Quantum Equality in the Courtroom

by

Magpie Michael Winslow

B.Sc., Montana State University, 2021

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

The College of Graduate Studies

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

April 2024

The following individuals certify that they have read, and recommend to the College of Graduate Studies for acceptance, a thesis/dissertation entitled:

QUANTUM EQUALITY IN THE COURTROOM

submitted by MAGPIE MICHAEL WINSLOW in partial fulfilment of the requirements of the degree of Master of Arts

Dr. Leanne ten Brinke, Irving K. Barber School of Arts and Sciences
Supervisor

Dr. Brian O'Connor, Irving K. Barber School of Arts and Sciences
Supervisory Committee Member

Dr. Douglas Scott, Department of Physics and Astronomy, UBC Vancouver
Supervisory Committee Member

Dr. Elena Nicoladis, Irving K. Barber School of Arts and Sciences
University Examiner

Dr. Madeleine Ransom, Simon Fraser University
External Examiner

Abstract

Survivors of sexual assault are often pitted against their alleged attacker in a contest of credibility. However, there is little research on the relationship between the survivor's credibility and the attacker's. The present study investigates this relationship using a model from quantum cognition called the QQ equality. This model predicts that the order of credibility judgements is important and that judgements about the survivor impact the alleged attacker's credibility to the same degree that judgements about the alleged attacker impact the survivor's credibility. A sample (n = 1493) of US adults were asked to judge the credibility of statements from both the survivor and attacker in hypothetical cases of sexual assault. Contrary to predictions, there was no effect of judgment order on the perceived credibility of either party. However, correlations between random subsamples did provide evidence of the QQ equality. These results suggest that the two judgements neither share a feature space nor exist in separate spaces; instead, they may exist in feature spaces with both unique and shared components. This study's findings suggest that further research into modelling feature spaces may provide insight into factors impacting a survivor's perceived credibility.

Lay Summary

While there is a fair amount of research on the perceived lack of credibility of sexual assault survivors, a majority of the studies look at survivors in isolation, even though people are often asked to compare the credibility of survivors against that of their attacker. To explore this relationship, the present study used a prediction from the field of quantum cognition called the QQ equality. Quantum cognition is a relatively new field that attempts to describe decision-making using the tools of quantum physics, with the QQ equality being a prediction about the relationship between similar questions. We found that changing the order in which people receive information can change who they think is credible. We proposed a new way for quantum cognition to describe the information people use in decision-making.

Preface

This research was conducted at the University of British Columbia Okanagan and was supervised by Dr. Leanne ten Brinke. For the present thesis, I was responsible for designing and developing the research paradigm, data preparation and analysis, and the written thesis. I also provided an outline of the statements presented to participants, and Dr. ten Brinke helped to flesh out the specifics. The Behavioural Research Ethics Board of the University of British Columbia's Okanagan Campus granted ethics approval for this research. The certificate approval number for the project is H22-02650. To date, the results of this study have not been published.

Table of Contents

Abstract . . . . **iii**

Lay Summary . . . . **iv**

Preface . . . . **v**

Table of Contents . . . . **vi**

List of Tables . . . . **viii**

List of Figures . . . . **x**

Acknowledgements . . . . **xi**

Dedication . . . . **xii**

Chapter 1: Introduction **1**
- 1.1 Sexual Assault in the Legal System 1
  - 1.1.1 Perceived Credibility of Sexual Assault Allegations . . 2
- 1.2 A New Perspective on Credibility Assessment 3
  - 1.2.1 Competing Narratives 4
  - 1.2.2 Rational Choice Theory 5
- 1.3 Quantum Cognition 6
- 1.4 Current Study 9

Chapter 2: Methods **11**
- 2.1 Participants 11
  - 2.1.1 Design & Materials 12
  - 2.1.2 Procedure 13
- 2.2 Analysis 15

Chapter 3: Results **18**

3.1 Question Order 18
3.2 Statement Placement 19
3.3 Subsampling 21

Chapter 4: Discussion 25
4.1 Quantum Cognition 26
4.2 Limitations and Future Directions 29
4.3 Conclusion 30

Bibliography 31

Appendices 38
Appendix A: Vignette and Questions 38
A.1 Vignettes 38
A.2 Questions 40
Appendix B: Task Instructions 42
B.1 Conditions 1 & 2 instructions 42
B.2 Conditions 3 instructions 42

List of Tables

Table 2.1 A table view of the different conditions in this study. “Survivor/Attacker Statement” here refers to the variation of the Mary and Bill statements used in that part of the study. “Survivor/Attacker Credibility” means that participants will assess that party’s credibility. 15

Table 2.2 Contingency table for the analysis of order effect. a represents the number of people that gave the response pattern for a given vignette in the first condition. b represents the number of people that gave the response pattern for a given vignette in the second condition. The subscripts give the response pattern with the response about the survivor first and the responses about the attacker second. For instance, “*YN*” is “*Yes-No*” and corresponds to thinking the survivor is credible and the attacker isn’t. Each column has a response pattern, while each row has a different question order. 16

Table 2.3 Contigency table for the analysis of the QQ equality. a represents the proportion of people that gave a specific response pattern in the first question order. b represents the proportion of people that gave a specific response pattern in the second question order. The subscripts represent the same response patterns in Table 2.2. “Major” refers to the total proportion of people who gave the same response to the first and second questions in that order (responded either yes-yes or no-no). “Minor” refers to the total proportion of people who responded differently to the first and second questions (either yes-no or no-yes). 17

Table 2.4 Contigency table for the analysis of the effect of information on the QQ equality. In this table, the variable c represents the number of people who responded with that response pattern for a given vignette in the third condition. "Mary - Bill" means that people determined Mary and Bill's. "Mary - Statement - Bill" means that Bill's statement was placed between credibility judgments. All other notations follow the conventions established previously. 17

Table 3.1 Response pattern frequencies for each vignette in Conditions 1 and 2. In the header, the first response is whether the participants thought the survivor was credible, and the second is whether they thought the attacker was credible ("*Yes-No*" corresponds to thinking the survivor is credible and the attacker isn't. 19

Table 3.2 Response pattern frequencies for each vignette in Conditions 1 and 3. In the header, the first response is whether the participants thought the survivor was credible, and the second is whether they thought the attacker was credible ("*Yes-No*" corresponds to thinking the survivor is credible and the attacker isn't.) 20

Table 3.3 Correlations and means from subsamples of the data. Results significant for $p \leq 0.01$ are bolded, and all correlations have a df = 98. The meaning of "Major," "Minor," and the subscripts for the means retain previously established meanings. 24

List of Figures

Figure 3.1	Scatter plots of context effects for the subsampled data of Conditions 1 and 2. The top row shows the relationship between the Yes-Yes response context effects and the No-No context effects. The bottom row shows the relationship between the Yes-No context effects and the No-Yes context effects.	22
Figure 3.2	Scatter plots of context effects for the subsampled data of Conditions 1 and 3. The top row shows the relationship between the Yes-Yes response context effects and the No-No context effects. The bottom row shows the relationship between the Yes-No context effects and the No-Yes context effects.	23

Acknowledgements

I can not begin to express how grateful I am to everyone who helped me bring this project over the finish line. I am thankful to Dr. Leanne ten Brinke, who gave me a chance and gave me the ability to try this weird idea. I would genuinely be in a worse place without your support and patience. I also need to thank Dr. Brian O'Connor not only for agreeing to be on my committee but also for his help with ideas for data analysis. The findings, and especially the partial feature space hypothesis, would not exist without your suggestions and advice. I am also so grateful to Dr. Douglas Scott for agreeing to try something new and supervise a psychology thesis. You've helped keep me honest about the implications of the work. Additionally, I want to thank Dr. Madeleine Ransom for agreeing to be my External Examiner as you posed some of the most intriguing questions about this work and Dr. Elena Nicoladis for being my neutral chair.

I could not have survived this thesis without the community I've met along the way. I'm so grateful to my entire cohort, to whom I owe the survival of my first year. I could not think of a better group of people to help introduce me to Canada. I am so lucky to have had friends like Kirthana, Iana, Natashaa, Rebecca, and Amané, who so often gave me the perspective and kind space I needed. I do not know what I would've done without you. I also need to thank my family for always being supportive of my weird theories. Finally, I need to thank Mx. Noble for so much, but also for teaching that math serves the theory, not the other way around.

Dedication

To the hope for better days. May they come to my siblings and all who need them.

We're here because we're here because we're here...

Chapter 1

Introduction

Sexual assault is a highly prevalent, rarely reported, and rarely tried crime (Rotenberg, 2017). When reported, it is often met with skepticism (Rotenberg, 2017). When tried, there is rarely reliable physical evidence, and it becomes a credibility contest between the survivor and the attacker (Vandervort, 2005). Previous research has identified situational, behavioural, and appearance-related variables that contribute to perceptions of the credibility of survivors and attackers (Bagby et al., 1994; Baker et al., 2021; Berry-Cabán et al., 2020; Golding et al., 2016; Lynch et al., 2020; Wasarhaley et al., 2012). However, what has not been considered is whether making credibility judgements about one party may impact judgements about the other party, ultimately influencing case outcomes. We draw upon the burgeoning field of quantum cognition to test whether a predictable pattern of order effects exists in credibility judgements around sexual assault.

1.1 Sexual Assault in the Legal System

It is estimated that only 6% of sexual assault cases are reported to the police (Cotter, 2021). In Canada, 34,242 cases of sexual assault were reported to the police in 2018 (Moreau, 2020), suggesting over half a million (570,700) cases of sexual assault occurred. Of the reported cases, 59% involved a person being accused, 43% of those resulted in charges, 49% of those went to trial, and 55% of those resulted in a conviction (Rotenberg, 2017). In other words, approximately half of cases are dropped at each step of the legal process, resulting in approximately 7% of accusations resulting in a conviction. While some of this attrition is from survivors choosing alternate routes for justice, a perceived lack of evidence is another major factor. Before reaching court, if the police or The Crown think there is not enough evidence to prove the alleged attacker guilty beyond a reasonable doubt, the charge will be changed to a less severe offence or be dropped altogether (Rotenberg, 2017). Sexual assault legal cases often have little tangible evidence and a critical determinant in whether sexual assault cases see trial, much less conviction, is the survivor's perceived credibility (Herlihy & Turner, 2015; Nitschke et al., 2019).

1.1.1 Perceived Credibility of Sexual Assault Allegations

Current research suggests that false sexual assault claims are rare. An estimated 5% of rape allegations in the United States are deemed unfounded (De Zutter et al., 2017). Less than 5% of rape cases brought before authorities were considered fraudulent in over twenty European countries (Kelly, 2010). In Canada, 8% of level one sexual assaults reported were classified as unfounded (Moreau, 2020). Despite such findings, McMillan (2018) found that police in the United Kingdom estimated 53% of rape allegations were false. The reasoning the police used for this estimate was often founded in assumptions of how "real" survivors of sexual assault ought to act (McMillan, 2018). The assumptions themselves were based on sexual assault myths and included how injured a survivor should be, the amount of detail in their statement, and the sobriety and mental health of the survivor, amongst others (McMillan, 2018). The effects of these assumptions carry over to lawyers who have also been found to use rape myths to determine the viability of a case, no matter their level of experience (Krahé et al., 2008). Judges will also rely on unfounded assumptions around the impact of traumatic memories on behaviour to inform their rulings (Herlihy & Turner, 2015). At all levels of the legal system, research suggests that survivors' credibility is a function of how those judging them think they should act.

To understand the credibility of sexual assault survivors, previous research has attempted to unpack how a variety of behavioural and appearance cues can influence judgements. For survivors who are perceived as women, appearance cues such as being perceived as overweight (Clarke & Stermac, 2011) or unattractive (Maeder et al., 2015) are related to being disbelieved. Behaviourally, a survivor's credibility can be tied to whether they present a distressed affect, whether their actions are congruent with expectations around sexual assault, and the degree to which they trigger other people's compassion (Ask & Landström, 2010; Kaufmann et al., 2003; Nitschke et al., 2019; Winkel & Koppelaar, 1991). The use of affect to assess credibility is known as the emotional survivor effect and is present regardless of professional experience or presentation medium (Nitschke et al., 2019; Winkel & Koppelaar, 1991).

Additionally, whether or not a survivor has a consistent narrative is often used to determine credibility. Consistency as a way to assess credibility assumes that, if a person's story changes, they are not telling the truth and is often used by both the police and judges to determine the credibility of sexual assault survivors (Herlihy & Turner, 2015; Kelly, 2010; McMillan, 2018). In reality, while traumatic memories tend to be more vivid than memories of other events, they are also highly susceptible to distortion (Strange & Takarangi, 2015). As a result, survivors have often been advised against seeking therapy until after the trial (Herlihy & Turner, 2015).

While these various biases are common, they do not lead to accurate assessments of a survivor's credibility. In Peace et al. (2012), participants were presented with two authentic and two false narratives about sexual assault. Not only were participants unable to accurately determine the credibility of the statements, but they also performed significantly worse than chance. Further, the more emotionally intense the narrative, the less likely participants were to judge the statement's credibility correctly. If the biases at play were simply inaccurate, then one would expect participants to perform as well as chance; however, these results suggest that biases may be in opposition to reality or that there may be further processes at play in biasing participants' responses.

1.2 A New Perspective on Credibility Assessment

These previous studies assume that people take a set of discrete observational cues to evaluate credibility through semi-rational processes. For example, Maeder et al. (2015) discusses multiple arguments participants could use to determine the credibility of the survivor and attacker from their respective attractiveness and race. Clarke and Stermac (2011) suggest different ways a survivor's weight may factor into various sexual assault myths by presenting potential ways participants could use the survivor's weight in their reasoning process. Lynch et al. (2020) discusses how the perceived gender of the survivor and attacker, as well as the cost of a date, influence each party's credibility.

This research is an attempt to step back and understand if these judgements can be influenced by something other than observational cues. Previous research has used quantum cognition to describe and predict the ways that the placement of information and the order of questions influence the credibility of political candidates (Wang & Busemeyer, 2013; Wang et al., 2014). This prediction called the QQ equality (Wang & Busemeyer, 2013), concerns pairs of related questions, such as judgments about the credibility of a survivor versus an attacker. This research will test if the QQ equality holds for these credibility judgements.

1.2.1 Competing Narratives

In the literature reviewed thus far, most studies have focused on the survivor's perceived credibility. Everything from the presentation medium (Nitschke et al., 2019) to the survivor's appearance and description (Clarke & Stermac, 2011; Maeder et al., 2015) are considered and varied, but the attacker is rarely accounted for. This does not reflect the environmental reality of sexual assault cases, because the attacker's narrative is often presented alongside the survivor's in public discourse (Casselman & Tankersley, 2019; Dale, 2022; Kessler, 2018; Peters, 2022; Urquhart, 2022). In Canadian law, once evidence has shown a lack of consent on the survivor's part, the burden of evidence lies on the alleged attacker to show they made a "mistake of fact" and were unaware that consent had not been given or had been withdrawn (Vandervort, 2005). This implies that the alleged attacker's credibility is just as important as the survivor's in deciding sexual assault cases.

The limited research that has presented both attacker and survivor narratives has found that situational context, attacker and survivor characteristics, and judge demographics are associated with perceived credibility. For example, in date-rape scenarios, attackers were perceived as more credible in their denials when they were of higher status and paid for more expensive dates (Lynch et al., 2020). Men, in general, tend to be more lenient on male attackers, finding them more credible when compared to female participants' judgements (Golding et al., 2016; Lynch et al., 2020; Wasarhaley et al., 2012). Additionally, attackers described as obese were considered more credible than those who were described without referencing weight (Yamawaki et al., 2018). When looking at the dyad of survivor and attacker, cases of interracial rape are more likely to be believed than intraracial cases (George & Martínez, 2002). In general, there is an understanding that situational and behavioural factors impact attacker credibility; however, there is little to no study of the relationship between the credibility of the attacker and the survivor.

Most studies that consider both parties' perceived credibility treat them as independent measures. That is, the attacker's perceived credibility is treated as if it has no bearing on the perceived credibility of the survivor. This is not a trivial assumption as it implies that the alleged attacker's verdict is independent of perceptions of the survivor's statement and goes against the prevailing framework used by legal decision-makers: rational choice theory (Jaeger & Trueblood, 2018).

1.2.2 Rational Choice Theory

Rational choice theory assumes that humans will make entirely rational choices based on the available evidence. Specifically, when people are uncertain about a decision, they will collect information and apply the assumptions of classical probability theory to come to a decision (Becker, 1976; Jaeger & Trueblood, 2018). This theory, and its underlying logic, has been the basis for many policies and is often used when making legal decisions (Jaeger & Trueblood, 2018), despite research finding that it is not generally applicable (Simon, 1955; Tversky & Kahneman, 1974, 1983). These inaccuracies have led legal scholars instead to develop behavioural law and economics (BLE: Jaeger & Trueblood, 2018) (Jaeger & Trueblood, 2018). BLE can be seen as an extension of rational choice theory since it uses research on heuristics and biases to describe specific instances where people do not follow the predictions of rational choice theory. However, BLE is severely limited in its predictive capabilities because it does not specify a mechanism to explain why people act irrationally, instead saying that people can be irrational in the ways described by heuristics and bias research (Jaeger & Trueblood, 2018). Because of this limitation, rational choice theory is the primary tool for making predictions in legal theory.

While the typical assumption in statistical tests is that the survivor and attacker's credibility are independent, applying rational choice theory predicts the opposite. Assuming the survivor and the attacker contradict each other, only one statement can be true, and the other must be false; thus, the two credibility scores should have a strong negative correlation. Applying BLE would predict the same negative correlation, with the addition that factors not directly related to the legal case, such as race, ethnicity, gender, affect, etc., may bias the credibility scores.

The limitations of rational choice theory and BLE have recently led some scholars to propose the theory of quantum cognition. This theory is based on the mathematical principles developed for quantum mechanics and can predict multiple different heuristics and biases (Pothos & Busemeyer, 2022). When applied to legal cases around sexual assault, quantum cognition can provide a unique perspective to understand credibility judgements.

1.3 Quantum Cognition

There is a fundamental difference between quantum cognition descriptions of uncertainty and measurement processes and the descriptions provided by rational choice theory. In rational choice theory, people deal with uncertainty by following the rules and assumptions of classical probability, specifically Bayes' theorem (Bruza et al., 2015). Bayes' theorem broadly describes how new evidence should impact the likelihood of a given event when reasoning using formal logic. Further, it assumes that the events are well-defined and cannot occur concurrently. By modelling cognition with Bayes' theorem, rational choice assumes that people effectively use formal logic to update individual beliefs using specific pieces of evidence.

While classical probability works with well-defined sets of distinct events, quantum mechanics is designed to work with more general, indeterminant systems (Busemeyer et al., 2011). In quantum mechanics, phenomena such as the position and momentum of an object only have specific values after being measured, otherwise existing in multiple places and speeds at the same time (Griffiths & Schroeter, 2018). Measurements then alter the phenomenon being measured and impact future measurements of related phenomena (Heisenberg, 1983). When applied to human cognition, one interpretation is that a person's answer to a question only solidifies once they respond to that question. This interpretation suggests that before answering, the person's mental state is diffused over the possible responses, with a preference for certain options that make them more likely.

A clear example of the differences between the classical perspective of rational choice theory and quantum theory is the famous Linda problem first described in Kahneman et al. (1982). In this experiment, Linda is described as having been a philosophy student who was "deeply concerned with issues of discrimination and social justice" and had "participated in anti-nuclear demonstrations." Participants were then asked if she was more likely to either be a bank teller or be a bank teller who is active in the feminist movement. If people use classical probability to reason, participants should say Linda is more likely to be just a bank teller. In classical probability theory, the probability of two things occurring (being a bank teller and an activist) will always be less than the probability of either occurring individually (only being a bank teller or only being an activist). However, Kahneman et al. (1982) found that people primarily said Linda was more likely to be both a bank teller and active in the feminist movement instead of saying she was solely a bank teller. To explain this discrepancy, Busemeyer et al. (2011) applied a quantum probability framework, which took into account that Linda's description was more similar to being both active in the feminist movement and a bank teller than just being a bank teller. The tools that come from quantum probability allowed for a more accurate description of how people use information to answer questions.

Much of the experimental and theoretical support for quantum cognition has come from applying quantum models to situations where classical models and rational choice are inaccurate. Li et al. (2020) compared quantum and classical models' ability to predict people's decision-making. Aerts and Gabora (2005) used quantum principles to explain conceptual combinations and related fallacies. Generally, quantum cognition has been used to describe and quantitatively predict numerous heuristics and fallacies (Jaeger & Trueblood, 2018; Pothos & Busemeyer, 2022). The findings around the impact of question order described in Wang and Busemeyer (2013) are of particular interest for this research.

The impact of the order of measurements on results is one of the early effects described by quantum mechanics (Griffiths & Schroeter, 2018; Heisenberg, 1983). Wang and Busemeyer (2013) used what they referred to as the "law of reciprocity" in quantum mechanics. This law states that the probability of transitioning from state x to state y is the same as the probability of transitioning from state y to state x. In terms of dichotomous questions, this means that the likelihood of someone who said Yes to question A will then say Yes to question B is equal to the probability that of someone who said Yes to question B will then say Yes to question A. In other words, the impact of answering one question on responses to the other is independent of the order in which the questions are asked. Wang and Busemeyer (2013) used this law to model pairs of questions and arrived at the following equations,

$$p(A_y B_y) + p(A_n B_n) = p(B_y A_y) + p(B_n A_n), \tag{1.1}$$

$$p(A_y B_n) + p(A_n B_y) = p(B_y A_n) + p(B_n A_y). \tag{1.2}$$

In these equations, $p(x)$ gives the probability of x, A and B are questions A and B, subscript y indicates responding yes to a question, and subscript n means responding no. The two equations form the quantum question order equality (the QQ equality). In plain language, the first equation states that when you have two sequential, dichotomous questions on a similar topic, the probability of saying yes to both questions plus the probability of saying no to both questions is constant regardless of question order. Classical probability does not produce an equivalent prediction as it does not model the impact of answering a question on mental states. To test this relationship, Wang et al. (2014) analyzed over 70 surveys and research papers, finding the predicted relationship in all applicable surveys. This is one of the most strongly supported a priori predictions in quantum psychology and supports the theory's underlying principles.

According to Wang and Busemeyer (2013), this relationship assumes no information is presented between the first and second questions. If information comes after the first question, people's responses to the second question will be altered by the new information, not just their response to the first question. As such, adding information between the two questions means that the QQ equality is no longer expected to occur. To test this, Wang et al. (2014) tested the QQ equality for surveys where information was presented between the questions and did not find the equality present, as predicted.

Returning to the sexual assault literature, the QQ equality may be able to shed light on the relationship between the perceived credibility of the survivor and the attacker. If the survivor and attacker's credibility are related, quantum cognition predicts that answering a question about the credibility of one will impact the credibility of the other, producing an order effect. This order effect should follow the QQ equality such that the portion of participants who judge both parties as credible plus the portion who judge both as not credible will be the same whether they are judging the survivor or the attacker first. In other words, the order effect will be constrained by the QQ equality. While previous research has focused on the factors relating to the survivor and attacker credibility separately, to our knowledge, this is the first study to test if these judgements are connected and to use quantum cognition to make predictions about the nature of this relationship.

If the QQ equality holds in this context, then quantum cognition could be used more broadly in legal theory and challenge rational choice and BLE. If quantum cognition applies to legal decision-making, it could change our understanding of sequential line-ups, factfinder's final judgements, and the use of biases in legal knowledge as a whole (Jaeger & Trueblood, 2018). If the QQ equality does not hold, that could suggest previously unpredicted constraints on the QQ equality, which would challenge our current understanding of quantum cognition.

1.4 Current Study

This study will attempt to determine whether the relationship between the credibility of the survivor and attacker follows the relationship predicted by the QQ equality. As such, there are three main hypotheses:

H1: Successive judgements of the survivor's and attacker's credibility will produce an order effect.

H2: This order effect will follow the QQ equality.

H3: The QQ equality will not hold when new information is presented between questions.

H1 is a prerequisite for meaningful results because the QQ equality is trivially true if there is no significant order effect. At a high level, if there is no order effect, then the order the questions are in does not matter. Thus, the probability of a given response pattern will not be impacted by the order of questions, and the QQ equality will always be true. As such, support for H1 is necessary to meaningfully test for the QQ equality. Support for H2 would suggest that credibility assessments follow a similar pattern as the QQ equality. Support for H3 would indicate that the pattern in H2 requires the same conditions as the QQ equality. Support for all three hypotheses would suggest that these credibility judgements follow the QQ equality.

The tests for the QQ equality used for H2 and H3 was the χ^2 test used in Wang et al. (2014) and described in that paper's supplemental material. This χ^2 test is specifically designed to test if the QQ equality is not present in the data, meaning that a significant χ^2 value indicates that the data does not follow the QQ equality. A common practice in quantum cognition where the lack of significance when testing if a model does not apply to data is interpreted to mean that data follows the model. In the case of Wang et al. (2014), they interpreted a χ^2 with $p > 0.05$ to mean that the QQ equality did not apply to the data. However, this can not be concluded from a p-value significance test and requires the calculation of a Bayes factor instead.

The p-value is defined to be the probability of an experiment producing a test statistic value equal to or greater than the observed value, assuming the null hypothesis is true (mathematically, the p-value represents $p(D|H_0)$ where D represents the observed data and H_0 is the null hypothesis). This can not be turned into a statement about the probability of the null hypothesis because it assumes that the null hypothesis is true (**Connor2017**; **Pollard1987**). A non-significant χ^2 value in the χ^2 test for the absence of the QQ equality is not equivalent to a significant value in a hypothetical test for the presence of the QQ equality. To test for the presence of the QQ equality using the χ^2 from Wang et al. (2014), a Bayes factor needs to be calculated. The Bayes factor originates in Bayesian statistics and allows for the comparison of two hypotheses. It is interpreted as the odds ratio of the likelihood of finding a result given one hypothesis versus the likelihood of finding the same result given a different hypothesis (**Connor2017**). In the case of this study, the likelihood under the null hypothesis is over the likelihood under the alternative hypothesis. Since the Bayes factor compares how likely a result is under two different assumed hypotheses, it allows us to perform null hypothesis testing.

In this study, most p-values are accompanied by the Bayes factor for that test statistic presented as $BF_{01} = x$. The subscript of the variable tells you how to interpret the Bayes factor where BF_{01} is the null hypothesis for the statistic over the alternative hypothesis, while BF_{10} is the alternative hypothesis over the null. Each Bayes factor is calculated such that a larger value indicates more substantial support for the experimental hypothesis. If it is a test for the QQ equality, for instance, then a larger Bayes factor means more robust support for the presence of the QQ equality in the data. Finally, a loose rule of thumb for interpreting a Bayes factor is that $1 < BF_{10} \leq 3$ is weak support for the hypothesis, $3 < BF_{10} < 10$ is moderate support for the hypothesis, $10 \leq BF_{10}$ is strong support the hypothesis, and $0 < BF_{10} < 1$ is evidence against the hypothesis with small values suggesting strong evidence(**Van2021**).

Chapter 2

Methods

2.1 Participants

We recruited 1493 participants aged 14 to 79 (mean = 40.4) from the United States using the online platform Prolific in exchange for monetary compensation. There were moderately more men ($n = 792$) than women ($n = 621$). The sample was predominately cisgender ($n = 1378$), with a minority of people identifying as trans or nonbinary ($n = 81$). Most participants were White ($n = 1101$), with people identifying as the following in descending order: Black/African American ($n = 175$), Hispanic ($n = 144$), Asian ($n = 139$), Native American/Alaskan Native ($n = 18$), Native Hawaiian/Pacific Islander ($n = 12$).

This sample size was estimated beforehand using previous studies. The studies used in Wang et al. (2014) were primarily studies with an average of ≈1,000 participants with an average of ≈ 500 participants in each condition. Since this study used three conditions, we needed approximately 1500 participants to have an equivalent number of participants in each condition. Since the analysis consists primarily of the novel statistical test developed in Wang et al. (2014), a full power analysis to determine the sample size was not in the scope of this study. However, as the novel test achieves the same effect as a χ^2 test of independence, a power analysis of that statistic can also provide an estimate for the sample size.

Similar to Wang et al. (2014), support for H1 requires a significant order effect with $DF = 3$ and support for H2 requires a non-significant result for the QQ equality with $DF = 1$. The reduction in degrees of freedom between testing for an order effect and testing for the QQ equality is because Equation 1.1 and 1.2 each function as separate constraints on the data, reducing the degrees of freedom. A significant test result for H2 suggests that the two sides of the equality are significantly different, supporting the idea that the QQ equality was not present. We tested if the QQ equality did not apply to the data by testing for a small difference in the two sides of the equation, which required powering the study for a small effect size for the QQ equality. However, since the test of the order effect has more degrees of freedom, a study that can find a small order effect can also find a small difference in the QQ equality. According to G*Power v3.1, to find an order effect of effect size $w = 0.1$ with a power of 0.8, the study needed a sample of 1091 participants, or approximately 500 participants in each condition (Faul et al., 2007). This generally aligns with the studies used in Wang et al. (2014) and confirms that we needed around 1500 participants in this study.

2.1.1 Design & Materials

We used a between-subject design with two phases. The first phase had two conditions that changed the credibility assessment order. The second phase consisted of one condition where new information was presented between the questions. Credibility was evaluated as a dichotomous measure with two levels: the person is either credible or not.

Participants were presented with three vignettes describing three different assaults. Each vignette consisted of a written statement from a survivor and a written statement from the attacker. The first vignette was adapted from Messman-Moore and Brown (2006) where the survivor was named Mary and the attacker was named Bill. The second and third vignettes were written for this study, with Dakota and Bill being the survivors and Alex and Alice being the attackers, respectively. To ensure each vignette was different, we varied four factors present in the Mary-Bill vignette, which are known to impact bias in sexual assault cases. To ensure each vignette had similar levels of ambiguity and bias, when one factor was changed to increase participant bias, others were changed to decrease bias. The strength of each factor's relationship to participant bias was approximated from the results of the studies and the number of studies that did not find an effect from the variable. The four variables were implied gender identity (Davies et al., 2009; George & Martínez, 2002; Pica et al., 2020, 2021; Sommer et al., 2016), prior relationship (D. J. Angelone et al., 2012; Bouffard, 2000; Campbell et al., 2009; Morabito et al., 2019; O'Neal, 2019; Sommer et al., 2016), intoxication (Schuller & Stewart, 2000; Wall & Schuller, 2000), and whether or not the survivor attempted to resist the attacker (D. Angelone et al., 2018; Davies et al., 2009; George & Martínez, 2002; Rogers et al., 2014). The text of all three vignettes and the corresponding questions can be found in Appendix A.

The initial vignette from Messman-Moore and Brown (2006) can be described as follows: the survivor (Mary) was implied to be a woman, and the attacker (Bill) was implied to be a man, they had just met, Mary was intoxicated, Bill was sober, and Mary did not describe resisting. In this case, previous research suggests that the fact that Mary was a woman likely biased participants in her favour (Sommer et al., 2016) while the fact that she did not resist biased participants against her (Davies et al., 2009; Rogers et al., 2014; Schuller & Stewart, 2000). Previous research suggests that the fact that Bill was male likely produced some participant bias for him (Pica et al., 2020; Sommer et al., 2016). The fact that Mary was sober but Bill was intoxicated likely produced some bias against Bill (Wall & Schuller, 2000).

The second vignette featured a survivor named Dakota and an attacker named Alex. It was described as a third date where neither was significantly intoxicated. Neither had an implied gender, and the survivor described attempting to resist the attack. Previous research suggests that the length of the relationship, the resistance, and the lack of alcohol would produce an effect in the survivor's favour (Bouffard, 2000; Campbell et al., 2009; Davies et al., 2009; Morabito et al., 2019; Rogers et al., 2014; Wall & Schuller, 2000).

The third vignette described a male survivor (Frank) and a female attacker (Alice). They were described as previously being friends, with both being moderately drunk and Frank attempting to resist Alice verbally. The literature generally describes male survivors as having low credibility (Davies et al., 2009; Sommer et al., 2016) and female attackers being perceived as more credible than male attackers (Sommer et al., 2016). Finally, the literature suggests that both parties being moderately intoxicated likely produced some bias towards the survivor (Wall & Schuller, 2000).

Finally, a personality assessment based on the five-factor model of personality was used as a distractor task between vignettes. Specifically, we used a shortened version of the International Personality Item Pool—Five-Factor Model measure (mini-IPIP) (Donnellan et al., 2006) because it provided a substantial number of questions on unrelated topics to distract participants.

2.1.2 Procedure

In phase one, participants were directed to a Qualtrics survey, where they provided fully informed consent before receiving instructions for the task (See Appendix B for all instructions provided to participants). Next, they were randomly assigned to one of two order conditions and moved into the first credibility assessment procedure. Both conditions started by reading both statements in the Mary-Bill vignette. After reading both statements, participants in the first condition ($n = 500$) were shown a new screen asking about Mary's credibility. Upon answering, they were shown a new screen that asked about Bill's credibility. Participants in the second condition ($n = 496$) were presented with the same vignettes in the same order. However, Condition 2 answered the questions in reverse order, judging Bill's credibility before Mary's. After judging Mary and Bill, both conditions were presented with the first ten questions of the mini-IPIP as a distractor task.

Participants then repeated this credibility assessment procedure for the Dakota-Alex vignette; however, participants in Condition 1 rated the attacker's (Alex's) credibility before the survivor's (Dakota's) credibility, with Condition 2 making their ratings in the opposite order, starting with Dakota. After both ratings, participants were presented with another distractor task consisting of the last ten questions of the mini-IPIP before moving on to the final credibility assessment procedure on the Frank-Alice vignette. In this final procedure, participants in Condition 1 were asked about the credibility of the survivor (Frank) before the credibility of the attacker (Alice). In contrast, participants in Condition 2 started with Alice's credibility before moving on to Frank's.

Finally, participants provided demographic information, were debriefed, and thanked for their time. The entire procedure for this study is depicted in Table 2.1. The counterbalancing of question orders between vignettes and the distractor tasks was to reduce carryover effects from previous vignettes and judgements.

Phase 2 tested if breaking the assumptions of the QQ equality impacted participants' response patterns. This phase involved Condition 3 ($n = 497$), which followed the same structure as Condition 1. However, instead of reading both vignettes and then making credibility judgements, participants made credibility judgements directly after reading the corresponding vignettes. For instance, participants read Mary's vignette, gave their credibility assessment of Mary, and then read Bill's statement and assessed him. Refer to Table 2.1 to see the entire procedure for this condition.

Table 2.1: A table view of the different conditions in this study. "Survivor/Attacker Statement" here refers to the variation of the Mary and Bill statements used in that part of the study. "Survivor/Attacker Credibility" means that participants will assess that party's credibility.

Condition 1	Condition 2	Condition 3
Survivor Statement 1	Survivor Statement 1	Survivor Statement 1
Attacker Statement 1	Attacker Statement 1	Survivor 1 Credibility
Survivor 1 Credibility	Attacker 1 Credibility	Attacker Statement 1
Attacker 1 Credibility	Survivor 1 Credibility	Attacker 1 Credibility
mini-IPIP 1-10	mini-IPIP 1-10	mini-IPIP 1-10
Survivor Statement 2	Survivor Statement 2	Attacker Statement 2
Attacker Statement 2	Attacker Statement 2	Attacker 2 Credibility
Attacker 2 Credibility	Survivor 2 Credibility	Survivor Statement 2
Survivor 2 Credibility	Attacker 2 Credibility	Survivor 2 Credibility
mini-IPIP 11-20	mini-IPIP 11-20	mini-IPIP 11-20
Survivor Statement 3	Survivor Statement 3	Survivor Statement 3
Attacker Statement 3	Attacker Statement 3	Survivor 3 Credibility
Survivor 3 Credibility	Attacker 3 Credibility	Attacker Statement 3
Attacker 3 Credibility	Survivor 3 Credibility	Attacker 3 Credibility

2.2 Analysis

To test for an order effect (H1), we used the χ^2 independence test from Wang and Busemeyer (2013) on contingency Table 2.2. This test was applied to the answers for each pair of credibility assessments. To determine if H1 is meaningfully supported, we calculated both a p-value and a Bayes factor for our χ^2 value. To calculate the Bayes factor, we applied the piecewise approximation of the Bayes factor described by Wagenmakers (2022).

Table 2.2: Contingency table for the analysis of order effect. a represents the number of people that gave the response pattern for a given vignette in the first condition. b represents the number of people that gave the response pattern for a given vignette in the second condition. The subscripts give the response pattern with the response about the survivor first and the responses about the attacker second. For instance, "*YN*" is "*Yes-No*" and corresponds to thinking the survivor is credible and the attacker isn't. Each column has a response pattern, while each row has a different question order.

Credibility Order	Yes - Yes	Yes - No	No - Yes	No - No
Survivor - Attacker	a_{yy}	a_{yn}	a_{ny}	a_{nn}
Attacker - Survivor	b_{yy}	b_{yn}	b_{ny}	b_{nn}

The test for the QQ equality (H2) used Table 2.3 with the data from Conditions 1 and 2. The "Major" column refers to the "major diagonal" or the proportion of people who said yes to both questions plus the proportion of people who said no to both questions (see equation 1.1). "Minor" refers to the "minor diagonal" or the proportion of people who said yes to the first question and no to the second plus the proportion of people who said no to the first question and then yes to the second (see equation 1.2). The QQ equality predicts that the major and minor sum will stay constant regardless of order (Wang et al., 2014), which aligns with the null hypothesis of the independence test. Because we tested for the null hypotheses, we calculated the Bayes factor and conducted a p-value significance test. Support for the QQ equality occurs when the Bayes factor supports the null hypothesis and there is a non-significant p-value. This analysis was performed for each vignette.

Table 2.3: Contigency table for the analysis of the QQ equality. a represents the proportion of people that gave a specific response pattern in the first question order. b represents the proportion of people that gave a specific response pattern in the second question order. The subscripts represent the same response patterns in Table 2.2. "Major" refers to the total proportion of people who gave the same response to the first and second questions in that order (responded either yes-yes or no-no). "Minor" refers to the total proportion of people who responded differently to the first and second questions (either yes-no or no-yes).

Credibility judgement Order	Major	Minor
Survivor - Attacker	$a_{yy} + a_{nn}$	$a_{yn} + a_{ny}$
Attacker - Survivor	$b_{yy} + b_{nn}$	$b_{yn} + b_{ny}$

Finally, to determine if new information breaks the QQ equality and test H3, we performed the same χ^2 but on Conditions 1 and 3 as depicted in Table 2.4. Since Condition 3 breaks the assumption that the questions directly follow each other, we would not expect the QQ equality to hold. Thus, H3 is supported if the Bayes factor supports the alternate hypothesis and there is a significant p-value. This analysis was also performed independently for each vignette.

Table 2.4: Contigency table for the analysis of the effect of information on the QQ equality. In this table, the variable c represents the number of people who responded with that response pattern for a given vignette in the third condition. "Mary - Bill" means that people determined Mary and Bill's. "Mary - Statement - Bill" means that Bill's statement was placed between credibility judgments. All other notations follow the conventions established previously.

Statement Placement	Major	Minor
Statements - Survivor - Attacker	$a_{yy} + a_{nn}$	$a_{yn} + a_{ny}$
Statement - Survivor - Statement - Attacker	$c_{yy} + c_{nn}$	$c_{yn} + c_{ny}$

Chapter 3

Results

Across all vignettes, 75.5% of people thought the survivor was credible, and 56.4% thought the attacker was credible. When judging Mary and Bill in the first vignette, 74.5% of participants overall said that Mary (the survivor) was credible, and 60.7% thought Bill (the attacker) was credible. In the second vignette about Dakota and Alex, 81.6% said Dakota (the survivor) was credible, and 55.0% said Alex (the attacker) was. In the final vignette on Frank and Alice, 70.2% said Frank (the survivor) was credible, and 53.5% said Alice (the attacker) was.

When participants judged the credibility of the survivor first after reading both statements (Condition 1), 76.0% of participants thought at least one survivor was credible, and 50.0% thought at least one attacker was credible. When the judgment order flipped (Condition 2), the percentages became 76.5% and 51.0%, respectively. Across all judgements that were made immediately following their corresponding statement (Condition 3), 73.9% thought the survivor was credible, and 68.3% thought the attacker was. The breakdown of response frequencies by condition and vignette can be found in Tables 3.1 and 3.2.

3.1 Question Order

Using the data in Table 3.1, the test described in Wang and Busemeyer (2013) was used to test if an order effect was present in successive credibility judgements as predicted by H1. In all three vignettes, there was substantial evidence that changing the question order from baseline did not impact credibility judgements, $\chi^2 = 2.44$ ($p = 0.49$, $BF_{10} = 0.0384$) for the Mary-Bill vignette, $\chi^2 = 2.26$ ($p = 0.54$, $BF_{10} = 0.037$) for the Dakota-Alex vignette, and $\chi^2 = 1.24$ ($p = 0.74$, $BF_{10} = 0.0341$) for the Frank-Alice vignette. As such, we did not find support for H1.

To test H2 and determine if the data followed the QQ equality, the χ^2 test was performed on the same data but structured according to Table 2.3. The test found that the major and minor sums did not vary with question order in any vignette, $\chi^2 = 0.67$ ($p = 0.41$, $BF_{01} = 23.33$) for the Mary-Bill

Table 3.1: Response pattern frequencies for each vignette in Conditions 1 and 2. In the header, the first response is whether the participants thought the survivor was credible, and the second is whether they thought the attacker was credible ("*Yes-No*" corresponds to thinking the survivor is credible and the attacker isn't.

Condtion	**Yes-Yes**	**Yes-No**	**No-Yes**	**No-No**
Vignette 1				
Condition 1	167	185	136	12
Condition 2	176	114	192	14
Vignette 2				
Condition 1	152	264	70	14
Condition 2	136	276	64	19
Vignette 3				
Condition 1	130	241	113	15
Condition 2	131	227	118	20

vignette, $\chi^2 = 0.40$ ($p = 0.52$, $BF_{01} = 26.84$) for the Dakota-Alex vignette, and $\chi^2 = 0.23$ ($p = 0.63$, $BF_{01} = 28.13$) for the Frank-Alice vignette. These results can equally be explained by the lack of an order effect and the presence of the QQ equality, and thus, they do not provide evidence for or against H2.

3.2 Statement Placement

Using the data in Table 3.2, the χ^2 test found that participants' responses were significantly different when judgements were made immediately after their corresponding statement (Condition 3) instead of responding after reading both statements (Condition 1). Specifically, more participants said both parties were credible when judging Mary and Bill and when judging Dakota and Alex, compared to the initial condition, which had the same question order but presented both statements first, with $\chi^2 = 13.58$ ($p < 0.01$, $BF_{10} = 2.98$) and $\chi^2 = 154.4$ ($p < 0.001$, $BF_{10} > 100$) re-

Table 3.2: Response pattern frequencies for each vignette in Conditions 1 and 3. In the header, the first response is whether the participants thought the survivor was credible, and the second is whether they thought the attacker was credible ("*Yes-No*" corresponds to thinking the survivor is credible and the attacker isn't.)

Condtion	**Yes-Yes**	**Yes-No**	**No-Yes**	**No-No**
Vignette 1				
Condition 1	167	185	136	12
Condition 3	215	177	97	8
Vignette 2				
Condition 1	152	264	70	14
Condition 3	304	83	94	13
Vignette 3				
Condition 1	130	241	113	15
Condition 3	157	161	148	29

spectively. For Frank and Alice, more participants said that only Frank's statement was credible in the initial condition when they had read both statements before making a judgement as opposed to having questions interspersed between statements with $\chi^2 = 27.80$ ($p < 0.001$, $BF_{10} > 100$).

The χ^2 test was then used on Table 2.4 to test if placing a vignette in between the questions resulted in data which did not follow the QQ equality, in accordance with H3. This test found strong evidence that the major and minor sums changed if a statement was placed inbetween questions as opposed to presenting both statements before the questions, $\chi^2 = 8.53$ ($p < 0.01$, $BF_{01} = 0.03$) for the Mary-Bill vignette, $\chi^2 = 97.00$ ($p < 0.001$, $BF_{01} < 0.003$) for the Dakota-Alex vignette, and $\chi^2 = 8.13$ ($p < 0.01$, $BF_{01} = 0.04$) for the Frank-Alice vignette. These results suggest that the QQ equality did not apply to our data when new information was presented between questions, supporting H3.

3.3 Subsampling

Since question order did not impact participants' responses, a supplemental analysis based on Wang et al. (2014) was performed to test for the presence of the QQ equality in the data (H2). By comparing multiple studies, Wang et al. (2014) showed that the QQ equality impacts the change in response pattern frequencies that result from different question orders (referred to as a context effect). Specifically, they found that the context effect for responding "Yes" to both questions (Yes-Yes context effects) had a very large, negative correlation with the context effect for responding "No" to both questions (No-No context effects) when the QQ equality was present in the data. The same was found for the correlation between the context effect for responding "Yes" and then "No" (Yes-No context effects) and the context effect for responding "No" and then "Yes" (No-Yes context effects). Wang et al. (2014) also found that the above correlations were centred at or near zero. If the QQ equality is present in this study, then similar results should be found for the impact of question order on response frequencies but not for the effect of the arrangement of statements and questions.

To determine if similar correlations appear in the present study, 100 samples of 200 participants were created for each experimental condition using random sampling. The relationships between context effects resulting from question order and the arrangement of questions and statements are shown in Figures 3.1 and 3.2, respectively, with the correlations and means displayed in Table 3.3.

For subsamples of the first two conditions, there were no significant correlations between Yes-Yes and No-No context effects (major context effects) produced by changing the question order, $r(98) = 0.062$ ($p = 0.5398$, $BF_{10} = 0.117$) for subsamples from the Mary-Bill vignette, $r(98) = -0.003$ ($p = 0.9789$, $BF_{10} = 0.1$) for subsamples of the Dakota-Alex vignette, and $r(98) = -0.237$ ($p = 0.01778$, $BF_{10} = 1.87$) for subsamples of the Frank-Alice vignette. The small correlations between Yes-Yes and No-No context effects (major context effects) are likely due to the low frequency of No-No responses in the original data, as this resulted in limited variance in the No-No context effects and thus small correlations. However, there were moderate to large correlations between the Yes-No and No-Yes context effects (minor context effects), $r(98) = -0.432$ ($p < 0.01$, $BF_{10} > 100$) for subsamples from the Mary-Bill vignette, $r(98) = -0.362$ ($p < 0.01$, $BF_{10} > 100$) for subsamples of the Dakota-Alex vignette, and $r(98) = -0.565$ ($p < 0.01$, $BF_{10} > 100$) for subsamples of the Frank-Alice vignette. While these results are not as large as the QQ equality predicts, they do suggest a significant,

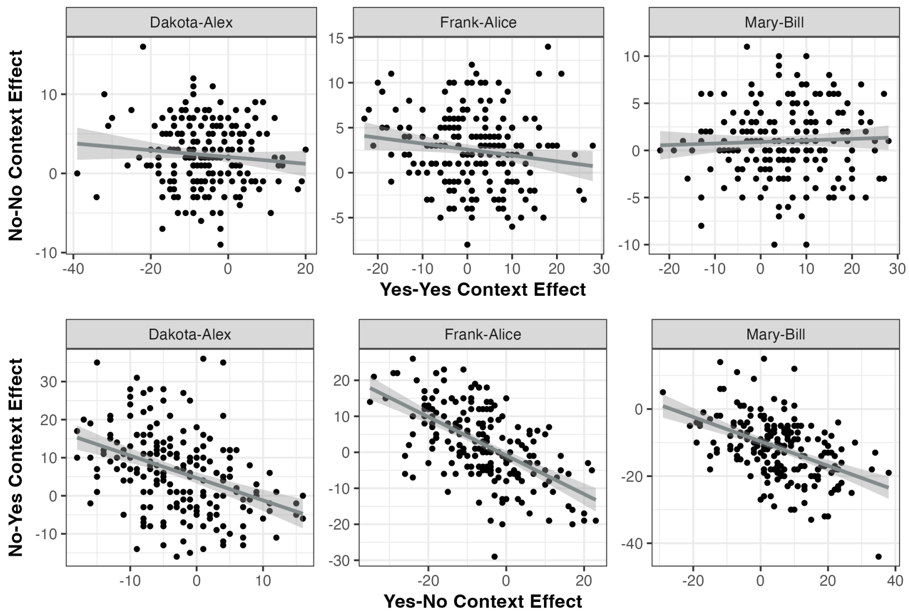

Figure 3.1: Scatter plots of context effects for the subsampled data of Conditions 1 and 2. The top row shows the relationship between the Yes-Yes response context effects and the No-No context effects. The bottom row shows the relationship between the Yes-No context effects and the No-Yes context effects.

negative relationship between context effects produced by changing the question order.

To determine if the subsamples are centred near zero, we calculated the mean context effect for all major and minor pairs to approximate the center point for the correlated data (see Table 3.3 for each pair of means). If the data was centred at the origin, then the origin is expected to be closer to the approximate center than 95% of subsamples. Quantitatively, this means the distance between the origin and the pair of mean subsample context effects should be less than the distance between the paired mean context effects and 95% of the subsample context effects. Since only the minor context effects had significant correlations, the test was only performed on those data. For the Mary-Bill vignette, the distance between the center and the origin was $D_m = 9.426$ and was closer than 64% of subsamples, which had an average distance of $M_d = 12.413$ ($\text{SD}_d = 7.057$). For the Dakota-Alex vignette, the distance from the origin was $D_m = 7.754$, which was closer than 63% of

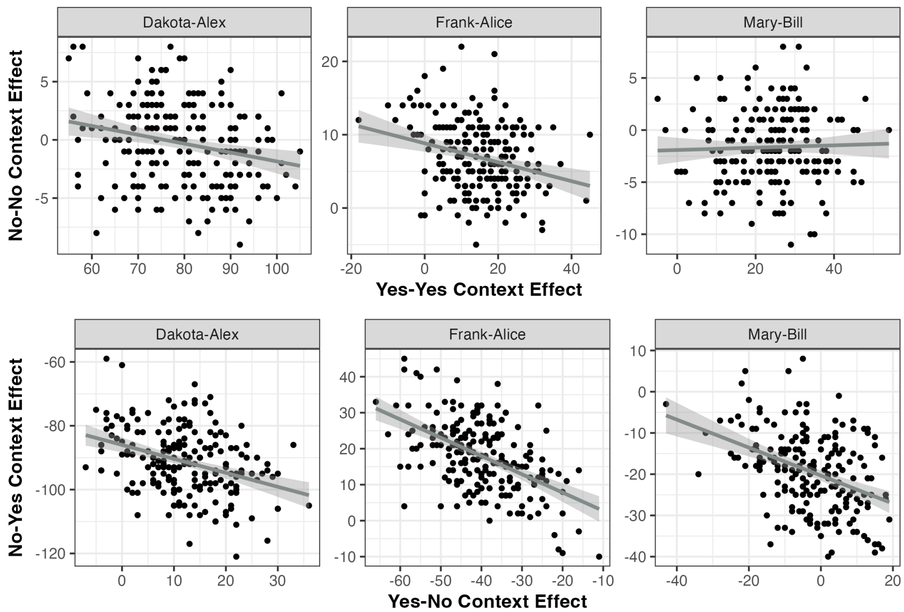

Figure 3.2: Scatter plots of context effects for the subsampled data of Conditions 1 and 3. The top row shows the relationship between the Yes-Yes response context effects and the No-No context effects. The bottom row shows the relationship between the Yes-No context effects and the No-Yes context effects.

samples at an average distance of $M_d = 10.83616$ ($\mathrm{SD}_d = 10.83616$). Finally, for the Frank-Alice vignette, the distance from the origin was $D_m = 6.935$, closer than 76% of samples whose average distance from the center was $M_d = 11.348$ ($\mathrm{SD}_d = 6.102$). These results suggest that the correlations were not significantly centred at zero; however, they also suggest that they are not significantly different than zero. Instead, it indicates that the centers of the correlations were within range of the origin, providing partial support for H2, that QQ equality applies to the first and second conditions.

The context effects produced by interspersing the statements with credibility judgements (Condition 3) compared to making both judgements after reading all statements (Condition 1) have a similar pattern of correlations to the previous set of context effects. Similar to the context effects from changing the order of questions, the correlations between major contexts effects were either small or non-significant, $r(98) = 0.062$ ($p = 0.5324$, $BF_{10} = 0.117$) for subsamples from the Mary-Bill vignette, $r(98) = -0.0328$

Table 3.3: Correlations and means from subsamples of the data. Results significant for $p \leq 0.01$ are bolded, and all correlations have a df = 98. The meaning of "Major," "Minor," and the subscripts for the means retain previously established meanings.

Vignette	**Major**	**Minor**
Question Order		
Mary-Bill	0.062 ($M_{yy} = 20.34$, $M_{nn} = -1.53$)	**-0.432** ($M_{yn} = -4.07$, $M_{ny} = -14.74$)
Dakota-Alex	-0.003 ($M_{yy} = 62.24$, $M_{nn} = -0.42$)	**-0.362** ($M_{yn} = 9.96$, $M_{ny} = -71.78$)
Frank-Alice	-0.237 ($M_{yy} = 11.42$, $M_{nn} = 6.22$)	**-0.565** ($M_{yn} = -32.63$, $M_{ny} = 14.99$)
Statement Placement		
Mary-Bill	-0.0632 ($M_{yy} = 16.24$, $M_{nn} = -2.27$)	**-0.484** ($M_{yn} = -6.11$, $M_{ny} = -7.86$)
Dakota-Alex	-0.0328 ($M_{yy} = 69.08$, $M_{nn} = -2.09$)	**-0.359** ($M_{yn} = -79.20$, $M_{ny} = 12.21$)
Frank-Alice	**-0.259** ($M_{yy} = 11.22$, $M_{nn} = 3.62$)	**-0.540** ($M_{yn} = 11.69$, $M_{ny} = -26.53$)

($p = 0.746$, $BF_{10} = 0.108$) for subsamples of the Dakota-Alex vignette, and $r(98) = -0.259$ ($p < 0.01$, $BF_{10} = 3.66$) for subsamples of the Frank-Alice vignette. Following the results from the previous set of context effects, the correlations between minor context effects were moderate to large, $r(98) = -0.484$ ($p < 0.01$, $BF_{10} > 100$) for subsamples from the Mary-Bill vignette, $r(98) = -0.359$ ($p < 0.01$, $BF_{10} > 100$) for subsamples of the Dakota-Alex vignette, and $r(98) = -0.540$ ($p < 0.01$, $BF_{10} > 100$) for subsamples of the Frank-Alice vignette.

Comparing these data to the origin, we find that the distance from the center for the Mary-Bill vignette is $D_m = 15.29158$ which is closer than 23% of subsamples which had an average distance of $M_d = 11.28911$ ($\text{SD}_d = 6.430167$). For the Dakota-Alex vignette, the distance between the center and origin was $D_m = 72.46772$, which is larger than any other subsample with a mean distance of $M_d = 10.30154$ ($\text{SD}_d = 5.414926$). Finally, for the Frank-Alice vignette, distance from the origin was $D_m = 35.90845$, which was farther than any other subsample considering the mean distance was $M_d = 9.927761$ ($\text{SD}_d = 6.448127$). These results suggest that the origin is farther from the center than a majority of subsamples and significantly different for all but the Mary-Bill vignette. This thus provides further support for the QQ equality not applying to the first and third vignettes, which is in line with H3.

Chapter 4

Discussion

This study tested if a prediction from quantum cognition, the QQ equality, applies to credibility judgements of attackers and survivors in three cases of sexual assault. Contrary to previous studies and H1, question order did not impact participant's credibility judgements. Although this makes testing the QQ equality challenging, a novel subsampling procedure was used to show that the correlations between question orders were consistent with the QQ equality (H2). Additionally, altering the relative placement of statements and questions affected response patterns without producing the QQ equality, as predicted by H3. This sub-sampling procedure also illustrated that the subsample correlations between different arrangements of statements and questions were inconsistent with the QQ equality, providing more support for H3.

Our finding that 75.5% of participants found the survivor credible seems to contradict previous findings that show survivors not typically believed. However, across all vignettes, an average of approximately 40% of respondents thought the survivor's narrative was the sole credible account. This is more in line with the results of Rotenberg (2017) where approximately 50% of cases were consistently dropped at each step of the legal process. Our findings suggest that cases were not necessarily settled because the survivor lacked credibility, but that the combination of the survivor's and alleged attacker's statements produced reasonable doubt if a mistake of fact occurred. This idea would follow from the reasoning presented in Rotenberg (2017).

While changing the order of questions had little impact, responses after reading both perspectives differed from responses produced when reading only one perspective at a time. According to Table 3.2, most participants who read both statements before making any credibility judgments reported that the survivor was credible and the attacker was not. However, participants who evaluated the survivor based only on their statement and subsequently read about and judged the attacker tended to assess both parties as credible. While the mechanism behind this change is beyond the scope of this study, it may be that breaking up the statements and questions works to decouple the people's credibility judgements, in line with the predictions from quantum cognition (Wang & Busemeyer, 2013). Specifically, quantum cognition predicts that putting information between questions means that a person's mental state for the second question is based on the new information presented, not the previous question. As such, judgements about the attacker would primarily be based on the attacker's statement and conclusions about the survivor could only be influenced by the survivor's statement. On a practical level, this implies that the survivor's experience would not contextualize the attacker's statement, thus potentially resulting in more perceived uncertainty about who to believe.

4.1 Quantum Cognition

Outside of their application, these findings also have significant implications for the field of quantum cognition. In opposition to previous results, this study did not find an effect of order on response rates. Prior research suggests that questions that draw on the same pool of information produce an order effect that follows the QQ equality (Pothos & Busemeyer, 2022; Wang & Busemeyer, 2013; Wang et al., 2014). If there is no order effect, then quantum cognition predicts that the questions are unrelated. However, the primary source of information for both credibility judgements was a pair of statements concerning the same incident, which contained a notable amount of redundant information. If the judgements were unrelated, we would not expect correlations between subsamples to follow the QQ equality. To explain these results, we suggest that the size of the order effect was diminished because the feature spaces for each credibility judgement had shared and unique components.

In quantum cognition, all the information and environmental features used in a decision form a feature space (an interpretation of a Hilbert space), which forms the basis for all descriptions of the decision moving forward (Busemeyer et al., 2011; Wang & Busemeyer, 2013). The outcomes of the decision are seen as different mutually exclusive combinations of some subset of features described mathematically as orthogonal subspaces (Bruza et al., 2015; Busemeyer et al., 2011; Pothos & Busemeyer, 2022; Wang & Busemeyer, 2013). After making the decision, a person's mental state moves to the subspace corresponding to that outcome. In this study, Mary's credibility is defined as a feature space based on information from both her and Bill's statements. Within this feature space, there are two subspaces: one corresponding to deciding Mary's statement was credible and one corresponding to deciding her statement was not credible. The features that compose these subspaces correspond to the information that supports the conclusion, such as the parts of Mary's statement that make her appear truthful or the parts of Bill's statement that make the participant suspicious of his narrative. When coming to a conclusion, the participant's mental state changes from an arbitrary combination of features to a combination of features which fall within the subspace for their conclusion.

When two questions share a feature space, the change in mental state resulting from answering a question produces an order effect. When participants decide that Bill is credible after their decision about Mary, their mental states move from the subspace corresponding to thinking that Mary is credible to the subspace for thinking that Bill is credible with a specific probability. However, if participants say Bill is credible before being asked about Mary, they will instead move from a random mental state to the corresponding subspace with a different probability, producing an order effect. Since this study did not find an order effect, quantum cognition would typically describe these questions as existing in different feature spaces. However, subsamples of this study's data showed some evidence of the QQ equality and the response rates changed if participants read both statements first instead of in sequence, suggesting that information from both vignettes is used in the question feature spaces. These results cannot thus be wholly described as the credibility judgments having entirely separate feature spaces nor as fully sharing feature spaces.

Conceptually, two main experimental factors differentiate this study from previous ones similar to Wang et al. (2014): the relationships being evaluated and the amount of information available. With respect to the first factor, a survivor's credibility is likely influenced by some features that do not impact the attacker's credibility, while the question pairs in Wang et al. (2014) are likely influenced by most of the same features. The pairs of questions analyzed in Wang et al. (2014) often focused on either two connected groups (i.e. president/vice president, democrats/republicans, etc.), a pair of related beliefs (i.e. affirmative action for minorities/women, abortion in different contexts, etc.), or a pair of inverse opinions (i.e. was political candidate A unfair towards candidate B versus was B unfair to A). In each case, the two questions are likely to be evaluated on similar features using similar criteria and biases, making them good candidates for sharing a feature space. Alternatively, survivors and attackers are conflicting parties evaluated on different grounds using different biases while still influencing each other (Clarke & Stermac, 2011; Maeder et al., 2015; Pica et al., 2020, 2021; Vandervort, 2005; Wall & Schuller, 2000). For example, survivors are seen as more credible in cases of interracial sexual assault, while alleged attackers are seen as less credible (George & Martínez, 2002). This suggests that the feature spaces for Mary and Bill's credibility will likely be far more distinct than those analyzed in previous research while still sharing some features.

As for the second factor, participants' responses in this study likely had far less to draw on than those analyzed in Wang et al. (2014). The questions in Wang et al. (2014) focused on prominent political figures and the main political issues of the time. This means that the answers would likely be based on a large well of knowledge, opinion, and personal experience upon which to build an ample feature space to answer the question. Conversely, this study asked participants about specific examples of sexual assault and provided vignettes with a limited set of features. While participants may have drawn from their pre-existing notions of sexual assault, there were qualitative differences in each vignette's response pattern, suggesting that the feature spaces contained at least some of the information from the vignettes. This would indicate that participants had a smaller feature space to use in decision-making, potentially impacting their responses.

Traditionally, quantum cognition has assumed that decisions influence each other when they exist wholly in the same feature spaces. Question A only impacts Question B when the answers are based on different arrangements of the same information. This is based on quantum mechanics in which Hilbert spaces describe components of our physical space, making this assumption trivial. However, this assumption becomes non-trivial in psychology because information is more fluid than physical dimensions and thus can be arranged in different manners. For example, when buying a smartphone and a desktop computer, a person may consider the storage size and processing speed for both items but will not likely think about the computer's battery life since that is not applicable to desktop computers. These judgements are not independent of each other, and it would not make sense to describe them using wholly unique feature spaces because there are many related features. However, you could not describe them as sharing a feature space either, since there are features that apply to the phone that are not related to the desktop computer. As such, the simplest solution is for the two to share some components of the feature space but not others.

While the credibility of survivors and attackers appear to be related (Clarke & Stermac, 2011; Lynch et al., 2020; Maeder et al., 2015; Pica et al., 2020; Vandervort, 2005; Wall & Schuller, 2000), these findings suggest that judgments may depend on both shared and unique features. This means the QQ equality would only apply to the components that use features shared by both judgements. Judgements about Mary would change part of the participant's understanding of Bill's credibility, but not all of it. In general, the size of the order effect would now depend on the number of features used in both judgements relative to the number used in a single decision. Thus, order effects are then more likely to be obscured in a small feature space since each feature represents a more significant percentage of the total, and fewer unique features are required to obscure the order effect.

These unique factors mean that each judgement, and thus each response option, has a degree of unique variance, which would diminish the correlations between context effects, as seen in this study. On the level of response options, this hypothesis allows the subspace for agreeing with the survivor, for instance, to be composed of a different number of unique factors than the subspace for agreeing with the attacker. This unequal variance would then suggest different amounts of unique variance in the context effects for the various pairs of responses, thus suggesting that the correlations would deviate from zero, as again seen in this study. In general, if this hypothesis is valid, it could indicate that feature spaces and quantum cognition models can contain more nuance and produce more variable predictions than previously thought.

4.2 Limitations and Future Directions

One of the most significant problems in quantum cognition is predicting when one question will impact the next (Pothos & Busemeyer, 2022). While we can make predictions assuming that two questions will interact, it is difficult to know when those situations occur. This study suggests that it is not a question of whether two decisions interact but how much. If the partial feature space hypothesis is true, then quantum interactions depend on the amount of information used in a decision and what proportion of it is used in the next. This could be tested using a similar experimental design to this study using a fully between-subject design where, instead of having three different vignettes, participants are presented with the same vignette with differing amounts of information and overlap. Even if it is not supported, the methodology of this study may offer insight into the spaces and subspaces involved in decision-making by using vignettes to control the amount of available information.

Outside of quantum cognition, the small scope of this study limited our conclusions about credibility judgements in sexual assault cases. Future studies can systematically vary the biases present in each vignette to test if the placement of information relative to the questions has a significant impact. Because we presented participants with written statements isolated from their broader context and allowed only dichotomous judgements, participants had a limited data set to work with and could not provide a detailed response. Further studies can test if these results hold when participants have more realistic information for their judgement (i.e. police statements, photographs or videos of both parties, witness statements, etc.) and when participants are not limited to a dichotomous decision. Finally, while previous research has shown that decision-making biases for sexual assault cases do not vary with a person's level of expertise, it is unclear whether these results generalize to those communities because the study only sampled from the general population.

4.3 Conclusion

In exploring credibility judgements around sexual assault, this study found initial support for applying the ideas of quantum cognition to the judgements of the survivor and the alleged attacker. While participants' responses were not impacted by question order, there was evidence for the QQ equality in correlations between subsamples of the data. Additionally, there was an effect of statement order on participants' responses, which suggested that hearing both perspectives of the assault before making a judgement could reduce the use of biases. In general, this study functions as a proof of concept of using the tools of quantum cognition to understand the ways sexual assault survivors are perceived. Future research is required for a more robust exploration of this approach and may provide new avenues to test and grow both quantum cognition and credibility research.

Bibliography

Aerts, D., & Gabora, L. (2005). A theory of concepts and their combinations i: The structure of the sets of contexts and properties. *Kybernetes*.

Angelone, D. J., Mitchell, D., & Lucente, L. (2012). Predicting perceptions of date rape: An examination of perpetrator motivation, relationship length, and gender role beliefs. *Journal of interpersonal violence*, *27*(13), 2582–2602.

Angelone, D., Mitchell, D., & Smith, D. (2018). The influence of gender ideology, victim resistance, and spiking a drink on acquaintance rape attributions. *Journal of interpersonal violence*, *33*(20), 3186–3210.

Ask, K., & Landström, S. (2010). Why emotions matter: Expectancy violation and affective response mediate the emotional victim effect. *Law and human behavior*, *34*(5), 392–401.

Bagby, R. M., Parker, J. D., Rector, N. A., & Kalemba, V. (1994). Racial prejudice in the canadian legal system: Juror decisions in a simulated rape trial. *Law and Human Behavior*, *18*(3), 339.

Baker, D. R., Golding, J. M., & Levi, M. M. (2021). Juror perceptions of excited utterance hearsay testimony in an adult sexual assault case. *Psychology, Crime & Law*, 1–17.

Becker, G. S. (1976). *The economic approach to human behavior* (Vol. 803). University of Chicago press.

Berry-Cabán, C. S., Orchowski, L. M., Wimsatt, M., Winstead, T. L., Klaric, J., Prisock, K., Metzger, E., & Kazemi, D. (2020). Perceived and collective norms associated with sexual violence among male soldiers. *Journal of family violence*, *35*(4), 339–347.

Bouffard, J. A. (2000). Predicting type of sexual assault case closure from victim, suspect, and case characteristics. *Journal of criminal justice*, *28*(6), 527–542.

Bruza, P. D., Wang, Z., & Busemeyer, J. R. (2015). Quantum cognition: A new theoretical approach to psychology. *Trends in cognitive sciences*, *19*(7), 383–393.

Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological review*, *118*(2), 193.

Campbell, R., Patterson, D., Bybee, D., & Dworkin, E. R. (2009). Predicting sexual assault prosecution outcomes: The role of medical forensic evidence collected by sexual assault nurse examiners. *Criminal justice and behavior*, *36*(7), 712–727.

Casselman, B., & Tankersley, J. (2019). Harvard suspends roland fryer, star economist, after sexual harassment claims. *The New York Times*. https://www.nytimes.com/2019/07/10/business/economy/roland-fryer-harvard.html

Clarke, A. K., & Stermac, L. (2011). The influence of stereotypical beliefs, participant gender, and survivor weight on sexual assault response. *Journal of Interpersonal Violence*, *26*(11), 2285–2302.

Cotter, A. (2021). Criminal victimization in canada, 2019. *Juristat. Statistics Canada Catalogue*, (85-002).

Dale, M. (2022). Bill cosby, nbc sued by 5 women in new sexual assault case. *CBC*. https://www.cbc.ca/news/entertainment/bill-cosby-sexual-assault-allegations-1.6676564

Davies, M., Rogers, P., & Whitelegg, L. (2009). Effects of victim gender, victim sexual orientation, victim response and respondent gender on judgements of blame in a hypothetical adolescent rape. *Legal and Criminological Psychology*, *14*(2), 331–338.

De Zutter, A., Horselenberg, R., & van Koppen, P. J. (2017). The prevalence of false allegations of rape in the united states from 2006-2010. *Journal of Forensic Psychology*, *2*(2), 1–5.

Donnellan, M., Oswald, F., Baird, B., & Lucas, R. (2006). The mini-ipip scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological assessment*, *18*, 192–203. https://doi.org/10.1037/1040-3590.18.2.192

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191.

George, W. H., & Martínez, L. J. (2002). Victim blaming in rape: Effects of victim and perpetrator race, type of rape, and participant racism. *Psychology of Women Quarterly*, *26*(2), 110–119.

Golding, J. M., Lynch, K. R., & Wasarhaley, N. E. (2016). Impeaching rape victims in criminal court: Does concurrent civil action hurt justice? *Journal of interpersonal violence*, *31*(19), 3129–3149.

Griffiths, D. J., & Schroeter, D. F. (2018). *Introduction to quantum mechanics*. Cambridge university press.

Heisenberg, W. (1983). *The actual content of quantum theoretical kinematics and mechanics* (tech. rep.).

Herlihy, J., & Turner, S. (2015). Untested assumptions: Psychological research and credibility assessment in legal decision-making. *European journal of psychotraumatology*, *6*(1), 27380.

Jaeger, C. B., & Trueblood, J. S. (2018). Thinking quantum: A new perspective on decisionmaking in law. *Fla. St. UL Rev.*, *46*, 733.

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.

Kaufmann, G., Drevland, G. C., Wessel, E., Overskeid, G., & Magnussen, S. (2003). The importance of being earnest: Displayed emotions and witness credibility. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *17*(1), 21–34.

Kelly, L. (2010). The (in) credible words of women: False allegations in european rape research. *Violence Against Women*, *16*(12), 1345–1355.

Kessler, G. (2018). Brett kavanaugh and allegations of sexual misconduct: The complete list. *The Washington Post*. https://www.washingtonpost.com/politics/2018/09/27/brett-kavanaugh-allegations-sexual-misconduct-complete-list/

Krahé, B., Temkin, J., Bieneck, S., & Berger, A. (2008). Prospective lawyers' rape stereotypes and schematic decision making about rape cases. *Psychology, Crime & Law*, *14*(5), 461–479.

Li, J.-A., Dong, D., Wei, Z., Liu, Y., Pan, Y., Nori, F., & Zhang, X. (2020). Quantum reinforcement learning during human decision-making. *Nature human behaviour*, *4*(3), 294–307.

Lynch, K. R., Jewell, J. A., Wasarhaley, N. E., Golding, J. M., & Renzetti, C. M. (2020). Great sexpectations: The impact of participant gender, defendant desirability, and date cost on attributions of a date rape victim and defendant. *Journal of interpersonal violence*, *35*(17-18), 3437–3461.

Maeder, E. M., Yamamoto, S., & Saliba, P. (2015). The influence of defendant race and victim physical attractiveness on juror decision-making in a sexual assault trial. *Psychology, Crime & Law*, *21*(1), 62–79.

McMillan, L. (2018). Police officers' perceptions of false allegations of rape. *Journal of Gender Studies*, *27*(1), 9–21.

Messman-Moore, T. L., & Brown, A. L. (2006). Risk perception, rape, and sexual revictimization: A prospective study of college women. *Psychology of women Quarterly*, *30*(2), 159–172.

Morabito, M. S., Pattavina, A., & Williams, L. M. (2019). It all just piles up: Challenges to victim credibility accumulate to influence sexual assault case processing. *Journal of Interpersonal Violence*, *34*(15), 3151–3170.

Moreau, G. (2020). Police-reported hate crime in canada, 2018. *Juristat: Canadian Centre for Justice Statistics*, 1–31.

Nitschke, F. T., McKimmie, B. M., & Vanman, E. J. (2019). A meta-analysis of the emotional victim effect for female adult rape complainants: Does complainant distress influence credibility? *Psychological Bulletin*, *145*(10), 953.

O'Neal, E. N. (2019). "victim is not credible": The influence of rape culture on police perceptions of sexual assault complainants. *Justice Quarterly*, *36*(1), 127–160.

Peace, K. A., Porter, S., & Almon, D. F. (2012). Sidetracked by emotion: Observers' ability to discriminate genuine and fabricated sexual assault allegations. *Legal and Criminological Psychology*, *17*(2), 322–335.

Peters, J. (2022). Former b.c. mayor charged with sexual assault alleged to have occurred while he was in office. *CBC*. https://www.cbc.ca/news/canada/british-columbia/basran-sexual-assault-kelowna-mayor-1.6678013

Pica, E., Sheahan, C., & Pozzulo, J. (2020). "but he'sa star football player!": How social status influences mock jurors' perceptions in a sexual assault case. *Journal of interpersonal violence*, *35*(19-20), 3963–3985.

Pica, E., Sheahan, C. L., & Pozzulo, J. (2021). Mock jurors' perceptions of sexual assault on a university campus. *Journal of interpersonal violence*, *36*(9-10), NP5447–NP5465.

Pothos, E. M., & Busemeyer, J. R. (2022). Quantum cognition. *Annual review of psychology*, *73*, 749–778.

Rogers, P., Lowe, M., & Boardman, M. (2014). The roles of victim symptomology, victim resistance and respondent gender on perceptions of a hypothetical child sexual abuse case. *Journal of Forensic Practice*, *16*(1), 18–31.

Rotenberg, C. (2017). From arrest to conviction: Court outcomes of police-reported sexual assaults in canada, 2009 to 2014. *Juristat: Canadian Centre for Justice Statistics*, 1–57.

Schuller, R. A., & Stewart, A. (2000). Police responses to sexual assault complaints: The role of perpetrator/complainant intoxication. *Law and human behavior*, *24*(5), 535.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, *69*(1), 99–118.

Sommer, S., Reynolds, J. J., & Kehn, A. (2016). Mock juror perceptions of rape victims: Impact of case characteristics and individual differences. *Journal of interpersonal violence*, *31*(17), 2847–2866.

Strange, D., & Takarangi, M. K. (2015). Memory distortion for traumatic events: The role of mental imagery. *Frontiers in psychiatry*, *6*, 27.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, *90*(4), 293.

Urquhart, M. (2022). Worker alleges she was sexually assaulted by union rep investigating her harassment complaint. *CBC*. https://www.cbc.ca/news/canada/new-brunswick/lawsuit-sexual-assault-union-official-1.6677298

Vandervort, L. (2005). Sexual assault: Availability of the defence of belief in consent. *Can. B. Rev.*, *84*, 89.

Wagenmakers, E.-J. (2022). Approximate objective bayes factors from p-values and sample size: The 3p$\sqrt{n}$ rule.

Wall, A.-M., & Schuller, R. A. (2000). Sexual assault and defendant/victim intoxication: Jurors' perceptions of guilt 1. *Journal of Applied Social Psychology*, *30*(2), 253–274.

Wang, Z., & Busemeyer, J. R. (2013). A quantum question order model supported by empirical tests of an a priori and precise prediction. *Topics in cognitive science*, *5*(4), 689–710.

Wang, Z., Solloway, T., Shiffrin, R. M., & Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proceedings of the National Academy of Sciences*, *111*(26), 9431–9436.

Wasarhaley, N. E., Simcic, T. A., & Golding, J. M. (2012). Mock juror perception of sexual assault nurse examiner testimony. *Violence and Victims*, *27*(4), 500–511.

Winkel, F. W., & Koppelaar, L. (1991). Rape victims' style of self-presentation and secondary victimization by the environment: An experiment. *Journal of Interpersonal Violence*, *6*(1), 29–40.

Yamawaki, N., Riley, C., Rasmussen, C., & Cook, M. (2018). The effects of obesity myths on perceptions of sexual assault victims and perpetrators' credibility. *Journal of interpersonal violence*, *33*(4), 662–685.

Appendices

Appendix A: Vignette and Questions

Vignettes

Mary's Account

"Bill and I attended a party that a mutual acquaintance threw on Friday night. We did not know each other previously, but met that night at the party and talked throughout the evening. I drank quite a bit of alcohol, and Bill offered to give me a ride home. I accepted. When we got to my apartment building, Bill walked me up to my apartment and gave me a kiss goodnight. I invited Bill into my apartment to watch a movie. Bill accepted, so I put on a movie. During the movie we started making out. We undressed and continued making out. I told Bill I did not want to have sex but was enjoying making out with him. Later I felt Bill's penis penetrate my vagina. I told him to stop, but he did not. The next day I went to the hospital and reported the event."

Bill's Account

"I met Mary at a friend's party on Friday night. This was the first time we met, but we really hit it off and spent a lot of time flirting. I noticed that Mary had been drinking so I thought it would be safe and courteous to offer her a ride home. She accepted. When we got to Mary's apartment building, I walked her up to her apartment and kissed her goodnight. Mary invited me in and I accepted. Then Mary put on a movie for us to watch. Instead of watching the movie we were making out. We were both really into it and so we undressed each other and continued making out. Mary even said aloud she was really enjoying hooking up. We ended up having sex. She never said anything to suggest she didn't enjoy it. I didn't even suspect anything was wrong until the next day when the police came to my apartment."

Dakota's Account

"Jordyn and I attended met up for a third date on Saturday night. We had met up for a coffee date and dinner over the past few weeks. This time we went bowling down the street from my house and we were hungry afterward, so we picked up a couple slices of pizza and I invited them up to my apartment to eat since we were nearby. We were chatting on the couch when they kissed me. We were having a good time – I was ok with the kiss, but then Jordyn put their hand on my inner thigh. I pushed their hand away a couple times but they kept doing it. Then they put their hand down my pants and started touching me. They were rough. I didn't like it and I tried to push them off but they held me down. Eventually I gave up and just hoped that it would be over quickly. I kinda pretended I was tired and was falling asleep. After about 20 minutes Jordyn got a phone call from their roommate and had to go because there was some sort of emergency. I was so grateful that the call distracted them. I told my roommate about it when she got home and she told me that I should report it to police, so that's when we came here to give my statement."

Jordyn's Account

"I met Dakota on a dating app about a month ago and we went out on a few dates and had a good time. On Saturday we decided to meet up at a bowling alley and we were having fun. We were both pretty terrible at bowling so it was funny. After a couple hours we were getting hungry so we grabbed a slice of pizza at Mario's Pizzeria and took it to Dakota's apartment to eat since it was just a block or two away. We were hanging out on the couch and started making out. Dakota seemed into it. They were definitely kissing me back. I thought we really liked each other. We were kinda getting into it. Like, we hadn't taken any clothes off yet, but we were both feeling each other up if you know what I mean. I thought we might go further but then my phone rang and wouldn't stop ringing. It turns out it was my roommate. He was calling to tell me that a pipe burst at my apartment and he really needed my help to try to clean up this huge mess. Anyway, I had to leave. I'm just totally shocked that I'm here because I thought Dakota was really cool and into me so I really don't know what happened."

Frank's Account

"Alice and I went to a comedy club – The Laugh Track – on Wednesday night with a bunch of our mutual friends. There were, like 8 of us there, but most people had upcoming exams and stuff so everyone went home early except for Alice and I who went to a dive bar next door. We were there a long time and had both drunk quite a bit. We were feeling pretty tipsy—laughing, having fun. Probably around midnight, I said that we should leave. But I had to go to the bathroom first. I had drank a few beers, you know. I walked to the bathroom at the back of the bar and turned around to find that Alice had followed me in there. I asked her what the hell she was doing—I have a girlfriend and she knows that. I told her to get out of there and that I just liked her as a friend. I kept saying 'no' and 'this isn't a good idea'. But, she didn't leave. She kinda cornered me and started grabbing at my zipper. Next thing I knew she was going down on me. I didn't want to push her off—I don't hit girls, but I told her no and she didn't listen. I didn't really know what else to do."

Alice's Account

"Frank and I and a bunch of friends went to a comedy show on Wednesday. It started at, like, 8pm and was over by 10pm. I was hoping that more people would come out for drinks afterward but a bunch of people had tests and assignments to work on so they didn't want to be out late. It was just Frank and I that went out afterward. We had a few beers, probably three or so each, and were chatting. It was fun, easy. Frank is really funny. We've always had a good time together but this time he seemed kinda flirty. He and his girlfriend had a fight last week and I think that's basically over now. He kinda looked back at me on his way to the bathroom and it seemed like an invitation. I was drunk and feeling pretty bold so I followed him back there. I gave him head and he seemed to like it. He definitely acted that way. I'm a little embarrassed about it now, but it wasn't like a big deal. We both just kinda got carried away."

Questions

Do you believe Mary's Statement?

Yes or *No*

Do you believe Bill's Statement?

Yes or *No*

Appendix B: Task Instructions

Instructions provided to participants before starting the task

Conditions 1 & 2 instructions

On the next page you will find the account of the incident provided from the alleged victim (Mary) and alleged perpetrator (Bill).

Please read each account carefully. You will not be able to proceed to the following page for 60 seconds. On the following page, you will be asked to answer questions about the individuals involved and yourself.

This sequence of events (reading accounts and answering questions) will repeat two more times afterwards, meaning you will read about three incidents in total.

Condition 3 instructions

On the next page you will find the account of the incident provided from the alleged victim (Mary).

Please read each account carefully. You will not be able to proceed to the following page for 30 seconds. On the following page, you will be asked to answer questions about the individual.

You will then be presented with the account of the same incident provided from the alleged attacker (Bill) followed by questions about that individual. Finally you will be presented with questions about yourself.

This sequence of events (reading accounts and answering questions) will repeat two more times afterwards, meaning you will read about three inci-

dents in total.