

Mad and / or bad? Jurors' attitudes towards women and men who plead insanity

by

Kendra Jennie McLaughlin

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Psychology

Carleton University

Ottawa, Ontario

©2020

Kendra Jennie McLaughlin

Abstract

Women are more likely to be perceived as having a mental disorder than men are (McGlynn, Megas, & Benson, 1976). Accordingly, legal decision-makers are more likely to attribute a woman offender's actions to mental illness in comparison to offenders who are men in insanity trials (see Yourstone, Lindholm, & Svenson, 2008). The purpose of this dissertation was to examine mock jury deliberations in a fabricated Not Criminally Responsible on Account of Mental Disorder case. I first examined the impact of defendant gender on jurors' expressions of stereotype content (warmth and competence words) and affect. I used an exhaustive Stereotype Content domain dictionary to guide my directed quantitative content analysis of mock jurors' group deliberations. I used the Linguistic Inquiry Word Count program (LIWC; see Pennebaker, Francis, & Booth, 2001) to comb deliberation transcripts to examine mock jurors' affect towards the defendant (based on the language they used). Second, I examined how juror gender relates to verdict decisions; third, I examined how juror gender relates to speaking roles in deliberations. Fourth, I conducted a thematic analysis of the deliberations and examined how themes related to defendant and juror gender. Overall, these studies did not find significant differences in jurors' use of stereotype content language or affect for men and women defendants. Moreover, I did not find a significant difference in the deliberation styles of women and men jurors. Through the thematic analysis, I found that jurors were generally focused on the mental health status of the defendant and the legitimacy of the NCRMD plea. The present research is of particular importance in Canada, where there is generally no procedural allowance for psycho-legal scholars' questioning of jurors about their social attitudes (e.g., about women) before the trial and about their deliberations after the trial. As such, this dissertation provides a unique and

exploratory “inside look” into how prospective jurors’ stereotypes and prejudices about the defendant’s gender may factor into deliberations in NCRMD trials.

Acknowledgements

I would like to extend my utmost appreciation to the many kind and brilliant people who supported me in completing my dissertation.

To my advisor, Dr. Evelyn Maeder, thank you so very kindly for your dedication, guidance, and continued support. Working under your supervision has been tremendously inspiring and enriching. I cannot imagine having a better advisor and mentor for my Ph.D. studies. To my committee members, Dr. Shelley Brown, Dr. Aaron Doyle, Dr. Marina Milyavskaya, and Dr. Regina Schuller, thank you for your valuable questions and helpful comments.

To my dear friends, Dr. Susan Yamamoto, Gabrielle Antaya, Kirstie Gibson, and Sarah Kanhai, I am grateful for your support, kindness, and all your advice throughout my university journey. Thanks also to all the wonderful lab mates I have had over these past five years, your assistance in collecting data has been beyond helpful. I consider myself incredibly fortunate to have been part of such supportive and intellectually curious group of folks.

To my psychiatrist, thank you for saving my life and helping me persevere through this endeavor. I now understand and know that bipolar disorder is only part of who I am and need not circumscribe my life.

To my loving family, thanks for all your jokes, guidance, and praise throughout the years. I think it's fair to say that I "got an education". Thank you all for supporting me and helping me believe that I can do this. Special thanks to my parents, Lise and Kent McLaughlin, for encouraging me to always be curious about the world and to fight for justice for all people.

To my brilliant partner, Dr. Michael Bueckert, your compassion, love, and dedication to social justice have helped me become a better academic and activist. I am so thankful to have you in my life.

This research was funded by the Social Sciences and Humanities Research Council.

Dedication

To my late Godfather, John McLaughlin, who always inspired me to learn and taught me to embrace life's infinite possibilities. I dedicate this thesis to you, John. Yoicks! And Away!!!

Table of Contents

Abstract.....	i
Acknowledgements	iii
Dedication	iii
List of Tables	viii
List of Figures.....	ix
List of Appendices.....	x
Background	1
Introduction.....	2
The not criminally responsible on account of mental disorder defence.....	5
Public attitudes towards the insanity defence.....	7
Attitudes	9
Measuring attitudes.....	12
Gender	13
Gender differences in psychiatric diagnoses.....	15
Gender in the courtroom	17
Defendant gender.....	17
Juror gender and speech style.....	22
Juror gender and verdict decisions.....	24
Juror gender, defendant gender, and verdict decisions.....	25
Jury decision-making.....	26
Theoretical models of jury decision-making.....	29
Project Overview	31
Method	31
Participants	31
Materials	36
Procedure	38
Analysis Overview	40
Data Preparation	42
Deliberation segmentation.....	44
Stereotype Content Domain dictionary.....	44

Results	46
Study 1: Group Analyses	46
Study 1a(i): Defendant gender and jurors' stereotype content expressions.	48
Study 1a(ii): Defendant gender and jurors' affect expressions.	48
Study 1a(iii): Jurors' stereotype content and affect expressions.	49
Study 1b: Juror gender and speech style.	50
Discussion	51
Implications.....	52
Limitations	53
Study 2: Individual Analyses	54
Study 2a: Juror and individual verdict.....	54
Study 2b: Juror gender, defendant gender, and individual verdicts.	55
Discussion	55
Implications.....	56
Limitations	56
Study 3: Quantitative Content Analysis and Results	56
Speaking turns/talkativeness.....	57
Speech interruptions	59
Sentence Tags.....	60
Assertive and affiliative language	61
Juror gender, influence, and leadership.....	62
Discussion	62
Implications.....	64
Limitations	64
Study 4: Thematic Analysis	65
Method	67
Establishing trustworthiness	70
Establishing saturation	71
Results	71
Defendant 'faking'.....	71

Diagnosis timing.....	72
Defendant’s contrived ‘story’.....	73
Family’s role and responsibility.....	73
Temporality and rationality of mental illness.....	76
Defendant-victim relationship.....	80
Psychiatrist’s credibility	81
Psychiatrist as ‘hired gun’.....	82
Psychiatrist’s expertise.....	82
Psychiatrist’s naivety.....	83
Legal-lay discourse.....	86
Jurors’ personal experiences.....	86
Meta-deliberation.....	88
Insufficient evidence.....	90
More experts needed.....	90
Acquaintance testimony needed.....	91
Verdict outcome.....	92
Incarceration focus.....	92
Consequences of incarceration.....	93
Discussion	95
Implications.....	97
Limitations	97
Study 5: Quantitative content analysis of theme relevance	98
Defendant gender’s relation to themes.....	99
Juror gender’s relation to themes.....	100
Discussion	100
Implications.....	101
Limitations	102
General Discussion.....	102
Implications.....	105
Defendant gender and jury decision-making.....	105
Attitudes in jury decision-making.....	106
Jurors’ legal inaccuracies and the misapplication of law.....	106
Juror gender and communication.....	108

Mixed-methods design	108
Limitations	111
Future directions	113
Conclusion	114
References	115

List of Tables

Table 1 <i>Demographics of jurors</i>	34
Table 2 <i>Descriptives of Stereotype Content and affective word use</i>	47
Table 3 <i>Group analyses outcomes</i>	51
Table 4 <i>Juror gender and verdicts</i>	55
Table 5 <i>Juror and defendant gender and verdict</i>	55
Table 6 <i>Sample descriptives using independent sample t-tests</i>	63
Table 7 <i>Theme and defendant gender</i>	99
Table 8 <i>Theme and juror gender</i>	100
Table 9 <i>Coded Stereotype Content dictionary</i>	150

List of Figures

Figure 1 Legal decision-making lab deliberation room **40**

List of Appendices

Appendix A. Recruitment notice on kijiji.com and initial screening	155
Appendix B. Informed consent online	156
Appendix C. Informed consent in lab	159
Appendix D. Online eligibility forms	162
Appendix E. In-lab welcome script	164
Appendix F. Pre-trial instructions script.....	165
Appendix G. Trial transcript.....	167
Appendix H. Individual verdict	177
Appendix I. Pre-deliberation instructions script	178
Appendix J. Debriefing session script	183
Appendix K. Group verdict form.....	184

Background

In 2002, Andrea Yates, an American woman, pled not guilty by reason of insanity (NGRI) to the murder of her five children. In her first trial, Yates was found guilty of capital murder; however, when the defence appealed this ruling on the basis of a legal error, she was found NGRI (Yates v. Texas, 171 S.W.3d 215). In 2011, Guy Turcotte, a Canadian man, pled not criminally responsible on account of mental disorder (NCRMD) to the murder of his two children (R v. Turcotte, 2013, QCCA 1916; educaloi.qc.ca, 2015). Turcotte was found NCRMD and thus spent 46 months in psychiatric care (educaloi.qc.ca, 2011). The prosecution successfully appealed the NCRMD verdict on the grounds that the trial judge had erred in law by not considering that the mental illness Turcotte suffered may have been self-induced (e.g., methanol intoxication; R. v. Turcotte, 2013, QCCA 1916). Turcotte was thereafter released from the mental institution and retried; he was then found guilty of second-degree murder (CBC News, 2015). The public construals of the Yates and Turcotte cases were very different: while Yates was cast as a ‘mother-gone-mad’ (see Galanti, 2003), the Turcotte case was construed as evidence for the need of tough-on-crime policies and laws (Duval, 2016). I propose that these different public narratives concerning the defendants may be the result of the Canadian and U.S. publics holding stereotypes about how gender relates to crime and mental disorders. In this dissertation, I examine whether and how jurors discuss stereotypes about gender, criminality, and mental disorders during their deliberations for a

man or woman defendant¹ who enters a plea of NCRMD for a charge of second-degree murder.

Introduction

There is no female mind. The brain is not an organ of sex. As well speak of a female liver.

—Charlotte Perkins Gilman, *Women and Economics* (1898)

Personal responsibility for one's behaviours is the basis of criminal law (Hacker & Frym, 1949). The legal standards for reasonable behaviour presume that all persons are sane or do not have a mental disorder that may influence their behaviours (Criminal Code, 1985). That is, the defendant is presumed to have intended to commit a criminal act (see Goldstein, Morse, Shapiro, & Weiner, 2003). The defendant's mental state at the time of the act (i.e., forethought, intention, and responsibility to commit the act) is integral to criminal law and can be summarised as *mens rea*, or 'guilty mind' (Malle & Nelson, 2003). The defendant's conduct (or physical actions) at

¹ In the context of my dissertation, I refer to gender in a limited manner: I refer only to adults, men or women, specifically white men or white women. Gender-stereotypic names (i.e., Gary or Gretchen) and gender pronouns (i.e., he and his or she and her) indicate the defendant's gender in the trial transcript. These names and pronouns are used to elicit the mock jurors' stereotypes about gender. Since I am interested in the social behaviours, perceptions, and expectations of the mock jurors, I refer to the defendant's gender (i.e., man or woman) rather than the defendant's sex (i.e., male or female; Posey, 2016; Pryzgodna & Chrisler, 2000). Throughout this thesis, I accordingly refer to the defendant as either the man defendant or the woman defendant. Similarly, I refer to the mock juror as man mock juror or woman mock juror and man offender or woman offender. Importantly, within feminist theory, many disagree on the construct(s) of gender and how to best understand it; no unified definition of gender exists (Haslanger, 2000). I understand gender as a socially constructed generalisation and stereotyped expectations of roles, norms, and traits (e.g., temperament, interests, and status) of those perceived to be men/boys or women/girls (see Millett, 1969), as well as the psychological and 'performative' individually determined gender identity and expression (see Butler, 1990).

the time of the act (i.e., a criminal act or the omission of an act) can be summarised as *actus reus* (*R v. Long*, 1990; Criminal Code, 1985, s. 21[1]). *Mens rea* and *actus reus* are the components required for the defendant to be held criminally responsible for a crime (*R. v. Long*, 1990). In cases where a defendant is deemed to not have had *mens rea*, it is unjust to hold the defendant accountable for his or her act (Gerber, 1984; Morse, 1985; see Monahan, 1973); he or she may thus be found not criminally responsible on account of mental disorder (NCRMD). The NCRMD defence is the Canadian version of the insanity defence (Verdun-Jones, 1994).

Researchers have found that most people, including mental health care staff, are significantly more concerned about the potential for violence of men offenders who were found insane than woman offenders who were found insane (see Coontz, Lidz, & Mulvey, 1994). Moreover, Yourstone et al. (2008) realised that clinicians and psychology students (but not judges) perceived a woman defendant's case as significantly more indicative of legal insanity than a man defendant's case, even when their cases were identical (see also Commerson, 2003). The public's perceptions of men and women's mental disorders and criminality are concerning because, as per the law, equality is the basis of justice and human rights (see UN General Assembly, 1948). If legal decision-makers' perceptions of men and women NCRMD defendants mirror those of the public and psychiatric professionals, these perceptions (e.g., attitudes and emotions) may, in turn, influence their verdicts for the defendant on trial. Furthermore, these gender-based perceptions may threaten the defendant's right to a fair and equal trial (Sygel, Sturup, For, Edberg, Gavazzeni, Howner, Persson, & Kristiansson, 2017). It is therefore important to understand how legal decision-makers (in this study, prospective jurors) discuss men and women defendants' cases.

Although numerous studies have measured how defendant gender influences jurors' verdicts (e.g., Bornstein & Muller, 2001; Nagal & Hagan, 1983; Quas et al., 2002; Russell, Ragatz, & Kraus, 2012; Spohn & Spears, 1997), those that measure jurors' specific stereotypes and emotions about defendant gender in insanity trials are far less common (Breheney et al. 2007; Mossière & Maeder, 2016). In fact, while research has demonstrated that gender stereotypes significantly impact jurors' perceptions of a woman defendant's guilt (Maeder & Dempsey, 2013), the literature has not yet examined which specific stereotypes (e.g., warmth and competence stereotypes; Fiske, Cuddy, Glick, & Xu, 2002) jurors consult when determining the criminal responsibility of men and women defendants. Furthermore, the literature has not yet examined which emotions are associated with stereotypes about the defendant within the context of jury deliberations in insanity trials.

In the jury decision-making literature, researchers have extrapolated findings based on individual jurors' verdict decisions to an entire group of jurors (Tabak, Kletke, & Knight, 2013). Few studies use full 12-person juries when examining the influence of defendant characteristics (or other extralegal variables) on verdicts. Importantly, a gender difference may exist in the contributions of men and women jurors to the deliberation; researchers have found that men and women use language differently and generally have different goals when speaking (Glenn Wiley & Eskilson, 1985). Research also reveals that in mixed-gender mock juries, men jurors are more likely to offer suggestions, opinions, and task-based directives than women jurors (see Devine, 2012). The purpose of this project is therefore twofold: First, I aim to test the tripartite model of attitudes (e.g., stereotypes, prejudices, and behaviours; Breckler, 1984) in the legal decision-making context; second, I examine jurors' gender roles (e.g., men and women's speaking roles) in deliberation.

In the following section, I first address the components of the NCRMD defence and the literature on the public perceptions of the insanity defence. I then discuss the psychological bases of gender attitudes and their relation to jury decision-making. Finally, I describe the studies that have examined mock jurors' perceptions of defendant gender, mental disorders, and criminality and how these attitudes relate to jurors' verdict decisions in NCRMD trials.

The not criminally responsible on account of mental disorder defence

Nearly all Anglo-American jurisprudences (e.g., Canada, the United States of America, the United Kingdom, and Australia) recognise a form of the insanity defence (Hacker & Frym, 1949). 'Insanity' is a legal rather than a psychiatric term. Specifically, the insanity defence is pled by a defendant who claims that he or she should not be held criminally responsible for his or her act due to his or her severe mental disorder, which precluded him or her from appreciating the wrongfulness of his or her act. Notably, the defendant is not automatically eligible to claim the insanity defence if he or she had a mental disorder at the time of the act; for the defendant to be found insane, the defence must demonstrate that the defendant's mental disorder is directly related to his or her behaviours during the crime and precluded him or her from appreciating the wrongfulness of his or her act (see Hacker & Frym, 1949). That is, the defence must prove that it is more likely than not that the defendant did not have *mens rea* when committing the act. One of the earliest tests for insanity in England was the 'wild beast' test (*Rex v. Arnold*, 1724), which states that for a man to be acquitted on the grounds of insanity, he must be 'totally deprived of his understanding and memory, and . . . not know what he is doing, no more than an infant, than a brute, or a wild beast, [as] such a one is never the object of punishment'.

The contemporary Anglo-American insanity defence was largely shaped by the 1843 trial of Daniel M'Naghten in England (see Allnutt, Samuels, & O'Driscoll, 2007, Quen, 1981). This

test for insanity rule proposes that the defendant may be acquitted due to insanity if he or she can prove that it is more likely than not that,

[...] at the time of committing the act, the accused was labouring under such a defect (disease) of the mind as not to know the nature and quality of the act he was doing; or if he did know it, that he did not know the act he was doing was wrong (see Williams, 1957).

The M’Naghten test is the guideline for the insanity defence that was adopted into the Canadian Criminal Code (1985). Section 16 of the Canadian Criminal Code states as follows:

(1) No person is criminally responsible for an act committed or an omission made while suffering from a mental disorder that rendered the person incapable of appreciating the nature and quality of the act or omission or of knowing that it was wrong. (2) Every person is presumed to not suffer from a mental disorder so as to be exempt from criminal responsibility by virtue of subsection (1), until the contrary is proved on the balance of probabilities. (3) The burden of proof that an accused was suffering from a mental disorder so as to be exempt from criminal responsibility is on the party that raises the issue.

Importantly, the defendant who pleads NCRMD must convince the triers of fact (e.g., the jury) that it is more likely than not that his or her mental disorder influenced his or her actions and ability to control or appreciate his or her actions. The law defines a mental disorder as ‘any illness, disorder, or abnormal condition that impairs the human mind and its functioning’ (National Judicial Institute, 2011).

The judge provides jurors with these legal guidelines on the standard for insanity prior to deliberation. Jurors are expected to accurately follow these guidelines and make an impartial

legal judgement (Finkel, 1995; see Finkel & Handel, 1989; Louden & Skeem, 2007).

Nevertheless, jurors tend to have limited comprehension of or adherence to the legal guidelines for insanity (Finkel, 1995). Jurors' limited adherence to these guidelines may be due to their legally inaccurate folk conceptions of the insanity defence and stereotypes about insanity defendants (Smith, 1993a; Malle & Nelson, 2003).

Public attitudes towards the insanity defence. The public are largely suspicious of the insanity defence (see Silver, Ciricione, & Steadman, 1994) and often erroneously believe that the defendants who plead insanity are immediately released into society after having been found insane. For example, Hans (1986) interviewed 330 participants and found that while the public tended to support the insanity defence in an abstract sense (e.g., people with severe mental disorders should receive treatment), many of the same participants overwhelmingly perceived that the insanity defence requires substantial reform. The participants in Hans's (1986) study may have accordingly advocated for insanity defence reform because they believed the defence was a get-out-of-jail-free plea. Moreover, the public have argued that the insanity defence is a legal loophole (Hans, 1986) and overused (Bloechl, Vitacco, Neumann, & Erickson, 2007). The public also commonly fear that the use of the insanity defence may embolden other defendants to overuse and misuse the defence to avoid incarceration (Hans & Slater, 1983; see Perlin, 1996; Silver, Ciricione, & Steadman, 1994; Skeem & Golding, 2001). In fact, of all those tried annually in Canadian provinces and territories, 1% are found NCRMD (Miladinovic & Lukassen, 2014). Of those found NCRMD in British Columbia, Ontario, and Quebec between 2000 and 2005 ($N = 1,800$), 15.6% were women (Nicholls, Crocker, Seto, Wilson, Charette, & Côté, 2015). According to data collected from Canadian policing agencies based on crimes

reported to the police in 2015, women (23.7%) are less likely to be accused of a violation of the Criminal Code than men (76.3%; Hotton Mahony, Jacob, & Hobson, 2017).

Despite the limited cases in which NCRMD is applied, the public hold negative attitudes towards the insanity defence. These negative attitudes may be rooted in the public's attitudes towards mentally ill defendants more broadly. The public tend to construe those with mental disorders (especially schizophrenia) as dangerous and unpredictable (Martin, Pescosolido, & Tuch, 2000). Such perceptions are potentially based on generalisations (i.e., stereotypes) about the causes and consequences of having a mental disorder (Angermeyer & Matschinger, 2003). According to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, a mental disorder is:

a syndrome characterized by clinically significant disturbance to an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities (American Psychiatric Association, 2013).

While much of the public understands that mental disorders are caused by the individual's biology and psychology, they nonetheless tend to perceive mental disorders as a character flaw of the person with the mental illness (Martin et al., 2000). Many are also unwilling to interact with people with mental disorders (Silton, Flannelly, Milstein, & Vaaler, 2011). Seeking to limit personal contact with people with mental disorders or seeking social distance from them is a common response of those who hold negative stereotypes about mental disorders (Pescosolido, 2013). For instance, 72.2% of Americans responded that they would be unwilling to allow someone with schizophrenia to marry into their families, and 64.1% reported that they would be

unwilling to work with a person with schizophrenia on a job (Martin et al., 2000; Pescosolido, 2013). Such stereotypes about people with mental disorders can result in systemic discrimination regarding employment (Link, 1987) and housing (see Overton & Medina, 2008). This discrimination consequentially impacts the lives and social opportunities of such underprivileged persons (see Burt, Simons, & Gibbons, 2012).

Perlin (1991; 1996; 2003; 2009; 2013) has argued that the prevalence of negative attitudes towards the insanity defence is due to the pervasive negative bias and discrimination directed towards people with mental disorders. Additionally, Perlin (2013) has contended that negative attitudes towards mentally ill persons are rooted in the stereotypical view that most people with mental disorders are uniquely irrational, deviant, and dangerous. Stereotypes about persons with mental disorders and the insanity defence may also be related to legal decision-makers' attitudes and behaviours towards defendants claiming NCRMD. Mock jurors' attitudes are therefore an essential component of the legal decision-making process, and it is thus essential to comprehend how attitudes can influence behaviours.

Attitudes

Attitudes are 'an evaluative integration of cognitions and affect experienced in relation to an object' (Crano & Prislin, 2006, p. 347). Attitudes are furthermore referred to as the 'psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor' (Eagly & Chaiken, 1993, p.1; Eagly & Chaiken, 2007). According to attitude theory (Fishbein & Ajzen, 1975; Haddock & Zanna, 1993; Krech & Crutchfield, 1948), attitudes entail three (tripartite) components: the cognitive, the affective, and the behavioural (see Breckler, 1984). These components are correlated (Ajzen, 1989; Breckler, 1984). The cognitive component comprises the thoughts directed at the attitude target; the affective component

encompasses the emotions towards the attitude target; and the behavioural component involves the actions directed towards the attitude target (Eagly & Mladinic, 1989). For example, if the attitude target is women, the cognitive component is analogous to the stereotypes (or attributes) one would ascribe to women (e.g., women are weak; see Eagly & Mladinic, 1989); the affective component is analogous to the emotions one may feel towards women (e.g., pity or empathy; see Cuddy, Fiske, & Glick, 2007; see also Ajzen, 1989); and the behavioural component is analogous to the actions towards women (e.g., a man insisting on carrying heavy groceries for a woman; see Ajzen, 1989). According to Zebrowitz (1996), 'we will not understand *group stereotypes* unless we also know who gets placed in what categories and why, and what attributes are associated with the various categories and why' (p.80). In a recent paper, Fiske (2018) demonstrates the continued strength of the stereotype content model (SCM; Fiske, Cuddy, & Glick, 2007) in explaining the social cognitive systems that people follow.

The SCM (Fiske et al., 2002) has been developed as a theory of social perceptions and describes the basic elements of stereotypes that people associate with others. Fiske et al. (2002) have demonstrated that most stereotypes are comprised of warmth and competence, which are based on the target's trustworthiness (i.e., warmth) and capacity to harm (i.e., competence). The first dimension of SCM is warmth (focused on competition for resources and interdependence). Stereotypes describing warmth include 'friendly', 'well-intentioned', and 'trustworthy' (Fiske et al., 2002). Specifically, individuals or group members who are characterised as warm are perceived as noncompetitive and nonthreatening. The second dimension of SCM is competence (focused on social status), which encompasses stereotypes such as 'confident', 'capable', and 'intelligent' (Fiske et al., 2002). Individuals or group members who are characterised as competent are perceived as powerful and able to control their actions. The theory behind the

dimensions of SCM initially built upon the work on social impression formation of Bakan (1966), who developed a theory of stereotypes based on *communion* and *agency*; these refer to the fundamental characteristics of human attitudes and behaviours. Communion refers to social connectedness (e.g., love, affiliation, and friendliness), while agency refers to individuality (e.g., dominance, power, and status; Gurtman, 2009). Furthermore, Rosenberg, Nelson, and Vivekananthan (1968) proposed the dichotomisation of stereotypes, referring to the categories as *social good-bad* (e.g., relevant adjectives include warm, sociable, good-natured, and their opposites) and *intellectual good-bad* (e.g., relevant adjectives include intelligent, scientific, determined, and their opposites).

The SCM (Fiske et al., 2002) is also a descriptive theory that illustrates how stereotypes combine and lead to group-specific prejudiced responses (Eagly, 2004; see Eagly & Karau, 2002). These stereotypes also lead to affective responses (e.g., anger and pity) that are directed at a person or group of people based on an overall evaluation or assumptions about the social group itself (Dovidio et al., 2010). Emotional reactions of fear and anger (i.e., negative affect) towards a stereotype target are generally based on the perceived level of warmth associated with the stereotype target, regardless of competence levels (Cuddy et al., 2007). Specifically, stereotype targets that are perceived as high in warmth are less likely to elicit fear or anger in perceivers, whereas stereotype targets that are perceived as low in warmth are more likely to elicit both fear and anger in perceivers.

Researchers have found that when stereotype targets are perceived to be high in both competence and warmth, they elicit positive emotions in perceivers (Fiske, Cuddy, & Glick, 2006). Perceivers are most likely to experience negative emotions towards the target when the target is stereotyped as being low in competence and warmth (Fiske et al., 2006). When a

stereotype target is perceived as being characterised by a mix of high and low levels of warmth and competence, perceivers will likely experience ambivalent emotions (i.e., a mix of positive and negative emotions) towards the stereotype target (Fiske et al., 2006; see Neal, Guadagno, Eno, & Brodsky, 2012).

Measuring attitudes. An attitude is a latent construct, and so are its components. Researchers have thus operationally defined and measured these components in several ways (Ajzen, 1989). One method of studying attitudes is to examine the stereotypes that people hold about the attitude target (e.g., women); to do this, one can look to the adjectives that people use when describing the attitude target (Ajzen, 1989). For instance, to understand the cognitive component of attitudes (the properties that characterise people's beliefs about groups of people), researchers have largely used descriptive research methods such as semantic differentials (Osgood, 1952; Osgood, Suci, & Tanenbaum, 1957). Semantic differentials (SD; Osgood, 1952) are a tool used to measure judgements (positive or negative attitudes) towards an attitude object (e.g., a person, a group of people, or ideas). Moreover, by asking participants to rate the attitude target on a series of bipolar scales of contrasting adjectives (e.g., easy-difficult, irrelevant-important, and kind-cruel), the semantic differential tool is aimed at measuring the cognitive (i.e., beliefs or stereotypes) component of attitudes (Ajzen, 1989). The three main factors of the SD are *evaluation* (i.e., the extent to which the attitude object is good-bad), *potency* (i.e., the extent to which the attitude object is strong-weak), and *activity* (i.e., the extent to which the attitude object is active-passive).

In an experiment, Eagly and Mladinic (1989) asked students from the United States to answer questionnaires assessing their attitudes towards men and women (and people of the Democratic and Republican parties as control items) on a semantic differential scale, rating

stereotype targets as pleasant-unpleasant and nice-awful). Participants were also asked to list five attributes that they think are stereotypically associated with the four attitude objects (e.g., free response items) and rate how positively or negatively they feel about the attitude object. Eagly and Mladinic (1989) found a significant correlation between positive stereotypes (e.g., warm and kind) towards an attitude target (e.g., women) and positive emotions (e.g., feeling that the stereotype target is good, positive, and pleasant). Similarly, negative stereotypes were associated with experiencing negative emotions towards the attitude target. The semantic differential method of evaluating the cognitive component of attitudes has also been integrated into the widely popular SCM (Fiske et al., 2002) of social perceptions (Kervyn, Fiske, & Yzerbyt, 2013). For example, people who are perceived as being low in competence but high in warmth (e.g., the elderly) may be treated in a patronising but sympathetic manner (i.e., paternalism; see Fiske et al., 2002; Fiske, 2012). Researchers have confirmed that individuals' perceptions of targets' warmth and competence are contingent upon the gender of the target (Cuddy, Fiske, & Glick, 2008; see Eagly & Steffen, 1984); women have been found to be construed as higher in warmth than men, and men tend to be construed as higher in competence than women (see Cuddy et al., 2008). It is therefore important to consider how the gender of the defendant may influence jurors' perceptions of the defendant as well.

Gender

Gender is relevant to most, if not all, institutions (e.g., health care, government, and legal institutions) due to patriarchy (Millett, 1969). Patriarchy is generally a system of social organisation in which men rule and dominate (socially, politically, and economically) and women are subordinate. Patriarchy is accordingly enacted in and reinforced by gender roles and stereotypes (Glick et al., 2000). Gender roles are social expectations for behaviours deemed

appropriate for men and women. Furthermore, gender roles are generally based on the stereotype that women are physically and psychologically weaker and less aggressive than men (Glick & Fiske, 2000; see Weare, 2013). Notwithstanding these stereotypes about men and women's behaviours and abilities, Hyde's (1984) meta-analysis has demonstrated that overall, men and women are more similar than different (or at least only marginally different) in their characteristics and capacities (e.g., aggression, $d = 0.50$). Despite the similarities of their characteristics, men and women are nonetheless treated differently by institutions, such as the criminal justice system.

Institutions such as the criminal justice system are structured to maintain this gender hierarchy between men and women (Millett, 1969). Glick et al. (2000) have argued that patriarchy creates and allows for hostile sexism—the negative attitudes and behaviours directed at women. Specifically, women who behave in a non-gender stereotypic manner may face hostile sexism from other women and men (Glick & Fiske, 1997). Conversely, women who behave in a gender stereotypic manner are treated with benevolent sexism—condescending attitudes and behaviours directed towards women (Glick & Fiske, 1997). Both forms of sexism entail attitudes and behaviours aimed at maintaining the patriarchy; they penalise women whose behaviours contradict gender stereotypes and patronise those who adhere to gender stereotypes.

The influences of hostile and benevolent sexism are apparent in legal decision-making. For instance, Herzog and Oreg (2008) demonstrated that women offenders received more lenient judgments than men offenders, so long as the women offenders were also perceived as feminine (Herzog & Oreg, 2008). This leniency effect was particularly notable when the study participants also held benevolent sexist opinions about women (the participants in the study exhibited very low ratings of hostile sexist opinions; Herzog & Oreg, 2008). Similarly, the chivalry/paternalism

hypothesis proposes that women offenders receive more lenient media attention than men offenders (Grabe, Trager, Lear, & Rauch, 2006). Chivalry/paternalism, similar to benevolent sexism, aims to 'protect' women, be it from society or other men.

Doctors and other medical professionals have long been responsible for determining whether women's malaise and 'inappropriate' behaviours should be attributed to mental disorder ('madness') or poor character ('badness'; Appignaesi, 2008; Chesler, 2005; see Laberge, Morin, & Armony, 1997; Ussher, 1991). Psychiatrists and psychological professionals have arguably served as gatekeepers for determining which emotions, attitudes, and behaviours are considered normal or deviant (Szasz, 1974; Ussher, 1991). For example, women who were perceived as violating gender role norms and expectations (e.g., single women and educated women) were once commonly targeted for psychiatric institutionalisation (Chesler, 2005; Ussher, 1991/2013).

Gender differences in psychiatric diagnoses. Gender differences in psychiatric diagnoses and treatment have been longstanding topics of controversy and consequence (Eriksen & Kress, 2008; Kaplan, 1983). While the psychiatric treatments provided to women may no longer explicitly enforce gender stereotypes (see Showalter, 1985), implicit sexism continues to be evident in the psychiatric diagnoses of women seeking psychiatric treatment. For instance, women are more likely than men to be diagnosed with a mental disorder and prescribed psychotropic drugs (Astbury, 2001; Grigoriadis & Erlick Robinson, 2007; see Stoppard, 2000).

To diagnose any person with a mental disorder, clinicians and psychiatrists typically rely on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), which outlines the symptoms a client must present to be diagnosed with a mental disorder. Recently, the DSM-5 (American Psychiatric Association [APA], 2013) has added an important note indicating that a client's gender can influence their symptomology and diagnosis. Research suggests that the

psychiatric diagnoses men and women receive are often gender contingent. Women are also more likely than men to be perceived by clinicians and psychiatric decision-makers as dependent, helpless, and pitiful persons (Boysen et al., 2014; Commerson, 2003; Ottati, Bodenhausen, & Newman, 2005). Specifically, in societies where gender role norms are pronounced and sexism is explicit, mental health patients are more likely to be diagnosed with gender-stereotypic mental disorders; specifically, women are more likely to be diagnosed with internalising disorders such as anxiety and depression, while men are more likely to be diagnosed with externalising disorders such as antisocial personality disorder and substance use disorders (Eaton et al., 2012; Seedat et al., 2009).

Moreover, while men and women are diagnosed with schizophrenia at equivalent rates (Government of Canada, 2006; Sygel et al., 2017; Veysey, 2015), researchers have found that a stereotypical association may exist between masculinity and schizophrenia (Boysen et al., 2014). This association may be due to overlapping stereotypes about the dangerousness of men and persons with mental disorders. The fact that masculinity and schizophrenia are perceived to overlap is important, considering that men defendants (71.8%) are more likely to be diagnosed with psychotic spectrum disorders (e.g., schizophrenia) than women defendants (66.2%) pleading NCRMD (Nicholls et al., 2015). However, the Canadian women who were found NCRMD were as likely as men to be charged with crimes against a person (Nicholls, Crocker, Seto, Wilson, Charette, & Côté, 2015). Furthermore, women who were found NCRMD were significantly more likely to have caused or attempted to cause death (9.6%) than the men who were found NCRMD (6.3%; Nicholls et al., 2015; see Veysey, 2015). Despite these statistics, the degree to which mentally ill persons are stereotyped as potentially violent may be contingent upon gender. Forensic psychiatric professionals' perceptions about an inpatient's gender, mental

disorder, and dangerousness may also influence their perceptions of the inpatient's criminality. For example, Allen (1987) found that forensic psychiatrists perceived men who committed violent crimes as legally and morally guilty, whereas psychiatrists perceived women who committed violent crimes as having little to no liability for their acts due to their lack of moral blameworthiness. Importantly, no significant difference in the violence perpetrated by men and women forensic psychiatric inpatients while in hospital exists (Nicholls, Brink, Greaves, Lussier, & Verdun-Jones, 2009; Nicholls, Ogloff, & Douglas, 2004).

Gender in the courtroom

Defendant gender. Legal decision-makers' gender stereotypes are apparent in the courtroom, though the effect of gender stereotypes on verdicts is varied (Abrams, Viki, Masser, & Bohner, 2003; Masser, Lee, & McKimmie, 2010; McKimmie, Masser, & Bongiorno, 2014; Viki & Abrams, 2002; Viki, Abrams, & Masser, 2004). Two meta-analyses on the influence of various defendant characteristics (including gender) on mock legal decision-makers' judgments have found that women defendants are not significantly less likely than men defendants to be found guilty ($r = 0.02$, 95% CI = $-0.04/0.08$ [Devine & Caughlin, 2014] and $r = -0.04$ [Mazzella & Feingold, 1994]). Herzog and Oreg (2008) demonstrated that legal decision-makers are more lenient towards a woman defendant than a man defendant, so long as the woman defendant is also perceived as feminine (i.e., high warmth).

The mixed effects of defendant gender on verdict decisions may suggest that moderators influence jurors' perceptions of men and women defendants' guilt. Researchers have theorised that women offenders receive leniency from the courts due to the chivalry or paternalism of the triers of fact (see Moulds, 1978; Nagel & Hagan, 1983). Jurors' gender role attitudes about how men and women ought to act may thus influence their verdict decisions. Legal decision-makers'

verdicts may be influenced by their stereotypes of the man or woman defendant. For instance, jurors who perceive a woman defendant as weak and vulnerable—and importantly, without agency (e.g., has no control over her criminal actions; Moulds, 1978)—may be more lenient towards her. Part of this leniency may also be based on the fact that women offenders are perceived as less physically dangerous or threatening to society than men offenders (see Allen, 1987, p. 101; Schnittker, 2000).

In a review of the limited literature on the impact of defendant gender on juror verdicts, McKimmie and Masser (2010) highlighted that the research on defendant gender and jurors' verdicts presents paradoxical results: research suggests that being a woman defendant can lead jurors to be lenient towards her but can also result in harsher attitudes and judgments of her in comparison to a man defendant. For example, although a woman defendant may match the stereotype of the irrational person, her actions may be attributed to her poor character and her violation of gender stereotypes. Indeed, women offenders who violate gender norm expectations (e.g., victimise their children) are depicted negatively in news media as well (Grabe et al., 2006). This harshness towards women who violate gender roles may be because these women are perceived as threatening to society, specifically to the patriarchy (see Herzog & Oreg, 2008). That is, women offenders violate gender stereotypes, and their behaviour is threatening to the status quo of women as docile and communal (see Glick & Fiske, 2011). In fact, Maeder and Dempsey (2013) found that when mock jurors perceived a woman defendant as more feminine (more gender stereotypic), they tended to perceive her as more credible and likeable, and in turn, they were less certain of her guilt.

Strub and McKimmie (2016) proposed that jurors' perceptions of the defendant's congruence or incongruence with offender and gender stereotypes influence their final verdict

decisions. To test this hypothesis, Strub and McKimmie (2016) presented mock jurors with a fictional court transcript that detailed a case where the (man or woman) defendant was charged with murder. Strub and McKimmie (2016) manipulated the defendant's femininity or masculinity (i.e., gender stereotypes) by changing the adjectives used to describe the target's behaviour (e.g., passive for feminine and domineering for masculine) and then asked the mock jurors to describe their stereotypes of a typical murder perpetrator (e.g., offender stereotypes). They found that regardless of whether men defendants matched gender stereotypes of men, they were more likely to be found guilty than women defendants; women defendants were evaluated less positively when they were described using masculine traits compared to feminine traits (Strub & McKimmie, 2016). Strub and McKimmie's (2016) results ultimately suggest that women defendants are compared against gender stereotypes, whereas men defendants are compared against the offender stereotype. Jurors' perceptions of a typical murder perpetrator varied more when the defendant was a woman than when the defendant was a man (Strub & McKimmie, 2016). Furthermore, researchers have proposed that men defendants tend to be found guilty more often than women defendants and suggest that this difference is potentially due to masculine behaviours being associated with criminality (Hendree & Nicks, 2000; Steffensmeier & Kramer, 1982). Men defendants may therefore be perceived as more responsible for their acts than women defendants because their behaviours are more likely to match legal decision-makers' gender and criminality stereotypes (see McKimmie, Masters, Masser, Schuller, & Terry, 2013).

In an experimental study, McKimmie et al. (2013) asked participants to read a 12-page trial transcript that described a service station robbery scenario. The defendant's gender was manipulated by changing the gender-specific pronouns and first name (Mrs Ann or Mr John)

throughout the transcript. The case strength was also manipulated by varying the details of the case in the witness testimony (i.e., a scenario with confident and reliable eyewitness testimony or a scenario with questionable eyewitness testimony). After reading one of the trial transcripts, participants were asked to provide a verdict (i.e., guilty or not guilty) to rate the convincingness of the evidence and whether men or women are more likely to commit aggravated robbery. McKimmie et al. (2013) found that mock jurors considered the strength of the evidence against the men defendants (i.e., stereotypical offenders) but not the women defendants (i.e., counter-stereotypical offenders). These findings are consistent with the continuum model of impression formation (Fiske & Neuberg, 1990), which purports that stereotypes about a target vary as a function of the characteristics of the stereotype target; in the case of my dissertation, this may suggest that jurors express different stereotypes (e.g., gender stereotypes or criminal stereotypes) when deliberating the man or woman defendant's case.

Moreover, within the context of defendant gender in insanity trials, Breheney et al. (2007) hypothesised that mock jurors would be more likely to find a woman defendant (in comparison to a man defendant) not guilty by reason of insanity (NGRI). To test this hypothesis, Breheney et al. (2007) asked 258 undergraduate college students (i.e., mock jurors) to read a trial vignette that described the discovery of a homicide, the crime scene, police testimony, and the psychological evaluation of the defendant. The descriptions of the defendant's gender (e.g., man or woman) and mental disorder (e.g., depression, schizophrenia, bipolar disorder, post-traumatic stress disorder, situational stressors, or substance-induced disorder) in an insanity trial were experimentally varied, and the mock jurors' verdict decisions were measured. Breheney et al. (2007) found that regardless of mental disorder conditions, the women defendants were significantly more likely to be found guilty than the men defendants. Moreover, these women

defendants were perceived as significantly more responsible for and in control of their criminal actions than the men defendants (Breheney et al., 2007).

Within the Canadian context, Mossière and Maeder (2016) found no main effect of defendant gender on mock jurors' verdict in an NCRMD trial. They did, however, find a significant effect of defendant gender on the attributions that mock jurors made about the defendant who was charged with second-degree murder. Specifically, mock jurors made more internal attributions for the woman defendant compared to the man defendant, indicating that participants were more likely to attribute the criminal act to the woman defendant's personality (Mossière & Maeder, 2016).

The research on defendant gender and legal decision-making collectively indicates that the outcomes of defendant gender on verdict may vary and that the variance in verdict decisions may be contingent upon the different stereotypes that jurors consult when evaluating men and women defendants. That is, stereotypes about gender and crime significantly impact jury decision-making (McKimmie et al., 2013). In this study, I have thus examined which stereotypes and which emotions (Study 1) jurors express while deliberating a man or woman defendant's case. I was interested in examining how the expressions of stereotypes and emotions are related; according to the SCM, the expression of certain stereotypes should correlate to specific emotions. I have also examined whether a relationship between stereotype and emotion expression exists in jury deliberations. Within the context of the present studies, I have explored how jurors' stereotypes about the defendant relate to the emotions they feel while deliberating and their final verdict decisions. Prior to my research, no research had examined the specific stereotype content (e.g., warmth and competence) that jurors consult when evaluating men and women defendants who have pled NCRMD. Based on the literature on gender stereotypes and

related emotional and behavioural outcomes, I proposed that when mock jurors perceive the woman defendant as acting in accordance with gender stereotypes, mock jurors may be more likely to express stereotypes about high warmth and low competence, and less anger or fear, than when they perceive the woman defendant as acting against gender stereotypes (see Herzog & Oreg, 2008). Furthermore, I have proposed that the man defendant in an NCRMD case may be more likely to match legal decision-makers' stereotype of a dangerous offender than the woman defendant in an NCRMD case.

The gender stereotypes that jurors may ascribe to defendants (e.g., warm, not competent, and associative) may also be ascribed to fellow jurors: Jurors may expect the women jurors to also generally be more docile and empathetic. Moreover, the gender stereotypes that jurors may ascribe to the man defendant (e.g., cold and competent) may also be ascribed to the men jurors. That is, jurors may expect the men jurors to be assertive and indifferent. Importantly, jurors' gender may influence how they perceive and process trial information; specifically, men and women jurors may discuss the trial evidence in different manners (e.g., assertive versus tentative language use). In the following sections, I outline the literature on how juror gender may influence not only verdict decisions but also jurors' speech styles and juror dynamics.

Juror gender and speech style. Women tend to use language to foster connection, understanding, and inclusivity when speaking with an individual or a group of people (Wood, 1995). Women accordingly tend to use tentative language more often than men (Lakoff, 1975). For instance, women are far more likely to use verbal hedging (e.g., 'I'm not an expert, but . . .'), tag a question onto a statement (e.g., 'That's incorrect, isn't it?'), and use minimal response queues (e.g., 'yeah', 'ummm', and 'hmmm') than men (Lakoff, 1975). Overall, women's conversations can be described as 'interactive', where turn-taking is indistinct, as women tend to

speak in an inclusive manner (e.g., by asking questions and drawing others into the conversation; Wood, 1995).

In a meta-analysis of adults' language use, Leaper and Ayres (2007) found significant, though small effect sizes, which demonstrated that men tend to be more talkative (e.g., talk more often and for longer in a group, $d = -0.14$, 95% CI = $-0.19/-0.08$) and more assertive (e.g., more confident in their statements, $d = 0.09$, 95% CI = $0.02/0.15$) than women. However, men are also less affiliative (e.g., make fewer attempts at fostering connections with other speakers, $d = 0.12$, 95% CI = $0.06/0.18$) than women. Men's language use tends to be aimed at exerting control over the conversation, preserving their independence by challenging others' statements, and enhancing their status (e.g., their skills, knowledge, and abilities) within the group (Wood, 1995). Overall, men's conversations can be described as 'forceful and dominative'; men tend to use others' statements as starting points and interrupt others to reroute conversations to make their own points (see Wood, 1995).

Whereas women's speech tends to be aimed at fostering an associative or collaborative dialogue to problem solve, men's speech tends to be aimed at controlling the group's discussion (e.g., telling others what they should or should not do or think) to problem solve. Notably, researchers have found that women tend to be more tentative in discussions in mixed-sex dyads (i.e., women speaking with men) than in same-sex dyads, conversely men are generally assertive in their discussions (Carli, 1990).

Considering the consistent differences in men and women's speech styles, I anticipated that juror gender may relate to how jurors speak in their deliberations. Specifically, I anticipated that juror gender may relate to the assertiveness (e.g., directives, disagreement, and negative speech; see Leaper & Smith, 2004) and tentativeness (e.g., hedging, minimal listening responses,

praise, and responsiveness; see Leaper & Smith, 2004) of their utterances during deliberation. In Study 1d, I thus examined how juror gender relates to speech styles (e.g., tentative or assertive language) in jury deliberation. I hypothesised that women jurors may be more likely than men jurors to make tentative statements; conversely, I anticipated that men may make more assertive statements than women in jury deliberations. To date, limited research has examined how the speech styles of jurors relate to jury decision-making and jurors' verdicts (Marder, 1987).

Juror gender and verdict decisions. Researchers have generally found that juror gender is significantly (though variably) related to jurors' final verdict decisions (Breheney et al., 2007; Pozzulo et al., 2010). Importantly, the influence of juror gender on verdict decisions seems to vary depending on the crime for which the defendant has been charged. Much of the literature that examines how juror gender relates to verdict decisions has centred on sexual offence cases (Guy & Edens, 2003; McNamara, Vattano, & Viney, 1993; Schutte & Hosch, 1997). In such cases, women jurors tend to be more likely than men jurors to find the defendant guilty (see Schutte & Hosch, 1997). For example, Pozzulo et al. (2010) found that women jurors deliver higher guilt ratings to the defendant (e.g., are more certain of the defendant's guilt) than men jurors in a sexual assault of a child case. Similarly, Dunlap, Hodell, Golding, and Wasarhaley (2012) found that women jurors were significantly more likely than men jurors to find the defendant guilty of stalking a former intimate partner. In an experiment aimed at examining jury decision-making in insanity trials, Mossière and Maeder (2015) found that juror gender was related to jurors' mental illness attitudes, case judgements, and final verdict decisions; women jurors (compared to men jurors) tended to have less negative attitudes about mental illness and were more certain that the defendant was guilty. Mossière and Maeder's (2015) findings are consistent with those of Breheney and colleagues (2007), who found that women jurors

(compared to men jurors) were more sympathetic to the mental illness of the defendant and saw the mental illness as influencing the defendant's act.

These studies collectively demonstrate that the influence of juror gender on verdicts tends to be significant, though varied, per case (e.g., sex offence trials, capital trials, and insanity trials; Devine, 2012). While the literature consistently demonstrates a relationship between juror gender and verdict decisions, most, if not all, of this research has involved only individual juror verdicts (i.e., no deliberation). For Study 2, I therefore examined how juror gender relates to jurors' individual verdicts.

Juror gender, defendant gender, and verdict decisions. Jurors' verdicts may additionally be influenced by the interaction between defendant and juror gender based on social identity. According to social identity theory (SIT; Tajfel & Turner, 1979; 1986), every person one interacts with can be categorised as either belonging to the same social group as oneself (i.e., in-group) or a different group that one is not part of (i.e., out-group; Ellemers & Haslam, 2011). For instance, a person who identifies as a woman may perceive other women as part of her gender in-group and men as part of her gender out-group. Part of a person's self-esteem comes from the social groups (e.g., gender) they identify with (see Trepte & Loy, 2017). People are motivated to feel positively about their in-group because a positive sense of the social group is integral to the self-esteem of the individual (i.e., collective self-esteem; see Crocker & Luhtanen, 1990; Luhtanen & Crocker, 1992). This motivation to maintain a positive sense of self and in-group can lead to in-group favouritism or leniency (Tajfel & Turner, 1979; 1986). For example, a woman's opinion of a woman who commits a crime may be less negative than her opinion of a man who commits the same crime. Furthermore, to distinguish the in-group from the out-group, individuals may also be motivated to denigrate out-group members to psychologically distance

the out-group from their in-groups (see Tajfel & Turner, 1979). In-group favouritism and out-group negativity are collectively referred to as in-group bias (Taylor & Hosch, 2004).

In-group bias may be particularly relevant in the context of jury decision-making in which jurors are determining the guilt of an in-group or out-group member. The social similarity or difference (e.g., in terms of social identity) between the individual juror and the defendant may influence the juror's perceptions of and verdict for the defendant. Women jurors may accordingly be more lenient towards women defendants, and men jurors may be more lenient towards men defendants, because they are of the same social group (i.e., gender; Billig, 1976; Tajfel & Billic, 1974; Turner, 1975).

In Study 2, I therefore examined how defendant and juror gender interact to influence jurors' individual verdicts. I anticipated that women jurors may be more likely than men jurors to find a woman defendant NCRMD (i.e., in-group favouritism) and a man defendant guilty (i.e., out-group negativity). Similarly, I anticipated that men jurors may be more likely than women jurors to find the man defendant NCRMD (i.e., in-group favouritism) and the woman defendant guilty (i.e., out-group negativity).

Jury decision-making

One of the first empirical examinations of the importance of juries in criminal trials was the Chicago Jury Project (Broeder, 1959; Kalven & Zeisel, 1966). This project involved many studies and methodologies aimed at better understanding how juries function in the United States, such as the large-scale interview study in which researchers analysed more than 1,500 jurors from 225 criminal trials in Brooklyn and Chicago (Broeder, 1959). The researchers of the Chicago Jury Project (Broeder, 1959; Kalven & Zeisel, 1966) also developed the first experimental jury paradigm study, wherein they presented 98 mock criminal juries with one of a

few different tape recordings of a mock trial. This study was the first to demonstrate how manipulated independent variables influence mock jurors' final verdicts (Broeder, 1959).

Since the Chicago Jury Project (Broeder, 1959), several studies have used a similar jury simulation paradigm (Bornstein, Golding, Neuschatz, Kimbrough, Reed, Magyarics, & Luecht, 2017; Wiener, Krauss, & Lieberman, 2011) to understand how jurors perceive defendants in various cases and how their verdicts are affected by these perceptions (see Devine, 2012). To date, many of these jury simulation studies have used student samples rather than community, non-student samples (Bornstein et al., 2017). Some concern and criticism that these jury studies may lack ecological validity has emerged because real juries would be unlikely to contain many students (e.g., due to the hardship jury duty would present for students; see the Ontario Ministry of the Attorney General, 2018). Nevertheless, despite the demographic differences in student and non-student samples (e.g., age), researchers have found no significant difference in the verdict decisions of student and non-student jury samples (see Bornstein et al., 2017; Maeder, Yamamoto, & McManus, 2017).

A common difference between the jury research paradigm and actual juries is that jury research typically studies individual mock jurors rather than jurors working together (e.g., deliberation). Research that compares group decision-making to individual decision-making (Clark, Stephenson, & Kniveton, 1990; Patry, 2008; Smith, Bushouse, & Lord, 2010; Vollrath, Sheppard, Hinsz, & Davis, 1989) has demonstrated that participants who provided individual verdicts remembered evidence less accurately than participants who made group verdict decisions (Clark et al., 1990). The pre-deliberation (or individual) verdicts common to jury decision-making research may thus be somewhat misleading; jurors' perceptions of the defendant and verdicts are likely to change based on group deliberation and the verdict

preferences of other jurors (Salerno & Diamond, 2010). Moreover, mock jurors may change their perceptions of the defendant post-deliberation (e.g., become more punitive towards the defendant; Lynch & Haney, 2009) and become more or less confident in their individual verdicts depending on the number of mock jurors who agree or disagree with their pre-deliberation verdicts (Hannaford-Agor, Hans, Mott, & Munsterman, 2002). This study aims to also broaden the jury decision-making literature from the individual verdict tradition to a more externally and ecologically valid group deliberation.

Researchers have suggested simulating jury decision-making by using focus groups (Tabak & Klettke, 2014), which involve open (or semi-structured) discussions in which participants openly share and deliberate a topic of social significance (Rothwell, Anderson, & Botkin, 2016). Focus groups are generally aimed at gathering in-depth information about people's attitudes on a topic (e.g., the insanity defence). Through the deliberation, jurors revise and mould their perceptions of the case and the defendant based on other jurors' understandings of the case; as a group, the jury 'makes sense of' the case and comes to a collectively formed verdict decision (Tabak & Klettke, 2014).

Researchers have found that while jurors tend to individually form verdicts prior to deliberation, they do not tend to hold these initial decisions very strongly; jurors' tentative verdicts become more solidified only after the jury's deliberation (Hannaford-Agor, Hans, Mott, & Munsterman, 2000; Salerno & Diamond, 2010). Jury decision-making research has demonstrated that jurors consider extralegal information (e.g., the defendant's race and gender) when making judgments about legal issues (Ford, 1986; Devine & Caughlin, 2014; Mazzella & Feingold, 1994; Visher, 1987), yet few studies have examined how jury deliberation can influence jurors' use of stereotypes. For instance, to my knowledge, no research has examined

whether jurors express gender stereotypes when deliberating a man or woman defendant's case. According to theoretical models of jury decision-making, jurors' stereotypes about the defendant, as well as their own identities, significantly influence their final verdict decisions (Devine, 2012). The following section contains more detailed information on the process of jury decision-making and how information about the trial, the defendant, and the juror are all integrated to form a final verdict decision.

Theoretical models of jury decision-making. A notable descriptive model aimed at explaining the process of jury decision-making is the story model (Pennington & Hastie, 1986; 1992). This theoretical starting point for the process of jury decision-making outlines the cause-and-effect relationship between how jurors process trial evidence in a narrative fashion (i.e., story building) and how they reach a verdict decision (Pennington & Hastie, 1986). The story model consists of three phases: The first phase is story construction, which involves jurors fitting trial evidence and case facts with their own mental representations of the crime and the defendant based on their own life experiences (e.g., attitudes about crime) and overall perceptions of the trial (Groscup & Tallon, 2009). Attitudes are thus an important part of jury decision-making (though they are extralegal and should not influence jurors' decisions) and influence how jurors find the defendant. For the second phase, Pennington and Hastie (1986) proposed that jurors process the judge's instructions about the verdict options and the components of the law that must be considered when making a verdict decision. As to the last phase, story classification, Pennington and Hastie (1986) found that jurors tend to fit their collective story construction about the evidence and the case facts to their understandings of the law and verdict options to reach a final verdict decision.

Jurors ultimately search for a ‘goodness of fit’ or convergence between their story of the evidence and the verdict. Pennington and Hastie (1986) have found that jurors form spontaneous mental representations or stories to explain the events of the case; importantly, jurors are not making legal decisions based purely on the rational and standardised processing of only the evidence presented at trial. The story a jury creates is influenced by the group members’ life experiences, attitudes, and perceptions of the trial (Devine & Caughlin, 2014). Notably, jurors may not have a uniform story about the trial. That is, each juror on a jury may hold different individual cognitive structures or stories about the trial and may thus form different individual verdict decisions. A recent theoretical model that accounts for both group- and individual-level decision-making processes and story construction is Devine’s (2012) Director’s Cut Model. This integrative model expands on Pennington and Hastie’s (1986) Story Model.

Whereas the Story Model asserts that jurors form a story based on evidence and judges’ instructions that they are presented during trial, the Director’s Cut Model expands on this by considering how jurors also initially draw from both their own (e.g., race, gender, and trust in the legal system) and the defendant’s characteristics (e.g., race, gender, and prior criminal record) to make a verdict decision (Devine, 2012; Devine & Caughlin, 2014) and to construct a story about the trial that shapes their verdict decision. That is, the Director’s Cut Model proposes that jurors initially perceive the trial and verdict information through the lenses of their personal and defendant characteristics, rather than simply construing a story about the evidence on trial that is largely free of personal experience. Importantly, men and women tend to have different attitudes about mental illness and the insanity defence (Mossière & Maeder, 2015). I therefore propose that men and women may not only deliberate differently (e.g., speech style differences) but also tend to have different individual verdict decisions for defendants in NCRMD cases.

Project Overview

This dissertation aims to examine how gender influences jury decision-making in deliberations for insanity trials. To realise this aim, I analysed mock jury deliberation transcripts (gathered from audio and video footage of deliberations). These deliberations involved community members (i.e., mock jurors) discussing the details of a second-degree murder trial, describing either a man or woman defendant's case involving an NCRMD claim due to schizophrenia. Specifically, Studies 1a and 1b focused primarily on jurors' use of stereotype content words and affect expression during deliberations, while Studies 2a and 2b examined the relationship between juror and defendant gender and jurors' individual verdicts. Studies 3, 4, and 5 are exploratory analyses of the content and themes of the deliberations; the thematic analysis explored the components of the themes that emerged from the deliberation data and the quantitative content analysis assessed the frequencies of theme discussions.

Method

Participants

The participants ($N = 77$; 53.20% women, 46.80% men) were community members from the National Capital Region of Canada. The average age of the participants was 29.84 years old ($SD = 11.97$ years old). Most participants were White (64.90%), 10.40% were Black, 9.10% identified as another race, 6.50% were Indigenous, 2.60% were Asian, 2.6% were Middle Eastern, 2.60% were Hispanic/Latinx, and 1.30% identified as South Asian.

These participants were recruited in the Spring of 2016 and in the Spring of 2017. As there had been a risk of participants communicating to other potential jurors about this study and its purpose, I halted participant recruitment from July to December 2016 so as to limit the influence of social threats (e.g., diffusion) on the internal validity of the woman defendant

condition. All participants were recruited via Kijiji.com (an online classified advertising service) and were paid \$40.00 for their participation in the mock jury deliberation.

For each of the 10 mock juries (five mock juries for the man defendant; five mock juries for the woman defendant), I recruited 12 participants. However, due to no-shows, there were fewer than 12 participants in each of the mock juries. There are certainly downsides to having juries with fewer than 12 jurors. Small juries tend to be less diverse in perspectives and attitudes about the case than larger juries (Bornstein & Greene, 2017). Sommers (2006) found that diverse juries (e.g., jurors of different races and ethnicities vs. all-White juries) tend to consider more case facts and evidence and are more likely to recall the evidence more accurately. Research on jury size also has demonstrated that 12-person juries (in comparison to 6-person juries) spend slightly more time deliberating (Brunell, Dave, & Morgan, 2009; Saks & Marti, 1997).

Nevertheless, larger juries (which are more likely to have heterogeneous perspectives) have greater difficulty coming to a consensus on a verdict and are thus more likely to be “hung” than smaller (and more homogenous juries) in criminal trials (unweighted mean is $Z = 3.23$, $p = .0006$, $d = .107$, weighted mean is $Z = 2.92$, $p = .0018$, $d = .071$; Saks & Marti, 1997). Importantly, researchers have found that jury size does not have a significant biasing effect on jury verdict (i.e., larger juries are not more or less likely to be lenient) (unweighted mean is $Z = -.404$, $p = .655$, $r = -.158$, weighted mean is $Z = .643$, $p = .261$, $r = .040$; Saks & Marti, 1997).

Accordingly, due to the experimental and funding requirements of the present research, I proceeded with the mock jury deliberations when there were fewer than 12 mock jurors present (minimum jury size included 5 mock jurors). The number of jurors varied per jury ($M = 7.70$, $SD = 2.31$). There were fewer participants who deliberated the woman defendant’s case ($n = 37$)

than the man defendant's case ($n = 40$); however, these sample sizes were not significantly different, $t(75) = -1.32, p = .19$.

The average age of the mock jurors (see Table 1) did not significantly differ by defendant gender condition, $t(75) = .59, p = .56$. The mock jurors who deliberated the woman defendant's case were on average 28.54 years old ($SD = 9.859$ years). The mock jurors who deliberated the man defendant's case were on average 30.98 years old ($SD = 13.59$ years).

The gender composition of the juries in the two conditions were not significantly different, $\chi^2(1) = .352, p = .553$. On average the juries deliberating the woman defendant's case comprised 56.80% women and 43.20% men. On average the juries deliberating the man defendant's case comprised 50.00% women and 50.00% men.

The race composition of the juries in the two conditions was not significantly different, $\chi^2(7) = 9.02, p = .251$. The race makeup of the mock juries in the woman defendant condition was: 56.80% White, 10.80% "Other" (e.g., Mixed), 8.10% Black, 8.10% Indigenous, 5.40% Asian, 5.40% Hispanic/Latinx, and 5.40% Middle Eastern. The race makeup of the mock juries in the man defendant condition was: 72.50% White, 12.50% Black, 7.50% "Other" (e.g., Mixed), 5.00% Indigenous, and 2.50% South Asian.

Table 1 *Demographics of jurors*

Jury Code	Number of Jurors	Defendant Gender	Group Verdict	Mean Juror Age (years)	Juror Gender	Juror Race
C	9	Man	Hung	$M = 40.44,$ $SD = 16.78$	Man = 44.40% Woman = 55.60%	Black = 11.10% White = 77.80% Indigenous = 11.10%
F	6	Man	NCRMD	$M = 30.00,$ $SD = 13.97$	Man = 50.00% Woman = 50.00%	White = 83.30% Other = 16.70%
G	6	Man	Hung	$M = 23.83,$ $SD = 7.20$	Man = 66.70% Woman = 33.30%	Black = 50.00% White = 50.00%
H	12	Man	Hung	$M = 22.67,$ $SD = 4.59$	Man = 41.70% Woman = 58.30%	Black = 8.30% South Asian = 8.30% White = 75.00% Other = 8.30%
J	7	Man	Guilty	$M = 40.00$ $SD = 12.65$	Man = 57.10% Woman = 42.90%	White = 71.40% Indigenous = 14.30% Other = 14.30%

K	10	Woman	Hung	$M = 26.60,$ $SD = 6.80$	Man = 40.00% Woman = 60.00%	Black = 10.00% Middle Eastern = 10.00% Hispanic/Latinx = 10.00% White = 60.00% Indigenous = 10.00%
L	9	Woman	Hung	$M = 33.33$ $SD = 15.41$	Man = 33.30% Woman = 66.70%	Asian = 11.10% Middle Eastern = 11.10% White = 66.70% Other = 11.10%
M	5	Woman	Guilty	$M = 24.25,$ $SD = 2.06$	Man = 60.00% Woman = 40.00%	White = 60.00% Indigenous = 20.00% Other = 20.00%
N	5	Woman	Hung	$M = 28.40,$ $SD = 11.13$	Man = 40.00% Woman = 60.00%	Asian = 20.00% Black = 20.00% White = 40.00% Other = 20.00%
O	8	Woman	Guilty	$M = 27.71,$ $SD = 5.12$	Man = 50.00% Woman = 50.00%	Black = 12.50% Hispanic/Latinx = 12.50% White = 50.00% Indigenous = 12.50% Other = 12.50%

Materials

Recruitment advertisement. The recruitment took place online via Kijiji.com (an online classified advertising service). Participants used their own computers to find the study advertisement, which was categorized under “jobs, general work” on Kijiji.com and was titled “You be the Jury”. The advertisement included a brief description of the study’s tasks (see Appendix A).

Questionnaire about qualifications for jury service. This brief questionnaire included questions to screen participants for qualification to participate in the study’s deliberation. Participants were asked to select “Yes” or “No” for the following questions: (e.g., “*Are you a Canadian citizen? Are you 18 years of age or older? Have you ever been convicted of an indictable offence?*”) (Appendix D). These qualification requirements were based on the questionnaire about qualifications for jury service of Ontario (*Juries Act*, R.S.O 1990).

Pre-trial instructions. These instructions were adapted from the Model Jury Instructions (CJC-NJI, 2011, A. 5.1) for the essential elements of the charges against the defendant, the requirements for NCRMD, and the burden of proof (see Appendix F).

Trial transcript. The trial transcript (see Appendix G) described the case against the defendant (Gretchen Miller or Gary Miller) who was charged for second-degree murder in the killing of their roommate (Dennis Hughes). The 9-page trial transcript was modelled after a previously used transcript (Yamamoto, 2019). The trial transcript included: the Crown’s and the defence’s opening statements; the Crown’s witness, police officer Mark Hanes, who arrested the defendant at the defendant’s mother’s home and who reported the behaviours of the defendant upon arrest; the defence’s witness, psychiatrist Dr. Devin Cassady, who examined the

defendant's mental state after the act, and who testified that the defendant likely has paranoid schizophrenia and was likely experiencing Capgras Delusion at the time of the act. The psychiatrist explained that Capgras delusions may occur in patients with paranoid schizophrenia and that a key feature of this delusion is the belief that family or friends have been replaced by identical looking imposters.

The transcript also included the cross-examinations of these witnesses by the Crown and/or defence. Lastly the transcript included the Crown's closing statement (i.e., the defendant is not mentally ill, rather they are criminal) and the defence's closing statement (i.e., the defendant was not in control of their actions and should not be held accountable or penalized). Notably, all case facts and testimony remained consistent across the woman and man defendant's trials. The defendant's names (Gretchen Miller or Gary Miller) and pronouns were varied so as to manipulate defendant gender.

Individual verdict. This form requested the mock jurors' individual verdicts for the defendant (i.e., guilty, not guilty², NCRMD) (see Appendix H).

Pre-deliberation instructions. These instructions included information on the legal definition of NCRMD (Criminal Code, 1985) and information on which questions mock jurors should consider when determining the criminal responsibility of the defendant (CJC-NJI, 2011, E.16.1, s.16) (see Appendix I). Lastly, the pre-deliberation instructions, based off of the Model Jury Instructions from the National Judicial Institute (CJC-NJI, 2011, C. 5.3, s.105 – s.108),

² It should be noted that a verdict of not guilty would indicate jurors' misunderstanding of the law (or a potential attempt at jury nullification), as jurors were instructed to either find the defendant guilty or NCRMD.

included directions on the logistics and goal of the deliberation (see Appendix H). Jurors also were asked to provide their group verdict in writing (see Appendix J).

Procedure

Participants from the general public were recruited from the Greater Capital Region of Canada via Kijiji.com (an online classified advertising service). The recruitment page advertised that for their participation in the complete deliberation study titled “Jury Deliberation and the Not Criminally Responsible on Account of Mental Disorder Defence,” each participant would be compensated \$40.00 for their time (Appendix A). From the advertisement, jurors were redirected to the informed consent form by a query string (Qualtrics’ algorithms for sending information between surveys in their survey development suite) (Appendix B). Jurors were then asked jury qualification questions (Appendix D). After consenting and qualifying to participate, participants were asked to provide their e-mail address. The query string used each e-mail address to automatically generate and send each participant a unique and randomized mock juror code number and the directions to the deliberation room and the date of the mock deliberation.

When participants arrived at the deliberation lab, they were each assigned a seat number and were asked to communicate with each other using these labels for the sake of confidentiality; these numbers were on the walls behind the jurors’ seats (see Figure 1). Thereafter, I read the welcome script (Appendix E) to participants describing their general role as mock jurors. Participants were then asked to read and sign an informed consent form (Appendix C), which further described the study. I then read the pre-trial instructions (Appendix F) and asked the participants to individually read the trial transcript (Appendix G) and provide an individual and confidential verdict (Appendix H).

Thereafter, I read the pre-deliberation model jury instructions (Appendix I) to the participants to inform them of their role and tasks as mock jurors. Specifically, I explained the elements of the insanity defence and the legal standard for finding a defendant NCRMD. Lastly, I asked the jury to provide a verdict in writing within the 45 minutes allotted (i.e., guilty, not guilty, NCRMD, or could not reach a verdict); I also explained that if the mock jurors could not reach a unanimous verdict that the jury would be declared “hung” (Appendix J). Participants then deliberated the case as a group for up to 45 minutes (deliberations were audio and video recorded). The time allotted to jurors to deliberate in jury experiments has varied greatly (from 7 minutes to an unlimited amount of time; Devine, Clayton, Dunford, Seying, & Pryce, 2001). The time limitation in the present study was selected based on the time that was allotted in Yamamoto’s (2019) study of jury decision-making. Yamamoto (2019) selected 45 minutes based on Sommers’ (2006) report describing the average deliberation time (in sexual assault trials) as ranging from 38 to 50 minutes. Moreover, given practical (e.g., transcription and coding time) and funding limitations for this project, 45 minutes was deemed to be a fair deliberation time. Prior to leaving the lab, participants were debriefed, and each participant was compensated with \$40.00 and thanked for their time.



Figure 1 Legal decision-making lab deliberation room

Apparatus. The jury deliberation lab has a one-way mirror, through which I observed the mock deliberations. The microphones (in the ceiling) and surveillance camera (on the walls on either side of the room) recorded the deliberations for later transcription (see Figure 1).

Analysis Overview

For the current study, mock jurors were asked to deliberate one of two trials (woman defendant or man defendant) and provide individual and group verdicts. Mock jury deliberations were audio and video recorded for later transcription; these transcripts were analyzed for themes of mock jurors' perceptions of the defendant's warmth or competence (SCM; Fiske, Cuddy, Glick & Xu, 2002; Wojciszke, 2005). Study 1a and Study 1b were aimed at describing if and how mock jurors' stereotypes and emotions towards an insanity defendant varies as a function of the defendant's gender; Study 1c was aimed at examining if the stereotypes and emotions expressed in deliberation are significantly correlated; Study 1d was aimed at describing how

women and men mock jurors' speech acts (e.g., assertive or tentative) differ. Study 2a and Study 2b were aimed at examining how the gender of the defendant as well as of the juror influences each juror's individual verdict decisions. Studies 3, 4, and 5 were exploratory studies and analyzed the content and themes of jurors' discussions.

First, the transcripts were analyzed for juror expressions of the defendant's "warmth" (e.g., the defendant was *helpful*, *likeable*, and/or *supportive*) and expressions of the defendant's "competence" (e.g., the defendant was *efficient*, *active*, *skillful*) (see Table 8). There are also several reverse-coded items that imply "low" competence (e.g., *stupid*) and "low" warmth (e.g., *callous*), which raters may code as pertaining to either warmth or competence stereotypes. Each of the mock jury deliberations were combed for any and all expressions of the defendant's warmth and competence (as defined by the Stereotype Content Dictionaries for high warmth, low warmth, high competence, low competence words) by using the Linguistic Inquiry Word Count program (LIWC; see Pennebaker, Francis, & Booth, 2001). LIWC is a software program that has two primary features: the processing component and the dictionaries. The LIWC program can open a series of texts and compare the texts word per word to the dictionary file (Tausczik & Pennebaker, 2010). For example, the program could read the sentence "The dog was happy" and then compare "the", "dog", "was", and "happy" to its dictionaries. If a word is in the 11,900-word dictionary, the word is coded as an affect word (e.g., relating to anxiety, anger, positive emotion), a social word (e.g., relating to family, friend), a core drive and needs word (e.g., relating to affiliation, achievement, power), etc. The LIWC program also allows researchers to add their own dictionaries (i.e., the SCM dictionaries). After the body of text (e.g., deliberation transcripts) are compared to LIWCs dictionaries, the program provided a metric for use of stereotype content words to all the words used in the deliberations. The higher the metric number

is (i.e., match between the dictionary words and transcript), the more relevant a stereotype content category (e.g., high warmth, low warmth, high competence, low competence) was to the deliberation.

I also used LIWC to examine mock jurors' affect towards the defendant (based on the language they used). That is, there is an internal dictionary in LIWC that categorizes specific words as indicating positive affect and negative affect. LIWC thus highlights and provides a metric of the overall affect in a body of text (e.g., proportion of text that demonstrates positive or negative affect). According to Fiske and colleagues (2002), the stereotypes perceivers associate with a target significantly relate to their emotional and behavioural responses towards the stereotype target. For example, if a defendant is perceived as not competent and not warm, mock jurors may be inclined to feel negative affect towards the defendant (Fiske et al., 2002). I also examined and coded the transcripts for mock jurors' speech acts (e.g., assertive, tentative).

For the present study, I conducted a few comparison tests. First, I examined if there is a significant difference between the warmth and competence words (DV) used for women defendants and for men defendants (IV). That is, I compared the number of stereotype content words (both warmth and competence) that appear in the woman and the man defendants' deliberations. This was to determine if there is a significant difference in the expression of these stereotypes per defendant gender. Second, I compared the positive and negative affect expressed in deliberations (DV) for the woman defendant and for the man defendant (IV). Third, I compared how men and women jurors find the man and woman defendant (e.g., juror and defendant influence on verdict decisions). Lastly, I compared the speech acts (e.g., tentative, assertive statements) of women and men jurors in deliberations.

Data Preparation

Content analysis is a method of analyzing written, audio, and video data (Cole, 1988). Specifically, this method of analysis involves the “objective, systematic and quantitative description of manifest content of communications” (Berelson, 1952, p. 18; Krippendorff, 2004). Manifest content (Kondracki, Wellman, & Amundson, 2002) are the descriptive data of a conversation (e.g., specific words used in conversation rather than the researcher’s interpretation of the words’ meanings). Content analysis is a systematic process of quantifying text and qualitative information (e.g., transcribed speech; Berelson, 1952; Krippendorff, 2004). Accordingly, each of the mock jury deliberations has been audio and video recorded and transcribed by independent research assistants who were blind to my hypotheses; all of mock jurors’ verbalizations are transcribed verbatim and literally, and mock jurors’ pauses and audible behaviours (e.g., sighing) are also transcribed (see Zhang & Wildemuth, 2009).

Broadly, there are two approaches to analyzing the content of conversations: inductive (theory is developed from the data’s themes) and deductive (theory directs the categories of discussion that are explored in the data) (Elo & Kyngäs, 2008). Inductive content analysis is useful when the researchers do not have concrete hypotheses (e.g., open-coding) or when there is limited literature on the topic of study (Elo & Kyngäs, 2008). Conversely, the deductive content analysis research strategy uses a theory-driven approach that guides the development of a coding manual (Elo & Kyngäs, 2008). The aim of deductive content analysis is to condense an entire discussion (e.g., jury deliberation in insanity trials) into pre-determined, theory-based categories (e.g., stereotype content) of information (e.g., jurors’ stereotypes and attitudes towards women defendants; see Elo & Kyngäs, 2008).

For the present study, I conducted a both a deductive and an exploratory inductive content analysis (as well as a thematic analysis) and analyzed the manifest content of the mock

jury deliberations. That is, I analyzed the words that mock jurors express when deliberating the man or the woman defendant (e.g., stereotypes), based on the theory that outlines my project and hypotheses. Based on gender stereotypes and attitudes about people with mental illness, I expected to find that mock jurors would express notably different stereotype themes (e.g., warmth, competence stereotypes), emotional expressions, and speech acts in the jury deliberations as a function of defendant gender and juror gender, respectively.

Deliberation segmentation. The unit of analysis for this deliberation study was each individual sentence jurors speak. This segmentation permitted the examination of the transcripts for the presence (or absence) of codes of warmth stereotypes, competence stereotypes, positive/negative affect, tentative/assertive language (see Chi, 1997). Accordingly, prior to analyzing the data, I segmented the deliberations to the level of the jurors' utterances (sentences). That is, each sentence each juror speaks was considered an utterance and was analyzed for stereotype content, affect, tentative language, and assertive language. Each utterance was nested within the speaker (i.e., labelled as J1, J2..., J12) and within the jury in which the juror spoke (i.e., label the utterance as occurring in Jury A, Jury B, etc.) (see Hamilton & Hunter, 1985).

Stereotype Content Domain dictionary. I have developed an exhaustive list of stereotype content traits (123 trait words; see Table 9), which I refer to as the Stereotype Content Dictionary (SCD). The development of the SCD involved generating as many warmth and competence adjectives, synonyms, and antonyms as possible by using dictionaries and thesauruses. This list also comprises stereotype traits from a few different papers on stereotype content theory and stereotype traits (Bergsieker, Leslie, Constantine, & Fiske, 2012; Fiske, Cuddy, Glick, & Xu, 2002; Zanna & Hamilton, 1972). The validity of these warmth and competence adjectives were determined by three independent raters.

The raters read this list of words and categorized each as a trait related to high/low warmth, to high/low competence, or to neither stereotype. There were no “filler” words in the list of stereotype adjectives, though I explained to coders that they can select the “neither” option in lieu of “warmth” or “competence” if the adjective does not apply to either of the stereotype content. After the raters had independently coded the list of stereotype traits, I reviewed their codings and assessed the inter-rater agreement on the codings. Inter-rater reliability (Intra class-correlation (ICC) = .93) indicated that there was a strong consistency in raters’ evaluations of warmth or competence words (see McHugh, 2012). There were four words (i.e., conventional, evasive, impulsive, materialistic) with low agreement; as such, these controversial trait words were re-evaluated and removed from the SCD (see Zhang & Wildmuth, 2009). Afterwards, I combed the deliberation trial transcripts to examine if mock jurors described the defendant using any of the warmth- or competence-specific words. The trial transcripts were first scanned and coded using the Linguistic Inquiry Word Count (LIWC; Pennebaker, Francis, & Booth, 2001). This program provided a metric of stereotype content use in the deliberation transcripts and highlight where the words appear in the deliberations.

I then read the stereotypes in context to ensure that the most words are indeed relevant to the warmth or competence categories (see Graham, Haidt, & Nosek, 2009). For example, I conducted a content analysis to ensure that words such as “crazy” do in fact occur in the context of describing the defendant (e.g., “she is clearly *crazy*”) rather than trial unrelated information (e.g., “the weather we received has been *crazy* nice”). In this example, the word *crazy* may contribute to the competence stereotype count in both contexts, but one of these is not relevant to the deliberation and thus should not contribute to the overall stereotype content counts. I used a similar method as Graham and colleagues (2009) to ensure the “contextual validated usage” of

my SCD words. I selected the 2-3 sentences that surround each of the usages of the stereotype words and scored the passage with 1 if it supported or was relevant to warmth or competence stereotype, and a score of 0 if the contextual usage of the stereotype word was unclear or irrelevant to the stereotype content.³

Results

Study 1: Group Analyses

Study 1 concerned the examination of mock jurors' attitudes and emotions towards the men and women defendants pleading NCRMD. To this end, I examined mock jurors' use of stereotype content and affective language. Each of the deliberation utterances were segmented into individual sentences and the segmented transcript were codified as containing (or not containing) stereotype content words (based off the SCD) by using LIWC (Pennebaker, Francis, & Booth, 2001). These codes of (e.g., warmth present/warmth not present) stereotype content were then analyzed using a multi-level linear model (MLM) using SPSS. This model nested the individual utterances (micro-level variable of interest) within the juror and within the individual juries (macro-level variable of interest) to form a 2-level MLM aimed at examining if the stereotype content and affect significantly varies between the juries for a woman defendant and juries for a man defendant (see Table 2 for descriptives).

³ There was very limited (and often no) mention of SCM dictionary words or affect language across deliberations. As such, all analyses using these variables had minor effects across deliberations.

Table 2 Descriptives of Stereotype Content and affective word use

		Average Stereotype Content and Affective words expressed											
		High warmth		Low warmth		High competence		Low competence		Tentative speech		Assertive Speech	
Defendant Gender	Juror Gender	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Woman	Women	0.12	0.25	0.04	0.10	0.03	0.06	0.007	0.03	2.73	1.17	0.94	0.58
	Men	0.13	0.17	0.04	0.11	0.01	0.03	0.006	0.03	2.64	0.89	1.13	0.67
Man	Women	0.19	0.42	0.02	0.03	0.04	0.08	0.00	0.00	3.14	2.31	0.95	0.74
	Men	0.104	0.18	0.007	0.03	0.05	0.09	0.005	0.01	2.59	1.94	0.76	0.50

Study 1a(i): Defendant gender and jurors' stereotype content expressions. The random intercept effect of high warmth stereotype language and jury was non-significant, $u_{0j} = 0.0036$, $Z = 0.55$, $p = .58$. Moreover, jurors were not significantly more likely to express high warmth stereotypes for a woman defendant ($M = 0.12$, $SD = 0.21$) than for a man defendant ($M = 0.14$, $SD = 0.32$), $t(8.13) = 0.010$, $p = .922$, Power = 0.09, Cohen's $d = 0.07$ (nil effect size).⁴ ⁵

The random intercept effect of high competence language and jury was non-significant, $u_{0j} = 0.0010$, $Z = 1.38$, $p = .169$. Moreover, jurors were not significantly more likely to express high competence stereotypes for a man defendant ($M = 0.04$, $SD = 0.08$) than for a woman defendant ($M = 0.02$, $SD = 0.05$), $t(8.59) = 0.850$, $p = .42$, Power = 0.35, Cohen's $d = 0.29$ (small effect size).

Study 1a(ii): Defendant gender and jurors' affect expressions. The random intercept effect of negative emotion language and jury was non-significant, $u_{0j} = 0.067$, $Z = 0.76$, $p = .45$. Moreover, jurors were not significantly more likely to express more negative emotion language for a man defendant ($M = 1.36$, $SD = 0.89$) than for a woman ($M = 1.67$, $SD = 0.85$), $t(6.92) = -1.22$, $p = .26$, Power = 0.36, Cohen's $d = 0.36$ (small effect size).

The random intercept effect of positive emotion language use and jury was non-significant, $u_{0j} = 0.000$. Moreover, jurors were not significantly more likely to express more positive emotion language for a woman defendant ($M = 1.24$, $SD = 0.61$) than for a man

⁴ Cohen's d was calculated as $M_{\text{man defendant}} - M_{\text{woman defendant}} / SD_{\text{pooled}}$

⁵ The magnitude of Cohen's d effect sizes were interpreted as "nil" when they were between 0.0-0.2, "small" when between 0.2-0.5, "medium" when between 0.5-0.8, and "large" when > 0.8 (McLeod, 2019).

defendant ($M = 1.54$, $SD = 1.26$), $t(75) = 1.29$, $p = .20$, Power = 0.37. Cohen's $d = 0.30$ (small effect size).

Study 1a(iii): Jurors' stereotype content and affect expressions. The context in which stereotype content traits are expressed are essential to determining the valence of these trait expressions (e.g., negative or positive affect associated with stereotype expression) (see Barrett, Mesquita, & Gendron, 2011). That is, the context of these stereotype expressions allows the linking of the jurors' expressions to the relevant affective and behavioural outcomes (see Mackie, Devos, & Smith, 2000). I examined the ICC values for these variables of interest and found that the $ICC < .1$, meaning that it is not necessary to conduct a nested model for this study (see Vajargah & Nikbakht, 2015). I used LIWC to code the segmented utterances of the deliberation transcripts (i.e., *target words*) to those of its internal dictionaries (e.g., affective processes dictionary: positive emotion, negative emotion) to provide general metrics on the emotional tones of the utterances in the man and woman defendant's deliberations. The LIWC program allows the researcher to do a text analysis that efficiently and effectively highlights the cognitive and emotional components of mock jurors' speech (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007). I used LIWC to highlight the expression of stereotype content words throughout each of the juror's (overall) utterances and provide a metric of the juror's overall use of high warmth, low warmth, high competence, and low competence stereotype words. I also used LIWC to quantify the individual juror's overall expression of negative or positive affect. I then used these scores as data points to conduct a correlation analysis in SPSS.

There was not a significant relationship between the use of high warmth stereotypes and expressions of positive affect, $r(75) = -0.13$, $p = 0.26$. There was not a significant relationship between the use of low warmth stereotypes and expressions of negative affect, $r(75) = 0.11$, $p =$

0.32. There was not a significant relationship between the use of high competence stereotypes and expressions of negative affect, $r(75) = -0.02, p = 0.90$. There was not a significant relationship between the use of low competence stereotypes and expressions of positive affect, $r(75) = 0.05, p = 0.69$.

Study 1b: Juror gender and speech style. I segmented each of the deliberation utterances into individual sentences and the segmented transcripts were codified as containing (or not containing) tentative speech words or assertive speech words (using the internal dictionary of LIWC); these codes were then analyzed using a multi-level linear model (MLM) using SPSS software. This model nested the individual utterances (micro-level variable of interest) within the juror and also nested the utterance within the juries (macro-level variable of interest) to form a 2-level MLM aimed at examining if the use of tentative or assertive speech style significantly varies between the woman and man jurors.

The random intercept effect of tentative language use and jury was non-significant, $u_{0j} = 0.69, Z = 1.53, p = .13$. Moreover, women jurors ($M = 2.92, SD = 1.80$) were not significantly more likely to express more tentative language than were men jurors ($M = 2.61, SD = 1.54$), $t(67.94) = 0.49, p = .49, \text{Power} = 0.13, \text{Cohen's } d = 0.19$ (small effect size).

The random intercept effect of assertive language use and jury was non-significant, $u_{0j} = 0.097, Z = 1.41, p = 0.16$. Moreover, men jurors ($M = 0.92, SD = 0.61$) were not significantly more likely to express assertive language than were woman jurors ($M = 0.94, SD = 0.65$), $t(67.06) = 0.058, p = .95, \text{Power} = 0.07, \text{Cohen's } d = 0.03$ (nil effect size).

Table 3 *Group analyses outcomes*

Group outcome variable	<i>t</i>	<i>p</i>	<i>d</i>
High warmth	0.010	.922	0.07
High competence	0.850	.42	0.29
Negative emotion	-1.22	.26	0.36
Positive emotion	1.29	.20	0.30
Individual variables	<i>t</i>	<i>p</i>	<i>d</i>
Tentative language	0.49	.49	0.19
Assertive language	0.058	.95	0.03

Discussion

There was not a significant impact of defendant gender on the stereotypes and affect jurors expressed in deliberations. This is not to say that jurors did not rely on stereotypes regarding their and the defendant's gender when making their decisions in insanity trials, rather that stereotypes about gender were not explicitly expressed when deliberating. Specifically, jurors may not verbally express Stereotype Content and affect language when deliberating men and women defendants who plead NCRMD. People are often reluctant to explicitly express prejudices in public and may only reveal their sexism in "safe" circumstances where there is potential that their bias can be attributed to something other than gender; people typically mask their sexist attitudes due to their social desirability concerns (see Swim & Cohen, 1997). As such, jurors who want to appear egalitarian despite their sexist attitudes may tend to use coded language or phrases to express stereotypes and attitudes towards women and men defendants rather than explicitly referring to the woman defendant with sexist stereotypes. Alternatively, jurors may feel ambivalent sexism (Glick & Fiske, 1997) towards women defendants (mixed

feelings) and thus their expression of sexist stereotypes may be more covert and imperceptible to a quantitative analysis using a stereotype taxonomy.

Moreover, defendant and juror gender did not significantly influence the emotions jurors explicitly expressed (via language). Again, it is very likely that jurors felt an assortment of emotions while deliberating, as emotions are relevant to all actions and discussions (Barbalet, 2002). Nevertheless, jurors' emotion expressions did not significantly vary for the man and woman defendant cases. This may be the case because jurors' foci and emotions in deliberations were fairly consistent across deliberations (e.g., mental illness faking), and the gender of the defendant may have had limited impact on jurors' perceptions of the defendant's NCRMD plea. It may have also been the case that the affective word corpus (built into LIWC) did not encompass language that would be relevant to group member emotions expressions in jury deliberations. Alternatively, jury deliberation may involve repressing emotional expressions due to being in a traditionally rational legal environment in which jurors are unwilling or restricted in their verbal emotional expressions (see Kelly & Barsade, 2001).

Lastly, there was not a significant difference in the tentative and assertive speech styles of men and women jurors. This finding is similar to the current literature on gender and speech. Leaper and Robnett (2011) conducted a meta-analysis testing for gender differences and moderators in tentative speech and found that women are slightly more likely to use tentative speech. Moreover, Leaper and Ayres (2007) found that men are marginally more likely to use assertive speech. The non-significant findings in the present study may also be due to there being very limited overall instances of tentative- and assertive-coded language in the deliberations.

Implications

The Stereotype Content Model (SCM; Fiske et al., 2002) may not be applicable to the NCRMD context and may also have limited utility in open ended discussion contexts. Jurors may be generally unlikely to use adjectives related to gender to describe the defendant. Jurors may also be hostile and unwilling to consider finding the defendant NCRMD regardless of the defendant's gender. These results also call into question the use of Language Inquiry Word Count tool (LIWC; Pennebaker, Francis, & Booth, 2001) for analysis of group deliberations. The utility of this tool in jury deliberations may be limited due to the idiosyncratic style of speech and short quotes of individual jurors.

Limitations

The proposed SCM had limited utility in interpreting the deliberations, since jurors did not tend to express exact SCM adjectives (this may be because the dictionary was not extensive enough, or because SCM is typically a survey measure). Importantly, Cuddy, Fiske, and Glick (2008) cautioned that while equating experimenter-chosen trait lists (e.g., SCM dictionary) with "person perception" may be empirically sound, there are issues with ecological validity in doing so. Specifically, such approaches may not mirror how perceivers spontaneously use SCM traits (i.e., warmth and competence and their related emotions). The SCM framework has typically studied participants' perceived warmth and competence of different people and social groups by asking participants to list "important social groups", and to then classify these social groups using a list of researcher-created SCM traits (Cuddy et al., 2008). Conversely, I measured the spontaneous occurrences of these stereotype adjectives in discussions; while this process was more ecologically valid, it had limited empirical utility.

Because there were so few instances of SCM and affective language in deliberations, I was not able to draw clear inferential relationships between stereotypes, emotions, and

discrimination in NCRMD trials. It may be the case that because this information occurred so infrequently across all juries that it appears that jurors' attitudes towards men and women who plead NCRMD are the same. Moreover, the overall sample size of the present study may have been too small to detect differences in jurors' stereotypes and affects. I posit that gender and sexism may still be related to jury decision-making, just less so in the context of insanity trials. Future studies should also include an analysis of the facial and body expressions of emotions of jurors during deliberation in addition to their verbal discussion (see Mondada, 2016).

Study 2: Individual Analyses

Study 2 concerned the examination of mock jurors' verdict decisions as a function of jurors' and defendants' genders. To examine this, I conducted Chi-square analyses that examined mock jurors' individual verdicts (and juror and defendant genders).

Study 2a: Juror and individual verdict. A chi-square test for independence was performed to examine the relationship between juror gender and verdict decisions (see Table 4). The relationship between these variables was non-significant, $\chi^2(2, N = 75) = 0.490, p = .484, V = 0.081, p = .484$. An analysis of the odds ratio of finding the defendant guilty revealed that women jurors were 1.389 times more likely to find the defendant guilty than a man juror was.⁶ However, considering the small sample size, this finding is likely negligible.

⁶ There was one participant who responded "not guilty" and one participant who did not provide an individual verdict, as such their responses were omitted from the analyses here; their responses indicated a misunderstanding of the law and jury instructions.

Table 4 *Juror gender and verdicts*

Juror gender	Verdict		
	NCRMD	Guilty	Total
Man	16 (47.10%)	18 (52.90%)	34 (100.00%)
Woman	16 (39.00%)	25 (61.00%)	41 (100.00%)

Study 2b: Juror gender, defendant gender, and individual verdicts. A power analysis for this model demonstrated low power (0.15). A loglinear analysis was conducted to examine the relationship of defendant gender and juror gender on jurors' verdict decisions (see Table 4). The three-way loglinear analysis did not produce a significant model. The likelihood ratio of this model was $\chi^2(7) = 5.093$, $p = .649$, meaning that the model is not a good fit and there is not a significant relationship between defendant gender, juror gender, and the verdict jurors provide, nor is there a significant main effect of juror gender or defendant gender on individual verdicts.

Table 5 *Juror and defendant gender and verdict*

Defendant gender	Juror gender	Verdict		
		NCRMD	Guilty	Total
Woman	Woman	8 (38.10%)	13 (61.90%)	21 (100.00%)
	Man	9 (60.00%)	6 (40.00%)	15 (100.00%)
Man	Woman	8 (40.00%)	12 (60.00%)	20 (100.00%)
	Man	7 (36.50%)	12 (63.20%)	19 (100.00%)

Discussion

The present study findings were all null, and thus may not mirror those of past research on the impact of juror gender in insanity trials. For example, Breheny and colleagues (2007) found that women jurors were more likely to accept that (attribute less responsibility) the defendant's mental illness impacted their actions than were men jurors. The present study's lack of a significant effect of defendant and juror gender on jurors' verdicts may fit in with the

contradictory and limited literature on legal decision-making in insanity trials. For example, while some researchers found that women defendants are significantly more likely than men to be found guilty when pleading insanity for the same charge (Breheney et al., 2007) other researchers found no such effect (Mossière & Maeder, 2016). Accordingly, the present study's findings of a null effect of defendant gender on verdict may be consistent with insanity jury decision-making literature in the Canadian context.

Implications

The present study demonstrates the difficulty of quantifying qualitative data. While I did not find a direct relationship between defendant gender (nor juror gender) and verdict in this study, there may be variables beyond those I proposed to examine that are affecting jurors' decision-making processes and verdict. That is, a qualitative analysis of deliberations may demonstrate that men and women jurors differently reach their verdicts for men and women defendants (see Study 5).

Limitations

The individual juror and verdict sample size was modest ($N = 77$) and therefore, the study was underpowered. As such, I am not able to make assertions on the relationship between juror and defendant gender and verdict outcomes.

Study 3: Quantitative Content Analysis and Results

Study 3 concerned the examination of mock jurors' deliberation behaviours (e.g., turn-taking and leadership roles). To this end, I examined and coded the deliberation transcripts for mock jurors' interruptions of other jurors, their talkativeness, language type (i.e., assertive, affiliative), and jurors' perceptions of the deliberations' influencers and leaders. The transcripts were read at least two times to get a "sense" of what jurors were discussing during their

deliberations. That is, I was interested in exploring the case facts and evidence that were most commonly discussed, jurors' primary concerns with the cases, and how jurors interacted with each other via speech acts (e.g., interruptions, sentence tags) during deliberation. After reading the transcripts several times and becoming familiar with the data, I segmented the transcripts into juror utterances (i.e., individual sentences). At any point where there was a period, dash, or ellipses, I labelled the sentence prior to it as an utterance. As such, each sentence in the transcript became an utterance and was coded as being uttered by either a man or woman juror (the gender of each juror was recorded from jurors' self-reported gender identification in the demographics form). I selected this method of segmentation as it was the most consistent and parsimonious manner to organize jurors' utterances⁷. Of note, the ICC values of each of the following research questions have been examined and are all below the .1 cut off value for analysing data using a nested model (Vajargah & Nikbaknt, 2015). As such I did not conduct nested examination of the following research questions.

Speaking turns/talkativeness

In a meta-analysis of gender differences in talkativeness, Leaper and Ayres (2007) found that men were significantly more likely than women to be talkative, though this is a small effect size ($d = -0.14$). In the present study, I segmented the data in each of the deliberations, and labelled each utterance as having been spoken by a man juror or a woman juror. Since each

⁷ To the best of my knowledge, there is not a consistent manner in which jury deliberation researchers have parsed or segmented qualitative data for quantitative analysis. I utilized punctuation to demarcate units of analysis, as this was the most parsimonious and simple method to consistently apply across my data set and was roughly similar to the method utilized by Green, Hayman, and Motyl (2008). While I am aware that there are a number of methods to grammatically parse and organize large bodies of text (Bengtsson, 2016; see Watkins, 2017), tokenizing language and lexical analysis (as in computer science), this is beyond the scope of my study.

utterance was generally similar in length, this metric of speech time could be used to measure the overall speaking time (turns, or talkativeness) of men and women jurors. Across the 10 deliberations, men spoke slightly more than women jurors. Of the overall 7010 speaking turns across the 10 deliberations, women jurors spoke 47.67% of the speaking turns ($M=81.51$, $SD = 75.92$) and men jurors spoke 52.33% of the speaking turns ($M = 101.92$, $SD = 101.13$), though this difference in speaking turns was not statistically significant, $t(75) = -1.01$, $p = .32$, Power = 0.15, Cohen's $d = 0.22$ (small effect size) (see Table 5).

Moreover, in the man defendant cases, men jurors spoke 54.24% of speaking turns ($M = 84.50$, $SD = 83.52$), whereas women jurors spoke 45.76% of speaking turns ($M = 71.30$, $SD = 83.99$). This difference in speaking turns was not significant, $t(38) = -0.50$, $p = .62$, Power = 0.07, Cohen's $d = 0.16$ (small effect size). In the woman defendant cases, men and women spoke at more equal rates, where men jurors spoke 50.80% of speaking turns ($M=123.69$, $SD = 118.84$) and women jurors spoke 49.20% of speaking turns ($M = 91.24$, $SD = 67.98$). This difference was not a statistically significant one, $t(35) = -1.05$, $p = .30$, Power = 0.17, Cohen's $d = 0.34$ (small effect size).⁸ These null findings are contrary to the meta-analytic findings of Leaper and Ayres (2007), who found that across 70 independent samples with a total N of 4,385 participants, men were more talkative (though small effect size) than women ($d = -.14$, $p < .01$). When examining potential mediators of speaker gender and talkativeness, Leaper and Ayres (2007) also found that there was only a negligible relationship ($d = -0.13$, $p < .05$) between speaker gender and talkativeness in studies in which participants were in a deliberation role.

⁸ Of note, the jurors spoke more in the woman defendant condition than in the man defendant condition. This effect is non-significant and has small effect size, $t(75) = 1.14$, $p = .26$, Power = 0.35, Cohen's $d = 0.26$ (small effect size).

Speech interruptions

The segmentation of data in the sentence format also allowed me to conduct a quantitative content analysis that examined jurors' speaking interruptions and interrupting speech acts. The dashes in the deliberations indicated that a juror's utterance was interrupted (either by self or another juror). As such, I combed the transcripts and coded any instance where a juror was interrupted by another juror, and where a juror was interrupting. I was particularly interested in whether there would be a gender difference in jurors' interrupted/interrupting utterances. Past research has posited that men are more likely to interrupt others' speech, and women are more likely to be interrupted (Rosenblum, 1986). However, more recent research on gender and conversational styles has found that this is not the case. Specifically, James and Clarke (1993) found that men and women do not significantly differ in the ways they simultaneously talk (i.e., interrupt, interrupted). My study results in the present study support the latter findings about gender and simultaneous speech. Specifically, based on the counts of co-occurring instances of interruption (e.g., one juror interrupting another jurors' utterance as indicated by a - in the transcript), there was not a significant difference in women jurors ($M = 5.24$, $SD = 7.70$) and men ($M = 5.78$, $SD = 6.85$) jurors interrupting other jurors, $t(75) = -0.32$, $p = .75$, Power = 0.06, Cohen's $d = 0.07$ (nil effect size). In fact, men jurors were more likely to be interrupted than women jurors, though there was also not a significant difference in the instances of men ($M = 7.64$, $SD = 9.22$) and women ($M = 5.32$, $SD = 7.29$) jurors being interrupted, $t(75) = -1.23$, $p = .22$, Power = 0.22, Cohen's $d = 0.28$ (small effect size).

In the man defendant cases, there was not a significant difference in women ($M = 5.95$, $SD = 9.74$) and men ($M = 4.00$, $SD = 5.56$) jurors' interruptions of other jurors' speech, $t(38) = 0.78$, $p = .44$, Power = 0.11, Cohen's $d = 0.24$ (small effect size). In the woman defendant cases, there was not a significant difference in men ($M = 8.00$, $SD = 7.80$) and women ($M = 4.57$, $SD =$

5.24) jurors' interruptions of other jurors' speech, $t(35) = -1.59, p = .12$, Power = 0.33, Cohen's $d = 0.52$ (medium effect size).

In the man defendant cases, there was not a significant difference in men ($M = 6.90, SD = 10.26$) and women ($M = 5.25, SD = 7.69$) jurors being interrupted by other jurors, $t(38) = 0.78, p = 0.44$, Power = 0.09, Cohen's $d = 0.18$ (small effect size). In the woman defendant cases, there was not a significant difference in men ($M = 8.56, SD = 7.97$) and women ($M = 5.38, SD = 7.08$) jurors being interrupted by other jurors, $t(35) = -1.28, p = 0.21$, Power = 0.23, Cohen's $d = 0.42$ (small-medium effect size).

Sentence Tags

At the utterance-level of analysis, I also examined jurors use of sentence tags or verbal hedging (e.g., "right?", "you know?"). Lakoff (1975) controversially argued that verbal hedging is more common in women as it mirrors broader societally expected gender roles for women. Specifically, verbal hedging is conceptualized as a means to maintain relationships with other speakers and express tentativeness, characteristics that are stereotyped as feminine (e.g., caring and affiliative). However, Dixon and Foster (1996) found no such effect of speaker gender on their use of hedging language (e.g., "you know"). In the present study, I did not find a significant difference in the amount of sentence tags used by men ($M = 0.72, SD = 1.60$) and women ($M = 0.90, SD = 1.76$) jurors, $t(75) = 0.47, p = .64$, Power = 0.08, Cohen's $d = 0.11$ (small effect size). Jurors' use of sentence tags also did not significantly differ by defendant gender.

In the woman defendant case, women jurors ($M = 1.14, SD = 2.20$) did not use significantly more sentence tags than men jurors ($M = 1.13, SD = 2.09$), $t(35) = 0.025, p = .98$, Power = 0.05, Cohen's $d = 0.005$ (nil effect size). There also was not a significant difference in men ($M = 0.40, SD = 0.99$) and women ($M = 0.65, SD = 1.14$) jurors' use of sentence tags in the man defendant case, $t(38) = 0.74, p = .46$, Power = 0.10, Cohen's $d = 0.23$ (small effect size).

Assertive and affiliative language

Researchers have posited that men and women use different styles of speech. Men tend to use slightly more assertive language (e.g., cognitive statements) in comparison to women, and women tend to use slightly more affiliative language (e.g., feeling statements) in comparison to men (see Leaper & Aryes, 2007, for a review). In terms of the content of the individual utterances in the present study, I coded for “feeling” and “cognitive” statements. Specifically, I conducted a word search for “I feel”, “I think”, “I know” and coded the associated sentences as a feeling statement or a cognitive statement. Interestingly, women ($M = 4.78$, $SD = 4.99$) and men ($M = 5.86$, $SD = 6.97$) were equally likely to make cognitive statements, $t(75) = -0.79$, $p = .43$, Power = 0.12, Cohen’s $d = 0.18$. (small effect size) There also was not a significant difference in the amount of feeling statements men ($M = 0.67$, $SD = 1.17$) and women ($M = 1.66$, $SD = 3.09$) jurors made, $t(75) = 1.82$, $p = .07$, Power = 0.46, Cohen’s $d = 0.43$ (medium effect size).

Moreover, in woman defendant cases, there was not a significant difference in the cognitive statements made by women jurors ($M = 4.29$, $SD = 4.48$) and men jurors ($M = 5.50$, $SD = 5.91$), $t(35) = -0.71$, $p = .48$, Power = 0.11, Cohen’s $d = 0.23$ (small effect size). In the man defendant cases, there was not a significant difference in the cognitive statements made by women jurors ($M = 5.30$, $SD = 5.54$) and men jurors ($M = 6.15$, $SD = 7.86$), $t(38) = -0.40$, $p = .70$, Power = 0.07, Cohen’s $d = 0.12$ (small effect size).

In woman defendant cases, there was not a significant difference in the feeling statements made by women jurors ($M = 1.90$, $SD = 3.27$) and men jurors ($M = 0.31$, $SD = 0.70$), $t(35) = 1.91$, $p = .065$, Power = 0.50, Cohen’s $d = 0.67$ (medium-large effect size). Moreover, in the man defendant cases, there was not a significant difference in the feeling statements made by women jurors ($M = 1.40$, $SD = 2.95$) and men jurors ($M = 0.95$, $SD = 1.39$), $t(38) = 0.62$, $p = .54$, Power = 0.09, Cohen’s $d = 0.20$ (small effect size).

Juror gender, influence, and leadership

Participants were asked to report whom they perceived as the most influential juror in deliberation, and to report which juror they perceived to be the leader of the deliberation. I recorded jurors' responses and noted the gender of the jurors who were perceived as influential and as leaders. My results demonstrated that there was not a significant difference in jurors' perceptions of juror gender and influence, $\chi^2 = 2.38, p = .12$. Interestingly, men jurors were more likely to be perceived as the leaders of deliberations than were women jurors across the deliberations, $\chi^2 = 7.05, p = .008$. These seemingly contradictory results may be due to the fact that men and women are equally likely to be perceived as effective (i.e., influential) leaders (Eagly, Karau, & Makhijani, 1995), though men matched jurors' stereotype for an agentic leader and thus were more likely to be labelled as such (see Bongiorno, Bain, & David, 2013).

Discussion

Past research has proposed that women and men speak in a manner that is congruent with gender stereotypes about agency and docility (Lakoff, 1975). The past literature on gender and communication grew out of feminist concerns with overcoming systemic sexism (Popp, Donovan, Crawford, Marsh, & Peele, 2003). While the popular *differences model* argues that men and women are significantly different in terms of their psychology is well known, Hyde's (2005; Hyde and Plant, 1995) meta-analysis demonstrated that women and men are more psychologically similar than they are different. The findings of Study 3 of the present research mirror the literature on the absence of gender differences in psychology and communication styles (Leaper & Aryes, 2007).

Table 6 *Sample descriptives using independent sample t-tests*

	Juror Gender	Overall				Woman Defendant				Man Defendant			
		<i>M</i>	<i>SD</i>	<i>t</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>d</i>
Speaking turns/ talkativeness	Men	101.92	101.13			123.69	118.84			84.50	83.52		
	Women	81.51	75.92	-1.01	0.22	91.24	67.98	-1.05	0.34	71.30	83.99	-0.50	0.16
Speech interrupting	Men	5.78	6.85			8.00	7.80			4.00	5.56		
	Women	5.24	7.70	-0.32	0.07	4.57	5.24	-1.59	0.52	5.95	9.74	0.78	0.24
Speech Interrupted	Men	7.64	9.22			8.56	7.97			6.90	10.26		
	Women	5.32	7.29	-1.23	0.28	5.38	7.08	-1.28	0.42	5.25	7.69	0.78	0.18
Sentence tags	Men	0.72	1.60			1.13	2.09			0.40	0.99		
	Women	0.90	1.76	0.47	0.11	1.14	2.20	0.025	0.005	0.65	1.14	0.74	0.23
Cognitive Statements	Men	5.86	6.97			5.50	5.91			6.15	7.86		
	Women	4.78	4.99	-0.79	0.18	4.29	4.48	-0.71	0.23	5.30	5.54	-0.40	0.12
Feeling Statements	Men	0.67	1.17			0.31	0.70			0.95	1.39		
	Women	1.66	3.09	1.82	0.43	1.90	3.27	1.91	0.67	1.40	2.95	0.62	0.20

Specifically, I found null effects throughout Study 3 which may suggest that men and women jurors are more similar than different in the manner they deliberate (e.g., similar speaking turns, interruptions, and similarly perceived as influencer). This is not to suggest that men and women communicate in mirrored and egalitarian manners in all contexts (Aries, 1987), however these findings may suggest that within the jury deliberation room, jurors' roles and discussions are more parallel than disparate.

Implications

In the present study, there was not a significant relationship between juror gender and speech style. The lack of a significant relationship between these variables may be due to a moderating impact of discussion context and speaker roles. Leaper, Anderson, and Sanders (1998) found that the context in which people are speaking impacts on the speech styles they use. Specifically, in task-oriented settings (e.g., jury deliberation room), assertive speech may broadly be more common for men and women (Leaper & Aryes, 2007). The present study corroborates this theory, as women and men generally used more assertive speech (i.e., cognitive statements) than affiliative speech (i.e., feeling statements) in jury deliberations. According to a paired-sample *t*-test, jurors in this study were significantly more likely to use assertive speech ($M = 5.29$, $SD = 5.98$) than affiliative speech ($M = 1.19$, $SD = 2.43$), $t(76) = 6.32$, $p < .01$, Cohen's $d = 0.90$ (large effect), Power = 0.96 across deliberations (these significant effect sizes should be understood within the context of this study's small sample size, thus the implications of these results are tempered).

Limitations

The deliberations in the present study took place in a setting unfamiliar to the jurors in a room in which they were clearly being recorded. Jurors' egalitarian interactions may have been a result of their being aware that they were being watched and evaluated (e.g., social desirability).

Study 4: Thematic Analysis

Descriptive, qualitative analysis allows for the exploration of the research topic when there is limited literature and results of other studies to rely on to understand the issue (Creswell, 2013). This research paradigm is thus appropriate considering the very limited number of studies examining jury decision-making at a group level (and no studies examining the intersection of gender and mental illness stereotypes in the jury box). Moreover, as my quantitative hypotheses were not supported by the data, I decided to qualitatively and inductively explore what jurors *were* talking about (e.g., case facts, attitudes expressed) and the arguments of these deliberations (i.e., themes).

According to Denzin and Lincoln (2011), qualitative research:

“...involves an interpretive, naturalistic approach to the world. This means that qualitative researchers study things in their natural setting, attempting to make sense of, or interpret, phenomena in terms of the meaning people bring to them” (p. 3).

According to Creswell (2013), qualitative research is appropriate when a problem or issue needs to be explored, particularly if the group or population and variables of interest cannot be easily quantified and measured (e.g., juror attitudes and deliberation processes). The mission of qualitative research is thus to find meaning and understanding of a topic or interest (using its own specific epistemological foundations), rather than test truths and predict outcomes (Myers, 2000).

Since the ‘narrative turn’ of the social sciences in the 1980s (Polletta, Chen, Gardner, & Motes, 2011), it has become increasingly important and evident that language is an important symbolic resource that reflects and creates social reality (see Berger & Luckmann, 1966). Specifically, I perceive language as mirroring reality, thus the language people use represents their attitudes and ideas (see Braun & Clarke, 2018). Talk, and more specifically, stories and narratives, are particularly relevant in the jury deliberation room as jurors are forming stories about the case (Devine, 2012), are putting forth their own extra-legal stories, and are evaluating the defendant’s “story”. The language jurors utilize form their and other jurors’ opinions and their ultimate verdict decisions. In the present study, I am particularly interested in uncovering which ideas or themes are salient and discussed during deliberations.

There are a variety of qualitative methods that can be used to understand jurors’ notions of the case (e.g., narrative analysis, thematic analysis, phenomenological research, ethnographies). I have decided to use the thematic analysis approach to studying jurors’ deliberations. Thematic analysis (TA) is a “method of systematically identifying, organizing, and offering insights into patterns of meaning (themes) across a data set” (Braun & Clarke, 2018, p. 57). Braun and Clarke (2018) also explain that TA is a useful method to understanding and to making sense of commonalities in the way a topic/idea is talked about (e.g., in deliberation). I chose to use the thematic analysis methodology because it allows careful and systematic analysis of large bodies of written text for which I did not have a specific research question (other than, what are jurors saying when deliberating about women and men defendants in insanity trials?). Thematic analysis allows for a primarily inductive examination of what is in the data (bottom-up approach to analysis), with themes being formed by the content of the data, which can then be tied to theory (Braun & Clarke, 2018). According to Marshall and Rossman (1999: p. 150),

thematic analysis of qualitative data involves: “bringing order, structure and interpretation to mass collected data...It is the search for general statements about relationships among categories of data...it is the search of data to identify content”. These categories of information (or themes) were developed from an inductive analysis as I proceeded through reading and familiarizing myself with the data.

Method

I followed Braun and Clarke’s (2006) outline on how to conduct a thorough thematic analysis. In their article, Braun and Clarke (2006) broadly outline six phases of conducting a thematic analysis (e.g., generating initial codes, searching for themes, reviewing themes...). For my analysis, I was interested in the semantic-level themes (rather than the latent-level themes) that emerged from the data. That is, I examined the explicit and surface meanings of the jurors’ utterances and as such summarized and described seven themes of talking points that occurred across the 10 deliberations. All analyses were conducted using the verbatim transcriptions of the 10 deliberations of my research project. All deliberation transcripts were uploaded to and coded using ATLAS.ti (version 8.4.2 (974)). Below I outline the steps I followed when conducting the thematic analysis (and a content analysis), after which I describe the themes that were the most prominent across the deliberations (see Braun & Clarke, 2006).

1. Familiarize self with data (as was done during initial content analysis): This step is accomplished by reading the deliberations several times each prior to and during coding. Braun and Clarke (2006) recommend reading the full data set (10 deliberations) at least once prior to coding. I uploaded each of the 10 deliberations into ATLAS.ti, where I read each of the 10 deliberations once prior to coding, and then several more times while coding and recoding the deliberations. Prior to analysis, I also “tagged” each of the data segments with a label of juror

code and juror gender (using the ATLAS.ti coding function). I also read each deliberation and wrote abridged versions of transcripts on a line by line basis so that I could quickly overview the ideas of the deliberation and have notes on the ideas that were most salient in the deliberation.

2. Generating initial codes: Once I was familiar with the data, I had a general impression of the main ideas of each of the deliberations. According to Nowell, Norris, White, and Moules (2017), qualitative coding allows the researcher to simplify and focus on specific characteristics of the data set; this process allows the researcher to move from an unstructured data set to a systematic and organized set of ideas found in the data. I utilized ATLAS.ti to aid me in sorting and organizing the large data set and the complex network of codings I had derived. Over the course of coding the approximately 7,000 juror utterances across the 10 deliberations, I created 532 “raw” codes. These codes each had explicit boundaries. After my final review of the codes, some codes were redundant and were thus merged in a later step of analysis (Nowell et al., 2017). Each segment of deliberation was coded as belonging to one of the 532 raw codes (segments were at times deemed to fit into more than one code). My coding process was inductive and hierarchical, allowing me to analyze the data for word use (for the content analysis) and ideas/themes (for the thematic analysis). Therefore, the codes and related themes that emerge from the data are closely tied and not fitted to any pre-existing coding frames or researcher preconceptions (Braun & Clarke, 2006). That is to say, the thematic analysis is “data-driven” (Nowell et al., 2017).

3. Searching for themes / 4. Reviewing quality of themes: According to Desantis and Ugarriza (2000), “A theme is an abstract entity that brings meaning and identity to a recurrent experience and its variant manifestations. As such, a theme captures and unifies the nature or basis of the experience into a meaningful whole” (p. 362). Accordingly, after completing the

initial coding round of the thematic analysis, I printed out the list of my codes and began to group codes that seemed conceptually similar (e.g., “need more witnesses”, with “need more experts”, and “need more evidence” or “defendant’s a liar with a “story”” with “others would have noticed illness”). I re-read these themes and the codes contained and re-evaluated the fit of each of the codes to ensure that they were not redundant within the theme or repeated across themes. While these themes are not dependent or quantifiable measures, they are refined through a review process to capture important ideas throughout the deliberations (Braun & Clarke, 2006). Themes allow researchers to understand and tie together the evidentiary information and narratives that are most significant to jurors in NCRMD trials. While reviewing and creating themes, there were codes that did not fit in any of the themes that emerged; as such, these were placed in a “miscellaneous” theme to temporarily house these codes (Nowell et al., 2017). By the end of this phase, I had a good idea of the themes and the story they form about the data set (see Braun & Clarke, 2006).

5. Defining and naming themes: This phase of analysis involved defining and naming the codes in a manner that was detailed and specific to the individual theme (Braun & Clarke, 2006). The title of the theme should be somewhat descriptive and illustrative of the main idea of the theme (Braun & Clarke, 2006). The seven categories/themes have been labelled as “Defendant relationships”, “Legal-lay discourse”, “Defendant ‘faking’”, “Temporality and rationality of mental illness”, “Psychiatrist’s credibility”, “Insufficient evidence”, and “Verdict outcome”. Notably, there is undoubtable overlap in the constructs and ideas of these themes; nevertheless, the themes offer an interpretation and map of the points of discussions jurors had across deliberations. Considering that I was a solo researcher on this project, I read through the codings

and themes within the data twice, and I consulted my advisor, Dr. Evelyn Maeder, to ensure that the themes were sufficiently clear and comprehensive.

6. Producing the report. The findings below provide a descriptive elaboration and theoretical explanation for each of the themes that emerged across the 10 deliberations of the present study.

Establishing trustworthiness

Lincoln and Guba (1985) have defined qualitative research trustworthiness as consisting of the criteria of credibility, transferability (i.e., generalizability), dependability (i.e., research process is logical and clearly documented), and confirmability; these mirror the quantitative research criteria of validity and reliability, respectively. These criteria comprise a widely accepted and easily recognized manner of examining research trustworthiness (Nowell, Norris, White, & Moules, 2017).

Credibility refers to the extent to which research participants' views and responses are accurately represented by the researcher (Tobin & Begley, 2004). Lincoln and Guba (1985) recommend peer debriefing as an external check on the research process (e.g., reviewing process and findings with research advisor).

Transferability is attained by the researcher providing rich and thorough descriptions of the research methods and findings so that other researchers can determine if the study's findings may transfer to their own research study (Lincoln & Guba, 1985).

Dependability is met when the researcher makes the research process clear and traceable (e.g., provides a clear and systematic delineation of and rationale for analysis decisions) (Lincoln & Guba, 1985).

Confirmability is met when the prior criteria are met (i.e., credibility, transferability, and dependability). Confirmability is also met when the researcher clearly and explicitly demonstrates how the conclusions and interpretations are derived from the data (Nowell et al., 2017).

Establishing saturation

Saturation is a critical and yet somewhat elusive benchmark on recommended sample size for qualitative research. According to Glaser and Strauss (1967, p. 65) data saturation is met when:

“no additional data are being found whereby the (researcher) can develop properties of the category. As he sees similar instances over and over again, the researcher becomes empirically confident that a category is saturated”.

Similarly, Guest, Bunce, and Johnson (2006) broadly define data saturation as “the point in data collection and analysis when new information produces little to no change in the codebook” (p. 65). Importantly, Morse (1995, p. 147) cautioned that while “saturation is the key to excellent qualitative work...there is no published guidelines or tests of adequacy for estimating the sample size required to reach saturation”. As such, it can be challenging to determine when a data set has reached the point of saturation (or data redundancy). In a study of focus group interviews and data saturation, Guest, Namey, and McKenna (2017) empirically tested how many group interviews were required to reach saturation; they found that 90% of all themes were discoverable within three to six focus groups.

Results

Defendant ‘faking’

This theme centres on jurors' concern that the defendant was not mentally ill and was thus falsely pleading NCRMD. Jurors expressed this worry in a number of narratives, many of which focused on the timeline of the defendant's diagnosis and history of mental illness. Jurors utilised these narratives to justify their perceptions that the defendant was faking mental illness as a means to be found NCR (which they also typically perceived as being 'an easy out' of prison or a 'danger to society').

Diagnosis timing. The timing of the diagnosis as being 'two weeks after arrest' was a major point of consternation for jurors across cases. Juror 3 on Jury N (i.e., J3N) posited:

What was between the incident and those two weeks that led him [the psychiatrist] to diagnose her as schizophrenia? ... What happened between that... Incident and a month- And a month and a half before he diagnosed her? What was that whole time period of six weeks for him to diagnose her and state. Why did he diagnose her with schizophrenia if he had a month to diagnose her with that? What was going on that whole month? (J3N)

J3N's worry about the timeline of the mental illness, act, and diagnosis was a common theme, typically applied to cast doubt on the defendant's past and current mental illness status and diagnosis. In Jury K, jurors tried to understand why J8K was not willing to find the defendant NCRMD. To answer their questions, J8K indicated his concern about the defendant's lack of mental health history:

J2K: Is there anything that would change your minds?

J9K: Like what kind of information is missing for you to decide that she is mentally ill?

J8K: The woman has no signs of schizophrenia from previous-

J3K: We don't have that information...

J8K: There's no history of mental illness, the girl's a grown adult.

Defendant's contrived 'story'. At times, jurors were explicit about their perception that the defendant was malingering. At times, jurors outright referred to the defendant's mental illness diagnosis as a 'story', so as to portray the defendant as lying about being mentally ill.

I personally don't buy the story. I think it's extremely convenient. And therefore, that's why I stand the way I stand. ... It's just, it's just, I don't know, like, the, the fact that it's just two weeks and, like, it's just a very convenient story. (J6G)

Jurors in Jury O also characterised the defendant's diagnosis and plea of NCRMD as stereotypical (or 'cliché') to suggest the defendant was malingering. J7O noted, 'She had two weeks to build up the rest of the story.' J2O added, 'Very clichéd story. Not that it is untrue, but it is also what is present in like, entertainment, and yeah, movie series.'

Family's role and responsibility. Jurors were notably concerned about the defendant's mother and friends (not) noticing the defendant's mental illness prior to the act. This point of

inquiry was also typically utilised to cast doubt on the expert testimony about the defendant's mental illness diagnosis or to substantiate why the juror would not vote NCRMD:

Like, cuz [sic] if someone is, like if someone is so mentally ill, that they're like, have schizophrenia, I think that, like, you wouldn't even try to get them, like, I feel, you wouldn't be able to live on your own, and, how would no one in your family know about it? (J2G)

Moreover, J5 in Jury M was particularly concerned that nobody in the defendant's life testified that they had noted the defendant's 'weird' behaviours, suggesting that J5M had a specific stereotype for the progression and manifestation of mental illness:

So, if they had, say, her mom come in and say, 'Yeah, she's been acting kind of weird, she's been talking strange', or anything like that. ... Something to support that it's been a recurring thing, or that she's having issues that even if she never went and got a proper diagnosis, there's still somebody there to say yeah, something's up with her, we- You know, she made some weird comments months ago, or something like that. So, I can't, I can't say that, uhh...She's not responsible. (J5M)

As Perlin (1991; 1996; 2003; 2009; 2013) has argued, the public reaction to the insanity defence is a litmus test for the broader society's perspectives of people living with mental illness – specifically those who come in contact with the law. People perceive the insanity defence to be misused and in need of reform (e.g., the defendant is not actually mentally ill) (Hans, 1986), and

the public generally has negative perspectives of mentally ill persons (see Angermeyer & Matschinger, 2003). It thus makes sense that across the deliberations in the present study, the jurors spent considerable time and effort deciphering whether the defendant was truly mental ill and thus eligible for the NCRMD verdict.

Moreover, according to Finkel (1995, p. 63), ‘jurors do not so much find reality as *construct* it’, regardless of the legal information they are provided. For example, jurors look to their own experiences and knowledge about mental illness and the NCRMD defence when making their verdict decisions (Skeem & Golding, 2001). In the deliberations in the present study, jurors often seemed to consult their notions of mental illness history and trajectory. That is, they assumed a long, gradual, and linear history of mental illness in people, with clear cut time stamps of increasingly disturbing signs and symptoms of illness. Skeem and Golding (2001) have found that jurors tend to refer to three general prototypes when conceptualising and evaluating defendants who plead insanity. These prototypes emphasise severe mental disability, ‘moral insanity’, and the defendant’s state of mind at the time of the offence (Skeem & Golding, 2001). All of these prototypes were reflected in the present study. Within the context of the ‘defendant faking’ theme, jurors seemed particularly focused on the severity of the defendant’s mental illness and how the illness impaired his or her rationality and actions.

Contradictorily, jurors also expected the defendant and people with mental illness to have exceptional mental illness histories (e.g., family noticing unusual behaviours) that were not ‘cliché’. Jurors thus seemed to have an ambivalent perspective of mental illness and how it impacts the individual inside and outside the courtroom. This is in line with Kopera, Suszek, Bonar, Myszk, Gmaj, Illgen, and Wojnar’s (2015) findings that attitudes towards people with mental illness are ambivalent among mental health professionals and medical students,

regardless of their contact with mentally ill people. Jurors' focus on the importance of others noticing the defendant's mental illness was potentially due to the public's skewed perspective on how mental illness is recognised (e.g., expecting unusual behavioural evidence of mental illness) (see Angermeyer & Dietrich, 2006).

Jurors may also have focused on the potential of defendant malingering because they perceived the NCRMD verdict as not punitive and were concerned that the defendant would immediately be released into the public (e.g., insanity myth, see Perlin, 1996). The worries about defendant malingering and mental illness history perceptions are in line with what other researchers have found about the public's general attitudes towards mentally ill people, violence (Stuart, 2003), and the insanity defence (Hans, 1986). Specifically, jurors appeared to be expressing stereotypes of dangerousness (e.g., unpredictable and violent) associated with mentally ill people, and especially people with schizophrenia (see Angermeyer, Buyantugs, Kenzine, & Matschinger, 2004; Angermeyer & Dietrich, 2006). As such, jurors' focus on the defendant's potential to 'fake' his or her mental illness may indicate their fear of and negative attitudes towards mentally ill people and the NCRMD defence.

Temporality and rationality of mental illness

This theme centres on the variables that jurors considered when trying to make sense of the defendant's state of mind at the time of the act. Jurors may have used this evidence to shape a narrative that the defendant was/was not mentally ill. Most of the case facts thus held ambiguous meanings – depending on the juror, they could demonstrate the defendant being mentally ill or not mentally ill. This theme is closely related to the themes of 'defendant faking' and 'psychiatrist's credibility'. That is, the defendant's state of mind prior to, during, and after the act constituted jurors' understanding of the defendant's criminal liability or blameworthiness (i.e.,

mens rea). For example, J5L asked, 'At that point in time, when she stabbed him... Was she schizophrenic?' J8L added, 'Do you think there is enough evidence... to that? Do you think it is more likely than not that she is schizophrenic, at that time?'

There was also substantial discussion on whether the defendant's actions were rational or irrational, which was likely aimed at determining the defendant's mental illness status, as jurors seemed to assume that one cannot be rational and mentally ill.

They have the delusions, and decrease in intensity over time, and they're more in and out of reality. Like, it changes in phases; it's not like a constant delusion... But, for me with the evidence presented I simply can't make a judgment of his state of mind at the time of the murder. For me, that act's rational, and that would indicate that he wasn't experiencing powerful delusions, if [he] was carrying out a rational act. And that's all that there is at the time. Because I am not going to take his word, even if the psychologist [is]. (J2C)

Lastly, the defendant's tolerance for psychiatric medication was a common point of discussion brought forward to support the notion that as the defendant appeared rational at trial, this was a result of his or her tolerance of the anti-psychotic medication. In other words, the defendant's tolerance of the medication indicated that he or she was truly mentally ill (and should be found NCRMD).

Another thing is the fact that he is on anti-psychotic medication and he's conscious, I do believe he's schizophrenic. Most anti-psychotic on a person who doesn't have a chemical imbalance in their brain, they'd probably stay in a corner for a few weeks until it wears off.

(J1C)

The defendant's state of mind at the time of the act (i.e., *mens rea*) is of import in most criminal cases (Butler, 2006), and is particularly salient to jurors when the defendant pleads insanity. In addition, Skeem and Golding (2001) found that the defendant's state of mind at the time of the offence was a significant prototype in jurors' conceptualisation of that defendant's guilt. Specifically, jurors in Skeem and Golding's (2001) analysis of insanity prototypes emphasised severe functional impairment of the defendant's mental illness and a longstanding illness (e.g., chronic and long history of mental illness that cannot be controlled by medication). That is, jurors seemed to trust that the defendant was mentally ill (e.g., illogical and delusional) at the time of the act if he or she had a prolonged and chronic history of mental illness. Similarly, in the present study, the jurors' perceptions of the defendant's *mens rea* seemed to be contingent upon their belief in the defendant's history of mental illness.

According to Malle and Nelson (2003), jurors' perceptions of a defendant's intention, motive, and forethought (i.e., *mens rea*) are often based on jurors' folk concept of the defendant's *mens rea*. In the present study, the jurors' folk construct of *mens rea* seemed to imply that mental illness precludes a person from acting rationally, and that acting rationally precludes a person from having a mental illness. Similarly, Bortolotti (2013) has argued that the diagnostic criteria for several psychiatric disorders tend to preclude rationality. In other words, to be found mentally ill, a person must demonstrate that he or she is experiencing fractured

rationality. The phenomenon of delusions (which Bortolotti [2009] defines as ‘unjustified beliefs and inaccurate narratives’) is largely determined by the behavioural manifestation of the disorder (Bortolotti, 2013). Accordingly, it is perhaps not surprising that jurors also perceive mental illness and rationality to be mutually exclusive. Moreover, that the defendant’s rationality and mental illness are evaluated based on his or her behaviours.

In the present study, jurors’ perceptions of the defendant’s mental illness and *mens rea* appeared to be based on determining the defendant’s state of mind not just at the time of the act, but also prior to the act (history of mental illness) and after the act (defendant’s reaction to psychiatric medication). The jurors’ focus on the temporality of the defendant’s mental illness is congruent with psychiatric discourses and practices with regard to the ‘chronicity’ (i.e., symptoms persisting constantly) of a mental disease (von Peter, 2009). Time is perceived as an objective measure of the legitimacy of the defendant’s claim of mental illness (von Peter, 2009). Jurors deliberated on their expectations of the temporal progression of delusions, and they often seemed to expect the defendant with mental illness to have chronic signs of delusions (e.g., disturbed behaviour and cognitions) prior to and after the act.

Jurors’ reliance on temporality (based on their imposed temporal frame of the defendant’s acts) of the defendant’s schizophrenia-related behaviours may imply their perception that time is an objective (and therefore reliable) metric of the truthfulness of the defendant’s diagnosis (see von Peter, 2009). Jurors’ perceptions of time (especially the temporality of crime) may thus be an important component of jury decision-making to consider, as Crawford (2015) has argued that

time affects and informs the experience and perception of justice and our understanding of regulatory practices in diverse ways. Feeling secure demands not only the absence of direct

harms in the moment but also a reasonable confidence that the conditions assuring our safety will continue into the foreseeable future. (p. 472)

Furthermore, jurors often hold negative and fearful perceptions of the defendant pleading NCRMD's potential for future harm (see Perlin, 1997). Accordingly, time may be of particular interest to jurors in NCRMD cases.

Defendant-victim relationship

The 'defendant-victim relationship' theme refers to jurors' justifications for the defendant's violence. Specifically, this theme contains elements addressing the dynamics of the defendant and victim's relationship and the history of victim-defendant violence (not necessarily related to mental illness):

in terms of evidence, like before she killed her roommate. There was no, like, history of, at least no obvious history, from what I could see, of her being violent, or you couldn't really figure out what there even was to the argument. (J2K)

Moreover, J3C indicated that the relationship (i.e., perceived friendship) of the defendant and victim was significant in determining the jury's verdict. Specifically, the jurors perceived the defendant as being mentally ill because of how he interacted with his roommate (i.e., the victim) in a non-normative, violent manner:

Umm what sort of decided it for me, the type of mental illness he was supposed to have and that was one where he harmed his friend. So that tells me that [he is the] person's [the defendant's] friend. And if he's friends with him, I think there has to be something very wrong there. (J3C)

The literature on jurors' perceptions of victim-defendant relationships and crime has demonstrated divergent results (Gross & Mauro, 1984; Simon, 1996a). Some studies have found that stranger (in comparison to non-stranger) offences receive longer sentences (Gross & Mauro, 1984; Simon, 1996a), while others have reported no effect of the victim-defendant relationship on sentence length (Simon, 1996b). Moreover, some researchers have observed that the degree of intimacy in the relationship influences jurors' perceptions of guilt (Dawson, 2004). Specifically, the more intimate the relationship (in comparison to stranger relationship cases), the more likely jurors will be to recommend the death penalty (though only for female victims) (Gillespie, Loughran, Smith, Fogel, & Bjerregaard, 2014). As such, while I cannot make an inferential statement on the effect of jurors' perceptions of the defendant-victim relationship on their verdict decisions, my research demonstrates that this relationship may be a notable component of jurors' deliberations in NCRMD trials.

Psychiatrist's credibility

This theme centres on jurors' perceptions of the psychiatrist's (lack of) authority and reliability in determining that the defendant was mentally ill. Jurors discussed the possibility that the psychiatrist could have been duped into diagnosing the defendant with schizophrenia in order for the defendant to avoid prison with an NCRMD verdict. They also talked about the psychiatrist's trustworthiness, expertise, and testimony.

Psychiatrist as ‘hired gun’. Some jurors were concerned that the psychiatrist was a ‘hired gun’ who had been brought into court by the defence to convince the jurors to erroneously find the defendant NCRMD. Juror 11 in Jury H made this point very bluntly: ‘Professionals can be paid off...If they needed help convincing the jury, the lawyers, the Crown, whatever, about um, the individual being mentally disordered? Um, so I’m just thinking if that would play into this scenario here.’

The primary responsibility of an expert witness (here, the psychiatrist in the trial) is to present empirical evidence to assist the triers of facts in the court in making legal decisions (Shuman & Greenberg, 2003). In Canada, the Supreme Court has ruled that the admissibility of expert testimony (or evidence) is governed by four factors: the testimony and evidence must have relevance to the case; they must be necessary in assisting the triers of fact; there must be no exclusionary rules attached to the evidence; and the expert must be properly qualified (*R v. Mohan*, 1994). Ultimately, the *Mohan* ruling ensures that expert opinion must be beyond the knowledge of the average person; the subject matter (testimony) must allow triers of fact to make an educated verdict (Glancy & Bradford, 2007). The Court later stated that evidence and expert testimony should not be admitted when there is danger of misuse and distortion of the legal decision-making processes of the judge and jury (Glancy & Bradford, 2007). That being said, in this study the juries were generally sceptical of expert testimony and not adept at deciphering the strength and relevance of psychological expert testimony.

Psychiatrist’s expertise. Some jurors were more trusting in the psychiatrist’s testimony and specifically pointed to the expert’s years of diagnostic experience as justification for trusting the expert’s testimony:

J1G: I feel like I believe in the doctor's diagnosis of the schizophrenia because the doctor has 20 years' experience.

J12H: If you have 20 years' experience in the field you're going to know what to look for so, I, I would say it's more likely than not that if it's that psychiatrist's professional opinion [of] that...case, I would be likely to believe that.

Psychiatrist's naivety. Considering the jurors' focus on whether or not the defendant was mentally ill, another important point of debate concerned the psychiatrist's naivety and whether or not the psychiatrist was '(un)trickable' – that is, whether the psychiatrist would have been able to tell if the defendant was malingering. For example, in Jury N, jurors discussed whether the psychiatrist was able to diagnose effectively if the defendant was malingering (the jurors assumed the defendant was not mentally ill and was faking mental illness to receive the NCRMD verdict in lieu of a guilty verdict):

J2N: I said that he is the person who makes the best decision when you're diagnosing in the whole room. So, in the whole courtroom, of all the people that could diagnose someone with mental disorder, he is the person who does it best. I didn't say he does it perfect.

J3N: But he just stated that there's a possibility of her tricking him. I'm still stuck on her being guilty because he stated that he is-

J2N: But the point of picking him to do the diagnosis is that he is the person who has the least chance of being tricked.

According to Jorm (2000), much of the public is not able to recognise and identify the symptoms and characteristics of specific mental disorders and mental distress. Moreover, the public has notably different perspectives on the causes and appropriate/effective treatment of mental illness. That is, the public generally has limited or poor 'mental health literacy' (Jorm, 2000). As such, jurors do not typically have a strong grasp on most of the scientific and complex evidence that is presented by psychological expert witnesses. They also have difficulty distinguishing between strong and weak evidence and reliable testimony (Cutler, Penrod, & Dexter, 1989).

Jurors' misunderstanding of expert testimony may be the reason why they focus on the characteristics of the expert witness (e.g., years of experience, psychiatrist's ethics) rather than the details of the psychiatric testimony. Researchers have found that jurors are influenced by extra-legal factors such as the expert's gender (McKimmie, Newton, Terry, & Schuller, 2004; Neal & Brodsky, 2008), the evidence's complexity (Cooper, Bennett, & Sukel, 1996; Cooper & Neuhaus, 2000), the expert's appearance (Memon & Shuman, 1998), and the expert's credentials and area of expertise (see Cutler et al., 1989; see also Kovera, McAuliff, & Hebert, 1999).

When expert witnesses use specialised language, it is likely that jurors will find their testimony difficult to understand (Cecil, Hans, & Wiggins, 1991) and will in turn use more peripheral cues to influence their legal decisions (Chaiken, 1980), such as the appearance of the expert or knowledge of the wages paid to the expert for his or her testimony. Accordingly, jurors may refer to their common-sense (i.e., lay knowledge) notions about the narrative proposed by the psychiatrist regarding the defendant's mental health and crime in his or her testimony. That is, jurors refer to and apply the knowledge with which they are most familiar (regarding mental illness and expert witnesses) to determine whether the expert's narrative is plausible (Ward,

1997). It may be because psychiatric testimony, diagnosis, and mental illness are largely misunderstood by the public that jurors rely on extra-legal factors concerning the expert and defendant when discussing the psychiatrist's credibility. In the present study, jurors seemed to be particularly sceptical of the psychiatrist's ability to detect defendant malingering. Jurors also countered such worries with arguments about the psychiatrist's years of experience (long career being indicative of greater knowledge and expertise).

Combining the roles of professional in psychological sciences and expert witness on trial is highly complex and ethically challenging, especially when the lawyers expect the psychologist to act as a 'hired gun' – an expert witness who is hired to convince the jury based on his or her credentials and not necessarily the soundness of his or her report (Boccaccini & Brodsky, 2002; Gemberling & Cramer, 2014). Psychological expert witnesses have come under fire by politicians, the media, and even other psychologists. For example, Dr Margaret A. Hagen, a psychologist studying visual perception, published a controversial book critiquing what she perceives as the overuse of psychological expert witnesses in all court cases (Hagen, 1997). Hagen is an avid proponent of eliminating all psychiatric testimony from courtrooms, as she believes that this testimony is based on subjective opinions, intuition, and 'junk science.' She refers to psychologists at trial as 'wizards' or 'whores of the courts' who will say anything is based on science for money. The crux of her argument is that whenever someone does something criminal, psychiatrists and psychologists are too often brought to court to convince the jury that

the defendant is mentally ill and should not go to prison, and ultimately such testimony should be banned from the courtroom (Hagen, 1997).⁹

While some experts may be motivated (e.g., by reputation or money) to present data in a manner that is favourable to the party that hired them regardless of the data, it is important to recognise that manipulating or purposely biasing testimony is legally prohibited (see Glancy & Bradford, 2007). Hence, the use of a hired gun expert witness should not be of serious concern to the courts. Moreover, research on jury decision-making has found that experts who are known to be paid high wages for their testimony are perceived as hired guns and are in turn not liked or trusted by jurors (Cooper & Neuhaus, 2000). As such, seeking a hired gun to testify defeats the purpose of this expert.

Legal-lay discourse

This theme centres on jurors' legal- and deliberation-related talking points, jurors' considerations of the standards for a verdict decision, and jurors' reflections on personal experiences.

Jurors' personal experiences. The legal-lay discourse theme is a broad one and also includes jurors' reflections on information outside the case facts. For example, when establishing their interest and role in the deliberation, J1C expressed their personal experience surrounding

⁹ In 1995, a New Mexico state senator named Duncan Scott put forth a bill amendment that would require psychological expert witnesses to wear wizard garb and a large beard while providing expert testimony to the courts. Senator Duncan's opinion was that there was a serious overuse of psychiatric/psychological professionals in legal trials and possibly that they had too much influence over jurors; interestingly, this amendment passed unanimously within the senate, but was never made law (Lamar, 2012).

mental illness and health: 'I've worked primarily with people with mental illness and action problems for 35 years, I guess. So, I have a personal interest in this I guess.'

Jurors are generally selected at random from a community (though they were self-selected in this study) and thus bring a variety of lived experiences, attitudes, and opinions to the jury deliberating room. Personal experiences, or anecdotes, are powerful rhetorical tools. Hence it is perhaps not surprising that jurors refer to such experiences when deliberating. I propose that jurors' expression of their lived experiences (e.g., working with people with mental illness, studying criminology) serves as a form of rhetorical device aimed at defending their legal position (e.g., guilty or NCRMD), and at swaying other jurors to side with them (see Heffer, 2005). For example, J70 expressed their personal experience studying criminology to establish their knowledge of the criminal justice system:

Yeah, um. Like I don't know a lot. But, like, I am doing, I am taking criminology at university. So, like, I just know that like, the whole like, the um, the treatment system in jail... It's not really... Good, like it doesn't really, like, a lot of what mentally, like insane people need. They need more treatment, but they are not given it. (J70)

Not all jurors were receptive to such rhetorical devices, however, which might be why some also redirected deliberation towards the jury instructions (e.g., burden of proof, sentencing choice). Jurors who referred to the law during deliberation may have trusted in the legitimacy of the law and may thus have seen referring to the law in jury decision-making as an objective and influential rhetorical device.

Meta-deliberation. Jurors also reflected on their role as jurors. Meta-deliberation occurred when jurors checked back into the group task at hand; for example, their legal responsibility was (not) determining the defendant's sentence. In Jury F's (as well as in other juries') deliberation, jurors noted that they were not in a position to request a sentence for the defendant:

J6F: I don't know if it's our job to kind of like...

J4F: No, we don't decide what happens like –

J3F: The judge – You can make like a suggestion to the judge, but you can't be like, he's guilty and we want him to be like put here, so...

J4F: No, no, no, no, no.

J2F: No, but yeah, we can't decide the sentence.

Moreover, jurors also brought the task at hand back to the attention of other jurors when the deliberation seemed to take a tangential turn, or when it seemed as if jurors were extrapolating or making assumptions about case facts:

Well, the burden of proof in this case is enough for these criteria, so it could be used to influence future juries. I don't know if that is necessarily a factor towards our decision-making, or whether it should be a factor for our decision-making, because it is not one of the questions that we are legally here to answer. (J8L)

According to the procedural justice model of public compliance with the law, people who believe in the legitimacy of legal institutions are more likely to comply with legal guidelines (Tyler, 2017). Those who are most likely to utilise and be swayed by legal discourse may also have strong right-wing authoritarian attitudes. That is, they are more likely to submit to authorities, which they perceive as established and legitimate (e.g., the law); they are generally aggressive, especially towards those who are sanctioned by established authorities (e.g., defendants), and they have a strong tendency to adhere to social conventions that are perceived to be endorsed by society and authorities (Altemeyer & Altemeyer, 1996). According to Altemeyer and Hunsberger (1992), right-wing authoritarians tend to be highly punitive: they tend to recommend longer and harsher prison sentences for a wide variety of crimes than people who are not right-wing authoritarians (Altemeyer & Hunsberger, 1992). As such, jurors who mobilise legal discourse may do so because they personally are swayed by it and perceive it as the most legitimate argument.

I argue that the referencing of legal instructions is a means to sway jurors' verdicts, and that it is possible that jurors reference the law (at times erroneously, as in J8L's comments on the burden of proof) to lend credence to their arguments and perceptions of the case. In this sense, jurors utilise legal language as an instrument. That is, the use of legal language in jury decision-making is not a neutral act; it is a powerful tool that can influence others' perspectives of the case (see Stygall, 1994). According to Heffer (2005), the type of discourse found in jury trials is called legal-lay discourse. This discourse encompasses the language that jurors use to make sense of the case either via reconstruction of personal experiences or a detached logical analysis (Heffer, 2005).

Moreover, Smith (1993b) has argued that jurors often have abstracted understandings and representations of the law (e.g., reasonable doubt standard, definition of NCRMD) and may thus use 'naïve representations of legal concepts' as information about the law. Finkel's (1995) theory of common-sense justice proposes that jurors make well-reasoned arguments and judgments of what they believe to be just and fair based on their intuitive notions of the defendant and the law. Thus, common-sense justice is 'what ordinary people think the law ought to be' (Finkel, 1999, p. 669). Notwithstanding potentially naïve understandings of the law, the jurors in the present study appeared to present their interpretations of the law as truth, with the goal of influencing other jurors' verdicts. According to critical discourse analysis (CDA), it is important to account for the relationship between discourse and social power (van Dijk, 2013) when analysing discussions. In the context of jury deliberations, it is possible that jurors utilise legal discourse (erroneously or not) as a means to position themselves as an authority based on their familiarity with privileged information (e.g., legal instructions, legal prongs for NCRMD) (see van Dijk, 2013). As such, it may be the case that jurors utilise both lived experiences and legal discourse as rhetorical devices to influence other jurors' verdict decisions.

Insufficient evidence

This theme represents jurors' hesitance in finding the defendant NCRMD. Jurors expressed that there was 'not enough evidence' (e.g., not enough witnesses, experts, physical evidence, mental illness history) in the defendant's case. They often fell back on the fact that they could not indicate which specific evidence demonstrated the defendant's responsibility or mental illness. Thus, this theme is broad in the evidence jurors perceived as 'missing'.

More experts needed. Some jurors expressed a desire for more experts to testify in the trial so that they could obtain a sense of the reliability of the psychiatrist's testimony:

I felt like there definitely wasn't enough evidence to say that one hundred percent that she did have this mental illness. Umm...And it umm...To me it doesn't make a difference whether the wait time before seeing the doctor was two weeks or whatever, I feel like, umm...She probably should have seen, probably like at least another doctor just to get like a different opinion. (J9K)

While jurors may want more experts to testify during the trial, Levett and Kovera (2008) have found that opposing expert testimony can create scepticism in jurors about all the expert testimony and lead to more guilty verdicts. Moreover, jurors who must process greater amounts of information tend to use heuristics (e.g., stereotypes) and misinterpret evidence, potentially making their verdict all the more biased and difficult to reach (see Horowitz & Bordens, 2000; see also Horowitz, Bordens, Victor, Bourgeois, & ForsterLee, 2001; Tamborini, Huang, Mastro, & Nabashi-Nakahara, 2007).

Acquaintance testimony needed. Some jurors expressed a desire to hear testimony from more people in the defendant's life (also relating to the jurors' perceptions of the importance of the defendant's relationships in determining his or her mental health status).

J2H: I don't know if there's enough evidence. Like, they needed more witnesses, if they had a couple people that he knew, like from his life maybe.

J5H: Yeah, I would've liked more of a profile too, from people that know him, even people just saying, like, 'I thought he was pretty normal except like...' except people who would notice this kind of like, disorganisation...

Importantly, the concept of evidence sufficiency is addressed by the court, not the jury (James, 1961). Moreover, a case will not go to the jury unless it is deemed that there is sufficient evidence for jurors to deliberate. Nevertheless, the perception that the defence did not have enough evidence was a highly common comment and point of discussion among the jurors. Calder, Insko, and Yandell (1974) found that jurors perceived the amount of evidence that a party had as indicative of the strength of the party's case. Tanford and Penrod (1986) also found that jurors talked about evidence sufficiency in simulated deliberations. Interestingly, the jurors' discussion about evidence sufficiency in the present study seemed to indicate that they would be more confident in their verdict if there was more information. This notion seemed to apply to the defence's case, whereby jurors often justified a guilty verdict by indicating evidence insufficiency for an NCRMD verdict. This notion of 'more evidence is better' is incongruent with Heuer and Penrod's (1994) finding that the more information jurors had to process in deliberation, the lower their confidence in their verdict would be.

Verdict outcome

This theme centres on jurors' concerns about the outcomes for the defendant and for society when finding the defendant NCRMD or guilty. Much of the discussion focused on the potential treatment (i.e., medical) of the defendant either in prison or in a psychiatric hospital.

Incarceration focus. Jurors were generally interested in the defendant being incarcerated in some form, though their reasons varied. Certain jurors perceived the defendant as lying about

mental illness and as being dangerous, while others saw the defendant as mentally ill and needing treatment, but still potentially dangerous. J3F commented, 'Like, regardless if he's still gonna [sic] get some form of punishment, he probably doesn't want to live in a mental health facility, so he'll still get some form of punishment while he's getting something beneficial out of it.'

The outcome of the verdict (e.g., prison or psychiatric hospital) was a particularly salient topic to jurors across deliberations. According to Roberts and Golding (1991), jurors' construals of defendant responsibility and their attitudes towards the insanity defence are strong predictors of their final verdict decisions. However, this focus on sentencing is a legal inaccuracy and may indicate the jurors' misunderstanding of their role as jurors (to provide a verdict, not a sentence). Moreover, jurors' focus on incarceration (in prison or hospital) may indicate their concern for how a sentence affects society's safety.

Consequences of incarceration. Jurors repeatedly referred to the potential societal consequences of finding the defendant either NCRMD or guilty. Jurors' interests in punishment and incarceration are culturally relevant and significant. In Jury H, Juror 12 discussed the actual consequences of finding a defendant NCRMD. Specifically, this juror corrected the stereotype that defendants who are found NCRMD are immediately released back into society.

[He] uses paranoid schizophrenia as a defence. If you're looking for an easy out to not go to jail, because that means you're going to be hospitalised, you're going to be put on medication, you're going to have many visits to the psychiatrist probably for a very, very long time. You'd honestly be better off serving your time without having to take anti-psychotic meds when you don't actually have, you know a psychotic disorder. That would

really mess a person up, so, if it was something that, you know, where like, okay, ‘he had two weeks after this happened, he researched, he came up with a good defence, he’s a great actor’, that wouldn’t be the first thing I would pick if that were me. (J12H)

In Jury K, jurors discussed what options were the most humane and helpful to the defendant’s rehabilitation.

J7K: She's guilty, whether she's going through the criminal system to jail or whether she's going through the mental illness system.

J10K: The difference is if you commit a murder, you go to jail, you get usually a life sentence...

J5K: You might [get] life for a second degree...

J10K: Yeah it is more of a chance of relief and rehabilitation going through the psychiatric system, rather than the prison system.

According to Roberts and Hough’s (2005) *penal escalation hypothesis*, people who perceive prisons as not unpleasant or harsh enough will call for more punitive laws and stricter treatment of people who break the law. Wozniak (2016) provides support for this theory: specifically, people’s perceptions of prison’s harshness on offenders relates to their attitudes about punishment, police, crime prevention, and prisoner re-entry programmes. In the present study, these perceptions about prison and desire for punishment seemed particularly pertinent

considering the jurors' strong focus on the conditions of the prison and/or psychiatric hospital when determining their verdict for the defendant.

The fact that jurors in the present study viewed the psychiatric hospital as an 'easy out' suggests their limited knowledge of the workings of psychiatric hospitals and prisons. The perception that prisons and psychiatric hospitals are too lenient and easy on prisoners may be due to most people's limited direct contact and personal experiences with correctional institutions and prison conditions (Wozniak, 2016). I argue that the public's generally punitive attitudes towards NCRMD defendants can be attributed to their general acceptance of insanity myths (e.g., NCRMD defendants are immediately released into society; Perlin, 2016) and their fear of mentally ill people (see Angermeyer & Dietrich, 2006). Such attitudes are also propagated by public officials in Canada, such as Ontario's Premier Doug Ford. In response to the disappearance of a client who was found NCRMD and permitted to leave the premises of Toronto's Centre for Addiction and Mental Health on a day pass, Ford (Goodfield, 2019) questioned how a 'nutcase' who was 'crazy, crazy' could leave the campus:

There are certain things that are a no-brainer. I could ask a 10-year-old kid should you let this guy loose on the street after he chopped up his roommate and a 10-year-old kid would say 'absolutely not'. – Ontario Premier Doug Ford (D'Mello, 2019)

Such attitudes and reactions towards the release of defendants found NCRMD seem to mirror and substantiate negative public perceptions of NCRMD defendants, mentally ill people, and the harshness of the criminal justice system.

Discussion

In general, jurors were sceptical of the defendant's claim of mental illness and its relationship to his or her act. This scepticism is congruent with the literature on the public's attitudes towards mental illness more broadly (Angermeyer & Matschinger, 2003; Martin, Pescosolido, & Tuch, 2000). Namely, the public is largely suspicious and untrusting of mentally ill people. It may be that jurors are afraid of these people and of the seeming randomness of mental illness. These factors may be threatening to jurors who perceive the act of the defendant as dangerous and erratic. In this study, jurors' elicitation of the defendant's relationships and the psychiatrist's credibility were potentially aimed at trying to understand how to control mentally ill people and their potential for violence. Jurors expected (or at least wanted to believe) that mental illness would be gradual and perceptible long before resulting in violence. It is possible that the jurors felt a loss of control in their ability to keep safe from dangerous mentally ill people, also leading them to question the legitimacy of the NCRMD defence.

Jurors' focus on the NCRMD defence was best characterised by their repeated expression that there was a 'lack of evidence' for the defendant's mental illness and their concerns about the verdict outcomes (i.e., they assumed that the defendant would immediately be released into society if found NCRMD). Jurors' suspicion of the NCRMD defence may thus have been an extension of their fear of mentally ill defendants. That is, jurors indicated being reluctant to find the defendant NCRMD because of 'insufficient evidence' and potential 'verdict outcomes', while potentially meaning that they feared that there would be a violent mentally ill person in society.

Importantly, jurors utilised two types of rhetoric to make the above points: 'legal' discourse and 'lay'/personal experience discourse. These rhetorical devices were possibly both aimed at characterising themselves as experts on the case evidence, so as to sway other jurors'

verdicts. Jurors' receptiveness to these rhetoric devices was varied and may have been contingent upon their own right-wing authoritarian or procedural justice attitudes, and/or their belief in lived experiences.

Implications

Overall, the identified themes suggest that the most important factor in jury deliberation in NCRMD trials is not jurors' gender stereotypes, but rather their broad stereotypes about mental illness and the insanity defence. Regardless of the defendants' gender (see Study 5), jurors discussed their mental illness status and plea first and foremost. This does not mean that the defendants' gender was not a contributing factor in jurors' final verdicts, but rather that gender stereotypes were not explicitly and prominently expressed in deliberations. Collectively, these themes signify that the defendant's mental health and the legitimacy of the NCRMD verdict are the most salient themes in NCRMD trial deliberations. From these, we can glean that in NCRMD deliberations, jurors not only question the defendant's mental health (e.g., defendant faking, temporality and rationality of mental illness) but also consider the legitimacy of the NCRMD standard and verdict (e.g., insufficient evidence, psychiatrist credibility).

These findings may imply that jurors are far more willing to explicitly express ideas on mental illness and the NCRMD law than gender stereotypes and affects in NCRMD deliberations. This is logical considering the task at hand for jurors. Nevertheless, jurors' interpretations of the law and of the defendant's mental illness may be biased. Jurors' characterisations of the defendant are at times incongruent with the evidence with which they are presented.

Limitations

The aim of Study 4 was to provide an in-depth perspective and understanding of the content of juror discussions. Following a qualitative thematic analysis, I examined the most prevalent themes and ideas present across the 10 deliberations. While thematic analysis is a useful and flexible method, its flexibility can lead to inconsistency and issues with consistency in theme development (Holloway & Todres, 2003; Nowell et al., 2017). Within the present study, these issues were addressed by applying a rigorous step-by-step programme of analysis (Braun & Clarke, 2006), with particular attention paid to establishing the trustworthiness of the results.

Furthermore, research validity is a particular challenge in qualitative research, because the data are based on analysis methods that are subject to researcher bias (see Olson, McAllister, Grinnell, Gehrke Walters, & Appunn, 2016). Triangulation is one of the most widely used methods to increase the validity of qualitative research (Lincoln & Guba, 1985). This process involves using multiple data points (events, times, locations, participants), researchers, and analysis methods to generate findings (Olson et al., 2016), allowing researchers to see if the results of the different sources produce similar results. If so, the validity of the findings is strengthened. Due to the time and funding restrictions of the present study, it would have been particularly difficult (and unlikely) to reach the exact same codes for each of the data segments. Nevertheless, I analysed jury decision-making in NCRMD trials from a variety of perspectives (e.g., several juries and different methodologies), which demonstrated similar results (e.g., no difference in jurors' gender roles or gender stereotype language and affect towards the defendants). Hence, the thematic analysis findings of the present study are likely trustworthy.

Study 5: Quantitative content analysis of theme relevance

Study 5 concerned the examination of the relationship between juror and defendant genders on deliberation discussion themes. To this end, I conducted two *z*-tests on theme relevance and defendant/juror gender.

Defendant gender's relation to themes

In general, all seven codes were relevant to the man and woman defendants' case deliberations; there were few differences between them in how often the components of the themes were raised. The table below (Table 6) provides an overview of the number of theme-relevant coded quotations for the man defendant deliberations in comparison to the woman defendant deliberations.

A noted gender difference was found in the theme of 'insufficient evidence'. Specifically, this theme was more likely to emerge from deliberations for the woman defendant than those for the man defendant; this was a statistically significant difference, $z = -3.39$, $p = .0007$. The remaining themes were not discussed at significantly different frequencies for the man and woman defendants. These analyses were calculated by aggregating the number of theme-related codes. Moreover, I calculated the proportion of these codes to the total number of theme related codes, and lastly I compared these theme proportions for the woman defendant versus man defendant cases (similarly comparing these code counts across the woman and man juror cases).

Table 7 *Theme and defendant gender*

Theme	<i>z</i>	<i>p</i>
Defendant faking mental illness	-0.035	0.97
Temporality and rationality of mental illness	1.94	0.05
Defendant-victim relationship	-0.27	0.79
Psychiatrist credibility	-0.25	0.80

Legal-lay discourse	0.18	0.86
Insufficient evidence	-3.39	0.0007**
Verdict outcome	-1.43	0.15

** $p < 0.01$

Juror gender's relation to themes

I also conducted a z -test for each of the individual themes to compare the themes that female and male jurors discussed. The results demonstrated that women jurors were more likely to consider the defendant 'faking' and the outcome of the verdict than were men jurors, $z = -2.43$, $p = .0015$. Moreover, women jurors were also more likely than men jurors to talk about verdict outcomes, $z = -2.67$, $p = .008$. There were no significant differences in the frequency with which the two genders discussed the amount of evidence, the deliberation, the mental illness focus, the defendant's relationships, the psychiatrist's credibility, or the verdict outcome (see Table 7).

Table 8 *Theme and juror gender*

Theme	z	p
Defendant faking mental illness	-2.43	0.015*
Temporality and rationality of mental illness	0.91	0.36
Defendant-victim relationship	-0.36	0.72
Psychiatrist credibility	0.26	0.79
Legal-lay discourse	-1.36	0.17
Insufficient evidence	-1.32	0.19
Verdict outcome	-2.67	0.008**

* $p < 0.05$

** $p < 0.01$

Discussion

Jurors expressed worries about ‘insufficient evidence’ more often when the defendant was a woman than when the defendant was a man. This theme specifically contains juror concerns about the number of witnesses (wanting more witnesses), perceptions of there being insufficient evidence about mental illness, the need for more expert testimony, and jurors’ view that there generally was not enough evidence to find the defendant NCRMD.

Jurors’ expression of this theme during female defendant deliberations may indicate that they were more ambivalent in their attitudes towards women defendants in insanity trials than they were towards men. Similarly, perhaps jurors were more certain in their attitudes and thus verdicts when the defendant was a man, and they were thus not as concerned about the amount of evidence. According to Maeder and Dempsey (2012), defendant gender influences jurors’ certainty of defendant guilt via their perceptions of the defendant’s femininity, credibility, and likability. Hence, in the present study, jurors’ desire for more evidence in the woman defendant (in comparison to the man defendant) cases may indicate that they were experiencing more uncertainty about the appropriate verdict for female defendants. Specifically, jurors might have been seeking information on the woman defendant’s character and personality.

Furthermore, juror gender was related to which themes were perceived as most relevant. Women jurors were more likely to discuss the defendant ‘faking’ mental illness and the verdict outcomes theme than men jurors were. These foci may have been the result of women jurors’ greater reliance on insanity myths about defendants pleading NCRMD. Specifically, women jurors may have feared violence from a mentally ill defendant who would be re-released into society. This fear makes sense when considering that women’s risk of violent victimisation is 20% higher than men’s (Perreault, 2015).

Implications

These results generally demonstrate that there is little significant difference between the themes that are evoked in deliberations for male and female defendants, or between what women and men jurors discuss in NCRMD deliberations. This is congruent with other findings in the present study, which demonstrated that jurors use few to no expressions of gender stereotypes in NCRMD deliberations. Thus, the results indicate that when it comes to the insanity defence, defendant and juror gender have a marginal impact on verdicts and deliberations. Importantly, across all deliberations, themes surrounding sexism and anti-NCRMD attitudes predominated over gender stereotypes.

Limitations

Ultimately, the results may imply that juror and defendant gender has marginal influence on jurors' focus on the defendant's mental illness and the NCRMD defence. These modest and null effects may be due to the small sample size of the present study and the overlap in all jurors' topics of discussion.

General Discussion

'Is not this insanity plea becoming rather common? Is it not so common that the reader confidently expects to see it offered in every criminal case that comes before the courts?... Really, what we want now, is not laws against crime, but a law against insanity. There is where the true evil lies.'

Mark Twain, *A New Crime—Legislation Needed* (1870)

This dissertation has explored the role of defendant and juror gender in mock jurors' verdict decisions and deliberation content in a fabricated NCRMD case involving paranoid schizophrenia. Study 1 entailed a multilevel model analysis of jurors' stereotype and affect

utterances, aimed at identifying jurors' gender biases. Study 2 analysed the relationship between gender and verdict. Next, Study 3 more closely examined jurors' deliberations through a content analysis of keywords and concepts (e.g., talkativeness, interruptions, sentence tags) and an analysis of how these relate to defendant and juror gender. Study 4 consisted of a qualitative thematic analysis of jurors' discussion. Lastly, Study 5 assessed discussion theme and gender.

The introduction opened with a quotation from Charlotte Perkins-Gilman: 'There is no female mind. The brain is not an organ of sex. As well speak of a female liver.' My dissertation addressed whether jurors believe that men and women's psychology and criminality are fundamentally different in the context of NCRMD trials. Specifically, I explored whether jurors' gender stereotypes and emotions significantly vary as a function of defendant gender. According to Breheny et al. (2007), 'gender matters in the insanity defence'. My primary interest was examining the stereotypes and affect on which mock jurors rely in attempting to defend their positions. That is, I wanted to apply the SCM theory (Fiske et al., 2002; e.g., warmth and competence stereotypes) to the jury deliberation context. I also sought to determine whether the SCM related to behavioural outcomes (as per the tripartite model of attitudes; Breckler, 1984). In Studies 1 and 2, mock juror utterances featured virtually no mention of warmth or competence stereotype content or affect expression. Accordingly, I cannot presume to answer whether jurors rely on gender stereotypes when making legal decisions in NCRMD trials. However, the data do shed light on jurors' relative emphasis on mental illness and the legitimacy of the insanity defence in NCRMD trials (Studies 4 and 5).

The introduction also presented Devine's (2012) Director's Cut Model, an integrative model that identifies how jurors initially draw from both their own (e.g., race, gender, and trust in the legal system) and the defendant's (e.g., race, gender, and prior criminal record)

characteristics to make a verdict decision (Devine, 2012; Devine & Caughlin, 2014) and to construct a story about the trial that shapes that decision. The director's cut model proposes that jurors initially perceive the trial and verdict information through the lenses of their personal and defendant characteristics. Although this model was not the central focus of my studies, the results do support this theory. Namely, jurors did appear to rely on their pre-existing attitudes when constructing their verdict decisions. These preconceived notions of the NCRMD defence were found in the thematic analysis I conducted for Study 4. On the other hand, while men and women tend to have different attitudes about mental illness and the insanity defence (Mossière & Maeder, 2015), I did not find a significant effect of juror gender on jurors' verdicts; the lack of significant effects may be due to this study being underpowered. The null effects may also be due to jurors' general suspicion of defendants who plead NCRMD. Moreover, the lack of gender effect on verdicts may thus be an extension of jurors' focus on the defendant's mental health rather than the defendant's gender. This focus was observed in Study 4, where jurors expressed that they were reluctant to find the defendant NCRMD because of 'insufficient evidence' and potential 'sentencing outcomes'. Interestingly, men and women jurors deliberated in a similar manner in the NCRMD trials. This similarity may be due to the context and speaker roles in deliberations (see Leaper, Anderson, & Sanders, 1998), or to women and men being more psychologically similar than different (Hyde, 2005). Lastly, in Study 5, while there was a great deal of overlap in the thematic foci of jurors' deliberations for men and women defendants, jurors did tend to focus more on insufficient evidence when the defendant was a woman (compared to a man). This dissimilarity in deliberation may indicate jurors' greater uncertainty about the appropriate verdict for women defendants compared to the man defendant. Ultimately, these findings suggest that defendant and juror gender may have a marginal impact on verdicts

and deliberations in NCRMD trials. Moreover, across all deliberations, jurors seemed to be more concerned with the defendant's mental health than with the defendant's gender.

Implications

At the beginning of this study, I questioned why the public construals of Andrea Yates and Guy Turcotte, two people who pled insanity for killing their children, were so disparate: Yates was cast as a 'mother-gone-mad' (see Galanti, 2003), whereas the Turcotte case was publicly construed as evidence of the need for tough-on-crime policies and laws (Duval, 2016). I posited that these narratives may have been the result of public stereotypes about how gender relates to crime and mental disorders; I thought that the defendants' gender may have influenced how the public perceived their crimes and their pleas of insanity. The present programme of research has demonstrated a far less clear-cut answer.

Defendant gender and jury decision-making. I proposed that gender is relevant to all institutions, including the criminal justice system. Though I did not find the hypothesised gender effects (on verdicts, stereotype content language, and affect), I posit that gender (of the defendants and the jurors) is nevertheless important to jurors' decision-making processes. While there were no significant verdict differences in the present study between men and women defendants, there were gender differences in the themes discussed in deliberation (Study 5). Furthermore, jurors tended to focus on insufficient evidence of mental illness in the female defendants' cases (more than in the male defendants' cases; Study 5). This may suggest that jurors attribute more criminal responsibility to women defendants. Thus, this dissertation contributes to the jury decision-making literature in NCRMD trials by demonstrating the attitudes towards defendant gender in deliberation. This research may suggest that while gender

attitudes are not explicitly expressed during deliberation (e.g., SCM and affect words), these attitudes can impact jurors' deliberation narratives.

Attitudes in jury decision-making. In the present study, few of the attitudes that jurors demonstrated directly referenced stereotype content (SCM; Fiske et al., 2002). This finding may imply that while defendant gender could be a relevant factor contributing to jurors' verdicts in other case types (e.g., sexual assault cases), in insanity trials jurors are overtly focused on the evidence surrounding the defendant's mental illness/health and the legitimacy of the NCRMD law. Specifically, jurors questioned whether the defendant was truly mentally ill and often sought more evidence of a history of his or her mental illness. They also questioned whether the NCRMD verdict was appropriate in the case (regardless of the defendant's gender). The proposed SCM had limited utility in interpreting the deliberations, since jurors did not tend to express exact SCM adjectives (this may be because the dictionary was not extensive enough, or because SCM is typically a survey measure). Moreover, because jurors employed SCM- and affect-related adjectives so infrequently, I did not find a significant relationship between SCM and affect. Descriptive measures such as semantic differentials (quantitative measure) may thus have limited applicability in group discussions comprising multiple persons' conflicting attitudes (see Eagly & Mladinic, 1989).

Jurors' legal inaccuracies and the misapplication of law. In NCRMD trials, the judge provides jurors with legal guidelines on the standard for insanity prior to deliberation (as was done in the present study). Jurors are then expected to accurately follow these guidelines and make an impartial legal judgement (Finkel, 1995; see Finkel & Handel, 1989; Loudon & Skeem, 2007). In the present study, jurors seemed to have limited comprehension of and adherence to the legal guidelines for insanity and to predominantly focus on their own conceptions of the

legitimacy of the insanity defence and of the defendants' pleas. Smith (1993a) has suggested that jurors tend to rely on inaccurate folk conceptions of the insanity defence and stereotypes about insanity defendants. The present study supports this notion by demonstrating that jurors broadly rely on their own and other jurors' sense of expertise on mental illness and law (e.g., lay-legal discourses). Importantly, research on jurors' reliance on their own perceptions of what the law means or what it ought to mean has demonstrated that this dependence on the self can lead to misapplications of the law or to jury nullification (see Weinstein, 1992; McKnight, 2013).

Nullification occurs when jurors decline to follow the law presented to them by a judge and instead follow their own sense of fairness and of the law, even though a case would seem to allow for no other verdict than guilty (or NCRMD). That is, they choose not to apply the law that they are prescribed by the judge (due to perceiving it as unjust) and instead supplant the judge's instructions with their own sense of the law (see Weinstein, 1992). This hidden process (i.e., jurors are not instructed on this option and it is therefore often unknown to them) is a lawful one based in the autonomy and legitimacy of the jury as a legal decision-making body (see Weinstein, 1992). McKnight (2013) has argued that jury nullification 'serves as a check on government power by adding a level of discretionary review by allowing common human experience to temper oft-times rigid application of the law' (p. 1104). Finkel (1990) has noted that insanity trials are noteworthy instances where jurors depart from the law (*de jure*), and that there is limited understanding of *why* this occurs in insanity trials. Moreover, Finkel (1990) has emphasised the importance of empirically examining how jurors construct and understand insanity and therefore accept or deny the insanity plea. This dissertation contributes to the literature by demonstrating that jurors largely construct the insanity defence as a legal loophole

and doubt the legitimacy of the defendants' claims of NCRMD, regardless of the defendant's gender.

Juror gender and communication. According to Marder (1987), women's participation rate in group discussions is lower than men's, and women need to speak more and men need to listen more for there to be equal and fair participation. The present dissertation does not support this position: women and men mock jurors seemed to participate and listen (e.g., not interrupt) at similar rates in jury deliberation. Researchers have found that while there is an expectation that men and women will communicate differently, they do not significantly vary in the types of language content and communication styles they use (Garrett, Meyers, & West, 1997). This dissertation contributes to the divisive literature on gender and communication style by demonstrating that there are more similarities than differences in the manner in which women and men participate and communicate in jury deliberations.

Mixed-methods design. While I had originally proposed to conduct a quantitative analysis of jury deliberations, I decided to also conduct a qualitative thematic analysis of the jury deliberations to understand why jurors expressed so few SCM and affective words. The use of mixed methods in this dissertation allowed for a holistic examination of how jurors perceive men and women who plead NCRMD. De Lisle (2011) has argued that qualitative findings can add value to quantitative results by contributing new and conflicting perspectives on the latter. The qualitative findings in the present study demonstrated that the jurors did not tend to use gender-specific SCM or affective language, potentially because they were focused on the mental illness and legitimacy of the NCRMD plea across all deliberations, regardless of the defendant's gender.

Historically, jury decision-making research has studied mock jurors at the individual level. The present study is amongst a growing body of work in jury decision-making which

integrates group deliberation as a component of analysis (Lynch & Haney, 2009; Peter-Hagene, 2019; Yamamoto, 2019). The aim of including a deliberation component is to ensure the verisimilitude of jury decision-making in a real court. Many of the studies on jury decision-making and NCRMD attitudes and verdict have utilised quantitative methodologies consisting of an individually administered and completed trial transcript, and attitude and verdict questionnaires. This dissertation extends the jury decision-making research paradigm by integrating a deliberation component. This paradigm is increasingly becoming the ideal in jury decision-making research, as it most closely simulates jury decision-making. It allows for rich data that can be analysed quantitatively and qualitatively (e.g., using mixed methods, see Johnson, Onwuegbuzie, & Turner, 2007).

While not all researchers will be receptive to mixed-methods research (e.g. Giddings & Grant, 2007; Giddings, 2006; see also De Lisle, 2011), I see many benefits to it in the context of jury decision-making deliberation research. Specifically, quantitative methods can allow researchers to test for group differences in verdicts and to analyse juror demographics and decision-making processes and outcomes, while qualitative methods can enable an analysis of the most salient ideas and attitudes that occur in jury discussion (e.g., ideas that may not be congruent with posited theory and hypotheses). According to McKim (2017), it is important to ask whether mixed methods will add more value to the research design than a single method could. Hurmerinta-Peltomäki and Nummela (2006) found that the use of mixed methods adds value by increasing validity and assisting in knowledge creation. Moreover, it allows a deeper, broader understanding of the topic of interest than studies using a single methodology (McKim, 2017). According to Creswell and Plano Clark (2007), choosing mixed methods combines the strengths of each methodology and minimises their weaknesses.

It is also important to consider the added resources, time, and expertise needed when conducting mixed-method research. For example, this research requires extra time to collect and analyse the sets and types of data (Creswell & Plano Clark, 2011), as well as knowledge of a variety of research methods. Ultimately, because jurors' discussions are often spontaneous and may not closely conform to quantitative questionnaire materials (as was seen in the present study), I see a great deal of added value in combining the traditional jury paradigm of individual questionnaires and verdict outcomes with jury discussion utilising both quantitative and qualitative analyses of these data.

My dissertation contributes to the psycho-legal literature on jury decision-making in insanity trials by illustrating mock jurors' perceptions and attitudes towards the defendants' cases and pleas of NCRMD. Whereas previous research on public attitudes towards the insanity defence has significantly demonstrated the generalised distrust towards the insanity defence (and towards mentally ill people; Perlin, 1996; 2016), the present study confirms that these attitudes are an important component of jurors' decision-making processes in deliberation. Specifically, my dissertation demonstrates that jurors are focused on and question the defendant's mental health and the applicability of the insanity defence, regardless of the defendant's gender.

Based on these findings, it may be of particular importance for lawyers and psychiatrists to persuade jurors of the defendant's history and progression of mental illness. Results from the thematic analysis in this dissertation underscore seven important areas of focus for jurors in insanity trials. In general, these themes imply that the 'madness' label trumps any mention and impact of gender stereotypes in jury deliberation. Moreover, jurors seemed to be concerned for the safety of society, particularly when they considered the possibility of the defendant being found NCRMD (i.e., perception that the defendant would be released into society and harm

again). Thus, jurors' reluctance to find the defendant NCRMD in many cases may indicate their reliance on sanist biases and insanity stereotypes, and their interest in risk-mitigation via a guilty verdict. One possible method to remedy jurors' tendency to depend on biases regarding mental illness in NCRMD trials is to make these biases salient to them (via education on biases' effects on attitudes; see Rudman, Ashmore, & Gary, 2001). To limit jurors' reliance on stereotypes, defence lawyers and psychiatrists (through opening/closing statements and expert testimony) should also consider emphasising the sanist stereotypes and insanity myths that jurors tend to hold against mentally ill defendants.

Limitations

There are general limitations to the present studies that warrant mentioning and consideration. Firstly, the exploratory nature and modest sample size of these studies temper the ability to make certain statements on the impact of juror and defendant gender on jury decision-making in NCRMD trials. The sample size also restricts the possibility to make significant quantitative and empirical statements on the influence of defendant and juror gender on jury verdicts. I conducted a post-hoc power analysis using *G*Power* (Erdfelder, Faul, & Buchner, 1996) to check whether my non-significant results were due to a lack of statistical power. The sample size required for the recommended $\pi = 0.80$ (see Cohen, 1988) and a moderate effect size of $d = 0.50$ (a large effect size according to Cohen's (1977) effect size conventions) was 128 participants. Thus, the 77-participant sample in the present study was generally too small (post-hoc observed $\pi < 0.50$ across the studies) for the recommended power and effect sizes. As such, the majority of the quantitative results concerning the impact of gender on jury decision-making were null, and there are potential Type II errors in the results. Future studies interested in

individual- and group-level impacts of jury deliberation on verdicts should utilise a larger sample size.

On the other hand, another aim of this study was to produce a rich and exploratory (rather than a generalisable) examination of a novel area of research, for which the smaller sample size was appropriate and fitting (see Sandelowski, 1995; Bowen, 2008). The qualitative analysis in Study 4 provided a rigorous, descriptive, and interpretive overview of the information that was most important and salient to jurors across NCRMD trials. Moreover, this study has met the milestone of saturation; by monitoring each of the five deliberations (for the woman defendant) there did not appear to be any additional information emerging. It is important to note that this method of monitoring saturation involves a degree of subjective assessment (Fusch & Ness, 2015). Nevertheless, my finding of saturation was consistent with what Guest et al. (2017) found when empirically testing saturation points in focus groups (i.e., 90% of theme information was discoverable within three to six focus groups).

The qualitative analysis in the present study allowed for an examination of the subtleties and complexities of jury deliberation, revealing information that would have been missed had I solely relied on my quantitative enquiries. However, like all approaches to research, there are undoubtable limitations to applying a qualitative thematic analysis. The most prominent is that the quality of the analysis depends on the skills of the researcher, and the analysis is more easily influenced by the researcher's biases (Anderson, 2010). Nonetheless, I trust that the results of the thematic analysis provide a powerful and compelling story of *how* and *what* jurors discuss during NCRMD deliberations. Despite the ongoing 'paradigm wars' between qualitative and quantitative research (Kelle, 2006), I am confident that the present study represents a convincing argument for the scientific benefits of using mixed methods when studying jury decision-

making. Combining these methods can help counteract their individual limitations and result in a strengthened coherent and holistic representation of jurors' attitudes in insanity trials (see Kelle, 2006).

Lastly, it is also worth noting the limitations of a simulated jury study. Mainly, the reduced gravity of the case (i.e., the trial did not have real world implications and outcomes), the lesser complexity of the trial facts, the single and limited deliberation session (45 minutes), not requiring a unanimous verdict, and the physical space of the deliberation (i.e., observation room with two-way mirror, microphones and cameras) resulting from the experimental setting. Nevertheless, these are important factors that limit (as much as possible) the impact of extraneous factors on jurors decision-making processes (see Lynch & Haney, 2011; see also Peter-Hagene, 2019)

Future directions

While my studies aimed to examine the intersectional representations of gender and mental illness for people who plead NCRMD, the words that jurors employed did not fit the stereotype content theory's (Fiske et al., 2002) taxonomy of gender stereotype and emotions. Jurors may not generally express their attitudes using stereotype-congruent adjectives. After conducting this study, I argue that jurors' attitudes are better gleaned from the rich narrative and stories they construct and mobilise during deliberations. Interestingly, while each jury deliberation in this study had an idiosyncratic pathway to a final verdict, jurors' narratives about mental illness and the NCRMD defence were consistent across all deliberations, regardless of the defendants' and jurors' genders. From these studies, it appears that mental illness stereotypes may trump those about gender in jurors' deliberations about the NCRMD defendant. Future studies should examine the best way to test how and when the madness label trumps the effects

of gender, race, and ethnicity stereotypes in jury decision-making. Furthermore, switching the focus on one stereotype over others has potentially interesting effects on jurors' perceptions of the defendant. Future work could also aim to address how to parse the impact of these various labels and stereotypes.

Conclusion

This dissertation provided a glimpse into jurors' discussion foci in NCRMD trial deliberations for men and women defendants. Overall, I found that the defendants' gender may not be as (explicitly) salient in jurors' deliberations as the defendants' mental illness and their NCRMD plea. As such, these studies may suggest that gender does not matter (as much as mental illness) in the insanity trial. Moreover, these studies demonstrate the power of mixed methodology in elucidating which attitudes and stereotypes are mobilised by jurors in a group decision-making setting. This programme of research also highlights the limitations of solely relying on a single methodological paradigm when trying to understand the rich and complex processes of jury decision-making. According to the law of instrument, relying on a single method to examine an issue of interest can have considerable consequences. As the adage goes, 'Give a small boy a hammer, and he will find that everything he encounters needs pounding'. In other words, there are true limitations to only using quantitative (or qualitative) methods to address questions about jury decision-making, and great benefits to methodological pluralism (see Bauer, Gaskell, & Allum, 2000). I hope that the present study has demonstrated the rich and holistic scientific benefits of using a multi-method approach. For, if I had not incorporated a thematic analysis, I would not have been able to determine whether jurors perceive women and men who plead insanity as mad and/or bad. It seems that jurors may view these defendants as both mad *and* bad.

References

- Abrams, D., Viki, G. T., Masser, B., & Bohner, G. (2003). Perceptions of stranger and acquaintance rape: The role of benevolent and hostile sexism in victim blame and rape proclivity. *Journal of Personality and Social Psychology, 84*(1), 111-125.
- Allen, H. (1987). *Justice Unbalanced: Gender, Psychiatry and Judicial Decisions*, Milton Keynes:
- Allnutt, S., Samuels, A., & O'Driscoll, C. (2007). The insanity defence: from wild beasts to M'Naghten. *Australasian Psychiatry, 15*(4), 292-298.
- Altemeyer, B., & Hunsberger, B. (1992). Authoritarianism, religious fundamentalism, quest, and prejudice. *The International Journal for The Psychology of Religion, 2*(2), 113-133.
- Altemeyer, R. A., & Altemeyer, B. (1996). *The Authoritarian Specter*. Harvard University Press.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- Anderson, C. (2010). Presenting and evaluating qualitative research. *American Journal of Pharmaceutical Education, 74*(8), 141.
- Angemeyer, M.C., Matschinger, H. (2003). Public beliefs about schizophrenic and depression: similarities and differences. *Social Psychiatry and Psychiatric Epidemiology, 38*, 526-534.
- Angermeyer, M. C., & Dietrich, S. (2006). Public beliefs about and attitudes towards people with mental illness: a review of population studies. *Acta Psychiatrica Scandinavica, 113*(3), 163-179.

- Angermeyer, M. C., & Matschinger, H. (2003). The stigma of mental illness: Effects of labelling on public attitudes towards people with mental disorder. *Acta Psychiatrica Scandinavica*, *108*(4), 304-309.
- Angermeyer, M. C., Buyantugs, L., Kenzine, D. V., & Matschinger, H. (2004). Effects of labelling on public attitudes towards people with schizophrenia: are there cultural differences?. *Acta Psychiatrica Scandinavica*, *109*(6), 420-425.
- Appignanesi, L. (2011). *Mad, bad and sad: A history of women and the mind doctors from 1800 to the present*. Hachette UK.
- Astbury, J. (2001) Gender disparities in mental health. In: *Mental health. Ministerial Round Tables 2001, 54th World Health Assemble*. World Health Organization, Geneva, Switzerland.
- Barbalet, J. (2002). Introduction: Why emotions are crucial. *The Sociological Review*, *50*(2), 1-9.
- Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, *20*(5), 286-290.
- Bauer, M. W., Gaskell, G., & Allum, N. C. (2000). Quality, quantity and knowledge interests: Avoiding confusions. *Qualitative Researching with Text, Image and Sound: A Practical Handbook*, 3-17.
- Beckham, C. M., Spray, B. J., & Pietz, C. A. (2007). Jurors' locus of control and defendants' attractiveness in death penalty sentencing. *The Journal of Social Psychology*, *147*(3), 285-298.
- Bengtsson, M. How to plan and perform a qualitative study using content analysis. *NursingPlus Open* 2, 8-14.

- Berelson, B. (1952). *Content analysis in communication research*. New York, NY, US: Free Press.
- Berger, P. L., & Luckman, T. 1966. *The Social Construction of Reality*. London: Penguin Press.
- Bergsieker, H. B., Leslie, L. M., Constantine, V. S., & Fiske, S. T. (2012). Stereotyping by omission: eliminate the negative, accentuate the positive. *Journal of Personality and Social Psychology, 102*(6), 1214.
- Billig, M. (1976). *Social psychology and intergroup relations*. London, Academic Press.
- Blais, J., & Forth, A. E. (2014). Prosecution-retained versus court-appointed experts: Comparing and contrasting risk assessment reports in preventative detention hearings. *Law and Human Behavior, 38*(6), 531.
- Bloechl, A. L., Vitacco, M. J., Neumann, C. S., & Erickson, S. E. (2007). An empirical investigation of insanity defense attitudes: Exploring factors related to bias. *International Journal of Law and Psychiatry, 30*(2), 153-161.
- Boccaccini, M. T., & Brodsky, S. L. (2002). Believability of expert and lay witnesses: Implications for trial consultation. *Professional Psychology: Research and Practice, 33*(4), 384.
- Bongiorno, R., Bain, P. G., & David, B. (2014). If you're going to be a leader, at least act like it! Prejudice towards women who are tentative in leader roles. *British Journal of Social Psychology, 53*(2), 217-234.
- Bornstein, B. H., & Greene, E. (2017). *The jury under fire: Myth, controversy, and reform*. Oxford University Press.

- Bornstein, B. H., & Muller, S. L. (2001). The credibility of recovered memory testimony: Exploring the effects of alleged victim and perpetrator gender. *Child Abuse & Neglect*, 25(11), 1415-1426.
- Bornstein, B. H., Golding, J. M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior*, 41(1), 13.
- Bortolotti, L. (2013). Rationality and sanity: the role of rationality judgements in understanding psychiatric disorders. *The Oxford Handbook of Philosophy and Psychiatry*. Oxford University Press, Oxford, 480-496.
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qualitative Research*, 8(1), 137-152.
- Boysen, G., Ebersole, A., Casner, R., Nykhala, C. (2014). Gendered Mental Disorders: Masculine and Feminine Stereotypes about Mental Disorders and Their Relation to Stigma. *The Journal of Social Psychology*, 154, 546-565.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Breckler, S.J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology* 47:1191–1205.
- Breheney, C., Groscup, J., & Galietta, M. (2007). Gender matters in the insanity defense. *Law & Psychology Review*, 31, 93.
- Broeder, D. W. (1959). The University of Chicago jury project. *Nebraska Law Review*, 38, 744.

- Brunell, T. L., Dave, C., & Morgan, N. C. (2009). Factors affecting the length of time a jury deliberates: Case characteristics and jury composition. *Review of Law & Economics*, 5(1), 555-578.
- Burt, C. H., Simons, R. L., & Gibbons, F. X. (2012). Racial discrimination, ethnic-racial socialization, and crime: A micro-sociological model of risk and resilience. *American Sociological Review*, 77(4), 648-677.
- Butler, B. (2006). NGRI Revisited: Venirepersons' Attitudes Toward the Insanity Defense 1. *Journal of Applied Social Psychology*, 36(8), 1833-1847.
- Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. Routledge: New York, NY.
- Calder, B. J., Insko, C. A., & Yandell, B. (1974). The relation of cognitive and memorial processes to persuasion in a simulated jury trial. *Journal of Applied Social Psychology*, 4(1), 62-93.
- Carli, L. L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology*, 59(5), 941.
- CBC News. (2015). Guy Turcotte found guilty of 2nd-degree murder in deaths of his 2 children. Retrieved from: <http://www.cbc.ca/news/canada/montreal/guy-turcotte-trial-verdict-1.3330322>
- Cecil, Joe S.; Hans, Valerie P.; and Wiggins, Elizabeth C., "Citizen Comprehension of Difficult Issues: Lessons from Civil Jury Trials" (1991). *Cornell Law Faculty Publications*. 414.
- Centre for Mental Health and Addiction. (2013). *Mental Health and Criminal Justice Policy Framework* (pp. 1-18). Toronto.

- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752.
- Charette, Y., Crocker, A., Seto, M., Salem, L., Nicholls, T., & Caulet, M. (2015). The National Trajectory Project of Individuals Found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 4: Criminal Recidivism. *The Canadian Journal of Psychiatry*, 60(3), 127-134.
- Chesler, P. (2005). *Women and madness*. Palgrave Macmillan.
- Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The journal of the learning sciences*, 6(3), 271-315.
- Clark, N. K., Stephenson, G. M., & Kniveton, B. H. (1990). Social remembering: Quantitative aspects individual and collaborative remembering police officers and students. *British Journal of Psychology*, 81(1), 73-94.
- Clarke, V., & Braun, V. (2018). Using thematic analysis in counselling and psychotherapy research: A critical reflection. *Counselling and Psychotherapy Research*, 18(2), 107-110.
- Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic
- Cohen, J. (1977). *Statistical Power Analysis for The Behavioral Sciences*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Cole, F. L. (1988). Content analysis: process and application. *Clinical Nurse Specialist*, 2(1), 53-57.

- Commerson, S. (2003). Villain or Victim-Myths, Gender, and the Insanity Defense. *Developments in Mental Health Law*, 22(1).
- Coontz, P. D., Lidz, C. W., & Mulvey, E. P. (1994). Gender and the assessment of dangerousness in the psychiatric emergency room. *International Journal of Law and Psychiatry* 17(4), 369-376
- Cooper, J., & Neuhaus, I. M. (2000). The “hired gun” effect: Assessing the effect of pay, frequency of testifying, and credentials on the perception of expert testimony. *Law and Human Behavior*, 24(2), 149-171.
- Cooper, J., Bennett, E. A., & Sukel, H. L. (1996). Complex scientific testimony: How do jurors make decisions?. *Law and Human Behavior*, 20(4), 379-394.
- Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review Psychology*, 57, 345-374.
- Crawford, A. (2015). Temporality in restorative justice: On time, timing and time-consciousness. *Theoretical Criminology*, 19(4), 470-490.
- Creswell, J. W. (2013). *Qualitative Inquiry & Research Design Choosing among Five Approaches* (3rd ed.). Thousand Oaks, CA SAGE
- Creswell, J., & Plano Clark, V. (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. (2nd ed.). Thousand Oaks, CA: Sage.
- Oaks, CA: Sage. *Criminal Code, R.S.C.* (1985). c. C-46

- Crocker, A., Charette, Y., Seto, M., Nicholls, T., Côté, G., & Caulet, M. (2015). The National Trajectory Project of Individuals Found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 3: Trajectories and Outcomes through the Forensic System. *The Canadian Journal Of Psychiatry*, 60(3), 117-126.
- Crocker, A., Nicholls, T., Seto, M., Charette, Y., Côté, G., & Caulet, M. (2015). The National Trajectory Project of Individuals Found Not Criminally Responsible on Account of Mental Disorder in Canada. Part 2: The People behind the Label. *The Canadian Journal Of Psychiatry*, 60(3), 106-116.
- Crocker, J., & Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of personality and social psychology*, 58(1), 60.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of Personality and social Psychology*, 92(4), 631.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40, 61-149.
- Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1989). The eyewitness, the expert psychologist, and the jury. *Law and Human Behavior*, 13(3), 311-332.
- Daudistel, H. C., Hosch, H. M., Holmes, M. D., & Graves, J. B. (1999). Effects of Defendant Ethnicity on Juries' Dispositions of Felony Cases 1. *Journal of Applied Social Psychology*, 29(2), 317-336.
- Dawson, M. (2004). Rethinking the boundaries of intimacy at the end of the century: The role of victim-defendant relationship in criminal justice decision-making over time. *Law & Society Review*, 38(1), 105-138.

- Denzin, N. K., & Lincoln, Y. S. (Eds.). (2011). *The Sage Handbook of Qualitative Research*. Sage.
- De Lisle, J. (2011). The benefits and challenges of mixing methods and methodologies: Lessons learnt from implementing qualitatively led mixed methods research designs in Trinidad and Tobago. *Caribbean Curriculum 18*, 87–120.
- DeSantis, L., & Ugarriza, D. N. (2000). The concept of theme as used in qualitative nursing research. *Western Journal of Nursing Research, 22*(3), 351-372.
- Devine, D. J. (2012). *Jury Decision-Making: The State of the Science*. New York and London: New York University Press.
- Devine, D. J., & Caughlin, D. E. (2014). Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments. *Psychology, Public Policy, and Law, 20*(2), 109-134.
- Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law, 7*(3), 622-727.
- D'Mello, C. (2019). Ontario government to join external review of CAMH amid recent disappearances. *CTV News Toronto*. Retrieved from:
<https://toronto.ctvnews.ca/mobile/ontario-government-to-join-external-review-of-camh-amid-recent-disappearances-1.4538343?cache=yes?clipId=375756>
- Dixon, J.A., and Foster, D.H. 1997. Gender and hedging: From sex differences to situated practice. *Journal of Psycholinguistic Research, 26*(1): 89–107.

- Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V. M. (2010). Prejudice, stereotyping, and discrimination: Theoretical and empirical overview. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *SAGE Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 3–28). London: Sage.
- Dunlap, E. E., Hodell, E. C., Golding, J. M., & Wasarhaley, N. E. (2012). Mock jurors' perception of stalking: The impact of gender and expressed fear. *Sex Roles, 66*(5-6), 405-417.
- Duval, M.-C. (2016). Comment les medias couvrent-ils les causes de justices?: L'affaire Guy Turcotte sous la loupe. Unpublished Master's Thesis.
- Eagly, A. H. (2004). Few women at the top: How role incongruity produces prejudice and the glass ceiling. In: D. van Knippenberg & M. A. Hogg (Eds), *Leadership and Power: Identity Processes in Groups and Organizations*. London: Sage
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25*(5), 582-602.
- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15*(4), 543-558.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46*(4), 735.
- Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin, 117*(1), 125.

- Eagly, A.H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review* 109(3), 573-598.
- Eaton, N., Keyes, K., Krueger, R., Balsis, S., Skodol, A., & Markon, K. et al. (2012). An invariant dimensional liability model of gender differences in mental disorder prevalence: Evidence from a national sample. *Journal Of Abnormal Psychology*, 121(1), 282-288.
- Educaloi.qc.ca (2015). Affaire Guy Turcotte: Le Directeur aux poursuites criminelles et pénales du Québec porte la décision en appel. Retrieved from:
<https://www.educaloi.qc.ca/nouvelles/affaire-guy-turcotte-le-directeur-aux-poursuites-criminelles-et-penales-du-quebec-porte-la>
- Ellemers, N., & Haslam, S. A. (2011). Social identity theory. *Handbook of Theories of Social Psychology*, 2, 379-398.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107-115.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28(1), 1-11.
- Eriksen, K., & Kress, V. E. (2008). Gender and diagnosis: Struggles and suggestions for counselors. *Journal of Counseling & Development*, 86(2), 152-162.
- Finkel, N. J. (1990). De facto departures from insanity instructions. *Law and Human Behavior*, 14(2), 105-122.
- Finkel, N. J. (1995). *Commonsense Justice: Jurors' Notions of the Law*. Cambridge: Harvard University
- Finkel, N. J. (1999). Commonsense justice, culpability, and punishment. *Hofstra Law Review*, 28, 1-706.

- Finkel, N., & Handel, S. (1989). How jurors construe "insanity." *Law and Human Behavior, 13*(1), 41-59.
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Fiske, S. T. (2012). Journey to the edges: Social structures and neural maps of inter-group processes. *British Journal of Social Psychology, 51*(1), 1-12.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In *Advances in Experimental Social Psychology* (Vol. 23, pp. 1-74). Academic Press.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology 82*(6), 878-902.
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.
- Ford, M. C. (1986). The role of extralegal factors in jury verdicts. *The Justice System Journal, 16*-39.
- Fusch, P., & Ness, L. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report, 20*, 1208–1416.
- Galanti, M. (2003). The Andrea Yates trial: What is wrong with this picture. *Cardozo Women's Law Journal, 9*(2), 345-360.
- Garrett, D., Meyers, R., & West, L. (1997). Sex differences and consumer complaints: do men and women communicate differently when they complain to customer service

- representatives?. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 10.
- Gemberling, T. M., & Cramer, R. J. (2014). Expert testimony on sensitive myth-ridden topics: Ethics and recommendations for psychological professionals. *Professional Psychology: Research and Practice*, 45(2), 120.
- Giddings, L. S. (2006). Mixed-methods research: Positivism dressed in drag?. *Journal of Research in Nursing*, 11(3), 195-203.
- Giddings, L. S., & Grant, B. M. (2007). A Trojan horse for positivism?: A critique of mixed methods research. *Advances in Nursing Science*, 30(1), 52-60.
- Gerber, R. J. (1984). *The Insanity Defense*. Port Washington, New York: Associated Press.
- Gillespie, L. K., Loughran, T. A., Smith, M. D., Fogel, S. J., & Bjerregaard, B. (2014). Exploring the role of victim sex, victim conduct, and victim–defendant relationship in capital punishment sentencing. *Homicide Studies*, 18(2), 175-195.
- Glancy, G. D., & Bradford, J. M. (2007). The admissibility of expert evidence in Canada. *The Journal of the American Academy of Psychiatry and the Law*, 35(3), 350-356.
- Glenn Wiley, M., & Eskilson, A. (1985). Speech style, gender stereotypes, and corporate success: What if women talk more like men?. *Sex Roles*, 12(9-10), 993-1007.
- Glick, P., & Fiske, S. T. (1997). Hostile and benevolent sexism: Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly*, 21(1), 119-135.
- Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, 56(2), 109.
- Glick, P., & Fiske, S. T. (2011). Ambivalent sexism revisited. *Psychology of Women Quarterly*, 35(3), 530-535.

- Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., Adetoun, B., Osagie, J. E., Akande, A., Aloa, A., Annetje, B., Willemsen, T. M., Chipeta, K., Dardenne, B., Dijksterhuis, A., Wigboldus, D., Eckes, T., Six-Martena, I., Expósito, F., Moya, M., Foddy, M., Kim, H.-J., Lameiras, M., Sotelo, M. J., Mucchi-Faina, A., Romani, M., Sakalli, N., Udegbe, B., Yamamoto, M., Ui, M., Ferrira, M. C., & López, W., L. (2000). Beyond prejudice as simple antipathy: hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*(5), 763.
- Glick, P., Fiske, S. T., & Mladinic, A. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*, 763–775.
- Goldstein, A. M., Morse, S. J., Shapiro, D.L. & Weiner, I. B. (2003), *Handbook of Psychology: Forensic Psychology*. Hoboken, New Jersey, United States: John Wiley & Sons, Inc.
- Goodfield, K. (2019a). Review launched after Ford questions how ‘nutcase’ disappeared from CAMH. *CTV News Toronto*. Retrieved from: <https://toronto.ctvnews.ca/reviews-launched-after-ford-questions-how-nutcase-disappeared-from-camh-1.4513651>
- Government of Canada. (2006). *The Human Face of Mental Health and Mental Illness in Canada 2006* (pp. 1-203). Ottawa.
- Grabe, M. E., Trager, K. D., Lear, M., & Rauch, J. (2006). Gender in crime news: A case study test of the chivalry hypothesis. *Mass Communication and Society, 9*(2), 137-163.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029.

- Greene, E., Hayman, K., & Motyl, M. (2008). "Shouldn't we consider...?" Jury discussions of forbidden topics and effects on damage awards. *Psychology, Public Policy, and Law* 14(3), 194-222.
- Grigoriadis, S., & Robinson, G. (2007). Gender Issues in Depression. *Annals of Clinical Psychiatry*, 19(4), 247-255.
- Groscup, J., & Tallon, J. (2009). Theoretical models of jury decision-making. Psychology, crime and law series. Jury psychology: Social aspects of trial processes: *Psychology in the courtroom*, 1, 41-65.
- Gross, S. R., & Mauro, R. (1984). Patterns of death: An analysis of racial disparities in capital sentencing and homicide victimization. *Stanford Law Review*, 37, 27.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82.
- Guest, G., Namey, E., & McKenna, K. (2017). How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods*, 29(1), 3-22.
- Guy, L. S., & Edens, J. F. (2003). Juror decision-making in a mock sexually violent predator trial: Gender differences in the impact of divergent types of expert testimony. *Behavioral Sciences & the Law*, 21(2), 215-237.
- Hacker, F., & Frym, M. (1949). The Legal Concept of Insanity and the Treatment of Criminal Impulses. *California Law Review*, 37(4), 575.
- Haddock, G., & Zanna, M. P. (1993). Predicting prejudicial attitudes: The importance of affect, cognition, and the feeling-belief dimension. *Advances in Consumer Research*, 20, 315-18.

- Hagen, M. A. (1997). *Whores of The Court: The Fraud of Psychiatric Testimony and The Rape of American Justice*. HarperCollins Publishers.
- Hamilton, M. A., & Hunter, J. E. (1985). Analyzing utterances as the observational unit. *Human Communication Research*, 12(2), 285-294.
- Hannaford-Agor, P. L., Hans, V. P., Mott, N. L., & Munsterman, G. T. (2002). *Are hung juries a problem*. Williamsburg, VA: National Center for State Courts.
- Hans, V. P. (1986). An analysis of public attitudes toward the insanity defense. *Criminology*, 24(2), 393-414.
- Hans, V. P., & Slater, D. (1983). John Hinckley, Jr. and the insanity defense: The public's verdict. *Public Opinion Quarterly*, 47(2), 202-212.
- Haslanger, S., (2000). Gender and race: (What) are they? (What) do we want them to be? *Noûs*, 34(1), 31-55.
- Heffer, C. (2005). *The Language of Jury Trial: A Corpus-Aided Analysis of Legal-Lay Discourse*. Springer.
- Hendree, P. K., & Nicks, S. D. (2000). Perceptions of violent crimes in relation to sex of perpetrator. *Perceptual and Motor Skills*, 90(2), 485-488.
- Herzog, S., & Oreg, S. (2008). Chivalry and the Moderating Effect of Ambivalent Sexism: Individual Differences in Crime Seriousness Judgments. *Law & Society Review*, 42(1), 45-74.
- Heuer, L., & Penrod, S. (1994). Trial complexity. *Law and Human Behavior*, 18(1), 29-51.
- Holloway, I., & Todres, L. (2003). The status of method: flexibility, consistency and coherence. *Qualitative Research*, 3(3), 345-357.

- Horowitz, I. A., & Bordens, K. S. (2000). The consolidation of plaintiffs: The effects of number of plaintiffs on jurors' liability decisions, damage awards, and cognitive processing of evidence. *Journal of Applied Psychology, 85*(6), 909.
- Horowitz, I. A., Bordens, K. S., Victor, E., Bourgeois, M. J., & ForsterLee, L. (2001). The effects of complexity on jurors' verdicts and construction of evidence. *Journal of Applied Psychology, 86*(4), 641.
- Hotton Mahony, T., Jacob, J., & Hobson, H. (2017). Women and the criminal justice system. Women in Canada: A Gender-based Statistical Report. Catalogue no. 89-503-X ISSN 1719-4407
- Hough, M., & Roberts, J. (2005). *Understanding Public Attitudes to Criminal Justice*. McGraw-Hill Education (UK).
- Hurmerinta-Peltomäki, L., & Nummela, N. (2006). Mixed methods in international business research: A value-added perspective. *Management International Review, 46*(4), 439-459.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581.
- Hyde, J. S., & Plant, E. A. (1995). Magnitude of psychological gender differences: Another side to the story. *American Psychologist, 50*, 159–161.
- James, F. (1961). Sufficiency of the Evidence and Jury-Control Devices Available Before Verdict. *Virginia Law Review, 47*, 218.
- James, D., & Clarke, S. (1993). Women, men, and interruptions. *Gender and Conversational Interaction, 231-280*.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133.

- Jorm, A. F. (2000). Mental health literacy: Public knowledge and beliefs about mental disorders. *The British Journal of Psychiatry*, 177(5), 396-401.
- Juries Act, R.S.O.* 1990, c. J.3
- Kalven, H., & Zeisel, H. (1966). *The American Jury*. Chicago: University of Chicago Press.
- Kaplan, M. (1983). A woman's view of DSM-III. *American Psychologist*, 38(7), 786-792.
- Kelle, U. (2006). Combining qualitative and quantitative methods in research practice: purposes and advantages. *Qualitative Research in Psychology*, 3(4), 293-311.
- Kelly, J. R., & Barsade, S. G. (2001). Mood and emotions in small groups and work teams. *Organizational Behavior and Human Decision Processes*, 86(1), 99-130.
- Kerr, N. L., Hymes, R. W., Anderson, A. B., & Weathers, J. E. (1995). Defendant-juror similarity and mock juror judgments. *Law and Human Behavior*, 19(6), 545-567.
- Kervyn, N., Fiske, S., & Yzerbyt, V. (2015). Forecasting the primary dimension of social perception. *Social Psychology* 46(1), 36-45.
- Kondracki, N. L., Wellman, N. S., & Amundson, D. R. (2002). Content analysis: review of methods and their applications in nutrition education. *Journal of Nutrition Education and Behavior*, 34(4), 224-230.
- Kopera, M., Suszek, H., Bonar, E., Myszk, M., Gmaj, B., Ilgen, M., & Wojnar, M. (2015). Evaluating explicit and implicit stigma of mental illness in mental health professionals and medical students. *Community Mental Health Journal*, 51(5), 628-634.
- Kovera, M. B., McAuliff, B. D., & Hebert, K. S. (1999). Reasoning about scientific evidence: Effects of juror gender and evidence quality on juror decisions in a hostile work environment case. *Journal of Applied Psychology*, 84(3), 362.

- Krech, D., & Crutchfield, R. S. (1948). *Theory and problems of social psychology*. New York, NY, US: McGraw-Hill.
- Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research, 30*(3), 411-433.
- Laberge, D., Morin, D., Armony, V. (1997). Les représentations sexuées dans les discours d'experts psychiatres. *Déviances et Société, 21*(3), 251-272.
- Lakoff, G. (1975). Hedges: A study in meaning criteria and the logic of fuzzy concepts. In *Contemporary Research in Philosophical Logic and Linguistic Semantics* (pp. 221-271). Springer, Dordrecht.
- Lamar, C. (2012). In 1995, New Mexico voted on a bill requiring psychologists to dress as wizards. Retrieved from: <http://io9.gizmodo.com/5882671/in-1995-new-mexico-voted-on-a-bill-requiring-psychologists-to-dress-as-wizards> [March 9, 2016].
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: the importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology, 93*(2), 234.
- Leaper, C., & Ayres, M. M. (2007). A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *Personality and Social Psychology Review, 11*(4), 328-363.
- Leaper, C., & Robnett, R. D. (2011). Women are more likely than men to use tentative language, aren't they? A meta-analysis testing for gender differences and moderators. *Psychology of Women Quarterly, 35*(1), 129-142.

- Leaper, C., & Smith, T. E. (2004). A meta-analytic review of gender variations in children's language use: talkativeness, affiliative speech, and assertive speech. *Developmental Psychology, 40*(6), 993.
- Leaper, C., Anderson, K. J., & Sanders, P. (1998). Moderators of gender effects on parents' talk to their children: A meta-analysis. *Developmental Psychology, 34*(1), 3.
- Levett, L. M., & Kovera, M. B. (2008). The effectiveness of opposing expert witnesses for educating jurors about unreliable expert evidence. *Law and Human Behavior, 32*(4), 363-374.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage.
- Link, B. G. (1987). Understanding labeling effects in the area of mental disorders: An assessment of the effects of expectations of rejection. *American Sociological Review, 52*, 96-112.
- Louden, J. E., & Skeem, J. L. (2007). Constructing insanity: jurors' prototypes, attitudes, and legal decision-making. *Behavioral Sciences & the Law, 25*(4), 449-470.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and social psychology bulletin, 18*(3), 302-318.
- Lynch, M., & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension, and discrimination. *Law and Human Behavior, 33*(6), 481.
- Lynch, M., & Haney, C. (2011). Mapping the racial bias of the white male capital juror: Jury composition and the "empathic divide". *Law & Society Review, 45*(1), 69-102.
- Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology, 79*(4), 602.

- Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology, 79*(4), 602.
- Maeder, E. M., & Dempsey, J. L. (2013). A likely story? The influence of type of alibi and defendant gender on juror decision-making. *Psychiatry, Psychology and Law, 20*(4), 543-552.
- Maeder, E. M., Yamamoto, S., & McManus, L. A. (2018). Methodology matters: Comparing sample types and data collection methods in a juror decision-making study on the influence of defendant race. *Psychology, Crime & Law, 24*(7), 687-702.
- Maeder, E., Dempsey, J., & Pozzulo, J. (2012). Behind the veil of juror decision making: Testing the effects of Muslim veils and defendant race in the courtroom. *Criminal Justice and Behavior, 39*(5), 666-678.
- Malle, B. F., & Nelson, S. E. (2003). Judging *mens rea*: The tension between folk concepts and legal concepts of intentionality. *Behavioral Sciences & the Law, 21*(5), 563-580.
- Marder, N. S. (1987). Gender dynamics and jury deliberations. *Yale Law Journal 3*(6), 593-612.
- Marshall, C., & Rossman, G. B. (1999). *Designing Qualitative Research* (3rd ed.). Thousand Oaks, CA: Sage.
- Martin, J. K., Pescosolido, B. A., & Tuch, S. A. (2000). Of fear and loathing: The role of “disturbing behavior”, labels, and causal attributions in shaping public attitudes toward people with mental illness. *Journal of Health and Social Behavior, 208-223*.
- Masser, B., Lee, K., & McKimmie, B. M. (2010). Bad woman, bad victim? Disentangling the effects of victim stereotypicality, gender stereotypicality and benevolent sexism on acquaintance rape victim blame. *Sex Roles, 62*(7-8), 494-504.

- Mazzella, R., & Feingold, A. (1994). The Effects of Physical Attractiveness, Race, Socioeconomic Status, and Gender of Defendants and Victims on Judgments of Mock Jurors: A Meta-Analysis 1. *Journal of Applied Social Psychology, 24*(15), 1315-1338.
- McLeod, S. A. (2019, July 10). What does effect size tell you? Simply Psychology. <https://www.simplypsychology.org/effect-size.html>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276-282.
- McKim, C. A. (2017). The value of mixed methods research: A mixed methods study. *Journal of Mixed Methods Research, 11*(2), 202-222.
- McKimmie, B. M., & Masser, B. M. (2010). The effect of gender in the courtroom. *Forensic psychology: Concepts, Debates and Practice, 2*, 95-122.
- McKimmie, B. M., Masser, B. M., & Bongiorno, R. (2014). What counts as rape? The effect of offense prototypes, victim stereotypes, and participant gender on how the complainant and defendant are perceived. *Journal of Interpersonal Violence, 29*(12), 2273-2303.
- McKimmie, B. M., Masters, J. M., Masser, B. M., Schuller, R. A., & Terry, D. J. (2013). Stereotypical and counterstereotypical defendants: Who is he and what was the case against her?. *Psychology, Public Policy, and Law, 19*(3), 343.
- McKimmie, B. M., Newton, C. J., Terry, D. J., & Schuller, R. A. (2004). Jurors' responses to expert witness testimony: The effects of gender stereotypes. *Group Processes & Intergroup Relations, 7*(2), 131-143.
- McKnight, A. (2013). Jury Nullification as a Tool to Balance the Demands of Law and Justice. *BYU Law Review, 1103*.

- McNamara, K., Vattano, F., & Viney, W. (1993). Verdict, sentencing, and certainty as a function of sex of juror and amount of evidence in a simulated rape trial. *Psychological Reports*, 72(2), 575-583.
- Memon, A., & Shuman, D. W. (1998). Juror perception of experts in civil disputes: The role of race and gender. *Law & Psychology Review*, 22, 179.
- Miladinovic, Z., & Lukassen, J. (2014). Verdicts of not criminally responsible on account of mental disorder in adult criminal courts, 2005/2006-2011/2012. Statistics Canada.
- Millett, K. (1969). *Sexual politics*. New York: Granada.
- Monahan, J., & Hood, G. L. (1976). Psychologically disordered and criminal offenders: Perceptions of their volition and responsibility. *Correctional Psychologist*, 3(2), 123-134.
- Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20(3), 336-366.
- Morse, S. J. (1985). Excusing the crazy: The insanity defense reconsidered. *State of California Law Review*, 58, 777.
- Mossière, A., & Maeder, E. M. (2015). Defendant mental illness and juror decision-making: A comparison of sample types. *International Journal of Law and Psychiatry*, 42, 58-66.
- Mossière, A., & Maeder, E. M. (2016). Juror decision making in not criminally responsible on account of mental disorder trials: Effects of defendant gender and mental illness type. *International Journal of Law and Psychiatry*, 49, 47-54.
- Moulds, E. F. (1978). Chivalry and paternalism: Disparities of treatment in the criminal justice system. *Western Political Quarterly*, 31(3), 416-430.
- Myers, M. (2000). Qualitative research and the generalizability question: Standing firm with Proteus. *The Qualitative Report*, 4(3), 9.

- Nagel, I., & Hagan, J. (1983). Gender and Crime: Offense Patterns and Criminal Court Sanctions. *Crime And Justice*, 4, 91-144.
- National Judicial Institute/Institut National de la Magistrature (CJC-NJI). (2011). Model Jury Instructions. Retrieved from: https://www.cjc-ccm.gc.ca/english/lawyers_en.asp?selMenu=lawyers_juryinstruction_en.asp
- Neal, T. M., & Brodsky, S. L. (2008). Expert witness credibility as a function of eye contact behavior and gender. *Criminal Justice and Behavior*, 35(12), 1515-1526.
- Neal, T. M., Guadagno, R. E., Eno, C. A., & Brodsky, S. L. (2012). Warmth and competence on the witness stand: Implications for the credibility of male and female expert witnesses. *Journal of the American Academy of Psychiatry and the Law Online*, 40(4), 488-497.
- Nicholls, T. L., Brink, J., Greaves, C., Lussier, P., & Verdun-Jones, S. (2009). Forensic psychiatric inpatients and aggression: An exploration of incidence, prevalence, severity, and interventions by gender. *International Journal of Law and Psychiatry*, 32(1), 23-30.
- Nicholls, T. L., Crocker, A. G., Seto, M. C., Wilson, C. M., Charette, Y., & Côté, G. (2015). The national trajectory project of individuals found not criminally responsible on account of mental disorder. Part 5: how essential are gender-specific forensic psychiatric services?. *The Canadian Journal of Psychiatry*, 60(3), 135-145.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1).
- Olson, J. D., McAllister, C., Grinnell, L. D., Gehrke Walters, K., & Appunn, F. (2016). Applying constant comparative method with multiple investigators and inter-coder reliability. *The Qualitative Report*, 21(1), 26-42.

- Ontario Ministry of the Attorney General. (2018). General information about jury duty.
Retrieved from: <https://www.attorneygeneral.jus.gov.on.ca/english/courts/jury/>
- Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49(3), 197-237.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Oxford, England: University of Illinois Press.
- Ottati, V., Bodenhausen, G. V., & Newman, L. S. (2005). Social psychological models of mental illness stigma. In P.W. Corrigan (Ed.) *On the Stigma of Mental Illness: Practical Strategies for Research and Social Change*, (pp. 99-128). Washington, DC: American Psychological Association.
- Overton, S. L., & Medina, S. L. (2008). The stigma of mental illness. *Journal of Counseling & Development*, 86(2), 143-151.
- Patry, M. W. (2008). Attractive but guilty: Deliberation and the physical attractiveness bias. *Psychological Reports*, 102(3), 727-733.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007. Austin, TX.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001*. Mahway: Lawrence Erlbaum Associates, 71.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of personality and social psychology*, 51(2), 242.
- Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of personality and social psychology*, 62(2), 189.

- Perkins Gilman, C. (1898). *Women and Economics: A Study of the Economic Relation Between Men and Women as a Factor in Social Evolution*. Reprinted (1966). New York: Harper TorchBooks.
- Perlin, M. L. (1991). Morality and pretextuality, psychiatry and law: Of 'ordinary common sense', heuristic reasoning, and cognitive dissonance. *Bulletin of American Academy of Psychiatry and the Law*, 19(2), 131-150.
- Perlin, M. L. (1996). Myths, realities, and the political world: The anthropology of insanity defense attitudes. *Journal of the American Academy of Psychiatry and the Law Online*, 24(1), 5-26.
- Perlin, M. L. (1997). Where the winds hit heavy on the borderline: mental disability law, theory and practice, us and them. *Loyola of Los Angeles Law Review*, 31, 775.
- Perlin, M. L. (2003). She breaks just like a little girl: Neonaticide, the insanity defense, and the irrelevance of ordinary common sense. *William and Mary Journal of Women and the Law* 10(001), 1-31.
- Perlin, M. L. (2009). "His brain has been mismanaged with great skill": How will jurors respond to neuroimaging testimony in defence cases? *Akron Law Review*, 442, 886-914.
- Perlin, M. L. (2013). Sanism and the law. *Virtual Mentor*, 15(10), 878-885.
- Perlin, M. L. (2016). *The Insanity Defense: Nine Myths that Will Not Go Away. The Insanity Defense: Multidisciplinary Views on Its History, Trends, and Controversies*. In M. D. White, (Ed.). Available at SSRN: <https://ssrn.com/abstract=2715577>
- Pescosolido, B. A. (2013). The public stigma of mental illness: What do we think; what do we know; what can we prove?. *Journal of Health and Social behavior*, 54(1), 1-21.

- Peter-Hagene, L. (2019). Jurors' cognitive depletion and performance during jury deliberation as a function of jury diversity and defendant race. *Law and Human Behavior, 43*(3), 232.
- Polletta, F., Chen, P. C. B., Gardner, B. G., & Motes, A. (2011). The sociology of storytelling. *Annual Review of Sociology, 37*, 109-130.
- Popp, D., Donovan, R. A., Crawford, M., Marsh, K. L., & Peele, M. (2003). Gender, race, and speech style stereotypes. *Sex Roles, 48*(7-8), 317-325.
- Posey, L. (2016). Gender. *Critical Quarterly, 58*(3).
- Pozzulo, J. D., Dempsey, J., Maeder, E., & Allen, L. (2010). The effects of victim gender, defendant gender, and defendant age on juror decision making. *Criminal Justice and Behavior, 37*(1), 47-63.
- Pryzgod, J., & Chrisler, J. C. (2000). Definitions of gender and sex: The subtleties of meaning. *Sex Roles 43*(7-8), 553-569.
- Quas, J. A., Bottoms, B. L., Haegerich, T. M., & Nysse-Carris, K. L. (2002). Effects of Victim, Defendant, and Juror Gender on Decisions in Child Sexual Assault Cases 1. *Journal of Applied Social Psychology, 32*(10), 1993-2021.
- Quen, J. M. (1981). Anglo-American concepts of criminal responsibility: A brief history. In S. J. Hucker, C. D. Webster and M H. Ben-Aron (Eds.). *Mental Disorder and Criminal Responsibility*. Toronto: Butterworths.
- R v. Turcotte*, 2013, QCCA 1916
- R. v. Long* (1990), 61 C.C.C. (3d) 156
- R. v. Mohan*, [1994] 2 S.C.R. 9.
- Raudenbush, S.W., Bryk, A.S, & Congdon, R. (2017). HLM 7.03 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.

- Roberts, C. F., & Golding, S. L. (1991). The social construction of criminal responsibility and insanity. *Law and Human Behavior, 15*(4), 349-376.
- Rosenblum, K. E. (1986). Revelatory or purposive?. Making sense of a 'female register'. *Semiotica, 59*(1-2), 157-170.
- Rothwell, E., Anderson, R., & Botkin, J. R. (2016). Deliberative discussion focus groups. *Qualitative Health Research, 26*(6), 734-740.
- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: the malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81*(5), 856.
- Russell, B., Ragatz, L., & Kraus, S. W. (2012). Expert testimony of the battered person syndrome, defendant gender, and sexual orientation in a case of duress: Evaluating legal decisions. *Journal of Family Violence, 27*(7), 659-670.
- Saks, M. J., & Marti, M. W. (1997). A meta-analysis of the effects of jury size. *Law and Human Behavior, 21*(5), 451-467.
- Salerno, J. M., & Diamond, S. S. (2010). The promise of a cognitive perspective on jury deliberation. *Psychonomic Bulletin & Review, 17*(2), 174-179.
- Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health, 18*(2), 179-183.
- Schnittker, J. (2000). Gender and Reactions to Psychological Problems: An Examination of Social Tolerance and Perceived Dangerousness. *Journal Of Health And Social Behavior, 41*(2), 224.
- Schutte, J. W., & Hosch, H. M. (1997). Gender differences in sexual assault verdicts: A meta-analysis. *Journal of Social Behavior and Personality, 12*(3), 761.

- Seedat, S., Scott, K., Angermeyer, M., Berglund, P., Bromet, E., & Brugha, T. et al. (2009). Cross-National Associations Between Gender and Mental Disorders in the World Health Organization World Mental Health Surveys. *Archives Of General Psychiatry*, 66(7), 785.
- Showalter, E. (1985). *Female malady: Women, madness and English culture* (pp. 1830-1980). Virago Press.
- Silton, N. R., Flannelly, K. J., Milstein, G., & Vaaler, M. L. (2011). Stigma in America: has anything changed?: Impact of perceptions of mental illness and dangerousness on the desire for social distance: 1996 and 2006. *The Journal of Nervous and Mental Disease*, 199(6), 361-366.
- Silver, E., Cirincione, C., & Steadman, H. J. (1994). Demythologizing inaccurate perceptions of the insanity defense. *Law and Human Behavior*, 18(1), 63-70.
- Simon, L. M. J. (1996a). Legal treatment of the victim-offender relationship in crimes of violence. *Journal of Interpersonal Violence*, 11, 94-106.
- Simon, L. M. J. (1996b). The effect of the victim-offender relationship on the sentence length of violent offenders. *Journal of Crime & Justice*, 19, 129-148.
- Skeem, J. L., & Golding, S. L. (2001). Describing jurors' personal conceptions of insanity and their relationship to case judgments. *Psychology, Public Policy, and Law*, 7(3), 561.
- Smith, C. M., Bushouse, E., & Lord, J. (2010). Individual and group performance on insight problems: The effects of experimentally induced fixation. *Group Processes & Intergroup Relations*, 13(1), 91-99.
- Smith, L. D. (1993a). To cure those afflicted with the disease of insanity: Thomas Bakewell and Spring Vale Asylum. *History of psychiatry*, 4(13), 107-127.

- Smith, V. L. (1993b). When prior knowledge and law collide—Helping jurors use the law. *Law and Human Behavior, 17*, 507-536.
- Sommers, S. R. (2006). On racial diversity and group decision making: identifying multiple effects of racial composition on jury deliberations. *Journal of Personality and Social Psychology, 90*(4), 597.
- Sommers, S. R., & Ellsworth, P. C. (2009). “Race salience” in juror decision-making: misconceptions, clarifications, and unanswered questions. *Behavioral Sciences & the Law, 27*(4), 599-609.
- Spohn, C., & Spears, J. (1997). Gender and Case Processing Decisions. *Women & Criminal Justice, 8*(3), 29-59.
- Steffensmeier, D., & Kramer, J. H. (1982). Sex-based differences in the sentencing of adult criminal defendants: An empirical test and theoretical overview. *Sociology & Social Research, 66*(3), 289-304.
- Stoppard, J. M. (2000). *Understanding depression*. New Jersey, USA: Clays Ltd.
- Strub, T., & McKimmie, B. M. (2016). Sugar and spice and all things nice: The role of gender stereotypes in jurors’ perceptions of criminal defendants. *Psychiatry, Psychology and Law, 23*(4), 487-498.
- Stuart, H. (2003). Violence and mental illness: an overview. *World Psychiatry, 2*(2), 121.
- Stygall, G. (1994). *Trial Language: Differential Discourse Processing and Discursive Formation* (Vol. 26). John Benjamins Publishing.
- Swim, J. K., & Cohen, L. L. (1997). Overt, covert, and subtle sexism: A comparison between the attitudes toward women and modern sexism scales. *Psychology of Women Quarterly, 21*(1), 103-118.

- Sygel, K., Sturup, J., Fors, U., Edberg, H., Gavazzeni, J., Howner, K., Persson, M., & Kristiansson. (2017). The effect of gender on the outcome of forensic psychiatric assessment in Sweden: A case vignette study. *Criminal Behaviour and Mental Health*, 27(2), 124-135.
- Szasz, T. S. (1974). The myth of psychotherapy. *American Journal of Psychotherapy*, 28(4), 517-526.
- Tabak, S. J., & Kletke, B. (2014). Mock jury attitudes towards credibility, age, and guilt in a fictional child sexual assault scenario. *Australian Journal of Psychology* 66, 47-55.
- Tabak, S. J., Kletke, B., & Knight, T. (2013). Simulated jury decision making in online focus groups. *Qualitative Research Journal*, 13, 102-113.
- Tajfel, H., & Billic, M. (1974). Familiarity and categorization in intergroup behavior. *Journal of Experimental Social Psychology*, 10(2), 159-170.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations*, 33(47), 74.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 724).
- Tamborini, R., Huang, R. H., Mastro, D., & Nabashi-Nakahara, R. (2007). The influence of race, heuristics, and information load on judgments of guilt and innocence. *Communication Studies*, 58(4), 341-358.
- Tanford, S., & Penrod, S. (1986). Jury deliberations: Discussion content and influence processes in jury decision making. *Journal of Applied Social Psychology*, 16(4), 322-347.

- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology, 29*(1), 24-54.
- Taylor, T. S., & Hosch, H. M. (2004). An examination of jury verdicts for evidence of a similarity-leniency effect, an out-group punitiveness effect or a black sheep effect. *Law and Human Behavior, 28*(5), 587-598.
- Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. *Journal of Advanced Nursing, 48*(4), 388-396.
- Trepte, S., & Loy, L. S. (2017). Social identity theory and self-categorization theory. In P. Rössler, C. A. Hoffner, L. van Zoonen (Eds.) *The International Encyclopedia of Media Effects*. London: John Wiley & Sons, Inc.
- Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. *European journal of social psychology, 5*(1), 1-34.
- Twain, M. (April 16, 1870). A New Crime: Legislation Needed. *Buffalo Express*.
- Tyler, T. (2017). Procedural justice and policing: A rush to judgment?. *Annual Review of Law and Social Science, 13*, 29-53.
- United Nations General Assembly. (1948). *The universal declaration of human rights*. Geneva: United Nations.
- Ussher, J. M. (1991). *Women's madness: Misogyny or mental illness?*. University of Massachusetts Press.
- Ussher, J. M. (2013). Diagnosing difficult women and pathologising femininity: Gender bias in psychiatric nosology. *Feminism & Psychology, 23*(1), 63-69.

- Vajargah, F., Nikbakht, M., 2015. Application Reamlmodel And Determining Cut Off Of Icc By Multi-Level Model Based On Markov Chains Simulation In Health. *Indian Journal of Fundamental Applied Life Science* 5, 1432–1448.
- Van Dijk, T. A. (2013). Discourse, power and access. In *Texts and Practices* (pp. 93-113). Routledge.
- Verdun-Jones, S. N. (1994). The insanity defence in Canada: Setting a new course. *International Journal of Law and Psychiatry*, 17, 175-189.
- Veysey, B. M. (2015). Gender role incongruence and the adjudication of criminal responsibility. *Albany Law Review* 78(3), 1087-1107.
- Viki, G. T., & Abrams, D. (2002). But she was unfaithful: Benevolent sexism and reactions to rape victims who violate traditional gender role expectations. *Sex Roles*, 47(5-6), 289-293.
- Viki, G. T., Abrams, D., & Masser, B. (2004). Evaluating stranger and acquaintance rape: The role of benevolent sexism in perpetrator blame and recommended sentence length. *Law and Human Behavior*, 28(3), 295-303.
- Visher, C. A. (1987). Juror decision making. *Law and Human Behavior*, 11(1), 1-17.
- Vollrath, D. A., Sheppard, B. H., Hinsz, V. B., & Davis, J. H. (1989). Memory performance by decision-making groups and individuals. *Organizational Behavior and Human Decision Processes*, 43(3), 289-300.
- Von Peter, S. (2010). The temporality of “chronic” mental illness. *Culture, Medicine, and Psychiatry*, 34(1), 13-28.
- Ward, T. (1997). Law, common sense and the authority of science: expert witnesses and criminal insanity in England, CA. 1840-1940. *Social & Legal Studies*, 6(3), 343-362.

- Watkins, D. C. (2017). Rapid and rigorous qualitative data analysis: The “RADaR” technique for applied research. *International Journal of Qualitative Methods* 16, 1-9.
- Weare, S. (2013). “The Mad”, “The Bad”, “The Victim”: Gendered Constructions of Women Who Kill within the Criminal Justice System. *Laws*, 2(3), 337-361.
- Weinstein, J. B. (1992). Considering jury nullification: When may and should a jury reject the law to do justice. *American Criminal Law Review*, 30, 239.
- Wiener, R. L., Krauss, D. A., & Lieberman, J. D. (2011). Mock jury research: Where do we go from here?. *Behavioral Sciences & the Law*, 29(3), 467-479.
- Williams, J. H. (1957). Defect of Reason from Disease of the Mind. *The Modern Law Review*, 55-57.
- Wojciszke, B. (2005). Affective concomitants of information on morality and competence. *European psychologist*, 10(1), 60-70.
- Wood, J. T. (1995). *Gendered interaction: Masculine and feminine styles of verbal communication*. New York: Wadsworth.
- Wozniak, K. H. (2016). Perceptions of prison and punitive attitudes: A test of the penal escalation hypothesis. *Criminal Justice Review*, 41(3), 352-371.
- Yamamoto, S. (2019). What’s in the box?: Punishment and insanity in the Canadian jury deliberation room. Unpublished Dissertation.
- Yourstone, J., Lindholm, T., Grann, M., & Svenson, O. (2008). Evidence of gender bias in legal insanity evaluations: A case vignette study of clinicians, judges and students. *Nordic journal of psychiatry*, 62(4), 273-278.
- Zanna, M. P., & Hamilton, D. L. (1972). Attribute dimensions and patterns of trait inferences. *Psychonomic Science*, 27(6), 353-354.

Zhang, Y., & Wildemuth, B. M. (2016). Qualitative analysis of content. *Applications of Social Research Methods to Questions in Information and Library Science*, 318.

Table 9 Coded Stereotype Content dictionary

Adjective	High Competence	Low Competence	High Warmth	Low warmth	Neither
Addict		X			
Aggressive				X	
Agreeable			X		
Alert	X				
Aloof				X	
Amiable			X		
Argumentative				X	
Arrogant				X	
Artistic	X				
Athletic	X				
Benevolent			X		
Bitter				X	
Boastful				X	
Brilliant	X				
Brutal				X	
Calculating	X				
Callous				X	
Capable	X				
Caring			X		
Cheerful			X		
Clever	X				
Cold				X	
Compassionate			X		

Competent	X				
Conceited				X	
Concerned			X		
Conventional					X
Courteous			X		
Cowardly		X			
Crafty	X				
Creative	X				
Criminal				X	
Cruel				X	
Cunning	X				
Deceitful				X	
Dumb		X			
Educated	X				
Efficient	X				
Empathetic			X		
Evasive					X
Evil				X	
Faithful			X		
Forgiving			X		
Friendly			X		
Generous			X		
Genial			X		
Gentle			X		
Genuine			X		
Good-natured			X		

Gracious			X		
Greedy				X	
Happy			X		
Hateful				X	
Heartless				X	
Honest			X		
Hostile				X	
Adjective	High Competence	Low Competence	High Warmth	Low Warmth	Neither
Humourless				X	
Ignorant		X			
Imaginative	X				
Imitative		X			
Impulsive					X
Incompetent		X			
Inconsiderate				X	
Independent	X				
Industrious	X				
Inept		X			
Insensitive				X	
Intelligent	X				
Intense					X
Jovial			X		
Kind			X		
Knowledgeable	X				

Lazy		X			
Likeable			X		
Loud				X	
Manipulative				X	
Materialistic					X
Meditative					X
Methodical	X				
Naïve		X			
Nasty				X	
Neat	X				
Nice			X		
Perceptive	X				
Persistent	X				
Phony				X	
Pleasant			X		
Practical	X				
Proficient	X				
Progressive	X				
Quick-tempered				X	
Quiet				X	
Radical				X	
Rational	X				
Reserved				X	
Resourceful	X				
Revengeful				X	

Rhythmic	X				
Scientific	X				
Sensitive			X		
Shallow				X	
Shrewd	X				
Simple		X			
Simple-minded		X			
Skilled	X				
Skillful	X				
Slick	X				
Slow		X			
Sly				X	
Smart	X				
Sophisticated	X				
Straightforward	X				
Strong	X				
Stupid		X			
Suave			X		
Suggestible		X			
Superstitious		X			
Suspicious				X	
Uneducated		X			
Unreliable		X			
Vulnerable		X			
Warm			X		

Appendix A. Recruitment notice on kijiji.com and initial screening



We are recruiting Canadian jury-eligible participants for a study “Jury Deliberation and the Not Criminally Responsible on Account of Mental Disorder Defence”. Participants must be Canadian citizens, at least 18 years of age, who have not been convicted of an indictable offence without receiving an official pardon. Participants will first be asked to fill out a brief online questionnaire (5 minutes). Participants will be asked to later come to Carleton University and read a transcript from a trial including an allegation of murder, then take part in a jury deliberation with up to 14 other participants. This study is being conducted by Dr. Evelyn Maeder from the Institute of Criminology and Criminal Justice at Carleton University, and PhD student Kendra McLaughlin from the Department of Psychology at Carleton University.

The study will take approximately 90 minutes and participants will be awarded \$40 in person.

Participants must be available to travel to Carleton University on a Saturday. This study has been cleared by the Carleton University Research Ethics Board – B (#16-023).

Interested participants should click the link below for eligibility and further information.

[Click here to begin.](#)

Appendix B. Informed consent online

The purpose of an informed consent is to make sure that you understand the purpose of the study and your involvement as a participant. The informed consent must include enough information regarding the study for you to be able to make a well-informed decision regarding whether or not you would like to partake in this study.

Title: Jury Deliberation and the Not Criminally Responsible on Account of Mental Disorder Defence

Research Personnel: This study is being conducted by Dr. Evelyn Maeder from the Institute of Criminology and Criminal Justice (evelyn.maeder@carleton.ca Tel. 613-520-2600 ext. 2421) at Carleton University in Ottawa, Ontario, Canada, and PhD student Kendra McLaughlin (kendra.mclaughlin@carleton.ca) from the Department of Psychology at Carleton University.

Concerns: Should you have any ethical concerns regarding this study, please contact Dr. Shelley Brown (Chair, Carleton University Research Ethics Board – B, Shelley_brown@carleton.ca, 613-520-2600 ext. 1505). For other concerns, please contact the Research Compliance Office at ethics@carleton.ca.

Purpose: This is a study to evaluate opinions about the Not Criminally Responsible on Account of Mental Disorder Defence and about punishment in general.

Task Requirements:

Participants must be Canadian citizens, at least 18 years of age, who have not been convicted of an indictable offence without receiving an official pardon. Please note that participants found to be ineligible for the study will not receive compensation.

Phase 1. This is part one of a two-part study. If you agree to participate in part one, then following this consent form you will be asked to answer some demographics questions to

ensure eligibility, (approximately 5 minutes). To receive compensation, you must take part in both Phase 1 and Phase 2.

Phase 2. The second phase of the study will take place at Carleton University. Participants will be asked to read a short trial transcript and to have a group discussion (no longer than 45 minutes) with up to 14 other participants, which will be video/audio recorded and observed through a two-way mirror. When you arrive at the lab we will need to obtain your consent to be video/audio recorded. Compensation for participating in both phases of the study is \$40 in cash, which will be provided in person immediately following participation in Phase 2 at Carleton University. Full participation in Phase 2 of the study will take up to approximately 90 minutes.

Potential Risk/Discomfort: You may refrain from answering any questions on the questionnaire if you are uncomfortable or otherwise do not want to and you will still be eligible to participate in Phase 2 of the study.

Right to Withdraw: In addition, at any time, you may discontinue your involvement in this study without penalty. However, if you withdraw prior to Phase 2 of the study, we cannot provide you with compensation.

Anonymity/Confidentiality: The data collected are strictly confidential. Your name is not associated with the responses you provide. We collect data through the software Qualtrics, which uses servers with multiple layers of security to protect the privacy of the data (e.g., encrypted websites and pass-word protected storage). Please note that Qualtrics is hosted by a server located in the USA. The United States Patriot Act permits U.S. law enforcement officials, for the purpose of an anti-terrorism investigation, to seek a court order that allows access to the personal records of any person without that person's knowledge. In view of this we cannot absolutely guarantee the full confidentiality and anonymity of your data. With your consent to participate in this study you acknowledge this. These data will only be used for research at Carleton University. After completing Phase 1 you will be asked to provide your email address in a separate survey. However, because we will need to correspond by email and assign you a juror code, it is possible that your email address will be associated with your responses until up to 2 weeks after Phase 2 of the study. This information will be kept strictly confidential and we will only use your email address for the purposes of providing information about the study.

This study has been cleared by the Carleton University Research Ethics Board – B (#16-023).

I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential with respect to my personal identity. By checking this box, I'm indicating that I agree to participate in this study.

I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential with respect to my personal identity. By checking this box, I'm indicating that I disagree to participate in this study.

Continue to survey =>

Appendix C. Informed consent in lab

Informed Consent: Phase 2

The purpose of an informed consent is to make sure that you understand the purpose of the study and your involvement as a participant. The informed consent must include enough information regarding the study for you to be able to make a well-informed decision regarding whether or not you would like to partake in this study.

Title: Jury Deliberation and the Not Criminally Responsible on Account of Mental Disorder Defence

Research Personnel: This study is being conducted by Dr. Evelyn Maeder from the Institute of Criminology and Criminal Justice (evelyn.maeder@carleton.ca Tel. 613-520-2600 ext. 2421) at Carleton University in Ottawa, Ontario, Canada, and PhD student Kendra McLaughlin (kendra.mclaughlin@carleton.ca) from the Department of Psychology at Carleton University.

Concerns: Should you have any ethical concerns regarding this study, please contact Dr. Shelley Brown (Chair, Carleton University Research Ethics Board – B, Shelley_brown@carleton.ca, 613-520-2600 ext. 1505). For other concerns, please contact the Research Compliance Office at ethics@carleton.ca.

Purpose: This is a study to evaluate opinions about the Not Criminally Responsible on Account of Mental Disorder Defence and about punishment in general.

Task Requirements: You will first read a short trial transcript (which will take about 15 minutes). We will then ask you to answer a brief questionnaire about what you have read. After that, we will ask you to begin group deliberation. Your task will be to come to a unanimous verdict decision. If after 45 minutes you cannot reach a verdict, then the jury will be declared “hung”, a term that describes a jury that cannot come to a unanimous decision. After deliberation is complete, we will have a 5-minute debriefing session, during which I will provide you with additional information and answer any questions you might have. Compensation for participating in both phases of the study is \$40 in cash, which will be

provided in person immediately following participation in Phase 2 at Carleton University. Full participation in Phase 2 of the study will take up to approximately 90 minutes.

Potential Risk/Discomfort: There is a possibility you may feel uncomfortable with the nature of the trial transcript, which involves a second-degree murder charge and testimony about mental illness.

Right to Withdraw: In addition, at any time, you may discontinue your involvement in this study without penalty. If you no longer wish to participate you can simply exit the room and knock on the door of the room on the immediate right, where we will assist you promptly. However, the video and audio data will be retained for later analysis, and so you cannot withdraw your data after you have already participated.

Anonymity/Confidentiality: The data collected are strictly confidential. Your name is not associated with the responses you provide. However, it is important to understand that you will be video/audio recorded, and these data will be retained indefinitely. In view of this we cannot absolutely guarantee the anonymity of your data. However, audio/video recorded deliberation data will not be stored online in order to reduce risk of data breach. Contact information will not be associated with these data in any way. The data will be kept in our lab facility on one password protected computer. Only coders/trusted colleagues will have access to these data. With your consent to participate in this study you acknowledge this. These data will only be used for research at Carleton University.

This study has been cleared by the Carleton University Research Ethics Board – B (#16-023).

Please check one of the options below:

I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential with respect to my personal identity. By checking this box, I'm indicating that I agree to participate in this study.

OR

I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential with respect to my personal identity. By checking this box, I'm indicating that I disagree to participate in this study.

Print Name: _____

Signature: _____

I understand that I will be video and audio recorded, and that once group deliberation begins I will not be able to withdraw these data. I understand that I may discontinue participation at any time and still receive full compensation, and that these data will be kept strictly confidential.

Please check one of the options below:

I consent to be audio/video recorded

I do not consent to be audio/video recorded

Appendix D. Online eligibility forms

Thank you for your interest in our study! Before we can schedule you for a session, we just need to make sure that you are eligible to participate. Please answer the following questions:

Are you a Canadian citizen?

Yes ___ No ___

Are you 18 years of age or older?

Yes ___ No ___

Have you ever been convicted of an indictable offence? If you have been convicted but you received a formal pardon, please select 'No'. Indictable offenses are different from more minor "summary" offences (which carry sentences of fewer than 6 months and fines of less than \$2000). Examples include theft, treason, murder, piracy, and robbery.

Yes ___ No ___



Ineligible Message

I'm sorry; it looks as though you are not eligible to participate in this study. We are looking for participants who could get called for jury duty in Canada (i.e., Canadian citizens at least 18 years of age, who have not been convicted of an indictable offence).

Thank you very much for your time and interest. If you have any questions please don't hesitate to contact us: carletonjury@gmail.com.

Eligible Message

Thank you for your response. It looks as though you are eligible to participate in this study. If you would like to participate, you will be asked to travel to Carleton University, at which point you will read a brief trial transcript (about 15 minutes). You will then be asked to discuss the case and reach a unanimous verdict with up to 14 other mock jurors. Full participation in this study will take approximately 90 minutes. The deliberation session will be audio and video recorded. These data will be kept strictly confidential (stored offline in a locked lab facility). We will need to obtain your consent for this when you arrive at the lab.

Before you come to the lab, we will need you to fill out 2 brief questionnaires. To receive the \$40 cash compensation, we will need you to fill out these questionnaires and travel to Carleton for the study; there is no compensation for the questionnaires alone. After you fill out these questionnaires, you will receive a confirmation email for the deliberation study.

Appendix E. In-lab welcome script

Welcome and thank you for agreeing to participate in our study entitled “Jury Deliberation and the Not Criminally Responsible on Account of Mental Disorder Defence”. Today, we are asking you to play the role of a juror. You will first read a short trial transcript (which will take about 15 minutes). We will then ask you to answer a brief questionnaire about what you have read. After that, we will ask you to begin group deliberation. Your task will be to come to a unanimous verdict decision. If after 45 minutes you cannot reach a verdict, then the jury will be declared “hung”, a term that describes a jury that cannot come to a unanimous decision. After deliberation is complete, we will have a 5-minute debriefing session, during which I will provide you with additional information and answer any questions you might have. Please note that the correspondence associating your juror code with your email address will be deleted from our email account, and that information will be separated from our database as well.

There is a possibility you may feel uncomfortable with the nature of the trial transcript, which involves a second-degree murder charge and testimony about mental illness.

It is important for you to understand that this study will be audio and video recorded. We will also observe the study through this mirror. Your participation in this study is completely voluntary, and you have the right to refrain from answering any questions that you so choose. Your participation signifies that you consent to this, and that you understand that the video will be kept strictly confidential. You have the right to discontinue participation in this study at any point and you will still receive full compensation. If you no longer wish to participate you can simply exit the room and knock on the door of the room on the immediate right, where we will assist you promptly.

Once deliberations begin, although you may discontinue your involvement, we will retain the video footage for later analysis. These data will only be reported in aggregate form. Before the study begins, we will ask you to complete an informed consent form (show the form).

You will notice that you have ‘name’ cards with your juror code on them. These are for the purpose of communicating with each other during the deliberation and to protect your identity given that you will be video and audio recorded. We kindly ask that you refrain from the use of cell phones throughout the study except in case of an emergency

We will now ask you to read this informed consent form (volunteer hands out informed consent form in folders) that will further explain the purpose and procedures of this study. Once you have completed this task, please close your folder. We will come back in a couple of minutes.

Appendix F. Pre-trial instructions script

We will now provide some legal instructions and ask you to read the trial transcript and answer some questions individually. Once you have completed these tasks, please close your booklet.

Gretchen Miller is charged with second-degree murder. The charge reads: You must find Gretchen Miller not guilty of second-degree murder unless the Crown has proven beyond a reasonable doubt that Miller is the person who committed the offence on the date and in the place described in the indictment. Specifically, the Crown must prove each of the following essential elements beyond a reasonable doubt:

1. That Gretchen Miller committed an unlawful act;
2. That Gretchen Miller's unlawful act caused Dennis Hughes's death; and
3. That Gretchen Miller had the intent required for murder.

If you are satisfied beyond a reasonable doubt of all these essential elements, you must find Gretchen Miller guilty of second-degree murder.

The defendant, Gretchen Miller claims to have been mentally ill at the time that the crime charged in the indictment was allegedly committed. Since the law does not hold a person criminally accountable for his or her conduct while mentally ill (not able to appreciate), not criminally responsible on account of mental disorder can be used as a defence against the crime charged. The accountability of defendant Miller at the time of the alleged offence is, therefore, a question which you must decide.

To be found not criminally responsible, the defence must prove by clear and convincing evidence that:

1. Gretchen Miller had a severe mental disorder or defect at the time that the acts constituting the crime were committed, and
2. As a result of this severe mental disorder or defect, Gretchen Miller was not able to understand what she was doing or to understand that what she was doing was wrong.

The defendant must prove that she was mentally ill at the time of the offence by clear and convincing evidence; that is, the defendant must show that it is highly probable that she was mentally ill at that time.

In making your decision, you may consider evidence of the defendant's mental condition before or after the crime charged and you may consider not only the statements and opinions of any experts who have testified, but also all of the other evidence received in the case.

If you find that the defendant committed the acts described in the essential elements of Second-Degree Murder, but that the defendant was not criminally responsible at the time that the acts were committed, you must find the defendant Not Criminally Responsible on account of Mental Disorder. Applying all of the other instructions given to you, you may find defendant Miller guilty if you decide Miller does not meet the requirements of Not Criminally Responsible on account of Mental Disorder. Even though the defendant has raised the issue of mental illness, the Crown still has the burden of proving all the essential elements of the offence beyond a reasonable doubt.

(Volunteer hands out trial transcripts and verdict forms) You may now read the trial transcript.

Appendix G. Trial transcript

Trial Transcript

R v. Miller

Alleged Crime: Second-degree murder

Victim: Dennis Hughes

Defendant: Gretchen Miller

D.O.B.: March 6th, 1989

Arrested: December 10th, 2014

Crown Opening Statement

On the night of December 10th, 2014, Gretchen Miller stabbed and killed Dennis Hughes. The facts meet the elements of second-degree murder; this is not in dispute by either side. The defendant further, by her own admission, knew that she was committing murder, and that murder is illegal. Thus, her action and admission satisfy all of the elements necessary to convict her of second-degree murder.

The reason we are here today is to determine whether the defendant is going to be held responsible for her actions. In pleading Not Criminally Responsible on account of Mental Disorder, the Defence actually assumes the burden of proof. It is true that the Crown always assumes the burden of proof in establishing guilt versus innocence, because the defendant is assumed to be innocent until proven guilty. However, the law has established that the defendant is assumed to be sane unless the Defence can prove otherwise. In other words, the burden for proving criminal responsibility rests with the Defence, not with the Crown. In this case, the Defence would have you believe that at the time of the offence, the defendant, Miller, was not capable of appreciating the quality of her act, of understanding that the act was morally wrong. But while you're listening to this testimony, I urge you to remember that the simplest explanation tends to be the right one. After you have seen the evidence you will understand that Ms. Miller is simply a violent person who snapped on her roommate during an argument that got so heated a neighbor overheard and called the police.

Although you will hear psychiatric testimony that, according to the words of the defendant, she was insane at the time of the offence, the Crown will show that actions speak louder than words in this case. You will hear testimony that the defendant stole money from the victim, and was preparing to skip town to avoid apprehension by the police. These are the actions of a person who knows she did something wrong but doesn't want to answer for it. Members of the jury, while Ms. Miller may have done an act that you and I believe only a sick individual could do, she was legally sane when she did so. Don't accept the Defence's farfetched fiction, the only evidence about which comes from a violent offender's claims, but instead hold her responsible. We ask you to return the only verdict appropriate in this case, Guilty.

Defence Opening Statement

Members of the jury, Gretchen Miller suffers from a very severe mental illness known as paranoid schizophrenia. The composed, seemingly rational person you see before you is a product of anti-psychotic medication. However, you have all heard the idiom, 'never judge a book by its cover.' To judge the inner workings of Ms. Miller's mind and mental illness based on her external appearance while she is on anti-psychotic medication is a grave mistake. You will hear medical testimony showing that the defendant has been positively diagnosed with paranoid schizophrenia.

Yes, it is true that Ms. Miller understood the legality of her act, but her actions were, in her mind, truly justified. The rationality of her belief, given her diagnosis of paranoid schizophrenia, is not the legal question to be decided here today. If a person kills another, and knows it is illegal but under an insane delusion that the salvation of the human race depends on it, then her action might be 'legally wrong', but it is not 'wrong' if we mean 'morally wrong'. The law is clear about this, members of the jury. Even if she knew it was against the law to kill another person, a mental disease caused her to think it was the only option. If you, the jury, feel that Ms. Miller, in her own mind, believed her actions were not morally wrong, then you must find that the defendant is Not Criminally Responsible on Account of Mental Disorder.

To return a verdict of Not Criminally Responsible you do not need to understand what she believed; surely, no sane person would believe that her loved ones have been replaced by alien imposters. What you do need to understand is why she believed what she did. The answer to that point is a severe mental illness or disease known as paranoid schizophrenia. It was only because of this mental illness that she stabbed Mr. Hughes. Keeping these facts in mind, the Defence calls on you to return a verdict of Not Criminally Responsible on account of Mental Disorder.

Crown Witness, Officer Mark Hanes

Crown: Can you please state your name and occupation?

Hanes: My name is Mark Hanes, and I am the police officer who was the first to arrive on scene on the day in question. I was also the officer that later arrested Gretchen Miller.

Crown: Can you please describe the events that took place on the evening of December 10th?

Hanes: At approximately 8:40 pm I responded to a 911 call from a neighbour about a disturbance at the apartment of Dennis Hughes and Gretchen Miller. When I arrived, the door was open, and the victim, Mr. Hughes, was lying on the floor in the kitchen. I could see that Mr. Hughes had lost a lot of blood, and I immediately called for medical assistance. Mr. Hughes was pronounced dead shortly after arriving at the hospital, and the cause of death was noted as multiple stab wounds to the neck and chest.

Crown: What happened next?

Hanes: We interviewed the neighbour who had called 911. We learned from this person that Mr. Hughes had a roommate, Gretchen Miller.

Crown: Was the defendant, Gretchen Miller, there at that time?

Hanes: No she was not.

Crown: When and where did you find Ms. Miller?

Hanes: At approximately 9:15 we found Ms. Miller at the home of her mother, Mrs. Miller.

Crown: Can you describe what happened next?

Hanes: We arrived at the home of Mrs. Miller, and identified ourselves as police officers. Mrs. Miller indicated that Ms. Miller was in her childhood bedroom. When we went to her room, it was evident the defendant was quickly attempting to pack some belongings. We let Ms. Miller know that we needed to ask her some questions pertaining to Dennis Hughes.

Crown: Did the defendant comply with your instructions?

Hanes: No she did not. Ms. Miller attempted to flee through the bedroom window, which was on the ground level of the house. At that time we apprehended Ms. Miller and took her in for questioning.

Crown: Did you find anything of note at Mrs. Miller's home?

Hanes: Yes. We found a butcher's knife, which we later identified as the murder weapon. Ms. Miller had cleaned the knife in the bathroom sink. We also recovered a wallet, which contained \$200 cash and several cards; it belonged to the victim, Mr. Hughes.

Crown: So, to summarize, Ms. Miller was attempting to pack her belongings, and to avoid capture?

Hanes: Yes.

Defence Cross-examination:

Defence: What was Ms. Miller's demeanor at the time you arrived at her house?

Hanes: She seemed frantic, and unsettled by sudden police presence.

Defence: Did she say anything?

Hanes: She shouted something to the effect of: "Get away from me, don't let them take me".

Defence Witness, Dr. Devin Cassady

Defence: Can you please state your name and occupation for the court?

Cassady: I'm Dr. Devin Cassady. I'm a psychiatrist working at the Forensic Psychiatric Institute.

Defence: What are your credentials?

Cassady: I am a medical doctor, and a Fellow of the Royal College of Physicians. I've been practicing psychiatry for over 20 years now.

Defence: Have you spoken extensively with the defendant, Gretchen Miller?

Cassady: Yes. I conducted a full psychiatric assessment of Ms. Miller.

Defence: What did you learn from this assessment?

Cassady: Based on a psychiatric and medical history, a standardized questionnaire, and my own more detailed interview, it is my professional opinion that Ms. Miller meets the diagnostic criteria for schizophrenia, paranoid type.

Defence: Can you describe for the courts what exactly 'schizophrenia' is?

Cassady: Schizophrenia is a severe brain disorder in which people interpret reality abnormally. Schizophrenia may result in some combination of hallucinations, delusions, and extremely disordered thinking and behaviour. Its origins are not yet fully understood by scientists, but its potential debilitating effects are well documented.

Defence: Could you please tell the jury some details about how someone is diagnosed with schizophrenia, and what that means?

Cassady: Well, first you must rule out other mental health disorders and determine that the symptoms aren't due to substance abuse, medication, or a medical condition. In addition, a person must have at least two of a specific set of symptoms outlined in the Diagnostic and Statistical Manual of Mental Disorder (also called the 'DSM'), and those symptoms would be present for most of the time during a one-month period, with some level of disturbance being present over six months. We look for things like delusions, hallucinations, disorganized speech (indicating disorganized thinking), and extremely disorganized behaviour.

Defence: What did you learn, during your assessment, about Ms. Miller's behaviour on the day in question?

Cassady: Ms. Miller suffers from what is called "Capgras Delusion", which can occur in patients with paranoid schizophrenia. The key feature of this delusion is that the patient believes that her loved ones have been replaced by identical looking imposters. Ms. Miller indicated to me that she believed that an alien imposter, who had transplanted a chip into her

brain, had replaced her roommate. This chip, she believed, was responsible for her hearing of Mr. Hughes's voice even when he was not present. She told me she suspected that aliens were conspiring to take over the planet, and that the Mr. Hughes imposter was attempting to extract information from her mind. She remarked to me that she began to suspect this was the case a couple of months prior, when she came home to find that Mr. Hughes had moved the TV to a different spot in the room.

Defence: In your discussions with Ms. Miller about the night in question, what did she tell you?

Cassady: She recalled that she and Mr. Hughes were talking in the kitchen, and that she heard a knock on the door. She believed that Mr. Hughes intended to take her away to a secret facility that night, and that she had to kill him to get away.

Defence: In your opinion, is Ms. Miller trying to mislead you into believing she has schizophrenia?

Cassady: No I do not. Ms. Miller presented with classic symptoms of schizophrenia, and in particular, Capgras Delusion.

Crown Cross-Examination

Crown: Are you an expert in deception, Dr. Cassady?

Cassady: No, I am not. But I have many years experience treating real illnesses, and the ability to detect malingering can be part of the job.

Crown: Is that because people sometimes lie, and try to trick their doctor into diagnosing them with an illness?

Cassady: It's a possibility, but to my knowledge it is not all that common. Capgras Delusion specifically is not necessarily well known to laypeople, so it wouldn't really be something one would fabricate easily.

Crown: But, wasn't your diagnosis just based on what the defendant told you, after some time had passed following the incident?

Cassady: I conducted a full psychiatric assessment. From this assessment, it is my opinion, based on many years of experience, that Ms. Miller's behaviour is consistent with a diagnosis of schizophrenia.

Crown: But in fact, you didn't even interview Ms. Miller until two full weeks after the crime had occurred, isn't that right?

Cassady: Yes that's correct.

Crown: Do you think that is enough time for someone to research the symptoms of schizophrenia, or research the insanity defence?

Cassady: I don't really know – I couldn't speak to the defendant's activities during that time.

Crown: Yet you can be confident about her mental state at the time of the crime? That she has this bizarre, specific delusion based only on a description of her activities?

Cassady: I conducted a full psychiatric assessment. That means that I had to take into account a lot of factors, not just Ms. Miller's word. I looked for things like certain speech patterns, emotional expression, thinking and perception spanning the months leading up to the incident and at the time of assessment. We're not just looking for what patients say, but how they behave – it's not as simple as just making up stories.

Crown: So, you're saying that the defendant believed aliens replaced her roommate and she had to escape quickly, but she still felt she had time to take the victim's wallet? Was that part of this 'delusion' as well?

Cassady: The point is that Ms. Miller's behaviours were erratic, frantic, because she was under the influence of paranoid delusions. For example, she indicated that if the police captured her, the aliens could get to her easily. Although the behaviour is irrational to a person who is well, it is reasonable to suspect that taking the wallet somehow played into those delusions.

Crown: Speaking of psychiatric history, to your knowledge, has the defendant ever been hospitalized for paranoid delusions before the incident?

Cassady: No she has not.

Crown: Thank you Dr. Cassady, that's all I have for you today.

Defence Re-direct

Defence: Dr. Cassady is it surprising to you that Ms. Miller would not have spent time in a psychiatric facility?

Cassady: Not necessarily. Among women, onset of schizophrenia typically occurs during mid to late 20's. Even if Ms. Miller began experiencing disturbances before the incident, she would not likely have understood the need to seek treatment. Having moved out of her family home, her family would not per se be able to intervene either.

Crown Closing Statement

Mr. Hughes woke up on December 10th, excited to finish his last exam before the holidays and soon go home to friends and family. He was a good student, and had many exciting plans in store, but instead, his life was cut short. I would like to remind you, ladies and gentlemen, that Ms. Miller does not deny intentionally ending Mr. Hughes's life. So, you don't need to take my word for it, instead you can take this information directly from the defendant. The disturbing truth is that Ms. Miller is a dangerous, cold-blooded killer. Frustrated with her roommate, she snapped and violently silenced Mr. Hughes. She knew that it was illegal, and she knew she would get in trouble. We can clearly see this because she took some quick cash from the victim, fled the scene, and even cleaned the murder weapon. Once she got caught red handed she had to come up with a good story. Members of the jury, do not fall for her fanciful story. What is likely: that the defendant suddenly experienced paranoid delusions even though we have no evidence of this? The defence is so insistent that these outlandish beliefs excuse Ms. Miller from criminal responsibility, and yet you didn't hear from a single witness who could attest to any strange behaviour in the months leading up to the crime. Or, is it more plausible that she is just a violent person who lost her cool when she argued with her roommate one too many times? Not one piece of evidence, besides the defendant's own account, that she had a mental disease at the time of the event. What we do have is overwhelming evidence of second-degree murder: the body of Mr. Hughes, a murder weapon, the defendant's belongings in Ms. Miller's room, and even a direct admission. While the psychiatrist you heard from might not be able to tell when someone is faking an illness, I have every confidence that you can, members of the jury, and that you will return the correct verdict in this case: Guilty.

Defence Closing Statement

This is a very tragic case, ladies and gentlemen; there is no doubt that. Ms. Miller also deeply feels the loss of her best friend and roommate, Mr. Hughes. The real culprit here is mental illness. We have shown you beyond a shadow of a doubt that Ms. Miller could not appreciate the nature and quality of her act. You heard testimony from a very experienced doctor describing an undisputed diagnosis of paranoid schizophrenia. The Crown would have you believe that because Ms. Miller's account is so bizarre, it can only be a piece of fiction. But after hearing Dr. Cassidy's testimony, you can understand that the reason these beliefs sound so far-fetched to you or me is because they came from a mind that is unwell. Don't fall into the trap of attempting to understand Ms. Miller's delusion. The Crown is trying to distract you from the real legal issue at hand: namely, whether Ms. Miller believed these things because of a mental illness. Trust in an expert's full assessment that was based on a lot more information than you've heard in this case. It was based on years of training, experience, and study of mental health. Let us not punish Ms. Miller for being the unlucky recipient of a mental disease that consumed her life and left her in fear for it. What Ms. Miller really needs to receive is psychiatric care from trained medical professionals who understand how the brain works. I trust you, members of the jury, to follow the law in this case. The law tells us that if a person could not appreciate the nature and quality of her act, then she will not be held criminally responsible, plain and simple. This doesn't mean she can just walk out of here, it just means that we can't blame her for something that was so beyond her control. This isn't a case of evil, but rather it is a case of illness. While Ms. Miller did not have the choice to act rationally at the time of the crime, you have a choice here and now; the rational one is to find the defendant Not Criminally Responsible on Account of Mental Disorder.

Appendix H. Individual verdict

1. How do you find the defendant?

Guilty

Not Guilty

Not Criminally Responsible on Account of Mental Disorder

Appendix I. Pre-deliberation instructions script

We will now instruct you on the law and then ask you to begin your deliberation. Again, if you no longer wish to participate at any point you, or you if you require any assistance, you can wave or simply exit the room and knock on the door of the room right next to us [motion to mirror].

I will now instruct you on the issue of mental disorder.

You will consider this issue only if you are satisfied beyond a reasonable doubt that Gretchen Miller stabbed Dennis Hughes, causing his death, and that she had the intent required for murder. If you are not satisfied beyond a reasonable doubt that Gretchen Miller committed the act, then she is entitled to an acquittal and there is no need to consider mental disorder.

A person is not criminally responsible if he or she suffered from a mental disorder at the time of the act and, as a result, was not capable of appreciating the nature and quality of her act or of knowing the act was wrong. The issue has been raised whether Gretchen Miller is exempt from criminal responsibility on this basis.

Every person is presumed not to suffer from a mental disorder. Exemption from criminal responsibility on this basis must be proved. I will now instruct you on a special rule on the burden of proof in relation to mental disorder.

Gretchen Miller has raised the issue of mental disorder. You will recall that I earlier instructed you that the burden of proof beyond a reasonable doubt was on the Crown and that Gretchen Miller was not required to prove anything. The issue of mental disorder is an exception to this rule. Gretchen Miller must prove that it is more likely than not that she is exempt from criminal responsibility due to mental disorder at the time the offence was committed. This is a lower standard than proof beyond a reasonable doubt.

To decide whether Gretchen Miller is exempt from criminal responsibility by reason of mental disorder, ask yourselves the following questions:

1. Is it more likely than not that Gretchen Miller was suffering from a mental disorder at the time of the act?
2. Is it more likely than not that Gretchen Miller's mental disorder made her incapable at the time either of (a) appreciating the nature and quality of the act or (b) knowing the act was wrong?

I will now review each of these questions with you.

First - Is it more likely than not that Gretchen Miller was suffering from a mental disorder at the time she committed the act?

A mental disorder is a disease of the mind. A disease of the mind is any illness, disorder, or abnormal condition that impairs the human mind and its functioning. A disease of the mind does not include a self-induced state caused by alcohol or drugs, or transitory mental states, such as hysteria or concussion.

This is the legal definition of disease of the mind and it is the definition you must apply, not any other definition that may have been used by counsel or the experts. I tell you as a matter of law that paranoid schizophrenia is a disease of the mind. It is for you to decide whether it is more likely than not that Gretchen Miller was suffering from paranoid schizophrenia at the time of the act.

If your answer to this question is yes, go on to the next question. If your answer is no, then Gretchen Miller is not exempt from criminal responsibility due to mental disorder.

Second - Is it more likely than not that Gretchen Miller's mental disorder made her incapable at the time either of:

- (a) appreciating the nature and quality of the act or,
- (b) knowing the act was wrong?

1. A person does not appreciate the nature and quality of an act if she does not know what she is doing, or does not foresee and understand the consequences of her act. The consequences refer only to the physical consequences, not the legal consequences.
2. Next, ask yourselves whether the mental disorder deprived Gretchen Miller of the capacity to decide rationally whether the act was wrong and, therefore, to make a rational

choice about whether to do it. “Wrong” means morally wrong, judged by the everyday standard of the ordinary person. It does not mean legally wrong, and it does not mean wrong according to Gretchen Miller’s own personal moral beliefs.

To reach the special verdict of not criminally responsible, it is not necessary that you find that Gretchen Miller was incapable by reason of mental disorder of both appreciating the nature and quality of her act and knowing that the act was wrong. As long as each of you finds that her mental disorder made her incapable of either one or the other, you do not have to agree on which one.

Ask yourselves whether it is more likely than not that Gretchen Miller’s mental disorder at the time made her incapable either of:

- (a) appreciating the nature and quality of the act or,
- (b) incapable of knowing the act was wrong.

If the answer to this question is yes, then you must find Gretchen Miller not criminally responsible by reason of mental disorder.

If you do not find that Gretchen Miller was suffering from a mental disorder, or if she was, it was not one that prevented her from appreciating the nature and quality of her act or from knowing it was wrong, you must still consider the evidence relating to mental disorder along with all the other evidence when you determine whether the Crown has proved the intent required for second degree murder beyond a reasonable doubt.

I will ask you to choose one juror to act as your foreperson. The foreperson will chair your discussions, and announce your verdict at the end of the case. Get to know each other a little before you choose your foreperson.

A verdict, whether of guilty, not guilty, or Not Criminally Responsible on Account of Mental Disorder, is the unanimous decision of the jury. To return a verdict on a count requires that all of you agree on your verdict. While your verdict on any count must be unanimous, your route to the verdict need not be.

You should make every reasonable effort, however, to reach a verdict. Consult with one another. Express your own views. Listen to the views of others. Discuss your differences with an open mind. Try your best to decide this case.

I am now handing you a verdict sheet. If you reach a unanimous verdict your foreperson should record it on your verdict sheet and notify me. We will come back to receive it.

If you cannot reach a unanimous verdict you should notify me in writing.

It is your duty to consult with one another and to try to reach a just verdict according to the law. Your foreperson will preside and assist you in the orderly discussion of the issues. You should each have the opportunity to express your own points of view without being unnecessarily repetitive. When you are discussing the issues, you should listen attentively to what your fellow jurors have to say. Approach your duties in a rational way and put your own points of view forward in a calm and reasonable manner. Avoid taking firm positions too early in your deliberations. Consider the views of your fellow jurors with an open mind before reaching your own decision.

We will not have a written transcript of the evidence available for you to review when you discuss your decision in this case. I think you will find that your collective memory of the evidence is good.

(To be read by research assistant) We will now ask you to begin your deliberation. Again, if you no longer wish to participate at any point, or if you require assistance, you can simply exit the room and knock on the door to the immediate right. We kindly ask that you refrain from the use of cell phones throughout the deliberation except in case of an emergency. I will come back in 45 minutes, or as soon as you reach a unanimous verdict.

(After 45 minutes)

Were you able to reach a unanimous verdict?

a. (If yes): How do you find the defendant?

b. (If no): In the event that a jury is unable to reach a unanimous verdict, the court usually declares a mistrial.

(Begin debriefing script)

Appendix J. Debriefing session script

Thank you all for your participation. This debriefing session will clarify the purpose of our study and why we are interested in this issue. The purpose of this study was to understand how jurors process and discuss information in a Not Criminally Responsible on Account of Mental Disorder trial. The trial transcript that you read was not an actual case, but instead was fabricated for the purpose of this study, and modeled after several different cases.

What were we trying to learn in this study?

Previous research has shown that jurors have somewhat negative views of the insanity defence, which might mean that they are hesitant to apply it, even in appropriate cases. However, because this trial was fabricated, there is no right or wrong verdict decision. Rather, we are interested in what types of information people think is relevant to discuss during deliberation. We are also interested in finding out if Canadian jurors have differing attitudes towards male and female defendants in Not Criminally Responsible on Account of Mental Disorder trials. To our knowledge, this is the first NCRMD deliberation study in Canada, and such a study can give us insight into how Canadians feel about the insanity defence and when they think it can be used.

We do not have specific predictions about your discussion or verdict decision, but we will look at whether there are any themes that emerge across 10 different mock juries. If you would like to receive an email summarizing these results, please write your email address on this sheet.

If you have any questions, I can answer them now, or you can email me questions any time. You may take this contact information home with you in case you have any questions about this study, about ethical concerns, or any other concerns.

The study is now complete. Thank you very much for participating; we value your time and opinions.

(Provide compensation and contact information sheet).

Appendix K. Group verdict form

GRETCHEN MILLER,

Defendant

DATE: _____

TIME: _____

CASE: R v. Miller

VERDICT FORM

COUNT ONE

Murder in the Second-Degree

_____ Guilty

_____ Not Guilty

_____ Not Criminally Responsible

on Account of Mental Disorder

_____ Unable to reach a verdict

Please sign and date the form and tell the assistant that
the jury has reached a verdict:

DD/MM/YYYY

Initials of Foreperson