

Comparative study of three training methods for enhancing process control performance: Emphasis shift training, situation awareness training, and drill and practice

Dina Burkolter^{a,*}, Annette Kluge^a, Jürgen Sauer^b, Sandrina Ritzmann^c

^aBusiness and Organizational Psychology, Department of Computational and Cognitive Sciences, Faculty of Engineering, University of Duisburg-Essen, Lotharstraße 65, D-47057 Duisburg, Germany

^bDepartment of Psychology, University of Fribourg, Rue de Faucigny 2, CH-1700 Fribourg, Switzerland

^cResearch Institute for Organizational Psychology, University of St. Gallen, Varnbühlstrasse 19, CH-9000 St. Gallen, Switzerland

Three training methods to improve attention management skills in process control were compared. Forty students from technical disciplines participated in a five-hour module of emphasis shift training (EST), EST combined with situation awareness training (EST/SA), and drill and practice (D&P) on a simulated process control task. Participants were then tested three times for 45 min each (immediately after training, two weeks after training, and six weeks after training) for system control performance and diagnostic performance on familiar and unfamiliar fault states. D&P led to superior diagnostic performance on familiar system faults. EST/SA training supported the diagnosis of novel system faults. EST was less effective than expected for system control performance. Implications for training design in process control are discussed.

1. Introduction

Highly automated installations such as refineries or nuclear power plants involve extremely complex, dynamic process control tasks that require personnel to monitor and control the system and to detect, diagnose, and rectify malfunctions or make repairs (Kluge, Sauer, Schüler, & Burkolter, 2009). These tasks demand different kinds of attention from process control operators such as focused, divided, and selective attention (Wickens & McCarley, 2008). But as crucial as attention capacities are to process control performance, they are limited. The operator "would gain most if he or she could fully attend to all elements, at all times. However, such full attention is not possible. Hence, some priorities and tradeoffs must be established along with attention allocation strategies" (Gopher, 1996, p. 28). In other words, process control operators must learn to apportion their attention strategically. Attention is strongly associated with mental workload. Workload is on the one hand determined by exogenous task demands such as task difficulty and task priority, and on the other hand by endogenous supply of attentional or processing resources to support information processing (Tsang & Vidulich, 2006). Our research goal was to support novice learners by providing them with attention-management strategies in order to reduce their mental workload.

Highly complex and demanding tasks such as process control can present demands which are difficult to cope with (Gopher, Weil, & Siegel, 1989). Therefore, especially in the early stages of training, learners should be supported by providing them with attention-management strategies.

1.1. Three training approaches to improve attention management and performance

Adapting, applying, and comparing auspicious training approaches drawn mostly from aviation, in this study we seek to improve attention, attention management, and, ultimately, performance on highly complex, dynamic process control tasks. In the following, we describe the approaches, their underlying theoretical concepts and why they are selected for learning a process control task. We introduce the training approaches of emphasis shift training (EST), situation awareness (SA) training, and drill and practice (D&P) (see Table 1) and summarize research findings on their effectiveness.

1.1.1. Emphasis shift training (EST)

The first training approach we selected for our research on attention management in process control was EST. It was originally developed by Gopher et al. (1989) to sharpen the ability to cope with highly demanding tasks and, especially, to strengthen attention management and the control of attentional resources. In EST,

* Corresponding author. Tel.: +49 203 3794207; fax: +49 203 3793411.
E-mail address: dina.burkolter@uni-due.de (D. Burkolter).

multiple changes in the emphasis (priority) on components of a task are introduced, but the whole task is left intact. Only the attention status of the subtasks is changed. Hence, EST is a part-task training approach. Key constructs of EST are strategies of performance, response schemas and the voluntary control of attention. Strategies and organized sets of response schemas are central to complex tasks. A strategy is a distinct approach of an individual to cope with the set of subgoals of a task. Strategies are controlled at the beginning, but may become high-level schemas that can be triggered automatically with training and practice. Once a schema is developed, the operation of it is assumed to require few attentional resources. Hence, attentional resources are freed for other tasks (Gopher et al., 1989). This is important, as operators performing complex tasks are required to coordinate many complex action sequences and subtasks in parallel. To support the development of strategies and schemas, Gopher et al. introduce the idea of voluntary control of attentional resources. There is theoretical and empirical evidence that attention control and attention management can be treated as a skill, and thus can be improved by training (Gopher, Weil, & Bareket, 1994). Gopher et al. (1989), for example, showed that spontaneous strategies developed by learners to try to cope with complexity were not very successful. In contrast, in EST, in which learners were provided with strategies, participants showed higher performance. Hence, trainees can be provided with strategies both to reduce mental workload and to improve performance.

EST is assumed to prepare participants for another challenge in process control. Some process control tasks, such as shutdown, start-up, and fault-finding, require completion only at certain intervals (Sauer, Hockey, & Wastell, 2000). Skill components are called upon not only in practiced, familiar situations but in novel, unfamiliar ones as well (Kluge et al., 2009). Training of unexpected, novel fault states should focus on attention-management strategies, because these strategies are central to responding to novel situations (Shebilske, Goettl, & Garland, 2000).

EST has thus far been used in different contexts, such as complex and dynamic environments (Space Fortress game), flying with a helmet-mounted display, touch-typing skills, and basketball (Gopher, 2007), but not to a process control task. EST makes it possible to resolve difficulties known from traditional part-task training approaches (Gopher et al., 1989). EST has been effective overall, especially for strengthening attention-management strategies (Gopher et al., 1994; Shebilske et al., 2000). EST has improved transfer of skills to new and changed tasks. However, EST's effectiveness has

usually been tested at the end of training (Gopher, 2007), not after an extended retention interval. Therefore, there is a need to test whether EST can support skill retention.

Gopher et al. (1989) explain EST's effectiveness in terms of load reduction that permits a person to increase the resources invested in the learning of other tasks. They maintain that EST helps participants broaden their perspective of their given task, expand their knowledge about the efficiency of their own resources, and gain flexibility in adopting different modes of response that suit their individual capabilities.

1.1.2. SA training

Attention is also critical to achieving SA (Endsley, 1995b; Tsang & Vidulich, 2006). SA is understood to be the perception and comprehension of elements in the environment and the projection of their status in the near future. Research on SA goes back to aviation (e.g. Endsley, 1995b), but in recent years other fields have followed, including process control (e.g. Hogg, Follesø, Strand-Volden, & Torralba, 1995). In process control, operators have to monitor plant states, alarm screens and panels, and to observe the state of numerous system parameters and patterns among them in order to gather information about the functioning of the system and future process state changes (Endsley, 1995b; Vicente, Mumaw, & Roth, 2004). Limited attention capacities, such as lapses in attention and the constraints on the ability to accurately perceive several items in parallel, present a major limit to SA (Endsley, 1995b; Wickens & McCarley, 2008). Schemas can support individuals to develop SA in that they are mechanisms for directing attention in the perception process. SA is achieved by recognizing critical cues in the environment that will map to key characteristics of the schema or mental model. Schemas and mental models are developed through training and experience by noticing recurrent situational components and causal relationships (Endsley, 1995b). As SA is critical to process control performance (Endsley, 1995b; Wickens & McCarley, 2008), we selected SA training as a second training approach. To improve individuals' SA, training of attention sharing and task management strategies is recommended (Endsley & Robertson, 2000). By combining EST with SA training, an approach to enhance SA and performance both directly (with the SA training) and indirectly by training attention and task management to support SA (with the EST) was chosen.

Relatively few programs include the evaluation of SA training (Endsley & Robertson, 2000). However, Saus et al. (2006) have found empirical evidence substantiating the effects of the SA training

Table 1
Comparison of emphasis shift training, situation awareness training, and drill and practice.

Description	Emphasis shift training (Gopher, 2007; Gopher et al., 1989)	Situation awareness training (Endsley and Robertson 2000; Saus et al., 2006)	Drill and practice (Carlson et al., 1989; Ericsson et al., 1993)
Rationale	Learning through priority changes on subcomponents of a whole task	Learning through randomized "freezing" of a task with situation awareness questions and debriefing	Learning through rehearsal, repetition and practice of a task
Original application area	Aviation, complex tasks	Aviation, police, complex tasks	Nondynamic tasks, motor tasks
Empirical findings	Enhancement of attention-management strategies, useful for transfer to novel situations	Useful for individual situation awareness as well as performance, but few empirical studies on effects of situation awareness training	Useful for procedural tasks and longer retention intervals for familiar situations
Explanations for training effectiveness	Reduction of load allowing to invest more resources in learning other tasks	Improved competence to make decisions and project events in the future	Reduction of load on working memory
	Broader perspective of task, better knowledge of own resources	Enhanced mental models allowing for better understanding of situations	Increased speed of component processes and restructuring in the use of working memory
Usefulness assumed for	System control performance	System control performance	Diagnostic performance of practiced faults
	Diagnostic performance of novel faults	Diagnostic performance of novel faults	

they designed for students of a police university on the basis of a “freezing” technique coupled with debriefing. This approach calls for randomly stopping, or freezing, a simulated task, posing the participant questions about the three levels of SA (Endsley, 1995b), then debriefing the individual. Because little is known about either SA training or the attendant retention intervals, we also investigate both matters in this study.

Explanations for the effectiveness of SA training are that it improves the ability to make timely and effective decisions and to anticipate events (Saus et al., 2006). In addition, SA training facilitates the construction of mental models that improve the understanding of both the importance of various situations and the resources relating to them (Endsley & Robertson, 2000).

1.1.3. Drill and practice (D&P)

Lastly, practice can aid attention performance by proceduralizing or automating a task in order to free resources for another task (Wickens & McCarley, 2008). As individuals continuously practice a task, gradual improvement in time-sharing performance and divided attention has been observed. As soon as one task has been automatized, attentional resources can be applied to other tasks. These changes are ascribed to two processes. First, interference between tasks depends on the demands of the tasks for a limited supply of mental resources. Second, the resource demand of a task decreases with practice until resource-free automaticity is reached (Wickens & McCarley, 2008). We therefore conducted drill and practice (D&P) training as a third approach for process control. D&P facilitates learning through rehearsal of a task in order to achieve a desired level of proficiency (Cannon-Bowers, Rhodenizer, Salas, & Bowers, 1998). Similarly in the EST approach, novice learners are provided with strategies instead of trying to cope with the task on their own. In D&P, learners are provided with a clear strategy to cope with the task and guided in a step-by-step manner through the steps of the task. Thereby, a learner’s attention is guided to the accurate execution of the task steps instead of having the learner divide attention between finding a strategy and performing the task at the same time. As with EST, D&P is expected to reduce the learner’s mental workload, especially in initial learning.

So far, D&P has been applied mainly for nondynamic and relatively easy cognitive tasks (Shute & Gawlick, 1995). Research has shown practice to be effective at improving accuracy and speed of performance on perceptual, motor, and cognitive tasks (Ericsson, Krampe, & Tesch-Romer, 1993). Practice has also proved to be a more complex training approach than originally thought (Cannon-Bowers et al., 1998; Schmidt & Bjork, 1992). D&P has been especially effective as a method for training people to diagnose practiced fault states (Kluge & Burkolter, 2008; Shute & Gawlick, 1995), though procedural skills demand more practice time than psychomotor skills do (Ginzburg & Dar-El, 2000). Moreover, studies involving initial training and refresher training have shown D&P to support skill retention through the repetition of training exercises (Hagman & Rose, 1983; Schendel & Hagman, 1982; Shute & Gawlick, 1995).

The effectiveness of D&P is attributed primarily to the method’s acceleration of component processes and to concomitant restructuring in the use of working memory. It is also attributed to the reduction of the load on working memory that is essential for carrying out cognitive processes (Carlson, Sullivan, & Schneider, 1989).

1.2. The present study

We empirically evaluated the EST, EST/SA, and D&P approaches for their effectiveness over a retention interval of several weeks and in practiced and novel fault states. Because D&P had already

been successfully employed (Kluge & Burkolter, 2008), it served as a baseline against which to assess EST and EST/SA. We conducted EST on two main subtasks of process control—(a) system control and stabilization and (b) diagnostic performance—shifting the emphasis between them. This procedure, known as “double manipulation,” has been shown to optimize EST (Gopher et al., 1989). In the EST/SA training, EST was supplemented by SA training based on the freezing-and-debriefing technique (see Saus et al., 2006). We used a simulated multitask environment that corresponded to a process control environment. The simulation thus involved the two main subtasks of process control named above.

We derived four assumptions from the literature and studies cited in this study (see Sections 1.1.1–1.1.3). First, the EST/SA group would be more successful than the other groups at developing and maintaining SA, since they receive SA training based on the freezing-and-debriefing technique described above, aimed directly at improving SA.

Second, EST and EST/SA would be more effective than D&P at developing participants’ performance on system control. We suggest that, as EST enhances attention-management strategies, participants will be better able to pay careful attention to the state of the parameters and the detection of deviances from the target range, as needed for good system control and stabilization performance.

Third, D&P would be more effective than the two other training methods for developing participants’ performance on diagnosing familiar fault states. This assumption is based on research which has shown that D&P supported skill retention, especially for practiced fault states.

Fourth, EST and EST/SA would be more effective for developing participants’ performance on diagnosing novel fault states than on diagnosing practiced ones. This assumption is backed first by empirical studies employing EST, which was shown to improve the transfer of skill to new and changed tasks. Second, the training of attention-management strategies is assumed to improve responsiveness to unexpected, novel fault states. Regarding SA training, we suggest that SA training has the potential to improve the diagnostic performance of unfamiliar fault states, for operators with sound SA might detect and understand abnormal situations earlier than they otherwise would. Moreover, they are trained in predicting future states that can evolve out of abnormal states.

2. Method

2.1. Design

A $3 \times 3 \times 2$ mixed factorial design was employed. Training as a between-participants variable varied at three levels (EST, EST/SA, and D&P). Time of measurement as a within-participant variable was taken in three separate testing sessions ($test_0$, $test_{2w}$, and $test_{6w}$). Fault type as a within-participants variable varied at two levels (practiced and novel faults).

2.2. Participants

Forty-eight university students (four female) participated in the experimental study. They were all enrolled in a program leading to a Bachelor of Science degree in a technical field of study (aeronautics, engineering, or electrical engineering) at universities of applied sciences in the Greater Zurich Area, Switzerland. The students were paid 100 CHF (approximately US \$90.00) for participation in all three parts of the experiment. Participants were randomly allocated to the three training methods at each location. Forty students (83.3%; four female) completed all three parts of the experiment. Training groups did not differ significantly regarding

drop-outs ($H(2)=1.336, p > 0.05$). The mean age of the participants was 24.7 years ($SD = 4.0$). There were no significant differences between the mean ages of the training groups ($F(2, 39) = 0.874, p > 0.05$).

2.3. Simulated process control environment

In an introductory training module and subsequent testing sessions, we used a computerized process control task simulated by the Cabin Air Management System (CAMS; for details, see Sauer, Wastell, & Hockey, 2000). CAMS models a life support system onboard a spacecraft. Five parameters (O_2 , CO_2 , cabin pressure, temperature, and humidity) are kept in a predefined zone by automatic controllers. The operator's task is to intervene if necessary. This individual must complete two primary tasks (system control and fault diagnosis, see Sections 2.5.3 and 2.5.4) and two secondary tasks (prospective memory and reaction time). CAMS records the actions carried out by the operator.

2.4. Training methods

Three different types of training were given: EST (13 participants), combined EST and SA training (14 participants), and D&P (13 participants). As is often the case with training that involves complex systems such as refinery simulators, training was carried out in small groups typically consisting of four to six participants supervised by one instructor in order to ensure sound supervision and effective learning (Kriedemann, 2008).

All training had the same general introduction to CAMS and focused on the same five system faults. Training material given to all training groups consisted of an illustration of CAMS, its components, and controls (with notations and translations); a CAMS manual; and an instruction manual. The CAMS manual described the main components, systems and controllers of CAMS, the tasks of the participants and 16 different system faults. For every system fault, a description of the fault was given, the symptoms were de-

scribed and the intervention steps (system control and fault-finding) were depicted. The instruction manual (see Fig. 1) was designed to guide the participant through the training of the five system faults. The instruction manual was based on the CAMS manual and included a screenshot of CAMS during the fault state and descriptions of symptoms and intervention procedures for fault diagnosis and repair and for control and stabilization of the system. The instructions and the number of exercises per system fault differed from one training group to the next, but the duration of training was the same for all groups. Whereas the EST and SA group worked with the same instruction manual, the D&P group received a different instruction manual (see Sections 2.4.1 and 2.4.2).

The first training block (see also Fig. 3) was a general introduction to CAMS and the corresponding manual. The participants received 10 min of multimedia instruction about CAMS, its features, and the primary and secondary tasks involved in the experiment. Participants followed the multimedia instructions individually on their computers using earphones. They were then given a few minutes to explore the system on their own (e.g. looking at the system components, trying out controllers). A short presentation and an exercise introducing the CAMS manual followed. The presentation was given by the instructor and aimed to prepare participants to use the manual by describing its content, structure, and function. (In the exercise, participants were asked to find a certain system fault ("On which page can the system fault 'vent stuck on' be found?"), and to describe the system fault and its symptoms. Answers were discussed with the instructor.

The second training block introduced all participants to five system faults: (a) a leak in an oxygen (O_2) valve, (b) a cooler set point failure, (c) a block in a mixer valve, (d) a carbon dioxide (CO_2) set point failure, and (e) a nitrogen set point failure. Selection of system faults was based on a hierarchical task and an analysis of sub-goal templates (see Burkolter, Kluge, Schüler, Sauer, & Ritzmann, 2007). System faults were randomly allocated over parameter, type of faults, and the difficulty of the repair procedure. The participants

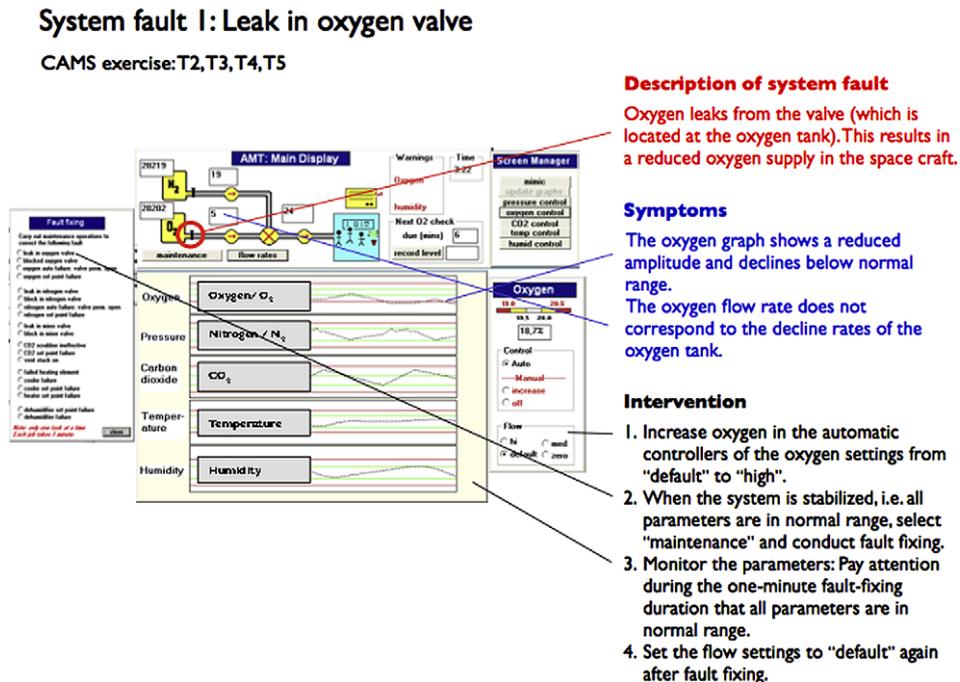


Fig. 1. Instruction manual for the D&P group depicting one of five system fault descriptions (translated from German). On the left side, a screenshot of the simulation CAMS during the fault state is displayed. On the right side, the system fault, the symptoms, and the steps of intervention are described. Faults, symptoms and intervention steps are indicated on the screenshot with straight lines.

practiced the five systems faults by following the instruction manual. Each first exercise of a system fault was conducted using a pause button that is provided in CAMS. Thus, the participants had the possibility to pause CAMS and refer to the instruction manual for further intervention steps.

Participants worked individually on their own computers by following the instruction manual and the instructions provided by the instructor. The participants were given the possibility to ask questions at any time during the training.

2.4.1. Emphasis shift training (EST)

At the beginning of the second training block, the instructor explained the principles of EST by the means of a presentation. The two main tasks of CAMS, the principle of changing the emphasis between the two, and the idea of EST (to learn to manage more than one task at the same time) were pointed out. Then, a 10-min introduction to system control and stabilization was provided along with a corresponding exercise (increasing three different parameters by adjusting controls and observing what happens). Thereafter, exercises on the five system faults to be trained followed. The instructor trained the EST group by changing the emphasis on the two main tasks of process control (system control and fault-finding). In each exercise, one of these main tasks was alternately emphasized. That is, participants practiced the actions of only one main task and did not have to execute the other interventions. The instruction manual of the EST group indicated the main task to be emphasized, and the proper steps were marked with a red arrow and boldfaced letters. Information on the other task appeared in light coloring (see Fig. 2). In all other respects, the instruction manual was identical to the one the D&P group received. The members of the EST group practiced each system fault (SF) three times: twice one fault separately and once together ($2 \times SF_1$, $2 \times SF_2$, $1 \times SF_1$ and 2 together; $2 \times SF_3$, $2 \times SF_4$, $2 \times SF_5$, $1 \times SF_{3, 4, 5}$ together).

2.4.2. EST combined with SA training (EST/SA)

The EST/SA group and EST group received the same instruction manual and the same exercises with changing emphasis. Additionally, the EST/SA group was given an SA training.

As in the EST group, the instructor explained the principles of EST by means of a presentation. Then, a 10-min introduction to system control and stabilization was provided along with a corresponding exercise. This was followed by exercises on the first system fault to be trained. Participants were trained by changing the emphasis on the two main tasks of process control. Then, in contrast to EST, a brief presentation introducing the concept of SA was given by the instructor. The concept of SA, the three levels of SA (Endsley, 1995b), and its relevance for CAMS performance were explained to the participants. From the third system fault on, SA was practiced through the freezing technique with debriefing (as described by Saus et al., 2006). The simulation was frozen one time during each exercise, and participants received open questions about two parameters based on Endsley's (1995b) concept of SA. The first item concerned perception of the situation (e.g. "How is the status of temperature?"); the second item concerned comprehension and future actions ("What does this mean? What needs to be done now?"). After completing the short questionnaire, the participants discussed responses with the instructor. The points at which the simulation was frozen were distributed over the beginning, middle, and end of a system fault. The questions took all parameters into account.

2.4.3. Drill and practice (D&P)

The D&P group extensively practiced system control and fault-finding by repeating several different drills. The participants were told to follow the steps of intervention closely as described in the instruction manual (see Fig. 1), which explained the various steps of system control and stabilization, fault diagnosis, and repair. For each system fault, they were first given 3 min to read about it and to memorize the intervention steps. They then practiced

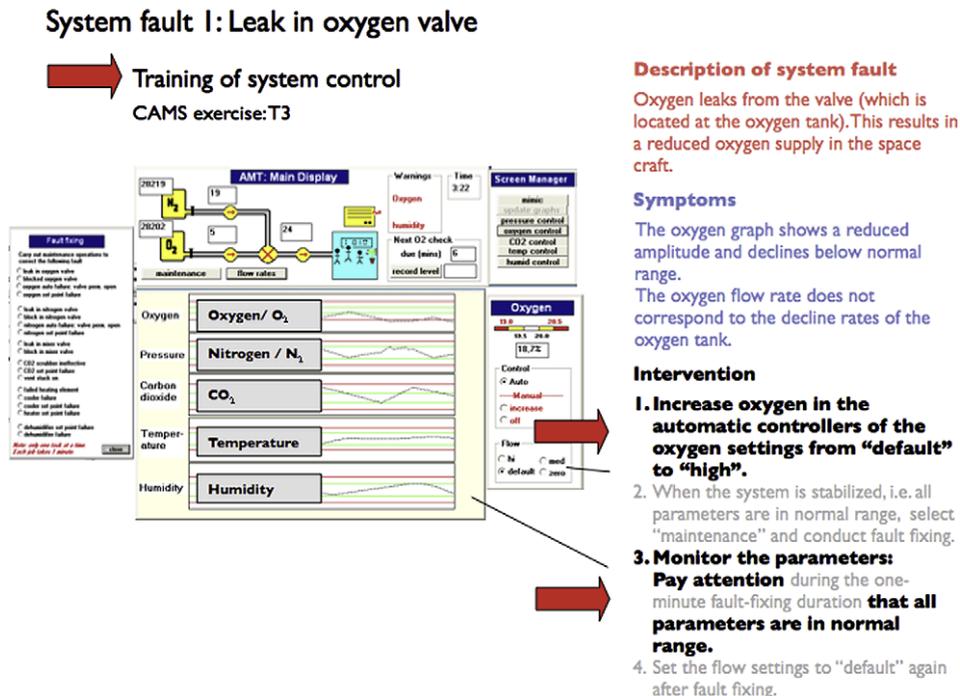


Fig. 2. Instruction manual for the EST and EST/SA group depicting one of five system faults (translated from German). This is an exercise emphasizing the training of system control, which is indicated at the top left for the participants. The steps of intervention that concern system control are in boldface and marked with a red arrow. Information on the other task, i.e. fault-finding (as well as description and symptoms), appear in light coloring.

each system fault with its intervention steps a total of five times—four times each fault separately, then once together with at least one other system fault (SF) ($4 \times SF_1$, $4 \times SF_2$, $1 \times SF_1$ and 2 together; $4 \times SF_3$, $4 \times SF_4$, $4 \times SF_5$, $1 \times SF_{3, 4, 5}$ together).

2.5. Measures

2.5.1. Reactions to training

After the training module the participants rated five reactions to the training they had just experienced. The first three—mental effort, anxiety, and fatigue—were each rated in response to a single item each. The item for anxiety, for example, read “How do you feel right now?” and was rated on a scale from 0 (e.g. calm) to 100 (e.g. tense) (see Sauer, Wastell, & Hockey, 2000, for details). The last two reactions were motivation and self-efficacy. They were rated on a six-point scale ranging from 1 (I totally agree) to 6 (I do not agree at all). The participants’ responses relating to post-training motivation was elicited by the following four items: “The task appealed to me.” “I would like to participate again in a training like this.” “I was motivated to accomplish the CAMS task.” “I find CAMS interesting and fascinating” (Cronbach’s Alpha = 0.86). The degree of self-efficacy that trainees felt about understanding and controlling CAMS was measured through their responses to four items (e.g. “I feel up to the tasks of CAMS.”) (Cronbach’s Alpha = 0.87; see Kluge, 2008, for details).

2.5.2. Situation awareness

To check whether SA training was improving SA, we measured the latter index by means of the SA Control Room Inventory (Hogg et al., 1995), which was adapted for application to CAMS. The inventory was developed for use in process control research based on the Situation Awareness Global Assessment Technique (Endsley, 1995a). During the testing sessions, the simulation was frozen six times at irregular intervals. The participants were instructed to switch off their screens so that they could not see the current system state. Questions about the status of the system were then presented. In accordance with Endsley’s (1995b) concept of SA, one item concerned the perception of the current situation (“How is the status of humidity?”). The response alternatives were “below normal range,” “within normal range,” and “above normal range.” Another item covered comprehension of the current situation and prediction of future status (“How do you think the course of humidity will develop over the next 10 s? Provided that no intervention to the system is undertaken, the parameter will ... “decline,” “remain stable,” or “increase” in 10 s.”). SA was measured once in the first CAMS test run and once in the second CAMS test run. There were twelve questions in total (six measurements with each two items). The responses were then compared to logged CAMS data to determine whether the responses by the participants were correct (for this approach see Hogg et al., 1995, p. 2411).

2.5.3. System control failures

One of the primary tasks of the operator was to maintain five key parameters within normal range. If one or more of the key parameters departed from normal range, the operator needed to intervene by adjusting automatic controllers or adapting manual control. The duration of the parameters’ deviation from normal range was measured in seconds and converted into percentages.

2.5.4. Fault-finding

The other primary task concerned fault diagnosis and repair with the assistance of the maintenance facility. In the event of a system fault, the operator had to identify the cause by carrying out appropriate tests. There were two measures of diagnostic performance: the percentage of incorrect diagnoses (diagnostic accu-

racy) and the number of seconds the operator needed in order to identify the fault correctly (diagnostic speed).

2.5.5. Knowledge tests

Knowledge was assessed with an adapted version of two existing knowledge tests on CAMS (see Sauer, Burkolter, Kluge, Ritzmann, & Schüler, 2008). Structural knowledge was measured with a method described by Meyer (2008). Findings concerning knowledge are reported in a separate article (Burkolter, Meyer, Kluge, & Sauer, in press).

2.6. Procedure

Fig. 3 summarizes the experimental procedure. There were three parts. The first consisted of questionnaires and a training module (about 4¼ h) followed immediately by a 45-min testing session ($test_0$). The second part was a 45-min testing session two weeks after the training module ($test_{2w}$). The third part was a 45-min testing session six weeks after the training module ($test_{6w}$). Because of organizational constraints stemming from the differences between the schedules of university terms, the retention interval between the testing sessions could not be identically long (i.e. two weeks between $test_0$ and $test_{2w}$ and four weeks between $test_{2w}$ and $test_{6w}$).

Upon arrival at the experimental facility, all participants spent 35–45 min completing questionnaires on cognitive ability, cognitive flexibility, personality, and motivation. This pretraining testing did not include a testing session on the CAMS task to rule out initial differences of the training groups in process control performance. However, CAMS is an artificial task which is not commercially available, cannot be retrieved from the internet and is not known outside the scientific community. Although CAMS is situated in the context of spaceflight, it does not correspond directly to real-world physical principles, but has its own rules and interconnections. Therefore, previous knowledge on spaceflight and related knowledge is unlikely to be very useful for CAMS performance, and thus the likelihood of initial differences of training groups are assumed to be low. However, we controlled for differences between training groups regarding cognitive and personality variables.

The training module was equally long for all three training groups. It lasted approximately 3¼ h (including one 5-min and one 20-min break). The number of training exercises varied, however. The D&P group performed five exercises per system fault, for the core idea of D&P is to provide a good deal practice on the task. During the SA training of the EST/SA group, the members of the EST group were given a comparable cognitive task. They heard a talk (about recruitment criteria for astronauts; see Sauer et al., 2008), viewed part of a documentary, and like the members of the EST/SA group, answered questions about the material to which they had just been exposed (about 35 min). $test_0$ covered all five practiced fault states in addition to fault states that the participants had not previously encountered (block in nitrogen valve and dehumidifier set point failure).

$test_{2w}$ and $test_{6w}$ were identical for all participants. After a brief introduction to refresh knowledge on the experimental task, participants worked with CAMS and were tested for SA during either the first or second part of the test (see Section 2.5.2). $test_{2w}$ was followed by the knowledge tests, which took 30–40 min. The second and third testing sessions each included three familiar and three novel fault states. For an overview of the system faults employed in training and testing sessions, see Table 2. Participants were given no advance information about the order and time of appearance of faults. The CAMS manual was available to participants during the testing sessions. After the final testing session, participants were debriefed about the three training methods

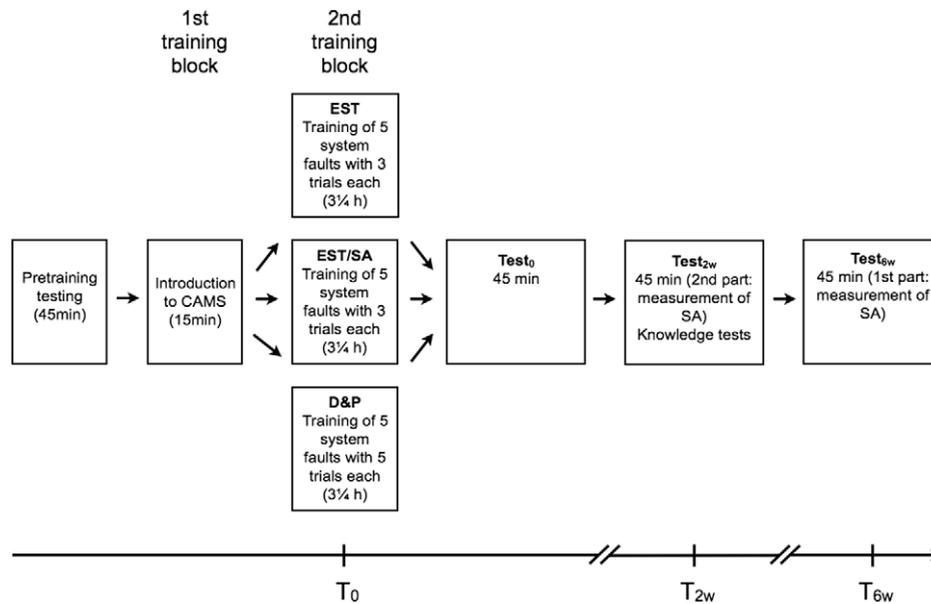


Fig. 3. The experimental procedure. EST: emphasis shift training; EST/SA: EST and situation awareness (SA) training; D&P: drill and practice. Test₀: test immediately after training; Test_{2w}: test two weeks after training; Test_{6w}: test six weeks after training.

and given the opportunity to ask questions about the design of the experiment.

During the first part of the experiment in which the training took place, it was assured that participants in the different training groups could not interact or learn from each other. The training sessions were conducted in different rooms or on different days. Thus, in the particular part of the experiment in which the actual training intervention took place, participants could not interact with each other. In the two weeks between the first and second as well as the four weeks between the second and third sessions, however, participants could have had the opportunity to interact. Unfortunately, this could not have been prevented, since participants were enrolled at the same universities and (partly) attending the same courses. The latter is associated with the fact that we conducted the experiment at the particular universities of the participants instead of inviting them to a lab. However, it was assured that none of the instruction material including the CAMS program was available to the participants outside the experimental sessions by collecting all material at the end of the sessions. Moreover, there is no single best “solution” to the CAMS task that could have been shared by participants, but during testing sessions, the CAMS manual was available to all participants.

Table 2
Overview of system faults employed in training and testing sessions.

Fault type	Training	Test ₀	Test _{2w}	Test _{6w}
<i>Practiced faults</i>				
Leak in oxygen (O ₂) valve	×	×	×	
Cooler set point failure	×	×	×	×
Block in mixer valve	×	×	×	
CO ₂ set point failure	×	×		×
Nitrogen set point failure	×	×		×
<i>Novel faults</i>				
Block in nitrogen valve		×		
Dehumidifier set point failure		×		
Nitrogen valve permanently open			×	
CO ₂ scrubber ineffective			×	
Oxygen valve permanently open			×	
Block in oxygen (O ₂) valve				×
Cooler failure				×
Leak in mixer valve				×

2.7. Assessment of control variables

We controlled for cognitive ability, cognitive flexibility, conscientiousness, and pretraining motivation, which are seen as relevant for training situations (see Colquitt, LePine, & Noe, 2000). These variables were employed to control for possible “unhappy randomization” (Mohr, 1995). Cognitive ability was assessed with the Wonderlic Personnel Test (Wonderlic Inc., 2002) cognitive flexibility, with the Cognitive Flexibility Inventory (Spiro, Feltovich, & Coulson, 1996). Big-Five Markers (Saucier, 1994) were employed to measure conscientiousness as a personality trait. Pretraining motivation was assessed with one item (“Please indicate how motivated you are to participate in this training? 0%, 20%, 40%, 60%, 80%, 100%.”). As Table 3 shows, participants in the three groups did not differ significantly with regard to control variables. Thus, there was no unhappy randomization.

3. Results

We performed mixed ANOVAs. If the assumption of sphericity was violated, we corrected degrees of freedom by using Greenhouse-Geisser estimates of sphericity. We drew planned contrasts in specific ways (see Loftus, 1996). The first contrast was between the D&P group and the EST/SA group, the purpose was to compare the two groups that had had only exercises and SA training and no further task. We focused the second contrast on the EST and EST/SA groups so as to ascertain whether the additional SA training enhanced process control performance beyond the effects of EST. Interaction effects were broken down into interaction contrasts, as proposed by Gamst, Meyers, and Guarino (2008).

Wickens (1998) states that low sample size and high variance may increase the probability of a Type II error and that, ergonomically speaking, Type II statistical errors can be as important as Type I errors. To avoid Type II statistical errors, we also report *p*-values at the 10% level.

3.1. Reactions to training

Training groups differed significantly in their ratings of mental effort, anxiety, and fatigue (see Table 4). Planned contrasts of effort

revealed that the D&P group differed significantly from the EST/SA group ($p < 0.05$), with the D&P group investing the most effort in the task. In terms of anxiety and fatigue, planned contrasts revealed that the EST group was less tense and tired than the EST/SA group ($p < 0.1$). However, participants did not differ significantly on either post-training motivation or self-efficacy ratings.

3.2. Situation awareness

Surprisingly, the participants of the three training groups did not differ significantly on their SA performance tested at T_{2w} and T_{6w} ($F(2, 35) = 0.097, p > 0.10, \eta_p^2 = 0.00$). Neither a significant main effect of time ($F(1, 35) = 0.463, p > 0.10, \eta_p^2 = 0.01$) nor a significant interaction effect of group and time was found ($F(2, 35) = 0.524, p > 0.10, \eta_p^2 = 0.03$). Thus, our first assumption was not supported by the data.

3.3. System control failures

An inspection of descriptive statistics (see Table 5) suggests that there was a main effect of training and that the EST group performed better in system control than the other two groups did. However, the three-way mixed ANOVA failed to support this interpretation ($F(2, 37) = 1.49, p > 0.10, \eta_p^2 = 0.07$). In contradiction of the second assumption there was no significant main effect of training. A significant interaction effect between fault type and training group was observed, however, ($F(2, 37) = 3.37, p < 0.05, \eta_p^2 = 0.15$). This interaction indicates that training groups differed in their performance depending on the type of fault (either practiced or novel). As depicted in Fig. 4, the D&P and EST groups performed better than the EST/SA group during practiced faults, and the EST/SA group performed better during novel faults. Interaction contrasts showed a significant interaction for the D&P group and EST/SA group ($F(1, 25) = 8.992, p < 0.01, \eta_p^2 = 0.27$) and for the EST and EST/SA group ($F(1, 25) = 2.895, p = 0.10, \eta_p^2 = 0.1$). Analysis

also revealed a significant main effect of both time ($F(1.73, 37) = 35.08, p < 0.01, \eta_p^2 = 0.49$) and fault type ($F(1, 37) = 5.861, p < 0.05, \eta_p^2 = 0.14$). Contrasts regarding the main effect of time showed that performance differed significantly between T_0 and T_{2w} and between T_{2w} and T_{6w} ($p < 0.001$), with the best performance occurring at T_{2w} .

3.4. Diagnostic performance

A three-way mixed ANOVA with diagnostic accuracy was performed (see Table 6). Analysis revealed no significant effect of training group on performance ($F(2, 37) = 0.795, p > 0.10, \eta_p^2 = 0.04$). Confirming the assumptions, however, there was a significant interaction effect between fault type and training ($F(2, 37) = 2.72, p < 0.10, \eta_p^2 = 0.13$), indicating that the performance of the training groups differed in fault type. The interaction graph (see Fig. 5) displays that D&P resulted in better performance on practiced faults than on novel faults, whereas EST/SA resulted in better performance on novel faults than on practiced faults. Interaction contrasts showed a significant interaction for the D&P group and EST/SA group ($F(1, 25) = 4.072, p < 0.10, \eta_p^2 = 0.14$), confirming the assumption that EST/SA aided diagnosis of novel faults. A significant main effect of time was found as well ($F(2, 74) = 6.45, p < 0.01, \eta_p^2 = 0.15$). Contrasts showed that performance at T_0 and T_{2w} and at T_{2w} and T_{6w} differed significantly ($p < 0.01$), with the poorest performance generally occurring at T_{2w} . There was no main effect of fault type ($F(1, 37) = 0.06, p > 0.10, \eta_p^2 = 0.00$).

On the second measure of diagnostic performance (diagnostic speed), the results of the three-way mixed ANOVA resembled those relating to diagnostic accuracy (see Table 6). We observed no significant main effect of training ($F(2, 37) = 0.51, p > 0.10, \eta_p^2 = 0.03$), but, as with diagnostic accuracy, the interaction between fault type and group was significant ($F(2, 74) = 3.87, p < 0.05, \eta_p^2 = 0.17$). This interaction effect indicated that training groups differed significantly in diagnostic speed regarding fault

Table 3
Descriptive statistics (*M, SD*) on control variables as a function of training group.

Control variables	Drill and practice	EST	EST/SA	<i>F/p</i>
Cognitive ability (0–50)	25.71 (4.68)	26.25 (6.26)	24.88 (6.04)	$F(2, 45) = 0.232, p > 0.05$
Cognitive flexibility (–4 to +4)	1.00 (1.22)	0.84 (0.90)	0.53 (1.07)	$F(2, 46) = 0.814, p > 0.05$
Conscientiousness (1–9)	6.84 (0.95)	6.52 (1.00)	6.78 (1.30)	$F(2, 46) = 0.360, p > 0.05$
Pretraining motivation (%)	80.00 (13.59)	80.0 (16.90)	76.47 (19.02)	$F(2, 45) = 0.236, p > 0.05$

Table 4
Descriptive statistics on reactions to training as a function of training group.

Reactions to training	Drill and practice <i>M (SD)</i>	EST <i>M (SD)</i>	EST/SA <i>M (SD)</i>	<i>F/p</i>
Effort (0–100)	68.93 (20.0)	48.00 (21.17)	50.83 (26.13)	$F(2, 47) = 3.644, p < 0.05, \eta_p^2 = 0.14$
Anxiety (0–100)	47.79 (23.73)	27.59 (19.53)	40.56 (23.47)	$F(2, 47) = 3.198, p < 0.1, \eta_p^2 = 0.12$
Fatigue (0–100)	68.96 (26.17)	47.06 (26.47)	64.17 (26.70)	$F(2, 47) = 2.936, p < 0.1, \eta_p^2 = 0.12$
Motivation (1–6)	4.13 (0.81)	4.21 (1.04)	3.68 (1.27)	$F(2, 45) = 1.147, p > 0.1, \eta_p^2 = 0.05$
Self-efficacy (1–6)	2.73 (0.94)	2.75 (0.94)	3.19 (1.13)	$F(2, 47) = 1.120, p > 0.1, \eta_p^2 = 0.05$

Table 5
System control failures (in percentages) as a function of training and fault type (*SD* in parentheses).

Fault type	Drill and practice			EST			EST/SA		
	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}
Practiced	11.6 (9.0)	3.5 (2.3)	8.6 (5.5)	9.5 (8.0)	4.0 (6.0)	6.4 (4.0)	17.6 (14.1)	5.1 (5.7)	9.8 (6.1)
Novel	8.8 (6.9)	7.4 (6.3)	17.5 (3.6)	6.6 (2.3)	4.3 (3.1)	15.5 (5.3)	8.4 (4.6)	6.2 (4.9)	15.7 (6.9)

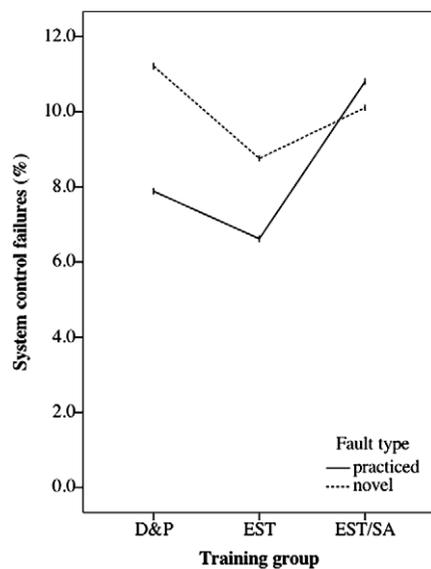


Fig. 4. Interaction between training group and fault type for system control performance.

type (Fig. 5). Interaction contrasts showed a significant interaction for the D&P group and EST/SA group ($F(1, 25) = 7.286, p < 0.05, \eta_p^2 = 0.23$). Moreover, a significant main effect of time ($F(2, 74) = 15.3, p < 0.001, \eta_p^2 = 0.29$) and fault type was observed ($F(1, 74) = 6.66, p < 0.05, \eta_p^2 = 0.15$). Contrasts for main effect of time revealed that performance at T_0 and T_{2w} and at T_{2w} and T_{6w} differed significantly ($p < 0.001$), with the poorest performance generally occurring at T_{2w} .

4. Discussion

Using the training methods EST, EST/SA, and D&P, we examined attention skills and process control performance in familiar and unfamiliar situations over a retention interval of several weeks. We aimed to support novices learning a highly complex and demanding task by providing them with attention-management strategies in order to reduce their mental workload. D&P was successful at enhancing diagnostic performance on familiar system faults, and EST/SA training supported the diagnosis of novel system faults. All in all, EST and EST/SA did not support system control performance as strongly as we had assumed they would.

D&P proved effective at increasing the speed and accuracy with which participants found familiar fault states. Thus, it seems that D&P was successful at providing trainees with clear strategies to cope with the high demands of the task. Participants were guided through the steps of a task so that attentional resources could be applied to learning instead of finding a strategy to cope with the task on their own. The finding that D&P improved diagnostic performance on practiced fault states confirms earlier results of research involving the CAMS task in a comparable experimental

setting (Kluge & Burkolter, 2008). In that experiment, error training, procedure-based training with error-relevant heuristics, and D&P were compared for their effectiveness at enhancing process control performance. D&P emerged clearly as the most effective method for developing the skill of diagnosing familiar fault states, even after retention intervals of 9 and 13 weeks. These findings suggest that the use of D&P was broadened from nondynamic and rather easy cognitive tasks (Shute & Gawlick, 1995) to more dynamic and highly complex cognitive tasks. However, the findings also show that the effectiveness of D&P was confined to familiar fault states. In other words, participants were prepared to apply acquired skills but limited in their ability to adapt skills to new situations. This finding supports the contention that training methods concentrating on teaching procedures rather than teaching knowledge have a restricted range of transfer (Hockey, Sauer, & Wastell, 2007). Lastly, participants in the D&P group were not less motivated than participants in the other two groups, although D&P requires repetitive work on a task. This result might be of special interest to instructors.

Possible explanations why the experiment did not confirm some of the assumptions regarding EST are discussed. The aim of the EST was to enhance attention management by introducing emphasis changes on components of a task. Thus, participants were provided with strategies to reduce mental workload and improve performance. The selection of EST as a training approach for process control was carefully based on empirical evidence, theoretical considerations, and its successful use in different contexts (Gopher, 2007; Gopher et al., 1989). Even so, our experiment showed EST to have only slightly positive interaction effects and did not detect any clear advantage for EST. We note, however, that the present study was more complex in task and design than previous ones on this subject. First, we used a complex and dynamic process control environment that entailed fewer psychomotor and higher cognitive demands than have tasks previously posed in like settings. CAMS involves task management activities such as monitoring to detect deviance and changes, actions to stabilize the system, retrieving information, diagnosing, planning, forming rules, and evaluating actions (Burkolter et al., 2007; Ormerod, Richardson, & Shepherd, 1998). The Space Fortress game employed by Gopher et al. (1989) involved manual control and discrete, precise motor responses, visual scanning and monitoring, memory requirements, and decision-making. Similarities between these tasks may have been overestimated. Second, we used a multifactorial design that included three different points in time. Previous findings concerned positive effects on performance at the end of training, not skill retention over several weeks. All in all, the “take-the-best” application and transfer of EST to a different field seemed to entail unexpected difficulties, and it was not as successful in this study as in previous ones (see Gopher, 2007). Methodologically, it could be argued for a larger sample, for small sample sizes increase the likelihood of incorrectly concluding that there is no statistical difference (Cook, Campbell, & Peracchio, 1990). However, the statistical significance one might gain with a larger sample will not necessarily improve the experiment’s practical significance.

Table 6
Diagnostic performance as a function of training and fault type (SD in parentheses).

Fault type	D&P			EST			EST/SA		
	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}
<i>Diagnostic errors (%)</i>									
Practiced	21.5 (12.8)	51.3 (17.3)	33.3 (19.2)	40.0 (29.4)	59.0 (30.9)	41.0 (30.9)	50.0 (35.7)	57.1 (27.5)	52.4 (31.3)
Novel	46.2 (24.7)	51.3 (35.0)	46.2 (32.0)	38.5 (30.0)	51.3 (25.9)	38.5 (18.5)	53.6 (23.7)	52.4 (38.6)	35.7 (33.2)
<i>Diagnostic speed (s)</i>									
Practiced	223.4 (58.9)	307.8 (59.4)	274.4 (70.5)	241.7 (87.4)	347.5 (83.4)	290.8 (92.9)	292.0 (89.5)	332.1 (76.6)	327.1 (94.0)
Novel	300.6 (74.5)	336.5 (84.1)	358.3 (80.7)	271.3 (65.4)	310.3 (89.0)	330.4 (70.0)	298.4 (80.0)	337.6 (115.5)	319.9 (95.5)

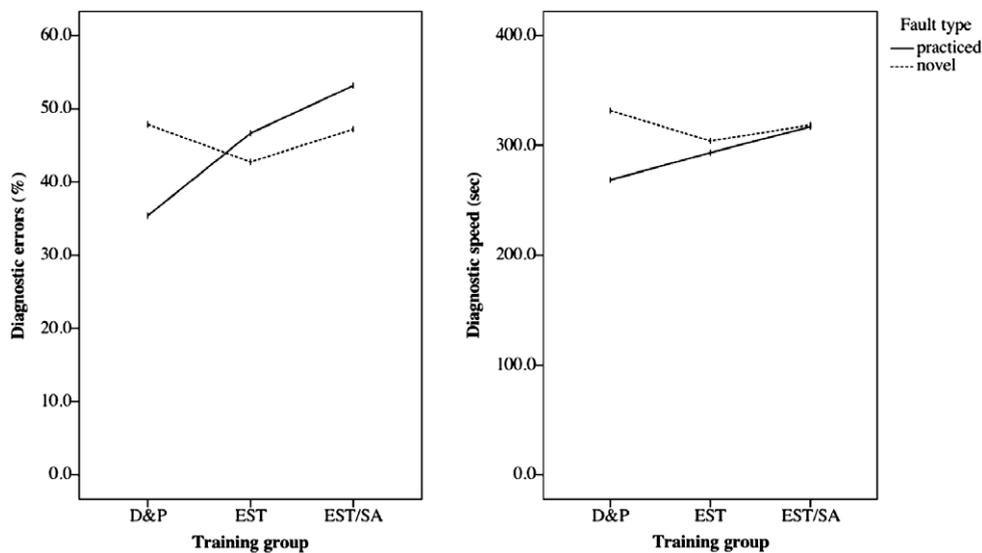


Fig. 5. Interaction between training group and fault type for diagnostic errors (left) and speed (right).

The EST/SA approach showed positive effects for fault-finding in novel situations, but it was not as effective as we assumed it would be for system control and SA. By combining EST and SA training, we aimed both to enhance SA and performance and to support SA by improving attention management. There are reasons to believe that the SA approach might have interfered with EST, an effect especially challenging to novices. In EST, procedures are taught through rule-based instruction, which only implicitly communicates properties of the system (Rasmussen, 1990). To answer the questions relating to SA, however, participants were required to anticipate system states. That is, they had to shift from a rule-based to a knowledge-based level. The combination of training methods that require cognitive processing at different levels may therefore have been too challenging for novices. It seems that the expected positive effects of EST and positive effects of SA were neutralized rather than compounded, at least at this stage of learning. One could consider delaying EST/SA training (see Schneider, 1985) until, say, D&P has helped participants firmly establish a procedure for the main tasks. This sequencing could enhance attention performance by automating a task and thereby freeing attention resources for accomplishing another task (Wickens & McCarley, 2008).

Regarding the measurement of SA, Vidulich (2003) assumes a rich interplay between specific memory of the current situation and a skilled individual's long-term memory. Long-term working memory is assumed to serve as the basis to answer questions during the freeze of the simulation. Possibly, participants had not yet developed enough expertise as a basis to answer the SA questions. With respect to methodological issues, results showed that the SA measurement did not significantly differ from one training group to the next. This finding might be an indication of validity issues with the SA measure we employed—based on the SA Control Room Inventory (Hogg et al., 1995). We note that all the participants scored relatively low regardless of what training group they were in, suggesting a floor effect. Further research employing the novel SA measure is needed to test this possibility.

Surprisingly, performance on system control decreased from the second testing session (two weeks after training) to the third (six weeks after training). By contrast, diagnostic performance was poorer at the second testing session and improved at the third one. These results concerning skill retention after training were similar for all training groups which is an interesting finding, especially because previous research on skill retention has focused mainly on single tasks (see Arthur, Bennett, Stanush, & McNelly,

1998). By contrast, the two tasks in our study—system control and fault-finding—had to be accomplished at the same time. Therefore, findings from single-task studies (e.g. Arthur et al., 1998) might not directly apply for transfer to dual tasks. As abilities of operators in the two main tasks of process control, system control and diagnostic performance are independent (Landeweerd, 1979), it might be possible that skill decay of the two tasks is also dissimilar. This assumption is supported by findings from two experiments with the same simulated process control task (Kluge & Burkolter, 2008; Burkolter, Kluge, & Brand, 2009), in which a similar pattern of results was observed. Whereas system control performance decreased from a first (9 weeks after training and directly after training, respectively) to a second testing session (13 weeks after training and one week after training, respectively), diagnostic performance increased from the first to the second testing session. While these effects have not yet been investigated in detail, we speculate whether participants might have concentrated more on the system control task if training had not taken place long ago. However, at later testing points, when participants might have felt that there were shortcomings in remembering the task, they possibly concentrated more on diagnosis, for which they could find specific information not only on intervention but also on the description of system faults and the symptoms in the manual. However, this issue needs further analysis and research.

Some limitations regarding the study procedure should be pointed out. We conducted the study with students, and not with experienced operators working in process control environments. However, we did invite engineering students to participate in our experiment in order to enhance transfer of study results to process control. These students participated voluntarily in the study as an extracurricular activity, which might indicate a high level of motivation. On the other hand, the extensive training might also have implied an additional workload on them. Moreover, participants did not practice the task between the experimental sessions, which might not directly apply to real-world settings, where operators usually work between training sessions and thus also gain experience between training sessions. Moreover, the process control task was new to the participants, therefore entailing initial learning of complex skills. Further research should determine to what extent such results are transferable to further stages of learning and training.

In summary, the present study aimed at contributing to research of workplace training by applying training methods from

fields such as aviation, police, and nondynamic, motor tasks (e.g. Gopher et al., 1989; Saus et al., 2006) to a process control environment in order to extend established training research findings to a novel work environment. The problem of limited attention capacities has been discussed with respect to complex tasks (Gopher, 1996) and the limitation they present for SA (Endsley, 1995b; Wickens & McCarley, 2008). We aimed to enhance this work by analyzing the training of attention management and allocation of limited attention resources in process control.

This study confirmed D&P as a promising approach for teaching novices to successfully diagnose familiar fault states in process control for up to several weeks after training. We recommend that further research identifies the conditions under which D&P is also effective with experienced operators. The present experiment might serve as a starting point for detailed investigation intended to gather further evidence about the effectiveness of EST and EST/SA training in process control. That work could be a promising step in the effort to provide operators with effective training designed to improve attention management and address the problem of limited attention capacities.

Acknowledgment

This project has been funded by the Swiss National Science Foundation (No. P001-106354).

References

- Arthur, W., Bennett, W., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11*, 57–101.
- Burkolter, D., Kluge, A., & Brand, M. (2009). Cognitive and personality variables of operators: Individual characteristics and process control performance. In D. de Waard, F. L. J. Godthelp, & K. A. Brookhuis (Eds.), *Human factors, security and safety* (pp. 77–88). Maastricht: Shaker Publishing.
- Burkolter, D., Meyer, B., Kluge, A., & Sauer, J. (in press). Assessment of structural knowledge as a training outcome in process control environments. *Human Factors*.
- Burkolter, D., Kluge, A., Schüler, K., Sauer, J., & Ritzmann, S. (2007). Cognitive requirement analysis to derive training models for controlling complex systems. In D. de Waard, G. R. J. Hockey, P. Nickel, & K. A. Brookhuis (Eds.), *Human factors issues in complex system performance* (pp. 475–484). Maastricht: Shaker Publishing.
- Cannon-Bowers, J. A., Rhodenizer, L., Salas, E., & Bowers, C. A. (1998). A framework for understanding prepractice conditions and their impact on learning. *Personnel Psychology, 51*, 291–320.
- Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 517–526.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*, 678–707.
- Cook, T., Campbell, D., & Peracchio, L. (1990). Quasi experimentation. *Handbook of industrial and organizational psychology* (2nd ed.) (Vol. 1, pp. 491–576). Palo Alto, CA, US: Consulting Psychologists Press.
- Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*, 65–84.
- Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.
- Endsley, M. R., & Robertson, M. M. (2000). Training for situation awareness in individuals and teams. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 349–365). Mahwah: Lawrence Erlbaum Associates Publishers.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.
- Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). *Analysis of variance designs. A conceptual and computational approach with SPSS and SAS*. Cambridge: Cambridge University Press.
- Ginzburg, S., & Dar-El, E. M. (2000). Skill retention and relearning - a proposed cyclical model. *Journal of Workplace Learning, 12*, 327–332.
- Gopher, D. (1996). Attention control: Explorations of the work of an executive controller. *Cognitive Brain Research, 5*, 23–28.
- Gopher, D. (2007). Emphasis change as a training protocol for high-demand tasks. In A. F. Kramer, D. A. Wiegman, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 209–224). New York: Oxford University Press.
- Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36*, 387–405.
- Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica, 71*, 147–177.
- Hagman, J. D., & Rose, A. M. (1983). Retention of military tasks: A review. *Human Factors, 25*, 199–213.
- Hockey, G. R. J., Sauer, J., & Wastell, D. G. (2007). Adaptability of training in simulated process control: Comparison of knowledge- and rule-based guidance under task changes and environmental stress. *Human Factors, 49*, 158–174.
- Hogg, D. N., Follesø, K., Strand-Volden, F., & Torralba, B. (1995). Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics, 38*, 2394–2413.
- Kluge, A. (2008). What you train is what you get? Task requirements and training methods in complex problem-solving. *Computers in Human Behavior, 24*, 284–308.
- Kluge, A., & Burkolter, D. (2008). Training principles for process control tasks and innovative approaches to training. In *Joint international symposium of ISSNP 2008/CSEP 2008/ISOFIC 2008: Toward innovative nuclear safety and simulation technology* (Vol. 2, pp. 99–106). Harbin, China: Harbin University Press.
- Kluge, A., Sauer, J., Schüler, K., & Burkolter, D. (2009). Designing training for process control simulators: A review of empirical findings and current practices. *Theoretical Issues in Ergonomics Science, 10*, 489–509.
- Kriedemann, I. (2008). *Training of process control at the PCK refinery*. 11. Simulator training meeting, May 2008, Vienna, Austria.
- Landeweerd, J. A. (1979). Internal representation of a process, fault diagnosis and fault correction. *Ergonomics, 22*, 1343–1351.
- Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5*, 161–171.
- Meyer, B. (2008). *The effects of computer-elicited structural and group knowledge on complex problem solving performance - An application of two computer-based tools for knowledge elicitation*. Open access doctoral dissertation, Humboldt University Press, Berlin.
- Mohr, L. B. (1995). *Impact analysis for program evaluation* (2nd ed.). Thousand Oaks: SAGE.
- Ormerod, T. C., Richardson, J., & Shepherd, A. (1998). Enhancing the usability of a task analysis method: A notation and environment for requirements specification. *Ergonomics, 41*, 1642–1663.
- Rasmussen, J. (1990). Mental models and the control of action in complex environments. In D. Ackermann & M. J. Tauber (Eds.), *Mental models and human-computer interaction 1* (pp. 41–69). North-Holland: Elsevier.
- Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar big-five markers. *Journal of Personality Assessment, 63*, 506–516.
- Sauer, J., Burkolter, D., Kluge, A., Ritzmann, S., & Schüler, K. (2008). The effects of heuristic rule training on operator performance in a simulated process control environment. *Ergonomics, 51*, 953–967.
- Sauer, J., Hockey, G. R. J., & Wastell, D. G. (2000). Effects of training on short- and long-term skill retention in a complex multiple-task environment. *Ergonomics, 43*, 2043–2064.
- Sauer, J., Wastell, D., & Hockey, G. R. J. (2000). A conceptual framework for designing micro-worlds for complex work domains: A case study of the Cabin Air Management System. *Computers in Human Behavior, 16*, 45–58.
- Saus, E.-R., Johnson, B. H., Eid, J., Riisem, P. K., Andersen, R., & Thayer, J. F. (2006). The effect of brief situational awareness training in a police shooting simulator: An experimental study. *Military Psychology, 18*, 3–21.
- Schendel, J. D., & Hagman, J. D. (1982). On sustaining procedural skills over a prolonged retention interval. *Journal of Applied Psychology, 67*, 605–610.
- Schmidt, R., & Bjork, R. (1992). New conceptualizations of practice. Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.
- Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors, 27*, 285–300.
- Shebilske, W. L., Goettl, B. P., & Garland, D. J. (2000). Situation awareness, automaticity and training. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 303–323). Mahwah: Lawrence Erlbaum Associates Publishers.
- Shute, V. J., & Gawlick, L. A. (1995). Practice effects on skill acquisition, learning outcome, retention, and sensitivity to relearning. *Human Factors, 37*, 781–803.
- Spiro, R. J., Feltovich, P. J., & Coulson, R. L. (1996). Two epistemic world-views: Prefigurative schemas and learning in complex domains. *Applied Cognitive Psychology, 10*, S51–S61.
- Tsang, P. S., & Vidulich, M. A. (2006). Mental workload and situation awareness. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 243–268). New York: John Wiley & Sons.
- Vicente, K. J., Mumaw, R. J., & Roth, E. M. (2004). Operator monitoring in a complex dynamic work environment: A qualitative cognitive model based on field observations. *Theoretical Issues in Ergonomics Science, 5*, 359–384.
- Vidulich, M. A. (2003). Mental workload and situation awareness: Essential concepts for aviation psychology practice. In P. S. Tsang & M. A. Vidulich (Eds.), *Principles and practice of aviation psychology* (pp. 115–146). Mahwah: Lawrence Erlbaum Associates.
- Wickens, C. D. (1998). Common sense statistics. *Ergonomics in Design, October*, 18–22.
- Wickens, C. D., & McCarley, J. S. (2008). *Applied attention theory*. Boca Raton: CRC Press.
- Wonderlic. (2002). *Wonderlic personnel test*. Libertyville: Wonderlic Inc.