Journal of Computer Assisted Learning (2003) **19**, 9-20

Supporting scientific discovery learning in a simulation environment

D.J. Reid, J. Zhang & Q. Chen
The University of Manchester, Tsinghua University, Beijing & Beijing Normal University

Abstract Until recent times, most studies on supporting simulation-based scientific discovery learning adopted the *ad hoc* strategies-oriented approach. This paper makes a systematic analysis of the internal conditions of scientific discovery learning to propose a triple scheme for learning support design that includes *interpretative support (IS)*, *experimental support (ES)*, and *reflective support (RS)*. The experiment was conducted with 78 students (aged from 12 to 13 years) to examine the effects of the IS and ES using a 2x2 between-subjects design. The main results were: significant effects were observed for IS on the post-test of intuitive understanding, flexible application and knowledge integration; no main effect was demonstrated for ES, and there was a marginally significant interactive effect for ES and IS on the intuitive understanding test. A process analysis showed that the successful learners had designed more well-controlled experiments than the failing ones. Learning support in a simulation environment should be directed toward the three perspectives to invite meaningful, systematic and reflective discovery learning.

Keywords: Control group; Discovery learning; Learning environment; Physics; Secondary; Simulation

Introduction

In the past decade, the research on discovery learning has evolved from *concept discovery learning* towards more sophisticated and *authentic scientific discovery learning* characterised by the need to design scientific experiments (van Joolingen & de Jong, 1997). Since computer simulation has the capacity to provide learners with an exploratory learning environment, it has been regarded as a powerful tool for scientific discovery learning (SDL). A growing number of studies have focused on SDL through computer simulation within a constructivist paradigm. However, as was pointed out in the reviews by Lee (1999) and de Jong & van Joolingen (1998), a lot of research comparing the effects of simulation-based learning to more traditional modes of learning finds little persuasive evidence in its favour. The question arises, why does simulation-based learning, that involves learners in active inquiry, not improve learning outcomes more consistently? One explanation lies in the wide

Accepted 25 June 2002

Correspondence: Jianwei Zhang, Educational Technology Centre, Tsinghua University, Beijing 100084
Email: zhangjw@tsinghua.edu.cn

range of difficulties learners may have in dealing with discovery learning processes. De Jong & van Joolingen (1998) classified the difficulties that learners may encounter into four categories:

- difficulties in generating and adapting hypotheses;
- poorly designed experiments;
- difficulties in data interpretation, and
- problems regarding the regulation of discovery learning.

Despite its potential in stimulating constructive learning activities, the simulation-based learning environment cannot guarantee effective learning without sufficient support ('scaffolding') for discovery learning activities.

In order to promote effective discovery learning, a number of studies have been conducted to help learners with particular strategies from specific aspects of the learning processes. For example, some researchers developed supportive methods to help generate hypotheses in simulation-based discovery learning (Shute & Glaser, 1990; Njoo & de Jong, 1993; Quinn & Alessi, 1994). Others have looked at the issues connected with experimental design (Leutner, 1993), planning (Tabak *et al*., 1996) and access to an appropriate knowledge base (Lewis *et al*., 1993). So far, most studies have adopted an *ad hoc* support strategies-oriented approach, intended to propose pieces of specific support according to learners' difficulties in particular aspects and examine the effects of the proposed learning support. No study was found to have made a systematic analysis of the internal conditions that determine the effectiveness of SDL and depict a proper scheme for support design in context.

Scientific discovery learning is a typical form of constructive learning based on problem solving activities involving the design and implementation of scientific experiments. Scientific discovery is usually interpreted as the processes of mindful coordination between hypothesised theories and evidence collected by experiments (Klahr & Dunbar, 1988; Kuhn *et al*., 1992). SDL is a knowledge construction approach that is based on scientific discovery activities. Three main interlocked spheres exist in the processes of effective SDL (see Zhang, 2000):

- problem representation and hypothesis generation, which heavily relies on the activating and mapping of prior knowledge and the meaning-making activities;
- testing hypotheses with valid experiments; and
- reflective abstraction and integration of the discovery experiences.

Taking all these perspectives into account, it is hypothesised that three interrelated conditions may determine the effectiveness of SDL to a great extent. These are:

- *The meaningfulness of discovery processes*: Learners need to activate their prior knowledge and map that onto the problem being addressed to help representing the problem and generating appropriate hypotheses and understandings.
- *The systematicy and logicality of discovery activities*: Effective discovery learning involves proper scientific reasoning, systematic manipulations of the variables, and qualified designs and implementations of experiments.
- *The reflective generalisation* over the discovery processes, which means the self-monitoring of the discovery processes and the reflective abstraction and integration of the discovered rules and principles.

The four categories of difficulties that learners may encounter during SDL, which have been summarised by de Jong & van Joolingen (1998), can all be attributed to the limitations in these three conditions. According to the three hypothesised conditions, three types of learning support can be designed and geared towards the three spheres:

- *interpretative support (IS)* that helps learners with knowledge access and activation, the generation of appropriate hypotheses, and the construction of coherent understandings;
- *experimental support (ES)* that scaffolds learners in the systematic and logical design of scientific experiments, the prediction and observation of outcomes, and the drawing of reasonable conclusions, and
- *reflective support (RS)* that increases learners' self-awareness of the learning processes and prompts their reflective abstraction and integration of their discoveries.

So far, most studies designed to support SDL focus on the impacts of certain specific support strategies, most of which are directed towards systematic, logical experiment and discovery activities (e.g. Rivers & Vockell, 1987; Njoo & de Jong, 1993; Tabak *et al*., 1996). More studies need to be carried out to examine the effects of the three types of learning support to propose a comprehensive scheme for learning support design. Within the triple scheme of learning support design for simulation-based SDL, an experiment was set up to examine the effect of the Reflective Support (Zhang, 2000). The purpose of the present study was to make an experimental investigation on the effects of interpretative and experimental support on simulation-based SDL. Based on the above theoretical analysis, it was expected that the experimental support should be able to enhance learners' systematic and valid experimental activities and hence manifest prominent effects on the discovery of the underlying rules. The interpretative support should be able to increase the meaningfulness of the discovery processes and hence promote the understanding, integration and flexible application of the discovered rules.

Methodology

The simulation learning environment

The topic chosen for the simulation was floating and sinking where the learners were required to explore the upthrust on objects floating in water. Their task was to discover which one or more of three given factors (shape, mass and volume) were related to the size of the upthrust on a floating object. Learners often hold misconceptions about this phenomenon, assuming that the size of the upthrust depends on the shape or volume of the object. Actually, the upthrust equals the weight of the object when it is floating because these two forces (weight and upthrust) are balanced whilst the object remains stationary. The size of the upthrust depends only on the mass of the object. This topic is one of the core issues in secondary science learning and has the structure of a scientific discovery task.

The simulation adopted paired-instance design that requires learners to construct a pair of experiments at a time, so that they could contrast the outcomes of two instances conveniently. For example, in order to examine the effect of the volume of object, learners can drag two objects of the same shape (e.g. ball) to the top of the left and right container, set the values on the left and right to keep their masses the same, and vary the volumes (see Fig. 1). Then they can click the 'RUN' button to see whether the upthrusts will be different or not. For all the learners, a data sheet was provided on screen to record and display the value of the input and output variables in each pair of experiments. A permanent button 'Main Steps' was prepared to remind learners of the main steps in an experiment, which involve selecting and dragging objects, deciding their values, running the experiments, observing outcomes, and clicking the 'NEW' button to start another trial. The learning environment contained two kinds of support: experimental and interpretative.

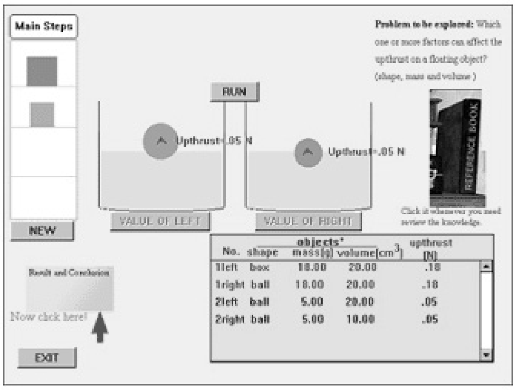

Fig. 1. Interface of the simulation

Experimental Support (ES). The ES included four specific treatments in order to help learners conduct systematic and valid experiments. These treatments were: explanations about scientific experimental design: in the introductory phase of the simulation, the programme gave learners the general explanations about scientific experimental design (particularly about 'varying one thing at a time'); identification of the objectives of each trial (before designing each pair of experiments, learners were required to identify their objectives by ticking the variables they wanted to examine; predictions and comparison (learners were required to predict which of the two specified objects would have a larger upthrust before running the experiments, and to check their predictions against the outcomes after the experiments); finally, conclusions of their new discovery against an experiment structure table showing the comparisons of the input and output variables between the two objects in a pair of experiments. Learners could access the last three supports by clicking a button (e.g. the 'Result and Conclusion' button in Fig. 1).

Interpretative Support (IS). The IS, which was intended to support learners to conduct meaningful discovery learning and generate coherent understandings, consisted of three measures:

- Activating prior knowledge: A multiple-choice question was offered in the introductory section of the simulation programme to activate learners' prior knowledge about balanced forces, which asked '*for an object floating in water, which force or forces are acting on it?*'
- General analysis of the problem: a multiple-choice question requiring students to select the factor(s) that are relevant to upthrust (without feedback) was given prior to the discovery processes to prompt the students to make a general analysis of the factor(s) that are relevant to the size of the upthrust on a floating object.
- Access to knowledge base: Helping learners access to the relevant knowledge during the discovery learning processes by providing a permanent button, *reference book*, which contained the descriptions of the concepts of weight, balanced forces, motion, as well as the basic meaning of upthrust.

The simulation programme was written in such a way that it registered learners' manipulations during the learning processes and wrote a logfile for each student.

Research design

In order to investigate the effects of ES and IS, a 2 (ES/no ES) by 2 (IS/no IS) between-subjects design was used to compare four versions of basically the same simulation environment: ES but no IS, IS but no ES, ES & IS, and no support.

Logfiles were used to analyse how learners proceeded in using the ES and IS.

Subjects

Subjects were 78 boys from Year 8 (age from 12 to 13 years) of a suburban boys comprehensive school in the UK. They were randomly distributed in four groups: ES but no IS (n = 20), IS but no ES (n = 20), ES & IS (n = 20), and no support (n = 18). A pre-test indicated no significant difference among the four groups in their topic knowledge, background knowledge, or experience with computers.

Tests

In order to gauge the effects of the learning supports on various aspects and levels of learning outcomes, four aspects of the outcomes were assessed in the post-test.

Principle knowledge. This was assessed by seven multiple-choice items. One item focused on the general principle about the factor(s) that can affect the upthrust on the object floating in water. The others concerned specific principles underpinning the phenomenon.

Intuitive understanding. Five multiple-choice questions were designed to measure learners' intuitive understanding, which is regarded as an important goal in SDL (de Jong *et al*., 1999; Swaak & de Jong, 2001), especially when conceptual change is desired. Using pictures, these items showed pairs of objects with different combinations of shapes, masses, and volumes and asked learners to predict how their upthrusts would compare in size.

Flexible Application. Eight multiple-choice items were written to determine how well learners could transfer the knowledge to new situations (e.g. the upthrust on a boat floating in a lake). These questions were more flexible, requiring the adaptation and integration of learners' knowledge. Flexible application is one of the crucial objectives of constructive learning, however, it seldom appears in studies of SDL.

Integration of knowledge. Two types of items were developed to assess associations between the discovered rule and learners' prior knowledge, especially the key concept of balanced forces. The first type took the form of 'association-rating' questions that are one of the response formats in concept map tests (Ruiz-Primo & Shavelson, 1996). Learners were required to indicate whether each of the concepts was related to their understanding about upthrust or not, and to give a brief justification if they responded with 'yes'. Two concepts including 'weight' and 'balanced forces' were set as target items, and three other concepts were included as filler items. Both learners' ticking and justification were used to decide the associations in their knowledge. The second type of item involved four 'instance-clustering' tasks (similar to Chi *et al*., 1981). A picture showing a beach ball floating in a swimming pool was set as the prototype instance. The other four instances were displayed to ask the learners to identify if each of the instances was similar to the prototype instance and explain why they thought so. The knowledge association was inferred according to whether they clustered the instances on the basis of the deep structures (force pattern) or merely in term of their surface features (e.g. moving or stationary, being in water or on a desk). Since the test of knowledge integration is relative new in format, reliability analysis was conducted particularly for this test, which revealed an acceptable internal consistency coefficient (α = 0.69).

Three items in each of the first two categories (principle knowledge and intuitive understanding) of the post-test were used in a pre-test to examine learners' prior topic knowledge about upthrust. The pre-test also encompassed the items about relevant background knowledge and their experience of using computers.

Procedure

After a pilot study in a second school, the pre-test was administered to all students in written format one week before the formal experiment. The study took place in a computer laboratory equipped with 20 networked Pentium computers. The students were required to accomplish the following tasks individually:

- *Warm-up.* Students worked with a tutorial version of the simulation programme. Two researchers were present to answer their questions about the programme. This stage lasted approximately 10 minutes.
- *Problem presentation*. The students were asked to explore which one or more of the factors among shape, mass and volume are related to the upthrust on an object floating in water. A brief description of the problem was available on the top-right corner of the screen throughout the discovery process.
- *Exploration*. Students were reminded that their task was to discover the rule on the basis of sufficient evidence through simulated experiments.

The post-test, also in written format, was administered immediately after the completion of the exploration. A total of 30 minutes was allotted for the post-test.

Results

Comparison between the pre-test and post-test results

As has been mentioned in the method section, the pre-test examined learners' prior topic knowledge about upthrust using the six items coming from the post-tests of principle knowledge and intuitive understanding. The mean scores of the four groups of learners on the pre-test and the post-test are presented in Table 1.

Table 1. Achievement on topic knowledge

Condition		Pre-test *s.d.*	Pre-test *m*	Post-test *s.d.*	Post-test *m*	Effect size
ES	IS	1.90	2.10	7.20	2.46	2.32
	no IS	2.47	2.18	6.47	2.70	1.64
No ES	IS	2.89	2.08	6.44	2.79	1.46
	no IS	2.94	2.14	5.76	2.73	1.16
Total		2.53	2.12	6.50	2.66	1.66

The repeated measures MANOVA, using the ES and IS as between-subjects factors, indicated a significant improvement on the post-test scores ($F_{1,68} = 88.41$, $p < 0.001$). There was also a marginally significant interaction between the tests and ES, $F_{1,68} = 3.07$, $p = 0.08$. The students with the ES made a greater improvement compared to those without the ES.

The effects of the ES and IS on the post-test

Table 2 shows the means and standard deviations of different groups on the four categories of the post-test.

Table 2. The means and standard deviations of four types of post-tests

Condition		Principle knowledge *m*	Principle knowledge *s.d.*	Intuitive understanding *m*	Intuitive understanding *s.d.*	Flexible application *m*	Flexible application *s.d.*	Integration of knowledge *m*	Integration of knowledge *s.d.*
ES	IS	4.00	2.70	6.80	1.77	4.50	1.47	4.75	3.19
	no IS	2.60	2.41	4.90	2.38	3.50	1.64	2.40	1.96
No ES	IS	3.75	2.65	5.70	2.27	4.85	1.36	4.58	2.83
	no IS	3.50	2.43	5.56	2.61	3.56	2.06	3.11	2.08
Total		3.72	2.51	5.74	2.33	4.12	1.77	3.71	2.72

Principle knowledge. Students with the ES scored a little higher on the principle knowledge test than those without the ES, and the same trend existed for IS (see Table 2). However, there was no significant main effect or interaction for ES or IS.

Intuitive understanding. Table 3 shows the result of the ANOVA for this test. There was a significant main effect with IS ($p < 0.05$). The two groups with the IS outperformed those who didn't receive this support (see Table 2). There was a marginally significant interaction between IS and ES ($p = 0.09$). IS indicates a significant effect among the students with the ES ($F_{1,75} = 7.08$, $p = 0.01$), whilst no such effect was found among the no ES groups ($p > 0.10$) (see Fig. 2).

Table 3. ANOVA of ES and IS on intuitive understanding test

Source of variance	*SS*	d.f.	*F*	*Sig.*
ES	1.18	1	0.23	0.63
IS	21.30	1	4.13	0.04
ES x IS	14.99	1	2.91	0.09
Residual	381.64	74		

Flexible application. The fourth column in Table 2 gives the achievements of the groups on this test. A two-way ANOVA showed a very significant main effect for IS, $F_{1,74} = 8.85$, $p < 0.01$, which was in favour of the role of the IS. There was no significant effect concerning ES or the interaction between ES and IS ($p > 0.10$)

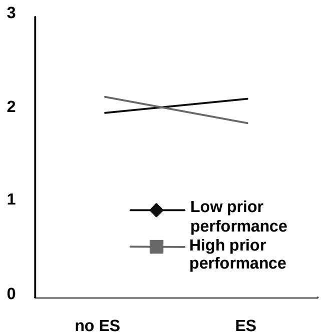

Fig. 3. The interaction of ES and prior science performance level on Index II

The integration of knowledge. For the items in this category, a full score of 2 was given when students' judgement and justification for an item were both clearly in favour of the knowledge association. A score of 1 indicated that the student had made a correct judgement but given only a vague justification. The maximum score was 12. As anticipated, IS indicated a quite significant positive main effect on the knowledge integration test (see Table 2), $F_{1,73} = 10.79$, $p < 0.01$. Again, no significant effect was found for ES or the interaction between ES and IS ($p > 0.10$).

Process analysis

Using the data provided by the logfiles, an analysis was made to see how the students had interacted with the simulation environment and how they had used the supports.

Number of experiments and time on exploration. A maximum time of 35 minutes and a minimum number of five pairs of experiments had been set for the exploration session. On average students conducted 10.83 (*s.d.* = 5.63) pairs of valid experiments. The students who received the ES performed significantly fewer experiments than those without the ES, $F_{1,73} = 20.91$, $p < 0.01$. There was no significant difference concerning IS ($p > 0.10$). The average time spent on the exploration was 23.26 (*s.d.* = 6.93) minutes. Significant main effects were found regarding ES and IS on the time spent on exploration ($F_{1,74} = 10.42$, $p < 0.01$ for ES and $F_{1,74} = 3.47$, $p < 0.10$ for (IS). However, there was no significant correlation between the post-test scores and the time on exploration or the number of experiments ($p > 0.10$). All the trends reported were also observed when the ANOVAs were performed using the time and the number of experiments as covariates. So it can be inferred that the effects of

ES and IS on the post-tests cannot be ascribed to the differences in time and number of the experiments performed by the individuals.

Frequencies of using the supports. The use of the noncompulsory supports was analysed. The mean frequency of using the '*Reference Book*' (a component of the IS) was 1.63, and the mean time spent on it was 80 seconds. Students chose to summarise their discovery (a component of the ES) in 78.64% of their experiments. There was no significant correlation between the post-test scores and the frequencies of using these supports ($p > 0.10$).

Evaluation of learners' experiments. Using the data collected in the logfiles, an analysis was made to see how the students had designed their experiments. 'Change one thing at a time' is an important principle in scientific experiment. Unfortunately, learners often vary many variables in one experiment (Glaser *et al*., 1992; de Jong & van Joolingen, 1998). Focusing on the principle of variable control in scientific experiment, three indices were adopted to evaluate the learners' experiment designs.

Index I: The percent of well-controlled experiments. As has been mentioned in the methodology section, this simulation used paired-instance design that required learners to conduct a pair of experiments at a time, so that they could contrast the outcomes of two instances. Index I indicates the percentage of the well-controlled experiments in which one and only one factor (shape, mass, or volume) was different between the left and right side instances (see Fig. 1 as an example).

Index II: Average number of variables varied in each pair of experiments. This is a looser criterion displaying how many variables among shape, mass and volume were varied in each pair of experiments on average.

Index III: Focused examination of the three variables. A pair of experiments were identified as having undergone a 'focused examination' of certain variable if that variable was the only variable varied in that pair of experiments. For each variable, a full score of 2 was given when it had been examined by at least two pairs of experiments at different levels of the controlled variables. Score 1 indicated that the variable had been examined by only one pair of experiments or by more than one pair of experiments but at the constant levels of controlled variables. Sequentially, score 0 meant that no experiment had been focused on this variable at all. An average score across the three variables was used in the final analyses. Among the three indices, Index III is the strictest in that it was the only index that indicated the distribution of well-controlled experiments across the three variables.

In comparing the experimental quality of the ES to the no-ES groups, no significant difference was found in the *T*-tests using ES as the grouping variable ($p > 0.05$). In order to analyse the effect of ES among high and low level students, the students' scores in a recent examination in science class were collected. The top and bottom 25 students were sorted out and formed the high and low science achievement group. Table 4 gives the evaluation of their experiments. ANOVAs of the three indices, using ES and science achievement level as independent variables, displayed a significant interaction on Index II ($F_{1,44} = 4.48$, $p<0.05$) (Fig. 3). Simple effect analysis demonstrated that ES had a marginally significant effect among the high science achievement students ($F_{1,45} = 4.07$, $p = 0.05$). Those with the ES varied much fewer variables in each pair experiments, whilst no significant effect was found among the low science achievement students. ES and science achievement level had no significant effect on Index I and III.

Table 4. The indices evaluating learners' experiments

Groups		Index I *m*(%)	*s.d.*	Index II *m*	*s.d.*	Index III *m*	*s.d.*
ES	High	36.80	0.23	1.86	0.39	0.71	0.54
	Low	34.00	0.12	2.12	0.21	0.61	0.24
No ES	High	28.30	0.15	2.15	0.33	0.70	0.25
	Low	33.33	0.20	1.97	0.41	0.70	0.26
Total		33.53	0.18	2.01	0.36	0.68	0.36

In order to explore the relation-ship between the quality of experiments and the result of learning, the students were put in two groups according to whether they had discovered the correct rule by the end of the experiment. This could be identified by their responses to one of the items in the principle knowledge test. Table 5 gives the evaluation of the experiments conducted by the successful and failure group. The successful group exceeded the failure group on all the three indices, with significant or marginally significant differences on Index I ($p = 0.01$) and Index III ($p < 0.10$).

Comparing the post-tests and the experiment design, the score of the principle knowledge test was significantly correlated to Index I (Pearson $r = 0.24$, $p < 0.05$) and Index II (Pearson $r = -0.25$, $p < 0.05$). There was also a marginally significant correlation between the principle knowledge test and Index III (Pearson $r = -0.20$, $p < 0.10$).

Table 5. Differences between groups on three indices

Groups	Index I *m*(%)	*s.d.*	Index II *m*	*s.d.*	Index III *m*	*s.d.*
Success ($n = 44$)	36.44	0.19	1.94	0.40	0.76	0.41
Failure ($n = 32$)	26.03	0.15	2.06	0.31	0.59	0.36
Significance	$T(75) = 2.54$ $p = 0.01$		$T(74) = -1.45$ $p = 0.15$		$T(75) = 1.86$ $p = 0.07$	

Note: 2 students were excluded because they didn't respond to this item.

Discussion

This study examined the effects of experimental support and interpretative support in simulation-based SDL. As was demonstrated by the pre-test and post-test scores, learners of the four groups all benefited significantly from the simulation-based learning processes. As far as the effects of the treatments are concerned, there was no significant effect for ES or IS on the post-test of principle knowledge. This seems to be consistent with the opinion that a formalised knowledge test is not a sensitive indicator of the effect of simulation-based learning (de Jong *et al.*, 1999). The IS manifested prominent main effects on the tests of intuitive understanding, flexible application and knowledge integration. The expected interactive effect between IS and ES was observable on the intuitive understanding test. It seems that the IS can foster learners' intuitive understanding when the ES is also available. This is congruent with the statement of Okada & Simon (1997) that students need to participate actively in both crucial experiments and explanatory activities in order to discover the right mechanism.

Overall, the function of the IS has been demonstrated as highly significant in this study. The IS was designed to support discovery learning by activating the relevant knowledge in the learners' memories, enhancing the problem representation and hypothesis generation based on their existing background knowledge, eliciting more explanation activities toward the experiments, and promoting the access of knowledge base. This has been found to be helpful for learners to construct a more elaborate, coherent and deep understanding about the explored domain. Hence, on the intuitive understanding test, the students who received the IS could come up with a more accurate insightful intuition in specific situations. On the knowledge integration test, the IS groups exhibited more elaborate associations between their understanding of upthrust and their prior knowledge about balanced forces. When confronted with novel problem situations in the application test, these students could generate more appropriate solutions by generalising and transforming the principle derived from their experiments. For instance, learners' explorations were focused on the upthrust of objects floating in water. One of the questions in the flexible application test asked: *Two objects are placed in two beakers half full of two kinds of liquids: water and oil, and they are both floating. How will the upthrusts on them compare?* It was more likely for students in IS groups to generalise their discovery to this situation because they could assimilate it as a new instance of the more fundamental conception, balanced forces.

All these results support one of the major assumptions in the study: the integrative meaning-making perspective plays a critical role in the SDL process, and hence it should be one of the key targets for instructional support. As Ausubel *et al*. (1978, p. 519–564) argued, the effect of discovery learning depends on the meaningfulness of the discovery experience. SDL does not end with the discovery of one or two pieces of rule, but is intended to incorporate the findings into learners' profound, elaborate, and coherent knowledge structures, and help learners developing their own 'ideas' on the scientific phenomena. To a large extent, these learning outcomes will rely on learners' explanatory and interpretative meaning-making activities. Apart from the access and activation of relevant knowledge which has already been addressed in some studies (Leutner, 1993; de Jong *et al*., 1999), instructional support in this sphere needs to do more to foster learners' explanatory and interpretative activities, such as representing the problem in relation to the relevant knowledge, explaining the experimental outcomes using their fundamental knowledge, and synthesising their discovery and generating a integrative understanding.

The overall effect of the ES was not so clear as that of the IS. An effect was only observed in the comparison of the post-test with the pre-test, and also somewhat in the interaction with IS on the intuitive understanding test. As can be seen from the process analysis, the ES groups outperformed the no-ES groups on two of the three indices evaluating their experiments, but with no significant difference being found. The ES in this study included a number of elements such as of the explanation about experiment design (especially 'varying one thing at a time'), the prompts about identifying the objective of each experiment, predicting and observing outcomes, and summarising their discoveries. However, these treatments were still not supportive enough to improve the learners' experimental activities or learning outcomes. This result disagrees with the findings of Rivers & Vockell (1987) that providing learners with general experimentation hints before their exploration could promote their experimentation abilities, as well as the conclusion of Swaak *et al*. (1998) that the experimental support in the form of assignments had clear effect on SDL. One possible explanation is that the students in the present study were 12–13-year-olds, whilst most of the other studies in this field were performed among college or high school students. In an earlier survey conducted in UK, Archenhold *et al*. (1988) documented that it was difficult for 13-year-olds to control the variables properly during the investigation process. A relevant result in the present study is the interaction between ES and science achievement level on Index II, which implies that learners with high science achievement could benefit more from the ES. Another possible reason is that selecting the supports and answering the questions during the learning processes might have distracted learners from their investigation and therefore cancelled the effect of the ES.

However, the importance of qualified experimental activities was verified from another aspect by the analysis of the relationship between the quality of experiment and the discovery result. Learners who had discovered the right rule surpassed the failing learners on all the three indices of experiment design. There was also a clear correlation between the indices and the principle knowledge test. As is emphasised by most researchers in this field, the perspective of scientific reasoning and experimental activities does count a lot in the SDL process. The outcome depends heavily on learners' reasoning and experiment activities such as generation and adaptation of hypothesis, systematic and focused manipulation of input variables, prediction and observation of outcomes, and drawing conclusions from experiments. The inefficiencies in the above activities can hamper successful discovery.

Conclusion

In conclusion, this study proposed a triple scheme of learning support design for simulation-based SDL and examined the effects of the experimental and interpretative support. The overall results support the importance of the meaningfulness and the logicality of the discovery processes, and imply that learning support, either bedded within simulation software or provided by a human tutor in classroom environment, should be directed toward these perspectives to invite meaningful and systematic discovery learning. However, it is important to note that there are many uncertain problems behind the triple learning support scheme for simulation-based discovery learning. More repetitive examinations of the three types of learning support are needed. Especially, future research needs to re-design the treatments in the ES and re-examine its effect on simulation-based discovery learning among learners with different capabilities.

Acknowledgements

The work reported here was part of the outcomes of an Academic Link between Beijing Normal University and the University of Manchester supported by British Council. The authors thank the British Council (Beijing office) for their support and efficient work and also Professor Margaret Cox of King's College, London, and the two anonymous reviewers for their valuable comments and criticism.

References

Archenhold, F., Bell, J., Donnelly, J., Johnson, S. & Welford, G. (1988) *Science at Age 15: a Review of APU Survey Findings, 1980–1984*. HMSO, London.

Ausubel, D.P., Novak, J.D. & Hanesian, H. (1978). *Educational Psychology: a Cognitive View*. Holt, Rinehart and Winston, New York.

Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science*, **11**, 2, 121-152.

Glaser, R., Schauble, L., Raghavan, K. & Zeitz, C. (1992) Scientific reasoning across different domains. In *Computer –Based Learning Environments and Problem Solving* (eds. E. de Corte, M.C. Linn, H. Mandel & L. Verschaffel), pp. 345–373. Springer-Verlag, Berlin.

de Jong, T., Martin, E., Zamarro, J., Esquembre, F., Swaak, J. & van Joolingen, W.R. (1999) The integration of computer simulation and learning support: An example from physics domain of collisions. *Journal of Research in Science Teaching*, **36**, 5, 597–615.

de Jong, T. & van Joolingen, W.R. (1998) Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, **68**, 2, 179–201.

van Joolingen, W.R. & de Jong, T. (1997) An extended dual search space model of scientific discovery learning. *Instructional Science*, **25**, 5, 307–346.

Klahr, D. & Dunbar, K. (1988) Dual space search during scientific reasoning. *Cognitive Science*, **12**, 1, 1–48.

Kuhn, D., Schauble, L. & Garcia-Mila, M. (1992) Cross-domain development of scientific reasoning. *Cognition and Instruction*, **9**, 4, 285-327.

Lee, J. (1999) Effectiveness of computer-based instructional simulation: a meta analysis. *International Journal of Instructional Media*, **26**, 1, 71–85.

Leutner, D. (1993) Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional supports. *Learning and Instruction*, **3**, 2, 113–132.

Lewis, E.L., Stern, J.L. & Linn, M.C. (1993) The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology*, **33**, 1, 45–58.

Njoo, M. & de Jong, T. (1993) Exploratory learning with a computer simulation for control theory: learning processes and instructional support. *Journal of Research in Science Teaching*, **30**, 8, 821–844.

Okada, T. & Simon, H.A. (1997) Collaborative discovery in a scientific domain. *Cognitive Sciences*, **21**, 2, 109–146.

Quinn, J. & Alessi, S. (1994) The effects of simulation complexity and hypothesis generation strategy on learning. *Journal of Research on Computing in Education*, **27**, 75–91.

Rivers, R.H. & Vockell, E. (1987) Computer simulations to simulate scientific problem solving. *Journal of Research in Science Teaching*, **24**, 5, 403–415.

Ruiz-Primo, M.A. & Shavelson, R.J. (1996) Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, **33**, 6, 569–600.

Shute, V.J. & Glaser, R. (1990) A large scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments*, **1**, 1, 51–77.

Swaak, J. & de Jong, T. (2001) Discovery Simulations and the Assessment of Intuitive Knowledge. *Journal of Computer Assisted Learning*, **17**, 3, 284–294.

Swaak, J., van Joolingen, W.R., Wouter, R. & de Jong, T. (1998) Support for simulation-based learning: The effects of model progression and assignments on definitional and intuitive knowledge. *Learning and Instruction*, **5**, 3, 235–252.

Tabak, I., Smith, B.K., Sandoval, W.A. & Reister, B.J. (1996) Combining general and domain-specific strategic supports for biological inquiry. In *Intelligent Tutoring Systems* (eds. C. Frasson, G. Gauthier & A. Lesgold), pp. 288–297. Springer-Verlag, Berlin.

Zhang, J. (2000) *Simulation-based Scientific Discovery Learning: The internal conditions and the facilitative learning supports*. A PhD Dissertation. Beijing Normal University.