Exploring the Correspondence Between General Correctional Programming and Inmate Misconduct Using a Time-Course Framework

Shannon J. Linning¹, Ian A. Silver²,³, and Jordan Papp⁴

Abstract
Inmate misconduct continues to threaten safety and order within correctional institutions. Yet few studies have examined its longitudinal nature. In this paper we explore the correspondence between correctional programming and inmate misconduct. To do this, we draw from Linning et al.’s time-course framework devised to improve the design and evaluation of interventions by considering effects that can occur before, during, and after programming. We provide the first empirical demonstration of their framework using prisoner misconduct data collected from all Ohio prisons between January 2008 and June 2012. A cross-lagged panel analysis provides support for the use of a time-course framework. Results show that misconduct decreased during programming. However, we observed increases in misconduct prior to and following exposure to programming. Our results suggest that future work needs to improve our understanding of causal mechanisms of inmate misconduct and when their effects are expected.

Keywords
prisoner misconduct, correctional programming, program evaluation, time-course framework, initial backfire, anticipatory effects, residual effects

¹Simon Fraser University, Burnaby, BC, Canada
²Rowan University, Glassboro, NJ, USA
³University of Cincinnati, Cincinnati, OH, USA
⁴University of Michigan, Ann Arbor, MI, USA

Corresponding Author:
Shannon J. Linning, School of Criminology, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada.
Email: shannon_linning@sfu.ca
Introduction

In recent decades, we have seen an increase in rehabilitative programming administered by correctional institutions (MacKenzie & Lattimore, 2018). The goal has been to provide resources and training to offenders that reduces their likelihood of recidivism post-release and makes them productive members of society (Lipsey & Cullen, 2007). Yet correctional institutions are simultaneously tasked with providing security to staff and inmates, maintaining order, and assuring sentences are carried out (Bottoms, 1999). Accomplishing these is important to keep staff and inmates safe, keep prison costs in check, and to provide an environment conducive to successful offender rehabilitation.

Maintaining order continues to be a challenge. At an institutional level, prisons deal with thousands of disciplinary infractions annually that collectively impose an immense financial burden (Tewksbury et al., 2014). Inmate misconduct, such as fighting, selling drugs, and disobeying staff orders, undermines rehabilitative efforts (Steiner & Wooldredge, 2014). Some studies suggest that inmates who engage in prison misconduct are more likely to recidivate post-release thus undermining the rehabilitative goals of prisons (Cochran & Mears, 2017; Silver & Nedelec, 2018a).

Although many studies have examined the correlates of inmate misconduct, most are cross-sectional and few evaluate the influence of programming (French & Gendreau, 2006; Steiner et al., 2014). There is an emerging literature showing the importance of assessing misconduct in a longitudinal manner (e.g., Cihan et al., 2017; Cochran, 2012; Silver & Nedelec, 2018b), but it is still in infant stages. To date, no studies have examined the longitudinal correspondence between prison programming and misconduct. It is possible that correctional programming can influence the time in which inmates engage in misconduct. For instance, inmates often apply for programming. In anticipation of the program, they may go on their best behavior with hopes of gaining entry. As such, prisons may experience fewer incidents of misconduct in the weeks before a program starts. However, as we will discuss, programming can also influence offender behavior at other time periods, namely during and after. Thus, it is worthwhile to assess misconduct and programming at different time periods. This is the focus of our paper. To do this, we draw upon Linning et al.’s (2019) time-course framework.

Linning et al. (2019) argued that although interventions can reap desired results overall, more subtle effects can occur at different times, namely before, during, and after implementation. Consequently, they proposed a time-course framework to guide the design and evaluation of interventions in criminal justice, including correctional programming. Finding additional tools that can help practitioners develop practices that make order maintenance and correctional programming compatible are invaluable. Our paper provides the first known demonstration of this time-course framework. Specifically, we explore various temporal effects concerning prisoner misconduct associated with correctional programming using longitudinal data collected from prisons in Ohio. We examine whether programming influences inmate misconduct in custody before, during, and after programming.
Inmate Misconduct and Correctional Programming

A key objective for correctional administrators is to maintain order within the facilities they are operating. This enables them to keep both staff and inmates safe. It also creates environments conducive to the correctional rehabilitative process. However, inmate misconduct—namely “the violation of within-prison policies” (Silver & Nedelec, 2018b, p. 17)—continues to pose a serious threat to correctional institutions. Such violations can range from refusing to comply with an order from staff to possession of drugs to assaulting a guard or another inmate (Berk et al., 2006; Langan & Pelissier, 2001). Addressing misconduct can also be costly. Some studies suggest that it costs nearly $1,000 per incident to address (Lovell & Jemelka, 1996), and even more expenses are incurred if institutions must factor in liabilities, litigation, and the need to implement higher security measures for problematic prisoners (Tewksbury et al., 2014).

Despite the recent shift toward using rehabilitative approaches and correctional programming (Lipsy & Cullen, 2007), “evaluations of the effect of treatment programs on prison misconduct have been rare” (French & Gendreau, 2006, p. 189). Correctional administrators have used various tools in attempts to reduce misconduct, including classification tools, structured routines, and intervention programs (Steiner et al., 2014). However, the latter is less common. Most correctional programs are focused on alternative goals (French & Gendreau, 2006). Some seek to reduce recidivism using cognitive behavioral techniques to correct thinking errors and develop anger management skills (Andrews & Bonta, 2010; Milkman & Wanberg, 2007). Other programs, such as vocational or educational ones, seek to produce individuals with marketable skills that can make them productive members of society upon release (Pompoco et al., 2017). In addition, some programs have been designed to improve inmates’ lives while in custody and provide them with constructive activities to engage in throughout the day (Strimple, 2003).

This raises an important policy issue. Although many correctional programs seek to reduce problematic behaviors (e.g., rehabilitative programs) and increase employable skills (e.g., vocational and educational programs), they may simultaneously influence instances of prison misconduct which undermines the safety and order goals of these institutions. For instance, Wooldredge (1998) explains that prisons might observe increases in misconducts—such as failing to follow an order—because inmates are spending more time with correctional staff. Spending more time participating in programming increases the amount of time inmates interact with staff which could increase the likelihood that they get caught for infractions. Steiner et al.’s (2014) systematic review of the prison misconduct research revealed mixed results for the influence of institutional routines and experiences. Specifically, although inmates involved in prison work assignments and religion-based programs committed less misconduct, those involved in other forms of prison programming engaged in misconduct more often. A systematic review of research has found several factors that correlate with misconduct, such as: age, prior record, history of misconduct, and security level of the prison (Steiner et al., 2014). Although these
studies have provided great insights into the causes and correlates of misconduct, further inquiry is needed. For instance, the research on misconduct overwhelmingly uses cross-sectional data (Steiner et al., 2014).

There is an emerging literature that has assessed misconduct in longitudinal terms, but it is still in infant stages. These studies range in scope from assessing the association between misconduct and pre-prison inmate characteristics (Cihan et al., 2017), visitation (Cochran, 2012), and cognitive abilities (Silver & Nedelec, 2018b). While the results from these studies suggest that individual- and institutional-level factors can influence inmate misconduct over time, only a limited number of associations have been empirically examined. The results imply that offender behavior can fluctuate over time. But more research on the longitudinal characteristics of inmate misconduct is needed.

Despite these advancements, no study has examined the longitudinal correspondence between prison programming and inmate misconduct. To do this, we draw from Linning et al.’s (2019) framework for understanding the longitudinal time-course of interventions and programming. They argue that it is theoretically possible for offenders to change their behavior before, during, and after interventions are implemented. This could explain why the existing cross-sectional literature contains mixed results regarding the association between programming and misconduct. The following section describes their time-course framework and its applicability to correctional programming.

The Longitudinal Time-Course of Interventions

The crime prevention literature has found evidence of various temporal effects caused by interventions that are either favorable or unfavorable. Prior to intervention start dates, we have observed reductions in crime, namely anticipatory benefits (Smith et al., 2002). However, increases in crime prior to intervention start dates are also possible (i.e., anticipatory backfire). During interventions we observe either reductions (i.e., initial deterrence; see Sherman, 1990) or increases (i.e., initial backfire; see Linning & Eck, 2018) in offending. Lastly, researchers have observed reductions in offending from interventions even after they have ceased, namely residual deterrence (Sherman, 1990). Increases in offending following interventions (i.e., residual backfire) are also possible (Linning et al., 2019).

Linning et al. (2019) point out that despite knowing about temporal effects that materialize before, during, and after an intervention is implemented, researchers typically study them in isolation. If, for example, a study tests for the presence of residual deterrence, it usually does not simultaneously assess whether anticipatory benefits occurred. Instead, they typically focus on single before and after observations surrounding an intervention. Linning et al. (2019) argue that failing to account for these temporal effects leads to a mis-estimation of the true impact of an intervention. For instance, if an intervention generates anticipatory benefits because of pre-intervention publicity (see Smith et al., 2002), but evaluators do not factor in its impact, an intervention may appear less effective if their pre-test observations are collected while
anticipatory benefits are in effect. Crime is already decreasing because of the intervention, just before it formally begins. As a result, the difference between the pre- and post-test observations appear smaller.

As such, Linning et al. (2019) argue that practitioners and evaluators should consider whether effects are possible before (i.e., anticipatory benefits or backfire), during (i.e., initial deterrence or backfire), or after (i.e., residual deterrence or backfire) an intervention. They call this the time-course of an intervention. They merged their arguments for considering the temporal component of interventions with the principles presented in Johnson et al.’s (2015) EMMIE scheme to create their time-course framework. EMMIE is an acronym used to highlight the importance of the various dimensions of an intervention that practitioners should consider. They include: 1. the overall effect of an intervention, 2. the mechanisms triggered by the policy, 3. the moderators/context that triggers mechanisms in a favorable way, 4. how to implement the policy, and 5. the economic costs and inputs needed to put a program into action (see Johnson et al., 2015, p. 463). The crux of Johnson et al.’s (2015) argument is that practitioners and evaluators should consider dimensions of interventions above and beyond effect sizes. While these are informative, they only provide part of the picture. We also need to understand elements such as the mechanisms being triggered and contextual factors needed to trigger them in beneficial (as opposed to harmful) ways. This might include more in-depth qualitative descriptions of how the intervention was carried out as well as how it is believed to have generated the resulting effect.

Linning et al. (2019) expanded upon EMMIE by adding a temporal component. They argued that different mechanisms can be triggered at different times throughout an intervention’s time-course (i.e., before, during, or after) leading to fluctuations in offender behavior over time. And depending on how a program is implemented either benefit or backfire could occur during each time period. As the authors of both papers argue, these frameworks can be modified for use in any area of criminal justice research. Here, we see the utility of these frameworks for correctional research. Correctional programming is also designed to modify offender behavior and could generate varying effects at different points during a program’s time-course.

To illustrate, consider a hypothetical example: a correctional program using cognitive behavioral therapy (CBT) to reduce recidivism. Let us assume an institution decides to design its own anti-bullying program using CBT techniques. While staff develop the program, they also post flyers to advertise the upcoming sessions that will be offered and encourage inmates to sign up. Let us also assume that the program takes place twice per week for 3 months and involves inmates assembling in a classroom to complete worksheets and engage in group discussion about effectively responding to bullying behavior.

Figure 1 provides an intervention time-course for the possible effects that might occur from this CBT program. To start, anticipatory benefits prior to the intervention are possible. This could occur in two ways. First, an officer motivation mechanism could be triggered (see Smith et al., 2002, p. 79). That is, if staff are designing the program themselves, they may be heavily invested in and/or convinced of its success. They might enthusiastically promote the program while interacting with inmates touting its
benefits. They may convince inmates that it is in their best interest to participate. Consequently, inmates might improve their behavior with hopes of gaining enrollment into the program. Second, the publicity effects mechanism (Smith et al., 2002, p. 79) could be triggered. Even in cases where inmates were not approached or encouraged by staff to enroll in the program, the posted flyers could generate a similar reduction in misconduct. Similarly, inmates discussing the program with one another could increase awareness. Interest in gaining entry into the program could compel them to behave accordingly so as not to compromise their eligibility for enrollment. This leads to a decline in misconduct prior to the program start date.

Once the program starts, this downward anticipatory trend in misconducts may not persist. It is possible that bringing offenders together could create new conflicts that would not have occurred had they not come into contact with each other to participate in the program (Wooldredge, 1998). It is also possible that bringing offenders together creates an opportunity for them to learn from one another. Particularly, lower risk offenders learning adverse behaviors from higher risk ones (Lowenkamp & Latessa, 2004). We call this an increased contact mechanism that generates backfire and increases misconduct. Similarly, participation in the program increases the amount of interaction that offenders have with correctional staff and the possibility of disobeying an order. One might think of this as an increased staff exposure mechanism. Thus, misconduct increases during programming.

That said, if the program was successful in helping inmates identify and modify thinking errors, they may become more equipped to manage antagonistic behavior from others. Instead of retaliating against an instigator, the inmate may be able to mitigate the situation without an aggressive response. As such, the corrected thinking

Figure 1. Time-course of inmate misconduct surrounding hypothetical cognitive behavioral therapy-based programming.
errors mechanism will be triggered. But given the progressive, cumulative nature of CBT-based programming, it may take time for inmates to develop these skills (Bonta & Andrews, 2016). Thus, beneficial effects may not be observed until after the program is completed, hence the residual deterrence effect shown afterward.

As this hypothetical example illustrates, it is theoretically possible for us to expect misconduct to occur at different times surrounding programming. From a theoretical standpoint, we believe this time-course framework shows promise for research in correctional programming. Not only could it help practitioners design more effective programs, but it could also help researchers better evaluate their effectiveness. It is a practical tool to help operationalize correctional programming and anticipate various fluctuations in misconduct. Such information can enhance practitioners’ ability to reduce risk within prisons and anticipate their associated costs. It also provides further support for the need to assess inmate behavior in a longitudinal manner (Silver & Nedelec, 2018b). However, the time-course framework has yet to be empirically applied. Here, we provide the first known empirical demonstration of its potential use in corrections. Specifically, we were able to explore the general utility of the framework in evaluating misconduct across all programs used by the Ohio Department of Rehabilitation and Correction (ODRC) over 3 years. If temporal variations in misconduct exist across a wide range of programs, we believe it would provide the opening of an important discussion on how to improve the design and evaluation of specific correctional programming and understand potential consequences.

Methods

Sample

To examine the influence of programming on inmate misconduct, we used data collected during the Evaluation of Ohio’s Prison Programs grant. All inmates incarcerated between January 2008 and June 2012 participated in the study. The Evaluation of Ohio’s Prison Programs grant assessed inmate programming, inmate misconduct, and inmate recidivism within the Ohio prison system. The data—which were retrospectively collected from official records—captured information regarding the individual- and system-level characteristics associated with programming, misconduct, and recidivism for 88,621 individuals (105,945 cases). For a more thorough review of the Evaluation of Ohio’s Prison Programs study, please see Latessa et al. (2015), Pompoco et al. (2017), and Lugo et al. (2019). In the overall sample (N = 88,621), the majority of individuals incarcerated between January 2008 and June 2012 were males (86%), of Caucasian descent (66%), not of Hispanic descent (98%), and not married (91%). Moreover, the average age at intake was 32 years of age (min, max = 15, 90), the average number of prior incarcerations was approximately one (min, max = 0, 16), and the average security classification was low moderate security (min, max = 1, 5). From the original sample, we created the analytical sample by limiting it to those who served 3 or more years in prison and removing cases that had incomplete data. The final analytical sample consisted of 25,014 individuals. The majority of individuals in the analytical sample were male (91%), of Caucasian descent (54%), not of Hispanic descent...
International Journal of Offender Therapy and Comparative Criminology 66(2-3)

(97%), and not married (92%). Additionally, the average individual in the analytical sample was 25 to 34 years of age, had zero or 1 to 2 prior prison sentences, and was classified in the low moderate security group.

Longitudinally Standardizing the Data

To examine the temporal effects of prison programming on prison misconduct, we implemented a longitudinal standardization process (Silver & Nedelec, 2018b). Most programs offered by the ODRC are shorter than 3 months. Thus, we created 3-month blocks labeled as ‘months from admission blocks’ (MFABs) conditioned upon the inmates’ admission date and release date. This allowed us to be relatively certain that enough time was captured within a single MFAB for an inmate to complete all or the majority of a program.3 The MFABs are not a variable (i.e., construct) included in the models, but rather a conceptual method for standardizing the time varying constructs. In essence, using the language of a traditional longitudinal dataset, the MFABs represent waves or rounds of data. Each inmate’s sentence was broken down into these 3-month blocks between their admission date and release date.

To provide an example of the MFAB standardization process, an inmate serving a prison sentence from June 1st 2008 to November 1st 2008 would have two MFABs of data associated with him/her (i.e., MFAB 1 and MFAB 2). Alternatively, an inmate serving a prison sentence from January 1st 2009 to January 1st 2011 would have eight MFABs of data associated with him/her (i.e., MFAB 1 through MFAB 8). Consequently, the creation of these MFABs permitted us to create standardized waves of data between inmates that served sentences at different times.

We standardized the data for the first 3 years of the inmates’ sentences only for those who served at least 3 years in prison (25,014 inmates served at least 3 years in prison). Specifically, 12 MFABs of data were created for each inmate included in the analytical sample. We standardized the data for the first 3 years of the inmates’ sentences because the number of inmates in the analytical sample (N=25,014) would be drastically reduced through listwise deletion needed for the analytical strategy.4 As such, an inmate serving from January 1st 2008 to January 1st 2011 would be included in the dataset (served at least 3 years in prison) and would have 12 MFABs of data associated with him/her, and an inmate serving from June 21st 2009 to June 21st 2012 would be included in the dataset (served at least 3 years in prison) and would have 12 MFABs of data associated with him/her. Moreover, even though these inmates’ admission dates are distinct (January 1st 2008 and June 21st 2009, respectively), the first MFAB would be consistent across inmates, representing the first 3 month period from their admission (i.e., January 1st 2008 to March 1st 2008 and June 21st 2009 to August 21st 2009, respectively).

Time-Varying Measures

Prison Programming. We used the difference between the admission date and the program start date to identify the specific MFAB in which individuals began their
programming. For instance, if an inmate started programming during the fourth month from admission s/he received a score of 1 during MFAB 2 and a 0 during MFAB 1 and MFABs 3 through 12. The standardization process allowed us to create 12 programming measures (an MFAB-specific measure) that captured the specific 3-month period—since the beginning of their sentence—that an inmate began programming.\(^5\) We aggregated the programming information at each MFAB to capture the frequency of programming at each MFAB. As such, an individual that received six programs during the fourth month from admission received a score of 6 during MFAB 2 of data and a 0 during MFAB 1 and MFABs 3 to 12 of data.

It should be noted that all programs authorized by the ODRC were included within the standardized longitudinal measure of sentence specific programming. This includes, but is not limited to, Thinking for a Change (T4C), From Bondage to Freedom (a religious program), Alcoholics Anonymous (AA) and Narcotics Anonymous (NA), among others. Most of the programs included within the measure do not adhere to the principles of cognitive behavioral therapy (Latessa et al., 2015). Overall, 528 different programs were offered at least once during the study period. Of those, 35 (or 7\%) were reentry approved and 24 (or 4.5\%) were explicitly designed with the goal of reducing recidivism.

**Guilty misconduct convictions.** We used guilty misconduct convictions as our outcome variable. Following the prior scholarship, we conceptualized guilty misconduct as an infraction that was subjected to a Rules Infraction Board (RIB) hearing, which resulted in a guilty finding and formal punishment (Lugo et al., 2019; Pompoco et al., 2017; Steiner & Wooldredge, 2014). We identified guilty misconduct across all MFABs in two ways. First, we classified the guilty misconduct data within each of the 12 MFABs by calculating the difference between the admission date and the guilty misconduct date. Second, we aggregated the number of guilty misconducts at each MFAB of data collection to capture the frequency of misconduct during each 3-month period.

**Time-Invariant Measures**

We included seven time-invariant measures as control variables in the cross-lagged panel analysis described below. The time-invariant control variables were selected following the guidance of the prior literature (e.g., Lugo et al., 2019; Pompoco et al., 2017; Steiner & Wooldredge, 2014), as well as for the theoretical and empirical influence of the measures of interest (e.g., Silver & Nedelec, 2018b). These control variables were *male* (0 = Female; 1 = Male), *non-White* (0 = White; 1 = non-White), *Hispanic* (0 = non-Hispanic; 1 = Hispanic), *prior prison sentences categories* (0 = 0; 1 = 1–2; 2 = 3 or more), *security level* (1 = minimum security; 2 = low/moderate security; 3 = moderate security; 4 = maximum security; 5 = super maximum security), *age categories* (1 = 15–24; 2 = 25–34; 3 = 35–44; 4 = 45–55; 5 = 55–80), and *married* at incarceration (0 = No; 1 = Yes).
Analytical Strategy

We used a two-step analytical strategy to estimate the temporal variations in misconduct associated with prison programming. First, we calculated descriptive statistics for all of the variables of interest (see Table 1).
Second, we estimated a cross-lagged panel analysis. Similar to a fixed effects model, this analytical technique allowed us to estimate the simultaneous impact of programming and misconduct on the likelihood of receiving subsequent programming or engaging in misconduct at later time periods while controlling for any indirect effects from earlier MFABs. Briefly, a cross-lagged panel analysis is a longitudinal modeling strategy used to examine highly interdependent relationships (Kearney, 2017; Selig & Little, 2012; Shingles, 1976). Inmate programming and misconduct are highly interrelated (French & Gendreau, 2006). Thus, we need to control for these effects in preceding time periods. Cross-lagged panel analysis has been identified as one of the foremost modeling strategies to parse out the effects of interest while controlling the effects of unobserved mechanisms (Shingles, 1976). Specifically, the analysis evaluates between-individual change in misconduct and programming and the effect of that change on subsequent misconduct and programming (in a cross-lagged manner), while controlling for any correlated residual error across equations.

The cross-lagged panel analysis generates four separate regression and covariance specifications to assess effects at each MFAB. To control for correlated error between lagged endogenous constructs, we specified residual covariances for programming and misconducts between each MFAB. This allows the model to control for any indirect effects of programming and misconduct at earlier MFABs. It also isolates the direct effects of programming and misconduct at one MFAB on programming and misconduct at the subsequent MFABs (e.g., variables at MFAB 3 predict MFAB 4). All of the paths in the cross-lagged panel analysis controlled for the seven time-invariant constructs described above. We estimated the cross-lagged panel analysis with weighted least squares estimator with robust standard errors from the Lavaan package in R (Kline, 2016; Rosseel, 2012). Appendix A provides the full results (Supplemental Table A1) associated with the cross-lagged panel analysis.

Results

Figure 2 provides the results of the cross-lagged panel analysis showing the association between programming and misconduct. The model generates separate path models to assess whether benefits or backfire occurred before, during, or after programming throughout the study period. The global fit statistics for the model indicated that the specification of a cross lagged panel analysis appeared to fit the data well and within the standards described by Hu and Bentler (1999; $X^2 = 322.897^{*}$; CFI = 0.99, RMSEA [90% CI] = 0.004 [0.003, 0.005]; SRMR = 0.008; $N = 25,014$; $p < .05$).

First, we assessed the paths for misconduct before programming. In ten of the eleven MFABs, the analysis reveals a positive association between misconduct and subsequent programming. This indicates that either people who engage in more misconduct are subsequently participating in programming, and the inverse, those who are not engaging in misconduct are less likely to participate in programming. This suggests a possible anticipatory backfire effect. Second, when measured at the same wave (i.e., during), programming and misconduct have a statistically significant negative covariance for nine of the twelve MFABs. The negative covariance indicates that inmates who participate in programming experience reductions in misconduct during
the same time period. This suggests an initial deterrence effect. Third, the results illustrated that programming had a positive association with subsequent misconduct on all of the cross paths (i.e., the regression model of programming on subsequent misconduct). This provides evidence of residual backfire. For example, a positive association between programming at MFAB 5 and misconduct at MFAB 6 ($b = 0.02$; $p < .05$) was observed in the cross-lagged panel analysis. The 11 positive associations suggest that individuals exposed to inmate programming experienced spikes in misconduct in subsequent MFABs.

In addition to these findings it is important to highlight that programming positively covaried with subsequent programming, and misconduct positively covaried with subsequent misconduct between each MFAB. This finding suggests that continuity in programming and misconduct exists while adjusting for the interrelated effects of programming on misconduct and misconduct on programming. Furthermore, the interrelated effects of programming on misconduct and misconduct on programming exist while adjusting for the continuity in programming and misconduct.

**Discussion**

In this paper we examined the longitudinal correspondence between correctional programming and inmate misconduct. Informed by Linning et al.’s (2019) time-course framework, we explored the possibility of varying temporal misconduct effects that may be associated with programming. Instead of using a before-after approach to evaluate misconduct, we assessed the possible effects that could materialize before (i.e., anticipatory benefit or backfire), during (i.e., initial deterrence or backfire), and after (i.e., residual deterrence or backfire) programming.
Consistent with the literature, our data show that misconduct declined over the full course of inmates’ sentences (Cochran & Mears, 2017; Steiner & Wooldredge, 2008). However, we also observed some notable effects that deviated from this overall trend, particularly when we assessed misconduct immediately before and after programming. We found declines in misconduct during programming, but evidence of increased misconduct before and after programming. Results from our cross-lagged panel analysis suggest a possible anticipatory backfire effect (i.e., backfire before programming). In ten of the eleven cross-paths, we found significant positive associations between misconduct and programming. Conversely, we found evidence of initial deterrence (i.e., benefit during programming). Our analyses revealed a negative covariance between misconduct and programming in nine of the twelve time periods assessed. This suggests that inmates participating in programming experience simultaneously lower incidences of misconduct. However, this beneficial effect appears to only be temporary. When we assessed misconduct in the time periods following programming, our models suggested that a positive relationship exists. That is, participating in programming at one time period was positively associated with misconduct in the following one. This suggests that a residual backfire effect occurred (i.e., backfire after programming).

We cannot speak to the precise mechanisms that may have given rise to our results because we analyzed an aggregate sample of programs. However, we can speculate as to why certain effects materialized at certain times. First, anticipatory backfire may have arisen out of policies in Ohio that make certain treatment programs mandatory under certain circumstances. For example, educational programming is mandatory when an inmate enters a facility, is of a certain age, and does not have a high school diploma or equivalent. The goal of these policies is to give inmates the basic education that is needed for most employment opportunities (Pompoco et al., 2017). While these programs are well-intentioned, they may lead to anticipatory backfire because they are required of inmates. If an inmate does not want to participate in such a program s/he may engage in misconduct hoping to avoid entering programming.

Second, initial deterrence could have been the result of a couple different mechanisms. One reason could be that the time inmates spent in programming was structured where they engaged in prosocial activities, which took away idle time that can facilitate misconduct (Vuk & Doležal, 2019). Another reason might be that the skills that inmates learn in treatment are effective at reducing engagement in misconduct in the short run because the skills had been recently learned and practiced.

Lastly, residual backfire could have occurred because inmates feel the pains of imprisonment once they return to mundane prison life (see Cochran, 2012). The lack of programming may also lead to a lack of motivation to refrain from engaging in misconduct because there is no reward (i.e., staying in programming) for good behavior. Alternatively, the skills learned in programming that help inmates avoid programming may decay over time. Importantly though, while these explanations may be plausible, they unfortunately cannot be tested and are only offered as speculation that may help to facilitate future research on the topic.
Our findings should also be interpreted with three caveats in mind. First, although the goal of the current study was to examine the correspondence between general programming and inmate misconduct, the inclusion of programs designed for non-behavioral change purposes (e.g., education, employability, and religiousness) likely influenced the observed pattern of findings. Specifically, most of the programs included in the analysis do not align with cognitive behavioral therapy (CBT), the best practice known to achieve prosocial behavior change (Latessa et al., 2015). As such, while these non-CBT programs do provide inmates with benefits (e.g., high school diplomas, resume builders, and access to prosocial actors), it is possible that they have limited effects on long term behavior change.

Second, the results of the current study are based on a subsample of inmates that served at least 3-years in prison. This was done because of the need to have a sample that could be examined longitudinally across a sufficient time period needed for our analytic strategy. While the analytical sample proved useful, most inmates in prison serve sentences shorter than 3 years. As such, while the analytical sample emulated that of the total sample, the individuals captured in the analytical sample were likely incarcerated for more serious offenses than the general prison population. Future research should examine these relationships using a sample of individuals that were incarcerated for less than 3 years to see whether findings differ.

Third, due to data limitations we are unable to determine what type of misconduct was committed by the individuals in the dataset. We can only determine that the misconducts committed were serious enough to result in a guilty sanction by the rules infraction board. Future studies should examine whether these temporal fluctuations differ depending on the severity of misconduct committed, facility security level, criminogenic risk of the inmate, and the timing of the treatment program during the inmate’s sentence (Papp et al., 2019; Steiner et al., 2014).

Our findings also provide several important policy implications. First, we should be thinking of correctional programming in longitudinal terms. While most correctional evaluations carry out before-after designs (Dowden & Andrews, 2000; Hoberman, 2016), our results show that we should factor in the possibility that offender behavior may change at different time periods in response to programming. Changes in inmate behavior can occur at different time periods surrounding programming. Thus, Linning et al.’s (2019) time-course framework could be of great use for program design and evaluation in corrections moving forward.

Second, we need more extensive research to understand the mechanisms triggered by correctional programming. Our results indicate that inmate misconduct can fluctuate over time. But as Johnson et al. (2015) suggest, the next step is to understand the mechanisms causing these fluctuations. To do this, we need to conduct individual assessments of specific programs using the time-course framework. Different programs have varying goals, and the resulting experiences that inmates receive can lead them to respond differently. Though we were not able to test the mechanisms of specific programs in this study, future studies should examine individual programs using this framework. Individual programs may produce varying temporal effects that are distinct from the aggregate results found here. Researchers should keep in mind that it
is also possible that incentives for a specific program could influence the level of misconduct occurring in institutions. For instance, many prisons allow inmates to earn time-served credits for participating in reentry approved/CBT-based programs (Demleitner, 2017). In such facilities, anticipatory benefits may materialize when CBT-based programs are offered compared to others. The time-served credit would incentivize inmates to behave themselves even more than if a vocational or educational program were offered. Future studies should consider these differences, particularly whenever comparisons between programs are made.

The third implication of our study rests in how we determine program effectiveness. Our results indicate that there was reduced misconduct during programming, but increased misconduct before and after programming. If an individual program generated similar outcomes, would we consider this program effective? Practitioners would need to consider whether to use such a program. As we use the time-course framework to increase our understanding of how programs work, we can predict when increases in misconducts will likely begin and end. This could allow us to consider the possible options available to mitigate any temporary backfire and when to implement them (e.g., temporarily increasing staff for more supervision before and after programming).

The suggestion that we should conduct longitudinal studies of correctional programs is not new (Cihan et al., 2017; Cochran, 2012; Silver & Nedelec, 2018b). But this time-course framework has shown potential to help us think in these terms. While collecting multiple pre- and post-test observations is optimal, sometimes resource constraints prevent us from doing so. We believe this time-course framework still has use in those instances. For starters, it would assist with program design. If practitioners consider how inmates might respond to interventions at different times, it could help us predict when benefits and backfires may occur. If a program is effective overall but we predict that a temporary backfire could occur, measures can be taken (e.g., temporary changes in staffing for increased supervision) to mitigate these effects. This should help reduce harm to staff and inmates. Then from an evaluation standpoint, it can assist us in deciding when to collect our observations. Even if resources only permit the collection of one pre-test and one post-test observation, researchers need to strategically decide when to carry this out. For instance, if anticipatory benefits are expected because of a *publicity effects* mechanism, then evaluators should consider collecting their pre-test observation before a program is advertised to inmates. One could argue that any anticipatory benefits are attributed to programming, just in a more indirect way. Thus, a pre-test observation collected while anticipatory benefits are occurring will be lower making the program appear less effective.

Overall, our results provide support for Linning et al.'s (2019) time-course framework. In terms of correctional programming and prisoner misconduct, various temporal effects can materialize at different times. This evidence suggests that we should keep thinking about program design and evaluation in terms of their time-course. Future work should carry out program-specific evaluations of misconduct trends using similar statistical techniques. Doing so will improve our understanding of the causal mechanisms of misconduct to help maintain safety and order within correctional institutions.
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Ohio Department of Rehabilitation and Correction. All conclusions or opinions expressed are those of the authors and not of the Ohio Department of Rehabilitation and Correction.

ORCID iD
Shannon J. Linning https://orcid.org/0000-0002-3027-1667

Supplemental Material
Supplemental material for this article is available online.

Notes
1. The divergence between the number of individuals and the number of cases is a function of multiple sentences. Approximately 20,000 inmates were incarcerated in the Ohio Department of Rehabilitation and Correction (ODRC) more than once during the study period. Only initial convictions were retained for the current study. We focused on initial convictions because we wanted to limit the cumulative effects of programming and misconduct across incarceration for the limited number of individuals that did return to prison.
2. The analytical data in the current study is balanced, which in longitudinal analyses indicates that all participants had complete data (i.e., non-missing) across the 12 MFABs (Kearney, 2017).
3. Due to data limitations, we were unable to determine the average length of every program administered by ODRC. We, however, do know that majority of programs offered by ODRC are shorter than 3 months. Nevertheless, this data limitation hinders our ability to estimate the programming overlap between MFABs.
4. Only a limited number of inmates in the total sample served more than 3 years in prison.
5. To state otherwise, the process allowed us to isolate the effects of programming during an individual’s sentence rather than observe the effects of a single program at a single time point. It should be noted that a substantial sum of inmates did not likely receive the same program at a single time point (e.g., January 1st, 2008), but received programming at the same point during their own sentence.

References
Andrews, D. A., & Bonta, J. (2010). The psychology of criminal conduct (5th ed.). Matthew Bender & Company.
Berk, R. A., Kriegler, B., & Baek, J. H. (2006). Forecasting dangerous inmate misconduct: An application of ensemble statistical procedures. Journal of Quantitative Criminology, 22(2), 131–145.
Bonta, J., & Andrews, D. A. (2016). The psychology of criminal conduct. Routledge.
Bottoms, A. E. (1999). Interpersonal violence and social order in prisons. *Crime and Justice*, 26, 205–281.

Cihan, A., Sorensen, J., & Chism, K. A. (2017). Analyzing the offending activity of inmates: Trajectories of offense seriousness, escalation, and de-escalation. *Journal of Criminal Justice, 50*, 12–18.

Cochran, J. C. (2012). The ties that bind or the ties that break: Examining the relationship between visitation and prisoner misconduct. *Journal of Criminal Justice, 40*(5), 433–440.

Cochran, J. C., & Mears, D. P. (2017). The path of least desistance: Inmate compliance and recidivism. *Justice Quarterly, 34*, 431–458.

Demleitner, N. (2017). Time off for good behavior. In K. Kerley (Ed.), *The encyclopedia of corrections* (pp. 1–4). John Wiley & Sons, Inc.

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology, 42*, 449–467.

French, S. A., & Gendreau, P. (2006). Reducing prison misconducts: What works! *Criminal Justice and Behavior, 33*(2), 185–218.

Hoberman, H. M. (2016). Forensic psychotherapy for sexual offenders: Has its effectiveness yet been demonstrated? In A. Phenix & H. M. Hoberman (Eds.), *Sexual offending: Predisposing antecedents, assessments and management* (pp. 605–666). Springer.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55.

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.

Kearney, M. W. (2017). Cross lagged panel analysis. In M. Allen (Ed.), *The SAGE encyclopedia of communication research methods* (pp. 313–319). Sage Publishers.

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Publications.

Langan, N. P., & Pelissier, B. M. (2001). The effect of drug treatment on inmate misconduct in federal prisons. *Journal of Offender Rehabilitation, 34*(2), 21–30.

Latessa, E., Lugo, M., Pompoco, A., Sullivan, C., & Wooldredge, J. (2015). *Evaluation of Ohio's prison programs*. Center for Crime Justice Research, University of Cincinnati.

Linning, S. J., Bowers, K. J., & Eck, J. E. (2019). Understanding the time-course of an intervention’s mechanisms: A framework for improving experiments and evaluations. *Journal of Experimental Criminology, 15*(4), 593–610. https://doi.org/10.1007/s11292-019-09367-0

Linning, S. J., & Eck, J. E. (2018). Weak intervention backfire and criminal hormesis: Why some otherwise effective crime prevention interventions can fail at low doses. *The British Journal of Criminology, 58*(2), 309–331.

Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science, 3*, 297–320.

Lovell, D., & Jemelka, R. (1996). When inmates misbehave: The costs of discipline. *The Prison Journal, 76*(2), 165–179.

Lowenkamp, C. T., & Latessa, E. J. (2004). Understanding the risk principle: How and why correctional interventions can harm low-risk offenders. *Topics in Community Corrections, 2004*, 3–8.

Lugo, M., Wooldredge, J. D., Pompoco, A., Sullivan, C., & Latessa, E. J. (2019). Assessing the impact of unit management programs on institutional misconduct and prison “returns”. *Justice Quarterly, 36*(1), 59–92.
MacKenzie, D. L., & Lattimore, P. K. (2018). To rehabilitate or not to rehabilitate: That is the question for corrections! Criminology & Public Policy, 17(2), 355–377.
Milkman, H. B., & Wanberg, K. W. (2007). Cognitive-behavioral treatment: A review and discussion for corrections professionals. US Department of Justice, National Institute of Corrections.
Papp, J., Wooldredge, J., & Pompcoco, A. (2019). Timing of prison programs and the odds of returning to prison. Corrections, 1–26.
Pompcoco, A., Wooldredge, J., Lugo, M., Sullivan, C., & Latessa, E. J. (2017). Reducing inmate misconduct and prison returns with facility education programs. Criminology & Public Policy, 16(2), 515–547.
Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). Journal of Statistical Software, 48, 1–36.
Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), Handbook of developmental research methods (pp. 265–278). The Guilford Press.
Sherman, L. W. (1990). Police crackdowns: Initial and residual deterrence. Crime and Justice, 12, 1–48.
Shingles, R. D. (1976). Causal inference in cross-lagged panel analysis. Political Methodology, 3, 95–133.
Silver, I. A., & Nedelec, J. L. (2018a). Ensnarement during imprisonment: Re-conceptualizing theoretically driven policies to address the association between within-prison sanctioning and recidivism. Criminology & Public Policy, 17(4), 1005–1035.
Silver, I. A., & Nedelec, J. L. (2018b). Cognitive abilities and antisocial behavior in prison: A longitudinal assessment using a large state-wide sample of prisoners. Intelligence, 71, 17–31.
Smith, M. J., Clarke, R. V., & Pease, K. (2002). Anticipatory benefits in crime prevention. Crime Prevention Studies, 13, 71–88.
Steiner, B., Butler, H. D., & Ellison, J. M. (2014). Causes and correlates of prison inmate misconduct: A systematic review of the evidence. Journal of Criminal Justice, 42(6), 462–470.
Steiner, B., & Wooldredge, J. (2008). Inmate versus environmental effects on prison rule violations. Criminal Justice and Behavior, 35(4), 438–456.
Steiner, B., & Wooldredge, J. (2014). Comparing self-report to official measures of inmate misconduct. Justice Quarterly, 31(6), 1074–1101.
Strimple, E. O. (2003). A history of prison inmate-animal interaction programs. American Behavioral Scientist, 47, 70–78.
Tewksbury, R., Connor, D. P., & Denney, A. S. (2014). Disciplinary infractions behind bars: An exploration of importation and deprivation theories. Criminal Justice Review, 39(2), 201–218.
Vuk, M., & Doležal, D. (2019). Idleness and inmate misconduct: A new perspective on time use and behavior in local jails. Deviant Behavior, 41(11), 1347–1369.
Wooldredge, J. D. (1998). Inmate lifestyles and opportunities for victimization. Journal of Research in Crime and Delinquency, 35(4), 480–502.