

The Assessment of Penological Effectiveness; The Rapsheet Model of Evaluation

John P. Conrad
Davis, California

RÉSUMÉ

L'évaluation des programmes pénologiques a compté trop sur les dossiers criminels. Bien que cette modèle d'évaluation soit convenable pour quelques programmes, pour la plupart, les résultats défigurent les réalités des bonnes programmes, souvent avec un effet malheureux; la programme est terminée. Je conduirai une critique d'une évaluation d'une programme de conseils en groupe dans une prison californienne. L'évaluation choisie suivit la modèle du "rapsheet," et la programme fut terminée. Des leçons de ce malheur inutile sont offertes. Quelques alternatives plus réalistes sont suggérées.

ABSTRACT

The evaluation of correctional programs has relied too much on "rapsheets." Although this model of evaluation is suitable for some programs, for most the results will distort the realities of good programs, often with the unfortunate effect of ending the program unnecessarily. I present here a critique of a well-known evaluation of group counseling in a California prison. The evaluation followed the "rapsheet" model and the program was terminated. Lessons from this needless misfortune are offered and some more realistic alternatives are suggested.

If any organized program offers a simpler method of evaluation than penology, I have not stumbled across it. When we want to know whether a penal system as a whole or any part thereof is effective, all we need to do is to collect the arrest records, (the "rapsheets"), of the criminals exposed to it, count the raps since the last exposure and draw conclusions that will seem incontrovertibly valid, even to a reasonably well-informed public. The model can take several forms, depending on what is to be evaluated. It lends itself to at least the appearances of controlled experiments, and sometimes, to rigorous and valid appraisals. The centre piece of the methodology is the rapsheet, a document that lends itself to a sharp dichotomy and an effortless interpretation. Because of rapsheet evaluations we know that "nothing works." We have known it for so long that anyone arguing to the contrary is identifiable to neo-conservative critics of criminal justice as a coddler of criminals.

The harm that this model has done to penological programs exceeds the benefits by a wide margin. Some iconoclast ought to explore this theme at much greater length than is available to me here. As an incentive to such a spoil-sport, I will offer a prime example from the abundant literature of penological assessment.

The Group Counseling Movement

In 1955, Norman Fenton, then a Deputy-Director of the California Department of Corrections and a clinical psychologist, proposed as an innovation in penological treatment a program of lay group counseling. Prison guards and other employees so inclined were invited to become counselors for small groups of prisoners. They were to submit to a short course of training conducted by Fenton himself and then to embark on ancillary careers as group counselors. Nobody was a full-time counselor; one hour a week was all that was asked, and that hour was paid as overtime.

Dr. Fenton struck a seductive theme. He suggested that the men and women of the Department had chosen their unattractive occupations out of a wish to help their fellow man, not out of an inclination to punish. A surprising number of guards succumbed to this lure — some, I suppose, because they liked the idea, some because they saw participation as a promising foothold on the long career ladder ahead. The Adult Authority, as the parole board was grandly designated in those years, decided to give the program its full support.

A great head of steam was built up. Prisoners got the message that the Department wished to help as well as to keep them locked up. There is some reason to believe that many of them had their doubts but signed up because of the appearances. Enrollment in group counseling rapidly filled the available groups. In some prisons there were waiting lists.

The objectives of group counseling, as enunciated by Fenton were,

“(1) To help prisoners adjust to the frustrations that are an unalterable part of life in an institution and in society.

“(2) To enable the clients to recognize the significance of emotional conflicts as underlying criminality.

“(3) To provide the opportunity for the client to learn from his peers about the social aspects of his personality.

“(4) To make possible a better understanding of make-believe, of phantasy, (*sic*) and of how costly may be behavioral responses to the antisocial content of daydreams.

“(5) [To improve] the emotional climate of the institution. (Fenton, 1961: 51-55).

As seen by the research team that was to evaluate this program with these vague goals and eclectic methods, the program “seems to be based . . . on certain tenets of Freudian psychology, Rogers’ nondirective counseling, group psychotherapy, social casework practice, and sociological studies of the social organization of the inmate community, (Kassebaum, Ward and Wilner, 1971: 61). The program was to be distinguished from group psychotherapy, as practiced by professional therapists, but the conceptual nature of the distinction was never clearly established. Over and over again it was stressed that the group setting was to be a situation of trust in which the “clients” were to feel free to discuss their own and each other’s feelings and attitudes. Everything that was said in the group was to remain in the room in which it was said. The relationship among the clients in the group

was to be a condition of mutual acceptance. It was important to Fenton that in his generous treatment model the members of these groups were neither prisoners, nor convicts, nor inmates: they were *clients*.

It is important here to note that these objectives were oriented to the prisoner's present condition. Nothing is said about the influence of the experience on his future adjustment to the world outside the prison. All that was *explicitly* hoped for was some honest introspection by the client, and an improvement in the prison environment. In the original conceptualization, at least, it was an artless sort of innovation, and it didn't cost much. Faced with increasing numbers of prisoners and few programs to keep them occupied, prison wardens gladly accepted the slight administrative burdens entailed by group counseling. If the program also reduced tensions that would be a welcome bonus.

The rapid spread of group counseling throughout the Department created on Fenton's staff a euphoria by which success was almost taken for granted.¹ Some of us, especially Fenton himself, thought of this simple innovation as the instrumentality that would revolutionize American penology. It was gratifying to learn of prison systems elsewhere in the country and in Europe, too, that were adopting group counseling as the means for transforming prisons from dungeons to facilities for the rehabilitation of offenders. Those of us who came to the Department as psychologists or social workers saw our choice of a penological career as justified at last. We could look ahead to the conduct of counseling and the training of lay counselors as the core of our professional practice.

Riding this wave into the future, it was clear to Fenton and to the rest of us that the next step was to conduct a rigorous evaluation of group counseling. With all the national and international attention the program was getting, it was not difficult to persuade the National Institute of Mental Health to consider arrangements for the support of an evaluation of the program. Three social scientists based at the University of California at Los Angeles, (UCLA), Gene Kassebaum, David Ward, and Daniel Wilner, assumed responsibility for undertaking the research. It is their study that I want to recall for review.

I do not know of any evaluation in penology that was more complex than the study completed by this industrious trio. It was a model of rigor. It also destroyed group counseling in prisons both as a concept and as a practice. Fifteen years later, I doubt that the lessons of its design, implementation and conclusions have yet been fully absorbed.

The Research Design

By a piece of good luck, the grant award coincided with the opening of a new medium security prison at San Luis Obispo, the California Men's Colony, (CMC). The superintendent of the new prison, John Klinger, had

* ¹ I write as one who was marginally engaged in the practice and a regular participant in staff meetings. For a while, I was liaison for the Department with the Kassebaum-Ward-Wilner research.

been Chief Deputy Director of the Department and in that capacity had made the plans for its architectural design. With a maximum capacity of 2400 prisoners, CMC was divided into four "quads," each holding 600 men. Each quad could be and was hermetically sealed from the others. Prisoners in different quads were in contact with each other only in the shops and schools, never in their living quarters. Each quad had its own staff.

This arrangement was ideal for a controlled experiment and the research group was quick to appreciate its advantages. Quad A was designated as an experimental locale for voluntary counseling in groups of 10-15, for which 450 of the population were eligible. The remaining 150 were assigned to "mandatory" large group counseling in which groups of 50 men met four mornings a week. Quad B was divided into two sections. One section of 300 men was required to attend group counseling in small groups twice weekly, while the other section of 300 comprised one sub-section of voluntary group counseling, and another subsection of voluntary controls. Quad C was the mandatory control group, in which no group counseling was allowed, but all other programs were available. Assignments to Quad C were randomly drawn from the same pool of eligible prisoners from which assignments to Quads A and B were made. Quad D was out of the study; it held some psychotics, prisoners who presented "management problems" and "troublesome homosexuals." Selection of men to be assigned to this quad was reserved to the central office classification staff.

Except for the men sent by the Sacramento office to Quad D, once a prisoner arrived at CMC he could not be transferred to another institution.² Assignment to one of the research quads was carefully randomized. Superintendent Kinger, a compulsive man, had agreed to maintain these policies in the interest of a uniform and controlled exposure of the population to the independent variables. He also had a lively appreciation of the advantages to himself and his staff of a stable population. For years CMC enjoyed a relatively quiet, non-violent, medium custody prison community, and at a time when several other California prisons were experiencing violent incidents that brought them world-wide notoriety.

CMC was opened to its guests during the period June-December 1961, and assignment of prisoners to quads followed the requirements of the research design. The study program began in January 1962 and continued through December 1963 when the last of the study subjects were released on parole. Post-release data were collected through June 1967. During the next two years, the data were analyzed and prepared for publication.

* ²This was an important deviation from Departmental practice; a traditional incentive to compliance was the prospect of transfer from a prison with a higher custodial security to one with lower requirements. To my knowledge, the effect of this exception to normal procedures upon prisoner morale has not been studied.

The Findings

From the theory of group counseling, as enunciated by Fenton, the research group abstracted three hypotheses:

- H₁ Participation in treatment results in lessened endorsement of the inmate code.
- H₂ Inmates who participated in the group counseling program will receive fewer disciplinary reports.
- H₃ Parolees who participated in the prison group counseling program will have lower recidivism rates than controls (Kassebaum, et al: 71-72).

To test H₁, questionnaires were administered to both experimentals and controls. An attempt was made to differentiate prisoners according to the typology proposed by Sykes 1958 but the research group had to conclude that this tidy paradigm did not fit the CMC population. The authors noted: "We found that we could assign very few men to inmate "types" either through classification schemes of other or through our own efforts." (Kassebaum, et al: 297). Nevertheless, the questionnaire responses did not establish any change in prisoner attitudes and values.

The disciplinary reports for each man in the study were systematically collected and analysed to test H₂. No significant differences were found between the men in the various experimental subgroups and those in the control group.

Much more attention was given to parole survival, the criterion for acceptance of H₃, but here, too, the differences between experimental and controls were not statistically significant. At the end of three years 51 percent of the entire study group were back in prison on new offenses or on "technical" violations of the conditions of parole.³ This was the rapsheet score that eventually sank group counseling, a program of which we hear very little in these times when the rehabilitative goals of penology have been eclipsed by the retributivist orientation of the "justice model." (Fogel and Hudson, 1981: 3-51). Social science established that group counseling didn't work, just as hardly any other penological intervention works.

Was This Evaluation Necessary?

The demise of group counseling was unfortunate and unnecessary. Between the hubris of the innovator and the insistence of the legislature's program analysis staff, an evaluation was inevitable. Dr. Fenton was as certain

* ³At that time, parolees were subject to 14 general conditions of parole, ranging from "cooperation with the parole officer" to unauthorised possession of firearms. In addition parolees with special problems might be subjected to special conditions, e.g., no contact with an estranged spouse. Violations of these conditions was sufficient cause for a return to prison "to finish term," but the Adult Authority did not always revoke parole on a technical violation.

as an incorrigible optimist can be that the impact of group counseling on those who participated would be reflected in dramatically superior performance after release. He had no qualms about subjecting his program to a rigorous ordeal by rapsheet.

The legislative analyst, Fred Lewe, had been a sceptic about the Department's treatment programs for many years. The discovery that methods existed for their evaluation by post-release recidivism gave him an instrument for the support of his doubts that anything useful could be done for convicts. Even though the group counseling enterprise was one of the least expensive innovations to be presented for his review, it had to be evaluated and the criterion for success was a statistically significant reduction of recidivism. I don't think he had in mind anything as elaborate as the Kassebaum-Ward-Wilner study, but he was delighted with the grant from a prestigious Federal agency, and not at all surprised at the disappointing results.

I maintain that this evaluation was misguided and inappropriate. Because the practitioner's over-confidence can again combine with the budget analyst's empiricism, I want to consider the shaky conceptual foundation for this massive, rigorous, and elegant evaluation. Though it's unlikely that anything so costly in time and funds will be undertaken again in the evaluation of a penal program, the lack of attention to conceptual foundations is a recurring problem in penological evaluation and, I daresay, in other public programs as well.

1. **The Wrong Goal:** Group counseling was committed to the notion that a daily hour-long session would affect those exposed to it so profoundly that their behavior after release on parole would be reflected in a lower rate of recidivism. By my hindsight, (and reasonable foresight that was not exercised), this goal was inherently incredible. Life is not like that, especially the lives of ex-convicts precariously situated in a society that barely tolerates their return. Unemployment or marginal employment, rejection by former friends, and the prospect that conventional living will be always inaccessible constitute a set of realities that face most released prisoners. For an increasing number of other men and women, the values that group counseling sought to instil and the processes by which those values were to be internalized were alien to the rough world of the underclass to which they would return. These realities would lead a parolee with the best of intentions to new offensive behavior. Combined, they are the basis of recidivism. The influence of some benign experience during the past years of incarceration is insufficient to offset present adversity.

As if this gloomy reality were not enough to shake Fenton's optimistic theory, we must also face the inability of social science to demonstrate the efficacy of psychotherapy by any objective criterion. Case histories of qualified successes and the continued demand for the many psychotherapies now available suggest a large number of believers, but rigorous proof of value has never been in sight. It was naïve to suppose that the efficacy of group counseling conducted by untrained guards could be objectively demonstrated even though professionally conducted psychotherapy had resisted any such assessment. Our naïveté cost us dearly.

2. **No Disseminable Package:** In an invaluable dialogue, Donald Campbell, that sagacious elder statesman of program evaluation, told his interlocutor, Kenneth Watson, that formal evaluations should be reserved for programs that constituted *disseminable packages*, i.e., "programs that can be replicated in other contexts with reasonably similar results . . . a textbook . . . the 55 miles-an-hour speed limit . . . the Japanese quality circles." (Campbell, 1986: 83-86). This was a term that had not crossed our horizon in those heady days. Nevertheless, I am on record in 1965 as suggesting that group counseling lacked the standardization that would have made the CMC model disseminable:

"From the first, there has been an eagerness to prove that this one or two hours of the week of the offender has enough impact to make a difference in the outcome of his confinement, regardless of what happens during that hour or two, and regardless of what happens during the rest of the week. The fact that the content of group counseling ranges from the near-didactic to the near permissive has seldom been faced squarely; neither has the question of the relative competence and motivation of the counselors, a variable of even greater significance." (Conrad, 1965: 246).

All that was disseminable in the CMC package was daily group meeting, conducted by a guard or some other institution employee, but not necessarily the same person. The group counseling coordinator remarked to one member of the research team that "we have a stable group if there's less than one leader change during a month." (Kassebaum, et al: 247). The training and supervision was exiguous, and poorly focused.⁴ There was no way of differentiating the men who thought they were getting something out of the group from those who thought they were not.

To borrow another phrase from Donald Campbell, this was not a "proud program." Replication of the administrative structure would be readily feasible, but the content was idiosyncratic with each counselor. Some counselors may well have been effective — there was no way of knowing. Some may have been actually destructive, and there was no way of knowing that, either. The appearance of disseminability was deceiving.

3. **Too Soon to Evaluate:** The California Men's Colony was a brand new institution, but the cultures that the guards and the prisoners brought with them were old, in many respects as old as the American prison. If the introduction of group counseling was to have an impact on the convict code or on the guards' stereotypes of their roles and the roles of the prisoners, a lot of time would have to pass. Suspicions, of which there were many, would have to be dissolved. The manipulation of group counseling by prisoners to impress the staff or the parole board would have to be offset; group participation should have been for the primary benefits intended,

⁴The CMC counselors were given a special course of training by a university psychologist consisting of a three-day didactic session followed by nine half-day sessions spaced at fortnightly intervals. It is not clear that there was any systematic training thereafter, but there were irregular contacts with the CMC group counseling supervisor.

but in many, perhaps in most cases, it never was. There can be no assurance that the program could ever be valued for the well-being of the participants, as Fenton intended. But if that outcome could ever be realised, years of group experience would have to erode the convict's "do your own number" culture. The program evaluated here lasted for two years, during which time counselors had to be accustomed to the counseling role and prisoners had to be accustomed to being counseled. There was never enough time to settle down.

What Should Have Been Expected?

In spite of the confusion, the woolly-minded theorising, the impossible expectations, and the slap-dash implementation, I think group counseling was a useful innovation in prison life, and I regret its demise at the hands of these expert evaluators. It was silly to suppose that a reduction in recidivism would result from the program, even if better administered than it was. What I regret is the opportunity to improve communications within the prison. The "solidary opposition" between guards and convicts, described so convincingly by Sykes, depends heavily on the stereotyped behavior dictated by the guard code as well as the convict code. Each side sees the other in terms of undifferentiated types, almost entirely negative. The notion that there could ever be *trust* between the two sides was the radical innovation proposed by Fenton, and well worth the test it never got. That the turmoil that infected other California prisons was never so severe at CMC may have been attributable to the change in prison culture. We shall never know. If the Department could have settled for a rapprochement between guards and prisoners — if indeed such a gain could have been achieved — realism might have won a small triumph. What is more, the group counseling movement might have continued and might have flourished elsewhere in luckier hands.

The Lesson

The Kassebaum-Ward-Wilner torpedo was only one of hundreds of evaluations which led to the "nothing works," destruction of the rehabilitative model of penology.⁵ The consequence of this imputation of futility has been generally demoralizing to persons working in prisons, probation, and parole. It is not too much to say that the gross underfunding of probation in many cities where probation is desperately needed and the withering of prison treatment programs almost everywhere can be blamed in large part on imprudent evaluations. Almost no industrial, educational or vocational

* ⁵Here it is necessary to cite the chaotic and almost unreadable compendium of failed penal programs collected by Douglas Lipton, Robert Martinson, and Judith Wild, *The Effectiveness of Correctional Treatment*, (New York, Praeger, 1975). The time frame for the study was January 1945 to December 1967. At the time of his death in 1978, Martinson was working on a second study of correctional treatment, having collected several hundred evaluations prepared after 1967. I understand that they led to a more optimistic view of penal treatment, but this study has not been brought to the point of publication.

training programs have been subjected to rapsheet evaluation, but the imputation of futility spreads even to them.

Innovative programs sometimes soared in the vigour of novelty with eager staff hoping to break new ground. Evaluation was required and the benefits were not clearly discernible, so programs came to an abrupt end. This trajectory from hope to disillusion is wholly unnecessary and, worse, irrational. It should not have been the fate of group counseling, and it need not be the model for assessment of any new program in prison management or treatment. The tyranny of the rapsheet must come to an end. In what follows, I shall discuss the preferable alternatives.

Lesson One: The Necessity for Theory: An innovative intervention is an independent variable whose existence depends on a theory. In the case of group counseling, the theory was loose, far too loose for systematic and "disseminable" implementation. There was no good reason to suppose that the vast culture changes called for by the first two hypotheses could have been accomplished within the two years of the project. There was even less reason, if possible, to believe that the impact of the counseling would carry over so powerfully as to show up in a significantly reduced rate of recidivism — the third hypothesis proposed for test.

What could be rationally expected of an experiment to test the administration of counseling by virtually untrained guards to prisoners accustomed to the ethos of the traditional prison culture? During the first two years, the stabilisation of the program, the elimination of unreliable counselors, and systematic, continuous training of adaptable employees would have been achievements enough. The measurements should have been limited to attendance and the responses to some very simple questionnaires.

Out of such an exploration, a theory might have emerged that would have indicated amplifications of the model that might have led to the selection of criteria by which early evaluations might have been made. I doubt that rapsheets would be theoretically appropriate at this stage. Measurements by other parameters would have to be tried and justified by the theory. Much harder work than pushing criminal histories through the computer, but a good deal more interesting, too.

The realisation of any innovation in penal programs must depend on a detectable change in the culture of the prison in which it is to be administered. The theory on which the innovation is based must specify the nature of the expected culture change and how that change will be achieved. If it is not achieved as expected, the assessor has a responsibility for discovering wherein the implementation failed or the theory was incomplete. The intellectual difficulties of such a task are formidable, but they must not be brushed aside on that account.

To sum up thus far, a program that is to be evaluated must be based on a theory, and it is the responsibility of the research team to require a rigorous statement of the theory before it prepares a design for the evaluation of the program. The criteria depend on reasoned expectations of the implementation at significant points in its development. Above all, the evaluation staff must think about what it sees and does. No evaluation should be performed by a research design taken off the shelf. Every program occurs in a different

setting in time and place and with different people involved in it in many different ways. All these differences make a difference in the application of the general theory of the intervention.

Lesson Two: The Necessity of Monitoring: A program must be stabilised before it can be evaluated. Note that in the CMC study, the turnover of counselors was high, perhaps as high as one change a month. It is not clear when this discovery was made, but confidence in the outcome of the study should have been sufficiently shaken to require a re-design of the evaluation. The program director should have been responsible for recording significant deviations from the program model and reporting them to the research team. Members of the research group should be alive to critical aspects of the program and conduct personal observations of its progress.

Time is always of the essence — in the case of evaluations of this kind, time must be open-ended. Too often decisions about the future of programs have to be made by the results of evaluations that are incomplete because there was not enough time to accumulate data for a stable, “disseminable” program. The program manager and the research team must agree on the stages of program development, the signs of completion, and the criteria appropriate for the evaluation of each stage. It simply isn’t enough to decide that two years will be enough — *rien ne va plus!*

Lesson Three: The Necessity for Discovery: Scientific endeavour is a sterile process if there are no additions to knowledge when an experiment is completed. It is a stultification of science to go to all the work that Kassebaum and his colleagues undertook and to arrive at no information more compelling than that an inefficient model of counseling failed of statistically positive results for three hypotheses. We are none the wiser about the impact of the independent variable on the institution, the attitudes of the staff toward it, or about what really went on in all that counseling. Here was a chance to burrow into the problems of treatment and culture changes resulting from attempts to administer it, but the focus on the rapsheet distracted attention from what would have been significantly more important.

* * *

Long ago, when I was chief of research for the Department of Corrections I was an innocent searching for sophisticated social scientists who could help us improve our methods of program evaluation, in which, even at that time, I had little confidence. One distinguished social scientist rebuffed me, saying that I was very much mistaken if I hoped to find any serious colleague in his department who would be willing to devote his time and energies to the evaluation of prison programs. Years later, federal and foundation money have changed all that; we have a lot of quite reputable scientists poring over rapsheets and printouts of rapsheet refinements and interpretations. The tyranny of the rapsheet continues, but understanding has not been increased, not by a jot or tittle.

* * *

References

- Conrad, John P. 1965. *Crime and its Correction*, Berkeley, The University of California Press.
- Fenton, Norman, 1961. *Group Counseling: A Preface To Its Use In Correctional and Welfare Agencies*, Sacramento, The Institute for the Study of Crime and Delinquency.
- Fogel, David and Hudson, Joe. 1981, *Justice As Fairness*, Cincinnati, Anderson.
- Kassebaum Gene; Ward, David; and Winer, Daniel; 1971. *Prison Treatment and Parole Survival: An Empirical Assessment*, New York, Wiley.
- "Programs, Experiments and Other Evaluations: An Interview With Donald Campbell," 1986. *The Canadian Journal of Program Evaluation*, Vol. 1, No. 1, April.
- Sykes, Gresham, 1958. *The Society of Captives*, Princeton, Princeton University Press.