

**CONTACT INFO FOR SAGE:** Anne Helsdingen anne.helsdingen@ou.nl  
PHONE: +31 619348915  
FAX: +31 338883638

**TOPIC CHOICE:** TRAINING, EDUCATION, INSTRUCTIONAL SYSTEMS

**KEYWORDS:**

command and control training, tactical decision-making, decision strategies, transfer of training

**RUNNING HEAD:** Critical Thinking Instruction

## **The Effects of Critical Thinking Instruction on Training Complex Decision Making**

**Anne S. Helsdingen**, Centre for Learning Sciences and Technologies, Open University of The Netherlands, Heerlen, The Netherlands, **Karel van den Bosch**, TNO Defence, Safety and Security, Soesterberg, The Netherlands, **Tamara van Gog**, Institute of Psychology, Erasmus University Rotterdam, The Netherlands, and **Jeroen J. G. van Merriënboer**, School of Health

### **Precis**

Explicit instruction in critical thinking processes such as story building, self-explanation and reflection can benefit training and transfer of military tactical decision-making for relative novice military decision makers.

**Objective:** Two field studies assessed the effects of critical thinking instruction on training and transfer of a complex decision-making skill.

**Background:** Critical thinking instruction is based on studies of how experienced decision makers approach complex problems.

**Methods:** Participants conducted scenario-based exercises in both simplified (Study 1) and high fidelity (Study 2) training environments. In both studies, half of the participants received instruction in critical thinking. The other half conducted the same exercises, but without critical thinking instruction. After the training, test scenarios were administered to both groups.

**Results:** The first study showed that critical thinking instruction enhanced decision outcomes, both during training and the test. In the second study, critical thinking instruction

benefited both decision outcomes and processes, specifically on the transfer to untrained problems.

**Conclusion:** The results suggest that critical thinking instruction improves decision strategy and enhances understanding of the general principles of the domain.

**Application:** The results of this study warrant the implementation of critical thinking instruction in training programs for professional decision makers that have to operate in complex and highly interactive, dynamic environments.

Correspondence should be addressed to Anne Helsdingen, CELSTEC, OUNL, P. O. Box 2960, 6401 DL Heerlen, The Netherlands; [anne.helsdingen@ou.nl](mailto:anne.helsdingen@ou.nl).

Address correspondence to Anne Helsdingen, CELSTEC, OUNL, P. O. Box 2960, 6401 DL Heerlen, The Netherlands; [anne.helsdingen@ou.nl](mailto:anne.helsdingen@ou.nl). *HUMAN FACTORS*, Vol. 52, No. 3, June 2010, pp. xxx–xxx. DOI: 10.1177/0018720809XXXXXX. Copyright © 2010, Human Factors and Ergonomics Society.

## INTRODUCTION

Studies on decision strategies of experienced decision makers in complex environments (e.g., military command and control, crisis management) have shown that experienced decision makers not only have a large body of knowledge, but also apply deliberate problem-solving strategies that differ significantly from novices' strategies (Endsley, Hoffman, Kaber, & Roth, 2007). When faced with a complex and non-routine problem, experienced decision makers collect and critically evaluate the available evidence, seek for consistency, and test assumptions underlying their assessment of the problem (Klein, Moon, & Hoffman, 2006). Cohen and colleagues (Cohen, Freeman, & Thompson, 1998; Cohen, Freeman, & Wolf, 1996; Freeman & Cohen, 1996) developed an instructional concept that combined instruction aimed at acquisition and application of domain knowledge with instruction on explicit problem-solving strategies: Critical thinking instruction.

Several evaluation studies showed promising results. After critical thinking instruction, trainees considered more observations, identified more cause and effect relations, provided better arguments, and made better decisions according to subject-matter experts (e.g., Cohen, Freeman, & Thompson, 1998; Freeman & Cohen, 1996). However, these studies had some limitations: they were conducted in simplified training environments, control groups did not receive any instruction or practice, and it was not established whether critical thinking instruction had effects on transfer of judgment skill to untrained tactical problems. Therefore, we conducted two field studies on the effects of critical thinking instruction on training and transfer of complex decision making in two training programs for military tactical command and control.

Because critical thinking instruction requires trainees to focus not just on the superficial aspects of situations or on mere observations, but on the underlying structures and causes, it is

expected to lead to deeper understanding of the training materials (cf. Berthold, Nückles & Renkl, 2007; Russel, 2002) and thereby to better transfer.

### **Study 1**

The first study was conducted in the Tactical Command Station of a ground-to-air defense battalion of the Royal Netherlands Air Force. The effects of critical thinking instruction on training processes and transfer test performance were assessed.

## **METHOD**

### **Participants and Design**

Sixteen officers of the Royal Netherlands Air Force participated voluntarily (15 male, 1 female; age  $M = 32.31$ ,  $SD = 3.81$ ). Prior experience in ground-to-air defense tasks ranged from one to five years ( $M = 3.69$ ,  $SD = 2.14$ ). The supervising project officer matched participants according to tactical education and experience, and randomly assigned participants from each pair to the critical thinking condition ( $n = 8$ ) or the control condition ( $n = 8$ ).

### **Materials**

*Practice scenarios.* Six practice scenarios were developed. The paper-and-pencil scenarios encompassed a starting point and a description of events specified in time. They required the participants to assess the geographical area, identify priorities in the area to be defended, assess threats and perform certain actions such as planning engagements of targets, deploying sensors and weapons, making a damage repair plan, planning for casualty transportation, reallocating resources, and making new priorities.

The scenario leader introduced the scenario events and provided feedback. Events were specified in minutes from the start (e.g., X = start scenario, event Y starts at X + 5 min., event Z at X + 10 min., etc.) and feedback moments were specified relative to participants' reaction to

events (e.g., after participants' assessment of the situation, provide feedback). Participants played the battle captain's role and the scenario leader all other roles. Each scenario took approximately 45 minutes to complete.

*Critical thinking instruction.* In the critical thinking condition, participants received a paper-based instruction, explaining the four steps in the process of critical thinking (Cohen, et al., 1998): (a) creating a story of the situation, (b) testing that story for conflicting or missing information, (c) evaluating its plausibility and finding alternative stories (i.e., contingency plans), and (d) quick consideration of the need to decide immediately or spend more time on the critical thinking process.

Participants were instructed to think aloud during the whole process (see Ericsson & Simon, 1993). Besides the textual explanation, a hand-out with a graphic representation of the four steps and two demonstrations of critical thinking in real-world situations were provided. In each demonstration, two scenario leaders showed how critical thinking should be used in the scenarios; one of them played the role of participant (i.e., the battle captain) and one of them played the other roles.

Feedback during the practice scenarios was focused on participants' critical thinking processes. Scenario leaders closely monitored participants' verbalizations of their decision-making processes and matched that with the required assessments as specified in the scenario. If they identified missing or wrong assessments, they would ask open-ended questions to encourage the participant to critically reflect on her/his own decision process, acknowledge the missing or wrong assessment, and take recuperating actions. When participants were adequately engaged in the critical thinking process, scenario leaders only made encouraging remarks (e.g.,

‘very well’ or ‘I understand’). Scenario leaders did not provide any other feedback or information.

*Control group instruction.* The control group participants were given a presentation of the organization and projects of the institute where two of the authors worked to engage them during the critical thinking instruction of the other group. They also received a short introduction regarding the purpose and context of the tactical practice scenarios they were about to receive. Furthermore, two example scenarios were demonstrated by two scenario leaders. These were being played as normal command post exercises, without critical thinking. During execution of the practice scenarios, participants received outcome feedback only (e.g., “that was a good assessment”, “you failed to destroy the target”).

*Test scenarios.* The test scenarios were two paper-and-pencil scenarios in which the participants again played the role of battle captain and the scenario leader covered all other roles. Neither outcome feedback nor process feedback was provided during the execution of the test scenarios. Participants were required to think aloud during execution of the test scenarios.

*Command & Control Process Measurement Tool (C2PMT):* Two types of performance measures were gathered during the practice and test scenarios: (1) outcome measures to assess the quantity and quality of the end result, and (2) process measures, to evaluate the strategies, steps and procedures used to accomplish the task. According to the C2PMT (Van Berlo & Schraagen, 2000), performance indicators were specified for both the outcomes and process measures, enabling the scenario leaders to observe and interpret the decision-making processes and outcomes.

The performance indicators were easily scored in terms of whether the behavior was observed or not; this was indicated in the respective column (Yes or No) and an overall grading

(ranging from 1 = very poor, to 10 = excellent) for that performance measure was determined (see Table 1). Prior to the experiment, scenario leaders had used the results of a pilot-study (using the same scenarios but with different participants) to reach a common understanding of assigning grades.

### **Procedure**

The experiment had a duration of approximately 10 hours, and was run in five daily two-hour sessions with one or two participants at the same time. In the first session, participants received the critical thinking instruction or general presentation followed by the two demonstration scenarios. In the three subsequent training sessions, two practice scenarios were played each time, with a 15-minute break between the scenarios.

The order of the six practice scenarios was randomly selected from all possible scenario sequences for each participant. Consecutive practice scenarios were facilitated by different scenario leaders, that is, participants practiced three scenarios with one scenario leader and the other three with another scenario leader.

The scenario leaders provided participants with all relevant scenario input, reacted to their decisions according to the scenario, provided feedback and evaluated performance. The participants were required to verbalize their thoughts during the scenarios to provide the scenario leader insight into their (critical) thinking processes, enabling the evaluation of participants' performance and provision of feedback. About two to three times during a practice scenario, the scenario leader provided feedback, during which the scenario was 'frozen' for a few minutes. Scoring of the performance indicators was done continuously during the scenario. Thus, when participants were *stating the right priorities* or *providing a logical argument for an assessment*, these performance indicators for the performance measure *quality of plans* were ticked off. These

indicators served as a support tool for the scenario leader to be able to grade (between 1 – 10) performance measures at the end of a scenario.

In the fifth session two test scenarios were played. Order and assignment of the test scenarios to scenario leaders was balanced. Again, all participants were asked to think aloud, however, no feedback was given. Participants' verbalizations of their (critical) thinking were used to grade performance.

### **Data Analysis**

Quality of actions and contingency plans were scored and aggregated into a measure of the decision outcome (Cronbach's  $\alpha = 0.76$ ). Similarly, information processing and argumentation were graded and aggregated into a measure of the decision process (Cronbach's  $\alpha = 0.93$ ).

## **RESULTS**

Measures of decision process and outcome were analyzed using a t-test for two dependent samples. In all analyses reported below, a significance level of .05 is set and *Hedge's g* with 95% Confidence Interval (CI) is reported as a measure of effect size, and because of small sample sizes it was corrected for bias using the method reported in Hedges and Olkin (page 80, 1985).

*Training.* The means of the decision process and outcome measures during training are presented in Figure 1. Paired t-tests showed that critical thinking had an effect on decision outcome,  $t(7) = -3.25, p = .01, g = 1.09$  [95% CI = .20; 1.98], with participants in the critical thinking group reaching higher grades ( $M = 5.65, SD = 1.62$ ) than participants in the control group ( $M = 3.78, SD = 0.31$ ). However, no significant effects were found of critical thinking on decision process,  $t(7) = 1.65, ns$ .

*Test.* The means of decision process and outcome measures are presented in Figure 1. Paired t-tests showed a small effect of critical thinking instruction on decision outcome,  $t(7) = -2.35$ ,  $p = .05$ ,  $g = 0.63$  [CI = -.08; 1.34], with the critical thinking group performing better ( $M = 4.94$ ,  $SD = 1.65$ ) than the control group ( $M = 3.44$ ,  $SD = 1.30$ ). Again, no effects were found of critical thinking on decision process measures,  $t(7) = 1.61$ , *ns*.

## DISCUSSION

Participants in the critical thinking condition made better decisions than the control group during training, and this seemed to carry over into the test. However, critical thinking did not benefit the processes of decision-making, only the decision outcomes. But although the process measures failed to show statistically significant effects of critical thinking, a trend can be observed from Figure 1 suggesting that participants from the critical thinking group obtained slightly higher grades than the control group.

It may have been that scenario leaders' scoring was biased by their knowledge of the instructional intervention, favoring the decision strategies of the control group because they did not receive critical thinking support. Nevertheless, the focus of the critical thinking approach on critical assumptions, conflict, missing information, and contingencies is likely to have resulted in the better decision outcomes, although with average grades around 5 (out of max 10), even the critical thinking group does not show adequate performance.

The results suggest that the critical thinking approach might be a useful tool for improving the quality of training tactical decision-making, as earlier explorative studies indicated (e.g. Cohen, et al. 1996; Freeman & Cohen, 1996; Klein et al., 1997). However, this study had some limitations.

First, scenario leaders' were the evaluators of the practice and test scenarios, so their scoring may have been biased. Second, the effect of the instruction has been studied in a simplified task environment. For reasons of transfer it is necessary to investigate whether critical thinking skills can be successfully trained in high-fidelity task environments. Finally, in this study the test tasks were rather similar to the training tasks. When critical thinking leads to deeper understanding of the training material, we would also expect benefits of this type of instruction for test tasks that are different from the training tasks. Hence, a second study was conducted to address these issues, using independent assessments and a more realistic task environment.

## **Study 2**

This study was conducted in the domain of *anti-air warfare* and *anti-surface warfare* at the Operational School of the Royal Netherlands Navy. Teams of one officer and one petty officer played single ship / single threat scenarios in a high fidelity tactical simulator.

## **Method**

### **Participants**

Sixteen officers of the Royal Netherlands Navy volunteered to participate in the experiment (15 male, 1 female; age  $M = 36.08$ ,  $SD = 6.07$ ), with prior sailing experience ranging from 2 to 17 years ( $M = 10.23$ ,  $SD = 5.57$ ). The supervising project officer matched participants according to their tactical education and experience, and assigned participants from each pair randomly to the critical-thinking condition ( $n = 8$ ) and control condition ( $n = 8$ ).

### **Materials**

*Practice scenarios.* Four anti-surface warfare practice scenarios of approximately 120 minutes each were developed. Two of these scenarios were paper-and-pencil scenarios, and two

were scenarios to be run in the high fidelity tactical simulator. All four scenarios had a structure similar to the scenarios in Study 1. Participants played the role of principal warfare officer and assistant and the scenario leader played all other roles. The assistant was responsible for tactical picture compilation, interacted with the operators or the system to gather all information and presented that graphically to the principal warfare officer, who interpreted the picture and decided on which actions (e.g., maneuvers, engagements) to take.

*Critical thinking instruction.* Identical to Study 1.

*Control group instruction.* Identical to Study 1.

*Test scenarios.* Two test scenarios were developed: (a) A post-test scenario that was a variation on the practice scenarios (i.e., anti-surface warfare), (b) a transfer scenario that was different from the practice scenarios on both superficial and structural level (anti-air warfare) scenarios. On board Royal Netherlands Navy frigates, air and surface warfare are different expertise areas and managed by different officers. Both test scenarios were played in the tactical simulator. No feedback was provided during execution of the test scenarios.

*Performance measures:* The same performance measures as in Study 1 were used. In addition, time management and team skills were graded (cf. Smith-Jentsch, Zeisig, Acton, & McPherson, 1998) but these were left out from our analyses. Because scenario leaders and evaluators were used to the official NATO scale, it was decided to use this 4-point scale which ranges from 1 = Unsatisfactory, 2 = Marginal, 3 = Satisfactory, to 4 = Excellent (see Van Berlo & Schraagen, 2000) to grade all performance measures.

## **Procedure**

Prior to the study, scenario leaders assigned to critical thinking teams were extensively briefed on the critical thinking instruction, as well as on how to support the teams in the

application of the critical thinking processes. Scenario leaders assigned to control teams were not informed about the concept of critical thinking and were instructed to provide outcome feedback. The experimenter instructed the two subject matter experts (SME) who evaluated the transfer tests on the scoring procedure and how to use the scale. These SME's were blind to the conditions and were not informed about the concept of critical thinking instruction nor of the purpose and design of the study. Again, scenarios of a pilot study were used to arrive at a common interpretation of performance indicators.

The experiment lasted approximately 20 hours and was run in five daily sessions of approximately four hours. In the first session, participants received the instruction to play the scenarios as a normal command and control exercise. Participants in the critical thinking group received a critical thinking tutorial, followed by a demonstration in which the scenario leader showed how critical thinking should be used in the scenarios. The control group witnessed a presentation of the institute where two of the authors work. After that, a demonstration was given of the same example scenario as the critical thinking group received, but without critical thinking. In the subsequent training sessions, two practice scenarios were played each day, with a 1-hour break in between the scenarios.

The order of practice scenarios was predefined and similar for all participants and scenario leaders were assigned to scenarios because specific expertise was required for specific scenarios. As in study 1, in the first two scenarios, scenario leaders provided participants with all relevant scenario input, reacted to their decisions according to the scenario, provided feedback and evaluated performance on the basis of participants' thinking-aloud. The two subsequent practice scenarios were played in the simulator, thus leaving only giving feedback and evaluating performance for the scenario leaders.

About four to five times during a practice scenario, the scenario leader provided feedback, and ‘froze’ the scenario for a few minutes. Scoring of the performance indicators was done during the scenario. The performance measures were graded four times during the scenario; markers in the scenario description prompted the scenario leader to do this.

In the fourth session, the test scenarios were played in the simulator. Order of the test scenarios was balanced per session: Thus, half of the participants played a transfer test scenario prior to the post-test scenario, and the other half vice versa. All participants were asked to think aloud to enable performance evaluation but no feedback was given. SME’s scored the performance indicators during execution of the scenario when a participant exhibited the behavior as in Study 1, and markers in the paper based scenario description prompted the evaluator to grade all performance measures at that particular moment, four times per scenario.

### **Data analysis**

Participants were graded individually. The mean grades on information processing, argumentation, quality of action plans, and quality of contingency plans were calculated for each team of two participants for the four training scenarios and test scenarios. Scores for information processing and argumentation were aggregated into a measure of decision process (*Cronbach’s*  $\alpha = 0.88$ ). Decision outcome was an aggregate measure of scores on quality of action and contingency plans (*Cronbach’s*  $\alpha = 0.84$ ).

## **RESULTS**

The effects of critical thinking on the decision process and outcome were analyzed using t-test for two dependent samples. In all analyses reported below, a significance level of .05 is set and Hedge’s *g* [with 95% CI] is reported as effect size measure and, because of the small sample size, corrected for bias using the method reported in Hedges and Olkin (page 80, 1985).

## Training

Analysis of the process,  $t(3) = -.31, ns$  and outcome measures,  $t(3) = -1.00, ns$ , showed no significant effects of critical thinking during training.

## Tests

Figure 2 presents the mean grades on all performance measures for both conditions in the tests. Paired t-tests showed that none of the performance measures on the post-test were affected by critical thinking instruction: decision process,  $t(3) = .61, ns$  and outcome  $t(3) = .67, ns$ .

Performance measures on the transfer test were significantly affected by critical thinking instruction. On the decision process, the critical thinking teams received significantly higher grades ( $M = 2.88, SD = 0.37$ ) than the control teams ( $M = 2.20, SD = 0.23$ ),  $t(3) = -4.23, p = .03$ ,  $g = 1.91$  [95% CI = .98; 2.83], and the decision outcomes of critical thinking teams were also higher ( $M = 2.42, SD = 0.36$ ) than those of the control teams ( $M = 1.69, SD = 0.39$ ),  $t(3) = -4.70, p = .02$ ,  $g = 1.67$  [95% CI = .76; 2.58] Using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) to correct for a false discovery rate, the effects on the performance measures remain significant.

## DISCUSSION

This study showed clear benefits of critical thinking instruction on transfer to a task that differed from the training scenarios. Thus, it seems that the critical thinking instruction has provided the participants with a better strategy to approach new problems in the transfer test scenario. The study failed to show benefit of critical thinking instruction on the post-test scenario. This difference with Study 1 may be because the tactical problems in the post-test scenario closely resembled the training scenario, providing participants with the opportunity to rely on their specific memories of similar cases they had experienced during training. As Cohen

et al. (1998) argued, critical thinking is not expected to increase decision quality for routine problems, probably because then there is no need to “search for patterns in the training material, restructuring of problem representations, or metacognitive processes” (Schnotz and Kürschner, 2007).

## **GENERAL DISCUSSION**

The two field studies reported in this paper examined the effects of critical thinking instruction on training complex decision making in command and control. With respect to participants’ performance, it was hypothesized that critical thinking instruction yields better training outcomes than training without critical thinking, that is, higher test performance. The studies support this hypothesis.

Our findings are consistent with those of studies on the effects of critical thinking instruction on command and control decision making (Freeman & Cohen, 1996), nursing (Forneris & Peden-McAlpine, 2006), medical diagnosis (Klein, 1998), and management (Yeo, 2007). Not only do participants show improved decision processes, their decision outcomes, measured in terms of e.g. the quality of plans, were also better. We have shown that critical thinking instruction can also be successfully applied in training decision making skills in interactive and dynamic high fidelity simulators and renders better results than merely conducting exercises in these simulator environments.

The performance gains can probably be partly ascribed to participants being able to apply a better decision strategy as a result of the critical thinking instruction (Cohen et al., 1996; Cohen et al. 1998; Freeman & Cohen, 1996). As Halpern (1997) has shown, when students are taught critical thinking skills, they are less likely to make the typical decision mistakes caused, for example, by confirmation bias.

That we found critical thinking instruction to have an effect on performance on an untrained transfer test scenario in Study 2 may be an indication that participants have gained a deeper understanding of the task content, which enabled them to solve new decision problems that differed from the training problems (cf. Van Merriënboer, 1997; Van Merriënboer & Paas, 1990). An interesting question for future research that cannot be answered based on our data, is exactly what aspects of critical thinking instruction lead to a deeper level of understanding. Possibly, the instruction triggers processes of self-explanation or reflection on the structure of the task (Chi, Basok, Lewis, Reimann & Glaser, 1989; Renkl, 1997). It might be interesting to disentangle the contributions of different processes to training and transfer.

There are some limitations of this field study: First of all, in study 1 the test performance evaluators were also the scenario leaders and thus informed about the experimental manipulation. Secondly, although the evaluators of test performance were blind to the experimental manipulation in study 2, scenario leaders and evaluators could not be balanced over training or test scenarios. This was because the specific training materials required subject matter experts to evaluate them, and because the scenario leaders instructing the control group had to remain uninformed about the critical thinking instruction to prevent them from introducing critical thinking elements in the control group instruction. This may have introduced some confound in the study. Thirdly, the small sample size, especially in study 2, limits the transferability of our findings to other training programs.

Nevertheless, in combination with previous studies, our field studies seem to warrant implementation of critical thinking instruction in training programs. However, implementation studies with larger numbers of participants are needed to provide answers to important questions, concerning how critical thinking instruction can be integrated into an existing curriculum, what

practice instructors need for successful application, and what the long term effects are on decision making performance.

### **KEY POINTS**

- Critical thinking instruction explicitly teaches strategies such as story building, self-explanation and reflection when training military tactical decision making
- Critical thinking instruction can be successfully taught to relative novice military decision makers
- Critical thinking instruction benefits test performance when compared to standard scenario based training, especially transfer to untrained tactical problems.

### **REFERENCES**

- Benjamini, Y., & Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal Of Royal the Statistical Society: Series B (Statistical Methodology)*. 57, 289-300.
- Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17, 564-577.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.
- Cohen, M. S., Freeman, J. T., & Thompson, B. T. (1998). Critical thinking skills in tactical decision making: a model and a training strategy. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 155-190). Washington, DC: American Psychological Association.

- Cohen, M. S., Freeman, J. T., & Wolf, S. (1996). Meta-recognition in time stressed decision making: Recognizing, critiquing, and correcting. *Human Factors*, 38, 206-219.
- Endsley, M. R., Hoffman, R., Kaber, D. B., & Roth, E. (2007). Cognitive engineering and decision making: An overview and future course. *Journal of Cognitive Engineering and Decision Making*, 1, 1-21.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.
- Forneris, S. G., & Peden-McAlpine, C. J. (2006). Contextual learning: A reflective learning intervention for nursing education. *International Journal of Nursing Education Scholarship*, 3, 1-18.
- Freeman, J. T., & Cohen, M. S. (1996). Training for complex decision-making: A test of instruction based on the recognition/metacognition model. *Proceedings of the 3rd International Command and Control Research and Technology Symposium 1996*. Retrieved May 12, 2005 from [www.cog-tech.com/papers/c2/c2\\_1996.pdf](http://www.cog-tech.com/papers/c2/c2_1996.pdf).
- Halpern, D. F. (1997). Students need 'reality checks'. *Education Digest*, 63, 24-26.
- Hedges, L., & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. New York: Academic Press
- Klein, G. A. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Klein, G. A., Moon, B. A., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, 21, 70-73.
- Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, 21, 1-29.

- Russell, A. (2002). The role of epistemic agency and knowledge building discourse to foster interprofessional practice in a Canadian hospital. Paper presented at the *American Educational Research Association Annual Conference, New Orleans*. Retrieved October 30, 2008 from [www.ikit.org/fulltext/2002AERAAnn.pdf](http://www.ikit.org/fulltext/2002AERAAnn.pdf).
- Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory, *Educational Psychology Review, 19*, 469–508.
- Smith-Jentsch, K. A., Zeisig, R. L., Acton, B., & McPherson, J. A. (1998). Team dimensional training: A strategy for guided team self correction. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 271-298). Washington, DC: American Psychological Association.
- Van Berlo, M. P. W., & Schraagen, J. M. C. (2000). A generic assessment tool for evaluating C2 exercises. In *Proceedings Interservice/Industry Training, Simulation and Education Conference (IITSEC) 2000*, Orlando FL (pp. 652-660 [CD ROM]). Arlington, VA: National Training and Simulation Association.
- Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology Publications.
- Van Merriënboer, J. J. G., & Paas, F. (1990). Automation and schema acquisition in learning elementary computer programming: Implications for the design of practice. *Computers in Human Behavior, 6*, 273-289.
- Yeo, R. K. (2007). (Re)viewing problem-based learning: An exploratory study on the perceptions of its applicability to the workplace. *Journal of Managerial Psychology, 22*, 369-391.

Anne S. Helsdingen received her PhD in Educational Technology from the Open University of The Netherlands in 2008. She is a researcher at the Centre for Learning Sciences and Technologies, Open University of The Netherlands, Heerlen

Karel van den Bosch received his PhD in Psychology from the Radboud University Nijmegen, The Netherlands, since 1991. He is a senior researcher at TNO Defence, Safety and Security, Soesterberg.

Tamara van Gog is an associate professor of educational psychology at the Institute of Psychology of the Erasmus University Rotterdam, The Netherlands. In 2006 she received her PhD in Educational Technology from the Open University of The Netherlands, Heerlen. At the time this work was conducted, Tamara van Gog was affiliated with the Centre for Learning Sciences and Technologies, Open University of The Netherlands, Heerlen.

Jeroen J.G. Van Merriënboer is a full professor in educational psychology at the School of Health Professions Education, University of Maastricht in The Netherlands. He received his PhD in Educational Technology from the University of Twente, Enschede, in 1990.

*Date received: August 14, 2009*

*Date accepted: June 2, 2010*

Table 1. *Performance measures and performance indicators*

Category	Performance measure	Indicator	Observed (Yes / No)	Grade
Process (story building & testing)	Information processing	Being able to distinguish between relevant – irrelevant information		
		Recognition of conflicting evidence		
		Identification of missing information		
	Argumentation	Recognition of cause and effect relationships		
		Incorporation of history of the situation		
		Explication of expectations for the future		
		Plausibility of these expectations		
		Recognition of uncertainties		
	Differentiation between observations and assumptions			
Outcome	Quality of actions	Concise		
		Unambiguous		
		Stating the right priorities		
		Conform expert solution		
		Logical result of assumptions and argumentation		
	Quality of contingency plans	Availability of a plan		
		Concise		
		Unambiguous		
	Logical result of (expressed) conflicting evidence or weak assumptions underlying the original plan.			

### FIGURE CAPTIONS

*Figure 1:* Means and *SD* of critical thinking group and control group scores during training and on the test, \* significant at 5% . (1= terrible, 2 = bad, 3 = very weak, 4 = weak, 5 =failing, 6 = passed, 7 = okay, 8 = good, 9 = excellent, 10 = perfect).

*Figure 2:* Means and *SD* of participants of the critical thinking group and control group during training, on the post-test and transfer test, \* significant at 5%. Trn = training, pst = post test, Tsf = Transfer. (1 = unsatisfactory, 2 = marginal, 3 = satisfactory, 4 = excellent)



