

Investigating the Relationship between Spatial Ability and Feedback Style in ITSs

Nancy Milik¹, Antonija Mitrovic¹ and Michael Grimley²

¹Intelligent Computer Tutoring Group, Department of Computer Science

²School of Educational Studies and Human Development

University of Canterbury, Christchurch, New Zealand

Nancy.milik@gmail.com

{tanja.mitrovic, michael.grimley}@canterbury.ac.nz

Abstract. Rapid and widespread development of computerised learning tools have proven the need for further exploration of the learners' personal characteristics in order to maximise the use of the current technology. In particular, this paper looks at the potential of accounting for spatial ability in ERM-Tutor; a constraint-based tutor that teaches logical database design. Our evaluation study shows no conclusive results to support a difference in effectiveness of the textual versus multimedia feedback presentation modes with respect to the students' spatial ability. However, we observed a number of trends indicating that matching the instruction presentation mode towards the students' spatial ability influences their perception of the system and motivation to use it, more than their learning gain.

1. Introduction

Intelligent Tutoring Systems (ITSs) are effective learning tools due to the adaptive pedagogical assistance they provide. They make decisions about the timing and content of teaching actions and feedback to each student based on their individual state. Students differ in their strategies, approaches, and capabilities for learning and processing cognitive information. Although it is evident that such personal characteristics play a vital role in the learning process, only a small number of studies have investigated the effects of accounting for them in ITSs. For example, Conati et al. [1] use the Five Factor personality traits (openness to experience, conscientiousness, extraversion, agreeableness and neuroticism) in representing different personality types and goal priority in a Dynamic Bayesian Network. This network is then used to maintain an assessment of the student's current emotional state. In contrast, EDUCE [2] uses the Multiple Intelligence learning characteristics (logical/mathematical, verbal/linguistic, visual/spatial and musical/rhythmic) in order to provide a customized learning path.

In this paper, we describe a project which focuses on the spatial ability, a psychometric construct [3] essential to activities related to spatial reasoning, such as the ability to manipulate images or spatial patterns into other arrangements [4]. Learners with high spatial abilities perform better with graphic or spatially-oriented

content than those with low spatial ability. It is worth noting, however, that a low spatial ability score is not a deficit; there is evidence that it can be improved through training and practice [5, 6]. Nevertheless, changing ITSs to accommodate low spatial ability learners could be more practical and beneficial for the system/domain's problem-solving task. That is, learners with different spatial abilities should receive different types of content.

This paper presents an approach to support learners' spatial ability in ERM-Tutor [7], a constraint-based ITS [8] that teaches logical database design (i.e. the algorithm for mapping conceptual to logical database schemas). We start by presenting some relevant work in Section 2. Section 3 gives an overview of the tests we used to measure spatial ability, while the following section presents ERM-Tutor and the modifications made in this project. We then describe the preliminary study and the results obtained in Section 5, followed by conclusions and future work in the final section.

2 Related Work

Personal characteristics are a major factor in learning. Many theories exist regarding how individuals process and encode information differently, such as Richard Mayer's theory of multimedia learning [9-13]. Mayer defines multimedia as the presentation of material using both words and pictures, and proposes that presenting verbal explanations alone in instructional situations is less conducive to learning for some students than presenting verbal explanations in conjunction with pictures [9]. Subsequently, he defines a multimedia instructional message as communication that makes use of our dual learning channel [14, 15] which is intended to foster learning.

Figure 1 shows a representation of the dual channel theory. One channel is dedicated to processing words, whether printed or spoken, and the other is for processing pictorial forms. Based on this assumption, along with the assumptions that each channel has a limited capacity and require active processing, Mayer defines the *Cognitive Theory of Multimedia Learning* [10]. The theory states that learning occurs when learners attend to relevant incoming information (sensory memory), select and organise important information and integrate it with their prior knowledge (working memory) into mental representations (long-term memory). Mayer argues that making use of both visual and auditory channels when presenting learning instructions aids in deep, or meaningful, learning, indicated by good retention and transfer performance. His rationale is that when presenting a message combining an image and text, the information is effectively being perceived and processed twice (once through each channel). Moreover, the words and pictures complement each other, aiding the learner to mentally encode and integrate the information.

It is evident, however, that learners have different cognitive styles and abilities. Some people learn better with visual methods of instruction, whereas others learn better with verbal methods of instruction. The question that arises is whether presenting the same instructional information, based the multimedia theory, is beneficial for both groups of people? Or does it overload the mental processing of some people or even confuse them? More importantly, if learners process information

differently, then how can an instructional environment be tailored to better suit their individual needs? Is it actually beneficial to customise digital instructional environments?

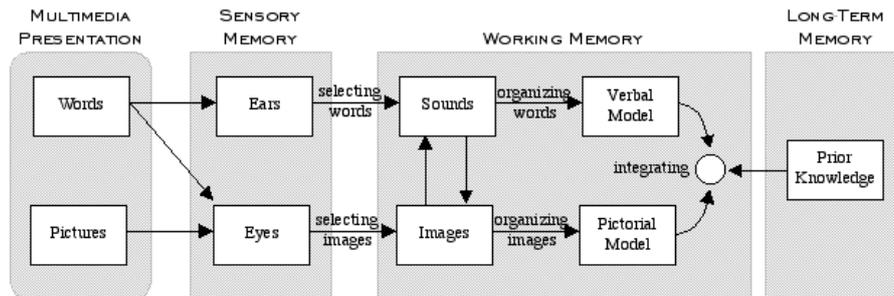


Fig. 1. Information processing via dual learning channels (Figure 3.2 from [9])

These are also some of the questions that Mayer considered. As the result, he documented a number of principles for designers of instructional environments to follow in order to make the maximum use of the learners' dual channels. For example, the *coherence principle* states that students learn better when extraneous words, pictures and sounds are excluded rather than included (i.e., presentations should be as clear and concise as possible to minimise mental processing overload). The principle that is of most interest to us however, is the *individual differences principle*, which states that “[multimedia] design effects are stronger for low-knowledge learners than for high-knowledge learners and for high spatial learners rather than from low spatial learners” (p. 161) [10]. This is because high-knowledge learners are able to use their prior knowledge to compensate for the cognitive processing needed to integrate the information received by the dual-channel. On the other hand, low-spatial learners must devote so much cognitive capacity to mentally integrate the information. Therefore, it is the combination of the learners' spatial ability and level of knowledge that influences their meaningful/deep learning.

3. Measuring Spatial Ability

Spatial ability is important in multimedia learning [10], as the learner needs to encode spatial information from sensory memory, maintain an internal representation in working memory, and perform spatial transformations in order to integrate the information in long term memory. There has been an interest in finding a correlation between individuals' spatial ability level and their gender and age. Studies investigating such correlation, for example testing spatial memory and spatial navigation through a novel environment, showed a male advantage for spatial performance, suggesting that spatial ability is one of the most reliable of all cognitive gender differences in humans [16], as well as an age related decline in performance [17].

The learners' spatial ability and the type of content representation directly affect the learners' cognitive load, level of concentration and motivation. Steinke et al. [18] investigated the usage of 3D models in a hypermedia learning system on plant and animal cell biology. They found that participants with high spatial ability levels spent more time on task-relevant content, whereas those with low spatial abilities spent more time with the 3D models. Low spatial ability participants experience more difficulties in using 3D models and are easier distracted from task-relevant content. More interestingly, high spatial ability participants had a more positive attitude towards 3D content, thus confirming that a high subjective involvement results in a positive influence on the knowledge gain.

Psychometric tests (such as [19-21]) used for determining spatial ability typically consist of paper-and-pencil tasks requiring inspecting, imagining or mentally transforming shapes or objects at the *figural* scale of space [22]. These tests do not provide a discrete value on the spatial ability scale, but rather a relative position within a sample group that determines high or low classifications. We explored short versions of two tests from the battery of cognitive tests developed by Ekstrom, French, and Harman [19]: a ten-item Paper Folding Test intended to evaluate a component of spatial ability called visualisation, and an eighty-item mental Card Rotation Test which evaluates spatial orientation. Each test has a three-minute time limit and is suitable for ages 13-18.

4. ERM-Tutor

Constraint-based tutors enhance learning in a variety of domains, such as database querying (SQL-Tutor [23]), database design (ER-Tutor [24]) and data normalization (NORMIT [25]). ERM-Tutor [7] is a tutor in which students practice the algorithm for mapping conceptual database schemas (i.e. ER diagrams) into relational schemas. Each step in the algorithm maps one ER concept by either creating a new relation or altering previously created relations by adding foreign keys and attributes [26].

The interface (Figure 2) enables students to view problems, work on their solutions and receive feedback. The problem-solving area is the main part of the page, and its general layout is the same for all steps. First, there is a short description of the student's task for that step. For example, for step two the task text reads "*Map all the weak entity types*". This is basically to remind the student what is required in this step, rather than be educational material in its own right. The problem is presented to the student as an ER diagram, but the student also has an option of seeing a textual description of the database, by clicking the *Problem Text* button. Underneath the diagram, brief instructions on what is expected in this step and how to use the input boxes to create or alter a table are presented. At any time, the student can view the solution developed so far by clicking the *Completed Tables* button. This pops up a window containing all the relations defined by the student.

The student creates or alters one relation at a time. Each step of the algorithm is broken into subtasks. For example, in step one, the student maps one regular entity type at a time, and the system checks the resulting relation before moving on to the next entity type. Figure 2 illustrates a situation when the student has mapped the

MEETING weak entity type, and has specified a relation (with the same name) with three attributes (*timing*, *id* and *description*). For each attribute, the student can specify whether it is a primary and/or foreign key. When the student completes the relation, he/she can request the system to check the solution. If there are any mistakes in the solution, ERM-Tutor provides feedback. In Figure 2, the system informs the student that there are some missing attributes representing foreign keys from the owners of the MEETING weak entity type. If the solution is correct, the student can move on to the next entity type, or to the following step of the algorithm.

The feedback/help area occupies the top right side of the screen. A help page is displayed by default when a task page is first displayed. This provides a textual description of how to use the interface. When the student submits a solution, the help page is replaced by the appropriate feedback. Based on the feedback level chosen (*hint*, *explanation*, *list all errors*, or *full solution*), the system informs the student whether their solution was correct or not, and provides hints for correcting the errors. The help page can be redisplayed any time by clicking on the *Help* button.

Influenced by Mayer's work, we created a new version of the system. The original ERM-Tutor only provides text-based feedback. Following the multimedia learning theory, we decided to incorporate a pictorial aspect in the messages; for each feedback message, we created a graphically annotated version.

The screenshot shows the ERM-Tutor interface. At the top, there is a navigation bar with buttons for 'Problem Text', 'Completed Tables', 'Change Problem', 'Help', and 'Logout'. Below this is a progress indicator showing steps 1 through 7, with Step 2 highlighted. The main content area is titled 'Step: 2. Map all the weak entity types'. It features an ER diagram with entities LECTURER, STUDENT, and MEETING. LECTURER and STUDENT are connected to MEETING via a relationship named MEETS. The MEETING entity has attributes Description and Timing. Below the diagram, there are instructions: 'Choose the entity you want to map, then specify each attribute you want to add to the table you have created, use the checkboxes if the attribute is a key or foreign key.' There is a 'Table attribute:' section with checkboxes for 'Key' and 'Foreign Key', and an 'Add attribute' button. The 'Current table:' section shows a table with columns 'id', 'timing', and 'description', and rows for 'Edit' and 'Delete' operations. The 'Relationship:' section shows 'MEETS' with a 'Delete relationship' button. At the bottom, there is a 'Feedback:' section with a dropdown menu set to 'List All Errors' and buttons for 'Check table' and 'Clear'. On the right side, there is a 'Feedback' area with a message: 'For this step you need to specify all the foreign keys from the owner entities.' Below this is a 'Questions' section with a text input field and a 'Submit question' button.

Fig. 2. Screenshot of ERM-Tutor

Being a constraint-based tutor, each feedback message in ERM-Tutor is associated with a constraint. In other words, each constraint has a feedback message which is displayed when the constraint is violated. Consequently, each message provides a hint on how to satisfy its particular constraint. To make the original and the newly created messages comparable, we kept the text identical in both versions. The only difference is the addition of a pictorial representation in the new version. Figure 3 shows the multimedia (text and picture) version of the second feedback message given in Figure 2. A total of 112 images were created, each corresponding to a single feedback message. In addition, ERM-Tutor was modified to cater for both versions of feedback and prepared for an evaluation study described in the following section.

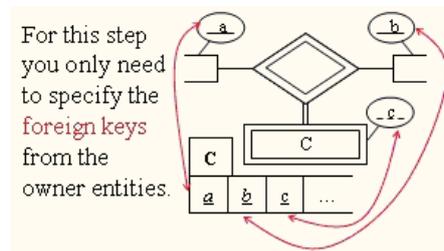


Fig. 3. An example feedback message in the multimedia form

5. Evaluation Study

We performed an evaluation study with students enrolled in an introductory database course at the University of Canterbury in March 2007. Our hypothesis is that students with a high spatial ability level will benefit more from multimedia feedback than students with a low spatial ability, given the same background knowledge. As each student's spatial ability level (either high or low, as opposed to the actual value) is determined relatively to the sample group, we conducted a preliminary study in the previous year to calculate the students' spatial ability median score and we made the assumption that the students' samples from both years will be comparable. This score was used as the threshold in classifying students as having high/low spatial abilities.

Each participant was allocated to one of the versions of the system, providing either textual or multimedia feedback. The experiment allows for a 2x2 comparison: textual messages for high (TH) and low spatial ability students (TL), and multimedia messages for high (MH) and low spatial ability students (ML).

The study was conducted in two sessions of scheduled labs on ER mapping, straight after students had attended lectures on the topic. Each participant attended one of the sessions, and worked with ERM-Tutor individually, solving problems at their own pace. At the start of a session, the students were given an information sheet describing the study, a consent form, and a pre-test on paper (with a maximal score of eight) consisting of four multichoice questions and a mapping question. To make the results of the pre-test and post-test comparable, two tests were used; students in the

first session used version A as the pre-test and version B as the post-test and students in the second session used the reverse.

When a student logged onto the system, they were presented with a set of instructions explaining the two spatial ability tests, and a sample problem. Additionally, for each test, they were asked to rate own ability on a scale of 1 to 5 before sitting the tests. They had three minutes to solve the problems in each test. Once the spatial tests were completed, or their time was up, the students were allocated to the appropriate version of the system and were asked to use it, solving as many problems as they would like. At the end of the session, students were asked to fill in a post-test and a questionnaire about the system. Finally, the students were encouraged to use the system at any time until the end of the course.

A total of 43 students submitted both pre and post tests. The mean score for all students on the pre-test was 4.3 (sd=2.2) and post-test was 5.2 (sd=2.2), resulting in a significant improvement in their performance ($t=3.4$, $p<0.001$). The scores for the different groupings of students are given in Table 1. The analysis indicated that there was statistically significant difference between the students' performance in the pre- and post- tests by those in the HT ($t=3.4$, $p<0.005$), LM ($t=2.0$, $p<0.05$) and HM ($t=1.8$, $p=0.0553$) groups. However, there was no significant difference for the LT ($t=0.2$, $p=0.4365$) group. A closer look at the LT group shows that its students have a higher pre-test score, with a mean of 5.8 (sd =1.1), and hence they improved the least in comparison with the other groups, scoring means of HT: 4.2 (2.4), LM: 3.6 (2.1) and HM: 4.2 (2.3). Although we hoped for an ideal setting of comparable groups, this imbalance in prior knowledge between the four groups was unavoidable.

Furthermore, ANOVA analyses across the pre and post- tests of the four groups did not yield any significant difference, indicating that all groups improved in a similar manner regardless of the feedback mode presented or their spatial ability. This suggests that presentation modes have similar influence on performance regardless of the spatial ability; that is, all students, whether low or high spatial, improved in performance regardless of the feedback mode they were given. We suspect however, that although ANOVA analysis on the pre-test did not indicate significant difference, the higher pre-test in the LT group has an influence on these statistical tests.

Table 1. Mean (*sd*) pre-test and post-test scores for all classifications of students

Classification	t-test: Paired Two Sample for Means				
	No.	Pre-test	Post-test	t-stat	P-value
LT	8	5.8 (1.1)	5.9 (2.5)	-0.2	0.4365
HT	12	4.2 (2.4)	5.7 (1.8)	-3.4	0.0029
LM	12	3.6 (2.1)	4.5 (2.1)	-2.0	0.0385
HM	11	4.2 (2.3)	5.1 (2.5)	-1.8	0.0553
Textual (LT HT)	20	4.8 (2.2)	5.8 (2.0)	-2.2	0.0185
Multimedia (LM HM)	23	3.9 (2.2)	4.9 (2.2)	-2.7	0.0070
Low (LT LM)	20	4.5 (2.1)	5.0 (2.2)	-1.4	0.0840
High (HT HM)	23	4.2 (2.3)	5.4 (2.2)	-3.6	0.0008
Matched (LT HM)	19	4.8 (2.2)	5.4 (2.5)	-1.4	0.0954
Unmatched (LM HT)	23	3.9 (2.2)	5.1 (2.0)	-3.8	0.0005

The tests for the rest of the groupings show either a statistically or marginally significant difference in students' performance between the pre- and post- tests scores. The p-values produced show that unmatched groupings scored a higher significant confidence than the matched groupings. A closer look at the figures show that the matched groupings had a higher pre-test mean score of 4.8 (sd = 2.2) than the unmatched groupings (mean=3.9,sd=2.2). This difference was verified as marginally significant using a two-sample assuming unequal variances t-test ($t=1.5$, $p=0.0759$). A further test within the matched groupings, comparing the pre-test scores between the LT and HM groups indicated a significant difference ($t=2.0$, $p<0.05$).

We did find, however, an interesting trend in the data after analysing the students' log files and their interaction with the system. We looked at the total time interacting with the system, number of attempted problems, number of solved problems, percentage of solved problems and the total number of attempts/student solutions submitted (Table 2). We found that the HM group had a consistently higher mean for all these types of interactions, followed by the HT group, then the LT group and lastly the LM group with the lowest mean. The four groups came out in the same order for all the types of interactions we examined. Although the difference in numbers is quite small and statistically insignificant, this trend is in line with Mayer's theory that high spatial students will benefit more from multimedia presentation. In other words, the high spatial students appreciated and used the system more when they received multimedia feedback messages, whereas the low spatial students were less inclined to use the system when they received the multimedia feedback messages.

Table 2. Summary of means (sd) of system interaction results

	LT	HT	LM	HM
Total time (min)	62 (27.3)	72.7 (37.3)	57.7 (29.8)	72.2 (31.8)
Attempted problems	6.9 (6.3)	8.8 (5.0)	4.8 (3.0)	9.1 (5.0)
Solved problems	4.4 (5.0)	6.5 (4.6)	3.2 (3.8)	7.2 (4.5)
% solved problems	61.2 (36.8)	69.4 (26.1)	53.4 (43.2)	72.9 (21.3)
Total attempts	118.4 (97.7)	130.1 (73.2)	69.7 (46.6)	148.5 (90.7)

The same trend is reflected by the students' perception of the system indicated by their subjective results. Table 3 shows the mean responses to the 1 to 5 Likert scale questions of the four groups, where 1 represents the most negative response and 5 the most positive response. Again although the difference is not statistically significant, it seems that the LM group consistently reported the lowest ratings for the system, finding it more difficult and less interesting than the other groups. An interpretation of this could be that because the LM group spent more cognitive effort processing the feedback messages and hence enjoyed ERM-Tutor the least.

Table 3. Summary of means (sd) of subjective results for ERM-Tutor

	LT	HT	LM	HM
Overall quality	3.8 (0.7)	3.7 (0.7)	3.0 (1.1)	3.7 (0.8)
Terrible-Wonderful	3.4 (0.9)	3.5 (0.9)	2.9 (0.8)	3.6 (0.7)
Difficult-Easy	3.3 (1.0)	3.4 (0.8)	2.8 (1.1)	3.0 (0.8)
Boring-Fun	3.2 (1.0)	3.4 (0.9)	3.3 (0.8)	3.2 (0.8)

6. Conclusions

This paper looks at the potential of incorporating a multimedia representation of ERM-Tutor's feedback messages and evaluating the impact of various styles of feedback messages on learning in respect to the students' spatial ability levels. In the evaluation study, we presented students with one of the two feedback presentation modes, either textual or multimedia. We analysed the students' performance with respect to their spatial ability level and feedback mode they received. The results indicate that all students improved in their domain knowledge after interacting with ERM-Tutor. However, we did not find statistically significant difference between the pre- and post- tests scores across the four groups (LT, HT, LM and HM). Although we allocated similar number of students to each group, we were unable to control for the students' prior existing knowledge that influence their gain scores between their pre- and post- tests.

We observed a number of trends in the collected data. In particular, there was a tendency for students with high spatial ability who received multimedia feedback to interact the most with the system, and students with low spatial ability to interact the least with the system. Moreover, there was no noticeable difference between students receiving textual feedback regardless of their spatial ability. These findings indicate that, in terms of interactions with the system, the textual feedback had the same effect on students, whereas the multimedia messages had a greater effect on the high spatial students than on the low spatial students. We also note that students in the matched groups enjoyed interacting with the system more and were more motivated than those who were not matched. On the other hand, students in the unmatched groups had a higher learning gain than those who received matched type of feedback, based on the post-test scores. Although our contributions towards accounting for the students' spatial ability lack statistically significant measures, there is evidence that matching the presentation of instruction towards the students spatial ability has an influence on their perception of the system and motivation to use it, more than their learning gain.

Our analyses suggest that although students have a range of spatial ability skills, their preferences could be different than their ability levels. It is therefore, worth further investigating whether students have a differing preference to their capabilities. If this is evident, then we suspect that allowing the students to choose their preferred feedback presentation mode would increase their motivation and influence a positive affective state.

Acknowledgements

The project presented in this paper has been supported by an ICTG MSc Scholarship to the first author. We thank all members of ICTG for their support and help.

References

1. Conati, C., Maclare, H.: Evaluating a Probabilistic Model of Student Affect. ITS 2004, LCNS, Vol. 3220, pp. 55-66, Springer, Heidelberg (2004)

2. Kelly, D., Tangney, B.: Predicting Learning Characteristics in a Multiple Intelligence Based Tutoring System. ITS 2004, LCNS, Vol. 3220, pp. 678-688, Springer, Heidelberg (2004)
3. Jensen, A.R., The G Factor: the Science of Mental Ability. *Psychology*, 10: p. 23 (1999)
4. Carroll, J., Human cognitive abilities. Cambridge University Press (1993)
5. Baenninger, M., Newcombe, N.: The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, 20(5), 327-344 (1989)
6. Vicente, K.J., Williges, R.C.: Accommodating individual differences in searching a hierarchical file system. *Man-Machine Studies*, 29(6), 647-668 (1988)
7. Milik, N., Marshall, M., Mitrovic, A. Responding to free-form student questions in ERM-Tutor. M. ITS 2006, pp. 707-709 (2006)
8. Ohlsson, S. Constraint-based Student Modeling. in *Student Modeling: the Key to Individualized Knowledge-based Instruction*. Berlin: Springer (1994)
9. Mayer, R., Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32, 1-19 (1997)
10. Mayer, R., Multi-media Learning. University of California (2001)
11. Mayer, R., Moreno, R.: Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43-52 (2003)
12. Mayer, R., Sobko, K., Mautone, P.: Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, 95(2), 419-425 (2003)
13. Moreno, R., Mayer, R.: Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91(2), 358-368 (1999)
14. Baddeley, A.D., Working Memory. Oxford University Press (1986)
15. Paivio, A., Mental Representations: A Dual Coding Approach. Oxford University Press (1986).
16. Moffat, S.D., Hampson, E., Hatzipantelis, M.: Navigation in a "virtual" maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution and Human Behavior*, 19(2), 73-87 (1998)
17. Moffat, S.D., Zonderman, A.B., Resnick, S.M.: Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, 22(5), 787-796 (2001)
18. Steinke, M., Huk, T., Floto, C.: The Influence of Cognitive Abilities and the Presence of 3D Models on the Use of Task Relevant Content in Hypermedia Learning Systems. 2004.
19. Ekstrom, R., French, J., Harman, H.: Manual for kit of factor referenced cognitive tests. Princeton (1997)
20. Bodner, G.M., Guay, R.B.: The Purdue Visualization of Rotations Test. *The Chemical Educator*, 2(4), 1-17 (1997)
21. Guay, R.B.: Purdue Spatial Visualization Test: Rotations. Purdue Research Foundation, West Lafayette, IN (1977)
22. Hegarty, M., et al., Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34, 151-176 (2006)
23. Mitrovic, A., A Knowledge-Based Teaching System for SQL. ED-MEDIA 1998. pp. 1027-1032 (1998)
24. Suraweera, P., An Intelligent Tutoring System for Entity Relationship Modelling. *Artificial Intelligence in Education*, 14(3), 375-417 (2004)
25. Mitrovic, A., NORMIT, a Web-enabled tutor for database normalization. Proc. ICCE 2002 pp. 1276-1280 (2002)
26. Elmasri, R. and S.B. Navathe, Fundamentals of database systems. Addison-Wesley Reading, Mass (2000).