

Jacek Woźniak
Olsztyn's College of Computer Sciences and Management
3, Artyleryjska Street, 10-165 Olsztyn, Poland
jkwozniak@tlen.pl

ON ROE AS A BETTER TRAINING MANAGEMENT TOOL¹

Abstract

The article describes arguments used in the discussion between proponents of ROI, and those who prefer relating to training goals („return on expectations”, the so-called ROE). It discusses consistency between the training practice and the directives of the ROI model. Several attempts at managing trainings basing on ROE are described. The similarities between research practice – a consequence of these two schools’ methodological postulates – are shown. We sketch a theory, which can be a basis for ROE training management, which features a scientific backing and nonrandomness in assessing training results equalling that of ROI’s. We show that the source of differences between these two approaches lies in a different concept of business training rather than measurement procedures (their being scientific, or their reliability), but i.

Key words: training, effectiveness of training, organizational development, ROI, soft-skills training.

Introduction

It is often said that trainings – like every other business activity – should be effective. It is also said that trainings should be related to the expectations of both the trainees and their organization. These two statements are slogans for two different approaches to training management – a school which would justify whether a training has sense with the use of indices of a project’s financial effectiveness (e.g. ROI), and a school which searches for other ways of ascertaining, whether the training was worthwhile. This second school of thought is currently less prominent in the literature on the subject, mainly because there is no widely accepted theory of the whole training process.

The goal of this paper is to report the arguments presented by both approaches. We will show that the weaknesses of the ROE approach are smaller than authors who argue for ROI for trainings would suggest. We argue that in the practice of research, the differences between these two approaches fade away. We will show that it is possible to quantify the effects of trainings and that the methodological requirements of ROE analyses are within the scientific standards accepted by social sciences. We will sketch the missing theory, which can be a basis for training management by ROE, although a full description of the theory, and the model of training management it is based on, is described elsewhere (Woźniak, in preparation).

The paper is organized as follows. First, we will present the theoretical framework of

¹ This is a changed version of a text delivered for the 5th Global Conference of Business and Economics, Cambridge 2006. A more elaborated Polish version of #3-#5 is now available in: *Efektywność źródłem bogactwa narodów*, ed.T.Dudycz, *Przedsiębiorczość i zarządzanie (SWSPiZ Łódź)*, vol.VII, no 1b, pp.380-393.

ROI for trainings and we will confront the requirements set within this framework with the practice of trainings in organizations. Next we will describe an approach to training management which draws on stakeholders' expectations concerning the training, to present the problems which this school of evaluative research is confronted with. We will use the example of a real-life training project to show how training practice resolves the tensions inherent in the ROE model. And finally we will sketch a theory which creates a basis for training management by ROE, which abides by the rules of science and assesses the results of training activities in a non-contingent manner. In analysing the assumptions of both approaches we will show that one must seek the sources of the differences between them not on the dimension of scientific practice, reliability or quality of measurement, but in a different vision of the role trainings play in the business organization.

ROI as an index for training evaluation

The theoretical frameworks for evaluating trainings practically ceased evolving since the concepts of the late fifties. It was then that D.Kirkpatrick proposed his model of four levels of training effectiveness (Kirkpatrick, 1994), which till today is the basis for thinking about training effectiveness (Alliger, 1989), (Bramley, 1996), (Alliger, 1997), (Ellis, 2005). Kirkpatrick introduced a taxonomy (Holton, 1996) of training results, differentiating between the effect of the training on participants (which is measured in two areas – reaction and learning), and its effect on the organization (measured by the transfer of skills to everyday practice, and the effects of this transfer for the organization). Kirkpatrick's theory is in its essence a positivist postulate for an external justification of the organization's results (i.e., the changes in an index of the organization's operations, observed by the researcher) (Woźniak, 2005k). The relationship between skills (or rather, knowledge) acquired, the will to try them out in practice, and the observed attempt at applying them, is a „theory” of this change² (Woźniak, 2005k). High values of these three explanatory variables are „proof” that the observed change in an important operational index is related to the training. „Everyone” must believe in the researcher's model of this change – the trainees wanted (as shown by the reaction measurement) to apply their newly gained knowledge (the learning measurement), tried to apply it in practice (the transfer measurement), and this brought the organization concrete effects (in the form of an observed change in an index on some practical dimension). Such reasoning – from will and skills, through application to effects – is naturally naive, as it is founded on the strong and unfounded assumption that all the remaining conditions affecting the organization's results are stable. It assumes that real causal relationships have been discovered – objective relationships, and not hypothetical ones from the Balanced Scorecard – so that the effects of different factors which operate in this stable situation can be separated out. It also overemphasizes the „power of words” – the conviction that people change their everyday behaviour, because they learned some content which they liked.

There are many alternative theories describing the chain of reasoning behind the effect of trainings on organizations. Phillips' anthropological research model, which leaves trainees the freedom to create this chain, is at the far end of this continuum. In J Phillips' model, it is the trainee who knows best how his work changed as a consequence of the training, and can assess the economic effects of this change (Phillips et al., 2002), (Woźniak, 2005w). Although

² Different interpretation of Kirkpatrick's theory contains (Alliger et al., 1989). They suggest correlations between training results on different levels. (Alliger et al., 1997) reveals substantial correlations between the various types of training results (against results from (Alliger et al., 1989)) and conclude that the reaction measures are as good surrogate of learning and transfer, as the job satisfaction for the job performance (Alliger et al., 1997, p.353).

the researcher then triangulates „raw data” from participants, the real basis for the analysis of economic effectiveness are opinions gathered from them.

All these research tools (be they quantitative, or – as in the case of Phillips – anthropological (Woźniak, 2005w)) imply that a concrete financial result can be indicated and – within the limits of common sense – ascribed to participating in the training. Although the costs of an analysis which would single out such a relationship can be significant and hence the economic value of the whole evaluation doubtful (Ellis, 2005), the model assumes such a figure can be singled out in real life³ and proved to be the effect of the training. In the opinion of practitioners such as the „ROI guru” J. Phillips, this approach has sense only in relation to a limited number of trainings (Phillips et al., 2002), but in the opinion of many who postulate the development of such methods of training management, the method is universally applicable and helpful in the rational management of the organization.

So it must come as a surprise that training management practice is inconsistent with universally recognized theory. Kirkpatrick himself disputed the prognostic value of the reaction measurement (which he equated with declarations of satisfaction with the training⁴). However, post-training questionnaires – which collect participants’ declarations concerning their attitude to the training – are a standard measurement in the training field (Training, 2004).

The training’s educational effect, which Kirkpatrick suggested could be observed using final examinations, is measured far less frequently. This may be justified by the almost nonexistent relationship between learning some content and long-term practical application (Woźniak, 2005w); but also by an immaturity of the training practice (Phillips et al., 2002). If practice, despite many years of appeals still does not want to conform to theory, then this is a social fact, requiring explanation.

Researchers have focused much of their energy on understanding the third level of training effectiveness – the factors that favour and hinder transfer (usually comprehended as the appearance in everyday life – i.e., at work – of behaviours consistent with the model taught) (Ford, 1997); (Taylor, 2005). Studies have been made of a significant number of variables and their interrelationships which may have an effect on transfer, with the application of refined statistical methods to training samples of various sizes and quality (Arthur et al., 2005). So it may come as a surprise that the influence of this literature on training practice is so limited – as exemplified by the fact that transfer is rarely measured in the US (Training, 2004).

It may be argued that from the perspective of organizational management practice, all this is research into an insignificant issue. Of importance for the organization – apart from the standards of correct behaviour decided on for strategic purposes – are not so much the way in which one acts, as the effects of this activity. This view – that the tasks of training and development are to justify their own effects on performance – has been the pivot of many discussions (Holton, 2002), but it is characteristic of the stance of ROI’s proponents and of contemporary approaches to HRM methods, that this view is self-evident. To recapitulate, although there has been almost half a century of research into training effectiveness since Kirkpatrick first formulated his theory, the ROI school of thought has not changed its main criteria for assessing a „good” theory of measurement effectiveness in trainings. The measurement has to supply a figure, which can be scientifically proved to be smaller than organizational „return” (profits). The fact that practice differs from postulates is for the ROI school reason for a critique of this practice, and not for revising its own assumptions. Training management practice

³ Kirkpatrick himself, when developing his theory, did not support such an approach.

⁴ He treated these more seriously than most later researchers, as he interpreted them as feedback from the client – a single criticism may be random but provokes thinking, whether it is not the symptom of a wider problem.

usually limits itself to a questionnaire on trainee opinions applied at the termination of the training⁵, which for the ROI school is not a rational activity, but a consequence of idleness or ignorance (Phillips et al., 2002).

Non-economic models of assessing the usefulness of trainings

Seeking economic justification for obviously reasonable activities in an organization seems a useless thing to do. A training is necessary if it will support changes, which the organization has to go through. The common sense feeling that an activity supports important changes is sufficient reason for that activity to be seen as worthwhile. If one adheres to such a pragmatic approach to training management, one needs only to find the tools which will help identify these important changes.

This model of thinking requires ascertaining what specific changes are necessary for improving organizational performance. This is not necessarily obvious and requires diagnosing the real needs of the main stakeholders engaged in the problem⁶ which sparked off thinking about the training. To collect and understand these needs, and also to distinguish between training and non-training aspects of the desired changes⁷, research termed „needs analysis” is conducted. This analysis is based on two types of data – the one concerning declared needs, the other a confrontation of stakeholders’ opinions with reality.

If training needs are equated with training „wishes”, then to evaluate a training means comparing pre-training expectations with an assessment of the results of the training, and the value of the training may be assessed with the use of the tools used for this comparison. Such an assessment of the benefits in relation to expectations (*Return on Expectation* – ROE) is based on the assumption that in studies on training effectiveness – as in any evaluative research – it is enough to ascertain that the outcomes are identical with goals. But these goals are to be set by the organization, not by the consultant who is in charge of preparing the training. The theory of responsive evaluation (Stake, 1975) (Bramley, 1996) is one of many that fit this model. Its name refers to the key factor in this approach – the training is to respond to the stakeholders’ view of their needs. This effectiveness evaluation model does not settle in what ways stakeholders’ opinions concerning their needs are formed and it is this weakness – as we shall show below – which is one of the sources of misunderstanding in discussions of this model.

In its exemplary form, as a result of the training evaluation process a report is presented to each of the stakeholders, describing the effects of the training for each of them – i.e., linking the training with issues which are important for each of the groups. The goal of this

⁵ US training practice of employers paid trainings can be analyzed on base of American Society of Training and Development year reports. Data vary yearly, but: level 1 measures are used for 72-83% of programs, level 2 – 32-53%; level 3 – 11-51%; level 4 – 6-44%. Other surveys show lower range – level 3 is 11-31%, and level 4 is 6-17% (Twitchell et al., 2002; p.87) For technical trainings (Twitchell et al., 2002) finds that percent of organizations using the level of training evaluation for more than 60% of technical programs are: 73% - 1.level; 47% - 2.level; 31% - 3.level; 21% - 4.level. Only 19% of organizations were attempting to use ROI to evaluate any technical training program.

⁶ On this level of generality we do not need to distinguish the kind of problem, which has been variously classified in the literature. (Boydell, 1999) distinguishes two dimensions – supporting, developing and perfecting the standard, and needs on the levels of individual, group and organization.

⁷ The differentiation between training-related and non-training needs is not purely analytical. For example, if students are unable to finish reading literature they have been set in the library’s working hours, we may extend the library’s working hours (a non-training, structural change) or train the students in faster reading skills. This slightly absurd example illustrates how structural changes may be substituted by changes brought about through training, and so shows the role played by decision makers, who determine what kind of changes the organization prefers.

analysis is to provide stakeholders with convincing arguments, that issues which were important for each group were duly included in the training. Research which prepares the way for such an assessment of whether goals have been met, needs positive and negative comments concerning the training from representatives of each group, and needs to single out the issues on which these groups differ. It is the researchers' task to isolate issues which should be the subject of further research. Two motives underlie the choice of research questions – the importance of the issue for the stakeholder group, and a well-founded concern that the given issue will not be assessed unequivocally. This second motive is not the same as obtaining an inconsistent opinion between stakeholders, as the researcher develops separate issues for analysis.

Such models of effectiveness assessment in the training field meet with two kinds of problems – technical (goal setting difficulties) and content-related (consistency of goals).

The first are a consequence of the drawbacks of a methodology which integrates individual preferences into a collective opinion, with the inherent costs and uncertainties concerning the value of training goals developed basing on consensus-type decisions. Consensus requires a high level of involvement from stakeholders who use this method to arrive at some kind of agreement. They have to be present at the same place and time, be ready to get involved in an interchange and to listen to each others' remarks, which requires their attention and a readiness to listen to others' argumentation. Achieving such a level of involvement in a large organization is less than probable. Beside the costs, a further drawback of the consensus method is the feeling of contingency participants have in relation to the group result. Everyone feels that a minor change in their position could bring about a change in the group result – and a change whose scale it is difficult to predict without making the attempt. Nor is there any tangible figure – i.e., data – which could be taken as an objective and prudent basis for the organization to act.

W. Trochim from Cornell University (Trochim, 1998) (www.conceptsystem.com) successfully manages to avoid these „technical” objections in his concept mapping method. This method quantifies the training assessment by relating the results of the training to expectations, or more precisely, to a map of these expectations created basing on actors' declarations. The process uses computer analysis techniques based on statistical grouping. The method deals skilfully with the problems of consensus building (Woźniak, 2006p). It does not require high involvement of actors, and overcomes the problem of their social and spatial dispersion. And although individual preferences are grouped by a formal procedure rather than by consensus, the method creates a collective preference, on the shape of which each stakeholder has some influence. The application of scientific paraphernalia such as computers and statistics rids this „consensus” of its aura of contingency, and avoids concerns over objectivity and precision, engendered by grouping different opinions through consensus (or meta-plan techniques) (Woźniak, 2006p).

Even in a paradigmatic situation this methodology is not completely irrefutable, as it can be used correctly only where the expectations of various stakeholder groups are fairly similar (Woźniak, 2006p). If stakeholders' opinions are contradictory, sensible training goals cannot be set without a more in-depth needs analysis. (Woźniak, 2004) describes an example, which vividly illustrates the need to distinguish between the two tasks – conducting a needs analysis and setting goals for the purposes of evaluating the training.

The goal of a needs analysis is to create a model of the problem, how it arises, how it is defined and how it will be eliminated. Moving from needs analysis to the decision that formulates training goals is not an analytical problem, but a decision made at the discretion of the main stakeholder, prior to the training. It is for management to decide that a proposed training will be a helpful tool in achieving its goals. So management defines its goals for the training, an activity which is to support others undertaken to achieve its managerial goals in a

given area.

This signifies that attaining consistency of expectations between actors is not the task either of the procedure of goal setting, or that of testing expectations. It is a task for needs analysis, which goes beyond testing expectations (i.e. collecting opinions or wishes) by researching facts and by formulating causal hypotheses concerning the source of the problem indicated by stakeholders, the source of its divergent interpretation, and the manner in which it can be eliminated.

This means that one of the tasks for needs analyses is to achieve consistency between key stakeholders' expectations. The development of such methods as conceptual mapping gives us the tools for creating quantifiable measurements of a training's results. As a result, achieving goals set for the training by the organization is given priority over the researcher's assessment of the training's economic effects. Although proof that goals have been attained does not have the standing of mathematical proof, the lack of one hundred percent correctness is not necessarily a shortcoming (Trochim, 1998).

Why – according to some – ROE is a better tool for managing training effectiveness than ROI

There are two main groups of arguments which attempt to show the greater worth of the ROE methodology of training management, over the ROI methodology.

Firstly, they dispute the condition that ROI for trainings has to be calculated. Naturally, being able to demonstrate a favourable ROI is useful from a persuasive perspective. However „education involves a series of human transactions that introduce ambiguity. Disentangling the variables that introduce ambiguity to achieve a level of scientific proof is difficult in most cases and arguably not cost-effective in others. As a result some programs may be able to demonstrate economic gains. Other programs may be unable to demonstrate gains, not because the program is at fault but because the unique features⁸ that make the proof of a financial result possible are not present. Such features include control over the implementation of the program, access to financial measures of outcomes⁹, and a clear relationship between the training program and any changes in business activity” (Trochim, 1998, p.286). In other words it is rarely possible to calculate the ROI, and in even fewer cases is this calculation profitable, so most good trainings cannot prove their effectiveness in this manner.

The second argument concerns the inadequacy of the paradigm demanding that a training's „accounts” be made in keeping with scientific method. „Training value is not an objective entity existing somewhere out there and waiting to be measured. Rather, value represents a constructed reality that depends on the views of the individuals or stakeholder groups associated with the process. Consequently, the construction of the meaning of value cannot come solely from an analysis of business activities but must include the constructed interpretations of value from multiple stakeholders” (Trochim, 1998; p. 287).

This argument goes beyond the belief that calculating ROI for trainings is in practice difficult, as it requires controlling too many variables. It states that no such concrete entity as the effect of a training – something like a hammer, or even an event described in a history manual – exists. Arguments are called on which question the objective judgement of history, as facts from social history are always given different values, depending on the perspective of

⁸ Another kind of argument referring to lack of control points to the team-orientation of organizational life, where the result is co-created by the necessary collective influence of other factors in various configurations of causal relationships (Dust, 2004), (Tobin, 2000).

⁹ (Dust, 2004) argues that in organizations, we have access only to figures relating direct costs and turnover, rather than figures associated with full costs and revenues (ENRON can be a good example).

the person who is doing the judging. And the postulate to depend on financial estimates for assessing value does not eliminate the differences between these perspectives, or the judgments that are based on them. We live in a world of socially constructed meanings, in which access to these meanings is enabled through negotiating these meanings by the actors in practical reality.

As a result, the proposed paradigm for evaluating trainings is based on the model of a complex problem, and not on that of a puzzle. This is because there are many „correct” answers to a problem; its „truths” consist of many compromises between alternative possibilities, and they allow for less-than-ideal „solutions” due to lack of resources or some other coincidence. However, this does not rule out the possibility of providing a means of measuring and communicating results to groups of stakeholders embroiled in the everyday workings of their organization, and disinterested in the detailed statistics of the internal problems of constructing the training.

To justify a training’s value, no more than two conditions must be met. Firstly, results should be provided, proving that expectations laid before the project which effected in this investment, were fulfilled. And secondly, the data on which further reasoning is based should be collected and processed in a manner that is methodologically sufficiently precise.

ROI or ROE – the problem of the scientific status of results

The thesis that there is no need to look for stronger proof than that admitted within the legal framework of argumentation, signifies looking for evidence and not proof in the mathematical sense. This is a fully conscious renunciation of the need for a complete and absolute irrefutability of analyses of the added value that a training brings. It does not mean that witnesses are to be called in to express their opinion about a given program, on the basis of which some independent judge will make a settlement, according to the principle of freedom of evidence appraisal¹⁰. For it is not the training management methodology – i.e. the need for a training needs analysis which will set reasonable goals for the training, and then relating the effects of the training to these goals – which is questioned, but the required ultimate level of explanation. A ROE-type study concludes with arguments showing that training goals have been achieved, for it is assumed that the training project operates within a framework of goals that have been set for it. These goals may be of an economic character; they may be directed towards changing some organizational index (e.g. the *Employee Satisfaction Index* or the *Customer Satisfaction Index*), or else towards teaching some procedures or behaviours, but they are all decided upon by the organization in the needs analysis phase.

This decision is not made in a social void – the organization makes it after rounds of negotiations between different groups. These negotiations are finally concluded with a decision concerning what the goals of the training are to be. It is for the decision makers to determine – basing on studies of the problem, stakeholders’ opinions and their own observations – if the explanation of negative phenomena and with it, the model of counteraction seems reasonable (Woźniak, 2005s) It is also they who decide what changes are to be expected following the training, as long as the causal model is not disrupted by additional circumstances

¹⁰ (Bramley, 1996) characterizes „the legal paradigm of justifying a training’s value” in such a manner, illustrating his argument with an example of research assessing the value of MBA programs provided by colleges in the US, and management programs provided by training companies in the UK (both studies from the late eighties). In both an evident problem was the lack of objectivisation – characteristic of the social life – of the opinions collected, which could serve to assess their value, and create an end result out of the partially contradictory pronouncements gathered.

(Woźniak, 20051). And it is they who have to be convinced that the expected effects are taking place, or – for identifiable reasons – have not been achieved.

This kind of reasoning shows that the problem of choosing a judge is a problem only in the social life studies, and not in an organization. Decisions within the organization have a natural arbiter, who must be convinced – the board. Formal power does not give infallibility but it gives the authority to resolve any conflicts, which may appear in the organization. And any settlement – despite fears to the contrary – is voluntaristic, as the opinions of other stakeholder groups and the activities they engage in act as safeguards against the arbitrariness of any decision maker's opinions.

In contemporary organizations it is impossible to force the participants of a worthless training to pronounce it useful and to prevent them from complaining about its being a waste of time. So although training goals do not always have to be compatible with the trainees' goals, a good training requires such a program and should be conducted in such a manner, that busy people will accept it as having been useful even if the training content was considered threatening¹¹, or insignificant¹² in everyday practice. Their opinions about whether the training was useful or valuable for them or their colleagues¹³ are not about the training's usefulness, but are an indicator of their feeling that they have increased their understanding of their work environment.

Measuring the reaction of trainees following the training (and after some time, when they have returned to work) is not done to prove that one of the training's effects is a will to apply theory to practice, but to find out whether an important stakeholder in the social life of the organization has accepted an action initiated by a decision of the managerial board. Participants accept the intent and how it is put into effect jointly, as opinion polls cannot justifiably be used to distinguish between „content” and how it is „transmitted”. Management's task was to choose a suitable supplier of trainings, who would diagnose the problem correctly and submit a correct solution, in the right way.

The problem of choosing a training, however, is not different from the problem of choosing a supplier of professional services. Trainers may proceed as called for by the rules of their profession, but this may not guarantee that the desired goals are achieved, even if the situational diagnosis was adequate, as particular circumstances independent of the supplier may prevent success.

Trainees' satisfaction is therefore an important indicator of whether the supplier's service is as it should be, but it does not signify that the problem the service was to solve has been solved. A training itself does not bring about changes in organizational reality, at the very most it can support changes (Woźniak, in preparation). Whether it supports or hinders these changes can be easily understood by observing and listening to trainees following the training. If the training was insignificant, trainees are usually well aware of their indifference towards the problems discussed, and express this in the post-training evaluations.

As we can see, this kind of reasoning is based on the conviction that reasonable and busy people benefit from a training which does not leave them indifferent. If the message that

¹¹ A good example may be trainings preparing for personal appraisal, communicating the introduction of a threatening system. The stronger the information that appraisals will be made basing on a forced distribution of scores, the more frequently the dimension “evaluate the training program” in post-training questionnaires is given less positive scores – the author's own research, on the basis of post-training questionnaires conducted in 1993-2003 (ca 100 trainings for different organizations).

¹² In Poland diversity-related projects introduced by American corporations based in Poland, are a typical example of training programs which trainees do not feel a personal need for, and cannot see the application of in their business reality.

¹³ These are typical dimensions, on which response is measured, directed not towards its affective but towards its utilitarian dimension.

was passed on by the training, and the form in which it was transmitted, are both accepted, we may conclude that „a good job was done”, even if the expected changes are not discernible. For the message was not a random one. It was chosen by management – which made its decision on the basis of the needs analysis and a causal model of intervention – as a tool used to support change (Woźniak, 20051).

To decide that “a sufficiently good job” was done despite the lack of changes, an explanation must be found for why none were observed. Such an explanation requires conducting another study, similar to the needs analysis. The symptoms and causes of the problem have to be identified from the perspective of stakeholders involved in it; behaviours observed and explained, and a causal model of the mechanisms of the problem constructed. Among others, this analysis should explain the relationship between the training and the problem, i.e. an answer should be created concerning how the training was useful and helped act in a desired manner, and what was missing, so that the desired behaviour did not become standard practise. In effect the analysis formulates recommendations on what should next be done to help implement the desired changes, and highlights the strong and weak points of the training project. To judge the training as „sufficiently good”, an assessment has to be made whether the weaknesses brought to light are an error which in the light of the previous needs analysis should have been prevented, or are a new circumstance, independent of the training.

The above description of further needs analyses associated with the problem may look like an extensive research procedure, but it takes place in most organizations informally, following on the training. It takes the form of a discussion by stakeholder groups involved in the problem, concerning the effects of the training and its effect on the problem under scrutiny. It is rarely formalized but is always conducted, if the training was an important event in the organization’s life. In this way, opinions concerning the professional activity (training) just terminated are negotiated between the stakeholder groups involved. These discussions are usually supported by other activities in the organization, as the desired change cannot be caused by a training in isolation, needing other organizational activities directed at bringing it about (Woźniak, in preparation).

As a result of these discussions, the opinion is formed that a good training was conducted. It is based on the strength of the trainees’ experience and the opinion they disseminate, on observations and informal interchanges between their superiors and other of the organization’s internal clients, on studies of the effect of the training and subsequent presentations of the results, and also on real changes in the indices targeted by the training. It is not just changes in business indices which are of significance for the assessment. Of equal importance are opinions that the training helped bring them about, in other words recognition by the community that the training was useful for the process of changing the organization.

As in every professional service, an assessment of its high quality is created by a combination of opinions including correct customer care (communication with the client’s representatives), opinions concerning input into solving the problem, and opinions concerning whether the rules of professional conduct were adhered to (Woźniak, 2006pk). All these opinions are the effect of negotiations between stakeholders, as the training service (as all consultancy) is not protected by professional corporations (of the legal or medical type). The result of these negotiations is final, but not random. Like in the case of a good judge, stakeholders are fairly accurate in settling which training was very bad and which very good. Participants do not have the professional competence to assess trainers’ behaviour in the terms of their professional conduct, but they are competent to assess their own new understanding of their organization. The rules of the game are to prepare and conduct the training in such a manner, that trainees will consider it worthwhile, in other words helpful for understanding practical reality.

We can accept (Woźniak, 20051) that the above verdict has a status similar to that of

results of scientific research conducted within the methodology of research into social movements, recognized and accepted since the past thirty years (Touraine, 1978, 2000). It is no more contingent than the acceptance of a model of social movements presented to their representatives, which is the confirmation criterion in Touraine's methodology. The more formalised the process of reporting training results – through observing trainees in practice, either by way of supportive coaching, or by way of measuring opinions or application of skills – the more the model of presenting results and obtaining acceptance for them on the part of informed stakeholders is reminiscent of the procedures of Touraine's methodology.

Conclusions

Although both approaches to testing training effectiveness consider themselves mutually exclusive and each argues that the other is not „sufficiently good”, our analysis shows that the differences between them are smaller than would seem so at first sight.

Both are able to formulate quantifiable opinions concerning training results, which can be related to goals set. In both, the scientific value of the measurements they create is equally irrefutable (as procedures are consistent with scientific practice in the social sciences) and equally dubious, as it is based on strong and unverifiable assumptions. In the case of the ROI model, it is assumed that one can estimate the isolated economic effect of the training – in other words the effect of other of the organization's parallel activities, as well as events occurring spontaneously in the environment, can be separated out. The ROE model assumes that the dialogues surrounding problems which the training was to change, is reliable. It should also be noted that research practice in both models differs insignificantly – one should talk with stakeholders, observe and understand what they do, observe organizational indices and try to explain them by referring to data amassed in other ways. Under certain special circumstances it is enough to hear only stakeholders' opinions (as in Trochim's model), or to observe changes in one of the economic indices (as in the popular understanding of the ROI model). Usually, satisfactory proof of the relationship between the training and the index it was to influence it is not possible without a more widespread study.

The real difference between these approaches becomes evident only when one comes to analyse assumptions underlying the training model.

The ROI model assumes that each activity in an organization is undertaken to change the economic outcome, and influences this outcome through a chain of causal relationships, registered in the form of a Balanced Scorecard. An effective activity will therefore change one of the indices, similarly to swallowing a pill (with cyanide) or producing a tool ready to be used. A training is effective if it provides such a pill and efficient, if its costs are smaller than the value of the effect it brings.

The ROE model assumes that the effect of the training belongs not so much to the sphere of objects, as to the sphere of meanings. It is not something substantial which exists at the end of a training, but is created in the course of dialogue between stakeholders, negotiated in different ways, and influences practice in different ways. For it is assumed that reasonable activities in social reality can be directed towards achieving various goals, and these can be so far removed from economic goals that the link between them is purely metaphorical. A common effort is necessary for every organizational effect, so of itself a training cannot cause anything (Tobin, 1998) (Dust, 2004), and only the dialogue that follows the training, conducted to support the training's effect or observe its consequences, can show up the effects of a new understanding of the professional role that participants built in the course of the training.

To summarise, the reason for rejecting the ROI approach – according to ROE's propo-

nents – lies not in the costs of its implementation and the necessity of controlling various environmental factors, but in the utopian character of its utility. A training itself does not bring effects, but is a step towards creating effects, made alongside other mutually supportive activities within the organization. Without activities undertaken to support its significance (Woźniak, in preparation), the training will be lost in the organizational dialogue that creates up-to-date meanings in the organization's social life.

Examples could naturally be found illustrating situations in which the ROI model is both possible and advisable. If the skills of using new machinery were taught, and following the training its participants perform well using the machinery, it is fully justified to evaluate the training according to the ROI model. On the other hand – if the goal of the training was increasing proactive attitudes in organizational communication, then treating such a training like a pill which will bring about change seems inappropriate, and in this case the ROE model will be a better tool for sorting out the good from the bad trainings. Further research should be conducted into the paradigmatic aspect of these examples, i.e. the suggestion that soft skills trainings should be assessed using the ROE rather than the ROI model.

Another direction for future research – other than research concerning type of training – revolves around the specifics of organizational context. The argumentation in our article (# 5) relates to knowledge-based organizations, but similar reasoning may be based on the relative strength of stakeholder groups. It is worth attempting to find other characteristics of the organization which may influence the decision which training evaluation model should be chosen.

REFERENCES:

1. Arthur Jr. W., Bennett Jr. W., Edens P.S., Bell S.T., (2003), Effectiveness of Training in Organizations: A Meta-Analysis of Design and Evaluation Features, *Journal of Applied Psychology*, vol. 88, no. 2, pp.234-245.
2. Alliger G.M., Janak E.A., (1989), Kirkpatrick's Level of Training Criteria: Thirty Years Later, *Personnel Psychology*, vol. 42, pp.331-342.
3. Alinger G.M., Tannenbaum S.I., Bennett W., Traver H., Shotland, (1997), A meta-analysis of relations among training criteria, *Personnel Psychology*, vol.50, iss.2, pp.341-358.
4. Boydell T., Leary M., (1999), Identifying Training Needs, Institute of Personnel and Development (Polish edition: (2001), *Identyfikacja potrzeb szkoleniowych*, Oficyna Ekonomiczna, Kraków)
5. Bramley P., (1996) *Evaluating Training Effectiveness*, McGraw-Hills, London (Polish edition (2002), *Ocena efektywności szkoleń*, Oficyna Ekonomiczna Kraków).
6. Dust B., (2004), The Myth of Training ROI, Chief Training Officer, no. 1; accessible on www.clomedia.com (as of 11.07.2004).
7. Ellis K., (2005), What's the ROI of ROI?, *Training*, vol. 42, no.1, pp. 6-8.
8. Ford J.K., Weissbein D.A., (1997), Transfer of training: An updated review and analysis, *Performance Improvement Quarterly*, vol. 10, no. 2, pp. 511-527.
9. Holton III E.F., (1996), The flawed four-level evaluation model, *Human Resource Development Quarterly*, vol. 7, no.1, pp.5-21.
10. Holton III E.F., (2002), Theoretical assumptions underlying the performance paradigm of human resource development, *Human Resource Development International*, vol.5, no.2, pp.198 - 214.
11. Kirkpatrick D., (1994), *Evaluating Training Programs*, Berret-Koehler, San Francisco.

12. Phillips J.J., Stone R.D., Pulliam Phillips P., (2002), *Human Resources Scorecard*, Butterworth-Heinemann.
13. Stake M.R.E.(ed.), (1975), *Evaluating the Art in Education: A Responsive Approach*, Columbus, Ohio.
14. Taylor P.J., Russ-Eft D.F., Chan D.W.L., (2005), A meta-analytic review of behaviour modelling training, *Journal of Applied Psychology*, vol. 90, no.4, pp. 692-709.
15. Tobin D.R., (1998) *The Fallacy of ROI Calculations*, *Corporate Learning Strategies*, accessible on www.tobincls.com (as of 20.06.2004).
16. Touraine A..A., (2000), *Method for Studying Social Actors*, *World Systems Research*, vol. 6, no.3, pp 900-918.
17. Touraine A.A., (1978), *La voix et le regard*, Editions du Seuil, Paris.
18. *Training management – a comprehensive analysis of employer-sponsored training in United States*, (2004), *Training*, vol. 41, no.11, (no author).
19. Trochim W.M.T., McLindena D.J., (1998), *From Puzzles to Problems: Assessing the Value of Education in a Business Context With Concept Mapping and Pattern Matching*, in: *Evaluation Systems and Processes*, ed. J. Phillips, Alexandria, pp.285-304.
20. Trochim W.M.T., McLindena D.J., (1998), *Getting to Parallel: Assessing the Return on Expectations of Training*, *Performance Improvement*, vol. 37, no. 8, pp.21-26.
21. Twitchell S., Holton III E.F. and Trott jr. J.W., (2002), *Technical Training Evaluation Practices in the United States*, *Performance Improvement Quarterly*, vol.13 no. 3, pp.84-109.
22. Woźniak J., (2004), *Analiza potrzeb szkoleniowych – procedura, nauka, czy sztuka sprzedaży*, in: *Przedsiębiorstwo i jego otoczenie w warunkach integracji europejskiej. Problematyka zasobów ludzkich*, ed. E. Niedzielski, Wyd.UWM, Olsztyn, pp. 183-192.
23. Woźniak J., (2005l), *Training Needs Analysis as a Method for the Development of a Quality Training*, in: *Economic Issues in Practice*, eds. B.Kryk, T.Bernet, Printgraph, Lama Project, Szczecin, pp.313-326.
24. Woźniak J., (2005s), *ROI dla szkoleń – utopia czy mit?*, in: *Polska gospodarka w UE*, eds. S. Pangsa-Kania S., G. Szczodrowski, Wyd.UG, Gdańsk, pp.284-295.
25. Woźniak J., (2005w), *Metoda Philippsa jako nowa nadzieja w badaniu efektywności szkoleń*, in: *Ewolucja zarządzania organizacjami. Doświadczenie i przyszłość*, ed. H. Król, WSiFiZ, Warszawa, pp.775-791.
26. Woźniak J., (2005k, in press), *Podejście zewnętrzne i wewnętrzne w badaniu efektywności szkoleń*, in: *Wybrane aspekty funkcjonowania gospodarki opartej na wiedzy*, ed. J. Korczak „Zeszyty Naukowe Wydziału Ekonomii i Zarządzania Politechniki Koszalińskiej”, chapter.4.9 (www.weiz.tu.koszalin.pl/konflo as of 1.06.2006).
27. Woźniak J., (2006p, in press), *Metoda mapowania pojęciowego jako narzędzie zarządzania efektywnością szkoleń*, *Zeszyty Naukowe Politechniki Świętokrzyskiej*, nr 32.
28. Woźniak J., (2006pk, in press), *O definiowaniu usług profesjonalnych*, *Zeszyty Naukowe Politechniki Świętokrzyskiej*, nr 32.
29. Woźniak J., *Szkolenia jako narzędzie rozwoju kadr* (in preparation).
30. www.conceptsystem.com, (as of 1.08.2004).