“What works?” revisited: new findings on criminal rehabilitation

JAMES Q. WILSON

Few articles appearing in this magazine have been as widely reprinted or as frequently cited as Robert Martinson’s “What Works?—Questions and Answers About Prison Reform,” published in 1974. Its major conclusion has become familiar to almost everyone even casually interested in crime control programs: “With few and isolated exceptions, the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism.” For politicians as well as for scholars, the message seemed clear—nothing works. In fact, the article was careful to point out that there were, scattered through the 231 studies that were reviewed by Martinson and his co-workers, hints of some reductions in criminality for some kinds of offenders under some circumstances. But these hints did not constitute, even generously interpreted, a clear and consistent pattern of success on which a public policy might be based.

There was little new in the 1974 article. In 1967, R.G. Hood had published in Europe a review that concluded that different ways of treating offenders generally led to similar, and not very encouraging, results. A year earlier, Walter C. Bailey published in this country a survey of 100 evaluations of correctional treatment programs that led him to the judgement that “evidence supporting the
efficacy of correctional treatment is slight, inconsistent, and of questionable reliability.” Indeed, such gloomy findings go back at least 30 years. In 1951, Edwin Powers and Helen Witmer reported on the results of the ambitious Cambridge-Somerville Youth Study, begun in 1939 as an effort to prevent delinquency by an intensive counselling program. Despite high hopes, they had to conclude that, by any measure, boys randomly assigned to counselling were as likely as similar boys left on their own to run a foul of the law.

Unlike these earlier studies, the Martinson article—based on a massive volume he had prepared in collaboration with Douglas Lipton and Judith Wilks—created a sensation. Partly it was the times: It appeared in the early 1970’s, after the optimism of the Great Society had been dashed and the enthusiasms of the 1960’s had moderated, but when politicians were still searching desperately for some response to the widespread public fear of street crime. Martinson did not discover that rehabilitation was of little value in dealing with crime so much as he administered a highly visible coup de grâce. By bringing out into the open the long-standing scholarly skepticism about most rehabilitation programs, he prepared the way for a revival of an interest in the deterrent, incapacitative, and retributive purposes of the criminal justice system.

But it was not just the times. During the 1960’s, there had developed in California a remarkable concentration of talent and energy devoted to finding and testing rehabilitation programs, especially ones designed to treat delinquents in the community. Marguerite Q. Warren, Ted Palmer, and others not only used advanced psychological testing to classify delinquents by personality type and employed skilled counsellors to provide intensive community supervision, they randomly assigned delinquents to the treatment and control groups in order to insure the best possible scientific evaluation of the results.

At first, these results were encouraging, so much so that the President’s Commission on Law Enforcement and Administration of Justice, in its 1967 report to Lyndon Johnson, endorsed the Community Treatment Program (CTP) of the California Youth Authority, describing it as having reduced delinquency (as measured by parole revocation) from 52 percent among youth who were incarcerated before release to 28 percent among those given intensive counselling in the community.

The Martinson article was particularly critical of these claims. In their re-analysis of the California data, Lipton, Martinson, and Wilks concluded that Warren and her colleagues had substantially under-
counted the number of offenses committed by the youth in the experimental community program. Apparently, probation officers assigned to these delinquents developed such close relations with their charges, and were so eager to see their program succeed, that they failed to report to the authorities a number of offenses committed by the experimentals, whereas youth assigned to the control groups had their offenses reported in the normal way by probation and parole authorities.

**California counterattacks**

Given the resources devoted to the California project and the publicity it had received, it is hardly surprising that its leaders counterattacked. Ted Palmer published in 1975 a rebuttal to the Martinson article, claiming that it overlooked or downplayed a number of success stories in the rehabilitation literature and that in particular it misrepresented the CTP. Palmer conceded that the youth in the experimental program had a number of offenses overlooked by counsellors, but argued that these were largely minor or technical violations, many of which were detected simply because the youth were under closer observation and some of which involved merely the failure to participate regularly in the intensive supervision program. Moreover, the Martinson review ended in 1967; if it had continued through 1973, Palmer said, the differences between experimentals and controls, at least for serious offenses, would have been clear.

Martinson responded vigorously to this challenge, and the battle was joined. In the midst of the verbal pyrotechnics of Palmer and Martinson—they were nothing if not spirited adversaries—a new and, as it turned out, more weighty voice was heard. Paul Lerman, a Rutgers sociologist, published a book-length evaluation of the CTP (as well as of the California probation subsidy program) in which he concluded, after a painstaking analysis of the published data, that “the CTP did not have an impact on youth behavior that differed significantly from the impact of the control program.”

Moreover, the “community” focus of the experimental program turned out to be somewhat exaggerated—in fact, the great majority of experimental youth were placed in detention at least once and many were detained repeatedly in order to maintain control over them. Indeed, the youth in the experimental “community” program

---

were more likely to be sent to detention centers than the control group supervised by regular parole officers. Finally, Lerman found strong evidence that, though the CTP had tried to match experimental and control groups by randomly assigning youth to each, over the many years the program operated the two groups began to differ markedly in their characteristics as persons dropped out of the program for one reason or another. In particular, the experimental group came to be composed disproportionately of persons who were older, had higher IQ's, and were diagnosed as "neurotic" (rather than as "power-oriented"). This intriguing finding, largely buried in an appendix to the Lerman book, raises issues to which we shall return presently.

Lerman had made many of these points earlier, in a 1968 article; he made them more elaborately in the 1975 book. Curiously, Palmer, who continued to protest against the Martinson article, appears to have taken little notice, at least publicly, of the Lerman criticisms. Palmer's book-length attack on Martinson and his reassertion of the claims of the CTP appeared in 1978; there is no mention of Lerman in it.2

Enter the National Research Council

While the debate in correctional journals raged, the public view, insofar as one can assess it from editorials, political speeches, and legislative initiatives, was that Martinson was right. Because of this widespread belief that "nothing works," the National Research Council, the applied research arm of the prestigious National Academy of Sciences, created in 1977 a Panel on Research on Rehabilitative Techniques, chaired by Professor Lee Sechrest, then of the Department of Psychology of Florida State University. The Panel was charged with reviewing existing evaluations of rehabilitative efforts to see if they provided a basis for drawing any conclusions about the effectiveness of these efforts. Its first report—on efforts to rehabilitate in correctional institutions—was issued in 1979; a second report, on community rehabilitation, will appear later.

Owing to the importance in the public debate of the review by Lipton, Martinson, and Wilks (LMW), that book was made the focus of the Panel's attention.3 In addition, the report examined

reviews that analyzed studies appearing after 1968, the cutoff date for the LMW review. Among the papers commissioned by the Panel was a detailed re-analysis of a sample of the studies analyzed by LMW, carried out by two scholars not identified with the on-going debate, Stephen Fienberg and Patricia Grambsch.

The conclusion of the Panel is easily stated: By and large, Martinson and his colleagues were right. More exactly, "The Panel concludes that Lipton, Martinson, and Wilks were reasonably accurate and fair in their appraisal of the rehabilitation literature." If they erred at all, it was in being overly generous. They were sometimes guilty of an excessively lenient assessment of the methodology of a given study. Moreover, the evaluations published since 1968 provide little evidence to reverse this verdict. For example, David F. Greenberg's 1977 review of the more recent studies comes to essentially the same conclusion as Martinson. S.R. Brody's survey in England on the institutional treatment of juvenile offenders agrees.

The Panel looked in particular at Palmer's argument that nearly half the studies cited by Martinson showed a rehabilitative effect. The Panel was not persuaded: "Palmer's optimistic view cannot be supported, in large part because his assessment accepts at face value the claim of the original authors about effects they detected, and in too many instances those claims were wrong or were over-interpretations of data." In any event, "we find little support for the charge that positive findings were overlooked."

The conclusion that Martinson was right does not mean that he or anyone else has proved that "nothing works," only that nobody has proved that "something works." There is always the chance, as the Panel noted, that rehabilitative methods now in use but not tested would, if tested, show a beneficial effect and that new methods yet to be tried will prove efficacious. (One such method will be discussed in a moment.)

Are some offenders amenable to treatment?

One unresolved issue is whether certain kinds of offenders are more amenable to rehabilitation than others. If this is the case, and if the amenable subjects are mixed together in a treatment group with non-amenable ones, then any reductions in criminality among the former might be masked by increases in criminality among the latter; the average (and misleading) result would be, "no change."

This view has been vigorously advanced by Daniel Glaser, among others. Writing in 1973, a year before the Martinson article appeared,
he pointed to evidence from a variety of evaluative studies suggesting that certain kinds of offenders were especially amenable to rehabilitation. They tended to be persons who could easily communicate, who had not found their prior criminal career to be especially rewarding, and who had not been greatly disappointed by their efforts to find legitimate alternatives to crime. The CTP, for example, made explicit use of a psychological classification scheme designed to differentiate among delinquents on the basis of their “interpersonal maturity level” and their particular mode of behavior. One such group was classified as having a relatively high level of maturity, by which is meant the members had an internalized set of standards and some regard for the opinions of others, but displayed as well neurotic tendencies—either feelings of guilt and anxiety or a proclivity to “acting-out.” The interpretation of the CTP data by Glaser, Palmer, and others was that these anxious, neurotic, guilt-stricken delinquents benefitted substantially from intensive counselling. Recognizing the criticisms already levelled by Lerman (in his 1968 article) at the CTP, Glaser felt that even allowing for counsellor bias the neurotics did substantially better in the treatment groups than they did when left alone in the control groups. Moreover, Glaser has argued that Lerman himself neglected the long-term effects of treatment on different types of delinquents.

If this is true, then those studies which show no change among treated offenders may include not only some “amenables” who commit less crime but some non-amendables who actually commit more crimes as a result of the treatment. And this is exactly what Palmer believed he found in the CTP data. In his 1978 book, he showed the monthly arrest rates for two kinds of offenders—the “conflicted” (by which he apparently means “neurotic”) and the “power-oriented” (by which he seems to mean those delinquents who lack an internalized set of conventional standards and either manipulate others or identify with the norms of a deviant group). Neurotic delinquents in the treatment group had a lower monthly arrest rate than neurotics in the control group, both during the early stages of the program and four years after discharge. “Power-oriented” offenders in the treatment groups, on the other hand, had a higher arrest rate than the power-oriented controls four years after discharge. Glaser had surmised that this increase in criminality among power-oriented delinquents in the treatment program arises because they learn from it how to manipulate their counsellors, obtain favors, win early release, and generally “con” the system. “Treating”
such persons—at least by means of verbal therapy—apparently makes society worse off.

An earlier study by Stuart Adams provides some confirmation for this point of view. In 1961 he described the “Pilot Intensive Counselling Organizations” (PICO) in California, aimed at reducing delinquency among older juvenile offenders. The eligible youth were first classified as “amenable” or “non-amenable” by the persons running the project. (Exactly how they reached these judgments is not clear.) Once classified, they were then randomly assigned to either a treatment or control group. (The treatment consisted of individual counselling sessions, once or twice a week for about nine months, carried on inside a correctional institution.) After nearly three years of observation, Adams discovered that the amenable delinquents who had been treated were much less likely than amenable delinquents who had not been treated to be returned to custody. On the other hand, delinquents judged non-amenable who were given counselling did much worse—indeed, they were more likely to be returned to jail than the non-amenables who had not been treated. In short, if you are amenable, treatment may make you less criminal; if you are not, treatment can make you more criminal. Adams found that the delinquents judged to be amenable were “bright, verbal, and anxious.” These characteristics are similar to those of the neurotics in the CTP.

This conclusion is consistent with a good deal of evidence about the effects of psychotherapy generally. Changing delinquents is not fundamentally different from changing law-abiding people: “Crime,” after all, is not a unique form of behavior; it is simply behavior that is against the law. The illegality of the behavior is no trivial matter, but illegality alone does not differentiate one action from many similar actions. For example, many (perhaps most) offenders tend to do poorly in school, to have emotional problems, to find it difficult to get along with parents and friends, and to drink a good deal of alcohol. They are generally a mess. But poor school work, strained peer relations, emotional stress, and drinking liquor are not illegal.

Psychologists have long argued over whether any form of therapy will help any kind of problem. H.J. Eysenck, in a famous pair of articles published in 1952 and 1965, claimed that there was little evidence that therapy did anybody much good. He was (and is) the Robert Martinson of psychotherapy. Of late, psychologists have questioned the sweeping nature of Eysenck’s claim. Mary Lee Smith and Gene V. Glass published in 1977 a comprehensive review of
nearly 400 controlled evaluations of therapy and counselling and concluded that the client was often better off being treated than not. However, they noted that the improvements were generally with respect to such matters as fear and self-esteem and much less often with respect to such matters as "adjustment" (under which heading, of course, one finds most criminal behavior). Smith and Glass also tried to measure what factors made some subjects more amenable to therapy than others. They were able to identify two statistically significant ones: whether the therapist resembled the client, and the IQ of the client. Brighter clients did better than duller ones.

Similarly, if we are to believe Lerman, the brighter (and more neurotic) delinquents remain in the CTP program longer than those with the opposite characteristics; thus, any improvement measured by Palmer in their law-abidingness may result either from their greater receptivity to therapy, or from their tendency over time to outnumber the more delinquent-prone youth, or both.

The possibility that some persons are amenable to criminal rehabilitation is intriguing but it is not yet clear how much to make of it. The National Research Council Panel took note of the issue but remains skeptical that we have any clear understanding of it. The CTP, the major source of claims about amenability, is methodologically flawed. The PICO project did not define amenability with any rigor. Classifying a criminal as "amenable" may only mean that a therapist has a good hunch as to who will cooperate with the program. But if the therapist cannot communicate to others the basis for that hunch or provide a clear explanation of its rationale, it is hard to see how it can be used routinely as the basis for classifying and treating offenders. Moreover, some difficult legal and ethical issues arise. Suppose we are able to differentiate, accurately, amenable from non-amenable offenders. Suppose further that the treatment from which the amenable will benefit is less restrictive, more benign, and shorter in duration than the conventional punishment to which non-amenable will be assigned. Should we allow the criminal justice system to be "nicer" to "amenable" offenders than to non-amenable ones, even though their offenses and prior records may be identical. (Of course, it may also turn out that the rehabilitative program is felt by the recipients to be more onerous than doing "straight time"; the issue, however, remains the same.)

Nevertheless, the possibility of identifying amenable subjects and aiming programs at them that work is sufficiently attractive as to merit intensive new research. Someone has even coined a short-
hand term to describe what we now suspect are the amenable subjects of therapy: YAVIN (young, anxious, verbal, intelligent, neurotic).

**Recidivism, rates, and restrictiveness**

The most dramatic new argument in the continuing debate over rehabilitation, however, comes from two authors who do not, at first glance, appear to be writing about rehabilitation at all. Charles A. Murray and Louis A. Cox, Jr., members of a private research organization, were retained to find out what happens to chronic delinquents in Chicago who are confronted by sanctions of varying degrees of restrictiveness.4

The Chicago authorities wanted to know if any of the programs offered in that city—ranging from commitment to a conventional juvenile reformatory, to newer programs that left the delinquent in the community or sent him to a wilderness program—changed the rate at which delinquents committed offenses. Such studies have been done many times, usually with the negative results reported by Martinson. But Murray and Cox redefined the outcome measure in a way that seems to make a striking difference. Until now, almost all students of recidivism “rates” or rehabilitation outcomes have measured the success or failure of a person by whether or not he was arrested for a new offense (or was convicted of a new offense, or had his parole revoked) after leaving the institution or completing the therapeutic program. “Success” was an either-or proposition: If you did not (within a stated time period) get into trouble again, you were a success; if you did get into trouble—even once—you were a “failure.” Though the evaluators of rehabilitation programs typically speak of “recidivism rates,” in fact they do not mean “rate” at all—they mean “percent who fail.” More accurately, they use “rate” in the sense of “proportion,” as in the “birth rate” or the “tax rate.” But there is a different meaning of rate: the *frequency* of behavior per unit of time. Even a cursory glance through the studies reviewed by Lipton, Martinson, and Wilks reveals that almost all of them use “recidivism rate” to mean “the proportion who fail.”

It was Murray’s and Cox’s happy thought to use rate in the sense of frequency and to calculate how many arrests per month were

---

charged against a given group of delinquents before and after being exposed to Chicago juvenile treatment programs, and to do so separately for each kind of program involved. They examined three groups of youth.

The first was composed of 317 serious delinquents. They had been arrested an average of 13 times prior to being sent to the Department of Corrections, which was when Murray and Cox first started to track them. They were young—the average age was 16—but active: They had been charged with 14 homicides, 23 rapes, over 300 assaults and 300 auto thefts, nearly 200 armed robberies, and over 700 burglaries. The boys entered the study by having been sentenced by the court to a state correctional institution where they served an average of about ten months. Murray and Cox followed them for (on the average) 17 months after their release.

By the conventional measure of recidivism, the results were typically discouraging—82 percent were rearrested. But the frequency with which they were arrested during the follow-up period fell dramatically—the monthly arrest rate (i.e., arrests per month per 100 boys) declined by about two-thirds. To be exact, the members of this group of hard-core delinquents were arrested 6.3 times each during the year before being sent away but only 2.9 times each during the 17 months on the street after release.

The second group consisted of 266 delinquents who were eligible to go to a state reformatory but who instead were diverted to one of several less custodial programs run by the Unified Delinquency Intervention Services (UDIS), a Cook County (Chicago) agency created to make available in a coordinated fashion noninstitutional, community-based programs for serious delinquents. Though chosen for these presumably more therapeutic programs, the UDIS delinquents had criminal records almost as severe as those sent to the regular reformatories—an average of over 13 arrests per boy, of which eight were for "index" (i.e., serious) offenses, including nine homicides, over 500 burglaries, and over 100 armed robberies. Nonetheless, since these youth were specially selected for the community-based programs, one would expect that in the opinion of probation officers, and probably in fact, they represented somewhat less dangerous, perhaps more amenable delinquents.

Despite the fact the UDIS group may have been thought more amenable to treatment, the reduction in their monthly arrest rates was less than it had been for the group sent to the reformatories (about 17 percent less). In general, UDIS did not do as well as the regular Department of Corrections. Even more interesting, Murray
and Cox found that the more restrictive the degree of supervision practiced by UDIS, the greater the reduction in arrest rates. Youths left in their homes or sent to wilderness camps showed the least reduction (though some reduction nonetheless); those placed in group homes in the community showed a greater reduction; and those put into out-of-town group homes, intensive-care residential programs, or sent to regular reformatories showed the greatest reduction. If this is true, it implies that how strictly the youth were supervised, rather than what therapeutic programs were available, had the greatest effect on the recidivism rate.

Ordinarily, we do not refer to the crime-reduction effects of confinement as “rehabilitation.” Technically, they are called the results of “special deterrence” (“special” in the sense that the person deterred is the specific individual who is the object of the intervention, and not the general delinquent population). “Rehabilitation” usually refers to interventions that are “nice,” benevolent, or well-intended, or that involve the provision of special services. A psychologist might say that rehabilitation involves “positive reinforcements” (such as counselling) rather than “negative reinforcements” (such as incarceration). Indeed, the National Research Council Panel defines rehabilitation as the result of “any planned intervention” that reduces further criminal activity, “whether that reduction is mediated by personality, behavior, abilities, attitudes, values, or other factors,” provided only that one excludes the effects of fear or intimidation, the latter being regarded as sources of special deterrence.

Although the distinction has a certain emotional appeal, it makes little sense either scientifically or behaviorally. Scientifically, there is no difference between a positive and negative inducement; behavioral psychologists long ago established that the two kinds of reinforcements have comparable effects. (It is not generally true that rewards will change behavior more than punishments, or vice versa.) Behaviorally, it is not clear that a criminal can tell the difference between rehabilitation and special deterrence if each involves a comparable degree of restriction. Rehabilitation can (and usually does) involve a substantial degree of coercion, even of intimidation (“be nice or you won’t get out,” “talk to the counsellor or stay in your cell,” “join the group discussion or run the risk of being locked up”). Behavior-modification therapy can involve the simultaneous use of positive reinforcers (“follow the rules and earn a token”) and negative ones (“break the rules and lose a token”). It might help the discussion of offender-oriented programs if the
distinction between rehabilitation and special deterrence were collapsed.

**Two questions**

The real issue raised by the Murray-Cox study is not, however, what to call the effect they observe, but whether they have actually observed any effect at all. A number of criticisms have been made of it, but two are of special importance. First, does the decline in arrests indicate a decline in actual criminality or merely an increase in skill at avoiding apprehension? Second, if there is an actual decline in criminality, might this not be explained by the maturation of youth—that is, growing out of crime as they become older? Andrew C. Gordon, Richard McCleary, and their colleagues made these and other criticisms in response to a preliminary report of the Murray-Cox findings. In their later, book-length treatment of the Chicago project, Murray and Cox responded.

The second criticism seems the easiest to answer. Murray and Cox were able to show that the decline in rearrest rates existed for all incarcerated serious delinquents regardless of age. As an additional check, the authors examined a third group—nearly 1,500 youth born in Chicago in 1960 and arrested at least once by the Chicago Police Department before their 17th birthdays. Since this group was chosen at random from all arrested youth of the same age, it naturally is made up primarily of less serious offenders. Indeed, only 3 percent of this group was ever referred to UDIS or the Department of Corrections. When the monthly arrest rates for this group were examined, the data showed a more or less steady increase throughout the teenage years. Being arrested or being placed on probation had no apparent effect on subsequent delinquency. By all the tests they used, therefore, the decline in arrest rates for those delinquents given strict supervision cannot be explained by the fact that they were simply getting older.

The other criticism is harder to answer. Strictly speaking, it is impossible to know whether arrest data are a reasonable approximation of the true crime rate. No one argues, of course, that every crime results in an arrest. All that is at issue is whether a more or less constant fraction of all crimes result in arrests. There are two possibilities—either having been arrested before draws police attention to the boy (he is “stigmatized” or “labeled”), thus making him more likely to be arrested for subsequent crimes, or the arrest and subsequent punishment increases his skills at avoiding detec-
tion (the system has served as a “school for criminals”), thus making him less likely to be arrested for a given offense.

Now it is obvious that the first of these two possibilities—the “labeling” effect of being arrested—cannot be true, for, we have seen, delinquents who are placed under supervision have their subsequent arrest rates decline. If the police “pick on” previously arrested youth, they either do so without making an arrest (by keeping an eye on “troublemakers,” for example) or they try harder to arrest them but find the youth are not committing as many crimes as before.

The other possibility—that boys become skilled at avoiding arrest—is impossible to disprove, but Murray and Cox raise some serious questions as to whether this gain in skills, if it occurs at all, could explain the decline in arrest rates. Perhaps their most telling argument is this: One must not only believe that correctional institutions are “schools of crime,” one must believe they are such excellent schools that they produce a two-thirds gain in arrest-avoiding skills. This would make reformatories and group homes the most competent educational institutions in the country, since no one has yet shown that conventional schools, with the best available educational technology, can produce comparable gains in learning non-criminal skills. And all this must be accomplished within the ten-month period that is the average length of detention. It is still possible, of course, that the “schools of crime” hypothesis is true, but it requires one to make some heroic assumptions in order to sustain it: that large numbers of boys learn more during ten months in a reformatory than they learn in ten years on the street; that the great majority, despite their statements to the contrary made to interviewers, increase their commitment to crime as a way of life (rather than as an occasionally profitable activity) as a result of incarceration; and that the object of their efforts when back on the street is to employ their sharpened skills at avoiding apprehension while committing relatively unprofitable crimes rather than attacking more profitable (and riskier) targets.

Though Murray and Cox make a persuasive case for the validity of their findings, it cannot be taken as a conclusive study. For one thing, we would like to know what happens to these delinquents over a much longer period. Most studies of rehabilitation suggest that any favorable effects tend to be extinguished by the passage of time (though such extinction usually appears within the first year). We would also like to know more about the kinds of offenses for which these persons were arrested, before and after court interven-
tion (perhaps they change the form of their criminal behavior in important ways). And above all, we would like to see such a study repeated in other settings by other scholars. It may even be possible to do this retrospectively, with data already in existence but never before analyzed using frequency of offending (rather than proportion of failures) as the measure of outcome.

In fact, long before the Murray-Cox study, LaMar T. Empey and Maynard L. Erickson had reported on the Provo Experiment in Utah, an effort to reduce delinquency that was evaluated by arrest rates before and after treatment—the same outcome measure used by Murray and Cox (indeed, the latters' book contains a foreword by Empey). The Provo Experiment was, in principle at least, an even better test of changing recidivism rates than the Chicago project because the former, unlike the latter, randomly assigned delinquent boys to either treatment or control groups and kept detailed records (in addition to before-and-after measures) of what actually happened to the boys in the treatment programs. The experimental program was community-based, but, unlike conventional probation or even group homes, involved an intensive level of participation in a supervised group discussion program, absence from which was promptly penalized by being locked up. The program was in time killed by community opposition (many persons thought it excessively punitive, others quarreled over who should pay for it). The four years worth of data which could be gathered, however, indicate that there have been substantial reductions in arrest rates that cannot be explained by maturation or social class differences for all boys. This was true of both those incarcerated and those left in the community, with the greatest reductions occurring among boys in the experimental programs. Though open to criticism, the Provo data provide some support for the view that, if one measures offense frequency, some kinds of programs involving fairly high degrees of restrictiveness and supervision may make some difference.

Thoughts for the future

The Murray-Cox and the Empey-Erickson studies are important, not only because they employ rates rather than proportions as the outcome measure, or even because they suggest that something might work, but also because they suggest that the study of deterrence and the study of rehabilitation must be merged—that, at least
for a given individual, they are the same thing. Until now, the two issues have been kept separate. It is not hard to understand why: Welfare and probation agencies administer “rehabilitation,” the police and wardens administer “deterrence”; advocates of rehabilitation think of themselves as “tender-minded,” advocates of deterrence see themselves as “tough-minded”; rehabilitation supposedly cures the “causes” of crime, while deterrence deals only with the temptations to crime; psychologists study rehabilitation, economists study deterrence. If Murray-Cox and Empey-Erickson are correct, these distinctions are artificial, if not entirely empty.

The common core of both perspectives is, or ideally ought to be, an interest in explaining individual differences in the propensity to commit crime, or changes in a single individual’s propensity over time. The stimuli confronting an individual can rarely be partitioned neatly into things tending to produce pain and those likely to produce pleasure; most situations in which we place persons, including criminals, contain elements of both. If explaining individual differences is our object, then studying individuals should be our method. Studies that try to measure the effect on whole societies of marginal changes in aggregate factors (such as the probability of being imprisoned, or the unemployment rate) are probably nearing the end of the line—even the formidable statistical methodologies now available are unlikely to overcome the gross deficiencies in data that we shall always face.

Policy makers need not embrace the substantive conclusions of Murray and Cox (though it is hard to see how they could reasonably be ignored) to appreciate the need to encourage local jurisdictions to look at the effect of a given program on the rate of behavior of a given set of offenders. If they do, they may well discover, as Murray and Cox feel they have discovered in Chicago, that for the serious, chronic delinquent, the strategy of minimal intervention—probation, or loosely supervised life in the community—fails to produce any desirable changes (whether one calls those changes deterrence or rehabilitation), whereas tighter, more restrictive forms of supervision (whether in the community or in an institution) may produce some of these desired changes, or at the very least not produce worse delinquency through “labeling” or “stigmatization.” It is hard to imagine a reason for not pursuing this line of inquiry.