A web-based tutoring system with styles-matching strategy for spatial geometric transformation

Hao-Chuan Wang [a,*], Tsai-Yen Li [b,*], Chun-Yen Chang [c,*]

[a] *Institute of Information Science, Academia Sinica, 128, Sec. 2, Academia Rd., Nankang, Taipei 115, Taiwan ROC*
[b] *Department of Computer Science, National Chengchi University, 64, Sec. 2, Zhinan Rd., Wenshan, Taipei 116, Taiwan ROC*
[c] *Department of Earth Sciences, National Taiwan Normal University, 88, Sec. 4, Tingjou Rd., Wenshan, Taipei 116, Taiwan ROC*

Abstract

It has been a major objective for researchers to develop computer systems that can effectively deliver instruction to learners. Therefore, how to incorporate instructional strategies in computer-assisted learning systems in a systematic manner deserves further investigation. In this paper, a style-matching strategy that attempts to match learning materials' styles to learners' latent traits is proposed and realized in a web-based tutoring system, called CooTutor. The mechanism of adaptive material selection takes learners' different spatial ability and learning styles as an integral learning profile into account, and performs traits-based personalization of learning experience. This system is specifically designed to conquer the difficulty of tutoring the topic on fundamental spatial geometry in conventional curriculums. By conducting empirical evaluation with a small group of students, it is found that CooTutor is generally beneficial to learning the domain, but the effect of the styles-matching mechanism remains inconclusive. The work aims to contribute to the community of adaptive hypermedia in providing an explorative example adopting the concern of individual difference for personalization. The system design, a usage scenario, and an exploratory evaluation are presented in this paper as implications for further studies.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Adaptive hypermedia; Web-based Learning; Personalization; Learning style; Spatial ability

* Corresponding authors.
E-mail addresses: haochuan@iis.sinica.edu.tw (H.-C. Wang), li@nccu.edu.tw (T.-Y. Li), changcy@cc.ntnu.edu.tw (C.-Y. Chang).

1. Introduction

In the past decades, researchers have built various types of computer-assisted instruction (CAI) tools to help teachers/learners to instruct/construct domain knowledge in a more effective way. Many examples are in the field of engineering education, such as a system for learning how to use electronic instruments (Huang and Lu, 2003) and an environment for learning computer programming (Daly and Horgan, 2004). It is evident that effective CAI tools would consider and incorporate appropriate instructional strategies for instructing particular domains, such as the strategy of simulating real tasks in the virtual world (e.g. the simulation of how electronic instruments work) and visualizing abstract concepts (e.g. the visualization of computer algorithms). Just as educational media studies have pointed out, without effective instructional methods, it seems pessimistic that digitalized content (even in a fancy multimedia form) presented on the computer screen would positively influence learning outcomes in comparison with conventionally educational settings, e.g. classrooms (Clark, 1994). Though the history of using computers to facilitate learning is long, how to systematically and principledly incorporate adequate instructional strategies into computer-based educational systems remains attractive and challenging to researchers.

In particular, with the tremendous development of the World Wide web (WWW), web-based instruction/learning has gradually become a new type of instruction/learning for teachers and learners. However, the fact is that most web-based learning today is undertaken upon dedicated learning or course management systems (LMS or CMS). The main responsibilities of these systems are to manage and deliver digitalized learning materials to learners. Learning materials, however, are not equivalent to successful learning experiences. Less instructional strategy or educational concern has been addressed at the system level in today's web-based learning paradigm. In other words, how to develop and deploy learning systems with good educational models on the web to disseminate successful learning experience has become a new and important issue worth further investigation.

One trend in contemporary educational practices is to pursue the goal of learner-centered learning. The paradigm of real-world education has been gradually shifting from objectivists' views of learning as knowledge transmission toward constructivist's views of learning as knowing and construction. Although the implementation of a learning environment by the means of pure constructivist approach may be still a great debate in educational practices (Chang and Tsai, 2005), a moderate and hybrid approach, learner-centered learning environment, emphasizing on individualized diagnosis and context-awareness, is generally accepted by the community as a sound alternative in amending traditional pedagogy (Bransford et al., 1999). The connection between the paradigm of learner-centered learning and the notion of personalization in human–computer interaction (HCI) is obvious. Generally speaking, personalization is the design of enabling systems to capture or infer the needs of each individual, and then to satisfy that needs in a given context (Riecken, 2000). For commercial applications, personalization benefits the business in providing customized services efficiently; while for the web-based learning paradigm, the design of personalization can push such systems a step forward, from content delivery systems to learner-centered learning environments. Works in

intelligent tutoring systems (ITS) and adaptive hypermedia (AH) (Brusilovsky, 2000; Brusilovsky, 2001), have been jointly providing a great basis to realize learner-centered learning environments, which can serve as the foundation in improving web-based learning.

The concept of spatial geometric transformation (or SGT for short) is an important foundation for computer graphics and mechanical engineering students. This topic is employed as an example domain in developing and evaluating the proposed system. SGT is about how to represent and compute 3D transformations (e.g. scaling, rotation and translation) of objects in the mathematical coordinate systems. A syllabus survey in computer graphics education shows that geometric transformations including 2D and 3D are the major topics that most computer graphics educators care about (Wolfe, 1999). However, it is observed that SGT learning would require learners to perform fluent spatial reasoning skill. That is, to construct mental images of the spatial configurations and subsequently manipulate these mentally visual imagery to learn spatial geometry topics (Wang et al., 2004). For learners who have insufficient spatial ability, it is inherently difficult for teachers to instruct them in the classroom without suitable instruction tools. Therefore, (1) how to develop suitable tools for instructing/learning SGT, and (2) how to tackle the fact that different learners are with different levels of spatial skills and learning styles, are the two major concerns, in terms of instruction, to be coped with in this research.

By extending the previous work (Wang et al., 2004), this work introduces a web-based tutoring system, called coordinate tutor (CooTutor), with additional integration of the instructional strategy of styles-matching. Researchers of the AH field have noted the potential of using learners' latent traits, e.g. spatial ability and learning styles, as indicators for effective personalization (Stach et al., 2004). But the discussion of considering learning traits in AH is just about to begin.

By considering the characteristics of SGT, CooTutor was initially designed as a visualization tool that can demonstrate SGT concepts with interactive 3D graphics. Along with the ongoing research efforts of AH and ITS, CooTutor aims to personalize its presentation adaptively to fit individual difference, including knowledge status and individual traits (i.e. spatial ability and learning styles). The work reported in the paper specifically focuses on using the styles-matching strategy to accommodate learners' traits.

This research presents the mechanism of adaptive material selection fulfilling the strategy of styles-matching in CooTutor. As a pilot exploration of this issue, an empirical evaluation was conducted to probe the influence of styles-matching on learners' SGT achievement and spatial ability enhancement. A total of 31 graduate-level students with computing majors have participated in this evaluation. Results, analysis, and implications of this evaluation are reported in this paper. From the current small-sample evaluation, CooTutor is generally beneficial to students on SGT learning, but the designed mechanism of material selection did not show a significant improvement comparing to other approach. As an explorative and developmental study on this issue, we have not found ample evidences to draw a dichotomous conclusion upon the effectiveness of this design. However, based on the conception of developmental research (Reeves, 1995), we aim to provide a comprehensive case study, including perspectives of system design, usage

scenario, evaluation methodology, and evaluation result, as an impetus and reference framework for following works.

The rest part of this paper is structured as follows. In Section 2, the theoretical background of learners' latent traits and possible instructional strategies are described. The background serves as the foundation for designing the mechanism of adaptive material selection. In Section 3, the overview of the CooTutor system is described. In Section 4, the mechanism of adaptive material selection will be introduced. Section 5 demonstrates a usage scenario showing how the system functions and interacts with learners. Section 6 presents the empirical evaluation, including the experimental design, results, analysis, and implications. Finally, Section 7 draws conclusions and future works of this research.

2. Learners' traits and instructional strategies

In our proposal, learners' latent traits are employed jointly to form an integral learning profile for each learner. In CooTutor, two major types of learners' traits are considered for adaptation. One is the psychometric construct, spatial visualization ability derived from the instrument of purdue visualization of rotation test (PVRT) proposed by Bodner et al. (Bodner and Guay, 1997); while the other is the Felder–Silverman learning styles model (Felder and Silverman, 1988) derived via the instrument called index of learning styles (ILS) proposed by Soloman and Felder, 2005. The two major traits and their corresponding strategies are described, respectively as follows.

2.1. Spatial visualization ability

Spatial ability is a psychometric construct that is recognized influential to activities related to spatial reasoning such as engineering activities and scientific thoughts (Bodner and Guay, 1997). In this study, we are interested in asking how to make use of spatial ability as a basis to adapt the presentation?

Two assumptions are given as instructional strategies. First, learners with different spatial ability should receive contents with different types of media representations as assistance. Second, the higher a learner's spatial ability is, the less degree of visualization she/he will need. For example, if 2D-based (i.e. texts and diagrams) and 3D-based illustrations (i.e. interactive 3D media) are both available for describing a concept, we could scaffold learners with low spatial ability by adopting 3D visualization. For learners with enough spatial reasoning skills, letting the learner practice to form and manipulate the mental image with abstract 2D-based representation is reasonable. With the design, the theory of cognitive scaffolding is considered for adapting the presentation to learners' spatial skills.

Besides, since several previous works pointing out that suitable spatial experiences, such as activities of engineering drawing or browsing computer-based 3D virtual environment would possibly improve learners' spatial ability (Woolf et al., 2003). This research also touches this point by conducting an empirical evaluation as described in Section 6.

2.2. Learning styles

Learning styles usually refer to different approaches learners would take to learn. A rather clear explanation on learning styles is as "… (learning styles are) strategies, or regular mental behaviors, habitually applied to learning, particularly deliberate educational learning, and built on her/his underlying potentials. (Draper, 2003)" Many endeavors are to identify types of learning style that can classify learners into distinguishable extremes of that type. For example, visual learners who intend to learn with pictorial representations and verbal learners who prefer to percept textual descriptions. Then visual and verbal styles of learning are proposed as two end points of an imagined dimension forming a continuum of visual/verbal learning style. Some other types of learning style, such as field dependent (FD)/field independent (FI) learners and sequential/global learners, have been frequently imported into education as the basis of instructional design.

By considering learning styles for instructional design, designers frequently take the approach of styles-matching to design the instruction. For example, apply more pictures and diagrams in learning materials prepared for visual learners. In brief, the strategy is to adapt the content or structure of the instruction with proper pedagogical/teaching styles to match learners' learning styles. CooTutor adopts the strategy of styles-matching as well. The Felder–Silverman learning styles model used in this research has formulated these pedagogical styles to cope with different type of learning styles (Felder and Silverman, 1988). However, note that some researchers argued that the effect of styles-matching is doubtable (Draper, 2003) while some studies suggested that the teachers should attempt to address all styles equally in the instruction and help learners to adapt themselves to learn in their less preferred modes (Felder et al., 2002).

It is still unclear that which type of instructional strategies would be better taken into computerized tutoring to tackle learning styles (e.g. matching of styles vs. addressing all styles equally). Although in conventional learning situations, such as classroom-based lecturing, it seems that less experimental result can reveal the effect of statistical interaction between matching and mismatching learning styles (Draper, 2003). However, in computer-based learning environments some studies have detected statistical interaction regarding the use of styles-matching (Rosati et al., 1998; Shute, 1993; Triantafillou et al., 2004). The instructional strategy seems potentially beneficial for designing ITS and AH systems. More instances and evaluations are required to address this issue.

3. Overview of the CooTutor system

The coordinate tutor (CooTutor) system is an adaptive web-based tutoring system with interactive 3D media for SGT learning. More details of the design and the underlying considerations of the system can be found in (Wang et al., 2004). In this section, a sketch of the system design is presented from a bird's eye view.

The system architecture of CooTutor is illustrated in Fig. 1. The system could be decomposed into two sides according to the web-based nature, server-side and client-side. At the server-side, several models, including domain, student and tutor models are explicitly
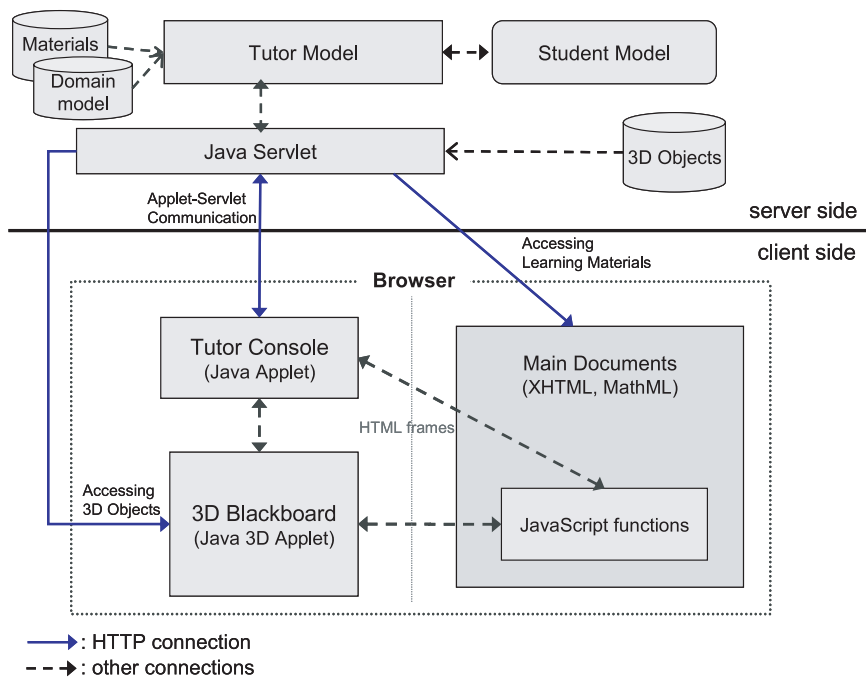
Fig. 1. Overview of the CooTutor system architecture.

designed and incorporated for performing adaptivity. At the client-side, three main parts of the user interface can be identified. They are the part of tutor console, main document area, and the 3D blackboard. The appearance of the user interface is shown in Fig. 2. Among these modules at the client-side, main document is the area used for presenting regular web pages. Tutor console is the module that mainly takes the responsibility of managing the communication between the server and the client. This module is also employed to interact with users to collect learners' information for adaptation.

3.1. Presenting SGT concepts with 3D blackboard

The 3D blackboard module is specifically designed for presenting 3D content, and is extremely useful for tutoring SGT concepts. In the 3D blackboard, learners are allowed to navigate the 3D environment by dragging the mouse to realize the spatial relation between objects and the coordinate system in a clearer manner. This type of 3D visualization could offer learners without sufficient spatial reasoning skills appropriate scaffolding.

Moreover, learners can operate interaction objects, such as buttons and input fields, embedded in main documents to interact with 3D objects shown in the 3D blackboard. By the theory of situated learning (Cognition and Technology Group at Vanderbilt, 1990; Tretiakov et al., 2003), the main document area is used to provide the context of the topics to be learned, and the 3D blackboard offers in-depth exploration of spatial concepts. The interaction between the main document and the 3D blackboard allows learners to explore

3D Blackboard — (label pointing to left panel)

Main Docs — (label pointing to right panel)

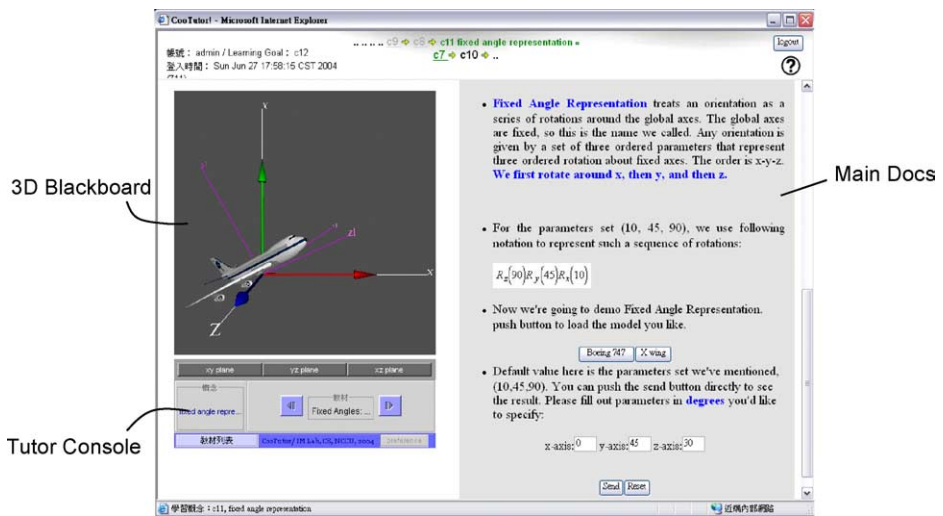**Tutor Console** — (label pointing to bottom panel)

Fig. 2. User interface with 3D blackboard presented at the client-side.

the scene and do experiments freely. This type of user interfaces provides an environment for learner-centered construction.

For example, in Fig. 2, the main document presents the textual description with interaction objects embedded, "Now we're going to demo Fixed Angle Representation. Push button to load the model you like…Please fill out parameters in degrees you'd like to specify…" Learners can read these descriptions, and then do experiments via interacting with these interaction objects to see how different spatial configurations would influence the geometric transformation of 3D objects.

3.2. Adaptivity in CooTutor

CooTutor adopts the course sequencing approach to attain adaptivity (Brusilovsky and Vassileva, 2003; Wang et al., 2004). In brief, the task is to select a set of learning materials for learners according to the student model and learning materials' features. Fig. 3 shows
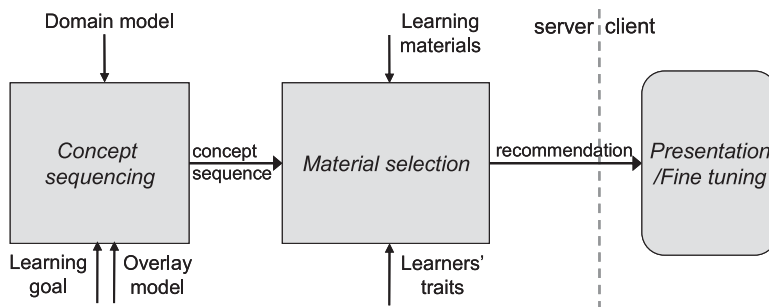

Fig. 3. The flow of adaptivity in CooTutor.

the flow of adaptivity in CooTutor. Three main levels exist: concept sequencing, adaptive material selection and client-side tuning.

Server-side decision making is divided into two levels. At the first level, the mechanism of concept sequencing is responsible for generating a sequence of concepts in accordance with learners' knowledge status. Or in other words, the objective of concept sequencing is to scaffold learners by navigating the knowledge space of the learning domain in a step-by-step means. If learners lack some required prior knowledge to learn a particular concept, they will be directed to learn those concepts first as proper complements. The concept sequencing algorithm used in CooTutor has been described in our previous work (Wang et al., 2004).

Once the sequence of concept has been determined, the next phase, i.e. adaptive material selection, proceeds. The task is to select learning materials from the content repository to illustrate the concepts to learners. Evidently, the styles-matching strategy could be applied in this phase to take learners' learning profiles which characterize learners' latent traits into consideration. This point is the main focus of this paper and will be described in next section.

4. Adaptive material selection in CooTutor

For each concept selected by the concept sequencing algorithm, various styles of learning materials could be possibly authored for illustrating the same concept. On the other hand, learners' latent traits, including spatial ability and learning styles, are modeled in the student model. Therefore, the mechanism of adaptive material selection is responsible for taking materials' pedagogical styles and learners' traits as inputs and performing the styles-matching strategy.

In CooTutor, adaptive material selection is abstracted as a task of information retrieval (IR). Importing the notion of IR, the information of learners' traits is used as the query, and each learning material's styles are employed as the materials' features. The styles-matching strategy can then be realized by computing the similarity between the query and the associated feature vector of learning materials (Baeze-Yates and Ribeiro-Neto, 1999). Note that the technique used in this research is still different from typical IR. While typical IR usually aims to retrieve documents at the level of keyword or phrases matching. In the case of styles-matching, information of latent styles associating with learners and materials is the main concern. Summarily, this type of IR is to retrieve documents at the level of pedagogical styles. In this section, features used to index learning materials, the process of query formulation, and the task of computing similarity are introduced, respectively.

4.1. Features of learning materials

In the SGT domain, learning materials are usually with different degree of abstraction (e.g. abstract mathematical descriptions vs. concrete and practical examples), different modalities of media representation (e.g. 2D-based web pages vs. interactive 3D visualization) and different types of learning activities (e.g. lecture vs. experiment).

Features	Values (numeric between 0 and 1)
Main representation	2D-based / 3D-based
Abstractness	Concrete / Abstract
Activity type	Lecture / Experiment
Level of details	(unidirectional)

Fig. 4. Features of learning materials.

Taking these features described above altogether, learning materials themselves are actually with various pedagogical styles. Fig. 4 shows features used to index learning materials.

Note that the features of main representation, abstractness, and activity type are bidirectional. For example, a research paper with plenty mathematical descriptions is thought to be more abstract. Then the feature of abstractness is assigned a higher value, such as 0.8. Similarly for concrete and practical learning materials, a lower value is assigned for the same feature. That is, only one field is employed to record such a bidirectional feature. This is viable since these bidirectional features are deemed to be continuous between the two ends of the dimension. The assignment of these values is done by the author and content provider of the learning materials.

4.2. Query formulation from learners' traits

As mentioned previously in Section 2, the instruments of PVRT and ILS are used to assess learners' spatial ability and learning styles, respectively. The ILS questionnaire consists of 40 question items. Four dimensions of learning styles can be assessed by the questionnaire. Each dimension of learning style is measured by 10 items evenly. The four dimensions are visual/verbal learning style, sequential/global learning style, sensing/intuitive learning style, and active/reflective learning style. Since this research intends to use spatial ability as features for adapting media representations that would also address the visual/verbal concern, items of the questionnaire targeting to measure the visual/verbal dimension of learning style are not used. Similarly, since it seems not viable to address the sequential/global learning style at the level of learning materials, items for this dimension of learning style are not employed either. After all, CooTutor assesses learners' spatial ability, sensing/intuitive learning style, and active/reflective learning style to form an integral learning profile for material selection.

The information of learners' traits is used to trigger the query of pedagogical styles. The process of query formulation is shown schematically in Fig. 5. By using learners' quantitative scores of spatial ability and learning style as external inputs, with default assumptions, the query of pedagogical styles is derived as the result. The default assumptions used in this case are enumerated here:

(1) Spatial ability: as mentioned in Section 2, the higher a learner's spatial ability is, the less degree of visualization she/he will need.
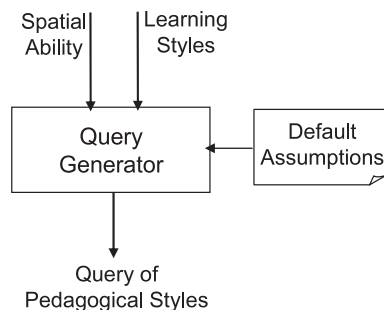
Fig. 5. Query formulation with default assumptions.

(2) Sensing/intuitive learning style: sensing learners would prefer concrete or practical learning materials, while intuitive learners prefer theoretical and abstract materials, such as mathematical descriptions.
(3) Active/reflective learning style: active learners would prefer doing experiments, while reflective learners would prefer learning materials in the form of typical lecture.
(4) Level-of-details is set to medium (numeric value 0.5) initially.

The pedagogical heuristics described by Soloman and Felder, 2005 associated with ILS as shown above are incorporated in order to make the decision more sensible and pedagogically valid.

The form of queries is illustrated below, which is represented as a vector consisting of seven elements:

$Q = \langle$ is_2D, is_3D, is_concrete, is_abstract, is_lecture, is_experiment, level_of_details $\rangle$

Each element has a numerical value varies from 0 to 1. The following rules of complement are established to assure that the representation is robust if any element is indexed as zero in computing similarity:

is_2D + is_3D = 1,
is_concrete + is_abstract = 1, and
is_lecture + is_experiment = 1.

The job of the query generator shown in Fig. 5 is to transform learners' traits into the query Q. The process of inference is a simple matching. The default assumption 1 shown above is applied to determine the value of elements is_2D and is_3D; the default assumption 2 is applied to is_concrete and is_abstract; the default assumption 3 is applied to is_lecture and is_experiment; while finally default assumption 4 is applied to level_of_details. Note that the representation of query has the same format as what will be described next—the feature vector of learning materials.

4.3. Computing similarity for material selection

In typical IR, it is quite popular to use the cosine measure between vectors as the similarity measure. A detailed comparison of different methods of measuring similarity

has been presented in Strehl and Ghosh, 2003. By considering the characteristics of our task, here we intend to use extended Jaccard coefficient to measure the similarity between content objects (Han and Kamber, 2001; Strehl and Ghosh, 2003). Notably, the extended Jaccard coefficient has the advantage of being capable of handling binary and numerical values at the same time. Since the condition is not unusual in the task of annotating styles on learning materials, the use of extended Jaccard coefficient in the task was grounded.

Though we currently employ the extended Jaccard coefficient as the similarity measurement, it is possible to incorporate other metrics to represent similarity, such as the Euclidean distance metric. From the clustering point of view, a meaningful similarity measure should be robust to transformations natural to the problem domain in the feature space (Strehl, 2002). However, in modeling the problem of styles-matching as clustering with mixed information source, we do not have such prior knowledge in judging which similarity measure would be the best to the first attempt. Nevertheless, it seems that extended Jaccard coefficient is generally feasible in terms of scaling- and translation-robustness (Strehl, 2002), so we experimentally use the metric in this work. Domain-specific factors of this design decision merit further investigation.

To derive the measure, each object should be represented as a feature vector. In our case, features of learning materials are transformed into the feature vector which consists of seven elements:

$$M = \langle \text{is\_2D, is\_3D, is\_concrete, is\_abstract, is\_lecture, is\_experiment, level\_of\_details} \rangle$$

Evidently, this feature vector is of the same form as the query. Each element of the vector has a numerical value varies from 0 to 1. The first two elements is_2D and is_3D stem from the feature of Main_representation shown in Fig. 4. For example, the feature-value pair, Main_representation=0.8, could be transformed as is_2D=0.2 and is_3D=0.8. Similarly, elements is_concrete and is_abstract are transformed from the feature of Abstractness; is_lecture and is_experiment are from the feature of Activity_type. Finally, no transformation was done for the feature of Level-of-details shown in Fig. 4 due to its unidirectional nature.

Given the query vector Q and learning material's feature vector M, the similarity measure using extended Jaccard coefficient is computed as:

$$S_{Q,M} = S_{\text{Jaccard}}(Q, M) = \frac{Q^{\text{T}}M}{Q^{\text{T}}Q + M^{\text{T}}M - Q^{\text{T}}M} \tag{1}$$

The higher the measure derived by Eq. (1), the more similar it is between the learning material M and the query Q. In other words, this learning material will be better matching to learners' traits and ordered with a higher priority. By repeatedly computing the similarity measures of the query and all candidate learning materials, a threshold could be set to divide learning materials into two categories, 'recommended' and 'not recommended'. That is:

$$S_{\text{threshold}} = \{h | 0 \leq h \leq 1\} \tag{2}$$

Note that such a threshold is set imperially by considering characteristics of the set of learning materials, the learning domain, and the pedagogical strategies intended to be applied. Then, the category C_M of each learning material M can be determined by:

$$C_M = \begin{cases} \text{recommended,} & \text{if } S_{Q,M} \geq S_{\text{threshold}} \\ \text{not recommemded,} & \text{otherwise} \end{cases} \qquad (3)$$

It is recognized that besides spatial ability and learning styles, there are still other types of features could be imported into this framework for selecting materials in terms of pedagogical styles and strategies. The mechanism of adaptive material selection is open to other descriptions of learners in reflecting various educational concerns at the system level.

4.4. Tuning the decision with learners' feedback

Since the mechanism of material selection described above only takes the measured scores of learners' traits and default assumptions as the basis of decision making, the generated query is just a rough stereotype, and may be improper to reflect learners' actual preference of styles. Therefore, CooTutor will directly interact with learners by asking them if they satisfy with the presentation style after a session is finished. A simple question like 'Do you want to see more pictures?' will be presented to learners. Learners' feedbacks are then employed to tune the query vector for the next round of material selection.
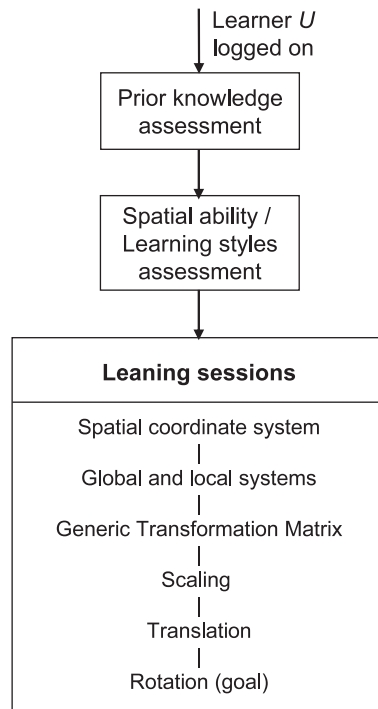
Learner *U* logged on

Prior knowledge assessment

Spatial ability / Learning styles assessment

Leaning sessions

Spatial coordinate system
|
Global and local systems
|
Generic Transformation Matrix
|
Scaling
|
Translation
|
Rotation (goal)

Fig. 6. Major events learner U will meet by setting the concept of 'Rotation around single axis' as the learning goal.

5. Usage scenario

Before describing the empirical evaluation of the system, here we first present an example of how the adaptive mechanism works, specifically how learners with different traits, including spatial ability and learning styles, would be tutored adaptively in CooTutor. A usage scenario is described to demonstrate how the system functions.

Assume that a learner U has logged in CooTutor. U's learning goal is to learn the concept of 'Rotation around a single axis'. The major events that she/he will meet are illustrated in Fig. 6. The system then arranges a concept sequence (i.e. learning plan) consisting of six concepts, as the box of learning sessions shown in Fig. 6. Note that concepts known by the learner will not be included into the plan. Assume that learner U's spatial ability score is assessed via the PVRT test as a normalized score of 0.7 (i.e. qualitatively speaking, a learner with sufficient spatial reasoning skill); sensing/intuitive learning style assessed by the learning style questionnaire is 0.3 (i.e. a sensing-apt learner); active/reflective learning style is 0.5 (i.e. a balanced learner in terms of the dimension of learning style). Then the following query is generated:

$$Q = \langle \text{is\_2D}, \text{is\_3D}, \text{is\_concrete}, \text{is\_abstract}, \text{is\_lecture}, \text{is\_experiment}, \text{level\_of\_details} \rangle$$
$$= \langle 0.7, 0.3, 0.7, 0.3, 0.5, 0.5, 0.5 \rangle$$

By employing the mechanism of material selection, the similarity measures of learning materials are computed for ranking these materials. Table 1 illustrates all learning materials in the content repository relating to the planned learning sequence. Assume that the threshold is set as 0.7. Note that besides the order of presentation is ranked according to the similarity scores, some of the learning materials are not recommended, as shown in Table 1. For example, learning material #26 is not recommended. Its feature vector is represented as:

$$M = \langle \text{is\_2D}, \text{is\_3D}, \text{is\_concrete}, \text{is\_abstract}, \text{is\_lecture}, \text{is\_experiment}, \text{level\_of\_details} \rangle$$
$$= \langle 0.8, 0.2, 0.1, 0.9, 0.8, 0.2, 0.5 \rangle$$

Table 1
Learning materials with feature values and similarity measures

Concept	ID	Main_rep.	Abstractness	Activity_type	Similarity	Recommended
Spatial coordinate	#27	0.2	0.2	0.2	0.90	T
system	#25	0.2	0.8	0.2	0.71	T
	#26*	0.2	0.9	0.2	0.65	F
Global and local	#31	0.4	0.3	0.2	0.90	T
sys.	#30	0.4	0.3	0.2	0.90	T
Generic Trans.	#29	0.2	0.7	0.2	0.77	T
Scaling	#53	0.2	0.8	0.2	0.71	T
	#33*	0.8	0.2	0.9	0.68	F
Translation	#52	0.2	0.8	0.2	0.71	T
	#32*	0.8	0.2	0.9	0.68	F
Rotation around	#36	0.2	0.3	0.2	0.91	T
single axis	#35	0.2	0.4	0.2	0.90	T
	#37*	0.8	0.2	0.9	0.68	F
	#34*	0.2	0.9	0.2	0.65	F

Rows are marked with asterisks(*) if that learning material is not recommended. ($S_{\text{threshold}} = 0.7$).

The similarity measure between *M* and *Q* is then computed as 0.65, which is lower than the threshold, so the learning material would not be recommended by the system.

CooTutor also incorporates the design of adaptive navigation support (Brusilovsky, 2001) in order to make the system capable of adapting to learners' knowledge status, and to let learners introspect their own learning progress. Fig. 7 illustrates the navigation bar of CooTutor, as the enlarged part of the figure shows. In the navigation bar, according to the learner's learning progress, some links will be activated, while some will be disabled. Activated links are colored with metaphor as the guidance. For example, grayed links, like 'C14' shown in Fig. 7, imply that the concepts have already been known by the learner. If the learner clicks on the link of 'C14', she/he will be directed to a review session of concept C14. Yellow links, like 'C6' in the figure, indicate that the learner has viewed materials of this concept but did not perform well in a quiz, so a remedy session will be presented. Note that the function of adaptive navigation support aims to provide facilitation upon inter-concept navigation. For intra-concept navigation of each specific concept, the mechanism of adaptive material selection highlighted turns to play the major role in personalization.

For each specific concept in the concept sequence, learning materials associated with that concept are retrieved and ranked. The prioritized order of presentation is shown in Table 1. As mentioned above, materials with similarity scores lower than the threshold are labeled as 'un-recommended'. Once the learner finished viewing recommended materials
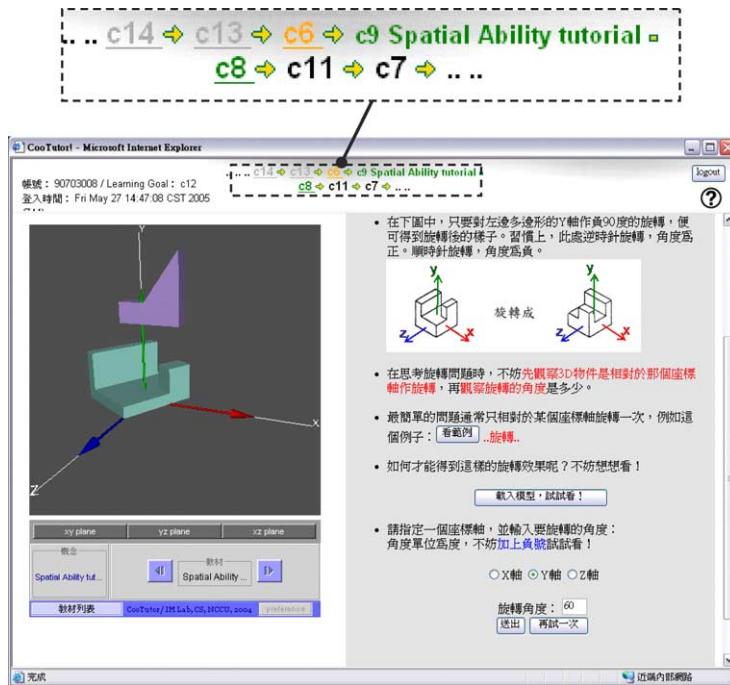

Fig. 7. The function of adaptive navigation support offered by CooTutor. Hyperlinks are colored with metaphors as guidance.
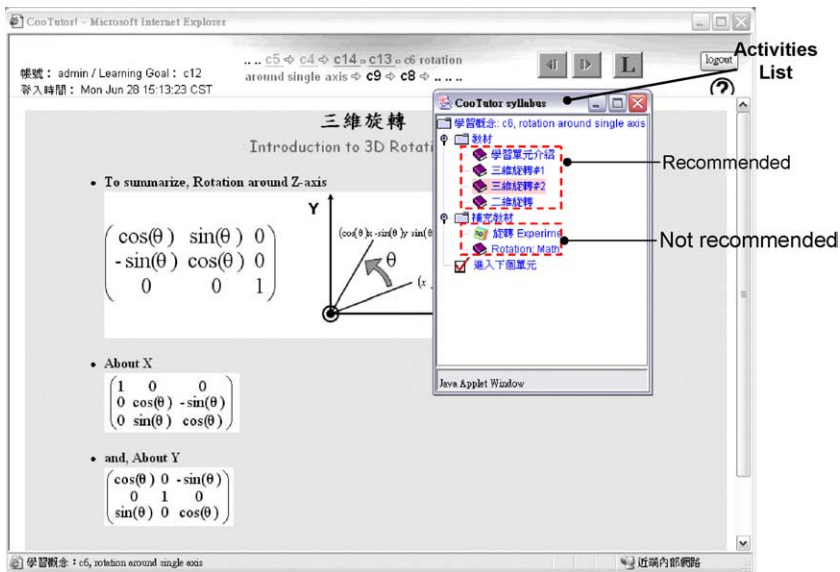
Fig. 8. An activity list presented to the learner showing the recommendation.

for a specific concept, recommended materials for the next concept would be fetched and presented subsequently.

The learner can also choose to browse the hyperspace by themselves within certain scope. On visiting each concept, CooTutor would prepare an activity list for the learner to navigate learning materials without a specific sequence. Fig. 8 shows the activity list presented to the learner. By adopting the activity list, the learner can choose to see un-recommended contents that are initially veiled by the system. The design is considered helpful to let the learner retain certain flexibility in interacting with the system.

6. Evaluation

6.1. Design of the evaluation

This exploratory empirical evaluation aims to investigate the effectiveness of the styles-matching strategy in CooTutor. Four different versions of CooTutor with different methods on material selection were used in this evaluation. This evaluation adopted a pre-test/post-test comparison-group quasi-experimental design (Chang, 2002). These four versions of system differ from each other on the strategy of material selection.

The evaluation was held in June 2004 at National Chengchi University (NCCU), Taiwan. Totally 31 graduate-level participants majored in Computer Science or Management Information Systems (master program) from NCCU have participated in the evaluation. All of them have learned fundamental linear algebra and computer graphics. They were organized in four groups, and each group was assigned to use one

version of CooTutor. The process of grouping is double-blinded. All participants did not know what version of CooTutor they were using. And the researcher did not know the participants' pre-test scores of spatial ability and SGT achievement test. So the researcher could not purposely prefer any group to others.

The whole duration of the intervention lasted for three weeks. Participants are asked to log in the system, take the pre-tests, view all learning materials, and finally be tested by post-tests. By considering that web-based learning is naturally a type of self-paced learning, participants were not enforced to operate the system at a specific time and fixed duration. They were only informed by the researcher that they can use the system at any moment they want before a specific due date. All of the four groups were assigned the same learning goal—the concept of 'Gimbal Lock' in SGT. And the same learning plan with identical concept sequence consisting of 14 concepts was used by all groups. All of the four groups received partial degree of adaptation, including concept sequencing and adaptive navigation support. The treatment of the evaluation is thus with or without adaptation in terms of spatial ability and learning styles. Fig. 9 depicts the whole process.

Table 2 summarizes how these four versions of CooTutor differ. Note that the number of participants in each group differs. Especially the last two groups shown in the table only have four and five participants respectively. This is because the main interest of this research is actually upon the first two groups (i.e. the LS and PreAuthor groups), but the effect of the last group is also suspected. In order to address both issues, this research choose to assign a large portion of participants into the first two groups, but keep a small portion of them in the last two for references.

Among these four versions, LS is the version that employs the mechanism of adaptive material selection. And the score of threshold of recommendation, $S_{threshold}$ was set as 0.6. Therefore, learning materials were selected and ranked adaptively based on participants' traits. Note that since learners' spatial ability and learning styles would vary, so the size of hyperspace would vary by employing the mechanism of adaptive material selection. In other words, some inappropriate materials were filtered out for a particular learner.
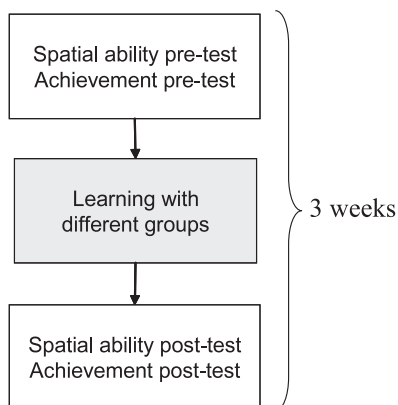

Fig. 9. Design of the evaluation.

Table 2
Four versions of CooTutor compared in the evaluation

Group	# Of participants	Strategy of selection	Adaptive ranking?	Size of hyperspace (# of materials)
LS	10	Styles-matching	Yes	Varies
PreAuthor	12	Manually pre-selected	No	16
NoFilter	4	No filtering	No	33
MisLS	5	Styles-mismatching	Mis-match	Varies

PreAuthor is the version that does not use the adaptive mechanism. Instead, a domain expert of SGT selected a fixed set of learning materials for the system a priori. This version could be thought as the group without traits-based adaptivity. Totally 16 learning materials were pre-selected for learning these 14 concepts. The third group, NoFilter, is the group that offers the participants all available materials stored in the content repository which now holds 33 learning materials. That is, no filtering or selection was done to reduce the size of hyperspace.

The last version, MisLS is the version that is designed to probe what if learning materials with improper pedagogical styles regarding learners' learning styles (i.e. styles-mismatching) were presented to the learner. For better factor control, the research chooses to only mismatch the element is_lecture and is_experiment of Q. The strategy is to exchange the values of is_lecture and is_experiment of the query. Five participants with extreme learning styles were specifically dispatched to this group. Therefore, some decisions have been made to form groups strategically. This is why a pre-test on spatial ability and achievement was necessary which may help compare the effects and interpret the results of different groups with fair.

6.2. Measuring instruments

There were three types of scores measured in this evaluation. They are learners' spatial ability, achievement on the topic of spatial geometric transformation (SGT), and learners' attitudes on CooTutor. For measuring spatial ability, the PVRT test was employed. For measuring learners' achievement on SGT, a self-compiled achievement test consisting of seven items was authored and integrated into CooTutor. And for measuring learners' attitudes on the system, an attitude questionnaire consisting of 15 single-choice questions was used.

In the previous work of PVRT (Bodner and Guay, 1997), the estimated reliability of 0.78–0.80 was reported by using Kuder-Richardson 20 formula (KR-20). In this study, the estimated reliability of the PVRT by using KR-20 is reported here for comparison: 0.69 for the pre-test and 0.75 for the post-test.

The 7-item SGT achievement test was compiled and employed on both pre- and post-test. The estimated KR-20 coefficient is 0.59 for the pre-test and 0.28 for the post-test. It is suspected that two main factors are subject to the scenario of low reliability. First, the number of items of this test is rather small. It is naturally difficult to achieve high reliability for a test with few items. Second, the homogeneity between participants is

high. In other words, small variation is there between individuals. As the evaluation result that will be described later reveals, all participants performed quite well on the achievement post-test. The amount of such participants has closely reached the limitation of measurement of this instrument. When the distribution is highly skewed like this case, it is likely that the coefficient was reduced because the underlying computation of reliability coefficient connects to the distribution of measurement results (De Gruijter et al., 2003). It is interesting to note that at one hand the low reliability coefficient on post-test seems to suggest that the result is not reliable enough. But on the other hand, by taking the global observation on results of both pre- and post-test, it is very likely that learning with CooTutor has substantially changed the distribution of the participants toward the high score area. This observation suggest that, the instrument itself is probably not that un-reliable (KR-20 0.59 on the pre-test), but the improvement of SGT understanding has contributed side-effects to computing the reliability coefficient of post-test.

The 15-item attitudes questionnaire is used to assess learners' attitudes toward CooTutor. This questionnaire adopts the six point Likert-type scoring method. For each question item, there are six response options extending from 'strongly agree' to 'strongly disagree' and are scored from 6 to 1 correspondingly. The reliability coefficient by using the method of Cronbach's alpha is estimated as 0.96. In order to probe participants' multi-dimensional attitudes toward the system, partial items were grouped and analyzed separately for assessing a specific type of attitude that one would be interested in. Three sub-dimensions of the survey were assessed and identified. They are: (1) learners' attitude toward the adaptive guidance provided by the user interface, (2) attitudes toward the learning materials recommended by the system, and (3) learners' engagement and motivation on learning SGT.

6.3. Data analysis

In this study, the objective is to detect if any difference of the effects existed between the four groups. Several issues are considered in the analysis:

(1) The unbalanced degree of freedom: note that the number of participants (i.e. degree of freedom of statistics) is not equal between the four groups. For the groups this research mainly focus on—LS and PreAuthor groups, each group was assigned around 10 participants. For groups intended to be as references, each group was assigned just around five participants. Under this scenario of unbalanced degree of freedom among different populations, it is not tenable to compare the results of all groups by applying a single statistical significance test (Aron and Aron, 2000).

(2) Covariates involved in the evaluation: for comparing the post-test scores of PVRT and achievement, it is necessary to consider how covariates can be used to enhance the validity of the analysis result. The Pearson correlation of PVRT pre-test and SGT achievement pre-test is not statistical significant although one may initially suspect that they should correlate significantly. It is interesting to note that a similar scenario also happened in the domain of Chemistry (Bodner and Guay, 1997). It is suggested that though spatial ability and the SGT domain (or chemistry) are recognized to be

highly correlated. If the test instrument of the domain tends to measure the accumulation of domain knowledge but not the use of spatial reasoning skills, the correlation between PVRT score and achievement score would seem to be insignificant (Bodner and Guay, 1997).

(3) The needs of comparing effect size between groups: researchers have noted the insufficiency of using only statistical significance testing to interpret experimental data (Cohen, 1988; Daniel, 1998; McLean and Ernest, 1998). This is mainly because the computation of statistical significance is related to the sample size involved in the analysis. For a small size of samples, like this study, achieving statistical significance is inherently more difficult than for a large one. Some researchers even recognized that "an SST (statistical significance testing) is largely a test of whether or not the sample is large." (Daniel, 1998, p26). Since in this evaluation, the sample size is rather small. The coefficient of effect size may offer more informative implication of the data in this case. As McLean and Ernest (1998, p17) described, "the effect size gives an estimate of the noteworthiness of the results". Meanwhile since effect sizes are standardized statistical scores, reporting these analyses may help following studies to conduct meta-analysis between different studies.

Taking these concerns altogether, the effect sizes of gain scores between post- and pre-tests are the main statistics the research focuses on. For the results of LS and PreAuthor groups, the method for analyzing covariance (ANCOVA) is also applied.

6.4. Results

Tables 3–5 present the results of PVRT test, SGT achievement test, and attitude questionnaire, respectively. For tests consisting of both pre- and post-tests, i.e. the PVRT test and achievement test, a paired 2-tailed t test was performed to compare the difference of means between post- and pre- tests. Meanwhile, the effect size is specifically computed by using Cohen's d coefficient. Since Cohen's d coefficient (and other types of effect size measure) is a standardized score, some criterion is required to judge and conceptualize the result. From the literature, researchers have suggested to use such a criterion. That is, for Cohen's d coefficient, 0.2 is a small effect size; 0.5 implies medium size; 0.8 and above

Table 3
Statistics of pre- and post- PVRT scores with effect size coefficients

	Pre_PVRT		Post_PVRT		Post vs. pre t-test, p	Effect size, d
	Mean	SD	Mean	SD		
LS ($n=10$)	15.60	3.03	16.60	2.12	.437	.383
PreAuthor ($n=12$)	13.25	3.22	16.17	2.13	.013**	1.069[‡]
NoFilter ($n=4$)	17.00	2.16	18.25	2.22	.194	.570[†]
MisLS ($n=5$)	16.60	1.95	13.20	5.81	.252	-0.785[△]
Overall	15.03	3.14	16.10	3.18	.173	.337

*$p<.1$ **$p<.05$. ES: ‡, large; †, medium; △, negatively large.

Table 4
Statistics of pre- and post- SGT achievement scores with effect size coefficients

	Pre_achievement		Post_achievement		Post vs. pre t-test, p	Effect size, d
	Mean	SD	Mean	SD		
LS ($n=10$)	3.90	1.85	4.90	1.52	.052*	.589[†]
PreAuthor ($n=12$)	4.67	1.72	5.33	1.23	.166	.445[†]
NoFilter ($n=4$)	4.00	1.63	5.25	1.26	.015**	.857[‡]
MisLS ($n=5$)	4.40	1.82	4.60	0.89	.749	.140[♦]
Overall	4.29	1.72	5.07	1.26	.003***	.514[†]

$*p<.1$ $**p<.05$ $***p<.01$. ES: ‡, large; †, medium; ♦, small.

indicates a large effect size (Cohen, 1988; Aron and Aron, 2000). By using the viewpoint of effect size, the results shown in Table 3−5 are described as follows.

6.4.1. Spatial ability enhancement

An ANCOVA analysis has been conducted upon the data of LS and PreAuthor. The post-test score of PVRT is the dependent variable and the pre-test score of PVRT is used as the covariate. The result, $F(1,19)=0.06$, is not statistically significant.

Back to Table 3, it is worth noting that the PreAuthor group reveals the best performance among all groups on spatial ability enhancement. The result of paired t-test reveals statistically significant ($p=0.013<0.05$), and the effect size is quite large ($d=1.069$). Though the LS group did not reveal strong effectiveness on this task, but on the other hand the MisLS group performs quite worse on the post-test, and the effect size, $d=-0.785$, is very large on the inverse (i.e. negative) direction. No similar scenario can be found in other groups.

6.4.2. SGT achievement

The ANCOVA analysis on comparing LS and PreAuthor shows no statistical significance as well. Similar to the scenario of examining the data of spatial ability enhancement, the pre-test score of SGT achievement is used as the covariate in the analysis. The result is also insignificant: $F(1,19)=0.034$. That is, these two groups seemed to perform equally well on the task of enhancing SGT achievement.

Table 5
Results of participants' responses on the attitude questionnaire

	Guidance		Recommendation		Engagement		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LS ($n=10$)	3.83	1.12	3.35	1.23	3.70	1.31	3.65	1.15
PreAuthor ($n=12$)	4.06	.93	3.54	1.16	4.08	0.66	4.01	.75
NoFilter ($n=4$)	4.25	1.23	3.50	.71	3.94	1.01	4.05	.90
MisLS ($n=5$)	4.33	1.00	4.30	1.20	5.10	0.98	4.77	.99
Overall ($n=31$)	4.05	1.03	3.60	1.14	4.11	1.06	4.02	.98
LS vs. PreAuthor t-test, p	.628($d=.2$)		.71($d=.16$)		.384($d=.36$)		.387($d=.37$)	

From Table 4, the LS group performed a little bit better than the PreAuthor group. From the view of effect size, the d coefficient is 0.589 for the LS group which is a medium effect size, and 0.445 for the PreAuthor group which is very close to a medium one. The best group is the NoFilter group, the result of paired t-test comparing the means of post- and pre- tests indicates statistical significance ($p = 0.015 < 0.05$). Its effect size, $d = 0.857$ is a large one. Finally, for the MisLS group, though there shows an increasing scenario of the post-test comparing to the pre-test. However, the gain effect is very small. The effect size $d = 0.140$ is relatively less significant comparing with the effects of improvement shown by other groups and the overall performance.

6.4.3. Participants' attitude on CooTutor

Table 5 shows the result of participants' attitude on CooTutor. The questionnaire adopts the 6-point Likert-type scoring method. That is, for each question item, 1 is the lowest score, and 6 is the highest one. Therefore, the middle score of each item is $(1 + 6)/2 = 3.5$. It can be observed that in Table 5, most results exceed the middle point indicating that participants have positive attitude on the system.

A 2-tailed independent t-test is conducted upon the scores of LS and PreAuthor groups. For the scores of guidance, recommendation, learning engagement, and the overall score, no statistical significance is found. The effect size d of the comparison (i.e. LS vs. PreAuthor) on each category is all small. Tough it seems interesting that the LS group got somewhat lower scores than others on each category, but the difference is small regarding the variation. It is difficult to judge if the underlying attitude of this group's participants is evidently different to other groups'.

6.5. Discussion

The discussion starts from investigating the result of the questionnaire. Using an attitude questionnaire to evaluate the system is a practice widely applied. To compile or answer a questionnaire is much easier than instruments like reliable achievement tests or standardized psychometric tests. However, it is inevitably that such type of measurement might be quite inaccurate and unreliable as questioned by researchers (Yu et al., 2000). The lesson learned is clearly that this type of self-reported data should be coped with cautiously. From Table 3, it can be found that the MisLS group reveals a degree of decrease on the PVRT post-test comparing to their high scores on the pre-test. Clearly, this scenario cannot be simply interpreted as that the participants become less-competent after using this version of system. It is suspected that this scenario may be largely caused by the attitude underlying them, especially, the attitude of not willing to give their best effort on the post-test. If the deduction is true, then the questionnaire is supposed to manifest such a situation. However, the result shows that MisLS's questionnaire score is still quite high. For the category of engagement, MisLS even shows the highest score among all the groups. Therefore, from the perspective of research methodology, it is suggested that attitude questionnaire is better to be employed along with other instruments or data source (e.g. web usage mining). Simply using self-reported data from questionnaire might be inappropriate in introspecting learners' underlying attitude.
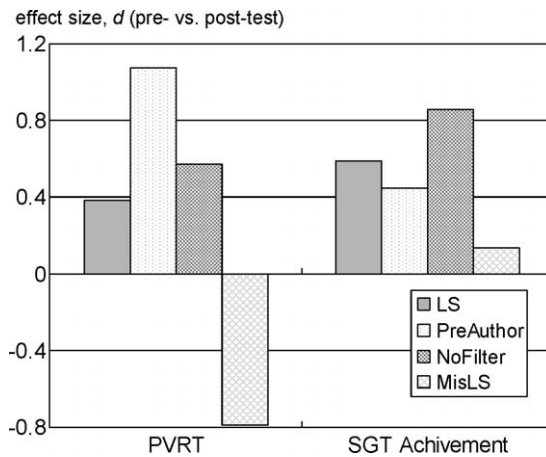
Fig. 10. A comparison of learning outcomes between different groups using two measurements (PVRT and SGT achievement test).

Regarding the evaluation results, Fig. 10 compares the learning outcomes of the four groups in terms of difference between post- and pre-test scores using PVRT and the SGT achievement test. It is quite evident that the design of CooTutor with 3D visualization and interactivity is beneficial to SGT learning in principle. Most students, disregarding which version of CooTutor they used, have achieved significant improvement in terms of spatial visualization ability and SGT's domain knowledge. However, the mechanism of adaptive material selection does not outperform other designs—especially the version intended to be compared with, a set of learning materials selected by a human teacher. It is also worth noting that the situation of styles mismatching, as shown by the result of MisLs group, may yield negative effects on learning, specifically for those learners with extreme learning styles.

The evaluation seems to suggest that regardless the method used in material selection, except MisLS, CooTutor is generally helpful to SGT learning. An interesting view is to relate this result to dual coding theory (DCT) (Paivio, 1991; Najjar, 1996). DCT proposes a model of memory and information processing for human cognition, in which two separate channels specialize in processing verbal information (e.g. text and speech) and visual information (e.g. pictures). The two channels are generally independent, but concurrent multimedia information processed through both channels is possible, which is the case called referential processing (Paivio, 1991; Najjar, 1996). Najjar indicated that computer-based multimedia may help learning specifically when multimedia can support referential processing and mutual-interference between multimedia is controlled (Najjar, 1996). Evidently, CooTutor has been providing an environment to support referential processing of verbal (i.e. text) and visual (i.e. 3D graphics) information for SGT learning. Also, the two types of information are integrated as a compact component of presentation via the means of interactive hyperlinking between the two media, such that the situation of mutual-interference (i.e. distraction caused by multimedia presentation) seems insignificant in CooTutor. Therefore, it may be reasonable to conjecture that probably the

matching of learning profiles and materials' presentation styles is minor in SGT learning, while domain-specific and cognitively informed concerns, such as the design of anchored instruction and interactive 3D graphics for SGT learning, may actually dominate learning effectiveness.

To summarize, two major findings revealed in this evaluation are: (1) CooTutor seems beneficial to most students, regardless which version they were actually using, and (2) MisLS users seem relatively demotivated in using the system. We conjecture that probably domain-specific concerns are of major influence in SGT learning. The effectiveness of styles-matching remains inconclusive, but when learners have certain extreme leaning styles, mismatching of styles may be still painful for learning. This situation merit subsequent works in learning environment and design experiment to resolve the issue. However, since various perspectives of this work, including the sample size employed in the evaluation, the design and control of interventions, the measuring instrumentation, and the design of the system are still at the early stage of exploration, refinement of designs and replications of the evaluation to further attest the value of trait-based adaptivity are necessary. The result of effect sizes reported in this paper can be employed for the purpose of meta-analyses and comparison along with other empirical studies adopting similar designs of system architecture and evaluation in the future.

7. Conclusion

In this paper, the instructional strategy of styles-matching is incorporated into the adaptive web-based learning system, CooTutor, for tutoring spatial geometric concepts. The system is designed to retrieve learning materials with appropriate pedagogical styles. Learners with different degrees of spatial reasoning skills and learning styles can then be tutored adaptively. CooTutor is a flexible scheme open to other specifications of latent traits. Therefore, Cootutor can be employed as a test-bed for exploring newly conceived educational concerns in terms of measurable traits. It is also possible to use CooTutor to assist theoretical works in exploring and verifying specific psychological constructs (e.g. a particular dimension of learning style) with adequate design experiments.

An empirical evaluation was conducted to evaluate CooTutor. By the evaluation result, it is found that CooTutor with 3D visualization and interactivity is generally beneficial to SGT learning. Although the mechanism of adaptive material selection fulfilling styles-matching strategy does not outperform typical one-size-fits-all designs, as shown in the work, it still seems hopeful in preventing from the situation of mismatching. A future research direction informed by this result is to unpack the interplay of domain-specific and domain-independent instructional concerns in designing effective learning environments. Specifically, we intend to incorporate ideas and techniques stemming from the area of science education to explore issues of learners' prior conception and problem-solving abilities situated in this type of learning environments. Recent studies of learning styles also suggested other strategies for learning styles (Felder et al., 2002), such as to offer needed scaffolding to learners and let learners try to learn from non-preferred styles. In other words, besides the strategy of styles-matching, it may be beneficial to try other approaches in addressing learners' latent traits. Another future goal in improving the

system design is to focus on the part of diagnosis and remediation that help learners become capable of learning from non-preferred or unfamiliar situations. The development of a user modeling framework for creative problem-solving abilities which aims to probe learners' alternative conceptions on solving scientific problems is now underway in our team (Wang et al., 2005a; Wang et al., 2005b) and may hopefully play a key role in our future works.

Acknowledgements

References

Aron, A., Aron, E.N., 2000. Statistics for Psychology, second ed. Prentice-Hall, Englewood Cliffs, NJ.

Baeze-Yates, R., Ribeiro-Neto, B., 1999. Modern Information Retrieval. Addison-Wesley, Reading, MA.

Bodner, G.M., Guay, R.B., 1997. The Purdue visualization of rotations test, *The Chemical Educator*, 2(4). Available from: http://chemeducator.org/bibs/0002004/00020138.htm.

Bransford, J.D., Brown, A.L. (Eds.), 1999. How People Learn: Brain, Mind, Experience, and School. National Academy Press, Washington, DC.

Brusilovsky, P. 2000. Adaptive Hypermedia: from Intelligent Tutoring Systems to Web-based education. Proceedings of Intelligent Tutoring Systems Conference, pp. 1–7.

Brusilovsky, P., 2001. Adaptive hypermedia. User Modeling and User-Adapted Interaction 11, 87–110.

Brusilovsky, P., Vassileva, J., 2003. Course sequencing techniques for large-scale web-based education. International Journal of Continuing Engineering Education and Lifelong Learning 13 (1–2), 75–94.

Chang, C.-Y., 2002. The impact of different forms of multimedia CAI on students' science achievement. Innovations in Education and Teaching International 39 (4), 280–288.

Chang, C.-Y., Tsai, C.-C., 2005. The interplay between different forms of CAI and students' preferences of learning environment in the secondary science class. Science Education 89 (5), 707–724.

Clark, R.E., 1994. Media will never influence learning. Educational Technology Research and Development 42 (2), 21–29.

Cognition and Technology Group at Vanderbilt, 1990. Anchored instruction and its relationship to situated cognition. Educational Researcher 19 (6), 2–10.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Lawrence Erlbaum, New Jersey.

Daly, C., Horgan, J.M., 2004. An automated learning system for Java programming. IEEE Transactions on Education 47 (1), 10–17.

Daniel, L.G., 1998. Statistical significance testing: a historical overview of misuse and misinterpretation with implication for the editorial policies of educational journals. Research in the Schools 5, 23–32.

De Gruijter, D.N.M., van der Kamp, L.J. Th., 2003. Statistical Test Theory for Education and Psychology. Retrieved May 27, 2005, Available from: http://icloniis.iclon.leidenuniv.nl/gruijter/.

Draper, S., 2003. Learning styles notes. Psychology Department, University of Glasgow Retrieved May 27, 2005, Available from: http://www.psy.gla.ac.uk/~steve/lstyles.html.

Felder, R.M., Silverman, L.K., 1988. Learning and teaching styles in engineering education. Engineering Education 78 (7), 674–681.

Felder, R., Felder, G.N., Dietz, E.J., 2002. The effects of personality type on engineering student performance and attitudes. Journal of Engineering Education 9 (1), 3–17.

Han, J., Kamber, M., 2001. Data Mining: Concepts and Techniques. Academic Press, London.

Huang, H.-P., Lu, C.-H., 2003. Java-based distance learning environment for electronic instruments. IEEE Transactions on Education 46 (1), 88–94.

McLean, J.E., Ernest, J.M., 1998. The role of statistical significance testing in educational research. Research in the Schools 5, 15–22.

Najjar, L.J., 1996. Multimedia information and learning. Journal of Educational Multimedia and Hypermedia 5, 129–150.

Paivio, A., 1991. Dual coding theory: retrospect and current status. Canadian Journal of Psychology 45, 255–287.

Reeves, T.C., 1995. Questioning the questions of instructional technology research. Proceedings of the Annual Conference of the Association for Educational Communications and Technology, Research and Theory Division, Anaheim, CA, pp. 459–470.

Riecken, D., 2000. Personalized view of personalization. Communication of the ACM 43 (8), 27–28.

Rosati, P., Dean, R.K., Rodman, S.M., 1998. A study of the relationship between students' learning styles and instructors' lecture styles. IEEE Transactions on Education 31 (3), 208–212.

Shute, V.J., 1993. A comparison of learning environments: all that glitters. Computers as Cognitive Tools. Erlbaum, Hillsdale, NJ, pp. 47–73.

Soloman, B.A., Felder, R.M. Index of Learning Styles. Retrieved May, 27, 2005, Available from: www.ncsu.edu/felder-public/ILSpage.html.

Stach, N., Cristea, A., De Bra, P, 2004. Authoring of learning styles in Adaptive Hypermedia. Proceedings of World Wide Web Conference, NY, USA.

Strehl, A., 2002. Relationship-based clustering and cluster ensembles for high-dimensional data mining, PhD Dissertation, The University of Texas at Austin.

Strehl, A., Ghosh, J., 2003. Relationship-based Clustering and Visualization for High-Dimensional Data Mining, INFORMS Journal on Computing, pp. 208–230.

Tretiakov, A., Kinshuk, Tretiakov, T., 2003. Designing multimedia support for situated learning, Proceedings of Third IEEE International Conference on Advanced Learning Technologies, pp. 32–36.

Triantafillou, E., Pomportsis, A., Demetriadis, S., Georgiadou, E., 2004. The value of adaptivity based on cognitive style: an empirical study. British Journal of Educational Technology, 35(1), 95–106.

Wang, H.-C., Li, T.-Y., Chang, C.-Y., 2004. Adaptive presentation for effective Web-based learning of 3D content. Proceedings of fourth IEEE International Conference on Advanced Learning Technologies, Joensuu, Finland, pp. 136–140.

Wang, H.-C., Li, T.-Y., Chang, C.-Y., 2005a. A User Modeling framework for exploring creative problem solving ability. Proceedings of 12th International Conference on Artificial Intelligence in Education, Amsterdam, The Netherlands, pp. 941–943.

Wang, H.-C., Chang, C.-Y, Li, T.-Y., 2005b. Automated scoring for creative problem-solving ability with ideation-exploration modeling. Proceedings of 13th International Conference on Computers in Education, Singapore.

Wolfe, R., 1999. A syllabus survey: examining the state of current practice in introductory computer graphics courses. ACM SIGGRAPH Computer Graphics 33 (1), 32–33.

Woolf, B., Romoser, M., Bergeron, D., Fisher, D., 2003. Tutoring 3-Dimensional visual skills: dynamic adaptation to cognitive skill. Proceedings of 11th International Conference on Artificial Intelligence in Education, pp. 515–517.

Yu, C.H., Onlund, B., DiGangi, S., Jannasch-Pennell, A., 2000. Estimating the reliability of self-reported data for Web-based instruction. Proceedings of the 2000 Association for Educational Communications and Technology Convention, California, USA. Retrieved August 25, 2005, Available from: http://seamonkey.ed.asu.edu/~alex/pub/pub.html.