

Running head: DYNAMIC TASK SELECTION IN A FLIGHT MANAGEMENT SYSTEM
TRAINING

This is a pre-print of the article that was published as:

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.

Dynamic Task Selection in Flight Management System Training

Ron J. C. M. Salden, Fred Paas

Open University of the Netherlands

Jelke van der Pal

National Aerospace Laboratory NLR, Netherlands

Jeroen J. G. van Merriënboer

Open University of the Netherlands

Correspondence concerning this article should be addressed to Ron J.C.M. Salden, Open University of the Netherlands, Educational Technology Expertise Center (OTEC), P.O. Box 2960, NL-6401 DL Heerlen, The Netherlands. E-mail may be sent to ron.salden@ou.nl
phone: 0031-45-5762398, fax: 0031-45-5762907

Abstract

The effects of three task selection methods on test performance and training efficiency in a computer-based Flight Management System (FMS) training were investigated. A fixed condition was compared to a learner control condition, and a condition using the participants' self-rated performance and mental effort. Although, the experimental conditions revealed more positive training effects, no differences were found for training efficiency and test performance. A follow-up study did not confirm the alternative hypothesis that these results were caused by the higher amount of tasks in the fixed condition. Extra analyses suggested that the quality of self-rating needs to be considered in future research.

Keywords: dynamic task selection, personalized training, mental efficiency, learner control

Dynamic Task Selection in a Flight Management System Training

In the aviation domain, continuous efforts are directed at increasing the cockpit automation. An automated cockpit has a significant effect on the pilot's tasks and demands additional competencies from the cockpit crew. The Flight Management System (FMS) is one of the core systems in an automated cockpit, which can control an entire flight from take off to landing. Since the mid-1990's FMS training programs have started to use realistic computer-based FMS simulations (e.g., Casner, 1995; Lyall, Vint, Niemczyk, Wilson, & Funk, 1998). Computer-based training may prepare and enhance the pilot's automation related skills and make time spent on expensive part-task trainers and full flight simulators more effective. Whereas, it may allow experienced pilots to practice on new FMS systems in a free-play fashion, novice pilots can be given more support, for instance, by adapting the complexity of learning tasks to their experience level. It can be expected that such personalized training can make FMS training more efficient.

An instructional design model for training complex cognitive skills was developed by van Merriënboer (1997). This four component instructional design model (4C/ID-model) provides methods and techniques for analyzing a complex cognitive skill into its constituent skills and their relationships. Furthermore, it supports analyses of the different knowledge structures that may be helpful or are required to be able to perform the constituent skills. Lastly, it also provides methods and techniques for designing a training blueprint with a sequence of whole task practice situations that support integration and coordination of the constituent skills. The 4C/ID-model has recently been applied within the aviation domain in a European-Community funded project called Advanced Design Approach for Personalized Training interactive tools (ADAPT^{IT}). In this project a software tool supporting the 4C/ID-methodology was developed and used to design new 4C/ID-based training programs for the Swedish Air Traffic Controller Training Academy (SATSA) and for aircraft

maintenance training in Piaggio Aero Industries (de Croock, Paas, Schlanbusch, & van Merriënboer, 2002).

The current study examines several training methods that were designed according to the 4C/ID-model (van Merriënboer, 1997; van Merriënboer, Clark, & de Croock, 2002). This model offers a training design that presents students with a predetermined order and complexity of learning tasks in such a way that their cognitive capacity is optimally used. Further efficiency may be reached by providing a personalized and adaptive training trajectory, in which learning tasks are selected during training based upon the performance and needs of the individual learner. Especially for training complex cognitive skills, the use of such a ‘dynamic’ ability to optimize task order and complexity for the individual trainee is believed to be strongly related to increased training efficiency (Salden, Paas, & van Merriënboer, in press; Salden, Paas, Broers, & van Merriënboer, 2004).

Research in the context of cognitive load theory (for an overview see Paas, Renkl, & Sweller, 2003, 2004) has shown that cognitive load is a crucial factor in the training of complex cognitive skills. The combination of cognitive load and performance measures is considered a superior estimate of a learner’s cognitive demands that can be used in the dynamic selection of learning tasks. For example, when two trainees achieve the same performance scores, measures of cognitive load might be able to reveal differences in training efficiency otherwise unnoticed. While one of them might have experienced a very high cognitive load and needed to work laboriously through a very effortful process, the other person might have experienced a low cognitive load and reached the same performance level with a minimum of effort. In a personalized task selection method this information could be used to present a less difficult next task to the first person than to the second person. Because the combination of performance and cognitive load measures provides a clear picture of the state of the student’s cognitive system at a certain moment in training, dynamically selected learning tasks fit well to the cognitive schemata a student has acquired. The individual capacity of a student is

taken into account, hence leading to high training efficiency. To obtain a good indication of the cognitive load that is imposed on a person's cognitive system, mental effort measurements may be used (Paas & van Merriënboer, 1993).

Paas and van Merriënboer (1993; see also Paas, Tuovinen, Tabbers, & van Gerven, 2003) 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 two studies of Camp, Paas, Rikers, and van Merriënboer (2001), and Salden et al. (2004). Both studies compared four methods of task selection in the Air Traffic Control (ATC) domain. In the first method, tasks were presented in a fixed, predetermined simple-to-complex sequence designed according to the 4C/ID-model. 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 automated task selection, learner control may offer another form of dynamic task selection because it gives the students control over what learning tasks they want to practice next. The theoretical claim for learner control is that trainees are capable of selecting appropriate tasks to practice and can avoid a possible overload of their cognitive system, thereby increasing the effectiveness and efficiency of learning (e.g., Borsook & Higginbotham-Wheat, 1991; Niemic, Sikorski, & Walberg, 1996; Steinberg, 1977, 1989; Williams, 1993). However, according to Bell and Kozlowski (2002), positive effects of learner control on learning can only be expected if instructional materials are designed in such a way that they provide learners with a level of control they are able to handle. Support for this claim was found in recent studies (van Merriënboer, Schuurman, de Croock, & Paas, 2002; Salden et al., 2004), which showed that learners who are given an appropriate level of control over task selection are well able to select their own learning tasks.

In the current study, participants with some flying background but no FMS knowledge are divided into three conditions of FMS training. The fixed condition is a control condition in which learning tasks are presented in a predetermined order based on increasing complexity of learning tasks. This condition will be compared with two personalized experimental conditions, in which learners either have to select the learning task themselves, i.e. the learner control condition, or in which the learning tasks are selected by the training program using a combination of the learners' self-rated performance and mental effort, i.e. the mental efficiency condition.

In agreement with Camp et al. (2001) and Salden et al. (2004), it is hypothesized that personalized task selection leads to more efficient training and better test performance than non-dynamic task selection. The differences between both dynamic conditions, which personalize the task selection either by using the mental efficiency or by allowing the learners to determine their own training sequence, are explored.

Experiment 1

Method

Participants

Thirty-one students of a higher education school for aviation (three women and twenty-eight men, $M = 20.1$ years, $SD = 2.69$) participated in this study. All students were in the same college year, had a similar amount of flight time in flight simulators and were novices with regard to the FMS. The students were randomly assigned to three experimental conditions: A fixed condition ($n = 10$), a learner control condition ($n = 10$), and a mental efficiency condition ($n = 11$). All participants were in good health and had normal or corrected-to-normal vision. They volunteered to participate in this study and were not paid for their participation.

Materials

FMS simulation. The training software was based on a realistic computer simulation of a Boeing 747 FMS developed by the National Aerospace Laboratory NLR. The training software ran on an IBM-compatible PC (Pentium III, 533 MHz) using an IBM 17-inch SVGA monitor. Figure 1 depicts the interface of the FMS program.

INSERT FIGURE 1 ABOUT HERE

Learning tasks. In the training, the participants were confronted with learning tasks, which presented flight information of a certain route from airport A to airport B that learners had to program into the FMS simulation. A simulated flight had to be executed after entering all information. At certain points during the task, changes in the flight route were required and made it necessary for the trainees to adjust the original flight route. Possible changes consisted of an

alteration in arrival data (e.g., a new Standard Terminal Arrival Route), a new runway, or a diversion to another airport.

Prior to training, the thirty-two learning tasks were categorized into eight levels (four tasks per level) that specified the complexity of the tasks based on three complexity factors: The amount of data to be programmed into the FMS program, the number of changes in flight route, and the amount of time pressure. Values on these factors were added to determine the overall complexity of a learning task. Similar factors have proven to be good indicators of task complexity in several experiments with a Air Traffic Control training (Camp et al., 2001; Salden, Paas, & van Merriënboer, in press; Salden, et al., 2004). Please note that due to the novice level of the participants the main focus was on learning to handle the basic FMS functions. The participants were not presented with more in depth aspects such as lateral and vertical navigation in relation to new waypoints. In case of missing waypoints, the task information presented the participants with the exact lateral and vertical specifications of a waypoint which had to be programmed into the FMS.

Task selection. The selection of tasks differed between the three experimental conditions. In the fixed condition, participants received a total of 16 learning tasks with two randomly chosen tasks of each of the eight complexity levels (1, 1, 2, 2, 3, 3, etc.). These 16 learning tasks were presented in a predetermined order from low to high complexity. This sequencing of gradually increasing complexity is a main element in the 4C/ID model. Previous studies (Camp et al, 2001; Salden et al., 2004) have shown that presenting two tasks of each complexity level ensures sufficient variation of all task elements. In the learner control condition, participants received an overview of all learning tasks with an indication of their complexity and could choose which task to practice next. Thus, these learners had maximum freedom to determine their own training sequence.

INSERT TABLE 1 ABOUT HERE

The use of subjective rating scales has been repeatedly proven to be a reliable measure of cognitive load (for an overview see Paas, Tuovinen, Tabbers, & van Gerven, 2003). Furthermore, in a recent study various subjective mental workload methods including the NASA Task Load Index, the Subjective Workload Assessment Technique, and the Workload Profile were compared to each other and the overall conclusion was that all methods showed high validity (Rubio, Diaz, Martin, & Puente, 2004). For our purpose of task selection, a method was required which could be processed fast enough to avoid delays in selection of new learning tasks. Therefore we choose a subjective rating scale to measure mental effort.

In the mental efficiency condition, task selection was based on participants' self-ratings of performance and mental effort on a 5-point rating scale. These subjective performance and mental effort scores were used to determine mental efficiency (Paas & van Merriënboer, 1993; Paas et al., 2003). If the subjective performance score (see Y axis of Figure 2) was higher than the mental effort score (X axis) this was interpreted as a high mental efficiency; the learner performed relatively well and invested less mental effort than could be expected on the basis of his or her subjective performance. If the subjective performance score was lower than the mental effort score this was interpreted as a low mental efficiency; the learner performed relatively low and invested more mental effort than could be expected on the basis of his or her subjective performance. As indicated in Table 1, the complexity of the next learning task was determined on the basis of this argumentation. The subjective performance and mental effort scores in this Table correspond with the self-ratings of the individual student. For instance, if a participant had a mental effort score of 2 and a subjective performance score of 5, task complexity was increased with three levels (+3). Note that participants in the mental efficiency condition had a maximum possible step size of 4 complexity levels, while participants in the learner control condition had a maximum possible step size of 7 complexity levels

(i.e., directly from the lowest to the highest complexity level or vice versa). The maximum step size of 4 complexity levels in the mental efficiency was based on previous studies (Camp et al., 2001; Salden et al., 2004) which showed this to be an effective maximum step size. The maximum size of 7 complexity levels in the learner control condition was a consequence of allowing the students to form their own training sequence, even if this led to ineffective decisions.

INSERT FIGURE 2 ABOUT HERE

Three possible outcomes could lead to the completion of the training. First, participants finished practice after working on 16 learning tasks. This number of tasks ensures sufficient variation over the eight complexity levels. Second, participants finished practice after their self-ratings equaled the preset performance (≥ 3) and mental effort (≤ 3) criteria for two successive learning tasks at the highest complexity level. Third, participants finished practice after they performed all four available learning tasks at the highest complexity level. Note that in the two latter cases a participant could complete the training after working on less than 16 learning tasks.

Test tasks. After the training, the participants had to perform a test that consisted of five test tasks, which were different from the learning tasks in two ways: (1) the amount of data that had to be programmed into the FMS simulation was increased, and (2) the number of changes in the flight route was higher.

Objective scoring of performance. To be able to compare the training and test performance between the experimental conditions, the objective performance of all participants was scored after completion of the experiment. For each flight, scores on a 5-point scale were given on four performance variables: (a) the given number of commands; (b) the number of correct commands; (c) the time on task, and (d) the time needed to process a change in flight route data. For all variables, a

score of 1 indicated a very low performance and a score of 5 indicated a very high performance. The mean score of these four variables was used to compare performance between the experimental conditions.

Procedure

All participants were given a paper-based 10-page introduction to the training and the use of the FMS simulation which presented all the information required for the training, including examples of how to enter commands into the FMS. Participants were free to consult this introduction during the entire training session. After they had read the introduction, the training started and a learning task at the first complexity level was presented. Subsequent tasks then depended on the experimental condition. In the fixed condition, participants received another task at complexity level 1, then two tasks at level 2, and so forth; in the learner control condition, the participants could choose whatever task they preferred next; and in the mental efficiency condition, the next task was based on their subjective performance and mental effort self-ratings (see Table 1 for the applied step size). All participants could continue with the next learning task as soon as they had completed the previous task, meaning that differences in training time could occur in all conditions. The participants performed the five test tasks immediately after they completed the learning tasks. The whole experiment took about three hours.

Results and Discussion

First, the results on the training phase are reported. The mean number of learning tasks, step sizes, and total training time are given for each condition to provide insight in the task selection process. Furthermore, the results for training performance and mental effort are given. Second, the results for performance, mental effort, and training efficiency are provided for the test phase. One-sample and independent *t*-tests, ANOVAs, ANCOVAs and planned comparisons were used to analyze the data. Means and standard deviations are presented in Table 2.

INSERT TABLE 2 ABOUT HERE

Training Phase

Training effects. Because the number of learning tasks was preset at 16 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 the learner control condition ($M = 6.50$, $SD = 1.35$; $t(19) = -4.3$, $p < .001$) and the mental efficiency condition ($M = 7.27$, $SD = 1.19$; $t(20) = -4.6$, $p < .001$) needed substantially less than the 16 learning tasks in the fixed condition to complete the training. The comparison between the learner control and mental efficiency condition showed no difference in number of learning tasks ($t = 1.4$).

The absolute step size in complexity level was also preset for the fixed condition, at one level in complexity per two tasks. In total the participants made 15 steps between 16 tasks, with 7 steps to a higher complexity level ($M = .47$ (7/15)). For the learner control and mental efficiency condition, steps between complexity levels could be negative or positive, corresponding to easier or more difficult tasks, respectively. One-sample *t*-tests on the absolute step size showed that participants in both the learner control condition ($M = .95$, $SD = .07$; $t(19) = 4.1$, $p < .01$) and the mental efficiency condition ($M = .93$, $SD = .09$; $t(20) = 4.3$, $p < .001$) made larger absolute steps between complexity levels than participants in the fixed condition. The step sizes between the learner control and mental efficiency condition did not differ ($t < 1$).

A significant effect was found for total training time, $F(2, 28) = 28.37$, $MSE = 444.40$, $p < .001$, $\eta^2 = .67$. Planned comparisons showed that the participants in the fixed condition ($M = 149.60$, $SD = 22.77$) needed more time to complete the training ($t(28) = 6.37$, $p < .001$) than the participants in the learner control and mental efficiency conditions ($M = 98.02$, $SD = 18.63$). Furthermore, the

participants in the learner control condition needed less training time than those in the mental efficiency condition ($t(28) = -4.20, p < .001$).

Performance and mental effort. A significant effect was found for training performance ($F(2, 28) = 15.00, MSE = .08, p < .001, \eta^2 = .52$). The performance score of the fixed condition ($M = 3.27, SD = .28$) was higher ($t(28) = 5.47, p < .001$) than the mean performance score of the learner control and mental efficiency conditions ($M = 2.67, SD = .28$). The comparison between the learner control and mental efficiency condition showed no difference in performance ($t < 1$). No significant effects were found on the mental effort during training ($F < 1$).

Test Phase

During the training, participants in the fixed condition worked on much more tasks, made smaller steps between complexity levels, and needed more training time than participants in the learner control and mental efficiency conditions. The number of tasks and the total training time could easily explain possible differences between conditions on the test tasks. Therefore, number of learning tasks and total training time are included as covariates in the subsequent analyses.

Performance and mental effort. Using an ANCOVA with number of learning tasks and total training time as covariates, no effects were found on performance and mental effort ($F_s < 1$).

Training efficiency. The training efficiency was determined using the following formula (Paas & van Merriënboer, 1993; Paas et al., 2003):

$$E = \frac{P - ME}{\sqrt{2}}$$

In this formula, E = mental efficiency, P = test performance, and ME = mental effort during training.

Using an ANCOVA with number of learning tasks and total training time as covariates, no effect found for Method of Task Selection, ($F < 1$). The estimated marginal means and standard deviations are provided in Table 2.

In conclusion, participants in the learner control and mental efficiency conditions worked on less learning tasks, made larger steps between complexity levels and needed less time to complete the training than the participants in the fixed condition. Performance during training was higher for the fixed condition, which can easily be explained by the prolonged training time. No differences were found on test performance, mental effort during the test phase, and training efficiency.

To control for the high number of learning tasks in the fixed condition (16) a second experiment was conducted, comparing a mental efficiency condition to a fixed condition with only 8 learning tasks. It was expected that this would limit the difference between conditions during the training phase, and possibly show the expected positive effect of dynamic task selection on test performance.

Experiment 2

Method

Participants

Twenty students of the same higher education school for aviation as in the first experiment (6 women and 14 men, $M = 23.8$ years, $SD = 4.12$) participated in this study. All students were in the same college year, had a similar amount of flight time in flight simulators and were novices with regard to the FMS. The students were randomly assigned to the two experimental conditions: A fixed condition ($n = 10$) and a mental efficiency condition ($n = 10$). All participants were in good health and had normal or corrected-to-normal vision. They received € 15 (approximately \$ 18) for their participation.

Materials and Procedure

The materials were the same as in Experiment 1. The only difference is that the number of learning tasks in the fixed condition was downsized from 16 to 8, resulting in only one task per

complexity level. The procedure was identical to the procedure in Experiment 1 and participants received the same test tasks.

Results and Discussion

First, the results on the training phase are reported. The mean number of learning tasks, step sizes, and total training time are given for each condition to provide insight in the task selection process. Furthermore, the results for training performance and mental effort are given. Second, the results for performance, mental effort, and training efficiency are provided for the test phase. One sample *t*-tests, ANOVAs, and ANCOVAs were used to analyze the data. Means and standard deviations are presented in Table 3.

INSERT TABLE 3 ABOUT HERE

Training Phase

Training effects. Because the number of learning tasks was preset at 8 in the fixed condition, one-sample *t*-tests were used to compare the number of tasks with the mental efficiency condition. The *t*-test showed that the mental efficiency condition ($M = 6.28$, $SD = 1.48$; $t(19) = -2.9$, $p < .01$) needed less than the 8 learning tasks in the fixed condition to complete the training.

The absolute step size in complexity level was also preset for the fixed condition, at one step in complexity per task. In total, the participants made 7 steps between 8 tasks ($M = .88$ (7/8)). For the mental efficiency condition, steps between complexity levels could be negative or positive, corresponding to easier or more difficult tasks, respectively. A one-sample *t*-test showed that the participants in the mental efficiency condition ($M = .96$, $SD = .44$; $t(19) = 3.6$, $p < .01$) made larger steps between complexity levels than those in the fixed condition. With regard to total training time, no effect was found for Method of Task Selection, ($F = 1.29$).

Performance and mental effort. No effects of Method of Task Selection were found on the training variables mental effort ($F < 1$) and performance ($F = 1.19$).

Test Phase

An ANCOVA with number of learning tasks as a covariate showed no effects on mental effort ($F = 1.05$) and performance ($F < 1$). An ANCOVA with number of learning tasks as a covariate showed no effect ($F = 2.67$) on training efficiency. The estimated marginal means and standard deviations are provided in Table 3.

Additional Analyses for Experiment 1 and 2 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 (Bjork, 1999; 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 ($F(2, 18) = 71.6, MSE = .03, p < .001$) was performed on the differences between objective and subjective performance scores. Three groups of self-raters were identified: Good self-raters ($n = 6$), average self-raters ($n = 9$), and bad self-raters ($n = 6$). 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 ($M = 3.41, SD = .46$) attained a higher test performance than the bad self-raters ($\chi^2 = 7.21, p < .01; M = 2.89, SD = .24$). 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 ($\chi^2 < 5.04, p < .05$) than the bad self-raters. No effects were found on the test variable mental effort ($F = 1.8$). The

means of the fixed group and the good self-raters and bad self-raters in the mental efficiency group are depicted in Figure 3.

INSERT FIGURE 3 ABOUT HERE

General Discussion

The main hypothesis of the first experiment that dynamic task selection leads to more efficient training and better test performance than non-dynamic task selection was not confirmed. Although the participants in the learner control and mental efficiency conditions progressed through training more effectively than the participants in the fixed condition, training performance was highest in the fixed condition. This effect was explained by the prolonged training time. Using number of learning tasks and total training time as covariates, no effects were found for test performance, mental effort on the test, and training efficiency. Participants in the learner control condition resembled the participants in the mental efficiency condition to a high extent. The only difference found between these two conditions was that the participants in the learner control condition required less time to complete the training than the mental efficiency condition.

The difference between the conditions in the number of learning tasks in the first experiment was avoided in the second experiment by decreasing the fixed number of learning tasks in the control condition. This reduction of the number of learning tasks in the second experiment was expected to show the positive effect of dynamic task selection on test performance. The participants in the mental efficiency condition still needed less tasks to complete training and made larger jumps than the participants in the fixed condition. No effects in support of the mental efficiency condition were found on mental effort, test performance and training efficiency.

Previous studies (for an overview see Paas, Tuovinen, Tabbers, & van Gerven, 2003) have shown that participants are well capable to rate their own mental effort, but the current study shows that rating one's own performance appears to be more difficult. From observing the participants in the mental efficiency conditions in both experiments, it was hypothesized that the absence of clear beneficial effects for this condition might have been caused by the low quality of the self-ratings. In other words, some of the participants would overrate their own performance compared to their actual objective performance. When comparing the fixed conditions to two subgroups of self-raters, effects were found on test performance. It was shown that the participants in the fixed condition attained a higher performance than the bad self-raters but no difference was found between the participants in the fixed condition and the good self-raters. The good self-raters attained a higher test performance than the bad self-raters. Based on these results, it seems plausible that the low quality of the self-ratings has confounded the results in both experiments. Since most participants were not very skilled at rating their own performance, the personalized training sequence they were presented with was not optimal. The overrating of their performance led them to receive more difficult tasks than they should have been given. Considering that 67% of the participants overrated their own performance it seems advisable to give a self-assessment training before letting student use it in the actual experiment.

In addition to the confounding effects of the self-ratings, there are some possible explanations that can be given for the lack of beneficial effects for personalized task selection.

For instance, it can be observed that the performance scores for all conditions were relatively high. Although additional analyses revealed no ceiling or floor effects, an alternative explanation is that the range of complexity used in this study was too limited. This suggestion is further strengthened by the relatively low levels of invested mental effort in all conditions.

The basic operations on the FMS are very recurrent since one always has to give certain commands in order to be able to execute a flight. The complexity in use can be increased when one has to deal with the FMS in a more authentic cockpit situation. The interaction of the FMS with many displays and control panels is important to achieve situational awareness for the trainee pilot. While the aim of our training was to familiarize the participants with the FMS, the scope of the training in terms of complexity might have been too limited due to the high level of repetitiveness. In accordance with the previous line of reasoning a larger range of complexity in the materials might have resulted in larger differences in performance and mental effort on both training and test phases.

Another explanation could be found in the use of the 4C/ID-model. Like in previous experiments (Camp et al., 2001; Salden et al., 2004) this model was used to create the fixed conditions. The design guidelines in this model recommend a steady increase in complexity during training. While participants in the personalized conditions often complete training faster, few beneficial effects were found on test performance and training efficiency. Whereas previous studies on Air Traffic Control (ATC) training were able to identify at least some beneficial effects, the combination of the 4C/ID-model and the recurrent nature of the FMS skills might deem personalized training methods unnecessary.

In conclusion, it seems necessary and appropriate to conduct a third experiment which addresses the aforementioned shortcomings in order to make a greater contribution to the aviation industry and to research in the area of training. The results of the two studies in this article only moderately supported the idea that adapting training to the individual needs of the student makes training more efficient. While the mental efficiency condition proved to be an effective training method, it did not prove to be efficient as well. In contrast, the fixed condition proved not to be the most effective yet did prove to be a fairly efficient training method. The current combined research

on the efficiency method so far, has shown that future research on the efficiency method is needed to fully grasp its benefits and shortcomings.

References

- 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.
- Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Cognitive regulation of performance: Interaction of theory and application* (pp. 435-459). Cambridge, MA: The MIT Press.
- 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., & 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.
- Casner, S. M. (1995). A personal CBT for the Boeing 737 flight management system. In R.S. Jensen, & L.A. Rakovan (Eds.), *Proceedings of the 8th Symposium on Aviation Psychology*. Columbus, OH: The Ohio State University.
- De Croock, M. B. M., Paas, F., Schlanbusch, H., & van Merriënboer, J. J. G. (2002). ADAPT^{IT}: Tools for training design and evaluation. *Educational Technology, Research & Development*, *50*, 47-58.
- Lyall, B., Vint, R., Niemczyk, M., Wilson, J., & Funk, K. (1998, December). *Training approaches and considerations for automated aircraft: A summary of training development experiences* (AAR-100). Tempe, AZ: Research Integrations, Inc.
- 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*(2), 157-174.

- 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.
- Paas, F., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 51-71.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1-4.
- Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32*, 1-8.
- Paas, F., Tuovinen, J. E., Tabbers, H. K., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71.
- Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profiles methods. *Applied Psychology: an International Review, 53*, 61-86.
- 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. (in press). A comparison of task selection approaches. *Computers in Human Behavior*.
- Salden, R. J. C. M., Paas, F., & van Merriënboer, J. J. G. (2004). *Personalized task selection in Air Traffic Control: Effects on training efficiency and transfer*. Manuscript submitted for publication.

- Steinberg, E. R. (1977). Review of student control in computer-assisted instruction. *Journal of Computer-Based Instruction*, 3, 84-90.
- Steinberg, E. R. (1989). Cognition and learner control: A literature review, 1977-1988. *Journal of Computer-Based Instruction*, 16, 117-121.
- 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., de Croock, M. B. M., & Jelsma, O. (1997). The transfer paradox: Effects of contextual interference on retention and transfer performance of a complex cognitive skill. *Perceptual and Motor Skills*, 84, 784-786.
- Van Merriënboer, J. J. G., Clark, R. E., & de Croock, M. B. M. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology, Research and Development*, 50 (2), 39-64.
- 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.
- Williams, M. D. (1993). A comprehensive review of learner-control: The role of learner characteristics. In M. R. Simonson & A. Dristen (Eds.), *Proceedings of the Annual Conference of the Association for Educational Communications and Technology* (pp. 1083-1114). New Orleans, LA: Association for Educational Communications and Technology.

Applications / So what section

Personalized training is a challenging goal for the coming decade. With more insights from cognitive science and motivation studies on the one hand, and better performance assessment techniques on the other, personalized training may actually be achievable even without having the best live instructors available. The benefits may have a huge impact on training costs, certainly in aviation training, as -overall- less training time can be expected, training may be more motivating for those that do not fit the label of ‘average student’. This study has focused on mental efficiency and learner control during a FMS programming familiarization course. The lesser training time required in these personalized conditions was clearly demonstrated, but no clear difference between the two task selection principles were found. The quality of self-rating may have been a factor, as well as the task complexity itself. For the trainees, who were aviation interested students, but not pilot students, a simplified part-task training was developed which focused on the recurrent skills of FMS programming. Using a more elaborate training in which the FMS task is combined with more navigational tasks, in which both recurrent and non-recurrent skills need to be applied, the effects of mental efficiency are expected to be stronger and the impact on regular aviation training could be extensive.

Personalized FMS training may be particularly interesting to provide in an early stage of the pilot student education. Before the pilot actually engages airline training, basic FMS skills may be acquired and smoothly integrate into the competencies needed for airline aircraft type rating and line oriented flight training.

Table 1

Selection Table Indicating Step Size in Complexity between Learning tasks

Mental effort	Subjective Performance				
	1	2	3	4	5
1	0	1	2	3	4
2	-1	0	1	2	3
3	-2	-1	0	1	2
4	-3	-2	-1	0	1
5	-4	-3	-2	-1	0

Table 2

Overview of Results Experiment 1

Dependent variables	Method of Task Selection					
	Fixed		Learner control		Mental efficiency	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Training phase</i>						
Total <i>N</i> of learning tasks	16	-	6.50	1.35	7.27	1.19
Step size	7/15	-	.95	.07	.93	.09
Training time	149.60	22.77	78.69	11.64	117.35	25.61
Mental effort	2.24	.52	2.13	.56	2.38	.41
Performance	3.27	.28	2.64	.38	2.70	.18
<i>Test phase ^a</i>						
Mental effort	2.41	.93	1.61	.53	2.22	.49
Performance	2.89	.38	3.16	.22	3.21	.20
Training efficiency	-1.09	1.17	.62	.67	.43	.61

^a Estimated marginal means are presented with number of learning tasks and total training time as covariates.

Table 3

Overview of Results Experiment 2

Dependent variables	Method of Task Selection			
	Fixed		Mental efficiency	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Training phase</i>				
Total <i>N</i> of learning tasks	8	-	6.20	1.48
Step size	7/8	-	.96	.44
Training time	79.97	18.32	70.71	18.17
Mental effort	2.40	.46	2.27	.46
Performance	3.26	.24	3.40	.34
<i>Test phase ^a</i>				
Mental effort	1.76	.18	2.08	.18
Performance	3.45	.15	3.14	.15
Training efficiency	.46	.30	-.46	.30

^a Estimated marginal means are presented with number of learning tasks and as covariate.

Figure Captions

Figure 1. Interface of the FMS training program.

Figure 2. Representation of the effect of mental efficiency on the selection of the complexity of the next learning task.

Figure 3. Histogram of test performance of fixed condition, good self-raters in the mental efficiency condition and bad self-raters in the mental efficiency condition.





