critically evaluate the impact those decisions have on the final model specification. Caliendo and Kopeinig (2005) suggest trying out a number of approaches to matching groups and if they results shift based on the matching technique it would be important to further investigate what might be causing the disparity in the results.

Finally, it is important to note that there were a relatively small proportion of non-group home girls that received any of the other programs as compared to girls who received group home treatment (refer back to Table 4 in the Methods chapter). This underrepresentation of non-group home girls in other programs is primarily driving the reduction in matching girls in Model B. Other studies have reported similar reductions in sample size due to PSM (Day et al., 2014), as previously discussed there were significant differences between the matched and unmatched girls. In particular, unmatched sample group home girls had higher criminogenic risk scores and received more programming than the matched sample of group home girls. Model A did not include the additional pre-treatment programming covariates and therefore was able to retain approximately 98% of the treatment sample after the matching procedure.

Defining Gender-Responsive Programming

There is no one clear or agreed upon definition of gender-responsivity. As reviewed in this study, there are a number of core elements related to gender-responsive programming content and programming context, but there are no set standards on the number of elements that comprise a gender-responsive program, the intensity of those elements, the appropriate dosage, or any type of standardized measure to determine if a program is gender-responsive and/or the extent to which it is gender responsive. This is both a limitation to the current study (e.g., an inability to measurably determine the degree of gender-responsiveness in the intervention) as

well as a recommended direction for future research (e.g., a need to develop and empirically evaluate such standards).

The “what works” literature on juvenile justice intervention has extensive information on optimal program dosage and the most effective program components (Lipsey, 2009). Hubbard and Matthews (2008) attempted to reconcile the gender-responsivity and “what works” literature, but since then few steps have been taken to do so. For example, researchers might develop specific criteria and validate measures to assess the extent to which juvenile justice programming for girls is gender-responsive. The current literature includes an extensive list of criteria of important factors related to content elements that programming should address as well as contextual factors related to the programming environment, staff, and general structure of services (refer back to Table 1 for list of criteria). However, there is no standardized way to determine (1) if a program is gender-responsive, or (2) the level or extent to which a program is gender-responsive.

The primary goals of court systems traditionally aim to reduce recidivism among juvenile offenders, and while the goals of gender-responsive intervention include that aim, they appear to be much broader and inclusive of additional outcomes (Hubbard & Matthews, 2008; Kerig & Schindler, 2013; Walker et al., 2012). In particular, by targeting girls’ specific needs in the group homes—such as offering intensive family and relational support services, a focus on mental health and self-esteem, and substance abuse within community-based context—girls may have not fared better in terms of a reduction in recidivism, but placement in the group home may have contributed to other positive outcomes that are not captured by solely examining recidivism rates. For example, the group home intervention may have increased the safety of girls while they were in placement and address previous traumas through the individual and group

therapeutic treatment provided in the program. Given that the main focus of the Girls Moving On curriculum is on building and maintaining healthy relationships, measuring outcomes related to perceived improvement in girls' relationships with their family, peers, and romantic partners is also important. Finally, while in the group home, girls receive support related to their education such as transportation to school, structured time to complete schoolwork, and tutoring.

Measuring the impact of the program on academic outcomes (e.g., increased school attendance during and after treatment, improvement in grades). Additional research is needed to tease apart the effectiveness of gender-responsive intervention and in what ways that gender-responsivity is different from—or adds something over and above—quality services for all youth in the system.

Recommendations for Future Research

Future research should examine the actual exit date from court. This will allow a more accurate representation of the control group's re-offense rate. For example, the control group should only include youth who have been dismissed from probation and recidivism should be measured in one- and two-year increments following that exact release date in order to get a true picture of the impact of court intervention. In addition, exit dates from the court will allow for examination of the lag time between girls' exit from the group home and their exit from court. The similar recidivism rates between the final treatment and control groups could be due to a lack of controlling for the time under court supervision (e.g., girls are still under court supervision after exit from the group home for an unknown period of time). This is an important consideration since youth appear to be at highest risk for re-offense while on probation (e.g., within their first two years under court jurisdiction) (see Barnes et al., 2015).

In addition to providing a more accurate representation of the impact of court intervention on girls' recidivism rates, obtaining exact exit dates from court would allow for

more advanced statistical techniques such as survival analysis, also referred to as event history analysis, to measure time to recidivism. Survival analysis is an appropriate method for analyzing time-to-event data (for recent examples in the juvenile justice literature see Asscher et al., 2014; Bright et al., 2014; Olver et al., 2012; Ryan et al., 2014). For example, a Cox proportional hazards regression model would allow for testing the differences in survival times (e.g., time between program exit and re-offense) between the treatment and control groups and provide both the likelihood of re-offense and the rate at which re-offense occurs. However, it was not feasible to perform a survival analysis with the current format of the data, because there are not absolute exit dates from which to measure risk to reoffend as there are for the treatment group.

Alternative outcomes and measures should also be considered in future research. Other studies examining gender-responsive programming have reported outcomes related to academic achievement, school attendance, improvement in interpersonal relationships, and other variations for measuring recidivism such as re-offense type (e.g., violent) or number of future petitions (Belgrave et al., 2002; Brown & Block, 2001; Burke et al., 2003; Davidson et al., 2011; Gordon, 2004; Irvine, 2005; Walsh et al., 2002; Weiss & Nicholson, 1998). Other outcomes of interest to examine include mental health needs, substance abuse, and contextual, community-level variables (e.g., access to community-based resources). For example, it would be of interest to examine if girls in the treatment group had fewer petitions to court after program exit than girls in the control group. The researchers evaluating Hawaii's Girls Court operationalized re-offending as number of future law violations and found that girls who received the treatment had significantly fewer violations than the matched comparison group (Davidson et al., 2011). Davidson and colleagues also examined number of days spent on the run and found that girls who received the treatment spent significantly fewer days on the run than the comparison group.

The study showed variation in the differences in the overall recidivism rates for each model and there could be differences in the number of new petitions and/or the severity of the offenses. For example, group home girl recidivists had an average of 1.58 ($SD = 1.10$) court petitions following their exit from placement. Unfortunately, due to the creation of a new time frame, it is not possible to calculate the exact number of petitions non-group girls received.

Subsequent analyses of different program combinations might also be explored. Since only 22 out of 172 girls who received the group home *only* received the group home, it is difficult to perform any further analyses on the effectiveness of the group home for the subpopulation of girls who only received that one program. However, it might be advantageous to examine specific combinations of additional services group home girls received (e.g., examining the effectiveness of the group home in addition to another program). Two of the other court-run programs, Intensive Probation Services and Family Support Services, appear to have some program elements that overlap with the gender-responsive criteria. For example, Family Support Services provides two juvenile court officers to the family—one for the girl and one for the parents—who work together to help improve family relationships and reduce stressors and barriers to a healthy relationships between the parent and juvenile. Intensive Probation Services includes additional counseling and more time and contact between the girl and her court officer. Gender-responsivity includes building a therapeutic alliance between the girl and her court officer (Matthews & Hubbard 2008). This additional contact can foster the development of trust and build a therapeutic alliance. As well, some of the residential facilities that girls' were placed in may have a similar impact. However, there is a limited number of girls across each of the various placements therefore sample size would limit these analyses.

Finally, future evaluation research on gender-responsivity should incorporate qualitative data. Exploring girls' narratives on their experiences while under court supervision and receiving programming could provide rich, detailed insight into the effectiveness of the program beyond measuring recidivism rates.

Conclusions

The results from the two propensity score models revealed different sets of results. Model A indicated that the group home reduced recidivism rates two years following release from the program. After adding additional covariates to the model and creating a hypothetical exit date for the control group, Model B did not reveal the same results. Model B indicated that there was not significant decrease in recidivism rates for girls who received the group home compared to girls who did not receive the group home treatment with similar criminogenic risk profiles and pre-treatment services.

The utility of gender-responsive intervention is still important to continue to monitor and investigate given the empirical literature on the topic is in its infancy. The findings from Model B were contrary to the assumptions that gender-responsive programs would decrease recidivism for girls, however, because this evaluation research is still in its infancy, there are not set standards or curricula considered “best practice” to evaluate. Therefore researchers and practitioners need to continually evaluate community-based programs for girls and use the evaluation findings to alter and improve those programs. There are also no set standards in the methodological literature on the best way to address the issues of variance in time frames when creating propensity score matched models.

This study adds to the body of literature on gender-responsive services – an area of much needed rigorous evaluative research in terms of what works for girls given the consistent call by

researchers for gender-responsive programming for female youth (Chesney-Lind et al., 2008; Kerig & Schindler, 2013; Zahn et al., 2009; Walker et al., 2012). In sum, the findings in this study add to the limited, yet growing, body of research on gender-responsive services in juvenile justice practice and highlights the importance of continuing to assess for gender-specific needs of youth involved with the court and evaluate the effectiveness of services that address these gender-based needs.

APPENDICES

Appendix A: Literature Review Table of Gender-Responsive Evaluation Studies

Table 14. Literature Review Table of Gender-Responsive Evaluation Studies

Study (Program)	Methods and Outcomes	Core Gender-Responsive Elements
Belgrave, 2002 (Project Naja)	<ul style="list-style-type: none"> Girls (n= 147) assigned to treatment or control group depending on which school they were recruited from—data were collected pre- and post-intervention Outcomes: At the end of the program girls' alcohol usage decreased and improved peer relationships and cultural values 	<ul style="list-style-type: none"> Culturally-responsive with a focus on ethnic identity and gender roles Self-esteem (e.g., physical appearance, happiness/satisfaction) Mental health (e.g., anxiety) Provides support among girls Develop girls' leadership skills Girls developing long-term supportive relationships with their female mentors
Brown & Block, 2001 (Project Chrysalis)	<ul style="list-style-type: none"> Girls recruited through advertisements and randomly assigned to the treatment or control group Surveyed at entry, exit, and one and two-year follow-up periods Outcomes: Short-term effects (program completion) included decreasing risky sexual behavior and long-term effects (one- and two-year follow-up) included decreasing substance use and risk for suicide 	<ul style="list-style-type: none"> Risky sexual behavior Substance abuse Suicidality Mental health needs Addresses trauma and abuse history Education about health (in general and specific to girls' sexuality) Provides support among girls Addresses sexism/racism and harassment

Table 14 (cont'd).

Burke et al., 2003 (Working to Insure and Nurture Girls Success)	<ul style="list-style-type: none">• Randomized controlled trial comparing program placement (n =171) to traditional probation services (n = 570)• Girls ages 12-18 with minimal involvement in the juvenile justice system• Increased protective factors and reduced risk factors among program participants and lower recidivism rates 6 months following program release (4% v. 6%), but higher recidivism rates 12 and 18 months following (15% v. 11% and 18% v. 15%)• Girls in the program were more likely to attend school during the program and receive fewer failing grades than the control group, but there no differences in the academic variables at the 6 month follow-up	<ul style="list-style-type: none">• Intensive probation with home and community-based services• Addresses safety• Communication and relationship with family• Truancy• Substance abuse• Education about health (in general and specific to girls' sexuality)• Job/career training and support• Wraparound service delivery
Davidson et al., 2011 (Honolulu's Girls' Court)	<ul style="list-style-type: none">• Five cohorts of girls in Girls' Court (n = 70) and a matched comparison group (n = 70) based on demographics, risk factors, and history of system involvement• Followed girls from program start until the 21st birthday (recidivism follow-up for the cohorts ranged from 1 year to 4.6 years)• Outcome: Girls court girls had significantly fewer law violations and significantly less days on the run than the comparison group	<ul style="list-style-type: none">• Intensive case management• Health and sexuality education• Family relationship building – parent-daughter therapy groups (parents must be involved in treatment)• Individualized therapy – particularly to address trauma• Life skills training

Table 14 (cont'd)

Day et al., 2014 (Gender-Responsive Residential Placements)	<ul style="list-style-type: none">• Residential facilities (eight total) in Connecticut• Propensity score matching girls (n = 73) in the gender-responsive facilities and girls in the traditional facilities (n = 73)• Dependent variable: Recidivism at 6 and 12 months after program release (measured from program release using event-history analyses)• Independent variable: Treatment condition• Matched variables: Prior offenses, prior detention, MAYSI-2 assessment variables, age, race/ethnicity	<ul style="list-style-type: none">• Daily motivation sessions• Self-esteem/self-affirming• Safety• Building positive relationships with other females (girls in treatment and staff)• Levels of privileges, problem behavior control (not identified as gender-responsive)
Gallagher, 2005 (Practice Academic Cultural Educational Center)	<ul style="list-style-type: none">• Pretest-posttest with no control group• Girls ages 12-18 primarily referred to the program through the juvenile court• Outcomes: Improved academically (e.g., moving up grad levels, obtaining GEDs, enrolling in college or obtaining employment)• Girls self-reported decreases in drug and alcohol use• Girls' recidivism rates were 9% one year following program completion	<ul style="list-style-type: none">• Day treatment programming focused on prevention and early intervention and completing high school• Individualized treatment plans• Utilize in-depth needs assessments• Involving family in the treatment• Home visits at least once each month
Gordon, 2004 (AMICUS Girls' Restorative Program)	<ul style="list-style-type: none">• Pretest-posttest with no control group• Chronic female juvenile offenders ages 14-21• Girls self-reported better understanding of their actions, increased compassion, remorse, self-awareness, motivation, maturity, optimism, and less destructive behavior• Girls self-reported improved family and peer relationship	<ul style="list-style-type: none">• Promotes accountability, inner change and healing through restorative justice circles

Table 14 (cont'd)

Houston, 2006 (Girls Empowered to Move Sucessfully)	<ul style="list-style-type: none">• Pretest-posttest with no control group• Juvenile justice involved girls ages 13-16• Outcomes: Over a one year period of time girls self-reported desired changes in antisocial behavior and social competencies	<ul style="list-style-type: none">• Mentorship-based program matching girls with adult female mentors• Mentors tracked the girls progress in terms of sentence completion• Girls developing long-term supportive relationships with their female mentors
Hueffner et al., n.d. (Girls and Boys Town USA)	<ul style="list-style-type: none">• Pretest-posttest with no control group• Female juveniles offenders ages 11-18 referred from court systems• Outcomes: Nearly two-thirds of girls were released to less restrictive environments; the study provides 4.5 year follow-up data in which 37.7% reoffended within that period of time	<ul style="list-style-type: none">• Short-term residential placement• Teaching-Family model – focused on life skills• Provides safety and structure
Irvine, 2005 (Girls Circle)	<ul style="list-style-type: none">• Pretest-posttest with no control group• Detained girls ages 9-18• Girls self-reported significant improvement in self-efficacy, body image, and social support	<ul style="list-style-type: none">• Weekly support groups focused on empathic communication as well as positive relationships, body image, skill-building for achieving goals, and self-efficacy• Relational-cultural empowerment model
Kirk & Griffith, 2004 (Holistic Enrichment for At-Risk Teens)	<ul style="list-style-type: none">• Pretest-posttest with a control group• Girls ages 12-18 in residential placements for substance related issues• Improved in use of and perceived social support, peer and family relationships, and education (e.g., more engaged in school and improved grades)	<ul style="list-style-type: none">• Feminist and relational theories underpin treatment• Substance abuse treatment• Therapeutic community environment• Focused on improving family relationships

Table 14 (cont'd)

Wolf et al., 2009 (Reaffirming Young Sisters' Excellence)	<ul style="list-style-type: none">• Randomized controlled trial comparing program placement (n = 249) to traditional probation services (n = 84)• Adjudicated girls, predominantly Black and age 11-17• No differences in recidivism rates at 6, 12, and 18 month follow-up; however, the treatment condition had a lower growth rate of recidivism between the 12 and 18 month follow-up (18% v. 39%) and had less serious re-offense charges• African American girls in the program reoffended at lower rates than all the other racial/ethnic groups in the program	<ul style="list-style-type: none">• Alternative probation services to address disproportionate minority contact• JCO home visits• Individualized case planning with a focus on culturally-appropriate services• Life skills training – leadership skills• Therapy• Teen pregnancy services and education around healthy sexuality• Job/career training and support
Walsh et al., 2002 (The Earls court Girls Connection)	<ul style="list-style-type: none">• Pretest-posttest with no control group• Girls entered the program after being referred by the schools or by their parents• Outcomes: At both 6- and 12- month follow-up girls' conduct disorder decreased and girls' relationships with others improved as reported by the girl and her caregiver	<ul style="list-style-type: none">• Addressing the mental health needs of girls• Improving relationships with peers, family, and school personnel• Teaching healthy sexuality• Leadership skills• Provides girl with a mentor to build a long-term relationships with
Weiss & Nicholson, 1998 (Friendly PEERsuasion – Girls Inc.)	<ul style="list-style-type: none">• Randomized controlled trial with 354 girls (ages 11 to 14) assessed at the beginning of the program and three times after receiving the program• Outcomes: Reduced incidence of drinking, taking substances, and being involved with substance-using peer groups; more effective for the younger program participants	<ul style="list-style-type: none">• Preventive focus on substance abuse and peer relationships• Empowering girls to choose healthy alternatives to substances• Focus on leadership skills

Table 14 (cont'd)

Williams et al., 2002 (Movimiento Ascendencia)	<ul style="list-style-type: none">• Program participants (n = 61) and comparison participants (n = 61) were gathered through the schools and the juvenile justice system ages 8-19• Program participants were randomly sampled and comparison sample used a snowball design for interviews/surveys at entry and exit• Outcomes: Greater reduction in delinquency and increase in grades among program participants at the 12-month follow-up	<ul style="list-style-type: none">• Provide positive alternatives to gang involvement and substance use• Providing a safe environment for girls• Self-defense skills• Focus on self-esteem• Culturally-based case management• Involving parents in the treatment• Pregnancy prevention• Addresses sexism/racism and harassment• Provides girl with a mentor to build a long-term relationships with• Education about health (in general and specific to girls' sexuality)• Job/career training and support
Wisconsin Department of Corrections, 2005 (Southern Oaks Girls School Stepping Up Program)	<ul style="list-style-type: none">• Pretest-posttest with a control group• Girls ages 13-19 with mental health needs• Fewer adult arrests (0.6 v 4.4 offenses/each) and transfers to state mental health institutions• Educational improvement (e.g., 11 of 23 girls completed their high school diplomas) among program participants	<ul style="list-style-type: none">• Correctional treatment facility designed to meet girls' individualized needs• Mental health focus• Development of independent living skills• Fostering healthy relationships

Appendix B: Youth Level of Service/Case Management Inventory (YLS/CMI)

Youth Level of Service/Case Management Inventory (YLS/CMI) Items by Subscale

-
1. **Prior/Current Offenses:** Three or More Prior Convictions
 2. **Prior/Current Offenses:** Two or More Failures to Comply
 3. **Prior/Current Offenses:** Prior Probation
 4. **Prior/Current Offenses:** Prior Custody
 5. **Prior/Current Offenses:** Three or More Current Convictions
 6. **Education:** Low Achievement
 7. **Education:** Problems with Teachers
 8. **Education:** Problems with Peers
 9. **Education:** Disruptive Classroom Behavior
 10. **Education:** Disruptive Behavior on School Property
 11. **Education:** Truancy
 12. **Leisure/Recreation:** Lack of Organized Activities
 13. **Leisure/Recreation:** Could Make Better Use of Time
 14. **Leisure/Recreation:** No Personal Interests
 15. **Peer Relations:** Lack of Positive Peer Acquaintances
 16. **Peer Relations:** Lack of Positive Friends
 17. **Peer Relations:** Some Delinquent Peer Acquaintances
 18. **Peer Relations:** Some Delinquent Friends
 19. **Substance Abuse:** Occasional Drug Use
 20. **Substance Abuse:** Chronic Drug Use
 21. **Substance Abuse:** Chronic Alcohol Use
 22. **Substance Abuse:** Substance Abuse Interferes with Life
 23. **Substance Abuse:** Substance Use Linked to Offense(s)
 24. **Family & Parenting:** Inadequate Supervision
 25. **Family & Parenting:** Difficulty in Controlling Behavior
 26. **Family & Parenting:** Inappropriate Discipline
 27. **Family & Parenting:** Inconsistent Parenting
 28. **Family & Parenting:** Poor Relations (Father-Youth)
 29. **Family & Parenting:** Poor Relations (Mother-Youth)
 30. **Attitudes & Orientation:** Not Seeking Help
 31. **Attitudes & Orientation:** Actively Rejecting Help
 32. **Attitudes & Orientation:** Defies Authority
 33. **Attitudes & Orientation:** Antisocial/Pro-criminal Attitudes
 34. **Attitudes & Orientation:** Callous, Little Concern for Others
 35. **Personality & Behavior:** Short Attention Span
 36. **Personality & Behavior:** Poor Frustration Tolerance
 37. **Personality & Behavior:** Verbally Aggressive/Verbally Intimidating
 38. **Personality & Behavior:** Explosive Episodes
 39. **Personality & Behavior:** Physically Aggressive
 40. **Personality & Behavior:** Inadequate Guilt Feelings
 41. **Personality & Behavior:** Inflated Self-Esteem
 42. *Unemployment/Not Looking for Work
-

*Note: The variable Unemployment/Not looking for Work was omitted from the measure. This item was not relevant to this sample due to age and had no variation.

REFERENCES

REFERENCES

- Acoca, L. (1998). Outside/inside: The violation of American girls at home, on the streets and in the juvenile justice system. *Crime & Delinquency*.
- Acoca, L., & Dedel, K. (1998). *No place to hide: Understanding and meeting the needs of girls in the California juvenile justice system*. San Francisco: National Council on Crime and Delinquency.
- American Bar Association and National Bar Association (2001). *Justice by Gender: The Lack of Appropriate Prevention, Diversion, and Treatment Alternatives for Girls in the Justice System*. Washington, DC: American Bar Association and National Bar Association.
- Anderson, V. R., & Davidson, W. S. (2015). *Gender-responsivity in the juvenile justice system*. Manuscript in preparation.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) Model: Does Adding the Good Lives Model Contribute to Effective Crime Prevention? *Criminal Justice and Behavior*, 38(7), 735-755.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28(3), 369-404.
- Asscher, J. J., Deković, M., Manders, W., van der Laan, P. H., Prins, P. J., van Arum, S., & Dutch MST Cost-Effectiveness Study Group. (2014). Sustainability of the effects of multisystemic therapy for juvenile delinquents in The Netherlands: effects on delinquency and recidivism. *Journal of Experimental Criminology*, 10(2), 227-243.
- Austin, P. C. (2011). Comparing paired vs. non-paired statistical methods of analyses when making inferences about absolute risk reductions in propensity-score matched samples. *Statistics in Medicine*, 30, 1292-1301.
- Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of multisystemic therapy and functional family therapy effectiveness: A multiyear statewide propensity score matching analysis of juvenile offenders. *Criminal Justice and Behavior*, 41(9), 1033-1056.
- Baines, M., & Alder, C. (1996). Are girls more difficult to work with? Youth workers' perspectives in juvenile justice and related areas. *Crime & Delinquency*, 42(3), 467-485.

- Barnes, A. R., Campbell, N., Anderson, V., Campbell, C., Onifade, E., & Davidson, W. (2015). Relative predictive validity of a post-court involvement assessment. Manuscript submitted for publication.
- Bechtel, K., Lowenkamp, C. T., & Latessa, E. (2007). Assessing the risk of re-offending for juvenile offenders using the Youth Level of Service/Case Management Inventory. *Journal of Offender Rehabilitation*, 45(3-4), 85-108.
- Belgrave, F. Z. (2002). Relational theory and cultural enhancement interventions for African American adolescent girls. *Public Health Reports*, 117, S76.
- Belknap, R. (2001). *The Invisible Woman: Gender, Crime and Justice*. Belmont, CA: Wadsworth Publishing 2001.
- Belknap, J., Gaarder, E., Holsinger, K., McDaniels-Wilson, C., & Cady, B. (2010). Using girls' voices and words to study their problems. In M. Kerr, H. Stattin, R. C. M. E. Engels, G. Overbeek, & A. Andershed (Eds.), *Understanding girls' problem behavior: How girls' delinquency develops in the context of maturity and health, co-occurring problems, and relationships* (pp. 92-116). West Sussex, UK: Wiley.
- Belknap, J. and Holsinger, K. (2006). The gendered nature of risk factors for delinquency. *Feminist Criminology*, 1(1), 48-71.
- Belknap, J., Holsinger, K., & Dunn, M. (1997). Understanding incarcerated girls: The results of a focus group study. *Prison Journal*, 77(4), 381-404.
- Bell, K., Terzian, M. A., & Moore, K. A. (2012). What works for male children and adolescents: Lessons from experimental evaluations of programs and interventions. Washington: Child Trends.
- Bloom, B., Owen, B., & Covington, S. (2004). Gender-responsive strategies: Research, practice, and guiding principles for women offenders (NIC No. 018017). Washington, DC: National Institute of Corrections.
- Bloom, B., Owen, B., Deschenes, E. P., & Rosenbaum, J. (2002). Moving toward justice for female juvenile offenders in the new millennium: Modeling gender-specific policies and programs. *Journal of Contemporary Criminal Justice*, 18, 37-56.
- Bloom, B., Owen B., Rosenbaum, J., & Deschenes. E. P. (2003). Focusing on girls and young women. *Women & Criminal Justice*, 14(2-3), 117-136.
- Bond-Maupin, L., Maupin, J. R., & Leisenring, A. (2002). Girls' delinquency and the justice implications of intake workers' perspectives. *Women & Criminal Justice*, 13(2-3), 51-77.

- Burke, C., & Pennell, S. (2001). *Breaking Cycles evaluation: A comprehensive approach to youthful offenders*. San Diego Association of Governments.
- Bright, C. L., Hurley, S., & Barth, R. P. (2014). Gender Differences in Outcomes of Juvenile Court-Involved Youth Following Intensive In-Home Services. *Journal of the Society for Social Work and Research*, 5(1), 23-44.
- Brown, K. J., & Block, A. J. (2001). Evaluation of Project Chrysalis: A school-based intervention to reduce negative consequences of abuse. *The Journal of Early Adolescence*, 21(3), 325-353.
- Brubaker, S. J., & Fox, K. C. (2010). Urban African American girls at risk: An exploratory study of service needs and provision. *Youth Violence and Juvenile Justice*, 8(3), 250-265.
- Burke, C., Keaton, S., & Pennell, S. (2003). *Addressing the Gender-specific Needs of Girls: An Evaluation of San Diego's WINGS Program: Board of Corrections Final Report*. SANDAG.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72.
- Cauffman, E., Farruggia, S. P., & Goldweber, A. (2008). Bad boys or poor parents: Relations to female juvenile delinquency. *Journal of Research on Adolescence*, 18(4), 699-712.
- Chesney-Lind, M. (2001). What about the girls? Delinquency programming as if gender mattered. *Corrections Today*, 38-45.
- Chesney-Lind, M., & Irwin, K. (2008). *Beyond bad girls: Gender, violence and hype*. New York: Routledge.
- Chesney-Lind, M. & Morash, M. (2011). *Feminist Theories of Crime*. Ashgate Publishing Group: United Kingdom.
- Chesney-Lind, M., Morash, M., & Stevens, T. (2008). Girls troubles, girls' delinquency, and gender responsive programming: A review. *Australian & New Zealand Journal of Criminology*, 41, 162-189.
- Chesney-Lind, M., & Shelden, R. G. (2004). *Girls, delinquency and juvenile justice*. (3rd ed.). Belmont, CA: Wadsworth.
- Comstock, D. L., Hammer, T. R., Strentzsch, J., Cannon, K., Parsons, J., & II, G. S. (2008). Relational-cultural theory: A framework for bridging relational, multicultural, and social justice competencies. *Journal of Counseling & Development*, 86(3), 279-287.

- Covington, S. (2000). Helping women recover: Gender-specific treatment for substance abuse in community corrections. In M. McMahon (Ed.), *Assessment to Assistance: Programs for Women in Community Corrections* (pp. 139-170). Lanham, MD: American Correctional Association.
- Covington, S. S. (2008). Women and addiction: A trauma-informed approach. *Journal of Psychoactive Drugs*, 40(5), 377-385.
- Covington, S. S., Burke, C., Keaton, S., & Norcott, C. (2008). Evaluation of a trauma-informed and gender-responsive intervention for women in drug treatment. *Journal of psychoactive drugs*, 40(sup5), 387-398.
- CT Touchstone Program Description* (2001). North American Family Institute. Retrieved from http://legacy.nafi.com/program_detail.htm?id=8.
- Davis, C. P. (2007). At-risk girls and delinquency: Career pathways. *Crime and Delinquency*, 53(3), 408-35.
- Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *Review of Law and Women's Studies*, 2, 11-52.
- Daly, K., & Chesney-Lind, M. (1988). Feminism and criminology. *Justice Quarterly*, 5(4), 497-538.
- Daly, K., & Stubbs, J. (2006). Feminist engagement with restorative justice. *Theoretical Criminology*, 10(1), 9-28.
- Davidson, J. T., Pasko, L., & Chesney-Lind, M. (2011). "She's Way Too Good to Lose": An Evaluation of Honolulu's Girls Court. *Women & Criminal Justice*, 21(4), 308-327.
- Day, J. C., Zahn, M. A., & Tichavsky, L. P. (2014). What works for whom? The effects of gender responsive programming on girls and boys in secure detention. *Journal of Research in Crime and Delinquency*, 0022427814538033.
- Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, 84(1), 151-161.
- Department of Juvenile Services. (January 2012). *Recidivism Report*. Baltimore, MD: Maryland Department of Juvenile Services.
- Development Services Group, Inc.* (2010). *Group Homes*. Retrieved from http://www.ojjdp.gov/mpg/litreviews/Group_Homes.pdf.

- Fagan, A. A., & Lindsey, A. M. (2014). Gender differences in the effectiveness of delinquency prevention programs: What can be learned from experimental research? *Criminal Justice and Behavior*, 41(9), 1057-1078.
- Feld, B. C. (2009). Violent girls or relabeled status offenders? An alternative interpretation of the data. *Crime & Delinquency*, 55(2), 241-265.
- Fields, E., Farmer, E. M. Z., Apperson, J., Mustillo, S., & Simmers, D. (2006). Treatment and post treatment effects of a residential treatment using a re-education model. *Behavioral Disorders*, 31(3), 312-322.
- Flores, A. W., Travis, L. F., & Latessa, E. J. (2003). Case classification for juvenile corrections: An assessment of the Youth Level of Service/Case Management Inventory (YLS/CMI)(May). Cincinnati, OH: Center for Criminal Justice Research, Division of Criminal Justice, University of Cincinnati.
- Foley, A. (2008). The current state of sex-specific delinquency programming. *Journal of Criminal Justice*, 36, 262-269.
- Ford, J. D., Elhai, J. D., Connor, D. F., & Frueh, B. C. (2010). Poly-victimization and risk of posttraumatic, depressive, and substance use disorders and involvement in delinquency in a national sample of adolescents. *Journal of Adolescent Health*, 46(6), 545-552.
- Freitas, K. & Chesney-Lind, M. (2001). Difference doesn't mean difficult: Practitioners talk about working with girls. *Women, Girls & Criminal Justice*, 2(5), August/September.
- Gaarder, E., Rodriguez, N., & Zatz, M. (2004). Criers, liars, manipulators: Probation officers' views of girls. *Justice Quarterly*, 21, 547-578.
- Garcia, C. A., & Lane, J. (2012). Dealing with the fall-Out: Identifying and addressing the role that relationship strain plays in the lives of girls in the juvenile justice system. *Journal of Criminal Justice*, 40(3), 259-267.
- Garcia, C. A. & Lane, J. (2010). Looking in the Rearview Mirror: What Incarcerated Women Think Girls Need From The System. *Feminist Criminology*, 5(3) 227-243.
- Georgetown Center on Poverty, Inequality and Public Policy (2012). *Improving the Juvenile Justice System for Girls: Lessons from the States*. (1st ed.) Watson, L. & Edelman, P.
- Gilligan, C. (1982). *In a different voice*. Harvard University Press.
- Girls Incorporated. (1996). *Prevention and parity: Girls in juvenile justice*. Indianapolis, IN: Girls Incorporated National Resource Center.

- Goodkind, S. (2005). Gender-specific services in the juvenile justice system: A critical examination. *Affilia*, 20, 52–70.
- Gordon, K. G. (2004). *From Corrections to Connections: A Report on the AMICUS Girls Restorative Program: a Gender-specific Restorative Practices Program for Serious and Chronic Juvenile Female Offenders*. AMICUS, Incorporated.
- Greene, Peters, & Associates. (1997). *Guiding principles for promising female programming: An inventory of best practices*. Nashville, TN: Office of Juvenile Justice and Delinquency Prevention.
- Guo, S., & Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications*, 11, Sage Publications.
- Hansen, B. B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science*, 23, 219-236.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15, 199-236.
- Hoge, R. D. (2001). *The juvenile offender* (Vol. 5). Springer Science & Business Media.
- Holsinger, K. (2000). Feminist perspectives on female offending: Examining real girls' lives. *Women and Criminal Justice*, 12(1), 23-52.
- Hooper, S. R., Murphy, J., Devaney, A., & Hultman, T. (2000). Ecological outcomes of adolescents in a psychoeducational residential treatment facility. *American Journal of Orthopsychiatry*, 70(4), 491–500.
- Hoyt, S., & Scherer, D. G. (1998). Female juvenile delinquency: misunderstood by the juvenile justice system, neglected by social science. *Law and human behavior*, 22(1), 81.
- Hubbard, D. J., & Matthews, B. (2008). Reconciling the differences between the “gender-responsive” and the “what works” literatures to improve services for girls. *Crime and Delinquency*, 54(2), 225-258.
- Irvine, A. (2005). Girls Circle summary of outcomes for girls in the juvenile justice system. *Report submitted by Cere Policy Research of the Girls Circle Association, Santa Cruz, CA*.
- Javdani, S., Sadeh, N., & Verona, E. (2011). Gendered social forces: A review of the impact of institutionalized factors on women and girls’ criminal justice trajectories, *Psychology, Public Policy, and Law*, 17(2), 161-211.

- Jones, R. J., & Timbers, G. D. (2003). Minimizing the need for physical restraint and seclusion in residential youth care through skill-based treatment programming. *Families in Society*, 84(1), 21–29.
- Jordan, J. V. (Ed.). (2013). *The Power of Connection: Recent Developments in Relational-cultural Theory*. Routledge.
- Jordan, J. V., & Hartling, L. M. (2002). New developments in relational-cultural theory. In M. Ballou & L. S. Brown (Eds.), *Rethinking mental health and disorder* (pp. 48-70). New York, NY: Guilford.
- Kerig, P. K., & Schindler, S. R. (2013). Engendering the evidence base: A critical review of the conceptual and empirical foundations of gender-responsive interventions for girls' delinquency. *Laws*, 2(3), 244-282.
- Kirigin, K. A., Braukmann, C. J., Atwater, J. D., & Wolf, M. M. (1982). An evaluation of teaching-family (achievement place) group homes for juvenile offenders. *Journal of Applied Behavior Analysis*, 15(1), 1-16.
- Laidler, K. J., & Hunt, G. (2001). Accomplishing femininity among the girls in the gang. *British Journal of Criminology*, 41(4), 656-678.
- Larzelere, R. E., Daly, E. L., Davis, J. L., Chmelka, M. B., & Handwerk, M. L. (2004). Outcome evaluation of Girls and Boys Town's Family Home Program. *Education and Treatment of Children*, 27(2), 130–149.
- Leeman, L. W., Gibbs, J. C., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. *Aggressive Behavior*, 19, 281–292.
- Leichtman, M. (2006).
- Leitz, L. (2003). Girl fights: exploring females' resistance to educational structures. *International Journal of Sociology and Social Policy*, 23(11), 15-46.
- Lewis, R. E. (2005). The effectiveness of Families First services: An experimental study. *Children and Youth Services Review*, 27, 499-509.
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders*, 4(2), 124-147.
- Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annu. Rev. Law Soc. Sci.*, 3, 297-320.
- Lipsey, M. W., & Howell, J. C. (2012). Research-based guidelines for juvenile justice programs. *Justice Research and Policy*, 14(1), 17-34.

- Lipsey, M. W., Howell, J., Kelly, M., Chapman, G., & Carver, D. (2010). Improving the effectiveness of juvenile justice programs: A new perspective on evidence-based practice. Georgetown University. Washington DC: Center for Juvenile Justice Reform.
- Mallicoat, S. L. (2007). Gendered justice: Attributional differences between males and females in the juvenile courts. *Feminist Criminology*, 2(1), 4-30.
- Matthews, B., & Hubbard, D. J. (2008). Moving ahead: Five essential elements for working effectively with girls. *Journal of Criminal Justice*, 36, 494-502.
- McCurdy, B. L., & McIntyre, E. K. (2004). And what about residential ...? Re- conceptualizing residential treatment as a stop-gap service for youth with emotional and behavioral disorders. *Behavioral Interventions*, 19, 137-158.
- Messina, N., Grella, C. E., Cartier, J., & Torres, S. (2010). A randomized experimental study of gender-responsive substance abuse treatment for women in prison. *Journal of substance abuse treatment*, 38(2), 97-107.
- Miller, J. B. (1976). *Toward a new psychology of women*. Boston, MA: Beacon Press.
- Morash, M. (2010). *Women on probation and parole: A feminist critique of community programs and services*. Boston, MA: Northeastern University Press.
- Morgan, M., & Patton, P. (2002). Gender-responsive programming in the justice system- Oregon's guidelines for effective programming for girls. *Fed. Probation*, 66, 57.
- Olver, M. E., Stockdale, K. C., & Wong, S. C. (2012). Short and long-term prediction of recidivism using the youth level of service/case management inventory in a sample of serious young offenders. *Law and human behavior*, 36(4), 331.
- Onifade, E., Davidson, W., Campbell, C., Turke, G., Malinowski, J., & Turner, K. (2008). Predicting recidivism in probationers with the youth level of service case management inventory (YLS/CMI). *Criminal Justice and Behavior*, 35(4), 474-483.
- Onifade, E., Nyandoro, A. S., Davidson, W. S., & Campbell, C. (2009). Truancy and patterns of criminogenic risk in a young offender population. *Youth violence and juvenile justice*, 8(1), 3-18.
- Orbis Partners (2014). *Girls...Moving On: A Program for Criminal Justice Involved Girls*. Retrieved from <http://orbispartners.com/wp-content/uploads/2014/07/GirlsMovingOn-DataSheet.pdf>.
- Owen, B., & Bloom, B. (1998). *Modeling gender specific services in juvenile justice: Final report to the Office of Criminal Justice Planning*. Sacramento, CA: OCJP.

- Pasko, L. (2006). *The female juvenile offender in Hawaii: Understanding gender differences in arrests, adjudications, and social characteristics of juvenile offenders*. Research and Statistics Branch, Crime Prevention and Justice Assistance Division, Department of the Attorney General.
- Puzzanchera, C., Adams, B., & Sickmund, M. (2010). Juvenile court statistics 2006-2007. *National Center for Juvenile Justice*.
- Ravoira, L., Graziano, J., & Patino-Lydia, V. (2012). Urgent Work: Developing a Gender-Responsive Approach for Girls in the Juvenile Justice System. *Universitas Psychologica*, 11(4), 1181.
- Recidivism Among Juvenile Delinquents and Offenders Released from Residential Care in 2008* (October 2011). Office of Child and Family Services: New York. Retrieved from http://www.ocfs.state.ny.us/main/detention_reform/Recidivism%20fact%20sheetpdf.
- Rivard, J. C., Bloom, S. L., McCorkle, D., & Abramovitz, R. (2005). Preliminary results of a study examining the implementation and effects of a trauma recovery framework for youths in residential treatment. *Therapeutic Community*, 26(1), 83–96.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33-38.
- Ryan, J. P., Abrams, L. S., & Huang, H. (2014). First-Time Violent Juvenile Offenders: Probation, Placement, and Recidivism. *Social Work Research*, svu004.
- Ryan, J. P., Marshall, J. M., Herz, D., & Hernandez, P. M. (2008). Juvenile delinquency in child welfare: Investigating group home effects. *Children and Youth Services Review*, 30(9), 1088-1099. doi:10.1016/j.childyouth.2008.02.004
- Salisbury, E. J., Van Voorhis, P., Spiropoulos, G. V. (2009). The predictive validity of a gender-responsive needs assessment: An exploratory study. *Crime & Delinquency*, 55(4), 550-585.
- Schaffner, L. (2007). *Girls in trouble with the law*. New Brunswick, NJ: Rutgers University Press.
- Schmidt, F., Campbell, M. A., & Houlding, C. (2011). Comparative analyses of the YLS/CMI, SAVRY, and PCL: YV in adolescent offenders: A 10-year follow-up into adulthood. *Youth Violence and Juvenile Justice*, 9(1), 23-42.
- Schram, P. J., Koons-Witt, B. A., & Morash, M. (2004). Management strategies when working with female prisoners. *Women & Criminal Justice*, 15:2, 25-50.

- Schwartz, J., & Steffensmeier, D. (2012). Stability and change in girls' delinquency and the gender gap: Trends in violence and alcohol offending across multiple sources of evidence. In *Delinquent Girls* (pp. 3-23). Springer New York.
- Shadish, W. R. (2013). Propensity score analysis: Promise, reality and irrational exuberance. *Journal of Experimental Criminology*, 9(2), 129-144.
- Smith, P. & Smith, W. A. (2005). Experiencing community through the eyes of young female offenders. *Journal of Contemporary Criminal Justice*, 21(4), 364-385.
- Smith, J. A., & Todd, P. E. (2001). Reconciling conflicting evidence on the performance of propensity-score matching methods. *American Economic Review*, 112-118.
- Snyder, H., & Sickmund, M. (1999). *Juvenile offenders and victims: 1999 National Report*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention, US Department of Justice
- Stevens, T., Morash, M., & Chesney-Lind, M. (2011). Are girls getting tougher, or are we getting tougher on girls? Probability of arrest and juvenile court oversight in 1980 and 2000. *Justice Quarterly*, 28(5), 718-744.
- Strom, K., J., Warner, T. D., Tichavsky, L., Zahn, M. A. (2010). Policing juveniles: Domestic violence arrest policies, gender, and police response to child-parent violence. *Crime and Delinquency*, doi: 10.1177/0011128710376293
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25, 1-21.
- Teplin, L. A., Abram, K. M., McClelland, G. M., Dulcan, M. K., & Mericle, A. A. (2002). Psychiatric disorders in youth in juvenile detention. *Archives of General Psychiatry*, 59(12), 1133-1143.
- Thoemmes, F., & Kim, E. S. (2011). A Systematic Review of Propensity Score Methods in the Social Sciences. *Multivariate Behavioral Research*, 46, 90-118.
- Valentine Foundation and Woman's Way. (1990). *A conversation about girls* (pamphlet). Byrn Mawr, PA: Valentine Foundation.
- Voorhis, P. V. (2012). On behalf of women offenders. *Criminology & Public Policy*, 11(2), 111-145.
- Van Voorhis, P., Wright, E., Salisbury, E., & Bauman, A. (2010). Women's risk factors and their contributions to existing risk/needs assessment: The current status of a gender responsive supplement, *Criminal Justice Behavior*, 37, 261-288.

- Walker, S. C., Munro, A., & Sullivan-Colglazier, C. (2012). Principles in practice: A multistate study of gender-responsive reforms in the juvenile justice system. *Crime & Delinquency*, doi: 10.1177/0011128712449712.
- Walsh, M.M., Pepler, D.J., & Levene, K.S. (2002). A model intervention for girls with disruptive behaviour problems: The Earls court Girls Connection. *Canadian Journal of Counseling*, 36(4):297–311.
- Watson, L., & Edelman, P. (2012). Improving the juvenile justice system for girls: Lessons from the states. Washington, DC: Georgetown Center on Poverty, Inequality and Public Policy.
- Wattanaporn, K. A., & Holtfreter, K. (2014). The impact of feminist pathways research on gender-responsive policy and practice. *Feminist Criminology*, 1557085113519491.
- Weiss, F. L., & Nicholson, H. J. (1998). Friendly PEERsuasion against substance use: The Girls Incorporated model and evaluation. *Drugs & Society*, 12(1/2), 7–22.
- Williams, K., Curry, D., & Cohen, M. L. (2002). Gang prevention programs for female adolescents: An evaluation. In W.L. Reed & S.H. Decker (Eds), *Responding to gangs: Evaluation and research* (pp. 225–263). Washington, DC: US Department of Justice, National Institute of Justice.
- Zahn, M., Day, J. C., Mihalic, S. F., & Tichavsky, L. (2009). Determining what works for girls in the juvenile justice system: A summary of evaluation evidence. *Crime and Delinquency*, 55(2), 266–293.
- Zahn, M., Hawkins, S., Chiancone, J., & Whitworth, A. (2008). *The Girls Study Group—Charting the Way to Delinquency Prevention for Girls*. Bulletin. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- Zimring, F. E. (2002). The common thread: Diversion in the jurisprudence of juvenile courts. *A century of juvenile justice*, 142, 144–45.