Discerning Evidence in Civil Sexual Assault Cases

by

Kristen McCowan

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2020 by the
Graduate Supervisory Committee:

Tess M.S. Neal, Chair
Kelly Davis
Jessica Salerno

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

Biases have been studied in many legal contexts, including sexual assault cases. Sexual assault cases are complex because there are many stages that biases can come into play and have lasting effects on the rest of the case proceedings. One aspect that has not been widely explored is how people perceive institutions' liability in sexual assault cases based on an obligation to create non-discriminating environments for members and employees according to laws like Title VII and Title IX. The current project focused on how and why cognitive biases affect laypeople's judgment. Specifically, laypeople's ability to discern the strength of evidence in civil sexual assault cases against institutions. This was addressed in a series of two studies, with samples collected from Prolific Academic (n = 90) and Arizona State University students (n = 188) for Study 1 (N = 278), and Prolific Academic in Study 2 (N = 449). Both studies used Latin-square design methods, with within and between subject elements, looking at how confirmation bias influenced decisions about whether an institution demonstrated negligence, and thus liability, in the way they responded to sexual assault allegations within their institution. Results from these studies suggest that jurors are overall accurately able to differentiate between weak and strong cases. However, consistent with previous literature, jurors may be susceptible to confirmation bias from outside information (e.g., news stories) and negatively influenced by their personal attitudes (e.g., rape myth acceptance). Given the increased attention of the Me Too movement, these results provide an initial insight into how individuals may be judging these types of cases against institutions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Discerning Evidence Strength in Civil Sexual Assault Cases

Our justice system places substantial responsibility in the hands of laypeople; such as expecting jurors to discern evidence to reach a verdict. Do they accurately assess evidence strength? There is evidence to suggest jurors are able to discriminate between weak and strong cases (Leippe, Eisenstadt, Rauch, & Seib, 2004); however, others have found evidence that jurors may overweight weak evidence (Smith, Bull, & Holliday, 2011), and perhaps may not assign enough weight to strong evidence (Kaye & Koehler, 1991). While these studies provide initial evidence for a potential problem in how jurors judge cases, it is important to understand when and why this may be happening in order to help them make the best use of evidence when reaching a verdict.

There is a certain amount of trust that the decisions made in these circumstances are impartial, and made objectively (U. S. Const. amend. IV); however, previous research (e.g., Harley, 2007) suggests this is not always the case. Cognitive and emotional biases may interfere with legal decision processes and play a role in why evidence is not always accurately weighed. One legal area which has been less studied than others is how people make decisions and to what degree their decisions might be affected by biases in cases of sexual assault, especially those that are pursued as civil cases against institutions.

Sex Discrimination in the Law

Discrimination based on sex has a longstanding history in the law, specifically Title VII (Civil Rights Act of 1964) and Title IX (Education Amendments Act of 1972, 2018). Title VII prohibits employment discrimination based on race, color, religion, sex, and national origin. Title IX similarly prohibits people from being excluded from participation in, be denied the benefits of, or be subjected to discrimination under any

1

education program or activity receiving Federal financial assistance. Essentially both of these laws serve to prevent differential treatment in the workforce or educational domain based on a person's sex. Sexual harassment, and in turn any attempted or completed acts of sexual assault, is prohibited under these acts. Thus, companies and schools are under obligation to provide certain measures to help ensure a safe environment for their members in effort to prevent, or appropriately handle, these circumstances.

Furthering these protections, a Dear Colleague Letter from the U.S. Department of Education's Office of Civil Rights in accordance to Title IX (Ali, 2011) guides and addresses the need for schools to uphold proper measures and protocols. This letter clearly states that educational institutions must: 1. Disseminate a notice of nondiscrimination; 2. Designate at least one employee to coordinate its efforts to comply with and carry out its responsibilities under Title IX; and 3. Adopt and publish grievance procedures providing for prompt and equitable resolutions of student and employee sex discrimination complaints. Schools must promptly investigate any complaints. Their investigation is also separate from any law enforcement investigation, and a law enforcement investigation does not relieve the school of its obligation. While companies under Title VII are not faced with as much liability, the company can be found liable if it knew, or should have known about the harassment and failed to take prompt and appropriate corrective action. These are just a few examples outlined by the law that set precedent for institutions being responsible for preventing discrimination, including sexual harassment and assault.

The #MeToo movement began as a movement against sexual harassment and assault, originally used on social media by sexual harassment survivor Tarana Burke. The

purpose of this movement is to empower women, and become a way for victims to tell their experience with sexual violence and stand in solidarity with other survivors, demonstrating the magnitude of the problem (Garcia, 2017). Since the #MeToo movement began, there have been more women and men coming forward with their stories, and an undeniable increase in sexual harassment and discrimination complaints (Chiwaya, 2018). Filing a complaint is far from the end of the road for these victims and institutions, and it seems that the road to justice in these cases is not easy. One common theme in media attention has focused on the legal path; showing how victims face challenges proving harassment occurred (Winter, 2018). One potential hurdle in these cases that has been a relative pattern in juror decision-making is the influence of biases and heuristics that may come into play and impact decisions (e.g., Süssenbach, Albrecht, & Bohner, 2017).

Cognitive Heuristics

Cognitive heuristics are mental shortcuts people use to cope with the vast amount of information they face all the time in order to make judgments and decisions quickly (Tversky & Kahneman, 1974). These heuristics are associated with fast, automatic cognitive processing. This type of processing is called Type I processing according to a model of cognition called the Dual Processing Model (Evans, 1984). This can be compared to the Type II processing aspect of the model. Type II processing is associated with analytical and deliberate cognitive processing. The fast nature of Type I processing may be beneficial and necessary in situations that require immediate decisions. However, the efficiency of, and lack of detailed attention in this kind of cognitive processing can also lead to errors and biases. For example, you might see a woman at a hospital with an

3

employee badge and scrubs, and make a quick stereotype-based inference that she is a nurse. This may be true and an effective way to make a fast judgment of the situation; however, the efficiency of Type I processes would lead you to err if she was a doctor. Had you used your more deliberative Type II processing, you may have noticed that her scrubs were the same color as the other doctors but not the nurses, analytically concluding that she is also a doctor.

Confirmation bias is one type of heuristic that explains the tendency for people to favor and seek out information that confirms existing beliefs and expectations (Nickerson, 1998). This phenomenon is a natural and automatic part of human processing. Similar to the above example, we rely on easily accessed expectations of a situation in order to make inferences about it. This tendency can be harmless and useful, as using previous information to inform future inferences in similar situations can often speed up and lead to correct judgments much of the time. However, it can also lead people to seek out information that only confirms their beliefs. For example, if someone is opposed to vaccines, they may interpret news stories on vaccines differently than people who support vaccinations, paying more attention to reasons against getting vaccinated. In both cases, we are influenced by preconceptions when we interpret and make decisions on current stimuli.

Bressan and Dal Martello's (2002) study found that when participants were led to believe that the people in two pictures were related, they rated the two people's features more similar compared to subjects that did not have background on whether they were related. Charman, Gregory, and Carlucci (2009) found similar results in legal decisions where mock jurors' perceptions of a defendant's guilt predicted how similar they rated

that defendant to a facial composite (i.e. a sketch of what the person who committed the crime was described to look like).

Additional studies have shown how confirmation bias interferes in numerous other domains such as interrogations (Narchet, Meissner, & Russano, 2011; Hill, Memom, & McGeorge, 2008), forensic analyses (Nakhaeizadeh, Dror, & Morgan, 2014; Dror, Charlton, & Peron, 2006), and eyewitness accounts (Hasel & Kassin, 2009). Overall these studies show similar patterns of people behaving in ways that confirmed expectations and prior beliefs: interrogators changed their tactics when they were primed to believe a suspect was guilty (compared to innocent), forensic analysts came to differing conclusions on information based on the adversarial side by whom they were hired (defense vs. prosecution), and when eyewitnesses were given information suggesting a suspect was guilty they incorrectly changed their identification decision to confirm the new information.

Biases in Sexual Assault Contexts

Several factors have been found to be related to perceptions of sexual assault. One of the most consistent findings in the sexual assault perception literature is the effect of rape myth endorsement on perceptions of sexual assault. Rape myths are prejudicial, stereotyped, and often false beliefs people have about sexual assault, perpetrators, and victims (Burt, 1980). People who endorse these myths are more likely to excuse rape and aggressive behavior, while blaming the victim for what happened. Hammond, Berry, and Rodriguez (2011) showed a significant correlation between rape myth acceptance and guilt attribution. Participants who endorsed more rape myths (e.g. "she asked for it") indicated a lower perception of guilt and were less likely to find the defendant guilty.

An investigation into cases with unanalyzed rape kits showed that police investigations were also influenced by rape myths, and whether the case fit the mold of what "real" rape is (e.g., assailant was a stranger). In the cases reviewed, officers on average performed only 3.4 out of 10 procedurally-expected investigative steps (Shaw et al., 2017; Shaw, Campbell, & Cain, 2016). When providing reasons why the investigation suffered, officers revealed victim blaming attitudes or reasoned that the victim made it difficult to investigate at all. Statistics from the Department of Justice (DOJ; 2012-2016) bring these studies into context. According to DOJ reports, only about 23% of sexual assaults were reported to police. Of those cases reported to police, only about 19% lead to arrest, and only 10% of those arrests lead to a conviction. These statistics suggest that this type of crime may be under-represented in the law, possibly because of the often-perceived ambiguous nature of sexual assaults – a context ripe for bias.

When stimuli is ambiguous, or there is more room for interpretation, heuristics and biases seem to play a larger role in decision processes (Braman & Nelson, 2007; Dunning et al., 1989). People are inclined to pay more attention to confirmatory and easily accessible information, but only to the extent that it is reasonable and justifiable. Lange et al. (2011) found evidence for confirmation bias in how people interpret auditory speech, but only when the recording was low quality, ambiguous, and could be easily misinterpreted (e.g. ripped vs. raped). This has important implications in sexual assault cases where a lot of evidence is "he-said-she-said" and difficult to prove. For example, a majority of victims know their assailant prior to the alleged assault, and if no major physical evidence is present, it might be reasonable for the assailant to claim there was consent.

In light of the recent #MeToo movement and high profile cases of sexual assault and harassment, sexual assault has received increased attention, creating awareness of how widespread the issue is. A common complaint that has not yet been the subject of much experimental inquiry is the perception of systematic covering-up of assaults. For example, Hollywood performers have been accused of running "enterprises," and being protected by their team (e.g., managers, bodyguards, drivers, personal assistants; see Meisner, Buckley, & Crepeau, 2019) and their high-power status (Cooney, 2017). Churches have covered up for sexual assaults within their congregations (Davis, 2002), universities have failed to investigate reports - especially with regard to high-profile athletics programs, and Olympic faculty have been exposed for years of misconduct (Evans, Alesia, & Kwiatkowski, 2016). What these cases have in common is an overall theme of powerful institutions excusing assault and using positions of power to cover up these issues.

Richards (2019) examined data from cases of reported incidents of sexual assault at higher-education institutions, finding that few incidents that were reported to Title IX coordinators lead to formal Title IX complaints. Of those that did, a majority resulted in offering services to the victims, rather than sanctions for the perpetrator. This begs the question of when institutions should be held accountable, and whether people have the same kinds of perceptual and cognitive biases toward institutions' liability as they do for individual perpetrators in criminal sexual assault cases. Institutions have an obligation to protect members from sexual misconduct, but there is flexibility in how institutions create and implement policies. This flexibility, similar to the ambiguity in assault cases

and in accordance to previously cited work, may make these cases especially susceptible to cognitive and personal biases interfering with decisions.

The Current Project

In a series of two studies, we aimed to test confirmation bias in the context of people's perceptions of institutions' liability in sexual assault cases. Expanding on what is known on decision making in sexual assault cases, the studies looked at juror decisions in the context of civil cases against institutions for their role in sexual assault, deciding whether an institution had created a safe environment for members and sufficiently investigated reports. Both studies involved participants reading cases of differing strength (e.g., a strong vs. weak case against the institution), looking at whether participants were able to accurately discern the strength of the civil cases. Confirmation bias was tested through another manipulated variable, having participants read mock news articles that contained statements meant to sway participants' judgments on the civil case. Both studies were preregistered on the Open Science Framework, with all materials and data accessible here: https://osf.io/dzk2h/?view_only=5042982dc1e34821945ad17ae9f63578.

Study 1

This study aimed to test whether participants would be susceptible to confirmation bias when making judgments about a civil case, relying on a mixed within- and between-subjects design. All participants read three case vignettes of differing strength, portraying a strong, weak, or neutral (control) civil case against an institution. The neutral case provided no evidence for or against liability. Prior to reading the strong and weak case vignettes, participants were randomly assigned to either read a confirmation-bias inducing article or a control article. The confirmation-bias inducing

article contained a statement that the case against the institution was strong or weak, consistent with the actual strength of the case. The control article had the same information, but without the statement about the strength of the case and was not meant to bias judgments. In this study, the confirmation bias was always consistent with the case strength; therefore, if participants received an article that contained confirmation bias it was intended to bias their judgments in the "correct" direction, rather than possibly lead to erroneous judgments. Consistent with the sexual assault literature, we also looked at Rape Myth Acceptance to see if these pre-existing attitudes also influenced civil case judgments.

Hypotheses

We pre-registered three main hypotheses for Study 1 on the Open Science Framework. (1) We hypothesized a main effect of the case strength condition, such that participants would overall be more likely to find the institution liable in the "Strong" condition compared to in the "Weak" and "Control" conditions. (2) We also predicted a main effect of the confirmation bias condition, such that participants in the "Strong" condition would be more likely to find the institution liable if they were given an article containing biasing information compared to those given a control article, and participants in the "Weak" condition would be more likely to find the institution not liable if they were given an article containing biasing information compared to those given a control article. And (3) we predicted that participants with high Rape Myth Acceptance would be less likely to find the institution liable, and that Rape Myth Acceptance would interact with case strength such that this effect would be more pronounced in the "Control" condition compared to the "Strong" and "Weak" conditions.

Pilot Study A

The civil cases participants read contained several pieces of case facts that provided evidence for or against the defendant (i.e., the institution). The purpose of this pilot was to test the strong and weak case facts to ensure that they were perceived as intended.

Methods.

Design and procedure. Participants first read a brief summary of the parties involved in the case and the applicable law upon which they were to base their judgements (see Appendix A). Participants were then shown a series of nine case facts and asked to rate each one a 10 point scale (*1 = very weak, 10 = very strong),* with 10 indicating that it provided strong evidence in favor of the Plaintiff's case (see Table 1). People were randomly assigned to view either the strong or weak version of each case fact.

Materials.

Case facts. Nine case facts were designed to be embedded in the case vignettes that would remain neutral and ambiguous, that is not provide strong evidence for or against the defendant. To create these nine facts, when possible we referred to Title VII and Title IX standards and expectations. The nine facts were based on general consensus among universities and the Title IX: Dear Colleague Letter (Ali, 2011) about appropriate investigative procedures. In addition, to get a sense of common themes in these trials and what types of case facts are presented, we referenced eight real case summaries against institutions similar to our study (*Buslon v. Pasl Soccer LLC, 2014; Denson v. The Corporation of the President of the Church of Jesus Christ of Latter-Day Saints, 2018;*

Doe v. Baylor University, 2017; Doe v. The University of North Texas Board of Regents,

2017; Harris v. CBS News Communications Inc., 2018; Hernandez v. Baylor University,

2016; Huang v. St. John's Evangelical Lutheran Church, 2012; and *Langley v. Faulkner*

University, 2012). We used this information to craft the facts of our experimental

vignettes.

Across these different sources, general consensus about appropriate investigative

policies and procedures, upon which our materials are based, included (1) having a

procedure in place or someone to whom members are clearly able to report the assault in

order to get an investigation started, (2) offering counseling services or directing them to

where to seek help, (3) keeping the victim informed about the state of the investigation,

(4) conducting an investigation in an appropriate amount of time (e.g. many universities

state that assault cases should be resolved within 60 days). Other problems that arise in

cases also involve (5) the institution creating a hostile environment by ignoring previous

allegations, (6) hiring someone who had a concerning history of violence or assault, (7)

not thoroughly investigating claims before hiring someone, (8) allowing the assailant to

remain at the institution, or (9) not working with the victim to ensure they would not have

to come into contact with the assailant while at the institution.

We created language to map onto all nine of these elements for the vignettes to be

strong and weak. The weaker case strength condition vignettes had versions of the case

facts that provided a weak case against the institution (favored the defendant) and the

strong case strength vignettes had versions that provided a strong case against the

institution (favored the plaintiff). For example, the evidentiary statement in the "strong"

condition vignette for whether an institution has a proper protocol or person in charge of

11

handling the investigation was: "The defendant did not have a set procedure in place for handling allegations, so it was not clear to the plaintiff who she was supposed to report to, or who was in charge of handling the investigation." The corresponding evidentiary statement in the "weak" condition was: "The defendant directed the plaintiff to their office where someone in charge of handling allegations would be able to formally handle the report."

Participants. Participants ($N = 134$; Age: 20-57, $M_{Age} = 26.2$; 78% female) were undergraduate students collected from an Arizona State University online forensic psychology course. They received extra credit for their participation and took on average 44 minutes to complete the study.

Results and discussion. Overall, the nine strong case fact statements were perceived as relatively strong ($M = 7.78$, SD = 2.44; see Table 2 for descriptive statistics for each case fact), and had left skewed distributions (Figure 1). The weak case fact statements, however, were perceived as relatively neutral ($M = 5.64$, SD = 2.83); see Table 2), and were somewhat uniformly distributed (Figure 1). The strong case facts were overall perceived as significantly stronger than the weak case facts, $t(1167.2) = 14.08$, p < .001. Based on these findings we kept the strong statements and slightly modified some of the weak statements in an effort to make then even weaker, and then moved on with Study 1.

Table 1.

Descriptive Statistics for Pilot A Case Facts.

Case Facts	*M*	*SD*
Strong Case Facts		
1. The defendant did not have a set procedure in place for handling allegations, so it was not clear to the plaintiff who she was supposed to report to, or who was in charge of handling the investigation.	7.47	2.26
2. Plaintiff was not made aware of any counseling services available.	7.23	2.91
3. After a couple weeks, she had not been made aware of any updates in the investigation. She had tried to contact them but could not get a hold of anyone. After a week she still had not heard back from anyone.	7.8	2.3
4. After 148 days, the institution had not concluded the investigation, nor had they indicated whether it was still under review.	8.34	2.24
5. There had been previous complaints of harassment at the institution, but no investigations were initiated for any of the complaints.	8.48	2.17
6. There were suspicious prior allegations against the alleged assailant for acting violently towards others.	8.47	1.77
7. The defendant claims to have performed a brief investigation into the alleged assailant's history based on previous concerns, but did not further investigate any alarming findings that they were made aware of during the brief investigation.	7.32	2.41
8. Following the report and allegations, the alleged assailant was able to return full time.	7.32	2.72
9. The institution offered no accommodations to the plaintiff to ensure she would have no contact with the alleged assailant.	7.67	2.66
Weak Case Facts	*M*	*SD*
1. The defendant directed the plaintiff to their office where someone in charge of handling allegations would be able to formally handle the report.	6.12	2.71
2. Plaintiff was given contact information for counseling services that the institution compensates members for in circumstances of stress and trauma.	6.31	2.87
3. The plaintiff was given several updates throughout the investigation and was told immediately when the case had concluded.	5.46	3.01
4. After 14 days the institution had concluded the investigation and notified the Plaintiff.	5.53	2.46
5. This seemed to be the first complaint of harassment at the institution.	4.71	2.83

6. The alleged assailant had no prior allegations against him.	4.74	2.87
7. The defendant had performed a brief investigation into the assailant's history, and the alleged assailant had a clean history.	5.43	2.59
8. Following the investigation, the assailant was suspended indefinitely.	5.97	3.02
9. The institution enacted a no contact policy so that the plaintiff would not have any form of contact with the assailant should he be able to return.	6.56	2.56

Note. The mean and standard deviation of case facts (1 = *very weak,* 10 = *very strong*).
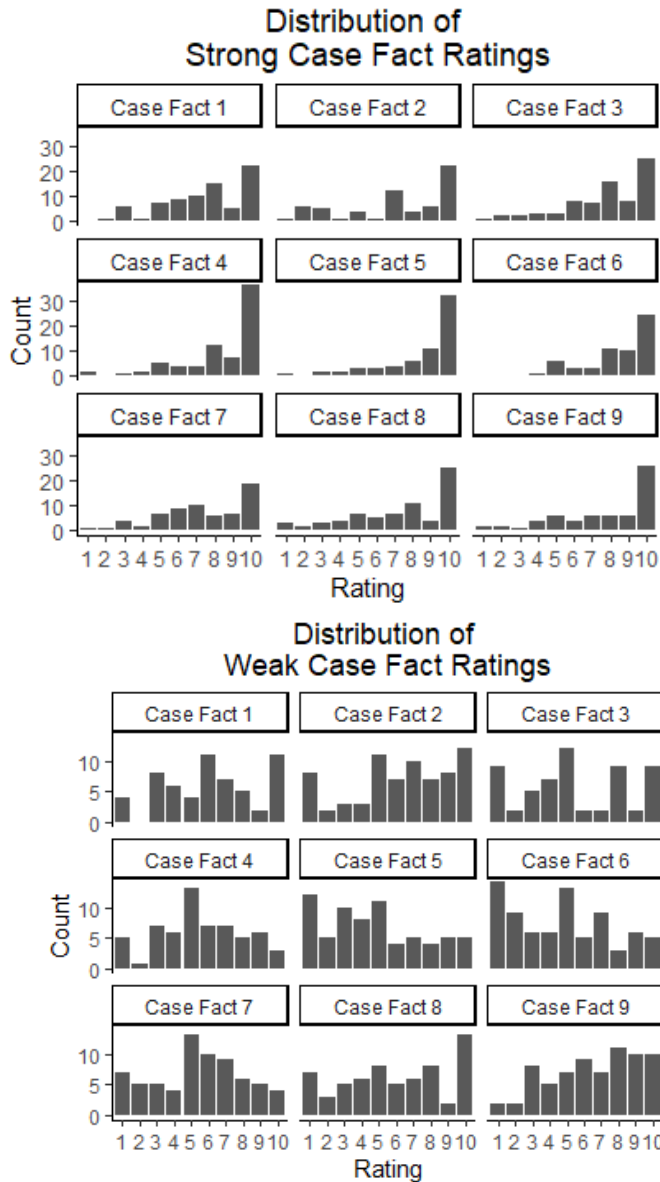

Figure 1. Count Distribution of Ratings for Pilot A Case Facts.

14

Study 1 Methods

    Participants. Our intended sample size (260) was determined by conducting a power analysis for repeated measures, within-between interactions, with enough power to detect an estimated effect size of Cohen's f .10, with an α of .05, and 80% power. Participants ($N = 278$, $M_{Age} = 23.74$, 63% Female) were 18 years or older, and had never been convicted of a felony, which are some of the requirements for serving on a jury. Due to faculty concern of depriving certain students from the opportunity to receive extra credit, we were not able to limit our sample to those who were U.S. citizens in order to fully satisfy the requirements of a jury-eligible sample. Participants were recruited through Arizona State University's West Campus SONA system ($n = 188$), receiving extra credit for their participation; and Prolific Academic ($n = 90$), an online survey platform where participants received $6.50/hour for their participation. People took on average 44 minutes to complete the study.

    While payment rates through Prolific Academic are higher than other online platforms (e.g., Amazon's Mechanical Turk (MTurk)), there is evidence suggesting that offering higher rewards incentivizes participants to put in more effort (Lovett, Bajaba, Lovett, & Simmering, 2018). Additionally, using Prolific, compared to more-often used platforms such as MTurk, offers additional benefits of having more naïve and diverse participants (Peer, Samat, Brandimarte, & Acquisti, 2017).

    Eight participants from the Prolific Academic sample were excluded for failing at least one of the two attention checks assessing whether jurors were adequately reading through the questions and study. The attention checks were multiple choice questions asking participants to select a specific answer, to ensure they were paying attention.

Participants took on average 44 minutes to complete they survey. They self-identified as 57.2% White, 12.2% Hispanic (white), 11.1% Hispanic (non-white), 6.1% African, 5.4% mixed, 2.9% as other, 2.1% Asian, 0.3% Native American, and 0.3% Pacific Islander. Participants' political views averaged 3.29 on a scale from 1 (Strongly Liberal) to 7 (Strongly Conservative).

Design. Study 1 was a 3 (case strength: strong vs. weak. vs. neutral) x 3 (confirmation bias information: strong vs. weak vs. control), mixed factor, Latin square design (see Figure 2). The within subject element of this study was the strength of the case. The strength of the case referred to whether the case facts were indicative of a weak or strong case against the institution. The neutral case was intended to act in part as a control, where participants received an ambiguous case with distractor information (e.g., the victim's age), and only general information on the plaintiff's experience reporting the alleged assault to the institution (e.g., how the institution questioned the plaintiff and commenced an investigation, without going into detail about the investigation and subsequent proceedings). Participants read and answered questions related to all three case strength conditions, in a counterbalanced order. The confirmation bias statements were between subjects, and were embedded in articles that participants read prior to the cases (more information to follow).
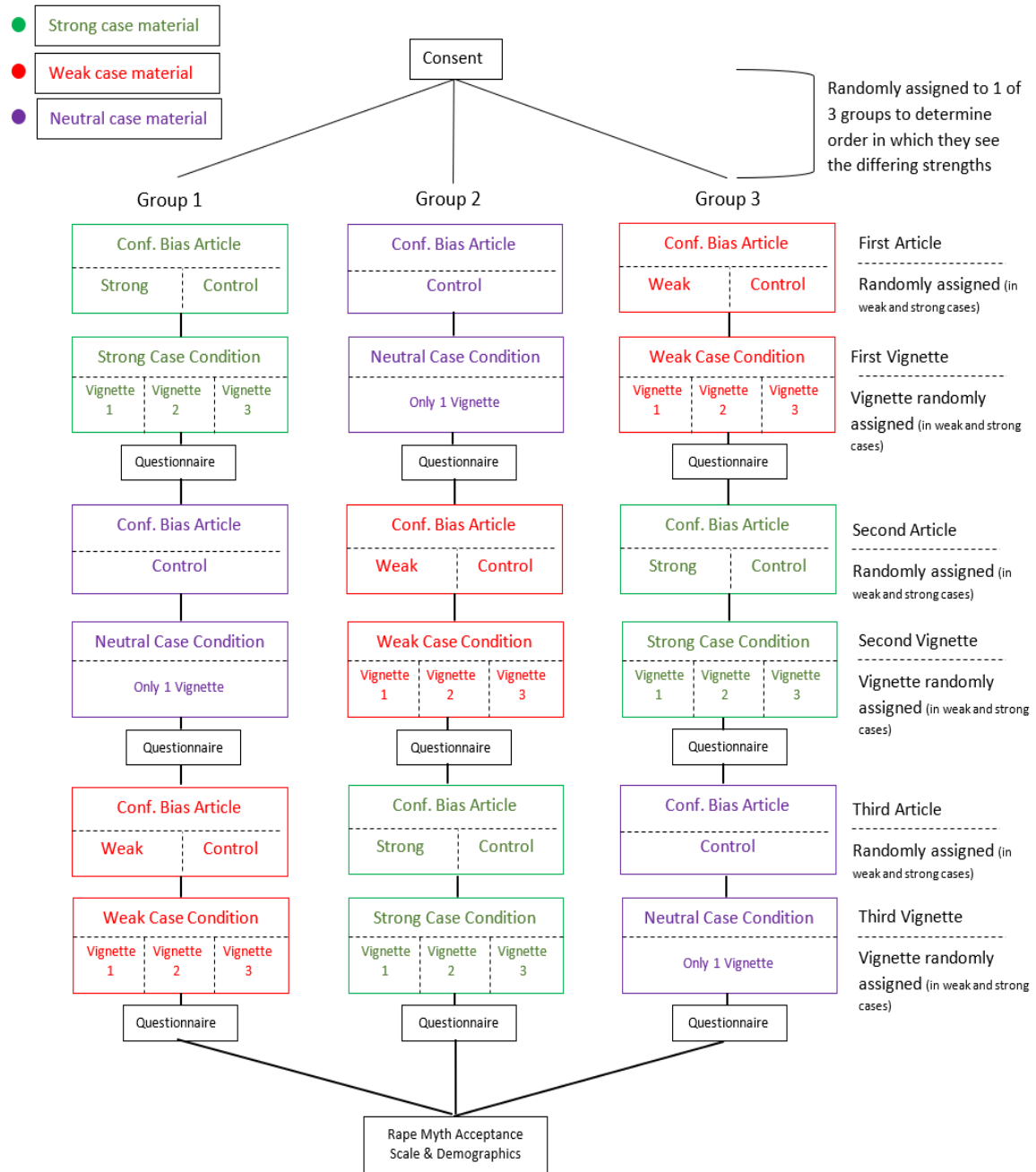
Figure 2. Flowchart depicting Study 1 procedure. Participants read an article, a case, and completed a questionnaire three times. The three columns indicate whether they were in Group 1, Group 2, or Group 3; and the colors indicate the case strength condition they were in at that time (counterbalanced). The right labels indicate the randomization of the article they saw (control vs. confirmation bias) and the vignette randomization for the

weak and strong cases (the neutral condition always utilized the control articles, and only had one vignette option).

Procedure. Participants read a consent form before beginning the study. Participants read three case vignettes, one for each case strength condition (strong, weak, and neutral). They were randomly assigned to one of the three groups, counter-balancing the order in which they saw the three case strength conditions (Group 1: strong, weak, neutral; Group 2: weak, neutral, strong; Group 3: neutral, strong, weak).

Prior to reading each case vignette, they read a mock news article containing a statement with confirmation bias (strong, weak, or control). In the strong and weak cases participants were randomly assigned to either read an article containing confirmation bias that corresponded to the condition (weak article prior to the weak case, strong article prior to the strong case), or the control article. Prior to the control vignette, participants always read an article that contained no confirmation bias (control). They then read the case vignette, and then filled out a questionnaire for the dependent measures. This process was completed a total of three times. After the participant read and completed the third and final vignette and questionnaire, they filled out the Illinois Rape Myth Acceptance Scale and demographic questions.

Materials.

Vignette. Participants read one-page vignettes describing a civil case against an institution for their role in a sexual assault that happened within their institution. The negligence at question related to an institution's obligation to provide a safe environment for members, including appropriately overseeing employees and effectively following up with any reports of misconduct. The case focused on the institution's policies (e.g., who

handles reports and allegations, hiring protocols) and procedures following the report (e.g., how long the investigation took). To focus on the institution's negligence rather than the assault itself, minimal details about the assault were provided, and it was not explicitly stated whether the alleged assailant was tried or found guilty.

The case presented evidence regarding policies and procedures in the context of four possible institutions: Church, University, Soccer Association, or Hollywood Network. These institutions were chosen to represent four domains that have received attention in the media for their involvement in sexual assault allegations. The type of institution in which the study material for each vignette was embedded was randomly assigned, as the primary concern was not differences between institutions, but rather if results would generalize across institutions. The names of all the institutions were created using random name and letter sequence generators, and do not reflect the names of any actual institutions (see Appendix B1).

Each vignette began with a general introduction of the case describing what the participant must decide and the specific rules that may be applied to the case in order to determine negligence and liability. Following the introduction was the body of the case. The body of the case consisted of case facts as well as randomly piped in details (e.g., Plaintiff name, date and location of alleged assault; see Appendix B1). The piped in details served to provide background details of the case and distract participants from the actual purpose of the study. These randomized details did not provide any evidence for or against the institution. The substance of the vignettes were a randomized selection of the pilot-tested case facts that provided evidence for or against the institution.

Beginning Instructions. The vignette began by providing an overview of negligence and what must be proven in order to find the institution liable. In order to prove negligence a plaintiff must prove the following are more likely to be true than not true: 1. Defendant was negligent; 2. Plaintiff was injured; 3. Defendant's negligence caused plaintiff's injuries. Negligence can be proved in various ways. We included six examples of negligence (e.g., "the Defendant failed to promptly and appropriately investigate and respond to the reported assault;" "the Defendant committed fraudulent concealment or nondisclosure"), stating that they must prove that one or more are true.

Case Facts. Each vignette in the strong and weak case conditions had five key case facts regarding the institution's actions/policies, randomly chosen from the nine possible case facts of corresponding strength: strong and weak. The control case strength vignette did not contain any of these useful case facts, instead it contained only distractor information (i.e., the institution and the plaintiff's relationship to the alleged assailant). (see Appendix B2 for case fact versions).

There were a total of seven different vignette versions across the three conditions (see Figure 3 for what case facts the seven vignette options were comprised of). In the neutral condition, there was only one control vignette option since this condition's vignette did not contain any of the case facts. There were three strong and three weak condition vignettes options, each containing a different combination of five of the nine possible case facts. Participants were randomly assigned to one vignette option for both of these case strength conditions. This procedure was designed to ensure each case remained similar in the type of evidence they contained, while minimizing possible carryover effects and contrast effects. By creating three options for them to see at each of

20

the three different case strength vignettes, it was unlikely that they would see the same combination of facts. In addition to this procedural design, variables such as the institution, Plaintiff's name, and age were randomly assigned as distractors for all three case strength condition vignettes.

Strong Case Condition

Randomly assigned to view

5 case facts (out of the 9 possible) randomly grouped to form 3 vignette options.

| Vignette 1 | Vignette 2 | Vignette 3 |
| Case Facts: 4, 5, 6, 7, 9 | Case Facts: 3, 4, 5, 7, 8 | Case Facts: 1, 2, 6, 7, 8 |

Neutral Case Condition

| Vignette [1 option] |
| Case Facts: none |

Weak Case Condition

Randomly assigned to view

| Vignette 1 | Vignette 2 | Vignette 3 |
| Case Facts: 4, 5, 6, 7, 9 | Case Facts: 3, 4, 5, 7, 8 | Case Facts: 1, 2, 6, 7, 8 |

Figure 3. Flowchart depicting Study 1 vignette design. Process of how each of the seven vignette options were designed and assigned to the Neutral, Strong, and Weak conditions. The diamond indicates the case strength condition. For each condition that participants were in, they read only one total vignette, randomly assigned to them through Qualtrics functions. Following the arrows down, the strong and weak conditions had three possible vignettes that participants could see. In order to determine what case facts would be used in each of the case vignettes I used randomly number generators. Using the random number generator, I first generated 6 groups of five case facts (out of the nine total case facts), so that each vignette in the strong and weak condition's had five total case facts. The six groups that were generated are represented by the bottom ovals (note the neutral only had one option that included no case facts so did not require a randomization process or multiple vignette options). Then they were randomly assigned to a condition (strong or weak) so that each condition had three vignette options that were comprised of these randomized case fact combinations. The bottom row indicates the case facts that were in each vignette. For example, vignette 1 in the strong condition has case facts 4, 5, 6, 7, and 9. Refer to Appendix B2 for what case fact corresponds with what number in each of the conditions.

22

Randomized Variables. Each vignette contained distractor variables related to the case that were randomly embedded into the body of the vignette. These variables served to make the cases somewhat unique so that participants would not see the exact same information throughout the three case strength conditions while reducing confounds. The variables that were randomly embedded included: the plaintiff's name and age, the state where the alleged assault occurred, the date of the alleged assault, the alleged assailant's name, how long it took the plaintiff to report the alleged assault to the institution, the plaintiff's relationship to the defendant, who the plaintiff reported the alleged assault to, and the alleged assailant's affiliation at the institution. For each of these variables, there were three to five options (see Appendix B1 for a list of options for each of the variables). Participants were randomly assigned to see one of the options. For example, a participant could be randomly given the Plaintiff name "Tammy Moore" in the strong condition, and the name "Rachel Hill" in the weak condition; whereas another participant may be randomly assigned the Plaintiff name "Rachel Hill" in the strong condition, and "Tammy Moore" in the weak condition.

The plaintiff's name was randomly piped in from a list of five possible female names, taken from a random female name generator. The plaintiff's age was randomly piped in and ranged from 18 to 25. The state locations of the alleged assault were randomly piped in from a list of four possible states, picked to represent a variety of political ideological preferences (e.g., conservative vs. liberal vs. middle-ground). The dates (month, day, year) of the alleged assault were randomly piped in from a list of five dates, formulated from a random date generator, making the year somewhat recent. How long it took the plaintiff to report the alleged assault was randomly piped in ranging from

23

1-4 days, or a week. The plaintiff's relationship to the assailant (i.e. how well they knew each other and the extent of their interactions) were randomly piped in from a list of four possible options. Who the plaintiff reported to and the assailant's affiliation were randomly piped in from a list of three options. To remain consistent with the structure of certain institutions, the options for these two variables differed based on the institution. For example, if a participant was given a case against the church, the plaintiff may have reported to a priest, whereas against the university they may have reported to a professor. Similarly, in a case against the church, the alleged assailant may have been an altar server, whereas against the university the alleged assailant may have been a student.

These details provide information about the case and incident itself, but do not provide any concrete evidence for or against the institution. While factors such as how long it took the plaintiff to report an assault or how well they knew the alleged assailant may impact judgements towards the assault itself (Spohn & Tellis, 2012), they do not have any relevance to how the institution handled the matter, and should not therefore play a role in jurors' judgements of these particular cases against the institutions. See Appendix B3 for a full example of a vignette.

Confirmation Bias Articles. The confirmation bias statements were embedded in a mock news article. The news article was a brief story related to an event inspired by the #MeToo movement. The confirmation bias statements at the end of the articles were intended to sway the participant's perception of the actual case they subsequently read to be strong or weak. For example, the strong confirmation statement was, "Compared to other similar cases, people at this rally agreed that the upcoming civil lawsuit seemed to favor the plaintiff, as people thought the institution seemed to have done little to act in a

timely and fair manner in their investigations to accommodate the parties involved." The

corresponding weak confirmation statement was, "Compared to other similar cases,

people at this rally agreed that the upcoming civil lawsuit seemed to favor the defendant,

as people seemed to think the institution did everything it could to act in a timely and fair

manner in their investigations to accommodate the parties involved." In the control

article, there was no information that was intended to skew the participant's views of the

case. See Appendix B4 for the articles participants read.

Measures. (Appendix B5)

Dependent Measures Questionnaire. We constructed a series of questionnaires to

assess perception of liability, measured dichotomously (liable vs. not liable) and along a

7 point Likert scale (1 = *not at all liable*, 7 = *extremely liable*); the strength of each

parties' case, asked for defendant and plaintiff along a 10 point scale (1 = *very weak*, 10 =

very strong); confidence in decision (1 = *not at all confident*, 10 = *very confident*); and an

open-ended brief justification for their decision. We also asked participants how likely

they think a similar case would happen again at that institution along a 7 point Likert

scale (1 = *very unlikely*, 7 = *very likely*). They were also asked to rate the strength of each

of the case facts (*1 = very weak, 10 = very strong*), or indicate that they did not receive

that case fact.

Trust in Institutions. To measure participants' general trust in the institutions, we

selected five items from a previous measure of trust in institutions (Hamm et al., 2013;

PytlikZillig et al. 2016). The items asked participants the extent to which they

agreed/disagreed with statements related to the institution's actions and ethics (e.g., "The

procedures followed by this type of institution are ethical"). These were rated on a 7-

point scale (1 = *strongly disagree*, 7 = *strong agree*), with one item reverse scored ("This type of institution is overly influenced by special interest groups"). Using Cronbach's alpha test for reliability, the scale had good internal consistency, α = .82.

Rape Myth Acceptance. The Updated Illinois Rape Myth Acceptance Scale (IRMA) was used to assess participants' attitudes and beliefs about rape, specifically whether they tend to endorse rape myths. The updated version of the Illinois Rape Myth Acceptance Scale has good internal consistency (Cronbach's = .87; McMahon & Farmer, 2011; in the current study α = .93.). The scale consists of 22 items, rated on a five-point Likert scale (1 = *strongly agree*, 5 = *strongly disagree*). Participants rated their agreement with a series of statements such as, "If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand," and, "If a girl doesn't physically fight back, you can't really say it was rape." Higher scores indicated greater rejection of rape myths. For the purpose of analysis interpretation, we reverse scored each item so that higher scores indicated greater rape myth acceptance. We then calculated each participant's mean score to use for analyses.

Study 1 Results

As part of our design, participants were randomly shown one of three possible vignettes to read for the strong case, and one of three possible vignettes for the weak. For example, one participant in the strong condition could read a vignette comprised of the strong case facts: 4, 5, 6, 7, and 9; whereas another participant could read the strong vignette that comprised of the strong case facts: 1, 2, 6, 7, and 8. The three vignette options in the strong condition were all intended to depict a strong case against the institution (rendering more liable verdicts); and the weak all intended to depict a weak

case against the institution. However, since within the conditions, the three vignette options did include a different combination of case facts, we performed chi-square tests of independence to test whether within the strong and weak case conditions, certain vignettes/combinations of facts were perceived as stronger or weaker than others (i.e., were people more likely to find the institution liable in the strong condition if they were given vignette option 1 vs 2 vs 3). Across both the strong and weak conditions, the chi-square tests were not statistically significant, $x^2(2, 277) = 4.52$, $p = .10$, $x^2(2, 277) = 1.00$, $p = .61$, respectively, allowing us to conclude that there was not a statistically significant difference in verdict between the three vignette options in either condition. That is, within both the strong and weak case the three vignette options were not significantly different, suggesting the combinations of case facts were all perceived as similar in strength when it came to verdict. See Table 2 for descriptive statistics for the case facts used in Study 1.

Table 2.

Descriptive Statistics for Study 1 and Study 2 Case Facts.

Case Fact	Strong Case				Neutral Case				Weak Case			
	Study 1		Study 2		Study 1		Study 2		Study 1		Study 2	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	5.85	3.13	6.63	3.07	-	-	5.47	2.75	5.06	2.79	4.46	2.93
2	5.49	3.40	6.45	3.32	-	-	4.93	2.79	4.99	2.81	4.43	2.94
3	5.81	3.45	6.98	3.23	-	-	4.98	2.49	4.99	3.09	4.22	3.26
4	5.98	3.47	7.01	3.19	-	-	5.34	2.77	4.88	2.93	4.36	3.14

5	6.90	2.99	7.29	2.84	-	-	5.36	2.52	4.42	2.42	4.04	2.93
6	6.67	2.94	7.50	2.76	-	-	5.76	2.70	4.12	2.99	3.20	2.86
7	5.80	3.08	6.84	2.88	-	-	5.06	2.65	4.50	2.89	3.90	2.96
8	7.04	3.03	7.41	2.98	-	-	4.87	3.05	4.94	3.05	4.13	3.14
9	6.09	3.44	7.16	3.21	-	-	4.87	3.01	5.28	3.30	4.44	3.45

Note. Mean and standard deviations for the case facts rated in Study 1 and Study 2. They were rated 1 = very weak, 10 = very strong; or participants indicated "I was not given this information." The neutral condition in Study 1 did not contain case facts so we did not provide descriptives for this group. See Appendix B2 and C1 for which case fact corresponds to each number in each group.

To test the main effect of case strength, we modeled a within-subject logistic regression, with verdict (0 = Not Liable, 1 = Liable) regressed on case strength condition (reference group = neutral condition). Running the model for Type II Wald chi-square effects, there was a significant main effect, $x^2(2, 277) = 62.08$, $p < .001$. In partial support of our hypothesis, participants were significantly more likely to vote liable in the strong condition (74.4% voted liable), $b = 1.62$, $SE = 0.21$, $p < .001$, compared to the neutral condition (40.6%). However, contrary to hypotheses, in the weak condition participants were also significantly more likely to vote liable (54% voted liable), $b = 0.63$, $SE = 0.18$, $p < .001$, compared to the neutral condition (40.6%) (see Figure 4). For these and subsequent logistic regression results, the beta refers to the log odds of choosing a 1 (liable). Converting to odds ratio, participants were about 4.9 times more likely to return liable verdicts in the strong case compared to the neutral case, and 1.9 times more likely in the weak case compared to the neutral. These results were confirmed with an ordinal regression looking at participants' responses on the 7-point Likert

responses to the extent of liability across conditions, such that participants rated the institution as more liable ($M = 5.08$) in the strong case, $b = 1.48$, $SE = 0.16$, $p < .001$, compared to in the control case ($M = 3.73$). However, there was no significant difference between the weak ($M = 3.56$) and control case on extent of liability, $b = -0.19$, $SE = .15$, $p = .20$.
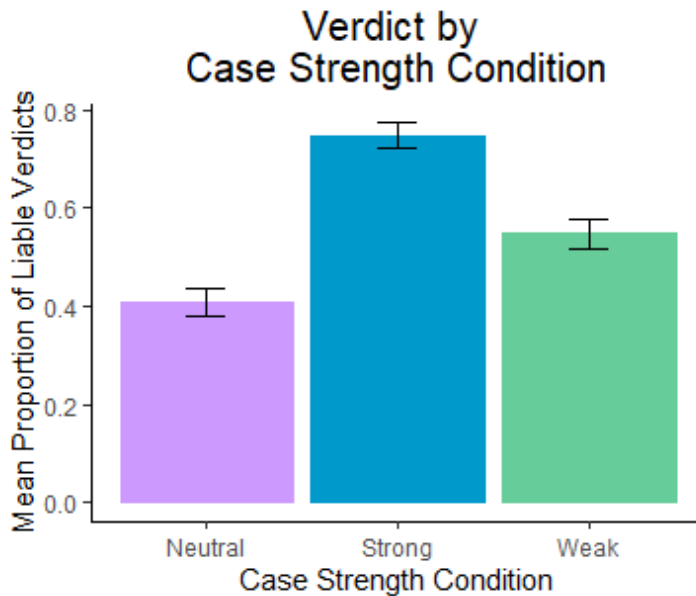
Figure 4. Study 1 Verdict by Case Strength Condition. Error bars represent standard errors. There were significant differences between all conditions at $p < .001$.

Given the within subject Latin-square design, we also tested for order effects, to see if the effect of case strength on verdict differed based on the order in which they saw the conditions. To test for this we included the Group number they were in (which determined the order they saw the cases) to the logistic regression model as an interaction term. To get an overall model effect, we ran the model for Wald chi-square effects. There

was no significant interaction between the group they were in and the case strength condition predicting verdict, $x^2(4, 277) = 7.15$, $p = .13$.

Next, we assessed whether the confirmation bias articles participants were exposed to (i.e., whether they received confirmation bias vs the control) influenced their judgements of the cases, and whether this differed by case strength. Due to the design of the study, and the fact that participants in the neutral case condition did not receive different confirmation bias articles, and those in the strong and weak case strength conditions received different confirmation bias statements (strong vs. weak, respectively), we did not include the neutral case strength in the analyses. Instead, we tested for the effects of confirmation bias by performing two separate chi-square tests in the strong and weak case strength conditions, as these were the only two conditions that would allow for an effect. Contrary to our hypothesis, there was no effect of confirmation bias in either the strong, $x^2(1, 277) = 0.87$, $p = 35$, or weak, $x^2(1, 277) = 0.38$, $p = .54$, case strength conditions. That is, verdicts in these conditions did not significantly differ based on whether they received confirmation bias or not.

Next, we explored the associations of Rape Myth Acceptance and verdict. Overall, participants tended to have low average rape myth acceptance scores ($M = 1.70$), with a one-way ANOVA revealing that males tended to endorse rape myths ($M = 2.03$) significantly more than females ($M = 1.52$), $F(1, 266) = 43.56$, $p < .001$, $\eta_p^2 = .14$. We then conducted a logistic regression with verdict regressed on participants' mean centered Rape Myth Acceptance scores. There was a significant main effect of rape myth acceptance, $R^2 = .01$, $b = -0.35$, $SE = 0.12$, $p = .003$, where participants who endorsed more rape myths were less likely to find the institution liable. There was also an

30

interaction, such that RMA remained a significant predictor of verdict in the strong case, $R^2 = .04$, $b = -0.57$, $SE = 0.21$, p = .01, but not in the neutral and weak cases, $R^2 = .01$, $b = -0.26$, $SE = 0.20$, $p = .19$, $R^2 = .01$, $b = -0.30$, $SE = 0.19$, $p = .11$, respectively (see Figure 5). Performing two separate ANOVAs, Rape Myth Acceptance scores did not significantly differ between those who received a strong vs. control confirmation bias article in the strong case strength condition, $F(1, 267) = 0$, $p = .99$, or based on the confirmation article they saw in the weak case, $F(1, 267) = .001$, $p = .97$.
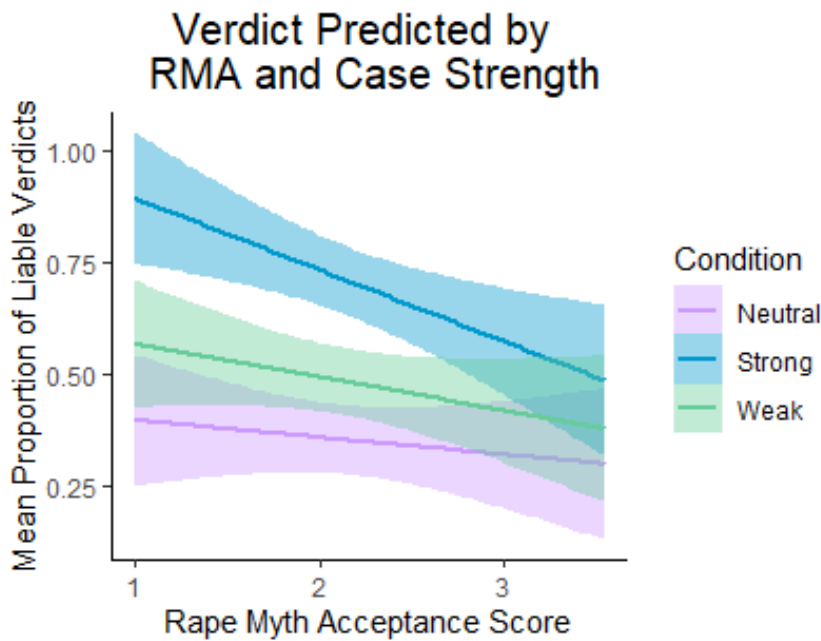
Figure 5. The proportion of liable verdicts predicted by participant's mean rape myth acceptance (RMA) scores in each case strength condition. Proportion of liable verdicts increases along the y axis. Higher scores on the RMA indicate greater acceptance of rape myths.

Exploratory Analyses. To test the effect of gender on verdict, we conducted a logistic regression. There was a significant main effect of gender, with males overall less likely to find the institution liable (48% males voted liable), compared to females (61%),

$R^2 = .02$, $b = -0.55$, SE $= 0.16$, $p < .001$. This remained true across all case strength conditions, as Wald chi-square results indicated there was no significant interaction, $x^2(2, 277) = 0.35$, $p = 84$. This was also not mediated by the fact that males tended to endorse rape myths more than females, as a mediation analysis suggested there was no significant indirect effect, $b = 0.07$, $SE = .04$, $p = .06$.

Performing a logistic regression, with condition and type of institution interacting to predict verdict, and running the models for overall Wald chi-square effects, the effect of condition did not significantly differ based on the type of institution they received, $x^2(6, 277) = 11.38$, $p = .08$. A logistic regression with condition and trust in institutions interacting to predict verdict, revealed there was also no main effect of trust in institutions, $x^2(1, 277) = 0.47$, $p = .50$. In addition, a two-way ANOVA test with condition and article interacting to predict confidence, revealed that participants' confidence in their verdict did not significantly differ by case strength condition, $F(2, 277) = 1.36$, $p = .26$, $\eta_p^2 = .01$, nor the type of article they received (confirmation bias vs. control), $F(2, 277) = 0.51$, $p = .60$, $\eta_p^2 = .001$.

We also tested for demographic differences between the Prolific Academic subjects and the Arizona State University student subjects (see Table 3 for demographic information broken down by sample source). While there were differences in some demographics between the two samples, performing a logistic regression, with case strength and sample source (i.e., Prolific Academic vs. ASU) predicting verdict, the effects of case strength on verdict did not differ significantly between the two samples, $x^2(2, 277) = 1.62$, $p = .45$. In addition, entering the demographic variables into the model as predictors, and running for Wald chi-square effects of interactions, the effect of case

32

strength on verdict did not differ based on age, $x^2(2, 277) = 3.33$, $p = .19$; race, $x^2(32, 277) = 5.30$, $p = .86$; or political ideology, $x^2(2, 277) = 1.68$, $p = .43$. However, there was a small main effect of political ideology, such that participants who endorsed more conservative views were overall less likely to find the institution liable, $R^2 = .01$, $b = -0.10$, $SE = 0.05$, $p = .04$.

Table 3.

Study 1 Demographics

Characteristic	Prolific Academic (n = 90)			Arizona State University (n = 188)		
	n	*M(SD)*	%	*n*	*M(SD)*	%
Gender***	90			183		
Female	44		48.9%	131		71.6%
Male	46		51.1%	52		28.4%
Race*	90			188		
African American	2		2.2 %	4		2.1%
Asian	4		4.4 %	13		6.9%
Hispanic (non-white)	2		2.2 %	29		15.4%
Hispanic (white)	6		6.7 %	28		14.9%
Mixed	3		3.3 %	12		6.4%
Native American	0		0%	1		.53%
Other	3		3.3 %	5		2.6%
Pacific Islander	1		1.1%	0		0%
White	68		75.6%	91		48.4%
NA	1		1.1%	5		2.6%
Age***	90	28.90 (9.78)		181	21.18 (5.06)	
Political Ideology	89	3.11 (1.56)		180	3.38 (1.52)	

33

RMA	89	1.79	180	1.65
		(.66)		(.65)

Note. The mean, standard deviations, and percentages of demographic characteristics between the Prolific Academic and ASU samples. With statistical significance indicating whether the two samples differed from each other for that characteristic. *p < .05, ***p < .001.

Study 1 Discussion

Consistent with hypotheses, jurors were overall able to accurately differentiate between weak and strong cases. Specifically, participants seemed to overall find institutions rightfully liable in cases that provided strong evidence against them. Compared to the strong evidence participants tended to not hold institutions liable when there was weak evidence against them; however, there were still a large number of people who did hold the institution liable when given weak evidence (54%). This may be due to people being sensitive to the topic of sexual assault and not wanting to seem too lenient in these cases. This suggests that jurors may be able to accurately differentiate between when there is strong evidence against an institution, or weak evidence against them, but still seem to overall lean towards finding an institution liable.

Contrary to hypotheses and findings from previous literature (Ruva & LeVasseu, 2012), participants did not seem to be influenced by outside information that could lead to confirmation bias. These findings have positive implications if jurors are in fact able to judge strong vs weak cases appropriately, returning verdicts in accordance to the strength of the evidence and not being unjustifiably motivated by outside sources and opinions.

While participants seemed to be protected against the effects of confirmation bias (i.e., information from mock news articles), they tended to be influenced by their own personal attitudes. This may pose an even larger potential threat to the justice system, as

34

it can be difficult to glean insight into, much less try to change, the pre-existing beliefs people hold. The effects of rape myth acceptance, and exploratory effects of gender, suggest that those who endorse rape myths, and men, tend to be less likely to find an institution liable, even when the case provides strong evidence against an institution. This finding calls into question how courts may address and account for these variables.

Limitations. While Study 1 provided some insight into what may be influencing verdicts in these types of cases, there were limitations. The null effects of confirmation bias could be due to the fact that the confirmation bias was always meant to sway participants to the "correct" verdict. If a participant was shown a strong case, we would expect (and hope) that they would lean towards a liable verdict, and thus being shown that confirmation bias statement would not have changed their verdict to show an effect. Instead, it would have further pushed them towards that same liable verdict. Further, since the confirmation bias was never meant to sway judgments to be inconsistent with the actual strength of the case, it was only able to provide evidence that confirmation bias could improve decisions, not whether it could lead to flawed judgments. For example, in the strong condition, it was unclear whether jurors would be impacted by information that contradicts the correct decision (liable). If participants read an article that was meant to sway participants to view a case as pro-defense (i.e., a weak case against an institution), prior to a strong case, would participants be influenced by this outside information and erroneously vote not liable?

Another limitation was with regard to the neutral case. It was perceived as weaker than both the strong and weak case; instead of being an ambiguous, middle-ground case as intended. This could be due to the lack of evidence for participants to base their

35

decisions on. It is possible that providing participants with no useful case facts may have caused them to feel there was no rational reason to find the institution liable. Therefore, based on the "innocent until proven guilty" framework they returned "not liable" verdicts. It is also important to note that while participants were given no relevant case facts, people in the neutral condition still often rated the case facts as if they were given that information, instead of choosing the appropriate "I was not given this information" response. This could be due to the question not being phrased clearly enough, or that participants were not reading the case closely enough to pick up on case details, the way we would have hoped. To address these limitations, we designed a follow-up Study 2.

<p align="center">**Study 2**</p>

The aim of Study 2 to improve upon the previous design. Study 2 used similar methods as Study 1, but instead of the neutral case strength condition having no case facts, we provided participants with neutral modifications of the case facts to provide a more ambiguous case for participants to judge. In addition, they read either the strong, weak, or control confirmation bias article, counterbalanced prior to each case. Therefore, in this study the confirmation bias could also contradict the actual strength of the case (e.g., a confirmation bias suggesting a weak case, followed by the strong case vignette), possibly resulting in more biased decisions in these instances.

Design

Study 2 was a 3 (case strength: strong vs. weak vs. neutral) x 3 (confirmation bias information: strong vs weak vs control), mixed factor, Latin square design (see Figure 6). The within subject element of this study was the strength of the case – whether the case facts/evidence were indicative of a weak case against the institution, a strong case, or a

neutral case. Participants saw and responded to all three case strength conditions in counterbalanced order: participants were randomly assigned to one of three groups to determine the order in which they saw the three different case vignettes.

	Time 1		Time 2		Time 3	
	Case Strength	**Article Confirmation**	**Case Strength**	**Article Confirmation**	**Case Strength**	**Article Confirmation**
Group 1	Strong	Strong	Neutral	Weak	Weak	Control
Group 2	Neutral	Control	Weak	Strong	Strong	Weak
Group 3	Weak	Weak	Strong	Control	Neutral	Strong

F*igure 6*. Study 2 design. Participants were randomly assigned to one of three groups counterbalancing the order they saw the conditions. Each participant saw all 3 case strengths counterbalanced. The article containing confirmation bias information that they read before each case was also counterbalanced.

We also manipulated the confirmation bias information participants saw in the articles. The articles contained a statement that was meant to sway the participant to view the case as strong or weak; or they saw a control article that contained no confirmation related information. Contrary to Study 1, where participants could only see a strong or control article prior to the strong case, a control article prior to the neutral case, and a weak or control article prior to the weak case, participants in Study 2 could see a strong, weak, or control article prior to all three case strength conditions. The confirmation bias article they saw before each one differed based on the group they were randomly assigned to.

In Group 1, participants read the strong confirmation bias article prior to the strong case strength vignette, a weak article prior to the neutral case, and the control article prior to the weak case; participants in Group 2 read the weak article prior to the strong case, the control article prior to the neutral case, and the strong article prior to the weak case; and participants in Group 3 read the control article prior to the strong case, the strong article prior to the neutral case, and the weak article prior to the weak case. Counterbalancing the order in which they saw the confirmation bias articles made it so participants could see confirmation bias articles that contradicted the case (e.g., a strong confirmation bias article prior to a weak case), lined up with the case (e.g., a strong confirmation bias article prior to a strong case), as well as being exposed to the control that did not contradict or line up with the case.

Hypotheses

We pre-registered four main hypothesizes for the current study on the Open Science Framework. (1) We predicted a main effect of case strength on judgments (e.g., verdict), such that participants would be able to discern between strong and weak cases. However, (2) we also predicted that jurors would be vulnerable to cognitive biases, and would make judgments in accordance with confirmation bias information they were presented with prior to the case. For example, we expected people who received articles with "strong" confirmation bias information would be more likely to perceive the institution as liable, compared to those who received weak, or no confirmation bias information. (3) We also predicted an interaction between the strength of the case a participant read and the confirmation bias they were exposed to prior to reading the case. Specifically, we expected the effect of confirmation bias to be more prominent in the

38

neutral case strength condition when there was more ambiguity on what the "correct" verdict was. And (4), we expected a main effect of Rape Myth Acceptance (RMA), where people who endorsed more rape myths would be less likely to find the institution liable, regardless of condition.

Pilot Study B

The purpose of the third pilot was to test the neutral versions of the nine case facts to ensure that they were being perceived as moderate/neutral.

Methods.

Participants. Participants ($N = 89$, 59% female) were collected from Amazon's Mechanical Turk, an online paid subject pool, through which we used the integrated TurkPrime platform for efficiency. Participants were paid $0.25 and on average took about 7 minutes to complete the study. Thirteen participants were excluded for failing an attention check that asked participants to type in the seventh word of a sentence.

Procedure. Participants first read a brief summary of the parties involved in the case and the applicable law they were to base their judgements on. They were then shown the nine neutral case facts, in a randomized order for each participant, and asked to rate each one a 7 point scale (*1 = does not prove negligence at all, 7 = strongly proves negligence*).

Results. The nine statements were rated as generally moderate ($M = 3.57$, $SD = 1.85$; see Table 4 for descriptive statistics), and were mostly normally or uniformly distributed (Figure 7). Based on the mean response results we moved forward using them as the neutral case fact versions in Study 2.

Table 4.

Descriptive Statistics for Pilot B Case Fact Ratings

Case Facts	M	SD
1. Although they had a procedure in place, the institution was not sure who specifically the Plaintiff should report to.	4.90	1.71
2. The plaintiff was given several websites where she could find information on how to obtain counseling services.	3.13	1.82
3. After a couple weeks the plaintiff was informed that the investigation was still ongoing.	3.04	1.62
4. After 60 days the institution had concluded the investigation.	3.22	1.78
5. There had been minor previous complaints of harassment at the institution but the institution investigated a majority of these cases.	4.09	1.74
6. The alleged assailant had warnings at prior workplaces, but no history of violence.	3.30	1.67
7. The defendant performed a brief investigation into the assailant's past based on previous concerns, but found nothing conclusive.	4.25	1.71
8. Following the allegations, the alleged assailant was suspended until further investigation could be conducted.	2.79	1.80
9. The institution offered to change the alleged assailant's schedule to accommodate the Plaintiff so they would not have contact while at work.	3.38	1.82

Note. Mean and standard deviation for the nine case facts rated in Pilot B on a 7-point scale (1 = *very weak*, 7 = *very strong*).
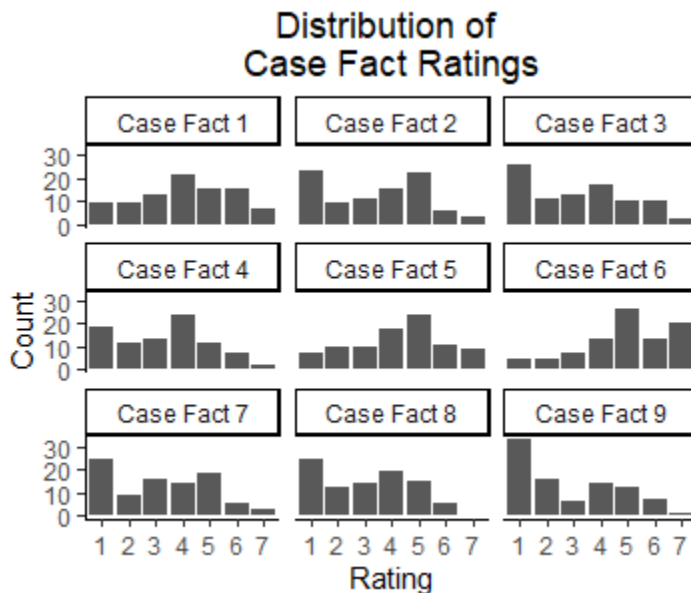

Figure 7. Count Distribution of Ratings for Pilot B Case Facts.

Study 2 Methods

Participants. Participants ($N = 449$, 53% female, $M_{Age} = 35.95$) were adult U.S. citizens. This intended sample size (n = 321) was determined by conducting a power analysis for repeated measure, within-between interactions (comparing across 3 groups), with enough power to detect an estimated effect size of Cohen's f .10, $\alpha = .05$, and 80% power. Participants were recruited through Prolific Academic, an online paid subject pool, and paid on average $10.54 per hour for their participation. On average, they took 21 minutes to complete the study.

Procedure. Participants read a consent form before beginning the study through Prolific Academic. Participants read three case vignettes, one for each case strength condition (strong, weak, and neutral). They were randomly assigned to one of the three groups, determining the order in which they saw the three case strength conditions (Group 1: strong, weak, neutral; Group 2: weak, neutral, strong; Group 3: neutral, strong, weak). Prior to reading each case vignette they read a mock news article containing a statement with confirmation bias (strong, weak, or control). They then read a case vignette, and following the case vignette, the participants filled out a questionnaire for the dependent measures. They read two more vignettes (and an article prior to each one), answering the same dependent measure questionnaire after each.

Materials.

Vignette. Participants read one-page vignettes describing a civil case against an institution for their role in a sexual assault that happened within their institution. The case fact versions that were used in the strong and weak case strength conditions were the same as in Study 1. For the neutral condition we modified the same nine case facts used

in the strong and weak conditions for the neutral case fact versions, which were pilot

tested in Pilot Study C. There was a total of nine different vignettes (three for each case

strength condition), each containing five out of the nine possible case facts. Similar to

Study 1, participants were randomly assigned to one vignette version for each of the three

conditions (see Figure 8) for a breakdown of the case fact combinations for each vignette
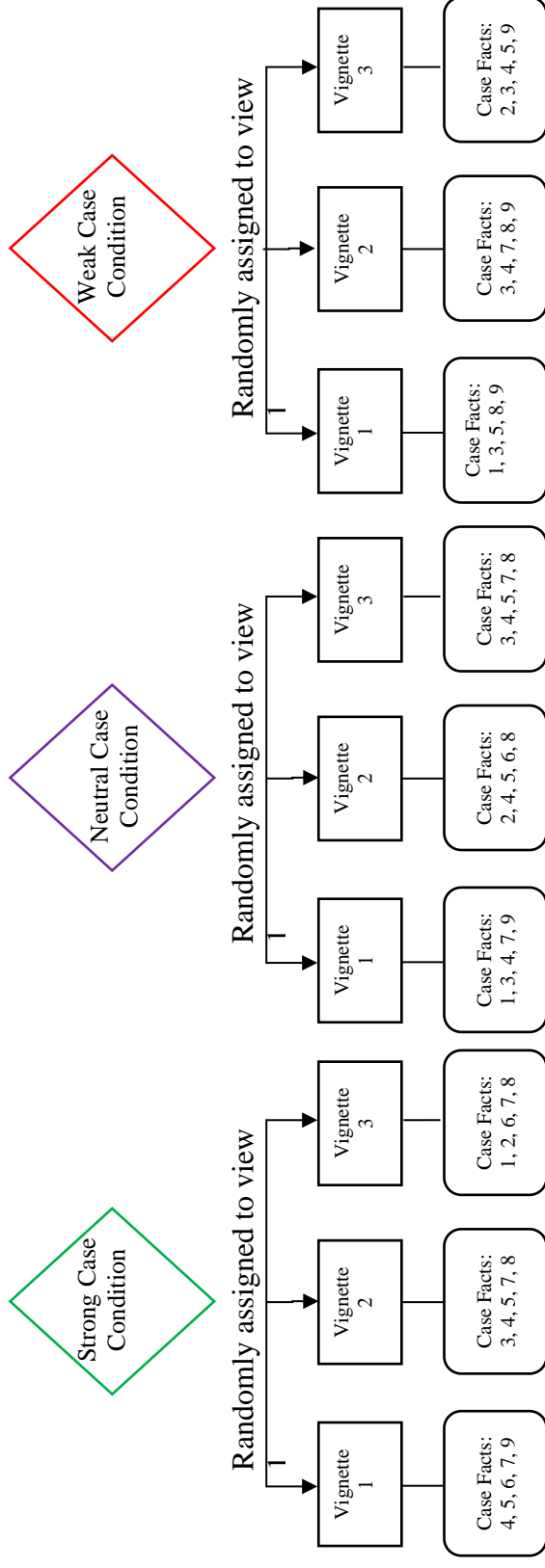
across the three conditions).

Figure 8. Flowchart depicting Study 2 vignette design. Process of how each of the nine vignette options were designed and assigned to the Neutral, Strong, and Weak conditions. The diamond indicates the case strength condition. For each condition that participants were in, they read only one total vignette, randomly assigned to them through Qualtrics functions. Following the arrows down, the conditions had three possible vignettes that participants could see. The bottom row indicates the case facts that were in each vignette. The vignette case fact combinations for the strong condition remained the same as Study 1. The case facts that made up each vignette in the neutral and weak conditions were randomly drawn using the same methods as Study 1 (i.e., random number combination generators). For example, vignette 1 in the neutral condition had case facts 1, 3, 4, 7, and 9. Refer to Appendix C1 for what case fact corresponds with what number in each of the conditions.

The same distractor variables (e.g., names, institution, location) were randomly assigned for each vignette. However, for Study 2, in order to simplify the design setup, who the plaintiff reported to, and the assailant's affiliation to the institution were removed. In order to make it even less likely that a participant would see the exact same plaintiff or assailant name more than once, we also separated the first and last names. There were five possible first names, and five possible last names for both the plaintiff and assailant. While it was possible that participants could see the same plaintiff first name more than once, it was improbable they would see the same first and last name combination.

Confirmation Bias Articles. The confirmation bias statements were embedded in a mock news article about an event inspired by the #MeToo movement, and were the same as we used for Study 1.

Measures. We used the same dependent measures from Study 1 minus the Rape Myth Acceptance Scale (which was an accidental oversight) and Trust in Institution measure (which was purposeful to reduce the length and expense of the questionnaire).

Study 2 Results

Prior to our analyses, we again tested the three vignette option in each of the three case strength conditions to ensure that in each condition all three options were being perceived as similar in strength. We performed chi-square tests of independence to test whether the vignette option participants saw in the conditions had an unintended effect on their verdicts (i.e., were people more likely to find the institution liable in the strong condition if they were given vignette option 1 vs 2 or 3). Across the strong, neutral, and weak conditions, the chi-square tests were not statistically significant, $x^2(2, 449) = 0.32$,

$p = .85$, $x^2(2, 449) = 0.68$, $p = .71$, $x^2(2, 449) = 0.88$, $p = .64$, respectively, allowing us to

conclude that there was not a statistically significant difference in verdict between the

three vignette options for all conditions. In addition, descriptive statistics for the case

facts used in Study 2 can be found in Table 2.

To test the main effect of case strength, we modeled a logistic regression, with

verdict (0 = Not Liable, 1 = Liable) regressed on case strength condition (reference group

= neutral condition), allowing intercept to vary by subject. Running the model for Type II

Wald chi-square effects, there was a significant main effect, $x^2(2, 449) = 207.57$, $p <$

.001. In support of our hypotheses, participants were able to differentiate between the

differing case strengths, such that participants were significantly more likely to vote

liable in the strong condition (82.8% voted liable), $b = 2.54$, $SE = .21$, $p < .001$, and

significantly less likely in the weak condition (23.6% liable), $b = -0.94$, $SE = .17$, $p <$

.001, compared to the neutral condition (39.4% liable). This finding was confirmed with

an ordinal regression looking at participants responses to the extent of liability across

conditions, such that participants rated the institution as more liable in the strong case, (M

= 5.57), $b = 2.45$, $SE = 0.15$, $p < .001$, and less liable in the weak case (M = 2.82), $b = -$

0.10, $SE = 0.12$, $p < .001$, compared to in the neutral case ($M = 3.62$).

To test whether the confirmation bias articles participants were exposed to (i.e.,

whether they received confirmation bias vs the control) influenced their judgements of

the cases, we modeled a logistic regression with confirmation bias and the case strength

interacting to predict verdict, allowing the intercept to vary for each participant. Running

the model for Type II Wald chi-square effects, there was a significant overall interaction,

$x^2(4, 449) = 15.25$, $p = .004$.

There is a significant interaction between receiving a strong article, compared to a control article, prior to a case with strong evidence, compared to neutral evidence, $b = 0.95$, $SE = .48$, $p = .049$. When provided a strong article, compared to a control article, prior to the neutral case, there is no significant effect, $b = 0.22$, $SE = .30$, $p = .46$. However, when provided a strong article, compared to a neutral article, prior to a strong case, there is a significant effect, $b = -0.72$, $SE = .35$, $p = .04$; such that, participants were less likely to vote liable in the strong case condition, when they received a strong article (76% liable verdicts), compared to a control article (85% liable). There is no significant effect of reading a weak article, compared to a control article prior to a neutral case, $b = -0.43$, $SE = .31$, $p = .16$. In addition, there was no significant interaction when receiving a weak article, compared to a control article, prior to the strong vs. neutral case, $b = 0.55$, $SE = .51$, $p = .28$, allowing us to conclude that the effect of the weak article, compared to the control, is also not significant for the strong case strength.

There was also no significant interaction for the strong vs. control article, prior to the weak vs. neutral case, $b = -0.52$, $SE = .48$, p $= .28$. Since the effect of the strong vs. control article was not significant in the neutral case, we can conclude that the strong article, compared to the control, also did not have a significant effect in the weak case. There was a significant interaction between the weak vs. control article prior to the weak vs. neutral case, $b = 1.13$, $SE = .47$, $p = .01$. As noted earlier, the effect of the weak, compared to control article, in the neutral case was not significant. However, when participants received a weak article, compared to control, prior to the weak case, there was a significant effect, $b = 0.68$, $SE = .32$, $p = .03$; such that, participants were more

likely to vote liable when given the weak article, compared to the neutral article, prior to the weak case condition.
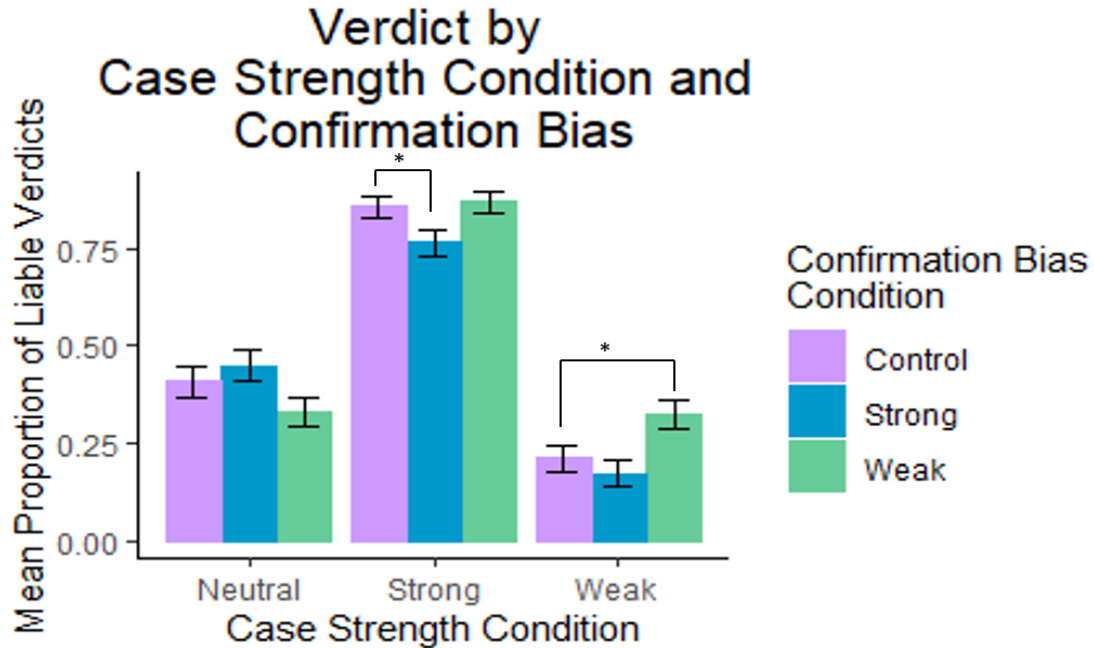
Figure 9. Verdict by Case Strength Condition and Confirmation Bias. Error bars represent standard errors. *p < .05.

Exploratory Analyses. Contrary to Study 1 results, modeling a logistic regression, with gender and condition interacting to predict verdict, running for Wald chi-square results, the effects of condition did not differ based on gender, $x^2(2, 449) = 4.42$, $p = .11$. Running a model with condition and the type of institution interacting, the effects of condition on verdict also did not differ based on the type of institution that the case was against, $x^2(6, 449) = 5.15$, $p = .52$. A two-way ANOVA with condition and article interacting to predict participant's confidence in their verdict showed a significant interaction, $F(4, 446) = 8.94 \, 4$, $p < .001$, $\eta_p^2 = .02$. To probe the interaction we split the

data by condition and ran three separate one-way ANOVAs with confirmation bias article predicting confidence ratings. There were significant differences in the strong case, $F(2, 446) = 13.95$, $p < .001$, $\eta_p^2 = .06$, and weak case, $F(2, 446) = 8.66$, $p < .001$, $\eta_p^2 = .04$, but no significant differences in the neutral case strength condition, $F(2, 446) = 0.21$, $p = .81$, $\eta_p^2 = .001$.

We performed Tukey HSD post hoc tests for the two significant effects. In the strong case, participants were more confident in their verdicts when they read a control article ($M = 5.60$) and weak article ($M = 5.57$), compared to the strong article ($M = 4.84$), $p < .001$ in both comparisons. There was no significant difference in confidence ratings between the control and weak articles, $p = .98$. In the weak case, participants were more confident in their verdicts when they read a control ($M = 5.32$) and strong article ($M = 5.21$), compared to the weak article ($M = 4.62$), $p < .001$, $p = .003$, respectively. There was no significant difference between the control and strong articles, p = .84. Due to the counterintuitive findings, where participants where actually less confident in their verdicts when they received a confirmation bias in accordance with the appropriate verdict, we ran a two-way ANOVA for both the strong and weak case strength conditions to see if these findings could be explained by whether the participant returned the appropriate verdict in each of the conditions. The effect of article on confidence in the weak case did not differ based on their verdict, $F(2, 443) = 0.81$, $p = .44$, $\eta_p^2 = .004$. That is, participants who appropriately voted not liable given the weak evidence strength, were not significantly more confident in their verdicts than those who did not vote appropriately in the weak condition.

However, in the strong case, there was a significant interaction of verdict and confirmation bias article on confidence, $F(2, 443) = 19.90$, $p = .002$, $\eta_p^2 = .03$. To probe the interaction we split the data by verdict and performed one-way ANOVAs to test whether the relationship between article and confidence differed based on if they returned the appropriate verdict or not. The effect of article on confidence ratings was no longer significant when looking at those who voted not liable in the strong case condition, $F(2, 74) = 0.89$, $p = .41$, $\eta_p^2 = .02$. There was a significant effect of article on confidence ratings when they appropriately returned liable verdicts, $F(2, 369) = 19.32$, $p < .001$, $\eta_p^2 = .10$; such that participants were significantly more confident in their verdict when they received a weak article ($M = 5.75$), and control article ($M = 5.71$), compared to a strong article ($M = 4.83$), $p < .001$ in both comparisons. There was no significant difference in confidence between the control and weak articles, $p = .96$. Therefore, the effect of confirmation bias article on confidence in the strong case strength condition may be partially explained by whether they returned an appropriate verdict. That is, only when they voted liable (as the evidence suggested) were participants more confident in their verdicts based on the confirmation bias article they were shown.

Study 2 Discussion

Study 2 replicated part of Study 1 findings, giving additional evidence that jurors can accurately discern evidence strength. However, contrary to Study 1, Study 2 results suggest that jurors are susceptible to being influenced by outside sources, in this case mock news articles. While the articles did not contain useful information relevant to the case, this unrelated information nevertheless biased judgments. It is important to note though that while they did have a significant effect, it was not in a direction that one

would expect. Rather than showing evidence of confirmation bias (e.g., participants being more likely to vote not liable when given a confirmation bias article suggesting a weak case against an institution), the articles were having the opposite effect. Participants instead were more likely to vote liable in a weak case if they were given a confirmation bias article suggesting a weak case. These results could be due to learning effects and correcting for their bias (see DeCoster & Claypool, 2004 for instances of correction). Due to the within subject element, participants may have picked up on the purpose of the articles' confirmation bias statements, and overcorrected for it in order to prevent themselves from being influenced by their bias.

The effect that the confirmation bias had on jurors' confidence in their judgments also was not in line with what one would intuitively expect. One may hypothesize that if a juror has information that is in line with the case they are judging (e.g., an article suggesting a strong case against an institution, followed by a strong case against an institution) they would be more confident in their decision, however this was not the case. When participants were given an article that was in agreement with the actual strength of the case they were judging, they were actually less confident in their verdicts compared to when it contradicted the case. This could potentially be explained by whether they returned the appropriate verdict when given a strong case, showing that when given weak and control articles prior to the strong case, they were more confident in their verdict when they returned the appropriate liable verdict. However, this pattern was not able to explain the difference in confidence ratings in the weak case strength condition. Given these results were not found in Study 1, and are overall counterintuitive, they should be further replicated before drawing strong conclusions from them.

Limitations. This study was susceptible to human error in forming the survey flow on Qualtrics. While the study itself was implemented as intended, the Rape Myth Acceptance (RMA) scale and in-depth demographics were unintentionally left out of the study, so we were unable to discover whether the RMA findings from Study 1 would replicate in Study 2. In addition, we do not have the kinds of demographics we would have liked to have had for this study. Although Prolific Academic has a wide range of participants, a majority of participants identify as white/Caucasian (73%), with only 5% mixed, 5% Latino/Hispanic, and less than 5% for all other ethnicities. Therefore, it is unclear whether our participants reflect diverse backgrounds. In addition, Prolific Academic does not allow researchers to filter based on prior felonies, so although we would have preferred to have a jury-eligible sample, it is possible that some our participants may not fully have met the jury-eligible criteria.

General Discussion

From a theoretical standpoint, this study has the potential to advance our understanding of the role confirmation biases play in civil cases, and whether jurors may be influenced by outside information and their personal beliefs. Both of the current studies suggest that jurors may be able to accurately judge cases and return appropriate verdicts based on strength of case facts. However, this result was limited in Study 1, where majority of participants (54%) were still voting liable when given weak evidence. These studies also reveal limitations in people's judgments, and suggest jurors may be unduly influenced by factors unrelated to the case.

Study 1 presents findings consistent with other literature looking at the effects of rape myths in court. Rape myth acceptance has been shown to affect sexual assault

investigations (e.g., Shaw et al., 2017) and legal proceedings (e.g., Hammond, Berry, & Rodriguez, 2011), but it was unclear whether these attitudes also impacted how people judge an institution's responsibility in the assault and investigative proceedings. This study provides initial, and somewhat novel, evidence that implications of these attitudes extend beyond criminal cases regarding the assault itself into adjacent realms. Similar to the mentioned studies, participants who endorsed rape myths in Study 1 were less likely to find an institution liable. This is not surprising, given if someone endorses rape myths, they likely endorse victim blaming attitudes (e.g., "if a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of control"). These victim-blaming attitudes, and general myths that undermine the severity of sexual assault likely could also lead to someone not feeling the need to hold an institution responsible for their role if they do not view the assault itself as serious. Exploratory analyses in Study 1 suggest men may also be more likely to endorse rape myths, and less likely to find an institution liable compared to women. While the effect of gender has been found in other studies (e.g., Hammond, Berry, & Rodriguez, 2011), this gender-related finding and the effect of institution type failed to replicate in Study 2, casting some doubt about the robustness of these findings.

Study 2 showed evidence that jurors were influenced by mock news articles unrelated to the case facts. However, these effects were not in line with confirmation bias, so it is unclear what mechanisms were at play. It is possible that the mock news articles put an emphasis on the #MeToo movement and therefore made participants inclined to support the plaintiff disregarding the confirmation bias statements altogether. It is also possible that the nature of sexual assault (vs. harassment which seems typical in

these institutional cases) is more extreme, and jurors may overall be more motivated to find those involved liable. Future directions may focus on the broader impact of media on #MeToo movements, including articles that depict it in a more negative light, and less extreme cases of harassment that may be harder to prove.

However, the confirmation bias article manipulations did have limitations. These articles were again always in the context of a Me Too movement related event. Therefore, it is possible that the topic of the movement itself made participants judge the institutions more harshly, regardless of what the confirmation statement was. In line with this, another limitation of failing to conduct manipulation checks for the articles. Due to this, it is unclear whether the articles and confirmation bias statements had the intended effect (i.e., whether the strong confirmation bias articles were eliciting pro-plaintiff attitudes). In addition, although participants were forced to stay on the article page for 45 seconds in Study 2, it is still possible that participants did not fully read the articles in order to be effected by the confirmation bias statement provided.

While the current project contained limitations, and the results only present an initial insight into how these cases are judged, there are still important implications for studies looking at these types of cases. Looking at claims against institutions has important real-world implications because of how widespread the issue is, how difficult sexual allegations are to prove, and the power institutions hold. One would hope that the increased media attention to the #MeToo movement would make people more aware of the issues, give victims a voice, and give these cases more credibility. However, media coverage is not all positive. People may be exposed to a variety of perspectives, some painting the #MeToo movement in a positive light (e.g., Bennett, 2018) and others telling

the narrative that the movement has gone too far, and lead to unfounded "less serious" accusations (e.g., Williams, 2019).

This study, and previous literature (Ruva & LeVasseur, 2012), suggest that jurors may take information they hear from news and other media sources (including information unrelated to a given case) into the courtroom, allowing it to influence their verdicts and perceptions of a given case. If the results from this study reflect true juror tendencies, there may be negative consequences to this media coverage, with it having unjust impacts on case outcomes. On the one hand it may lead to victims being painted as untrustworthy; and on the other it could interfere with the "innocent until proven guilty" concept, forcing the defendant to instead prove their innocence.

Overall, creating a framework to understand the role of rape myth acceptance and the possible effects of media and confirmation bias will allow us to work on a larger model, looking at whether people are influenced by evidence and biases that are introduced during trials, and the degree to which their pre-existing biases (e.g., rape myth acceptance) affects perceptions of institutions' liability. The more we understand the mechanisms at play in these case outcomes, the more we can move towards ensuring that victims are treated fairly and justice is served, and that institutions sufficiently protect their members and are sufficiently protected against false allegations.

REFERENCES

Ali, R. (2011, April 4). Dear colleague letter. Washington, DC: U.S. Department of Education, Office for Civil Rights.

Bennett, J. (2018, June 28). After #MeToo, the Ripple Effect. New York Times. Retrieved from https://www.nytimes.com/2018/06/28/arts/what-is-next-metoo-movement.html

Braman, E., & Nelson, T. E. (2007). Mechanism of motivated reasoning? Analogical perception in discrimination disputes. *American Journal of Political Science*, *51*, 940-956. https://doi.org/10.1111/j.1540-5907.2007.00290.x

Bressan, P., & Dal Martello, M. F. (2002). Talis pater, talis filius: Perceived resemblance and the belief in genetic relatedness. *Psychological Science*, 13, 213-218. http://doi.org.ezproxy1.lib.asu.edu/10.1111/1467-9280.00440

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38(2), 217–230. https://doi.org/10.1037/0022-3514.38.2.217

Buslon v. Seattle Impact FC, (Wash.Super. 2014).

Charman, S. D., Gregory, A. H., & Carlucci, M. (2009). Exploring the diagnostic utility of facial composites: Beliefs of guilt can bias perceived similarity between composite and suspect. *Journal of Experimental Psychology: Applied*, *15*, 76-90. http://doi.org.ezproxy1.lib.asu.edu/10.1037/a0014682

Chiwaya, N. (2018, October 11). New data on #MeToo's first year shows 'undeniable' impact. *NBC News*. Retrieved from https://www.nbcnews.com/news/us-news/new-data-metoo-s-first-year-shows-undeniable-impact-n918821

Civil Rights Act of 1964 § 7, 42 U.S.C. § 2000e et seq (1964).

Cooney, S. (2017, Nov 09). Harvey Weinstein, Kevin Spacey: More Men Accused of Assault. *Time*. Retrieved from time.com

Davis, J. (2002). Spotlight Church abuse report: Church allowed abuse by priest for years. *Boston Globe*. Retrieved from bostonglobe.com

DeCoster, J., & Claypool, H. M. (2004). A Meta-Analysis of priming effects on impression formation supporting a general model of informational biases. *Personality and Social Psychology Review*, *8*, 2–27. https://doi.org/10.1207/S15327957PSPR0801_1

Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, National Crime Victimization Survey, 2010-2016 (2017).

Denson v. The Corporation of the President of the Church of Jesus Christ of Latter-Day Saints, (D. Utah. 2018).

Doe v. Baylor University, (W.D. Texas, 2017).

Doe vs. University of N. Texas Bd. of Regents, (E.D. Tex. 2017).

Dror, I.E., Charlton, D., & Peron, A.E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International, 156*, 74-78. https://doi.org/0.1016/j.forsciint.2005.10.1017

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*, 1082-1090. http://doi.org.ezproxy1.lib.asu.edu/10.1037/0022-3514.57.6.1082

Education Amendments Act of 1972, 20 U.S.C. §§1681 – 1688 (1972).

Evans, J. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75*, 451–468. https://doi.org/10.1111/j.2044-8295.1984.tb01915.x

Evans, T., Alesia, M., and Kwiatkowski, M. (2016). "A 20-year toll: 368 gymnasts allege sexual exploitation". *Indianapolis Star*. Retrieved 17 Feb 2018.

Garcia, S. E. (2017, Oct. 20) The Woman Who Created #MeToo Long Before Hashtags. *The New York Times*. Retrieved from https://www.nytimes.com/2017/10/20/us/me-too-movement-tarana-burke.html

Hamm, J.A., PytlikZillig, L.M., Herian, M. N., Bornstein, B.H., Tomkins, A.J. & Hoffman, L. (2013). Deconstructing confidence in state courts. *Journal of Trust Research, 3*, 11-31.

Hammond, E. M., Berry, M. A., & Rodriguez, D. N. (2011). The influence of rape myth acceptance, sexual attitudes, and belief in a just world on attributions of responsibility in a date rape scenario. *Legal and Criminological Psychology*, *16*, 242-252. http://doi.org.ezproxy1.lib.asu.edu/10.1348/135532510X499887

Harley, E. M. (2007). Hindsight bias in legal decision making. *Social Cognition*, 25, 48–63. https://doi.org/10.1521/soco.2007.25.1.48

Hasel, L. E., & Kassin, S. M. (2009). On the presumption of evidentiary independence: Can confessions corrupt eyewitness identifications? *Psychological Science*, *20*, 122–126. https://doi.org/10.1111/j.1467-9280.2008.02262.x

Harris v. CBS News Communications Inc., (N.Y. Sup. 2018).

Hernandex v. Baylor University, (W.D. Tex., 2016).

Hill, C., Memon, A., & McGeorge, P. (2008). The role of confirmation bias in suspect interviews: A systematic evaluation. *Legal and Criminological Psychology*, *13*, 357-371. http://doi.org.ezproxy1.lib.asu.edu/10.1348/135532507X238682

Huang v. St. John's Evangelical Lutheran Church, (N.Y. Sup. 2012).

Jacobs, J. (2020, January 12). #MeToo Cases' New Legal Battleground: Defamation Lawsuits. The New York Times. Retrieved from https://www.nytimes.com/2020/01/12/arts/defamation-me-too.html

Judd, J. (2020, January 12). #MeToo Cases' New Legal Battleground: Defamation Lawsuits. *The New York Times*. Retrieved from https://www.nytimes.com/2020/01/12/arts/defamation-me-too.html

Kaye, D., & Koehler, J. (1991). Can Jurors Understand Probabilistic Evidence? *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *154*(1), 75-81. https://doi.org/10.2307/2982696

Lange, N. D., Thomas, R. P., Dana, J., & Dawes, R. M. (2011). Contextual biases in the interpretation of auditory evidence. *Law and Human Behavior*, *35*, 178–187. http://doi.org/10.1007/s10979-010-9226-4

Langley v. Faulkner University, (M.D. Ala. 2012).

Leippe, M. R., Eisenstadt, D., Rauch, S. M., & Seib, H. M. (2004). Timing of eyewitness expert testimony, jurors' need for cognition, and case strength as determinants of trial verdicts. Journal of Applied Psychology, 89(3), 524-541. doi:http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/0021-9010.89.3.524

Lovett, M., Bajaba, S., Lovett, M., & Simmering, M. J. (2018). Data quality from crowdsourced surveys: A mixed method inquiry into perceptions of Amazon's Mechanical Turk Masters. *Applied Psychology: An International Review*, *67,* 339-366. https://doi.org/10.1111/apps.12124

Meisner, J., Buckley, M., & Crepeau, M. (2019, July 12). R. Kelly hit with federal indictments in New York, Chicago; faces new racketeering, sex crime charges, allegations he paid to recover sex tapes and cover up conduct. *Chicago Tribune*. Retrieved from https://www.chicagotribune.com/news/criminal-justice/ct-r-kelly-arrested-federal-charges-20190712-6ghntysw3zf3lpncn4owcfzyje-story.html

Nakhaeizadeh, S., Dror, I. E., & Morgan, R. M. (2014). Cognitive bias in forensic anthropology: Visual assessment of skeletal remains is susceptible to confirmation bias. *Science & Justice*, *54*, 208-214. https://doi.org/10.1016/j.scijus.2013.11.003

Narchet, F. M., Meissner, C. A., & Russano, M. B. (2011). Modeling the influence of investigator bias on the elicitation of true and false confessions. *Law and Human Behavior*, *35(6),* 452–465. https://doi.org/10.1007/s10979-010-9257-x

Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175-220. https://doi.org/10.1037/1089-2680.2.2.175

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153-163. https://doi.org/10.1016/j.jesp.2017.01.006

PytlikZillig, L. M., Hamm, J. A., Shockley, E., Herian, M., Neal, T. M. S., Kimbrough, C., Tomkins, A. j., & Bornstein, B. H. (2016). The dimensionality of trust-relevant constructs in four institutional domains: Results from confirmatory factor analyses. *Journal of Trust Research, 6*, 111-150. https://doi.org/10.1080/21515581.2016.1151359

Richards, T. N. (2019). No evidence of "weaponized Title IX" here: An empirical assessment of sexual misconduct reporting, case processing, and outcomes. *Law and Human Behavior*, *43*, 180-192. https://doi.org/10.1037/lhb0000316

Ruva, C., & LeVasseur, M. (2012). Behind closed doors: The effect of pretrial publicity on jury deliberations. Psychology. Crime & Law. 431-452. https://doi.org/10.1080/1068316X.2010.502120.

Shaw, J., Campbell, R. and Cain, D. (2016), The view from inside the system: How police explain their response to sexual assault. *Am J Community Psychol*ogy, *58*, 446-462. https://doi.org/10.1002/ajcp.12096

Shaw, J., Campbell, R., Cain, D., & Feeney, H. (2017). Beyond surveys and scales: How rape myths manifest in sexual assault police records. *Psychology of Violence*, *7*, 602-614. https://doi.org.ezproxy1.lib.asu.edu/10.1037/vio0000072

Smith, L., Bull, R., & Holliday, R. (2011). Understanding juror perceptions of forensic evidence: Investigating the impact of case context on perceptions of forensic evidence strength. *Journal of Forensic Sciences, 56*, 409-14. https://doi.org.10.1111/j.1556-4029.2010.01671.x

Spohn, C., & Tellis, K. (2012). The Criminal Justice System's Response to Sexual Violence. Violence Against Women, 18(2), 169–192. https://doi.org/10.1177/1077801212440020

Süssenbach, P., Albrecht, S., & Bohner, G. (2017). Implicit judgments of rape cases: An experiment on determinants and consequences of implicit evaluations in a rape case. *Psychology, Crime, & Law, 23*, 291-304. https://dx.doi.org/10.1080/1068316X.2016.1247160

Tversky, A., Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*. *59*, 251–278.

U. S. Const. amend. IV.

Walmsley, S., & Gilbey, A. (2016). Cognitive biases in visual pilots' weather-related decision making. *Applied Cognitive Psychology*, *30*, 532-543. https://doi.org.ezproxy1.lib.asu.edu/10.1002/acp.3225

Williams, J. (2019, October 10). Two years of #MeToo: a disaster for men and women. Spiked. Retrieved from https://www.spiked-online.com/2019/10/10/two-years-of-metoo-a-disaster-for-men-and-women/

Winter, R. J. (2018). In Harvey Weinstein's Wake. *Monitor on Psychology*, *49*, 25. http://www.apa.org/monitor/2018/04/jn

APPENDIX A

PILOT A MATERIALS

Appendix A1.

General Summary:

This case arises from Defendant's response both before and after an alleged

sexual assault. The party who filed this lawsuit is called the plaintiff. The plaintiff seeks

damages from the institution, who is the defendant in this case, for injuries that she

claims the defendant caused her. The specific tort in question is that of negligence.

Specifically, the plaintiff seeks compensation for emotional distress damages caused by

the defendant's failure to create a safe environment, negligent supervision, and overall

handling of the investigation. You must decide what the facts are in this case. The term

injury here forth will refer to said emotional distress damages.

Juror Instructions and law to be applied:

It is your job as jurors to determine whether the institution demonstrated negligent

actions in relation to this alleged assault, in breach of the duties the institution owed to its

members.

Plaintiff must prove the following are more likely to be true than not true:

1. Defendant was negligent; 2. Plaintiff was injured; 3. Defendant's negligence

caused plaintiff's injuries

Negligence is the failure to use reasonable care. Negligence may consist of action

or inaction.

To prove the first element, that the defendant was negligent, the plaintiff must

prove that one or more of the following are more likely to be true than not true:

1) The Defendant showed deliberate indifference to complaints; OR

2)   The Defendant failed to promptly and appropriately investigate and respond to the reported assault; OR

3)   The Defendant committed fraudulent concealment or nondisclosure; OR

4)   The Defendant failed to appropriately supervise members resulting in a hostile and dangerous environment; OR

5)   The Defendant's actions fell well below standard of care required for the reasonable person; OR

6)   The Defendant's policies and selective conduct code enforcement resulted in a discriminatory environment.

Plaintiff must also prove that the defendant's negligence was a cause of Plaintiff's injury. Negligence causes an injury if it helps produce the injury and if the injury would not have happened without the negligence.

APPENDIX B

STUDY 1 MATERIALS

Appendix B1. Participants saw each case in the context of either a university, church, Hollywood television network, or soccer league. The names that were formulated for each were:

1. University: Murphy College (randomly selected a last name that was not tied to a real university)

2. Church: St. Agrecius Church (selected a saint that was not already tied to a common church)

3. Hollywood: PNT News (randomly generated a string of three letters that was not tied to a real network station)

4. Soccer League: a conference of the Women's American Soccer League (meant to specify a women's league while remaining an institution at large)

In addition, for each of the randomized variables there were several options for each variable that could be randomly piped in. The options for each variable are listed. All variable options remained consistent across institutions unless otherwise specified.

Plaintiff Name. The names of the defendant were selected from a random name generator website that generates a first and last name from a database of common names. To keep gender consistent, female names were used for the plaintiffs.

1. Tammy Moore

2. Rachel Hill

3. Margaret Evans

4. Helen Washington

5. Melissa Parker

Plaintiff Age. Assigned random integer 18-25 using Qualtrics' embedded data pipe function.

Location. These were selected based on political party orientation (conservative vs. liberal). Texas = rep/conservative, California = dem/liberal, AZ + OH have mixed affiliations leaning towards conservative and liberal, respectively. This information was gathered from the following website: https://news.gallup.com/poll/226643/2017-party-affiliation-state.aspx.

1. Texas

2. Arizona

3. California

4. Ohio

Date. These were randomly picked out of 577 possible dates between January 1, 2017 and July 31, 2018.

1. February 3, 2017

2. February 15, 2017

3. April 27, 2017

4. May 26, 2017

5. July 20, 2017

Assailant Name. The names of the alleged assailant were selected from a random name generator website that generates a first and last name from a database of common names. To keep gender consistent, male names were used for the alleged assailants.

1. Peter Washington

2. Eugene Ward

3. Daniel Cooper

4. Keith Price

5. Anthony Bell

Alleged Assailant's Affiliation. The options for the alleged assailant's role/affiliation to the institution were created based on the institution – this was done to keep a coherent story within the vignettes that remained consistent with what one would expect in different institutions. There were three options for each institution.

University: 1. Research Assistant, 2. Student, 3. Teaching Assistant

Church: 1. Cantor, 2. Liturgist, 3.Altar Server

Hollywood: 1. Network employee, 2. Journalist for the network, 3. Talk show host for the network

League: 1. Team Manager, 2. Assistant Coach, 3. Trainer

Elapsed Time. The time it took the plaintiff to report the alleged assault to the institution were selected to remain somewhat prompt.

1. A day

2. A week

3. Two days

4. Three days

5. Four days

Reported to. Who the plaintiff reported the alleged assault to were created based on the institution – this was done to keep a coherent story within the vignettes that remained consistent with what one would expect in different institutions. There were three options for each institution.

University: 1. University Counselor, 2. Professor, 3. Dean

Church: 1. Parish Administrator, 2. Deacon, 3. Priest

Hollywood: 1. Human Resource Director, 2. Network Executive, 3. Project Manager

League: 1. League Director, 2. Head Coach, 3. League Manager

Relationship. The plaintiff's relationship to the defendant (i.e. how well they knew each other and the extent of their interactions) were designed to reflect a range of options you would expect among co-worker and peer relationships.

1. Had some, but limited, contact with in the past

2. Had known through the institution and considered an acquaintance

3. Had interacted with several times and known through organized group events

4. Had developed a close friendship with over the months

Appendix B2. The strong and weak case facts that were embedded in the strong and weak case vignettes for Study 1.

Strong Case Facts. Case facts used in the strong case vignettes.

1. The defendant did not have a set procedure in place for handling allegations, so it was not clear to the plaintiff who she was supposed to report to, or who was in charge of handling the investigation.

2. The plaintiff was not made aware of any counseling services available.

3. After a couple weeks, she had not been made aware of any updates in the investigation. She had tried to contact them but could not get a hold of anyone. After a week she still had not heard back from anyone.

4. After 148 days, the institution had not concluded the investigation, nor had they indicated whether it was still under review.

5. There had been previous complaints of harassment at the institution, but no investigations were initiated for any of the complaints.

6. There were suspicious prior allegations against the alleged assailant for acting violently towards others.

7. The defendant claims to have performed a brief investigation into the alleged assailant's history based on previous concerns, but did not further investigate any alarming findings that they were made aware of during the brief investigation.

8. Following the report and allegations, the alleged assailant was able to return full time.

9. The institution offered no accommodations to the plaintiff to ensure she would have no contact with the alleged assailant.

Weak Case Facts. Case facts used in the weak case vignettes.

1. The defendant directed the plaintiff to their office where someone in charge of handling allegations would be able to formally handle the report.

2. Plaintiff was given contact information for counseling services that the institution compensates members for in circumstances of stress and trauma.

3. The plaintiff was given several updates throughout the investigation and was told immediately when the case had concluded.

4. After 14 days the institution had concluded the investigation and notified the Plaintiff.

5. This seemed to be the first complaint of harassment at the institution.

6. The alleged assailant had no prior allegations against him.

7. The defendant had performed a brief investigation into the assailant's history, and the alleged assailant had a clean history.

8. Following the investigation, the assailant was suspended indefinitely.

9. The institution enacted a no contact policy so that the plaintiff would not have any form of contact with the assailant on their property.

Appendix B3. Example of a vignette in each of the three case strength conditions. Where the bold words are piped in with one of the options corresponding to the randomized variable option (see Appendix C5 for the options in each of the variables). The italicized sentences are the case facts that made up that vignette.

Strong Case Vignette 1. Vignette option 1 in the strong condition embedded in the context of a case against a church as the institution.

Plaintiff Name was a Plaintiff Age year old, member of St. Agrecius Church, incorporated under the laws of the State, at its location in Location at the time of the reported assault. The victim stated that on Data she was sexually assaulted by Assailant Name, who was Affiliation at the time. Elapsed Time after the event, she went to a Reported To to report the assault. After 148 days, the institution had not concluded the investigation, nor had they indicated whether it was still under review. The institution's personnel questioned the plaintiff on her relationship with the assailant, who she Relationship. She stated that all interactions she had with the

alleged assailant had been platonic. After hearing her account of what happened, the institution commenced an initial investigation. There had been previous complaints of harassment at the institution, but no investigations were initiated for any of the complaints. During the investigation process, the institution looked into the assailant and plaintiff. There were suspicious prior allegations against the alleged assailant for acting violently towards others. The defendant claims to have performed a brief investigation into the alleged assailant's history based on previous concerns, but did not further investigate any alarming findings that they were made aware of during the brief investigation. The institution offered no accommodations to the plaintiff to ensure she would have no contact with the alleged assailant.

Weak Case Vignette 2. Vignette option 2 in the weak condition embedded in the context of a case against a television network as the institution.

Plaintiff Name was a Plaintiff Age year old, member of PNT News, a national television network and domestic business corporation, at its location in Location at the time of the reported assault. The victim stated that on Date she was sexually assaulted by Assailant Name, who was a Affiliation at the time. Elapsed Time after the event, she went to a Reported To to report the assault. The plaintiff was given several updates throughout the investigation and was told immediately when the case had concluded. The institution's personnel questioned the plaintiff on her relationship with the assailant, who she Relationship. She stated that all interactions she had with the alleged assailant had been platonic. After hearing her account of what happened, the institution commenced an initial investigation. After 14 days the institution had

concluded the investigation and notified the Plaintiff. During the investigation

process, the institution looked into the assailant and plaintiff. The defendant had

performed a brief investigation into the assailant's history, and the alleged assailant

had a clean history. Following the investigation, the assailant was suspended

indefinitely. The institution enacted a no contact policy so that the plaintiff would not

have any form of contact with the assailant on their property.

Neutral Case Vignette. The vignette used in the neutral condition embedded in the

context of a case against a university as the institution. This case contained all

distractor information still, but did not provide participants with any of the case facts.

Plaintiff Name was a Plaintiff Age year old, member of Murphy College, an

accredited bachelor degree-granting educational institution, at its location in Location

at the time of the reported assault. The victim stated that on Date she was sexually

assaulted by Assailant Name, who was a Affiliation at the time. Elapsed Time after

the event, she went to a Reported To to report the assault. The institution's personnel

questioned the plaintiff on her relationship with the assailant, who she Relationship.

She stated that all interactions she had with the alleged assailant had been platonic.

After hearing her account of what happened, the institution commenced an initial

investigation. During the investigation process, the institution looked into the

assailant and plaintiff.

Appendix B4. The strong, weak, and control confirmation bias articles that participants would read prior to each case.

The Hubbard Post

Issue 23   May 18, 2018

Debate Amongst Politicians on New #MeToo Policy

By Tomas Campbell

The rising #MeToo movement has brought with it many reactions on whether current policies do enough to protect against continued harassment. The issue of whether to enact stricter policies has been a topic of conversation among Congress members. As it currently stands, institutions have little guidance in how to proceed with harassment allegations. Supporters of this motion argue that clearer policies will hold members to a higher standard, and make it more difficult to defend any questionable activity going on in their institution. However, those against stricter policies see a possibility that the stronger policies may also have a backlash effect, or no effect at all. Institutions cannot watch over all members, and while it is their job to create a safe environment, it is unrealistic to expect complete oversight of everything going on. Creating too-strict policies may create guidelines that are too difficult to meet, leaving institutions vulnerable to unwarranted lawsuits.

The control article that all participants read prior to the neutral case.

The Anton Gazette

#MeToo Rally Amidst Claims Against Large Institution

By Francis Beck

A recent #MeToo rally took place in Washington D.C., attracting large crowds. Some participated in support of victim advocacy, whereas others participated in opposition. The rally was in response to a lawsuit against a large institution for poorly handling a sexual assault allegation reported at their institution. The #MeToo supporters organized the rally to bring attention to the issue of how allegations of sexual assault are managed. Since the #MeToo movement began in 2017, increasingly larger numbers of people have been making allegations of sexual misconduct. Supporters of the rally question whether parties are being held responsible enough. Some attendees reported they believed big-name corporations and influential people were still unfairly protected. Rally protestors made statements about how the #MeToo movement was blowing the issue out of proportion. Other protesters argued that some people make allegations because they believe they will benefit from filing unfounded suits against the accused people, and in some cases, against institutions that allegers argue failed to protect them. After interviews with both rally supporters and protesters, it was apparent that the rally was more about the #MeToo movement itself than it was about the particular lawsuit that instigated this rally. The lawsuit that instigated this rally is closely related in its fact pattern to an upcoming civil lawsuit against a different institution. Compared to other similar cases, people at this rally agreed that the upcoming civil lawsuit seemed to favor the plaintiff, as people thought the institution seemed to have done little to act in a timely and fair manner in their investigations to accommodate the parties involved.

The article participants read if assigned to read the strong confirmation bias (prior to the strong case). Those who were assigned to read the control article prior to the strong case read the exact same article with the last sentence removed.

The Patton Herald

#MeToo Movement Inspires Local Artists

By Suzanne Carpenter

A #MeToo exhibit that took place in an art gallery in downtown Los Angeles, California showed graphic artwork depicting sexual assault. One piece in particular attracted a lot of attention. The artist created a set of paintings telling the story of the backlash she experienced after coming forward about being assaulted, criticizing the lack of support she was offered. The piece sparked controversial reactions from the other artists and exhibit attendees on opening night about whether the artistic representation fairly reflected the perspectives of everyone involved. Some people talked about how it took the attention away from the assailant and placed unfair blame on an outside party's involvement (those people to whom she reported the assault). Others discussed the value of the subjective perspective of the victim, regardless of how poorly it reflected on the assailant and the party to whom she reported the assault.

The article participants read if assigned to read the weak confirmation bias (prior to the weak case). Those who were assigned to read the control article prior to the weak case read the exact same article with the last sentence removed.

Appendix B5. Measures used in Study 1.

Dependent Measures.

1. The plaintiff proved negligence

    a. 1 = strongly disagree, 7 = strongly agree

2. The plaintiff proved injury

    a. 1 = strongly disagree, 7 = strongly agree

3. I find the defendant_____.

    a. Not Liable

    b. Liable

4. If 3b (only displayed if answered Liable to Question 3)

    a. What damages would you award the plaintiff?

        i. $_____

5. How confident are you in your decision?

    a. 1 = not at all confident, 7 = very confident

6. How liable do you believe the institution is?

    a. 1 = not at all liable, 7 = very liable

7. Overall, how strong is the prosecution's case? _____

    a. 1 = very weak, 7 = very strong

8. Overall, how strong is the defense's case? _____

    a. 1 = very weak, 7 = very strong

9. What influenced your decision most? Please check all boxes that apply.

    a. Precedent: if so please briefly explain _____

    b. Case evidence alone

c. Other applicable law: if so please specify _____

d. Unsure

e. Other: please specify _____

10. How likely do you think it is that a similar case would happen again at this

institution?

    a. 1 = very unlikely, 7 = very likely

11. Please rate the strength of the following pieces of information, in regards to how

strong they are in ability to prove the plaintiff's case.

    *The following were rated 1 = very weak, 10 = very strong; or participants

    indicated "I was not given this information"

    a. The victim's relationship to the assailant.

    b. How long it took the plaintiff to report the assault.

    c. The institution's policy and procedure on who to report the assault to.

    d. The counseling services the institution offered.

    e. The extent to which the institution updated the plaintiff.

    f. How long it took the institution to investigate the plaintiff's report.

    g. The institution's history of harassment complaints.

    h. The assailant's history of complaints.

    i. The institution's investigation into the assailant's history.

    j. Whether the assailant was able to return to work.

    k. The actions the institution took in regards to the plaintiff's possible future

    contact with the assailant.

Trust in Institutions. Participants rated the following five items on a 7-point likert scale (1 = strongly disagree, 7 = strongly agree)

> Please answer the following questions in relation to your general attitudes toward that type of institution (i.e. Universities, Churches, Hollywood, or Sport leagues).
>
> 1. The procedures followed by this type of institution are ethical.
> 2. This type of institution uses their power appropriately.
> 3. I believe this type of institution shares my values.
> 4. This type of institution is overly influenced by special interest groups.
> 5. For the most part, the work done by this type of institution is out of care and concern for the general population.

Updated Illinois Rape Myth Acceptance Scale. 22 item questionnaire measured on a 5 point likert scale (1 = strongly agree, 5 = strongly disagree). We reverse scored these items so that higher scores would indicate higher rape myth acceptance.

1. If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of control.
2. When girls go to parties wearing slutty clothes, they are asking for trouble.
3. If a girl goes to a room alone with a guy at a party, it is her own fault if she is raped.
4. If a girl acts like a slut, eventually she is going to get into trouble.
5. When girls are raped, it's often because the way they said "no" was unclear.
6. If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex.
7. When guys rape, it is usually because of their strong desire for sex.

8. Guys don't usually intend to force sex on a girl, but sometimes they get too sexually carried away.

9. Rape happens when a guy's sex drive gets out of control.

10. If a guy is drunk, he might rape someone unintentionally.

11. It shouldn't be considered rape if a guy is drunk and didn't realize what he was doing.

12. If both people are drunk, it can't be rape.

13. If a girl doesn't physically resist sex—even if protesting verbally—it can't he considered rape.

14. If a girl doesn't physically fight back, you can't really say it was tape.

15. Rape probably didn't happen if the girl has no bruises or marks.

16. If the accused "rapist" doesn't have a weapon, you really can't call it a rape.

17. If a girl doesn't say "no" she can't claim rape.

18. A lot of times, girls who say they were raped agreed to have sex and then regret it.

19. Rape accusations are often used as a way of getting back at guys.

20. A lot of times, girls who say they were raped often led the guy on and then had regrets.

21. A lot of times, girls who claim they were raped just have emotional problems.

22. Girls who are caught cheating on their boyfriends sometimes claim that it was a rape.

APPENDIX C

STUDY 2 MATERIALS

Appendix C1. The strong, neutral, and weak case facts used in the corresponding case vignettes for Study 2.

Strong Case Facts. Case facts used in the strong case vignettes.

1. The defendant did not have a set procedure in place for handling allegations, so it was not clear to the plaintiff who she was supposed to report to, or who was in charge of handling the investigation.

2. Plaintiff was not made aware of counseling services available.

3. After a couple weeks, she had not been made aware of any updates in the investigation. She had tried to contact them but could not get a hold of anyone. After a week she still had not heard back from anyone.

4. After 148 days, the institution had not concluded the investigation, nor had they indicated whether it was still under review.

5. There had been previous complaints of harassment at the institution, but no investigations were initiated.

6. There were suspicious prior allegations against the alleged assailant for displaying aggressive behavior in the workplace.

7. The defendant claims to have performed a brief investigation into the alleged assailant's history before he was hired, but did not appear to further investigate any potential concerns that they were made aware of during the brief investigation.

8. Following the report and allegations, the alleged assailant was able to return full time.

9. The institution offered no accommodations to the plaintiff to ensure she would have no contact with the alleged assailant.

Neutral Case Facts. Case facts used in the neutral case vignettes.

1. Although the institution had a procedure in place for handling allegations of sexual assault, it was unclear who specifically the Plaintiff should report the incident to.

2. The plaintiff was given several websites where she could find information on how to obtain counseling services.

3. After a couple weeks the plaintiff was informed that the investigation was still ongoing.

4. After 60 days the institution had concluded the investigation.

5. There had been minor previous complaints of harassment at the institution but the institution investigated a majority of these cases.

6. The alleged assailant had warnings at prior workplaces, but no history of violence.

7. The defendant performed a brief investigation into the assailant's past based on previous concerns, but found nothing conclusive.

8. Following the allegations, the alleged assailant was suspended until further investigation could be conducted.

9. The institution offered to change the alleged assailant's schedule to accommodate the Plaintiff so they would not have contact while at the institution.

Weak Case Facts. Case facts used in the weak case vignettes.

1. The defendant directed the plaintiff to their office where someone in charge of handling allegations would be able to formally handle the report.

2. Plaintiff was given contact information for counseling services that the institution compensates members for in circumstances of stress and trauma.

3. The plaintiff was given several updates throughout the investigation and was told immediately when the case had concluded.

4. After 14 days the institution had concluded the investigation and notified the Plaintiff.

5. This seemed to be the first complaint of harassment at the institution.

6. The alleged assailant had no prior allegations against him.

7. The defendant had performed a brief investigation into the assailant's history, and the alleged assailant had a clean history.

8. Following the investigation, the assailant was suspended indefinitely.

9. The institution enacted a no contact policy so that the plaintiff would not have any form of contact with the assailant on their property.

APPENDIX D

IRB APPROVAL

EXEMPTION GRANTED

Tess Neal
Social and Behavioral Sciences, School of (SSBS)
602/543-5680
Tess.Neal@asu.edu

Dear Tess Neal:

On 1/29/2019 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Discerning Evidence Strength in Sexual Assault Liability Cases
Investigator:	Tess Neal
IRB ID:	STUDY00009559
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• Judge Consent Form, Category: Consent Form; • Measures, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Juror Consent Form, Category: Consent Form; • Judge Recruitment Form, Category: Recruitment Materials; • IRB Protocol, Category: IRB Protocol; • Juror Recruitment Form, Category: Recruitment Materials; • Vignette, Category: Other (to reflect anything not captured above);

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 1/29/2019.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).