How Generalizable are Findings from Police Surveys? A Review of Multi-Agency Studies*

Erin M. Kearns
School of Criminology and Criminal Justice
National Counterterrorism Innovation, Technology and Education (NCITE) Center
University of Nebraska at Omaha, Omaha, NE
ORCiD: 0000-0002-7895-9129; Twitter: @KearnsErinM

Justin Nix
School of Criminology and Criminal Justice
University of Nebraska at Omaha, Omaha, NE
ORCiD: 0000-0002-3812-8590; Twitter: @jnixy

Forthcoming at: *Police Practice and Research* (Special Issue on Replicating & Reproducing Policing Research)

DOI: 10.1080/15614263.2023.2214282

Corresponding Author: Erin Kearns, PhD. Email: ekearns@unomaha.edu.

Word Count: 7,018

Funding: None.

Disclosure Statement: The authors declare no conflict of interest.

Acknowledgements: We thank Hyunin Baek, Carol Ann Sharo, Brandon Tregle, and Taylor Windsheimer for providing invaluable research assistance on this project. We are grateful to the studies' original authors for providing additional analyses so we could conduct the comparisons found in this manuscript. We are also grateful to Dick Adams who, when we requested additional analyses from his 2002 article, tracked down the floppy disks on which the data were saved and enlisted the help of a colleague to convert the SAS files to SPSS. Unfortunately, the data and syntax could not be recovered – but the willingness to help is commendable.

Abstract

Policing scholars frequently use surveys to understand officer attitudes and behavioral intentions. Yet, it is difficult to gain access to one – let alone multiple – agencies. Thus, officer surveys often reflect views in a single department, making it unclear how generalizable the findings are. For the present study, we conducted an exploratory review of articles published in 16 criminology and policing journals from 2000 to 2017. We identified 600 studies that involved surveying one or more samples of police officers. From this list, we set out to determine: (1) how often authors administered their surveys to more than one sample, and (2) when surveys *were* administered to more than one sample, how often were results consistent across samples? We found eighty-seven (14.5%) articles that involved collecting survey data from multiple agencies, though only 29 (4.8% overall, 33.3% of multi-agencies studies) met our inclusion criteria. Importantly, only 15 studies could be analyzed as some authors no longer had data, could not share data, or did not response to our emails. Results were fully consistent across samples in just one published study. In the other studies, findings partially replicated—though sometimes results were in the opposite direction across departments. Thus, replication is critical before policy is created from single-agency surveys.

Key Words: policing; survey; attitudes; external validity; replication

Introduction

In the last few decades, policing scholars have increasingly used survey research methods to understand officers' perceptions and attitudes (Nix et al., 2019). For example, through officer surveys, researchers have gained a better understanding of what motivates people to pursue a career in policing (White et al., 2010), how officers view their agency once they have started working (Wolfe & Nix, 2017), and perceptions of support seeking and willingness to engage mental health service among law enforcement officers (Karaffa & Koch, 2016; Lambert & Steinke, 2015). Similarly, scholars have examined variation in officer perceptions of policing crime types including traffic violations (Bates et al, 2015), sexual assault (Wentz & Archbold, 2012), terrorism (Kearns, 2018; Muibu, 2022), and interpersonal violence (Garcia et al., 2014). Further, amidst tense relationships between law enforcement and the public – particularly over the last decade – policing survey studies have provided insights on officer views on topics including citizen review (Lee et al., 2017), community policing (Kearns, 2017), use of force (Ingram et al., 2014), de-policing (Nix et al., 2018), and body-worn cameras (Gaub et al., 2016).

Studies using officer surveys certainly contribute to what we know about policing. Yet, it often remains unknown how generalizable findings are to other agencies. To collect data from officers in any given department, researchers need the chief's permission. The reality, unfortunately, is that it is often difficult to gain access to one – let alone multiple – agencies. Police can be distrustful of scholars and may not see the benefit of participating in academic research (Reiner, 2000; Skogan, 2015). As a result, it has become quite common for scholars to publish their findings from surveys administered to single agencies. Complicating matters, there is growing consensus across multiple disciplines that replication is vital to the scientific endeavor (Murayama et al., 2014). Yet, replication is rare in criminology and criminal justice – with

replication studies making up between 0.45% (Pridemore et al., 2018) and 2.34% of published articles (McNeeley & Warner, 2015).

Presumably scholars shy away from replications studies because they are more difficult to publish. For example, Galiani et al. (2017, p. 2) found that "In a survey of editors, almost all responded they would in principle publish a replication study that overturned the results of the original study, but only 29% responded that they would consider publishing a replication study that confirmed the original study results." This is problematic with over 17,000 state and local law enforcement agencies of various sizes, geographies, populations served, political landscapes, et cetera in the United States where – presumably – what matters in a major police department on the East Coast might not translate to a small-town department in the Midwest. Such heterogeneity might result in "hidden moderators" (see e.g., Van Bavel et al., 2016; c.f. Inbar, 2016; Olsson-Collentine et al., 2020) preventing results from a study of one agency from fully replicating in a separate study of any other agency. Indeed, it is well-documented that police organizations can influence the subculture – and the behaviors – of the officers they employ (Chan, 1996; Crank, 1990; Fyfe, 1988; Wilson, 1968; White, 2001). Accordingly, the unknown generalizability of single-agency surveys deserves more attention than a throwaway "limitation" line in the discussion section.

We endeavored to accomplish the following in our exploratory study: (1) identify the percentage of studies using police surveys published between 2000 and 2017 that collected data from more than one department and (2) identify how often the results were consistent across the agencies among the studies that did include multiple departments. First, we outline the importance of replication in criminology and criminal justice research – particularly that on policing and officer perceptions. We next describe the sample frame, data, coding, and analytical approach.

Finally, we conclude with a discussion of the findings and their limitations, policy implications stemming from these findings, and future research directions.

Background

Particularly over the last decade, the "open science movement" has pushed to make scientific research more transparent and accessible within scientific communities and across broader society. Amidst this, organizations like the Center for Open Science have been created with the mission to "increase openness, integrity, and reproducibility of research."[1] One core tenant of reproducible scientific research is that scholars can and should *replicate* their studies. The push to replicate social science research is nothing new. Nearly 30 years ago, Gary King (1995) called for replication in political science with his recommendations on how the field – from faculty and students to funders and editors – could incentivize and implement this process. More recently, psychology has faced a so-called "replication crisis" as the field grappled with canonical papers failing to replicate and what that meant for the discipline (Bohannon, 2015; Francis, 2012; Maxwell et al., 2015; Shrout & Rogers, 2018).

As Pridemore et al. (2018) note, criminology has yet to face the same challenges to our knowledge bases through systematic or larger scale replications. In a more extensive look at the literature, Pridemore et al. (2018) found that only 0.45% of published articles in criminology were replication studies – a far smaller percentage than the 2.34% that McNeeley and Warner (2015) reported. Notably, three quarters of these replications were successful and another 15% produced mixed results, which is promising for the future of replication work in criminology and criminal justice. Further, replication has received more attention within the field over the last handful of years. For example, key points from a session on replication at the 2017 American Society of

[1] https://www.cos.io/about/mission

Criminology annual meeting now appear in the *Journal of Experimental Criminology* (Farrington et al., 2019), the *Journal of Contemporary Criminal Justice* published a special issue on replication in 2018 (Savolainen & VanEseltine, 2018), and *Police Practice and Research* is currently publishing a special issue on replication within policing studies.

Amidst the increasing emphasis on replication, David Weisburd (in Farrington et al., 2019, p. 388) noted that "replication is important to science because it improves our ability to define the impacts of programs or practices. But it is a mistake to begin with the idea that the replications will bring us certainty." Within the field, some areas of study may be more likely to produce replicable findings due to the nature of data available, methodological control, homogeneity of groups being studied, and relative stability in societal and cultural precursors over time. However, as both Robert Boruch and Friederich Lösel (in Farrington et al., 2019) state, there are numerous cultural, societal, political, temporal, and other changes over time that make it unlikely that findings will replicate over time and place. It is reasonable to expect that this may be particularly the case for policing survey research given: a) the influence that organizational culture can have on officer attitudes and behaviors (Shupard & Kearns, 2019; Silver et al., 2017), b) the mandate for police to be responsive to the unique needs of the community they serve; and c) the sheer heterogeneity in size, geography, demographics, et cetera across law enforcement agencies.

Blaskovits et al. (2018) aimed to conduct a full replication of Telep and Lum's (2014) examination of U.S. police officers' receptivity to empirical research using a Canadian sample. Findings across these two studies were similar in several ways, though Blaskovits et al.'s (2018) sample also indicated greater openness to evidence-based policing. This singular example of a full replication attempt within policing survey research is encouraging. However, more attempts to replicate policing survey research are certainly needed to better understand whether the similar

findings across agencies is a broader trend, limited to certain topics or issues, or a more anomalous finding. The current study aims to answer this question by identifying how often survey results were consistent across multiple agencies among studies published between 2000 and 2017 in 15 criminology and policing journals.

Methods

Sampling Frame and Data

To begin, we adopted Nix et al.'s (2019) sampling frame that included a total of 16 journals – eight policing-focused journals and eight prestigious general criminology and criminal justice journals. We searched the eight policing-focused journals listed on the American Society of Criminology's Division of Policing website (https://ascpolicing.org/members-corner/): *European Journal of Policing Studies, International Journal of Police Science & Management*, *Journal of Police and Criminal Psychology*, *Police Practice & Research*, *Police Quarterly*, *Policing & Society, Policing: A Journal of Police and Practice*, and *Policing: An International Journal*. We also searched eight prestigious general criminology journals (see Nix et al., 2019 for a full discussion): *British Journal of Criminology*, *Crime & Delinquency*, *Criminal Justice and Behavior*, *Criminology*, *Journal of Criminal Justice*, *Journal of Quantitative Criminology*, *Journal of Research in Crime and Delinquency*, and *Justice Quarterly*. As Nix and colleagues (2019) noted, online access to *European Journal of Policing Studies* was not available and thus they excluded this journal from their study. For the present study, we were able to find abstracts for all 127 articles published in the journal between 2012 and 2017 but could not find full text of many of the articles; thus, we also excluded this journal from our study.

Our first inclusion criterion was that studies involved survey data collection from police officers. We searched each of these journals for articles using survey data from police officers that

were published between 2000 and 2017 (including "online first"). We identified 600 articles that met this criterion. Our next inclusion criterion was that the study collected data from multiple departments, which eliminated most studies, leaving 87 (14.5% of total).

For our analytic sample, we focused on articles where comparisons of officer survey data was both logical and feasible. First, given cross-national variation in policing practices and organizational structures, we excluded articles (n = 14) where data were collected from officers in different countries. Second, several studies used samples of officers across many departments with only a small number of officers in any one department participating. Since department-level samples could not be identified in these cases, these studies were excluded from our sample (n = 11). Third, a few studies did systematically sample from multiple departments, but the sample sizes in at least one of the departments was too small for department-specific comparisons to be sufficiently powered. Thus, between-department comparisons were not statistically feasible, so these studies were also excluded (n = 10). Finally, many policing studies come from the Project on Policing Neighborhoods (POPN) data which includes officer interviews and ethnographic observations. We excluded POPN studies wherein the variables come from a source other than surveys, since survey data is the focus of this project (n = 23). In sum, we only included studies where: a) data came from officer surveys, including experimental, cross-sectional, and panel data, b) the researcher(s) systematically sampled officers in at least two departments, c) departments were in the same country, and d) department-specific analyses were feasible and sufficiently powered. See Figure 1 for a flow chart summarizing our sampling approach and exclusion criteria.

[FIGURE 1 HERE]

Taking all the above into account, our final analytic sample is comprised of the 29 studies (4.8% of total; 33.3% of studies including multiple departments) that met our inclusion criteria

across articles published in these journals between 2000 and 2017. Table 1 presents each of the journals included in this study and – for each journal – the number of articles published in each between 2000 and 2017 that surveyed police officers (criterion 1), the number of those articles that surveyed officers in multiple departments (criterion 2), and the number of articles that met all our inclusion criteria and thus were in the final sample.

[TABLE 1 HERE]

Procedure

Once we identified the 29 studies that met our inclusion criteria, each of the present authors separately reviewed each study to determine whether it provided sufficient data to compare results across departments. Seven (24.1%) of the published manuscripts did provide sufficient information for us to compare findings across departments while 22 (75.9%) did not. We next contacted the corresponding author for each of these 22 manuscripts via email to request the results disaggregated by department.[2] The body of those emails was as follows:

> Hi Dr. [LAST NAME],
>
> A colleague and I are examining studies that have collected survey data from officers in multiple police departments. Your study [TITLE] in [*JOURNAL*] is one of the articles that we would like to include. If possible, could you provide us with [INSERT WHAT IS NEEDED FROM THIS STUDY]? Please accept my apologies if I missed this information in the text.
>
> Thanks!
>
> [AUTHOR NAMES REDACTED]

Of the 22 authors whom we contacted, eight provided the data that we requested, 10 responded that they either no longer had the data or had agreed not to disaggregate it as a condition of IRB or department participation, and four did not respond to our initial or follow up email. Thus,

[2] In some cases, the corresponding author had moved to a different institution, so we contacted those individuals at their current place of employment.

we were able to compare the replicability of findings across departments for 15 studies (2.5% of total; 17.2% of studies on multiple departments; 51.7% of studies that meet our inclusion criteria).

Analytic Plan

After we either disaggregated each study's results based on information in the published manuscript or received disaggregated results from the study's corresponding author, we then identified four main components of each study. First, we identified the research question(s) for each study. Second, we noted key aspects of each sample including: the number of departments, the sample size for each department, the response rate, and any notable similarities or differences described across departments. Third, we identified the dependent and independent variables in each study. Finally, where findings were disaggregated by department, we examined if each result was a) consistent in significance & direction, b) inconsistent in significance only, or c) inconsistent in significance & direction. We then compared the findings to see where there are consistencies or inconsistencies—essentially examining the replicability of findings across agencies within a single study.

Full replication materials for the tables and figures in the present manuscript will be posted on the authors' websites upon publication. We will also post full replication materials for two of the 15 studies analyzed here ([REDACTED]; [REDACTED]), but we do not have permission to share the original replication materials for the other 13 articles analyzed in this manuscript.

Results

The main objective of this study was to examine the prevalence of consistent findings in surveys of police officers across departments. Of the 600 published studies using surveys of police that we identified, few (n=87; 14.5%) collected data from officers in multiple departments. Even

fewer studies (n=29; 4.8%) systematically collected officer-level data across multiple departments with samples large enough to compare statistically.

As shown in Figure 2, there has been an increase in the number of articles published using officer surveys over time – with a particular uptick in 2016 and 2017 – across the 18 years in our sample. In the years immediately preceding this, community-police relations were widely and hotly debated in the aftermath of high-profile incidents where police used deadly force against unarmed Black citizens across the United States (Nix & Wolfe, 2018), and concurrently, there were marked decreases in public confidence in police.[3] During the same period, President Obama convened the Task Force on 21st Century Policing (Ramsey & Robinson, 2015) and body-worn cameras were being increasingly adopted by law enforcement agencies around the country (see Jennings et al., 2014) among other shifts. Thus, the increased academic focus on understanding how police view the communities they serve, their responsibilities and changes to them, their departments, and other aspects of policing logically follows. Of note, however, is that a similar uptick in the percentage of studies that surveyed multiple agencies does not occur during this same time.

[FIGURE 2]

In Figure 3, we disaggregate by journal the number of articles published in each 2-year interval using a single versus multiple agencies. There are notable differences in the frequency that research on officer attitudes and behavioral intentions have been published across the 15 journals in our sample – seven of which are policing-focused and eight of which are general criminology and criminal justice journals. Perhaps unsurprisingly, most articles using officer survey data – regardless of whether a single agency or multiple agencies were included – have been published

[3] Jones, J.M. (2015). In U.S., confidence in police lowest in 22 years. Gallup. Retrieved from: http://www.gallup.com/poll/183704/confidence-police-lowest-years.aspx.

in policing-focused journals (end of second row and full third row). In policing-focused journals, there appears to be some variation in frequency of publishing officer survey papers both by outlet and over time. Across more general criminology and criminal justice focused journals, however, there seems to be more between-outlet variation than temporal. See the online appendix for the underlying data used to create Figure 3.

[FIGURE 3]

As Table 2 shows, there were 15 studies that met our inclusion criteria and that we could also either examine between-department differences based on the published results or the authors provided us with supplemental analyses upon request. Table 2 notes the topic of the study, number of departments surveyed, the sample sizes for each department, the range of response rates across departments, a brief description of the similarities and differences between departments, and how consistent the findings were across departments. Each of the 15 articles surveyed between two and four (M=2.67; SD=0.82) agencies. Broadly, most of the studies can be grouped into studying one of the following four topics: views on evidence-based policing, intelligence-led policing, police misconduct, or police culture. At the department-level, sample sizes ranged from N=59 to 523 (M=199.85; SD=125.13) and response rates ranged from 7.3% to 98.5%.

[TABLE 2 HERE]

Across the 15 survey-based studies of police officers in multiple agencies where between-department comparisons are possible, the analytic strategies vary both in the statistical techniques used and the number of analyses conducted. For example, one study presents only descriptive statistics (Telep & Lum, 2014) while another presents only exploratory factor analysis (Domingues & Machado, 2017). One includes ranking variables alongside mean comparisons (Westmarland & Rowe, 2018), while others present means comparisons across numerous variables (Darroch &

Mazerolle, 2015; Ivkovich & Khechumyan, 2013; Paoline, 2004; Sherwood, 2000), and others still rely primarily (Paoline et al., 2000) – or solely – upon regression (Chu & Sun, 2014; Darroch & Mazerolle, 2013; Gaub et al., 2016; Kearns, 2017; Nix et al., 2017; Sun & Chu, 2008; Telep, 2017). Accordingly, it is challenging to simply answer "how often do results replicate?" To address this, the far-right column of Table 2 notes each study's analytical approach and succinctly summarizes between-department comparisons of findings.

As shown in Table 2, results only completely replicated in one study (#15) which conducted exploratory factor analysis in two departments. Domingues and Machado (2017) surveyed officers in two departments to examine sources of officers' stress. Using exploratory factor analysis where data were both pooled and then disaggregated by department, the same three-factor model was suggested with similar factor loadings. In another study (#10), the findings nearly replicate fully. Nix et al. (2017) administered a survey experiment in two agencies to test the causal effects of suspect demeanor on police officers' willingness to use procedural justice. In both departments, officers indicated it was less important to treat disrespectful suspects with procedural fairness. However, whereas the authors used just one measure of demeanor in their first survey (disrespect), in the second survey they included separate measures for verbal and symbolic disrespect – with only the latter significantly reducing officers' perceived importance of using procedural justice.

Six studies rely solely (#4, #5, #8, #10, #11) or partially (#12) on regression analyses. Five of the six studies have at least one coefficient that is significant in the same direction across departments while one fails to replicate a single result across departments. All six studies also have coefficients that are in the same direction, even if they are statistically significant in some departments but not others. Further, in most (83.3%) of these studies, at least one of the coefficients

is also significant in multiple departments but in opposite directions. One study (#1) presented descriptive information about pre- and post-deployment attitudes toward body worn cameras and compared percent change in those attitudes, which varied across the three departments. Another study (#2) presented descriptive analyses, which are largely similar across departments though one agency scored higher in some areas. In study (#6), officers were asked to rank the seriousness of offenses, which is largely consistent across agencies. This study was also one of the seven (along with #3, #7, #9, #12, #13, #14) that used between-department mean comparisons across variables. In each of these studies, there were significant between-department differences for some variables but not others. Sometimes one agency had consistently different views, other times this was not the case.

In sum, the authors of these 15 studies more often than not found that at least some of their results replicated across departments. Very few police survey studies collect data from multiple departments and even fewer do so systematically and with large enough samples to compare findings between departments. Across the 15 studies where between-department comparisons of results were possible, only one study replicated completely while another nearly replicated fully. Findings replicated partially for most of the other studies, though sometimes results were significant in the opposite direction. Finally, one study failed to replicate a single significant finding across departments.

Discussion

At the onset of this project, we aimed to first identify the percentage of studies using police surveys published between 2000 and 2017 that collected data from more than one department and then identify how often the results were consistent across the agencies among the studies that did include multiple departments. Overall, a relatively small percentage (14.5%) of the published

articles collected data from multiple departments where replication would be possible. However, for nearly half of these articles, there was not enough information in the text for us to disaggregate the analyses ourselves and the author(s) were unable to provide those analyses or did not respond to our email request. While Blaskovits et al.'s (2018) replication of Telep and Lum (2014) was promising, only two of the 15 articles in our final analytic sample showed a similar degree of replication. Though results were mixed for most of the other studies – and there is reason to believe that replication may be lower in policing studies given the influence of organizational culture (Paoline, 2021) – this does highlight the need for both better reporting on articles that are published to enable future studies such as the present one and more replication efforts such as Blaskovits et al.'s (2018).

With the "open science movement" and emphasis on replication across social science disciplines (King, 1995; Simons, 2014), it is certainly time to take seriously how criminology and criminal justice will address these topics. We are still at the early stages of our disciplinary discussions on replication (Farrington, et al. 2019; Pridemore et al., 2018). Some concerns – like the so-called "file drawer problem" – are universal, so other disciplines can provide instructive guidance. For example, editors of economics journals are far more reluctant to publish replication studies that confirm the original findings than ones that overturn them, which disincentivizes researchers from pursuing these studies (Galiani et al., 2017). In a related vein, a recent survey of criminologists found that over half of those who responded to the survey – the vast majority of whom were mid-career or senior scholars – admit to underreporting results. Further, almost half admit to omitting non-significant studies or variables, which highlights the pressures to publish significant findings (Chin et al., 2021). As this pertains to replication concerns for policing survey research – and to the field more broadly – David Farrington (in Farrington et al., 2019, p. 379)

called for a "*Journal of Criminological Replication*" that would publish both successful and failed replication attempts and perhaps this is a suggestion that the discipline should take seriously.

We are also faced with some unique challenges due to the sensitive nature of some of the topics we study, and the access required to study them. As is true for many subfields within criminology and criminal justice, data collection from law enforcement often requires explicit access and permission from those in charge of the agency. Thus, it is difficult and time consuming to collect primary data from police officers in a single department. And, no doubt, it is markedly more challenging to carry out officer-level data collection across multiple departments. To suggest that this be the standard for research on officer attitudes and behavioral intentions would preclude scholars from undertaking valuable research. That said, when researchers are able to collect data from officers in multiple agencies for the same study, there are minimum reporting standards that they should follow about both the samples themselves (see Nix et al., 2019) and with the analyses pooled and disaggregated by department whenever possible (see Shupard & Kearns, 2019 for an example). The current prevailing practice of simply controlling for agency with dummy variable(s) can – and does – mask important differences between departments. To be clear, we are not suggesting that scholars should *not* control for agency in their regression analyses. Rather, we suggest that they include supplemental analyses that disaggregate samples by agency to examine whether their findings are invariant across departments, which contributes to the burgeoning replication efforts within policing research and criminology and criminal justice more broadly.

References

Bates, L. J., Antrobus, E., Bennett, S., & Martin, P. (2015). Comparing police and public perceptions of a routine traffic encounter. *Police Quarterly*, *18*(4), 442-468.

Blaskovits, B., Bennell, C., Huey, L., Kalyal, H., Walker, T., & Javala, S. (2018). A Canadian replication of Telep and Lum's (2014) examination of police officers' receptivity to empirical research. *Policing and Society*.

Bohannon, J. (2015). Many psychology papers fail replication test.

Chan, J. (1996). Changing police culture. *British Journal of Criminology*, *36*(1), 109-134.

Chin, J. M., Pickett, J. T., Vazire, S., & Holcombe, A. O. (2021). Questionable research practices and open science in quantitative criminology. *Journal of Quantitative Criminology*, 1-31.

Chu, D. C., & Sun, I. Y. (2014). Reactive versus proactive attitudes toward domestic violence: A comparison of Taiwanese male and female police officers. Crime & delinquency, 60(2), 216-237.

Crank, J. P. (1990). The influence of environmental and organizational factors on police style in urban and rural environments. *Journal of Research in Crime and Delinquency*, *27*(2), 166-189.

Darroch, S., & Mazerolle, L. (2013). Intelligence-led policing: A comparative analysis of organizational factors influencing innovation uptake. Police quarterly, 16(1), 3-37.

Darroch, S., & Mazerolle, L. (2015). Intelligence-led policing: a comparative analysis of community context influencing innovation uptake. Policing and society, 25(1), 1-24.

Domingues, I. P., & Machado, J. C. (2017). The loosely coupled factors of organizational stress in police forces. *Policing: An International Journal of Police Strategies & Management*.

Gaub, J. E., Choate, D. E., Todak, N., Katz, C. M., & White, M. D. (2016). Officer perceptions of body-worn cameras before and after deployment: A study of three departments. Police quarterly, 19(3), 275-302.

Farrington, D. P., Lösel, F., Boruch, R. F., Gottfredson, D. C., Mazerolle, L., Sherman, L. W., & Weisburd, D. (2019). Advancing knowledge about replication in criminology. *Journal of Experimental Criminology*, *15*(3), 373-396.

Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science, 7(*6), 585-594.

Fyfe, J. J. (1988). Police use of deadly force: Research and reform. *Justice Quarterly, 5*(2), 165-205.

Galiani, S., Gertler, P., & Romero, M. (2017). *Incentives for replication in economics* (No. w23576). National Bureau of Economic Research.

Gaub, J. E., Choate, D. E., Todak, N., Katz, C. M., & White, M. D. (2016). Officer perceptions of body-worn cameras before and after deployment: A study of three departments. *Police Quarterly*, *19*(3), 275-302.

Gracia, E., García, F., & Lila, M. (2014). Male police officers' law enforcement preferences in cases of intimate partner violence versus non-intimate interpersonal violence: Do sexist attitudes and empathy matter?. *Criminal Justice and Behavior*, *41*(10), 1195-1213.

Inbar, Y. (2016). Association between contextual dependence and replicability in psychology may be spurious. *Proceedings of the National Academy of Sciences*, *113*(34), E4933-E4934.

Ingram, J. R., Weidner, R. R., Paoline III, E. A., & Terrill, W. (2014). Supervisory influences on officers' perceptions of less lethal force policy: A multilevel analysis. *Policing: An International Journal of Police Strategies & Management*.

Ivkovich, S. K., & Khechumyan, A. (2013). The state of police integrity in Armenia: Findings from the police integrity survey. Policing: An International Journal of Police Strategies & Management.

Jennings, W. G., Fridell, L. A., & Lynch, M. D. (2014). Cops and cameras: Officer perceptions of the use of body-worn cameras in law enforcement. *Journal of Criminal Justice*, *42*(6), 549-556.

Karaffa, K. M., & Koch, J. M. (2016). Stigma, pluralistic ignorance, and attitudes toward seeking mental health services among police officers. *Criminal Justice and Behavior*, *43*(6), 759-777.

Kearns, E. M. (2017). Why are some officers more supportive of community policing with minorities than others? *Justice Quarterly*, *34*(7), 1213-1245.

Kearns, E. M. (2020). Exploring officer views of community policing in counterterrorism. *Police practice and research*, *21*(1), 18-32.

King, G. (1995). Replication, replication. *PS: Political Science & Politics*, *28*(3), 444-452.

Lambert, A. D., & Steinke, C. M. (2015). Negative perceptions of asking for support in law enforcement: Potential impact on benefit avoidance. *International Journal of Police Science & Management*, *17*(2), 134-144.

Lee, H. D., Collins, P. A., Hsieh, M. L., Boateng, F. D., & Brody, D. (2017). Officer attitudes toward citizen review and professional accountability. *International Journal of Police Science & Management*, *19*(2), 63-71.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean?. *American Psychologist, 70*(6), 487.

McNeeley, S., & Warner, J. J. (2015). Replication in criminology: A necessary practice. *European Journal of Criminology*, *12*(5), 581-597.

Muibu, D. (2022). Police Officer Experiences with Community Policing and Views on Counterterrorism in Somalia. *The Journal of the Middle East and Africa*, 1-28.

Murayama, K., Pekrun, R., & Fiedler, K. (2014). Research practices that can prevent an inflation of false-positive rates. *Personality and Social Psychology Review*, *18*(2), 107-118.

Nix, J., Pickett, J. T., Baek, H., & Alpert, G. P. (2019). Police research, officer surveys, and response rates. *Policing and Society*, *29*(5), 530-550.

Nix, J., Pickett, J. T., Wolfe, S. E., & Campbell, B. A. (2017). Demeanor, race, and police perceptions of procedural justice: Evidence from two randomized experiments. Justice quarterly, 34(7), 1154-1183.

Nix, J., Wolfe, S. E., & Campbell, B. A. (2018). Command-level police officers' perceptions of the "war on cops" and de-policing. *Justice Quarterly*, *35*(1), 33-54.

Olsson-Collentine, A., Wicherts, J. M., & van Assen, M. A. (2020). Heterogeneity in direct replications in psychology and its association with effect size. *Psychological Bulletin*, *146*(10), 922.

Paoline III, E. A. (2004). Shedding light on police culture: An examination of officers' occupational attitudes. Police quarterly, 7(2), 205-236.

Paoline, E. A. (2021). (Re)taking stock: Toward a richer understanding of police culture. In R. Dunham, G. Alpert, & K. McLean (Eds.), *Critical Issues in Policing: Contemporary Readings* (8th ed., pp. 159-186). Waveland Press.

Paoline III, E. A., Myers, S. M., & Worden, R. E. (2000). Police culture, individualism, and community policing: Evidence from two police departments. Justice quarterly, 17(3), 575-605.

Pridemore, W. A., Makel, M. C., & Plucker, J. A. (2018). Replication in criminology and the social sciences. *Annual Review of Criminology*, *1*, 19-38.

Ramsey, C., & Robinson, L. (2015). *Interim report of the President's Task Force on 21st Century Policing*. Washington, DC: Office of Community Oriented Policing Services.

Reiner, R., 2000. Police research. In Roy D. King and Emma Wincup, eds. *Doing research on crime and justice*. New York, NY: Oxford University Press.

Sherwood, C. W. (2000). Job design, community policing, and higher education: A tale of two cities. Police Quarterly, 3(2), 191-212.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology, 69*(1), 487-510.

Shupard, M., & Kearns, E. M. (2019). Why do officers support community policing? A cross-departmental and cross-temporal comparison. *Behavioral Sciences & the Law, 37*(6), 665-680.

Silver, J. R., Roche, S. P., Bilach, T. J., & Bontrager Ryon, S. (2017). Traditional police culture, use of force, and procedural justice: Investigating individual, organizational, and contextual factors. *Justice Quarterly, 34*(7), 1272–1309.

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, *9*(1), 76-80.

Skogan, W. G., 2015. Surveying police officers. In. Michale D. Maltz and Stephen K. Rice, eds. *Envisioning criminology: Researchers on research as a process of discovery.* New York, NY: Springer.

Sun, I. Y., & Chu, D. C. (2008). Gender differences in policing: An analysis of Taiwanese officers' attitudes. Police Practice and Research: An International Journal, 9(5), 431-443.

Telep, C. W. (2017). Police officer receptivity to research and evidence-based policing: Examining variability within and across agencies. Crime & delinquency, 63(8), 976-999.

Telep, C. W., & Lum, C. (2014). The receptivity of officers to empirical research and evidence-based policing: An examination of survey data from three agencies. *Police quarterly*, *17*(4), 359-385.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Reply to Inbar: Contextual sensitivity helps explain the reproducibility gap between social and cognitive psychology. *Proceedings of the National Academy of Sciences*, *113*(34), E4935-E4936.

Wentz, E., & Archbold, C. A. (2012). Police perceptions of sexual assault victims: Exploring the intra-female gender hostility thesis. *Police Quarterly*, *15*(1), 25-44.

Westmarland, L., & Rowe, M. (2018). Police ethics and integrity: can a new code overturn the blue code?. *Policing and society*, *28*(7), 854-870.

White, M. D. (2001). Controlling police decisions to use deadly force: Reexamining the importance of administrative policy. Crime & Delinquency, 47(1), 131-151

White, M. D., Cooper, J. A., Saunders, J., & Raganella, A. J. (2010). Motivations for becoming a police officer: Re-assessing officer attitudes and job satisfaction after six years on the street. *Journal of Criminal Justice*, *38*(4), 520-530.

Wilson, J. Q. (1968). *Varieties of police behavior: The management of law and order in eight communities*. Harvard University Press.

Wolfe, S. E., & Nix, J. (2017). Police officers' trust in their agency: Does self-legitimacy protect against supervisor procedural injustice?. *Criminal Justice and Behavior*, *44*(5), 717-732.

Table 1. List of Journal and Articles Included at Each Stage of Coding

	Criterion 1		Criterion 2		Final Sample		
	# of Articles using Surveys	**% of all Articles using Surveys**	**# of Articles using Surveys from Multiple Departments**	**% of Articles from Criterion 1**	**# of Articles**	**% of Articles from Criterion 1**	**% of Articles from Criterion 2**
British Journal of Criminology	6	1.0%	0	0.0%	0	0.0%	---
Crime & Delinquency	18	3.0%	5	27.8%	3	16.7%	60.0%
Criminal Justice and Behavior	24	4.0%	2	8.3%	0	0.0%	0.0%
Criminology	8	1.3%	6	75.0%	0	0.0%	0.0%
International Journal of Police Science & Management	69	11.5%	3	4.3%	0	0.0%	0.0%
Journal of Criminal Justice	37	6.2%	3	8.1%	1	2.7%	33.3%
Journal of Police and Criminal Psychology	55	9.2%	3	5.5%	2	3.6%	66.7%
Journal of Quantitative Criminology	0	0.0%	0	---	0	---	---
Journal of Research in Crime and Delinquency	2	0.3%	2	100%	0	0.0%	0.0%
Justice Quarterly	16	2.7%	9	56.3%	3	18.8%	33.3%
Police Practice & Research	59	9.8%	6	10.2%	2	3.4%	33.3%
Police Quarterly	87	14.5%	14	16.1%	8	9.2%	57.1%
Policing & Society	40	6.7%	6	15.0%	3	7.5%	50.0%
Policing: A Journal of Police and Practice	27	4.5%	1	3.7%	0	0.0%	0.0%
Policing: An International Journal	152	25.3%	27	17.8%	7	4.6%	25.9%
TOTAL ARTICLES	**600**	**100%**	**87**	**14.5%**	**29**	**4.8%**	**33.3%**

Table 2. Replicability of Findings across Multiple Departments

Article	Research Area	Sampling				Replicability of Findings
		# Depts	N	RR	Dept. Notes	
1	Perceptions of body worn cameras	3	106, 153, 205	77% - 92%	2 close, 1 different	Descriptive & percent change comparisons. Pre- and post-deployment officers in Phoenix were more negative about BWCs than officers in either Spokane or Tempe. Pre to post attitudes comparisons across departments show consistent significance & direction, same direction but not all significant, and significant in opposite direction
2	Familiarity with evidence-based policing	3	523, 343, 94	36% - 78%	2 states, 2 capitals	Descriptive analyses. Similar findings across departments, though smaller rural agency scored higher in some areas.
3	Relative uptake of intelligence-led policing	4	67, 75, 59, 85	70% - 82%	Strong v. weak innovation	Mean comparisons. No difference between strong uptake sites on some outcomes and differences on others. No differences regardless of innovation level on some outcomes as well.
4	Receptivity to evidence-based policing	4	523, 276, 71, 122	38% - 77.5%	Diverse in size of force, jurisdiction & geography	Regression models. Coefficients a mix of consistent in significance & direction and consistent in direction but not all significant.
5	Domestic violence views by gender	2	159, 113	88.3% - 95.6%	2 largest cities in Taiwan	Regression models. None of the independent variables have a significant relationship with the dependent variable in both departments, though there is some consistency in non-significant independent variables.
6	Perceptions of police misconduct	3	102, 110, 308	7.3% - 41.1%	2 large cities w/rural parts & 1 small rural in different areas	Ranking offense seriousness nearly consistent across agencies. Mean likelihood of reporting varies by officer's rank and agency.

7	Uptake of intelligence-led policing innovation	4	67, 75, 59, 86	70.3% - 82.4%	Strong v. weak innovation	Mean comparisons. No differences on some variables. Significant differences on other variables though no consistently lower agency.
8	Views on policing by gender	2	206, 130	>90% - >95%	2 largest cities in Taiwan	Regression models. Coefficients a mix of consistent in significance & direction and consistent in direction but not all significant.
9	Perceptions of police misconduct	2	232, 236	79% - 88%	Large capital v. rural, more junior force	Mean comparisons. Mostly similar views on rule violations & seriousness. Capital force indicated greater willingness to report and punish offenses.
10	Perceptions of procedural justice, race, and demeanor	2	242, 251	20% - 70%	2 Southeastern cities: 1 large, half white; 1 mid-sized, majority white	Regression models. Coefficients a mix of consistent in significance & direction and consistent in direction but not all significant.
11	Support for community policing with minorities	3	417, 135, 161	94.9% - 98.5%	All in DC metro area, varied size & % rural	Regression models. Coefficients a mix of consistent in significance & direction and consistent in direction but not all significant.
12	Views that are part of "police culture" as per conventional wisdom	2	398, 240	93% - 98%	Both near national average for % of minority and female officers	Mean comparisons. No difference on some variables. Significant differences on other variables though no consistently lower agency. Regression models. Coefficients a mix of consistent in significance & direction, consistent in direction but not all significant, and significant in opposite directions.
13	Enrichment of the policing job	2	171, 243	42.4% - 47.7%	Both urban; smaller one also more decentralized	Mean comparisons. Significant differences between departments on most core job characteristics. Differences by rank vary between departments. Largely no differences by education in either department.

14	Views on boundaries of occupational culture of police	2	398, 240	93% - 98%	Both near national average for % of minority and female officers	Mean comparisons. Between-agency differences on some variables but not others, though no consistently lower agency. Direction of correlates for each cluster the same for each agency, though not always same significance.
15	Sources of officers' organization stress	2	210, 303	Not listed	One dept was younger, had fewer years of exercise, and fewer members	EFA. In both departments, a three-factor model is suggested with the same three factors and similar factor loadings.

Figure 1. Article Identification Flow Diagram

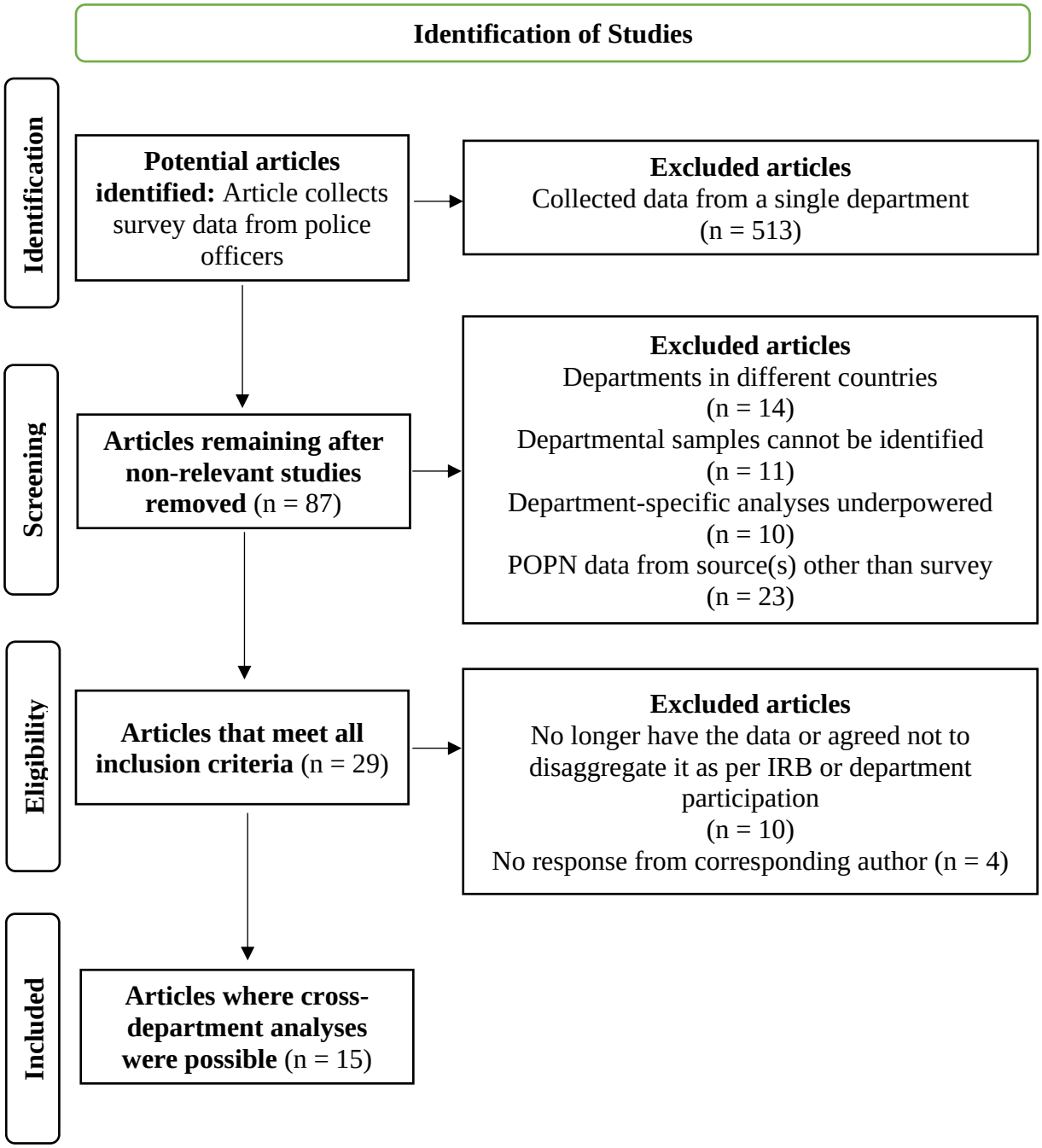
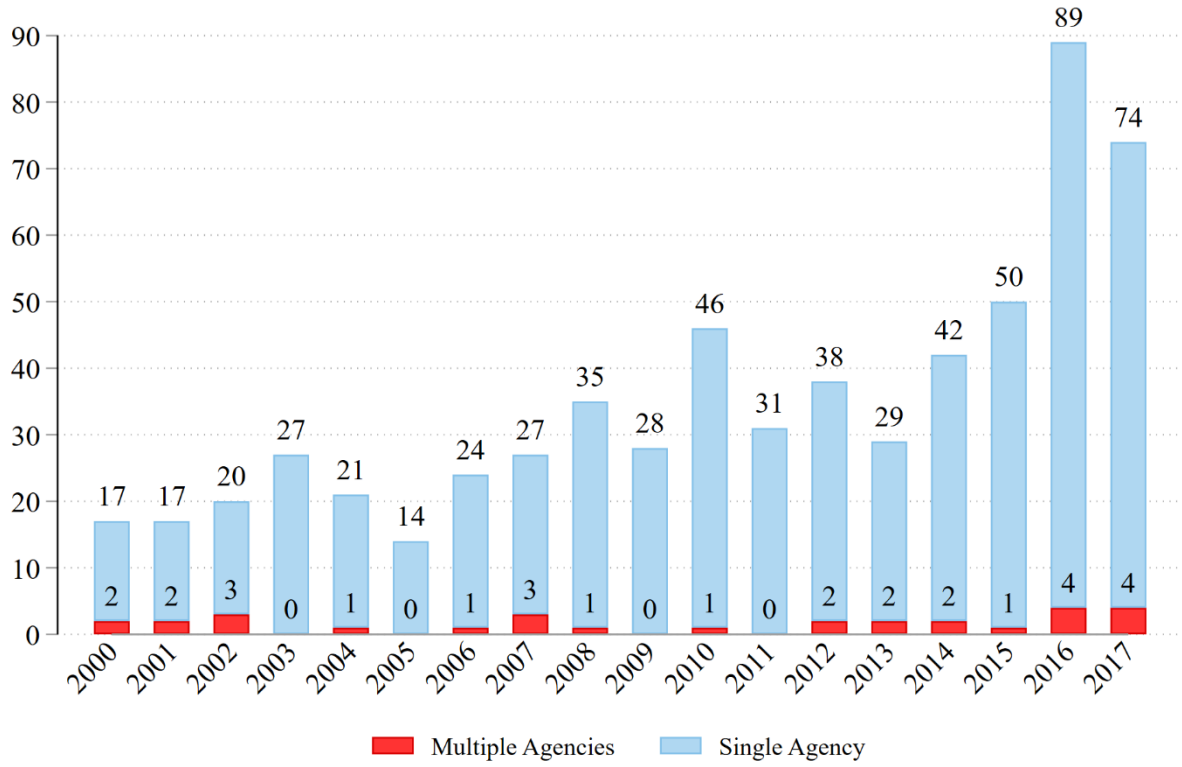
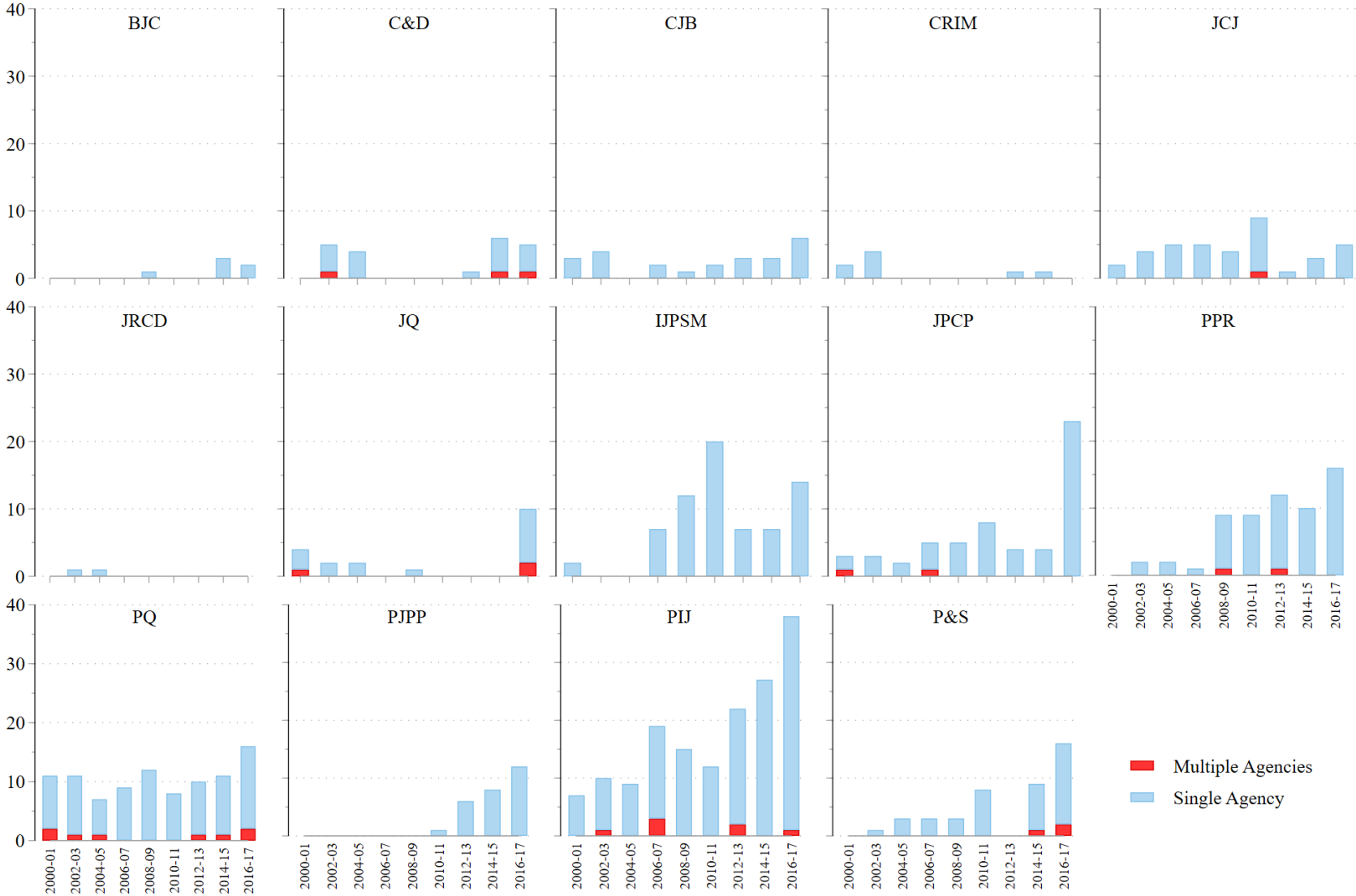


Figure 2. Number of officer survey articles published per year 2000–2017.



Note: Multiple agency articles refer to the 29 articles that meet the inclusion criteria for this study.

Figure 3. Number of officer survey articles published per 2-year increments by journal 2000–2017



Notes: Multiple agency articles refer to the 29 articles that meet the inclusion criteria for this study. Journal of Quantitative Criminology did not publish any articles using officer surveys between 2000 and 2017.

APPENDIX

TABLE A1. Number of Officer Survey Articles Published 2000-2017

Journal	2000-01		2002-03		2004-05		2006-07		2008-09		2010-11		2012-13		2014-15		2016-17		**TOTAL**	
	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S	**M**	**S**
BJC	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	2	**0**	**6**
C&D	0	0	1	4	0	4	0	0	0	0	0	0	0	1	1	5	1	4	**3**	**18**
CJB	0	3	0	4	0	0	0	2	0	1	0	2	0	3	0	3	0	6	**0**	**24**
CRIM	0	2	0	4	0	0	0	0	0	0	0	0	0	1	0	1	0	0	**0**	**8**
JCJ	0	2	0	4	0	5	0	5	0	4	1	8	0	1	0	3	0	5	**1**	**37**
JRCD	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**2**
JQ	1	3	0	2	0	2	0	0	0	1	0	0	0	0	0	0	2	8	**3**	**16**
IJPSM	0	2	0	0	0	0	0	7	0	12	0	20	0	7	0	7	0	14	**0**	**69**
JPCP	1	2	0	3	0	2	1	4	0	5	0	8	0	4	0	4	0	23	**2**	**55**
PPR	0	0	0	2	0	2	0	1	1	8	0	9	1	11	0	10	0	16	**2**	**59**
PQ	2	9	1	10	1	6	0	9	0	12	0	8	1	9	1	10	2	14	**8**	**87**
PJPP	0	0	0	0	0	0	0	0	0	0	0	1	0	6	0	8	0	12	**0**	**27**
PIJ	0	7	1	9	0	9	3	16	0	15	0	12	2	20	0	27	1	37	**7**	**152**
P&S	0	0	0	1	0	3	0	3	0	3	0	8	0	0	1	8	2	14	**3**	**40**
TOTAL	**4**	**30**	**3**	**44**	**1**	**34**	**4**	**47**	**1**	**62**	**1**	**76**	**4**	**63**	**3**	**89**	**8**	**155**	**29**	**600**

NOTE: "M" = multiple agency survey, "S" = single agency survey