*PERIODICA POLYTECHNICA SER. SOC. MAN. SCI. VOL. 7, NO. 2, PP. 105–118 (1999)*

# A METHODOLOGY FOR ASSESSING AND DEVELOPING TEAMWORK IN COGNITIVELY DEMANDING JOBS

Miklós ANTALOVITS and Lajos IZSÓ

Department of Ergonomics and Psychology
Budapest University of Technology and Economics
H–1521 Budapest, Hungary
Phone: (36-1) 463-2654, Fax: (36-1) 463-2106
E-mail: antalovits@erg.bme.hu

Received: Oct. 9, 1999

## Abstract

In cognitively demanding jobs such as control room operators in nuclear power plants, chemical works and oil refineries, pilots and flight traffic operators, dispatchers, fire brigades, rescue services, surgeons, police officers, etc., it is a crucial question that the members of the team have appropriate professional competence as well as all sensory-motor, cognitive, communicational and social skills necessary to identifying, diagnosing critical states and adequate intervening. Simulated malfunctions or other events potentially offer excellent opportunities for establishing, refreshing and improving these skills, provided that the simulation is embedded into an appropriate pedagogical and psychological environment. For efficient teamwork both individual and group learning correct and informative feedback have an outstanding importance: it provides crew members with a clear and relatively objective picture about team work effectiveness as well as quality and quantity of individual contribution to it. For these purposes a computer-supported methodology and a set of tools have been developed ('COSMOS'), by which the instructor involves the operator team into the evaluation of a training session, thus supporting formation of more realistic self-evaluation, more uniform ways of seeing things and shared mental models. This paper is a report on a case study carried out with nuclear power plant operators.

*Keywords:* team assessment, team development, cognitive demands, simulator training, computer-support.

## 1. Introduction

In the nuclear power industry recently a very explicit demand has been formulated that – in accordance with the 'Systematic Approach to Training' (SAT) philosophy introduced by IAEA – a training simulator has to have an effective learning and evaluating environment capable of developing technical, communicational and co-operational skills as well as evaluating both individual and crew performance as objectively as it is possible. To be able to meet these requirements trainees have to be provided with carefully designed informative feedback for technical and social learning and in addition to the usual evaluation by instructors the possibilities of operators' self-evaluation also have to be utilised. If conducted properly, self-evaluation could be a powerful tool to increase objectivity, to increase operators' self-knowledge and their knowledge about their fellow operators, to improve communication skills within the group, etc.

## 2. Fundamentals of the COSMOS Methodology

### *2.1. The Basic Philosophy*

Our previous experiences have proved that the short evaluating sessions immediately after completing simulator training are potentially very unique and psychologically extremely valuable situations, which, by properly designed methods, can be effectively utilised to increase crews' preparedness and – indirectly – to also improve the safety of operation. The basic problem has been, however, the lack of such methods.

The situation immediately after a cognitively demanding simulator session can be characterised by the following: (1) the experiences and memories of crew members are still quite vivid and fresh concerning the details of simulated malfunctions, their own and the fellow operators' activities, (2) in addition to the factual memories, they still keep the emotional flavour of the situation related mainly to success or failure in form of tensions requiring for acting out, (3) besides, they still have quite definite opinions – should these be correct or incorrect – about expected roles and actual effectiveness of individual operators, (4) video recordings and computer protocols are still available as objective references for discussion and debates. So a very intensive learning process can be started involving not only technological knowledge and experience about their own and the others' task, but also concerning group norms, communication skills, co-operation and leadership effectiveness. The results of this accelerated learning are more realistic and more uniform opinions and knowledge about situations and problems to be solved, risks, expected roles of crew members during emergency situations and optimal group behaviour.

### *2.2. The Main Steps of COSMOS Methodology*

Based on the experiences described above a new method called **COSMOS** (**CO**mputer **S**upported **M**ethod for **O**perators' **S**elf-assessment) has been developed. It has eight main steps as follows.

*(1) Carefully designing training scenarios in advance that include identification of what are called 'key situations' of the simulated emergency*

Those elements of simulated malfunctions (operating conditions) that presumably will play a determining role in the decision made by the trainees occupying the various operator positions must be identified in advance. These elements define the 'key situations' for which the trainees later – immediately after the training session – perform a retrospective group and individual self-assessment of their performance.

*(2) Defining the key situations of the scenario and the trainees for the different operator posts to the computer right before the actual session*
As a part of preparing for the simulation training the instructor puts the key situations of the scenario and the names of trainees in different operator posts into the COSMOS software.

*(3) Carrying out simulation training session; recording and observing crew behaviour*
All relevant information that characterises operators' activity, including interventions, behaviour, communication within the crew and between the crew and the instructor, behaviour of the team leader and group climate is recorded by means of audio/video tape, computer log, observation, and other methods.

*(4) A short discussion of the main events actually occurred during the current exercise and redefining key situations if necessary*
Immediately after the training session the trainees and the instructor briefly discuss the main events of the session, redefining the key situations if necessary.

*(5) Determining and assessing perceived relative difficulties in key situations, with operators comparing pairs of situations and examining concordance within the crew*
They then identify and compare the perceived difficulties in pairs of key situations, comparing in such a way only two at a time. The computer calculates the coefficients of intra-individual consistency, and group concordance data that are then projected on a large screen. The consistency coefficients are defined as

$$K = 100 - \frac{2400a}{k_{\max}^3 - k_{\max}} \qquad \text{if } k_{\max} \text{ is odd, and}$$
$$K = 100 - \frac{2400a}{k_{\max}^3 - 4k_{\max}} \qquad \text{if } k_{\max} \text{ is even.}$$

It did happen, however, that assessments of key situations were inconsistent. For example, if an operator judged key situation 1 to be more difficult than key situation 2, key situation 2 more difficult than key situation 3, and then key situation 3 is more difficult than key situation 1, the last assessment in this three-part chain, or triad, contradicted the initial assessment. That contradiction is that we refer to hereafter as a *decision loop,* or *inconsistent triad.* In the formulas $a$ is the actual number of decision loops (inconsistent triads) characterising individual decision contradictions, and $k_{\max}$ is the maximal number of key situations considered (actually 6 or 7).

Following this procedure the degree of group concordance is tested on the basis of Kendall U statistics (KENDALL, 1948). The formula adapted to our case is

$$U = \frac{\sum_{k_s,k_0} c_{k_s,k_0}^2}{\binom{n}{2}\binom{k_{\max}}{2} - \frac{n+1}{n-1}},$$

where $k_s$ and $k_0$ are the row and column indices of the aggregated difficulty matrix, respectively, $n$ is the number of operators (in our case $n = 5$) and $c$ is the element of aggregated difficulty matrix in the corresponding row and column.

If all coefficients of individual consistency and group concordance are acceptable – high enough – the overall rank order is computed on the basis of mathematical model. If this is not the case, the assessment is repeated to increase either individual consistency or group concordance.

*(6) Assessing the fellow trainees' and their own expected roles (in short also called Involvement) and actual performances (in short also called Effectivity) in each key situation*

Having compared perceived difficulties of key situations, the trainees conduct individual situation-by-situation or operator-by-operator evaluation and self-assessment of expected roles (on a three- degree scale: small, medium, large) and performance (on a five-degree scale: unacceptable, poor, medium, good, excellent). Summary characteristics (and when necessary and justified also individual characteristics) are presented to the crew in graphic form on a large projector screen. If warranted by the results of discussion, the assessment is repeated to increase objectivity and the degree of agreement within the group. The setting of the group self-assessment of performance and the individual self-assessment of performance is pictured in *Fig.* 1.

*Fig. 1.* The setting of a COSMOS assessment and self-assessment session

*(7) Playing back a two-to-five minute video recording of the most critical key situation as a basis for making self-assessment about their own behaviour*

Having the crew members view the recording of what they have judged to be the most difficult situation helps them recall the situation and refresh their memories. After watching the video clip, the crew members evaluate their effectivity in this critical situation along three dimensions: (1) information-gathering, (2) decision-making, and (3) co-operation. In addition to that they also assess their satisfaction with themselves on a five-degree scale. Operators receive global graphic feedback from these assessments in addition to summarised opinions formulated by their fellow operators.

*(8) Reviewing results and giving detailed feedback to the trainees in a discussion moderated by the instructor, concluding remarks*
The instructor embeds the use of COSMOS into the process of evaluating performance. In other words, this method is never used in isolation. Rather, the instructor applies it as a flexible set of tools subordinated to pedagogical and didactic goals as a form of reinforcement. The instructor can also use the method to provoke debate by focusing on conflicting viewpoints and opinions, depending on the pedagogical context. In addition to the varied graphic feedback formats designed to inform operators, detailed and sophisticated numerical and tabular information about group and individual self-assessment are made available to the instructor by the mathematical model underlying the COSMOS method, the purpose being to deepen that person's understanding of, and insight into, group behaviour, dynamics, attitudes, and norms.

A summary flow chart of the use of the COSMOS method can be seen in *Fig.* 2.

## 3. A Case Study

### *3.1. Background*

After carrying out video-supported laboratory pilot studies with the COSMOS method – see References – involving members of simulator training staff, the method has also been tested with ten real trainee operator crews in the frame of their regular refresher simulator training sessions in April – May of 1996. The aims of this series of experiments were to gain practical experiences about the implementation of COSMOS method and also to prepare regular use of self-assessment in simulator training session.

The operator crews were made of the following five operators:

- BE Block Electrician,
- SS Shift Supervisor (head of crew),
- RO Reactor Operator,
- TM Turbine Mechanic,
- TO Turbine Operator.

The scenario of a simulated emergency situation has been developed by the staff of Simulator Centre based upon the following six cognitively demanding 'key situations' (*Table 1*).

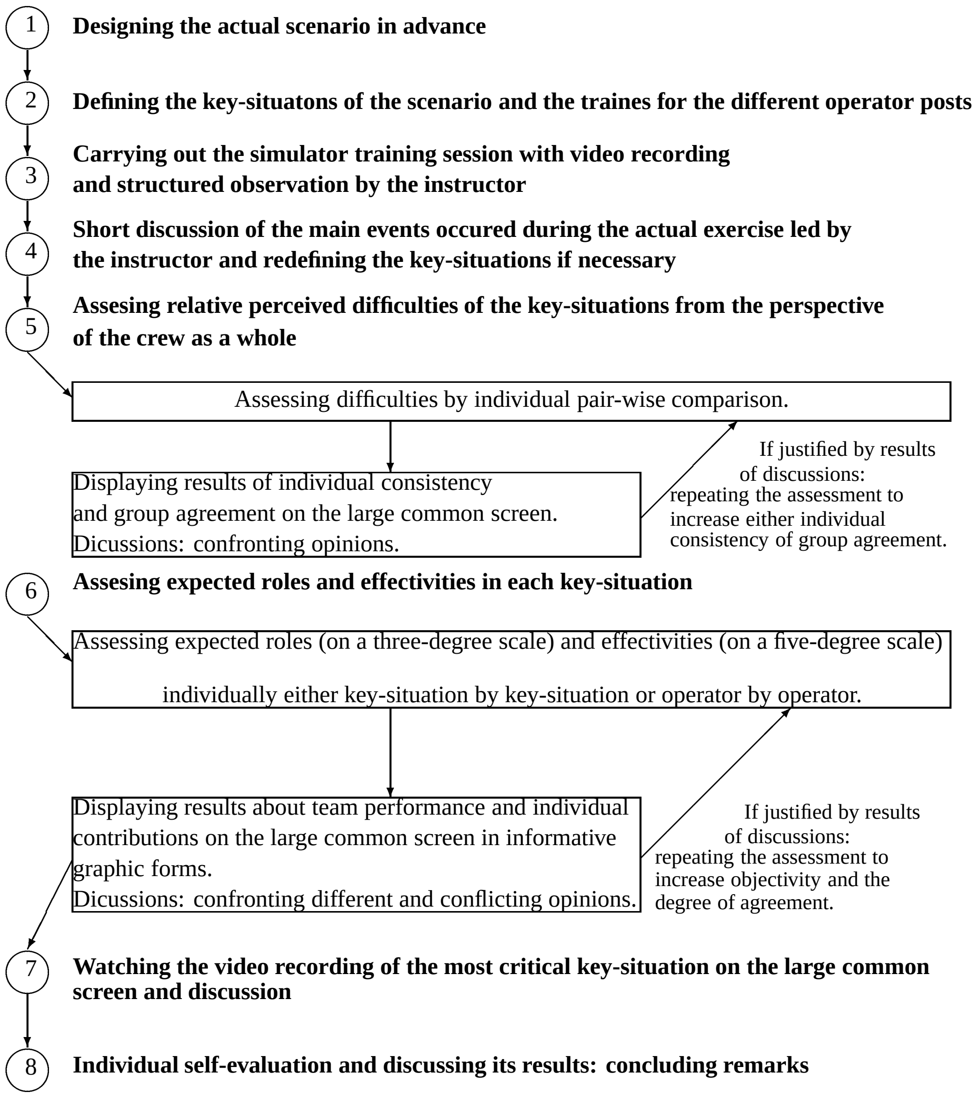


*Fig. 2.* Flow chart of COSMOS method

*Table 1*. The six cognitively demanding ‘key situations’ in the case study

| Code of ‘key situation’ | Required actions from the crew for solving problems emerged in ‘key situations’ |
|---|---|
| K1 | Realising trip of Main Circulating Pump; stabilising power. |
| K2 | Realising Steam Generator rapture; identifying leakage. |
| K3 | Keeping Feed-water Pumps in operation, keeping water level. |
| K4 | Realising that the loop can not be isolated; decreasing primary circuit pressure. |
| K5 | Adjusting appropriate cooling speed, avoiding reactor shut down. |
| K6 | Avoiding spreading radioactivity; isolating Steam Generator. |

### *3.2. Results*

The main experiences and results of the series of ‘experimental’ sessions are summarised in this section. The results of the very first session were not recorded due to a technical error.

a) The great majority of operators were able to make consistent decisions concerning perceived difficulties of key situations during their pair-wise comparison (*Fig.* 3).

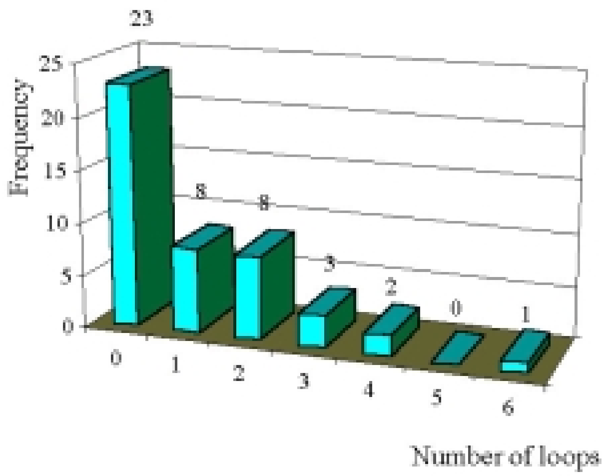


*Fig. 3*. Frequency of all operators’ loop numbers

A ‘loop’ represents an inconsistent decision triad, which is a direct measure of operators’ inconsistency. It can be seen in the figure, that about 90% of operators – 39 out of 45 – made decisions with less than three loops.

b) With the exception of one crew (see crew No. 4 on *Fig.* 4) there was also significant agreement in the judgement on the order of perceived relative difficulties of key situations.

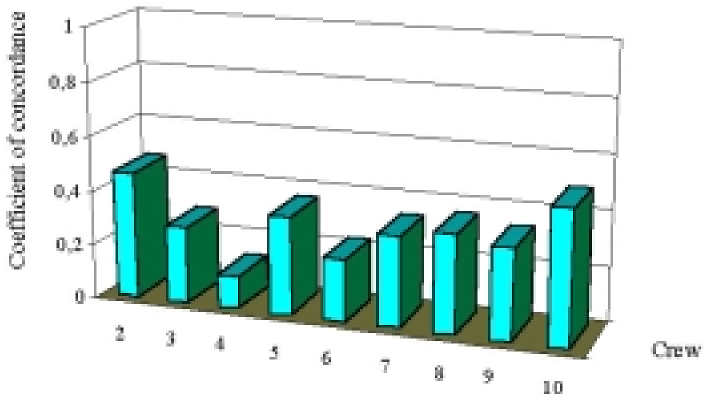


*Fig. 4.* Coefficients of concordance of crews

c) The order of judged relative perceived difficulties was rather similar in all crews (see relatively low SD-s in *Fig.* 5). As judgements on difficulties were made in agreement within all crews but one, this order of perceived difficulties can be taken as general and characteristic of the great majority of the operating personnel.

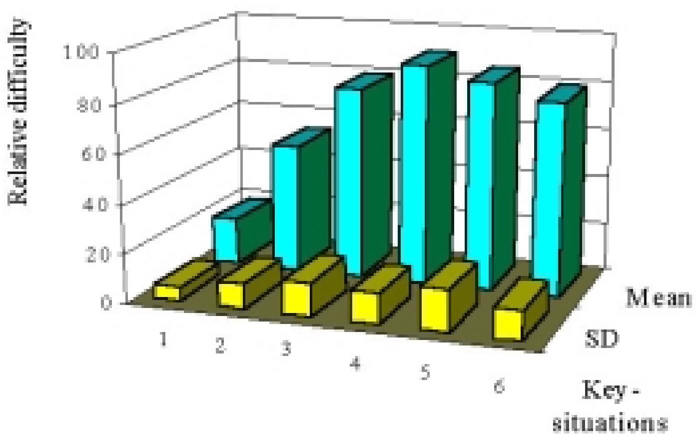


*Fig. 5.* Mean and SD of relative perceived difficulties of situations

d) A closer look to the group decision matrix of crew No. 4 with unacceptable low degree of agreement in *Table* 2 (see also crew No. 4 on *Fig.* 4) revealed that the actual causes of global disagreement could be attributed mainly to different opinions about Key situation 2 (K2), because in column K2 there are one 2 and three 3s. In this matrix a 2 (or 3) represents that Key situation corresponding to the row of that 2 (or 3) is judged to be more difficult by 2 (or 3) operators out of the five than Key situation corresponding to the column of that 2 (or 3). As frequencies of judgements distributed in a 2:3 ratio thus represent the highest possible disagreement in a 5-member crew, the root cause of disagreement probably has something to do with Key situation 2.

*Table 2.* Aggregated matrix of difficulties of Crew No. 4 showing unacceptably low degree of agreement ($U = 0.12$, $p > 0.05$), attributable mainly to judgements on Key situation 2

| Key situation | K1 | K2 | K3 | K4 | K5 | K6 | $\sum$ |
|---|---|---|---|---|---|---|---|
| K1 | – | 1 | 1 | 0 | 1 | 1 | 4 |
| K2 | 4 | – | 2 | 2 | 3 | 2 | 13 |
| K3 | 4 | 3 | – | 1 | 4 | 2 | 14 |
| K4 | 5 | 3 | 4 | – | 4 | 4 | 20 |
| K5 | 4 | 2 | 1 | 1 | – | 1 | 9 |
| K6 | 4 | 3 | 3 | 1 | 4 | – | 15 |

e) Delving further into the level of individual operators of Crew No. 4 analysing individual matrices of difficulties identified – among other things – that though both RO (*Table* 3) and TM (*Table* 4) had individually consistent opinions they were of largely conflicting opinions.

*Table 3.* Individual matrix of difficulties of RO of Crew No. 4

| Key situation | K1 | K2 | K3 | K4 | K5 | K6 | $\sum$ |
|---|---|---|---|---|---|---|---|
| K1 | – | 0 | 0 | 0 | 0 | 0 | 0 |
| K2 | 1 | – | 0 | 0 | 1 | 1 | 3 |
| K3 | 1 | 1 | – | 1 | 1 | 1 | 5 |
| K4 | 1 | 1 | 0 | – | 1 | 1 | 4 |
| K5 | 1 | 0 | 0 | 0 | – | 1 | 2 |
| K6 | 1 | 0 | 0 | 0 | 0 | – | 1 |

It was proved that though the instruction of COSMOS asked for assessing difficulties from the point of view of the crew as a whole, RO and TM in this case clearly made their assessments from their particular narrow professional view: RO's decisions reflected a definitely primary circuit character, while TM's decisions had a secondary circuit flavour. As in our view the basic unit of safety on the Human Factors side is not the individual operator but the crew as a whole, we designed the

*Table 4.* Individual matrix of difficulties of TM of Crew No. 4

| Key situation | K1 | K2 | K3 | K4 | K5 | K6 | ∑ |
|---|---|---|---|---|---|---|---|
| K1 | – | 1 | 1 | 0 | 1 | 1 | 4 |
| K2 | 0 | – | 0 | 0 | 0 | 0 | 0 |
| K3 | 0 | 1 | – | 0 | 1 | 0 | 2 |
| K4 | 1 | 1 | 1 | – | 1 | 1 | 5 |
| K5 | 0 | 1 | 0 | 0 | – | 0 | 1 |
| K6 | 0 | 1 | 1 | 0 | 1 | – | 3 |

COSMOS so that it can detect too narrow professional thinking and reasoning and can modify it in a more desirable direction.

f) On the basis of relative difficulties it was possible to identify some details of technological problems that represented major difficulties for the crews and so was also possible to design special teaching programs to increase operators' preparedness in this field.

g) On the basis of distribution of decision loops along key situations it was possible to identify ambiguously defined key situation which had different meaning to different crew members (*Fig.* 6).

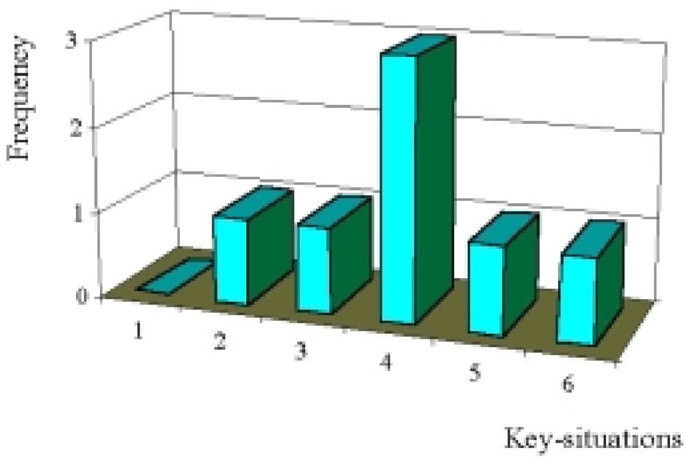


*Fig. 6.* Ambiguity in the interpretation of situations by the operators indicated by the number of loops

h) In accordance with the expectation it was proved that Shift Supervisors (SS) – heads of crews – were the most consistent in their judgements (*Fig.* 7).

i) On the basis of frequencies of judgements distributed in 2:3 ratio – the

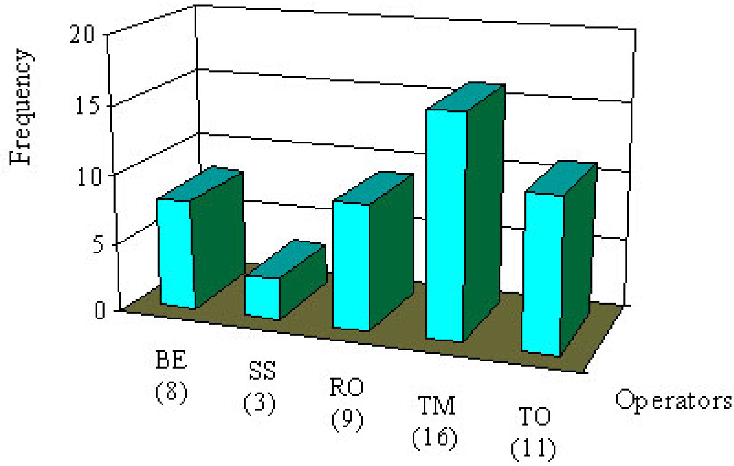


*Fig. 7.* Level of consistency in assessing key situations' perceived difficulties indicated by the number of loops made by various operators

possible highest disagreement in a 5-member crew – it was possible to identify those key situations on which crew members had different – but individually consistent – opinions (*Fig.* 8).

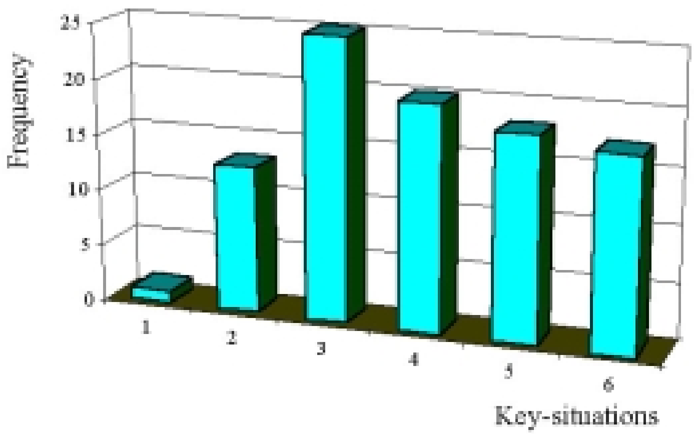


*Fig. 8.* Frequency of opinions distributed in 2:3 ratio as a result of pair-wise comparison of difficulties of situations

j) Different operator posts received different average expected roles. One extreme is the BE post, assessed consistently as having low expected roles, which have the methodological consequence that training scenarios should be designed so that BEs also have important tasks. TOs have relatively high SD which may have indicated different points of view within different crews.

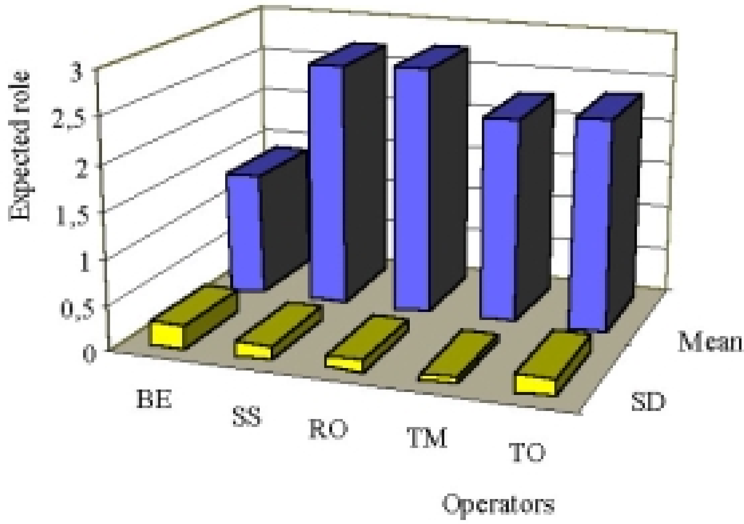


*Fig. 9*. Mean and SD of expected roles of different operators assessed by the members of different crews

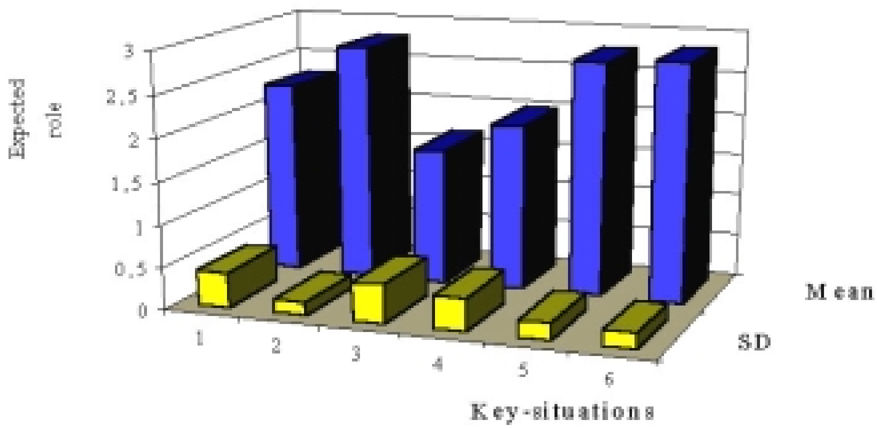


*Fig. 10*. Mean and SD of expected roles of TOs as a function of key situation assessed by the members of different crews

k) Going a bit deeper from global data about TOs presented in *Fig.*9, expected roles of TOs were analysed as a function of key situation (*Fig.* 10). It can be mentioned, for example, that in Key situation 3 the expected role of TOs was assessed the lowest but with the highest SD, which in the light of some technological details was possible to explain. This type of differences in perceived expected roles may also reveal different interpretations of the same procedure in different crews. Some procedures definitely allow different interpretations.

l) In general, crew members were able to utilise feedback from the computer and as a consequence it sometimes modified their original opinions.

As a conclusion it can be stated that the COSMOS method has proved to be advantageously usable during the first series of real application during simulator training: it helped to identify sources of problems of misunderstandings as well as disagreements, and also provided operators with quick and meaningful feedback supporting a sped up learning process both in the technological domain and in the social context.

## References


[1] ANTALOVITS, M. – IZSÓ, L. (1999): Self-assessment and Learning in Nuclear Power Plant Simulation Training. In: Misumi, Wilpert, Miller (eds), Nuclear Safety: A Human Factors Perspective, pp. 243–256. Taylor & Francis, London.

[2] ANTALOVITS, M. – IZSÓ, L. – JENEI, CS. (1995): New Method for Performance Assessment in NPP Training Simulator. *Seventh European Congress on Work and Organizational Psychology, Győr, Hungary. Proceedings,* p. A84. (19–22 April, 1995).

[3] ANTALOVITS, M. – IZSÓ, L. – JENEI, CS. (1995): New Method for Performance Assessment in NPP Training Simulator. *Seventh European Congress on Work and Organizational Psychology, Győr, Hungary. Proceedings,* p. A84. (19–22 April, 1995).

[4] ANTALOVITS, M. – IZSÓ, L. – TAKÁCS, I. (1995): A Frame for Evaluation and Self-evaluation of NPP Personnel. *Seventh European Congress on Work and Organizational Psychology, Győr, Hungary. Proceedings,* p. A14. (19–22 April, 1995).

[5] HAJTMAN, B. – IZSÓ, L. – RADINSZKY, A. (1990): Multivariate Ranking Based on Subjective Classifications. *XVth International Biometric Conference* (2–6. July 1990), Budapest.

[6] IAEA (International Atomic Energy Agency) 1989: Guidebook on Training to Establish and Maintain the Qualification and Competence of Nuclear Power Plant Operations Personnel (IAEA-TECDOC-525). Vienna, IAEA.

[7] IAEA (International Atomic Energy Agency) 1994: Role and Responsibilities of Management in NPP Personnel Training and Competence. Presentations for the seminar of the IAEA model project in Hungary: Strengthening training for operational safety at Paks NPP (IAEA-TC-UN/9/01/1). Vienna, IAEA.

[8] IZSÓ, L. – ANTALOVITS, M. (1994): What Are the Basic Aims, Criteria and Methods of Developing Application Methodology for a Full Scope Simulator? *First International Conference on Human Factors Research in Nuclear Power Operation, Berlin, Germany. Conference Report,* V. Session: Simulation, pp. 1–2. (October 31 – November 2, 1994).

[9] IZSÓ, L. – ANTALOVITS, M. (1996): Developing Computer Supported Method for Operators' Self-assessment ('COSMOS') for improving team work. *Research and Study on Human Factors Workshop.* Tokyo Electric Power Company Nuclear Power Research and Development Centre, Yokohama (January 17, 1996).

[10] IZSÓ, L. – ANTALOVITS, M. (1997): Results of a Validation Study of the Method COSMOS in NPP Simulator Sessions. Paper presented at the *13th Triennial Congress of the International Ergonomics Association. Proceedings* Vol. 7 pp. 231–233, June 29 – July 4, 1997, Tampere, Finland.

[11] KENDALL, M. G. (1948): Rank Correlation Methods. Griffin, London, 1948.