

Protocol for a Campbell Collaboration Systematic Review:

Effects of Face-to-Face Restorative Justice  
for Personal Victim Crimes:

Submitted to the  
Campbell Crime and Justice Group  
Campbell Collaboration

*Supported by the*  
Smith Richardson Foundation “Test-Bed” Project

by

Heather Strang  
Centre for Restorative Justice  
Australian National University

Lawrence W. Sherman  
Jerry Lee Center of Criminology  
University of Pennsylvania

Submitted to the  
Campbell Crime and Justice Group  
Campbell Collaboration

And  
Smith Richardson Foundation

Draft # 4  
November 2004

## Cover Sheet

### Title

Effects of Face-to-Face Restorative Justice for Personal Victim Crimes.

### Reviewers

Heather Strang, Ph.D.  
Director, Centre for Restorative Justice  
Research School of Social Sciences  
Australian National University  
Canberra, ACT 0200  
Australia  
Tel: ++61 2 6125 3287  
Fax: ++ 61 2 125 4933  
Email: [heather.strang@anu.edu.au](mailto:heather.strang@anu.edu.au)

Lawrence W. Sherman, Ph.D.  
Director and Albert M. Greenfield Professor  
Jerry Lee Center of Criminology  
University of Pennsylvania  
3814 Walnut Street  
Philadelphia PA 19104-6197  
USA  
Tel: ++ 1 215 898 9216  
Fax: ++ 1 215 898 0864  
Email: [lws@sas.upenn.edu](mailto:lws@sas.upenn.edu)

### Sources of Support

Jerry Lee Center Foundation and the Smith Richardson Foundation.

## **Background**

Restorative justice is a recent name given to community practices for responding to crime that are several thousand years old (Braithwaite, 1998). These practices have continued in many nations, including China and several Middle East countries, but they diminished in the North Atlantic world during the 19<sup>th</sup> century with a new emphasis after the French Revolution on equity of treatment under the law ('nullum crimen sine lege'). But even in Europe a tradition of informal disposals such as East Germany's 'social courts' – bodies of ordinary citizens that dealt with minor offences and disputes by mediation and shaming but without imposing substantive punishment – continued in some countries (M. Killias, personal communication, November 2, 2004).

Since the 1980s, they have reappeared in a variety of forms, all designed to remedy various criticisms of the formal criminal justice system. These forms include victim-offender mediation, sentencing circles, family group conferences, and other names unique to certain localities or ethnicities. These practices function in similar ways, whether they are 1) ongoing traditional practices that never stopped, such as the *Sulla* in the Middle East (Blumenfeld, 2002), 2) newly emergent practices that have been received as brand new inventions, such as diversionary conferencing in Canberra Australia, or 3) recognized forms of indigenous justice, such as Maori family group conferences in New Zealand, which were incorporated into the legislative reform of the formal juvenile justice system in 1989.

The varieties of restorative justice are substantial. The full scope of the concept has been used to describe almost any effort to enable offenders to try to restore or repair some of the damage they have caused to victims or communities by their crimes. This review attempts to isolate what may be the most theoretically powerful form that these efforts take: face-to-face restorative justice. We define this form of restorative justice as any

1. face-to-face discussion between a restorative justice facilitator or mediator and
2. anyone affected by a crime
3. with or about offenders who have accepted the responsibility for having caused the crime
4. addressing the harm they have caused and
5. possible ways the offenders can repair that harm

Among restorative justice programs that meet the test of this definition are Canadian peacemaking sentencing circles. These are usually attended by victims, offenders and the entire affected community and also include members of the mainstream justice system – defence, prosecution and the judiciary. It is also possible for victims to participate in these circles when offenders opt out, or for communities to conduct them without the participation of either victims or offenders. Victim–offender mediation is another form of restorative justice that is usually, though not always, face–to–face. Here a neutral mediator meets with victims and offenders, either together or separately and usually without their supporters, to discuss and resolve the harm caused by the offence.

The proposed review will focus on the subset of face–to–face restorative justice that has been studied most carefully and extensively. That subset is known as conferencing, where as many as possible of the principals involved in the crime meet in the presence of their families, friends or both, even though it is possible to hold a conference without supporters and even without the offender or without the direct victim<sup>1</sup>. At the conference, any victims (or their representatives) present have the opportunity to describe

---

<sup>1</sup> All the conferences in the trials included in this Review had at least an offender present, and a direct victim as well wherever possible.

the full extent of the harm a crime has caused, offenders are required to listen to the victims (or their representatives) and to understand the consequences of their own actions, and all participants are invited to deliberate about what actions the offender could take to repair the harm. The pre-condition of such a conference is that the offender does not dispute the fact that he is responsible for the harm caused, and that the conference cannot and will not become a trial to determine what happened.

The review is further limited to the effects of face-to-face restorative justice conferencing programs for personal victim crimes only. Offences without a direct victim, such as drunk driving without an accident or large-store shoplifting, lack apparent harm to individuals and the absence of anyone in the conference who has experienced such harm results in diminished emotional power. This alters the way in which these conferences unfold and renders them sufficiently different in their content and dynamics as to be inappropriate for inclusion in the review.

Restorative conferencing has strong theoretical connections to Braithwaite's theory of reintegrative shaming (1989), Tyler's theory of procedural justice (1990; Tyler and Huo, 2002), Sherman's (1993) theory of defiance, and Braithwaite's (2002) theory of responsive regulation. Nonetheless, there is no causal theory that describes the exact mechanisms by which face-to-face restorative justice is intended to work. Thus the review of treatment effects must be approached as a testing of policy rather than theory.

Restorative justice theory and practice have developed with enthusiasm over a short space of time, with many optimistic claims being made about its advantages over formal justice processes by its proponents. But rigorous evaluation of the effectiveness of restorative programs on clearly identified outcome measures has tended to lag behind. Numerous small-scale process evaluations have resulted encouraging results regarding participants' perceptions of fairness and satisfaction, but few so far have attempted the more demanding task of measuring reoffending behavior and comparing it with a control group of any kind. The proposed review will be limited to studies that have employed a randomized design to answer the question: what is the effect of restorative justice

conferencing for personal victim crimes on repeat offending by the offenders assigned to such conferences?

As the attached *Registry of Randomized Controlled Trials in Face-to-Face Restorative Justice* (Appendix 1) shows, at least 11 RCTs of restorative justice have been designed to answer this question. Of these, three have at least preliminary findings and eight more are underway. These RCTs vary on many dimensions, each of which poses a challenge to a systematic review of effects that integrates the findings of diverse tests.

Perhaps the most important dimension of variation is the location of RJ and comparison to control at almost every point of the criminal justice process:

- 1) where a decision has been made not to prosecute, and a caution will be administered with or without restorative justice (RCTs # 6 and 7).
- 2) diversion from prosecution to a restorative justice conference that will leave no criminal conviction record (RCTs # 1-3).
- 3) after conviction for a crime but before a sentence is imposed by a Magistrate or Judge (RCTs # 4, 5, 8, 9).
- 4) after conviction for a crime and a community sentence to be supervised by probation officers (RCT # 10).
- 5) after conviction for a crime and a sentence to imprisonment, but before the offender has been released from prison (RCT # 11).

The Registry also shows that these RCTs have examined, or are currently examining, a wide range of offense types, separately and in various combinations with the same random assignment sequences:

- 1) committing property crimes that generally involve no contact with victims (RCTs # 2, 6, 9).
- 2) committing burglary that often involves contact with victims (RCT # 4)
- 3) committing violent crimes only (RCTs # 1, 7, 8, 10, 11)
- 4) committing robbery only (RCT # 5)
- 5) committing either minor violent or minor property crimes (RCT # 3)

The registry further shows that different pools of offender characteristics have been defined as criteria for eligibility for inclusion in the various RCTs, primarily based on the age of the accused:

- 1) first offenders aged 7-17 (# 3)
- 2) any offenders aged 10-17 (# 2, 6, 7)
- 3) any offenders under 30 (# 1)
- 4) any offenders 18 and over (# 4, 5, 8, 9, 10, 11)

Further variability across the RCTs include research design elements that can have major effects on the content of the conclusions drawn. These issues are addressed in the methodology section (4) below.

## **Objectives**

The objective of this review is to derive the most unbiased estimates of the effectiveness of restorative justice programs in reducing repeat offending, and in satisfying victims that justice has been done. Because the RCTs to date have been extremely heterogeneous with respect to different kinds of offences and offenders and different stages of the criminal justice process, the proposed review will disaggregate and regroup findings to answer a wide range of questions about RJ structured to create sensitivity to the heterogeneity of the RCTs.

## **Methodology**

This section describes at the outset the methodological challenge faced by this review, and summarizes our proposed solution. It then describes the dimensions on which the 11 registered RCTS of RJ vary (three completed and eight in-progress). Anticipating that we will find few additional RCTs of this relatively new program, the variable designs of the 11 registered RCTs are described as the basis for a plan to make the most of diverse tests of a similar treatment.

## 1. Systematic Reviews and Inconsistent Designs in Crime and Justice Evaluations

Evaluations of restorative justice pose a challenge to systematic reviews, one that is found in many program evaluations in crime and justice. That challenge is inconsistency of design and analysis, even among randomized controlled trials. The dimensions of inconsistency include, but are not limited to:

- Measures of the dependent variable
- Analysis of the dependent variable
- Time period of analysis of the dependent variable
- Offense and offender characteristics
- Nature of the control treatment to which the experimental treatment is compared
- Percentage of cases in which treatment is delivered, and taken up, as randomly assigned

Two solutions to inconsistency are commonly employed in systematic reviews, but these solutions alone may fail to reveal the underlying truth about treatment effects. While we propose to employ both of those standard solutions this review, we also propose to add a methodological test that may provide a general caution about these solutions in crime and justice reviews. The standard solutions are 1) lowest common denominator analysis, and 2) disaggregation of findings by sample and design characteristics. The methodological test is a comparison of reviews of identical sets of studies using lowest common denominator analysis versus optimal design analysis.

**1.1. *Lowest Common Denominator Reviews.*** By *lowest common denominator*, we mean the features that can be found in all randomized controlled trials (RCTs) of restorative justice that allow comparative analysis of effect sizes—even if these features do not offer the optimal means for drawing causal inference from the RCTs, individually or collectively. An example of lowest common denominator analysis is comparing prevalence of repeat offending (percentage of offenders with repeat offenses) because it

is the only outcome found in all completed RCTs, even though four of the seven have data on offense frequency.

**1.2. Disaggregation Reviews.** By *disaggregation* of findings, we mean grouping all RCTs that meet different criteria for eligibility for a review on a precisely framed, conditional question about the effects of the treatment. An example of disaggregation analysis is examining only RCTs that test restorative justice on violent crimes, as distinct from all RCTs testing restorative justice on all types of crimes. The distribution of magnitude and direction of effect sizes may vary so much from one disaggregation to the next that any conclusions about the complete population of RCTs would be misleading as to the conclusions about the effects of the treatment under tightly defined specific conditions.

**1.3. Methodological Test: Lowest Common Denominator vs. Optimal Design Analysis.** By *optimal design analysis*, we mean an RCT design that features the strongest logic available for the reduction of bias in the interpretation of treatment effects. The lowest common denominator solution may achieve the largest possible universe of eligible reviews, but at the price of increasing bias in the interpretation of results. If a direct comparison of a subsample of RCTs allowing optimal design analysis to the results of using lowest common denominator analysis with the same subsample shows very different results, that test would suggest the dangers of relying on the lowest common denominator analysis in reaching conclusions about treatment effects.

The specific methodological test we propose is based on our initial examination of the five RCTs in restorative justice reported in the program of the American Society of Criminology from 1997 to date. That examination reveals that while three of the RCTs report *after-only prevalence differences* in offending rates, two of the RCTs report *before-after frequency differences* in offending rates. Because all five samples are relatively small, heterogeneous, and subject to gaps of varying sizes between intention-to-treat and treatment-as-delivered, the use of after-only differences increases the risk of selection bias by baseline differences in offending rates between treatment groups. The

use of prevalence rather than frequency of offending measures also decreases the sensitivity of the test, given the fact that a large proportion of all offending is never detected and officially recorded. If, as we argue on logical grounds, a before-after frequency analysis is the optimal test of treatment effects, and the after-only prevalence analysis is at least less effective in eliminating selection bias, then the conclusions of a lowest common denominator analysis of after-only prevalence effects of treatment would have to be considered as limited in value.

1.4. *Current Findings vs. Registered RCTs Underway.* One more reason for adding the methodological test to the initial review is to encourage all future RCTs to employ the optimal design, rather than the lowest common denominator (LCD). Over-emphasis on the LCD results could have the perverse effect of encouraging future RCTs to focus on that design, rather than to raise the design quality of restorative justice experiments. The proposed review design is intended to accommodate the future findings of the 8 registered ongoing trials, as well as findings of any other RCTs we may locate or which may be launched after the first publication of this review.

## **2. Criteria for considering studies for this review**

Studies will be included that use a randomized design to examine the effectiveness for personal victim crimes of a face-to-face restorative justice ‘conferencing’ program. This model entails a meeting between offenders and victims (where the latter can be identified), in the presence of their family and friends and under the auspices of a trained facilitator, where the purpose is to discuss the harm caused by the crime and what needs to be done to repair it. It excludes mediation programs involving only the victim and offender, where the mediator negotiates between the two parties who may never actually meet. The major dependent variable must concern reoffending.

Studies that relate to forms of restorative justice other than conferencing and that do not use a random assignment design will be excluded. For example, Schneider’s (1986) randomised controlled trial is excluded because it concerned a mediation program that did not have the content of a conferencing program described above. Moore and

Forsythe's (1995) evaluation of the Wagga conferencing program is excluded because it did not entail random assignment.

In addition, the reviewers have decided to exclude studies where fewer than two-thirds of offenders randomly assigned to restorative justice actually received restorative justice. Consequently, the two otherwise-eligible Bethlehem studies will be excluded; the Bethlehem property experiment achieved an implementation rate of 48.5 percent and the violence experiment a rate of 31.6 percent.

Decisions about inclusion or exclusion from the Review of published studies will be based initially on report titles, then on abstracts and finally on full reports of ostensibly eligible studies. Reviewers will extract information from the full text of the studies, when published reports are available, and directly from the investigators undertaking the studies where reports are not yet published or where some data are not available in published form. The information will be in accordance with the checklist prepared by Farrington and Petrosino (2000).

The relevance decisions will be made by the reviewers jointly.

### **3. Search strategies for identification of studies**

Evaluation in restorative justice is a recent development because these programs are themselves a recent innovation in justice. We will use formal search tools, including the databases listed below, but anticipate that they will not yield as much information as other less formal sources. These will include narrative and empirical reviews of literature that examine effects on reoffending, bibliographies of restorative justice programs and direct contact with leading researchers.

We will search the following databases from 1986 to the present (the terms 'restorative justice' and 'conferencing' were not in use prior to 1986):

- C2-SPECTR
- National Criminal Justice Reference Service (NCJRS)

- Criminal Justice Abstracts
- Sociological Abstracts
- Criminal Justice Periodicals Index
- Dissertation Abstracts
- Social Science Abstracts

The search terms to be used will be:

Restorative *and* justice

*or*

Conferenc/e/ing

*with*

Reoffending

Recidivism

Evaluation

#### **4. Description of methods used in primary research**

The studies included in the review draw their sample from offenders who have admitted their offence, and who are randomly assigned to be dealt with either by a criminal justice process appropriate for the offence (such as court) or by a restorative justice program of the kind described above, whether alone or in conjunction with the usual criminal justice process. Measurement of reoffending will be based on official records, usually police records of re-arrest.

For each study, the following will be recorded where the information is available and will be noted where it is not available:

- Comparability of experimental and control groups
- Violation of randomization procedures
- Case attrition from the initial sample.

Official records of reoffending and self-reported reoffending will be reported separately.

Where available, the data will be organized by incidence (the percentage of the

experimental and control groups who reoffended), prevalence (the average number of offences in each group), severity of reoffending (the percentage of each group that later committed violent crimes) and ‘time to failure’.

## **5. Criteria for determination of independent findings**

Usually in the studies that meet the criteria for inclusion in this Review, reoffending is not the only outcome measure. Often in these studies an equal amount of attention is given to measurement of victim satisfaction with the restorative justice process, compared with normal treatment. These measures relate to the same *cases* as those involving the offenders whose reoffending rates will be measured, but clearly concern the victims in these cases, not the offenders.

In those studies where repeated measures are made relating to reoffending at several points in time, we would expect some variability and we will report them accordingly. We do not intend to conduct a meta-analysis as most studies will have been designed to differentiate the effects of the restorative justice intervention for different kinds of offenders and offences.

The Review is designed easily to incorporate both additional studies as they are completed and studies that may have been missed by the reviewers. This will be via two mechanisms: first, the Registry will disaggregate basic features of the study – offence type, offender age etc – that will allow users readily to identify studies relevant for their purpose and allow the reviewers readily to add new studies; second, forest graphs will show both individual studies and aggregated studies depending on where we are commenting on the specific features of any study or wishing to compare effectiveness of the intervention across studies.

Definition of success and failure will be on the basis of crime reduction and victim satisfaction. The effect size metric will be Cohen’s D.

## **6. Statistical procedures and conventions**

The Review will report the statistical procedures used in each study. There will be no attempt to aggregate the statistical findings owing to the heterogeneity in the procedures and designs in the studies reviewed. However, we will note the direction of the results in each study. When the studies report on missing data, that information will be included. Every effort will be made to discover directly from the study authors what happened to missing cases etc.

### **7. Treatment of qualitative research**

The primary criterion for selection of studies for inclusion in the review will be a randomised design. If qualitative research is included in these studies, it will be reported as well.

### **8. Grouping of Effects by Baseline Sample Characteristics at Point of Random Assignment**

1. All RJ
2. All diversion from prosecution (RCTs # 1-3, 6, 7)
3. All juvenile court experiments (RCT # 2)
4. All treatment-as-delivered rates at 90 percent of intention-to-treat (ITT) cases (no RCTs so far completed)
5. All property crime (RCTs # 2, 6, 9)
6. All violent crime (RCTs # 1, 5, 7, 8, 10, 11)
7. All RJ ITT cases
8. All ITT cases with few priors (cut point to be decided)
9. All ITT cases with many priors (cut point to be decided)
10. All ITT cases with direct victim present (post random assignment)

### **Timeframe**

It is anticipated that the Review will be completed within three months of the acceptance of the Protocol.

## **Plans for Updating the Review**

Co-reviewer Heather Strang will be responsible for updating the Review every two years.

## **Acknowledgments**

No additional individuals beyond the two reviewers listed on the cover sheet will contribute to the preparation of the Protocol.

## **Conflict of Interest**

The reviewers have conducted most of the studies to be reported upon, but do not see this as a conflict of interest for reviewing purposes.

## **References**

Blumenfeld L (2002). Revenge. A Story of Hope. New York, Simon & Schuster.

Braithwaite J. (1989). Crime, Shame and Reintegration. Cambridge, Cambridge University Press.

Braithwaite J. (1998) “Restorative Justice” In Michael Tonry, editor, The Handbook of Crime and Punishment. Oxford: Oxford University Press.

Braithwaite J. (2002) Restorative Justice and Responsive Regulation. Oxford: Oxford University Press.

Moore D, L Forsythe (with T O'Connell). (1995). A New Approach to Juvenile Justice: An Evaluation of Family Conferencing in Wagga Wagga, Charles Sturt University, Wagga Wagga.

Schneider A. (1986). 'Restitution and recidivism rates of juvenile offenders: results from four experimental studies', Criminology, 24:533–52.

Sherman, L. W. (1993). 'Defiance, deterrence and irrelevance: a theory of the criminal sanction', Journal of Research in Crime and Delinquency, 30:445–473.

Tyler, T. (1990). Why People Obey the Law. New Haven: Yale University Press.

Tyler, Tom and Yuen J. Huo (2002). Trust in the Law: Encouraging Public Cooperation with the Police and Courts. N.Y.: Russell Sage Foundation.

## **Tables and Figures**

See attached Registry