Note:
This is a translation of the RSK recommendation entitled
"Bewertung der Wirksamkeit von Maßnahmen zur Vermeidung der Wiederholung von Ereignissen"
In case of discrepancies between the English translation and the German original, the original shall prevail.

RSK recommendation
(512th meeting of the Reactor Safety Commission (RSK) on 22/23 October 2019)

Assessing the effectiveness of measures to prevent recurrence of events

RECOMMENDATION

1 Introduction and consultations

The RSK Committee on REACTOR OPERATION (RB) regularly deals with the assessment of national and international events. In-depth discussions on individual events often include the presentation of the results of integrated event analyses. In the related consultations, the RB Committee has come to the conclusion that there is a potential for improvement with regard to the derivation of appropriate measures and the effectiveness control of defined measures. For example, in the case of recognisable missing appropriateness of defined measures with regard to contributing factors and for checking the effectiveness of defined measures against recurrence, specific inquiries of the RB Committee could often not be answered or could only be answered unsatisfactorily.

The RB Committee therefore looked at this issue in more detail. For this purpose, at the 246th meeting of the RSK Committee RB on 17 November 2016, a working group on integrated event analysis – effectiveness evaluation (GANZHEITLICHE EREIGNISANALYSE – WIRKSAMKEITSEVALUATION – AG GEAW) was charged with preparing a draft statement. At the first meeting of the working group on 20 March 2017, the working group members agreed on gathering information on how effectiveness tests and evaluations should be carried out and which national and international requirements and specifications exist for carrying out effectiveness evaluations for a subsequent comparison with the practical approach taken by the operators. In dealing with the current practice of establishing processes for the derivation of corrective actions and their effectiveness evaluation, the working group decided at its second meeting on 27 April 2017 to ask VGB Powertech e. V. (VGB) for a presentation on the current practice at the plants. The requested presentation by VGB took place at the third meeting of the working group on 10 October 2017 and was based on an RSK catalogue of questions sent in advance. At the fourth and fifth meeting of the working group on 7 December 2017 and 1 February 2018, the working group prepared a first draft statement which was adopted by the RB Committee at its 262nd meeting on 18 October 2018. The draft statement adopted by the RB Committee was first submitted to the RSK for discussion at its 511th meeting on 4 September 2019 and was adopted at its 512th meeting on 22/23 October 2019.

2 Assessment criteria

2.1 Relevant documents

In a first step, the relevant documents of the IAEA as well as further guiding documents, such as the Safety Requirements for Nuclear Power Plants as well as the RSK recommendation "Guideline for the performance of integrated event analyses" [1] and the VGB guideline on the integrated event analysis [2] were analysed with regard to whether they describe sufficiently precisely how an effectiveness evaluation is to be performed. The following documents were taken into account:

- IAEA-TECDOC-1458 Effective corrective actions to enhance operational safety of nuclear installations,
- IAEA-TECDOC-1580 Best Practices in the Utilization and Dissemination of Operating Experience at Nuclear Power Plants,
- IAEA-TECDOC-1653 Best Practices in the Management of an Operating Experience Programme at Nuclear Power Plants,
- IAEA Safety Standards Series No. SSR2/2 (Rev. 1) Safety of Nuclear Power Plants: Commissioning and Operation,
- Sicherheitsanforderungen an Kernkraftwerke (SiAnf) *(Safety Requirements for Nuclear Power Plants)*: Appendix 1(2)/1(3),
- KTA 1402: "Integriertes Managementsystem zum sicheren Betrieb von Kernkraftwerken" *(Integrated Management System for the Safe Operation of Nuclear Power Plants)*,
- Richtlinie über die Gewährleistung der notwendigen Kenntnisse der beim Betrieb von Kernkraftwerken sonst tätigen Personen *(Guideline relating to the assurance of the necessary knowledge of the persons otherwise engaged in the operation of nuclear power plants)*,
- RSK recommendation "Leitfaden für die Durchführung von ganzheitlichen Ereignisanalysen" *(Guideline for the performance of integrated event analyses)*, and
- VGB Powertech e. V.: "Leitfaden Ganzheitliche Ereignisanalyse" *(Guideline on the integrated event analysis)* as of October 2015).

Table 1 shows the main references to the effectiveness evaluation from the IAEA documents, the SiAnf, KTA 1402, the RSK recommendation on the performance of integrated event analyses and the VGB guideline on the integrated event analysis:

Table 1: Requirements and descriptions from national and international (IAEA) documents

Document	Relevant passages and/or specifications in the respective document
SSR 2/2 Rev 1	Requirement 24, para 5,30: *"Corrective actions shall be…reviewed for their effectiveness"*
IAEA TEC DOC 1458	Effective Corrective Actions (in particular Section 5.5. Tracking effectiveness of corrective actions)
IAEA TEC DOC 1580	Best Practices in the Utilization and Dissemination of Operating Experience at Nuclear Power Plants (in particular Section 3.4. Utilization of OE to improve human performance and Section 3.5. Utilization of OE in the decision making process)
IAEA TEC DOC 1653	Best Practices in the Management of an Operating Experience Programme at Nuclear Power Plants (in particular Section 2.4. Programme results and effectiveness reviews)
SiAnf	1 (2) Responsibility of the plant management: *"Within the framework of its responsibility, the plant management shall fulfil in particular the following requirements:* *.../"* 9. Registration, evaluation, utilisation and communication of internal and external experiences. 1 (3) Integrated management system (IMS): *"Regulations shall be made at least for the following processes:* */../"* execution of the operating experience feedback, planning and execution of internal and external communication
KTA 1402	3.6 *"Plan-Do-Check-Act Cycle* *With regard to a continuous improvement, the plan-do-check-act cycle (PDCA cycle) shall be applied to all relevant operational activities, to partial and entire processes and to the management system as a whole.* *Note:* *The systematic approach to be applied in this context comprises the planning, analysis and description of operational procedures, the implementation of specified guidelines, checking the effectiveness on the basis of definite criteria and, as far as necessary, the corrections directed at optimizing these procedures.* 5.13.2 Plant-internal experience feedback: (4) *"Modification measures shall be specified on the basis of the analysis results. The implementation of these measures shall be surveilled and their effectiveness evaluated through suitable methods."* (5) *"The causes determined from these analyses shall be subjected to a trend analysis in order to enable an early detection and prevention of a possibly occurring accumulation of these causes."*
GMBI – Guideline relating to the	3.3 Knowledge transfer and maintenance, ensuring effectiveness for own and external personnel, recognition

assurance of the necessary knowledge of the persons otherwise engaged in the operation of nuclear power plants	C. Ensuring effectiveness: *“”...The control of the safety-oriented behaviour at the workplace by the respective superior also serves to prove the effectiveness”*
VGB – Guideline “Integrated event analysis” of October 2015”	11. Implementation of measures and effectiveness control C. Ensuring effectiveness: *“”...The control of the safety-oriented behaviour at the workplace by the respective superior also serves to prove the effectiveness”* 11.2 Effectiveness control of the implemented measures: *“The effectiveness of the implemented measure is to be controlled in a predefined time window and in a predefined manner.”* 11.2 Effectiveness control of the implemented measures: *“If the measure does not have the intended effect, an appropriate measure identification / decision process is to be re-launched and documented.”*
RSK recommendation Guideline for the performance of integrated event analyses	9 Evaluation of the event analysis: *“If the event analysis is considered to be completed, the management orders the implementation of the individual proposed measures and arranges for the control of the implementation and effectiveness.”* 11 Implementation of measures and effectiveness control: *“The implementation of the defined measures is to be monitored with regard to their* • *execution,* • *completeness, and* • *correctness.”* 11.1 Control of measure implementation: *“If deviations are found in the implementation of measures, these must be justified in writing and, if necessary, an appropriate measure identification / decision process is to be re-launched.”* 11.2 Effectiveness control of the implemented measures: *“The effectiveness of the implemented measure is to be controlled in a predefined time window and in a predefined manner. If the measure does not have the intended effect, an appropriate measure identification / decision process is to be re-launched.”*

IAEA-TECDOC-1458 “Effective corrective actions to enhance operational safety of nuclear installations” gives a comprehensive insight into the necessary standards. The “Corrective Actions Programme”, which is presented in the annex to IAEA-TECDOC-1458 in the form of a flow chart, contains a recommendation for the implementation of an effectiveness evaluation of initiated measures from events at all organisational levels of a nuclear power plant. The indicators contained in IAEA-TECDOC-1458 show one of the possibilities which are suitable for checking the effectiveness of the experience feedback process.

Chapter 1 of the Safety Requirements for Nuclear Power Plants, for example, stipulates that the company management shall ensure that the internal and external feedback of experience, changes in the state of the art in science and technology and in proven international safety practices including the associated information arranged by authorities is systematically registered, evaluated and documented in a process of the management system. Chapter 1 (2), requires that the plant management shall ensure, among other things, registration, evaluation, utilisation and communication of internal and external experience. Here, the plant management shall be mindful that regarding the internal feedback of experience, information about nearmiss events shall be given special attention. The above requirements are further specified in KTA 1402 “Integrated Management System for the Safe Operation of Nuclear Power Plants”.

With regard to the internal flow of experience, root cause analyses are to be carried out, taking into account all aspects of the areas of man, technology and organisation and their interaction. The procedure for conducting the analyses is to be regulated.

Furthermore, improvement measures are to be defined on the basis of the analysis results, their implementation is to be monitored and the causes identified from the analysis are to be subjected to trend monitoring in order to detect and eliminate accumulation of causes in time.

Also with regard to the RSK recommendation “Guideline for the performance of integrated event analyses” and the VGB guideline on the integrated event analysis, the RSK comes to the conclusion that both the RSK recommendation on the integrated event analysis and the VGB guideline on the integrated event analysis already contain general specifications for checking the effectiveness of derived measures. For the professional planning, structuring and performance of effectiveness evaluations, the expert knowledge on the state of the art in science and technology with regard to evaluation methods, tools and packages of measures is also decisive.

2.2 Effectiveness evaluation methods

The working group evaluated recent scientific publications with regard to effectiveness evaluations. These publications show learning cycles that organisations have to go through in order to learn collectively from events as a socio-technical system. The state of the art in science and technology provides questionnaire instruments that allow assessing an organisation's learning ability. Especially relevant seems to be the requirement that experiences made and measures implemented should also bring about sustained and permanent changes in the behaviour of all those organisation members who had not been directly involved in the event.

In this respect, it has to be taken into account that the appropriate fit between the corrective action and the underlying event cause or contributing factor must be the central starting point. If measures are implemented that do not address the actual cause, a mere evaluation of the measure (Has the measure been implemented? Were the staff members satisfied with the training?) cannot lead to a statement about the effect of the measure with regard to the cause to be addressed.

In the context of an integrated event analysis, a thorough evaluation of effectiveness is therefore dependent on the completeness of a root cause analysis, which is the prerequisite for the deduction of appropriate corrective actions.

Particularly in the area of qualification and leadership development, the effectiveness of corrective actions is to be monitored, tracked and assessed in accordance with state-of-the-art evaluation methods at four different levels.

- Level 1 (reaction level): level of satisfaction, acceptance of the training,
- Level 2 (learning level): extent to which knowledge, skills and attitudes are acquired and maintained,
- Level 3 (behavioural level): degree of work-related transfer and application, and
- Level 4 (result level): degree of (quantifiable) change in the organisation.

For these four levels, appropriate points in time are to be determined at which the effectiveness of the derived measures to achieve the targeted goal is checked. The points in time can vary significantly between days (e.g. Level 1), months (e.g. Level 2/3) or years (e.g. Level 4). For example, the reaction level can be checked immediately at the end of a training course. What has been learned can also be tested immediately after the end of the training. However, the permanent application of what has been learned and its transfer to work with its varying work environment conditions can only be assessed after weeks or months. Finally, the extent to which the entire organisation has learned from events can only be measured after several months or even years. Transfer means the application and retention of what has been learned in training or in a training course at work as well as the adaptation of what has been learned to situations not addressed in training and thus to new situations.

The appropriate timing for the evaluation levels is derived from the transfer model of Baldwin and Ford (1988) and Baldwin, Ford and Blume (2009, Figure 1). Positive transfer of training refers to the extent to which learning that has taken place in training is transferred to the work activity and contributes to significant changes in (safety-related) work performance.

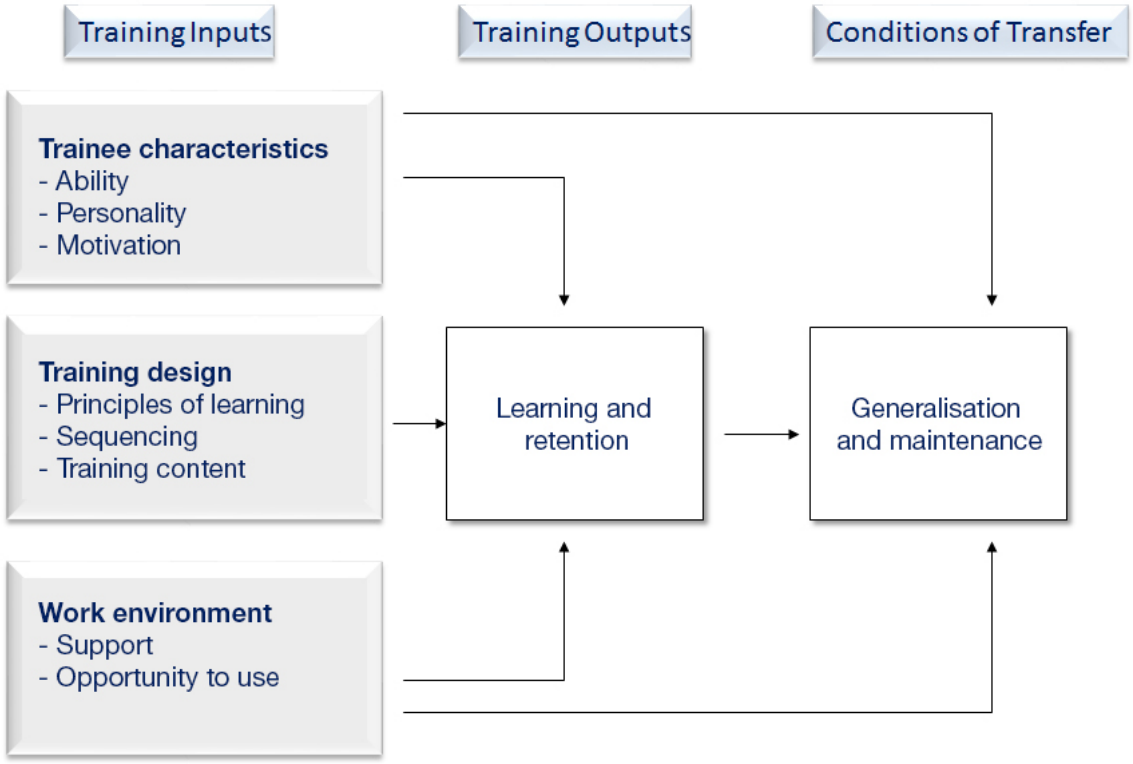

Figure 1: Baldwin and Ford's transfer of training model (1988)

This model illustrates that evaluation must not take place at the end of the training (e.g. in the sense of a final feedback) but that transfer can only be referred to when the actual application of the training content has been measured at a point in time in the context of work *after* the training.

Boundary conditions that promote the transfer of training are actively requesting and reaffirming the application of the training contents by superiors and colleagues and the support of the application through organisational and technical additional measures (e.g. equipment, checklists, time resources made available). The assessment of the effect of corrective actions can serve two objectives. On the one hand, it can be carried out in order to be able to intervene for adjustment or optimisation. On the other hand, it also serves as a final proof of the effect.

From a scientific perspective, the framework model for achieving these two objectives is formative and summative evaluation.

In formative evaluation, continuous feedback is already obtained *during* the development, piloting and implementation of a measure, and the extent to which the objectives of the measure are adequately considered is reflected upon with the participants. Within the scope of a formative evaluation, it is for example examined with training providers to what extent an existing training offer needs to be adapted to the objectives. For example, training documents are checked for comprehensibility by the participants *in advance* in the sense of a pre-test, or possible transfer obstacles are discussed with the participants *prior to* the training. The findings

of the formative evaluation can then be used immediately and already during the development or implementation phase to optimise measures.

Summative evaluation is the measurement of the effect *after* completion of the measure, e.g. after a training measure has been carried out and with direct reference to the previously defined training objectives. A focus on summative evaluation entails the risk that it will only be determined afterwards that a measure has no or not the defined effect and that the development of a new measure must then be started.

3 Status report by the operators

VGB was asked for a report on how the practical implementation of the effectiveness evaluation within the scope of the implementation of measures related to the integrated event analysis is carried out as commonly practised by the operators. For this purpose, a catalogue of questions was developed which was to pre-structure the VGB answers and to narrow down the focus of interest. The questions of the RSK answered by VGB during a hearing on 10.10.2017 were as follows:

- Which practical examples from the plants exist for the evaluation of the effectiveness of

  - technical measures?
  - organisational measures?
  - behavioural measures?
  - measures related to technical qualification?

- How are execution, correctness and completeness as well as the effectiveness of a measure evaluated in practice?

- How is the time window in which the effect is observed/controlled defined in practice?

- What practical examples are there of indicators used by operators to assess a measure as “effective”?

- Which indicators are used to detect the recurrence of similar events with similar contributing factors?

- In case the desired effect does not occur, how is this addressed in practice?

- In practice, how is the process of identifying measures and making decisions reconsidered and revised if a measure does not have the intended effect?

- How is the effectiveness evaluation process itself reflected for its effect in practice?

- In practice, which bodies discuss/decide on the effectiveness of measures?

- What specifications on “effectiveness controls” are there in the process descriptions? In which processes are such specifications included?

VGB explained in its presentation that more detailed information on the performance of effectiveness controls had been made available in the plants in written or software-based form, e.g. in the form of a newly developed guideline in the SOL-VE software. The selection of the methodology/effectiveness check to be applied was based on the proposal of the experienced event analyst with the involvement of a decision-making body. These paper- and software-based tools queried, among other things, how the effectiveness of measures can be demonstrated and whether a suitable indicator including target value is available for the assessment of effectiveness.

The presented screenshots of the software-based support showed the process of tracking in particular, i.e. the scheduling of a measure (controller, agreed date for control, control status, actual control date, control note) by a responsible person (see Figure 2). With reference to the SOL-VE software, VGB reported that in 2015, the VGB guideline on the integrated event analysis had been supplemented by another document, which specifically addressed measures for effectiveness checks and which were implemented accordingly by the operators. All sites use the VGB guideline on the integrated event analysis and the supplemented guideline for effectiveness checks, which was integrated into the SOL-VE software in spring 2017.

The VGB guideline on effectiveness checks has not been made available to the RSK.

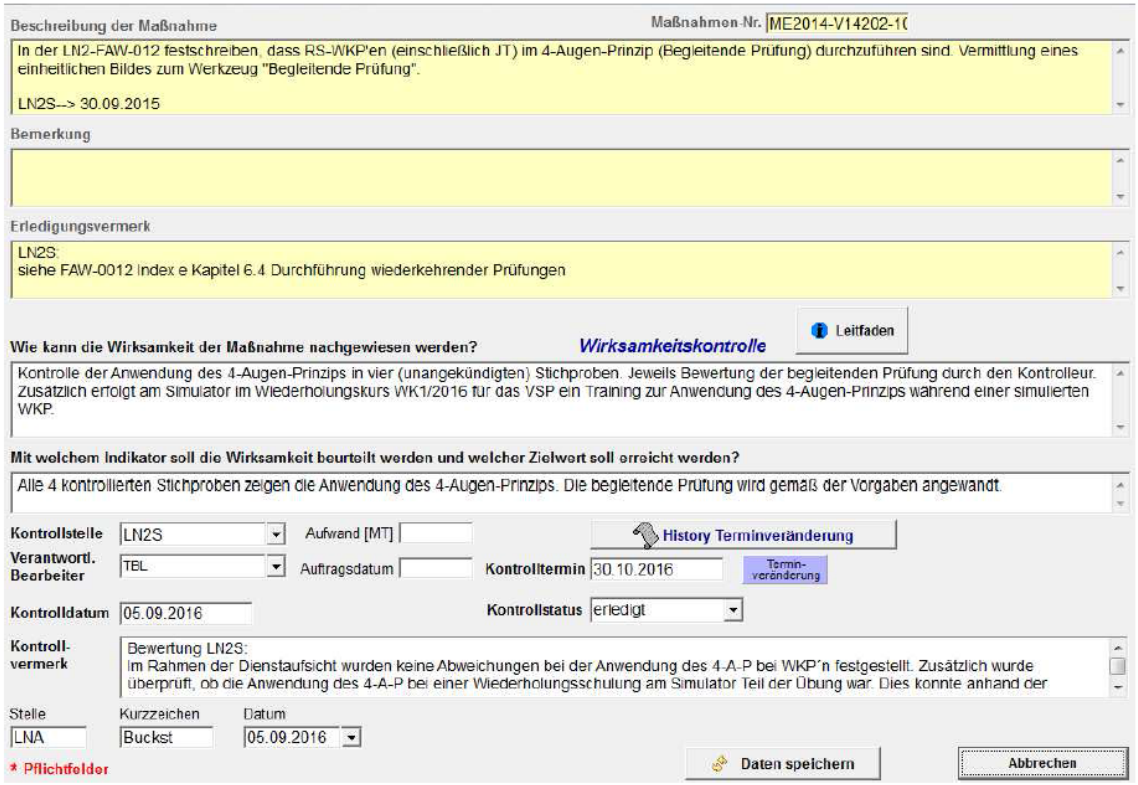

Figure 2: Software-based support in effectiveness checks (slide 13) according to VGB

The evaluation of effectiveness comprised three stages:

1 Assessment of measures prior to their implementation, (analysts, measure-related workshop, decision-making body) and follow-up/control of the implementation according to SMART[1] criteria.

2 Assessment of measures after their implementation (classic measure effectiveness check):

2.1 Implementation control of the measures regarding correct execution and completeness.

2.2 Checking the effectiveness of the implemented measures
- Control according to a reasonable time window.
- Has the measure reached its objective or will it reach the objective at a later time?
- Type of check according to the type of measure (technical, organisational, behavioural).

3 Assessment of the entire process with regard to the effectiveness of measures (long-term management task), for example in:

- management review,
- VGB-SBS (VGB safety culture assessment system),
- audits, peer reviews, and
- exchange of experience of nuclear safety officers.

VGB presented several examples of effectiveness checks of technical, technical/organisational, organisational and behavioural measures and measures related to technical qualification carried out in practice.

The organisational measures described were evaluated for their effectiveness, for example in the context of meetings and observations during on-site implementation, e.g. through unannounced spot checks and observations during simulator training.

Behavioural measures were evaluated through observations (e.g. by “managers in the field”: MIF) including supervision. In addition, various examples of training feedback were presented, e.g. feedback from training participants directly after a training course and evaluated by the trainer or instructor, interviewing persons in charge of the training, the participants’ assessment regarding the relevance of training contents.

The following indicators to assess a measure as effective were mentioned:

- successful implementation of ISIs,
- interviews,
- meetings (establishing common understanding),
- monitoring of activities (MIF),
- no event recurrence,
- results of special audits,

[1] S= Specific, M = Measurable, A= Acceptable, R = Realistic, T= Time-bound

- application of trained procedures in practice (MIF),
- feedback from briefings and debriefings,
- assessment by experts,
- simulating similar situations using the simulator, including evaluation,
- results of the VGB-SBS for comprehensive change.

The definition of the time window in which the effect is considered was dependent on the respective event and in relation to its significance, the intervals at which the activity is carried out, the point in time at which experience can be available and the appropriate specification, e.g. what was to be observed during the operating period.

The effectiveness of corrective actions was discussed in the following "bodies": event analysis, affected staff within the framework of measure-related workshops, management team, feedback from the team (e.g. in the form of debriefings after activities have been carried out), decision-making body.

Based on the knowledge why a measure does not work, optimisation of the original measure took place, if necessary:

- provision of further background information to increase the acceptance of a measure,
- cancellation of the measure, if dispensable, and
- determination of new measures to replace them.

If the measure did not have the intended effect, an appropriate measure identification / decision process was to be re-launched (VGB guideline on the integrated event analysis), i.e. as a rule, return to the event analysis and again passing through the process of the integrated event analysis for this partial aspect was necessary.

The following indicators were used to identify the recurrence of similar events:

- technical qualification of the event analysts / event analysis,
- consistent analysis with regard to operating experience, predecessors, etc. for every integrated event analysis,
- events collected in database → comparisons/evaluations possible – coding,
- failure statistics / ageing management, and
- trend analyses.

The effectiveness of the process for evaluation of the effectiveness of corrective actions was not completely checked, but individual elements were checked in audits, management reviews, VGB-SBS, peer reviews and within the framework of the exchange of experience of nuclear safety officers.

In the concluding remarks, the VGB stated that the effectiveness check was in principle considered a good means to check the actual effect of measures resulting from the analyses. In practice, however, there were challenges with regard to the aspects mentioned below. The available time horizon for proving that the measure is successful in preventing recurrence of an event was possibly too short (e.g. in view of the remaining lifetime

of the plant) and therefore the question may arise as to whether such a measure was useful. In addition, it was desirable that the measures for checking effectiveness should not exceed the expenditure and effort associated with the measure itself. If the effort required to evaluate the effectiveness of measures was disproportionate, acceptance would be reduced. From this, VGB deduces that an effectiveness check measure should be in a reasonable proportion to the measure, i.e. in the case of a minor technical measure a quick check. Moreover, formal effectiveness checks should be carried out in the future in the case of in-depth analyses or selected basic analyses.

According to VGB, an "effectiveness check measure" is sometimes perceived difficult or impossible to define, e.g. for the measure "insertion of an additional control column". In addition, significant M/O measures should be checked in the future, if possible, with established tools such as audits, MiF and the VGB safety culture assessment system.

When asked, the VGB explained that the processes for effectiveness evaluation are not yet anchored in the management system.

4 Assessment

4.1 Requirements from rules and regulations

With regard to the evaluation of the documents referred to in Chapter 2, the RSK comes to the conclusion that the national and international regulations, guidelines or standards sufficiently describe the basic requirements for performing an adequate effectiveness evaluation in generic form. However, the RSK sees the need to further specify the existing generic requirements for effectiveness evaluation and therefore derives the following recommendations.

4.2 Implementation of the state of the art in science and technology in the field of effectiveness evaluation

In their presentation, the operators explained that from their point of view, too, it is necessary to evaluate the effectiveness of corrective actions. The operators have shown that they use different methods to assess the effectiveness of corrective actions, depending on the respective circumstances.

Regarding the indicators mentioned by the VGB by means of which the effectiveness of measures should be assessed, it is stated from the point of view of the RSK, however, that methods (interviews, observations) and results (e.g. no event recurrence) are mixed here. In this context, the method (observation, interviews, assessments by experts) only represents the "measurement procedure" where no distinction is made between survey or measurement procedures on the one hand and measurement results on the other hand. Thus, the list includes not only indicators but also methods for the collection of indicators.

The RSK acknowledges that the operators basically take the requirements from the rules and regulations into account. Nevertheless, taking into account the advanced state of the art in evaluation science, tools and

packages of measures, the RSK comes to the conclusion that there is further potential for improvement with regard to the evaluation methods applied by the operators.

This concerns, for example,

- taking account of the distinction between formative and summative evaluation and the existing contextual connection, which can be optimised from the point of view of the RSK,

- the methods used, with a focus on their information value for the assessment of effectiveness as well as their specific features to be taken into account with regard to the appropriate timing of the evaluation of an effect, and

- better use of results as feedback information in the ongoing process of designing and adapting measures to achieve the intended effect.

Furthermore, the design and scope of the evaluation of a corrective action to prevent the recurrence of events should not be based on the scope of the corrective action but on their safety relevance.

In current practice, evaluation measures in the field of human-related measures are often designed in such a way that they seek confirmation of the intended effect (e.g. by collecting reinforcing feedback from the training participants directly after the training). In contrast, an open effectiveness check, which aims at detecting a possible lack of effect, reduces the risk of the so-called confirmation bias. In this context, confirmation bias means that the person responsible for checking effectiveness may tend to consider and seek only those pieces of information that confirm his or her expectation of successful implementation of measures.

In contrast, the observations of behavioural changes in simulation environments (e.g. plant simulator, practical training on occupational safety, field simulator, component models (mock-up), etc.) are in the RSK's view well suited as training feedback and to check the effectiveness of a measure. Here, however, the transfer from the simulation environment to the real environment has to be taken into account.

The RSK also sees potential for improvement in the area of "pattern/trend recognition" and in the "derivation of general measures". This concerns e.g. the integration of information from different sources and experience of the past.

In summary, it can be stated that the RSK sees potential for improvement with regard to the methods used in current practice to verify the achievement of set objectives on all four evaluation levels defined above in section 2 by means of appropriate evaluation tools (questionnaire, observations), methods (qualitative, quantitative) and measures (e.g. time series designs), taking into account the state of the art in science and technology.

The RSK is of the opinion that especially human- and organisation-related measures require high competence for the selection of appropriate evaluation tools, methods and packages of measures. The combination of qualitative and quantitative methods such to achieve the set objectives and the use of existing procedures

require a balanced mix of methods. If, for example, management tools such as “managers in the field” are used for evaluation purposes, these should be adapted and adjusted to the evaluation objective, e.g. with a special briefing of the managers and be supported by other e.g. quantitative tools. Furthermore, experience shows that only using indicators in the management system specifically for the assessment of effectiveness is insufficient.

The RSK therefore recommends updating knowledge on the state of the art in science and technology with respect to evaluation methods, tools and packages of measures available within the organisation in a systematic and structured manner (see recommendation 1).

On the basis of KTA 1402 (in particular Chapter 3 and Section 5.13), the processes for root cause analyses and for the subsequent implementation of derived corrective actions up to checking the effectiveness of defined improvement measures are also to be implemented in the management system, taking into account the state of the art in science and technology (see recommendation 2).

To prevent confirmation bias, a clear distinction should be made between the person(s) responsible for the implementation of the defined corrective action and the person(s) responsible for the effectiveness evaluation of this corrective action (see recommendation 3).

Furthermore, from the point of view of the RSK, not only changes initiated on the basis of event analyses but all changes that may lead to safety-relevant impacts (e.g. organisational and technical changes or desired changes in behaviour) should be systematically checked according to the principles of the PDCA cycle to ensure that the objective that the change does not have a detrimental effect on safety is achieved.

5 Recommendations

Recommendation 1: For effective and efficient implementation of effectiveness evaluations of defined corrective and improvement measures and integration of effectiveness evaluations into the continuous improvement process, knowledge on the state of the art in science and technology as regards evaluation methods, tools and packages of measures is to be developed and maintained within the organisation in a systematic and structured manner. Suitable tools to support the implementation of effective and efficient effectiveness evaluations are to be created and kept up to date in the organisation.

Recommendation 2: On the basis of nuclear safety standard KTA 1402 (in particular Chapter 3 and Section 5.13), the process for root cause analyses and for the subsequent implementation of derived corrective actions up to checking the effectiveness of defined improvement measures is to be implemented in the management system, taking into account recommendation 1.

Recommendation 3: To prevent confirmation bias, a clear distinction is to be made between the person(s) responsible for the performance of the defined corrective action and the person(s) responsible for the effectiveness evaluation of this corrective action.

6 References

[1] Leitfaden für die Durchführung von ganzheitlichen Ereignisanalysen
RSK-Empfehlung vom 06.11.2014

[2] VGB Powertech e. V.: „Leitfaden Ganzheitliche Ereignisanalyse“ (Stand: Oktober 2015)

[3] Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick’s levels of training criteria: Thirty years later. Personnel Psychology, 42, 331-342

[4] Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. Personnel Psychology, 50, 341-358

[5] Arthur, W., Bennett, 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, 88(2), 234-245

[6] Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. Personnel Psychology, 41(1), 63-105

[7] Baldwin, T.T., Ford, K.J. & Blume, B. D. (2009). Transfer of Training 1988-2008: An updated Review and Agenda for future research. International Review of Industrial and Organizational Psychology, 24, 41-70

[8] Bergmann, B. & Sonntag, Kh. (2006). Transfer erworbener Kompetenzen in den Arbeitsalltag. In Kh. Sonntag (Hrsg.), Personalentwicklung in Organisationen. 3. Auflage (S. 355-388). Göttingen: Hogrefe

[9] 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

[10] Carroll, J. S., & Fahlbruch, B. (2011). “The gift of failure: New approaches to analyzing and learning from events and near-misses.” Honoring the contributions of Bernhard Wilpert. Safety science, 49(1), 1-4

[11] Dechy, N., Dien, Y., Drupsteen, L., Felicio, A., Cunha, C., Roed-Larsen, S., & Vetere Arellano, A. L. (2015). Barriers to learning from incidents and accidents. European Safety Reliability and Data Association (ESReDa) http://repository.tudelft.nl/view/tno/uuid:8aa15026-1f7c-4d77-8e91-3b60ff777ca9/, retrieved 30.3.2017

[12] Drupsteen, L., & Guldenmund, F. W. (2014). What is learning? A review of the safety literature to define learning from incidents, accidents and disasters. Journal of Contingencies and Crisis Management, 22(2), 81-96

[13] Drupsteen, L., & Hasle, P. (2014). Why do organizations not learn from incidents? Bottlenecks, causes and conditions for a failure to effectively learn. Accident Analysis & Prevention, 72, 351-358

[14] Drupsteen, L., & Wybo, J. L. (2015). Assessing propensity to learn from safety-related events. Safety Science, 71, 28-38

[15] Drupsteen, L., Groeneweg, J., & Zwetsloot, G. I. (2013). Critical steps in learning from incidents: using learning potential in the process from reporting an incident to accident prevention. International journal of occupational safety and ergonomics, 19(1), 63-77

[16] Elsbecker, G. & Kluge, A. (2013). Können Arbeitssicherheitsseminare (allein) das Verhalten nachhaltig verändern? Beitrag auf dem Frühjahrskongress der GfA, in Krefeld, 28.2.-1.3.2013, p. 275-278. Dortmund: GfA-Press

[17] Fahlbruch, B., & Schöbel, M. (2011). SOL–Safety through organizational learning: A method for event analysis. Safety Science, 49(1), 27-31

[18] Gollwitzer, M., & Jäger, R. J. (2007). Evaluation: Workbook. Weinheim, Basel: Beltz Verlag

[19] Gotcheva, N., Oedewald, P., Wahlström, M., Macchi, L., Osvalder, A. L., & Alm, H. (2016). Cultural features of design and shared learning for safety: A Nordic nuclear industry perspective. Safety Science, 81, 90-98

[20] Hagemann, V., Herbstreit, F., Kehren, C., Chittamadathil, J., Wolfertz, S., Dirkmann, D. Kluge, A. & Peters, J. (2017). Teaching medical students non-technical skills for emergencies improves such skills but not simulated patient outcome. International Journal of Medical Education

[21] Hayes, J., & Maslen, S. (2015). Knowing stories that matter: learning for effective safety decision-making. Journal of Risk Research, 18(6), 714-726

[22] Hesketh, B.(1997). Dilemmas in Training for Transfer and Retention. Applied Psychology: AAn International Review, 46 (4), 317-386

[23] Hochholdinger, S. & Schaper, N. (2007). Trainingsevaluation und Transfersicherung. In H. Schuler & Kh. Sonntag (Hrsg.). Handbuch der Arbeits- und Organisationspsychologie (S.625-633). Göttingen: Hogrefe

[24] Hochholdinger, S., Rowold, J., & Schaper, N. (2008). Ansätze zur Trainings- und Transferevaluation. In Evaluation und Transfersicherung betrieblicher Trainings (S. 30-53). Göttingen: Hogrefe

[25] IAEA Safety Standards, Leadership and Management for Safety – General Safety Requirements, No. GSR Part 2; ISBN 978-92-0-104516-4 (2016)

[26] Imaizumi, T., Takeda, Y., & Kusukami, K. (2014). Development of a Support Tool to Learn Lessons from Incidents Experienced by Others. JR East Technical Review, 30

[27] Kirkpatrick, D. L. (1987). More Evaluating Training Programs: A collection of articles from Training and Development Journal. Alexandria: American Society for Training and Development

[28] Kluge, A. & Burkolter, D. (2013). Training for Cognitive Readiness: Research Issues and Experimental Designs. Journal of Cognitive Engineering and Decision Making, 7, 96 -118

[29] Kluge, A. (2014) The acquisition of knowledge and skills for task work and team work to control complex technical systems. A cognitive and macroergonomics Perspective. Dortrecht: Springer

[30] Kluge, A. & Frank, B. (2014). Counteracting skill decay: Four refresher interventions and their effect on skill retention in a simulated process control task. Ergonomics. 57(2), 175-190

[31] Kraiger, K. (2002). Decision-based evaluation. In K. Kraiger (Ed.), Creating, implementing, and managing effective training and development (pp. 331-375). San Francisco: Jossey-Bass

[32] Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology, 78(2), 311-328

[33] Lemke, S.G. (1995) Transfermanagement. Göttingen: Verlag für angewandte Psychologie

[34] Littlejohn, A., Margaryan, A., Vojt, G., & Lukic, D. (2017). Learning from Incidents Questionnaire (LFIQ): The validation of an instrument designed to measure the quality of learning from incidents in organisations. Safety Science. 99, 3. 80-93

[35] Moorman, C., & Miner, A. S. (1998). Organizational improvisation and organizational memory. Academy of Management Review, 23(4), 698-723

[36] Pidgeon, N., & O'Leary, M. (2000). Man-made disasters: why technology and organizations (sometimes) fail. Safety Science, 34(1), 15-30

[37] Pilbeam, C., Doherty, N., Davidson, R., & Denyer, D. (2016). Safety leadership practices for organizational safety compliance: Developing a research agenda from a review of the literature. Safety Science, 86, 110-121

[38] Ramanujam, R., & Goodman, P. S. (2011). The challenge of collective learning from event analysis. Safety science, 49(1), 83-89

[39] Ritzmann, S., Hagemann, V. & Kluge, A. (2014). The Training Evaluation Inventory (TEI) - Evaluation of Training Design and Measurement of Training Outcomes for Predicting Training Success. Vocations and Learning, 7 (1), 41-73

[40] Salas, E., Wilson, K. A., Priest, H. A., & Guthrie, J. W. (2006). Design, delivery, and evaluation of training systems. In G. Salvendy (Ed.), Handbook of human factors and ergonomics (pp. 472-512). Hoboken, NJ: J. Wiley

[41] Shelton, A., & Alliger, G. M. (1993). Who's afraid of level 4 evaluation? A practical approach. Training and Development, 47, 43-46

[42] Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. Journal of Applied Psychology, 93(2), 280-295

[43] Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure. Academy of Management Learning & Education, 9, 169-191

[44] Solga, M. (2005). Management des Lerntransfers. In J. Ryschka, M. Solga & A. Mattenklott (Hrsg.) Praxishandbuch Personalentwicklung. Instrumente, Konzepte, Beispiele (S. 269-3292). Wiesbaden: Gabler

[45] Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. Human Resource Management Review, 20(2), 115-131

[46] Wottawa, H., & Thierau, H. (2003). Lehrbuch Evaluation. Bern: Verlag Hans Huber