

This is a pre-print of:

Taminiau, E. M. C., Kester, L., Corbalan, G., Alessi, S. M., Moxnes, E., Gijsselaers, W., Kirschner, P. A., & Van Merriënboer, J. J. G. (2013). Why Advice on Task Selection May Hamper Learning in On-Demand Education. *Computer in Human Behavior*, 29, 145-154. doi: 10.1016/j.chb.2012.07.28

Copyright Elsevier Ltd; available online at

<http://www.sciencedirect.com/science/article/pii/S0747563212002312>

Why Advice on Task Selection May Hamper Learning in On-Demand Education

Elisabeth M. C. Taminiau<sup>a</sup>, Liesbeth Kester<sup>a</sup>, Gemma Corbalan<sup>b</sup>, Stephen M. Alessi<sup>c</sup>, Erling Moxnes<sup>d</sup>, Wim H. Gijsselaers<sup>e</sup>, Paul A. Kirschner<sup>a</sup>, Jeroen J. G. Van Merriënboer<sup>f</sup>

<sup>a</sup> Centre for Learning Sciences and Technologies, Open University of The Netherlands

<sup>b</sup> The Netherlands Institute for Curriculum Development

<sup>c</sup> College of Education, The University of Iowa,

<sup>d</sup> Department of Geography, University of Bergen

<sup>e</sup> School of Business and Economics, Maastricht University

<sup>f</sup> School of Health Professions Education, Maastricht University

#### Author's note

Correspondence concerning this article should be addressed to Elisabeth M. C. Taminiau, Centre for Learning Sciences and Technologies, Open University of The Netherlands, P.O.

Box 2960, 6401 DL Heerlen, The Netherlands. Tel.: +31 45 5762358, E:

Bettine.Taminiau@ou.nl

### **Abstract**

In on-demand education, learners are required to plan their own learning trajectory by selecting suitable learning tasks. A positive effect on learning is expected when learners select tasks that help them fulfil their individual learning needs. However, the selection of suitable tasks is a difficult process for learners with little domain knowledge and suboptimal task-selection skills. A common solution for helping learners deal with on-demand education and develop domain-specific skills is to give them advice on task selection. In a randomized experiment, learners ( $N = 30$ ) worked on learning tasks in the domain of system dynamics and received either advice or no advice on the selection of new learning tasks. Surprisingly, the no-advice group outperformed the advice group on a post-test measuring domain-specific skills. It is concluded that giving advice on task selection prevents learners from thinking about how the process of task selection works. The advice seems to *supplant* rather than support their considerations why they should perform the advised task, which results in negative effects on learning. Implications for future research on giving advice in on-demand education are discussed.

*Keywords:* Advice, Task selection, On-demand education

## 1. Introduction

On-demand education requires learners to set their own learning trajectory to adapt it to their individual learning needs (Hannafin, 1984; Van Merriënboer et al., 2006). This learning trajectory is set in an iterative cycle of performing a task, assessing the quality of their own performance, and selecting a new task to perform. Theoretically, on-demand education (1) could help learners develop domain-specific skills by following this adapted learning trajectory, and (2) could increase learners' motivation, feelings of responsibility, and self-efficacy (Corbalan, Kester, & Van Merriënboer, 2009; Kinzie & Sullivan, 1989; Schnackenberg & Sullivan, 2000; Topping, 2003). However, in the literature divergent results are found regarding the effects of on-demand education on learning (e.g., Corbalan et al., 2009; Ross, Morrison, & O'Dell, 1989; Shute, Gawlick, & Gluck, 1998; Williams, 1996). Some studies found that such a learner-controlled environment has positive effects on learning while others demonstrate that a system-controlled environment, in which the system selects suitable tasks and not the learners themselves, is more effective. Finally, there are also many studies that find that there is no difference between these two types of environments (for an overview of studies, see Williams, 1996).

Why an on-demand educational environment does not always bring the expected positive results might be explained by the prior knowledge of the learners. Novices in on-demand education may not have developed the task-selection skills and domain knowledge to handle the environment appropriately (Clarebout & Elen, 2009; Corbalan, Van Merriënboer, & Kicken, 2010; Gay, 1986). They might be unfamiliar with the relevant task-selection aspects to consider for selecting a next task (e.g., quality of performance on previous tasks, invested mental effort) and might omit to combine this information with characteristics of available tasks (e.g., difficulty, available support) to interpret what an appropriate next task

would be to select (Van Merriënboer et al., 2006). In addition, low prior knowledge of the domain to-be-learned might hamper learners to understand the organization of the information presented in the environment (Gay, 1986), for example, the difference between the different difficulty levels of tasks and how these different tasks are structured in the environment. The usual large number of tasks in an on-demand educational environment that is available for task selection probably does not make things easier. Learners can easily become overwhelmed by this large number of tasks (Corbalan, Kester, & Van Merriënboer, 2008; Iyengar & Lepper, 2000; Schwartz, 2004). The extensive amount of choice can cause learners to become frustrated and de-motivated. In addition, novices without any prior knowledge might select suboptimal tasks (Schwartz, 2004) that do not fit their learning needs, and, consequently, their learning in on-demand education might be hampered (Corbalan et al., 2008).

### 1.1 The task-selection process

To set a suitable trajectory in an on-demand learning environment, learners must carry out an appropriate task-selection process. This process includes recognizing relevant task-selection aspects and making use of these aspects to select a suitable next task. Relevant task-selection aspects are (1) the level of difficulty and support of the last task performed, (2) the quality of the learner's past performance, and (3) the learner's mental effort (i.e., cognitive resources) invested to perform the last task (Gray, 2003; Kalyuga, 2006; Kostons, Van Gog, & Paas, 2012; Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013). The difficulty and support level of the last task performed can serve as a starting point from which the next task can be selected. Because learners in an on-demand learning environment work independently on the complete learning cycle (i.e., assess-select-perform) without any help of a system or a teacher, learners should derive the quality of their past performance from *self*-assessments in order to determine how their performance can be improved (Garrison, 1997). Finally, the invested mental effort during task performance should be considered. The amount

of effort invested in a certain task in combination with the quality of the performance indicates a learner's level of expertise (Kalyuga, 2006; Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013). For instance, if one learner invests more effort to solve a learning task but performs the task equally successful as a second learner who has invested less mental effort, the first learner's expertise level is lower. So, the first learner should select a less complex next task to perform than the second learner.

When the learners recognize what the relevant task aspects are, they need to combine their performance and mental effort with the relevant task aspects to decide what an appropriate next task would be to improve their expertise level (Van Merriënboer et al., 2006). For example, when a learner has performed a difficult task with a low level of support, and self-assessed that this task was not performed successfully but did require a high amount of mental effort, it should lead him or her to select a new task that is less difficult and/or with a higher amount of support.

Since novices often do not have the skills and knowledge to appropriately execute this task-selection process they may not profit from the benefits of an on-demand environment (Corbalan et al., 2010). They need information on which task they should select to set an appropriate learning pathway to improve their learning. Researchers recommend providing them with advice on task selection (Hannafin, 1984; Kalyuga, 2009; Ross et al. 1989; Kirschner, Sweller, & Clark, 2006) to provide enough information for them to know how to select a suitable task to gain domain knowledge and to prevent the impediment of learning how the task-selection process works. These researchers suggest this advice might help learners to improve domain-specific skills development.

### 1.3 Task-selection advice

Ross et al. (1989) argues that task-selection advice may help learners to deal with on-demand education when it guides them to a suitable task selection providing specific

information what task aspects to select for selecting a task according to their individual learning needs. Therefore, it might help learners to get acquainted with and focus on the relevant task-selection aspects. Furthermore, this might prevent learners from getting overwhelmed by a large number of tasks to choose from because helping learners to focus on the relevant task-selection aspects might help them to indicate a subset of suitable tasks with an appropriate level of difficulty and support to choose from, hereby limiting the amount of choice (Shin, Schallert, & Savenye, 1994).

Other researchers (Hannafin, 1984; Kalyuga, 2009; Tattersall et al., 2005; Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013; Van Merriënboer et al., 2006) suggest that task-selection advice in general might contribute to setting an appropriate learning pathway in which learners can work on their points-of-improvement and improve their development of domain-specific skills, by helping them to select a suitable next task that optimally fits their learning needs. Therefore, an on-demand educational environment with task-selection advice might enable novices to deal with the unfamiliar environment and improve learning.

However, little research tested if such an educational environment with task-selection advice has the expected effects on learners' development of domain-specific skills. One of the few empirical studies that did test such an environment with task-selection advice is a study of Bell and Kozlowski (2002). They tested what the effects of 'adaptive guidance' on learning are in an on-demand environment in which learners had to learn skills on tracking targets on a radar system. This adaptive guidance can be seen as task-selection advice to help learners make appropriate task selections, since it provided information on what tasks needed to be selected in order to improve performance. In addition, the adaptive guidance structured the task database (i.e., the tasks the students could make a selection from) with different learning topics to be learned. It turned out that learners who studied with the guidance set a better

learning pathway and gained more strategic domain knowledge than learners who studied without the guidance. From this study, however, it does not become clear whether these results can be ascribed to either structuring the task database or to the information concerning the tasks to select in order to improve performance. Moreover, the adapted guidance was based on the learner's performance as assessed by the system. We argue that in order to teach novices how to handle the learning cycle of assessment–selection–performance in an on-demand educational environment entirely on their own they should learn to use their *self*-assessments to base their task selection on. When advice is based on expert-assessments learners might become accustomed to the system providing the assessment and not focus on their self-assessments when selecting a next task. Expert-assessments might thus prevent learners from investing enough effort in their self-assessments, which would make it more difficult for them to appropriately deal with the cycle of assessment–selection–performance when they do not receive this advice. Therefore, the advice should be based on their self-assessments.

Kicken, Brand-Gruwel, Van Merriënboer, and Slot (2009) have conducted another study, the results of which indicate that task-selection advice can have positive effects on learning. Two groups of learners had to deal with an on-demand learning environment in the hairdressing domain, one group received task-selection advice and one group did not. All learners used a development portfolio to assess their performance and select new tasks. Learners in the condition with advice selected more suitable tasks and showed higher improvement of domain-specific skills than learners in the other condition. However, the learners with advice received not only advice on task selection but also advice on how to improve the quality of their self-assessments. Due to this additional advice, it is not clear whether the results can be ascribed to the task-selection advice, to the self-assessment advice, or to their combination of both. To study the effects of task-selection advice on the domain-

specific skills of novices in an on-demand educational environment, a control condition is needed that does not provide any help to deal with the environment and an advice condition with task-selection advice but without self-assessment advice.

#### 1.4 The current study

The present study tested whether an on-demand educational environment with task-selection advice helps novices improve development of domain-specific skills. An advisory model has been designed and implemented, directing the learners' attention to the difficulty and support level of the last performed task and their self-assessment and mental-effort scores (i.e., relevant task-selection aspects) in order to help them determine their level of expertise and formulate their learning needs. In addition, the advice includes a short recommendation that includes what level of difficulty and level of support are best to select for the next task. This straightforward task-selection advice points out a subset of suitable tasks, and, therefore, might also decrease the overwhelming effect of a large task database for novices. It is hypothesized that learners who receive task-selection advice will develop more domain-specific skills and knowledge, and, therefore, will show higher test results than learners without this advice.

## 2. Method

### 2.1 Participants

Bachelor students from two international Schools of Business and Economics in The Netherlands were invited to participate in the experiment. In total 30 students participated (14 females and 16 males, mean age = 21.20 years;  $SD = 2.02$ ). They received a payment in Euros after they had finished the experiment. All participants earned at least € 10 and could increase their payment with an additional amount of up to € 30, depending on their scores in the test phase.

The participants were randomly assigned to the experimental conditions: Task



selection without advice ( $n = 15$ ) and task selection with advice ( $n = 15$ ). They had no experience with selecting tasks in an on-demand learning environment. All materials were written in English because participants had different nationalities and participated in educational programmes taught in English.

## 2.2 Materials

The participants worked for approximately four hours in an electronic environment. This environment contained a 1) prior knowledge test, 2) introduction to the experiment, 3) a learning environment with a task database with learning tasks from which participants could select their tasks, a self-assessment screen, and a task-selection screen with or without advice on task selection (depending on the condition), and 4) test tasks. The learning domain was system dynamics, which is a methodology to enhance learning about the behavior of a complex system over time (Richardson & Pugh, 1981). A complex dynamic system is simplified by depicting it in a model using flow(s) and/or stock(s) and depicting the behavior of the system over time in graphs.

### 2.2.1 Prior knowledge test

The participants were tested on their knowledge of system dynamics to make sure they were novices. Participants had to perform a prior knowledge test which contained 12 multiple-choice questions regarding basic concepts of the learning domain. This test showed no significant difference in prior knowledge between conditions,  $p = .448$ .

### 2.2.2 Introduction

The introduction contained practical information about the experiment and the learning environment, information about the various difficulty levels and support levels of to-be-selected learning tasks, and an explanation of some basic elements in the system dynamics domain (e.g., symbols used for stocks and flows).

### 2.2.3 Learning tasks and task database

In system dynamics, the complex dynamic system is first depicted in a model, and then the behavior over time of the elements in the model is depicted in a table to be able to draw graphs of the flow(s) and/or stock(s). Research on system dynamics has shown that many participants have difficulties learning the conceptual knowledge and modeling skills necessary to draw the system-dynamics models and graphs (Cronin & Gonzalez, 2007; Cronin, Gonzalez, & Sterman, 2009).

The system-dynamic learning tasks are designed following a whole-task approach that requires learners to acquire all components of domain-specific skills in a real-life, authentic task in which all their domain knowledge and skills need to be combined (Van Merriënboer & Kester, 2008). This approach leads to high learning outcomes and better transfer results. Therefore, all learning tasks contained a case description of a system-dynamics system and three assignments to practice all components of the system-dynamics skill in each task. An example of a case description is: ‘A printing office prints posters for a client who will use the posters to advertise for an upcoming festival. The office already has 100 festival posters printed. Every hour the office prints 50 festival posters. From the 8<sup>th</sup> hour on, the printing office begins to also print posters for another client parallel to the festival posters. Therefore, from this hour on they have less capacity to print the festival posters and the office can only print 25 festival posters each hour.’

\*\*\*insert Figure 1 about here\*\*\*

The first assignment was always to depict the case description as a system dynamics model (see Figure 1). A model had to be built by dragging system-dynamics elements from a legend and connecting them to other elements to form a model which depicted the flows, stocks, and relationships between them. Moreover, the participants had to connect a label to each element in which they provided a description of it. The second assignment was to fill out a table with descriptions of the behaviour of each flow and stock over time. The third

assignment contained three sub assignments, each of which consisted of a multiple-choice question with four answer options to choose the correct graph. Participants had to explain their choice to check for guessing. The first sub assignment was to choose the correct inflow/outflow graph. The second sub assignment requested the participants to choose the correct net flow graph. And the third sub assignment asked them to choose the correct stock graph.

Each participant had to perform 9 learning tasks in total. Every learning task had to be selected of a task database. The task database contained 81 learning tasks organized following the whole-task approach ordering tasks gradually from simple-to-complex whole tasks (Van Merriënboer & Kester, 2008). The four-component instructional design model (4C/ID-model; Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013) provides guidelines for designing on-demand learning environments and use the whole-task approach while preventing learners to become cognitively overloaded (Merrill, 2002). The model indicates to structure the task database clearly by distinctively ordering tasks by their difficulty and support levels. Hence, when participants had to select a task, they received a task-selection screen, which contained the task database in which the tasks were structured by distinctive difficulty and support levels (see left part of Figure 2). The following question was provided: “Select the task you think is best to perform next. Click on a radio button representing a task to see what its title and description is. The numbers indicate the sequence in which you selected the previous tasks”.

\*\*\*insert Figure 2 about here\*\*\*

After a task was selected by the participant it appeared on the computer screen step by step and had to be performed step by step. As soon as the first step of the task was completed the second step was shown. This process continued until all steps were completed. The answers given on the previously performed steps were logged and could not be changed to

prevent the participants from changing their answers on prior steps when learning from the following steps.

A learning task could be at one of three *difficulty levels* (i.e., easy, moderate, difficult). Difficulty levels were implemented to give participants the opportunity to perform tasks of increasing difficulty. The differences between the difficulty levels were derived from studies in the field of system dynamics showing that the presence of a stock often results in counter-intuitive behavior (Cronin & Gonzalez, 2007; Cronin et al., 2009). That makes models containing a stock more difficult than models without a stock, even in simple models. If one or more flows are accumulating into or out of a stock, it turns out that this accumulation is making it a difficult problem to solve. Therefore, the difficulty levels were implemented as follows. “Easy” learning tasks concerned a model with an inflow, an outflow, and a residual flow, but without a stock so that the difficult accumulation principle is not required. “Moderate” learning tasks concerned a model with one inflow and one stock, so that the included accumulation principle makes them more difficult. “Difficult” learning tasks concerned a model with an inflow, a stock, and an outflow, so that these are even more difficult because there are now two flows accumulating.

The tasks of each difficulty level could be one of three *support levels* (i.e., high, medium, low). The support levels were designed following the scaffolding principle of the 4C/ID-model (Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013). The model proved to be useful for designing support levels in previous studies of Corbalan, Kester, & Van Merriënboer (2009b, 2011) and Kicken et al. (2009). In the support level “high”, the task was a worked-out example in which all solution steps and correct answers were provided to the participant. In the support level “medium”, a completion task was provided in which some solution steps including the correct answers were provided and some solution steps and correct answers were not, so that the participants had to come up with the solution steps for

the missing answers and had to fill in the missing answers themselves. The support level “low” provided a conventional task in which no solution steps and no correct answers were given, so that the participants had to come up with all the solution steps and had to give all answers by themselves. For each combination of difficulty level and support level nine tasks were included in the task database; these tasks only varied in their case descriptions (note that, in principle, a learner thus had the opportunity to complete the whole experiment by selecting nine learning tasks with the same difficulty level and the same support level).

After each learning task, a self-assessment screen showed both the given and the correct answers for each step, so that participants could compare their own answers with the correct answers. On the basis of this information, the participants received the question: ‘Assess your understanding of/ performance on the task filling in the rating scale’, with ‘understanding of’ when a participant worked on a worked-out example and ‘performance on’ when a participant performed a completion tasks or conventional task. To self-assess their task performance the participants filled in a five-point rating scale (i.e., from 1 = ‘bad performance’ to 5 = ‘good performance’). Moreover, the participants received the mental-effort question: ‘How much mental effort did you invest to study/complete/perform the task?’, with ‘study’, ‘complete’, and ‘perform’ if the participants just performed a worked-out example, a completion task, or a conventional task respectively. They had to answer the question using a five-point rating scale (i.e., from 1 = ‘very low’ to 5 = ‘very high’; Paas, Tuovinen, Tabbers, & Van Gerven, 2003).

In the learning phase, all screens contained the same menu at the top of the screen (see Figure 1) with links to different parts of the introduction of the experiment: (a) the *task introduction* with basic concepts of system dynamics, (b) the *task aspects*, that is, a short explanation of the different difficulty levels and support levels as well as a description that there were 9 tasks with different case descriptions for each combination of difficulty level and

support level, (c) an explanation of the *learning environment* containing information on the learning-task contents, the mental effort question, the self-assessment screen, and the task-selection screen, and (d) a link to a list with *performance standards*, which specify the criteria for acceptable performance (Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013).

#### 2.2.4 Task-selection advice

In the condition without advice participants received the task-selection screen without any added information (see the left part of Figure 2). Learners in the condition with advice on task selection received additional information, that is, the advice. This task-selection screen with additional advice is depicted at the right part of Figure 2.

The advice provided a sentence giving straightforward advice on what level of difficulty and level of support to select for the next task. The advice was generated by an algorithm using the participants' self-assessment and mental-effort scores (see the Appendix). The sentence with advice indicated how much the difficulty level and support level should be decreased/increased for a suitable next task selection, for example, "You are advised to select a task with one lower difficulty level and one higher support level than your previous task". In addition, the column in the task database with the subset of appropriate tasks as indicated by the sentence with advice was coloured to prevent the participant from making mistakes in translating the advice into an actual task selection. Still the participants were free to deviate from the advice, even if they selected nine times a task with the same difficulty level and the same support level.

### 2.3 Measurements

#### 2.3.1 Test tasks

To measure acquired system-dynamics skills, participants performed two test tasks. These test tasks were comparable with the most difficult tasks without support as used in the learning phase. Hence, the results on these test tasks show if the participants had been able to

learn system-dynamics skills in the learning environment.

Each test task contained three assignments. First, the “Model” assignment required the participants to draw a model that fits the case description. But, in contrast to the learning tasks, they did not receive any support and could not use a legend to drag the elements from to build the model. The participants had to perform this assignment with paper and pencil. Second, the “Table” assignment required participants to fill in an empty table just as in the learning tasks of support level low. Third, the “Graph” assignment required the participants to draw the graphs themselves instead of answering multiple-choice questions as in the learning tasks. The participants received a graph-drawing computer-programme with an introduction to use the programme. They had to fill in the most crucial points of a flow or stock to draw the characteristic graph. Afterwards they could check what the graph looked like and change it if they wanted to.

### 2.3.2. Knowledge test

To test knowledge of system-dynamics aspects that were described in the learning environment the learners had to perform a multiple-choice questionnaire. This questionnaire tested their knowledge of stock-flow problems, models, graphs, the definition of system dynamics, and the meaning of some system dynamics elements. The test included 25 questions and had a Cronbach’s alpha of .723.

### 2.4. Procedure

The participants were tested in groups of 1 to 7 persons in different sessions and each participant worked individually. During the experiment each participant sat behind a computer screen at least two meters away from others or separated from others by a wall to prevent them from interrupting one another or copying each others answers.

The participants completed the entire experiment in one session. The learning environment was designed in such a way that they were guided through the experiment in

which first the introduction was provided, second the learning phase, and third the test phase. They worked at their own pace. All actions in the electrical environment and the time spent on each task were logged.

First, all participants received the introduction to the experiment. Moreover, they were informed that in the learning phase they should learn about system dynamics by performing 9 self-selected learning tasks in total, and that their knowledge and skills would be tested in the test phase.

Second, the learning phase started. The learning environment provided the task-selection screen to all participants. After selecting and performing a learning task, they answered the mental-effort question and received the self-assessment screen to assess their performance using the self-assessment rating scale. After the self-assessment screen, the task-selection screen was provided again and they selected a subsequent task. This cycle of task selection, task performance, and self-assessment was repeated in this order until they had selected 9 learning tasks, performed them, and assessed their task performances. Participants who were in the condition with advice received advice in the task-selection screen each time they had to select a task. When they started the learning phase they had to select a task for the first time. Since no self-assessment and mental-effort scores were available to give advice for selecting the first learning task, the advice for all participants in the advice condition was to select a task of easy difficulty and high support. The participants in the condition without advice had to select a next task (including the first task) without any advice. During the learning phase all participants were allowed to look at the menu with the links at the top of screen of the learning environment at any time and as often as they wanted. After completing 9 tasks in the learning phase, they could take a 10-minute break in which they were not allowed to talk with each other and should not disturb other participants still working.

Third, after the break the participants continued with the test phase. In this phase they



could not check the links in the menu anymore. All participants performed the two test tasks without any support or advice. The sequence of providing the test tasks to them was randomised. After having performed the test tasks they were given the knowledge test. The sequence of the questions of the knowledge test was randomised for each participant too.

### 2.5. Scoring

To check if the participants in the advice condition actually followed the advice or ignored it, the percentage of times they followed the advice and thus selected one task of the subset of suitable tasks indicated by this advice was calculated. In addition, it was checked if this percentage was different from the percentage of times the participants in the no-advice condition selected a task that would have been part of the subset of suitable tasks indicated by the advice if it was provided to them.

To test if an on-demand educational environment with task-selection advice can help novices to develop domain-specific skills the two conditions were compared regarding their performance on the learning tasks and test tasks. A system-dynamics expert assessed all learning tasks, except for the learning tasks of support level high since these tasks were worked-out examples (i.e., learners had not to provide any input for these tasks). The completion and conventional task performances were assessed on each solution step of a learning task. The maximum number of points a participant could receive for a solution step was different for the tasks of different difficulty and support levels, depending on (a) the number of flows and/or stocks included in the case description, and (b) the support level because more answers had to be given in tasks with lower support. To be able to compare the learning-task performances between the conditions, the expert was using the same five-point rating scale as the self-assessment rating scale in the self-assessment screen in the learning environment. Therefore, the maximum number of points of an expert-assessment score was 5 points for each solution step. In addition, the mean expert-assessment was calculated for each

task performance and for each participant's performance in the learning phase

The number of learning tasks performed for each combination of difficulty level and support level was analysed to see if there was a difference in number of performed learning tasks of particular levels. If there was such a difference it could have affected the performance on the tests. These data are analysed using the chi-square Goodness-of-Fit test.

To see if the conditions differed in the *quality* of their self-assessments, the self-assessments scores of the participants were compared to the expert-assessments scores. The quality of self-assessments could influence the results in the learning phase and the test phase because the advice was based on these self-assessments, therefore, quality of the given advice would have been higher if the quality of self-assessments was higher. The quality of self-assessments was analysed by correlating the self-assessment scores with expert-assessment scores. Kendall's tau was used because the number of participants is low.

The computer programme calculated the knowledge test scores for which each participant could receive a maximum of 25 points. The performance on the test tasks was assessed by the system-dynamics expert. For the performance on the test tasks the participants could receive a maximum of 80.75 points, broken down in 20 points for the "Model" assignment, 31.5 points for the "Table" assignment, and 29.25 points for the "Graph" assignment. For 8 participants their performance on the test tasks was also scored by a second system-dynamics expert to calculate the inter-rater reliability which was  $r = .985$ , indicating very high agreement.

### **3. Results**

#### **3.1. Learning phase**

First, it has been checked if the participants in the advice condition followed the advice and if they selected more tasks of the advised subset of tasks than the participants without advice – assuming that these participants would have received the advice. For the

advice group, it turns out that 76% of the selected tasks were indeed tasks of the subset indicated by the advice. For participants in the no-advice condition, 47% of the selected tasks were tasks that would have been part of the subset of suitable tasks if advice had been given. Hence, participants with advice selected significantly more tasks that were within the advised subset of tasks than participants in the no-advice condition ( $p < .0001$ ).

A comparison of expert-assessment scores with self-assessment scores show a correlation between the self-assessment scores and the expert-assessment scores of  $r = .183$  ( $p = .020$ ) for the condition with advice, and  $r = .251$  ( $p = .002$ ) for the condition without advice. These correlations are low and thus suggest that the overall quality of the self-assessments is low. The correlations between the self-assessment scores and expert-assessment scores are not significantly different from each other between both conditions ( $p = .629$ ), indicating that there is no difference in the quality of the self-assessment of the participants in the conditions with and without advice.

Table 1 presents for both conditions the results on performance, mental effort, and time on task in the learning phase.

\*\*\*insert Table 1 about here\*\*\*

There is no significant difference between the conditions on performance on the learning tasks,  $t(27) = .81$ ,  $p = .424$ , with a Cohen's  $d$  of .31. In both conditions participants performed satisfactory in the learning phase. Moreover there is no significant difference on reported mental effort,  $t(28) = .40$ ,  $p = .695$ , with a Cohen's  $d$  of .15, and time on task,  $t(28) = .45$ ,  $p = .660$ , with a Cohen's  $d$  of -.17. Regarding the number of learning tasks performed for each combination of difficulty level and support level, there are also no significant differences between the conditions,  $\chi^2(8, N = 270) = 1.77$ ,  $p > .05$ .

Since time on task and mental effort in the learning phase can have an effect on learning in the learning phase and, therefore, on performance in the test phase, they are

included as covariates in the analyses of the knowledge test and test-task performance.

### 3.2. Test phase

Table 2 presents for each condition the results on the knowledge test and on the test tasks, including the results on each separate assignment of the test tasks.

\*\*\*insert Table 2 about here\*\*\*

An ANCOVA on the knowledge test scores does not show a significant difference between the conditions,  $F(1, 26) = 0.04$ ,  $MSE = 6.87$ ,  $p = .836$ . Both conditions scored satisfactory on the knowledge test. Regarding the performance on the test tasks and every test-task assignment the participants in the no-advice condition performed in general good and the participants in the advice condition performed satisfactory. An ANCOVA on the test-task scores shows a clear trend in favour of the no-advice group,  $F(1, 26) = 4.18$ ,  $MSE = 147.01$ ,  $p = .051$ ,  $\eta_p^2 = .14$ . Participants without advice score somewhat higher on the test tasks ( $M = 67.90$ ,  $SD = 10.47$ ) than participants with advice ( $M = 60.21$ ,  $SD = 15.48$ ). The results on each separate assignment show that the participants without advice draw better models ( $M = 16.93$ ,  $SD = 2.99$ ) than the participants with advice ( $M = 14.42$ ,  $SD = 3.97$ ),  $F(1, 26) = 7.92$ ,  $MSE = 8.08$ ,  $p = .009$ ,  $\eta_p^2 = .23$ , and also draw better graphs ( $M = 24.26$ ,  $SD = 3.60$ ) than the participants with advice ( $M = 20.64$ ,  $SD = 6.94$ ),  $F(1, 26) = 4.30$ ,  $MSE = 28.68$ ,  $p = .048$ ,  $\eta_p^2 = .14$ . The difference for performing the “Table” assignment is in the same direction but not statistically significant,  $F(1,26) = 0.96$ ,  $MSE = 33.65$ ,  $p = .336$ . Overall, these results show that the participants without advice outperform the participants with advice.

## 4. Discussion

This study investigated the effects of giving task-selection advice on the development of system-dynamics skills in a full on-demand educational environment. Participants were not familiar with selecting their own learning tasks and were novices in the system-dynamics domain.

The hypothesis that learners who receive advice on task selection will have higher test results was not confirmed. On the contrary, for the “Model” and “Graph” assignments of the test tasks, the condition without advice outperformed the condition with advice; for the “Table” assignment, no difference between the conditions was found. The test-task “Table” assignment was identical to the learning-task “Table” assignment of support level low. The test-task “Model” and “Graph” assignments, however, were *not* identical to the learning-task “Model” and “Graph” assignments, because additional and unfamiliar solution steps were required. In the test-task “Model” assignment the learners had to draw the model themselves, thus, without a legend from which they could drag the system-dynamics elements to construct the model. In the test-task “Graph” assignment, the learners had to draw the graph themselves instead of choosing it of four possible graphs. Thus, learners in the no-advice condition outperformed learners in the advice condition on those test tasks that required them to perform *unfamiliar* solution steps.

Because the number of performed learning tasks of particular difficulty levels and support levels, the time-on-task in the learning phase, and the quality of the self-assessments did not differ between conditions, these variables cannot easily explain why learners not receiving advice reached higher test performance. The finding that learners without advice are better able to perform unfamiliar additional solution steps seems to indicate that they gained deeper understanding of the strategies and processes of solving system-dynamics problems. One plausible explanation might be that learners without advice considered their task selection more thoughtfully in order to select a task that best fitted their learning needs and, therefore, focussed more strongly on fulfilling these learning needs while performing the task – leading to deeper comprehension and understanding. Learners with advice, on the other hand, might have mindlessly followed the recommendation and selected one task of the subset of advised tasks without considering why they should perform precisely this particular task

and how it might help them to fulfil their learning needs. The marking of the subset of advised tasks might contribute to this. Perhaps it might even have caused learners to only look at the marked column and select a next task of it without even looking at the recommendation. They probably, without considering why they selected the task, continued performing the selected task without focusing on improvement of performance. Moreover, although the results on mental effort in the learning phase do not show significant differences between the conditions, learners without advice invested somewhat more effort in performing the learning tasks than learners with advice. This might also support the explanation that learners without advice were more involved in the learning process and invested more effort in understanding system-dynamics problem solving (Paas, Tuovinen, Van Merriënboer, & Darabi, 2005).

The results on the knowledge test showed no significant differences between the conditions. This result replicates the findings on the test-task “Table” assignment and confirms that learners in both conditions gained the same amount of system-dynamics knowledge necessary to perform familiar tasks or answer multiple-choice questions on their gained knowledge on system dynamics. The sole difference between conditions seems to be on outcomes requiring deeper understanding of system-dynamics problem solving, although we should be a bit cautious interpreting the results because the effect-sizes indicate rather small differences and the number of participants was low. Therefore, future research should use a larger group of participants.

Yet, the results clearly show that the used advisory model did not provide effective advice to better enable novices develop domain-specific skills. The advisory model might have provided too much direction by explicitly indicating which tasks are suitable to select. The given advice did not encourage learners to become involved in the decision-making process of task selection (Moreno & Mayer, 2007) and, consequently, did not trigger them to consider how this process works and to reach a deeper understanding of the system-dynamics

problem solving process. So, learners probably did not focus on relevant task-selection aspects but mindlessly followed the given advice without considering why this specific task would fit their learning needs. Moreover, the advisory model did not inform learners on how relevant task-selection aspects needed to be combined for determining the best task to select. This is in line with the warning some researchers give for providing *insufficient* advice (Kirschner et al., 2006), which might not improve or might even hamper the development of domain-specific skills. Not knowing why advised tasks are actually advised might have led to misconceptions (Moreno & Mayer, 2007); for example, when a learner – incorrectly – thinks the advice is mainly based on his or her invested mental effort, he or she will mainly focus on the effort invested in subsequent learning tasks and not on improvement of specific skills and self-assessments.

Moreover, it needs to be considered that perhaps the algorithm used for the advice might not have constantly delivered suitable step sizes. Although the algorithm is based on the main guidelines of the 4C/ID-model (Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013) it did not calculate to provide optimal subset of tasks for all possible situations. For example, the algorithm might have provided step sizes that caused learners to be advised to select tasks with the same difficulty and support level repeatedly. This might have de-motivated the learners. Or if a learner did select a difficult task with low support for the first time instead of selecting an advised easy task with high support it might have been too difficult to perform. When the learner had assessed his/her performance as ‘bad’ and indicated to have invested a high amount of mental effort, the suitable advice would have indicated to select an easy task with high support. However, because of the maximum available step sizes the algorithm could only have indicated a maximum of four step sizes to decrease. In addition, the advice indicated to select a medium task with medium support which might again be too difficult for the learner to perform. So, the advice might not always

have been appropriate for the learners. Therefore, in the future research should be done in which advice is provided to learners which is based on an algorithm adjusted to these kinds of situations which can lower the quality of the advice or which is not based on an algorithm at all.

To help learners learn about the process of task-selection, it might be better not to provide advice that straightforwardly indicates which tasks are suitable to select because this invites learners to mindlessly follow the advice. A better option might be to give advice that contains information on how the process of task selection works (Gay, 1986; Kinzie, 1990), telling the learners which task-selection aspects should be taken into account and how these aspects should be used to select a suitable task. For example, more effective advice might explain how self-assessments and mental-effort measures might help to select a future task, but leave it up to the learners to make the actual task selection.

Finally, the results show that learners in the condition with advice selected more tasks of the advised subset of tasks than learners in the condition without advice – assuming that they would have received this advice. Thus, although the self-assessments in both conditions were of low quality, giving the advice had the intended effects of selecting more tasks of the advised subset of tasks. Yet, the advice based on a self-assessment of low quality will be itself of low quality. To prevent inaccurate self-assessments, indirectly hampering the beneficial effects of task-selection advice, in addition to advice on task-selection advice on self-assessment could be added. Self-assessment advice could provide learners with standards of acceptable performance (Kicken et al., 2008) and continuous information on how to improve performance (Moreno & Mayer, 2007; Tennyson, 1980; Tennyson & Buttrey, 1980). If advice triggers learners to compare their own performance with given standards, this will help them improve their self-assessment skills, monitor their progress, and, in addition, set a suitable learning trajectory adapted to their learning needs (Bell & Kozlowski, 2002).



The fact that learners without the advice did have a low self-assessment quality too but did outperform the learners with the advice, might indicate that the learners without the advice did not take their self-assessment into account when selecting their tasks and performing them. The suggested self-assessment advice might also help these learners to improve their self-assessment quality and indicate that they should consider their self-assessments for appropriate task-selections. Hence, their task-selection quality might be higher resulting in a learning trajectory that suits their learning needs better. Therefore, self-assessment advice might also contribute to their improvement of learning. This points out the necessity for future study on advice on self-assessment and task-selection to improve domain-specific skills in an on-demand environment.

Bell and Kozlowski (2002) also argued that next to task-selection advice, ordering the topics of the tasks in a ramped sequence can help learners select an appropriate task. Moreover, Gay (1986) showed that structuring an environment helps learners with low-prior knowledge to improve learning. The learning tasks in the database in our study were ordered following the whole-task approach from simple tasks to complex tasks, and, therefore, helped learners in both conditions to set an appropriate pathway and improve learning (Van Merriënboer & Kester, 2008). Moreover, the tasks were ordered distinctively in difficulty levels and support levels helping all learners to recognize the two relevant task-selection aspects (Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013). This might have helped learners ‘without’ advice to select suitable tasks. In future research, the effects of task-database structure should be investigated in further detail.

In conclusion, the reported experiment showed that task-selection advice did not help but even hampered novices' development of system-dynamics problem solving skills. So, it seems that straightforward advice should better not be used in on-demand education because it prevents learners to consider *why* they should select a specific task (i.e., consider how the

task might contribute to improving their performance), which may negatively affect their understanding of the required problem-solving process and thus transfer of learning. Future research is needed to investigate whether advice that encourages learners to *think* about their task selections has more positive effects on the development of domain-specific skills.

**References**

- 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. doi: 10.1111/j.1744-6570.2002.tb00111.x
- Clarebout, G., & Elen, J. (2009). Benefits of inserting support devices in electronic learning environments. *Computers in Human Behavior, 25*, 804-810. doi:10.1016/j.chb.2008.07.006
- Corbalan, G., Kester, L., & Van Merriënboer, J. J. G. (2008). Selecting learning tasks: Effects of adaptation and shared control on efficiency and task involvement. *Contemporary Educational Psychology, 33*, 733-756. doi:10.1016/j.cedpsych.2008.02.003
- Corbalan, G., Kester, L., & Van Merriënboer, J. J. G. (2009a). Combining shared control with variability over surface features: Effects on transfer test performance and task involvement. *Computers in Human Behavior, 25*, 290-298. doi:10.1016/j.chb.2008.12.009
- Corbalan, G., Kester, L., & van Merrienboer, J. J. G. (2009b). Dynamic task selection: Effects of feedback and learner control on efficiency and motivation. *Learning and Instruction, 19*, 455-465.
- Corbalan, G., Kester, L., & van Merrienboer, J. J. G. (2011). Learner-controlled selection of tasks with different surface and structural features: Effects on transfer and efficiency. *Computers in Human Behavior, 27*, 76-81.
- Corbalan, G., Van Merriënboer, J. J. G., & Kicken, W. (2010). Shared control over taskselection: A way out of the self-directed learning paradox? *Technology, Instruction, Cognition and Learning, 8*, 119-136.
- Cronin, M. A., & Gonzalez, C. (2007). Understanding the building blocks of dynamic systems. *System Dynamics Review, 23*, 1-17. doi:10.1002/sdr.356

- Cronin, M. A., Gonzalez, C., & Serman, J. D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes*, *108*, 116-130. doi:10.1016/j.obhdp.2008.03.003
- Garrison, D. R. (1997). Self-directed learning: toward a comprehensive model. *Adult Education Quarterly*, *48*, 18-33.
- Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. *Journal of Educational Psychology*, *78*, 225-227.
- Gray, E. (2003). *Conscious choices: A model for self-directed learning*. Upper Saddle River, New Jersey: Prentice Hall.
- Hannafin, M. J. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction. *Journal of Instructional Development*, *7*(3), 6-10.
- Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*, 995-1006. doi:10.1037//0022-3514.79.6.995
- Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. *Applied Cognitive Psychology*, *20*, 333-342. Doi:10.1002/acp.1249
- Kalyuga, S. (2009). Instructional designs for the development of transferable knowledge and skills: A cognitive load perspective. *Computers in Human Behavior*, *25*, 332-338. doi:10.1016/j.chb.2008.12.019
- Kicken, W., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Slot, W. (2009). The effects of portfolio-based advice on the development of self-directed learning skills in secondary vocational education. *Educational Technology Research and Development*, *57*, 439–460. doi:10.1007/s11423-009-9111-3

- Kinzie, M. B. (1990). Requirements and benefits of effective interactive instruction: Learner control, self-regulation, and continuing motivation. *Educational Technology Research and Development, 38*(1), 5-21.
- Kinzie, M. B., & Sullivan, H.J. (1989). Continuing motivation, learner control, and CAI. *Educational Technology Research and Development, 37*(2), 5-14.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.  
doi:10.1111.169.8810
- Kostons, D., Van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22*, 121-132. doi:10.1016/j.learninstruc.2011.08.004
- Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43-59.
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. Special Issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review, 19*, 309-326. doi: 10.1007/s10648-007-9047-2
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71. doi: 10.1207/s15326985EP3801-8
- Paas, F., Tuovinen, J. E., Van Merriënboer, J. J. G., & Darabi, A. A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development, 53*(3), 25-34.
- Richardson, G. P., & Pugh, A. L. (1981). *Introduction to system dynamics modeling*. Portland,

OR: Productivity Press.

- Ross, S. M., Morrison, G. R., & O'Dell, J. K. (1989). Uses and effects of learner control of context and instructional support in computer-based instruction. *Educational Technology Research and Development, 37*(4), 29-39.
- Schnackenberg, H. L., & Sullivan, H. J. (2000). Learner control over full and lean computer-based instruction under differing ability levels. *Educational Technology, Research and Development, 48*(2), 19-35. doi:10.1007/BF02313399
- Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: HarperCollins.
- Shin, E. C., Schallert, D. L., & Savenye, W. C. (1994). Effects of learner control, advisement, and prior knowledge on young students' learning in a hypertext environment. *Educational Technology Research and Development, 42*(1), 33-46.
- Shute, V. J., Gawlick, L. A., & Gluck, K. A. (1998). The effects of practice and learner control on short- and long-term gain and efficiency. *Human Factors, 40*, 296-310.
- Tattersall, C., Manderveld, J., Van den Berg, B., Van Es, R., Janssen, J., & Koper, R. (2005). Self organising wayfinding support for lifelong learners. *Education and Information Technologies, 10*, 109-121.
- Tennyson, R. D. (1980). Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. *Journal of Educational Psychology, 72*, 525-532.
- Tennyson, R. D., & Buttrey, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. *Educational Communication and Technology Journal, 28*, 169-176.
- Topping, K. (2003). Self and peer assessment in school and university: reliability, validity and utility. In M. Segers, F. Dochy & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55-87). Dordrecht, The

Netherlands: Kluwer Academic Publishers.

- 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., & Kester, L. (2008). Whole-task models in education. In J. M. Spector, M. D. Merrill, J. J. G. Van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational Communications and technology* (3<sup>rd</sup> ed., pp. 441-456). New York: Routledge/Taylor & Francis Group.
- Van Merriënboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning* (2<sup>nd</sup> Rev Ed.). New York: Taylor & Francis.
- Van Merriënboer, J. J. G., Sluijsmans, D., Corbalan, G., Kalyuga, S., Paas, F., & Tattersall, C. (2006). Performance assessment and learning task selection in environments for complex learning. In J. Elen & E. Clark (Eds.), *Handling complexity in learning environments: Theory and research* (pp. 201-220). Oxford, UK: Elsevier.
- Williams, M. D. (1996). On-demand and instructional technologies. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 957-983). New York: Simon & Schuster Macmillan.

### Appendix A. Algorithm For Providing Task-Selection Advice

To give advice on task selection, nine complexity levels were distinguished for each combination of difficulty level and support level (see Table A.1)

Table A.1

*All complexity levels distinguished for each combination of difficulty level and support level.*

Complexity level	Difficulty level	Support level
1	Easy	High
2	Easy	Medium
3	Easy	Low
4	Moderate	High
5	Moderate	Medium
6	Moderate	Low
7	Difficult	High
8	Difficult	Medium
9	Difficult	Low

Advice on the complexity level of the next to-be-selected task was based on the participant's self-assessment score and mental-effort score on the last performed task, according to Table A.2.

Table A.2

*The Algorithm for Providing Advice On Task Selection With All Possible Options For Each Combination Of the Participant's Self-Assessment(SA) And Mental Effort (ME).*

SA	1	2	3	4	5
ME					
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

The first row indicates the score on the self-assessment rating scale (from 1 'bad' to 5 'good') and the first column indicates the score on the mental-effort rating scale (from 1 'very



low' to 5 'very high)'. Each combination of self-assessment and mental effort leads to a step size that was added to or subtracted from the complexity level of the previous task, depending on the step size being positive or negative, respectively. This led to the complexity level of the next task to be advised.

For example, if a participant had just performed a task of moderate difficulty level and high support (i.e., complexity level 4), and assessed his or her performance with a mean score of 'good' (i.e., a rating of 5) and a mental effort score of 'neither low nor high' (i.e., a rating of 3), the algorithm provided a step size of +2. This step size was added to complexity level 4 so that a next task was advised with complexity level 6, that is, a task with a moderate difficulty level and a low support level. The desired complexity level indicated by the algorithm was automatically transformed into a sentence formulating the advice. The sentences were constructed of this general sentence which contains all options on advised difficulty and support levels: 'You are advised to select a [two levels less/one level less/an equally/one level more] difficult task with [two times higher/one higher/equal/one lower/two times lower] level of support.'. The 4C/ID-model describes that it is important for participants to first test if they are able to successfully perform a conventional task on one particular difficulty level before continuing to the next difficulty level (Van Merriënboer, 1997; Van Merriënboer & Kirschner, 2013). Therefore, the rule was implemented which prohibited the advisory sentence to advise the participants to proceed to the next difficulty level before they had successfully performed a conventional task on the current difficulty level (i.e., the advice was then to perform tasks of the same difficulty level and with low support). This also implies that when participants performed a conventional task and the algorithm provided a positive step size, the participants were advised to select only one level more difficult task.

Table 1

*Means and Standard Deviations on Performance, Mental Effort, and Time on Task in the Learning Phase, and Mean Number of Selected Tasks for Each Combination of Difficulty Level and Support Level*

	No advice on task selection		Advice on task selection	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Performance	4.01	0.48	3.84	0.65
Mental effort	2.86	0.54	2.78	0.58
Time on task	501.67	182.61	532.87	200.86

---



---

 (seconds)
 

---

Table 2

*Means and Standard Deviations of the Performance Results on the Knowledge Test and the Retention Test for Each Condition Adjusted to Time on Task and Mental Effort of the Learning Phase*

	No advice on task selection		Advice on task selection	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Knowledge test	17.30	2.91	17.10	2.67
Retention test				
Overall	67.90	10.47	60.21	15.48
Model	16.93	2.99	14.42	3.97
Table	26.70	4.94	25.15	6.97
Graph	24.26	3.60	20.64	6.94

### **Figure Captions**

*Figure 1.* A screenshot of the first assignment of each learning task, showing the legend (top) and a system dynamics model constructed by the learner (bottom)

*Figure 2.* Two screenshots of the task-selection screen. The left screenshot is of the task-selection screen for the no-advice condition and the right screenshot is of the task-selection screen for the advice condition