HELP DESIGN AND HELP USE IN A COMPUTER BASED LEARNING ENVIRONMENT

J. C. Sakdavong, F. Amadieu, N. Huet and C. Dupeyrat
CLLE-LTC, UMR CNRS 5263, Université Toulouse le Mirail, 5 Allée A. Machado
31058 Toulouse Cedex 1, Toulouse, France

Keywords: Learning, Distance education, Computer supported education, ICT, AIED, Modelling, Help, Guidance, Metacognition, Metacognitive guidance, Multiagent system.

Abstract: This paper presents the framework of elaboration of help in a text-processing computer-based learning environment. A multiagent software system have been developed including help. Then it was tested in participants who had to learn autonomously and for which we analyzed the behavioral pattern of using help. Results showed a lack of regulation and highlight the need for adding a metacognitive guidance enabling learner to efficiently use help.

1 INTRODUCTION

The following study has been sponsored by the French National Agency for Research (ANR) for a period of four years from 2006 to 2011 (ref. CEAGMATIC).

For several years, the use of new technologies in higher education and vocational training were widely developed but did not inevitably entail an improvement of learning. Several studies showed that computer learning environments lead to poor learning (e.g., Dillon & Gabbard, 1998; Shapiro & Niederhauser, 2004). Furthermore, there is also a very high dropout rate in distance training environments (Lebel, 1995).

One of the reasons for this failure is the lack of a fine analysis of the sequence of learning, the sub-tasks required to acquire the contents of knowledge. From an ergonomic point of view the learning programs using new technologies were too often elaborated without modelling the task of learning that is without an analysis of the cognitive demands, the types of potential errors associated with a modelling of the activity of the learner (e.g., Boucheix, 2003; Leplat, 2002; Samurcay & Pastré, 2004).

In the present study, we will present how we elaborated the learning sequence taking into account the previous points.

Another reason which can contribute to this failure lies in the fact that the educational and technological tools require learners to take charge of their own cognitive learning in autonomy for which they are not all necessarily prepared (Azevedo & Cromley 2004). This problem occurs particularly in distance training (Lebel, 1995). To resolve these difficulties of autonomous learning, it may be useful to provide learner with helps (for example, glossaries, work-out exercises) for learning in their own (Baker, Puustinen & Lund, 2002). However, learning outcomes partly depend on the quality of the help provided (Puustinen, Volckaert-Legrier, Coquin & Bernicot, 2009). If the help does not meet the learners’ needs, it will not be used or will not improve learning. Taking into account an analysis of errors made on a task in help designing enables to meet learner’s needs and increase the relevance of the help provided. In the present study we will describe how we elaborated help in this sense. Research showed that help is often unsuited for learners’ needs and rarely or inappropriately used (e.g., Aleven et al., 2003; Puustinen et al., 2009; Roll et al., 2005).

Furthermore, the appropriate use of help is considered in the literature as a self-regulated strategy (Puustinen, 1998). Self-regulation requires metacognitive knowledge (e.g., Flavell, 1979), the knowledge that individuals have of their own knowledge, perceptions of difficulty, utility and cost and the relative efficiency of strategies. According
to some authors (e.g., Nelson-Le-Gall, 1981), an awareness of the need of help-seeking followed by the decision of seeking help depends on an individual’s metacognitive knowledge about the utility and cost of help-seeking. According to literature on metacognition, metacognitive knowledge or perceptions of the utility and cost of using strategies enabling a gain in performance is positively associated to strategy use (e.g., Pressley et al. 1985; Escribe & Huet, 2005). In an interactive learning environment of statistics, we found that the more students perceived the utility of on-line course, of worked-out problem, the more they used it (Noury, Huet, Escribe, & Narciss, 2008).

However, even when help is perceived as useful, in some studies it was found that help is seldom used (Hofer & al., 1996).

Furthermore, studies on help-seeking and learning showed divergent results. Some authors found that an increase of help used was not associated with an increase in performance (e.g., Clarebout & Elen, 2009), others found a positive relationship between the use of help and performance (Jiang, Elen & Clarebout, 2009). These divergent results reveal that the process of help-seeking is complex. The help-seeking process can be considered as a continuum from low to high self-regulation (Puustinen & Rouet, 2009). Several characteristics of self-regulated help-seekers or users have been identified (Puustinen et al. In press). A high self-regulated help-seeker is illustrated by: (1) a high awareness of the need for help as indicated by a high relation between confidence judgment in the accuracy of the answer and performance (2) the high rate of time before selecting help as an indicator of the necessity of the need for help; (3) an instrumental help selection (asking only for an explanation or a general procedure or solving principle and trying to solve the problem by oneself) rather than an executive help (that is ready-made answers); (4) the ability to re-invest the first use of help to the similar exercises.

Within the context of Interactive Learning Environment (ILE), the aim of the present study is 1) to show how we have elaborated help on the basis of a previous analysis of errors made by the learners; 2) to explore the use of help in the new device and in relation with perceptions of help utility and performance.

2 METHOD

2.1 Materials

2.1.1 The Text-processing Computer-based Learning Environment

The environment consisted in a software system for learning how to use word processor formatting and style-sheets. The learning material consisted of screen casts dealing with word processor’s procedures and application exercises. Within this environment; the participants were not allowed to transgress the sequence. This environment resulted from a previous exploratory study carried out on 80 students (Sakdavong, Huet & Adreit, 2009). The instructional screen casts were designed especially for the experiment. Each screen cast was composed of a narration from the teacher and either an explanation screen or a video of a word processor’s procedure. These screen casts dealt with formatting and using style-sheets. They were composed of explanations of declarative and procedural knowledge accompanied by examples. The screen casts duration was about half an hour.

After viewing the screen casts, participants had to deal with three exercises. These exercises were provided by the computer-based learning environment through two windows: one presented the subject of the exercise and the other was a fully functional generic word processor. All actions from the participants in these two windows were recorded by the computer system.

Each exercise was composed of two questions: an initial question and a checking question.

Exercise 2 was more complex than exercise 1 in that it required to master all of the procedures for formatting and using style-sheets.

After reading the subject of a question, each participant had to answer to it in the word processor window.

For example, in the second exercise, the word processor was presenting a small text from Molière and the initial question was:

You have to modify the “default” style in order that paragraphs are in Times New Roman font, size of 10 points, in italics and with justified paragraphs. You have to modify the “title 3” style in order that ...
Once a participant had validated her/his answers, she or he had to assess how confident he or she was in the accuracy of the answer and then, received a feedback. If the answer was right, the participant had to go on to the checking question. If wrong, the participant was invited to accept or refuse to get help. According to her/his choice, the participant was getting help or not, then she/he had to go to the checking question.

This checking question differed from the initial question in the following aspect: it was not associated with help or feedback. It had two roles: if the participant had failed the initial question, the checking question tested if the participant had transferred the chosen help; if the participant had been successful with the initial question, the checking question tested if she or he really knew how to solve the question or if she/he succeed by chance. A great care has been taken to ensure the two questions were equivalent in difficulty.

2.1.2 The Provided Instrumental and Executive Helps

Help was only provided after a failure on the initial question of each exercise.

Figure 2 shows the general scheme of each exercise.

If the participant accepted to receive help, she/he was invited to select one of the four following options: (1) “Go back to see the lesson in relation with the error I did”; (2) “Read a text describing the error I did”; (3) “Read a text describing the general procedure to solve the exercise”; (4) “Watch a video showing the exercise’s solution”. Then participant had to go to the checking question.

Help #1 to #3 were instrumental helps (Arbreton, 1998) and help #4 was an executive help.

Participant could only choose exactly one help once she/he accepted to get help.

The instrumental help #1 was generally a link to one of the first instructional screencasts. The instrumental help #2 was a simple text identifying the error but not how to solve it. The instrumental help #3 was the correct procedure to answer the question, presented as an algorithm.

The executive help #4 was a screencast of the correct solving procedure. This screencast was without audio commentary.

The errors with their associated help have been selected thanks to our previous exploratory study carried out on 80 students (Sakdavong, Huet & Adreit, 2009): we made a statistical analysis of the most common and recurrent errors in a word processor, and have selected the most difficult exercises and interesting errors. Then we have designed a general scheme of help in two steps: first step to know if participant wanted help; second step to know which type of help she or he would like.

The computer environment device was a software system not for learning how to use word processor formatting and style-sheets but to record and analyse participant’s behavior in order to provide data for the psychologist and automatic feedback for the participant.

This device has been implemented as a multi-agent system (Wooldridge, 2002). Thanks to this choice, we had software elements (agents) able
intrinsically to observe the activity and to produce a
behavior, also to communicate between them. The
system was very dynamic: creating agents during the
learner’s behavior (for example, creating a new
helping agent when an exercise starts) or modifying
in real time the behavior of agents (for example, an
helping agent can change of behavior according to
an evolution of the learner’s profile) were very easy.

We have used the framework JADE
(Bellefemine, Cairo, Trucco and Rimassa, 2004) and
programmed the agents in the Java language.

One of the agents role was to record the learner’s
behavior as a sequence of actions (the activity
graph), we called it “Historical” agent. It records the
activity with respect to the full teaching scenario (for
example, it records which help the participant has
chosen in case of failure in solving a question; it
records the consultation time of help or the
processing time of one exercise), to the word
processor agent (for example, the learner selects a
paragraph then clicks on the shortcut button
“centering the paragraph”).

A second important agent role was the
“Scrutinizer” one, it made it possible to observe and
analyze the activity of learning. These agents are in
charge to identify characteristic behaviors. Such
agents are created specifically for each exercise.
They had a mechanism of subscription which
enabled them to receive from the “Historical” agent
the sequences of actions which they were in charge
to analyze. According to their analysis, they created
“Helper” agents or will communicate with the
existing “Helper” agents. They could also
communicate with the “Profile” agent in charge to
dynamically adapt the profile of the learner.

The “Helper” agents were the most important
ones: they provided the assistance by giving
feedback, proposing and displaying help. The
combination of “Scrutinizer” agents and “Helper”
agents while they gave a typical feedback were
acting as classical work on intelligent tutoring (e.g.,
Wood, Bruner& Ross, 1976) in which the tutor
decide which help providing on the basis of an
analysis of frequently made errors (Puustinen,

In order to allow our team to analyze each
participant results, the device was generating XML
and CSV files containing all records.

2.2 Measures

2.2.1 Prior Knowledge Test

In order to control previous experiences and
familiarity with text processing programs, each
participant was asked to report what were the text
processing programs he/she knew and their
frequency of use. The frequencies of use were very
extended among participants (from many times a
day to sometimes a year).

In addition, they were given a pre-test of seven
multiple choice questions to determine their level of
prior declarative knowledge in principles of text
processing programs (5 answer choices including the
choice “I don’t know” to avoid a random answer
bias). The questions dealt with shared procedures
and functions of text processing programs. They
focused on formatting and using style-sheets with
text processing programs. Each correct answer was
scored by 1 point. An example of a prior knowledge
question was as follows: Justifying text means
(Expected answer: Aligning text on both left and
right sides). Cronbach’s Alpha for the prior
knowledge test was 0.60. The mean proportion of
correct answers of 49.23 % (SD = 26.00) highlighted
a large heterogeneity of knowledge within the
participants group.

2.2.2 Confidence Judgment and Perceived
Utility of Help

After finishing the question and before receiving a
feedback, the participant had to assess how
confident he or she was in the accuracy of the
answer on scale from 0 (not at all confident) to 100
(very confident).

For each consulted type of help, five questions
assessed the perceived utility of the concerned type
of help. The questions were designed on the basis of
existing measures of perceived utility (Davis, 1989;
Venkatesh, Morris, Davis, & Davis, 2003). The
questions were adapted to the learning and problem
solving aspects of the task. Participants rated their
answer on a scale from 0 “not at all” to 100
“totally”. A mean score was computed on the basis
of the five questions.

2.2.3 Actual use of Help and Solving
Behaviours

On-line recording tools were used to get data of the
exercise solving behaviors: accuracy of the answer,
time to decide to accept or refuse help, solve
exercise, solving procedure, chosen type of help
(executive or instrumental.), acceptance or refusal of
help. The recording tools were the agents presented
above. Each student action was recorded and/or
processed by the system. The system was reactive to
these actions and the system’s actions were recorded too.

2.3 Procedure

Participants were tested in groups in a spacious computer room, in order to prevent them seeing what the other students were doing. The duration was approximately 90 min. They performed individually the tasks and they were not allowed to use material other than the computer-based learning environment.

Participants first filled out the template assessing their past experiences with text processing program use and their declarative knowledge about formatting and using style-sheets. Next, the learning step consisted in watching the instructional screencasts about formatting and style sheets. The main part of the experiment was the problem solving step. Participants were instructed to perform three types of exercises implying formatting and use of style sheets. In order to familiarize participants with the problem solving tasks, the exercises device and the types of available help were presented through a screencast. This exercise was designed to promote participants’ familiarity with the exerciser environment. The first exercise required formatting procedure of a given text. The second exercise required to use and change a style sheet to process a given text. Each exercise was made of two equivalent group of questions. Each participant was first getting the subject of the first question, and then was allowed to try to answer to it with the integrated text processing program. When participants indicated that they had finished, the system decided if the exercise was correctly solved according to the grid we have designed. In case of wrong solving, the system proposed participants to use help. Participants were allowed to reject help. Participants accepting help had to make a choice among different types of help. After consulting help, participants dealt with the second question of the current exercise (see figure 2).

At the final step, participants indicated their perceived utility of the environment’s help.

3 RESULTS

3.1 Descriptive Statistics

Table 1 provides descriptive statistics for performance and help use on the overall sample (n=65).

Concerning performance, results showed that in the initial question of exercise 1, a little more than fifty percent of participants succeed. Among those who failed, 44.83% refused help (i.e., they expressed that they do not need or do not want help); 55.17% accepted to select help. Concerning the initial question of the second exercise, most of the participants failed, and among them 76% accepted help, 24% refused help.

Table 1: Descriptive statistics for performance and help use for the initial question of exercises.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td>36</td>
<td>55.4</td>
</tr>
<tr>
<td>Failure</td>
<td>29</td>
<td>44.6</td>
</tr>
<tr>
<td>Help refusal</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>Help selection</td>
<td>16</td>
<td>24.6</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td>15</td>
<td>23.1</td>
</tr>
<tr>
<td>Failure</td>
<td>50</td>
<td>76.9</td>
</tr>
<tr>
<td>Help refusal</td>
<td>12</td>
<td>18.5</td>
</tr>
<tr>
<td>Help selection</td>
<td>38</td>
<td>58.5</td>
</tr>
</tbody>
</table>

Concerning the kind of help selected, nearly all of the participants selected an instrumental help: 14 out 16 for exercise 1 and 35 out 38 for exercise 2. Concerning the kind of error made in the initial question of exercise 1, most of the participants did an error of forgetting a sub-goal (13 participants out of 29); some participants (10 out 29) did an error not identified by the help system (i.e., an error for which no specific help had been elaborated, only general help could be provided) and a minority did a wrong choice of procedure (6 out 29).

Concerning the initial question of exercise 2, most of participants (31 out 50) did a wrong choice of procedure, the others forgot sub-goal (11 out 50) or made an error not identified by the system (8 out 50).

3.2 Efficacy of Help used

In exercise 1, the acceptance or the refusal of help was not related to performance; \( \chi^2 = .144; p >.05. \) Among those who refused help, 69.2% failed again in the checking question; 30.8% succeeded. Among those who accepted help, only 37.5% took benefit from the help by succeeding in the checking question. The other participants, although they selected a help, they did not take benefit from it and failed again in the checking question (62.5%).
Similarly to the previous exercise, in exercise 2, the acceptance or the refusal of help was not related to performance; $\chi^2 = .93; p > .05$. Among those who refused help, 92% failed again in the checking question; 8% succeeded. Among those who accepted help, only 20.6% took benefit from the help by succeeding in the checking question. The other participants, although they selected an help, they did not take benefit from it and failed again in the checking question (79.4%).

Overall, these results showed that the help selection did not lead the majority of learners to succeed in a similar question.

A mean rate of the perceived utility was computed for each learner who consulted help ($N = 45$ participants who used at least once a help in at least one of the two exercises). The learners generally perceived a positive utility of the consulted help ($M = 61.14; SD = 24.68$).

For exercise 2 (checking question), those who succeeded after consulting help displayed a higher perception mean of the utility of the help ($M = 72.71; SD = 18.65$) than those who failed ($M = 57.72; SD = 24.96$), $t(42) = 1.94, p < .05$. Because, the score of perceived utility was assessed after executing all of the exercises we did not compute the relationship between performance at exercise 1 (checking question) and the utility perceived of the help used in all of the exercises.

However, high perceived utility of help was linked to a high confidence-judgment of success to solve the checking exercise $r(45) = .47, p < .01$. In other words, a positive confidence judgment of performance was associated to a high perceived utility of help consulted to perform the checking question.

### 3.3 Individual Differences in Self-regulated Learning

The analysis of individual differences in self-regulation of behavior on the most complex exercise (exercise 2) revealed that:

- 15 participants succeed both the initial and the checking question.
- 4 participants succeeded in the initial question but failed in the checking question.
  - Among the participants who failed in the initial question, 12 participants refused the help and 11 out of 12 failed in the checking question, considered as bad regulated help-seekers.
  - Among the participants who failed in the initial question, 34 participants accepted help and only 7 succeeded in the checking question, showing that they were self-regulated help-seekers. The other 27 were bad regulated help-seekers.

These behavioral patterns showed that the majority of the participants lack of self-regulated skills and consequently cannot appropriately use the helps. Analysis of variance were computed between the three groups who failed (see Table 2) (we excluded one group who was composed of only one participant who failed at the initial question, refused help and succeeded at the checking question) on the mean time duration of the decision to accept or refuse help on the exercise 2. Results were not significant, $F(2,42) = 1.30, p > .05$. The self-regulated group ($N = 7$) did not differ from the other two groups. The five behavioral pattern did not differ on the confidence judgment in the answer at exercise 2 (initial question), $F(4,59) = 1.27, p > .05$ (see Table 2).

### Table 2: Behavioral pattern and descriptive statistics of confidence judgment in the answer at exercise 2.

<table>
<thead>
<tr>
<th>Behavioral pattern</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failed initial q:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Help refusal-failure checking q.</td>
<td>11</td>
<td>56.18</td>
<td>36.42</td>
</tr>
<tr>
<td>Help accepted-failure checking q.</td>
<td>27</td>
<td>54.63</td>
<td>31.33</td>
</tr>
<tr>
<td>Help accepted-success checking q.</td>
<td>7</td>
<td>56.43</td>
<td>26.64</td>
</tr>
<tr>
<td>Success initial q:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success checking q</td>
<td>15</td>
<td>75.33</td>
<td>22.17</td>
</tr>
<tr>
<td>Failure checking q</td>
<td>4</td>
<td>63.25</td>
<td>33.60</td>
</tr>
</tbody>
</table>

Note. q: question

### 4 CONCLUSIONS

The aim of this study was to present the design of the helps provided by a multi-agent system based on a previous analysis of errors. Then, a second aim was to explore how learners behaviour in a device with help and especially are they self-regulated help-seekers?

Results showed that taking charge of one’s learning, even with a computer-based learning environment with relevant helps, is not easy for most of the learners. They lack of accurate monitoring: judgment of confidence scores should be less high when they failed. An important amount of them refused help when needed, and consequently failed. Some of them used help but inefficiently although they selected an instrumental help, which is
considered as an adaptive help for learning. They perceived the utility of the help but did not take benefit from it.

These results are in the line of those found in the literature and highlight the need for helping learners to select help when needed and to be self-regulated help-seeking. In this perspective, the next step of our study will be to conceive a metacognitive guidance. If learners do not use correctly the helps, it is because they have a lack of metacognitive abilities: they do not regulate correctly their learning behavior.

Two types of metacognitive guidance will be proposed (double-blind, 2009):

- A suggested guidance that the learner can accept or refuse
- An imposed guidance if the system identifies a recurrent metacognitive mistake or lack.

This made the principal originality of our approach from the point of view of psychology.

Then:

- The two types of metacognitive guidance will be included in the multi-agent system as new helping agents
- The online course including guidance will be tested over many groups of learners and the multi-agent system will again record all the learners’ behavior into graphs
- Psychologists will analyze the graphs of learners’ behavior in order to check if our metacognitive helps are useful.

REFERENCES


Sakdavong, J.C., Huet, N., Adreit, F.