

How to Measure Safety Training Effectiveness? Towards a More Reliable Model to Overcome Evaluation Issues in Safety Training

Michela Vignoli^{*a}, Laura Punnett^b, Marco Depolo^a

^aDepartment of Psychology, University of Bologna, Via Filippo Re 6, 40126, Bologna

^bDepartment of Work Environment, University of Massachusetts, USA, One University Avenue, 01854, Lowell, MA.
michela.vignoli@unibo.it

Recently, safety training evaluation has received increasing attention. This is due both to the new Italian government guidelines and laws, and to the fact that a growing number of companies have started to perceive safety training as an investment. Many models have been published in the scientific literature on how to evaluate training in general, but less attention has been paid specifically to safety interventions and safety training. The aim of this paper is to develop a new model, which could help organizations in evaluating the effectiveness of training activities using both qualitative and quantitative data. The model is composed of different tools and it is based on the mixed methods approach, which combines quantitative and qualitative measures to provide a better understanding of research problems. Using this approach and collecting a large amount of data, it could become possible to create a national public database, which could be used to on a large scale to monitor the effectiveness of training activities and improve them.

1. Introduction

Safety at work has been a field of strong interest for many decades. The published data from the 15 Europe state Members, in 2010, reported 3,882,633 non-fatal injuries and 3,882 fatal injuries. In Italy in the same year 775,374 non-fatal injuries and 980 fatal injuries have been recorded (INAIL, 2011). Despite in the last years the occurrence of injuries follows a steady but slight decrease (that could be due in part to the economic crisis), fatal injuries in particular report very high rates compared to EU-15 countries. Furthermore, INAIL estimated that in 2010 almost 164.000 injuries were "invisible" (not reported to the authority), (INAIL, 2011), due probably to the position of workers without formal labor contracts and thus without public insurance of their health. To counteract this kind of phenomenon, the legislative bodies of different countries have formulated laws aiming to reduce risks associated with safety and health hazards at work. Taking the example of Italy, a series of laws from no. 626 of 1994 up to Law 81/2008 (in which were co-ordinated the various Italian and European standards), have affirmed the need to follow a different approach to the topic of safety at work. As well as the modification of some measures related to various dangerous jobs, there was a change that begun from a system essentially based on prevention technology, which passed through a comprehensive security system focuses on the man rather than the machine, and came to the personnel, the worker who became the true protagonist (Cavadi, 2007). This last Decree imposed the obligation on companies to train their employees on safety themes and their workers to be an active part of the training process, as protagonists and promoters of safety in their own workplace. In the last years, also the Agreement between state and Regions (2011) focused the attention on safety training interventions, stating that preference should be given to interactive methodologies through a balance between lectures and exercises, problem solving, simulation of work context problems, and e-learning activities. Although in Italy and in other countries safety training is mandatory, this activity could also be seen from an organizational point of view as investment. In fact, allocating financial resources for the training is a way to expect earnings, which could be both social (healthier and more aware workers) and monetary (in term of less injuries and less productivity loss). Furthermore safety

training is related to other important aspects such as environment. In fact, as reported by Jabbour (2013), the environmental maturity of the organizations could be influenced by organizational and human factors such as motivation and training of the workers. For those reasons it is important to evaluate the effectiveness of those training activities both to enhance the health and awareness of workers and to assure a return on investment of the organization. Thus, the aim of this paper is to present a new way to test safety training which is useful for two main reasons: a) to facilitate evaluation of effectiveness in the short term and in the long term; b) to collect data that could be generalizable in order to create a more wide database of training effectiveness.

2. How to evaluate safety trainings in organizations

The literature offers many models of training evaluation. Fraccaroli and Vergani (2004), tried to formulate a brief review of these kinds of approaches, based on seminal authors (e.g. Patton, 2002). Those authors noticed a broad consensus in distinguishing different kind of approaches. Some are highly structured and oriented to issues of measurement reliability, whereas others recognize the complexity of the evaluation process and aim to promote a shared interpretation among the actors involved.

The main approaches described are the following:

- a) "Classical" (with an emphasis on the aims of the social program and its achievements);
- b) "Scientific" (evaluation as an experimental approach, mainly based on quantitative measures);
- c) "Comparative" (ratio of costs and benefits between different programs);
- d) "Overall judgment" (qualitative assessment of a program);
- e) "Decision making" (the evaluation has to produce information useful for decision-makers);
- f) "Systemic" (analyzing relationships between inputs and outputs of a program to produce better results);
- g) "User-centered" (giving information useful for individuals involved in the program, specially stakeholders);
- h) "Goal-free" (emphasis on the unexpected consequences, not on pre-defined objectives); and
- i) "Responsive" (focus on the relationship between members of the system being evaluated). The last three models are more recent and are considered as emerging models (Scaratti and Regalia, 2000).

While the scientific debate concerning evaluation issues continually evolves, most peer-reviewed studies follow a scientific/quantitative approach (as defined by Fraccaroli and Vergani, 2004). In particular, the most widely used is the Kirkpatrick four-level model (1965). The four levels defined are: a) reaction (initial responses to training in terms of perceived utility and satisfaction; b) learning (the knowledge acquired by the participants); c) behavior (mainly how training contents are applied to the organizational contexts); and d) results (the consequences that training could have on workers and thus the organization).

Concerning workplace safety more specifically, a growing number of scholars have started to formulate new ways to design and evaluate safety training or more generally organizational interventions related to workers' safety. Pedersen and colleagues (2012a) postulated the Realistic Evaluation approach (RE) as a new way to design and evaluate occupational safety interventions. This approach acknowledges that criteria well-recognized in academic context (e.g. Cochrane criteria or randomized controlled trials (RCT)) ensure many advantages such as high internal and external validity and allow causality inferences (Victoria, Habicht and Bryce, 2004). However, it is important also to take into account the contexts and contents of research concerning workers, organizations and workplaces. For example, ethical issues may make it difficult to assign workers to the experimental and control samples within the same organization.

The realistic approach and the classical approach both start with an analysis of the issue proposed. However, RE differs from RCT in the kind of studies proposed (multiple and complex, compared to simpler and discrete) and in the aim (an explanation of what happened, compared to a final judgement) (Pawson et al., 2005). Finally, RE gives more importance to qualitative measures. An important strength of this kind of evaluation is that policy-makers can get a wider comprehension of the variables involved in interventions and programs, and why they really worked or not. That is different from getting information on interventions and programs only in terms of statistical significance: such an outcome is crucial for scientific research, but less useful for decision-making. The model proposed by Pedersen et al. (2012a) provides that to infer a causal outcome between the intervention and the outcome it is necessary to understand the mechanism that links them and the context in which the relationships occurs (Pawson et al., 2005). The authors add that mere safety interventions are not presumed to have causal powers in themselves, but the context and the mechanisms are among the factors which can initiate the causal relationships. Concerning the context, the stakeholders related to the interventions are also an important factor to the good success of the safety intervention. The outcome of the intervention can vary accordingly to the intervention, the context, the mechanisms, and the interplay between these. The outcome can be categorized as positive or negative, expected and unexpected. These arguments created an interesting recent debate in the *Safety*

Science journal. In fact, after the publication of the Realistic Evaluation model, Verbeek and Ruotsalainen (2012) replied with some arguments - such as the possibility to use cluster randomization - to overcome ethical issues related to control groups, concluding that the Cochrane Systematic Review remains the most powerful, valid and transparent method for compiling studies' results. Pedersen and colleagues (2012b) responded with various argumentations such as many studies have highlighted the limits of RCT, the Campbell Collaboration includes a more variety of study design and that the value of analyzing the mechanism that explains why an intervention worked or not. Although this debate was not focused only on safety training, but more in general on safety interventions, it highlighted the importance of defining how to evaluate safety training in a valid and meaningful way. Concerning the Italian context, despite safety training evaluation is not widespread, there are some ongoing experiences documented in the literature (e.g. Saracino et al., 2012), while worldwide many authors are putting their attention also on issues relating the ongoing web-technologies (e.g. Grigorov and colleagues, 2012). Relating to those theoretical arguments, the literature suggests two additional aspects that should be considered in evaluating the scope and impact safety training: a) level of engagement of the training; b) the distinction between supervisors and employees.

2.1 Level of engagement of the training

Two important reviews have argued that the level of engagement of the training could have an impact on safety and health outcomes. The first one (Burke et al., 2006), tried to figure out what is the relative effectiveness of different methods of safety and health training with respect to three relevant outcomes: a) safety knowledge (self-rated or by test of knowledge); b) safety performance (self-, supervisor-, or observer rating of safety-related behavior); and c) safety and health outcomes (measurement of accident, illness or injuries). The authors proposed three categories of engagement in safety training: least engaging (comprehends lectures, video and pamphlets, written materials); moderately engaging (feedback interventions, performance information provided in small groups, computer-based interactive instruction); and most engaging (hands-on demonstration associated with behavioural stimulation, active participation of the trainee). The results showed that the more engaging is the training, the greater are the effects on knowledge acquisition. Furthermore, as the level of engagement in training increases, training will have greater impact on reduction of negative safety and health outcomes. The findings concerning safety performance were more ambiguous but provided some support for the effectiveness of more engaging training methods. Overall, Burke and colleagues (2006) argued that in safety training the most engaging methods were approximately three times more effective than the least engaging ones in promoting knowledge and skill acquisition. The second review developed more recently by Robson and colleagues (2012) uses a new conceptual model, which postulates that training, is supposed to have an immediate effect on outcomes such as knowledge, attitudes, and behavioural intentions. These outcomes potentially impact on job behaviours and hazards, which in turn affect outcomes measured in the longer term, such as workplace injuries and illnesses. The model also indicates that various aspects of training, trainees, and work environment determine these effects. In Robson's review (2012), training interventions have been categorized by level of engagement, following the classification from Burke and colleagues (2006). The categorization of outcome interventions instead, is quite different as Robson's review distinguishes an additional category (attitudes and beliefs) that includes variables such as perceived risk, self-efficacy and behavioural intentions. Robson concluded that the effectiveness of OHS intervention on knowledge, attitudes, and beliefs could not be confirmed, because of the small numbers of studies. Regarding the effect on behaviours, the effectiveness of interventions is strong, but for health outcomes the results were inconsistent. Overall, these two reviews showed that the more engaging are safety training methods, the more they are effective in reducing negative consequences such as accidents. However, looking at the relationship between engagement level and safety outcome, the two reviews' results also highlighted some differences such as for example the effect on behavioural outcomes. The contradictory nature of the results could be due to two possible reasons. The first one is that the reviews took in consideration different sets of studies. The second reason is that there are a large number of heterogeneous variables that could influence the effectiveness of the evaluation so it is very difficult to understand what are the variables that combining with the intervention process influence the outcomes measured.

2.2 Distinction between employees, supervisors and management

The distinction between employees, supervisor and management is suggested from different fields: both the academic one and the legislative one. Concerning the scientific literature, it is widely demonstrated (e.g. Marsh et al., 1998) how management commitment could have a strong impact in determining the success of a safety intervention. Furthermore also many safety measure scales (e.g. Zohar et al., 2005;

Brondino et al., 2013) distinguish with different subscales the perceptions related to workers, supervisors and management. Regarding the legislative field, the Agreement between State and Regions provide different mandatory training contents related to the role in the organization. In particular the Agreement distinguish between workers, safety supervisors, management and the employer in the defining the duration and the contents of the training.

3. Aim

Starting from the arguments outlined above, we aim to develop a new model to help organizations (public or private) to evaluate safety training activities in a way that is in line with the literature but also time-efficient. A useful approach, which is becoming more acknowledged in recent literature, is the mixed methods approach where qualitative data could integrate quantitative data. Creswell and Plano Clark (2005) define it as “A research design with philosophical assumption as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analysing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone.” (p.5). Furthermore, we believe that the mixed methods approach, where qualitative data could be integrated with quantitative data, provides information that is inherently meaningful to the members of an organization and therefore may result in a higher degree of motivation to implement and strengthen a program.

4. Safety training evaluation model

4.1 Timing

In the training effectiveness evaluation, short versus long term evaluation is a core theme. The concept that best describes this theme is the transfer, which is defined by Blume and colleagues (2010) as the extent to which the learning that results from a training experience transfers to the job and leads to meaningful changes in work performance. The importance of the transfer is showed in many studies that reported that acquiring new skills doesn't guarantee a behavior change. For example Burke and Hutchins (2007) showed that only the 50 % of training activities produce a behavior change. To monitor this process this model provides to collect data on three different times: before the training start (pre), after the training (post), and after 6 months (follow-up). Those measurement times could help in analyzing whether there are changes before and after the training, to understand the transfer process in the passing time, and to understand whether the organization provides the conditions within each worker can use the new knowledge obtained.

4.2 Tools

The proposed model relies on quantitative and qualitative data collected by several different tools, which are briefly described in the following sections.

4.2.1 Questionnaires

Questionnaires are useful instrument to understand workers' personal perceptions. The questionnaire proposed is composed by variables which are investigated in different times (pre, post and follow up). According to the literature described before, those variables should be related to the knowledge, the attitudes and beliefs, the behaviors and the health outcomes. In addition to this it is important to evaluate the level of engagement of the worker and the training in general (e.g. whether it worked or not also respect to the expectations). Recently Grohmann and Kauffeld (2013) created the Questionnaire for Professional Training Evaluation (Q4TE) based on an analysis of the literature on the safety training evaluation. This questionnaire has psychometrically sound properties and it is widely applicable among different training contents, so it allows for comparisons of training programs within and between different organizations. Furthermore, it appears interesting to test the instrument in the Italian context.

4.2.2 Interviews

In line with the arguments proposed before, to better understand how training activities are effective or not, a random sample of workers will be interviewed. The interviews will focus on the main dimensions investigated by the questionnaire: satisfaction with and utility of the training, the participant's perceived knowledge acquisition, the practical utility of what has been learnt during the training and the impact of this

knowledge on individual and organizational results, the engagement level of the training and the safety climate.

4.2.3 Organizational Training Summary

A sheet fulfilled by the organization to summarize data concerning the training content (related to the categories defined by the Agreement between State and Region) and the activities played during the training (e.g. simulations of job activities, lectures etc.).

4.2.4 Organizational Database Recording

Organizational data recording respect to injuries, accidents, near misses provide an important tool to monitor the safety indices in the organization.

5. Application of the model

In this section we provide a brief example of an application of the model, which is still in progress (data has been collected at the pre and post time so far). Data available right now concern a safety training on asbestos with 14 managers. The mean age was 34.27 years (s.d.= 9.44), while the organizational seniority was 7.75 (s.d.= 6.58). Most of them were men (72.7%) and Italian (90.9%). We asked the workers to fulfill the questionnaire before the training (pre), after the training (post) and after 6 months (follow up). After the training we randomly choose three workers to participate in an interview with the researcher. The preliminary results of this application showed that despite the small sample (N=14) a statistically significant difference has been found (paired sample t-test) between the expectation of the workers on the course before starting, and the perceptions of the course when it is concluded. In particular, the workers overcame the previous expectations respect to the clarity of the content ($p=.012$; mean-pre=4.18, mean-post=4.36) and the importance of the training course ($p=.006$; mean-pre=4.36, mean-post=4.55). Relating to this, the interviews confirmed and enriched those results. Managers interviewed told that the most important variable that contributed to the training success was the engaging level of the trainers and the subsequent participatory climate. This was confirmed also by the significant correlation ($p=.04$ $r=.653$) between the level of engagement perceived by the workers and the knowledge acquisition. Another noteworthy result obtained thank to the integration between quantitative and qualitative measures concern an important variable influencing the transfer into the workplace, i.e. the applicability of the learned concepts on the workplace. Indeed, despite the result is not statistically significant, the average mean between pre and post is very large (1.54 before the training and 2.63 after the training). Furthermore, although the interviews confirmed that managers really enjoyed the training course, two of them reported that the contents of the training couldn't be easily transferred into their job for different reasons (because their type of work has not been considered and this resulted in a difficulty in trying to modify their safety behavior. As the project is in progress, the follow up data will be collected six months after the end.

6. Conclusions

The paper described the literature state of the art on how to evaluate safety training activities in organizational contexts. Starting from this point, we developed a new practical model composed by different tools that could help the researcher and the organizations in better understanding why a training activity is effective or not and what are the variables that could best describe the success or the failure of a training activity. Therefore we provided an example in progress of a possible application. Preliminary results showed that using both qualitative and quantitative data was a useful way to collect more in-depth information, and reduce possible bias that could lead to a wrong interpretation of the results. After a large amount of data collected from different organizations, thresholds and categories could be created to help both the organizations in understanding the level of effectiveness of the training activities, and to create a national public database aiming to monitor the effectiveness of training activities. In doing this it is important to differentiate between in-presence training and e-learning training, as they can induce differences respect to workers' engagement. In fact, different kind of training could present peculiar challenges (e.g. Grigorov et al., 2012), which could influence the training success (Pedersen et al. 2012)

References

Blume B.D., Ford J.K., Baldwin T.T., Huang J.L., 2010, Transfer of Training: A meta-analytic review, *Journal of Management*, 36(4), 1065-1105, DOI: 10.1177/0149206309352880

- Brondino M., Pasini M., da Silva S.C.A., Development and validation of an integrated Organizational Safety Climate Questionnaire with multilevel confirmatory factor analysis, *Quality and Quantity*, 47(4), 2191-2223, DOI: 10.1007/s11135-011-9651-6
- Burke L.A., Hutchins H.M., 2007, Training transfer: An integrative literature review, *Human Resource Development Review*, 6, 263-296, DOI: 10.1177/1534484307303035
- Burke M.J., Sarpy S., Smith-Crowe K., Chan-Serafin S., Salvador R.O., Islam G., 2006, Relative Effectiveness of Worker Safety and Health Training Methods, *American Journal Of Public Health*, 96(2), 315-324, DOI: 10.2105/AJPH.2004.059840
- Cavadi G., Ed., 2007, *Safety Psychology*. Cortina Libreria, Milan, Italy (in Italian).
- Creswell J.W., Plano-Clark, W.L., Eds., 2007, *Designing and conducting mixed methods research*. Sage, Thousand Oaks, CA
- Fraccaroli, F., Vergani, A., Eds., 2004, *Valutare gli interventi formativi*. Carrocci Editore, Rome, Italy.
- Grigorov, A., Angelov, A., Detcheva, E. Varbanov, P.S., 2012, Web technologies for interactive graphical e-learning tools in engineering education, *Chemical Engineering Transactions*, 29, 1597-1602, DOI: 10.3303/CET1229267
- Grohmann A., Kauffeld S., 2013, Evaluating training programs: development and correlates of the Questionnaire for Professional Training Evaluation, 17(2), 135-155, DOI: 10.1111/jttd.12005
- INAIL (Istituto Nazionale per l'Assicurazione contro gli Infortuni sul Lavoro), 2012, *Rapporto Annuale 2011* <www.inail.it> accessed 19.11.2013
- Jabbour, C.J.C., Environmental training and environmental management maturity of Brazilian companies with ISO14001: empirical evidence, *Journal of Cleaner Production* (2013), DOI: 10.1016/j.jclepro.2013.10.039
- Kirkpatrick D.L., 1959, Techniques for evaluating training programs, *Journal of American Society of Training Directors*, 13(3), 21–26.
- Marsh T.W., Davies R., Phillips R.A., Duff A.R., Robertson I.T., Weyman A., Cooper M.D., 1998, The role of management commitment in determining the success of a behavioral safety intervention, *Journal of the institution of Occupational Safety & Health*, 2(2), 45-56.
- Patton M.Q., Ed., 2002, *Qualitative Research and Evaluation Methods*. Sage, Thousand Oaks, CA.
- Pawson R., Greenhalgh T., Harvey G., Walshe K., 2005, Realist review – a new method of systematic review designed for complex policy interventions, *Journal of health services research & policy*, 10(1), 21–34, DOI: 10.1258/1355819054308530
- Pedersen L.M., Nielsen K.J., Kines P., 2012a, Realistic evaluation as a new way to design and evaluate occupational safety intervention, *Safety Science*, 50, 48-54, DOI:10.1016/j.ssci.2011.06.010
- Pedersen L.M., Nielsen K.J., Kines, P., 2012b, Reply to letter regarding Realistic evaluation as anew way to design and evaluate occupational safety interventions. *Safety Science*, 50, 1153-1154, DOI: 10.1016/j.ssci.2011.11.016 pp. 1153-1154.
- Robson L.S., Stephenson C.M., Schulte P.A., Amick B.C., Irvin, E.L., Eggerth D.E., Chan S., Bielecky A.R., Wang A.M., Heidotting T.L., Peters R.H., Clarke, J.A., Cullen K., Rotunda, C.J., Grubb P.L., 2012, A systematic review of the effectiveness of occupational health and safety training, *Scandinavian Journal of Work and Environmental Health*, 38(3), 193-208, DOI:10.5271/sjweh.3259
- Saracino A., Spadoni G., Curcuruto M., Guglielmi D., Bocci M., Cimarelli M., Dottori E., Violante F. , 2012, A New Model for Evaluating Occupational Health and Safety Management Systems (OHSMS), *Chemical Engineering Transactions*, 26, 519-424, DOI: 10.3303/CET1226087
- Scaratti G., Regalia C., 2000, *Approcci metodologici alla valutazione dei servizi*. In Regalia C., Bruno A. (Eds.), *Valutazione e qualità dei servizi* (pp. 43-84). Unicopli, Milan, Italy
- Scriven M., 1967, *The Methodology of Evaluation*. In Tyler R.W. et al., Eds., *Perspective of Curriculum Evaluation* (pp. 39-83). Rand McNally, Chicago, IL, USA.
- Stretcher B.M., Davis W.A., Eds., 1987, *How to focus on Evaluation*. Sage, Newbury Park, CA, USA.
- Verbeek J., Ruotsalainen J, 2012, Letter to the editor re Pedersen et al., *Safety Science*, 50, 1152. DOI: 10.1016/j.ssci.2011.11.022
- Victoria C.G., Habicht J.P., Bryce J., 2004, Evidence-based public health: moving beyond randomized trials, *American Journal of Public Health*, 94(3), 400-405, DOI: 10.2105/AJPH.94.3.400
- Zohar D., Luria G., 2005, A multilevel model of safety climate: cross-level relationships between organization and group level climates, *Journal of Applied Psychology*, 90(4), 616-628, DOI: 10.1037/0021-9010.90.4.616