

グラフを利用した仮説検証プロセス学習支援システムの開発と評価

中池 竜一[†] 三輪 和久^{††}Development and Evaluation of a Tutoring System for Hypothesis
Testing Processes with GraphsRyuichi NAKAIKE[†] and Kazuhisa MIWA^{††}

あらまし 認知心理学領域における研究では、仮説検証プロセスは、仮説形成・実験計画・結果考察という三つの主要なプロセスに分けることができる。これら三つのプロセスの一貫性を維持することは、科学的に発見を行うために必要である。本研究では、一貫した仮説検証プロセスの学習を支援するために、4種類のグラフ利用活動を含む、計算機上にシミュレートされた実験環境を提案した。それは、(1) 仮説形成の後に結果予想グラフを描くこと、(2) 実験計画・実施後に得られた結果のグラフを描き、(3) システムから提示される模範的なグラフを閲覧すること、そして(4) 結果考察の前に、それらグラフの比較を行うことである。実験的な検討の結果、グラフの利用を通して仮説検証プロセスを練習することにより、一貫したプロセスが実施できるようになることが示唆された。

キーワード 学習環境、仮説検証、マイクロワールド学習、科学教育

1. ま え が き

実験を通じた仮説検証プロセスは、科学的な探求手法として一般的なものであり、認知心理学領域では、科学的発見の背後に存在する認知プロセスとして関心もたれている。

仮説検証プロセスは、大きく、(1) 仮説形成段階、(2) 実験計画段階、(3) 結果解釈段階に分けることができる[1]。既存の知識から仮説を形成した後、その仮説を検証するために必要な実験を計画する。そして、実験によって得た結果から、仮説が確認されたのかを判断する。このように相互依存する3段階のプロセスは、仮説検証プロセスを通じた発見のために一貫させることが重要となる。

この科学的発見のプロセスを学習させるための最も一般的な方法は、現実の実験室で実際にそのプロセスを体験させることであり、計算機上の仮想的なマイクロワールドで体験的に学習させる試みもいくつか存在

する[2], [3]。

我々もまた、(2) 実験計画段階における実験変数統制という観点から、実験計画スキルの学習を支援するシステムである仮想心理実験室 VPL (Virtual-Psychology-Laboratory) を提案した[4]。そして、実験デザインの組織性を記述するスキーマに基づいて図式化された実験行動履歴を閲覧しながら、VPL を用いて仮想的な心理実験室内で実験を繰り返すことにより、学習者の実験計画スキルが向上することが確認された[5]。しかし、VPL 学習で行われた実験活動を詳細に検討すると、計画された実験自体は実験変数が統制された適切なものであっても、仮説を検証できない実験を計画したり、結果の解釈を誤ったり、仮説の確認/反証の判断を誤るなど、仮説検証プロセス全体では一貫性を欠く行動を示す学習者がかなり多く存在した。

それでは、どのような支援により一貫した仮説検証プロセスの学習が可能になるだろうか。本論文では、仮説検証プロセスを検討することで新たな支援方法を提案する。

例えば、ある植物が生長するのに必要な条件に関して、仮説検証プロセスを通して調べる場合を考える。ここでは、“観測変数”(植物の背丈など;従属変数と呼ばれる場合もある)に影響を与えると考えられる実

[†] 京都大学大学院教育学研究科, 京都市

Graduate School of Education, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan

^{††} 名古屋大学大学院情報科学研究科, 名古屋市

Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya-shi, 464-8601 Japan

験変数を“要因”（与える水の量や肥料の量・日照量など；独立変数と呼ばれる場合もある）と呼び、その要因のとり得る値を“水準”と呼ぶ。

最初の (1) 仮説形成段階で「肥料の量が多いほど生長するだろう」という仮説が形成されたものとする。この仮説は、考えられる数多くの要因の中から肥料の量という要因に着目し、その着目要因と観測変数の間の関係について仮説を形成している。次の (2) 実験計画段階では、この仮説を検証するための実験を計画する。着目要因である肥料の量についていくつかの水準を設定し、複数の“実験条件”を作成することで一つの“実験”を計画する。例えば、肥料なし・肥料少量・肥料多量という三つの実験条件（群）を作成することで一つの実験が計画される。ここで注意すべきことは、着目要因のみを操作し、それ以外の要因に関しては統制しなければならないことである。なぜならば、複数の要因が無計画に変動する実験において特徴的な実験結果を得たとしても、その結果が着目要因の変化に起因するものか判断することができないためである。

計画された実験を実施した後、(3) 結果解釈段階では結果の解釈及び仮説の考察が行われる。各実験条件ごとに得られた実験結果を比較・分析し、操作要因により観測変数がどのように変化したのかを調べ、その結果を最初の仮説に対応づけて、仮説が確証/反証されているのか判断しなくてはならない。例えば肥料の量が多いほど植物の背丈が大きくなることが実験で確かめられたならば、ここでの仮説は確証されたことになる。

以上より、一貫した仮説検証プロセスを実現させるためには、それぞれの段階で中心となる着目要因・操作要因・観測変数に注意を向けさせ、それらを適切に関係づけさせることが必要であることが分かる。

本研究では、そのためにグラフの利用を提案する。なぜならば、グラフは着目要因と観測変数を目に見える形で表象させることができるためである。例えば、二つの着目要因（肥料の量と水の量）と観測変数（植物の背丈）の関係を折れ線グラフで表現した場合を考える。折れ線グラフでは、観測変数である植物の背丈の数値が縦軸に割り当てられる。そして、一方の着目要因である肥料の量の水準が横軸に並べられ、もう片方の水の量の水準が異なる種類の折れ線（直線や点線など）によって表現される。

本研究では、“予想グラフ”、“結果グラフ”、そして“模範グラフ”という三つの折れ線グラフと、それら三

つの折れ線グラフを用いた以下の四つの活動を行う。学習者にグラフを与えて閲覧させるだけでなく、形成した仮説や実験実施後に得られた実験結果を学習者にグラフの形で表現させたり、学習者が作成した複数のグラフを比較させることで、仮説検証プロセスにおける着目要因（操作要因）と観測変数に明示的に注意を向けさせることができると考えられる。

- 仮説形成後に“予想グラフ”作成
- 実験終了後に“結果グラフ”作成
- 作成した“結果グラフ”とシステムが示す“模範グラフ”を比較
- 考察時に“予想グラフ”と“結果グラフ”を比較

以上の活動を用いて実際に仮説検証プロセスを体験することで、一貫した仮説検証プロセスの学習が促進するのを実験的に検証する。

なお、現実場面では、ここでの仮説（一般には実験仮説と呼ばれる）を形成するにあたり、そもそも何を独立変数として扱えばよいのかが不明である場合がほとんどである。そのため研究者は、実験に先立って“要因の切出し”という課題に取り組みなくてはならない。しかし本システムでは、この課題を取り扱っていない。本システムの学習状況では、学習者に対して操作要因のリストが与えられ、学習者はその中から実験で操作する要因を選択することになる。そのため本システムの利用は、“候補となる要因が明らかになった後に、それらの要因から独立変数を選択し、計画的な実験を実施する”という状況に限定される。

また、本システムで取り上げた折れ線グラフは基本的に 2 要因までの実験計画しか扱うことができないため、本システムでは 2 要因以下の実験計画の学習支援を行うことを想定している。現実には 3 要因以上の複雑な関係は数多く存在していると考えられるが、実際の実験において意味ある情報を取り出せるのは多くの場合 2 要因実験までがほとんどであるといわれており、その意味ではこの制限は妥当と考えられる。

2. システム概要

本システムは Microsoft Windows 上に Allegro Common Lisp を用いて実装されており、仮説・実験計画・考察などを入力するメインウィンドウ（図 1）と、グラフ描画を支援するグラフ描画ウィンドウ（図 2）から構成される。学習者は、プランクトンがどのような条件下で最も繁殖するのかを、実験を通して解明することが求められる。メインウィンドウは以下のよう



図 1 メインウィンドウ

Fig. 1 A main-window of the system.

に分かれており、この番号の順に仮説検証プロセスを遂行する。

- (1a) 仮説入力欄 (図 1 左上段)
- (1b) 実験結果予想グラフ描画欄 (図 1 右上段)
- (2a) 実験計画欄 (図 1 左中段)
- (2b) 実験結果グラフ描画欄 (図 1 右中段)
- (2c) 模範グラフ表示欄 (図 1 右下段)
- (3) 結果考察入力欄 (図 1 左下段)

具体的には、まず、(1a) 仮説入力欄に形成した仮説を入力し、(1b) 実験結果予想グラフ描画欄をクリックすることでグラフ描画ウィンドウ (図 2) を呼び出す。グラフ描画ウィンドウでは、横軸の目盛の数 (水準数: 図 2 では “2”), 横軸ラベルの選択 (要因名: 図 2 では “栄養分”), 横軸目盛りラベルの選択 (水準名: 図 2 では “0” と “High”), 縦軸数値の最大値の選択 (図 2 では “1000”) を行った後、折れ線グラフの描画を行うことができる。このグラフ描画ウィンドウでは、複数本の折れ線を含む折れ線グラフを描画することが可能である。

なお、学習者には縦軸の数値 (定量値) に関して特別な指示を与えなかった。そのため、実際に学習者が作成したグラフを見ると、定量値として最大の値を選択する者がほとんどであった。本システムでは、“従属変数の値を決定する独立変数の関数” のような具体的で定量的な解を求めるのではなく、“従属変数と独立変数の間に正の相関がある” といった定性的な解を求めている。そのため、縦軸の数値は絶対値ではなく相対値が重視されることから、ここで用いた方法は妥当であると考えている。

予想グラフの描画終了後、そのグラフは (1b) 実験結果予想グラフ描画欄に描画される。

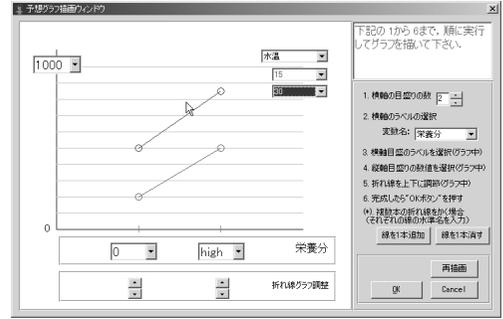


図 2 グラフ描画ウィンドウ

Fig. 2 A graph-window for drawing a graph.

(2a) 実験計画欄では、課題で用意された複数の要因を操作し、組み合わせることで実験を計画する。例えば、図 1 では、用意された五つの要因 (栄養量・水温・酸素含有量・水深・塩分濃度; 各要因 4 段階の水準を設定可能) の各水準を選択することで一つの実験条件を設定する。このプランクトン問題では、考えられるすべての要因の水準の組合せ、すなわち実験条件の種類は、 $1024 (= 4 \times 4 \times 4 \times 4 \times 4)$ 通り存在する。そのような実験条件を複数組み合わせることで、ここでの一つの“実験”を計画する。ただし、本システムでは、1 度の実験で組み合わせることができる実験条件の数を四つまでに制限している。現実場面でも、実験コストのために実験条件の数は最小限であることが望ましい。この制限により、限られた調査回数を有効に使った、より熟慮された実験計画が行われるようになると考えられる。

実験計画終了後、システムにより実験が自動的に実行され、その結果が (2a) 実験計画欄の中の実験結果表示部に追記される。実験結果、ここではプランクトン繁殖後の個体数は、システム内部にあらかじめ用意された関数を用いて求められる。その実験結果を参照しながら、(2b) 実験結果グラフ描画欄からグラフ描画ウィンドウを呼び出して結果グラフを描画する。描画終了後、そのグラフは (2b) 実験結果グラフ描画欄に縮小表示される。

結果グラフ描画後、(2c) 模範グラフ表示欄に、実験結果から描画可能なグラフの典型例がシステムから提示される。計画された実験の中で水準が変化している要因を操作要因とみなし、グラフの横軸にその操作要因を、縦軸に実験結果である繁殖後の個体数を対応づける。操作要因が二つある場合は、横軸と折れ線の種

類にそれぞれを対応づける．どちらを横軸（若しくは折れ線の種類）に対応づけるかの判断は、学習者が直前に描画した実験結果グラフに従う．ただし、実験計画が混乱してグラフ描画が不可能な場合、混乱した実験計画である可能性をシステムが学習に指摘する．

最後に、学習者は、(3) 結果考察入力欄において、(1b) 予想グラフと (2b) 結果グラフの比較を行いながら結果の解釈及び仮説の考察を行うよう求められる．

学習者は、このような一連の仮説検証プロセス（1回の実験に該当する）を何度も繰り返し行うことで、プランクトンの繁殖を規定する要因の解明を行う．

3. 評価実験

本研究では、仮説検証プロセスの学習支援として、グラフを利用した仮説検証プロセスの実施を可能とするシステムを提案し構築した．このようなシステムを用いて学習することにより仮説検証プロセスに関する理解が促進し、一貫した仮説検証プロセスの実施が可能になるのかを実験的に検討する．

3.1 実験手続き

[実験計画] 大学学部生 45 名が集団での実験と個別の実験に参加した．実験は約 3 週間にわたって実施された．実験は、学習フェーズの前後にプレ・ポストテストを行う実験デザインを採用した．学習効果の検討のために、本システム（以降 graph system と呼ぶ）と、graph system からグラフ描画機能を取り除いたシステム（以降 simple system と呼ぶ）を用意し、実験参加者 45 名を、学習時に simple system を用いる統制群 22 名と graph system を用いる実験群 23 名に分けた．

[プレテスト] プレテストは、実験参加者全員が simple system を用いて行った．仮説検証プロセスに関する知識、実験結果をグラフで表現する方法、そして simple system の使用方法を事前知識として教示した．課題は水中バクテリアの繁殖を解明し説明すること（どのような条件下で最も繁殖するか）で、設定された五つの要因（温度・酸素含有量・栄養量・照射量・pH；各要因は四つの水準をもつ）のうち、三つの要因（温度・栄養量・pH）に着目し、これらの要因が繁殖にどのような影響を与えるかを実験を通して解明するよう求めた．実験可能な回数、すなわち仮説検証プロセスの回数は 5 回までに制限し、実験終了後、解答用紙に説明を書くよう求めた．

[学習フェーズ] 学習フェーズは、日を改めて、個別

に行った．実際の学習の前に事前知識の復習を行い、実験群のみ、graph system におけるグラフ描画機能の利用方法を教示した．統制群の参加者は simple system を、実験群の参加者は graph system を用いて、水中プランクトンの繁殖を解明するよう求めた．この課題において設定された要因は五つ（栄養量・水温・酸素含有量・水深・塩分濃度；各要因は四つの水準をもつ）で、このうち三つの要因（水温・酸素含有量・水深）に着目して解明するよう求めた．実験回数は 6 回までに制限した．

[ポストテスト] ポストテストは、更に日を改めて、個別に行った．事前知識の復習を行った後、実験参加者全員が simple system を用いて、ウイルスの繁殖の解明を行った．具体的には、課題において設定された四つの要因（湿度・照射量・pH・温度；各要因は四つの水準をもつ）について実験を通して解明するよう求めた．ただし、プレテストや学習段階とは異なり、どの要因に着目するかの教示は行わなかった．学習フェーズと同様に、6 回の実験が終了した時点でシステムの使用を終了させた．

3.2 テスト課題の相違

両テストと学習フェーズの課題は、実質的に同じ内容の解が用意された．具体的には、ある一つの要因について繁殖数に対する主効果（例：pH の値は高い方が繁殖しやすい）があること、そしてその他の二つの要因について交互作用（例：湿度が高いときには温度が低い方がよく、湿度が低いときには温度が高いほうがよい）があることである．主効果に関する解と交互作用に関する解を比較すると、後者は二つの要因に着目しなければならないため、発見がより困難である．

課題の問題文に焦点を当てると、プレテストと学習フェーズについては、実験可能な回数（5 回と 6 回）や、要因の名前などの相違はあるが、用意された要因の数（ともに五つ）や、その中で特に着目するよう求めた要因の数（ともに三つ）などは同じで、基本的にほぼ同じ構造であるとみなすことができる．それに対して、ポストテストの問題では、用意された要因が四つで、着目する要因についての言及はなく、より難しい問題となっている．また、ポストテストの問題文中には交互作用に関する解の存在を示唆した文章が含まれている．

4. 分析

分析は、3 回の実験すべてに参加し、system の使用

や解答用紙への記入に不備のなかった 40 名（統制群 20 名，実験群 20 名）について行った。

なお，プレテストではほぼ全員が 60 分以内にシステムを使った実験活動を終了した。一方学習フェーズでは，グラフ群の平均が 68 分，統制群の平均が 43 分であった。ポストテストでは，グラフ群の平均が 39 分，統制群の平均が 44 分であった。

ここではまず，仮説検証プロセスの一貫性を定量的に示すために指標を提案する。

4.1 分析の指標

はじめに述べたように，仮説検証プロセスは，(1) 仮説形成段階，(2) 実験計画段階，(3) 結果考察段階に大きく分けることができる。これらの一貫性を示す指標として，次のような六つの指標を提案する。指標 A, B, C は，プロセスの各段階を単独で見たときの論理的な正しさを判断することで確認され，指標 D, E, F は，隣接する二つの段階の間で一貫性を判断することで確認される。これら指標はすべて，適/不適のいずれかの値をとるものとする。

指標 A：仮説としての妥当性 この指標は，仮説の言及内容の正しさではなく，言及が仮説として妥当かどうかを示すものである。ここでの妥当な仮説とは，一つ（あるいは複数）の要因と，パフォーマンスの間の関係についての言及である。例えば「温度が高ければ繁殖しやすいだろう」という仮説は実験によって確かめることが可能であるため，それが実験の結果正しいかどうかとは関係なく，仮説として妥当と判断する。それに対し，仮説を言及しなかったり「どの要因が繁殖に影響を与えるのか調べたい」などのあいまいな言及は仮説として妥当ではないと判断する。

指標 B：実験計画としての妥当性 仮説の確証/反証を確かめるためには，要因を統制して実験を計画しなくてはならない。実験を計画する際に用いられる重要な方略として，いくつかの実験変数統制方略 CVS (Control-Variables-Strategy) が知られている。例えば，一つの要因の主効果を調査可能な VOTAT (Varying-One-Thing-at-A-Time) 方略 [6] や，二つ以上の要因の交互作用を調査可能な FD (Factorial-Design) 方略などがある。これらの方略を用いた実験は，特徴的な観測変数の変化を要因の操作に関連づけることができ，確実な知見を得ることができるため，実験計画として妥当である。それに対して，要因を統制せず一度にほとんどすべての要因を操作して実験を計画する場合がある。このような実験では，たとえ特

徴的な結果を得ることができたとしても，その結果がどの要因を操作して得られた結果なのか明らかでないため，確実な知見を得ることができない。そのため，このような混乱した実験計画は実験計画として妥当ではない。以上より，実験計画が VOTAT 若しくは FD であるとき，本指標を適と判断する。

指標 C：考察としての妥当性 この指標は，考察時の言及が考察として望ましいかどうかを判断することで確認される。主に，考察入力欄の文章内で論理的な矛盾が生じていないかを判断の基準とする。判断は考察時の言及のみを対象とし，考察の根拠となる実験結果は考慮しない。

指標 D：仮説 実験計画間の一貫性 一貫した仮説検証プロセスでは，仮説を検証できる実験を計画することが重要である。言い換えると，仮説における着目要因と実験計画時の操作要因が一致しなくてはならない。そのため，着目要因と操作要因の一致がこの指標の判断基準となる。

指標 E：実験計画 考察間の一貫性 一貫した仮説検証プロセスでは，実験結果の解釈が実験で得られたデータに基づき適正に行われていることが重要である。そのため，結果考察欄に書かれた内容が，実験計画欄で行われた実験を解釈することで得ることが可能かどうかをこの指標の判断基準とする。

指標 F：仮説 考察間の一貫性 仮説検証プロセスは，文字どおり，仮説を検証するために行われる。そのため，仮説が確認されたかどうかを考察時に言及されなければならない。考察入力欄に，仮説の確証/反証に関する言及がない場合や，仮説における着目要因以外の要因に関して言及されていた場合には，この指標に関して不適とみなす。

4.2 指標による分析

以降では，プレ・ポストテスト成績の比較に基づき，仮説検証プロセスの一貫性の改善について分析を行う。

4.2.1 プレテストの分析結果

指標による分析：前述の六つの指標を用いて仮説検証プロセスを評価するため，各指標について，それぞれの実験参加者ごとに，行われたすべての仮説検証プロセス数（すなわち実験数）において“適”と判断された数の割合をもとめた。例えば，計 5 回の仮説検証プロセスが行われたプレテストで，仮説形成（指標 A）に関して，4 回が適，1 回が不適と判断されたとする。この場合，指標 A の値は， $0.8 (= 4/5)$ となる。それぞれの指標ごとに両実験条件群の平均値を求め，まと

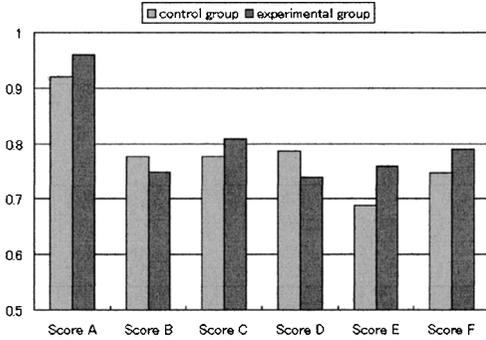


図 3 6 指標によるテスト成績 (プレテスト)
Fig.3 Averages of six evaluation scores (pre-test).

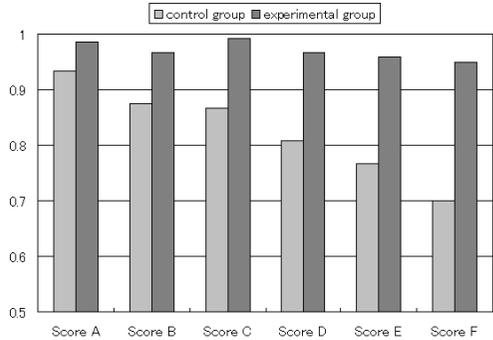


図 4 6 指標によるテスト成績 (ポストテスト)
Fig.4 Averages of six evaluation scores (post-test).

表 1 完全に一貫していた者 (プレテスト)

Table 1 Number of completely consistent participants throughout five processes (pre-test).

	一貫	一貫せず
統制群	6	14
実験群	4	16

表 2 完全に一貫していた者 (ポストテスト)

Table 2 Number of completely consistent participants throughout six processes (post-test).

	一貫	一貫せず
統制群	4	16
実験群	15	5

めたものが図 3 である。

いずれの指標においても両実験条件群間に統計的な差は存在しなかった (両側検定: 指標 A, $t(38) = 0.76$, $n.s.$; 指標 B, $t(38) = 0.30$, $n.s.$; 指標 C, $t(38) = 0.35$, $n.s.$; 指標 D, $t(38) = 0.48$, $n.s.$; 指標 E, $t(38) = 0.79$, $n.s.$; 指標 F, $t(38) = 0.42$, $n.s.$).

すべての仮説検証プロセスが一貫していた者の人数: 1 回の仮説検証プロセスで、六つの指標すべてについて“適”と判断されたとき、その仮説検証プロセスが完全に一貫していたとみなす。計 5 回の仮説検証プロセスのすべてが一貫している者の人数を表 1 に示す。2 × 2 の直接確率計算の結果、統計的な偏りはなかった (両側検定: $p = 0.72$, $n.s.$)。

4.2.2 ポストテストの分析結果

指標による分析: プレテストの分析と同様に、それぞれの指標ごとに両実験条件群の平均値を求め、まとめたものが図 4 である。

それぞれの指標について t 検定を実施したところ、指標 A は有意差がなく (両側検定: $t(38) = 1.46$, $n.s.$)、指標 B は有意傾向 (両側検定: $t(38) = 1.99$, $0.05 < p < .10$)、指標 C, D, E, F は有意差があった (両側検定: 指標 C, $t(38) = 2.66$, $p < .05$; 指標 D, $t(38) = 2.94$, $p < .01$; 指標 E, $t(38) = 2.98$, $p < .01$; 指標 F, $t(38) = 3.73$, $p < .01$)。

すべての仮説検証プロセスが一貫していた者の人数: 計 6 回の仮説検証プロセスすべてが一貫している

者の人数を、表 2 に示す。2 × 2 の直接確率計算の結果、偏りが存在した (両側検定: $p = 0.02$, $p < .05$)。すなわち、実験群の方が、すべての仮説検証プロセスが一貫した者が多かった。

4.2.3 プレ・ポスト分析のまとめ

六つの指標と、すべての仮説検証プロセスが一貫していた者の人数という二つの観点から分析を行った結果、プレテストについては、両方の観点で条件群間に差がなかった。つまり、両群の実験参加者は、一貫した仮説検証プロセスを実施する習熟度という点で等質であったと考えることができる。プレテストにおける両群の等質性が示されたため、システムの学習効果をポストテストの成績により検討することができる。

ポストテストの分析では、六つの指標に関して、指標 A を除くすべての指標で統制群よりも実験群の成績がよく、完全に一貫した仮説検証プロセスが実施できた人数に関して実験群の方が多かった。例外である指標 A に関しては、両群ともにプレテストの段階で、既に成績が 9 割を超えていたことから、ポストテストにおいて成績の差が現れなかったと考えられる。

以上の分析より、一貫した仮説検証プロセスの学習において、グラフを利用した本システムの有効性が確認された。

4.3 学習段階の分析

4.3.1 学習段階の分析結果

学習段階では、どのような仮説検証プロセスがたど

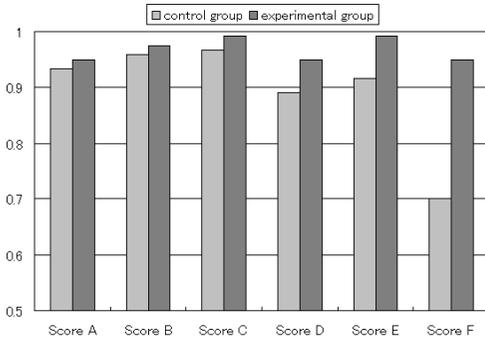


図 5 6 指標による成績 (学習段階)
Fig. 5 Averages of six evaluation scores (learning).

表 3 完全に一貫していた者 (学習段階)

Table 3 Number of completely consistent participants throughout six processes (learning).

	一貫	一貫せず
統制群	8	12
実験群	11	9

られ、どのような学習が行われたのだろうか。プレ・ポストテストと同様の観点から検討する。

指標による分析：前述の六つの指標ごとに、両実験条件群の平均値を求め、まとめたものが図 5 である。6 指標の平均値について、それぞれ t 検定を実施したところ、指標 A, B, C, D に関して有意差はなく (両側検定：指標 A, $t(38) = 0.23, n.s.$; 指標 B, $t(38) = 0.52, n.s.$; 指標 C, $t(38) = 0.92, n.s.$; 指標 D, $t(38) = 1.00, n.s.$)、指標 E, F についてのみ有意差があった (両側検定：指標 E, $t(38) = 2.20, p < .05$; 指標 F, $t(38) = 3.86, p < .01$)。

すべての仮説検証プロセスが一貫していた者の人数：計 6 回の仮説検証プロセスすべてが一貫している者の人数は、表 3 のようになった。2 × 2 の直接確率計算の結果、統計的な偏りはなかった (両側検定： $p = 0.53, n.s.$)。

4.3.2 学習段階の分析のまとめ

学習段階では、実験参加者は統制群と実験群に分かれ、仮説検証プロセスを実施した。分析の結果、グラフを用いることで実験結果の解釈 (指標 E) の支援と考察における仮説の正誤判断 (指標 F) の学習支援が行われたことが示された。前者は、結果グラフを作成し、作成したグラフを結果解釈時に一覧することが結果解釈を支援したためと考えられ、後者は、予想グラフと結果グラフの比較をすることが、仮説の正誤判断に役立つためと考えられる。

指標 A で差がなかったのは、プレ・ポストテストの分析において述べたように、仮説形成について両群ともに最初から問題なかったためである。指標 B, C, D で差がなかったのは、グラフを利用した実験群だけでなく、利用しなかった統制群も等しく成績が向上したためである。具体的には、プレテスト—学習段階 × 統制群—実験群の 2 × 2 の ANOVA を実施した結果、交互作用がなく (指標 B, $F(1, 38) = 0.26, n.s.$; 指標 C, $F(1, 38) = 0.00, n.s.$; 指標 D, $F(1, 38) = 1.16, n.s.$)、プレテスト—学習段階条件の主効果のみ有意であった (指標 B, $F(1, 38) = 22.35, p < .01$; 指標 C, $F(1, 38) = 19.08, p < .01$; 指標 D, $F(1, 38) = 10.23, p < .01$)。グラフ利用の有無と関係なく成績が向上していることから、グラフ利用以外の学習支援が有効であったと考えられる。可能性の一つとして、同じ課題に繰り返し取り組むことによる学習効果が挙げられる。プレテストの課題と学習段階の課題は、ほぼ同じ内容であり、その繰返しによる学習効果は大きいと予想される。この点は、指標 B (実験計画段階) の成績向上に関しては、特別な支援が行われない統制群であっても仮想的な実験室の中で同じ課題について実験活動 (仮説検証プロセス) を繰り返すことで、VOTAT 方略や FD 方略などを用いた実験を計画する割合が増加するという先行研究の知見とも整合的である [5]。また、仮説を検証できる実験を計画できているかどうかを判断基準とする指標 D は、完全な実験計画ができていないことが必要条件であるため、指標 B に大きく依存している。そのため、指標 B の成績向上に伴って指標 D も向上したと考えられる。

学習段階で指標 B, C, D の成績に差がなかったにもかかわらず、ポストテストに差が現れた理由として、課題の難易度が増したことが挙げられる。ポストテストの課題はプレテスト・学習段階の課題と本質的に同じだが、着目すべき要因の指定がないため、実験参加者自らが着目要因を決定しなければならない。

また、ポストテストの課題には、「どうやら二つの要因が相互に関連しつつ繁殖に影響を与えているのですが詳しいことは分かっていません」という交互作用の解の存在を示唆する文章が含まれている。そのため実験参加者は一度の実験で複数の要因に着目し、交互作用の解を発見できるより高度な実験計画 (前述の FD 方略等) に取り組もうとしたと考えられる。それを裏づけるデータとして、学習段階とポストテストについて、二つ以上の要因に着目した仮説の割合を求

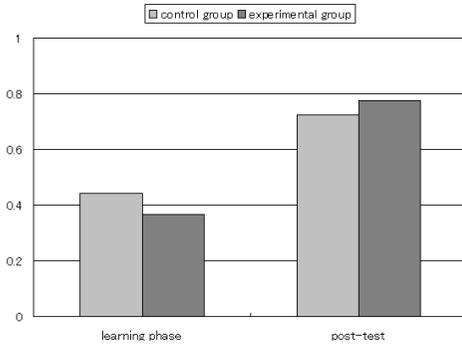


図 6 二つ以上の要因に着目した仮説の割合
Fig.6 Ratio of hypothesis formation of multiple focused factors.

めたのが図 6 である。2 × 2 の分散分析を行ったところ、学習段階—ポストテスト条件のみ主効果があった ($F(1, 38) = 32.54, p < .01$)。図 6 から分かるように、学習段階では 4 割程度であるのに対し、ポストテストでは 8 割近い実験で二つ以上の要因に着目している。

以上から、グラフを利用した本システムで学習することで、より複雑な課題において一貫した仮説検証プロセスを実施できるようになった可能性が示唆される。

5. 総合考察

5.1 解を発見できた人数

先に述べたように、プレ・ポストテストの両方の課題で、(a) 主効果、(b) 交互作用の二つの解が設定されている。それぞれの解を発見できた人数を実験条件群ごとにまとめたものが表 4 である。すべての段階において、解 (a)、(b) のそれぞれについて 1 × 2 の直接確率計算を行った結果、統計的な偏りはなかった (両側検定: プレテスト (a), $p = 0.87, n.s.$; プレテスト (b), $p = 0.38, n.s.$; 学習段階 (a), $p = 0.87, n.s.$; 学習段階 (b), $p = 1.00, n.s.$; ポストテスト (a), $p = 1.00, n.s.$; ポストテスト (b), $p = 0.66, n.s.$)。

科学的発見における学習目標は、science contents と how to do science に関する 2 種類があるといわれている。言い換えるならば、対象領域に存在する科学的概念 (ここではバクテリアや植物の繁殖) と発見を行う科学的手法 (ここでは仮説検証プロセス) の 2 種類の学習目標である。本学習システムでは前者ではなく後者を学習の目的としており、前者に関連するバクテリア等の繁殖の挙動は、システム開発時に便宜的に

表 4 解を発見できた人数

Table 4 Number of participants who found the solutions.

	プレテスト		学習段階		ポストテスト	
	解 (a)	解 (b)	解 (a)	解 (b)	解 (a)	解 (b)
統制群	18	4	16	10	16	9
実験群	16	1	18	9	16	12

仮定されたものにすぎない。

前者の学習における評価項目である解の発見割合に関してはここで示したように差がなかったが、後者の学習における評価指標に関しては既に示したように有意な向上が見られた。したがって、科学的探求プロセスの学習という本システムがターゲットとする学習支援に関しては、本システムの有効性が支持されたものと考えられる。

5.2 グラフによるデータの理解

グラフ理解に関するモデルとして、Pinker はグラフ情報を表現する複数の知識表象を提案している [7]。例えば、グラフ図から読み取ったオブジェクト間の関係などを構造的に記述した visual description や、あらかじめ知識としてもっているグラフ記述規則に関する graph schema、そして、グラフから読み取った情報 (与える肥料の量が X なら繁殖後の個体数は Y であるなど) を記述した conceptual message などがこれにあたる。グラフの基本的な理解、例えば、そのグラフが折れ線グラフであり、肥料の量と植物の背丈に関するデータを示していることへの理解は、visual description と graph schema の対応付けを通して行われる。対応付けの基準となるのは、軸を中心としてデータの属性を表現するという graph schema 中のグラフの基本的な性質と、visual description 中のデータの属性に関する記述 (縦・横軸に付記されたラベルや軸に沿って並べられた複数の水準名等) である。これらの対応付けによって、グラフの物理的次元 (縦軸・横軸・折れ線の種類等) に表現されたデータの属性 (肥料の量や植物の背丈等) が読み取られ、グラフの基本的な理解が行われる。

Pinker のモデルによれば、学習者はグラフを操作することにより、データの属性、すなわち観測変数と着目要因に注意が向けるようになる。これにより学習者は、一貫した仮説検証プロセスを遂行することができるようになったと考えられる。

ここで注意したいのは、グラフ作成能力の向上は、本システムの直接的な目的ではないという点である。

グラフ作成は、観測変数や着目要因に注意を向けさせ、仮説検証プロセスにおいて一貫させることの重要性に気づかせるための手段である。例えば、グラフ機能のない simple system を用いるポストテストで実験参加者全員にメモ用紙を渡したが、実験条件群の実験参加者であっても、仮説検証プロセスの実施中にグラフをメモ用紙に描画しなかった。このことは、グラフを作成するようになったために仮説検証プロセスが一貫するようになったのではなく、グラフ作成活動を通して着目要因や観測変数の一貫性を理解することで仮説検証プロセスが一貫するようになったことを示唆している。

5.3 振返り行動による学習

学習段階で行われた仮説検証プロセスを詳細に検討すると、実験群では、グラフ利用を通してプロセスの様々な失敗を即座に訂正していた。例えば、仮説における着目要因と実験における操作要因の不一致という失敗をしたある学習者は、その実験の考察記入欄に、予想グラフと結果グラフの比較を通してその誤りに気づいた旨を記述していた。つまり、グラフを利用することで、自らの行動の修正という振返り行動が促進されていることが分かる。

自己の認知状態や理解の状態を振返り・修正する活動はリフレクションと呼ばれ、学習時にそのような活動を頻繁に行うことで理解がより深まると考えられる [8]。しかし学習者にとって、学習時に行った自分のプロセスは明示的でなく振り返ることが困難であるため、何らかの支援が必要である [9]。その支援手段の一つが、プロセスの明示的な提示である。例えば Schauble らは、仮想的な実験室における実験プロセスを記述する表記法として、Discovery and Reflection Notation (DARN) を提案している [10]。また、Anderson らは、幾何学の学習を支援する Geometry Tutor システムにおいて、問題の初期状態からゴールに至るまでの解の導出過程を proof goal として提示している [11]。

本システムにおける複数のグラフは、学習者が行った仮説検証プロセスの各段階を明示的に提示することによって学習者のリフレクションを支援している。一般にグラフは、データを整理し、その傾向を読み取るのに有効であると理解されている。そのためグラフは、これまで提案されてきた様々な学習支援システムにおいて、データを学習者にフィードバックする方法の一つとして用いられてきた。例えば、BGuILE (Biology Guided Inquiry Learning Environments) プロジェ

クトにおいて利用される The Galapagos Finches というソフトウェアは、解明対象である鳥 (finch) の様々なデータをグラフの形で提示する [3]。

本システムがこれらシステムと異なる点は、データの representation としてグラフを利用したのではなく、Pinker のモデルで述べられた、グラフ利用時に行われる「グラフの軸・折れ線の種類」と「実験要因」との対応付けという認知的活動に注目し、仮説検証プロセスの要因の一貫性に関するリフレクションを促進する認知的道具としてグラフを利用した点である。

6. む す び

本研究では、仮説検証プロセスにおける、仮説形成・実験計画・結果考察の各段階のつながり、すなわち仮説検証プロセスの一貫性に着目した。一貫した仮説検証プロセスの学習を支援するために、グラフを利用する学習支援システムを提案し、実験的に評価を行った。システムがもつ仮想的な実験室の中で、学習者が、グラフ作成による結果の予測・グラフ作成による結果の解釈・システムが返す模範グラフによる作成したグラフの正誤判断・予測グラフと結果グラフの比較という4種類のグラフ利用を行いながら仮説検証プロセスを練習することで、一貫した仮説検証プロセスを実施できるようになることが分かった。

本システムでは、グラフの作成・閲覧・比較という様々なグラフ利用が行われている。本研究の評価実験によりこれら複数のグラフ利用活動による学習効果を示されたが、その学習効果がどのグラフ利用活動によるものなのか、あるいはいずれか複数の利用によるものなのかは特定されていない。この特定は重要な問題であり、今後の研究課題である。

文 献

- [1] D. Klahr, *Exploring Science: The Cognition and Development of Discovery Processes*, MIT Press, London, 2000.
- [2] D. Schwartz, S. Brophy, X. Lin, and J. Bransford, "Flexibly adaptive instructional design, A case study from an educational psychology course," *Educational Technology Research and Development*, vol.47, no.2, pp.39-59, 1999.
- [3] I. Tabak and B. Reiser, "Complementary roles of software-based scaffolding and teacher-student interactions in inquiry learning," *Proc. Computer Supported Collaborative Learning '97*, ed. N.M.R. Hall and N. Enyedy, pp.289-308, Toronto, Ontario, Canada, 1997.
- [4] 中池竜一, 三輪和久, "探索的実験行動を記述する仮想心理

- 実験室の構築,” 人工知能誌, vol.17, no.4, pp.490-499, 2002.
- [5] K. Miwa and R. Nakaïke, “Learning meta experimental skills in a simulated psychology laboratory,” ICCE, pp.1182-1190 2003. in press.
- [6] J. Tschirgi, “Sensible reasoning: A hypothesis about hypothesis,” Child development, vol.51, pp.1-10, 1980.
- [7] S. Pinker, “A theory of graph comprehension,” in Artificial Intelligence, and the Future of Testing, pp.73-126, Lawrence Erlbaum, Hillsdale, NJ, 1990.
- [8] 三輪和久, 齊藤ひとみ, “学習科学に基づく学習/教育支援システムの設計と実現—リフレクションに基づく学習支援を題材として,” 教育システム情報学会誌, vol.21, pp.145-156, 2004.
- [9] X. Lin, C. Hmelo, C. Kinzer, and T. Secules, “Designing technology to support reflection,” Educational Technology Research and Development, vol.47, no.3, pp.43-62, 1999.
- [10] L. Schauble, K. Raghavan, and R. Glaser, “The discovery and reflection notation: A graphical trace for supporting self-regulation in computer-based laboratories,” in Computers as cognitive tools, pp.319-337, Mahwah, Erlbaum, NJ, 1993.
- [11] J.R. Anderson, C.F. Boyle, and B.J. Brian, “Intelligent tutoring systems,” Science, vol.228, pp.465-462, 1985.

(平成 16 年 5 月 21 日受付, 9 月 15 日再受付)



中池 竜一

1998 名工大・工・知能情報システム卒 .
2000 名古屋大学大学院人間情報学研究科
前期課程了 . 2004 より京都大学教育学
研究科助手 . 認知科学, 教育工学の研究に
取り組む . 探索的な問題解決時に行われる
人間の高次思考過程の解明とその知見の教育
的な応用に興味をもっている . 日本認知科学会,
人工知能学会,
教育システム情報学会各会員 .



三輪 和久 (正員)

1984 名大・工卒 . 1989 同大学院工学
研究科博士課程了 (情報工学専攻) . 工博 .
1989 同大学情報処理教育センター助手,
1993 同大学院人間情報学研究科助教
授を経て, 2004 より名古屋大学大学院情報
科学研究科メディア科学専攻教授 . 1991 ~
1992, 米国 Carnegie Mellon University, Dept. of Psychol-
ogy, visiting assistant professor . 認知科学,
人工知能, 教育
工学の研究に従事 . とりわけ, 発見, 創造,
洞察, 協同など,
人間の高次思考過程に興味がある .