2022

Beyond the Binary Male: A Systematic Review of Gender-responsive Modifications to Risk Assessment Tools

Devin Randle

Pepperdine University

Graduate School of Education and Psychology

BEYOND THE BINARY MALE: A SYSTEMATIC REVIEW OF GENDER-RESPONSIVE MODIFICATIONS TO RISK ASSESSMENT TOOLS

A clinical dissertation submitted in partial satisfaction

of the requirements for the degree of

Doctor of Psychology

by

Devin Randle

July, 2022

Susan Hall, J.D., Ph.D. – Dissertation Chairperson

This clinical dissertation, written by

Devin Randle

under the guidance of a Faculty Committee and approved by its members, has been submitted to and accepted by the Graduate Faculty in partial fulfillment of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

Doctoral Committee:

Susan Hall, J.D., Ph.D., Chairperson

Carolyn Keatinge, Ph.D.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

ACKNOWLEDGMENTS viii

VITA ix

ABSTRACT xii

Chapter 1: Literature Review 1

- Statement of the Problem 1
- Historical and Contemporary Risk Assessment 2
- Social Perspectives on Gender 11
- Psychological Perspectives on Gender 15
- Gender-Responsive Modifications to Risk Assessment Tools 28
- Summary of Proposed Need for Gender-Responsive Modifications 31
- Research Questions 34

Chapter 2: Methodology 35

- Systematic Review Approach 35
- Eligibility Criteria 36
- Workflow for Selection, Screening, and Quality Appraisal 37
- Search, Screening, and Selection Processes 37
- Quality Appraisal of Individual Studies 41
- Data Collection and Extraction 46
- Data Management, Synthesis and Analysis Plan, and Reporting of Results 48

Chapter 3: Results 53

- Overview of Included Studies 54
- Results of Individual and Overall Study Quality Appraisal 58
- Findings Organized by Research Questions 59
- Conclusion 65

Page

Chapter 4: Discussion 67

Implications for Practice and Policy 68
Implications for Research 72
Limitations 74
Summary: Contributions and Concluding Remarks 77

REFERENCES 79

APPENDIX A: Search Records 89

APPENDIX B: Dissertation Database 91

APPENDIX C: Code Manual 94

APPENDIX D: Inclusion/Exclusion Criteria Checklist 97

APPENDIX E: MMAT Tool 99

APPENDIX F: CEBM Tool 101

APPENDIX G: Data Collection And Extraction Form 104

APPENDIX H: Table Of Included Studies 107

APPENDIX I: Clinician Resource 112

APPENDIX J: GPS IRB Approval Notice 117

LIST OF TABLES

Page

Table 1: Results of Implemented Modifications, as Organized by Tool 61

Table 2: Results of Implemented Modifications of "Adjacent Tools" 62

Table 3: Results of Suggested Modifications, as Organized by Category 64

Table 4: Author Recommendations of Gender-Responsive Tools for Use 71

LIST OF FIGURES

Page

Figure 1: PRISMA Workflow Diagram 54

ACKNOWLEDGMENTS

I would like to extend my sincerest gratitude to my family, friends, professors, and supervisors for their consistent support and sage wisdom throughout this process. My parents have always believed in my passion and my abilities, supporting me through each challenge this process has presented. My husband has remained my biggest advocate, always providing empathy and motivation. My friends have provided me balance throughout my degree, reminding me the importance of self-care. My professors have encouraged me to search for my passion and to embrace flexibility. I would particularly like to thank Dr. Keatinge for agreeing to serve on my committee and for always encouraging her students to expand their knowledge and experience. Lastly, I would like to thank Dr. Hall for her mentorship and encouragement throughout my time in this program. I greatly value the respect and grace that you have consistently provided throughout this process, and I look forward to all that is to come.

VITA

EDUCATION:

Pepperdine University, Graduate School of Education and Psychology **May 2022**
Los Angeles, CA, *Doctor of Psychology in Clinical Psychology* (APA-Accredited)

- Dissertation: "Beyond the Binary Male: A Systematic Review of Gender-responsive Modifications to Risk Assessment Tools"
    - Committee: Susan Hall, PhD, JD (Chair) and Carolyn Keatinge, PhD

George Washington University, **Columbian College of Arts and Sciences** **June 2017**
Washington, D.C., *Master of Arts (MA) in Forensic Psychology*

University of Arizona, **College of Science** **May 2012**
Tucson, AZ, *Bachelor of Arts (BA) in Psychology with Honors, Cum Laude*

CLINICAL EXPERIENCE:

Pre-Doctoral Intern, APA-Accredited Internship **August 2021 – August 2022**
Mendota Mental Health Institute, Madison, WI

- *Risk Assessment Rotation; Supervisor: Stefany Trevino, PsyD*
- *Intensive Treatment Unit Rotation; Supervisor: Karyn Gust-Brey, PhD*
- *Psychotherapy Minor Rotation; Supervisor: Julian Horowitz, PhD*

Doctoral Student Extern **September 2020 – July 2021**
Department of State Hospitals – Metropolitan (DSH-M), Norwalk, CA
Supervisors: Kiera Himsl, PsyD & Narae Lee, PhD

Doctoral Student Extern **August 2020 – June 2020**
Ventura Youth Correctional Facility, Camarillo, CA
Supervisor: James Morrison, PhD

Doctoral Student Extern **July 2020 – October 2020**
UCLA Depression Grand Challenge, Innovative Treatment Network, Screening and Treatment for Anxiety and Depression, Los Angeles, CA
Supervisor: Rafael Rose, PhD

Doctoral Student Extern **August 2019 – July 2020**
Olive View-UCLA Medical Center, Forensic Psychology Inpatient Unit, Sylmar, CA
Supervisor: Kele Kirschenbaum, PsyD

Doctoral Student Assessment Extern **April 2019 – August 2020**
Sports Concussion Institute, Culver City, CA
Supervisors: Tony Strickland, PhD & Katherine Kruser, PsyD

Doctoral Student Extern **July 2019**
Rich & Associates, Friendship Island: A Social Skills Summer Camp, Los Angeles, CA
Supervisor: Seth Shaffer, PhD

Assessment Volunteer **May 2019**
Boys Hope Girls Hope, Irvine, CA
Supervisor: Carolyn Keatinge, PhD

Clinical Psychology Trainee **August 2018 – December 2020**
Pepperdine University, Graduate School of Education & Psychology, Los Angeles, CA
Supervisor: Joan Rosenberg, PhD

Mental Health Worker **July 2017 - March 2018**
Cummins Behavioral Health Systems, Indianapolis, IN
Supervisors: Erika Klosterhoff MS, EdS & Lindsey Bland, MA

Forensic Psychology M.A. Practicum Student **February 2016 - December 2016**
Corrections Corporation of America, Correctional Treatment Facility, Washington, D.C.
Supervisor: Ronald Klein, PhD

Psychology & Legal B.A. Practicum Student **July 2011 - January 2012**
Pima County Attorney's Office, Tucson, AZ
Supervisor: Chad Marchand, MS

Psychology B.A. Practicum Student **January 2009 - December 2009**
University of Arizona, *College of Science*, Tucson, AZ
Supervisor: Julie Feldman, PhD

RESEARCH EXPERIENCE:

Primary Investigator **January 2019 – April 2022**
Pepperdine University, *Graduate School of Education & Psychology*, Los Angeles, CA
Committee: Susan Hall, PhD, JD (Chair) and Carolyn Keatinge, PhD
Dissertation: "Beyond the Binary Male: A Systematic Review of Gender-responsive Modifications to Risk Assessment Tools"

Research Assistant **January 2019 – July 2019**
Pepperdine University, *Graduate School of Education & Psychology*, Los Angeles, CA
Supervisors: Edward Shafranske, PhD and Carol Falender, PhD

Research Assistant & Lab Member **May 2011 - May 2012**
University of Arizona, *College of Science*, *Forensic Psychology Lab,* Tucson, AZ
Supervisors: Judith V. Becker, PhD & Emily Mackelprang

Primary Investigator **August 2010 - May 2012**
University of Arizona, *College of Science*, *Forensic Psychology Lab,* Tucson, AZ
Supervisor: Judith V. Becker, PhD
Honors Thesis: Response to Sexual Offender Notifications Within the Student Community

TEACHING & SUPERVISION EXPERIENCE:

Peer Supervisor **August 2021 – August 2022**
Mendota Mental Health Institute, Madison, WI
Supervisors: Karyn Gust-Brey, PhD & Stefany Trevino, PsyD

Peer Supervisor **September 2020 – August 2021**
Pepperdine University, *Graduate School of Education and Psychology*, Los Angeles, CA
Supervisor: Joan Rosenberg, PhD

Teaching Assistant **September 2019 – December 2020**
Pepperdine University, *Graduate School of Education and Psychology*, Los Angeles, CA
Cognitive Assessment, Personality Assessment, Advanced Assessment
Supervisors: Carolyn Keatinge, PhD and Susan Himelstein, PhD

Preceptor **January 2012 – May 2012**
University of Arizona, *College of Science*, Tucson, AZ
Supervisor: Judith V. Becker, PhD

GRANTS, FELLOWSHIPS, HONORS, & AWARDS**:**

- **Evelyn D. Blake Endowed Scholarship** (2018-2020)
- **Honors Research Grant** (2011)
- **Term Honors** (2008-2017): *Dean's List (with Distinction), First Level Honors, Academic Year Distinction, and Honorable Mention awarded consistently. SOPHOS Sophomore Honorary member.*

PRESENTATIONS & PUBLICATIONS

- Randle, D., Hall, S. (2022, March). Gender-Responsive Modifications to Risk and Need Assessment Tools. Presentation at the Institute of Violence, Abuse, & Trauma's 19th International Summit on Preventing, Assessing, & Treating Trauma Across the Lifespan, Honolulu, HI.
- Yañez, J. J., Kruser, K, Randle, D. (2020, Fall). The Magnification of Health Disparities on Chronic Health Conditions by COVID-19. *The California Psychologist, Volume 53* (4), 21-23.
- Perkins, A., Becker, JV., Randle, D., and Nogami, M. (2013, October). New Neighbors: Promoting Cautious Compassion towards Sex Offenders. Presentation at the ATSA Annual Conference, Chicago, IL.
- Randle, D., Becker, JV., Perkins, A., and Mackelprang, E. (2012, May). Response to Sexual Offender Notifications Within the University of Arizona Student Community. Presentation at the College of Science Honors Research Forum, Tucson, AZ.
- Randle, D., Becker, JV., Perkins, A., and Mackelprang, E. (2012, May). Response to Sexual Offender Notifications Within the University of Arizona Student Community. Presentation at the Honors College Grant Forum, Tucson, AZ.

ABSTRACT

The American Psychological Association's (APA) ethical guidelines identify that psychological tests should only be utilized on populations for which reliability and validity have been established. Likewise, the APA specifies gender to be nonbinary, encouraging clinicians to utilize inclusive and accurate measures. Following such guidelines when conducting violence risk assessments with women and individuals who identify as non-binary is currently challenging because a majority of violence risk assessment tools utilize a male normative sample. Applying such tools to populations beyond a cisgender male can result in misclassifications of individual risk. When risk is misclassified, improper interventions are utilized, which leads to an increase in the probability of recidivism. In response, research has begun to explore gender-responsive modifications to violence risk assessment tools; however, these options have yet to be synthesized in one place. Therefore, this systematic literature review was the first study to identify gender-responsive modifications and which risk assessment tools have implemented them, what suggestions have been made for relevant future research, and how this literature defines gender. Through a comprehensive and detailed review, nine databases were examined, ultimately identifying 59 literature sources that were analyzed using deductive and inductive coding and narrative synthesis. In total, this study identified nine unique means by which violence risk assessment tools may become gender-responsive, alongside 13 violence risk assessment tools that have implemented these modifications. Likewise, this study's results also identified 13 total suggested modifications, split across four categories of potential implementation. Recommendations for clinical use and research on these measures and modifications are discussed. Finally, of note, since none of the synthesized resources examined non-binary gender, with only two resources even mentioning gender outside of binary terms, it is

critical that researchers work to create appropriate violence risk assessment measures for this population.

Keywords: gender, assessment, gender-responsive, risk and need assessment, violence risk assessment, violence risk, non-binary

Chapter 1: Literature Review

Statement of the Problem

As practicing clinicians, psychologists are guided by the ethical guidelines of the American Psychological Association (APA). For the last two decades, the APA has established the expectation that psychologists "are expected to use tests that have been demonstrated to be reliable and valid for the population being tested" (Chrisler & McCreary, 2010, p. 66). Further, recognizing that gender is no longer defined on a binary, APA has expanded these guidelines to encourage clinicians to "abandon binary gender measurement in favor of accurate and inclusionary measures" (APA, 2015, p. 4).

When a clinician is working with a forensic population, or a population involved within the legal system, oftentimes an assessment of risk is pertinent in classification or treatment (Forensic Psychology, 2019). This systematic review focused upon assessment instruments utilized within an adult, forensic population that measure a person's potential to inflict harm upon others. The tools intended to measure this risk may be classified by a variety of names, including violence risk assessment, violence prediction, or risk and need assessment, each with the intention of reducing recidivism by examining various individual factors (National Center for State Courts, 2014). While these examined factors remain diversified amongst the identified tools, such as criminogenic needs (i.e., factors directly related to recidivism), static risk factors (i.e., unchanging factors related to risk), and dynamic risk factors (i.e., changeable factors related to risk), they will henceforth be inclusively referred to as risk assessment tools.

Risk assessment tools are highly useful instruments for clinicians. A limitation in their use is that the gender of the normative populations used to develop the tools are often that of the binary male. Yet, no synthesized resources are available for clinicians to reference which risk

assessment tools may be used on a binary female or nonbinary population. The lack of available synthesized information on gender and violence risk assessment presents a glaring issue, as tools that are improperly used can lead to misclassification of an offender, which can ultimately increase the probability of recidivism (Taxman, 2016). The goal of this dissertation, therefore, was to conduct a systematic review to organize the risk assessment tools that have incorporated gender-responsive components to expand their normative population to encompass genders outside of the male binary. This goal, matched with a call for proposed modifications, a more nuanced understanding of gender in the assessment literature, and for ethically inclusive clinical practice, drives the rationale for this systematic review.

In order to address gender-responsive modifications made to risk assessment tools, this chapter first reviews the initial need for and creation of these tools, details the evolution of risk assessment tools, describes the population and setting for which they are utilized, and notes the limitations of their current use. It then details how both society and psychology are adopting the gender spectrum (in lieu of the gender binary), and the role that gender plays in violence risk and risk and need assessment. The rationale and background for the present study are then summarized with focal research questions for the systematic review of the literature that was conducted for this dissertation.

Historical and Contemporary Risk Assessment

This section describes the history of risk assessment, including the generational evolution of risk assessment tools; examines contemporary risk assessment tools; introduces ethical considerations when examining gender within risk and need factors; and highlights gender differences within research when addressing risk.

History of Risk Assessment

The assessment of an individual's risk to society was not formally addressed until the late nineteenth century when Cesare Lombroso, a renowned criminologist, developed a theory believed to identify future violent criminals based solely upon their physical features (Ackerman, 2010). When individuals committed crimes, they were then punished based upon their unique features, not the crimes themselves, and removed from society through detainment (Ackerman, 2010). It was not until the turn of the twentieth century that society began to shift focus from containing offenders to rehabilitating them, which held as the dominating treatment approach through approximately the next seven decades (Conroy & Murrie, 2007).

By the 1930s, medical explanations began to take prominence in assessing criminal behavior. Attempts to identify moral insanity after the publication of Harvey Cleckley's *The Mask of Insanity* widely propelled the definition of a psychopath (Ackerman, 2010). Legal systems quickly adhered to this newfound "science," enacting laws and statutes that were deemed compassionate and reasonable to identify and treat those that had been labeled as morally insane or psychopathic (Conroy & Murrie, 2007).

By the 1960s, the indefinite terms of hospitalization began to fall away as the movement towards compassion and reasonable treatment was actualized. Hospitalization was replaced by community-based treatment as people in society called into question the condemnation of individuals labeled as deviant (Conroy & Murrie, 2007). The groundbreaking legal case of *Baxstrom v. Herold* (1966) addressed an offender's constitutional right to be released or conserved upon the completion of a criminal sentence. This case resulted in the release of numerous other offenders who may still have had a mental illness but were deemed to be sane upon the completion of their sentences (Warren & Supreme Court of The United States, 1965).

This legal and societal shift sparked a series of longitudinal studies following offenders who had been deemed "high risk" but had never reoffended, which revealed an alarmingly high rate of false-positive predictions, which contributed to the continuation of unjust incarcerations (Ackerman, 2010). Clinicians retrospectively determined that, during this time, when a clinician evaluated a client who was deemed to be of significant risk to society, this assessment was only accurate approximately one-third of the time (Melton et al., 2007).

The recognition that false positives resulted in the curtailment of individual liberties eventually led to the development of future generations of risk assessment tools. In the 1980s, risk assessment tools "emerged as a dominant theme for institutional and community corrections, which contributed to an expanding number of tools in the public domain and proprietary tools" (Taxman, 2016, p. 3). With this increased breadth of tools, the behaviors that were predicted also began to expand (Taxman, 2016). The utilization of risk assessment tools to assess for violence and guide management has continued to hold prioritization within criminal justice and forensic environments, particularly when addressing imminent risk (Ramesh et al., 2018). By the early 2000s, risk assessment was posited "as a core organizing concept of the Western world" (McSherry, 2004, p. 1).

Generations of Risk Assessment Development

Throughout the literature, four generations of risk assessment development span these aforementioned societal phenomena, resulting in unique additions from each era. They are most commonly organized as follows:

- 1st: Professional judgment
- 2nd: Actuarial assessment
- 3rd: Use of static and dynamic risk factors (structured professional judgment)

- 4th: Addition of case management factors (structured professional judgment)

In the early 1900s, the first generation was largely reliant upon the subjectivity of professional judgment offered by clinical and correctional staff based on training and experience. While valuable, neither standardization nor structure was incorporated in the formulation of this opinion, leaving the original judgment of risk privy to human error and biases (Bonta & Wormith, 2008).

The creation of the first actuarial assessment instrument in 1936 marked the beginning of the second generation of risk assessment. These tools allowed for the incorporation of structure and objectivity due to the statistical analyses of static factors (i.e., historical and unchanging factors, such as criminal history) that resulted in numeric formulations of risk. This generation continued to gain momentum, and with it, further iterations of risk assessment modalities began to emerge (Bonta & Wormith, 2008).

The third generation developed through the movement towards risk and need responsivity, leading to the inclusion of both static and dynamic risk factors (i.e., changeable factors, such as engagement in substance use or negative peer associations). This inclusion began to remove the focus solely from historic and immutable focal points towards factors that may shift for the individual, allowing for treatment recommendations in the reduction of risk (Bonta & Wormith, 2008).

Lastly, the fourth generation of risk assessment continues to build upon these ideas by integrating a case management component into the assessment of both static and dynamic factors. Instruments from the second, third, and fourth generation of risk assessment development remain relevant to modern use and are further detailed in the following section (Bonta & Wormith, 2008).

Contemporary Risk Assessment

Today, the ideal risk assessment evaluation balances the inclusion of evidence-based risk tools, static and dynamic factors, and areas of need and intermittent re-assessments of the individual (Lanier & DeVall, 2018). As such, this process will be inclusively referred to as risk assessment, and its three aforementioned components are described next, followed by the different types of risk assessment tools and schools of thought/models.

Static risk factors and dynamic risk factors are both ideally examined within a thorough risk assessment (Melton et al., 2007). Both have been associated with the prediction of violence, such as criminal record, social and interpersonal factors, mental health diagnoses and history, substance use, and more (McSherry, 2004). By also basing risk assessment and treatment on dynamic factors, the clinician is assuming that "offender's risk is not static, fixed, or immutable, but rather is dynamic and changeable and that positive changes can occur on treatment targets" (Otto & Douglas, 2011, p. 127). A mistake often made by clinicians is approaching risk to be a singular trait; when in reality, it is a complex combination of objective and subjective traits held by the individual, as well as factors of the environment (Melton et al., 2007). Risk assessment tools that are formulated from the appraisal of these static and dynamic risk factors, otherwise known as criminogenic needs, aid in identifying which offenders require higher supervision, which offenders would benefit from treatment programs, which are most likely to recidivate, and which offenders simply require shifts to their situational needs (Andrews et al., 2012). While the inclusion of these varied factors provides a well-rounded approach to risk assessment, clinicians must remain aware of factors that influence judgments of dangerousness. These may include variability in the definition of legally relevant risk (henceforth referred to as risk), the variety of supporting literature, judgment biases, and prior legal precedent (Melton et al., 2007).

Encompassing these components, a majority of contemporary risk assessment tools fall under five main types: “clinical assessment,” “anamnestic assessment,” “structured professional judgment,” “actuarial assessment,” and “adjusted actuarial assessment” (Melton et al., 2007, p. 307). Actuarial assessments, from the second generation, have historically been utilized for their reliable ratings and predictive accuracy achieved through derived formulas. However, these formulas are generally composed of static factors, or those that do not change over time (Melton et al., 2007). Since assessment tools that focus on static risk factors are based upon past, unamenable events, Otto and Douglas (2011) note:

> They are not suitable for assessing the current functioning of the individual or informing important clinical or management activities such as treatment to reduce the risk for sexual recidivism, assessment of changes in risk, and community supervision of the individual. (p. 127)

To amend this limitation, many clinicians now prefer to utilize risk assessment tools that incorporate structured professional judgment. This method calls for the clinician to conduct a structured clinical assessment to identify and synthesize both static and dynamic factors that “have established empirical correlations with violence” (Melton et al., 2007, p. 309) with the clinician’s lived experience and expertise, combining both clinical judgments with empirical data. It has been touted as the “gold standard approach [as] recommended by the Department of Health” (Roychowdhury & Adshead, 2014, p. 13). Due to the inclusion of both static and dynamic factors, structured professional judgment instruments may branch both the third and fourth generation of risk assessment development (Bonta & Wormith, 2008).

Clinical assessment largely remains unstructured, consisting of a clinician examining both static and dynamic factors, but with the clinician subjectively deciding what is relevant to the assessment (Otto, 2000). Anamnestic assessment also relies upon the clinician’s individual identification of important themes across factors, leaving each clinician free to decide what holds

weight within the assessment (Otto, 2000). Due to the variability that may exist within these instruments, largely believed to fall within the first generation of risk assessment development, when examining the most commonly used risk assessment tools for general recidivism, violent recidivism, and sexual offense recidivism, a majority of the identified tools fall under either an actuarial or structured professional judgment instrument type (Melton et al., 2007).

The field remains segmented on models of recidivism under which these results may be interpreted. This separation is most frequently identified through the prediction model, which expresses "frequency and probability estimates," and the prevention model, which expresses "categorial estimates of risk (e.g., low, moderate, high), including a listing of relevant risk factors...and risk management strategies" (de Ruiter & Kaser-Boyd, 2015, p. 246). Although the prevention model is historically deemed preferable by forensic psychologists and members of the court, a mixed model approach is still supported by practicing clinicians (de Ruiter & Kaser-Boyd, 2015). Both types of models allow for quantitative data to shape the reduction of recidivism by providing directions to correctional systems on the prioritization of resources, in lieu of the dated assumption that nothing was effective, a bias historically used when addressing recidivism (Taxman, 2016).

Ethical Considerations in Risk Assessment

Given the substantial implications of the determination of risk on both society and forensic practice, ethical considerations are key in the risk assessment process. While it is important to note that violence risk assessment tools can be utilized in a variety of environments, this systematic review remains focused on those of a forensic nature. A clinician's involvement in forensic science, particularly the prediction of violence, remains a controversial topic, namely due to the potentially significant effects upon the offender's freedom, or even life itself (Barbaro,

2019). Legal commitments associated with risk can be both short term, such as temporary civil commitment, or long term, such as rearrest or rehospitalization (Melton et al., 2007). Given that violence risk assessment is pertinent in nearly all forensic psychiatric settings, researchers have encouraged clinicians to treat this with the same weight as other clinical interventions (Roychowdhury & Adshead, 2014). An elevated level of competence is required when administering these instruments, involving years of education, training, experience, and strict adherence to ethical codes (Barbaro, 2019). Considering the significant outcome of these decisions, clinicians often face ethical dilemmas when estimating risk.

Furthermore, in a forensic environment, consent versus court-ordered referrals must be taken into account (Roychowdhury & Adshead, 2014). The clinician must also consider the disclosure of an offender's confidential communication, weighing confidentiality with risk to self, others, and society (McSherry, 2004). The *Tarasoff v. The Regents of the University of California* (1976) case still instructs mental health practitioners on warning intended victim(s) when a formal threat to harm another is made (de Ruiter & Kaser-Boyd, 2015). Errors in judgment can occur throughout the administration of risk assessment tools, particularly in forensic settings (e.g., prison, jails, and hospitals) where offenders are removed from typical social environments and contexts in which future recidivism may occur (Melton et al., 2007). Due to the complexity of this field, as it is based upon both historical factors and ever-shifting human nature, methodological limitations and biases of human error can lead to contradictions in the research and clinical practice (Melton et al, 2007).

Judgment errors are often attributed to cultural factors, as "mental health professionals who conduct such evaluations are typically white and hail from middle- to high-income families, while persons of color and lower socioeconomic status are overrepresented among examinees"

(Melton et al., 2007, p. 301). In striking contrast, persons of color are glaringly overrepresented within many forensic samples, as these often stem from justice systems that are deeply ingrained with their own biased, systemic structure (Davis et al., 2021). When data stems from a racially biased system, this ultimately leads to biases within the resulting risk scores (Eckhouse et al., 2018). This has increased the call for research examining the various layers of systemic bias reflected throughout the criminal justice system. While some researchers have suggested that the utilization of expert witness testimony may have a mitigating factor on a jury's perception of race influencing risk, the data largely remains inconclusive and requires further exploration (Davis et al., 2021).

Societal values can affect clinical judgment, as researchers have found western ideals to emphasize false negatives in lieu of false positives (i.e., it is worse to free someone who is dangerous than to detain someone who is not dangerous; Roychowdhury & Adshead, 2014). Compounding this ethical consideration is the outcome of false negatives that widely affect mental health staff, institutions, policy makers, and society; whilst false positives generally affect the sole individual who received an inflated statistic related to their level of risk (Roychowdhury & Adshead, 2014). Likewise, societal pressures and current risks in media foci can influence bias, such as heightened public anxiety, inadvertently exerting pressure upon researchers to seek predisposing features within patients that may imply biological predispositions to criminal behavior (Wasserman, 2014). With the public spotlight on forensic work, increasingly driven through true crime television shows and podcasts and dubbed the "CSI effect," the general public expects more advanced contributions than what the field can offer at this point in time (Barbaro, 2019).

Gender discrepancies further drive judgment errors, as researchers repeatedly illustrate the glaring disregard for any mention of gender throughout the literature, despite gender being a crucial component for normative data and despite the growing number of females in the forensic sciences (Barbaro, 2019; Melton et al., 2007). Likewise, normative standards are often based upon men-in-power's social behavior, which leads to feminist views that androcentrism dominates psychological standards, and results in male "stereotypic behavior" to be viewed as positive, whilst female's behaviors tend to be "defined negatively" (Chrisler & McCreary, 2010, p. 76). Otto and Douglas (2011) detailed the importance of considering how risk factors differ between binary genders, as some risk factors may be universal, but many will hold higher prevalence in one group versus another.

Social Perspectives on Gender

A crucial component of understanding the necessity of gender-responsive modifications to risk assessment tools is to understand society's ever-developing definition of gender. This section defines gender and its embodiment within individuals' lives, differentiates gender from sex and sexual preference, and addresses the spectrum of gender.

Gender is no longer ruled by the binary idea of male versus female. By eliminating this dichotomous view, the range through which masculinity and femininity are experienced can receive justice, inclusive to variations of background, socio-economic status, generation, and ethnicity that all affect gender expression (Chrisler & McCreary, 2010). The idea of gender now incorporates an expanded notion of self-identity and societal identity, addressing not only the way an individual sees oneself but how others view the individual (Brill & Kenney, 2016). These identities take shape in roles an individual is expected to fulfill, interests an individual is assumed to harbor, and what membership an individual is allocated amongst peers in society

(Brabender & Mihura, 2016). Many of these ideas are represented within feminist theory, which acknowledges the reality of oppression through masculinity's longstanding normative role and posits strategies to shift this dynamic by affirming that gender matters in the way that individuals psychologically, socially, culturally, and historically identify and interact (Taxman, 2016). Messages of gender permeate one's life, starting with childhood toys, following individuals through schooling, bombardments of the media and community, following one's development through religion and culture, and all the while, slowly integrating into almost every aspect of self and society (Brill & Kenney, 2016).

When considering definitions, the terms *gender* and *sex* are often misclassified as interchangeable, "disregard[ing] the nonaggregate differences among women and between men and women" (Hannah-Moffat, 2009, p. 214). *Sex* is innate and biological in nature, generally related to reproductive organs, hormones, and physical makeup, and determines legal classification; while *gender* is formed by society (Bloom et al., 2003). Sex is generally "assigned at birth due to these characteristics, often miss[ing] not-visible intersex configurations…[or] complex genitalia" (Brill & Kenney, 2016, p. 43).

Likewise, it is a common misperception that one's gender identity or expression automates a stereotyped sexual orientation (Brill & Kenney, 2016). However, "gender is about who we are, and sexual orientation is about who we are attracted to physically and/or romantically" (Brill & Kenney, 2016, p. 61). Sexual orientation similarly lives in contained categories of heterosexual, gay, and bisexual, instead of the reality of a "fluid-continuum model of attraction to one or both sexes" (Brabender & Mihura, 2016, p. 506).

Society often exercises labels by *assumed gender*, a term based upon the *assumed sex*, inherently assigns males the gender of boy/man and pronouns of he/him, while females are

assigned the gender of girl/woman and pronouns of she/her (Brill & Kenney, 2016). The embracing of this assigned role is considered *gender conformity* (Brabender & Mihura, 2016). As society continues to embrace gender-expansive pronouns, language continues to shift to embrace that which best defines an individual's identity. The use of gender-expansive pronouns (such as they/them) have become more common identifiers for individuals that do not claim a binary sex.

Gender identity is the internalized experience of one's gender; it is not chosen nor determined by society, but "emerges from within" (Brill & Kenney, 2016, p. 50). Issues of conforming gender identity to one's physical characteristics had previously been addressed by asking individuals to "prove to the medical practitioner their gender identities" (Brabender & Mihura, 2016, p. 745) and often assigned a pathological ascription.

In contrast, *gender expression* is an externalized embodiment of one's gender, often shaped by societal influences such as dress, hairstyle, behavior, and interactions (Brill & Kenney, 2016). Gender expression is composed of:

> *Gender presentation* (how we present our self to others and what they see whether we intend them to see us that way or not); *gender norms* (the rules and ideas about how a culture thinks males and females should look); and *gender roles* (the behaviors we expect, and think are appropriate for each gender). (Brill & Kenney, 2016, p. 54)

The antiquated idea of masculine versus feminine is shifted by gender expression's conformity, or lack thereof, to gender stereotypes (Brabender & Mihura, 2016). Stereotypes are formed from expectations of social groups and are generally flexible in nature, such as intelligence, height, preferences, etc. (Ellemers, 2018). This adaptability is frequently lost when addressing *gender stereotypes*, often maintaining its binary composition, with individuals preserving a "chronically salient, relatively fixed, and easily polarized" view on gender (Ellemers, 2018, p. 277).

Progressing from the binary idea of gender, gender is now defined along a spectrum, encompassing the holistic nature of human development (Brill & Kenney, 2016). The *gender spectrum* is identified as the:

> Concept that gender exists beyond a simple man/woman binary model, but instead exists on a continuum. Some people fall towards more masculine or feminine aspects, some people move fluidly along the spectrum, and some exist off the spectrum entirely. (PFLAG, 2021)

When one's assigned sex matches gender identity, this is considered *cisgender*; while *transgender* refers to one's gender identity not matching one's assigned sex (Brabender & Mihura, 2016). *Gender-expansive* and *nonbinary* are alternate terms used in addressing the removal from binary classification, which is even embodied through this cisgender and transgender categorization (Brill & Kenney, 2016). The term nonbinary, which will be utilized throughout this dissertation, reduces the grand spectrum of gender, but remains a digestible identifier in our current society (Motmans et al., 2019).

While the academic research specifically addressing mental health matters of nonbinary individuals remains limited, the data suggests that nonbinary individuals are at higher risk for discrimination and negative outcomes (i.e., anxiety, depression, and suicidality; Rider et al., 2020; Young, 2020). Pursuing specialized training to gain advanced competencies in working with nonbinary populations are highly recommended. It was only in the recent past that transgender health was conceptualized from a disease model (i.e., a deviation from natural development) in lieu of the modern, and more appropriate, transgender health model, which is that of diverse identity, with association to traumatic stressors stemming from the imposed stigmas that continue to permeate our culture (Motmans et al., 2019). Notwithstanding, when experiences of perceived discrimination based upon gender identity are removed, research

continues to show nonbinary youth remain at a great risk to negative mental health outcomes (Price-Feeney et al., 2020).

Gender affirmation interventions are being made available through social, legal, political, and medical means (Fontanari et al., 2020). Research shows that, when transgender and gender nonbinary youth engaged with these services, results included further promotion of gender positivity and lowered rates of negative mental health outcomes (Fontanari et al., 2020). Influential entities continue to push for the positive promotion of transgender health, with various countries adopting options for nonbinary gender identities, the World Health Organization removing the classification of transgender as a mental disorder, and discussion of the new addition of the Diagnostic and Statistical Manual of Mental Disorders- Fifth Edition (DSM-5) removing gender dysphoria from its next edition (Motmans et al., 2019). With the recent release of the DSM-5-TR, gender dysphoria has not been removed, but has been modified with culturally sensitive language (e.g., "desired gender" changed to "experienced gender" and "cross-sex medical procedure" changed to "gender-affirming medical procedure"; American Psychiatric Association, 2022).

Psychological Perspectives on Gender

As society's definitions and ideas about gender progress, so must the way that the psychological community addresses gender. Biases often emerge through practice, whether present in intake, assessment, diagnosis, psychotherapy, psychological assessment, report writing, or other clinical duties (Brabender & Mihura, 2016). It is the job of the clinician to remain aware of these biases, both within themselves and within our society, since all patients are entitled to competent interventions and assessment (Brabender & Mihura, 2016). Clinical interventions can occur across treatment and practice, including micro- (e.g., utilizing preferred

pronouns), mezzo- (e.g., providing gender-informed trainings to peers regarding interventions, evaluations, or programming), and macro-level interventions (e.g., advocating for gender-inclusivity through means such as policy directives or research; Matsuno, 2019).

In order to examine trends in gender research, Chrisler and McCreary (2010) conducted a content analysis of two of the largest and longest-standing academic journals for community psychology, spanning the years of 1973-2007. While some countries, such as Canada, had already began to incorporate gendered issues within their research initiatives by 1990 (Task Force on Federally Sentenced Women, 1990), Chrisler and McCreary (2010) found a sharp rise in the number of articles addressing gendered research beginning in 1998, signifying a significant increase to publications addressing gendered differences. While research has continued to analyze biases, barriers, and discrimination faced by non-binary-males, it is still widely lacking in recommendations for the effective reduction of these issues (Belingheri et al., 2021).

In order to better understand gender-responsive modifications to risk assessment tools, the psychological and forensic fields' approaches to gender-responsivity must first be examined. This subsection discusses forensic psychology and gender, gender-responsivity with female offenders, gender in psychological assessment, gender in violence risk assessment, and how gender is embodied within forensic psychological research.

Forensic Psychology and Gender

Historically, forensic psychology tended to determine gender based on biological sex, exemplified by the past practice of determining which correctional facility a transgender individual would be assigned to based solely upon the individual's genitalia (Chrisler & McCreary, 2010). This error of gender equating to biological sex was mirrored within legal

policies, as exemplified by the Federal Bureau of Investigation's 1992 classification of rape only applying if the victim and the perpetrator were of two separate genders (Chrisler & McCreary, 2010). Departments of corrections rarely incorporated feminine needs into basic protocols, overlooking female-specific needs such as hygiene and child visitation considerations (Bloom et al., 2003).

The misinformed oversimplification of gender has slowly begun to shift, with legal jurisdictions beginning to determine gender based off of "psychological, sexual, and social identities" (Chrisler & McCreary, 2010, p. 588). Expanding upon this progress, research has found gender differences to exist amongst "each stage in the criminal justice process" (Bloom et al., 2004, p. 29).

As a result, the turn of the century brought the development of and advocacy for gender-informed criminal-justice programs (Gobeil et al., 2016). A decade later, in 2010, the United Nations created the first international standards for adult, female prisoners ("the Bangkok Rules"), calling for research on "the causes of women's imprisonment, the characteristics of women in prison, and the impact [this has] on children" (Taxman, 2016, p. 237). This work specified that classification methods address gender-specific needs and that individualized treatment plans address gender-specific factors of rehabilitation and reintegration to be implemented (Taxman, 2016). As of 2014, the National Institute of Corrections reported that "73% of [responding] jurisdictions have developed some gender-responsive policies for their female offenders involving health care, programming, allowable properties, searches, and restraints" (King & Foley, 2014, p. 3). As of 2015, California became the first state to fund the sex affirmation surgery of a transgender inmate, leading to the expansion of policy in provisions for transgender female inmates that are incarcerated in a men's facility (Associated Press, 2017).

Despite these advances, gender-responsive correctional changes are often not driven by policy or correctional directives, leading to "tremendous inconsistency in the management of women offenders" (King & Foley, 2014, p. 1). When correctional institutions have created gender-responsive shifts to policy, these shifts are historically viewed as a deviation from the standard assumption of genders being interchangeable and may not incorporate the appropriate training required (Gray et al., 2013). Yet, shifts toward gender-responsivity continue to be supported as the literature builds on the deleterious effects that the stress caused by the perception of gender-based stigmas and discrimination on both mental and physical health (Brabender & Mihura, 2016; Vigod & Rochon, 2020).

Gender-responsivity with Female Offenders

Alongside the psychological literature's general tendency to not address gender outside of the binary male, research shows that, in general, gender minorities are "disproportionally affected by psychopathology," inclusive to people living with anxious, depressive, and psychotic disorders (Brabender & Mihura, 2016, p. 656). Similarly, in research based on Western female culture, many assumptions were made about pathways to pathological behaviors, such as the extreme underestimation of sexual abuse encountered by females and the victim-blaming practices that often accompanied the treatment of sexual abuse survivors (Chrisler & McCreary, 2010).

In an effort to validate, or invalidate, these findings, quantitative research on female offenders has increased over approximately the last two decades, reidentifying what was previously considered gender-specific needs to that of gender-specific risk factors (Taxman, 2016). This refocus allowed the view of female offenders to shift, previously being viewed as a

"needy, deficient, and/or poorly adjusted" population (Kruttschnitt & Gartner, 2003, p. 21), to that of a high-risk population that requires specified services and interventions (Taxman, 2016).

In order for truly gender-responsive risk assessment tools to be developed, this rich baseline for defined gender-specific risk factors, environments, and practices must be established (Taxman, 2016). Taxman (2016) acknowledged that the most recent wave of gender-responsive research has been focused upon binary females; however, she also accentuated the benefits of gender-responsive research for *all* genders, as gender-specific risk factors are present for binary males, binary females, and those that identify as gender-expansive. In order for psychology to appropriately address gender, research must remain "contextually grounded and ecologically relevant" to accurately address an individual's resources to best determine treatment modalities (Chrisler & McCreary, 2010, p. 631).

Gender in Psychological Assessment

Historically, a non-gendered, or gender-neutral, approach was thought to be ideal in psychological assessment, particularly within cognitive tests (Brabender & Mihura, 2016; Chrisler & McCreary, 2010). Although this approach might be said to work under the guise of eliminating gender bias, it in fact assumes that the spectrum of genders would fall equally amongst the continuum of the skill being measured, with some theorists going as far as to name gender as irrelevant (Brabender & Mihura, 2016; Chrisler & McCreary, 2010).

Furthermore, even stereotypes of binary gender do not equate upon said continuum, as measured traits are often viewed through a "gender-based lens" (Brabender & Mihura, 2016, p. 14). For example, Brabender and Mihura (2016) described a study examining stereotypes within aggression. The primary subject, a self-identified female client with an aggressive style in the workplace, scored lower on an aggression scale with a non-gendered population norm as

compared to a gendered one. When interviewed, the client's colleagues expected the client's female gender role to insinuate a lower expressed level of aggression. Brabender and Mihura (2016) remind us that "[a]ny discussion of gender - be it a group or an individual - must reference the environment because social environments see individual behaviors through the lens of gender" (p. 14).

The lack of gender research across all assessment tools has been attributed to two forms of bias - that of assessment bias and criterion bias (Chrisler & McCreary, 2010). Assessment bias is recognized as "the same behaviors or symptoms being rated differently based on the gender of the individual who exhibits them" (Chrisler & McCreary, 2010, p. 73), whereas criterion bias is rooted on a deeper level, as the "construct or pathology is based on a gender-biased model" (Chrisler & McCreary, 2010, p. 73). This issue of biases in test construction inherently leads to flaws in statistical inferences and estimated measurement errors, as well as proclivity towards male-dominated standardization samples; "thus, any gender bias in the constructs can be expected to carry over into the tests" (Chrisler & McCreary, 2010, p. 81).

Unfortunately, these biases are bolstered by the substantial lack of literature addressing psychological assessment with any gender outside of the binary norm, with instruments often being "based disproportionately on [binary] male and female samples" (Brabender & Mihura, 2016, p. 23). When psychological assessments do address binary gender, the level of complexity present in the gender differences is apparent, shown through differences in endorsed personality traits (such as on the NEO-PI-R), the attempted creation of gender-based norms (as was done with the MMPI), items attempting to address gender expression, but conflating this with sexual orientation (also present in the MMPI), or assumptions in historical experiences being gender-biased (argued to be present in the MMPI-2's F and Fp scales; Brabender & Mihura, 2016;

Chrisler & McCreary, 2010). Even heavily used forensic tools, such as the Psychopathy Checklist-Revised (PCL-R) may not be diagnostically accurate for females, as the data was originally normed upon male prisoners who were deemed to be psychopathic (Chrisler & McCreary, 2010).

Gender in Violence Risk Assessment

Risk does not have one universal definition, and the predictive validity of a risk assessment may shift based on the length of detainment in question, the severity of the historical violence, and subjective details of each case (Melton et al., 2007). The highest predictive utility of a risk assessment tool is achieved when risk factors and protective factors are taken into account, the exposure and sensitivity to which differ amongst genders (Otto & Douglas, 2011). By focusing on these differences, the necessity of utilizing assessments for their intended, normative population becomes clear. By utilizing a risk assessment tool that is responsive to the person as a whole, including their gender identity, the clinician can gather subjective empirical data to see each individual through a holistic lens and create a full idea of the true risk present within the offender's life (Melton et al., 2007). Given the necessary empirical data required to promote gender-responsive modifications to psychological assessment, they can be viewed as outdated in comparison to the policies adopted by the psychological field as a whole, which are inclusive to the modern idea of expansive gender (Chrisler & McCreary, 2010).

When examining gender in psychological assessments, the individual's self-assignment of gender, the assumed perception of the individual's gender, and cultural labels of gender all affect biases and assessment procedures. Although only the first tends to directly be inquired within an assessment, all three have the ability to lead to biases in the interpretation of the assessment data (Chrisler & McCreary, 2010). These considerations are being incorporated into

nation-wide policies in efforts to eliminate biases, and therefore errors, in assessing for risk. The National Institute of Corrections released a bulletin detailing the importance of revising existing protocols to address the lack of gender-responsive, or gender-informed, policies in order to "create a more effective and efficient correctional approach" (p. 1), in a research-driven initiative that:

> Takes into consideration the unique pathways that lead women to commit crimes and are trauma-informed, strengths-based, and culturally competent. The term [gender-responsive] is not meant to be exclusive, but rather stresses the importance of recognizing the gendered differences for both men and women in psychological development, socialization, exposure to trauma, and cultural, racial, and class-based experiences. (p. 2)

Different pathways to crime result in different needs for the offenders, most of which are ignored when a non-gender-responsive approach is utilized (King & Foley, 2014). Brown (2017) noted that binary female offenders "score higher on virtually all gender-responsive factors," leaving the debate not for integration of these factors in "women-centered corrections philosophy," (p. 6) but for these factors being actively incorporated into risk assessment tools.

There is agreement among criminology and criminal justice researchers that gender differences in crime exist; however, there is a divergence of opinion when addressing the formulation of these differences (Taxman, 2016). When examining binary gender, some researchers substantiate the innate differences between a female's life and a male's life prior to offending, implying that the pathways to criminal behavior, as well as expressions and means of violence, therefore differ by gender (Taxman, 2016; Otto & Douglas, 2011). Some factors that have been associated with binary gender pathways include experiencing abuse, exposure to violence, substance abuse, or homelessness experienced throughout the offender's life. Other factors are considered non-gendered or appear to affect both binary genders equally, such as crimes that are economically motivated (Otto & Douglas, 2011). The goal then becomes incorporating these different factors into the individual's assessment and treatment to achieve

optimal interventions for each offender, while still adhering to evidence-based interventions (Taxman, 2016).

Terms Addressing Gender in Risk Assessment

Although the terms gender-responsive and gender-informed are used interchangeably throughout the literature, feminist theorists advocate the use of gender-responsive, as it is thought to imply empowerment over passive acceptance and inclusion (Taxman, 2016). A gender-responsive approach is multidimensional and based on theoretical perspectives that address the aforementioned gendered pathways to build self-efficacy (Bloom et al., 2003). In treatment, this approach utilizes empathic and empowering behavioral interventions, supported by social learning theory and feminist philosophies, to create optimal responses (Taxman, 2016).

When addressing gender, three types of risk factors primarily emerge: gender-neutral, gender-salient, and gender-specific. Gender-neutral risk factors have the same predictive accuracy for both binary genders, inclusive to the domains of criminal peers and attitudes, employment, and community functioning (Brown, 2017). The existing bodies of research are largely composed of gender-neutral risk factors, with progressive studies showing that some factors, largely touted as "primary criminogenic needs for all offenders" are actually less predictive for binary female offenders (Taxman, 2016, p. 233). Bloom et al. (2004) found that "programs or policies declared 'genderless' or 'gender-neutral' are in fact male-based" (p. 36), highlighting the prevalence of sexism in current practices.

In order to correct this issue, research is beginning to endorse the inclusion of risk factors that consider gender differences (i.e., those that are gender-responsive; Brown, 2017). Some researchers have taken into account gender-salient factors, entailing that some risk factors have a stronger prediction for one gender over others, while other researchers are supporting gender

specificity, detailing risk factors solely entailed for the identified gender (Andrews et al., 2012). A broad amount of data exists for gender-neutral risk and correctional treatments, but there has historically been a severe lack of empirical evidence for gender-specific or gender-salient approaches to risk assessment (Andrews et al., 2012). Even when taking into account feminist theorists, "gender neutrality [remains] the rule and gender specificity is the exception" (Andrews et al., 2012, p. 119). Brown (2017) proposes that by promoting mixed-gender samples in further research to prove that specific factors are gender-salient or gender-specific, research identifying gender-responsive factors of recidivism will be bolstered (Brown, 2017).

Gender within Forensic Psychological Research

Given all of the aforementioned factors, it is imperative to ensure that the assessment tools chosen to determine risk are as precise of a fit to the individual in question as possible. The American Psychological Association's Ethics Code has, since 2002, identified that "psychologists are expected to use tests that have been demonstrated to be reliable and valid for the population being tested" (Chrisler & McCreary, 2010, p. 66). Bloom et al. (2004) posed the question of gender's effect on criminal behavior and criminal justice processes, asking readers to:

> Consider this: if gender played no role in criminal behavior and criminal justice processing, then 51.1% of those arrested, convicted, and incarcerated could be expected to be women, as that figure represents the proportion of women in the general population. (p. 13)

However, this is not the case, and the staggering number of binary male offenders is seen to eclipse the needs of other genders (Bloom et al., 2003). Although the binary male holds the majority within the correctional system, the Bureau of Justice Statistics reports a profound shift has occurred with the number of incarcerated binary females, increasing by over 700% between the years of 1980 to 2019 (Incarcerated Women and Girls, 2020).

Prior to the turn of the century, a majority of research was either conducted solely on men, or gender was not factored into the given sample (Cobbina, 2010), with studies often failing to disaggregate data based upon the represented genders (Gobeil et al., 2016). Such studies leave researchers with convoluted data, which was deemed excusable in the past due to the minority representation of females in the overall offender population worldwide (Gobeil et al., 2010). This smaller base rate inherently leads to smaller sample sizes, making gendered treatment effects harder to pinpoint (Blanchette & Brown, 2006). When applied to a non-binary population, this research becomes even more scarce. Despite this difficulty, research supports that pathways and motives for violence differ vastly amongst genders, to the point that "even if [women] commit the same type of crime... [the crime will be] qualitatively and quantitatively different from men's" (Hannah-Moffat, 2009, p. 212).

It is unquestionable that males are drastically overrepresented in samples of criminal offenders (Andrews et al., 2012). This overrepresentation, therefore, leaves the lesser-represented genders, or anything other than the non-binary male, with less attention, and assumedly, fewer resources tailored to the gender's specific needs. Bloom et al. (2003, p. vi) concluded that the "criminal justice system often has difficulty applying to women offenders' policies and procedures that have been designed for male offenders," leaving the system inadequate in correctly assessing the risks or needs for non-binary-male genders.

It is also pertinent to consider the international progress of gender-responsive assessment research. Canada has notably "led the way in the development of offender assessment," particularly that addressing gender (Caulfield, 2010, p. 322), with the Canadian government initiating the aforementioned Task Force on Federally Sentenced Women nearly three decades ago, examining the management and policies addressing binary female offenders and mandating

changes to encourage female empowerment in their correctional systems (Government of Canada, 2019).

A point of consideration is that, historically, the limited number of female offenders represented in intra-study comparisons of correctional systems resulted in confidence intervals that are drastically lower than that of their male counterparts (Andrews et al., 2012). Having a binary gender underrepresented in the data is egregious in its own right, but when consideration for non-binary genders is taken into account, there appears to be a "forced choice" to choose one of the binary labels, which in turn fuels interpretation and statistical analysis to only be represented by two genders (Chrisler & McCreary, 2010, p. 65). Chrisler and McCreary (2010) identified that "most psychological tests ignore the issues of people who define their gender differently than their biological sex…transgendered [or gender-expansive] people...have neither guidance nor visibility of standardized psychological tests" (p. 65).

Hannah-Moffat (2009) critiqued criminological and correctional disciplines for their treatment of non-binary-male offenders as "afterthoughts," and for the "uncritical use of male norms in the management and treatment" (p. 212) of offenders. This critique provides the foreground for a critical reevaluation of the role gender plays in the development of risk assessment tools, and necessitates not only a gender-informed instrument, but one that was intentionally created for the population on which it will be utilized (Hannah-Moffat, 2009). Andrews et al. (2012) framed the importance of evaluating risk factors through a gender-responsive framework, not because a factor doesn't exist across genders, but because the discrepancies between opportunities, equality, and discrimination provide different social advantages for different genders. According to Andrews et al. (2012), "it would be absurd to

deny the existence of gendered social disadvantage or the importance of efforts to reduce it" (p. 117).

Nearly a decade ago (i.e., 2008), in the first large-scale study unquestionably supporting gender specificity, Brown and Motiuk examined individual items of predictive validity. This study discovered that a greater percentage of risk factors used to predict criminal recidivism (53%) held predictive evidence of gender specificity (e.g., mental illness or history of victimization), whilst a lesser percentage (47%) held predictive evidence of gender neutrality (e.g., offender demographics or criminal history; Andrews et al., 2012; Marshall & Miller, 2018). The authors concluded that gender-neutral measures have the capacity to become gender-informed, should adequate gender-responsivity shifts to measures occur; however, further primary research is required (Brown & Motiuk, 2008). Despite years of this data existing, only small steps have been made towards equality in policy and practice.

The current research on gender-responsivity focuses on global queries, comparing entire risk assessment tools to one another instead of homing in on which aspects of individual assessment tools may work better for non-binary-male genders (Brown, 2017). There are a variety of smaller modifications that affect the responsivity of any given measure, such as normative populations, included factors, the scales that these factors form, or methods of interpreting the resulting data. When risk assessment tools are not built to measure gender-specific factors, they are not likely to capture these factors within the results, and therefore lead to frequent misclassification of risk (Taxman, 2016). When risk is misclassified, the offender is then not able to receive the proper interventions for treatment, leading to an increased probability of recidivism (Taxman, 2016). This literature directly emphasizes that the improper use of a tool

can not only have a negative effect on the individual being assessed, but the community within which they exist may be placed at greater risk.

In order to balance the gender discrepancy amongst normative samples, Brabender and Mihura (2016) suggested that, in lieu of available normative test data being at the discretion of individual constructors, a consensus should be reached amongst practicing constructors and assessors to determine the minimization of gender-based biases. This would potentially be done by deciding which circumstances gender-specific or gender-salient information should or could be utilized (Brabender & Mihura, 2016).

Gender-Responsive Modifications to Risk Assessment Tools

The direction that risk assessment development has taken, in an attempt to be more inclusive to individuals outside of the binary, cisgender male, has been towards that of gender neutrality (Andrews et al., 2012). Yet, as previously mentioned, to adopt a gender-neutral approach is to make the erroneous decision that all circumstances and individual characteristics that ultimately predict criminal behavior are similar enough to be deemed neutral between males and females (Andrews et al., 2012). Bloom et al. (2003) identified that “although many respondents discussed differences between men and women offenders in terms of needs and risks to institutional and public safety, few states have incorporated these differences in objective classification instruments” (p. 17). In general, research demonstrates that the predictive validity of an assessment tool is largely reliant upon its use within the intended population (Taxman, 2016). With risk assessment tools specifically, the “quality of prediction is the primary concern,” inclusive to the “selection, weighting, and tailoring” of items to the population on which they will be used, or “gender-responsive calibration" (Taxman, 2016, p. 356). In Chrisler and McCreary’s publication from 2010, they exemplify the critical approach researchers take in

attempted gender-responsive adaptations, describing the utilization of norms that claimed to be gender-based (as is utilized in the Minnesota Multiphasic Personality Inventory [both MMPI and MMPI-2]), gender-neutral scales that ignore any differences between male versus female respondents (as is used within the Level of Service Inventory - Revised [LSI-R]), and the blatant omission of gender as an explicit risk factor (as is the choice within the Post-Conviction Risk Assessment, PCRA; p. 69).

Out of the aforementioned approaches, the utilization of gender-based norms, appears to be the most commonly used. However, the interpretations that are provided are still non-gendered, meaning that "the same interpretation is provided for different levels of behavior, an obvious source of bias" (Chrisler & McCreary, 2010, p. 69). A limitation of attempting to validate a preexisting risk assessment tool with a population outside of the established normative population is that it does not shift the empirically and theoretically androcentric basis on which the assessment was initially developed and therefore is not an equivalent measure when utilized outside of its normative population (Hannah-Moffat, 2009). Taxman (2016) argues that in order "to create a criminal justice environment that reflects the realities of justice-involved women, while simultaneously addressing their individual needs and strengths, traditional, male-based risk assessments are simply insufficient" (p. 222).

However, these are considerations that must still be taken into account. By foregoing the use of gender-responsive risk factors in the creation of risk assessment tools, the risk of misclassifying the non-binary-male offender increases (Taxman, 2016). This "overclassification, which assigns women to higher risk levels than is behaviorally warranted, frequently occurs with male-based institutional custody classification instruments, such as the NIC Model Prisons classification" (Taxman, 2016, p. 233). However, to ignore all prior established violence risk

assessment tools would also be a disservice to our referral questions and the patients that they so vastly impact. Brown (2017) identified that "there is no agreed upon statistical litmus test for ascertaining whether gender differences exist" (p. 8), as currently, researchers often choose to "disaggregate the data by gender and conduct separate analyses for each" (p. 8). Although this is not ideal, it is representative of forward-thinking shifts to account for gender in some capacity.

An Example of Gender-Responsivity within Contemporary Risk Assessment

Through the research conducted for this background and rationale, one example of a risk assessment tool continued to surface that incorporated ideal gender-responsive modifications. This was the Women's Risk/Needs Assessment (WRNA) suite, which consists of an individual instrument and a trailer that may be utilized with pre-existing violence risk assessment instruments. Through the completion of this systematic review, we intend to find violence risk assessment tools that may be classified as gender-responsive (or the suggestions that are made in order for tools to accommodate gender-responsivity) and to gather these suggestions to benefit both practicing clinicians and future research endeavors exploring this topic.

The Women's Risk/Needs Assessment (WRNA) suite, developed in 2008 as a collaboration between the National Institute of Corrections and the University of Cincinnati, consists of both a formal assessment instrument and a trailer (the Women's Risk/Needs Assessment Trailer, WRNA-T; Women's Risk Needs Assessment Research, 2019). The suite is described to contain "the only validated, peer-reviewed risk/need instruments in the public domain specifically designed by and for system-impacted women" (Women's Risk Needs Assessment Research, 2019, p. 1). It is composed of both gender-neutral and gender-responsive risk factors, which have not been established as female-salient or female-specific but have still been shown to accurately predict female offenders' recidivism (Brown, 2017). Focal points of

trauma, interpersonal relationships, stress associated with parenting, and safety, matched with a strength-based focus, have been correlated with positive treatment outcomes (Taxman, 2016). The WRNA-T is often paired with other widely utilized risk assessment tools, such as the Level of Service Inventory - Revised (LSI-R) and the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), supplying additional scales that "provide contextual details that contribute to gender responsivity" (Taxman, 2016, p. 347).

Taxman (2016) argues that the most accurate appraisal of a woman's risk will be achieved by utilizing risk assessment instruments that are built solely for women, as the inclusion of statistically predictive gender-responsive risk factors is the only way to "see true and accurate measures of women's criminogenic risk" (p. 232). The WRNA suite fulfills this ideal; however, this is not the reality for a wide variety of assessment tools. Therefore, this dissertation's systematic review focused on risk assessment tools for adults that have incorporated gender-responsive modifications.

Summary of Proposed Need for Gender-Responsive Modifications

The need for conducting this systematic review arose among the limited, yet at times conflictual, research on gender-responsive risk assessment tools. When examining gender-responsive modifications to psychological assessments, a majority of the literature focuses upon binary male normative samples, largely due to their gender majority within correctional systems, but also due to the assumptions of male normative data applying equally to females (Andrews et al., 2012; Bloom et al, 2003; Hannah-Moffat, 2009; NIC, 2014; Taxman, 2016). The emergence of binary-female-driven data in the literature accentuates the fact that rarely any literature addresses gender outside of the binary norm (Brabender & Mihura, 2016).

As a shift in awareness for this need has occurred over approximately the last two decades, the path to examining non-binary male offenders has been slow to form. A majority of the contributing literature focuses upon gender-neutral interventions, with few studies attending to solely females, much less any non-binary gendered individuals (Gobeil et al., 2016). These gender-neutral risk instruments, based upon gender-neutral risk factors, "appear to be less valid for women offenders, particularly those who follow gendered pathways to offending" (Taxman, 2016, p. 232).

Gender-responsive pathways, often determining exposure to gender-responsive risk factors, instead provide a more holistic view of the individual, allowing the clinician a clear idea of the true risk present within the offender's life (Melton et al., 2007). The research found that by incorporating gender-responsive pathways and considerations into institutional and community settings, better outcomes were found for both binary male and female genders (Bloom et al., 2003).

Amongst the significant rise in gender-based research, starting in 1998 (Chrisler & McCreary, 2010), the National Institute of Corrections (NIC) began a multi-year endeavor (1999-2002) to expand correctional institutions' knowledge of gender-responsive interventions and management of female offenders (Bloom et al., 2003). The NIC noted a variety of conditions that required the initiation of gender-responsive policy development, including:

- growth in the number of incarcerated females was faster than that of males,
- gender-responsive training of employees was viewed to be lacking,
- significant growth in research on females across mental health fields,
- and public interest in gender-responsivity, specifically for children of the incarcerated (Bloom et al., 2003).

With the United Nations' creation of "the Bangkok Rules" establishing mandatory research on women's criminal pathways and risk factors (Taxman, 2016) and the National Institute of Corrections' 2014 survey reporting that nearly three-quarters of the jurisdictions in their survey had created policies that were gender-responsive for those offenders that identify as female (King & Foley, 2014), these changes have started to become more promising. With the drive for further empirical data and policy changes, strong bases for risk assessment tools may now be available for researchers. On a national level, program evaluations of gender-responsivity have been activated, including "Moving On," "the Women Offender Case Management Model," and "Seeking Safety," which have demonstrated the "effectiveness of adult correctional-based interventions for women targeting their substance abuse and psychological well-being" (Taxman, 2016, p. 235).

Further research surrounding gender issues within risk remains a necessity (Andrews et al., 2012). When accommodations to the existing risk assessment process are added, it is imperative that they are reflective of that gender's risk factors. This research tends to be limited in both "theoretical and conceptual scope," even for the binary female population (Taxman, 2016, p. 228). As has been identified, the best risk appraisal for a specific gender will occur through the utilization of a risk assessment tool that was built for that gender, as it was "never scientifically defensible to add women and stir" (Taxman, 2016, p. 228).

With the aforementioned development of the WRNA suite, an assessment suite that was built for female offenders, research has supported that other existing risk assessment tools may be under classifying binary female offender's risk (Taxman, 2016). Despite the lower base rate, the drastic proportional increase in female offenders over the last four decades necessitates the

same ethical and legal obligations apply to managing and treating the presented risk (Taxman, 2016; Incarcerated Women and Girls, 2020).

Risk assessment tools carry some of the highest consequential outcomes of any psychological assessments (Melton et al., 2007), and addressing their normative values is pertinent on both an ethical and moral level for practitioners (Otto & Douglas, 2011; Melton et al., 2007). As identified by Chrisler and McCreary (2010), "forensic researchers have an obligation to adhere to the best possible methodological standards and to the practice and policy standards of the field to minimize [the] risk that bad science causes harm to the public" (p. 589). Hence, it is the clinician's responsibility to administer a risk assessment tool with a normative population that, if not matched to, at least takes into consideration the individual's gender.

Research Questions

It was the goal of this dissertation, therefore, to systematically review all risk assessment tools that have incorporated gender-responsive components to expand their normative population to encompass genders outside of the binary male. This systematic research review potentially lays the groundwork for clinicians to reference when utilizing risk assessment tools for a non-binary-male patient. The research questions for the dissertation included:

1. According to the available literature, what gender-responsive modifications have been made to which risk assessment tools?

2. Which gender-responsive modifications have been proposed for future changes?

3. How does the aforementioned literature define gender in discussing these gender-responsive modifications?

Chapter 2: Methodology

Systematic Review Approach

The goal of a systematic review is to provide a "methodical, replicable, and transparent approach" (Siddaway et al., 2019, p. 749) to comprehensively gather all existing data on a particular research subject in order to critique the extent and quality of available information. When done properly, a systematic review is able to link evidence and theory, synthesizing an extractable idea upon which future research can build (Siddaway et al., 2019). The five-stage process of an integrative review includes: "(1) problem formulation, (2) data collection or literature search, (3) evaluation of data, (4) data analysis, and (5) interpretation and presentation of results" (Russell, 2005, p. 8).

This systematic review examined three research questions: (1) According to the available literature, what gender-responsive modifications have been made to which violence risk assessment tools and risk/need assessment tools? (2) Which gender-responsive modifications have been proposed for future changes? (3) How does the aforementioned literature define gender in discussing these gender-responsive modifications?

Given the novel nature of this review's research questions, a wide variety of evidence was evaluated in order to cultivate an encompassing view of the current research that is able to "capture the context, processes, and subjective elements of the topic" (Whittemore & Knafl, 2005, p. 552). Therefore, an integrative review approach appeared to be the most appropriate due to its incorporation of both quantitative and qualitative empirical studies, as well as applicable theoretical literature (Whittemore & Knafl, 2005). This technique allows researchers to examine what information is known, the quality of this information, as well as gaps in the literature, and allows researchers to propose potential steps to progress research (Russell, 2005).

Eligibility Criteria

In this systematic review, publication sources that were eligible for inclusion included published peer-reviewed journal articles and Canadian government documents. While government entities within the United States have started to research this pertinent field as well, Canada has remained a frontrunner in these matters with the development of widely accepted and utilized models and assessments addressing gender-responsivity, publishing government documents addressing each research project as it unfurls (Gobeil & Blanchette, 2007; Van Voorhis et al., 2010). Therefore, in deciding to expand beyond the inclusion of peer-reviewed and published articles, the unparalleled quality of novel research expanded within the Canadian government deemed these available resources to be adequate for inclusion as well. Sources must also have been published in English and had a publication date no earlier than 1998 in order to be eligible for inclusion.

The sources chosen to be examined within this systematic review must have contained research variables inclusive to assessment tools addressing violent risk or risk/need, gender-responsivity, and a forensic population. A forensic population is defined by the American Psychological Association to be "composed broadly of individuals who may present with a psychiatric diagnosis or may have other characteristics that are relevant to a clinical-legal decision and who are involved with the judicial system" (Forensic Psychology, 2019, p. 1). Forensic populations are often divided into either "civil" or "criminal" populations, inferring the litigation or legal proceedings with which the individual is involved (Forensic Psychology, 2019, p. 1). As this definition of a forensic population remains broad, individuals defined as such may be associated with a variety of institutions or judicial bodies, likely addressing pre-, post-, and/or current incarceration or legal issues.

The sources chosen to be examined within this systematic review must have contained risk and need assessment tools that had either already been modified or had suggestions for modification specifically addressing individuals who identified their gender to be outside of a binary, cisgender male. This criterion was inclusive to any other identification amongst the spectrum of gender identities, which may include those who identify as binary cisgender female, transgender, bigender or trigender, non-gendered, other-gendered, or any genderfluid identification. The sex of the individuals can be male or female, and there is no limitation to sexual preference. Individuals addressed must be adults, defined as 18 years of age or older. The individual's status of involvement within the legal system must abide by the aforementioned definition of a forensic population. The data examined spanned from 1998 to present, as made pertinent by Chrisler and McCreary's (2010) content analysis of published gender research.

Workflow for Selection, Screening, and Quality Appraisal

The order in which the selection, screening, and quality appraisal processes occurred is depicted in the "PRISMA Workflow Diagram" below. It includes the selection stage, the screening stage, the quality appraisal stage, and the data extraction/inclusion stage.

Search, Screening, and Selection Processes

The sources chosen to be examined within this systematic review were obtained from a variety of psychological databases, criminal justice databases, academic databases, legal journals, textbooks, government publications, and presentations. The online databases that were originally identified included: WorldCat, PsycINFO, PsychARTICLES, HeinOnline, Academic Search Complete, ProQuest: GenderWatch, ProQuest: Sociological Abstracts, Sage Journals, and PubMed. However, ProQuest Sociological Abstracts was discarded as a database upon completion of screening, due to each retrieved article being an abstract of the original article.

In order to acquire the proper sources fitting the aforementioned criteria, the original Boolean search string was: *gender* AND ("risk and need" OR "risk-need" OR "RNA" OR "R.N.A." OR "risk/need") AND assessment* AND ("offender*" OR "prison*" OR "ex-offender*" OR "forensic*" OR "correction*" OR "criminal justice")*. However, upon completing an initial pilot review of this search within two randomly chosen databases (PsycINFO and WorldCat), we realized that this search string required refinement in order to garner more results addressing expansive views of gender, ideally encompassing that beyond a binary view, as well as a more encompassing view of the types of risk assessments that we wished to address.

Therefore, the updated Boolean search became: *gender* AND ("risk and need" OR "risk-need" OR "RNA" OR "R.N.A." OR "risk/need" OR "risk assessment" OR "violen* risk*" OR "violen* assess*") AND assessment* AND ("non-binar*" OR "gender expan*" OR "gender-expan*" OR "transgender*" OR "trans*" OR "Gender-Inform*" OR "Gender-Respon*" OR "gender-specific" OR "gender specific*" OR "gender respons*" OR "gender informed*" OR "gender fluid*" OR "gender-fluid*" OR "gender divers*" OR "gender-divers*" OR "genderqueer*" OR "gender-non-conform*" OR "gender non-conform*" OR "gender incongruen*" OR "gender-incongruent*") AND ("offender*" OR "prison*" OR "ex-offender*" OR "forensic*" OR "correction*" OR "criminal justice")* was used to identify key facets within articles that were pertinent to this systematic review. When applicable, the specified parameters were “peer reviewed” and had a custom publication range from the dates of “1998 to present.” This Boolean search, with the specified parameters, was entered into each database. The process of recording these individualized search terms, utilizing them in a search plan for the aforementioned databases, and record of which databases and search terms have been used was included within the Search Record (Appendix A).

The resulting studies then entered the first step of screening, removing any noticed duplicates. After these duplicates were excluded, the remaining articles were then further screened based on a series of steps. This screening process included an examination of the title, abstract, and overview of the full article for the following inclusion/exclusion criteria:

- Articles must have been published in the English language.
- Articles must have been published in 1998 or later.
- Articles must have addressed an adult population (18 years or older).
- Article must have been peer-reviewed.
- A violence risk assessment or risk/need assessment tool addressing a forensic population must have been identified within the article.
- The study must have addressed a gender-responsive modification to the identified assessment tool (e.g., changes to administration, changes to the development of the tool, creating a tool solely for non-binary males, etc.).

In order to verify that both the primary researcher and the research assistant were applying the inclusion and exclusion criteria in the same manner, a secondary pilot review was conducted. Ten articles were chosen by the primary researcher after reviewing the title and abstract that appeared to contain information relevant to this systematic review. Both the primary researcher and the research assistant were provided these articles, along with the Inclusion/Exclusion Criteria Checklist sheet (see Appendix D) to complete independently. The results were then directly compared, with disagreement occurring between the two researchers on three of the ten articles. After discussion, both researchers agreed on the status of inclusion or exclusion of the three articles, garnering a detailed understanding of both the material and the process for both parties. The items that required further discussion or clarification included: the

nature of the identified tool (i.e., was the tool a violence risk assessment or a risk and need assessment tool); the nature of the risk addressed (i.e., violence towards others, excluding tools that are solely examining sexual violence or violence towards the self); and the nature of the addressed modification (i.e., was gender-responsivity actually discussed within this resource). Due to this exercise, both of the researchers were able to clarify their understanding of the inclusion and exclusion criteria. It was collaboratively determined that, should any questions arise as to whether a specific resource may or may not be worthy of inclusion, a verification option would be added to the inclusion and exclusion form to indicate that further screening and/or discussion was deemed necessary. This allowed for the screening mechanism to be of co-collaboration between the primary researcher and the research assistant. This pilot process was deemed essential by the participating researchers, as it inherently allowed for greater clarity and, ultimately, an assumed elevation in interrater agreement with the utilization of the screener.

When applied to the full screening process, the initial examining researcher would determine if a specific resource was worthy of inclusion, should be excluded, or if it required further verification. Each of the articles that were deemed to be worthy of inclusion or required further verification were then double verified by an alternate researcher, either the primary investigator or a research assistant. In this process of double verification, the researchers agreed to meet on a biweekly basis in order to address any discrepancies. In total, 96 resources were double verified, 94 of which were originally found to be worthy of inclusion, and two of which had been marked for verification. In total, 66 additional resources were double screened, but determined not to meet criteria for inclusion. This was due to the resource not mentioned gender-responsive modifications (33 resources); no violence risk assessment or risk and need assessment tool identified (15 resources); the gender-responsive modifications mentioned were for

treatment, policy, or programming (11 resources); females were only identified as the victim (1 resource); and the types of risk addressed were outside of limitations defined (i.e., risk for substance use, risk for sexual deviancy, risk for self-harm, or health risks; 6 resources).

While all attempts were made to control for biases, it is important to acknowledge that the primary researcher had additional knowledge regarding the subject matter due to specialized education and clinical experience. The nature of a power differential between a primary researcher and a research assistant was openly discussed, and the research assistant was encouraged to vocalize any contrary ideas or opinions. At no point was an agreement unable to be reached after discussion. All included and excluded articles were noted in the Dissertation Database (Appendix B) under the "Screening and Selection" tab.

Quality Appraisal of Individual Studies

Once all of the sources were screened for inclusion into the systematic review, the quality of each included study was then critiqued. This critique assessed overall quality using two tools: the Mixed Methods Appraisal Tool (MMAT) and the University of Oxford's Centre for Evidence-Based Medicine's Systematic Reviews Critical Appraisal Sheet (CEBM Tool). The goal of this step was to assess the quality of each individual study to determine if that study is worthy of undergoing data extraction, the next step of the systematic review process. In an attempt to control for any bias, particularly due to the limited resources appearing to address the research questions, a third researcher (henceforth known as the "QA research assistant") was brought in to appraise the quality of the resources that had been screened thus far. A thorough training was provided by the primary researcher to utilize each of the appraisal tools. A pilot trial of five studies was conducted, including resources that would utilize both the MMAT and the CEBM tool. The primary researcher and the QA research assistant retained consensus on the

inclusion of all but one resource. After discussion, consensus was reached. This allowed for a thorough understanding of the material and tools. The primary researcher and QA research assistant likewise held biweekly meetings in order to address any questions or concerns that had surfaced.

The MMAT is a critical appraisal tool designed to gauge the quality of research in five different categories of studies, including: qualitative research, randomized controlled trials, non-randomized studies, quantitative descriptive studies, and mixed method studies. While the MMAT was originally created in 2006, it was reiterated in both 2011 and 2018 (Hong et al., 2018). Since it was developed approximately 15 years ago, it has gained widespread traction in the research community and has been utilized within a multitude of systematic reviews (Hong et al., 2018). This tool was chosen for use in the present dissertation study because of the MMAT's supportive studies exhibiting high levels of interrater reliability, content validity, usability, and efficiency (Hong et al., 2018).

For the purposes of this systematic review, we utilized the most recent version of this appraisal tool. While the 2018 updated version encouraged users to take a qualitative approach to analyzing the quality of the study, the authors noted that the pre-existing method of scoring quantitatively was still adequate (Hong et al., 2018). Therefore, for the ease of coding, each study was separately examined in a holistic manner while still assigned a numerical value to determine inclusion or exclusion, following the identified coding instructions as detailed within the previous edition.

The MMAT consists of two sections, Part I is a two-item checklist and Part II consists of five tailored criteria, sorted by category of study design. The two items within Part I consisted of "Are there clear research questions" and "do the collected data allow you to address the research

questions?" For details on Part II, please see Appendix E to examine each individual methodological quality criteria as organized by category. In order to progress to Part II, both items in Part I must have been answered with a "Yes." For the remaining criteria in Part II, criteria can qualify as "Yes," "No," or "Can't Tell." Only criteria that received a "Yes" was awarded points. If all five criteria received affirmative responses, then the study was granted 100/100. Therefore, in order for a study to be included within our systematic review, only one criterion was allowed to be marked as "Can't Tell" or "No", or a minimum of 80/100 points.

As the MMAT specifically defines its parameters of use to be that for empirical studies, an alternate validated appraisal tool was needed to assess systematic reviews in the current study. Therefore, we chose the CEBM Tool to assess the small amount of systematic reviews that had passed screening. The University of Oxford's Centre for Evidence-Based Medicine has created a number of critical appraisal tools to ensure that researchers are critically appraising the reliability, validity, importance, and applicability of clinical evidence. The tool addressing Systematic Reviews specifically has been translated into six different languages (Systematic Reviews Critical Appraisal Sheet, n.d.). The CEBM Tool similarly consists of one initial screening question ("Is the main question clearly stated") followed by five criteria that can receive a response of "Yes," "No," or "Unclear." These criteria include items addressing the likelihood of relevant studies being missed, the appropriateness of selection criteria, the validity of included studies given the research question(s), and the similarity of results. In order for the study to be considered, the initial screening question must have been answered in the affirmative. For each affirmative response on the remaining criteria, 20/100 points were awarded. Therefore, if all five of the secondary criteria received affirmative responses, then the study was granted 100/100. In order for a study to be included within our systematic review, only one criterion was

allowed to be marked as “Can’t Tell” or “No”, or a minimum of 80/100 points. Please refer to the aforementioned Appendix F to further examine the evaluative criteria.

The areas of research to be included within this systematic review entailed a broader range of sources than just empirical studies and systematic reviews. Given the novelty of this area of research, literature reviews were deemed to be important contributors of data within the available sources. Unfortunately, neither existing quality appraisal tool fits the literature review source type. The CEBM Tool addresses quality by inquiring about the encompassing nature of the review, the inclusion criteria, and the quality appraisal process of that specific review; therefore, given the nature of literature review process, any literature review would have automatically resulted in a disqualification from the screening process if the CEBM Tool was used. This researcher determined that the best course of action to determine quality for literature reviews was to adhere to the original inclusion criteria mandated in the primary screening process. Therefore, if a literature review had previously passed inclusion criteria, it was deemed acceptable for inclusion in this systematic review, despite the limitation of not having a quality appraisal check.

Likewise, additional resources, such as government publications, were not appropriate for the available quality appraisal tools, as they most closely resembled literature reviews. Therefore, the judgment for inclusion took into consideration the creator of the source. It was assumed that government agencies were publishing appropriately researched literature, given the weight and influence that their publicized opinion may hold.

The information from both appraisal tools, as well as the outlier resources, was then collected and coded onto the Dissertation Database (see Appendix B) under the “Quality Appraisal Review” tab. For an article to pass the critical appraisal process, it must have received

a score of 80 or 100, as both tools scored in increments of 20. Those that are 80 or higher were added to the "Included Sources" portion to move on to Data Extraction, while those that are less than 80 were moved to the "Excluded Sources" tab of the Dissertation Database. The outlier resources (e.g., the literature reviews and government publications), were placed in an additional section to acknowledge inclusion. To ensure that this process was thorough and objective, an intermittent peer review occurred for all articles at this stage. Due to potential subjectivity in the quality appraisal process, it was not expected that the peer reviewer and the principal reviewer would be in exact concordance in assigning values. After the initial coding had occurred, any reconciliation between differences in quality ratings were discussed and an agreement was reached between raters.

Although this process did not fully eliminate biases, it allowed for less discretion amongst individual contributors. An audit trail was kept, documenting any disagreement amongst reviewers and the resolution process of such disagreements. Reconciliation during the quality appraisal process was required on 9 of the resources due to questions regarding the fulfillment of criteria. Each of these was reviewed by the primary researcher, thoroughly discussed between the primary researcher and the QA research assistant, and consensus was reached across each resource.

Ultimately, 96 total studies were appraised for quality. Eighty-four studies underwent evaluation with the MMAT tool, 3 studies underwent evaluation with the CEBM tool, and 9 resources were determined to qualify as literature reviews (including the two government resources). The CEBM tool deemed the 3 systematic reviews to be inappropriate for inclusion, largely due to lack of clarity in their own inclusion or exclusion criteria and the perceived likelihood that relevant studies were missed due to the limitations of the search process and criteria. Of the

screened MMAT tools, 3 of the studies were deemed inappropriate for inclusion, due to lack of clarity in participant selection, lack of clarity in methods of statistical analysis, and lack of accountability for confounding factors. Therefore, 90 studies total entered the final screening process. For further details, please refer to the Results section (Results of Individual and Overall Study Quality subsection).

Final Screening Process

At this point in the process, it was noted that further duplicate resources required screening from the included resources. This was largely due to additional authors being noted, different publication years being noted, or the human error or misspelled titles. After this screening was conducted, 27 more resources were removed. The remaining 63 resources were then thoroughly reviewed to verify full adherence to the identified criteria. This step further eliminated 4 resources, all of which had appeared to address gender-responsive modifications to risk assessment tools but had ultimately failed to fulfill this specified criterion due lack of true modifications either being suggested or examined. A total of 59 resources passed the final screening process.

Data Collection and Extraction

The variables included within the data collection process highlighted the research questions previously posed: (1) According to the available literature, what gender-responsive modifications have been made to which risk assessment tools? (2) Which gender-responsive modifications have been proposed for future changes? (3) How does the aforementioned literature define gender in discussing these gender-responsive modifications?

In order to garner the relevant findings for these questions, sources that underwent both the screening for title, abstract, and inclusion/exclusion criteria, as well as the quality analysis

process, were then read fully, for the purposes of completing the Data Collection and Extraction Form (Appendix G). At this point, any remaining duplicates were also removed from the sources.

The data extraction form included a wide variety of categories that were meant to capture each of the important points to address in gathering relevant data. In total, the ten categories include: origin information, general information, details of inclusion/exclusion, design characteristics and methodological features, assessment of research variables, study participant characteristics, and relevant quotes, as detailed below:

- Origin Information addressed Document ID, Database Used, Year Published, Type of Publication, Journal, and Country of Research
- General Information included Full Document Title, Authors, In-text Citation, and Full Citation
- Details of Inclusion/Exclusion outlined the Date Completed, if the article appeared to address implemented and researched modifications or if the article only suggested modifications, and any relevant notes to the inclusion/exclusion screening process
- Design Characteristics and Methodological Features detailed the Date Completed, Identified Design, QA Score, and any relevant notes to the quality appraisal process
- Study Participant Characteristics defined Sample Size and Sample Characteristics (Gender, Ethnicity, Age, Forensic Status, Location, etc.)
- Assessment of Research Variables asked: How was gender defined? What risk assessment tools were identified? Suggested gender-responsive modifications? Implemented gender-responsive modifications or direct research conducted regarding gender-responsive modifications to identified risk assessment?

- Relevant Quotes organized all important highlighted quotes from the article as to contribute further to the discussion of the data

The information from each of these forms was then compiled into a master sheet called the Dissertation Database (see Appendix B).

Data Management, Synthesis and Analysis Plan, and Reporting of Results

Data Management

As previously mentioned, the studies that passed the initial screening and achieved the cut-off (80+/100) for the inclusion criteria within the Quality Appraisal forms, as well as the outlier resources that were deemed worthy of inclusion, underwent the Data Extraction process. The information from all of these steps was entered into the Dissertation Database Excel spreadsheet on Google Drive (see Appendix B), which held information from each article in one place.

The Dissertation Database (see Appendix B) contained four tabs, labeled “Screening and Selection,” “Quality Appraisal Review,” “Included Sources,” and “Excluded Sources.” In order to move on to the next tab, the source must have “passed” the previous tab for inclusion. The “Included Sources” tab held all of the different data points as collected from both the Quality Appraisal process and the Data Extraction process. This also acted as the Evidence Base for Research Questions, as all data was included in one page. The “Excluded Sources” tab held only the identifying information of the study (Document ID, Database Used, Authors, Year Published, and Reason for Exclusion). In total, after the two duplicate screenings, inclusion and exclusion screening, and quality appraisal screening, 1692 resources were excluded, and 59 resources remained for full synthesis and analysis.

Synthesis and Analysis

Following the initial quality appraisal, the data collection and extraction process allowed for the accumulation of relevant data to be gathered into the Dissertation Database (see Appendix B). This database was then appraised as both a "Rating of Evidence Reviewed" and as an "Overall Quality of the Systematic Review." The Rating of the Body of Evidence occurred in two phases. Phase one involved viewing the quality appraisal numbers across the board and averaging the overall number to see how strong the articles utilized for the systematic review truly were. Phase two consisted of clustering and comparing relevant data that aided in answering the three research questions. Since this was an integrative systematic review, a qualitative metasynthesis was chosen as the best method for synthesizing information.

Qualitative metasynthesis allows for the researcher to identify specific research questions that guided the search for relevant qualitative evidence. It is a deliberate, systematic process by which the data undergoes selection, appraisal, and refinement in order to best present the evidence for the research questions. This process allows for rigorous synthesis to meld with educated interpretation, presenting a finished product that extends the application and the meaning of the given data (Erwin et al., 2011). By coding the data through a systematic process, themes and patterns are able to be identified for analysis. The benefits of using qualitative coding include improved validity of the analysis due to structured processes, increased awareness of biases due to defined guidelines for coding, and increased transparency for fellow researchers (Essential Guide to Coding Qualitative Data, n.d.).

For this systematic review, a combined approach was utilized, incorporating both inductive and deductive coding processes. After completing the literature review relevant to the background and rationale, specific research questions were identified. This created the

foundation for a collection of codes through which the articles would be screened, including: the way gender was defined (binary versus non-binary), the types of risk assessment tools that would be screened (violent risk, general risk, risk/need), and some forms of the modifications made apparent through the literature review (different cut-off scores, development "from the ground up," etc.; see Code Manual in Appendix C for further detail).

By initially developing this loose collection of deductive coding, the researchers were then able to identify relevant information from the screened sources. While the coding criteria went through some refinement during data extraction, evolving the definitions of the categories, it was largely led by the deductive coding process. After the data was acquired, the primary researcher was able to use knowledge acquired throughout the dissertation process to categorize different patterns and themes by utilizing higher order coding obtained as a result of the original deductive coding, further organizing the extracted information (Saldaña, 2013). While some preconceived notions existed regarding the data that would be found, the researcher remained open to both themes and narratives that began to emerge through the data extraction process. Allowing this flexibility, the deductive coding process provided preconceived categories by which the information could be organized, while the inductive coding process allowed for the creation of new categories of coding to wholly encompass the discovered information. Given the exploratory and novel nature of this research, the combined approach was deemed appropriate to fully encompass the available data into this systematic review.

Through this combined coding process, the initial cycle of coding largely entailed Descriptive Coding, or the summation of content, encapsulating the themes of the foraged data (Saldaña, 2013). After this step was completed, the researchers were left with a variety of data points that was then organized through Thematic Analysis. This process allowed for recurrent

themes and patterns to be discovered within the data while also grouping content of similar coding together (Given, 2008).

The remaining findings from the data extraction were then compared by publication information, methodological features, study participant characteristics, and setting characteristics. Any patterns that surfaced were examined to see if relevant conclusions or inferences could be drawn from the gathered information. The coded information was then reviewed as data to help synthesize qualitative and quantitative information. As mentioned, this combined coding process documented different patterns, lacking information, or anomalies that arose in the data and allowed for an open-minded and thorough examination of the selected literature. In order to limit biases that were potentially present in the data analysis process (e.g., expecting certain results), a detailed audit trail was kept, as recorded within the Dissertation Database's Quality Appraisal section (Appendix B). Any biases that appeared to have been present were then detailed in the Limitations section of the finalized systematic review.

After final review of the included articles, and thereby the risk assessment tools that were identified, three tools were specifically noted to be highly relevant to violence risk assessment, but not directly measuring this quality. The limitations of these remaining included resources are addressed within Chapter 4; however, this information is separated from the final results into their own chart for future researchers to reference. These resources were not removed from the results overall due to their composure (i.e., the questions addressed within each of these three contained violence risk questions) and how commonly these tools are utilized within violence risk assessments.

Reporting of Results

Analyzed data from the appraisal of the Dissertation Database, including responses to the research questions, as well as the combined coding process, were then presented in narrative form in the Results section. Also, visual representations of thematic contact were formed into charts and graphs as found throughout the Results section and within the accompanying appendices.

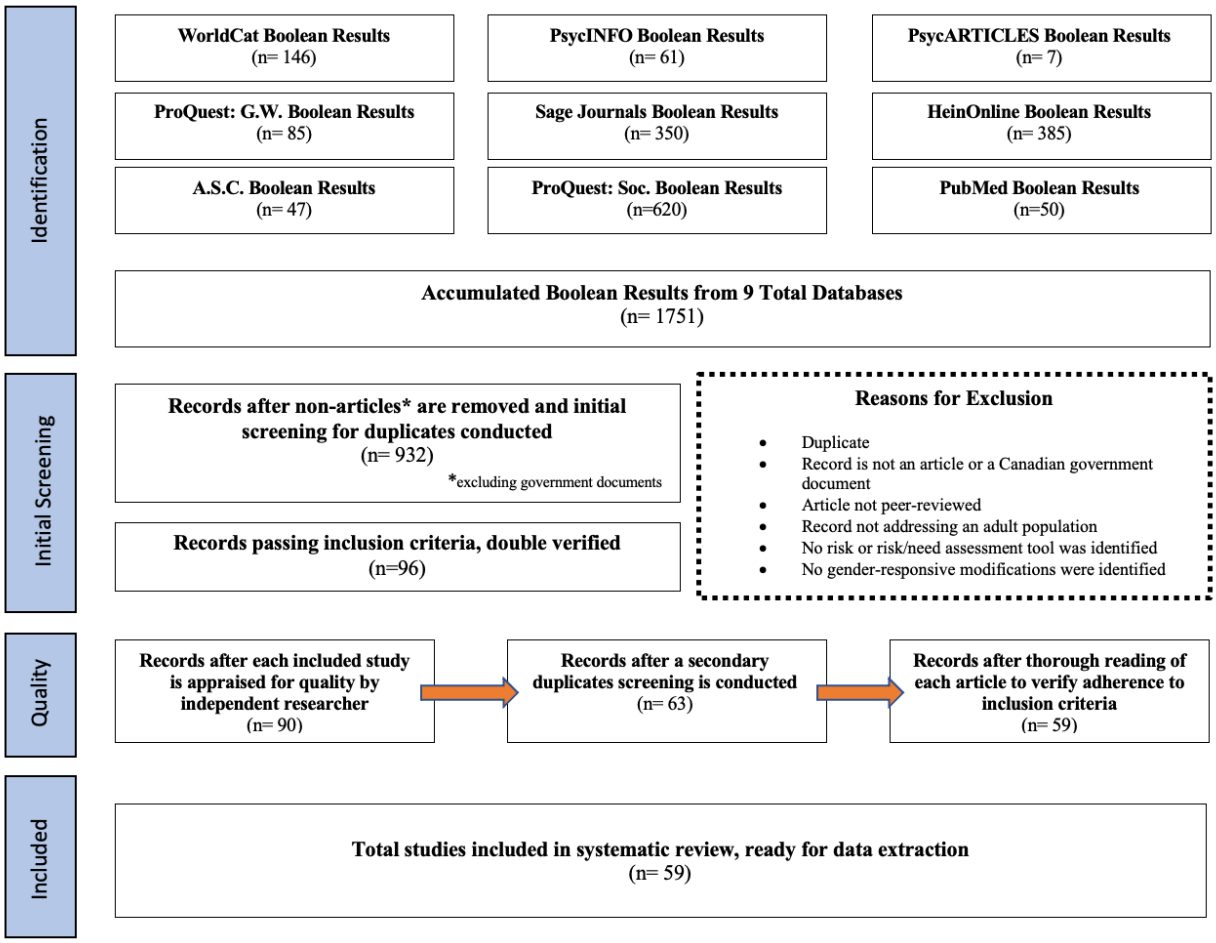
Chapter 3: Results

This results section provides an overview of the 59 included studies in this systematic review of the literature on gender-responsive risk assessment tools, as well as how the results of the analysis addressed the study's three research questions (1) According to the available literature, what gender-responsive modifications have been made to which risk assessment tools? (2) Which gender-responsive modifications have been proposed for future changes? and (3) How does the aforementioned literature define gender in discussing these gender-responsive modifications?

The screening process, as identified within the PRISMA Diagram (see Figure 1 below), refined the original 1751 sources into 59 included studies, each passing inclusion/exclusion criteria and containing information relevant to our assessment of research variables. Upon completing the coding of the data from these eligible articles and documents, the results of the deductive and inductive coding and thematic analyses were conducted. This results section summarizes each of the defined research questions through both visual representations and descriptive explanations.

Figure 1

PRISMA Workflow Diagram



Overview of Included Studies

After completing the full study selection process, 59 studies were included within the analysis portion of the systematic review. Of these, there were 46 non-randomized peer-reviewed articles, 9 literature reviews (2 of which were government published documents addressing policies), 1 qualitative peer-reviewed article, 1 quantitative descriptive peer-reviewed article, and 1 peer-reviewed randomized controlled trial.

Of these studies detailing gender-responsive modifications to risk and risk/need assessment tools, there were 51 implemented modifications, 19 researched modifications, and 50

suggested modifications. 16 of these 59 studies included multiple modalities of modifications, resulting in the increased sum of each type of modification. While statistical analyses were not independently utilized, the researchers calculated frequencies throughout the analysis process in order to identify patterns and themes within the data. The following identified patterns are organized by geolocation, database, and journal.

As organized in respect to frequency, the locations from which these resources originated spanned across:

- the United States (33 modifications);
- Canada (17 modifications);
- the Netherlands (3 modifications);
- New Zealand (3 modifications);
- the United Kingdom (3 modifications);
- Australia (1 modification);
- Finland (1 modification);
- Germany (1 modification).

Addressing databases, the most frequently cited publications were found within:

- HeinOnline (30 sources);
- Sage Journals Online (13 sources);
- PsycINFO (7 sources);
- PubMed (4 sources);
- WorldCat (3 sources);
- Academic Search Complete (2 sources).

Regarding journals, the most frequently cited publications were found within:

- The Criminal Justice and Behavior Journal (30 sources);
- The International Journal of Forensic Mental Health (3 sources);
- Government publications (3 sources);
- The Law and Human Behavior Journal, Crime and Delinquency Journal, and Journal of Contemporary Criminal Justice (2 sources each);
- Social Work in Public Health; Behavioral Sciences and the Law; the Journal of Interpersonal Violence; Federal Probation; Criminal Justice Policy Review; Feminist Criminology; Sexual Abuse; Assessment; Probation Journal; International Journal of Offender Therapy and Comparative Criminology; Criminology and Public Policy; The Prison Journal; Psychiatry, Psychology, and Law; the Journal of Forensic Psychiatry and Psychology; Psychology, Crime, and Law; and the Journal of Sexual Aggression (1 source each).

Addressing overall demographics, 23 of the 59 studies specified examining only a binary female population while none of the studies examined a non-binary population. The remaining were of a binary male and female sample. The average age across the pertinent studies was 33 years old.

When specified, average sample size across the pertinent studies was 8,426, while the median sample size was 645. Available sample sizes ranged from 73 to 128,183.

Addressed ethnicities within these sources, rounded to the nearest tenth of a percentage, included:

- White/Caucasian (19.4 to 79.6%);
- Indigenous/Aboriginal/Native (0.2 to 60.0%);
- Black/African (3.0 to 56.3%);

- Hispanic/Latino (1.0 to 31.5%);
- Asian (0.4 to 4.3%);
- East Indian (2.5%);
- Pacific Populations (8.0%);
- Dutch Origin (84.4 to 87.2%);
- Other Minority (0.3 to 11.3%).

Of note, not all studies addressed each of the aforementioned ethnicities; therefore, only stated percentages were included.

The noted forensic statuses of the populations included:

- Incarcerated individuals (such as federal, prison, Department of Corrections, or jail, spanning all levels of security) with various offenses or convictions (such as felonies and misdemeanors, inclusive to seriously violent, substance-use related, sexual, and homicide);
- Individuals that were pre-sentence, post-incarceration, or commitment;
- Individuals found to be not guilty by reason of insanity (NGRI);
- Forensic Psychiatric Hospitals patients;
- Court-Mandated Residential Treatment Facility patients;
- Individuals on probation, community supervision, or parole.

The locations from which the examined populations hailed spanned the United States (including Arkansas, California, Colorado, Connecticut, Florida, Illinois, Michigan, Minnesota, Missouri, New Mexico, New York, Ohio, Oregon, Texas, Washington, Wisconsin, the "Southwest," "Eastern" states, the "Midwest," and "Southern" states), Canada, the Netherlands, New Zealand, Australia, England and Wales, Finland, Germany, and Pakistan.

For more specific details regarding each of the aforementioned demographics, please refer to Appendix H.

Results of Individual and Overall Study Quality Appraisal

Of the 59 included studies, 9 were identified as literature reviews, with two of these identified as government documents, and the remaining 50 were analyzed for individual quality appraisal with information from the MMAT tool. All included studies must have passed the standard of "80% or above." Two of the included articles (both non-randomized trials) included two different studies, resulting in multiple assigned quality appraisal scores within each article. This resulted in 37 of the included studies fulfilling all quality appraisal criteria, scoring 100%, while 11 studies missed one criteria point through omission or lack of clarity, scoring 80%, and the two remaining resources that held dual studies were averaged for an additional 80% and 90%, respectively (i.e., 80% and 100% averaged to 90% while the alternate resource scored 80% on both studies). Ten of the 80% studies, along with the two dual-study resources, fell within the Non-Randomized Trial categorization, while the remaining study was identified as a Randomized Controlled Trial. These points were missed for: lack of complete outcome data, lack of clarity surrounding the accountability of confounding variables, or lack of clarity regarding the intervention administration procedure. Details of this process can be found within Appendix B.

As previously mentioned, the resulting quality appraisal numbers were averaged to determine the overall strength of the articles utilized for this systematic review; referred to as overall quality appraisal. This analysis resulted in 37 resources scoring 100/100, 1 resource scoring 90/100, and 12 resources scoring 80/100. The remaining 9 resources did not receive a

QA score. By averaging the assigned scores, the rating of the body of evidence was found to be 95/100.

Findings Organized by Research Questions

According to the Available Literature, What Gender-responsive Modifications Have Been Made to Which Risk Assessment Tools?

This research question was inductively organized into both implemented and researched gender-responsive modifications for risk/risk-need assessment tools. In this study, researched modifications means that, while the modification itself has not been directly implemented, it has surpassed the stage of suggestion and the research either has, or is currently being, conducted in order for implementation to occur. While researched modifications may not yet have been applied, the novelty of this topic honors the progress made through peer-reviewed research. There was a total of 22 implemented modifications across 13 identified risk assessment tools and 13 researched gender-responsive modifications across 5 identified risk assessment tools identified in this dissertation. It is relevant to note that further tools may be within the research process for gender-responsive modifications, but many of these tools were left either unnamed or unidentified.

The present systematic review of 59 sources found the following gender-responsive modifications to have been directly implemented across a variety of risk/risk-need assessment tools, as organized by frequency of description across the sources: developed a gender-responsive tool from the ground up (26 sources), included gender-specific risk factors (7 sources), developed gender-specific risk assessment guidelines from the ground up (5 sources), differed cut-off scores by gender (4 sources), developed a gender-responsive trailer (3 sources), included gender-informed scales (2 sources), used gender-specific risk models (2 sources),

varied coding by gender (1 sources), and used gender-based norms (1 sources). Please refer to Appendix C for further definitions of these categories.

Likewise, this review found the following gender-responsive modifications to have been researched in application to risk assessment tools, as organized by frequency: develop a gender-responsive tool from the ground up (4 research projects), include gender-specific risk factors (4 research projects), develop a gender-responsive trailer (3 research projects), differ cut-off scores by gender (3 research projects), include gender-informed scales (3 research projects), develop gender-specific risk assessment guidelines from the ground up (1 research project).

Across these implemented and researched modifications, the assessment tools included: Female Additional Manual (FAM), the Women's Risk Needs Assessment (WRNA) Suite, the Security Reclassification Scale for Women (SRSW), the Service Planning Instrument (SPIn) Suite, the Level of Service (LS) Suite, The Historical Clinical Risk Management – 20, Version 3 (HCR-20 v3), the Missouri Women's Risk Assessment, the Inventory of Need Pretrial Screening Tool (ION), the COMPAS Risk & Need Assessment System, the Ohio Risk Assessment System (ORAS), the Minnesota Screening Tool Assessing Recidivism 2.0 (MnSTARR), the Static Risk and Offender Needs Guide (STRONG) Assessment System, and the Dynamic Risk Assessment for Offender Re-entry (DRAOR). A number of researched modifications were also identified, without formal affiliations with a specified risk/risk-need assessment instrument.

Below is a detailed table including each identified risk/risk-need assessment tool and the modification that was implemented or researched. The researched modifications are included at the end, organized either by the unnamed, "general" research, or by the temporary name that the researchers had provided the tool as it remained in creation.

Table 1

Results of Implemented Modifications, as Organized by Tool

Assessment Tool	Modifications
COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) Risk & Need Assessment System	Include gender-specific risk factors
	Include gender-informed scales
Dynamic Risk Assessment for Offender Re-entry (DRAOR)	Include gender-specific risk factors
Female Additional Manual (FAM)	Develop gender-specific risk assessment guidelines from ground up
Historical Clinical Risk Management - 20, Version 3 (HCR-20 v3)	Develop gender-specific risk assessment guidelines from ground up
Inventory of Need Pretrial Screening Tool (ION)	Develop gender-specific tool from ground up
	Include gender-informed scales
Level of Service (LS) Suite	Include gender-specific risk factors
	Different cut-off scores by gender
	Use of gender-based norms
Minnesota Screening Tool Assessing Recidivism 2.0 (MnSTARR)	Use of gender-specific risk model
Missouri Women's Risk Assessment	Develop gender-responsive tool from ground up
Ohio Risk Assessment System (ORAS)	Different cut-off scores by gender
Security Reclassification Scale for Women (SRSW)	Develop gender-responsive tool from ground up
	Include gender-informed scales
Service Planning Instrument (SPIn) Suite	Develop gender-responsive tool from ground up
	Include gender-specific risk factors
	Different cut-off scores by gender
Static Risk and Offender Needs Guide (STRONG) Assessment System	Use of gender-specific risk model
	Different cut-off scores by gender
	Include gender-specific risk factors
Women's Risk Needs Assessment (WRNA) Suite	Develop gender-responsive tool from ground up
	Include gender-specific risk factors
	Develop gender-responsive trailer

Three tools that were related to violence risk assessment, but not directly measuring these qualities themselves, remained included within these findings yet separately denoted as "adjacent tools." This inclusion and demarcation was due to the frequency of their use, as exemplified by the Static-99R, and the Hare Psychopathy Checklist - Revised (PCL-R), or to their representation of context-specific tools, as is the case with the Correctional Mental Health Screen for Females (CMHS-F). Accordingly, each of these tools are frequently utilized within comprehensive batteries of risk, even if they themselves and on their own are not direct measures of violence risk. The Static-99R specifically examines sexual recidivism, but in doing so, assesses for convictions of non-sexual violence (Hanson & Anderson, 2021). Likewise, the PCL-R assesses for psychopathy and is frequently included in both batteries as an independent tool and as a scoring factor embedded within other violence risk assessment tools (e.g., the Violence Risk Appraisal Guide-Revised; Hare, 2017). Alternatively, the CMHS-F was developed for use in correctional institutions specifically to assess for the mental health risks and needs of an individual and are often included in initial screenings of inmates as a determination of potential risk for disciplinary and adjustment problems while incarcerated (Ford et al., 2007). Table 2 lists the tools that had remained included throughout this screening process, and are often utilized in adjacent means to violence risk assessment:

Table 2

Results of Implemented Modifications of "Adjacent Tools"

Assessment Tool	**Modifications**
Correctional Mental Health Screen for Females (CMHS-F)	Develop gender-responsive tool from ground up
Hare Psychopathy Checklist - Revised (PCL-R)	Develop gender-specific risk assessment guidelines from ground up
	Different cut-off scores by gender
Static-99R	Different "coding" by gender

Which Gender-responsive Modifications Have Been Proposed for Future Changes?

Across the 59 sources, 13 different suggestions for gender-responsive modifications to risk/risk-need assessment tools were located. Four higher order categories developed through thematic analysis captured the nature of the suggestions: the development of a tool, the interpretation of a tool, the improvement of a tool that involves further research, and an extraneous category of modifications that could occur either during the development or the improvement of a tool. This latter category was created due to the lack of specificity by each resource's author as to specifically how or when this proposed modification would happen. Given the nature of the modification, it was determined that this could occur during either stage. Below, Table 3 details these four categorizations, 13 suggestions, and the frequency with which each of these were suggested within the literature.

Table 3

Results of Suggested Modifications, as Organized by Category

Categorization of Suggestion	**Actual Suggestion**	**Suggestions in Literature**
Development of a tool	Develop a gender-responsive tool from the ground up	13
	Develop a gender-responsive trailer	1
	Develop gender-specific risk assessment guidelines	1
	Include gender-specific risk and/or risk/needs factors	15
	Include gender-informed scales	1
Interpretation of a tool	Differ cut-off scores by gender	5
	Gender-specific modifications in interpretation	1
Improvement of a tool that involves further research	Improve statistical reliability and validity by gender	3
	Use gender-specific risk models	1
	Adjust risk categories to be more inclusive to gender	1
	Use gender as a predictor or risk assessment item	1
Modifications that can occur either in the development or improvement of a tool	Include larger female sample sizes	4
	Reweigh or recalibrate tool to be gender-responsive	3

How Does the Aforementioned Literature Define Gender in Discussing These Gender-responsive Modifications?

Of all included research questions, the least amount of data pertained to how gender was defined within the 59 sources that made it through the screening and quality appraisal process. No sources specifically defined gender beyond the binary male or female within their data set(s). Only two studies acknowledged the existence of non-binary aspects within gender, with this information remaining within the literature review or discussion section. First, de Vogel and Nicholls (2016) referred to the lack of research surrounding transgender individuals, particularly related to violence. Second, Hannah-Moffat (2009) explicated that studies "often categorically characterize gender as a 'sex' variable and disregard the nonaggregate differences among women and between men and women. Such a binary approach rests on specific assumptions about the nature of 'gender'" (p. 214).

Thus, these two studies demonstrated awareness about the social construct of gender and the range of characteristics and identities beyond the male and female binary. Despite this recognition, the targeted population in these two studies remained that of the binary male and females; the same was true for all of the other 57 sources. Also, neither of these sources defined gender in specific relation to their gender-responsive modifications.

Conclusion

The goal of this chapter was to present the results of the deductive and inductive analyses of the 59 studies on gender-responsive modifications to risk and need assessment tools. Careful and detailed categorizations were made of these findings to address the three research questions pertaining to modifications of risk and need assessment tools. By organizing the data in a meaningful and digestible manner, it is hoped that the reader will easily access information

related to gender-responsive modification types, tools that have incorporated these changes, suggestions for future changes, and their associated research articles. Of note, the quality or amount of actual use of these modifications, tools, and suggestions was not a part of the research questions or the collected data.

To make more meaning of this data, the results must be considered within the context of their practical use. The following chapter details the field's actual use of the various modifications or tools noted in the results and provides opinions regarding the quality and suggested applications of this information, to inform future research and practice.

Chapter 4: Discussion

This systematic review focused on identifying gender-responsive modifications and which risk assessment tools have implemented these modifications, what suggestions have been made for relevant future research, and how this literature defined gender. This literature review was the first study to identify and synthesize this information. Through the use of a comprehensive and detailed review process, nine databases were carefully searched for relevant content, ultimately resulting in 59 synthesized resources. In total, nine unique means by which violence risk assessment tools may become gender-responsive were identified, alongside 13 violence risk assessment tools that have implemented these modifications (see Table 1). In regards to suggested modifications, four independent categories were identified, consisting of 13 total suggestions (see Table 3). Lastly, none of the aforementioned resources addressed non-binary gender, with only two of these resources even mentioning gender outside of binary (i.e., male and female) terms; a significant finding given the APA's ethical guidelines that specify gender to be non-binary. Each of these findings has the potential to carry an important foundational step in furthering psychology's understanding of gender-specific violence risk assessment.

Discussion of Findings

This discussion section seeks to extend the impact of these results by exploring the information's practical use. Due to the foundational nature of this study, the quality or amount of actual use of these gender-responsive risk assessment tools remained outside the scope of the project. Therefore, a retrospective consideration of these tools' practical use will be presented to facilitate the use of this systematic review as a resource for clinicians, researchers, and

policymakers. Thus, this chapter is intended to discuss these findings in practical use and to provide opinions regarding implications for practice and policy, future research, and theory.

Implications for Practice and Policy

This systematic review was intended to provide a tool for clinicians to reference when working with a forensic population that identifies beyond the binary male in assessing for violence risk and/or criminogenic need. While the nuances of assessment tool selection include various types of risk (of note, suicide risk or tools solely addressing sexual violence were not included in the present study) and a number of identity variables, the aforementioned literature highlights the importance of representation, particularly in gender. Whereas gender remains the crux of this reviewed research, assessment tool selection must also take into account other static and dynamic factors. The following information is reflective of gender specifically.

When assessing for violence risk and criminogenic need, a clinician will often formulate a subjective battery that encapsulates a variety of different clinical factors in order to garner the most accurate picture of the individual's risk factors and future potential to recidivate. When seeking risk assessment tools that are responsive to the identity markers of a specific individual, many clinicians will initially seek a tool that fulfills the same functional need as an original, gender-neutral tool. As previously mentioned, the two primary types of tools utilized within violence risk assessment include actuarial tools, which estimate specific probability of an individual engaging in future violence through the use of fixed algorithms that are interpreted through inductive processes, or structured professional judgment (SPJ) tools, which incorporate these evidence-based guidelines with professional discretion (Hart et al., 2007; Melton et al., 2007).

During the assessment tool selection process, clinicians must also consider the quality and validity of the gender-responsive modification. Hierarchically, tools with normed populations that most closely resemble the individual identified for assessment are likely to provide the most accurate representation of the individual's violence risk and criminogenic need. For example, it is inherent that a gender-specific tool that is created for a binary female population will be the preferred tool in regard to validity for cisgender women (e.g., WRNA Suite). A gender-specific tool is created most often by incorporating gender-responsive risk factors, including gender-informed scales, and identifying methods of interpretation based on the identified gender (e.g., COMPAS, ION, and LS Suite, respectively).

However, the frequency with which a modification is created should not be confused with the ideal method of utilizing a normative population that is representative of the individual with whom it will be utilized. For example, some tools may only identify specific cut-off scores for alternate populations (e.g., ORAS). While alternate cut-off scores remain a valid modification, it must be supported by the appropriate research in order to be widely applied. Without properly enacting the validation research necessary to explore the predictive validity for a specific population (i.e., binary female offenders), researchers caution against the use of said modifications (Blanchette & Brown, 2006). When adjusting cut-off scores, the most common deleterious effects to the validity of the measure can include misclassification of the individual or an imbalance of the tool's specificity to sensitivity (Blanchette & Brown, 2006; Roehl et al., 2005).

We see an example of researchers warning clinicians that the gender-responsive modification of differing coding cut-offs by gender is insufficient within the Static-99R (a tool measuring sexual recidivism that is commonly utilized within risk assessment batteries). This

tool is not fully supported for use with binary females by the community, largely due to a published study examining the association between individual items on the Static-99R and the sexual recidivism of binary females (Marshall et al., 2021). Since Marshall et al. (2021)'s study displayed an overwhelming lack of association between these items and the women's recorded sexual recidivism, the author's noted that the tool was not suitable for use with this specific population.

While this systematic review was initially built upon the ideas of ensuring empirical data remains current with societal progress, there may be areas where this progress is inappropriately applied or utilized in practice prior to proper validation. Likewise, risk assessment tools are often based upon small sample sizes, particularly when accounting for gender. While violence risk assessment tools ultimately perform better than other risk assessments, they continue to have lower positive predictive values and have been found to function better when applied to individuals deemed to be of lower risk (Fazel et al., 2012). Therefore, a natural limitation within this type of tool is the accuracy of predicting specific instances to occur within a multitude of potentially impactful stimuli, inclusive to environments, events, and external individuals.

Thus, this author recommends the following gender-responsive risk assessment tools, which fall under the highest level of modification quality, namely a tool that has been developed specifically for a binary female population and has been thoroughly researched. Both stand-alone tools and supplements are included in the following list, categorized by their classification as an actuarial or structured professional judgment type of risk assessment:

Table 4

Author Recommendations of Gender-Responsive Tools for Use

Risk Assessment Type	**Gender-Responsive Risk Assessment Tool**
Actuarial	WRNA, WRNA-T (to supplement the LSI-R), SRSW, SPIn-W
Structured Professional Judgment	FAM (to supplement the HCR-20)

With the organization of the totality of information found within this systematic review, as seen in Appendix I, clinicians can search by gender-responsive modification or name of assessment tool. The tabular presentation allows for ease of access to the results of this systematic review in a variety of applications. Clinicians may then utilize the provided information to guide their choices by further researching the population upon which the tool was normed and the current research on the tool itself.

Although this systematic review examined the specific domain of violence risk assessment, gender-responsivity directly impacts case management and programming across institutions. By incorporating the relevant pathways and criminogenic needs that are likewise addressed in the creation of these assessment tools, society can further direct efforts to related avenues within our correctional and forensic facilities. Based on the resources created in the present study, potential changes in mental health policies could also be recommended. Such changes are necessary given the fields of clinical and forensic psychology's need to remain current with the individuals and societies they serve.

While ultimately a call for research can be concluded from the sparse information available (see next subsection), there may also be the potential for mental health policies to be affected on a variety of levels (i.e., institutional versus larger ordinances). For example, the

Council of State Justice Center's (2021) resource guide for adopting gender-responsive approaches for women in the justice system provide direct potential implications of these shifts:

> Gender-responsive and trauma-informed policies, practices, and programs recognize that women have distinct histories, pathways to offending, and experiences in the criminal justice system. These approaches address issues that may contribute to women's involvement in the justice system, such as domestic violence, abuse, and victimization; family and relationships; trauma; and poverty, mental illnesses, and substance use disorders. (p. 3)

Finally, a primary goal of this systematic review was to bring attention to the disparity between society's definition of gender and psychological assessment's incorporation of gender as a factor of consideration. While a wide variety of markers define an individual's identity or intersectional identities, the societal construct of gender, whether binary or non-binary, has a significant impact on an individual potential path to criminality and criminogenic needs. Again, it was significant that only two studies included within this systematic review even mentioned the concept of non-binary gender. Still, it is the hope that such considerations may be included both within forensic psychological practice and the field as a whole.

Implications for Research

Even though a strength of this systematic review is that it provides a referential source to which clinicians may refer when conducting risk assessments with non-binary-male populations, it also points to the need for more research regarding gender-responsivity in risk assessment tools. This limitation was highlighted through the aforementioned findings, with many articles largely noting the small sample sizes of non-binary-male offenders.

One way to remedy this problem could be to include larger samples of binary female populations within future research. This may be enacted either in accompaniment with binary male populations or through solely examining binary female populations. For example, a future research endeavor could focus upon a binary female population within a correctional or forensic

setting. Through a hypothetical longitudinal study, a specified risk assessment tool could be administered to a group of inmates or patients prior to their release dates. These individuals would then be followed for a set amount of time to record if any violent actions, violations of supervision, or general recidivism were to occur. The correlation between individual items or collective scores on the assessment would then be directly compared to the longitudinal violence and recidivism data. However, continuing to solely study binary genders remains a flawed approach.

Instead, addressing non-binary genders in risk assessment would bridge a brazen gap in the literature and would comport with ethical practice standards. As previously mentioned, the APA's (2015) ethical guidelines "stipulate that psychologists should (a) recognize gender as nonbinary and (b) abandon binary gender measurement in favor of accurate and inclusionary measures" (p. 4). Given that none of the included tools identified within this study define gender beyond "male" and "female," it is pertinent that non-binary populations begin to be examined with efforts to create normative populations through which existing risk assessment tools and novel tools may be validated.

Cameron and Stinson (2019) detailed the consequences and ethical violations, as defined by the American Psychological Association (APA), in utilizing binary gender in demographic measures, and provided suggestions for future research to evolve the field of psychological research. Cameron and Stinson (2019) cautioned that when a non-binary individual participating in psychological research is forced to choose between binary gender identifiers, this experience fuels prejudice and discrimination. Further, this forced choice leads to false data collection, as non-binary individuals are not able to accurately account for their gender within these limited factors, which can threaten the validity of the psychological study.

One way to be more gender-inclusive is through the use of open-ended questions of gender identification. Another way is to use a gender-inclusive measure. If one is not available, the researcher may maintain the reliability and validity of the research by including a thorough description of participants' gender identities as well as a thorough methodological description of how data from nonbinary participants were treated (Cameron & Stinson, 2019). While the research and tools may not currently be available to adequately represent individuals that do not identify as one of the binary genders, clinicians must account for such factors within their evaluations. It is a rarity for the normed population of a tool to represent an individual's identity in full in direct practice. However, including these limitations within one's interpretation of violence risk and criminogenic need is exceedingly important when accounting for how an individual may reintegrate into society.

Due to the limited nature of the data examining gender responsivity in risk assessment tools, extant studies may often lack critical methodologies of statistical precision to uphold empirical validity (Saxena et al., 2014). Given the limitations identified within the variety of included studies, there appear to be challenges in assessing the clinical significance between gender-neutral and gender-responsive interventions, such as unreliable validity or misinformed conclusions being drawn (Gobeil et al., 2016). Therefore, this appears to be an area in desperate need of further research, allowing grace while the gap between social science and society narrows.

Limitations

The limitations of this study occurred in both the methodology and through the nature of the research questions themselves. As identified by Whittemore and Knafl (2005), an integrative systematic review can be easily biased, as methods of coding and synthesizing information can

be poorly defined. Therefore, the detail within the methodology section was intended to be as well-defined as possible, particularly since work in this project involved multiple researchers, and so that the replicability of the subjective structure would be elucidated. In order to control for systematic biases or errors, the use of an intermittent peer reviewer and an audit trail aided in balancing the subjectivity of the quality appraisal and data extraction forms.

In addressing limitations within the methods section, it is important to note that these researchers utilized quality appraisal tools for both mixed method resources and systematic reviews. However, nine of the included resources were categorized as literature reviews, inclusive to government publications. These resources did not undergo a formal quality appraisal process aside from the established inclusion and exclusion criteria for this systematic review, including the resource having undergone a peer review process. A majority of researched quality appraisal tools appeared to best address systematic reviews and required a defined rationale for the included articles, which none of the included literature reviews appeared to have. While varied quality of included resources is expected in an integrative systematic review examining a budding topic in the field of psychology, all possible standards should be upheld. Therefore, for future researchers, curating a standardized quality appraisal process for the inclusion of literature reviews may be an option, given this writer's hope that further research will continue to surface, broadening the scope and availability of relevant information.

Also, the chosen search criteria specifically targeted either violence risk or risk *and* need assessment tools. Some gender-responsive tools solely measuring criminogenic need (such as the Gender Informed Needs Assessment, or GINA) were not included within these findings (The Council of State Governments Justice Center, 2021). Likewise, only sources that had been published in the English language were included within this study. This inclusion criteria may

have affected both the breadth of the information included as well as the bias towards Eurocentric ideas and demographics reflected within these results.

While the primary researcher utilized the aid of two research assistants during this process, formal interrater reliability checks were not conducted. If disagreements regarding the criteria of a source were encountered, they were met with open discussion and were all resolved without needing to involve a further party or mediator. The primary researcher acknowledges that this may be the result of a power imbalance between herself and the research assistants. However, all three involved parties maintained respectful terms, audit trails, and detailed notes regarding decisions made to avoid potential biases having detrimental effects upon the research.

Likewise, this systematic review chose to solely focus upon gender and not sex or sexuality, as these terms are frequently conflated within the community. While there are many other variables that hold extreme importance to the identity of an individual, gender is the aspect chosen for this specific dissertation. Future research would best include a variety of focal points, not limited to sexuality, ethnicity, cultural and religious values, and many other pertinent points of identity as well as their intersectionality.

As this systematic review examined a topic that is a recently developed field of research, the search process yielded a relatively low amount of data. As previously mentioned in the literature review, the research that appears to exist outside of the binary male gender addresses the binary female gender. The extracted data solely applied to binary females, with only two included articles mentioning non-binary genders but not including such individuals within the study itself. The lack of non-binary data was a pre-research assumption; however, this lack of evidence does not conclude that the difference does not hold significance but may instead be reflective of the field of psychological assessment's early stages of change towards gender

existing on a full spectrum. While future researchers are encouraged to expand their limited definition of gender, the present study is still believed to be a progressive step towards modifying normative populations from the current majority of binary male.

Summary: Contributions and Concluding Remarks

This systematic review contributes to the fields of forensic and clinical psychology by providing clinicians and researchers with a novel comprehensive resource that identifies which contemporary forensic risk and risk/need assessment tools have been modified to accommodate binary female or gender-expansive normative populations. After conducting the literature review necessary to complete the background and rationale section of this systematic review, as well as the full systematic review, this author did not find any extant resource that accomplished this goal. To meet this need, this systematic review created a detailed resource identifying the nature of any implemented or proposed gender-responsive modifications. It is hoped that this novel and needed synthesis of accurate risk and risk/need assessment tools will, therefore, make them more easily accessible for use and study with members of the forensic population outside of the binary male gender.

While not all tools are able to be widely applied amongst different populations, this resource may also serve as a launching point for further gender-responsive tools to be developed. Future researchers can identify relevant tools in existence and model steps of modification from these examples. By also examining how these modifications have defined "gender," bases for further research can be provided.

In conclusion, it is hoped that the results of this dissertation have the potential to aid future research and practitioners in their work with forensic individuals identifying outside of the male binary gender. With the impact to social liberties that each risk assessment may address, it

is ever important to ensure that the clinician's role remains adherent to ethical models of best practices. By further investigating and expanding the available resources by which nonbinary or binary female individuals may identify, the psychological field will continue to maintain adherence to societal expectations and due justices, particularly within forensic systems.

References

Ackerman, M. J. (2010). *Essentials of Forensic Psychological Assessment,* (2nd ed.). John Wiley & Sons.

American Psychiatric Association (2022). DSM-5-TR Gender Dysphoria Fact Sheet. https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/DSM-5-TR/APA-DSM5TR-GenderDysphoria.pdf

American Psychological Association (2015). Guidelines for psychological practice with transgender and gender non-conforming people. *American Psychologist*, *70*, 832–864.

Andrews, D. A., Guzzo, L., Raynor, P., Rowe, R. C., Rettinger, L. J., Brews, A., & Wormith, J. S. (2012). Are the major risk/need factors predictive of both female and male reoffending? *International Journal of Offender Therapy and Comparative Criminology*, *56*(1), 113–133. https://doi.org/10.1177/0306624x10395716

Associated Press (2017, January 6). California murder convict becomes first U.S. inmate to have state-funded sex reassignment surgery. *Los Angeles Times*. https://www.latimes.com/local/lanow/la-me-ln-inmate-sex-reassignment-20170106- story.html.

Barbaro, A. (2019). Women in forensics: An international overview. *Forensic Science International: Synergy*, *1*, 137–139. https://doi.org/10.1016/j.fsisyn.2019.06.047

Belingheri, P., Chiarello, F., Fronzetti Colladon, A., & Rovelli, P. (2021). Twenty years of gender equality research: A scoping review based on a new semantic indicator. *PLOS ONE*, *16*(9). https://doi.org/10.1371/journal.pone.0256474

Blanchette, K., & Brown, S. L. (2006). *The assessment and treatment of women offenders: An integrative perspective*. John Wiley.

Bloom, B. E., Owen, B. A., & Covington, S. (2003). *Gender-responsive strategies: Research,*

practice, and guiding principles for women offenders. U.S. Department of Justice, National Institute of Corrections.

Bonta, J., & Wormith, S. J. (2008). Risk and need assessment. In G. McIvor & P. Raynor (Eds.), *Developments in social work with offenders*. (pp. 131-152). Jessica Kingsley Publishers.

Brabender, V., & Mihura, J. L. (2016). *Handbook of gender and sexuality in psychological assessment*. Routledge, Taylor & Francis Group.

Brill, S. A., & Kenney, L. (2016). *The transgender teen: A handbook for parents and professionals supporting transgender and non-binary teens*. Cleis Press.

Brown, S. (2017). *A Review of the women offender risk/need research: In search of gender-neutral, women-salient and women-specific risk factors* (Research Report R-386). Correctional Service of Canada.

Brown, S. & Motiuk, L. (2008). *Using dynamic risk factors to predict criminal recidivism in a sample of male and female offenders*. [Powerpoint slides]. https://carleton.ca/gendercrimelab/wp-content/uploads/CPA_2008-DFIA.pdf

Cameron, J. J., & Stinson, D. A. (2019). Gender (mis)measurement: Guidelines for respecting gender diversity in psychological research. *Social and Personality Psychology Compass*, *13*(11). https://doi.org/10.1111/spc3.12506

Caulfield, L. (2010). Rethinking the assessment of female offenders. *Howard Journal of Criminal Justice*, *49*(4), 315-327. https://doi.org/10.1111j.1468-2311.2010.00625.x

Chrisler, J. C., & McCreary, D. R. (2010). *Handbook of gender research in psychology, volume 2: Gender research in social and applied psychology*. Springer New York.

Cobbina, J. E. (2010). Reintegration success and failure: Factors impacting reintegration among

incarcerated and formerly incarcerated women. *Journal of Offender Rehabilitation*, *49*(3), 210–232. https://doi.org/10.1080/10509671003666602

Conroy, M. A., & Murrie, D. C. (2007). *Forensic assessment of violence risk: A guide for risk assessment and risk management*. John Wiley & Sons.

Davis, R. M., Batastini, A. B., Sacco, D., Dahlen, E. R., & Jones, A. C. (2021). Does race matter? An examination of defendant race on legal decision making in the context of actuarial violence risk assessments. *Legal and Criminological Psychology*. https://doi.org/10.1111/lcrp.12204

de Ruiter, C. D., & Kaser-Boyd, N. (2015). *Forensic psychological assessment in practice: Case studies*. Routledge.

de Vogel, V., & Nicholls, T. L. (2016). Gender matters: An introduction to the special issues on women and girls. *International Journal of Forensic Mental Health*, *15*(1), 1–25. https://doi.org/10.1080/14999013.2016.1141439

Delve Tool. *Essential guide to coding qualitative data.* Delve. https://delvetool.com/guide

Eckhouse, L., Lum, K., Conti-Cook, C., & Ciccolini, J. (2018). Layers of bias: A unified approach for understanding problems with risk assessment. *Criminal Justice and Behavior*, *46*(2), 185–209. https://doi.org/10.1177/0093854818811379

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*(1), 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Erwin, E. J., Brotherson, M. J., & Summers, J. A. (2011). Understanding qualitative metasynthesis. *Journal of Early Intervention*, *33*(3), 186–200. https://doi.org/10.1177/1053815111425493

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to

predict violence and antisocial behaviour in 73 samples involving 24,827 people: Systematic Review and meta-analysis. *The BMJ*, *345*(jul24 2). https://doi.org/10.1136/bmj.e4692

Fontanari, A. M., Vilanova, F., Schneider, M. A., Chinazzo, I., Soll, B. M., Schwarz, K., Lobato, M. I., & Brandelli Costa, A. (2020). Gender affirmation is associated with transgender and gender nonbinary youth mental health improvement. *LGBT Health*, *7*(5), 237–247. https://doi.org/10.1089/lgbt.2019.0046

Ford, J., Trestman, R. L., Osher, F., Scott, J. E., Steadman, H. J., & Robbins, P. C. (2007, May). *Mental health screens for corrections*. National Institute of Justice. https://nij.ojp.gov/library/publications/mental-health-screens-corrections

Forensic Psychology. (2019). https://www.apa.org/ed/graduate/specialize/forensic.

Given, L. M. (2008). *The Sage encyclopedia of qualitative research methods*. Sage.

Gobeil, R., & Blanchette, K. (2007). Revalidation of gender-informed security reclassification scale for women inmates. *Journal of Contemporary Criminal Justice*, *23*(4), 296-309. https://doi.org/10.1177/1043986207309411

Gobeil, R., Blanchette, K., & Stewart, L. (2016). A meta-analytic review of correctional interventions for women offenders. *Criminal Justice and Behavior*, *43*(3), 301–322. https://doi.org/10.1177/0093854815621100.

Government of Canada, P. S. and P. C. (2019). *Women offender assessment: Can gender-informed variables improve risk prediction?* https://publications.gc.ca/collections/collection_2019/scc-csc/PS83-3-413-eng.pdf

Gray, N. S., Noaks, L., & Laing, J. M. (2013). *Criminal justice, mental health and the politics of risk*. Routledge-Cavendish.

Hannah-Moffat, K. (2009). Gridlock or mutability: Reconsidering "gender" and risk assessment. *Criminology & Public Policy*, *8*(1), 209–219. https://doi.org/10.1111/j.1745-9133.2009.00549.x

Hanson, R. K., & Anderson, D. (2021). Static-99R: An empirical-actuarial risk tool for adult males with a history of sexual offending. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (pp. 106–130). Routledge/Taylor & Francis Group.

Hare, R. D. (2017). *Hare psychopathy checklist-revised (Pcl-R), 2nd Edition*. MHS, Multi-Health Systems.

Hart, S., Michie, C., & Cooke, D. (2007). Precision of actuarial risk assessment instruments: Evaluating the 'margins of error' of group v. individual predictions of violence. *British Journal of Psychiatry,190*(S49), S60-S65. https://doi.org/10.1192/bjp.190.5.s60

Hong, Q. N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice*, *24*(3), 459–467. https://doi.org/10.1111/jep.12884

Hong, Q.N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I. (2018). Mixed Methods Appraisal Tool (MMAT). Registration of Copyright (#1148552), Canadian Intellectual Property Office, Industry Canada.

Incarcerated Women and Girls. The Sentencing Project. (2020, November 24). https://www.sentencingproject.org/publications/incarcerated-women- and-girls/

King, E. & Foley, J. (2014) *Gender-Responsive policy development in corrections: What we know and roadmaps for change*. National Institute of Corrections. (2014)

Kruttschnitt, C., & Gartner, R. (2003). Women's imprisonment. *Crime and Justice*, *30*, 1–81. https://doi.org/10.1086/652228

Lanier, C.L. & DeVall, K. (2018, April). *Understanding areas of risk and need among female offenders: Translating research to practice.* Presented at the National TASC Annual Conference, St. Petersburg, FL.

Marshall, E. A., & Miller, H. A. (2018). Examining gender-specific and gender-neutral risk factors in women who sexually offend. *Criminal Justice and Behavior*, *46*(4), 511–527. https://doi.org/10.1177/0093854818796872

Marshall, E., Miller, H. A., Cortoni, F., & Helmus, L. M. (2021). The Static-99r is not valid for women: Predictive validity in 739 females who have sexually offended. *Sexual Abuse*, *33*(6), 631–653. https://doi.org/10.1177/1079063220940303

Matsuno, E. (2019). Nonbinary-affirming psychological interventions. *Cognitive and Behavioral Practice*, *26*(4), 617–628. https://doi.org/10.1016/j.cbpra.2018.09.003

McSherry, B. (2004). Risk assessment by mental health professionals and the prevention of future violent behaviour. *Trends & Issues in Crime and Criminal Justice*. https://doi.org/10.1037/e620752012-001

Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers*. Guilford.

Motmans, J., Nieder, T. O., & Bouman, W. P. (2019). Transforming the paradigm of nonbinary transgender health: A field in transition. *International Journal of Transgenderism*, *20*(2-3), 119–125. https://doi.org/10.1080/15532739.2019.1640514

National Center for State Courts. (2014). *Offender risk & needs assessment instruments: A*

primer for courts. https://www.ncsc.org/__data/assets/pdf_file/0018/26226/bja-rna-final-report_combined-files-8-22-14.pdf.

Otto, R. K. (2000). Assessing and managing violence risk in outpatient settings. *Journal of Clinical Psychology*, *56*(10), 1239–1262. https://doi.org/10.1002/1097-4679(200010)56:10<1239::aid-jclp2>3.0.co;2-j

Otto, R. K. & Douglas, K. S. (2011). *Handbook of violence risk assessment (international perspectives on forensic mental health)*. Routledge.

PFLAG national glossary of terms. PFLAG. (2021, October 29). https://pflag.org/glossary

Price-Feeney, M., Green, A. E., & Dorison, S. (2020). Understanding the mental health of transgender and nonbinary youth. *Journal of Adolescent Health*, *66*(6), 684–690. https://doi.org/10.1016/j.jadohealth.2019.11.314

Ramesh, T., Igoumenou, A., Vazquez Montes, M., & Fazel, S. (2018). Use of risk assessment instruments to predict violence in forensic psychiatric hospitals: A systematic review and meta-analysis. *European Psychiatry*, *52*, 47–53. https://doi.org/10.1016/j.eurpsy.2018.02.007

Rider, G. N., Vencill, J. A., Berg, D. R., Becker-Warner, R., Candelario-Pérez, L., & Spencer, K. G. (2020). The Gender Affirmative Lifespan Approach (GALA): A framework for competent clinical care with nonbinary clients. *Non-Binary and Genderqueer Genders*, 157–170. https://doi.org/10.4324/9781003015888-15

Roehl, J., O'Sullivan, C., Webster, D., & Campbell, J. (2005). Intimate partner violence risk assessment validation study: The RAVE study practitioner summary and recommendations: Validation of tools for assessing risk from violent intimate partners. *PsycEXTRA Dataset*. https://doi.org/10.1037/e515672006-001

Roychowdhury, A., & Adshead, G. (2014). Violence risk assessment as a medical intervention: Ethical tensions. *The Psychiatric Bulletin*, *38*(2), 75–82. https://doi.org/10.1192/pb.bp.113.043315

Russell, C. L. (2005). An overview of the integrative research review. *Progress in transplantation*, *15*(1), 8–13. https://doi.org/10.1177/152692480501500102

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Sage.

Saxena, P., Messina, N. J., & Grella, C. E. (2014). Who benefits from gender-responsive treatment? Accounting for abuse history on longitudinal outcomes for women in prison. *Criminal Justice and Behavior*, *41*, 417-432.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*(1), 747–770. https://doi.org/ 10.1146/annurev-psych-010418-102803

Systematic Reviews Critical Appraisal Sheet. *Centre for Evidence Based Medicine, University of Oxford.* https://www.cebm.ox.ac.uk/resources/ebm-tools/ critical-appraisal-tools

Task Force on Federally Sentenced Women (1990). Creating choices: Report of the Task Force on Federally Sentenced Women. Correctional Service Canada.

Taxman, F. S. (2016). *Handbook on risk and need assessment: Theory and practice (the ASC division on corrections & sentencing handbook series 1).* Routledge, Taylor & Francis Group.

The Council of State Governments Justice Center, Fleming, E., Upton, A., Lopez Wright, F.,

Wurzburg, S., & Ney, B., *Adopting a Gender-Responsive Approach for Women in the Justice System: A Resource Guide* (2021). CSG Justice Center. https://csgjusticecenter.org/wp-content/uploads/2021/08/Adopting-Gender-Responsive-Approach-for-Women-in-Justice-System.pdf.

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 gender-responsive supplement. *Criminal Justice and Behavior*, *37*(3), 261-288. https://doi.org/10.1177/0093854809357442

Vigod, S. N., & Rochon, P. A. (2020). The impact of gender discrimination on a woman's mental health. *EClinicalMedicine*, *20*, 100311. https://doi.org/10.1016/j.eclinm.2020.100311

Wasserman, D. (2014). Ethical and policy issues in genetic prediction of violence: Implications for clinicians. *Current Genetic Medicine Reports*, *2*(4), 216–222. https://doi.org/10.1007/s40142-014-0054-6

Women's Risk Needs Assessment Research. (2019) Women's Risk Needs Assessment Research (WRNA) - College of Social Work at the University of Utah. https://socialwork.utah.edu/research/ucjc/wrna/index.php

Warren, E. & Supreme Court of The United States. (1965) *U.S. Reports: Baxstrom v. Herold, 383 U.S. 107.* [Periodical] https://www.loc.gov/item/usrep383107/.

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.

Wright, E., Van Voorhis, P., Bauman, A., & Salisbury, E. (2008). Gender-Responsive Risk/Needs Assessment: Final Report. *Minnesota Department of Corrections*.

Young, E. (2020). Mental Health. In *They/Them/Their: A guide to nonbinary and genderqueer identities* (p. 142). essay, Jessica Kingsley Publishers.

APPENDIX A

Search Records

FORMING the SEARCH STRINGS

Search Term ID#	Primary Term	Synonyms/Alternate Forms	Notes
1			
2			
3			
4			

SEARCH PLAN

Full Search ID#	Search Type	Databases or Sources	Search Term ID#s	Boolean Search String	Fields to Search	Specifiers	Notes

SEARCH DOCUMENTATION RECORD

Search Date	Full Search ID#	Type of Search	Database /Resource	Search Term ID#s	Boolean Search String	Fields Searched	Search Specifier: YEARS	Search Specifier: PUBLICATION	Search Specifier: PEER REVIEWED	Search Specifier: OTHER	# of Records Found	Notes

APPENDIX B

Dissertation Database

SCREENING and SELECTION

Source Number	Type of Resource	Saved to Google Drive?	Authors	Year	Title	Database/Source	Screener 1: Include? (Y/N/V) (Date) (Initials)	Screener 2: Include? (Y/N) (Date) (Initials)	Final Decision – Include? (Y/N & Date)	Submit for QA?	Notes

QUALITY APPRAISAL

Source Number	Author	Year	Title	Yes(A) or Yes(B)?	Date of Quality Appraisal	Type of Study	Screening Tool Used	QA Percentage	Notes

	Origin Information					General Information			
Doc ID#	Database	Year Published	Type of Publication	Journal	Country/ies of Research	Full Document Title	Authors	In-text Citation *(Authors, Year)*	Full Citation

Inclusion/Exclusion Process			Design Characteristics & Methodological Features					
Date IECC Completed	YES(A) or YES(B)?	Appendix B Notes	Date Completed	Identified Design	Sample Size	Sample Characteristics	QA Score	QA Notes

ASSESSMENT OF RESEARCH VARIABLES							Modifications to Risk Assessments (i)				
Date of Data Collection & Extraction	How was gender defined?	G-R	G-S	G-I	G-N	Type of verbiage used in discussing modifications? *(G-R, G-I, G-N, G-SI)*	How are the assessments identified? *(VRA/RNA)*	Risk Assessment Tool Identified	Modification suggested, implemented, or both?	What was the modification?	What pg. was this found on?	Notes for DC&E

Modifications to Risk		Modifications to Risk			
Modification suggested or enacted?	What pg. was this found on?	Modification suggested or enacted?	What pg. was this found on?	Quote #1	Quote #2

EXCLUDED SOURCES

Source Number	Database/Source	Authors	Year	Reason for Exclusion (Screening 1)	Type of Resource	Duplicate Removal

APPENDIX C

Code Manual

"Developed a gender-responsive tool from the ground up"

Created an original risk assessment tool, including conducting supporting research, utilizing a specific gender for the sample population, and conducting validation studies.

E.g.: A research team secures funding to create an entirely new risk assessment tool solely based on a binary female population.

"Included gender-specific risk factors"

Individual items included in the risk assessment tool are related to risk factors that have been identified to be gender-specific in nature.

E.g.: The risk assessment tool includes items that have been specifically found to correlate with a binary female's pathway to crime.

"Developed gender-specific risk assessment guidelines from the ground up"

An addition created for a specific, existing risk assessment tool that, when utilized, increases the validity of the tool with the specified population. These guidelines often include both novel risk factors that reflect gender-specific issues and expansion of existing risk factors to be gender-responsive. These additions may also expand the areas of risk that are examined.

E.g.: When gender-responsive guidelines are utilized, the evaluator receives information on gender-responsive items (i.e., childcare, prostitution) that may expand the resulting assessment of risk (i.e., the inclusion of victimization).

"Differed cut-off scores by gender"

Based on validation research, the participant's gender identification may result in a higher or lower cut-off score than the original sample population. This may apply to individual items or to the overall outcome measure of the tool.

E.g.: On a specific assessment, a binary female might have a higher cut-off score than a binary male (i.e., they must surpass the threshold of "7" on an item, where a male would only have to surpass the threshold of "5").

"Developed a gender-responsive trailer"

An addition created to supplement existing risk assessment tools with items related to gender-responsive issues. These items, in turn, form additional scales to be included within the final opinion of risk. Trailers can be a combination of both interview questions and a self-report survey.

E.g.: By using a trailer, items may be added to the assessment that examine family conflict and relationship issues.

"Included gender-informed scales"

A grouping of gender-informed items create a gender-informed scale, as determined through literature reviews, clinical experience, and validation testing.

E.g.: Individual items that may be related to specified areas of gender-specific pathways may gather to form an overarching scale that addresses a specific area of concern or need.

"Used gender-specific risk models"

A method of incorporating realistic views of gender-specific risk and need items to capture the constellation of factors that may be represented within an individual. This allows for a holistic view of the individual's circumstances and how they may have arrived at their current position in society.

E.g.: One of the many pathways to crime could be victimization as a child. Research shows that female juvenile delinquents have a higher victimization rate of neglect and abuse as compared to their male counterparts[1]*. This, in turn, becomes a risk model in which somebody may explain a gender-specific aspect of the individual's pathway to crime.*

"Different coding by gender"

Risk assessment tools may account for gender by assigning various numbers or weights to a specified gender.

E.g.: When scoring a measure, an individual who has the gender identity of "binary female" may receive one extra point than an individual who has the gender identity of "binary male."

"Used gender-based norms"

When normative values are specifically established for that gender.

E.g.: When validation studies determine that the recidivism rate for a binary female differs from that of the population rate, a gender-based norm is established for the tool's future use with binary females.

[1] Asscher, J. J., Van der Put, C. E., & Stams, G. J. (2015). Gender differences in the impact of abuse and neglect victimization on adolescent offending behavior. *Journal of Family Violence*, *30*(2), 215–225. https://doi.org/10.1007/s10896-014-9668-4

APPENDIX D

Inclusion/Exclusion Criteria Checklist

Inclusion/Exclusion Screening Criteria Checklist

Article #:

Article Title:

Primary Author:

Criteria:

- ❑ Article published in English
- ❑ Article published in 1998 or after: ____
- ❑ Adult population (18 y.o. +)
- ❑ Article peer-reviewed
  - *Was the article pre-screened as peer-reviewed in the database?* *Y* *N*
  - *If "No," how did you determine?* _____
- ❑ Violence risk/Risk-need assessment(s) listed:

  _____
- ❑ Gender inclusive/responsive/informed modification(s) to VRA/RNA(s) discussed in article
  - *Was this:* ***(A)*** *actually implemented* *or* ***(B)*** *proposed?*

Notes: _____

_____

Do we include this article? YES (a) YES (b) NO VERIFY

Date Completed: _____ **Who Completed**: _____

Date Verified: _____ ***Who Verified***: _____

APPENDIX E

MMAT Tool

Part I: Mixed Methods Appraisal Tool (MMAT), version 2018

Category of study designs	Methodological quality criteria	Responses			
		Yes	No	Can't tell	Comments
Screening questions (for all types)	S1. Are there clear research questions?				
	S2. Do the collected data allow to address the research questions?				
	Further appraisal may not be feasible or appropriate when the answer is 'No' or 'Can't tell' to one or both screening questions.				
1. Qualitative	1.1. Is the qualitative approach appropriate to answer the research question?				
	1.2. Are the qualitative data collection methods adequate to address the research question?				
	1.3. Are the findings adequately derived from the data?				
	1.4. Is the interpretation of results sufficiently substantiated by data?				
	1.5. Is there coherence between qualitative data sources, collection, analysis and interpretation?				
2. Quantitative randomized controlled trials	2.1. Is randomization appropriately performed?				
	2.2. Are the groups comparable at baseline?				
	2.3. Are there complete outcome data?				
	2.4. Are outcome assessors blinded to the intervention provided?				
	2.5 Did the participants adhere to the assigned intervention?				
3. Quantitative non-randomized	3.1. Are the participants representative of the target population?				
	3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)?				
	3.3. Are there complete outcome data?				
	3.4. Are the confounders accounted for in the design and analysis?				
	3.5. During the study period, is the intervention administered (or exposure occurred) as intended?				
4. Quantitative descriptive	4.1. Is the sampling strategy relevant to address the research question?				
	4.2. Is the sample representative of the target population?				
	4.3. Are the measurements appropriate?				
	4.4. Is the risk of nonresponse bias low?				
	4.5. Is the statistical analysis appropriate to answer the research question?				
5. Mixed methods	5.1. Is there an adequate rationale for using a mixed methods design to address the research question?				
	5.2. Are the different components of the study effectively integrated to answer the research question?				
	5.3. Are the outputs of the integration of qualitative and quantitative components adequately interpreted?				
	5.4. Are divergences and inconsistencies between quantitative and qualitative results adequately addressed?				
	5.5. Do the different components of the study adhere to the quality criteria of each tradition of the methods involved?				

APPENDIX F

CEBM Tool

SYSTEMATIC REVIEW

Are the results of the review valid?

What question (PICO) did the systematic review address?

What is best?

The main question being addressed should be clearly stated. The exposure, such as a therapy or diagnostic test, and the outcome(s) of interest will often be expressed in terms of a simple relationship.

Where do I find the information?

The Title, Abstract or final paragraph of the Introduction should clearly state the question. If you still cannot ascertain what the focused question is after reading these sections, search for another paper!

In this paper

Yes ☐ No ☐ Unclear ☐

Comment:

F - Is it unlikely that important, relevant studies were missed?

What is best?

The starting point for a comprehensive search for all relevant studies is the major bibliographic databases (eg Medline, Cochrane, EMBASE, etc) but should also include a search of reference lists from relevant studies and contact with experts, particularly to inquire about unpublished studies. The search should not be limited to English language only. The search strategy should include both MESH terms and text words.

Where do I find the information?

The Methods section should describe the search strategy, including the terms used, in some detail. The Results section will outline the number of titles and abstracts reviewed, the number of full-text studies retrieved, and the number of studies excluded together with the reasons for exclusion. This information may be presented in a figure or flow chart.

In this paper

Yes ☐ No ☐ Unclear ☐

Comment:

A - Were the criteria used to select articles for inclusion appropriate?

What is best?

The inclusion or exclusion of studies in a systematic review should be clearly defined *a priori*. The eligibility criteria used should specify the patients, interventions or exposures and outcomes of interest. In many cases the type of study design will also be a key component of the eligibility criteria.

Where do I find the information?

The Methods section should describe in detail the inclusion and exclusion criteria. Normally, this will include the study design.

In this paper

Yes ☐ No ☐ Unclear ☐

Comment:

A - Were the included studies sufficiently valid for the type of question asked?

What is best?

The article should describe how the quality of each study was assessed using predetermined quality criteria appropriate to the type of clinical question (e.g., randomization, blinding and completeness of follow-up)

Where do I find the information?

The Methods section should describe the assessment of quality and the criteria used. The Results section should provide information on the quality of the individual studies.

In this paper

Yes ☐ No ☐ Unclear ☐

Comment:

APPENDIX G

Data Collection and Extraction Form

Data Collection & Extraction Form

Document ID#	Database	Year Published
Type of Publication	**Institution conducting research/review**	**Country of Publishing**

Full Document Title

Authors *(Last, First)*

In-text Citation *(Authors, Year)*

Full Citation

IECC

Date IECC Completed	
Verification Required?	
YES(A) or YES(B)?	
Appendix B Notes	

Quality Appraisal

Date QA Completed	
Identified Method	
QA Score	
QA Notes	

Assessment of Research Variables

Date of Data Extraction		
How was gender defined? *Binary or non-binary?*		p. ooo
Type of verbiage used in discussing modifications? *G-R, G-I, G-N, G-S?*		p. ooo
How are the assessments identified? *VRA/RNA?*		p. ooo

Modified Risk-Assessment	Modifications suggested or enacted?	What was the modification?	What page is this found on?
			p. ooo
			p. ooo
			p. ooo
			p. ooo

Relevant/Impactful Quotes:

-

APPENDIX H

Table of Included Studies

	Year	Design	Sample Size	Sample Characteristics					Reference
				Sample Gender (binary)(%age rounded to nearest whole)	**Ethnicities** (%ages averaged if multiple trials, nearest whole)	**Average Age**	**Forensic Status** (type of institution, security level, legal status, etc.)	**Location**	
1	2011	Non-Randomized	2,423	Female (15%), Male (85%)	Unspecified	Adults and Minors	Department of Corrections	Canada	Andrews, D. A., Guzzo, L., Raynor, P., Rowe, R. C., Rettinger, L. J., Brews, A., & Wormith, J. S. (2011). Are the major risk/need factors predictive of both female and male reoffending? *International Journal of Offender Therapy and Comparative Criminology*, *56*(1), 113–133. https://doi.org/10.1177/0306624x10395716
2	2013	Non-Randomized	5104	Female (23%), Male (77%)	60% White, 40% Black	Adult	Incarceration (Prison)	United States (Wisconsin)	Baskin-Sommers, A. R., Baskin, D. R., Sommers, I. B., & Newman, J. P. (2013). The intersectionality of sex, race, and psychopathology in predicting violent crimes. *Criminal Justice and Behavior*, *40*(10), 1068–1091. https://doi.org/10.1177/0093854813485412
3	2016	Non-Randomized	506	Female (7%), Male (93%)	Unspecified	32	Probation	Pakistan (Lahore, Kasur, Sheikhupura, & Nankana Sahib of Lahore division)	Bhutta, M. H., & Wormith, J. S. (2016). An examination of a risk/needs assessment instrument and its relation to religiosity and recidivism among probationers in a Muslim culture. *Criminal Justice and Behavior*, *43*(2), 204–229. https://doi.org/10.1177/0093854815604011
4	2007	Non-Randomized	172	Female (100%)	60% Caucasian, 26% Aboriginal, 14% Other	31	Incarceration (Prison; 2-20 year sentences)	Canada	Blanchette, K., & Taylor, K. N. (2007). Development and field test of a gender-informed security reclassification scale for female offenders. *Criminal Justice and Behavior*, *34*(3), 362–379. https://doi.org/10.1177/0093854806290162
			323		56% Caucasian, 35% Aboriginal, 5% African American, 3% Other	33	Incarceration (Federal, minimum/medium/maximum security)		
5	2012	Non-Randomized	718	Female (100%)	30% Latina, 25% African American, 33% White, 2% Asian:, 4% Native American, 6% "Other"	37	Parole	United States (California)	Brennan, T., Breitenbach, M., Dieterich, W., Salisbury, E. J., & van Voorhis, P. (2012). Women's pathways to serious and habitual crime. *Criminal Justice and Behavior, 39*(11), 1481–1508. https://doi.org/10.1177/0093854812456737
6	2009	Non-Randomized	2,328	Female (20%), Male (80%)	76% White, 15% African American, 7% Latino, 2% "Other"	Adult	Probation (Misdemeanors)	United States (Eastern Region)	Brennan, T., Dieterich, W., & Ehret, B. (2009). Evaluating the predictive validity of the compas risk and needs assessment system. *Criminal Justice and Behavior*, *36*(1), 21–40. https://doi.org/10.1177/0093854808326545
7	2016	Non-Randomized	31,583	Female (19%), Male (81%)	**Female:** 24% "Non-White"; **Male**: 32% "Non-White;"	Adult	Community Supervision	United States (Washington)	Campbell, C. M. (2016). It's not technically a crime: investigating the relationship between technical violations and new crime. *Criminal Justice Policy Review*, *27*(7), 643–667. https://doi.org/10.1177/0887403414553098
									CanBenschoten, S. (2008). Risk/Needs Assessment: Is This
9	2018	Non-Randomized	15727	Female (23%), Male (77%)	**Female**: 54% White, 34% Black, 11% Hispanic, 1% Other; **Male**: 48% White, 34% Black, 16% Hispanic, 2% Other	Adult	Probation and Incarceration (Jail)	United States (Southwestern Region)	Caudy, M. S., Tillyer, M. S., & Tillyer, R. (2018). Jail versus probation: A gender-specific test of differential effectiveness and moderators of sanction effects. *Criminal Justice and Behavior*, *45*(7), 949–968. https://doi.org/10.1177/0093854818766375
10	2010	Literature Review	N/A	N/A	Unspecified	Adult	Unspecified	Unspecified	Caulfield, L. (2010). Rethinking the assessment of female offenders. *The Howard Journal of Criminal Justice*. https://doi.org/10.1111/j.1468-2311.2010.00625.x
11	2016	Non-Randomized	2181	Female (100%)	Unspecified	30	Incarceration (Prison)	United States (Oregon)	Davidson, M., Sorensen, J. R., & Reidy, T. J. (2016). Gender-responsiveness in corrections: Estimating female inmate misconduct risk using the Personality Assessment Inventory (PAI). *Law and Human Behavior*, *40*(1), 72–81. https://doi.org/10.1037/lhb0000157
12	2016	Literature Review	N/A	N/A	Unspecified	Adults and Minors	Unspecified	Unspecified	de Vogel, V., & Nicholls, T. L. (2016). Gender matters: An introduction to the special issues on women and girls. *International Journal of Forensic Mental Health*, *15*(1), 1–25. https://doi.org/10.1080/14999013.2016.1141439
13	2019	Non-Randomized	78	Female (100%)	87% Dutch Origin	36	Forensic Psychiatric Hospital Patient	Netherlands	de Vogel, V., Bruggeman, M., & Lancel, M. (2019). Gender-sensitive violence risk assessment: Predictive validity of six tools in female forensic psychiatric patients. *Criminal Justice and Behavior*, *46*(4), 528–549. https://doi.org/10.1177/0093854818824135
14	2016	Non-Randomized	550	Female (50%), Male (50%)	84% Born in Netherlands	35	Forensic Psychiatric Hospitals	Dutch/Netherlands (Van der Hoeven, Oldenkotte Psychiatric Hospital, Woenselse Poort and Assen)	de Vogel, V., Stam, J., Bouman, Y. H., Ter Horst, P., & Lancel, M. (2016). Violent women: A Multicentre study into gender differences in forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology*, *27*(2), 145–168. https://doi.org/10.1080/14789949.2015.1102312
15	2019	Non-Randomized	1,266	Female (100%)	36% Indigenous	35	Incarceration (Federal)	Canada	Derkzen, D., Wardrop, K., & Wanamaker, K. (2019). *Women offender assessment: Can gender-informed variables improve risk prediction?* (Research Report R-413). Ottawa, Ontario: Correctional Service of Canada.
16	2020	Non-Randomized	39,355	Female (10%), Male (90%)	Unspecified	Adult	Post Incarceration (Prison)	United States (Minnesota)	Duwe, G. (2020). The development and validation of a classification system predicting severe and frequent prison misconduct. *The Prison Journal*, *100*(2), 173–200. https://doi.org/10.1177/0032885510894587
17	2012	Quantitative Descriptive	80	Female (100%)	Unspecified	36	Pre-sentencing Forensic Psychiatric Evaluation	Germany	Eisenbarth, H., Osterheider, M., Nedopil, N., & Stadtland, C. (2012). Recidivism in female offenders: PCL-R lifestyle factor and VRAG show predictive validity in a German sample. *Behavioral Sciences & the Law*, *30*(5), 575–584. https://doi.org/10.1002/bsl.2013
18	2007	Non-Randomized	2196	Female (31%), Male (69%)	41% White, not Hispanic: 38% Black 20% Latino or Hispanic, 1% Asian American, <1% American Indian	Adult	Incarceration (Jail)	United States (Connecticut)	Ford, J. D., Trestman, R. L., Wiesbrock, V., & Zhang, W. (2007). Development and validation of a brief mental health screening instrument for newly incarcerated adults. *Assessment*, *14*(3), 279–299. https://doi.org/10.1177/1073191107302944

19	2018	Non-Randomized	266	Female (39%), Male (61%)	**Female**: 44% White, 49% African American, 3% Hispanic:, 5% "Other"; **Male**: 49% White, 49% African American, 1% Hispanic, 1% "Other;"	**Male:** 31; **Female:** 27	Pre-trial Defendants	United States (Ohio)	Gehring, K. S. (2018). A direct test of pathways theory. *Feminist Criminology*, *13*(2), 115–137. https://doi.org/10.1177/1557085116646195
20	2014	Non-Randomized	266	Female (39%), Male (61%)	**Female**: 44% White, **Male**: 49% White, noting the rest to be "non-White"	**Female:** 27; **Male**: 31	Pre-trial Defendants	United States (Ohio)	Gehring, K. S., & van Voorhis, P. (2014). Needs and pretrial failure: Additional risk factors for female and male pretrial defendants. *Criminal Justice and Behavior*, *41*(8), 943–970. https://doi.org/10.1177/0093854814538022
21	2007	Non-Randomized	298	Female (100%)	43% White, 49% Aboriginal, 4% African Canadian, 3% "Visible Minority," 2% "Unknown"	34	Incarceration (Federal)	Canada	Gobeil, R., & Blanchette, K. (2007). Revalidation of a gender-informed security reclassification scale for women inmates. *Journal of Contemporary Criminal Justice*, *23*(4), 296–309. https://doi.org/10.1177/1043986207309411
22	2013	Non-Randomized	591	Female (100%)	56% Caucasian, 35% Aboriginal, 5% Black, 3% "Other/Unknown"	33	Medium Security Incarceration	Canada	Gobeil, R., Blanchette, K., & Barrett, M. R. (2013). The impact of security placement on female offenders' institutional behavior. *Crime & Delinquency*, *59*(8), 1211–1233. https://doi.org/10.1177/0011128708330102
23	2016	Non-Randomized	124	Female (19%), Male (81%)	**Female:** Ethnic Minority: 78%, Black/African: 39%, Hispanic: 22%, Asian: 0%, Other: 17%, Caucasian: 22%; **Male**: Ethnic Minority 81%, Black/African: 51%, Hispanic: 17%, Asian: 4%, Other: 8%, Caucasian: 19%	**Female**: 41; **Male**: 46	NGRI Acquittees	United States (New York State)	Green, D., Schneider, M., Griswold, H., Belfi, B., Herrera, M., & DeBlasi, A. (2016). A comparison of the HCR-20 v3 among male and female insanity acquittees: A retrospective file study. *International Journal of Forensic Mental Health*, *15*(1), 48–64. https://doi.org/10.1080/14999013.2015.1134726
24	2014	Non-Randomized	725	Female (100%)	57% Caucasian, 19% Aboriginal, 19% "Other" Minority Groups	37	Sentenced female offenders, released into the community	Canada	Greiner, L. E., Law, M. A., & Brown, S. L. (2014). Using dynamic factors to predict recidivism among women. *Criminal Justice and Behavior*, *42*(5), 457–480. https://doi.org/10.1177/0093854814553222
25	2017	Non-Randomized	47,970	Female (18%), Male (82%)	Unspecified	Adult	Supervised (felony and misdemeanor)	United States (Washington)	Hamilton, Z., Campagna, M., Tollefsbol, E., van Wormer, J., & Barnoski, R. (2017). A more consistent application of the RNR model: The Strong-R needs assessment. *Criminal Justice and Behavior*, *44*(2), 261–292. https://doi.org/10.1177/0093854816678032
26	2016	Non-Randomized	44,010	Female (20%), Male (80%)	Unspecified	Adult	Supervised (felony)	United States (Washington)	Hamilton, Z., Kigerl, A., Campagna, M., Barnoski, R., Lee, S., van Wormer, J., & Block, L. (2016). Designed to fit: The development and validation of the Strong-R recidivism risk assessment. *Criminal Justice and Behavior*, *43*(2), 230–263. https://doi.org/10.1177/0093854815615633
27	2009	Literature Review	N/A	N/A	Unspecified	Adult	Unspecified	Unspecified	Hannah-Moffat, K. (2009). Gridlock or mutability: Reconsidering "gender" and risk assessment. *Criminology & Public Policy*, *8*(1), 209–219. https://doi.org/10.1111/j.1745-9133.2009.00549.x
28	2016	Literature Review	N/A	N/A	Unspecified	Adult	Unspecified	Unspecified	Hannah-Moffat, K. (2016). A conceptual kaleidoscope: Contemplating 'dynamic structural risk' and an uncoupling of risk from need. *Dynamic Risk Factors*, 32–45. https://doi.org/10.4324/9781315266039-3
29	2014	Literature Review	N/A	N/A	Unspecified	Adult	Transition from Prison to Community Initiative (TPCI)	United States (Michigan & Arkansas)	Holtfreter, K., & Wattanaporn, K. A. (2014). The transition from prison to community initiative: An examination of gender responsiveness for female offender reentry. *Criminal Justice and Behavior*, *41*(1), 41–57. https://doi.org/10.1177/0093854813504406
30	2011	Randomized Controlled Trials	71,122	Female (16%), Male (84%)	20% Indigenous, 80% non-Indigenous	Adult	Community Supervision and Custodial	Australia (New South Wales)	Hsu, C.-I., Caputi, P., & Byrne, M. K. (2011). The level of service inventory-revised (LSI-R) and Australian offenders: factor structure, sensitivity, and specificity. *Criminal Justice and Behavior*, *38*(6), 600–618. https://doi.org/10.1177/0093854811402583
31	2015	Non-Randomized	3,656	Female (19%), Male (81%)	67% White, 17% Aboriginal, 4% Black, 3% Asian, 3% East Indian, 7% "Other"	33	Community Sentences	Canada (Alberta)	Jones, N. J., Brown, S. L., Robinson, D., & Frey, D. (2015). Incorporating strengths into quantitative assessments of criminal risk for adult offenders: The service planning instrument. *Criminal Justice and Behavior*, *42*(3), 321–338. https://doi.org/10.1177/0093854814547041
32	2001	Non-Randomized	442	Female (28%), Male (72%)	**Female:** 57% "Non-White," 43% White; **Male**: 50% "Non-White," 50% White;	30	Incarceration (Prison)	United States (Midwestern Region)	Lowenkamp, C. T., Holsinger, A. M., & Latessa, E. J. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior*, *28*(5), 543–563. https://doi.org/10.1177/0093854801028005001
33	2013	Non-Randomized	427	Female (36%), Male (64%)	**Female:** 71% White, African American 25%, 4% Other; **Male:** 44% White, 46% African American, 10% Other	34	Incarceration (Prison; Implending release)	United States (Ohio)	Makarios, M., & Latessa, E. J. (2013). Developing a risk and needs assessment instrument for prison inmates. *Criminal Justice and Behavior*, *40*(12), 1449–1471. https://doi.org/10.1177/0093854813496240
34	2010	Non-Randomized	2052	Female/Male	Unspecified	35	Adult Parole Authority Jurisdiction	United States (Ohio)	Makarios, M., Steiner, B., & Travis, L. F. (2010). Examining the predictors of recidivism among men and women released from prison in Ohio. *Criminal Justice and Behavior*, *37*(12), 1377–1391. https://doi.org/10.1177/0093854810382876
35	2009	Non-Randomized	1105	Female (6%), Male (94%)	**Female:** 67% White, 19% Black, 4% Asian, 10% Hispanic, 4% Native; **Male**: 72% White, 20% Black, 3% Asian, 10% Hispanic, 4% Native	**Female:** 37; **Male**: 39	Post Incarceration (serious violent offenses)	United States (Washington)	Manchak, S. M., Skeem, J. L., Douglas, K. S., & Siranosian, M. (2009). Does gender moderate the predictive utility of the level of service inventory—revised (LSI-R) for serious violent offenders? *Criminal Justice and Behavior*, *36*(5), 425–442. https://doi.org/10.1177/0093854809333058

36	2009	Non-Randomized	1,105	Female (7%), Male (93%)	**Female: 67%** White, 19% Black, 4% Asian, 10% Hispanic, 4% Native; **Male**: 72% White, 20% Black, 3% Asian, 11% Hispanic, 4% Native;	**Male:** 39; **Female:** 37	Post-Incarceration (violent offenses)	United States (Washington)	Manchak, S. M., Skeem, J. L., Douglas, K. S., & Siranosian, M. (2009). Does gender moderate the predictive utility of the level of service inventory—revised (LSI-R) for serious violent offenders? *Criminal Justice and Behavior*, *36*(5), 425–442. https://doi.org/10.1177/0093854809333058
37	2018	Non-Randomized	225	Female (100%)	64% White; 36% Non-White	31	Pre-Sentence, Post-Sentence, and Post Incarceration	United States (Texas)	Marshall, E. A., & Miller, H. A. (2018). Examining gender-specific and gender-neutral risk factors in women who sexually offend. *Criminal Justice and Behavior*, *46*(4), 511–527. https://doi.org/10.1177/0093854818796872
38	2019	Non-Randomized	506	Female (100%)	Caucasian: 59%; Hispanic: 21%; Black 19%; Other: 1%	30	Incarceration (Prison)	United States (Texas)	Marshall, E. A., & Miller, H. A. (2019). Arbitrary decision making in the absence of evidence: An examination of factors related to treatment selection and recidivism for female sexual offenders. *Journal of Sexual Aggression*, *26*(2), 178–192. https://doi.org/10.1080/13552600.2019.1611961
39	2020	Non-Randomized	739	Female (100%)	58% White, 20% Hispanic, 19% Black, 3% Other	38	Incarceration (sexual offense)	United States (Texas)	Marshall, E., Miller, H. A., Cortoni, F., & Helmus, L. M. (2020). The static-99r is not valid for women: Predictive validity in 739 females who have sexually offended. *Sexual Abuse*, *33*(6), 631–653. https://doi.org/10.1177/1079063220940303
40	2019	Non-Randomized	128,183	Female (9%), Male (91%)	**Female:** 47% White, 3% Hispanic; **Male**: 42% White, 6% Hispanic	**Female:** 35; **Male**: 33	Post-Incarceration	United States (Florida)	Miller, A. D., Jones, M. S., & Schleifer, C. (2019). The overall and gendered effects of postrelease supervision on recidivism: A propensity score analysis. *Criminal Justice and Behavior*, *46*(7), 1020–1043. https://doi.org/10.1177/0093854819847697
41	2007	Non-Randomized	150	Female (100%)	81% White ethnic origin, 11% African Caribbean, 6% Asian, 1% "Other"	Adult and Minors	Incarceration	England	Palmer, E. J., & Hollin, C. R. (2007). The level of service inventory-revised with English women prisoners: a needs and reconviction analysis. *Criminal Justice and Behavior*, *34*(8), 971–984. https://doi.org/10.1177/0093854807300819
42	2014	Non-Randomized	801	Female (100%)	Unspecified	28	Community Sentences	England and Wales	Palmer, E. J., Hatcher, R. M., McGuire, J., & Hollin, C. R. (2014). Cognitive skills programs for female offenders in the community: Effect on reconviction. *Criminal Justice and Behavior*, *42*(4), 345–360. https://doi.org/10.1177/0093854814552099
43	2007	Literature Review	N/A	N/A	**Female:** 59% Maori, 35% NZ European, 6% Pacific Peoples; **Male:** 52% Maori, 39% NZ European, 8% Pacific Peoples	Discusses "under 25" age group, Adult	Incarceration (Prison)	New Zealand	Poels, V. (2007). Risk assessment of recidivism of violent and sexual female offenders. *Psychiatry, Psychology and Law*, *14*(2), 227–250. https://doi.org/10.1375/pplt.14.2.227
44	2010	Non-Randomized	411	Female (100%)	74% Caucasian, 12% Aboriginal, 10% Black, 4% Another Minority Group	30	Incarceration and Community Supervision	Canada (Ontario)	Rettinger, L. J., & Andrews, D. A. (2010). General risk and need, gender specificity, and the recidivism of female offenders. *Criminal Justice and Behavior*, *37*(1), 29–46. https://doi.org/10.1177/0093854809349438
45	2009	Non-Randomized	313	Female (100%)	68% White, 30% African American, 1% Asian, 1% Hispanic/Latino, <1% Biracial	Adult	Probation	United States (Missouri)	Salisbury, E. J., & Van Voorhis, P. (2009). Gendered pathways: A quantitative investigation of women probationers' paths to incarceration. *Criminal Justice and Behavior*, *36*(6), 541–566. https://doi.org/10.1177/0093854809334076
46	2009	Non-Randomized	156	Female (100%)	52% White, 30% Black, 16% Hispanic, 2% Native American	Adult	Minimum and Medium Security Incarceration	United States (Western Region)	Salisbury, E. J., Van Voorhis, P., & Spiropoulos, G. V. (2009). The predictive validity of a gender-responsive needs assessment: An exploratory study. *Crime & Delinquency*, *55*(4), 550–585. https://doi.org/10.1177/0011128707308102
47	2020	Non-Randomized	1,100	Female (50%), Male (50%)	**Female:** 42% Maori, 48% European, 7% Pacific Peoples, 3% Other; **Male: 43%** Maori, 45% European, 8% Pacific Peoples, 4% Other	**Female:** 35; **Male:** 34	Community Supervision Sentence	New Zealand	Scanlan, J. M., Yesberg, J. A., Fortune, C.-A., & Polaschek, D. L. (2020). Predicting women's recidivism using the dynamic risk assessment for offender re-entry: Preliminary evidence of predictive validity with community-sentenced women using a "gender-neutral" risk measure. *Criminal Justice and Behavior*, *47*(3), 251–270. https://doi.org/10.1177/0093854819896387
48	2000	Literature Review	N/A	N/A	Discusses Aboriginal population	Adult	Unspecified	Canada	Shaw, M., & Hannah-Moffat, K. (2000). Gender, diversity and risk assessment in Canadian corrections. *Probation Journal*, *47*(3), 163–172. https://doi.org/10.1177/026455050004700301
49	2016	Non-Randomized	14,310	Female (50%), Male (50%)	41% African American	38	Incarceration (Substance Use-related conviction)	United States ("Federal")	Skeem, J. L., Monahan, J., & Lowenkamp, C. T. (2016). Gender, risk assessment, and sanctioning: The cost of treating women like men. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2718460
50	2014	Non-Randomized	106	Female (41%), Male (59%)	75% Caucasian	35	Post Incarceration, Post Forensic Psychiatric Hospital Patient, and Probation	Canada (Greater Vancouver Area)	Strub, D. S., Douglas, K. S., & Nicholls, T. L. (2014). The validity of version 3 of the HCR-20 violence risk assessment scheme amongst offenders and civil psychiatric patients. *International Journal of Forensic Mental Health*, *13*(2), 148–159. https://doi.org/10.1080/14999013.2014.911785
51	2010	Non-Randomized	1626	Female (100%)	Caucasian	34	Prision, Probation, and Pre-release	United States	Van Voorhis, P., Wright, E. M., Salisbury, E., & Bauman, A. (2010). Women's risk factors and their contributions to existing risk/needs assessment. *Criminal Justice and Behavior*, *37*(3), 261–288. https://doi.org/10.1177/0093854809357442
52	2017	Non-Randomized	375	Female (38%), Male (62%)	47% Caucasian/non-Hispanic, 41% African American, 12% Hispanic	Adult	Probation	United States (Southern Region)	Vaske, J. C., Gehring, K. S., & Lovins, B. (2017). Gender differences in the measurement of criminal thinking. *Criminal Justice and Behavior*, *44*(3), 395–415. https://doi.org/10.1177/0093854816677311
53	2011	Non-Randomized	645	Female (14%), Male (86%)	Unspecified	Adult	Post-sentencing (Homicide) Forensic Psychiatric Evaluation	Finland	Weizmann-Henelius, G., Grönroos, L. P. M., Putkonen, H., Eronen, M., Lindberg, N., & Häkkänen-Nyholm, H. (2011). Gender-specific risk factors for intimate partner homicide. *Journal of Interpersonal Violence*, *27*(8), 1519–1539. https://doi.org/10.1177/0886260511425793

54	2012	Qualitative	4164	Female (100%)	37% White, 31% Black, 21% Hispanic, and 11% "Other"	35	Prison and Community-Based Reentry Programs	United States (Illinois, California, & New Mexico)	White, G. D. (2012). Gender-responsive programs in U.S. prisons: Implications for change. *Social Work in Public Health*, *27*(3), 283–300. https://doi.org/10.1080/19371918.2012.629875
55	2019	Non-Randomized	77 73	Female (42%), Male (58%) Female (45%), Male (55%)	49% European American, 28% Hispanic American, 16% African American 38% European American, 32% Hispanic American, 19% African American	Adult	Court Mandated Residential Substance Use Treatment Facility	United States (Texas)	Williams, M. M., Rogers, R., & Hartigan, S. E. (2019). The validity of the picts-SV and its effectiveness with positive impression management: An investigation in a court-mandated substance use treatment facility. *Criminal Justice and Behavior*, *47*(1), 80–98. https://doi.org/10.1177/0093854819879733
56	2007	Non-Randomized	272	Female (100%)	80% White, 20% African American, <1% Asian, <1% Indian	33	Incarceration (Newly Admitted)	United States (Missouri)	Wright, E. M., Salisbury, E. J., & Van Voorhis, P. (2007). Predicting the prison misconducts of women offenders: The importance of gender-responsive needs. *Journal of Contemporary Criminal Justice*, *23*(4), 310–340. https://doi.org/10.1177/1043986207309595
57	2012	Literature Review	N/A	N/A	Unspecified	Adult	Unspecified	Unspecified	Wright, E. M., Van Voorhis, P., Salisbury, E. J., & Bauman, A. (2012). Gender-responsive lessons learned and policy implications for women in prison. *Criminal Justice and Behavior*, *39*(12), 1612–1632. https://doi.org/10.1177/0093854812451088
58	2015	Non-Randomized	266	Female (50%), Male (50%)	56% Maori, 40% NZ European, 4% Other	35	Parole	New Zealand	Yesberg, J. A., Scanlan, J. M., Hanby, L. J., Serin, R. C., & Polaschek, D. L. (2015). Predicting women's recidivism: validating a dynamic community-based 'gender-neutral' tool. *Probation Journal*, *62*(1), 33–48. https://doi.org/10.1177/0264550514562851
59	2013	Non-Randomized	664	Female (100%)	25% Aboriginal	33	Post Incarceration	Canada	Zakaria, D., Allenby, K., Derkzen, D., & Jones, N. (2013). *Preliminary development of a dynamic risk assessment tool for women: An examination of gender neutral and gender specific variables*. Research Report R-280. Ottawa, ON: Correctional Service of Canada.

APPENDIX I

Clinician Resource

Clinician Resource as Organized by Assessment Tool

Assessment Tool	Modifications	Studies Addressing Modifications
COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) Risk & Need Assessment System	Include gender-specific risk factors	**Implemented:** (Gehring, 2018) **Researched:** (Holtfreter & Wattanaporn, 2014)
	Include gender-informed scales	**Implemented**: (Hamilton et al., 2016)
Dynamic Risk Assessment for Offender Re-entry (DRAOR)	Include gender-specific risk factors	**Implemented**: (Yesberg et al., 2015)
Female Additional Manual (FAM)	Develop gender-specific risk assessment guidelines from ground up	**Implemented:** (de Vogel et al., 2019), (de Vogel et al., 2016), (de Vogel & Nicholls, 2016) **Researched:** (Green et al., 2016)
Historical Clinical Risk Management - 20, Version 3 (HCR-20 v3)	Develop gender-specific risk assessment guidelines from ground up	**Implemented**: (Strub et al., 2014)
Inventory of Need Pretrial Screening Tool (ION)	Develop gender-specific tool from ground up	**Implemented**: (Gehring, 2018), (Gehring & Van Voorhis, 2014)
	Include gender-informed scales	**Researched**: (Gehring & Van Voorhis, 2014)
Level of Service (LS) Suite	Include gender-specific risk factors	**Implemented:** (Hsu et al., 2011), (Caulfield, 2010) **Researched:** (Yesberg et al., 2015)
	Different cut-off scores by gender	**Implemented**: (Miller et al., 2019)
	Use of gender-based norms	**Implemented**: (Poels, 2007)
Minnesota Screening Tool Assessing Recidivism 2.0 (MnSTARR)	Use of gender-specific risk model	**Implemented**: (Hamilton et al., 2016)
Missouri Women's Risk Assessment	Develop gender-responsive tool from ground up	**Implemented**: (Wright et al., 2007)
Ohio Risk Assessment System (ORAS)	Different cut-off scores by gender	**Implemented**: (Hamilton et al., 2016)
Security Reclassification Scale for Women (SRSW)	Develop gender-responsive tool from ground up	**Implemented**: (Gobeil et al., 2013), (Gobeil & Blanchette, 2007), (Caulfield, 2010), (Blanchette & Taylor, 2007), (Blanchette & Taylor, 2007), (de Vogel & Nicholls, 2016)
	Include gender-informed scales	**Implemented**: (Caulfield, 2010)
Service Planning Instrument (SPIn) Suite	Develop gender-responsive tool from ground up	**Implemented**: (Greiner et al., 2014), (Derkzen, 2019)

	Include gender-specific risk factors	**Implemented**: (Gehring, 2018)
	Different cut-off scores by gender	**Researched**: (Jones et al., 2015)
Static Risk and Offender Needs Guide (STRONG) Assessment System	Use of gender-specific risk model	**Implemented**: (Hamilton et al., 2016)
	Different cut-off scores by gender	**Implemented**: (Hamilton et al., 2016), (Hamilton et al., 2017)
	Include gender-specific risk factors	**Implemented**: (Hamilton et al., 2017)
Women's Risk Needs Assessment (WRNA) Suite	Develop gender-responsive tool from ground up	**Implemented**: (Makarios et al., 2010), (de Vogel et al., 2019), (Vaske et al., 2017), (Brennan et al., 2012), (Wright et al., 2012), (Salisbury & Van Voorhis, 2009), (Gehring, 2018), (Hamilton et al., 2016), (Hamilton et al., 2017), (Miller et al., 2019), (Gehring & Van Voorhis, 2014), (Holtfreter & Wattanaporn, 2014), (Makarios & Latessa, 2013), (de Vogel & Nicholls, 2016), (Derkzen, 2019)
	Include gender-specific risk factors	**Implemented**: (Gehring, 2018)
	Develop gender-responsive trailer	**Implemented**: (Hamilton et al., 2017), (Palmer et al., 2014), (de Vogel & Nicholls, 2016)

Clinician Resource as Organized by Modification

Modifications	Assessment Tool	Studies Addressing Modifications
Develop gender-responsive tool from ground up	Women's Risk Needs Assessment (WRNA) Suite	**Implemented**: (Makarios et al., 2010), (de Vogel et al., 2019), (Vaske et al., 2017), (Brennan et al., 2012), (Wright et al., 2012), (Salisbury & Van Voorhis, 2009), (Gehring, 2018), (Hamilton et al., 2016), (Hamilton et al., 2017), (Miller et al., 2019), (Gehring & Van Voorhis, 2014), (Holtfreter & Wattanaporn, 2014), (Makarios & Latessa, 2013), (de Vogel & Nicholls, 2016), (Derkzen, 2019)
	Security Reclassification Scale for Women (SRSW)	**Implemented**: (Gobeil et al., 2013), (Gobeil & Blanchette, 2007), (Caulfield, 2010), (Blanchette & Taylor, 2007), (Blanchette & Taylor, 2007), (de Vogel & Nicholls, 2016)
	Service Planning Instrument (SPIn) Suite	**Implemented**: (Greiner et al., 2014), (Derkzen, 2019)
	The Missouri Women's Risk Assessment	**Implemented**: (Wright et al., 2007)
	Inventory of Need Pretrial Screening Tool (ION)	**Implemented**: (Gehring, 2018), (Gehring & Van Voorhis, 2014)
	Correctional Mental Health Screen for Females (CMHS-F)	**Implemented**: (Ford et al., 2007)
	Unnamed	**Researched**: (Van Voorhis et al., 2010), (Duwe, 2020), (Zakaria et al., 2013), (Derkzen, 2019)
Include gender-specific risk factors	Women's Risk Needs Assessment (WRNA) Suite	**Implemented:** (Gehring, 2018)
	Service Planning Instrument (SPIn) Suite	**Implemented:** (Gehring, 2018)
	Level of Service (LS) Suite	**Implemented:** (Hsu et al., 2011), (Caulfield, 2010) **Researched:** (Yesberg et al., 2015)
	COMPAS Risk & Need Assessment System	**Implemented:** (Gehring, 2018) **Researched:** (Holtfreter & Wattanaporn, 2014)
	Static Risk and Offender Needs Guide (STRONG) Assessment System	**Implemented:** (Hamilton et al., 2017)
	Dynamic Risk Assessment for Offender Re-entry (DRAOR)	**Implemented:** (Yesberg et al., 2015)
	Unnamed	**Researched:** (Shaw & Hannah-Moffat, 2000)
	General	**Researched:** (Hannah-Moffat, 2016)
Develop gender-specific risk assessment guidelines from ground up	Female Additional Manual (FAM)	**Implemented:** (de Vogel et al., 2019), (de Vogel et al., 2016), (de Vogel & Nicholls, 2016) **Researched:** (Green et al., 2016)
	Hare Psychopathy Checklist - Revised (PCL-R)	**Implemented:** (Shaw & Hannah-Moffat, 2000)
	Historical Clinical Risk Management - 20, Version 3 (HCR-20 v3)	**Implemented:** (Strub et al., 2014)
Different cut-off scores by gender	Service Planning Instrument (SPIn) Suite	**Researched**: (Jones et al., 2015)

	Level of Service (LS) Suite	**Implemented**: (Miller et al., 2019)
	Ohio Risk Assessment System (ORAS)	**Implemented**: (Hamilton et al., 2016)
	Static Risk and Offender Needs Guide (STRONG) Assessment System	**Implemented**: (Hamilton et al., 2016), (Hamilton et al., 2017)
	Hare Psychopathy Checklist - Revised (PCL-R)	**Researched**: (Baskin-Sommers et al., 2013), (Poels, 2007)
Develop gender-responsive trailer	Women's Risk Needs Assessment (WRNA) Suite	**Implemented:** (Hamilton et al., 2017), (Palmer et al., 2014), (de Vogel & Nicholls, 2016)
	Gender-Responsive "Trailer"	**Researched:** (Wright et al., 2007)
	Unnamed	**Researched:** (Van Voorhis et al., 2010)
	General	**Researched:** (Salisbury et al., 2009)
Include gender-informed scales	Security Reclassification Scale for Women (SRSW)	**Implemented**: (Caulfield, 2010)
	Inventory of Need Pretrial Screening Tool (ION)	**Researched**: (Gehring & Van Voorhis, 2014)
	COMPAS Risk & Need Assessment System	**Implemented**: (Hamilton et al., 2016)
	"Gender-Responsive Risk/Needs scale"	**Researched**: (Wright et al., 2007)
	General	**Researched**: (Salisbury et al., 2009)
Use of gender-specific risk model	Minnesota Screening Tool Assessing Recidivism 2.0 (MnSTARR)	**Implemented:** (Hamilton et al., 2016)
	Static Risk and Offender Needs Guide (STRONG) Assessment System	**Implemented:** (Hamilton et al., 2016)
Different "coding" by gender	Static-99R	**Implemented**: (Marshall et al., 2020)
Use of gender-based norms	Level of Service (LS) Suite	**Implemented**: (Poels, 2007)

APPENDIX J

GPS IRB Approval Notice

February 25, 2020

Protocol #: **02252020**

Project Title: **A Systematic Review of Gender-Responsive Modifications to Risk and Need Assessment Tools**

Dear Devin:

Thank you for submitting a "GPS IRB Non-Human Subjects Notification Form" for *A Systematic Review of Gender-Responsive Modifications to Risk and Need Assessment Tools* project to Pepperdine University's Institutional Review Board (IRB) for review. The IRB has reviewed your submitted form and all ancillary materials. Upon review, the IRB has determined that the above titled project meets the requirements for *non-human subject research* under the federal regulations 45 CFR 46.101 that govern the protection of human subjects.

Your research must be conducted according to the form that was submitted to the IRB. If changes to the approved project occur, you will be required to submit *either* a new "GPS IRB Non-Human Subjects Notification Form" or an IRB application via the eProtocol system (httpw://irb.pepperdine.edu) to the Institutional Review Board.

A goal of the IRB is to prevent negative occurrences during any research study. However, despite our best intent, unforeseen circumstances or events may arise during the research. If an unexpected situation or adverse event happens during your investigation, please notify the IRB as soon as possible. We will ask for a complete explanation of the event and your response. Other actions also may be required depending on the nature of the event. Details regarding the timeframe in which adverse events must be reported to the IRB and documenting the adverse event can be found in the *Pepperdine University Protection of Human Participants in Research: Policies and Procedures Manual* at https://community.pepperdine.edu/irb/policies/.

Please refer to the protocol number denoted above in all further communication or correspondence related to this approval.

On behalf of the IRB, we wish you success in this scholarly pursuit.

Sincerely,

Institutional Review Board (IRB)
Pepperdine University

cc: Mrs. Katy Carr, Assistant Provost for Research
Dr. Judy Ho, Graduate School of Education and Psychology IRB Chair