ReadIT – a developing and testing method for training reading comprehension strategies

O. Kivinen, L. Galli and M.-T. Kaarakainen

Research Unit for the Sociology of Education (RUSE), FI-20014 University of Turku, Finland

By developing a web-based intelligent learning system called ReadIT the RUSE group attempts to integrate the most promising reading comprehension strategies as an automatic part of the learners' reading process. ReadIT collects log file information from students' behaviour (e.g. time use, transitions, used strategies and help-use activities). In the case study, log files, exercise answers and survey data from 15-year-old comprehensive school pupils (N=108) were analysed and ten different performance groups with respective pedagogical decisions were created. The goal is that by using ReadIT both the poor learners can improve their reading comprehension skills and the skilled learners have access to more demanding challenges.

Keywords Reading comprehension strategies; differentiation of learning materials; personalization methods

1. Introduction

Reading comprehension is an elementary component of skilled learning in the current knowledge society. Reading comprehension refers to the understanding of information presented in written form. The level of understanding can be risen via reading comprehension strategies. For instance PISA 2009 results show that awareness of reading comprehension strategies correlated with positive scores in the reading test more than reading widely for pleasure [1].

Skilled readers read frequently and they are capable of reading different kinds of text in appropriate ways. According to their occurrence in the reading process, reading strategies can be classified into before reading, during reading and after reading strategies. When it comes to before reading strategies, it has been found that skilled readers understand their reading goals, skim the text in advance and make predictions and questions about the given topic. Skilled readers typically read the text from front to back, but they can be selective in doing so. Competent readers are also able to make conscious inferences during reading and try to figure out how the text relates to their prior knowledge. Moreover they try to deal with inconsistencies in the text as well as with unfamiliar words and concepts. With regard to after reading strategies, skilled readers typically construct and revise summaries of the text. [2–4.]

Different levels of reading comprehension [e.g. 5, 6] dates back to Barrett's taxonomy [7]. The idea is that readers must first obtain a threshold of competence in literal comprehension before they are able to reach inferential or evaluative understanding of written texts. Literal comprehension refers to straightforward understanding of the text (facts, vocabulary, etc.). Inferential understanding involves more than just literal understanding. It requires readers to interact with the text, to use their prior knowledge while reading and to recognize the relationships that exist among events, characters and objects. In the case of evaluative comprehension, learners need to make judgements – overall or critical – when the text invites them to take their own experiences and values into consideration. [8.]

In our RoSA-lab we are developing a web-based intelligent learning system, i.e. ReadIT, which allows students to improve their reading comprehension skills. ReadIT directs students in predicting, activating prior knowledge, self-questioning, clarifying difficult concepts and note-making. Adapting to student activities, ReadIT guides students' learning enhancing practices by differentiating learning materials and tasks corresponding with each student's current level. The intention is that with the use of ReadIT, both the poor learners can improve their skills and the skilled learners receive demanding challenges that suit their needs.

2. Personalization in intelligent learning environments

The goal of a user-adaptive learning system is via automatic personalization to provide students with appropriate guidance without requiring them to ask for it explicitly. Automatic personalization is sometimes presented as customization. But let us state that it is important to make a difference between these two concepts as concerns the control of the creation of user profiles and the control of the presentation of interface elements to the user. In the customization process, the users partake in the creation of the profile by specifying their preferences manually. In the case of personalization, the user profiles are created and updated automatically by the system. [9, 10.] Personalization solutions lean often on users' preferences or on user profiles based, for instance, on background data such as age, gender, or education level. In order to construct profiles for each user, data is collected by monitoring user activity, for example monitoring browsing behaviour. Nowadays modifiable and dynamic user profiles are preferred. [11–13.] Educational data mining connects statistical methods, machine-learning algorithms and data mining techniques to handle the educational data. It converts raw data from educational systems into useful information that is used for
educational research and practice. Educational data mining utilizes available data mining techniques (e.g. classification, clustering, association-rule mining, sequential mining and text mining), but also visualization techniques, which are not data mining techniques in a strict sense. [14.] Applying mining techniques on web-usage enables the building of recommender systems for e-learning purposes, utilizing web server logs (user browsing history) as input and data mining techniques (like association rules and clustering) to provide recommendations as output [15]. Though personalization based on user preferences, interests and browsing behaviour individualizes learning environments, it cannot promote differentiated learning. The student pool with heterogeneous learning abilities and differences in learning performance needs appropriately differentiated learning materials, tasks and learning goals. To be successful, personalization has to take learners’ current knowledge level into account. Chen et al. have introduced a genetic algorithm for creating personalized learning paths by matching the learner's abilities and the difficulty level of the learning materials to each other; personalized learning path guidance is preferred instead of the freely lead browsing learning mode. [16, 17]

Many intelligent learning systems provide assistance or context-sensitive hints only when students request them. Recognizing the need for help is said to be a metacognitive skill requiring students to be able to monitor their own learning process and understanding. For example, it can be difficult to recognize whether an error is just a slip, caused by poor knowledge, or the result of guessing. [18.] In their study on learners' help-seeking behaviour, Aleven and Koedinger [19] found that students did not use the offered intelligent help facility very often. Therefore they suggest that systems should support students in learning how to utilize hints or guidance during the learning process. The task of the intelligent learning systems is to guide students to seek help in time, and to use it in problem solving. Otherwise the worst option that those who need the most help are the least likely to receive it in time will be realized.

“Gaming the system” means that students attempt “to succeed in an educational environment by exploiting the properties of the system rather than by learning the material and trying to use that knowledge to answer correctly” [20]. In gaming the system, students are for instance asking repeatedly for help trying to get the tutor to give the correct answers, apply systematic guessing, and/or click every multiple-choice checkbox until the tutor identifies a correct answer and allows the student to proceed. [21.] According to Baker et al. [20, 22] gaming the system may lead to poor learning outcomes; students having the most difficulties tend to game most harmfully. By eliminating gaming, it is possible to improve lower-performing students' learning in particular.

According to Klašnja-Milicević et al. [23] the learning styles can be utilized in finding learners' individual traits, weaknesses and strengths, and in creating a basis for differentiated teaching, learning and exercises. However, there is a lot of discussion and criticism concerning the available learning style models [24]. They have for instance said to lead to labelling and to have too rigid conception about students' abilities. Coffield et al. suggest that learning styles are more compatible with self-development instruments than assessment tools. [25.] According to Garner [26], learners' improved knowledge of different perspectives and stages of the learning process more likely promotes self-development than that of the identification with some particular learning style.

There is no proven recipe for the application of learning styles in adaptation [27]. Carmona et al. [28] connected learning styles with the students' knowledge level and their preferences. However, they noticed that the learning styles are not well applicable due to poor measurement instruments. An additional problem is that the students' learning styles are not constantly updated by the actual information about their behaviour. The module used for adaptation is unable to adapt itself. Therefore both the user module and the adaptive module are necessary to build in a way that they adapt themselves constantly in light of new information. From our point of view, traditional learning style models cannot offer apt solutions for intelligent learning environment adaptation. Learning style models are most likely more comport with self-development functions than personalizing. It is important to match the learning materials and various needs of students to each other.

Knowledge tracing procedures have been used to encode the cognitive mastery of skills tutored. The knowledge tracing model is simple and has been used in various intelligent tutoring systems. [29, 30.] In Arroyo’s et al. [30] model addressing students' help requests, optimizing the ‘problem difficulty level’ and timing are key elements. The difficulty level is crucial to the adaptive behaviour of the intelligent learning application. Therefore the learning material, problem difficulty and students' abilities have to be matched. These, together with the information acquired from students' learning behaviour (as transitions) in e-learning context, create a basis for differentiated learning and appropriately guided learning paths.

3. ReadIT

ReadIT is a web-based intelligent learning system to improve students' reading comprehension skills. Figure 1 represents an overview of the system architecture and the types of information collected into the database. ReadIT’s six phases are introduction, previewing, reading, testing, self-evaluation and feedback. In the first phase, learners get short instructions on how to proceed with ReadIT. In the preview phase, ReadIT guides learners to use pre-reading strategies (predicting, activating prior knowledge and self-questioning). By showing the main heading of the text, ReadIT encourages students to predict the contents of the text. After that, it presents the introduction of the text and the subtitles. Pupils are asked to activate their prior knowledge by writing down what they already know about the subject.
In the preview phase, the system also encourages learners to write their questions about the topic. In the reading phase, ReadIT shows the whole text and encourages learners to write their own notes and to clarify difficult words. ReadIT contains a notes-area and includes hyperlinks from potentially difficult words to Wikipedia articles. Having read the given text and written their notes, learners get tasks in ReadIT's test phase. Tasks are given in the form of multiple-choice questions, divided into literal, inferential and evaluative types. In answering the questions, learners can have their notes available, but the original text is not available anymore. Next, in the self-evaluation phase, learners answer a questionnaire concerning their reading habits. In the final phase, ReadIT gives performance feedback to the learners.

ReadIT collects the log file information from students' behaviour (time use, transitions, used strategies, and help-use activities) automatically. The goal of the full-developed ReadIT is to match students' current skills to the text-material and to the task difficulty levels. In its final form, the adaptation module will adapt students' learning process by differentiating learning materials and learning tasks to fit the students' current level. Krüger, Merceron and Wolf [31] state that though learning software are not usually designed for data analysis and mining, there is no doubt that educational data mining requires this kind of data. We have therefore taken this into account from the beginning of the ReadIT's development process.

4. The case study

In our case study, we have tested our preliminary version of ReadIT on Finnish ninth graders. The tests were performed in May 2011. The empirical data consists of log files, exercise answers and survey data from 15-year-old comprehensive school pupils. Altogether 111 pupils participated in the study. Three pupils were identified representing the “gaming the system”-type of behaviour. They turned up to be a significant exception and were removed from the data. Thus the final number of participants is 108. In our study, 61 of the pupils were girls. All the participants spoke Finnish as their native language and the exercise was conducted in Finnish.

The questions in the test phase were ordered into three comprehension levels from literal to evaluative. The literal questions required learners to provide information that was explicitly stated in the text they read. The inferential questions required learners to make appropriate inferences about the text. The evaluative questions required learners to make their own judgements about the text. Almost all pupils (97 %) achieved the literal level. Two thirds (68 %) achieved the inferential level and only a fifth (21 %) the evaluative level.

The pupils were classified by their performance level into three groups on the basis of the test points received (min 0, max 10). Those who achieved points higher than one standard deviation ($\sigma = 1.44$) from the mean ($\bar{x} = 7.74$) were classified as high-performers (17 %) and those who got points lower than one standard deviation from the mean...
belonged to the low-performing group (29%). More than half (54%) of the pupils received values within one standard deviation of the mean. In our classification they belonged to the middle-group. There are differences in the performance level groups in what measures pupils within the different groups used the offered strategies (Table 1). The high-performers utilized most of the offered strategies. Low-performing pupils typically predicted the topic and wrote their own notes and one third of them used the self-questioning strategy and clarified difficult concepts. It should be noted that only 13 percent of the low-performing pupils used their notes while answering the questions.

Table 1 Pupils by their performance level and the type of strategies used.

<table>
<thead>
<tr>
<th>Performance level</th>
<th>N</th>
<th>Prediction</th>
<th>Background knowledge</th>
<th>Self-questioning</th>
<th>Writing notes</th>
<th>Clarifying concepts</th>
<th>Using notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>18</td>
<td>89%</td>
<td>83%</td>
<td>56%</td>
<td>94%</td>
<td>67%</td>
<td>72%</td>
</tr>
<tr>
<td>Middle</td>
<td>59</td>
<td>81%</td>
<td>64%</td>
<td>46%</td>
<td>85%</td>
<td>56%</td>
<td>46%</td>
</tr>
<tr>
<td>Low</td>
<td>31</td>
<td>87%</td>
<td>52%</td>
<td>36%</td>
<td>77%</td>
<td>39%</td>
<td>13%</td>
</tr>
</tbody>
</table>

The average time to go through the exercise was 28.18 minutes the standard deviation being 10.62 minutes (min 4.35, max 62.03). Those who were quicker than one standard deviation from the mean got the classification fast (20%) and those who were slower than one standard deviation were classified as slow (12%). About 68 percent of the pupils (the average group) lay between these two ends. The fact that the majority of slow-performing pupils used all the offered strategies explains why they used a lot of time. Only about a fourth of the fast pupils made their own questions about the topic and clarified difficult concepts in the text. The majority (69%) of the slow pupils used their notes while answering the questions while only 14 percent of the fast pupils did likewise.

Table 2 Pupils by their performance time and the type of strategies used.

<table>
<thead>
<tr>
<th>Performance time</th>
<th>N</th>
<th>Prediction</th>
<th>Background knowledge</th>
<th>Self-questioning</th>
<th>Writing notes</th>
<th>Clarifying concepts</th>
<th>Using notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast</td>
<td>22</td>
<td>82%</td>
<td>50%</td>
<td>27%</td>
<td>64%</td>
<td>23%</td>
<td>14%</td>
</tr>
<tr>
<td>Average</td>
<td>73</td>
<td>84%</td>
<td>64%</td>
<td>44%</td>
<td>88%</td>
<td>59%</td>
<td>44%</td>
</tr>
<tr>
<td>Slow</td>
<td>13</td>
<td>92%</td>
<td>85%</td>
<td>77%</td>
<td>100%</td>
<td>69%</td>
<td>69%</td>
</tr>
</tbody>
</table>

When it comes to transitions, ReadIT records the time point and the frequency of the transitions that users make within the system. Table 3 describes transitions by the level and the time of performance. Variance analysis (ANOVA) shows that differences between the performance level groups were statistically significant (F 5.909, df 2, p= .004). The result indicates that the high-performers utilized the offered strategies and other elements in the ReadIT-system more than two other groups. For this particular analysis, we had to remove three pupils from the data, because the transition numbers deviated from the normal distribution assumption. High-performers used plenty of strategies and produced a great deal of written notes, which increased the number of transitions in their log files. However, the performance level groups still remain quite heterogeneous. For example, in the high-performing group, there are pupils who accomplish the process very fast and contrariwise pupils who demand more time to go through the system.

Because of the heterogeneity of the groups, we analysed transitions also by used time. Levene’s test indicated that due to the inequality of variances in our samples, we had to use Kruskall-Wallis analysis of variances by ranks. Differences between the groups divided by used time were statistically significant (p= .000). Slow pupils used a lot of strategies writing long notes and in many cases double-checked their answers, which increases their transitions. Fast and high-performing students have not used the offered strategies as much, resulting in a smaller amount of transitions. It is important to notice that in the fast group there are also low-performers to whom the material has been so difficult that they have not been able to understand it from the very beginning, and that is seen as a small amount of transitions.

Table 3 Transitions by the three performance levels and the length of performance.

<table>
<thead>
<tr>
<th>Performance level</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>16</td>
<td>101.31</td>
<td>26.934</td>
</tr>
<tr>
<td>Middle</td>
<td>58</td>
<td>76.00</td>
<td>35.184</td>
</tr>
<tr>
<td>Low</td>
<td>31</td>
<td>66.90</td>
<td>30.518</td>
</tr>
<tr>
<td>Total</td>
<td>105</td>
<td>77.17</td>
<td>34.258</td>
</tr>
</tbody>
</table>

By demonstrating three tentative cases of the significance of adaptation below, we also describe three types of distinct learning paths. Figure 2 presents an example, where the learning material is too easy, which means that learners require upwards differentiation. Figure 3 represents a situation in which the challenge level is optimal to the pupil. Figure 4 shows a situation where the material is too difficult and downwards differentiation is needed.
The case where the material was too easy (Figure 2) describes how the pupil moved linearly through the system, read the text quite fast and did not clarify any concepts. He wrote notes, but did not use them in the test phase. He spent 16 minutes on ReadIT and got 9.5 points from the test phase. By looking at the transitions he made, we can conclude that the material and tasks were too easy. The difficulty level of the learning material for this kind of pupils must of course be raised.

Figure 2  Transitions in a situation where the learning material is too easy.

Figure 3 represents a situation of optimal difficulty level. The pupil read the text from front to back, backtracking and rereading several pages. He clarified five concepts, looked at the given examples and wrote notes about the text, concepts and examples. He used his notes in the test phase and checked his own answers before moving to the self-evaluation phase. He spent 23 minutes in the ReadIT-system and got 8 points. In this case, the pupil utilized appropriate strategies. From his performance points and learning process, we could draw the conclusion that the demand level was suitable for him, in other words, ReadIT offered appropriate challenges to this particular pupil.

Figure 3  Transitions in a situation where the difficulty level is optimal.

Figure 4 shows the transitions of a low-performing pupil. She read the text once from front to back, reread some pages and clarified three difficult concepts. After that, she reread two pages, looked carefully at an example and wrote long notes about it. Almost all the notes she wrote concerned this single example. Then she reread pages 3–8. She did not, however, use her notes during the test phase. She received a low score of 4.5 points and her performance time was ©FORMATEX 2011
31 minutes. In this case, the text-material was most likely too challenging, which led to low performance, backtracking and rereading actions and to writing practically all her notes from a single example. The difficulty level of the learning material for pupils like her must be lowered. Overall these three case examples demonstrate the different needs of the personalization features.

One central goal in this preliminary study is to find out the initial performance groups to create a basis for adaptation in order to improve pupils outperforming. Table 4 demonstrates ten performance groups and the respective pedagogical decisions for the adaptation module to adjust ReadIT-system functions to different pupils and their needs. The performance points do not define differences between employed strategies alone. Another meaningful factor is the use of time. Each level performance varied according to the time that different pupils needed to go through the process. That is why we define groups both by performance time and performance points. This classification creates the basis for ReadIT's adaptation module in the forthcoming stages of development and research. More examination of the sequences of transitions is still inevitable for ensuring the proper actions of personalization functions.

Table 4  Ten performance groups and the respective pedagogical decisions.

<table>
<thead>
<tr>
<th>Time</th>
<th>Performance</th>
<th>Performance group</th>
<th>Pedagogical decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High-performers:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>fast</td>
<td>Raise the demand level</td>
</tr>
<tr>
<td>t &lt;= x - 2σ</td>
<td>p &gt;= x + σ**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x + σ &lt; t &lt; x - σ</td>
<td>p &gt;= x + σ</td>
<td>average</td>
<td>Raise the demand level</td>
</tr>
<tr>
<td>t &gt;= x + σ</td>
<td>p &gt;= x + σ</td>
<td>slow</td>
<td>Raise the demand level and eliminate unnecessary help-use and double-checking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Middle-group:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>fast</td>
<td>Retain the demand level</td>
</tr>
<tr>
<td>t &lt;= x - 2σ</td>
<td>x - 2σ &lt; p &lt;= x + σ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x + σ &lt; t &lt; x - σ</td>
<td>x - σ &lt; p &lt;= x + σ</td>
<td>average</td>
<td>Retain the demand level</td>
</tr>
<tr>
<td>t &gt;= x + σ</td>
<td>x - σ &lt; p &lt;= x + σ</td>
<td>slow</td>
<td>Retain the demand level and eliminate double-checking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-performers:</td>
<td></td>
</tr>
<tr>
<td>x - 2σ &lt; t &lt;= x - σ</td>
<td>x - 2σ &lt; p &lt;= x - σ</td>
<td>fast</td>
<td>Lower the demand level, give instructions about strategies and require to utilize them and to answer the questions</td>
</tr>
<tr>
<td>x + σ &lt; t &lt; x - σ</td>
<td>p &lt;= x - σ</td>
<td>average</td>
<td>Lower the demand level and give instructions about strategies</td>
</tr>
<tr>
<td>t &gt;= x - σ</td>
<td>p &lt;= x - σ</td>
<td>slow</td>
<td>Lower the demand level and give instructions about strategies</td>
</tr>
<tr>
<td>t &lt; x - 2σ</td>
<td>p &lt; x - 2σ</td>
<td>gaming the system</td>
<td>Lower the demand level, eliminate gaming, give instructions about strategies and require to utilize them and to answer the questions</td>
</tr>
</tbody>
</table>

* t = time used (preview phase + process phase + test phase), ** p = performance points
5. Conclusions

In our case study high-performing pupils used reading comprehension strategies more than low-performing pupils. According to Jeong et al. [32] high-performing students exhibit patterns of behaviours that indicate good learning strategies. It is consequently essential to provide instruction for lower performing learners in the strategy-use.

In their studies Niezetal et al. [33] have defined six different types of transitions: linear, jumping, retrying, searching, transitioning, and backtracking. Students in the different performance groups employ different learning behavioural patterns when using intelligent learning environments. Jeong et al. [34] found in their study that high-performing students moved more linearly through the system than low-performers. Floundering, typical to low-performers, indicated certain backtracking actions. In our case study, higher performing students made more transitions than lower performers. This however indicates the use of the offered strategies and other ReadIT-system functions and was characteristic to successful learning. High-performing pupils used several strategies and wrote long notes, which increased their transitions in the log files. On the other hand, a great amount of transitions could indicate backtracking actions and floundering as well. In order to find out which transition cases indicate effective learning and which cases are related to problems in the learning process, transitions must be analysed more closely.

According to Baker et al. [22] gaming the system is the second frequent off-task behaviour after off-task conversation. Students who game the system perform poorer than average students. When it comes to low-performers, those who did not game the system learned more than those who gamed the system. [22.] Detecting gaming is an important element of the adaptation module in particular when it comes to low-performing students. Our preliminary case study indicates that it is likely that boys especially will be motivated to learn with systems like ReadIT. In fact against general expectations [e.g. 1], boys outperformed girls in our study in both performance points and time. Of course this preliminary but interesting result needs more analysing.

Students utilize the system in different ways. Different behaviour within the system expresses learners' goals, needs and preferences. According to Ben-Zadok et al. [35] many cognitive and meta-cognitive aspects of learning could be reflected by the traces they leave in the log files. The understanding of various paths enhances our knowledge of the online learning process. Automatic personalization of the learning material based on learners' behavioural patterns makes it possible to give individualized instructions and feedback to every pupil. Our adaptation module bases at the pupils' performance level and performance time within the ReadIT-system. This is still just the first part of the personalization. The second part will lean on the perceived difficulties of the material and tasks. The final goal is to match these in a way that promotes every pupil's learning. As it has been highlighted, Arroyo et al. [30], for example, have contributed beneficial work with examining problem difficulties. In light of this, our next goal is to further develop the system particularly related to the perceived difficulties; the above-mentioned matching principles are also on the next agenda.

References


