The influence of police treatment and decision-making on perceptions of procedural justice: a Field Study

Terpstra, B.L.; Wijck, P.W. van

Article

Journal of Research in Crime and Delinquency
1-34

DOI: 10.1177/00224278211030968
journals.sagepub.com/home/jrc
SAGE

The Influence of Police Treatment and Decision-making on Perceptions of Procedural Justice: A Field Study

Bo L. Terpstra[1] and Peter W. van Wijck[1]

Abstract

Objectives: This study examines whether police behavior that signals higher quality of treatment or decision-making leads to higher perceived procedural justice. *Methods:* Analyses are based on data collected during police traffic controls of moped drivers in two Dutch cities over a period of six months. Police behavior was measured through systematic social observation (SSO), and data on perceived procedural justice were collected through face-to-face interviews immediately after the encounters. Linear regression analysis with bootstrap estimates was used (n = 218), with an overall perceived procedural justice scale as the dependent variable in all regressions. Independent variables included an overall observed procedural justice index and four separate scales of police treatment and decision-making. *Results:* We find no evidence that police behavior that signals fairer treatment or decision-making leads to higher perceived procedural justice.

[1] Leiden Law School, Leiden University, Netherlands

Corresponding Author:
Bo L. Terpstra, Leiden Law School, Leiden University, PO Box 9520, 2300 RA Leiden, Netherlands.
Email: b.l.terpstra@law.leidenuniv.nl

Conclusions: Our findings add to the currently very limited empirical evidence on an important question, and raise questions about a central idea, that more procedurally just treatment and decision making by authorities leads to an increase in perceived procedural justice and enhanced compliance. The first of these requires more research.

Keywords
procedural justice, behavior, perceptions, field research

Introduction

In recent years, an increasing number of studies have been published on the fairness of procedures used by the police and other authorities. Overall these studies find that if citizens feel that they are treated more fairly by legal authorities, they ascribe more legitimacy to justice institutions and tend to be more inclined to abide by the law and to cooperate (Murphy 2005; Tyler 1990; Winter and May 2001). The research on this relationship and the fairness of these procedures, termed "procedural justice" (Cropanzano and Ambrose 2001), suggests that perceptions are based on two related components: quality of treatment and quality of decision making (Blader and Tyler 2003; Gau 2014; Reisig, Bratton, and Gertz 2007; Reisig and Lloyd 2009; Sunshine and Tyler 2003; Tankebe 2009a; Tyler 1990, 2003). The research seems to imply that an improvement in the quality of treatment and decision-making by police officers leads citizens more likely to view the police as a legitimate institution, and in turn, are more likely to comply with the law and cooperate with police.

However, studies on procedural justice and compliance are generally based on survey data, so refer to perceived procedural justice rather than to actual treatment and decision-making by the police, thus essentially being about what individuals say about how they were treated rather than being about how they were actually treated, so this conclusion cannot be clearly drawn. Although one would expect that higher quality of treatment and decision-making results in higher perceived procedural justice, research on the relationship of actual behavior to perceptions of it is limited (Nagin and Telep 2017). Establishing whether actual police treatment and decision-making influence perceived procedural justice, requires study of the relationship between data on police behavior and data on citizen perceptions. Due to the labor-intensity of the field-research necessary for this, the

current body of research on this relationship is very limited and, in the studies that exist, the results are not consistent (Nagin and Telep 2017). This inconsistency leads to fundamentally different conclusions. Mazerolle et al. (2013), for example, conclude that short police-citizen interactions in traffic stops can be highly influential on perceptions of procedural justice, while Worden and McLean (2017) conclude that it would be surprising if one single interaction such as a traffic stop materially altered perceptions of procedural justice.

The main purpose of the present study is to extend the research on the relationship between police behavior and perceptions of procedural justice by answering the following research question: to what extent does police behavior that signals higher quality of treatment or decision-making lead to higher perceived procedural justice? To answer this, we investigated interactions between police officers and citizens at police traffic controls of moped drivers in two Dutch cities over a period of six months, using instruments derived from previous studies to collect data on both perceived procedural justice and on treatment and decision-making by police officers. Data on perceived procedural justice were collected using questionnaires taken from the literature (Gau 2014; Jackson et al. 2012; Sunshine and Tyler 2003; Tyler 1990), and data on actual treatment and decision-making were collected using a systematic observation protocol taken from the literature (Jonathan-Zamir, Mastrofski, and Moyal 2015).

In the next section, we present a short review of previous research on perceived procedural justice and the relationship of these perceptions with the quality of treatment and decision-making by the police. Following that, we present a more detailed description of the current study, a description of the data and the plan of analysis, the results, and conclude with a discussion of the implications and limitations.

Prior Research

This section presents an overview of prior research regarding the relation between variations in the quality of treatment and decision-making and perceived procedural justice. Generally, a distinction is made between four ingredients of procedural justice: (1) participation, (2) neutrality, (3) dignity and respect and (4) trust in the motives of the police.

First, we discuss two studies that systematically observed the four ingredients and constructed a procedural justice index (Dai, Frank, and Sun (2011), Jonathan-Zamir et al. (2015)). The strength of these studies is that

an observation protocol is used and that a validated instrument is developed. The weakness is the lack of a subjective assessment of procedural justice.

Second, we discuss three experimental studies investigating the relation between police behavior and perceptions of procedural justice (Mazerolle et al. (2012), MacQueen and Bradford (2015), Sahin et al. (2017)). These studies compare an explicitly procedural just treatment with a business-as-usual treatment. For the procedural just treatment, police officers use a concise script. The strength of the studies is the explicit experimental design. An important weakness is the use of a short script, implying that there was limited capacity to capture the full range of a procedurally just encounter.

Third, a study that combines observational data and data on perceptions of procedural justice (Worden and McLean (2017)) is discussed. The combination of these types of data is the strength of this study. The main weakness is that data on perceptions of procedural justice are based on a survey administered after two to five weeks. This led to a low response rate and potentially a less accurate reproduction of the encounter.

This study aims to build on the strengths of previous studies while avoiding the weaknesses.

The Role of Procedural Justice

Demonstrating that people are more willing to defer to unfavorable court decisions when they feel that the court procedures used to arrive at these outcomes are perceived as fair, Thibaut and Walker (1975) discussed the meaning of procedural justice in terms of control over the outcome. In their instrumental model, people seek maximal attainment of favorable outcomes and prefer fair procedures because these procedures are most likely to provide favorable (economic) outcomes in the long run (Cropanzano and Ambrose 2001).

Lind and Tyler (1988) proposed a different view on the role of procedural justice. In their group-value model, a procedurally just treatment emphasizes the perception of a shared group membership; and how authorities communicate with members of a group conveys information about the status of those members (Smith et al. 1998; Tyler and Lind 1992). Here, a procedurally just treatment sends the message that people are valued by society (Lind and Tyler 1988), strengthening the justification for obedience to an authority. The acceptance of an authority, or more specifically, the "belief that legal authorities are entitled to be obeyed and that the individual ought to defer to their judgments," is known as legitimacy (Tyler and Huo

2002: xiv). Legitimacy, in turn, leads to more respect for laws, rules and regulations issued by the authority, and the obligation to comply with these laws and cooperate with authorities (Blader and Tyler 2003; Jackson et al. 2012; Sunshine and Tyler 2003; Tyler and Fagan 2006).

Many studies confirm the importance of procedural justice, that people are more inclined to cooperate with the police and abide by the law when they feel treated in a fair, respectful and impartial manner (Hertogh 2015; Hough, Jackson, and Bradford 2013; McCluskey 2003; Murphy, Hinds, and Fleming 2008; Sunshine and Tyler 2003; Tankebe 2009b; Tyler 2004; Tyler and Wakslak 2004).

Procedural Justice Ingredients

Procedural justice is generally thought to be based on information on the quality of treatment and on the quality of decision-making. Quality of treatment involves people's assessments about whether, and to what extent, they believe police treat citizens with dignity and respect, while quality of decision making refers to people's perceptions of police as reaching decisions based on objective indicators such as facts, law, and reason rather than on personal beliefs (Gau 2011).

Tyler (2004), Schulhofer, Tyler, and Huq (2011) and Mazerolle et al. (2014) propose that four essential ingredients make up the quality dimensions of procedural justice. The first is citizen participation in the proceedings prior to an authority reaching a decision. According to Goodman-Delahunty (2010), decision-making processes are viewed as fairer when citizens are given the opportunity to voice their views and opinions. This opportunity is generally characterized as "participation" or "voice." The second is perceived neutrality of the authority in his/her decision-making, with neutral behavior signaling that police are playing by the rules set forth in the law, so indicating unbiased decisions and a fair decision-making process (Huq, Tyler, and Schulhofer 2011; Tyler 2004). The third ingredient is whether or not the authority showed dignity and respect throughout the interaction. According to Tyler and Lind (1992), "dignity and respect" is the core ingredient to procedural justice. The underlying hypothesis is that when people are treated with respect, politeness and dignity, evaluations of fair treatment, so of procedural justice, improve. The fourth ingredient is whether or not the authority conveyed trustworthy motives. Tyler (2004, 2008) proposes that citizens infer the fairness of police treatment from the motives they are able to understand from what they observe. In this reasoning, when an authority

shows care for the wellbeing of a citizen and society at large, its treatment is likely to be viewed as fairer.

Police Behavior and Perceived Procedural Justice

As the field-studies necessary to study the relationship between police behavior and perceptions of procedural justice are labor-intensive, studies investigating police behavior in terms of the four procedural justice ingredients are scarce.[1] There are, however, a few. The first, a study by Dai et al. (2011), found that, in terms of police demeanor and citizen voice, the impact of procedurally fair behavior of the police was to significantly increase citizen behaviors of respect and compliance toward the police (though the impact of other procedurally just behavior by police had a less consistent effect on citizen behavior). The second study, conducted by Jonathan-Zamir et al. (2015), of 233 police-citizens encounters between June and December 2011 in Everdene, a small suburban American city, used an observation protocol to systematically observe the four ingredients of procedural justice, (1) participation, (2) neutrality, (3) dignity and respect and (4) trust in the motives of the police. Based on the scores on these categories, the authors develop an "overall procedural justice index." As they found this index correlates significantly with observed satisfaction with the police handling of the situation, they argue it supports the validity of their measurement approach.

The results from Dai et al. (2011) and Jonathan-Zamir et al. (2015), that procedural justice increases satisfaction and cooperation with the police, are similar to the studies based on survey instruments. Both studies, however, lack subjective survey assessments on procedural justice, making it impossible to investigate to what extent higher quality in treatment and decision making by the police leads to higher perceived procedural justice.

The relation between police behavior and perceptions of procedural justice has also been investigated in a number of experimental studies. The first is the Queensland Community Engagement Trial (QCET) by Mazerolle et al. (2012). The second is a replication of the QCET study performed by MacQueen and Bradford (2015), the Scotland Community Engagement Trial. The third was an experiment conducted by Sahin et al. (2017) with the help of Turkish police.

The studies focus on police behavior during traffic controls. The setting in the three studies is slightly different: Drivers at Random Breath Test stationary operations (a routine-alcohol check), drivers stopped at routine vehicle stops, and drivers stopped by traffic officers for speeding violations.

The three studies are characterized by an experimental design. The experiment group received a "procedurally just" treatment based on a concise script, and the control group received a "business-as-usual" treatment. To investigate differences in perceived procedural justice, surveys were used. In the first two studies the survey was distributed to all drivers at the end of the encounter and the drivers were also provided with a stamped address envelope and asked to return the survey. In the last study drivers were interviewed after completion of the traffic stop.

The studies lead to contrasting results on the relation between police behavior and perceptions. Mazerolle et al. (2012) and Sahin et al. (2017) find that drivers in the procedural justice treatment condition scored significantly higher on perceived procedural justice than the drivers in the business-as-usual condition.[2] MacQueen and Bradford (2015) however, find that procedurally-just police vehicle stops decreased citizen trust in police officers and reduced satisfaction with police conduct compared to routine police vehicle stops. One potential explanation for the difference in the findings, is the that in the first two studies drivers were stopped at routine checks, whereas in the last study drivers were stopped by traffic officers for speeding violations. The contrasting results can also be explained by a difference in policing context between the different countries.

But it also draws attention to some important lessons for research. Traffic controls, such as alcohol checks, generally lead to short encounters between drivers and police officers. It appears to be very difficult to incorporate the full range of the key procedural justice ingredients into a short experimental script. The use of extensive scripts incorporating variations in the ingredients, would lead to encounters that take substantially longer than business-as-usual. Even with concise scripts, the duration of the fair treatment tends to be longer than the duration under BAU conditions, hence (small) differences in perceived procedural justice between the groups may be caused by the duration factor rather than the procedural justice elements from the script. Another lesson from these studies is that the use of paper surveys leads to a low response rate and tends be biased toward no-offenders. Furthermore, if there is a time-interval between the event and the survey, that may affect the answers in the survey.

Overall, the experimental studies do not appear to be very successful in combining data on the full range of the key ingredients of procedural justice and data on perception of procedural justice.

There appears to be only one study that successfully combines data on the key procedural justice ingredients and data on perceived procedural

justice. This is a study by Worden and McLean (2017).[3] During police patrols in Schenectady, New York, survey data acquired from 411 citizens combined with observational data made with in-car cameras, revealed a significant relation between scales that represent the officers' procedural (in)justice behavior and perceptions of procedural justice, although the variation in police behavior only accounted for 12 percent of the variations in procedural justice perceptions. When further controls are added for the nature of the situation and officers' exercise of authority, the estimated effects of the relationship between procedural justice behavior and perceptions of procedural justice disappear, though procedural injustice still has a small effect.

That study also has some limitations. An important limitation is that the low response rate (10.3 percent) may lead to sampling bias. Furthermore, the surveys were administered two to five weeks after the encounter with the police, making it difficult to determine if the survey scores are an accurate reproduction of the details of the encounter. It is problematic to determine if the variations in perceptions found were caused by the recent encounter with the police or were more representative of other influences, such as opinions from peers when talking about the encounter or pre-existing attitudes and beliefs formed on the basis of previous encounters with the police, (social) media, friends and family, or other more recent events.

The Current Study

As discussed above, there is little evidence that higher quality of police treatment and decision-making leads to higher levels of perceived procedural justice. By combining data on the essential ingredients of procedural justice of police behavior with data on citizen perceptions of procedural justice, our study investigated the relation between treatment and decision-making by police officers on the one hand, and perceived procedural justice on the other. Following Mazerolle et al. (2012), we focused on police-citizens encounters at routine traffic controls. In consultation with the Traffic department of The Hague unit of the Dutch National Police, the options to set up a field study were assessed. The option of conducting a classic experiment with the full range of the key procedural justice ingredients were limited because it would involve either longer or more varied scripts, both of which would increase the chance of within group variation in the delivery of the treatment. Since we wanted to observe the full range of procedural justice ingredients of police behavior, we used the systematic

social observation (SSO) method used by Jonathan-Zamir et al. (2015) to observe treatment and decision-making. In other words, rather than systematically varying the treatment of moped drivers, we systematically observed actual variations in police behavior. We did this by using four previously validated scales of police treatment and decision-making extracted from earlier work. Because our observations took place in a setting with encounters substantially longer in duration than random breath tests studied by Mazerolle et al. (2012), we were able to study the full range of procedural justice ingredients.

Information on perceived procedural justice was gathered through questionnaires, administered directly after the traffic controls.

In summary, our study builds on the scarce empirical research where procedural justice is studied in the context of traffic controls. Specifically, we focus on moped traffic control checks. This enables us to observe the full range of procedural justice ingredients using validated scales. It is, of course, an open question whether the results we find in the context of moped traffic controls can be generalized to traffic controls in general or broader contexts.

Set-up

Our field research focuses on traffic controls of mopeds and their drivers. Mopeds are two-wheeled motorized vehicles that can be operated by persons over 16 years of age with a valid driving license. Dutch traffic law distinguishes two kinds of mopeds: mopeds with a top speed of 25 km per hour that can be operated without a helmet and mopeds with a top speed of 45 km per hour for which wearing a helmet is compulsory.

In the Netherlands, the National Police regularly set up traffic control check-points for mopeds where they check for a number of traffic law violations: driving a vehicle with a higher top speed than allowed, driving without a valid driving license or insurance, driving under the influence of alcohol, driving without proper lighting, using a mobile phone while driving, and driving without a helmet when required. The nature of these routine checks makes them an appropriate setting for SSO-research because they take approximately five minutes, thus relatively short but substantially longer than, for example, random breath tests. In addition, the drivers stopped include both compliant and non-compliant drivers, and variation in the length of the encounters is limited due to the fact that all mopeds are thoroughly inspected.

Two different locations were selected for our research: "Wassenaarseweg" in Leiden and "1ste Stationsstraat" in Zoetermeer. Both these cities are part of the urban agglomeration in the west of the Netherlands, halfway between Amsterdam and Rotterdam. They were selected because of they are comparable in terms of the population of interest (people driving mopeds), the number of moped drivers passing the location, and the average number of traffic violations per driver stopped by the police.

The research was conducted from January 19, 2017 until August 2, 2017. On average three or four police officers were present at a traffic control check point, and one or two additional officers driving around the checkpoint in approximately a two-mile radius. After being stopped or pulled over, drivers were asked for their license and insurance papers. All mopeds were checked for defects. After visual inspection, all mopeds were placed on a roller test bench to determine the top speed. In the case of detection of a traffic law violation, drivers received a sanction.

During the above standard procedure, the interaction between police officers and drivers was observed by researchers of Leiden University. After the above procedure finished, the drivers were informed by the police that researchers of Leiden University were present at the location, inviting them to participate in a survey.

Perceived Procedural Justice

Perceived procedural justice was measured directly after the traffic control check, using a survey conducted by a pool of eight trained interviewers, student-assistants studying criminology or law at Leiden Law School, three or four interviewers per control. All interviewers received 4 hours of training on how to conduct the survey and how to interpret the questions.

To ensure that participants were able to disclose all information, the surveys, which were administered through verbal interviews on average seven minutes long, were conducted approximately 50 meters from the traffic control check. The survey covered a wide range of topics in the field of procedural justice, using questions derived from previous research (Gau 2014; Sunshine and Tyler 2003; Tyler 1990), related both to the traffic control that had just taken place as well as to previous encounters with the police. Most responses were measured using a seven-point Likert scale (answers ranging from 1 to 7, where 1 is "totally disagree" and 7 "totally agree"). The survey was tested and slightly modified after two pilot traffic controls in November 2016. The main reasons for the modifications were that two items were not representative of the situation of moped checks, one

item was difficult to interpret for drivers, and two items were highly correlated with other items ($r > .95$, $p < .001$) so, due to time restrictions, were omitted.

To construct an overall perceived procedural justice scale, we calculated the average of the following six (Likert scale) items: (1) "The officer treated me with respect," (2) "The officer treated me fairly," (3) "The officer took the time to listen to what I had to say," (4) "The officer treated me the same as other people," (5) "The officer made decisions on the basis of the facts of the situation, and not on her/his personal opinions," and (6) "The officer explained her/his actions and decisions to me."

Police Treatment and Decision-making

The observations of treatment and decision-making by the police were performed using a systematic social observation protocol (SSO) by student-assistants who also conducted the surveys. To allow observers to overhear conversations without influencing them, for each check, two to three observers were placed at a distance of at least five meters, on average seven meters. All observers received six hours of training on how to score the systematic observation-protocol. To reduce the potential problem of different scoring methods, inter-observer differences were intensively studied and discussed during this training. These differences were tested during the pilot traffic controls in November 2016 and found to be negligible. This was confirmed during the main phase of the field research, in which some drivers were randomly selected to be observed by multiple observers. Due to the nature of the checks, all interactions between police and drivers lasted longer than one minute.

The observation protocols are derived from Jonathan-Zamir et al. (2015), and bear similarities to protocols used by Worden and McLean (2017) and McCluskey (2003) applied to traffic encounters as well as to a broader range of police-citizen encounters. Based on decades of SSO research, Jonathan-Zamir et al. (2015) developed a systematic observation protocol that assesses items that aim to capture police behaviors that make citizens feel that they have been treated fairly.[4] Based on these items, they constructed four scales of police treatment and decision-making, based on the four essential ingredients that, according to previous research, constitute the quality dimensions of procedural justice: (1) participation, (2) neutrality, (3) dignity and respect and (4) trust in the motives of the police. In the following sub-sections, we discuss these four scales in more detail.

Participation. Jonathan-Zamir et al. (2015) based their construct of participation on observable choices made by police-officers. Following their definition and observation-items, our observers recorded whether citizens were asked for information or viewpoints, and whether they provided information or viewpoints. The "interest" the officer showed in the information provided was also recorded by looking at confirmatory and non-confirmatory behaviors such as nodding, humming, summarizing, carrying out other activities during the interaction and ignoring information provided. The items were coded and grouped as follows:

> Participation = The officer asked for information/viewpoint (0 = no; 1 = yes) + The citizen provided information/viewpoint (0 = no; 1 = yes) × The officer expressed interest in the information/viewpoint (on a scale ranging from 0 to 3, where 0 = dismissive listener; 1 = inattentive listener; 2 = passive listener; 3 = active listener, as defined in the coding protocols).

This formula resulted in a participation scale ranging from 0 (very low) to 4 (very high).

Neutrality. To construct a measure of neutrality, Jonathan-Zamir et al. (2015) used three types of items: the desire for a balanced information-gathering process, the absence of any obvious indication of decision-making bias based upon personal characteristics, and transparency of decision-making by articulating the reasons for the officer's choices. In our study, we used the same observation-items. For example, if an officer explains to a citizen why the traffic control is being conducted, or explicit statements are made that stress the neutrality of the officers in question. We constructed the neutrality measure as follows:

> Neutrality = Officer indicated s/he would seek all viewpoints about the matter at hand (0 = no; 1 = yes) + Officer indicated s/he would not make a decision about what to do until s/he had gathered all the necessary information (0 = no; 1 = yes) + Officer did not indicate that his/her decisions in this situation were influenced by the personal characteristics (race, age, sex) of anyone present (0 = no; 1 = yes) + Officer explained why the police became involved in the situation (0 = no; 1 = yes) + Officer explained why s/he chose to resolve the situation as s/he did (0 = no; 1 = yes).

This formula resulted in a neutrality scale ranging from 0 (very low) to 5 (very high).

Dignity and respect. The scale measuring dignity and respect was originally constructed by Jonathan-Zamir et al. (2015) with independent measures of respect and disrespect. Due to the lack of observations in the disrespect category, they created a single dignity measure of respect. We extended this by observing speech and gestures indicating (dis)respect, such as using a loud voice, interruptions and belittling remarks as indications of disrespect and greetings, compliments, jovial gestures, saying "thank you," good-humored and friendly remarks as indications of respect. The duration or frequency of such actions during the encounter (brief/intermittent/dominant) was also noted, resulting in the following scale:

> *Dignity = To what extent did the officer behave respectfully toward the citizen?* (on a scale ranging from 0 to 4, where 0 = Officer showed disrespect; 1 = Officer showed neither respect nor disrespect—"business-like" behavior; 2 = Officer showed brief respect; 3 = Officer showed intermittent respect; 4 = Officer showed dominant respect).

The scale ranges from 0 (disrespect) to 4 (dominant respect), with higher scores indicating higher levels of dignity and respectful behaviors by police officers.

Trustworthy motives: showing care and concern. To construct a concept reflecting trustworthy motives, Jonathan-Zamir et al. (2015) used observation items that note when police officers provide something to citizens that they requested or would unambiguously perceive as beneficial. These behaviors indicate care and concern, reflecting higher levels of trustworthy motives. Police can exhibit such care and concern in several ways: an officer can comfort a citizen, can promise to give the citizen's situation special attention, tell or ask the citizen to call if the citizen's problem recurs, or—at the officer's initiative—provide information or physical assistance, or contact an agency for assistance on the citizen's behalf. The concept is constructed as follows:

> Trust in the motives of the decision-maker: Showing care and concern = The officer asked the citizen about his/her well-being or asked others in a way that the citizen observed it (0 = no; 1 = yes) + The officer offered comfort or reassurance to the citizen (0 = no; 1 = yes) + The officer provided or promised to exert control or influence over another person for the citizen (0 = no; 1 = yes) + The officer filed a report or promised to file a report for the citizen (0 = no; 1 = yes) + The officer acted or promised to act on behalf

of the citizen with a government agency or private organization (0 = no; 1 = yes) + The officer provided/arranged or promised to provide/arrange physical assistance to the citizen (0 = no; 1 = yes) + The officer provided or promised to provide advice on how the citizen could handle the situation or deal with the problem (0 = no; 1 = yes).

The scale depicting trustworthy motives ranges from 0 (very low) to 7 (very high).

Overall observed procedural justice behavior index. Following Jonathan-Zamir et al. (2015), we also developed a composite index based on the four separate indices of police treatment and decision-making described above. This composite index is intended to be a broad assessment of the officer's behavior, its antecedents and its outcomes. The four separate scales were averaged into an overall observed procedural justice index.

Description of the Data

In the period between January 19, 2017 and August 2, 2017, 687 moped drivers were stopped at traffic control checks, 299 of whom participated in the survey (43.5 percent response rate). Of the 687 drivers stopped, 590 were observed. Not all drivers who participated were also observed as, on several occasions, the number of drivers stopped exceeded the number of observers present. Ultimately, 218 of the collected surveys could be matched to an observation and were included in our sample. Of these matches, 210 drivers were observed once, and four drivers were observed twice. Not all respondents who completed the interview answered every question. More specifically, with only four exceptions, the missing data relates to the questionnaire item about the police officer taking the time to listen, which was unanswered in 25 of 218 questionnaires. According to Little's multivariate-test, ($\chi^2(26) = 12.740$, $p = .986$), on all missing data, the likelihood of missingness depends neither on the observed data nor on the missing data. Consequently, due to the reduced sample size, ignoring missing data will increase the SE of the sample estimates rather than introducing bias (Dong and Peng 2013). To respond to this, missing data was substituted using the expectation maximization algorithm (Dempster, Laird, and Rubin 1977), based on all questionnaire-items on procedural justice and 50 iterations. This algorithm provides unbiased parameter estimates and improves statistical power of analyses when only a very small part of that data is missing (in this case 2.2 percent) (Enders 2001; Scheffer 2002).[5]

Table 1. Descriptive Statistics of the Sample of Drivers that Were Observed and Participated in the Survey (N = 218).

Variable	Score
Age in years: mean (SD)	33.0 (16.1)
Sex: % male	56.4
Member of ethnic minority group: %[a]	17.1
Household income: median class[b]	€ 20,000–30,000
Education: median class[c]	High school 2
Sanctioned by police during current traffic control: % yes	21.1

[a]The Central Bureau of Statistics in the Netherlands defines a person with a migration background as someone with (at least) one of his/her parents born abroad. [b]Income was measured by asking respondents to classify their gross household income in 2016: € 0–10,000, €10,000–20,000, € 20,000–30,000, € 30,000–50,000, € 50.000+ and unknown. [c]With respect to their education respondents were asked about the highest achieved level of schooling, which was then classified as: elementary, vocational, high school levels 1, 2 and 3, college/university and unknown.

The details of the observed population who participated in the survey are presented in Table 1. Interestingly, the descriptives of the total observed population (N = 590) are similar to the descriptives of the sample that was observed and participated in the survey (N = 218). For example, in the total observed sample, the ratio of offenders to non-offenders was 18.1 percent, compared to 21.1 percent in the sample of observed drivers who also participated in the survey ($\chi^2(1) = .192$, $p = .340$); and the ratio of males to females in the total observed sample was 58 percent, compared to 56.4 percent in the sample that also participated in the survey ($\chi^2(1) = .155$, $p = .693$). Kruskal-Wallis Tests were also conducted to examine whether observed police behavior differed in the total observed population compared to the sample with drivers that were observed and participated in the survey. No significant differences in participation ($\chi^2(1) = 2.171$, $p = .141$), neutrality, ($\chi^2(1) = .0951$, $p = .758$), dignity and respect ($\chi^2(1) = .120$, $p = .729$), and trustworthy motives ($\chi^2(1) = .594$, $p = .441$) were found. Based on these tests, we conclude that there are no systematic differences between the observed population and the population that participated in the survey.

Perceived Procedural Justice

The dimensionality of the perceived procedural justice scale was examined using different techniques. Table 2 shows the correlations and descriptives

Table 2. Correlation Matrix and Descriptive Statistics of the Six Items of Perceived Procedural Justice and the Perceived Procedural Justice Scale (N = 218).

	1	2	3	4	5	6	7
1. Perceived procedural justice index	1						
Items							
2. The officer treated me with respect	.768*	1					
3. The officer treated me fairly	.803*	.758*	1				
4. The officer took the time to listen to what I had to say	.794*	.552*	.565*	1			
5. The officer treated me the same as other people	.645*	.324*	.360*	.432*	1		
6. The officer made decisions on the basis of the facts of the situation, and not on her/his personal opinions	.740*	.437*	.542*	.529*	.450*	1	
7. The officer explained her/his actions and decisions to me	.671*	.380*	.412*	.444*	.225*	.390*	1
Range	2.33–7	1–7	1–7	1–7	1–7	1–7	1–7
M	6.29	6.42	6.44	6.19	6.18	6.40	6.11
SD	.732	.981	.862	.998	1,121	.864	1,192

*$p < .01$.

of the items on perceived procedural justice used in the questionnaire, together with the overall procedural justice scale. The mean inter-item correlation for the items is .453 (range: .225 to .798). Mean-item total correlation is .737 (range: .645 to .803). This suggests that all elements are well presented by the overall scale. A third indicator used to gauge the internal consistency of the perceived procedural justice scale, Cronbach's coefficient α, is .819 in this case. Acceptable values of α range from 0.70 to 0.95 (Nunnally and Bernstein 1994).

Factor-analytic techniques were used to further investigate whether the six survey items loaded on the perceived procedural justice scale. We used principal axis factor analysis because it corrects for measurement error by using more conservative score reliability estimates (Velicer and Jackson 1990). The Kaiser-Meyer-Olkin measure of sampling adequacy is .817, indicating that the data are appropriate for factor-analytic techniques (Comrey and Lee 2013). The factor results indicate a one factor solution: a single factor with an eigenvalue ($\lambda = 3.295$) above the Kaiser-Guttman

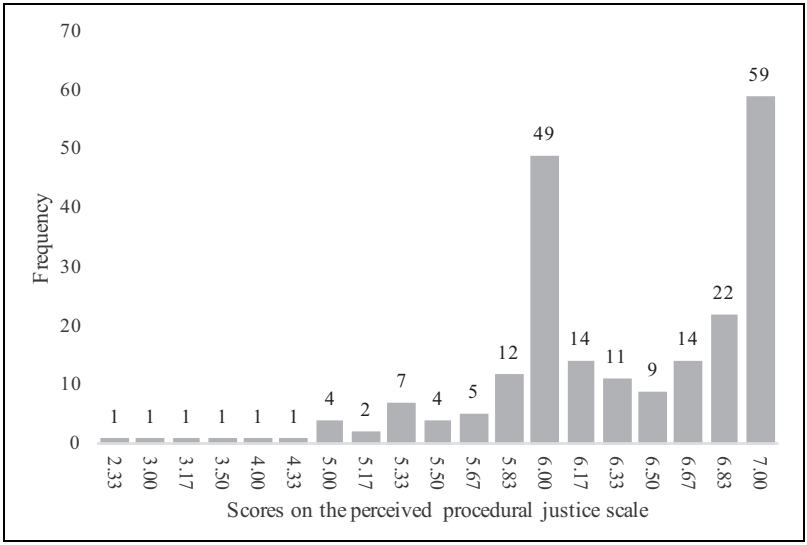

Figure 1. Frequency distribution of scores on the perceived procedural justice scale.

criterium ($\lambda > 1$) and a scree plot supporting this conclusion. The techniques we used to investigate the dimensionality of the perceived procedural justice scale all indicate one dimension.

Figure 1 depicts the distribution of the scores on the perceived procedural justice scale, based on the six questionnaire items, showing a negatively skewed distribution with a relatively high mean. Although comparison of this pattern of perceived procedural justice to those reported in previous research on police-citizen contacts is complicated by differences in sampling, for example the reason for contact with the police, overall, it appears that citizens' subjective experiences in our sample are similar to those reported in previous research on routine traffic stops.

Police Treatment and Decision-making

The distribution of the observation-scores of the four categories of police behavior are shown in Figure 2. Most observations of "participation," are in the categories "high and very high." We see no need to alter the construct. The distribution of "neutrality" has most observations in categories "very low" and "low." In the category "very high," there is only one observation.

For the purpose of our study, we regrouped the categories and merged "very high" with "high." The distribution of "dignity and respect" shows most of the scores in the category "dominant respect," a single observation in the category "brief respect," and the absence of scores in the category "disrespect." A more detailed overview of the different items used to construct the four categories of observed police behavior is provided in Table 3. These details do not fundamentally alter the construct. We did merge brief respect with business-like respect.

The construct of "trustworthy motives" of the decision-maker is of more concern. In Figure 2, we see that the majority of the scores is in the category "very low." The reason can be seen in Table 3, where we see that two of the items used in the construct have not been observed in our study. In addition, for the observed behaviors that did occur during our study, we see that the only item of significance concerns advice on handling the situation. Due to the low number of observations in the category "moderate," we merged this with the category "low."

The distribution of the scores on the overall observed procedural justice scale, based on the four separate indices of police treatment and decision-making described above, is depicted in Figure 3. Note that that the "observed procedural justice scale" is based on observations by researchers using an observation protocol, whereas the "perceived procedural justice scale" is based on perceptions of citizens as revealed in survey research.

Jonathan-Zamir et al. (2015) argue persuasively that, rather than reflecting an underlying construct, the four ingredients form a construct, which implies that they are not expected to develop from a single latent variable. The various behaviors are viewed as tapping different facets of treatment and decision-making, and are not expected to be intercorrelated. Consequently, the dimensionality analysis is restricted to polychoric correlation coefficients for the four constructs of police behavior together with the overall observed procedural justice scale (Muthén and Kaplan 1985).

The results in Table 4 show mostly low and insignificant inter-item correlations (range −.074 to .364), and medium to strong item total correlations (range .470 to .793). The Kaiser-Meyer-Olkin measure of sampling adequacy is .461, indicating that, overall, the four constructs have too little in common to warrant a factor analysis (Comrey and Lee 2013). Since this supports the view that the four ingredients are not reflective of an underlying construct, we find no reason to deviate from the four categories proposed in previous research (Jonathan-Zamir et al. 2015; Schulhofer et al. 2011; Sunshine and Tyler 2003; Tyler 2004).

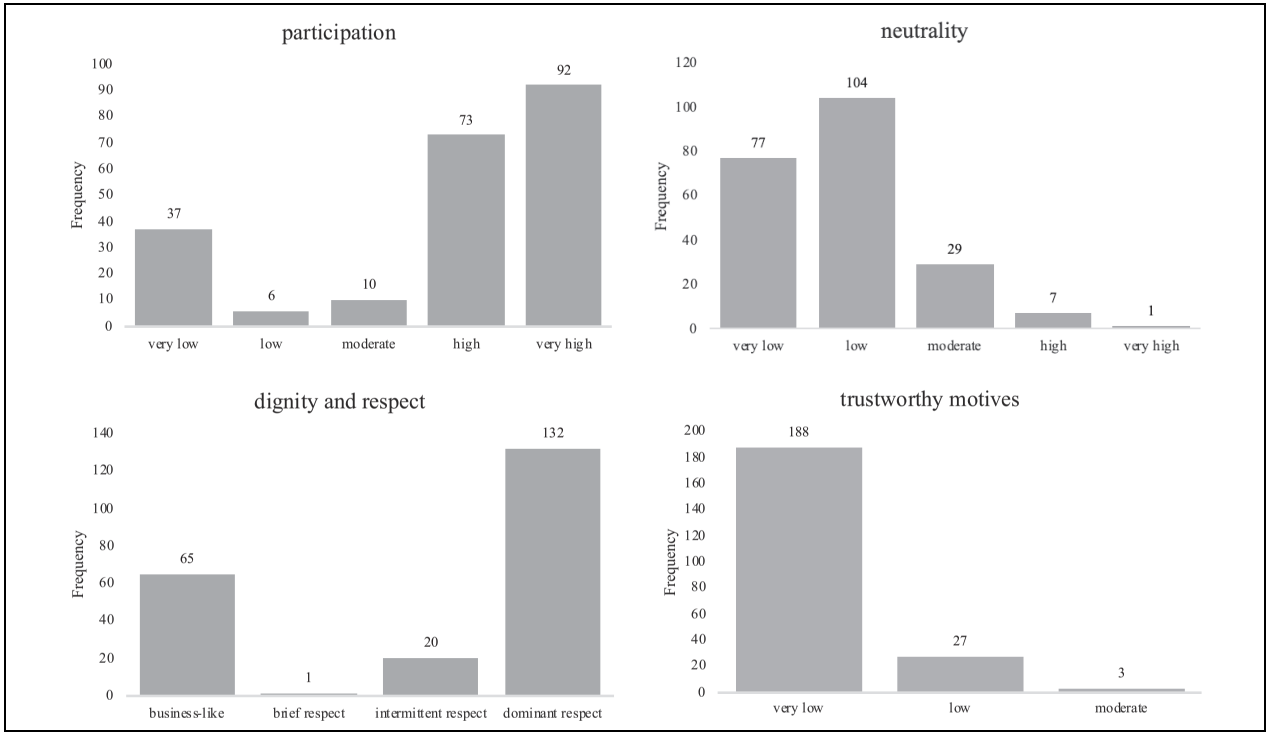

Figure 2. Frequency distributions of scores on "participation," "neutrality," "dignity and respect" and "trust in the motives of the decision-maker: showing care and concern" in police-citizen encounters.

Table 3. Descriptive Statistics of the Individual Observation Items Composing the Four Categories of Police Behavior (N = 218).

	Values	%
Participation		
Officer asked the citizen to provide information/viewpoint	Yes	76.6
Citizen provided information/viewpoint	Yes	81.2
Officer expressed interest in information/viewpoint	Dismissive	.5
	Inattentive	2.8
	Passive	52.7
	Active	44.0
Neutrality		
Officer expressed desire to hear all viewpoints	Yes	5.0
Officer indicated he would not make a decision about what to do until s/he had gathered all the necessary information	Yes	3.7
Officer indicated that his decisions in this situation were influenced by the personal characteristics (race, age, sex) of anyone present (reversed)	Yes	.9
Officer explained why the police carries out routine moped checks	Yes	17.0
Officer explained why s/he chose to resolve the situation ass/he did	Yes	59.3
Dignity and respect		
Officer showed respectful behaviors to this citizen during the encounter	Yes	70.2
Duration of the officer's respectful behaviors	Brief	1.4
	Intermittent	33.5
	Dominant	65.1
Officer showed disrespectful behaviors to this citizen during the encounter	Yes	.0
Trustworthy motives: Showing care and concern		
Officer asked about citizen's well-being	Yes	.5
Officer offered comfort or reassurance to this citizen	Yes	1.8
Officer provided or promised to exert control or influence over another person for the citizen	Yes	.5
Officer filed a report or promised to file a report for the citizen	Yes	.5
Officer acted or promised to act on behalf of the citizen with a government agency or private organization	Yes	.0
Officer provided/arranged or promised to provide/arrange physical assistance to the citizen	Yes	.0
Officer provided or promised to provide advice handling the situation/problem	Yes	11.9

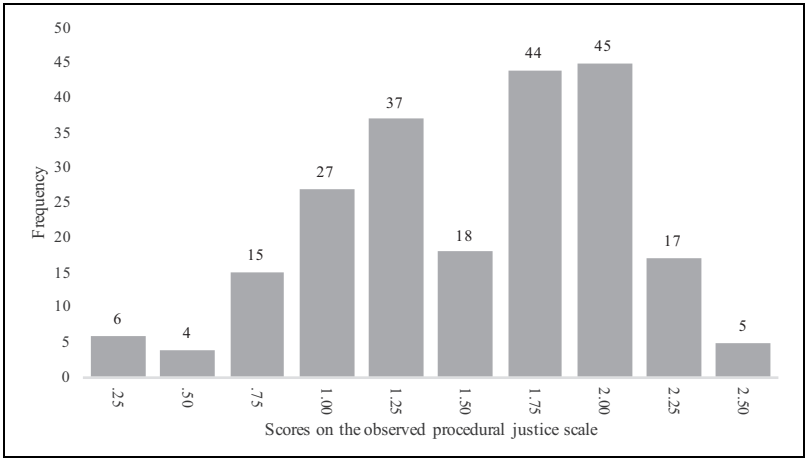

Figure 3. Frequency distribution of scores on the overall observed procedural justice scale.

To reflect previous research by Worden and McLean (2017) and research by Jonathan-Zamir et al. (2015), on which we based our observation protocol, we also retain the overall observed procedural justice scale used in their research.

Plan of Analysis

In order to answer the research question "To what extent does police behavior that signals higher quality of treatment or decision-making lead to higher perceived procedural justice?," we used linear regression analysis. The most commonly used regression technique, Ordinary Least Squares (OLS), requires that residuals are random and normally distributed (Field 2013) but this assumption does not hold true in our analyses. Since a transformation of the data did not solve the problem, we used bootstrapping, a nonparametric approach to effect-size estimation and hypothesis testing that makes no assumptions about the shape of the distributions of the variables or the sampling distribution of the statistic (Efron 1982).[6] The results presented in the next section are therefore based on 1,000 bootstrap iterations using bias-corrected and accelerated (BCa) bootstrap intervals (Efron and Narasimhan 2020).

The dependent variable in all regressions is the overall perceived procedural justice scale. In our first analysis, the independent variables are

Table 4. Polychoric Correlation Coefficients and Descriptive Statistics of the Four Categories of Police Behavior and the Overall Observed Procedural Justice Scale (N = 218).

	1	2	3	4	5
1. Obs overall PJ	1				
Sub-indices					
2. Participation	.793*	1			
3. Neutrality (revised)	.542*	.191	1		
4. Dignity and respect (revised)	.648*	.319*	−.026	1	
5. Trustworthy motives: care and concern	.470*	.051	.364*	−.074	1
Range	0–2.5	0–4	0–3	0–3	0–1
M	1.526	.281	.850	2.30	.140
SD	.523	1.442	.784	.906	.345

* $p < .05$.

the four scales of police treatment and decision-making: participation, neutrality, dignity and respect and trustworthy motives. All four scales were coded using dummy variables with the lowest category as the reference category, i.e. for participation, the category "very low" is the reference category, and four dummy variables represent the categories "low," "moderate," "high" and "very high." Similarly, for neutrality and trustworthy motives, "very low" is the reference category. For dignity and respect, "business-like" acts as reference category.

If one or more of the estimated parameters of these dummies proves to be significant, it is evidence that variations in treatment and decision-making by police officers affect perceived procedural justice. Based on previous research, we expected the parameters to be positive, i.e. when police officers exhibit more behavior that transmits signals of fairer treatment and decision-making, we expected perceived procedural justice to increase.

We also performed a second regression in which the independent variable is the overall observed procedural justice index. This index is useful to obtain a broad assessment of the officer's behavior (Jonathan-Zamir et al. 2015). We expected this relationship to be positive, i.e. when police officers exhibit overall more behavior that transmits signals of fairer treatment and decision-making, we expected perceived procedural justice to increase.

Both the first and the second regression were performed with and without covariates on age, sex, income, education and a dummy variable that depicts whether or not a driver was sanctioned during the traffic control checks. The sanction dummy was added because previous research has

shown that perceptions of procedural justice can be attenuated by the outcome of an encounter with the police (Worden and McLean 2017).

A statistical power analysis was performed using G*Power 3.1 (Faul et al. 2009) to determine the minimal detectable effect (MDE) identifiable by our study. With an α of .05 and power of 0.80, the MDE (f^2) with our sample size ($N = 218$) ranges between .0363 for the model with the single overall observed procedural justice scale and .0836 for the regression with the four procedural justice scales including covariates. Thus, depending on the model, we are able to identify small ($f^2 \geq .02$) or medium ($f^2 \geq .15$) effect sizes (Cohen 1988).

Results

In this section, we discuss the results of the regressions. We first present the regression results with the four scales of police treatment and decision-making as independent variables, then we present the results of the regressions with the overall observed procedural justice as independent variable. Our results do not support the idea that higher quality of police treatment and decision-making leads to higher levels of perceived procedural justice. In Table 5, the results of the regression with the four scales of police treatment and decision-making (Model A) show that most relevant coefficients are insignificant. We find a significant relationship only between neutrality and perceived procedural justice. This specific relationship is not consistent with our expectations. When the neutrality of treatment and decision-making by police officers is low, compared to it being very low, drivers' perception of procedural justice declines. This indicates that drivers perceive a slight improvement in neutrality from the lowest level of neutral behavior as a signal that they are being treated less procedurally fairly.

Importantly, the proportion of variance of the regression that is explained is relatively small. Only 8.5 percent of the variation in perception can be explained by the variation in actual treatment and decision-making, and just 4.1 percent when looking at the adjusted R-squared value. This implies that the vast majority of perception of treatment by the police and, more specifically, perceived procedural justice is determined by factors other than the elements of procedural justice observed in this study.

In model B in Table 5, the sanction dummy and demographic and socio-economic characteristics are included.[7] However, the relationships between the added covariates and the perception of procedural justice are all insignificant, causing a larger loss in degrees of freedom compared to the

Table 5. Regression Results with "Perceived Procedural Justice" as Dependent Variable and the Four Scales of Police Treatment and Decision-making as Independent Variables (N = 218).

	Model A						Model B					
					BCa 95% C.L.						BCa 95% C.L.	
	B	Bias	S.E.	Sig.	Lower	Upper	B	Bias	S.E.	Sig.	Lower	Upper
Constant	6.530	−.010	.155	.001	6.238	6.794	6.534	−.012	.218	.001	6.088	6.949
Observed participation low	−.192	.011	.339	.543	−.875	.505	−.149	.026	.276	.569	−.801	.460
Observed participation moderate	.226	.004	.201	.245	−.203	.638	.270	.004	.220	.201	−.197	.707
Observed participation high	−.110	.015	.142	.442	−.425	.240	−.138	.018	.146	.349	−.434	.182
Observed participation very high	−.092	.009	.130	.485	−.326	.199	−.059	.004	.137	.687	−.307	.226
Observed neutrality low	−.301	.004	.103	.011	−.518	−.083	−.301	.013	.111	.018	−.565	−.028
Observed neutrality moderate	.072	.010	.132	.605	−.177	.355	.081	.008	.136	.549	−.192	.381
Observed neutrality high	−.168	.010	.295	.562	−.770	.379	−.063	.008	.309	.843	−.694	.529
Observed respect intermittent	.244	.003	.170	.149	−.089	.563	.256	.000	.177	.155	−.111	.612
Observed respect dominant	−.012	.006	.128	.910	−.298	.284	−.043	.002	.134	.755	−.333	.237
Observed trust in motives low	−.256	−.011	.171	.135	−.627	.044	−.286	−.013	.176	.095	−.655	.012
Demographic and socio-economic characteristics							....	....	....	....	....	....
R^2				.085						.160		
adj. R^2				.041						.050		
F-value				1.925						1.461		
p				.044						.081		

Note. Estimated parameters are based on 1,000 bootstrap iterations using bias-corrected and accelerated (BCa) bootstrap intervals.

loss in sum of squared errors, hence a lower F-value. Adding the covariates to the model does not cause a better fit.

The results in Table 6 show that when we take the overall observed procedural justice index as independent variable, the results do not change. As with the different categories of behavior, a broad assessment of the officer's behavior also does not significantly influence perceptions of procedural justice.

Discussion and Conclusion

A considerable volume of research has shown that citizens are more likely to comply with rules and regulations and to cooperate with the police when they believe that the police act in a procedurally just manner. However, little is known about the relationship between how people are treated and perceptions of procedural justice. Investigating this requires data on both police behavior and perceptions of procedural justice. We therefore investigated interactions between police officers and citizens, here moped drivers, at police traffic controls in two Dutch cities over a period of six months. We collected data on police behavior using systematic social observation, and data on perceived procedural justice using a survey administered directly after the traffic controls. Both of the methods, systematic observation and of survey items, were derived from previously validated research.

In police-citizen encounters at routine traffic controls, we found no evidence that police behavior that signals fairer treatment or decision-making leads to higher perceived procedural justice. Conversely, when police behavior that signals neutrality, we found that drivers perceive a slight improvement in neutrality from the lowest level of neutral behavior as a signal that they are being treated less procedurally fairly. Our results on police treatment of moped drivers are in line with previous research by Worden and McLean (2017) on the relationship between police behavior and perceptions of procedural justice. Based on a more diverse sample of encounters, they concluded that police behavior in a single encounter does not substantially influence perceptions of procedural justice. As discussed before, Worden and McLean may have been influenced by the low response rate and by the fact that they gathered their survey data two to five weeks after the interaction between policy officers and citizens. As, on the one hand, memory decay may give rise to random errors and, on the other hand, events after the encounter with the police, such as discussions with peers of the encounter, may influence the recollection of the encounter, we tried to

Table 6. Regression Results with "Perceived Procedural Justice" as Dependent Variable and the Overall Observed Procedural Justice Index as the Independent Variable (N = 218).

	Model A						Model B					
					BCa 95% C.L.						BCa 95% C.L.	
	B	Bias	S.E.	Sig.	Lower	Upper	B	Bias	S.E.	Sig.	Lower	Upper
Constant	6.548	.003	.143	.001	6.238	6.857	6.542	.010	.225	.001	6.084	7.036
Overall observed procedural justice	−.163	−.002	.093	.079	−.353	.019	−.155	−.002	.101	.116	−.364	.039
Demographic and socio-economic characteristics							....	....	....	....	....	....
R^2				.013						.084		
adj. R^2				.009						.011		
F-value				2.924						1.156		
p				.089						.306		

Note. Estimated parameters are based on 1,000 bootstrap iterations using bias-corrected and accelerated (BCa) bootstrap intervals.

minimize such potential problems by administering our surveys immediately after the traffic control check.

Our results can probably be attributed to the high ratings of perceived procedural justice, even when officers' behavior represents low-to-moderate levels of quality of treatment and decision-making. This implies that once a certain level of perceived procedural justice is reached, better quality of treatment or decision-making cannot improve citizens' subjective assessments very much, and other factors become more important in further enhancing the perception of procedural justice. As Gau (2014) noted, these elements can consist of pre-existing attitudes and beliefs that have formed based on previous encounters with the police, (social) media, friends and family, or other socialization processes.

The study also has limitations. The first is that the setting of the field study was neither longitudinal nor a true experiment. This limits the control over interference from variables that were not included in our analysis, such as pre-existing beliefs about the police. A second limitation is the external validity of the results. Our findings are based on the behaviors of Dutch police officers during routine moped traffic control checks. This setting is well suited for observing the full range of procedural justice ingredients of police behavior, since an encounter takes five minutes on average and the sample of drivers stopped, consisting of both offenders and non-offenders. However, the specific setting of moped drivers makes it difficult to extrapolate our outcomes to formulate a general theory on the relationship between treatment and decision-making on the one hand, and perceived procedural justice on the other.

These limitations, however, do not override the fact that, with our study we intended to add to the literature on an underexposed element in procedural justice research, i.e. the relation between specific categories of behavior of the police and perceived procedural justice. In a real-life setting in which we were able systematically observe the full range of procedural justice ingredients of police-behavior and decision-making, with a high response rate and the absence of an offender-bias, we did not find that higher quality of police treatment and decision-making leads to higher levels of perceived procedural justice.

Our findings raise questions about one of the main ideas in the procedural justice literature: that more procedural just treatment and decision making by authorities leads citizens more likely to view the police as a legitimate institution, and in turn, are more likely to comply with the laws and cooperate with police. A single encounter with police may be less important than assumed in shaping the pathway from procedural justice

perceptions to compliance. This does not imply that police officers should not be concerned with respectful treatment, voice, trustworthiness or neutrality, rather that we need to further investigate how these behaviors can contribute to the accumulation of influences on perceptions of procedural justice.

Two lines of future research on this relationship are likely to be fruitful. The first is more research based on the full range of procedural justice ingredients of police behavior combined with perceptions on procedural justice. The main improvement of SSO-research over experimental studies is its ability to incorporate all procedural justice ingredients of behavior without asking too much of the police officers involved. Results from different settings and larger sample sizes: different settings may contribute to a better understanding of the conditions under which police behavior can influence perceptions, and larger sample sizes could contribute by being able to detect smaller differences at the margin. The second line of research would be to use multiple points of measurement over time to accurately investigate how changes in perceptions due to police encounters are influenced by other elements such as pre-existing attitudes and beliefs, (social) media and friends and family.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD


Bo L. Terpstra https://orcid.org/0000-0001-9455-6912

Notes

1. For a good overview, see Jonathan-Zamir et al. (2015).
2. Different articles based on the same data of the Queensland Community Engagement Trial show comparable results (Mazerolle et al. 2013; Murphy, Mazerolle, and Bennett 2014; Sargeant et al. 2016).
3. A study by Willits, Makin, and Koslicki (2019) also combines procedural justice behavior data with survey data on procedural justice but lacks statistical power due to the limited number of respondents.

4. By following the method by Jonathan-Zamir et al. (2015), we recognize that the focus is on behaviors that indicate procedurally just treatment. Although previous research has shown that negative experiences have a greater impact on judgements of encounters with the police (Skogan 2006), our study is not aimed at procedural injustice, rather we investigate, using previously validated instruments, the extent to which police behavior that signals higher quality of treatment or decision-making leads to higher perceived procedural justice.
5. Alternative methods of handling missing data, such as full information maximum-likelihood (FIML) and multiple imputation (MI), have been applied to the data and resulted in comparable results. EM was chosen because it allows for data imputation independently of model estimations.
6. Different transformations of the dependent variables were also applied, but all possible solutions still violated the normality assumption of normally distributed residuals. Dichotomization of the dependent variable was also considered but not executed because it often yields misleading results (MacCallum et al. 2002).
7. For reasons of space, in Table 5 we have omitted the estimates for the demographic and socio-economic characteristics, which were mostly insignificant, and followed a rather erratic pattern as far as they were significant.

References

Blader, Steven L. and Tom R. Tyler. 2003. "A Four-Component Model of Procedural Justice: Defining the Meaning of a "Fair" Process.'" *Personality and Social Psychology Bulletin* 29(6):747-58.

Cohen, J. M. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates.

Comrey, Andrew L. and Howard B. Lee. 2013. *A First Course in Factor Analysis*. London: Psychology Press.

Cropanzano, Russell and Maureen Ambrose. 2001. "Procedural and Distributive Justice Are More Similar than You Think: A Monistic Perspective and a Research Agenda." Pp. 119-51 in *Advances in Organization Justice*, edited by J. Greenberg and R. Cropanzano. Palo Alto, CA: Stanford University Press.

Dai, Mengyan, James Frank, and Ivan Sun. 2011. "Procedural Justice during Police-Citizen Encounters: The Effects of Process-Based Policing on Citizen Compliance and Demeanor." *Journal of Criminal Justice* 39(2):159-68. doi: 10.1016/j.jcrimjus.2011.01.004.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data Via the *EM* Algorithm." *Journal of the Royal Statistical Society: Series B (Methodological)* 39(1):1-22. doi: 10.1111/j.2517-6161.1977.tb01600.x.

Dong, Yiran and Chao-Ying Joanne Peng. 2013. "Principled Missing Data Methods for Researchers." *SpringerPlus* 2(1):222. doi: 10.1186/2193-1801-2-222.

Efron, Bradley. 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Efron, Bradley and Balasubramanian Narasimhan. 2020. "The Automatic Construction of Bootstrap Confidence Intervals." *Journal of Computational and Graphical Statistics* 1-12. doi: 10.1080/10618600.2020.1714633.

Enders, Craig K. 2001. "A Primer on Maximum Likelihood Algorithms Available for Use with Missing Data." *Structural Equation Modeling: A Multidisciplinary Journal* 8(1):128-41. doi: 10.1207/S15328007SEM0801_7.

Faul, Franz, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. 'Statistical Power Analyses Using G*Power 3.1: Tests for Correlation and Regression Analyses'. *Behavior Research Methods* 41(4):1149-60. doi: 10.3758/BRM.41.4.1149.

Field, Andy. 2013. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. Thousand Oaks, CA: Sage Publications.

Gau, Jacinta M. 2011. "The Convergent and Discriminant Validity of Procedural Justice and Police Legitimacy: An Empirical Test of Core Theoretical Propositions." *Journal of Criminal Justice* 39(6):489-98. doi: 10.1016/j.jcrimjus.2011.09.004.

Gau, Jacinta M. 2014. "Procedural Justice and Police Legitimacy: A Test of Measurement and Structure." *American Journal of Criminal Justice* 39(2):187-205. doi: 10.1007/s12103-013-9220-8.

Goodman-Delahunty, J. 2010. "Four Ingredients: New Recipes for Procedural Justice in Australian Policing." *Policing* 4(4):403-10. doi: 10.1093/police/paq041.

Hertogh, Marc. 2015. "What Moves Joe Driver? How Perceptions of Legitimacy Shape Regulatory Compliance among Dutch Traffic Offenders." *International Journal of Law, Crime and Justice* 43:214-34.

Hough, M., J. Jackson, and B. Bradford. 2013. 'Legitimacy, Trust and Compliance: An Empirical Test of Procedural Justice Theory Using the European Social Survey." in *Legitimacy and Criminal Justice: An International Exploration*, edited by J. Tankebe and A. Liebling. Oxford, UK: Oxford University Press.

Huq, Aziz Z., Tom R. Tyler, and Stephen J. Schulhofer. 2011. "Why Does the Public Cooperate with Law Enforcement? The Influence of the Purposes and Targets of Policing." *Psychology, Public Policy, and Law* 17(3):419-50. doi: 10.1037/a0023367.

Jackson, J., B. Bradford, M. Hough, and K. H. Murray. 2012. "Compliance with the Law and Policing by Consent: Notes on Police and Legal Legitimacy." Pp. 29-49 in *Legitimacy and Compliance in Criminal Justice*, edited by A. Crawford and A. Hucklesby. Abingdon Thames: Routledge.

Jackson, J., B. Bradford, M. Hough, A. Myhill, P. Quinton, and Tom R. Tyler. 2012. "Why Do People Comply with the Law? Legitimacy and the Influence of Legal Institutions." *British Journal of Criminology* 52(6):1051-71. doi: 10.1093/bjc/azs032.

Jonathan-Zamir, Tal, Stephen D. Mastrofski, and Shomron Moyal. 2015. "Measuring Procedural Justice in Police-Citizen Encounters." *Justice Quarterly* 32(5):845-71. doi: 10.1080/07418825.2013.845677.

Lind, E. Allan and Tom R. Tyler. 1988. *The Social Psychology of Procedural Justice*. Berlin: Springer.

MacCallum, Robert C., Shaobo Zhang, Kristopher J. Preacher, and Derek D. Rucker. 2002. "On the Practice of Dichotomization of Quantitative Variables.' *Psychological Methods* 7(1):19-40. doi: 10.1037/1082-989X.7.1.19.

MacQueen, Sarah and Ben Bradford. 2015. "Enhancing Public Trust and Police Legitimacy during Road Traffic Encounters: Results from a Randomised Controlled Trial in Scotland." *Journal of Experimental Criminology* 11(3): 419-43. doi: 10.1007/s11292-015-9240-0.

Mazerolle, Lorraine, Emma Antrobus, Sarah Bennett, and Tom R. Tyler. 2013. "Shaping Citizen Perceptions of Police Legitimacy: A Randomized Field Trial of Procedural Justice: Shaping Citizen Perceptions of Police." *Criminology* 51(1):33-63. doi: 10.1111/j.1745-9125.2012.00289.x.

Mazerolle, Lorraine, Sarah Bennett, Emma Antrobus, and Elizabeth Eggins. 2012. "Procedural Justice, Routine Encounters and Citizen Perceptions of Police: Main Findings from the Queensland Community Engagement Trial (QCET)." *Journal of Experimental Criminology* 8(4):343-67. doi: 10.1007/s11292-012-9160-1.

Mazerolle, Lorraine, Elise Sargeant, Adrian Cherney, Sarah Bennett, Kristina Murphy, Emma Antrobus, and Peter Martin. 2014. *Procedural Justice and Legitimacy in Policing*. Cham: Springer International Publishing.

McCluskey, John David. 2003. *Police Requests for Compliance: Coercive and Procedurally Just Tactics*. New York: LFB Scholarly Pub.

Murphy, Kristina. 2005. "Regulating More Effectively: The Relationship between Procedural Justice, Legitimacy, and Tax Non-Compliance." *Journal of Law and Society* 32(4):562-89. doi: 10.1111/j.1467-6478.2005.00338.x.

Murphy, Kristina, L. Hinds, and J. Fleming. 2008. "Encouraging Public Cooperation and Support for Police." *Policing and Society* 18(2):136-55. doi: 10.1080/10439460802008660.

Murphy, Kristina, Lorraine Mazerolle, and Sarah Bennett. 2014. "Promoting Trust in Police: Findings from a Randomised Experimental Field Trial of Procedural Justice Policing." *Policing and Society* 24(4):405-24. doi: 10.1080/10439463.2013.862246.

Muthén, Bengt and David Kaplan. 1985. "A Comparison of Some Methodologies for the Factor Analysis of Non-Normal Likert Variables." *British Journal of Mathematical and Statistical Psychology* 38(2):171-89. doi: 10.1111/j.2044-8317.1985.tb00832.x.

Nagin, Daniel S. and Cody W. Telep. 2017. "Procedural Justice and Legal Compliance." *Annual Review of Law and Social Science* 13(1):5-28. doi: 10.1146/annurev-lawsocsci-110316-113310.

Nunnally, Jum and I. H. Bernstein. 1994. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill.

Reisig, M. D., J. Bratton, and M. G. Gertz. 2007. "The Construct Validity and Refinement of Process-Based Policing Measures." *Criminal Justice and Behavior* 34(8):1005-28. doi: 10.1177/0093854807301275.

Reisig, M. D. and C. Lloyd. 2009. "Procedural Justice, Police Legitimacy, and Helping the Police Fight Crime: Results from a Survey of Jamaican Adolescents." *Police Quarterly* 12(1):42-62. doi: 10.1177/1098611108327311.

Sahin, Nusret, Anthony A. Braga, Robert Apel, and Rod K. Brunson. 2017. "The Impact of Procedurally-Just Policing on Citizen Perceptions of Police during Traffic Stops: The Adana Randomized Controlled Trial." *Journal of Quantitative Criminology* 33(4):701-26. doi: 10.1007/s10940-016-9308-7.

Sargeant, Elise, Emma Antrobus, Kristina Murphy, Sarah Bennett, and Lorraine Mazerolle. 2016. "Social Identity and Procedural Justice in Police Encounters with the Public: Results from a Randomised Controlled Trial." *Policing and Society* 26(7):789-803. doi: 10.1080/10439463.2014.989159.

Scheffer, Judi. 2002. "Dealing with Missing Data." Pp. 153-60 in *Research Letters in the Information and Mathematical Sciences*. New Zealand: Massey University.

Schulhofer, Stephen J., Tom R. Tyler, and Aziz Z. Huq. 2011. "American Policing at a Crossroads: Unsustainable Policies and the Procedural Justice Alternative." *The Journal of Criminal Law and Criminology* 101(2):335-374.

Skogan, Wesley G. 2006. "Asymmetry in the Impact of Encounters with Police." *Policing and Society* 16(2):99-126. doi: 10.1080/10439460600662098.

Smith, Heather J., Tom R. Tyler, Yuen J. Huo, Daniel J. Ortiz, and E. Allan Lind. 1998. "The Self-Relevant Implications of the Group-Value Model: Group Membership, Self-Worth, and Treatment Quality." *Journal of Experimental Social Psychology* 34(5):470-93. doi: 10.1006/jesp.1998.1360.

Sunshine, Jason and Tom R. Tyler. 2003. "The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing." *Law & Society Review* 37(3):513-48. doi: 10.1111/1540-5893.3703002.

Tankebe, Justice. 2009a. "Policing, Procedural Fairness and Public Behaviour: A Review and Critique." *International Journal of Police Science & Management* 11(1):8-19. doi: 10.1350/ijps.2009.11.1.105.

Tankebe, Justice. 2009b. "Public Cooperation with the Police in Ghana: Does Procedural Fairness Matter?' *Criminology* 47(4):1265-93.

Thibaut, J. and L. Walker. 1975. *Procedural Justice*. Hillsdale, NJ.: Lawrence Erlbaum.

Tyler, Tom. 2008. 'Psychology and Institutional Design'. *Review of Law & Economics* 4(3):801-87. doi: 10.2202/1555-5879.1233.

Tyler, Tom R. 1990. *Why People Obey the Law*. Princeton, NJ: Princeton University Press.

Tyler, Tom R. 2003. "Procedural Justice, Legitimacy, and the Effective Rule of Law." *Crime and Justice: A Review of Research* 30:283.

Tyler, Tom R. 2004. "Enhancing Police Legitimacy." *The ANNALS of the American Academy of Political and Social Science* 593(1):84-99. doi: 10.1177/0002716203262627.

Tyler, Tom R. and E. Allan Lind. 1992. "A Relational Model of Authority in Groups." Pp. 115-91 in *Advances in Experimental Social Psychology, Vol. 25*, edited by J. M. Olson and M. P. Zanna. San Diego, CA: Academic Press.

Tyler, Tom R. and Jeffrey Fagan. 2006. "Legitimacy and Cooperation: Why Do People Help the Police Fight Crime in Their Communities?" *SSRN Electronic Journal*. doi: 10.2139/ssrn.887737.

Tyler, Tom R. and Yuen Huo. 2002. *Trust in the Law: Encouraging Public Cooperation with the Police and Courts*. New York: Russell Sage Foundation.

Tyler, Tom R. and Cheryl J. Wakslak. 2004. "Profiling and Police Legitimacy: Procedural Justice, Attributions of Motive, and Acceptance of Police Authority." *Criminology* 42(2):253-82. doi: 10.1111/j.1745-9125.2004.tb00520.x.

Velicer, Wayne F. and Douglas N. Jackson. 1990. "Component Analysis versus Common Factor Analysis: Some Further Observations." *Multivariate Behavioral Research* 25(1):97-114. doi: 10.1207/s15327906mbr2501_12.

Willits, Dale, David Makin, and Wendy Koslicki. 2019. *Procedural Justice and Traffic Stops: Analyzing Police and Community Perspectives through Survey and Recorded Observational Research. 2018-AG-1906*. Washington, DC: Department of Criminal Justice and Criminology, Washington State University.

Winter, Søren C. and Peter J. May. 2001. "Motivation for Compliance with Environmental Regulations." *Journal of Policy Analysis and Management* 20(4):675-98. doi: 10.1002/pam.1023.

Worden, Robert E. and Sarah J. McLean. 2017. *Mirage of Police Reform: Procedural Justice and Police Legitimacy*. Berkeley: University of California Press.

Author Biographies

Bo Terpstra is a lecturer at Leiden Law School with a special interest in behavioral economics. His current research focuses on procedural justice, legitimacy and the probability of apprehension.

Peter van Wijck is an associate professor of law and economics at Leiden Law School. His research is focused on the economic analysis of criminal justice, dispute resolution, tort and competition law.