Received for publication: 27.10.2013. Review paper
UDK: 343.8

EVALUATION OF INTERVENTION PROGRAMMES IN PENOLOGY

Milko Mejovšek
University of Zagreb
Faculty of Education and Rehabilitation Sciences

SUMMARY

The main purposes of evaluation of intervention programmes are to check their effects and to improve them. Scientific checking of program effects in penology is usually performed using quasi-experimental design with two groups (treatment and control) in two time points: before (pretest) and after program application in treatment group (posttest). On the basis of obtained difference in criterion variable (in penology the most often recidivism) between treatment and control group in posttest the conclusion about program efficiency is made. The difference is in the most cases presented as the shift in common standard deviation (Cohen d) or as the correlation coefficient. Meta-analysis is often used in intervention programmes evaluation. Evaluations performed using treatment and control groups show only mean effects, but neglect individual differences. According to the principle of treatment individualisation, evaluation of programme effects in practical work with offenders should be individualized in the way that in defined time points planned and obtained results are compared, and when necessary some other modalities of program or a new program might be used. This could be done by cybernetic model.

Key words: *penology, intervention programme, evaluation, cybernetic model*

INTRODUCTION

In every day life many things are being evaluated (measured): for example, the quality of goods and services offered to general public, the quality of educational programmes and prevention and therapy programmes in the field of healthcare, and all with the view of ensuring quality and greater efficacy. In the field of education, healthcare, social welfare, employment, criminology and many others the intention is to guarantee quality of intervention programmes. Intervention programmes are different forms of systematic activity that attempt to solve the existing problem, i.e. improve an unsatisfactory situation in a certain field. In general, the society intervenes when something is not good, when something could go wrong (prevention programmes) or when it has already gone wrong (therapy and treatment programmes). We can take the fields of social pathology and criminality as examples. Prevention programmes target at children and youth in risk of developing violent behaviour, drug dependence and delinquent behaviour tendency. Treatment or therapy programmes target at those who have already become violent, addicted to drugs and delinquent. The same logics can be applied for the field of health care. Prevention programmes target at preserving health when it is at risk, and therapy programmes at treatment of illnesses when the damage or disorder have already occurred. In the field of welfare intervention programmes present different forms of social support and assistance for at risk population. In the field of employment intervention programmes target at fast professional retraining of unemployed population with the view of enabling their employment in professions that are at that point eligible on the labour market. In penology intervention programmes help offenders to change and successfully reintegrate into the society.

To what extent is evaluation a scientific problem, and to what extent is it a practical problem, is a question that experts do not completely agree on. According to one group of experts evaluation of intervention programmes should have its foundations in science and should be based on scientific research methods. Although it concerns practical work and is developed for the purpose of practical

work, the evaluation should in principle fulfil the prescribed criteria for scientific research. Criteria that are set for evaluation procedures form a tight connection between these procedures and applicable and fundamental research (Kulenović, 1996). Gal et al. (2003) still mention some differences between scientific and evaluation research: evaluation research is conducted with the view to make a strategic, i.e. political decision, and the goal of scientific research is to explain a certain phenomenon; evaluation research targets at specific programmes, and scientific at general scientific knowledge; evaluation research has the a measuring character, and scientific aims at identifying the essence of the phenomenon that is studied. This question reached its culmination in the "clash" of two renowned scientists, Campbell and Cronbach. The former advocated the thesis of scientific-based evaluation and the letter of practice-based evaluation.

From historical perspective, evaluation studies were first conducted in the field of education (at the beginning of the 20th century), and boomed after World War II. In the 70-ties of the past century in USA they become a separate field within social sciences, when programme evaluation started developing as profession. Associations of evaluators and agencies that offer professional services of evaluation were being founded and journals dealing with the topic of evaluation were being issued. Scientific and professional conferences were being held. However, the question of formal education of evaluators still remained. This is a profession where people of different educational backgrounds in the field of social sciences come together.

World-renowned scientists Donald Campbell and Lee Cronbach played an important role in the field of evaluation of intervention programmes. Campbell was an advocate of social experiments. Campbell (1969, 1991) writes about a society that experiments with social programmes. In these real experiments, as he states, big, randomly chosen groups of participants for the treatment and the control group would be formed, where the interference factors would vary at random and the groups would be practically completely equal (randomized field experiments). So, these experiments would have a high level of internal value. Cronbach (1982), however, considers that programme evaluation is more a skill and less science and that it should serve the function of making political decisions. Cronbach does not favour experiment but advocates the opinion that other research designs can be used depending on the purpose of evaluation and evaluation questions. He attaches more importance to external than internal validity (generalization of results). Although the initial positions of these two eminent scientists were pretty contrary, later a sort of compromise or convergence of views was reached. Alongside experiment and rigorous quantitative scientific methodology, Campbell accepts the application of qualitative methodology and when it questions quantitative results they should be examined. Cronbach on the other hand accepted the legitimacy of rigorous scientific methodology when it is necessary and possible to conduct (Cronbach, 1982; Rossi et al., 2004). Rossi et al. (2004) suggest that these opposing viewpoints need to connect, although it is not easy to do so. It is really about connecting science and practice. Scientific research in social sciences is complex and lengthy. The decision makers related to programmes want the information to be presented quickly and in a simple manner. The evaluators should therefore make sure that the information is scientifically based, and in choice of evaluation design they should think about evaluation research costs, speed that is expected and the fact that results need to be clear and understandable to those who need the information. A certain compromise is usually imposed, but one should not back away from the minimum requirements of scientific research.

The question is: Who is interested in evaluation of intervention programmes? In her answer to this question Arlena Fink (1995) lists the following interested beneficiaries of evaluation: government, programme authors, creators of policies in certain fields (government committees, institutes, agencies), programme financiers and scientists. It is interesting that Fink did not include on the list all those who are directly included in a programme. These are programme managers, performers and beneficiaries. The above listed parties should also be interested in intervention programme evaluation, although they sometimes and for different reasons oppose evaluation, especially those involved in programme execution.

An important question is: What are the reasons for conducting evaluation? Posavac and Carey (1989) mention the following: to obtain official quality certificate (accreditation), to enable funding, to be able to answer the questions about the programme (its quality), to enable programme selection, if there are more, to enable development and improvement of the existing programmes, to find out about unintended programme effects.

What are the main objectives of intervention programmes evaluation? Chelimsky (1977, according to Rossi et al., 2004) mentions three basic purposes

of evaluation: programme improvement, definition of programme value (efficacy) and enhancement of knowledge about the programme. The evaluation with the first mentioned objective is called formative evaluation. The evaluation of this type should be fast, concrete and useful for direct application. The evaluator is constantly communicating with the parties interested in evaluation, often in an informal way. The second type of evaluation is called summative evaluation. The goal of this evaluation is to make the final decision on the "faith"of the programme by determining programme efficacy. This type of evaluation should be carried out seriously, observing all principles of scientific research. The evaluator in this type of evaluation communicates with the interested parties formally, through evaluation reports. The third type of evaluation is meant for completely new programme models that are in the phase of development and checking. It is often performed on academic level and results are announced on scientific conferences and in scientific journals. This is in fact a scientific check of new approaches. In the summative evaluation an independent evaluator (which is not employed by the institution that runs the programme) is hired. In formative evaluation, which can be conducted also by a competent person employed by the institution conducting the programme, programme performers have a cooperative role (this is the so called participative evaluation). This role enables programme performers to contribute significantly to the improvement of the programme with their insights.

Writing about social problems, Shadish et al. (1991) point out that there are many problems in practice that make the implementation of evaluation difficult. An idealized, rational situation would be to have social problems clearly defined, potential solutions generated and some implemented in practice and evaluated and to acquire knowledge on successful solutions and disseminate it to decision makers. However, it does not function in practical life, because social problems are badly defined, the interested parties do not agree on priorities, programme goals are general, programme changes yield weak effects, or decision making on political level is diffuse. Besides, policy managers, programme managers and programme performers do not attach enough importance to evaluators' suggestions, and small and variable programme effects lead the evaluator into danger to make wrong negative conclusions about the programme.

The main activities in evaluation of an intervention programme concern finding the answers to the so called evaluation questions. These are important issues regarding programme evaluation.

There are some differences between authors regarding this subject. Fink (1995) lists the following evaluation questions: To what extent have the programme goals been reached? What kind of characteristics of persons and/or groups resulted from the programme? With which persons and/or groups has the programme been most effective? How long do the effects of the programme last? Which programme specificities (activities, situations, management styles) have proven to be most effective? To what extent are the goals and programme activities applicable to other persons and situations? What are the financial effects of the programme? To what extent have the changes in social, political and financial circumstances affected programme support and results?

Rossi et al. (2004) mention a smaller number of evaluation questions. These questions can be considered main evaluation questions:

1. What is the need for the programme?
2. Is the programme based on theory?
3. Is the programme executed in the way it was planned?
4. What are programme results?
5. What is the price of the programme?

These are actually key questions regarding evaluation and they form logical phases of a comprehensive programme evaluation. This means that programme evaluation should be conducted in this order. What follows is a more detailed explanation of these evaluation questions.

The first evaluation question refers to the condition or the status of potential programme beneficiaries that the programme would like to involve, and the definition of the need for the programme. The definition of the need for the programme contains the following: definition of a degree to which the population is endangered, type of intervention needed, its duration and how it should be provided. Here different sources can be used: statistical data, expert opinions and results of scientific and professional studies.

The second evaluation question concerns theoretical foundations of the programme. A good programme should have a theoretical concept: it should be clear what are the causes of the problem and how the problem can be solved. It is essential to know what is the reasoning, i.e. the logics of the programme and how do the activities of the programme help solve the existing problem. Programme evaluators should know what is the theoretical basis of the programme and should assess which theory is

good and to which extent it can serve in creation of the programme. In social sciences phenomena are extremely complex, and so are behaviour, personality and other characteristics of the beneficiary that the programme is trying to change. Very often it is not just one theory but more theories that explain a certain characteristic of the beneficiary that the programme is trying to change. The so called eclectic programmes are therefore not rare. These are programmes based on more theories. Programme author combines notions from more theories in a practical way, expecting the beneficiary to accept the programme well and the programme to yield results. In this case one can only guess about the causes of changes achieved by the programme. In this case it is difficult to propose programme improvements, because it is not completely clear what happens when the programme is applied. Accordingly, it would be good for every programme to have clear theoretical foundations.

The third evaluation question is programme implementation. This question regards programme execution and is in literature often called process programme evaluation. The main question that arises here is: Is the programme implemented in harmony with original ideas? In this context the control of programme implementation (monitoring) is important. This is where data on frequency and duration of the meetings between programme performer and programme beneficiary are gathered. Apart from that, the basic question asked here is the question of programme implementation quality. Do programme beneficiaries get the level and quality of the programme that was envisaged? Is everything that is important for the smooth running of the programme in place? Do the performers conduct prescribed tasks in a good-quality way? Are all necessary things, equipment and alike available? Are the beneficiaries satisfied with the activities conducted in the framework of the programme? The beneficiaries assess the satisfaction with the service provided and the satisfaction with programme performers. Programme implementation evaluation is very important, because once it is determined that the implementation programme is not yielding the desired effects, in searching for possible failure causes one should definitely re-analyse the gathered data on programme implementation mode. (Milas, 2005).

Within programme process control, one should gather data on programme implementation continuously, which in practice means on daily basis. It is the only way to reach a conclusion on the type of service that the beneficiaries were provided with and to what extent. It is impossible to evaluate the programme without these data. It is also important to know who are the beneficiaries of the programme. This is where the questions of programme appropriateness for the selected beneficiaries and the question of whether the programme meets the needs of its beneficiaries and to what extent arises?

It is important for programme implementation data gathering to be well-organized. Special forms can be used for that purpose or it is even better to organize data into previously prepared tables on personal computers. Data on every beneficiary and activities implemented with the beneficiary (provided services) are recorded. One can also record data on changes for which it was established that they had resulted from the programme.

Here we should also mention a question that often arises in the field of education, health care, social protection, pedagogy and re-education of children, youth and persons with deviant and delinquent behaviour. This is the question of system evaluation, i.e. its potentials and it regards the evaluation of conditions for programme implementation. The question considers the extent to which the system is organized to conduct the required activities and tasks and how the system functions. Are experts for programme implementation involved and is funding available? Do institutions, working concepts, professional standards and legal regulations in a certain field of work exist? Thus, the goal is to asses if all prerequisites for successful programme implementation are in place. These assessments are carried out by experts in different fields who in case it is necessary propose taking certain actions the goal of which would be to remove the defined shortcomings and achieve positive changes. One such comprehensive evaluation of the functioning of the system of institutional social care for children and youth was conducted by Žižak and Koller-Trbović (1999).

The fourth and most important evaluation question regards programme results evaluation and includes achievement of programme goals. Have the planned results been achieved, i.e. did the planned changes regarding programme beneficiaries occur? For example, has the delinquency rate in local community been reduced? This question refers to standards or success criteria and helps with making conclusions about the degree of programme efficacy. The standards or criteria of efficacy should be precise (Fink, 1995). The standard can be defined by the percentage of the desired improvement that has been achieved, through expert assessment or as statisti-

cally significant difference between the respondent group in which the programme has been implemented (experimental or treatment group) and the group where it has not been implemented (control group). In quantitative evaluation standards are numerically defined. For example, for juvenile delinquents the measures of success can be: reoffending, but also regular school attendance and success at school. Differences in recidivism, regular school attendance and success at school between the treatment and the control group of respondents are compared.

The most common form of evaluation occurs after programme termination. Data on programme results are gathered once the programme is officially completed. For example, after the release of prisoners after they had served their sentence, data on recidivism are gathered in postpenal period, which should indicate the efficacy of the penological treatment. These data can be compared to recidivism rate recorded for a certain type of institution, or comparisons can be made on the basis of crime types or characteristics of the perpetrator. In this case we are talking about a nonexperimental approach. Or, if a number of inmates was included in a special programme, in postpenal period data on recidivism are gathered for those who have participated in the programme (treatment group) and for the group that has not been included in the programme (control group). In this case we are talking about an experimental, i.e. quasi-experimental approach.

Apart from evaluation conducted after programme completion, the programme can also be evaluated during its implementation. This implies programme results evaluation during programme implementation. Programme evaluation conducted during programme implementation has one important advantage over programme evaluation after the programme is completed, which is the possibility to change the programme if it is not yielding desired results. In this type of evaluation feedback on programme effects is most important. The programme can be implemented in groups or individually, but evaluation of results should in the first place be individual. This type of intervention programmes can be conducted in practice according to cybernetics principles.

The collection of quantitative data on effects of a programme can be difficult and time-consuming. The main advantage of qualitative data on personal experiences of beneficiaries regarding programme effects is that this type of data can be collected faster and can sooner be made available to programme decision makers (Posavac and Carey, 1989). Qualitative evaluation is useful because one can analyse things that are difficult to grasp in quantitative evaluation, such as, as it was mentioned, personal experiences of beneficiaries regarding their participation in the programme. Patton (2002) states that qualitative research focuses on programme design and the way programme is executed, while quantitative research focuses on programme effects evaluation. Qualitative evaluation concentrates more on formative evaluation when the goal is to develop a programme, and quantitative concentrates on summative evaluation when the decision has to be made on whether the programme should continue or it should be terminated. Qualitative evaluation uses interviews to research into personal experiences of beneficiaries and observe programme activities with the view of making amendments and improvements. Qualitative evaluation can be used to find out many details which are difficult or impossible to grasp in a quantitative way. Quantitative evaluation is prevailingly group evaluation and it consists of testing the importance of difference between the treatment and the control group according to average group results, and qualitative evaluation is more individualized and contains case descriptions. Patton (2002) advocates detailed case descriptions where the beneficiary reports on programme effects, contacts with performers, performers' efforts to include other experts, contacts between performers and family members and on his or her own personal progress. Programme performer should have access to description of beneficiary's life before the programme, beneficiary's reactions to the programme and information about the life of the beneficiary after the programme. The evaluator also needs these data to be able to assess the programme as objectively as possible. Data of this sort would improve the individualized approach to beneficiaries and make it possible to meet their needs in best possible way. Data of qualitative type can present an important supplement to quantitative data. Unfortunately, qualitative approach in programme evaluation is not used sufficiently in penology. Similar situation can be found in case of individual case design which is used to quantitatively evaluate programme efficacy for individuals.

The fifth evaluation question links programme efficacy to programme costs. The question that is asked here can for example be: have the funds that have been invested into the prevention programme targeting at children and youth in risk been returned to the society and individuals through reduced damage caused by delinquent behaviour of children and

youth? The achieved results are compared with the programme costs. There are two possibilities here:

1. financial savings achieved through the programme (cost-benefit analysis)
2. the ratio of price and programme efficacy (cost-effectiveness analysis).

The first question is how much money is saved by including children and youth in local community into delinquency prevention programmes. The savings are calculated by deducting the programme price from the price of damage that children and youth could do if they had not been included into the prevention programme. The second question is what are the costs of a certain programme that can achieve certain results. The comparison of price of two programmes that can yield similar or different results falls under this question too. The fifth evaluation question is complex because it is difficult to calculate the exact price for the majority of intervention programmes, and an even bigger problem is to calculate savings that would be made by a certain programme. When this evaluation question is dealt with within programme evaluation, the evaluation becomes complex, and if it is dealt with separately, it requires two evaluations - programme and financial effects evaluation (Fink, 1995).

Out of five evaluation questions the most important is the fourth question, which regards programme results, i.e. programme goals realization. The evaluation of achieved results can be: group or individual. Group evaluation is conducted in programme beneficiary groups. Three approaches can be applied here: nonexperimental, quazi-experimental and experimental. Nonexperimental approach consists of collection of data on programme results for only one group. This is the group that participated in the programme. There are two possibilities in this case: testing after programme implementation and testing before and after programme implementation. If we have data on programme results after the programme has been implemented, we can only get information on whether the achieved results are satisfactory or not in comparison to expectations and achievement standards that were defined. The second approach which makes the comparison of the situation regarding beneficiary's behaviour before and after the programme possible, is better. Accordingly, it is possible to draw conclusions on possible improvements influenced by the programme. In order for the conclusion on programme effects to be more complete and precise, it is necessary to have the control group of respondents which has not been involved in the programme subject to evaluation. However, if different kinds of measuring were made on one group of clients before and after programme implementation, it is possible to make relatively reliable conclusions on programme efficacy. This type of design can be categorized as quasi-experimental design. To conclude, the lack of control group can be compensated with a larger number of measurements, which in turn allows for more reliable conclusions on programme efficacy.

Experimental and quasi-experimental approach in programme evaluation are in principle used apart from treatment also for control group of respondents. Treatment group respondents have been exposed to the programme subject to evaluation and respondents of control group have not been exposed to that programme (they have either not been exposed to any kind of programme or they have been exposed to standard programme which is conducted in a standard way). The advantage of the experimental design is that it is possible to make conclusions about the influence of the independent variable (programme) on the dependent (criterion) variable (treatment effects, changes in behaviour) with the greatest certainty. Only a well-planned and conducted experiment allows for cause-effect conclusions to be made. Quasi-experimental approach is more often conducted in natural environment which often does not allow for observance of strict control standards that are applied for experiments. It is mainly a problem of inclusion of respondents into the treatment and control group and limited possibilities of equalization of groups, and the problem of inequality of treatment and control groups of respondents presents the most important shortcoming of quasiexperiment. In experiment the choice of respondents is random, and than in the pretest phase the treatment and control group are made equal in dependent variable and variables that could beside the independent variable influence the dependent variable (moderator variable). Nonexperimental and quasi-experimental approach are used in penology for intervention programmes evaluation. In penitentiaries, clinics, schools and other institutions it is a problem to form completely equivalent experimental and control groups regarding all important characteristics, and especially when it interferes with the activities conducted in these institutions. Kazadin (1980) provides a vivid description of the situation: In clinical work, especially in clinical institutions, researchers are not in position to move patients in order to meet the requirements of a real experiment, they are forced to work within administrative, bureaucratic and sometimes even antiexperimental frameworks.

There are a number of statistical procedures used to assess the efficacy of intervention programme which is implemented in the treatment group of respondents: t-test, variance analysis, point-biserial correlation coefficient, phi correlation coefficient, proportions ratio and percentages comparison. Point-biserial correlation coefficient can be used to show success of a programme when the results in the dependent variable are continuous, and the respondents' group membership is a discontinuous variable. Phi correlation coefficient can be used to express success of a programme when both variables are discontinuous. In statistics this coefficient is called biserial correlation coefficient. Here the dependent variable consists of two categories: successful and unsuccessful. When all participants of the programme are successful and all non-participants are unsuccessful the correlation is maximum and positive (+1,0). When the situation is completely opposite, the correlation is maximum and negative (-1,0), and when control and treatment groups are equal in the independent variable after the completion of the intervention programme, the correlation is zero (0,0). Since this is a 2x2 frequencies table chi-squared test can be calculated and it can be used to obtain the phi-coefficient. In this table rows are treatment and control group of respondents and columns are success and failure in the dependent variable (e.g. in penology recidivists and non-recidivists). The ratio of proportions is the ratio of success in the treatment group shown as the proportion of relationship of success and failure divided by the ratio of success in the control group. This can be used to express the number of times by which success in the treatment group is bigger in comparison to the control group. Comparison of percentages can also be used to express success of the programme. For example, how high was the percentage of recidivism in juvenile delinquents included in the programme in the six month after the completion of the programme in comparison to the recidivism percentage of juvenile delinquents who were not included in the programme.

EVALUATION OF INTERVENTION PROGRAMMES IN PENOLOGY

Efficacy of intervention programmes in penology is not easy to check. There are a number of factors that should be controlled (especially variable moderator) and which can easily get out of control in nonexperimental and quasi-experimental approaches.

In penology a great number of evaluations of institutional and community-based treatments of youth and adults have been carried out. Very often the general ("official") treatment programme, the sort of which is conducted in different forms of community-based and institutional treatments in nonexperimental way, without control groups of respondents, was evaluated. The goal of such evaluations was to gather data via the so called efficacy variables which test the level of integration into social community (recidivism, employment, education, social and pathological behaviour forms, relationship with family, and membership in deviant groups) in postpenal period. Conclusions on efficacy of certain treatment type were made based on these data. The majority of these studies identified modest treatment effects.

Comparisons of institutional and community-based treatments identified that institutional treatment was in principle less effective. There are a number of reasons for it. Offenders who have committed more serious delicts are placed in institutions. Recidivists are more common among that type of offenders. It has been proven that these offenders were growing up in unfavourable circumstances. They become offenders at younger age. Apart from the mentioned, institutional placement has a number of disadvantages. These are: isolation from the natural environment, concentration of serious offenders, persons prone to violent and socio-pathological behaviour, exploitation and abuse of the weak, persons who are bad role models for others confined in a small place, as well as different restrictions and deprivations, bad accommodation and crowdedness. These are all the reasons why institutional placement is avoided when ever there is a possibility of imposing a community-based sanction. Today, for adults it is most commonly probation.

What were the reasons for modest effects of penology treatment? In the first place it was the general value and non-standardized value of treatment programmes which were not adjusted to special needs of offenders, weak motivation of offenders for active participation in the treatment, lack of expertise of the treatment staff, insufficient material conditions, errors in classification of offenders and unsystematic tracking of changes brought about by the treatment.

In the past thirty years, the situation has become slightly better, because specific programmes tailored for specific characteristics of offenders have been developed and evaluations of these programmes have been conducted with better quality, where alongside treatment there were also control groups of respondents included and the most modern statistical and mathematical methods of data processing such as meta-analysis have been introduced.

The seriousness of the problem can be proven by negative results of some of evaluation studies in the past. For example, prevention programme known as Cambridge-Somerville, a programme meant for high-risk boys in Boston proved to be a complete failure in a check conducted thirty years after programme completion. In a great number of criterion (dependent) variables the control group did better than the treatment group (according to Kulenović, 1996). Among evaluations of penology programmes, a negative evaluation by Robert Martinson (1974) received wide publicity. Unfavourable situation regarding intervention programmes did not only exist in penology. In psychology literature Eysenck's opinion on inefficacy of psychoanalytical theory of neurotic disorders (Eysenck, 1996) is often cited. Eysenck, basing his statement on a number of evaluation studies, says that roughly speaking two-thirds of neurotic patients show improvement within two years after the symptoms first occurred without any therapy.

A very thorough evaluation of penology treatment programmes was conducted by Robert Martinson and his associates (1974), who concluded, under the influence of repressive climate in American justice system at the time, in a systematic overview of studies on treatment programme efficacy that were available to him, that with very few exceptions, great majority of treatment programmes did not have any effects. These were the 70-ties when there was a culmination of dissatisfaction with effects of correction programmes in the justice system. Especially severe criticism of correction programmes was published by Robert Martinson (1974) in an article titled: "What works? – questions and answers about prison reform" in which he gives very negative assessments of treatment effects after analysing a greater number of studies that evaluated different treatment types. A phrase "nothing works" has ever since been frequently in use.

This is a very detailed analysis of carefully selected studies on treatment efficacy. A total of 231 studies which involved the control group of respondents, originated from USA and other countries and were published in English in the period between 1945 and 1967 were selected. Studies that did not have the control group were not used, neither were those where the treatment procedure was not clearly defined, respondents samples too small, where data on treatment effects were not gathered in a reliable way and those where treatment effects were under the influence of external factors. Martinson conducted this analysis with his associates for the Government's committee for combating crime in the state of New York and his study consisted of 1400 pages. Apart from recidivism, other criteria of treatment efficacy (dependent variables) were selected:

1. adaptation to correctional institution
2. success in vocational training
3. success in education
4. changes in personality and attitudes and
5. level of integration into the wider social community.

The article Martinson published in 1974 presented results of evaluation of different types of institutional and community-based treatments only for the dependent variable of recidivism, which is the basic criterion variable for assessment of treatment programme efficacy for juvenile and adult offenders. The following was analysed:

1. educational and vocational programmes for juveniles and adults
2. individual and group counselling
3. supportive environment programmes
4. medical programmes
5. length of sentence and level of institution security
6. community-based programmes.

According to author's opinion the results were devastating. With the exception of a few community-based treatment programmes, the effects of treatment programmes were practically negligible. It should be noted here that within the covered period some treatment programmes were not applied that later on proved the best, first of all this refers to cognitive-behavioural and multimodal treatment programmes (types of treatments that apply various treatment forms, i.e. treatment approaches). At the end of the article, Martinson recommends the return to retributional sentence model. One thing that should be noted separately is ambiguity, i.e. contradiction of results even in these carefully selected studies in the area of same programme types. While in some studies results were better in treatment groups of respondents, in others they did not do better or they were even worse in relation to control groups of respondents. This shows that many of the selected studies were burdened with serious methodological flaws. It should be mentioned that methodology shortcomings, although smaller, are still a serious problem in this type of research even nowadays. It should be noted here that it is not an easy job to organize and implement evaluation research with good-quality methodol-

ogy. The main problem lies in making the treatment and the control group equal in pretest (before the programme is applied to the treatment group) in dependent variable (variables) and all other variables that could apart from the independent variable (programme that is being evaluated) influence the dependent variable. These are moderator variables that are a big problem in evaluation studies. Since in quasi-experimental design treatment and control groups are not completely equal, it is definitely an unfavourable circumstance which makes the control of the influence of moderator variables difficult.

Andrews and Bonta (2006) mention three main objections to Martinson's article:

1. studies with negative conclusions were accepted without remarks
2. studies with positive conclusions were exposed to criticism in terms of clarity and reliability of the criteria for assessment of efficacy, clarity of research and theoretical foundations of the programme
3. the mentioned reasons could have reduced programme effects and not necessarily increased them (for example, reduced reliability of criteria for programme efficacy assessment can reduce, rather than increase the effects).

Different illogicalities can be found in Robert Martinson's article. Conceptually equal programmes have proven to be successful as well as unsuccessful. The author himself warns about it, stating that offenders differ greatly; what works for one group does not work for the other. Also, the question of stability of programme effects is unclear. Here, the author himself was not sure about how long the programme effects should last for the programme to be considered successful. The main illogicality, if we can call it that way, is author's negativistic approach, which is mentioned as the first objection by Andrews and Bonta, in which he approaches all evaluations with positive conclusions with criticism, looking for faults which are sometimes assumptions, rather than facts. However, Martinson is right in case of a number of programmes when he warns that different circumstances and influences on results attributed to those programmes are not controlled, like for example influences of other types of programmes that offenders participated in (he probably referred to moderator variables here).

Around that time, in former Yugoslavia a comprehensive evaluation of treatment programmes efficacy that juvenile delinquents had been exposed to was carried out and titled: "Efficacy of criminal sanctions for juvenile delinquents with special reference to recidivism in juveniles." Total sample of respondents consisted of 1342 juvenile delinquents that were imposed the following sanctions: reprimand, disciplinary centre, increased parents' supervision, increased supervision by guardians, educational institution, correctional facility and juvenile prison. However, the check of treatment efficacy in postpenal period was conducted for a part of the total sample of 459 respondents. Data were collected using following measuring instruments:

1. B-series (Z. Bujas)
2. Battery SVPN-1 (M. Reuchlin and E. Vallin; adaptation by A. Matić et al.)
3. Battery SVPN-2 (M. Reuchlin and E. Valin; adaptation by A. Matić et al.)
4. Revised beta series (C.E. Kellogg et al.)
5. Battery Gvertos (I. Ignjatović et al.)
6. Battery 18 PF (K. Momirović)
7. Battery 16 PF (R.B. Cattell)
8. Battery MPI (H.J. Eysenck)
9. Subordination scale (S) (V. Kovačević)
10. Battery ACK (M. Mraković)
11. Political conservatism scale Besk (BES) (D. Radovanović and Lj. Stojić)
12. Efficacy variables (VE)
13. Sociology variables (SV).

Efficacy variables (VE) include data on recidivism, socio-pathological behaviour, relation to own appearance, success at work or school, attitudes towards family, attitudes towards deviant groups and social activities. Data on social position of respondents were collected using the instrument of Sociology Variable (SV). The goal was to use this instrument to collect data on conditions that juvenile delinquents grew up in and data on postpenal period.

The main topics are education and/or employment (e.g. number of completed grades, changing schools or work organizations), juvenile's family (e.g. family structure, family relationships, financial status), parental reactions, (e.g. they advise juveniles, scold them, beat them), social care centre interventions (e.g. financial assistance, assistance in continuation of education or employment) and environmental conditions (e.g. crime rate, existence of deviant groups, organized gathering of youth in the place of residence). In the analysis of treatment effects, beside the Efficacy Variable and Sociology Variable data on cognitive and conative characteristics of respondents were used. A summarized overview of some papers from the project follows.

Kovačević et al. (1974) analysed relations between imposed sanctions and behaviour in postpenal period. Tracking time was between 12 and 14 months. The results showed that more severe sanctions are followed by a higher recidivism rate. Thus, a stronger social intervention was followed by recidivism by juveniles. The same was determined for socio-pathological forms of behaviour. The same tendency was determined in case of employment, i.e. that juveniles are employed less after institutional treatment; those who are employed are less satisfied with their job and less successful. Also, after institutional treatment, when compared to community-based treatment, the respondents have worse relationships with their families. Permanent affiliation with deviant groups is almost proportional to the severity of the sanction. Involvement in social and sporting organizations was lower than before the treatment. A general conclusion is that efficacy of sanctions, especially institutional sanctions, is low.

Hošek et al. (1974) interpreted the connection between imposed sanctions and social characteristics of juveniles in postpenal period. In general, it can be concluded that more severe sanctions are followed by more unfavourable social characteristics. In the first place it can be seen in possibilities or lack of possibilities of further education or employment. Furthermore, these are also bad relationships in the family and insufficient support. Connected to that is also low financial status and cultural level of the family and existence of socio-pathological phenomena. The more severe the sanction, the more obvious are unfavourable family conditions. Finally, one should add the lowefficacy of social welfare centres, which did not react in a number of cases, when help was needed.

Mejovšek et al. (1974) analyzed the connection between the variables of efficacy and sociological variables in postpenal period. As expected, unfavourable sociological characteristics are connected to weaker effects in efficacy variables. Unfavourable conditions in the area of education and/or employment, bad relationships in the family and other disadvantages connected to family, are followed by a higher likelihood of recidivism, failure at school and/or work, socio-pathological forms of behaviour and affiliation to deviant groups.

Momirović et al. (1974) interpreted the relations between cognitive and conative characteristics of respondents and efficacy variables. The results showed that persons with increased cognitive capabilities and persons who do not suffer from personality disorders are more successful at integration into a social community. It can thereby be concluded that conative characteristics are of somewhat greater importance. It especially refers to tendency for aggressive forms of behaviour, which is the most serious counterindication for successful social integration.

From the papers described it can be concluded that success or failure are not only a consequence of treatment but also psychological and sociological characteristics of juveniles after treatment.

A bit later, evaluation of institutional treatment of juveniles was carried out on the territory of Croatia. The research project was titled: "Test of success of institutional treatment for juveniles on the territory of the Socialist Republic of Croatia". A sample of 628 respondents of both sexes, who were released from educational and correctional institutions between 1972 and 1975, was selected. The postpenal period was between 3,5 and 7,5 years. The age of respondents at the moment the research was conducted varied between 18 and 29 years. The following measuring instruments were applied in the research:

1. General information about respondent
2. Efficacy variables
3. Treatment variables
4. Demographic, social and economic variables
5. Superego scale from Cattell's personality questionnaire 16 PF
6. Authoritarianism scale (Eysenck scale modified by M. Mraković).

Efficacy Variables and Demographic, Social and Economic variables were constructed after the model of Efficacy Variables (EV) and Sociological Variables (SV) from the project previously described. Treatment variables refer to: education and professional education during treatment, conditions in which educational activities took place, structure of staff in the treatment, optional activities, contacts between respondents and parents and respondents and welfare staff.

A summarized overview of some papers from the project follows.

Uzelac (1982) analysed efficacy of resocialization in postpenal period according to the type of institution and respondent's sex. The results of this analysis confirm the findings of the previous research, according to which the efficacy of the treatment decreases as sanctions become more severe. Here a comparison between educational and correctional institutions is made. In a number

of variables treatment effects are less favourable in case of correctional institutions: higher recidivism rate, shorter time period between release from institution and recidivism, more frequent school drop out, higher unemployment rate, increased presence of aggressive behaviour, vagrancy and gambling. The success of resocialization is greater with female respondents, and especially low recidivism rate in postpenal period is highlighted.

Bujanović Pastuović and Bašić (1982) analysed the link between efficacy variables and treatment variables. The results of their analysis showed that there is a link between absence of recidivism, regular school or work attendance, school or work satisfaction, satisfaction of others with respondents' school or work, satisfaction of respondents with institutional placement and orderly appearance of respondents, satisfaction with a smaller number of residents in educational groups and no changing between educational groups, satisfaction with teachers as educators, vocational training in harmony with respondents' choice of profession in a better learning environment and with better equipment as well as contacts with family through mutual visits by parents and juveniles.

Mejovšek (1982) interpreted the correlation between efficacy variables and demographic, social and economic variables. The main characteristics of a successful resocialization, which is manifested in absence of criminal and offensive activities, success at workplace or continuation of education, good level of integration into family and secondary social groups, absence of socio-pathological phenomena and avoidance of deviant groups, are positively associated with good conditions in primary social environment, which means higher social, educational, economic and cultural status of parents and other family members, positive relations in primary social environment and absence of deviant groups. In this way, this project too yielded the same results as in the previously described project. Results show, that if the social environment in which the juvenile lived after the treatment (and probably before the treatment) is more favourable, the probability that resocialization would be successful in postpenal period is higher. This piece of information, however, brings treatment efficacy into question.

Mejovšek and Kovačević (1982) analysed the relationship between the efficacy of resocialization in the postpenal period and variables of authoritarianism and superego. It is expected that well-socialized persons would have a higher level of superego and authoritarianism, where the increased superego is more desirable than authoritarianism, because it includes the components of rationality, humaneness and flexibility. Results show that successfully resocialized juveniles have authoritarian attitudes that are better pronounced than attitudes based on the power of the superego. A conclusion has been made that institutional treatment is more directed towards authoritarian attitudes and less towards attitudes based on the power of the superego.

Žižak (1982), as part of the same project, analysed the connection between treatment variables and juveniles' relation towards family in postpenal period. A canonical correlation analysis was conducted which showed that on the level of canonical factors there was no significant connection. In two canonical correlation analyses that included treatment variables (Bujanović Pastuović and Bašić, 1982 and Žižak, 1982), a low canonical correlation was obtained in the first case, and it was absent in the second case. In canonical analyses that did not include treatment variables, canonical links were significant and significantly higher. These findings could be interpreted as being against the treatment effects, or the way these effects were measured. Specifically, it is well-known that when a single variable is not reliably measured its correlation with other variables is underestimated.

Analysis conducted by Kovačević and Mejvošek (1985) on the bases of the same project shows that treatment effect is questionable. A hierarchical factor analysis with all variables together (a total of 186 variables) was performed. At the highest level of generalization (in the third row) three orthogonal (independent) factors were obtained, which were interpreted as prosocial behaviour in postpenal period, general factor of wider range and two factors of narrower range, which refer to educational and pedagogical work during institutional treatment. Prosocial behaviour is defined by abandonment of delinquent activity, success at school/workplace, avoidance of deviant groups, avoidance of social and pathological forms of behaviour, acceptance of social standards on authoritarian level and partly on superego level, favourable circumstances for passive social status, and exploitation of the potential for development of active social status. Respondent's passive social status is the one that is secured by his or her family, in the first place parents, and active social status is the one that respondent creates independently, through own activity. Two factors that refer to institutional treatment show that pedagogical and educational activities are not a single process, and that institutional treatment

has no or minimum influence on the behaviour of the educated party in the postpenal period.

Evaluation of general ("official") treatment programmes for juvenile delinquents in the two described projects was conducted according to nonexperimental design, which means without control groups of respondents. In the above described papers that resulted from these evaluation projects we see the effect of moderator variables, a thing that is usually neglected in evaluation studies. Moderator variable influences the dependent variable (in penology most often recidivism) apart from the independent variable (intervention programme). Anastasi (1968) most commonly states the following as moderator variables:

1. sex
2. age
3. educational level
4. social and economic status
5. interests and
6. motivation.

From the short description of papers from two evaluation projects, it is evident that the following appear as moderator variables: sex, educational level, social and economic status, cognitive abilities, conative characteristics and social attitudes of respondents. Influence of moderator variables should be excluded from the results of evaluation studies in order to get "pure" effects of the evaluated programme. This can in principle be done in two ways. One is that treatment and control group are equal in all potential moderator variables. The other way is statistical, where potential moderator variables are neutralized by partialization. Of course, under condition that data on potential moderator variables are gathered. The risk of moderator variables in evaluation studies can be removed only through careful planning of evaluation studies, according to experimental, i.e. quasi-experimental design, where potential moderator variables and ways of neutralizing their effects should be considered.

Antonija Žižak (2001) made an interesting attempt to evaluate institutional treatment of children and youth with behaviour disorders using the Psychoeducational model that was proposed by Brendtro and Ness (1983, according to Žižak, 2001). This is an eclectic model that connects educational activities with a variety of psychological processes. The model contains six dimensions, i.e. it is based on six guidelines, which form the bases of a good treatment programme: importance of interpersonal relationships, contextuality of evaluation, integrality of behaviour, humane component of learning, crises means opportunity and practice is pragmatic. The main goal was actually to construct instruments according to the model which would serve for collection of data on treatment and evaluation of treatment effects. The paper is interesting for two reasons; first is a question of theoretical models of treatment of offenders and second is a design of good-quality instruments for evaluation of treatment programme effects. The question of theoretical model, i.e. theoretical models for individual groups of offenders has even in recent times not been definitely resolved (Polaschek, 2012), and as for good-quality evaluation instruments the situation is not much better, because it is related to the first question.

Meta-analysis plays an important role in evaluation research. By applying statistical and mathematical methods in meta-analysis the efficacy of intervention programmes that have the same purpose is assessed. Of special importance is the comparison of programmes with different efficacy, which can help detect the characteristics of successful programmes. Firstly, we will describe four well-known, classical meta-analyses of programmes, that were meant for offenders, and secondly we will describe two more recent meta-analyses by a group of Canadian psychologists led by Andrews (Andrews, Dowden and Gendreau, 1999, according to Andrews and Bonta, 2006; Andrews and Dowden, 2005). Andrews and associates carried out a number of meta-analyses of efficacy of penology programmes. They are renowned for their Risk-Need-Responsivity Model of offender rehabilitation, which they have been developing for more than twenty years and which is at the moment one of most valued theoretical models in the field of offender rehabilitation.

Meta-analysis uses a statistical parameter of effect size. It is the difference between treatment and control group in the field of common standard deviation (Cohen's d) or correlation coefficient as criterion of success of the intervention programme in the treatment group (described in the introductory part of the text).

Whitehead and Lab (1989) conducted a strict selection in order to choose 50 studies that dealt with evaluation of community-based and institutional treatment programmes for juvenile delinquents. These studies involved both the treatment and the control group of respondents, and the treatment programme was clearly described. According to the assessment of meta-analysis' authors the results

were devastating. It should be mentioned here that the authors had set a relatively strict requirement for the phi correlation coefficient to be at least 0,20 for the programme to be considered effective, which is maybe too high a number for programmes of this sort. Institutional treatment, as expected, had proven to be considerably worse than community-based treatment. Treatment programmes targeting at change of behaviour (programmes where prosocial behaviour was encouraged) had not proven to be better than the rest, although these were the expectations based on previous studies. The authors detected lower treatment efficacy for studies where the choice of respondents for the treatment and control group was random, and higher for studies where the choice was not random, which can be interpreted by stating that partiality in selection of treatment and control group could influence the results. Whitehead and Lab are a good example of authors who "have set the bar too high", which consequently leads them to pessimistic conclusions. Besides, in primary studies selected for meta-analysis, cognitive-behavioural programmes were not included, which were proven to be most successful (Andrews et al., 1990).

Andrews et al. (1990) carried out meta-analysis on 154 treatment evaluations conducted on samples of juvenile and adult delinquents. This meta-analysis is one of the first papers that published the theoretical model these Canadian authors advocate, known as Risk-Need-Responsivity, which are the three fundamental principles of offender rehabilitation model, which was checked in a large number of studies. According to that model, the most intensive treatment should be provided to offenders in highest risk of re-committing the offence, the treatment should be directed to criminogenic needs, dynamic factors that direct offenders to committing crimes (e.g. antisocial attitudes and orientation, socialising with persons from criminal milieu, antisocial personality, drug abuse) and the treatment should be adapted to learning styles and needs of offenders. The third principle is general and resembles much the principle of individualization or differentiation of treatments. A form of behaviour therapy and/or cognitive behaviour therapy that best suits individuals should be selected. This should be governed by behaviour and social learning principles, interpersonal influence, development of skills and cognitive change (restructuring). One should thereby use modelling, gradation, practicing, role-playing and confirmation, and provide necessary resources and detailed verbal guidance and explanations. A clinical approach should be used with offenders.

In this meta-analysis, 30 studies dealt only with court procedures and decisions without treatment (e.g. warning, court surveillance, probation, imprisonment). The comparison of these 30 studies and 124 studies in which different treatment forms were evaluated, showed that greater effects in reduction of recidivism were accomplished in the latter studies, where treatment programmes were evaluated. In the meta-analysis of 124 studies which evaluated treatment programmes, the authors started from three basic principles of their theoretical model of rehabilitation, according to which the programmes should be directed to offenders in higher risk, their criminogenic needs and adapted to their learning styles. The results of meta-analysis confirmed all three principles of the theoretical model. Of the total of 54 programmes that satisfied the principles, in 38 programmes the phi coefficient of correlation was at least 0,20, and the average phi coefficient of correlation was 0,30. Just like in the previous meta-analysis the institutional treatment had proven to be less successful. In institutional environment even the good programmes adjusted well to the mentioned principles were less effective. The programmes that did not meet these principles yielded very bad results in institutions. According to the authors, the negativities of the institutional environment weaken even the effects of best-designed treatment programmes. Unlike Whitehead and Lab, Andrews and others are more optimistic about the treatment and consider that the effects of the treatment programme exist. The programmes that meet all three principles of the described model have shown noteworthy effects, which amounted to around 30% less recidivism in treatment groups in comparison to control groups of respondents.

Lipsey (1992) conducted a very comprehensive meta-analysis of efficacy evaluation of treatment programmes for juvenile delinquents, which included more than four hundred evaluations. Lipsey did not limit himself only to the published, but he also included the unpublished papers. The results of meta-analysis showed reduction in recidivism in treatment groups by 10% on average. The best treatment programmes were the ones directed to acquisition of skills (e.g. communication) and control and modification of behaviour. These programmes showed recidivism reduction by at least 20%. Punishing methods of coercion and intimidation in institutions showed considerable increase in recidivism compared with the control group (around 25%).

Lipsey was also interested in how treatment programmes affect other variables of treatment efficacy (dependent or criterion variables): attitudes and

personality traits, interpersonal adaptation, school attendance, success at school and professional education. The results showed that the effect of treatment programmes is mostly more pronounced in these variables, than in recidivism reduction and it is most pronounced in variables of attitudes and personality traits. Accordingly, criminal recidivism does not necessarily have to mean that the treatment programmes are a failure and one should not conclude about treatment programme success only on the basis of criminal recidivism. Criminal recidivism can be a consequence of specific situation or unfavourable circumstances, however the treatment programme has nevertheless brought about positive changes, which is often ignored.

Just like Andrews and others, Lipsey too has an optimistic attitude to effects of treatment programmes. According to these authors the real question is not whether penology treatment programmes are effective or not, or necessary or not, but which programmes and for which groups of offenders yield better or weaker results?

Antonowicz and Ross (1994) studied the literature on evaluation studies in penology and singled out a number of components they assumed could play an important role in programme efficacy. After that, they conducted a strict selection and chose 44 treatment programme evaluations that were methodologically correct, had the control group and where the dependent variable was recidivism. In 20 evaluations the treatment programme was effective, which means that the treatment group had achieved significantly better results in the dependent variable than the control group of respondents (tested with chi-square test). The authors used these 20 evaluations in which the treatment programmes had proven to be successful and the 24 evaluations in which the programmes had proven to be unsuccessful, to check the selected components. Each component was tested to determine the frequency of the programme with the mentioned component for the successful and unsuccessful programmes. The test of significance of differences was performed through chi-square test. The results showed that in only six components there is a significantly higher prevalence of programmes with these components among successful programmes, than among unsuccessful programmes. The six components were as follows: theoretical foundations of the programme on cognitive-behavioural model, multifaceted programming (variety of treatment programmes), focus on criminogenic needs, responsivity i.e. adaptation of the programme to the learning styles and capabilities of offenders (two principles taken over from the theoretical model of Andrews and associates), role-playing/modelling and training in socio-cognitive skills. In interpreting the results, the authors state that when developing treatment programmes one should definitely start from a theoretical model of delinquent behaviour or crime theory. The authors believe that the most appropriate model for explaining delinquent behaviour is the cognitive-behavioural model, according to which delinquent behaviour occurs because of the wrong way of thinking, weak or non-existing behaviour control, undeveloped social skills and wrong habits. As offenders are very heterogeneous population, treatment programmes that are diverse, that is, including a larger number of different procedures and methods, should be more effective. The authors support the opinion of Andrews and his associates on the importance of directing the treatment programme to criminogenic needs of offenders and adjustment to of the programme to learning styles, capabilities and other characteristics of offenders. Role-playing and modelling are important, because they allow practice of prosocial behaviour models and formation of desirable social habits. The same applies for the training of social and cognitive skills. These programmes should develop new ways of problem solving, perception and interpretation of events in different social situations, as well as influence the establishing of better behaviour control.

Andrews and Bonta (2006) state in their book "The Psychology of Criminal Conduct", which was published in several editions, that at Carleton University in Canada there is a database in which data on evaluation studies of intervention programmes are entered cumulatively. They mention that almost 400 primary evaluation studies have already been collected. This database was used for a number of meta-analyses. In meta-analysis conducted with 374 primary evaluation studies that was performed by Andrews, Dowden and Gendreau (1999, according to Andrews and Bonta, 2006) an average effect size of 0,08 was established and expressed as biserial (phi) coefficient of correlation, which points to the average of 8% less recidivism in the treatment group. Average percentages for recidivism were also calculated and amounted to 46% of recidivism in treatment groups and 54% of recidivism in control groups. Although the average value is not high, it still shows that the treatment is effective. However, what causes more concern is a very high efficacy variability, i.e. inefficacy of the treatment programme which is between -0,43 and

+0,83 (expressed through the values of the biserial coefficients of correlation). This very high variability without a doubt includes the methodological shortcomings of individual evaluations and not only differences in quality of intervention programmes. This also includes the earlier mentioned differences in psychological and sociological characteristics of juveniles and adults in postpenal period (moderator variables that mask the real effects of intervention programmes). It is also well-known that recidivism does not necessarily have to mean that the intervention programme is unsuccessful, but it can also be a consequence of unfavourable circumstances.

From the described meta-analyses one can notice that there are considerable differences in programme efficacy and that one can single out elements that make up a good programme. By analyzing the sizes of different coefficients one could conclude that effects of treatment programme meant for offenders are not remarkable. However, if we start from the point of view that delinquency is a complex and serious social problem, then even tiny improvements could be considered success. It can be concluded from the above mentioned that absolute and unambiguous success criteria do not exist, but that the final assessment to a certain degree depends on the complexity and seriousness of the social problem one is trying to solve. The author of this article believes that greater effects of intervention programmes could be achieved trough individualized approach and permanent control of the effects accomplished on the bases of feed back from the cybernetic model (Mejvošek, 1986, 1998).

Andrews and Dowden (2005) conducted meta-analysis which partly supports the above mentioned and in which the main subject is integratedness of penology programmes. Under good programme integratedness the authors understand good-quality programme management and monitoring of direct programme effects, theoretical foundations of the programme, programme implementation through well-trained practitioners who possess the skills of interpersonal work, supervision of clinical type, programmes that have manuals, programme that have an adequate (sufficient) duration, recent programmes, programmes conducted in small treatment groups and those in which evaluation is conducted. Meta-analysis was conducted on 273 primary studies, the goal of which was to check the efficacy of different penology programmes. Results show that those programmes which have been based on Risk-Need-Responsivity model and which contain the mentioned components of programme integratedness accomplish best effects in fighting recidivism. The problem is only that a relatively small number of primary studies have data on programme integratedness. Although the Risk-Need-Responsivity model by Andrews and associates is considered to be one of the best theoretical offender rehabilitation models, criticism is also present. Polaschek (2012), apart from lauding the model, also states that the third model principle, the principle of responsivity, is not developed enough, it is too general and does not say anything about how you can motivate the offenders to get actively involved in the treatment programme and make them give up the delinquent behaviour. Besides, the model has not yet responded to the needs of practice in clearly designed programmes for individual groups of offenders. The model is general and does not offer solutions for specific situations and design of concrete programmes. Accordingly, the theoretical model is in general terms acceptable, and it is expected from the authors to give recommendations for the design of specific programmes for the needs of practice.

In cybernetic model of penology treatment (Mejvošek, 1986, 1998) the analysis of feedback about the effects of treatment is undertaken at several transition control points (control points in timeline). The analysis of feedback in transitive treatment points is conducted for every inmate and for every characteristic that is the subject of the treatment. Efficient monitoring of the treatment process requires a great deal of information about each inmate and these pieces of information have to be processed and analysed quickly, for the treatment continues and can not be stopped for them to be studied in peace and slowly. This problem can only be solved with the use of personal computers.

The cybernetic model of penological treatment is a regulatory system which uses cybernetic principles in an attempt to change different characteristics of inmates, from socially undesirable and unacceptable to socially desirable and acceptable. The regulation functions on the principle of negative feedback loop; it attempts to reduce the difference between the initial situation and the desired situation, i.e. between the initial and final state of characteristics that are the object of change, under the influence of the treatment.

In human organism a number of physiology functions are regulated after the principles of cybernetics with negative feedback loop. In the system with the negative feedback loop the difference between initial, undesirable state of organism and desirable (normal) state of organism is diminished.

For example, if there is an increased level of carbon dioxide in organism, different physiological mechanisms that are supposed to excrete the excess of carbon dioxide from the organism are activated. As the level of carbon dioxide decreases the activity of these mechanisms gradually weakens, until the point when normal level of carbon dioxide in organism is established, when the activity ceases completely. Blood pressure, level of sugar in blood and other physiological processes and organism conditions are regulated in organism in the same way. At the system exit there is a sensory device (sensor) which uses the loop to send information about the momentary state of characteristics that need to be regulated back to the system entrance. If the difference between the existing and desired state is above the acceptable difference level, the regulatory system is activated with the task of bringing the differencewithin tolerable level.

In cybernetic model of penological treatment the initial state (input) is the state of characteristics that will be changed in the treatment, at the beginning of the treatment. The desired state (output) is the desired, planned state of these characteristics at the end of the treatment. As the treatment evolves, the difference between the initial and final state should be reduced, if the treatment programme is effective.

It is necessary to collect data about the characteristics that will be the object of the treatment before the treatment beginning. The level at which the characteristics of inmates are at the beginning of the treatment is the starting point, initial point, or baseline for the future evaluation of the treatment programme effects. The overall treatment period is divided into several three-month intervals (or other time intervals). Thus, between the initial and final point of the treatment programme there are several transition points in three-month intervals. Maybe the three-month intervals are too long. It all depends on the expected speed of changes. When we consider personality and behaviour characteristics which require more time, the three-month intervals are appropriate. If we consider behaviour characteristics where changes happen faster it could be a one-month interval.

To illustrate the possible assessment variables two examples can be used. The example from the penitentiary for the variables which were used to evaluate adaptation to the penological treatment (according to Mejovšek, 1992):

1. work performance
2. amount of fallout at workplace
3. handling machines, tools and equipment
4. relationship with other inmates at workplace
5. relationship with immediate superior officials at workplace
6. success in professional development
7. success in leisure activities
8. relationship with immediate superior officials in leisure activities
9. attitude to equipment and material in leisure activities
10. activity in inmate self-government
11. position in educational group (integratedness)
12. rewards and benefits
13. disciplinary measures
14. specialist interventions (psychiatrist, physician or psychologist)
15. overall assessment of behaviour in the last month in regard to the previous month.

Example of treatment areas from a part of the Questionnaire for evaluation of institutional treatment for youth with behavioural disorders (according to Žižak et al., 2001):

1. motivation for positive changes
2. self-control of behaviour
3. self-image
4. relationship with parents
5. relationship with brothers and sisters – wider family
6. relationship with educators
7. relation to institution – professional staff and social group climate
8. relation to assets – material assets
9. relation to non-material values
10. attitude to work and learning
11. attitude to other important activities
12. relation to institution – status, rules, organization of life
13. attitude towards own future
14. attitude towards other juveniles
15. development of habits and interests
16. development of social and communication skills
17. relation to local community.

Changes in behaviour of inmates and juveniles can be assessed by social pedagogists, pedagogists and social workers, and other professionals working in penitentiaries or institutions for juvenile offenders (including those running the programme outside correctional institutions, in case of community-based sanctions) after a training. Psychologists will be more involved in studying personality changes, using standardized measuring instruments. Here the data about inmates collected by the Diagnostics Centre, Prison Administration and the Ministry of Justice of the

Republic of Croatia should be used. All data about inmates should be unified into one system, which needs to be constantly updated, even in the postpenal period, in order to obtain information about the duration of the accomplished changes and to help former inmates with integration into the wider social community. For juveniles data from juvenile courts and social care centres should be used.

How can we define the level of risk, i.e. need for treatment for an inmate? There are good measuring instruments for classification of inmates according to the level of risk and need for treatment. LSI-R, the Level of Service Inventory-Revised by Andrews and Bonte (1995) is a good example of good quality measuring instrument of the kind. This instrument is meant for risk assessment and need for treatment for adult offenders. Such and similar measuring instruments, appropriate for classification of inmates according to risk level and need for treatment, can also be used for assessment of treatment programmes. For assessment of risk and need for treatment for juveniles an instrument of similar concept is YLS/CMI, the Youth Level of Service/ Case Management Inventory (Hoge and Andrews, 1994) and its updated version (Hoge, Andrews and Leshield, 2002).

Changes in personality or behaviour can also be established through tests, questionnaires, assessment scales and systematic observation. LSI-R and YLS/SMI, as well as other instruments of similar purpose can be used not only for classification of inmates and youth according to the risk level and need for treatment, but also as an instrument of measuring changes under the influence of the treatment, i.e. reduction of risk and the need for treatment. Data collection and analysis should be carried out in every transition point of the treatment.

Test of efficacy of intervention programmes in penology is in principle conducted in groups. Hereby, individual is neglected and average effects are obtained. If we start from the basic requirement of treatment individualization in penology (Mejovšek, 1989), and a similar requirement applies for non-penological programmes too, the evaluation of programme effects should be individual. Although programmes in penology are mostly conducted in groups, their evaluation in practice should be individualized. Scientific, quantitative programme evaluation is for statistical reasons conducted in groups, but in practical work with persons involved in intervention programmes, it should be individualized, so the best possible effects for each individual person are achieved. This can be achieved in different modalities of individual programmes or through inclusion of individual persons or smaller groups into special programmes, if the applied programme is not yielding the desired individual effects. The evaluation of effects of intervention programme on individual level should be conducted during intervention treatment, in order to enable timely reaction, in case it is established that the desired changes are not occurring. According to cybernetic model (Mejovšek, 1986, 1998) at certain points in time the planned and achieved results are compared for every individual and changes in programme are implemented, if necessary. This implies the introduction of alternative programme modalities, if those exist, or introduction of new programmes.

Today, in the era of comprehensive informatization, it is astonishing that information technology is not used much in implementation and evaluation of penological intervention programmes. Individual data are stored electronically, starting from the baseline data to the final data obtained at the end of the programme (or even later). Electronic recordings enable not only data storage, but also different statistical elaborations of the collected data. This approach raises the quality of work of programme implementers, because they have data on success of the programme for every beneficiary and allow for the programme changes when desired effects are not accomplished. A systematic data collection about treatment programme effects can enhance the objectivity of rewarding those who were successful in treatment and thereby enhance their motivation to persevere. As regards programme evaluation, this approach is of great assistance to programme evaluators, because they have access to data on programme effects for every individual beneficiary.

How to get access to data in individual monitoring? An approach that should not present serious problems to programme implementers in practice is the assessment of defined characteristics (or their measuring with standardized measuring instruments whenever that is possible) at defined points in time, as was described earlier. If necessary, data comparison can be made not only in regard to the pervious point in time, but also in regard to the baseline. The collected data can be statistically processed, for the needs of programme implementation, but also for the needs of programme evaluation.

REFERENCES

Anastasi, A. (1968): Psychological testing (third edition). New York: MacMillan.

Andrews, D.A., Zinger, I., Hoge, R.D., Bonta, J., Gendreau, P., Cullen, F.T. (1990): Does correctional treatment work? A clinically relevent and psychologically informed meta-analysis. Criminology, 28, 369-404.

Andrews, D.A., Bonta, J. (1995): LSI-R The level of service inventory-revised. Manual. North Tonawanda: Multi-Health Systems.

Andrews, D.A., Dowden, C. (2005): Managing correctional treatment for reduced recidivism: A meta-analytic review of programme integrity. Legal and criminological psychology, 10, 173-187.

Andrews, D.A., Bonta, J. (2006): The psychology of criminal conduct (fourth edition). Cincinnati: Anderson.

Antonowitz, D.H., Ross, R.R. (1994): Essential components of successful rehabilitation programs for offenders. International journal of offender therapy and comparative criminology, 38, 97-104.

Bujanović Pastuović, R., Bašić, J. (1982): Povezanost efikasnosti resocijalizacije i zavodskog tretmana maloljetnika s poremećajima u ponašanju u SR Hrvatskoj. Defektologija, 18, 145-151.

Campbell, D.T. (1969): Reforms as experiments. American Psychologist, 24, 409-429.

Campbell, D.T. (1991): Methods for the experimenting society. Evaluation Practice, 12, 223-260.

Cronbach, L.J. (1982): Designing evaluations of educational and social programs. San Francisco: Jossey-Bass.

Eysenck, H.J. (1966): The effects of psychotherapy. New York: The International Science Press.

Fink, A. (1995): Evaluation for education and psychology. Thousand Oaks: Sage.

Gall, M.D., Gall, J.P., and Borg, W.R. (2003): Educational research (seventh edition). Boston: Pearson Education.

Hoge, R.D., Andrews, D.A. (1994): The Youth Level of Service/Case Management Inventory. Description and evaluation. Otawa: Carleton University.

Hoge R.D., Andrews, D.A., Leschield, A.W. (2002): The Youth Level of Service/Case Management Inventory. Toronto: Mental Health System.

Hošek, A., Petrović, K., Momirović, K. (1974): Neke relacije između sankcija izrečenih maloljetnim počiniocima krivičnih djela i njihovih sociooloških karakteristika u postpenalnom razdoblju. Defektologija, 10, 49-92.

Kazdin, A.E. (1980): Research design in clinical psychology. New York: Harper and Row.

Kovačević, V., Singer, M., Momirović, K. (1974): Relacije između sankcija izrečenih maloljetnim izvršiocima krivičnih djela i njihovog ponašanja u postpenalnom razdoblju. Defektologija, 10, 4-48.

Kovačević, V., Mejovšek, M. (1985): Uspješnost resocijalizacije nakon zavodskog tretmana. Defektologija, 21, 1-24.

Kulenović, A. (1996): Evaluacija psihosocijalnih intervencija. U: Stres, trauma, oporavak (ur. J. Pregrad), 269-291. Zagreb: Društvo za psihološku pomoć.

Lipsey, M.W. (1992): The effect of treatment on juvenile delinquents: Results from meta-analysis. In: Psychology and law. International perspectives (eds. F. Losel, D. Bender, T. Bliesener), 131-143. Berlin: Walter de Gruyter.

Martison, R. (1974): What works? – questions and answers about prison reform. The public interest, 35, 22-54.

Mejovšek, M., Horga, S., Momirović, K. (1974): Kanoničke relacije sociooloških karakteristika i efikasnosti resocijalizacije u postpenalnom razdoblju. Defektologija, 10, 93-117.

Mejovšek, M. (1982): Povezanost efikasnosti resocijalizacije i socijalnog statusa maloljetnika nakon zavodskog tretmana. Defektologija, 18, 123-133.

Mejovšek, M., Kovačević, V. (1982): Povezanost efikasnosti resocijalizacije s nekim osnovnim socijalnim stavovima maloljetnika s delinkventnim ponašanjem. Defektologija, 18, 135-143.

Mejovšek, M. (1986): Kibernetički model penološkog tretmana. Penološke teme, 1, 145-152.

Mejovšek, M. (1989): Evaluacija institucionalnog penološkog tretmana. Penološke teme, 4, 1-7.

Mejovšek, M. (1992): Povezanost kognitivnih i konativnih karakteristika osuđenih osoba i adaptacija na penološki tretman. Penološke teme, 7, 71-75.

Mejovšek, M. (1998): Kibernetički model penološkog tretmana. Hrvatska revija za rehabilitacijska istraživanja, 34, 31-37.

Milas, G. (2005): Istraživačke metode u psihologiji i drugim društvenim znanostima. Jastrebarsko: Naklada Slap.

Momirović, K., Viskić Štalec, N., Mejovšek, M. (1974): Relacije kognitivnih i konativnih karakteristika maloljetnih delinkvenata i efikasnost resocijalizacije nakon penalnog tretmana. Defektologija, 10, 155-173.

Patton, M.Q. (2002): Qualitative research and evaluation methods (third edition). Thousand Oaks: Sage.

Polaschek, D. L. L. (2012): An appraisal of the risk-need-responsivity (RNR) model of offender rehabilitation and its application in correctional treatment. Legal and criminological psychology, 17, 1-17.

Posavac, E.J., Carey, R.G. (1989): Program evaluation. Methods and case studies (third edition). Englewood Cliffs: Prentice Hall.

Rossi, P.H., Lipsey, M.W., Freeman, H.E. (2004): Evaluation. A systematic approach (seventh edition). Thousand Oaks: Sage.

Shadish, W.R., Cook, T.D., Leviton, L.C. (1991): Foundations of program evaluation. Newbury Park: Sage.

Uzelac, S. (1982): Uspješnost resocijalizacije maloljetnih delinkvenata u različitim odgojnim institucijama SR Hrvatske. Defektologija, 18, 161-167.

Whitehead, J.T., Lab, S.P. (1989): A meta-analysis of juvenile correctional treatment. Journal of research in crime and delinquency, 26, 276-295.

Žižak, A. (1982): Odnos maloljetnih delinkveneta prema porodici nakon zavodskog tretmana. Defektologija, 18, 153-159.

Žižak, A., Koller Trbović, N. (1999): Odgoj i tretman u institucijama socijalne skrbi. Zagreb: Edukacijsko-rehabilitacijski fakultet.

Žižak, A., Koller Trbović, N., Lebedina Manzoni, M. (2001): Od rizika do intervencije (prilozi). Zagreb: Edukacijsko-rehabilitacijski fakultet.

Žižak, A. (2001): Evaluacija institucionalnog tretmana. U: Žižak, A., Koller-Trbović, N., Lebedina-Manzoni, M. (Ur.): Od rizika do intervencije, 117-153, Zagreb. Edukacijsko-rehabilitacijski fakultet.