Available online at www.sciencedirect.com
ScienceDirect
Computers in Human Behavior 23 (2007) 1943–1957
Computers in Human Behavior
www.elsevier.com/locate/comphumbeh

The comparative efficacy of 2D- versus 3D-based media design for influencing spatial visualization skills

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

[a] *Department of Earth Sciences, National Taiwan Normal University, 88, Section 4, Ting-Chou Road, Taipei 11677, Taiwan, ROC*
[b] *National Chengchi University, Taipei, Taiwan, ROC*

Available online 20 March 2006

Abstract

This study explored the effects of 2D- versus 3D-based media representations on the influence of the spatial visualization ability of undergraduate science majors. A pre-test/post-test comparison-group experiment was conducted with 23 participants involved in the study. Participating students were randomly assigned either to the interactive 3D media representation group ($n = 13$) or the conventional 2D media representation group ($n = 10$); learning materials in both groups deliver the same information to students, but employ different media representations. All the activities were performed in a self-paced, web-based instructional system. The results of ANCOVA analysis showed statistically insignificant difference between groups in terms of students' post-test scores on the spatial visualization ability test with the students' pre-test scores as the covariate. However, a medium effect size was observed in favor of the 3D group in terms of practical significance. As a pilot study with a small sample size aiming to probe the research direction of this problem, the result of medium-sized effect magnitude is likely to implicate that the discrepancy of different representational design on students' performance of spatial ability assessment is noteworthy. Future study of this nature appears to merit further replications and investigations.

Keywords: Spatial ability; Media representation

* Corresponding authors. Tel.: +886 229354393; fax: +886 229333315 (C.-Y. Chang).
E-mail addresses: haochuan@iis.sinica.edu.tw (H.-C. Wang), changcy@ntnu.edu.tw (C.-Y. Chang).

0747-5632/$ - see front matter
doi:10.1016/j.chb.2006.02.004

1. Introduction

Spatial ability refers to a group of cognitive functions and aptitudes relevant to spatial tasks which is considered as an important factor in human intelligence, one which is essential to several scientific and engineering activities (Bodner & Guay, 1997; Lajoie, 2003; Olkun, 2003). Several studies have noted the high correlation between spatial ability and learners' achievement in various domains, such as chemistry (Bodner & Guay, 1997), engineering drawing (Potter & Merwe, 2001), and medical surgery (Eyal & Tendick, 2001). In everyday life situations, as some early works cited by Lajoie (2003) revealed, spatial ability is essential to activities like driving and orienting vehicles in unfamiliar physical surroundings. Meanwhile, it is also likely that spatial ability may influence the meta-cognitive learning of science and engineering. A few preliminary empirical works recently reported that learners with different degrees of spatial ability are likely to hold different attitude toward multimedia instructions and learning activities that include videos and three-dimensional computer animations (Huk, Steinke, & Floto, 2003). It appears that spatial ability seems to influence not only students' capacities for performing scientific or engineering skills, but may also take effects on students' perception of learning activities at the meta-cognitive and affective levels.

Empirical works have frequently detected the correlations between spatial ability and achievements on domains relevant to spatial manipulations. Fundamental research on analyzing spatial aptitude and developing appropriate instrumentation has also clarified factors underlying spatial ability (Bodner & Guay, 1997; Lajoie, 2003; Shepard & Metzler, 1971). The role of spatial ability in successful science and engineering activities is likely to be prominent. However, it is a missing ring that, until now, the statistical effect of aptitude-treatment interaction (ATI) between instructions and spatial ability is still understudied, which results in a dilemma that no evidence-based recommendations can be found regarding how to apply abundant findings of spatial aptitude into educational practices (Kyllonen & Lajoie, 2003; Lajoie, 2003). Huk et al.'s (2003) finding of spatial aptitude by courseware-design interaction is likely to be an alternative ATI effect that the statistical interaction between students' spatial ability and their attitudes toward different designs of the e-learning courseware module was revealed, although the targeted measure is not students' achievements in terms of domain knowledge acquisition. Nevertheless, studies of the relation of spatial ability and instruction seem not mature enough to direct real applications. Two main open questions include (1) how to use spatial ability in instructional design for science/engineering learning, and (2) how to develop it (Lajoie, 2003). The later question is within the focus of this work.

Researchers have suggested that spatial ability can be enhanced and taught by some instructional designs (Alias, Black, & Gray, 2002; Kwon, 2003; Lajoie, 2003; Potter & Merwe, 2001; Woolf, Romoser, Bergeron, & Fisher, 2003). Some works demonstrated that instructions using computer-based 3D visualizations can provide learners with adequate spatial experiences for developing their spatial ability (Kwon, 2003; Woolf et al., 2003). However, few empirical studies have established the causal relationships in greater depth. Moreover, no studies have explored the effects of two-dimensional (2D) versus three-dimensional (3D) media representations on the influence of the spatial ability of undergraduate students.

After reviewing meta-analyses and other studies of instructional media's influence on learning, Richard Clark (1983, 1985) concluded that there are no learning benefits to be gained from employing any specific medium for the purpose of instruction. He went on to argue that most media or computer-aided instruction (CAI) research, which compared CAI with conventional means of instruction or other media-based means, has suffered from inherently flawed methodologies. He also made the claim that media are only instructional vehicles which do not influence students' achievement or learning (Clark, 1994, 2001). Therefore, according to Clark's claim (Clark, 1994), only instructional methods can influence learning; media by themselves cannot do so. On the other side, empirical works on multimedia learning have reported some interesting results that confirmed the dual-coding hypothesis (Mayer & Anderson, 1991; Najjar, 1996). When multimedia instruction is designed properly that information encoded in multiple media representations is well integrated and permits cognitively referential processing to happen, it is likely to result in additive learning effects (Mayer & Anderson, 1991; Najjar, 1996; Paivio, 1990). The arguments around instructional media and multimedia learning seem to suggest very different positions about multimedia use in instruction, however, in our view, the core question is still about: under what condition and in what form will multimedia instruction become effective, such as an effective instructional method can be realized or cognitive referential processing can be triggered.

Modern multimedia technologies are matured enough allowing educators to have more options in designing the instructions. For the development of computer-based tutorials for spatial visualization skills, the comparison of 2D and 3D representations is considered necessary and informative in informing future design decision and the literature of multimedia learning. Specifically, the features of 3D navigation, interactivity, and simulation offered by the 3D media may help science and engineering educators to actualize novel instructional methods. The purpose of this study was to carry out such a comparison. Since spatial ability is a complex construct in which several similar but distinct skills may involve, it is necessary to clarify and restrict the scope of this study. Two major categories of spatial ability identified by previous works are spatial rotation and spatial visualization (Bodner & Guay, 1997). The later is considered to be more difficult and challenging. This study employed the Purdue visualization and rotation test (PVRT) in which adequate validity and reliability on assessing spatial visualization ability have been established, as the measuring instrumentation in the study. Accordingly, a web-based learning environment featuring interactive 3D visualization for instructing spatial visualization ability was designed to meet the purpose of the study.

2. Purpose of the study

The purpose of this study was to investigate the comparative effects of using web-based tutorials differentiated in including either 2D representation or interactive 3D representation, on the influence of undergraduate students' spatial visualization ability. The comparative effects of different media representations were investigated and analyzed in detail.

3. Methodology

This section describes the web-based tutorials used in this study, the characteristics of our participants, the experimental design, the measurement used, and the methods of data analysis employed in this study.

3.1. CooTutor: web-based tutorials for spatial ability

The CooTutor (Coordinate Tutor) system is a web-based system incorporating multiple media representations aimed at facilitating learners' spatial reasoning. Details of the system and its other revisions are described in Wang, Li, and Chang (2004, in press). Fig. 1 gives an overview depicting the system architecture of the web-based tutorials designed and used in this study. On the client side, three main elements form the user interface as the content presentation area. The main document area presents the main learning materials and contents, including textual and mathematical contents. These learning materials are typically authored in HTML format for easy incorporation of 3D media representation. By embedding scripting codes for 3D presentations (e.g., procedural programs of 3D animation) into them, these HTML documents are enriched with 3D contents programmed in the form of embedded JavaScript codes. When users interact with the user interface on-line, the system will activate the scripting codes and provide 3D visualization in the 3D blackboard module. These functions are implemented by using Java 3D technology and FastScript3D toolkits (Koehler, 2003). All learning resources on the server side are delivered through the HTTP protocol to the client side.

Fig. 2 shows the user interface featuring interactive 3D representation. The 3D blackboard is a module that provides 3D visualization and interactivity to help learners attain better reasoning with spatial concepts. Learners are allowed to interact with the 3D blackboard by dragging the mouse using different modes of navigation, e.g., zooming, panning and rotating, to view 3D models with different perspectives for better understanding of the

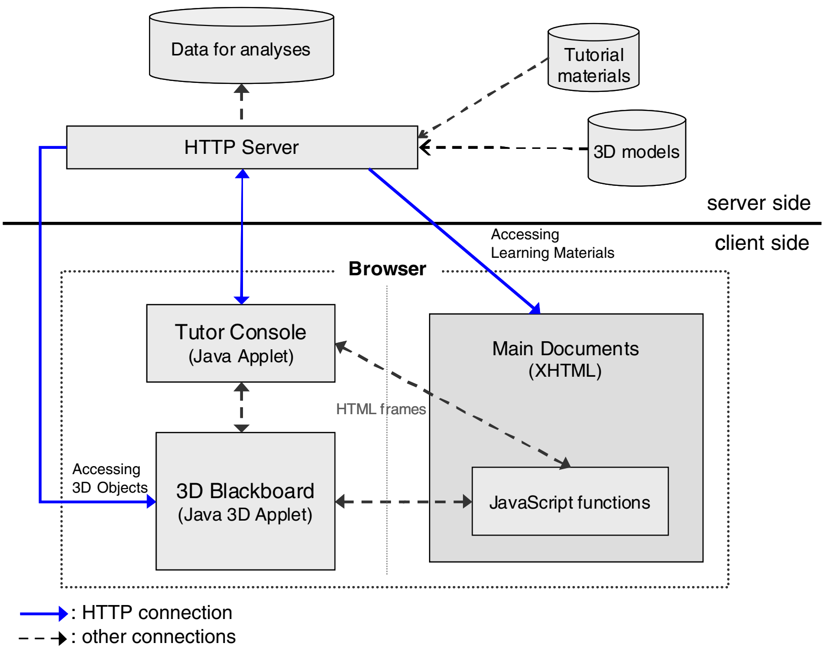

Fig. 1. System architecture of the web-based tutorials.

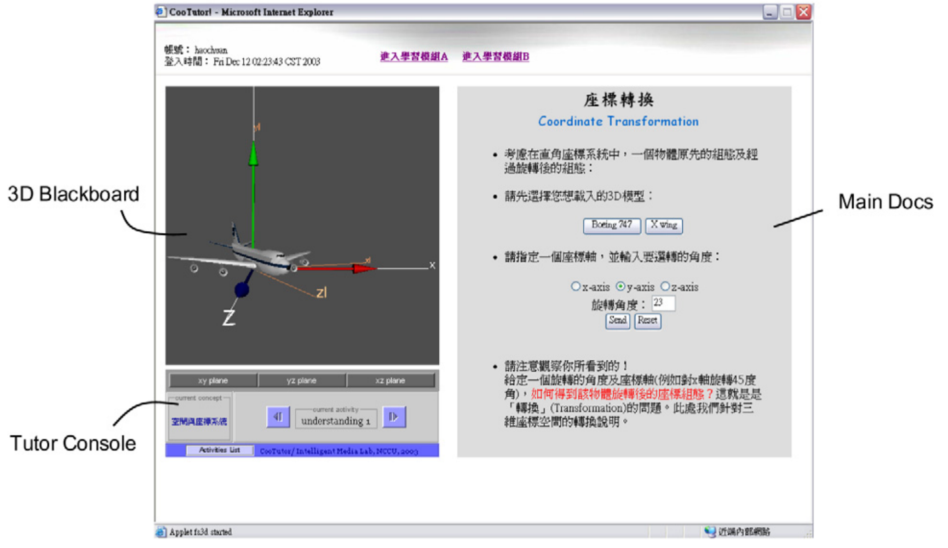

Fig. 2. User interface with 3D blackboard.

contents. Besides, there is also internal interaction between textual documents and the 3D blackboard. Learners can read the textual content in the main document area and then directly manipulate objects in 3D scenes by interacting with user interface objects (e.g., buttons, input fields) embedded in the HTML documents. The linkages between textual descriptions and 3D instantiations can be authored and programmed according to the need for understanding particular topics. Students can engage in activities of modifying parameters for three-dimensional transformations, such as translating and rotations, to see simulated animation results and to construct their understanding of the spatial tasks accordingly. The design of the user interface was informed by the theory of situated learning and anchored instruction in terms of providing authentic situations for students to explore (Cognition & Technology Group at Vanderbilt, 1990; Tretiakov, Kinshuk, & Tretiakov, 2003). The main document area is used to provide a context narrating instructional content and the 3D blackboard provides an authentic and visible representation for the abstract dynamic processes of spatial transformation of 3D objects. Tutorials can be authored to provide instructions for particular concepts situated in and demonstrated by 3D visualizations; while the interaction design between the main document and the 3D blackboard allows learners to explore the 3D world and conduct experiments of spatial transformation freely. This type of user interface aims to provide an environment which reshapes the contents to be delivered by web-based instruction: in addition to static documents, we now have dynamic learning activities for promoting students' active learning.

3.2. Participants in the experiment

The experiment was conducted in January 2004, in a computer classroom at the College of Science, National Taiwan Normal University. The duration of the experiment was approximately 60 min including an overview of the experiment, the pre-test, the self-paced learning sessions, and the post-test. Each participant had to complete the pre-test, the self-paced learning session (using the system), and the post-test. There were a total of 23 undergraduate students involved in this experiment. All of the participants major in Earth Sciences. Note that the small sample size is a major limitation of this study in discussing and generalizing its results.

3.3. Experimental design

A pre-test/post-test comparison-group experimental design was employed, similar to the one adopted by one of the article authors, Chang (2002, 2003). There were two versions of the learning materials, using different representations (2D versus 3D) delivering the same information for comparison. The overview of the experimental design is shown here schematically:

Group 1: S_1 R T_1 S_2
Group 2: S_1 R T_2 S_2
R = random assignment of students
T_1, T_2 = experimental treatments, self-paced learning with different representations
S_1, S_2 = pre-test and post-test, both using the PVRT spatial visualization ability test

We designed a web-based version of the Purdue visualization of rotation test (PVRT) proposed by Bodner and Guay (1997); this web-based PVRT was integrated into our system to assess all participants' spatial visualization ability on both the pre-test and post-test. Details of the instrument will be discussed in the following section.

After the participants entered the system, all of them were immediately tested by using an identical web-based PVRT. Then they were randomly assigned to either the 2D group (comparison group) or 3D group (experimental group) by a computer program. This assignment was independent of their performance on the pre-test. The comparison group ($n = 10$) received learning materials with conventional 2D presentations while the experimental group ($n = 13$) received materials with interactive 3D representations.

Note that the duration of the learning sessions was not strictly limited; participants were only reminded that they are allowed to browse the tutorial freely, repeating whenever they liked, within a 35-min period (or less if they preferred); then they would be assessed by the post-test. Even so, participants were still controlled by the system and asked to browse all of the contents at least once before they could get permission to leave the tutorial and enter the post-test. The system asks the participants to confirm their readiness before entering the post-test session. The experiment was designed based on the assumption that web-based learning is a self-directed and self-paced activity. This scenario is thus very different from what happens in a conventional classroom, where students are assumed to "learn" within a given period of time, even if they do not really learn. Note that chance-of-learning is also a variable in this experiment: this factor is the main reason for controlling duration in conventional educational studies, and so our experimental design also controlled duration by this design. The time spent by each student in the learning session—i.e., the amount of time each learner spent after finishing the pre-test and before entering the post-test—was also recorded and reported.

3.4. Learning materials with different representations

Two versions of tutorials were put into the experimental design. These two versions, 2D-based and 3D-based ones, both comprise seven HTML pages delivering the same information. The contents, using tutorial narrations and examples, were intended to instruct skills for handling spatial visualization tasks and let students construct their own senses of spatial transformation.

Participants in the 3D group were allowed to manipulate 3D objects presented in the 3D blackboard. On the other hand, for the 2D-based version, only texts and static pictures are presented. Possible influencing factors like color style and textual description were designed to be as close to the 3D-based version as possible.

It is worth noting that the architecture of the 3D version features the use of hyperlinking between textual narrations and animated visualization/simulations. In our experiment the student could press a button labeled "Press to show the demo" embedded in the Web page, and see the 3D animation presented in the 3D blackboard. Our study attempted to mimic this function and embedded the same buttons in the 2D-based version. But in the 2D-based version, there was no 3D model and no animation due to the limitations of 2D representation. Instead, after the student pressed buttons, previously hidden pictorial illustrations would become visible immediately. The screenshots of these two versions are shown and differentiated in Fig. 3, in which (a) is the 2D-based version, and (b) is the 3D-based version.

3.5. Measuring instruments

The web-based PVRT was transplanted from the original version proposed by Bodner and Guay (1997), which aims to measure the ability of spatial visualization. The original version was designed as a paper–pencil test comprising 20 items. In order to restrict the participants using analytical answering techniques upon these test items, a time limit of 10 min for the test is also enforced. The estimated reliability coefficient of 0.78–0.80 by using Kuder–Richardson 20 formula (KR-20) was reported by Bodner and Guay (1997).

The web-based PVRT developed in this study consists of all 20 items of the paper-based version. For the convenience of answering, the layout of these 20 items was arranged as two HTML pages where each page shows 10 items. The sequence of items also conformed to the paper-based version. Also, a time limit of 10 min was enforced strictly. A timer was embedded in the web-based PVRT shown to participants for reminding the rest of time. Fig. 4 shows the screenshot of web-based PVRT.

In order to prevent participants answering the post-test by rote or depending on their impressions of the pre-test, the order of selections for each item of the post-test was rearranged. The web-based PVRT was highly consistent on both pre- and post-tests. For the pre-test, the KR-20 coefficient was 0.80; and for the post-test, a coefficient of 0.88 was derived.

3.6. Data analysis

In order to examine and compare the effects of these two groups, an analysis of covariance (ANCOVA) was conducted on the post-test scores with the pre-test scores as the

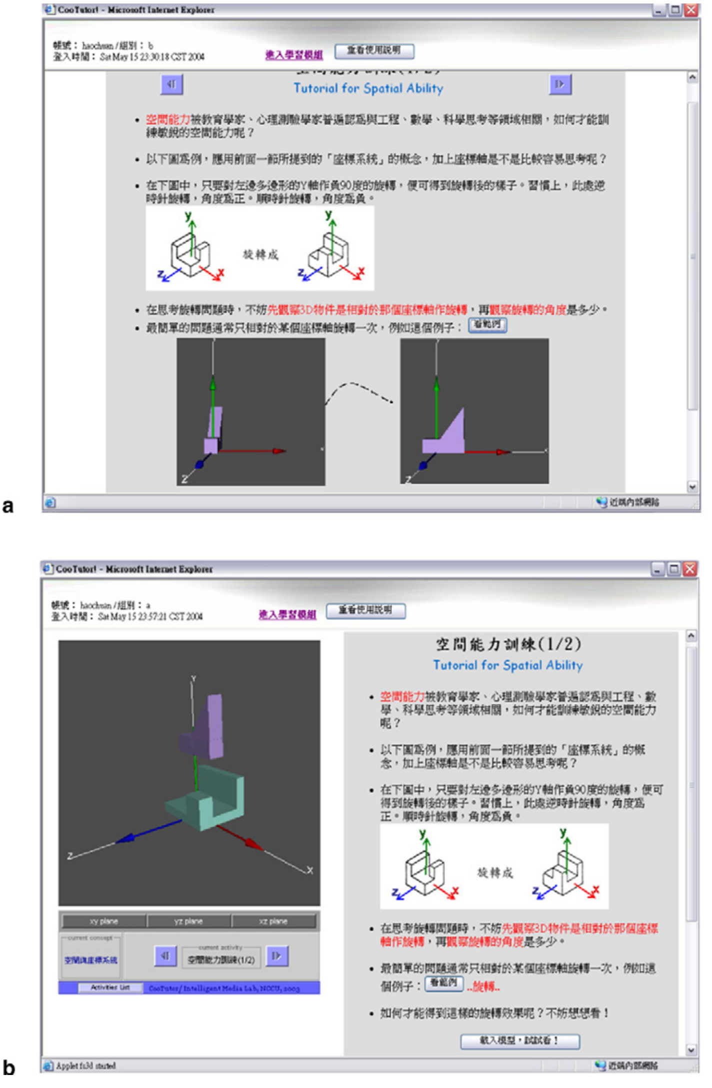

Fig. 3. The screenshots for learning materials with different media representations: (a) 2D-based version and (b) interactive 3D-based version.

covariate. Thus, the independent variable of the analysis was the groups; the dependent variable was the post-test scores; and the covariate was the pre-test scores.

Before performing the analysis, several underlying assumptions of ANCOVA were first examined. The assumptions include normality, homogeneity of variance, and homogeneity of within-group regression. The Kolmogorov–Smirnov tests revealed that both pre- and post-test scores for both two groups conforming to normality. The Levene's test of equality also assured that there was homogeneity of variances between two groups. However, the result of homogeneity of within-group regression was not strictly conformed. Though this point makes the process of ANCOVA somewhat informal, it is interesting and

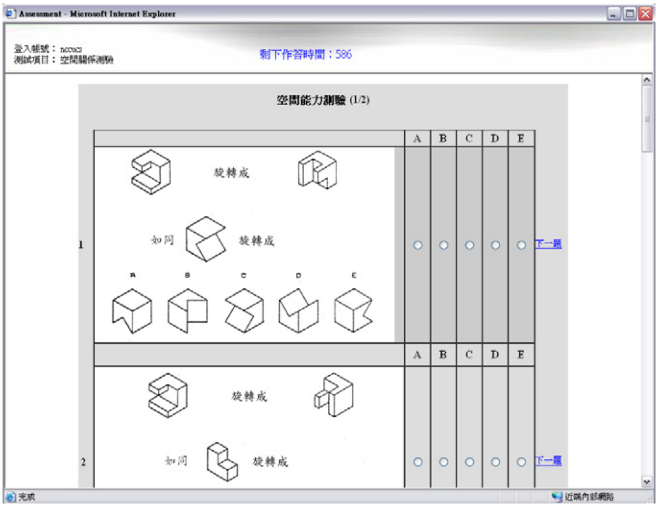

Fig. 4. The screenshot of the web-based PVRT.

noteworthy by concerning the design and purpose of this evaluation. The heterogeneity of regression of these two groups regarding their pre- and post-test scores evidently informs the influence of the treatment is inconsistent between the two groups. That is, this implicates that the influence of the two media designs is unlikely the same. More discussions about the result and its implication will be carried on in the following sections. The ANCOVA analysis was undertaken by using SPSS 11.0 (Statistical Package for Social Sciences).

4. Results

Table 1 shows the descriptive statistics of participants' pre- and post-test scores assessed by the Web-based PVRT. Note that the score of the PVRT test represents the number of items the participant answered correctly. Since there were 20 items in the PVRT totally, the highest possible score was 20. If the participant skipped any item of the test, then that item was treated as answered incorrectly.

As mentioned previously, the duration that each participant stayed in the learning session was also recorded. The data were, 2D group: Mean = 4.13 min (SD = 2.87) and 3D group: Mean = 8.33 min (SD = 4.03). An independent 2-tailed t-test was performed, and

Table 1
Descriptive statistics of participants' pre- and post-test scores, number of participants: 2D group ($n = 10$), 3D group ($n = 13$)

Pre-test scores		Post-test scores	
2D group Mean (SD)	3D group Mean (SD)	2D group Mean (SD)	3D group Mean (SD)
13.80 (3.94)	15.08 (3.73)	12.80 (4.59)	15.46 (4.82)
Overall, $n = 23$ 14.52 (3.79)		Overall, $n = 23$ 14.30 (4.81)	

it shows that there existed statistical significance ($p = 0.008$, smaller than 0.05). It could be inferred, therefore, that participants of different groups appeared to perform distinct browsing behaviors as a result of the current study.

The ANCOVA method allows us to eliminate the difference of pre-test scores between groups, and then the adjusted post-test scores revealing the real effects of the experimental treatment could be derived. Table 2 shows the result of adjusted post-test scores and the hypothesis test. The adjusted means with 95% confidence intervals of the result are also illustrated in Fig. 5.

The adjusted post-test score of the 3D group is higher than the 2D group. However, the result was not statistically significant, $F(1,20) = 1.1$, $p = 0.307$. It is noteworthy that the overall mean of pre-test scores as shown in Table 1 was 14.52. 2D group's adjusted mean of post-test score is less than the overall mean of pre-test, while conversely, the 3D group reveals an increase comparing to that value. The inclination between pre- and post-test looks quite different between 2D and 3D groups.

On the other hand, the effect size ($\eta^2 = 0.052$ or $f = 0.234$) according to the criteria of effect size proposed by Cohen (1988, p. 286) is very close to a medium effect size. As McLean and Ernest (1998, p. 17) described, "the effect size gives an estimate of the noteworthiness of the results." The result derived via the analysis could be a valued signal for informing future directions of investigation.

Table 2
The results of ANCOVA with pre-test scores as the covariate

Adjusted post-test scores		ANCOVA $F(1,20)$
2D group Mean (SEM)	3D group Mean (SEM)	
13.238 (1.35)	15.124 (1.18)	1.10 ($p = 0.307$, $\eta^2 = 0.052$)

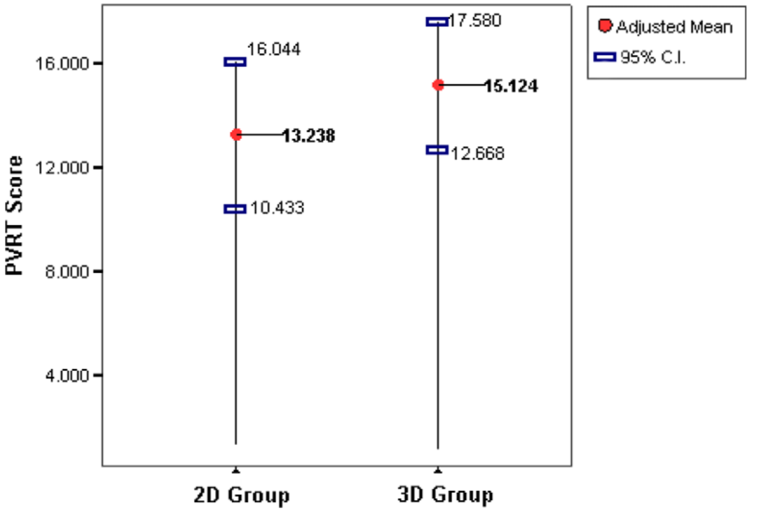

Fig. 5. The adjusted mean of post-test scores with 95% C.I.

5. Discussion and implications

In this section, the first issue to be discussed is the interpretation of the result of ANCOVA analysis. Researchers have noted the insufficiency of using only the result of statistical significance testing in statistical inference (Cohen, 1988; Daniel, 1998; McLean & Ernest, 1998). One of the main reasons is that the computation of statistical significance is related to the sample size involved in the analysis. For a small size of samples, like this experiment, to achieve the statistical significance is inherently more difficult than for a large one even if both have the same effect size. Conversely, it is quite possible to observe a statistical significance with a large sample size, even if there was little effect actually. Some researchers even considered that "a SST (statistical significance testing) is largely a test of whether or not the sample is large." (Daniel, 1998, p. 26).

Statistical significance and the coefficient of effect size offer different aspects of information. In this case, with a rather small sample size, the result of statistical insignificance is more likely to happen. The result of medium effect size may signify the possibility of finding statistical significance of this comparison with a future replication of study with a larger sample size.

As mentioned previously, the experiment aimed to conform to the actual scenario of web-based learning, and thus adopt an alternative design on offering students equivalent chance of learning. That is, it lets students to determine the time they want to leave the learning session. In this study, students in different groups revealed very different behaviors regarding time duration they spent in using the program. Since the factors underlying students' determination of when to leave the tutorial were not very clear, explanation of the statistical significance on learning duration should be undertaken cautiously. The difference may due to the differential features in the two groups of whether to enable 3D animation and 3D navigation or not. Interestingly, it was not so clear and may be beyond the scope of this study whether staying long is good or bad. On the one hand, one may treat this as an implication that students preferred 3D representations and were willing to spend their time to learn with 3D. However, it is also possible that students felt perplexed by 3D related features and thus spent more time on trying the task of 3D manipulation, but which may not really promote the acquisition of spatial visualization skills. Roughly speaking, there is still a gap between observing how students behave and understanding what students think or whether students learn in this case. Rigorous hypothesis testing regarding this issue would require further studies recording and analyzing students' sophisticated usage behaviors in a multiple representations environment like this. However, analyzing and addressing the implications of behavioral data in terms of underlying cognitive functions, e.g., spatial aptitude or even learning, are challenging research topics (Jonsson et al., 2005). Recent research initiatives developing analysis techniques like educational data mining (Jonsson et al., 2005) and log analyses (Koedinger & Mathan, 2004; Wang et al., 2005) may help answer our research question in the future.

The scenario of inhomogeneity of regression described in the section of data analysis is considered as a clue suggesting that different versions of media design influence learners in different ways. From the descriptive statistics shown in Table 1, we can take a closer look at this situation. It is interesting to find that there were inverse inclinations between the two groups. One may initially conjecture that the post-test scores may naturally become higher than the pre-test scores due to the effect of impression or getting familiar with the test. However, this is not the fact observed in this study. For the 2D group, the mean of post-test scores is even worse than the mean of pre-test, whereas for the 3D group, the result shows positive increase on post-test. The scenario happened to the 2D group was similar to the situation reported by Branoff and Connolly (1999) that participants involved in such pre-test/post-test experimental design may not always score higher on the post-test. Specifically, when the process of pre-test/treatments/post-test is contiguous within a short period of time like this study, it is suspected that participants may easily feel bored on the post-test and unmotivated, and not give their best efforts. From this point of view, such result might also imply participants' different attitudes underlying the two groups. Since the treatment of the two groups was both conducted in the same classroom at the same period of time, the positive increase of post-test scores, though limited, revealed by the 3D group is likely to be valuable. However, due to the lack of measurements indicating students' attitudes toward the learning activities or even the experimental activities at the meta-level, the current study cannot establish the connection between attitudes and performance in the spatial visualization test. In following studies, it would be interesting to investigate the relationships between these two factors. Specifically, Huk et al. (2003) indicated that students' spatial ability may influence their attitude toward some learning activities; while our data seemed to imply that learning activities based on different media representations may take some effect on students' attitudes toward the post-test. Mayer and Massa (2003) have established the relationship between spatial aptitude and learning preference in which spatial ability is correlated with students' multimedia learning preference (e.g., verbal learning or visual learning). Our hypothesis, which is informed by Mayer and Massa (2003), conjectures that students' spatial aptitude, learning preference, motivation and the context they situated are mutually interfered. Studying the statistical interaction of spatial aptitude by learning-achievement or even the interaction of spatial aptitude by learning-preference by motivation by learning-achievement will further clarify related issues, contribute to the study of ATI effect and produce generalizable results regarding the use of spatial ability in adaptive instruction.

At a higher level, as proposed in the introduction, the study attempted to shed some light on the question of what kind of multimedia learning will be an effective one in terms of instructional design for spatial visualization skills. The current study, with the statistical insignificant result, remains inconclusive of whether additional features of the 3D representation, such as 3D navigation, interactivity, and simulation bring students additional learning benefits. Regarding the media and method argument (Clark, 1994; Clark, 2001), it is inconclusive whether 3D features are necessary instructional methods for spatial ability. Also, given the result that observable increase on spatial ability scores for students in both groups is small, whether spatial ability is trainable by our multimedia tutorials is not evidenced in this study. However, a medium effect size of the difference between groups on adjusted post-test, different using behaviors, and the possibly disparate attitude underlay students were found or inferred.

The results from this study, though seems to be limited by its small sample size, also provided some valued implications for future works. Three major findings of this study are: (1) No statistical significance was found between students' post-test scores in two groups, but a medium effect size was observed. Future replications of study with a large sample size may be able to derive statistically significant result for the same hypothesis. (2) Students in the two groups are likely to possess different browsing behaviors based on the time duration they spent in using the system. It merits further studies to measure and analyze students' behavior in the context of multimedia learning. Recent research endeavors of behavior analyses for computer-based learning may provide helpful examples to explain behavioral data in terms of cognitive functions (Jonsson et al., 2005; Koedinger & Mathan, 2004). (3) The differential direction of the gain scores (i.e., post-test score subtract pre-test score) for the 2D and 3D groups seems likely to suggest students' attitudes toward the post-test was influenced by learning activities. This hypothesis requires further studies with reliable measures and adequate experimental design to detect the aptitude-treatment interaction (ATI) effect.

6. Conclusion

As a pilot study, the comparative experimental evaluation with 23 participants has been conducted. Though the result was not statistically significant ($p = 0.307$), but a practical significance regarding the difference between the two groups via the ANCOVA analysis has been observed ($f = 0.234$, a medium effect size). The result implies approximately that different modalities of media representations (2D and 3D) are likely to influence students in different ways. By removing the research limitation of a small sample size in this study, replications of studies with more samples will be beneficial to address related issues in a clearer manner. Moreover, it would be interesting to look at the differences in 2D and 3D media representations in terms of web usage behavior, attitude, engagement and fun. This series of studies also aim to shed the light on the issue of how to employ technologies to help learners learn in an effective and joyful means.

Acknowledgements

This study is supported in part by National Science Council, ROC under contracts No. NSC 94-2524-S003-014 and NSC 93-2213-E-004-001. The authors thank the editor and anonymous reviewers for their insightful comments on an earlier version of the manuscript.

References

Alias, M., Black, T. R., & Gray, D. E. (2002). Effect of instructions on spatial visualization ability in civil engineering students. *International Educational Journal, 3*(1), 1–12.

Bodner, G. M., & Guay, R. B. (1997). The purdue visualization of rotation test. *The Chemical Educator, 2*(4).

Branoff, T., & Connolly, P. E. (1999). The addition of coordinate axes to the Purdue spatial visualization test-visualization of rotations: a study at two universities. In *Proceedings of the American Society of Engineering Education Annual Conference and Exposition*.

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

Chang, C.-Y. (2003). Teaching earth sciences: should we implement teacher-directed or student-controlled CAI in the secondary classroom? *International Journal of Science Education, 25*, 427–438.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*, 445–459.

Clark, R. E. (1985). Confounding in educational computing research. *Journal of Educational Computing Research, 1*, 137–148.

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

Clark, R. E. (2001). What is next in the media and methods debate? In R. E. Clark (Ed.), *Learning from Media* (pp. 327–337). Greenwich Conn: Information Age Publishers Inc.

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* (2nd ed.). Hillsdale, NJ: Erlbaum.

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*(2), 23–32.

Eyal, R., & Tendick, F. (2001). Spatial ability and learning the use of an angled laparoscope in a virtual environment. In *Proceedings of the annual medicine meets virtual reality conference* (pp. 146–151).

Huk, T., Steinke, M., & Floto, C. (2003). The influence of visual spatial ability on the attitude of users towards high-quality 3D animations in hypermedia learning environments. In *Proceedings of the world conference on e-learning in Corp., Govt., Health., and Higher Education* (pp. 1038–1041). Phoenix, Arizona, USA.

Jonsson, A., Johns, J., Mehranian, H., Arroyo, I., Woolf, B., Barto, A., et al. (2005). Evaluating the feasibility of learning student models from data. *Proceedings of the Workshop on Educational Data Mining at AAAI-2005*.

Koedinger, K. R., & Mathan, S. (2004). Distinguishing qualitatively different kinds of learning using log files and learning curves. In *Proceedings of the ITS 2004 Log Analysis Workshop*. Maceio, Brazil.

Koehler, P. K. (2003). FastScript3D: a JavaScript companion to Java3D. In *Proceedings of Annual Conference of ACM SIGGRAPH*.

Kwon, O. N. (2003). Fostering spatial visualization ability through web-based virtual-reality program and paper-based program. *Lecture Notes in Computer Science, 2713*, 701–706.

Kyllonen, P. C., & Lajoie, S. P. (2003). Reassessing aptitude: introduction to a special issue in honor of Richard E. Snow. *Educational Psychologist, 38*(2), 79–83.

Lajoie, S. P. (2003). Individual differences in spatial ability: developing technologies to increase strategy awareness and skills. *Educational Psychologist, 38*(2), 115–125.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: an experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*(4), 484–490.

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*(4), 833–846.

McLean, J. E., & Ernest, J. M. (1998). The role of statistical significance testing in educational research. *Research in the Schools, 5*(2), 15–22.

Najjar, L. J. (1996). Multimedia information and learning. *Journal of Educational Multimedia and Hypermedia, 5*(2), 129–150.

Olkun, S., (2003). Making connections: improving spatial abilities with engineering drawing activities. *International Journal of Mathematics Teaching and Learning*, Center for Innovation in Mathematics Teaching, University of Plymouth, UK.

Paivio, A. (1990). *Mental representations: a dual coding approach*. New York: Oxford University Press.

Potter, C., & Merwe, E. (2001). Spatial ability, visual imagery and academic performance in engineering graphics. In *Proceedings of the international conference on engineering education*. Oslo/Bergen, Norway.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.

Tretiakov, A., Kinshuk, & Tretiakov, T. (2003). Designing multimedia support for situated learning. In *Paper presented at the IEEE international conference on advanced learning technologies* (pp. 32–36). Athens, Greece.

Wang, H.-C., Li, T.-Y., & Chang, C.-Y. (2004). Adaptive presentation for effective Web-based learning of 3D content. In *Proceedings of the IEEE international conference on advanced learning technologies* (pp. 136–140). Joensuu, Finland.

Wang, H.-C., Li, T.-Y. Chang, C.-Y. (in press). A web-based tutoring system with styles-matching strategy for spatial geometric transformation. *Interacting with Computers*.

Wang, H.-C., Lu, C.-H., Yang, J.-Y., Hu, H.-W., Chiou, G.-F., Chiang, Y.-T., et al. (2005). An empirical exploration of using Wiki in an English as a second language course. In *Proceedings of the fifth IEEE international conference on advanced learning technologies* (pp. 155–157). Kaohsiung, Taiwan.

Woolf, B., Romoser, M., Bergeron, D., & Fisher, D. (2003). Tutoring three-dimensional visual skills: dynamic adaptation to cognitive level. In *Proceedings of the eleventh international conference on artificial intelligence in education*. Sydney, Australia.

Hao-Chuan Wang, M.S., is now a research assistant of National Taiwan Normal University, Taiwan. He received his M.S. degree in Computer Science from the Department of Computer Science at National Chengchi University, Taipei, Taiwan in 2004. His main research interests include adaptive hypermedia systems, technology-enhanced learning and human–computer interaction.

Chun-Yen Chang, Ph.D., is a professor and chair in the Department of Earth Sciences at the National Taiwan Normal University, Taipei, Taiwan. His major research interests include computer-assisted instruction, problem-solving models, inquiry teaching, and pedagogical understandings and conceptual change instruction in the area of science education.

Tsai-Yen Li, Ph.D., is an associate professor in the Computer Science Department of National Chengchi University, Taipei, Taiwan. Professor Li received his M.S. and Ph.D. from Stanford University in 1992 and 1995, respectively. His current research interests include motion planning, computer animation, virtual environment systems, intelligent user interface, autonomous agents, and e-learning systems.