

# 仮想心理実験室における学習者の行動変容過程

三輪和久  
石井成郎  
斎藤ひとみ  
中池竜一

名古屋大学大学院人間情報学研究科

464-8601 名古屋市千種区不老町

E-mail: {miwa, ishii, hitomi, nakaike}@cog.human.nagoya-u.ac.jp

あらまし コンピュータ上に仮想心理実験室（VPL: Virtual Psychology Laboratory）を構築した。VPLでは、発見科学に関する心理学において、これまで伝統的に用いられてきたWasonの2-4-6課題を、二人の被験者が解決する過程をシミュレートする。学習者は、Wason課題を用いて人間の協同問題解決を研究する実験心理学者として、授業に参加する。VPLを利用して、3回の授業を実施した。その結果、実験デザインの構造化、仮説の正確性、発見される解の一般性といった多様な側面で、顕著なパフォーマンスの向上が認められた。

キーワード 仮想心理実験室 学習 科学的発見 実験

## Transition Process of Learner's Exploratory Behavior in a Virtual Psychology Laboratory

Kazuhisa MIWA, Norio ISHII, Hitomi SAITO, and Ryuichi NAKAIKE

Graduate School of Human Informatics, Nagoya University  
Furo-cho, Nagoya, 464-8601 Japan

**Abstract** We have constructed a virtual psychology laboratory (called VPL) on a computer. VPL simulates the process of pair subjects collaboratively solving Wason's 2-4-6 task, which has been traditionally used in the field of the psychology of discovery science. Participants take part in the class as an experimental psychologist who studies human collaborative problem solving using Wason's task. We conducted three sessions of classes using VPL. As a result, we confirmed the improvement of various types of performances such as the construction of experimental design, the degree of correctness of hypotheses the participants formed, and the generality of targets they found.

**Key words** Virtual psychology laboratory, Learning, Scientific discovery, and Experimentation

## 1. はじめに

複雑なシステムの挙動を理解することは、科学的探究の典型的な目的の1つである。例えば、心理学では、人間を1つのシステムと考え、実験を通して、そのシステムの動作を規定する要因を明らかにしようとする。

心理学的実験を行うための知識は多岐にわたるが、もっとも重要なスキルとして、実験変数を統制する能力、CVS (the Control of Variables Strategy) があげられる。Klahrらは、おもちゃのビークルを制御するプログラムの文法の学習を題材として、小学生から大学生、大学院生に至る様々な被験者のCVSの能力を、実験的に明らかにしてきている (Klahr, 2000)。さらに、CVSの学習に関して、実験室で得られた知見を、実際の教育場面に適用することを試みている (Klahr, et al., 2001)。

さて、Shunnらは、SPL (Simulated Psychology Lab) と呼ばれるコンピュータ上に実現された実験環境を用いて、大学生と心理学研究者の実験遂行能力の比較を行い、心理学的実験を行うための一般的知識と領域固有知識の差異について論じている (Shunn, 1999)。Shunnらはさらに、このSPLを、実験計画スキルの学習環境として用いることを提案している。

しかし、SPLでは、あらかじめ2つの理論(仮説)が学習者に与えられ、学習者は、その2つの理論のうちいずれが正当であることを判別するような実験を計画することが要請される。従って、理論(仮説)自体を形成するプロセスは無視されている。また、システムは実際に人間の認知プロセスをシミュレートするのではなく、入力されたパラメータの値に基づいて、単純な関数によってパフォーマンスを出力する。

従って、実験環境のリアリティにはある一定の制限がある。

本論文では、このSPLに比して、より複雑で、かつリアルな実験環境を用いて、学習者の実験プロセスを観察した。以下では、本実験環境を、VPL (Virtual Psychology Laboratory) と呼ぶ。

## 2. 仮想心理実験室

### 2.1 VPL: Virtual Psychology Laboratory

VPLでは、二つのプロダクションシステムが、協同して、Wasonの2-4-6課題を解決する (Wason, 1960)。ここで強調しておきたいことは、VPLが扱うこのテーマ自体が、この数十年を通して、人間の協同的発見を研究する心理学者の間で実際に検討されてきた、非常に現実的な題材であるということである (Gorman, 1992; Laughlin, 1996; Newstead & Evans, 1995)。さらに、本シミュレータのパフォーマンスは、人間のパフォーマンスをよく反映したものであることが確かめられてきていることも重要である (Miwa, 2001)。

図1は、VPLのインターフェースを示している。「コントローラ」は、シミュレーションの開始や終了、各ウインドウの表示/非表示等を管理する。学習者は、「要因入力ウインドウ」から、実験要因を設定する。「シミュレーションウインドウ」には、2つのプロダクションシステムが、Wason課題を解決するプロセスが表示される。「結果ウインドウ」には、シミュレーションの最終結果が表示される。「サマリーウインドウ」には、それまでの実験結果の一覧が表示される。

表1は、シミュレータの操作要因の一覧を示す。

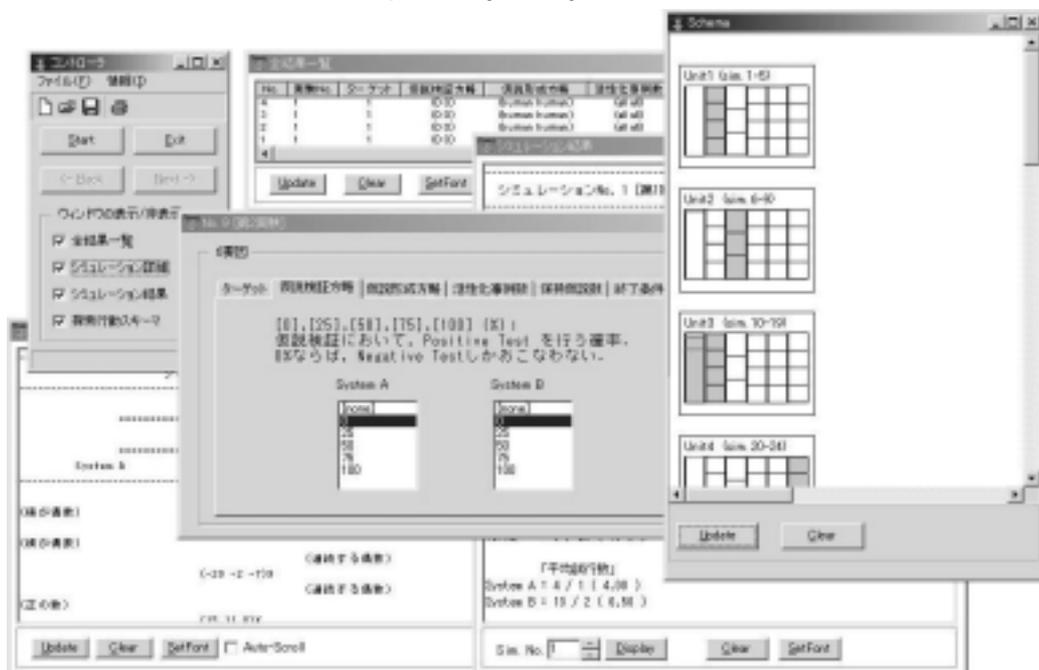


図1 システムのインターフェース

表 1 実験要因

Factors	Levels
Target [T]	[1] - [35] Thirty-five kinds of targets that were used in the experiment. For example, Target #1 is "ascending numbers"; Target #35 is "three different numbers".
Hypothesis testing strategies [HT]	[0], [25], [50], [75], [100] The probability of conducting positive tests in generating instances. [100] and [0] mean that systems always conduct positive tests and negative tests, respectively.
Hypothesis formation strategies [HF]	[human], [random], [specific], [general] [human] means that systems generate hypotheses as humans do. [random]: generating hypotheses randomly. [specific]: generating specific hypotheses prior to general ones. [general]: generating general hypotheses prior to specific ones.
# of activated instances [AI]	[all], [6], [5], [4], [3] The number of instances that can be activated at once in the working memory when generating hypotheses.
# of maintained hypotheses [RH]	[all], [5], [4], [3], [2] The number of rejected hypotheses that can be maintained in the working memory.
Condition for terminating the search [TE]	[all], [5], [4], [3], [2] The number of continuous confirmations when systems terminate the search. [2] means when a hypothesis is continuously confirmed two times, systems recognize the hypothesis as the solution, and terminate the search.

なお、実験では、活性化事例数と保持仮説数の水準は、それぞれ“all”に固定されており、参加者は、合計4つの要因を操作することができた。

シミュレータのパフォーマンスは、多様な要因により規定される。ターゲットの性質と仮説検証方略の間に顕著な相互作用があること、ワーキングメモリのキャパシティ（活性化事例数、保持仮説数）に主効果があることなどは、実際の心理学実験において得られてきた知見によく一貫している（Klayman, 1987）。

## 2.2 実験

**被験者:** 20名の大学生が、授業の一環として、実験に参加した。

**事前知識:** 参加者は、事前に、Wason 課題の手続き、およびこれらの発見課題を用いた実験室研究の意義について学習していた。

参加者は、事前に、著者らが作成した実験教材としての論文を読んだ。論文には、単独被験者が、Wason 課題を解決した実験の結果が示されていた。実験結果は、仮説検証方略とターゲットの性質の間に相互作用が存在することを示していた。参加者は、これらの知見を事前知識として、実験に参加した。

**手続き:** 実験は、1週間間隔で、3回にわたって行なわれた。各回の実験時間は、約1時間であった。参加者は、各授業の最後に、一連の実験を通して発見した知見を報告した。

参加者は、実験シートに、まず (1) 実験の目的(何を調べようとしているのか)、(2) 実験結果の予想、(3) 要因操作に関する実験計画を記入した後に、シミュレータを操作して一連の実験を

実行する。実行結果を得た後に、(4) 実験結果の考察を記入する。参加者は、この一連の手続きを繰り返す。

**プリ-ポストテスト:** 一連の授業の前後には、参加者の実験要因操作の基本的能力を測定するためのプリテスト、およびポストテストが行なわれた。

## 3. 実験結果

### 3.1 チャンキング行動

参加者は、図2に示された実験空間を操作して実験を行なう(前述のように、活性化事例数、保持仮説数の2つの要因は、それぞれ“all”に固定されている)。

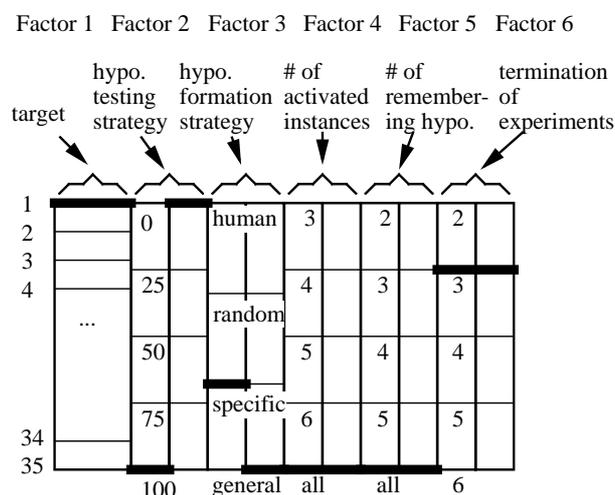


図2 実験空間

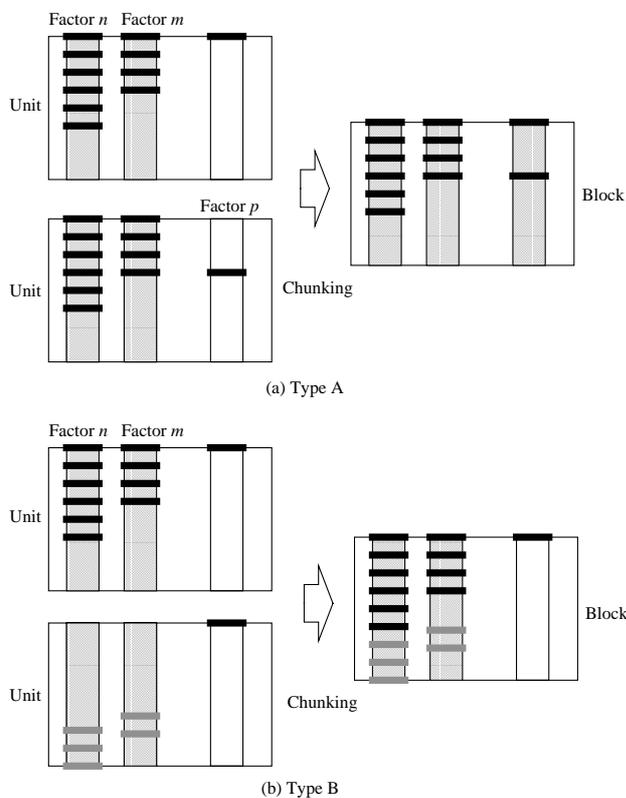


図3 チャンキング行動

参加者は、各要因の特定の水準の組み合わせによって行なわれる個々の実験をまとめて、すなわちチャンキングして、1単位の実験群を構成してゆくと考えられる (Miwa, 2000)。もっとも基本的なチャンクは、図3の左列に例示されるような、ある特定の要因、この例では、要因nと要因mを操作した一連の実験群である。このチャンクを、“ユニット”と呼ぶ。

次に、以下の二つの方法 (図3の(a)タイプAと(b)タイプB)により、複数のユニットをまとめて、より高次のチャンクを構成することができる。

まず、要因n, mをの操作を維持しながら、他の要因pの水準を一つずつずらしてゆくような行動である (図3(a)参照)。これらは、要因n, m, pの3要因実験として、一つのチャンクと見なすことができる。ここでは、要因n, m, p以外の要因が固定されているということが重要である。

次に、特定の要因n, mの全ての水準の全ての組み合わせを一度に探索できない場合には、それを複数のユニットに分解して探索することが考えられる (図3(b)参照)。この場合も、それらのユニットは一つのチャンクと見なすことができる。ここでは、要因n, m以外の要因が固定されているということが重要である。

このような2つの方法によって、複数のユニットからより大きなチャンクを構成することが可能である。このより高次のチャンクを、“ブロック”と呼ぶ。

ここで、ブロック数/実験数を、チャンキング

率と定義する。もし、個々の実験が完全にバラバラに行なわれれば、1実験=1ユニット=1ブロックとなり、チャンキング率は1となる。

例えば、図3(a)の例では、6(要因n)×4(要因m)×2(要因p)=48の実験が一つのブロックを構成しているため、チャンキング率は、 $1/48=0.021$ である。一方、図3(b)の例では、 $6 \times 4 \times 3 = 30$ の実験がブロックを構成しているため、チャンキング率は、 $1/30=0.033$ である。

チャンキング率が小さくなればなるほど、参加者は、その実験行動において、より大きなチャンクを構成することが可能であったことを意味する。すなわち、より小さなチャンキング率は、参加者がより構造的な実験を計画し、実施していたことを示唆している。

図4は、3回の全ての授業に参加した、20名中16名の参加者の各回の授業ごとのチャンキング率の平均を示している。分散分析の結果、クラスの主効果が優位であった ( $p < .01$ )。以上は、参加者が、実験を繰り返すことを通して、より大きなチャンクを構成することが可能になったことを示している。

### 3.2 要因操作

より大きなチャンクを構成してゆく様子は、参加者が操作した要因数の推移にも見てとることができる。

図5は、1要因、2要因、3要因以上が操作されていたブロックの、全ブロックに対する割合の平均を示したものである (参加者16名)。分散分析の結果、クラスと要因数の間に相互作用が存在する傾向が認められた ( $p < .1$ )。図5は、実験の進行に従って、参加者が、より多様な要因を組み合わせた実験を行なうようになっていったことを示している。

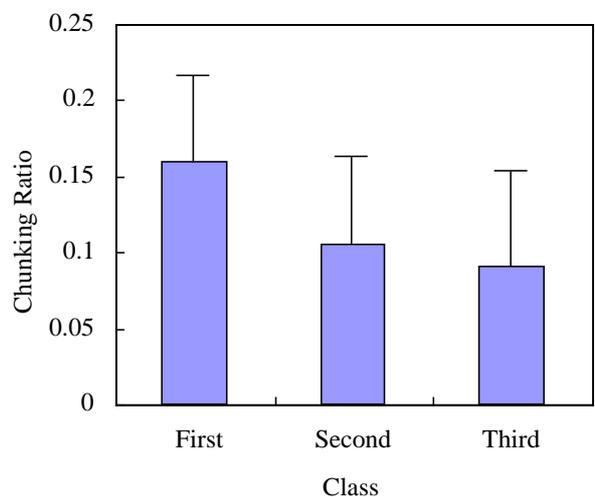


図4 チャンキング率の推移

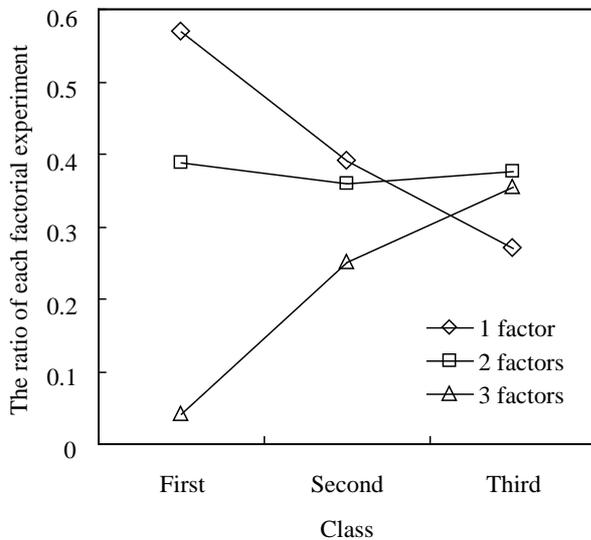


図5 実験要因数の推移

### 3.3 仮説

次に、参加者が形成した仮説に着目する。参加者は、一連の実験を行なう前に、実験計画シートに実験結果の予想を記入した。その際に、その予想の確信度を、1～5の5段階で評価している。

さらに、シミュレータを用いた実験の後、その実験結果の考察を記入する時に、合わせて予想的中度を、同じく5段階で評価している。

図6は、初回と3回目の授業における、実験実験前に記入された確信度の平均、および実験後に記入された的中度の平均である（参加者16名）。分散分析の結果、クラスと評価（確信度と的中度）の間に相互作用が存在した（ $p < .01$ ）。

図6より、初回から3回目の授業に移行するにつれて、確信度は改善されていないのに対して、的中度が改善されていることがわかる。確信度

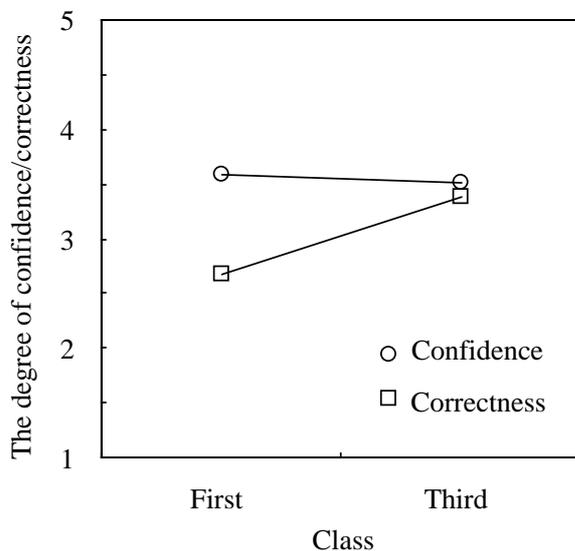


図6 仮説の確信度 / 的中度の推移

は、仮説に対する主観的蓋然性の評価であったのに対して、的中度は、仮説の客観的正確さを反映していると考えられる。

確信度が一定であることは、実験の前半と後半で、参加者が形成した仮説の内容が質的に変化していなかったことを示唆する一方で、的中度が改善されることは、より正確な仮説を形成することができるようになったことを意味していると考えられる。

### 3.4 解決

次に、実験を通して、参加者が発見した解に着目する。解をその一般性の観点から分類する。

ここでは、ある要因（もしくは要因群）とシステムのパフォーマンスの関係に言及しているものを、一般的な解と定義する。例えば、「特殊なターゲットの場合ポジティブテストが有効で、一般的なターゲットの場合には、ネガティブテストが有効である」は、ターゲットと仮説検証方略の二つの要因とシステムのパフォーマンスの関係に言及しているので、一般的な解の例である。

一方、ある要因の特定の水準とシステムのパフォーマンスの関係に言及しているものを、特殊な解と定義する。例えば、「ターゲット27に関しては、どの仮説形成方略を用いても、ネガティブテストが有効だった」は、ターゲットという要因の一つの水準に対する限定的な知見を述べているので、特殊な解の例である。

図7は、1回目の授業と3回目の授業における、特殊な解、および一般的な解の種類数の平均を示したものである。分散分析の結果、クラスと解の性質（一般・特殊）の間に相互作用が認められた。図7より、参加者は、実験が進行するにつれて、より一般的な解に到達してゆくことがわかる。

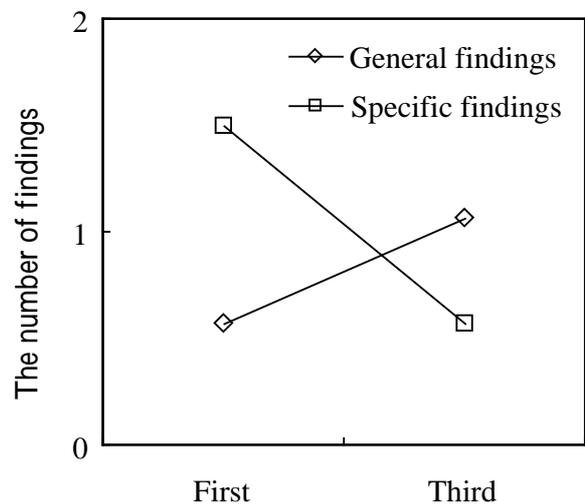


図7 発見された解の種類の変化

### 3.5 知識獲得

最後に、実験の前後に行なわれたプリテスト、ポストテストの結果を分析することにより、参加者が、実験計画に関する一般的な知識をいかに獲得したかを検討する。プリテストでは、参加者は、バクテリアが発生する要因（温度と湿度）を同定する実験を計画することが求められた。ポストテストでは、プランクトンの発生要因を同定する、同型の問題が用いられた。

テストに書かれた参加者の回答は、大きく二つのタイプに分類された。一つは、片方の要因を固定して他方の要因を変化させる（温度を一定にして湿度を変化させ、次に湿度を一定にして温度を変化させる）ことにより、バクテリアの繁殖への影響を見る方略。これを、“1要因組み合わせ実験”と呼ぶ。もう一つの方法は、二つの要因を同時にコントロールする“2要因実験”である。前者の方法では、二つの要因間の相互作用を発見することはできないが、後者はそれを見つけることができるという意味で、後者の方法の方が優れていると考えられる。

表2は、実験に参加した20名の参加者のプリテストとポストテストの成績である。1名は、明らかに混乱した回答を示していたので、除外してある。フィッシャーの正確検定の結果、度数の偏りに有意傾向が認められた ( $p < .1$ )

表2より、参加者は、VPLでの実験を繰り返し行なうことを通して、より適切な実験計画を行なうことができるようになったことがわかる。

### 4. 考察と結論

本実験では、参加者に対して、教師からの教授を一切行っていない。参加者は、他者からの評価なしに、繰り返し自分で実験を計画し、システムからのフィードバックを得ながら、3回の授業を経験した。にも関わらず、実験デザインの構造化、仮説の正確性、発見される解の性質といった多様な側面で、顕著なパフォーマンスの向上が認められることは興味深い。

このパフォーマンスの向上が、CVSに代表される、一般的な実験計画スキルの向上からもた

らされたものなのか、単に実験の進展に従って探索する問題空間に関する知識が増えたためにもたらされたものなのかは、判然としない。ただし、プリテストからポストテストの成績の向上は、何らかの意味で、一般的なスキルが獲得されたことを示唆している。

今後の課題としては、一つは、実験システムとしてのVPLの利用が考えられる。例えば、ノービスとエキスパートの実験プロセスの異なりを明らかにすることや、事前知識の影響等の研究を考えることが可能である。もう一つの発展は、学習システムとしてのVPLの利用である。例えば、ここで用いたチャンキングの考え方に基づき、実験プロセスを学習者にフィードバックすることにより、効果的な実験スキル学習支援環境を構築することなどを考えている。

### 参考文献

- 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.
- Klahr, D., Chen, Z., & Toth, E. (2001). From cognition to instruction to cognition: a case study in elementary school science instruction. In Crowley, K, et al. (Eds.), *Designing for science: implications from everyday, classroom, and professional settings*. NJ: LEA.
- Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211-228.
- Laughlin, P. (1996). Group decision making and collective induction. In Witte, E., & Davis, J. (Eds.), *Understanding group behavior*, Vol. 1. NJ: LEA.
- Miwa, K. (2000). Human Discovery Processes Based on Searching Experiments in Virtual Psychological Research Environment. *LNAI*, 1967, 225-239.
- 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.
- Newstead, S., & Evans, J. (Eds.). (1995). *Perspectives on Thinking and Reasoning*. UK: Lawrence Erlbaum Associates Ltd.
- Shunn, C., & Anderson, J. (1999). The generality /specificity of expertise in scientific reasoning. *Cognitive Science*, 23, 337-370.
- Wason, P. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly journal of experimental psychology*, 12, 129-140.

表2 プリテストとポストテストの成績

	2 factors	combinational 1 factor
Pre Test	7	12
Post Test	12	7