Spring 2024

Masculine vs. Feminine Women: Verdict, Blame, and Punishment Outcomes in the Legal System

Alaina Helmerichs

MASCULINE VS. FEMININE WOMEN: VERDICT, BLAME, AND PUNISHMENT OUTCOMES IN THE LEGAL SYSTEM

by

ALAINA HELMERICHS

(Under the Direction of Amy Hackney)

ABSTRACT

It is commonly recognized that sentencing disparities exist between men and women who are accused of the same crime—with men often receiving harsher judgements than women. However, research explaining why this disparity exists is incomplete. Research suggests that higher levels of facial masculinity in men is tied to higher levels of criminality (Estrada-Reynolds et al., 2017); however, little research looks at this pattern amongst female defendants. The current study investigated whether the gendered appearance (masculine vs. feminine) of female defendants in different types of cases (child negligence vs. medical malpractice) affects mock jurors' judgments about verdict, punishment, and internal blame attribution. Mock jurors were presented with photographs of the defendant, which were either manipulated to be high in physical characteristics of warmth (feminine) or high in physical characteristics of competence (masculine). Participants then read a trial summary and answered questions about their judgments. The results from Study 1 suggested that with the proper manipulation of gendered appearance, how masculine or feminine the defendant appears could influence mock jurors' trial judgments. However, Study 2 found no effect of the defendant's gendered appearance on participants' trial judgments. The type of case, however, did influence how much the defendant was assigned to pay in damages. The lack of significant findings regarding gendered appearance could be due to inadequate power or because the gendered appearance manipulations were not

salient enough. These findings suggest that masculinity may not be tied to guilt and harsher judgments amongst female defendants in the same way that it is with male defendants.

INDEX WORDS: Jury decision-making, Gender bias, Defendant gender

MASCULINE VS. FEMININE WOMEN: VERDICT, BLAME, AND PUNISHMENT OUTCOMES IN THE LEGAL SYSTEM

by

ALAINA HELMERICHS

B.S., North Central College, 2022

A Thesis Submitted to the Graduate Faculty of Georgia Southern University

in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

COLLEGE OF BEHAVIORAL AND SOCIAL SCIENCES

MASCULINE VS. FEMININE WOMEN: VERDICT, BLAME, AND PUNISHMENT OUTCOMES IN THE LEGAL SYSTEM

by

ALAINA HELMERICHS

Major Professor: Amy Hackney

Committee: Michael Nielsen

C. Thresa Yancey

Brenda Sims Blackwell

Electronic Version Approved:

May 2024

ACKNOWLEDGMENTS

I would like to thank everyone who played a role in completing this research. First and foremost, I would like to thank Dr. Amy Hackney for all her time, effort, and expertise—I am a better researcher because of her. Thank you to Dr. Mike Nielsen, Dr. Thresa Yancey, and Dr. Brenda Sims Blackwell for all your time and expertise serving on my committee. Thank you to my research assistants, Lindsay Eggen, Natalie Leandre, and Audrey Molter. I truly appreciate all the time you have spent helping me finish this project, and I have loved working with each of you. Finally, thank you to my partner, John, and all of our cats for all of your support.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Existing inequalities within the legal system have continued to come to light both within psychological and criminological research as well as in mass media. Specifically, sentencing disparities between men and women across the United States have long been recognized and favor female defendants. For example, a study conducted on data from Arizona found that male offenders were twice as likely to be incarcerated than female offenders, and women were overall more likely than men to receive probation than any jail time (Johnston et al., 1987). More recent research finds that men are 74% more likely to be incarcerated than women for the same crime (Doerner & Demuth, 2014).

However, there is a lack of clarity on what exactly caused this sentencing disparity between men and women in the first place and why the incarceration rate for women has been increasing. In order to provide a possible answer for these things, it is important to first determine what underlying biases may be at play when juries are determining their verdicts in these cases. Thus, the current research investigated gendered appearance (masculinity vs. femininity) as one possible factor to explain why men and women receive vastly different outcomes for committing the same crimes.

Stereotype Content Model

The stereotype content model serves as an explanation for how people may make judgments involving other people with regard to how warm and competent they are perceived as being (Fiske et al., 2002). “Warmth” refers to how kind and friendly someone appears while “competent” refers to how capable or knowledgeable someone appears. Fiske and colleagues argue that everyone falls somewhere on the dimensions of both competence and warmth. People

with high competence tend to be perceived as being of a higher status and are therefore judged more positively than people with lower competence. Those who are perceived as being high in both competence and warmth are often admired (e.g., allies). On the other hand, people feel contempt towards those who are perceived to be low in both competence and warmth (e.g., poor people). There are instances in which people may be high in one characteristic and low in the other. When people are high in warmth but lack competence, they are pitied (e.g., housewives). Yet when people are high in competence but lack warmth, they are envied (e.g., rich people).

It is argued that the stereotype content model can explain the biases and prejudices existing toward various groups of people. For example, there are smart-but-cold and warm-but-dumb stereotypes that exist about the different genders (Fiske, 2012). Fiske argues that men have a smart-but-cold stereotype because they are seen as capable but not very kind, while women have a warm-but-dumb stereotype because they are seen as kind but not very capable. Research seems to back up this claim. When participants are asked to categorize photos of both men and women as intelligent, capable, warm, or friendly, they are more likely to rate men as intelligent and capable and women as warm and friendly (Wen et al., 2020). Similarly, people rate female surgeons as more warm and less competent while also rating male surgeons as more competent and less warm (Ashton-James et al., 2019). Interestingly, when there was no gender assigned to the surgeons, they were rated as high in both competence and warmth, suggesting that the associated gender impacts perceptions.

Gender vs. Gendered Appearance

Because it is known that jurors may be consciously or unconsciously taking the defendant's characteristics into consideration when determining a verdict, it is plausible that jurors' focus on defendant characteristics contributes to the gender disparity in sentencing. While

this sentencing disparity is currently discussed in the literature and public discourse in a way that suggests someone's gender is to blame for this inequality, it is also possible that how masculine or feminine someone looks is responsible for this disparity, rather than their gender per se.

There is currently a lack of research on how a defendant's gendered appearance (masculine vs. feminine) can affect their outcome in a trial, however, some available research supports this idea. When mock jurors read descriptions of a defendant in a murder case, male defendants were much more likely to be found guilty of murder than female defendants (Strub & McKimmie, 2016). However, when female defendants were described in a stereotypically masculine way (e.g., "strong" & "courageous"), they were more likely to be found guilty than when described in a stereotypically feminine way (e.g., "gentle" & "caring"). Similarly, when participants are presented with an assault case and faced with a line-up scenario made up of men, they pick out the men with the most masculine appearances to be the criminal (Estrada-Reynolds et al., 2017), suggesting that masculinity in men could be a cue to criminality.

In another mock jury study, this time involving a duress case, heterosexual men were most likely to be found guilty, followed by gay men, lesbian women, and finally heterosexual women (Russell et al., 2012). This pattern of findings may be explained by other research, suggesting people perceive gay men as less masculine and more feminine than straight men while also perceiving lesbian women as more masculine and less feminine than straight women (Blashill & Powlishta, 2009). In other words, as perceived femininity increased and perceived masculinity decreased across the target categories of gender and sexual orientation, perceptions of criminality decreased.

Furthermore, outside of a mock jury or mock line-up scenario, when participants are asked to pick which faces have a criminal appearance and which faces have a remorseful

appearance, the more masculine faces are seen as criminal (Funk et al., 2017). Specifically, in Funk et al.'s study, "masculine faces" were faces with a more prominent chin, small eyes, a low brow, and darker pigmentation, and these faces were perceived as more criminal.

While it could be the gender of a person creating the current sentencing disparity, previous research demonstrates that masculinity is associated with criminality. Thus, when people see a masculine appearing person, they may perceive that person as being guiltier than a feminine appearing person, regardless of their gender. Although there is some existing research supporting this claim, previous studies mainly focus on male defendants. Research that focuses on female defendants manipulates masculinity and femininity through written descriptions. However, there is a lack of research manipulating the gendered appearance of female defendants through their actual physical appearance. Thus, one purpose of the current research was to analyze whether this pattern persists with facial masculinity and femininity of female defendants, which would represent a more realistic scenario for how jurors judge defendants.

Stereotype Incongruency

Stereotype incongruency is a phenomenon in which a person's actions are inconsistent with the stereotype that someone holds about them, which often leads to some sort of bias or discrimination. Applying stereotype incongruency to stereotypes about gendered appearance, a woman would be acting in a stereotypically incongruent way if she did something "masculine," such as repairing a car. Similarly, a man would be acting in a stereotypically incongruent way if he did something "feminine," such as taking care of a sick loved one. When presented with stereotypically incongruent actions happening from both men and women, participants demonstrated more biased processing for female stereotype-inconsistency versus male stereotype-inconsistency (Sekaquaptewa & Espinoza, 2004). This means that people are more

judgmental when women act in stereotypically masculine ways than when men act in stereotypically feminine ways.

The difference in how participants judge the male versus female targets is believed to be due to the status of the target. Lower status individuals (e.g., women) have more constraints on what is expected of them and typically have more stereotypes attached to them compared to higher status groups (Glick et al., 1995). Higher status individuals (e.g., men) are thought to be judged less overall in these scenarios due to society expecting them to be the ones who set the social norms. For example, Fiske (1998) discusses how people automatically think of a White man when someone says "person," thus exemplifying that White men are the "default."

Current Investigation and Hypotheses

Since prior research suggests the possibility of a defendant's gendered appearance affecting how jurors judge them, the current studies manipulate the gendered appearance (masculine vs. feminine) of female defendants along with the gendered stereotype of the type of civil case (medical malpractice vs. child negligence). Civil cases were used in the current research due to a lack of research in this area. The purpose of this research was to determine if a female defendant's gendered appearance affects juror judgments and whether stereotype incongruency of the civil case type impacts juror judgements. Based on previous findings that masculine features on men signal criminality to other people (Estrada-Reynolds et al., 2017), I hypothesized that participants would be more confident in a guilty verdict, assign more damages for the defendant to pay, and assign more internal blame to the more masculine-appearing female defendants compared to the more feminine-appearing female defendants. Drawing on the stereotype incongruency perspective, I hypothesized that participants would find the defendant guiltier when the gendered stereotype of the case violates the gendered stereotype of the

defendant. This means that harsher judgments were expected across all dependent measures for the feminine woman accused of child neglect and when a masculine woman is accused of medical malpractice. However, because Sekaquaptewa and Espinoza (2004) concluded women are judged more harshly when acting in stereotypically masculine ways, I expected that this pattern would be more extreme for the masculine-appearing woman accused of medical malpractice.

Study One was developed as a pilot study to test the construct validity of the three dependent measures as well as the facial stimuli for the defendants. Specifically, this study explored whether the mock jurors of the intended population actually interpret the more masculine faces as more masculine and the more feminine faces as more feminine. While the faces were pre-rated on masculinity and femininity by Walker et al. (2018), this was in a different country, meaning that the faces may not be interpreted the same way in the United States. I present study one first, followed by study two, which tested my hypotheses.

CHAPTER 2

STUDY 1

Method

Participants

Participants in this study were 27 undergraduate students from a large Southeastern university recruited via the university's participant recruitment system and flyers posted around campus. Those who signed up through the university's recruitment system were all psychology students and received class credit for their research participation. Those who signed up by responding to flyers received a $17 e-gift card. In order to be eligible to serve on a real jury in the United States, people must be at least 18 years old and a citizen of the United States. Thus, all participants were verified to be citizens of the United States and at least 18 years old in order to reflect a jury-eligible population. The participants' ages ranged from 18 to 25 years old (M = 20.12, SD = 1.79). Approximately 50.0% of participants were White, and 57.7% of participants identified as men. On average, participants listed their political beliefs as being neutral. It is important to note that one participant was excluded from data analysis due to answering an attention check incorrectly. See Table 1 for more information regarding participant demographics.

Table 1

Demographic Characteristics of the Participants in Study 1

Participants (N = 26)	*n*	%
Sex		
Male	14	53.8
Female	11	42.3
Prefer not to say	1	3.8
Gender		
Man	15	57.7
Woman	10	38.5
Non-binary	1	3.8
Transgender	1	3.8
Race/Ethnicity		
American Indian or Alaskan Native	1	3.8
Black or African American	10	38.5
White	13	50.0
Hispanic	3	11.5
Prefer not to say	1	3.8
Political beliefs		
Very conservative	1	3.8
Conservative	2	7.7
Somewhat conservative	5	19.2
Neutral	9	34.6
Somewhat liberal	2	7.7
Liberal	6	23.1
Very liberal	1	3.8

Note. One participant chose both "Man" and "Transgender" for their gender identity.

Materials

Trial Stimuli

Participants were randomly assigned to one of two possible trial summaries: a medical malpractice case or a child negligence case. Both cases had the same amount of evidence presented and were nearly the same length. The medical malpractice case was based on Lindquist v. Dengel (1979), and created to represent a lack of competence. The child negligence case was based on Williams v. State (2007), and created to represent a lack of warmth. Fiske (2012) demonstrated that women are expected to be high in warmth and lack competence while men are expected to be low in warmth and high in competence. Thus, the medical malpractice case violates masculine gender stereotypes while the child negligence case violates feminine gender stereotypes.

In Lindquist v. Dengel (1979), the real medical malpractice case, a patient went to their general practitioner for symptoms of bronchitis. The general practitioner took x-rays, which showed signs of tuberculosis. The doctor then suggested to the patient that a saliva test and additional x-rays be done to confirm this diagnosis. The patient claimed that nothing was done to actually do any of the follow-up tests, but the general practitioner says otherwise. The doctor claims a container was given to the patient for the saliva test, and that the patient was told to schedule additional x-rays. Over a year later, the patient ended up hospitalized for tuberculosis. During the hospitalization, a different doctor performed a partial lung removal that left the patient unable to perform their job any longer. The patient claimed that this could have been prevented if the general practitioner had performed those confirmatory tests when they were first suggested. The jury in this case found the doctor and patient both 50% responsible.

Some elements were changed for the current research to make the case facts more ambiguous to the participants in this study. The edited version presented to study participants

depicted all of the doctors as women to ensure that gender stereotypes were not at play when envisioning a doctor. Additionally, the case presented to the mock jurors included testimony from a receptionist from the doctor's office stating that she called the patient to schedule an appointment for the x-rays. This testimony further noted that the doctor cancelled this appointment, and after this, the doctor's office was unable to reach the patient to reschedule the x-rays. Furthermore, in the version presented to the mock jurors, an expert witness, who was a pulmonologist, testified saying that she did not think the partial lung removal was necessary and she would have tried drug therapy first.

In Williams v. State (2007), the real child negligence case, a mother took her two children from their grandmother's house (with working utilities) to her boyfriend's duplex (which lacked working utilities) for the night while she went out with her friends. Since the utilities were not working, they lit a candle for light. The two children and the boyfriend eventually fell asleep, but the boyfriend woke up to the children screaming. The candle had caused a fire, and the children ended up passing away in the fire.

Again, elements of this case description were altered to keep the case more neutral and to more closely match the severity level presented in the medical malpractice case. Specifically, in the case presented to participants, the mother, the defendant, was a nurse working the night shift instead of going out with friends. Additionally, the children did not die, but instead played with the flame of a candle and ended up with severe burns as a result. In this presentation, the grandmother served as a witness who claimed that she warned the mother that it could be dangerous to have the children spend the night somewhere without utilities, and further indicated she told her daughter she would be happy to keep the children with her until the morning.

Face Stimuli

Participants were presented with a photo of a defendant. Since prior research shows that the masculinity and femininity of faces can be manipulated through changing the levels of warmth and competence facial characteristics (Walker et al., 2018; Wen et al., 2020), each photo was manipulated to be high in warmth (more feminine-presenting) or high in competence (more masculine-presenting). These photos were taken from the Basal Face Database, where they were previously validated on these measures of warmth and competence (Walker et al., 2018). Participants saw a photo of one of three possible women. However, there were two versions of each woman's face: a masculine version and a feminine version, resulting in six possible photos that participants could see. Each participant was randomly assigned to one of the six photos.

Confidence in Guilt

Confidence in guilt was measured using a direct question of whether the defendant was guilty or not guilty, followed by asking mock jurors how confident they were in the verdict they chose, ranging from 0% to 100%. A confidence in guilt variable was then created by combining the answers to these two questions. A guilty verdict was re-coded as a "1," and a not guilty verdict was re-coded as a "-1." Then, this number was multiplied by the percentage that each participant chose for how confident they were in their verdict. Thus, the scoring on this variable ranged from -100% confident in a guilty verdict to 100% confident in a guilty verdict. Thus, if a participant said the defendant was guilty and then said they were 75% confident in that verdict, they would have a re-coded score of 75% confident in guilt. Likewise, if a participant said the defendant was not guilty and then said they were 75% confident in that verdict, they would have a re-coded score of -75% confident in guilt.

Punishment

If a mock juror chose a guilty verdict, they were asked to assign punishment to the defendant through damages (e.g., *In this case, how much should the defendant, if guilty, pay in damages to the plaintiff?*) with answers on a sliding scale ranging from $0 to $10,000. This range was determined by the range that would be applicable in the real world during these types of court cases.

Internal and External Blame Attributions

Internal blame attribution was measured using 12 items based on questions from the Revised Gudjonsson Blame Attribution Inventory (Gudjonsson & Singh, 1989) (e.g., *The defendant was in full control of what happened*). All items were measured on Likert scales ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*). Items 1, 3, 4, 6, and 7 were reverse scored and scores were averaged for the overall measure of internal blame attribution. The original scales include both mental-element attributions and guilt-feeling attributions. For the mental-element attributions, the Cronbach's alpha is .79, and for the guilt-feeling attributions, the Cronbach's alpha is .81 (Cima et al., 2006). In the current study, a mixture of both mental-element and guilt-feeling attributions were used for the internal blame scale, and the Cronbach's alpha for this scale was calculated to be .49.

External blame attribution was measured using 12 questions based on the Revised Gudjonsson Blame Attribution Inventory (e.g., *In this case, someone besides the defendant was largely to blame for what happened*) all measured on Likert scales ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*). Items 1, 2, 4, 9, and 12 were reverse scored and the scores were averaged for the overall measure of external blame attribution. The original scale includes more than 12 items, and past research has calculated the Cronbach's alpha for the full scale to be .77

(Cima et al., 2006). In the current study, Cronbach's alpha for the external blame measures was calculated to be .79.

Since the Revised Gudjonsson Blame Attribution Inventory is a self-report tool that forensic psychologists use with inmates, some items on the inventory were not applicable to a mock-jury situation. The items that were applicable were reworded to be from the point of view of a juror judging a defendant rather than being a self-report measure.

Attention Checks, Manipulation Checks, and Demographics

Two multiple-choice questions were included to make sure participants were paying attention. These questions asked what type of case they read and the gender of the defendant. An additional multiple-choice question was included to ensure that the manipulation of the defendant's gendered appearance was successful (e.g., *How masculine or feminine does the defendant look?*), which had answers ranging from 1 (*Very masculine*) to 7 (*Very feminine*). This question was presented alongside the photo of the defendant. These questions were followed by demographic questions about the participants (e.g., sex, gender, age, race, and political beliefs).

Procedure

The study took place in a private psychology lab with a maximum of 3 participants in the lab completing the study at once. When entering the lab, participants silenced their cell phones and placed their belongings somewhere where they would not be distracted. After the researcher explained their role in the study, participants opened a Qualtrics link and read an informed consent form. After agreeing to participate in the study, participants viewed one of six randomly assigned photos of the defendant along with one of two randomly assigned types of cases followed by a transcription of the trial. Then, participants answered questionnaires regarding their verdict, confidence in their verdict, internal and external blame attributions, and if they

chose a guilty verdict, participants assigned an amount of money for the defendant to pay in damages. The end of the questionnaire included attention and manipulation checks along with demographic questions. After the questionnaire was completed, participants were presented with a list of resources they could use, if needed, and the researcher debriefed them on the purpose of the study while providing an opportunity for them to ask questions.

Results

Effects on Verdict

In order to make sure the type of case was not altering the verdict (guilty vs. not guilty) chosen by the participants, a crosstabs was conducted to see the frequency of each verdict across the types of cases (medical malpractice vs. child negligence). The crosstabs showed that the verdicts were even across cases with six not guilty verdicts and seven guilty verdicts in both the medical malpractice and child negligence trials. Another crosstabs was conducted to determine if the frequency of each verdict was approximately the same or different for each type of gendered appearance (masculine vs. feminine). While there did not appear to be much difference, descriptive statistics showed slightly more guilty verdicts for the masculine appearing defendants (eight) than the feminine appearing defendants (six).

Facial Manipulation Interpretations

Because the main purpose of this study was to determine if the different facial manipulations of the three defendants were interpreted in the way they were meant to be, a *t*-test was conducted to determine if there was a significant difference in how masculine or feminine the participants rated the feminine photos versus the masculine photos. Participants did not rate the feminine appearing defendants ($M = 5.93$, $SD = .62$) as appearing significantly more

feminine differently from the masculine appearing defendants (M = 5.25, SD = 1.14), $t(24)$ = 1.93, p = 0.07, Cohen's d = 0.74

Since the feminine and masculine photo manipulations did not seem to be interpreted significantly differently, a *t*-test was conducted for the masculine and feminine version of each of the three women's faces. This was done to see if any of the women were overall being rated much differently than the others.

Face 1

While there was no significant difference between the feminine and masculine versions of Face 1, the feminine photos were rated as being between "feminine" and "highly feminine" (M = 6.25, SD = .50) while the masculine photos were rated as being between "slightly feminine and feminine" (M = 5.20, SD = 1.30), $t(7)$ = 1.51, p = 0.18, Cohen's d = 1.06.

Face 2

For Face 2, there was virtually no difference in how participants rated the feminine photos (M = 5.80, SD = .83) and the masculine photos (M = 5.75, SD = .50), $t(7)$ = .105, p = 0.92, Cohen's d = 0.07. The different photos were both rated as being between "somewhat feminine" and "feminine," suggesting that the manipulation did not work properly with these photos.

Face 3

For Face 3, there was again no significant difference in the ratings of the different versions, but participants did rate the feminine photos (M = 5.80, SD = .44) as being between "somewhat feminine" and "feminine" while the masculine photos (M = 4.67, SD = 1.53) were rated as being between "neither masculine nor feminine" and "slightly feminine," $t(6)$ = 1.63, p = 0.16, Cohen's d = 1.01. Similar to Face 1, this suggests that participants did notice some

gendered difference between the different versions of Face 3. Since Face 2 was rated quite differently than the other faces, Face 2 was dropped from study two in an attempt to ensure the faces are properly manipulating the femininity and masculinity of the defendants.

Punishment

Descriptive statistics were conducted to determine if the scale for measuring damages should be edited. Participants were only asked to assign punishment (in the form of damages) if they chose a guilty verdict. The minimum amount of damages that participants assigned for the defendant to pay was $0 with the maximum amount assigned being $10,000. There were no significant main or interaction effects found regarding the amount of damages mock jurors assigned the defendant to pay (See Table 2). The masculine defendants (*M* = $5, 214.29, *SD* = $3, 533.85) were assigned a similar amount of damages to pay as the feminine defendants (*M* = $3, 037.20, *SD* = $2, 699.54). Additionally, a similar amount of damages were assigned in the medical malpractice case (*M* = $6,007.17, *SD* = $2,901.29) compared to the child negligence case (*M* = $2,607.17, *SD* = $2868.28). There was no sign of either a ceiling or floor effect, as participants used both endpoints of the scale as well as a healthy amount of numbers above and below the midpoint.

Table 2

ANOVA Results for Punishment Across Conditions

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	188948376.96	1	188948376.96	20.63	.002	.721
Defendant Gendered Appearance	8108082.84	1	8108082.84	.885	.374	.100
Case Type	24840424.49	1	24840424.49	2.71	.138	.253
Defendant Gendered Appearance x Case Type	2332600.96	1	2332600.96	.255	.627	.031
Error	73259157.17	8	9157394.65			

Internal and External Blame Attributions

Although the internal and external blame attribution measures were based on an existing questionnaire, not all of the statements were included from this questionnaire and all statements were reworded to fit a jury decision-making situation. Thus, Cronbach's alpha was calculated for the internal blame measures ($\alpha = 0.49$) as well as the external blame measures ($\alpha = 0.79$). While the internal reliability of the internal blame measures is fairly low, this could be affected by the small sample size. If it continues to be low in the final study, items 1 or 4 may be dropped from analysis to increase internal reliability.

Discussion

The results from Study 1 indicate that people may in fact see a more masculine appearing defendant and see that as a sign of guilt. While there was not a significant difference between masculine and feminine defendants and rates of a guilty verdict, there were substantively more

guilty verdicts overall for the masculine-appearing defendants compared to the feminine-appearing defendants. This aligns with past research, suggesting that the more masculine a person looks, the more criminality is cued in a person's brain (Estrada-Reynolds et al., 2017). These findings also align with Strub and McKimmie's (2016) conclusion that female defendants described as being more masculine are more likely to be found guilty compared to female defendants described as feminine. The current study's results indicate that it is possible people pick up on masculine characteristics in a defendant and unconsciously focus on these characteristics when determining trial verdicts, specifically yielding harsher outcomes.

The current study's results further suggest that the proper manipulation of facial masculinity (or femininity) can significantly impact how mock jurors judge the defendant in a trial. Future research can further confirm this finding by replication without the Face 2 photos. Participants did not distinguish a difference in the gendered appearance between the masculine and feminine versions of the photo, indicating that the photo manipulations themselves were not effective. Thus, Study 2 sought to solve this issue by only using the photos of Face 1 and Face 3 for further investigation.

CHAPTER 3
STUDY 2
Method

The methodology and data analysis plan for this study was pre-registered. The pre-registration document can be found here: https://aspredicted.org/ki32n.pdf.

Participants

According to a power analysis conducted through G*Power, at least 128 participants were needed in order to have sufficient power (Cohen's $f = .25$; $\alpha = .05$; $1 - \beta \geq .80$). Participants for this study were 128 undergraduate students at a large Southeastern university who completed this study for either class credit, extra credit, or a $20 Amazon gift card. Participants were recruited using the university's participant recruitment system as well as posting flyers around campus and having professors advertise the study to their classes. One notable difference between the participants in Study 1 and Study 2 is that Study 1 participants were all psychology students whereas Study 2 participants were from any major. All participants fit the jury-eligible criteria of being at least 18 years old and a U.S. citizen. Of the 128 participants, seven were excluded from data analysis due to either completing the study too quickly ($n = 5$) or failing an attention check ($n = 2$). The participants' ages ranged from 18 to 53 ($M = 20.53$, $SD = 3.55$). Approximately 66.1% of participants identified as White, and 65.3% of participants identified as a woman. Most participants identified their political beliefs as neutral or some level of conservative identification. Further information about the demographics of participants can be found in Table 3.

Table 3

Demographic Characteristics of the Participants in Study 2

Participants (N = 121)	*n*	%
Sex		
Male	37	30.6
Female	83	68.6
Gender		
Man	38	31.4
Woman	79	65.3
Non-binary	3	2.5
Genderfluid	1	0.8
Race/Ethnicity		
American Indian or Alaskan Native	2	1.7
Asian or Asian American	3	2.5
Black or African American	36	29.8
White	80	66.1
Hispanic	9	7.4
Another race	1	0.8
Prefer not to say	2	1.7
Political beliefs		
Very conservative	5	4.1
Conservative	22	18.2
Somewhat conservative	21	17.4
Neutral	40	33.1
Somewhat liberal	19	15.7
Liberal	10	8.3
Very liberal	4	3.3

Materials

Trial Stimuli

Participants were randomly assigned to either read the trial transcription of the medical malpractice case or the child negligence case. Because the results from Study 1 did not show that the type of case had an effect on the dependent measures, these cases were the same cases presented in Study 1, which were adapted from Lindquist v. Dengel (1979) and Williams v. State (2007), respectively. Recall that the medical malpractice case was meant to violate masculine gender stereotypes by representing a lack of competence while the child negligence case was meant to violate feminine gender stereotypes by representing a lack of warmth.

Face Stimuli

Because the results from Study 1 suggested that the masculine and feminine versions of Face 2 were not perceived in the same way as Face 1 or Face 3, the Face 2 photos were not used in Study 2. Participants were presented with one of four randomly assigned photos meant to represent the defendant. The photo each participant received was either the masculine or feminine version of either Face 1 or Face 3, which were taken from the Basal Face Database (Walker et al., 2018) and manipulated to represent either heightened facial characteristics of warmth (more feminine) or heightened facial characteristics of competence (more masculine).

Confidence in Guilt

Confidence in guilt was measured the same way as in Study 1. This involved combining the direct question of whether the participant thinks the defendant is guilty (re-coded as “1”) or not guilty (re-coded as “-1”) and multiplying that by how confident the participant is in their verdict (ranging from 0% to 100%). After creating the “confidence in guilt” variable, the

potential responses ranged from -100% confident in a guilty verdict to 100% confident in a guilty verdict.

Punishment

Due to Study 1 showing that participants used the full length of the scale when asked how much the defendant should pay in damages, punishment was measured in the same way with answers on a slider scale ranging from $0 to $10,000.

Two additional questions were included in Study 2. First, participants were presented with an open-ended question asking them to explain why they chose the specific amount of damages that they did. Participants were then presented with another open-ended question asking them who they believe will receive the money.

Internal and External Blame Attributions

While there was low internal reliability amongst the internal blame attribution items retrieved from the Revised Gudjonsson Blame Attribution Inventory (Gudjonsson & Singh, 1989) in Study 1, the same 12 internal blame attributions were included in this study. As with Study 1, items 1 and 4 were problematic to the internal reliability of the scale. Thus, both items were dropped. Item 11 was also dropped due to a low internal reliability of the scale with this item included ($\alpha = 0.57$). After dropping items 1, 4, and 11 in the current study, there was a sufficient Cronbach's alpha ($\alpha = 0.73$). These attributions were scored in the same way as Study 1. The internal reliability for these attributions was calculated for each condition: (a) the child negligence case with a masculine-appearing defendant ($\alpha = 0.66$), (b) the child negligence case with a feminine-appearing defendant ($\alpha = 0.82$), (c) the medical malpractice case with a masculine-appearing defendant ($\alpha = 0.66$), and (d) the medical malpractice case with a feminine-appearing defendant ($\alpha = 0.78$).

The internal validity of the external blame attributions in Study 1 was high, so the same questions were used (α = 0.78). These attributions were also scored in the same way as Study 1. The internal reliability for these attributions was calculated for each condition: (a) the child negligence case with a masculine-appearing defendant (α = 0.64), (b) the child negligence case with a feminine-appearing defendant (α = 0.71), (c) the medical malpractice case with a masculine-appearing defendant (α = 0.88), and (d) the medical malpractice case with a feminine-appearing defendant (α = 0.79).

Reasoning for Verdict

An additional short response question was included where participants relayed why they chose a guilty or not guilty verdict. This allowed mock jurors to explain what they based their verdict on to provide more insight into their decision-making process.

Defendant Characteristics

Participants were shown the defendant's photo again while being asked to explain, in years, how old they perceived her to be. The feminine version of Face 1 had the youngest average age rating (*M* = 26.73), followed by the masculine version of Face 1 (*M* = 27.16), the masculine version of Face 3 (*M* = 28.19), and then the feminine version of Face 3 (*M* = 29.13). Additionally, participants were asked to rate how competent, skilled, warm, friendly, and conscientious the defendant appeared to be ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*).

Attention Checks, Manipulation Checks, and Demographics

The same two multiple-choice attention check questions asking the participants what type of case they saw and what the gender of the defendant was were included in Study 2. Participants were asked to rate how masculine or feminine the defendant appears to be in order to make sure

the photos were actually manipulating their gendered appearance. Demographic questions asking the participants' sex, gender, age, race/ethnicity, and political beliefs were included at the end of the questionnaire.

Procedure

Participants entered the lab either alone or with up to two other participants. After the researcher provided them with general information about what they would be doing, the participants began the study on Qualtrics on the lab computers. After reading the informed consent and choosing to participate, participants then saw one of four randomly assigned photos of the defendant that was presented alongside the type of case (either medical malpractice or child negligence), which was also randomly assigned. Mock jurors then read through a trial transcription of the evidence presented in whichever type of case they were randomly assigned. After the trial transcription was presented, participants completed a questionnaire asking for their chosen verdict, how much money the defendant should pay in damages (if guilty), how confident they were in their verdict, internal blame attributions, and external blame attributions. When they completed the dependent measures, participants answered attention checks, manipulation checks, and demographic questions while also explaining why they chose the verdict they chose. Before finishing, a separate survey was completed where participants were able to identify the type of incentive they signed up to receive. A short, written debriefing was then presented to them followed by the researcher providing further information on the study's aims.

Results

Preliminary Analyses

Bivariate correlations were conducted across all of the dependent measures as well as external blame and the participants' ratings of the defendant's characteristics (competent, skilled, warm, friendly, and conscientious). The correlations can be seen in Table 4.

Table 4

Correlation Table Across the Dependent Measures and Defendant Characteristic Ratings

Variable	*n*	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. Confidence in Guilt	121	-8.69	68.82	-									
2. Damages	54	$3,949.74	$3,404.88	.487**	-								
3. Internal Blame	117	3.60	0.88	.655**	.596**	-							
4. External Blame	114	3.82	0.55	-.445**	-.180	-.581**	-						
5. Competent	121	5.24	1.18	-.283**	-.171	-.500**	.209*	-					
6. Skilled	119	5.23	1.12	-.146	-.049	-.290**	.197*	.597**	-				
7. Warm	121	4.62	1.07	-.090	-.079	-.276**	.057	.328**	.326**	-			
8. Friendly	120	4.86	1.11	-.121	.000	-.244**	.016	.325**	.274**	.770**	-		

Facial Manipulations

First, an independent samples *t*-test was conducted to determine if the gendered appearance manipulation worked in the intended way. The feminine versions of the defendant photos ($M = 5.70$, $SD = 0.95$) were rated as significantly more feminine than the masculine

versions of the defendant photos (M = 5.18, SD = 1.31), $t(119) = -2.51$, $p = .01$, Cohen's $d = -.46$. While these results support that the manipulation of gendered appearance worked as intended, this is a small-to-moderate effect size.

Confirmatory Hypothesis Testing

Confidence in Guilt

When examining confidence in guilt, none of the findings supported my hypotheses. Confidence in guilt scores could range from -100, indicating 100% confidence in a not guilty verdict to +100, indicating 100% confidence in a guilty verdict. Scores near 0 indicate little confidence in the chosen verdict. There was no significant effect of gendered appearance on confidence in guilt, $F(1, 117) = 0.001$, $p = .98$, $\eta_p^2 < .001$; participants' ratings of guilt for the feminine defendants (M = -8.79, SD = 68.59) were similar to ratings for the masculine defendants (M = -8.58, SD = 69.64). There was also no significant effect of case type on confidence in guilt, $F(1, 117) = 2.71$, $p = .10$, $\eta_p^2 = .02$; participants' ratings of guilt were similar between the child negligence case (M = 1.54, SD = 67.17) and the medical malpractice case (M = -19.08, SD = 69.47). There was also no interaction between gendered appearance and case type, $F(1, 117) = .072$, $p = .79$, $\eta_p^2 = .001$. See Table 5.

Table 5

ANOVA Results for Confidence in Guilt Across Conditions

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	9332.90	1	9332.90	1.97	.163	.017
Defendant Gendered Appearance	4.84	1	4.84	.001	.975	.000
Case Type	12834.30	1	12834.30	2.71	.103	.023
Defendant Gendered Appearance x Case Type	342.80	1	342.80	.072	.789	.001
Error	555164.73	117	4745.00			

Punishment

Participants could award damages of 0 to 10,000 dollars. Contrary to my prediction, there was no significant effect of gendered appearance on the amount of damages assigned for the defendant to pay, $F(1, 50) = 0.14$, $p = .71$, $\eta_p^2 = .003$; participants assigned similar amount of damages for both the masculine (M = \$3,819.04, SD = \$3,542.42) and feminine defendants (M = \$4,080.44, SD = \$3,323.86) to pay. However, there was a significant effect of case type on the amount of damages assigned for the defendant to pay, $F(1, 50) = 53.58$, $p < .001$, $\eta_p^2 = .52$; participants assigned more damages in the medical malpractice case (M = \$6,961.90, SD = \$2,667.98) compared to the child negligence case (M = \$2,032.91, SD = \$2,238.10). There was no interaction between gendered appearance and case type, $F(1, 50) = 1.59$, $p = .21$, $\eta_p^2 = .03$. See Table 6.

Table 6

ANOVA Results for Punishment Across Conditions

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	1029397158.1	1	1029397158.1	177.01	<.001	.780
Defendant Gendered Appearance	836583.71	1	836583.71	.144	.706	.003
Case Type	311572165.92	1	311572165.92	53.58	<.001	.517
Defendant Gendered Appearance x Case Type	9219470.42	1	9219470.42	1.59	.214	.031
Error	290768440.80	50	5815368.82			

Internal Blame Attribution

When looking at internal blame attribution, none of the findings were as predicted. There was no significant effect of gendered appearance, $F(1, 113) = 0.13$, $p = .72$, $\eta_p^2 = .001$; participants attributed similar levels of internal blame to both the masculine ($M = 3.63$, $SD = 0.78$) and feminine defendants ($M = 3.57$, $SD = 0.98$). There was no significant effect of case type on internal blame attribution, $F(1, 113) = 0.01$, $p = .98$, $\eta_p^2 < .001$; internal blame attribution was similar amongst both the medical malpractice ($M = 3.59$, $SD = 0.87$) and child negligence cases ($M = 3.60$, $SD = 0.90$). There was also no interaction between gendered appearance and case type, $F(1, 113) = 0.81$, $p = .37$, $\eta_p^2 = .007$. See Table 7.

Table 7

ANOVA Results for Internal Blame Attribution Across Conditions

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	1511.72	1	1511.72	1906.44	<.001	.944
Defendant Gendered Appearance	.099	1	.099	.125	.724	.001
Case Type	.001	1	.001	.001	.977	.000
Defendant Gendered Appearance x Case Type	.643	1	.643	.811	.370	.007
Error	89.60	113	.793			

Exploratory Analyses

Confirmatory hypothesis testing demonstrated that the primary hypotheses were not supported. To better understand how participants received the manipulations and made decisions about the defendant, a series of exploratory analyses were conducted.

External Blame Attribution

When looking at external blame attribution, there was no significant effect of gendered appearance, $F(1, 110) = 0.14$, $p = .71$, $\eta_p^2 = .001$; participants attributed similar levels of external blame across cases with both the masculine ($M = 3.80$, $SD = 0.55$) and feminine defendants ($M = 3.83$, $SD = 0.55$). There was no significant effect of case type on external blame attribution, $F(1, 110) = 2.77$, $p = .10$, $\eta_p^2 = .03$; participants attributed similar levels of external blame across both the medical malpractice ($M = 3.91$, $SD = 0.60$) and child negligence cases ($M = 3.74$, $SD = 0.49$).

There was also no interaction between gendered appearance and case type, $F(1, 110) = 0.06$, $p = .81$, $\eta_p^2 = .001$. See Table 8.

Table 8

ANOVA Results for External Blame Attribution Across Conditions

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	1657.16	1	1657.16	5460.18	<.001	.980
Defendant Gendered Appearance	.043	1	.043	.141	.708	.001
Case Type	.842	1	.842	2.77	.099	.025
Defendant Gendered Appearance x Case Type	.018	1	.018	.058	.810	.001
Error	33.39	110	.303			

Facial Manipulations

Since there was no effect of gendered appearance for the dependent measures, two between-subjects *t*-tests were conducted to determine if there were any differences between how participants interpreted the gendered appearance manipulations of Face 1 versus Face 3. There was no significant difference of gendered appearance for Face 1, $t(60) = -1.42$, $p = .12$, Cohen's $d = -.36$; participants did not rate the gendered appearance of the masculine defendant ($M = 5.55$, $SD = 1.21$) to be any different from the feminine defendant ($M = 5.90$, $SD = 0.70$). However, there was a significant difference of gendered appearance for Face 3, $t(57) = -2.21$, $p = .03$, Cohen's $d = -.58$.; participants rated the feminine defendant ($M = 5.50$, $SD = 1.14$) as significantly more feminine than the masculine defendant ($M = 4.79$, $SD = 1.32$).

Exploratory Testing with Face Three

Factorial ANOVAs were conducted again for each of the dependent measures for just the participants who saw cases with versions of Face 3. Findings were very similar to the initial findings with the only difference being that there was a main effect of case type on confidence in guilt, $F(1, 55) = 7.01$, $p = .01$, $\eta_p^2 = .11$; mock jurors were more confident in a guilty verdict in the child negligence case ($M = 20.21$, $SD = 62.71$) compared to the medical malpractice case ($M = -26.30$, $SD = 65.80$).

Participant Gender

An ANOVA was conducted to determine if the gender of the participants was related to how masculine or feminine the mock jurors rated the defendant to be for each person (Face 1 vs. Face 3). Across both versions of Face 1, there was a main effect of participant gender on the defendant's gendered appearance rating, $F(1, 55) = 6.67$, $p = .01$, $\eta_p^2 = .11$; men rated Face 1 photos to be significantly more feminine ($M = 6.22$, $SD = 0.55$) than women ($M = 5.49$, $SD = 1.10$). Male participants rated the masculine ($M = 6.25$, $SD = 0.71$) and feminine ($M = 6.20$, $SD = 0.42$) versions of Face 1 as being nearly identical to each other, whereas female participants rated the feminine version ($M = 5.72$, $SD = 0.83$) as being a little more feminine than the masculine version ($M = 5.30$, $SD = 1.26$). The differences here can be seen in Figure 1. There was no interaction between the defendant's gendered appearance and participant gender on the appearance ratings of the mock jurors, $F(1, 55) = 0.72$, $p = .40$, $\eta_p^2 = .01$. See Table 9.

Figure 1

Ratings of Defendant Gendered Appearance Based on Participant Gender for Face 1

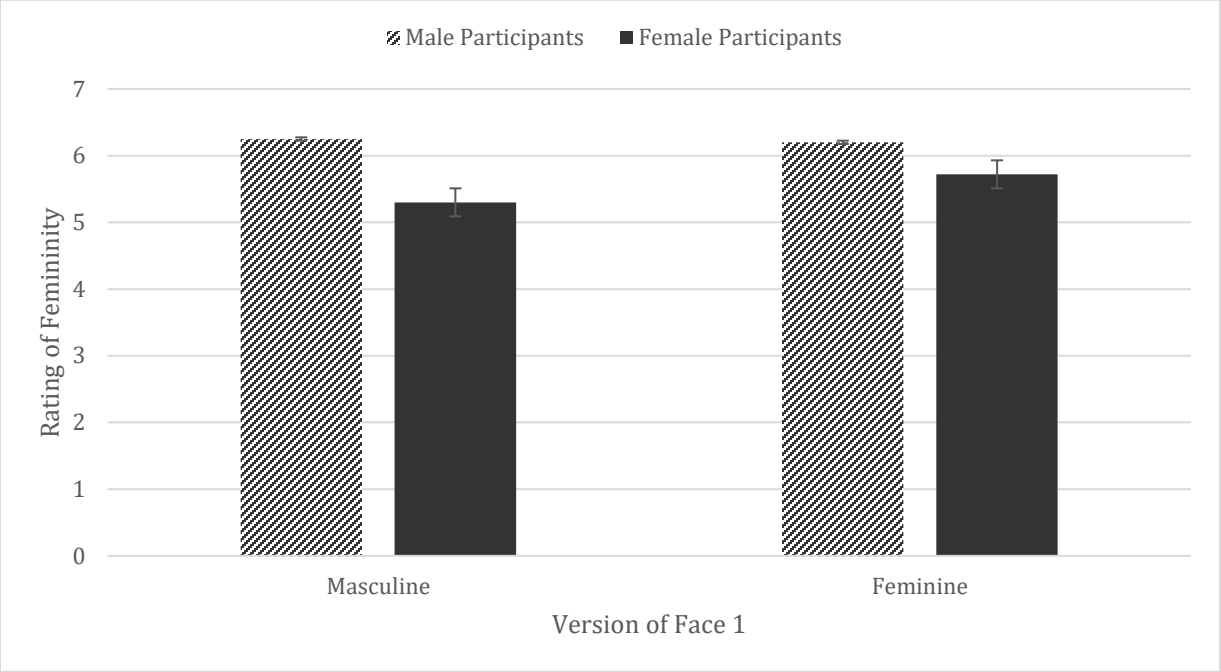

Table 9

ANOVA Results for Participant Gender and Defendant Gendered Appearance on Face 1

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	1700.90	1	1700.90	1813.66	<.001	.971
Participant Gender	6.25	1	6.25	6.67	.013	.108
Defendant Gendered Appearance	.418	1	4.18	.445	.507	.008
Participant Gender x Defendant Gendered Appearance	.676	1	.676	.720	.400	.013
Error	51.58	55	.938			

Across both versions of Face 3, there was a main effect of participant gender on the appearance ratings of the defendants, $F(1, 54) = 10.04$, $p = .003$, $\eta_p^2 = .16$; men ($M = 5.80$, $SD = 0.95$) rated the Face 3 photos as being significantly more feminine than women ($M = 4.79$, $SD = 1.30$). Male participants rated the masculine ($M = 5.60$, $SD = 1.17$) and feminine versions ($M = 6.00$, $SD = 0.67$) of Face 3 as being much more similar to each other compared to how female participants rated the masculine ($M = 4.37$, $SD = 1.21$) and feminine versions ($M = 5.21$, $SD = 1.27$) of Face 3. The differences here can be seen in Figure 2. There was no interaction between gendered appearance and participant gender on the appearance ratings of the mock jurors, $F(1, 54) = 0.48$, $p = .49$, $\eta_p^2 = .01$. See Table 10.

Figure 2

Ratings of Defendant Gendered Appearance Based on Participant Gender for Face 3

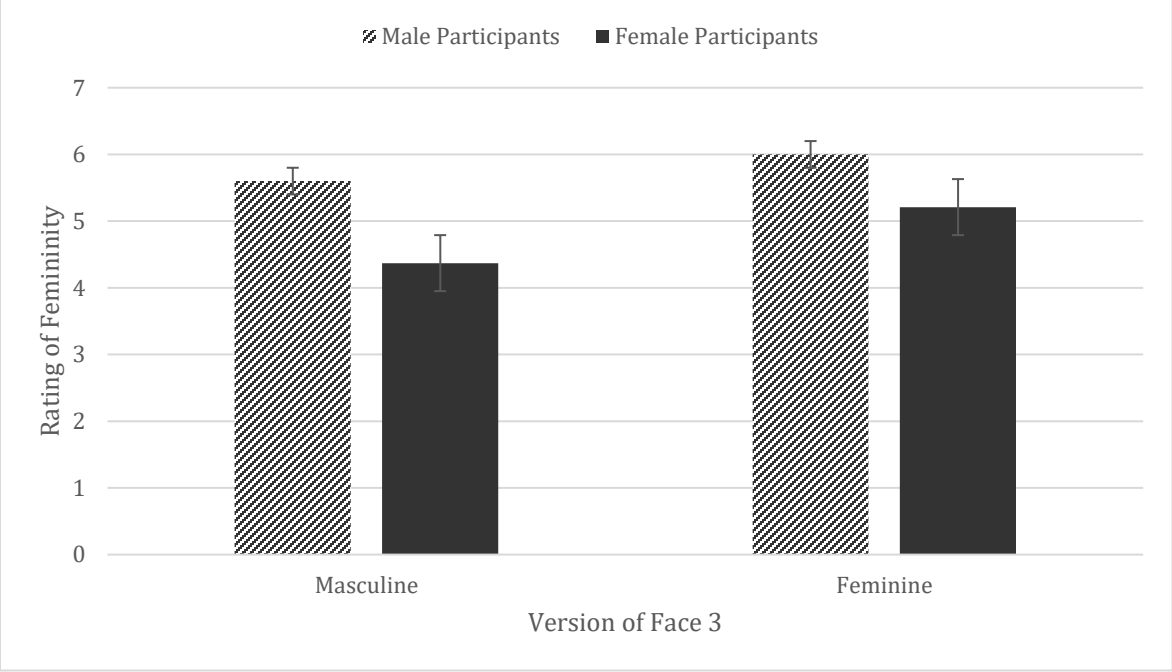

Table 10

ANOVA Results for Participant Gender and Defendant Gendered Appearance on Face 3

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	1469.38	1	1469.38	1102.36	<.001	.225
Participant Gender	13.38	1	13.38	10.04	.003	.953
Defendant Gendered Appearance	5.05	1	5.05	3.79	.057	.066
Participant Gender x Defendant Gendered Appearance	.640	1	.640	.480	.491	.009
Error	71.98	54	1.33			

Additional analyses were conducted to determine if there were any three-way interactions between participant gender, defendant gendered appearance, and case type on each dependent variable. However, no three-way interactions were found. See Tables 11, 12, and 13.

Table 11

Three-Way ANOVA Results for Confidence in Guilt

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	221133.95	1	221133.95	4.48	.037	.039
Participant Gender	17707.59	1	17707.59	3.76	.055	.033
Defendant Gendered Appearance	310.81	1	310.81	.066	.798	.001
Case Type	10361.11	1	10361.11	2.20	.141	.020
Participant Gender x Defendant Gendered Appearance	68.88	1	68.88	.015	.904	.000
Participant Gender x Case Type	1434.59	1	1434.59	.304	.582	.003
Defendant Gendered Appearance x Case Type	527.37	1	527.37	.112	.739	.001
Part. Gender x Def. Gendered Appearance x Case Type	486.39	1	486.39	.103	.749	.001
Error	514068.97	109	4716.23			

Table 12

Three-Way ANOVA Results for Punishment

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	808748996.19	1	808748996.19	137.22	<.001	.761
Participant Gender	11681329.84	1	11681329.84	1.98	.166	.044
Defendant Gendered Appearance	288814.48	1	288814.48	.049	.826	.001
Case Type	287091001.94	1	287091001.94	48.71	<.001	.531
Participant Gender x Defendant Gendered Appearance	1222673.84	1	1222673.84	.207	.651	.005
Participant Gender x Case Type	19461412.67	1	19461412.67	3.30	.076	.071
Defendant Gendered Appearance x Case Type	8203676.00	1	8203676.00	1.39	.245	.031
Part. Gender x Def. Gendered Appearance x Case Type	528323.07	1	528323.07	.090	.766	.002
Error	253428970.72	43	253428970.72			

Table 13

Three-Way ANOVA Results for Internal Blame Attribution

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	1262.86	1	1262.86	1545.07	<.001	.936
Participant Gender	.746	1	.746	.912	.342	.009
Defendant Gendered Appearance	.304	1	.304	.372	.543	.003
Case Type	.015	1	.015	.019	.892	.000
Participant Gender x Defendant Gendered Appearance	1.14	1	1.14	1.39	.240	.013
Participant Gender x Case Type	.001	1	.001	.001	.973	.000
Defendant Gendered Appearance x Case Type	.711	1	.711	.870	.353	.008
Part. Gender x Def. Gendered Appearance x Case Type	.003	1	.003	.004	.951	.000
Error	86.64	106	.817			

Defendant Personal Characteristics

More ANOVAs were conducted to investigate if defendant gendered appearance and case type influenced participants' ratings of defendant characteristics. Ratings of warmth and friendliness were averaged to create one measure, which was used as a factor in the first ANOVA. There was a main effect of case type, no main effect of gendered appearance, and no

interaction (See Table 14). Participants who saw the child negligence case (*M* = 4.96, *SD* = 1.00) rated the defendant as being more warm/friendly than participants who saw the medical malpractice case (*M* = 4.52, *SD* = 1.02). Mock jurors presented with the masculine-appearing defendant (*M* = 4.57, *SD* = 1.04) had similar ratings of warmth/friendliness compared to those who saw the feminine-appearing defendant (*M* = 4.93, *SD* = 1.00).

Table 14

ANOVA Results Using Warmth/Friendliness as the Dependent Measure

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	2692.19	1	2692.19	2686.51	<.001	.959
Gendered Appearance	3.87	1	3.87	3.86	.052	.032
Case Type	5.70	1	5.70	5.69	.019	.047
Gendered Appearance x Case Type	.011	1	.011	.011	.916	.000
Error	116.25	116	1.00			

Similarly, ratings of competence and skill were averaged to create one measure, which was used as a factor in the second ANOVA. There were no significant findings (See Table 15). Participants who saw the child negligence case (*M* = 5.23, *SD* = 0.99) had similar ratings of the defendant's competence/skill compared to those who saw the medical malpractice case (*M* = 5.24, *SD* = 1.08). Mock jurors who were presented with the masculine-appearing defendant (*M* = 5.25, *SD* = 1.08) had similar ratings of competence/skill compared to those who saw the feminine-appearing defendant (*M* = 5.22, *SD* = 0.98).

Table 15

ANOVA Results Using Competence/Skill as the Dependent Measure

Predictor	Sum of Squares	Df	Mean Square	F	P	η_p^2
Intercept	3256.65	1	3256.65	3002.82	<.001	.963
Gendered Appearance	.028	1	.028	.025	.874	.000
Case Type	.006	1	.006	.005	.943	.000
Gendered Appearance x Case Type	.644	1	.644	.594	.443	.005
Error	124.72	115	1.09			

Conscientiousness was used as a factor in the third ANOVA. There were no main effects found, but there was an interaction between gendered appearance and case type on ratings of conscientiousness (See Table 16). Participants who saw the child negligence case (*M* = 4.89, *SD* = 1.18) had similar ratings of the defendant's conscientiousness compared to those who saw the medical malpractice case (*M* = 4.85, *SD* = 1.20). Mock jurors who were presented with the masculine-appearing defendant (*M* = 4.83, *SD* = 1.14) had similar ratings of their conscientiousness compared to those who were presented with the feminine-appearing defendant (*M* = 4.90, *SD* = 1.25). Simple effects tests were done to evaluate the interaction effect; however, there were no significant findings. The masculine-appearing defendant was rated similarly in conscientiousness when accused of medical malpractice (*M* = 5.07, *SD* = 1.17) as when accused of child negligence (*M* = 4.60, *SD* = 1.07). Additionally, the feminine-appearing defendant was rated similarly in conscientiousness when accused of child negligence (*M* = 5.16, *SD* = 1.24) as when accused of medical malpractice (*M* = 4.63, *SD* = 1.22). Within the child negligence case,

participants rated the feminine-appearing defendant (*M* = 5.16, *SD* = 0.21) similarly on levels of conscientiousness compared to the masculine-appearing defendant (*M* = 4.60, *SD* = 0.22). Within the medical malpractice case, participants also rated the feminine-appearing defendant (*M* = 4.63, *SD* = 0.22) similarly on levels of conscientiousness compared to the masculine-appearing defendant (*M* = 5.07, *SD* = 0.22).

Table 16

ANOVA Results Using Conscientiousness as the Dependent Measure

Predictor	Sum of Squares	*df*	Mean Square	F	*p*	η_p^2
Intercept	2863.66	1	2863.66	2065.31	<.001	.946
Gendered Appearance	.124	1	.124	.089	.766	.001
Case Type	.028	1	.028	.020	.886	.000
Gendered Appearance x Case Type	7.48	1	7.48	5.40	.022	.044
Error	162.23	117	1.39			

Discussion

While there were some main effects found regarding case type, there were no significant effects found regarding the gendered appearance of the defendant and the dependent variables of confidence in guilt, punishment, and internal blame attribution. This contradicts past research finding that more masculine defendants tend to be associated with harsher verdicts (Estrada-Reynolds et al., 2017; Strub & McKimmie, 2016). However, there are several plausible explanations of the current findings that may benefit future research.

While the current research focused solely on the appearance of female defendants, past research has solely studied the appearance of male defendants or has focused on written descriptions of masculinity and femininity rather than appearance. Thus, it is possible that the perceived masculinity - criminality pattern only exists amongst male defendants. Alternatively, it could be that this pattern exists across both male and female defendants but that the bias is focused more on the actions and characteristics of the defendant that are provided to them rather than just their physical appearance.

Another explanation could be that the gender of the mock jurors makes a difference in how the facial manipulations are perceived. While no three-way interactions were found between participant gender, defendant gendered appearance, and case type, there were differences in how male and female participants perceived the gendered appearance of each defendant photo. This implies that participant gender may moderate perceptions of the masculine and feminine defendants by case type. However, adequate testing of this three-way interaction requires a much larger sample size.

One last reason that no significant effects of gendered appearance were found could be a general issue of power. According to G*Power, a sample size of at least 128 participants was required ($1 - \beta \geq .80$). While 128 participants completed Study 2, seven participants were excluded from data analysis, potentially resulting in insufficient power. There is a possibility that a significant effect would be found if there were enough participants to detect a smaller effect size.

CHAPTER 4

GENERAL DISCUSSION

The purpose of the current research was to investigate whether the gendered appearance (masculine vs. feminine) of female defendants and the type of case (child negligence vs. medical malpractice) affect mock jurors' judgments of confidence in guilt, punishment, and internal blame attribution in a mock jury trial. I hypothesized that the more masculine-appearing female defendants would elicit harsher judgments across all three of the dependent variables. Additionally, I predicted that the defendant would receive harsher judgments when the gendered stereotype of the case violated the gendered stereotype of the defendant's appearance (with the masculine-appearing woman accused of medical malpractice receiving the harshest judgments). However, these hypotheses were not supported.

One possibility for the lack of significant findings could be that the case facts distracted participants from the gendered appearance manipulation. Both cases involved injury to a human, and similarly to real world cases tried by a jury, who to blame was unclear. Thus, the severity and complexity of the cases may have distracted mock jurors from the defendant photos. Evidence of this is seen in participant responses when they explained the reasoning behind their verdict. For instance, a participant who saw the medical malpractice case stated their reasoning was, "*Due to Dr. Wilson's negligence of knowledge, Ms. Miller will have to live an abnormal life. Ms. Miller will feel the emotional and physical pain every day for the rest of her life*." A participant who saw the child negligence case stated, "*I think that Michelle doesn't seem like a fully neglectful parent, however for the burns to be so severe it involved the children being left alone for an extended period of time.*"

Additionally, the photo of the defendant was only shown once at the beginning of the study before any case facts were provided. The photo was visible while participants were completing the dependent measures; however, if the photo was presented along with the case facts, it is possible that participants would have associated the defendant's appearance more with the severity of the case facts. Research on juror decision making supports a story telling model (Pennington & Hastie, 1993) in which jurors use available information to create a plausible story of what most likely happened. If participants received the facial manipulations as they processed case evidence, the story constructed may have changed.

Additionally, since the effect size for gendered appearance was small in the preliminary analyses, this supports that the photo manipulation may not have been strong enough to produce the desired effects. This problem could potentially be solved by showing the photos along with the case facts or using different photos with stronger manipulations for feminine and masculine characteristics.

The number of participants was smaller than needed for adequate power to detect an effect if one was present. However, more data are being collected beyond what is presented here in order to test the results with a sufficient number of participants. If an adequate sample size does lead to significant results, this research could provide one possible explanation for why men and women are punished differently for the same crimes.

The gender of the mock jurors may also potentially explain the lack of significant findings, as male and female participants rated the masculine and feminine versions of each face differently. With sufficient power, a three-way interaction may be detected between participant gender, case type, and defendant gendered appearance. If future research finds this to be the case, these findings would have important implications for who is being selected to serve on a jury.

Although the hypotheses of this study were not supported, there are a few things to note regarding the correlational data in Table 1. Regardless of case type and gendered appearance, internal blame is strongly associated with both confidence in guilt and damages. This suggests that each of the three dependent measures is related to one another. The personality traits (e.g., competent and warm) are also significantly correlated with internal blame across case type and gendered appearance. However, participants answered the internal blame questions before rating the defendant's characteristics. The order of materials may have affected how these answers related to one another because counterbalancing was not utilized.

Ratings of the defendant's appearance were related to how warm participants rated the defendant to be. However, this was a small correlation. There is a possibility these findings are due to all of the defendants being women, meaning that stereotypically, all of them could be rated similarly on levels of warmth.

The ratings for the defendant characteristics (competent, skilled, conscientious, warm, and friendly) are all above the midpoint on average, which was unexpected. It is possible that these ratings would be different if the cases presented were criminal in nature rather than civil or if mock jurors were judging male defendants rather than female defendants. Another potential explanation for these findings is that the mock jurors may have rated all of the defendants (despite their gendered appearance) as high in both characteristics relating to warmth and competence due to them viewing the defendants as being part of their in-group. According to Fiske et al.'s (2002) work, people are rated as being high in both competence and warmth if they are viewed as being part of one's in-group. The defendant photos depicted White women somewhere around their twenties or thirties, which is similar to the demographics of the participants. There is evidence of participants viewing the defendant as being part of their in-

group through their qualitative data (e.g., "*As a mother, we make decisions where we may not know the ending result but we make them to the best of our ability").* While this would support the in-group favoritism aspect of the stereotype content model, these correlations do not support Fiske's (2012) smart-but-cold stereotype about men or the warm-but-dumb stereotype about women.

As for the predicted interaction hypothesis, one reason that this was not supported may be that participants simply did not detect the gendered appearance of the face enough. If the face was not being processed enough, the manipulation was strong enough. Thus, we would not really be able to find an interaction with case type even if an interaction did exist.

Limitations

Along with the aforementioned reasons for non-significant findings, there are a handful of limitations that should be considered regarding the current research. First, the participants in these studies were undergraduate students at the same university. Thus, while the participants were determined to be jury-eligible, they do not accurately reflect the population likely to be serving on a jury. In fact, it is not uncommon for undergraduate students to be exempt from jury duty due to their responsibilities as students. Additionally, in an actual trial, there would likely be parents of some sort on a jury, especially mothers, who would be selected by the attorneys. These parents may be much more guided by the role expectations (masculine vs. feminine) that were not met by the defendant in the current study. However, while some of the undergraduate participants may be parents, it is likely that the majority of them are not in a parenting role. Future research could attempt to solve this limitation by recruiting participants from a community population rather than an undergraduate population.

Another limitation is how the trial material was presented. In a real trial, the jurors are face-to-face with the people involved in the trial. This means that jurors would actively be viewing the testimonies rather than just reading transcripts. Thus, the mock jurors in this study may not have taken their role as seriously as they would have in real life jury situations. Future research could avoid this problem by presenting the trial stimuli as a recorded video where the participants are actively seeing testimony. On the other hand, the qualitative data collected from participants suggest that they were reasonably invested in providing a just outcome. Data collection occurred in person, and the length of the qualitative comments that participants left suggest that they were likely more invested than a typical experimental study.

Future Directions

Although the gendered appearance of the defendants in the current research did not affect the trial outcomes, it is important that research continues in this area. Specifically, research should continue to try to explain why there is a sentencing disparity between men and women. To better understand where that gendered bias seems to present itself, future research could take into consideration participant gender along with the gendered appearance of both male and female defendants. Researchers should also consider using more realistic trial stimuli, such as a video recording of the trial proceedings. This would allow for the defendant's face to be present during the entirety of the trial while ensuring their face is associated with their testimony.

Additionally, using a larger sample size from a community sample rather than an undergraduate sample may be beneficial in making results more generalizable to the jury-eligible population within the United States. All of these suggestions would allow for a better understanding of what type of factors jurors take into consideration while determining the verdict

in a trial, whether they are aware of it or not. The results from research in this area could be applied to real-life jury settings to create a more equitable jury for all defendants.

Conclusions

While this research did not result in significant findings regarding the gendered appearance of defendants, it adds a new finding to the existing literature regarding laypeople having gendered biases about defendants. Prior research established that amongst male defendants, participants will pick out the most masculine-appearing face to be the criminal (Estrada-Reynolds et al., 2017). There seems to be evidence that this pattern may persist for female defendants (Strub & McKimmie, 2016); however, researchers have focused on written descriptions of masculinity and femininity rather than facial manipulations. The current research may have failed to detect this pattern amongst facial manipulations of female defendants due to a lack of power, distracting case facts, or failing to consider participant gender as a variable. However, further research is needed to control for these aspects and more accurately conclude whether the gendered appearance of female defendants makes a difference in the outcome of a trial.

REFERENCES

Ashton-James, C. E., Tybur, J. M., Grieber, V., & Costa, D. (2019). Stereotypes about surgeon warmth and competence: The role of surgeon gender. *PLOS ONE, 14*(2). https://doi.org/10.1371/journal.pone.0211890

Bergeron, C. E., & McKelvie, S. J. (2004). Effects of defendant age on severity of punishment for different crimes. *The Journal of Social Psychology, 144*(1), 75-90. 10.3200/SOCP.144.1.75-90

Blashill, A. J., & Powlishta, K. K. (2009). Gay stereotypes: The use of sexual orientation as a cue for gender-related attributes. *Sex Roles, 61,* 783-793. https://doi.org/10.1007/s11199-009-9684-7

Cima, M., Merckelbach, H., Butt, C., Kremer, K., Knauer, E., & Schellbach-Matties, R. (2006). It was not me: Attribution of blame for criminal acts in psychiatric offenders. *Forensic Science International, 168,* 143-147. 0.1016/j.forsciint.2006.07.004

Doerner, J. K., & Demuth, S. (2014). Gender and sentencing in the federal courts: Are women treated more leniently? *Criminal Justice Policy Review*, *25*(2), 242-269. https://doi.org/10.1177/0887403412466877

Estrada-Reynolds, V. C., Reynolds, J. J., McCrea, S. M., & Freng, S. (2017). I don't like the cut of your jib: Perceived facial masculinity as a cue to criminality. *Psychiatry, Psychology and Law, 24*(3), 392-409. http://dx.doi.org/10.1080/13218719.2016.1247639

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. Download PDF

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype

content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902. 10.1037//0022-3514.82.6.878

Fiske, S. (2012). Managing ambivalent prejudices: The smart-but-cold, and the warm-but-dumb stereotypes. *The ANNALS of the American Academy of Political and Social Science, 639*(1), 33-48. https://doi.org/10.1177/0002716211418444

Funk, F., Walker, M., & Todorov, A. (2017). Modelling perceptions of criminality and remorse from faces using a data-driven computational approach. *Cognition and Emotion, 31*(7), 1431-1443. http://dx.doi.org/10.1080/02699931.2016.1227305

Glick, P., Wilk, K., & Perreault, M. (1995). Images of occupations: Components of gender and status in occupational stereotypes. *Sex Roles, 32*(9-10), 565-582. https://doi.org/10.1007/BF01544212

Gudjonsson, G. H., & Singh, K. K. (1989). Revised Gudjonsson Blame Attribution Inventory. *PsycTESTS*. https://doi.org/10.1037/t13556-000

Johnston, J. B., Kennedy, T. D., & Shuman, G. (1987). Gender differences in the sentencing of felony offenders. *Federal Probation*, *51*(1), 49-55. https://www.ojp.gov/pdffiles1/Digitization/105796NCJRS.pdf

Lindquist v. Dengel, 20 Wn. App. 630, 20 Wash. App. 630, 581 P.2d 177 (1978). https://casetext.com/case/lindquist-v-dengel-1

Pennington, N., & Hastie, R. (1993). The story model for juror decision-making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192-222). Cambridge University Press.

Petsko, C. D., & Bodenhausen, G. V. (2017). Race-crime congruency effects revisited: Do we

take defendants' sexual orientation into account? *Social Psychological and Personality Science*, *10*(1), 73-81. https://doi.org/10.1177/1948550617736111

Russell, B., Ragatze, L., & Kraus, S. W. (2012). Expert testimony of the battered person syndrome, defendant gender, and sexual orientation in a case of duress: Evaluating legal decisions. *Journal of Family Violence*, *27*(7), 659-670. https://doi.org/10.1007/s10896-012-9459-8

Sekaquaptewa, D., & Espinoza, P. (2004). Biased processing of stereotype-incongruency is greater for low than high status groups. *Journal of Experimental Social Psychology*, *40*, 128-135. 10.1016/S0022-1031(03)00093-3

Strub, T., & McKimmie, B. M. (2016). Sugar and spice and all things nice: The role of gender stereotypes in jurors' perceptions of criminal defendants. *Psychiatry, Psychology and Law*, *23*(4), 487-498. http://dx.doi.org/10.1080/13218719.2015.1080151

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Walker, M., Schönborn, S., Greifeneder, R., & Vetter, T. (2018). The Basel Face Database: A validated set of photographs reflecting systematic differences in Big Two and Big Five personality dimensions. *PLoS One*, *13*(3). 10.1371/journal.pone.0193190

Wen, F., Zuo, B., Ma, S., Xu, Y., Coley, J. D., & Wang, Y. (2020). Do we see masculine faces as competent and feminine faces as warm? Effects of sexual dimorphism on facial perception. *Evolutionary Psychology*, *18*(4). https://doi.org/10.1177/1474704920980642

Williams v. State, 235 S.W.3d 742 (2007). https://casetext.com/case/williams-v-state-518?q=williams%20v.%20state%202007&sort=relevance&p=1&type=case

APPENDIX A

MEDICAL MALPRACTICE CASE MANIPULATION

Miller v. Wilson, 2021

CASE FACTS
Defendant Name: Michelle Wilson
Type of Case: Medical Malpractice
Case Number: 001027120

Case Summary:

In February of 2019, Lisa Miller saw Dr. Michelle Wilson, a general practitioner, because Lisa had some symptoms that appeared to be bronchitis. During the appointment, Dr. Wilson ordered an x-ray to be done on Lisa's chest. The x-ray was sent to Dr. Sarah Anderson to examine. This is when Dr. Anderson saw signs of Lisa having tuberculosis. Thus, it was suggested that Lisa have additional chest x-rays in 3 months as well as a mucus sample taken to confirm the diagnosis. Dr. Wilson relayed this information to Lisa. In April of 2019, these suggestions were discussed again during a routine physical exam. Lisa saw Dr. Wilson again in May of 2019 for an unrelated issue, however, nothing was done in regard to the tuberculosis diagnosis. In October of 2020, Lisa saw Dr. Wilson once again for symptoms of pneumonia. Treatment was given to Lisa, but the treatment was not successful. Lisa then needed to be hospitalized for her persistent pneumonia. While in the hospital, Lisa received test results confirming the tuberculosis diagnosis. After this diagnosis, Lisa was under Dr. Anderson's care for treatment of tuberculosis. Dr. Anderson ended up removing part of Lisa's lung, which cured her of tuberculosis.

Lisa Miller is filing a medical malpractice case against Dr. Wilson. Lisa states that while she no longer has tuberculosis, the delayed diagnosis is what led to part of her lung needing to be removed. Lisa claims that Dr. Wilson did not diagnose her properly to begin with, which led her to suffer the consequences of a partial lung removal. Lisa claims that the partial lung removal has caused a significant loss in salary due to her no longer being able to perform her previous job. Lisa states that if Dr. Wilson had provided a diagnosis earlier, the partial lung removal would not have been necessary.

Trial Transcription:

DEFENSE ATTORNEY: Dr. Wilson, do you recall your consultation with Ms. Miller in February of 2019?

DEFENDANT: Somewhat, yes. I do not remember every detail about it, but I remember that she was presenting with symptoms of bronchitis.

DEFENSE ATTORNEY: What did you do to treat the symptoms of bronchitis?

DEFENDANT: I ordered x-rays and sent them to another doctor for analysis.

DEFENSE ATTORNEY: Why were the x-rays examined by a different doctor?

DEFENDANT: I am a general practitioner, so I sent them to a specialist to be analyzed.

DEFENSE ATTORNEY: Do you recall what the x-rays showed?

DEFENDANT: Dr. Anderson said that she had signs of tuberculosis and suggested that I test a mucus sample and also have additional x-rays done in the next few months to confirm the diagnosis.

DEFENSE ATTORNEY: Did you relay this information to Ms. Miller?

DEFENDANT: Yes, I passed this information along to her.

DEFENSE ATTORNEY: When did you pass this information along to her?

DEFENDANT: I told her at the same time that I gave her the x-ray results, so in February of 2019.

DEFENSE ATTORNEY: Thank you, Dr. Wilson.

Direct Examination of the Eyewitness, Jackie Olson:

PROSECUTOR: Please state your name for the record.

WITNESS: Jackie Olson.

PROSECUTOR: How do you know the defendant?

WITNESS: I am a receptionist at the doctor's office Michelle Wilson works at.

PROSECUTOR: What are some of your duties as a receptionist in the office?

WITNESS: I help keep track of patient files and schedule appointments.

PROSECUTOR: Do you recall scheduling an appointment for Ms. Miller to receive follow-up x-rays?

WITNESS: Yes, I remember calling her sometime after her original appointment to schedule x-rays.

PROSECUTOR: Approximately how long after receiving the x-ray results do you think this happened?

WITNESS: I don't recall the exact timing, but we typically call to schedule those appointments within 2 weeks of receiving results like that.

PROSECUTOR: So there was no appointment scheduled the day that Ms. Miller was contacted about the results?

WITNESS: No, we scheduled it later on.

PROSECUTOR: And if there are any appointment cancellations, do you also deal with that?

WITNESS: Yes, I contact patients reminding them of upcoming appointments and also deal with any cancellations or rescheduling that needs to happen.

PROSECUTOR: Do you recall an appointment cancellation for Ms. Miller?

WITNESS: Yes, I remember calling Ms. Miller to tell her that we had to cancel the appointment due to Dr. Wilson having to be out of the office the day her appointment was scheduled.

PROSECUTOR: So, what you are saying is that Ms. Miller's follow-up appointment wasn't scheduled until sometime after the day the x-ray results came in and then Dr. Wilson had to cancel that appointment, is that correct?

WITNESS: Yes, that is correct.

PROSECUTOR: Thank you, Ms. Olson.

Cross Examination of the Eyewitness, Jackie Olson:

DEFENSE ATTORNEY: Ms. Olson, you said that some of your duties involve appointment cancellations and rescheduling, correct?

WITNESS: Yes, that is correct.

DEFENSE ATTORNEY: How long have you worked with Dr. Wilson?

WITNESS: For about 3 years.

DEFENSE ATTORNEY: During that time, approximately how often would you say that Dr. Wilson has been the one to cancel the appointment rather than the patient?

WITNESS: It has happened a few times, but it is pretty rare for that to happen.

DEFENSE ATTORNEY: Are you aware of whether Ms. Miller's follow-up appointment was ever rescheduled?

WITNESS: To my knowledge, it was not rescheduled.

DEFENSE ATTORNEY: Did you continue to reach out to Ms. Miller to reschedule the appointment?

WITNESS: Yes, I reached out a couple times and left voicemails messages but was unable to reach her.

DEFENSE ATTORNEY: Thank you, Ms. Olson.

Direct Examination of the Expert Witness, Dr. Melissa Johnson:

DEFENSE ATTORNEY: Please state your name for the record.

EXPERT WITNESS: Melissa Johnson.

DEFENSE ATTORNEY: What is your occupation, Dr. Johnson?

EXPERT WITNESS: I am a pulmonologist. This means that I am a doctor that specializes in examining and treating the respiratory system.

DEFENSE ATTORNEY: How long have you been a pulmonologist?

EXPERT WITNESS: For 7 years.

DEFENSE ATTORNEY: What qualifies you for that position?

EXPERT WITNESS: I completed medical school and a residency and fellowship at the University of New Mexico. I am now a licensed pulmonologist in the state of Georgia.

DEFENSE ATTORNEY: As a pulmonologist, have you examined patients with tuberculosis before?

EXPERT WITNESS: Yes, I have examined and treated patients with tuberculosis before.

DEFENSE ATTORNEY: Prior to arriving here today, did you take a look at the x-rays of Ms. Miller's lungs?

EXPERT WITNESS: Yes, I was given copies of the x-rays to analyze.

DEFENSE ATTORNEY: What are some treatment options that you would see fit for Ms. Miller based on these x-rays?

EXPERT WITNESS: In this case, the first option would be drug therapy. If that did not work, then surgical removal of part of the lung would be the last option.

DEFENSE ATTORNEY: Is there anything about the x-rays that suggest to you that Ms. Miller needed to have part of her lung surgically removed instead of trying drug therapy?

EXPERT WITNESS: No, there is nothing about the x-rays that suggest drug therapy wouldn't work.

DEFENSE ATTORNEY: Thank you, Dr. Johnson. I have no further questions.

Cross Examination of the Expert Witness, Dr. Melissa Johnson:

PROSECUTOR: Dr. Johnson, you stated that you have examined and treated patients with tuberculosis before, is that correct?

EXPERT WITNESS: Yes, that is correct.

PROSECUTOR: Approximately how many times would you say you have worked with patients with tuberculosis before?

EXPERT WITNESS: I would say between 25 and 40 times throughout my career.

PROSECUTOR: If it was suggested by another doctor that additional steps be taken to confirm a tuberculosis diagnosis, in what timeframe would you typically get that set up?

EXPERT WITNESS: I would typically get the mucus sample collected and schedule the additional x-rays within a week of receiving that suggestion.

PROSECUTOR: And why do you typically get that done within a week?

EXPERT WITNESS: Tuberculosis is a serious medical condition that could continue to spread, so getting all of that confirmed quickly can reduce the spread of it.

PROSECUTOR: How would you say that a partial lung removal, caused by tuberculosis, would affect someone's daily life?

EXPERT WITNESS: It is common for people to experience shortness of breath, especially with any sort of physical activity. If it is severe enough, some people may also require oxygen tanks.

PROSECUTOR: So would you say this surgery could impair someone's daily living?

EXPERT WITNESS: Yes, it could be harder for someone who underwent this type of surgery to maintain the same daily lifestyle they previously had.

PROSECUTOR: Thank you, Dr. Johnson.

Closing Statements:

The defense states that the prosecution has failed to prove by a preponderance of evidence that the defendant should be found guilty since she did reach out to Ms. Miller about needing to run some follow-up tests, mentioned the follow-up tests in another appointment, and reached out multiple times to reschedule the follow-up tests without any response from Ms. Miller. The defense additionally argues that the partial lung removal was not deemed necessary to treat Ms. Miller's tuberculosis because Dr. Anderson could have tried drug therapy to see if it worked before resorting to the partial lung removal.

The prosecution states that they have proven by a preponderance of evidence that the defendant should be found guilty since she failed to provide proper care for her patient, as she was aware that she needed to do further testing and failed to do so in a timely manner, while also canceling the patient's appointment and failing to reschedule it. The prosecution argues that Ms. Miller's appointment should have been scheduled as soon as possible and that due to Dr. Wilson's failure to provide care in a timely manner, Ms. Miller is unable to work her previous job.

Legal Definition:

A medical malpractice case must involve **all** of the three following elements:

- A standard of care violation
    - A medical professional failed to adhere to the standard of care, or acceptable medical practice, that is expected of them when treating patients.
- Causal of an injury
    - An injury occurred to the patient that would not have otherwise occurred. The patient will need to provide enough evidence to show that the medical professional's actions, or lack thereof, is what led to this specific injury.
- Causal of significant damage
    - Due to their injury, the patient has had to endure a significant degree of monetary damage. The patient will need to provide enough evidence to show that the injury they endured resulted in expensive care afterwards.

APPENDIX B

CHILD NEGLIGENCE CASE MANIPULATION

State v. Wilson, 2021

CASE FACTS
Defendant Name: Michelle Wilson
Type of Case: Child Negligence
Case Number: 001027120

Case Summary:

On the evening of February 19, 2019, Michelle Wilson and her two children were at her mother, Lisa Miller's house, where they lived. Michelle and her children were picked up by Michelle's boyfriend, Ryan, and taken to his apartment, where he planned to watch over the children while Michelle worked the night shift. After arriving at the house, Michelle changed clothes and left for work. Ryan's apartment did not have working utilities at the time, so the children were dressed in multiple layers and candles were lit for warmth in the bedroom while they slept. Ryan checked on the children multiple times before going to bed around 9:00 pm. He was awakened around 11:00 pm from the children crying. This is when he noticed they had marks on their arms and hands, which were later determined to be severe third-degree burns from playing with the melted candle wax. The children were immediately taken to the emergency room and treated for their burns, which is when Michelle met back up with them. Child protective services were called to investigate by the medical professionals in the emergency room due to two children from the same household both having severe burns.

The state of Georgia is filing a child negligence case against Michelle Wilson. They state that Michelle should not have moved the children from her house, which had working utilities, to her boyfriend's house for the night, which lacked working utilities at the time. The state claims that Michelle was negligent in providing proper care for her children, as she left them in an unsafe environment, which ultimately led to their injuries.

Trial Transcription:

DEFENSE ATTORNEY: Ms. Wilson, what were you doing on the evening of February 19, 2021?

DEFENDANT: I am a certified nursing assistant at a nursing home, so I was working the night shift.

DEFENSE ATTORNEY: What did you do with your children while you were working?

DEFENDANT: They were staying with my boyfriend, Ryan, for the night.

DEFENSE ATTORNEY: Were you with Ryan the whole day?

DEFENDANT: We spent the morning at my mother's house, where we live. I knew that I was going to be working the night shift a few weeks before, so Ryan and I had planned that we would spend the day with him and that he would take care of the kids while I worked that night.

DEFENSE ATTORNEY: Has Ryan watched your children in the past?

DEFENDANT: Yes, we have been together for 3 years, so he has spent plenty of time with them. It's usually either him or my mother who watches the kids when I work at night.

DEFENSE ATTORNEY: Have your kids ever been injured when staying with Ryan before?

DEFENDANT: No, they have never had injuries like this before.

DEFENSE ATTORNEY: Is there a reason you did not have your mother watch the children that evening?

DEFENDANT: No, we had just planned to spend the day with Ryan in advance, so I thought it would be easier if he just watched them that night.

DEFENSE ATTORNEY: Thank you, Ms. Wilson.

Direct Examination of the Eyewitness, Lisa Miller:

PROSECUTOR: Please state your name for the record.

WITNESS: Lisa Miller.

PROSECUTOR: How do you know the defendant?

WITNESS: I am Michelle Wilson's mother and the children's grandmother.

PROSECUTOR: Where were you on the evening of February 19, 2021?

WITNESS: I spent the evening at home.

PROSECUTOR: Where do the defendant and her children live?

WITNESS: They live with me at my home.

PROSECUTOR: How long have they lived with you?

WITNESS: For the last 2 years.

PROSECUTOR: Were you asked to watch the children that night?

WITNESS: No, Michelle did not ask me to watch them.

PROSECUTOR: Is there a reason you were not asked to watch the children?

WITNESS: Not that I know of.

PROSECUTOR: Were you aware that the defendant's boyfriend was going to watch the children that night?

WITNESS: Yes, I was aware.

PROSECUTOR: Did you say anything when the defendant told you she was taking the children there?

WITNESS: I told her that it may not be the best idea because he was having issues with his utilities.

PROSECUTOR: Thank you, Ms. Miller.

Cross Examination of the Eyewitness, Lisa Miller:

DEFENSE ATTORNEY: Ms. Miller, have you ever doubted Ms. Wilson's capability as a mother before?

WITNESS: No, I haven't.

DEFENSE ATTORNEY: How long have you known her boyfriend, Ryan?

WITNESS: For about 3 years.

DEFENSE ATTORNEY: Would you say you have a good relationship with him?

WITNESS: Yes, I would say so.

DEFENSE ATTORNEY: Are you aware of any other times that he has been responsible for the children?

WITNESS: Yes, he has watched them quite a few times.

DEFENSE ATTORNEY: On those occasions, are you aware of any other incidents that have happened regarding the children?

WITNESS: No, not that I'm aware of.

DEFENSE ATTORNEY: Thank you, Ms. Miller.

Direct Examination of the Expert Witness, Dr. Melissa Johnson:

DEFENSE ATTORNEY: Please state your name for the record.

EXPERT WITNESS: Melissa Johnson.

DEFENSE ATTORNEY: What is your occupation, Dr. Johnson?

EXPERT WITNESS: I am a forensic dermatologist. This means that I am a doctor that specializes in determining the cause of injuries to places like the skin.

DEFENSE ATTORNEY: How long have you been a forensic dermatologist?

EXPERT WITNESS: For 7 years.

DEFENSE ATTORNEY: What qualifies you for that position?

EXPERT WITNESS: I completed medical school and a residency and fellowship at the University of New Mexico. I am now a licensed dermatologist in the state of Georgia.

DEFENSE ATTORNEY: As a dermatologist, have you examined burns?

EXPERT WITNESS: Yes, I have examined and treated burns before.

DEFENSE ATTORNEY: Prior to arriving here today, did you take a look at photos of the burns that the children endured?

EXPERT WITNESS: Yes, I was given photos of their burns to analyze.

DEFENSE ATTORNEY: Are these typical burns that you would see that result from touching something like hot candle wax?

EXPERT WITNESS: Yes, these look like burns that would typically occur in that situation.

DEFENSE ATTORNEY: Is there anything about the burns that suggest to you that someone inflicted these wounds upon the children?

EXPERT WITNESS: No, there is nothing about the burns that suggest someone else caused them.

DEFENSE ATTORNEY: Thank you, Dr. Johnson. I have no further questions.

Cross Examination of the Expert Witness, Dr. Melissa Johnson:

PROSECUTOR: Dr. Johnson, you stated that you have examined and treated burns before, is that correct?

EXPERT WITNESS: Yes, that is correct.

PROSECUTOR: Approximately how many times would you say you have examined burns like this before?

EXPERT WITNESS: I would say between 25 and 40 times throughout my career.

PROSECUTOR: How severe would you say the burns were in the photo of the children?

EXPERT WITNESS: They show signs of being third-degree burns, which are very severe.

PROSECUTOR: And what tells you that these are third-degree burns?

EXPERT WITNESS: You can see that the burns are fairly deep and affect more than just the top two layers of the skin. Their skin also appears to be somewhat leather-like in texture, which is a sign of third-degree burns.

PROSECUTOR: How long would you estimate that it would take for these burns to occur from candle wax?

EXPERT WITNESS: It would have to be some sort of extended period of time. More than just a couple minutes.

PROSECUTOR: What kind of treatment do these types of burns require?

EXPERT WITNESS: They would require immediate treatment. Each burn is a little different, but it could be anywhere from just cleaning and removing dead skin to receiving an IV or getting a skin graft to replace the wounded area.

PROSECUTOR: Thank you, Ms. Johnson.

Closing Statements:

The defense states that the prosecution has failed to prove by a preponderance of evidence that the defendant should be found guilty since her boyfriend had watched the children many times before without anything happening to them, there was no reason to question his capability as a proper babysitter, and there is no evidence suggesting that he inflicted the wounds upon the children.

The prosecution states that they have proven by a preponderance of evidence that the defendant should be found guilty, since she failed to provide proper supervision for her children,as she was

aware that her boyfriend's utilities may not have been working, which then resulted in the childrens' severe injuries. The prosecution argues that the children should not have been left unsupervised around candles for an extended period of time and could have stayed at home with Lisa Miller, where they would have had working utilities.

Legal Definition:

A child negligence case must involve **one** of the three following elements:

- Failure of a parent or caregiver to provide suitable care, education, or other necessity that is required for a child's physical, mental, or emotional wellbeing
- Failure of a parent or caregiver to provide a child with proper supervision
- Abandonment of a child by a parent or legal guardian

APPENDIX C

DEPENDENT MEASURES

Do you find the defendant guilty or not guilty?

a. Guilty

b. Not guilty

How confident are you in your verdict?

(Presented on a slider scale ranging from 0%-100%)

In this case, how much should the defendant, if guilty, pay in damages to the plaintiff?

(Presented on a slider scale ranging from $0-$10,000)

What is your reasoning for assigning the amount of damages that you chose?

_____

Based upon your understanding of the case, if damages were awarded, who would receive the money?

Why did you pick the verdict (guilty vs. not guilty) that you chose?

_____

What type of case did you read about?

a. Medical malpractice

b. Child negligence

c. Invasion of privacy

d. Other

The defendant was a:

a. Man

b. Woman

c. I don’t remember

How masculine or feminine does the defendant look?

a. Very masculine

b. Masculine

c. Somewhat masculine

d. Neither masculine nor feminine

e. Somewhat feminine

f. Feminine

g. Very feminine

How old does the defendant look? Please enter an age in years below.

Please indicate the extent to which you agree the defendant is:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Competent	○	○	○	○	○	○	○
Skilled	○	○	○	○	○	○	○
Warm	○	○	○	○	○	○	○
Friendly	○	○	○	○	○	○	○
Conscientious	○	○	○	○	○	○	○

APPENDIX D

DEMOGRAPHIC QUESTIONS

What is your sex?

a. Male

b. Female

c. Prefer not to say

What is your gender or gender identity? (Select all that apply)

a. Man

b. Woman

c. Non-binary

d. Transgender

e. Genderqueer

f. Agender

g. Another gender _____

h. Prefer not to say

What is your age (in years)?

_____

What is your race or racial identity? (Select all that apply)

a. American Indian or Alaskan Native

b. Asian or Asian American

c. Black or African American

d. White

e. Hispanic

f. Native Hawaiian or other Pacific Islander

g. Another race _____

h. Prefer not to say

What best describes your political beliefs?

a. Very conservative

b. Conservative

c. Somewhat conservative

d. Neutral

e. Somewhat liberal

f. Liberal

g. Very liberal