

RE-EVALUATING REACTIVE POLICING

A MINORITY REPORT

VRIJE UNIVERSITEIT

RE-EVALUATING REACTIVE POLICING
A MINORITY REPORT

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof. dr. Jeroen J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op [defenseDay] [DATE] om [TIME xx.xx] uur
in [de aula / het auditorium] van de universiteit,
De Boelelaan 1105

door

Tim Valentijn Verlaan

geboren te Utrecht

promotoren: prof.dr. S. Ruiter
prof.dr. R.D. van der Mei

This dissertation was approved by

Promoter:

Prof. dr. Stijn Ruiter	NSCR / Vrije Universiteit Amsterdam, The Netherlands
------------------------	---

Promoter:

Prof. dr. Rob van der Mei	CWI / Vrije Universiteit Amsterdam, The Netherlands
---------------------------	--

Dissertation Committee

Chairman:

Rector Magnificus, Prof. dr. Jeroen J.G. Geurts	Vrije Universiteit Amsterdam, The Netherlands
--	--

Committee:

Prof. dr. Sandjai Bhulai	Vrije Universiteit Amsterdam, The Netherlands
Prof. dr. Wim Bernasco	NSCR / Universiteit van Amsterdam, Amsterdam, The Netherlands
Prof. dr. Michael Townsley	Griffith University, Brisbane, Australia
Prof. dr. Cynthia Lum	George Mason University, Fairfax, United States of America
Dr. Toby Davis	University of Leeds, Leeds, United Kingdom

Keywords: Evidence-based Policing, Proactive Policing, Reactive Policing

Printed by:

Cover:

The author set this thesis in L^AT_EX using the XY and YX fonts.

ISBN: xxxxx

An electronic version of this thesis is available at <https://www.vu.nl/>.

© 2025, Tim Verlaan, Amsterdam, the Netherlands.



VRIJE
UNIVERSITEIT
AMSTERDAM



Nederlands Studiecentrum
Criminaliteit en Rechtshandhaving

This research was carried out within the What Works in Policing research programme of the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) and funded by the Netherlands Police. The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Copyright ©

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author.

What can be asserted without evidence can also be dismissed without evidence.

Christopher Hitchens

There are no solutions. There are only trade-offs.

Thomas Sowell

CONTENTS

Introduction	1
A Tale of Chiefs and Towers	1
Outline of This Introductory Chapter	2
The Role of Police in Society.	2
The Role of the Police in the Netherlands	4
A Pragmatic Approach to Policing	5
The Standard Model of Policing and Proactive and Reactive Policing	6
Deterrence Theory	6
The Pragmatic Policing Model	7
Imbalance in the Evidence Base Detailed	8
Outline of This Thesis	10
Contributions.	11
Guiding Principles	11
References	13
List of Publications	17
I Revisiting Reactive Policing	19
1 Predictors of Police Response Time: A Scoping Review	21
1.1 Introduction	22
1.1.1 Police Response Time	22
1.2 Methods	23
1.2.1 Screening Criteria	23
1.2.2 Information Sources	24
1.2.3 Search Query.	24
1.2.4 Selection Phase.	25
1.2.5 Data Extraction.	26
1.2.6 Data Conversion and Interpretation	26
1.2.7 Critical Appraisal of Research Methodologies	27
1.3 Results	28
1.3.1 Empirically Tested Predictors of Police Response Time	28
1.3.2 Categories of Predictor Variables	29
1.3.3 Challenges to Synthesizing the Findings	29
1.4 Strategy for Synthesizing the Findings	31
1.4.1 Narrative Synthesis.	31

1.5	Discussion	34
1.5.1	Discussion of the Literature.	34
1.5.2	Discussion of the Results	35
1.5.3	Limitations.	36
1.5.4	Recommendations for Future Research Into Police Response Time	37
1.5.5	Towards a Causal Model of Police Response Time	39
	References	41
A	Appendix: <i>Scoping Review</i>	45
A.1	Flow Chart	45
A.2	Included Studies	46
A.3	Data Extraction Form.	47
A.4	Data Conversion and Interpretation Form	48
A.5	Hybrid Synthesis	49
A.6	Causal Model of Police Response Time	51
2	On Re-evaluating Rapid Response	53
2.1	Introduction	54
2.2	Early Research on Rapid Response	55
2.2.1	The Kansas City Response Time Analysis Study	56
2.2.2	The Police Executive Research Forum (Replication) Study	57
2.2.3	Summarizing Findings of Early Rapid Response Studies	58
2.3	The Standard View of Rapid Response	58
2.4	A Critique of the Standard View	59
2.4.1	No (Indirect) Effect	59
2.4.2	Strong Evidence	60
2.4.3	Large Samples	60
2.4.4	Careful Measurements: Reliability and Validity.	60
2.5	Contemporary Validity	63
2.5.1	Contemporary Validity of Proportion of Involvement Crimes	63
2.5.2	Contemporary Validity of Citizen Reporting Delays.	65
2.5.3	Conclusion of the Critique	65
2.6	Subsequent Research on the Effects of Rapid Response on Arrest	67
2.7	Concluding on the State of Rapid Response Research	69
2.8	Future Research on the Impact of Rapid Response	70
	References	73
3	The Maryland Report: <i>Constructing the Proactive Policing Paradigm</i>	77
3.1	Evaluating the Standard Model of Policing	78
3.2	Impact of These Landmark Evaluation Studies	79
3.3	The Maryland Report.	80
3.4	The Maryland Report on the Standard Model of Policing	80
3.4.1	Rapid Response	80
3.4.2	Random Patrol	83
3.4.3	Directed Patrol	84

3.5	Aftermath of the Maryland Report	86
3.6	Conclusion	87
	References	88
II	Investigating Police Influence on Response Times	91
4	The Effect of Police Deployment Strategy on Emergency Response Times: An Agent-Based Modelling Investigation	93
4.1	Introduction	94
4.2	Methodology	98
4.2.1	Primary Data	99
4.2.2	Agent-Based Model Development	99
4.2.3	Model Robustness	103
4.2.4	Experimental Setup.	103
4.2.5	Model Outputs and Data Analysis	103
4.3	Results	105
4.3.1	Average Effect of Deployment Strategies	105
4.3.2	Moderating Effect of Police Capacity	108
4.3.3	Moderating Effect of Urbanisation	108
4.3.4	Moderating Effect of Call Volume	108
4.4	Summary.	110
4.5	Discussion	111
4.5.1	Limitations.	112
4.5.2	Future Research and Broader Implications	112
	References	113
B	Appendix: ABM Investigation	117
B.1	Input Data	117
B.2	Data Preprocessing	118
B.3	Histogram of Priority 1 Call Volume per Shift per front line team	119
B.4	Histogram of Police Capacity per Shift per front line team	120
III	Mapping Reactive Demand and Supply	121
5	On the Use of Inferential Statistics On Administrative Police Data	123
5.1	Introduction	124
5.2	Understanding Police Data	124
5.2.1	Apparent Populations.	125
5.2.2	Approach of the Field	127
5.3	Evidence from the Field	127
5.4	The Consequences	129
5.4.1	Wrongful Suggestions of Generalisability of the Observed Differences/Effects	129
5.4.2	Undervaluing of Observed Differences/Effects	130
5.4.3	Obscuring of the Observed Differences/Effects	130

5.5	Hypothesised Drivers of Misuse	131
5.6	Proposed Remedies for Misuse	132
5.7	Concluding Remarks	133
5.8	Acknowledgements	133
	References	134
6	Operationalizing Deployment Time In Police Calls-For-Service	137
6.1	Background	138
6.2	Data & Methods	139
6.2.1	Unit-Level Dispatch Data	139
6.2.2	Operationalisation of ‘Time Spent’	139
6.2.3	Call Classification	140
6.3	Methods	141
6.4	Results	141
6.4.1	Time Spent	143
6.4.2	Multi-Dispatch Incident Types	143
6.5	Discussion	145
6.6	Availability of Data and Materials	146
	References	147
7	Describing the Scale and Composition of Calls for Police Service: A Replication and Extension Using Open Data	149
7.1	Background	150
7.2	Describing Demand	151
7.2.1	Ratcliffe (2021).	152
7.2.2	Lum et al. (2022).	152
7.3	Data	153
7.3.1	Detroit Calls For Service	153
7.4	Methods	154
7.4.1	Descriptive Statistics	154
7.4.2	Spatio-Temporal Patterning	154
7.5	Results	155
7.5.1	Scale and composition	155
7.5.2	Time Distributions	156
7.5.3	Spatio-Temporal Patterning	157
7.6	Discussion	161
	References	164
C	Appendix: Scale and Composition of Police Calls-for-Service	167
C.1	Raw Count Maps	168
C.2	Time-on-Scene Distributions	169
8	Reactive Police Deployment on the Street	171
8.1	A Multitude of Tasks	173
8.2	Differences Between Frontline Teams	174
8.3	Reflections on Findings	175
8.3.1	Recognition	175

8.3.2	Training	176
8.3.3	Internal Collaboration	176
8.3.4	Local Steering	176

IV Evaluating Proactive Policing in the Netherlands 177

9 Evaluating Place-Based Policing in Amsterdam: A *Quasi-Experimental Study of Calls-for-Service Levels and Public Perceptions Using a Micro-Synthetic Control Framework* 179

9.1	Introduction	180
9.2	Hypotheses.	182
9.3	Background Information on the Intervention	183
9.4	Study Area	185
9.5	Data	186
9.5.1	NCVS Data ('VMAA')	186
9.5.2	Variables.	187
9.6	Methodology.	188
9.6.1	Difference-in-difference Setup	188
9.6.2	Difference-in-difference Analysis Using Synthetic Controls	189
9.6.3	Experimental Variable: Police Presence.	190
9.6.4	Outcome Metrics for Survey Data and Calls-for-Service.	190
9.6.5	Inference Criteria.	191
9.6.6	Robustness Check: Survey Mode Effect Analysis	192
9.7	Results	193
9.7.1	Effects of the Intervention on Calls-for-service Counts	193
9.7.2	Effects of the Intervention as Measured by the NCVS	193
9.8	Mechanisms	195
9.8.1	Implementation.	196
9.9	Conclusions	198
9.9.1	Limitations.	199
9.10	Discussion	199
	References	202

D Appendix: *Evaluating Place-Based Policing in Amsterdam* 205

D.1	Descriptive Statistics of Survey Data	206
D.2	Deviations From the Pre-registration	210
D.3	Survey Data Forest Plots	211
D.4	Supplementary Analysis Results	215

Discussion 217

Part I – Revisiting Rapid Response	217
Summary of Chapters 1–3	217
Discussion of Part I.	218
Implications for Practice	218
Implications for (Future) Research	219

Part II – Investigating Police Influence on Response Times	220
Summary of Chapter 4	220
Discussion of Part II	220
Implications for Practice	220
Implications for (Future) Research	221
Part III - Mapping Reactive Demand and Supply	222
Summary of Chapters 5–8	222
Discussion of Part III.	223
Implications for Practice	223
Implications for (Future) Research	223
Part IV – Evaluating Proactive Policing in the Netherlands	224
Summary of Chapter 9	224
Discussion of Part IV.	225
Implications for Practice	225
Implications for (Future) Research	226
Concluding Remarks	227
References	228
Bibliography	231

INTRODUCTION

A TALE OF CHIEFS AND TOWERS

Imagine a group of police chiefs approaching the ivory towers of academia, seeking guidance. The chiefs have each been granted an additional thousand officers to deploy over the next years in their jurisdictions. Faced with this unprecedented opportunity, they are confronted with a critical decision: How should they allocate this new manpower to resolve and prevent as much crime and disorder as possible, while also maximally helping the people in need of police attention and maintaining optimal police legitimacy?

The options before them: They could invest in more proactive policing, for example by increasing police patrol in crime hot spots aiming to prevent future crime and disorder. Alternatively, they could strengthen reactive policing, enhancing the force's ability to respond rapidly to crimes in-progress or improving investigative capacity to solve crimes after they occur, and in the process possibly also prevent future crime in the process.

The chiefs are not interested in ideological debates or untested theories; they simply want an evidence-based assessment of the marginal benefits of each option. How much additional deterrence would an extra officer assigned to proactive patrol yield? How many more crimes would be solved for every extra detective? And what would solving more crimes through investigative efforts contribute to crime prevention? What combination of strategies would optimise public safety while maintaining police legitimacy?

After decades of policing research, the field should be able to provide empirically motivated insights into the comparative effectiveness of all available policing tasks and strategies in both resolving and deterring crime and maintaining citizen satisfaction and police legitimacy.

Yet, when the chiefs arrive at the ivory towers, it becomes evident that these questions cannot be answered with the existing evidence base. Virtually all systematic reviews and meta-analyses in policing research focus on proactive strategies. These studies rigorously test how proactive strategies perform compared to business-as-usual, but business-as-usual itself is consistently poorly defined and examined. Moreover, there is no comparable body of research innovating or systematically evaluating the effectiveness of reactive policing strategies – such as rapid response or investigative follow-up – on either crime prevention or crime resolution. To the chiefs, this absence of systematic research into reactive policing strategies is particularly striking given that reactive policing constitutes the majority of what their departments actually do.

This thesis aims to bridge the gap between academic research and operational decision-making by (re-)evaluating reactive policing and contributing to a more comprehensive and practical evidence base for policing.

OUTLINE OF THIS INTRODUCTORY CHAPTER

As illustrated in the tale of the police chiefs, this thesis starts from the simple observation that there is a significant imbalance in research being done into the various tasks and strategies the police carry out on a daily basis. While police departments spend around half of their time on *reactive policing* (e.g., *rapid response* and *reactive investigation*) strategies (Lum, Koper, Wu et al., 2020), virtually no (quasi-)experimental research has examined the effectiveness of reactive strategies in resolving or preventing crime and disorder. In contrast, all of the dominant *proactive policing strategies* – i.e., *hot spot policing* (Braga, Turchan, Papachristos & Hureau, 2019), *problem-oriented policing* (Hinkle et al., 2020), *community policing* (Gill et al., 2014), *focussed deterrence* (Braga, Weisburd & Turchan, 2019), *stop-and-frisk* (Petersen et al., 2023), *broken-windows policing* (Braga et al., 2024) – have been systematically reviewed for these effects using meta-analysis and yield, on average, 38 (quasi-)experimental studies per strategy.

This discrepancy suggests that contemporary policing research exists within what I term a Proactive Policing Paradigm (PPP) *Proactive Policing Paradigm*, wherein extensive empirical scrutiny has been directed at proactive strategies, while reactive policing remains an empirical blind spot. The chapters in this thesis, in one way or another, aim to contribute towards remedying this blind spot.

However, before outlining the rest of this thesis, this chapter starts with a broad introduction to policing – beginning with a broad overview of theoretical perspectives on the role of the police in society. Tracing the evolution of these ideas from political philosophers of the Enlightenment to 20th-century sociologists and criminologists, I show how our understanding of what policing entails became increasingly broad and ambiguous. In response to this, I propose a simple, confined, and pragmatic perspective on policing. I then outline the primary aims of the police from this pragmatic perspective, examining both the policing strategies employed to achieve these aims and the mechanisms through which these strategies would yield their effects. Finally, I introduce the nine studies in this thesis, and how each contributes to addressing the historical oversight of reactive policing research.

A note to the reader: As this thesis is written in the Netherlands, working primarily with data from the Netherlands Police, special attention is given throughout this thesis to this context.

THE ROLE OF POLICE IN SOCIETY

Much ink has been spilt on the role of the police in society, and yet the topic remains a source of considerable conceptual fuzziness. In my reading of this literature, a large part of this ambiguity stems from something analogous to David Hume's (1739) *is/ought* distinction: historically, political philosophers during the Enlightenment begin with a theoretical perspective, defining what law enforcement ought to do, while more recent scholars take an empirical approach, examining what law enforcement *actually does*. This divide leads to contradictions, as empirical researchers highlight discrepancies between idealised models and the realities of law enforcement, adding caveats to normative conceptualisations of law enforcement in society. To illustrate this, I outline a brief historical evolution of thought on policing through the contributions of influential early thinkers.

Thomas Hobbes (1651) in *Leviathan* provided one of the earliest justifications for state

enforcement, arguing that without a powerful authority to impose order, society would descend into a chaotic 'war of all against all'. His vision of the social contract positioned the state as the ultimate enforcer of laws, ensuring compliance through coercion when necessary. Yet, this absolute authority provoked concerns about unchecked power, leading to John Locke's (1689) more restrained view in *Two Treatises of Government*. While he accepted that government was necessary to maintain order, Locke insisted that its primary role was to protect individual rights – life, liberty, and property – rather than exert dominance over its citizens. His scepticism of absolute rule laid the foundation for a more limited and accountable state apparatus.

Building on these ideas, Charles de Montesquieu (1748) in *The Spirit of the Laws* introduced the separation of powers, arguing that law-making, enforcement, and adjudication should be divided (i.e., *trias politica*) to prevent tyranny. By placing policing within the executive branch, he formalised its role as distinct from legislative and judicial functions, reinforcing that law enforcement should operate within legal constraints rather than outside of them. However, as states expanded their power, Jean-Jacques Rousseau in *The Social Contract* (1762) voiced concerns that even a legally bound executive could become oppressive. He warned that law enforcement, if detached from the people's will, could serve the interests of the powerful rather than the public. His work reflected growing anxieties in 18th-century Europe about the potential for state-led oppression, particularly in policing dissent and controlling populations rather than protecting them.

Cesare Beccaria's *On Crimes and Punishments* (1764) laid the foundation for deterrence theory by arguing that punishment should be swift, certain, and proportionate to the crime. Beccaria famously asserted that the deterrent effect of punishment arises primarily from its certainty rather than its severity, challenging the prevailing belief that harsh penalties were necessary for crime control. His work marked a departure from retributive justice, advocating instead for a rational legal system designed to prevent crime through predictable and measured enforcement.

Jeremy Bentham's *An Introduction to the Principles of Morals and Legislation* (1789) expanded on Beccaria's ideas by embedding deterrence within a broader utilitarian framework. Bentham argued that crime could be controlled through a calculated system of rewards and punishments that maximised overall societal well-being. This work laid the groundwork for what would later be known as rational choice theory in criminology (Cornish & Clarke, 1986). His concept of the *Panopticon* (1789) extended this logic by suggesting that continuous surveillance, or even the perception of being watched, could serve as an effective deterrent by increasing the perceived certainty of punishment. Unlike Rousseau, who feared state overreach, Bentham viewed law enforcement as a necessary mechanism to regulate behaviour rationally. His ideas laid the groundwork for modern debates on proactive policing, deterrence-based policies, and the ethical implications of state surveillance.

By the early 20th century, Max Weber in *Politics as a Vocation* (1919) formalised the idea that the state maintains order through its monopoly on legitimate violence. His framework positioned the police as the principal institution through which states exercise this monopoly, distinguishing them from other forms of coercion by emphasizing their legitimacy and legal grounding. While Weber helped define policing as a core function of modern governance, sociologists like Michael Banton (1964) and Albert J. Reiss Jr. (1971) refined this further by linking policing to its social function rather than just its coercive capacity. Banton emphasised

that policing was not merely about enforcing laws but also about managing conflicts and maintaining order in ways that reflect societal values. Reiss Jr. built on this by situating policing as a broader mechanism of social control, highlighting the role of discretion and institutional authority. He framed the police as key actors in managing interactions between the state and the public.

Over time, the evolving actions of police officers in the field began to redefine what policing itself was expected to be. Banton and Reiss Jr. were among the first to expand the conceptualisation of policing by observing officers in practice, shifting the focus from law enforcement to broader social regulation. The influential sociologist Egon Bittner pushed this further by studying extreme policing contexts, such as on Skid Row, where officers managed marginalised populations outside conventional criminal justice roles. His now-famous definition of police work as handling ‘something-that-ought-not-to-be-happening-and-about-which-somebody-had-better-do-something-now’ (Bittner, 1990, p. 249) encapsulated the fluid and situational nature of their role. While these insights deepened our understanding of policing, they also fuelled mission creep – the gradual expansion of police responsibilities beyond their original scope. What police did in exceptional cases increasingly shaped expectations of what they should do, leading to a vision of policing as an all-purpose solution for a growing range of social problems. This has resulted in descriptions of policing so expansive that some now characterise it as the institution responsible for ‘unf@cking people’s problems’ (Huey & Johnston, 2023).

THE ROLE OF THE POLICE IN THE NETHERLANDS

This mission creep can be clearly discerned within the Dutch policing context. In the past year, several high-ranking Dutch officials have expressed concerns about this expanding role of the police. Each of these officials referred to the police as a *‘duizenddingendoekje’* – a Dutch term implying that the police have become a jack-of-all-trades, tasked with handling a growing array of social issues beyond their core duties.

The former chief of the Netherlands Police, stated that ‘political problems are threatening to become police problems’, as the force is increasingly used to fill gaps left by other institutions (Regioburgemeesters.nl, 2024). Similarly, the minister of justice has pointed out that while police officers are often the first point of contact in crisis situations, they are not always the most appropriate responders, particularly in cases involving mental health (van Gemert, 2025). Lastly, the chairman of the Public Prosecution Office has also raised this issue, stating that the police are being burdened with too many non-criminal responsibilities, leading to backlogs in criminal investigations and leaving citizens without proper recourse when reporting crimes (Binnenlandsbestuur.nl, 2024).

Recognizing the risks of this mission creep, the Netherlands Police have recently outlined a shift in their strategic agenda, aiming to move away from being a *duizenddingendoekje* and refocus on ‘core policing tasks’ (Nationale Politie, 2025). However, despite this stated intention, the Strategic Agenda 2025-2030 provides no concrete definition of what these core tasks actually entail. Instead, it again describes the core function of the police in broad, abstract terms: ‘The core function of the police is “to be present” in order to act when necessary. We are present in society 24 hours a day, 7 days a week to set boundaries, to protect, and to empower.’ (Nationale Politie, 2025, p. 14)

This lack of specificity is, at least in part, a consequence of the legal framework within

which the police operate. Article 3 of the Police Act prescribes the role of the police as follows: ‘The police are tasked, in subordination to the competent authority and in accordance with applicable legal provisions, with ensuring the effective enforcement of the legal order and providing assistance to those in need.’ (Rijksoverheid, 2012)

It is the inclusion of the phrase ‘providing assistance to those in need’ that has opened the door to expansive and loosely defined interpretations of police responsibilities. While, in practice, the police are not held accountable for fulfilling this role in its broadest sense: There are countless instances of people in need of assistance, suffering, or even dying in the Netherlands where the police play no role whatsoever – with other institutions such as public healthcare, religious organisations, and civil society stepping in instead.

I argue that this broad and ambiguous mission hinders the formulation of objective and valid measures, which are essential for accountability, evidence-based evaluation of police effectiveness and innovation.

A PRAGMATIC APPROACH TO POLICING

In response to this ambiguity, this thesis proposes a simple, confined and pragmatic perspective on the role of the police – one that is neither theoretical nor empirical in nature, but instead grounded in practical necessity and societal expectations. From this perspective, the police should be understood as the primary institution responsible for: (1) resolving transpired crime and disorder, (2) preventing future crime and disorder, and (3) doing so with the highest possible level of public approval, cooperation and legitimacy.

This approach does not predefine police tasks based on abstract principles but instead asks: What (combination of) strategies and activities most effectively and efficiently achieve these goals? Whatever best serves crime control, public order, public approval, cooperation, and police legitimacy should define the role of the police.

It is true, as is stated in Article 3 of the Police Act, that police sometimes perform tasks unrelated to the resolution and prevention of crime and disorder, such as public assistance (e.g., rescuing the proverbial cat from a tree). The position of the pragmatic approach to policing on this is straightforward: if performing the task aligns with public expectations at that moment, does not interfere with more urgent policing responsibilities, and failing to complete the task would undermine police legitimacy, the task at hand should be carried out. However, the police might simultaneously work to alter this public expectation.

A useful example illustrates this principle: In the Netherlands, it has long been standard practice that when a (minor) car accident occurred, the police were called to oversee the situation and aid in filling in the insurance forms. The public felt a need for police attention in this situation, and the police provided. However, recognising that such incidents did not require police intervention, authorities gradually implemented an informational campaign advising the public to call only in cases of clear disorder or danger. The pragmatic approach supports – and even prescribes – this sequence: initially meeting public expectations where feasible to uphold legitimacy and citizen satisfaction, while simultaneously working to reshape them to better align with policing priorities.

THE STANDARD MODEL OF POLICING AND PROACTIVE AND REACTIVE POLICING

So, based on the pragmatic approach, police should primarily endeavour to resolve and prevent crime and disorder, but how? The way the police have historically sought to prevent and resolve crime and provide public assistance is captured in the *Standard Model of Policing* (Frydl & Skogan, 2004), which is described to consist of three primary policing activities – rapid response, reactive investigation and random patrol. These three activities can be distributed over two distinct modes of policing: reactive and proactive policing (Reiss, 1971; Reiss & Bordua, 1964).

Reactive policing is generally understood as police actions taken *in response to* crime and disorder incidents and citizen requests for service more generally. It consists of two core tasks: (1) rapid response to incidents, where officers are dispatched to emergencies or reported crimes, and (2) reactive investigations, which involve gathering evidence and identifying offenders after an incident has occurred. The primary goal of reactive policing is to *resolve* crime and *restore* order by addressing incidents after they have happened, doing so through enforcement – such as arrests, fines, or other legal actions – and conflict resolution, which includes mediation, de-escalation, and referrals to appropriate services. Moreover, reactive policing can have the added benefit of *preventing* future crime and disorder by physically constraining perpetrators, reinforcing legal norms and demonstrating the consequences of criminal behaviour to both the perpetrator and the wider public.

Proactive policing, in contrast, is generally understood as police-initiated activities aiming to *prevent* crime and disorder. Officers engage in tactics such as surveillance, in the form of (random) patrol, or through more novel police strategies such as hot spot policing.

The underlying mechanism through which policing achieves this preventive effect is referred to as deterrence.

DETERRENCE THEORY

Deterrence theory posits that individuals refrain from committing crimes when they perceive that the costs of offending outweigh the potential benefits. As previously discussed, the foundational elements of this theory – certainty, severity, and celerity of punishment – were first articulated by Beccaria in 1764. Bentham (1789) later expanded on these ideas by embedding them in a broader utilitarian framework, arguing that individuals make rational decisions by weighing the expected costs and benefits of their actions.

After two centuries of scholarly disinterest, deterrence theory experienced a significant revival in the mid-20th century, largely driven by the work of economist Gary Becker (1968) and sociologist Jack P. Gibbs (1968). Becker's economic model of crime mirrored Bentham's utilitarian logic, treating offenders as rational actors who weigh the expected costs and benefits of both crime and lawful alternatives. Like Beccaria, Becker argued that among the costs of crime, certainty was more important than severity in shaping offender decision-making (Becker, 1968, p. 178). Gibbs, in contrast, focused on empirical testing, systematically examining the relationship between punishment certainty, severity, and crime rates. His work provided the first statistical evidence supporting the deterrent effect of legal sanctions, offering a crucial empirical foundation for modern deterrence theory.

Contemporary deterrence theory generally understands deterrence as acting through three distinct mechanisms: incapacitation, specific deterrence, and general deterrence (Nagin, 2013, p. 200). Incapacitation reduces crime by physically restricting offenders, typically through imprisonment, preventing further offences while they are confined. Specific deterrence discourages reoffending by increasing an offender's perceived cost of crime through first-hand *experience* of punishment, shaping its views on certainty, severity, and celerity. General deterrence discourages potential offenders through the perceived awareness and thus the threat of (the certainty, severity and celerity of) punishment.

However, incapacitation is sometimes argued to be apart from deterrence, as the nature of these mechanisms is quite distinct. While incapacitation can be thought of as an exclusion mechanism, general and specific deterrence are best thought of as an update mechanism: allowing individuals to recalibrate their perceived cost of crime (in terms of certainty, severity and celerity) based on direct experience or awareness of enforcement actions.

It should be noted that not all policing strategies engage all three deterrence mechanisms. Proactive policing strategies, by definition, do not result in punishment. As such, they generally do not facilitate an update of the severity and celerity of punishment. Instead they increase the perceived cost of crime/disorder through an increased sense of the *certainty* of punishment.

THE PRAGMATIC POLICING MODEL

To synthesise the key points discussed, I have constructed a preliminary model of the pragmatic approach to policing in Figure 1. The figure systematically outlines the core functions of policing as described in the pragmatic approach to policing: to deter and resolve crime and disorder, while maintaining legitimacy. Here, all police activities not related to crime and disorder (Cardiopulmonary Resuscitation (CPR) response, public assistance, etc.) are understood as functioning to uphold its legitimacy in the eyes of the public.

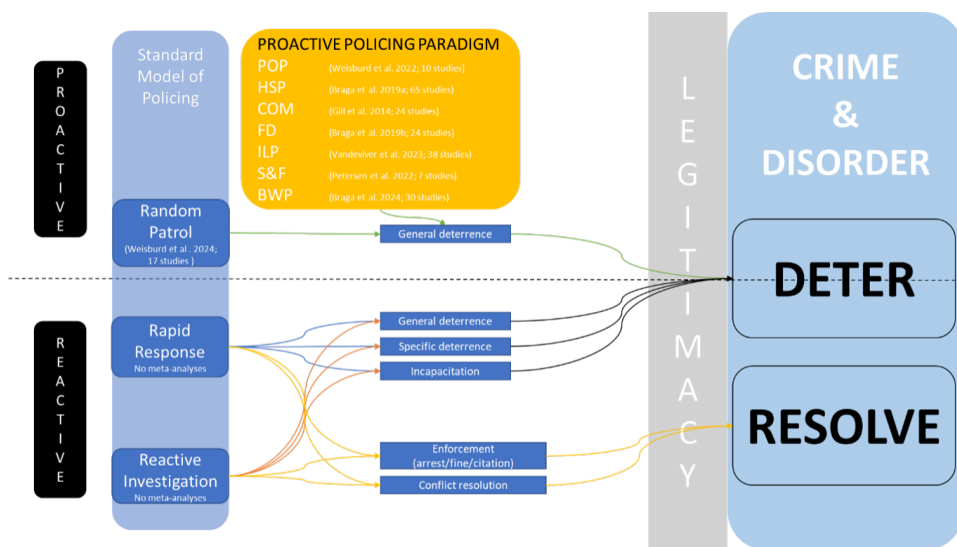


Figure 1: A model of the pragmatic approach to policing

The figure also details the three mechanisms of deterrence – incapacitation, specific deterrence, and general deterrence – which have historically been captured in three core policing activities: rapid response, reactive investigation, and random patrol – commonly referred to as the Standard Model of Policing.

As illustrated by the causal arrows, the two reactive activities (rapid response and reactive investigation) contribute to incapacitation, specific deterrence, and general deterrence. In contrast, the proactive activity – random patrol and the subsequent proactive policing strategies – operate through general deterrence, specifically by increasing the perceived *certainty* of punishment among the public. The figure also highlights two mechanisms through which reactive policing strategies contribute to crime resolution: enforcement actions (e.g., arrests, fines, citations) and conflict resolution.

Beyond illustrating the pragmatic approach to policing, Figure 1 also serves a secondary purpose: to detail the evolution from the mainly reactive Standard Model of Policing into the new Proactive Policing Paradigm by highlighting the uneven growth of the empirical evidence base – represented through systematic reviews – for both reactive and proactive policing strategies.

IMBALANCE IN THE EVIDENCE BASE DETAILED

Over the past five decades, research on proactive policing has expanded considerably, with all major proactive strategies undergoing rigorous (quasi-)experimental study and systematic review. In the latest iterations of these reviews, we observe: 17 (quasi-)experimental studies on the general deterrent effect of *random patrol* (D. Weisburd et al., 2024), 39 on *problem-oriented policing* (Hinkle et al., 2020), 65 on *hot spot policing* (Braga, Turchan, Papachristos & Hureau, 2019), 24 on *community policing* (Gill et al., 2014), 24 on *focussed deterrence* (Braga & Dusseault, 2018), 40 on *stop-and-frisk* (Petersen et al., 2023), and 56 on *broken-windows policing* (Braga et al., 2024) – with the vast majority of these studies taking place in the US and, to a lesser extent, in the UK.

This substantial evidence base for proactive strategies stands in stark contrast to the total absence of systematic reviews on the crime deterrent and crime resolution effects of reactive policing strategies, despite the fact that reactive strategies – namely, rapid response and reactive investigations – remain the primary activities of most police agencies (D. Weisburd & Majumdar, 2018).

Without comprehensive literature searches which are part of systematic reviews, it is difficult to know the state of the literature. My own targeted literature search yielded only a few (quasi-)experimental studies directly testing the crime resolution effects of reactive investigation (i.e., Braga & Dusseault, 2018; Braga, Turchan & Barao, 2019; Roman et al., 2009) and one on the crime prevention effects of reactive investigation (Fox & Farrington, 2015). Moreover, it yielded no (quasi-)experimental studies directly testing the crime resolution or crime prevention effects of rapid response.

Nevertheless, a partial evidence base for the deterrent effects of reactive strategies might be constructed based on the effects of the mechanisms through which they operate: incapacitation, specific deterrence and general deterrence.

Most studies on incapacitation focus on the effects of incarceration (e.g., Piquero & Blumstein, 2007) with some focussing on the incapacitative effects of alternative sanctions such as alcohol locks (see Elder et al., 2011) or electronic monitoring (see Di Tella et al.,

2009). However, there is also incapacitation that occurs fully on the police side of the judicial spectrum such as the sober-up cell/drunken tank on which little to no empirical (experimental) research is done. As such, the current evidence base provides only partial insights into the incapacitation effect of (reactive) policing.

Much of the literature on specific deterrence also focuses on the penitentiary – i.e., sentencing and incarceration (see Loeffler & Nagin, 2021) – rather than the police side of the judicial spectrum. Research on the specific deterrent effects of police arrest exists but remains narrow. There is quite extensive literature on the deterrent effects of arrest for juvenile offences (e.g., Klein, 1986); domestic violence arrests (e.g., Berk et al., 1992; Sherman & Berk, 1984); and DUI offences (e.g., Tay, 2005; Weatherburn & Moffatt, 2011). There are also some other studies looking into the specific deterrent effect in other crime types, but these generally lack replication, for example for prostitution-related arrests (Brewer et al., 2006). A general observation is that the literature assumes specific deterrence occurs only when an arrest is made. For example, Sherman et al. (1997, Chapter 8, p. 11) reasons that there is ‘strong but indirect’ evidence that rapid response will not reduce crime, *because* it has very limited effects on arrest rates. However, it is plausible that (the stress of) near-arrest is experienced as a form of (psychological and immediate) punishment and could also deter crime: A suspect who narrowly avoids capture from a rapid police response may be deterred from reoffending. Similarly, an ongoing reactive investigation might cause the suspect to ‘lay low’, (temporarily) deterring him from committing crime.

The literature on the general deterrent effect of policing is extensive but bifurcated, with separate strands examining deterrence resulting from reactive and proactive strategies. Research on the general deterrent effect of reactive policing primarily consists of correlational studies exploring the relationship between arrest rates and crime rates (see Sherman et al., 1997, Chapter 8, p. 17). These studies offer only weak evidence for a general deterrent effect. In fact, these studies are generally unable to disentangle the effects of incapacitation, specific deterrence and general deterrence. A more rigorous approach was taken by Ariel et al. (2019), who experimentally tested a carrot-and-stick conversation intervention on prolific offenders and assessed both its specific deterrent effect on the targeted offender and its general deterrent effect on the offender’s direct criminal network, as identified through police records. However, this line of inquiry has seen limited replication to date.

In contrast, research on the general deterrent effect of proactive policing is dominated by systematic reviews and meta-analyses of (quasi-)experimental intervention studies. While these studies do not appear, at face value, to shed light on the general deterrent effects of reactive policing, this distinction may be misleading. To the best of my knowledge, none of these reviews attempt to disentangle the general deterrent effects of the proactive components of an intervention from those of its reactive components. In practice, an increased proactive police presence in a neighbourhood is likely to enhance not only visibility and preventive action but also reactive capacity – for example, by enabling quicker response times and improving contextual understanding of incidents. It is therefore reasonable to assume that the reactive dimension of proactive interventions contributes, at least in part, to their observed general deterrent effects.

OUTLINE OF THIS THESIS

As stated throughout this introduction, this thesis starts from the observation that there is a significant imbalance in research being done into the various strategies and underlying tasks the police carry out on a daily basis. In reality, this was an insight gained during the analysis of Chapter 1 (a scoping review) of this thesis.

Originally, the first part of this thesis asked: *What are the effects and determinants of rapid response?* (Verlaan & Ruiter, 2023a). The initial search yielded two findings: (1) that contemporary research on the effects of rapid response is largely absent, and (2) that studying both the effects and the determinants of police response time in a single analysis would be suboptimal. In response, this part reports on a subsequent scoping review focusing specifically on the determinants of response time and then poses two additional questions: *Is the dismissal of rapid response as ineffective in resolving (and thus) preventing crime justified? And how did the current lack of research on the effects of rapid response come about?* It finds the (contemporary) dismissal to not be justified based on the empirical data. Moreover, this part identifies the Maryland Report as the main driver of what I have termed the *Proactive Policing Paradigm*. It concludes that renewed research into the effects of rapid response is both necessary and overdue.

The second part of this thesis addresses an assertion in the Maryland Report suggesting that even if rapid response was effective at resolving and thereby preventing crime, the police would have limited means to improve response times. As Sherman puts it: ‘Cutting police travel time for such crime from 5 to 2.5 minutes could require a doubling of the police force’ (1997, Chapter 8, p. 11). In response, this second part asks: *What are the effects of police deployment strategy, and police capacity, urbanisation and call volume on police response time?* Through detailed and highly realistic agent-based modelling simulation experiments, it finds that police have substantial influence on improving police travel times through police deployment strategy, while police capacity, call volume, and the degree of urbanisation play a much more limited role. It concludes that renewed research into the effects of rapid response are not only necessary and overdue, but also relevant and feasible.

Part three aims to start this renewed interest in reactive policing. As a first step, it maps the current demand and supply of reactive policing on the street in terms of scale and composition. It asks: *What does reactive policing consist of in terms of scale and composition? How can it be measured appropriately? And how does it vary across jurisdictions?* Based on the first academic analysis of nationwide Calls for Service (CFS) and police dispatch data, it finds that reactive policing in the Netherlands (as well as in the city of Detroit) consists of roughly 50% crime-related and 50% disorder-related CFS. Moreover, these findings on the composition of reactive supply seem to be robust to various operationalisations of the *time-spent* measure, even less precise ones. Lastly, we show that in the Netherlands, there is substantial variation in the composition of various call categories among the various police jurisdictions. Both at the highest (*Police Regional Unit/Politie Eenheid*) as well as the lowest (*Frontline Team/Basisteam*) jurisdiction level. It concludes that, as a result, police strategies should be highly context-specific.

The fourth part of this thesis serves as a reminder that this thesis does not advocate for an exclusive focus on reactive policing but for a balanced research agenda that considers both reactive and proactive strategies. While previous chapters reassess and refine our understanding of reactive policing, this section highlights the lack of (quasi-)experimental

research on proactive policing in the Netherlands. It asks: *What are the effects of a proactive policing intervention on crime, disorder, and public perceptions in the Haarlemmerbuurt, Amsterdam, the Netherlands?* It finds that intensified police presence can reduce crime and disorder, but that these effects are highly localised. Moreover, crime and disorder appear to be displaced rather than actually being prevented. It also finds that only a small percentage of the surveyed residents noticed the increased proactive presence and that, during the intervention, public perceptions of the neighbourhood, the police and the municipality decreased. It also underscores the challenge of isolating proactive effects from reactive mechanisms. It concludes that policing research must move beyond the uncritical promotion of proactive policing and adopt a more evidence-based approach that systematically considers the role and interplay of both reactive and proactive strategies.

Lastly, it should be noted that this thesis does not (claim to) resolve the purported evidence gap. The scope of the issue is too large to be addressed in a single project. Nevertheless, the studies presented here provide a first step in re-evaluating and better understanding reactive policing. By doing so, this thesis seeks to lay the groundwork for a more comprehensive, evidence-based approach to policing.

CONTRIBUTIONS

This thesis offers several distinct contributions to the field of policing research. First, it brings a rare non-Anglophone perspective by focusing on the Netherlands – a wealthy, low-crime and low-disorder country without widespread firearm access, marked by a robust social safety net, progressive drug policies, and high institutional trust. This context contrasts sharply with the predominantly US – and UK – centred research that forms the core of the current evidence base. Second, it (in part) draws on nationwide CFS and dispatch data, allowing for the examination of both urban and rural policing, and enabling analysis of variance and comparisons across and within regional units and frontline teams. This stands in contrast to most policing studies, which mostly focus on single large, high-crime cities. Third, this research is written from an outsider's perspective. With a background in the natural sciences – rather than in criminology, law, or policing – as such, this research began from a position of unfamiliarity and healthy scepticism toward the uncritical embrace of established police practices and the proactive policing paradigm. In that sense, and as the subtitle suggests, this thesis is a *minority report*: one that challenges established assumptions and offers an alternative vision of what policing (research) is and what it might be.

GUIDING PRINCIPLES

This introduction started with a tale of police chiefs seeking guidance at the ivory towers of academia. The tale critically examines the current state of the evidence on the effects of various police activities and strategies. At the same time, it evokes the image of two separate worlds: practice and research. Although the rise of Evidence-Based Policing (EBP) has done much to bridge this divide, the gap remains. This thesis, through close collaboration, formal partnerships, and even personal connections, aims to be both guided by and useful to police practice – while maintaining health scholarly independence.

Moreover, this thesis incorporates key principles from the EBP paradigm (Sherman, 2017) such as the Maryland Scientific Methods Scale (MSMS) (Farrington et al., 2002) to assess internal validity and from the Realist Evaluation (RE) (Pawson & Tilley, 1997), such as the

Effect, Mechanism, Moderator, Implementation, and Economic cost (EMMIE) framework (Johnson et al., 2015) to ensure a holistic evaluation of empirical studies. By integrating these methodological standards, this thesis endeavours to uphold rigour, replicability, and transparency, contributing to a more robust and reliable evidence base for policing research.

This thesis also reflects a commitment to methodological breadth. It brings together a diverse set of research genres – from a scoping review and a critical essay to descriptive, methodological, simulation and quasi-experimental intervention studies. Across these, a wide range of methods is applied: from statistical modelling, data visualisation, and agent-based simulation, to ethnographic observation, semi-structured interviews, and survey research – each chosen not for comfort or methodological symmetry, but for their ability to best answer the research question at hand.

Lastly, all chapters in this thesis adhere, where applicable, to the principles of Open Science. This means that empirical studies were pre-registered on the Open Science Framework (OSF) with clearly defined hypotheses, maximum transparency regarding data use, and openly accessible, well-documented code and dedication to replicability. The (quasi-)experimental study on police presence went a step further by using synthetic data to pre-construct the R code for analysing the intervention. While minor adjustments were made, the chapter explicitly details every modification to ensure full transparency.

REFERENCES

- 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 and Criminology*, 109.
- Banton, M. (1964). *The policeman in the community*. Tavistock.
- Beccaria, C. (1764). *On Crimes and Punishments and Other Writings* (R. Bellamy, Ed.; R. Davies, V. Cox & R. Bellamy, Trans.; Vol. 1995). Cambridge University Press.
- Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2), 169–217. <https://doi.org/10.1086/259394>
- Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation* (J. H. Burns & H. L. A. Hart, Eds.; Vol. 1996). Clarendon Press.
- Berk, R. A., Campbell, A., Klap, R., & Western, B. (1992). The Deterrent Effect of Arrest in Incidents of Domestic Violence: A Bayesian Analysis of Four Field Experiments. *American Sociological Review*, 57(5), 698. <https://doi.org/10.2307/2095923>
- Binnenlandsbestuur.nl. 2024. Veranderende verhoudingen in de lokale driehoek. <https://www.binnenlandsbestuur.nl/juridisch/de-politie-een-duizenddingendoekje-geworden> Boonstra, Wouter.
- Bittner, E. (1990). *Aspects of police work*. Northeastern University Press.
- Braga, A. A., & Dusseault, D. (2018). Can Homicide Detectives Improve Homicide Clearance Rates? *Crime & Delinquency*, 64(3), 283–315. <https://doi.org/10.1177/0011128716679164>
- Braga, A. A., Schnell, C., & Welsh, B. C. (2024). Disorder policing to reduce crime: An updated systematic review and meta-analysis. *Criminology & Public Policy*, 23(3), 745–775. <https://doi.org/10.1111/1745-9133.12667>
- Braga, A. A., Turchan, B., & Barao, L. (2019). The Influence of Investigative Resources on Homicide Clearances. *Journal of Quantitative Criminology*, 35(2), 337–364. <https://doi.org/10.1007/s10940-018-9386-9>
- Braga, A. A., Turchan, B., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1046>
- Braga, A. A., Weisburd, D., & Turchan, B. (2019). Focused deterrence strategies effects on crime: A systematic review. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1051>
- Brewer, D. D., Potterat, J. J., Muth, S. Q., & Robert, J. M. (2006). A Large Specific Deterrent Effect of Arrest for Patronizing a Prostitute. *PLOS ONE*, 1(1), e60. <https://doi.org/10.1371/JOURNAL.PONE.0000060>
- Cornish, D. B., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending* (D. B. Cornish & R. V. Clarke, Eds.). Springer-Verlag.
- Di Tella, R., Diaz, F., Quintana, J., Buffa, S., Canepa, M., Castillos, J., Di Tella, R., & Schargrodsky, E. (2009). Criminal Recidivism after Prison and Electronic Monitoring. <https://doi.org/10.3386/W15602>
- Elder, R. W., Voas, R., Beirness, D., Shults, R. A., Sleet, D. A., Nichols, J. L., & Compton, R. (2011). Effectiveness of ignition interlocks for preventing alcohol-impaired driving and alcohol-related crashes: a Community Guide systematic review. *American*

- Journal of Preventive Medicine*, 40(3), 362–376. <https://doi.org/https://doi.org/10.1016/j.amepre.2010.11.012>
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In D. P. Farrington, D. C. Gottfredson, L. W. Sherman & B. C. Welsh (Eds.), *Evidence-based crime prevention* (1st ed., pp. 13–21). Routledge.
- Fox, B. H., & Farrington, D. P. (2015). An Experimental Evaluation on the Utility of Burglary Profiles Applied in Active Police Investigations. *Criminal Justice and Behavior*, 42(2), 156–175. <https://doi.org/10.1177/0093854814548446>
- Frydl, K., & Skogan, W. (Eds.). (2004, April). *Fairness and Effectiveness in Policing*. The National Academies Press. <https://doi.org/10.17226/10419>
- Gibbs, J. P. (1968). Crime, punishment, and deterrence. *The Southwestern Social Science Quarterly*.
- Gill, C., Weisburd, D., Telep, C. W., Vitter, Z., & Bennett, T. (2014). Community-oriented policing to reduce crime, disorder and fear and increase satisfaction and legitimacy among citizens: a systematic review. *Journal of Experimental Criminology*, 10(4), 399–428. <https://doi.org/10.1007/s11292-014-9210-y>
- Hinkle, J. C., Weisburd, D., Telep, C. W., & Petersen, K. (2020). Problem-oriented policing for reducing crime and disorder: An updated systematic review and meta-analysis. *Campbell Systematic Reviews*, 16(2). <https://doi.org/10.1002/cl2.1089>
- Hobbes, T. (1651). *Leviathan* (E. Curley, Ed.; Modern Edition, Vol. 2012). Hackett Publishing.
- Huey, L., & Johnston, S. (2023). Unf@cking People's Problems: A Theory of Policing. *CrimRxiv*. <https://doi.org/10.21428/CB6AB371.389F4506>
- Hume, D. (1739). *A Treatise of Human Nature* (D. F. Norton & M. J. Norton, Eds.; Vol. 2000). Oxford University Press.
- Johnson, S. D., Tilley, N., & Bowers, K. J. (2015). Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology*, 11(3), 459–473. <https://doi.org/10.1007/S11292-015-9238-7>
- Klein, M. W. (1986). Labeling Theory and Delinquency Policy: An Experimental Test. *Criminal Justice and Behavior*, 13(1), 47–79. <https://doi.org/10.1177/0093854886013001004>
- Locke, J. (1689). *Two Treatises of Government* (P. Laslett, Ed.; Vol. 1988). Cambridge University Press.
- Loeffler, C. E., & Nagin, D. S. (2021). The Impact of Incarceration on Recidivism. *Annual Review of Criminology*, 5, 133–152. <https://doi.org/10.1146/annurev-criminol-030920-112506>
- Lum, C., Koper, C. S., Wu, X., Johnson, W., & Stoltz, M. (2020). Examining the Empirical Realities of Proactive Policing Through Systematic Observations and Computer-Aided Dispatch Data. *Police Quarterly*, 23(3), 283–310. <https://doi.org/10.1177/1098611119896081>
- Montesquieu. (1748). *The Spirit of the Laws* (A. M. Cohler, B. C. Miller & H. S. Stone, Eds.; Vol. 1989). Cambridge University Press.
- Nagin, D. S. (2013). Deterrence in the Twenty-First Century. *Crime and Justice*, 42(1), 199–263. <https://doi.org/10.1086/670398>
- Nationale Politie. (2025, March). *Strategische Agenda Politie 2025-2030: Stevig Staàn in Deze Tijd* (tech. rep.).

- Pawson, R., & Tilley, N. (1997). *Realist evaluation*. SAGE Publications Ltd.
- 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). <https://doi.org/10.1002/cl2.1302>
- Piquero, A. R., & Blumstein, A. (2007). Does Incapacitation Reduce Crime? *Journal of Quantitative Criminology*, 23(4), 267–285. <https://doi.org/10.1007/s10940-007-9030-6>
- Regioburgemeesters.nl. 2024. Afscheidsinterview korpschef Henk van Essen - 'Politieke problemen dreigen steeds meer politieke problemen te worden'. <https://www.regioburgemeesters.nl/actueel/?id=901>
- Reiss, A. J. (1971). *The police and the public*. Yale University Press.
- Reiss, A. J., & Bordua, D. J. (1964). Environment and organization: A perspective on the police.
- Rijksoverheid. 2012. Politiewet, Artikel 3. *Overheid.nl*.
- Roman, J. K., Reid, S. E., Chalfin, A. J., Knight, C. R., Roman, J. K., Reid, S. E., & Chalfin, A. J. (2009). The DNA field experiment: a randomized trial of the cost-effectiveness of using DNA to solve property crimes. *Journal of Experimental Criminology* 2009 5:4, 5(4), 345–369. <https://doi.org/10.1007/S11292-009-9086-4>
- Rousseau, J.-J. (1762). *The Social Contract and Other Later Political Writings* (V. Gourevitch, Ed.; Vol. 1997). Cambridge University Press.
- Sherman, L. W. (2017). Evidence-based policing.
- Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American sociological review*, 49(2), 261–272. <https://doi.org/10.2307/2095575>
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997, January). *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress* (tech. rep.). NIJ. Washington, DC.
- Tay, R. (2005). General and Specific Deterrent Effects of Traffic Enforcement: Do we have to Catch Offenders to Reduce Crashes? *Journal of Transport Economics and Policy*, 39(2), 209–224.
- van Gemert, J. 2025. Verward gedrag van een naaste kan veel zorgen veroorzaken, ervaart ook Joke bij zoon Kees: 'Familie vaak eerste die problemen ziet'. <https://eenvandaag.avrotros.nl/item/verward-gedrag-van-een-naaste-kan-veel-zorgen-veroorzaken-ervaart-ook-joke-bij-zoon-kees-familie-vaak-eerste-die-problemen-ziet/EenVandaag>.
- Verlaan, T., & Ruiter, S. (2023a). *PROTOCOL-Determinants and effects of police response time: a scoping review*. <https://doi.org/10.17605/OSF.IO/HU2E9>
- Weatherburn, D., & Moffatt, S. (2011). The Specific Deterrent Effect of Higher Fines on Drink-Driving Offenders. *The British Journal of Criminology*, 51(5), 789–803. <https://doi.org/10.1093/BJC/AZR043>
- Weber, M. (1919). *The Vocation Lectures: Science as a Vocation; Politics as a Vocation* (D. Owen & T. B. Strong, Eds.; R. Livingstone, Trans.; Vol. 2004). Hackett Publishing.
- Weisburd, D., & Majumdar, M. K. (Eds.). (2018, February). *Proactive Policing*. National Academies Press. <https://doi.org/10.17226/24928>

- Weisburd, D., Petersen, K., Telep, C. W., & Fay, S. A. (2024). Can increasing preventive patrol in large geographic areas reduce crime? A systematic review and meta-analysis. *Criminology & Public Policy*, 23(3), 721–743. <https://doi.org/10.1111/1745-9133.12665>

LIST OF PUBLICATIONS

ACADEMIC PUBLICATIONS

1. **Verlaan, T.**, & Ruiter, S. (2023). *Predictors of Police Response Time: a Scoping Review*. *Crime Science*, 12(19), 1–23, BioMed Central Ltd. <https://doi.org/10.1186/s40163-023-00194-3>
2. Langton, S., Ruiter, S., & **Verlaan, T.** (2023). *Describing the Scale and Composition of Calls for Police Service: a Replication and Extension using Open Data*. *Police Practice and Research*, 24(5), 523–538, Taylor & Francis. <https://doi.org/10.1080/15614263.2022.2102494>
3. Langton, S., **Verlaan, T.**, & Ruiter, S. (2023). *Operationalizing Deployment Time in Police Calls for Service*. *Crime Science*, 12(20), 1–8, BioMed Central Ltd. <https://doi.org/10.1186/s40163-023-00198-z>
4. **Verlaan, T.**, Langton, S., & Ruiter, S. (2023). *Reactieve Politie-inzet op Straat*. *Tijdschrift voor de Politie*, 4, 36–40. <https://www.researchgate.net/publication/378043226>
5. **Verlaan, T.**, & Langton, S. (2024). *On the use of Inferential Statistics on Administrative Police Data*. In *The Crime Data Handbook*, 197–210. Bristol University Press. <https://doi.org/10.51952/9781529232073.ch014>
6. **Verlaan, T.**, & Ruiter, S. (2025). *Evaluating Place-Based Policing in Amsterdam: A Quasi-Experimental Study of Calls-for-Service Levels and Public Perceptions using a Micro-Synthetic Control Framework*. *Journal of Experimental Criminology*, 21(3), xxx-xxx, Springer Nature. <https://doi.org/10.1007/s11292-025-09678-5>

MANUSCRIPTS UNDER REVIEW

7. **Verlaan, T.**, Birks, D., & van der Mei, R.D. (2025). *What is the Effect of Police Deployment Strategy on Emergency Response Times? An Agent-Based Modelling Investigation*.
8. **Verlaan, T.** (2025). *On Re-evaluating Reactive Policing*.
9. **Verlaan, T.** (2025). *The Maryland Report: Constructing the Proactive Policing Paradigm*

POPULAR PUBLICATIONS

10. **Verlaan, T.** (2025). *Meer blauw op straat? Dat helpt nauwelijks om criminaliteit terug te dringen*. *de Volkskrant*. [Op-Ed]. <https://www.volkskrant.nl/columns-opinie/meer-blauw-op-straat-dat-helpt-nauwelijks-om-criminaliteit-terug-te-dringen~bb4891d5/>
11. Hamelink, R., Spruijt, M., & **Verlaan, T.**, van Jole, L. (2025). *Blaauw op Straat. Het Haarlemmerbuurt Experiment*. [Documentary]. <https://www.at5.nl/nieuws/e32383/>

I

REVISITING REACTIVE POLICING

1

PREDICTORS OF POLICE RESPONSE TIME

A SCOPING REVIEW

Background As rapid response has been a key policing strategy for police departments around the globe, so has police response time been a key performance indicator. This scoping review maps and assesses the variables that predict police response time.

Methods This review considers empirical studies, written in english, that include quantitative data from which an association between the outcome variable police response time and any predictor can be observed or derived. This review provides both a narrative synthesis as well as what we termed a hybrid synthesis, a novel way of synthesizing a large quantitative dataset which is considered too rich for a mere narrative synthesis and yet does not allow for meta-analysis.

Results The search, screening and selection process yielded 39 studies, which presented 630 associations between 122 unique predictor variables and police response time. In order to present the results in a digestible way, we classified these into categories and subcategories. All methodological steps and the findings are made public: <https://github.com/timverlaan/prt>.

Conclusions Most of the conclusion and discussion focuses on lessons learned and recommendations for future research, as it proved hard to draw any definitive conclusions on causal factors related to police response time. We recommend that future studies clearly describe mechanisms, focus on the components of police response time (reporting time, dispatch time, travel time – or a combination of these), attempt to standardise predictors and outcome variables, and we call for more research into reporting time. We conclude this review with a first attempt at deriving a causal model of police response time from the subcategories of predictor variables we observed in the empirical studies included in this review.

1

1.1 INTRODUCTION

For decades, *rapid response* has been a key strategy for police departments around the globe. There are two general reasons for this. First, to fulfill the police's lifeline function, to be physically present when citizens are in direct need of help. Second, because faster response times to citizens' calls-for-service are commonly believed (and empirically suggested) to be associated with increased *witness availability* (e.g., Cordner et al., 1983), *on-scene arrests* (e.g., Cihan et al., 2012), *crime clearance rates* (e.g., Blanes i Vidal & Kirchmaier, 2018), *crime deterrence* (e.g., S. Weisburd, 2021), *victim satisfaction* (e.g., Brandl & Horvath, 1991), *victim wellbeing* (e.g., B. B. Brown & Harris, 1989), and overall *citizen satisfaction* (e.g., Frank et al., 2005). As it is also an easy to track quantitative measure, *police response time* is one of the key performance indicators in policing. Police departments around the globe measure, assess, and publicly share their response times and aim to improve them. Therefore, knowledge of the factors that determine police response time is of great value to police in particular and society at large. Although there is a large body of research on predictors⁹ of police response time, so far, there has been no research that gathered, mapped and synthesised the findings. This review aims to fill this gap by answering the following research questions:

1. What are the empirically tested predictor variables of police response time?
2. What conclusions can be drawn about the direction, size and statistical significance of the effects of these predictor variables on police response time?

The type of literature review carried out here is a scoping review, which is defined as 'a form of knowledge synthesis that addresses an exploratory research question aimed at mapping key concepts, types of evidence, and gaps in research related to a defined area or field by systematically searching, selecting, and synthesizing existing knowledge.' (Colquhoun et al., 2014). We used the PRISMA-ScR checklist (Tricco et al., 2018) and first developed a protocol, which we pre-registered before starting the research (<https://osf.io/hu2e9>). This review retrieves English written literature – scientific and grey – that presents quantitative, correlational data, between the outcome variable *police response time* and any empirically tested predictor variable. This study critically appraises the methodologies and evidence presented in the primary studies to assess if they support causal inference regarding the empirically tested predictor variables. We aim to quantitatively synthesise the effects of predictor variables on police response time and derive a comprehensive causal model.

1.1.1 POLICE RESPONSE TIME

In this review, *police response time* is considered to consist of three components: *reporting time*, *dispatch time* and *travel time*. *Reporting time* starts the moment an incident takes place and ends when it is notified to the police. *Dispatch time* starts when the police are notified of the incident and ends as soon as a patrol car is assigned to the call. *Dispatch time* can be subdivided into *queue time* and *handling time*. *Travel time* starts when the incident is assigned to a patrol officer or car(s) and ends when the first police officer arrives on scene. *Travel time* is determined by two factors: *travel distance* and *average travel speed*. Police response time can thus be expressed mathematically as in (1.1).

⁹This study originally set out to identify and quantify the *determinants* of police response time. However, in order to be in line with the evidence base of the findings, the term *predictors* rather than *determinants* is used.

$$t_{\text{response}} = \Delta t_{\text{incident reported}} + \Delta t_{\text{response initiated}} + \frac{\Delta x}{\bar{v}} \quad (1.1)$$

Figure 1.1 is an extension of Spelman and Brown (1981) and displays the three stages of police response time schematically. Throughout this review, we refer to these three components of police response time. Any predictor variable with a causal link to police response time has to affect police response time through one of these components. The addition of this layer of complexity to police response time is needed to better understand the mechanisms through which predictor variables affect police response time and to consequently allow for better policies and interventions regarding police response time.

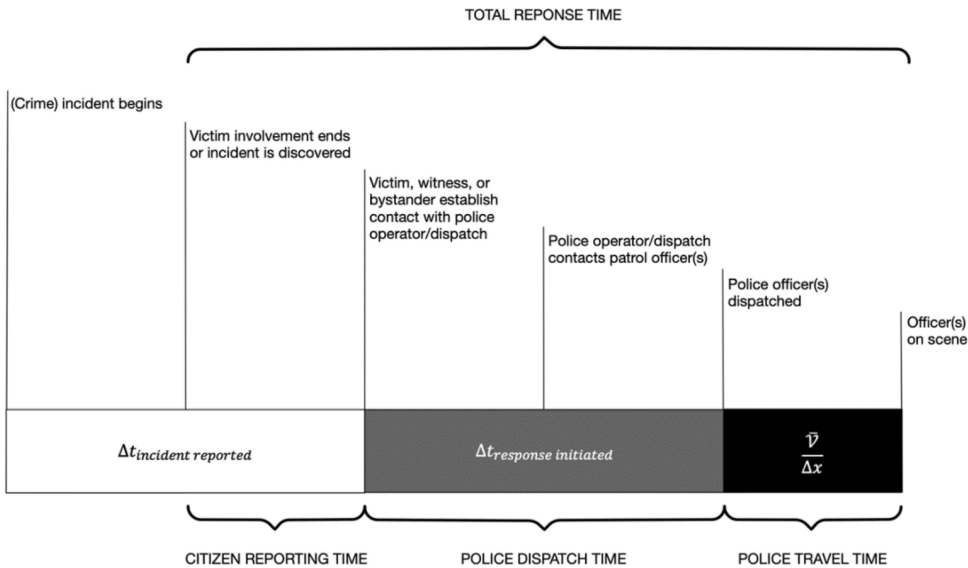


Figure 1.1: The components of police response time-adjusted from Spelman and Brown (1981)

1.2 METHODS

This methods section attempts to be as explicit as possible about the followed procedures without being excessively minute. Additional details (e.g., the exact queries used for each database) can be found in the referenced pre-registration protocol.

1.2.1 SCREENING CRITERIA

This review only considers empirical studies, written in English, that include quantitative data from which an association (whether causal or not) between the outcome variable *police response time* and any predictor variable can be observed or derived. We made no further a priori selection based on research design, literature type (academic or grey), geographical location, or year of publication; all literature until January 2021 (which is when we carried out the literature search) was considered.

1.2.2 INFORMATION SOURCES

The primary databases consulted were: Scopus, EBSCO/ Criminal Justice Abstracts, Clarivate Analytics/Web of Science Core Collection, ProQuest/International Bibliography of the Social Science (IBSS), EBSCO/APA PsycINFO, ProQuest/Criminal Justice Database, ScienceDirect, Wiley Online Library, and JSTOR. The secondary databases consulted were: Google Scholar, National Criminal Justice Reference Service (US), Global Policing Database.

1.2.3 SEARCH QUERY

After a preliminary search, we decided on the broad search query (1) below, consisting of various instances of the concepts: *police* and *response time*. Both terms were separated by the operator W/25, meaning that the terms should occur within 25 words of each other in a text. The limit of 25 words was chosen as it resembles the upper bound of the mean word count of written sentences. Moreover, it allows to retrieve studies in which 'response time' refers to 'police response time' without the words having to be in that exact sequence and without including studies which include both terms, but in which these are unlinked. After the initial search phase, it was concluded the synonyms 'rapid response' and 'prompt response' should have also been included. Hence, a second search was carried out using query (2).

1. ('police' OR 'law enforc*') W/25 ('response tim*' OR 'arrival tim*')
2. ('police' OR 'law enforc*') W/25 ('rapid response' OR 'prompt response')
3. ('police' OR 'law enforc*') W/25 ('response tim*' OR 'arrival tim*' OR 'rapid response' OR 'prompt response')

Titles and abstracts obtained from the queries of the primary databases were screened, supplemented with hand-searched articles from the secondary databases. The flow chart in Appendix A.1 visualises the entire data collection process. Table 1.1 displays the number of studies each query yielded in each primary database. Note that the Criminal Justice Database was not available to the authors during the first round of queries. For that reason, we combined the first two queries into query (3), which was used for the second round CJD retrieval.

Table 1.1: Data retrieval per database

	Scopus	CJA	WoS	IBSS	PsychINFO	CJD	SD	Wiley	JSTOR	Total
Retrieval1	330	103	64	118	23	N/A	191	426	1559	2814
Retrieval2	89	29	24	102	6	1239	1593	152	769	4003

SCREENING PHASE

The retrieved dataset of studies from the primary databases was first cleaned and enriched with the study DOIs and using the update function of the Mendeley reference manager. This step greatly increased the number of abstracts that were retrieved. The set was subsequently deduplicated using Mendeley's *duplicate removal* function. Next, the remaining unique study titles and abstracts were imported into ASReview, an active learning software package created to screen literature for systematic and/or scoping reviews (van de Schoot et al., 2021). Using ASReview, the first author (TV) coded the documents based on title and abstract as

‘maybe relevant’ or ‘irrelevant’. Documents were coded ‘maybe relevant’ when the title and abstract suggested the study concerned empirical research linking one or more predictor variables to *police response time*. All other studies were marked as ‘irrelevant’. In case the title or abstract contained insufficient information, the document was marked as ‘maybe relevant’ in this part of the screening phase.

Using a machine learning algorithm on the user inputs, ASReview constantly presents what it predicts to be the most relevant document. This way, the user does not have to classify the entire set of documents, but can stop screening after a predefined set of consecutive documents are coded as ‘irrelevant’. Any stopping criterion is arbitrary, but some are empirically shown to be sensical (e.g., Scherhag & Burgard, 2023). We settled on the stopping criterion of 100 consecutive ‘irrelevant’ articles as advised to us in an informal meeting with the developers of the ASReview program. We used the following ASReview settings: classifier: ‘Naïve Bayes’; feature extraction: ‘tf-idf’; and query strategy: ‘Max’, and we started the ASReview screening phase by providing the algorithm one relevant (Salimbene & Zhang, 2020) and one irrelevant (Zhu et al., 2020) study.

As the data retrieval was done in two rounds (adding the terms *rapid response*, *prompt response* and the CJD database in the second round), the ASReview screening phase was also done in two rounds. ASReview round one coded 111 studies (4.8%) as ‘maybe’ relevant, and 597 (26%) as ‘irrelevant’ before the stopping criterion was reached. ASReview round two coded an additional 11 studies as ‘maybe relevant’.

Additionally, Google Scholar, National Criminal Justice Reference Service, Global Policing Database were hand searched. For all the studies that were retrieved in these secondary databases a judgement was made – in the browser, based on the title/abstract – if they were ‘maybe relevant’ or ‘irrelevant’.

1.2.4 SELECTION PHASE

All documents marked as ‘maybe relevant’ during the screening phase progressed to the full-text selection phase in which eligibility was assessed. Studies were considered eligible if they concerned empirical research, had *police response time* as the outcome variable, and reported quantitative data from which an association between the outcome variable police response time and a predictor variable can be observed or derived. If a study failed on any of these criteria, it was excluded. The first violated criteria observed was noted for each discarded study. The first author (TV) did the full-text eligibility check for all literature that passed the screening test. The selection phase resulted in 37 eligible studies. Next, forward and backward searches were carried out with these 37 studies as input, which resulted in 2 additional studies. Both these additional studies were not present in the original ASReview screening. Appendix A.2 contains the list of included studies with some metadata. The ‘search overview’ file on the public GitHub makes the selection process transparent (<https://github.com/timverlaan/prt>).

1.2.5 DATA EXTRACTION

In order to extract the relevant information from the 39 studies, we created a data extraction form in Microsoft Excel. It was used to extract data specific on the outcome variable *police response time* and all observed predictor variables. Appendix A.3 presents what data were extracted. In order to increase reproducibility of the data extraction, ATLAS.ti was used to code the *research location, time period, sampling strategy*, the presented *correlational & descriptive data* and the conclusion in the primary studies. The data extraction form is available on the GitHub page.

1.2.6 DATA CONVERSION AND INTERPRETATION

PREDICTOR VARIABLE CATEGORIES

Because of the large number of unique predictor variables identified in the extracted data, we started with a data reduction step by classifying all predictor variables in categories and subcategories. This data reduction step helps with presenting the results in a digestible way. Table 1.2 specifies what additional information was coded per variable observed in the primary studies.

Table 1.2: Variables added to the extracted information

Code name	Explanation
IV_label	Generalisation of independent variable name
variable category	<i>Kinematics, exogenous, human involvement, incident characteristics, innovation, police strategies</i>
variable subcategory	<i>Day of the week, demand, road type, time of day, traffic, weather type, caller characteristics, community characteristics, victim characteristics, witness characteristics, incident characteristics, incident type, AGDS, AVL, emergency button, response unit type, travel distance, travel speed, call priority, department size, differential response system, jurisdiction area, no. officers per response unit, non-emergency number, on duty officers, patrol strategy, shift, workload</i>
Regression type	<i>Simple or multiple</i>
Effect direction	<i>+/-/none (for nominal and dichotomous predictor variables always none)</i>
Effect size interpretation	<i>Very small, small, medium, large (see Appendix A.4)</i>
Statistical significance	<i>*/**/***/non-applicable (i.e. population data)</i>
Internal validity	<i>Level I to V (see section Critical appraisal of research methodologies)</i>

STRENGTH OF THE ASSOCIATIONS

Some (models in the) primary studies did not directly present associations, or presented it in coefficients that do not directly allow for effect size calculation. In the first case, we attempted to extract descriptive statistics that allowed us to derive the size of the association ourselves. We converted multiple observed group means of population studies to percentage difference/change and sample means & standard deviations, t-statistics and z-statistics into Cohen's D where possible. In the second case, we converted the primary study's metric into a metric that allowed for effect size interpretation. We converted unstandardised B coefficients into a new metric B^* by dividing it by the mean *response time* of the *calls for service* to make it analogous to the percentage change metric. If no mean was presented, we converted the unstandardised B coefficients into standardised Beta coefficients by extracting the standard deviation of both the outcome and the predictor variables. For the one study that presented results of an ordinal logit regression (Kayes et al., 2019), we converted the unstandardised

B coefficients into odds ratios. In order to interpret the observed and converted Cohen's D, Pearson's R and odds ratios, we relied on Cohen's (2013) interpretation of effect sizes. In order to interpret the effect size of percentage difference/change measures, we posit our own interpretation and scale the interpretation of the unstandardised B* coefficient and the standardised beta coefficient analogous to this. We propose the following interpretation of the effect size of percentual change we believe to be reasonable for the context of police response time: 0–10% change, *very small*; 10–30% change, *small*; 30–50%, *medium*; > 50%, *large*. All these data conversion and interpretation steps are detailed in Appendix A.4.

DIRECTION OF THE ASSOCIATIONS

In our synthesis of the results, we discuss the direction of the association between scale/ordinal predictor variables and the outcome variable police response time. By a *positive* relation we mean that a unit/level increase of the scale/ordinal predictor variable corresponds with an increase in *police response time*, and thus police response is slower. By a *negative* relation we mean that a unit/level increase of the scale/ordinal predictor variable, *police response time* decreases, and thus police response is faster. Interpreting *direction* of associations for nominal and dichotomous predictor variables is often meaningless. For example, different *crime types* (*property crime*, *violent crime*, etc.) might correspond with different police response times, but there is no logical direction of the association. For this reason, we only report whether there is very small/small/medium/large systematic variation in police response time between categories of those nominal variables, but more details can be found in the data extraction file published on GitHub. For some dichotomous variables though, especially those that capture pre/ post interventions, the direction of the association has an intuitive interpretation. In such cases, this will be elaborated on.

1.2.7 CRITICAL APPRAISAL OF RESEARCH METHODOLOGIES

In order to judge what conclusions can be credibly drawn regarding the presented associations between the outcome variable *police response time* and the predictors, the retrieved literature was critically appraised based on research design. We classified these using the Maryland Scientific Methods Scale (MSMS) (Farrington et al., 2002), as presented in Table 1.3. The higher the level of the research design, the more trust can be placed in the internal validity of the observed associations. Strong causal claims are only warranted for research level V designs.

Table 1.3: Maryland Scientific Methods Scale interpreted to suit the studies under review

Internal validity	Description
Level I	Correlation analysis, no control group nor attempt to establish counterfactual
Level II	'Before & after comparison' or comparison group without balancing covariates
Level III	Difference in Difference, attempt at balancing (OLS, matching) but uncontrolled differences likely remain
Level IV	Explicitly reasoned and proper balancing (OLS, matching), biases discussed but not addressed
Level V	RCTs, no selective sample attrition

1.3 RESULTS

1.3.1 EMPIRICALLY TESTED PREDICTORS OF POLICE RESPONSE TIME

Together, the 39 included studies yielded 630 empirically tested associations between predictor variables and (a component of) police response time. However, over two-thirds of these predictors originate from the three studies by Bieck (1978, 1980a, 1980b), which are based on relatively small sample sizes. The 630 associations relate to 122 unique predictor variables. Because it is impossible to present and discuss all these unique predictors in this article, we will present some descriptive statistics and refer the interested reader to the data extraction file available on GitHub. Of the 630 tested associations, 18% ($n = 113$) related to scale predictor variables, 6% ($n = 40$) to ordinal predictor variables, and 76% ($n = 477$) to nominal predictor variables, the latter of which predominately represented *incident type*. The outcome variable *police response time* was operationalised in the following ways: 142 associations were tested for the component *reporting time* (23%), 81 for *dispatch time* (13%), 159 for *travel time* (25%), 237 for *dispatch time + travel time* (38%), and 11 for *reporting time + dispatch time + travel time* (2%). *Police response time* was virtually always measured as a scale variable (99%), which was operationalised as *minutes per call for service/incident* in most cases (95%). Police response time data mostly originated from US Police computer aided dispatch (CAD) systems, or predecessors of CAD. Only a few (older) studies used response times gathered by officers and observers (e.g., the Kansas City response time studies). Figure 1.2 shows a Sankey plot that presents the categories of predictors found in each study and against which component(s) of *police response time* the predictors variables within these categories were tested.

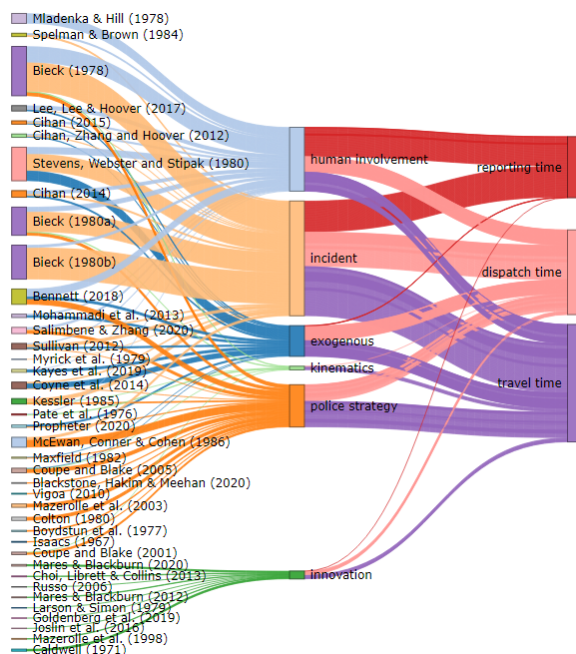


Figure 1.2: Sankey plot of studies, categories and response time components. The Sankey diagram is available more clearly as an interactive visualisation at: https://timverlaan.github.io/prt/sankey_prt.html

1.3.2 CATEGORIES OF PREDICTOR VARIABLES

Capturing the 122 unique predictor variables in intuitive categories and subcategories – which make the analysis more tangible as well as do justice to the individual predictors – is no trivial task. We have based our categorisation on the perceived (and sometimes described) mechanisms through which the predictor variables are assumed to affect (a component of) *police response time*. Table 1.4 presents the resulting categories and subcategories of this effort, next to the number of empirically studied associations within these (sub)categories and the measurement level of the predictor variables found within these (sub)categories.

Table 1.4: Categorisation of predictor variables

Category	N _{cat}	Subcategory	N _{subcat}	Unique N	Measurement levels observed
Kinematics	9	Travel distance	8	3	Dichotomous, scale
		Travel speed	1	1	Scale
Exogenous	71	Day of the week	8	2	Dichotomous, scale
		Demand	5	1	Scale
		Road type	12	6	Dichotomous, nominal, ordinal
		Time of day	36	3	Dichotomous, nominal
		Traffic	4	2	Dichotomous
		Weather type	6	4	Dichotomous
Human involvement	150	Caller characteristics	39	16	Dichotomous, scale
		Community characteristics	87	23	Scale, dichotomous, nominal
		Victim characteristics	23	10	Dichotomous, nominal, ordinal, scale
		Witness characteristics	1	1	Nominal
Incident	276	Incident characteristics	27	9	Ordinal, dichotomous, nominal
		Incident type	249	8	Nominal, dichotomous
Innovation	19	AGDS	11	1	Dichotomous
		AVL	2	1	Dichotomous
		Emergency button	1	1	Dichotomous
		Response unit type	5	1	Nominal
Police strategies	93	Call priority	34	1	Ordinal, dichotomous, nominal
		Department size	7	3	Scale
		Differential response system	1	1	Dichotomous
		Jurisdiction area size	4	2	Scale
		No. officers per response unit	6	6	Dichotomous, nominal
		Non-emergency number	8	1	Dichotomous
		On duty officers	2	1	Scale
		Patrol strategy	16	11	Dichotomous, ordinal
		Shift	7	1	Nominal
		Workload	7	5	Scale

1.3.3 CHALLENGES TO SYNTHESIZING THE FINDINGS

After analysing the included literature, we concluded that quantitative synthesis (in the form of meta-analysis) would not be possible. We identified five types of heterogeneity that hinder quantitative synthesis, here presented in no particular order: First, large heterogeneity in the operationalisation of predictor variables. For nominal and ordinal variables, this concerns heterogeneity in: the number of categories, the values these categories represent, coding system (e.g., dummy coding or deviation coding), and – in the case of dummy coding – the reference category selected. For scale variables, this concerns heterogeneity in the unit of analysis. The following two examples illustrate this heterogeneity in operationalisation:

1. *Call priority* (ordinal) has been analysed in many different operationalisations – e.g., *call priority ranging between 1 and 3* (Salimbene & Zhang, 2020) or *1 and 5* (Colton, 1980); *emergency/non-emergency* (Isaacs, 1967). While all clearly describe the same concept, they are also clearly measured using different categories. All priority levels correspond to varying target response times and even within categories, target response times vary (e.g., priority 1 can mean ‘within 10 min’ or ‘within 15 min’ in different jurisdictions). Lastly, the studies vary in the use of reference categories, it either being the lowest priority level (e.g., Salimbene & Zhang, 2020) or the highest priority level (Bennett, 2018).
2. *Demand (scale)* is generally captured by call volume. This variable has been operationalised as *number of calls for service received in the same hour as the call* (Bennett, 2018); *the number of calls for service received in the same day as the call* (Sullivan, 2012); or *the average number of calls for service per hour* (Maxfield, 1982).

Second, large heterogeneity in the constitution of (regression) models. The majority of the predictor variables are tested in multiple (linear) regression models with varying sets of control and interest variables. Obviously, the inclusion and/or omission of relevant predictors variables (can) greatly influence the size, direction and statistical significance of predictor variables’ parameter estimates.

Third, heterogeneity in the application of log-transformation of the outcome and (one) predictor variable. For 378 associations (60%), the outcome variable *police response time* (or one of its components) was log-transformed before it was analysed; in all other cases, it was not. In four regression models, the predictor variable *call rate* was also log-transformed.

Fourth, large heterogeneity in the scope of the different models. Many of the models analyse *police response time* in relation to a subset of all calls for service – e.g., those related to *domestic violence* (i.e., Lee et al., 2017), *in-progress burglary* (e.g., Coupe & Blake, 2005), *in-progress assault* (i.e., Cihan, 2015), or *all part I crimes* (e.g., Kessler, 1985). Although we have not witnessed it being articulated in the studies, the underlying assumption seems to be that there are fundamental differences between these subsets in how the predictor variables relate to police response time. Unfortunately, none of the studies explicitly tested this assumption and the extracted data do not allow us to test it ourselves.

Fifth, large heterogeneity in the metrics in which research findings are presented, e.g., Pearson’s R, Cohen’s D, unstandardised B coefficients, Chi-squared, t-statistic, Z-statistic, F-statistic, mean difference, median difference, Kendall’s tau and Beta coefficients. In many studies, a multitude of these metrics was presented without clear directions on what to base the conclusions on. While these metrics can mostly be converted into one another, the necessary information to do so was not always present or easily identifiable. Moreover, not all interpretations of the effect size of these metrics could be converted into one another. This is especially unfortunate, because (a component of) police response time is always expressed in unit of time (seconds, minutes, or hours, or a combination thereof). Standardised time differences/change (e.g., dividing the time difference by mean or median time) would therefore present the ideal metric for synthesis.

1.4 STRATEGY FOR SYNTHESIZING THE FINDINGS

Given these obstacles, we deemed it impossible to quantitatively synthesise the results in the form of a meta-analysis. Instead, we present the results in a hybrid manner: we count the number of times parameter estimates for (sub) categories of predictor variables were in a certain direction and of a specific size. This hybrid synthesis, as presented in Appendix A.5, is our best attempt to present the wealth of data for 630 associations found in 39 studies in a way that is both digestible and all-encompassing. We refer anyone interested in the intricacies of the data to consult the full data extraction file on GitHub. In the next section, we give a narrative synthesis of the main findings presented in Appendix A.5.

1.4.1 NARRATIVE SYNTHESIS

KINEMATICS

Average travel speed and *travel distance* can be reasoned to entirely explain police travel time. Any model that incorporates both accurately should thus – in theory – leave no room for other predictor variables. Only one study tested the effect of *travel speed* on *police response time* (Pate et al., 1976). Surprisingly, only a very small relation was found:

... there was little correlation between driving speed and response time. However, the speed limits generally ranged only from 20 to 35 miles per hour. With so little variation among reported driving speeds, the low statistical correlations between driving speed and response time should not be construed to indicate that driving speeds do not affect response time. The method of estimating driving speeds may not be sufficiently reliable, and the actual range of speeds may be too narrow to permit the accuracy required for correlation analysis. (Pate et al., 1976, p. 28)

The relation between *travel distance* and *travel time* was analysed in three older studies, upon which five articles were based (i.e., Bieck, 1978, 1980b; Blake & Coupe, 2001; Coupe & Blake, 2005; Kessler, 1985; Pate et al., 1976). *Travel distance* was mostly measured in kilometres (scale variable). However, some studies also recoded it into binary variables *car in beat of incident* (Kessler, 1985) and *officers' location is in beat of incident* and included it – in tandem with the default scale variable – into a regression model (Bieck, 1978, 1980b). These recoded variables of essentially the same information generally had a smaller effect size and very likely interfered with the effect of *travel distance*. The studies that looked at *travel distance* found small to large positive effects on *travel time*.

EXOGENOUS PREDICTOR VARIABLES

Exogenous variables all have – with the exception of *demand* – a very small association with police response time. *Demand*, although not tested using designs with a high internal validity, can easily be understood to have a large effect on *police response time*. A high demand for police response likely congests the dispatch centre, which would increase *queuing times*. Besides, if a fixed number of available officers need to respond to a larger volume of calls-for-service, it will probably take them longer to respond because dealing with one call will increase the time before they can leave to respond to the next call. Between *police response time* and *time of day* and *day of the week* we observed some relations of a large size.

However, these can be attributed to lack of control variables in the underlying models. For example, the one *large* relationship was found for 3.00–3.59 pm, which was the same time period the police unit under investigation had a *shift change*, but shift change itself was not accounted for in the model (Stevens et al., 1980).

HUMAN INVOLVEMENT PREDICTOR VARIABLES

Human involvement variables show very small to small statistical relations to police response time. The direction also varies. One striking finding deserves further discussion. *Community characteristic* predictor variable call rate – operationalised as the *number of calls for service per 1,000 citizen per neighbourhood per year* – shows a large relationship twice (Cihan, 2014, 2015). These parameter estimates seem anomalous in light of the other evidence that call rate shows two small negative and one very small positive relation to police response time. Because we also cannot think of a theoretical mechanism that would explain why call rate – in its current operationalisation – would have such a large relationship to *police response time*, we urge the reader to interpret these large parameter estimates presented in Cihan (2014, 2015) with caution.

INCIDENT PREDICTOR VARIABLES

Incident type is by far the most analysed predictor variable. However, most of these analyses are based on observations from the Kansas City response time study (i.e., Bieck, 1978, 1980a, 1980b; Kessler, 1985), which was a relatively small study (sample sizes varied in the various presented regression models, max = 2,800, mean = 669.5, SD = 444.6) and the study has been criticised as a ‘highly non-random sample’ (Blanes i Vidal & Kirchmaier, 2018). Besides, *call priority level* has not been considered in the majority of the regression models based on the Kansas City response time study. The only evidence that sheds some light on the effect of *incident type* controlling for *call priority level* is found in Salimbene and Zhang (2020). They report only very small differences between the reference category disorder crime and incident categories *violent crime*, *property crime* and *other crime* types. *Incident characteristics* can roughly be broken down into four types: (*degree of*) *injury*, *weapon involvement*, *in progress* and *source of the call for service* (911 or automated alarm system). Controlling for *call priority level*, *weapon involvement* and *source of the call for service* both have very small relations to *police response time*. Without controlling for *call priority level*, some (other) studies find the relationships for (*degree of*) *injury* to range from very small positive to large negative, and those for *in-progress* to range from very small to large negative. However, we do not place too much trust in these findings because *call priority level* was not controlled for.

INNOVATIONS

We labeled four predictor variables as *innovations*: *acoustic gunshot detection system* (AGDS), *automatic vehicle location systems* (AVL), *emergency button*, and *response unit type*. AGDSs are systems that should directly affect *reporting time*. Surprisingly, they have not yet been analysed against reporting time. Instead, AGDSs show very small to small relations to the combined components dispatch time and travel time. In two studies by Mares and Blackburn (2012, 2021), this association also changes direction over time. At first, calls from the AGDS received slightly faster dispatch and travel times, while they later received slower dispatch and travel times. Dennis Mares argued (in personal communication) that

this might be explained by *response fatigue*. AGDS-generated calls for service mostly turn out to lead to no suspects or arrests and are often believed to be false positives. AVLs were analysed in two studies, using data gathered in 1974 (G. C. Larson & Simon, 1979) and 1999 (Russo, 2006). AVLs are meant to improve travel times by enabling dispatching of the closest car to an incident. Despite this clear mechanism, the systems had very small relations to *police response times*. Possibly, the technology at the time was not advanced enough to outperform status quo. It should be noted that the studies by Russo (2006) and G. C. Larson and Simon (1979) have a low internal validity level and results should thus be interpreted with caution. The *emergency button* was only researched in one study with a relatively low internal validity score. Nevertheless, the results showed a large negative relation to *police response time* (reporting + dispatch + travel time), meaning the intervention lowered *police response time*. This shows the potential for innovations that are expected to affect *reporting time*, the component of *police response time* which Spelman and Brown (1981, p. 55) suggested to be the largest contributor to *police response time*. *Response unit type* measures the difference in response time between two types of police cars in the late 1960s. Because the context of these findings is so outdated, we refrain from drawing any conclusions.

POLICE STRATEGIES

Police strategies is by far the broadest category we used. It includes: *call priority*, *workload*, *department size*, *differential response system*, *jurisdiction area size*, *number of officers per response unit*, *implementation of a non-emergency number*, *number of on-duty officers*, *patrol strategy*, and *shift*. The common denominator of all these predictors is that all are – in some way – measures that reflect decisions on how police operate. *Call priority* has a very small to large positive (positive due to inverse order of priority level, level 1 being highest) effect on *police response time*. There are some instances where the primary results suggested a negative relation, but this was entirely due to the use of a different reference category (referenced against highest instead of lowest priority level). *Call priority* mostly shows a large effect and from analysing the results in context we deem this to be accurate. *Workload* is suggested to have varying relations with police response time for different priority levels. For high priority calls, it seems there is virtually always enough capacity, and *workload* is thus not a limiting factor. In contrast, variation in *police response time* for low priority calls seems to be explained for a large part by *workload*. *Department size* (–), *jurisdiction area size* (+), *number of officers per response unit* (–), *number of units responding* (–) and *number of on-duty officers* (–) all show very small to small relations with *police response time*, with their respective direction reported between brackets. Within subcategory department size, the predictor variable *number of officers employed* had a negative direction while the predictor variable *number of citizens employed* showed a negative direction once and a positive direction twice. *Shift* captures the difference between the morning, night, and evening shift (while controlling for other variables), and these showed very small to small relations to response times. Coupe and Blake (2005) report that *shift change* has a significant effect on response times, as was also hinted at in Stevens et al. (1980). Unfortunately, the effect size and direction of *shift change* is not mentioned in the study, only its p-value (Coupe & Blake, 2005).

1.5 DISCUSSION

First, we discuss the general characteristics of the reviewed literature. Second, we discuss the results as presented in Appendix A.5 and the narrative synthesis. Third, we discuss the limitations of this review. Finally, we present recommendation for future research into police response time and our attempt towards building a causal model.

1.5.1 DISCUSSION OF THE LITERATURE

Looking at the general characteristics of the included studies, a few things immediately stand out. First, the vast majority of studies (90%) originate from the United States (and most from large urban areas), with only three studies from the United Kingdom and one from Iran. Without sufficient research from other countries, it remains unclear how the urban and US-centric findings generalise. In countries where predictor variables have a greater range or variance, they might also differently affect *police response times*. For example, while *road type* did not prove to be a limiting factor to *police response time* in the large cities studied in the United States, it is easily imagined that mountainous dirt roads in rural parts of the world will constitute a limiting factor to *police response time*.

Second, research interest into police response time seems to have come in two waves, with a quiet period between 1985 and 1995. Studies from the first wave mainly focus on *incident* predictor variables. Studies from the second wave, especially more recently, mainly focus on whether and how *police response times* vary with *human involvement* variables. Remarkably, some important predictor variables like *travel distance* and (average) *travel speed* were controlled for in the first wave – when measuring them was very difficult and less reliable – more often than in the second (current) wave when measuring them – and thus controlling for them – is far easier and more reliable.

Third, since 1977, all data used in the included studies were extracted from computer aided dispatch (CAD) systems from (local) police departments or other such large-scale databases. These CAD data should be considered population data for the police department they are extracted from. With the exception of Bennett (2018), none of the studies included in this review choose the local police departments from which they extract data through a sampling procedure. Instead, police departments are selected based on convenience and/or other criteria. In the absence of a predefined sample space and explicit sampling procedure, the results of these studies should not be interpreted as generalizing to any larger spatial unit such as state, country or continent. As the studies are based on population data that themselves do not generalise to a larger superpopulation, there is actually no need for the use of p-values or inferential statistics in general: all the variation and effects observed are, by definition, the real variation and effects. In other words, these studies provide stronger evidence than they give themselves credit for, although it is evidence that does not present a clear basis for generalisation beyond the study area.

Fourth, the literature fails to build towards the common goal of gaining a comprehensive understanding of determinants of (and their effects on) police response time. Perceived mechanisms and causal models are scarcely communicated, and if they are, not in an explicit manner (i.e. in non-specific text rather than in causal diagrams). The models seem not to get gradually improved based on insights from prior research. Overall, the literature seems to be stuck in a paradigm of correlation rather than trying to work towards a causal model.

1.5.2 DISCUSSION OF THE RESULTS

In total, the relationship between 122 unique predictor variables and (a component of) police response time was analysed. Given all the obstacles on synthesizing the findings mentioned before, it is not possible to draw any definitive conclusions on the effects of predictor variable (sub)categories. Nevertheless, we will discuss some observations and lessons learned from the synthesis of each variable (sub)category.

Kinematics directly determines travel time and therefore have a causal link with police response time. More effort should be directed to accurately measuring and including *travel distance* or (*average*) *travel speed* in studies of police response time. Note though that including both in a regression model will leave no room for other predictors. Having those measures might actually shift the research interest from the explanation of police response time as such to explaining why police officers drive at different speeds or are positioned at different distances from incident locations.

Exogenous variables (except *demand*) have negligible associations with police response time in the contexts where they have been studied. Only in very extreme cases can variables like *weather* or *traffic* be expected to have a noteworthy effect on police response time. For this reason, we consider the inclusion of these *exogenous variables* (except *demand*) in a causal model as optional in urban areas of most developed countries.

The association between *human involvement variables* and police response time is potentially contentious and easily linked to societal debates around ‘Defund the Police’ and ‘Black Lives Matter’. It is, therefore, vital that the research into this is rigorous in its design, cautious in its claims and clear on the distinction between correlation and causation. Nevertheless, most studies that have looked into this make bold claims we feel are not warranted by the research presented. The most authoritative article on *human involvement variables* (Bennett, 2018) retrieved CAD data from 40 police agencies and measured the effect of many *human involvement variables* on dispatch and travel time. Effect sizes and direction varied greatly, which caused Bennett (2018, p. 30) to write the following: ‘Notice that in every case the standard deviation is much larger than the mean, in most cases by a factor of three or more. While the sign of the means mostly correspond to the estimates from the combined regression, in every case there are estimates on both sides of zero.’

From this we conclude that the associations of human involvement variables with police response time are far from homogenous or clear, and scholars should be careful to draw any conclusions. More (rigorous) research is needed.

Incident Variables show varying effect sizes, from very small to large. However, in most of the models in which these varying findings are observed *call priority level* is not controlled for. *Call priority level* is determined exclusively by incident variables and itself has a direct causal mechanism to police response time. Unsurprisingly, *incident variables* were found to have only a very small effect on police response time once *call priority level* was controlled for. We believe that the variation in police response time for different incident variables – not explained by *call priority level* – is indeed very small. Nevertheless, the mechanisms behind these small differences present an interesting avenue for future research. All current *innovations* aim to shorten reporting time by either bypassing citizen action or facilitating it. Even though the evidence base is either currently lacking or very small, we believe these innovations can actually have large effects on shortening reporting time and, consequently, police response time.

1

Regarding the subcategory *police strategies*, we limit our discussion to a few predictor variables that are suggested to – empirically and/or theoretically – have a large effect on police response time: workload (consisting of *demand* and *police capacity*), *call priority level*, and *police discretion*. Of these variables, *call priority level* has the largest effect as it directly dictates police actions. Any variation among incidents of the same priority level can thus be attributed to actions not dictated. In part it will be related to *police discretion*, which can influence police response time in three ways: police dispatchers handling calls for service faster/slower, patrol officers travelling to incidents faster/slower, or patrol officers positioning themselves so they are closer to/further away from (future) incidents. *Workload* mainly determines when a response is initiated and therefore has a greater effect on calls with lower priority.

1.5.3 LIMITATIONS

In this section I discuss the limitations of the data gathering and analysis, and the limitations of the manner in which the findings are presented.

LIMITATIONS OF DATA GATHERING

The first limitation of this scoping review is that the search terms did not include the separate components of police response time (i.e., reporting time, dispatch time and travel time) – the importance of which was only realised during the analysis of the retrieved data. Nevertheless, we doubt that any relevant study only uses the component terms without any of the various search terms we included for police response time (i.e., ‘rapid response’, ‘prompt response’, ‘arrival time’, and ‘response time’). A second limitation of this literature review is that no double collecting, screening, selecting and coding of the literature was done. While this undeniably lowers the reliability of the findings, we are confident that – largely thanks to our open and systematic way of analysing and interpreting the data – the results meet the standards of reproducibility and reliability. We believe saturation was reached in searching for relevant literature as both the forward and backward searches only yielded a single extra study – neither of which was found to be present in the literature retrieved from primary databases. Third, the disproportionate concentration of data emanating from the United States, coupled with an inadequate representation of data from other geographical origins, likely has implications for generalisability of this review’s findings. While articulating these specific implications is challenging, our preceding discussion attempts to illuminate the potential ramifications arising from the limited diversity observed in the study regions.

LIMITATIONS OF DATA ANALYSIS

In order to present the results in a digestible manner, we presented the observed statistical relations on a four-point scale (very small/small/medium/large) in Appendix A.5. Assigning those values for Cohen’s D, Pearson’s R and odds ratios was straightforward using Cohen’s (2013) classification or extrapolations of this work (see Appendix A.4). However, for unstandardised B coefficients, standardised Beta coefficients and percentual change, we could not rely on existing rules of thumb and had to decide on what constitutes a very small/small/medium/ large statistical relation. Such judgment calls are inherently arbitrary; using different cut-off values might lead to different results. We also wanted to provide some indication of the (average) internal validity of the scientific evidence per subcategory

by using the Maryland Scientific Methods Scale. All these steps should be taken as a first attempt to present the large data collection effort in a way that is still digestible by a reader. We encourage the reader to critically assess these choices or to even suggest other ways of synthesizing large quantitative data collections without solely relying on a narrative synthesis. As systematic reviews are periodically updated, we hope constructive criticism is able to improve the validity of our interpretations for updated versions.

LIMITATIONS OF THE FINDINGS

In many ways, this study suffers from attempting to standardise the analysis of data that proved to be too heterogeneous to be captured in such a rigorous manner. As a result, the article might now read as a critique of prior work in this field rather than having an exclusive focus on uncovering the empirically tested determinants of police response time and their effect size, direction and significance. Although we are aware of this limitation, we believe this is a necessary assessment of the research evidence and the article presents a better way forward than any (less critical) alternative would offer.

1.5.4 RECOMMENDATIONS FOR FUTURE RESEARCH INTO POLICE RESPONSE TIME

In order to help police response time research progress – and in response to the obstacles to quantitative synthesis we had to deal with – we have four general recommendations for future research.

First, we argue that the research into police response time should shift from the current correlational paradigm to a causal paradigm. Many exploratory studies have now been done and the field would benefit from research that works towards a causal model of police response time. Moreover, this shift would likely result in higher MSMS scores than are currently witnessed. Obtaining all the data required to properly test such a causal model can be challenging and might well be beyond the possibilities of many research projects. Nevertheless, clearly defining and communicating the perceived causal model and pitfalls in testing it should be possible for all research projects and would already be a big step forward for this line of research.

Second, we argue that each *call priority level* corresponds to a distinct *police response time* distribution. Therefore, we recommend to analyse the different call priority levels separately. Moreover, in case the goal is to investigate how specific determinants limit faster police response time, we would even argue that researchers should only analyse variables if they can be assumed to present a limiting factor. Only when the police are aimed at realizing the shortest possible response time, all variables can be assumed to present a limiting factor. Therefore, we recommend – if the goal is to investigate how specific determinants limit faster police response times – to only analyse the highest priority levels calls.

Third, research into police response time would benefit greatly from more standardisation in operationalisation of both predictor and outcome variables. Such standardisation is preferably reached through consensus among experts. In order to start this conversation, we present some suggestions on four matters we deem most in need of standardisation in the current literature: transformations, measure of central tendency, coding, standardisation for effect size interpretation:

1. The majority of the 630 associations were tested against log-transformed components of the outcome variable *police response time*. In most of the studies that used log-transformation, it was argued useful for making the outcome variable less skewed, or, to be beneficial for regression analysis as it improves normality of the error distribution (e.g., Cihan et al., 2012, p. 316; Lee et al., 2017, p. 67). We argue against log-transformation for these reasons and believe skewedness of the data should be investigated rather than treated a nuisance. We judge much of the skewedness of police response time data to be a result of measurement error – i.e., by police officers forgetting to log their *on-scene arrival*, resulting in a number of very large *police response times*. A more appropriate response to the observed skewedness would thus be *outlier detection*, as employed by Choi et al. (2014).
2. If outlier detection does not suffice to curb large skewedness of the data, we argue that the *median* would be the better measure of central tendency over the *mean*, which is now more commonly used. To illustrate: in the study by Bieck (1978, p. 140), the *mean* police response time (reporting + dispatch + travel) is reported to be 3 h and 57 min, while the *median* response time is reported to be 18 min and 50 s; we argue in such cases the median better reflects the central tendency of police response times.
3. With the exception of Stevens et al. (1980), which uses deviation coding, all studies that included nominal variables in multiple regression models use dummy coding. This can only be done with an arbitrary reference category, and we observed over twenty different reference categories used for the predictor *incident type* alone. We argue that deviation coding is preferable as it would improve comparability and facilitate synthesis of findings.
4. In order to make normative judgements on the size of observed difference or associations, many studies standardise the associations by using the standard deviation of the predictor and/or outcome variable distribution. For example, Coupe and Blake (2005, p. 249) report: 'As expected, response times were strongly associated with response distances (Pearson's $r = +0.47$, $p < 0.01$)'. We argue, instead of presenting (on standard deviation) standardised correlation measures and its standardised interpretation (e.g., 'strong association'), that it is generally more meaningful to present the associations in the intuitive units it is measured in and interpret effect sizes in context, for example, by comparing the unstandardised differences/effects to the median response time.

Fourth, more research should be conducted into *reporting time*. Several of the earliest studies (Bieck, 1978, 1980a, 1980b; Spelman & Brown, 1981) show that reporting time constitutes the largest share of total police response time and claim it should thus receive most focus when aiming to improve response times. To illustrate, a quote from Spelman and Brown (1981):

Most of the time, the [reporting] delays are so substantial that even our fastest response to the crime will be ineffective in producing arrests. In short, we have focused on using high technology dispatching equipment and sophisticated deployment schemes to reduce police response time, when we should have focused on reducing citizen delays. (Spelman & Brown, 1981, xi–xii)

While the evidence base upon which these claims are made is rather slim, may be outdated and originates from a very specific setting (five cities in the USA), we judge *reporting*

time to be a very promising and often neglected research topic, especially when the aim of the research is to contribute to improving *police response time*. However, it is difficult to accurately measure reporting time. We believe more research into AGDSs would provide a good start, as AGDSs automatically record timestamps for each incident (e.g., *shots fired* or *shooting*), which could then be compared to the reporting timestamps of calls-for-service in order to obtain an accurate reporting time. Furthermore, the only intervention tested in relation to *reporting time* (i.e., *emergency button*) showed a very large effect of on shortening overall police response time (Joslin et al., 2016).

1.5.5 TOWARDS A CAUSAL MODEL OF POLICE RESPONSE TIME

Although the methodologies of the reviewed studies do not allow for causal inference, the predictor variables studied could well be determinants of police response time. In fact, for many of the predictor variables, a mechanism by which they influence (a component of) police response time can easily be reasoned. This is most evident for *travel distance* and *travel speed* as these fully determine travel time, but it also applies to other predictor variables.

Because most studies neglect to communicate the (perceived) mechanisms, much remains unknown about the causal model of *police response time*. In order to work towards a causal model of police response time, we present a first conceptualisation of a causal model that consists of the subcategories identified from the empirical literature included in this review. Appendix A.6 depicts this first conceptualisation in the form of a causal directed acyclic graph (Textor et al., 2016). The different categories are depicted using different colors. Next to the subcategories we used to synthesise the literature, we included two mediating variables (*coverage* and *willingness to help*) and one moderating variable (personal bias/ preference) in the model in order to complete the perceived mechanisms in the model.

The causal model presented in Appendix A.6 is by no means a finished product; it is likely underdetermined (e.g., regarding *reporting time*) and at the same time oversaturated (e.g., because of the inclusion of very specific subcategories). Nevertheless, it provides a first mapping of the current state of the literature and could be used to guide future scientific inquiries into police response time and show how these will be related. It can be helpful in identifying the variables that should be controlled for when trying to estimate specific total, direct, or indirect effects of a certain (set of) predictor variable(s) on a component of police response time. Based on this first causal model, we would like to make one final recommendation:

Instead of investigating the effects of variables on overall police response time, we argue it is better to focus on a certain component of *police response time*. For example, studies that investigated the effects of *human involvement* variables on overall *police response time* often find very small to small relationships: certain communities would receive slightly slower or faster response times. Conclusions are then drawn – though never very explicitly – that these findings suggest racial biases among police personnel. However, with the empirical designs used, it is often unclear whether the observed disparities are explained by travel speed, travel distance or handling time. Disparities in travel distance would have to be attributed to a focus on specific locations within a police jurisdiction and such decisions are generally not made at the level of individual police officers. Studies would present stronger evidence for racial bias among police officers if they could show disparities in the (mean, median or max) *travel*

speed of patrol officers or *call handling time* of dispatchers. In other words, more specific causal models and analyses are easier to communicate, control for confounding, guard for omitted variable bias and generally less susceptible to noise and therefore, provide a stronger evidence base and contribute more to the development of the field.

ACKNOWLEDGEMENTS

We thank Kreg Purcell, librarian at the *OJP Response Center & Library Services*, for his assistance in locating and obtaining various studies included in this scoping review.

AVAILABILITY OF DATA AND MATERIALS

Can be retrieved from: <https://github.com/timverlaan/prt>.

REFERENCES

- Bennett, D. S. (2018, November). *Police Response Times to Calls for Service: Fragmentation, Community Characteristics, and Efficiency* [Doctoral dissertation, Stanford Institute for Economic Policy Research].
- Bieck, W. H. (1977b). *Response Time Analysis: Part I Crime Analysis (Vol. II)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1978). *Response Time Analysis: Methodology (Vol. I)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1980a). *Response Time Analysis: Noncrime Analysis (Vol. IV)* (tech. rep.). National Institute of Justice. Kansas City, MO.
- Bieck, W. H. (1980b). *Response Time Analysis: Part II Crime Analysis (Vol. III)* (tech. rep.). National Institute of Justice. Kansas City.
- Blackstone, E. A., Hakim, S., & Meehan, B. (2020). Burglary reduction and improved police performance through private alarm response. *International Review of Law and Economics*, 63. <https://doi.org/10.1016/j.irle.2020.105930>
- Blake, L., & Coupe, R. T. (2001). The Impact of Single and Two-Officer Patrols on Catching Burglars in the Act: A Critique of the Audit Commission's Reports on Youth Justice. *British Journal of Criminology*, 41(2), 381–396. <https://doi.org/10.1093/bjc/41.2.381>
- Blanes i Vidal, J., & Kirchmaier, T. (2018). The Effect of Police Response Time on Crime Clearance Rates. *The Review of Economic Studies*, 85(2), 855–891. <https://doi.org/10.1093/restud/rdx044>
- Boydston, J. E., Sherry, M. E., & Moelter, N. P. (1977). *Patrol staffing in San Diego: One- or two-officer units*. (tech. rep.). National Policing Institute. Washington, DC.
- Brandl, S. G., & Horvath, F. (1991). Crime-victim evaluation of police investigative performance. *Journal of Criminal Justice*, 19(3), 293–305. [https://doi.org/10.1016/0047-2352\(91\)90008-J](https://doi.org/10.1016/0047-2352(91)90008-J)
- Brown, B. B., & Harris, P. B. (1989). Residential burglary victimization: Reactions to the invasion of a primary territory. *Journal of Environmental Psychology*, 9(2), 119–132. [https://doi.org/10.1016/S0272-4944\(89\)80003-9](https://doi.org/10.1016/S0272-4944(89)80003-9)
- Caldwell, R. (1971). *Optimal distribution policy for mobile police patrols* [Doctoral dissertation, Durham University].
- Choi, K. S., Librett, M., & Collins, T. J. (2014). An empirical evaluation: gunshot detection system and its effectiveness on police practices. *Police Practice and Research*, 15(1), 48–61. <https://doi.org/10.1080/15614263.2013.800671>
- Cihan, A. (2014). Social disorganization and police performance to burglary calls: A tale of two cities. *Policing*, 37(2), 340–354. <https://doi.org/10.1108/PIJPSM-06-2013-0058>
- Cihan, A. (2015). Examining the neighborhood effects on police performance to assault calls. *Police Practice and Research*, 16(5), 391–401. <https://doi.org/10.1080/15614263.2014.927765>
- Cihan, A., Zhang, Y., & Hoover, L. (2012). Police Response Time to In-Progress Burglary. *Police Quarterly*, 15(3), 308–327. <https://doi.org/10.1177/1098611112447753>
- Cohen, J. (2013, May). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>

- 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>
- Colton, K. W. (1980). *Police and Computer Technology: The Case of the San Diego Computer-Aided Dispatch System* (tech. rep. No. 1).
- Cordner, G. W., Greene, J. R., & Bynum, T. S. (1983). *The Sooner the Better: Some Effects of Police Response Time* (tech. rep.). National Institute of Justice. Washington, DC.
- Coupe, R. T., & Blake, L. (2005). The effects of patrol workloads and response strength on arrests at burglary emergencies. *Journal of Criminal Justice*, 33(3), 239–255. <https://doi.org/10.1016/j.jcrimjus.2005.02.004>
- Coyne, C. J., Hall, A. R., McLaughlin, P. A., & Zerkle, A. (2014). A hidden cost of war: the impact of mobilizing reserve troops on emergency response times. *161*(3), 289–303. <https://doi.org/10.1007/sl>
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In D. P. Farrington, D. C. Gottfredson, L. W. Sherman & B. C. Welsh (Eds.), *Evidence-based crime prevention* (1st ed., pp. 13–21). Routledge.
- Frank, J., Smith, B. W., & Novak, K. J. (2005). Exploring the Basis of Citizens' Attitudes Toward the Police. *Police Quarterly*, 8(2), 206–228. <https://doi.org/10.1177/1098611103258955>
- Goldenberg, A., Rattigan, D., Dalton, M., Gaughan, J. P., Thomson, J. S., Remick, K., Butts, C., & Hazelton, J. P. (2019). Use of ShotSpotter detection technology decreases prehospital time for patients sustaining gunshot wounds. *Journal of Trauma and Acute Care Surgery*, 87(6), 1253–1259. <https://doi.org/10.1097/TA.0000000000002483>
- Isaacs, H. H. (1967). *A Study of Communications, Crimes and Arrest in a Metropolitan Police Department. Appendix B of Science and Technology Task Force Report*. (tech. rep.). U.S. Government. Washington DC.
- Joslin, J. D., Goldberger, D., Johnson, L., & Waltz, D. P. (2016). Use of the Vocera Communications Badge Improves Public Safety Response Times. *Emergency Medicine International*, 2016, 1–3. <https://doi.org/10.1155/2016/7158268>
- Kayes, M. I., Al-Deek, H., Sandt, A., Carrick, G., Staves, C., Gamaleldin, G., Faruk, O., & Al-Sahili, O. (2019). Characteristics of Law Enforcement Response to Wrong-Way Driving Events in Florida. *Transportation Research Record*, 2673(9), 514–524. <https://doi.org/10.1177/0361198119846456>
- Kessler, D. A. (1985). One-or two-officer cars? A perspective from Kansas City. *Journal of Criminal Justice*, 13(1), 49–64. [https://doi.org/10.1016/0047-2352\(85\)90025-X](https://doi.org/10.1016/0047-2352(85)90025-X)
- Larson, G. C., & Simon, J. W. (1979, February). *Evaluation of a police AVM system: A study of the St. Louis experience 1976-1977* (tech. rep. No. 1). National Institute of Law Enforcement and Criminal Justice.
- Lee, J. S., Lee, J., & Hoover, L. T. (2017). What Conditions Affect Police Response Time? Examining Situational and Neighborhood Factors. *Police Quarterly*, 20(1), 61–80. <https://doi.org/10.1177/1098611116657327>

- Mares, D., & Blackburn, E. (2012). Evaluating the Effectiveness of an Acoustic Gunshot Location System in St. Louis, MO. *Policing*, 6(1), 26–42. <https://doi.org/10.1093/police/par056>
- Mares, D., & Blackburn, E. (2021). Acoustic gunshot detection systems: a quasi-experimental evaluation in St. Louis, MO. *Journal of Experimental Criminology*, 17(2), 193–215. <https://doi.org/10.1007/s11292-019-09405-x>
- Maxfield, M. G. (1982). Service Time, Dispatch Time, and Demand for Police Services: Helping More by Serving Less. *Public Administration Review*, 42(3), 252. <https://doi.org/10.2307/976012>
- Mazerolle, L., Rogan, D., Frank, J., Famega, C., & Eck, J. E. (2003). *Managing Citizen Calls to the Police: An Assessment of Non-Emergency Call Systems*. President, Statistical Analysis for Law Enforcement Strategies (tech. rep.).
- Mazerolle, L. G., Watkins, C., Rogan, D., & Frank, J. (1998). Using Gunshot Detection Systems in Police Departments. *Police Quarterly*, 1(2), 21–49. <https://doi.org/10.1177/109861119800100202>
- Mcewen, J. T., Connors, E. F., & Cohen, M. I. (1986). *Evaluation of the Differential Police Response Field Test* (tech. rep.).
- Mladenka, K. R., & Hill, K. Q. (1978). The Distribution of Urban Police Services. *The Journal of Politics*, 40(1), 112–133. <https://doi.org/10.2307/2129978>
- Mohammadi, M., Martiniuk, A. L. C., Ansari-Moghaddam, A., Rad, M., Rashedi, F., Ghjasemi, A., & Ahangari, H. (2013). Police response time to road crashes in south-east of Iran. *Journal of the Pakistan Medical Association*, 63, 1523–1527.
- Myrick, J. A. (1979). *Evaluation of the role of Police in the EMS system* (tech. rep.). Georgia Institute of Technology. Atlanta, Georgia.
- Pate, A. M., Ferrara, A., Bowers, R. A., & Lorence, J. (1976, January). *Police Response Time: Its Determinants and Effects* (tech. rep.). National Policing Institute. Washington, DC.
- Propheter, G. (2020). Do urban sports facilities have unique social costs? An analysis of event-related congestion on police response time. *International Journal of Urban Sciences*, 24(2), 271–281. <https://doi.org/10.1080/12265934.2019.1625805>
- Russo, C. W. (2006). *AVL and Response Time Reduction: Image and Reality* [Doctoral dissertation, University of Central Florida].
- Salimbene, N. A., & Zhang, Y. (2020). An examination of organizational and community effects on police response time. *Policing: An International Journal*, 43(6), 935–946. <https://doi.org/10.1108/PIJPSM-04-2020-0063>
- Scherhag, J., & Burgard, T. (2023). Performance of Semi-Automated Screening Using Rayyan and ASReview: A Retrospective Analysis of Potential Work Reduction and Different Stopping Rules. *Big Data & Research Syntheses*.
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Stevens, J. M., Webster, T. C., & Stipak, B. (1980). *Response Time: Role in Assessing Police Performance* (tech. rep. No. 3).
- Sullivan, E. (2012). *The Effects of Crime Incident Characteristics and Neighborhood Structure on Police Response Time* [Doctoral dissertation, Arizona State University].

- Textor, J., van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. *International Journal of Epidemiology*, 45(6), 1887–1894. <https://doi.org/10.1093/IJE/DYW341>
- Tricco, A. C., Lillie, E., Zarin, W., O’Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdem, 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>
- Vigoa, O. F. (2010). *Comparative Analysis between Routine Preventative Uniformed Comparative Analysis between Routine Preventative Uniformed Police Patrols and Crime Reduction and Calls for Service Police Patrols and Crime Reduction and Calls for Service* (tech. rep.).
- Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, 103(2), 280–293. https://doi.org/10.1162/rest{_}_}a{_}_}00889
- Zhu, R., Hu, X., Li, X., Ye, H., & Jia, N. (2020). Modeling and Risk Analysis of Chemical Terrorist Attacks: A Bayesian Network Method. *International Journal of Environmental Research and Public Health*, 17(6), 2051. <https://doi.org/10.3390/ijerph17062051>



APPENDIX: SCOPING REVIEW

A.1 FLOW CHART

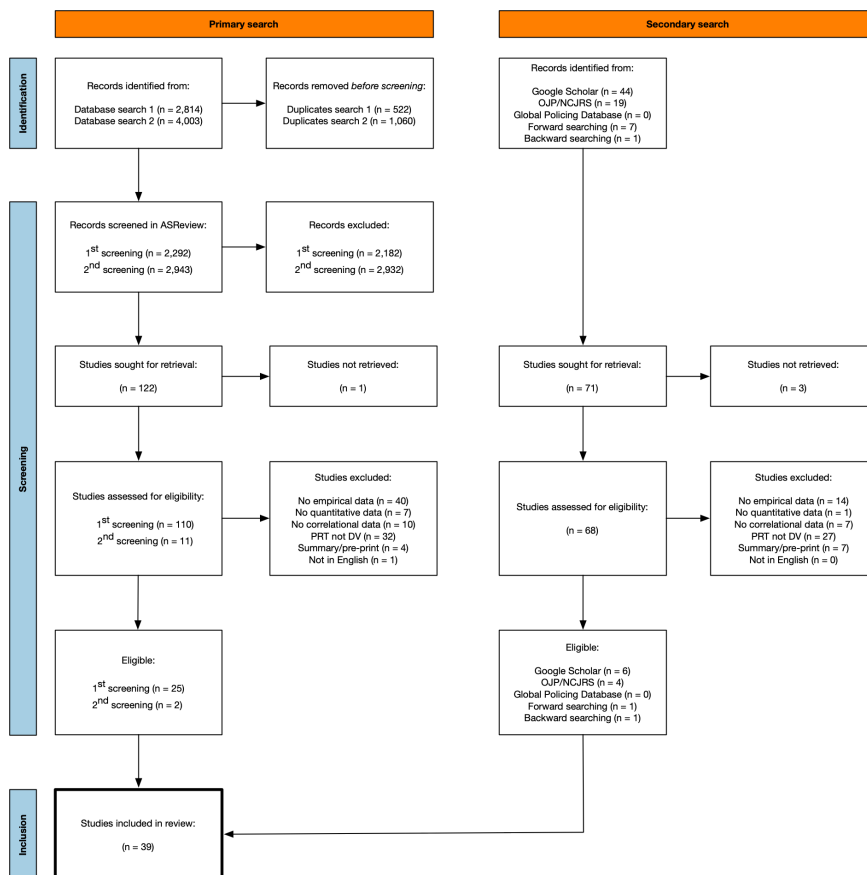


Figure A.1: Flow chart

A.2 INCLUDED STUDIES

Table A.1: Included studies

Authors	Year	Location	Design	Eval?	Evaluated	Study type	Sampling technique
Bennett	2018	40 jurisdictions in USA	observational	No	–	Sample	stratified random sampling
Bieck	1977b	Kansas City (USA)	observational	No	–	Sample	non-random sampling
Bieck	1980b	Kansas City (USA)	observational	No	–	Sample	non-random sampling
Bieck	1980a	Kansas City (USA)	observational	No	–	Sample	non-random sampling
Blackstone, Hakim & Meehan	2020	Salt Lake City (USA)	observational	Yes	Verified Response	Population	–
Blake & Coupe	2001	West Midlands (UK)	observational	Yes	One or two-officer car	Sample	disproportionate stratified sampling
Boydston et al.	1977	San Diego (USA)	experimental	Yes	One or two-officer car	Sample	random sampling
Caldwell	1971	Durham (UK)	observational	No	–	Population	–
Choi, Librett & Collins	2014	Brockton (USA)	quasi-experimental	Yes	AGDS	Population	–
Cihan	2015	Houston (USA)	observational	No	–	Population	–
Cihan	2014	Houston, Dallas (USA)	observational	No	–	Population	–
Cihan, Zhang & Hoover	2012	Houston (USA)	observational	No	–	Population	–
Colton	1980	San Diego (USA)	observational	No	–	Population	–
Coupe & Blake	2005	West Midlands (UK)	observational	No	–	Sample	disproportionate stratified sampling
Coyne et al.	2014	Country-wide USA	observational	Yes	Mobilizing reserve troops	Population	–
Goldenberg et al.	2019	Camden (USA)	observational	Yes	AGDS	Population	–
Isaacs	1967	Los Angeles (USA)	observational	No	–	Population	–
Joslin et al.	2016	Syracuse (USA)	observational	Yes	Vocera badge	Population	–
Kayes et al.	2019	Florida (USA)	observational	No	–	Population	–
Kessler	1985	Kansas City (USA)	observational	Yes	One or two-officer car	Sample	non-random sampling
Larson & Simon	1979	St Louis (USA)	experimental	Yes	AVM	Population	–
Lee, Lee & Hoover	2017	Houston (USA)	observational	No	–	Population	–
Mares & Blackburn	2021	St Louis (USA)	quasi-experimental	Yes	AGDS	Population	–
Mares & Blackburn	2012	St Louis (USA)	quasi-experimental	Yes	AGDS	Population	–
Maxfield	1982	San Francisco (USA)	observational	No	–	Population	–
Mazerolle et al.	2003	Dallas (USA)	observational	Yes	Non-emergence call system	Population	–
Mazerolle et al.	1998	Dallas (USA)	experimental	Yes	AGDS	Population	–
McEwan, Conner & Cohen	1986	Garden Grove; Greensboro; Toledo (USA)	experimental	Yes	Differential Police Response	Population	–
Mladenka & Hill	1978	Houston (USA)	observational	No	–	Sample	random sampling
Mohammadi et al.	2013	Sistan and Baluchistan (IRAN)	observational	No	–	Population	–
Myrick et al.	1979	Tucker (USA)	observational	No	–	Population	–
Pate et al.	1976	Kansas City (USA)	observational	No	–	Sample	non-random sampling
Propheter	2020	Sacramento (USA)	observational	Yes	Presence of Urban Sport Facilities	Population	–
Russo	2006	Altamonte Springs (USA)	observational	Yes	AVL	Population	–
Salimbene & Zhang	2020	Northeast Texas (USA)	observational	No	–	Population	–
Spelman & Brown	1981	Jacksonville, Peoria, Rochester, San Diego (USA)	observational	No	–	Sample	stratified random sampling
Stevens, Webster & Stipak	1980	York (USA)	observational	No	–	Sample	disproportionate stratified sampling
Sullivan	2012	Southwest Sky (USA)	observational	No	–	Population	–
Vigoa	2010	Doral (USA)	observational	No	–	Population	–



A.3 DATA EXTRACTION FORM

Table A.2: Overview of extracted variable coding and documentation

Column name	Explanation
Independent variable (iv)	Predictor variable name as presented in the primary study
Iv dummy	
Iv label	
Iv type	Predictor variable type: <i>nominal, dichotomous, ordinal</i> or <i>scale</i>
Iv description	A description of what the predictor variable entails
Iv reference category	If <i>nominal</i> or <i>ordinal</i> : IV reference category is noted here.
Iv unit	If <i>scale</i> : IV unit of analysis is noted here
Iv transformation	<i>None, Ln</i>
Dv component	<i>Reporting, dispatch</i> or <i>travel time</i> ; or any combination
Dv type	The outcome variable type: <i>nominal, ordinal</i> or <i>scale</i>
Dv unit	Unit of analysis of police response time (usually <i>min/incident</i>)
Dv transformation	<i>None, Ln</i>
Mechanism described	
Counterfactual/balancing	
Iv function	
Study reference	Authors (year published)
Source	Table or page where extracted data was found
Correlation test	Statistical test used to gage the correlation between IV and DV
Conclusion study	The conclusion the study draws based on the correlational data
Sample/population	Does the study concern population or sample data?
Sample space	To what geographic unit does this data generalise?
Research design	<i>Observational/quasi-experimental/experimental</i>
Intervention research	<i>Yes/no</i>
Evaluated	Description of what (IV) is evaluated
Cad data?	<i>Yes/no</i>
Alternative dataset	If not CAD, what then?
Sampling strategy	Sampling strategy for analysis unit of analysis (usually <i>CFSs</i>)
Sample info	Additional information on sampling choices specified
Sample priority level 1	
N (sample)	Sample size
Sample unit	Unit of analysis
Unit scope	What subset of CFSs are analysed?
Metric	Metric in which the association is expressed
Value	Value of the metric
Standard error	Standard error of the value of the metric
Significance	p-value of the value of the metric
Date of data extraction	The starting and ending day, month and year of data extraction
Descriptives	Descriptive statistics IV such as mean, standard deviation, median, degrees of freedom are collected for studies using population data

A.4 DATA CONVERSION AND INTERPRETATION FORM

Table A.3: Metric conversion table

Input metric(s)	Condition	Conversion	Output metric
Population Means	Population study	$\frac{\mu_{\text{post}} - \mu_{\text{pre}}}{\mu_{\text{pre}}}$	% Change
Sample: mean, standard deviations and size	Sample study	$\frac{(\bar{X}_{\text{post}} - \bar{X}_{\text{pre}})}{\sqrt{\frac{(N-1) \cdot SD_{\text{post}}^2 + (N-1) \cdot SD_{\text{pre}}^2}{(N+1) + (N-1)}}}$	Cohen's D
Z-statistic, sample size	—	$2 \cdot \frac{z}{\sqrt{N}}$	Cohen's D
t-statistic	—	$2 \cdot \frac{t}{\sqrt{N-1}}$	Cohen's D
Unstandardised B, standard deviations	Standard deviations provided	$B \cdot \frac{SD_X}{SD_Y}$	Beta
Unstandardised B, Grand mean	Grand mean provided	$\frac{B}{\text{Mean}}$	B*
Unstandardised B	Logit regression	e^B	Odds ratio

Table A.4: Metric interpretation table

Input metric	<i>Very small</i>	<i>Small</i>	<i>Medium</i>	<i>Large</i>	Interpretation based on:
Cohen's D	< ± 0.2	[± 0.2, ± 0.5]	[± 0.5, ± 0.8]	> ± 0.8	Cohen (2013)
Pearson's R	< 0.1	[0.1, 0.3]	[0.3, 0.5]	> 0.5	Cohen (2013)
Odds Ratio	< 1.44	[1.44, 2.48]	[2.48, 4.27]	> 4.27	Conversion from Cohen (2013)
% Change	< 0.1	[0.1, 0.3]	[0.3, 0.5]	> 0.5	We posit these as reasonable interpretations for % changes in <i>police response times</i>
B*	< 0.1	[0.1, 0.3]	[0.3, 0.5]	> 0.5	Analogous to % Change
Beta	< 0.1	[0.1, 0.3]	[0.3, 0.5]	> 0.5	Analogous to B*, under the assumption of equal standard deviations between the predictor and outcome variable

A.5 HYBRID SYNTHESIS



Table A.5: Hybrid Synthesis of Predictor Variable Subcategory Results

Category	Subcategory	Effect size interpretation	Direction	N	Mean internal validity
Exogenous	Day of the week	Large	None	1	3.00
	Day of the week	Very small	None	7	3.00
	Demand	Large	+	1	1.00
	Demand	Medium	+	1	3.00
	Demand	Very small	+	3	3.00
	Road type	Very small	–	3	3.00
	Road type	Very small	None	9	3.00
	Time of day	Large	None	1	1.00
	Time of day	Medium	None	6	1.00
	Time of day	Small	None	11	1.18
	Time of day	Very small	None	18	1.89
	Traffic	Very small	None	4	3.00
	Weather type	Very small	+	2	3.00
	Weather type	Very small	None	4	3.00
Human involvement	Caller characteristics	Large	None	1	1.00
	Caller characteristics	Small	None	18	3.00
	Caller characteristics	Very small	None	20	3.00
	Community characteristics	Large	–	1	3.00
	Community characteristics	Large	+	2	3.00
	Community characteristics	medium	–	3	3.00
	Community characteristics	Medium	+	1	3.00
	Community characteristics	Small	–	14	2.00
	Community characteristics	Small	+	12	1.67
	Community characteristics	Small	None	5	1.00
	Community characteristics	Very small	–	20	2.70
	Community characteristics	Very small	+	14	2.43
	Community characteristics	Very small	None	15	1.27
	Victim characteristics	Large	None	1	1.00
	Victim characteristics	Medium	None	1	1.00
	Victim characteristics	Small	–	3	3.00
	Victim characteristics	Small	None	3	3.00
	Victim characteristics	Very small	–	3	3.00
	Victim characteristics	Very small	+	6	3.00
	Victim characteristics	Very small	None	6	2.33
	Witness characteristics	Small	None	1	1.00
Incident	Incident characteristics	Large	None	4	2.00
	Incident characteristics	Medium	–	3	1.00
	Incident characteristics	Medium	None	4	1.50
	Incident characteristics	Small	–	3	1.67
	Incident characteristics	Small	+	1	3.00
	Incident characteristics	Small	None	2	2.50
	Incident characteristics	Very small	–	1	3.00
	Incident characteristics	Very small	+	2	2.00
	Incident characteristics	Very small	None	7	2.71
	Incident type	Large	None	10	1.60
	Incident type	Medium	None	16	1.25
	Incident type	Small	None	100	1.52
Innovation	Incident type	Very small	None	123	1.54
	AGDS	Medium	None	4	2.00

Category	Subcategory	Effect size interpretation	Direction	N	Mean internal validity
	AGDS	Small	None	2	2.00
	AGDS	Very small	None	5	2.00
	AVL	Very small	None	2	2.00
	Emergency button	Large	None	1	2.00
	Response unit type	Medium	None	2	2.00
	Response unit type	Small	None	2	2.00
	Response unit type	Very small	None	1	2.00
Kinematics	Travel distance	Large	+	2	2.00
	Travel distance	Medium	+	2	1.00
	Travel distance	Small	+	1	3.00
	Travel distance	Small	None	3	3.00
	Travel speed	Very small	+	1	1.00
Police strategy	Call priority	Large	–	2	3.00
	Call priority	Large	+	8	1.25
	Call priority	Large	None	11	1.00
	Call priority	Medium	–	1	1.00
	Call priority	Medium	+	1	1.00
	Call priority	Medium	None	2	1.00
	Call priority	Small	–	2	2.00
	Call priority	Small	+	1	1.00
	Call priority	Small	None	3	1.00
	Call priority	Very small	–	2	1.00
	Call priority	Very small	+	1	1.00
	Department size	Small	–	2	3.00
	Department size	Very small	–	2	3.00
	Department size	Very small	+	2	3.00
	Department size	Very small	None	1	3.00
	Differential response system	Large	None	1	1.00
	Differential response system	Very small	None	3	1.00
	Jurisdiction area size	Small	+	1	1.00
	Jurisdiction area size	Very small	None	3	3.00
	No. officers per response unit	Small	None	4	2.00
	No. officers per response unit	Very small	None	5	2.00
	Non-emergency number	Medium	None	2	2.00
	Non-emergency number	Small	None	1	2.00
	Non-emergency number	Very small	None	5	2.00
	On duty officers	Very small	None	2	3.00
	Patrol strategy	Medium	–	1	1.00
	Patrol strategy	Small	–	1	3.00
	Patrol strategy	Small	None	5	3.00
	Patrol strategy	Very small	None	9	3.00
	Shift	Small	None	1	3.00
	Shift	Very small	None	1	3.00
	Workload	Large	+	2	1.50
	Workload	Medium	+	1	1.00
	Workload	Small	+	2	1.50
	Workload	Very small	+	1	3.00
	Workload	Very small	None	1	3.00

A.6 CAUSAL MODEL OF POLICE RESPONSE TIME

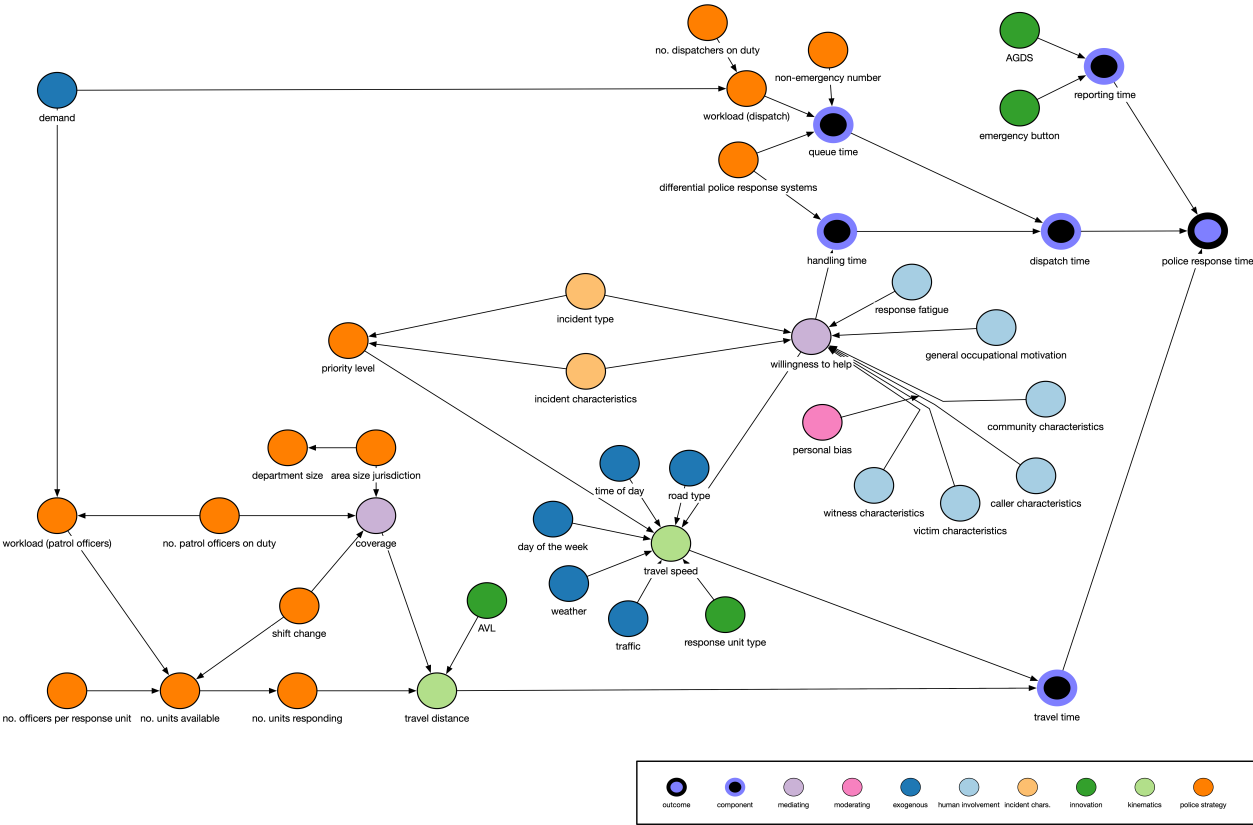


Figure A.2: Causal model of police response time

2

2

ON RE-EVALUATING RAPID RESPONSE

The standard view in policing research holds that rapid response has no impact on crime clearance. This perspective is primarily based on two landmark studies: the Kansas City Response Time Analysis Study (Bieck, 1977a) and the Police Executive Research Forum replication study (Spelman & Brown, 1981). While these studies, along with earlier research, found that when response is truly rapid, arrest rates increase, they dismissed the broader effectiveness of rapid response based on two key assumptions: (1) that only a small proportion of crimes are eligible for its benefits, as most are discovered too late, and; (2) that long citizen reporting delays negate any potential gains.

Through a critical reassessment of these studies, I challenge these assumptions. I argue that the dismissal of rapid response is based on weak evidence: flawed research designs, small samples, unreliable measurements, and an outdated empirical reality. Given these limitations, I contend that the standard view on rapid response should be reconsidered. Contrary to Sherman's (1997) assertion that further research would be 'a waste of tax dollars', there is ample empirical and theoretical justification for rigorously reassessing its potential impact.

2.1 INTRODUCTION

Modern policing can be described as consisting of two distinct forms: reactive and proactive policing (D. Weisburd & Majumdar, 2018). Reactive policing focuses on clearing crime and resolving disorder and, through incapacitation and specific and general deterrence, deterring future crime and disorder. Proactive policing, through general deterrence, solely focuses on deterring future crime and disorder. In practice, these two forms of policing are translated into three main activities, sometimes described as the three R's of the Standard Model of Policing: random patrol (proactive); rapid response (reactive); and reactive investigations (reactive) (Berkow, 2011). However, in the 1970s, four seminal studies in the US cast doubt on the effectiveness of all three of these pillars of the Standard Model of Policing.

First, the 1974 Kansas City Preventive Patrol Experiment (KCPPE) convinced police scholars and officials that preventive patrol did not reduce crime or disorder (Kelling, Pate, Dieckman & Brown, 1974). Second, the 1977 Kansas City Response Time Analysis Study (KCRTAS) (Bieck, 1977a) – and the subsequent observational replication study by the Police Executive Research Forum (PERF) with data from four more cities (Spelman & Brown, 1981) – convinced police officials and academics that 1) most crimes are discovered after the perpetrator has fled the scene, making rapid response ‘largely irrelevant’ and 2) for the minority of calls in which rapid response can make a difference, ‘... the impact of police response is often nullified by delays in citizen reporting.’ (Bieck, 1977a, p. 24). Lastly, the Rand Corporations’ 1977 Criminal Investigation Process (CIP) observational study ‘... provocatively concluded that solvability factors and uniformed patrol officers seemed more important to case resolution than investigative work.’ (Chaiken et al., 1977; as cited in Prince et al., 2021, p. 685).

Although research findings raised substantial doubts about all three pillars, the response of the scientific community varied considerably. Since its publication, the design, execution and the results of the KCPPE have been thoroughly debated – up to as recently as 2023 (see Farrington, 1983; Fienberg et al., 1976; R. C. Larson, 1975, 1982; Pate et al., 1975; Risman, 1980; Sherman & Weisburd, 1995; D. Weisburd et al., 2023; Zimring, 1978).

This debate, specifically the paper by Sherman and Weisburd (1995), led to novel police strategies (i.e., hot spot policing) and new evaluation designs and has translated into much follow-up research. Recent systematic reviews (i.e., Braga, Turchan et al., 2019; D. Weisburd et al., 2024) have identified over forty subsequent experimental and quasi-experimental studies focusing on the deterrent effects of proactive preventive patrol in various geographic scopes. In contrast, there has been no debate on the design and results of the studies on rapid response and reactive investigation. Subsequently, there have been no studies done focusing on the deterrent effects of rapid response and reactive investigation and few on the crime-clearance effects of these strategies (Prince et al., 2021; Verlaan & Ruiter, 2023a).

This essay aims to start this debate on rapid response. It does so by first reviewing the literature up to and including the KCRTAS and PERF studies, focusing on their main findings regarding the effect of response times on arrest and/or clearance rates. We show how nuanced findings and conditional effectiveness observed in these studies morphed over time into an oversimplified view that denounces rapid response as having no effect on arrest, and by extension, crime rates. By providing a critical analysis of the KCRTAS and PERF studies, we show this ‘standard view on rapid response’ to be inaccurate and outdated. Afterwards, we discuss recent advances in the study of rapid response. Based on this overview and

discussion of the available research on the effects of rapid response on arrest rates, we argue for the need for a re-evaluation of the rapid response strategy. And I propose a research design to do so.

2.2 EARLY RESEARCH ON RAPID RESPONSE

2

The first study on the effects of rapid response on arrest and clearance rates was conducted in Los Angeles in 1967 by Herbert Isaacs. The research identified a positive correlation between faster response times and higher arrest probabilities: one-minute police responses had a 30% arrest rate, compared to 25% for five-minute responses. However, the study was criticised (Spelman & Brown, 1981, pp. 18–23) for: aggregating results over all crime types; including non-response related arrests; not being able to distinguish between the direction of the effect (do fast response times increase probability of arrest, or does the probability of arrest increase fast response times?); unreliable response time estimates; and small sample size (30 on-scene arrests over 70 crime calls).

From the mid-1970s onward, several studies attempted to replicate Isaacs' findings in other cities with larger sample sizes. Clawson and Chang (1977) conducted a study in Seattle, using a larger, incident-level sample (507 on-scene arrests from 2,532 crime calls between April 1974 and March 1975). Their results supported Isaacs' conclusions but provided more nuance: no significant effect for dispatch time alone, a small effect for travel time, and a moderate effect for dispatch and travel times combined: 35% arrest rate for one-minute police response, 24% for 5-minute response. In 1978, Tarr replicated Clawson and Chang's study with an earlier dataset from Seattle (April 1973 to March 1974), analysing 1,214 arrests from 2,532 crime calls. Tarr confirmed the previous findings (and reported even higher statistical significance) and the observation by Clawson and Chang that the impact of rapid response diminished rapidly as time passed, with the highest likelihood of arrest occurring when response times were three minutes or less.

Beyond the Seattle studies, three more studies presented findings on rapid response, albeit as secondary results. First, in 1974, Holliday – while developing a methodology for radio car planning – analysed 6,114 calls (resulting 417 arrests) from New York and found '... there was a trend toward increasing arrests with shorter travel times' (Holliday, 1974, p. 31). Second, W. J. Brown and Lipsett (1976) studied the effect of police travel speed on response time by equipping a single patrol car in Ottawa with a speed recorder. After analysing 183 calls to which the patrol car responded, they state to find no correlation between response time (dispatch and travel times) and the (very broad) category 'outcome of calls', which contained: suspect/vehicle gone; all in order; settled with advice; settled by arrest/charge; summons issued; further investigation required (see p. 198). However, no information is provided on what percentage of calls concerned crime calls, the number of arrests realised, or how the analysis was carried out. Third, in 1976, Pate et al. analysed data from the KCPPE and found no significant relationship between police travel time and 'outcome of encounter'. However, to test this, investigators used data from three different surveys: the encounter survey, the 1972 community survey, and the 1973 community survey, which included 5, 10, and 14 arrests, respectively. Measured against today's standards, none of the secondary findings in these three studies demonstrate the methodological rigor, sample size or data quality necessary to draw reliable conclusions.

2.2.1 THE KANSAS CITY RESPONSE TIME ANALYSIS STUDY

During the same period when many of these studies were being conducted, a significant observational research effort was launched in Kansas City, Missouri, to examine the effects of rapid response on arrest rates ¹. Funded by the National Institute of Law Enforcement and Criminal Justice, the Kansas City Response Time Analysis Study ran from March 1975 to January 1976 ².

This landmark study aimed to produce valid and more reliable findings by adopting a realistic and comprehensive approach to studying police response time. Unlike previous research, data collection included not only dispatch and travel time but also citizen reporting time. Figure 2.1 presents a copy of the original Figure 2 – 1, detailing these response time components.

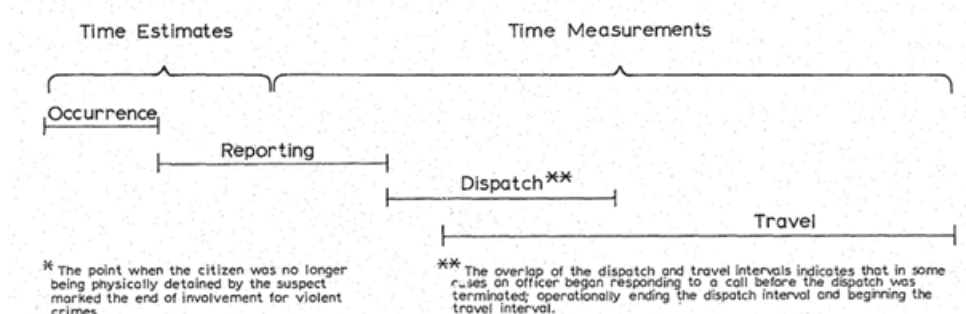


Figure 2.1: Conceptual model of response time components and response time intervals of the total response time continuum – adjusted from Figure 2-1 in Bieck (1977b, p. 15)

Moreover, the study conducted separate analyses based on the FBI Uniform Crime Report (UCR) distinction of violent and non-violent, and Part I and Part II crimes. It also introduced a distinction between *discovery crimes* (incidents identified after the perpetrators had left) and *involvement crimes* (incidents witnessed by a victim or bystander). Violent crimes were inherently classified as involvement crimes, while non-violent crimes could fall into either category, depending on the circumstances. Table 1 details these classifications and the crime types which they can contain, as well as their proportions in the Kansas City study. Lastly, in order to focus on ‘response-related arrests’, the analysis excludes arrests where suspects: (1) were apprehended by private citizens or security guards before police arrival, (2) were immobilised by injuries, (3) were identified through personal information (e.g., name or address) provided by victims or witnesses, (4) or voluntarily surrendered to police.

The study confirmed the earlier findings of Isaacs, Clawson & Chang, and Tarr, concluding that: ‘Rapid reporting can increase the chances of making an on-scene arrest for all types of Part I involvement crimes.’ (Bieck, 1977a, p. 23) Results showed that, depending on the crime type, arrests rates increased between 10 and 40% when fast responses could be realised. However, the conclusion also stressed two novel points, which later came to define the study.

¹And *witness availability, injury and citizen satisfaction*. However, we do not discuss these within the scope of the article.

²A study by Bertram and Vargo (1976) is also often cited in this literature, however, as it shows preliminary results of a small subsample of the KCRTAS (69 robberies leading to 6 arrests), we do not discuss it here separately.

Table 2.1: Incident classifications and crime types

		Involvement		Discovery	
		%	crime types	%	crime types
PART I (N = 949)	Violent	23%	Homicide; Rape; Robbery; Aggravated assault	0%	N/A
	Non-violent	14%	Burglary; Larceny; Motor vehicle theft	63%	Burglary; Larceny; Motor vehicle theft
PART II (N = 196)	Violent	30%	Non-aggravated assault	0%	N/A
	Non-violent	38%	Vandalism; Forgery, fraud, embezzlement	32%	Vandalism; Forgery, fraud, embezzlement

First, a large proportion of Part I crime calls (63%) was classified as discovery crimes, meaning they are not susceptible to the impact of rapid response, as they are discovered after the perpetrator has already fled the scene. Second, half of involvement crime calls (those crimes that are susceptible to rapid response) are only reported after more than 5 minutes, which stunts the potential impact of a swift response. Moreover, it was argued that these delays occur due to reasons that are not easily remedied.

2.2.2 THE POLICE EXECUTIVE RESEARCH FORUM (REPLICATION) STUDY

These results, particularly the findings on long citizen reporting delays, were met with scepticism by police chiefs in other cities, who questioned whether such delays also occurred in their jurisdictions or if the findings were unique to Kansas City (Spelman & Brown, 1981, p. xi). This prompted the Police Executive Research Forum to carry out a replication study across four cities in different states: Jacksonville, Florida; Peoria, Illinois; Rochester, New York; and San Diego, California (Spelman & Brown, 1981). The study focussed only on citizen delays in the reporting of Part I crimes and the effect these delays have on response-related on-scene arrest. In each city, data were collected every other day, between April 21, 1979 and January 15, 1980. In all four cities, the findings from Kansas City and those of Isaacs, Clawson & Chang, and Tarr on the effect of rapid response on arrests were once more replicated: the study found that the 13% involvement crimes that were reported in-progress had a 35% probability of arrest. This likelihood dropped sharply to around 7% for crimes reported 5 minutes after the event (Spelman & Brown, 1981, p. xxiii).

However, the two novel points highlighted in the Kansas City report were also replicated, and showed even more starkly: (1) on average, 75% of Part I crimes were classified as discovery crimes, meaning they were not susceptible to the effects of rapid police response (73% in FL, 69% in IL, 71% in NY, and 83% in CA) (p. 72); (2) the median reporting delay for involvement crimes ranged from 4 to 5.5 minutes across the cities, with an average of 46% of these crimes reported after 5 minutes. Notably, the report states that the likelihood of arrest after a 5-minute delay was almost the same as after a 60-minute delay – low. Furthermore, the report states that police have limited ability to reduce reporting delays, as citizens often delay for what they consider valid reasons which are unlikely to change.

Spelman and Brown conclude:

The findings are clear. Most serious crimes reported to the police are discovery crimes for which there is virtually no chance for response-related arrests. For the remaining crimes, those in which there is citizen involvement, the citizen must call the police within one minute, or the likelihood of response-related arrest drops dramatically.

2.2.3 SUMMARIZING FINDINGS OF EARLY RAPID RESPONSE STUDIES

In summary, early studies consistently found that faster police response times were associated with higher arrest probabilities, though the effectiveness was conditional on factors like call type and characteristic (involvement or discovery), whether it was a response-related on-scene arrest and on citizen reporting delays. These patterns are summarised in Table 2.2.

Table 2.2: Arrest rates for n th-minute response (component; reporting (R), dispatch (D), travel (T)) time in early rapid response studies

Study	Compo.	N _{calls}	N _{arrest}	0 th	1 st	2 nd	3 rd	4 th	5 th
Isaacs	D&T	70 (crime)	30	–	30%	27%	26%	25%	25%
Clawson & Chang	D&T	2532 (crime)	507	–	35%	27%	27%	25%	24%
Tarr	D&T	2532 (crime)	1214	32%	26%	24%	24%	26%	22%
Holliday	T	6114 (crime)	417	11%	9%	5%	5%	11%	5%
Brown & Lipsett	T	183 (all)	?	–	–	–	–	–	–
Pate et al.	T	340 (all)	16	–	–	–	–	–	–
Betram & Vargo	T	69 (robberies)	6	–	–	–	–	–	–
Bieck	R&D&T	943 (involv.)	27 (rros)	–	44%	37%	32%	31%	–
Spelman & Brown	R	913 (involv.)	259	35%	18%	10%	–	–	7%

2.3 THE STANDARD VIEW OF RAPID RESPONSE

However, over time, the findings from these early studies on rapid response were gradually simplified as they were disseminated through academic and policy discourse. While the original research emphasised the conditional nature of rapid response effectiveness – distinguishing between involvement and discovery crimes and highlighting the impact of citizen reporting delays – subsequent interpretations increasingly fixated on its limitations without fully accounting for these complexities.

This resulted in a standard view of rapid response that inaccurately characterised it as having no effect on arrest rates or crime levels. One of the earliest formulation of this standard view can be found in Lawrence Sherman's chapter, 'Policing for Crime Prevention' in the highly influential report to the US congress, *Preventing crime: What works, what doesn't, what's promising*:

The conclusion that reduced response time will not reduce crime is based on strong but indirect evidence. The evidence is strong because it is based on large samples, careful measurement, and a replicated research design in five diverse cities showing little variation in arrest rates by police travel time, the main factor that tax dollars can affect. It is indirect because an experimental test of the

effects of reduced police travel time on city-wide arrest and crime rates has never been conducted. Yet there is neither empirical nor theoretical justification for such an expensive test. Given the strong evidence of citizen delays in reporting involvement crimes, and the small proportion of serious crimes that feature direct victim-offender involvement, further tests of this theory seem to be a waste of tax dollars. (Sherman et al., 1997, p. 11)

The view articulated here by Sherman has profoundly shaped the discourse on rapid response, becoming a cornerstone of policing literature over the past few decades. It has been echoed (in one form or another) in many influential academic texts on the topic, for example: the Consensus Study Report *Fairness and Effectiveness in Policing* for the National Research Council (2004, pp. 226–227), a highly cited article on what police can do to reduce crime, disorder and fear in the *Annals of the American Academy of Political and Social Science* by D. Weisburd and Eck (2004, p. 50), a highly influential review of deterrence by Nagin (2013, p. 237), a foundational article on the rise of evidence-based policing by Sherman (2013, p. 405), and the recent Consensus Study Report *Proactive Policing* for the National Academy of Sciences, Engineering, and Medicine (2018, p. 28).

To challenge this standard view, I will present a detailed review of the Kansas City and Forum Cities response time analysis studies, which serve as its empirical foundation. By revisiting the original data and methodologies underlying these studies, I demonstrate that the limitations of rapid response have been overstated or are outdated. Moreover, I argue that – contrary to Sherman’s conclusions – there is ample empirical and theoretical justification for renewed (experimental) research into its effects on arrests and crime levels.

2.4 A CRITIQUE OF THE STANDARD VIEW

In critiquing the standard view of rapid response, we first deal with the claims that the Kansas City and Forum cities studies provide strong evidence, based on large samples, of the lack of an (indirect) effect of police response time on crime levels. Afterwards, in order to critique the statement that these studies are based on careful measurements, we turn to a more expansive critique of these studies.

2.4.1 NO (INDIRECT) EFFECT

As demonstrated in the discussed literature, this premise directly contradicts all findings of earlier studies. Research by Isaacs (1967), Clawson and Chang (1977), Tarr (1978), and the investigators in Kansas City consistently reported a relationship between police response time and arrest rates. While Spelman and Brown (1981) did not directly measure police response times (focusing solely on citizen reporting delays), they explicitly state that for property crimes reported in-progress (i.e., with zero reporting delay), ‘chances were excellent of effecting a response-related arrest if police response was prompt’ (1981, p. 68). Similarly, for violent crimes reported in-progress, they conclude that ‘fast police response would very likely result in arrest’ (1981, p. 70).

While the proportion of cases susceptible to rapid response may be limited (I will later argue this remains to be seen), these findings nevertheless indicate that response time has a measurable effect on arrest rates, and by extension, crime levels.

2.4.2 STRONG EVIDENCE

Sherman describes the evidence supporting the lack of a relationship between police response time and arrest rates (and crime levels) as ‘strong’. However, Sherman’s own application of (what would later become known as) the MSMS reveals a significant contradiction to this classification. In the report to Congress, he assigns findings from both the Kansas City and Forum Cities studies a Level 2 internal validity score (Sherman et al., 1997, Chapter 8, p. 12). According to the MSMS – co-authored by Sherman – Level 2 studies should be classified as ‘inadequate and uninterpretable’ (Farrington et al., 2002, Chapter 2, p. 17).

Furthermore, I would even argue that a Level 2 classification for these studies is difficult to justify. According to the MSMS, the key distinguishing feature of Level 2 studies is their ability to establish causal order. Yet the Kansas City or Forum Cities studies fail to establish such causal order. This failure is explicitly acknowledged by Spelman and Brown, who write: ‘A distinct but related problem that has not been handled by researchers is the problem of two-directional effects’ (1981, Appendix C, p. 117).

2.4.3 LARGE SAMPLES

The Kansas City study included 949 Part I crimes and 196 Part II crimes in its dataset. However, only 27 response-related on-scene arrests were recorded for Part I crimes and just 11 arrests for Part II crimes. The study by Spelman and Brown included data from Jacksonville, Peoria, Rochester, and San Diego. The total number of crime calls analysed was 913, 508, 614, and 765, respectively. Despite oversampling response-related on-scene arrests, these samples were limited to 259 in Jacksonville, 63 in Peoria, 75 in Rochester, and 249 in San Diego.

Given these figures, we can conclude that the sample sizes used to determine the effect of response time on arrest rates in these studies are relatively small – especially by today’s standards.

2.4.4 CAREFUL MEASUREMENTS: RELIABILITY AND VALIDITY

To evaluate the claim of ‘careful measurements’, we must take a closer look at the methodologies employed in the Kansas City and Forum Cities studies. Reading the reports, one gets the impression that these measurements were indeed taken with diligence. However, the claim of ‘careful measurements’ here also conveys the impression that they are reliable and valid.

I argue that this is not the case. Moreover, we argue that their validity in modern society is severely limited. In the following paragraphs, we evaluate the reliability of three key measurements: reporting time, crime classification (involvement/discovery), and response-related on-scene arrests. Afterwards, we will assess the contemporary validity of the measures for citizen reporting delay and proportion of involvement crimes in the context of modern societies.

RELIABILITY OF REPORTING TIME IN KANSAS CITY AND THE FORUM CITIES

Reporting time is notoriously difficult to measure because determining the exact moment a crime occurs is often challenging, and, even when it can be pinpointed, it is rarely recorded by anyone involved. As a result, both KCRTAS and PERF study had to rely on the best available data: estimates of response delays provided by victims, witnesses, and bystanders. These estimates were collected during follow-up interviews which took place, on average,

two weeks after the crime occurrence. To assess the reliability of these estimates, researchers also asked respondents to estimate a known time interval: the time between when the police were called and when the first officers arrived.

In the Kansas City study, researchers found an R^2 value of 0.02, indicating that only 2% of the variation in actual response time could be explained by perceived response time – concluding in response that ‘it suggests that citizen perceptions of times cannot be considered reliable’ (Bieck, 1978, p. 92). In the PERF study, researchers found that only 4.3% of the variation in actual response time could be attributed to perceived response time (Spelman & Brown, 1981, Appendix B, p. 82). The authors noted that while the average of a large number of estimates would be close to the actual duration, individual estimates varied widely. For instance, for a 6-minute response time, two-thirds of the estimates fell between 3.5 and 8.5 minutes, with 5% of estimates exceeding 11.5 minutes or being less than 30 seconds (Spelman & Brown, 1981, Appendix B, p. 86). Moreover, the researchers describe systematic biases towards rounding off to the nearest minute or the nearest five minutes.

These inefficiencies of the estimates led Kansas City researchers to note ‘... citizen perceptions of times cannot be considered reliable’ (Bieck, 1977b, p. 92). For the Forum cities, Spelman and Brown explored various methods to improve these estimates, such as adjustment of individual estimates and the deletion of outliers. However, they ultimately conclude that ‘no worthwhile adjustment is possible’ (Spelman & Brown, 1981, Appendix B, p. 106) and ‘identification and deletion of outliers is impractical’ (Spelman & Brown, 1981, Appendix B, p. 112). Hence, we deem the reporting measurements not to be reliable.

RELIABILITY OF THE PROPORTION OF INVOLVEMENT CRIME IN THE FORUM CITIES

In both the Forum cities and Kansas City, if a citizen saw, heard, or became involved in the crime at any point during its commission, and if the citizen subsequently reported the crime to the police, it was classified as an involvement crime. However, if the citizen saw, heard or became involved in the crime but did not report it to the police (immediately), and the crime was later discovered and reported, it was classified as a discovery crime (Bieck, 1977b, p. 2). This process opens the door to underestimation of involvement crimes from two sources: (1) incomplete data registrations by reporting officers and (2) failures to report crimes by witnesses.

1. In Kansas City, the classification of crimes as either involvement or discovery was made by trained field observers at the scene, and the accuracy of these classifications was regarded as high. In contrast, in the Forum cities, the categorisation was based on crime reports which varied greatly as a result of variations in ‘instructions given to officers [filling them in; author’s insertion]’ (Spelman & Brown, 1981, 254, Appendix E) ³. On-site research assistants reviewed these crime reports and, if explicitly registered that the caller saw, heard or was involved, classified the crime as ‘involvement’. In any

³Crime reports consisted of ‘all of the various reporting forms used in the study sites to record information about the incidents of criminal activity’ (Spelman & Brown, 1981, Appendix A, p. 20). As no copies of these forms are present in the Appendix A and ‘involvement’ originated from the Kansas City study, it is unlikely this categorisation was present in these existing forms. Instead, it is likely that the research assistant on site made a judgment call on whether a crime concerned a ‘discovery’ or an ‘involvement’ crime, based on the registrations made by the officer(s) handling the scene.

other case, the call was classified as 'discovered'⁴. We deem it likely that missing information on crime reports caused both inconsistent estimates in discovery crime proportion among the Forum cities (73% in Florida, 69% in Illinois, 71% in New York, and 83% in California), as well as an overestimation when compared to the more accurate measurements in the Kansas City study, where 62% of crimes were classified as discovery crimes.

2. This classification process implies that the underestimation of involvement crimes is proportional to the number of crimes which were witnessed, but were never reported (or not before being discovered by someone else). Given that this potential bias is present in both studies and both studies collected detailed reasons for reporting delays from hundreds of citizens, some inferences could be drawn regarding the extent of this threat: Researchers in Kansas City observed at least six patterns of reporting delay that likely also accounted for failures to report completely: delay due to apathy (6.5% of cases), delay due to being unsure about police assistance (7.4%), delay due to contacting security (5.1%), delay due to public communication problems (22.2%), delay due to police communication problems (6.3%), delay due to fear or emotional shock (10.5%). Although these percentages are not mutually exclusive; their relative abundances suggest a substantial potential threat to underestimate the proportion of involvement crimes, and thus overestimate the proportion of discovery crimes.

RELIABILITY OF RESPONSE-RELATED ON-SCENE ARRESTS IN THE FORUM CITIES

The same underlying issue of incomplete data registration also affects the construct validity of response-related on-scene arrests. In Kansas City, data on response-related on-scene arrests were collected by trained field observers at the scene of the crime, which minimised the risk of misclassification. Again, in contrast, in the Forum cities, crime reports were the sole source of information. If the crime report mentioned one or more of the excluded types of arrests – such as citizen arrests, the perpetrator being immobilised, identified, or having surrendered – the crime was considered a nonresponse-related on-scene arrest, in any other case, it was considered a response-related on-scene arrest. As a result of this method, Spelman and Brown stated, 'The percentage of crimes resulting in response-related arrest is overestimated in our study' (Spelman & Brown, 1981, Appendix D, p. 144).

The degree of overestimation varies between the Forum cities, likely due to differences in how officers filled in crime reports and the instructions they received. In Kansas City, 24% of on-scene arrests were classified as response-related (38 out of 161), while in the Forum cities, this percentage ranged from 59% to 69%. Given the similarities between Kansas City and the Forum cities in other aspects of the study, this discrepancy suggests an overestimation of response-related arrests by a factor of more than two in the Forum cities.

⁴Although discovery rates were based on crime reports, sampled respondents were asked explicitly in the interview: 'Did you see, hear or become involved in this crime at any time while it was happening? No; Yes (Skip to Q.19) – INVOLVEMENT CRIME)' (Appendix G, Data Collection Instrument, Section 1, Question 12). It is unclear why the involvement/discovery crime proportions obtained in this question are given priority over the judgement calls based on crime reports or, at least, recorded anywhere in the report. The absence is particularly noticed in Table 1 in the report.

2.5 CONTEMPORARY VALIDITY

At the time of the KCRTAS study and its replication in the four PERF cities, the technological landscape was vastly different from today. Likely no citizen had access to mobile phones, only about 40% of U.S. households had access to a fixed telephone line (World Bank, 2023), and public pay phones were still commonplace. Only one of the police departments (Peoria) used a central police emergency number (911) and none had access to Automatic Vehicle Location (AVL) systems, Global Positioning System (GPS) navigation, Computer-Aided Dispatch (CAD) systems, Caller ID, or any form of personal computers. Personal security systems for homes, shops, and cars were far less advanced and adopted, and the ‘crime-drop’, especially for (discovery) property crimes, had not yet occurred. Today, modern mobile phones with GPS and high-resolution cameras enable citizens to report crimes instantly and provide crucial real-time information. These technological advancements have transformed the environment in which crimes are reported and responded to, altering the mechanisms and moderators observed in the original studies.

In this section, we will examine the findings, mechanisms, and moderators from the original studies and assess their contemporary validity in light of these technological changes. In particular, we assess the validity of the claims that (1) the majority of crimes are discovery crimes and thus not susceptible to be impacted by rapid response, and (2) that for the involvement crimes, reporting delay by citizens blunts any potential effect of rapid response.

2.5.1 CONTEMPORARY VALIDITY OF PROPORTION OF INVOLVEMENT CRIMES

Since the 1970s, there has been no research (to my knowledge) collecting data on the proportions between discovery and involvement crimes, making it impossible to directly determine their current ratios. However, we do know the development of the UCR Part I crime categories in these studies. And we do know that since the turn of the millennium, the overall crime rate has dropped significantly – a phenomenon often referred to as the ‘crime-drop’. Based on this information, we attempt here to approximate these ratios.

Table 2.3 is based on Table A-4 from Appendix A in Spelman and Brown (1981, p. 9) and provides the crime counts and rates per 100,000 for each UCR Part I crime category across various cities for 1976, which were at the time drawn from official FBI UCR statistics. Table 2.3 also includes the crime rates per 100,000 for 2019, based on statistics from the same FBI UCR, and the percentual changes between 1976 and 2019 for each crime category in these cities. For Kansas City, the total crime-drop comes out to 38%. However, interestingly, if we look at the change in violent crime and property crime, we see that this crime-drop is fully reliant on the drop in property crime, and that violent crime has actually increased over the intermediate timeframe. As all discovery crime is property crime, and all involvement crime is violent crime, this logically has implications for their proportions, and thus the share of involvement crime.

To better understand the contemporary proportions of discovery and involvement crimes, we conducted a thought experiment: What would the share of involvement crimes look like if the KCRTAS had been conducted in 2019? To answer this, we make a single key assumption: within each Part I crime category (rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft), the proportion of crimes classified as discovery or involvement remains constant over time. Using the percentage changes in crime rates between 1976 and 2019, we infer the current proportions of discovery and involvement crimes for each category.

Table 2.3: Crime rates per 100,000 in Kansas City in 1976 and in 2019 according to the FBI UCR statistics

Crime Category	Crime Type	1976 Rate	2019 Rate	% Change
Violent	Rape	69	70	+1.40%
	Robbery	475.3	290.9	-38.80%
	Aggravated Assault	537.1	1040.2	+93.70%
	Subtotal	1081.4	1401.1	+29.56%
Property	Burglary	2490.1	619	-75.10%
	Larceny	4303	2470.5	-42.60%
	Motor Vehicle Theft	663.1	766.4	+15.60%
	Subtotal	7456.2	3855.9	-48.29%
Total		8537.6	5257	-38.42%

In both the 1976 and 2019 studies, violent crimes (forcible rape, robbery, aggravated assault) are consistently categorised as involvement crimes. Property crimes (burglary, larceny, motor vehicle theft) can fall into either discovery or involvement categories, depending on whether the crime was witnessed or discovered after the fact. Table 2.4, based on Table 1-1 from Bieck (1977b, p. 3), outlines the number and percentage of involvement crimes for each crime category in Kansas City in 1976.

Table 2.4: Discovery and Involvement Part I crime counts in 1976

Crime Type	Crime Count	Discovery	Involvement	% Involvement
Homicide	0	0	0	-
Forcible Rape	10	0	10	100%
Robbery	127	0	127	100%
Aggravated Assault	84	0	84	100%
Burglary	352	302 ⁵	40	11%
Larceny	297	206	91	31%
Motor Vehicle Theft	79	74	5	6%
Total	949	582	357	38%

The results of this thought experiment are summarised in Table 2.5, which presents the inferred discovery and involvement counts for Part I crimes in Kansas City for 2019. Following this thought experiment, the proportion of involvement crimes in a contemporary Kansas City Response Time Analysis Study would be 53%, instead of the 38% observed in 1976. This shift would make involvement crimes the majority in 2019. We view this estimate as conservative, given the advancements in surveillance and communication technologies that have likely increased the detection of property crimes in progress. Further research, leveraging modern data collection methods and surveillance technologies, would provide a more precise understanding of these dynamics. Nevertheless, these dynamics do point at a bigger share of involvement crimes, and thus of a bigger potential impact for rapid response.

Table 2.5: Inferred Discovery and Involvement Part I Crime Counts in 2019

Crime Type	Count ₁₉₇₆	% Change	Count ₂₀₁₉	% Involv. ₂₀₁₉	Involv. ₂₀₁₉	Disc. ₂₀₁₉
Forcible Rape	10	+1.4%	10	100%	10	0
Robbery	127	-38.8%	77	100%	77	0
Aggr. Assault	84	93.70%	163	100%	163	0
Burglary	352	-75.1%	88	11%	10	78
Larceny	297	-42.6%	169	31%	52	117
Motor Veh. Theft	79	15.6%	91	6%	5	86
Total	949	–	599	53%	319	280

2.5.2 CONTEMPORARY VALIDITY OF CITIZEN REPORTING DELAYS

The KCRTAS study specifies eight patterns and five problems in reporting crime to the police. Below, we discuss the contemporary validity of these issues. The eight patterns, detailed in Table 2.6, provide insight into common behaviours that led to delays in contacting the police. The table includes pattern descriptions, the frequency of their occurrence in Part I and Part II crimes, and their level of validity in modern society.

Besides patterns, the report also identifies five ‘problems’. To set the stage, it is worth noting these five problems were at the time described as ‘unavoidable hindrances’ (Bieck, 1977b, p. 84). Table 2.7 details these problems, their descriptions, and the frequency of their occurrence in Part I and Part II crimes, and their level of validity in modern society.

2.5.3 CONCLUSION OF THE CRITIQUE

Reviewing the nearly one and a half thousand pages written on the KCRTAS and PERF studies reveals a tremendous research effort. As such, it is understandable why these two studies have long been regarded as the ‘best available evidence’ on the effects of rapid response. However, given the discussed issues around reliability and contemporary validity of the key measurements (citizen reporting delay, proportion of involvement crime, and response-related on-scene arrests), and because of their weak correlational designs, these studies must ultimately be recognised for what they are: offering weak and outdated evidence.

Fortunately, research into rapid response and (the effects of) police response time has not completely halted. Here, we provide a discussion of the research on the effects of police response time on arrest/crime clearance rates, published since the KCRTAS and PERF studies.

Table 2.6: Patterns in Reporting Crime, source: Bieck, 1977b, p. 82; Bieck, 1980b, p. 65

Pattern	Description	Part I	Part II	Contemporary Validity
Talking to another person	Delayed while seeking advice, permission to use a landline, or additional information before contacting police.	47%	40%	Moderate – While advice-seeking persists, mobile phones reduce the need to ask for permission and enable direct contact with the police.
Investigating the incident scene	Delayed while assessing damage, searching for missing property, or determining what happened.	18%	6%	Low - People may still investigate, but mobile phones allow reporting while assessing the situation.
Telephoning another person	Delayed while calling others for advice, information, or support before contacting police.	10%	9%	Moderate - While advice-seeking persists, increased awareness of emergency numbers reduces the need for intermediary calls.
Waiting or observing the situation	Delayed to gather more information or to decide whether police intervention was necessary.	9%	3%	Moderate - Citizens can now observe and report simultaneously using mobile technology.
Unsure about police assistance	Delayed due to doubts about whether police intervention would be helpful or necessary.	7%	8%	Unclear – Public campaigns may have had influence, although not clear.
Chasing the suspect	Delayed while pursuing the suspect, sometimes leading to apprehension.	7%	8%	Moderate - Mobile phones enable simultaneous reporting and pursuit, reducing delays.
Apathy	Delayed because the caller didn't consider the incident important or didn't want responsibility.	7%	11%	Unclear - Changes in societal attitudes toward responsibility and importance are difficult to assess.
Contacting security	Delayed due to company policies requiring consultation with security guards or supervisors before police.	5%	3%	Unclear – We are unaware of research into this.

Table 2.7: Problems in Reporting Crime, source: (Bieck, 1977b, p. 82; Bieck, 1980b, p. 65)

Problem	Description	Part I	Part II	Contemporary Validity
Public communication problems	Delays caused by lack of access to phones, inoperative phones, or busy lines, requiring individuals to find payphones.	22%	18%	Low - Mobile phones provide immediate access, eliminating reliance on payphones or landlines.
Uninformed or misinformed	Witnesses not immediately aware of a crime or mistakenly assuming someone else reported it.	11%	9%	Moderate - Mobile phones allow citizens to stay present at the scene, reducing the likelihood of miscommunications.
Fear or emotional shock	Fear of reprisal or emotional shock leading to hesitation or inability to report promptly.	11%	13%	High - Emotional reactions are unlikely to have changed much.
Police communication problems	Issues such as unanswered calls, busy lines, or slow connections when contacting police.	6%	5%	Low - Modern 911 systems and improved dispatch technology ensure faster and more reliable communication.
Injury	Victims or witnesses delayed due to being injured and/or immobilised.	6%	2%	Low - Mobile phones allow for requesting medical/police assistance, while immobilised/injury.

2.6 SUBSEQUENT RESEARCH ON THE EFFECTS OF RAPID RESPONSE ON ARREST

In 1983, Cordner et al. published a study in response to the findings in the KCRTA and PERF studies. The study uses months instead of individual incidents as the unit of analysis and adds a crude estimate for ‘discovery time’ to arrive at what it terms ‘true response time’. The study regresses monthly on-scene arrest for all index crimes and for burglary separately against monthly average primary response time. The results showed a significant moderate effect for index crimes in general, and a smaller and statistically insignificant effect for burglaries. Due to the use of aggregated data and the crude estimation of discovery time, the study is considered to have a weak research design, which likely contributed to its limited impact in the literature.

Since the early 2000s, Timothy Coupe together with various co-authors has published numerous analysis of the effect of police response time on arrests rates for in-progress burglaries (see Blake & Coupe, 2001; Coupe, 2016, 2019; Coupe & Blake, 2005, 2011; Coupe & Girling, 2001; Coupe & Griffiths, 2000). These studies consistently find significant effects of varying magnitudes, with Coupe’s work emphasizing how the size of the effect depends on differing response conditions – such as the number of officers responding – and incident characteristics – such as the time of day. For instance, Coupe found that when two or more police units arrive on the scene within five minutes, the probability of arrest increases from 9.5% to 27%. However, all analyses appear to be based on the same dataset collected in the mid-1990s in the West Midlands (UK), limiting the additional value of the successive analyses.

In 2012, Cihan et al. published a study that can be viewed as marking the beginning of a new era in research on police response times. Analysing 4,917 in-progress burglaries, resulting in 399 arrests (8%), this study was the first to fully utilise a large administrative dataset of individual incidents obtained from a computer-aided dispatch system without sampling. Descriptive statistics revealed that 75% of incidents were responded to within 5 minutes and approximately half of the incidents were responded to (dispatch and travel time) within 3.5 minutes. The analysis found a significant relationship between police response time and arrests, with an odds ratio of 0.91, meaning that for every minute the police arrive sooner, the probability of an arrest increases by 9%.

In 2018, Blanes i Vidal and Kirchmaier published a study using advanced econometric method on population-level police registration data from the Greater Manchester Police. This comprehensive administrative dataset contained over 300,000 calls for service, both priority 1 and 2, of which 38% were classified as ‘cleared’⁶.

⁶In a footnote in the paper described as:

‘Although we refer to “crime clearance”, the corresponding terminology in the U.K. is “crime detection”. The official definition of “detection” is as follows: “A sanctioned detection occurs when (1) a notifiable offence (crime) has been committed and recorded; (2) a suspect has been identified and is aware of the detection; (3) the Crime Prosecution Service evidential test is satisfied; (4) the victim has been informed that the offence has been detected, and; (5) the suspect has been charged, reported for summons, or cautioned, been issued with a penalty notice for disorder or the offence has been taken into consideration when an offender is sentenced.” See <http://data.london.gov.uk/dataset/percentage-detected-and-sanctioned-offences-borough>’ (Blanes i Vidal & Kirchmaier, 2018, p. 863)

In order to account for endogeneity in police response times ⁷ the author exploited (as an instrument) discontinuities in response times around police territorial division boundaries ⁸. The authors find that faster response times were strongly associated with higher crime clearance rates. Their estimates indicated that a 10% decrease in response time led to a 4.7 percentage point increase in clearance rates. Interestingly, the effect was not limited to very short response times, which had been observed in earlier studies and which could be regarded as the main argument against generalised rapid response to calls for service. In terms of research design, this study would rank as a Level 4 on the Maryland Scale of Scientific Methods for internal validity.

Like the paper by Blanes i Vidal & Kirchmaier, there are two more papers that use advanced econometrics methods on police data. However, these studies only indirectly report on the effect of rapid response on crime clearance rates. The first is by Mastrobuoni (2019), who analyses data from commercial robberies in Milan during 2009 and 2010. Using the quasi-experimental setting of police shift changes – periods when officers leave their beats to travel to headquarters – the study examines the impact of reduced prompt police presence on crime outcomes. Mastrobuoni finds that clearance rates for robberies decline sharply during these shifts change intervals, dropping from 13.6% to 9.1%. This reduction is implicitly linked to delays in police response.

The second is by S. Weisburd (2021) who investigates the broader effects of police presence on crime rates using detailed administrative data from the Dallas Police Department. Employing an instrumental variable approach, Weisburd examines changes in police presence driven by officers being diverted to respond to 911 calls, which she attributed to the ‘rapid response philosophy’. She finds that a 10% reduction in police presence at a location leads to a 7% increase in crime at the abandoned beat. However, Weisburd also notes that as police presence increases proportionally in beats with 911 incidents, the overall effect on crime is neutral.

In 2023, Rief and Huff revisited the role of police response time on arrest outcomes, leveraging data collected during a body-worn camera (BWC) experiment involving over 1.2 million priority 1, 2 and 3 crime calls (nested within 810 officers) and over 350,000 arrests (28.70%). They employ hierarchical logistic regression models and surprisingly find a significant odds ratio of 1.02, meaning there was an increase in the odds of arrest of 2% for every minute police arrive later on the scene.

Although the study’s findings contradict all earlier research, it offers no clear mechanism for this anomaly. A potential explanation lies in that the study does not differentiate between

⁷ ‘Crimes assigned a higher priority could be those with an ex ante higher or lower clearance difficulty. Furthermore, these crimes are likely to receive more police investigative resources ex-post.’ (p. 857)

⁸ The instrument is described in more detail in the paper as follows: ‘To credibly identify causal effects, we first take advantage of a particular feature of our police force: the fact that, when a call for service is received, the responding officer often departs from the station where they are based, rather than from a random point along a patrolling route. Therefore, crime scenes closer to a response station are reached more quickly following a call for service. Unobserved determinants of clearance difficulty at the area level might, however, correlate with distance to the response station. To account for this, we exploit the partition of the Greater Manchester territory into eleven operationally distinct divisions. This implies that crime scenes within a small local area, but on different sides of a division boundary, are served by separate response stations, which may be at very different respective distances. In our empirical specifications we control for the “local area” by introducing a large number of geographical cell indicators, each representing an area of 0.185 km² (556 m by 332 m). Variation in distance to the division response station, which we use as an instrument for response time, is then largely due to crime scenes in the same geographical cell falling on separate sides of division boundaries.’ (Blanes i Vidal & Kirchmaier, 2018, p. 857)

response-related arrests and other arrests, such as those involving known suspects; crimes where the suspect is known are more likely to result in arrests but also tend to have longer response times, as these arrests often rely more on careful planning and communication rather than rapid response. If such non-response-related arrests comprise a substantial portion of the total arrest pool, this could bias the observed effect, leading to a spurious positive association. Given that 80% of analysed calls are priority 2 or 3⁹ – priority levels for which the Police do not expect rapid response to make a worthwhile difference on clearing the crime – and that the odds of arrest are 13% higher for priority 3 calls than for priority 1 calls, it is plausible such bias exists. This bias is exacerbated by the study's reliance on linear models, which assume a constant relationship between response time and the log-odds of arrest. This linear assumption not only oversimplifies the well-documented non-linear relationship where the greatest effects of rapid response occur within the first few minutes and diminish thereafter. It also effectively allows non-response-related arrests to influence the estimated effect of rapid response cases, artificially inflating the observed relationship.

As such, while this is an intriguing finding – particularly in the context of BWCs – on the relationship between arrest rate and police response time for various crimes, it has limited validity in evaluating the causal effect of rapid response on response-related, on-scene arrests. If anything, this study highlights the continued need for informed, highly specific (non-linear) analyses into the effects of rapid response on arrest – an approach championed in the landmark studies of the 1970s.

2.7 CONCLUDING ON THE STATE OF RAPID RESPONSE RESEARCH

Early research from the 1970s into rapid response provided foundational insights to effects of rapid response on arrest rates but was constrained by small sample sizes, imprecise measurements, and weak research designs. While they highlighted key dynamics – such as the impact of citizen reporting delays and the proportion of crimes susceptible to rapid response – their methodological limitations prevent definitive conclusions. Moreover, these studies are now outdated, reflecting a technological and policing landscape that no longer exists.

More recent studies, in contrast, leverage large administrative datasets and advanced statistical techniques. However, the analysis of these large administrative datasets come with a cost, namely the loss of specificity required to isolate the causal effects of rapid response. This limitation, as we have argued, can lead to spurious relationships. Despite these shortcomings, both old and new research provide enough indications that rapid response can influence arrest rates to justify renewed investigation.

The gold standard for testing causal claims is the randomised controlled trial (RCT). However, an RCT on rapid response is difficult to implement, expensive, and raises ethical concerns; as delaying emergency response times for experimental purposes is not justifiable. Given these challenges, we propose a non-experimental research design that builds on lessons from past studies while using modern data and methodologies to provide the most rigorous possible evidence on the impact of rapid response on arrest.

⁹Moreover, even a proportion of priority 1 calls may not depend on rapid response if the perpetrator's identity is already known, such as in domestic violence cases.

2.8 FUTURE RESEARCH ON THE IMPACT OF RAPID RESPONSE

Future research on the impact of rapid response should begin with a clear understanding of the causal mechanisms underlying police response time and arrest outcomes. To facilitate this, we propose a preliminary causal framework informed by the reviewed research and represented as a directed acyclic graph (DAG) constructed using the DAGitty software package (Textor et al., 2016).

As outlined in Figure 2.2, an arrest ultimately depends on whether police are in a position to apprehend a perceived perpetrator and whether they decide to make an arrest. Response time influences both of these processes. Faster response times increase the likelihood of catching a suspect either red-handed or while fleeing. At the same time, faster responses may also increase the likelihood of arrest due to heightened officer adrenaline or urgency-driven decision-making. Response time itself is shaped by two primary forces: police officer discretion, which includes driving faster or prioritizing calls, and structural response constraints, such as call priority levels, the use of lights and sirens, and the number of responding officers. While structural response constraints can often be directly measured and accounted for, officer discretion remains largely unobservable, introducing significant endogeneity concerns.

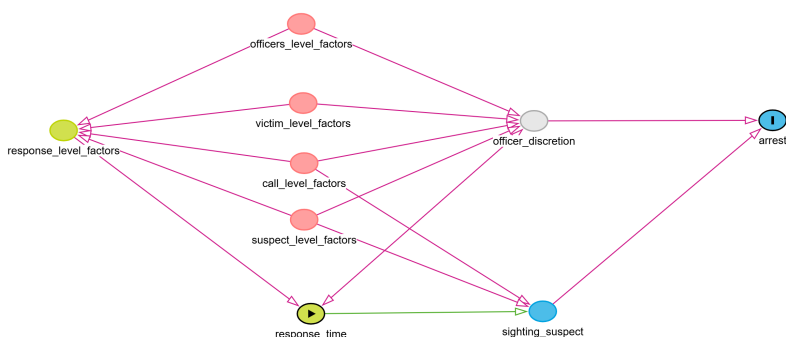


Figure 2.2: DAG of arrest

The relationships outlined in Figure 2.2 create biasing paths that cannot be fully resolved through statistical adjustment. As a result, neither the total nor the direct effect of response time on arrest can be estimated without bias. However, Figure 2.3 highlights that the effect of rapid response on whether police catch perpetrators can be estimated without bias.

To estimate this effect, analyses must control for factors that increase the likelihood of apprehension regardless of response time, such as call level factors (e.g., whether suspect is more or less likely to be aware of responding police), and suspect level factors (e.g., whether the suspect's identity is known, whether the suspect is injured, or whether they surrender voluntarily). Focusing on suspect apprehension rather than arrest introduces greater causal clarity by reducing the role of discretionary factors in the analysis. It also mitigates another key source of bias found in previous studies: arrests that occur at the scene but are unrelated to response time, such as when a bystander intervenes and the offender is subsequently arrested. Additionally, this approach eliminates the need for complex econometric models

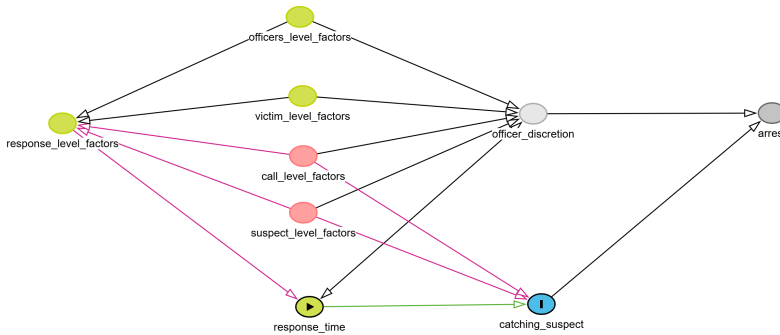


Figure 2.3: DAG of suspect being caught

designed to correct for response-time endogeneity stemming from sources such as those identified by Blanes i Vidal and Kirchmaier (2018).

Given these insights, I argue that research focus should shift from how response time affects arrest rates to the variable that directly responds to rapid response and determines the feasibility of arrest: whether police sight and catch suspects. This new focus also leads to two novel fundamental research questions that need to be answered in order to properly assess the potential of rapid response: (1) what is the causal effect of faster police response times on catching suspects, and (2) to what extent can police departments feasibly reduce response times? To explore these questions, I conclude by outlining three key considerations for improving the study of the first – focusing on data, analytical approach, and presentation of results – followed by a methodological suggestion for addressing the second.

First, accurate measurement of response time is essential, particularly given that the effect of rapid response declines sharply within the first few minutes. Current CAD systems often lack the precision required for robust analysis. Integrating response vehicle GPS data, which provides more reliable timestamps, would improve measurement accuracy. In the Netherlands, for example, police vehicles log GPS signals when the ignition key is removed, providing a more objective indicator of arrival time. Improving measurement precision would also allow for second-by-second analysis, offering greater granularity in understanding the impact of rapid response.

Second, when analysing the effects of response time on catching suspects, only in-progress incidents should be included to avoid the otherwise insurmountable challenge of measuring citizen reporting delays for completed crimes¹⁰. By focusing on suspect apprehension rather than arrest, there is no longer a need to exclude cases where arrests are not directly tied to response time. However, analyses should still control for cases where suspects are apprehended for reasons unrelated to police response, such as self-surrender or third-party interventions. It is also necessary to account for the non-linearity of the relationship between response time and apprehension probability. Empirical research has consistently shown that a 30-second delay when response time is one minute is far more consequential than the same delay when response time is five minutes. A linear model cannot adequately capture this

¹⁰Note: this does not mean that the effect of rapid response only pertains to this subsample, merely that the effect can best be estimated by analysing this restricted sample

diminishing effect. Instead, response times should be binned into discrete intervals before being incorporated into regression models (or alternatively transformed using logarithmic or quadratic functions).

2 Third, findings should be presented in a way that policymakers can readily interpret. Odds ratios and reference group coding obscure the absolute effects of response time and other explanatory variables, making it difficult to assess the real-world implications of statistical findings. This issue is evident in two examples. Cihan et al. (2012) report that each additional minute of response time is associated with a 9% decrease in the odds of arrest, but because the baseline probability of arrest is low, this translates to only a 1 percentage point decrease in absolute probability – far less striking than the odds ratio suggests. Similarly, Rief and Huff (2023) for example report that night shifts have an arrest odds ratio of 0.72 compared to morning shifts, implying a lower likelihood of arrest. However, without knowing the absolute arrest probability for either shift, the true extent of this difference remains unclear. A more effective way to present results is through marginal effects and deviation coding. Marginal effects translate model estimates into absolute probability changes, making it easier to understand the real-world consequences of response delays. Deviation coding avoids the reliance on an arbitrary reference group and instead estimates the effect of each category independently, providing a clearer picture of how different conditions influence arrest probability. Clear communication can also be achieved visually, as was commonly done in the 1970s, by plotting arrest probabilities across response time bins with a fitted logistic regression curve.

Lastly, a separate but equally important question is the degree to which police departments can realistically reduce response times. Here, agent-based modelling offers a powerful tool. By simulating real-world contexts – such as jurisdictional geography, street networks, speed limits, and resource allocation – these models can approximate the degrees of freedom police departments have to influence response times. An agent-based model (ABM) also allows for testing various scenarios under different levels of police capacity, offering insights into the resource trade-offs involved. For example, such models could evaluate whether, as Sherman et al. (1997, Chapter 8, p. 11) conjectured, halving response times would indeed require a doubling of the police force. A stylistic example of how such research may be carried out is suggested by Birks and Townsley (2018).

REFERENCES

- Berkow, M. 2011. Lecture to the National Police Academy.
- Bertram, D. K., & Vargo, A. (1976). Response time analysis study - preliminary findings on robbery in Kansas City. *Police Chief*, 43(5), 74–77.
- Bieck, W. H. (1977a). *Response Time Analysis: Executive summary* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1977b). *Response Time Analysis: Part I Crime Analysis (Vol. II)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1978). *Response Time Analysis: Methodology (Vol. I)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1980b). *Response Time Analysis: Part II Crime Analysis (Vol. III)* (tech. rep.). National Institute of Justice. Kansas City.
- Birks, D., & Townsley, M. (2018). Agent-based decision-support systems for crime scientists. In R. Wortley, A. Sidebottom, N. Tilley & G. Laycock (Eds.), *Routledge handbook of crime science* (pp. 334–349). Routledge. <https://doi.org/https://doi.org/10.4324/9780203431405-25>
- Blake, L., & Coupe, R. T. (2001). The Impact of Single and Two-Officer Patrols on Catching Burglars in the Act: A Critique of the Audit Commission's Reports on Youth Justice. *British Journal of Criminology*, 41(2), 381–396. <https://doi.org/10.1093/bjc/41.2.381>
- Blanes i Vidal, J., & Kirchmaier, T. (2018). The Effect of Police Response Time on Crime Clearance Rates. *The Review of Economic Studies*, 85(2), 855–891. <https://doi.org/10.1093/restud/rdx044>
- Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis. *Journal of Experimental Criminology*, 15(3), 289–311. <https://doi.org/10.1007/s11292-019-09372-3>
- Brown, W. J., & Lipsett, F. (1976). Response speeds and response times of urban police patrol cars in Ottawa, Canada. *Journal of Criminal Justice*, 4(3), 191–202. [https://doi.org/10.1016/0047-2352\(76\)90002-7](https://doi.org/10.1016/0047-2352(76)90002-7)
- Chaiken, J. M., Greenwood, P. W., & Petersilia, J. R. (1977). The criminal investigation process: A summary report. *Policy Analysis*, 3(2), 187–217.
- Cihan, A., Zhang, Y., & Hoover, L. (2012). Police Response Time to In-Progress Burglary. *Police Quarterly*, 15(3), 308–327. <https://doi.org/10.1177/1098611112447753>
- Clawson, C., & Chang, S. K. (1977). The relationship of response delays and arrest rates. *Journal of Police Science and Administration*, 5(1), 53–68.
- Cordner, G. W., Greene, J. R., & Bynum, T. S. (1983). *The Sooner the Better: Some Effects of Police Response Time* (tech. rep.). National Institute of Justice. Washington, DC.
- Coupe, R. T. (2016). Evaluating the effects of resources and solvability on burglary detection. *Policing and Society*, 26(5), 563–587. <https://doi.org/10.1080/10439463.2014.989155>
- Coupe, R. T. (2019, January). Investigative Activities, Resources and Burglary Detection. In *Crime solvability factors* (pp. 367–389). Springer International Publishing. https://doi.org/10.1007/978-3-030-17160-5_{_}16

- Coupe, R. T., & Blake, L. (2005). The effects of patrol workloads and response strength on arrests at burglary emergencies. *Journal of Criminal Justice*, 33(3), 239–255. <https://doi.org/10.1016/j.jcrimjus.2005.02.004>
- Coupe, R. T., & Blake, L. (2011). The effects of target characteristics on the sighting and arrest of offenders at burglary emergencies. *Security Journal*, 24(2), 157–178. <https://doi.org/10.1057/sj.2009.23>
- Coupe, R. T., & Girling, A. (2001). Modelling police success in catching burglars in the act. *Omega*, 29(1), 19–27. [https://doi.org/10.1016/S0305-0483\(00\)00030-X](https://doi.org/10.1016/S0305-0483(00)00030-X)
- Coupe, R. T., & Griffiths, M. (2000). Catching Offenders in the Act: An Empirical Study of Police Effectiveness in Handling ‘Immediate Response’ Residential Burglary. *International Journal of the Sociology of Law*, 28(2), 163–176. <https://doi.org/10.1006/ijsl.2000.0120>
- Farrington, D. P. (1983). Randomized Experiments on Crime and Justice. *Crime and Justice*, 4, 257–308. <https://doi.org/10.1086/449091>
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In D. P. Farrington, D. C. Gottfredson, L. W. Sherman & B. C. Welsh (Eds.), *Evidence-based crime prevention* (1st ed., pp. 13–21). Routledge.
- Fienberg, S., Larntz, K., & Reiss, A. J. (1976). Redesigning the Kansas City preventive patrol experiment. *Evaluation*, 3(1-2), 124–131.
- Frydl, K., & Skogan, W. (Eds.). (2004, April). *Fairness and Effectiveness in Policing*. The National Academies Press. <https://doi.org/10.17226/10419>
- Holliday, L. P. (1974). *A methodology for radio car planning* (tech. rep.). RAND. Rand Institute.
- Isaacs, H. H. (1967). *A Study of Communications, Crimes and Arrest in a Metropolitan Police Department. Appendix B of Science and Technology Task Force Report*. (tech. rep.). U.S. Government. Washington DC.
- Kelling, G. L., Pate, A. M., Dieckman, D., & Brown, C. (1974). *The Kansas City Preventive Patrol Experiment: A Summary Report* (tech. rep.). Police Foundation. Washinton D.C.
- Larson, R. C. (1975). What happened to patrol operations in Kansas City? A review of the Kansas City preventive patrol experiment. *Journal of Criminal Justice*, 3(4), 267–297. [https://doi.org/10.1016/0047-2352\(75\)90034-3](https://doi.org/10.1016/0047-2352(75)90034-3)
- Larson, R. C. (1982). Critiquing Critiques: Another Word on the Kansas City Preventive Patrol Experiment. *Evaluation Review*, 6(2), 285–293. <https://doi.org/10.1177/0193841X8200600209>
- Mastrobuoni, G. (2019). Police disruption and performance: Evidence from recurrent redeployments within a city. *Journal of Public Economics*, 176, 18–31. <https://doi.org/10.1016/j.jpubeco.2019.05.003>
- Nagin, D. S. (2013). Deterrence in the Twenty-First Century. *Crime and Justice*, 42(1), 199–263. <https://doi.org/10.1086/670398>
- Pate, A. M., Ferrara, A., Bowers, R. A., & Lorence, J. (1976, January). *Police Response Time: Its Determinants and Effects* (tech. rep.). National Policing Institute. Washington, DC.

- Pate, A. M., Kelling, G. L., & Brown, C. (1975). A response to “what happened to patrol operations in Kansas City?” *Journal of Criminal Justice*, 3(4), 299–320. [https://doi.org/10.1016/0047-2352\(75\)90035-5](https://doi.org/10.1016/0047-2352(75)90035-5)
- Prince, H., Lum, C., & Koper, C. S. (2021). Effective police investigative practices: an evidence-assessment of the research. *Policing: An International Journal*, 44(4), 683–707. <https://doi.org/10.1108/PIJPSM-04-2021-0054>
- Rief, R., & Huff, J. (2023). Revisiting the influence of police response time: Examining the effects of response time on arrest and how it varies by call type. *Journal of Criminal Justice*, 84(84), 1–11. <https://doi.org/10.1016/j.jcrimjus.2022.102025>
- Risman, B. J. (1980). The Kansas City Preventive Patrol Experiment: A Continuing Debate. *Evaluation Review*, 4(6), 802–808. <https://doi.org/10.1177/0193841X8000400605>
- Sherman, L. W. (2013). The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking. *Crime and Justice*, 42(1), 377–451. <https://doi.org/10.1086/670819>
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997, January). *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress* (tech. rep.). NIJ. Washington, DC.
- Sherman, L. W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime “hot spots”: A randomized, controlled trial. *Justice Quarterly*, 12(4), 625–648. <https://doi.org/10.1080/07418829500096221>
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Tarr, D. P. (1978). Analysis of response delays and arrest rates. *Journal of Police Science & Administration*, 6(4), 429–451.
- Textor, J., van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. *International Journal of Epidemiology*, 45(6), 1887–1894. <https://doi.org/10.1093/IJE/DYW341>
- Verlaan, T., & Ruiter, S. (2023a). *PROTOCOL-Determinants and effects of police response time: a scoping review*. <https://doi.org/10.17605/OSF.IO/HU2E9>
- Weisburd, D., & Eck, J. E. (2004). What Can Police Do to Reduce Crime, Disorder, and Fear? *Annals of the American Academy of Political and Social Science*, 593(4), 42–65. <https://doi.org/10.1177/0002716203262548>
- Weisburd, D., & Majumdar, M. K. (Eds.). (2018, February). *Proactive Policing*. National Academies Press. <https://doi.org/10.17226/24928>
- Weisburd, D., Petersen, K., Telep, C. W., & Fay, S. A. (2024). Can increasing preventive patrol in large geographic areas reduce crime? A systematic review and meta-analysis. *Criminology & Public Policy*, 23(3), 721–743. <https://doi.org/10.1111/1745-9133.12665>
- Weisburd, D., Wilson, D. B., Petersen, K., & Telep, C. W. (2023). Does police patrol in large areas prevent crime? Revisiting the Kansas City Preventive Patrol Experiment. *Criminology & Public Policy*, 22(3), 543–560. <https://doi.org/10.1111/1745-9133.12623>
- Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, 103(2), 280–293. https://doi.org/10.1162/rest{_}_a{_}_00889

World Bank. 2023. Fixed telephone subscriptions (per 100 people). <https://data.worldbank.org/indicator/IT.MLT.MAIN.P2> *World Bank Open Data*.

Zimring, F. E. (1978). Policy Experiments in General Deterrence: 1970-1975. In *Deterrence and incapacitation estimating the effects of criminal sanctions on crime rates* (pp. 140–186). The National Academies Press. <https://doi.org/10.17226/18561-2>

3

3

THE MARYLAND REPORT

CONSTRUCTING THE PROACTIVE POLICING PARADIGM

This chapter provides a critical examination of the influential report Preventing Crime: What Works, What Doesn't, What's Promising – commonly referred to as the Maryland Report – and its role in constructing what is here termed the Proactive Policing Paradigm. Focusing specifically on a section of Chapter 8, Policing for Crime Prevention, the discussion centres on the report's evaluation of both reactive and proactive policing strategies, specifically rapid response, random patrol, and directed patrol. By closely examining the evidence cited and the application of methodological standards, the chapter highlights a pattern of inconsistent treatment between these strategies. It argues that these inconsistencies contributed to flawed conclusions – particularly the strong endorsement of proactive approaches and the premature dismissal of reactive ones. The chapter concludes that a renewed and systematic examination of the evidence base for reactive policing strategies is warranted, and that such a reappraisal would contribute to a more balanced understanding within contemporary evidence-based policing.

3.1 EVALUATING THE STANDARD MODEL OF POLICING

The 1970s marked a turning point in policing research and practice. The crime-resolving and (by extension) preventive effectiveness of random patrol, rapid response and reactive investigation – the three core activities of what would later become known as the *Standard Model of Policing* (Frydl & Skogan, 2004) – were fundamentally challenged in four landmark studies (Bieck, 1977a; Chaiken et al., 1977; Kelling, Pate, Dieckman & Brown, 1974; Spelman & Brown, 1981).

3

The first of these, the *Kansas City Preventive Patrol Experiment* (KCPPE), tested the deterrent effect of visible police patrols by randomly assigning three levels of patrol presence – eliminated patrols, routine patrols, and increased patrols – to fifteen police beats (Kelling, Pate, Dieckman & Brown, 1974). After a year, results revealed no significant difference in crime levels between these three beat conditions. Moreover, there was no effect on fear of crime; no effect on citizen attitude towards or satisfaction with the police; and no effect on police response times. In the aftermath of the KCPPE, the view that random police patrol does not deter crime became widespread (Felson, 1994, 10–11; Gottfredson & Hirschi, 1990, p. 270; Klockars, 1983, p. 130; Skolnick & Bayley, 1986, p. 4).

The second study, the *Kansas City Response Time Analysis* (KCRTA) study, examined the relationship between police response times and arrest rates. Consistent with earlier research (Clawson & Chang, 1977; Holliday, 1974; Isaacs, 1967; Tarr, 1978), the study found that faster response times correlated with higher arrest rates. However, it also introduced two critical insights that undermined the case for rapid response as a crime resolution and prevention tactic. First, the majority (63%) of Part I crime calls were classified as ‘discovery crimes’ – incidents reported after the perpetrator had already fled – rendering response time irrelevant. Second, for the remaining crimes where police intervention could influence outcomes (i.e., involvement crimes), half of all victims delayed reporting for more than five minutes, severely limiting the potential impact of rapid response.

The findings of the KCRTA study were received with scepticism by the wider police community, arguing that the results, particularly the long citizen reporting delay, may be specific to Kansas City (Spelman & Brown, 1981). To address this concern, the National Institute of Justice funded the Police Executive Research Forum to carry out a replication study across four cities in different states: Jacksonville, Florida; Peoria, Illinois; Rochester, New York; and San Diego, California (Spelman & Brown, 1981). The replication confirmed the original study’s findings: while shorter police response times correlated with higher arrest rates, long citizen reporting delays remained a persistent issue, and most calls concerned discovery crimes. These results reinforced the conclusion that rapid response would not reduce overall crime rates.

Finally, the RAND Corporation’s 1977 *Criminal Investigation Process* (CIP) study, which also relied heavily on data from Kansas City, cast further doubt on the effectiveness of reactive police investigations (Chaiken et al., 1977). The study found that 30% of all cleared crimes resulted from on-scene arrests, while another 50% were solved through suspect identification. An additional 17% were cleared through actions taken by patrol officers, members of the public, or routine investigative procedures that could be performed by clerical staff. In contrast, only about 3% of cases were solved using specialised investigative techniques. These findings challenged the assumption that traditional detective work played a decisive role in solving crimes.

3.2 IMPACT OF THESE LANDMARK EVALUATION STUDIES

The KCRTA received some criticism in a subsection of a report titled ‘Synthesizing and Extending the results of Police Patrol Studies’ (R. C. Larson & Cahn, 1985), but this critique received recognition nor traction. The subsequent replication study by Spelman and Brown (1981) did also not result in further debate. In contrast, the RAND study and the KCPPE have ‘provoked more comment, controversy and debate about police matters than any other single research project’ (Hayes, 1987, p. 101).

The RAND study sparked widespread debate, including a critique (Gates & Knowles, 1976) and subsequent rebuttal (Greenwood et al., 1976) in *The Police Chief*. Beyond academic discussions, the study had a substantial impact on policing policy and practice. In response to its findings, the National Institute of Criminal Justice and Law Enforcement (NILECJ) designated *Managing Criminal Investigations* (MCI) as the principal police subject in its National Criminal Justice Executive Training Programme (NCJETP) in 1976 (Hayes, 1987, p. 106). Over the next three years, more than 700 police executives from over 400 departments participated in the program, with as many as 90 percent implementing changes in their investigative processes upon returning to their agencies (Miron et al., 1979). In addition to national training efforts, the MCI program was field-tested in five jurisdictions: Birmingham, Montgomery County, Rochester, St. Paul, and Santa Monica – though no effect evaluations of these studies were published. As a result, research on the crime-resolving effects of reactive investigations remained limited, and studies on its deterrent effects appear non-existent (Prince et al., 2021).

The KCPPE was widely debated between the report’s authors and experts from the field, with criticisms focusing on its design, execution, and the interpretation of its results from its publication until 1983 (see Farrington, 1983; Fienberg et al., 1976; R. C. Larson, 1975; Pate et al., 1975; Risman, 1980; Zimring, 1978). The debate was revisited in 1986 by Sherman when he used the KCPPE to illustrate the Evaluation Problem in Policing, arguing that the low number of experimental and non-experimental sites led to a statistical bias favouring the null hypothesis.

In 1995, Sherman & Weisburd introduced hot spot policing, which focused proactive policing efforts in (micro-)places with relatively high crime concentrations referred to as ‘hot spots’ (Sherman et al., 1989), and published the findings of four large policing experiments in crime hot spots (Sherman & Rogan, 1995; Sherman & Weisburd, 1995; D. Weisburd & Green, 1995). Hot spot policing not only allowed for more targeted and potentially effective policing, but it also addressed the Evaluation Problem by increasing the number of experimental and non-experimental units without diluting crime levels to unmeasurable degrees. Its introduction facilitated a surge in rigorous (quasi-)experimental research into proactive policing interventions. By February 2017, the latest iteration of an ongoing meta-analysis identified 65 studies examining the deterrent effects of proactive policing in crime hot spots. The building dominance of proactive policing research culminated in the 1997 Maryland Report, which not only reinforced this trend but actively discouraged further exploration of reactive strategies. Leading to what I have termed the *Proactive Policing Paradigm*.

3.3 THE MARYLAND REPORT

In 1996, the U.S. National Institute of Justice (NIJ), mandated by Congress, commissioned a report on the ‘comprehensive evaluation of the effectiveness’ of over 500 federally funded crime prevention programs, with a combined budget exceeding \$3 billion. This report was required to be ‘independent in nature’ and to ‘employ rigorous and scientifically recognised standards and methodologies’ (U.S. Congress, 1995). Led and largely authored by Lawrence Sherman, the report – often referred to as the *Maryland Report* – became the most widely cited text on the effectiveness of policing strategies to date with around 3000 citations spread over the report and the subsequent research brief on Google Scholar (Sherman et al., 1997).

In many ways, the Maryland Report represented a major step forward in the professionalisation of both policing practice and policy evaluation. The evidentiary framework employed in the report – now known as the Maryland Scientific Methods Scale (Farrington et al., 2002) – is still widely used for evaluating (crime prevention) policies and is central to the increasingly popular evidence-based policing paradigm.

Nevertheless, in this chapter I argue that the report also contains inconsistencies in applying these evidentiary standards to reactive and proactive policing strategies, with subsequent errors in drawing conclusion as a result. These errors led the report to strongly promote focussed proactive policing strategies as the only way to achieve ‘substantial reduction in national rates of serious crime’ (Sherman et al., 1997, p. v) and denounce (research into) the crime preventive effects of reactive policing strategies, calling further research efforts a ‘waste of tax dollars’ asserting they lacked ‘empirical and theoretical justification’ (Sherman et al., 1997, Chapter 8, p. 11), which contributed to a field-wide tunnel vision that disproportionately prioritised proactive policing research.

These are strong claims, which I will substantiate through a critical evaluation of the report, focusing specifically on a section of Chapter 8, ‘Policing for Crime Prevention’, authored by Lawrence Sherman.

3.4 THE MARYLAND REPORT ON THE STANDARD MODEL OF POLICING

In this section, I present a chronological critique of the paragraphs evaluating the policies of Rapid Response, Random Patrol and Directed Patrol. Each paragraph in the report consists of a text supplemented with a table, in which the findings are summarised and the evaluation studies are classified based on (what is now known as) the Maryland Scientific Methods Scale. Both the text and the summary table (of which a exact copy is provided) are subjected to scrutiny.

3.4.1 RAPID RESPONSE

In the section ‘Rapid Response to 911’, the Maryland report presents two early studies supporting the hypothesis that shorter response times increase arrest rates (Clawson & Chang, 1977; Isaacs, 1967) and contrasts these with three studies that it claims fail to confirm this relationship (Bieck, 1977c; Pate et al., 1976; Spelman & Brown, 1981). To summarise the findings, the report presents them in a table (see Figure 3.1) that assigns a scientific methods score to each study and categorises its main conclusions.

Studies	Scientific Methods Score	Findings
Isaacs, 1967	1	Shorter police travel time, more arrests
Clawson and Chang, 1977	1	Same as preceding
Pate et al., 1976	1	Police travel time unrelated to arrest
Kansas City (MO) Police, 1977	2	Same as preceding, most crimes
Spelman and Brown, 1981	2	Same as preceding

Figure 3.1: Summary table of Rapid Response studies (Figure 8 - 2 in the official report)

First of all, it should be noted that two studies (Cordner et al., 1983; Tarr, 1978) are left out of the discussion, both of which strongly support the hypothesised relationship between shorter response times and more arrests. The omission of the Tarr study is particularly striking as (1) it provided the largest arrest sample size of all studies (in fact, larger than the arrest sample of all studies that refute the relationship, combined) and (2) was discussed and cited in Spelman and Brown (1981), which was included.

Second, both the scientific methods score and the summarised findings (Figure 3.1) of the studies that were included contain multiple imprecisions:

1. Both the Isaacs (1967) and the Clawson and Chang (1977) study tested for police dispatch and travel time, not just police travel time.
2. The article by Pate et al. (1976) is cited to find 'Police travel time unrelated to arrest'. However, the conclusion in the article itself reads 'The results from all three surveys suggest no significant relationship between response time and outcome' (1976, p. 34, emphasis added), arrest being only a subcategory of outcome. If we do look exclusively at arrests, the study presents findings from three surveys, which, taken together, report 11 arrests made for response times between 1 and 5 minutes (response times are based on rough estimates of travel time by citizen in this study), and 5 arrests for response times between 5 and 10 minutes. Given the extremely low sample size, it is no wonder we do not find statistically significant results. However, if the results can be inferred to provide any insights into a relationship, it is certainly in support of the hypothesis, not against, as presented here.
3. The Kansas City study by Bieck (1977c) and the PERF study Spelman and Brown (1981) are cited as 'same as preceding, most crimes'. However, Pate et al. (1976) only analyses (estimates of) travel time, while Bieck analyses (combinations of) reporting time, dispatch time and travel time and Spelman & Brown solely focus their analysis on reporting time.
4. Bieck (1977c) does not find 'police travel times unrelated to arrests, most crimes'. The report only analyses the relationship for involvement crimes, asserting (but not testing) that discovery crimes (which indeed constitute the majority of crime calls – 63%) are not susceptible to rapid response because of too long (and unknown/unknowable) reporting times. On involvement crimes, the report states: 'Rapid reporting can

increase the chances of making an on-scene arrest for all types of Part I involvement crimes.’ (Bieck, 1977a, p. 23). Table 3.1 illustrates that while the relationship between travel time and response-related arrests differs in strength per crime type/group, it is clearly present for all crimes and in some, very strong.

Table 3.1: Response-related arrest probability for different crime type (groups) per travel time bin. Derived from Figures 3–9 and 3–10 (Bieck, 1977b) and Figure 3–5 (Bieck, 1980b)

Crime type	1-min	2-min	3-min	4-min	5-min	6-min	7-min	8-min	9-min	10-min
Forgery, fraud, embezzlement	90%	71%	50%	33%	23%	14%	8%	4%	1%	0%
Non-aggravated assault & vandalism	27%	15%	11%	7%	5%	3%	1%	1%	0%	0%
Burglary	80%	63%	52%	43%	37%	32%	28%	24%	20%	17%
Burglary, larceny & auto thefts	58%	45%	36%	29%	24%	20%	16%	13%	10%	8%
Rape, robbery & aggravated assault	10%	8%	7%	6%	6%	6%	5%	5%	5%	5%

5. Similarly, in their replication study, Spelman and Brown (1981) write: for property crimes reported in-progress, ‘chances were excellent of effecting a response-related arrest if police response was prompt’ (p. 68). Also, for violent crimes reported in-progress, they conclude that ‘fast police response would very likely result in arrest’ (p. 70). However, it is true that once again the argument is made that most calls are not susceptible to rapid response as they are reported after-the-fact, or reported with a too long delay (mainly due to telephone communication restrictions in the 1970s).
6. The Maryland report classifies the KCRTA study (Bieck, 1977c) and the PERF study (Spelman & Brown, 1981), as level 2 studies, while classifying the other included studies as level 1. However, while it is true that the KCRTA and PERF studies contained a more diligent and accurate data collection, the scientific methods score is not based on data quality, but research design. The KCRTA and PERF studies use the same correlational designs as previous studies and also fail to establish causal order; the differentiating factor for level 2 studies as explicated by the Maryland report (Sherman et al., 1997, Chapter 2, p. 19). The failure establish causal order is explicitly acknowledged by Spelman and Brown (1981, Chapter Appendix B, p. 117), who write: ‘A distinct but related problem that has not been handled by researchers is the problem of two-directional effects.’ As such, these studies should also be classified as level 1.

Fourth, the Maryland report writes ‘average reporting time for involvement crimes is 41 minutes’ (Sherman et al., 1997, Chapter 8, p. 11). In contrast, the KCRTA authors themselves write: ‘Because the skewness is reflected in the means and standard deviations of the response times reported, the median time is suggested as more representative of the time taken to report, dispatch, and respond to a crime incident.’ The average reporting time for involvement crimes based on the median, is 5 minutes – a stark difference from 41 minutes reported in the Maryland report.

Fifth, the Maryland report writes: ‘Cutting police travel time for such crime from 5 to 2.5 minutes could require a doubling of the police force’ (Sherman et al., 1997, Chapter 8, p. 11). However, no source is cited for this assertion and no research appears to have been done to substantiate the claim.

Sixth, the Maryland report writes:

The conclusion that reduced response time will not reduce crime is based on strong but indirect evidence. The evidence is strong because it is based on large samples, careful measurement, and a replicated research design in five diverse cities showing little variation in arrest rates by police travel time, the main factor that tax dollars can affect. (Sherman et al., 1997, Chapter 8, p. 10)

However, first, even if we accept the reports classification of these studies as level 2, Cook and Campbell (1979) would categorise such studies as ‘inadequate and uninterpretable’ – a classification that is later explicitly adopted in the formalisation of the Maryland Scientific Methods Scale (Farrington et al., 2002, p. 17). This is a far cry from ‘strong evidence’. Second, the sample size of response-related on-scene arrests in Kansas City was only 27 for Part I crimes and 11 for Part II crimes. In the replication cities, the sample sizes were: 259, 63, 75, and 249. As the reports analyses these samples across thirty one-minute response time bins, it is difficult to maintain that these are ‘large’ samples. Third, while it is true that the proportion of discovery crimes and citizen reporting delays was replicated across five cities, the relationship between shorter response times and increased arrest rates was replicated in all twelve cities. Lastly, it is difficult to support the claim that there is ‘little variation in arrest rates by police travel time’ based on the data presented in Table 3.1.

The inconsistencies and inaccuracies outlined above all appear within the approximately 750 words dedicated to *Rapid Response* in the report. At the end of this section, the report reaches the following conclusion on the evidence concerning the effects of rapid response:

It is indirect because an experimental test of the effects of reduced police travel time on city-wide arrest and crime rates has never been conducted. Yet there is neither empirical nor theoretical justification for such an expensive test. Given the strong evidence of citizen delays in reporting involvement crimes, and the small proportion of serious crimes that feature direct victim-offender involvement, further tests of this theory seem to be a waste of tax dollars. 1997, Chapter 8, p.11

This dismissal of Rapid Response by the Maryland report as a strategy that ‘doesn’t work’ is remarkable not only because of the reasons outlined above, but also because it contradicts the explicit evidentiary standards it sets itself for classifying a strategy as ineffective. According to these standards, a practice must be discredited by at least two level-3 studies (Sherman et al., 1997, p. ix) – a threshold that is clearly not met in this case.

3.4.2 RANDOM PATROL

Perhaps even more striking is the contrast between this unqualified rejection of *Rapid Response* and the far more cautious treatment of *Random Patrol*, despite the latter being supported by stronger level-3 studies (see Figure 3.2). In discussing these studies, Sherman takes a far more critical stance, citing multiple methodological limitations:

... failure to measure the actual differences in patrol dosage and the possible lack of them (R. C. Larson, 1975), its inadequate statistical power to detect

large percentage differences in crime as not due to chance (Fienberg et al., 1976), and its failure to assign patrol dosage at random (Farrington, 1983). . . . The weakness of the evidence is even greater for the one study claiming to find a crime prevention effect from random patrols not focused on crime risks (Trojanowicz, 1986). The design of this study was limited to recorded crime and calls for service, with no victimisation surveys. 1997, Chapter 8, p. 13

Studies	Scientific Methods Score	Findings
Kelling et al., 1974	3	No difference in crime by N of police cars assigned
Police Foundation, 1981	3	No difference in crime by N of foot patrol assigned
Trojanowicz, 1986	3	Foot patrol areas had fewer crimes than controls, but no significance tests reported.

Figure 3.2: Summary table of Random Patrol studies (Figure 8 - 3 in the official report)

3.4.3 DIRECTED PATROL

Moreover, when we turn to the report's evaluation of *Directed Patrol*, we do not see the same level of scrutiny and methodological scepticism applied elsewhere, as I will outline below. Figure 3.3 shows the original studies presented in the summary table in the Maryland Report.

Studies	Scientific Methods Score	Findings
Press, 1971	3	40% more police, reductions of outdoors crime
Chaiken et al., 1975; Chaiken, 1978	3	Police on subways at night reduced crime
Dahman, 1975	2	More police, reductions of outdoors crime
Schnelle et al., 1977	2	400% more patrol, less Part I crime
Sherman and Weisburd, 1995	5	100% more patrol, less observed hot spot crime
Koper, 1995	4	Longer patrol visits, longer post-visit crime-free time
Reiss, 1995 Review:		
Barker et al., 1993	2	Squad focused on hot spots, where street crime dropped
Burney, 1990	2	Saturation patrols, reduced street crime

Figure 3.3: Summary table of Directed Patrol studies (Figure 8 - 4 in the official report)

However, of these original nine studies, Chaiken (1978) does not present any new empirical data but merely report on findings from Chaiken et al. (1975) and Koper (1995) does not introduce new empirical evidence but rather provides additional analysis of data collected in the level-5 study by Sherman and Weisburd (1995). As such, these do not contribute independent findings regarding the effects of directed patrol. Moreover, the reports by Burney (1990) and Barker et al. (1993) constitute mainly short summaries of findings and lack sufficient detail to allow for any meaningful assessment of methodological quality, and arguably should not have been included in an evaluation report that claims to consider only ‘completed scientific studies’ (Chapter 1, p. 4) ‘meeting minimal standards of scientific rigour’ (Chapter 1, p. 6).

However, even allowing for Barker et al. (1993) and Burney (1990), this leaves seven studies that do contribute new empirical evidence (i.e., Barker et al., 1993; Burney, 1990; Chaiken et al., 1974; Dahmann, 1975; Press, 1971; Schnelle et al., 1977; Sherman & Weisburd, 1995).

Among these, Sherman and Weisburd (1995) stands out as the prototypical rigorous hot spot policing study, fully warranting its level-5 rating on the Maryland Scientific Methods Scale (MSMS): Unlike the random patrol studies critiqued in the report, this study has (1) sufficient statistical power to detect substantial effects, (2) assigns patrol dosage randomly, and (3) is able to measure actual patrol variation (although it also recorded periods with no dosage difference).

However, and notably absent from the reports’ discussion, the six other studies (i.e., Barker et al., 1993; Burney, 1990; Chaiken et al., 1974; Dahmann, 1975; Press, 1971; Schnelle et al., 1977) fail to meet these same three criteria. Moreover, all studies except Press (1971) – but including Sherman and Weisburd (1995) – rely solely on registered crime data rather than victimisation surveys; a limitation which was treated as disqualifying for Random Patrol studies, but not so for Directed Patrol studies.

Yet, more fundamentally, the report fails to clearly operationalise what counts as ‘directed’ in Directed Patrol. While it associates the strategy with hot spots and hot times, no concrete definition or threshold is provided.

The contemporary operationalisation of this, laid out in the Campbell systematic review of hot spot policing (Braga, Turchan, Papachristos & Hureau, 2019), includes only studies that target ‘places’ smaller than neighbourhoods. Of the studies cited in the report, only Sherman and Weisburd (1995) meet this stricter definition of directed place-based intervention and is thus the only one included in this systematic review.

Now, if we consider the ‘level of direction’ in terms of the geographic scale of patrol deployment, we can see why:

- In the study by Press (1971), police manpower was increased by 40% in the 20th precinct of New York City.
- In Chaiken et al. (1974), the Transit Police Force in NYC – tasked with patrolling every NYC subway station and train at night – was nearly tripled, from 1219 to over 3100 patrol officers.
- Dahmann (1975) reports on three different crime interventions in: 3 (out of 40) precincts in Denver; in 3 (out of 6) districts in Cleveland; 6 (out of 490) Pauly blocks in St Louis.

- Schnelle et al. (1977), reports a 398% increase in patrol movement in 4 (out of 33) patrol zone in the metropolitan Nashville area.
- Sherman and Weisburd (1995) allocated around twice as much police presence in 55 crime hot spots.
- Burney (1990) reports on police initiatives in 1 borough (out of 32) in the Greater London region.
- Barker et al. (1993) reports on a policing initiative in Battersea, a district within the adjacent borough of Wandsworth.

In short, the spatial focus of proactive presence in the ‘directed patrol’ studies ranged from half a city (Cleveland), to one of the world’s largest public transport systems (NYC metro), to micro-places (hot spots) in Minneapolis.

Compare this to the geographical units used in the studies classified as *random patrol studies*: 5 beats in Kansas City (Kelling, Pate, Dieckman & Brown. Charles, 1974); 4 beats in Newark (Pate, 1981); and 14 neighbourhoods in Flint (Trojanowicz, 1986).

Based on this comparison, it is not possible to argue that the classification of studies as ‘directed patrol’ versus ‘random patrol’ was based on the level of geographic specificity of patrol. If anything, it could be argued that the ‘random patrol’ studies were directed at smaller geographic places. Currently, as it stands, the only consistent difference appears to be that studies classified as Directed Patrol reported positive findings, whereas those in the Random Patrol category, did not.

3.5 AFTERMATH OF THE MARYLAND REPORT

The impact of the Maryland Report is hard to understate, according the website of the University of Maryland, it led to the founding of the Campbell Collaboration as an international review organisation for evaluating crime prevention programs, as well as the Academy of Experimental Criminology, the ASC Division of Experimental Criminology, and the Societies of Evidence-Based Policing in the UK, Australia, New Zealand, Canada and the US (Criminology and Criminal Justice Department | University of Maryland, n.d.).

What is more, if we examine the subsequent consensus reports on police evaluation research published by the National Academy of Sciences, Engineering, and Medicine (Frydl & Skogan, 2004; D. Weisburd & Majimundar, 2018), it becomes clear that the Maryland Report had its intended influence on the research agenda. Since its publication, the two most prominent reactive policing strategies – rapid response and reactive investigation – have received little rigorous scientific scrutiny. In the most recent NASEM report, appropriately titled *Proactive Policing*, the Kansas City Response Time Analysis and the PERF replication study are still cited as the best available evidence on rapid response. Similarly, the RAND Corporation’s *Criminal Investigation Process* (CIP) study (1977) remains the primary source cited for evaluating reactive investigations.

3.6 CONCLUSION

It is important to begin by acknowledging that this analysis focuses on a single section of a much larger research effort. The findings presented here do not necessarily reflect the quality or rigour of other parts of the Maryland Report. However, it should be stated that no selective focus or ‘cherry-picking’ for a section with perceived mistakes has taken place; rather, this section was chosen because of its central relevance to research on the Standard Model of Policing.

That said, a close reading of this subsection of Chapter 8, Policing for Crime Prevention, reveals significant inconsistencies between the evidentiary standards outlined in the report and the manner in which they are applied. The analysis has shown that this relatively short section contains a sizeable number of methodological oversights, inaccuracies, and misclassifications. Notably, all these errors appear to systematically favour proactive and directed strategies, while downplaying or prematurely dismissing reactive strategies.

Given the influence of the Maryland Report on the trajectory of policing research and policy – particularly in shaping what I have called the Proactive Policing Paradigm – this is a consequential finding. The report not only catalysed the rise of proactive policing strategies but also contributed to the near-total stagnation of empirical research on the effects of reactive policing approaches; rapid response and reactive investigation. This lopsided development in the evidence base is not the inevitable result of empirical findings, but rather the outcome of selective interpretation and uneven application of methodological criteria. As such, I argue that a systematic re-evaluation of reactive policing strategies is both necessary and overdue.

REFERENCES

- Barker, M., Geraghty, J., Webb, B., & Key, T. (1993). *The Prevention of Street Robbery*, Home Office Police Department.
- Bieck, W. H. (1977a). *Response Time Analysis: Executive summary* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1977b). *Response Time Analysis: Part I Crime Analysis (Vol. II)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1977c). *Response Time Analysis: Synopsis* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1980b). *Response Time Analysis: Part II Crime Analysis (Vol. III)* (tech. rep.). National Institute of Justice. Kansas City.
- Braga, A. A., Turchan, B., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1046>
- Burney, E. (1990). *Putting street crime in its place: a report to the Community/Police Consultative group for Lambeth*. Goldsmith's College.
- Chaiken, J. M. (1978). What Is Known About Deterrent Effects of Police Activities. In J. A. Cramer (Ed.), *Preventing crime* (pp. 109–135, Vol. 10). Sage Publications.
- Chaiken, J. M., Greenwood, P. W., & Petersilia, J. R. (1977). The criminal investigation process: A summary report. *Policy Analysis*, 3(2), 187–217.
- Chaiken, J. M., Lawless, M., & Stevenson, K. (1975). The Impact of Police Activity on Crime: Robberies on the New York City Subway System. *Urban Analysis*, 3, 173–205.
- Chaiken, J. M., Lawless, M., & Stevenson, K. A. (1974, January). *The Impact of Police Activity on Crime: Robberies on the New York City Subway System* (tech. rep.). The New York Rand Institute. Santa Monica, CA, RAND Corporation.
- Clawson, C., & Chang, S. K. (1977). The relationship of response delays and arrest rates. *Journal of Police Science and Administration*, 5(1), 53–68.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin.
- Cordner, G. W., Greene, J. R., & Bynum, T. S. (1983). *The Sooner the Better: Some Effects of Police Response Time* (tech. rep.). National Institute of Justice. Washington, DC.
- Dahmann, J. S. (1975, November). *Examination of police patrol effectiveness - high impact anti-crime program* (tech. rep.). National Institute of Law Enforcement and Criminal Justice. Washington.
- Farrington, D. P. (1983). Randomized Experiments on Crime and Justice. *Crime and Justice*, 4, 257–308. <https://doi.org/10.1086/449091>
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In D. P. Farrington, D. C. Gottfredson, L. W. Sherman & B. C. Welsh (Eds.), *Evidence-based crime prevention* (1st ed., pp. 13–21). Routledge.
- Felson, M. (1994). *Crime and everyday life*. Sage.
- Fienberg, S., Larntz, K., & Reiss, A. J. (1976). Redesigning the Kansas City preventive patrol experiment. *Evaluation*, 3(1-2), 124–131.
- Frydl, K., & Skogan, W. (Eds.). (2004, April). *Fairness and Effectiveness in Policing*. The National Academies Press. <https://doi.org/10.17226/10419>

- Gates, D., & Knowles, L. (1976). Evaluation of the Rand Corporation's Analysis of the Criminal Investigation Process. *The Police Chief*, 43(7), 20–77.
- Gottfredson, M. R., & Hirschi, T. (1990). A General Theory of Crime. *A General Theory of Crime*. <https://doi.org/10.1515/9781503621794>
- Greenwood, P., Chaiken, J., & Peterselia, J. (1976). Response to: An evaluation of the RAND Corporation's analysis of the criminal investigation process". *The Police Chief*, 43(12), 62–71.
- Hayes, C. (1987). The Impact of Recent Research on the Detective Role. *The Police Journal: Theory, Practice and Principles*, 60(2), 97–111. <https://doi.org/10.1177/0032258X8706000202>
- Holliday, L. P. (1974). *A methodology for radio car planning* (tech. rep.). RAND. Rand Institute.
- Isaacs, H. H. (1967). *A Study of Communications, Crimes and Arrest in a Metropolitan Police Department. Appendix B of Science and Technology Task Force Report*. (tech. rep.). U.S. Government. Washington DC.
- Kelling, G. L., Pate, A. M., Dieckman, D., & Brown, C. (1974). *The Kansas City Preventive Patrol Experiment: A Summary Report* (tech. rep.). Police Foundation. Washinton D.C.
- Kelling, G. L., Pate, A. M., Dieckman, D., & Brown, Charles. (1974). *The Kansas City Preventive Patrol Experiment Technical report*. Police Foundation.
- Klockars, C. B. (1983). *Thinking about Police: Contemporary Readings*. McGraw-Hill.
- Koper, C. S. (1995). Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, 12(4), 649–672. <https://doi.org/10.1080/07418829500096231>
- Larson, R. C. (1975). What happened to patrol operations in Kansas City? A review of the Kansas City preventive patrol experiment. *Journal of Criminal Justice*, 3(4), 267–297. [https://doi.org/10.1016/0047-2352\(75\)90034-3](https://doi.org/10.1016/0047-2352(75)90034-3)
- Larson, R. C., & Cahn, M. F. (1985). *Synthesizing and Extending the Results of Police Patrol Studies* (tech. rep.). National Institute of Justice.
- Miron, H., Wasserman, R., & Rickard, T. (1979). *Managing criminal investigations: a handbook*.
- Pate, A. M. (1981). *The Newark Foot Patrol Experiment* (tech. rep.). Police Foundation. Washington.
- Pate, A. M., Ferrara, A., Bowers, R. A., & Lorence, J. (1976, January). *Police Response Time: Its Determinants and Effects* (tech. rep.). National Policing Institute. Washington, DC.
- Pate, A. M., Kelling, G. L., & Brown, C. (1975). A response to “what happened to patrol operations in Kansas City?” *Journal of Criminal Justice*, 3(4), 299–320. [https://doi.org/10.1016/0047-2352\(75\)90035-5](https://doi.org/10.1016/0047-2352(75)90035-5)
- Press, S. J. (1971). *Some effects of an increase in police manpower in the 20th precinct of New York city* (tech. rep.). New York City-Rand Institute. New York City. <https://doi.org/20.500.12592/nxwwjj>
- Prince, H., Lum, C., & Koper, C. S. (2021). Effective police investigative practices: an evidence-assessment of the research. *Policing: An International Journal*, 44(4), 683–707. <https://doi.org/10.1108/PIJPSM-04-2021-0054>

- Risman, B. J. (1980). The Kansas City Preventive Patrol Experiment: A Continuing Debate. *Evaluation Review*, 4(6), 802–808. <https://doi.org/10.1177/0193841X8000400605>
- Schnelle, J. F., Kirchner, R. E., Casey, J. D., Usselton, P. H., & Mcnees, M. P. (1977). Patrol evaluation research: A multiple-baseline analysis of saturation police patrolling during day and night hours. *Journal of applied behavior analysis*, 10, 33–40.
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place*. *Criminology*, 27(1), 27–56. <https://doi.org/10.1111/j.1745-9125.1989.tb00862.x>
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997, January). *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress* (tech. rep.). NIJ. Washington, DC.
- Sherman, L. W., & Rogan, D. P. (1995). Effects of gun seizures on gun violence: “Hot spots” patrol in Kansas City. *Justice Quarterly*, 12(4), 673–693. <https://doi.org/10.1080/07418829500096241>
- Sherman, L. W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime “hot spots”: A randomized, controlled trial. *Justice Quarterly*, 12(4), 625–648. <https://doi.org/10.1080/07418829500096221>
- Skolnick, J. H., & Bayley, D. H. (1986). *The new blue line: Police innovation in six American cities*. Simon; Schuster.
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Tarr, D. P. (1978). Analysis of response delays and arrest rates. *Journal of Police Science & Administration*, 6(4), 429–451.
- Trojanowicz, R. C. (1986). Evaluating a neighborhood foot patrol program: The Flint, Michigan project. In D. Rosenbaum (Ed.), *Community crime prevention: Does it work?* (1st ed., p. 157). Sage.
- U.S. Congress, H. o. R. (1995). *Making appropriations for the Departments of Commerce, Justice, and State, the Judiciary, and related agencies for the fiscal year ending September 30, 1996, and for other purposes (House Report 104-378, Sec. 116)* (tech. rep.). Washington.
- Weisburd, D., & Green, L. (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly*, 12(4), 711–735. <https://doi.org/10.1080/07418829500096261>
- Weisburd, D., & Majimundar, M. K. (Eds.). (2018, February). *Proactive Policing*. National Academies Press. <https://doi.org/10.17226/24928>
- Zimring, F. E. (1978). Policy Experiments in General Deterrence: 1970-1975. In *Deterrence and incapacitation estimating the effects of criminal sanctions on crime rates* (pp. 140–186). The National Academies Press. <https://doi.org/10.17226/18561-2>

II

INVESTIGATING POLICE INFLUENCE ON RESPONSE TIMES

4

THE EFFECT OF POLICE DEPLOYMENT STRATEGY ON EMERGENCY RESPONSE TIMES

4

AN AGENT-BASED MODELLING INVESTIGATION

Objectives This study investigates the impact of three police deployment strategies on emergency response times using agent-based modelling (ABM). Specifically, it evaluates the effectiveness of random patrol, stationary deployment (optimal spreading), and last location deployment (idling at last incident location). It further examines how key variables – urbanisation, call volume, and police capacity – moderate these effects.

Methods A detailed ABM was developed using NetLogo, integrating real-world data: historical calls for service (CFS), jurisdiction shapefiles, and street network data from the Netherlands. The model simulated police travel and response dynamics across 300 runs, varying deployment strategies, urbanisation levels, call volumes, and police capacities. Outputs were analysed to assess response times, fast response rates (< 3 minutes), and late response rates (> 13 minutes). Methods were pre-registered at <https://osf.io/yrwdp/>.

Results On average, stationary deployment reduced response times by 35% ($SD \pm 14\%$), increased fast responses by 74% ($SD \pm 40\%$), and decreased late responses by 66% ($SD \pm 33\%$) compared to random patrol. Last location deployment also outperformed random patrol, reducing response times by 13% ($SD \pm 9\%$), increasing fast responses by 22% ($SD \pm 14\%$), and reducing late responses by 42% ($SD \pm 36\%$). The advantages of stationary and last location deployment were most pronounced in rural areas and at lower police capacities. Urbanisation reduced the performance gap between strategies, while higher call volumes modestly diminished the relative benefits of stationary deployment.

Conclusions This study highlights the significant impact of police deployment strategies on response times and rapid interventions. These findings underscore the need for further research on rapid response. The modular ABM framework offers a valuable tool for adapting investigations to different policing contexts, enhancing external validity.

4.1 INTRODUCTION

The Standard Model of Policing consists of two core activities for police officers on the street: *rapid response* and *random patrol* (Frydl & Skogan, 2004). While many different models of policing have been proposed and implemented in law enforcement agencies worldwide, these two activities still remain central to most departments. Nevertheless, their effectiveness has long been subject to scrutiny, particularly since the 1970s, when a series of landmark studies questioned the effectiveness of both of these core activities.

First of all, The Kansas City Preventive Patrol Experiment (Kelling, Pate, Dieckman & Brown. Charles, 1974) cast doubt on the effectiveness of random patrol, concluding that it had minimal impact on crime or citizen perceptions of safety. Soon after, the Kansas City Response Time Analysis Study (Bieck, 1977a) and its subsequent replication study in four more cities (Spelman & Brown, 1981) challenged the utility of rapid response, asserting that most crimes are either discovered long after the perpetrator has fled or are not reported quickly enough by citizens for a rapid police response to be effective at increasing the chance of apprehension.

While both pillars of the Standard Model faced substantial critique, the academic response varied widely. The Kansas City Preventive Patrol Experiment spurred a robust debate (see Farrington, 1983; Fienberg et al., 1976; R. C. Larson, 1975, 1982; Pate et al., 1975; Risman, 1980; Sherman & Weisburd, 1995; D. Weisburd et al., 2023; Zimring, 1978), leading to the development of innovative strategies like hot spots policing (Sherman & Weisburd, 1995) and numerous follow-up studies evaluating its deterrent effects (see Braga, Turchan, Papachristos & Hureau, 2019). Conversely, the critique of rapid response spurred very little debate and led to a strong dismissal of rapid response and its potential benefits in the influential *Preventing Crime: What Works, What Doesn't, What's Promising* report – commonly referred to as the Maryland Report (Sherman et al., 1997). The report argued that rapid response does not have an effect on arrest rates and, by extension, crime levels and that response times are hard to improve as: ‘Cutting police travel time for such crime from 5 to 2.5 minutes could require a doubling of the police force.’ (Sherman et al., 1997, Chapter 8, p. 11) Moreover, further experimental investigation into the effects of rapid response was discouraged as: ‘... there is neither empirical nor theoretical justification for such expensive test ... further tests of this theory seem to be a waste of tax dollars.’ (Sherman et al., 1997, Chapter 8, p. 11)

As a result, experimental or quasi-experimental research into the rapid response strategy has never materialised. Some correlational studies have explored the predictors of response time (for an overview, see Verlaan & Ruiter, 2023b) and its effect on arrest likelihood, especially for in-progress burglaries (e.g., Cihan et al., 2012; Coupe & Blake, 2005). More recently, several studies in the economics literature have used methods from econometrics – such as instrumental variable – to estimate the causal effect of faster response on crime outcomes (Blanes i Vidal & Kirchmaier, 2018; Mastrobuoni, 2019; S. Weisburd, 2021). Although these studies generally report positive effects, their findings have not sparked significant follow-up work within criminology (with the exception of Rief & Huff, 2023).

However, a recent essay by [Author] argues that this neglect is unjustified, and that the prevailing view of rapid response is both outdated and overly simplistic. The essay reviews historical evidence showing that the effectiveness of rapid response is highly conditional, with substantial benefits observed for fast responses (i.e., under three minutes). Moreover, the essay highlighted that the (historically small) proportion of crimes amenable to rapid

response has likely increased due to declines in property crime rates (known as *the crime drop*), and might now even constitute the majority of crime calls. Lastly, it argues that recent advancements in communication technologies have mitigated many of the primary reasons for citizen reporting delays (such as: *finding a payphone* or *asking permission to use someone's landline*), further enhancing the potential impact of rapid response. These insights underscore the need for renewed empirical investigation.

Nevertheless, large-scale experimental research to explore police response times remains impractical due to high costs, logistical complexities, and ethical concerns. Additionally, lingering doubts persist about whether police agencies can meaningfully improve response times, further complicating the justification for such resource-intensive studies. Before embarking on large-scale experiments, it is vital to address these doubts and explore the dynamics underpinning police agency in improving response times.

To this end, computational methods provide a powerful alternative for studying emergency response times, allowing for controlled experimentation without real-world risks or costs. Over the years, various modelling approaches have been applied to different aspects of police operations, including vehicle selection (Dunnett et al., 2019); patrol beat design (Curtin et al., 2010; Kwak & Leavitt, 1984; R. C. Larson, 1974; Mitchell, 1972; Sacks, 2000; Zhang & Brown, 2012); patrol route design (Leigh et al., 2019); police station location optimisation (Aly & Litwhiler, 1979; Chow et al., 2015; R. C. Larson, 1974); police staffing (Taylor & Huxley, 1989); and patrol deployment distribution over beats (Chenevoy, 2022). However, with the exception of some research into ambulance response times (Jagtenberg et al., 2015, 2017; van Barneveld et al., 2016, 2018), no computational investigation has been done into the effect of *deployment strategies* on emergency response times – while this is identified as the primary variable under police control to increase police travel time.

Moreover, with the exception of Chenevoy (2022), these studies relied on equation-based models, which, while effective for optimisation problems, often require high levels of abstraction that limit their ability to capture real-world variability. More flexible computational methods, such as agent-based modelling (ABM), allow for greater realism by explicitly simulating individual police units, emergency calls, and their interactions within a spatially explicit environment. Unlike equation-based models, ABMs can capture real-life complexity and heterogeneity as they incorporate feedback loops, non-linearity, and path dependency.

ABMs have been used to investigate a wide range of crime problems and policing responses. Many of these studies are primarily theory-driven, using ABMs to test the generative sufficiency of criminological mechanisms. For example, Birks et al. (2012, 2014) demonstrate how routine activity theory, rational choice theory, and crime pattern theory can jointly explain the emergence of core empirical regularities such as repeat victimisation, spatial clustering, and journey-to-crime patterns. Birks and Davies (2017) similarly test competing hypotheses about the influence of street network structure on crime risk. E. R. Groff (2007, 2008) employ ABMs to assess how spatial and temporal constraints on offender movement shape robbery opportunities, while P. J. Brantingham and Tita (2008) and Bosse and Gerritsen (2008) examine the formation and displacement of crime hot spots.

Other studies demonstrate how ABMs can be used to simulate crime prevention strategies under hypothetical or artificial conditions. For instance, E. Groff and Birks (2008) explores the potential of simulation to evaluate alternative crime prevention interventions, while Malleson (2012) models burglary risks in Leeds in the context of urban regeneration. More

recently, D. Weisburd et al. (2017) and Wooditch (2023) have used agent-based simulations to assess the area-wide effects of hot spots policing and the potential gains of reallocating officers' uncommitted time to high-crime areas.

Despite this breadth, most ABM studies remain primarily exploratory or theoretical. To assess the potential for police agencies to meaningfully reduce response times, a more decision-support oriented approach is required – one that links simulation outcomes to actionable organisational choices under realistic operational constraints. The primary state-of-the-art example of such an approach is the work of Chenevoy (2022), who developed a simulation-based optimisation framework combining an agent-based model of patrol activities with a genetic algorithm to improve both response efficiency and crime deterrence. Her model is designed to support strategic decision-making by exploring deployment alternatives and identifying effective configurations in complex urban environments.

While Chenevoy's work represents a significant advance, it primarily addresses the strategic (long-term) allocation of police resources, optimising last location deployment patterns across broader spatial areas. The present study complements this by focusing on the tactical implications of deployment strategy – such as coverage schemes and dispatch protocols – at finer temporal and spatial scales. Rather than optimising configurations, it provides a flexible, empirically grounded simulation environment for evaluating the performance of alternative deployment concepts.

Specifically, the present study employs a highly detailed ABM to investigate the effects of police *deployment strategies* – the primary variable under police influence – on emergency response times, while also considering the moderating roles of key variables. The methodology expands on the chapter “Agent-based decision-support systems for crime scientists” by Birks and Townsley (2018), who demonstrated how ABMs can simulate police response dynamics and test the impact of various interventions using the agent-based modelling software NetLogo. While their approach relied on a simplified representation (i.e., imaginary geographic context and simulated crime data), the present study enhances the realism and granularity of the simulation by integrating real historical emergency calls for service (CFS) data, police jurisdiction shapefiles, and detailed street network data. Additionally, it incorporates observed dispatch dynamics derived from ethnographic research, all within the NetLogo framework (Wilensky, 1999).

In the Netherlands, as in many other parts of the world, deployment mostly takes the form of motorised patrol carried out by a one- or two-officer team. Motorised patrol ranges from undirected (often called *random*) to fully directed, and may even include optimally timed stops in hot spots (Koper, 1995). However, sometimes more stationary forms of deployment are also used, an example is the use of temporary mobile police units in hot spots or near the house of a repeat-victim.

As such, this paper evaluates the impact of three distinct police deployment strategies – *random patrol deployment* (i.e., randomly, and consecutively, selecting target nodes to travel towards), *stationary deployment* (i.e., continuously returning to a predetermined, optimally central, node in the network), *last location deployment* (i.e., remaining at the node of the last CFS) – on police response times through simulation. In the random patrol strategy, police units sequentially and randomly select nodes within their jurisdiction's street network, travel to these nodes via the shortest path, and idle for 15 minutes before repeating the process unless dispatched to an emergency call. The stationary deployment strategy involves

clustering nodes within a jurisdiction and assigning each police unit to the most central node in its cluster; when not responding to incidents, the units return to and remain at their designated (optimally distributed) stations. The last location deployment strategy positions police units at the location of their most recently handled emergency call. This approach assumes that recent incidents serve as a reasonable predictor for future calls, as crime and emergency events often exhibit spatio-temporal patterns.

Besides the effect of these three different police deployment strategies, we also investigate the effect of three variables suggested by Verlaan and Ruiter (2023b) to moderate the effect between these police deployment strategies and police response time. These moderating variables are: *police capacity*, *call volume* and *degree of urbanisation of the front line team*, the lowest level geographical unit of the Netherlands Police with a police station; analogous to the US *precinct*.

This aim leads to the development of the following research questions:

1. What is the impact of police deployment strategies on emergency response times?
 - a. How do urbanisation, call volume, and police capacity moderate these effects?

To systematically address these questions, we have developed six falsifiable hypotheses. These hypotheses are as follows:

1. *Stationary deployment* yields better response times than *random patrol deployment* operationalised as lower average response time, higher percentage fast responses (*i.e.* < 3 minutes) and lower percentage late responses (*i.e.* > 13 minutes).
2. *Last location deployment* yields better response times than *random patrol deployment* operationalised as lower average response time, higher percentage fast responses (*i.e.* < 3 minutes) and lower percentage late responses (*i.e.* > 13 minutes).
3. *Stationary deployment* yields better response times than *last location deployment* operationalised as lower average response time, higher percentage fast responses (*i.e.* < 3 minutes) and lower percentage late responses (*i.e.* > 13 minutes).
4. As *police capacity* increases, operationalised as the number of virtual police response units serving each front line team, the relative effect on *average response time*, percentages of *fast responses* and percentages late responses within *stationary* -, *last location* -, and *random patrol deployment* decreases.
5. As the *urbanisation* of the front line teams increase, operationalised based on its street length in meters per square kilometres, the relative difference in *average response time*, percentages *fast responses* and percentages *late responses* within *stationary* -, *last location* -, and *random patrol deployment* decreases.
6. As *call volume* increases, operationalised as high call volume shift or normal call volume shift, the relative difference in *average response time*, percentages of *fast responses* and percentages *late responses* within *stationary* -, *last location* -, and *random patrol deployment* decreases.

In answering these questions and testing these hypotheses, this research aims to provide actionable insights for both academic and practical contexts. By leveraging real-world data, including historical priority 1 CFS, jurisdictional shapefiles, and detailed street network data, the study seeks to bridge the gap between theoretical exploration and practical application. Furthermore, it aligns with recent calls to revisit long-standing assumptions about rapid response and its role in modern policing.

Importantly this study focuses specifically on how deployment strategy and moderating variables influence outcomes in reactive policing, while acknowledging that officers on the street engage in a range of other activities, including proactive patrol – an area that has received comparatively greater empirical attention in recent years (Braga, Turchan, Papachristos & Hureau, 2019). We argue that the simulation approach used here is particularly well suited to studying reactive policing, as empirical experimentation in this area would require direct manipulation of police responses to real incidents, presenting significant practical and ethical challenges. Moreover, a key strength of the simulation method is its ability to isolate reactive policing from other core activities, something that would also be exceptionally difficult using traditional empirical approaches. Future research could of course apply similar approaches to explore proactive policing strategies and further models could combine simulation of additional police activities to better understand the relative opportunity cost impacts of each on the others. For instance, integrating models of proactive patrol, community engagement, or investigative work could provide a more holistic view of how police resources are allocated and the trade-offs involved.

4.2 METHODOLOGY

This study follows a data-driven simulation approach, integrating real-world police data, geographical information, and agent-based modelling to assess the impact of deployment strategies on response times. The ABM developed simulates daily police incidents and responses in the District Oost-Utrecht over a full year, operating on a second-by-second basis. Using real-world police data on incident timing, location, and recorded response actions, the model replicates observed response patterns while systematically varying police deployment strategy and capacity. To ensure the behavioural realism of the model, the development process was informed not only by administrative data but also by extensive fieldwork, including systematic observation in dispatch centres, ride-alongs with patrol units, and follow-up interviews with dispatchers and officers. By grounding the simulation in both data and practice, the model generates synthetic counterfactuals that allow for a structured evaluation of different strategies under controlled conditions. This approach enables direct comparisons of response effectiveness across key performance metrics, such as average response time and the proportion of rapid responses.

The methodology first describes the primary data sources, including police incident and dispatch records alongside geographical data used to simulate police movement. Next, it details the ABM environment, specifying agent behaviour, deployment strategies, and dispatch dynamics. The simulation design is then outlined, including variations in police capacity and strategy. Finally, the data extraction and analysis section explains how response time metrics are generated within the ABM, extracted, and analysed using descriptive statistics and regression models.

4.2.1 PRIMARY DATA

The GMS data used in the ABM is obtained from the Netherlands Police and covers the whole of District Utrecht-Oost for the year 2019. The data are spread over two databases. The first database contains information about the incident, the second contains information about the police dispatches to these incidents; one incident can have multiple units dispatched to it. Table B.1 and Table B.2 in the Appendix represent the relevant pseudonymised variables which are used from each database.

In addition to police records, geographical data were incorporated to ensure realistic simulation of police movement and response times. Jurisdiction shapefiles were used to define the boundaries of each front line team, while street network data were retrieved from OpenStreetMap (OSM) via OSMnx (Boeing, 2017). The road network was filtered to include only drivable roads, with travel speeds and road restrictions applied accordingly. Table B.3 details the characteristics of the street network data.

PRIMARY DATA PRE-PROCESSING

The pre-processing code is openly available on the open GitHub page: [anonymised]. Here, we provide a short summary of the pre-processing. We begin with a number of filtering steps in order to obtain a representative sample of incidents, and all individual dispatches to these various incidents. First, we exclude incidents that were initiated by the police (as this constitutes proactive policing). Second, we remove incidents which did not receive a physical dispatch, but were, for example, handled over the phone or considered *false alarm*. Third, we remove dispatches to incidents by non-standard response units, e.g. helicopter, military police units etcetera. These steps leave 34,593 incidents and 50,694 individual dispatches to model in the ABM. However, there are a few more filtering steps needed to ensure data quality and representativeness: Fourth, we exclude dispatches that contain timestamp logs in the hour in which, in the Netherlands, the clock is set to daylight-saving hours: 2019-10-27 02:00:00 until 03:00:00. Fifth, we exclude dispatches for which *no cleared* timestamp is present in the data. Sixth, we exclude dispatches for which the duration between being dispatched to a call and clearing a call is less than five minutes, as these are often false alarms, resolved before police arrival, or otherwise erroneous. Seventh, we exclude dispatches which last longer than the duration of a shift; eight hours. Eighth, we exclude dispatched which arrive within one minute of being dispatched to a call. Lastly, we exclude calls for which the duration between arrival and clearing is less than one minute. After these last filtering steps, we are left with 30,380 (88%) incidents and 38,893 (77%) individual dispatches to these incidents. Table B.4 in the Appendix summarises these steps, together with the reason behind each filtering step and its effect on the size of the database. Lastly, we took steps to infer missing data in the arrived timestamp, see the open code for details.

4.2.2 AGENT-BASED MODEL DEVELOPMENT

This section describes the development of the ABM, including the model environment, agent behaviour, deployment strategies, dispatch dynamics, and experimental setup.

MODEL ENVIRONMENT

The agent-based model is built in NetLogo 6.2.2 (Wilensky, 1999) and simulates traversable street networks for the four front line teams within District Oost Utrecht. Shapefiles of

police districts and front line teams were sourced from the Netherlands Police, while street network data was retrieved from OpenStreetMap using OSMnx (Boeing, 2017). The network was constructed with the 'drive' specification, ensuring only drivable roads and paths were included. Each road segment in the network is characterised by parameters such as average travel speed for cars, directional flow, and true length in meters. For the Netherlands, these parameters in OSM are highly accurate and complete. The temporal dimension of the simulation is structured in ticks, where each tick corresponds to one real-world second. A graphical representation of the model is shown in Figure 4.1.

4

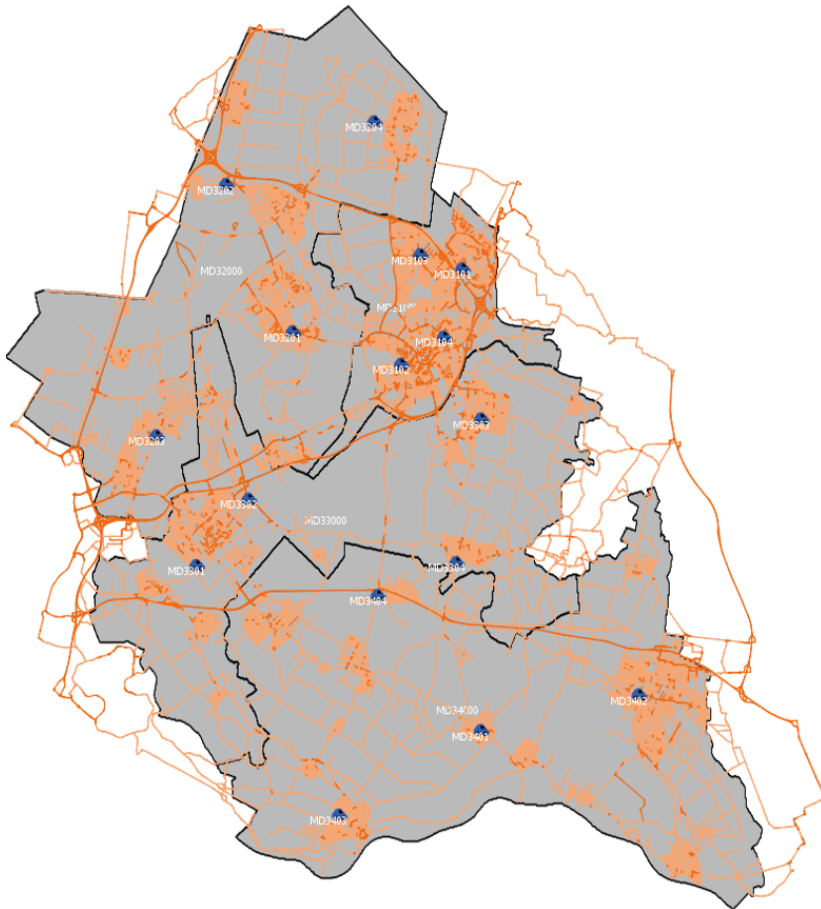


Figure 4.1: Graphical representation of the ABM model - instantiated with police capacity = 4

AGENTS: CALLS-FOR-SERVICE

The model incorporates real historical calls-for-service (CFS) data from the year 2019 within the jurisdiction of District Oost-Utrecht, which includes the four *front line teams*. These calls provide the foundation for simulating police response dynamics and are pre-processed to ensure representativeness and compatibility with the ABM framework. The details of the

CFS data are summarised in Table 4.1.

Each CFS is assigned a pseudonymised incident ID, enabling linkage with the original data while ensuring privacy. Incidents are also categorised by their priority level as assigned at the time of the call. To simplify the simulation, calls with lower priority levels (3, 4, and 5) are re-coded as priority level 2 due to their low frequency in the dataset. This adjustment maintains model efficiency without compromising realism.

CFS timestamps, indicating when the call was received by the dispatch centre, are operationalised as seconds elapsed since the start of 2019. This conversion aligns the data with the temporal structure of the ABM, where each tick corresponds to one second in real-world time.

The model further integrates operational details to accurately simulate resource allocation. This includes the number of police units deployed to each incident (*reactive_supply*) and a list of on-scene durations (*on_scene_times*), capturing the time each unit remains engaged with the incident. Additionally, spatial precision is ensured by mapping each incident to specific x and y coordinates (*Xcor* and *Ycor*) within the NetLogo world, aligning the incidents with the simulated street network.

Table 4.1: Pre-processed CFS data input in the ABM

Variable name	Description
Incident_id	Unique incident identifier
Priority_level	Priority level of the incident
Start_incident	Timestamp of incident recorded by dispatch centre, operationalised as seconds since start of the year 2019
Reactive_supply	Number of virtual police units to respond to the incident
On_scene_times	List of on-scene times (<i>arrived_cleared</i>)
Xcor	Integer representing the x coordinate of the incident on the street network, in the format of the NetLogo world
Ycor	Integer representing the y coordinate of the incident on the street network, in the format of the NetLogo world

AGENTS: VIRTUAL POLICE RESPONSE UNITS

The agent-based model simulates a population of virtual police response units to represent the operational capabilities of law enforcement within each front line team. These agents act as individual police units, dynamically responding to CFS and interacting with the simulated environment in real time.

The number of virtual police units per front line team varies between 1 and 10, depending on the instantiation of the simulation run. This variation allows for an exploration of the moderating effects of differing police capacities on the relationship between police strategy and response times. At the start of the simulation (tick = 0), all agents are instantiated and deployed within the street network of their respective front line team jurisdictions.

For simplicity and to focus on response dynamics, the model does not include mechanisms for returning police units to their stations at the end of shifts. Additionally, the size of the police force remains constant throughout the simulation, irrespective of typical fluctuations in staffing levels during day and night shifts. While these simplifications may reduce fidelity in certain contexts, they allow for a controlled examination of deployment strategies and response outcomes without the added complexity of shift-based staffing adjustments.

Each virtual police unit is equipped with a set of behavioural rules that govern their movement, response to dispatches, and idle behaviour. These rules vary depending on the active deployment strategy, ensuring the agents accurately represent the operational logic of different patrol methods. By incorporating individual decision-making processes and interactions with the environment, the agents are capable of capturing emergent patterns in response times and resource allocation under various scenarios.

MODEL DYNAMICS: CALL BEHAVIOUR

The behaviour of CFS is driven by real-world historical data. Each call becomes active at the simulation tick corresponding to its *start_incident* timestamp, representing the time the call was received by the dispatch centre.

When a call becomes active, the model assesses the availability of the required number of police units (*reactive_supply*) within the front line team. If units are available, they are dispatched to the incident. When fewer units are available than required, the maximum number of available units is dispatched, and the remaining demand is added to a dispatch queue. Calls in the queue are prioritised by urgency (priority level), with waiting time determining the order of dispatch for calls of the same priority. This queuing system follows a first-in, first-out strategy for calls of equal priority.

Once dispatched, police units travel along the shortest path to the incident location, observing road speed limits (always travelling at the max travel speed of a road section) and restrictions. Upon arrival of the first unit, the response time is recorded as the elapsed time since *start_incident*. The units then remain on scene for the duration specified in the call data (*on_scene_times*), during which they are unavailable to respond to other calls. However, units can be redirected mid-response to handle higher-priority incidents. In such cases, the interrupted call re-enters the dispatch queue, and the interrupted on-scene time is preserved for completion in the future.

MODEL DYNAMICS: AGENT BEHAVIOUR

The actions of virtual police response units are divided into two phases: their behaviour during response to active calls and their behaviour during idle periods.

When responding to a call, agents traverse the street network to reach the incident location via the shortest possible route, calculated based on road speed and travel time. Upon arrival, agents remain at the scene for the allocated time before becoming available for new assignments. During this response phase, agents can be dynamically reallocated if a higher-priority call arises. This ensures that police response prioritises the most urgent incidents.

During idle periods, the behaviour of agents is determined by the deployment strategy assigned in the simulation. Under the *random* deployment strategy, agents sequentially and randomly select nodes within their front line team jurisdiction and travel to these nodes, idling at each location for 15 minutes before repeating the process. This pattern mimics undirected patrol operations commonly observed in real-world policing.

The *stationary deployment* strategy assigns agents to optimally distributed locations within their front line team. Street network nodes are divided into n clusters (n being equal to the number of initialised virtual response units in the particular simulation run) using k-means clustering, and agents are stationed at the node with the highest closeness centrality within each cluster. When not responding to incidents, agents return to their designated

stations via the shortest path and remain there until dispatched to another call. This strategy reflects a more structured and centralised approach to resource deployment.

Under the *last location deployment* strategy agents remain idle at the location of their most recently handled call for service. This approach seeks to leverage the temporal and spatial clustering of incidents by assuming that recent call locations serve as effective predictors for future demand, and provides an alternative to centralised or random patrol patterns.

4.2.3 MODEL ROBUSTNESS

The ABM presented in this study is deliberately designed to be straightforward and transparent (apart from the open code, a visual representation of the model – in the form of a flow diagram – is available on OSF. The model relies heavily on real data and operational mechanisms observed and documented during extensive fieldwork.

The only substantive model parameter that requires an assumption is vehicle speed. We model this using road-specific speed limits from OpenStreetMap (OSM), which is known to be highly accurate in the Netherlands. Agents follow the maximum allowed speed for each road segment. We chose not to perform formal sensitivity analysis on this parameter, as we consider it the most defensible and realistic approximation available and the value most central to the phenomenon being modelled.

4.2.4 EXPERIMENTAL SETUP

Each experiment simulates the same calls-for-service within District Utrecht-Oost for the exact same 365 days (in 2019) on a second-by-second basis. The simulation experiments systematically vary the key independent variables: *police deployment strategy* and *police capacity*. Each experiment is run under a combination of these conditions, allowing for a structured assessment of their effects on response times. Police deployment strategy has three levels: random patrol, stationary deployment, and last location deployment. Police capacity varies from one to ten units per front line team. The range of 1 to 10 units per front line team was selected to capture both typical and extreme staffing scenarios. Empirical data (see Figure 6) show that most front line teams operate with three to seven active units per shift. Including values outside this range enables the model to explore the impact of under- and over-capacity conditions on response outcomes, which may be relevant for resource planning or stress testing deployment strategies.

These two parameters result in thirty unique simulation scenarios. The full set of simulation parameters is summarised in Table 4.2. The *last location* and *stationary* deployment models are fully deterministic, and the *random* deployment models is highly deterministic with only minor influences of stochasticity. Hence, multiple random seed reruns of the models were found to be unnecessary.

4.2.5 MODEL OUTPUTS AND DATA ANALYSIS

MODEL OUTPUT DATA

Each simulation run generates a dataset capturing all dispatch-level events, structured similarly to the original police dispatch input data (see Table B.2). These generated data include a record of each dispatch event, capturing the police unit involved, the corresponding incident, a unique dispatch and key time stamps. This results in 30 output datasets, each corresponding to one of the unique simulation runs. Table 4.3 details these output data.

Table 4.2: Simulation parameters

Simulation parameters	Description	Operationalisation
Police deployment strategy	A navigation strategy for police units not currently responding to a call	random patrol deployment; stationary deployment; last location deployment
Police capacity	The number of virtual police response units per front line team in the simulation	1–10 (increment: 1)

Table 4.3: Structure of ABM dispatch level output data

Call sign	Incident ID	Dispatch ID	Timestamp	Status
Police call sign	Unique ID	Unique ID	In ticks (seconds since start 2019)	Levels = dispatched, arrived, cleared, cancelled

DATA ANALYSIS APPROACH

The analysis focuses on quantifying response times and evaluating the effects of different deployment strategies under varying conditions. Following Verlaan and Ruiter (2023b), response time is operationalised as the elapsed time (in seconds) between the dispatch timestamp and the arrival of the first unit on the scene. To test the hypotheses, response times are analysed using descriptive statistics and visualised per front line team and police capacity level. As preregistered, falsification of the hypotheses is attempted based on mean and median response times, as well as distributions of fast responses (under three minutes) and late responses (over 13 minutes). The results are compared across deployment strategies, while controlling for urbanisation level, call volume, and police capacity. The variables included in the analysis and their sources are detailed in Table 4.4.

Table 4.4: Dependent and independent variables

Variables	Operationalisation	Source
Response time	(s)	ABM
Deployment strategy	Random; Stationary; Last location	ABM
Urbanisation rank	standardised (0–100) front line team rank: <i>meters of streets/km²</i>	GIS
Police capacity	Police units/shift	ABM
Call volume	Calls for service/shift	GMS

HYPOTHESIS TESTING AND EXPLORATORY ANALYSIS

The results section systematically attempts falsification of the hypotheses using descriptive plots, regression and statistical summaries. The analysis first examines response time distributions for each strategy per front line team and police capacity level. Following Verlaan and Ruiter (2023b), and in order to ensure comparability, findings are only interpreted for priority 1 calls, where all independent variables can be considered limiting factors.

Falsification of hypotheses 1–3 is conducted under all conditions of police capacity and per front line team but only within the ‘normal’ call volume condition. Hypothesis 4 is tested across all deployment strategies and front line teams but again restricted to normal

call volumes. Hypothesis 5 is evaluated across all strategies and police capacities, while hypothesis 6 is tested under all strategy and front line team conditions. Since different scenarios may yield different results, hypotheses may be falsified under some conditions but not others.

All simulation code, pre-processing scripts, and output visualisations will be made openly available on the project's GitHub repository.

4.3 RESULTS

The results presented below are based on Figures 4.2, 4.3 and 4.4. These figures provide insights into the outcome metrics of mean response time, percentage of fast responses (< 3 minutes), and percentage of late responses (> 13 minutes) for each of the scenarios (combination of jurisdiction and initialised police capacity level). More specifically, Figure 4.2 shows the response time distribution for each scenario, Figure 4.3 details the mean response time, percentage fast responses, and percentage late responses per scenario, and Figure 4.4 investigates the relative impact of incremental increases in call volume.

To provide context to these results, we first examine the empirical distribution of priority 1 call volume per shift and police capacity per shift for each real-world front line team, as presented in Figure B.1 and Figure B.2 (in Appendix C). A substantial proportion of shifts – approximately a quarter to a third – receive no priority one calls. Among shifts that do, most shifts receiving around six calls per shift ($SD = 5$). Meanwhile, police capacity per shift shows less variation within each front line team, with most shifts operating at five police units ($SD = 2$). These distributions inform the conditions under which different deployment strategies are evaluated in the subsequent analyses.

4.3.1 AVERAGE EFFECT OF DEPLOYMENT STRATEGIES

Across all scenarios, stationary deployment consistently outperformed both last location and random patrol strategies in terms of mean response time, percentage of fast responses (< 3 minutes), and percentage of late responses (> 13 minutes). As detailed in Figure 4.2, on average among the 40 different scenarios (4 front line teams $\times$ 10 capacity levels), stationary deployment reduced mean response times by 35% ($\pm 14\%$) compared to random patrol. Similarly, it achieved substantially higher rates of fast responses (74% $\pm$ 40%) and reduced late responses by 66% ($\pm 33\%$). Last location deployment also exhibited substantial improvements over random patrol, particularly in scenarios with higher police capacity. Mean response times were reduced by 13% ($\pm 9\%$), with fast response rates increasing by 22% ($\pm 14\%$) and late response rates decreasing by 42% ($\pm 36\%$).

Figure 4.3 illustrates the potential of stationary deployment in absolute terms: in the urban context of front line team Amersfoort, under the average capacity of 5 police units, *stationary* deployment enables police to (theoretically) respond to 75% of all priority 1 calls in under 3 minutes – approximately twice as many as under *last location* deployment (38%) or random patrol (32%). Even with police capacity doubled to 10 units, *last location* deployment would barely be able to respond to 65%, and *random* patrol would still fail to reach 50% of such calls within the same timeframe.

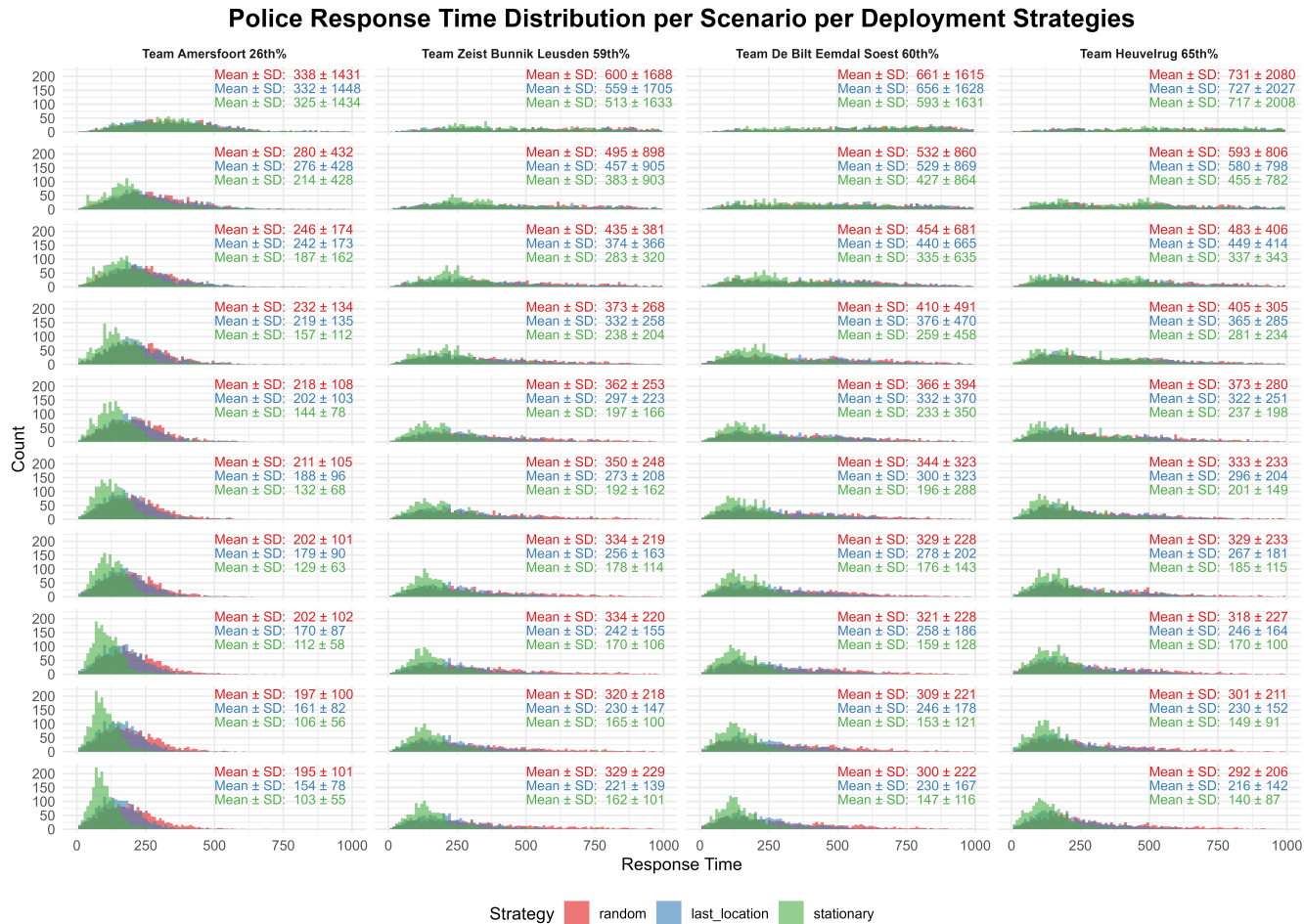


Figure 4.2: Police Response Time Distribution per Deployment Strategy

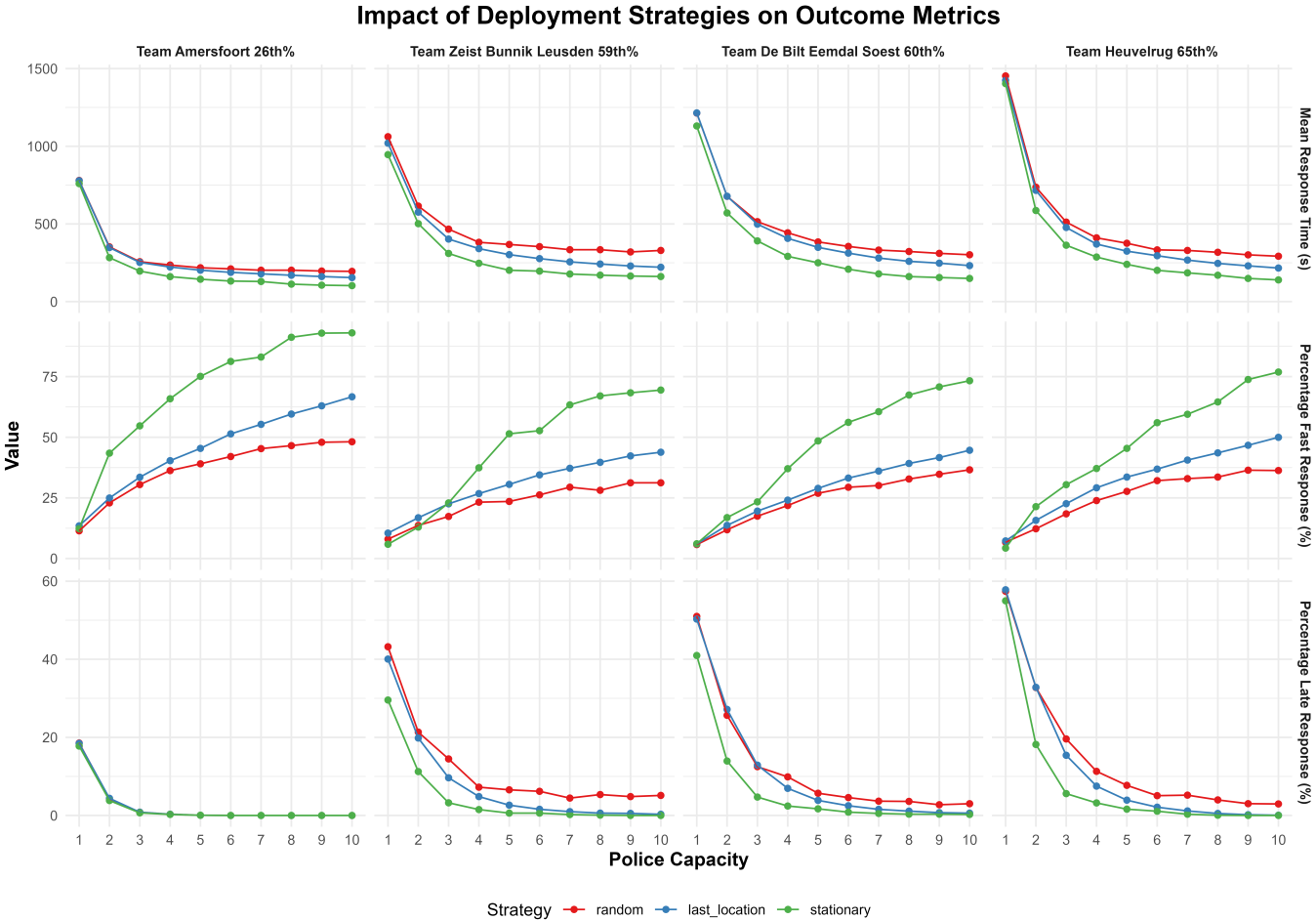


Figure 4.3: Impact of Deployment Strategies on Outcome Metrics

4.3.2 MODERATING EFFECT OF POLICE CAPACITY

As police capacity increased, mean response time and the percentage of fast and late responses improved in absolute terms across all deployment strategies. However, in line with hypothesis 4, both the relative as well as the absolute marginal improvement per additional police unit diminished. Nevertheless, as is evident from Figure 4.3, the marginal benefit of each additional police unit is larger under *stationary* deployment than under *last location* and *random* deployment for all three measures.

For *percentage of late response*, we observe saturation (meaning all calls are serviced with the timeframe of 13 minutes) across all front line teams within the currently tested levels of police capacity. For *percentage of fast response*, this saturation does not yet take place, although it is close for front line team Amersfoort under stationary deployment. In contrast, for *random patrol*, it is not clear if it will ever reach saturation, as the curve appears to asymptote below the 50% line.

Ultimately, across virtually all scenarios and metrics, stationary deployment with a police capacity of four outperforms last location and random with a police capacity of ten.

4.3.3 MODERATING EFFECT OF URBANISATION

Across the board, it is evident that the more rural a jurisdiction is (all else being equal), the higher the response times and higher the percentage of late responses. However, this is mainly due to having a higher baseline, the marginal effects per additional police capacity does not appear to vary per front line team over any of the strategies. It is noteworthy that we do not witness the same dynamic for fast response. In fact, while we do observe substantially higher percentages of fast response in the urban jurisdiction of front line team Amersfoort, there is no observable difference between the other more rural jurisdictions for similar police capacity levels.

4.3.4 MODERATING EFFECT OF CALL VOLUME

Figure 4.4 details the effect of (the empirical fluctuation of) call volume on the three outcome metrics. The effect of call volume on response metrics is not strictly monotonic across strategies and jurisdictions. While mean response time and late response rates show some discernible increases as call volume rises, the relationship is less clear for fast responses. The impact is also more pronounced under stationary deployment, likely because its efficiency depends on pre-positioning, which becomes less relevant when calls accumulate, forcing units into a queue. As queueing increases and units have reduced idling time, deployment strategy plays a diminishing role in performance, effectively eroding the advantage of better strategies.

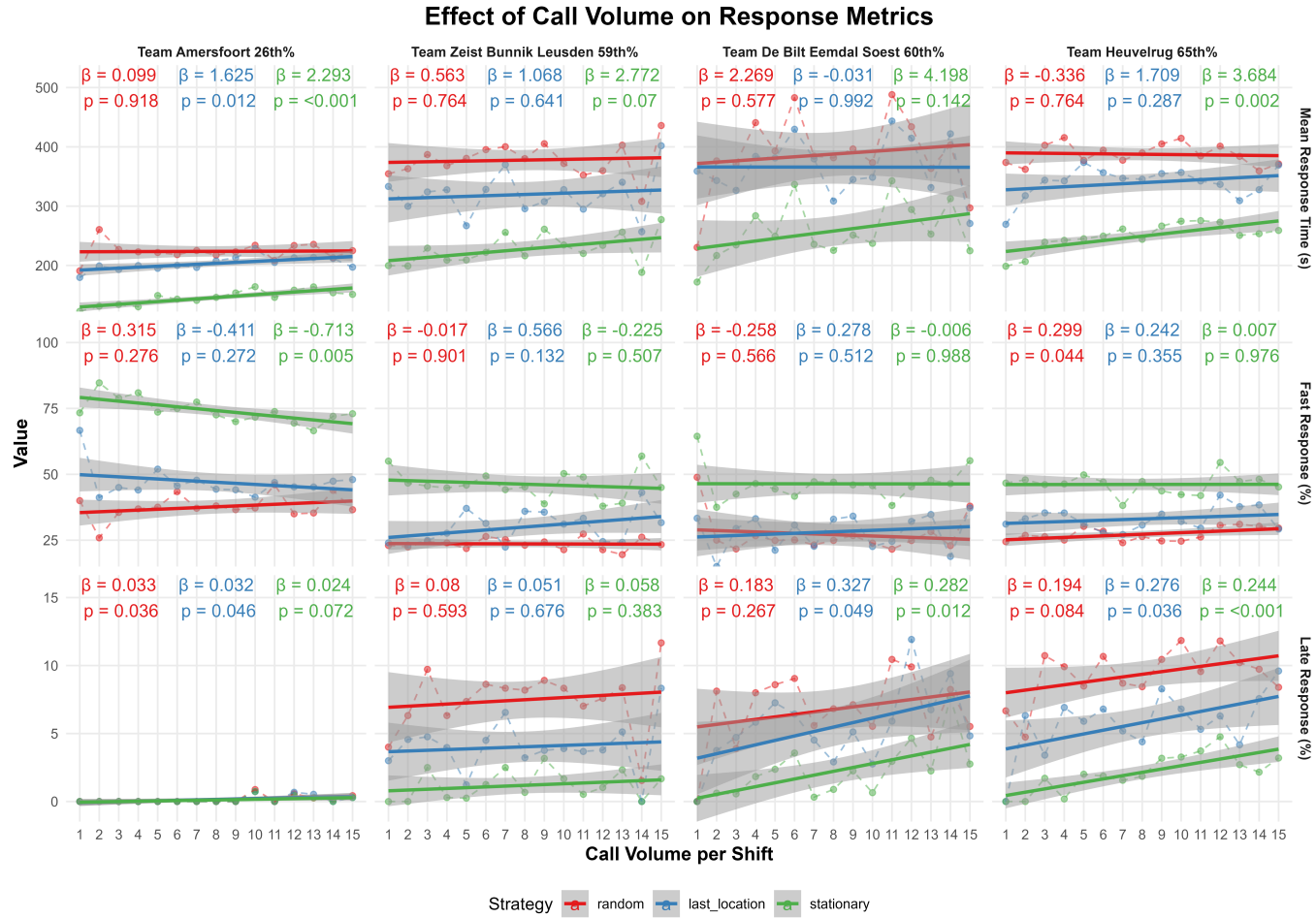


Figure 4.4: Impact of Call Volume per front line team per Deployment Strategy (for police capacity 3 - 7)

As such, at higher call volumes, the differences between deployment strategies diminish, as reduced idle time and increased queuing constrain the advantages of more effective strategies.

For realistic policing conditions – where front line teams typically operate with three to seven units per shift (see Figure B.2) and priority 1 call volumes rarely exceed fifteen per shift (see Figure B.1) – the overall effect of call volume is limited:

As call volume increases, mean response times also rise – by approximately 10–20 seconds under random patrol, 20–30 seconds under last location deployment, and up to 40 seconds under stationary deployment, where high demand begins to erode its advantage. In contrast, the proportion of fast responses declines by 10–15 percentage points under random patrol, 7–12 under last location, and 5–10 under stationary deployment and late responses increase by 3–5 points for random patrol, 2–4 for last location, and only 1–3 for stationary – reinforcing the stationary strategy's relative resilience even under pressure.

Crucially, stationary deployment at the highest observed call volumes still outperforms random and last location deployment at the lowest call volumes. Even when facing the most demanding conditions, stationary deployment achieves substantially lower response times, higher fast response rates, and fewer late responses than the other strategies under optimal conditions. This highlights that deployment strategy is a far greater determinant of performance than variations in call volume, reinforcing the importance of an effective positioning approach.

4

4.4 SUMMARY

The findings strongly support hypotheses 1–3: stationary deployment consistently outperforms both last location and random patrol across all response metrics. Even under the highest call volumes, it achieves faster response times, more fast responses, and fewer late responses than last location and random patrol under optimal conditions.

Hypothesis 4 is also confirmed, as marginal gains from additional police capacity diminish across all strategies, though stationary deployment benefits most. Hypothesis 5 receives partial support – while urban areas see lower response times and fewer late responses, the effect on fast response rates is less clear.

The results for hypothesis 6 (call volume effects) are mixed. While response times and late responses increase with more calls, the effect remains relatively minor within realistic policing conditions. More importantly, as call volume rises, differences between strategies narrow, particularly for stationary deployment, where its effectiveness is increasingly nullified by workload saturation.

Ultimately, deployment strategy is the dominant factor in determining response times, outweighing the effects of urbanisation, call volume and police capacity.

4.5 DISCUSSION

This study set out to investigate the impact of police deployment strategies – stationary deployment, last location deployment, and random patrol deployment – on emergency response times and to examine how these effects are moderated by police capacity, urbanisation, and call volume. By employing a detailed ABM that integrates real-world CFS data, geographical information, and dispatch dynamics, we have provided actionable insights into the dynamics of police response times under varying conditions.

The findings reveal that stationary deployment – whereby response vehicles relocate themselves to a central location within their respective jurisdiction after resolving each incident – consistently outperforms both last location deployment and random patrol across all key metrics of response time, percentage of fast responses (< 3 minutes), and percentage of late responses (> 13 minutes). This superiority is particularly evident in fast responses, where it produces a large and decisive impact.

These findings contribute to the reinvigorated academic debate surrounding rapid response in modern policing. The results demonstrate a substantial potential for the police to reduce response times and increase fast responses through the selection of deployment strategies. This challenges Sherman's (1997, Chapter 8, p. 11) conjecture that cutting response times in half may require doubling the police force. Instead, the choice of deployment strategy shows a much larger effect size, underscoring the agency that police departments have in improving response times without significant increases in resources. Coupled with rapid response literature on the effect of fast responses on on-scene arrests (see Bieck, 1977a; Blanes i Vidal & Kirchmaier, 2018; Cihan et al., 2012; Clawson & Chang, 1977; Isaacs, 1967; Spelman & Brown, 1981; Tarr, 1978), these findings indicate that police departments can substantially improve their arrest and clearance rates through their deployment strategy.

Beyond theoretical contributions, these findings also hold direct implications for practice. Stationary deployment is particularly effective for either increasing fast responses or limiting late responses, making it valuable in understaffed conditions where meeting on-time response quotas – such as the 90% threshold mandated in countries like the Netherlands – is critical.

From an academic perspective, the study underscores the potential of ABM as a decision-support tool for exploring complex policing dynamics. By explicitly simulating individual police units, CFS, and their interactions within a spatially explicit environment, the ABM provides a level of granularity and realism that is difficult to achieve with traditional equation-based models.

A final theoretical insight emerged from the performance of last location deployment – whereby officers remain at the location of the most recently resolved incident until they are required to respond to the next. While this deployment strategy performs poorly under low police capacities, it demonstrates increasing effectiveness as capacity rises. To us, this pattern suggests a fundamental misalignment between the cyclical nature of hotspots – where recent incidents predict future activity – and the sequential nature of police response. The assumption that the location of the last call serves as an effective predictor of the next call breaks down when intervening calls from other hotspots redirect police units. This sequential reality introduces delays that disrupt the predictive power of the strategy. However, as police capacity increases, this misalignment diminishes, as more units are available to simultaneously address calls from multiple hotspots.

4.5.1 LIMITATIONS

While this study provides valuable insights, several limitations should be acknowledged. First, like all models, the ABM's external validity is inherently constrained by the context of its design. The simulation focuses on Dutch front line teams, with specific geographic, demographic, and operational characteristics. While these findings may generalise to similar contexts, their applicability to other policing systems (e.g., in the US or UK) requires further exploration.

Second, the model incorporates several simplifying assumptions to enhance computational efficiency and isolate the effects of deployment strategies. For instance, the model does not account for shifts or the practice of returning police units to stations at the end of a shift. Similarly, non-reactive policing activities, such as proactive patrols or community engagement, are not simulated, potentially limiting the holistic applicability of the findings.

This focus on reactive policing means the model does not capture potential trade-offs or synergies between emergency response and other police functions, such as proactive patrol or community engagement. As such, the results should be interpreted as highlighting response time dynamics in isolation and ignorant of the proactive task of policing.

4.5.2 FUTURE RESEARCH AND BROADER IMPLICATIONS

The demonstrated influence of deployment strategies on response times underscores the need for renewed empirical investigations into rapid response, a topic long side-lined in policing research. While randomised controlled trials (RCTs) may remain impractical due to cost and ethical concerns, quasi-experimental and observational studies offer viable alternatives to advance this line of inquiry. As [anonymised] highlights, these methods provide rigorous and actionable insights, addressing many of the constraints associated with large-scale experimentation.

Additionally, the modular nature of the ABM developed in this study allows for its adaptation to diverse policing contexts. Future applications in different geographical and operational settings, such as urban US precincts or rural UK constabularies, would enhance the external validity of the findings and identify context-specific dynamics. By extending the model's scope, we hope to support a broader understanding of rapid response and inform evidence-based strategies for improving public safety.

Further research could also explore the potential synergy between stationary deployment and hot spots policing. Rather than relocating units purely based on spatial centrality, future models could incorporate spatial and temporal crime and call-for-service patterns to identify optimal fixed node placements. This would allow agencies to evaluate whether positioning patrol units near chronic high-demand locations could achieve both efficient emergency response and increased deterrence. By integrating insights from spatial crime analysis with deployment strategy design, such models could bridge the gap between proactive and reactive policing priorities.

In conclusion, this study demonstrates the substantial potential for police agencies to influence response times through strategic deployment, challenging longstanding assumptions and providing a strong foundation for further research and policy development. The findings emphasise that renewed focus on rapid response is not only feasible but overdue, offering a clear opportunity to bridge the gap between theoretical insights and practical advancements in policing.

REFERENCES

- Aly, A. A., & Litwhiler, D. W. (1979). Police Briefing Stations: A Location Problem. *A I I E Transactions*, 11(1), 12–22. <https://doi.org/10.1080/05695557908974395>
- Bieck, W. H. (1977a). *Response Time Analysis: Executive summary* (tech. rep.). National Institute of Justice. Kansas City.
- Birks, D., & Davies, T. (2017). Street Network Structure and Crime Risk: An Agent-Based Investigation of the Encounter and Enclosure Hypotheses. *Criminology*, 55(4), 900–937. <https://doi.org/10.1111/1745-9125.12163>
- Birks, D., & Townsley, M. (2018). Agent-based decision-support systems for crime scientists. In R. Wortley, A. Sidebottom, N. Tilley & G. Laycock (Eds.), *Routledge handbook of crime science* (pp. 334–349). Routledge. <https://doi.org/https://doi.org/10.4324/9780203431405-25>
- Birks, D., Townsley, M., & Stewart, A. (2012). Generative Explanations of Crime: Using Simulation to Test Criminological Theory. *Criminology*, 50(1), 221–254. <https://doi.org/10.1111/J.1745-9125.2011.00258.X>
- Birks, D., Townsley, M., & Stewart, A. (2014). Emergent Regularities of Interpersonal Victimization. *Journal of Research in Crime and Delinquency*, 51(1), 119–140. <https://doi.org/10.1177/0022427813487353>
- Blanes i Vidal, J., & Kirchmaier, T. (2018). The Effect of Police Response Time on Crime Clearance Rates. *The Review of Economic Studies*, 85(2), 855–891. <https://doi.org/10.1093/restud/rdx044>
- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>
- Bosse, T., & Gerritsen, C. (2008). Agent-based simulation of the spatial dynamics of crime: on the interplay between criminal hot spots and reputation. In L. Padgham, D. C. Parkes, J. Mueller, S. Parsons & L. Cao (Eds.), *Proceedings of the seventh international joint conference on autonomous agents and multi-agent systems (aamas'08)* (pp. 1129–1136). ACM Press. <https://doi.org/10.5555/1402298.1402378>
- Braga, A. A., Turchan, B., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1046>
- Brantingham, P. J., & Tita, G. (2008). Offender Mobility and Crime Pattern Formation from First Principles. In *Artificial crime analysis systems* (pp. 193–208). IGI Global. <https://doi.org/10.4018/978-1-59904-591-7.ch010>
- Chenevoy, N. L. (2022). *Determining optimal police patrol deployments: a simulation-based optimisation approach combining agent-based modelling and genetic algorithms* [Doctoral dissertation, University of Leeds].
- Chow, A. H. F., Cheung, C. Y., & Yoon, H. T. (2015). Optimization of Police Facility Locationing. *Transportation Research Record: Journal of the Transportation Research Board*, 2528(1), 60–68. <https://doi.org/10.3141/2528-07>
- Cihan, A., Zhang, Y., & Hoover, L. (2012). Police Response Time to In-Progress Burglary. *Police Quarterly*, 15(3), 308–327. <https://doi.org/10.1177/1098611112447753>
- Clawson, C., & Chang, S. K. (1977). The relationship of response delays and arrest rates. *Journal of Police Science and Administration*, 5(1), 53–68.

- Coupe, R. T., & Blake, L. (2005). The effects of patrol workloads and response strength on arrests at burglary emergencies. *Journal of Criminal Justice*, 33(3), 239–255. <https://doi.org/10.1016/j.jcrimjus.2005.02.004>
- Curtin, K. M., Hayslett-McCall, K., & Qiu, F. (2010). Determining optimal police patrol areas with maximal covering and backup covering location models. *Networks and Spatial Economics*, 10(1), 125–145. <https://doi.org/10.1007/S11067-007-9035-6>
- Dunnett, S., Leigh, J., & Jackson, L. (2019). Optimising police dispatch for incident response in real time. *Journal of the Operational Research Society*, 70(2), 269–279. <https://doi.org/10.1080/01605682.2018.1434401>
- Farrington, D. P. (1983). Randomized Experiments on Crime and Justice. *Crime and Justice*, 4, 257–308. <https://doi.org/10.1086/449091>
- Fienberg, S., Larntz, K., & Reiss, A. J. (1976). Redesigning the Kansas City preventive patrol experiment. *Evaluation*, 3(1-2), 124–131.
- Frydl, K., & Skogan, W. (Eds.). (2004, April). *Fairness and Effectiveness in Policing*. The National Academies Press. <https://doi.org/10.17226/10419>
- Groff, E., & Birks, D. (2008). Simulating Crime Prevention Strategies: A Look at the Possibilities. *Policing*, 2(2), 175–184. <https://doi.org/10.1093/police/pan020>
- Groff, E. R. (2007). Simulation for theory testing and experimentation: An example using routine activity theory and street robbery. *Journal of Quantitative Criminology*, 23(2), 75–103. <https://doi.org/10.1007/S10940-006-9021-Z>
- Groff, E. R. (2008). Adding the Temporal and Spatial Aspects of Routine Activities: A Further Test of Routine Activity Theory. *Security Journal*, 21(1-2), 95–116. <https://doi.org/10.1057/palgrave.sj.8350070>
- Isaacs, H. H. (1967). *A Study of Communications, Crimes and Arrest in a Metropolitan Police Department. Appendix B of Science and Technology Task Force Report*. (tech. rep.). U.S. Government. Washington DC.
- Jagtenberg, C. J., Bhulai, S., & van der Mei, R. D. (2015). An efficient heuristic for real-time ambulance redeployment. *Operations Research for Health Care*, 4, 27–35. <https://doi.org/10.1016/j.orhc.2015.01.001>
- Jagtenberg, C. J., Bhulai, S., & van der Mei, R. D. (2017). Dynamic ambulance dispatching: is the closest-idle policy always optimal? *Health Care Management Science*, 20(4), 517–531. <https://doi.org/10.1007/s10729-016-9368-0>
- Kelling, G. L., Pate, A. M., Dieckman, D., & Brown, Charles. (1974). *The Kansas City Preventive Patrol Experiment Technical report*. Police Foundation.
- Koper, C. S. (1995). Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, 12(4), 649–672. <https://doi.org/10.1080/07418829500096231>
- Kwak, N. K., & Leavitt, M. B. (1984). Police Patrol Beat Design: Allocation of Effort and Evaluation of Expected Performance. *Decision Sciences*, 15(3), 421–433. <https://doi.org/10.1111/j.1540-5915.1984.tb01227.x>
- Larson, R. C. (1974). A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, 1(1), 67–95. [https://doi.org/10.1016/0305-0548\(74\)90076-8](https://doi.org/10.1016/0305-0548(74)90076-8)

- Larson, R. C. (1975). What happened to patrol operations in Kansas City? A review of the Kansas City preventive patrol experiment. *Journal of Criminal Justice*, 3(4), 267–297. [https://doi.org/10.1016/0047-2352\(75\)90034-3](https://doi.org/10.1016/0047-2352(75)90034-3)
- Larson, R. C. (1982). Critiquing Critiques: Another Word on the Kansas City Preventive Patrol Experiment. *Evaluation Review*, 6(2), 285–293. <https://doi.org/10.1177/0193841X8200600209>
- Leigh, J., Dunnett, S., & Jackson, L. (2019). Predictive police patrolling to target hotspots and cover response demand. *Annals of Operations Research*, 283(1-2), 395–410. <https://doi.org/10.1007/S10479-017-2528-X>
- Malleson, N. (2012). Using Agent-Based Models to Simulate Crime. In *Agent-based models of geographical systems* (pp. 411–434). Springer Netherlands. https://doi.org/10.1007/978-90-481-8927-4{_}_}19
- Mastrobuoni, G. (2019). Police disruption and performance: Evidence from recurrent redeployments within a city. *Journal of Public Economics*, 176, 18–31. <https://doi.org/10.1016/j.jpubeco.2019.05.003>
- Mitchell, P. S. (1972). Optimal Selection of Police Patrol Beats. *The Journal of Criminal Law, Criminology, and Police Science*, 63(4), 577. <https://doi.org/10.2307/1141814>
- Pate, A. M., Kelling, G. L., & Brown, C. (1975). A response to “what happened to patrol operations in Kansas City?” *Journal of Criminal Justice*, 3(4), 299–320. [https://doi.org/10.1016/0047-2352\(75\)90035-5](https://doi.org/10.1016/0047-2352(75)90035-5)
- Rief, R., & Huff, J. (2023). Revisiting the influence of police response time: Examining the effects of response time on arrest and how it varies by call type. *Journal of Criminal Justice*, 84(84), 1–11. <https://doi.org/10.1016/j.jcrimjus.2022.102025>
- Risman, B. J. (1980). The Kansas City Preventive Patrol Experiment: A Continuing Debate. *Evaluation Review*, 4(6), 802–808. <https://doi.org/10.1177/0193841X8000400605>
- Sacks, S. R. (2000). Optimal Spatial Deployment of Police Patrol Cars. *Social Science Computer Review*, 18(1), 40–55. <https://doi.org/10.1177/089443930001800103>
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997, January). *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress* (tech. rep.). NIJ. Washington, DC.
- Sherman, L. W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime “hot spots”: A randomized, controlled trial. *Justice Quarterly*, 12(4), 625–648. <https://doi.org/10.1080/07418829500096221>
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Tarr, D. P. (1978). Analysis of response delays and arrest rates. *Journal of Police Science & Administration*, 6(4), 429–451.
- Taylor, P. E., & Huxley, S. J. (1989). A Break from Tradition for the San Francisco Police: Patrol Officer Scheduling Using an Optimization-Based Decision Support System. *Interfaces*, 19(1), 4–24. <https://doi.org/10.1287/inte.19.1.4>
- van Barneveld, T., Bhulai, S., & van der Mei, R. D. (2016). The effect of ambulance relocations on the performance of ambulance service providers. *European Journal of Operational Research*, 252(1), 257–269. <https://doi.org/10.1016/j.ejor.2015.12.022>
- van Barneveld, T., Jagtenberg, C., Bhulai, S., & van der Mei, R. (2018). Real-time ambulance relocation: Assessing real-time redeployment strategies for ambulance relocation.

- Socio-Economic Planning Sciences*, 62, 129–142. <https://doi.org/10.1016/j.seps.2017.11.001>
- Verlaan, T., & Ruiter, S. (2023b). Predictors of police response time: a scoping review. *Crime Science*, 12(1), 19. <https://doi.org/10.1186/s40163-023-00194-3>
- Weisburd, D., Braga, A. A., Groff, E. R., & Wooditch, A. (2017). Can Hot Spots Policing Reduce Crime in Urban Areas? An Agent-based Simulation. *Criminology*, 55(1), 137–173. <https://doi.org/10.1111/1745-9125.12131>
- Weisburd, D., Wilson, D. B., Petersen, K., & Telep, C. W. (2023). Does police patrol in large areas prevent crime? Revisiting the Kansas City Preventive Patrol Experiment. *Criminology & Public Policy*, 22(3), 543–560. <https://doi.org/10.1111/1745-9133.12623>
- Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, 103(2), 280–293. https://doi.org/10.1162/rest{_}a{_}00889
- Wilensky, U. 1999. Netlogo. <http://ccl.northwestern.edu/netlogo/>.
- Wooditch, A. (2023). The Benefits of Patrol Officers Using Unallocated Time for Everyday Crime Prevention. *Journal of Quantitative Criminology*, 39(1), 161–185. <https://doi.org/10.1007/S10940-021-09527-4>
- Zhang, Y., & Brown, D. E. (2012). Police patrol district design using agent-based simulation and GIS. *2012 IEEE International Conference on Intelligence and Security Informatics*, 54–59. <https://doi.org/10.1109/ISI.2012.6284091>
- Zimring, F. E. (1978). Policy Experiments in General Deterrence: 1970-1975. In *Deterrence and incapacitation estimating the effects of criminal sanctions on crime rates* (pp. 140–186). The National Academies Press. <https://doi.org/10.17226/18561-2>

B

B

APPENDIX: *ABM INVESTIGATION*

B.1 INPUT DATA

Table B.1: GMS calls-for-service data

Variable name	Description
Incident ID	Unique identifier of each incident
Dispatch ID	Unique identifier of each dispatch
Call_sign	Unique identifier of police response unit vehicle
Incident_start_timestamp	Timestamp of when the call was received.
Incident_closure_type	Captures response to CFS, e.g., fake call or handled over phone
Incident_priority_level	Priority level of call – range: 5 Levels (1,2,3,4,5)
T_x_coord_aanrij	X coordinate of the location
T_y_coord_aanrij	Y coordinate of the location

Table B.2: GMS dispatch data

Variable name	Description
Incident ID	Unique identifier of each incident
Dispatch ID	Unique identifier of each dispatch
Timestamp	Timestamp of status logging (Format: yyyy-mm-dd hh:mm:ss)
Status	Status logged by unit (4 Levels: dispatched, on-way, arrived, cleared)

Table B.3: OSM ‘drive’ street network data within District Oost-Utrecht

Variable name	Description
Road_ID	Unique identifier for each road segment within the network
Average_Speed	Average travel speed for cars on the road segment (km/h)
Directionality	Indicates whether the road is one-way or two-way
Length	True length of the road segment (meters)

B.2 DATA PREPROCESSING

Table B.4: Pre-processing steps and their impacts on incidents and dispatches

Filter		Incidents	Dispatches	Rationale for filtering step
District Oost-Utrecht 2019		72,649 (100%)	113,959 (100%)	The dataset encompasses four front line teams, both rural and urban
Reactive calls		56,118 (77%)	90,695 (80%)	This ABM focuses exclusively on reactive policing, necessitating the exclusion of proactive or preventive calls
Received dispatch		38,457 (53%)	67,983 (60%)	Certain CFS are false alarms or are resolved without police dispatch
Received police response unit		34,593 (48%)	50,694 (44%)	Some CFS are handled by nonstandard units, such as Military Police, which fall outside the scope of this study
Summer-time		34,573 (48%)	50,664 (44%)	The transition from summer to wintertime results in clock adjustments that create timestamp inconsistencies; these cases were removed to prevent data integrity issues
<i>cleared</i> missing		33,664 (46%)	45,426 (40%)	Incidents lacking a <i>cleared</i> timestamp were excluded, as the duration of police presence at the scene could not be determined
<i>dispatched</i> to <i>cleared</i> > 5 min		32,882 (45%)	43,687 (38%)	Calls with a very short duration often lack complete timestamps and are likely false alarms, resolved before police arrival, or otherwise erroneous
<i>dispatched</i> to <i>cleared</i> < 8 h		32,880 (45%)	43,682 (38%)	A police shift lasts eight hours, this step removes only a minimal number of cases (five dispatches and two incidents)
<i>dispatched</i> to <i>arrived</i> > 1 min		32,099 (44%)	41,990 (37%)	These cases are likely instances of cars already at the scene or other unrepresentative events
<i>arrived</i> to <i>cleared</i> > 1 min		30,380 (42%)	38,893 (34%)	These cases are likely the result of data entry errors or are otherwise unrepresentative of typical police responses

B.3 HISTOGRAM OF PRIORITY 1 CALL VOLUME PER SHIFT PER FRONT LINE TEAM

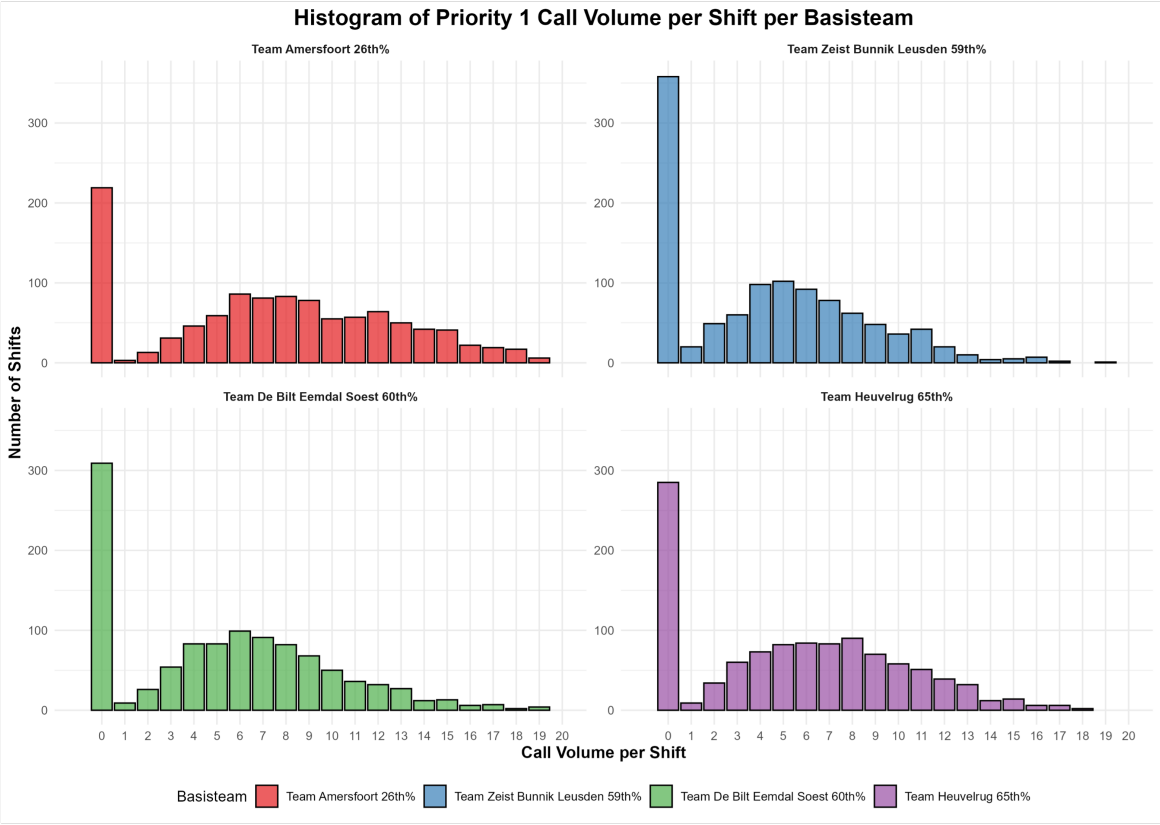


Figure B.1: Histogram of Priority 1 Call Volume per Shift per front line team

B.4 HISTOGRAM OF POLICE CAPACITY PER SHIFT PER FRONT LINE TEAM

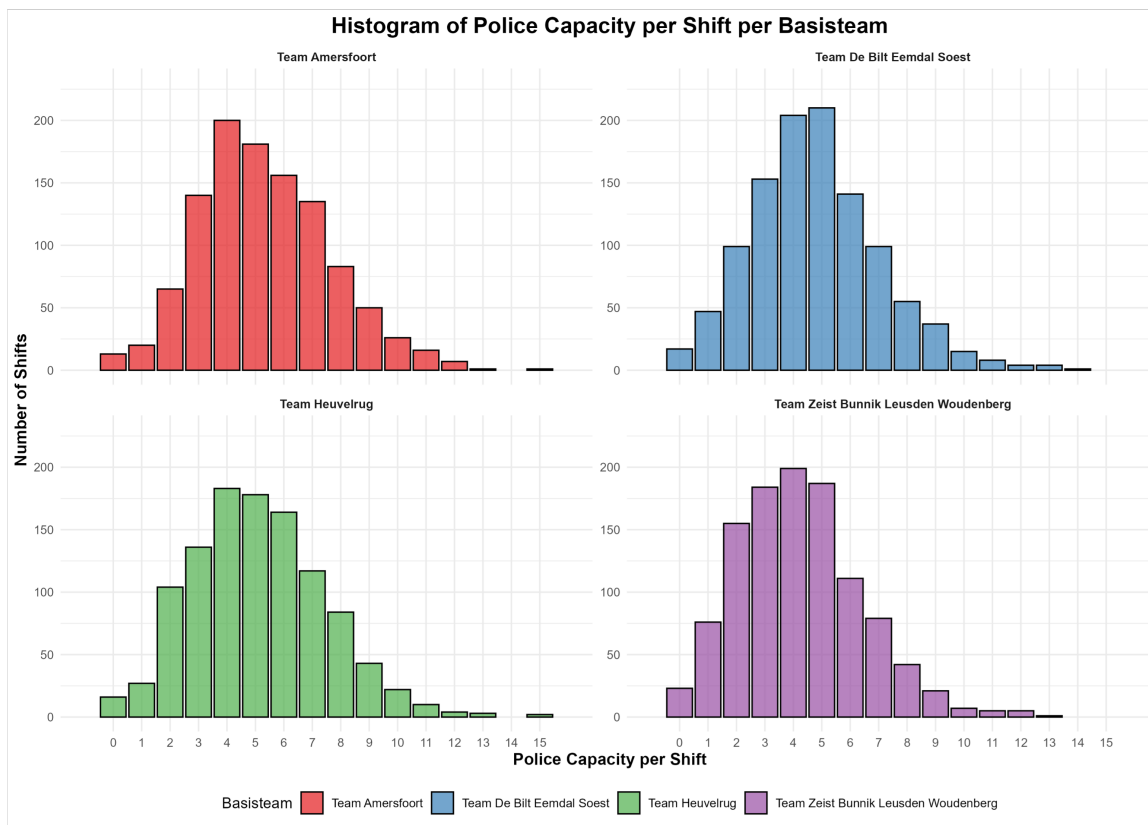


Figure B.2: Histogram of Police Capacity per Shift per front line team (only including police units active within the shift)

III

MAPPING REACTIVE DEMAND AND SUPPLY

5

ON THE USE OF INFERENTIAL STATISTICS ON ADMINISTRATIVE POLICE DATA

5

Do not make exact, what you are not willing to make precise.

This chapter examines the widespread use of inferential statistics – such as p-values and confidence intervals – on administrative police data. While inferential statistics have enabled researchers to make stronger claims about generalisability, their application to police data, which often comprise full population datasets rather than probability samples, is conceptually problematic. Drawing on a review of recent articles in leading criminology journals, we show that most studies employing inferential statistics on police data do so without explicitly defining the population to which they generalise, leading to subtle but significant issues. These include wrongful suggestions of generalisability, undervaluing of observed effects, and obscuring of findings by decontextualising results. The chapter concludes by proposing remedies: prioritising descriptive statistics when working with population data, requiring explicit statement of intended populations when inferential statistics are used, and encouraging open debate on the appropriate use of normative judgments in crime and policing research. Our aim is not to reject inferential statistics outright, but to advocate for more transparent, committed, and contextually appropriate use when analysing police administrative data.

5.1 INTRODUCTION

This chapter is written for academics in the field of crime and policing studies who work with administrative police data, are familiar with and employ p-values/confidence intervals in their research, but for who *inferential statistics* in and of itself is not an intrinsic or research interest. For the sake of accessibility and digestibility, we have refrained from starting this chapter with a long and technical exposé of the statistical concepts touched on in this chapter¹; we trust our reader to have sufficient grasp of the concepts discussed to follow the reasoning as presented. Instead, we suffice with a more general introduction of inferential statistics.

Before the inception of inferential statistics, researchers were designated to imprecise statements about the generalisability of their findings, for example when Laplace wrote (commenting on the *human sex ratio*) ‘It is natural to conclude that these possibilities are very nearly in the same ratio’ (Laplace, 1781). With the formalisation of measures of statistical significance, i.e., p-values and confidence intervals, researchers were empowered to make much bolder claims, e.g. ‘There is a 1% chance [i.e., $p = 0.01$] that I would observe an effect as extreme (or more extreme) than effect X in my sample, assuming that the actual effect in the population is 0’. In essence, inferential statistics provided researchers with tools to transform the vague language of generalisability into an exact science; immensely powerful tools which since have been wielded in great numbers and with great effect, but which are also not without negative consequences if misused.

With this chapter we hope to start a discussion about the use of *inferential statistics* on administrative police data. We will be arguing for a general caution against the use of inferential statistics when working with this type of data. The epigraph to this chapter is ‘Do not make exact, what you are not willing to make precise’. It signifies that our quarrel is not with p-values, confidence intervals or inferential statistics itself, but rather, with its uncommitted application: application without explicitly stating the population to which is generalised. In the event that we fail in convincing readers to refrain from applying inferential statistics to the datasets we describe, we at least hope to encourage them to be explicit in defining their perceived population. The conclusions drawn from, and indeed actions informed by, research on police data are too important and impactful to be left to vague implications.

5.2 UNDERSTANDING POLICE DATA

Our focus in this chapter is administrative police data (hereafter referred to as *police data*), which we understand to be all data collected by police that carry any information about crimes, incidents and police-citizen interaction more broadly. Two key examples of such data would be police-recorded crime data and emergency calls for service records (e.g., Computer Aided Dispatch data). Such police data are widely used in criminological research, largely due to their accessibility under open data licences, their spatial and temporal specificity, and for the very practical reason that they free researchers from primary data collection and the resource constraints that come with that. We recognise that these data often: involve non-criminal issues, do not represent all actual crime (i.e., ‘the dark figure of crime’), and

¹We understand this might mean some readers are left wanting for an in-depth statement on our statistical position. To this end, we have published a more technical elaboration on the open GitHub page https://github.com/timverlaan/Police_administrative_data.

have many issues regarding (in)accuracy and (in)completeness. However, all these important and valid critiques – together with other debates around inferential statistics, e.g., p-hacking – are outside of the scope of this chapter.

Working with police data requires an understanding of the origin and nature of the data. By origin, we mean the data generating process that resulted in the dataset at hand. By nature, we mean the characteristics of the data that determine the scope and the weight of the evidence it is able to support; these being a direct consequence of the origin of the data.

So, let us have a look at the origin and nature of police data. Most commonly, police data are made available to researchers as if appearing from a black box, either following a formal application or through download from an open access data portal. The data are gathered by the police apparatus, which is aided in doing so by the citizenry. While police officers do encounter crime while out on patrol, the majority of the crime they record is reported to them by the public. Crime reported to the police – for example through a telephone call – is documented, entered in a database and saved for later consultation or analysis². Now, how can we conceptualise this administrative data-generating process so that we can understand it in scientific terms?

We propose viewing the data-generation process of police data as a longitudinal census survey that is continually distributed to the public. This distribution is not carried out in the active, but rather in the passive sense: the distribution is facilitated by the public's understanding of its ever-present existence. At each time interval, respondents – comprised of the entire population of a geographically demarcated area (i.e., the police jurisdiction) – are queried as to whether a crime and/or incident is taking place and, if so, to report it (and its characteristics) to the police.

In this conceptualisation, all crime that is *not* reported to the police can be viewed as *nonresponse*. No *new crimes* or incidents can be collected: all datapoints that could potentially be added to the dataset at a later time – e.g., through some hypothetical independent data collection effort by researchers – have initially been queried by the original data generating process and should thus be understood as *late response* rather than *new response*. Consequently, in this simplified conceptualisation, the difference between *reported crime* and *actual crime* can thus be understood as resulting entirely from *participation bias*.

5.2.1 APPARENT POPULATIONS

What follows from this conceptualisation, is that the data gathered by the police are all the data there are; no additional data could be collected, even in principle. Datasets with properties such as these are what Berk et al. (1995, p. 421) term 'apparent populations', and they present two main conceptual problems for statistical inference: 'First, the data are not generated by probability sampling or random assignment. Second, and far more important for this paper, an apparent population is the result of a data generation mechanism that produces only a single batch of data.' (Berk et al., 1995, p. 422).

When confronted with *apparent population* police data, scholars can have either of two responses, which we have termed the *superpopulation approach* and the *population approach*. Either imagine a *superpopulation* from which the data is assumed to represent a

²We acknowledge that reality presents more complications: recording of crime and incidents by the police is an imperfect process with many inaccuracies and mistakes. However, as these are not at odds with the conceptualisation or the points made in this contribution, we shall ignore these here.

hypothetical random sample and apply conventional inferential statistics, or treat the data as a true population, and thus refrain from employing inferential statistics and rely solely on descriptive statistics. Both these responses have epistemological underpinnings which are implicitly adhered to, however, we shall leave discussion of these to Berk et al. (1995). Instead, we will here investigate the practical meaning of these responses when working with police data and attempt to trace their implications.

SUPERPOPULATION APPROACH

We differentiate between three categories of superpopulations: the geographical, the temporal and the – for lack of a better term – latent superpopulation. In order to illuminate these categories, we imagine having obtained all police-recorded crime data from the Dallas Police Department (Texas, USA) for the year 2023.

The geographical superpopulations we may imagine for this data include, in decreasing levels of credibility, police-recorded crime data from: all Texan cities; all cities in the USA; the whole of the USA; the whole World; or even further: the known Universe; all possible Universes.

5

Temporal superpopulations are less intuitive as a part of the temporal scale (i.e., the future) is, by definition, impossible to be sampled in the present. Nevertheless, the general idea is that the future, past and present are imagined as a superpopulation of which the past represents a random sample. In our example, and again in decreasing levels of credibility, the police data from Dallas 2023 would be taken as a random sample of all police-recorded crime data from: Dallas 2022–2024, Dallas 2023–2030, or Dallas in all the years to come.

Lastly, the latent superpopulation we may imagine for this data are all crimes that occurred in Dallas 2023 – including those crimes not recorded by the police. In other words, in our example, the latent superpopulation views the police-recorded crime as a random sample of all police-recorded crimes plus all non-police-recorded crimes.

Whether any – or a combination – of these superpopulations are appropriate is entirely down to the context of the situation. Nevertheless, we feel comfortable in making two general comments on these categories of superpopulations.

First, not only can a (careful) hierarchy of credibility be made *within* categories, but also *among* categories. We would like to express this carefully by stating that, in general, city A in year X can be expected to be more like, and therefore to constitute a better sample upon which to base inference, city A in year X+1, than city B in year X. In other words, *a priori*, we place more trust in the appropriateness of a generalisation solely along the temporal dimension rather than in a generalisation – of the same magnitude – solely along the geographic dimension. Moreover, in talking to colleagues during the formation of this argument, we have understood the temporal dimension to be their main inherent axis of generalisation – although it must be added that none had explicitly formulated this in their work.

Second, but building on the previous point, neither we nor our colleagues have witnessed – in the literature presented here nor in our daily work within the field – a study on police data employing inferential statistics (p-values/confidence intervals) which explicitly claims to generalise to any – or a combination – of these superpopulations. This does not mean that such cases do not exist but, rather, it indicates that such superpopulations are not routinely invoked explicitly within our field.

POPULATION APPROACH

The population approach takes a very different point of departure, reasoned by Berk et al. (1995):

... the data may be treated as a (fixed) population. ... statistical inference is moot. Descriptive summaries of the population can be used. In effect, we assume the data could not have been different because of sampling error. And the data from the apparent population become data from a true population. (Berk et al., 1995, p. 425)

In other words, any claim of generalisation beyond the scope of the population itself through the application of inferential statistics is inappropriate and without meaning, and is therefore forfeited. Staying with our example, any conclusions drawn from the Dallas 2023 police data generalise only to Dallas, 2023.

5.2.2 APPROACH OF THE FIELD

Within the field, we observe an overwhelming preference, albeit implicitly, for the first response: using inferential statistics on police data for a given year and study region, as if these represent a random sample from a clearly defined population. In order to support this claim, we shall now present some empirical findings from a mini scoping review. Subsequently, we will discuss three negative consequences that we observed in reviewing the literature linked to this response.

5.3 EVIDENCE FROM THE FIELD

In order to investigate the extent of the (by us) purported problem, we reviewed the most recent issues (as of 26 July 2022) of eight journals: *Annual Review of Criminology*, *Criminology*, *Crime Science*, *Journal of Quantitative Criminology*, *Journal of Experimental Criminology*, *Police Quarterly*, *Crime & Delinquency*, and *Police Practice and Research*. We selected these journals as they are relevant to the field of crime and policing research and are widely read and cited. However, we want to stress that our findings are limited and should not be generalised to the entire field of crime and policing research. We simply aim to demonstrate that the issues identified have occurred in recent, widely-read and widely-cited crime and policing journals.

We inspected the latest issue of each journal for any articles employing inferential statistics on administrative police data. Of these eight journals, the *Annual Review of Criminology* and *Criminology* did not publish any articles in the examined issue that met this criterion. The examined issue of the remaining six journals yielded a total of 16 articles which did meet this criterion, they are presented in Table 1.

We analysed these articles for their use of police data and its origin and nature. Moreover, we checked if they stated any superpopulations to which was being generalised. Additionally, we coded the (possible) negative consequences of the application of inferential statistics. The one negative consequence we purported *a priori* was the *wrongful suggestion of generalisability*, the two other negative consequences (*undervaluing* and *obscuring*) have been identified and coded inductively. Coding was conducted by both authors, but final decisions were made by the first author. All coding is found on the open GitHub repository: https://github.com/timverlaan/Police_administrative_data.

After finalizing the coding, we concluded that all but one (i.e., Salerno-Ferraro & Jung, 2022) of the 16 retrieved articles do not meet the key requirement for statistical inference: a random sample from a clearly defined population. Furthermore, none of the retrieved articles explicitly state the (super)population to which findings could be generalised.³ We shall now present descriptions of the negative consequences we have identified in the literature.

Table 5.1: Included and examined articles

Title	Author(s) & Year	Journal & Volume, Issue
Re-Opening Dallas: A Short-Term Evaluation of COVID-19 Regulations and Crime	Riddell et al. (2022)	<i>Crime & Delinquency</i> 68(8)
Temporal Variations in Calls for Police Service During COVID-19: Evidence from China	Dai et al. (2022)	<i>Crime & Delinquency</i> 68(8)
Skimmed at the Pump: How COVID-19 Lockdowns Increased Gas Pump Skimming	Belshaw et al. (2022)	<i>Crime & Delinquency</i> 68(8)
A Case Study of Family Violence During COVID-19 in San Antonio	Leal et al. (2022)	<i>Crime & Delinquency</i> 68(8)
Different places, different problems: profiles of crime and disorder at residential parcels	O'Brien, Ristea et al. (2022)	<i>Crime Science</i> 11
The Effects of the Summer All Out Foot Patrol Initiative in New York City: a difference-in-differences approach	Bilach et al. (2022)	<i>Journal of Experimental Criminology</i> 18(2)
A randomised controlled trial of the impact of body-worn camera activation on the outcomes of individual incidents	Huff et al. (2022)	<i>Journal of Experimental Criminology</i> 18(2)
Pushed out of the education system: using a natural experiment to evaluate consequences for boys	Larsen et al. (2022)	<i>Journal of Experimental Criminology</i> 18(2)
How and Why is Crime More Concentrated in Some Neighborhoods than Others? A New Dimension to Community Crime	O'Brien, Ciomek and Tucker (2022)	<i>Journal of Quantitative Criminology</i> 38(2)
Measuring the Impacts of Everyday Police Proactive Activities: Tackling the Endogeneity Problem	Wu et al. (2022)	<i>Journal of Quantitative Criminology</i> 38(2)
Gang Graffiti, Group Process, and Gang Violence	Hughes et al. (2022)	<i>Journal of Quantitative Criminology</i> 38(2)
The Effect of a Death-in-Police-Custody Incident on Community Reliance on the Police	Moyer (2022)	<i>Journal of Quantitative Criminology</i> 38(2)
Evaluating the 4th Circuit's decision to limit officer use of Tasers: a descriptive and time-series approach	Boehme et al. (2022)	<i>Police Practice and Research</i> 23(5)
To charge or not to charge? Police decisions in Canadian sexual assault cases and the relevance of rape myths	Salerno-Ferraro and Jung (2022)	<i>Police Practice and Research</i> 23(5)
Operation provide: a multi-agency response to increasing police engagement in cases of intimate partner violence during the COVID-19 pandemic	Halford and Smith (2022)	<i>Police Practice and Research</i> 23(5)
Unpacking Sequential Actions Within Use of Force Incidents	Tillyer (2022)	<i>Police Quarterly</i> 25(2)

³Here, we would note that some of studies might conceptualise populations and use inferential statistics for the purposes of evaluating an experimental manipulation, rather than solely for generalisation across geographic or temporal superpopulations. The former example might include quasi-natural experiments, such as COVID-19 lockdowns, in which pre and post-lockdown periods are conceptualised as two random samples which (might) have been drawn from different populations. However, we do not observe a clear distinction or explanation of this usage in the literature examined here.

5.4 THE CONSEQUENCES

First, a disclaimer: here we demonstrate the negative consequences we witnessed through specific examples from the examined literature. These examples do not function to reflect the overall severity of the issues, they simply serve to illustrate the different negative consequences we identified in a clear and concise manner. In other words, they are selected for clarification purposes, not for generalisation purposes.

5.4.1 WRONGFUL SUGGESTIONS OF GENERALISABILITY OF THE OBSERVED DIFFERENCES/EFFECTS

First, unwarranted use of inferential statistics leads to a wrongful suggestion of generalisability of the observed differences/effects. There are two ways in which studies make suggestions of generalisability: explicitly and implicitly. Explicit suggestions of generalisability can then also be made in two ways, bold and subtle. A bold and explicit suggestion of generalisability would constitute, for example ‘Our findings from city X generalise to cities Y and Z’. As previously stated, we never witness such explicit and bold suggestions of generalisability. We do witness explicit but subtle suggestions of generalisability in the literature, as illustrated in the following quote: ‘Findings from this study provide important insight to the potential negative (or positive) consequences of the Court’s decision. This may help inform other jurisdictions regarding whether to restrict officer use of CEDs.’ (Boehme et al., 2022, p. 524) And in quotes supplemented with various degrees of caveating:

Last, the findings here hold valuable implications for problem properties policies and interventions . . . In sum, it is important not to overgeneralise the precise results of this study – the specific types of parcels and neighborhoods identified – but to concentrate on their overarching story. (O’Brien, Ristea et al., 2022, p. 10)

Our findings are significant for frontline police officers and the government, allowing them to learn from the past and better prepare for future exceptional events . . . Finally, one limitation of this study is the generalisability of its results, given that we only focused on a single city. (Dai et al., 2022, pp. 1200–1201)

Judging the legitimacy of the above suggestions of generalisability – whether they should be considered *wrongful* suggestions of generalisability or not – is challenging. On the one hand, aspiring to generalisations is at the heart of all (social) science and in many cases the assumption that ‘observations on a phenomenon in one place & time inform on that phenomenon in another place & time’ is a reasonable one to defend, and aspire to, and one we encourage. On the other hand, if these subtle suggestions of generalisability are made while discussing observed differences/effects which are subjected to exact calculations – in the form of p-values or confidence intervals – of precisely how likely they are to inform a completely unspecified other place and time, as the above-mentioned studies do, then – to us – the waters turn murky.

We consider all use of inferential statistics to make at least an implicit suggestion of generalisability: the moment inferential statistics are employed on a dataset, this dataset is brought into a paradigm in which it is taken to be a sample of a (perceived or actual) population, regardless of whether this population is made explicit. Within this paradigm, the

dataset is thus taken to make some suggestion of generalisability, even if no such claim is made. An example of this is found in the paper by Leal et al. (2022, p. 1174), in which we identify none of the negative consequences except for the (very) subtle implicit generalisation of including – what we believe to be, although it is not explicitly stated as such – a confidence interval around the Loess curve.

5.4.2 UNDERVALUING OF OBSERVED DIFFERENCES/EFFECTS

Second, the use of inferential statistics on findings that do not require validation by the sigil of ‘statistical significance’ leads to undervaluing of observed differences/effects. By *undervaluing* we mean the dismissal of observed differences/effects as *non-existent*, *non-significant*, or *not worthy of discussing*. Within the examined literature, we witnessed *undervaluing* in three forms: (1) research that denies the existence of observed difference/effect (e.g., ‘no association is observed’) even though difference/effect of a non-zero size is observed (e.g., Huff et al., 2022; Hughes et al., 2022; Larsen et al., 2022); (2) research that discusses but subsequently disqualifies observed differences/effects because they are *not significant* (e.g., Bilach et al., 2022; Dai et al., 2022; Riddell et al., 2022); (3) research that does not discuss any observed differences/effects that are qualified as *non-significant* (e.g., Boehme et al., 2022; Huff et al., 2022; Moyer, 2022). Undervaluing takes place because the *tools of inferential statistics* find the observed differences/effects not pronounced enough to be considered sufficiently likely (compared to a predetermined α -level) to be observed in some unstated superpopulation. We would suggest that the findings resulting from the analysis of large administrative police datasets are always valuable within their own scope of analysis (i.e. the jurisdiction and time period). They may – depending on the context – also present grounds for generalisations beyond their own scope of analysis, but, if not, this should never be taken as grounds to dismiss the observed differences/effects within their own scope of analysis.

5.4.3 OBSCURING OF THE OBSERVED DIFFERENCES/EFFECTS

As a last negative consequence, we argue that the *inferential statistics* frame also leads to less precise and less informative conclusions. The reason for this is that this frame inherently assumes a *scope of interest* (i.e., the population) different from the *scope of analysis* (i.e., the sample). Within this frame, in order to facilitate translation of the findings from the *scope of analysis* to the *scope of interest*, findings are presented in percentages, coefficients and indices. These percentages, coefficients and indices carry more information as they – and opposed to the observed counts – are assumed to be the same for both the *scope of interest* and the *scope of analysis*. In other words, percentages, coefficients and indices trade a loss in contextual clarity for gain in transferability/generalisability.

However, for studies based on population data, the *scope of analysis* is the *scope of interest*. Thus, in this case, nothing is to be gained in transferability/generalisability and percentages, coefficients and indices serve only to obscure the actual observed differences or effects. To illustrate, take the overall conclusion by Larsen et al. (2022):

Focusing on potential criminogenic effects of the reform, our empirical results suggest that the reform led to a 2 percentage points higher cumulative rate of criminal charges 9 months after finishing ninth grade (from a baseline of 7.2% in the pre-reform group, corresponding to a 28% relative increase in criminal charges). However, we find no reform effect on cumulative offending rates after 12 months. Therefore, our results suggest that educational enrolment itself may decrease criminal behaviour among low-achieving boys in the very short run (in the middle of the school year) but not in the slightly longer run (towards the end of the school year). In conjunction with a series of robustness checks, which support our main results, we argue that there is a short-term impact of the 2015 reform on criminal charges – although statistically small. (Larsen et al., 2022, p. 429)

Grasping the actual magnitude of the reform's effect directly after reading might be challenging even for trained experts. For a policy maker skimming research articles and reports, this might be near impossible.

Behind the phrases '2 percentage points higher cumulative rate', 'baseline of 7.2%' and '28% relative increase' are of course raw counts of how many Danish teenage boys entered into crime as a result of the intervention in the educational system. As virtually all boys who are affected by this intervention are accounted for in the study, we are here dealing with population data and the scope of interest thus is the scope of analysis. Hence, we argue that such conclusions could be more precise and informative from a descriptive frame, for example by presenting the raw counts ⁴ of how many boys enter into crime earlier than could be expected based on the previous year, as a result of the educational reform.

5

5.5 HYPOTHESISED DRIVERS OF MISUSE

We hypothesise three drivers of misuse. First, as other authors – focusing on development, educational and sports research (Ballinger, 2011, pp. 16–17; B. G. Gibbs et al., 2015, pp. 119–120; B. G. Gibbs et al., 2017, p. 218) – also have, we identify the main driver of misuse of inferential statistics to be the uncritical application of statistical techniques. We would like to add that this uncritical application may be facilitated by easy-to-use statistical software (such as SPSS) that – by default – assume data to be *sample data* and automatically present p-values with all output.

Second, we believe part of the misuse can be attributed to the ever-increasing need for generalisability and focus on impact of scientific research – mainly promoted by academic institutions, journals, and reviewers but derived from demands made by governmental policymaking and research funding organisations. We ourselves, as well as many of the people we have spoken to, recognise the general call to include p-values in our research and comment on the 'generalisability' of the findings. While we encourage researchers to comment on how their findings might inform other places and times, this should be done without false signals (and without meaningless precision) such as unwarranted p-values.

Third, we hypothesise that inferential statistics is often wrongfully used as a proxy for normative judgements in the absence of normative guidelines. For example, let's say

⁴Which, after some data wrangling, turn out to be: 6 boys after 3 months, 19 boys after 6 months, 27 boys after 9 months and a remaining 5 boys after 12 months.

City X experienced a 2% drop in crime-rate; is this difference big? Little? Should it be considered within natural variation or is it anomalous and to be investigated further? Instead of confronting these difficult questions, research often – though maybe not purposefully – sidesteps by resorting to inferential statistics; letting statistical significance stand in for normative significance: ‘City X experienced a significant drop in crime-rate’.

To further problematise this second hypothesised driver of misuse, we propose that this need for normative judgements is itself a consequence of the *inferential frame*; the cure is causing the ailment. As we have argued, the inferential frame – using percentages, coefficients and indices – decontextualises findings in order to be able to translate them to the context of the *scope of interest*. In the absence of the original context and the intuitive interpretation – and in the absence of a clearly defined new context – scholars struggle to (re)interpret their findings. The common reflex to this conundrum has been to employ decontextualised interpretations of standardised differences/effects, such as those present in Cohen’s (2013 seminal work *Statistical Power Analysis for the Behavioral Sciences*.

5

5.6 PROPOSED REMEDIES FOR MISUSE

In order to reduce the – according to us – unwarranted use of inferential statistics, and overcome the consequent negative consequences described in this article, we propose three remedies.

First, we join, amongst others, Goodchild (2009), Shalev (2007), and Ballinger (2011) in calling for a general caution of inferential approaches, and a consequent revaluing of descriptive statistics, when studying social phenomena through administrative population datasets. Quoting Goodchild (2009):

‘Why are we not content simply to describe specific parts of the heterogenous world that we see around us, using the rigorous methods of science, and why do we insist instead that inferences be made about some poorly conceived and nonexistent hypothetical world? (Goodchild, 2009, p. 415)’

Second, while we are not opposed to sophisticated statistical techniques, we do believe that researchers should be committed in their application; we cannot have our cake and eat it too. Therefore, we echo the call of B. G. Gibbs et al. (2017, p. 218) for academic journals to ‘... adopt or revise standards to make the interpretation of inferential statistics and p-values, such as defining the intended inferred population, a required part of reporting in analytic methods sections’.

Third, we call for the opening up of a debate around the need for making normative judgements, and the appropriate ways of doing so. We propose to use the earlier mentioned GitHub page as the (initial) place for this discussion. We would like our opening contribution to this debate to be: The field of crime and policing research is as applied and contextual as it comes; if we are to make normative judgements on our findings, we should never do so devoid of context.

5.7 CONCLUDING REMARKS

We would like to reiterate that ours is not a plea against the use of inferential statistics, it is a plea against unwarranted and/or uncommitted statistical inference. The core of our argument is that: if research is unable to provide a precise and explicit conceptualisation of the perceived (or actual) (super)population, we as a field should not accept exact figures (p-values/confidence intervals) regarding generalisability of findings.

5.8 ACKNOWLEDGEMENTS

This research was carried out within the What Works in Policing research programme of the NSCR (Netherlands Institute for the Study of Crime and Law Enforcement) and funded by the Netherlands Police. We thank the anonymous reviewers for their useful comments. We thank Henk Elffers, Wouter Steenbeek, Jose Pina-Sánchez, Daniel Birks, Arjan Blokland, and Edward Kleemans for their discussions, disagreements, encouragements and feedback on earlier versions of the text.

REFERENCES

- Ballinger, C. (2011). Why Inferential Statistics are Inappropriate for Development Studies and How the Same Data Can be Better Used. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1775002>
- Belshaw, S. H., Nodeland, B., Pratt, T., & Leeper Piquero, N. (2022). Skimmed at the Pump: How COVID-19 Lockdowns Increased Gas Pump Skimming. *Crime & Delinquency*, 68(8), 1207–1222. <https://doi.org/10.1177/00111287221081028>
- Berk, R. A., Western, B., & Weiss, R. E. (1995). Statistical Inference for Apparent Populations. *Sociological Methodology*, 25, 421. <https://doi.org/10.2307/271073>
- Bilach, T. J., Roche, S. P., & Wawro, G. J. (2022). The effects of the Summer All Out Foot Patrol Initiative in New York City: a difference-in-differences approach. *Journal of Experimental Criminology*, 18(2), 209–244. <https://doi.org/10.1007/s11292-020-09445-8>
- Boehme, H. M., Martin, A., & Kaminski, R. J. (2022). Evaluating the 4th Circuit's decision to limit officer use of Tasers: a descriptive and time-series approach. *Police Practice and Research*, 23(5), 523–538. <https://doi.org/10.1080/15614263.2021.1982713>
- Cohen, J. (2013, May). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Dai, M., Xia, Y., & Han, R. (2022). Temporal Variations in Calls for Police Service During COVID-19: Evidence From China. *Crime & Delinquency*, 68(8), 1183–1206. <https://doi.org/10.1177/00111287221074942>
- Gibbs, B. G., Shafer, K., & Dufur, M. J. (2015). Why infer? The use and misuse of population data in sport research. *International Review for the Sociology of Sport*, 50(1), 115–121. <https://doi.org/10.1177/1012690212469019>
- Gibbs, B. G., Shafer, K., & Miles, A. (2017). Inferential statistics and the use of administrative data in US educational research. *International Journal of Research & Method in Education*, 40(2), 214–220. <https://doi.org/10.1080/1743727X.2015.1113249>
- Goodchild, M. F. (2009). What Problem? Spatial Autocorrelation and Geographic Information Science. *Geographical Analysis*, 41(4), 411–417. <https://doi.org/10.1111/j.1538-4632.2009.00769.x>
- Halford, E., & Smith, J. (2022). Operation provide: a multi-agency response to increasing police engagement in cases of intimate partner violence during the COVID-19 pandemic. *Police Practice and Research*, 23(5), 600–613. <https://doi.org/10.1080/15614263.2022.2033621>
- Huff, J., Katz, C. M., & Hedberg, E. C. (2022). A randomized controlled trial of the impact of body-worn camera activation on the outcomes of individual incidents. *Journal of Experimental Criminology*, 18(2), 247–272. <https://doi.org/10.1007/s11292-020-09448-5>
- Hughes, L. A., Schaible, L. M., & Kephart, T. (2022). Gang Graffiti, Group Process, and Gang Violence. *Journal of Quantitative Criminology*, 38(2), 365–384. <https://doi.org/10.1007/s10940-021-09507-8>
- Laplace, P. (1781). Mémoire sur les probabilités. *Mémoires de l'Académie Royale des Sciences de Paris*, 1778, 227–332.
- Larsen, K. S., Andersen, L. H., & Larsen, B. Ø. (2022). Pushed out of the education system: using a natural experiment to evaluate consequences for boys. *Journal of*

- Experimental Criminology*, 18(2), 411–438. <https://doi.org/10.1007/s11292-020-09444-9>
- Leal, W. E., Piquero, A. R., Kurland, J., Piquero, N. L., & Gloyd, E. L. (2022). A Case Study of Family Violence During COVID-19 in San Antonio. *Crime & Delinquency*, 68(8), 1161–1182. <https://doi.org/10.1177/001112872111064781>
- Moyer, R. A. (2022). The Effect of a Death-in-Police-Custody Incident on Community Reliance on the Police. *Journal of Quantitative Criminology*, 38(2), 459–482. <https://doi.org/10.1007/s10940-021-09504-x>
- O'Brien, D. T., Ciomek, A., & Tucker, R. (2022). How and Why is Crime More Concentrated in Some Neighborhoods than Others?: A New Dimension to Community Crime. *Journal of Quantitative Criminology*, 38(2), 295–321. <https://doi.org/10.1007/s10940-021-09495-9>
- O'Brien, D. T., Ristea, A., Hangen, F., & Tucker, R. (2022). Different places, different problems: profiles of crime and disorder at residential parcels. *Crime Science*, 11(1), 4. <https://doi.org/10.1186/s40163-022-00165-0>
- Riddell, J. R., Piquero, A. R., Kaukinen, C., Bishopp, S. A., Piquero, N. L., Narvey, C. S., & Iesue, L. (2022). Re-Opening Dallas: A Short-Term Evaluation of COVID-19 Regulations and Crime. *Crime & Delinquency*, 68(8), 1137–1160. <https://doi.org/10.1177/001112872111054718>
- Salerno-Ferraro, A. C., & Jung, S. (2022). To charge or not to charge? Police decisions in Canadian sexual assault cases and the relevance of rape myths. *Police Practice and Research*, 23(5), 539–552. <https://doi.org/10.1080/15614263.2021.2002150>
- Shalev, M. (2007). Limits and Alternatives to Multiple Regression in Comparative Research. In *Capitalisms compared* (pp. 261–308). [https://doi.org/10.1016/S0195-6310\(06\)24006-7](https://doi.org/10.1016/S0195-6310(06)24006-7)
- Tillyer, R. (2022). Unpacking Sequential Actions Within Use of Force Incidents. *Police Quarterly*, 25(2), 178–200. <https://doi.org/10.1177/109861112111049549>
- Wu, X., Koper, C., & Lum, C. (2022). Measuring the Impacts of Everyday Police Proactive Activities: Tackling the Endogeneity Problem. *Journal of Quantitative Criminology*, 38(2), 343–363. <https://doi.org/10.1007/s10940-021-09496-8>

6

OPERATIONALIZING DEPLOYMENT TIME IN POLICE CALLS-FOR-SERVICE

6

Analyses of emergency calls for service data in the United States suggest that around 50% of dispatched police deployment time is spent on crime-related incidents. The remainder of time is spent in a social service capacity: attending well-being checks and resolving disturbances, for instance. These findings have made a considerable contribution to the discourse around public perceptions of the police and the distribution of public funds towards (or away) from law enforcement. Yet, an outstanding issue remains. No investigation has been undertaken into whether findings are robust to the different ways in which ‘time spent’ is operationalised in these studies. Using dispatch data for Amsterdam during 2019, this study compares three operationalisations of ‘time spent’. Additionally, in order to provide some context on the potential mechanisms through which these different operationalisations might yield different results, we report on dispatch numbers per incident category and provide an initial exploration into ‘multi-dispatch’ incident types. We find that general proportional breakdowns are fairly robust to the time measure used. However, for some incident categories (e.g. Health) and incident types (e.g. Shootings), analysed in isolation, the results are not robust to the different operationalisations. We propose that the mechanism explaining this lack of robustness can be traced to the high dispatch numbers for specific incident categories and types, particularly those with an imminent threat to life.

Preregistration: *This study has been preregistered under the title: Scale and composition of emergency reactive police demand in Amsterdam, Netherlands (<https://osf.io/qgww6/>).*

6.1 BACKGROUND

Recent years have seen a number of studies investigate the scale ('how much?') and composition ('what type?') of demand for and supply of police services. In other words: what kinds of incidents do the public request police assistance for, and how much time is consumed by attending these incidents? Findings from the US have consistently demonstrated that a considerable amount of dispatched police deployment time is consumed by incidents which do not involve crime, such as resolving community issues, attending mental health crises and dealing with traffic incidents (Langton, Ruiter & Verlaan, 2023; Lum & Koper, 2021; Ratcliffe, 2021). Amidst calls for a radical reform and defunding of police forces, these studies have raised an important discussion around the demand for and supply of policing services. With a considerable proportion (if not, the majority) of dispatched police deployment time consumed in response to requests for public assistance that do not (directly) involve crime, a rapid reduction in their capacity to meet this demand might have detrimental effects on public safety and well-being.

While findings have made a contribution to the discourse, a pressing issue remains: there has been no scrutiny of how 'police dispatched deployment time' is measured. We consider this issue to be particularly critical given the attention that some research has been given in mainstream media (Lum & Koper, 2021) and the impact that this might have on public sentiment and government policy. Lum et al. (2022), Ratcliffe (2021) and Langton, Ruiter and Verlaan (2023) operationalised dispatched deployment time at the incident level: time measurements were taken from when the first police unit arrived on the scene, to when the last unit cleared the scene. We broadly define this measure as 'first in last out' (FILO). Ratcliffe (2021) provided a weighting to the operationalisation by multiplying 'FILO' by the number officers that attended the incident. As visualised in Fig. 6.1, using a stylised example for a multi-dispatch incident, the operationalisations of Lum et al. (2022) and Langton, Ruiter and Verlaan (2023) would underestimate the total dispatched deployment time, while Ratcliffe (2021) would overestimate it.

However, it is unclear how these different operationalisations might affect the proportional breakdowns of 'time spent.' If the underestimation or overestimation is uniformly distributed across all incident categories, the analytical outcomes will be robust to different operationalisations. As we have access to unit-level dispatch data, we can examine whether this is the case, or if any operationalisations introduce enough bias to compromise robustness. For clarity, we define 'robust' findings as those that consistently answer a research question, regardless of the time measurement used. This means that the level of robustness can vary depending on the specific research question.

With this research gap in mind, in this contribution we systematically compare different operationalisations of 'dispatched deployment time' using both incident and unit-level dispatch data in order to ascertain the impact on proportional breakdowns of time spent. We also report descriptive statistics on minutes consumed and the concentration of time according to incident classifications. This part of the study has been preregistered on an Open Science Framework repository. To better grasp the prevalence of multi-dispatch incidents and their potential impact on time spent measures, we report descriptive statistics on dispatch numbers per incident class and visualise the distribution underlying the top 10 'multi-dispatch' incident types in our data. Lastly, we share insights from our experience of double-coding an existing classification of police incidents (Ratcliffe, 2021).

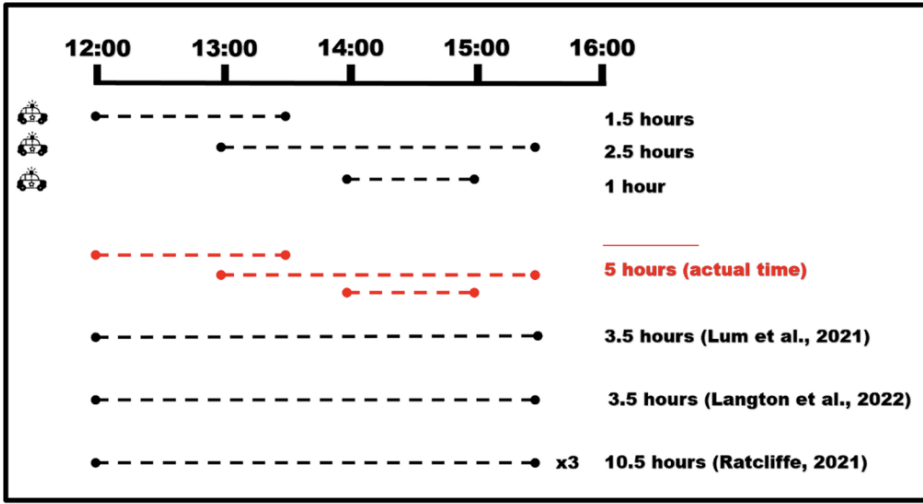


Figure 6.1: Visualizing the different operationalisations of time spent. For the stylised example, we assume all operationalisations use the same start time

6.2 DATA & METHODS

6.2.1 UNIT-LEVEL DISPATCH DATA

The data for this study are unit-level emergency dispatch records for the Amsterdam Police Unit during 2019. For this study, we only include incidents that occurred within the boundaries of the city itself and for which police units were dispatched. We refer readers to the preregistration document for further details on how the raw data were handled in preparation for analysis.

6.2.2 OPERATIONALISATION OF ‘TIME SPENT’

Besides the stated differences between existing studies in how time spent is operationalised, studies can also use different ‘start’ points. Ratcliffe et al. (2021) used *dispatch* as the first timestamp, thereby including *travel time* in the operationalisation of time spent. Lum et al. (2022) exclude *travel time* by using the *on-scene* timestamp as the start time. This is justified on the basis that it excludes potential *idle time* (‘response times vary dramatically across different call types and are actually much longer than laypersons believe because officers do not immediately respond to many calls, given their low priority’, p. 267) and for what we would consider measurement error (‘officers may be assigned to calls while on other calls, or can be reassigned to higher priority calls during a current call assignment’, footnote 14).

In the Netherlands, the dispatch system cannot link units to multiple calls, and the data are therefore not susceptible to that source of measurement error. Moreover, as agency over response prioritisation mainly lies with the dispatch centre, we can expect to see a much lower *idle time*. Nevertheless, the system does register an *on-way* time stamp for when the unit signals that they have begun their response to an incident. Indeed, the mean amount of time that elapses between *dispatch* and *on-way* is short (around 2 min) even

across different incident types (see supplementary materials). We do not make any definitive claim as to what constitutes the ‘most appropriate’ start time. However, we do have the data available to test the impact of these different start times, in addition to the unit or dispatch-level operationalisations discussed above. Here, we would note that in our data 30% of responses for *on-way* are missing. This can occur due to technical malfunctions, time-pressure, forgetting and/or non-compliance.

We report on three operationalisations of time spent, as summarised in Table 6.1.⁰ These three measures cover both existing FILO measures (weighted by the number of units and unweighted) in addition to a unit-level cumulative sum. In all three operationalisations, we define the start point as *dispatch* although we report results for the same three operationalisations using both *on-way* and *on-scene* start points in the supplementary materials (nine in total). As such, we provide a comprehensive overview of the various different operationalisations available.¹

Table 6.1: Description of the different time operationalisations

Time measure	Description
Dispatched to clear (Cumulative)	Cumulative timespan between unit dispatched and unit cleared across all response units for an incident
Dispatched to clear (standard FILO)	Timespan between first unit dispatched and final unit cleared
Dispatched to clear (FILO*UNITS)	Timespan between first unit arrived and final unit cleared, multiplied by the number of units who have been present on scene

6.2.3 CALL CLASSIFICATION

Call handlers classify calls according to a pre-specified list of incident types. We excluded those classifications which were clearly not initiated by the public. This left us with around 110,000 unique incidents during 2019 which required just over 220,000 dispatches. In the interests of consistency and international comparison, we then classified these existing incident types according to the demand classification first used by Ratcliffe (2021) and later Langton, Ruiter and Verlaan (2023), namely: crime, health, traffic, community and quality of life. The classification procedure was undertaken by two raters independently, as outlined in the preregistration document. The degree of inter-rater reliability between the two raters was assessed using the Krippendorff’s alpha statistic (Krippendorff, 2011). The result (0.701, $p<0.05$) exceeded the threshold stated in the preregistration document. As a result, the classification was considered appropriate and reliable enough for the Dutch context. Disagreements were resolved through group discussion together with the third author (author 3 initials). This highlighted a consistent issue, namely that of the demand classification ‘suspicious situation’. The raters could not definitively agree on whether this

⁰Following a comment by an anonymous reviewer, we have changed the word ‘individual’ to ‘all’ in the description of cumulative measures for clarification, compared to the preregistration document.

¹In the preregistration document, we stated that we would only report the various operationalisations with the *on-way* start point due to the concerns raised by Lum et al. (2022). At that stage, we did not realise that issues over idle time and measurement error were not a concern in the Dutch context. As mentioned, this calculation is still reported in the supplementary materials.

incident type could be classifiable under the Ratcliffe (2021) schema. A decision was made to create a new class for this incident type.

6.3 METHODS

We refer readers to the preregistration document for a complete summary of the methods deployed to answer our main question, namely, how robust findings on proportional breakdowns of time are to different operationalisations of ‘dispatched deployment time’. We also report on dispatch numbers using descriptive statistics and an accompanying visualisation on the distribution according to specific incident types. This second part of the analysis serves as a preliminary exploration into multi-dispatch incidents, given their potential role in determining dispatched deployment time calculations (see Fig. 6.1), and has not been preregistered.

6.4 RESULTS

Demand breakdown The overall breakdown of time according to each of the three operationalisations is summarised in Table 6.2. This highlights a number of initial insights. First, similar to findings in the US, in Amsterdam the public rely on the police for a diverse array of issues which are, more often than not, non-criminal. Taking the unit-level measure of ‘cumulative dispatched to clearance’ as an example, the 31% of crime incidents consumed 40% of total dispatched police deployment time during the year. Quality of life consumed 19% of time, suspicious situations 18%, traffic 15%, health 7% and community just 1%.

Table 6.2: Breakdown of counts and proportions of time spent (dispatched to clear)

Demand class	Incident count	% incident count	% Cumulative	% FILO	% FILO*units
Community	1,166	1.06	0.54	0.66	0.38
Crime	34,019	30.99	40.11	38.51	37.93
Health	6,043	5.51	7.17	7.78	11.90
Quality of life	30,484	27.77	19.43	21.59	17.56
Suspicious situation	22,864	20.83	17.59	17.42	15.34
Traffic	15,193	13.84	15.16	14.05	16.89

That said, we do observe some differences in the proportional breakdown according to the different operationalisations of time. For instance, using the cumulative measure, 7% of total dispatched deployment time is consumed by health incidents, but using the weighted FILO measure, this figure is 12%. Nevertheless, the impact of using different time measures is minimal in the sense that it does not change the finding that Amsterdam police spend the majority of their reactive dispatched deployment time resolving non-crime issues. The overall ranking of the demand classifications based on this relative proportion of ‘time spent’ is also the same, irrespective of the operationalisation used, with the exception of the weighted FILO measure. Here, traffic and suspicious situation swap place. We observe these findings irrespective of the start time used (see supplementary materials).

A proportional breakdown of time for the specific incident types nested within these broad demand classifications is visualised in Fig. 6.2. For brevity, we visualise the cumulative dispatched to scene clearance measure in isolation. Comparable graphics for all measures are available in the supplementary materials. Again, the picture is similar irrespective of the

measure used. A considerable proportion of time spent attending crime incidents is consumed by violence and theft offences. The vast majority of time consumed by traffic-related duties involve accidents. Time spent on quality of life incidents is dominated by resolving nuisance calls (e.g., noisy neighbors) and conflicts (e.g., verbal confrontations). The largest health incident category is vague: unwell and sickness. Unfortunately, we cannot differentiate between physical ill-health (e.g., a fall) and mental ill-health, although we can see some distinctions with regards to accidents and suicide attempts. We know from observational work in the dispatch centre itself that ‘suspicious situation – person’ is often used for confused or vulnerable people. In the Netherlands, dispatch centres assign nearby police, ambulance and fire service to respond to CPR by default, hence the substantial amount of time consumed by such incidents.

6

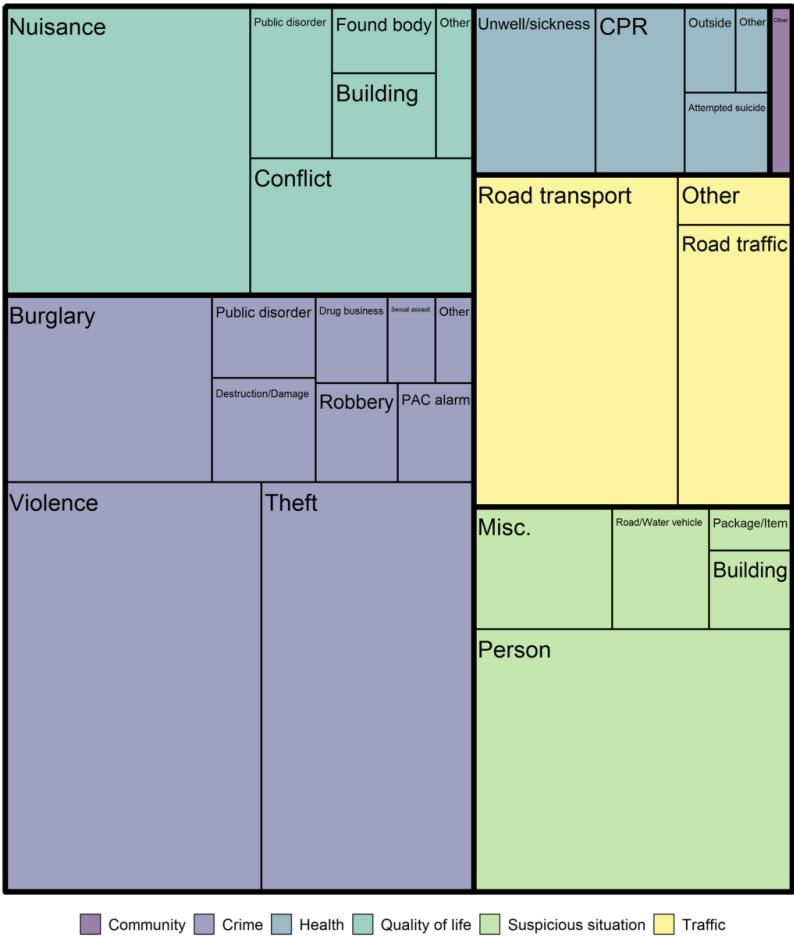


Figure 6.2: Proportional breakdown of dispatched deployment time, operationalised using the ‘dispatched to clear’ measure

6.4.1 TIME SPENT

In Table 6.3, descriptive statistics summarise the amount of time (in minutes) typically consumed by the different demand classifications. Given the skew in the data, we focus on the median. Incidents involving crime and health typically consume a comparable amount of time, with a median of 43 and 42 min respectively. Community issues, quality of life cases, and investigating suspicious situations consume less time (19, 22 and 25 min, respectively). Traffic incidents, on median, take a little longer to resolve (32 min). The equivalent statistics for all time measures are available in the supplementary materials.

Table 6.3: Descriptive statistics about the amount of time (in minutes) consumed

Demand class	Min	Max	Mean	Median	SD
Community	0	571.60	34.33	19.03	52.50
Crime	0	6006.98	87.23	43.45	173.21
Health	0	3739.10	87.75	42.15	160.65
Quality of life	0	14638.42	47.16	22.12	154.96
Suspicious situation	0	10486.23	56.91	24.92	134.68
Traffic	0	15919.25	73.80	32.07	196.61

In Fig. 6.3, we use Lorenz curves to visualise the concentration of dispatched police deployment time according to each of the main three ‘time spent’ measures. Irrespective of the time measure used, we see that a considerable proportion of dispatched police deployment time in Amsterdam is consumed by a small number of incidents. Using the cumulative measure as an example (in blue), we can state that 25% of total dispatched deployment time is consumed by just 2% of incidents and 50% of time is consumed by 10% of incidents.

We observe different degrees of concentration depending on the method for operationalizing time spent: the standard FILO measure is *least* concentrated while the weighted FILO measure is *most* concentrated. This finding is consistent with the skew demonstrated in Fig 6.1. Standard FILO does not account for the number of units: incidents for which there were multiple dispatches do not contribute any more to concentration than an incident for which only one unit attended. By contrast, the weighted measure accounts for units, but the multiplication with FILO (rather than unit-level time) contributes to a higher degree of concentration.

6.4.2 MULTI-DISPATCH INCIDENT TYPES

Descriptive statistics on the numbers of dispatches by demand classification are reported in Table 6.4. As expected, a large proportion of dispatches during the year were made in response to crime incidents (32%) although as reflected in the median, often only one unit is dispatched (noting that in the Dutch context this will often be a car containing two officers). The same is true for community, quality of life and suspicious situations. Traffic incidents have a median of two dispatches (this will often involve closing roads due to a traffic accident). Due to the CPR response procedure noted earlier, health incidents have the highest median number of dispatches with three. Perhaps most interestingly in terms of multi-dispatch incident types, the maximum number of dispatched units by demand classification is considerable. These range from community incidents involving eleven dispatches, to quality of life incidents for which 62 dispatches were made.

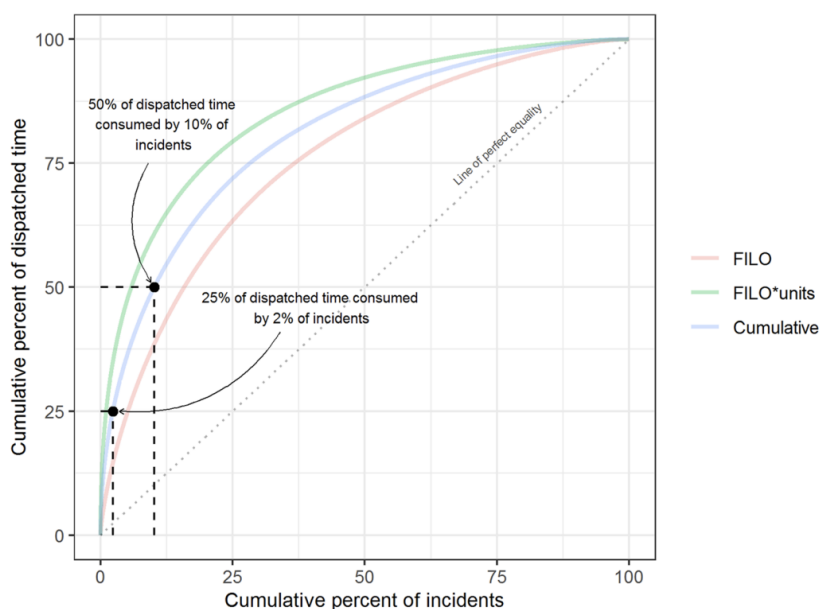


Figure 6.3: Lorenz curves for each of the main time measures with an example of the thresholds highlighted for the 'dispatched to clear' measure

Table 6.4: Descriptive statistics about the number of dispatches by incident type

Demand class	Total dispatches	% dispatches	Min	Max	Mean	Median	SD
Community	1,623	1	1	11	1.39	1	0.80
Crime	68,811	32	1	44	2.02	1	2.05
Health	21,587	10	1	32	3.57	3	2.37
Quality of life	47,455	22	1	62	1.56	1	1.55
Suspicious situation	40,499	19	1	32	1.77	1	1.47
Traffic	33,386	16	1	32	2.20	2	2.18

To further drill-down into multi-dispatch incident types, we visualise the distribution of dispatch counts for specific call types (see Fig. 6.4). These are the top-10 call types in terms of the median number of dispatches after having excluded those call types with less than twenty dispatches throughout the year. Eight out of these ten call types involve an imminent threat to life: serious violence, fires, CPR, and vehicle or people in water. The remaining two involve robbery (commercial and residential). In the Netherlands, incidents can be assigned three different priorities (1 – most serious; 3 – least serious). All ten of these incident types are priority-1.²

²Priority levels 4 and 5 do exist in the system but are rarely used in practice. Such incidents account for 6% of incidents in our study period.

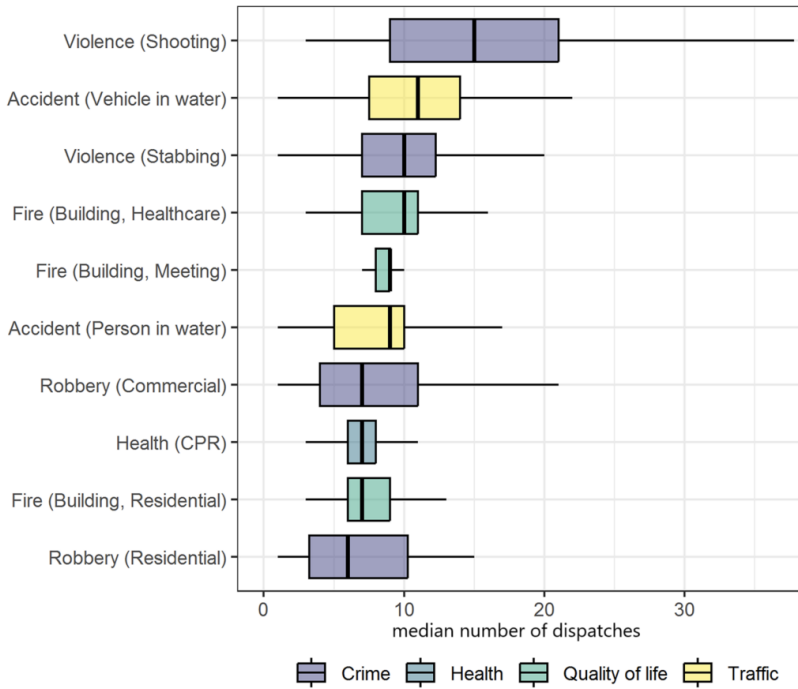


Figure 6.4: Top-10 incident types by the median number of dispatches. Note that we focus on the interquartile range in isolation, given that the minimum and maximum values are already reported

6.5 DISCUSSION

In this short contribution, we set out to investigate to what extent proportional breakdowns of police ‘time spent’ are robust to the different operationalisations of *dispatched deployment time*. The proportion of total time spent on crime varies between 38 and 40% across the three measures. The ranking of different demand classifications by proportions of total time remains the same for cumulative (unit-level) and FILO (dispatch-level) operationalisations, irrespective of how we define the ‘start’ of deployment (dispatched, on way, or on scene). We observe only a minimally diverging ranking for the weighted FILO (Table 6.4).

Our descriptive exploration of dispatch numbers shows that there is variation between demand classifications in terms of the average (mean or median) number of responding units. Incidents involving health, for example, typically receive more dispatches compared to other incident types. In turn, we see a jump in the proportion of time consumed by Health when using the weighted FILO measure.

Only with unit-level data do we have the opportunity to calculate *actual* dispatched deployment time. But, many studies do not have access to the unit-level data required to calculate the cumulative operationalisation. Based on our data in Amsterdam, if the goal of the analysis is to investigate the general breakdown of dispatched deployment time per demand class, the findings are fairly robust to the various operationalisations. Especially when research questions pertain to ranking of the categories based on relative time spent.

Conversely, if the goal of the analysis is to investigate demand classes or incident types for their individual absolute and/or proportional size, not all demand classes and incident types will be robust to the various operationalisations. Namely, classes which receive systematically more dispatched units (e.g. Health) will be under- and overrepresented in absolute terms (like also in relative terms) by, respectively, the standard and weighted FILO measures used in existing US studies.

We make no claim of generalisability to the US, from which the research originated. Nevertheless, the underlying issues demonstrated here over the usage of CAD for the purposes of calculating dispatched deployment time are still highly relevant, and we certainly encourage future research to engage with the limitations and biases of the operationalisation(s) used, ideally with reference to the prevalence of multi-dispatch incidents in the data, and in the context of the study region (e.g., dispatch practices).

Although we have shed new light on the various operationalisations of dispatched deployment time, some important factors remain unknown. We would propose using GPS tracking data of police units (e.g., used for different purposes in Dau et al. (2023)) to capture unit movements towards an incident that were not initiated by or communicated to the dispatch centre, such as those initiated after overhearing about an incident on the radio. These movements might constitute (unnecessary) flocking but are unmeasurable using the data used here, and difficult to determine using CAD data Lum et al. (2022). Even then, neither data sources fully capture the context of multi-dispatch incidents. These investigations would also benefit from interviews and observations of dispatch centre personnel and officers to better understand the conditions under which flocking might occur. We hope that these findings encourage further investigations in this area.

Lastly, as an additional point, we would like to highlight that – as our double-coding exercise demonstrated – recoding police incident types into broader a demand classification is no straightforward task. While we achieved an inter-coder reliability that surpassed the -preregistered threshold, the process was far from perfect. We suggest that all future research follow the methodology we have described in the preregistration protocol to make this imperfect process at least as transparent as possible.

6.6 AVAILABILITY OF DATA AND MATERIALS

The data used for this study cannot be made publicly available for reasons of confidentiality. Additional materials and results, including the preregistration document, can be found on the Open Science Framework repository: <https://osf.io/qgwwv6/>.

REFERENCES

- Dau, P. M., Dewinter, M., Vander Beken, T., Witlox, F., & Vandeviver, C. (2023). Police equilibrium: a glocal assessment of the demand and supply of police. *Policing and Society*, 33(5), 537–554. <https://doi.org/10.1080/10439463.2022.2149750>
- Krippendorff, K. (2011). *Computing Krippendorff's alpha-reliability* (tech. rep.). University of Pennsylvania.
- Langton, S., Ruiter, S., & Verlaan, T. (2023). Describing the scale and composition of calls for police service: a replication and extension using open data. *Police Practice and Research*, 24(5), 523–538. <https://doi.org/10.1080/15614263.2022.2102494>
- Lum, C., & Koper, C. S. 2021. Guest post: Shifting mental health calls away from police will not greatly affect their workload. *The Washington Post*.
- Lum, C., Koper, C. S., & Wu, X. (2022). Can We Really Defund the Police? A Nine-Agency Study of Police Response to Calls for Service. *Police Quarterly*, 25(3), 255–280. <https://doi.org/10.1177/10986111211035002>
- Ratcliffe, J. H. (2021). Policing and public health calls for service in Philadelphia. *Crime Science*, 10(1), 5. <https://doi.org/10.1186/s40163-021-00141-0>
- Ratcliffe, J. H., Taylor, R. B., Askey, A. P., Thomas, K., Grasso, J., Bethel, K. J., Fisher, R., & Koehnlein, J. (2021). The Philadelphia predictive policing experiment. *Journal of Experimental Criminology*, 17(1), 15–41. <https://doi.org/10.1007/s11292-019-09400-2>

7

DESCRIBING THE SCALE AND COMPOSITION OF CALLS FOR POLICE SERVICE

A REPLICATION AND EXTENSION USING OPEN DATA

This paper describes the scale and composition of emergency demand for police services in Detroit, United States. The contribution is made in replication and extension of analyses reported elsewhere in the United States. Findings indicate that police spend a considerable proportion of time performing a social service function. Just 51% of the total deployed time responding to 911 calls is consumed by crime incidents. The remainder is spent on quality of life (16%), traffic (15%), health (7%), community (5%), and proactive (4%) duties. A small number of incidents consume a disproportionately large amount of police officer time. Emergency demand is concentrated in time and space, and can differ between types of demand. The findings further highlight the potential implications of radically reforming police forces in the United States. The data and code used here are openly available for reproduction, reuse, and scrutiny.

7

7.1 BACKGROUND

Amidst austerity measures, growing public expectation, and scrutiny, understanding the public demand for police services has become a priority among evidence-based policing researchers and practitioners (Boulton et al., 2017). Without a grasp on the scale and composition of police demand, we are likely to observe suboptimal and inequitable outcomes for the public, the inappropriate distribution of public funds, and unnecessary strain on officers (Ellison et al., 2021; Lum et al., 2022). Understanding the characteristics of public demand for the police has become particularly pertinent following recent calls to rethink, and in some cases, radically reform, the role and reach of contemporary police forces (Lum et al., 2022). Despite the importance of this topic, and the public interest invested in the findings, we are currently lacking open replications and extensions of existing research. This study represents the first such investigation using open data from Detroit, United States, to replicate and extend recent analyses presented by Ratcliffe (2021) and Lum et al. (2022). It also presents open study materials which can be used to reproduce the results and replicate analyses across different contexts.

Ratcliffe (2021) sought to describe the complexity and diversity of public demand for police services in Philadelphia, United States (US), during 2019. The study was motivated by the recognition that police spend a considerable proportion of deployed resource resolving incidents which fall outside of the traditional crime-fighting role of the police. Instead, the origins of public demand for the police can often be attributed to a lack of supply and/or accessibility failure in other organizations, such as those providing support for people requiring (mental) health assistance (van Dijk & Crofts, 2017; Wood et al., 2021). Ratcliffe reported that just 40% of public emergency calls for service during the year were crime incidents. These calls consumed 55% of the total officer shift time. The remaining public demand from calls involved community issues (10% of calls, 7% of time), medical/public health incidents (7% of calls, 9% of time), proactive policing (8% of calls, 5% of time), quality of life (24% of calls, 14% of time), and traffic duties (11% of calls and 11% of time). Comparable results were subsequently reported across multiple anonymized police jurisdictions in the US (Lum et al., 2022); see also, Lum and Koper (2021) in *The Washington Post*.

These descriptive findings provided data-driven evidence to suggest that the police indeed spend considerable amounts of time dealing with incidents in a social rather than solely crime-fighting capacity. This has implications to and beyond policing practice. The findings directly contribute to the fierce and politically charged debates around the role and responsibilities of police officers in contemporary society. Both Ratcliffe (2021) and Lum et al. (2022) were published amidst the backdrop of the killing of George Floyd in May 2020 and subsequent ‘defund the police’ protests across the US. Evidence which demonstrates that the public relies on the police for social duties beyond law enforcement suggests that a radical reduction in their role and responsibilities (e.g., through defunding) would increase demand on other social services. Should such radical reform take place without first understanding the scale and composition of public demand for the police, and without preemptively providing appropriate substitutes, there will be no supply to meet this shift in demand. Thus, Ratcliffe (2021) and Lum et al. (2022) contributed important empirical data to the (ongoing) discussion on police reform in the US.

Here, we propose an extension of Ratcliffe (2021) and Lum et al. (2022). This extension

has two principal components. First, we replicate the main analyses of these existing investigations into the scale and composition of police demand using open study materials. To date, investigations have been conducted on data which are not publicly available and without the materials (e.g., code) to reproduce the results. Open science practices have been identified as fundamental to the aims of evidence-based policing and vital in ensuring that in the long term, real-life decisions are based on a concrete and replicable pool of evidence (Bennell & Blaskovits, 2019). Yet, reproductions and replications are rare. This has been demonstrated in domestic policing research (Huey & Bennell, 2017) and criminology more broadly (McNeeley & Warner, 2015). The necessity for a fully open investigation into the scale and composition of police demand is particularly important given the public nature of the debate around police reform and the considerable implications of redefining the role of the police without transparent empirical investigation. It is reasonable for police practitioners and the public to expect a thorough, open, and replicable investigation into police demand before academia concludes that the issue is resolved – certainly, before findings are used to change or introduce police practice or public policy. Open materials have the additional benefit of being available for reuse in other jurisdictions, helping build a reliable and comparable evidence-base around the scale and composition of police demand across different contexts.

Second, we extend the analyses used by Ratcliffe (2021) and Lum et al. (2022). This includes a comparison between measures of officer time with and without the response time. While Lum and colleagues raised concerns about the inclusion of the response time in measures of officer time, largely due to low-priority calls not necessitating a rapid response, the precise impact of their usage on the same data is unclear. We also provide a comprehensive account of the variability in the time consumed by different demand classifications (e.g., crime, community, health). These statistics shed light on the degree to which a disproportionately large amount of officer time is consumed by just a small number of calls – insight which can be used to assess how a reduction in demand, or a shift away from calls for police service towards other social services, might free-up officer time. In addition, we report on the spatial and temporal patterning of each demand classification. Distinctions in these patterns across demand types would be indicative of similarly distinct demand-generating mechanisms (e.g., built environment). Such concentrations in time and space might also inform deployment decisions based on officer expertise and experience, and/or collaboration with other services, such as mental health professionals (e.g., White & Weisburd, 2018).

With these points in mind, this paper describes the emergency demand for police services in a large city, Detroit, US, using open data and open code. All materials are publicly available (see https://github.com/langtonhugh/demand_detroit). It does so in replication, and where appropriate extension, of the main findings reported by Ratcliffe (2021) and later Lum et al. (2022). In this way, the contribution serves to test the robustness of existing findings; provides the materials required for reuse, reproduction, and replication; and provides additional insight through new analyses.

7.2 DESCRIBING DEMAND

Given the aims of this study, this section first provides a brief outline of the data and methods used by Ratcliffe (2021) and Lum et al. (2022) to describe the scale and composition of emergency reactive police demand.

7.2.1 RATCLIFFE (2021)

Using computer-aided dispatch (CAD) data from the Philadelphia Police Department (PPD), the Ratcliffe (2021) study provided a breakdown of calls for police service during 2019. Calls were subset to exclude those which did not require officer dispatch or did not originate from the public. CAD codes, discrete descriptions of incidents based on information available to the dispatcher, were categorized into six broad classifications: community issues, crime, medical/public health, proactive policing, quality of life, and traffic duties. The classification scheme was used to report on the frequency and proportional breakdown of calls relating to each type of demand throughout 2019.

Recognizing that frequency counts of calls for service do not necessarily reflect the amount of police resource consumed by public demand, raw counts were augmented with the time spent by officers on each of the six classifications of demand. This information permits the calculation of a proportional breakdown of police time consumed by the different demand classifications throughout the year. The raw numbers were reported in a table. A visual representation of call types nested within each demand classification was summarized in a tree map graphic. The time committed by officers was calculated from when the call was dispatched to an officer to when the officer indicated that the event had been closed. Ratcliffe accounted for the number of officers attending each incident when calculating the amount of time consumed. The aggregate time consumed by different demand classifications was reported using a single count of officer shifts (7.5 hours) over the year.

The spatio-temporal patterning of this demand was summarized in two visuals, focusing specifically on those calls relating to medical/public health. A heatmap showed the mean call volume during hours of the day and days of the week, identifying peak times for demand. A geographic map visualized the spatial patterning of these call count concentrations using kernel density estimation in order to identify hotspots of police demand involving medical/public health issues. To supplement these findings, Ratcliffe also reported on the differences between the initial CAD incident type classification and the final disposition using a Sankey diagram – again focusing specifically on medical and public health calls.

7.2.2 LUM ET AL. (2022)

To describe the scale and composition of emergency police demand across the US more broadly, Lum et al. (2022) used CAD data from nine different police agencies. The agencies themselves were anonymized. Individual incidents were categorized into 14 different demand classifications, namely, alarms, disorders, vice, domestic, follow-ups and service requests, mental, medical, missing persons, violence, interpersonal-other, property, suspicions, traffic, and admin, agency, and non-crime. The authors reported a proportional breakdown of call counts and police time spent across the nine agencies according to the 14 demand classifications, including the average (mean). For each agency, the authors also reported the average time spent on each demand classification in minutes, and the average across the nine agencies. No descriptive statistics on the range or spread of these figures was reported.

The ‘time spent’ measure used was the *time on scene* (i.e., excluding the response time). The authors considered this to be preferable since low priority calls may have a non-urgent (slow) response due to the nature of the call, and officers may be reassigned to other duties while on an assignment. Lum and colleagues calculated the *time on scene* as the time from the first officer’s arrival to when the call was reported as resolved. In this way, the calculation

did not account for the additional time consumed by multi-officer responses (e.g., backup). This was justified for a number of reasons, among them, avoiding the over-inflation of time estimates due to multiple officers attending non-serious incidents, and limitations with the CAD data. A proportional breakdown of the disposition (assist only, report written or arrest citation) was provided for all 14 call types. No spatial or temporal findings were reported.

7.3 DATA

7.3.1 DETROIT CALLS FOR SERVICE

The data used for this study cover the city of Detroit in Michigan, US. The City of Detroit authority publish 911 emergency calls for service data through their Open Data Portal.⁰ In alignment with the two existing studies, the data used here exclude those calls initiated by an officer in order to capture public-initiated calls. Each individual incident has a corresponding timestamp and location coordinates to define it in time and space. Data was a subset of 2019.

For each incident, the police response time and time on the scene are reported. The former measures the time from dispatch to when the first officer arrives on-scene.¹ The latter mimics the measure used by Lum et al. (2022), namely, the amount of time which elapses between when the first officer arrives and when the last unit clears the scene. Here, the data permit a distinction between the time spent on calls with and without the response time. We refer to the sum of the response time and time on scene as the *total deployed time*. Calls which had a response time of negative values or missing were dropped from analyses. Calls which had a *time on scene* of 0 minutes, negative values or missing were also excluded. As noted, Lum et al. (2022) considered the time on scene measure to be preferable. For completeness, we report times with and without the response time – defined as travel time.

Each incident has a *call description* variable which describes the nature of the call. Call descriptions involving administrative duties (e.g., ‘start of shift information’), completely unknown problems, and non-deployment (e.g., ‘employee call in/time off’) were removed. This left 207 unique call description categories. In the interests of parsimony and ease of interpretation, these categories were recoded into 99 broader call descriptions.² This included the removal of incidents involving transporting prisoners and executing warrants, which were deemed to have been erroneously flagged as not being officer-initiated. Each call description was then categorized into the six demand-type classifications used in Ratcliffe (2021), namely, crime, community, medical/public health, proactive, quality of life, and traffic. A total of 2.8% of calls were deemed unclassifiable. For each of the six demand classifications, an ‘other’ incident type was generated for those incidents which consumed less than 0.2% of the total deployed time during the year. This left a total of 49 unique categories for the reported breakdowns, capturing 260,000 individual public calls for service in Detroit during 2019. No information on the final disposition of calls is provided in the open data.

⁰Accessible via <https://data.detroitmi.gov/>.

¹In the raw data, the time between dispatch and first on-scene arrival is referred to as *travel time*. The *intake* and *dispatch time* are reported; however, we did not consider these measures to reflect deployment time. We return to this point in the discussion.

²We refer readers to the corresponding GitHub repository for further details on how call descriptions were combined (https://github.com/langtonhugh/demand_detroit). Readers unfamiliar with R code can review the reference table entitled ‘categorization summary’.

7.4 METHODS

The methods used largely replicate those used by Ratcliffe (2021) and Lum et al. (2022), and where appropriate, extensions. First, descriptive statistics and a treemap graphic provide an account of the scale and composition of police demand during the year. Second, visualizations summarize the spatial and temporal patterning of this demand.

7.4.1 DESCRIPTIVE STATISTICS

Summary statistics describe the overall composition of demand according to the six demand-type classifications. This includes call frequency counts, a proportional breakdown of call counts, and a proportional breakdown of police time. Both *total deployed time* (response time + time on scene) and *time on scene* measures of police resource are reported. The proportional breakdown of *time on scene* is visualized using a treemap graphic for each incident type nested within the demand classifications. The equivalent graphic for *total deployed time* is available in the corresponding GitHub repository.

A series of descriptive statistics summarize the time on scene consumed by each demand type. For comprehensiveness and transparency, we report the mean, median, standard deviation, and range, and in the Appendix (Figure C.2), a histogram of the underlying distribution. This sheds light on how long a ‘typical’ call might take while noting a probable skew in the distribution. Such a skew would suggest that averages of time (e.g., Lum et al., 2022) or aggregate measures of time (e.g., Ratcliffe, 2021) in isolation are insufficient and potentially mask important underlying variation, namely, that a small number of high-demand calls drag up the mean and inflate aggregate measures of time. To investigate this distribution, we plot Lorenz curves for each demand classification. Lorenz curves plot the cumulative distribution of a variable (in this case, time on scene) against the cumulative distribution of a unit (in this case, incidents; Bernasco & Steenbeek, 2017). By way of example, we report on these figures according to thresholds of 25% and 50% of time spent. Here, we recognize that the thresholds are arbitrary, but their usage provides an interpretable metric for demonstrating the degree to which a disproportionately large amount of officer time might be consumed by just a small number of calls.

7.4.2 SPATIO-TEMPORAL PATTERNING

The temporal patterning of the six demand-type classifications are visualized using a heatmap of mean call counts, aggregated by day of the week, and hour of the day. To summarize the spatial patterning, call location coordinates are aggregated to synthetic grid cells using spatial joins. Grid cells are defined as 1000 ft² (305 m²) – a size deemed appropriate based on a balance between capturing localized variation and interpretable visualizations. Two percent of incidents are excluded from the spatial visualizations due to incomplete geographic information. To accurately convey the resource deployed to each area during the year, maps visualize the aggregate *time on scene* (in hours) within each grid cell over the course of the year for each demand-type classification. The equivalent raw count maps are available in the Appendix, Figure C.1. In this way, we visualize the spatio-temporal patterning of demand in a manner comparable to Ratcliffe (2021), while adding a comparison across different demand classifications and two different measures of demand.

7.5 RESULTS

7.5.1 SCALE AND COMPOSITION

The frequency counts and proportional breakdowns for each demand-type classification are reported in Table 7.1. This demonstrates that in Detroit, a considerable proportion of calls for service, and in turn, time spent by officers, is committed to a diverse array of (often non-criminal) issues. Just 44% of emergency calls for service related to crime incidents. These crime incidents consume 51% of time on the scene throughout the year. The remainder of time on scene is spent dealing with traffic duties (16%), quality of life incidents (15%), public and mental health (7%), community issues (5%), and proactive policing (4%). The differences in the proportional breakdown between time on scene and total deployment time, which includes response time, are marginal. This indicates that the response time – as measured by travel time – plays only a minor role in determining the scale and composition of the time consumed.

Table 7.1: Breakdown of frequency counts and time spent on each demand type. Total deployed time is the sum of the time on scene and response time

Demand type	Count	Count (%)	Time on scene (%)	Total deployed time (%)
Community	14,466	5.59	5.06	5.15
Crime	114,800	44.36	51.40	50.55
Health	19,873	7.68	7.10	7.19
Proactive	15,663	6.05	3.89	4.04
Quality of life	54,702	21.14	14.66	15.54
Traffic	32,147	12.42	15.63	15.27
Unclassified	7,135	2.76	2.25	2.26

The proportional breakdown of time spent on scene for each incident type, grouped by the six demand classifications, is visualized in Figure 7.2.³ This figure demonstrates the diversity and complexity of public demand for police services in Detroit. Time consumed on the scene of crime incidents is largely comprised of domestic incidents, burglary, and assault. For traffic incidents, police time was mostly spent attending to accidents, while for quality of life, attending to disturbances consumed the most time on scene. Calls relating to public health involved mental health incidents, overdoses, (threats of) suicide, and well-being checks. Proactive incidents largely involved investigations, while time consumed on community issues consisted mostly of missing/found persons and animal incidents (e.g., aggressive dogs).

³The purpose of the graphic is to provide a broad summary, in replication of Ratcliffe’s visual. The numbers underlying the graphic are reported in the corresponding GitHub repository.

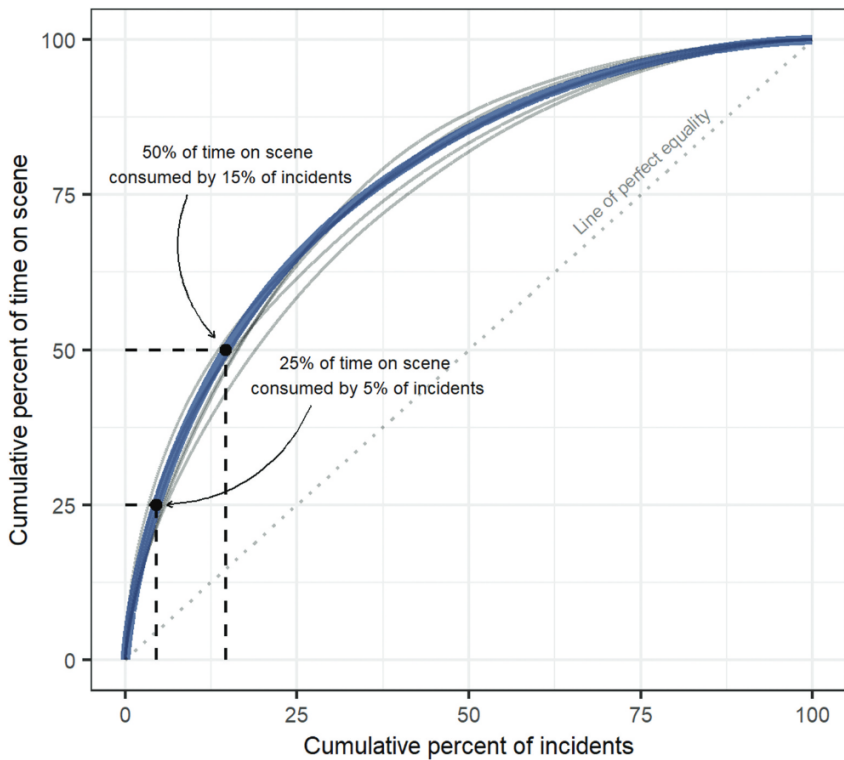


Figure 7.1: Lorenz curve for time spent on scene across all incidents (in blue). Lorenz curves in grey for each of the six demand classifications

7

7.5.2 TIME DISTRIBUTIONS

To further quantify the resources consumed, Table 7.2 summarizes the time spent on the scene for each demand classification in minutes. Histograms of the underlying distributions are reported in the Appendix. Table 7.2 demonstrates a number of things. First, there is a modest amount of variation in the typical time at the scene of different types of emergency calls. For instance, the mean time spent on the scene of traffic incidents (55 minutes) and crime (51 minutes) is considerably more than that of quality of life issues (31 minutes) or proactive activity (28 minutes). These differences will be a result of a combination of factors, including the complexity and severity of incidents. Second, there is considerable variation in the time consumed by incidents *within* each demand type. While many calls are resolved in minutes, a handful consume hours of officer time on the scene, resulting in a highly skewed distribution. This also highlights how mean or aggregate summary statistics of police time as a measure of demand can mask (important) underlying variation.

Figure 7.1 visualizes the disproportionate share of time spent on a small number of calls explicitly using Lorenz curves for each demand classification. These can be compared against the reference line of perfect equality. Here, the line of perfect equality represents a situation in which officer time is distributed evenly across all calls. In reality, we can see that

Table 7.2: Descriptive statistics of minutes spent on scene for each demand-type classification

Demand type	Mean	Median	Min	Max	SD
Community	40.0	24.6	0.1	847.7	48.0
Crime	51.0	30.4	0.1	987.0	64.1
Health	40.9	29.1	0.1	897.4	45.2
Proactive	28.4	16.5	0.1	907.9	47.4
Quality of life	30.5	19.7	0.1	885.6	41.1
Traffic	55.0	30.0	0.1	901.1	65.0

this is not the case: the Lorenz curves for each demand classification deviate considerably from perfect equality. The curves for each classification are almost indistinguishable, so we focus on the curve for all calls (in blue). This shows that 50% of the time officers spent on the scene during 2019 was consumed by just 15% of calls, and 25% of time on scene was consumed by just 5% of all calls. In other words, a small number of emergency calls for service account for a disproportionately large amount of police resources out on the street. This finding is consistent (and comparable) across all six classifications of demand.

7.5.3 SPATIO-TEMPORAL PATTERNING

Mean incident counts by day of the week and by hour of day, for each demand classification, are visualized in Figure 7.3. This demonstrates that the different types of public demand for police services, as indicated by the six demand classifications, can have a distinct temporal patterning. Community calls appear to concentrate on weekdays, around lunchtime, when other classes of demand are low. Crime and quality of life-related calls both appear to occur in the evening and early hours of the morning. Traffic incidents concentrate in the evenings, particularly on a Friday. Proactive policing does not have a distinct hourly patterning, but this could be attributed to the low hourly counts. Calls relating to health issues, the primary interest of Ratcliffe (2021), appear to largely take place in the evening and early morning.

The spatial patterning of police demand in Detroit, defined as the aggregate time spent on the scene throughout the year, is visualized for each demand classification in Figure 7.4.⁴ This demonstrates that the time police spent attending to incidents is highly concentrated in space and can vary between different demand types. The police spend a considerable amount of time around the Wayne State University (WSU) and Midtown attending to community, proactive and quality of life calls for service. Officers spend vast amounts of time on crime-related calls in or near hospitals, most notably, the Henry Ford Hospital. Police time spent attending to health-related calls concentrates in and around WSU and Midtown, but also among grid cells containing mental health and well-being facilities outside of the city center. Time on traffic incidents has a distinct spatial patterning along arterial roads into Downtown.

⁴Noting that the aggregate time spent on scene may be influenced by outliers, the raw count maps and underlying distributions are reported in the Appendix. These indicate that 'high time' outliers do not seriously misrepresent the aggregate figures in Figure 7.4.

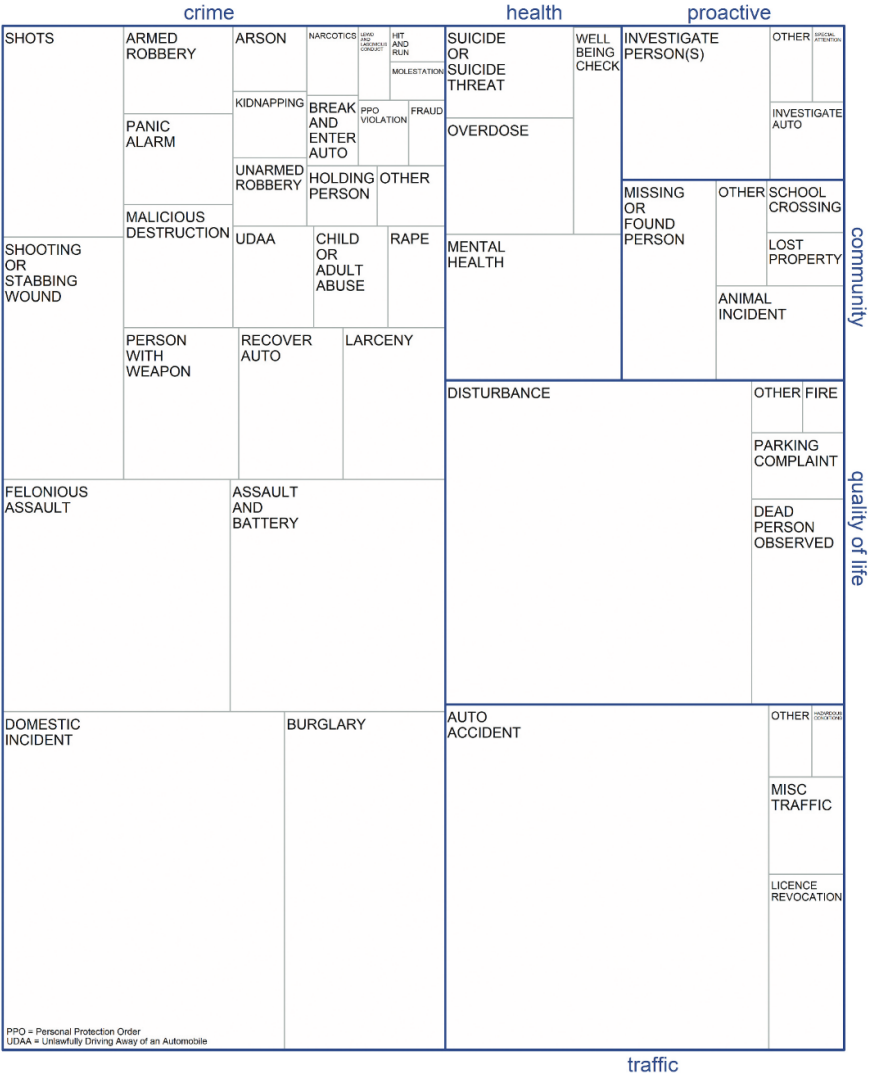


Figure 7.2: Proportional breakdown of police time consumed over the year, defined as time spent on the scene, for each call category

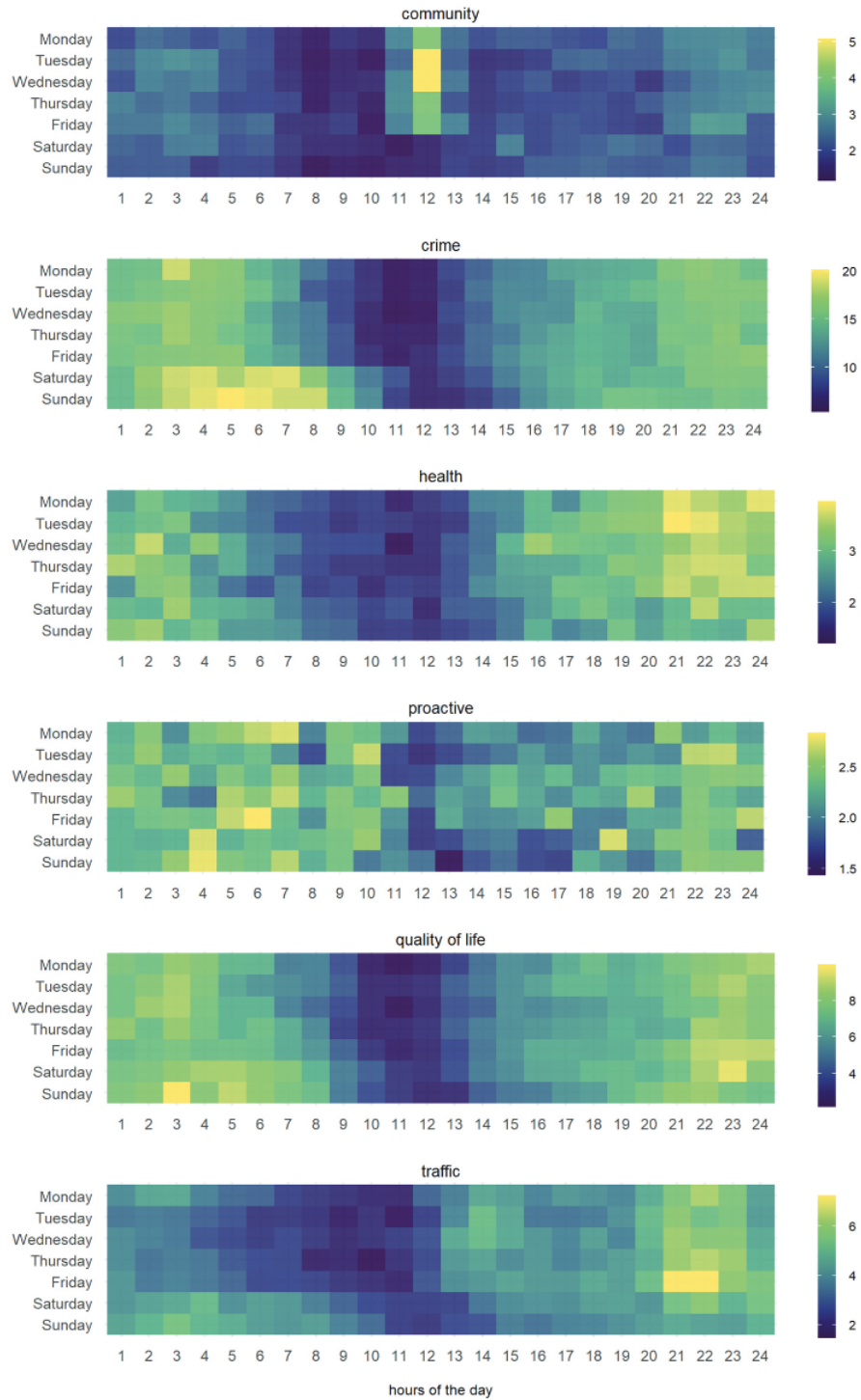


Figure 7.3: Mean call counts by day and hour, for each demand type

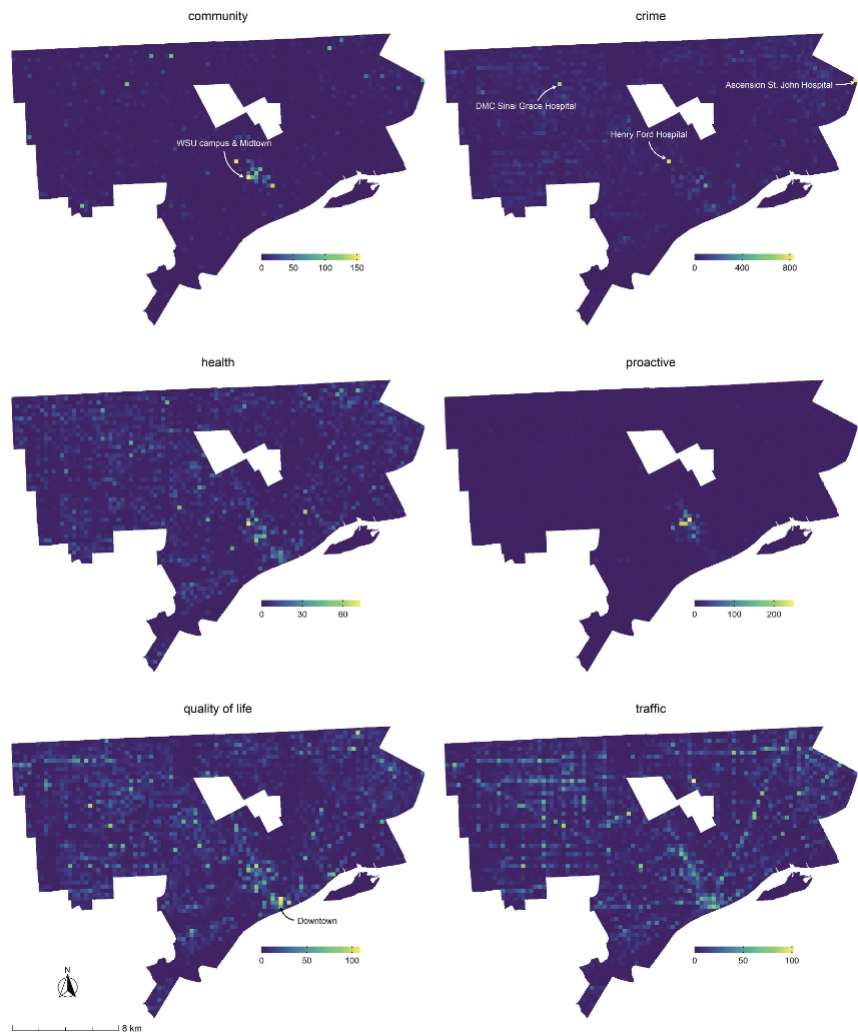


Figure 7.4: Spatial patterning of the aggregate time spent on scene (in hours) for each demand type

7.6 DISCUSSION

In replication and extension of the main findings reported by Ratcliffe (2021) and Lum et al. (2022), this contribution has described the scale and composition of calls for police service in Detroit, US. Broadly speaking, the results presented here provide robustness to those reported by Ratcliffe (2021) and Lum et al. (2022). There is substantial empirical evidence to demonstrate that, in multiple jurisdictions, the public rely on the police for both social and crime-fighting functions. The evidence highlights the potential implications of radically reforming police forces in the United States. A reduction in the capacity of the police to meet emergency demand, without first providing appropriate, capable, and accessible replacement emergency services, could have considerable repercussions for public safety and well-being. This contribution provides the first such demonstration using open data and open code, opening up analyses for scrutiny, reuse, and extension by the public, academics, and/or police practitioners. We encourage usage of the materials provided with this study to replicate analyses across other jurisdictions. In doing so, we hope to see a consistent and reliable evidence-base across different contexts for the purposes of comparison and generalization.

In this study, calls were categorized according to the six classes used by Ratcliffe (2021). The overall composition of police demand in terms of call volume and proportional time spent is remarkably similar. A considerable volume of calls for service, and in turn, a considerable proportion of deployed police time, relate to incident types which do not fall under the traditional ‘crime-fighting’ role of the police. Instead, the police fulfill a variety of other social service functions, including the resolution of public (mental) health incidents and resolving community issues. That said, descriptive statistics which drill down into the underlying distribution of officer time consumed by each demand type provide additional insight which would otherwise be overlooked. First, it is clear that there is considerable variation in the amount of time consumed both between and within demand classifications. On a methodological point, this shows how average (e.g., mean times) or aggregate (e.g., sum times) measures might mask the underlying variation. Second, and relatedly, these extensions demonstrate that a disproportionately large amount of police time is spent on a small number of emergency calls. Fifty percent of total time on scene is consumed by just 15% of calls. High-demand calls will likely be complex and/or life-threatening incidents or calls which require officers to secure a scene for extended periods of time. A targeted examination of these calls might shed light on the extent to which such high-intensity demands could be reduced or better resolved through or in concert with other public services.

In this regard, this study raises useful discussion points around the different measures of officer time. First, it is clear that findings on the scale and composition of police demand are only marginally sensitive to whether measures of officer time include the response time – defined as the time from dispatch to when the first officer arrives on-scene. The overall findings are almost identical between the two measures. Here, we would highlight that our response time measure did not include dispatch time, but rather, travel time in isolation. Including dispatch time would increase the gap between the two measures: low-priority calls will be held in the dispatch queue for longer, inflating the overall response time, while high-priority calls will be dispatched rapidly. Since certain calls will have higher priority responses (e.g., violent crime), this gap is likely to differ between demand classifications. We excluded dispatch time on the basis that it does not capture deployment, but a specific investigation into its impact would be straightforward using the open materials provided.

Second, and relatedly, we would reiterate that the ‘time on scene’ measure reported here, in alignment with Lum et al. (2022), refers to the amount of time which elapses between when the first officer arrives and when the last unit clears the scene. This differs from that of Ratcliffe (2021) which accounted for the number of officers attending the scene. Lum et al. (2022) note that including backup officers in the time calculation might introduce complications. For instance, non-serious incidents might receive backup unnecessarily, particularly during periods of low-demand, which would inflate the time consumed and overstate the resources required to resolve such an incident. Although the overall composition of police demand is comparable between Philadelphia (Ratcliffe, 2021) and Detroit, the true impact of these measures remains unclear. We propose that future research investigate these distinctions. This would require CAD (or equivalent) data, timestamped at each stage of the response (dispatch, scene arrival, scene departure), for individual police units nested within calls. This level of granularity would permit the calculation and comparison of multiple measures of officer time, as well as an investigation into the extent to which backups and/or ‘flocking’ might influence such measures. In sum, further investigations into different measures of ‘police time’ and the corresponding impact on our assessment of the scale and composition of police demand are necessary before we can claim that the evidence-base is sufficient.

The spatial and temporal patterning of demand for police services in Detroit demonstrates that different demands often originate from different places and can concentrate during different times of the day. Given the specialized training and experience required to deal with certain forms of police demand (e.g., well-being checks, suicide threats), forces might consider accounting for this variation when assigning personnel to shifts and patrol areas, or organizing co-patrols with other social services. We would emphasize that space-time hotspot policing and co-responder initiatives should be tailored for specific demand types, including disaggregated crime types. Pilot studies are already underway in this area, including the policing of mental health issues within hotspots characterized by violence and drug crimes (White & Weisburd, 2018). This is supported by recent research which has shown that, at least in the Canadian context, mental-health-related calls for service can have a different spatial patterning to crime (Vaughan et al., 2016) but that there is considerable overlap when it comes to violence (Hodgkinson & Andresen, 2019). The co-dependence of specific criminal and non-criminal forms of demand in time and space would support moves to deepen collaboration between different social services, rather than bringing about a radical shift in supply away from the police in isolation.

The spatio-temporal findings presented here might prompt further investigations into the underlying environmental factors generating different forms of police demand. Recent research has found that the explanatory frameworks underpinning environmental criminology (e.g., place attractors) are relevant for non-criminal forms of demand, such as those calls involving emotionally disturbed persons (Vaughan et al., 2016). That said, we know little about the nuances of how different environmental characteristics might drive different forms of demand in the same study area. The study materials made available here could be extended using open data resources, such as Open Street Map, to provide a theoretically driven investigation into the urban character (e.g., place attractors) of high-demand areas according to the six demand classifications. Based on the distinct temporal patterning of different demand types, such analyses should be time-sensitive, capturing, for instance, the

ambient population flows of different places throughout the day, or the opening times of risky facilities. To date, ambient population analyses have focused on crime-specific calls for service (Andresen & Brantingham, 2007) which, as we have seen, constitute only part of the emergency demands placed on police. These advancements would move analyses beyond descriptive accounts of specific study regions to a generalizable causal model of police demand.

The study is not without shortcomings. First, while the Detroit data permit an open and reproducible case site, the raw incident data do not make a distinction between the initial call classification and the final disposition. This would permit a comparison, as conducted by Ratcliffe (2021), into the extent to which the call handler's initial incident classification changes following the attendance of officers at the scene. Here, we cannot conduct such an investigation. Second, findings presented here, and those reported by Ratcliffe (2021) and Lum et al. (2022), rely on the reliability and validity of police incident classifications. These are imperfect. Recent evidence from the United Kingdom and Canada indicates that police under-record the proportion of emergency calls involving persons with mental ill-health, which in turn underestimates the amount of police time consumed by such incidents (Koziarski et al., 2022; Langton et al., 2021). Investigations into recording practices and call handler decision-making will be necessary to further understand this issue across US police jurisdictions (see Lum, Koper, Stoltz et al., 2020; Simpson, 2021).

REFERENCES

- Andresen, M. A., & Brantingham, P. J. (2007). Hot spots of crime in Vancouver and their relationship with population characteristics. *CrimRxiv*. <https://doi.org/10.21428/cb6ab371.8b4a90b7>
- Bennell, C., & Blaskovits, B. (2019). The case for open police research. In G. Tong, B. Blaskovits & N. Schneider (Eds.), *Evidence-based policing: An introduction* (pp. 117–131). Routledge.
- Bernasco, W., & Steenbeek, W. (2017). More Places than Crimes: Implications for Evaluating the Law of Crime Concentration at Place. *Journal of Quantitative Criminology*, 33(3), 451–467. <https://doi.org/10.1007/s10940-016-9324-7>
- Boulton, L., McManus, M., Metcalfe, L., Brian, D., & Dawson, I. (2017). Calls for police service. *The Police Journal: Theory, Practice and Principles*, 90(1), 70–85. <https://doi.org/10.1177/0032258X16671032>
- Ellison, M., Bannister, J., Lee, W. D., & Haleem, M. S. (2021). Understanding policing demand and deployment through the lens of the city and with the application of big data. *Urban Studies*, 58(15), 3157–3175. <https://doi.org/10.1177/0042098020981007>
- Hodgkinson, T., & Andresen, M. A. (2019). Understanding the Spatial Patterns of Police Activity and Mental Health in a Canadian City. *Journal of Contemporary Criminal Justice*, 35(2), 221–240. <https://doi.org/10.1177/1043986219842014>
- Huey, L., & Bennell, C. (2017). Replication and Reproduction in Canadian Policing Research: A Note. *Canadian Journal of Criminology and Criminal Justice*, 59(1), 123–138. <https://doi.org/10.3138/cjccj.2016.E09>
- Koziarski, J., Ferguson, L., & Huey, L. (2022). Shedding Light on the Dark Figure of Police Mental Health Calls for Service. *Policing: A Journal of Policy and Practice*, 16(4), 696–706. <https://doi.org/10.1093/police/paac006>
- Langton, S., Bannister, J., Ellison, M., Haleem, M. S., & Krzemieniewska-Nandwani, K. (2021). Policing and Mental ill-health: Using Big Data to Assess the Scale and Severity of, and the Frontline Resources Committed to, mental ill-health-related calls-for-service. *Policing: A Journal of Policy and Practice*, 15(3), 1963–1976. <https://doi.org/10.1093/police/paab035>
- Lum, C., & Koper, C. S. 2021. Guest post: Shifting mental health calls away from police will not greatly affect their workload. *The Washington Post*.
- Lum, C., Koper, C. S., Stoltz, M., Goodier, M., Johnson, W., Prince, H., & Wu, X. (2020). Constrained Gatekeepers of the Criminal Justice Footprint: A Systematic Social Observation Study of 9-1-1 Calltakers and Dispatchers. *Justice Quarterly*, 1–23. <https://doi.org/10.1080/07418825.2020.1834604>
- Lum, C., Koper, C. S., & Wu, X. (2022). Can We Really Defund the Police? A Nine-Agency Study of Police Response to Calls for Service. *Police Quarterly*, 25(3), 255–280. <https://doi.org/10.1177/10986111211035002>
- McNeeley, S., & Warner, J. J. (2015). Replication in criminology: A necessary practice. *European Journal of Criminology*, 12(5), 581–597. <https://doi.org/10.1177/1477370815578197>
- Ratcliffe, J. H. (2021). Policing and public health calls for service in Philadelphia. *Crime Science*, 10(1), 5. <https://doi.org/10.1186/s40163-021-00141-0>

- Simpson, R. (2021). Calling the Police: Dispatchers as Important Interpreters and Manufacturers of Calls for Service Data. *Policing: A Journal of Policy and Practice*, 15(2), 1537–1545. <https://doi.org/10.1093/police/paaa040>
- van Dijk, A., & Crofts, N. (2017). Law enforcement and public health as an emerging field. *Policing and Society*, 27(3), 261–275. <https://doi.org/10.1080/10439463.2016.1219735>
- Vaughan, A. D., Hewitt, A. N., Andresen, M. A., & Brantingham, P. L. (2016). Exploring the Role of the Environmental Context in the Spatial Distribution of Calls-for-Service Associated with Emotionally Disturbed Persons. *Policing*, 10(2), 121–133. <https://doi.org/10.1093/police/pav040>
- White, C., & Weisburd, D. (2018). A Co-Responder Model for Policing Mental Health Problems at Crime Hot Spots: Findings from a Pilot Project. *Policing: A Journal of Policy and Practice*, 12(2), 194–209. <https://doi.org/10.1093/police/pax010>
- Wood, J. D., Watson, A. C., & Barber, C. (2021). What can we expect of police in the face of deficient mental health systems? Qualitative insights from Chicago police officers. *Journal of Psychiatric and Mental Health Nursing*, 28(1), 28–42. <https://doi.org/10.1111/jpm.12691>

C

C

***APPENDIX: SCALE AND
COMPOSITION OF POLICE
CALLS-FOR-SERVICE***

C.1 RAW COUNT MAPS

C



Figure C.1: Spatial patterning of the raw call counts for each demand classification

C.2 TIME-ON-SCENE DISTRIBUTIONS

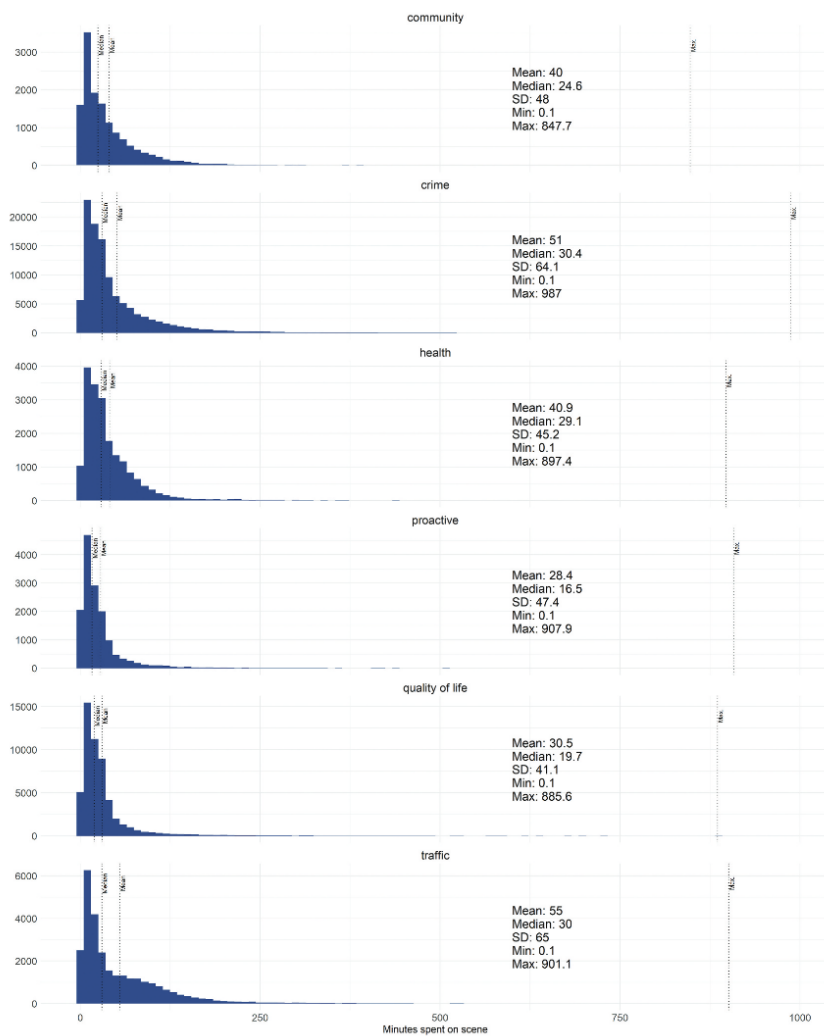


Figure C.2: Distribution and descriptive statistics of time spent on scene by demand classification at the call level

8

REACTIVE POLICE DEPLOYMENT ON THE STREET

Every day, thousands of police officers are active on the streets in the Netherlands. Much of the work these officers carry out is driven by requests for assistance from the public. This is known as reactive police deployment. The types of help citizens seek from the police are highly diverse and span a much broader range of tasks than the image of the police as merely crime fighters might suggest. In this article, we examine the nature of these public requests for assistance as they are received via the 112-emergency number. We analyse how much time is spent on them and how this differs between frontline teams and regional units.

To map reactive police deployment, we use data from the Integrated Dispatch Centre System (GMS) from 2019 – the most recent and complete calendar year not affected by the Covid-19 pandemic. In total, nearly one million incidents were recorded for which the various dispatch centres dispatched one or more police vehicles. The first two columns of Table 8.1 (*Number of incidents and Percentage of incidents*) show how these are distributed across different incident categories in the GMS. About a third of incidents relate to the *Living Environment*, and a quarter to *Public Safety and Order*. Incidents regarding *Property Crime*, *Accidents*, or *Traffic* each account for approximately 10 percent of reactive police deployment, while incidents concerning *Alarms*, *Fires*, and *Health* are far less common.

Table 8.1: Distribution of incidents and total time spent across incident categories

Incident category	Number of incidents	% of incidents	% of time spent
Alarm	50,446	5.2	2.3
Property Crime	113,054	11.6	15.8
Fire	24,119	2.5	3.3
Health	26,084	2.7	5.5
Living Environment	325,196	33.4	21.3
Accident	90,640	9.3	14.4
Safety and Public Order	238,628	24.5	30.1
Traffic	105,454	10.8	7.3
Total	973,621	100.0	100.0

The GMS system also includes detailed timestamps of when calls were *received* by the dispatch centre, when operators issued assignments, when police vehicles were *en route*, when they arrived on scene, and when they *cleared* the incident again (see Figure 8.1). It is important to note that these data are neither complete nor entirely error-free. Nevertheless, our analysis, based on these data, provides the most comprehensive overview currently available. We assess the data as sufficiently reliable to offer valuable insights into – and for – policing practice.

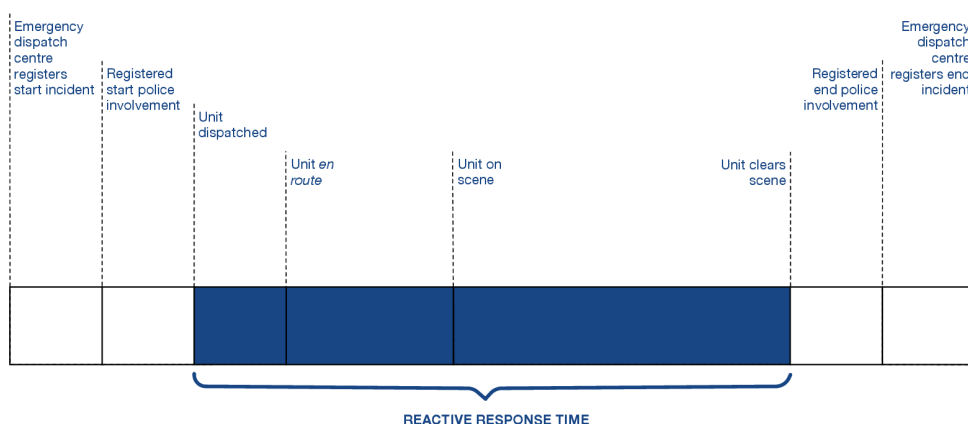


Figure 8.1: Timestamps per incident and police vehicle in the GMS data

By calculating the difference between the time the dispatch was issued and the time officers reported available again, we can compute the duration of reactive deployment per police vehicle. The total time spent per incident is then the sum of time spent by all vehicles involved. Administrative tasks following incidents are not included.

8.1 A MULTITUDE OF TASKS

The total reactive deployment in 2019 analysed here adds up to 37,437 days. The final column of Table 8.1 (Percentage of time spent) shows the relative time investment across the various incident categories. Notably, the time percentages deviate significantly from the distribution of incidents. For instance, nearly half of all incidents fall under the categories *Alarm*, *Living Environment*, or *Traffic* (49.4%), yet they account for less than a third (30.9%) of total time spent. Conversely, incidents categorised as *Property Crime*, *Fire*, *Health*, *Accident*, and *Safety and Public Order* make up half of all incidents (50.6%) but account for over two-thirds (69.1%) of the time spent.

Figure 8.2 shows the relative distribution of total time spent on reactive deployment for specific incident types within broader incident categories. The overall rectangle represents all hours spent on reactive policing in 2019. The smaller, coloured rectangles represent the different categories, with their size indicating their share of total time. Crime-related incidents are primarily found in the categories *Safety* and *Public Order* and *Property Crime*. In total, less than half (45.9%) of all reactive deployment time in 2019 was spent on incidents in these categories. This highlights the incomplete nature of the image of the police as solely crime fighters: reactive deployment turns out to be far more diverse. Although suspicious situations, violent incidents, thefts, and burglaries demand substantial time, many hours are also spent dealing with disorder related calls such as nuisance, disputes, traffic incidents, and medical resuscitations.

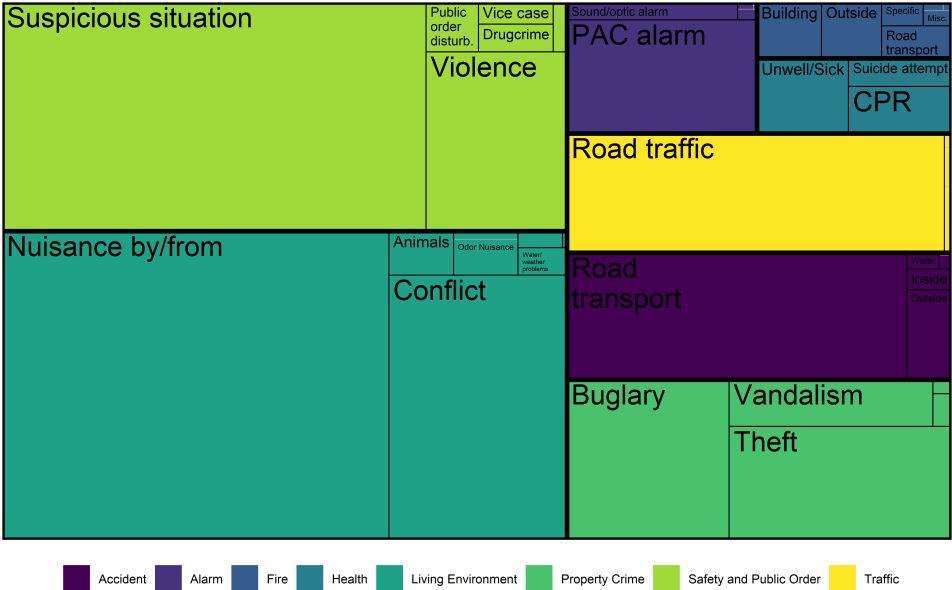


Figure 8.2: Relative proportion of total time spent on reactive police deployment in 2019

Figure 8.2 offers a national overview of the diverse workload on the streets. If a random officer were assigned to emergency duties at a random time and location, this figure suggests the mix of tasks they could expect to face.

8.2 DIFFERENCES BETWEEN FRONTLINE TEAMS

Behind the national picture lie significant differences. By linking incident locations to the jurisdictions of frontline teams, we can show (Figure 8.3) the extent to which reactive deployment differs across regional units and, within them, between frontline teams. Variation is greater within some regional units (e.g., Amsterdam and Midden-Nederland) than others (e.g., Noord-Nederland and Oost-Brabant).

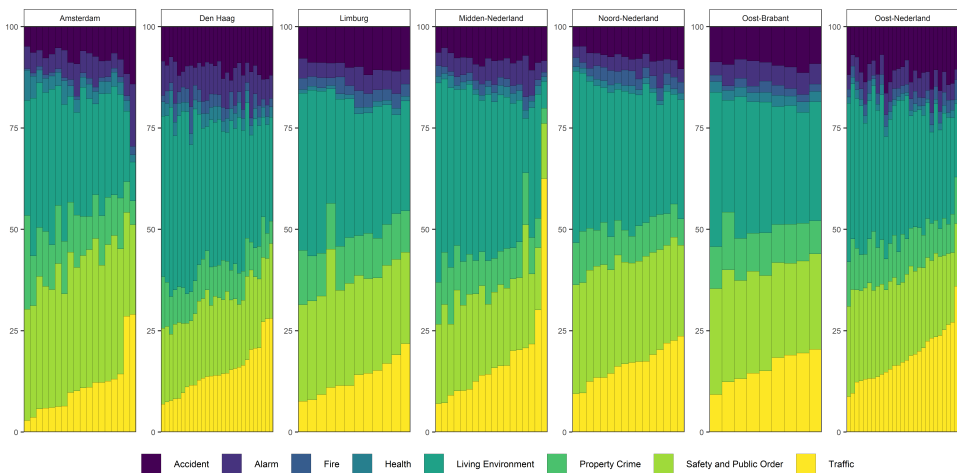


Figure 8.3: Differences in reactive deployment between frontline teams

8

Based on Figure 8.3, we could say that a random officer in Noord-Nederland unit is more likely to spend time on traffic-related incidents than one in the Amsterdam unit. However, the within-district variation is large. Some frontline teams in Amsterdam spend more time on traffic incidents than certain teams in Noord-Nederland. Thus, there is no clear distinction between officers' task profiles across regional units; within-unit differences are simply too substantial.

Figure 8.4 illustrates just how wide this variation among frontline teams can be. Particularly in the categories *Living Environment*, *Safety and Public Order*, and *Traffic*, there is a huge spread in the share of deployment time. Some frontline teams are clear outliers: in one *Midden-Nederland* team, almost three-quarters of time is spent on *Safety and Public Order*, while another team in the same unit spends over a third of their time on traffic-related incidents. While variation is least for *Alarms and Fires*, even here some teams spend less than 1%, and others up to 7% and 8% respectively. The share for *Health* and *Traffic* ranges from about 1% to over 14% and 39%, respectively. The ranges for *Property Crime*, *Living Environment*, *Accidents*, and *Safety and Public Order* are even broader. In short, the tasks faced by officers in emergency response vary greatly between frontline teams.

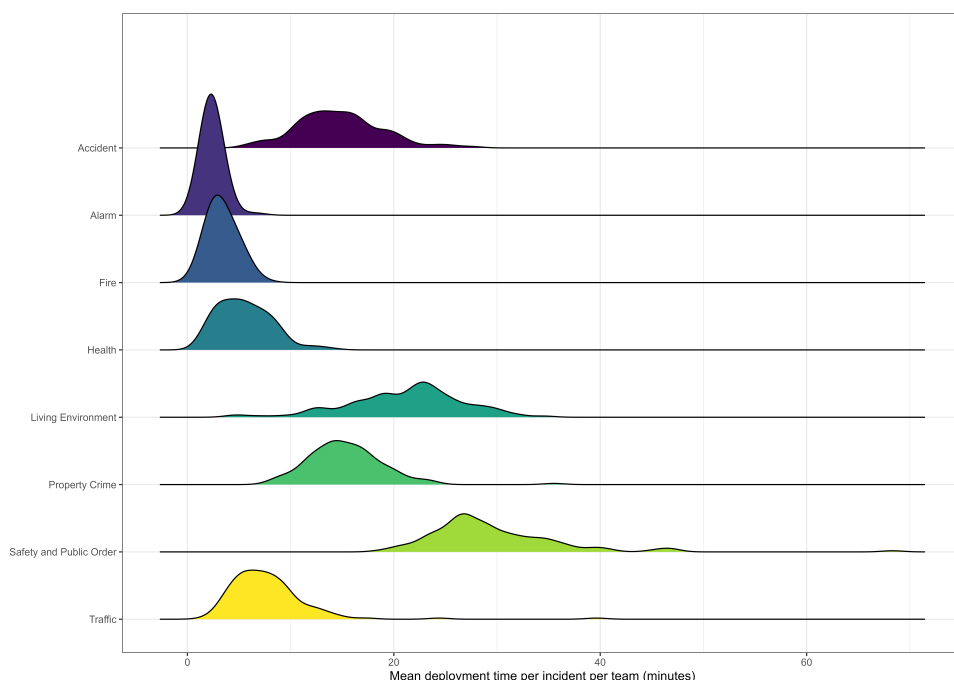


Figure 8.4: Variation in the share of incident categories across frontline teams

The graphs do not explain the causes of these differences. At this stage, we can only speculate. They may stem from differences in the socio-demographic makeup and degree of urbanisation in each team's jurisdiction. Cities have higher population densities, possibly leading to more interpersonal conflict. Additionally, large cities like Amsterdam, Rotterdam, The Hague, and Utrecht face considerable tourism, which may correlate with specific incident types. It is also possible that the police themselves influence the nature of reactive deployment. Given the still fragmented reality of Dutch policing, local differences in police culture, training, working style, or focus could shape how 112 calls are interpreted and handled. Operators at different dispatch centres may categorise similar incidents differently.

8

8.3 REFLECTIONS ON FINDINGS

We draw four key lessons from this analysis: recognition, training, internal collaboration, and local steering.

8.3.1 RECOGNITION

The diversity of incidents shown in Figure 8.2 will come as no surprise to frontline officers. However, given the strong variation between teams, the national picture may not reflect individual experiences. Still, the findings may reinforce the sense of complexity in police work and help communicate this complexity to others.

8.3.2 TRAINING

The variation in task profiles raises the question of whether training should be tailored accordingly. Teams that spend nearly half their time on *Safety and Public Order* will require different skills than those focused on *Traffic* issues. Since officers can never predict the next call, they must remain generalists to some extent. Yet the wide variety in tasks suggests a need for differentiated expertise and training.

8.3.3 INTERNAL COLLABORATION

Insights from analyses like this can help identify which teams have similar or divergent task profiles. This can inform the sharing of best practices, targeted recruitment, and mutual understanding of local issues within the national police organisation.

8.3.4 LOCAL STEERING

Over the past decade, the Dutch police have become increasingly centralised. Yet, when the National Police was established, it was acknowledged that policing is complex and context-dependent, requiring local decision-making. Our findings reinforce that view. Given the considerable variation across teams and districts, there is only limited room for top-down prioritisation.

IV

EVALUATING PROACTIVE POLICING IN THE NETHERLANDS

9

EVALUATING PLACE-BASED POLICING IN AMSTERDAM

A QUASI-EXPERIMENTAL STUDY OF CALLS-FOR-SERVICE LEVELS AND PUBLIC PERCEPTIONS USING A MICRO-SYNTHETIC CONTROL FRAMEWORK

Objectives Evaluate the impact of a place-based, community and problem-oriented policing intervention in Amsterdam, the Netherlands.

Methods This evaluation is guided by the EMMIE framework, and its methods and hypotheses are pre-registered on the OpenScienceFramework: <https://osf.io/7kyr3/>. The effect is estimated using a quasi-experimental, difference-in-difference design, comparing treatment areas (high-treatment, low-treatment, and catchment area) to synthetic control groups constructed from comparable neighbourhoods in Amsterdam. Outcome measures included the number of fast responses (<3 min), crime-, disorder-, traffic- and accident-related calls-for-service counts, and survey-based assessments of perceived neighbourhood disorder, fear of crime, victimisation, and citizen satisfaction with police and municipality.

Results In line with pre-registered hypotheses, we observed a substantial decrease in crime- ($RIRR = 0.77$) and disorder-related calls-for-service ($RIRR = 0.84$) in the intervention area and a small but significant increase in fear of crime in the low-treatment area ($SMD = -0.11$). Contrary to pre-registered hypotheses, we observed a substantial and statistically significant increase in perceived neighbourhood disorder ($SMD = -0.31$) and decrease in citizen satisfaction with police in the neighbourhood ($SMD = -0.42$) and citizen satisfaction with the municipality ($SMD = -0.47$) in the high-treatment area. Over all treatment areas (referred to a full-treatment: high, low and catchment combined), the intervention yielded a small but significant increase in fear of crime ($SMD = -0.07$) compared to synthetic controls.

Conclusions This study yielded mixed results. We observed a substantial decrease in crime- and disorder-related calls-for-service in the high-treatment area. However, observed over the full-treatment area, the reduction disappears, indicating crime displacement dynamics. Moreover, we observed worsened public perceptions in the high-treatment, low-treatment and the catchment area; highlighting a disconnect between improved local conditions and worsened public perceptions in the high-treatment area.

9.1 INTRODUCTION

Over the past decades, a substantial body of criminological research has emphasised that crime is not uniformly distributed across urban landscapes but instead clusters in specific high-activity areas, or ‘hot spots’ (Braga et al., 2010; P. L. Brantingham & Brantingham, 2017; E. R. Groff et al., 2010; Sherman & Weisburd, 1995; D. Weisburd, 2015). This knowledge has driven the development of focused policing strategies, such as hot spots policing, which concentrate police resources and activities within these small geographic areas to reduce crime more effectively (Sherman & Weisburd, 1995). Evidence suggests that this strategy can yield meaningful crime prevention benefits, as meta-analysis of hot spot policing evaluations has shown reductions for various crime types (Braga, Turchan, Papachristos & Hureau, 2019). However, these results are also not uniform, as the same meta-analysis also includes numerous evaluations that show null and even negative effects for such interventions. Moreover, the extent to which these results generalise across different settings and policing cultures, what the driving mechanisms are and what moderators influence whether the results are positive or negative, remain important questions.

Most research on hot spots policing interventions originates from large cities in the USA and (to a lesser extent) UK, with limited evidence available from (Western) European cities, which differ significantly in terms of urban design, crime rates, policing strategies and practices, and legal frameworks. This study helps address this gap by evaluating the impact of a place-based, intensified police presence intervention in Amsterdam, the Netherlands. The intervention implemented a combination of community-oriented and problem-oriented policing strategies, focusing on a segment of a kilometre-long contiguous shopping street (*Haarlemmerstraat/Haarlemmerdijk*) leading onto a public square (*Haarlemmerplein*) – places identified as micro-level hot spots of crime and disorder located in the ‘high-treatment’ area.

The intervention was designed, implemented, and evaluated together with frontline team (the lowest level of police jurisdiction in the Netherlands) *Centrum-Jordaan* of the Amsterdam Police. The frontline team had previously adopted a place-based focused approach (what they called ‘cluster-oriented policing’) by dividing the (already small, ± 2.5 square km) frontline team jurisdiction into six clusters. Officers working for the frontline team would be assigned to a cluster and, during discretionary time, focus their proactive policing efforts within their respective clusters. However, two years after the implementation, the adoption was considered a failure by local leadership responsible for the implementation. Illustrative of this is the fact that none of officers participating in the intervention could name the clusters they were supposed to focus their efforts at for the past two years.

The present intervention was designed to evaluate what the maximum potential effect of such a micro-level, place-based policing approach might be within the means of this particular frontline team. This was done by administering the maximum dosage of proactive policing the frontline team could support without endangering its emergency response capability within the other five clusters of their jurisdiction. Together with frontline team leadership, researchers selected the aforementioned crime and disorder hot spot within one of the six clusters and increased proactive policing in it for four months (March 23rd, 2024 until July 16th, 2024).

The intervention incorporated elements from both community and problem-oriented policing, and it was characterised by a large degree of street-level police officer discretion.

To prepare for the intervention, officers participated in a series of activities, including a training session on community and problem-oriented policing and ‘speed dating’ with partner organisations (e.g., public sanitation team, social safety team, municipal area representatives, municipal enforcement team, housing cooperation team) to facilitate collaboration. Officers also visited the hot spot areas, where they conducted both scheduled and impromptu interviews with residents and shopkeepers, gaining first-hand insights into local concerns. Following this, there was a group session to systematically identify and prioritise pressing issues in the intervention area. During this session, three primary issues emerged: (1) persistent disorder and crime linked to a small group of repeat offenders with addiction and homelessness issues, (2) traffic hazards posed by (electric) bikes, and (3) lacking police-public relations.

The effects of the intervention on these three issues are analysed based on eleven key outcome measures, namely *the number of fast responses (<3 min)*, *crime-, disorder-, traffic- and accidents-related calls-for-service (CFS) counts*; and standardised *perceived neighbourhood disorder, fear of crime*, (self-reported) *victimisation, citizen satisfaction with the police (local and general)*, and *citizen satisfaction with the municipality* survey scores.

This quasi-experimental evaluation employs the micro-synthetic control method (Robbins & Davenport, 2021), an approach for estimating treatment effects in non-randomised micro settings based on the synthetic control framework by Abadie et al. (2010). By constructing synthetic controls using comparable neighbourhoods in Amsterdam that did not receive the intervention, this study provides a robust counterfactual to estimate the intervention’s effects.

Hot spots policing and place-based interventions have often raised concerns regarding crime displacement – where offenders would simply relocate their activities to nearby areas (Eck, 1993; Reppetto, 1976). However, recent reviews and meta-analyses suggest that such interventions can also result in a diffusion of benefits, where crime reduction effects extend beyond the targeted areas (Guerette & Bowers, 2009; Telep et al., 2014). Our study investigates this phenomenon by the effect of the intervention in a so-called catchment area. Selection of an appropriate catchment area is a hard and much-debated problem (Ratcliffe et al., 2011). For this study, we selected as the catchment area all neighbourhoods directly contiguous to the three neighbourhoods in which our intervention street and public square lie. However, on its own, we judge this standard way of defining the catchment area to be insufficient to account for the possible negative consequences of this intervention, especially because we could consider police deployment to be a zero-sum game: all participating officers deployed in the treatment area are effectively diverted from the other neighbourhoods of the frontline team’s jurisdiction. In order for the intervention to be considered effective, the positive effect in the treatment area has to then outweigh the potential negative effects in the catchment area *and* the neighbourhoods from which participating officers were diverted, in this study referred to as the low-treatment area.

To systematically evaluate the outcomes of the intervention, we pre-registered a set of hypotheses addressing its impact across various treatment areas and outcome measures. These hypotheses reflect the intervention’s expected effects on crime, disorder, and community perceptions in the high-treatment area, while also accounting for potential (negative) externalities, such as resource diversion and crime displacement/diffusion of benefits, in other areas under study. By comparing high-, low-treatment, catchment, and control areas, the hypotheses allow for a comprehensive assessment of both the intended benefits and unintended consequences of the intervention.

9.2 HYPOTHESES

For each of the treatment areas, we pre-registered a separate hypothesis:

1. The intervention favours high-treatment area compared to synthetic control, meaning crime-, disorder-, accidents-, and traffic-related calls-for-service counts and standardised survey scores for perceived neighbourhood disorder, fear of crime, and victimisation will be lower and number of fast responses (< 3 min), and citizen satisfaction with police (local and general) and municipality will be higher in the high-treatment area compared to the synthetic control.
2. The intervention favours synthetic control compared to low-treatment area, meaning crime-, disorder-, accidents-, and traffic-related calls-for-service counts and standardised survey scores for perceived neighbourhood disorder, fear of crime, and victimisation will be higher and number of fast response (< 3 min), and citizen satisfaction with police (local and general) and municipality will be lower in the low-treatment area compared to synthetic control.
3. The intervention favours synthetic control compared to catchment area, meaning crime-, disorder-, accidents-, and traffic-related calls-for-service counts and standardised survey scores for perceived neighbourhood disorder, fear of crime, and victimisation will be higher and number of fast response (< 3 min), and citizen satisfaction with police (local and general) and municipality will be lower in the catchment area compared to synthetic control.
4. The intervention favours full-treatment area (high, low, catchment combined) compared to the synthetic control, meaning crime-, disorder-, accidents-, and traffic-related calls-for-service counts and standardised survey scores for perceived neighbourhood disorder, fear of crime, and victimisation will be lower and number of fast response (< 3 min), and citizen satisfaction with police (local and general) and municipality will be higher in the full-treatment area compared to synthetic control.

Each hypothesis should be interpreted as encompassing eleven distinct sub-hypotheses, corresponding to each specific outcome variable under investigation. This means that for every hypothesis, the falsification process will be applied individually to each of the eleven outcome variables.

This study is guided by the EMMIE framework – Effect, Mechanism, Moderator, Implementation, and Economic. Beyond examining the effects of the intervention, we also aim to understand the mechanisms underlying the results and the context in which the intervention was implemented. We also provide a detailed account of the implementation process and consider the broader implications of reallocating police resources. While the EMMIE framework shapes our approach, we do not conduct formal moderator analyses, as this would greatly increase the (already large) number of comparisons and the risk of Type I errors. Similarly, we do not perform a formal economic analysis. However, we do pay attention to the associated costs as well as the benefits of the intervention.

In summary, this study aims to contribute to the broader literature on hot spots policing by providing empirical evidence from an underrepresented study area. While most hot spot policing evaluation studies primarily focus on the effects on crime and disorder, this study

expands the scope by also examining the intervention's impact on public perceptions, offering a more comprehensive perspective. It also adds to the growing body of research utilising (micro) synthetic control methods to produce counterfactual estimates in quasi-experimental evaluations of place-based interventions. The hypotheses, methods, and analysis techniques of this quasi-experimental evaluation have been pre-registered on the OpenScienceFramework website (<https://osf.io/7kyr3>). The findings are expected to inform both policy and practice, offering insights into the effectiveness of intensified police presence in an urban hot spot outside the US context and contributing to the global discourse on effective crime prevention strategies.

9.3 BACKGROUND INFORMATION ON THE INTERVENTION

The intervention was developed following an internal evaluation of 'cluster-oriented policing' by the Centrum-Jordaan frontline team in Amsterdam. In 2022, team leadership introduced this approach hoping that a smaller geographic focus area would benefit street-level cops' contextual knowledge and therefore their proactive policing efficacy. Implementation was done by dividing their jurisdiction into six clusters and assigning each officer to a specific cluster. Officers were encouraged to focus their discretionary time on community engagement and problem-solving tasks within their assigned clusters, allowing for better familiarity with local communities and specific area concerns. Despite these aims, two years after the initial implementation, the team's leadership concluded that the framework had not been adopted by the frontline officers, resulting in no impact on cluster specific proactive policing efforts.

When considering the reimplementation of this place-based policing model, the frontline team leadership sought scientific insights into its potential maximum effectiveness within their jurisdiction. In response, this intervention and evaluation were collaboratively designed by team leadership and researchers to rigorously assess the upper limits of such an approach within the resource constraints of the local frontline team. Designed as a quasi-experiment, the intervention aimed to deploy the maximum feasible dosage of proactive police presence at selected hot spots within one of the clusters. This concentration of resources in a single high-treatment area was intended to amplify any potential effects, thereby increasing statistical power to detect meaningful changes across various outcomes. This reallocation incurred no additional staffing costs, as the intervention only leveraged existing resources.

The selection of the intervention neighbourhoods was informed by a rise in complaints from residents, prompting the municipality to initiate a series of neighbourhood consultation sessions at the end of 2023. These consultations revealed concerns about growing nuisances caused by individuals facing housing and drug-related issues, increased littering, and an increasingly hazardous traffic environment due to the rising use of electric bicycles and increasing discontent with municipality and the police. In response to these concerns, and as a direct outcome of the consultations, the Haarlemmerbuurt-Oost, Haarlemmerbuurt-West, and Planciusbuurt-Zuid neighbourhoods were chosen as the intervention area. Together, these neighbourhoods are referred to as the high-treatment area and encompass key hot spots, including Haarlemmerstraat, Haarlemmerdijk, and Haarlemmerplein, which were identified as micro-level zones of heightened crime and disorder.

The current intervention and its evaluation were modelled after a similar study conducted over 40 years ago in Amsterdam (Spickenheuer, 1983). The study then focused on enhancing police visibility and fostering community relations through foot patrols and public crime

prevention activities around the *Osdorpplein*, a prominent historical square similar to the Haarlemmerplein. Key lessons drawn from this earlier intervention study included having officers participate only part-time to maintain morale and connection with non-participating colleagues, conducting the intervention during the summer months to align with higher crime rates and pleasant conditions for foot patrols, and allowing officers autonomy in designing their own (micro) interventions. The Osdorp project ran for six months and gained widespread recognition, with over two-thirds of residents aware of the police team's increased presence. The findings were mixed and less rigorously tested, but were suggestive of improved community-police relations. These insights were adapted to the present-day context and directly informed the current intervention's design.

In the high-treatment area, dedicated proactive police presence was substantially elevated and a temporary police station was realised in a vacant shopping property for the duration of the intervention at a total rent cost of €6,000 for the four months. The temporary police station was, due to strict regulations around police stations, not open to the public, but could be used by officers for breaks, administrative and planning duties. During weekdays, a minimum of two police officers were present during both the morning (07:00—16:30) and evening (13:30—23:00) shifts, resulting in a three-hour overlap where at least four officers were on duty. On weekends, the presence was maintained with a minimum of two officers during the day shift (08:00—17:30) without additional morning or night shifts. These officers were relieved from routine call-for-service duties, except for high-priority emergency calls (priority 1) or those within the high-treatment area. This arrangement allowed them to focus on community policing and problem-solving tasks instead of on their activities under normal circumstances: mainly patrol and rapid response across the entire frontline team Centrum-Jordaan.

Two weeks before the start of the intervention, eighteen officers received a one-day training in community policing, with an emphasis on approaches to engaging with citizens, and in problem-oriented policing, specifically on the SARA (Scanning, Analysis, Response, and Assessment) model. The afternoon also featured a professional speed-date session in which officers individually met with around fifteen representatives from partner organisations (e.g., public sanitation team, social safety team, municipal area representatives, municipal enforcement team, housing cooperation team). The next day, all participating officers attended a preparatory kick-off day in which they visited the designated high-treatment neighbourhoods, specifically the hot spots: the public square (*Haarlemmerplein*) and the contiguous shopping street (*Haarlemmerdijk* and *Haarlemmerstraat*). During this visit, officers interviewed random individuals on the streets and made pre-arranged visits to residents and shopkeepers who shared their local experiences and concerns.

Afterwards, all participating officers convened for a group session where the insights from these interactions were gathered and systematically organised. This process led to the identification of three primary issues: traffic safety; crime and disorder caused by individuals with housing, mental health and addiction problems; and lacking public-police relations. Based on these themes, officers were organised into pairs and tasked with developing problem-oriented policing interventions. While they were encouraged to utilise the SARA model to guide their planning, officers had complete freedom to design their interventions as long as they adhered to specific boundary conditions: the interventions had to be feasible within the high-treatment area, conducted in uniform on foot or by bike (no motor vehicle

use), include concrete actions, and involve some sort of appropriate data collection.

Unfortunately, however, none of the pairs succeeded in producing a specific and actionable intervention plan. Consequently, it was decided that the first month of the intervention would focus on the community policing aspect: familiarising officers with the local context, residents, and prevalent issues. They would also begin with mapping individuals facing issues related to addiction, mental health, and housing to build a more comprehensive understanding of the community's needs.

To support these efforts, the police allocated €3000 in discretionary funding, which officers could use for resources they deemed beneficial for the intervention, such as: hiring a coffee stand and barista on the public square to encourage informal interactions with the public. During this first month, officers were advised to avoid repressive actions (such as fines and arrests) as much as possible. This initial focus on engagement was intended to prevent sudden enforcement efforts from displacing individuals with housing and addiction issues into neighbouring jurisdictions, which could compromise the intervention's potential for a lasting positive impact. Besides this advice, individual police officers were fully in-charge. This 'self-directing ability' of officers was a core principle of the cluster-oriented policing approach, empowering them to identify and address issues autonomously.

9.4 STUDY AREA

The study area comprises three contiguous small neighbourhoods in Amsterdam's central district: Haarlemmerbuurt-Oost, Haarlemmerbuurt-West, and Planciusbuurt-Zuid. These neighbourhoods are organised around a main thoroughfare – divided into Haarlemmerstraat and Haarlemmerdijk – that extends to the historic public square, Haarlemmerplein. This street functions as one of Amsterdam's primary non-motorised traffic routes, connecting the central district to the Westerpark and surrounding neighbourhoods in West Amsterdam. Haarlemmerstraat and Haarlemmerdijk are vibrant with storefronts along their one-kilometre stretch, hosting a variety of small businesses, boutiques, tradesmen, restaurants, cafés, and coffeeshops. Historically, many shop owners resided above their shops, though this pattern has declined in recent years. Originally a working-class area, these neighbourhoods have undergone significant gentrification, attracting new residents and businesses over recent decades. Table 9.1 provides some descriptive statistics of all areas involved in this evaluation.

Table 9.1: Characteristics of all treatment areas (source: CBS)

	High	Low	Catchment	Full	Control (unweighted)
Area size (km ²)	0.24	2.09	1.76	3.11	70.1
Population size	4055	31,210	17,390	39,405	542,310
Population density (p./km ²)	16,896	14,933	9881	12,670	7737
Social rent (%)	3.45	2.53	2.88	2.79	3.76
Married (%)	20.6	21.4	22.5	21.2	22.2
Non-European immigrants (%)	27.0	26.2	28.0	27.2	35.4
Age 0–25 (%)	17.9	19.0	19.0	18.9	24.5
Age 25–65 (%)	63.9	63.1	62.7	63.3	62.4
Age 65+ (%)	18.4	17.9	18.4	17.8	13.1

9.5 DATA

The evaluation of the four hypotheses is based on two primary data sources used to analyse the intervention's effects and potential mechanisms: calls-for-service (CFS) data (and computer-aided dispatch (CAD) data), and the National Crime Victimization Survey (NCVS), in the Amsterdam region called the 'Veiligheidsmonitor Amsterdam-Amstelland' (VMAA). Additionally, we carried out semi-structured interviews with community members and participating officers halfway through the experiment and at the end of the intervention period in order to obtain information about the actual implementation and day-to-day activities carried out in the intervention.

9.5.1 NCVS DATA ('VMAA')

The VMAA survey, conducted three times annually (in March, July, and November) by the Research and Statistics department of the Municipality of Amsterdam (in cooperation with CBS and Ipsos I&O), gathers information on residents' (above age 15) experiences with crime, perceptions of disorder, and satisfaction with both municipal and police services. The VMAA uses a stratified sampling approach to achieve a minimum respondent count per district and to compensate for low response rates in certain districts in previous waves of the survey. However, due to the VMAA's relatively small sample size at the neighbourhood-level – yielding approximately 10 respondents per neighbourhood per wave – its scope would have been too limited to obtain statistically robust estimates for smaller geographic areas. In order to increase statistical power of the survey, a shortened version (65 survey items) of the VMAA, referred to as the VN survey, was conducted in cooperation with the Research and Statistics department of the Municipality of Amsterdam within the three neighbourhoods comprising the intervention's high-treatment area and three out of the 24 neighbourhoods in the low-treatment areas. This survey was administered twice alongside the regular VMAA survey, in March and July (pre- and post-intervention).

For the post-intervention VN survey, a dozen additional questions were added which pertain specifically to the intervention. A total of 1602 households were invited to the VN survey's baseline measure, yielding 257 responses (17% response rate), and a similar number, 1620 households, were invited for the post-intervention measure, yielding 265 responses (17%). Among the baseline respondents, 120 opted in for the post measure (July), which accounted for 66 of the post-intervention sample size. It should be noted here, that besides the shorter form of the VN survey, also the invitation letter (or email) was different (see OSF repository for copies of these invitation letters).

CFS AND CAD DATA ('GMS')

All CFS and CAD data was obtained for the Police Unit Amsterdam (which encapsulated the inner-city and various suburbs) from 01-2017 until 10-2024. The data contained all calls-for-services to the emergency phone number ('112') and the potential subsequent dispatches (and response times) to these calls.

SEMI-STRUCTURED INTERVIEW WITH PUBLIC AND PARTICIPATING POLICE OFFICERS

Semi-structured interviews provide qualitative insights into the intervention's day-to-day activities and participant perspectives. Conducted halfway through the intervention and at the end, these interviews engaged both community members and police officers. Interviews

with officers provide information on the practical implementation of the intervention and capture any adjustments or challenges faced in maintaining intervention fidelity. Additionally, interviews with shopkeepers and residents shed light on the public experience of the implementation.

9.5.2 VARIABLES

The study evaluates eleven outcome variables derived from the call-for-service (CFS) and survey data, chosen to capture the intervention’s effects on the primary issues identified in the hot spot: traffic safety; crime and disorder caused by individuals with housing, mental health and addiction problems; and lacking public-police relations.

Calls-for-service (CFS) counts are employed to assess the effect on the former two issues. This evaluation thus focuses on a subset of CFS classifications, specifically excluding the classifications: alarm-, health-, and fire-related calls. Among the included classifications – crime, disorder, accident, and traffic – CFS counts are not uniformly distributed. Crime-related calls constitute the majority at 50%, followed by disorder-related calls at 36%, while accident- and traffic-related calls each account for 7%. These proportions highlight the differing frequencies of the issues addressed in this evaluation. As such, the emphasis is placed on crime and disorder as key areas of concern.

Survey-based outcome measures are employed to assess the interventions effect on public-police relations. Survey-based outcome measures include *perceived neighbourhood disorder*, *fear of crime*, *victimisation*, *citizen satisfaction with the police (local and general)*, and *citizen satisfaction with the municipality*.

Some survey items are transformed for intuitive interpretation: positive values reflect improvements in treatment areas compared to the synthetic control, while negative values signify deterioration compared to the synthetic control. This consistent coding across survey outcome variables simplifies comparisons and allows for a clear, unified approach to interpreting the intervention’s effects.

In Table 9.2, the descriptive statistics (means and standard deviations) for the calls-for-service counts per month and the number of fast responses per month are presented for the (unweighted) control group at the post-intervention timepoint (meaning august 2024). Due to its large size, the descriptive statistics table for the survey data is presented in Appendix section D.1. It presents the survey item, question, answers categories, the Likert scale range, and the mean and standard deviation from the unweighted control group at the post-intervention time point. This table also provides insight into which survey items were reverse-coded in order to align all variable in direction. The reported standard deviations are used to calculate standardised mean differences.

Table 9.2: Descriptive statistics for calls-for-service and fast responses per month per neighbourhood (control group at post measurement timepoint)

Category	Mean (% of total)	SD
Crime	11.48 (50%)	10.98
Disorder	8.32 (36%)	8.30
Accident	1.52 (7%)	1.89
Traffic	1.52 (7%)	2.79
Fast response count	1.23	1.92

9.6 METHODOLOGY

This methodology section is based on the pre-registration document which was submitted before accessing any of the data. The pre-registration also contains data simulation and data analysis R scripts which were closely adhered to – any deviations from the pre-registration are discussed in Appendix section D.2.

9.6.1 DIFFERENCE-IN-DIFFERENCE SETUP

This study employs a quasi-experimental design using repeated measures neighbourhood-level data from the three-times-a-year ‘Veiligheidsmonitor’ (VMAA) crime and victimisation survey, alongside monthly count data from fast responses and crime-, disorder-, traffic-, and accidents-related CFS/CAD data. In order to estimate the effect of the intervention, we employ a difference-in-difference framework. Meaning, we compare the change in the outcome variables in the treatment areas before and after the intervention with the change in the outcome variables before and after the intervention in a control area. In order to obtain a representative control area, we apply a (micro) synthetic control framework, which will be detailed later in this methodology section.

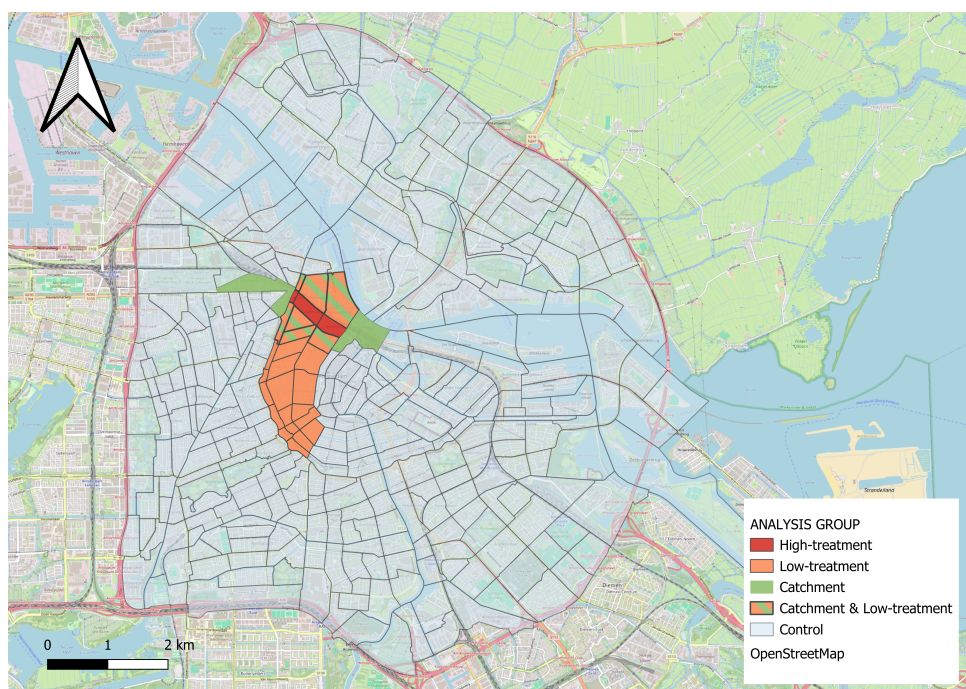


Figure 9.1: Neighbourhoods in Amsterdam categorised as treatment and control areas

We distinguish four treatment areas: high-treatment, low-treatment, catchment, and full-treatment (which constitutes the high, low and catchment area). Figure 9.1 portrays the neighbourhoods that make up the four treatment areas and the neighbourhoods in light blue which are weighted in the synthetic control approach to constitute a representative control

area. The high-treatment area consists of the three neighbourhoods in which the selected crime and disorder hot spots were located: the public square *Haarlemmerplein* 0.0043 km² and the six contiguous street segments, three on *Haarlemmerdijk* (approximately 3 × 150 m) and three on *Haarlemmerstraat* (approximately 3 × 100 m).

This is consistent with the size of hot spots typically studied in the (US) literature. For example, Braga, Turchan, Papachristos and Hureau (2019) note that most hot spots policing studies focus on street segments, intersections, or similarly small units, often under 0.01 km². Thus, while Dutch neighbourhood units may be larger, the actual intervention area in this study falls well within the common spatial scale of US-based place-based policing initiatives.

The low-treatment area consists of the all other neighbourhoods in the Amsterdam Police frontline team Centrum-Jordaan, as they will – due to reallocation of resources to the high-treatment area – receive less proactive police presence as a result of the intervention. The catchment area consists of all the neighbourhoods which are adjacent to the high-treatment area. All neighbourhoods within the inner-city of Amsterdam, here defined as within the A10 highway ring, are eligible to be included in the synthetic control group.

9.6.2 DIFFERENCE-IN-DIFFERENCE ANALYSIS USING SYNTHETIC CONTROLS

The synthetic control method (SCM) is a quantitative technique designed to create a synthetic control unit as a weighted combination of untreated units, representing an estimated counterfactual for treatment groups had they not received the intervention (Abadie et al., 2010). This approach is particularly advantageous when the intervention is applied to only a single or a limited number of units, making traditional randomised controls or other comparison methods infeasible.

To understand SCM, let Y_{ijt} represent the value for outcome i in (neighbourhood) unit j at time t , where $Y_{ijt}(1)$ and $Y_{ijt}(0)$ denote the potential outcomes under treatment and control, respectively. The causal effect α_{ijt} for a treated unit at a specific time point is given by the following:

$$\alpha_{ijt} = Y_{ijt}(1) - Y_{ijt}(0) \quad (9.1)$$

Since $Y_{ijt}(0)$ is unobservable, SCM approximates this by constructing a synthetic unit that mirrors the pre-treatment characteristics and trends of the treated unit as closely as possible. For SCM to work effectively, weights w_j , with j indexing units in the donor pool, are assigned such that the weighted average of outcomes and covariates in the donor pool resembles the treatment unit before the intervention. This is expressed formally as follows:

$$\sum_{j=1}^{J_0} w_j Y_{ijt} \approx Y_{ijt} \quad \text{for } t \leq T_0 \quad (9.2)$$

Where J_0 represents the number of untreated units in the donor pool, and T_0 is the last pre-treatment time period.

Micro Synthetic Control Method (MSCM) adapts SCM to handle multiple treated units and micro-level data (Robbins et al., 2017). In this evaluation, MSCM is implemented using the *microsynth* package in R (Robbins & Davenport, 2021). For more technical details on the weighting of control units, we refer to the documentation of this R package.

The analysis employs a *microsynth* model for the combined calls-for-service (CFS) data (crime, disorder, accident, traffic), constructing a single synthetic control group for comparison to each treatment area (high, low, catch, full). For *fast response count*, to improve accuracy, a separate synthetic control is constructed for the comparison to each treatment area: high, low, catch, full. In all cases, the synthetic control constructed by matching on 84 months of pre-intervention CFS count data.

For survey data, separate synthetic controls are constructed for each survey item, as a combined model was found to be insufficiently accurate – meaning constructed pre-trends of treatment and the constructed synthetic control deviated substantially based on visual inspection. In constructing each synthetic control, neighbourhood-level means for age, sex, and educational level are also employed in matching using the ‘match.out.min’ parameter, and the population-level rate of non-western migration background is included as a matching covariate in the ‘match.covar.min’ parameter. In all cases, the synthetic control constructed was matched based on 4 waves of pre-intervention survey data.

In both analyses, the constructed synthetic control, in sum, consists of an equal amount of control neighbourhoods as there are neighbourhoods in the specific treatment area: high (3), low (24), catch (10), full (27).

9.6.3 EXPERIMENTAL VARIABLE: POLICE PRESENCE

The core experimental variable in this study is the intensity of proactive police presence, although no direct data on actual officer presence (such as GPS tracking, patrol logs, or deployment audits) were available for the pre- or post-intervention periods. However, as the intervention was designed to employ the full proactive policing capacity of the Centrum-Jordaan frontline team – ordinarily spread across the entire jurisdiction – within the high-treatment area, we can estimate the increased levels of police presence.

This high-treatment area comprises approximately one-tenth of the total jurisdiction by surface area (0.24 km² out of 2.33 km²). Given this reallocation, and assuming proactive capacity was previously distributed more or less evenly across space, we conservatively estimate that the intervention led to a four- to fivefold increase in proactive police presence per square kilometre within the high-treatment area.

This structural reallocation serves as the treatment input of the intervention and is the presumed mechanism by which the observed effects on calls-for-service patterns and public perceptions may occur.

9.6.4 OUTCOME METRICS FOR SURVEY DATA AND CALLS-FOR-SERVICE

To evaluate the intervention’s impact, we analyse both calls-for-service (CFS) and survey-based data, transforming raw outputs into interpretable outcome metrics.

For the survey data, the *microsynth* function provides the sum of each outcome variable across all neighbourhoods within both the treatment and synthetic control groups. We calculate the relative mean difference by subtracting the synthetic control’s cumulative value from the treatment groups cumulative value, then dividing by the total number of neighbourhoods in the combined groups. This measure is standardised by dividing it by the standard deviation of the unweighted control groups at the post-treatment timepoint, measure commonly known as Glass’s D (Glass, 1976).

Construct-level measures – such as fear of crime, perceived neighbourhood disorder, self-reported victimisation, satisfaction with the municipality, and satisfaction with the police – are derived from individual survey items using a meta-analytical random-effects model. This approach reduces dimensionality and accounts for item-level variability, ensuring robustness. For the survey items, confidence intervals are calculated based on the pooled sample sizes of the treatment and control groups per survey item.

For the CFS data, we employ the relative incident rate ratio (RIRR) to assess the proportional change in incident rates between the treatment and control areas (Wilson, 2022). The RIRR, conceptually analogous to a difference-in-differences effect size, compares the relative changes in rates rather than raw counts to control for potential differences in sampling frames (e.g., time period, population size, geographic area). The RIRR is constructed by first calculating the incident rate ratio (IRR) for both the treatment and control groups. The IRR is the ratio of post-intervention to pre-intervention rates, defined as the count of incidents (x) divided by the units of the sampling frame (t). Specifically, the IRR for each group is calculated as follows:

$$IRR = \frac{x_1/t_1}{x_0/t_0} \quad (9.3)$$

where x_0 and x_1 represent pre- and post-intervention counts, and t_0 and t_1 denote the respective sampling frames. However, as we assume equal sampling frames for t_0 and t_1 , these cancel out. The RIRR is then computed as the ratio of the IRRs for the treatment and control groups:

$$RIRR = \frac{IRR_{\text{treatment}}}{IRR_{\text{control}}} \quad (9.4)$$

The null hypothesis asserts that this ratio equals 1, indicating no relative change between groups. Furthermore, permutation test p-values and confidence intervals are calculated to provide robustness checks for the observed RIRRs. For both survey and calls-for-service outcomes, pre-intervention values are taken at the end of March, and post-intervention values are assessed at the end of July, resulting in a consistent four-month evaluation period.

9.6.5 INFERENCE CRITERIA

Given that all hypotheses are directional, one-tailed tests are applied. Inferences are made at the construct level, covering *perceived neighbourhood disorder*, *fear of crime*, *victimisation*, *citizen satisfaction with the municipality*, and *citizen satisfaction with police (local and general)*. Each hypothesis is tested across these six constructs, resulting in 24 comparisons. To address the resulting *multiple comparisons problem*, where conducting many tests increases the risk of obtaining significant results by chance (Type I error). To reduce this risk, we set a stricter alpha level (0.01) for comparisons in lower-priority treatment areas while maintaining the standard 0.05 level for the high-treatment area, to also reduce Type II errors.

For our calls-for-service outcome variables, inferential statistics based on sampling assumptions are not applicable because these are population data rather than a sample (see Verlaan & Langton, 2024). For this reason, we used placebo tests as pseudo-experiments. These permutation tests involve reassigning intervention labels to untreated areas and calculating the expected range of effect sizes under the null hypothesis (see Robbins &

Davenport, 2021). The resulting permutation p-value is derived as the percentile rank of the observed effect within the distribution of placebo effects. Specifically, for the high-treatment area, the p-value reflects the proportion of placebo effects with a lower relative incident rate ratio than the observed RIRR. For other areas, it reflects the proportion of placebo effects with a higher RIRR.

Unlike traditional p-values, permutation p-values do not evaluate whether the population effect is precisely zero. Instead, they indicate whether the observed effect deviates significantly from the range of natural variability captured by the placebo distribution. This approach provides a robust benchmark for determining whether the observed effects are distinct from random chance in untreated areas. Consequently, effects are described as statistically distinct from placebo distributions rather than statistically significant, better aligning with the interpretive framework of permutation testing.

9.6.6 ROBUSTNESS CHECK: SURVEY MODE EFFECT ANALYSIS

To assess potential survey mode effects, we employed the Wilcoxon rank-sum test to compare each survey item between the VMAA and VN groups. To address the issue of multiple comparisons, we adjusted p-values using the Benjamini-Hochberg (BH) method, considering results with adjusted p-values below 0.05 as statistically significant. The analysis revealed numerous significant differences, with responses from the VN survey uniformly being more negative across items than responses from the VMAA survey.

In regression terms, this pattern indicates a significant intercept difference between the same items across the two different surveys. The key concern, however, is whether this difference introduces bias in the interaction term: the difference-in-difference (DiD) estimator. Since the VN survey was conducted in both the pre- and post-intervention waves and showed a consistent negative bias in both, the impact on the DiD estimates might be minimal or even non-significant. To verify this, we estimated linear (or logistic, for binary outcome variables) regression models with an interaction term between survey wave (pre/post) and survey mode (VMAA/VN). Any variable with a statistically significant interaction effect would have been excluded from further analysis. However, because we observed no such differences, we retained all data from both the VMAA and VN surveys for all variables in the synthetic control analysis.

9.7 RESULTS

9.7.1 EFFECTS OF THE INTERVENTION ON CALLS-FOR-SERVICE COUNTS

Table 9.3 summarises the findings of the micro synthetic control analysis for the fast response dispatch count and each calls-for-service category and treatment group. The results, expressed as relative incident rate ratios (RIRR), provide insight into the intervention’s impact.

Table 9.3: Calls-for-service count results, measured after intervention (RIRR with permutation-derived p-values)

Variable	High RIRR (permval)	Low RIRR (permval)	Catch RIRR (permval)	Full RIRR (permval)
Crime	0.77 (p = .032)	1.11 (p = .778)	1.02 (p = .284)	1.00 (p = .162)
Disorder	0.84 (p = .085)	1.23 (p = .492)	1.02 (p = .136)	1.10 (p = .138)
Accident	1.02 (p = .419)	1.02 (p = .994)	1.17 (p = .732)	1.10 (p = .992)
Traffic	1.40 (p = .605)	1.12 (p = .770)	0.98 (p = .258)	1.05 (p = .342)
Fast responses	1.63 (p = .264)	0.99 (p = .780)	1.01 (p = .792)	1.06 (p = .540)

In the high-treatment area, the intervention resulted in a statistically substantial reduction in crime-related calls-for-service, with a relative incident rate ratio (RIRR) of 0.77 (p = 0.032). This indicates that the incident rate ratio of crime-related calls was 77% of that of the synthetic control group, an effect unlikely to have occurred by chance alone, as shown by the permutation p-value. The effects on other call types were either less pronounced or much more likely to be within natural variation (disorder-related calls decreased: RIRR = .84, p = .085; traffic-related calls increased: RIRR = 1.40, p = .605; accident-related calls remained unchanged: RIRR = 1.02, p = .419, and the number of fast responses increased: RIRR = 1.63, p = .264).

9.7.2 EFFECTS OF THE INTERVENTION AS MEASURED BY THE NCVS

The results of this evaluation highlight several significant findings across the different treatment areas. The impact of the intensified police presence and community and problem-oriented policing strategies was assessed using standardised mean differences (Glass’s Glass’s *D*) to quantify changes in public perception. Table 9.4 presents a summary of the intervention results found in Appendix section D.3. Statistically significant findings are given an asterisk (*).

Contrary to the pre-registered hypothesis, the high-treatment area experienced a significant and substantial increase in *perceived neighbourhood disorder* (–.31 [95% CI –.49, –.13]), and decrease in *citizen satisfaction with the police in general* (–.42 [95% CI –.75, –.09]) and *citizen satisfaction with the municipality* (–.47 [95% CI –.76, –.18]). These findings suggest that while the increased police presence aimed to foster trust and satisfaction, it may have inadvertently led to negative perceptions. Potential factors for these counterintuitive results could include perceived over-policing, lack of meaningful engagement, or unmet community expectations or rapid local deterioration of the situation. These elements will be explored further in the Mechanisms section.

The evaluation also observed a non-significant reduction in fear of crime (SMD –.10 [95% CI –.24, .05]), victimisation (SMD .11 [95% CI –.19, .41]) and satisfaction with the police in the neighbourhood (SMD –.14 [95% CI –.46, .18]). The fact these findings are

non-significant can either be due to a lack of an actual effect, or because of the low statistical power in the high-treatment area, due a small sample sizes in the synthetic control groups.

Table 9.4: Results for survey-based constructs per treatment area

Variable	High [SMD $\pm$ 95% CI]	Low [SMD $\pm$ 99% CI]	Catch [SMD $\pm$ 99% CI]	Full [SMD $\pm$ 99% CI]
Perceived neigh. disorder	-0.31* [-0.49; -0.13]	-0.05 [-0.17; 0.07]	-0.03 [-0.13; 0.07]	-0.11 [-0.23; 0.01]
Fear of crime	-0.10 [-0.24; 0.05]	-0.11** [-0.19; -0.02]	-0.05 [-0.17; 0.06]	-0.07** [-0.13; -0.01]
Victimisation (self-reported)	0.11 [-0.19; 0.41]	0.07 [-0.10; 0.23]	0.05 [-0.11; 0.22]	0.03 [-0.13; 0.18]
Satisfaction with Police (general)	-0.42* [-0.75; -0.09]	-0.06 [-0.24; 0.11]	-0.03 [-0.23; 0.17]	-0.11 [-0.24; 0.03]
Satisfaction with Police (neigh.)	-0.14 [-0.46; 0.18]	-0.10 [-0.22; 0.03]	0.01 [-0.13; 0.16]	-0.10 [-0.20; 0.00]
Satisfaction with Municipality	-0.47* [-0.76; -0.18]	-0.03 [-0.16; 0.10]	-0.00 [-0.21; 0.20]	-0.01 [-0.11; 0.09]

The low-treatment area demonstrated a significant, albeit small, increase in fear of crime, with an SMD of -0.11 [99% CI $-0.19, -0.02$]. In the catchment area, none of the observed variables yield a standardised mean difference statistically (α -level = .01) significantly different from zero. The full model – measuring the overall effect of the intervention on the high, low and catchment area combined, compared to a synthetic control area – found a small but significant negative effect on fear of crime. These significant findings from the low-treatment, catchment and full-treatment area all align with the pre-registered hypotheses. However, all other variables yield small and non-significant findings. Again, this could be due to a lack of an actual effect, or because of the low statistical power in identifying small effect sizes. It should be noted for example that we never observe statistically significant intervention effects for victimisation. This is mainly because it is so rare, that the analysis is too underpowered to detect any meaningful changes.

SUPPLEMENTARY ANALYSIS: PUBLIC AWARENESS AND JUDGEMENT OF THE INTERVENTION

The survey data indicate that the intervention was not widely noticed or strongly regarded by the public. Only a small proportion of respondents (8%) were aware of the additional police deployment as a result of the intervention. Similarly, awareness of the temporary police station at Haarlemmerdijk was minimal, with just 8% of respondents knowing about its existence. When asked about the impact of the increased police presence, responses were mixed. While 14% (strongly) agreed that it led to less crime and disorder, 22% of respondents were neutral, and 11% expressed disagreement. Notably, a significant proportion (47%) of respondents indicated that they did not know what the effect could be. These findings align with the general lack of strong opinions about the intervention, as 21% of respondents were neutral about the extra police presence overall, and only one-third viewed it as positive or very positive. These results suggest that the intervention lacked visibility and resonance with the community, with limited awareness and lukewarm opinions about its effectiveness in reducing crime and disorder. Full supplemental survey item results are presented in Appendix section D.4.

9.8 MECHANISMS

This paragraph explores three potential mechanisms for the unexpected outcomes observed in the high-treatment area, particularly the paradoxical combination of substantially reduced crime and disorder calls-for-service and a significant decrease in perceived neighbourhood disorder alongside declining citizen satisfaction with the local police and municipality. We consider: (1) the ‘history’ threat (Cook & Campbell, 1979), (2) adverse effects of the intervention itself, and (3) the influence of a dominant local narrative.

THE ‘HISTORY’ THREAT

The ‘history’ threat refers to external events or shifts in local conditions coinciding with the intervention period, which may influence outcome measures. If local conditions worsened during the intervention period while remaining stable or improving in the synthetic control group, they could account for the observed decline in citizen satisfaction.

Qualitative data from resident interviews initially seem to support this mechanism. Several residents reported perceiving worsening problems in their neighbourhood. For example, one respondent remarked, ‘I haven’t noticed an increase in police, but I have seen more drunk and homeless people and drug use. More aggression, antisocial behaviour, and fat bikes speeding around. This gives an increasingly unsafe feeling’. However, this interpretation conflicts with the calls-for-service (CFS) data, which showed a substantial reduction in crime- and marginally so in disorder-related calls-for-service, making this mechanism an unlikely explanation.

ADVERSE CONSEQUENCES OF THE INTERVENTION

The second mechanism considers that the intervention itself may have unintentionally caused dissatisfaction and/or decreased willingness to report crimes due to visible actions or perceived inactions by the police. While the intervention was designed to improve citizen perceptions through police visibility and foster positive community relations, this may have had the reverse effect.

However, given the limited awareness about the intervention (as laid out in the supplementary analysis), it is unlikely that a negative impact of the intervention itself would be responsible for the substantial decrease in perceptions or a sudden decreased willingness to report crimes.

DOMINANT LOCAL NARRATIVE

The third mechanism considers the role of dominant local narratives in shaping citizen perceptions, irrespective of objective changes in neighbourhood conditions. These narratives, shaped by community discourse or media, can overshadow objective improvements, such as the documented relative decline in calls-for-service. This dynamic was evident in interviews with local residents and shopkeepers. For example, one shopkeeper expressed frustration, stating that the police had ‘moved out of the communities as a result of centralisation’, leaving her unable to ‘walk into a police station to file a complaint or talk to an officer’. This is striking given the close proximity of a police station (800 meters away) and the regional Police Headquarters (less than 2 kilometres away) for this respondent’s shop. Such perceptions, seemingly detached from the actual presence of police resources, highlight the entrenched nature of (local) narratives. These narratives appear resistant to evidence of

improved conditions, such as the 23% reduction in crime-related calls-for-service, and may undermine efforts to positively influence public perceptions. This mechanism provides the most plausible explanation for the paradox of improved crime metrics alongside declining public satisfaction.

9.8.1 IMPLEMENTATION

From the outset of the intervention, there was noticeable friction regarding the emphasis on 'self-directed' police efforts. In the problem-oriented policing workshop, approximately half of the participants voiced scepticism about being responsible for analysing and defining issues, as well as designing and implementing appropriate police responses. Some officers argued that this preparatory work could have been completed by police leadership or the researchers involved, allowing them to focus solely on execution. During the group session, translating the insights and objectives collected during the site visit into actionable plans also proved challenging for the officers.

Despite these challenges, many aspects of the intervention's rollout were met with positive feedback. The speed-dating session with representatives from partner organisations was seen as informative, valuable, and enjoyable by both the officers and the representatives. Although the allocated time of 10 minutes per round was considered brief, participants actively exchanged contact information and planned to continue discussions beyond the session. Some partners were even invited to spend a day at the temporary police station within the intervention zone, fostering stronger working relationships. Additionally, visiting the hot spot, engaging with residents, passers-by, and shopkeepers, and learning about the neighbourhood, the community and the local issues were unanimously acknowledged as beneficial. The subsequent group session, which gathered and systematised these insights, was also well received.

One and a half month after the intervention began, interviews were conducted with all participating officers to assess their actions, the implementation process, and their experiences. Officers reported positive outcomes, including successful engagement with residents, data collection, and mapping of key individuals in the community, particularly those experiencing mental health issues, addiction, and homelessness. Local entrepreneurs appreciated the officers' gradual introductions and explanations of the planned activities for the coming months. However, the officers noted that this phase of engagement reached saturation relatively quickly. They realised that if everyone continued to approach local entrepreneurs individually, it would lead to redundancy and inefficiency.

Moreover, the officers also observed that the known repeat-offender group quickly became aware of the increased police presence and moved out of the area, which they stated (temporarily) reduced disturbances. This left the officers feeling that they had achieved their initial objectives and were uncertain about what to do next. Many struggled to create plans and felt there was insufficient guidance. Although the intervention's self-directed approach aimed for officers to organise and coordinate among themselves, this proved difficult. Communication, such as end-of-shift debriefings and updates, was inconsistent. While some officers used a dedicated Blackberry group to share updates and tasks, most did not, resulting in poor information transfer. Officers often resorted to patrolling without a clear plan, simply responding to whatever they encountered.

To address these challenges, a plan was developed based on the interviews. Leadership

figures were appointed from within the group to oversee the three main focus areas: traffic safety in the street and square, disturbances from the group of repeat offenders, and fostering better community relations. Two officers were assigned as theme leaders for each focus area, responsible for coordinating efforts and ensuring updates were communicated more effectively in the group. Each participating officer was assigned into these thematic sub-groups, so they knew their primary responsibilities. This top-down structure aimed to provide enough organisation for the officers to manage themselves effectively and maintain self-directed, coordinated actions.

The final evaluation, conducted through interviews with all participating officers, revealed that despite the mid-evaluation adjustments and the introduction of thematic leadership, officers continued to find self-directed work challenging. Few concrete, systematic approaches were established. If self-organisation took place, this was always with an ‘event’ as the intended outcome. For example, one pair of officers spent approximately four months developing a plan they referred to as a ‘one-day service corridor’ which aimed to bring all chain partners together to engage with members of the group of repeat offenders in a coordinated way. However, despite significant time and effort, this plan never materialised into an actual event.

The events that were realised are as follows: an exposition stand educating the community on recognising signs of drug labs and in-house cannabis farms and a stand focussing on burglary prevention that highlighted effective lock types and security measures. Both activities took place on the middle of the square during the local farmers market and were largely arranged by Sergeant Van Jole, who was overseeing the intervention. Additionally, two senior officers took on a project they called ‘New Perspectives’, which involved inspecting local businesses to identify potential money-laundering activities. Though innovative, none of these events aligned directly with the three main focus areas of the intervention. Moreover, despite frequent encouragements and discussions on how to utilise the allocated funds, the discretionary budget of €3000 ultimately went unused, underscoring a reluctance or lack of initiative within the team to creatively deploy available resources.

Officers expressed that they still struggled with finding proactive tasks and relied heavily on undirected patrolling. Efforts to enhance traffic safety were minimal, with only a few days of traffic checks conducted and no roller test bench obtained to address the nuisance caused by electric bikes, despite initial plans. A systematic approach to the repeat-offender group also fell short. Although the officers initially mapped out key individuals and met relevant chain partners during the speed-dating session – including a municipal representative for repeat offenders – follow-up communication was lacking or non-existent. Despite the expressed willingness of these partners to be involved, the group did not maintain contact or build collaborative efforts.

From the outset of the intervention, and increasingly as it progressed, the project was met with negative perceptions from the broader police department. Non-participating officers expressed concerns that the high-treatment area received disproportionately high levels of police presence, which they felt could not be justified given the demand for police resources elsewhere. Many perceived that they were left to ‘shoulder the burden’ in other parts of the frontline team, which became relatively understaffed as a result. This sentiment was frequently communicated to project leadership and participating officers, leading to noticeable morale challenges within the group and to motivation waning over time.

Surprisingly, during a final public group evaluation in the presence of police leadership, officers' collective feedback was quite positive. This contrasted with the more critical views expressed when speaking to individual officers in informal talks and interviews. A sense of group dynamics seemed to influence the positive presentation of results, masking underlying feelings of discomfort and the perception of a lack of a clear vision. Leadership acknowledged that providing further top-down direction would be challenging and unsustainable in a scenario intended to be self-directed. Ultimately, it seemed this project reached its practical limits.

9.9 CONCLUSIONS

This study evaluated the effects of a place-based, community, and problem-oriented policing intervention in Amsterdam using a quasi-experimental design and the micro synthetic control method. The intervention yielded mixed results. While the high-treatment area experienced a significant reduction in crime-related calls-for-service (accounting for 50% of all calls-for-service), public perceptions deteriorated, with significant decreases in satisfaction with local police and municipal services, alongside a negative impact on perceived neighbourhood disorder.

In contrast, the low-treatment area, which experienced reduced police presence due to resource diversion to the high-treatment area, showed a small but significant negative effect on fear of crime. It also exhibited increases in all types of calls-for-service, though none were statistically distinguishable from natural variation. The catchment area showed no statistically significant or substantial changes in any outcome measures.

The overall model, which included the high-, low-, and catchment-treatment areas, only revealed a small but significant negative effect on fear of crime. It also showed minor increases in all types of calls-for-service and the number of fast responses, though all remained within natural variation. Table 9.5 summarises the pre-registered hypotheses and their outcomes in terms of being supported, not supported, or inconclusive.

Table 9.5: Preregistered hypotheses (positive/negative effect of intervention) per treatment area compared to synthetic control: supported, not supported or inconclusive

	High (positive effect)	Low (negative effect)	Catch (negative effect)	Full (positive effect)
Perceived neigh. disorder	Not supported	Inconclusive	Inconclusive	Inconclusive
Fear of crime	Inconclusive	Supported	Inconclusive	Not supported
Victimisation	Inconclusive	Inconclusive	Inconclusive	Inconclusive
Satisf. with Police (gen.)	Not supported	Inconclusive	Inconclusive	Inconclusive
Satisf. with Police (neigh.)	Inconclusive	Inconclusive	Inconclusive	Inconclusive
Satisf. with Municipality	Not supported	Inconclusive	Inconclusive	Inconclusive
Crime	Supported	Inconclusive	Inconclusive	Inconclusive
Traffic	Inconclusive	Inconclusive	Inconclusive	Inconclusive
Disorder	Inconclusive	Inconclusive	Inconclusive	Inconclusive
Accident	Inconclusive	Inconclusive	Inconclusive	Inconclusive
Fast responses	Inconclusive	Inconclusive	Inconclusive	Inconclusive

9.9.1 LIMITATIONS

This study faces several limitations that affect the strength of its conclusions, stemming from the research design, the evaluation problem, and the specific challenges encountered in constructing synthetic control groups.

First, the quasi-experimental design used in this study inherently limits internal validity. While it employs synthetic control methods to approximate counterfactual scenarios, the design ranks at Level 3 on the Maryland Scientific Methods Scale, falling short of the internal validity offered by randomised experiments, which achieve Level 5. Despite efforts to enhance transparency through rigorous preregistration and a priori hypothesis formulation, the inability to fully eliminate confounding factors remains a key limitation.

Second, this study exemplifies the ‘evaluation problem’, as described by Sherman (1986, p. 362). Crime events, particularly at the neighbourhood level, are inherently sparse, leading to high percentual variability in key outcome measures such as victimisation rates, traffic-related calls-for-service, and fast response counts. For instance, the intervention areas often recorded zero incidents in many months, making it difficult to identify true treatment effects amidst random fluctuations. Even large percentual changes, such as the observed 60% increase in fast response counts in the high-treatment area, could not be statistically distinguished from natural variation. This inherent sparsity complicates statistical testing, as small absolute differences can easily occur by chance, undermining confidence in the results.

Finally, the evaluation problem significantly impacts the construction of accurate synthetic control groups. The high-treatment area, composed of only three neighbourhoods, posed additional challenges for the synthetic control algorithm. The limited pool of neighbourhoods constrained the ability to generate an optimal match, further exacerbating the issue of high percentual variability. As a result, even if the synthetic controls had been perfectly constructed, the variability in the data would still have made it difficult to interpret observed differences as anything other than random noise.

These limitations collectively highlight the challenges inherent in evaluating policing interventions using neighbourhood-level data and quasi-experimental designs. While the study offers valuable insights, the findings should be interpreted with caution, recognising the constraints imposed by the methodological and contextual factors.

9.10 DISCUSSION

This conclusion section provides several key insights into the intervention and its outcomes. First, the intervention demonstrates that even a significant increase in police presence can largely go unnoticed by the public. Survey data revealed that the majority of residents were unaware of the additional police deployment or the temporary police station, underscoring the limitations of visibility alone in fostering public recognition or engagement.

Second, the effects of the intervention on traffic safety could not be reliably assessed due to the sparsity of traffic- and accident-related calls-for-service data. While the high-treatment area showed a 40% increase in traffic-related calls compared to the synthetic control, a similar or larger increase was observed in 39% of untreated placebo groups. This degree of natural variation makes it impossible to discern whether the intervention had any genuine impact on traffic safety.

Third, the intervention achieved a substantial reduction in crime and disorder. However, this localised effect was completely offset when the catchment area and the rest of the

jurisdiction were also considered. This nullification suggests a displacement of crime, likely caused by (1) the known group of repeat-offenders relocating their activities and (2) increased crime opportunities elsewhere due to reduced police presence. The lack of a comprehensive or durable strategy on the part of the participating officers to address the root causes of the repeat-offenders' behaviour – such as improved support, deterrence, or incapacitation – predictably led to a continuation of their activities in other areas.

Fourth, despite a measurable decrease in crime- and disorder-related calls-for-service within the high-treatment area, public perceptions of the neighbourhood, the police, and the municipality deteriorated. These findings align with concerns raised in the literature about the potential adverse effects of hot spots policing on citizen satisfaction (Rinehart Kochel, 2011; Rosenbaum, 2006). However, because only a small proportion of respondents reported being aware of the intervention, the observed changes in perception are unlikely to reflect direct reactions to police behaviour during the intervention itself. Instead, qualitative interview data suggest that a dominant and increasingly negative local narrative may have persisted independently of the intervention's progress in reducing crime and disorder. This points to a potential disconnect between the effectiveness of policing efforts and how they are perceived by the public. Rather than being driven by police performance, public sentiment may be shaped by broader contextual factors. This interpretation is consistent with findings from other recent (quasi-)experimental evaluations of hot spots policing, which observed statistically significant reductions in crime or disorder but no corresponding improvements in public perceptions (Ratcliffe et al., 2015; D. Weisburd et al., 2011). However, it contrasts with an earlier review of the literature by Zhao et al. (2002), which found that public satisfaction was strongly associated with the success of fear of crime reduction programmes. A key difference between the present intervention and some of the programs reviewed by Zhao et al. (2002) may lie in the absence of a public awareness campaign. In this study, such a campaign was deliberately omitted to isolate the effect of intensified police presence. Future research could explore the independent impact of communication strategies on perceptions of safety, satisfaction, and police legitimacy. For example, an experimental design might vary the intensity of a public awareness campaign about a standard hot spots intervention in both a subset of treatment and control neighbourhoods, to assess the potential placebo effect of such campaigns on public perceptions.

Fifth, from a cost-benefit perspective, the intervention yielded no discernible benefits beyond the high-treatment area and resulted in additional financial and resource costs. These included €6000 for the temporary police station's rent plus the extra costs (wages) associated with the organisation and management of the intervention.

Lastly, we propose a theoretical insight that emerged during this study: deterrent effects in increased policing interventions may arise not only from proactive mechanisms – such as heightened visibility and increased patrols – but also from reactive mechanisms, such as rapid response to incidents. Traditionally, deterrence in hot spot policing interventions has been understood as driven by proactive strategies that elevate the perceived risk of apprehension. However, our data reveal a 63% increase in fast response counts in the high-treatment area. Due to the limited power of this intervention – stemming from sparse data on fast response count data within treatment and synthetic control areas – this observed change in fast response counts cannot be distinguished from natural variation. Nonetheless, the most conservative interpretation of this finding is that it does not falsify the hypothesis that

reactive mechanisms contribute to changes in crime and disorder rates. Currently, the vast majority of studies on hot spot policing interventions fail to account for reactive elements. Therefore, future research should disentangle the respective contributions of proactive and reactive policing mechanisms. Addressing this gap would enhance our understanding of how policing interventions influence offender behaviour and improve policing strategies.

ACKNOWLEDGEMENTS

We extend our gratitude to the Amsterdam Police frontline team Centrum-Jordaan for their collaboration, particularly Sergeant Lars van Jole and the participating officers. We also thank Sara de Wilde, Anne Huijzer, Iris Regeer, Farah Makrani, and Jeroen Slot from the Amsterdam Municipality Statistics Department for their role in survey data collection and access. We thank research assistants Roos Hamelink, Chiara van den Engh, and Annemerijn Weijers for their help in organising (RH) and collecting interview data (CE&AW).

DATA AVAILABILITY

The National Crime Victimization Survey (VMAA) data used in this study is accessible from CBS Netherlands, subject to their access policies. The VN survey data and Calls-for-Service (CFS) data are not publicly available due to privacy and security restrictions but can be accessed on-site at the Secured Analytics Lab (SAL) at the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) in Amsterdam. Access to these datasets is subject to institutional data access policies and procedures. For further inquiries, please contact the corresponding author.

REFERENCES

- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>
- Bieck, W. H. (1977a). *Response Time Analysis: Executive summary* (tech. rep.). National Institute of Justice. Kansas City.
- Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2010). The Concentration and Stability of Gun Violence at Micro Places in Boston, 1980–2008. *Journal of Quantitative Criminology*, 26(1), 33–53. <https://doi.org/10.1007/s10940-009-9082-x>
- Braga, A. A., Turchan, B., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1046>
- Brantingham, P. L., & Brantingham, P. J. (2017). Environment, Routine, and Situation: Toward a Pattern Theory of Crime. *Routine Activity and Rational Choice*, 259–294. <https://doi.org/10.4324/9781315128788-12>
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin.
- Eck, J. E. (1993). The Threat of Crime Displacement. *Problem Solving Quarterly*.
- Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*, 5(10), 3. <https://doi.org/10.2307/1174772>
- Groff, E. R., Weisburd, D., & Yang, S.-M. (2010). Is it Important to Examine Crime Trends at a Local “Micro” Level?: A Longitudinal Analysis of Street to Street Variability in Crime Trajectories. *Journal of Quantitative Criminology*, 26(1), 7–32. <https://doi.org/10.1007/s10940-009-9081-y>
- Guerette, R. T., & Bowers, K. J. (2009). Assessing the extent of crime displacement and diffusion of benefits: A review of situational crime prevention evaluations. *Criminology*, 47(4), 1331–1368. <https://doi.org/10.1111/j.1745-9125.2009.00177.x>
- Ratcliffe, J. H., Groff, E. R., Sorg, E. T., & Haberman, C. P. (2015). Citizens' reactions to hot spots policing: Impacts on perceptions of crime, disorder, safety and police. *Journal of Experimental Criminology*, 11(3), 393–417. <https://doi.org/10.1007/s11292-015-9230-2>
- Ratcliffe, J. H., Taniguchi, T., Groff, E. R., & Wood, J. D. (2011). The Philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, 49(3), 795–831. <https://doi.org/10.1111/j.1745-9125.2011.00240.x>
- Reppetto, T. A. (1976). Crime Prevention and the Displacement Phenomenon. *Crime & Delinquency*, 22(2), 166–177. <https://doi.org/10.1177/001112877602200204>
- Rinehart Kochel, T. (2011). Constructing Hot Spots Policing: Unexamined Consequences for Disadvantaged Populations and for Police Legitimacy. *Criminal Justice Policy Review*, 22(3), 350–374. <https://doi.org/10.1177/0887403410376233>
- Robbins, M. W., & Davenport, S. (2021). microsynth: Synthetic Control Methods for Disaggregated and Micro-Level Data in R. *Journal of Statistical Software*, 97(2). <https://doi.org/10.18637/jss.v097.i02>

- Robbins, M. W., Saunders, J., & Kilmer, B. (2017). A Framework for Synthetic Control Methods With High-Dimensional, Micro-Level Data: Evaluating a Neighborhood-Specific Crime Intervention. *Journal of the American Statistical Association*, 112(517), 109–126. <https://doi.org/10.1080/01621459.2016.1213634>
- Rosenbaum, D. P. (2006, May). The limits of hot spots policing. In *Police innovation* (pp. 245–264). Cambridge University Press. <https://doi.org/10.1017/CBO9780511489334.013>
- Sherman, L. W. (1986). Policing Communities: What Works? *Crime and Justice*, 8, 343–386. <https://doi.org/10.1086/449127>
- Sherman, L. W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime “hot spots”: A randomized, controlled trial. *Justice Quarterly*, 12(4), 625–648. <https://doi.org/10.1080/07418829500096221>
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Spickenheuer, H. (1983). *Voetsurveillance en preventievoorzichting in Amsterdam-Osdrorp: een onderzoek naar de effecten op de criminaliteitsontwikkeling*. Staatsuitgeverij.
- Telep, C. W., Weisburd, D., Gill, C. E., Vitter, Z., & Teichman, D. (2014). Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, 10(4), 515–548. <https://doi.org/10.1007/s11292-014-9208-5>
- Verlaan, T., & Langton, S. (2024, April). On the Use of Inferential Statistics on Administrative Police Data. In *The crime data handbook* (pp. 197–210). Bristol University Press. <https://doi.org/10.51952/9781529232073.ch014>
- Weisburd, D. (2015). The Law of Crime Concentration and the Criminology of Place. *Criminology*, 53(2), 133–157. <https://doi.org/10.1111/1745-9125.12070>
- Weisburd, D., Hinkle, J. C., Famega, C., & Ready, J. (2011). The possible “backfire” effects of hot spots policing: an experimental assessment of impacts on legitimacy, fear and collective efficacy. *Journal of Experimental Criminology*, 7(4), 297–320. <https://doi.org/10.1007/s11292-011-9130-z>
- Wilson, D. B. (2022). The Relative Incident Rate Ratio Effect Size for Count-Based Impact Evaluations: When an Odds Ratio is Not an Odds Ratio. *Journal of Quantitative Criminology*, 38(2), 323–341. <https://doi.org/10.1007/s10940-021-09494-w>
- Zhao, J., Schneider, M., & Thurman, Q. (2002). The effect of police presence on public fear reduction and satisfaction: A review of the literature. *The Justice Professional*, 15(3), 273–299. <https://doi.org/10.1080/0888431021000049471>

D

APPENDIX: EVALUATING PLACE-BASED POLICING IN AMSTERDAM

D

D.1 DESCRIPTIVE STATISTICS OF SURVEY DATA

Table D.1: Descriptive statistics of survey items used in the analysis

Survey item	Survey question	Likert Scale Labels	Range	Mean	SD
Perceived change in disorder	Do you think your neighbourhood has improved, deteriorated, or remained the same in the past 12 months?	Deteriorated, Stayed the same, Improved	1:3	1.97	0.61
Liveability score	On a scale of 1 to 10, how pleasant is it to live in your neighbourhood?	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	1:10	7.61	1.38
Nuisance from litter on the street	Trash on the street: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.01	0.72
Nuisance from vandalised street furniture	Vandalised street furniture, e.g., trash cans or benches: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.55	0.64
Nuisance from public drunkenness	Drunk people on the street: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.63	0.61
Nuisance from confused persons	Confused people: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.58	0.64
Nuisance from drug use	Drug use, e.g., on the street or at coffee shops: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.7	0.59
Nuisance from drug dealing	Drug dealing: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.76	0.55
Nuisance from neighbours	Nuisance from neighbours: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.68	0.58
Nuisance from street harassment	People being harassed on the street: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.8	0.5
Nuisance from loitering youth	Loitering youth: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.55	0.68
Nuisance from noise	Noise nuisance: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.4	0.75
Nuisance from odour	Odour nuisance: Do you experience a lot, a little, or no nuisance?	A lot of nuisance, A little nuisance, No nuisance	3:1	2.78	0.53
Nuisance from disrespectful treatment	This question is about disrespectful behaviour. This is behaviour that exceeds the boundaries of decency. How often are you treated disrespectfully by strangers on the street in your neighbourhood?	Often, Sometimes, Rarely, Never	4:1	2.99	0.87
Frequency of feeling unsafe (neighbourhood)	How often do you feel unsafe in your neighbourhood?	Often, Sometimes, Rarely, Never	4:1	3.55	0.85

Survey item	Survey question	Likert Scale Labels	Range	Mean	SD
Not opening the door due to safety concerns	Does it ever happen that you do not open the door in the evening because you do not feel safe?	Often, Sometimes, Rarely, Never	4:1	3.06	1.02
Taking detours to avoid unsafe areas	Does it ever happen that you take a detour in your neighbourhood to avoid unsafe places?	Often, Sometimes, Rarely, Never	4:1	3.39	0.89
Feeling unsafe walking at night	Does it ever happen that you feel unsafe when you walk on the street in your neighbourhood in the evening?	Often, Sometimes, Rarely, Never	4:1	3.24	0.91
Feeling unsafe while home alone	Does it ever happen that you feel unsafe when you are home alone in the evening?	Often, Sometimes, Rarely, Never	4:1	3.66	0.67
Fear of becoming a crime victim	Are you ever afraid that you will become a victim of crime in your neighbourhood?	Often, Sometimes, Rarely, Never	4:1	3.31	0.84
Perception of neighbourhood crime level	Do you think there is a lot, little, or no crime in your neighbourhood?	A lot, A little, None	3:1	1.97	0.56
Perceived change in neighbourhood crime level	Do you think crime in your neighbourhood has increased, decreased, or remained the same in the past 12 months?	Increased, Stayed the same, Decreased	3:1	1.88	0.48
neighbourhood safety rating	Can you give a score for the safety in your neighbourhood?	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	1:10	7.21	1.45
Frequency of feeling unsafe (general)	How often do you feel unsafe?	Often, Sometimes, Rarely, Never	4:1	3.26	0.91
Violence	Have you been a victim of violence in your neighbourhood in the past 12 months?	Yes, No	2:1	1.98	0.13
Pickpocketing	Have you been a victim of pickpocketing in your neighbourhood in the past 12 months?	Yes, No	2:1	1.96	0.18
Theft	Have you been a victim of theft in your neighbourhood in the past 12 months?	Yes, No	2:1	1.78	0.41
Vandalism	Have you been a victim of vandalism in your neighbourhood in the past 12 months?	Yes, No	2:1	1.86	0.35
Burglary	Have you been a victim of burglary in your neighbourhood in the past 12 months?	Yes, No	2:1	1.96	0.19
Satisfaction with last police contact	How satisfied or dissatisfied are you with this contact?	Very dissatisfied, Dissatisfied, Neutral, Satisfied, Very satisfied	1:5	3.72	1.12
Satisfaction with police	How satisfied or dissatisfied are you with the functioning of the police in your neighbourhood?	Very dissatisfied, Dissatisfied, Neutral, Satisfied, Very satisfied	1:5	3.45	0.83
Police visibility	How often do you see the police in your neighbourhood?	Never, Rarely, Sometimes, Often	1:4	2.64	0.86
Satisfaction with police visibility	How satisfied or dissatisfied are you with this visibility?	Very dissatisfied, Dissatisfied, Neutral, Satisfied, Very satisfied	1:5	3.2	0.98

Survey item	Survey question	Likert Scale Labels	Range	Mean	SD
Police protection for residents	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police provide protection to the residents of this neighbourhood.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.41	0.66
Police contact with residents	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police have little contact with the residents of this neighbourhood.	Strongly agree, Agree, Neutral, Disagree, Strongly disagree	5:1	1.72	0.69
Police cooperation with residents	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police work well with the residents of this neighbourhood.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.09	0.61
Police handling of issues	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police handle problems in this neighbourhood well.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.11	0.61
Police fining of violations	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police issue too few fines.	Strongly agree, Agree, Neutral, Disagree, Strongly disagree	5:1	1.96	0.7
Police taking residents seriously	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police in the neighbourhood take you seriously.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.39	0.58
Police approachability	To what extent do you agree or disagree with the following statements about the police in your neighbourhood: The police in the neighbourhood are easily approachable.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.27	0.7
Satisfaction with police	How satisfied or dissatisfied are you with the functioning of the police in general?	Very dissatisfied, Dissatisfied, Neutral, Satisfied, Very satisfied	1:5	3.51	0.83
Police approachability	To what extent do you agree or disagree with the following statements about the police in general: The police are easily approachable.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.46	0.68
Police information provision	To what extent do you agree or disagree with the following statements about the police in general: The police inform the citizens.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.12	0.71
Police handling of crime	To what extent do you agree or disagree with the following statements about the police in general: The police handle crime well.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.15	0.7
Satisfaction with handling of safety	How satisfied or dissatisfied are you with the functioning of your municipality in terms of handling liveability and safety?	Very dissatisfied, Dissatisfied, Neutral, Satisfied, Very satisfied	1:5	3.29	0.91
Municipality's effort to improve safety	To what extent do you agree or disagree with the following statements about the functioning of your municipality: The municipality is committed to liveability and safety in the neighbourhood.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.41	0.71

Survey item	Survey question	Likert Scale Labels	Range	Mean	SD
Municipality's provision of safety information	To what extent do you agree or disagree with the following statements about the functioning of your municipality: The municipality informs the neighbourhood about the approach to liveability and safety in the neighbourhood.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.33	0.76
Municipality's involvement of citizens	To what extent do you agree or disagree with the following statements about the functioning of your municipality: The municipality involves the neighbourhood in the approach to liveability and safety in the neighbourhood.	Strongly disagree, Disagree, Neutral, Agree, Strongly agree	1:5	2.26	0.74
Visibility of municipal enforcers in neighbourhood	How often do you see municipal enforcers: In your neighbourhood?	Never, Rarely, Sometimes, Often	1:4	2.33	0.95
Visibility of municipal enforcers elsewhere	How often do you see municipal enforcers: Elsewhere in your municipality?	Never, Rarely, Sometimes, Often	1:4	3.05	0.78
Satisfaction with municipal enforcers	How satisfied or dissatisfied are you with the functioning of municipal enforcers?	Very dissatisfied, Dissatisfied, Neutral, Satisfied, Very satisfied	1:5	3.17	0.78
Leaving lights on when no one is home	How often do you take the following measures: I leave the light on when no one is home in the evening.	Always, Rarely, Sometimes, Often, Never	5:1	3.62	1.45
Parking bike in guarded parking	How often do you take the following measures: I park my bike in a guarded bike parking if possible.	Always, Rarely, Sometimes, Often, Never	5:1	2.91	1.54
Removing valuables from car	How often do you take the following measures: I take valuable items out of the car.	Always, Rarely, Sometimes, Often, Never	5:1	1.41	0.86
Leaving valuables at home	How often do you take the following measures: When I go somewhere, I leave valuable items at home to prevent street theft.	Always, Rarely, Sometimes, Often, Never	5:1	3.2	1.39
Installing extra security locks	How often do you take the following measures: Extra security locks or bolts on windows and/or (external) doors?	Yes, No	2:1	1.5	0.5
Using outdoor lighting with sensor	How often do you take the following measures: Outdoor lighting with a sensor?	Yes, No	2:1	1.8	0.4

D.2 DEVIATIONS FROM THE PRE-REGISTRATION

The analyses deviate from the pre-registration in several key ways:

1. The pre-registration specified district-level instead of neighbourhood-level aggregation for synthetic control areas. However, analyses revealed that district-level controls did not yield accurate synthetic control groups. Neighbourhood-level controls were employed instead to ensure a better match and more reliable comparisons, although providing somewhat lower statistical power.
2. While the pre-registration planned to use calls-for-service rates per 100,000 inhabitants, the microsynth package works with cumulative counts rather than rates for its calculations. As a result, we used raw CFS counts to ensure compatibility with the software.
3. The pre-registration identified average emergency response time as a key outcome variable. However, this metric proved too volatile to reliably estimate a counterfactual. Literature also highlights the importance of the number of ‘fast responses’ (< 3 minutes) in influencing outcomes such as arrest rates (Bieck, 1977a; Spelman & Brown, 1981). Consequently, we focused on fast response counts instead.
4. BVH data (police registration data) is left out as it could not be obtained in time for analysis.
5. The pre-registration incorrectly stated that the percentage of non-Western migrants would be included in the *match.covar* parameter in the microsynth function. This was corrected to *match.covar.min*, in line with the intended approach for covariates.
6. The pre-registration lacked a detailed plan for addressing survey mode effects. While it mentioned comparing survey results between the VMAA and VN survey modes, it did not specify analysing the interaction effect between survey wave (pre/post) and survey mode. In this analysis, we used regression models to examine interaction effects, ensuring survey mode biases did not influence the difference-in-difference (DiD) estimates. No significant interaction effects were found, justifying the retention of data from both survey modes.
7. The pre-registration indicated that standardised mean differences (SMDs) would use the pooled standard deviation (SD) of the treated and synthetic control groups. However, given the multiple treatment groups and different sets of synthetic control groups for each survey item, this approach resulted in inconsistencies. Instead, we standardised using the SD of the unweighted control group at the post-intervention timepoint. This change effectively shifted from Hedges’ *g* to Glass’s *D*, yielding consistent standardisation and allowing reverse engineering of mean differences using descriptive statistics provided in Appendix section D.1.
8. To better capture the broader spatial impacts of the intervention, the final analysis incorporated a full model that included all areas potentially affected (i.e., high-treatment, low-treatment, and catchment areas). This expanded the scope beyond the originally specified high- and low-treatment areas. Simultaneously, hypotheses focusing on the intersection of the catchment and low-treatment areas were removed, as they were deemed superfluous and provided minimal additional information compared to the full model. The full model was ultimately more informative for evaluating whether crime displacement or a diffusion of benefits occurred.

D.3 SURVEY DATA FOREST PLOTS

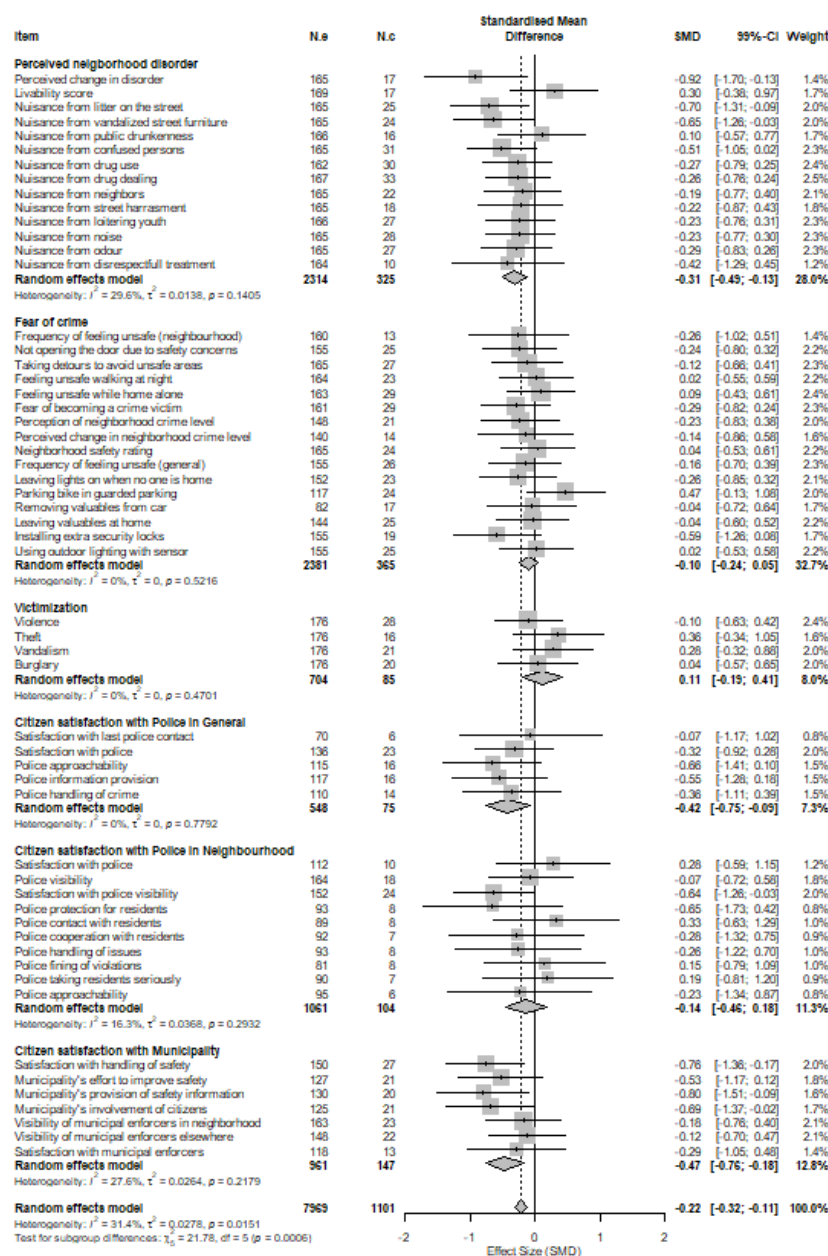


Figure D.1: Forest plot of standardised treatment effects in the high-treatment area

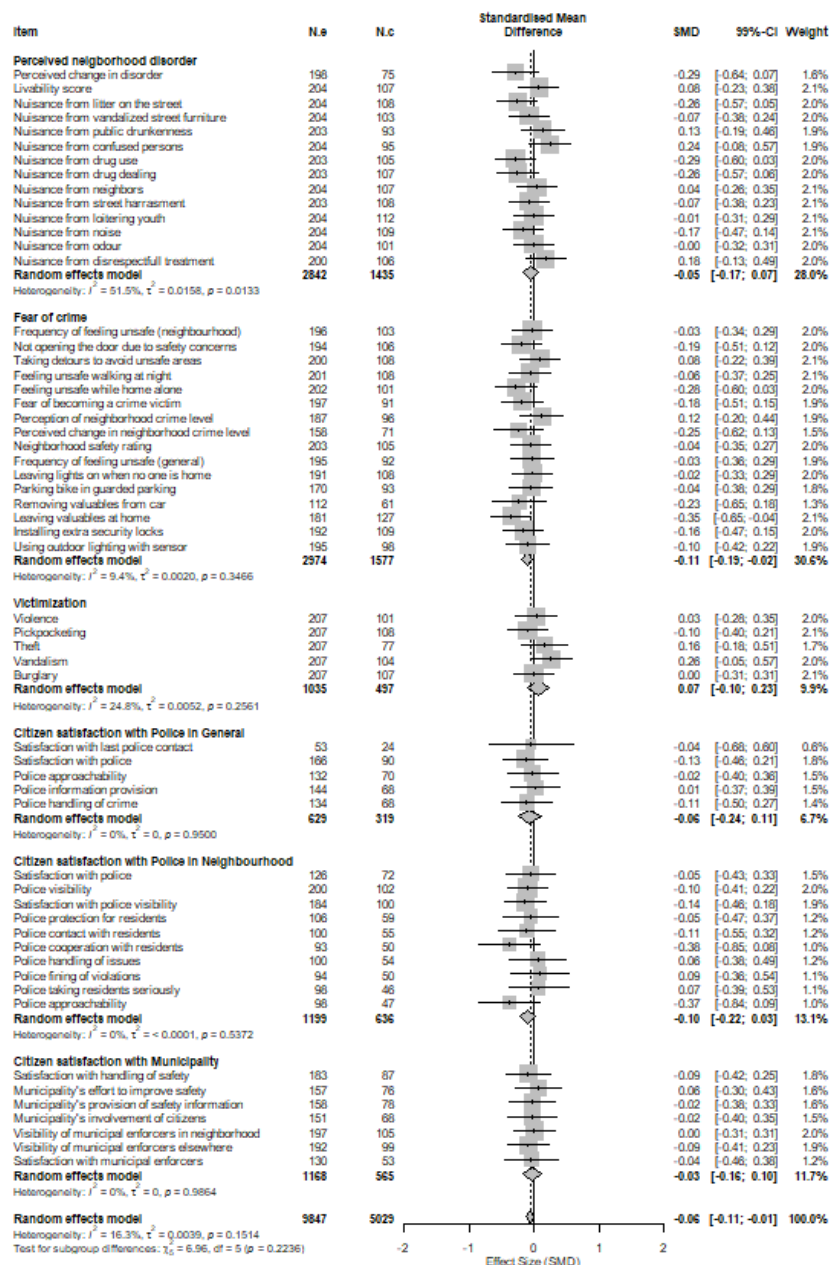


Figure D.2: Forest plot of standardised treatment effects in the low-treatment area

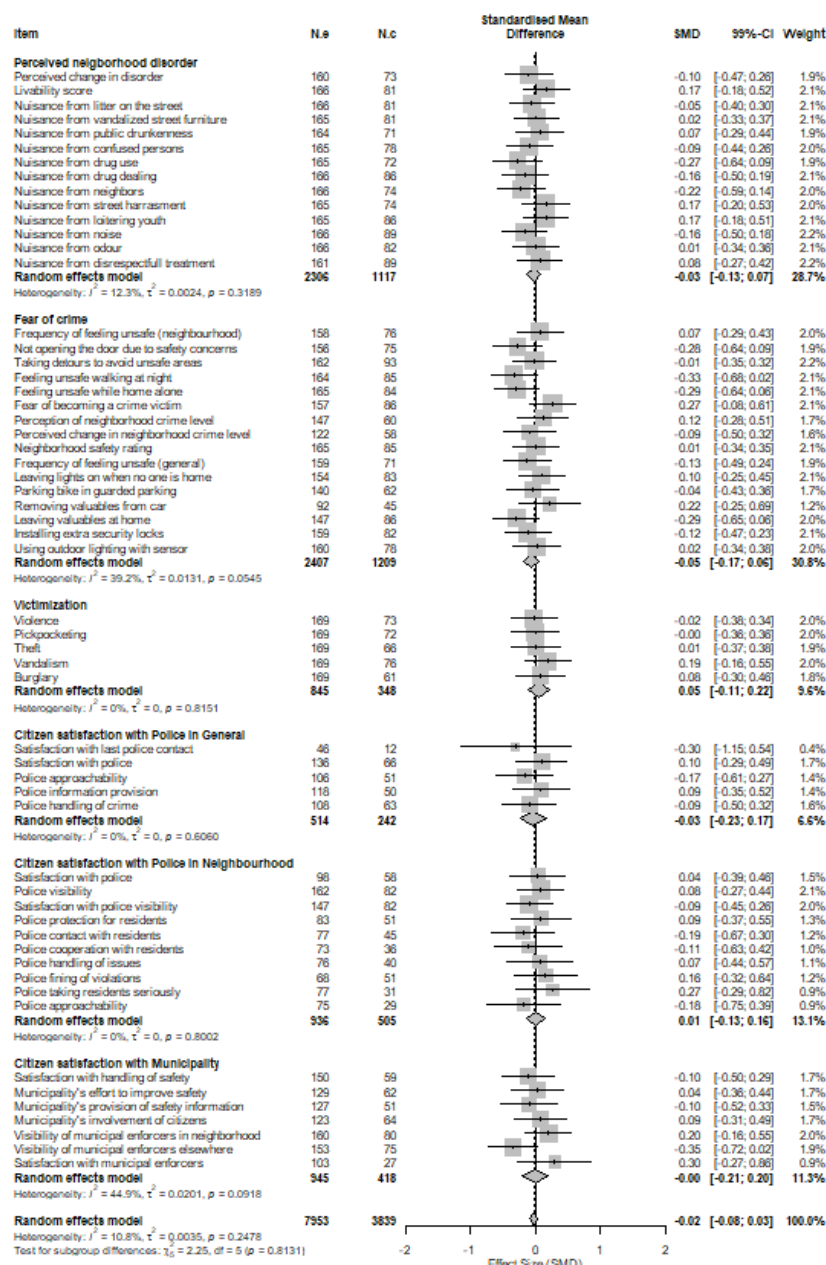


Figure D.3: Forest plot of standardised treatment effects in the catchment area

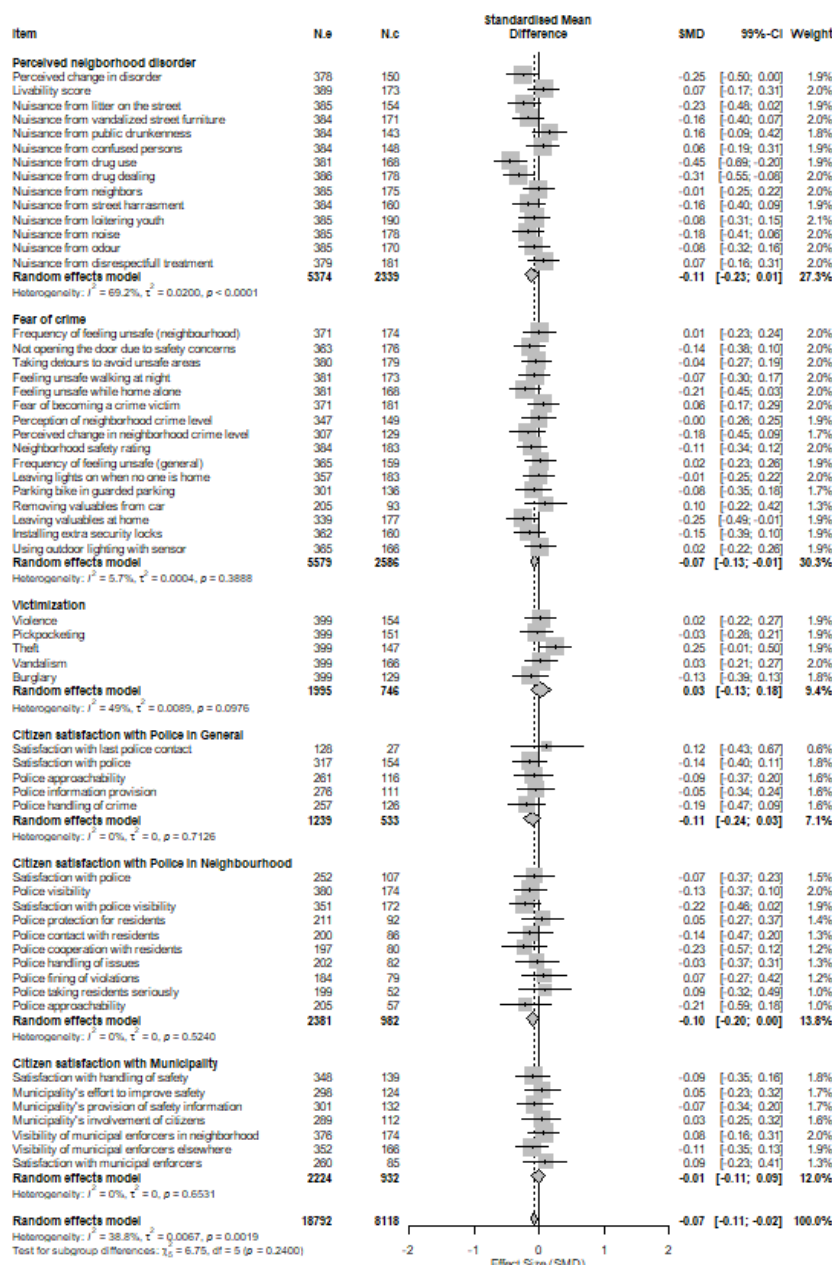


Figure D.4: Forest plot of standardised treatment effects in the full-treatment area

D.4 SUPPLEMENTARY ANALYSIS RESULTS

Table D.2: Supplementary analysis results

Question	1	2	3	4	5	6	7	8	9
How many days do you typically visit Haarlemmerplein?	0 (2.9%)	1 (15.0%)	2 (11.7%)	3 (12.4%)	4 (10.9%)	5 (6.6%)	6 (6.2%)	7 (22.6%)	Don't know (5.1%)
In the past 4 months, have you seen more or less police around Haarlemmerplein compared to last year?	Much more (1.1%)	More (17.9%)	No more/ No less (47.8%)	Less (1.8%)	Much less (0%)				Don't know (24.8%)
Were you aware of the temporary, additional police presence at Haarlemmerplein before this survey?	Yes (8.8%)	No (81.8%)							Don't know (2.9%)
What do you think about the extra police presence at Haarlemmerplein over the past 4 months?	Very positive (6.9%)	Positive (26.6%)	Neutral (21.2%)	Negative (2.9%)	Very negative (0.4%)				Don't know (35.4%)
To what extent do you agree with the statement: 'More police presence has led to less crime and disorder at Haarlemmerplein'?	Strongly agree (4.4%)	Agree (9.1%)	Neutral (21.9%)	Disagree (6.2%)	Strongly disagree (4.4%)				Don't know (47.4%)
Were you aware of the temporary police station at Haarlemmerdijk before this survey?	Yes (8.4%)	No (82.5%)							Don't know (2.6%)

DISCUSSION

This thesis began with a tale of police chiefs approaching the ivory towers of academia. It questioned the current state of the evidence and its inability to provide the chiefs with sufficient evidence-based insights in allocating 1,000 additional police officers. This final chapter turns that same critical eye on the thesis itself: *What evidence does it provide? How might this evidence aid the police chiefs in making evidence-based decisions? And what are the implications for the academic field of policing studies – the proverbial ivory tower?*

The first point to note here is that this thesis does not come close to resolving the purported imbalance in the evidence base for reactive and proactive policing strategies. As such, a comprehensive analysis of all policing strategies and tasks, and the marginal benefit of allocating officers across them, remains out of reach. Still, this thesis hopes to have contributed to the notion that such evidence-based decision-making is both necessary and possible, perhaps even within the career span of our and many other young, eager and evidence-based police chiefs. While practice, policy and research continue to work towards this goal, this thesis does offer several implications worth discussing. Each part of the discussion begins with a short summary of the relevant chapters, followed by a short discussion and reflection on its implications for practice (personified in the police chiefs) and research (caricatured as the ivory tower). Where appropriate, I also outline directions for future research.

PART I – REVISITING RAPID RESPONSE

SUMMARY OF CHAPTERS 1–3

The Chapter 1 (originally published in *Crime Science*) presented a scoping review on the predictors of police response time (Verlaan & Ruiter, 2023b). The review gathered and synthesised 39 empirical studies from both academic and grey literature, identifying 122 unique predictor variables and classifying them into the categories: kinematics, exogenous factors, human involvement, incident characteristics, innovations, and police strategies. The direction, magnitude, significance, and level of internal validity of the observed relationships are summarised in an extensive table and the relationships in a preliminary causal model. The review identified a lack of research into (predictors of) reporting time. Moreover, the initially preregistered search – including both studies on the determinants and effects of police response time (Verlaan & Ruiter, 2023a) – identified a lack of (contemporary) research into the effects of rapid response, resulting in the investigations presented in Chapter 2 and 3.

Chapter 2 provided a historical analysis of how the lack of empirical research into reactive policing – and the overwhelming focus on proactive policing – came about. It started from a series of landmark papers in the US (mostly Kansas City) in the 1970s, and traced the findings to the first large NASAM report, the Maryland Report. It provides a critical empirical analysis the Maryland Report and argues that the dismissal of reactive policing in favour of proactive policing was based on a limited and selectively interpreted and synthesised

evidence base. It concludes that reactive policing should be re-evaluated, starting with a critical re-evaluation of the evidentiary foundation of the dismissal of reactive strategies.

Chapter 3 of this thesis took this in-depth, critical look at the evidentiary foundation of the dismissal of one of the reactive strategies (i.e., rapid response): the 1977 Kansas City Response Time Analysis study (Bieck, 1977a) and its 1981 Police Executive Research Forum replication study (Spelman & Brown, 1981). It argued that the dismissal of rapid response is based on weak evidence: flawed research designs, small samples, unreliable measurements, and an outdated empirical reality. Given these limitations, it contended that the standard view on rapid response, specifically its crime resolution effect, should be reconsidered.

DISCUSSION OF PART I

This first part of this thesis drives an initial wedge in the orthodoxy of the Proactive Policing Paradigm. It synthesises the empirical evidence on predictors of police response times, and investigates the lack of research on the effects of a primary reactive policing strategy: rapid response. It shows that the existing evidence does not justify excluding rapid response from the repertoire of effective policing tactics. What remains is that, across studies, faster response times are associated with higher arrest rates. This relationship is strong but highly crime-specific and strongly non-linear – once response times (including reporting delay) exceed five minutes, marginal benefits largely disappear. Moreover, rapid response is also empirically associated with higher witness availability (e.g., Cordner et al., 1983), crime clearance rates (e.g., Blanes i Vidal & Kirchmaier, 2018), victim satisfaction (e.g., Brandl & Horvath, 1991), victim wellbeing (e.g., B. B. Brown & Harris, 1989), and overall citizen satisfaction (e.g., Frank et al., 2005).

IMPLICATIONS FOR PRACTICE

For the chiefs, these findings suggest that officers assigned to rapid response are not necessarily officers wasted in working towards goals outlined in the pragmatic model of policing – namely, the resolution and prevention of crime and disorder and maintenance of legitimacy and public approval – provided they can respond quickly to ongoing incidents and that citizens are not delaying in reporting. To support this, it would be wise to assign some officers to trained information roles, responsible for systematically monitoring and analysing response times and reporting delays. Identifying local bottlenecks in reporting, dispatch, and travel could improve deployment efficiency and stimulate innovation. Two examples illustrate the potential of such innovations. First, a study introducing personal emergency buttons in a hospital emergency department with frequent incidents of violence reported an improvement in average police response time from three minutes to one (Joslin et al., 2016). Second, a public-private partnership aimed at reducing false alarm responses resulted in an 87% annual reduction in police alarm calls, a 26% decline in burglaries, and a roughly 50% faster response to overall priority 1 calls (Blackstone et al., 2020). Innovation could also be pursued at the policy level. In the Dutch context, the current benchmark requires that 90% of priority 1 calls be responded to within 15 minutes. However, given the steep decline in arrest likelihood after five minutes, it may be also meaningful to track the proportion of responses occurring within five minutes for ongoing crime calls. Establishing a benchmark around this faster threshold could help align performance metrics more closely with outcome-oriented policing goals.

IMPLICATIONS FOR (FUTURE) RESEARCH

For academia, these findings expose a key vulnerability in the evidence-based paradigm: its reliance on (systematic) review studies to offer an authoritative account of the state of knowledge, upon which future scholarship is built. This reliance carries two distinct risks. First, historical errors may go uncorrected for decades. Second, changes in empirical reality may render earlier conclusions outdated, even if the original studies were correct in drawing them. This is most clearly illustrated by the Maryland Report (Sherman et al., 1997, 2400+ citations), which drew strong conclusions about the (in)effectiveness of rapid response without sufficiently acknowledging the specific conditions that made it ineffective at the time – namely, that in the 1970s most crimes were (property crimes and) discovered after the fact rather than witnessed in real time, and that citizen reporting delay was substantial, largely because of the absence of mobile phones. These conclusions were then (understandably) echoed in other highly influential works, such as D. Weisburd and Eck (2004, 1300+ citations), even though by that time property crime had already declined sharply and mobile phone ownership had reached around 65% of the population (Pew Research Center, 2024). This recursive pattern of citation is sometimes explicitly acknowledged – as in Nagin (2013, 1200+ citations), who states: ‘The discussion that follows draws heavily on two excellent reviews of this research by D. Weisburd and Eck (2004) and Braga (2008)’ (Nagin, 2013, p. 236). As a result of this recursive pattern, foundational conclusions can persist long after their original empirical justification has eroded.

Besides this point, this part highlights three significant gaps in the research literature: (1) a lack of knowledge on the causal determinants and effects of police response times; (2) a lack of knowledge on the scope, magnitude and mechanisms behind citizen reporting delay; and (3) a lack of research into (promising) innovations in improving (all components of) police response time.

The first gap – limited knowledge of the causal determinants and effects of rapid response – is admittedly challenging to address, as conducting field experiments in this domain raises serious ethical and practical concerns. Nevertheless, recent advances in econometric methods offer promising alternatives. Techniques such as instrumental variable (IV) estimation and regression discontinuity designs (RDD) can yield causal insights, even in the absence of randomised interventions. While such approaches to studying police response time have seen increasing use in economics (e.g., Blanes i Vidal & Kirchmaier, 2018), they remain underutilised in criminology and policing research. A more systematic engagement with this type of research is both necessary and long overdue within criminology. Concrete suggestions for advancing this research are discussed in detail in the concluding section of Chapter 1 and Chapter 2.

The second gap – the lack of a current understanding of citizen reporting delay – is particularly striking. Since the 1970s, to the best of my knowledge, no empirical investigation into this phenomenon has been conducted, leaving the field reliant on findings from an era when payphones were commonplace and mobile phones a rarity. Admittedly, reporting delay is a difficult construct to measure: the precise start time of an incident can be conceptually ambiguous (for instance, does an assault incident begin at the moment of verbal confrontation, or at the point of first physical harm?) and even with a clear beginning, accurate timestamps are difficult to obtain. However, recent technological developments open up promising new avenues for empirical inquiry. Automatic gunshot detection systems (AGDS) provide data on

a clearly delineated type of incident along with a precise timestamp and location. Similarly, (inner-city) CCTV footage documents timestamped records of offences and subsequent help-seeking behaviour. In the Netherlands, research using such footage is expanding rapidly through the work of, among others, the Video Group at the NSCR (see Lindegaard & Bernasco, 2018; Philpot et al., 2019, 2020; van Bruchem et al., 2024). Cross-referencing data from these systems with calls-for-service data may not only quantify the scope and magnitude of reporting delays but also shed light on the underlying mechanisms – such as bystander hesitation, victim confusion, or contextual barriers to reporting.

The third gap – the lack of (research on) innovations to improve police response time – is equally noteworthy. While much of the recent innovation in policing has focused on proactive strategies (e.g., *PredPol*), reactive policing – particularly rapid response – has seen far less recent development, both in terms of operational experimentation and empirical evaluation. Nevertheless, the limited innovations that have been tested yield promising results (e.g., Blackstone et al., 2020; Joslin et al., 2016).

PART II – INVESTIGATING POLICE INFLUENCE ON RESPONSE TIMES

SUMMARY OF CHAPTER 4

Chapter 4 presented an agent-based modelling (ABM) study investigating the effects of police deployment strategy on response times – specifically, average response time, percentage of fast responses, and percentage of late responses. This was the first study to use empirical data to simulate the effects of deployment on police response times. The model compared three strategies: random patrol, strategic positioning, and idling at the location of the most recent call-for-service. It also explored how these effects varied with police capacity, urbanisation level, and call volume per shift. The findings showed that strategic positioning significantly outperformed the other strategies on all outcome measures, particularly in increasing the share of fast responses. By contrast, changes in police capacity had a much smaller impact.

DISCUSSION OF PART II

This second part of the thesis challenges the notion – put forward in the Maryland Report – that even if rapid responses were effective, the police would have little ability to improve travel times. Specifically, it scrutinises the claim that halving police travel time would require a doubling of the police force (Sherman et al., 1997), refuting the idea that substantial improvements in response time are only achievable through unrealistically large increases in police resources. Instead, it shows that strategic deployment – without expanding the workforce – can substantially improve responsiveness. This is a crucial insight, as the police's ability to influence response time has been a precondition to studying its effects. The more so, because accurately studying response times in experimental settings could be a very costly and ethically infeasible endeavour.

IMPLICATIONS FOR PRACTICE

For the police chiefs, these findings suggest that operational decisions – particularly those concerning deployment strategy – offer real leverage in improving the speed of police

response and decreasing the likelihood of late responses and improving the likelihood of fast responses.

Reducing late responses is especially salient in the Dutch context, where the police operate under mayoral authority and where national benchmarks require 90% of priority 1 calls to be answered within 15 minutes. In practice, this benchmark is met in only about 10% of municipalities (RTLNieuws, 2025), and some 40% of Dutch mayors report dissatisfaction with response times (Terpstra et al., 2015). Strategic deployment may offer a cost-effective lever for addressing these concerns.

That said, police leaders should exercise caution. The ABM, while realistic, remains a simulation. It offers strong internal validity – unlike real-world experiments, agents in the model follow instructions perfectly and without deviation – but its external validity still requires empirical testing. Moreover, the model optimises for response time alone. It does not account for other policing goals, such as maintaining a proactive presence in high-crime areas. As such, its primary value lies not in prescribing daily deployment decisions but in two other areas: (1) guiding future research, including experimental or quasi-experimental tests of deployment strategies in practice; and (2) informing contingency planning – for example, during understaffed shifts in rural jurisdictions – where preserving the integrity of the response system becomes especially critical.

IMPLICATIONS FOR (FUTURE) RESEARCH

For academia, this part of the thesis offers three main contributions. First, it demonstrates the value of agent-based modelling (ABM) as a method for investigating operational strategies that are otherwise difficult to examine through field experiments. Given the ethical and logistical challenges of experimentally manipulating response strategies in real-world settings, ABMs offer a promising route to high-internal-validity simulations grounded in empirical data. Second, the findings challenge long-held assumptions about the (in)flexibility of police responsiveness – assumptions that have largely gone untested. Third, the results show that key contextual factors such as capacity, call volume, and level of urbanisation do not determine response time in a fixed way, but interact with strategic deployment choices.

These findings call for greater scholarly attention to the how of reactive policing. More specifically, to me, they invite a bold next step: the first experimental study of a reactive policing model, which I envision to be as follows:

Rather than patrolling streets or crime hot spots, officers are stationed in mobile police stations at strategically chosen locations across a jurisdiction, much alike how contemporary ambulance deployment is structured. These locations should be optimised using street network centrality and coverage while taking into account variability in reactive demand. Importantly, they should also be selected for their potential to support community interaction – for example, near town squares or large convenience stores.

Each mobile unit would be staffed by a pair of officers tasked with two roles: rapid response and reactive investigation. First, they must be ready to respond rapidly (by car) to urgent calls for service within their designated area. Second, during call-free periods, they manage crime reports within their zone. This could include general investigation, identifying patterns across incidents, gathering additional information from victims or witnesses, updating complainants on case progress, or explaining why a case was closed.

I would suggest to first pilot this approach in a more rural police district as it would directly

address a frequently voiced concern among the Dutch public following the nationalisation of the police: the disappearance of local police stations. In terms of visibility, a stationary unit positioned in a location with high foot traffic may be more effective than traditional mobile patrols. Moreover, a stationary unit might also have a larger and more lasting deterrent effect than individual mobile police cars (Blesse & Diegmann, 2022; Fondevila et al., 2021). As such, this approach could not only improve police response times – helping more municipalities meet the national 15-minute benchmark and increasing the proportion of fast responses, potentially leading to more on-scene arrests and identified witnesses, and maybe even deterring more crime than realised under a proactive policing regime. In addition, it offers a potential remedy to a persistent issue in the Dutch policing context: the growing backlog of unresolved or unaddressed crime reports – approximately 40% of complainants report never receiving feedback on their crime report, and another 35% are informed that no investigation will be pursued (Kester, 2023). Recent research in the Netherlands indicates that victim satisfaction with police services is best predicted by variables within police control, namely regarding how they handle crime reports (Jansen et al., 2024). But most importantly, it would offer a counterweight to the exclusive focus on proactive policing interventions.

PART III - MAPPING REACTIVE DEMAND AND SUPPLY

SUMMARY OF CHAPTERS 5–8

Chapter 5 (originally published as chapter in *The Crime Data Handbook*) offers a discussion of the use of inferential statistics on police administrative data, emphasising their status as population data rather than samples, and the implications this holds for statistical analysis (Verlaan & Langton, 2024). The chapter also reports on a short scoping review of recent studies which shows statistical inference is frequently applied to such population data without justification. The chapter discusses this practice and advises greater transparency in inferential assumptions, a reconsideration of alternative analytical approaches, and a stronger emphasis on descriptive statistics in police research – as was followed in the chapters of this thesis.

Chapter 6 (originally published in *Crime Science*) examines the measurement of reactive policing through police administrative data, focusing on different operationalisations of dispatched deployment time and their impact on estimating the scale and composition of police work (Langton, Verlaan & Ruiter, 2023). Among other things, this chapter shows the dispatched deployment time to be fairly robust to the different operationalisations. While the most granular data containing timestamps at the unit level are more accurate, incident level timestamps can also provide a close approximation of the composition of dispatched deployment time.

Chapter 7 (originally published in *Police Practice and Research*) reported on a replication of work by Ratcliffe (2021) and Lum et al. (2022) on the scale and composition of the reactive supply of policing (Langton, Ruiter & Verlaan, 2023). Using open source calls-for-service data from Detroit 2019 and operationalizing deployment time based on incident level timestamps, the study found that police spend about 50% of their time on crime-related calls, 46% of their time on disorder-related calls, and 4% of their time was registered as self-initiated/proactive.

Chapter 8 (translated from an article published in Dutch practice-oriented journal

Tijdschrift voor de Politie) describes the scale and composition of the reactive supply of police for the whole of the Netherlands based on highly-detailed unit-level timestamps (Verlaan et al., 2023). The composition of reactive supply is constructed for the vast majority of police frontline teams and nested within their overarching regional units, showing the variations both among and within police regions. The chapter reflects on this context-dependent nature of (reactive) policing and its implications for police practice.

DISCUSSION OF PART III

This part of the thesis contributes to a more grounded understanding of what reactive policing actually entails in practice. It provides empirical estimates of the scale and composition of reactive deployment, drawing on data from both a major US city and the whole of the Netherlands policing context. Across contexts, roughly half of police time is spent on crime-related calls, the other half on disorder-related issues – while only a small fraction is spent on registered proactive activities. Moreover, the studies demonstrate that measures of deployment time are relatively robust to different levels of data granularity, with incident-level timestamps serving as a reasonable proxy in the absence of unit-level data.

IMPLICATIONS FOR PRACTICE

For the police chiefs, this part highlights four key points. First, it affirms the complexity and diversity of reactive policing. Second, the variation in task composition across teams suggests training could be tailored to local demand – for example, prioritising public order skills in some areas, traffic or investigative skills in others. Third, insights into team-level variation can inform internal coordination, such as targeted knowledge sharing and more strategic staff deployment. Finally, these insights support more context-sensitive allocation decisions, tailored to local variations in demand. The mapping of reactive deployment at both national and team levels in the Netherlands shows that one-size-fits-all strategies are unlikely to be effective – local conditions require local solutions.

To support effective decision-making, the (national) police could ensure that these insights are continuously made available to local jurisdictions. Patterns in reactive demand and supply are subject to (seasonal) changes, yet remain highly informative for local strategic decision-making. Fortunately, monitoring these patterns is entirely feasible using standard (administrative) calls-for-service and dispatch data. With minimal effort, they can be visualised through straightforward dashboards that track the scale and composition of reactive demand and deployment over time, enabling local teams to adjust strategies based on up-to-date operational realities.

IMPLICATIONS FOR (FUTURE) RESEARCH

For academia, this part of the thesis (Chapter 5) first calls attention to a key methodological issue: the routine but often unjustified use of inferential statistics in research based on administrative police data. These datasets typically represent full populations rather than samples, rendering traditional inferential tools – such as p-values and confidence intervals – unnecessary and, in many cases, misleading. The uncritical use of inference in this context carries several risks: (1) it wrongfully suggests generalisability; (2) it may obscure or undervalue substantively meaningful results that fail to meet statistical significance thresholds; and (3) it can reduce the clarity and accessibility of findings by prioritising abstract statistical

markers over concrete, interpretable measures. Instead, this thesis advocates for a more deliberate and transparent use of descriptive statistics, which are better suited to the nature of administrative data and more effective in supporting policy and operational decision-making. Where inferential methods are used, researchers should clearly define the target population, articulate the purpose of statistical testing, and demonstrate its added value beyond what descriptive analysis can already reveal.

Beyond this methodological point, part III highlights the broader need for empirical research on the demand and supply side of policing. While much policing scholarship focuses on interventions or outcomes, comparatively few studies map how police time is actually spent. Such foundational work is essential to ground evidence-based policing in real operational conditions. Moreover, these studies show that even coarse data – such as incident-level timestamps – can yield valuable insights when interpreted carefully and descriptively. This opens up research opportunities even in jurisdictions with limited access to granular data.

In terms of future research, findings on reactive policing are also highly relevant for understanding the effectiveness of proactive policing strategies. In particular, it is important for future research to examine how variations in reactive demand and supply interact with the outcomes of proactive interventions. For instance, are there systematic – potentially non-linear – differences in the crime-preventive effects of proactive policing strategies across jurisdictions with lower overall levels of crime and disorder, or across those with similar volumes but different compositions of incidents? Currently, most proactive interventions have been tested in large U.S. cities with relatively high levels of crime and disorder and subsequently synthesised through meta-analyses with limited attention to contextual variation. A more granular understanding of crime and disorder composition, combined with techniques such as moderator analysis, may offer valuable insights into the mechanisms and moderators that shape the effectiveness of proactive policing programmes.

PART IV – EVALUATING PROACTIVE POLICING IN THE NETHERLANDS

SUMMARY OF CHAPTER 9

Chapter 9 presented a quasi-experimental evaluation of a place-based policing intervention in Amsterdam's Haarlemmerbuurt. The intervention was co-designed with the local frontline team and aimed to maximise proactive presence in a small hotspot area using existing personnel and resources. Using the micro synthetic control method, the study examined the effects on crime-, disorder-, and traffic-related calls-for-service levels and public perceptions using the standardised and long running national victim survey (*VeiligheidsMonitor*).

The intervention realised an estimated 4 to 5 times increase police presence in the neighbourhood and led to a localised (23%) reduction in crime-related calls-for-service in the high-treatment area. However, there were signs of displacement to surrounding areas and the rest of the frontline team jurisdiction. Moreover, during the intervention period, public perceptions worsened: residents in the treated area reported increased concerns about disorder and a sharper decline in satisfaction with the police and municipality compared to the synthetic control group. The chapter also examined the implementation process in detail. Despite initial enthusiasm, frontline officers experienced difficulty and sometimes

even frustration with the ‘self-directing’ character of the intervention and proactive policing in general.

DISCUSSION OF PART IV

This part provides a practice-embedded test of place-based policing in a European urban context. While the intervention achieved measurable reductions in crime-related calls, it did not translate into net benefits due to displacement effects. Moreover, local public perceptions deteriorated. These findings echo previous findings that crime reductions do not always align with improved public trust or perceived safety. Interviews with participating police officers reveal how police organisational dynamics – particularly the challenges of self-directed work and internal politics – can shape implementation and limit impact. Furthermore, the evaluation also raises the important but underexplored issue of reactive mechanisms, which may have contributed to crime reduction, yet are typically omitted from evaluations of proactive policing interventions.

IMPLICATIONS FOR PRACTICE

For police chiefs, these findings raise important questions about the assumed benefits of proactive presence – or *blauw op straat* (blue on the street), as it is referred to in Dutch public and political discourse. Until now, the perceived effectiveness of proactive policing has rested primarily on studies from large, high-crime urban centres in the United States and the United Kingdom. With this study, the Netherlands Police finally has some robust findings from its own context – the first since the Spickenheuer (1983) evaluation.

These findings are particularly timely. In its recent strategic agenda, the Dutch police have announced an ambition to shift away from reactive work and invest more heavily in proactive presence within neighbourhoods with the goal to improve visibility, strengthen the relationship between citizens and the police, and enhance the organisation’s information position (Nationale Politie, 2025).

However, as I also argued in a recent op-ed in *de Volkskrant*, our findings do not support these expectations. We found that only 8% of residents noticed the increased police presence. Moreover, public satisfaction with both the police and the municipality declined during the intervention period. Also, we received no indication that the increased presence yielded concrete intelligence that led to tangible outcomes, such as arrests or more focused follow-up.

This evaluation is, of course, only a first step. If the Netherlands Police wish to make genuinely evidence-based strategic decisions, many more studies will be needed. Nonetheless, decisions must be made in the meantime – and at present, this study offers the most robust, albeit limited, empirical basis for doing so within the Dutch context. As such, and as I argued in the op-ed, if police leadership aims to strengthen its information position and build public trust, a more promising strategy may lie not in increasing visible patrols but in consistently and promptly following up on crime reports. As mentioned earlier, research by Kester (2023) found that 40% of Dutch citizens who file a report never receive any response from the police. Addressing that gap may ultimately do more to enhance legitimacy than putting more ‘Blue on the Street’ (Jansen et al., 2024).

IMPLICATIONS FOR (FUTURE) RESEARCH

This study advances the evaluation of proactive policing in several important ways. It applies the EMMIE framework, not only assessing effectiveness but also exploring underlying mechanisms, implementation processes, and contextual factors. The study's pre-registration and pre-specified analysis code strengthen transparency and reduce analytic flexibility – practices still rare in policing research. The use of a (micro) synthetic control method provides a robust counterfactual in the absence of randomisation, while qualitative interviews with participating frontline officers, local shopkeepers and residents, offer valuable insight into implementation challenges, including role ambiguity, self-direction ability, and organisational politics. Uniquely, the study also captures both objective and subjective outcome domains: reductions in crime-related calls are weighed alongside declines in public trust and satisfaction, reflecting a more holistic view of proactive police impact. Moreover, by evaluating a proactive policing intervention in a relatively low-crime European city, the study broadens the geographic and criminological scope of evidence, moving beyond the typical U.S.-centred, high-crime contexts that dominate the literature.

At the same time, this study prompts a broader reflection on what remains unknown in the evaluation of proactive hot spot policing interventions. While this approach has become the dominant evidence-based policing strategy, the underlying evidence base remains narrower and more contextually constrained than often acknowledged. For example, the systematic review on hot spot policing aggregates highly heterogeneous interventions – ranging from foot patrols and traffic stops to CCTV and drug raids – into a single effect size, despite distinct aims, moderators and mechanisms (Braga, Turchan, Papachristos & Hureau, 2019). Contextual variation is often noted but seldom investigated: the vast majority of studies are conducted in large, high-crime U.S. cities, with little attention to whether, how, or why effects might differ in smaller, less criminogenic, or more centralised policing environments.

Moreover, the core mechanisms of deterrence are often underspecified. It remains unclear whether observed effects stem from general or specific deterrence, and whether they arise from proactive visibility (mainly increasing certainty of punishment), or through reactive policing working through incapacitation or specific deterrence. Dosage – how often, for how long, and in what configurations patrols are deployed – is not often measured in ways that allow cross-study comparison or cumulative learning. Potential placebo effects – on both the public (as a result of a public awareness campaign) and police officers (as a result of using an exciting new technology) – are almost never addressed. Nor are opportunity costs: what other policing functions (e.g., follow-up investigations, community engagement) are being displaced by increased patrol?

These gaps leave important questions unanswered. How does proactive policing perform across different contexts and crime types? What are the precise mechanisms of action? Under what conditions does proactive policing reduce crime, and when might it erode public trust or misallocate scarce police resources? A more careful, mechanism-aware, and context-sensitive research agenda is needed – one that moves beyond the question of whether proactive policing 'works', to more insightful questions of how, for whom, and at what cost.

CONCLUDING REMARKS

As this discussion began by acknowledging; a comprehensive analysis of all policing strategies and tasks – and the marginal benefit of allocating officers across them – still remains out of reach. The current evidence base is too fragmented, historically skewed toward proactive strategies, and limited in its attention to context, mechanisms, costs and trade-offs.

Nonetheless, strategic decisions must be made – and are being made. These decisions cannot wait for perfect evidence. In practice, they are based on the best available insights, even if those insights are partial, provisional, or emerging from new directions such as the ones explored in this thesis. While this thesis does not resolve the imbalance in the evidence base, it does offer several contributions: a re-evaluation of reactive policing; new empirical tools and insights to better understand police responsiveness; improved descriptions of reactive supply and demand; and the first robust evaluation of proactive presence in the Netherlands since the 1980s.

Together, these contributions aim to shift the conversation from simply asking ‘what works’ to asking: how does it work, in what context, at what dosage, in what combination, and at what cost? Only through such a lens can police leaders make informed, realistic, and legitimate choices.

In that light, I end this thesis with the same quote that opened it – one that feels especially apt for the position police chiefs find themselves in: making decisions under pressure, with ambitious goals, limited resources, and incomplete knowledge.

There are no solutions, there are only trade-offs. – Thomas Sowell

REFERENCES

- Bieck, W. H. (1977a). *Response Time Analysis: Executive summary* (tech. rep.). National Institute of Justice. Kansas City.
- Blackstone, E. A., Hakim, S., & Meehan, B. (2020). Burglary reduction and improved police performance through private alarm response. *International Review of Law and Economics*, 63. <https://doi.org/10.1016/j.irle.2020.105930>
- Blanes i Vidal, J., & Kirchmaier, T. (2018). The Effect of Police Response Time on Crime Clearance Rates. *The Review of Economic Studies*, 85(2), 855–891. <https://doi.org/10.1093/restud/rdx044>
- Blesse, S., & Diegmann, A. (2022). The place-based effects of police stations on crime: Evidence from station closures. *Journal of Public Economics*, 207, 104605. <https://doi.org/10.1016/J.JPUBECO.2022.104605>
- Braga, A. A. (2008). *Police Enforcement Strategies to Prevent Crime in Hot Spot Areas*. (tech. rep.). U.S. Department of Justice. Washington, D.C.
- Braga, A. A., Turchan, B., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1046>
- Brandl, S. G., & Horvath, F. (1991). Crime-victim evaluation of police investigative performance. *Journal of Criminal Justice*, 19(3), 293–305. [https://doi.org/10.1016/0047-2352\(91\)90008-J](https://doi.org/10.1016/0047-2352(91)90008-J)
- Brown, B. B., & Harris, P. B. (1989). Residential burglary victimization: Reactions to the invasion of a primary territory. *Journal of Environmental Psychology*, 9(2), 119–132. [https://doi.org/10.1016/S0272-4944\(89\)80003-9](https://doi.org/10.1016/S0272-4944(89)80003-9)
- Cordner, G. W., Greene, J. R., & Bynum, T. S. (1983). *The Sooner the Better: Some Effects of Police Response Time* (tech. rep.). National Institute of Justice. Washington, DC.
- Fondevila, G., Vilalta-Perdomo, C., Galindo Pérez, M. C., & Cafferata, F. G. (2021). Crime deterrent effect of police stations. *Applied Geography*, 134, 102518. <https://doi.org/10.1016/J.APGEOG.2021.102518>
- Frank, J., Smith, B. W., & Novak, K. J. (2005). Exploring the Basis of Citizens' Attitudes Toward the Police. *Police Quarterly*, 8(2), 206–228. <https://doi.org/10.1177/1098611103258955>
- Jansen, R. P., Ruiter, S., & van Steden, R. (2024). Crime reporting and victim satisfaction with the police: a large-scale study among victims of crime in the Netherlands. *Crime Science*, 13(1), 1–11. <https://doi.org/10.1186/S40163-024-00231-9>
- Joslin, J. D., Goldberger, D., Johnson, L., & Waltz, D. P. (2016). Use of the Vocera Communications Badge Improves Public Safety Response Times. *Emergency Medicine International*, 2016, 1–3. <https://doi.org/10.1155/2016/7158268>
- Kester, J. 2023. Bijna helft slachtoffers misdrijf doet geen aangifte bij politie: weinig vertrouwen in pakkans daders. <https://eenvandaag.avrotros.nl/panels/opiniepanel/alle-uitslagen/item/bijna-helft-slachtoffers-misdrijf-doet-geen-aangifte-bij-politie-weinig-vertrouwen-in-pakkans-daders/EenVandaag>.
- Langton, S., Ruiter, S., & Verlaan, T. (2023). Describing the scale and composition of calls for police service: a replication and extension using open data. *Police Practice and Research*, 24(5), 523–538. <https://doi.org/10.1080/15614263.2022.2102494>

- Langton, S., Verlaan, T., & Ruiter, S. (2023). Operationalizing deployment time in police calls for service. *Crime Science*, 12(1), 20. <https://doi.org/10.1186/s40163-023-00198-z>
- Lindegaard, M. R., & Bernasco, W. (2018). Lessons Learned from Crime Caught on Camera. *Journal of Research in Crime and Delinquency*, 55(1), 155–186. <https://doi.org/10.1177/0022427817727830>
- Lum, C., Koper, C. S., & Wu, X. (2022). Can We Really Defund the Police? A Nine-Agency Study of Police Response to Calls for Service. *Police Quarterly*, 25(3), 255–280. <https://doi.org/10.1177/10986111211035002>
- Nagin, D. S. (2013). Deterrence in the Twenty-First Century. *Crime and Justice*, 42(1), 199–263. <https://doi.org/10.1086/670398>
- Nationale Politie. (2025, March). *Strategische Agenda Politie 2025-2030: Stevig Staàn in Deze Tijd* (tech. rep.).
- Pew Research Center. 2024. Demographics of Mobile Device Ownership and Adoption in the United States. <https://www.pewresearch.org/internet/fact-sheet/mobile/>
- Philpot, R., Liebst, L. S., Levine, M., Bernasco, W., & Lindegaard, M. R. (2020). Would I be helped? Cross-national CCTV footage shows that intervention is the norm in public conflicts. *American Psychologist*, 75(1), 66–75. <https://doi.org/10.1037/amp0000469>
- Philpot, R., Liebst, L. S., Møller, K. K., Lindegaard, M. R., & Levine, M. (2019). Capturing violence in the night-time economy: A review of established and emerging methodologies. *Aggression and Violent Behavior*, 46, 56–65. <https://doi.org/10.1016/j.avb.2019.02.004>
- Ratcliffe, J. H. (2021). Policing and public health calls for service in Philadelphia. *Crime Science*, 10(1), 5. <https://doi.org/10.1186/s40163-021-00141-0>
- RTLNieuws. 2025. Politie haalt in 307 gemeenten eigen norm voor aanrijtijden bij noodgevalen niet. <https://www.rtl.nl/nieuws/onderzoek/artikel/5497691/politie-brandweer-aanrijtijden-gemeenten-spoed-noodgeval> Jasper Bunschoek.
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997, January). *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress* (tech. rep.). NIJ. Washington, DC.
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Spickenheuer, H. (1983). *Voetsurveillance en preventievoorzichting in Amsterdam-Osdorp: een onderzoek naar de effecten op de criminaliteitsontwikkeling*. Staatsuitgeverij.
- Terpstra, J., Foekens, P., & van Stokkum, P. (2015, July). *Burgemeesters over hun Nationale Politie* (tech. rep.). Radboud Universiteit Nijmegen. Den Haag.
- van Bruchem, M., Proost, K., van Ruysseveldt, J., & Lindegaard, M. R. (2024). Bystander actions during police work on the street: officer perspectives. *Police Practice and Research*. <https://doi.org/10.1080/15614263.2024.2363335>
- Verlaan, T., & Langton, S. (2024, April). On the Use of Inferential Statistics on Administrative Police Data. In *The crime data handbook* (pp. 197–210). Bristol University Press. <https://doi.org/10.51952/9781529232073.ch014>
- Verlaan, T., Langton, S., & Ruiter, S. (2023). Reactieve politie-inzet op straat. *Tijdschrift voor de Politie*, (4), 36–40.

- Verlaan, T., & Ruiter, S. (2023a). *PROTOCOL-Determinants and effects of police response time: a scoping review*. <https://doi.org/10.17605/OSF.IO/HU2E9>
- Verlaan, T., & Ruiter, S. (2023b). Predictors of police response time: a scoping review. *Crime Science*, 12(1), 19. <https://doi.org/10.1186/s40163-023-00194-3>
- Weisburd, D., & Eck, J. E. (2004). What Can Police Do to Reduce Crime, Disorder, and Fear? *Annals of the American Academy of Political and Social Science*, 593(4), 42–65. <https://doi.org/10.1177/0002716203262548>

BIBLIOGRAPHY

- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>
- Aly, A. A., & Litwhiler, D. W. (1979). Police Briefing Stations: A Location Problem. *A I I E Transactions*, 11(1), 12–22. <https://doi.org/10.1080/05695557908974395>
- Andresen, M. A., & Brantingham, P. J. (2007). Hot spots of crime in Vancouver and their relationship with population characteristics. *CrimRxiv*. <https://doi.org/10.21428/cb6ab371.8b4a90b7>
- 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 and Criminology*, 109.
- Ballinger, C. (2011). Why Inferential Statistics are Inappropriate for Development Studies and How the Same Data Can be Better Used. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1775002>
- Banton, M. (1964). *The policeman in the community*. Tavistock.
- Barker, M., Geraghty, J., Webb, B., & Key, T. (1993). *The Prevention of Street Robbery*, Home Office Police Department.
- Beccaria, C. (1764). *On Crimes and Punishments and Other Writings* (R. Bellamy, Ed.; R. Davies, V. Cox & R. Bellamy, Trans.; Vol. 1995). Cambridge University Press.
- Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2), 169–217. <https://doi.org/10.1086/259394>
- Belshaw, S. H., Nodland, B., Pratt, T., & Leeper Piquero, N. (2022). Skimmed at the Pump: How COVID-19 Lockdowns Increased Gas Pump Skimming. *Crime & Delinquency*, 68(8), 1207–1222. <https://doi.org/10.1177/00111287221081028>
- Bennell, C., & Blaskovits, B. (2019). The case for open police research. In G. Tong, B. Blaskovits & N. Schneider (Eds.), *Evidence-based policing: An introduction* (pp. 117–131). Routledge.
- Bennett, D. S. (2018, November). *Police Response Times to Calls for Service: Fragmentation, Community Characteristics, and Efficiency* [Doctoral dissertation, Stanford Institute for Economic Policy Research].
- Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation* (J. H. Burns & H. L. A. Hart, Eds.; Vol. 1996). Clarendon Press.
- Berk, R. A., Campbell, A., Klap, R., & Western, B. (1992). The Deterrent Effect of Arrest in Incidents of Domestic Violence: A Bayesian Analysis of Four Field Experiments. *American Sociological Review*, 57(5), 698. <https://doi.org/10.2307/2095923>
- Berk, R. A., Western, B., & Weiss, R. E. (1995). Statistical Inference for Apparent Populations. *Sociological Methodology*, 25, 421. <https://doi.org/10.2307/271073>

- Berkow, M. 2011. Lecture to the National Police Academy.
- Bernasco, W., & Steenbeek, W. (2017). More Places than Crimes: Implications for Evaluating the Law of Crime Concentration at Place. *Journal of Quantitative Criminology*, 33(3), 451–467. <https://doi.org/10.1007/s10940-016-9324-7>
- Bertram, D. K., & Vargo, A. (1976). Response time analysis study - preliminary findings on robbery in Kansas City. *Police Chief*, 43(5), 74–77.
- Bieck, W. H. (1977a). *Response Time Analysis: Executive summary* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1977b). *Response Time Analysis: Part I Crime Analysis (Vol. II)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1977c). *Response Time Analysis: Synopsis* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1978). *Response Time Analysis: Methodology (Vol. I)* (tech. rep.). National Institute of Justice. Kansas City.
- Bieck, W. H. (1980a). *Response Time Analysis: Noncrime Analysis (Vol. IV)* (tech. rep.). National Institute of Justice. Kansas City, MO.
- Bieck, W. H. (1980b). *Response Time Analysis: Part II Crime Analysis (Vol. III)* (tech. rep.). National Institute of Justice. Kansas City.
- Bilach, T. J., Roche, S. P., & Wawro, G. J. (2022). The effects of the Summer All Out Foot Patrol Initiative in New York City: a difference-in-differences approach. *Journal of Experimental Criminology*, 18(2), 209–244. <https://doi.org/10.1007/s11292-020-09445-8>
- Binnenlandsbestuur.nl. 2024. Veranderende verhoudingen in de lokale driehoek. <https://www.binnenlandsbestuur.nl/juridisch/de-politie-een-duizenddingendoekje-geworden> Boonstra, Wouter.
- Birks, D., & Davies, T. (2017). Street Network Structure and Crime Risk: An Agent-Based Investigation of the Encounter and Enclosure Hypotheses. *Criminology*, 55(4), 900–937. <https://doi.org/10.1111/1745-9125.12163>
- Birks, D., & Townsley, M. (2018). Agent-based decision-support systems for crime scientists. In R. Wortley, A. Sidebottom, N. Tilley & G. Laycock (Eds.), *Routledge handbook of crime science* (pp. 334–349). Routledge. <https://doi.org/https://doi.org/10.4324/9780203431405-25>
- Birks, D., Townsley, M., & Stewart, A. (2012). Generative Explanations of Crime: Using Simulation to Test Criminological Theory. *Criminology*, 50(1), 221–254. <https://doi.org/10.1111/J.1745-9125.2011.00258.X>
- Birks, D., Townsley, M., & Stewart, A. (2014). Emergent Regularities of Interpersonal Victimization. *Journal of Research in Crime and Delinquency*, 51(1), 119–140. <https://doi.org/10.1177/0022427813487353>
- Bittner, E. (1990). *Aspects of police work*. Northeastern University Press.
- Blackstone, E. A., Hakim, S., & Meehan, B. (2020). Burglary reduction and improved police performance through private alarm response. *International Review of Law and Economics*, 63. <https://doi.org/10.1016/j.irl.2020.105930>
- Blake, L., & Coupe, R. T. (2001). The Impact of Single and Two-Officer Patrols on Catching Burglars in the Act: A Critique of the Audit Commission's Reports on Youth Justice.

- British Journal of Criminology*, 41(2), 381–396. <https://doi.org/10.1093/bjc/41.2.381>
- Blanes i Vidal, J., & Kirchmaier, T. (2018). The Effect of Police Response Time on Crime Clearance Rates. *The Review of Economic Studies*, 85(2), 855–891. <https://doi.org/10.1093/restud/rdx044>
- Blesse, S., & Diegmann, A. (2022). The place-based effects of police stations on crime: Evidence from station closures. *Journal of Public Economics*, 207, 104605. <https://doi.org/10.1016/J.JPUBECO.2022.104605>
- Boehme, H. M., Martin, A., & Kaminski, R. J. (2022). Evaluating the 4th Circuit's decision to limit officer use of Tasers: a descriptive and time-series approach. *Police Practice and Research*, 23(5), 523–538. <https://doi.org/10.1080/15614263.2021.1982713>
- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>
- Bosse, T., & Gerritsen, C. (2008). Agent-based simulation of the spatial dynamics of crime: on the interplay between criminal hot spots and reputation. In L. Padgham, D. C. Parkes, J. Mueller, S. Parsons & L. Cao (Eds.), *Proceedings of the seventh international joint conference on autonomous agents and multi-agent systems (aamas'08)* (pp. 1129–1136). ACM Press. <https://doi.org/10.5555/1402298.1402378>
- Boulton, L., McManus, M., Metcalfe, L., Brian, D., & Dawson, I. (2017). Calls for police service. *The Police Journal: Theory, Practice and Principles*, 90(1), 70–85. <https://doi.org/10.1177/0032258X16671032>
- Boydston, J. E., Sherry, M. E., & Moelter, N. P. (1977). *Patrol staffing in San Diego: One- or two-officer units*. (tech. rep.). National Policing Institute. Washington, DC.
- Braga, A. A. (2008). *Police Enforcement Strategies to Prevent Crime in Hot Spot Areas*. (tech. rep.). U.S. Department of Justice. Washington, D.C.
- Braga, A. A., & Dusseault, D. (2018). Can Homicide Detectives Improve Homicide Clearance Rates? *Crime & Delinquency*, 64(3), 283–315. <https://doi.org/10.1177/0011128716679164>
- Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2010). The Concentration and Stability of Gun Violence at Micro Places in Boston, 1980–2008. *Journal of Quantitative Criminology*, 26(1), 33–53. <https://doi.org/10.1007/s10940-009-9082-x>
- Braga, A. A., Schnell, C., & Welsh, B. C. (2024). Disorder policing to reduce crime: An updated systematic review and meta-analysis. *Criminology & Public Policy*, 23(3), 745–775. <https://doi.org/10.1111/1745-9133.12667>
- Braga, A. A., Turchan, B., & Barao, L. (2019). The Influence of Investigative Resources on Homicide Clearances. *Journal of Quantitative Criminology*, 35(2), 337–364. <https://doi.org/10.1007/s10940-018-9386-9>
- Braga, A. A., Turchan, B., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/cl2.1046>
- Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis. *Journal of Experimental Criminology*, 15(3), 289–311. <https://doi.org/10.1007/s11292-019-09372-3>

- Braga, A. A., Weisburd, D., & Turchan, B. (2019). Focused deterrence strategies effects on crime: A systematic review. *Campbell Systematic Reviews*, 15(3). <https://doi.org/10.1002/c12.1051>
- Brandl, S. G., & Horvath, F. (1991). Crime-victim evaluation of police investigative performance. *Journal of Criminal Justice*, 19(3), 293–305. [https://doi.org/10.1016/0047-2352\(91\)90008-J](https://doi.org/10.1016/0047-2352(91)90008-J)
- Brantingham, P. J., & Tita, G. (2008). Offender Mobility and Crime Pattern Formation from First Principles. In *Artificial crime analysis systems* (pp. 193–208). IGI Global. <https://doi.org/10.4018/978-1-59904-591-7.ch010>
- Brantingham, P. L., & Brantingham, P. J. (2017). Environment, Routine, and Situation: Toward a Pattern Theory of Crime. *Routine Activity and Rational Choice*, 259–294. <https://doi.org/10.4324/9781315128788-12>
- Brewer, D. D., Potterat, J. J., Muth, S. Q., & Robert, J. M. (2006). A Large Specific Deterrent Effect of Arrest for Patronizing a Prostitute. *PLOS ONE*, 1(1), e60. <https://doi.org/10.1371/JOURNAL.PONE.0000060>
- Brown, B. B., & Harris, P. B. (1989). Residential burglary victimization: Reactions to the invasion of a primary territory. *Journal of Environmental Psychology*, 9(2), 119–132. [https://doi.org/10.1016/S0272-4944\(89\)80003-9](https://doi.org/10.1016/S0272-4944(89)80003-9)
- Brown, W. J., & Lipsett, F. (1976). Response speeds and response times of urban police patrol cars in Ottawa, Canada. *Journal of Criminal Justice*, 4(3), 191–202. [https://doi.org/10.1016/0047-2352\(76\)90002-7](https://doi.org/10.1016/0047-2352(76)90002-7)
- Burney, E. (1990). *Putting street crime in its place: a report to the Community/Police Consultative group for Lambeth*. Goldsmith's College.
- Caldwell, R. (1971). *Optimal distribution policy for mobile police patrols* [Doctoral dissertation, Durham University].
- Chaiken, J. M. (1978). What Is Known About Deterrent Effects of Police Activities. In J. A. Cramer (Ed.), *Preventing crime* (pp. 109–135, Vol. 10). Sage Publications.
- Chaiken, J. M., Greenwood, P. W., & Petersilia, J. R. (1977). The criminal investigation process: A summary report. *Policy Analysis*, 3(2), 187–217.
- Chaiken, J. M., Lawless, M., & Stevenson, K. (1975). The Impact of Police Activity on Crime: Robberies on the New York City Subway System. *Urban Analysis*, 3, 173–205.
- Chaiken, J. M., Lawless, M., & Stevenson, K. A. (1974, January). *The Impact of Police Activity on Crime: Robberies on the New York City Subway System* (tech. rep.). The New York Rand Institute. Santa Monica, CA, RAND Corporation.
- Chenevoy, N. L. (2022). *Determining optimal police patrol deployments: a simulation-based optimisation approach combining agent-based modelling and genetic algorithms* [Doctoral dissertation, University of Leeds].
- Choi, K. S., Librett, M., & Collins, T. J. (2014). An empirical evaluation: gunshot detection system and its effectiveness on police practices. *Police Practice and Research*, 15(1), 48–61. <https://doi.org/10.1080/15614263.2013.800671>
- Chow, A. H. F., Cheung, C. Y., & Yoon, H. T. (2015). Optimization of Police Facility Locationing. *Transportation Research Record: Journal of the Transportation Research Board*, 2528(1), 60–68. <https://doi.org/10.3141/2528-07>
- Cihan, A. (2014). Social disorganization and police performance to burglary calls: A tale of two cities. *Policing*, 37(2), 340–354. <https://doi.org/10.1108/PIJPSM-06-2013-0058>

- Cihan, A. (2015). Examining the neighborhood effects on police performance to assault calls. *Police Practice and Research*, 16(5), 391–401. <https://doi.org/10.1080/15614263.2014.927765>
- Cihan, A., Zhang, Y., & Hoover, L. (2012). Police Response Time to In-Progress Burglary. *Police Quarterly*, 15(3), 308–327. <https://doi.org/10.1177/1098611112447753>
- Clawson, C., & Chang, S. K. (1977). The relationship of response delays and arrest rates. *Journal of Police Science and Administration*, 5(1), 53–68.
- Cohen, J. (2013, May). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- 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>
- Colton, K. W. (1980). *Police and Computer Technology: The Case of the San Diego Computer-Aided Dispatch System* (tech. rep. No. 1).
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin.
- Cordner, G. W., Greene, J. R., & Bynum, T. S. (1983). *The Sooner the Better: Some Effects of Police Response Time* (tech. rep.). National Institute of Justice. Washington, DC.
- Cornish, D. B., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending* (D. B. Cornish & R. V. Clarke, Eds.). Springer-Verlag.
- Coupe, R. T. (2016). Evaluating the effects of resources and solvability on burglary detection. *Policing and Society*, 26(5), 563–587. <https://doi.org/10.1080/10439463.2014.989155>
- Coupe, R. T. (2019, January). Investigative Activities, Resources and Burglary Detection. In *Crime solvability factors* (pp. 367–389). Springer International Publishing. https://doi.org/10.1007/978-3-030-17160-5_{_}16
- Coupe, R. T., & Blake, L. (2005). The effects of patrol workloads and response strength on arrests at burglary emergencies. *Journal of Criminal Justice*, 33(3), 239–255. <https://doi.org/10.1016/j.jcrimjus.2005.02.004>
- Coupe, R. T., & Blake, L. (2011). The effects of target characteristics on the sighting and arrest of offenders at burglary emergencies. *Security Journal*, 24(2), 157–178. <https://doi.org/10.1057/sj.2009.23>
- Coupe, R. T., & Girling, A. (2001). Modelling police success in catching burglars in the act. *Omega*, 29(1), 19–27. [https://doi.org/10.1016/S0305-0483\(00\)00030-X](https://doi.org/10.1016/S0305-0483(00)00030-X)
- Coupe, R. T., & Griffiths, M. (2000). Catching Offenders in the Act: An Empirical Study of Police Effectiveness in Handling 'Immediate Response' Residential Burglary. *International Journal of the Sociology of Law*, 28(2), 163–176. <https://doi.org/10.1006/ijsl.2000.0120>
- Coyne, C. J., Hall, A. R., McLaughlin, P. A., & Zerkle, A. (2014). A hidden cost of war: the impact of mobilizing reserve troops on emergency response times. *161*(3), 289–303. <https://doi.org/10.1007/sl>
- Curtin, K. M., Hayslett-McCall, K., & Qiu, F. (2010). Determining optimal police patrol areas with maximal covering and backup covering location models. *Networks and Spatial Economics*, 10(1), 125–145. <https://doi.org/10.1007/S11067-007-9035-6>

- Dahmann, J. S. (1975, November). *Examination of police patrol effectiveness - high impact anti-crime program* (tech. rep.). National Institute of Law Enforcement and Criminal Justice. Washington.
- Dai, M., Xia, Y., & Han, R. (2022). Temporal Variations in Calls for Police Service During COVID-19: Evidence From China. *Crime & Delinquency*, 68(8), 1183–1206. <https://doi.org/10.1177/00111287221074942>
- Dau, P. M., Dewinter, M., Vander Beken, T., Witlox, F., & Vandeviver, C. (2023). Police equilibrium: a glocal assessment of the demand and supply of police. *Policing and Society*, 33(5), 537–554. <https://doi.org/10.1080/10439463.2022.2149750>
- Di Tella, R., Diaz, F., Quintana, J., Buffa, S., Canepa, M., Castillos, J., Di Tella, R., & Schargrodsky, E. (2009). Criminal Recidivism after Prison and Electronic Monitoring. <https://doi.org/10.3386/W15602>
- Dunnett, S., Leigh, J., & Jackson, L. (2019). Optimising police dispatch for incident response in real time. *Journal of the Operational Research Society*, 70(2), 269–279. <https://doi.org/10.1080/01605682.2018.1434401>
- Eck, J. E. (1993). The Threat of Crime Displacement. *Problem Solving Quarterly*.
- Elder, R. W., Voas, R., Beirness, D., Shults, R. A., Sleet, D. A., Nichols, J. L., & Compton, R. (2011). Effectiveness of ignition interlocks for preventing alcohol-impaired driving and alcohol-related crashes: a Community Guide systematic review. *American Journal of Preventive Medicine*, 40(3), 362–376. <https://doi.org/https://doi.org/10.1016/j.amepre.2010.11.012>
- Ellison, M., Bannister, J., Lee, W. D., & Haleem, M. S. (2021). Understanding policing demand and deployment through the lens of the city and with the application of big data. *Urban Studies*, 58(15), 3157–3175. <https://doi.org/10.1177/0042098020981007>
- Farrington, D. P. (1983). Randomized Experiments on Crime and Justice. *Crime and Justice*, 4, 257–308. <https://doi.org/10.1086/449091>
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In D. P. Farrington, D. C. Gottfredson, L. W. Sherman & B. C. Welsh (Eds.), *Evidence-based crime prevention* (1st ed., pp. 13–21). Routledge.
- Felson, M. (1994). *Crime and everyday life*. Sage.
- Fienberg, S., Larntz, K., & Reiss, A. J. (1976). Redesigning the Kansas City preventive patrol experiment. *Evaluation*, 3(1-2), 124–131.
- Fondevila, G., Vilalta-Perdomo, C., Galindo Pérez, M. C., & Cafferata, F. G. (2021). Crime deterrent effect of police stations. *Applied Geography*, 134, 102518. <https://doi.org/10.1016/J.APGEOG.2021.102518>
- Fox, B. H., & Farrington, D. P. (2015). An Experimental Evaluation on the Utility of Burglary Profiles Applied in Active Police Investigations. *Criminal Justice and Behavior*, 42(2), 156–175. <https://doi.org/10.1177/0093854814548446>
- Frank, J., Smith, B. W., & Novak, K. J. (2005). Exploring the Basis of Citizens' Attitudes Toward the Police. *Police Quarterly*, 8(2), 206–228. <https://doi.org/10.1177/1098611103258955>
- Frydl, K., & Skogan, W. (Eds.). (2004, April). *Fairness and Effectiveness in Policing*. The National Academies Press. <https://doi.org/10.17226/10419>
- Gates, D., & Knowles, L. (1976). Evaluation of the Rand Corporation's Analysis of the Criminal Investigation Process. *The Police Chief*, 43(7), 20–77.

- Gibbs, B. G., Shafer, K., & Dufur, M. J. (2015). Why infer? The use and misuse of population data in sport research. *International Review for the Sociology of Sport*, 50(1), 115–121. <https://doi.org/10.1177/1012690212469019>
- Gibbs, B. G., Shafer, K., & Miles, A. (2017). Inferential statistics and the use of administrative data in US educational research. *International Journal of Research & Method in Education*, 40(2), 214–220. <https://doi.org/10.1080/1743727X.2015.1113249>
- Gibbs, J. P. (1968). Crime, punishment, and deterrence. *The Southwestern Social Science Quarterly*.
- Gill, C., Weisburd, D., Telep, C. W., Vitter, Z., & Bennett, T. (2014). Community-oriented policing to reduce crime, disorder and fear and increase satisfaction and legitimacy among citizens: a systematic review. *Journal of Experimental Criminology*, 10(4), 399–428. <https://doi.org/10.1007/s11292-014-9210-y>
- Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*, 5(10), 3. <https://doi.org/10.2307/1174772>
- Goldenberg, A., Rattigan, D., Dalton, M., Gaughan, J. P., Thomson, J. S., Remick, K., Butts, C., & Hazelton, J. P. (2019). Use of ShotSpotter detection technology decreases prehospital time for patients sustaining gunshot wounds. *Journal of Trauma and Acute Care Surgery*, 87(6), 1253–1259. <https://doi.org/10.1097/TA.0000000000002483>
- Goodchild, M. F. (2009). What Problem? Spatial Autocorrelation and Geographic Information Science. *Geographical Analysis*, 41(4), 411–417. <https://doi.org/10.1111/j.1538-4632.2009.00769.x>
- Gottfredson, M. R., & Hirschi, T. (1990). A General Theory of Crime. *A General Theory of Crime*. <https://doi.org/10.1515/9781503621794>
- Greenwood, P., Chaiken, J., & Peterselia, J. (1976). Response to: An evaluation of the RAND Corporation's analysis of the criminal investigation process". *The Police Chief*, 43(12), 62–71.
- Groff, E., & Birks, D. (2008). Simulating Crime Prevention Strategies: A Look at the Possibilities. *Policing*, 2(2), 175–184. <https://doi.org/10.1093/police/pan020>
- Groff, E. R. (2007). Simulation for theory testing and experimentation: An example using routine activity theory and street robbery. *Journal of Quantitative Criminology*, 23(2), 75–103. <https://doi.org/10.1007/S10940-006-9021-Z>
- Groff, E. R. (2008). Adding the Temporal and Spatial Aspects of Routine Activities: A Further Test of Routine Activity Theory. *Security Journal*, 21(1-2), 95–116. <https://doi.org/10.1057/palgrave.sj.8350070>
- Groff, E. R., Weisburd, D., & Yang, S.-M. (2010). Is it Important to Examine Crime Trends at a Local “Micro” Level?: A Longitudinal Analysis of Street to Street Variability in Crime Trajectories. *Journal of Quantitative Criminology*, 26(1), 7–32. <https://doi.org/10.1007/s10940-009-9081-y>
- Guerette, R. T., & Bowers, K. J. (2009). Assessing the extent of crime displacement and diffusion of benefits: A review of situational crime prevention evaluations. *Criminology*, 47(4), 1331–1368. <https://doi.org/10.1111/j.1745-9125.2009.00177.x>
- Halford, E., & Smith, J. (2022). Operation provide: a multi-agency response to increasing police engagement in cases of intimate partner violence during the COVID-19

- pandemic. *Police Practice and Research*, 23(5), 600–613. <https://doi.org/10.1080/15614263.2022.2033621>
- Hayes, C. (1987). The Impact of Recent Research on the Detective Role. *The Police Journal: Theory, Practice and Principles*, 60(2), 97–111. <https://doi.org/10.1177/0032258X8706000202>
- Hinkle, J. C., Weisburd, D., Telep, C. W., & Petersen, K. (2020). Problem-oriented policing for reducing crime and disorder: An updated systematic review and meta-analysis. *Campbell Systematic Reviews*, 16(2). <https://doi.org/10.1002/cl2.1089>
- Hobbes, T. (1651). *Leviathan* (E. Curley, Ed.; Modern Edition, Vol. 2012). Hackett Publishing.
- Hodgkinson, T., & Andresen, M. A. (2019). Understanding the Spatial Patterns of Police Activity and Mental Health in a Canadian City. *Journal of Contemporary Criminal Justice*, 35(2), 221–240. <https://doi.org/10.1177/1043986219842014>
- Holliday, L. P. (1974). *A methodology for radio car planning* (tech. rep.). RAND. Rand Institute.
- Huey, L., & Bennell, C. (2017). Replication and Reproduction in Canadian Policing Research: A Note. *Canadian Journal of Criminology and Criminal Justice*, 59(1), 123–138. <https://doi.org/10.3138/cjccj.2016.E09>
- Huey, L., & Johnston, S. (2023). Unf@cking People's Problems: A Theory of Policing. *CrimRxiv*. <https://doi.org/10.21428/CB6AB371.389F4506>
- Huff, J., Katz, C. M., & Hedberg, E. C. (2022). A randomized controlled trial of the impact of body-worn camera activation on the outcomes of individual incidents. *Journal of Experimental Criminology*, 18(2), 247–272. <https://doi.org/10.1007/s11292-020-09448-5>
- Hughes, L. A., Schaible, L. M., & Kephart, T. (2022). Gang Graffiti, Group Process, and Gang Violence. *Journal of Quantitative Criminology*, 38(2), 365–384. <https://doi.org/10.1007/s10940-021-09507-8>
- Hume, D. (1739). *A Treatise of Human Nature* (D. F. Norton & M. J. Norton, Eds.; Vol. 2000). Oxford University Press.
- Isaacs, H. H. (1967). *A Study of Communications, Crimes and Arrest in a Metropolitan Police Department. Appendix B of Science and Technology Task Force Report*. (tech. rep.). U.S. Government. Washington DC.
- Jagtenberg, C. J., Bhulai, S., & van der Mei, R. D. (2015). An efficient heuristic for real-time ambulance redeployment. *Operations Research for Health Care*, 4, 27–35. <https://doi.org/10.1016/j.orhc.2015.01.001>
- Jagtenberg, C. J., Bhulai, S., & van der Mei, R. D. (2017). Dynamic ambulance dispatching: is the closest-idle policy always optimal? *Health Care Management Science*, 20(4), 517–531. <https://doi.org/10.1007/s10729-016-9368-0>
- Jansen, R. P., Ruiter, S., & van Steden, R. (2024). Crime reporting and victim satisfaction with the police: a large-scale study among victims of crime in the Netherlands. *Crime Science*, 13(1), 1–11. <https://doi.org/10.1186/S40163-024-00231-9>
- Johnson, S. D., Tilley, N., & Bowers, K. J. (2015). Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology*, 11(3), 459–473. <https://doi.org/10.1007/S11292-015-9238-7>

- Joslin, J. D., Goldberger, D., Johnson, L., & Waltz, D. P. (2016). Use of the Vocera Communications Badge Improves Public Safety Response Times. *Emergency Medicine International*, 2016, 1–3. <https://doi.org/10.1155/2016/7158268>
- Kayes, M. I., Al-Deek, H., Sandt, A., Carrick, G., Staves, C., Gamaleldin, G., Faruk, O., & Al-Sahili, O. (2019). Characteristics of Law Enforcement Response to Wrong-Way Driving Events in Florida. *Transportation Research Record*, 2673(9), 514–524. <https://doi.org/10.1177/0361198119846456>
- Kelling, G. L., Pate, A. M., Dieckman, D., & Brown, C. (1974). *The Kansas City Preventive Patrol Experiment: A Summary Report* (tech. rep.). Police Foundation. Washington D.C.
- Kelling, G. L., Pate, A. M., Dieckman, D., & Brown, Charles. (1974). *The Kansas City Preventive Patrol Experiment Technical report*. Police Foundation.
- Kessler, D. A. (1985). One-or two-officer cars? A perspective from Kansas City. *Journal of Criminal Justice*, 13(1), 49–64. [https://doi.org/10.1016/0047-2352\(85\)90025-X](https://doi.org/10.1016/0047-2352(85)90025-X)
- Kester, J. 2023. Bijna helft slachtoffers misdrijf doet geen aangifte bij politie: weinig vertrouwen in pakkans daders. <https://eenvandaag.avrotros.nl/panels/opiniepanel/alle-uitslagen/item/bijna-helft-slachtoffers-misdrijf-doe-geen-aangifte-bij-politie-weinig-vertrouwen-in-pakkans-daders/EenVandaag>.
- Klein, M. W. (1986). Labeling Theory and Delinquency Policy: An Experimental Test. *Criminal Justice and Behavior*, 13(1), 47–79. <https://doi.org/10.1177/0093854886013001004>
- Klockars, C. B. (1983). *Thinking about Police: Contemporary Readings*. McGraw-Hill.
- Koper, C. S. (1995). Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, 12(4), 649–672. <https://doi.org/10.1080/07418829500096231>
- Koziarski, J., Ferguson, L., & Huey, L. (2022). Shedding Light on the Dark Figure of Police Mental Health Calls for Service. *Policing: A Journal of Policy and Practice*, 16(4), 696–706. <https://doi.org/10.1093/police/paac006>
- Krippendorff, K. (2011). *Computing Krippendorff's alpha-reliability* (tech. rep.). University of Pennsylvania.
- Kwak, N. K., & Leavitt, M. B. (1984). Police Patrol Beat Design: Allocation of Effort and Evaluation of Expected Performance. *Decision Sciences*, 15(3), 421–433. <https://doi.org/10.1111/j.1540-5915.1984.tb01227.x>
- Langton, S., Bannister, J., Ellison, M., Haleem, M. S., & Krzemieniewska-Nandwani, K. (2021). Policing and Mental ill-health: Using Big Data to Assess the Scale and Severity of, and the Frontline Resources Committed to, mental ill-health-related calls-for-service. *Policing: A Journal of Policy and Practice*, 15(3), 1963–1976. <https://doi.org/10.1093/police/paab035>
- Langton, S., Ruiter, S., & Verlaan, T. (2023). Describing the scale and composition of calls for police service: a replication and extension using open data. *Police Practice and Research*, 24(5), 523–538. <https://doi.org/10.1080/15614263.2022.2102494>
- Langton, S., Verlaan, T., & Ruiter, S. (2023). Operationalizing deployment time in police calls for service. *Crime Science*, 12(1), 20. <https://doi.org/10.1186/s40163-023-00198-z>
- Laplace, P. (1781). Mémoire sur les probabilités. *Mémoires de l'Académie Royale des Sciences de Paris*, 1778, 227–332.

- Larsen, K. S., Andersen, L. H., & Larsen, B. Ø. (2022). Pushed out of the education system: using a natural experiment to evaluate consequences for boys. *Journal of Experimental Criminology*, 18(2), 411–438. <https://doi.org/10.1007/s11292-020-09444-9>
- Larson, G. C., & Simon, J. W. (1979, February). *Evaluation of a police AVM system: A study of the St. Louis experience 1976-1977* (tech. rep. No. 1). National Institute of Law Enforcement and Criminal Justice.
- Larson, R. C. (1974). A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, 1(1), 67–95. [https://doi.org/10.1016/0305-0548\(74\)90076-8](https://doi.org/10.1016/0305-0548(74)90076-8)
- Larson, R. C. (1975). What happened to patrol operations in Kansas City? A review of the Kansas City preventive patrol experiment. *Journal of Criminal Justice*, 3(4), 267–297. [https://doi.org/10.1016/0047-2352\(75\)90034-3](https://doi.org/10.1016/0047-2352(75)90034-3)
- Larson, R. C. (1982). Critiquing Critiques: Another Word on the Kansas City Preventive Patrol Experiment. *Evaluation Review*, 6(2), 285–293. <https://doi.org/10.1177/0193841X8200600209>
- Larson, R. C., & Cahn, M. F. (1985). *Synthesizing and Extending the Results of Police Patrol Studies* (tech. rep.). National Institute of Justice.
- Leal, W. E., Piquero, A. R., Kurland, J., Piquero, N. L., & Gloyd, E. L. (2022). A Case Study of Family Violence During COVID-19 in San Antonio. *Crime & Delinquency*, 68(8), 1161–1182. <https://doi.org/10.1177/00111287211064781>
- Lee, J. S., Lee, J., & Hoover, L. T. (2017). What Conditions Affect Police Response Time? Examining Situational and Neighborhood Factors. *Police Quarterly*, 20(1), 61–80. <https://doi.org/10.1177/1098611116657327>
- Leigh, J., Dunnett, S., & Jackson, L. (2019). Predictive police patrolling to target hotspots and cover response demand. *Annals of Operations Research*, 283(1-2), 395–410. <https://doi.org/10.1007/S10479-017-2528-X>
- Lindegaard, M. R., & Bernasco, W. (2018). Lessons Learned from Crime Caught on Camera. *Journal of Research in Crime and Delinquency*, 55(1), 155–186. <https://doi.org/10.1177/0022427817727830>
- Locke, J. (1689). *Two Treatises of Government* (P. Laslett, Ed.; Vol. 1988). Cambridge University Press.
- Loeffler, C. E., & Nagin, D. S. (2021). The Impact of Incarceration on Recidivism. *Annual Review of Criminology*, 5, 133–152. <https://doi.org/10.1146/annurev-criminol-030920-112506>
- Lum, C., & Koper, C. S. 2021. Guest post: Shifting mental health calls away from police will not greatly affect their workload. *The Washington Post*.
- Lum, C., Koper, C. S., Stoltz, M., Goodier, M., Johnson, W., Prince, H., & Wu, X. (2020). Constrained Gatekeepers of the Criminal Justice Footprint: A Systematic Social Observation Study of 9-1-1 Calltakers and Dispatchers. *Justice Quarterly*, 1–23. <https://doi.org/10.1080/07418825.2020.1834604>
- Lum, C., Koper, C. S., & Wu, X. (2022). Can We Really Defund the Police? A Nine-Agency Study of Police Response to Calls for Service. *Police Quarterly*, 25(3), 255–280. <https://doi.org/10.1177/10986111211035002>

- Lum, C., Koper, C. S., Wu, X., Johnson, W., & Stoltz, M. (2020). Examining the Empirical Realities of Proactive Policing Through Systematic Observations and Computer-Aided Dispatch Data. *Police Quarterly*, 23(3), 283–310. <https://doi.org/10.1177/1098611119896081>
- Malleson, N. (2012). Using Agent-Based Models to Simulate Crime. In *Agent-based models of geographical systems* (pp. 411–434). Springer Netherlands. https://doi.org/10.1007/978-90-481-8927-4{_}19
- Mares, D., & Blackburn, E. (2012). Evaluating the Effectiveness of an Acoustic Gunshot Location System in St. Louis, MO. *Policing*, 6(1), 26–42. <https://doi.org/10.1093/police/par056>
- Mares, D., & Blackburn, E. (2021). Acoustic gunshot detection systems: a quasi-experimental evaluation in St. Louis, MO. *Journal of Experimental Criminology*, 17(2), 193–215. <https://doi.org/10.1007/s11292-019-09405-x>
- Mastrobuoni, G. (2019). Police disruption and performance: Evidence from recurrent redeployments within a city. *Journal of Public Economics*, 176, 18–31. <https://doi.org/10.1016/j.jpubeco.2019.05.003>
- Maxfield, M. G. (1982). Service Time, Dispatch Time, and Demand for Police Services: Helping More by Serving Less. *Public Administration Review*, 42(3), 252. <https://doi.org/10.2307/976012>
- Mazerolle, L., Rogan, D., Frank, J., Famega, C., & Eck, J. E. (2003). *Managing Citizen Calls to the Police: An Assessment of Non-Emergency Call Systems*. President, Statistical Analysis for Law Enforcement Strategies (tech. rep.).
- Mazerolle, L. G., Watkins, C., Rogan, D., & Frank, J. (1998). Using Gunshot Detection Systems in Police Departments. *Police Quarterly*, 1(2), 21–49. <https://doi.org/10.1177/109861119800100202>
- Mcewen, J. T., Connors, E. F., & Cohen, M. I. (1986). *Evaluation of the Differential Police Response Field Test* (tech. rep.).
- McNeeley, S., & Warner, J. J. (2015). Replication in criminology: A necessary practice. *European Journal of Criminology*, 12(5), 581–597. <https://doi.org/10.1177/1477370815578197>
- Miron, H., Wasserman, R., & Rickard, T. (1979). Managing criminal investigations: a handbook.
- Mitchell, P. S. (1972). Optimal Selection of Police Patrol Beats. *The Journal of Criminal Law, Criminology, and Police Science*, 63(4), 577. <https://doi.org/10.2307/1141814>
- Mladenka, K. R., & Hill, K. Q. (1978). The Distribution of Urban Police Services. *The Journal of Politics*, 40(1), 112–133. <https://doi.org/10.2307/2129978>
- Mohammadi, M., Martiniuk, A. L. C., Ansari-Moghaddam, A., Rad, M., Rashedi, F., Ghjasemi, A., & Ahangari, H. (2013). Police response time to road crashes in south-east of Iran. *Journal of the Pakistan Medical Association*, 63, 1523–1527.
- Montesquieu. (1748). *The Spirit of the Laws* (A. M. Cohler, B. C. Miller & H. S. Stone, Eds.; Vol. 1989). Cambridge University Press.
- Moyer, R. A. (2022). The Effect of a Death-in-Police-Custody Incident on Community Reliance on the Police. *Journal of Quantitative Criminology*, 38(2), 459–482. <https://doi.org/10.1007/s10940-021-09504-x>

- Myrick, J. A. (1979). *Evaluation of the role of Police in the EMS system* (tech. rep.). Georgia Institute of Technology. Atlanta, Georgia.
- Nagin, D. S. (2013). Deterrence in the Twenty-First Century. *Crime and Justice*, 42(1), 199–263. <https://doi.org/10.1086/670398>
- Nationale Politie. (2025, March). *Strategische Agenda Politie 2025-2030: Stevig Staen in Deze Tijd* (tech. rep.).
- O'Brien, D. T., Ciomek, A., & Tucker, R. (2022). How and Why is Crime More Concentrated in Some Neighborhoods than Others?: A New Dimension to Community Crime. *Journal of Quantitative Criminology*, 38(2), 295–321. <https://doi.org/10.1007/s10940-021-09495-9>
- O'Brien, D. T., Ristea, A., Hangen, F., & Tucker, R. (2022). Different places, different problems: profiles of crime and disorder at residential parcels. *Crime Science*, 11(1), 4. <https://doi.org/10.1186/s40163-022-00165-0>
- Pate, A. M. (1981). *The Newark Foot Patrol Experiment* (tech. rep.). Police Foundation. Washington.
- Pate, A. M., Ferrara, A., Bowers, R. A., & Lorence, J. (1976, January). *Police Response Time: Its Determinants and Effects* (tech. rep.). National Policing Institute. Washington, DC.
- Pate, A. M., Kelling, G. L., & Brown, C. (1975). A response to “what happened to patrol operations in Kansas City?” *Journal of Criminal Justice*, 3(4), 299–320. [https://doi.org/10.1016/0047-2352\(75\)90035-5](https://doi.org/10.1016/0047-2352(75)90035-5)
- Pawson, R., & Tilley, N. (1997). *Realist evaluation*. SAGE Publications Ltd.
- 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). <https://doi.org/10.1002/cl2.1302>
- Pew Research Center. 2024. Demographics of Mobile Device Ownership and Adoption in the United States. <https://www.pewresearch.org/internet/fact-sheet/mobile/>
- Philpot, R., Liebst, L. S., Levine, M., Bernasco, W., & Lindegaard, M. R. (2020). Would I be helped? Cross-national CCTV footage shows that intervention is the norm in public conflicts. *American Psychologist*, 75(1), 66–75. <https://doi.org/10.1037/amp0000469>
- Philpot, R., Liebst, L. S., Møller, K. K., Lindegaard, M. R., & Levine, M. (2019). Capturing violence in the night-time economy: A review of established and emerging methodologies. *Aggression and Violent Behavior*, 46, 56–65. <https://doi.org/10.1016/j.avb.2019.02.004>
- Piquero, A. R., & Blumstein, A. (2007). Does Incapacitation Reduce Crime? *Journal of Quantitative Criminology*, 23(4), 267–285. <https://doi.org/10.1007/s10940-007-9030-6>
- Press, S. J. (1971). *Some effects of an increase in police manpower in the 20th precinct of New York city* (tech. rep.). New York City-Rand Institute. New York City. <https://doi.org/20.500.12592/nxwwjj>
- Prince, H., Lum, C., & Koper, C. S. (2021). Effective police investigative practices: an evidence-assessment of the research. *Policing: An International Journal*, 44(4), 683–707. <https://doi.org/10.1108/PIJPSM-04-2021-0054>

- Propheter, G. (2020). Do urban sports facilities have unique social costs? An analysis of event-related congestion on police response time. *International Journal of Urban Sciences*, 24(2), 271–281. <https://doi.org/10.1080/12265934.2019.1625805>
- Ratcliffe, J. H. (2021). Policing and public health calls for service in Philadelphia. *Crime Science*, 10(1), 5. <https://doi.org/10.1186/s40163-021-00141-0>
- Ratcliffe, J. H., Groff, E. R., Sorg, E. T., & Haberman, C. P. (2015). Citizens' reactions to hot spots policing: Impacts on perceptions of crime, disorder, safety and police. *Journal of Experimental Criminology*, 11(3), 393–417. <https://doi.org/10.1007/s11292-015-9230-2>
- Ratcliffe, J. H., Taniguchi, T., Groff, E. R., & Wood, J. D. (2011). The Philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, 49(3), 795–831. <https://doi.org/10.1111/j.1745-9125.2011.00240.x>
- Ratcliffe, J. H., Taylor, R. B., Askey, A. P., Thomas, K., Grasso, J., Bethel, K. J., Fisher, R., & Koehnlein, J. (2021). The Philadelphia predictive policing experiment. *Journal of Experimental Criminology*, 17(1), 15–41. <https://doi.org/10.1007/s11292-019-09400-2>
- Regioburgemeesters.nl. 2024. Afscheidsinterview korpschef Henk van Essen - 'Politieke problemen dreigen steeds meer politieproblemen te worden'. <https://www.regioburgemeesters.nl/actueel/?id=901>
- Reiss, A. J. (1971). *The police and the public*. Yale University Press.
- Reiss, A. J., & Bordua, D. J. (1964). Environment and organization: A perspective on the police.
- Reppetto, T. A. (1976). Crime Prevention and the Displacement Phenomenon. *Crime & Delinquency*, 22(2), 166–177. <https://doi.org/10.1177/001112877602200204>
- Riddell, J. R., Piquero, A. R., Kaukinen, C., Bishopp, S. A., Piquero, N. L., Narvey, C. S., & Iesue, L. (2022). Re-Opening Dallas: A Short-Term Evaluation of COVID-19 Regulations and Crime. *Crime & Delinquency*, 68(8), 1137–1160. <https://doi.org/10.1177/00111287211054718>
- Rief, R., & Huff, J. (2023). Revisiting the influence of police response time: Examining the effects of response time on arrest and how it varies by call type. *Journal of Criminal Justice*, 84(84), 1–11. <https://doi.org/10.1016/j.jcrimjus.2022.102025>
- Rijksoverheid. 2012. Politiewet, Artikel 3. *Overheid.nl*.
- Rinehart Kochel, T. (2011). Constructing Hot Spots Policing: Unexamined Consequences for Disadvantaged Populations and for Police Legitimacy. *Criminal Justice Policy Review*, 22(3), 350–374. <https://doi.org/10.1177/0887403410376233>
- Risman, B. J. (1980). The Kansas City Preventive Patrol Experiment: A Continuing Debate. *Evaluation Review*, 4(6), 802–808. <https://doi.org/10.1177/0193841X8000400605>
- Robbins, M. W., & Davenport, S. (2021). microsynth: Synthetic Control Methods for Disaggregated and Micro-Level Data in R. *Journal of Statistical Software*, 97(2). <https://doi.org/10.18637/jss.v097.i02>
- Robbins, M. W., Saunders, J., & Kilmer, B. (2017). A Framework for Synthetic Control Methods With High-Dimensional, Micro-Level Data: Evaluating a Neighborhood-Specific Crime Intervention. *Journal of the American Statistical Association*, 112(517), 109–126. <https://doi.org/10.1080/01621459.2016.1213634>

- Roman, J. K., Reid, S. E., Chalfin, A. J., Knight, C. R., Roman, J. K., Reid, S. E., & Chalfin, A. J. (2009). The DNA field experiment: a randomized trial of the cost-effectiveness of using DNA to solve property crimes. *Journal of Experimental Criminology* 2009 5:4, 5(4), 345–369. <https://doi.org/10.1007/S11292-009-9086-4>
- Rosenbaum, D. P. (2006, May). The limits of hot spots policing. In *Police innovation* (pp. 245–264). Cambridge University Press. <https://doi.org/10.1017/CBO9780511489334.013>
- Rousseau, J.-J. (1762). *The Social Contract and Other Later Political Writings* (V. Gourevitch, Ed.; Vol. 1997). Cambridge University Press.
- RTLNieuws. 2025. Politie haalt in 307 gemeenten eigen norm voor aanrijtijden bij noodgeval-
len niet. <https://www.rtl.nl/nieuws/onderzoek/artikel/5497691/politie-brandweer-aanrijtijden-gemeenten-spoed-noodgeval> Jasper Bunskoek.
- Russo, C. W. (2006). *AVL and Response Time Reduction: Image and Reality* [Doctoral dissertation, University of Central Florida].
- Sacks, S. R. (2000). Optimal Spatial Deployment of Police Patrol Cars. *Social Science Computer Review*, 18(1), 40–55. <https://doi.org/10.1177/089443930001800103>
- Salerno-Ferraro, A. C., & Jung, S. (2022). To charge or not to charge? Police decisions in Canadian sexual assault cases and the relevance of rape myths. *Police Practice and Research*, 23(5), 539–552. <https://doi.org/10.1080/15614263.2021.2002150>
- Salimbene, N. A., & Zhang, Y. (2020). An examination of organizational and community effects on police response time. *Policing: An International Journal*, 43(6), 935–946. <https://doi.org/10.1108/PIJPSM-04-2020-0063>
- Scherhag, J., & Burgard, T. (2023). Performance of Semi-Automated Screening Using Rayyan and ASReview: A Retrospective Analysis of Potential Work Reduction and Different Stopping Rules. *Big Data & Research Syntheses*.
- Schnelle, J. F., Kirchner, R. E., Casey, J. D., Uselton, P. H., & Mcnees, M. P. (1977). Patrol evaluation research: A multiple-baseline analysis of saturation police patrolling during day and night hours. *Journal of applied behavior analysis*, 10, 33–40.
- Shalev, M. (2007). Limits and Alternatives to Multiple Regression in Comparative Research. In *Capitalisms compared* (pp. 261–308). [https://doi.org/10.1016/S0195-6310\(06\)24006-7](https://doi.org/10.1016/S0195-6310(06)24006-7)
- Sherman, L. W. (1986). Policing Communities: What Works? *Crime and Justice*, 8, 343–386. <https://doi.org/10.1086/449127>
- Sherman, L. W. (2013). The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking. *Crime and Justice*, 42(1), 377–451. <https://doi.org/10.1086/670819>
- Sherman, L. W. (2017). Evidence-based policing.
- Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American sociological review*, 49(2), 261–272. <https://doi.org/10.2307/2095575>
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place*. *Criminology*, 27(1), 27–56. <https://doi.org/10.1111/j.1745-9125.1989.tb00862.x>
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997, January). *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress* (tech. rep.). NIJ. Washington, DC.

- Sherman, L. W., & Rogan, D. P. (1995). Effects of gun seizures on gun violence: “Hot spots” patrol in Kansas City. *Justice Quarterly*, 12(4), 673–693. <https://doi.org/10.1080/07418829500096241>
- Sherman, L. W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime “hot spots”: A randomized, controlled trial. *Justice Quarterly*, 12(4), 625–648. <https://doi.org/10.1080/07418829500096221>
- Simpson, R. (2021). Calling the Police: Dispatchers as Important Interpreters and Manufacturers of Calls for Service Data. *Policing: A Journal of Policy and Practice*, 15(2), 1537–1545. <https://doi.org/10.1093/police/paaa040>
- Skolnick, J. H., & Bayley, D. H. (1986). *The new blue line: Police innovation in six American cities*. Simon; Schuster.
- Spelman, W., & Brown, D. K. (1981). *Calling the Police: Citizen Reporting of Serious Crime* (tech. rep.). National Institute of Justice. Washington, DC.
- Spickenheuer, H. (1983). *Voetsurveillance en preventievoorzichting in Amsterdam-Osdrorp: een onderzoek naar de effecten op de criminaliteitsontwikkeling*. Staatsuitgeverij.
- Stevens, J. M., Webster, T. C., & Stipak, B. (1980). *Response Time: Role in Assessing Police Performance* (tech. rep. No. 3).
- Sullivan, E. (2012). *The Effects of Crime Incident Characteristics and Neighborhood Structure on Police Response Time* [Doctoral dissertation, Arizona State University].
- Tarr, D. P. (1978). Analysis of response delays and arrest rates. *Journal of Police Science & Administration*, 6(4), 429–451.
- Tay, R. (2005). General and Specific Deterrent Effects of Traffic Enforcement: Do we have to Catch Offenders to Reduce Crashes? *Journal of Transport Economics and Policy*, 39(2), 209–224.
- Taylor, P. E., & Huxley, S. J. (1989). A Break from Tradition for the San Francisco Police: Patrol Officer Scheduling Using an Optimization-Based Decision Support System. *Interfaces*, 19(1), 4–24. <https://doi.org/10.1287/inte.19.1.4>
- Telep, C. W., Weisburd, D., Gill, C. E., Vitter, Z., & Teichman, D. (2014). Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, 10(4), 515–548. <https://doi.org/10.1007/s11292-014-9208-5>
- Terpstra, J., Foekens, P., & van Stokkum, P. (2015, July). *Burgemeesters over hun Nationale Politie* (tech. rep.). Radboud Universiteit Nijmegen. Den Haag.
- Textor, J., van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. *International Journal of Epidemiology*, 45(6), 1887–1894. <https://doi.org/10.1093/IJE/DYW341>
- Tillyer, R. (2022). Unpacking Sequential Actions Within Use of Force Incidents. *Police Quarterly*, 25(2), 178–200. <https://doi.org/10.1177/10986111211049549>
- Tricco, A. C., Lillie, E., Zarin, W., O’Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>

- Trojanowicz, R. C. (1986). Evaluating a neighborhood foot patrol program: The Flint, Michigan project. In D. Rosenbaum (Ed.), *Community crime prevention: Does it work?* (1st ed., p. 157). Sage.
- U.S. Congress, H. o. R. (1995). *Making appropriations for the Departments of Commerce, Justice, and State, the Judiciary, and related agencies for the fiscal year ending September 30, 1996, and for other purposes (House Report 104-378, Sec. 116)* (tech. rep.). Washington.
- van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdem, 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 Barneveld, T., Bhulai, S., & van der Mei, R. D. (2016). The effect of ambulance relocations on the performance of ambulance service providers. *European Journal of Operational Research*, 252(1), 257–269. <https://doi.org/10.1016/j.ejor.2015.12.022>
- van Barneveld, T., Jagtenberg, C., Bhulai, S., & van der Mei, R. (2018). Real-time ambulance relocation: Assessing real-time redeployment strategies for ambulance relocation. *Socio-Economic Planning Sciences*, 62, 129–142. <https://doi.org/10.1016/j.seps.2017.11.001>
- van Bruchem, M., Proost, K., van Ruysseveldt, J., & Lindegaard, M. R. (2024). Bystander actions during police work on the street: officer perspectives. *Police Practice and Research*. <https://doi.org/10.1080/15614263.2024.2363335>
- van Dijk, A., & Crofts, N. (2017). Law enforcement and public health as an emerging field. *Policing and Society*, 27(3), 261–275. <https://doi.org/10.1080/10439463.2016.1219735>
- van Gemert, J. (2025). Verward gedrag van een naaste kan veel zorgen veroorzaken, ervaart ook Joke bij zoon Kees: 'Familie vaak eerste die problemen ziet'. <https://eenvandaag.avrotros.nl/item/verward-gedrag-van-een-naaste-kan-veel-zorgen-veroorzaken-ervaart-ook-joke-bij-zoon-kees-familie-vaak-eerste-die-problemen-ziet/EenVandaag>.
- Vaughan, A. D., Hewitt, A. N., Andresen, M. A., & Brantingham, P. L. (2016). Exploring the Role of the Environmental Context in the Spatial Distribution of Calls-for-Service Associated with Emotionally Disturbed Persons. *Policing*, 10(2), 121–133. <https://doi.org/10.1093/police/pav040>
- Verlaan, T., & Langton, S. (2024, April). On the Use of Inferential Statistics on Administrative Police Data. In *The crime data handbook* (pp. 197–210). Bristol University Press. <https://doi.org/10.51952/9781529232073.ch014>
- Verlaan, T., Langton, S., & Ruiter, S. (2023). Reactieve politie-inzet op straat. *Tijdschrift voor de Politie*, (4), 36–40.
- Verlaan, T., & Ruiter, S. (2023a). *PROTOCOL-Determinants and effects of police response time: a scoping review*. <https://doi.org/10.17605/OSF.IO/HU2E9>
- Verlaan, T., & Ruiter, S. (2023b). Predictors of police response time: a scoping review. *Crime Science*, 12(1), 19. <https://doi.org/10.1186/s40163-023-00194-3>
- Vigoa, O. F. (2010). *Comparative Analysis between Routine Preventative Uniformed Comparative Analysis between Routine Preventative Uniformed Police Patrols and Crime*

- Reduction and Calls for Service Police Patrols and Crime Reduction and Calls for Service* (tech. rep.).
- Weatherburn, D., & Moffatt, S. (2011). The Specific Deterrent Effect of Higher Fines on Drink-Driving Offenders. *The British Journal of Criminology*, 51(5), 789–803. <https://doi.org/10.1093/BJC/AZR043>
- Weber, M. (1919). *The Vocation Lectures: Science as a Vocation; Politics as a Vocation* (D. Owen & T. B. Strong, Eds.; R. Livingstone, Trans.; Vol. 2004). Hackett Publishing.
- Weisburd, D. (2015). The Law of Crime Concentration and the Criminology of Place. *Criminology*, 53(2), 133–157. <https://doi.org/10.1111/1745-9125.12070>
- Weisburd, D., Braga, A. A., Groff, E. R., & Wooditch, A. (2017). Can Hot Spots Policing Reduce Crime in Urban Areas? An Agent-based Simulation. *Criminology*, 55(1), 137–173. <https://doi.org/10.1111/1745-9125.12131>
- Weisburd, D., & Eck, J. E. (2004). What Can Police Do to Reduce Crime, Disorder, and Fear? *Annals of the American Academy of Political and Social Science*, 593(4), 42–65. <https://doi.org/10.1177/0002716203262548>
- Weisburd, D., & Green, L. (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly*, 12(4), 711–735. <https://doi.org/10.1080/07418829500096261>
- Weisburd, D., Hinkle, J. C., Famega, C., & Ready, J. (2011). The possible “backfire” effects of hot spots policing: an experimental assessment of impacts on legitimacy, fear and collective efficacy. *Journal of Experimental Criminology*, 7(4), 297–320. <https://doi.org/10.1007/s11292-011-9130-z>
- Weisburd, D., & Majimundar, M. K. (Eds.). (2018, February). *Proactive Policing*. National Academies Press. <https://doi.org/10.17226/24928>
- Weisburd, D., Petersen, K., Telep, C. W., & Fay, S. A. (2024). Can increasing preventive patrol in large geographic areas reduce crime? A systematic review and meta-analysis. *Criminology & Public Policy*, 23(3), 721–743. <https://doi.org/10.1111/1745-9133.12665>
- Weisburd, D., Wilson, D. B., Petersen, K., & Telep, C. W. (2023). Does police patrol in large areas prevent crime? Revisiting the Kansas City Preventive Patrol Experiment. *Criminology & Public Policy*, 22(3), 543–560. <https://doi.org/10.1111/1745-9133.12623>
- Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, 103(2), 280–293. https://doi.org/10.1162/rest{_}_a{_}_00889
- White, C., & Weisburd, D. (2018). A Co-Responder Model for Policing Mental Health Problems at Crime Hot Spots: Findings from a Pilot Project. *Policing: A Journal of Policy and Practice*, 12(2), 194–209. <https://doi.org/10.1093/police/pax010>
- Wilensky, U. 1999. Netlogo. <http://ccl.northwestern.edu/netlogo/>.
- Wilson, D. B. (2022). The Relative Incident Rate Ratio Effect Size for Count-Based Impact Evaluations: When an Odds Ratio is Not an Odds Ratio. *Journal of Quantitative Criminology*, 38(2), 323–341. <https://doi.org/10.1007/s10940-021-09494-w>
- Wood, J. D., Watson, A. C., & Barber, C. (2021). What can we expect of police in the face of deficient mental health systems? Qualitative insights from Chicago

- police officers. *Journal of Psychiatric and Mental Health Nursing*, 28(1), 28–42. <https://doi.org/10.1111/jpm.12691>
- Wooditch, A. (2023). The Benefits of Patrol Officers Using Unallocated Time for Everyday Crime Prevention. *Journal of Quantitative Criminology*, 39(1), 161–185. <https://doi.org/10.1007/S10940-021-09527-4>
- World Bank. 2023. Fixed telephone subscriptions (per 100 people). <https://data.worldbank.org/indicator/IT.MLT.MAIN.P2> *World Bank Open Data*.
- Wu, X., Koper, C., & Lum, C. (2022). Measuring the Impacts of Everyday Police Proactive Activities: Tackling the Endogeneity Problem. *Journal of Quantitative Criminology*, 38(2), 343–363. <https://doi.org/10.1007/s10940-021-09496-8>
- Zhang, Y., & Brown, D. E. (2012). Police patrol district design using agent-based simulation and GIS. *2012 IEEE International Conference on Intelligence and Security Informatics*, 54–59. <https://doi.org/10.1109/ISI.2012.6284091>
- Zhao, J., Schneider, M., & Thurman, Q. (2002). The effect of police presence on public fear reduction and satisfaction: A review of the literature. *The Justice Professional*, 15(3), 273–299. <https://doi.org/10.1080/0888431021000049471>
- Zhu, R., Hu, X., Li, X., Ye, H., & Jia, N. (2020). Modeling and Risk Analysis of Chemical Terrorist Attacks: A Bayesian Network Method. *International Journal of Environmental Research and Public Health*, 17(6), 2051. <https://doi.org/10.3390/ijerph17062051>
- Zimring, F. E. (1978). Policy Experiments in General Deterrence: 1970-1975. In *Deterrence and incapacitation estimating the effects of criminal sanctions on crime rates* (pp. 140–186). The National Academies Press. <https://doi.org/10.17226/18561-2>