QUESTIONS OF CHIVALRY IN POLICE-CITIZEN INTERACTIONS: AN
EXAMINATION OF PROCEDURAL JUSTICE

by

Emily Cahill
A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Arts
Criminology, Law and Society

Committee:

_____   Director

_____

_____

_____

_____   Department Chairperson

_____   Dean, College of Humanities
and Social Sciences

Date: _____   Spring Semester 2016
George Mason University
Fairfax, VA

Questions of Chivalry in Police-Citizen Interactions: An Examination of Procedural Justice

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts at George Mason University

by

Emily Cahill
Bachelor of Sciences
Old Dominion University, 2014

Director: Stephen D. Mastrofski, Professor
Department of Criminology, Law and Society

Spring Semester 2016
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my parents who have supported me from the very beginning.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

ABSTRACT

QUESTIONS OF CHIVALRY IN POLICE-CITIZEN INTERACTIONS: AN EXAMINATION OF PROCEDURAL JUSTICE

Emily Cahill, M.A.

George Mason University, 2016

Thesis Director: Dr. Stephen D. Mastrofski

This thesis examines whether a citizen's gender predicts the amount of procedural justice that he or she is shown by an officer during a police-citizen encounter. Possible gender effects are examined in the context of the chivalry and the selective chivalry hypotheses in order to determine whether females are shown more procedurally just treatment than males and whether a female citizen's treatment by police is contingent upon her conformity to traditional gender norms. Using multiple linear regression to analyze data collected during systematic social observations of 243 police-citizen interactions, no evidence of a significant gender effect on officers' procedural justice decision making was revealed. It was found, however, that a citizen's minority status, involvement in an encounter with a service-related primary problem, and negative attitude toward police at the beginning of an interaction all have a significant negative impact on procedural justice. Consideration of these results and of the predictors' effects on the individual elements of procedural justice suggests that chivalry theory may not be

applicable to procedural justice decision-making; that service-related encounters do not offer as many opportunities for officers to show neutrality as do other encounters; and that minority citizens do not benefit from as neutral of decision-making by the agents of the criminal justice system as do their White, non-Hispanic counterparts.

INTRODUCTION

Perhaps as tense as they have been since the Los Angeles Riots of 1992, police-community relations in America have been especially turbulent in recent years. High-profile cases like the fatal shooting of Michael Brown in Missouri and the arrest and subsequent death of Sandra Bland in Texas have led many Americans to become more vocal about their dissatisfaction with police practices, leaving frustrated police and community leaders turning to researchers for explanations about whether – and which – extralegal factors are actually influencing police decision-making and what steps can be taken to improve police-community relations and increase police legitimacy. A common recommendation to accomplish these goals has been to increase procedural justice (Mazerolle, Bennett, Davis, Sargeant, & Manning, 2013; President's Task Force on 21st Century Policing, 2015, Tyler, 2004).

Distinguished from other common criteria against which the public might evaluate police performance, procedural justice is defined as the fairness of police processes (Tyler, 2004). Related only to how police treat people during police-citizen encounters, procedural justice is unrelated to whether officers secure the outcomes desired by citizens with whom they interact or with how well they accomplish other performance-related goals like controlling crime in their beats. It is generally defined according to the following four elements, each of which relates only to the treatment of

citizens: participation, related to whether officers give citizens an opportunity to explain their side of or take on a situation; neutrality, related to whether officers make fair and objective decisions based on facts instead of personal biases; dignity, related to the politeness and respect shown to a citizen; and trustworthy motives, related to how officers show care and concern for citizens' well-being (Mastrofski, Jonathan-Zamir, Moyal, and Willis, 2016; Tyler, 2004). This concept of procedural justice, and the four elements it comprises, has become such an important consideration in recent research because it has been shown to increase both public satisfaction with police and police legitimacy – "the belief that police are entitled to call upon the public to follow the law and help combat crime and that members of the public have an obligation to engage in cooperative behaviors" (Tyler, 2004, pp. 86-87; Mazerolle et al., 2013).

Although there has been an upsurge in procedural justice research in recent years and there is a fair deal of extant research related to the influence of extralegal factors on police decision-making, there are still gaps in these studies that need filling. For instance, although demographic information is commonly considered in these kinds of studies, certain variables – like gender – are often included only as controls with other variables of greater interest and they are not examined in the context of a specific theoretical framework. Additionally, literature reviews and meta-analyses of studies on police decision-making published in the last fifteen years have reported mixed findings about the effects of citizen gender and, even when the findings have been more consistent, they have been in conflict with other evidence that has since been published (Lytle, 2014; Mastrofski et al., 2016; Skogan & Frydl, 2004; Worden & McLean, 2014). These

theoretical weaknesses and conflicting findings about how citizen characteristics influence police decision-making reveal the need for greater examination from a theoretical perspective that can help to explain the discrepancies. For this reason, the present study will examine specifically the effect of citizen gender on the amount of procedural justice shown by police and will do so in the context of the chivalry hypothesis, which suggests that females receive more favorable and lenient treatment from the criminal justice system than do their male counterparts (Herzog & Oreg, 2008; Moulds, 1978; Visher, 1983).

While many researchers might report findings of gendered leniency with little consideration for its implications, those who warn against the long-term consequences of gender disparity in the justice system reveal why it is important to identify and evaluate. Moulds (1978), for example, explains how lenient treatment of women by the justice system has "set a tone for the inferior regard of women held by much of society" (pp. 419). Herzog and Oreg (2008) provide evidence that beliefs that women should receive special treatment by the justice system are often coupled with sexist attitudes and that even if they are benevolent, "benevolent sexism also reinforces patriarchy by portraying women as weak and needy" (pp.50). Moulds (1978) goes as far as to say that the institutional view of women as weaker than men or in need of special treatment by the justice system causes "extensive personal psychological, social, economic, and political damage to the democratic notions of self-determination and equality" (pp.430). Consideration of a possible chivalry effect on the amount procedural justice shown to citizens will allow me to contribute to this discussion of disparate treatment of genders by

3

the justice system and its potential consequences.

The present study will use data collected during systematic social observations of 243 police-citizen interactions at two divisions of the Gallanton Police Department (pseudonym) and examine them in the context of the chivalry hypothesis to determine whether police show more procedural justice to females than to males. These interactions will also be examined in the context of the selective chivalry hypothesis – which specifies that females only receive chivalrous treatment from the justice system *if* they conform to traditional gender norms (Herzog & Oreg, 2008) – as I attempt to answer two more nuanced research questions: Will resistant behavior by women toward police have a greater influence on procedural justice than does the same behavior from men? Will female citizens who demonstrate a need for police assistance or intervention receive more procedural justice than male citizens who also demonstrate need?

The next section of this thesis will provide an overview of related literature and explain procedural justice and the chivalry and selective chivalry hypotheses in greater detail. Trends in the methods and findings of extant research will be examined, as will be the gaps and limitations that need to be addressed. An explanation of the methodology of the present study follows and will include information about the data collection sites, descriptions of the instruments used, and a discussion of my statistical model and the variables included in it. The statistical findings will then be reported and discussed and, finally, the last section will describe the implications of this research, how this study could have been improved, and what other research can take from it going forward.

LITERATURE REVIEW

This chapter begins with a more detailed explanation of procedural justice in the context of police discretion and its association with police legitimacy. Common themes and findings in procedural justice research are then described, as are the limitations of past studies. The benefit of using observational research in studies of procedural justice is also discussed and deeper explanations of the chivalry and selective chivalry hypotheses are provided along with descriptions of studies that have produced evidence supporting them. Other studies of gender effects in the criminal justice system are reviewed and at the end of this chapter are remarks about how the present study will attempt to overcome the limitations of extant research, as well as what I expect to find in my analysis based on existing evidence and theory.

Procedural Justice, Legitimacy, and Discretion

As stated in the previous section, recent strains on police-community relations in the United States have yielded an increasing demand for research about what police departments can do to improve them and regain trust of the American people. While these improvements may be of primary interest to local government and police department leaders, the federal government has also expressed concern and, in December of 2014, President Barack Obama enlisted a Task Force on 21st Century Policing to identify the policing practices that would be most effective in reducing crime while also

improving public trust (President's Task Force on 21st Century Policing, 2015). Two of the major themes among the recommendations in the final report submitted by this blue-ribbon task force were a conscious effort to remove implicit and explicit biases from police decision-making and, of even greater emphasis, to increase procedural justice. Pillar One of the task force's final report describes how adopting procedural justice as a "guiding principle for internal and external policies and practices to guide their interactions with rank and file officers and the citizens the serve" (President's Task Force on 21st Century Policing, 2015, p. 1) will allow departments to build trust and legitimacy with the public. To understand why procedural justice has come so highly recommended and how it is thought to contribute to legitimacy, however, one must have a clear understanding of what actually constitutes procedurally just police behavior.

Descriptions of the four elements of procedural justice are provided in the preceding section but behavioral examples are helpful, as well. The element of participation, for instance, is displayed when an officer asks citizens for their "side of the story" and when he or she listens to, rather than ignores or dismisses, the citizens' response. Neutrality is displayed when officers explain to citizens how their decisions were made based on evidence or facts rather than any personal characteristics of the parties involved; transparency and impartiality are key. Dignity is shown when an officer calls a citizen "ma'am" or "sir" or uses other polite speech. Finally, trustworthy motives are displayed when an officer expresses concern for a citizens' well being by asking how s/he is doing or offering him/her some kind of assistance (Tyler, 2004).

The four elements of procedural justice, manifested in the behaviors cited above

and others like them, have consistently been found to influence citizens' perceptions of procedural fairness (Tyler, 2004). Relatedly, it has been found that when citizens perceive procedurally fair or just treatment by police, they are more likely to believe that police deserve the authority bestowed upon them; when citizens feel that police authority is deserved, they likewise feel a sense of responsibility to cooperate with officers when given the opportunity (Hinds & Murphy, 2007; Tyler, 2004). Such is the argument that demonstrating procedural justice can increase police legitimacy and improve police-community relations. Importantly, though, while evidence of this relationship between procedural justice and legitimacy has been cited by researchers for more than a decade, it must be acknowledged that how and how much procedural justice is shown to citizens is a matter of police discretion (Mazerolle, et al.,2013; Gau, Corsaro, Stewart, & Brunson, 2012; Sunshine & Tyler, 2003).

An officer's discretion, or "the leeway that officers enjoy in selecting from more than one choice in carrying out their work," (Mastrofski, 2004, pp. 101) is perhaps the most widely discussed facet of American policing. While police officers in the same city or state may receive uniform training and be taught identical techniques and strategies, the heterogeneous nature of human beings ensures that no two police-citizen-encounters are identical and thus there is virtually always at least some room – and necessity – for police discretion. As Mastrofski and colleagues (2016) explain, decisions about how and how much procedural justice to show to a citizen is subject to police discretion since procedural justice itself is essentially another tool for achieving goals and completing the tasks of the officer's job. Not only can procedural justice be used to elicit compliance,

but, as Mastrofski et al. (2016) explain, it serves the expressive purpose of allowing officers to convey their feelings about a citizen.

Importantly, while all police discretion exists because there are not exhaustive policies or guidelines to shape an officer's behavior in every conceivable circumstance, the street-level discretion to use procedural justice is especially broad; very few formal controls exist to direct its use (Mastrofski et al., 2016). This immense freedom that officers have in deciding how much procedural justice they show citizens is one reason why its examination through research is so valuable; among other things, research can shed light on which citizen characteristics or situational variables of an encounter predict the amount of procedural justice displayed by an officer. Knowledge of which factors influence this discretionary decision, and to what extent they do so, will help police leaders and policy makers create guidelines that can begin to shape this discretion so departments can heed the advice of the President Obama's task force and others who suggest that procedural justice is such an effective way to improve police legitimacy and police-community relations.

Procedural Justice Research: Beginnings, Improvements, and Current Directions

Unfortunately, though police discretion itself has been widely researched for many years, the research has mostly been concerned with more coercive police actions, which is intuitive since an officer's authority to take coercive and proscriptive action is said to set him/her apart from other service providers (Bittner, 1974). Studies examining police use of force, as well as arrests, have been more prominent than those examining non-coercive police action (Dai & Nation, 2009; Rabe-Hemp, 2008; Skogan & Frydl,

2004; Worden & McLean, 2014). For this reason, studies of procedural justice coming into the central focus of policing research is a somewhat new phenomenon that gained momentum following the nationwide shift toward community policing – a philosophy and style of policing that aims to build partnerships between local law enforcement agencies and their associated communities in order to enlist citizen cooperation that will allow for proactive, problem-oriented policing and organizational structures that are conducive to it (U.S. Department of Justice [DOJ], 2014).

With this policing paradigm shift in the 1980s came the shift in researchers' focus to what motivates citizens to participate in community policing. This led to much examination of procedural justice as a possible influence on police legitimacy – a greater focus in early years than what influenced the discretionary decision to show procedural justice – and survey and interview data were widely used to explore this relationship. One early study, for example, was completed by Sunshine and Tyler (2003) who collected survey and interview data from two samples of New York City residents to examine what factors influence police legitimacy and whether police legitimacy itself influences public support and cooperation with police. In one survey, Likert scales of level of agreement were used to assess perceived obligation to obey police – a measurement of legitimacy – and general questions about perceived fairness were used to measure the citizens' perceptions of procedural justice from officers. For example, the item, "Do police treat people fairly," was used in one survey (Sunshine & Tyler, 2003).

While data from studies like these allow researchers to understand the respondents' perceptions, which are of great importance since legitimacy depends

9

entirely on how citizens *perceive* police authority and treatment, there are some inherent issues with early studies like these. For instance, although it can be seen from Sunshine and Tyler's (2003) data whether a citizen who believes that officers treat people fairly is also a citizen who believes that police deserve their authority, very little can be taken from this finding and applied to policy or departmental guidelines that will improve policing practice. These statements speak little to how officers should actually behave and they do not consider dosage – the amount of procedural justice displayed – at all. Additionally, there are reliability issues associated with asking different citizens general questions about whether treatment was "fair;" what one citizen considers fair may be viewed as injustice by another. In fact, Kulik, Lind, Ambrose, and MacCoun (1996) found evidence of gender differences in how fairness is defined and perceived by litigants involved in auto negligence lawsuits.

Fortunately, observation data has allowed procedural justice researchers to bypass some of these weaknesses associated with interview and survey data. Systematic social observation – the method of data collection used in the present study – is particularly valuable since observers have been trained to follow a predetermined set of rules and protocols for data collection and recording and have also been provided with specific definitions of sometimes-ambiguous terminology that will be used in the coding instruments, which often take the form of questionnaires to be completed by the observers (McCluskey, Parks, & Mastrofski, 2014). Unlike the surveys used by Sunshine and Tyler (2003), for example, observers would define a single definition of "fairness" in advance so that it is reported consistently.

In terms of maximizing the usefulness of the data collected, findings from observational research can be more easily applied to policing strategies since third-party observers are able to provide more objective measures of police behavior (Jonathan-Zamir et al., 2015). These objective measures help to determine which behaviors constitute procedural justice and how much of it is shown. Furthermore, observers are able to provide relevant context of police-citizen encounters, which must be considered when making statements about what influences an officer's discretionary decision to show procedural justice. Such context may include the physical setting of the encounter, the disposition of those involved, the original problem associated with the encounter, and how it was initiated. Aggravating factors like the presence of a belligerent citizen or an overwhelming number of bystanders may lead an officer to abandon his or her typical patient disposition while mitigating factors like an especially respectful citizen or a quick problem resolution may improve the officer's demeanor, leading to a greater than average display of procedural justice (DeJong, 2004; Mastrofski et al., 2002; Rabe-Hemp, 2008).

Likewise, officer and citizen demographics also influence procedural justice and have been included in past research, though gender has typically only been the main focus of studies examining how procedural justice varies by *officer* gender (Rabe-Hemp, 2008; Gale, 2001). Still, citizen gender has been considered in many observational studies, but it has generally been in a group of other variables acting as a control and, as was previously mentioned, in very little detail. For this reason, the most valuable contribution of the present research is offering an examination of gender differences in

the context of the chivalry and selective chivalry hypotheses, which are discussed below along with common gender-related findings in extant criminal justice research.

Chivalry in the Criminal Justice System

The differential treatment of males and females has certainly not been ignored in other research related to the criminal justice system. As with many studies of police discretion, however, it is often the more dramatic settings or outcomes that receive the most attention. For instance, arrest rate and sentencing disparities across citizen gender have been widely examined in order to determine whether one gender fares more favorably than the other when both are suspected of the same wrongdoing. Supporters of the chivalry hypothesis – sometimes referred to as paternalism – argue that females receive more favorable treatment from police, as well as the court system, which is reflected in a lower likelihood of arrest and in more lenient sentencing following conviction (Kulik et al., 1996; Visher, 1983). The premise of this chivalry thesis is that criminal justice procedures will systematically protect women because the chivalrous and paternalistic relationships that are said to exist between men and women in everyday life will also translate to the justice system where male-dominated police forces and courtrooms are dealing with female citizens. Since women are viewed as fragile, overly emotional, and thus less accountable than men and in greater need of paternalistic protection, male agents of the justice system will attempt to protect female offenders and offer them consideration and leniency that would not be afforded to their male counterparts (Herzog & Oreg, 2008; Rodriguez, Curry, & Lee, 2006).

Two import notes about the gender stereotypes that underlie this chivalry theory

must be made. First, while gender norms and the male-dominated justice system may make the chivalry dynamic more observable between a male authority and female suspect or citizen, evidence suggests that both men and women hold the underlying stereotypes of the chivalry bias, which allows it to operate between female authorities and female civilians, as well (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972; Glick & Fiske, 1996; Kulik et al., 1996). This is an important consideration for more modern research of a justice system with an increasing population of female police and court officials.

Second, as Embry and Lyons (2012) explain, the influence of these gender stereotypes and biases on the chivalry dynamic causes some to argue that the chivalrous or paternalistic treatment of women is not without conditions. In order for it to persist, the women involved must adhere to traditional gender roles and stereotypes whereby they are feminine, vulnerable, and submissive. If the woman deviates from her traditional female role, she is less likely to receive the special treatment (Herzog & Oreg, 2008). This conditional alternative to the chivalry hypothesis is referred to as selective chivalry, also sometimes called typicality,[1] and different researchers have interpreted it in slightly different ways. While some explain selective chivalry as the notion that preferential treatment is only available to women who do not violate gender expectations and that those who do violate them are treated "more harshly, perhaps similar to men" (Rodriguez

---

[1] In 1995, Farnworth and Teske differentiated between the typicality hypothesis and selective chivalry, referring to typicality as the notion that women are treated with chivalry only when their charges fit stereotypic female gender norms and to selective chivalry as this idea that the justice system extends chivalry disproportionately to White women. They found little support for either of these theses and recent research has since used the terms interchangeably and separated selective chivalry from Farnworth & Teske's (1995) racial qualification.

et al., 2006, pp.322), others equate selective chivalry to the evil woman hypothesis and claim that when a female's crime violates gender norms, she will be treated *more harshly* than a man who committed the same crime (Embry & Lyons, 2012).

These explanations of selective chivalry in the context of female offending explain why the limited research on selective chivalry has often focused on sentencing disparities, suggesting that leniency and consideration will only be shown to women who commit crimes that are "'typical' of females and stereotypic female gender roles" (Rodriguez et al., 2006, pp.322), such as shoplifting and check forgery. Such research has produced very mixed results, however, and offered greater support for the general chivalry hypothesis than selective chivalry specifically (Emrby & Lyons, 2012; Rodriquez et al., 2006). When considering the effects of adherence to (or violation of) gender roles and stereotypes in other forms – other than crime type – and in other settings – other than post-conviction – however, some support has been found.

For instance, in her study of drug arrest trends in the Narcotics Division of the Chicago Police Department (CPD), DeFleur (1975) found some evidence that officers are more likely to show leniency to female suspects when their behaviors, attitudes, or remarks conform to gender stereotypes. Seeking to identify the influence of biasing factors on official arrest rates in order to assess their validity for research, DeFleur (1975) examined over thirty years of drug arrest records, noting very sudden or extreme shifts in the volume and types of arrests (and arrestees) and used qualitative data to explain them. DeFleur (1975) studied police records and interviewed retired narcotics officers about the policies, procedures, and goals that may have influenced the shifting trends. She then

began field research at the Narcotics Division of the CPD, spending months as a participant observer, accompanying officers on surveillance missions, raids, and to the courtroom. During this time, she also attended what she calls "after work drink sessions" with the officers who would share with her their habits, values, and beliefs, as well as the kinds of pressures they felt from the community and organization to deal with certain kinds of drug offenders in certain ways (DeFluer, 1975).

While she writes that she cannot "quantify the amount of influence of these factors on the drug statistics studied" (DeFleur, 1975, pp.94), she describes their general impact and presents some evidence of selective chivalry. DeFleur (1975) found that "there was a tendency not to arrest females as often as males if they behaved in expected, stereotypic ways" (DeFleur, 1975, pp.101) like crying or claiming to have been led astray by men. The crying behavior conforms to the female stereotype of being emotional and perhaps the claims to have been led astray by men also contributed to the related view of women as being less accountable for their actions, as mentioned above. Further, DeFleur (1975) notes that she observed, and also learned in interviews with officers, that those females who were aggressive and hostile tended to be processed and arrested more often than those who behaved in traditional ways.

Likewise, Visher (1983) also claimed to find some support for this selective chivalry, though she took a somewhat different approach. Using observation data originally collected during a large-scale evaluation of police services at twenty-four American police departments in 1977, she compared the structure of the arrest process for 142 females to that of 643 males, estimating the relative effect of a number of personal

and situational variables (e.g. race, age, demeanor, presence of bystanders, etc.) on the probability of arrest.[2] She found only a 4% difference in the overall arrest percentage for males and females, showing little support for the general chivalry hypotheses, but noted that police officers (95% of whom were male in her sample) use different criteria for their arrest decisions for male and female suspects, considering a wider range of situational factors when dealing with males, and putting a greater emphasis on personal characteristics – like age and race – when dealing with females.

Visher (1983) suggested that this finding may reflect the influence of officers' gender expectations on arrest decisions of female suspects, and argued that women who did not conform to "typical middle-class standards of traditional female characteristics" (pp.22) were not afforded the same leniency as those who did conform. Visher (1983) defined this "standard" as White, older, and submissive, citing evidence of police tendencies to treat younger females more harshly than older females and arguing that Black females may be less likely to display traditional gender behaviors because they are more often the heads of their households, making them more independent and of equal or greater status position than the males with whom they live.

Although Visher (1983) did find that "young, Black, or hostile women received no preferential treatment, whereas older, White women who are calm and deferential toward the police [were] granted leniency" (pp. 23), she provided no research evidence that Black females actually behave in less traditionally feminine ways than White females

---

[2] Visher (1983) explains that her original dependent variable, arrest decision, violated the assumptions of the original least squares model since it was both dichotomous and highly skewed. Thus, the dependent variable became the probability of arrest – still related to a set of indepdent variables as in the ordinary least squares model – and her new model assumed the form of a probit probability density function.

or that this was the cause of the race effect in her study. Furthermore, although her finding that hostility increases a female citizen's probability of arrest supports selective chivalry's notion that women who violate gender norms (in this case showing hostility rather than the expected submissiveness) will not be shown leniency, the even larger effect of hostility on a male's probability of arrest makes this finding less compelling.

Other studies have found more convincing support for the general chivalry hypothesis, though they have also been limited mostly to decisions made in the courtroom. For instance, in their meta-analysis of studies of adult sentencing decisions published between 1991 and 2010, Bontrager, Barrick, and Stupi (2013) sought to determine whether sentencing outcomes are related to gender and whether the methodological characteristics of the studies impact their findings. From the 58 studies included in their analysis, 143 unique estimates of the relationship between gender and sentencing outcomes were provided. The four groups of sentencing outcomes included were the decision to incarcerate, the incarceration length, ordinal representations of sentence severity, and sentencing departures (Bontrager et al., 2013). To be included in the meta-analyses, the studies needed only to have examined the relationship between gender and one of these outcomes.

To account for the methodological differences, Bontrager et al. (2013) grouped the studies according to how many of these three control variables were included in their model: the offender's prior record, the offense type, and the offense severity. They then counted the number of unique gender and sentencing estimates in each of these four groups (those that controlled for 0, those that controlled for 1, those that controlled for 2,

and those that controlled for all 3) that yielded statistically significant support of the chivalry hypothesis. They found that even studies that controlled for all three variables produced estimates consistent with the chivalry hypothesis in 62% of cases and that this support was mostly related to the outcomes of incarceration decision, sentence length, and sentence departure (Bontrager et al., 2013).

Studies of sentencing decisions made prior to those included in Bontrager et al.'s (2013) meta-analysis found similar evidence that women receive more lenient sentences for the same crimes as their male peers, even when controlling for race, age, and criminal record (Daly & Bordt, 1995; Moulds, 1978). Further, marginal support of more favorable outcomes for women has been found in the civil courtroom, as well, where exit survey data from litigants in auto negligence lawsuits reveal that women reported receiving slightly higher rewards than men and perceiving more control during the arbitration process (Kulik et al., 1996).

While most of these studies referenced above specifically examined gender differences in discretionary justice system decisions and viewed them through the lenses of the chivalry or selective chivalry hypotheses, research on police officers' discretionary decisions to show procedural justice has yet to do the same. What these studies have found about the effects of citizen gender and police-provided procedural justice – though the relationship between these variables was neither framed in the context of the chivalry thesis or of central focus in these studies – is included below.

Gender Effects in Observational Procedural Justice Research

Using data collected during systematic social observations of 442 shifts at the Cincinnati Police Division in 1997 and 1998, Dai and Nation (2009) compared coercive and non-coercive police behaviors displayed by officers in situations involving domestic and non-domestic conflict. Although they were not measuring specifically the four elements of procedural justice, one of the non-coercive behaviors they measured – police care – used four indicators that are also included in the present study to measure the trustworthy motives element of procedural justice. Police care was a dummy variable in their analysis that measured whether an officer contacted an agency on the citizen's behalf, provided physical assistance to the citizen, provided information on how to deal with the problem, or comforted and reassured the citizen (Dai & Nation, 2009).

In addition to finding that citizen disrespect greatly influences officer behaviors and that police officers behave differently in situations involving domestic conflicts, Dai and Nation (2009) found that officers are more likely to provide care and assistance to female citizens than male citizens and that female citizens are also less likely to be subjected to verbal force or arrest when other things are equal. These findings related to police care suggest that the favorable treatment women receive from the justice system is not limited only to arrest or sentencing decisions.

More recently, Mastrofski et al. (2016) completed an observational study of police-citizen encounters at two departments to examine predictors of police-provided procedural justice specifically. They were especially interested in how the moral worthiness of the citizen, the circumstances that make procedural justice difficult (e.g. citizens at the scene in conflict with one another, a high number of bystanders present at

the scene, etc.), and the popular scripts for handling situations (e.g. those typically used in traffic stops) predicted procedural justice, but they examined the influence of officer and citizen characteristics in their model, as well. Importantly, Mastrofski et al. (2016) used systematic social observations for data collection and measured procedural justice as a single variable using Jonathan-Zamir et al.'s (2015) formative index, on which the one used in the present study is closely based.

Analysis of about 600 citizen interactions revealed that victims or people in need of assistance received more procedural justice than suspects and witnesses; the number of citizens present at the scene and the amount of "people work" (previous encounters) the officer had performed earlier in the shift had negative effects on the amount of procedural justice shown; traffic encounters were associated with higher levels of procedural justice than other types of encounters; and the officer's status as backup had a suppressing effect on procedural justice (Mastrofski et al., 2016). While male citizens were found to experience less procedural justice than female citizens, this relationship was not statistically significant, nor were the effects of any other officer or citizen characteristics. This finding about the relationship between citizen gender and procedural justice as a whole is quite different from what Dai and Nation (2009) found when examining only police care, which is related to but one of the four procedural justice elements.

Worden and McLean (2014) also used observation data – coded from dash-mounted camera recordings – to examine many of the same predictors included in Mastrofski et al.'s (2016) model, though they were concerned with how (if at all) departments are using procedural justice and whether citizens' self-reported perceptions

reflect the amount of procedural justice actually displayed to them by police. In Schenectady, New York, researchers trained in systematic social observation watched and coded 539 police-citizen encounters, which were sampled from approximately 1800 encounters about which citizens had already been interviewed (Worden & McLean, 2014). Using the coded observational data (rather than the interview data) for analysis, the researchers found that more procedural justice was shown when an encounter involved a violent crime or interpersonal conflict and when the citizen was Black. In contrast, procedural justice was lower when the citizen was a suspect or third party, and when he or she was resistant toward police. Finally, procedural *injustice* was found to be greater when the citizen was a suspect, disrespectful, passively resistant, or male (Worden & McLean, 2014). While Worden and McLean (2014) used a separate measure – procedural injustice – to account for discourteous behaviors shown by the officer, their finding that males would be subjected to more of these behaviors is consistent with the chivalry hypothesis.

Current Study Implications

The nature of the gender-related findings in the studies cited above and the ways in which they were examined reveal the need for further research. Although most of these studies reported gender-related findings that are consistent with the direction one could expect based on the chivalry hypothesis, there is inconsistency in the statistical significance of the results, especially among the procedural justice studies. Further, as was stated above, none of these studies of procedural justice focused specifically on gender effects or examined them within the theoretical context of the chivalry or selective

hypotheses, as have studies of other discretionary decision-making in the criminal justice system. Examining gender effects in these theoretical contexts would have led to more focused research questions and hypotheses (like numbers 2 and 3 below), as well as consideration of variables related to conformity or violation of gender norms that likely would have been excluded otherwise. Consideration of these variables could have helped to disentangle some of the inconsistencies in past gender-related findings.

The present study draws on methodological strengths of existing research and couples them with a stronger theoretical foundation in order to contribute more meaningful findings to the discussion of how discretionary procedural justice decisions are shaped and how/whether men and women are treated by the criminal justice system. Per the chivalry and selective chivalry hypotheses, I expect and will test the following:

- **Hypothesis 1**: Female citizens will receive more procedural justice than male citizens when other personal and situational variables are held constant.

- **Hypothesis 2**: Female citizens who demonstrate a need for police assistance will be shown more procedural justice than males who demonstrate a similar need when other personal and situational variables are held constant.

- **Hypothesis 3**: Female citizens who show resistant behavior toward police will receive less procedural justice than will male citizens who show similar resistant behavior when other personal and situational variables are held constant.

An important note must be made about these latter two, which are informed specifically by the selective chivalry hypothesis and will determine whether selective chivalry is at work by examining the interaction between gender and need, a condition that fits the

female stereotypes on which selective chivalry is built, and the interaction between gender and resistance, a behavior that violates gender norms and, thus, should compromise any potential consideration that may have been shown to a female citizen according to the general chivalry theory. For Hypothesis 3, in particular, it should be noted that the view of selective chivalry used by Embry & Lyons (2012) and others is assumed here and that it is expected that women who violate gender norms will simply not be treated the same way as men, as Rodriguez et al. (2006) suggested, but that they will be treated with *less* procedural fairness than men. Thus, while the general chivalry theory suggests that all women will receive more favorable treatment than all men, the expectation under selective chivalry is that only women who conform to traditional gender expectations will receive more favorable treatment than men and that those who violate gender expectations will receive less[3].

---

[3] Both the chivalry and selective chivalry hypotheses speak specifically to how the justice system treats females and make no statements about the treatment of males. Thus, no predictions are made here about how male adherence to or violation of gender norms will affect their treatment by the justice system.

METHOD

This chapter describes the data and methods used in the present observational study of the effects of citizen gender on the amount of procedural justice shown by police. First, the data source and collection sites are described, then explanations of the unit of analysis, sampling procedure, and coding instruments are provided. Next, all of the variables included in the regression model are discussed and an explanation of the statistical model is included in the final section.

Data Source

To test the aforementioned hypotheses, data from a systematic observation study of police-citizen encounters at the Gallanton Police Department (GPD) are used[4]. The GPD is a large department (over 1,000 full-time sworn personnel) that serves an ethnically diverse community and has somewhat lower violent and property crime rates than other large agencies in the United States. Data were collected at two divisions – Ellsworth and Camden (pseudonyms) – of the GPD in the months of August and September 2015.

Collection Sites

Systematic observations were conducted first in the Ellsworth Division of Gallanton and about one month later in the Camden Division. Both divisions serve

---

[4] Please see the disclaimer about this larger study in Appendix D.

populations that are made up of mostly racial/ethnic minorities. Although Ellsworth is more than double the geographic size of Camden and has a population that is about 1.5 times its size, Camden employs nearly 40% more patrol officers. Of all Ellsworth patrol officers who were eligible to be observed in this study, about 86.3% were male and 13.7% were female. Of those eligible for participation in Camden, 89.6% were male and 10.4% were female. Both Ellsworth and Camden experience higher crime than the average division in Gallanton, though this is mostly the case for Ellsworth because of the amount of what the FBI call's Part II crimes (e.g. simple assaults, disorderly conduct, drug crime, etc.) that take place in the division; Ellsworth is near the Gallanton average for Part I crimes (e.g. murder, robbery, forcible rape, etc.). Camden, however, has higher-than-average Part I crime and Part II crime compared to the rest of Gallanton.

For over a decade, department-wide efforts have been made by Gallanton police and leaders to improve public trust. The ultimate goals have included promoting and respecting civil rights, better handling use-of-force incidents, and increasing community outreach and transparency, among others. The two individual divisions also publish divisional goals statements and in Ellsworth, improving community relationships and increasing community trust are cited as top priorities. In Camden, improving specific facets of law-enforcement and crime-reduction performance (e.g. report writing, timely response, and crime analysis) are highlighted but there are no expressed goals about community trust or relationships beyond improving community education and involvement. Neither the department as a whole nor either of these divisions specifically cites procedural justice in its goals statement or provides specific details about how to

achieve its goals through methods that reflect a procedural justice perspective, but some

of the goals do reflect procedurally just values. With no expressed focus on – and no

formal definition of – procedural justice, however, Ellsworth and Camden are

comparable to other police agencies in the United States that may not be formally

promoting procedural justice initiatives specifically, but are still aiming to improve

police-community relations.[5]

Procedure

Unit of Analysis

The unit of analysis in the present study is the interaction that takes place between

an officer and each individual citizen involved during an encounter. For police contact

with one or more citizens to constitute an encounter, an officer must have had

communication that lasted longer than one minute in duration or included a minimum of

three verbal exchanges with at least one citizen present.[6] If multiple citizens were present

and the officer had sufficient contact (that would constitute an encounter) with all of

them, each citizen was involved in a police-citizen interaction (the unit of analysis) and

would have been recorded as an individual case. Thus, more than one citizen interaction

could have occurred during a single encounter, which happened in just over half of the

observed cases. Of the 243 police-citizen interactions included in this study, 46.1% of

---

[5] It should be noted that shortly after data collection was finished, the department at which these observations were completed introduced a new training program for officers that was more focused on specific procedural justice initiatives than many of the other steps it had taken in the past. While the term "procedural justice" was not used, there was a focus on encounter-level behaviors to improve public trust and many of them were closely related to procedural justice elements, especially dignity and participation.
[6] Other police or civilian service personnel (e.g. emergency medical staff) at the scene were not considered citizens.

citizens were involved in single-citizen encounters, 28.8% of citizens were involved in

two-citizen encounters, and 25.1% of citizens were involved in encounters with three or

more citizens.[7]

It must be noted, though, that in order for these police-citizen interactions to have

been included in the present study, they must have been of sufficient duration *and*

intensity for the citizen to have been a "primary participant," or person about whom

decisions could have been made, assistance could have been provided, or on whose

behalf police might have taken action. Thus, per the one-minute/three-exchange standard

described above, an officer may have been engaged in an encounter with a citizen if s/he

spoke to the citizen for more than one minute or had three or more verbal exchanges with

the citizen, but if these were such superficial exchanges that they did not give an officer

the opportunity to make decisions about, provide assistance to, or otherwise act on behalf

of the citizen, there was not sufficient *intensity* for that citizen to constitute a primary

participant. In this case, observers would not have recorded procedural-justice-related

details on the coding forms and thus, the dependent variable in this study would not have

been measured. An example of such a case might be when an officer is at a pawnshop

taking a report about stolen goods from the shop owner. While the owner is gathering

paperwork for the officer, a customer briefly chats with the officer about a guitar hanging

on the wall. Even if the officer and customer had three or more exchanges, the limited

nature of the conversation did not give the officer an opportunity to take action or make

---

[7] After completing the main analysis, a sensitivity test was completed to determine whether involvement in
encounters with large groups (more than three) of citizens affected the amount of procedural justice shown
to a single citizen by the primary officer. The results of this test can be found in Table 6 in Appendix C.

decisions on behalf of the citizen and s/he would not have been a primary participant.

To be clear, however, the officer need only the *opportunity* to provide assistance or make decisions on behalf of a citizen for him/her to be a primary participant; the officer does not necessarily need to *take* the potential action or *make* a potential decision. Another example may be helpful: Assume that the observed officer reports to a two-car traffic accident in a residential neighborhood. When she reports to the scene, she determines whether anyone is hurt and speaks to both drivers about what happened. The drivers, Citizen 1 (C1) and Citizen 2 (C2), and the single passenger, Citizen 3 (C3), all insist that they are fine, though shaken up, and that they do not want any medical attention. While O1 speaks to the three and helps C1 and C2 exchange insurance information, a man who is walking his dog passes by and says hello to the officer who waves to the man as he continues down the road. In this example, C1, C2, and C3 all had sufficient communication with the officer for each to have experienced their own police-citizen interaction and since they all could have potentially been assisted (or had decisions made on their behalf) by the officer, the officer had the opportunity to display procedural justice to them and thus they all would have been included in analysis as individual cases. This is true even though the officer only assisted C1 and C2 (by helping them to exchange insurance information) while C3 refused assistance (in this case, the officer's beckoning of an ambulance or other medical help). In contrast, since the man walking his dog had insufficient contact with the officer to have been included, he would have been considered a bystander – someone without the need or desire for assistance or

any other involvement in the situation – instead of a citizen participant (hereinafter referred to only as citizens).

This study's sample of 243 police-citizen interactions (or individual cases) came from a total of 182 encounters and 55 observation sessions (of typically 6-hour duration). The 243 cases were derived from a data frame that originally included 445 police-citizen interactions, 239 encounters, and 57 observation sessions, meaning that just under half of the citizens with whom the observed officers had interactions did not have sufficient contact with those officers to have been viewed as "primary participants." Procedural justice-related data likewise would not have been recorded for these 202 individuals and they were thus excluded from the final data frame.

Sampling

During this preliminary stage of the greater study from which these data are taken, the researchers aimed to observe as many patrol officers as possible – who were not on special assignment (e.g. specialized motorcycle enforcement patrol or K-9 units) – in each of the two divisions. To accomplish this goal, all eligible officers were added to a digital data frame and sorted with a computerized randomization tool. This digital data frame included information about when officers were scheduled and which ones were on vacation or in court on a given day. Before beginning an observation session – which either took place during the first or last six hours of an officer's twelve-hour shift – an observer would be assigned whichever officer was at the top of the list, was scheduled during the current watch (shift), and was not otherwise unavailable (i.e. in court or on vacation). If the observer reported for observation and the assigned officer was not on

duty for any other reason (e.g. s/he had gone home sick), the next eligible officer on the list was assigned to the observer and the original officer returned to his/her place on the list, where s/he would remain until s/he was available for observation.

The officer who was actually observed (not necessarily the one to whom the observer was originally assigned) is referred to as O1. This abbreviation is used in the narratives written by observers, in the coded forms completed by observers, and will be used below to distinguish between the observed officer and other officers present during police-citizen encounters. Relatedly, it must be noted that although an observer was only assigned to a single officer (O1) during a given observation session, officers at the observed divisions complete patrol shifts as two-officer units so O1 almost always had a partner (O2) present. Although some officers choose to patrol with the same partners whenever possible, it is also quite common for officers to ride with different members of their squad on consecutive shifts. For this reason, although an officer would not have been selected as an O1 more than once, s/he may have been present during multiple observation sessions, as s/he may have been the partner to one or more observed O1s. Being present as the partner to an O1 (observed officer) did not affect an officer's place in the data frame whatsoever, though his or her actions during the observation session presumably influenced the observed outcomes and were documented in some capacity in both the written narratives and the coded data. To understand how the actions of O1 and his/her partner (O2) during police-citizen encounters were observed and documented, the general structure of the data collection and recording process must be explained.

Observer Training and Coding Instruments

30

As stated above, systematic social observations (SSO) were used for data collection and this form of field research requires observers to follow a predetermined set of detailed rules and protocols for recording observations (McCluskey et al., 2014). To learn these rules and protocols, the observers, three of whom were male and two of whom were female,[8] spent a week in Gallanton about two months prior to research collection learning about the coding instruments and about which aspects of the police-citizen encounters were of greatest interest. All five researchers had at least a four-year college degree, four holding degrees related to law or criminal justice, and all five had experience with criminal justice or law-related research, two having prior experience observing police in the field.

During their week of training in Gallanton, the observers learned protocols for observations (e.g. where to stand to see and hear interactions, what to do in emergency situations), learned how to write detailed narrative accounts of their observations – the qualitative component of the larger study, and carefully examined the coding forms as a group, taking time to discuss any confusion or ambiguities. Individually and as a group, the observers viewed videos of police-citizen encounters and practiced coding their observations of these videos. To ensure coder reliability, one of the primary investigators led group discussions about the videos and polled the observers about what they reported on the coding forms. When there were discrepancies, they discussed them and clarified the appropriate responses. The observers also completed one brief ride-along (about two

---

[8] Although nine observers (four male and five female) collected data during this preliminary stage, data from only five observers are used in the present study. Data from the other four observers were not included due to unavailability at the time of data analysis for this thesis and due to an insufficient number of interactions collected.

hours in duration) each with a supervisor from either division (some observers rode in Ellsworth while others rode in Camden) in order to learn a bit about the geography, crime, and typical patrol style in the respective division. When the observers returned to Gallanton about two months later to begin data collection, they spent two more days reviewing protocols and getting reacquainted with the coding forms, which were undergoing final changes.

Chronologically, the first of four forms completed by the observers, which will be referred to as the SSO Light form, was completed on a Microsoft Surface tablet while the observers were in the field collecting data. After observing an encounter and (optionally) taking brief hand-written notes, observers documented very specific details about individual events on these forms. Demographic data about citizens, information about the number of police present, and times of events were coded on these forms, which were generally completed inside of the patrol car between encounters. An SSO Light Form was completed for every single police-citizen encounter and was later used, in combination with hand-written notes and the more lengthy narrative accounts, to complete the remaining three forms.

Full Encounter Forms were completed for every police-citizen encounter, defined, as stated above, as contact with a citizen(s) that was longer than one-minute in duration or included at least three verbal exchanges, and included more detailed information than what was recorded on the SSO Light Form. These forms were completed after observation and, generally, after narrative accounts were written, as well. Unique identifiers on both forms (Full Encounter and SSO Light) allowed the two to be linked

for subsequent analysis.

The Full Citizen Form, the longest of all four coding forms, was completed for every single citizen observed in a police encounter (according to the one-minute or three-exchange standard described above). Thus, the details coded on these lengthy forms describe each unique police-citizen interaction, more than one of which may have taken place during a single encounter if multiple citizens were present (e.g. at the scene of a traffic accident in which an officer interacted with multiple drivers). Details about specific exchanges between the officer and citizen, like whether they spoke the same language and whether they were engaged in various levels of physical or verbal conflict, were coded on these forms, as were all of the items on the formative index used to measure procedural justice in the present analysis. Lastly, Full Ride Forms were completed for each observed shift. Basic information about observation time and O1 and O2 were coded on this form. Data from these Full Ride Forms were not used in the present analysis but their unique identifiers allowed them to be linked to the other forms.

Relevant Coded Data

As stated above, observers were assigned to a single officer, O1, but his or her partner was also generally present and interacting with citizens during encounters. In the event that O1 and his/her partner separated during an encounter and the observer was unable to see and hear both of them interacting with citizens, the observer was instructed to remain with O1. In cases in which O1 and O2 were both present, however, observers were instructed to report in the narrative and coding forms who acted as the "primary officer" in each individual encounter.

The primary officer was defined as the officer with whom the citizens in a given encounter had the most significant interaction, a standard usually evaluated according to the duration and intensity of the officer's interaction with that citizen; examples of significant interactions include an officer providing assistance to a citizen, questioning him/her, and making decisions concerning the citizen (e.g. arrest decisions) among other things. Identification of the primary officer was important because most of the procedural justice-related questions on the coding form asked about behaviors and actions of this primary officer, though some referred specifically to O1 or to police more generally. In some cases, another officer at the scene acted as the decision-making leader (meaning s/he ultimately decided how to resolve the situation, or at least had the greatest influence on choosing the resolution), but either O1 or O2 was always identified as a primary officer for coding purposes since observers were instructed to stay with O1 and may not have seen other officers' interactions at the encounter if they were not within view/earshot. If O1 and his/her partner operated in an equal capacity, observers defaulted to O1 as the primary officer and answered the related questions accordingly.

These coding guidelines were established in order to manage the unique challenges presented when SSO is used to collect data about individual officers who patrol as partners. It was the aim of the researchers to capture the most relevant interactions between police and the citizens present at encounters, as well as to include the most meaningful demonstrations of procedural justice (or failures to demonstrate it), which would not be possible if only O1's actions were reported on the coding forms. For example, if O1's partner (O2) initiated contact with a citizen, conversed with the citizen,

and subsequently decided to arrest the citizen while O1 mostly served as backup, documenting only O1's actions would grossly misrepresent what actually transpired. Thus, observing and coding details about the primary officer's behavior is sensible.

It must be noted, though, that when reporting the actions of the primary officer rather than O1 exclusively, it is possible that an officer's procedurally just (or unjust) actions will be observed and reported during more than one observation session if s/he acted as the primary officer when s/he was observed as O1, as well as when s/he was the partner of another O1. Behaviors of 5 of the 55 unique O1s (9.1%) were recorded during separate observation sessions, as these officers acted as the primary officers in encounters during which they participated as both O1 and O2. In contrast, there were several other instances in which an officer was sampled and observed as O1 but no data related to his or her procedurally just behaviors were recorded because s/he never acted as the primary officer. This was true for 6 of the 55 (10.9%) unique O1s observed. It is worth noting, though, that in 4 of these instances, the observation session only included 1 or 2 encounters that met this study's criteria for inclusion in the final data frame.

Variables

Dependent Variable

The dependent variable in this study is, of course, the amount of procedural justice displayed to citizens by the primary officer during police-citizen interactions. Procedural justice was measured as a formative index based closely on the one constructed and validated by Jonathan-Zamir et al. (2015), which was previously used by Mastrofski et al. (2016) to measure procedural justice in another observational study of

police behavior. The index was based on the average, standardized score (0-100) of the four elements of procedural justice – participation, neutrality, dignity, and trustworthy motives – to ensure that each element received equal weighting in the overall measure. While slight differences in the instrument used for data collection in the present study did not allow for identical calculations of each element, the indicators used were kept as similar as possible.[9] The indicators used to measure the element of participation, for instance, were related to the officer's effort to solicit information from the citizen about his/her view of a situation and the way in which the officer listened to the citizen's response if one was provided. An officer who asked for the citizen's side of the story and then listened well to his/her response (by paying close attention and/or asking questions, for example) would have encouraged more citizen participation than an officer who never asked the citizen to contribute and/or ignored or dismissed the citizen's response. Thus, the first officer would have scored higher on the participation index and on the procedural justice index overall, other things (elements) being equal.

Neutrality was also measured with a series of dichotomous indicators, which were related to the effort made by the officer to explain his or involvement in an encounter and how s/he would or did make an impartial decision about the handling of a given problem. Such an effort would include explaining to the citizen justifications for his/her (the officer's) actions based on the law, evidence, public safety, or the citizen's welfare rather than personal characteristics of anyone involved. An officer who explained his/her actions and communicated his or her objectivity would have scored higher on the element

---

[9] See appendix A for details.

of neutrality and would have shown more procedural justice than one who did not bother to provide the citizen with any explanations, other things being equal.

The element of trustworthy motives was measured primarily with dichotomous indicators related to behaviors demonstrating care and concern for the citizens involved (e.g. whether the officer offered comfort or assurance to the citizen or whether the officer provided or arranged physical assistance to/for the citizen). The more of these caring behaviors were shown by an officer, the higher s/he would have scored on the trustworthy motives index and on the overall procedural justice index, other things being equal. Full details on the indicators included in the measurement of trustworthy motives – and the other elements – can be found in Appendix A, but a key difference in the measurement of the dignity element in the present study and in the original index created by Jonathan-Zamir et al. (2015) must be acknowledged.

While Jonathan-Zamir et al. (2015) recoded relevant items and grouped them into a 0-4 dignity scale that measured the extent to which an officer behaved respectfully toward a citizen – as determined by whether disrespect was shown at all and by how much of the encounter was spent showing respect – the instrument used is the present study did not capture similar information about respect duration. For this reason, as can be seen in Appendix A, the dignity score is calculated by subtracting the number of disrespectful behaviors shown to a citizen from the number of respectful behaviors shown to a citizen during an encounter. Because calculation of the dignity element in this way allowed for possible dignity scores ranging from -7 to 7, the origin was shifted by 7 points (7 was added to each score) before standardization. This allowed for the relative

standing of the dignity score in each unique case to be maintained but avoided the possibility of averaging a negative standardized element score into the overall procedural justice index, which would made interpretation very difficult. It is worth noting that while there were a few negative dignity scores before shifting the origin, the distribution was not symmetrical at all and the vast majority of the original scores were positive. In fact, 216 citizens (88.9%) were shown only respect, 13 citizens (5.35%) were shown only disrespect, and 14 citizens (5.76%) were shown both.

After shifting the origin of the dignity score and standardizing it into a 0-100 metric, the standardized dignity score was added to other three standardized element scores and the total was divided by four (the number of procedural justice elements) according to the method introduced by Jonathan-Zamir et al. (2015). The aim of using a measure of procedural justice that was as similar to theirs as possible was to enjoy the same benefits of a formative index that they discussed when they introduced it: to assess the total amount of procedural justice citizens receive from police, even if all of the procedurally just behaviors are not highly correlated,[10] as is required in a reflective index like those commonly used to measure procedural justice in surveys and other past research (Jonathan-Zamir et al., 2015).

While this method of measuring procedural justice does assign equal weight to all of the elements, it is not obvious whether and how different weights should be assigned based on existing research. There is some evidence that the quality of interpersonal treatment (dignity and trustworthy motives) influence assessments of procedural fairness

---

[10] The bivariate correlations of the procedural justice index from this study and its constituent elements can be found in Table 5 of Appendix B.

more so than the quality of decision-making (participation and neutrality), especially if the interpersonal treatment is negative (e.g. when officers show disrespectful or mean behavior to citizens), but the extent to which these elements have a greater influence has not been measured (Jonathan-Zamir et al., 2015; Tyler & Fagan, 2008; Worden & McLean, 2014). Furthermore, since procedural justice theory also makes no suggestion that any one element carries more weight than the others and no research provides specific guidance about alternative ways to assign weight, to do so here would be arbitrary (Tyler, 2004).

Independent Variables

The key independent variable in this study is citizen gender, which was reported by observers according to judgments based on a citizen's physical appearance, explicit remarks about his/her gender, or information provided by the officer according to his/her prior knowledge of the citizen or other available sources of information like the CAD system. The sample of 243 police-citizen interactions included a total of 103 females (42.4%) and 140 males (57.6%). In order to isolate the influence of a citizen's gender on the amount of procedural justice that s/he received, a number of other personal and situational characteristics that have been shown to influence procedural justice and police decision-making in existing research are included in analysis. Each of these independent variables can be found in Table 1. Tests for multicollinearity were completed to determine whether any of these variables were so highly correlated that their mutual inclusion in the model would compromise the accuracy of the individual predictors' coefficients. Multicollinearity diagnostics from SPSS revealed no collinearity issues.

Citizen characteristics like age, race/ethnicity, and wealth are considered in the

model, as these variables have been shown to affect officer decision making and

procedural justice, even if the reported effects have been mixed (Dai & Nation, 2009;

Lytle, 2014; Worden & McLean, 2014).[11] Like citizen gender, these characteristics were

reported according to observer judgments based on a citizen's physical appearance,

residence or possessions (in the case of wealth), or any other information available. A

citizen's role in an encounter is also included in the model in one of three categories:

Category 1 roles include those in need of some kind of assistance (victims, service

recipients, and helpless people), Category 2 roles include those in need of investigation

(suspects and disputants), and Category 3 (the reference category) roles include those

who may have some sort of information (witnesses, third parties, and unknowns).

It is expected that citizens filling Category 1 roles will be shown more procedural

justice than others, as they are the least blameworthy, likely the least threatening to the

officer or public safety, and perhaps likely to evoke the most sympathy, as Mastrofski et

al. (2016) explain in their discussion of citizens' moral worthiness. Attribution theory

suggests that a person's – in this case a police officer's – perception of the cause of an

event "affects [their] judgments about, and behavior toward, individuals involved in that

event" (Foley & Terrill, 2008, pp. 194). Thus, when an officer encounters a suspect –

who would fill the Category 2 role – s/he will likely assign blame to that suspect and treat

him/her with less procedural justice, which, Mastrofski et al. (2016) argue, is because the

---

[11] The gender and race/ethnicity of the primary officer are not included in the main model because of the small number of officers in this sample. They were introduced to the regression model during sensitivity testing and the results can be found in Table 8 in Appendix C.

officer views the suspect as less deserving or worthy. In contrast, victims and helpless

people – filling Category 1 roles – are viewed as more deserving and will be shown more

procedural justice (Foley & Terrill, 2008; Mastrofski et al., 2016). Witnesses and third

parties – those filling Category 3 roles – on the other hand, may be shown less procedural

justice since they are considered "peripheral" (Mastrofski et al., 2016, pp. 122) to the

encounter and are thus "given less attention" (Worden & McLean, 2014, pp. 7-20).

Evidence from recent research has supported these expectations of how citizen roles

predict procedurally fair treatment (Mastrofski et al., 2016; Worden & McLean, 2014).

Situational variables that make encounters more complex or dangerous for

officers to handle are also included in the model, as they can make it more difficult for

officers to display procedural justice; when officers are preoccupied with their own safety

or the safety of others, interpersonal treatment and displays of fairness become less of a

priority (Mastrofski et al., 2016; Muir, 1977). As Muir (1977) explains, certain conditions

increase the likelihood of "critical incidents" (pp.59), or instances in which a citizen has

or could have an advantage over an officer. Such conditions include those in which a

citizen is "likely to suffer a less severe injury than he could inflict on the [officer];" is

"less vulnerable to injury" than is the officer; is "less remorseful about injuries he

caused" than is the officer; and is "less aware of any injuries he caused" than is the

officer (Muir, 1977, pp.59). Variables included in the model that could contribute to these

conditions are whether a citizen is intoxicated; whether a citizen is showing signs of a

mental disorder; whether the citizen was in conflict with another citizen present; and the

number of citizens present at a scene upon or shortly after O1's arrival. Because this

variable was highly skewed right (a mean of 5.31 and a median of 3 with 72% of cases reporting 5 or less citizens present), a square root transformation was used to reduce the effect of outliers.

Whether a citizen showed any level of resistance toward police is another variable that could increase the likelihood of a critical incident, decreasing an officer's ability to show procedural justice. The resistance variable included in the model was derived from a question on the coding instrument that inquired about the level of verbal or physical conflict a citizen had with police. These levels included calm verbal disagreements or resistance toward police (no threats or aggressive gestures); agitated verbal disagreements or resistance toward police (no physical contact); flight from police; physical resistance to police contact or use of force; displays of weapons in a threatening manner toward police; using or attempting to use a weapon toward police; and initiating physical contact with police. Any one (or more) of these would constitute a score of 1 on the dichotomous resistance variable but it should be noted that the two highest levels of resistance in any of the cases included in this study were flight from police and physical resistance to police contact or use of force, each of which only occurred one time.

Whether a citizen showed disrespect and whether s/he had a negative attitude toward police at the beginning of the encounter are also included in the model, though not because they necessarily create conditions conducive to the aforementioned critical incidents. Instead, these variables are associated with what Mastrofski, Reisig, and McCluskey (2002) call the citizen's "self-presentation" (pp. 522) and they influence where an officer views an individual in Black's (1995) normative social space. According

to Black (1995), how the law behaves – in this case, how police treat citizens – depends

on where the individuals involved in a legal transaction are positioned in social space.

Along the normative dimension, citizens who are disrespectful or disobedient to police

are placed in a negative normative space and are more likely to be treated disrespectfully,

or with less procedural justice, by police (Mastrofski et al., 2002). While the instrument

used in the present study did not capture whether the citizen initiated disrespect in a

police-citizen interaction, only a citizen's attitude at the beginning of the encounter is

included, as a negative attitude later in the encounter could reflect a response to (rather

than predictor of) treatment by police. Per Black's theory and existing research, it is

expected that both citizen disrespect and negative attitudes toward police will have

negative effects on procedural justice (Dai & Nation, 2009; Mastrofski et al., 2016;

Worden & McLean, 2014).

The primary problem type associated with an encounter is also considered and is

included in the model as a series of dummy variables that represent the following

categories: interpersonal conflict, disorder and nonviolent crime, violent crime, service

problems, and traffic problems (the reference category). Primary problems associated

with an encounter, or the nature of police contact, have often been explored as predictors

of both officer and citizen behavior, with problem seriousness of especially great interest

(Dai, Frank, & Sun, 2011; Mastrofski, Snipes, & Supina, 1996; McCluskey, Mastrofski,

& Parks, 1999). In this case, problem seriousness is related to the potential consequences

that a citizen may face with the legal system (e.g. citation, arrest, etc.) as a result of what

transpires during the encounter (McCluskey et al., 1999). Some researchers argue that

this problem seriousness will affect a citizen's decision-making, especially in terms of engaging in rebellious or disrespectful behaviors, because more serious problems have higher stakes; when citizens are more invested (the stakes are higher) they are more willing to face the consequences of refusing compliance and showing disrespect to the officers. (Mastrofski et al., 1996; McCluskey et al., 1999). Others have argued a possible alternative, however, that these citizen behaviors are actually indirect responses to problem seriousness and that they are, in a sense, reactions to the procedural justice shown to them, which tends to be less in situations of greater seriousness (Dai et al., 2011). Put differently, police show less procedural fairness in more serious encounters and this affects citizens' behavioral decisions (Dai et al., 2011). To account for this possibility that the seriousness of the problem predicts procedural justice independently of how it predicts citizen behavior (e.g. police will show less procedural justice in more serious encounters, even if citizens are not rebellious or disrespectful), these primary problem types have been included in the model. It is expected that citizens involved in the more serious problems (e.g. those involving crimes) will be shown less procedural justice than those involved in less serious problems (e.g. traffic stops and requests for service).

Lastly, a citizen's need is considered and is operationalized as whether a citizen summoned police to an encounter or appeared to be injured or ill. This need variable will be used with gender to create an interaction term, described below, to test selective chivalry and thus it is meant to reflect the kind of need that may require paternalistic help without including kinds that could simultaneously violate gender norms or be conducive

to Muir's (1977) critical incidents described above. For this reason, although they could also reveal a citizen's need for assistance, mental illness and whether a citizen is intoxicated are only included in the model as separate variables, not in this measure of need. Both of these conditions may reveal a citizen's potential threat to society's welfare – not just his or her own – and intoxication could be associated with a level of blameworthiness that would not appropriately reflect the type of vulnerability that this need variable is meant to capture (Worden & McLean, 2014). Furthermore, mixed findings about whether women are more or less likely than men to serve prison time for drug crimes (as well as whether their sentences differ) and evidence that masculine norms encourage drinking, especially at a level that would result in intoxication, suggest that signs of intoxication could violate female gender norms, making them a poor indicator of the need that will later be used to test selective chivalry (Iwamoto, Cheng, Lee, Takamatsu, & Gordon, 2011; Rodriguez et al., 2006; The Sentencing Project, 2007).

It should be noted that there are a few other variables that were measured by the coding instrument and have been shown to have statistically significant effects on procedural justice that were not included in this study. For instance, whether an officer ran code to an encounter (used lights and sirens in emergency response), a potential measure of an officer's expectation of urgency and danger, as well as the amount of "people work" an officer had completed at the time of a given encounter (the number of previous encounters an officer was involved with during his/her shift) were excluded. Due to the small sample size, maintaining a low number of independent variables was a priority. Since the present study focuses primarily on citizen characteristics and

behaviors, those variables (e.g. citizen gender, citizen attitude, whether the citizen was resistant) were chosen for inclusion in the model over others that are less theoretically relevant (e.g. the officer's earlier "people work"). Although preliminary bivariate analysis revealed that these excluded variables had only marginal relationships with the dependent variable,[12] indicating that they are not significant predictors of procedural justice, their exclusion is still associated with a limitation of this analysis, to be discussed in greater detail below.

---

[12] Sensitivity tests completed after the main analysis also revealed that when added to the base regression model, these two variables had no statistically significant effects on procedural justice.

Table 1: Independent Variables

The variable	Description
Citizen characteristics	
Citizen's gender	1 = Female (n = 103) 0 = Male (n = 140)
Citizen's age	1 = 21 or older (n = 203) 0 = Younger than 21 (n = 40)
Citizen's race/ethnicity	1 = Minority: Black/Hispanic/Asian/Pacific Islander/ Native American/Mixed Race/Other (n = 204) 0 = White, Non-Hispanic (n = 39)
Citizen's wealth	1 = Appeared poor (n = 52) 0 = Appeared middle class or above (n = 191)
Citizen's role in encounter	
Category 1	1 = Yes (victim, service recipient, helpless person; n = 79) 0 = Other role (n = 164)
Category 2	1 = Yes (suspect, disputant; n = 121) 0 = Other role (n = 122)
Category 3	1 = Yes (third party, witness, unknown; n = 43); reference category 0 = Other role (n = 200)
Situational variables	
Primary problem type	
Interpersonal conflict	1 = Yes (Domestic or non-domestic dispute; n = 63) 0 = Other problems (n = 180)
Disorder and nonviolent crime	1 = Yes (public nuisance, suspicious person or condition, drug violation, weapons violation, theft, motor vehicle theft, warrant to be served for nonviolent crime, burglary; n = 60) 0 = Other problems (n = 188)
Violent crime	1 = Yes (assault, general child abuse, processing/booking for violent crime, homicide; n = 43) 0 = Other problems (n = 201)
Traffic problems	1 = Yes (traffic accident, moving/vehicle violation, DUI, or other traffic problems; n = 44); reference category 0 = Other problems (n = 199)
Service problems	1 = Yes (general request for service, mental disorder, problems with money/credit/documents, misc.; n = 33) 0 = Other problems (n = 206)

Table 1: Independent Variables (continued)

The variable	Description
No. of citizens on scene	Number of citizens present upon/shortly after O1's arrival on scene (square root); M=2.05; Range=1-5.32; SD=1.06; n =243
Citizen is intoxicated	1 = Yes, showed any sign(s) of alcohol/drug use (n = 38) 0 = No (n = 205)
Citizen mental disorder	1 = Yes, showed any sign(s) of a mental disorder (n = 23) 0 = No (n = 220)
Citizen in conflict with another citizen	1 = Yes, any type of conflict with citizen present (n = 47) 0 = No (n = 196)
Citizen disrespectful to police	1 = Yes, showed any disrespectful behavior (n = 46) 0 = No (n = 197)
Citizen attitude toward police	1 = Negative attitude toward police at the beginning of an encounter (n = 36) 0 = Neutral or positive attitude toward police at beginning of an encounter (n = 207)
Citizen was resistant to police	1 = Yes, any level of resistance (n = 52) 0 = No (n = 191)
Citizen showed need	1 = Yes, summoned police or was injured/ill (n = 86) 0 = No (n = 157)

Selective Chivalry Interaction Terms

Citizen gender x Resistance	1 = Female and resistant (n = 17) 0 = Female and not resistant, male and resistant, or male and not resistant (n = 226)
Citizen gender x Need	1 = Female and showing need (n = 42) 0 = Female not showing need, male showing need, or male not showing need (n = 201)

Statistical Analysis

Multiple regression is used to test the effects of the aforementioned variables on the amount of procedural justice shown to a citizen. To test the main effect of citizen gender on procedural justice, all of the citizen characteristics and situational variables are analyzed in the base model. To test selective chivalry, two interaction terms will be added to this base model individually in subsequent tests. First, a Citizen Gender X Need interaction term (calculated by multiplying the Citizen Gender variable by the Need variable) will be introduced to determine whether a female showing need to police has a greater influence on procedural justice than a male showing need. Next, the Citizen Gender X Need interaction term will be removed and a Citizen Gender X Resistance interaction term (calculated by multiplying the Citizen Gender variable by the Resistance variable) will be added to the model to determine whether resistant behaviors from females have a greater influence on procedural justice than similar behaviors from males.

Methodological Limitations

Before moving on to the findings from these analyses, some limitations of the method described above must be acknowledged. Several of these limitations can be traced to the original coding instrument and to the kind of data that it collected. Since the instrument was originally designed for and used in a larger project that was not focused on gender disparity in police treatment of citizens, there were several variables that would be valuable for this study that were not measured, particularly those related to citizens' conformity to gender norms. The physical appearance of female citizens (beyond their apparent wealth), what kinds of cars they were driving if they were involved in traffic

stops, and whether they cried or behaved in other ways that would be considered

traditionally feminine would have been valuable for considerations of selective chivalry.

Additionally, the durations of each procedurally just behavior (e.g. how long the officer

spent showing care and concern to a citizen, how much of the encounter was spent

showing respect to a citizen, etc.) would have been desired in order to create a procedural

justice index identical to the one created by Jonathan-Zamir et al. (2015). Unfortunately

these durations were not captured well on the instrument, which required some

adjustments like that of the measurement of the dignity element as described above. Still,

however, since the instrument was designed to measure procedural justice, most of the

indicators included in the Jonathan-Zamir et al. (2015) index were recorded.

In addition to those related to the coding instrument, there are also limitations

associated with the use of observational data collection. Like any other method, there are

concerns with error, which can be introduced intentionally (by cheating or shirking) or

unintentionally (through unconscious biases or reactivity) by observers (McCluskey et

al., 2014). Unfortunately, there is little research about the frequency of these types of

error in SSO and the evidence that does exist has not been consistent (McCluskey et al.,

2014). Thus, until more systematic assessments have been completed, one cannot assume

that error is a greater concern in SSO than in other methods and the benefits of SSO (e.g.

increased reliability, greater objectivity of a third party observer, and usefulness for

informing policing strategies, as described above) may outweigh these risks.

Still, however, there are weaknesses associated with the observational method that

are particularly apparent at departments that employ two-officer units, as did the two

divisions observed in the present study. While the coding instrument was designed to capture the behaviors of the primary officer in O1's unit (O1 or O2) during a given encounter, there is certainly the possibility that some important information was not captured since a single observer cannot account for every action taken by two officers; since observers were instructed to remain with their assigned officer (O1) if s/he and her/his partner separated during an encounter, any actions taken by O2 out of vision or earshot of the observer would not have been recorded. This is an important issue to consider during research design of future observational studies of procedural justice at departments with two-officer units.

Finally, there are analytical limitations associated with the small sample size, which will be addressed in greater detail in the Discussion section below. The low number of interactions and unique officers observed did not allow for the use of hierarchical linear modeling, which would have been the preferred statistical approach to analyzing the nested cases in the present study. Furthermore, the small sample size limits the sensitivity of the analysis, increasing the possibility that some effects will go undetected and forcing the researcher to exclude some potential variables in order to keep the number of predictors low. As stated above, whether officers ran code to an encounter and the amount of people work completed by an officer prior to a given encounter were both excluded from the present study in order to manage the number of independent variables included in the model. In the future, studies with larger sample sizes and those specifically designed to answer gender-related research questions will be able to avoid some of these limitations.

RESULTS

This section will provide descriptive statistics of this study's sample and will describe the results of the aforementioned regression analyses. Results related to each hypothesis will be presented in the order that they were listed above and they will be interpreted and discussed in detail in the subsequent Discussion section.

Sample Descriptive Statistics

As stated above, 243 police-citizen interactions – 140 with male citizens and 103 with female citizens – are included in this study. The majority of the citizens involved were over twenty-one years old (n=203, 83.5%) and most were racial/ethnic minorities (n=204, 84%). About half of the citizens involved in these encounters filled Category 2 roles – suspects and disputants (n=121, 49.8%) – and just over one third filled Category 1 roles – victims, service recipients, and helpless people (n=79, 32.5%). The distribution of primary problem types was fairly even with each of the five types applying to at least 15.2% of cases but no more than 25.9% of cases. The most common type of primary problem was interpersonal conflict (n=63, 25.9%), which involved domestic and non-domestic disputes. Descriptive statistics related to the remaining independent variables can be found in Table 1 above.

Hypotheses 1: Citizen Gender and Procedural Justice

The first hypothesis tested was based on general chivalry theory and stated:

Female citizens will receive more procedural justice than male citizens

when other personal and situational variables are held constant.

To test this hypothesis, a multiple linear regression was calculated to predict the amount

of procedural justice shown to citizens based on the independent variables in Table 1,

with the exception of the primary problem type reference category (traffic problems), the

citizen role reference category (Category 3 roles), and the two interaction terms related to

selective chivalry (Citizen Gender X Resistance and Citizen Gender X Need), which

were added in subsequent analyses. The model was statistically significant and explained

8.1% of the variance in the amount of procedural justice shown, which is less than has

been found in other models (Mastrofski et al., 2016; Worden & McLean, 2014).

Although female citizens were on average shown 2.05% more procedural justice

than their male counterparts, this relationship was not statistically significant and support

for Hypothesis 1 was not found. In fact, the only significant predictors of procedural

justice were a citizen's racial/ethnic minority status (Beta = -6.2, $p < .05$), his or her

involvement in an encounter with a service-related primary problem (as compared to

traffic-related primary problems) (Beta = -8.78, $p < .05$), and his or her negative attitude

about police involvement at the beginning of their interaction (Beta = -11.51, $p < .01$).

Each of these predictors had a negative effect on procedural justice but no other citizen

characteristics or situational variables showed significant effects. The full results from

this regression can be found in Table 2 below and will be discussed in greater detail in

the subsequent Discussion section.

Table 2: Procedural Justice Regression (Base Model of Main Effects)

	B	SE	β
Constant	71.32	5.87	
Citizen characteristics			
Female citizen	2.05	2.28	.06
Citizen over 21	-1.49	3.27	-.03
Citizen racial/ethnic minority	-6.20*	3.06	-.13
Citizen appear poor	-5.71^	3.12	-.14
Citizen's role in encounter			
Category 1	4.27	3.27	.12
Category 2	.96	3.16	.03
Category 3			
Situational variables			
Primary problem type			
Interpersonal conflict	-2.91	3.49	-.08
Disorder and nonviolent crime	-6.34^	3.65	-.16
Violent crime	-1.08	3.91	-.02
Traffic problems			
Service problems	-8.78*	4.07	-.18
No. of citizens on scene (SQRT)	-.7	1.10	-.04
Citizen is intoxicated	1.31	3.32	.03
Citizen appeared to have mental disorder	2.76	4.16	.05
Citizen in conflict with another citizen	-3.25	3.03	-.08
Citizen was disrespectful to police	2.66	3.89	.06
Citizen negative attitude toward police	-11.51**	3.69	-.24
Citizen was resistant to police	-3.43	3.70	-.08
Citizen showed need	-.37	2.46	-.01

Adjusted R^2 = .081**

N = 243
$^\wedge p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Hypothesis 2: The Interaction of Citizen Gender and Need

The second hypothesis tested was based on the selective chivalry hypothesis and stated:

> *Female citizens who demonstrate a need for police assistance will be*
>
> *shown more procedural justice than males who demonstrate a similar*
>
> *need when other personal and situational variables are held constant.*

As described above, a Citizen Gender X Need interaction term was added to the base model to test this hypothesis. Examination of the change in R^2 (the amount of variance in the dependent variable explained by the predictors in the model) and the effect of the interaction term on the model, neither of which were statistically significant, revealed that the effect of a citizen's need on procedural justice does not vary significantly by citizen gender. Thus, no support for Hypothesis 2 was found. This result can be seen in Table 3 below.

Table 3: Examination of Citizen Gender X Need Interaction Effect

	Step 1: Base Model				Step 2: Citizen Gender X Need Interaction			
	b	SE	β	t	b	SE	β	t
Female Citizen	2.05	2.28	.06	.90	4.68	2.88	.14	1.62
Need	-.37	2.46	-.01	-.15	2.83	3.26	.08	.87
Female Citizen X Need					-6.79	4.55	-.15	-1.49
Adjusted R²			.081**				.086**	
Δ in R²				.15				.008
Sig. F Change				.005				.138
Constant	71.32	5.87		12.14	70.20	5.90		11.89

N = 243
^p < .10; *p < .05; **p < .01; ***p < .001

57

Hypothesis 3: The Interaction of Citizen Gender and Resistance

The third hypothesis tested was also based on selective chivalry and stated:

> *Female citizens who show resistant behavior toward police will receive*
>
> *less procedural justice than will male citizens who show similar resistant*
>
> *behavior when other personal and situational variables are held constant.*

To test this hypothesis, the Citizen Gender X Need interaction term described above was removed from the regression model and the Citizen Gender X Resistance interaction term was added. As was the case with the interaction between Citizen Gender and Need, however, no significant interaction effect was found. As can be seen in Table 4 below, the change in R^2 when the Citizen Gender X Resistance interaction term was added to the model was not statistically significant, indicating that the effect of citizen resistance on procedural justice does not vary significantly by citizen gender. Hypothesis 3 was not supported.

Table 4: Examination of Citizen Gender X Resistance Interaction Effect

	Step 1: Base Model				Step 2: Add Citizen Gender X Resistance Interaction			
	b	SE	β	t	b	SE	β	t
Female Citizen	2.05	2.28	.06	.90	1.91	2.50	.06	.76
Resistance	-3.43	3.70	-.08	-.93	-3.70	4.13	-.09	-.90
Female Citizen X Resistance					.84	5.76	.01	.15
Adjusted R²				.081**				.077**
Δ in R²				.15				.00
Sig. F Change				.005				.884
Constant	71.32	5.87		12.14	71.26	5.90		12.08

N = 243
^p < .10; *p < .05; **p < .01; ***p < .001

59

DISCUSSION

This section will interpret the findings reported in the preceding Results section and explain what they suggest about the applicability of the chivalry and selective hypotheses to decision-making by all agents of the justice system, as well as what they contribute to more general discussions of procedural justice. Some limitations of these findings are also described and suggestions for future research are provided.

General Chivalry Theory and Procedural Justice (Hypothesis 1)

Overall, the results from the regression analyses provided little support for the three expressed hypotheses. For Hypothesis 1, although females were shown on average more procedural justice than males in the model, this finding was not statistically significant and thus may have been due to chance. A sensitivity test that included regressing all of the predictors in the model on each of the four elements of procedural justice individually – as four separate dependent variables – also revealed that, on average, females are shown more of *each* procedural justice element than males but none of these individual relationships were significant either.[13] This further reinforces the notion that there is no gender effect on procedurally just treatment. Still, though no evidence of a gender effect was found, it would be imprudent to interpret these results as conclusive evidence that the justice system has finally abandoned its chivalrous past after

---

[13] See the full results of this sensitivity test in Table 9 of Appendix C.

decades of disparate sentencing and arrest decisions. Being that these findings were derived from analysis of such a small sample of interactions that were only observed at a single department, their generalizability is very limited and they should only contribute to the discussion of gendered procedural justice; they should not be taken as the final word.

Furthermore, the lack of evidence supporting Hypothesis 1 is possibly explained by the inapplicability of chivalry theory to the kinds of interactions that citizens had with agents of the justice system in the present study, particularly in terms of the participants and decision-making involved. Firstly, this study includes citizens filling all kinds of roles (e.g. victims, suspects, witnesses, third parties, etc.) while, as explained in the Literature Review section above, chivalry theory has generally been applied only in studies of female offenders or suspects. In those cases, the decision to be made by an agent of the justice system involves the citizen's arrest, conviction, or sentence, which introduces the need for chivalrous treatment; the underlying stereotypes of chivalry theory are that women are weaker, more emotional, and less accountable for their actions than men, and are thus in need of paternalistic protection and consideration from the justice system. In contrast, women who fill the roles of victims, witnesses, helpless people, or third parties are not facing the same undesirable outcomes from the justice system, and thus may not be in need of such protection or consideration.

This speaks to the possible inapplicability of chivalry theory to decisions involving procedural justice in general. As stated above, procedural justice is the fairness of the policing *process* and is unrelated to the outcome (e.g. an arrest) of a police-citizen interaction. It may be that the fairness of the process – or the quality of interpersonal

treatment and decision-making – is not something from which women need protection or for which they need special consideration, even when they are an offender or suspect. Thus, it is possible that the kind of consideration associated with leniency in arrest and sentencing decisions is not analogous to the kind of consideration associated with the procedurally fair treatment of all kinds of female citizens (e.g. witnesses, third parties, victims, etc.) and that the theory cannot be applied to *all* decisions made by police (e.g. discretionary decisions to show different amounts of procedural justice). Ultimately, it is possible that fair treatment and penal leniency cannot be equated under chivalry theory.

With all of this being said, however, researchers should not be discouraged from examining chivalry in the context of procedural justice decisions in the future. As stated above, the present study's small sample and limited setting restrict the generalizability of its findings and there are still many other matters to consider and ways to improve upon this research in order to make valuable contributions to the discussion of gendered treatment by agents of the justice system. For instance, researchers using larger samples could examine whether the underlying stereotypes of chivalry theory affect officers' decisions to show individual elements differently, which would be masked if only a single measure of procedural justice is used. One could make the argument, perhaps, that the view of women as weak, emotional, and unaccountable could lead officers to show them less respect or more *disrespect* than men, decreasing the dignity element, while also encouraging them to show more care and concern (as a paternalistic protection), increasing the element of trustworthy motives. To investigate this possibility, future

researchers could use four separate regression models where each procedural justice element is a single dependent variable.

This approach would be similar to that used by Visher (1983) who also ran separate models to examine more nuanced differences in the treatment of male and female suspects by police. Rather than exploring individual procedural justice elements in separate models, however, Visher (1983) examined the probability of arrest for males and the probability of arrest for females as separate dependent variables. While she also found that males and females were treated similarly during police-citizen interactions, she was able to examine how each of the predictors in her study affected the treatment of males and females differently, allowing her to make judgments about selective chivalry. Such an approach would also be desirable in future procedural justice studies with large enough sample sizes since it would allow researchers to avoid the reliance on interaction terms, decreasing the total number of predictors entered into a model and preserving some statistical power. In the present study, only interaction terms were used to detect selective chivalry and the associated results are discussed below.

Selective Chivalry and Procedural Justice (Hypotheses 2 & 3)

Just as there was virtually no evidence supporting Hypotheses 1, which was based on general chivalry theory, there was also a lack of evidence supporting Hypotheses 2 and 3, which were based on the selective chivalry hypothesis. When the two interaction terms used to test these hypotheses were individually added to the base model from main analysis, no significant interaction effects were found between citizen gender and need or resistance, indicating that females who show need or are resistant toward police are not

treated differently than males who do the same. During sensitivity testing, however, after all cases involving encounters that included four or more citizens were excluded from analysis, it was found that females who demonstrated need were actually shown 5.42% *less* procedural justice than males who demonstrated need or females who did not demonstrate need.[14] This contradicts Hypothesis 2 and the selective chivalry hypothesis, which would predict female citizens being shown more procedural justice after demonstrating need since they would be conforming to traditional gender norms whereby women are vulnerable and in need of protection. While these results may suggest that the size of encounters (based on the number of citizens involved) can affect how citizens demonstrating need are treated by police, they must be interpreted with great caution since they were derived from analysis of an even smaller sample (n=212) than that used during main analysis. Further, with little evidence of a general chivalry effect on procedural justice, it is difficult to draw conclusions about the more specific selective chivalry hypothesis.

Without considering any of the findings related to general chivalry theory, the marginal effects of the resistance and need interaction terms would contradict the selective chivalry hypothesis as Embry and Lyon's (2012) interpret it since they include the evil woman hypothesis, which suggests that women who violate gender norms will be punished *more* harshly than men. This goes beyond Rodriguez et al.'s (2006) interpretation, which is only a conditional extension of general chivalry theory that says women who violate gender norms will be punished more harshly than women who *do*

---

[14] See results of this sensitivity test in Table 7 of Appendix C.

conform, "perhaps *similar* [emphasis added] to men" (pp. 322). By Embry and Lyon's (2012) interpretation, resistant behavior by females should have had a more negative impact on procedural justice than resistant behavior by males, even if there was no evidence of a general chivalry effect. Since this relationship was not found, the results from this study did not support Embry and Lyon's (2012) interpretation.

In contrast, with no evidence of a general chivalry effect in the present study, Rodriguez et al.'s (2006) selective chivalry cannot even be tested. If women were not shown more consideration than men, there should be no difference between women who violate gender norms and women who conform to them since violators only risk losing an advantage that may have been awarded them based on their gender; they do not risk being punished more harshly than men. Relatedly, if the lack of evidence supporting Hypothesis 1 is interpreted to mean that chivalry theory does not apply to officers' decisions to show procedural justice in general (not only that chivalry was not shown in these specific instances), it must also be true that selective chivalry cannot be applied; if officers do not offer paternalistic consideration to females in procedural justice decision-making, violators of gender norms have nothing to lose in these encounters and neither conformity to nor violation of gender norms would affect the amount of procedural justice shown.

Although a lack of evidence supporting a general chivalry effect makes it difficult to make any kind of statements or judgments about selective chivalry in the present study, it is also possible that the need and resistance variables used are simply inadequate indicators of conformity to and violation of gender norms. As stated above, since the data

from this study were taken from a larger project that was not concerned with gender norms or chivalry, there were no other citizen behaviors or characteristics captured that could have allowed for more in-depth examination of the extent to which the citizens conformed to or violated gender norms. Other details about citizen behaviors related to the underlying stereotypes of selective chivalry such as a woman's weakness or emotional fragility would be valuable, particularly in light of DeFleur's (1975) findings. Whether a woman cried or appeared/indicated that she was emotionally or financially dependent on a male at the scene, for example, may have been more appropriate considerations than the need variable used here.

Overall, although these findings do contribute to the discussion of gender disparity in treatment by agents of the justice system, they cannot be taken as definitive evidence that women will receive the same amount of procedurally just treatment as men in similar encounters or that their violation of or conformity to gender stereotypes will not influence that treatment. It is possible that police are gravitating toward more equal treatment of women and that they are abandoning gender stereotypes – or at least resisting the influence of these stereotypes on their behavior – but more research is needed in order to make conclusive or generalizable statements.

Significant Predictors of Procedural Justice

The Negative Effect of Involvement in an Encounter with a Service Problem

Despite the lack of findings supporting my three hypotheses, there were other statistically significant relationships uncovered that should be discussed. As stated above, it was found that citizens involved in encounters with service-related primary problems

were shown 8.78% less procedural justice than those citizens involved in traffic problems (the reference category). Because the remaining three primary problem types (violent crime, nonviolent crime and disorder, and interpersonal conflict) had no such statistically significant negative effects on procedural justice and each of these three types were more serious than service problems, this finding contradicts the aforementioned expectation that less procedural justice would be shown to citizens involved in encounters with more serious problems. To better understand this unexpected negative effect, a sensitivity test was completed to determine the influence of a citizen's involvement in a service-related encounter on each of the four elements of procedural justice individually. This was done by regressing all of the independent variables in the base model against four new dependent variables – the four elements – instead of the full procedural justice index.[15]

It was found that the entire significant effect of the service primary problem type on procedural justice was attributed to its significant effect on the neutrality element; the service primary problem type had no statistically significant relationship with any one of the other three elements. To understand this finding, it is important to consider the indicators used to measure neutrality in the context of service-related encounters. The three indicators used are related to how/whether an officer explained his/her involvement in the encounter and how/whether the officer explained his/her objective decision-making, including whether s/he indicated that s/he would wait to make a final decision until s/he had gathered all necessary information.

It is possible that the nature of service-related problems simply do not present

---

[15] The results of these regression analyses can be found in Table 9 of Appendix C.

officers with the same opportunities to express these neutral behaviors (or the same *need* to express them) as do other kinds of problems at different encounters. For instance, service calls that would fall under this service primary problem type include transports (e.g. bringing home a person involved in an automobile accident) and police escorts. It seems intuitive that when citizens request these kinds of services from police, the officers will not *need* to explain why they became involved in the situation (e.g. "I am taking you home now because your car was totaled and you have no other safe way home"), especially since the citizen will know why officers are present if s/he is the one who made the request. Likewise, although there are certainly circumstances under which an officer would need to consider all relevant facts and perhaps explain his/her decision *not* to provide the requested service if such a decision was made, these instances are likely rare, as well; in many cases, an officer's presence at the scene of a service encounter suggests that s/he has already decided to provide the service to the citizen (e.g. arriving at a local church to provide an escort for a funeral procession) and there would be no need to provide much of an explanation.

In addition to providing insight about why citizens involved in service-related encounters are shown less procedural justice than citizens involved in traffic-related encounters, the finding that one element accounts for the entire statistically significant effect on procedural justice may reveal the value of examining the elements individually in the future. Doing so would not only help to uncover masked findings and explain nuanced effects, but it would also allow for more meaningful recommendations about the best way to increase procedural justice in order to improve citizen perceptions of police

68

legitimacy. For instance, if police leaders learn that citizens involved in service encounters are being shown less procedural justice than citizens involved in traffic encounters without being made aware of the more exclusive effect on neutrality, they may advise their patrol officers to show more procedurally just behaviors to these citizens across the board. Such a general recommendation could lead to increases in neutrality, but if officers were aware of the specific area of need, they could target it and more efficiently improve their procedural justice performance without unnecessarily changing other behaviors. Furthermore, they would not run the risk of simply showing more of the dignity or trustworthy motives elements, neither of which would help to improve the negative effect on neutrality.

This finding of a single element effect yielding a statistically significant relationship with the full procedural justice index may also speak further to the need for systematic assessments of how individual elements affect citizen perceptions of police performance and legitimacy. Such assessments will help researchers to assign different weights to certain elements if found to be necessary. As was previously mentioned, some researchers have found that interpersonal treatment (dignity and trustworthy motives) has a greater effect on perceptions of procedural justice than the quality of decision-making (participation and neutrality) (Jonathan-Zamir et al., 2015; Tyler & Fagan, 2008; Worden & McLean, 2014). While too little is known about the extent of this differential influence to have assigned varying weights to the elements in the present study, seeing one of the elements thought to have a lesser effect on perceptions of procedural justice determining the entire measure does raise some concerns about the equal weighting. Perhaps if the

two elements related to interpersonal treatment were assigned greater weight in the

overall measure, their relationship with the predictor would have been statistically

significant, as well, and the entire effect on procedural justice would have shifted. The

value of research that contributes to our increased understanding of individual element

effects in order to improve measurement cannot be overstated.

The Negative Effect of a Citizen's Negative Attitude

Another statistically significant finding that perhaps requires less explanation than

the one above is that citizens who had negative attitudes about police presence at the

beginning of their interactions with police were shown 11.51% less procedural justice

than those citizens who had neutral or positive attitudes.[16] This negative relationship was

expected according to Black's (1995) notion of normative social space described in the

Method section above. While it is surprising that the other predictor thought to influence

where citizens are placed in normative social space – whether a citizen was disrespectful

toward police – did not have a statistically significant effect on the amount of procedural

justice shown to citizens, the expected finding related to citizen attitudes is helpful to

police agencies.

It seems intuitive that a negative reception from another person with whom one is

interacting would influence one's disposition and behaviors as communication continues

with that person, but for police, understanding this effect is especially important so that it

---

[16] To determine whether the effect of the citizen's negative attitude on procedural justice was different for male and female citizens, a Citizen Gender X Citizen Attitude interaction term (calculated by multiplying the Citizen Gender variable by Citizen Attitude variable) was added to the base model. No significant interaction effect was found, suggesting that female citizens who show negative attitudes about police presence at the beginning of encounters are not shown more or less procedural justice than male citizens who do the same.

can be changed. Although the present study provides no information about what shaped the citizen's negative attitude toward the officers at the start of an encounter, it is not unreasonable to expect that negative perceptions of police will be manifested as negative attitudes. If this is the case, the very thing that suffers because of a citizen's negative attitude – the amount of procedural justice shown – may be an effective tool to prevent it. Unlike many of the other predictors included in the model, this predictor – citizen's negative attitude – can both affect and be *affected by* procedural justice. As Mazerolle et al. (2013), Gau et al. (2012), and Sunshine and Tyler (2003) have found, procedural justice can improve citizens' views of police as legitimate and can increase their satisfaction with police. Increased satisfaction and views of legitimacy will hopefully yield more positive attitudes about police presence and perhaps more positive community relations overall. Such is the ultimate goal of procedural justice, as described in the Introduction section above.

Of course, there are many other reasons why a citizen may have a negative attitude about police presence that are not specifically associated with their views of police as illegitimate or their satisfaction with police in a general sense. Further, since the attitudes measured in this study are demonstrated at the *beginning* of encounters, showing procedural justice in order to elicit more positive attitudes would mostly be a proactive effort for future encounters. Still, it is possible that meeting negative attitudes with more procedural justice will help to improve citizens' negative attitudes *during* encounters and it is good for police leadership to be aware of patrol officers' responses to citizens' initial attitudes, even if the underlying causes of those attitudes are unknown.

The Negative Effect of a Citizens' Racial/Ethnic Minority Status

The third and final statistically significant predictor of procedural justice found in the main regression model highlighted one area of note that was not as consistent with prior research. Based on the data in this sample, it was found that citizens of racial/ethnic minority groups were shown 6.2% less procedural justice than White, non-Hispanic citizens when all other variables were held constant.[17] This finding contradicts those of Worden and McLean (2014) and Mastrofski et al. (2016) – who conducted the other two most recent observational studies of procedural justice – who did not find negative race effects in their studies. In fact, Worden and McLean (2014) found that Black citizens were shown *more* procedural justice than their White counterparts.

Of course, one must consider the differences between the samples of both officers and citizens in these past studies and those of the present study. Worden & McLean (2014), for instance, collected data in two cities in upstate New York where at least 56% of each population was White and at least 91.8% of the sworn personnel at each department were White. In Mastrofski et al.'s (2016) study, observations were completed at two departments, one of which had a population of half minority citizens while the other had a population of only one third. Of the encounters observed, 43.2% included White citizens and 61% included White officers. In the present study, however, only 16% of the citizens encountered were White, non-Hispanic and only 40.7% of the encounters included White, non-Hispanic officers. The percentage of minority officers and citizens included in the present study are certainly higher than national averages (Reaves, 2015),

_____

[17] A sensitivity test was completed to determine whether there was a significant interaction between a citizen's race/ethnicity and his/her gender. No significant effect was found.

which again reveals the limited generalizability of these findings. Importantly, however, the inconsistent race/ethnicity-related findings from these three studies indicate that it may not be possible to generalize about race effects across departments. Furthermore, despite the limited generalizability, these surprising findings are worth exploring in greater depth to better understand *how* a citizen's minority status affects procedural justice. To do so, I examined the effects of a citizen's minority status on each of the four elements of procedural justice with four separate regression models, as was done to explore the negative effect of service problems described above.[18] Each element score was used as a dependent variable (replacing the full procedural justice index used as the dependent variable in the base model) in one of the four new regression models and it was found that a citizen's minority status *only* had a statistically significant effect on the element of neutrality. This is important for the same reasons cited above in the discussion of the similar finding related to service problems; future researchers may want to examine the effects of predictors on individual elements and police leaders can use these findings to provide tailored and more effective recommendations for performance improvement. Additionally, this finding contributes to broader discussions of disparate racial treatment in the criminal justice system.

While recent observational studies – that have also used single measures of procedural justice[19] – have not found negative race effects, a recent meta-analysis of the effects of suspect characteristics on arrest decisions found that Black, male, and Hispanic individuals are significantly more likely to be arrested than White, female, and non-

---

[18] The results from these four regression analyses can be found in Table 9 in Appendix C.
[19] Though Worden & McLean (2014) also used a separate measure of procedural *injustice*.

Hispanic individuals (Lytle, 2014). This meta-analysis included 54 unique and independent studies, 42 of which examined suspect race, and 15 of which examined suspect ethnicity. These unique studies differed in dates completed, duration, number of agencies included, methods used, sample sizes, and types of independent variables considered, but the race effects endured under most circumstances; Blacks were still more likely to be arrested than Whites, "regardless of whether a study accounted for the seriousness of the offense, the amount of evidence, whether a suspect was under the influence of drugs or alcohol, whether the suspect used a weapon, the demeanor of the suspect, and/or whether the suspect committed a crime in front of the officer during the encounter" (Lytle, 2014, pp. 595). Further, the race findings were consistent across study settings and regardless of whether any kind of police-citizen encounters were included in analysis or if only traffic stops were examined (Lytle, 2014).

While the control variables described above are certainly not an exhaustive list of which factors and circumstances can affect officers' arrest decisions, Lytle's (2014) findings do suggest that arrest decisions have been made based on personal characteristics like race, gender, and ethnicity. In this case, it is not surprising that the neutrality of decision-making in the present study is negatively affected by a citizen's minority status since minority citizens appear to receive less objective treatment from officers than their White counterparts. It cannot be assumed that citizens' personal characteristics will affect all officer decisions the same way or that officers will communicate all of the true influences on their decision-making (including personal biases) in ways that would be perceptible to citizens or observers, but the findings from

this study may be coupled with those presented by Lytle et al. (2014) to argue the possibility that citizens of racial/ethnic minorities are subjected to less objective decision-making by police than are White citizens. Based on this finding, more research should be conducted.

Non-Significant Results

Overall, a citizen's minority status, his/her involvement in a service-related encounter, and his/her negative attitude toward police were the only variables included in analysis that were found to significantly predict procedural justice at the .05 level. While both Mastrofski et al. (2016) and Worden and McLean (2014) found that citizens filling suspect and third party roles were shown less procedural justice than victims, citizen roles had no significant effects in the present study. Other situational variables that have been significant in previous studies – such as citizens' resistant or disrespectful behavior, whether s/he initiated police involvement, and how many citizens were present at the scene during an encounter – also had no significant effects here. These limited findings may be due to the small sample size that makes detection of effects more difficult, but they also speak to how meaningful are the relationships that were significant. Furthermore, just as the less common findings described above (e.g. the negative effects of service problems and citizen minority status) should be used to contribute to conversations rather than to act as definitive or general statements, the insignificant effects of many predictors on procedural justice should not be taken to mean that these variables do not or will not affect procedural justice in general or in the future; they too should add to growing conversations and research.

CONCLUSIONS

This study examined 243 police-citizen encounters in the context of chivalry theory and the selective chivalry hypothesis to determine whether females are shown more procedural justice than males during police-citizen interactions and whether any special treatment females are awarded is dependent upon their conformity to traditional gender norms. Although little evidence of differential treatment of female citizens was found, the results from this study provide insight about the applicability of these theories to this kind of decision-making by agents of the justice system. Further, the findings contribute to conversations about which personal and situational factors predict officer decision-making and about the possibility that these factors vary by setting.

Subsequent tests completed to better understand significant findings related to the negative effects of involvement in a service-related encounter, a citizen's negative attitude toward police at the beginning of an encounter, and a citizen's minority status also revealed potential weaknesses related to the current measurement of procedural justice and helped to orient future research of gendered treatment in the justice system and of procedural justice more generally. With the aim of overcoming the weaknesses and limitations of the present study, it is recommended that future researchers use larger sample sizes and conduct observations at multiple departments in order to obtain data that will be more sensitive to effects during analysis and that will produce more generalizable

findings. Future researchers should also design their studies with these gender questions in mind so that observers can capture more relevant data and record them on more specialized coding forms.

With continuing shifts in police focus toward procedural justice and community policing efforts, it is important to continue assessing the persistence and effects of past trends like the paternalistic treatment of women and the discriminatory treatment of racial/ethnic minorities that may still be woven into American society and policing. Further examinations of these patterns in procedurally just treatment by police and continued improvements with research design and the measurement of procedural justice will allow researchers, police leaders, and policy-makers to better under how police are making procedural justice decisions and how procedural justice performance can be improved. Ultimately, the fruit of these studies will be improved understanding of police decisions, improved guidelines for police behavior, improved perceptions of police legitimacy, and improved police-community relations.

MEASUREMENT OF PROCEDURAL JUSTICE IN POLICE BEHAVIOR (ADOPTED FROM JONATHAN- ZAMIR, MASTROFSKI, & MOYAL, 2015)

Participation = The primary officer asked the citizen to provide information/viewpoint regarding the situation/problem (0=no; 1=yes) + [Citizen provided information/viewpoint regarding the situation/problem (0=no; 1=yes) x The way in which the primary officer responded to the citizen's information/viewpoint (0=Poor Listener; 1=Good Listener)]

Neutrality = The primary officer indicated that s/he would not make a decision about what to do in this encounter until s/he had gathered all necessary information (0=no; 1=yes) + The primary officer explained to this citizen or in the presence of this citizen justifications for actions based on law and evidence, fact assertions, public safety, the community's benefit, or the citizen's welfare (0=no; 1=yes) + The primary officer explained to this citizen or in the presence of this citizen why s/he became involved in the situation (0=no;1=yes)

Dignity = [The primary officer called the citizen politely by name or title (0=no; 1=yes) + The primary officer used a calm, even-handed voice (0=no, 1=yes) + The primary officer let the citizen speak freely without interruption (0=no; 1=yes) + The primary officer used polite requests/acknowledgements (0=no; 1=yes) + The primary officer responded to the citizen's questions/comments politely (0=no; 1=yes) + The primary officer identified him/herself by name to the citizen (0=no/after being asked by the citizen; 1=yes, without being asked) + The primary officer showed other respectful behavior (0=no; 1=yes)] – [The primary officer used derogatory name calling, belittling remarks, slurs, or cursing (0=no; 1=yes) + The primary officer used an unnecessarily loud voice (0=no; 1=yes) + The primary officer interrupted the citizen without apology (0=no; 1=yes) + The primary officer used impolite/obscene gestures, spitting (0=no; 1=yes) + The primary officer sneered, rolled eyes, made jokes at citizen's expense (0=no; 1=yes) + The primary officer ignored the citizen who sought attention or spoke curtly to the citizen (0=no; 1=yes) + The primary officer showed other disrespectful behavior (0=no; 1=yes)]

Trust in the motives of the decision-maker = The primary officer described, advised or acknowledged to the citizen that s/he had legal rights as a complainant (0=no; 1=yes) + The primary officer described, advised or acknowledged to the citizen that s/he had legal rights as a suspect (0=no; 1=yes) + The primary officer asked about or showed concern for the citizen's well-being (0=no; 1=yes) + The primary officer offered comfort or assurance to the citizen (0=no; 1=yes) + The primary officer exerted control or influence

over another person for the citizen (0=no; 1=yes) + The primary officer acted on behalf of the citizen with a government agency or private entity (0=no; 1=yes) + The primary officer provided/arranged physical assistance to the citizen (0=no; 1=yes) + The primary officer provided advice on how the citizen could handle the situation or deal with the problem (0=no; 1=yes) + The primary officer encouraged the citizen to seek additional assistance from the police for this or future problems (0=no; 1=yes) + The primary officer empathized with the citizen's concerns (0=no;1=yes)

Table 5: Descriptive Statistics and Correlations of Procedural Justice Elements

	1	2	3	4	5
1. Procedural justice index	1				
2. Participation	.758***	1			
3. Neutrality	.739***	.304***	1		
4. Dignity	.628***	.403***	.240***	1	
5. Trustworthy motives	.607***	.276***	.228***	.450***	1
n	243	243	243	243	243
Range	7-86	0-100	0-100	21-93	0-80
M	59.02	85.6	52.67	70.12	27.7
SD	17.13	30.8	33.55	13.92	19.15

N = 243

$^\wedge p < .10$; $*p < .05$; $**p < .01$; $***p < .001$

Note: All variables were standardized into a 0 to 100 metric.

Range values are rounded to the nearest whole number.

APPENDIX C

Sensitivity Tests Completed Following Main Analysis

Removing Encounters with Four or More Citizens

As stated above, multiple citizen interactions occurred during a single encounter in just over half of the observed cases and except in a few very rare cases, officers were involved in multiple encounters during a single observation session. In order for researchers to account for the nested nature of cases like these, hierarchical linear modeling (HLM) is commonly used – as was the case in Mastrofski et al.'s (2016) recent observational study of police-citizen encounters – but unfortunately there were too few cases to use HLM in the present study. Thus, to determine whether a citizen's presence at an encounter with several other citizens affected the amount of procedural justice shown to him/her by the primary officer, a sensitivity test was completed that involved replicating the regression analysis from the base model with only cases involving three or fewer citizens and then comparing these results to those from the original analysis. When cases involving four or more citizens were excluded, the overall percentage of women included in the sample decreased by less than 1%; the gender make-up of the sample used in this sensitivity test (n = 212; 58% male and 42% female) closely matched the original sample (n = 243; 57.6% male and 42.4% female).

As can be seen in Tables 6 below, the adjusted R^2 in the new model decreased by about 1.8%, meaning that less variance in the dependent variable is explained by the

predictors in the model when the cases of encounters involving four or more citizens are excluded from analysis. Like the original model used in the main analysis above, however, a citizen's minority status and negative attitude toward police at the beginning of an encounter remained statistically significant and their relative negative effects on the amount of procedural justice shown barely changed at all. Unlike the original model, the negative effect of involvement in service-related encounters on procedural justice was no longer significant in this new model. This was the biggest change between the two models since no variables that were not significant in the original were found to be significant in this one. It should be noted, as well, that no cases involving service-related encounters were excluded from analysis in this model, but 10 of the 31 cases excluded (32.3%) were traffic-related. Since traffic-related encounters acted as the reference category for all primary problem types, this may help to explain why the effect of involvement in service-related encounters was no longer significant; perhaps the citizens involved in the traffic encounters that were excluded were shown less procedural justice than those maintained, making the contrast between service encounters and the reference category less perceptible. If true, this may suggest that citizens involved in traffic-related encounters that include four or more citizens are shown less procedural justice than those including three or fewer. This may be the case if the high number of citizens involved creates a more stressful environment for the officers, decreasing the amount of procedural justice they show as they focus more on managing a potentially dangerous encounter than showing fairness. Still, however, because these findings are derived from such a small

sample (n=212), conclusions should be drawn with caution and more research about

procedural justice in encounters with larger groups of citizens should be completed.

Table 6: Regression when Encounters of 4+ Citizens are Excluded

	B	SE	β
Constant	66.97	6.44	
Citizen characteristics			
Female citizen	2.95	2.45	.09
Citizen over 21	-.44	3.82	-.01
Citizen racial/ethnic minority	-6.24*	3.05	-.15
Citizen appeared poor	.61	3.48	.02
Citizen's role in encounter			
Category 1	2.54	3.54	.07
Category 2	.41	3.49	.01
Category 3			
Situational variables			
Primary problem type			
Interpersonal conflict	-.72	3.82	-.02
Disorder and nonviolent crime	-4.66	3.86	-.13
Violent crime	4.52	4.20	.10
Traffic problems			
Service problems	-6.27	4.20	-.14
No. of citizens on scene (SQRT)	-.31	1.14	-.02
Citizen is intoxicated	-1.39	3.81	-.03
Citizen appeared to have mental disorder	.12	4.19	.00
Citizen in conflict with another citizen	-2.26	3.09	-.05
Citizen was disrespectful to police	1.77	3.91	.04
Citizen negative attitude toward police	-10.37**	3.9	-.22
Citizen was resistant to police	-3.33	3.71	-.08
Citizen showed need	.05	2.48	.00
Adjusted R^2 = .063*			

N=243
^p < .10; *p < .05; **p < .01; ***p < .001

After duplicating the base model without the cases involving encounters of four or more citizens, the subsequent analyses that introduced the interaction terms Citizen Gender X Need and Citizen Gender X Resistance were also duplicated. While no significant interaction between citizen gender and resistance was found, as was the case in the analysis that included all 243 original cases, as well, a significant interaction between citizen gender and need *was* found in this smaller sample. As can be seen in Table 7 below, however, this significant effect was not in the direction predicted in Hypothesis 2 according to the selective chivalry hypothesis; it was found that females who demonstrated need were actually shown 5.42% *less* procedural justice than males who demonstrated need or females who did not demonstrate need.

Despite the finding that involvement in service-related encounters was no longer significant and that the interaction between citizen gender and need *was* significant when citizens involved in larger encounters were excluded from analyses, these results are fairly similar to those derived from the full original sample of 243 police-citizen interactions; a citizen's negative attitude and minority status still negatively affected the amount of procedural justice that s/he was shown and no other predictors had significant effects. This similarity between the original results and those found during sensitivity testing suggests that a citizen's involvement in a larger encounter does not lead an officer to treat him/her much differently than the officer would any other citizen. It seems, rather, that officers are more inclined to let a citizen's minority status or negative attitude influence the way that they treat a citizen, independent of whether that citizen is involved in a larger or small encounter.

Table 7: Examination of Citizen Gender X Need Interaction Effect when Encounters of 4+ Citizens are Excluded

	Step 1: Base Model with 4+ Citizen Encounter Excluded				Step 2: Add Citizen Gender X Need Interaction			
	b	SE	β	t	b	SE	β	t
Female Citizen	2.95	2.45	.09	1.21	7.54	3.17	.23	2.38
Need	.05	2.48	.00	.02	5.06	3.32	.15	1.53
Female Citizen X Need					-10.48*	4.68	-.25	-2.24
Adjusted R²				.063*				.082*
Δ in adjusted R²				.143				.022*
Sig. F Change				.03				.03
Constant	66.97	6.44		10.39	64.67	6.46		10.01

N = 243
^p < .10; *p < .05; **p < .01; ***p < .001

Consideration of Primary Officer Characteristics

Hierarchical linear modeling is also used to examine how specific officers treat citizens (which allows researchers to make statements about how officer characteristics, like gender and race/ethnicity, affect the amount of procedural justice shown) but, as stated above, there were simply too few cases to perform HLM here. Furthermore, because of the low number of unique officers in this study – 55 O1s where 46 (83.6%) were male and 9 (16.4%) were female and 52 O2s where 46 (88.5%) were male and 6 (11.5%) were female – it is difficult to make meaningful statements about how their characteristics affect procedural justice. Thus, the following results of a sensitivity test that involved adding the primary officer's gender and race/ethnicity individually to the regression model (without the interaction terms) must be viewed cautiously.

The sensitivity test revealed that although the primary officer's gender had no significant effect on the amount of procedural justice shown, his/her race/ethnicity was significant; citizens who encountered officers of racial/ethnic minorities were shown 6.48% less procedural justice than those who encountered White, Non-Hispanic officers.[20] Further, when the primary officer's race/ethnicity is introduced into the model, there is still a negative relationship between the citizen's minority status and the amount of procedural justice shown to him/her but it is no longer significant at the .05 level. In contrast, the significant relationships between procedural justice and service calls and

---

[20] Another sensitivity test was completed to determine whether there was a significant interaction effect between Citizen Race/Ethnicity and Primary Officer Race/Ethnicity, as well as Citizen Gender and Primary Officer Gender; no significant interaction effects were found.

procedural justice and citizen attitudes toward police endured when the primary officer's

minority status was added to the model. These results can be found in Table 8 below.

Table 8: Primary Officer Race/Ethnicity Added to Procedural Justice Regression

	B	SE	β
Constant	72.99	5.80	
Citizen characteristics			
Female citizen	2.50	2.25	.07
Citizen over 21	-1.53	3.21	-.03
Citizen racial/ethnic minority	-5.53^	3.02	-.12
Citizen appeared poor	-5.17^	3.08	-.12
Citizen's role in encounter			
Category 1	5.74^	3.26	.16
Category 2	2.08	3.13	.06
Category 3			
Situational variables			
Primary problem type			
Interpersonal conflict	-2.36	3.44	-.06
Disorder and nonviolent crime	-6.38^	3.59	-.16
Violent crime	-1.80	3.85	-.04
Traffic problems			
Service problems	-7.98*	4.01	-.16
No. of citizens on scene (SQRT)	-.74	1.08	-.05
Citizen is intoxicated	.95	3.27	.02
Citizen appeared to have mental disorder	1.63	4.11	.03
Citizen in conflict with another citizen	-3.14	2.98	-.07
Citizen was disrespectful to police	2.22	3.83	.05
Citizen negative attitude toward police	-11.27**	3.62	-.23
Citizen was resistant to police	-2.29	3.66	-.06
Citizen showed need	.27	2.42	.01
Primary officer's race/ethnicity	-6.48**	2.21	-.19
Adjusted R^2 = .111***			

N=243

^$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Service Problems and Citizen Minority Status Regressed on Individual Elements

   To better understand the nature of the significant relationship between the amount

of procedural justice shown to a citizen and his/her involvement in a service-related

encounter or his/her racial/ethnic minority status, the predictors in the base model were

regressed on the four individual elements of procedural justice. It was found that the

entire statistically significant effects of both involvement in service-related encounters

and a citizen's minority status on procedural justice were due to negative effects on the

element of neutrality. These results can be found in Table 9 below and are explained in

greater detail in the Discussion section above.

Table 9: Regression of Predictors on Individual Procedural Justice Elements

	Participation B (SE)	Neutrality B (SE)	Dignity B (SE)	Trustworthy Motives B (SE)
Constant	91.99 (10.34)	80.65 (11.49)	81.55 (4.50)	31.07 (6.29)
Citizen characteristics				
Female citizen	5.06 (4.02)	1.49 (4.47)	.86 (1.75)	.80 (2.45)
Citizen over 21	-1.46 (5.76)	-6.75 (6.39)	2.82 (2.5)	-.58 (3.50)
Citizen racial/ethnic minority	-4.95 (5.39)	-15.54** (5.99)	-2.03 (2.34)	-2.29 (3.28)
Citizen appeared poor	-5.61 (5.50)	-7.01 (6.11)	-8.85*** (2.39)	-1.36 (3.35)
Citizen's role in encounter				
Category 1	8.15 (5.77)	-2.87 (6.40)	-.54 (2.51)	12.34*** (3.51)
Category 2	1.53 (5.56)	10.24^ (6.17)	-4.95* (2.42)	-2.99 (3.38)
Category 3				
Situational variables				
Primary problem type				
Interpersonal conflict	.92 (6.15)	-11.25 (6.83)	-5.03^ (2.67)	3.72 (3.74)
Disorder and nonviolent crime	-1.09 (6.43)	-20.06** (7.14)	-.85 (2.8)	-3.36 (3.91)
Violent crime	4.91 (6.88)	-6.74 (7.65)	-.16 (2.99)	-2.34 (4.19)
Traffic problems				
Service problems	-3.22 (7.17)	-21.23** (7.96)	-5.55^ (3.12)	-5.11 (4.36)
No. of citizens on scene (SQRT)	.48 (1.93)	.28 (2.15)	-1.82* (.84)	-1.73 (1.18)
Citizen is intoxicated	1.15 (5.84)	1.49 (6.49)	.36 (2.54)	2.34 (3.55)
Citizen appeared to have mental disorder	-6.86 (7.33)	3.88 (8.14)	6.62* (3.19)	7.42^ (4.46)
Citizen in conflict with another citizen	-4.41 (5.34)	-5.17 (5.93)	-1.84 (2.32)	-1.57 (3.25)
Citizen was disrespectful to police	.73 (6.85)	10.87 (7.60)	.43 (2.98)	-1.41 (4.16)
Citizen negative attitude toward police	-21.81*** (6.49)	-11.48 (7.21)	-5.93* (2.82)	-6.83^ (3.95)
Citizen was resistant to police	-11.66^ (6.52)	.74 (7.24)	-4.53 (2.83)	1.71 (3.96)
Citizen showed need	.44 (4.32)	-3.75 (4.80)	-1.31 (1.88)	3.17 (2.63)
adj R^2	.119***	.084**	.185***	.157***

N = 243
^$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

APPENDIX D

REFERENCES

Bittner, E. (1974). Florence Nightingale in pursuit of Willie Sutton: A theory of the police. In C. B. Klockars & S. D. Mastrofski (Eds.), *Thinking about police: contemporary readings* (pp. 35-51). New York: McGraw-Hill.

Bontrager, S., Barrick, K., & Stupi, E. (2013). Gender and sentencing: A meta-analysis of contemporary research. Journal of Gender, Race & Justice, 16, 349–372. Retrieved from http://heinonline.org.mutex.gmu.edu/HOL/Page?handle=hein. journals/jgrj16&size=2&id=371

Broverman, I.K., Vogel, S.R., Broverman, D.M., Clarkson, F.E., & Rosenkrantz, P.S. (1972). Sex-role stereotypes: a current appraisal. *Journal of Social Issues, 28*, 59-78. doi: 10.1111/j.1540-4560.1972.tb00018.x

Dai, M. & Nation, D. (2009). Understanding non-coercive, procedurally fair behavior by the police during encounters. *International Journal of Law, Crime and Justice, 37*, 170-181. doi: 10.1016/j.ijlcj.2009.10.002

Dai, M., Frank, J., & Sun, I. (2011). Procedural justice during police-citizen encounters: the effects of process-based policing on citizen compliance and demeanor. *Journal of Criminal Justice, 39,* 159-168. doi:10.1016/j.jcrimjus.2011.01.004

Daly, K. & Bordt, R. (1995). Sex effects and sentencing: a review of statistical literature. *Justice Quarterly, 12*, 141-175. doi: 10.1080/07418829500092601

DeJong, C. (2004). Gender differences in officer attitude and behavior: providing comfort to citizens. *Women & Criminal Justice, 15*, 1-32. doi: 10.1300/J012v15n03_01

Embry, R. & Lyons, P.M. (2012). Sex-based sentencing: sentencing discrepancies between male and female sex offenders. *Feminist Criminology, 7*, 146-162. doi: 10.1177/1557085111430214

Farnworth, M. & Teske, R.H.C. (1995). Gender differences in felony court processing: three hypotheses of disparity. *Women & Criminal Justice, 6*, 23-44. doi: 10.1300/ J012v06n02_02

Foley, T. & Terrill, W. (2008). Police comfort and victims. *Victims and Offenders, 3*, 192-216. doi: 10.1080/15564880801938334

Gale, M.E. (2001). Calling in the Girl Scouts: feminist legal theory and police misconduct. *Loyola of Los Angeles Law Review, 34*. Retrieved from http://digitalcommons.lmu.edu/llr/vol34/iss2/8

Gao, J.M., Corsaro, N., Stewart, E.A., & Brunson, R.K. (2012). Examining macro-level impacts on procedural justice and police legitimacy. *Journal of Criminal Justice, 40*, 333-343. doi: 10.1016/j.jcrimjus.2012.05.002

Glick, P. & Fiske, S.T. (1996). The ambivalent sexism inventory: differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 3*, 491-512. doi: 10.1037/0022-3514.70.3.491

Herzog, S. & Oreg, S. (2008). Chivalry and the moderating effect of ambivalent sexism: individual differences in crime seriousness judgments. *Law & Society Review, 42*, 45-74. doi: 10.1111/j.1540-5893.2008.00334.x

Hinds, L. & Murphy, K. (2007). Public satisfaction with police: using procedural justice to improve police legitimacy. *Australian & New Zealand Journal of Criminology, 40*, 27-42. doi: 10.1375/acri.40.1.27

Iwamoto, D.K., Cheng, A., Lee, C.S., Takamatsu, S., & Gordon, D. (2011). "Man-ing" up and getting drunk: the role of masculine norms, alcohol intoxication, and alcohol-related problems among college men. *Addictive Behaviors, 36*, 906-911. doi:10.1016/j.addbeh.2011.04.005

Jonathan-Zamir, T., Mastrofski, S.D., & Moyal, S. (2015). Measuring procedural justice in police-citizen encounters. *Justice Quarterly, 32,* 845-871. doi: 10.1080/07418825.2013.845677

Larsen, A. (2012). Gendering criminal law: sentencing a mothering person with dependent children to a term of imprisonment. *Australian Journal of Gender and Law, 1*, (1). Retrieved from http://www.austlii.edu.au/au/journals/AUJlGendLaw/2012/3.pdf

Lytle, D.J. (2014). The effects of suspect characteristics on arrest: a meta-analysis. *Journal of Criminal Justice, 42*, 589-597. doi: 10.1016/j.jcrimjus.2014.10.001

Mastrofski, S.D. (2004). Controlling street-level police discretion. *Annals of the American Academy of Political and Social Science, 593*, 100-118. doi: 10.1177/0002716203262584

Mastrofski, S.D., Jonathan-Zamir, T., Moyal, S., & Willis, J.J. (2016). Predicting procedural justice in police-citizen encounters. *Criminal Justice and Behavior, 43,* 119-139. doi: 10.1177/0093854815613540

Mastrofski, S.D., Parks, R.B., & McCluskey, J.D. (2010). Systematic social observation. In A.R, Piquero & D. Weisburd (Eds.), *The handbook of quantitative criminology* (225-247). New York: Springer-Verlag.

Mastrofski, S.D., Reisig, M.D., & McCluskey, J.D. (2002). Police discretion toward the public: an encounter based analysis. *Criminology, 40*, 519-552. doi: 10.1111/ j.1745-9125.2002.tb00965.x

Mastrofski, S.D., Snipes, J.B., & Supina, A.E. (1996). Compliance on demand: the public's response to specific police requests. *Journal of Research in Crime and Delinquency, 33*, 269-305. doi: 10.1177/0022427896033003001

Mazerolle, L., Bennett, S., Davis, J., Sargeant E., & Manning, M. (2013). Procedural justice and police legitimacy: a systematic review of the research evidence. *Journal of Experimental Criminology, 9*, 245-274. doi: 10.1007/s11292-013-9175-2

McCluskey, J.D., Mastrofski, S.D., & Parks, R.B. (1999). To acquiesce or rebel: predicting citizen compliance with police requests. *Police Quarterly, 2*, 389-416. doi: 10.1177/109861119900200401

McCluskey, J.D., Parks, R.B., & Mastrofski, S.D. (2014). Systematic social observation in criminology in Handbook of Quantitative Criminology (12). Retrieved from http://link.springer.com/chapter/10.1007%2F978-0-387-77650-7_12#

Moulds, E.F. (1978). Chivalry and paternalism: disparities of treatment in the criminal justice system. *The Western Political Quarterly, 31*, 416-430. doi: 10.2307/ 447741

Muir, W. K. (1977). *Police: Streetcorner politicians*. Chicago, IL: University of Chicago Press.

President's Task Force on 21st Century Policing. 2015. *Final Report of the President's Task Force on 21st Century Policing*. Washington, DC: Office of Community Oriented Policing Services.

Rabe-Hemp, C. E. (2008). Female officers and the ethic of care: Does officer gender impact police behaviors? *Journal of Criminal Justice*, *3*, 426-434. doi: 10.1016/j.jcrimjus.2008.07.001

Reaves, B.A. (2015). Local police departments, 2013: personnel, policies, and practices. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, US Department of Justice.

Rodriguez, S.F., Curry, T.R., Lee, G. (2006). Gender differences in criminal sentencing: do effects vary across violent, property, and drug offenses? *Social Science Quarterly, 87*, 318-399. doi: 10.1111/j.1540-6237.2006.00383.x

Skogan, W. & Frydl, K. (Eds.). (2004). Fairness and effectiveness in policing: The evidence. Committee to Review Research on Police Policy and Practices. Washington, DC: The National Academies Press.

Sunshine, J. & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law and Society Review, 37*. Retrieved from http://www.uakron.edu/centers/conflict/docs/Sunshine.pdf

The Sentencing Project. (2007). Women in the criminal justice system: briefing sheets. Retrieved from http://www.sentencingproject.org/doc/publications/ womenincj_total.pdf

Tyler, T.R. (2004). Enhancing police legitimacy. *Annals of the American Academy of Political and Social Science, 589*, 84-99. Retrieved from http://www.jstor.org/ stable/4127668

Tyler, T.R. (2008). Legitimacy and cooperation: why do people help the police fight crime in their communities? *Ohio State Journal of Criminal Law, 6*. Retrieved from http://moritzlaw.osu.edu/osjcl/Articles/Volume6_1/Tyler-Fagan-PDF.pdf

U.S. Department of Justice. Office of Community Oriented Policing Services. (2014). *Community policing defined*. Retrieved from http://www.cops.usdoj.gov/pdf/vets-to-cops/e030917193-CP-Defined.pdf

Visher, C.A. (1983). Gender, police arrest decisions, and notions of chivalry. *Criminology, 21*, 5-28. doi: 10.1111/j.1745-9125.1983.tb00248.x

BIOGRAPHY

Emily Cahill graduated from Green Run High School, Virginia Beach, Virginia, in 2010. She graduated summa cum laude from Old Dominion University with a Bachelor of Sciences in Criminal Justice in 2014 and worked as a research consultant and Graduate Research Assistant while completing her Master of Arts at George Mason University.