Virtual training: Making reality work?

Johanna Bertram a,b,*, Johannes Moskaliuk b,1, Ulrike Cress a,2

a Knowledge Construction Lab, Knowledge Media Research Center, Schleichstr. 6, 72076 Tuebingen, Germany
b Department of Applied Cognitive Psychology and Media Psychology, University of Tuebingen, Schleichstr. 4, 72076 Tuebingen, Germany

Abstract

Team training seems to be crucial for the success of teams such as those in emergency services, the police or fire fighting. To carry out an operation successfully, intense training for complex collaborative tasks needs to be provided for all team members. In our study we applied a virtual training environment to train police personnel for complex collaborative tasks. The virtual training group was compared to a group with standard training and to a control group. The data show that the standard training resulted in more motivation, perceived value of the training and knowledge after the training session than virtual training. But with regard to the learning transfer measured by the behaviour in a real and complex situation, the virtual training was as good as the standard training. Both outperformed the control group.

1. Introduction

Emergency service personnel such as the police, fire fighters or the army have to be prepared for various operations, and continuous training is essential for their success and security. But training in real situations is very costly and complex, and only a small variety of real scenarios can be included. Virtual training environments (VTEs) are helpful in training for such complex collaborative tasks, especially if training in real situations is not possible (Moskaliuk et al., 2013a, 2013b). Especially services such as the police, army, emergency services or fire fighters can profit from virtual training to train for scenarios which preclude reality training because of cost, danger or effort (Rose et al., 2000). Often virtual training is the only option for such training (Romano & Brna, 2001). But there is a lack of empirical evidence about team training for complex collaborative tasks and their transfer to the work environment (Burke & Hutchins, 2007; Delise, Gorman, Brooks, Rentsch, & Steele-Johnson, 2010). Evaluative studies on virtual training are still rare (Chittaro & Ranon, 2007; Jou & Wang, 2013). It remains unclear if virtual training results in the knowledge that is expected of trainees and if the technology used for the training is efficient (Salas, Milham, & Bowers, 2003). A theory-driven design for a virtual training environment (VTE) seems to be crucial for successful learning and training in VTEs (Moskaliuk et al., 2013b), and there is evidence for positive training outcomes in these virtual environments (Moskaliuk et al., 2013a). But the questions remain as to how effective virtual team trainings are and whether the virtually acquired knowledge can be transferred to real situations. Clearly there is a need for empirical work concerning virtual team training for complex collaborative tasks.

In this paper we present a field study of a virtual training session for members of the police force using the VTE ViPOL. We show that knowledge, skills and attitudes (Salas & Cannon-Bowers, 2000) gained during the virtual training for complex collaborative tasks could in fact be transferred to real-world situations, comparable to a standard training. In the next sections we first give a brief introduction to the training context and summarize recent findings on team training and transfer of training content. In the empirical part of the paper we describe our field study that compares a virtual training group with two control groups (one with standard training and one with just written instruction and no training, as usually practiced by the police) and report the results. In conclusion we discuss the results and their impact on the effective implementation of virtual team training for complex collaborative tasks.

1.1. Virtual Training for the police force

This field study is part of a larger project conducted by the police training department of a German federal state. The overall goal of the project is to develop a VTE with the objective of providing adequate training for complex collaborative tasks (Moskaliuk et al., 2013b). Among other issues, the interactions of police ground forces with a helicopter crew was identified as an important
training issue that cannot be included in reality training because of high cost and great effort. The complex collaborative tasks involved in this particular issue are a need for team coordination of the ground forces, communication with the helicopter crew, integrating information from the helicopter crew into the ground operation, awareness of dangerous situations, coping with the sound of the helicopter and remembering knowledge and action courses. Thus, the interaction of ground forces with a helicopter crew can be described as a highly complex collaborative scenario, and the need for training in this case seems obvious. In this study we applied the VTE ViPOL to provide training for these collaborative interactions. The screenshot in Fig. 1 shows the VTE ViPOL, with virtual police officers, suspects, pedestrians, vehicles and the helicopter in a virtual town.

1.2. Training teams for complex collaborative tasks

Operations of action forces are characterized by “rapidly evolving and changing conditions, severe time compression, and high degree of ambiguity and uncertainty” (Cannon-Bowers & Salas, 2010, p. 18). In such situations the individuals have to work as a team, defined as “a set of two or more individuals who interact interdependently and adaptively toward a common goal or objective” (Cannon-Bowers & Salas, 1998, p. 83). The team task requires “multiple sources of information, task interdependencies, coordination among members, common and valued goals, specialized member roles and responsibilities, task-relevant knowledge, intensive communication, and adaptive strategies to help respond to change” (Paris, Salas, & Cannon-Bowers, 2000, p. 1052–1053). To solve the team task, each team member needs specific knowledge, skills and abilities (Delise et al., 2010). Team training needs to impart to trainees an understanding of multiple factors: of the equipment they use, of the demands of the task and its environmental effects, of the decision-making process, of their own role in the task, and finally, of the knowledge, skills and attitudes of their teammates (Cannon-Bowers & Salas, 2010). Team members have to know how to interact and communicate with each other and how to solve their task collaboratively (Cannon-Bowers & Salas, 2010). Decisions need to be made quickly, and therefore it is essential that all team members know what their own role is and what their teammates' knowledge, skills and attitudes are (Cannon-Bowers & Salas, 2010). Cannon-Bowers and Salas (1998) emphasize the need for collaboration of teammates and defined roles and functions within the team for effective teamwork. They describe team performance as a combination of jobs skills on the one hand, and “being an effective team member” (p. 84) on the other hand. For effective teamwork implicit coordination in the team, reached through shared mental models and shared understanding of a situation within team members (Cannon-Bowers & Salas, 1998), is very important. Not only performance has to be trained, but team skills (Dobson et al., 2001). Teamwork is dependent upon three competencies that should be considered in training: cognition/knowledge, behaviour/skills and attitudes (Salas & Cannon-Bowers, 2000). Cognition/knowledge refers to the relevant knowledge that each team member needs in order to perform several team tasks. Behaviour/skills refers to the skills team members should possess to perform adequately. Finally, the attitudes of each team member and how they approach the task are important for teamwork (Salas & Cannon-Bowers, 2000).

Providing feedback on learners' performance support their acquisition of procedural knowledge (Serge, Priest, Durlach, & Johnson, 2013). While training for complex team tasks, individual team members often do not have the cognitive resources for self-monitoring or the reflection of team processes (Tannenbaum, Smith-Jentsch, & Behson, 2010), but providing opportunity for such reflection would be important for efficient team training. Tannenbaum and colleagues (2010) propose a team learning circle (p. 251), starting with a pre-brief where roles, strategies and goals are defined, followed by a phase of team performance, which can consist of role play, work samples or simulations. The training should then be observed by the trainers, and each individual team member as well should monitor problems in the team process. Then the individual observations of trainees and trainers should be discussed in a post-action review, focusing on team performance and team processes (Tannenbaum et al., 2010).

Hackman and Morris (1974) underline the importance of the interaction process in a team. The link between input factors (such as individual skills, group factors or environmental factors) and performance outcomes can only be identified through the group interaction process (Hackman & Morris, 1974). To understand the relation between input and output in teamwork, the interaction process of a team needs to be analysed (Hackman & Morris, 1974). DeChurch and Mesmer-Magnus (2010) identified collective cognition as an important factor for team performance in their meta-analysis and found that team cognition has a positive relation to team process. Their study supports the theory that behavioural processes, motivational states and cognitive states are all crucial for team effectiveness. For the development of team cognition they suggest a team task analysis and a focus on training and leadership. But they point out that there is still a lack of research both on how team cognition forms and on explicit team processes (DeChurch & Mesmer-Magnus, 2010) Delise et al. (2010) showed an effect of team training on cognitive outcomes which was larger compared to that of team building.

To sum up, the literature on training for complex collaborative tasks highlights the importance of team interaction processes and implicit coordination within a team (Cannon-Bowers & Salas, 1998; Hackman & Morris, 1974), describes self-monitoring or the reflection of team processes as relevant for effective learning in teams (Tannenbaum et al., 2010) and identifies behaviour, motivation and cognition as important factors for team effectiveness (DeChurch & Mesmer-Magnus, 2010). Team training should include the knowledge, skills and attitudes relevant for the trained task (Salas & Cannon-Bowers, 2000). VTEs deliver various possibilities for the enhancement of training, such as replay functions or the swap of perspectives (Moskaliuk et al., 2013b). In our former work, we showed that a VTE can be used to provide training for complex collaborative tasks, and use of the VTE yields positive results in reactions, learning and behaviour (Moskaliuk et al., 2013a). But will the virtual training for a task be transferred to a complex collaborative scenario in a real situation when the team has to accomplish the task? Will the training prove to be effective if we not only measure individual knowledge but also team efficacy as indicators of effectiveness? In the next section we examine...
conditions for the transfer of virtually acquired knowledge to reality.

1.3. Transfer of training

Holton (1996) describes the relation between learning and individual performance and their transfer as dependent upon the motivation to learn, the transfer design and the transfer climate. Already in 1988, Baldwin and Ford (1988) had proposed a model of training transfer which included trainee characteristics, trainings design and work environment. Yamnill and McLean (2001), after a literature review on training theories, came to the conclusion that all three factors mentioned above contribute to understanding the transfer of training contents. They argue that transfer of training can be reached through three steps: linking the training to strategic goals of the organization, providing explicit performance goals and involving trainees in the design of training. Burke and Hutchins (2007) also base their meta-analysis of training transfer on the Baldwin and Ford model, because research on training transfer continues to refer to these same three factors mentioned above: individual, intervention, and environmental factors. They define transfer as “the application of trained knowledge and skills on the job” (p. 265, Burke & Hutchins, 2007). Concerning individual or learner characteristics, which positively affect transfer, their analysis delivered support for cognitive ability, self-efficacy, pre-training motivation, anxiety, perceived utility, career planning and organizational commitment. As important variables of intervention design, their results revealed positive transfer effects for learning goals, content relevance, practice and feedback, behavioural modelling and error-based examples. For the work environment they found transfer climate, supervisory support, peer support and opportunity to perform as key influences on transfer. Because they had already found a lot of supporting factors for training transfer, they focused further on formulating the need for research which would concentrate on the evaluation of transfer and performance outcomes of technology based training (Burke & Hutchins, 2007).

Some studies showed that VTEs can be used to provide training in specific skills, such as spatial perception, route learning or the reactions in certain emergency situations (Arthur, Hancock, & Chrysler, 1997; Gambirini, Cottle, Spagnoli, Varotto, & Mantovani, 2003; Lloyd, Persaud, & Powell, 2009; Rose et al., 2000) and that VTEs can be applied to evaluate individual differences in spatial cognition (Moriganti, Carassa, & Geminiani, 2007). There is also evidence that a 3D presentation of visual information in a virtual environment supports collaborative decision-making compared to a 2D presentation (Van Der Land, Schouten, Feldberg, Van Den Hooff, & Huysman, 2013). VTEs keep learners engaged and motivated (Monahan, McArdle, & Bertolotto, 2008) and have proven useful in pedagogical contexts (Mikropoulos & Natsis, 2011). They encourage a wide variety of different learning styles and support collaborative knowledge building and reflective thinking (Mikropoulos & Natsis, 2011) as well as experiential learning (Jarmyn, Traphagan, Mayrath, & Trivedi, 2009). But operations of the police, the army, or other emergency services are much more complex. Training must be not only for single reactions, but also for complex team collaboration.

The effects of a training can be enhanced through adequate instructions adapted to the pre-training, during-training and post-training phase (Vogel-Walcutt, Fiorella, & Malone, 2013). To foster transfer of such complex team tasks, Salas, Tannenbaum, Kraiger, and Smith-Jentsch (2012) propose debriefings during and/or after a training session where trainees reflect on the training. Furthermore, they highlight the importance of a job environment where the newly acquired knowledge and skills can be applied with a team leader who encourages the application of training contents (Salas et al., 2012). Johnson, Suriya, Won Yoon, Berrett, and La Fleur (2002) showed that virtual team performance resembles team performance in reality. Team performance in a VTE depends on whether procedures are established within the team, how conflicts are resolved and if collaboration occurs (Johnson et al., 2002). Bacon, Windall, and MacKinnon (2012) highlight the need for realistic virtual training, that is, training which simulates the information overload and pressure during an operation, especially when the training is for crisis management. During the training an emotional state should be invoked in the trainees similar to the emotional state they would be confronted with during an actual operation, because of the strong effect of emotions on decision-making, problem-solving and learning (Bacon et al., 2012). This is in line with the suggestion of Romano and Brna (2001) that virtual training for action forces such as fire-fighters should provide first-person experiences and situated learning without time for conscious thoughts, but with the possibility of gaining knowledge. Presence or immersion in the virtual world seem to be crucial for the development of first-person experiences (Winn, 1993). The knowledge gained is then “direct, personal, subjective and often tacit” (p. 2) and resembles experiences in the real world (Winn, 1993). A first-person perspective can enhance learning in a virtual environment (Lindgren, 2012). Mental representations acquired in a virtual world might be similar or close to representations of real objects (Arthur et al., 1997). Experiences which a trainee has in a VTE become recallable, along with experiences in the real world (Tichon, 2007) and VTEs enable the construction of context dependent knowledge (Mikropoulos & Natsis, 2011).

Kozak, Hancock, Arthur, and Chrysler (1993) argue that transfer from virtual training to reality is definitely enhanced when the context of the virtually trained task resembles the task in reality. Fidelity to real equipment and environment and structuring virtual training according to psychological factors seem to be important for transfer (Dorsey, Campbell, & Russell, 2009). With virtual simulations or games, realistic scenarios can be explored, and learning and performance can be enhanced (Salas et al., 2012). Also to facilitate transfer, the virtual training should be designed with specific training goals in mind (Salas et al., 2003). The integration of a VTE into an educational curriculum seems to be promising (Chittaro & Ranon, 2007), and didactic context seems to be important for effective implementation of virtual training (Mikropoulos & Natsis, 2011).

To examine whether training content is useful to the trainees, Salas et al. (2003) highlight the importance of assessing the trainees’ perception of relevance and value of a training. This kind of utility data predicts the transfer of knowledge and the impact of training on job performance (Alliger, Tannenbaum, Bennett, Traver, & Sholten, 1997). Learning measures alone do not show if the knowledge gained will be applied, but a positive attitude towards the training makes transfer more probable (Alliger et al., 1997). The motivation of participants to learn and improve is higher when they accept the training and see a personal relevance for their job (Ryan & Deci, 2000).

Models for the transfer of training have in common the three main conditions for transfer referred to earlier in this text: individual, intervention and environmental factors (Baldwin & Ford, 1988; Burke & Hutchins, 2007; Holton, 1996; Yamnill & McLean, 2001). For training transfer in general a positive attitude towards the training enhances transfer (individual), debriefings during and after training seem to be important (intervention), and the possibility to apply the new skills on the job also fosters transfer (environmental) (Alliger et al., 1997; Salas et al., 2012). To enhance transfer from virtual training to reality, first-person experiences seem to be crucial on the individual level, realistic training designed with special training goals in mind is important on the intervention level, and on the environmental level training should
be embedded in an educational curriculum (Bacon et al., 2012; Chittaro & Ranon, 2007; Mikropoulos & Natsis, 2011; Romano & Brna, 2001). In a search through the training transfer literature, a lack of studies on the transfer of virtual training to reality is apparent. Recent studies concentrate on the development and transfer of specific individual skills through virtual training, whereas other studies focus on team training of complex collaborative tasks but do not use virtual training. As there is evidence for the transfer of skills from virtual training to reality and for the efficiency of team training for complex collaborative tasks, we argue that a VTE, if carefully designed, can be used to train teams for complex collaborative tasks.

1.4. Evaluation of virtual training

The goals of our study are to evaluate a VTE with respect to its transfer to reality and compare it to other training methods. The model of Kirkpatrick (1976) proposes four factors to be evaluated as important outcomes of a training method: reactions, learning, behaviour and results. Reactions can be described as the satisfaction of trainees with the training (measured by self-reports). Learning means the knowledge and skill improvement of trainees through the training. Behaviour summarizes observables changes in performance. All organizational or personnel changes, like reduced costs or a better work climate, are encompassed in the results.

Salas et al. (2012) recommend the model of Kirkpatrick (1976) for evaluating training methods, because this model developed into a widely accepted, basic model for evaluation in organizations, and therefore can be used as a basis for comparing evaluations. They suggest measuring trainee reactions, learning, behaviour and results (Salas et al., 2012). We weighed these factors and decided to measure learning, knowledge transfer and reactions to the training. Our choice of outcomes to measure is compatible with the dimensions of teamwork as defined by Salas and Cannon-Bowers (2000): knowledge (learning), skills (knowledge transfer) and attitudes (reactions to the training).

In our study we measured learning with a knowledge acquisition pre- and post-test. Knowledge transfer was measured through expert interviews and judgments in a complex task performed in a real situation. Kirkpatrick (1976) suggests simply asking employees about their behaviour after training and using these responses to assess behavioural changes through the training. Alliger et al. (1997) argue that training transfer must be measured as on-the-job performance, and any other attempts to measure behaviour should only be classified as behavioural demonstration of knowledge acquired. This is in line with Delise et al. (2010) who found that only a few field studies on team training have been conducted and still only a small number of studies measure transfer to the work environment. They formulate the need for evaluative measures on the team level (Delise et al., 2010). In response to this and considering all of this previous work, in our study we combined individual measures of reactions and learning with team measures of transfer.

In evaluating reactions to the training, we distinguished between motivational outcomes and perceived value of training. Motivational outcomes describe whether the trainees accepted the training, were satisfied with the training and perceived the training content as relevant for their daily work. These subjective reactions are important to be assessed, because the motivation of participants to learn and improve is stronger when they accept the training and see the training as personally relevant for their job (Ryan & Deci, 2000). The perceived value of training focuses on the subjective learning success of the participants, whether they feel secure and prepared for emergency situations, and if they perceived the training as realistic. This kind of utility data is related to on-the-job performance and might even be more strongly related to transfer than measures of learning (Alliger et al., 1997). Furthermore, these variables might indicate if the trainees perceived the training as useful (Salas et al., 2003).

1.5. Research question

We argue that the VTE ViPOL can be used to train teams for complex collaborative tasks. To support this assertion, we try to answer pertinent research questions:

(1) Does the virtual training yield positive results in reactions, learning and transfer to a complex task in reality?

In further support of the argument we compare a virtual training group with a standard training group and a control group to answer related research questions:

(2) Do the virtual training and standard training groups outperform the control group?

(3) Does the virtual training group at least perform as good as the standard training group?

2. Method

2.1. Virtual training environment ViPOL

We used the virtual training environment ViPOL which was designed by the software company TriCAT together with the police training department of a German federal state. ViPOL is an adaptive desktop VTE that can be adjusted to different training needs, including such elements as weather conditions, landscape, town-scape, population or vehicles. Up to 12 trainees can train together at the same time and can use the VTE from widely distributed locations. Participants use the keyboard and mouse to navigate and headsets to communicate via simulated radio. Trainers can swap positions and change roles and functions of trainees, and can use a replay function for debriefings. For our study we chose as a training topic the interaction of ground forces with a helicopter crew (for more details see Moskaliuk et al., 2013b).

2.2. Design

We applied a between-subject design and assigned the subjects to a virtual training, a standard training or a control group. In the virtual training group the virtual training environment ViPOL was applied and embedded in a didactic concept by the police training unit (Moskaliuk et al., 2013b). The standard training was also designed by the police trainers and took place in a nearby training area. The control group spent their time with an unrelated task. As dependent variables after the training we measured reactions to the training (motivational outcomes and perceived value of the training), learning (knowledge test) and knowledge transfer to a complex task (judged by trainers and the helicopter crew) for all three groups (virtual, standard and control). To measure knowledge transfer to a complex task we designed two real-world scenarios, where the trainees had to solve a complex team task together with a real helicopter under realistic conditions. Thus our field study combines a profound experimental design with performance measurements under real world conditions. Despite the high costs and effort required by the helicopter operations, we were able to train and test 24 individuals. But because this meant 8 participants per condition, we provide descriptive results and not inferential statistics.
2.3. Dependent variables

The measurement of reaction to the training and learning follow our past study (Moskaliuk et al., 2013a). Reactions to the training were measured after the training and after the transfer to a complex task in a real situation. With regard to reactions, we distinguish between motivational outcomes and perceived value of training.

2.4. Reactions to the training – motivational outcomes

2.4.1. Acceptance

To assess the participants’ willingness to train we used three items from Kopp, Dvorak, and Mandl (2003) (sample item: “I would recommend to my colleagues that they attend the training”).

2.4.2. Satisfaction

Participants’ satisfaction with the training was measured with two individual items (e.g. “I am satisfied with the training”).

2.4.3. Relevance

To assess the perceived relevance of the training for the daily work of our participants, we used six individual items (sample item: “The training had a strong connection to my daily work”).

2.5. Reactions to the training – perceived value of training

2.5.1. Subjective learning success

The subjective impression of the trainees about having gained knowledge was assessed by four questions from an evaluation questionnaire by Kopp et al. (2003) (sample item: “Through the training I gained knowledge about the interaction with a helicopter crew”).

2.5.2. Assurance

Four individual items measured the subjective assurance of participants for future helicopter operations (sample item: “The training makes me feel prepared well for helicopter operations”).

2.5.3. Realism of training

The degree of perceived realism during the training and operation was assessed with four individual items (sample item: “I did not feel like I was in a training”).

2.6. Learning

We used a multiple choice knowledge test to measure how much knowledge the participants acquired through the training. The multiple choice knowledge test was developed together with police trainers and followed defined training goals and target performance levels. It contains 20 items which were pretested (Moskaliuk et al., 2013a). Knowledge acquisition was measured as a repeated measure before and after the training.

2.7. Transfer – complex task (2 Scenarios)

We designed a complex task to assess training transfer through performance in a real situation. All three groups took part in this task. The task was planned by the police trainers and divided into two scenarios. In scenario one the groups went by foot and had to find a target person in a nearby training area with the help of the helicopter crew. All participants knew this training area and the standard training took place in this same environment. In scenario two the groups went by car first and had to find the target person in unknown woodland while communicating with the helicopter crew. This environment was new to all participants, and scenario two was much more complex than scenario one. Both trainers and helicopter crew followed a script for each scenario to minimise the differences among the three groups. The helicopter crew did not know with which group they were communicating.

2.7.1. Transfer – judgement of trainers and helicopter crew

The trainers and helicopter crew judged the quality of the operation of each group with a questionnaire consisting of 6 items according to the identified training goals: knowledge, reactions, awareness of dangerous situations, and communication. Because of the difficulty in judging a group of 6–8 persons at the same time, they were asked to judge only the leader of the group. Example items were “The leader of the team lacked the knowledge to react adequately in this situation.” or “The communication between the team leader and the helicopter crew was problematic.” Items were judged on a 7-point likert scale.

2.7.2. Transfer – interviews with trainers and helicopter crew

The trainers and helicopter crew were interviewed after the real operation to get some qualitative information about the three groups. The interviews were recorded by video and analysed.

2.8. Procedure

The study encompassed two training days and a real operation for all participants two weeks after the training phase. The measurement of learning was conducted as a pre-test before the training and as a post-test after the training. The reactions to the training were measured after the training and after the complex task. The judgements and interviews of the trainers and helicopter crew (transfer) were assessed after the real operation. The procedure is summarized in Fig. 2.

The training started with an introduction to the study together with a questionnaire measuring demographic data and the knowledge acquisition pre-test. The two training groups (virtual and standard) attended a theoretical introduction on helicopter operation with basic information about how to handle such operations and videos of real operations. The introduction was based on defined training goals and target performance levels. The training for the standard training group was based on different training scenarios from reality. There was no helicopter present, but a trainer simulated the communication with a helicopter via radio.

The virtual training group was trained in the virtual training environment VIPOL and participated in different scenarios. The goal of the training was to prepare the trainee for a broad range of daily scenarios, where ground forces had to interact with a helicopter crew. In each scenario, the training team had to execute complex collaborative tasks that required continuous communication both with the helicopter crew and within the team, permanent coordination of different actions, quick decisions under great time pressure and the need to ensure one’s own and other team members’ safety. The scenarios differed in their complexity and started with simpler tasks, such as searching for a non-dangerous person. Other tasks were, for example, to search for an aggressive and possibly armed suspect or to chase a suspect on the run that expressed suicidal intentions. The difficulty of the scenarios were increased, weather conditions were made worse, and the complexity and riskiness of the tasks were increased. The virtual training scenarios followed the same structure as real operations of the police force. The problems and goals of each task were discussed before the scenario started. The control group was occupied with tasks unrelated to the training goals during the training time of the other groups. They only read a written hand-out about the interaction of ground forces with a helicopter crew, which is a standard procedure for untrainable operations or tasks. All groups followed the same schedule for breaks. After the training session the dependent variables...
(the learning post-test and subjective reactions to the training) were measured.

Two weeks later all three groups performed two complex tasks in reality and had to carry out collaborative operations. In this task ground forces had to interact with a helicopter crew to search for a target person in an unknown area. All groups had to solve the task under the same weather conditions during the daytime. A different team leader was chosen for each scenario. To provide a fair comparison of the groups, the trainers of each group nominated the two strongest participants of the team. These two participants then each took a leading part in the operation for either scenario one or scenario two. All groups carried out the same operation and collaborated with the same helicopter crew. The police trainers and the helicopter crew filled out a questionnaire about each training group directly after each scenario to assess their judgement. Additionally they were interviewed at the end of the day. After the operation all participants were asked to fill out a questionnaire about their reactions. The questionnaire did not differ between scenarios, the participants were asked about the whole operation, not divided into scenarios. This measure was parallel to the measurement of reactions to the whole operation after the training. To provide fair conditions for all groups and to avoid giving the participants of the virtual training group an unfair advantage due to the study, both the standard training group and the control group were given the opportunity to participate in a virtual training for helicopter operations after the study was completed.

2.9. Participants

In the study 23 police officers participated voluntarily and without payment, 12 males and 11 females, between the ages of 19–40, mean age 24 years (SD = 5.26). Participants were sorted into one of the three conditions. We made sure that the three groups were balanced according to competence (judged by the supervisors), work experience and gender of the participants. The virtual and standard training conditions consisted of 8 people each, the control condition consisted of 7 people. In the real operation, two participants of the virtual training condition missed because of illness, so the condition consisted of only 6 people for the real operation.

2.10. Data analysis

As mentioned above, the small sample size is due to the high costs and effort of this field study, as we used a real helicopter for the two scenarios measuring knowledge transfer in the three groups. Because of this situation, we based our interpretation of differences among the three conditions on descriptive results only.

3. Results

3.1. Reactions to the training

After the training, the means of the subjective reactions of the two training groups (virtual and standard training) to the training were compared. Results are summarized in Table 1 and Table 2. Cronbach’s alpha for all scales was between $z = .70$ and $z = .93$.

3.1.1. Motivational outcomes

As shown in Table 1, the standard condition produced higher results in acceptance of the training, satisfaction with the training and relevance of the training. The virtual training condition thus did not produce motivational outcomes as high as those of the standard condition.

Table 1
<table>
<thead>
<tr>
<th>Reactions to the training – motivational outcomes.</th>
<th>Standard condition</th>
<th>Virtual condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance (M (SD))</td>
<td>5.79 (0.73)</td>
<td>5.33 (1.51)</td>
</tr>
<tr>
<td>Satisfaction (M (SD))</td>
<td>6.19 (0.80)</td>
<td>5.19 (1.91)</td>
</tr>
<tr>
<td>Relevance (M (SD))</td>
<td>5.18 (0.67)</td>
<td>4.30 (1.33)</td>
</tr>
</tbody>
</table>

Table 2
<table>
<thead>
<tr>
<th>Reactions to the training – perceived value of training.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjective learning success (M (SD))</td>
</tr>
<tr>
<td>Subjective assurance (M (SD))</td>
</tr>
<tr>
<td>Realism of training (M (SD))</td>
</tr>
</tbody>
</table>

Table 3
<table>
<thead>
<tr>
<th>Reactions to the operation – motivational outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction (M (SD))</td>
</tr>
<tr>
<td>Relevance (M (SD))</td>
</tr>
</tbody>
</table>
3.1.2. Perceived value of training

As can be seen in Table 2, the standard condition produced higher results in subjective learning success, subjective assurance and realism of the training. The outcomes for the overall perceived value of training were thus higher in the standard condition than in the virtual training condition.

3.2. Learning

3.2.1. Knowledge acquisition

For all participants, a pre-training and post-training knowledge score was calculated, and the difference between the scores was computed. This knowledge score was controlled for whether or not the participants had read the handout about helicopter operations. Results are shown in Fig. 3.

In sum, we found evidence for our assumption that the two training groups outperform the control group. The standard training lead to better learning results than the virtual training.

3.3. Reactions to the operation

After accomplishing the complex task, which took place in reality two weeks after the training, all three groups (virtual training, standard training and control group) were compared with regard to their reactions to the training. Results are summarized in Table 3 and Table 4. Cronbach’s alpha for all scales was between \( \alpha = .41 \) and \( \alpha = .78 \) for the operation.

3.3.1. Motivational outcomes

With regard to satisfaction with the training, the virtual training condition showed the highest scores followed by the standard condition. The control condition reflected the lowest score in satisfaction. But the control condition scored highest in relevance of the complex task, where the other conditions showed lower scores.

3.3.2. Perceived value of operation

The virtual condition showed the highest score in subjective learning success, but there was no large difference from the other groups. A larger difference emerged concerning subjective assurance, where the virtual condition participants reported stronger feelings of assurance during the operation. In perception of realism of the operation, the control condition scored slightly lower than the two training conditions.

3.4. Transfer – complex task

3.4.1. Judgement of trainers and helicopter crew

The judgements of trainers and helicopter crew as to how each group had solved the complex task were compared for scenario 1 and scenario 2. The ratings of the trainers and the helicopter crew reached a marginally significant correlation of \( r = .66, p = .05 \). For scenario 1, we found that the standard condition group showed the best results in accomplishing the task during the real operation, in comparison to the control group and the virtual training group. There was no difference between the virtual and control conditions. In scenario 2, the virtual training condition showed the best results in comparison to the control group and to the standard training group. The standard and control conditions showed no observable difference. Means are shown in Fig. 4.

As scenario two was much more complex than scenario one and the environment was unknown to the trainees, the virtual training in this case seemed to have special advantages compared to the standard training.

3.4.2. Interviews of trainers and helicopter crew

We interviewed the trainers and helicopter crew after the operation to investigate the differences among the groups. In sum, we found out that all three groups were good at the end of the scenarios when they had to arrest the target person. This can be explained by the fact that they all had a similar level of experience in detention. But detention was only the last part of the scenarios. In the other parts of the scenarios, the control group was uncoordinated, the roles in the team seemed undefined, and they were identified as the weakest group in the operation. The standard and virtual training groups were both identified as having performed well in the operation. The standard group was very strong in communication and showed good team work with the helicopter crew. The virtual training group was rated best in the first part of the scenarios where they were led by the helicopter crew. The
leading ground officer was very easy to identify, because he/she regularly made the identifying signal. This seems to be an adaptation from the virtual training environment where the avatars made the signal all the time. The ground officers copied this behaviour, which was judged as very useful by the helicopter crew. Furthermore, the helicopter crew realised that the virtually trained officers were used to the noise produced by the helicopter and that they were not disturbed by it. The leading ground officer in the virtual group coordinated the group very well and gave commands to the team. The virtually trained team showed the best tactics in the operation and did not give up responsibility to the helicopter crew.

Thus, after the real operation we found evidence that the virtually trained team had some advantages. They showed better results in the more complex scenario, felt more secure during the operation, and were more satisfied with their performance than the other groups.

4. Discussion

Virtual training for complex collaborative tasks is a newly evolving topic in training research. Efficient training for complex team tasks should be based on knowledge, skills and attitudes that are relevant for the trained task (Salas & Cannon-Bowers, 2000). Individual, intervention and environmental factors should be taken into account to determine the best ways to foster transfer (Baldwin & Ford, 1988; Burke & Hutchins, 2007; Holton, 1996; Yamnill & McLean, 2001). In our study we combined findings on training for complex collaborative tasks with findings on transfer of training and training in VTEs. With the VTE ViPOL we tested our proposition that teams can be trained in a VTE for complex collaborative tasks. We trained a team of police officers for complex collaboration and interaction with a helicopter crew. Our aim was to evaluate the effectiveness of virtual training for complex collaborative tasks in a field study where we could compare the virtual training to standard training and to a control group. The evaluation focused on the measurement of reactions to the training, learning at the individual level, and transfer of training to a real situation at the team level. Results in the reactions and learning after the training session showed an advantage of the standard training. But after completion of the complex reality task, an interesting effect emerged: the participants who had virtual training showed higher scores in their satisfaction with the training and in their subjective assurance they had experienced throughout the training. Both the standard and virtual training groups yielded better results for transfer of training to the complex reality task than the control group. And in the more complex scenario of the reality task, the virtual training group outperformed the other groups. These results might be due to the fact that the virtual training provided a more realistic and dynamic training environment than the standard training. The virtually trained participants practiced with a virtual helicopter and therefore were used to the stressful elements of the helicopter. They had also trained in unknown terrain and had to cope with the complex tasks in a new environment. This might have led to their advantage in the more complex scenario during the real operation, providing some proof for the idea of Rose et al. (2000) who supposed that virtual training leads to more automatic real world performance. They conclude that, therefore, virtually trained participants might have more cognitive capacity for processing interfering tasks during a real operation. Our results might be promising for future research, since they indicate that the realism of virtual training seems to be important for effective transfer to reality, as already described by Bacon et al. (2012) and Kozak et al. (1993). We showed that a VTE can even outperform training in real situations when it virtually reproduces real-life training. Our results also support the ideas and findings of Tichon (2007) and Winn (1993) who described experiences in a VTE as recallable in reality and resembling experiences in the real world. After the virtual training with the VTE ViPOL our officers were prepared for real operations and could recall their training behaviour in a complex, stressful task.

Furthermore, our results yield evidence for the importance of evaluation on the team level. In our study the standard training group showed the best results in reactions to training and learning on the individual level after the training session. But the standard training team was still outperformed by the virtually trained team in a more complex task. These results indicate that individual measurements might not be related to better results on the team level and that learning measures might not be related to the transfer to complex work situations.

Our results are promising and deliver evidence for the effective use of VTEs for training in complex collaborative tasks, but our study has limitations. Our transfer measure (complex task in reality) needed a helicopter and came at high cost and great effort. We argue that this kind of experimental design under field conditions is really worthwhile, as most existing virtual training programs are not sufficiently evaluated. There is hardly any empirical work that directly compares different kinds of training or their impact on knowledge transfer and group performance. In the study presented in this paper, we used a simulated operation with a helicopter to measure knowledge transfer to a real situation. This use of real world behaviour to measure knowledge transfer goes beyond any existing training studies. Our results will provide meaningful insights for the discussion about the efficacy of virtual training.

Such field studies with randomized experimental variations are of great interest, as they provide results with high ecological validity. Because such studies incur very high costs and require a great deal of effort, it is a useful strategy to start with small sample sizes to gain preliminary experience from the field. Due to the huge effort involved in our field study, we could test only a small sample and thus report only descriptive data in this paper. Another limitation was that in the complex reality task, the trainers and helicopter crew judged the leading officer of each group, reducing our group size to two subjects per group in comparison to 6–8 subjects per group in the other measures. Our participants could not be assigned to the condition at random because of the small group size. Instead, we had to assign them according to their competencies to balance the three conditions. Since we were conducting a field study, we could not eliminate environmental influences, which condition is much different when compared to a laboratory study. We also did not measure any organizational results throughout the training. This might be very interesting for further research. But in spite of the limitations of our study, we finally deliver some results that reinforce the theory that there are advantages of virtual training for complex collaborative tasks, especially for the training of emergency service personnel such as the police. We have produced important data to start bridging the gap between the lack of field studies about virtual training and the evaluation of virtual training’s potential. Our work can be applied as a framework for further evaluative field studies.

5. Conclusion

In sum, the VTE ViPOL can be used to prepare police officers for operations that can hardly be trained for in reality. Even if there are no measurable differences in knowledge after a training session, a virtual training might lead to better performance in a task in reality. Our results furthermore indicate that measurement at the team level might prove to be an effective way to evaluate a training method, especially a virtual training method, and the connections
between the training and team effects must also be further investigated.

Acknowledgements

We thank all partners of the project ViPol; especially the Police training division of the federal state Baden Württemberg, Germany, the software company TriCAT (http://www.tricat.net) and the project Polizei-Online of the Ministry of the Interior of Baden Württemberg (http://www.polizei-online.de).

References


