

System Justification and Normative Influence: Jury Decision-Making in a Police Shooting Trial

Roxana Ehsani-Moghaddam

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

Master of Arts

in

Psychology

Carleton University

Ottawa, Ontario

©2021

Roxana Ehsani-Moghaddam

**Abstract**

This study investigated whether normative influence (i.e., arguments to conform to the group) was related to jurors' system justification (SJS) beliefs (i.e., beliefs that justify a racially disparaging societal status quo) and time pressure during the deliberation phase of a mock criminal trial. Given the traumatic colonial context that exists between Indigenous communities in Canada and the police, as well as the current disproportionalities of Indigenous people in the Canadian criminal justice system, jurors with lower SJS may be compelled to use normative influence to persuade other jurors to conform to their verdict preference for an Indigenous defendant who raised a claim of self-defence for the killing of a police officer. Further, past research has found a relation between time pressure and normative influence. Thus, lower SJS and time nearing the end of the deliberation were hypothesized to be related to greater normative discussion content. Deliberations were transcribed, coded, and analyzed for 11 mock juries ( $N = 83$  jurors) in a simulated first-degree murder trial. Findings did not support a relationship between normative influence and either SJS or time. This research may bear implications for our understanding of jury decision-making processes and how to instruct jurors.

## Acknowledgements

I would first like to acknowledge that Carleton University is situated on unceded and traditional Algonquin territory. Prior to the colonization of Turtle Island, the Indigenous peoples of this territory had their own established systems of education, government, law, order, and justice. I live on this land as a guest.

This thesis was a product of two years' worth of scholarship, meaningful mentorship, and invaluable friendship to which I am deeply grateful.

I must firstly express my sincerest appreciation to my supervisor and mentor, Dr. Evelyn Maeder: a trailblazer in her field who offered me her guidance, supported and enriched my ideas along the way, and never failed to make time for me despite all the people who depend on her every day. It is an understatement for me to say that you have inspired me and profoundly impacted my life. The gratitude I feel towards all that you have done for me cannot be so simply expressed.

With Dr. Maeder's supervision came the opportunity to become acquainted with my lab mates and cohort friends. It has been such a pleasure learning from you. You are all brilliant individuals. To Cassandra and Logan, thank you so much for your time and patience in helping me complete the coding portion of my thesis.

I must also highlight my sincerest gratitude towards Dr. Craig Leth-Steensen. There were countless times when I reached out to you for statistical help in a time crunch and under stress. Not only did you always answer my questions, but you always went above and beyond in ensuring that my statistical issues were resolved. Your expertise and graciousness helped make this thesis possible.

To my committee members, Dr. Kevin Nunes, Dr. Craig Bennell, Dr. Chad Danyluck, and Dr. Melanie Adrian, thank you for granting me your time and expertise.

To my loved ones, thank you for never failing to support me through all my endeavors. To my mother, Forough, and father, Behrouz, you have both shaped me into the woman I am today. You showed me what hard work and dedication look like. I love you both with all my heart and could never put into words how appreciative I am of all the sacrifices you have made for me in your life. I hope I have made you proud. To my partner, Alex Moore, thank you for your never-ending love and encouragement. To my dear friend and colleague, Heather Burke, you never failed to soothe my rumination and to bring joy the most stressful of times. To my beloved friend and prior colleague, Alice Shen, the passion and creativity you put into your work motivates me to emulate the same in my own work. Finally, to my sweet grandmother, Talat Bayani, you are my life's greatest inspiration.

The opportunity to complete this thesis is beholden to each of you.

**Table of Contents**

Abstract.....	ii
Acknowledgements.....	iii
List of Tables.....	v
List of Figures.....	vi
List of Appendices.....	vii
Introduction.....	1
Social Influence in Jury Deliberations.....	4
Influence Mode.....	6
System Justification Beliefs and Normative Influence.....	9
System Justification.....	12
Temporal Patterns of Influence.....	15
The Current Study.....	17
Methods.....	18
Participants.....	18
Materials.....	19
Trial Stimulus.....	19
Demographics Questionnaire.....	20
System Justification Scale.....	21
Procedure.....	22
Coding Deliberation Content.....	23
Results.....	30
General Characteristics.....	30
Assumption Testing.....	34
Counterbalancing.....	34
Hierarchical Linear Model.....	35
Discussion.....	40
Implications.....	50
Limitations.....	51
Future Directions.....	58
Conclusion.....	60
References.....	66

**List of Tables**

Table 1- Racial Breakdown for Mock Juror Participants.....	19
Table 2- Coding Scheme by Kaplan and Miller (1987).....	24
Table 3- Worst-Case Interrater Agreement on Individual Utterance Type Categories .....	27
Table 4- Best-Case Interrater Agreement on Individual Utterance Type Categories .....	28
Table 5- Worst-Case Interrater Agreement on Individual Content Type Categories .....	29
Table 6- Best-Case Interrater Agreement on Individual Content Type Categories.....	29
Table 7- General Characteristics of the Jury Groups .....	31
Table 8- Demographic Breakdown of the Jury Groups .....	32
Table 9- Utterance and Content Code Frequencies by Jury Group.....	33
Table 10- Estimates of Predictor Fixed Effects .....	37
Table 11- Three-Level Analysis Variance Components .....	38
Table 12- Two-Level Analysis Estimates of Predictor Fixed Effects .....	39
Table 13- Two-Level Analysis Variance Components .....	39
Table 14- Deliberation Content Proportions Compared to Hansen et al. (1993).....	44
Table 15- Ratio of Normative to Total Utterance Averages across Time Blocks Compared to Hansen et al. (1993) .....	49

**List of Figures**

Figure 1a- Scatter Plot of Normative to Total Utterances by Time Block .....	62
Figure 1b- Scatterplot of Ratio of Normative to Total Utterances by SJS Score.....	63
Figure 2a- Quantile-Quantile Scatterplot of Ratio of Normative to Total Utterances .....	64
Figure 2b- Quantile-Quantile Scatterplot of Arcsine Transformed Ratio of Normative to Total Utterances .....	65

**List of Appendices**

Appendix A: Demographics Questionnaire ..... 75  
Appendix B: System Justification Scale ..... 78  
Appendix C: Three-Level Hierarchical Linear Model Equations ..... 79

### System Justification and Normative Influence: Jury Decision-Making in a Police Shooting Trial

Indigenous people are overrepresented in all stages of the Canadian criminal justice system for both youth and adult offenders (Corrado et al., 2014; Department of Justice, 2019). For instance, while Indigenous people comprised 4.9% of the Canadian population according to the 2016 Census (Statistics Canada, 2017), they made up 26.4% of the total federal prison population in 2016-2017 (Zinger, 2017). The Indigenous prison population has been on a steady increase, such that Indigenous people in Canada made up 30.04% of the total federal prison population by 2020 and are predicted to comprise 33% of the federal prison population over the following three years (Kingsley, 2020). According to the Correctional Investigator, Dr. Ivan Zinger, surpassing this 30% figure is indicative of an “Indigenization” of the Canadian correctional system: “nothing short of a national travesty” (Kingsley, 2020).

Among researchers who study Indigenous overrepresentation in the criminal justice system, there is growing consensus that said overrepresentation is not a product of one single factor, such as disproportional offending or discriminatory police practices, and instead is likely a complex interplay of many factors: over-policing, biased laws, a greater level of offending, and social disorganization (Blagg, 2012; Corrado et al., 2014). Given this complex interplay of factors, the overrepresentation of Indigenous people in the criminal justice system must be discussed in the context of an inter-generationally traumatic, multi-century-long history of colonialism and ethnic cleansing to which Indigenous peoples in Canada were/are subject (Corrado et al., 2014).

However, jurors, the individuals who are responsible for verdict decision-making if a defendant chooses a trial by jury, may differ in their awareness of Indigenous issues and history. In one study by Schaepli and colleagues (2018), 42,916 first-year Ontario university students



scored an average of 24.28% on a test co-designed with First Nations, Métis, and Inuit educators on the knowledge they believed that students should have regarding colonialism and its relationship to Indigenous people and modern Canadian society. Schaepli and colleagues (2018) also found a standard deviation of 16.06% in test scores among this student sample, suggesting that these first-year university students differed greatly from one another on their knowledge base on Indigenous issues. Schaepli and colleagues' (2018) findings underscored the dearth of knowledge many Canadians have about the issues Indigenous Peoples of Canada face.

Juries are not composed of twelve blank slates, as all jurors bring distinct backgrounds to the deliberation room (Fraher, 1987). Given findings that Canadians differed greatly from one another on knowledge about Indigenous issues (Schaepli et al., 2018), one may expect that Canadian jurors also bring a diverse range of knowledgeability on Indigenous issues to the courtroom. Greater awareness of societal inequity has been related to fewer beliefs that support the societal status quo, as individuals who are aware of a disparaging system will be less likely to endorse beliefs supporting the legitimacy of that system (Saguy et al., 2008). Said beliefs, called system justification beliefs, serve to justify and maintain a racially disparaging social status quo (Jost et al., 2004). System justification beliefs strengthen social order legitimacy at the group level (Jost et al., 2004). Individually, these beliefs improve one's satisfaction with the social order in which one exists (Jost et al., 2003). Those who are high in system justification beliefs may not express concern for those who are disadvantaged by the system as some may not be aware that such a disadvantage exists. Given that a jury's verdict is intended to represent the judgement of the community, and communities are composed of individuals with diverse backgrounds (Fraher, 1987), jurors will necessarily vary on the extent to which they endorse

system justification beliefs. This study thus aims to examine the relation between the degree to which a juror endorses these beliefs and their deliberation discussion content.

Some researchers have tackled analyzing juror discussion content by audiotaping or videotaping mock jury deliberations and coding their statements to examine how deliberation content generates a group decision (e.g., Hastie et al., 1983; Holstein, 1985; Horowitz, 1985; Kaplan & Miller, 1987). Indeed, jury deliberations are ideal venues for studying various social psychology constructs, including influence, conformity, persuasion, and group norms, as juries place individuals of differing backgrounds in a group where they must work together to reach a unanimous decision. The goal of reaching a unanimous decision is a condition that could facilitate normative influence: arguments that aim to persuade an individual to conform to the decision preferences of others, thus facilitating group harmony and cohesion (Kaplan & Miller, 1987). In contrast, informational influence can be defined by arguments that facilitate the acceptance of information presented by others as reality. Conformity due to informational influence is attributed to the presentation of relevant arguments and factual evidence about the issue at hand (Kaplan & Miller, 1987).

In addition to juries' unanimous decision rule, some researchers have proposed that normative influence may be facilitated through a motivation to prevent the miscarriage of justice (Hansen et al., 1993). To my knowledge, the concept of a fear of injustice has not been operationalized, measured, or clearly defined. However, for the current purposes, fear of injustice may be described as an individual's wariness that another individual or group may experience unjust treatment or incur unfair penalties due to inequities within a societal system. Although system justification beliefs may not be equated to a fear of injustice, system justification beliefs may be one way to conceptualize fear of injustice for the current study given

the mock trial's context: an Indigenous defendant claiming self-defence against a police officer in the Canadian criminal justice system, within which Indigenous people are overrepresented. Speculatively, jurors who believe that the current Canadian system is racially disparaging may have more reason to fear that an injustice may occur than jurors who believe that the system is just and legitimate. Thus, the current study's preliminary investigation of system justification and normative influence in jury deliberations may set the stage for future studies to examine fear of injustice and social influence more directly.

### **Social Influence in Jury Deliberations**

Largely influenced by classic research paradigms pioneered by Sherif (1935) and Asch and Guetzkow (1951), the study of social influence is fundamental to the field of social psychology. Some psychologists have defined social psychology as the study of influence, where the cognitions and behaviours of individuals are influenced by the presence of others (Allport, 1954). Historically, the majority of research in influence has focused on conformity (Nail, 1986). Conformity can be defined as the act of changing one's behaviour or beliefs to align with the behaviours or beliefs of others (Cialdini & Goldstein, 2004), and it arises when an individual's position conflicts with another individual's expressed position (Allen, 1965).

In his studies examining the autokinetic effect, Sherif (1935) found that participants conformed to group norms in the presence of ambiguous stimuli. When participants were asked to estimate the distance that a projected spot of light moved over several trials, individuals in the group would consistently conform to the estimate of the majority. Sherif (1935) concluded that people would prefer to come to a group agreement than make an individual judgement. However, to account for the ambiguity of Sherif's (1935) paradigm, where there was no singular correct answer for the distance the spot of light travelled, Asch and Guetzkow (1951) devised a

paradigm in which participants in groups of eight with seven confederates would make a judgement about the length of a line. Confederates in the study would match a target line to one of three comparison lines, and the correct answer would be obvious as differences between the target line and incorrect comparison lines were large (between .5- and 1.75-inches difference between lines). However, the confederates in the study would answer incorrectly before the participant answered, thus creating a dilemma in which the participant's internal position would be in radical conflict with the group's external position. Asch and Guetzkow (1951) found that one-third of participants yielded to the majority, later explaining that they believed they may have been experiencing a perceptual distortion, did not have access to the same information as the other group members, or simply did not want to disagree with the majority. The findings by Sherif (1935) and Asch and Guetzkow (1951) thus underscored the power of social influence as they revealed that individuals are highly motivated to conform to the majority both in ambiguous and perceptually obvious situations.

Distinct from Sherif's (1935) paradigm, Asch and Guetzkow's (1951) study of conformity allowed for the internal state of the subject to be considered, distinguishing between public and private agreement with the majority. Researchers often distinguish between three types of conformity: conversion, or "true conformity", is the true internalization of another individual's position; compliance, or "expedient conformity", is the public agreement, but private disagreement with a position; and congruence is an existing public and private agreement with an external position (Nail, 1986). Conversion and compliance can also be considered types of movement conformity, in which an individual changes their position to align with an external position (Willis, 1963). Movement conformity can be distinguished from congruence conformity, in which no individual change is needed to conform (Willis, 1963). The present

study is specifically concerned with the process of movement conformity, or how jurors influence each other towards conformity through different influence modes.

### **Influence Mode**

Several studies have investigated the discussion content of jury deliberations and have found them to consist primarily of three different types of content: normative arguments, informational arguments, and procedural-legal discussion (Brodbeck et al., 2007; Kaplan et al., 1994; Tanford & Penrod, 1986). Normative influence involves arguments or pressure to conform in implicit or explicit forms (Kaplan et al., 1994). Under normative influence, individuals with opposing viewpoints tend to be motivated to conform to majority viewpoints as a means of pleasing others, gaining social approval, or avoiding peer rejection (Brodbeck et al., 2007). Informational influence involves arguments to accept information as factual (Kaplan et al., 1994). Under informational influence, changes in beliefs tend to result from the re-evaluation of one's opinions in light of newly learned information (Brodbeck et al., 2007). In essence, normative influence tends to produce compliance, while informational influence often results in conversion (Nail, 1986). Although not a mode of influence, procedural-legal discussion content involves discussions about the trial proceedings (i.e., deliberation guidelines, verdict reaching, and courtroom procedure) and legal facts (i.e., reasonable doubt, precedents, legal terminology, and criminal code; Tanford, & Penrod, 1986).

Recent research on influence modes in jury decision-making found interesting results on the effect of influence modes on movement conformity. Garcia and colleagues (2021) employed a 2 X 2 experimental design in which participants were placed in conditions of either private or public deliberation and majority or minority opinion faction. The private condition was intended to control against the presence of normative influence, whereas the public condition was

intended to facilitate both forms of influence. In examining mock jurors' voting behaviours and their ratings of evidence, Garcia and colleagues (2021) found that fewer than 22% of jurors changed their vote due to changes in their beliefs about the evidence, such that a small portion of movement conformity (i.e., changing a public position to align with the majority) was related to true belief change (i.e., conversion). Further, they found that jurors changed their votes at equal rates in public compared to private deliberations. These findings led Garcia and colleagues (2021) to conclude that even anonymous group decision-making conditions may not eliminate the presence of normative influence. Thus, not only was normative influence difficult to prevent in a group decision-making setting, but it may possibly be more attributable to juror agreement in deliberations than informational influence (Garcia et al., 2021).

When decisions must be made in groups, influence modes may either take the form of evidence relevant to the discussion question (i.e., informational influence), or arguments regarding the preferences of others within the group (i.e., normative influence; Kaplan et al., 1994). According to Kaplan and Miller's (1987) model of conditions determining influence processes in groups, two conditions – issue type and interactive goal type– determine whether informational or normative influence modes are used.

### ***Issue Type***

The types of issues on which groups make decisions may be considered as a continuum, with intellective issues being on one end of the continuum and judgmental issues on the other (Kaplan & Miller, 1987). Intellective issues are those for which there is a conclusive correct answer. Conversely, judgmental issues may be subjective and do not have a conclusive correct answer. Topics that involve intellective issues are thought to elicit informational influence as they require arguments based on facts and have a correct answer (Kaplan & Miller, 1987).

Judgmental issues are thought to elicit normative influence as moral issues may not have a clear answer and would involve satisfying group consensus (McGrath, 1984).

In the context of the criminal trial stimulus used in this study, jurors may perceive an intellectual issue to be whether the defendant shot the police officer, which can be confirmed by factual evidence presented in the trial video. Conversely, jurors may perceive the question of whether the defendant shot the officer in self-defence to be a judgmental issue, as the evidence is less definitive on whether the defendant was necessarily acting to preserve his life. Judgmental issues thus allow for greater speculation and moral discussion as jurors may not be satisfied in their resolution with evidence alone. Indeed, in a study by Kaplan and Miller (1987), content analysis of mock jury deliberations revealed that intellectual issues elicited more informational than normative influence, while normative influence was more common when juries discussed a judgmental issue. Thus, one may expect jurors to resort to a greater proportion of normative arguments in a discussion of whether an Indigenous defendant was “reasonable” (s.34 of the Criminal Code) in his use of self-defence when shooting a police officer.

### ***Interactive Goal Type***

In addition to the two issue types, there are two main group interactive goals: task goals and group goals. Task goals involve group discussions that emphasize the task at hand and the production of a factually correct decision (Kaplan & Miller, 1987). Group goals involve discussion conditions stressing concern for the group and its relations, where the best decision satisfies the most members and functions to enhance harmony and cohesion (Kaplan & Miller, 1987). Normative and informational arguments should be expected to be most effective in influencing group member judgements when the influence type is congruent with the relevant

goal type: informational influence will be most effective for task goals and normative influence will be most effective for group goals (Kaplan & Miller, 1987; Rugs & Kaplan, 1993).

Jury deliberations are useful settings for examining influence modes as both informational and normative influence modes are anticipated given the presence of both task and group goals (Hansen et al., 1993). Firstly, jury deliberations involve task goals favouring informational influence through the presence of facts (e.g., evidence and testimony) and the law (e.g., the burden of proof, judge's instructions, and legal procedures; Hansen et al., 1993). Jury deliberations may also involve group goals that elicit normative influence through the presence of other motivations, including reaching a verdict with which all jurors agree, the prevention of justice miscarriage, and meeting social expectations (Hansen et al., 1993; Hastie et al., 1983).

Most relevant to the present study is the motivation to prevent justice miscarriage, where greater intent to prevent justice miscarriage (eliciting group goals) has been proposed as potentially being associated with normative influence (Hansen et al., 1993). Given the racialized context of this mock trial, jurors who endorse fewer system justification beliefs may be more inclined to believe that the Canadian racial status quo disproportionately disparages Indigenous people, resulting in potential injustice against an Indigenous defendant. This raises the question of whether jurors who believe that a system is unjust will be more likely to engage in normative influence. This study aimed to investigate this question by analyzing the deliberation content of jurors' discussions leading up to the final group verdict.

### **System Justification Beliefs and Normative Influence**

The perception of a system as illegitimate may be the catalyst that empowers individuals to amend injustices perceived to be produced from said system through emotions that instill a sense of power, such as anger (Martorana et al., 2005). Speculatively, a juror who believes that



there are systemic racial disparities in Canada may be more likely to believe that a miscarriage of justice must be prevented against an individual whose race is disparaged by the status quo.

Therefore, system justification beliefs will be investigated as being potentially related to a jurors' greater use of normative arguments to try to sway fellow jurors towards their verdict position.

The mock trial video used in this study included an Indigenous defendant who shot a police officer during a police use of force encounter. Use of force encounters have garnered considerable attention in the media as cellphones allow bystanders to easily record and upload moments from these encounters (Brown, 2016). Further, the most recent Black Lives Matter protests that began as a reaction to the murder of George Floyd have placed the discussion of race and police brutality at the forefront of public attention<sup>1</sup>. However, while these protests began in the US and later spread to Canada and many other parts of the world, the issue of criminal justice and race is not unique to the US context.

Indeed, tensions between police and Indigenous people in Canada have resulted from many decades of unjust practices between the Government of Canada and Indigenous communities, often enacted by police. For instance, the Royal Canadian Mounted Police enforced cultural cleansing policies at the direction of the Canadian government by returning runaway Indigenous children to residential schools, a reality that continues to fuel tensions between Indigenous communities and police to this day (Kiedrowski et al., 2017). While some attempts have been made to address the horrific historical context of Canada's treatment of Indigenous people and the resulting tensions with police, including the implementation of the First Nations Policing Program (FNPP) in 1992, Indigenous people continue to be grossly

---

<sup>1</sup> The data being analyzed in the present study were collected prior to May 2020. The acceleration of the Black Lives Matter movement and the protests calling for justice for George Floyd did not precede the present study's data collection and thus would not have affected jurors' perceptions of police use of force encounters.

overrepresented in the Canadian criminal justice system (Kiedrowski et al., 2017). For example, Indigenous people comprised 4.9% of the Canadian population in 2016 (Statistics Canada, 2017), but they accounted for 26% of all provincial and territorial correctional service admissions (Department of Justice, 2019).

Further, Indigenous people are also overrepresented in police use of force incidents. In a report examining police use of force data in Ontario, Canada from the Special Investigations Unit (SIU), a civilian law enforcement agency, Wortley (2006) found that Indigenous people were overrepresented compared to their population size and other race categories. For example, Wortley (2006) found that while Indigenous people made up 1.7% of the provincial population at the time, they made up 7.1% of all civilians involved in SIU investigations, such that they were 4.2 times more likely to be involved in an SIU investigation than their population size would predict. Of the SIU cases involving Indigenous civilians, 57% were cases in which the harm in question was directly caused by police actions: a proportion that was second highest of all race categories, with Black Canadians having a higher proportion (64%). Further, the rate of cases involving Indigenous civilians in which harm was directly caused by police was 4.7 and 6.1 times the provincial and White rates, respectively. When considering police use of force specifically, Indigenous people represented 8.3% of all SIU investigations involving use of force incidents: 5.0 and 6.2 times the provincial and White rates, respectively. Indigenous people were also overrepresented in police fatal use of force encounters, such that they represented 8.1% of all police use of force deaths, despite comprising 1.7% of the provincial population at the time: a figure 4.8 and 7.2 times greater than the provincial and White rates, respectively. As a whole, these deeply concerning findings underline the gross overrepresentation of Indigenous people in fatal police encounters.

Given that the status quo for Indigenous people in Canada has been one marked by trauma and injustice, jurors who are aware of these injustices may hold fewer beliefs that justify the status quo of race relations in Canada. Indeed, one study found that when individuals were made aware of group disparities, they spoke more on power relations during ingroup discussions in hope of ending disparities through the promotion of equality between groups (Saguy et al., 2008). While awareness of Indigenous issues was not measured in the present study, jurors who believe that the Canadian societal status quo is racially disparaging may be more likely to discuss inequitable norms when an Indigenous person is involved in a criminal case. Thus, lower endorsement of system justification beliefs may be an individual characteristic related to differences in normative discussion content in a trial with an Indigenous defendant.

### **System Justification**

A juror who scores low in system justification, such that they reject the racial status quo in Canada, may be more likely to speak on societal norms of racial inequity in a police shooting trial with a racialized defendant. System justification theory posits that there exists a psychological motive for individuals to defend and justify the societal status quo (Jost & Banaji, 1994). For the individual, system-justifying beliefs function to decrease negative affect by increasing one's satisfaction with the status quo within which they live (Jost et al., 2003; Kluegel & Smith, 1983). At the group level, system justification functions to strengthen the legitimacy of the existing social order (Jost et al., 2004). The acceptance of system-justifying beliefs is associated with heightened in-group favouritism among members of advantaged groups (e.g., Whites) and heightened out-group favouritism among members of disadvantaged groups (e.g., Blacks; Jost et al., 2004).

The tendency to accept system-justifying beliefs may be predicted by several antecedent factors. Firstly, individuals with a heightened sense of needing to manage uncertainty may be more likely to endorse conservative ideologies and thus also more likely to endorse similar system-justifying ideologies (e.g., right-wing authoritarianism, social dominance orientation, economic system justification; Jost & Hunyady, 2005). More specifically, system justification has been related to uncertainty avoidance, need for order, intolerance of ambiguity, perception of a dangerous world, and fear of death (Jost & Hunyady, 2005). Alternatively, individuals higher in cognitive complexity and openness to new experiences are less likely to endorse system-justifying beliefs (Jost et al., 2003).

Unintuitively, group membership in an advantaged or disadvantaged group is not a reliable predictor of system justification, as disadvantaged group members also tend to endorse system-justifying beliefs (Jost et al., 2004). When two interdependent entities of unequal power dynamics must coexist, such as racialized groups in heterogeneous societies, peaceful coexistence is dependent on the reconciliation of power asymmetry through the acceptance of the status quo (Vallacher et al., 2011). Over time, both entities must internalize said status quo or risk social destabilization. Upsetting the group relationship equilibrium would reduce the coherence of the entities' attitudes towards one another and undermine the existing script between the two entities. Thus, any challenge to the status quo, such as an attempt to reinstate power to the disadvantaged entity, would tarnish the peace between the entities – an outcome that is potentially harmful to both parties (Vallacher et al., 2011). For this reason, members of disadvantaged groups may justify a power-imbalanced system to perpetuate feelings that life is fair, thereby reducing feelings of conflict with the advantaged group. When disadvantaged groups publicly express system dissatisfaction, assessments of unconscious system justification

suggest that their attitudes are consistent with the power asymmetric status quo (Jost et al., 2002). Consequently, members of disadvantaged groups who accept the status quo and reject egalitarian alternatives have been found to suffer from heightened levels of depression and lower levels of self-esteem (Jost & Thompson, 2000).

Rousing people from their acceptance of the status quo is difficult due to the presence of a system justification motivation, in which individuals are motivated to defend and justify their social systems (Jost et al., 2010; Jost et al., 2008; Kay et al., 2009). However, cultivating a sense of injustice and moral outrage is necessary for inciting collective action and social change (Jost & Kay, 2010; Wakslak et al., 2007). As demonstrated by Wakslak and colleagues (2007), system-justifying beliefs reduce emotional distress (e.g., negative affect, frustration, and guilt) because accepting the status quo reduces moral outrage. Moral outrage is a motivational prerequisite for reconciliation of injustice (Deutsch, 2006). Indeed, Wakslak and colleagues (2007) found that reductions in moral outrage, as induced by “rags-to-riches” themed stories, led to participants experiencing less motivation to help the disadvantaged, as indicated by reduced support for public benefit programs. Moral outrage thus results in efforts to remedy miscarriages of justice, such that a reduction in moral outrage due to a greater presence of system-justifying beliefs contributes to the withdrawal of support for social change (Jost & Hunyady, 2005).

The present study used a modified version of the system justification scale (SJS) to measure mock juror attitudes endorsing the racial status quo in Canada (Jost & Kay, 2005). Given findings that perceptions of inequity between groups were related to a greater willingness to speak on disparities as a means of promoting equity (Saguy et al., 2008), those who score lower on the SJS may have a greater motivation to discuss societal norms in a mock police shooting trial with an Indigenous defendant than those who score higher on the SJS, as those

who score higher perceive the system to be more just. A motivation to prevent injustice has been discussed as potentially eliciting normative discussion content, as fear of injustice is motivated by a desire to maintain social harmony and positive intergroup relations (Hastie et al., 1983). Racially disparaging systems create injustices, and thus those who believe that a system is not just may be more likely to be wary of a miscarriage of justice. While fear of injustice cannot be operationalized by the SJS, system justification beliefs may serve as a preliminary conceptualization of fear of injustice for the purposes of the current study.

### **Temporal Patterns of Influence**

Besides system justification beliefs, time pressure is another factor that may be related to a juror's likelihood of using normative or informational influence while deliberating. According to Kaplan and Miller's (1987) model of conditions determining influence processes in group discussion, informational and normative content should shift temporally throughout a decision-making process.

In support of this model, Hansen and colleagues (1993) conducted a study with 36 undergraduate participants in 6-person same-sex juries who acted as mock jurors for a realistic trial involving criminal negligence causing bodily harm. The discussion content of the deliberations was categorized and coded based on the coding scheme introduced by Kaplan and Miller (1987), where each speech turn (i.e., a sentence, sentence fragment, or series of sentences from one speaker) was coded as either informational, normative, procedural-legal, or uninterpretable. Deliberations were transcribed in three segments: the beginning segment as the period from 5 to 15 minutes after the start of the deliberation; the middle segment as the 10 minutes in the midpoint of the total deliberation; and the end segment as 10 minutes before the final verdict. Their results indicated that informational content comprised a higher percentage of

the discussion compared to normative content in all three deliberation segments. However, the proportion of normative to informational content increased from the beginning segment to the end segment: while the mean proportion of normative to informational speech turns in the beginning segment was .10, this proportion increased to .58 in the end segment. Further, the proportion of procedural-legal to informational discussion content also increased from .07 in the beginning segment to .81 by the end segment of the deliberation. While the duration of the mock deliberations varied between the juries, with the longest session being seven times the duration of the shortest, all six juries showed equivalent sequential patterns of discussion content proportions, with the proportion of normative and procedural-legal content increasing across the three distinct time segments (Hansen et al., 1993).

These results were in line with findings reported in Hastie and colleagues' (1983) book, *Inside the Jury*, indicating overall decreases in evidence-based discussion content and increases in normative and legal discussion content throughout the deliberation duration. Further, Kelly and colleagues (1997) found similar results in their study on the effect of time pressure on the emergence and effectiveness of either normative or informational influence. They found that normative influence was more likely to emerge and was more effective for groups under time pressure, such as the time pressure that may be present as a jury deliberation session is nearing an end. They also found that informational influence was more likely to emerge and be more effective under low time pressure conditions, such as at the beginning of a deliberation session when there is ample time remaining. Therefore, when examining the relation between system justification beliefs and normative influence in the current study, temporal changes in normative influence were accounted for by a time variable.

### **The Current Study**

The current study investigated the relation between system-justifying beliefs and normative discussion content in a jury deliberation for a trial with an Indigenous defendant. Normative discussion involves arguments of conformity and encourages group harmony (Kaplan et al., 1994). Informational discussion involves arguments based on factual information (Kaplan et al., 1994). Further, discussion about societal norms is considered normative discussion, while discussion about the facts of a trial is considered informational discussion (Kaplan & Miller, 1987). While several circumstantial and individual factors may be related to one's propensity to use normative influence, a motivation to prevent justice miscarriage has been cited as one potential factor that may be related to the use of normative influence (Hansen et al., 1993; Hastie et al., 1983). Given the criminal court system would necessarily be part of a racially disparaging social structure, jurors with lower endorsement of system justification beliefs may be wary of a racialized defendant being treated unfairly in a system they perceive as unjust. Indeed, greater awareness of inequitable norms has been found to be related to a greater likelihood of speaking on those norms (Saguy et al., 2008).

Time pressure may also influence the discussion content of group decision-making. Past research has shown that the proportion of normative influence discussion increased nearing the end of a deliberation session (Hansen et al., 1993) and that greater normative influence was exerted in the presence of time pressure (Kelly et al., 1997). Therefore, time was also considered in this study.

### **Hypothesis 1**

Greater endorsement of system justification beliefs can be operationalized by higher scores on the SJS scale. Normative content can be operationalized by a greater proportion of



normative discussion utterances to total utterances by an individual juror and within a jury cluster. A ratio of normative to total utterances provides the proportion of normative influence in a juror's speech, thus accounting for jurors who may have simply spoken more. The first hypothesis of this study predicted a negative association between SJS scores and normative discussion content, where jurors with lower scores on the SJS were predicted to discuss more normative content.

### **Hypothesis 2**

The 1-hour deliberation period was divided into three 20-minute time blocks representing the beginning, middle, and end of the deliberation. More time pressure was expected in the last time block compared to the first and second time blocks, with the least amount of time pressure expected in the first time block. Therefore, this study's second hypothesis posited that discussion in the third time block would be predictive of more normative discussion content for individual jurors and jury clusters compared to the first and second time blocks, with the first time block being predictive of the least normative discussion content.

### **Methods**

The present study analyzed a portion of data that had been collected through a larger study conducted at the Legal Decision-Making Laboratory at Carleton University in Ottawa, Ontario, Canada. This study received ethics clearance from the Carleton University Research Ethics Board-B (CUREB-B Clearance # 106258).

### **Participants**

A total of 83 university students participated in this study as mock jurors.<sup>2</sup> In line with the eligibility criteria for Canadian jury duty, all participants were at least 18 years of age, a

---

<sup>2</sup> Note: these data are drawn from a larger study conducted by the Legal Decision-Making Lab at Carleton University. Participants, measures, and procedures outside of the scope of this thesis are omitted from this section.

Canadian resident, had never been convicted of an indictable offence without receiving a record suspension, and were fluent in speaking, reading, and writing in English. Participants' ages ranged from 18 to 56 ( $M = 20.39$ ,  $SD = 5.86$ ). There were 53 women, 29 men, and one participant who identified as gender nonconforming. Table 1 includes the racial breakdown of the participants. All participants were enrolled in psychology courses at Carleton University for which they participated in psychology research projects through an online research tool, Sona Systems, for course credit.

**Table 1**

***Racial Breakdown for Mock Juror Participants***

Race	Number of Participants
Chinese	1
South Asian	8
Black	14
Filipino	1
Latin American	2
Arab	4
White	47
Indigenous Peoples of Canada	2
Other	4

**Materials**

***Trial Stimulus***

The mock trial video was professionally produced with paid actors and a film director on a courtroom set. The trial depicted in the video was based on the real case of *R. v. Gayle* (1994).

The trial video (approximately 40 minutes long) began with the judge stating the charge (first degree murder/murder of peace officer, Constable Nathan Hayes; s.231(4) of the Criminal Code) against the defendant, as well as procedural-legal instructions to the jury<sup>3</sup>. The Crown prosecutor and the Defence counsel provided opening remarks. Following opening remarks, the Crown and Defence direct- and cross-examined the following witnesses: Constable Brian Patterson, who discussed the nature of the foot patrol that preceded the murder and the events that followed, including the defendant punching Hayes in the chest and eluding police as he ran towards a parking garage exit; Dr. John Dougherty, a medical examiner who testified on the bullet wounds of Hayes and Whiteduck; Leah Kendall, a witness that was present in the parking garage area during the incident and spoke on the auditory distance between shots fired; and finally, the defendant, Justin Whiteduck, recounted the events and the emotions he experienced during the incident, arguing that he feared for his life because the officers punched and shot at him. The Crown and Defence presented closing remarks at the end of the trial, followed by the Crown's rebuttal to the Defence's closing remarks. Finally, the judge provided deliberation instructions to the jury.

### ***Demographics Questionnaire***

All participants completed a background form to collect information about personal demographics, including age, gender, race/ethnicity, citizenship, English language speaking

---

<sup>3</sup> Jury instructions were as follows: the jury must decide the case solely on the evidence presented in the courtroom; the defendant, Justin Whiteduck, is presumed innocent unless the Crown proves beyond a reasonable doubt that he is guilty; and the defendant is not guilty if he acted in self-defence. Jurors were also informed that the Crown is responsible for proving beyond a reasonable doubt that Whiteduck did not shoot Hayes in self-defence, such that one of the following requirements is met: force was not being used against the defendant; he was not protecting himself from the use of force; and the circumstances for self-defence were unreasonable. Further, self-defence does not apply if the person threatening to use force (i.e., Hayes) was required and authorized by the law to do so.

Jury instructions were obtained from the National Judicial Institute (n.d.).

abilities, and any history of indictable offences (see Appendix A). This questionnaire also served to assess participants' eligibility for jury duty.

### *System Justification Scale*

The present study adapted the SJS from Jost and Kay (2005) to measure system-justifying beliefs in jurors specific to the status quo of Canadian race relations (see Appendix B). The SJS scale from Jost and Kay (2005) involved eight items that measured beliefs about gender inequality in society. The SJS used in the current study removed the second item from Jost and Kay's (2005) SJS: "The division of labor in families generally operates as it should". All other items were modified to ask about racial rather than gender disparity, while retaining the wording of the original items as much as possible. For example, the first item in the scale read as "In general, relations between men and women are fair" in Jost and Kay's (2005) study and was modified to the following: "In general, relations between racial groups are fair".

Jurors rated their agreement with seven statements regarding systemic racial disparities in Canada on a scale from 1 to 5, where 1 represents "Strongly Disagree" and 5 represents "Strongly Agree". Sample items include "In general, relations between racial groups are fair" and "Most policies relating to race serve the greater good". Jurors who indicate a high degree of agreement with these statements show high degrees of ideological support for the status quo of race relations in Canada. Items 2 ("Race relations need to be radically reconstructed") and 6 ("Racism in society is getting worse every year") were reverse coded, such that higher ratings on these two items represented fewer system-justifying beliefs. Jost and Kay (2005) calculated the Cronbach's alpha value for the eight items of their SJS scale ( $\alpha = .79$ ). This study adapted Jost and Kay's (2005) scale, and thus calculated the Cronbach's alpha value of the seven adapted SJS items to determine the internal consistency of the current sample ( $\alpha = .74$ ). This value is

comparable to the value found by Jost and Kay (2005) and deemed acceptable given the recommendation that a Cronbach's alpha value of .70 or greater is acceptable for preliminary research (Peterson, 1994).

### **Procedure**

As the data used in this study were part of a larger study, including other questionnaire information outside of the scope of the present study. All participants received participant codes to maintain anonymity and confidentiality. All participants also received a number tag to place on the table at their seat, allowing jurors to refer to each other by their numbers as opposed to their names.

Students signed up for this study through Sona Systems. Upon arrival, participants listened to an introductory script and completed a signed informed consent form agreeing to participate in the study. The completion of the jury questionnaire, which included the SJS, was randomized to be completed either before the trial video or after the deliberation.

Counterbalancing allows for the investigation of any possible influences that the questionnaire could have on the deliberation process. Participants then watched the 40-minute mock trial video.

After jurors watched the mock trial video, they deliberated as a group for up to 1 hour to decide whether the defendant was guilty or not guilty of first-degree murder. Once the deliberation reached an hour or the jury reached a verdict, the jury completed a group verdict form indicating whether the jury found the defendant guilty or not guilty, or if they were hung. Finally, all jurors completed the demographics form and were debriefed on the purpose of the study. Participants received course credit through Sona Systems as compensation for their participation.

### ***Coding Deliberation Content***

All deliberation sessions were audio and video recorded, then professionally transcribed. Transcriptions were transformed into a codable dataset by segmenting jurors' speech into utterances, which were determined through the concept of an idea unit: a segment of speech spoken by one juror on one topic (Greene et al., 2008). The utterance could be of any length so long as the same juror focused on only one topic throughout the utterance.

Deliberation content was coded based on the coding scheme used by Kaplan and Miller (1987), a two-fold coding scheme in which the utterance type and the content type of an utterance was coded. The utterance type code involved coding an utterances as direct assertions, references to assertions by others, agreement with others, disagreement with others, or questions. Kaplan and Miller (1987) did not provide definitions or examples for the utterance type codes, so utterance type was coded based on the colloquial understanding of the utterance type codes (e.g., references to assertions by others would be considered utterances that mentioned or alluded to another juror's assertion). An utterance type code of "unclear" was also added to account for utterances that could not be readily categorized due to unintelligible speech that was not transcribable.

The content type code involved coding utterances as either testimony, inferences from testimony, values/norms, verdict preferences, nonspecific pressure, procedures, legal issues, or other based on specific criteria and examples (Table 2). These content type categories were considered as either normative, informational, or procedural-legal content based on the first category in which they were coded, as specified in Table 2. Utterances were considered normative if they involved discussion about values, norms, and verdict preferences (i.e., statements asserting one's preferences and implicit or explicit pressure to conform to the group).

Utterances were considered informational if they involved discussion about testimony or inferences from testimony (i.e., statements providing information or evidence presented in the mock trial video). Kaplan and Miller (1987) found virtually no procedural-legal statements and only a few statements that fit under the category of nonspecific pressure. Statements that fit under more than one content type were multiply coded.

**Table 2**

*Coding Scheme by Kaplan and Miller (1987)*

Utterance Type	Criteria	Example	Content Type
Testimony	Statements citing testimonial facts that were provided in the case description	“He was a heavy smoker”	Informational
Inferences from Testimony	Statements of facts not given in the case description but that were inferred from facts given	“Just the hospital bills alone from five days would be enormous”	Informational
Values/Norms	Statements of personal values or societal norms of right or wrong in actions and outcomes, or statements of appropriateness	“It’s wrong to put off repairs for so long.”	Normative
Verdict Preferences	Statements alluding to one's preference for the amount of the	“He should get more than that.”	Normative

	award, including both simple declarations and normative pressure to reach a specific verdict	“Do what the majority thinks is right.”	
Nonspecific Pressure	Statements applying pressure on actions and decisions, but not toward a particular decision	“If we can get this done, we can go home.”	Normative
Procedures	Statements about rules and procedures to be followed by the group	“Should we take a vote now?”	Procedural-Legal
Legal Issues	Statements about pertinent law	“If they give him the money, does it have to be used for that?”	Procedural-Legal
Other	Statements not included in the other categories and irrelevant to the trial	“What psych section are you in?”	N/A

---

**Determining Interrater Reliability Using Fleiss’ Kappa.** To calculate the interrater reliability of the utterance and content type codes, three raters independently coded the discussion content of three deliberations. Given the presence of three raters, Cohen’s (1960) kappa could not be computed and thus Fleiss’ kappa ( $\kappa$ ) was calculated as it accounts for multiple raters (Fleiss, 1971). Fleiss’  $\kappa$  was used to determine the acceptability of agreement between independent raters. Any coding disagreements between raters were verbally resolved to ensure consistency in coding.



Given utterances were multiply coded, the  $\kappa$  values for the utterance and content type codes could not be determined by a single  $\kappa$  value. Instead, a “Best-Worst”  $\kappa$  method was applied, in which two datasets were derived from the raters’ codes to depict two  $\kappa$  scenarios: a “worst” case, which involved codes with the lowest agreement between raters; and a “best” case, which involved codes with the greatest agreement between raters. For example, if the first rater coded an utterance as codes 1 and 2, the second rater coded an utterance as codes 1 and 3, and the third rater coded an utterance as only code 1, the worst-case scenario dataset would include the codes 2, 3, and 1, and the best-case scenario dataset would include codes 1, 1, and 1 for the first, second, and third rater, respectively. The computed  $\kappa$  from the worst-case dataset would signify the lowest possible value of the  $\kappa$  and the computed  $\kappa$  from the best-case dataset would signify the highest possible value of  $\kappa$ . Thus, a  $\kappa$  value range would be used to determine the acceptability of interrater agreement as opposed to a single value. This Best-Worst method thus allowed for the computation of  $\kappa$  values given the presence of multiple codes being assigned to one datapoint (i.e., utterance). To my knowledge, this method of calculating a range for  $\kappa$  given the presence of multiple coding has not been utilized by other published studies.

Landis and Koch (1977) provided the first ranges describing the relative strength of  $\kappa$ , which apply to both Cohen’s and Fleiss’  $\kappa$ . They state that while these divisions are arbitrary, they provide useful benchmarks for discussing  $\kappa$  values. Thus, based on Landis and Koch’s (1977) cut-offs for interrater agreement, a coefficient between .21 and .40 is fair, .41 and .60 is moderate, .61 and .80 is substantial, and .81 and 1.00 is almost perfect. For the purposes of the current study, a  $\kappa$  coefficient above .40 was considered acceptable and anything above .75 was considered excellent (Yamamoto, 2019) Once a  $\kappa$  value above .40 was reached, one rater completed the coding of the remaining deliberations (Yamamoto, 2019). If the overall  $\kappa$

coefficient was not acceptable at the low end of the  $\kappa$  range, then the three raters continued to verbally resolve coding disagreements and independently code until an acceptable worst-case scenario  $\kappa$  was reached.

The following Fleiss'  $\kappa$  values represent overall agreement between the three raters across codes. For the worst-case utterance type codes, there was acceptable overall agreement,  $\kappa = .430$  (95% CI, .406 to .453),  $p < .001$ . For the best-case utterance type codes, there was acceptable overall agreement,  $\kappa = .634$  (95% CI, .609 to .659),  $p < .001$ . For the worst-case content type code, there was acceptable overall agreement,  $\kappa = .587$ . (95% CI, .576. to .619),  $p < .001$ . For the best-case content type code, there was acceptable overall agreement,  $\kappa = .667$  (95% CI, .645 to .689),  $p < .001$ . See Table 3 and 4 for agreement of individual utterance type categories in the worst- and best-case scenario, respectively, and Table 5 and 6 for agreement of individual content type categories in the worst- and best-case scenario, respectively. Note that Table 3, 4, 5, and 6 also include the frequency of each code for all three raters, as a  $\kappa$  value below .40 was deemed sufficient for codes with particularly low frequencies so long as the overall agreement was still above .40. Indeed, low frequencies can decrease  $\kappa$  even if there is high agreement between raters (Xu & Lorber, 2014). Thus,  $\kappa$  values should be interpreted with caution given an individual category had low base rates.

**Table 3**

***Worst-Case Interrater Agreement on Individual Utterance Type Categories***

Utterance Type Code	$\kappa$	95% Confidence Interval		Rater 1 Code	Rater 2 Code	Rater 3 Code
		Lower Bound	Upper Bound	Frequency	Frequency	Frequency
Direct Assertion	.399**	.356	.441	398	375	392

Reference to	.010*	-.033	.052	42	97	18
Assertions by Others)						
Agreement with	.549**	.506	.591	88	89	77
Others						
Disagreement with	.143**	.100	.185	25	3	32
Others						
Questions	.683**	.641	.726	105	97	82
Unclear	.458**	.415	.500	53	50	110

Note. \*  $p > .05$ , \*\*  $p < .001$

**Table 4**

*Best-Case Interrater Agreement on Individual Utterance Type Categories*

Utterance Type Code	$\kappa$	95% Confidence Interval		Rater 1 Code Frequency	Rater 2 Code Frequency	Rater 3 Code Frequency
		Lower Bound	Upper Bound			
Direct Assertion	.645*	.602	.687	458	480	391
Reference to	.188*	.146	.230	32	23	19
Assertions by Others)						
Agreement with	.728*	.686	.771	74	67	76
Others						
Disagreement with	.210*	.168	.253	5	2	33
Others						
Questions	.843*	.801	.886	92	88	82
Unclear	.511*	.468	.553	50	51	110

Note. \*  $p < .001$

**Table 5**

*Worst-Case Interrater Agreement on Individual Content Type Categories*

Content Type Code	$\kappa$	95% Confidence Interval		Rater 1 Code Frequency	Rater 2 Code Frequency	Rater 3 Code Frequency
		Lower Bound	Upper Bound			
Testimony	.562*	.519	.604	50	65	32
Inference from Testimony	.692*	.650	.735	199	229	230
Values/Norms	.463*	.421	.506	15	9	12
Verdict Preferences	.533*	.491	.575	62	75	88
Nonspecific Pressure	.492*	.450	.535	16	5	11
Procedures	.365*	.322	.407	27	39	21
Legal Issues	.456*	.414	.499	80	107	37
Other	.661*	.619	.704	262	182	280

Note. \*  $p < .001$

**Table 6**

*Best-Case Interrater Agreement on Individual Content Type Categories*

Content Type Code	$\kappa$	95% Confidence Interval		Rater 1 Code Frequency	Rater 2 Code Frequency	Rater 3 Code Frequency
		Lower Bound	Upper Bound			
Testimony	.628*	.586	.671	46	60	32
Inference from Testimony	.741*	.699	.784	207	234	231

Values/Norms	.671*	.629	.714	13	11	10
Verdict Preferences	.680*	.637	.722	60	79	88
Nonspecific	.509*	.467	.551	15	5	11
Pressure						
Procedures	.573*	.530	.615	31	41	21
Legal Issues	.567*	.524	.609	74	99	37
Other	.666*	.624	.709	265	182	281

Note. \*  $p < .001$

## Results

HLM software 7 was used to compute the hierarchical linear modelling (HLM) analysis. All other analyses of the data (e.g., assumption tests, descriptive statistics, counterbalancing) were analyzed using IBM SPSS Statistics Version 25.

### General Characteristics

Table 7 displays the general characteristics of the jury groups and Table 8 displays the demographic breakdown of the jury groups. Deliberations ranged from 2 minutes and 48 seconds long to 59 minutes and 30 seconds long ( $M = 25:33$ ;  $SD = 20:16$ ). Jury F2 had the shortest deliberation but was not considered an outlier as jury H2 also had a deliberation under 4 minutes (3:36). Many juries decided on a group verdict before the allotted 60 minutes and did not deliberate into the second and/or third time blocks; thus, only time blocks relevant to each jury were retained (e.g., if a jury did not deliberate into the third time block, the dataset row representing the third time block for each juror in that jury were removed as they were not applicable).

Jury sizes ranged from 5 to 11 jurors. Juror 2 from jury K1 was removed from analysis due to a missing answer on an SJS item, causing their total SJS score to not be reliably compared to other jurors. Juror 11 from jury T1 did not speak throughout the deliberation and did not answer the SJS or demographic questions, including age; given the extent of missing data from this individual, they were removed from the analysis. Further, Juror 7 from Jury F1 was removed from analysis for not completing the age question and Juror 5 from Jury C2 was removed from analysis for answering that they were 16 years old. Thus, out of the 87 initial jurors involved in the 11 juries, 4 were removed from analysis, totalling 83 final jurors included in the analysis.

**Table 7*****General Characteristics of the Jury Groups***

<b>Jury</b>	<b>Number of Jurors</b>	<b>Deliberation Time (Minutes:Seconds)</b>	<b>SJS score <i>M (SD)</i></b>	<b>Verdict</b>
Jury 1 (E1)	6	41:27	2.917 (0.369)	Hung
Jury 2 (F1)*	7	46:53	3.146 (0.780)	Not Guilty
Jury 3 (I1)	11	8:57	3.023 (0.567)	Not Guilty
Jury 4 (K1)**	9	46:47	3.125 (0.488)	Hung
Jury 5 (P1)	9	10:05	2.694 (0.698)	Not Guilty
Jury 6 (R1)	9	59:30	2.361 (0.528)	Hung
Jury 7 (T1)***	11	32:02	3.075 (0.666)	Not Guilty
Jury 8 (V1)	5	14:14	2.525 (0.657)	Not Guilty
Jury 9 (C2)****	9	14:52	2.599 (0.413)	Not Guilty
Jury 10 (F2)	5	2:48	2.950 (0.553)	Not Guilty
Jury 11 (H2)	6	3:36	2.417 (0.637)	Not Guilty

*Note.* SJS items were rated on a 5-point scale, where 1 represented “Strongly Agree” and 5 represented “Strongly Disagree”. \*Jury F1 included 7 jurors, but juror 7 was removed from analysis so only 6 jurors were included. \*\*Jury K1 included 9 jurors, but juror 2 was removed from analysis so only 8 jurors were included. \*\*\*Jury T1 included 11 jurors, but juror 11 was removed from analysis so only 10 jurors were included. \*\*\*\* Jury C2 included 9 jurors, but juror 5 was removed from analysis so only 8 jurors were included.

**Table 8**

***Demographic Breakdown of the Jury Groups***

Jury	Age Mean (SD)	Gender (Count)			Racial Composition (Count)	
		Man	Woman	Other	White	Another Race
Jury 1 (E1)	26.17 (14.82)	4	2	0	2	4
Jury 2 (F1)	20.33 (4.06)	3	3	0	2	4
Jury 3 (I1)	19.36 (1.50)	6	5	0	3	8
Jury 4 (K1)	19.75 (3.26)	3	5	0	5	3
Jury 5 (P1)	19.22 (1.56)	3	5	1	4	5
Jury 6 (R1)	24.56 (10.97)	8	1	0	5	4
Jury 7 (T1)	18.60 (1.05)	6	4	0	9	1
Jury 8 (V1)	18.80 (1.30)	5	0	0	4	1
Jury 9 (C2)	19.50 (2.39)	8	0	0	6	2
Jury 10 (F2)	19.20 (1.64)	3	2	0	3	2
Jury 11 (H2)	19.33 (1.97)	5	1	0	4	2

Table 9 breaks down the utterance frequencies across juries. There was a total of 3763 utterances. Utterances were multiply coded: 3287 utterances received only one content type categorization, 421 utterances received two content type categorizations, and 55 utterances received three content type categorizations.

**Table 9***Utterance and Content Code Frequencies by Jury Group*

<b>Jury</b>	<b>Total Utterances</b>	<b>Content Type</b>			
		<b>Informational</b>	<b>Normative</b>	<b>Procedural/ Legal</b>	<b>Other</b>
Jury 1 (E1)	386	284 (73.6%)	70 (18.1%)	88 (22.8%)	41 (10.6%)
Jury 2 (F1)	559	218 (39.0%)	75 (13.4%)	93 (16.6%)	191 (34.2%)
Jury 3 (I1)	105	24 (22.9%)	18 (17.1%)	17 (16.2%)	54 (51.4%)
Jury 4 (K1)	863	571 (66.2%)	162 (18.8%)	144 (16.7%)	134 (15.5%)
Jury 5 (P1)	81	50 (61.7%)	31 (38.3%)	13 (16.0%)	14 (17.3%)
Jury 6 (R1)	966	620 (64.2%)	113 (11.7%)	72 (7.5%)	247 (25.6%)
Jury 7 (T1)	403	239 (59.3%)	54 (13.4%)	53 (13.2%)	102 (25.3%)
Jury 8 (V1)	148	98 (66.2%)	28 (18.9%)	38 (25.7%)	21 (14.2%)
Jury 9 (C2)	177	125 (70.6%)	43 (24.3%)	51 (28.8%)	13 (7.3%)
Jury 10 (F2)	28	7 (25.0%)	9 (32.1%)	15 (53.6%)	7 (25.0%)
Jury 11 (H2)	47	15 (31.9%)	6 (12.8%)	6 (12.8%)	20 (42.6%)

*Note.* Given utterances were multiply coded, the counts under content type do not sum to the total number of utterances for each jury, as one utterance may be counted more than once if they fit under multiple content type categories. As such, percentages do not sum to 100% as they represent the proportion of utterances that fit each respective content code, and some utterances fit under more than one content code.

When considering utterance types, 2693 utterances were categorized as direct assertions and 1070 utterances fell under the other utterance type categories (i.e., reference to an assertion, agreement with others, disagreement with others, questions, or unclear). A Bernoulli HLM was conducted to examine whether there was a difference in content type between direct assertions



and other types of utterances and there was no significant difference. Thus, all utterance types were included in the final analysis.

### **Assumption Testing**

HLM analysis was used to investigate the relation between time and SJS scores and the ratio of normative to total utterances. HLM assumptions of linearity and normality were examined prior to proceeding with the analysis.

The assumption of linearity was examined through visual analysis of scatterplots for the dependent variable, ratio of normative to total utterances, by time block and by SJS scores (see Figure 1a and 1b, respectively). A visual inspection did not illustrate a discernable pattern in the datapoints, suggesting the assumption of linearity was met.

The assumption of normality was examined through a visual analysis of Quantile-Quantile (Q-Q) scatterplots of the dependent variable, ratio of normative to total utterances. A visual inspection showed that most datapoints aggregated around the line, suggesting a normally distributed outcome variable (see Figure 2a). Given the proportional nature of the dependent variable, an arcsine transformation was computed to ensure homogeneity of variance. The arcsine transformed outcome variable, ratio of normative to total utterances, was used in the HLM analysis. The resulting Q-Q plot for the arcsine transformed proportion also suggests a normal distribution (see Figure 2b). Some outliers are visually present in the Q-Q plots. Outliers signify jurors whose deliberation content consisted mostly or entirely of normative utterances. Thus, removal of outliers would probably not be appropriate for the dependent variable.

### **Counterbalancing**

An independent samples *t*-test was conducted to examine whether there was a difference in SJS scores and normative discussion content among jurors who completed the questionnaire

before or after the trial. Regarding SJS scores, there was no significant difference between those who completed the questionnaire first ( $M = 19.78$ ,  $SD = 4.20$ ) and those who completed the trial first ( $M = 19.51$ ,  $SD = 5.13$ ),  $t(81) = 0.25$ ,  $p = .800$ . When examining the arcsine-transformed ratio of normative to total utterances, mock jurors who completed the questionnaire first ( $M = 1.00$ ,  $SD = 0.57$ ) had significantly greater ratios of normative to total utterances than those who completed the questionnaire after the trial ( $M = 0.62$ ,  $SD = 0.44$ ),  $t(81) = 3.41$ ,  $p = .001$ . While there was no statistically significant difference in total utterance counts between those who completed the questionnaire before ( $M = 22.88$ ,  $SD = 20.76$ ) versus after the trial ( $M = 24.31$ ,  $SD = 29.06$ ),  $t(149) = -0.34$ ,  $p = .737$ , jurors who completed the questionnaire first ( $M = 4.69$ ,  $SD = 4.30$ ) had higher normative utterance counts on average compared to those who completed the trial first ( $M = 3.26$ ,  $SD = 3.91$ ),  $t(149) = 2.12$ ,  $p = .036$ . Further, there was no difference in deliberation times between those who completed the questionnaire first ( $M = 23:10$ ,  $SD = 19:40$ ) and those who completed the trial first ( $M = 27:32$ ,  $SD = 22:24$ ),  $t(9) = -0.34$ ,  $p = .743$ . These findings indicate that while there were no order effects for the completion of the SJS before or after the trial, there appeared to be order effects for normative discussion that were unrelated to how much a juror spoke or how long a jury deliberated. The limitations associated with this finding are discussed further in the Limitations section.

### **Hierarchical Linear Model**

HLM was chosen to analyze this study's hypotheses given the nested nature of the data structure: individual jurors were nested in their jury groups, such that there was a hierarchy in the data with a possibility of within-group correlations between jurors of the same jury. Data that involve a hierarchy are best analyzed using statistical techniques that account for the hierarchy (Woltman et al., 2012), which was why an HLM analysis was chosen for this study. Of specific

interest was the relationship between juror ratio of normative to total utterances (Level 1 outcome variable) and juror SJS scores (Level 2 predictor variable), as well as the relation between time block (Level 1 predictor variable) and juror ratio of normative to total utterances (Level 1 outcome variable).

A memory check was conducted in which jurors were asked to recall the race of the defendant. Five jurors failed the memory check. Results were computed and compared for datasets in which jurors who failed the memory check were included and excluded. There was no difference in significance findings given the exclusion or inclusion of jurors who failed the memory check, so all jurors were retained.

The first step of the HLM analysis involved computing the null model to examine the variation in the ratio of normative to total utterances at each level of analysis before the addition of predictor variables. See the equations representing the null model in Appendix C. The null model components were used to compute the intraclass correlation (ICC) at the juror and jury levels using equation 1 and 2, respectively (see Appendix C).

Using the first equation, there was a juror-level ICC of .00086, signifying that .086% of variance in the ratio of normative to total utterances was between jurors. Using the second equation, there was a jury-level ICC of .1565, indicating that 15.65% of the variance in the ratio of normative to total utterances was between juries. Thus, the jury level accounted for much greater variance in the ratio of normative to total utterances than the juror level, which accounted for very little variation. The ICC found at the jury level was large enough to proceed with a three-level model. Jurors were nested within juries, and time blocks were nested within jurors; thus, despite a low ICC for jurors, jurors (Level 2) necessarily had to be included in the model.

Following the examination of the null model, grand-mean-centred SJS scores and time block were added to the model to examine predictor fixed effects. The covariate of time block, broken down into three 20-minute segments totaling a 60-minute-long deliberation, represented the first level of the model and totaled 151 units. The second level of the model, the 83 jurors, were nested within 11 juries, the third level of the model. See the full predictive model's equations at each level of analysis in Appendix C.

See Table 10 for the estimated regression coefficients of the model predictors. Non-significant fixed effects for the predictors of time block and SJS score and their interaction indicated that none of them individually or cumulatively described any variation in the ratio of normative to total utterances. Thus, the addition of the predictors of time block and SJS scores, as well as their interaction, did not contribute significantly to the model.

The variance components for the predictive model are presented in Table 11. A significant variance component for time block at the jury level ( $u_{10}$ ) indicated that the relation between the ratio of normative to total utterances and time block varied across juries. However, as indicated by the non-significant fixed effects findings for time block, the slope for time block was not significantly different from zero on average. All other variance components were not significant, suggesting that there was no significant juror or jury variance in any of the other fixed effects.

**Table 10**

*Estimates of Predictor Fixed Effects*

Source	Coefficient	Standard Error	<i>t</i> -ratio	<i>p</i> -value
Intercept	0.753307	0.060548	12.442	< 0.001
SJS	.009419	0.011258	0.837	.422

TimeBlock	-0.048925	0.128602	-0.380	.712
TimeBlock x SJS (Interaction)	-0.003387	0.013467	-0.251	.807

**Table 11**

*Three-Level Analysis Variance Components*

Random Effect		Standard Deviation	Variance Component	df	$\chi^2$	p-value
Level 1*	<i>e</i>	0.49074	0.24082			
Level 2*	<i>r</i> <sub>0</sub>	0.02087	0.00044	29	24.55498	>.500
	<i>r</i> <sub>1</sub>	0.12232	0.01496	29	21.55919	>.500
Level 3**	<i>u</i> <sub>00</sub>	0.07416	0.00550	4	5.31962	.255
	<i>u</i> <sub>01</sub>	0.01276	0.00016	4	2.35108	>.500
	<i>u</i> <sub>10</sub>	0.36323	0.13193	4	15.85262	0.004
	<i>u</i> <sub>11</sub>	0.00866	0.00007	4	3.05083	>.500

*Note.* Fixed effects and variance components are based on all the data. \*The  $\chi^2$  statistics reported above are based on only 39 of 83 juror units that had sufficient data for computation. \*\* The  $\chi^2$  statistics reported above are based on only 5 of 11 juries that had enough data for computation.

The HLM analysis only included 39 out of 83 jurors and 5 out of 11 juries due to insufficient data at the first level of analysis. While the first HLM level represented time, which was divided into three time blocks, 44 jurors and six juries had data for only the first time point as they did not deliberate past 20 minutes. Given the extent of data that was not included in the primary HLM analysis, a two-level HLM analysis was conducted in which the first level was dropped, thus leaving only the juror and jury levels of analysis. This analysis was conducted only to ensure that the extent of missing data from the primary three-level HLM did not compromise

findings. Total deliberation time for each jury was included in the analysis instead of time block. All other variables remained the same. All 83 jurors and 11 juries were retained once the time block level was removed. Findings from the two-level HLM did not deviate from the primary three-level HLM analysis, such that the jurors and juries excluded from the primary analysis due to insufficient data did not make a difference to the results (see Table 12 and 13).

**Table 12***Two-Level Analysis Estimates of Predictor Fixed Effects*

Source	Coefficient	Standard Error	t-ratio	p-value
Intercept	0.861922	0.103364	8.339	< 0.001
SJS	-0.000082	0.005358	-0.015	.988
Deliberation Time	0.001191	0.012429	0.096	.926
Deliberation Time x SJS (Interaction)	-0.000234	0.000666	-0.351	.734

**Table 13***Two-Level Analysis Variance Components*

Random Effect	Standard Deviation	Variance Component	df	$\chi^2$	p-value
Level 1* <i>r</i>	0.48795	0.23809			
Level 2* <i>u</i> <sub>0</sub>	0.28987	0.08402	9	23.41391	.006
<i>u</i> <sub>1</sub>	0.00525	0.00003	9	7.10907	>.500

As a whole, the findings from the HLM analysis were discordant with the current study's first and second hypotheses as the model indicated that neither SJS scores nor time block were significantly related to (i.e., did not account for a significant amount of variation in) the ratio of normative to total utterances.

### Discussion

The present study investigated the relation between jurors' system justification beliefs and normative influence. Normative influence is discussion that pushes for others to conform to a decision preference, facilitating greater group harmony and cohesion (Kaplan & Miller, 1987). System justification beliefs are those that justify a societal status quo (Jost et al., 2004), and in this context, a racially disparaging societal status quo. Given that the relationship between system justification beliefs and normative influence has not been previously researched to my knowledge, the present study aimed to provide preliminary findings by examining the deliberation content of 11 mock juries.

Speculatively, the presence of an Indigenous defendant may motivate jurors who endorse fewer system justification beliefs to facilitate group harmony and cohesion through normative arguments for the benefit of a defendant perceived to be disparaged by the societal status quo. Previous findings indicate that individuals who are aware of disparities between social groups are more eager to talk about imbalanced power relations in hope of ending injustice (Saguy et al., 2008). Moreover, a motivation to prevent injustice has been discussed as potentially inducing normative influence as normative influence stems from a motivation to maintain social harmony and good group relations (Hansen et al., 1993; Hastie et al., 1983). Although a motivation to prevent injustice was not directly investigated in the present study, one may speculate that jurors who believe the societal status quo to be disparaging towards the perceived racial group of the defendant may use arguments that facilitate group harmony and cohesion to prevent an injustice against that defendant. Thus, the first hypothesis of this study posited that there would be a negative association between endorsement of system justification beliefs and normative discussion content, such that jurors with lower SJS scores were expected to discuss a greater proportion of normative content.

In contrast with my first hypothesis, there was no significant relation between SJS scores and the ratio of normative to total utterances. Given the preliminary nature of the present study, my findings cannot be readily compared to previous findings. If a relation did exist between lower endorsement of system justification beliefs and greater use of normative influence, the present study may not have had sufficient statistical power to detect it (see the Limitations section below). However, if the results were accurate in signifying that system justification beliefs were not related to normative influence, then jurors who perceived the Canadian status quo of race relations to be unjust, compared to those who perceived it to be just, were not more likely to use normative arguments to sway fellow jurors towards their verdict preferences.

While the present study did not include discussions about race or societal status quos in the coding scheme (see the Limitations section), jurors did indeed discuss racially disparaging systems in the context of the trial. Examples of jurors speaking on the race of the defendant, police brutality in other cases, and the status quo of Canada disparaging Indigenous people were found in multiple deliberations. For instance, one juror in jury F1 said, “But, I mean, obviously they targeted him for being Aboriginal, right?” Then, the same jury later said, “No, I think [unclear] because he’s Aboriginal and that’s why they approached him,” in response to why the officers initially approached the defendant. A juror from jury R1 referenced their beliefs that the societal status quo is not equitable towards Indigenous peoples:

*Juror 5:* “But they mentioned that specifically he was an Aboriginal man. They could have just said he was a man. But they specified that he was Aboriginal, which I believe is racial profiling, which I'm sure a lot of people here understand about. Even I do.”



Another juror from jury K1 expressed, “So, if they used, like, pepper spray, it could have just been an excuse to beat a kid up because he’s Aboriginal ... that’s not that big of a stretch.” The same juror later said, “Let’s ... put ourselves in his shoes for just a second. You’re [alone], you’re Aboriginal, so, like, whether we either like it or not, like there is harassment and abuse against these people.” A different juror from the same jury, K1, also said, “With what’s been going on in the news and with, with him being Aboriginal, that really does change up the case... I definitely would have been scared for my life in that situation.” A separate exchange in jury K1 exemplifies discussion of police brutality and societal status quos of racial disparity:

*Juror 8:* “Well, I mean, if we’re talking about him being Aboriginal, we can kind of talk about, like, the case with, um, Trevon Martin and George Zimmerman. Like, he was just a kid walking home, right? And George Zimmerman... Well, he wasn’t a police officer, but he went up and killed this, this kid.”

*Juror 3:* “Yeah.”

*Juror 8:* “And, obviously, if you, like, a non-police officer could do that and get away with it, if an actual police officer comes up and you’re an Aboriginal, like, you’re a person of colour, a teenager, you’re going to freak out a little bit. Because that’s these cases.”

Some jurors, although comparatively fewer, covertly argued against the relevance of race in the case, as may be exemplified in an exchange in jury K1:

*Juror 7*: “I don’t know. Looking, looking at everything, all the cases that have been occurring, it’s really, it’s really scary when a cop approaches you.”

*Juror 3*: “Not if you’re not doing anything illegal.”

These examples of jurors discussing systems and the status quo of racial disparity for Indigenous peoples are not comprehensive and other jurors had similar discussions. However, as mentioned, the present study did not explicitly code for this type of discussion content. Thus, while a relation between system justification and normative influence was not found, this does not mean that jurors did not discuss issues of racial inequity. As discussed in the Limitations section, future studies examining jury deliberation content may account for this gap in coding.

While the present study did not find a significant relationship between system justification beliefs and normative influence, the influence mode discussion content and average SJS scores of jurors may be compared to previous studies. The codebook of the present study was first created by Kaplan and Miller (1987), but they did not provide the proportions of normative influence discussion content from their deliberations. Hansen and colleagues (1993) also used Kaplan and Miller’s (1987) codebook and did provide proportions of discussion content from their deliberations. See Table 12 for the proportions of discussion content for Hansen and colleagues (1993) compared to the current study.<sup>4</sup> Similar to Hansen and colleagues (1993), the current study found that informational influence made up over half of all deliberation content. This is perhaps unsurprising given that informational influence was

---

<sup>4</sup> Note that Hansen and colleagues (1993) provided frequencies of uninterpretable utterances that were included in their total utterance counts, which had to be subtracted from the frequencies of interpretable discussion content for the calculation of discussion content proportions. Further, Hansen and colleagues (1993) did not multiply code utterances, as was done in the current study; so, while the discussion content proportions from their study sum to 100%, the proportions from the current study sum to over 100%.

defined by statements citing testimonial facts or inferences from the testimony, and thus should be expected to dominate deliberation discussion as jurors are instructed to discuss evidence from the trial when deliberating. Further, for both the current study and Hansen and colleagues' (1993) study, normative influence made up the smallest proportion of discussion content. As a whole, the proportions of discussion content in the present study were comparable to the proportions from the Hansen and colleagues' (1993) study, which also used an undergraduate student sample and the Kaplan and Miller (1987) coding scheme.

**Table 14**

***Deliberation Content Proportions Compared to Hansen et al. (1993)***

Study	<i>Normative</i>	<i>Informational</i>	<i>Procedural-Legal</i>	<i>Other</i>
Current Study	19.91%	52.78%	20.89%	24.45%
Hansen et al. (1993)	12.80%	59.30%	14.97%	13.07%

As mentioned in the literature review, the current study investigated the means by which jurors produced movement conformity in their fellow jurors, which involves an individual changing their position to align with an external position (Willis, 1963). Normative influence tends to produce compliance, in which an individual publicly changes their position to please others, gain social approval, or avoid peer rejection, but may not privately agree with that position (Brodbeck et al., 2007; Nial, 1986). If jurors already largely agree with one another, then modes of influence are not needed to produce movement conformity and reach a unanimous verdict. Sixty-nine percent of jurors believed the defendant to be not guilty before deliberation and 86.2% of jurors believed the defendant to be not guilty after deliberation. Further, while three of the 11 juries in this study were hung as they could not reach a

unanimous decision, the remaining eight juries all decided on a not guilty verdict. While some 17.2% of jurors changed their position on the verdict, most jurors maintained their position throughout the deliberation.

This is comparable to the agreement found among juries in Hansen and colleagues' (1993) study. Five out of the six juries (83.33%) included in their study rendered final verdicts that reflected the predeliberation verdict majority. In the present study, the predeliberation verdict majority was not guilty, and eight out of 11 juries (72.73%) rendered a not guilty verdict. Thus, jurors' individual verdict preferences, as well as their final verdicts, indicate that jurors largely believed the defendant to be not guilty and juries overwhelmingly decided on a not guilty verdict. The present study did not analyze changes in verdict decision-making, and thus an in-depth analysis on the extent of movement conformity is beyond the scope of this study; however, comparisons of proportions of not guilty votes at the juror and jury levels indicates that there may not have been considerable disagreement between jurors on the innocence of the defendant. Given the overwhelming number of not guilty verdicts, the deliberations included in this study may not have been fertile ground for normative influence to occur. The extent of agreement on a not guilty verdict was surprising as a pilot study conducted prior to the current study revealed that there were relatively equal frequencies of guilty and not guilty votes for the trial stimulus. Thus, the overwhelming frequency of not guilty votes was not an anticipated result.

In considering system justification, there are no published studies that include an SJS modified to measure beliefs about racial inequality in society to my knowledge, but the average SJS scores from the present study may be cautiously compared to studies that use Jost and Kay's (2005) 8-item SJS. As a reminder, the present study used a modified 7-item SJS in which items are rated on a scale from 1, strongly agree, to 5, strongly disagree, where higher scores represent more perceived fairness of Canadian systems, or less perceived racial disparity. Items on the scale were averaged for each juror and

the mean of the average across 11 juries was 2.80 ( $SD = 0.664$ ). One study by Napier and colleagues (2020), which included 20,000 participants across 23 countries, asked about beliefs of system fairness using the SJS. Napier and colleagues' (2020) mean SJS score was 3.35 ( $SD = 1.26$ ). In a different mock jury study with 282 community members, Pals (2021) found an average gender related SJS score of 3.96 ( $SD = 1.80$ ). For both studies, their average SJS scores were higher than the average SJS score from the current study, and the SJS scores in their studies varied more greatly than in the current study (Napier et al., 2020; Pals, 2021). Finally, Phelan and Rudman (2011) utilized the SJS to measure beliefs about system fairness and legitimacy in 297 undergraduate students, and the 8 items were rated on a 9-point scale from 1, strongly agree, to 9, strongly disagree. They found a mean SJS score of 4.88 ( $SD = 1.18$ ), which is equivalent to a mean score of 2.71 if converted to a 5-point scale. Thus, their average SJS score was slightly less but comparable to the average SJS score found in the present study, and their participants' SJS scores varied more greatly than the jurors in this study.

The present study's average SJS score findings should be compared to Napier and colleagues' (2020), Pals' (2021), and Phelan and Rudman's (2011) findings with caution as this study used a modified version of the SJS; however, the mean SJS from this study was more similar to the mean SJS from Phelan and Rudman's (2011) study compared to Napier and colleagues' (2020) and Pals' (2021) studies, which, speculatively, may be due to the current study using an undergraduate sample, similar to Phelan and Rudman (2011). If this is indeed the case, then one may expect that undergraduate student samples may produce lower SJS scores overall compared to non-student samples. Future studies may compare student samples to non-student samples (e.g., community samples) to examine differences in average SJS scores between different sample types.

In addition to system justification beliefs, time was included as a predictor to account for a potential relationship between time pressure and normative discussion content. The variable of

time was investigated by dividing the 1-hour deliberation period into three 20-minute time blocks representing the beginning, middle, and end of the deliberation. Jurors were expected to experience greater time pressure nearing the end of the allotted deliberation time. Given Hansen and colleagues (1993) and Kelly and colleagues (1997) found that there was greater normative discussion in the presence of time pressure, the present study's second hypothesis posited that the third time block would be predictive of a greater proportion of normative discussion content than the first and second time blocks, with the first time block being predictive of the smallest proportion of normative discussion content comparatively.

Discordant with my second hypothesis, time block was not significantly related to the ratio of normative to total utterances such that jurors did not speak a greater proportion of normative utterances nearing the end of the deliberation. This finding runs contrary to past studies that found that normative influence was related to time pressure (Hansen et al., 1993; Hastie et al., 1983; Kelly et al., 1997). There may not have been sufficient statistical power to detect this relation if it did exist, as discussed further in the Limitations section. However, if a relation between time pressure and normative discussion truly did not exist in this study (and does not reflect a Type II error), the most likely explanation may be that time pressure was not experienced by the jurors. While each jury had up to 60 minutes to deliberate, 6 out of 11 juries made a verdict decision in the first 20 minutes (time block 1) and 1 out of 11 juries made a verdict decision between 20 and 40 minutes (time block 2). This means that only 4 out of 11 juries made a verdict decision after 40 minutes (time block 3), and only one jury used all of the deliberation time allotted. Thus, the basis of the second hypothesis, that time block would be related to normative influence due to time pressure, may not have been applicable as many juries in this study may not have experienced time pressure.

While Kelly and colleagues (1997) examined the relationship between normative influence and time pressure, they did it through a 2x2 study design in which time pressure was manipulated using instructions to participants and thus changes in discussion content were not examined over time. Instead, changes in discussion content over time may be compared to Hansen and colleagues' (1993) study. Hansen and colleagues (1993) operationalized their time blocks differently than the present study, such that their first time block represented 5 to 15 minutes after the start of the deliberation, the second time block was the 10 minute midpoint of the total deliberation, and the third time block was the 10 minutes before the final verdict. As previously mentioned, the current study divided the 60-minute allotted deliberation time into three 20-minute time blocks. Hansen and colleagues' (1993) three time blocks were each 10 minutes in length and the time at which the second and third time block began was relative to the total length of the deliberation. Their deliberations ranged from 23 to 166 minutes, such that jurors did not have a 60-minute deliberation limit. Further, the shortest deliberation in their sample was much longer than the shortest deliberation time in the current study, 2 minutes and 48 seconds. Thus, their normative to total utterance ratios across time blocks should be cautiously compared to those of the present study (see Table 15). The ratio of normative to total utterances increased from each progressive time block in Hansen and colleagues' (1993) study, perhaps indicating a steady increase in time pressure for jurors. In the present study, the ratio of normative to total utterances decreased from time block 1 to 2, then increased from time block 2 to 3, such that the hypothesized linear increase in time pressure could not be implicated. Contrasting findings between the current study and Hansen and colleagues' (1993) study may be a product of the differences in deliberation structures and the operationalizations of time blocks, as well as differences in the extent of time pressure felt by jurors.

**Table 15*****Ratio of Normative to Total Utterance Averages across Time Blocks Compared to Hansen et al. (1993)***

Study	Time Block 1	Time Block 2	Time Block 3
Current Study	.205	.169	.207
Hansen et al. (1993)	.077	.110	.187

*Note.* Ratio of normative to total utterances for Hansen et al. (1993) was calculated by dividing the frequency of normative utterances by the total number of intelligible utterances, such that the frequencies of unintelligible utterances were removed from the total number of utterances reported in their study.

As a whole, the findings of this study did not support the hypotheses that normative influence and the predictor variables of system justification and time block would be related. Despite my hypotheses not being supported, the current study's investigation of a potential relationship between system justification and normative influence was novel in nature. This study provided the first analysis on how system justification beliefs may be associated with influence modes in jury deliberation content, and thus a preliminary understanding into how beliefs about racial disparity in the Canadian societal status quo may be associated with how jurors try to influence each other towards conformity. Further investigation is warranted to understand the ways in which jurors' beliefs about a racially disparaging societal status quo and time pressure in a deliberation may be related to how jurors attempt to persuade others towards conformity. Future research may address the limitations of this study and consider the addition of several variables of interest to expand on existing knowledge about influence modes in jury decision-making.



## Implications

While the present study did not find a statistically significant relationship between jurors' system justification beliefs and their use of normative influence in a jury deliberation, the topic of this research has implications for our understanding of how jurors influence each other towards conformity. The current study hypothesized that given a trial with an Indigenous defendant, jurors who believe that the Canadian system is racially disparaging may be more likely to use normative influence than those who endorse beliefs that the system is just. While this hypothesis was not supported, investigating the relationship between jurors' beliefs about race and the influence modes used in deliberations may have practical implications for understanding how jurors persuade each other towards conformity given different trial contexts, such as trials with a racialized defendant.

If a trial's context is one in which jurors are more likely to use normative arguments rather than facts to influence others towards conformity, then a larger proportion of a deliberation may involve pressure to conform rather than evidence from the trial to reach a verdict. Indeed, normative influence has been cited as being related to compliance (i.e., public agreement and private disagreement) while informational influence has been cited as being related to conversion (i.e., public and private agreement, or true conformity; Nail, 1986). Ideally, jurors are making verdict decisions based on facts alone as opposed to subjective feelings of right and wrong.<sup>5</sup> While jurors are instructed to consider the evidence presented during a trial

---

<sup>5</sup> . Jurors are intended to represent the judgement of the community, and thus jurors should be expected to come from diverse backgrounds (Fraher, 1987). One may not realistically expect that a jury verdict could be made entirely objectively, and one may argue that the subjective perspectives that jurors bring into the courtroom are valuable as they are perspectives that represent the community. Thus, some level of subjectivity is expected in the courtroom; however, endorsing policies that would support evidence-based decision-making may aid in ensuring the fairness of trial proceedings for all parties involved.

when reaching a verdict decision, research has found that normative influence is difficult to eliminate even in an anonymized deliberation setting (Garcia et al., 2021).

A better understanding of what contexts produce greater normative influence in a deliberation may inform the courts about how to moderate jurors' use of normative influence so there is greater assurance that jurors are solely relying on evidence when making a decision. While normative influence is unlikely to be eliminated in any case, some practices may encourage greater reliance on the facts of a case, such as instruction that notetaking by jurors during the trial is permitted. As a whole, deepening our understanding of the contexts in which jurors rely on normative influence may allow us to advocate for policies that will facilitate greater reliance on evidence when reaching a verdict decision.

### **Limitations**

The present study had several limitations that should be addressed in future research, including limitations in sampling, consequentiality, potential priming effects, and the measurement of normative influence and system justification beliefs. Firstly, many researchers have expressed concern with the generalizability of an undergraduate student sample, and some have compared findings between undergraduate samples and community samples (Bornstein et al., 2017; Keller & Weiner, 2011; Maeder et al., 2018; Weiner et al., 2011). In a mock jury study by Keller and Weiner (2011) comparing 120 students and 99 community members, some similarities and some differences were found in how measures of bias (e.g., indicators of bias in homicide verdict judgement) were related to beliefs about culpability (e.g., beliefs about the defendant's criminal intentions) and verdict decision-making. Keller and Weiner (2011) concluded that researchers should consider their samples in the context of trial procedures, such as type of charge and standards of evidence. In a meta-analysis by Bornstein and colleagues

(2017) using 53 studies ( $N = 17,716$ ), verdicts, culpability ratings, and damage awards did not vary between student and community samples. Further, the differences found on sentencing and liability judgements rendered small effect sizes, and so Bornstein and colleagues (2017) concluded that the use of student samples may not pose much concern. Maeder and colleagues (2018) compared 506 student and 441 community member mock jurors in a study in which they varied the defendant's race. They found that while students were more likely than community members to convict overall (46% vs. 36%, respectively), there were no differences in verdict as a function of defendant race. Compared to community members, student samples also had a greater proportion of women, were younger, had less age and education variability, and were more ethnically diverse. Students were also more likely to fail a manipulation check indicating an Indigenous defendant's race (Maeder et al., 2018). Maeder and colleagues (2018) concluded that despite some differences between student and community member samples, student samples were not more likely than community samples to threaten ecological validity in a jury simulation in which defendant race was a factor. Indeed, student samples are efficient and are suitable for initial investigations in mock jury studies (Weiner et al., 2011); however, given the contradictory findings on whether student and community samples show differential findings in mock jury studies (Bornstein et al., 2017; Keller & Weiner, 2011; Maeder et al., 2018), certainty on the generalizability of findings may depend on comparing results between different sample types. Thus, future studies may attempt to replicate the current study with a community sample to examine differences in findings.

Another sampling limitation pertains to the sample size of the study as it relates to statistical power. The number of jurors and the number of juries represented data units in the HLM analysis at the juror and jury levels, respectively. At the first level of analysis, time block,

there were 151 units representing jurors at the first, second, and third deliberation time blocks. There were 83 units at the juror level and 11 units at the jury level. Post hoc analyses using Optimal Design (Raudenbush et al., 2011) indicated that, at the juror level, for a sample size of 85 participants, the minimum effect size for which there would be an 80% chance of detecting a relation between system justification and normative influence was .60. Further, for 11 juries with 8 jurors each and an ICC of .05, the minimum effect size for which there would be an 80% chance of detecting a relation between deliberation time and normative influence was .75. Thus, to achieve statistical power of 80% with the current study's sample size, a medium to large effect size would be required. However, if the effect size of the relation between normative influence and system justification beliefs was small, an insufficiently large sample size would mean that the statistical power was not adequately large to detect said small effect. Future studies should thus increase the sample size as a larger sample would serve to increase the power of the study, thus increasing the likelihood of detecting a relation between normative influence and the predictors of time and system justification if a relation exists.

Another limitation of this study was the issue of consequentiality: a limitation of ecological validity inherent to all mock jury studies in which jurors' verdicts lack consequence to a defendant (Bornstein & McCabe, 2005). Jury simulation studies may be realistic in reproducing the real-world characteristics of a trial and jury deliberation; however, regardless of how realistic a jury simulation may be, participants in a mock jury study will still be aware that they are participating in a simulation. If mock jurors are necessarily aware that their decisions carry no consequence to any individual, can their deliberation processes reliably inform researchers about the decision-making of real-life juries? According to a literature review by Bornstein and McCabe (2005) on the issue of consequentiality in mock jury research, few studies

have directly examined how varying the weight of consequences for mock jurors might influence their decision-making, and the few studies that have examined this issue produced inconsistent findings. Determining if and how consequentiality affects the validity of psycholegal research may bear significance to the courts' willingness to implicate research findings (Bornstein & McCabe, 2005). Given that the current study used a jury simulation, the issue of consequentiality should be considered when comparing the results of this study to real-world jury deliberation scenarios.

Another limitation of the current study was the potential priming effect of questionnaire administration. The current study counterbalanced the administration of the SJS questionnaire before and after the trial stimulus so that any possible influences of the questionnaire on the deliberation process could be examined. While there was no difference in SJS scores among jurors who completed the questionnaire versus trial first, jurors who completed the questionnaire before compared to after the trial had larger ratios of normative to total utterances. Further, not only were normative utterance counts higher for those who completed the questionnaire first compared to the those who completed the trial first, there was also no difference in deliberation time nor total utterance counts between the orders of questionnaire completion. Thus, differences in the ratio of normative to total utterances may not be attributable to how long a jury deliberated or how much a juror spoke. These findings suggest a potential priming effect of the questionnaire, where those who completed the questionnaire before the deliberation used more normative arguments than those who were exposed to the questionnaire after the verdicts were decided. Future studies may benefit from investigating the order of questionnaire administration by including it as a predictor variable to control for the potential priming effects of questionnaires on jury decision-making.

Additionally, there were several limitations in the transcription and coding of deliberations that may have resulted in limited reliability in the measurement of normative influence. Pertaining to the transcription of the deliberations, there may have been limitations in the accuracy of transcriptions due to the nature of transcribing audio. There were many instances of jurors speaking over each other or the speech of a juror being indiscernible. Thus, some utterances that could have been coded as normative may not have been transcribed at all. However, this limitation may not be improved through changes in methodology as some discussion content should be expected to be indiscernible in any conversation with interruptions and overtalking.

Outside of limitations to transcription, there were potential limitations relating to the coding of utterances. The codebook used in the present study was initially proposed by Kaplan and Miller (1987) to code for influence modes in jury deliberations. However, the trial stimulus used by Kaplan and Miller (1987) and the purpose of their study was largely different from the present study, and thus their codebook may not have encapsulated all types of normative discussion content that emerged in the present study. Kaplan and Miller's (1987) study involved a civil damage trial in which jurors were responsible for deciding the type of damage award for the plaintiff based on the actions of the defendant. Kaplan and Miller (1987) were interested in the influence mode differences that would arise based on the issue type (i.e., intellective or judgemental) and the assigned decision rule (i.e., unanimous or majority). The trial stimulus included in the present study was a criminal trial, not a civil trial, and jurors decided the verdict of a defendant, not the type of damage award. Normative influence was investigated in this study as relating to system justification beliefs, not issue type or assigned decision rule. Normative arguments relating to a juror's system justification beliefs may be expected to include discussion

about the norms of societal race relations in Canada. While one of the codes from Kaplan and Miller's (1987) codebook represented values and norms, this code was defined as statements of personal values or societal norms of right and wrong in actions and outcomes. This definition appears to relate to the defendant's actions in a civil case determining a damage award. Thus, the definition of this code does not readily apply to norms about race relations. Given the codebook used in the present study was designed to investigate qualitatively different variables using a different type of trial stimulus, the codebook may not adequately encapsulate the normative influence relevant to the present study. Thus, limitations of transcription accuracy and codebook fit may have restricted the precision of how normative influence was operationalized.

To address this coding limitation, future studies examining normative influence and system justification may modify the coding scheme by Kaplan and Miller (1987) to fit the context of a criminal trial with an Indigenous defendant. Modifying the coding scheme would ensure that normative discussion content relevant to the trial stimulus is represented in the codebook, thus reducing the likelihood that a normative utterance is miscoded. Of greatest relevance would be modifying the "Values/Norms" code definition by Kaplan and Miller (1987), such that values and norms about societal race relations are included.

Another potential modification of the coding methods would be to devise a second codebook that codes for system justification discussion content. This may involve a code defined by discussion about system justification beliefs, and further coding these utterances as beliefs that justify, reject, or are neutral about the status quo of race relations in Canada. Future studies that code for system justification content in jury deliberations should examine the relation between juror self-reported system justification beliefs and their propensity to voice these beliefs in deliberations, including the extent to which their self-reported system justification beliefs

align with their spoken beliefs. Further, influence mode coding may be included to examine the overlap between system justification discussion content and normative discussion content. This future direction would thus further dissect whether system justification beliefs are related to how jurors aim to influence each other towards conformity in a case with an Indigenous defendant.

Finally, there were several potential issues relating to the measurement of jurors' system justification beliefs. Firstly, the SJS asked jurors about their beliefs on race relations in Canada but did not specify Indigenous issues. From one perspective, jurors' beliefs about racial disparity may generalize to their beliefs about Indigenous experiences of disparity, as was initially assumed in this study. However, jurors' beliefs about racial disparity in Canada may not necessarily generalize to Indigenous peoples. As previously discussed in the literature review, jurors may differ from each other on their awareness of Indigenous issues (Schaepli et al., 2018). Some jurors who believe that the Canadian status quo is racially disparaging may not generalize these beliefs to Indigenous peoples as they may not know about the extent of societal disparities experienced by Indigenous peoples. Further, one may also speculate that even if a juror believes that there should be racial equality in Canada, they may not apply these beliefs to Indigenous peoples due to a negative racial bias, as opposed to a lack of knowledge about Indigenous issues. Thus, given the SJS used in this study asked about race relations in general, it may not accurately measure jurors' beliefs about the systemic and structural disparities experienced by the Indigenous Peoples of Canada.

To address this limitation, future research that involves an Indigenous defendant and examines jurors' system justification beliefs should modify the SJS to be specific to the societal disparities experienced by the Indigenous Peoples of Canada. For example, the SJS item "For people of colour, Canada is the best country in the world to live in" could be modified to "For



Indigenous peoples, Canada is the best country in the world to live in". Modifying the SJS to be specific to Indigenous peoples as opposed to racialized people in general could potentially increase the validity of measuring jurors' system justification beliefs specific to Indigenous issues and would thus be a more appropriate measurement given a trial stimulus that involves an Indigenous defendant.

### **Future Directions**

Given the preliminary nature of the present study, some future research directions may more closely examine a potential relation between system justification beliefs and normative influence. Outside of addressing the limitations of the present study, future research should compare results between the ratio and count method of measuring normative utterances. In this study, greater normative influence was operationalized by a higher ratio of normative to total utterances in a deliberation. Future research should compare results from the proportion of normative utterances in a deliberation to results in which greater normative influence is operationalized by a higher count of normative utterances. The present study used the ratio method to account for individual differences between jurors in how much they spoke throughout the deliberation. However, Saguy and colleagues (2008) found that individuals who perceived disparities between groups as illegitimate, and specifically those who sought to change the status quo, spoke more on power relations compared to individuals who did not perceive these disparities as illegitimate (Saguy et al., 2008). Thus, frequency of speech may be relevant to a juror's lack of endorsement of system justification beliefs. Future studies on system justification and normative influence should include frequency of speech in the analysis by comparing results from the ratio and count method of measuring normative utterances so that results from both methods may be investigated and compared.

Finally, future research should include several additional variables to analyze the relationship between normative influence and system justification, including juror knowledge about Indigenous issues, defendant race, and juror race. Juror knowledgeability of Indigenous issues is a covariate of interest as jurors with lower knowledgeability of Indigenous issues may be more likely to support the systems that disparage Indigenous peoples. Indeed, a greater awareness of societal inequity has been related to fewer beliefs that support the societal status quo, as those who are aware of a disparaging system will be less likely to hold beliefs that legitimize that system (Jost et al., 2004). Further, Shepherd (2012) found that when participants did not have a deep understanding of an issue, they were more likely to trust the government's management of that issue. Further, increasing awareness of the illegitimacy of a power incongruence between groups has been found to increase an individual's willingness to speak on said power incongruence in support of equality, even among advantaged group members who may be less motivated than disadvantaged group members to change the status quo (Saguy et al., 2008). Given these findings, greater knowledgeability of the disparities experienced by Indigenous communities may potentially be related to jurors' system justification beliefs and their willingness to talk about these disparities when a defendant is Indigenous. Thus, juror knowledgeability of Indigenous issues should be a variable included in future research examining system justification beliefs and jury decision-making in a trial involving an Indigenous defendant.

Future research examining system justification beliefs and normative influence may benefit from conditions in which the defendant's race varies. In the present study, all jurors were exposed to a trial stimulus in which the defendant was Indigenous. Adding Black or White defendant conditions to the existing Indigenous defendant condition included in this study would

permit the inclusion of a defendant race variable. For instance, one may examine whether there are differences in normative discussion content between defendant race conditions or whether an interaction exists between SJS scores and defendant race. Those with fewer system justification beliefs may be more likely to be motivated to prevent justice miscarriage in the presence of minority group disparity (Jost & Hunyady, 2005). Thus, comparing the relationship between system justification and normative influence across multiple defendant race conditions may allow for jurors' modes of eliciting conformity to be compared as a function of defendant race.

Finally, juror race should be considered in future studies examining system justification and normative influence. Saguy and colleagues (2008) found that disadvantaged group members were more inclined to discuss power imbalances between groups, while advantaged group members were more likely to discuss commonalities between groups. Jurors who have experienced disparities due to their perceived race may thus be more inclined to speak on issues related to racial inequality in society given the presence of a racialized defendant than individuals who have not experienced racial disparities. Future research should consider the inclusion of juror race as a variable in studying system justification and normative influence as there may be an interaction between juror race and system justification.

### **Conclusion**

As a whole, the ways in which jurors' beliefs about the status quo of race relations in Canada are related to how jurors influence each other towards conformity warrant further examination. While the preliminary findings of the present study did not support a relation between system justification beliefs or deliberation time block and normative discussion content, future research may address several limitations of this study and make relevant additions to study system justification and normative influence more deeply. Given the potential implications of

this research for understanding the jury decision-making process and in broadening our knowledge of conformity in jury decision-making, the preliminary findings of the present study would benefit from continued research.

Figure 1a

*Scatter Plot of Normative to Total Utterances by Time Block*

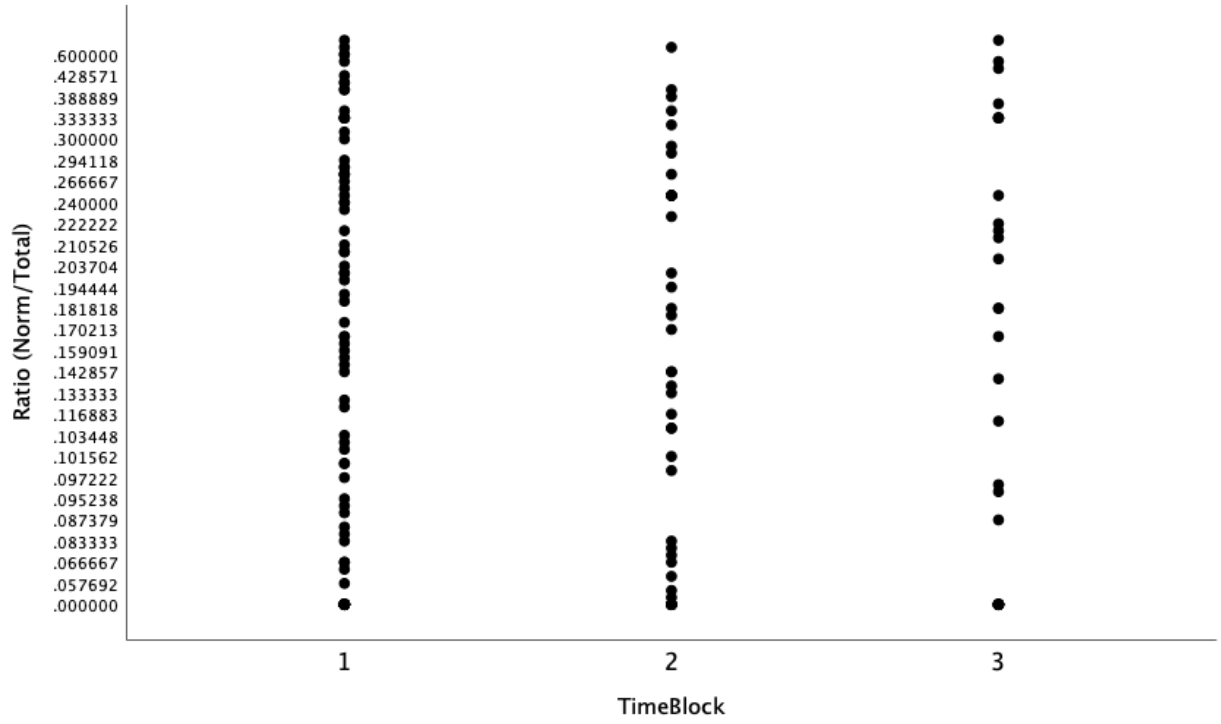


Figure 1b

*Scatterplot of Ratio of Normative to Total Utterances by SJS Score*

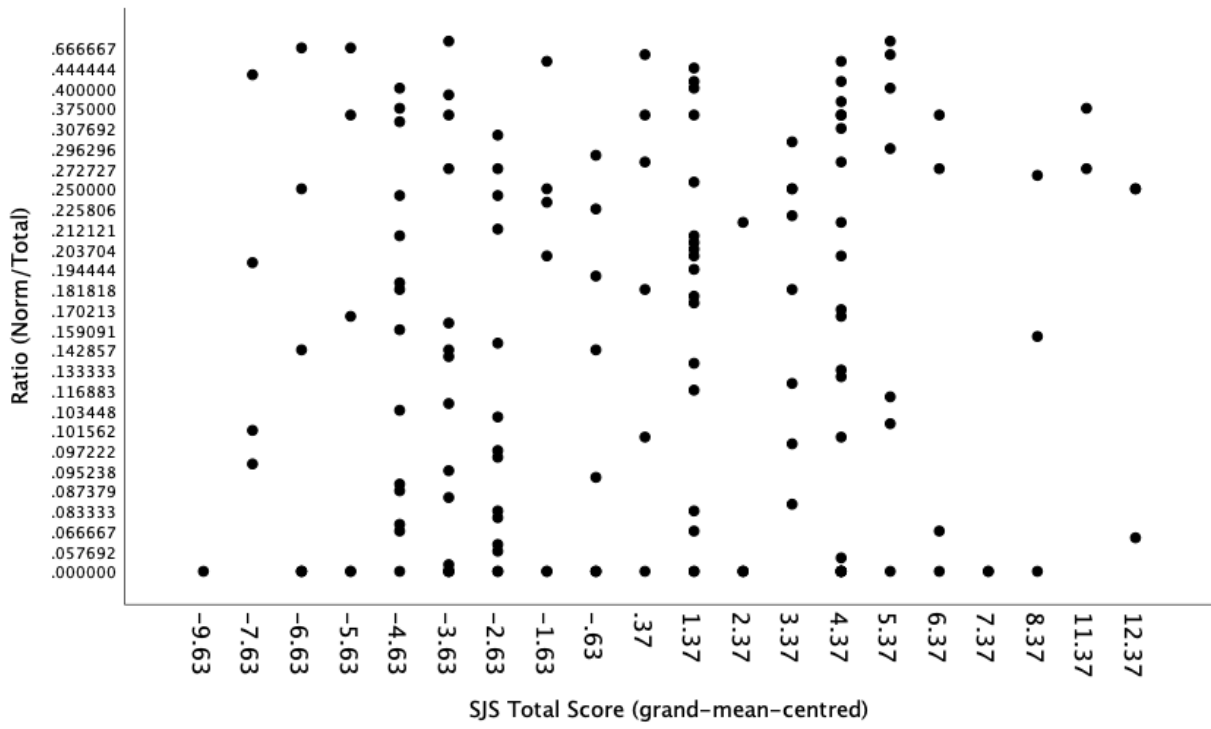


Figure 2a

*Quantile-Quantile Scatterplot of Ratio of Normative to Total Utterances*

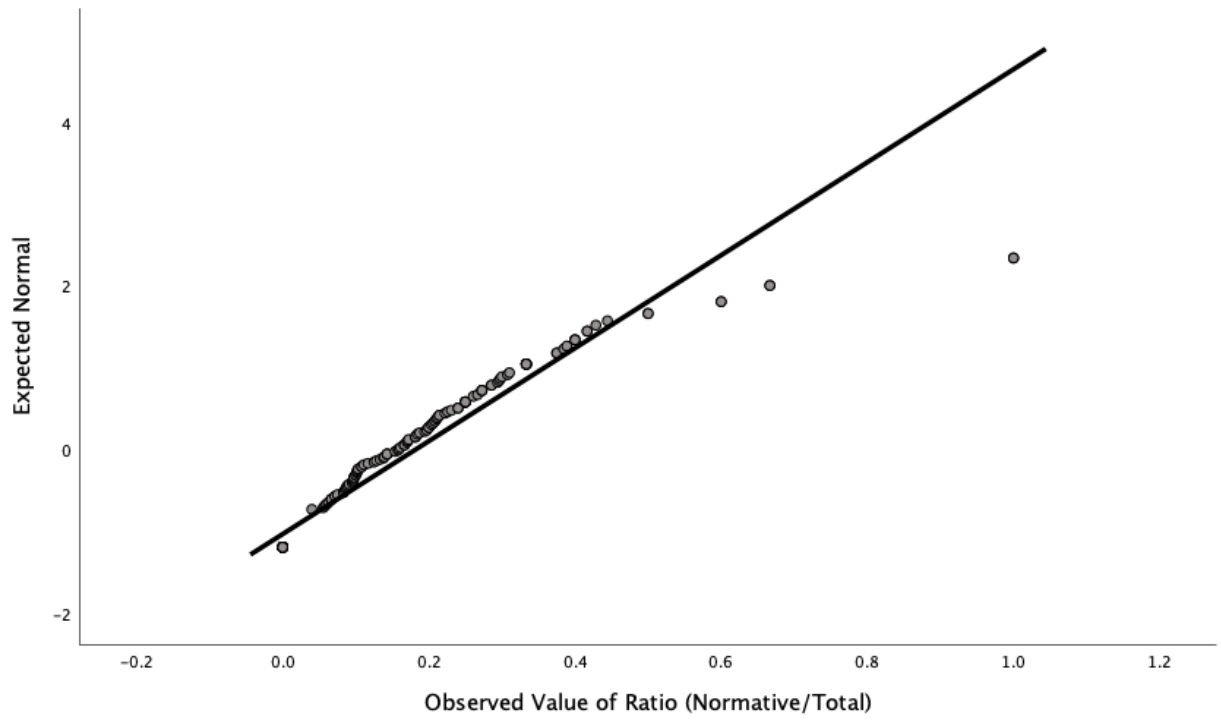
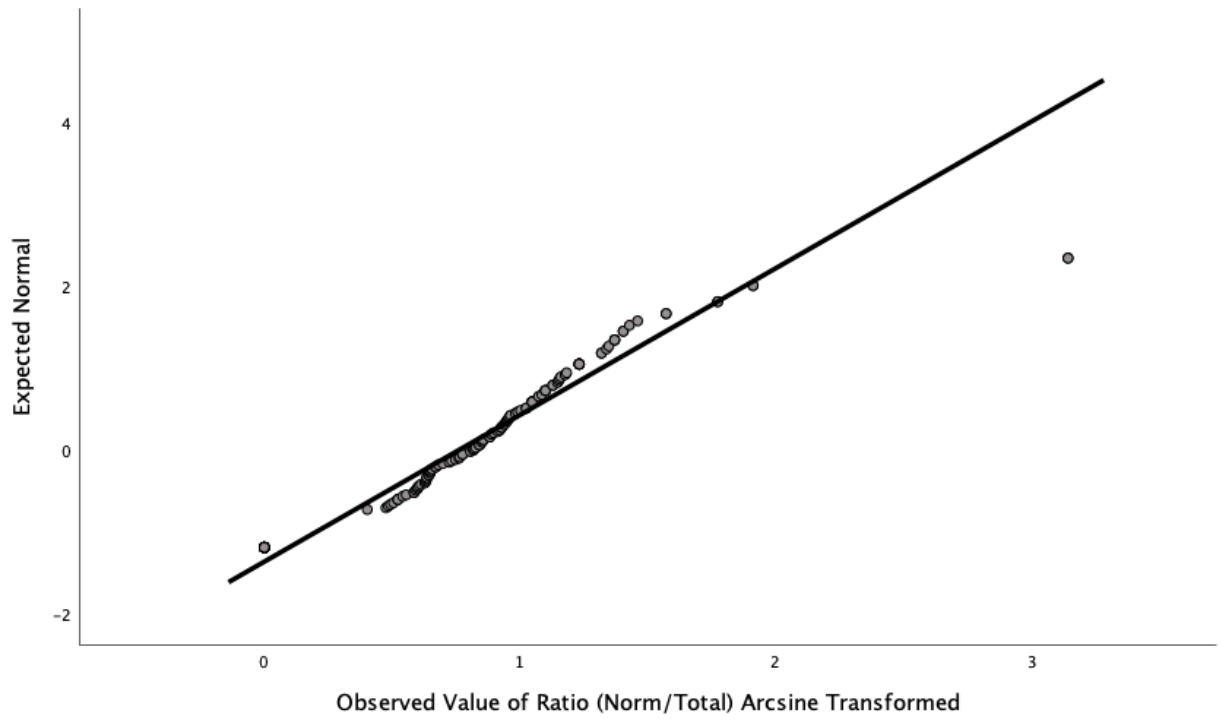


Figure 2b

*Quantile-Quantile Scatterplot of Arcsine Transformed Ratio of Normative to Total Utterances*





### References

- Allen, V. L. (1965). Situational factors in conformity. *Advances in Experimental Social Psychology*, 2, 133-175. [https://doi.org/10.1016/S0065-2601\(08\)60105-7](https://doi.org/10.1016/S0065-2601(08)60105-7)
- Allport, G. W. (1954). *The Nature of Prejudice*. Addison -Wesley Publishing Company.  
[https://faculty.washington.edu/caporaso/courses/203/readings/allport\\_Nature\\_of\\_prejudice.pdf](https://faculty.washington.edu/caporaso/courses/203/readings/allport_Nature_of_prejudice.pdf)
- Asch, S. E., & Guetzkow, H. (1951). Effects of group pressure upon the modification and distortion of judgements. *Organizational Influence Processes*, 295-303.  
<https://doi.org/10.1017/CBO9781107415324.004>
- Blagg, H. (2012). Re-imagining youth justice: Cultural contestation in the Kimberley region of Australia since the 1991 Royal Commission into Aboriginal Deaths in Custody. *Theoretical Criminology*, 16(4), 481-498. <https://doi.org/10.1177/1362480611436360>
- 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. <https://doi.org/10.1037/lhb0000223>
- Borstein, B. H., & McCabe, S. G. (2004). Jurors of the absurd? The role of consequentiality in jury simulation research. *Florida State University Law Review*, 32(2), 443-467.  
<https://ir.law.fsu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1260&context=lr>
- Brodbeck, F. C., Kerschreiter, R., Mojzisch, A., & Schulz-Hardt, S. (2007). Group decision making under conditions of distributed knowledge: The information asymmetries model. *Academy of Management Review* 32(2), 459-479.  
<https://doi.org/10.5465/AMR.2007.24351441>

- Brown, G. R. (2016). The blue line on thin ice: Police use of force modifications in the era of cameraphones and YouTube. *British Journal of Criminology*, *56*(2), 293-312.  
<https://doi.org/10.1093/bjc/azv052>
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591-621.  
<https://doi.org/10.1146/annurev.psych.55.090902.142015>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Corrado, R. R., Kuehn, S., & Margaritescu, I. (2014). Policy issues regarding the overrepresentation of incarcerated Aboriginal young offenders in a Canadian context. *Youth Justice*, *14*(1), 40-62. <https://doi.org/10.1177/1473225413520361>
- Criminal Code*, R.S.C., 1985, c. C-46, s. 34(1).
- Criminal Code*, R.S.C., 1985, c. C-46, s. 231(4).
- Department of Justice. (2019). *Spotlight on Gladue: Challenges, Experiences, and Possibilities in Canada's Criminal Justice System*. <https://www.justice.gc.ca/eng/rp-pr/jr/gladue/p2.html>
- Deutsch, M. (2006). A framework for thinking about oppression and its change. *Social Justice Research*, *19*(1), 7-41. <https://doi.org/10.1007/s11211-006-9998-3>
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological bulletin*, *76*(5), 378.  
<http://www.wpic.pitt.edu/research/biometrics/Publications/Biometrics%20Archives%20PDF/395-1971%20Fleiss0001.pdf>

- Fraher, R. M. (1987). Adjudicative facts, non-evidence facts, and permissible jury background information. *Indiana Law Journal*, 62(2), 333-353.  
<https://www.repository.law.indiana.edu/ilj/vol62/iss2/10>
- Garcia, R. J., Shaw, E. V., & Scurich, N. (2021). Normative and informational influence in group decision making: Effects of majority opinion and anonymity on voting behavior and belief change. *Group Dynamics: Theory, Research, and Practice*. Advance online publication. <http://doi.org/10.1037/gdn0000145>
- 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. <http://doi.org/10.1037/a0013486>
- Hansen, K. L., Schaefer, E. G., & Lawless, J. J. (1993). Temporal patterns of normative, informational, and procedural-legal discussion in jury deliberations. *Basic and Applied Social Psychology*, 14(1), 33-46. [https://doi.org/10.1207/s15324834basps1401\\_3](https://doi.org/10.1207/s15324834basps1401_3)
- Hastie, R., Penrod, S. D., & Pennington, N. (1983). *Inside the Jury*. Harvard University Press.  
<https://doi.org/10.4159/harvard.9780674865945>
- Holstein, J. A. (1985). Jurors' interpretations and jury decision making. *Law and Human Behavior*, 9(1), 83-100. <https://doi.org/10.1007/BF01044291>
- Horowitz, I. A. (1985). The effect of jury nullification instruction on verdicts and jury functioning in criminal trials. *Law and Human Behavior*, 9(1), 25-36.  
<https://doi.org/10.1007/BF01044287>
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33(1), 1-27.  
<https://doi.org/10.1111/j.2044-8309.1994.tb01008.x>

- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25*(6), 881-919. <https://doi.org/10.1111/j.1467-9221.2004.00402.x>
- Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129*(3), 339-375. <https://doi.org/10.1037/0033-2909.129.3.339>
- Jost, J. T., & Hunyady, O. (2005). Antecedents and consequences of system-justifying ideologies. *Current Directions in Psychological Science, 14*(5), 260-265. <https://doi.org/10.1111/j.0963-7214.2005.00377.x>
- Jost, J. T., & Kay, A. C. (2010). Social justice: History, theory, and research. *Handbook of Social Psychology, 1*, 1122-1165. <https://doi.org/10.1002/9780470561119.socpsy002030>
- Jost, J. T., Ledgerwood, A., & Hardin, C. D. (2008). Shared reality, system justification, and the relational basis of ideological beliefs. *Social and Personality Psychology Compass, 2*(1), 171-186. <https://doi.org/10.1111/j.1751-9004.2007.00056.x>
- Jost, J. T., Liviatan, I., van der Toorn, J., Ledgerwood, A., Mandisodza, A., & Nosek, B. A. (2010). System justification: How do we know it's motivated? *The Psychology of Justice and Legitimacy, 11*, 1-22. <https://doi.org/10.4324/9780203837658>
- Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology, 38*(6), 586-602. [https://doi.org/10.1016/S0022-1031\(02\)00505-X](https://doi.org/10.1016/S0022-1031(02)00505-X)
- Jost, J. T., & Thompson, E. P. (2000). Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among

- African Americans and European Americans. *Journal of Experimental Social Psychology*, 36(3), 209-232. <https://doi.org/10.1006/jesp.1999.1403>
- Kaplan, M. F., & Miller, C. E. (1987). Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *Journal of Personality and Social Psychology*, 53(2), 306-313. <https://doi.org/10.1037/0022-3514.53.2.306>
- Kaplan, M. F., Schaefer, E. G., & Zinkiewicz, L. (1994). Member preferences for discussion content in anticipated group decisions effects of type of issue and group interactive goal. *Basic and Applied Social Psychology*, 15(4), 489-508. [https://doi.org/10.1207/s15324834basp1504\\_6](https://doi.org/10.1207/s15324834basp1504_6)
- Kay, A. C., Czapliński, S., & Jost, J. T. (2009). Left-right ideological differences in system justification following exposure to complementary versus noncomplementary stereotype exemplars. *European Journal of Social Psychology*, 39(2), 290-298. <https://doi.org/10.1002/ejsp.500>
- Keller, S. R., & Wiener, R. L. (2011). What are we studying? Student jurors, community jurors, and construct validity. *Behavioral sciences & the law*, 29(3), 376-394. <https://doi.org/10.1002/bsl.971>
- Kelly, J. R., Jackson, J. W., & Hutson-Comeaux, S. L. (1997). The effects of time pressure and task differences on influence modes and accuracy in decision-making groups. *Personality and Social Psychology Bulletin*, 23(1), 10-22. <https://doi.org/10.1177/0146167297231002>
- Kiedrowski, J., Jones, N. A., & Ruddell, R. (2017). 'Set up to fail?' An analysis of self-administered Indigenous police services in Canada. *Police Practice and Research*, 18(6), 584-598. <https://doi.org/10.1080/15614263.2017.1363973>

- Kingsley, M. (2020, January 21). *Indigenous People in Federal Custody Surpasses 30% Correctional Investigator Issues Statement and Challenge*. <https://www.ocibec.gc.ca/cnt/comm/press/press20200121-eng.aspx>
- Kluegel, J. R., & Smith, E. R. (1983). Affirmative action attitudes: Effects of self-interest, racial affect, and stratification beliefs on whites' views. *Social Forces*, *61*(3), 797-824. <https://doi.org/10.1093/sf/61.3.797>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. <http://www.jstor.org/stable/2529310>
- 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. <https://doi.org/10.1080/1068316X.2017.1409895>
- Martorana, P. V., Galinsky, A. D., & Rao, H. (2005). From system justification to system condemnation: Antecedents of attempts to change power hierarchies. *Research on Managing Groups and Teams*, *7*, 283-313. [https://doi.org/10.1016/S1534-0856\(05\)07012-X](https://doi.org/10.1016/S1534-0856(05)07012-X)
- McGrath, J. (1984). *Groups: Interaction and Performance*. Prentice-Hall, Inc. <http://users.ece.utexas.edu/~perry/education/382v-s08/papers/mcgrath84.pdf>
- Nail, P. R. (1986). Toward an integration of some models and theories of social response. *Psychological Bulletin*, *100*(2), 190-206. <https://doi.org/10.1037/0033-2909.100.2.190>
- Napier, J. L., Suppes, A., & Bettinsoli, M. L. (2020). Denial of gender discrimination is associated with better subjective well-being among women: A system justification

- account. *European Journal of Social Psychology*, 50(6), 1191-1209.  
<https://doi.org/10.1002/ejsp.2702>
- National Judicial Institute. (n.d.) *Model Jury Instructions*. <https://www.nji-inm.ca/index.cfm/publications/model-jury-instructions/?langSwitch=en>
- Pals, A. (2021). Himpathy? The Impact of Defendant Social Status on Perceptions of a Rape Legal Case.  
[https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1191&context=psychology\\_etds](https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1191&context=psychology_etds)
- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21(2), 381-391. <https://doi.org/10.1086/209405>
- Phelan, J. E., & Rudman, L. A. (2011). System justification beliefs, affirmative action, and resistance to equal opportunity organizations. *Social Cognition*, 29(3), 376-390.  
<https://doi.org/10.1521/soco.2011.29.3.376>
- Raudenbush, S. W., Spybrook, J., Bloom, H., Congdon, R., Hill, C., & Martínez, A. (2011). Optimal Design Software for Multi-level and Longitudinal Research (Version 3.01) [Software]. Available from [www.wtgrantfoundation.org](http://www.wtgrantfoundation.org)
- Rugs, D., & Kaplan, M. F. (1993). Effectiveness of informational and normative influences in group decision making depends on the group interactive goal. *British Journal of Social Psychology*, 32(2), 147-158. <https://doi.org/10.1111/j.2044-8309.1993.tb00992.x>
- Saguy, T., Dovidio, J. F., & Pratto, F. (2008). Beyond contact: Intergroup contact in the context of power relations. *Personality and Social Psychology Bulletin*, 34(3), 432-445.  
<https://doi.org/10.1177/0146167207311200>
- Schaeffli, L., Godlewska, A., Korteweg, L., Coombs, A., Morcom, L., & Rose, J. (2018). What do first-year university students in Ontario, Canada, know about First Nations, Métis, and

- Inuit peoples and topics? *Canadian Journal of Education*, 41(3), 688-725.  
<https://files.eric.ed.gov/fulltext/EJ1192158.pdf>
- Sherif, M. (1935). A study on some social factors in perception. *Archives of Psychology*, 187, 1-60.  
[http://web.mit.edu/curhan/www/docs/Articles/15341\\_Readings/Influence\\_Compliance/Sherif\\_A\\_Study\\_of\\_Some\\_Social\\_Factors\\_\(1935\)\\_Arch\\_Psych.pdf](http://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Influence_Compliance/Sherif_A_Study_of_Some_Social_Factors_(1935)_Arch_Psych.pdf)
- Statistics Canada. (2017). *2016 Census topic: Aboriginal Peoples* (Catalogue number 98-404-X2016001). <https://www12.statcan.gc.ca/census-recensement/2016/rt-td/ap-pa-eng.cfm>
- 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.  
<https://doi.org/10.1111/j.1559-1816.1986.tb01144.x>
- Vallacher, R. R., Coleman, P. T., Nowak, A., & Bui-Wrzosinska, L. (2011). Rethinking intractable conflict: The perspective of dynamical systems. *Conflict, Interdependence, and Justice*, 11, 65-94. [https://doi.org/10.1007/978-1-4419-9994-8\\_4](https://doi.org/10.1007/978-1-4419-9994-8_4)
- Wakslak, C. J., Jost, J. T., Tyler, T. R., & Chen, E. S. (2007). Moral outrage mediates the dampening effect of system justification on support for redistributive social policies. *Psychological Science*, 18(3), 267-274. <https://doi.org/10.1111/j.1467-9280.2007.01887.x>
- Willis, R. H. (1963). Two dimensions of conformity-nonconformity. *Sociometry*, 26(4), 499-513.  
<https://doi.org/10.2307/2786152>
- Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology*, 8(1), 52-69.  
<http://dx.doi.org/10.20982/tqmp.08.1.p052>



- Wortley, S. (2006). Police use of force in Ontario: An examination of data from the special investigations unit. *Final report to the African Canadian Legal Clinic for Submission to the Ipperwash Inquiry*.  
[http://www.attorneygeneral.jus.gov.on.ca/inquiries/ipperwash/policy\\_part/projects/pdf/AfricanCanadianClinicIpperwashProject\\_SIUStudybyScotWortley.pdf](http://www.attorneygeneral.jus.gov.on.ca/inquiries/ipperwash/policy_part/projects/pdf/AfricanCanadianClinicIpperwashProject_SIUStudybyScotWortley.pdf)
- Yamamoto, S. (2019). *What's in the Box? Punishment and Insanity in the Canadian Jury Deliberation Room*. Thesis (Ph.D.) - Carleton University, 2019.
- Xu, S., & Lorber, M. F. (2014). Interrater agreement statistics with skewed data: Evaluation of alternatives to Cohen's kappa. *Journal of Consulting and Clinical Psychology, 82*(6), 1219-1227. <https://psycnet.apa.org/buy/2014-31345-001>
- Zinger, I. (2017). Annual report: Office of the correctional investigator 2016-2017. *Annual Report of the Office of the Correctional Investigator, 2016*. <https://www.ocibec.gc.ca/cnt/rpt/annrpt/annrpt20162017-eng.aspx>

## Appendix A

### Demographics Questionnaire

Finally, we would like to ask you a few questions about yourself, so that we can get a sense of who the people in our study are.

1. What is your age?
2. What is your gender?
  - Man
  - Woman
  - Trans
  - Prefer not to disclose
  - Or please specify:
3. What is your highest level of education that you have obtained?
  - Doctoral or professional degree
  - Master's degree
  - Bachelor's degree
  - Associate's degree
  - Postsecondary non-degree award
  - Some college, no degree
  - High school diploma or equivalent
  - Less than high school
4. What (if applicable) is your religious affiliation?
  - Protestant
  - Roman Catholic

- Mormon
- Orthodox
- Jewish
- Muslim
- Buddhist
- Hindu
- Atheist
- Agnostic
- Nothing
- Or please specify:

5. What is your racial/ethnic background?

- Chinese
- South Asian
- Black
- Filipino
- West Asian
- Latin American
- Arab
- White
- Southeast
- Indigenous Peoples of Canada
- Korean
- Japanese

- Please elaborate here:

6a. Please indicate where your political beliefs fall, using the scale below.

0      1      2      3      4      5      6      7      8      9      10

Liberal

Conservative

6b. What, if applicable, is your political affiliation?

7. Are you a Canadian citizen?

- No
- Yes

8. Can you read, write, and understand the English language?

- No
- Yes

9. Have you ever been convicted of an indictable offence?

- No
- Yes

**Appendix B**

Adapted from: Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology*, 88(3), 498-509.

This set of questions pertains to your views about *Canadian* society.

Please rate the following statements using a scale from 1 to 5, where 1 = Strongly Disagree, 3 = neutral, and 5 = Strongly Agree.

In general, relations between racial groups are fair.

Race relations need to be radically restructured.\*

For people of colour, Canada is the best country in the world to live in.

Most policies relating to race serve the greater good.

In Canada, everyone has a fair shot at wealth and happiness, regardless of race.

Racism in society is getting worse every year.\*

Society is set up so that people from all racial groups usually get what they deserve.

\*Item reverse-scored

## Appendix C

### Null Model

$$\text{Level 1 model: NormRatio}_{tij} = \pi_{0ij} + e_{tij}$$

$$\text{Level 2 model: } \pi_{0ij} = \beta_{00j} + r_{0ij}$$

$$\text{Level 3 model: } \beta_{00j} = \gamma_{000} + u_{00j}$$

### Equation 1

$$\frac{\tau_{00}}{\tau_{00} + \sigma^2}$$

### Equation 2

$$\frac{\tau_3^2}{\tau_2^2 + \tau_3^2 + \sigma^2}$$

### The Complete Model

$$\text{Level 1 model: NormRatio}_{tij} = \pi_{0ij} + \pi_{1ij} * (\text{TimeBlock}_{tij}) + e_{tij}$$

$$\text{Level 2 model: } \pi_{0ij} = \beta_{00j} + \beta_{01j} * (\text{SJS}_{ij}) + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j} * (\text{SJS}_{ij}) + r_{1ij}$$

$$\text{Level 3 model: } \beta_{00j} = \gamma_{000} + u_{00j}$$

$$\beta_{01j} = \gamma_{010} + u_{01j}$$

$$\beta_{10j} = \gamma_{100} + u_{10j}$$

$$\beta_{11j} = \gamma_{110} + u_{11j}$$