Gender and Risk Assessment in Juvenile Offenders: A Meta-Analysis

by

Natasha Pusch

A Thesis Presented in Partial Fulfillment
Of the Requirements for the Degree
Master of Science

Approved March 2016 by the
Graduate Supervisory Committee:

Kristy Holtfreter, Chair
Kevin Wright
Cody Telep

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

Although young males are still the primary perpetrators of juvenile crime, girls are increasingly coming into contact with the criminal justice system. While girls may have different pathways to crime and risks for recidivism than boys, their risk to reoffend is typically assessed using a gender-neutral tool that is based on social learning theory: a theory originally developed and tested on males. With the appropriateness of using gender-neutral tools to assess female criminality coming into question, a number of researchers have searched for a resolution. To date, mixed findings on the predictive validity of risk assessment tools have not provided any definitive answers. To help assess the predictive validity of the Youth Level of Service Inventory, separate meta-analyses were conducted for male and female juvenile offenders using previous studies. The mean effect sizes were compared in order to determine whether the predictive validity is similar for both males and females. With the exception of violent recidivism, results indicate that the YLS/CMI works equally well for male and female offenders. The implications of these findings for theory, research, and correctional policy are discussed.

TABLE OF CONTENTS

Page

LIST OF TABLES

INTRODUCTION

For many, the picture that comes to mind when thinking of a "typical" criminal is usually a young male. While it is true that young males do commit the majority of most types of crime, the proportion of crime that females are responsible for is increasing. This is also true of juvenile offenders. Boys still commit a greater proportion of crime than females, but statistics show that juvenile girls are becoming increasingly involved in the criminal justice system. For example, in 1992, girls made up of only 20% of juvenile arrests. In 2013, this increased to 29%. Moreover, a greater proportion of girls had contact with the courts, were sentenced to probation, and received out of home placements than before (Sherman & Balck, 2015). UCR data for juvenile offense types show that between the year 2001 and 2010, the proportion of girls arrested for larceny rose from 39.2% to 46%, and overall property crime arrests rose from 31.9% to 38.4% (Chesney-Lind & Shelden, 2013). While it may be that girls are committing more crimes, it is possible that girls are being treated more harshly by the criminal justice system than in the past (Stevens, Morash, & Chesney-Lind, 2011). Perhaps police officers are more willing to arrest girls who may have previously been given a warning. Whether there is an increase in juvenile females committing crimes or a change in the way they are treated, the salient fact is that more girls are coming in contact with the criminal justice system now than ever before.

So why is this important? For one, more arrests lead to more convictions, which leads to more females being sentenced to supervision either in the community or in a secure facility. These girls are given risk assessment tools in order to determine the level

of security needed while they are incarcerated, to determine which risk factors should be focused on in correctional treatment programs, and to determine their risk of re-offending once they are released from custody. The most widely used risk assessment tool in juvenile corrections, the Youth Level of Service Inventory (YLS), is used on both male and female juvenile offenders (Andrews & Bonta, 2006). It is based on social learning theory, which is purported to be a gender-neutral theory of crime. Gender-neutral theories do not look at gender differences and do not take into account the unique gendered pathways a juvenile may take into crime. For example, it has been found that abuse histories, including physical, emotional, and sexual abuse are much more prevalent in female juvenile offenders than in male juvenile offenders (Belknap & Holsinger, 2006; Bloom, Owen, Deschenes, & Rosenbaum, 2002; van der Put et al., 2014). Delinquent peer associations appear to be the strongest predictor of recidivism for boys, while girls seem to be much more influenced by factors related to the family (Bloom et al., 2002; Funk, 1999; van der Put, 2014). Since prior research suggests differences in the pathways to crime and the causes of recidivism for girls and boys , is it appropriate to use the exact same risk assessment tool for both genders? If this is in fact the case, we may be doing girls a disservice by incorrectly assessing their risks to reoffend based on criteria from a theory that was originally developed and tested with male samples.

The question of whether it is appropriate to use alleged gender-neutral risk assessment tools on female offenders is of upmost importance, and several other researchers have focused on this question. While empirical evidence suggests that adult females would benefit from a gender-specific risk assessment tool (Holtfreter & Cupp, 2007; Holtfreter & Morash, 2003; Holtfreter, Reisig, & Morash, 2004; Van Voorhis,

Wright, Salisbury, & Bauman, 2010), the topic has not been as extensively studied for juveniles, and findings are mixed. Some studies have found that gender-neutral tools do not accurately predict risks for juvenile females, and suggest that they may benefit from a gender-specific risk assessment tool (Baird et al., 2013; Schmidt, Campbell & Houlding, 2011). Other have found that the gender-neutral tool works equally well for both genders (Barnes, 2013; Campbell et al., 2014; Flores, 2013; Gossner, 2003; Jung & Rawana, 1999; Schmidt, Hoge, & Gomes, 2005; Stockdale, 2008; Thompson & McGrath, 2012), and it has even been found that the gender-neutral tool works better on females than males (Onifade, Davidson, & Campbell, 2009). To help address these conflicting results, a meta-analysis will be conducted using studies that have examined the predictive validity of the YLS risk assessment instrument. Comparing the mean effect sizes will determine if the tool predicts recidivism significantly better for one sex than the other.

THEORY AND RESEARCH

Theories used to explain juvenile offending typically fit under one of two broad schools of thought: mainstream—or gender-neutral—theories, and feminist—or gender-specific theories. The former general theories are purported to explain all crime at all times, and to do so equally well across sociodemographic groupings (e.g., age, race and ethnicity, gender). The latter, which were developed by feminist criminologists in response to the limitations of general theory, are based on the notion that males and females follow unique pathways to offending.

GENDER-NEUTRAL THEORIES OF CRIME

As noted at the outset, general or mainstream theories are considered gender-neutral. This includes many of the most well known general theories of crime, including

self-control theory (Gottfredson & Hirschi, 1990) and social control theory (Sampson & Laub, 1993) among others. This is also true for differential association theory. The main premise behind differential association theory is that behavior is learned from interactions with others. People that are exposed to delinquent others will be more likely to acquire definitions that are favorable to the violation of the law, whereas people who are exposed to others who have pro-social attitudes are likely to acquire definitions that are unfavorable to the violation of law. The people who learn from others that delinquency is favorable are the ones more likely to engage in it (Sutherland, 1947).

Smith and Paternoster (1987) argued that Sutherland's theory was clearly meant to explain male deviance rather than female deviance, and they tested differential association theory on marijuana use among adolescence. They found that associating with deviant others and holding definitions that are favorable to crime increased delinquency similarly for males and females. Although the causal mechanisms were found to work the same on both sexes, it could be that males are more exposed to delinquent peers and are therefore more likely to develop definitions that are favorable to deviance, thus explaining why males have a higher crime rate than females.

Much like differential association theory, social learning theory assumes that people learn criminal behavior. There are four important premises including *differential association*, which claims that deviance is learned through association with others who commit this deviance; *definitions*, in which the person holds values that are favorable to deviance; *differential reinforcement*, in which rewards or punishments are anticipated in committing deviance; and *imitation*, in which people engage in behaviors that they have witnessed others engaging in (Akers, 1998). Although gender was mentioned in the

theory, Akers did not discuss it at any length. Other criminologists discussed how gender might play a role in social learning theory.

Morash (1999) critiqued social learning theory and suggested some ways in which the theory could be improved to better reflect the experiences of female offending. One critique Morash (1999) had was with the conceptualization and measurement of crimes. She stated that instead of testing the theory using only quantitative data, qualitative data should be used as a supplement, since men and women may see acts of violence in very different ways. Sampson (1999) also critiqued the research design used by Akers, and mentioned that the theory does not seem to explain the role gender plays in crime.

Although Akers (1998) briefly mentioned the gender gap in crime by theorizing that males commit more crime because they are socialized to accept deviance more than females, who are socialized to conform, Morash (1999) offered a critique of this statement as well. She theorized that socialization is different dependent on race, culture, and socio-economic status, and opportunity and power is different for men and women. These factors may differently influence the variables that explain social learning theory.

Another theory closely related to social learning theory is the personal, inter-personal, and community-reinforcement perspective (PIC-R). This theory also looks at interactions with others and how they may influence deviant and pro-social behavior. Andrews and Bonta (2003) used this theoretical model in the development of the LSI-R risk assessment tool, and claim that the theory explains offending for all individuals, regardless of their demographics or pathways into crime.

Social bond theory is another gender-neutral theory that was originally developed and tested using an all-male sample. Hirschi (1969) theorized that deviance could be

avoided if social reinforcement could be developed through bonds to conventional others. Hirschi did not differentiate how the four elements of social bond (attachment, involvement, commitment and belief) could differ for males and females and therefore explain gender differences in crime.

However, other researchers have theorized ways in which these concepts may differ by sex. For example, there may be differences in the way boys and girls are raised, with parents encouraging girls to be more attached to others and boys to be more independent. This may cause the deterrent effect for deviance provided by conventional others to be stronger for girls than boys, explaining the higher crime rate among male offenders (Block, 1984). Another possible explanation for how social bond theory may work differently for boys than girls is way children are supervised. Perhaps, girls are supervised by parents more intensely than boys are. This monitoring may lead to the formation of a stronger parental bond for girls than for boys, therefore providing a greater deterrent effect (Hagan, Gillis, & Simpson, 1985).

GENDER-SPECIFIC THEORIES OF CRIME

Feminist criminologists have questioned the idea that general theories can account for offending equally well among males and females, arguing that there are unique factors and circumstances that facilitate entry into crime. Along these lines, one of the most influential perspectives has been Daly's (1992) pathways to crime. This organizing framework recognizes the differences in factors that may lead to crime between males and females, and also takes into account within-gender variation in offending patterns. Along these lines, there are five purported pathways to crime. According to Daly's research, the "street women" have troubled childhoods and run away from home to a life

on the streets. The "drug connected women" have family members involved in using or dealing drugs and got involved in drugs through them. The "harmed and harming women" experience abuse in childhood and lash out violently as a result of that. The "battered women" experience abuse from a romantic partner. The "economically motivated women" usually commit crimes out of financial motivation. Reisig, Holtfreter, and Morash (2006) found that the economically motivated women did not share unique experiences of abuse characterized by the other typologies, and the crimes of these offenders most closely resembled those of male offenders. Daly's feminist theory of crime suggests that females may have different pathways to crime, some of which are unique to the gender, and one (economically motivated) that closely resembles male offending.

Holsinger (2000) also warned against assuming that males and females have similar life experiences, and therefore, similar pathways into crime. Holsinger stated that the socialization of girls and boys is very different, with boys being encouraged to be more aggressive and girls using other means to resolve conflict. Hagan, Gillis, and Simpson (1985) used differences in socialization between males and females to discuss gender differences in delinquency. In power-control theory, Hagan, Gillis, and Simpson (1985) theorized that males have a higher rate of delinquency because of the way families are structured. Females are less dominant within the family, and this may cause deterrence to exert a stronger effect on females than males, as females may be more deterred by formal sanctions. This theory explains the gender gap in delinquency.

Like Daly (1992), Holsinger (2000) also mentioned the role that victimization plays in the pathways to female offending. Incarcerated girls and women have a higher

rate of abuse than girls and women in the general population, and girls in the juvenile justice system have a higher rate of abuse than boys in the juvenile justice system. An example of how abuse can lead to offending is running away from an abusive home, which is considered an offense for juvenile girls (Belknap and Holsinger, 2006).

Another perspective that links abuse as a pathway to female offending is the cycle of violence. Widom (1989) stated that violence is cyclical, and being abused increased the risk for delinquency and violent behavior later on life. Riveria and Widom (1990) found that abused and neglected girls were more likely than non-abused girls to commit violent offenses as adults, whereas abused boys were no more likely than non-abused boys to continue the cycle of violence as adults. This finding suggests that there are sex differences in abuse as a pathway to crime.

While there is no consensus on whether gender-neutral theories are able to adequately explain female offending, or whether we need theories that take the unique experiences of females in order to explain their offending, criminological theories have many uses in the criminal justice system. One use of theory is to guide the development of risk assessment tools, which will be discussed in the next section.

RISK ASSESSMENT

Risk assessment instruments are tools that are administered to offenders when they enter the criminal justice system. Forensic psychologists or other correctional employees administer these tools that are used to determine offenders' correctional placement and risk of offending while incarcerated, as well as predict recidivism after release. Several different risk assessment tools are currently being used in corrections. These vary according to the demographic characteristics of the offender, the jurisdiction,

and the type of crime. Some risk assessment tools survey the offender directly, and some survey people who the offender is associated with, such as parents or teachers. Some rely on observations of the offender by criminal justice professionals. Areas that risk assessment tools may tap into include mental health concerns, criminal history, family problems, and delinquent peer associations, among others (Andrew, Bonta, & Wormith, 2006).

Risk assessment tools that accurately predict recidivism once released are crucial to offender success. Andrews et al. (2006) discussed the "RNR" principles of risk assessment. "RNR" stands for risk, need, and responsivity, and they provide a guideline for using risk assessment tools on offenders. The first principle is the *risk* principle. This principle states that the level of intervention should match the risk that the offender has to reoffend. Offenders that are considered to be high risk to recidivate are the ones that should be focused on using interventions. Since a low risk offender has a smaller risk of committing another crime once released from incarceration, time and resources would be wasted on these offenders who would probably not recidivate even if they did not receive programming. It would be much wiser to devote time and money on administering programming to the offenders who pose a greater risk to reoffend.

The second principle is the *needs* principle. This principle states that interventions should focus on the specific needs of the offender. There are two categories of needs: dynamic needs and static needs. Dynamic needs are those that are amenable to treatment, such as substance use, pro-criminal attitudes, and deviant peer associations. These are contrasted by static needs, which cannot be changed with intervention. These include prior criminal history, childhood abuse, and family offending. The newest risk

assessment instruments focus on both dynamic and static needs (Andrews et al., 2006; van der Put et al, 2014), while older risk assessment tools generally focused on one or the other. The goal in focusing on both dynamic and static needs is to change what can be changed, and teach the offender coping strategies for things that cannot be changed.

The *responsivity* principle has two components: general responsivity and specific responsivity. General responsivity states that cognitive behavior therapy should be used in treatment in order to influence changes in behavior. Specific responsivity addresses how the treatment should be administered to the offender. This principle takes into consideration that not every offender responds to the same interventions, and the treatment should be administered to offenders in a way that has the greatest potential to be effective (Andrews et al., 2006). For example, having an illiterate offender attend a class that requires them to read textbooks and handouts will probably not have the desired effect. The offender may become frustrated with the difficulties caused by the method of administering the intervention, and the treatment would be unlikely to work. Assessing the offender before treatment and determining what has not worked for them in the past will ensure that the treatment is delivered in a way that the offender finds most receptive. This will give the offender the best chance at treatment success.

THE EVOLUTION OF RISK ASSESSMENT

Andrews et al. (2006) discussed the history of risk assessment instruments, which have evolved over the past few decades. The first generation of risk assessment tools used unstructured clinical judgments to assess offender risks. This approach relied on the subjective opinion of the psychologist administering the instrument, and this led to different psychologists having different assessments for the same offender. Not

surprisingly, the subjective nature of these assessments caused the first generation tools to have very weak predictive validity. A meta-analysis of first generation tools found that the mean predictive validity had an effect size of only r = .12, which is considered to have a low level of ability to predict recidivism (Andrews et al., 2006).

Second generation tools relied on actuarial assessments, which were more objective than the unstructured clinical judgments. These tools took weighted scores of different factors and assigned a risk level based on how many of these factors were present. These tools focused mainly on static risk factors, and left out the majority of dynamic factors. The second generation of risk assessment tools was a vast improvement over the first generation, with a meta-analysis showing a mean effect size of r = .42 for general recidivism and r = .39 for violent recidivism (Andrews et al., 2006).

The third generation risk assessment tools also employ actuarial assessments, but added more dynamic risk factors. They were also based on theory, whereas the second-generation tools were atheoretical. Fourth generation risk assessment tools, which are the most recent, include protective factors and integrate a case-management plan that follows the offender from intake to case closure. The third generation LSI-R and the fourth generation LSI/CMI have effect sizes that range from r = .22 to r = .63 (Andrews et al., 2006).

THE YOUTH LEVEL OF SERVICE INVENTORY

One of the most widely used tools in risk assessment is the Level of Service Inventory-Revised (LSI-R). The LSI-R is an alleged gender-neutral actuarial tool that quantifies subjective and objective characteristics using an interview administered by trained correctional staff in order to determine the risk level of the offender (Andrews &

Bonta, 1995). A tool for assessing youth offenders, called the Youth Level of Service/Case Management Inventory, was constructed from the LSI-R. The YLS/CMI has 42 risk items and relies on interviews and collateral information in order to predict future recidivism. The YLS/CMI has 8 subsections that each represent a criminogenic area of risk factors important for explaining youth offending: Prior and Current Offenses/Dispositions, Family Circumstances/Parenting, Education/Employment, Peer Relations, Substance Abuse, Leisure/Recreation, Personality/Behavior, and Attitudes/Orientation. Subsection scores are tallied, and administrators arrive at a risk level of low (0 to 8 items present), medium, (9 to 22 items present), high (23 to 34), or very high (35 to 42 items present), for each youth offender (Olver, Stockdale, & Wormith, 2009).

The YLS/CMI is theoretically based on social learning theory, discussed previously. Social learning theory is considered a general theory of crime, (i.e., purported to explain both male and female offending). A question that continues to plague risk-assessment research is whether or not it is appropriate to use risk assessment tools that are theoretically based on general theories of crime for female offenders, or whether gender-specific risk-assessment tools are needed. Perhaps males and females recidivate for different reasons, and separate risk assessment tools are needed in order to capture the unique risks of each gender.

Smith, Cullen, and Latessa (2009) conducted a meta-analysis to determine if the LSI-R predicted the recidivism of both male and female adult offenders. In their sample of 14,737 female offenders, they found that both sexes had similar effect sizes in the prediction of recidivism, suggesting that the LSI-R performs equally well for females as

it does for males. Smith et al. (2009) recommended that correctional officials use the LSI-R for both males and females. However, not all studies have found that the LSI-R was equally effective in predicting female recidivism. In contrast, Reisig, Holtfreter, and Morash (2006) tested the LSI-R on a sample of female offenders and found that it only predicted recidivism for females whose crimes resembled male crimes (economically motivated in Daly's theory) and that recidivism was not predicted accurately for the other gendered pathways.

Van Voorhis, Wright, Salisbury and Bauman (2010) addressed this concern by creating an add-on trailer to the LSI-R. Instead of creating a new risk assessment tool specifically for females, adding a trailer that focused on risks unique to female offenders was found to be more cost-effective and faster. Using the trailer along with the gender-neutral tool increased the predictive validity for female offenders. However, this trailer was not tested on juvenile offenders in the original study, so juvenile differences are still unknown. In this next section, empirical testing of the YLS/CMI will be discussed.

YLS/CMI – PREVIOUS FINDINGS

The YLS/CMI is used on different types of offenders and in different situations. It is used on both male and female offenders who have committed a variety of crimes, such as property crimes, status offenses, violent crimes, and even sexual offenses. It is used on offenders of different ethnicities and in different parts of the world. The question here is whether it is suitable to have one risk assessment tool for all these different circumstances. Much like the LSI-R, researchers have found conflicting results in using the YLS/CMI on female offenders.

The predictive validity is determined by how well the YLS/CMI predicts who will recidivate and who will not based on the risk level. A correct prediction would be that someone who is low risk does not reoffend, and an incorrect prediction is that someone who is low risk *does* reoffend. Several studies have found that there is no significant difference with how well the instrument predicts recidivism for girls and for boys (Barnes, 2013; Campbell et al., 2014; Flores, 2013; Gossner, 2003; Jung & Rawana, 1999; Schmidt, Hodge, Gomes, 2005; Stockdale, 2008; Thompson & McGrath, 2012). Vaswani and Merone (2014) also found that the YLS/CMI showed no significant differences in the predictive validity for males and females for general recidivism, but they found that the instrument predicted violent recidivism slightly better for females than it did for males. While Vitopolous, Peterson-Badali, and Skilling (2012) found that the YLS/CMI predicted recidivism equally well for males and females, they found that the ability to match offenders to recommended services was better for males than it was for females.

Although there are several studies that found the predictive validity to be significantly similar across gender, there are also studies that found differences between males and females. In looking at recidivism over a ten-year period, Schmidt, Campbell, and Houlding (2011) found that the YLS/CMI predicted recidivism for males better than it did for females. Baird et al. (2013) also found that the tool predicted recidivism better for boys than girls. While Olver, Stockdale and Wong (2012) found that the instrument worked equally well on some outcomes, they found that it predicted male recidivism better for others. While the study completed by Marshall, Egan, English, and Jones

(2006) looked at violent offending against correctional staff rather than recidivism, it revealed that the YLS/CMI was better at predicting male violence than female violence.

Several of the studies revealed no significant differences in the predictive validity across gender for the entire instrument, but there were subscale differences. Recall that the YLS/CMI has 8 different domains of risk factors. Anderson (2012) discovered that for boys, every subscale significantly predicted recidivism except for the family subscale. For girls, the family subscale was the only one that predicted recidivism. This suggests that different risk factors may be more important in explaining recidivism for each sex—an argument that would be more consistent with the views of feminist criminologists (Belknap & Holsinger, 2006).

Chu and colleagues (2014; 2015) also established that different subscales had different importance across gender and geographical area. Compared with youth offenders in Canada, Singaporean females had more risk factors than Singaporean males, while Canadian males scored higher than Canadian females in seven of the eight subscales. This suggests that along with differences in the YLS/CMI subscales across gender, there may be important cultural differences as well.

Race also showed inconsistent results across studies. While Bechtel, Lowencamp, and Latessa (2008) found that the tool worked best for males and was the least effective for white females, Onifade, Davidson, and Campbell (2009) discovered that the instrument worked the best for white females. They found that the YLS/CMI was least effective for black males. These differing findings suggest that we must continue testing the instrument on different offender populations in order to better understand its use.

To date, only two meta-analyses have been completed on juvenile risk assessment. Schwalbe (2008) looked specifically at gender, and found that effect sizes were almost identical for both genders. However, this meta-analysis looked at many different risk assessment tools, including ASSET, ARNA, and NCAR. In fact, out of 19 studies, only 3 of them looked at the YLS/CMI specifically. This was also the case for a meta-analysis done by Olver et al. (2009). Out of the 49 studies that they used in their analysis, only 22 examined youth adaptions of the LSI-R. While this meta-analysis did include gender comparisons, it was not the main focus of the meta-analysis. Only 13 of those 22 studies contained separate effect sizes for males and females. While the effect sizes were similar for both males and females, suggesting similar predictive validity for both sexes, the samples were mostly limited to Canadian youth, which calls into question their generalizability. The inconsistent findings in previous literature on gender differences as well as the use of various other risk assessment instruments and a focus on mainly Canadian samples in previous meta-analyses suggests the need for another, more up-to-date study.

CURRENT STUDY

The increasing involvement in the criminal justice system for girls proves that it is imperative that we learn all we can about gender differences. An essential component is determining if we can use the same risk assessment tool for girls and boys, or if we should develop a tool that focuses on the unique needs of girls. In doing this, the goal will be to increase the predictive validity of risk assessment and better predict who is a higher recidivism risk.

Although several meta-analyses on youth risk assessment tools have been conducted, a meta-analysis that focuses on the predictive validity of the YLS/CMI between sexes is needed for several reasons. First, existing meta-analyses are several years old, and there are more recent studies that would be able to be included in a new meta-analysis. As well, there is no current meta-analysis that focuses on only the different versions of the YLS/CMI and focuses specifically on gender differences. Previous meta-analyses used several different risk assessment tools, and only one focused specifically on gender. Another reason that a new meta-analysis is needed is that recent research on the topic has come out of many areas of the world. Schwalbe (2008) and Olver et al. (2009) focused mainly on American and Canadian samples. Including samples from other areas, such as Europe, Asia, and Australia, in a new meta-analysis may have interesting and useful findings. The fact that studies are still finding conflicting answers to the question of whether the predictive validity of the YLS/CMI is similar for both girls and boys suggests that a meta-analysis of studies that attempt to answer this question is needed.

HYPOTHESES

The present study is a meta-analysis of sex differences of the Youth Level of Service risk assessment instrument. For the purpose of this paper, the sole research question is whether the YLS/CMI predicts recidivism equally well for females and it does for males. The null hypothesis is that there are no significant differences in the predictive validity of the YLS/CMI between genders. The alternative hypothesis is that the YLS/CMI predicts recidivism significantly better for males than females.

DATA AND METHODS

SAMPLE

The first goal was to identify all current published and unpublished studies that looked at the predictive validity of any version of the Youth Level of Service inventory. Four criteria for inclusion were established. The first is that the study had to use any of the youth versions of the LSI risk assessment tool, including the YLS/CMI, the YLS/CMI/SV, and the YLS/CMI/AA. The second is that it had to assess the risk for recidivism using a longitudinal design. The third is that it had to use either Pearson's *r* or Area Under the Curve (AUC) in the analysis. The fourth is that there had to be separate effect sizes for both males and females. Studies that were dated November 2015 and earlier were used in the study.

To build the sample of studies used in this meta-analysis, both published and unpublished works were used. Publication bias is something that researchers should be aware of when conducting a meta-analysis. A concern is that the studies that are published are more likely to have significant findings, and those that do not have significant findings are not submitted to academic journals. Testing treatments is sometimes not done in an academic context, so these studies may be more likely to remain unpublished and therefore be more likely to suffer publication bias than those testing criminological theory, which are usually done by academics that have a greater interest in getting their work published (Pratt, 2010). The topic of risk assessment is probably not subject to much publication bias, because there is an interest in reporting whether a tool predicts the risk of recidivism well or not, regardless of the result.

However, it is still worth mentioning publication bias, as it is a concern that needs to be addressed when conducting a meta-analysis.

METHOD FOR LOCATING STUDIES

Potential studies were found by entering combinations of the words "youth level of service" "yls" "risk assessment" "juvenile" "predictive validity" and "recidivism" into Google Scholar in order to search academic journals. This was useful in order to find published studies, but was not effective for locating unpublished works. To accomplish this, the same keywords were typed into a university library search tool, which allows the user to choose if they want to search solely for theses and dissertations. ProQuest was then used to access these works.

In total, 65 studies that assessed the YLS/CMI (or LSI on juvenile offenders) were found. However, not all of these met the criteria for inclusion that were established. The first criterion for inclusion was that the study had to specifically look at a youth version of the LSI risk assessment tool, including the YLS/CMI, the YLS/CMI-SV, and the YLS/CMI-AA. Several studies had a juvenile sample, but used one of the adult versions of the LSI (Luong, 2007; Luong & Wormith, 2011), so these studies were excluded because they did not use a juvenile version.

The second criterion for inclusion was that the study had to assess the risk for recidivism using a longitudinal design. Studies were gathered that include any measure of recidivism, including re-arrest, new charge, new conviction, re-incarceration, and technical violation. Studies included different recidivism outcomes such as general recidivism, violent recidivism, and non-violent recidivism, among others. There was one study (Marshall, Egan, English, & Jones, 2006) that used the YLS/CMI to predict the risk

for offending against staff while incarcerating. Since the dependent variable was offending while incarcerated rather than recidivism once released, this study was not included.

The third criterion was that the study had to predict the risk for recidivism using either Pearson's *r* correlation or an AUC value. There were several studies that were located that examined the YLS/CMI and recidivism, but used other statistical methods, such as logistic regression. These mainly included studies that looked at risk domains and had a goal of comparing the coefficients of the separate risk domains between males and females, rather than determining the predictive validity of the tool. These were not included because they did not address the research question (Carnes & Martin, 2011; Costigan, 1999; Chu, Daffern, Thomas, & Lim, 2012; Illaqua, Coulson, Lombardo, & Nutbrown, 1999; Jara, Garcia-Gomis, & Villanueva, 2015; Jack, 2000; Jung, 1996; Taylor, 2009; Vitopoulos, 2011; Vitopoulos, Peterson-Badali, & Skilling, 2012).

The last criterion of inclusion was that the study had to include separate effect sizes for males and females. There were a large number of studies that met the first three criteria, but combined males and females when calculating the predictive validity of the YLS/CMI (Campbell, 2009; Catchepole & Gretton, 2003; Clarke, 2015; Cuervo & Villanueva, 2014; Guebert & Olver, 2014; Hilterman, Nicholls, & Nieuwenhuizen, 2013; Li, Chu, Goh, Ng, & Zeng, 2015; Livsey, 2005; McGrath & Thompson, 2012; Onifade, Barnes, Campbell, Anderson, Petersen, & Davidson, 2014; Onifade, Davidson, Livsey, Turke, Horton, Malinowski, Atkinson, & Wimberly, 2008; Onifade, Wilkins, Davidson, Campbell, & Petersen, 2011; Thompson & Pope, 2005; Peterson-Badali, Skilling, & Haquanee, 2015; Shepherd, Green, & Omobien, 2005; Shepherd, Singh, & Fullam, 2015;

Upperton & Thompson, 2007; Vieira, Skilling, & Peterson-Badali, 2009; Welsh, Schmidt, McKinnon, Chattha, & Meyers, 2008). Since the research question addresses gender differences, these studies that did not separate males and females were excluded from the meta-analyses.

After all the studies that did not meet one or more criteria for inclusion were excluded, 33 studies remained. However, two of these studies (Gossner, 2007; Takahashi, 2010) were unpublished works of studies that were later published. Since the samples were used in the published works and included in these meta-analyses, the unpublished works were not included so that each sample was only used once. Two studies with the same first author used the same sample, so one of the studies was excluded from the meta-analyses (Chu, Yu, Lee, Zeng, 2014). The total number of studies that were used in these meta-analyses was 30.

In order to ensure that no studies were missed, the reference list of each study was examined to see if there was another study cited that could possibly be included. Also, the list of articles that each particular study was cited in was examined on Google Scholar. Since it is very unlikely that there is a study that did not cite any of the important literature or previous studies on the YLS/CMI and was not cited in any future works, it is unlikely that any studies that could potentially meet the criteria for inclusion for these meta-analyses were missed.

Compared to the other meta-analyses on the YLS/CMI (Olver et al., 2009; Schwalbe, 2008), ten studies that were used in Olver et al. (2009) were included in these meta-analyses (Gossner & Wormith, 2007; Marczyk, 2002; McKinnon 2004; Morton, 2003; Onifade et al.; 2008; Rowe, 2002; Schmidt, Hoge, & Gomes, 2005; Stockdale,

2008; van de Ven, 2004; Viljoen et al.; 2008). Three studies that were used in the study completed by Schwalbe (2008) met all of the criteria for inclusion, and were therefore included in this study (Flores, 2004; Rowe, 2002; Schmidt, Hoge, & Gomes, 2005).

ANALYSIS

As stated before, studies that reported the effect size by either correlation (*r)*, or Area under the Curve (AUC) were used. AUC, which is the probability that a random drawing of a recidivist score will be higher than a non-recidivist score, is more commonly being used to predict accuracy in risk assessment. An AUC of 0.5 means complete chance prediction, an AUC of 0.55-0.63 means low predictive ability, an AUC of 0.64-0.70 means average predictive ability, and an AUC of more than 0.71 means good predictive ability (Rice & Harris, 2005, Olver et al., 2011). However, it was not possible to run the analysis using AUC measures. To run the analyses, the program *Comprehensive Meta-analysis* was used, and did not allow AUC scores to be used. To solve this problem, AUC scores were converted into point-bisereal correlation scores (*rpb*) using the table provided by Rice and Harris (2005). Because the table did not report every single conversion to three decimal places, the correlations were rounded to two decimal places once converted. Both previous meta-analyses on the topic (Schwalbe, 2008; Olver et al., 2009) also used *rpb* correlation rather than AUC scores. For the *r* statistics, an *r* of 0.1 represents a small effect size (low predictive validity) an *r* of 0.3 represents a moderate effect size (moderate predictive validity), and an r of 0.5 represents a large effect size (strong predictive validity) (Rice & Harris, 2005). Effect sizes were then entered into the computer program along with the study name and sample size. The

computer then calculated a mean effect size using the information given. Separate models were created for males and females.

When conducting a meta-analysis, either a fixed-effects or a random-effects procedure may be used. The choice between using the two types depends on the homogeneity of the effect size parameters (Hedges & Vevea, 1998). Since there was evidence that there was heterogeneity among the population effects in the studies used in the meta-analysis, the random-effects procedure was used.

RESULTS

DESCRIPTIVE STATISTICS

These studies included both all-male samples, all-female samples, and samples that included both females and males and reported effect sizes separately for both sexes. Studies were conducted in a wide variety of geographical locations, including the United States, Canada, Japan, Singapore, Australia, Scotland, Great Britain, and Germany. Samples were derived from incarcerated samples, probation samples, residential school samples, and mixed samples. While the majority of the samples were racially heterogeneous, one sample each was found for all-white males, all-black males, all-white females, and all-black females. As noted above, several different definitions of recidivism were used, as well as several different offense types. Follow up time for recidivism ranged from 6 months to ten years.

100 different effect sizes were found, ranging from one effect size in a study to 12 effect sizes in a study. Several studies had different effect sizes for recidivism type (general, violent, non-violent) and definition of recidivism (re-arrest, re-change, re-conviction). The effect size that corresponded with the recidivism outcome was used so

that each sample was not used more than once in each meta-analysis, ensuring that every offender in the meta-analysis was counted only once.

Two of the studies were used twice in several of the meta-analyses. This is because they had separate samples for the effect sizes. For example, one of the studies (Onifade et al., 2008) had separate samples for blacks and whites, so both were included in the meta-analyses. The other (Schmidt, Sinclair, & Thomasdottir, 2015) had separate samples for sexual offenders and non-sexual offenders, so both samples were used.

TABLE 1

Descriptive Statistics

	n	**%**
Publication Status		
Unpublished	10	32.3
Published	21	67.7
Geographical Location		
Canada	11	35.5
USA	10	32.3
Europe	3	9.7
Asia	4	12.9
Oceania	3	9.7
Gender Composition		
Male Only	11	35.5
Female Only	1	3.2
Mixed male/female	19	61.3
YLS Version Used		
YLS/CMI	28	90.3
YLS/CMI-SV	2	6.5
YLS/CMI-AA	1	3.2

GENERAL RECIDIVISM

The first meta-analysis used the measure of general recidivism. Only studies that included a measure of general recidivism were included. If they did not have a measure of general recidivism, they were excluded. The studies yielded 27 effect sizes for males, with a total male sample size of 17,059, which is shown in Table One. The studies

yielded 19 effect sizes for females, with a total sample size of 3,301, which is shown in Table 2. The mean effect size for males using point-bisereal correlation (*rpb*) is 0.272 with confidence intervals of 0.237 for the lower limit, 0.306 for the upper limit, and a Q-statistic of 109.957. For females, the mean effect size is *rpb* = 0.280, with confidence intervals of 0.227 for the lower limit, 0.332 for the upper limit, and a Q-statistic of 39.766. Since the confidence intervals overlap, this indicates that the YLS/CMI significantly predicts general recidivism in both male and female samples.

Slightly larger effect size were found in the meta-analysis conducted by Schwalbe (2008), with $r = 0.33$ for non-violent recidivism in male offenders, and $r = 0.36$ for female offenders.. However, it must be noted that this meta-analysis did not only contain studies of the YLS/CMI, but other risk assessment tools as well. In the meta-analysis conducted by Olver, Stockdale, and Wormith (2009), larger effect sizes were found as well, with an r of 0.319 for males and an r of 0.403 for females. While this meta-analysis only used the YLS/CMI as their risk assessment tool, the measure of recidivism was not defined, and therefore contained both violent and non-violent recidivism.

TABLE 2

General Recidivism- Males

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Anderson (2012)	0.17	0.103	0.235	4.943	0.000
Barnes (2013)	0.15	0.031	0.340	2.332	0.020
Bechtel, Lowencamp, Latessa (2007)	0.20	0/171	0.231	12.695	0.000
Caldwell, Dickenson (2009)	0.12	-0.030	0.265	1.568	0.117
Campbell et al. (2014)	0.18	0.079	0.273	3.482	0.001
Chu et al. (2015)	0.25	0.216	0.284	13.858	0.000
Flores (2013)	0.30	0.235	0.383	8.533	0.000
Gossner, Wormith (2007)	0.41	0.199	0.585	3.645	0.000
Greiger, Hosser (2014)	0.24	0.162	0.315	5.925	0.000
Marcyk (2002)	0.07	-0.133	0.256	0.673	0.501
MacKinnon (2004)	0.24	-0.006	0.459	1.912	0.056
Morton (2003)	0.12	-0.107	0.335	1.037	0.300
Olver, Stockdale, Wong (2011)	0.24	0.271	0.599	4.598	0.000
Onifade, et al. (2009) #1	0.23	0.130	0.325	4.443	0.000
Onifade, et al. (2009) #2	0.17	0.063	0.273	3.109	0.002
Rowe (2002)	0.40	0.315	0.479	8.526	0.000
Rennie, Dolan (2010)	0.32	0.160	0.464	3.810	0.000
Schmidt, Hoge, Gomes (2005)	0.25	0.010	0.452	2.043	0.011
Shepherd, Luebbers, Ogloff (2014)	0.38	0.216	0.500	5.277	0.000
Shepherd, et al. (2014)	0.34	0.202	0.465	4.644	0.000
Skowron (2004)	0.25	0.122	0.370	3.742	0.000
Stockdale (2008)	0.70	0.554	0.804	6.992	0.000
Takahashi, Mori, Kroner (2013)	0.34	0.249	0.425	6.957	0.000
Thompson, McGrath (2011)	0.27	0.239	0.300	16.531	0.000
Van de Ven (2004)	0.23	0.150	0.307	5.552	0.000
Vaswani, Merone (2014)	0.40	0.344	0.453	12.829	0.000
Viljoen, et al. (2009)	0.28	0.144	0.405	3.965	0.000
	0.272	**0.237**	**0.306**	**14.658**	**0.000**

TABLE 3

General Recidivism- Females

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Anderson (2012)	0.13	0.013	0.244	2.168	0.030
Barnes (2013)	0.21	-0.044	0.439	1.623	0.104
Bechtel, Lowencamp, Latessa (2007)	0.17	0.038	0.244	4.112	0.000
Campbell et al (2014)	0.35	0.22	0.468	5.037	0.000
Chu et al (2014)	0.26	0.154	0.360	4.685	0.000
Flores (2013)	0.26	0.113	0.395	3.429	0.001
Gossner, Wormith (2007)	0.55	0.155	0.793	2.624	0.009
McKinnon (2004)	0.14	-0.188	0.440	0.834	0.404
Olver, Stockdale, Wong (2012)	0.29	0.066	0.486	2.516	0.012
Onifade, et al. (2009) #1	0.38	0.232	0.511	4.784	0.000
Onifade, et al. (2009) #2	0.31	0.144	0.459	3.584	0.000
Onifade et al (2008)	0.16	-0.049	0.355	1.505	0.132
Rowe (2002)	0.60	0.439	0.724	6.122	0.000
Schmidt, Hoge, Gomes (2005)	0.14	-0.179	0.433	0.857	0.331
Schmidt, Campbell, Houlding (2011)	0.06	-0.252	0.361	0.370	0.711
Shepherd, et al. (2014)	0.29	-0.033	0.558	1.766	0.077
Stockdale (2008)	0.39	0.161	0.579	3.243	0.001
Thompson, McGrath (2011)	0.24	0.160	0.317	5.777	0.000
Van de Ven (2004)	0.25	0.112	0.379	3.502	0.000
Vaswani, Merone (2014)	0.38	0.260	0.488	5.866	0.000
	0.280	**0.227**	**0.332**	**9.905**	**0.000**

VIOLENT RECIDIVISM

The next meta-analysis used effect sizes for violent recidivism only. Only the studies that included an effect size for violent recidivism were included in this meta-analysis. This yielded 18 effect sizes for males (n=7,931) and 9 effect sizes for females (n=777). For males, the mean effect size was r = 0.301, with a lower limit of 0.252, an upper limit of 0.348, and a Q-statistic of 60.494. For females, the mean effect size was r = 0.387, with a lower limit of 0.275, an upper limit of 0.488, and a Q-statistic of 20.536. Since the confidence intervals overlap, we cannot say that the YLS/CMI predicts violent recidivism better for one gender than the other.

In the meta-analysis conducted by Schwalbe (2008), the effect sizes were slightly smaller. The *r* value for violent recidivism for males was 0.23, and the *r* value for the violent recidivism measure for females was 0.24. Again, this meta-analysis used several other tools other than the YSL/CMI, so different effect sizes from this meta-analysis are expected.

TABLE 4

Violent Recidivism – Males

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Caldwell, Dickenson (2009)	0.15	0.000	0.293	1.965	0.049
Chu et al. (2015)	0.21	0.175	0.244	11.574	0.000
Greiger, Hosser (2014)	0.26	0.183	0.334	6.442	0.000
McKinnon (2004)	0.23	-0.017	0.450	1.829	0.067
Morton (2003)	0.21	-0.015	0.414	1.834	0.067
Olver, Stockdale, Wong (2012)	0.53	0.366	0.662	5.599	0.000
Rowe (2002)	0.35	0.251	0.442	6.578	0.000
Rennie, Dolan (2010)	0.17	0.001	0.329	1.972	0.049
Schmidt et al. (2011)	0.26	0.028	0.465	2.194	0.028
Schmidt, et al. (2015)	0.51	0.401	0.605	7.978	0.000
Schmidt, et al. (2015)	0.35	0.217	0.471	4.930	0.000
Shepherd et al. (2014)	0.26	0.116	0.393	3.490	0.000
Skowron (2004)	0.27	0.143	0.388	4.078	0.000
Stockdale (2008)	0.51	0.309	0.667	4.537	0.000
Takahashi, Mori, Kroner (2013)	0.35	0.260	0.434	7.180	0.000
Van de Ven (2004)	0.19	0.109	0.269	4.535	0.000
Van de Ven (2004)	0.38	0.159	0.565	3.275	0.001
Vaswani, Merone (2014)	0.31	0.250	0.367	9.707	0.000
	0.301	**0.252**	**0.348**	**11.503**	**0.000**

TABLE 5

Violent Recidivism - Females

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
McKinnon (2004)	0.50	0.215	0.707	3.250	0.001
Olver, Stockdale, Wong (2012)	0.11	0.519	0.519	2.889	0.004
Rowe (2002)	0.60	0.439	0.724	6.122	0.000
Schmidt, et al (2011)	0.14	-0.175	0.429	0.869	0.385
Shepherd et al. (2014)	0.24	-0.086	0.520	1.448	0.148
Stockdale (2008)	0.40	0.173	0.587	3.336	0.001
Van de Ven (2004)	0.25	0.112	0.379	3.502	0.000
Van de Ven (2004)	0.68	0.429	0.834	4.387	0.000
Vaswani, Merone (2014)	0.33	0.206	0.443	5.027	0.000
	0.387	**0.275**	**0.488**	**6.388**	**0.000**

NON-VIOLENT RECIDIVISM

Next, separate analyses were conducted for non-violent recidivism for both males and females. Only a small number of studies included an effect size for non-violent recidivism (general recidivism scores were not used since these may have been violent crimes as well). For males, there were 8 studies that had non-violent effect sizes (n=4285), and for females there were 3 studies (n=428). For males, the mean effect size was $r = 0.324$, with a lower limit of 0.260, an upper limit of 0.395, and a Q-statistic of 19.044. For females, the mean effect size was $r = 0.178$, with a lower limit of 0.075, an upper limit of 0.267, and a Q-statistic of 2.15. Since the confidence intervals overlap, we cannot say that there are significant differences in the predictive validity of the YLS/CMI for non-violent recidivism across gender.

TABLE 6

Non-Violent Recidivism - Males

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Chu et al. (2015)	0.25	0.216	0.284	13.783	0.000
Olver, Stockdale, Wong (2012)	0.40	0.294	0.615	4.839	0.000
Rennie, Dolan (2010)	0.28	0.117	0.429	3.305	0.001
Schmidt, Campbell, Houlding (2011)	0.40	0.184	0.579	3.493	0.000
Schmidt, Sinclair, Thomasdottir (2015)	0.40	0.295	0.496	6.922	0.000
Schmidt, Sinclair, Thomasdottir (2015)	0.35	0.228	0.461	5.371	0.000
Takahashi, Mori, Kroner (2013)	0.35	0.260	0.434	7.180	0.000
Viljoen, Elkovitch, Scalora, Ullman (2009)	0.19	0.050	0.323	2.651	0.008
	0.324	**0.260**	**0.395**	**9.392**	**0.000**

TABLE 7

Non-Violent Recidivism – Females

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Chu et al. (2015)	0.16	0.050	0.266	2.842	0.004
Olver, Stockdale, Wong (2012)	0.31	0.088	0.503	2.701	0.007
Schmidt, Campbell, Houlding (2011)	0.05	-0.262	0.352	0.308	0.758
	0.178	**0.075**	**0.278**	**3.363**	**0.000**

NEW CHARGE OR ARREST

Next, analyses were conducted using the different definitions of recidivism. There were 4 different definitions used in the studies, including new charge, new arrest, new conviction, and new incarceration. Since there were a small amount of studies that included new arrest or new incarceration, new arrest was combined with new charge, and new incarceration was combined with new conviction. This created two different categories for definition of recidivism, arrest/charge, and conviction/incarceration, which is a more conservative definition of recidivism.

In the category of arrest/charge, there were 18 studies for males (n=4959) and 13 studies for females (n=1557). The males had an average effect size of r = 0.274, with a

lower confidence interval of 0.213, an upper confidence interval of 0.333, and a Q-statistic of 77.785. The females had an average effect size of 0.307, with a lower confidence interval of 0.234, an upper confidence interval of 0.376, and a Q-statistic of 25.357. Since the confidence intervals of males and females overlap, we cannot say that there are significant differences of the predictive validity of the YLS/CMI between males and females when the arrest/change definition of recidivism is used.

TABLE 8

New Charge or Arrest – Males

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Anderson (2012)	0.18	0.113	0.245	5.240	0.000
Barnes (2013)	0.19	0.031	0.340	2.332	0.020
Caldwell, Dickenson (2009)	0.11	-0.040	0.255	1.436	0.151
Campbell et al (2014)	0.18	0.079	0.278	3.462	0.001
Campbell et al (2014)	0.31	0.138	0.464	3.452	0.001
Gossner, Wormith (2007)	0.41	0.199	0.585	3.645	0.000
Marcyzk (2002)	0.07	-0.133	0.268	0.673	0.501
Morton (2003)	0.11	-0.177	0.326	0.950	0.342
Onifade, et al. (2014)	0.23	0.130	0.325	4.443	0.000
Onifade, et al (2014)	0.18	0.074	0.282	3.296	0.001
Shepherd, et al. (2014)	0.38	0.245	0.500	5.247	0.000
Shepherd et al. (2014)	0.38	0.246	0.500	5.277	0.000
Skowron (2004)	0.25	0.122	0.370	3.762	0.000
Stockdale (2008)	0.70	0.554	0.804	6.992	0.000
Vaswani, Merone (2014)	0.40	0.344	0.453	12.829	0.000
Van de Ven (2004)	0.23	0.150	0.307	5.522	0.000
Van de Ven (2004)	0.38	0.159	0.565	3.275	0.000
Viljoen, et al. (2009)	0.28	0.144	0.405	3.965	0.000
	0.274	**0.213**	**0.333**	**8.497**	**0.000**

TABLE 9

New Charge or Arrest – Female

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Anderson (2012)	0.13	0.013	0.244	2.168	0.030
Barnes (2013)	0.21	-0.044	0.439	1.623	0.104
Campbell et al. (2014)	0.35	0.220	0.468	5.037	0.000
Campbell et al. (2014)	0.26	0.064	0.437	2.580	0.010
Gossner, Wormith (2007)	0.55	0.155	0.793	2.624	0.009
Onifade et al. (2014)	0.38	0.232	0.511	4.784	0.000
Onifade, et al. (2014)	0.30	0.133	0.450	3.461	0.001
Onifade et al. (2008)	0.16	-0.049	0.355	1.505	0.132
Shepherd et al (2014)	0.26	-0.065	0.535	1.574	0.115
Stockdale (2008)	0.39	0.161	0.579	3.243	0.001
Vaswani, Merone (2014)	0.38	0.260	0.488	5.866	0.000
Van de Ven (2004)	0.25	0.112	0.379	3.502	0.000
Van de Ven (2004)	0.68	0.429	0.834	4.387	0.000
	0.307	**0.234**	**0.376**	**7.921**	**0.000**

NEW CONVICTION OR RE-INCARCERATION

For the conviction/re-incarceration definition of recidivism, there were 12 studies on males (n=12,860) and 8 studies on females (n=2,006) that were included in the meta-analyses. For males, the average effect size was r = 0.288, with a lower confidence interval of 0.249, an upper confidence interval of 0.327, and a Q-statistic of 43.122. For females, the average effect size was r = 0.278, with a lower confidence interval of 0.217, an upper confidence interval of 0.338, and a Q-statistic of 12.360. Since the confidence intervals overlapped for males and females, we cannot claim that there is a significant difference in the predictive validity of the YLS/CMI for males and females when the definition of conviction/re-incarceration is used.

TABLE 10

New Conviction or Re-Incarceration – Males

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-value
Bechtel, Lowencamp, Latessa (2007)	0.20	0.171	0.231	12.695	0.000
Chu et al. (2015)	0.26	0.226	0.293	14.448	0.000
Flores (2013)	0.30	0.235	0.363	8.633	0.000
Gossner, Wormith (2007)	0.34	0.119	0.529	2.963	0.003
Greiger, Hosser (2014)	0.26	0.183	0.334	6.442	0.000
Rowe (2002)	0.40	0.305	0.487	7.626	0.000
Rennie, Dolan (2010)	0.30	0.138	0.446	3.556	0.000
Takahashi, Mori, Kroner (2013)	0.38	0.292	0.462	7.860	0.000
Thompson, McGrath (2011)	0.28	0.247	0.313	15.773	0.000
Olver, Stockdale, Wong (2012)	0.50	0.330	0.639	5.211	0.000
Van de Ven (2004)	0.20	0.119	0.278	4.780	0.000
Van de Ven (2004)	0.29	0.059	0.491	2.444	0.015
	0.288	**0.249**	**0.327**	**43.122**	**0.000**

TABLE 11

Reconviction or Re-incarceration – Females

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Bechtel, Lowencamp, Latessa (2007)	0.17	0.088	0.244	4.112	0.000
Chu et al. (2015)	0.30	0.196	0.398	5.450	0.000
Flores (2013)	0.33	0.188	0.458	4.417	0.000
Gossner, Wormith (2007)	0.46	0.035	0.744	2.110	0.035
Rowe (2002)	0.45	0.257	0.609	4.281	0.000
Thompson, McGrath (2011)	0.25	0.171	0.326	6.023	0.000
Olver, Stockdale, Wong (2012)	0.31	0.088	0.503	2.701	0.007
Van de Ven (2004)	0.29	0.154	0.415	4.094	0.000
	0.278	**0.217**	**0.338**	**12.360**	**0.000**

NORTH AMERICAN SAMPLES

Lastly, separate analyses were completed depending on where in the world the data were collected. While Schwalbe (2008) did not examine international comparisons in his meta-analysis, Olver, Stockdale, and Wormtih (2009) looked at differences between Canadian and non-Canadian studies and found that the youth adapted LSI had significantly greater predictive accuracy for Canadian youth than non-Canadian youth. However, they did not look at international differences for males and females separately.

Since there have been a number of recent studies that have used non-North American samples (Chu et al., 2015; Chu, Lee, Zeng, & Teoh, 2011; Greiger & Hosser, 2007; Shepherd, Luebbers, Ogloff, Fullam, & Dolan, 2014; Shepherd, Luebbers, & Ogloff, 2014; Takahashi, Mori, & Kroner, 2013; Thompson & McGrath, 2012; Vaswani & Merone, 2014), separate meta-analyses were conducted for males and females depending on whether they had a North American sample, or a sample from elsewhere in the world.

The North American meta-analyses included 22 studies with male effect sizes (n=9017), and 16 studies with female effect sizes (n=2164). The male studies had an average effect size of r = 0.262, with a lower confidence interval of 0.219, an upper confidence interval of 0.305, and a Q-statistic of 67.341. The female studies had an average effect size of r = 0.263, with a lower confidence interval of 0.197, an upper confidence interval of 0.325, and a Q-statistic of 30.745. Since the confidence intervals of males and females overlap, we cannot say that there are significant differences in the predictive validity of the YLS/CMI between males and females when we use North American samples.

TABLE 12

North American Samples – Male

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Anderson (2012)	0.18	0.113	0.245	5.240	0.000
Barnes (2013)	0.19	0.031	0.340	2.332	0.020
Bechtel, Lowencamp, Latessa (2007)	0.20	0.171	0.231	12.695	0.000
Caldwell, Dickenson (2009)	0.11	-0.040	0.255	1.436	0.151
Campbell et al. (2014)	1.18	0.079	0.278	3.462	0.001
Flores (2013)	0.30	0.235	0.363	8.633	0.000
Gossner, Wormith (2007)	0.41	0.199	0.585	3.645	0.000
Marczyk (2002)	0.07	-0.133	0.268	0.673	0.501
McKinnon (year)	0.25	0.004	0.467	1.995	0.046
Morton (2003)	0.12	-0.107	0.335	1.037	0.300
Olver, Stockdale, Wong (2012)	0.45	0.271	0.599	4.598	0.000
Onifade et al. (2014)	0.23	0.130	0.325	4.443	0.000
Onifade et al. (2014)	0.18	0.074	0.282	3.296	0.001
Schmidt, Hoge, Gomes (2005)	0.25	0.010	0.462	2.043	0.041
Schmidt et al. (2015)	0.40	0.184	0.579	3.493	0.000
Schmidt, et al.(2015)	0.38	0.273	0.478	6.537	0.000
Schmidt, Campbell, Houlding (2011)	0.32	0.196	0.434	4.874	0.000
Skowron (2004)	0.25	0.122	0.370	3.762	0.000
Stockdale (2008)	0.70	0.554	0.804	6.992	0.000
Van de Ven (2004)	0.23	0.150	0.307	5.522	0.000
Van de Ven (2004)	0.38	0.159	0.565	3.275	0.001
Viljoen et al.(2009)	0.28	0.144	0.405	3.965	0.000
	0.262	**0.219**	**0.305**	**11.359**	**0.000**

TABLE 13

North American Samples – Female

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Anderson (2012)	0.13	0.013	0.244	2.168	0.030
Barnes (2013)	0.21	-0.044	0.439	1.623	0.104
Bechtel, Lowencamp, Latessa (2007)	0.17	0.088	0.244	4.112	0.000
Campbell et al (2014)	0.35	0.220	0.468	5.037	0.000
Flores (2012)	0.26	0.113	0.395	3.429	0.001
Gossner, Wormith (2007)	0.55	0.155	0.793	2.642	0.009
McKinnon (year)	0.14	-0.188	0.440	0.834	0.404
Olver, Stockdale, Wong (2012)	0.29	0.066	0.468	2.516	0.012
Onifade, Davidson, Campbell (2014)	0.38	0.232	0.511	4.784	0.000
Onifade, Davidson, Campbell (2014)	0.30	0.133	0.450	3.461	0.001
Onifade et al. (2008)	0.16	-0.049	0.355	1.505	0.132
Schmidt, Hoge, Gomes (2005)	0.14	-0.179	0.433	0.857	0.391
Schmidt, Campbell, Houlding (2011)	0.05	-0.262	0.352	0.308	0.758
Stockdale (2008)	0.39	0.161	0.579	3.243	0.001
Van de Ven (2004)	0.25	0.112	0.379	3.502	0.000
Van de Ven (2004)	0.68	0.066	0.486	4.387	0.000
	0.263	**0.197**	**0.325**	**7.661**	**0.000**

NON-NORTH AMERICAN SAMPLES

The final analyses were done on non-North American samples, and included samples from Asia, Europe, and Australia. There were 9 samples that included males (n=8449) and 4 samples that included females (n=1128). The males had an average effect size of $r = 0.317$, with a lower confidence interval of 0.274, an upper confidence interval of 0.360, and a Q-statistic of 3.324. The females had an average effect size of $r = 0.293$, with a lower confidence interval of 0.233, an upper confidence interval of 0.350, and a Q-statistic of 3.324. Since these confidence intervals overlap, we cannot say that there are significant differences in the predictive validity of the YLS/CMI for males and females when we use non-North American samples. As well, the confidence levels between North American male samples and non-North American male samples overlap. This is also true for the female North American samples and the female Non-North American samples.

This means that the YLS/CMI does not work significantly better in one part of the world than another.

TABLE 14

Non-North American Samples – Male

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Chu, et al. (2015)	0.260	0.226	0.293	14.448	0.000
Chu, Ng, Fong, Teoh (2011)	0.250	0.060	0.422	2.567	0.010
Greiger, Hosser (2007)	0.260	0.183	0.334	6.442	0.000
Rennie, Dolan (2010)	0.300	0.138	0.446	3.556	0.000
Shepherd et al. (2014)	0.380	0.245	0.500	5.247	0.000
Shepherd et al. (2014)	0.380	0.246	0.500	5.27	0.000
Takahashi, Mori, Kroner (2013)	0.380	0.292	0.462	7.860	0.000
Thompson, McGrath (2012)	0.280	0.247	0.313	15.773	0.000
Vaswani, Merone (2014)	0.400	0.344	0.453	12.829	0.000
	0.317	**0.274**	**0.360**	**13.404**	**0.000**

TABLE 15

Non-North American Studies – Female

Study Name	Correlation	Lower Limit	Upper Limit	Z-Value	P-Value
Chu et al. (2015)	0.300	0.196	0.398	5.450	0.000
Shepherd et al. (2014)	0.260	-0.065	0.535	1.574	0.115
Thompson, McGrath (2012)	0.250	0.171	0.326	6.023	0.000
Vaswani, Merone (2014)	0.380	0.260	0.488	5.866	0.000
	0.293	**0.233**	**0.350**	**9.264**	**0.000**

DISCUSSION

Building on prior research, the current meta-analyses examined the performance of a purportedly gender-neutral risk assessment tool in predicting recidivism for male and female juvenile offenders. There were no significant differences between the predictive validity of the YSL/CMI and related instruments for male and female juvenile offenders.

THEORETICAL IMPLICATIONS

The results from these meta-analyses support the use of this instrument on both male and female juvenile offenders. This finding bolsters the previous meta-analysis on gender and juvenile risk assessment tools conducted by Schwalbe (2008). However, this is not to say that there are no gender differences in the causes of juvenile offending and no differences in the risk factors most important for predicting offending for each gender. As noted in Schwalbe (2008), actuarial risk assessment tools do not assume the underlying causes of crime. Rather, they add up the total of risk factors that may or may not be present and give a score that is then transformed into their likelihood to recidivate, which is either very high, high, medium, or low risk. We cannot say that male and female juvenile offenders have similar causes of crime, but rather, that the actuarial risk assessment tool *equally* predicts risk.

While the non-significant findings suggest that a risk assessment tool based on a gender-neutral theory of crime will work for both male and female juvenile offenders, this does not mean that gender-specific theories of crime are incorrect. Indeed, a more likely explanation is that gender differences in offending are more pronounced in adult offenders than their juvenile counterparts. Since there have been more significant findings on gender differences with studies using the adult versions of the LSI (Holtfreter & Cupp, 2007; Holtfreter & Morash, 2003; Reisig, Holtfreter, & Morash, 2006; Holtfreter, Reisig, & Morash, 2004) than in youth versions, it could be that gendered experiences in adulthood (i.e.. violent victimization and substance abuse) may lead to differences in offending, rather than assuming that these experiences occur in childhood or adolescence.

RESEARCH IMPLICATIONS

There are several limitations that must be noted when using a measure of recidivism as the dependent variable in a study. One of these is the wide range of follow-up times that were used in the studies used in this meta-analysis. These follow-up times ranged from a mere six months to over ten years. While it is true that most offenders who recidivate do so rather quickly after being released from custody (Schmidt et al., 2011), the range of follow up times may present a problem. It is possible that offenders either desist from crime temporarily after being released and then start back up, or an offender who is released from custody commits crimes that are not detected for a period of time before they are caught again. It is predicted that a study that looks at recidivism for 6 months will have a smaller percentage of recidivists than a study that follows offenders over a ten-year period.

Another limitation that comes with using recidivism is the fact that not all recidivism is detected. The studies used in this meta-analysis used an official measurement of recidivism, except for one single study that used a self-report measure. Official reports are usually conservative estimates for recidivism. While there is no reason to suspect that this would vary by gender, there is a large "dark figure" of crime, and many crimes do not come to the attention of the criminal justice system. Because of this, someone who is classified as a non-recidivist in one of these samples may actually have recidivated. It should be noted that prisons may be seen as schools for crime (Akers, 1998). It could be that the juvenile offenders who were incarcerated learned better methods for committing crimes, and thus were able to commit new offenses once released that were not detected by the criminal justice system.

A related issue that needs to be addressed is the unequal treatment of girls and boys in the criminal justice system. Generally, boys are more likely to be charged with violent offenses such as robbery and assault, and girls are more likely to be charged with status offenses, such as incorrigibility and running away. In fact, girls are three times more likely to be detained for a status offense, and four times more likely to be placed in a correctional facility for a status offense (Chesney-Lind & Shelden, 2013). It is interesting to note that boys and girls have similar self-report rates of committing status offenses, so the over-representation of official sanctions that girls experience is likely due to the differential treatment of males and females. While status offenses for male juveniles may be may be seen as "boys being boys" and they would not be subject to being arrested or charged, girls who commit the same offenses may be subject to official sanctions because it is the justice system's job to "parent" these girls (Chesney-Lind & Shelden, 2013; Gavazzi, Yarcheck & Chesney-Lind, 2006). These differences in dealing with status offenses may be reflected in the official recidivism data in the studies used in this meta-analysis. While boys may have not been arrested, charged, or convicted for certain behaviors, girls may have these offenses counted in their official recidivism.

Another factor that should be addressed concerning differential treatment by the criminal justice system is the fact that the studies used in this meta-analysis used samples from all over the world, and therefore represent many different cultures and views towards juvenile offenders. As Olver et al. (2009) stated, the Level of Service risk assessment tools were developed in Canada on Canadian samples, but are used internationally. Australia has adapted the YLS/CMI to include factors that are specifically Australian (Thompson & McGrath, 2012), but many other countries use the tool that was

developed in Canada. Differences in culture, justice systems, and treatment of juvenile offenders may lead to differences in the predictive validity.

Similarly, the predictive validity of individual scales was not examined in this study. As stated previously, even though the mean effect sizes were the same, this does not mean that there are no gender differences in the individual scales of the YLS/CMI. There is evidence that suggests that certain risk domains are of greater importance for one gender over the other. As previously stated, some studies found that the family domain was more important in predicting the recidivism of girls, and the peer domain was more important in explaining the recidivism of boys (Anderson, 2012). However, when looking at mean effect scores of the total instrument, these differences are not apparent. While it would have been a benefit to include a meta-analysis separated by domain, this was not possible in this project. There were only a handful of studies of the YLS/CMI that included separate analysis of the 8 different domains, so it was not possible to conduct a meta-analysis with so few studies. Hopefully, more studies will be conducted in the future, and one day, perhaps a meta-analysis that looks at the separate risk domains may be done.

POLICY IMPLICATIONS

The support for using one risk assessment tool in juvenile corrections rather than two separate tools that was found in this study can benefit correctional systems economically. It may save the correctional system money. Currently, most correctional resources are used to house offenders. A very small proportion is used for treatment programs for juvenile offenders (Justice Policy Institute, 2009). After paying for the costs

of housing offenders and paying correctional officers and other employees, there is very little left over for programming such as mental health services, education, and other intervention programs aimed at preventing recidivism. If it were found that a separate risk assessment tool was needed for juvenile female offenders, some of these scarce resources would be needed in order to develop, test, and implement this new risk assessment tool. This would most likely be taken out of the small percentage of funding that is reserved for programming. Because of the finding that the general tool works equally well among female juvenile offenders, these resources can now be used for other endeavors, which would hopefully include interventions aimed at the highest risk juvenile offenders in order to prevent future recidivism.

CONCLUSION

This study supports the use of the Youth Level of Service inventory on both male and female juvenile offenders. While this meta-analysis determined mean effect sizes using all previous studies that looked at the predictive validity of the YLS/CMI, it does not undermine those studies that do not have similar findings. These studies may contain important differences that were not captured in the calculations of overall mean effect sizes.

As rates of girls' involvement in the criminal justice system continue to climb, it is imperative to pay attention as to how they may differ from male offenders in both their pathways to crime and correctional risks and needs. While this meta-analysis determined that a popular risk assessment tool may be utilized for both genders, it did not explain the differences in male and female offending. Toward that end, future research focused on preventing juvenile offending must continue to examine the complex circumstances—

gendered and neutral—that are associated with both male and female entry into the criminal justice system.

REFERENCES

* = Used in the meta-analyses

Akers, R. L. (1998). Social Structure and Social Learning. *Los Angeles: Roxbury*.

Anderson, V. (2012). Juvenile risk assessment: An examination of gender. Unpublished Master's thesis. Michigan State University. *

Andrews, D.A. & Bonta, J. (1995). *The Level of Service-Inventory Revised*. Toronto, Canada. Multi-health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and delinquency*, *52*(1), 7-27.

Bechtel, K., Lowencamp, 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.*

Baird, C., Healy, T., Johnson, K., Bogie, A., Dankert, E. W., & Scharenbroch, C. (2013). A comparison of risk assessment instruments in juvenile justice. *Madison, WI: National Council on Crime and Delinquency*. *

Barnes, A.R. (2013). Youth level of service/case management inventory. The predictive validity of post-court involvement assessment. Unpublished doctoral dissertation, Michigan State University. *

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. *

Belknap, J., & Holsinger, K. (2006). The gendered nature of risk factors for delinquency. *Feminist Criminology*, *1*(1), 48-71.

Block, J.H. (1984). Sex role identity and ego development. Jossey-Bass

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*(1), 37-56.

Caldwell, M. F., & Dickinson, C. (2009). Sex offender registration and recidivism risk in juvenile sexual offenders. *Behavioral Sciences & the Law*, *27*(6), 941-956. *

Campbell, C.A. (2009). The exploration of a mixed method and empirical short form using the youth level of service/case management inventory (YLS/CMI). Unpublished master's thesis. Michigan State University, Lansing, MI.

Campbell, C., Onifade, E., Barnes, A., Peterson, J., Anderson, V., Davidson, W., & Gordon, D. (2014). Screening Offenders: The Exploration of a Youth Level of Service/Case Management Inventory (YLS/CMI) Brief Screener. *Journal of offender rehabilitation*, *53*(1), 19-34. *

Carnes, T., & Martin, S. (2011). Does the YLS/CMI help to predict recidivism? Report to the division of juvenile justice by Alaska judicial council and institute for social and economic research.

Catchpole, R.E. & Gretton, H.M. (2003). The predictive validity of risk assessment with violent young offenders: A 1 year examination of criminal outcome. *Criminal Justice and Behavior, 30*(6), 688-708.

Chesney-Lind, M., & Shelden, R.G. (2013). Girls, delinquency, and juvenile justice, 4th ed. John Wiley & Sons, Inc.

Chu, C.M., Daffern, M., Thomas, S., & Lim, Y.J. (2012). Violence risk and gang affiliation in youth offenders: a recidivism study. *Psychology, Crime, and Law, 18*(3), 299-315.

Chu, C. M., Lee, Y., Zeng, G., Yim, G., Tan, C. Y., Ang, Y., ... & Ruby, K. (2015). Assessing Youth Offenders in a Non-Western Context: The Predictive Validity of the YLS/CMI Ratings. *

Chu, C. M., Ng, K., Fong, J., & Teoh, J. (2011). Assessing youth who sexually offended: The predictive validity of the ERASOR, J-SOAP-II, and YLS/CMI in a non-Western context. *Sexual abuse: a journal of research and treatment*, 1079063211404250. *

Chu, C. M., Yu, H., Lee, Y., & Zeng, G. (2014). The utility of the YLS/CMI-SV for assessing youth offenders in Singapore. *Criminal justice and behavior*,*41*(12), 1437-1457.

Clarke, M. (2015). The examination of changes in dynamic risk factors and recidivism over time in young offenders. Unpublished doctoral dissertation. University of Toronto, Toronto, Canada.

Costigan, S. (1999). Critical evaluation of the long-term validity of risk/need assessment and its young offender typology. Unpublished master's thesis. Lakehead University, Thunder Bay, Canada.

Cuervo, K., & Villanueva, L. (2014). Analysis of risk and protective factors for recidivism in Spanish youth offenders. *International Journal of Offender Therapy and Comparative Criminology*. 1-17.

Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *S. Cal. Rev. L. & Women's Stud.*,*2*, 11.

Emeka, T. Q., & Sorensen, J. R. (2009). Female juvenile risk: is there a need for gendered assessment instruments? *Youth Violence and Juvenile Justice*.

Flores, A.W. (2013). Examining the youth level of service/case management inventory in the context of reliability, validity, equity, and utility: A 6-year evaluation. Unpublished doctoral dissertation, University of Cincinnati. *

Funk, S. J. (1999). Risk assessment for juveniles on probation a focus on gender. *Criminal Justice and Behavior*, *26*(1), 44-68.

Gavazzi, S. M., Yarcheck, C. M., & Chesney-Lind, M. (2006). Global risk indicators and the role of gender in a juvenile detention sample. *Criminal Justice and Behavior*, *33*(5), 597-612.

Gottfredson, M.R., & Hirschi, T. (1990). A general theory of crime. Stanford, CA: Stanford University Press

Guebert, A. F., & Olver, M.E. (2014). An examination of criminogenic needs, mental health concerns, and recidivism in a sample of violent young offenders: Implications for risk, need, and responsivity. *International Journal of Forensic Mental Health, 13*, 295-310.

Gossner, D. (2003) Validation of the youth level of service/case management inventory with Saskatchewan young offenders. Unpublished Master's thesis. *

Gossner, D., & Wormith, J. S. (2007). The prediction of recidivism among young offenders in Saskatchewan. *Canadian Journal of Police & Security Services*,*5*(1/2), 1-13. *

Grieger, L., & Hosser, D. (2014). Which Risk Factors are Really Predictive? An Analysis of Andrews and Bonta's "Central Eight" Risk Factors for Recidivism in German Youth Correctional Facility Inmates. *Criminal justice and behavior*,*41*(5), 613-634. *

Hagan, J., Gillis, A.R., & Simpson, J. (1985). The class structure of gender and delinquency: Towards a power-control theory of common delinquent behavior. *American Journal of Sociology, 90*(6), 1151-1178.

Hedges, L.V., & Vevea, J.L. (1998). Fixed- and random effects models in meta-analysis. *Psychological Methods 3*(4), 486-504.

Hilterman, E.B., Nicholls, T.L., & Nieuwenhuizen, C. V. (2013). Predicitive validity of risk assessments in juvenile offenders. Comparing the SAVRY, PCL:YV, and YLS/CMI with unstructured clinical judgments. Assessments.

Hirschi, T.H. (1969). *Causes of Delinquency*. Berkeley, CA: University of California.

Holsinger, K. (2000). Feminist perspectives on female offending. *Women and Criminal Justice 12*(1), 23-51.

Holtfreter, K., & Cupp, R. (2007). Gender and risk assessment the empirical status of the LSI-R for women. *Journal of Contemporary Criminal Justice*, *23*(4), 363-382.

Holtfreter, K., & Morash, M. (2003). The needs of women offenders. *Women and Criminal Justice, 14*(2-3), 137-160.

Holtfreter, K., Reisig, M., & Morash, M. (2004). Poverty, state capital, and recidivism among women offenders. *Criminology and Public Policy, 3*(2), 185-208.

Ilacqua, G.E., Coulson, G.E., Lombardo, D., & Nutbrown, V. (1999). Predictive validity of the young offender level of service inventory for criminal recidivism of male and female young offenders. *Psychological Reports, 84*(1), 1214-1218.

Jack, L.A. (2000). Psychopathy, risk/need factors, and psychiatric symptoms in high-risk youth: Relationships between variables and their link to recidivism. Unpublished doctoral dissertation. Simon Fraser University, Burnaby, Canada.

Jara, P., Garcia-Gomis, A., & Villanueva, L. (2015). Impact of type of intervention on youth reoffending: Are gender and risk level involved? *Psychiatry, Psychology, and Law, 0*(0), 1-9.

Justice Policy Institute (2009). The cost of confinement: Why good juvenile justice policies make good fiscal sense. Retrieved from: http://www.justicepolicy.org/images/upload/09_05_rep_costsofconfinement_jj_ps.pdf

Jung, S. (1996). Critical evaluation of the risk/need assessment with aboriginal young offenders in Northwestern Ontario. Unpublished master's thesis. Lakehead University, Thunder Bay, Canada.

Jung, S., & Rawana, E. P. (1999). Risk and need assessment of juvenile offenders. *Criminal Justice and Behavior*, *26*(1), 69-89.

Li, D., Chu, C.M., Goh, J.T., Ng, I.R., & Zeng, G. (2015). Impact of childhood mistreatment on recidivism in youth offenders: A matched control study. *Criminal Justice and Behavior, 42*(10), 990-1007.

Livsey, S.E. (2005). The the YLS a valid predictor of juvenile recidivism? Unpublished doctoral dissertation. Michigan State University, Lansing, MI.

Luong, D. (2007). Risk assessment and community management: The relationship between implementation quality and recidivism. Unpublished master's thesis. University of Saskatchewan, Saskatoon.

Luong, D., & Wormith, J.S. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior, 38*(12), 1177-1199.

Marcyzk, G.R. (2002). Predicting juvenile recidivism and validating juvenile risk factors in an urban environment. Unpublished doctoral dissertation. Medical College of Pennsylvania and Hahnemann University, Philadelphia, Pennsylvania. *

Marshall, J., Egan, V., English, M., & Jones, R. M. (2006). The relative validity of psychopathy versus risk/needs-based assessments in the prediction of adolescent offending behaviour. *Legal and Criminological Psychology*, *11*(2), 197-210. *

McGrath, A., & Thompson, A.P. (2012). The relative predictive validity of the static and dynamic domain scores in risk-need assessment of juvenile offenders. *Criminal Justice and Behavior, 39*(3), 250-263.

Morash, M. (1999) A consideration of gender in relation to social learning and social structure: A general theory of crime and deviance. *Theoretical Criminology, 3*(4), 451-462.

Morton, K.E. (2003). Psychometric properties of four risk assessment measures with male adolescent sexual offenders. Unpublished doctoral dissertation. Carleton University. *

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. *

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk Assessment With Young Offenders A Meta-Analysis of Three Assessment Measures. *Criminal Justice and Behavior*, *36*(4), 329-353.

Onifade, E., Barnes, A., Campbell, C., Anderson, V., Petersen, J., & Davidson, W. (2014). Juvenile offenders and experiences of neglect: The validity of the YLS/CMI with dual-status youth. *Children and Youth Services Review, 46*, 112-119.

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., Davidson, W., Livsey, S., Turke, G., Horton, C., Malinowski, J., Atkinson, D., & Wimberly, D. (2008). Risk assessment: Identifying patterns of risk in young offenders with the youth level of service/case management inventory. *Journal of Criminal Justice, 36*, 165-173.

Onifade, E., Wilkins, J., Davidson, W., Campbell, C., & Petersen, J. (2011). A comparative analysis of recidivism with propensity score matching of informal and formal juvenile probationers. *Journal of Offender Rehabilitation, 50*(8), 531-546.

Peterson-Badali, M., Skilling, T., & Haquanee, Z. (2015). Examining implementation of risk assessment in case management for youth in the justice system. *Criminal Justice and Behavior, 42*(3). 304-320.

Pratt, T. C. (2010). Meta-analysis in criminal justice and criminology: What it is, when it's useful, and what to watch out for. *Journal of Criminal Justice Education*, *21*(2), 152-168.

Reisig, M. D., Holtfreter, K., & Morash, M. (2006). Assessing recidivism risk across female pathways to crime. *Justice Quarterly*, *23*(3), 384-405.

Rennie, C., & Dolan, M. (2010). Predictive validity of the Youth Level of Service/Case Management Inventory in custody sample in England. *The Journal of Forensic Psychiatry & Psychology*, *21*(3), 407-425. *

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and human behavior*, *29*(5), 615.

Riveria, B., & Widom, C.S. (1990). Childhood victimization and violent offending. *Violence and Victims, 5*(1), 19-35.

Rowe, R.C. (2002). Predictors of criminal offending: Evaluating measures of risk/needs, psychopathy, and disruptive behavior disorders. Unpublished doctoral dissertation. Carleton University, Ottawa, Canada. *

Sampson, R.J. & Laub, J.H. (1993). Crime in the making: Pathways and turning points through life. Cambridge, MA: Harvard University Press.

Sampson, R. J. (1999). Techniques of research neutralization. *Theoretical Criminology, 3*, 438-451.

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. *

Schmidt, F., Hoge, R. D., & Gomes, L. (2005). Reliability and validity analyses of the youth level of service/case management inventory. *Criminal Justice and Behavior*, *32*(3), 329-344. *

Schmidt, F., Sinclair, S.M., Thomasdottir, S. (2015). Predictive validity of the youth level of service/case management inventory with youth who have committed sexual and non-sexual offenses. *Criminal Justice and Behavior, XX*(X), 1-18. *

Schwalbe, C. S. (2008). A meta-analysis of juvenile justice risk assessment instruments predictive validity by gender. *Criminal Justice and Behavior*, *35*(11), 1367-1381.

Shepherd, J.B., Green, K.R., & Omobien, E.O. (2005). Level of functioning and recidivism risk among adolescent offenders. *Adolescence, 40*(157), 23-32.

Shepherd, S. M., Luebbers, S., Ogloff, J. R., Fullam, R., & Dolan, M. (2014). The predictive validity of risk assessment approaches for young Australian offenders. *Psychiatry, Psychology and Law*, *21*(5), 801-817. *

Shepherd, S.M., Luebbers, S., & Ogloff, J.R. (2014). Are youth violence risk instruments interchangeable? Evaluating instrument convergence in a sample of incarcerated adolescent offenders. *Journal of Forensic Psychology Practice, 14*(1). 317-341. *

Shepherd, S.M., Singh, J.P., & Fullam, R. (2015). Does the youth level of service/case management inventory generalize across ethnicity? *International Journal of Forensic Mental Health, 14*, 193-204.

Sherman, F.T., and Balck, A. (2015) Gender injustice: system-level juvenile justice reforms for girls. *The National Crittenton Foundation*. Retrieved from: http://www.nationalcrittenton.org/wpcontent/uploads/2015/09/Gender_Injustice_Report.pdf?ed2f26df2d9c416fbddddd2330a778c6=vgzfzfkbg-vbvztgfm

Skowron, C. (2005). Differentiation and predictive factors in adolescent sexual offending. Unpublished doctoral dissertation. Carleton University, Ottawa, Canada. *

Smith, P., Cullen, F. T., & Latessa, E. J. (2009). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders*. *Criminology & Public Policy*, *8*(1), 183-208.

Smith, D.A., & Paternoster, R. (1987). The gender gap in theories of deviance: Issues and evidence. *Journal of Research in Crime and Delinquency, 24*(2), 140-172.

Stockdale, K.C. (2008). The validity and reliability of the violence risk youth version (VRS/YV). Unpublished doctoral dissertation. University of Saskatchewan. *

Stevens, T., Morash, M., & Chesney-Lind, M. (2011). Are girls getting tougher? Or are we just getting tougher on girls? Probability of arrest and juvenile court oversight in 1980 and 2000. *Justice Quarterly, 28*(5), 719-744.

Sutherland, E.H (1947). Principles of criminology, Third edition. Philadelphia, PA: J.B. Lippencott.

Takahashi, M., Mori, T., & Kroner, D. G. (2013). A cross-validation of the Youth Level of Service/Case Management Inventory (YLS/CMI) among Japanese juvenile offenders. *Law and human behavior*, *37*(6), 389. *

Taylor, C. (2009). Girls and boys, apples and oranges. A theoretically informed analysis of gender-specific predictors of delinquency. Unpublished doctoral dissertation. University of Cincinnati, Cincinnati, Ohio.

Thompson, A. P., & McGrath, A. (2012). Subgroup differences and implications for contemporary risk-need assessment with juvenile offenders. *Law and human behavior*, *36*(4), 345. *

Thompson, A.P., & Pope, Z. (2005). Assessing juvenile offenders: Preliminary data for the Australian adaption of the youth level of service/case management inventory. *Australian Psychiatrist, 40*(3), 207-214.

Upperton, R.A., & Thompson, A.P (2007). Predicting juvenile offender recidivism: Risk-need assessment and juvenile justice officers. *Psychiatry, Psychology, and Law, 14*(1), 138-146.

Van de Ven, J.T. (2004). Assessment of risk and need factors and service use in diverted youth. Unpublished doctoral dissertation, Carleton University, Ottawa, Canada. *

van der Put, C. E., Deković, M., Hoeve, M., Stams, G. J. J., van der Laan, P. H., & Langewouters, F. E. (2014). Risk assessment of girls are there any sex differences in risk factors for re-offending and in risk profiles?. *Crime & Delinquency*, *60*(7), 1033-1056.

Van Voorhis, P., Wright, E. M., 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 and Behavior*, *37*(3), 261-288.

Vaswani, N., & Merone, L. (2014). Are there risks with risk assessment? A study of the predictive accuracy of the Youth Level of Service–Case Management Inventory with young offenders in Scotland. *British Journal of Social Work*, *44*(8), 2163-2181. *

Vieira, T.A., Skilling, T.A., Peterson-Badali, M. (2009). Matching court-ordered services with treatment needs. *Criminal Justice and Behavior, 36(*4), 385-401.

Viljoen, J. L., Elkovitch, N., Scalora, M. J., & Ullman, D. (2009). Assessment of reoffense risk in adolescents who have committed sexual offenses: Predictive validity of the ERASOR, PCL: YV, YLS/CMI, and Static-99. *Criminal Justice and Behavior*. *

Vitopoulos, A.N. (2011). The efficacy of the risk-need-responsivity framework in guiding treatment for young female offenders. Unpublished master's thesis. University of Toronto, Toronto, Canada.

Vitopoulos, N. A., Peterson-Badali, M., & Skilling, T. A. (2012). The relationship between matching service to criminogenic need and recidivism in male and female youth examining the RNR principles in practice. *Criminal Justice and Behavior*, *39*(8), 1025-1041.

Welsh, J.L., Schmidt, F., McKinnon, L., Chattha, H.K., & Meyers, J.R. (2008). A comparative study of adolescent risk assessment instruments: Predictive and incremental validity. *Assessment, 15*(1). 104-115.

Wattanaporn, K.A., & Holtfreter, K. (2014). The impact of feminist pathways research on gender-responsive policy and practice. *Feminist Criminology, 9*(3), 191-207.

Widom, C.S. (1989). The cycle of violence. *Science, 244*(1), 160-166.

APPENDIX A

CODE SHEET

Codesheet - Predictive Validity of the YSL/I Meta-analysis

Model # _____ Data set # _____ Study # _____

Authors: _____

Article Name: _____

Publication Status: Published _____ Unpublished _____

Demographics

Percentage Male: _____ # Male _____

Percentage Female: _____ # Female _____

Racial Composition: 1= Heterogeneous 2= White Only 3= White and Black
4= Black Only 5= Native 6= Other

Age: _____

Origin of Sample: 1= United States 2= Canada 3= Europe
4= Asia 5= Oceania

Sampling Frame: 1 = Prison 2 = Probation 3 = Mixed Prison/Probation
4 = Residential School 5 = Mixed Prison/Residential School
6 = Referred to Mental Health Services 6 = Court Contact

Sampling Technique: 1= Full Sample 2= Random Sample 3= Non-random Sample

Measures

Version of YLS used: 1 = YLS/CMI 2 = YLS/CMI-SV 3 = YLS/CMI-AA

Definition of Recidivism: 1 = Re-arrest 2 = Re-conviction 3 = Re-Incarceration
4 = New Charge 5= Self Report

Type of Recidivism: 1 = General 2 = Non-Violent 3 = Violent 4 = Sexual
5 = Nonsexual violent 6 = Breach of Conditions

Follow-up Time for Recidivism:_____

Statistical Analysis

r Value: Male: _____ Female: _____

AUC: Male: _____ Female: _____

95% C.I.: Male: _____ Female: _____