Universiteit Leiden
The Netherlands

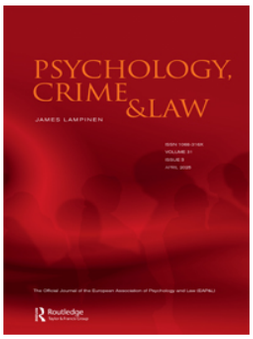
Mapping the theoretical pathways from police contact to criminal behavior: a scoping review

Op de Weegh, L.A.G.; Nivette, A.; Nägel, C.R.E.; Ruiter, S.

Citation
Op de Weegh, L. A. G., Nivette, A., Nägel, C. R. E., & Ruiter, S. (2025). Mapping the theoretical pathways from police contact to criminal behavior: a scoping review. *Psychology, Crime &Amp; Law*, 1-30. doi:10.1080/1068316X.2025.2474535

Version: Publisher's Version
License: Creative Commons CC BY 4.0 license
Downloaded from: https://hdl.handle.net/1887/4308346

Note: To cite this publication please use the final published version (if applicable).



Psychology, Crime & Law

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gpcl20

Mapping the theoretical pathways from police contact to criminal behavior: a scoping review

Amina op de Weegh, Amy Nivette, Christof Nägel & Stijn Ruiter

To cite this article: Amina op de Weegh, Amy Nivette, Christof Nägel & Stijn Ruiter (09 Mar 2025): Mapping the theoretical pathways from police contact to criminal behavior: a scoping review, Psychology, Crime & Law, DOI: 10.1080/1068316X.2025.2474535

To link to this article: https://doi.org/10.1080/1068316X.2025.2474535

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 09 Mar 2025.

Submit your article to this journal

Article views: 314

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gpcl20

PSYCHOLOGY, CRIME & LAW
https://doi.org/10.1080/1068316X.2025.2474535

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Mapping the theoretical pathways from police contact to criminal behavior: a scoping review

Amina op de Weegh [a], Amy Nivette [a], Christof Nägel [b] and Stijn Ruiter [a,c]

[a]Department of Sociology, Utrecht University, Utrecht, The Netherlands; [b]Institute of Sociology and Social Psychology, University of Cologne, Cologne, Germany; [c]Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, The Netherlands

ABSTRACT

Many studies aim to evaluate and explain how and when police contact leads to offending. However, research on this topic is diffuse, drawing from different theoretical frameworks, which has consequently produced many possible mechanisms explaining this association. The current study therefore aims to provide a starting point for organizing the literature on police contact and criminal behavior in order to highlight potential points for theoretical competition, integration, and testing. Specifically, we conducted a systematic scoping review ($N = 95$ studies) of all quantitative studies that have tested this relationship and assessed the theoretical frameworks used, measurement of key concepts, mediators and moderators, as well as methodological quality. Our results found that studies were primarily based on four major theoretical frameworks: procedural justice theory, strain theory, labeling theory and deterrence theory, with little integration or competition between them. The broad variety in measurements and the relatively small number of studies that were able to adequately address threats to internal validity demonstrated the difficulty in drawing conclusions about the strength of a given theory and where the weight of the evidence lies. We further discuss the need for the development of clear and falsifiable hypotheses, standardization of measurement, and attention to causal inference designs.

ARTICLE HISTORY
Received 27 November 2023
Accepted 22 February 2025

KEYWORDS
Criminology; police contact; criminal behavior; compliance

Introduction

A wide range of research has shown that contact with the police is associated with criminal behavior (Beardslee et al., 2019; Chenane et al., 2021; Kim & Lee, 2019; Liberman et al., 2014; Petersen et al., 2023; Slocum et al., 2016). Contact has been operationalized in a number of ways, including arrests, police-initiated contact, and evaluations of encounters with the police. For example, arrests have been linked with subsequent offending and

CONTACT A. Op de Weegh l.a.g.opdeweegh@uu.nl Department of Sociology, Utrecht University, Padualaan 14, Utrecht 3584 CH, The Netherlands

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1068316X.2025.2474535

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

substance use (e.g. Beardslee et al., 2019; Mitchell, 2016), and unfair encounters with police have been associated with lower compliance with the law (e.g. Slocum et al., 2016). While these studies vary in their theoretical and methodological approaches, at the core they generally aim to evaluate and explain how and when police contact can lead to offending. Typically, this focus is on 'negative' direct or vicarious contact with police, which can be broadly described as an interaction with the police that is characterized by negative consequences (e.g. arrest) and/or negative treatment (e.g. unfairness) for the individual. Understanding this relationship can be considered key to developing effective interventions to reduce the likelihood that individuals may engage in criminal behavior, as well as reduce recidivism among offenders (Mazerolle et al., 2021; Schulz, 2014).

However, research on this topic is diffuse, drawing from different theoretical frameworks, which has consequently produced a number of possible mechanisms explaining this association. These can include for example deterrence, labeling, strain, and procedural justice theories, some of which propose competing mechanisms. Based on deterrence theory, for example, one would expect that police contact results in an updating of potential risks and benefits of offending, which subsequently can reduce the likelihood of engaging in criminal behavior (Paternoster, 2010). Based on labeling theory, one would expect those who encounter the police may be labeled as deviant, which in turn may increase feelings of exclusion and contact with deviant peers and ultimately offending (Wiley et al., 2013). These studies do not always speak to each other, even holding different expectations about whether the main effect of police contact on criminal behavior is positive or negative. Furthermore, most evidence tends to be correlational, while causal evidence is weak or mixed (Nagin & Telep, 2020; Paternoster, 2010; Petersen et al., 2023; Walters & Bolger, 2019), and explains only a small proportion of variance (Weisburd & Piquero, 2008).

These issues are part of a broader problem facing criminological theory and research more generally (Agnew, 2011; Bernard & Snipes, 1996; Niemeyer et al., 2022; Proctor & Niemeyer, 2019). Specifically, critics argue that the field is plagued by theoretical fragmentation resulting in an 'overabundance of competing, contradicting, and empirically weak theories of crime' (Proctor & Niemeyer, 2019, p. 2). These different potential explanations make the task of falsification and knowledge consolidation especially difficult (Dooley & Goodison, 2020). One potential solution to this issue is to systematically assess possible theories and identify their mechanisms, conditions, and evidence to construct comparable schema that can then be integrated, falsified, or framed as competing (Proctor & Niemeyer, 2019).

The current study therefore aims to provide a starting point for organizing the diffuse literature on police contact and criminal behavior in order to highlight potential points for theoretical competition, integration, and testing. Specifically, we conduct a systematic scoping review of quantitative empirical research that has examined the relationship between contact with police and criminal behavior. A scoping review is a valuable evidence synthesis approach to identifying knowledge gaps, scope a body of literature, clarify concepts, or investigate research conduct (Munn et al., 2018). As we aim to conduct a broad and open-minded literature search, we have provided only brief examples of the content of specific criminological theories in our introduction. Only after the search, we will discuss in more depth the content of the theoretical perspectives that emerged from our search.

The current paper

The current review was conducted according to the PRISM-ScR checklist (Colquhoun et al., 2014) and best practice recommendations (Arksey & O'Malley, 2005), both of which aim to ensure the methodological rigor of scoping reviews. The purpose of our scoping review is to provide an overview of quantitative research that has examined the relationship between police contact and criminal outcomes. Specifically, we aim to:

(a) Investigate the range of criminal outcome variables to which police contact has empirically been related in previous research.
(b) Provide an overview of theoretically proposed mechanisms linking police contact with criminal outcome variables.
(c) Provide an overview of the variables used to assess those relationships (e.g. mediators, moderators, control variables).
(d) Critically assess the methodological and analytical approaches used to assess causal effects within the included studies.

Our goals are primarily theoretical and methodological. That is, we aim to create an overview of how studies have specified and tested the relationship between police contact and criminal behavior. Notably, we do not systematically collect information on effect sizes or findings of the studies in our sample; however, we do provide a more general overview of findings for relevant theoretical frameworks in the discussion.

Methods

Protocol and registration

To increase transparency and replicability of our results, we pre-registered a protocol through the Open Science Framework (OSF). The protocol is public and can be accessed through https://osf.io/xq92e/.

Eligibility criteria

Studies considered for review had to meet the following criteria:

(a) The study's main focus was on how police contact relates to criminal outcome variables. Studies that predict criminal outcomes, without including any type of police contact[1], were excluded. Studies that examined police contact in relation to non-criminal outcomes[2] were excluded as well.
(b) The study quantitatively tested how police contact relates to criminal outcome variables. Theoretical and review studies (e.g. systematic reviews or meta-analyses) were excluded.
(c) The study focused on individual-level interactions with the police. All studies using solely macro-level variables were excluded[3].
(d) The study was informative for answering at least one of the research questions.

The purpose of a scoping review is to identify all relevant literature (Arksey & O'Malley, 2005). As we aimed to provide a broad overview, we considered both academic and gray literature, and any output (e.g. book chapter, journal article, technical report) published in English. The study considered all years and months up to and including October 2021, as the search was conducted in November 2021. There were no geographical limitations.

Information sources

In collaboration with a librarian and drawing from the authors' knowledge of and expertise in (criminological) research, a search strategy was set-up. The primary electronic databases used were Sociological abstracts, Clarivate Analytics/Web of Science Core Collection, Scopus, Proquest/IBSS, and Ebsco/Criminal Justice Abstracts; APA PsycArticles; APA PsycInfo; Psychology and Behavioral Sciences Collection. We also consulted several secondary electronic databases, i.e. Google Scholar, Global Policing Database, and OpenGrey.

The search using primary databases was supplemented by articles searched by hand from the secondary databases. Additionally, the primary search was supplemented by forward and backward snowballing. Forward snowballing involves identifying potentially relevant articles that have cited the article found in the search. Backward snowballing involves scanning the reference lists of articles found in the search to identify potentially relevant articles. We also searched the secondary databases for articles published by authors that are prevalent in the literature (i.e. authors that published more than two relevant articles). Newly identified documents through either hand-searching, forward or backward snowballing, or specific author searches, were flagged for full-text screening.

Search

The following terms were used as index terms or free-text words:

- Police contact(s), police encounter(s), police confrontation(s), police interaction(s), police stops, arrests, stop-and-frisk, police-citizen interaction(s), police-citizen contact(s), police legitimacy
- Criminal behavio(u)r, deviant behavio(u)r, law-abiding behavio(u)r, law-violating behavio(u)r, delinquent behavio(u)r, illegal behavio(u)r, obey or obedience in relation to law, compliance in relation to law, delinquency, crime, violence, recidivism, homicide, offending, violation in relation to law, offenders, offenses, unlawfulness, terrorism, violent extremism

These terms were adjusted to create specific search queries for each database, as databases differ in operator specifics. Online Appendix A provides an overview of the exact search queries used per database, and the number of studies these searches yielded.

Selection and appraisal of sources of evidence

First stage screenings

The primary search yielded 38,156 results (see Figure 1). The documents retrieved from the search were imported into Zotero in order to easily remove duplicates. In total,

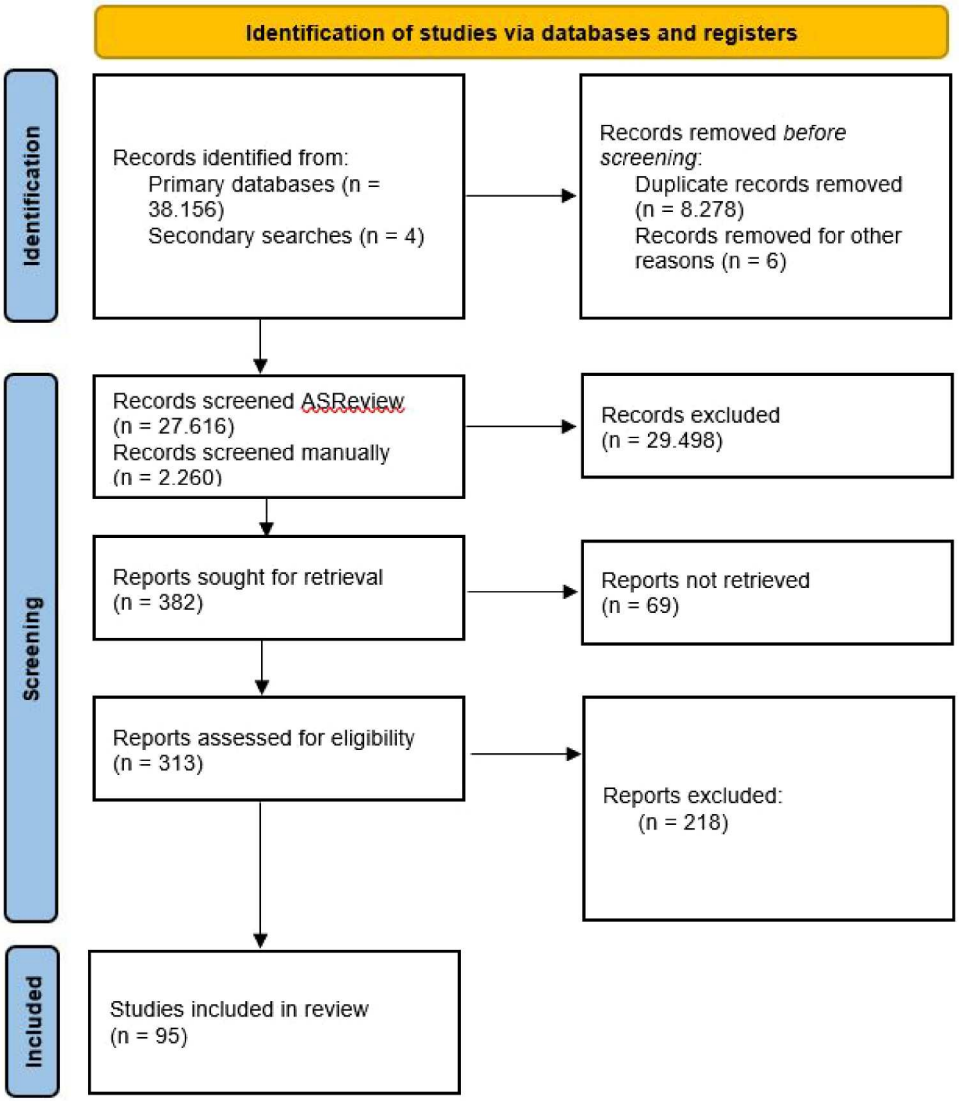


Figure 1. Flowchart.

8278 duplicates and 6 retracted articles were removed. The remaining documents were subsequently imported into ASReview for abstract screening. Because ASReview requires a title and abstract, documents that did not match these criteria were manually reviewed ($N = 2256$). ASReview is an active learning method, which uses the title and abstract of documents to present users with the most relevant literature in an ordered sequence (Van de Schoot et al., 2021). ASReview offers a much quicker way to select relevant literature and adds to transparency and reproducibility since each decision of the reviewer will be automatically logged in a log-file. Initially, the algorithm is informed by a small number of user-selected relevant and irrelevant documents, after which it is constantly updated based on the users' subsequent screening decisions.

All titles and abstracts were screened independently in ASReview by the first and second authors. Following Van de Schoot and colleagues (2021), both coders stopped screening after marking one hundred studies in a row as irrelevant in ASReview. The same eligibility criteria were used as in the primary search. Studies identified as relevant by either one of the authors proceeded for full-text review by the first author. Tests for inter-rater reliability were conducted and showed moderate inter-rater reliability (*Cohen's* $\kappa = .526$) (Langton, 2023; McHugh, 2012). Because of the moderate inter-rater reliability, every document marked relevant by either the first or second author was included in the second-stage screenings.

Second stage screenings

After completing the title – and abstract screening in ASReview, all documents marked as relevant either by the first or second author were exported into an Excel sheet. After removing duplicates, 378 unique documents remained for full-text screening by the first author. Thirty-nine documents could not be accessed, even when trying with different university libraries. When contact details of the authors could be found, the authors were emailed to ask for the full text. If the contact details could not be found, the author did not respond, or the author could not share the document, the document was excluded from the scoping review ($N = 30$). The secondary searched yielded four additional relevant documents. A random subsample (5%, $N = 18$) was independently reviewed by the second author. Since the first and second authors disagreed about one document, this document was discussed until a consensus on inclusion was reached. In total, 95 documents were marked as relevant to include in the scoping review.

Coding

After full-text screening, we moved on to in-depth coding and analysis of all remaining documents ($N = 95$). The final literature selection was imported into Nvivo12[4] for systematic coding by the first author. Initially, the documents were coded in a general way to highlight the information mentioned about the following topics: (1) General study details (type of data, year, and context of data collection), (2) Research aim(s), (i.e. the research questions and aims stated by the authors, such as extending theories to a new outcome or improving the methodology of prior research), (3) Proposed theoretical mechanisms, (4) Sample size and type of sampling strategy (i.e. probability or non-probability samples[5], or not enough information provided), (5) Population characteristics (i.e. oversampling based on gender, or targeting drug-users or youth specifically), (6) Measurement of police contact, criminal outcome(s), moderator(s), mediator(s), control variables, and (7) Statistical analysis (type of analysis and analysis strategy). We have used the Maryland Scientific Methods Scale (Maryland SMS) (Farrington et al., 2003) to assess the methodological quality of the studies included in our sample. Information on statistical analyses, such as the type of analysis that was conducted and the addition of control variables, was used to evaluate the design and quality of studies in relation to key elements of the SMS. The SMS largely focuses on threats to internal validity, and so analytical strategies related to for example the use of techniques to adjust for differences between the treated and untreated group, were extracted and evaluated.

The Maryland Scientific Methods Scale consists of five levels (Farrington et al., 2003). The first level is a correlation study without controlling for other variables. At this level, studies simply observe a correlation between an intervention and an outcome without attempting to control for other factors that might affect the results. At the second level, there is a temporal sequence without a control group. These studies establish a timeline where the intervention precedes the outcome. However, without a control group, it is difficult to directly attribute the changes to the intervention. At the third level, there is a control group included, but no pre-intervention measures. Studies at this level have both an experimental group and a control group. However, they do not measure outcomes before the intervention, making it hard to account for pre-existing differences between the groups. At the fourth level, studies have included a control group and measure outcomes both before and after the intervention for both groups. This design helps account for pre-existing differences and provides stronger evidence of causation. Level five studies are Randomized control trials (RCT). These studies randomly assign participants to either the experimental or control group. Randomization helps ensure the groups are equivalent at the start, allowing for the most robust conclusions about the intervention's causal impact.

In the second round of coding, these general codes were subsequently divided into meaningful subcodes. Both the coding scheme for the first and second rounds of coding were discussed and agreed upon by all authors (the full coding scheme is available at https://osf.io/xq92e/).

Results

Description of studies

Table 1 provides an overview of all 95 included studies, of which there were 85 journal articles, 7 dissertations, 2 books and 1 book chapter. All documents were published between 1974 and 2021. Data used in the studies were collected between 1945 and 2020. Together there were 103 tests of theories performed within these studies.[6] The studies were predominantly conducted in the American context ($N = 63$), while 7 studies were conducted in Australia, 10 in the United Kingdom, and 3 in Canada. The remaining documents were studies conducted in other countries around the world, including The Netherlands, Brazil, Norway, or Slovenia. About a third of the studies in the sample were based on cross-sectional data ($N = 32$), and another 31 studies relied on longitudinal panel data. The remainder of the studies used cohort ($N = 13$), experimental ($N = 9$), record-linkage ($N = 7$) or other longitudinal data, such as data from institutional files or pooled data from multiple sources.

We assessed the sampling strategy used in the included studies (see Table 1). Many studies ($N = 60$) used a random sample, namely a cluster- or stratified sampling strategy. About a third of the included studies ($N = 27$) used non-probability samples, mostly convenience or purposive sampling. For 12 studies in our sample, there was not enough information provided to determine which sampling strategy was used, or the information about sampling procedures was missing from the paper. It is noteworthy that many studies used a sample of adolescents or young adults up to age 30 ($N = 62$). Additionally, a noticeable number of studies were performed on a sample of offenders or participants

Table 1. Characteristics of studies included in the systematic review organized by four major theoretical frameworks.

	Total *N* (%)	Procedural Justice *N* (%)	Deterrence *N* (%)	Strain *N* (%)	Labeling *N* (%)
N tests of theories	103 (100)	42	22	7	32
Country of data collection					
USA	76 (79.2)	19 (45.2)	16 (72.7)	4 (57.1)	29 (90.6)
UK	13 (13.5)	8 (19.0)	3 (13.6)	*n.a.*	1 (3.1)
Australia	8 (8.3)	6 (14.3)	1 (4.5)	1 (14.3)	*n.a.*
Other	15 (15.6)	9 (21.4)	2 (9.1)	2 (28.6)	2 (6.25)
Sampling strategy					
Random	72 (75)	18 (42.9)	16 (72.7)	4 (57.1)	29 (90.6)
Non-random	30 (31.3)	20 (47.6)	5 (22.7)	2 (28.6)	1 (3.1)
Unknown	10 (10.4)	4 (9.5)	1 (4.5)	1 (14.3)	2 (6.25)
Type of data					
Cross-sectional	37 (38.5)	24 (57.1)	3 (13.6)	3 (42.9)	3 (9.4)
Record-linkage study	10 (10.4)	1 (2.4)	3 (13.6)	*n.a.*	5 (15.6)
Cohort study	12 (12.5)	1 (2.4)	3 (13.6)	2 (28.6)	5 (15.6)
Panel study	44 (45.8)	10 (23.8)	11 (50.0)	2 (28.6)	17 (53.1)
Experiment	11 (11.5)	6 (14.3)	3 (13.6)	1 (14.3)	1 (3.1)
Other longitudinal	4 (4.2)	3 (7.1)	*n.a.*	*n.a.*	1 (3.1)
N studies ≥ level 3 Maryland SMS	27 (28.1)	7 (16.7)	7 (31.8)	2 (28.6)	10 (31.3)

Notes. Numbers in the table refer to the number of tests of theories, not the number of studies. In the table we focus on the four most prominent theories; however, there are more theoretical ideas included in the sample.

who had previously been in contact with the police ($N = 21$). About one third ($N = 36$) of the included studies were based on a (predominantly) male sample, while 10 studies included predominantly female participants. A third of the studies ($N = 30$) oversampled migrants.

A large number of studies aimed to test new mechanisms to explain the relationship between police contact and criminal outcomes. In about a fifth of the studies, new mediators were tested ($N = 18$). Barkworth and Murphy (2015) for example aimed to test the mediating role of negative emotions within the procedural justice framework. Within the labeling framework, McGlynn-Wright (2014) explored whether the relationship between adolescents contact with the police and subsequent arrest is mediated by current involvement in illegal behavior. Additionally, quite a few studies ($N = 20$) focused on contextual effects by identifying and testing new moderators. For example, one study tested which environmental and behavioral factors moderate the relationship between legal processing and orientation towards delinquency (Ageton & Elliott, 1974), while another tested whether the relationship between arrest and subsequent substance use differed by race and gender (Hassett-Walker et al., 2017).

In addition to testing mediators and moderators, studies also focused on improving the way in which these arguments were empirically assessed ($N = 26$), for example by using longitudinal data, different statistical analyses, or a more representative sample. Some studies also aimed to extend existing theories to new outcomes ($N = 13$), such as by extending the procedural justice framework to outcomes such as intergroup or private violence (e.g. Gerber et al., 2018; Jackson et al., 2013). Another prevalent aim was to test theories in a different context ($N = 19$), such as in a new country or on a different sample. Next to that, a number of studies aimed to test for (racial/ethnic) group differences in the relationship between police contact and criminal outcomes ($N = 17$).

Description of theories

The studies included in our sample mention multiple explanatory frameworks, variously referred to as theories, models, approaches, and hypotheses. Twelve studies did not include a clear theoretical or explanatory framework or examined only risk and protective factors. Some of these explanatory frameworks are nested within broader theoretical frameworks, for example, the asymmetry hypothesis, invariance thesis, coercive and consensus approaches, group engagement/value models, and 'experience the expected' hypothesis are all related or used in some way to elaborate on procedural justice theory. There are no less than 20 different theories mentioned across papers, including strain, labeling, deterrence, procedural justice, self-control, control, social learning, equity, defiance, conflict, social support, age-graded theory of informal social control, interactional theory of delinquency, self-determination, differential association, reintegrative shaming, symbolic assailant, and dual taxonomy theory. However, on closer inspection, only a small number of these theories explicitly aim to explain the relationship between police contact and offending. Specifically, six theories focused on or included an explanation as to why police contact would lead to future offending behavior: procedural justice theory ($N = 42$), labeling theory ($N = 32$), deterrence theory ($N = 22$), general strain theory ($N = 7$), defiance theory ($N = 7$), and interactional and age-graded theories ($N = 2$). We shortly elaborate on these theoretical pathways below.

Procedural justice theory proposes that the way in which police treat citizens influences how they perceive the police and subsequently, whether they obey the law (Tyler, 2006). According to this theory, when police act in a procedurally just manner, by treating people with fairness, neutrality, and respect, individuals will be more likely to view the police as legitimate powerholders and therefore voluntarily comply with the law because they believe it is the right thing to do. The effect of police contact on offending therefore depends on the quality of treatment. Negative or unjust police contact (personal or vicarious) can lead to less legitimacy and more offending.

Deterrence theory assumes that individuals weigh the costs and benefits associated with breaking the law, and that individuals are expected to refrain from offending when the costs of crime outweigh the benefits they believe can be gained by crime (Ariel et al., 2019). The threat of punishment therefore has to be real, consequential and probable, and the punishment itself has to be severe. Specific deterrence theory predicts that contact with the police (e.g. an arrest) provides information about the costs of offending, which is then expected to factor into the decision-making process for future opportunities for crime (Raaijmakers et al., 2017). Police contact signals that the costs are high, and therefore outweigh the benefits, preventing future offending.

General strain theory suggests that police contact and unjust treatment by the police are types of subjective strain (noxious or negatively valued stimuli) that cause psychological distress and negative emotions (Barkworth & Murphy, 2015). Negative emotional responses, such as increased anger and fear, require some corrective or coping action, one of which may be criminal behavior (Agnew & White, 1992). Unjust treatment by police can elicit negative emotions that are particularly conducive to crime, such as anger (Agnew, 2001). According to strain theory, the relationship between police contact and offending depends on a number of conditions including the type and degree of contact (i.e. unjust treatment), the emergence of certain negative emotions

(i.e. anger), and the individual characteristics and opportunities that influence different forms of coping.

Labeling theory starts with the basic argument that interactions with the legal justice system lead to increased offending because of changes in one's self-concept (e.g. Motz et al., 2020; Wiley et al., 2013). Individuals who come into contact with the justice system (e.g. police) are being labeled as deviant by society, which in turn influences offending via three possible mechanisms. First, the labeling process can lead to exclusion and weakening of prosocial bonds. Individuals may encounter social rejection or experience discrimination which leads them to withdraw from prosocial bonds (Wiley et al., 2013). Second, the deviant label may cause changes in one's self-concept, whereby individuals may redefine themselves and adopt a deviant identity in accordance with the label. Third, alongside feelings of exclusion and the adoption of a deviant identity, labeled individuals may become more involved with deviant peers who may accept and reinforce deviant values.

Defiance theory aims to explain why the relationship between contact with the criminal justice system (e.g. police contact such as arrest) differs between individuals (Bouffard & Piquero, 2010). Defiance refers to the 'proud, shameless reaction to the administration of a criminal sanction' (Sherman, 1993, p. 459). The primary causal mechanism is the denial of shame; however, in order for police contact to lead to offending three other conditions must be met. Borrowing from procedural justice and labelling theories, the sanctions (e.g. arrest) must be perceived as unjust and stigmatizing, and drawing from reintegrative shaming theory (Braithwaite, 1989), the individual must be poorly bonded to prosocial others (Bouffard & Piquero, 2010).

Interactional theory (Thornberry, 1987) and *age-graded theory of informal social control* (Sampson & Laub, 1990) are broader theories of criminal behavior that specify similar mechanisms linking police contact and later offending. While there are important differences in the origins of criminal behavior, both theories broadly expect that police contact (e.g. arrest) can lead to the deterioration of social bonds and control over the life course, which in turn increases the likelihood of criminal behavior. However, in both cases, it is not clear to what extent involvement in crime leading up to police contact or contact itself is the cause of weakened bonds and subsequent criminal behavior. In case of the former, the relationship between police contact and future offending would be spurious, as involvement in criminal behavior, or some third common cause, would cause both police contact and the breakdown of social bonds (e.g. Skardhamar & Savolainen, 2014).

These theories were sometimes tested against each other, or variables reflecting other theoretical mechanisms were included as control variables. For presentation, we focus our description of results on the first four theories (i.e. procedural justice, deterrence, strain, labeling), as they were the most prevalent in our sample of literature. Table 2 shows how often four of the main theories in our sample were tested together in one study.[7] Certain theories, such as procedural justice theory, specifically contrast their assumptions and mechanisms with a competing theory, namely deterrence theory. Procedural justice theory proposes that compliance with the law (i.e. avoidance of criminal behavior) is *not* due to instrumental concerns with risks and costs of crime as portrayed in deterrence theory. Instead, most people comply out of a duty or felt obligation to obey stemming from the belief that those in power, including the police, have the right to rule and enforce the law. Other theories aim to provide points for integration, such as defiance

Table 2. Number of theories tested together in the sample.

Theory	1	2	3	4
1. Procedural justice	*n.a.*			
2. Deterrence	4	*n.a.*		
3. Strain	3	1	*n.a.*	
4. Labeling	1	13	2	*n.a.*

Notes. Numbers in the table refer to the number of times that multiple theories were tested in one study.

theory, labeling theory, and interactional theories, all of which suggest that stigmatizing contacts, combined with other processes, can weaken social bonds and increase the likelihood of future criminal behavior. What is important to highlight are the differences between theories regarding their expectation for the main relationship between police contact and criminal outcomes. For example, whereas deterrence theory predicts a negative relationship between police contact and crime, labeling theory predicts a positive relationship.

Measuring police contact

Table 3 summarizes the main pathways that were tested in each theoretical framework. As our inclusion criteria specified that studies should include a measure of actual police contact, we focus on the police contact measure and do not discuss other predictor variables included in our sample. Studies testing Procedural Justice theory very often operationalized police contact by the quality of treatment by police (90%), such as perceived procedural justice or perceived unjust use of force. Studies testing Deterrence or Labeling theory, relatively often included general measures of any police contact or arrest (Deterrence 74.3%; Labeling 84.2%). Strain theory was tested predominantly using the perceived quality of treatment by police (71.4%), and to a lesser extent a general measure of police contact (42.9%).

Within the sample, most measures of police contact were about police-initiated contacts, specifically asking when participants were stopped or arrested most recently. Most studies simply referred to such variables as 'police contact', but in some studies, a clear distinction was made between police- and citizen-initiated contacts (e.g. Trinkner et al., 2018). The studies in the sample also primarily asked about negative police contacts as the result of engagement in or suspected engagement in criminal acts. In few studies it was unclear whether the contact was positive or negative and police- or citizen-initiated. For example, Barkworth and Murphy (2015) asked participants to evaluate their most recent encounter with police, but did not specify what type of police contact (i.e. police- or citizen-initiated).

The time-period over which measures of police contact were asked also varied substantially. For example, experimental vignette studies measured participants' reactions immediately after being presented with the (vicarious) police contact scenario (e.g. Gerber et al., 2018; Van Petegem et al., 2021). Other studies measured whether participants had any police contact in the past year (Pollock et al., 2016), in the past eight years (Motz et al., 2020) or ever in their life (Chenane et al., 2021). Most measures of police contact were self-reported, in which participants were asked whether they had contact with the police over a certain period of time. In fewer instances (21.1%), official measures of arrests were used (e.g. Beardslee et al., 2019; Doherty et al., 2016).

Table 3. Summary of variables used as predictors, mediators, moderators, outcomes and controls organized by the four major theoretical frameworks.

	Total *N* (103)	Procedural Justice *N* (42)	Deterrence *N* (22)	Strain *N* (7)	Labeling *N* (32)
Predictors					
Police contact	44	10	13	3	17
Vicarious police contact	6	2	1	1	2
Deterrence[a]	12	4	6	*n.a.*	1
Evaluation police treatment	54	38	3	5	*n.a.*
Arrest	21	*n.a.*	7	*n.a.*	14
Incarceration	4	*n.a.*	2	*n.a.*	2
Situational[b]	10	5	3	*n.a.*	2
Mediators					
Individual	7	*n.a.*	1	*n.a.*	6
Psychological	23	8	3	3	7
Neighborhood	2	*n.a.*	1	*n.a.*	1
Legal attitudes	32	23	1	1	4
Social support	7	1	*n.a.*	1	5
Situational	3	1	*n.a.*	*n.a.*	*n.a.*
Moderators					
Individual	25	3	5	5	9
Psychological	9	5	1	1	*n.a.*
Neighborhood	7	1	2	*n.a.*	4
Legal attitudes	9	3	2	1	3
Social support	13	3	2	2	5
Situational	1	*n.a.*	1	*n.a.*	*n.a.*
Outcomes					
Offending	52	13	11	2	22
Vandalism	1	1	*n.a.*	*n.a.*	*n.a.*
Arrest	8	1	3	*n.a.*	3
Compliance	22	17	2	1	*n.a.*
Defiance towards police	2	2	*n.a.*	*n.a.*	*n.a.*
Substance use	11	3	2	2	3
Traffic offending	3	2	1	*n.a.*	*n.a.*
Dull compulsion	1	1	*n.a.*	*n.a.*	*n.a.*
(Support for) use of violence	20	7	4	2	4
Gang-membership	4	*n.a.*	2	*n.a.*	2
Anti-social behavior	4	2	*n.a.*	1	1
Controls					
Individual	103	37	19	7	31
Psychological	51	12	11	4	20
Neighborhood	32	6	4	3	14
Legal attitudes	26	11	4	1	6
Social support	58	12	10	4	25
Situational	13	4	3	1	4

Notes: Numbers in the table refer to the number of tests of theories, not the number of studies. In the table we focus on the four most prominent theories; however, there are more theoretical ideas included in the sample.
[a]These involve measurements regarding perceptions of getting caught or severity of punishment.
[b]These involve measurements asking about a specific interaction with the police and/or circumstances surrounding that interaction.

Pathways to criminal behavior: mediating mechanisms

Assessing the pathways connecting police contact to criminal outcome variables, about half of the studies (54.8%) that tested procedural justice theory included legal attitudes as a mediator (e.g. legitimacy beliefs, trust in police, legal cynicism) (e.g. Bello & Matshaba, 2020; Bradford et al., 2015; Jackson et al., 2012). This number is surprisingly low given that the basis of procedural justice theory is that procedural justice influences compliance through changes in legitimacy beliefs (Tyler, 2006). Additionally, Bello and Matshaba (2020) included dull compulsion as a mediator, arguing that citizens voluntarily comply

with the law out of a sense of powerlessness and fear rather than perceived legitimacy. Relatedly, Cavanagh et al. (2022) investigated whether the relationship between perceptions of police treatment and offending was mediated by legal cynicism. They argued that increased legal cynicism, potentially caused by feelings of procedural injustice, could make citizens feel justified in breaking the law. Both Hough et al. (2013) and Jackson et al. (2012) added perceived risk of sanctions as an additional mediating mechanism to the procedural justice model. While this may have been done in order to test procedural justice theory alongside deterrence theory, their theoretical justification was not clearly specified.

Psychological factors (e.g. negative affect, self-esteem) were most prevalent in studies testing strain (42.9%), labeling (21.9%), or procedural justice (19%) theory. For example, drawing from strain theory, both Barkworth and Murphy (2015) and Baron (2019) included negative emotions as a mediator between procedural justice and compliance behavior. Next to strain theory, Barkworth and Murphy (2015) also based the inclusion of negative affect as a mediator on equity theory, reintegrative shaming theory and defiance theory. Their argument was that procedurally unjust behavior is a form of strain, which might in turn cause negative emotional responses, such as anger or depression. Such negative emotional responses are considered conducive to crime, as crime facilitates a relief of such emotions and strain, and allows for revenge against the sources of strain (Agnew, 2001). Similarly, Del Toro et al. (2019) used strain theory to derive the hypothesis that psychological distress mediates the relationship between police stops and delinquency.

Another psychological factor that was often included as a mediator was identification. For example, Bradford et al. (2015) drew from the group-engagement model to argue that procedural fairness can strengthen identification with the group the police represent, which, in turn, can motivate the legitimation of group authorities (i.e. the police). When people identify more strongly with a group, they are motivated to support and legitimize its authorities, and to cooperate with authorities and on their behalf. In addition, greater commitment to the group might lead to greater aversion to behaviors that might endanger other group members (pp. 177–178). Similarly, Jackson (2015) included social identification in the 'classical' procedural justice model as a mediator. Based on Tyler's (2006) research, he expected that procedural justice strengthens one's identification with the role of 'good and law-abiding citizen', which then can motivate people to comply with the law.

Likewise, a number of studies testing labeling theory assessed psychological factors as mediators explaining the link between contact and criminal behavior. Van Gelder et al. (2020) drew from specific deterrence, self-control and labeling theories to test whether police (and school) sanctioning reduced self-control, and subsequently increased future offending. The authors argued that sanctions increase the probability of future crime because they reduce self-control, in direct and indirect ways, which can encourage short-term mindsets at the expense of considering the future. In a direct way, sanctions may lead offenders to internalize the label of offender, who are characterized by low self-control. Therefore, offenders will adapt to this label and act in ways that are impulsive and risk-seeking. Indirectly, the delinquency label that sticks to offenders may inhibit successful engagement in future-oriented activities, such as employment or education, which are important factors in desistance from crime (Van Gelder et al., 2020).

Measures of social support (i.e. parental or peer support) were most often included as mediators in studies testing labeling theory (15.6%). This percentage was lower than

expected, since labeling theory argues that police contact affects delinquency via three mechanisms: the development of a deviant identity, social exclusion and attenuation of prosocial bonds, and involvement with delinquent groups. These three mechanisms were included as mediators in a few studies in our sample (Wiley, 2015; Wiley et al., 2013). Whereas other studies only partly measure these mechanisms. For example, Kim and Lee (2019) examined to what extent deviant self-identity mediated the relationship between police contact and crime. Additionally, they tested whether educational characteristics (e.g. school commitment, educational aspiration) mediated the relationship, arguing that arrest can lead to time lost in school and a negative label among teachers, therefore increasing drop-out. Buchanan and Krohn (2020) further elaborated on potential mediating mechanisms within the labeling process, as they argued that one's self-concept consists of two components: self-esteem and self-identity. Whereas labeling theory specifies the role of changing self-identity as a mechanism, self-esteem may also be affected by deviant social labeling. In other words, police contact can lead to more negative reflected appraisals, which in turn lowers self-esteem.

Conditional effects: moderators

Among others, moderating variables included measures of social support (e.g. peer and parental relationships), legal attitudes (e.g. legal cynicism, legitimacy), and neighborhood characteristics (e.g. structural disadvantage, share of minorities, number of criminal activities). In studies testing procedural justice theory, a wide range of psychological moderators were tested (11.9%). This included for example self-control (Baron, 2019; Baron & Macdonald, 2020), affect intensity (Murphy, 2009), callous-unemotional traits (Walsh et al., 2019), and racial identification. For example, Gerber et al. (2018) tested whether the effect of perceived procedural justice on police legitimacy and support for both police and indigenous violence was stronger among indigenous peoples who strongly identified with their group. The authors showed that for those who identify strongly with their indigenous group, perceiving high procedural justice predicts greater police legitimacy, greater support for police violence, and lesser support for violence perpetrated by indigenous activists. Similarly, McLean (2017) tested whether ethnic identification moderated the impact of procedural injustice on offending.

For other theoretical frameworks, individual characteristics such as age, gender, race, or criminal history were tested most often (deterrence: 22.7%; strain: 71.4%; labeling: 28.1%). For example, Felson et al. (2005) tested whether reporting and arrest affected whether intimate partner assault was repeated, and whether the deterrent effect depended on individual characteristics, such as whether the offender had a prior history of violence against the victim, the type of assault, the offender's race, marital status, and socioeconomic status. Dennison and Finkeldey (2021) investigated whether the relationship between unfair police treatment, as a measure of strain, and offending behavior (e.g. drug-use), varied by race and ethnicity. Drawing from labeling theory, Augustyn et al. (2019) tested whether the effect of involuntary contact with the police on subsequent offending behavior varies across mother's and father's arrest histories. Chenane (2018) examined to what extent the effect of prior police contact on delinquency/violence was conditional on neighborhood context.

Measuring criminal behavior

Across theoretical frameworks, the outcome variables included were quite similar. These involved measures of offending or compliance with the law, substance use and (support for) the use of violence. Some examples of specific forms of criminal behavior that were used as outcome variables include self-reported speeding (Bates et al., 2016), re-offending (Felson et al., 2005), non-compliance regarding tax returns (Murphy, 2009), and committing violent crimes (Ward et al., 2014).

Within the procedural justice framework, offending (31%) and compliance (40.5%) were most often used as outcome variables. Measures of offending and compliance were quite similar; however, they are formulated differently: measures of offending asked how often participants violated the law, whereas compliance measured the extent to which respondents had *never* committed offenses or never intended to violate the law. The deterrence (50%) and labeling (68.8%) studies tended to use measures of offending more often compared to other theoretical frameworks. Studies testing strain theory most often included a measure of offending (28.6%), substance use (28.6%) or (support for) the use of violence (28.6%). However, studies that tested procedural justice (16.7%), deterrence (18.2%), and labeling theories (12.5%) also measured (support for) the use of violence as an outcome.

Similar to police contact, most criminal outcome measures were self-reported. There was considerable variation in the amount of time that has passed since the police contact. For example, Augustyn et al. (2019) measured offending by asking whether and how often participants offended since the last interview. They interviewed participants every year. Trinkner et al. (2018) measured compliance with the law by asking participants how often they had engaged in different illegal behaviors in the 5 years prior to their data collection. Barkworth and Murphy (2015) assessed respondents' willingness to obey the police and propensity for future compliance immediately after participants read the police contact vignettes.

Methodological assessment

We assessed the methodological quality of the studies included in our sample using the Maryland Scientific Methods Scale (Farrington et al., 2003). The Maryland SMS aims to assess the differences in methodological quality of criminological interventions, specifically focused on internal validity. The scale ranges from level 1 (cross-sectional correlations) to level 5 (Randomized Control Trials). Figure 2 shows the distribution of the Maryland SMS ratings across the major theoretical frameworks identified in this review. The largest share of studies in our sample ($N = 73$) where classified as level 2 on the Maryland SMS. These studies were either cross-sectional with 'adequate' control variables,[8] or longitudinal without an untreated comparison group.

Table 1 shows the number of studies classified as level 3 or above for the four major theoretical frameworks identified. Relatively speaking, deterrence was most often classified as level 3 or higher (31.8%), followed by labeling (31.3%) and strain (28.6%) theories. Procedural justice (16.7%) theory was classified as level 3 or above least often. A small proportion of our sample ($N = 13$) consisted of level 3 studies. It is important to emphasize here that only less than a third of the studies of all four major theoretical

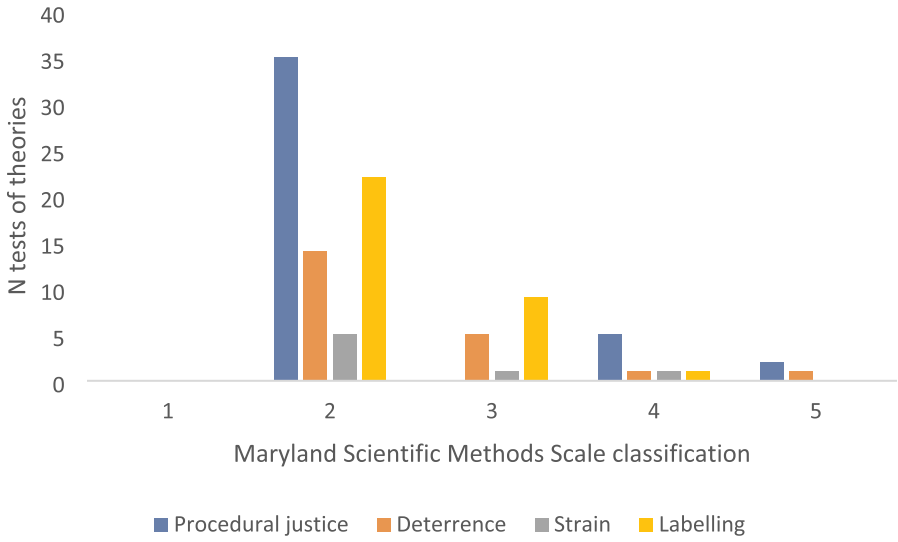


Figure 2. Distribution of Maryland Scientific Methods scale classification across theories.

frameworks could be classified as level 3 or higher. We reflect on this further in the discussion. Studies classified as level 3 were mainly longitudinal studies in which a treatment and control group were present. These studies used techniques such as propensity score matching to adjust for differences between the treated and untreated group. At this level, it is important that the groups are comparable, but it is likely that unobserved differences remain

Only a very small proportion of studies in our sample (9.5%) were classified at level 4 or level 5 on the Maryland SMS. Level 4 includes quasi-experimental studies, in which it can be credibly held that treatment and control groups differ only in their exposure to the random allocation or treatment. In our sample, we found that almost all of the six-level four studies (6.3%) were vignette-based experiments. With one exception (Sherman et al., 1992), all studies measured criminal behavior using intentions (e.g. intentions to comply in the future) or attitudes (e.g. to what extent they agreed with the use of violence by police or citizens). These vignette studies were only able to show very short-term and situational effects. Only three studies within our sample were classified as a level 5 randomized controlled trial. Unlike level 4 studies, here we found that two out of three studies tested long-term effects of the experimental manipulation (i.e. police contact). Ariel et al. (2019) examined the effects on criminal behavior effects up to 12 months post-contact, and Mazerolle et al. (2021) examined effects up to two years post-contact. Posch et al. (2021) examined short-term and situational effects, similar to most studies in level 4.

Control variables

Almost all studies ($N = 81$) included control variables, predominantly individual characteristics, such as age and education. Psychological characteristics (e.g. impulsivity, self-control) and social support factors (delinquent peers, living with their parents) were

also controlled for relatively often. Studies testing procedural justice theory most often controlled for legal attitudes (26.2%, e.g. obligation to obey police, trust in police), whereas studies testing labeling theory often controlled for neighborhood characteristics (43.8%, e.g. neighborhood income, percentage of immigrants in the neighborhood).

On average, about four to six control variables were included. However, in some cases, more than ten control variables were added to the model (e.g. Augustyn et al., 2019; Crutchfield et al., 2009; Dennison & Finkeldey, 2021). In general, we found that a proper justification for the inclusion of control variables was often lacking. Sometimes a reference was included to a study that showed a given variable was correlated with the dependent and/or independent variables, or which was the source of the original measurement (e.g. Bates et al., 2016; Felson et al., 2005). In a few cases, only the variables that were controlled for were listed, without specifying the measurement or providing justification (e.g. Pehrson et al., 2017).

Discussion

The aim of this paper was to systematically review studies and theories that have assessed the relationship between contact with the police and criminal behavior. By giving a thorough overview of the research landscape, we aim to encourage endeavors of theoretical integration and competition as well as to increase the methodological rigor of future research. Our discussion of the results will be structured as follows: (1) We will provide an outlook of the fragmented state of different theoretical 'schools of thought' that have assessed the relationship so far, and the ways in which future research could advance. (2) We will discuss to what extent the current state of research is able to determine whether the proposed link between contact with the police and criminal behavior is 'causal' and in which ways future studies need to be designed in order to answer this question as rigorously as possible. (3) We will discuss the limited standardization of designs, measurement, and concepts, which hinder the advancement of this strand of research and how we believe increased standardization can be achieved. (4) We discuss the limitations of our study.

This review has provided an overview of the abundance of theoretical approaches concerned with the link between police contact and criminal behavior. Our findings are in line with other critical papers, suggesting that there is a large number of theoretical perspectives and hypotheses within criminology, which at the moment are not adequately empirically tested and falsified (e.g. Bruinsma, 2016; Niemeyer et al., 2022; Proctor & Niemeyer, 2019). In line with other reviews (e.g. Gibbs, 1985; Niemeyer et al., 2022), we found that the current state of research on police contact is fragmented and siloed, which renders the formulation of policy-relevant implications difficult. We believe that there are generally two ways of moving forward. Firstly, and following a theory-integration paradigm (e.g. Bernard & Snipes, 1996), it could be argued that the 'competition' among different theories concerning the relationship in question is largely of empirical nature. Accordingly, researchers might rather be concerned with the amount of variation that can be explained by different theories as variables. This approach has been summarized as the 'common research technique of simply throwing a whole bunch of variables into one gigantic regression stew' (Bernard & Snipes, 1996, pp. 302–303). We believe that such a methodological paradigm, which can be

appropriate depending on the specific research question, can also give birth to both theoretical and methodological issues.

From a theoretical perceptive, we arguably need to test implications that are unique to particular theories, and to rule out alternative explanations in the vein of Popperian falsification. However, this always controlling for all 'theories as variables' approach also poses statistical problems. Overcontrolling for all potential paths might lead researchers to control for colliders that can induce a distorted association between the exposure and outcome (e.g. see Cinelli et al., 2022). For example, psychological distress might be caused by police contact but also by criminal behavior itself. Always controlling or selecting for this variable thus creates a bad control from a causal perspective (Angrist & Pischke, 2009). Essentially, we are comparing the effect of police contact on criminal behavior among people with psychological distress if we control for it. Similar problems might arise when we control for social exclusion (labeling theory). Variables should only be included in the 'gigantic regression stew' if we can be sure that both exposure and outcome do not have a colliding path to that variable. These issues of causality lead us to our second discussion point; is the relationship between police contact and criminal behavior actually causal?

Our review found that most of the studies in our sample were classified below level 3 of the Maryland Scientific Methods Scale, which means that most studies (76%) were not able to adequately address threats to internal validity and make causal claims about whether police contact causes subsequent involvement in criminal behavior. A higher proportion of studies testing deterrence, labeling, and strain theories used more rigorous quasi- or experimental designs that could account for threats to internal validity. Focusing only on level 3 studies or above, it is possible to draw some tentative conclusions regarding strengths and weaknesses of different theoretical mechanisms, as well as potential points for integration or competition. Almost all studies testing deterrence theory and mechanisms found that police contact (typically measured as arrest) led to an *increase* in the likelihood of subsequent criminal behavior. These findings are contrary to what was expected from deterrence theory.

Based on the notion of specific deterrence, the expectation was that contact leads to a re-evaluation of the risk of sanctions, and therefore a reduction in (re-)offending (Ariel et al., 2019). Studies that found these criminogenic effects generally concluded that this was attributable to labeling processes that occur as a result of contact with the police. Specifically, being stopped or arrested was associated with weaker bonds to society, stronger neutralizations and deviant identity, and more delinquent peers (e.g. Wiley et al., 2013; Wiley & Esbensen, 2016). Generally, these studies tended to focus on expectations of specific deterrence when police contact occurs in a reactive way (i.e. in response to arrests or convictions). However, one study made a distinction between reactive and preventative specific deterrence, arguing that the ways in which the threat of sanctions is delivered by an authority matter in increasing the perception that the threat is credible (i.e. swift, certain, consequences) and therefore deterring future criminal behavior (Ariel et al., 2019; Braga et al., 2018). In this case, the threat was communicated through face-to-face interactions between officers and offenders, the incentives for compliance and consequences for offending were made clear, and the police followed up through regular unannounced visits.

So, while reactive-specific deterrence may have enough evidence to suggest falsification, preventative specific deterrence may still be relevant for theoretical integration or

competition. Importantly, preventative deterrence can still induce labeling by drawing formal attention to individuals and their networks who are not currently involved in criminal activity (Rowe & Søgaard, 2020). More theoretical and empirical work is needed to evaluate whether and to what extent these labeling processes might be mitigated or amplified during these preventative contacts.

In line with previous research, the causal evidence for other theoretical frameworks, such as procedural justice, is more mixed (Nagin & Telep, 2020). These studies were diverse in focus, with some evaluating questions about the measurement and applicability of (alternative) theoretical mechanisms (Barkworth & Murphy, 2015; Posch et al., 2021; Van Petegem et al., 2021), and others assessing questions about reverse-causality (Trinkner et al., 2019). For example, attempts of integrating strain theory into the procedural justice model showed some evidence that procedural injustice is positively related to negative affect (Barkworth & Murphy, 2015; Dennison & Finkeldey, 2021). However, since the original police legitimacy mediator was not taken into account, it is not possible to draw conclusions about which theoretical mechanisms are at work.

Taken together, the evidence for procedural justice theory suggests that there is a potential for reverse-causality, and that any studies aiming to make causal inferences need to account for a wide range of factors influencing selection into treatment (police contact), in particular prior criminal behavior (Dennison & Finkeldey, 2021). Here there is also much potential for integration. Traditionally, procedural justice theory distinguishes between normative motives for compliance and more self-interested, instrumental motives, such as in deterrence theory (Tyler, 2006). However, these are merely assumptions, and not necessarily oppositional. Much empirical research suggests that both instrumental and normative motives correlate with compliance (Bradford et al., 2015). Considering the above discussion on preventative deterrence and labeling, it may be possible to integrate procedural justice into these theoretical frameworks. For example, given that preventative deterrence involves face-to-face police-citizen interaction, the quality of treatment by the officers may influence both the credibility of sanctions and feelings of trust and moral obligations to comply with the law. The quality of treatment may also work to minimize labeling effects that may lead to increases in criminal behavior. Future research should investigate these potential points for integration in order to identify and unravel the causal mechanisms that connect police contact to criminal behavior.

There was some evidence for a causal relationship between police contact and changes in attitudes, such as police legitimacy (Mazerolle et al., 2021; Trinkner & Cohn, 2014). However, causal evidence for the relationship between legitimacy and criminal behavior, and the relationship between procedural justice and criminal behavior, is still lacking (Nagin & Telep, 2020). One key limitation to establishing such causal evidence is the relatively large number of criminological studies using cross-sectional data. Cross-sectional data not only makes it difficult to determine long-term effects, but more importantly to satisfy the assumption of temporal ordering necessary to rule out threats to internal validity (Posch et al., 2021). One solution is to control for prior levels of the mediator and/or moderator and dependent variable in the model (Walters & Mandracchia, 2017). In order to establish causal mediation, there can be no interaction between the treatment and the mediator that affects the outcome (Posch et al., 2021). This is crucial since this interaction hinders the decomposition of total, direct and indirect

effects. Reliance on cross-sectional data can increase the possibility of an interaction between the treatment variable and mediator. Another key limitation is that most proposed mediators within our sample, and criminology in general, were concepts that reflect individuals' subjective perceptions and attitudes. It becomes difficult to establish temporal order, make a conceptual distinction between them, and to exclude any interaction between them (for example when subjective procedural justice is the treatment variable and police legitimacy is the mediator). In addition, certain attitudes are not randomly distributed across the population and tend to be relatively stable over time (e.g. perceptions of procedural justice, Augustyn, 2016). This feeds into another problem, namely the causal mediation assumption that there should be no unmeasured confounders (Posch et al., 2021). In order to ensure that this assumption is met, random assignment of the treatment variable and an exhaustive set of relevant pre-treatment covariates is required. Within our sample, we found very few studies that randomly assigned the treatment variable, performed matching techniques, or controlled for an exhaustive set of covariates. In order to assess causal effects, future studies should take care to achieve random or 'as-good-as-random' assignment of the treatment variable (i.e. police contact). While challenging, previous studies have shown that it is possible to randomly assign contact (e.g. Ariel et al., 2019) and types of contact (e.g. quality of treatment, Mazerolle et al., 2021).

Another crucial point to make is that the large variety of conceptualizations and operationalizations found in empirical research poses a major challenge to assessing support for the different pathways that link police contact to criminal behavior. Many concepts have been introduced and measured inconsistently across studies. For instance, procedural justice has been measured using different sets of components: some studies emphasize trustworthiness, respectful treatment, neutrality, and voice (e.g. Barkworth & Murphy, 2015; Tyler, 2006), while others focus only on voice and impartiality (e.g. Trinkner et al., 2019). This inconsistency is problematic for evaluating causal support for theoretical mechanisms, as it complicates comparisons across studies and limits the accuracy of meta-analyses.

The field would benefit greatly from the clarification of measurement, particularly with regard to police contact and criminal behavior outcomes. Researchers should clearly define key features of police interactions, such as who initiated the contact (police or citizen), the context of the encounter (e.g. arrest or random stops), and the behavior of the police during the interaction. However, it is crucial to emphasize that the field first needs to establish robust measures that rely on sound psychometric principles before standardization can occur. The current measures are inadequate – there is virtually no construct validation, widespread issues with discriminant validity, and significant bias introduced by drawing indicators from the same measurement source. These foundational issues severely undermine the reliability and validity of existing research (e.g. Elson et al., 2023; Graham, 2019). Addressing these problems is an essential first step; only after developing high-quality, psychometrically sound measures can standardization be meaningfully implemented. By taking these steps, the field can build a more rigorous foundation, facilitating accurate comparisons, robust meta-analyses, and a deeper understanding of the mechanisms linking police contact to criminal behavior.

Similarly, standardization of the measurement of crime is necessary. Within our sample, it was noticeable that authors referred to the concept of crime in many different ways (e.g.

offending, compliance with the law, rule-violating behavior, or criminal behavior). In line with Frankenhuis et al. (2023), we found that theories need more precision and clarity about what exactly they hypothesize, what is the scope of the theory, and how key concepts should be measured. The current lack of clarity gives space to an abundance of untestable and unfalsifiable theories and hypotheses, which prevents the field from moving forward. A good starting point is Frankenhuis et al. (2023, p. 14)'s six questions for evaluating theoretical transparency: (1) is each term defined? (2) Are all relations between terms specified? (3) Are all assumptions stated explicitly? (4) Has the theory been formalized? (5) Is the scope of the theory well specified? (6) Is the theory consistent across papers? Both authors and reviewers should keep these questions in mind when evaluating the development and assessment of theoretical mechanisms.

Limitations

Despite our rigorous and extensive search strategy, it is possible that we have missed studies on this topic. While we aimed to be comprehensive, the inherent limitations of search algorithms and the variability in study indexing and keywords might have led to the omission of relevant research. This potential gap could influence the overall conclusions drawn from our analysis, only if omitted studies were selective and have characteristics that would significantly change the results.

Secondly, our search led to a large sample of studies encompassing a wide array of theoretical pathways and variables. This extensive scope made it impossible to describe every aspect in detail. As a result, we summarized the key findings and made strategic decisions to group the information, primarily focusing on the four predominant theoretical perspectives within the sample. This approach, while necessary for manageability, might have overlooked some nuances and less common perspectives that could provide additional insights.

Thirdly, the coding strategy was designed based on extracting information from the studies that the authors considered relevant for comparison between studies. We extracted this information to answer our questions and would argue that others would have found the same information and thus findings if they started with the same set of questions as we did. However, other researchers may set off with different questions, and prioritize different aspects, therefore leading them to extract alternative information from the same set of studies. We encourage others to re-examine the data with their own perspectives and objectives in mind. This diversity in analysis can enrich the field and provide a more comprehensive understanding of the topic.

Fourth, we did not assess construct validity, but future research should systematically assess the measurement of all key theoretical constructs relevant for testing these relationships.

Conclusion

This scoping review aimed to create an inventory of quantitative empirical research that has examined the relationship between police contact and criminal behavior. Studies on this topic were primarily based on four major theoretical frameworks: procedural justice theory, strain theory, labeling theory and deterrence theory, with little integration or

competition between them. The broad variety in measurements and the relatively small number of studies that were able to adequately address threats to internal validity demonstrated the difficulty in drawing conclusions about the strength of a given theory and where the weight of the evidence lies. There is a pressing need for more clear and falsifiable hypotheses and attention to designs that allow for causal inferences. Furthermore, the measurement of key variables should be standardized in order to be comparable across theoretical frameworks. By cooperation and integration, instead of fragmentation, between schools of thought, criminologists can provide more rigorous explanations of crime as well as clear recommendations about crime prevention and reduction in society.

Notes

1. Only studies using a sample of which at least part of the participants had actual police contact were included. Studies measuring police contact by asking about attitudes towards or perceptions of the police, without there having been contact between at least part of the participants and the police, were excluded. Studies testing procedural justice theory, which only included general perceptions of the police without there being any indicator of participants having police contact were excluded as well.
2. As the focus of this scoping review is on criminal outcomes, studies solely measuring outcome variables such as perceived duty to obey the law, were excluded from the review. Studies measuring intentions for criminal behaviour were not excluded.
3. Multi-level studies were only included in case they considered specific individual level interactions with police and individual level outcomes.
4. Nvivo is a software program that is suitable for qualitative analysis, such as the analysis and coding of journal articles. We coded large parts of text based on our coding scheme, which was then analyzed both quantitatively (counting up codes) and qualitatively (assessing arguments in the text) for the results.
5. The following sampling strategies were defined as probability sampling strategies: cluster, community, disproportionate stratified, independent probability, and stratified sampling. Non-probability sampling strategies involved: convenience, mixed non-probability, purposive, and snowball sampling.
6. Tests of theories refer to empirical tests of theoretical arguments. Within one study, there can be multiple different theoretical arguments that are being tested simultaneously.
7. These involve studies in which the authors explicitly incorporate other theoretical concepts or tested against other theories.
8. It is difficult to assess precisely when a study has 'adequately' controlled for confounders, and the guidelines used in the Maryland scale are not entirely clear about the benchmark for 'adequate'. Since level 2 studies are still expected to have omitted variable bias, we take a lenient approach to classifying studies as level 2 instead of level 1.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the Netherlands Organization for Scientific Research (NWO) Talent Grant [grant number: VI.Vidi.191.135].

ORCID

Amina Op de Weegh http://orcid.org/0000-0003-2841-6721
Amy Nivette http://orcid.org/0000-0003-0597-3648
Christof Nägel http://orcid.org/0000-0003-2161-539X
Stijn Ruiter http://orcid.org/0000-0003-2872-2710

References

References ending with an asterisk (*) were included in the review.

Ageton, S. S., & Elliott, D. S. (1974). The effects of legal processing on delinquent orientations. *Social Problems*, *22*(1), 87–100. https://doi.org/10.2307/799569*

Agnew, R. (2001). Building on the foundation of general strain theory: Specifying the types of strain most likely to lead to crime and delinquency. *Journal of Research in Crime and Delinquency*, *38*(4), 319–361. https://doi.org/10.1177/0022427801038004001

Agnew, R. (2011). *Toward an unified criminology: Integrating assumptions about crime, people, and society*. New York University Press.

Agnew, R., & White, H. R. (1992). An empirical test of general strain theory. *Criminology*, *30*(4), 475–500. https://doi.org/10.1111/j.1745-9125.1992.tb01113.x

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Ariel, B., Englefield, A., & Denley, J. (2019). I heard it through the grapevine: A randomized controlled trial on the direct and vicarious effects of preventative specific deterrence initiatives in criminal networks. *Journal of Criminal Law & Criminology*, *109*(4), 819–868.

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Augustyn, M. B. (2016). Updating perceptions of (in) justice. *Journal of Research in Crime and Delinquency*, *53*(2), 255–286. https://doi.org/10.1177/0022427815616991

Augustyn, M. B., Ward, J. T., Krohn, M. D., & Dong, B. (2019). Criminal justice contact across generations: Assessing the intergenerational labeling hypothesis. *Journal of Developmental and Life-Course Criminology*, *5*(2), 137–175. https://doi.org/10.1007/s40865-019-00118-3*

Barkworth, J. M., & Murphy, K. (2015). Procedural justice policing and citizen compliance behaviour: The importance of emotion. *Psychology, Crime & law*, *21*(3), 254–273. https://doi.org/10.1080/1068316X.2014.951649*

Baron, S. W. (2019). Police strain, negative emotions, criminal propensity, and criminal coping. *American Journal of Criminal Justice*, *44*(6), 938–961. https://doi.org/10.1007/s12103-019-9471-0*

Baron, S. W., & Macdonald, S. (2020). Police contact, procedural injustice, and drug use. *Canadian Journal of Criminology and Criminal Justice*, *62*(1), 1–25. https://doi.org/10.3138/cjccj.2019-0028*

Bates, L., Allen, S., & Watson, B. (2016). The influence of the elements of procedural justice and speed camera enforcement on young novice driver self-reported speeding. *Accident Analysis & Prevention*, *92*, 34–42. https://doi.org/10.1016/j.aap.2016.03.023*

Beardslee, J., Miltimore, S., Fine, A., Frick, P. J., Steinberg, L., & Cauffman, E. (2019). Under the radar or under arrest: How is adolescent boys' first contact with the juvenile justice system related to future offending and arrests? *Law and Human Behavior*, *43*(4), 342. https://doi.org/10.1037/lhb0000340*

Bello, P. O., & Matshaba, T. D. (2020). What predicts university students' compliance with the law: Perceived legitimacy or dull compulsion? *Cogent Social Sciences*, *6*(1), 1803466. https://doi.org/10.1080/23311886.2020.1803466*

Bernard, T. J., & Snipes, J. B. (1996). Theoretical integration in criminology. *Crime and Justice*, *20*, 301–348. https://doi.org/10.1086/449245

Bouffard, L. A., & Piquero, N. L. (2010). Defiance theory and life course explanations of persistent offending. *Crime & Delinquency*, *56*(2), 227–252. https://doi.org/10.1177/0011128707311642*

Bradford, B., Hohl, K., Jackson, J., & MacQueen, S. (2015). Obeying the rules of the road: Procedural justice, social identity, and normative compliance. *Journal of Contemporary Criminal Justice*, *31*(2), 171–191. https://doi.org/10.1177/1043986214568833*

Braga, A. A., Weisburd, D., & Turchan, B. (2018). Focused deterrence strategies and crime control: An updated systematic review and meta-analysis of the empirical evidence. *Criminology & Public Policy*, *17*(1), 205–250. https://doi.org/10.1111/1745-9133.12353

Braithwaite, J. (1989). *Crime, shame and reintegration*. Cambridge University Press.

Bruinsma, G. (2016). Proliferation of crime causation theories in an era of fragmentation: Reflections on the current state of criminological theory. *European Journal of Criminology*, *13*(6), 659–676. https://doi.org/10.1177/1477370816667884

Buchanan, M., & Krohn, M. D. (2020). Does it matter if those who matter don't mind? Effects of gang versus delinquent peer group membership on labeling processes. *Criminology*, *58*(2), 280–306. https://doi.org/10.1111/1745-9125.12237*

Cavanagh, C., Fine, A., & Cauffman, E. (2022). How do adolescents develop legal cynicism? A test of legal socialization mechanisms among youth involved in the justice system. *Justice Quarterly*, *39*(3), 478–496. https://doi.org/10.1080/07418825.2020.1801805*

Chenane, J. L. (2018). *The effects of prior police contact and neighborhood context on delinquency* [Doctoral dissertation]. University of Nebraska at Omaha*.

Chenane, J. L., Wright, E. M., & Wang, Y. (2021). The effects of police contact and neighborhood context on delinquency and violence. *Victims & Offenders*, *16*(4), 495–518. https://doi.org/10.1080/15564886.2020.1815112*

Cinelli, C., Forney, A., & Pearl, J. (2022). A crash course in good and bad controls. *Sociological Methods & Research*, *0*(0). https://doi.org/10.1177/00491241221099552

Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., Kastner, M., & Moher, D. (2014). Scoping reviews: Time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology*, *67*(12), 1291–1294. https://doi.org/10.1016/j.jclinepi.2014.03.013

Crutchfield, R. D., Skinner, M. L., Haggerty, K. P., McGlynn, A., & Catalano, R. F. (2009). Racial disparities in early criminal justice involvement. *Race and Social Problems*, *1*(4), 218–230. https://doi.org/10.1007/s12552-009-9018-y*

Del Toro, J., Lloyd, T., Buchanan, K. S., Robins, S. J., Bencharit, L. Z., Smiedt, M. G., Reddy, K. S., Pouget, E. R., Kerrison, E. M., & Goff, P. A. (2019). The criminogenic and psychological effects of police stops on adolescent black and latino boys. *Proceedings of the National Academy of Sciences*, *116*(17), 8261–8268. https://doi.org/10.1073/pnas.1808976116*

Dennison, C. R., & Finkeldey, J. G. (2021). Self-reported experiences and consequences of unfair treatment by police. *Criminology*, *59*(2), 254–290. https://doi.org/10.1111/1745-9125.12269*

Doherty, E. E., Cwick, J. M., Green, K. M., & Ensminger, M. E. (2016). Examining the consequences of the "prevalent life events" of arrest and incarceration among an urban African-American cohort. *Justice Quarterly*, *33*(6), 970–999. https://doi.org/10.1080/07418825.2015.1016089*

Dong, B. (2018). *Analyzing developmental pathways of gang membership: Risk factors, turning points and consequences*. [Doctoral dissertation]. University of Florida*

Dong, B., & Krohn, M. D. (2017). The protective effects of family support on the relationship between official intervention and general delinquency across the life course. *Journal of Developmental and Life-Course Criminology*, *3*(1), 39–61. https://doi.org/10.1007/s40865-016-0051-4*

Dong, B., & Krohn, M. D. (2020). Sent home versus being arrested: The relative influence of school and police intervention on drug use. *Justice Quarterly*, *37*(6), 985–1011. https://doi.org/10.1080/07418825.2018.1561924*

Dooley, B. D., & Goodison, S. E. (2020). Falsification by atrophy: The Kuhnian process of rejecting theory in US criminology. *The British Journal of Criminology*, *60*(1), 24–44. https://doi.org/10.1093/bjc/azz026

Elson, M., Hussey, I., Alsalti, T., & Arslan, R. C. (2023). Psychological measures aren't toothbrushes. *Communications Psychology*, *1*(1), 25. https://doi.org/10.1038/s44271-023-00026-9

Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2003). The Maryland scientific methods scale. InLawrence W. Sherman, David P. Farrington, Brandon C. Welsh, & Doris Layton MacKenzie (Eds.), *Evidence-based crime prevention* (pp. 27–35). Routledge.

Felson, R. B., Ackerman, J. M., & Gallagher, C. A. (2005). Police intervention and the repeat of domestic assault. *Criminology*, *43*(3), 563–588. https://doi.org/10.1111/j.0011-1348.2005.00017.x*

Fitzgerald, R. T., & Carrington, P. J. (2011). Disproportionate minority contact in Canada: Police and visible minority youth. *Canadian Journal of Criminology and Criminal Justice*, *53*(4), 449–486. https://doi.org/10.3138/cjccj.53.4.449*

Frankenhuis, W., Panchanathan, K., & Smaldino, P. E. (2023). Strategic ambiguity in the social sciences. *Social Psychological Bulletin*, *18*, 1–25. https://doi.org/10.32872/spb.9923

Gerber, M. M., González, R., Carvacho, H., Jiménez-Moya, G., Moya, C., & Jackson, J. (2018). On the justification of intergroup violence: The roles of procedural justice, police legitimacy, and group identity in attitudes toward violence among indigenous people. *Psychology of Violence*, *8*(3), 379. https://doi.org/10.1037/vio0000177*

Gibbs, J. P. (1985). The methodology of theory construction in criminology. In R. F. Meier (Ed.), *Theoretical methods in criminology* (pp. 23–50). Sage Publication.

Goldstein, N. E., Kreimer, R., Guo, S., Le, T., Cole, L. M., NeMoyer, A., Burke, S., Kikuchi, G., Thomas, K., & Zhang, F. (2021). Preventing school-based arrest and recidivism through prearrest diversion: Outcomes of the Philadelphia police school diversion program. *Law and Human Behavior*, *45*(2), 165–178. https://doi.org/10.1037/lhb0000440*

Graham, A. (2019). *Measuring procedural justice: A case study in criminometrics* [Doctoral dissertation]. University of Cincinnati.

Green, K. M., Doherty, E. E., Sifat, M. S., & Ensminger, M. E. (2019). Explaining continuity in substance use: The role of criminal justice system involvement over the life course of an urban African American prospective cohort. *Drug and Alcohol Dependence*, *195*, 74–81. https://doi.org/10.1016/j.drugalcdep.2018.09.033*

Hassett-Walker, C., Walsemann, K., Bell, B., Fisk, C., Shadden, M., & Zhou, W. (2017). How does early adulthood arrest alter substance use behavior? Are there differential effects by race/ethnicity and gender? . *Journal of Developmental and Life-Course Criminology*, *3*(2), 196–220. https://doi.org/10.1007/s40865-017-0060-y*

Hemphill, S. A., Toumbourou, J. W., Herrenkohl, T. I., McMorris, B. J., & Catalano, R. F. (2006). The effect of school suspensions and arrests on subsequent adolescent antisocial behavior in Australia and the United States. *Journal of Adolescent Health*, *39*(5), 736–744. https://doi.org/10.1016/j.jadohealth.2006.05.010*

Hepburn, J. R. (1977). The impact of police intervention upon juvenile delinquents. *Criminology*, *15*(2), 235–262. https://doi.org/10.1111/j.1745-9125.1977.tb00064.x*

Horwitz, A., & Wasserman, M. (1979). The effect of social control on delinquent behavior: A longitudinal test. *Sociological Focus*, *12*(1), 53–70. https://doi.org/10.1080/00380237.1979.10570335*

Hough, M., Jackson, J., & Bradford, B. (2013). Legitimacy, trust and compliance: An empirical test of procedural justice theory using the European social survey. In J. Tankebe, & A. Liebling (Eds.), *Legitimacy and criminal justice: An international exploration* (pp. 326–352). Oxford University Press.*

Isom, D. (2016). Microaggressions, injustices, and racial identity: An empirical assessment of the theory of African American offending. *Journal of Contemporary Criminal Justice*, *32*(1), 27–59. https://doi.org/10.1177/1043986215607253*

Jackson, J. (2015). On the dual motivational force of legitimate authority. In Brian H. Bornstein & Alan J. Tomkins (Eds.), *Motivating Cooperation and Compliance with Authority* (pp. 145–166). Springer. https://doi.org/10.1007/978-3-319-16151-8_7*

Jackson, J., Bradford, B., Hough, M., Myhill, A., Quinton, P., & Tyler, T. R. (2012). Why do people comply with the law? Legitimacy and the Influence of Legal Institutions. *British Journal of Criminology*, *52*(6), 1051–1071.*

Jackson, J., Huq, A. Z., Bradford, B., & Tyler, T. R. (2013). Monopolizing force? Police legitimacy and public attitudes toward the acceptability of violence. *Psychology, Public Policy, and Law*, *19*(4), 479–497. https://doi.org/10.1037/a0033852*

Jorgensen, C. (2011). *Public perceptions matter: a procedural justice study examining an arrestee population*. [Doctoral dissertation], Arizona State University*.

Kaiser, K., & Reisig, M. D. (2019). Legal socialization and self-reported criminal offending: The role of procedural justice and legal orientations. *Journal of Quantitative Criminology*, *35*(1), 135–154. https://doi.org/10.1007/s10940-017-9375-4*

Keane, C., Gillis, A. R., & Hagan, J. (1989). Deterrence and amplification of juvenile delinquency by police contact: The importance of gender and risk-orientation. *The British Journal of Criminology*, *29*(4), 336–352. https://doi.org/10.1093/oxfordjournals.bjc.a047853*

Kim, J., & Lee, Y. (2019). Does it take a school? Revisiting the influence of first arrest on subsequent delinquency and educational attainment in a tolerant educational background. *Journal of Research in Crime and Delinquency*, *56*(2), 254–302. https://doi.org/10.1177/0022427818801053*

Klahm IV, C. F., Steiner, B., & Meade, B. (2017). Assessing the relationship between police use of force and inmate offending (rule violations). *Crime & Delinquency*, *63*(3), 267–295. https://doi.org/10.1177/0011128714558291*

Langton, S. (2023). asreview_irr. GitHub. https://github.com/langtonhugh/asreview_irr

Liberman, A. M., Kirk, D. S., & Kim, K. (2014). Labeling effects of first juvenile arrests: Secondary deviance and secondary sanctioning. *Criminology*, *52*(3), 345–370. https://doi.org/10.1111/1745-9125.12039*

Liu, J., Wu, G., & Boateng, F. D. (2020). Does procedural fairness matter for drug abusers to stop illicit drug use? Testing the applicability of the process-based model in a Chinese context. *Psychology, Crime & Law*, *26*(5), 507–526. https://doi.org/10.1080/1068316X.2019.1696802*

Maxwell, C. D., Moore, A. M., Fontaine, J., & Smith, J. (2020). Status influences on perceptions of procedural justice: A test of the group value model among intimate partner violence arrestees. *International Journal of Comparative and Applied Criminal Justice*, *44*(1–2), 29–45. https://doi.org/10.1080/01924036.2019.1575253*

Mazerolle, L., Antrobus, E., Cardwell, S. M., Piquero, A. R., & Bennett, S. (2021). Harmonizing legal socialization to reduce antisocial behavior: Results from a randomized field trial of truanting young people. *Justice Quarterly*, *38*(1), 152–179. https://doi.org/10.1080/07418825.2019.1618894*

McCarthy, M., Murphy, K., Sargeant, E., & Williamson, H. (2021). Policing COVID-19 physical distancing measures: Managing defiance and fostering compliance among individuals least likely to comply. *Policing and Society*, *31*(5), 601–620. https://doi.org/10.1080/10439463.2020.1869235*

McGlynn-Wright, A. (2014). *Labels for some: Long-term consequences of police encounters and arrests for African American and White youth* [Doctoral dissertation]. University of Washington*

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*, 276–282. https://doi.org/10.11613/BM.2012.031

McKinlay, A., James, V. L., & Grace, R. C. (2015). Development of an actuarial static risk model suitable for automatic scoring for predicting juvenile recidivism. *Legal and Criminological Psychology*, *20*(2), 288–305. https://doi.org/10.1111/lcrp.12024*

McLean, K. (2017). Ethnic identity, procedural justice, and offending: Does procedural justice work the same for everyone? *Crime & Delinquency*, *63*(10), 1314–1336. https://doi.org/10.1177/0011128715620429*

Meško, G., & Tankebe, J. (2015). *Trust and legitimacy in criminal justice*. Springer International Publishing.*

Mitchell, O. (2016). The effect of drug arrest on subsequent drug offending and social bonding. *Journal of Crime and Justice*, *39*(1), 174–188. https://doi.org/10.1080/0735648X.2015.1087145*

Motz, R. T., Barnes, J. C., Caspi, A., Arseneault, L., Cullen, F. T., Houts, R., Wertz, J., & Moffitt, T. E. (2020). Does contact with the justice system deter or promote future delinquency? Results from a longitudinal study of British adolescent twins. *Criminology*, *58*(2), 307–335. https://doi.org/10.1111/1745-9125.12236.*

Mowen, T. J., Brent, J. J., & Bares, K. J. (2018). How arrest impacts delinquency over time between and within individuals. *Youth Violence and Juvenile Justice*, *16*(4), 358–377. https://doi.org/10.1177/1541204017712560*

Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or

scoping review approach. *BMC Medical Research Methodology*, *18*(1), 1–7. https://doi.org/10.1186/s12874-017-0458-6

Murphy, K. (2009). Procedural justice and affect intensity: Understanding reactions to regulatory authorities. *Social Justice Research*, *22*(1), 1–30. https://doi.org/10.1007/s11211-008-0086-8*

Murphy, K. (2016). Turning defiance into compliance with procedural justice: Understanding reactions to regulatory encounters through motivational posturing. *Regulation & Governance*, *10*(1), 93–109. https://doi.org/10.1111/rego.12073*

Murray, J., Loeber, R., & Pardini, D. (2012). Parental involvement in the criminal justice system and the development of youth theft, marijuana use, depression, and poor academic performance. *Criminology*, *50*(1), 255–302. https://doi.org/10.1111/j.1745-9125.2011.00257.x*

Murray, K., McVie, S., Farren, D., Herlitz, L., Hough, M., & Norris, P. (2021). Procedural justice, compliance with the law and police stop-and-search: A study of young people in England and Scotland. *Policing and Society*, *31*(3), 263–282. https://doi.org/10.1080/10439463.2020.1711756*

Nagin, D. S., & Telep, C. W. (2020). Procedural justice and legal compliance: A revisionist perspective. *Criminology & Public Policy*, *19*(3), 761–786. https://doi.org/10.1111/1745-9133.12499

Niemeyer, R. E., Proctor, K. R., Schwartz, J. A., & Niemeyer, R. G. (2022). Are most published criminological research findings wrong? Taking stock of criminological research using a Bayesian simulation approach. *International Journal of Offender Therapy and Comparative Criminology*, 1–20.

Paternoster, R. (2010). How much do we really know about criminal deterrence? *Journal of Criminal Law & Criminology*, *100*(3), 765–824.

Pedersen, W. (2000). Crime and punishment in Norwegian mid-adolescents: A normal population study. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, *1*(1), 87–104. https://doi.org/10.1080/14043850050116282*

Pehrson, S., Devaney, L., Bryan, D., & Blaylock, D. L. (2017). Beyond group engagement: Multiple pathways from encounters with the police to cooperation and compliance in northern Ireland. *PLoS One*, *12*(9), e0184436. https://doi.org/10.1371/journal.pone.0184436*

Petersen, K., Weisburd, D., Fay, S., Eggins, E., & Mazerolle, L. (2023). Police stops to reduce crime: A systematic review and meta-analysis. *Campbell Systematic Reviews*, *19*(1), 1–42. https://doi.org/10.1002/cl2.1302.

Pina-Sánchez, J., & Brunton-Smith, I. (2021). Are We All equally persuaded by procedural justice?: Re-examining the invariance thesis using longitudinal data and random effects. *Journal of Developmental and Life-Course Criminology*, *7*(3), 1–24. https://doi.org/10.1007/s40865-021-00170-y*

Pollock, W., Menard, S., & Hill, M. C. (2016). Revisiting "measuring the problem" separate examination of police contact in serious and nonserious offenders. *Criminal Justice Review*, *41*(3), 294–317. https://doi.org/10.1177/0734016816648188*

Pollock, W., Oliver, W., & Menard, S. (2012). Measuring the problem: A national examination of disproportionate police contact in the United States. *Criminal Justice Review*, *37*(2), 153–173. https://doi.org/10.1177/0734016811436335*

Posch, K., Jackson, J., Bradford, B., & Macqueen, S. (2021). Truly free consent"? Clarifying the nature of police legitimacy using causal mediation analysis. *Journal of Experimental Criminology*, *17*(4), 563–595. https://doi.org/10.1007/s11292-020-09426-x*

Pósch, K. (2021). Prying open the black box of causality: A causal mediation analysis test of procedural justice policing. *Journal of Quantitative Criminology*, *37*(1), 217–245. https://doi.org/10.1007/s10940-020-09449-7

Proctor, K. R., & Niemeyer, R. E. (2019). *Mechanistic criminology*. Routledge.

Quinn, K., Walsh, J. L., & Dickson-Gomez, J. (2019). Multiple marginality and the variation in delinquency and substance use among adolescent gang members. *Substance Use & Misuse*, *54*(4), 612–627. https://doi.org/10.1080/10826084.2018.1528465*

Raaijmakers, E. A., Loughran, T. A., de Keijser, J. W., Nieuwbeerta, P., & Dirkzwager, A. J. (2017). Exploring the relationship between subjectively experienced severity of imprisonment and recidivism: A neglected element in testing deterrence theory. *Journal of Research in Crime and Delinquency*, *54*(1), 3–28. https://doi.org/10.1177/0022427816660588

Rowe, M., & Søgaard, T. F. (2020). 'Playing the man, not the ball': Targeting organised criminals, intelligence and the problems with pulling levers. *Policing and Society*, *30*(2), 120–135. https://doi.org/10.1080/10439463.2019.1603226

Sampson, R. J., & Laub, J. H. (1990). Crime and deviance over the life course: The salience of adult social bonds. *American Sociological Review*, *55*(5), 609–627. https://doi.org/10.2307/2095859.

Schulz, S. (2014). Individual differences in the deterrence process: Which individuals learn (most) from their offending experiences? *Journal of Quantitative Criminology*, *30*(2), 215–236. https://doi.org/10.1007/s10940-013-9201-6*

Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of Research in Crime and Delinquency*, *30*(4), 445–473. https://doi.org/10.1177/0022427893030004006

Sherman, L. W., Smith, D. A., Schmidt, J. D., & Rogan, D. P. (1992). Crime, punishment, and stake in conformity: Legal and informal control of domestic violence. *American Sociological Review*, *57*(5), 680–690. https://doi.org/10.2307/2095921*

Skardhamar, T., & Savolainen, J. (2014). Changes in criminal offending around the time of job entry: A study of employment and desistance. *Criminology*, *52*(2), 263–291. https://doi.org/10.1111/1745-9125.12037

Slocum, L. A., & Wiley, S. A. (2018). Experience of the expected?" Race and ethnicity differences in the effects of police contact on youth. *Criminology*, *56*(2), 402–432. https://doi.org/10.1111/1745-9125.12174*

Slocum, L. A., Wiley, S. A., & Esbensen, F. A. (2016). The importance of being satisfied: A longitudinal exploration of police contact, procedural injustice, and subsequent delinquency. *Criminal Justice and Behavior*, *43*(1), 7–26. https://doi.org/10.1177/0093854815609069*

Snipes, J. B., Maguire, E. R., & Tyler, D. H. (2019). The effects of procedural justice on civil disobedience: Evidence from protesters in three cities. *Journal of Crime and Justice*, *42*(1), 32–44. https://doi.org/10.1080/0735648X.2018.1559128*

Tankebe, J. (2009). Self-help, policing, and procedural justice: Ghanaian vigilantism and the rule of law. *Law & Society Review*, *43*(2), 245–270. https://doi.org/10.1111/j.1540-5893.2009.00372.x*

Tauchen, H. V., Witte, A. D., & Griesinger, H. (1994). Criminal deterrence: Revisiting the issue with a birth cohort. *The Review of Economics and Statistics*, *76*(3), 399–412. https://doi.org/10.2307/2109966*

Thomas, C. W., & Bishop, D. M. (1984). The effect of formal and informal sanctions on delinquency: A longitudinal comparison of labeling and deterrence theories. *Journal of Criminal Law & Criminology*, *75*(4), 1222–1245. https://doi.org/10.2307/1143384*

Thomas, J. M., Thomas, S. A., Burgason, K. A., & Wichinsky, L. C. (2014). Early contact with the criminal justice system and intellectual functioning as risk factors for violent and chronic adult offending. *Criminology, Criminal Justice, Law & Society*, *15*(1), 34–50.

Thompson, A. J., & Pickett, J. T. (2021). Asymmetry in process-based model relationships: A longitudinal study of adjudicated adolescents. *Criminology*, *59*(4), 585–609. https://doi.org/10.1111/1745-9125.12279*

Thornberry, T. P. (1987). Toward an interactional theory of delinquency. *Criminology*, *25*(4), 863–892. https://doi.org/10.1111/j.1745-9125.1987.tb00823.x

Thrane, L. E. (2003). *Predictors of police harassment and arrests among homeless and runaway adolescents*. [Doctoral dissertation]. Iowa State University*

Trinkner, R., & Cohn, E. S. (2014). Putting the "social" back in legal socialization: Procedural justice, legitimacy, and cynicism in legal and nonlegal authorities. *Law and Human Behavior*, *38*(6), 602–617. https://doi.org/10.1037/lhb0000107*

Trinkner, R., Jackson, J., & Tyler, T. R. (2018). Bounded authority: Expanding "appropriate" police behavior beyond procedural justice. *Law and Human Behavior*, *42*(3), 280–293. https://doi.org/10.1037/lhb0000285*

Trinkner, R., Mays, R. D., Cohn, E. S., Van Gundy, K. T., & Rebellon, C. J. (2019). Turning the corner on procedural justice theory: Exploring reverse causality with an experimental vignette in a longitudinal survey. *Journal of Experimental Criminology*, *15*(4), 661–671. https://doi.org/10.1007/s11292-019-09358-1*

Trinkner, R., Rodrigues, H., Piccirillo, D., Gifford, F. E., & Gomes, A. M. M. (2020). Legal socialisation in Brazil: Examining the generalisability of the procedural justice model. *International Journal of Comparative and Applied Criminal Justice*, *44*(1–2), 7–27. https://doi.org/10.1080/01924036.2019.1587710*

Tyler, D. H., Barak, M., Maguire, E. R., & Wells, W. (2018). The effects of procedural injustice on the use of violence against police by occupy Wall Street protesters. *Police Practice and Research*, *19*(2), 138–152. https://doi.org/10.1080/15614263.2018.1418153*

Tyler, T. R. (2006). *Why people obey the law*. Princeton University Press.*.

Valdimarsdóttir, M. (2020). *Examining the contextual effects of racial profiling, and the long-term consequences of punitive interventions: Testing labeling theory with the national longitudinal study of adolescent to adult health data* [Doctoral dissertation]. City University of New York*.

Van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, *3*(2), 125–133. https://doi.org/10.1038/s42256-020-00287-7

Van Domburgh, L., Vermeiren, R. R. J. M., Blokland, A. A., & Doreleijers, T. A. (2009). Delinquent development in Dutch childhood arrestees: Developmental trajectories, risk factors and co-morbidity with adverse outcomes during adolescence. *Journal of Abnormal Child Psychology*, *37*(1), 93–105. https://doi.org/10.1007/s10802-008-9260-6*

Van Gelder, J. L., Averdijk, M., Ribeaud, D., & Eisner, M. (2020). Sanctions, short-term mindsets, and delinquency: Reverse causality in a sample of high school youth. *Legal and Criminological Psychology*, *25*(2), 199–218. https://doi.org/10.1111/lcrp.12170*

Van Petegem, S., Trinkner, R., van der Kaap-Deeder, J., Antonietti, J. P., & Vansteenkiste, M. (2021). Police procedural justice and adolescents' internalization of the law: Integrating self-determination theory into legal socialization research. *Journal of Social Issues*, *77*(2), 336–366. https://doi.org/10.1111/josi.12425*

Walsh, H., Myers, T. D. W., Ray, J. V., Frick, P. J., Thornton, L. C., Steinberg, L., & Cauffman, E. (2019). Perceptions of police-juvenile contact predicts self-reported offending in adolescent males. *Psychology, Crime & Law*, *25*(10), 963–976. https://doi.org/10.1080/1068316X.2019.1597094*

Walters, G. D. (2018). Procedural justice, legitimacy beliefs, and moral disengagement in emerging adulthood: Explaining continuity and desistance in the moral model of criminal lifestyle development. *Law and Human Behavior*, *42*(1), 37–49. https://doi.org/10.1037/lhb0000266*

Walters, G. D., & Bolger, P. C. (2019). Procedural justice perceptions, legitimacy beliefs, and compliance with the law: A meta-analysis. *Journal of Experimental Criminology*, *15*(3), 341–372. https://doi.org/10.1007/s11292-018-9338-2

Walters, G. D., & Mandracchia, J. T. (2017). Testing criminological theory through causal mediation analysis: Current status and future directions. *Journal of Criminal Justice*, *49*, 53–64. https://doi.org/10.1016/j.jcrimjus.2017.02.002

Ward, J. T., Krohn, M. D., & Gibson, C. L. (2014). The effects of police contact on trajectories of violence: A group-based, propensity score matching analysis. *Journal of Interpersonal Violence*, *29*(3), 440–475. https://doi.org/10.1177/0886260513505219*

Weisburd, D., Morris, N. A., & Ready, J. (2008). Risk-focused policing at places: An experimental evaluation. *Justice Quarterly*, *25*(1), 163–200. https://doi.org/10.1080/07418820801954647*

Weisburd, D., & Piquero, A. R. (2008). How well do criminologists explain crime? *Statistical Modeling in Published Studies. Crime and Justice*, *37*(1), 453–502. https://doi.org/10.1086/524284

Wellford, C. F., & Aultman, M. G. (1981). Police contact and delinquency: A relevant evaluation. *Australian & New Zealand Journal of Criminology*, *14*(2), 110–118. https://doi.org/10.1177/000486588101400207*

Wiley, S. A. (2014). *The amplification of deviance following police contact: an examination of individual and neighborhood factors among a sample of youth* [Doctoral dissertation]. University of Missouri-Saint Louis.*

Wiley, S. A. (2015). Arrested development: Does the grade level at which juveniles experience arrest matter? *Journal of Developmental and Life-Course Criminology*, *1*(4), 411–433. https://doi.org/10.1007/s40865-015-0019-9*

Wiley, S. A., Carson, D. C., & Esbensen, F. A. (2017). Arrest and the amplification of deviance: Does gang membership moderate the relationship? *Justice Quarterly*, *34*(5), 788–817. https://doi.org/10.1080/07418825.2016.1226936*

Wiley, S. A., & Esbensen, F. A. (2016). The effect of police contact: Does official intervention result in deviance amplification? *Crime & Delinquency*, *62*(3), 283–307. https://doi.org/10.1177/0011128713492496*

Wiley, S. A., Slocum, L. A., & Esbensen, F. A. (2013). The unintended consequences of being stopped or arrested: An exploration of the labeling mechanisms through which police contact leads to subsequent delinquency. *Criminology*, *51*(4), 927–966. https://doi.org/10.1111/1745-9125.12024*