

A Microworld Learning for Psychology Experiments by Combining Real and Virtual Experiments

Kazuhisa Miwa, Norio Ishii, Hitomi Saito, and Ryuichi Nakaike

{miwa, ishii, hitomi, nakaike}@cog.human.nagoya-u.ac.jp
Graduate School of Human Informatics, Nagoya University
Nagoya, 464-8601 JAPAN

Abstract: In this paper, we propose a learning environment for psychology experiments by combining real and virtual experiments. Eighteen undergraduate students were actually enrolled in a class whose design was based on our learning program. By analyzing the participants' learning processes in the class, it was shown that our program worked well as a learning environment for psychology experiments that included various types of activities such as hypothesis formation, experiment planning, and interpretation of the experimental results. Comparing the results from the pre- and post- tests, we found that the participants learned to be guided, through their learning activities in this environment, to deeper understanding of the research objectives dealt with in the microworld. They also acquired knowledge about research methods, such as experiment designing skills, in order to understand research objectives. These benefits seemed to be brought about by the correspondence of the participants' problem solving activities in the real experiment to the training in the microworld learning.

Keywords: inquiry approach, psychology education, microworld learning

1. Introduction

The effectiveness of the inquiry approach has been well known for instructing students in scientific activities. In the approach, students are required to solve a given problem where they are placed in the role of a practicing scientist. There have been many trial applications based on this approach in scientific education; in this paper, we apply the inquiry approach in experimental psychology education.

When trying to do so, one problem is the cost of performing psychology experiments where many human subjects are needed to collect data. To solve this problem, we may use a microworld learning environment. Shunn and Anderson discussed the differences between general domain-independent and specific domain-dependent knowledge for conducting psychology experiments using a simulated psychology laboratory called SPL (Shunn & Anderson, 1999). They also proposed that such a simulated lab. could be used as a learning environment for tutoring experiment planning skills (Shunn & Anderson, 2001). However, their studies did not discuss the relationship between a computational microworld and the class design in which the microworld was used.

A key point for making the approach function effectively is how the learning environment and curricula are designed. When a microworld is used in a classroom based on the above approach, the issue of how microworld learning is performed in the total design of the class seems crucial. In this paper, we present an example of class design that places microworld learning in the context of global learning activities in the class.

The main theme set in this class was the psychology of human scientific discovery. In the class, issues related to laboratory studies of scientific discovery were discussed. As a learning activity in the class, the participants performed exercises of psychological experiments using Wason's 2-4-6 task, which has been widely used as an experimental task throughout the history of laboratory studies. The participants performed the experiments using a microworld learning environment developed by the authors instead of real psychological experiments in which human subjects participate.

Goals of the class

The two goals of the class were assumed as follows.

(1) Understanding the nature of human hypothesis testing strategies: Many issues are dealt with in the laboratory studies of scientific discovery. One of the most important issues is hypothesis testing strategy (Klayman & Ha, 1987). Humans tend to collect positive data when testing their hypotheses. This tendency is known as the positive test strategy. One goal of this class was to understand the possibilities and limitations of this bias appearing in human discovery processes.

(2) Learning skills for designing psychological experiments: The first goal above is related to scientific content. The second goal is to learn skills when doing the science to achieve the first goal. There are also many different kinds of

skills for doing science; this class focuses on CVS (Control Variable Strategy), which is the most important skill for designing psychological experiments.

Principles of class design

The following describes two core ideas for establishing the class design.

(1) Combining real and virtual experiments: Real experiments and virtual experiments in the microworld are combined to improve the effects of learning with the microworld. We intend to provide the participants with realistic situations in the virtual experiment given by the microworld based on their participation in the real experiment prior to its manipulation by the microworld.

(2) Gradual introductions of research objectives: Based on the ideas of the inquiry approach, the participants investigate the process of solving Wason's 2-4-6 task using the microworld. It is difficult for novice participants to engage in the entire mission. Therefore, we bring about progressive learning stages where first participants are required to solve a mission given by an instructor and then to move to an inquiry phase in which the participants find problems by themselves and solve them based on their own ideas.

2. Task, Environment, and Participants

2.1 Task

The experiment task set in this learning environment was a traditional discovery task, Wason's 2-4-6 task. Table 1 shows an example process of solving the 2-4-6 task. The standard procedure of the 2-4-6 task is as follows. All subjects are required to find a "target rule" of a relationship among three numerals. In the most popular situation, a set of three numerals, "2, 4, 6", is presented to subjects at the initial stage. The subjects form a "hypothesis" about the regularity of the numerals based on the presented set. The subjects then produce a new set of three numerals as an "experiment" and present it to the experimenter. The experimenter gives a Yes as feedback to the subjects if the set produced by the subjects is an instance of the target rule, or a No as feedback if it is not an instance of the target rule. The subjects continuously carry out experiments, receive feedback from each experiment, and search to find the target. The subjects propose a final hypothesis whenever they think they know what the rule is and they receive feedback on whether or not the hypothesis is correct.

Table 1: An example process of Solving Wason's task.

hypotheses	experiments	
-	2, 4, 6	Yes
continuous evens	4, 6, 8	Yes
continuous evens	20, 22, 24	No
divisors of 24	8, 8, 8	Yes
divisors of 24	18, 100, 2	No
divisors of 24	8, 6, 4	Yes
divisors of 24	24, 12, 8	No
single digits	1, 1, 5	Yes

2.2 Two important concepts

There are two important concepts dealt with in this class.

The nature of targets: we categorize the target rules used in our experiments from the viewpoint of their generality. We define targets as general targets if the proportion of their members (positive instances) to all instances (all sets of three numerals) in the search space is large. On the other hand, we define targets as specific targets if the same proportion is small. An example of the former type of target is "the product of three numerals is even" (where the proportions of target instances to all possible instances is 7/8) and an example of the latter type is "three evens" (where the proportion is 1/8).

Hypothesis testing: There are two types of hypothesis testing: a positive test and a negative test. The positive test is conducted in an instance where the subject expects there to be a target. That is, the positive test is a hypothesis test using a positive instance for a hypothesis. The negative test is, in contrast, a hypothesis test using a negative instance for a hypothesis. For example, if a hypothesis were about "ascending numbers", the positive test would use a sequence like "1, 3, 9"; the negative test would use a sequence like "1, 5, 2". In Table 1, the subject conducted 6

positive tests by using instances, "4, 6, 8", "20, 22, 24", "8, 8, 8", "8, 6, 4", "24, 12, 8", and "1, 1, 5", and one negative test by "18, 100, 2".

2.3 Microworld

The microworld used in this study was a computerized microworld called VPL (Virtual Psychology Laboratory), developed by the authors. VPL installs an agent, constructed as a production system model, to solve Wason's task. The participants' mission was to identify the relationship between the experimental factors and the performance of the agent's problem solving ability while manipulating factors such as the agent's problem solving strategies and cognitive capacities. In VPL, the following six parameters can be controlled by the participants: target rules used in experiments, hypothesis testing strategies, hypothesis formation strategies, the number of activated instances in the working memory, the number of maintained hypotheses in the working memory, and the conditions for terminating the experiments. More detailed descriptions of these factors can be seen in our preceding reports (Miwa et al., 2001; 2002). Figure 1 shows an example snapshot of VPL. The "Controller" manages the starting and ending of simulations and the appearance of each window. The participants set up experimental factors in the "Input Window". The "Simulation Window" presents a real time process of a production system solving the 2-4-6 task. The "Result Window" shows the final result of each simulation. The "Summary Window" summarizes the experimental results obtained by the preceding simulations.

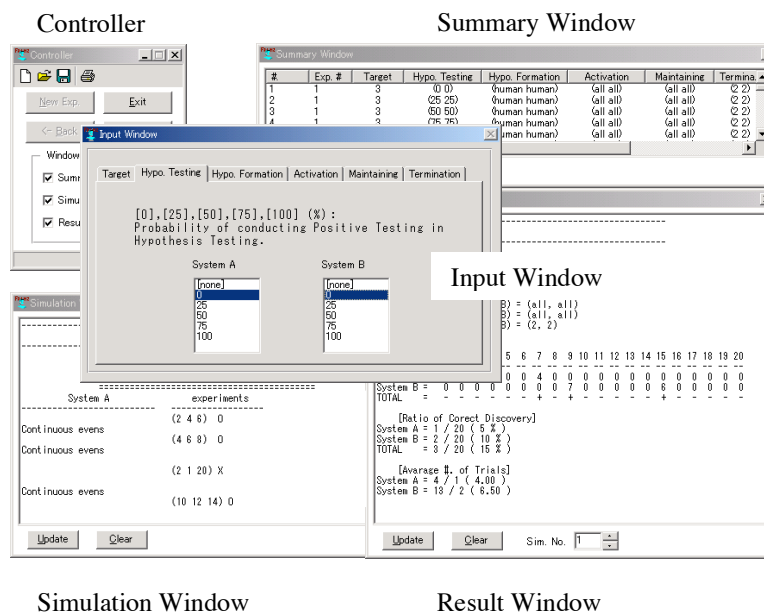


Figure 1: System's interface.

2.4 Participants

Eighteen undergraduate students participated in this class. The participants did not major in psychology; they were involved in school of science or school of engineering.

3. Learning Activities

In this learning program, the participants initially take part in a real psychology experiment in which they solve Wason's task by themselves. The task is also solved by the agent in VPL. Then, based on the experience, they perform virtual psychology experiments, manipulating the agent in VPL. In the following, we indicate four main phases of this learning program (see Figure 2) and the summary of the participants' learning activities.

Phase 1: Participation in a Real Experiment

First, the participants take part in a real psychology experiment and actually solve Wason's task. Each of the eighteen participants was required to locate two types of target rules. The target used in the first task was "the product of three numerals is 48", and that in the second task was "three different numbers." The former is an example of a specific target, and the latter is a general target. In the first task, eight of the eighteen participants found the correct target; in the second task, seven participants found it.

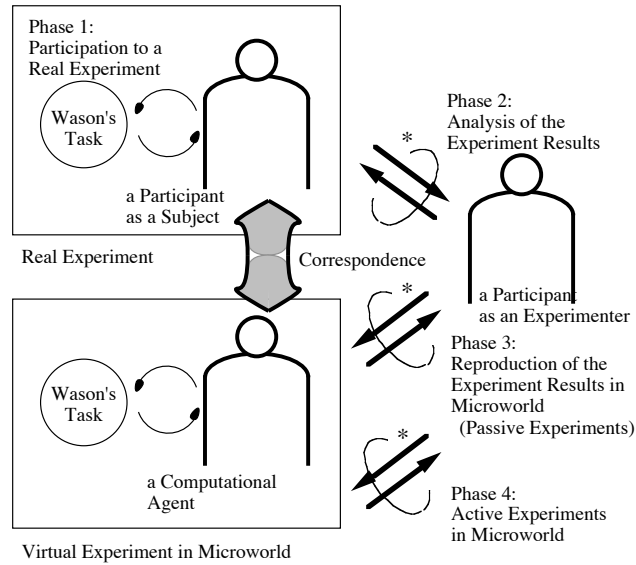


Figure 2: Four phases of the learning program.

Phase 2: Analysis of Experimental Results

Each participant receives the experimental results (i.e., the data from problem solving processes) of all participants including him/herself. They then analyze the psychological data as instructed by the tutor. That is, the participants were required to calculate, in each of the correct and incorrect solution groups, the proportion of positive tests while solving Wason's task. The result is shown in Figure 3. Sixteen of the eighteen participants took part in Phase 2. All sixteen participants indicated that positive tests are effective for finding a specific target, whereas they are ineffective for finding a general target by depicting a similar figure to Figure 3.

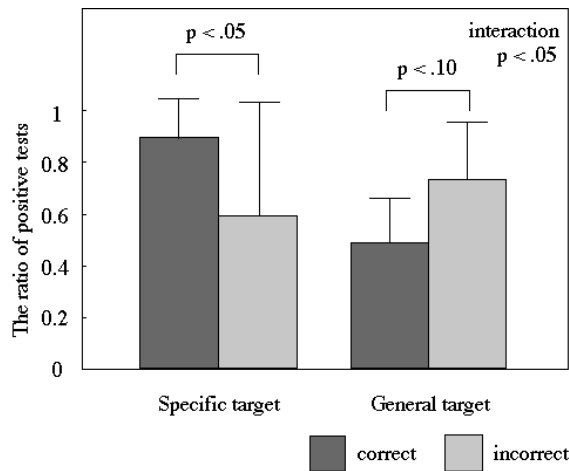


Figure 3: A real experiment in Phase 2.

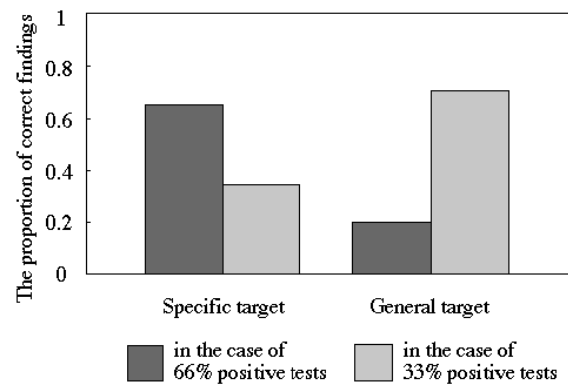


Figure 4: A virtual experiment result in Phase 3.

Phase 3: Reproduction of Experimental Results in a Microworld

The participants confirm the findings obtained in Phase 2 by manipulating the parameters of the agent in VPL. In Phase 3, the participants train themselves to manipulate VPL, and learn to understand the similarity between the real experiments they participated in and the virtual experiments in VPL.

The participants were instructed to confirm the findings indicated in Phase 2 by manipulating two of the agent's parameters: one relates to hypothesis testing strategies (the proportion of positive tests) and the other is for selecting tested targets (each target has its own generality). All those who participated in Phase 3 confirmed the common tendency responsible for the agent's performance (i.e., interaction between the two factors found in Phase 2) in the

virtual experiments in VPL. Figure 4 shows a sample result obtained by a participant, S7, from virtual experiments in VPL. We note here that (1) the problem solving performance (such as the proportion of correct findings) is obtained in VPL based on real simulations in which the agent in the microworld actually solves the problem, and (2) the topic discussed here (i.e., the interaction between targets' generality and hypothesis testing strategies) is not an artificial one, but has actually been investigated throughout the history of cognitive psychology of human scientific discovery (Gorman, 1992; Newstead & Evans, 1995).

Phase 4: Active Activities in a Microworld

Phase 4 is the inquiry learning phase. In Phase 4, the participants first individually set up their experimental objectives, then investigate the objectives while manipulating VPL. In Phase 4, the participants were required to reveal the relationship between the 6 parameters describing the agent's behavior (see 2.3) and the agent's problem solving performance (the proportion of finding a correct target rule) while manipulating VPL. We held two class sessions, each of which lasted for an hour and a half. After the two sessions, we returned the results of the experiments that each participant had conducted, and instructed them to report their findings (must not exceed 4 items) with supporting graphs.

Table 2 shows the factors identified by each of the participants that explain the agent's performance. The numbers in Table 2 indicate the order of importance of each finding as reported by the participants. In addition, the findings marked with a "*" denote interaction between two factors. For example, one participant, S1, pointed out that the most important factor determining the performance is the sixth parameter, which terminates the search, while the second most important factor is target generality, which is controlled by the first parameter. We excluded from Table 2 the cases where the graphs were inconsistent with the reported findings. We also excluded two subjects whose VPL manipulations included some seriously confusing processes.

Table 2: Findings identified by the participants.

	Target	Hypothesis testing	Hypothesis formation	# of activated instances	# of maintained hypotheses	Terminating the search	Others
S1	2					1	
S2			2, 4				
S3			1, 2				
S4	1*					1*, 2, 3, 4	
S5		1, 2					
S6		1					
S7	2*		1, 2*			3	
S8			1*, 2, 3				1*
S9		2, 3				1	
S10		1, 2, 3	1, 2				
S11		2		1*, 4	1*, 4	4	
S12							
S13	4		2	1		3	
S14					1		
S15			1, 2, 3		4		
S16				1	2		

As seen in Table 2, the participants detected various types of findings in Phase 4. However, almost all findings were valid when analyzed normatively: e.g., as the working memory capacity increases, the performance gets higher. This validity is given because in VPL, as mentioned earlier, the performance is given based on the results of the agent actually solving the problem. Psychology experiments include various types of processes such as hypothesis formation, experiment designing, and interpretation of experimental results (Klahr, 2000). Table 2 implies that VPL provides a learning environment where the participants are not forced to plan uniform experiments based on adhoc previously-given hypotheses, but permitted to experience spontaneous experiments based on various hypotheses individually formed by the participants themselves.

4. Comparison of Pre- and Post- tests

As mentioned in Section 1., we set up two goals: understandings science content and learning skills for doing science. In this section, we estimate each type of learning by comparing the pre-test with the post-test.

4.1 Understandings of human hypothesis testing

First, to estimate the participants' improvement in understanding the class topics, we tested to what extent the participants can understand the relation between their own strategies used in their problem solving processes and their performance in Phase 1 based on their findings obtained through the learning activities from Phase 2 to Phase 4.

The pre-test was performed between Phase 1 and Phase 2, and the post-test was performed after Phase 4; the results of both tests were later compared.

In each test, an experimental sheet was presented to each participant, on which was given each participant's problem solving process in the second task (finding a general target) in Phase 1. The participants who found the correct target were required to describe reasons why they reached the correct solution, while the participants who failed were required to state the reasons why they did not.

Table 3 shows which concepts each participant used to describe the explanations. Each concept is indicated and divided into two groups: the six factors controlled in VPL, and others. The former six factors are well-defined psychological concepts describing human problem solving processes whereas the latter factors are ambiguous concepts.

Table 3: Reasons for success or failure in finding the correct targets.

	Pre-test									Post-test										
	Manipulated Factors					Other Factors				Manipulated Factors					Other Factors					
	Target	Hypo. Testing	Hypo. Form.	# of Instance	# of Hypo.	Search	Hypo. revi.	Various Verifi.	Discon.	Others	Target	Hypo. Testing	Hypo. Form.	# of Instance	# of Hypo.	Search	Hypo. revi.	Various Verifi.	Discon.	Others
S1																				
S2						○	○	○			○	○				○			○	
S3							○				○	○				○				
S4		○									○	○					○			○
S5								○	○		○	○						○		
S6						○					○	○								
S7							○				○	○								
S8	○			○					○		○	○								
S9						○					○	○					○			
S10				○									○							
S11								○			○	○		○	○					
S12									○		○	○						○		
S13							○				○	○					○			
S14						○			○							○	○			
S15								○								○				
S16									○											○
Total	1	1	2	0	0	4	4	4	3	3	8	12	1	1	1	4	2	2	3	2

Table 3 shows that in the post-test, almost all participants explained the reasons for their success or failure using well-defined concepts, such as the hypothesis testing strategy and the targets' generality, which are two factors related to the class topics. Such explanations were not seen for the pre-test results.

The above results show that the participants learned to use psychological concepts dealt with in the microworld in a real context, such as understanding problem solving processes on their own. This means that our microworld learning brought about deep understandings of the research subjects for the participants.

4.2 Learning skills for designing psychological experiments

Next, we discuss knowledge acquisition for cognitive skills necessary for planning psychology experiments. Did the participants learn to employ more sophisticated experimental strategies through performing virtual experiments in VPL? To verify the effect, we conducted another set of pre- and post- tests using the following task.

The reproduction of bacteria is considered to be influenced by various factors such as temperature, humidity, illumination, pressure, and so on. The previous studies have suggested that the reproduction of a certain species of bacteria becomes more active at high temperatures. However, a new phenomenon has been observed where the reproduction of the bacteria is still active even when the temperature is low. In this situation, plan your experiments to detect the factors responsible for the reproduction of the bacteria.

This experiment requires the detection of interaction among multiple factors.

Experiment strategies described by the participants in each test were categorized into the following seven types.

Trial & Error: No organizational experiments were indicated.

All Combination: When considering n factors, all combinations of the factors (2^n combinations) were tested.

VOTAT: Only one factor was manipulated while other factors were fixed. This is regarded as the VOTAT (Vary One Thing At a Time) strategy (Chen & Klahr, 1999).

Multiple VOTAT: The VOTAT experiments were conducted with multiple factors. For example, first, temperature was searched using VOTAT, followed by a VOTAT search for humidity.

Fixing: While fixing a factor at one level (e.g., while maintaining low temperature), another factor (e.g., humidity) was manipulated.

Multiple Fixing: First, while fixing a factor (A) at one level, another factor (B) was manipulated. Then, while moving the fixed level in A to another level, the previously manipulated factor (B) was similarly manipulated (e.g., first, while maintaining low temperature, the humidity was manipulated. When the temperature was increased, the humidity was manipulated again).

Two Factors: Two factors were simultaneously manipulated.

Table 4 shows the strategies used by the participants in the pre- and post- tests. Each cell of the table is shaded in one of three colors, based on the degree of validity described as follows: (1) Poor: The "Trial & Error" strategy is inappropriate for experiment planning, and the "All Combination" strategy is also unsuitable because the increase in experiment combinations prevents its actual execution. (2) Fair: The "VOTAT", "Multi-VOTAT," and "Fixing" strategies are also incomplete in this context, because they cannot detect interaction between multiple factors even though they are valid for general experiment designing. (3) Good: The "Multi-Fixing" and "Two Factors" strategies are suitable for this experimental objective.

Table 4 indicates that a considerable number of participants moved to more valid experiment strategies from the pre-test to the post-test, meaning that training in this learning environment provided the participants with more sophisticated skills for planning psychology experiments. When we assign scores 1, 2, and 3 to Poor, Fair, and Good strategies, the mean scores of the pre-test and post-test are 0.94 and 1.38 respectively. A paired t-test reaches significance ($p < 0.05$, $t(15)=2.78$).

Table 4: Changes in experiment strategies employed by the participants.

	Pre-test	Post-test
S1	Two Factors, Multi-VOTAT	Two Factors, Multi-VOTAT
S2	Multi-Fixing	Two Factors, Multi-VOTAT
S3	VOTAT	Two Factors, Multi-VOTAT
S4	All Combination	Two Factors, Multi-VOTAT
S5	Multi-VOTAT	Multi-Fixing, VOTAT
S6	Multi-Fixing	Multi-Fixing, VOTAT
S7	Fixing	Multi-Fixing
S8	Fixing, VOTAT	Multi-VOTAT, VOTAT
S9	Trial & Error	Fixing, VOTAT
S10	Fixing, VOTAT	Fixing
S11	Fixing	Fixing
S12	All Combination	Fixing
S13	Fixing	Fixing, Multi-VOTAT
S14	Fixing, Multi-VOTAT	Multi-VOTAT
S15	Multi-VOTAT	Multi-VOTAT
S16	Trial & Error	Trial & Error

5. Discussions and Conclusions

Finally, we discuss another aspect of benefits brought about by the combination of real and virtual experiments in microworld learning.

Miller et al. highlighted a problem in microworld learning: that superficial structures, such as a narrow, task-specific goal of a game or a cover story brought about to increase its entertainment value, often prevent the participants from learning the primary knowledge and skills originally intended to be acquired (Miller et al., 1999). In this case, the participants do not understand the essential learning contents because of superficial and mechanical manipulation of the microworld.

In psychology experiments using a microworld, similar problems may appear. In virtual experiments, the participants manipulate strategies and abilities of an agent in a microworld as independent variables. However, those variables, which are supposed to have cognitive entities such as strategies and abilities, may be manipulated mechanically as meaningless numeral parameters, and in this case deeper learning, originally intended by a tutor, is not likely to be achieved.

In our learning program, the participants take part in a real psychology experiment in which they solve by themselves the task dealt with in the microworld before they begin learning activities in the microworld. Participation in a real experiment enables the participants realize correspondence between their task-solving experience and a task that employs an agent on a computer to solve it. This benefit may result in positive effects that reach beyond the superficial problematic learning that should be carefully removed from microworld learning.

Correspondence has the following two functions (also see Figure 2).

(1) From Participant to Agent: The first function is the correspondence from the participant to the agent. That is, it is expected that the participants think of parameter control as a meaningful manipulation of the agent's cognitive functions by matching the parameters' functions to their own problem solving activities.

(2) From Agent to Participant: The second function is the correspondence from the agent to the participant. This enables the participants to analyze and interpret their own problem solving processes based on the results of the virtual experiments conducted in the microworld.

In the post-test, five of the sixteen participants enrolled in the class analyzed their problem solving processes based on the results of the virtual experiments. Let us use a participant S10 as an example. The participant, S10, detected two factors responsible for the problem solving performance, the hypothesis formation strategies and the hypothesis testing strategies, by controlling those two factors in the virtual experiments in Phase 4 (see Table 2). S10 could not find the general target in the experiment in Phase 1. In the post-test, he mentioned the reason for his failure as follows: "The order of hypothesis formation was from specific to general (in my own problem solving process in the experiment of Phase 1). In the simulations, the proportion of correct findings was 0% for this hypothesis formation strategy. I failed because I began with a very specific hypothesis." S10 actually tried to relate the agent's behavior in VPL with his own problem solving activities, and analyzed their activities based on findings obtained in the virtual experiments.

This implies that the VPL used in our learning program could function as an environment for experiencing meta cognitive activities where the participants are guided to perform self-reflective activities with their own problem solving processes.

References

- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, 70, 1098-1120.
- Gorman, M. (1992). *Simulating science: heuristics, mental models, and technoscientific thinking*. Indiana university press.
- Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, Mass.: MIT Press.
- Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211-228.
- Miller, C., Lehman, J., & Koedinger, K. (1999). Goals and Learning in Microworlds. *Cognitive Science*, 23, 305-336.
- Miwa, K. (2001). Emergence of effects of collaboration in a simple discovery task. *Proceedings of the 23rd annual conference of the cognitive science society*, 645-650.
- Miwa, K., Ishii, N., Saito, H., & Nakaike, R. (2002). Changes in Learners' Exploratory Behavior in a Simulated Psychology Laboratory. *Proceedings of the 24th annual conference of the cognitive science society*. 667-672.
- Newstead, S., & Evans, J. (Eds.). (1995). *Perspectives on Thinking and Reasoning*. UK: Lawrence Erlbaum Associates Ltd.
- Schunn, C. D., & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science*, 23, 337-370.
- Schunn, C. D., & Anderson, J. R. (2001). Acquiring expertise in science: explorations of What When and How. In C. Crowley, et al. (Eds.), *Designing for science: implications from everyday, classroom, and professional settings*. Mahwah, NJ: LEA.