

# Dynamic Task Selection in Aviation Training

**Ron J. C. M. Salden (ron.salden@cs.cmu.edu)**

Human-Computer Interaction Institute

Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh, PA 15213, USA

**Fred Paas (fred.paas@ou.nl)**

Open University of the Netherlands

Educational Technology Expertise Centre (OTEC)

P.O. Box 2960, 6401 DL Heerlen, The Netherlands

**Jeroen van Merriënboer (jeroen.vanmerriënboer@ou.nl)**

Open University of the Netherlands

Educational Technology Expertise Centre (OTEC)

P.O. Box 2960, 6401 DL Heerlen, The Netherlands

## Abstract

While the aviation domain is exemplary for its complex cognitive skills, the pace of automation steadily increases making it crucial to train people as effectively as possible. Over the last three decades training programs have evolved a strong focus on personalized dynamic whole-tasks. Adapting training to the individual student's progress is believed to be strongly related to increased training efficiency (Salden, Paas, & van Merriënboer, 2006). Four studies investigated a variety of personalized training methods. Results confirm the hypothesis that personalized instruction can have beneficial effects for the training of complex cognitive skills.

Keywords: complex cognitive skills, personalized instruction, cognitive load, mental efficiency

## Introduction

Technical domains like the chemical industry and aviation incorporate a vast amount of complex cognitive skills in highly demanding working environments. Mistakes can lead to dangerous situations and high costs, yet the available training time in which the complex job skills have to be mastered, is limited. Efficient training that offers trainees a powerful learning environment seems mandatory to ensure that they can acquire skills quickly and adequately, and learn how to apply these skills flexibly in new situations and tasks.

One of the main characteristics of the aviation domain is that each task often contains new elements compared to the previous tasks. In other words, each new task can be considered as a transfer task in which the previously acquired knowledge needs to be applied differently. One should note that besides new elements, each learning task contains the basic skills that have to be acquired. Though the variability and complexity of the learning tasks increase during training, each task builds upon this basis.

During the last three decades, training methods and programs have evolved in three important ways (for an

overview see Salden, Paas, & van Merriënboer, 2006): from static to dynamic, from part-task based to whole-task based, and from group-based to personalized. Especially, the use of personalized selection of learning tasks is believed to be strongly related to increased training efficiency.

Flexible learning on the basis of meaningful learning tasks requires some form of dynamic task selection. An intelligent agent makes decisions about the most optimal learning-task sequence during the training or teaching process. In order to make appropriate decisions, information on the student's progress is used such as indications of the level of performance and the costs related to reaching this performance.

Although many Intelligent Tutoring Systems (ITS) have extended their capacity to adapt the selection of learning tasks to the individual learner's needs by incorporating student models that keep track of a student's performance history, we claim that they are lacking an important aspect of the learning process, namely, cognitive load. Although, the concept of cognitive load is sometimes measured (e.g., Kashihara, Hirashima, & Toyoda, 1995) it has never been used in ITSs as a determinant for task selection. There is no doubt that cognitive load is a crucial factor in the training of complex cognitive skills (e.g., Sweller, 1989; Sweller, van Merriënboer, & Paas, 1998), but usually, only performance measures such as speed and accuracy are used to select learning tasks.

From the viewpoint of cognitive load theory (Paas, Renkl, & Sweller, 2003), dynamic task selection can be superior to fixed task selection as it provides the possibility to adjust the training to the cognitive state of the learner, thereby controlling the load that is imposed on a learner's cognitive system. Although individual measures of performance and mental effort can be used as indicators of the cognitive demands a certain task places on the learner, the combination of both measures is considered a superior

estimate of the cognitive demands in the dynamic selection of training tasks. It is quite feasible for two people to attain the same performance levels, while one of them experiences a very high cognitive load and needs to work laboriously through a very effortful process, whereas the other person experiences a low cognitive load and reaches the same performance level with a minimum of effort. However, most people would agree that the next learning task should be less difficult for the first person than for the second person. Our claim is that task selection, and consequently training efficiency can be improved by taking the combination of performance and cognitive load measures into account. To obtain a good indication of the cognitive load that is imposed on a person's cognitive system, mental effort measurements are used.

Paas and van Merriënboer (1993) have developed a calculational approach for combining measures of mental effort and task performance that allows one to obtain information on the relative efficiency of instructional conditions. Based on Ahern and Beatty's (1979) efficiency view on learning, it is proposed that learners' behavior in a certain training condition is more efficient if (1) their performance is higher than might be expected on the basis of their invested mental effort, and/or (2) their invested mental effort is lower than might be expected on the basis of their performance. Thus, training conditions in which high performance is attained with a low mental effort investment are considered as 'high efficient'. 'Low efficient' conditions are characterized by a combination of low performance and high mental effort.

A first indication that the use of a combined performance and mental effort score can make personalized training more efficient was found in a study by Camp, Paas, Rikers, and van Merriënboer (2001). In this study four methods of task selection in the Air Traffic Control (ATC) domain were compared. In the first method, tasks were presented in a fixed, predetermined simple-to-complex sequence designed according to the 4C/ID-model (van Merriënboer, 1997). In the other three methods, the tasks were presented dynamically, based on either performance, mental effort, or the combination of both (i.e., mental efficiency). Results showed that dynamic task selection leads to more efficient training than non-dynamic task selection. However, dynamic task selection based on mental efficiency did not lead to more efficient training and better test performance than dynamic task selection based on performance or mental effort alone.

Besides such system-controlled task selection, learner-controlled selection may offer another form of personalized dynamic task selection because it gives the students control over what learning tasks they want to practice next. While a clear definition of learner control is missing, most studies in the field of computer-based training operationalize it in two ways: Either learners are given the option to request additional instructional material or they are given the option to bypass instructional material (Crooks & Klein, 1996). A third way was explored in the reported studies in which

students could either select the task complexity, or the learning task itself from the entire task database.

The basic theoretical claim for potential positive effects of learner control (i.e., personalized preference) is that trainees are able to select the appropriate tasks to practice while avoiding a possible overload of their cognitive system, thereby increasing the effectiveness and efficiency of learning (e.g., Borsook & Higginbotham-Wheat, 1991). However, several studies show that low-ability learners experience problems with the control they are given (e.g., Niemic, Sikorski, & Walberg, 1996; Steinberg, 1989). A possible explanation is that the given level of control is often not compatible with the learners' abilities.

According to Bell and Kozlowski (2002), it is critical to design instructional material that provides learners with a level of control they are able to handle. Furthermore, the 'expertise reversal effect' (e.g., Kalyuga, Ayres, Chandler, & Sweller, 2003) indicates that the trainees' increasing expertise level is probably the most important determinant for deciding on the appropriate level of freedom that is given to them. Support for this claim was found in recent studies (van Merriënboer, Schuurman, de Croock, & Paas, 2002; Salden, Paas, & van Merriënboer, in press), which showed that learners who are given an appropriate level of control over task selection are well able to select their own learning tasks.

## Research Questions

The main research question is how dynamic task selection can be used to optimize training programs, the learning process, and transfer test performance. More specific research questions focus on the different types of information that are required to effectively use dynamic task selection and on the role of the trainees themselves in this task selection process. For example, do performance measures contain sufficient information for dynamic task selection or are other measures, such as invested mental effort, also important to take into account? And to what extent are trainees able to fulfill an active role in the process of task selection?

These research questions were addressed in four experiments with the first two studies focusing on Air Traffic Control (ATC) learning materials and the latter two studies on learning materials for the cockpit automation of the Flight Management System (FMS).

## Calculations and Methodology

Throughout all four experiments performance was measured on 5-point scale (1 = very low; 5 = very high) concerning several performance variables. For the two ATC studies the mean performance was a combined measure of (a) number of commands, (b) time outside airway, (c) time without separation between airplanes, and (d) number of gate hits (i.e., safely directed airplanes to end point). Concerning the two FMS studies the mean performance consisted of (a) number of commands, (b) number of changes in flight route, and (c) amount of time pressure.

In all four experiment cognitive load was measured using a 5-point subjective rating scale (1 = very low; 5 = very high) on which students had to indicate their invested mental effort after the completion of each learning task.

The combination of the average performance score and the invested mental effort score of the last executed task was used to determine the complexity level of the next to be presented task. The score was found by filling in the performance and the mental effort scores in the efficiency formula:

$$\frac{\text{Performance} - \text{Mental Effort}}{\sqrt{2}}$$

When the efficiency score was smaller than zero, task complexity was decreased; and if the efficiency score was larger than zero, task complexity was increased.

## Experiment 1

The first study (Salden, Paas, Broers, & van Merriënboer, 2004) compared the differential effects of four task selection methods on training efficiency (e.g., training time and number of tasks needed to reach the exit performance level) and transfer test performance in a computer-based Air Traffic Control (ATC) training program. A non-dynamic condition, in which the learning tasks are presented to the participants in a fixed, predetermined sequence, was compared to three dynamic conditions. The dynamic conditions selected learning tasks on the basis of performance, mental effort, or mental efficiency (i.e., a combination of performance and mental effort). The participants were first given an introduction to the ATC field and had to complete a practice task before they could continue with the actual training program. All participants start with a task of the lowest complexity level and then continued with learning tasks that are selected according to the condition they worked in. After the training was completed, they were presented with ten transfer tasks.

## Results Experiment 1

Since several factors were fixed in the control condition they were excluded from three analyses: highest complexity level reached in training phase, absolute jump size between complexity levels, and total number of training tasks.

With regard to the *highest complexity level* ( $F(2,65) = 20.5$ ,  $MSE = 3.31$ ,  $p < .0001$ ,  $\eta^2 = .39$ ) the dynamic conditions differed significantly with the mental efficiency condition reaching a higher complexity level than both performance and mental effort conditions ( $t(65) = 2.72$ ,  $p < .01$ ). Furthermore, following a main effect ( $F(2,65) = 28.6$ ,  $MSE = .01$ ,  $p < .0001$ ,  $\eta^2 = .47$ ) the mental efficiency condition attained a larger *jump size* than the other two dynamic conditions ( $t(65) = 3.43$ ,  $p < .01$ ). Lastly, the efficiency condition did not execute less or more *training tasks* than the performance and mental effort conditions ( $t(65) = -.65$ ,  $p = .52$ ).

Analyses ( $F(3,87) = 42.6$ ,  $MSE = 225376.6$ ,  $p < .0001$ ,  $\eta^2 = .60$ ) between all four conditions revealed that the fixed condition needed more time to complete the training ( $t(87) = 7.92$ ,  $p < .0001$ ) than the three dynamic conditions. While no difference in performance was found between the fixed condition and the three dynamic conditions, analysis on mental effort ( $F(3,87) = 8.3$ ,  $MSE = .17$ ,  $p < .0001$ ,  $\eta^2 = .22$ ) showed that the fixed condition did invest more mental effort ( $t(87) = 3.48$ ,  $p < .001$ ) during training than the dynamic conditions.

Though no significant effects were found in performance or mental effort on the transfer test, an analysis on the training efficiency ( $F(3,87) = 7.3$ ,  $MSE = 1.21$ ,  $p < .0001$ ,  $\eta^2 = .20$ ) revealed that the fixed condition was less efficient ( $t(87) = -4.46$ ,  $p < .0001$ ) than the three dynamic conditions. There were no differences between the dynamic conditions regarding training efficiency.

## Experiment 2

In the second Air Traffic Control study (Salden et al., in press) two personalized methods were contrasted to yoked control conditions. In one personalized condition, task selection was based on a combination of performance and invested mental effort (i.e., mental efficiency); in the other personalized condition, the learner was free to select the complexity level of the next learning task (i.e., learner control). Furthermore, participants in both personalized conditions were matched to "yoked" participants in two control conditions. That is, each individualized training sequence of a participant in the mental efficiency condition or the learner control condition was also presented to a participant in the corresponding yoked control condition. Note that the yoked participant was presented with the training sequence of someone else, hence no personalization occurred in the yoked conditions. After an introduction to the ATC field, all participants were given a short pre-training before they started with the actual training program. After completion of the training all participants are presented with a two-fold transfer test consisting of ten transfer test tasks and a reaction time test.

## Results Experiment 2

Training effects were found between the two personalized conditions on highest complexity level attained ( $F(3,56) = 3.04$ ,  $MSE = 2.07$ ,  $p < 0.5$ ,  $\eta^2 = .14$ ) and highest jump size between tasks ( $F(3,56) = 5.27$ ,  $MSE = 0.03$ ,  $p < 0.1$ ,  $\eta^2 = .22$ ). The mental efficiency condition reached a higher complexity level ( $t(56) = 2.19$ ,  $p < .05$ ) and made larger jumps between complexity levels ( $t(56) = 2.69$ ,  $p < .05$ ) than the learner control condition.

While no effects were found for invested mental effort during training, a strong trend was found for training performance ( $F(3, 56) = 2.74$ ,  $MSE = 144.01$ ,  $p = .05$ ,  $\eta^2 = .13$ ). Both personalized conditions obtained a higher performance than their corresponding yoked conditions ( $t(56) = 2.25$ ,  $p < .05$ ). Furthermore, the mental efficiency

condition attained a higher performance score ( $t(56) = 2.44$ ,  $p < .05$ ) than the learner control condition.

No effects were found on the transfer tasks in terms of performance or mental efforts. However, the reaction time test revealed a main effect on conflict identification ( $F(3,56) = 8.18$ ,  $MSE = 28.18$ ,  $p < .0001$ ,  $\eta^2 = .31$ ). The personalized conditions (mental efficiency and learner control) made more correct conflict identifications ( $t(56) = 2.04$ ,  $p < .05$ ) than the yoked conditions. Furthermore, the learner control condition outperformed the mental efficiency condition by making more correct conflict identifications ( $t(56) = -3.58$ ,  $p < .01$ ). Lastly, while an analysis on training efficiency did not expose differences between the personalized conditions and their yoked conditions, it did reveal that the mental efficiency condition was less efficient ( $t(56) = -3.00$ ,  $p < .01$ ) than the learner control condition.

### Experiment 3

The third study (Salden, Paas, van der Pal, & van Merriënboer, 2006) examined the effects of three task selection methods on training efficiency and test performance in a computer-based training program for programming a Flight Management System (FMS). A non-dynamic condition, in which the learning tasks were presented to the participants in a fixed, predetermined sequence, was compared to two dynamic conditions. In the dynamic conditions, the learning tasks were either selected by the participants themselves (i.e., learner control) or by a task selection algorithm in the computer-based training program that used the participant's self-ratings for performance and mental effort. The participants in the learner control condition had total freedom in selecting the learning task they wanted to practice next. All participants were presented with five test tasks after completion of the training.

### Results Experiment 3

Because the number of learning tasks was preset in the fixed condition, one-sample t-tests were used to compare this number of tasks to those of the learner control and mental efficiency conditions. Both these dynamic conditions needed substantially less tasks ( $t(20) = -4.6$ ,  $p < .001$ ) than the fixed condition to complete the training. Furthermore, both dynamic conditions made larger jumps in complexity levels ( $t(20) = 4.3$ ,  $p < .001$ ) in the fixed condition. Following a main effect for training time ( $F(2,28) = 28.37$ ,  $MSE = 444.40$ ,  $p < .001$ ,  $\eta^2 = .67$ ) it was shown that the fixed condition needed more time ( $t(28) = 6.37$ ,  $p < .001$ ) to complete the training than both dynamic conditions. Lastly, the learner control condition needed less training time than the mental efficiency condition ( $t(28) = -4.20$ ,  $p < .001$ ).

With regard to training performance ( $F(2,28) = 15.00$ ,  $MSE = .08$ ,  $p < .001$ ,  $\eta^2 = .52$ ), the fixed condition obtained a higher score ( $t(28) = 5.47$ ,  $p < .001$ ) than both dynamic conditions. Furthermore, no differences were found on the invested mental effort during training.

Controlling for the number of learning tasks and total training time in the analyses for test performance, mental effort on test and training efficiency no effects were found.

### Experiment 4

Since the data from study 3 suggest that some participants systematically overrated their performance, the role of self-ratings was further investigated in a second FMS study (Salden, et al., 2006). More specifically, the fourth study investigated whether the higher amount of training time and the larger number of training tasks in the non-personalized condition confounded the results of the third study.

The non-dynamic fixed condition was again compared to a mental efficiency condition in which students assessed their own performance and mental effort. As in the study 3, five test tasks were given after the participants had completed the training.

### Results Experiment 4

Because the number of learning tasks was preset in the fixed condition, one-sample t-tests were used to compare the number of tasks with the mental efficiency condition. It was shown that the mental efficiency condition ( $t(19) = -2.9$ ,  $p < .01$ ,  $r = -.65$ ) needed less training tasks than the fixed condition to complete the training. Furthermore, the mental efficiency condition made larger jumps in complexity levels ( $t(19) = 3.6$ ,  $p < .01$ ,  $r = .14$ ) than the fixed condition. No effects were found for training performance and mental effort invested during training.

Controlling for the number of learning tasks no effects were found for mental effort on test, test performance, and training efficiency.

**Results Experiments 3 and 4 Combined** Experimenter's observations of the participants in the mental efficiency conditions of both experiments suggested that the absence of clear beneficial effects for this condition might have been caused by the poor quality of self-ratings of performance (e.g., Tousignant & DesMarchais, 2002). In particular, it seemed that some of the participants overrated their performance as compared to their objective performance scores. To test this alternative hypothesis, a K-means cluster analysis on the differences between objective and subjective performance scores identified three groups of self-raters: Good self-raters, average self-raters, and bad self-raters. The extreme groups (i.e., good and bad self-raters) were compared to the combined fixed conditions of both studies on the test variables mental effort and performance. Kruskal-Wallis tests revealed that the participants in the fixed condition attained a higher test performance than the bad self-raters ( $\chi^2 = 7.21$ ,  $p < .01$ ;  $M = 2.89$ ,  $SD = .24$ ,  $r = .10$ ). However, no difference was found between the fixed condition and the good self-raters ( $\chi^2 < 1$ ;  $M = 3.27$ ,  $SD = .22$ ). In addition, the good self-raters attained a higher test performance ( $\chi^2 = 5.04$ ,  $p < .05$ ,  $r = .19$ ) than the bad self-raters. These results revealed an important difference between good and bad self-raters, which might have

confounded possible beneficial effects of the efficiency method.

## General discussion

The results of the four studies lead to the following conclusions. First of all, personalized instruction can have beneficial effects for the training of complex cognitive skills. Although the mental efficiency method did not lead to superior test results, it showed training benefits in every study. Furthermore, students are capable to use learner control of learning task selection effectively as shown in Experiment 2, where the students who trained with learner control exhibited superior performance on a reaction time test. Whereas students seem able to deal with the given control, Studies 3 and 4 indicate that self-ratings should be used with caution. Because these students were novice learners with the FMS, it is conceivable that the novelty of the task at hand disabled their ability to judge their own performance. Of all students in these two studies, only 33% of the students were able to estimate their performance accurately.

## Limitations to Research

While personalized instruction can be beneficial, the four experiments also point out what might have limited possible effects of the training methods.

First of all, when taking the results of all studies into account, questions arise why the overall performance during both training and test phases seems higher than in comparative studies. Although additional analyses revealed no ceiling or floor effects, it might be that the range of complexity used in our studies was too limited.

This might imply that the overall complexity of the materials used could have been too low and suggest that possibly larger differences in performance and mental effort could have been found with more complex materials. Overall, the participants attained a slightly lower test performance than training performance, but the relatively high test performance scores suggest that they might have been able to execute even more complex tasks.

A further aspect that might have attributed to the limited effects of the training methods might be found in the efficiency method. While it was originally developed to estimate the efficiency of experimental conditions the current studies used it as a determinant for dynamic task selection. To use it for this purpose, the relation between performance and mental effort (i.e., efficiency) is estimated for each learning task based on the performance and mental effort scores of the last executed task. The optimization of the learning process might have been limited due to the fact that the efficiency method does not take the history of previous learning tasks and associated performance and mental effort scores into account.

Lastly, since the efficiency formula takes only performance and mental effort into account, it is insensitive to other important factors like motivation. However, an indication of the learner's motivation might be found in the

relationship between performance and mental effort. While a student who attains a low performance score but yet invests a high amount of mental effort is seen as low efficient according to the efficiency formula, the invested mental effort might also indicate that the student is highly motivated. In overview of the studies, the moderate levels of invested mental effort during training could indicate that motivation might decrease when trainees feel that they are not really challenged anymore.

## Implications

Automation should be used carefully in training programs, especially for novice learners who are easily overloaded with the complexity of an extensive work environment of an Air Traffic Controller or a pilot. It might be good for them to start with a simplified and less automated training environment and after having acquired the basic skills, to advance to more complex training programs.

In contrast to previous research, the studies have shown that students seem to be able to use learner control efficiently. Students who are given control over the learning tasks and their respective complexity level are able to create an effective training sequence. As long as the level of given control does not overload the students, they can shape their own training sequence. Further exploration of the level of given learner control, and of how to adapt the amount of control to the growing expertise of the learners during training, represents a promising line of future research.

While students are able to select an appropriate learning task in terms of complexity, the FMS studies show that the capacity of estimating the quality of one's own performance is lacking in most students. Since the students were novice learners, it is conceivable that the novelty of the task at hand disabled their ability to judge their own performance. While 66% of all students overestimated their performance, only 33% of the students were able to estimate their performance accurately.

For future research it would be interesting to investigate to what extent more advanced students are able to use self assessment. The 'expertise reversal effect' (e.g., Kalyuga, et al., 2003) shows that instructional materials should be adjusted to the level of learner expertise. The elaborated instructional materials that are helpful at the start of a training program might become redundant when the student has attained a higher level of expertise. Not only might such more advanced students be able to deal with higher levels of learner control but they might also be capable to use self assessment more accurately than the novice learners in our studies.

Also, in combination with self assessment, the use of peer assessment in novice students might lead to interesting effects. Research has shown that peer assessment positively influences the students' view on learning and assessment, improves learning satisfaction, and enhances clarity of the learning criteria (e.g., Sluijsmans, Moerkerke, Dochy, & van Merriënboer, 2001). Furthermore, by learning to assess their peers, the students reflect more on their own performance

(e.g., Sobral, 1997) and the awareness of the quality of their own performance improves (e.g., Anderson & Freiberg, 1995). More advanced students who have used peer assessment in early training phases might also be more capable in rating their own performance in a later training phase.

### Final Remarks

The current studies can be seen as a first attempt to investigate the possibilities, benefits, and limitations of personalized training methods that are based on an extensive instructional design model such as the 4C/ID model. To use an extensive instructional design model as the basis for training development and to adapt the actual training to the needs of the individual learner is something that has started only recently. Also, the additional use of the concept of cognitive load in the process of dynamic task selection is not to be found in many studies.

Though the studies have not delivered indisputable support for the claim that personalized training methods are more effective, they have shown that personalized instruction can have beneficial effects for the individual learner. While some questions are left unanswered and new ones have arisen, the studies give various leads and clues on how to proceed with the investigation of personalized training methods.

### References

Anderson, J. B., & Freiberg, H. J. (1995). Using self assessment as a reflective tool to enhance the student teaching experience. *Teacher Education Quarterly*, 22, 77-91.

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science*, 205, 1289-1292.

Bell, B. S., & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology*, 55, 267-306.

Borsook, T. K., & Higginbotham-Wheat, N. (1991). Interactivity: what is it and what can it do for computer-based instruction? *Educational Technology*, 4, 11-17.

Camp, G., Paas, F., Rikers, R. M. J. P., & van Merriënboer, J. J. G. (2001). Dynamic problem selection in air traffic control training: a comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior*, 17, 575-595.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23-31.

Kashihara, A., Hirashima, T., & Toyoda, J. (1995). A cognitive load application in tutoring. *Journal of User Modeling and User-Adapted Instruction*, 4, 279-303.

Niemic, R. P., Sikorski, C., & Walberg, H. J. (1996). Learner-control effects: A review of reviews and a meta-analysis. *Journal of Educational Computing Research*, 15, 157-174.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38, 1-4.

Paas, F., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors*, 35, 737-743.

Salden, R. J. C. M., Paas, F., Broers, N. J., & van Merriënboer, J. J. G. (2004). Mental effort and performance as determinants for the dynamic selection of learning tasks in Air Traffic Control training. *Instructional Science*, 32, 153-172.

Salden, R. J. C. M., Paas, F., & van Merriënboer, J. J. G. (2006). A comparison of approaches to learning task selection in the training of complex cognitive skills. *Computers in Human Behavior*, 22, 321-333.

Salden, R. J. C. M., Paas, F., & van Merriënboer, J. J. G. (in press). Personalised adaptive task selection in Air Traffic Control: Effects on training efficiency and transfer. *Learning & Instruction*.

Salden, R. J. C. M., Paas, F., van der Pal, J., & van Merriënboer, J. J. G. (2006). Dynamic task selection in Flight Management System training. *International Journal of Aviation Psychology*, 16, 157-174.

Sluijsmans D. M. A., Moerkerke, G., Dochy, F., & van Merriënboer, J. J. G. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation*, 27, 153-173.

Sobral, D. T. (1997). Improving learning skills: A self-help group approach. *Higher Education*, 33, 39-50.

Steinberg, E. R. (1989). Cognition and learner control: A literature review, 1977-1988. *Journal of Computer-Based Instruction*, 16, 117-121.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 81, 457-466.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.

Tousignant, M., & DesMarchais, J. E. (2002). Accuracy of student self-assessment ability compared to their own performance in a problem-based learning medical program: A correlation study. *Advances in Health Sciences Education*, 7, 19-27.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four component instructional design model*. Englewood Cliffs, NJ: Educational Technology Publications.

Van Merriënboer, J. J. G., Schuurman, J. G., de Croock, M. B. M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance, and training efficiency. *Learning and Instruction*, 12, 11-37.