Evaluation of Reentry-Based Restorative Justice

A Randomized Controlled Trial

Roger Jarjoura, Derrick Franke, Juliette Roddy, Paul Draus, Joseph Williams, Colton Currah, Ariel Roddy, Samuel E. DeWitt, and Rebecca Abbott

American Institutes for Research

University of Michigan-Dearborn

September 2025

This project was supported by Award Number: 2019-RY-BX-0002, awarded by the National Institue of Justice, Office of Justice Programs, U.S. Department of Justice. The opnions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily relect those of the Department of Justice.

Table of Contents

Exhibits

Introduction

Although research continues to find a strong relationship between family support and successful reentry, reentry strategies in the United States remain narrowly focused on those committing crime. Whether risk-based (surveillance and control) or needs-based (rehabilitation and services), the traditional reentry landscape leaves little room for targeting criminogenic risk factors beyond justice-involved individuals (JIIs) themselves. As the late criminologist Leslie Wilkins noted decades ago, "The problem of crime cannot be reduced to the problem of the criminal," (Wilkins, 1991). More recently, experts have called for an innovative reentry approach, one that is, by definition, more harm-, community-, and family-based, emphasizing accountability in the broader context of informal social control and support (Bazemore & Erbe, 2004).

Family involvement during reentry may vary with the relationship before, during, and after incarceration (Ganem & Agnew, 2007). A JII's history of family violence, substance use, or criminal behavior can discourage efforts to rebuild family relationships. Some familiar may even dissolve these relationships during incarceration (Hairston, 2003). Additionally, events leading up to incarceration can create harm within the family, with little opportunity to heal during the period of incarceration.

Unfortunately, the experience of incarceration can lead to a deterioration in relationships with family members. Families may have to travel long distances to a prison and face daunting security procedures. When they arrive, many find that the visits do not foster positive parent-child interaction (Hairston & Oliver, 2006). Because of changes in family structure and financial circumstances when a JII is out of the home, the roles that family members play may change, particularly relative to parenting (Braman & Wood, 2003) Such changes typically require some

renegotiation of roles when the JII returns to the family, which may contribute to stress and undermine the likelihood for effective reentry (Martinez, 2008).

One promising approach to addressing the challenges of reentry is restorative justice. Restorative justice (RJ) begins with the idea that crime primarily causes harm to people and relationships and, in turn, creates an obligation to repair that harm. Restorative conferencing (of which family group conferencing [FGC] is one type) brings together the people most affected by a crime to talk about what happened, learn how people have been affected, and decide what needs to be done to repair the harm.

Typically, RJ is understood as a "front-end" justice intervention whereby lower-risk or juvenile JIIs are diverted from court in favor of face-to-face meetings with stakeholders of the offense. A growing evidence base suggests crime-reducing and victim-healing potential for RJ interventions (McCold & Wachtel, 1998; McGarrell & Hipple, 2007; Sherman & Strang, 2007; Sherman et al., 2013). However, the same analyses suggest that RJ conferencing may yield positive results even with high-risk adult JIIs who commit serious and repeated offenses, and that these results can be achieved after conviction. Given the emotionally intense nature of RJ conferences, the reasoning is that crimes that cause more harm also produce more need for repair and therefore more opportunity for transformation. More recently, studies suggest improved crime control, cost savings, and victim benefits (such as reduced posttraumatic stress symptoms) when RJ is used with serious offenses and adult JIIs (Angel, 2005; Daly, 2006; Hayes & Daly, 2004). RJ practices have been shown to be promising and effective, as indicated on CrimeSolutions.gov (Bergseth & Bouffard, 2007; Rojek et al., 2003; Sherman et al., 2000; Umbreit & Coates, 1992).

The Current Study

This study draws on a prior small-scale analysis of FGC for adult male inmates nearing reentry from prison in Indianapolis, Indiana (Franke & Jarjoura, 2016). In that study, 17 high-risk inmates were randomly assigned to either the treatment condition (restorative conference with their family) or control condition (no conference). Three years after release, half the men in the control condition had been reincarcerated, and none of the men who received a conference had returned to prison. Six years after release, 83% of those in the control condition were back behind bars compared to only 30% of those in the treatment condition (Franke & Jarjoura, 2016). Although lowering recidivism by 50% is certainly promising, the small sample size limited researchers' ability to draw strong conclusions about the effectiveness of reentry-based RJ.

The purpose of this research is to study the capacity of FGC to enhance public safety by equipping returning citizens with the skills and insight to repair relationships and reintegrate successfully into their communities and families. The Supporting Restorative Reentry in Detroit (SRRD) program implements RJ through residential programs at a residential reentry center located in Macomb, a suburb of Detroit. Self-Help Addiction Rehabilitation (SHAR), a Detroit-based therapeutic community, in collaboration with the Michigan Department of Corrections (MDOC), is the setting for the work of SRRD. The current evaluation involves a randomized controlled trial (RCT) of FGC for high-risk, formerly incarcerated individuals still under MDOC supervision but currently residing in SHAR, with the goal of enhancing outcomes in two major areas: (1) reducing recidivism and (2) increasing quality of life for returning citizens, their families, and their destination communities.

The goal of this project is to conduct a rigorous evaluation of the effectiveness of FGC for returning JIIs and their families. Objectives of the study include conducting (1) a process evaluation that includes a focus on fidelity of implementation, (2) an impact evaluation using a randomized controlled design, and (3) a study of the costs and benefits of the intervention. The report that follows begins with a review of the relevant research literature, then describes the project design, methods, and implementation, presents the process, impact, and cost study results, and concludes with a discussion of implications for practice and policy

Review of Relevant Literature

Bushway and Paternoster (2013) suggested that criminal activity declines when individuals achieve a "structural break" in identity such that decisions, choices, and actions are shaped by a new conception of self. The new identity would provide a foundation from which different (i.e., prosocial) decisions, choices, and actions would arise. Bushway and Paternoster (2013) theorized that when a JII envisions a future noncriminal identity as more beneficial and attractive than the previous identity as a criminal and/or drug addict, then the person is in the process of desisting from criminal activity.

Research shows, however, that it may not be enough for JIIs to envision or even aspire to a more prosocial identity in the future. Instead, actions, choices, and decisions are best shaped by a personal commitment to their current identity. Adopting a new identity is likely to take time, but systematic change in behavior that translates into a reduction in recidivism is more likely once the person has practiced and embraced the new identity (Bushway & Paternoster, 2013).

Indeed, taking initial steps to change one's sense of identity will often require structural and social supports (Bushway & Paternoster, 2013). Structural supports may involve any number of reentry programs, treatment programs, or support groups (Kiecolt, 1994). Social supports can

include peer support, family support, or employment (Bushway & Paternoster, 2013). These supports reinforce and may even serve to maintain the returning JII's transformed identity.

JIIs face many barriers to embracing a new prosocial identity. Potential employers pay particular attention to the indicators of the previous identity (i.e., criminal record, gap in employment history). Peers may reinforce old patterns of behavior that remind the individual about who they were in the past. Family members may struggle to get past the pain and mistrust that they still feel because of the JII's previous offenses and incarceration and therefore may express skepticism about the likely success the individual will have in turning his or her life around.

A new identity may also evolve with the assistance of reintegration rituals, as described by Maruna (2011). When reintegration involves "atonement, forgiveness, redemption, and reconciliation," the process is likely to be successful (Braithwaite & Mugford, 1994). A process like FGC that provides an opportunity for the JII to make amends and achieve mutually satisfying reconciliation is one potentially powerful approach (Maruna, 2011). Signs of generativity also indicate that a person is adopting a new identity. The core concept of generativity is expanding one's perspective to think about others, rather than focusing only on him- or herself (McAdams & de St Aubin, 1998). Generative commitments (e.g., deciding to be a better parent or mentoring another person) provide a path by which a sense of purpose and meaning can be achieved, as well as a feeling of redemption. Acting generatively can also legitimize a JII's claim to changing their behavior (McNeill & Maruna, 2007).

Finding ways to build in social support for efforts to make amends and move forward are important, and the family is likely to be a key source of social support. Family support has been identified as among the most important protective factors for reducing the likelihood of

recidivism and increasing the likelihood of desistance and healthy adjustment (i.e., sobriety, addressing mental health issues) to life after incarceration (Sullivan et al., 2002). Families can experience a substantial burden due to the challenges that formerly incarcerated individuals face after their release. So, it is critical to maximize the likelihood that family relationships and the capacity of the family members to support the JII are both strong (Naser & La Vigne, 2006).

Little is known, however, about the mechanisms by which RJ programs reduce reoffending, or more broadly how RJ fits in the context of prisoner reintegration. Recent theoretical developments suggest moving beyond the traditional conceptualization of RJ as an isolated "program" and instead situate restorative principles within broader models of reentry. In both Bazemore and Stinchcomb's civic engagement model of reentry, and Bazemore and Maruna's restorative reentry model, emphasis is placed on processes such as identity change, generativity, and redemption (Bazemore & Maruna, 2009). The current study builds on the latter model.

Project Design and Implementation

The Intervention: Family Group Conferencing

FGC at SHAR follows the International Institute for Restorative Practices guidelines for best practices in RJ. This means the following:

- FGC is voluntary for all stakeholders (JII, family, and community).
- The JII must take full responsibility for his or her crime to be eligible.
- The crime must have caused harm to the family, and there must be a need to repair that harm.
- Conferences and preconference preparation are conducted by an International Institute for Restorative Practices–trained facilitator and follow a consistent, scripted model.

Two teams carried out the evaluation in collaboration with one another. The Design and Analysis Team designed data collection protocols and instruments, and the random assignment of participants to the treatment and control groups. The Intervention Team enrolled participants into the study, ensured the fidelity of the delivery of the FGC, and collected survey data from study participants. In Exhibit 1, we list the key tasks for each of the three phases of this study.

Exhibit 1. Study Phases and Tasks

Phase I: Pre randomization
Clients complete the intake process upon entering the SHAR facility, including orientation to FGC.
During first week at facility, clients are recruited for study: participants request to begin screening process.
Participants complete baseline survey.
Trained facilitators assess appropriateness of conferencing for client.
Conduct interviews with leadership team for SRRD and SHAR administrators.
Document process leading up to conferences for each participant.
Contact families to determine willingness to participate in FGC.
Assign participants randomly to FGC or control group.
Phase II: Post randomization/Pre release
Final preparations are made for conferences.
Conference takes place.
Restoration agreement made to help support client for an effective reentry.
Follow-up survey completed prior to release along with tracking form.
Preparation process for conferences and development of restoration agreements is documented.
Phase III: Post release
Perform follow-up interviews 6 months after release.
Analyze administrative recidivism records data.

When entering the SHAR center, individuals (referred to as clients) go through an intake process. In addition to introducing clients to rehabilitative programming such as substance abuse treatment, trained facilitators provide orientation to FGC. The orientation, like other phases of conference preparation, incorporates restorative elements itself (e.g., asking clients to reflect on the harm they caused to their families, using affective questions and statements, sitting in a

circle, using a talking piece, taking responsibility, and discussing making amends). The primary goal of these initial steps, referred to as *restorative practices*, is to initiate a reflective process in which clients can think deeply about the harm their actions have caused others and how they can begin to repair it. Restorative *justice*, on the other hand, refers specifically to the face-to-face conference itself. A client at SHAR could therefore receive some level of restorative practices and subsequently decline or be determined ineligible for restorative justice (FGC).

As shown in Exhibit 2, an individual must meet two conditions before enrolling in the evaluation. First, participants must volunteer and be determined eligible for conferencing. Second, the person's family members must also volunteer and be determined eligible. Random assignment occurs at this point. This design overcomes a common limitation in RJ research in which JIIs who volunteer for conferencing are compared to a control group of non-volunteers. Like the design in Indiana, randomization after screening and all preparatory restorative practices yields a pool of equally motivated individuals who have received the same dosage of any initial restorative practices.

Exhibit 2. Flow of Participants in Treatment and Control Groups

<table>
<tr><th colspan="2">Before Randomization</th></tr>
<tr><td colspan="2">Intake orientation.
Participants volunteer for FGC.
Family commits to participating in conference.</td></tr>
<tr><th colspan="2">After Randomization</th></tr>
<tr><th>Treatment Group</th><th>Control Group</th></tr>
<tr><td>Participates in FGC.
Completes a survey prior to release.
Completes interview 6 months after release.
Data from law enforcement and corrections collected for recidivism analysis.</td><td>Completes a survey before release.
Completes interview 6 months after release.
Data from law enforcement and corrections collected for recidivism analysis.</td></tr>
</table>

During FGC, a primary focus is that the JII is held accountable to the family. To be eligible for FGC, the JII must take full responsibility for the offense and be prepared to directly face the harm they have caused to their family. This model differs from some RJ efforts in which the victims, not the family, are involved in conferences. Often in the wake of crime and incarceration, the JII's family becomes a "hidden victim" otherwise ignored in the reentry process.

The International Institute for Restorative Practices recommends a formally structured conferencing process that covers three phases: what happened, how people were affected, and what needs to be done to repair the harm. Conferences typically last 1–3 hours, depending on the number of participants. In that time, JIIs can begin to develop empathy for those they have harmed, build a more prosocial identity for themselves, and create a realistic plan for their return. Conferences conclude with a restoration agreement outlining what the JII agrees to do after the conference to repair the harm.

Theory of Change

In this section we lay out the theory of change related to the use of FGC for those preparing for reentry. Bushway and Apel introduced the concept of signaling as it relates to the reentry process (Bushway & Apel, 2012). They suggest that JIIs may have ways of providing a signal that they are desisters so those thinking about giving them an opportunity (e.g., employers) can make smarter choices regarding whom to offer a second chance. Analogous to the scenario of the employers taking a chance on the JIIs that are signaling desistance, we suggest that family members who may be somewhat apprehensive about an FGC will set aside their apprehension when they see that the JII has signaled.

We believe that an individual's request for RJ can also signal desistance, particularly in the context of corrections/reentry. Many steps are needed for an RJ conference in prison to occur. First, JIIs must initiate the process by volunteering. Then they must convince the RJ coordinator that they volunteered for the right reasons. The RJ coordinator must reach out to the victims or families and present the request for their participation, underscoring that the request from the JII is genuine and that the conference itself is safe. Finally, the JII must prepare for and actually go through with the conference and not withdraw.

Experts agree that for FGC to be most effective, participation should be voluntary (i.e., not coerced) and motivated by a desire to repair the harm that has been caused. Yet, when RJ is offered as a diversion from court, for example, JIIs may choose to participate simply to avoid formal punishment and not from a desire to apologize or make amends. In these cases, the request to participate in RJ is not likely to be a potential signal for desistance.

When requests by the JII to participate in RJ within correctional facilities take place, this should more often fit the conception of signaling because formal punishment has already happened (by way of a prison sentence). Offering no additional incentives, such as time-cuts, should further increase confidence that JIIs who choose RJ are signaling desistance and not something else.

Before bringing parties together, facilitators determine the appropriateness of a conference. Because RJ focuses on repairing harm, the facilitator must establish that some level of harm was caused and that the harm should be repaired. If during the preparation phase, the facilitator learns that the offense harmed only the JII or if the family has already repaired the harm on their own, they need not proceed with conferencing. Facilitators are trained to look for signals that would jeopardize the RJ process. Maruna wrote, "The job of much social interaction is to differentiate

between ‘authentic’ signs and manipulation—discerning the genius from the con artist, the ‘gangsta’ from the ‘wannabe’ and so forth,” (Maruna, 2012). In RJ, the con artist might be the JII who shifts blame or volunteers for the wrong reasons. Do they really want to earn a second chance (authentic) or do they just want to get another visit from their family (manipulation)? Is the JII ready to talk honestly about his offense, listen actively to what his family has to say, and take action to make amends? These are the criteria by which the true signalers emerge in the preconference phase.

Does signaling cause desistance? If the individuals who signal are destined to become desisters, then the RJ conference should not make a difference. In other words, we should see similar recidivism outcomes for the treatment group (those assigned to participate in RJ conference) and the control group (those assigned to not receive a RJ conference). By contrast, if RJ is critical for desistance/redemption, then we should see a difference in outcomes for those in treatment group versus those in the control group. In this case, we expect that the control group subjects will not desist. In the treatment group, whether the individuals desist will be because of the effect of the RJ conference on the likelihood of desistance.

By the end of the conference, if successful, the family members become open to the notion that the JII is going to be a present and active member of the family with expectations that he or she will honor commitments and show respect and regard for others. The family now has a transformed impression of the JII that is more positive and more hopeful. Finally, the restoration agreement kicks off the willingness of the family to provide the kinds of informal support and assistance that will contribute to an effective reentry. We believe the conference itself is an initial step in a feedback loop that will further enhance the process.

In their development of a theory of identity related to desistance, Bushway and Paternoster pointed to the importance of behaviors that "structure and support" the decision to desist (Bushway & Paternoster, 2013). Such behaviors in the context of RJ in reentry may include steps to carry out the elements of the restoration agreement. These actions are potential indicators that the person is committed to desistance and may, in turn, further strengthen family relationships to provide additional motivation and incentives for solidifying the commitment to desistance.

Consistent with the belief that identity change is necessary for desistance to occur, Nakamura and Bucklen (2014) pointed to cognitive change as the necessary ingredient for solidifying a new identity. Cognitive change may lead the JII to engage in signaling. This signaling may be the result of changes in motivation and commitment, which are then further reinforced by the experiences in the RJ conference and the subsequent interactions with the family after the conference is concluded.

Sample and Methods

Study Sample

At the outset of this evaluation, we sought to recruit 300 residents from SHAR Macomb to participate in the study. However, between 2019 and 2024, the implementation of our randomized controlled trial faced several challenges, most notably disruptions caused by the COVID-19 pandemic. In early 2020, in-person programming at SHAR Macomb was suspended due to the COVID-19 pandemic, and the facility's population dropped significantly due to halted MDOC referrals and safety restrictions. Staff illness, remote work arrangements, and the inability to conduct face-to-face orientations and conferences further delayed recruitment and intervention delivery. These constraints persisted across multiple periods, requiring the research team to rethink its approach to recruitment, facilitation, and data collection.

To adapt, the team transitioned to virtual programming in 2021, implementing Zoom-based orientations, one-on-one volunteer sessions, and remote FGC. This shift was supported by grant-funded technology and close coordination with SHAR staff. Virtual conferencing not only maintained fidelity to the intervention model but also increased accessibility for family members who might not have been able to attend in person. Additional facilitators were trained, including individuals with lived experience, and recruitment was expanded to SHAR Detroit and the Detroit Recovery Project to diversify the sample and increase enrollment. Randomization procedures were also modified to increase the likelihood of treatment assignment from 50% of volunteers to 75%, increasing the number of volunteers exposed to the intervention. The virtual model proved effective in maintaining engagement and data integrity, and the expansion to additional sites helped mitigate earlier recruitment shortfalls. These adaptive strategies ensured the continuation of the study and strengthened its empirical foundation despite significant external challenges.

Of those who initially volunteered for the study (n=108), seven dropped out, leaving us with n=101 FGC volunteers who were successfully randomized. Approximately 75% of randomized participants were assigned to the treatment group (n=68), with the remainder placed in the control group (n=34). Prior to the administration of FGC, 21 members of the treatment group members declined to continue participating in the study or were lost to follow-up, leaving us with a treatment sample of n=47 who participated in FGC. To provide additional context for our comparisons of the treatment and control groups, we also incorporated a separate comparison group of SHAR residents who did not volunteer for FGC (n=160). Members of the comparison group did not participate in the surveys we administered to the treatment and control groups, but we did obtain recidivism data for this group. Some observations were lost from duplicate IDs

(n=4), respondents appearing in the survey data but not the recidivism data (n=5), respondents who died under supervision (n=7) and respondents we cannot observe for a fill 12-month follow-up window (n=1). Our final analytical sample included n=153 participants in the comparison group, n=30 in the control group, and n=39 in the treatment group. Exhibit 3 provides descriptive statistics for each of these groups.

Exhibit 3. Demographic Descriptives for Study Samples

Sample Characteristic	Comparison	Control	Treatment
N	153	30	39
Age (Mean (SD))	37.84 (10.00)	35.90 (8.46)	36.93 (8.84)
Sex (%)			
Male	105 (68.6)	18 (60.0)	25 (64.1)
Female	47 (30.7)	12 (40.0)	14 (35.9)
Unknown	1 (0.7)	0 (0.0)	0 (0.0)
Race (%)			
White	115 (75.2)	26 (86.7)	29 (74.4)
Black	22 (14.4)	2 (6.7)	7 (17.9)
Other	8 (5.2)	2 (6.7)	2 (5.1)
Unknown	8 (5.2)	0 (0.0)	1 (2.6)
Year (%)			
2019	24	43	33
2020	07	00	03
2021	21	17	15
2022	23	13	15
2023	19	20	26
2024	06	07	08

Methods

Our evaluation employs a multi-component design to rigorously assess the SRRD program. We structure the methods into three interrelated components: a process analysis, an impact analysis, and a cost analysis. Together, these components allow us to assess not only whether the conference improves outcomes for participants, but also how FGC participants experience their conference and the extent to which benefits outweigh costs. The following sections provided details for the samples, measures, and analytic methods we employ for each of these analyses.

Process Study

Our process evaluation is a critical component of the study, designed to understand how SHAR participants perceive their treatment throughout the conference and its helpfulness. By addressing this aspect of the FGC, the process evaluation will help to identify best practices and areas for improvement, ultimately contributing to the effectiveness and sustainability of the intervention.

The process evaluation will focus on a key research question: (P1) How do FGC participants experience and perceive the FGC as delivered by SRRD staff? This question will be explored through a combination of closed- and open-ended survey responses from FGC participants about their perceptions of the conference. This approach will ensure an understanding of FGC implementation and its impact on the program's outcomes through the lens of SHAR clients who participated in the FGC process.

Process Study Sample

The process study sample consists of SHAR residents who were randomly assigned to the treatment group and participated in the FGC (n=47). Within the pre-release survey for the treatment group, we included closed- and open-ended questions about their perception of the FGC. One individual in the treatment group refused to answer nearly all of the survey questions about the FGC, so our effective sample for the analysis of survey responses is 46 subjects.

Process Study Measures

Survey Data on FGC Perceptions. To capture treatment group members' perceptions of the FGC, we administered a series of survey questions focused on participants' experiences. Participants reflected on their own emotional responses including feelings of guilt, shame, forgiveness, and care, as well as whether they felt respected, judged, or stigmatized during the process. They also evaluated the fairness of the procedure, indicating whether their rights were respected, whether they were given the opportunity to express their views, and whether their input was taken into account in decision-making. In addition, participants assessed the conference's practical value, including whether it added to the treatment they received at SHAR, supported their transition back into the community, and reduced the likelihood of reoffending or relapse. Finally, participants rated the quality of facilitator performance, including preparation, listening, impartiality, and ability to create a safe and supportive environment. Open-ended questions further invited participants to describe how their attitudes and feelings changed after the conference, and how they believed the process would influence their reentry.

Impact Study

We designed our impact analyses to rigorously test whether FGC improves both intermediate and recidivism outcomes for returning citizens. Scholars such as Bushway and Paternoster (2013) have argued that desistance from crime is best understood as a process of identity transformation—shaped by supportive structures and opportunities for generativity—rather than a discrete event. Restorative justice approaches like FGC directly engage these mechanisms by fostering empathy, accountability, and reconciliation between returning citizens and their families. Our impact analysis therefore examines whether FGC can produce measurable improvements in psychosocial outcomes and recidivism.[1] We frame our analysis around four primary research questions:

Exhibit 4. Research Questions for Impact Evaluation

Research Question	Plan to Address Research Question
I1. Does participating in FGC result in significant positive changes in belief in redeemability, empathy, remorse, and accountability, and generativity as compared to individuals who did not participate in FGC?	Analysis of the difference in response values across the initial and pre-release surveys
I2. Does participating in the FGC result in significant reductions in criminal identity/core self compared to individuals who did not participate in FGC?	Analysis of the difference in response values across the initial and pre-release surveys
I3. Are FGC participants less likely to have a recidivism event compared to individuals who did not participate in FGC?	Analysis of MDOC data on recidivism events that occur within the state
I4. Do moderators account for significant variation in treatment effects for each of the relationships tested in questions I1–I3?	Analyses incorporating data on the social bonds/relationships of the participants

[1] We also intended to conduct a mediation analysis of recidivism outcomes as mediated through observed changes in the survey constructs from the initial to pre-release surveys. Unfortunately, these models were not estimable due to a low sample size and sparse recidivism events.

Impact Study: Survey Analysis

Survey Design and Implementation

We designed the survey component of the evaluation to capture the intermediate psychosocial outcomes that the literature identifies as central to the desistance process. Decades of criminological theory emphasize that offenders desist not only when external constraints are applied, but when internal identity shifts and stronger social supports take hold. Bushway and Paternoster (2013) argued that desistance occurs when individuals develop a new, prosocial self-concept that is more appealing and sustainable than a prior criminal identity, while Maruna (2001) highlighted the role of "reintegration rituals" that enable forgiveness, atonement, and the building of generative commitments. Drawing on these theoretical foundations, we administered a pre/post survey including four validated constructs to assess belief in redeemability, criminal identity, empathy/remorse/accountability, and generativity. For the full list of survey items organized by construct, please see Exhibit A1 in Appendix A. The following sections provide additional information about the individual survey constructs and our implementation of the survey.

To gain additional perspective on post-release wellbeing, we also planned to conduct six-month follow-up interviews for all individuals in the treatment and control groups. To facilitate these interviews, we collected information from study participants about two to three close contacts we could reach out to if we had difficulties contacting them for a follow-up interview. Following each JII's release from a SHAR facility, we conducted monthly check-ins with the volunteers and their contacts to ensure that our contact information remained valid. Despite our efforts to mitigate attrition for the six-month follow-up interviews, we were only

able to conduct 9 of the originally planned 69 interviews for our analytic sample – leaving us with an interview rate of roughly 13%. As such, we are unable to use these data to provide additional context for our impact analyses.

Survey Constructs

Four Confirmatory Factor Analysis (CFA) models were defined to assess four theorized constructs among the initial and pre-release survey items: belief in redeemability, criminal identity/core self, empathy/remorse/accountability, and generativity. All CFA models were estimated using the lavaan package in R, with robust maximum likelihood (MLR) to account for non-normality and Full Information Maximum Likelihood (FIML) to handle missing data. This approach ensures more accurate standard errors and retains cases with incomplete responses to preserve the available sample data. In the following sections, we provide a review of the research supporting each of these constructs and the results from our CFA models.

Belief in Redeemability. The Belief in Redeemability (BIR) construct measures the extent to which offenders believe they can successfully desist from crime and rebuild a valued role in society. Building on Maruna's (2001) "redemption script" framework, O'Sullivan and colleagues (2018) developed and tested a self-report scale derived from statements made by offenders about their chances of "going straight." Their study demonstrated that BIR is a measurable construct with meaningful variation across offenders, skewed toward optimistic beliefs. Using a combination of card-sorting and psychometric testing, O'Sullivan et al. (2018) found that 24 of the 37 items in the complete scale mapped onto three underlying dimensions of BIR: agency (e.g., "I am just a little piece in a big game"), belonging (e.g., "I can be a positive member of society"), and optimism (e.g., "I could be happy going straight").

The CFA model for the BIR construct included three first-order latent constructs—agency, belonging, and optimism—loading onto a higher-order BIR factor. While item loadings were generally statistically significant on the three first order factors, the standardized loadings for each were modest with only two items having items over .80 for belonging and optimism and agency having only one item over .8 indicating a modest ability of the measures to capture the first order constructs. While the first order factors loaded strongly onto the higher order factor (agency=.990, belonging=.95, and optimism=1.004) , overall model fit was suboptimal (Robust CFI = 0.763, Robust TLI = 0.736, RMSEA = 0.114; $\chi^2(227) = 570.95$, $p < .001$), suggesting potential areas for model refinement.

Criminal Identity. To measure the extent that SHAR participants in our sample identify with criminal attitudes/belief and internalize that into their own self-perception, we included multiple items (e.g., "Being a criminal is an important part of my self image") from the Measure of Criminal Social Identity (MSCI) (Boduszek et al., 2012). Prior research testing the MSCI construct with a sample of individuals who recidivated after release from prison supported a three-factor structure, including cognitive centrality (i.e. the importance of criminal identity), in-group affect (pride felt about belonging to a criminal group), and in-group ties (sense of personal connection with other individuals who identify as criminals) (Boduszek et al., 2012).

The criminal identity/core self CFA model shows poor fit between the model and the data validity (Robust CFI = .383, Robust TLI = .177, Robust RMSEA = .218, SRMR = .174; $\chi^2(27) = 163.31$, $p < .001$). The strongest standardized loadings of indicators for criminal core were items where the respondent reported others believing they deserve another chance (.763), belief they have changed for the better (.747), and people they care about believe they have changed for the better (.704). All other items showed a standardized loading below .2 indicating poor alignment

with the underlying construct. High modification indices point to correlation between the items in the factor not explained by the criminal core latent construct, potentially due to similar wording or overlapping item content. For example, the questions related to criminal and addict self-image have an MI=32.4.

Empathy/Remorse/Accountability. To account for changes in empathy, remorse, and accountability, we include survey items drawn from the Interpersonal Reactivity Index (IRI), a widely used and psychometrically validated measure of empathy (Pulos et al., 2004). Prior research indicates that the IRI captures emotional and cognitive aspects of empathy, making it suitable for assessing relates outcomes such as remorse (through empathic concern and perspective taking) and accountability (through recognition of the impact of one's action on others). To capture these dimensions, we adopted seven survey items from the IRI, including questions tapping into empathic concern/perspective taking (e.g., I understand how my actions have affected my family) and recognizing the impact of one's actions (e.g., There are relationships I damaged that need to be repaired).

The empathy/remorse/accountability CFA model demonstrated acceptable fit (Robust CFI = .928, Robust TLI = .892, Robust RMSEA = .102, SRMR = .069; $\chi^2(14) = 30.76$, $p = .006$). Item loadings show three strong indicators with moderate to high standardized loadings: putting themselves in their family's place (.953), trying to put themselves in their family's shoes (.866), and trying to see things from their family's perspective (.625). The rest of the items showed loadings at .416 and lower. These results suggest the empathy remorse construct is best captured by items emphasizing empathetic reflection and a family centered perspective.

Generativity. We use a reduced version of the Loyola Generativity Scale (LGS) (McAdams & de St. Aubin, 1992) to assess the extent to which SHAR participants in our sample are sensitive to the legacy they leave behind through a commitment to create and maintain positive contributions to society. We adopted five items from this scale (see Exhibit A1 in Appendix A for the complete list) based upon extant research findings indicating that formerly incarcerated individuals enact generativity through mentoring, advocacy, and peer support in an effort toward giving back to their communities, reconciling their pasts, and constructing new prosocial identities (Lebel et al., 2015). Incorporating these items from the LGS into the current survey allows us to test whether Family Group Conferencing fosters this orientation, supporting reentry by strengthening participants' sense of purpose, responsibility, and capacity to "give back" to family and community.

The generativity CFA model shows poor overall fit and limited construct validity (Robust CFI = .842, Robust TLI = .684, Robust RMSEA= .151, SRMR=.078 ; $\chi^2(5) = 19.65$, $p = .001$). The highest standardized item loadings were the sense of having made a meaningful difference (.811) or having created an impact (.779). The rest of the items had weak loadings of .402 or lower.

Summary of CFA Model Results

The primary aim of estimating our CFA models was not to validate the latent constructs of the established factors and their measures. Instead, our goal was to assess the factor model's fit for use as predictors. Overall, the results indicated some item sets functioning as expected in these established scales whereas some factors showed weaker internal consistency. Exhibit 5 provides summary statistics for each of the constructs for the total, control, and treatment samples.

Exhibit 5. Descriptive Statistics for Survey Measures

Post Pre Differences in Factor Scores	Total	Control	Treatment
Belief in Redeemability (Mean(n))	.002(58)	.058(20)	-.028(38)
Criminal Identity/Core Self (Mean(n))	-.364(57)	-.400(19)	-.346(38)
Empathy/Remorse/Accountability (Mean(n))	.090(58)	.090(20)	.090(38)
Generativity (Mean(n))	.022(58)	-.004(20)	.036(38)

Considering the weak internal consistency across these factors, using them as predictor or outcome variables may result in less stable or less detectable impact in our study. It is important to note that this is not a reflection on these established scales and their theorized constructs but more likely a direct result of our limited sample size. The suggested sample size for factor analysis is a ratio of 20 observations for every one parameter in your model, with a minimum of 100 (Kline, 2016) while our sample size was 58.

Analytical Method

Our analytic strategy builds directly on criminological research showing that empathy, remorse, and accountability, belief in redeemability, prosocial identity formation, and generativity are among the strongest predictors of successful reentry (Duwe, 2012; Lebel et al., 2015; Paternoster et al., 2016). For example, Bushway and Paternoster (2013) emphasized that returning citizens are more likely to desist when they envision a new noncriminal identity supported by strong social relationships, while Maruna (2001) described the importance of "reintegration rituals" that provide opportunities for forgiveness, atonement, and redemption. By measuring these constructs before and after the FGC, we can determine whether FGC accelerates these processes relative to a control group of similarly motivated returning citizens.

Accordingly, we computed change in standardized factor scores for each construct as pre-release minus initial survey scores and compared mean change between groups using two-sample, one-sided t-tests. Directional hypotheses were pre-specified: for belief in redeemability, empathy/remorse/accountability, prosocial identity formation, and generativity, the alternative tested whether the treatment group showed greater positive change than the control group; for criminal identity, the alternative tested whether the treatment group showed lower change (i.e., a larger decrease) than the control group, consistent with reductions in criminal self-concept. All tests were conducted on standardized score changes to place constructs on a common metric, and one-sided p-values are reported in the direction of these hypotheses.

Impact Study: Recidivism Analysis

Data and Measures

To address research question I3 (see Exhibit 4), we analyze official administrative data provided by the MDOC to assess the program's impact on recidivism. These records include information for all individuals in the treatment (n=47), control (n=34), and comparison groups (n=160) about new offenses or technical violations that resulted in a SHAR resident being reincarcerated or placed on probation from the time of their release from a SHAR facility until December 31, 2024. These data only include information about the first offense post-release and, as such, we are not able to measure the number of recidivism events in this analysis. For the purposes of this analysis, we only consider new offenses as a recidivism event and exclude new commitments to MDOC custody for technical violations.

Additionally, we focus on the first twelve months after release as that is the minimum amount of time we can observe many individuals in our sample. It is important to also note that MDOC provided an indicator for whether a member of the study passed away while under

community supervision (n=8). We removed these observations from our analysis as we are either unable to determine when the death occurred (n=7) or the death occurred earlier than twelve months post-release (n=1).

Finally, as is the case in many analyses using administrative recidivism data, we are unable to verify that all the individuals in our sample remain at risk for new offenses or technical violations during the entirety of the twelve month follow up window. For example, although we have information about deaths on probation, we lack this information for individuals in our sample who are not under community supervision. Similarly, we only have information about recidivism events that occur within Michigan and are unable to determine if individuals commit new offenses in a different state. In consideration of these limitations, our measure of recidivism is likely to be an underestimate of the true level of recidivism in this sample.

COMPAS Scores. All individuals who enter a SHAR facility undergo a risk assessment using Northpointe's COMPAS (Correctional Offender Management Profiling for Alternative Sanctions). This tool categorizes risk for reoffending as low, medium, or high. SHAR residents assessed at either medium or high risk were eligible to participate in this study. The predictive validity of COMPAS, relative to other prominent risk assessment instruments, has been demonstrated in validation studies (Blomberg et al., 2010; Brennan et al., 2009; Fass et al., 2008).

For the purposes of our analysis, we use four subscales from the COMPAS assessments: (1) general recidivism risk, (2) violent recidivism risk, (3) substance abuse needs, and (4) social isolation needs. The scales for recidivism risk incorporate information about the individuals' criminal history (e.g., number of prior arrests/convictions, age at first conviction) which is then

scored from values of 1 (lowest risk) to 10 (highest risk). MDOC then uses these values to categorize risk of general or violent offending into three categories: low (1 to 4), medium (5 to 7), and high (8 to 10). By contrast, values for the substance abuse and social isolation needs scores incorporates information about the individuals' substance abuse and treatment history (e.g., frequency of drug/alcohol use, prior treatment attempts) and their level of social isolation (e.g., social support, community involvement) which is then scored from values of 1 (lowest probability of need) to 10 (highest probability of need). MDOC then uses these values to categorize need into three categories: unlikely (1 to 4), probable (5 to 7), and highly probable (8 to 10). Exhibit 6 provides descriptive statistics for the treatment, control, and comparison samples, including the prevalence of new offenses and scores on each of the COMPAS subscales.

Exhibit 6. Descriptive Statistics for New Offenses and COMPAS Scores

Sample Characteristic	Comparison	Control	Treatment
N	153	30	39
COMPAS (Mean (SD))			
General recidivism risk	2.41 (0.60)	2.60 (0.50)	2.46 (0.64)
Violent recidivism risk	1.92 (0.73)	2.10 (0.80)	2.08 (0.77)
Substance abuse needs	2.71 (0.56)	2.77 (0.43)	2.82 (0.51)
Social isolation needs	1.84 (0.85)	1.67 (0.76)	1.85 (0.90)
New Recidivism Event in 12mos (%)	6.00	4.00	3.00
Type of Recidivism Event (%)			
Drug	55.56	0.00	0.00
Property	33.33	100	0.00
Violent	0.00	0.00	0.00
Other	11.11	0.00	100.00

Note. The Other recidivism event category includes: driving under the influence, public order offenses, and criminal traffic offenses (e.g., driving with a suspended license).

Across the three study groups (comparison = 153, control = 30, treatment = 39), participants showed similar profiles on COMPAS risk and needs assessments. Average general recidivism risk scores ranged from 2.41 to 2.60 (SD ≈ 0.5–0.6), while violent recidivism risk scores were slightly lower (1.92–2.10; SD ≈ 0.7–0.8). Substance abuse needs were rated highest across domains (2.71–2.82; SD ≈ 0.4–0.6), whereas social isolation needs were lowest (1.67–1.85; SD ≈ 0.8–0.9). New recidivism events within 12 months were uncommon, occurring in 6% of the comparison group, 4% of the control group, and 3% of the treatment group. Importantly, no violent recidivism was observed, with drug and property offenses comprising most events in the comparison and control groups, and all events in the treatment group categorized as "other."

Analytical Method

To address research question I3 (see Exhibit 4), we will use a logistic regression model predicting the presence of a recidivism event in the first twelve months after a SHAR residents release from the facility. We will conduct two group comparisons:

1) Treatment to Control group

2) Treatment to Comparison group

In the first contrast, both groups are randomly assigned, so we will simply include an indicator for group status (Treatment/Control) in the logistic regression model and a dichotomous indicator for if the individual joined the study before or after the COVID-19 pandemic began (i.e., pre-2020 = 0, post-2020 = 1). However, the contrast of the Treatment group to the Comparison group does not benefit from random assignment to both groups, as the Comparison group was selected separately from our randomization process. As such, we used an inverse probability of treatment weighting (IPTW) model to minimize pre-existing baseline

differences between the Treatment and Comparison groups. We provide post-weighting summary statistics in Exhibit 7. As is evident from these statistics, the Comparison and Treatment samples are nearly identical with respect to most characteristics, with a minor exception being subject race, though the standardized differences across racial categories is below 0.1, indicating strong balance across groups.

Exhibit 7. Weighted Balance Table

	Comparison	Control	Treatment
N	153	30	39
Age (Mean (SD))	37.15 (8.46)	35.90 (8.46)	36.93 (8.84)
Sex (%)			
Male	62.9	60.0	64.1
Female	36.9	40.0	35.9
Unknown	0.2	0.0	0.0
Race (%)			
White	78.3	86.7	74.4
Black	15.0	6.7	17.9
Other	4.7	6.7	5.1
Unknown	2.0	0.0	2.6
General recidivism risk	2.45 (0.59)	2.60 (0.50)	2.46 (0.64)
Violent recidivism risk	2.10 (0.75	2.10 (0.80)	2.08 (0.77)
Substance abuse needs	2.79 (0.45)	2.77 (0.43)	2.82 (0.51)
Social isolation needs	1.85 (0.84)	1.67 (0.76)	1.85 (0.90)

When IPTW models include relevant indicators that account for baseline differences between groups that do and do not receive treatment, they are able to substantially minimize selection bias in treatment effect estimates (Austin & Stuart, 2015). In the current context, however, we are limited to just seven indicators for our IPTW models: demographic characteristics (age, sex, and race) and the four COMPAS subscales (general/violent recidivism risk, social isolation, and substance abuse). This limits our ability to minimize selection bias to the extent that our observables are able to account for meaningful differences between groups

with respect to the probability that they volunteer for the FGC and engage in recidivism during the follow-up period. We return to this limitation in our discussion.

Impact Study: Moderation Analysis

In addition to examining direct effects, we extend our impact analysis by exploring whether the effectiveness of FGC varies across participant subgroups. In other words, we test moderation to determine *for whom* and *under what conditions* FGC is most effective. These analyses allow us to link theory with practice: if restorative conferencing functions through identity transformation and strengthened social bonds, then these intermediate outcomes should partially account for observed improvements in recidivism outcomes. By situating moderation and mediation alongside the direct impacts, we provide a more complete understanding of the processes through which restorative justice supports desistance and reentry success.

Moderation Sample

To conduct our moderation analysis, we need to use individuals from the Treatment and Control groups that have valid MDOC data on recidivism and data for both the initial and pre-release surveys. In the MDOC data, we began with 47 and 34 SHAR residents in the treatment and control groups, respectively. However, 30 participants (17 control, 13 treatment) did not provide data for both the initial and pre-release surveys. As such, our sample for the moderation analysis includes 34 treated and 17 control observations.

Moderation Analytical Method

To address research question I4 (see Exhibit 4), we estimated a series of moderation models predicting changes in four psychosocial constructs central to desistance and a binary indicator for whether the individual experienced a recidivism event during their 12-month

follow-up window. All moderation models include an indicator of perceived family support at the time of the initial survey as a moderating variable.

Perceived family support was operationalized as an additive index constructed from six items measured at the initial survey: (1) "How much do you think your family really cares about you?" (2) "How well do you think your family understands your thoughts and feelings?" (3) "How much does your family appreciate you?" (4) "How much can you rely on your family for help if you have a serious problem?" (5) "How much can you talk to your family about your worries?" and (6) "How much can you relax and be yourself around your family?" Response options ranged from "Very little" (1) to "Very much" (5). We combined these items into a composite scale (higher scores = stronger perceived family support), which was then mean-centered for inclusion in interaction models.

Cost Study

The cost study will address one research question: (C1) What are the comparative costs and benefits for each impact? We will address this question through a cost–benefit analysis that examines the dollar value of resources invested in the intervention compared with the dollar value of benefits. We will calculate the per-person cost of delivering the intervention through a resource cost model—a "bottom-up" method for developing cost and expenditure estimates for education, founded on the "ingredients approach" to cost-effectiveness (Levin & McEwan, 2001). This approach involves identifying the types and quantities of specific program resources and calculating their associated costs. We will calculate monetized benefits by referencing estimated effects of FGC on criminal justice system outcomes based on published estimates of the averted costs to the justice system attributed to prevention of recidivism (Hunt et al, 2017, Hunt et al., 2018, Rouhani et al., 2022).

Cost Study: Calculating Costs

For the cost–benefit analysis, we used published studies to generate credible estimates of justice system costs that can be paired with program impacts on recidivism. Hunt et al. (2017) provided national and state-level estimates of the judicial and legal costs of crime to taxpayers, including courts, prosecution, and public defense. Complementing this, Hunt et al. (2019) estimated equivalent national and state-level costs for law enforcement responses to reported crimes by type. Finally, we consulted Rouhani et al. (2023) to generate estimates for legal system costs associated with drug possession cases. Their analysis of court data from Baltimore found that the cost of prosecuting a drug possession case, including law enforcement and court staff costs, ranged from $1,642 to $9,554 (in 2019 dollars) per case.

We used the Costs of Responding to Crime[2] tool published by RAND (n.d.) and based on the Hunt et al studies (2017; 2019) to obtain national estimates of court and law enforcement costs (in 2010 dollars) for each of the following recidivism events we observed in our sample: aggravated assault ($11,635), burglary ($1,868), and multiple types of property crimes ($1,727). There were five other miscellaneous crime types in our data that we were unable to match to a validated cost in the extant literature or tool so we chose to instead use the lowest estimate available from the RAND tool for other property crimes ($1,727).[3] From Rouhani et al. (2023) we take only the processing costs associated with a drug possession case ($1,642 in 2019 dollars) which excludes incarceration costs to make it comparable to the estimates provided by Hunt et al. (2017; 2019). We then transform each estimate into 2024 dollars by calculating an inflation

[2] https://www.rand.org/pubs/tools/TLA517-1/tool.html
[3] This includes: driving under the influence of drugs/alcohol, family/custody related offenses, obstruction of justice/resisting arrest, minor criminal traffic offenses (e.g., driving with a suspended license), and unauthorized use of a vehicle.

factor based on the ratio of consumer price index (CPI) in 2024 compared to the year the estimates were originally created (2010 or 2019), then multiplying the original dollar amounts by this factor. Exhibit 8 provides a list of the original estimates from each source and their 2024 dollar amounts by crime type in our data.

Exhibit 8. Estimated Recidivism Event Costs

Offense Type	Original Cost (Year)	Inflation Factor	Estimated 2024 Cost
Aggravated Assault	$11,635 (2010)	1.43	$16,789.31
Burglary	$1,868 (2010)	1.43	$2,695.52
Other Property	$1,727 (2010)	1.43	$2,492,06
Drug Possession	$1,642 (2019)	1.23	$2,019.66

With respect to the costs of implementing FGC, we collected information about the amount of time the facilitator spent prepping for each conference, the amount of time each conference lasted, and the number of participants, including the individual from our treatment sample. We then calculated averages for each of these indicators, yielding an average of 2.96 (SD = 1.06) participants per conference, an average preparation time of 3.92 hours (SD = 1.10) per conference, and an average of 1.58 hours (SD = 0.50) length of each conference.

Given that nearly all of our conferences were held virtually, there were no costs for a conference room, travel, or for printing out materials. We also estimate the wage costs, inclusive of benefits, for the facilitator ($63.46 per hour) and other participants ($44.67 per hour) based upon average total employer compensation costs per hour data provided by the U.S. Bureau of Labor Statistics (2025) for state/local government and private industry employers.

Unfortunately, we were unable to acquire actual ranges of wage rates for potential MDOC employees who would facilitate conferences in place of AIR staff. Further, we do not

have available estimates for costs such as travel, meeting space, or other material costs associated with FGC should conferences be held in-person as opposed to virtually. We are also unable to estimate costs for certain types of crime events present in our data as there were no clear estimates readily available for each of the recidivism event types we observed in our sample. As such, our estimates should be interpreted with due caution, as they are likely to underestimate the actual costs of implementing FGC in practice. We revisit these limitations in our discussion.

Results

Process Study Results

P1: How do FGC participants experience and perceive the FGC as delivered by SRRD staff?

Perceptions During the FGC

In Exhibit 9 we report treatment sample means, medians, and standard deviations for thirteen questions about their perceptions of FGC and the discussions surrounding their offending during the conference. All response options ranged from 1 "Strongly Disagree" to 5 "Strongly Agree."

Emotional Responses. Participants' emotional responses show that the FGCs evoked meaningful feelings of accountability, but the strength of these emotions varied. Participants tended to agree that they felt ashamed (mean = 3.98; median = 4.0) and guilty (mean = 3.42; median = 4.0), suggesting that the process did encourage personal reflection and acknowledgment of wrongdoing. They also leaned toward agreement that their offense was wrong (mean = 3.57; median = 4.0), indicating moral recognition of their actions.

As one participant reflected, "*I see I'm not only hurting myself but the people I love.*" Another explained, "*I was more excited and relieved about the fact that I was able to have this family conference*," underscoring both accountability and a sense of relief in addressing their behavior.

In comparison, feelings of forgiveness were more muted. With a mean of 3.16 and median of 3.00, participants were neutral overall about whether they felt forgiven. This suggests that while the FGC successfully encouraged accountability, it may have been less effective in helping participants feel fully accepted or absolved by others in the process.

Exhibit 9. Perceptions During the FGC (n=46)

During The Conference:	Mean	Median	Std. Dev
I felt ashamed of myself	3.98	4.00	1.14
I felt forgiven for what I did	3.16	3.00	0.88
I felt guilty for what I did	3.42	4.00	0.72
I felt that my offense was wrong	3.57	4.00	1.59
I learned that there are people who care about me	1.69	2.00	0.47
I was treated like I was a bad person	1.61	1.00	0.91
I was treated like I would probably commit another offense	1.58	1.00	0.81
I was treated with respect	3.61	4.00	0.61
Others indicated that I had learned my lesson and deserved a second chance	3.22	3.00	0.84
People made negative judgments about the kind of person I am	1.70	1.50	0.87
People said that it was not like me to do something wrong	3.22	3.00	1.17
People talked about aspects of myself that they liked	3.36	3.00	0.68
People were polite to me	3.54	4.00	0.72

Care and Respect. The lowest scores centered on perceptions of care. The statement "I learned that there are people who care about me" had a very low mean (1.69), reflecting clear disagreement with the idea that the conference communicated care or support. This represents a

critical gap, as restorative justice aims not only to hold participants accountable but also to reinforce community connectedness.

By contrast, perceptions of respect and politeness were notably stronger. Participants tended to agree they were treated with respect (mean = 3.61; median = 4.0) and politeness (mean = 3.54; median = 4.0). Importantly, they generally disagreed that they were treated like a bad person (mean = 1.61) or as if they would likely reoffend (mean = 1.58). Similarly, they disagreed that people made negative judgments about them (mean = 1.70). Taken together, these findings suggest that while participants did not feel actively cared for, they also did not feel stigmatized or demeaned during the process.

Support and Second Chances. Responses to items about encouragement and recognition of positive traits were generally neutral. Participants neither agreed nor disagreed that others said they had learned their lesson (mean = 3.22), that the offense was out of character (mean = 3.22), or that others spoke positively about their qualities (mean = 3.36). These neutral responses suggest that some participants may have experienced supportive messages, but these were not consistent or strong across the group.

Still, open-ended responses reveal moments where participants did perceive encouragement. One described feeling "*way more positive*" after the FGC, while another shared, "*I am optimistic about my relationship with my daughter.*" Participants also expressed hope about how the conference could support their transition back to the community. One participant explained:

> "*I think it will help me with my transition coming back to the community. Being able to see how I hurt and harmed them, as well as explaining why I did the things I did. Being able to realize how my addiction affected them with my*

actions and consequences. Also be given an opportunity to make amends to my family."

Another echoed this theme of reconciliation and forward-looking change: "*It allowed me to get to the root of my problem and addiction by sharing my past experiences with my family while being able to make amends.*"

Perceptions of the FGC Process

In Exhibit 10 we report treatment sample means, medians, and standard deviations for eight questions about their perceptions of the FGC process. All response options ranged from 1 "Strongly Disagree" to 5 "Strongly Agree."

Exhibit 10. Perceptions of the FGC Process (n=46)

The Conference Process:	Mean	Median	Std. Dev.
Added a needed component to the treatment I received at SHAR	3.59	4.00	0.69
Allowed me to express my views	1.71	2.00	0.458
Respected my rights	1.72	2.00	0.46
Took account of what I had to say in deciding what should be done	2.64	3.00	0.53
Was fair to me	1.69	3.00	0.51
Will help my transition back to the community after leaving SHAR	2.70	3.00	0.51
Will keep me from breaking the law in the future	3.54	4.00	0.69
Will keep me from relapsing in the future	2.59	3.00	0.62

*Fairness, R*espect, and Voice. Participants reported low levels of agreement that the FGC process was fair or respectful or that it allowed them to share their views. Items such as "Allowed me to express my views" (mean = 1.71; median = 2.0), "Respected my rights" (mean = 1.72; median = 2.0), and "Was fair to me" (mean = 1.69; median = 3.0) scored well below the neutral midpoint

of the scale. This suggests that participants actively disagreed with the notion that they were treated fairly or given a meaningful voice in the process.

Treatment and Transition. When considering whether the FGC contributed positively to their treatment and reentry, participants were closer to neutral. The statement "Took account of what I had to say in deciding what should be done" (mean = 2.64; median = 3.0) suggests that, on average, participants neither agreed nor disagreed that their input was meaningfully incorporated, indicating mixed or uncertain experiences. Similarly, views on whether the FGC "will help my transition back to the community" (mean = 2.70; median = 3.0) or "will keep me from relapsing" (mean = 2.59; median = 3.0) hovered near the neutral point. These findings suggest that participants were ambivalent about the long-term usefulness of the FGC for their recovery and reentry, rather than clearly seeing it as supportive.

By contrast, participants expressed stronger agreement that the FGC "added a needed component to the treatment I received at SHAR" (mean = 3.59; median = 4.0), suggesting that while they were unsure of its impact on future outcomes, they did see it as an immediate and worthwhile addition to the treatment program.

Several participants, however, described positive shifts. One reflected, "*It has opened my eyes that this is really about me and that I needed to focus on my wants and needs that are important to my recovery.*" Another emphasized, "*I have lied to myself and my loved ones for so long and this conference was the perfect ice breaker that I needed. My family and I were able to identify so many key factors about myself and my behavior that will formulate the plan I need for change.*" Such comments suggest that for some individuals, the FGC was a meaningful step in treatment and planning for recovery, even if survey responses revealed more mixed or neutral perceptions.

Future Behavior Change. Participants were more optimistic about the FGC's role in deterring future crime. The statement "Will keep me from breaking the law in the future" (mean = 3.54; median = 4.0) was the second most strongly endorsed item in this section, falling above neutrality and into the range of agreement. This suggests that, despite dissatisfaction with fairness and voice, many participants nonetheless believed the process had the potential to influence their future law-abiding behavior.

Some open-ended responses echoed this sense of optimism. One participant stated, "*I think it will help us stay on the right path and work as a team with the one I love,*" while another added, "*It helps me because I finally made my amends that needed to be made even if he doesn't fully understand.*" These reflections highlight how participants linked the process to strengthened relationships and renewed commitments to lawful and prosocial behavior.

Perceptions about the Conference Facilitator

In Exhibit 1 we report treatment sample means, medians, and standard deviations for eight questions about their perceptions about the FGC facilitator. All response options ranged from 1 "Strongly Disagree" to 5 "Strongly Agree."

Support and Communication. Participants' views of the facilitator's ability to create a supportive environment were mixed. Ratings for "Allow enough time to say everything you needed to say" (mean = 2.68; median = 3.0) and "Create an environment where you felt safe and comfortable talking" (mean = 2.72; median = 3.0) fell near neutrality. This suggests that participants were divided or uncertain about whether the facilitator fostered open dialogue, with some feeling adequately supported while others did not.

Similarly, participants were neutral on whether the facilitator "did a good job preparing [them] for the conference" (mean = 2.74; median = 3.0) and "did a good job preventing attacking

language and lecturing" (mean = 2.57; median = 3.0). These results imply that facilitators may have met minimum expectations but did not consistently provide a strong sense of readiness or actively shape the tone of the discussions.

Several participants described how the process helped them feel more open. One explained, "*My sister and I can talk about things without arguing,*" while another noted, "*It shows me it is not too late to know my mom.*" These remarks underscore how facilitation sometimes enabled constructive communication, though not consistently across participants.

Exhibit 11: Perceptions of the FGC Facilitator (n=46)

Did the Facilitator of Your Conference:	Mean	Median	Std. Dev.
Allow enough time to say everything you needed to say?	2.68	3.00	0.56
Create an environment where you felt safe and comfortable talking?	2.72	3.00	0.50
Do a good job listening during your conference?	1.78	2.00	0.42
Do a good job preparing you for the conference?	2.74	3.00	0.54
Do a good job preventing attacking language and lecturing?	2.57	3.00	0.58
Interject their own thoughts or opinions about what should happen?	3.78	4.00	1.21
Seem genuinely interested in your own needs?	2.64	3.00	0.53

Active Listening and Interest in Needs. The lowest score in this section concerned listening. Participants generally disagreed that the facilitator "did a good job listening during [the] conference" (mean = 1.78; median = 2.0). This finding is particularly significant, as active listening is a core facilitator responsibility in restorative justice practice. Without strong perceptions of being listened to, participants may have felt that their voices were minimized or undervalued. Likewise, participants were neutral on whether the facilitator "seemed genuinely interested in [their] needs" (mean = 2.64; median = 3.0), suggesting ambivalence about the extent to which facilitators recognized and centered their perspectives.

Neutrality and Interjection of Opinions. One notable outlier was the item "Interject their own thoughts or opinions about what should happen" (mean = 3.78; median = 4.0). Unlike other items, this was rated closer to agreement, indicating that participants perceived facilitators as actively inserting their own views into the process. While facilitators sometimes provide structure and guidance, restorative justice models emphasize participant ownership of the outcomes. Thus, this finding raises a concern that facilitators may have inadvertently shifted the balance away from participant-driven dialogue.

For some, this dynamic was experienced as a chance to reset and commit to change. One participant reflected, "*I have something to live up to now; an agreement has been made, it seems.*" Another added, "*Help me think before I act.*" These remarks reflect a willingness to embrace change, though they also raise questions about whether change was participant-driven or shaped more heavily by facilitator direction.

Summary

Taken together, participants' experiences of the FGCs as delivered by SRRD staff were mixed and uneven. The conferences were moderately successful in fostering accountability, respect, and rehabilitation, as participants reported shame, guilt, recognition of wrongdoing, and a belief that the process could help prevent future offending. They also felt they were generally treated politely and without stigma.

However, participants were less likely to feel cared for, forgiven, or included in decision-making. Ratings near neutrality on many process and facilitator items indicate ambivalence, while low scores on fairness, respect for rights, and facilitator listening reveal shortcomings in

facilitation quality. The perception that facilitators interjected their own opinions compounds these concerns, suggesting a shift away from participant-led dialogue.

For SRRD staff, the findings underscore a dual imperative: build on the strengths of accountability and rehabilitation while strengthening the relational and procedural dimensions of FGC. This includes fostering genuine listening, increasing participant voice, demonstrating care, and ensuring that facilitators guide without dominating. By enhancing these dimensions, SRRD can better align FGC delivery with restorative justice principles and improve both participant experiences and long-term outcomes.

Impact Study – Survey Results

In the final sample for our survey analysis, we had 38 participants in the treatment group and 20 in the control group who took both the initial and pre-release survey. Our research questions I1 and I2, look at the difference between the pre and post survey data for the treatment and control groups, and ask did the treatment group show a greater improvement on scores empathy/remorse, social bonds/supports, generativity or belief in redeemability compared to the control group? And, did the treatment group show a greater decrease in the criminal identity score compared to the control group?

I1: Does participating in FGC result in significant increases in belief in redeemability, empathy, remorse, and accountability, and generativity as compared to individuals who did not participate in FGC?

Participants of the FGC did not show a significantly greater increase in belief in redeemability, empathy, remorse, and accountability, or generativity compared to the control group. Despite the small sample size, one score shows promising results. The treatment group had a near significant

increase on their standardized generativity scores. Participants in the treatment group showed a greater increase in generativity compared to the control group (Mtreated = 0.036, Mcontrol = −0.004, ΔM = 0.040, t(37.8) = 1.63, p = .056).

I2: Does participating in FGC result in significant reductions in criminal identity/core self compared to individuals who did not participate in FGC?

Both groups reported decreases in criminal identity from pre to post, with no significant difference in the change in means at p<.05 (Mtreated = 0.346, Mcontrol = −0.400, ΔM = 0.054, t(40.7) = .386, p = .649). For both groups, the decrease in criminal identity between pre and post was significant with a mean difference of –0.40(t(18) = -3.70, p = 0.0008) for the control, and –0.35(t(37) = -3.97, p = 0.0002) for the treatment group. This suggests that for those who volunteered for the FGC program (both treatment and control) time from release may impact criminal identity/core self more than participation in the FGC program itself.

Impact Study – Recidivism Results

To answer research question I3, we compared the likelihood of 12-month recidivism for the comparison group (n=153), control group (n=30) and treatment group (n=39). To adjust for baseline differences across the treatment and comparison groups, we used a propensity score weighting approach including the following covariates: age, age squared, sex, race, general recidivism risk, violent recidivism risk, substance abuse, and social isolation. Both full-sample and average treatment effect on the treated (ATT) weighting methods were implemented. The final unweighted sample was predominantly male (60-68%) and White (74–87%), with average ages ranging from 35.9 (Control) to 37.8 (Comparison). Risk scores for recidivism, substance

abuse, and social isolation were relatively consistent across groups, with no statistically significant differences observed in unweighted comparisons.

I3: Are FGC participants less likely to have a recidivism event compared to individuals who did not participate in FGC?

We used logistic regression models to estimate the effect of treatment and control group status on new recidivism events during the post-release 12-month follow-up window. We report estimates from the following comparisons in Exhibit 12: (1) Treatment vs. Control (unweighted), (2) Treatment vs. Comparison (weighted), and (3) Control vs. Comparison (weighted). Across all models, the treatment and control indicators were not statistically significant predictors of recidivism. For example, in the unweighted Treatment vs. Control model, the coefficient for treatment was –0.41 (p = .781), and in the weighted Treatment vs. Comparison model, the coefficient was –1.36 (p = .271). Similarly, the Control vs. Comparison model yielded a non-significant coefficient of -0.99 (p = .381).

Exhibit 12. Logistic Regression Results Predicting 12-Month Recidivism

Predictor	Treatment	p value	Treatment vs.	p	Control vs.	p value
(Intercept)	–20.37	.995	–1.45	.053	–2.47	<.001
treat	–0.41	.781	–1.36	.271	—	—
control	—	—	—	—	–0.99	.381
Year Post 2020	17.73	.996	–1.87	.032	0.28	.734

Comparison	Pseudo R^2	Residual Deviance	Null Deviance	BIC	Sample Size
Treatment vs Control	0.13	15.484	17.800	27.96	64
Treatment vs Comparison	0.121	93.207	106.013	—	180
Control vs Comparison	0.014	96.414	97.820	—	172

Model fit statistics varied across specifications. The unweighted Treatment vs. Control model showed a pseudo R^2 of 0.13, while the weighted Treatment vs. Comparison and Control vs.

Comparison models yielded pseudo R^2 values of 0.121 and 0.014, respectively. These results suggest modest explanatory power and underscore the importance of cautious interpretation, particularly given the small sample sizes in the treatment (n = 39) and control (n = 30) groups and the very low prevalence of recidivism events within these samples (n=2 events).

Impact Study – Moderation Results

I4: Do moderators account for significant variation in treatment effects for each of the relationships tested in questions I1–I3?

Belief in Redeemability

Exhibit 13. OLS Results Predicting Change in Belief in Redeemability Factor Score (n=51)

Predictor	Coefficient (SE)	Test Statistic	p value
(Intercept)	0.08 (0.13)	0.62	0.54
Treatment	-0.09 (0.16)	0.57	0.57
Social Bonds	-0.15 (0.16)	-0.96	0.34
Treatment X Social Bonds	0.10 (0.20)	0.54	0.60
Model Diagnostics			
F-Statistic (p-value)	0.58 (0.63)		
R^2	0.04		
Adjusted R^2	-0.03		

In the OLS model predicting change in belief in redeemability, neither the main effect of treatment at average baseline social bonds ($\beta = -0.09$, SE = 0.16, p = .57) nor the main effect of social bonds among controls ($\beta = -0.15$, SE = 0.16, p = .34) was statistically significant. The Treatment × Social Bonds interaction, which tests moderation, was also non-significant (β = 0.10, SE = 0.20, p = .60), indicating no evidence that baseline family social bonds condition the treatment effect within the observed range. Further, model fit was weak (F = 0.58, p = .63; R^2 = .04; adjusted $R^2 = -.03$). In summary, for belief in redeemability, we do not find support that baseline social bonds significantly moderate the treatment effect.

Empathy, Remorse, and Accountability

Exhibit 14. OLS Results Predicting Change in Empathy/Remorse/Accountability Factor Score (n=51)

Predictor	Coefficient (SE)	Test Statistic	p value
(Intercept)	0.09 (0.05)	1.86	0.07
Treatment	0.01 (0.06)	0.16	0.88
Social Bonds	0.02 (0.06)	0.29	0.77
Treatment X Social Bonds	-0.02 (0.07)	-0.25	0.80
Model Diagnostics			
F-Statistic (p-value)	0.04 (0.99)		
R^2	<0.01		
Adjusted R^2	-0.06		

In the OLS model for change in empathy, remorse, and accountability, the estimated treatment effect was near zero and not statistically distinguishable from the control mean (β = 0.01, SE = 0.06, p = .88). Additionally, social bonds measured at the initial survey were not associated with change among controls (β = 0.02, SE = 0.06, p = .77). Further, the interaction effect was small and imprecise (β = −0.02, SE = 0.07, p = .80), providing no indication that the treatment effect depends on family support. Model diagnostics indicate minimal explanatory power and essentially no overall fit (F = 0.04, p = .99; R^2 < .01; adjusted R^2 = −.06). Overall, we observe no evidence that social bonds moderate the effect of FGC on change in empathy, remorse, and accountability.

Generativity

In the OLS model of change in generativity, there is evidence of a modest treatment-associated gain. Evaluated at the mean of the centered family support index, the treatment group improved more than controls (β = 0.05, SE = 0.03, p = <.05). Social bonds at the time of the initial survey were not detectably related to change among controls (β = −0.02, SE = 0.02, p =

.37). The Treatment × Social Bonds interaction was positive but not significant (β = 0.03, SE = 0.03, p = .32), indicating no support for variation in the treatment effect across levels of family support.

Exhibit 15. OLS Results Predicting Change in Generativity Factor Score (n=51)

Predictor	Coefficient (SE)	Test Statistic	p value
(Intercept)	-0.02 (0.02)	-0.83	0.41
Treatment	0.05 (0.03)*	2.05	<0.05
Social Bonds	-0.02 (0.02)	-0.91	0.37
Treatment X Social Bonds	0.03 (0.03)	0.99	0.32
Model Diagnostics			
F-Statistic (p-value)	1.69 (0.18)		
R^2	0.10		
Adjusted R^2	0.04		

The constant suggests negligible change for controls at the centered moderator (−0.02 SD, p = .41). Model diagnostics point to limited explanatory power and a non-significant F-test (F = 1.69, p = .18; R^2 = .10; adjusted R^2 = .04). Overall, the pattern of results is consistent with a small average treatment-related increase in generativity that does not appear to depend on family bonds.

Criminal Identity/Core Self

In the OLS model for change in criminal identity/core self, the constant indicates a significant decrease in the factor score for the control group at the mean of the centered social bonds index (−0.42 SD, SE = 0.13, p < .01). Meanwhile, the estimated effect of FGC for the treatment group at that point was effectively zero and not significant (β = 0.01, SE = 0.16, p = .95).

Exhibit 16. OLS Results Predicting Change in Criminal Identity/Core Self Factor Score (n=51)

Predictor	Coefficient (SE)	Test Statistic	p value
(Intercept)	-0.42 (0.13)**	-3.21	<0.01
Social Bonds	-0.11 (0.15)	-0.73	0.47
Model Diagnostics			
R^2	0.05		

Among controls, social bonds at the time of the initial survey were not related to change (β = −0.11, SE = 0.15, p = .47). The interaction effect was positive but not statistically significant (β = 0.27, SE = 0.19, p = .16), providing little indication that the treatment effect varies by family support. Finally, model diagnostics show weak fit (F = 0.83, p = .48; R^2 = .05; adjusted R^2 = −.01). Overall, there is no evidence of moderation by social bonds for this outcome, and the average treatment–control difference at the centered moderator is indistinguishable from zero.

Recidivism Events

Exhibit 17. Logistic Regression Results Predicting Recidivism (n=51)

Predictor	Coefficient (SE)	Test Statistic	p value
(Intercept)	-3.23 (1.59)*	-2.03	0.04
Treatment	-0.41 (2.05)	-0.20	0.84
Social Bonds	-1.04 (1.33)	-0.78	0.44
Treatment X Social Bonds	1.65 (1.97)	0.84	0.40
Model Diagnostics			
Log likelihood	-7.64		
Pseudo R^2 (McFadden's)	0.07		
AIC	23.27		

In the logistic regression predicting a recidivism event within 12 months (n = 51), the treatment effect evaluated at the mean of the centered social bonds indicator was not statistically

significant (β = −0.41, SE = 2.05, p = .84). Additionally, reported social bonds at the time of the initial survey were not associated with the likelihood of recidivism among controls (β = −1.04, SE = 1.33, p = .44). The interaction effect was positive but insignificant (β = 1.65, SE = 1.97, p = .40), providing no evidence that the treatment effect varies across levels of social bonds in this sample. The intercept implies a predicted probability of approximately 3.9% for a recidivism event for controls at the centered moderator (β = −3.23, SE = 1.59, p = .04). Finally, model fit statistics indicate limited explanatory power (McFadden's R^2 = 0.07; log likelihood = −7.64; AIC = 23.27). Overall, neither a main treatment effect nor moderation by social bonds was detected for recidivism.

Cost Study Results

C1: Do the benefits of implementing FGC outweigh its costs?

To evaluate the benefits of FGC, we used a total sample of 72 FGC volunteers with full 12-month follow-up windows (n=34 control, n=38 treatment). We calculated the average costs of recidivism to be the summed cost of each recidivism event within each group divided by the total group sizes ($\overline{C}_C$ [control] = \$80.39, $\overline{C}_T$ [treatment] = \$69.22). We then estimated the average costs of implementing the FGC per individual in the treatment group (c_{avg} = \$557.94) and calculated lower (c_{lb} = \$167.72) and upper bound (c_{ub} = \$1137.56) costs for this estimate based upon twice the value of the standard deviations for estimates of the amount of time facilitators spent prepping the conference (mean = 3.92, sd = 1.10), the number of non-facilitator conference participants (mean = 2.96, sd = 1.06), and the length of the conferences (mean = 1.58, sd = 0.50). Using these estimates, we then calculated the benefit per person with respect to crime costs averted, the net present value of FGC per person (NPV), the benefit-cost ratio (BCR), and the

return on investment (ROI). Equations and values for each of these indicators are provided in Exhibit 18:

Exhibit 18. Person-level Costs and Benefits of FGC

Indicator	Equation	Values	Result
Benefit per-person	$B = \overline{C}_C - \overline{C}_T$	$\overline{C}_C$ = \$80.39; $\overline{C}_T$ = \$69.22	B = \$11.17
NPV	NPV = $B - c$	c_{lb} = \$167.72 c_{avg} = \$557.94 c_{ub} = \$1137.56	NPV_{lb} = \$(156.55)NPV_{avg} = \$(546.77)NPV_{ub} = \$(1126.39)
Benefit-Cost Ratio	$BCR = \frac{B}{c}$	-----	BCR_{lb} = 0.07 BCR_{avg} = 0.02 BCR_{ub} = 0.01
ROI	$ROI = \frac{B - c}{c}$	-----	ROI_{lb} = -0.93 ROI_{avg} = -0.98 ROI_{ub} = -0.99

When intervention costs were compared to benefits, the net present value (NPV) was consistently negative, ranging from –\$156.55 under the lowest cost scenario to –\$1,126.39 under the highest. Correspondingly, benefit–cost ratios (BCR) were far below 1 (0.07 to 0.01), and return on investment (ROI) values were strongly negative (–0.93 to –0.99). These results indicate that while the intervention generated some reduction in recidivism costs, the costs of program delivery substantially outweighed the economic benefits of crimes averted within the one-year follow-up window.

It is important to note, however, that our analysis lacked information about several potentially relevant outcomes, including the costs of crime to victims and any financial or social benefits of FGC participation, such as improved employment or residential stability. Moreover, the sample size was relatively small, and the number of recidivism events observed within the treatment (n = 1) and control (n = 2) groups was very low. As a result, our cost estimates rest on a limited set of observed events and should be interpreted with caution. While our calculations consistently indicated that the per-person costs of delivering FGC exceeded the modest economic

benefits from reduced recidivism costs, these findings are best viewed as preliminary. The exclusion of broader victim costs and potential long-term participant benefits, combined with the small sample size, substantially tempers the conclusion that FGC is not cost-beneficial in this context.

Discussion

Overview

In this study, we conducted a mixed-methods evaluation of Family Group Conferencing (FGC) for justice-involved individuals returning to the community. We combined a process study of implementation and participant experiences, an impact study using a randomized controlled design, and a cost study to assess the comparative costs and benefits of the intervention. This integrated approach allowed us to examine not only whether FGC had measurable effects on participants, but also how the program was delivered and whether it produced value from a system-level perspective.

Summary of Findings

Process Study. The process study highlighted both strengths and challenges in how FGC was implemented. Many participants described the conferences as meaningful opportunities to confront their behavior, recognize the harm they had caused, and take accountability in front of their families. Reports of shame, guilt, and recognition of wrongdoing suggest that the conferences were capable of activating important mechanisms of change that align with restorative justice principles. Participants also consistently noted being treated with respect and without overt stigma, an important finding given concerns about how justice-involved individuals often experience institutional processes. At the same time, the survey and qualitative feedback revealed mixed results on dimensions central to relational quality. Ratings of fairness,

inclusion in decision-making, and facilitator listening were uneven, with some participants indicating that facilitators spoke too much or imposed their own views rather than fostering participant-led dialogue. These patterns suggest that while FGC can successfully encourage accountability and respect, the delivery of the model in this setting varied in quality, with relational aspects less reliably achieved.

Impact Study. The randomized controlled trial component was designed to test whether FGC influenced psychosocial outcomes and recidivism. Across measures of empathy, generativity, and personal identity, the analyses found limited evidence of positive change, with no significant differences between the treatment and control groups. Importantly, the 12-month follow-up window yielded very few recidivism events—only one in the treatment group and two in the control group. This lack of statistical power severely constrained our ability to detect differences even if they existed. While the null findings may suggest limited impact, they are more likely reflective of the small sample size and short observation period than of the ineffectiveness of the model itself. Future evaluations with larger samples and longer follow-up periods are necessary to determine whether FGC can influence reoffending behavior in a measurable way.

Cost Study. The cost analysis provided additional perspective on the value of FGC as a reentry strategy. We estimated the average cost of program delivery at approximately $558 per participant, with plausible bounds ranging from $168 to over $1,100 depending on assumptions about preparation time, conference participation, and conference duration. When compared to the modest $11 per-person reduction in recidivism costs between the treatment and control groups, the program costs substantially outweighed the benefits. Net present values were negative under

all assumptions, and benefit–cost ratios remained well below one. These findings suggest that from a narrow perspective focused only on justice-system expenditures within a short time frame, FGC was not cost-beneficial. However, this analysis does not capture potential long-term or non-monetary benefits such as improvements in family relationships, enhanced legitimacy of justice processes, or reduced stigma, all of which may justify investment in restorative practices even when immediate cost savings are not realized.

Study Limitations. Several limitations should be considered when interpreting the findings of this evaluation. First, the sample size was modest, and the number of recidivism events observed across treatment and control groups was very small. This severely limited statistical power and constrained our ability to detect meaningful differences in outcomes, even if they existed. Second, implementation quality varied, with participants reporting inconsistent relational experiences of fairness, inclusion, and facilitator listening, suggesting that fidelity to restorative principles may not have been uniform. Third, the cost study relied on available data about staff time and participation, and while we included plausible lower and upper bounds, estimates remain sensitive to assumptions about preparation, facilitation, and overhead. In addition, the economic analysis focused narrowly on justice-system costs within a 12-month window. It did not account for other potential benefits of FGC, such as improved family relationships, greater victim satisfaction, or long-term social and economic stability. Taken together, these limitations mean that our results should be interpreted with caution and viewed as preliminary evidence rather than definitive conclusions about the effectiveness or efficiency of FGC.

Implications for Research and Policy

Our results align with prior work emphasizing both the potential and the limits of restorative justice. While FGC fostered accountability and some dimensions of respect, it did not produce detectable reductions in recidivism or clear economic benefits within the study period. Future evaluations should broaden the range of outcomes examined to include victim satisfaction, family relationships, and long-term social and economic stability, which may better capture the benefits of restorative approaches. Larger samples and extended follow-up periods are also needed to more fully assess effects on reoffending.

With respect to policy, our findings underscore the importance of investing in facilitator training to strengthen fairness, listening, and participant voice during conferences. Policymakers should also be cautious about evaluating restorative programs solely through the lens of short-term justice-system costs, as doing so may overlook broader gains in community well-being, legitimacy, and reintegration. Continued research is needed to clarify under what conditions and for whom FGC may be most effective as a reentry strategy.

References

Angel, C. M. (2005). *Crime victims meet their offenders: Testing the impact of restorative justice conferences on victims' post-traumatic stress symptoms*. University of Pennsylvania. Unpublished dissertation.

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine, 34*(28): 3661–3679.

Bazemore, G. & Erbe, C. (2004). Reintegration and Restorative Justice: Towards a Theory and Practice of Informal Social Control and Support. In S. Maruna and R. Immarigeon (Eds.), *After Crime and Punishment: Pathways to Offender Reintegration* (pp. 27–56). Devon, UK and Portland, OR: Willan Publishing; Lehman et al. (2002) The Three "R's" of Reentry. Justice Solutions. Washington, DC.

Bazemore, G., & Maruna, S. (2009). Restorative justice in the reentry context: Building new theory and expanding the evidence base. *Victims & Offenders, 4*(4), 375–384; Bazemore, G., & Stinchcomb, J. (2004). Civic engagement model of reentry: Involving community through service and restorative justice. *Federal Probation, 68*(2), 14–24.

Bergseth, K. J., & Bouffard, J. A. (2007). The long-term impact of restorative justice programming for juvenile offenders. *Journal of Criminal Justice, 35*(4), 433–451.

Blomberg, T., Bales, W., Mann, K., Meldrum, R., & Nedelec, J. (2010). *Validation of the COMPAS risk assessment classification instrument*. College of Criminology and Criminal Justice, Tallahassee, FL: Florida State University.

Braithwaite, J., & Mugford, S. (1994). Conditions of successful reintegration ceremonies: Dealing with juvenile offenders. *British Journal of Criminology, 34*(2), 139–171; Maruna, S. (2011). Reentry as a rite of passage. *Punishment & Society, 13*(1), 3.

Braman, D., & Wood, J. (2003). *From one generation to the next: How criminal sanctions are reshaping family life in urban America.* In *Prisoners once removed: The impact of incarceration and reentry on children, families, and communities*, (pp. 157–188). Washington, DC: Urban Institute Press; Travis, J., Cincotta McBride, E., & Solomon, A. L. (2005). *Families left behind: The hidden costs of incarceration and reentry.* Washington, DC: Urban Institute.

Brennan, M. A., Barnett, R. V., & McGrath, B. (2009). The intersection of youth and community development in Ireland and Florida: Building stronger communities through youth civic engagement. *Community Development*, *40*(4): 331–345.

Bushway, S. D., & Apel, R. (2012). Signaling perspective on employment-based reentry programming: Training completion as a desistance signal. *Criminology and Public Policy, 11*(1), 21–50.

Bushway, S. D., & Paternoster, R. (2013). Desistance from crime: A review and ideas for moving forward. In C. L. Gibson, & M. D. Krohn (Eds.), *Handbook of life-course criminology: Emerging trends and directions for future research* (pp. 213–231. New York, NY: Springer Science + Business Media.

Daly, K. (2006). Restorative justice and sexual assault: An archival study of court and conference cases. *British Journal of Criminology, 46*(2), 334–356.

Duwe, G. (2012). Can circles of support and accountability (COSA) work in the United States? Preliminary results from a randomized experiment in Minnesota. *Sexual Abuse: A journal of Research and Treatment*, *25*(2): 143–165.

Fass, T. L., Heilbrun, K., DeMatteo, D., & Fretz, R. (2008). The LSI-R and the COMPAS: Validation data on two risk-needs tools. *Criminal Justice and Behavior, 35*(9): 1095–1108.

Franke, D., & Jarjoura, R. (2016). *Why reentry-based restorative justice signals desistance: Further evidence from a randomized trial*. Paper presented at the annual meetings of the American Society of Criminology, New Orleans.

Ganem, N. M., & Agnew, R. (2007). Parenthood and adult criminal offending: The importance of relationship quality. *Journal of Criminal Justice, 35*(6), 630–643; Naser, R. L., & Visher, C. A. (2006). Family members' experiences with incarceration and reentry. *Western Criminology Review*, 7, 20–31.

Hairston, C. F., & Oliver, W. (2006). *Safe return: Domestic violence and prisoner reentry: Experiences of black women and men*. New York, NY: Vera Institute of Justice; Travis, J., Cincotta McBride, E., & Solomon, A. L. (2005). *Families left behind: The hidden costs of incarceration and reentry.* Washington, DC: Urban Institute.

Hayes, H., & Daly, K. (2004). Conferencing and re-offending in Queensland. *Australian & New Zealand Journal of Criminology, 37*(2), 167–191.

Hunt, P., Anderson, J., & Saunders, J. (2017). The price of justice: New national and state-level estimates of the judicial and legal costs of crime to taxpayers. *American Journal of Criminal Justice, 42*(2): 231–254.

Hunt, P. E., Saunders, J., & Kilmer, B. (2019). Estimates of law enforcement costs by crime type for benefit-cost analyses. *Journal of Benefit-Cost Analysis, 10*(1): 95–123.

Kiecolt, K. J. (1994). Stress and the decision to change oneself: A theoretical model. *Social Psychology Quarterly, 57*(1), 49–63.

Kline, R. B. (2016). Principles and Practice of Structural Equation Modeling (4th ed). Guilford publications.

LeBel, T. P., Richie, M., & Maruna, S. (2014). Helping others as a response to reconcile a criminal past: The role of the wounded healer in prisoner reentry programs. *Criminal Justice and Behavior, 42*(1), 108-120.

Martinez, D. J. (2008). *Homeward bound: Examining social support among former prisoners and their family members.* The University of Chicago: ProQuest Dissertations Publishing, 3309077.

Maruna, S. (2011). Reentry as a rite of passage. *Punishment & Society, 13*(1), 3.

Maruna, S. (2012). Elements of successful desistance signaling. *Criminology & Public Policy, 11*(1), 73–86.

McAdams, D. P., & de St Aubin, E. (1998). 'Introduction.' In D. P. McAdams & E. de St Aubin (Eds.), *Generativity and adult development: How and why we care for the next generation*. Washington, DC: American Psychological Association.

McCold, P, & Wachtel, B. (1998). "Community is not a place: A new look at community justice initiatives." *Contemporary Justice Review* 1.1 (pp. 71–85).

McGarrell, E. F., & Hipple, N. K. (2007). Family group conferencing and re-offending among first-time juvenile offenders: The Indianapolis experiment. *Justice Quarterly, 24*, 2221–2246.

McNeill, F., & Maruna, S. (2007). Giving up and giving back: Desistance, generativity and social work with offenders. In G. McIvor & P. Raynor (Eds.), *Developments in work with offenders.* London, UK: Routledge.

Nakamura, K., & Bucklen, K. B. (2014). Recidivism, redemption, and desistance: Understanding continuity and change in criminal offending and implications for interventions. *Sociology Compass, 8*(4), 384–397.

Naser, R. L., & La Vigne, N. G. (2006). Family support in the prisoner reentry process expectations and realities. *Journal of Offender Rehabilitation, 43*(1), 93–106; Visher, C., Kachnowski, V., La Vigne, N., & Travis, J. (2004). *Baltimore prisoners' experiences returning home.* Washington, DC: Urban Institute.

Paternoster, R., Bachman, R., Kerrison, E., O'Connell, D., & Smith, L. (2016). Desistance from crime and identity: An empirical test with survival time. *Criminal Justice and Behavior, 43*(9): 1204–1224.

Pulos, S., Elison, J., & Lennon, R. (2004). The hierarchical structure of the interpersonal reactivity index. *Social Behavior and Personality, 32*(4): 355–360.

RAND Corporation. (n.d.). *Costs of Responding to Crime: Police, Court, and Legal Services*. Retrieved August 8, 2025, from https://www.rand.org/pubs/tools/TLA517-1/tool.html

Rojek, D. G., Coverdill, J. E., & Fors, S. W. (2003). Effect of victim impact panels on DUI rearrest rates: A 5-year follow-up. *Criminology, 41*(4), 1319–1340.

Rouhani, S., Tomko, C., Silberzahn, B., Sherman, S. G., & Cepeda, J. (2023, April). *Estimating economic costs of prosecuting simple drug possession in Baltimore City*. Johns Hopkins Bloomberg School of Public Health.

Sherman, L. W. & Strang, H. (2007). *Restorative justice: The evidence*, London, UK: Smith Institute.

Sherman, L. W., Strang, H., & Woods, D. J. (2000). *Recidivism patterns in the Canberra reintegrative shaming experiments (RISE)* (pp. 276-295). Canberra: Centre for Restorative Justice, Research School of Social Sciences, Australian National University.

Strang, H., Sherman, L., Mayo-Wilson, E., Woods, D. J., & Ariel, B. (2013). Restorative Justice Conferencing (RJC) using face-to-face meetings of offenders and victims: Effects on offender recidivism and victim satisfaction. *Campbell Systematic Reviews,* 9 doi: 10.4073/csr.2013.10.

Sullivan, M. J. L., Rodgers, W. M., Wilson, P. M., Bell, G. J., Murray, T. C., & Fraser, S. N. (2002). An experimental investigation of the relation between catastrophizing and

activity intolerance. *Pain, 100*(1–2), 47–53; Visher, C. A., & Courtney, S. M. E. (2006). *Cleveland prisoners' experiences returning home.* Washington, DC: Urban Institute.

U.S. Bureau of Labor Statistics. (2025, March 14). *Employer costs for employee compensation – December 2024* (USD L-25-0335). https://www.bls.gov/news.release/archives/ecec_03142025.pdf

Umbreit, M. S., & Coates, R. B. (1992). The impact of mediating victim offender conflict: An analysis of programs in three states. *Juvenile & Family Court Journal, 43*(1), 21–28.

Wilkins, L.T. (1991). *Punishment, Crime, and Market Forces*. Brookfield, VT: Dartmouth.

Appendix A. Constructs and Survey Items

Exhibit A1. List of Survey Questions by Construct

Construct (Subscale)	Question	Response Scale
Belief in Redeemability (Agency)	I am basically a criminal and probably always will be.	Strongly Agree (1) to Strongly Disagree (5)
	I am just a little piece in a big game.	
	I am just condemned to a life of crime.	
	I believe I could "go straight."	
	I do not have a lot of control over the future.	
	I do not really know where my life is going.	
	I think I can put my criminal past behind me.	
	My future is in my hands.	
Belief in Redeemability (Belonging)	I am a part of the same world as everyone else	Strongly Agree (1) to Strongly Disagree (5)
	I am a victim of society	
	I can be a positive member of society	
	I do not think I will ever really be valued by society	
	I think I could play my part by helping other people in my community	
	I think people will respect me one day for doing my part in the community	
Belief in Redeemability (Optimism)	I believe "going straight" is possible AND I want to do it	Strongly Agree (1) to Strongly Disagree (5)
	I can make it in a straight world	
	I could be happy going straight	
	I could be happy in my life without doing crime again	
	I just cannot see myself settling down and being satisfied	
	I will never be able to enjoy a straight job	
	Stuff has happened in my life that means I just cannot go back to living straight	
	The only thing that would make a difference is winning the lottery	
	To be honest, I do not have a plan for the long term	

Construct (Subscale)	Question	Response Scale
Criminal Identity	People think I deserve another chance	Strongly Agree (1) to Strongly Disagree (5)
	I believe that I have changed for the better	
	The people I care about believe that I have changed for the better	
	Being a criminal has little to do with how I feel about myself	
	Being a criminal is an important part of my self-image	
	The fact that I am a criminal rarely enters my mind	
	Being an addict has little to do with how I feel about myself	
	Being an addict is an important part of my self-image	
	The fact that I am an addict rarely enters my mind	
Empathy, Remorse, and Accountability	I understand how my actions have affected my family	Strongly Agree (1) to Strongly Disagree (5)
	My family is as much to blame for what happened as I am	
	There are relationships I damaged that need to be repaired	
	I sometimes find it difficult to see things from my family's point of view	Does not describe me at all (1) to Describes me very well (5)
	I sometimes try to understand my family better by imagining how things look from their perspective	
	When I'm upset with my family, I usually try to "put myself in their shoes" for a while	
	Before criticizing my family, I try to imagine how I would feel if I were in their place	
Generativity	I do not feel like other people need me	Statement never applies (1) to Statement always applies (4)
	I feel as though I have made a difference in the lives of others	
	I have made and created things that have had an impact on other people	
	In general, my actions do not have a positive effect on others	
	I have a responsibility to improve the neighborhood where I live	

Construct (Subscale)	Question	Response Scale
Social Bonds	How do you feel about the way things are: In general between you and your close relatives?	Terrible (1) to Delighted (7); (N/A – 8)
	How do you feel about the way things are: Between you and your children?	
	How do you feel about the way things are: Between you and your partner?	
	How much do you think your family really cares about you?	Very little (1) to Very much (5)
	How well do you think your family understands your thoughts and feelings?	
	How much does your family appreciate you?	
	How much can you rely on your family for help if you have a serious problem?	
	How much can you talk to your family about your worries?	
	How much can you relax and be yourself around your family?	

About the American Institutes for Research®

Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit institution that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of education, health, and the workforce. AIR's work is driven by its mission to generate and use rigorous evidence that contributes to a better, more equitable world. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit AIR.ORG.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

XXXXX_MO/23