Do John Schools Really Decrease Recidivism?

A methodological critique of an evaluation of the San Francisco First Offender Prostitution Program

by Rachel Lovell and Ann Jordan

July 2012

Summary

A growing number of governments are creating “john schools” in the belief that providing men with information about prostitution will stop them from buying sex, which will in turn stop prostitution and trafficking. John schools typically offer men arrested for soliciting paid sex the opportunity (for a fee) to attend lectures by health experts, law enforcement and former sex workers in exchange for cleared arrest records if they are not re-arrested within a certain period of time. A 2008 examination of the San Francisco john school, “Final Report on the Evaluation of the First Offender Prostitution Program,” claims to be the first study to prove that attending a john school leads to a lower rate of recidivism or re-arrest (Shively et al.). Despite its claims, the report offers no reliable evidence that the john school classes reduce the rate of re-arrests.

This paper analyzes the methodology and data used in the San Francisco study and concludes that serious flaws in the research design led the researchers to claim a large drop in re-arrest rates that, in fact, occurred before the john school was implemented.
Introduction

Over the last 15 years, the issue of human trafficking has risen to the top of the global agenda for many governments, feminists, human rights advocates and social service providers. In the majority of countries, the focus has been on the trafficking of women and girls into forced prostitution. Many of the resulting efforts to prevent trafficking have, consequently, ignored the fact that a large percentage of trafficking and forced labor occurs in other sites (such as farms, homes, streets and businesses) and involves large numbers of men and boys.¹

Some individuals and groups believe that arresting men who solicit commercial sex will stop men from soliciting and also stop prostitution and trafficking into prostitution. They target male buyers as the “cause” instead of addressing the reasons why people decide to sell sex in the first place or why sex workers are unwilling or unable to find other work. This position has become quite widespread and has led to an imbalance of resources, with money, time and people dedicated to eliminating “demand” rather than addressing the social, economic and other circumstances that lead people to sell sex in the first place. A considerable amount of those resources have focused on “john schools” in which men arrested for soliciting or purchasing paid sex agree to participate in a class about the harms of and laws on prostitution in exchange for a cleared record after a certain period of time.

The first modern-day john school was the First Of-fender Prostitution Program (FOPP) in San Francisco. Today there are several dozen john schools in the U.S. and a few in other countries (e.g., Canada and South Korea). However, no study to date has documented a causal connection between a john school program and a decline in recidivism (re-arrest) rates among male purchasers of commercial sex or a reduction in the incidence of prostitution or human trafficking.²

Nonetheless, john schools continue to be promoted as a chief tool in anti-trafficking and anti-prostitution campaigns. Many continue to believe—despite the lack of evidence—that if john schools can prevent men from reoffending, then the schools are a success and have had a positive impact on stopping prostitution and trafficking. The most influential study cited in support of john schools is the previously mentioned 2008 “Final Report on the Evaluation of the First Offender Prostitution Program” (FOPP Report) that was conducted by Michael Shively et al. on the San Francisco john school. The authors claim that the FOPP Report

¹Forty-four percent of the people held in forced labor globally are men and boys (ILO, 15).

²A report on the Toronto john school “traces the ambiguous nature of the programme’s objectives by contrasting its widely promoted ‘educational’ and ‘constructive’ aims with the more punitive qualities that emerge in practice” (Fischer et al., 385). Another Toronto study asked men what they would do after the program but did not report actual recidivism rates (Wortley et al). A study of a defunct program in Portland, Oregon, found that “the rate of recidivism among men participating in the Portland program was low. However, recidivism among men who did not participate in the program was also low. The findings suggest that recidivism may not be a useful measure of effectiveness for programs aimed at men arrested for trying to hire street prostitutes” (Monto and Garcia, 1). A review of empirical literature on evaluation and recidivism concluded that programs have little to no impact (Kennedy et al.). Wilcox et al. found that “educative approaches, such as ‘John schools’, have demonstrated attitude change but have not changed behaviour” (Wilcox et al., ii).
is the first study to definitively prove that a john school reduced the rate of recidivism. They cite a reduction of 30 to 50 percent, depending on the type of statistical design used (Shively et al., 78, 81).

According to the FOPP Report, “[t]he First Offender Prostitution Program (FOPP) is designed to reduce the demand for commercial sex and human trafficking in San Francisco by educating men arrested for soliciting prostitutes (or “johns”) about the negative consequences of prostitution” (Shively et al., Abstract). “The ultimate goal is to decrease the demand for prostitution, and, hence, reduce the amount of human trafficking and sexual exploitation that occurs” (Shively et al., 23). The authors do not contend or attempt to prove that the San Francisco john school reduced prostitution or trafficking into prostitution. Instead, they say that the FOPP reduces recidivism and leave it to the readers to assume that a drop in recidivism may also have led to a reduction in prostitution and trafficking.

This paper critically examines Shively et al.’s claim of reduced recidivism by analyzing the underlying data and the statistical methodology. It demonstrates that, due to flaws in the methodology, the FOPP Report presents no evidence proving that the FOPP is effective at reducing rates of recidivism. Therefore, governments should exercise caution before adopting this or any similar john school program.

**Background on the San Francisco First Offender Prostitution Program (FOPP)**

The FOPP consists of three components:

1. Arrest of a first-time offender (someone with no previous domestic violence or sexual offense),
2. Participation in the john school classes (for a fee), and
3. Cleared record if the offender is not rearrested for a prostitution-related offense within a one-year period.

The FOPP study only focuses on one component—participation in the john school—and does not consider the impact of arrest or the cleared record. Selective omission of the other two extremely important elements of the process has implications for the validity of the research, which is discussed later in this paper. But, first, it is important to “unpack” the research claims about the alleged impact of the john school on recidivism.

**Claim that FOPP Reduced Recidivism Not Supported**

The FOPP Report has gained widespread support because of its claim that it is the first study to link a john school to a reduction in recidivism among men who solicit paid sex. It asserts that the San Francisco john school has cut the recidivism rate by approximately 30 to 50 percent (Shively et al., 78, 81). However, if one tunnels deeply into the report, the picture changes completely.
San Francisco recidivism rates dropped prior to the FOPP and rose seven years after the FOPP

Figure 1 was created using data compiled from Table 21 (p. 75) of the FOPP Report. It includes data ten years before and ten years after the implementation of the FOPP in 1995.

The Figure 1 data reveal that:

1. From 1985 to 1994, San Francisco had a significantly higher recidivism rate than the rest of California.4

2. In the two years before the FOPP was implemented, San Francisco’s recidivism rate dropped dramatically.

3. In 1995 when the FOPP was implemented, San Francisco and the rest of California had the same re-arrest rates (San Francisco, .061, and the rest of California, .060)

4. In the first year after the FOPP, San Francisco and the rest of California had a similar drop in recidivism rates.

5. From 2002 to 2005, San Francisco’s...
rate of recidivism was higher than that of the rest of California.

After the FOPP was implemented and until 2001, San Francisco’s re-arrest rates were almost the same as the rest of California’s. Thus, even if the Shively et al.’s report had been able to show that the FOPP classes had an impact during this period, the effect would have been insignificant since the rates for San Francisco and the rest of California were more or less the same. Even more interestingly, in 2001 San Francisco’s re-arrest rates suddenly increased while rates in the rest of California decreased. In other words, not only did San Francisco and California have the same re-arrest rates after the FOPP, San Francisco’s recidivism rates even increased after the FOPP while rates in the rest of California—without a FOPP (except for San Diego)—decreased. The FOPP study does not address or explain this surprising phenomenon.

So, given the data presented in Figure 1 above, how did the FOPP researchers attribute an approximately 50 percent drop in re-arrests to the FOPP? In essence, they averaged the data both 1985 to 1994 and from 1995 to 2005 (Table 22, p. 79 and Table 23, p. 82). As Figure 1 clearly demonstrates, this averaging allowed the researchers to include the dramatic decrease that had already started in 1993, two years before the FOPP actually began. By averaging the unusually high recidivism rates prior to the FOPP and the more normal recidivism rates after the FOPP, the results showed (1) that the pre-FOPP recidivism rate was higher than the actual rate in 1995 and (2) that part of the pre-FOPP decrease occurred post-FOPP.

In the following sections, we will unravel the methodologies the researchers used to achieve this result and reveal additional flaws that led to those unreliable and improbable conclusions.

San Francisco’s recidivism rates even increased after the FOPP while rates in the rest of California—without a FOPP—decreased.

The research does not measure actual recidivism rates of FOPP participants

It is important to note that the FOPP study does not measure the actual rate of recidivism of men who attended the john school as three of the authors stated. In order to assess the actual rate of recidivism after the FOPP was implemented, the researchers would have had to track the men who took the john school classes to see if they were ever re-arrested. They did not do this. Nor did they compare men who actually attended the john school with other, similar, first-time offenders who did not attend the john school. Instead, they used other methods—two quasi-experimental designs—in a way that produced unreliable or invalid results. It is important to remember these differences when interpreting the results discussed below.

Three of the FOPP Report authors claim that they “…compared re-arrest rates of men exposed to two conditions: 1. Arrest plus john school (FOPP participants) [and] 2. Arrest only” (Shively et al. presentation, slide 9). However, in the FOPP Report the authors explain in great detail (Shively et al., 72–76) that they did not have reliable data on FOPP participants and, therefore, had to compare men who would have been eligible in San Francisco with men who would have been eligible to participate in FOPP in the rest of California.
Randomized Control Trial (RCT): The Gold Standard

The ideal methodology for calculating the actual rates of recidivism of the men who attended the John school would involve a randomized control trial (RCT), which is the gold standard for assessing the effect of any program or intervention. In practice, cost, time and, often, the infeasibility of randomization make this standard difficult to achieve. In this instance, it would have required the researchers to:

1. Randomly assign the men who were first-time offenders for soliciting prostitution to either the treatment group (the men who attend the John school) or the control group (the men who do not attend the John school).

2. Compare recidivism rates for the two groups of men over time.

If there were statistically significant differences between the recidivism rates of the two groups, then the researchers would be able to reasonably claim a correlation between the John school (the “treatment” or the “intervention”) and the rates of recidivism. If the researchers were able to exclude all other factors that could have reasonably caused different recidivism rates, then they would also be able to claim that the John school caused the difference.

Due to the infeasibility of an RCT, the FOPP researchers used two quasi-experimental research designs: a Difference-in-Difference (DiD) design and a Regression Discontinuity (RD) design. These designs are similar to RCT’s because, if done correctly, they compare persons who participated in the intervention with those who did not. However, they lack the key RCT element—random assignment to the treatment and control groups. Nonetheless, through a variety of statistical techniques, these designs create a treatment and a control group without random assignment.

The FOPP study did not compare men who participated in the FOPP with men who could have participated. Instead, it attempts to create substitutes by treating San Francisco as the treatment group and the rest of California as the control group. This methodological design is weak because the two groups are very different, and the DiD and RD designs require the treatment and control groups to be very similar. A strong design would have compared first-time offenders who participated in the program with those who did not.

Methodological Violations In Quasi-Experimental Designs

Statistics don’t lie, but they can be used inappropriately. Both the DiD and RD models have certain assumptions that must be met in order to produce valid, conclusive results. To critique the soundness of the design requires extensive familiarity with relatively complex methodological statistical models. Without going into the statistical details of the models, the designs fail to meet two key methodological requirements.

We have chosen not to provide extensive detail on the statistical designs in this report because we feel the statistical details are distracting. These models are actually quite complex since, for example, they do not actually measure recidivism, they lack true control and treatment groups, and multiple time frames were investigated (1993 to 1994, 1992 to 1994, 1991 to 1994, etc.). In the case of the RD model, the statistical modeling is quite difficult to follow because the authors provide very little detail. However, if anyone would like the statistical particulars for these models, please contact the first author directly at rlovell@depaul.edu.
First, San Francisco and the rest of California did not have similar recidivism trends prior to implementation. A key requirement in the DiD design is that the treatment and control groups have similar recidivism trends prior to program implementation. To measure changes over time, the DiD model assumes two conditions:

1. The recidivism rate is relatively stable in both groups prior to implementation of the FOPP.

2. Trends in the recidivism rates for the two groups would have remained stable if the FOPP had not been implemented.

As shown in Figure 1, these conditions are not met. The San Francisco recidivism rate was not stable prior to the FOPP and the recidivism trends of the two groups were not the same prior to implementation of the FOPP. The consequence of this faulty comparison is that, even if Figure 1 did show a drop from 1995 to 1996 or later, there would have been no way to determine whether the claimed reduction was due to the FOPP or to other unexplained variations (also referred to as statistical noise) in the data or some other unidentified factor that started before the FOPP. The FOPP Report ignores these possibilities. Additionally, it is difficult to make the case that the recidivism rates would have remained stable in San Francisco if the FOPP had not been implemented, given the drastic fluctuations in its recidivism rates.

Second, the researchers assume that the only difference between San Francisco and the rest of California is the FOPP. Even if the FOPP study had produced evidence of a drop in re-arrest rates after the FOPP (which it did not), quasi-experimental designs require the two groups to be very similar in order to conclude that any difference in re-arrest rates after the program was implemented was due to the FOPP.

Are San Francisco and the rest of California really that similar? San Francisco is a city (and a very unique city at that) with its own culture and priorities. Why compare a unique city to the rest of the state? These two “groups” could be different in many ways. For example, San Francisco might treat offenders significantly different than the rest of California—it might dedicate more resources to arresting for solicitation and conduct more “decoy operations” (where female police officers pose as prostitutes) than the rest of the state, putting men at a higher risk of arrest and re-arrest even before the FOPP. Other researchers of the FOPP have established that “a direct correlation exists between the number of decoy operations, men arrested for soliciting prostitution, solicitors eligible for FOPP, and solicitors actually participating in the FOPP” (Management Audit, 13).

The FOPP Report authors did not consider this or any other possible differences between the two groups. They did not provide any theoretical justification for choosing the rest of the state as the control group. They even included San Diego in the control group (as part of “the rest of California”) despite the fact that the authors mention that San Diego actually has a john school (Shively et al., 82).

The San Francisco recidivism rate was not stable prior to the FOPP and the recidivism trends of the two groups were not the same prior to implementation of the FOPP.
Do John Schools Really Reduce Recidivism?

Also, the researchers did not incorporate a time lag into their analysis. In other words, they assumed that the FOPP impacted re-arrest rates almost immediately—in the same year that the program was implemented—actually, in March 1995. Even if the program had been implemented in January 1995, it would be reasonable to assume that an impact, if any, would take at least one year to become apparent.

In conclusion, since the FOPP research does not meet these basic methodological requirements, it does not and cannot demonstrate that the FOPP is related to or the cause of any drop in the recidivism rates of the FOPP participants.

The Devil Is in the Details

In addition to the methodological flaws discussed already, the report also contains other problems that the authors do not adequately address.

The claim of a 30–50 percent reduction in recidivism after a one-day program is counter to all known research

The authors claim that the FOPP reduced the recidivism rate in San Francisco by approximately one third to one half, depending on the type of statistical analysis conducted (Shively et al., 78, 81). That number sounds astonishing. The authors admit that this finding even surprised them. They note that “[the FOPP’s] design appeared to violate several of the principles of effective intervention with offenders that have been derived from more than 40 years of research.”

In contrast, just a year after the FOPP Report was released, the City and County of San Francisco conducted its own study of the FOPP and found that the program could not have reduced recidivism. In the 2009 Management Audit, the San Francisco Budget Office concluded that the “FOPP is not a sufficiently comprehensive program to be effective in reducing recidivism or assisting women to leave prostitution. … FOPP does not meet the National Institute of Justice’s criteria for programs to reduce recidivism, which includes providing intensive services, long-term intervention, and follow-up with men who solicit prostitution. Nor does FOPP provide sufficiently comprehensive services to assist women in leaving prostitution” (Management Audit, Letter to Board of Supervisors, 2-3). In other words, the FOPP is a seriously flawed program, notwithstanding the “success” claimed in the
Do John Schools Really Reduce Recidivism?

Additionally, even if the FOPP Report had shown any drop in re-arrest rates after 1995 (which it did not), it would not have been able to prove that the FOPP was responsible. The FOPP authors dismiss alternative explanations for the claimed decrease in the re-arrest rate (Shively et al., 83-86):

- They claim that since the FOPP did not teach men how to avoid re-arrest, it is “unlikely” that this caused the decline. They did not consider that, once arrested, men who continue purchasing sex may take great care to avoid re-arrest whether or not they are “taught” how.

- They claim that the FOPP does not push men to other jurisdictions because the research can “capture rearrest anywhere in California” (p. 84). However, the researchers did not “capture” re-arrests of the actual men who took the FOPP classes, only re-arrests in general.

- Lastly, they claim that the FOPP does not push prostitution online since displacement from the streets would be caused by arrest alone. This simply points out the reality that arrest has a significant impact on men’s (and sex workers’) behavior and that the FOPP classes had no additional impact.

Most importantly, all these explanations rest on the assumption that the data are correctly modeled, which we have shown is not the case. Therefore, the research does not prove that the FOPP reduced recidivism.

The research fails to account for the effect of the arrest and/or cleared record on recidivism

In addition, the researchers did not take into account other elements of the legal process in San Francisco. As stated previously, the FOPP process consists of three elements: (1) the arrest of a first-time offender, (2) the john school (the intervention) and (3) the cleared record after a year. The FOPP researchers only focused on the john school, ignoring the question of whether arrest or the promise of a cleared record alone could cause a reduction in recidivism rates among actual participants. The authors reached their conclusions about the effectiveness of the FOPP without even attempting to separate the arrest or cleared record effect from the john school effect.

The most important reason reported by the men for attending the FOPP was to avoid a criminal record—68 percent (Shively et al., 63). It seems logical that, if the men knew they could avoid a criminal record if they were not re-arrested for a year, they would have made every effort to avoid re-arrest. It is also logical to assume that participants simply saw the classes as something they had to do in order to qualify for the cleared record option. Participation in the classes would probably have very little, if any, impact on their future behavior because they had already decided not to re-offend before starting class.

When men were asked—before starting the classes—whether they would try to solicit sex in the future, 73 percent said they would not seek out pros-

---

The FOPP Report also claims that the program is supported by the fees the men paid (Shively et al., 88-93). However, the San Francisco Budget Office found that “FOPP program costs exceed fee revenues, especially costs incurred by the SFPD” (Management Audit, Letter to Board of Supervisors, 3). The fee would have to almost double from $1,000 to $1,908 and all participants would have to pay the full fee in order to cover all the program costs (Management Audit, Letter to Board of Supervisors, 3).
titutes. After the FOPP, 70 percent said they would not seek out prostitutes (Shively et al., 66). In other words, before the men even started the classes, they had already decided not to re-offend. This demonstrates that the arrest and the possibility of a clean record had the greatest impact on the men’s intention to change their behavior, not the john school. Obviously, without real data on the men’s actual behavior, these responses are not evidence of changed behavior. However, they do show that the john school had no influence on what they believed they would do in the future.

Many men in the classes paid little attention to the presentations. The report reveals that men slept, read newspapers or magazines or appeared “disengaged” (Shively et al., 55). This behavior is consistent with the view that most of the men had already decided to change their behavior and they knew that their record would be cleared in a year. They simply had to sit through the class and not be re-arrested in order to achieve their goal.

In a study of the Toronto john school, Wortley, Fischer and Webster noted that “the vast majority of respondents [in their study] claim that they will never use prostitutes again—even before they actually attend the educational component of the John School process. In other words, the John School appears to have no significant deterrent effect above and beyond arrest and subsequent criminal proceedings. … Indeed, the official recidivism rate produced by the Toronto John School (2.4%) is almost identical to the re-offence rate (2.0%) produced by regular criminal justice proceedings brought against male prostitution offenders in Vancouver [citation omitted]” (Wortley et al., 393-94, emphasis supplied).

It appears then that, if there is an effect on the men who entered the FOPP, it comes from being arrested and from knowing that their record will be cleared in a year if they attend the school and do not re-offend. Only arrests can actually be measured. The FOPP researchers address this issue: “One could argue that the arrest process alone—and not the john school program—could teach men how to avoid recapture, but arrest is a constant across all of the cases in our database. Since all of the men have the experience of arrest, something besides arrest must have produced the post-1995 drop in recidivism rates in San Francisco” (Shively et al., 84). This could be a valid assumption only if the research on recidivism in the study had produced reliable evidence of a drop in recidivism after the FOPP. The study’s serious flaws and lack of demonstrable evidence to support a link between the FOPP and the drop preclude the authors’ claim that “something besides arrest” caused the decline.

The more significant impact on men’s behavior might be the cleared record. What first-time offender would not pay a few hundred dollars to attend a “school” if, after a year, charges against him

---

**When men were asked—before starting the classes—whether they would try to solicit sex in the future, 73 percent said they would not seek out prostitutes. After the FOPP, 70 percent said they would not seek out prostitutes.**
are dropped? To determine whether the cleared record could have affected the recidivism rates, the researchers would have had to compare the FOPP with (1) another city with a john school that does not offer the cleared record at the end of a year and (2) another city without a john school that does offer a cleared record after a year. The FOPP researchers did not make either of these comparisons.

Conclusion

The FOPP Report fails to prove that recidivism rates decreased as a result of the FOPP for at least two major reasons: (1) data clearly demonstrate that the large drop in recidivism rates occurred pre-FOPP and (2) the research violates key methodological requirements for valid findings. Additionally, if the research had produced findings linking the FOPP to a reduced recidivism rate, it would still suffer from unreliability, for not dealing adequately or at all with alternative explanations for any alleged reduction.

Since the john school model is being replicated around the U.S. and exported to other countries, it is imperative that proper, methodologically sound evaluations of several different models of john schools be undertaken. The research should evaluate the schools for their impact on sex sellers and sex buyers and should consider whether such efforts meet concrete objectives such as reducing prostitution and human trafficking into prostitution.

If the only goal of the john school is to stop first-time sex buyers from re-offending, is that useful? The real goal of jurisdictions sincerely wanting to reduce prostitution should be to address the underlying social, economic, cultural and other reasons that lead people into prostitution in the first place and the impediments that prevent them from voluntarily abandoning prostitution. Simply putting men in prison or sending them to a john school will not lead to long-term solutions for sex workers or at-risk youth who turn to prostitution. The United States has been arresting clients for decades and yet prostitution flourishes. Perhaps it is time to recognize that law enforcement is not the solution to what is basically an economic problem (Weitzer, 2009).

Acknowledgements

This publication was made possible by the generous support of the Oak Foundation, an anonymous donor and DePaul University’s College of Liberal Arts and Social Sciences Dean’s Office.

The opinions, findings, conclusions or recommendations expressed in this publication are those of the SSRC and the Program on Human Trafficking and Forced Labor and do not necessarily reflect the views of its funders.

The authors would like to gratefully acknowledge SSRC staff Linda Levendusky and Jessica Speer for their invaluable assistance in preparing this report.

© Copyright 2012. American University Washington College of Law and DePaul University. In the spirit of the United Nations’ encouragement of collective efforts at the international level (Resolution 49/184), this document is placed in the public domain and put at the disposal of all interested persons to consult it or use it. Reproduction is authorized provided that: (1) the text is reproduced in its entirety without changes, (2) reproduction is for educational ends and not for commercial use and (3) credit is given to the Program on Human Trafficking and Forced Labor, Center for Human Rights and Humanitarian Law, American University Washington College of Law and the Social Science Research Center, DePaul University.
References


