

協調問題解決における仮説形成検証過程の分析

鈴木 裕利[†] 三輪 和久[†] 櫻井 桂一^{††}

An Analysis of Formation and Verification Processes of Hypotheses
in Cooperative Problem Solving

Yuri SUZUKI[†], Kazuhisa MIWA[†], and Keiichi SAKURAI^{††}

あらまし 本論文では協調問題解決の場面において、どのようなインタラクションが発生し、どのように仮説形成検証過程が進行していくのかを分析する。そこで、コンピュータ上に実現した概念同定課題を用いた実験を行い、2人の被験者が協力して課題を遂行するプロセスを詳細に検討した。実験結果の分析については、「インタラクティブな仮説形成検証モデル」を導入し、セマンティックプロトコル法を用いた。その結果は、以下のようによまとめられる。(1)成績の良いグループほど、2人の交互の発話が多く、活発なインタラクションが観察された。(2)特定の仮説に直接言及する発話の交換よりも、データの吟味や仮説の検証についての発話の交換が多いことが観察され、更に、成績との相関において、仮説の評価、検討にかかわる発話の重要性を指摘した。(3)一つの仮説の形成検証プロセスを長く続けることよりも、多種類の仮説を形成検証していくことが重要であると指摘した。

キーワード 協調問題解決, 仮説形成検証, インタラクション, 概念同定, プロトコル分析

1. ま え が き

近年、「認知科学」の分野において、「インタラクション (相互作用)」に関する研究が注目されている [1]。その重要性は、既に 1979 年に Norman が述べており [2]、彼の予言どおり、現在の「認知科学」にとって、避けて通れない課題となっている。発見や問題解決の領域においても、グループによる問題解決は、個人のそれにはないいくつかの要因が存在しており [3]、その研究の進展が待たれている。

発見や問題解決のプロセスに決まって現れる過程が、仮説形成検証過程である。仮説形成検証過程に関する研究は、科学的発見や推論の領域で、継続的な研究がなされてきておりその報告も多い。例えば、Klahr らのグループでは、仮説空間と実験空間を探索する過程として、仮説形成と仮説検証の過程を統一的な観点からとらえようとしている [4]~[6]。

しかし、旧来の研究は、個人の過程に着目されたもの

のがほとんどであった。近年、複数の人間による協調問題解決にまで、その関心を拡張しようとする試みがなされてきている。しかし、その多くは、個人と複数の人間の問題解決のパフォーマンスに焦点が当てられたものや、個別的なテーマのもとで、特定の関心のもとに行われたものがほとんどである。ここでは、例えば、仮説検証における「確認バイアス」の問題 [7]、[8]、発達心理学の観点から小学生を対象に行われた研究 [9]、若しくは社会心理学の立場から多数派、少数派の相互作用に焦点が当てられたものなど [10]~[12] がある。このように従来の研究では、仮説形成検証という、より一般的な文脈の中で、かつ 2 人の人間のインタラクションの違いに焦点を当てたものはまれである。

本研究では、簡単な概念同定課題を用いて、2 人の人間が協調的に課題を解決していくプロセスを詳細に検討する。具体的には、その問題解決のプロセスに、どのようなインタラクションが生じ、それが問題解決のパフォーマンスにどのような影響を与えているのかを明らかにする。

複数の学習者が共同して問題を解決する過程には、二つの仮説形成検証過程が、互いに補完し合いながら進むものと考えられる。本論文では、そのような「相

[†] 名古屋大学大学院人間情報学研究所, 名古屋市
Graduate School of Human Informatics, Nagoya University,
Nagoya-shi, 464-8601 Japan

^{††} 愛知県立大学情報科学部, 愛知県
Faculty of Information Science and Technology, Aichi Prefec-
tural University, Aichi-ken, 480-1198 Japan

補的構造を有した仮説形成検証過程」というフレームワークから、2人の学習者の仮説形成検証のプロセスを詳細にトレースし、その結果を分析することによって、上述の課題にこたえようとする。

以下2.では、実験方法と実験に使用した概念同定実験システムについて説明する。次に3.では、同システムを用いて行った実験の結果を示す。4.では、「インタラクティブな仮説形成検証モデル」を導入し、セマンティックプロトコル法を用いて結果を分析し、実験結果を考察する。最後に5.では、本論文をまとめ、今後の課題を述べる。

2. 実験

ここでは、本研究のために開発した概念同定実験システムと、実験の詳細について説明する。概念同定実験は、仮説形成検証過程について議論するために、認知心理学において古くから用いられている方法である[13]。

2.1 概念同定実験システム

概念同定を課題とする実験システムをパーソナルコンピュータ上に実現した。図1に、画面表示の一例を示す。

このシステムでは、被験者の入力情報に対し、実験システムが、ある「ルール」に従って情報を出力する。被験者には、自分の入力した情報と実験システムからの出力情報をもとに、この「ルール」を発見する

ことが課せられる。入力情報をメッセージ1、メッセージ2と呼び、実験システムからの出力情報を結果メッセージと呼ぶ。被験者は、このメッセージ1とメッセージ2を入力し実験システムから結果メッセージを得る(これを「試行」と呼ぶ)という一連の操作を行う。被験者は1問につき最大20試行まで行うことができる。

メッセージは、色、形、数という三つの次元の組合せからなる。色の次元は黄、赤、緑、青、形の次元は●、■、▲、★、数の次元は1、2、3、4というように、それぞれ4種類の値をもつ。よって、メッセージは合計64種類となる。例えば、図1の最上段の各メッセージは、左から順に、(黄、●、1)、(緑、■、2)、(黄、★、3)である。実際の実験システムでは、選択可能な64種類のメッセージが、図1に示した画面の右側に一覧で表示されている。被験者はマウスでクリックすることによって、メッセージ1とメッセージ2を入力することができる。

次に、本実験システムにおいて設定されたルールについて述べる。例えば、「結果メッセージの色がメッセージ1の色と同じか、又は、結果メッセージの数がメッセージ2の数と同じになる」というルールは、「(結果メッセージの色=メッセージ1の色) OR (結果メッセージの数=メッセージ2の数)」と記述できる。このルールは“OR”、すなわち選言概念として特徴づけられる。“OR”のほかには本システムが用いる概念は、例えば「結果メッセージの色がメッセージ2の色と同じになる」という単純概念、更に、否定概念(NOT)、及び、連言概念(AND)である。

2.2 実験方法

本実験に参加した被験者は、2人を1グループとする、10グループ20名である。以下に実験の実施手順を説明する。

各グループの2人の被験者に与えられる問題は、選言概念の問題である。被験者には「相談をして解答をすること」、「できる限り少ない試行回数で正解を見つけること」、「時間に関しては制限はないこと」が教示される。また、実験の終了は、被験者が実験者に対して、被験者のもつ仮説を表明し、正誤の確認を求め、その仮説に対して、実験者が正解であると認めた時点とする。したがって、被験者の発話の中に正解仮説が出現しても、実験者に確認を求めない限り実験は継続される。なお、これ以降で論じられる、パフォーマンスの指標としての「試行回数」は、実験の終了時点ま

問	メッセージ1	メッセージ2	結果メッセージ
1	●	■	★
2	■	●	●
3	■	★	★
4	●	●	▲

NOTE:
Each box in the actual display is colored such
as    

図1 実験システムの画面例

Fig. 1 An example display of the experimental system.

での試行回数を意味している。

各被験者に対して、個別的に、事前に3問の問題を与えている。過去に実施した単独被験者による予備的な実験 [14] においては、この事前の問題の難易度の差により、その後に続く本課題のパフォーマンスに偏りが見られた。つまり、難易度が低い問題（単純概念や連言概念）に比べて、事前に難易度が高い問題（否定概念）を与えた場合、被験者は本課題（選言概念）で高いパフォーマンスを示す傾向がある。そこで、ここでの実験では、本課題のパフォーマンスを適度に分散させるために、被験者を二つの群に分け、一方には事前に難易度の低い問題を与え、他方には難易度の高い問題を中心に与えている。

被験者の問題解決過程は、次のように記録される。まず、メッセージ1、メッセージ2と結果メッセージの履歴をフロッピーディスク上に自動的に記録する。また、ビデオカメラとテープレコーダにより、実験システム画面と被験者の会話を記録する。

3. 実験結果

ここではパフォーマンスに関する実験結果を示し、後に詳細に分析するプロトコルデータについて説明する。

3.1 試行回数及び解答時間

図2は、実験に参加した10グループのうち、9グループについて、試行回数と実験開始から終了までの解答時間の関係を示している。なお、残りの1グループ

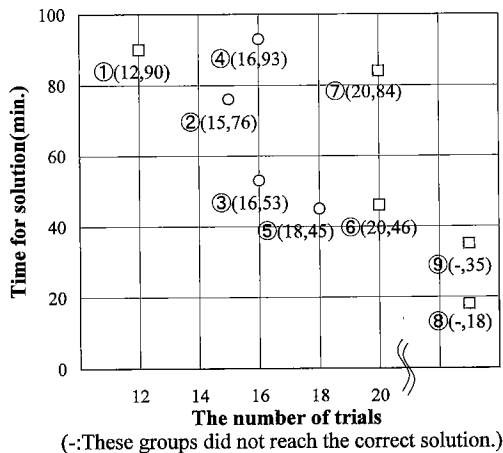


図2 実験結果（試行回数及び解答時間）
Fig.2 Experimental results (the number of trials and the time for solution).

プについては、非常に特殊な方略を用いて問題に取り組んだため、実験データから除外した。

図中の①から⑨の番号は、試行回数の少ないグループから順に番号付けをしたものであり、()内は左側が試行回数、右側が解答時間(分)である。また、グループ⑧⑨は、正解に至らなかったグループである。更に、事前に難易度の低い問題に取り組んだグループ(①⑥⑦⑧⑨)は「□」、難易度の高い問題に取り組んだグループ(②③④⑤)は「○」で示している。

図2からは次のことがわかる。事前問題の難易度の高いグループは試行回数が少なく、難易度の低いグループは①を除いて、最大回数(20回)を要するか、または、不正解となっている。これは、単独被験者を対象とした実験の傾向と一致している。解答時間については、大まかに見て、75分以上を要しているグループ(①②④⑦)と50分前後で終了しているグループ(③⑤⑥)に分かれる。不正解のグループ(⑧⑨)の所要時間も少ない。

3.2 プロトコルデータ

全グループについて、ビデオテープとカセットテープの記録から、2人の発話のプロトコルデータを作成した。具体例を表1に示す。以下では、グループ内において発話数の多い被験者を被験者A、少ない被験者を被験者Bと表す。各発話は、意味的なまとまりをもつ発話を1個の発話とみなして区切り、記述の単位とする。このように被験者の発話の意味的なまとまりに着目する方法は、セマンティックプロトコル法と呼ばれている [15], [16]。なお、表1の右側の欄については、次章で詳しく説明する。

なお、グループ当りの発話数の平均は680発話(最大1,276発話, 最低221発話)であった。

4. インタラクティブな仮説形成検証モデル

被験者の仮説形成検証プロセスを記述するために、事前に設定した分類カテゴリーに基づいて、プロトコルデータにラベル付けを行う方法が広く用いられている [17]~[19]。本研究においても同様の方法を用い、被験者内及び被験者間の仮説形成検証プロセスを分析する。以下では、「仮説」に対して、観察されるメッセージ1、メッセージ2、結果メッセージの組を「事例」と呼ぶ。

4.1 仮説形成検証プロセスに関するカテゴリー

一般に、仮説形成検証プロセスは「仮説形成」「仮

表 1 プロトコルデータ例
Table 1 Examples of the semantic protocols.

NO.	被験者 A	被験者 B	被験者 A				被験者 B				
			Oa	Pa	Ga	Ta	Tb	Gb	Pb	Ob	
105		形は関係ない									✓
106	ない、うん、形は関係ない	色はかん	✓								✓
107											
108	関係ないですよ、え、そうすると		✓								
109		数、数も									✓
110	関係ない		✓								
111		数も関係ない、ない									✓
112	はははは、全部ランダムじゃないみたいですねえ		✓								
113	はーん										
114	赤でしょお										
115	難しいですね、これ		✓								
116		同じ形がでてる									✓
117	うん			✓							
118		同じ色でてる									✓
119	ふふふ、同じかたちもでるなあ			✓							
120		同じ形はでる									✓
121		M1に関しては同じM2がでる									✓
122	ふん										
123		のときえっと、同じ数がそれぞれでていない									✓
124	うんうん、そうするとですねえ				✓						
125		と、またしても、NOT系か、答えは									✓
1256	ふたつふたつでしょ			✓							
1257	●で●がでる			✓							
1258		///									
1259	これ、2のメッセージが出てますよね、え、蓮うか			✓							
1260		そう、									✓
1261	2の			✓							
1262		あなたのいうように2の形がこれ出てるのね、そう									✓
1263	2の要素が出る、2の要素のなにが出てますか？			✓							
1264		なにが									✓
1265	2の要素の何が出てるかっていうとお、形ですよえ			✓							
1266		そう2の形がでてますよ									✓
1267	ええ、でどういうときに2の形がでるんですか			✓							
1268		それがわかってないんです									✓
1269	だから、M1の同じ色ともしくは、っていうことかしら					✓					
1270		だって、論理でしょ、									✓
1271	論理ですよ										
1272		まさかとは思うけど、ランダムで2番の形が優先される						✓			✓
1273	優先される、そういうことじゃないですか							✓			
1274		たまたま2番目の形を優先するんですか								✓	
1275	だからORなんですよ、とにかく、								✓		
1276	M1の色と、M2の形でORになってるんですよ、出方が								✓		

説評価」から構成される。更に前者は、「規則性の観察」と「仮の一般化の提案」、後者は「実験と観察による仮説の予測のテスト」に分類される [10]。

本研究では、このフレームワークを基本として、次のような分類を用いることとする。まず前述の「仮の一般化の提案」を、「仮説生成」期ととらえる。次に、「規則性の観察」を、明確な仮説を生成する前の「仮説準備」期とする。この仮説準備期には規則性をとらえるための事例の吟味と、完全な仮説には至らないが、ある規則性に言及する段階も含むものとする。更に、本課題においては、「仮説準備」期以前に、仮説や事例に対する非常にあいまいな言及も確認されるので、この時期を「事例観察」期ととらえて、仮説形成検証プロセスに加える。また仮説形成後に行われる「仮説評価」の過程を「仮説検証」期とする。このよ

うにして、ここでは仮説形成検証プロセスを「事例観察」期、「仮説準備」期、「仮説生成」期、「仮説検証」期の各サブプロセスに分類する。以下ではこの四つプロセスを O, P, G, T と表し、各々の定義を以下に示す。

(1) O (事例観察プロセス; observing process)

仮説、事例への不明確な言及を示す発話である。ある仮説の構成要素に言及するものの、非常にあいまいで明確な仮説に関する言及とは言いがたい発話を示す。例えば、「色が怪しい」という発話のように、「色」という要素に着目して事例を観察しているものの、仮説としては明確な意味をもたない内容の発話がこれにあたる。

また、事例に言及しているが、内容が明確ではない発話もここに分類される。例えば、結果メッセージが

出力された直後の「変わりましたね」という発話のように、何が変わったか明確でなく、何に対して言及しているのが不明確な場合である。

(2) P (仮説準備プロセス; preparatory process)

仮説形成の前段階として、着目する仮説や事例に、具体的に言及する発話を示す。例えば、仮説を意味する発話であるが、完全な仮説になっていない発話がこれに該当する。具体的には、「ANDかな?」というように、連言概念に関する仮説を意図しているが、どういふ要素の「AND」であるかが不明な場合である。また、「1, 2, 3, 1, 2」といふ発話のように、「数字」の要素に関して何らかの仮説を生成しようとして事例を吟味している場合もここに該当する。

(3) G (仮説生成プロセス; generating process)

明確な仮説を表明する発話や、以前に表明されている仮説を復唱したり、言い換えたりする発話を示す。例えば、「メッセージ1の色がいつも出る」といふ発話のように、「メッセージ1の色=結果メッセージの色」といふ単純概念のルールに関する仮説を明確に表現している場合である。

(4) T (仮説検証プロセス; testing process)

生成した仮説の検証や評価をする発話を示す。例えば、ある仮説を生成した後、新たに試行を行った結果を観察して、「全然違う」「おかしい」と発話する場合、あるいは、「この4はルールとあわないから困る」のように、具体的事例に言及して仮説を検証しているような場合である。つまり、事例に基づいて事前に表明している仮説を吟味、検証する場合がこれにあたる。

4.2 カテゴリーに基づいたコーディング例

4.1のカテゴリーに基づき、本実験の全プロトコルデータについてコーディングを行った。コーディングには、書き起こされた発話データに加え、ビデオカメラを使って記録された、あいづちや指さし等の非発話情報を補助データとして用いている。

ここでは、被験者A, Bの発話をそれぞれa, bの添字を用いて表示する。よって、各プロセスに該当する全発話は、O_a, P_a, G_a, T_a, O_b, P_b, G_b, T_bのいずれかにラベル付けされる。ラベル付けの例が、表1の右欄に示されている。

以下、表1に従って、具体的にコーディングの例を示す。

NO.105の被験者Bの発話「形は関係ない」は、

「形」といふ要素に着目はしているが仮説にはなっておらず、具体的事例にも言及していないので、O_bとする。同様な発話が、NO.106から112、及び115であり、O_aまたはO_bとコーディングされる。NO.113は独り言とみなしラベルは付けない。NO.114の「赤でしょ」は具体的事例に言及しているのでP_aとする。また、NO.116は被験者Bが「同じ形が出る」と具体的な事例の情報に言及し、被験者Aが「うん」と明らかに同意しているので、両者の発話はそれぞれ、P_b, P_aとする。NO.118から124も同様に、P_aまたはP_bとする。NO.125は、否定概念の仮説に言及しているが、その内容が明確でないのでP_bとする。

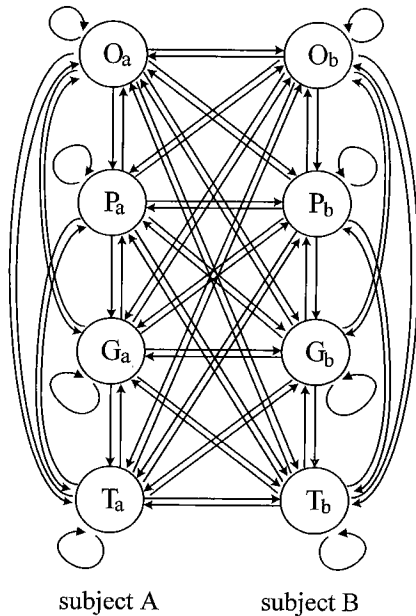
後半についても同様に、NO.1256から1268はP_aまたはP_bに分類する。ただし、NO.1258の「/ / /」は聞き取り不可能な発話であり、独り言とみなす。NO.1269は1267の発話を受けて、「M1の色、若しくは、M2の形」といふ具体的な仮説を生成している発話なのでG_aとする。NO.1270以降はNO.1269での仮説をもとに検証を行っているのでT_a, T_bに分類する。最後にNO.1276は先の仮説を復唱しているのでG_aとする。

以上のコーディングを行った結果、仮説形成検証に関する発話(O, P, G, Tのラベルの付いた発話)の全発話に対する割合は、平均して61.5% (最大75.2%, 最小45.2%)であった。

なお、コーディングの信頼性を検討するために、表1のデータに対して、他の2名の評価者によるコーディングを実施し比較・検討した。同一カテゴリーに分類された場合を「一致」、異なったカテゴリーに分類された場合を「不一致」、どのカテゴリーにも分類されなかった場合を「不能」とすると、「一致」がそれぞれ78.6%, 92.8%, 「不一致」が9.5%, 4.8%, 「不能」が11.9%, 2.4%となった。以上により、コーディングの妥当性を確認した。

4.3 インタラクションを含むカテゴリー間の遷移パターンモデル

さて、ここでは2人の被験者間のインタラクションを考慮するために、2者間の仮説形成検証プロセスの遷移パターンのモデルを考える。2者間のO, P, G, T相互の遷移パターンは64種類あり、図3のように示される。このモデルに基づき、全グループの全発話の遷移パターンを集計し、1試行当りに換算した結果を図4、及び図5に示す。図4は全グループの合



O: observing process P: preparatory process
G: generating process T: testing process

図3 インタラクティブな仮説形成検証モデル
Fig.3 A model of interactive formation and verification processes of hypotheses.

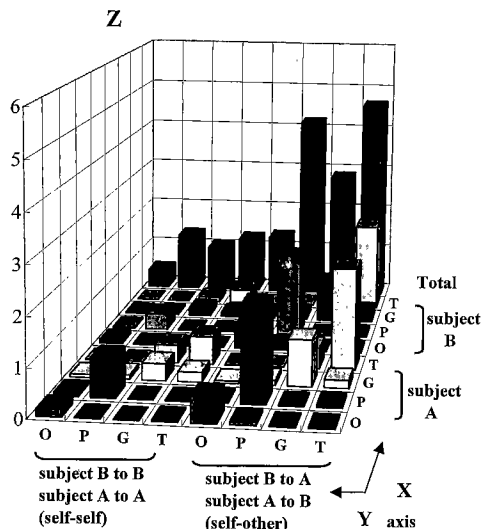


図4 四つのプロセス間の遷移パターン平均出現頻度 (平均)
Fig.4 Occuring frequency of transition patterns among the four processes (average).

計を平均した結果を示し、図5(a)~(i)は各グループ①~⑨の結果を示す。

図4, 図5の縦軸(Z軸)の数値は、各遷移パターンの出現数を各グループの試行回数で割った値、すなわち、1試行当りの出現数を示している。次にY軸は、発話が遷移する際に、先行する発話と後続する発話の話者が、同一被験者か相手方被験者かによって区別し、それぞれを更に先行発話の種類によって分類している。すなわち、Y軸の右側が、先行発話と後続発話の被験者が異なる場合の遷移の出現数 ($\$a \rightarrow \b , $\$b \rightarrow \a ; $\$ = O, P, G, T$) を示し、左側が、先行発話と続く発話の被験者が同一の場合の遷移の出現数 ($\$a \rightarrow \a , $\$b \rightarrow \b ; $\$ = O, P, G, T$) を示す。また、X軸については、被験者A先行の遷移の場合か、被験者B先行の遷移の場合かによって分けている。すなわちX軸の下段が、被験者Aの発話が先行する場合の遷移の出現数 ($\$a \rightarrow \$*$; $\$ = O, P, G, T, \# = a, b$) を示し、上段がBの発話が先行する場合の遷移の出現数 ($\$b \rightarrow \$*$; $\$ = O, P, G, T, \# = a, b$) の分類を示す。そして、グラフの背面には Total として、右半分には、先行発話と後続発話の被験者が異なる場合における遷移の、先行発話の種類ごと (O, P, G, T) の出現数の合計を示す。同様に先行発話と後続発話の被験者が同じ場合の合計を左半分に表示す。以下では、例えば、「Gプロセスからの遷移」、つまり「先行発話がGプロセスに分類された発話からの遷移」を、単に「Gプロセスの発話」と表すことにする。なお、Totalに関しては、後続発話の種類ごとに集計した合計も求めたが、先行発話ごとの場合とほぼ同様の結果が得られたので、図4, 図5においては省略する。

4.4 モデルによる結果の解釈

最初に、全グループを平均した図4から以下の知見が得られる。図4において、同一被験者の発話が続く場合(図4のX-Y平面の左半分)と交互に2人の被験者の発話が続く場合(図4のX-Y平面の右半分)を比較すると、後者が多いことは明白であり、一般的に2人の被験者間で情報が活発に交換されていることがわかる。また、仮説形成検証プロセスの面からは、具体的に仮説形成に言及するG(生成プロセス)の発話よりも、その前後に続く事例吟味や仮説検証を行うP(準備プロセス)やT(検証プロセス)の発話が多くなっていることも特徴的である。

次に、個々のグループに着目する。まず、成績の良

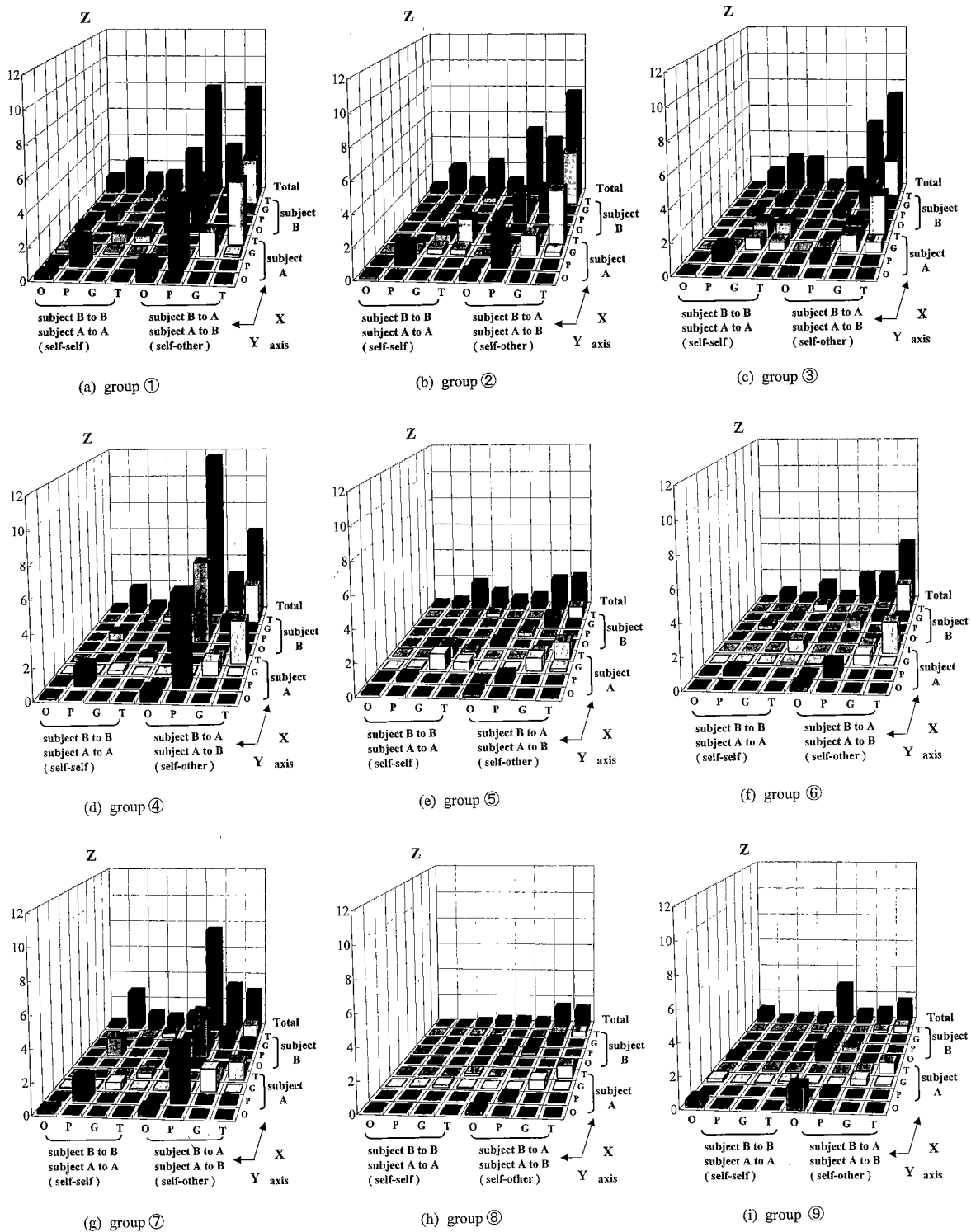


図 5 四つのプロセス間の遷移パターン出現頻度 (各グループ)

Fig. 5 Occurring frequency of transition patterns among the four processes (each group).

表 2 O, P, G, T プロセスの出現数と試行回数との相関
Table 2 The correlation between the number of O, P, G, T processes and the number of trials.

		self-self		self-other		total	
		r	p of t-tests	r	p of t-tests	r	p of t-tests
The number of O+P+G+T processes		-0.782	< 0.05	-0.818	< 0.01	-0.851	< 0.005
focusing on a former segment	The number of O processes	-0.453	= 0.22	-0.350	= 0.36	-0.382	= 0.31
	The number of P processes	-0.586	= 0.09	-0.523	= 0.15	-0.553	= 0.12
	The number of G processes	-0.467	= 0.21	-0.694	< 0.05	-0.660	= 0.05
	The number of T processes	-0.599	= 0.09	-0.894	< 0.005	-0.867	< 0.005
focusing on a latter segment	The number of O processes	-0.173	= 0.66	-0.307	= 0.42	-0.279	= 0.47
	The number of P processes	-0.566	= 0.11	-0.529	= 0.14	-0.555	= 0.12
	The number of G processes	-0.456	= 0.22	-0.632	= 0.07	-0.615	= 0.08
	The number of T processes	-0.714	< 0.05	-0.884	< 0.005	-0.874	< 0.005

いグループは、一般的な傾向として、同一被験者の連続する発話、2人の交互の発話ともに、数が多いことがわかる。更に、成績上位のグループ(図5(a)~(d))を見ると、Pの数が多いグループ④と、Tの数が多いグループ②、③、そしてP、Tの数がほぼ同じグループ①に特徴が分かれている。

逆に、正解に至らなかったグループ⑧、⑨の特徴は、各発話プロセスの総数が、正解に至ったグループよりも著しく少ないことである。また、グループ⑨についてはO(観察プロセス)の発話が顕著に多いという他のグループには見られない特徴がある。

より詳細に、1試行当りのO、P、G、T、各プロセスの数、及び、それら4種類のプロセスの総数と、試行回数との相関を求めた結果を表2に示す。表2には先行発話と後続発話の被験者が同じ場合(self-self)、異なる場合(self-other)、合計(total)に分けて相関係数を示している。最上段にはプロセスの総数、以下、先行発話に着目して各プロセス数を集計した場合(focusing on a former segment: 図4、5においてTotalとして表示)、後続発話に着目して各プロセス数を集計した場合(focusing on a latter segment)との相関係数を示している。

まず、O、P、G、T、各プロセスの総数(The number of O+P+G+T processes)と試行回数に負の相関が見られることがわかる。この傾向は先にも指摘したように、同一被験者の連続する発話、2人の交互の発話ともに、同様である。更にその中でも、特にTプロセスとの相関が著しく高い。このことから、まず、発話の中で仮説形成検証に関する発話が多くなることが良い成績につながると考えられる。そして、それら仮説形成検証発話の中でも、特に仮説検証に関

わる発話の重要性が示唆される。

以上の結果については、以下4.5の検討を踏まえて4.6で更に詳しく検討する。

4.5 遷移パターンモデルの系列化からの結果の解釈

4.4では、各カテゴリー間の遷移パターンの出現数に着目することにより検討を行ったが、これは2対間のローカルな遷移パターンといえる。実際には更にグローバルな遷移パターンが存在している。以下では、よりグローバルな遷移を「系列長」、「系列数」という観点から検討する。

一般に仮説形成検証プロセスは、「規則性の観察」→「仮の一般化の提案」→「実験・観察による仮説の予測のテスト」という順に処理されていくと考えられていることは前述した。つまり、一つの仮説は、「テスト」の結果、その仮説が否定された場合、再度「規則性の観察」に戻る、あるいは、「別の一般化の提案」に戻る。また、仮説が肯定されれば、他の事例を観察し「テスト」を続けるというような、ある一定の文脈に沿って処理されると考えられる。この文脈を、4.2に提案した遷移パターンモデルを用いて整理すると、O(観察)→P(準備)→G(生成)→T(検証)となる。この一定の文脈の存在を客観的に同定するために「系列」という考えを導入する。

まず、O→P→G→Tという文脈を一つの「系列」と定義する。そして、この文脈に逆行するプロセスが発生したときには(例えば、G→Pなど)、新しい系列が生じたとみなす。この一つの「系列」に含まれるO、P、G、Tといったカテゴリーの数を「系列長」と定義し、「系列」の個数を「系列数」と定義する。前述の表1の例では、NO.105から114の間に

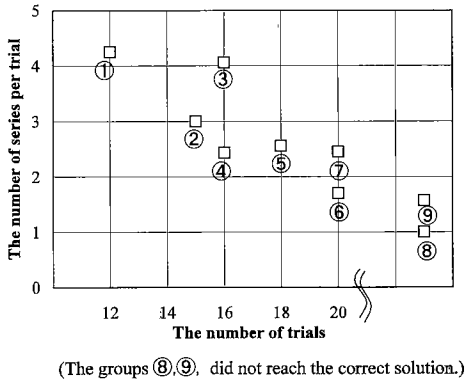


図 6(a) 系列数と試行回数との関係
Fig. 6(a) The correlation between the number of trials and the number of series.

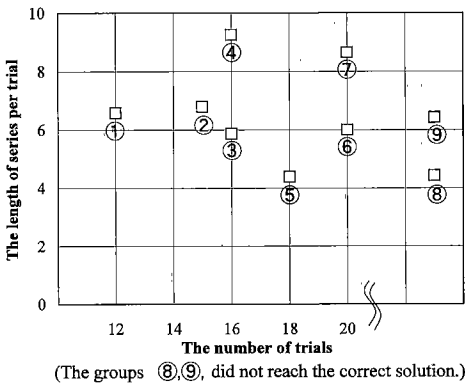


図 6(b) 系列長と試行回数との関係
Fig. 6(b) The correlation between the number of trials and the length of series.

系列長「9」の系列が1個存在し、同様にして、NO. 115 から 125 の間に系列長「10」の系列が1個存在する。よって、NO. 105 から 125 の間の系列数は「2」となる。

このように一定の文脈の存在が、表1におけるプロトコルデータのコーディング例にも観察される。更に、このような系列が存在しないと考えた場合、すなわち、すべてのプロセスがランダムに発生すると考えた場合の「系列長」の期待値は、Oプロセスを起点とするときには1.79、Pプロセスを起点とするとき1.59、同様にGプロセスからは1.44、Tプロセスからは1.33と求められる。以下に示されるように、実際に観察される系列長は、この理論値を大きく上回っており、このことは、O→P→G→Tという文脈の

存在を裏づけている。

さて、実験データから求めた、1試行当りの「系列数」、及び、「系列長」と、「試行回数」との関係を図6(a)、図6(b)に示す。図6(a)から、「系列数」と試行回数に負の相関があることがわかる(相関係数： $r = -0.875, p < 0.005$)。しかし図6(b)からわかるように、「系列長」と試行回数には相関は見られない($r = -0.174, p = 0.66$)。また、「系列長」と解答時間との相関係数を求めたところ、両者の間には正の相関が観察されたことから($r = 0.802, p < 0.05$)、一つの文脈が長くなるグループほど解答時間が長くなるという傾向があることがわかった。なお、正答に達しなかったグループ⑧⑨については、試行回数「21」として算出している。

以上から、試行回数が少ない、高いパフォーマンスを示したグループは、一つの仮説形成検証プロセスを長く続けるよりも、その課題に対応した適切な長さのプロセスを数多く繰り返していることがわかる。ただし、ここで取り上げている課題は、単純な概念同定課題であり、この知見の一般性の検証は、「課題の難易」と「適切な長さ」との関係の解明とともに今後の課題である。

4.6 考察

ここでは、4.4及び4.5で述べてきた知見について、先行する同様の研究と比較しながら考察する。

Klayman and Ha [20]は、単独被験者を対象とした一連の先行研究において、仮説形成検証プロセスにおける「反証と仮説の修正の重要性」を主張してきている。本研究における、「ある系列から次の系列への遷移」は、仮説の修正の機会と考えられ、系列数は仮説修正の回数を反映するととらえることができる。よって、4.5で得られた「より高いパフォーマンスを示したグループは、一つの仮説形成検証プロセスを長く続けるよりも、適切な長さのプロセスを数多く繰り返している」、すなわち、仮説形成検証プロセスにおける「系列数」の重要性を示した結果は、Klaymanらの知見が、協調的な仮説形成検証プロセスにおいても共通することを確認するものである。

Laughlinら [10]は大学生を対象とし、トランプのカードを利用したルール発見課題を用いて、個人と4人組における仮説形成検証プロセスの比較を行った。そして被験者がもつ仮説情報量と証拠情報量の差と成績との関係から、「個人に対するグループの優位性は、より多く仮説を生成できることよりも、証拠から仮説

を評価することがより優れるためである」と指摘している。ここでも、仮説検証プロセス (Tプロセス) の重要性を確認した本実験と同様の結果を得ている。Laughlin らの実験は、仮説だけを交換する群、事例 (本論文でいうメッセージ) だけを交換する群を設定するといった、被験者にとっては不自然な状況を設定しているのに対して、本論文では自由に問題に取り組みさせた後に、実験者が発話内容を分類することで、プロセスの特徴を分類し、同様の結論を得ているという意味において重要である。

個人と2人組における発見プロセスの違いを検討した Okada and Simon [21], Teasley [9] らの研究でも、個人に対するグループの優位性として、「説明活動」や「説明発話」の存在を指摘している。本研究においては、グループ対グループの比較において仮説評価発話と成績との間に高い相関が見られたことから、グループ間での比較においても、仮説を評価する発話が重要であることが新たに観察されたといえる。

5. むすび

本研究は、協調問題解決の場面におけるインタラクションや仮説生成プロセスの状況を、被験者グループごとに詳細な検討、比較を行った結果、次のようにまとめられる。

(1) 「インタラクティブな仮説形成検証モデル」に基づき、成績の良いグループほど、2人間の交互の発話が多く、活発なインタラクションが観察されることを明らかにした。

(2) 実際に特定の仮説に言及する「生成プロセス」に関する発話の交換よりも、その前後に発話される、事例を吟味する「準備プロセス」や仮説の検証を行う「検証プロセス」に関する発話の交換が多いことを示した。

(3) 特に「検証プロセス」の発話が多いグループほど、少ない試行回数で正解に達していることから、仮説の評価、検証にかかわるプロセスの重要性を指摘した。

(4) 仮説形成検証プロセスの遷移の面からは、一つの仮説の検証を長く続けるよりも、検証結果をもとに仮説を修正、あるいは、新たな仮説を生成し、いくつもの仮説を検証していくことの重要性を示唆した。

今後の課題は、インタラクションの環境をコントロールして、その協調作業への影響を検討することである。例えば、被験者間の物理的距離、時間的距離、情

報交換の形態等の違いによって、コンピュータを使用した複数の人間の知的プロセスにいかなる変化が起きるかを検討していくことを考えている。

謝辞 本研究を遂行するにあたり、貴重な御助言をいただきました、名古屋大学大学院人間情報学研究所、峯村吉泰教授、横澤暉教授、横井茂樹教授、寛一彦教授、齋藤洋典教授に感謝の意を表します。本研究の一部は、平成8年度文部省科学研究費補助金基盤研究(C) (No.08680395) および、基礎研究(C) (No.10680377) の援助を受けた。

文 献

- [1] 三宅なほみ, “かかわり合いの統一理論をめざして,” 安西祐一郎, 石崎 俊, 大津由起雄, 波多野諒余夫, 溝口文雄 (編), 認知科学ハンドブック, 第1編, pp.11-20, 共立出版, 東京, 1992.
- [2] D. A. Norman, “Twelve issues for cognitive science,” *Cognitive Science*, vol. 4, pp. 1-32, 1980. (戸田正直, 宮田義郎 (訳), 認知科学のための12の主題, 認知科学の展望, 産業図書, 東京, 1984.)
- [3] 亀田達也, “合議の知を求めて—グループの意思決定,” 日本認知学会 (編), “認知科学モノグラフ,” 第3巻, 共立出版, 東京, 1997.
- [4] J. Shrager and D. Klahr, “Instructionless learning about a complex device,” *International Journal of Man-Machine Studies*, vol. 25, pp. 153-189, 1986.
- [5] D. Klahr and K. Dunbar, “Dual space search during scientific reasoning,” *Cognitive Science*, vol. 12, pp. 1-48, 1988.
- [6] D. Klahr, A. Fay, and K. Dunbar, “Heuristics for scientific experimentation: A developmental study,” *Cognitive Psychology*, vol. 25, pp. 111-146, 1993.
- [7] M. E. Gorman, M. E. Gorman, R. M. Latta, and G. Cunningham, “How disconfirmatory, confirmatory and combined strategies affect group problem solving,” *British Journal of Psychology*, vol. 75, pp. 65-79, 1984.
- [8] M. E. Gorman, “How the possibility of error affects falsification on a task that models scientific problem solving,” *British Journal of Psychology*, vol. 77, pp. 85-96, 1986.
- [9] S. D. Teasley, “The roll of talk in Children’s peer collaborations,” *Developmental Psychology*, vol. 31, no. 2, pp. 207-220, 1995.
- [10] P. R. Laughlin and R. P. McGlynn, “Collective induction: Mutual group and individual influence by exchange of hypotheses and evidence,” *Journal of Experimental Social Psychology*, vol. 22, pp. 567-589, 1986.
- [11] P. R. Laughlin, “Collective induction: Group performance, social combination processes and mutual majority and minority influence,” *Journal of Personality and Social Psychology*, vol. 54, pp. 254-267, 1988.
- [12] P. R. Laughlin, “Collective versus individual induction:

Recognition of truth, rejection of error, and collective information processing,” *Journal of Personality and Social Psychology*, vol. 61, pp. 50-67, 1991.

- [13] J. R. Anderson, “認知心理学概論,” 富田達彦, 増井 透, 川崎恵里子, 岸 学(訳), 誠信書房, 東京, 1982.
- [14] 鈴木裕利, 三輪和久, 櫻井桂一, “協調問題解決下における『発見過程』の研究,” 人工知能学会研究会資料, SIG-HICG-9403-3, 1995.
- [15] J. R. Anderson, R. Farrell, and R. Sauer, “Learning to program in LISP,” *Cognitive Science*, vol. 8, pp. 87-129, 1984.
- [16] K. A. Ericsson and H. A. Simon, “Protocol analysis-verbal reports as data (revised edition),” MIT Press, Cambridge, MA, 1993.
- [17] W. C. Robertson, “Detection of cognitive structure with protocol data: Predicting Performance on physics transfer problems,” *Cognitive Science*, vol. 14, pp. 253-280, 1990.
- [18] A. C. Kruger, “The effect of peer and adult-child transactive discussions on moral reasoning,” *Merrill-Palmer Quarterly*, vol. 38, no. 2, pp. 191-211, 1992.
- [19] 森本志朗, 三輪和久, “階層のスキーマに基づく共同発見学習プロセスの分析,” *認知科学*, vol. 4, no. 2, pp. 3-19, 1997.
- [20] J. Klayman and Y.-W. Ha, “Confirmation, disconfirmation, and information in hypothesis testing,” *Psychological Review*, vol. 94, pp. 211-228, 1987.
- [21] T. Okada and H. A. Simon, “Collaborative discovery in a scientific domain,” *Cognitive Science*, vol. 21, no. 2, pp. 109-146, 1997.

(平成10年1月27日受付, 6月4日再受付)



櫻井 桂一 (正員)

昭48名大・工・電子卒, 昭52同大大学院博士(電気電子)課程中退, 昭52より, 同大プラズマ研究所助手, 昭63より, 同大情報処理教育センター助教授, 平5より, 愛知県立大学教授, 工博, プラズマ工学, 情報処理教育の研究に従事, 日本物理学会, 電気学会, プラズマ・核融合学会, 情報処理学会, 教育工学会各会員。



鈴木 裕利 (正員)

昭53関西大・工・管理卒, 同年(株)リコー入社, 昭58よりリコー情報システム(株), 昭60退職, 平5名大・法・法律卒, 平7同大大学院人間情報学研究所(情報処理理論講座)修士課程了, 現在, 同大大学院博士課程在学中, 認知科学, 情報処理教育の研究に従事, 情報処理学会, 人工知能学会, 認知科学会, 教育工学会各会員。



三輪 和久 (正員)

昭59名大・工・応物卒, 平1同大大学院博士(情報)課程了, 工博, 平1より, 同大情報処理教育センター助手, 平3~4カーネギーメロン大学(アメリカ)客員助教授, 平5より, 名大大学院人間情報学研究所認知情報論講座助教授, 認知科学, 人工知能, 教育工学の研究に従事, *Cognitive Science Society*, 人工知能学会, 認知科学会各会員。