Journal of Criminal Justice xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Criminal Justice

journal homepage: www.elsevier.com/locate/jcrimjus

Post-experimental follow-ups—Fade-out versus persistence effects: The Rialto police body-worn camera experiment four years on

Alex Sutherland[a,d], Barak Ariel[b,e,*], William Farrar[c,f], Randy De Anda[c]

[a] Communities, Safety & Justice RAND Europe, Westbrook Centre, Milton Road, Cambridge CB4 1YG, United Kingdom
[b] Institute of Criminology, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, United Kingdom
[c] Rialto Police Department, CA, United States
[d] Jerry Lee Centre of Experimental Criminology, Institute of Criminology, University of Cambridge, United Kingdom
[e] Institute of Criminology, Faculty of Law, Hebrew University, Mount Scopus, Jerusalem 91905, Israel
[f] Administration of Justice Instructor, Mt. San Jacinto College, San Jacinto, CA, United States

ARTICLE INFO

Keywords:
Body-worn cameras
Fade-out effect
Persistent effect
Randomized controlled trials
Police
Natural experiment

ABSTRACT

Purpose: Under certain conditions, experimental treatment effects result in behavioral modifications that persist beyond the study period, at times, even after the interventions are discontinued. On the other hand, there are interventions that generate brief, short-term effects that "fade out" once the manipulation is withdrawn or when the in-study follow-up period is completed. These scenarios are context specific.
Methods: This study reports the results from a three-year post-experimental follow-up from the world's first randomized controlled trial of police body-worn cameras.
Results: The results show that initial falls in rates of complaints against police and police use of force during arrest were sustained during the four years following the cameras being introduced.
Conclusions: The findings suggest that police officers do not become habituated to the effect of the body-worn cameras, and that persistence rather than fade-out effects may characterize this emerging technology.

1. Introduction

The Rialto Police Department was the first police department in the world to participate in a randomized controlled trial of police body-worn cameras. That study, known as the "Rialto experiment," was first published in 2014 (Ariel, Farrar, & Sutherland, 2015), and quickly gained attention following a renewed focus on critical incidents involving officers' shootings in the United States, which sadly continues to this day. Concerns with police accountability, police legitimacy, and use of force in police–public contacts have led to two intertwined phenomena: public upheaval on the one hand (Ransby, 2015), and de-policing (i.e., police withdrawal from proactive engagement with the public; see Oliver, 2015; Pyrooz, Decker, Wolfe, & Shjarback, 2016), on the other. From both sides of this spectrum, body-worn cameras were proposed as a potent solution. Civil liberties organizations such as the American Civil Liberties Union (ACLU) have promoted the use of body-worn cameras to increase the accountability of armed police officers (Stanley, 2013). The police profession pushed for mass rollout as a strategy to reduce some of the tensions with minority groups that recently surfaced, as well as to provide much-needed evidence on police–public encounters (see Lum, Koper, Merola, Scherer, & Reioux, 2015).

The Rialto Experiment (Ariel et al., 2015) provided evidence on the benefits of body-worn cameras in three major ways: first, the study suggested that using body-worn cameras causes a reduction of about 50% in the use of police force compared with control conditions. It also suggested a dramatic reduction in complaints lodged against Rialto police officers, of > 90%, compared with the year prior to the experiment. Finally, the study suggested that the benefits of the equipment justify the costs, with about a 4:1 ratio (see also Ariel, 2016).

There are at least two critical questions about the findings from the Rialto Experiment and they are both linked to the issue of study validity (Shadish, Cook, & Campbell, 2002). First, are the findings from the Rialto Experiment replicable in other settings? Rialto might have been "special" in some way; therefore, the conclusions may have been susceptible to a site selection bias (Allcott, 2015), as Rialto is just one police department from the "universe" of police departments. If this is the case, the findings would not be generalizable. However, this question has been at least partly answered through the Cambridge University Replication Study (CURS) (Ariel et al., 2016b, 2016c, 2016a; see also Drover & Ariel, 2015; Henstock & Ariel, 2017). CURS used an identical methodology to Rialto in a dozen other jurisdictions in English-speaking police departments. CURS discovered virtually identical

* Corresponding author.
E-mail addresses: as2140@cam.ac.uk (A. Sutherland), ba285@cam.ac.uk, barak.ariel@mail.huji.ac.il (B. Ariel), wfarrar@msjc.edu (W. Farrar), ltlojack@aol.com (R. De Anda).

http://dx.doi.org/10.1016/j.jcrimjus.2017.09.008
Received 24 August 2017; Received in revised form 24 September 2017; Accepted 24 September 2017

trends in terms of complaints against the police: an *average* overall reduction of 93% on a year-to-year comparison ($Z = -3.234$; $p \leq 0.001$; between-sites variation $Q = 4.905$; $p = 0.428$). In terms of the use of force, a similar pattern emerged, with significant reductions on a between-groups basis (SMD = −0.346; SE = 0.137; 95% CI −0.614 to −0.077), however only in sites that were characterized by high treatment fidelity (Ariel et al., 2016c; see also Slothower, Sherman, & Neyroud, 2015).

The second critical question deals with what we may call "Fade-out" effects: is the impact of body-worn cameras time dependent, and will the rate of complaints and/or use of force regress back to a pre-implementation mean, as if the body-worn cameras were never introduced? The present report is meant to deal with this practical as well as theoretical question, which underpins a key causal mechanism behind body-worn cameras. In other words, do officers (and suspects) eventually become desensitized to being videoed by a camera during interactions (Ariel, Sutherland, Henstock, Young, & Sosinski, 2017), limiting the effects of body-worn cameras in the longer term? As the findings suggest, responses to these queries both provide deeper insight into self-awareness theory and have direct implications for how to design interventions more optimally.

1.1. Study follow-up periods: taking a longer view

It is widely accepted that randomized controlled trials should have "appropriate" follow-up periods. There are known concerns with the lack of completeness of follow-up *during the in-trial period* of experiments—that is, during or immediately after an intervention has been administered to experimental units. These are mainly issues associated with biased causal estimates of the treatment effect and threats to the statistical power of the test (Juni, Altman, & Egger, 2001; Moher, Schulz, & Altman, 2001). There is no "recommended" in-trial follow-up period, and some flexibility is needed depending on intervention type and discipline. The majority of experiments follow up on the study participants during the grant lifecycle—usually not > 6, 12, or 24 months after the last case was randomly allocated into the study conditions (see Farrington & Welsh, 2005; for more on cascaded random allocation sequences, see Ariel & Farrington, 2010; Wittes, 2002).

In spite of the often substantial costs associated with experiments, it is rare that follow-ups are conducted beyond the life of the original study. This is a concern because "a treatment response restricted to this brief 'in-trial' period can potentially underestimate the long-term benefits of treatment and also may fail to detect delayed hazards" (Llewellyn-Bennett, Bowman, & Bulbulia, 2016: 1). This means that our knowledge about effectiveness is typically limited to the short-term, covering one or two years' post-allocation of units into treatment conditions at most. These issues have been noted particularly in decision-making and education studies (see Allcott & Rogers, 2014; Protzko, 2015, respectively), but there are no apparent reasons they would not also characterize experiments in criminology.

Studies that did measure medium- and long-run effects of interventions have provided critical insight into various interventions.[1] These studies were able to unravel "legacy effects" (Ford, Murray, McCowan, & Packard, 2016), as well as delayed hazards, which are likely to materialize only years after participants were exposed to the treatment (see Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Leventhal, Fauth, & Brooks-Gunn, 2005; Schweinhart et al., 2005; Sherman & Harris, 2013, 2015). For example, the Milwaukee Domestic Violence Experiment (MilDVE) found, with a 23-year follow-up of domestic violence arrests, that death was more prevalent in treatment compared with control groups (Sherman & Harris, 2013, 2015; see also Harris, Polans, Mazeika, & Sherman, 2016). Compared to control cases, victims whose partners were arrested and jailed (than if warned and allowed to remain at home) had a 60% greater risk of all-cause mortality ($p = 0.037$, 95% CI = risk ratio of 1:1.024 to 1:2.628). At 23 years after enrolment, suspects assigned to arrest were almost three times more likely to have died of homicide (at 2.25% of suspects) than suspects assigned to a warning (at 0.81%), a small to moderate effect size ($d = 0.39$; $p = 0.096$; relative risk ratio = 2.79:1; 90% CI = 1.0007 to 7.7696). These findings would not be known with the relatively short follow-up period of the original experiment (Sherman, 1990; Sherman, Smith, Schmidt, & Rogan, 1992).

To our knowledge, Schweinhart et al. (2005) carried out the longest follow-up in a criminological intervention study, on the "Perry Preschool Program". Children were followed up over 40 years after attending a cognitively oriented preschool program aimed to increase thinking and reasoning abilities and school achievement and the children in the program, compared with control children, showed 35% fewer arrests. Treatment children eventually worked harder, were less likely to commit a crime, and participated in many fewer social pathologies than did control group members. Rightly so, their follow-up study was termed "lifetime effects" (see also Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010).

Two additional examples are noteworthy. First, Olds et al. (1998) conducted a 15-year follow-up of the effect of nurse home visitation on children's criminal and antisocial behavior. Their study has shown that the children of visited mothers were arrested at a significantly (54%) lower rate than the children of nonvisited mothers. Second, Henggeler, Clingempeel, Brondino, and Pickrel (2002) had a four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. Analyses demonstrated significant long-term treatment effects for aggressive criminal activity (0.15 versus 0.57 convictions per year) but not for property crimes. We note these additional yet rare studies as they indicate the limited extent of medium and long-term follow-up periods in our field.

1.2. Persistence, durability, and "fade-out" effects

The motivation for long-term follow-ups of interventions is to understand if treatment effects "persist" or "fade out" over time.[2] Allcott and Rogers (2014:3) differentiate between sustained treatment/control differences when treatments are continued for long periods of time ('durability'), and if treatment effects are observed even after interventions have discontinued ('persistence'). To illustrate, it has been shown that early childhood interventions are beneficial during, or immediately after, the intervention has been administered; however, as children move on to poorer quality schools after early childhood intervention, the treatment effect vanishes. Protzko (2015) conducted a meta-analysis of 39 randomized controlled trials aimed at increasing children's IQ scores and investigated whether the effects were durable and persistent. The meta-analysis shows that after an intervention that successfully raised intelligence scores, the effects reduce to nil [effect size immediately after the intervention completed $d = 0.523$ (95% CI = 0.451 to 0.666); over time $b = -0.132$/year (95% CI = −0.243 to −0.021)]. Protzko (2015) suggests that these reductions occur because those in the experimental group lose their IQ advantage over time. It may also be the case that control cases "catch up" with treatment cases in the long run. Although the end result is a nil difference, whether a jump in IQ then reduces in the treatment group, or control group eventually catch up have quite different

[1] Some of these measured effects were reviewed by Allcott and Rogers, 2014: p. 6 [internal references omitted]): "exercise, smoking, weight loss, water conservation, academic performance, voting, charitable donations, labor effort."

[2] Sherman (1990) discusses some of these "after-the-fact" phenomena; within the context of policing, he refers to these as "residual deterrence" and "deterrence decay."