



HOKKAIDO UNIVERSITY

Title	仮説検証過程の適応的变化に関する実験的検討
Author(s)	眞嶋, 良全
Degree Grantor	北海道大学
Degree Name	博士(行動科学)
Dissertation Number	甲第4961号
Issue Date	2000-03-24
DOI	https://doi.org/10.11501/3168527
Doc URL	https://hdl.handle.net/2115/51634
Type	doctoral thesis
File Information	000000353718.pdf



仮説検証過程の適応的变化に関する実験的検討

眞嶋良全

①

学位申請論文

仮説検証過程の適応的変化に関する実験的検討

北海道大学大学院文学研究科

行動科学専攻博士後期課程

眞嶋 良全

目次

はじめに	1
第1章 仮説検証過程に関する心理学的研究概観	5
1.1 仮説検証研究におけるさまざまなアプローチ	5
1.2 規範論的アプローチによる研究	8
1.2.1 確証バイアス (confirmation bias)	9
1.3 適応論的アプローチによる研究	12
1.3.1 確証型検証方略	13
1.3.2 論理的診断性方略	19
1.3.2.1 二重目標課題	20
1.3.2.2 反事実的推論	22
1.3.3 確率論的診断性方略	24
1.3.4 重要エラー検出・最小化方略	28
1.3.5 決定論的アプローチ	30
1.4 仮説検証方略の規定因	33
1.4.1 課題を通じた検証過程の段階的变化	33
1.4.2 受動的情報収集手続きが方略選択に与える影響	34
1.4.2.1 検証事例を自ら選択できない状況 — 受容方略課題	35
1.4.2.2 仮説評価課題における推論の特徴	36
1.4.2.3 仮説検証における情報の解釈・評価	42
1.4.3 受動的情報収集手続きを含む課題における推論特徴の検討	44
1.5 第1章のまとめ	45
第2章 仮説検証過程の動態的変容分析	46
2.1 第1実験 — 規則発見課題における検証方略の段階的变化	46
2.2 方法	46
2.3 結果	50

2.4 考察 — 仮説検証過程の動態的変容プロセス	59
2.5 第2章のまとめ	62
第3章 受動的情報収集手続きを含む課題における方略の変化	64
3.1 仮説検証における受動的情報収集	64
3.2 第2実験 — 受容方略型課題における仮説検証過程	66
3.2.1 方法	67
3.2.2 結果	69
3.2.3 考察	83
3.3 第3実験 — 仮説評価課題における推論の特徴	85
3.3.1 方法	86
3.3.2 結果	88
3.3.3 考察	101
3.4 受動的情報収集手続きによる方略選択の変化	104
3.5 第3章のまとめ	107
第4章 反証型検証の選択を促進する要因	109
4.1 第4実験 — 反証型仮説否定の獲得か、一般的方略としての選択 傾向のシフトか	109
4.2 方法	113
4.3 結果	116
4.4 考察	124
4.4.1 反証型検証の選択に影響する要因	124
4.4.2 仮説検証のダイナミックな変容過程	126
4.5 第4章のまとめ	129
第5章 仮説評価課題における評価者独自の推論過程	131
5.1 第5実験 — 参照学習過程としての仮説評価課題	131
5.2 方法	132
5.3 結果	135

5.4 考察	145
5.4.1 仮説評価課題における評価者独自の推論過程	145
5.4.2 評価課題における事例選択傾向 — 時系列に沿った検証方略の ダイナミックな変化	147
5.4.3 呈示事例の適切性, および情動的価値の評価	150
5.5 第5章のまとめ	153
第6章 仮説検証過程で得られた情報の評価	154
6.1 第6実験 — 4種類の情報に対する評価の比較	154
6.2 方法	158
6.3 結果	161
6.4 考察	166
6.4.1 本実験で得られた結果	166
6.4.2 検証方略の時系列に沿った動情的変容という考え方の一般性	168
6.5 第6章のまとめ	170
第7章 課題の種類と反応傾向についての考察	171
7.1 仮想生物課題と2-4-6問題の構造的差異	170
7.2 仮想生物課題と2-4-6問題における事例選択傾向の違い	172
7.3 仮想生物課題と2-4-6問題における検証事例の適切性評価	178
7.4 第7章のまとめ	180
第8章 総合的考察	182
8.1 本研究で得られた知見とその意義	182
8.1.1 仮説検証過程の動情的側面	182
8.1.2 受動的情報収集手続きの検証方略選択に対する影響	186
8.1.3 事例選択方略モデルの精緻化	187
8.1.4 人の仮説検証方略の選択における適応的合理性	191
8.2 未解決の問題と今後の課題	193
8.2.1 協調的解決過程の検討	193

8.2.2 多面的な観点からの検証過程の検討	196
8.2.3 仮説生成過程の探求 — 科学における発見の文脈	199
8.2.4 検証方略選択における個人差	201
8.3 結語	202
文献	205
謝辞	212

はじめに

われわれが日常的に行う様々な判断，推論，意思決定は，われわれの持つ信念・知識体系に大きく依存する．この信念・知識体系は，多くの経験からの帰納的一般化によって形成されることが多い．この帰納的一般化において中核的な役割を果たすのが，仮説検証 (hypothesis testing) という認知的活動である．狭義の仮説検証は，帰納的推論 (inductive reasoning) の下位コンポーネントとして位置付けられ，様々な情報から生成された仮説の真偽を決定する方法論のことを指すが，広義の仮説検証は，仮説生成 → テスト → 仮説の更新 → テストが繰り返される一連のプロセスを指し，帰納的推論そのものとほぼ同義であると考えられる．本研究では，仮説検証を後者の意味，すなわち仮説の生成からテスト，さらに結果に応じた仮説の更新が繰り返されるダイナミックな認知的プロセスとして捉える．

仮説検証，あるいは帰納的推論は，元来，科学における方法論の問題として，特に科学哲学の領域において検討が重ねられてきており，その歴史は 16 世紀のフランシスコ・ベーコンの時代にまでさかのぼるといわれている (内井, 1995)．その後，近代科学の発展とともに，科学の方法論は，さらに洗練されてきた．しかし，確率判断における確率論，演繹的推論における形式論理学，問題解決におけるアルゴリズムなどと同様に，科学的探求における「科学哲学的な科学方法論」とは，あくまで科学的探求過程の理想像に過ぎず，実際に人がこの理想的方法論に従っている保証はない．実際に人が行っている仮説検証

過程の特徴を検討すべく始められた、仮説検証に関する心理学的研究は、その40年あまりの歴史においてめざましい発展を遂げてきた。

本研究の主要な目的の第一は、仮説検証方略が時間を追って変化するプロセスをとらえ、記述することである。従来の仮説検証研究では、被験者の反応、特に検証事例の選択傾向を全体的な反応パターンとして静的 (static) にとらえる傾向があった。しかし、本来仮説検証とは、その過程の中で得られた情報を元に、必要であれば新たな可能性を模索するという極めて動的 (dynamic) な認知的活動であり、それを全体的な傾向として平均化してしまうことは、この認知プロセスの重要な側面を検討の対象から外してしまうことになる。本研究では、このような観点にたち、被験者の仮説検証過程が、どのような段階を経てどのように変化するかを、時系列に沿って分析、検討する。

本研究の目的の第二は、これまでの研究で言われてきたような確証型検証の多用傾向、すなわち確証バイアスの仮説検証における位置付けを再検討することである。これまでのところ、他者の仮説を評価する仮説評価課題においては、一般的な仮説検証課題においてみられる確証型検証の多用傾向が軽減され、反証型検証が用いられやすくなることが知られている (Kareev & Halberstadt, 1993; Farris & Revlin, 1989; Tschirgi, 1980)。この結果に関しては、仮説生成が要求されないことによる認知的負荷の軽減という説明がなされているが、本研究では、むしろ、この評価課題と検証課題における課題手続き上の重要な相違が影響を与えると考える。仮説評価課題は、その名の示す通り、他者の検証過程を観察する課題であり、一般的な仮説検証課題における仮説生成、検証事例の

選択が能動的に行われないという特徴を持つ。本研究では、この受動的情報収集活動にこそ、評価課題と検証課題の間に見られた反応パターンの違いを生じさせる原因となるものがあると考ええる。

以上のような目的を達成するため、本研究では 6 つの実験を行った。本論文の構成を概略すると、第 1 章では、これまでの仮説検証過程に関する研究、中でも心理学的研究の成果を概観し、その問題点を指摘する。特に、前述の、検証過程の段階的变化を分析するという視点の重要性 (動態的変容分析パラダイム) について考察する。また、仮説検証研究における課題のバリエーションとして、規則発見課題、受容方略型課題、仮説評価課題という 3 種類の課題における違いを構造的に比較し、特に受容方略型課題、仮説評価課題の持つ受動的情報収集手続きが検証過程に与える影響について考察する。

第 2 章の第 1 実験では、動態的変容分析パラダイムの考えに基づき、規則発見課題における被験者の反応を、時系列に沿って分析する。また、従来の研究ではあまり注目されてこなかった、仮説の取捨選択の過程についても、同様に動態的変容分析パラダイムを適用して分析する。

第 3 章の第 2, 第 3 実験では、受動的情報収集手続きが検証過程に与える影響を調べるが、第 2 実験では受容方略型課題を、第 3 実験では仮説評価課題を用い、それぞれの課題における推論過程、特に事例選択傾向の特徴を吟味する。また、得られた結果に対して、反証型仮説否定情報の獲得と、一般的傾向としての時系列に沿った段階的な方略の変化という 2 つの点からの解釈を試みる。また、続く第 4 章の第 4 実験では、このうちいずれがより強力な規定因である

かについて検討する。

第5章の第5実験では、第3実験で見られた、検証過程で得られる情報の価値評価についてさらに詳しく調べると共に、評価課題における反応の一般性を確かめるために、伝統的な2-4-6問題を仮説評価課題に転用した課題を用いた実験によって検討する。また、第6章の第6実験では、仮説検証過程で得られる様々な情報に対する評価を、系統的に調べることができるよう修正を加えた課題を用い、情報の価値評価の過程を分析する。

第7章では、本実験で使用した、カバーストーリーが異なる2種類の課題、すなわち第1～第4実験で用いられた仮想生物課題、および第5、第6実験で用いられた2-4-6問題における検証方略の違いを、両課題における構造的差異という観点から考察する。

第8章では、第1～第6実験で得られた結果をまとめ、仮説検証方略の選択に影響を及ぼす諸要因、特に時系列に沿った段階的な方略の変化、受動的情報収集手続きの効果について総合的に考察する。さらに、本研究において検討の対象とならなかった話題を取り上げ、その解決を含む、今後の発展的研究の方向性について提案する。

第1章 仮説検証過程に関する心理学的研究概観

我々の持つ信念・知識体系の構築において中心的な役割を果たすのが、仮説検証過程である。仮説検証過程に関する心理学的研究は、人が仮説に沿う確証事例を多用することを示した Wason (1960) に端を発し、その後様々な研究が行なわれている。本章では、これら一連の研究を、その立脚点の違いから、規範論的アプローチと適応論的アプローチに分類し、それらの知見を概観する。また、従来の研究が抱えている問題点についても考察する。

1.1 仮説検証研究におけるさまざまなアプローチ

仮説検証に関する研究は、大別すると前述した規範論的アプローチ、および適応論的アプローチに分けられる。規範論的アプローチでは、反証主義 (falsificationism; Popper, 1959) という規範的モデルとの対比から、人の仮説検証過程の合理性を議論することが中心的な話題とされてきた (e.g. Mynatt et al., 1977, 1978; Snyder & Swann, 1978; Wason, 1960)。これら一連の規範論的アプローチ研究では、人が反証主義の方法論に従わず、仮説に沿った確証事例を多用する確証バイアスと呼ばれる傾向を持つことが明らかとなり、人間の非合理性を示す証拠として扱われてきた。

それに対して適応論的アプローチでは、人がさまざまな状況に応じた推論を行っているということを主張してきた。例えば、Klayman & Ha (1987) は、人間の非合理性の証拠として考えられてきた確証バイアスを、確証型検証方略

(positive test strategy) という適応的ヒューリスティックスとしてとらえ直すべきであると述べ、多くの状況において確証事例による検証が合理的な選択となることをベイズの定理 (Bayes' theorem) の考え方に基づいて議論した。また、前述の研究と並行する形で進められた別の研究では、特に規則発見課題で正解に到達した被験者が、仮説およびそれと競合する対立仮説とを区別するのに最も適した検証事例を選択する診断性方略 (diagnostic strategy) を用いているということが示された (Baron & Beattie, & Hershey, 1988; Farris & Revlin, 1989; Freedman, 1992; Gorman, Stafford, & Gorman, 1987; Klayman & Ha, 1989; Oaksford & Chater, 1994a; Platt, 1964; Skov & Sherman, 1986; Slovic et al., 1992; Trope & Bassok, 1982, 1983; Tukey, 1986; Tweney et al., 1980; Wason & Johnson-Laird, 1972; Wharton, Cheng, & Wickens, 1993)。さらに、Freidrich (1993) は、検証方略の選択における動機づけの側面を強調し、被験者がそれぞれの状況で設定している目標という要因を重視する「重要エラー検出・低減モデル (Primary Error Detection and Minimization: PEDMIN)」を提唱した。これらの適応論的アプローチに共通する主張は、被験者が状況に応じた推論を行うために用いている方略が、外見上は確証事例の多用としてあらわれ、2-4-6問題のような特定の状況において特異的に課題解決の障害になりうるという考え方である。

また、近年の傾向として、検証事例の選択を一種の意思決定状況として捉え、検証の結果得られるであろう情報の認識的効用 (epistemic utility) を重要視する研究も見られるようになってきた (e.g. Kirby, 1994; Oaksford & Chater, 1994b; Poletiek, 1995)。これらの研究は、仮説検証研究における決定論的アプローチ

と呼ぶことができるであろう。この決定論的アプローチは、認識的効用の変化に伴って事例選択も変化することを予測し、状況に応じた推論を行うという基本的主張を持つ適応論的アプローチの一環として捉えることが可能である。

本研究は、当該の問題状況に応じて、人の仮説検証方略の選択が適応的に変化することを主張する適応論的アプローチの立場に立ち、そのダイナミックな変容過程を分析することを目的とする。本研究では、検証方略の変化をもたらす要因として、検証過程全体を通じたフェーズの移行、および被験者が仮説検証状況において情報を獲得する手続きないし手段の2つに注目する。本研究において第1に取り上げる要因は、検証過程を通じたフェーズの移行である。規範論的アプローチ、適応論的アプローチ、あるいは決定論的アプローチに基づく従来の研究のほとんど全ては、被験者の事例選択を、全体的かつスタティックに記述する傾向があり、本来ダイナミックに進行するであろう認知プロセスの重要な側面が取り上げられてこなかったという問題点を持っている。例えば現実場面で行われる科学的探求では、観察されたデータに基づいて生成された複数の仮説群の中からもっとも可能性が高そうなものを1つ抽出し、さらなるデータ収集によってその仮説をより洗練されたものへと更新していくという方向性を持った推論が進行すると考えられる。このとき、複数の可能な仮説群から1つを抽出する場合と、抽出された仮説を改訂していく場合とでは、異なる方法が用いられるという可能性は十分に考えられる。しかし、従来のような全体的分析では、このように段階を経て変化するプロセスを十分に記述することはできないため、本研究では、検証過程のダイナミズムを積極的に検討する

動態的変容分析パラダイムを提唱し、認知的プロセスの進行過程を時間軸に沿って分析することを試みる。このことによって、人がどのように生成された仮説の真偽を決定し、結論を導くのかを、より現実場面に即した形で記述することが可能になると思われる。

本研究において第2に取り上げる要因は、情報を獲得する手段の違いである。仮説検証研究で一般的に用いられる規則発見課題では、被験者自身が仮説を生成し、検証事例を選択するという手続きが採用される。一方、仮説や検証事例が自発的に生成・選択されない状況として仮説評価課題 (hypothesis evaluation task)、受容方略課題 (reception strategy task) と呼ばれる課題が存在する。これらの、受動的情報収集手続きを有する課題における推論の特徴を分析することは、人の状況適応的な推論過程の理解に大きな貢献を果たすことであろう。本研究では、この2つの課題状況下における推論の特徴について詳細に検討する。

本研究の基本的立場、および本研究の目的は以上に述べた通りである。以下の節では、従来の仮説検証研究を概観し、そこで得られた知見、未解決の問題点について詳述する。

1.2 規範論的アプローチによる研究

仮説検証過程に関する規範論的アプローチは、検証過程の規範的モデルとして、科学哲学における反証主義を採用するという共通点を持つ。反証主義とは、科学的探求において、仮説を否定する可能性が高い反証事例を用いるべきであると主張する考え方である。一般に科学的理論は「全ての A は B である」とい

う全称命題の形をとるが、全称命題が真であることを証明するためには、潜在的に可能な事例を全て列挙する必要がある (完全枚挙法)。例えば、「すべてのカラスは黒い」という命題が真であることを証明するためには、過去、現在、未来において存在しうる全てのカラスを調べ、その色が黒いことを確かめなくてはならない。しかし、現実的にはそのような方法が実行できないことは明らかであり、確証型検証によって全称命題の真偽を決定することは事実上不可能であるといって良い。一方で、全称命題はたった一つの否定事例によって偽であることが証明される。例えば、前述のカラスの例では、アルビノのカラスは、「全てのカラスは黒い」という命題を否定する。そのため、科学的探求の試みにおいては、仮説を否定する試みを続け、仮説が否定されない限りにおいて、暫定的にこれが正しいものであると扱うべきであるというのが、反証主義の主張する規範的な科学的方法論である。

1.2.1 確証バイアス (confirmation bias)

Wason (1960) は、2 - 4 - 6 問題あるいは規則発見課題と呼ばれる課題を用いて、人が用いる仮説検証方略の記述を試みた。2 - 4 - 6 問題とは、以下のような課題である。被験者は、「2 4 6」という 3 つの数字から成る数字列を呈示され、この数字列を構成する規則を発見するよう求められる。この時、被験者は、仮説を自由に生成することができ、また仮説を検証する事例も自由に選択することができた。被験者が選択した検証事例に対するフィードバックはその都度与えられ、仮説に絶対的な自信が持てるようになった段階で、被験

者は実験者にその仮説を報告した。この実験の結果から明らかになったことは、被験者は反証主義が主張するような仮説を否定する試みをするのではなく、仮説を支持する可能性の高い確証事例を多く用いた検証を行う傾向があることであった。Wason 自身は、この傾向を検証バイアス (verification bias) と呼んだが、後に確証バイアス (confirmation bias) と呼ばれ、人の非合理性を示す証拠であると考えられた (Mynatt et al., 1977, 1978; Snyder & Swann, 1978)。2 - 4 - 6 問題において、被験者が発見すべき規則は「上昇系列 (ascending numbers)」というものであったが、多くの被験者は初期事例である「2 4 6」という数字列が発見すべき規則に含まれるターゲット集合事例であることから、「連続する偶数」「公差が 2 の等差数列」といった仮説を生成することが多かった。被験者は、前者の仮説を検証するためには「8 10 12」、後者の仮説を検証するためには「1 3 5」といった事例を用いることが多かった。これらの事例は、仮説に沿った確証事例であり、この確証事例を多用することから確証バイアスと呼ばれてきた。なぜこの事例選択がバイアス、すなわち規範からの逸脱となるかという点、例えば、「連続する偶数」という仮説に対する「8 10 12」という検証事例は、仮説に沿った確証事例であり、この検証事例によって仮説が否定されることはない。このような事例のみによって仮説を検証した場合、誤った結論を導く結果となる。実際に、Wason (1960) の実験に参加した被験者の多くが、正しくない結論を実験者に報告する結果となった。

Mynatt et al. (1977, 1978) は、この確証バイアス傾向が、2 - 4 - 6 問題とい

う極めて抽象的な課題に限定されるものではないことを示すために、人工空間 (artificial universe) 課題と呼ばれる、より現実在即した課題を用いて被験者の仮説検証方略を分析した。この課題は、コンピューターディスプレイ上に+でマークされた位置から発射される粒子 (particle) の運動を説明する規則を発見するというものであった。この人工空間はいくつかのディスプレイに分割され、あるディスプレイには (1) 灰色の三角形の領域、別のディスプレイには、(2) 黒い三角形及びそれに隣接する灰色の円形の領域が存在した。規則は、「粒子は灰色の領域に近づくと停止する」というものであった。被験者は、何が粒子を停止させるかを調べるために、2つのディスプレイのうち1つを選択することができた。この研究において明らかになったことは、例えば、「三角形の領域」という仮説を立てた被験者は、(2) のような他の灰色の図形がある方よりも、(1) の灰色の三角形のみがあるディスプレイを選択する傾向があり、「灰色の図形」というような対立仮説を立てないということであった。すなわち、一般的に被験者は、対立仮説の検証が可能な選択肢を選ばなかったのである。これにより、Mynatt et al. は、Wason (1960) が示した確証傾向が、2-4-6問題のような抽象的課題に限定されたものではないことを指摘した。

さらに、Snyder & Swann (1978) は、社会的相互作用の中で行われる、他者の個人的属性についての知識を得ようとする試みの中で用いられる仮説検証方略を分析した。被験者の半分は対象人物が外向的であるという仮説を検証し、残りの半分は対象人物が内向的であるという仮説を検証するよう求められた。また被験者は、仮説を検証するために必要だと思われる質問を、あらかじめ用

意されていた 26 個の質問項目の中から 12 個選択するように教示された (インタビュー・パラダイム)。26 個の質問項目は、11 個の外向型質問 (対象人物が外向的であったなら肯定的回答が期待される質問)、10 個の内向型質問 (対象人物が内向的であったなら肯定的回答が期待される質問)、5 個の外向性－内向性とは無関連な中性的質問の 3 種類で構成されていた。この実験の結果、被験者は、質問に対する肯定的回答が仮説を支持するような質問を選択する傾向があることが確かめられた。すなわち、外向的であるという仮説を検証するよう求められた被験者は外向型の質問を多く選択し、一方、内向的であるという仮説を検証するよう求められた被験者は内向型の質問を多く選択することが示された。例えば、「対象人物が外向的であるかどうか」を確かめるために、「パーティーに行くのは好きですか」や「パーティーが活気づくようにするためには、あなたは何をしますか」といった外向型の質問がなされる傾向が見られた。また、対象人物の属する集団における外向的な人の割合を変化させたり、適切な検証方略を採用した場合に報酬を与えると教示することでモチベーションを高めても、確証的証拠を求める傾向はなくなることが示された。この結果から、確証バイアスは、2－4－6 問題や人工空間課題のような科学的探求状況のみならず、日常的な状況においても生じることが示された。

1.3 適応論的アプローチによる研究

確証バイアスという名が示すとおり、この確証事例を多用する傾向は、規範に対する逸脱、人間の非合理性の証拠として扱われてきた。しかし、近年、この

確証傾向を、現実場面への適応を重視した方略としてとらえなおそうとする試みが進められている。このような一連の試みを仮説検証に対する適応論的アプローチと呼ぶこととする。この適応論的アプローチに共通している点としては、規範論的アプローチとは異なり、人が従うべき規範としての反証主義という前提を持たないことが挙げられるであろう。むしろ、このアプローチは、仮説検証過程における反応傾向は、日常的場面に対する適応的合理性を反映していると主張する。

仮説検証研究の適応論的アプローチには、確証事例を多用する傾向を肯定的な側面からとらえなおした確証型検証方略 (positive test strategy) を主張する研究、競合する 2 つの仮説間の区別を最も効率的に行う診断性方略 (diagnostic strategy) を主張する研究、当該の状況における目標という動機的側面を強調する PEDMIN モデルを主張する研究、検証事例の選択を意思決定状況としてとらえ、検証事例の持つ情報的価値、すなわち認識的効用 (epistemic utility) を重要視する研究が存在する。以下の節で、それぞれについて詳述する。

1.3.1 確証型検証方略

ヒューリスティックス (heuristics) は、一般に問題を解決する場合に、正解にたどり着く保証はないものの、認知的負荷が低く実行が容易な解決方略であると言える。Klayman & Ha (1987) は、Wason (1960) などが示した確証バイアスは、単なる規範からの逸脱としてよりは、むしろ、適応的合理性を重視したヒューリスティックスとしてとらえられるべきであることを指摘し、これを確証

型検証方略 (positive test strategy) と呼んだ¹。

彼らは、2 - 4 - 6 問題 (Wason, 1960) において典型的に生成される初期仮説の種類に注目した。ここで、仮説に対する確証事例の集合 (仮説集合) を H ，真の規則に対する正事例の集合 (規則集合) を T とすると、 H と T の間には、図 1-1 に示されるような状況が存在しうる。ところで、2 - 4 - 6 問題で初期事例として与えられる「2 4 6」という数字列から典型的に生成される初期仮説は、「連続した偶数」「公差 2 の等差数列」などである。従って、この問題においては、被験者が典型的に生成する初期仮説は、「上昇系列」という規則の下位集合となりやすくなっており、その関係は、 $H \subset T$ と表現される (図 1-1 a) に相当)。この状況の下で、仮説の確証事例のみを用いて検証を行うと、不十分な仮説が否定されることがなく、正しい規則に到達することはできない。しかし、仮説の確証事例を用いた検証の方が効率よい検証となる状況も考えられる。例えば、仮説集合と規則集合が部分的にオーバーラップする b) のような状況では、仮説の確証事例を用いた検証によって仮説を否定する結果を得る可能性がある。これは例えば、2 - 4 - 6 問題で被験者が発見すべき規則が「3 つの偶数」で、被験者の立てた仮説が「2 ずつ増える数」といったような状況である。ここで問題となるのは、仮説の確証事例の集合 (H) と反証事例の集

¹ 本来 positive test strategy は、正事例検証方略とでも訳すべき用語であるが、正事例という言葉は、正のフィードバックを生じた事例 (positive instance) との区別が困難になるため、以後確証型検証方略という用語で統一する。

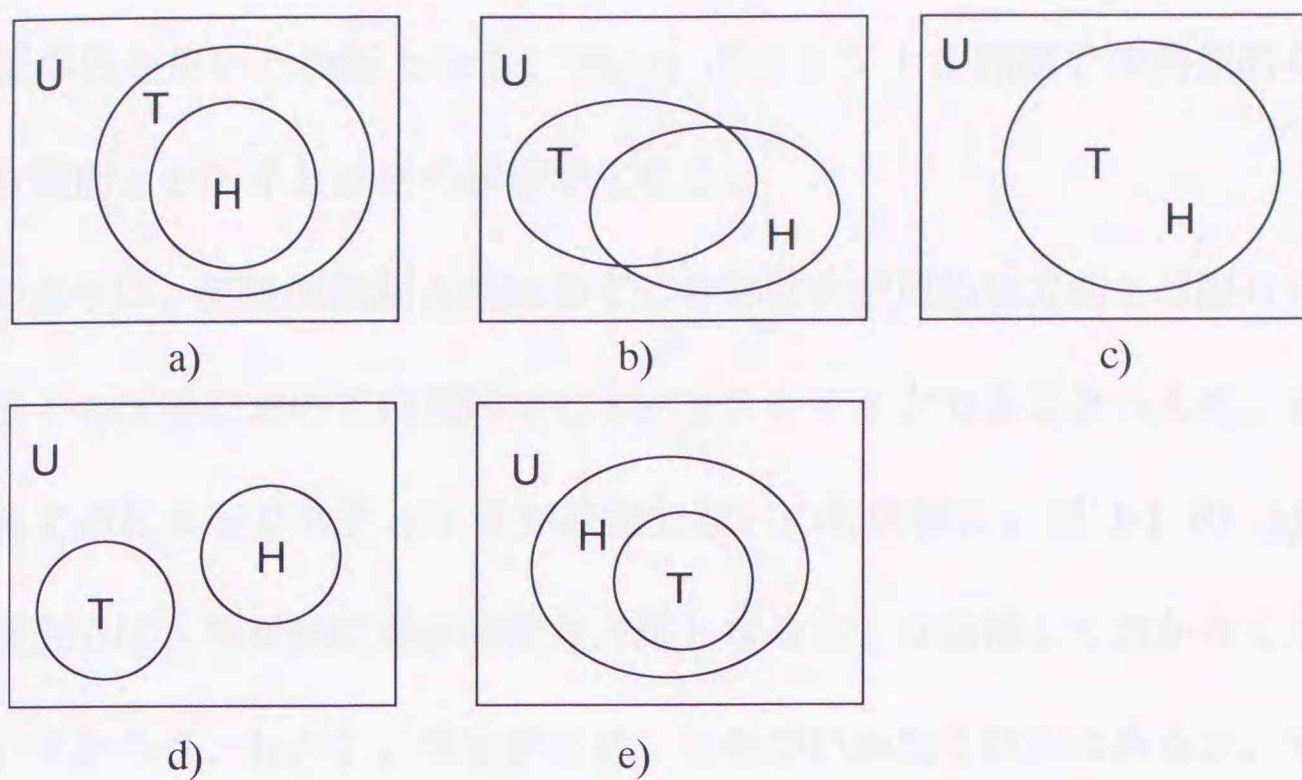


図1-1 規則集合と仮説集合の論理的構造のタイプ

(U : 全体集合, T : 規則集合, H : 仮説集合)

合 ($not-H$) の相対的な大きさの違いであるが、日常的には $not-H$ の方が大きい方が普通であると考えられる。そのような状況の下では、 $not-H$ を検索して、それが T に含まれるかどうかを確かめるよりは、 H を検索し、それが $not-T$ に含まれるような場合がないかどうかを確かめるほうが効率のよい検証であるといえる。さらに 図 1-1 の a) とは逆に、仮説集合が規則集合の上位集合となっている 図 1-1 の e) のような状況では、 $not-H$ を用いた検証を行った場合、決して仮説の予測に反した結果を得られないことになる。例えば、2-4-6 問題において、実験者が設定している真の規則が「連続する偶数」であり、被験者の立てた仮説が「2 ずつ増える数」であるといったような状況がこれにあたる。この場合、仮説が誤っていることを証明する唯一の方法が、仮説に対す

る確証事例を用いた検証となり，Wason のオリジナル課題での典型的な初期仮説と規則との関係とは逆の状態が生じる。

このように，確証型検証方略は必ずしも常に非合理的な方略とは限らず，むしろ多くの状況において合理的なヒューリスティックであるといえる。もちろん，全てのヒューリスティックが万能でないのと同様に，図 1-1 の (a) のような状況では，特異的に非合理的な方略となることは指摘しておかなくてはならないであろう。しかし，現実的には，自分がいかなる状況にあるか，すなわち仮説集合と規則集合の関係がどのようなになっているかを事前に知り得る方法はない。仮説検証において決定的に重要なのは，仮説に沿わない事例を用いて検証することではなく，仮説が誤っている場合に，それを決定的に示す証拠を得やすい方略を選択することである。オリジナルの 2-4-6 問題においては，仮説の反証事例を用いた方略がそれに該当したが，常にそれが成り立つとは限らないし，現実的に仮説集合が規則集合の下位集合となるような状況，すなわち図 1-1 の a) のような状況がどれほど多いかも疑問が残る。

そこで，Klayman & Ha (1987) は，仮説を否定する可能性が高い検証を選択する方略について，以下の 4 つの指標を用いて定式化した。

$p(t)$ ：全体集合のうち規則集合に含まれる事例のベースレート

$p(h)$ ：全体集合のうち仮説集合に含まれる事例の比率

$z^+ = p(\text{not-}t | h)$ ：正の予測が誤っている比率

$z^- = p(t | \text{not-}h)$ ：負の予測が誤っている比率

上記の， z^+ ， z^- は，仮説の誤りを示す指標であるといえる。ここで，

$$z^+ = z^- \times \frac{p(\text{not} - h)}{p(h)} + \left(1 - \frac{p(t)}{p(h)}\right) \quad (1)$$

という関係が存在する。このとき、被験者が生成する仮説が、少なくともその集合の大きさにおいて規則集合と変わらない、すなわち、被験者が規則集合の大きさを過小評価、あるいは過大評価しないと仮定すると、 $p(h) = p(t)$ と考えて良い。この場合、式 (1)は、

$$z^+ = \frac{p(\text{not} - t)}{p(t)} \times z^- \quad (2)$$

と置き換えることができる。したがって、 $p(t) < .5$ であるならば、 $z^+ > z^-$ 、すなわち、確証型検証 (+H test) により仮説の誤りが示される方が、反証型検証 (-H test) により仮説の誤りが示される場合より確率が高くなる。これは、全体集合に対する規則集合の大きさが比較的小さい現象 (minority phenomenon) に関して仮説検証を行う場合、確証型検証の方が、反証型検証よりも仮説に対する否定を得やすいことを示している。現実場面における仮説検証は、ほとんどの場合、この minority phenomenon に関するものであるといえる。例えば、Klayman & Ha (1987) の挙げた例では、惑星を持つ恒星の比率は 1/3 に過ぎなく、精神分裂病の患者は全人口の 1%に満たなく、AIDS の発症率は 10^{-4} から 10^{-5} であり、また Wason (1960) の 2-4-6 問題における「上昇系列」という規則の事例でさえ、全体の 1/6 にしかならないことが示されている。また、 $p(t)$ が .5 を越えるような場合には、被験者はターゲットと非ターゲットを入れ替えて扱うということが知られている (c.f. Bourne & Guy, 1968;

Einhorn & Hogarth, 1986). これらのことから、確証型検証方略は、多くの場合において仮説が誤っていることを示しうる適応的合理性をもつヒューリスティックであるといえる。

Klayman & Ha (1989) は、以上の考え方に基づいて、確証型検証方略についてさらなる実証的検討を加えた。彼らは、2-4-6問題を題材とし、典型的な初期仮説が規則に対して「下位集合 (embedded: 図 1-1 の a) に相当)」、「部分的重なり (overlapping: 図 1-1 の b) に相当)」、「上位集合 (surrounding: 図 1-1 の e) に相当)」となるように修正した課題を用いて、被験者がどのような検証方略を用いるかを調べた。例えば、「2 4 6」という数字列が与えられた場合に典型的に生成される仮説は、「連続する偶数」や「2 ずつ増える数」である。このとき、被験者が発見すべき規則は、下位集合条件では「上昇系列」、部分的重なり条件では「一桁の数字」、上位集合条件では「一の位が 2 4 6 である偶数」であった。Klayman & Ha は、このような課題を用いて、被験者が検証に用いる事例を分析した。その結果、全体的傾向として、確証型検証が多く使用され (約 66%)、反証型検証はきわめて少なかった (7%)。この結果は、Wason (1960) をはじめとする先行研究と同様に、被験者は基本的に確証事例による検証を多用する傾向をもっていることを示している。また、典型的な初期仮説と発見すべき規則の関係についての操作は効果がなく、全ての条件で確証型検証に対する強い選好が見られた。確証型検証に過度に依存することは、ターゲット事例の見落とし (式 (1) の z) の検出の失敗を招くものであるが、多くの場合は、 z^+ の方が z より目につきやすく、かつ重要であると考えられ

る (Klayman & Ha, 1987, 1989). 一方, 21%の試行では, 仮説と競合する対立仮説を立て, 両者の間の比較検討を行う診断性方略 (diagnostic strategy) と呼ばれる検証法が用いられることが示された. 特に, 正解到達者と未到達者を比較した場合, 到達者の方が確証型検証を行う回数が相対的に少なく, 代わりに対立仮説の検討を行う診断性方略が多いこと, 上位集合条件では確証型検証の数が増え, 結果として仮説の否定の数がほかの条件に比べて多いことなどが示された. ここで, 診断性方略という新たな概念が出てきたが, 以下の節でこの方略について検討する.

1.3.2 論理的診断性方略

確証事例による検証を積極的かつ肯定的に評価する確証型検証ヒューリスティックとは異なる観点から, 確証バイアスというとらえ方に異論を唱える一連の研究がある. これらの研究では, 人は情報の持つ診断的価値に敏感であるということが指摘されており, このような方略を診断性方略 (diagnostic testing strategy) と呼んでいる. 診断性方略の研究では, 科学的探求過程を題材にした競合仮説間の論理的構造に注目するものと, 社会的認知を題材にした特性の出現頻度の確率的分布に注目するものの2通りの流れがあり, それぞれが一定の成果を挙げている.

競合する仮説間の論理的構造という観点からみた診断性方略は, 競合する2つの仮説に対して排他的な検証を行い, 一方を棄却するという方略である (Freedman, 1992; Klayman & Ha, 1989; Platt, 1964; Wason & Jonson-Laird, 1972).

確認バイアスにせよ、確認型検証方略にせよ、単一の仮説の検討における検証事例の選択傾向について述べたもので、その有効性は発見すべきターゲットとの関係で決定され、独立に検討することができない (三輪, 1999)。しかし、診断性方略は、2 つの相反する仮説を同時に検討し、さらに一方の仮説に対しては確認事例、もう一方の仮説に対しては反証事例となるようなテストを選択するため、フィードバックが正負のいずれであっても、かならずどちらかの仮説が否定されることとなる。そのため、結果として仮説の否定が多く得られることとなり、極めて有効な検証法となる。しかし、この方略は、外見上は仮説に対する確認事例を選択しているように見えるため、確認バイアスとの区別をつけにくい。競合する二つの仮説に対する検証が、極めて有効であることを示した研究としては、二重目標課題 (dual goal task) と呼ばれる、2 - 4 - 6 問題の変形課題を用いたものがある (Gorman et al., 1987; Tukey, 1986; Tweney et al., 1980; Wharton et al., 1993)。以下に、二重目標課題研究の成果を概観する。

1.3.2.1 二重目標課題

Tweney et al. (1980) は、2 - 4 - 6 問題の DAX-MED ヴァージョンと呼ばれる変形課題を用いて、この規則発見課題において課題解決が促進されることを示した。この課題は、基本的手続きや課題の論理構造はオリジナルの 2 - 4 - 6 問題と同一であったが、フィードバックがイエス・ノーの形ではなく、それぞれ DAX, MED というラベルで与えられたこと、それに伴い、DAX に対応する規則を発見するという教示が与えられたことが異なっていた。ここ

で、DAX に対応する規則は、オリジナルの規則と同じ「上昇系列」であり、MED はそれ以外の事例、すなわち「上昇系列」の補集合に対応した。このような状況の下では、被験者は確証型検証を多用し、反証型検証をほとんど行わなかったが、DAX, MED カテゴリーのそれぞれに対応する規則を別々なものとして扱い、それぞれに対して確証型検証を行った結果、全体的な成績は向上するという結果が見られた。この課題では、一方の仮説に対する確証は、他方に対する反証となるため、表面的には確証型検証を多用しているように見えるものの、それぞれの仮説に対して確証的な検証を行うために、一方の仮説については確証型と反証型の混合戦略を適用することになる。このように2つの仮説を考慮に入れた診断性方略を用いた結果、成績が向上したと考えられる。

なぜ二重目標課題において成績が向上するのかについては、情報量の差、および目標の補完性という2通りの説明が考えられている。情報量の差に原因を求める考え方は、単一目標課題であるオリジナルの2-4-6問題に比べて、二重目標課題の方が、純粹に獲得される情報量が多いために解決が促進されると主張する (Farris & Revlin, 1989; Gorman et al., 1987; Tweney et al., 1980; Wetherick, 1962)。一方、目標補完性説は、単に情報量が多いかどうかということではなく、単一目標課題で見過ごされる情報の利益に注目する (Evans, 1989; Gorman et al., 1987; Klayman & Ha, 1987; Politzer, 1986; Tukey, 1986; Tweney et al., 1980; Wharton et al., 1993)。これは、補完的規則に関する情報を集めることで、単一目標では得られない情報が獲得され、解決が促進されることを主張する説明である。従って、被験者が目標の補完性に気づかないと、単一

目標課題と二重目標課題のパフォーマンスに差がなくなることが予測される。この問題を検討するため、Wharton et al. (1993) は、単一目標課題、および DAX と MED の間の論理的構造をあいまいにした二重目標課題の間の比較 (実験 1)、単一目標課題、および DAX と MED が論理的に補完的構造にあることを明言した二重目標課題の間の比較 (実験 2) を行った。その結果、前者の場合は二重目標であっても成績は改善されず、後者の場合は二重目標において成績の改善が見られることが示された。この結果は情報量説を否定し、目標補完性説を支持している。すなわち、二重目標課題において解決が促進されたのは、2 つの補完的仮説を同時に検討することで、単一目標課題では得られない情報を獲得することによるものであるといえる。

1.3.2.2 反事実的推論

二重目標課題は、競合する仮説間で診断的 (diagnostic) なテストを行うことによって、全般的な課題解決の促進を導くものであった。ところで、オリジナルの 2-4-6 問題では、1 回目の規則報告において正しい規則を報告できた被験者が極めて低いことが知られているが、少ないながらも 1 回目の報告において正しい規則を挙げることはできた被験者も存在する。このような正解到達者は、未到達者とどのような違いがあるのであろうか。この問題について調べた興味深い研究がある。

Farris & Revlin (1989) は、2-4-6 問題で正解に達した被験者が、反事実的推論 (counterfactual reasoning) を用いていることを指摘した。この方略は、

まず初期事例「2 4 6」を観察した後で、例えば「偶数」という仮説を生成した場合に、続いてこの仮説と補完的な関係にある「奇数」仮説を考えることから始まる。そして、この仮説が真であると仮定して、この補完的仮説に対して確証型検証を実施する。もし、補完的仮説に対する確証型検証がイエスという結果を導いたならば、オリジナルの仮説（ここでは「偶数」）は否定されることになる。もし、結果がノーであったならば、オリジナルの仮説が正しいという可能性が高まり、さらにこの手続きが繰り返されることになる (Farris & Revlin, 1989; Oaksford & Chater, 1994a)。この方略は、Platt (1956) が *strong inference* と呼んだ、複数の競合する仮説を同時に検証する方略であり、前述の診断性方略と類似した方略であるといえる。

この反事実的推論は一種の背理法 (*reductio ad absurdum*) のようなものであると考えることができる。反事実的推論においてイエス反応が得られた場合、補完的仮説である「奇数」仮説のもっともらしさが高くなるように思えるが、この仮説は、初期事例から生成された「偶数」仮説の補集合であるために、もともと生成時点で誤りであることが確定しており、新たな別の仮説が生成されることとなる。しかし、この反事実的推論については、論理学的な意味での補完的仮説 (*complementary hypothesis*) と、反対仮説 (*opposite hypothesis*) とを混同していることが指摘されている (Oaksford & Chater, 1994a)。前述の例でいえば、「偶数」仮説と「奇数」仮説は、論理的な意味で補完的関係にあるとは言えない。これは、例えば、「2 5 8」という事例は、どちらの仮説に対しても確証事例とはなりえないことから明らかである。Oaksford & Chater

(1994a) は、この論理的矛盾を指摘した上で、反事実的推論方略を修正した反復性反事実的推論方略 (iterative counterfactual strategy) によって被験者の反応をとらえなおすことを試みた。この方略は、元々の反事実的推論において、イエスという結果を導く事例に焦点を当てている。前述の例における「奇数」という補完的仮説に対して、「3 5 7」という検証を行った結果、イエスという結果が返ってきた場合は、オリジナルの「偶数」仮説（「3 5 7」により否定）と補完的仮説である「奇数」仮説（「2 4 6」により否定）の組み合わせは共に棄却される。しかし、この補完的仮説に対して行われた検証は、新しい仮説を生成するための、付加的な情報源として利用される。ここで「2 4 6」と「3 5 7」は、共に規則集合の事例であるため、両者に共通した特性が抽出され、新たな別の仮説が生成される。新しい仮説は、さらに別の反事実的推論により検証され、この手続きが反復される。このように、反事実的推論は、いわゆる反証型方略が想定するような否定式 (*modus tollens*) ではなく、背理法の手法に基づいて行われる仮説否定の試みであるといえる。

このように、論理的側面からみた診断性方略に共通しているのは、規則発見に成功する被験者が、2 つかそれ以上の仮説を同時に心内に保持し、課題解決を促進しているという考え方である。

1.3.3 確率論的診断性方略

確率論的診断性方略の存在を主張する研究は、人が用いる検証方略が、仮説の予測する事例が当該の特性を持つ確率という情報を考慮に入れた最適な方略で

あり、人は情報の持つ診断的価値に敏感であると主張する。これらの研究は、Snyder & Swann (1978) から始まる、社会的認知を題材に用いたインタビュー・パラダイム課題を採用しており、確率論的診断性方略は、仮説と対立仮説の下での出現頻度が最も異なる特性を質問として選択する、すなわち、正のフィードバックが仮説を確認するか対立仮説を確認するかに関わらず、両仮説を最も良く区別する質問を選択する方法であると定義される (e.g. Baron et al., 1988; Skov & Sherman, 1986; Sloviaczek et al., 1992; Trope & Bassok, 1982, 1983)。

Skov & Sherman (1986) は、被験者が仮説検証において、まず診断性の高い質問を選択する傾向がある、すなわち診断性に対する関心がデータの選択を誘導することを示した。彼らは、それまでの仮説検証研究において指摘された確認型方略には、(1) 考慮している仮説のみに関連する情報検索を行なう傾向、(2) 検証者が情報を収集する際に、仮説が真となるように解釈を歪める傾向、(3) 仮説のもっともらしさを過大評価させるような情報検索を行なう傾向という3種類が混同されていること、またそのような確認型方略とは別に診断的情報を検索する方略が存在することを指摘した。

(1) の意味での確認型方略は、仮説の下での出現頻度が高く、対立仮説の下での出現頻度の低い特性に関する質問を行なうという傾向である。この検証に対する正の反応、すなわちその特性を持っているというフィードバックは仮説を確認し、負のフィードバックは対立仮説を確認する。この論理に従うと、例えば事例を分類するカテゴリーがグループ I、II と 2 つあるような場合に、グループ I のメンバーの 50% が持っており、グループ II のメンバーは 10% しか

持っていないような特性は、ある事例がグループ I に属するという仮説を検証するために用いられるということが予測される。ここで、正のフィードバックは仮説に対する確証となるが、両グループの集合の大きさが等しい場合は、全体集合において負のフィードバックが占める割合の方が多いことが予想され、仮説が否定される場合が多くなると考えられる。

確証型方略の (2) は、情報検索というよりは情報解釈の方略であるといえる。この問題に関連する研究として Lord, Ross, & Lepper (1979) は、死刑に対する被験者の初期信念が客観的には中性的な情報の解釈を歪めることを指摘している。彼らは、死刑廃止論者と死刑存続論者に、死刑の犯罪抑止効果に関する 2 種類の調査報告を与え、被験者がその情報をどのように解釈し、また信念がどのように変化するかを検討した。一方の報告書には、死刑の犯罪抑止効果を支持する証拠が含まれ、もう一方の報告書には死刑の犯罪抑止効果を否定する証拠が含まれていた。また、それぞれについて、調査の種類として同じ州での継時比較を行なった研究と異なる州での共時比較を行なった研究を用意した。被験者の半分は、継時比較－死刑存続支持研究と、共時比較－死刑廃止支持研究を与えられ、残りの被験者は継時比較－死刑廃止支持研究と、共時比較－死刑存続支持研究を与えられた。すなわち、死刑廃止論者にも死刑存続論者にも、それぞれ半分ずつ、考えを支持する継時比較研究と考えに反する共時比較研究という組み合わせか、その逆が提示された。この研究の結果、被験者は自分の主張に沿った研究に対しては、調査の種類が継時比較か共時比較かに関わりなく、死刑の効果について重要な証拠を提供する良い研究であると判断する一方

で、自分自身の主張に反する研究に対しては、調査手法がどうであれ研究方法上の欠陥を多く指摘する傾向が示された。また、調査報告書を読むことによって、被験者の態度は更に両極化するという結果が示された。すなわち、どちらの立場の被験者も、自らの主張に対して相反する証拠を提示されたにも関わらず、当初の自分の主張をさらに強くするといえる。この結果は、確率論的に言えば、規範に違反している。すなわち、自らの主張に対して肯定、否定双方の結果が得られた場合、少なくとも自らの主張に対する確信度が高くなることは理論的にはありえないため、この結果は情報解釈のバイアスであると言える。

Skov & Sherman (1986) はまた、Trope & Bassok (1982, 1983) らによる一連の研究において示された診断的方略についても論じている。Skov & Sherman は、過去の診断的方略の研究 (Trope & Bassok, 1982, 1983) が持つ方法論的問題点を解決した上で、特定のオブジェクトがグループ I, II のいずれに属するかを決定するためにどのような方略を用いるかを調べた。彼らは特に、2 グループ間で出現頻度の大きく異なる特性に関する質問、すなわち診断性が高い質問項目を選ぶか、あるいは特定のグループにおいて出現頻度が極めて高い (あるいは逆に低い) 特性に関する質問、すなわち確証的な質問項目を選ぶかどうかという点に注目した。その結果、質問項目の診断性が主要な規定因であること、確証的であるかどうかは副次的ではあるが効果を持っていることが示された。これは、例えば、同じ診断性の高い質問項目の中でも、確証的な質問項目の方がより高く好まれるという形で見られ、否定的情報の処理の難しさによるものであると考えられた。

1.3.4 重要エラー検出・最小化方略

適応論的観点から見た仮説検証における方略選択の記述的モデルとしては、以上に挙げてきた確証型検証方略や診断性方略とは異なり、検証者が当該状況において持っている目標に注目したものもある。Friedrich (1993) は、人の仮説検証方略が、単なる確証傾向をもつというよりはむしろ、現在の目標の下で最も重要かつ避けるべきエラーを最小化するという方向性を持っていると主張した。彼は、この傾向を重要エラーの検出及び最小化 (primary error detection and minimization; PEDMIN) と呼び、人が日常的に行う仮説検証においては、ターゲットに属すると思っていた事例が実際には属さないというエラー、すなわち仮説確証事例－規則負事例 (false positive: $+H-T$) の検出と回避が主要な目標となることが多いことを指摘した。例えば、企業の人事担当者が、セールスマンの採用にあたって、「外向的な人ならば、セールスマンとして成功する」という仮説を持っているという状況を考える。通常、企業の人事担当者にとって重要かつ避けるべきエラーとなるのは、セールスマンとして成功しない人を採用してしまうことであり、この場合では外向的な人がセールスマンとして成功しない、 $+H-T$ 事例に当たる。この状況では、セールスマンとして成功しない外向的な者の検出が主要な目標となるため、外見上は確証型検証が行われているように見えるはずである。この場合は、仮説が真の規則に対して「十分」であればよく、「必要」であるかどうかはそれほど問題とはならない。しかし、一方で、仮説の必要性が重要になる状況もあり、その状況下で何が主要な関心

事になるかによって、外見的な検証方略が変化することがありうる。例えば、前述した企業の人事担当者の例で、「男性ならば、セールスマンとして成功する」という仮説が保持されていたとしよう。また、この人物が採用に際して雇用機会均等法に注意を払わなければならない状況下にあったとすると、男性を採用してセールスマンとして成功するかどうかを見ること以外に、女性でセールスマンとして成功する者がいるかどうかの確認も重要な目標となる。この場合、セールスマンとして成功する女性、すなわち仮説反証事例—規則正事例 (false negative: $-H+T$) の見落としが重要となるため、女性の確認が行われることとなる。これは、一見すると反証型方略を採用しているかのように見えるだろう。

Friedrich (1993) は、仮説検証研究において観察された、ある意味で一貫性の欠けた行動パターンが、この重要エラーの検出及び最小化という枠組みの下で統一的に把握できると主張した。この枠組みでは5つの仮定が挙げられているが、そのうち重要と思われるのは、以下の3つである。(1) 人は通常、真実の発見 (truth detection) よりも誤りの低減 (error reduction) に関心を持つ素人功利主義者 (lay pragmatist) である。(2) 上記のような関心のために、人は誤りの低減を目標とした質問を選択する、すなわち十分性 (sufficiency) の検証を重要視する。(3) 素人の検証方略は、検証の文脈に応じて、系統的・予測的に変化する。このモデルに従うと、これまでの仮説検証研究において観察された確証型検証の多用は、当面の状況にうまく対処するために、とりあえず十分な仮説を構築しようとする試みの現れであると解釈される。また、これまでの

研究で反証型検証の選択数が極端に少なかったのは、被験者が仮説の必要性をそれほど重要視していなかった結果によるものであると考えられる。

1.3.5 決定論的アプローチ

仮説検証過程を適応論的側面からとらえようとする研究の流れと並行して、仮説検証過程、中でも検証事例の選択を一種の意思決定場面と考え、決定理論の概念を導入してモデル化しようとする試みもなされている。この一連の研究では、過去の仮説検証に関する研究が、論理的な確証と反証の概念に依存しており、反証を意図した方略が確証的な方略よりも優れているとされてきたこと、さらにこの考え方は、実験者が「反証的」あるいは「確証的」とラベル付けする検証行動と被験者の意図との間連づけが希薄であるといった問題点を抱えていることが指摘されている (Klayman & Ha, 1987; 眞嶋, 瀧川, 1997; McDonald, 1992; Poletiek, 1995)。確率論的な概念であるベイズの定理を導入したアプローチを用いることで、検証事例の選択に加え、アポステオリな検証結果の解釈をも考慮に入れることができ、過去の研究で扱われてこなかった、検証事例の選択と結果の解釈との間の関係を説明することができる。Poletiek (1995) は、それまでの確率論の概念を導入した仮説検証の研究では、確率的な刺激材料が用いられているため、被験者に統計的な推論を要求するという不利な点を併せ持つことを指摘した上で、極力確率的な刺激を用いることなく、検証行動の中でベイズの定理がどのように用いられうるか、このモデルが直感的に理解可能であるかどうかを示すことを試みた。

ベイズの定理によると、ある観察の頻度が、仮説が与えられた時の頻度に比べて低ければ低いほど、それが観察されたときに仮説に対する強い支持を与えるとされる。これは、以下のように定式化される。仮説 H により予測される観察 A, B があった時に、 A を観察した後の方が B を観察した後よりも信念の更新の度合いが高くなるのは、

$$\frac{p(H|A)}{p(H)} > \frac{p(H|B)}{p(H)} \quad (3)$$

すなわち、 $p(H|A) > p(H|B)$ の時である。ベイズの定理より、以下の式が導き出される。

$$\frac{p(A|H)}{p(A)} > \frac{p(B|H)}{p(B)} \quad (4)$$

これは、仮説が与えられたときの A の頻度に比べて、 A の観察頻度が低い (B と比べて) こと、言いかえると相対的にまれな現象である方が、それが観察されたときにより強く仮説を支持することを意味する。例えば、「あるメロンはおいしい」という仮説を考えてみよう (c.f. Poletiek, 1995)。ここで、おいしい (*tasty*) メロンは黄色 (*yellow*) でかつ甘い匂い (*sweet*) がすることを知っていたとすると、 $p(\text{yellow} | \text{tasty}) = p(\text{sweet} | \text{tasty}) = 1$ となる。ここで、ほとんどのメロンの色が黄色 ($p(\text{yellow}) \doteq 1$) であったなら、色という次元に注目し、あるメロンが黄色であるという結果を得たとしても、そのメロンがおいしいメロンであるという仮説には強力な支持を提供し得ない。逆に、メロンの半分だけが甘い匂いがする ($p(\text{sweet}) = .5$) ことがわかっているとすれば、

甘い匂いという観察は、相対的に強い支持を与えることになる。従って、この場合は、色による検証よりも匂いによる検証の方が、仮説の真偽について強力な証拠を提供することができる。

Poletiek (1995) は、直接的な統計的推論を避け、なおかつベイズ的推論過程が、仮説検証において果たす役割について検討するために、Wason (1960) の 2-4-6 問題に基づき、仮説を支持する結果を得る確率の高い検証と低い検証のどちらが好まれるか、さらに被験者がどのようにそれらを解釈するかを信念の更新という意味で観察する実験課題を用いた。その結果、被験者は仮説を支持する確率の高い検証法を好み、この確証的結果を仮説に対する弱い証拠であるとはみなさなかった。これは、被験者が、検証前の確証的結果の確率と、検証後にこれらのテストが持つ証拠としての価値が負の相関関係にあることに気づいていなかったことを意味する。この結果から、被験者が、前述のベイズ流規範モデルに従っていないということがいえる。検証者が自らの検証結果の解釈に関して示すバイアスは、以下の 2 つが考えられている (Poletiek, 1995)。

まず第一に、確証的結果を高い確率で得る検証法を用いるとき、その結果の証拠的価値を過大評価する傾向がある。第二に、確証的結果を得る確率が低い検証法を用いるときは、確証的結果の価値を過小評価する傾向がある。仮説確証型の検証法は、一般に好まれる傾向があるため、検証結果の証拠的価値の過大評価は頻繁に起こりうると考えられる。

1.4 仮説検証方略の規定因

これまでの仮説検証研究では、被験者が、少なくとも表面上は当該の仮説に対する確証型検証を多用しているように見えることを示してきた。しかしそれは、必ずしも単純な確証型検証というわけではなく、さまざまな状況に適応的に対応した結果の産物であると考えられる (Farris & Revlin, 1989; Friedrich, 1993; Klayman & Ha, 1987; Oaksford & Chater, 1994a; Wharton et al., 1993)。これらの研究は、被験者が単に外見上確証型検証をしているように見えるだけであり、実際は状況に応じた適応的方略を採用していることを主張する。本研究も基本的な立場として、この適応論的アプローチに立ち、被験者の方略使用を規定する要因の探求を行う。本研究では、被験者の方略選択を規定する要因として、時系列に沿った検証過程の段階的变化、および情報収集手続きの違いによる変化について取り上げる。

1.4.1 課題を通じた検証方略の段階的变化

これまでの仮説検証研究 (Gorman & Gorman, 1984; Gorman et al., 1987; Klayman & Ha, 1987, 1989; Mynatt et al., 1977, 1978; Skov & Sherman, 1986; Tukey, 1986; Tweney et al., 1980; Wason, 1960; Wetherick, 1962; Wharton et al., 1993) には、仮説検証過程全体を通して、確証事例および反証事例がどの程度使用されたかによって仮説検証方略を記述しようとする傾向があった。これは仮説検証過程を全体的、スタティックにとらえる研究であるといえる。一方、現実の仮説検証過程は、仮説の生成 → テスト → 仮説の更新 → テストというように、データ

からの仮説の構築および更新と、その仮説に対する検証とが交互に繰り返される、極めてダイナミックなプロセスである。そのプロセスにおいて、常に確証事例のみによる検証を行ったり、あるいは逆に反証事例のみによって検証を行うということは考えにくく、むしろ、両方の事例をうまく併用していると考えたほうが自然である。そして、確証事例・反証事例も、常にある一定の割合で選択されるのではなく、最初は複数の可能な仮説群から抽出された1つの仮説に対する確信度を高めるために確証型検証を行い、後にその仮説に不備がないかどうかを確認し、より洗練されたものにするために確証型検証と反証型検証を併用するということも考えられる。

このような事例選択傾向のシフトを観察するためには、従来のように、両事例の全体的な選択比率を求めるのではなく、解決に至る全過程をいくつかの段階に分け、それぞれの段階における事例の選択比率を算出し、相互に比較することが必要となるであろう。本研究では、この動態的変容分析パラダイムを、規則発見課題における被験者の反応の分析に適用し、方略選択が時間を追ってどのように変化するかを検討する。

1.4.2 受動的情報収集手続きが方略選択に与える影響

本研究では、以上に挙げた時系列に沿った方略選択の変化に加えて、被験者の情報収集方法の違いに起因する方略選択の変化にも注目する。仮説検証過程の研究には、科学的探求状況をシミュレートした規則発見課題が用いられることが多い。ここで、科学的探求状況が、どのような下位プロセスによって構成さ

れるかを考えてみると、(1) 状況の理解、(2) 仮説の生成、(3) 検証事例の選択、(4) 検証結果の解釈、(5) 必要に応じた仮説の更新という要素から構成され、これらが動的に繰り返されるプロセスであると言える。多くの仮説検証研究で用いられてきた規則発見課題は、これら5つの要素を全て含んでいるが、課題によっては、これらの要素を含まないものが存在する。以下に、そのような課題の例として、受容方略課題 (reception strategy task)、および仮説評価課題 (hypothesis evaluation task) を挙げる。

1.4.2.1 検証事例を自ら選択できない状況 — 受容方略課題

受容方略課題は、Bruner, Goodnow, & Austin (1956) が用いたことで有名な概念達成課題である。彼らは、概念を達成する課題の手続きとして選択方略 (selection strategy) と受容方略 (reception strategy) の2つを挙げ、両手続き下における推論過程の違いを検討した。彼らが用いた課題は、4次元のそれぞれが3つの属性をもつ事例集合を2つに区分する概念的規則を発見するものであった。選択方略課題では、この概念的規則の発見に際し、被験者自身が事例を1つずつ選択し、どちらの集合に属するかを尋ねることができた。しかし、受容方略課題では、事例の選択は被験者の自由意思にはよらず、実験者からの事例呈示を受動的に観察することしかできなかった。その意味で、受容方略課題とは、現実世界における観察学習に相当すると言えるであろう。多くの日常的な信念や態度が、このような観察により形成されと考えられることから、その重要性は明らかである。

Halberstadt & Kareev (1995) は、規則発見課題において、予め用意された事例の中から情報を獲得する受容方略型事例獲得手続きと、自由に事例を選択できる通常の選択方略型事例生成手続きのいずれかを被験者が自ら選択できる課題を用いて、受容方略型手続きが常に選択方略型手続きより先に選択され、受容方略型手続きから選択方略型手続きへの移行は、正事例を得た後に生じやすいことを示した。また、事例の生成においても、直前の試行で正事例を得た場合に、次に少ない数の次元に注目した、統制された事例生成を行なう傾向 (fine-tuning) が強く、逆に負事例を得たり、直前の試行が受容方略型手続きであった場合に、より多くの次元の属性値を変化させた事例生成を行なう傾向 (looking-for-target-items) が見られた。しかし、仮説検証という研究領域において、受容方略型手続き下での推論について検討した研究は、この他には全く見られない。

1.4.2.2 仮説評価課題における推論の特徴

前述の受容方略課題が、検証事例を自由に選択できない仮説検証であるのに対し、仮説評価課題は、仮説を生成することなく、検証事例を選択することもない、客観的立場から他者の検証過程を観察する状況である。「岡目八目」という言葉の示す通り、このような客観的立場から評価が行われる場合は、被験者はより規範に近い反応を示すことが知られている (Farris & Revlin, 1989; Kareev & Halberstadt, 1993; Tschirgi, 1980)。

Tschirgi (1980) は、日常的場面での因果推論を題材に用い、物語中に登場す

る人物が行う仮説検証過程を被験者に評価させた。被験者は、表 1-1 のような状況を与えられ、登場人物が物語の結末に影響するであろう 3 つの要因の重要性に関する仮説を検証するという状況において、その人物がどのように仮説を検証すべきかを指摘した。被験者は a. ~ c. で示される選択肢のうち、登場人物がとるべき行動はどれかを選択した。ここで、a. ~ c. には、CA, HOTAT, VOTAT という別名が記されているが、CA (上記の a) は、全ての次元の属性値を変化させるという選択肢、HOTAT (b) は、重要であると考えられる「説明の長さ」という次元はそのままに、その他の次元の属性値を変化させる選択肢、VOTAT (c) は、HOTAT と反対に、重要である「説明の長さ」という次元のみを変化させ、その他は変化させないという選択肢である。この場合は、VOTAT がいわゆる反証型検証に相当し、論理的に正しい選択肢であると言える。Tschirgi (1980) は、成人被験者が全体的に VOTAT を多く選択することを示した。この結果については、通常の仮説検証課題との手続きの差、すなわち仮説生成手続きの有無に原因があると考えられた。通常の仮説検証課題、すなわち規則発見課題では、被験者に対して仮説の生成と評価の両方が求められる (Baron, 1985; Weimer, 1977) が、評価課題では仮説の評価が求められるものの、生成は要求されないため、課題遂行にかかわる認知的負荷が低いと考えられる。

Farris & Revlin (1989) は基本的には Tschirgi (1980) の主張に賛同しつつ、仮説評価課題における認知的負荷の低さに加えて、記述中の出来事の結末の個人的望ましさという要因の影響を吟味した。Farris & Revlin (1989) は、Tschirgi

表1-1 日常的因果推論を題材に用いた仮説評価課題の例

マークは、仕事の面接に行って、新しいアプローチを試した。

- ・ 長々と説明するのではなく簡潔に答えた
- ・ スポーツシャツとスラックスではなくスーツを着て行った
- ・ 面接時間ギリギリではなく、15 分前に着くように出かけた

面接はうまく行かず、あっけなく、また非生産的だった。マークは、面接に落ちた理由は、簡潔すぎる答えがいけなかったと考えた。彼は、スーツを着ること、15 分前に行くことは問題ないと考えた。

彼のこの考えが正しいかどうかを確かめるためには、どうしたら良いか？

- a. 次の面接では、彼は長い説明をし、スポーツシャツとスラックスを着て、面接時間ちょうどに着くべきだ。

(Change All: CA)

- b. 次の面接では、彼はもう一度簡潔な説明をこころがけ、スポーツシャツとスラックスを着て、面接時間ちょうどに着くべきだ。

(Hold One Thing At a Time: HOTAT)

- c. 次の面接では、彼は簡潔にではなく長い説明をするようにするが、スーツを着ること、15 分前に着く事は今まで通りにするべきだ。

(Vary One Thing At a Time: VOTAT)

(1980) の用いた課題に基づいて仮説評価課題における反応を検討した結果、Tschirgi (1980) と同様に、評価課題では反証型検証に相当する VOTAT の選択が多いこと、またこの結果は記述中の結末の望ましさによって影響され、結末が個人的に望ましくない場合は VOTAT が選択され、逆に結末が個人的に望ましい場合には、HOTAT の選択が増加することを示した。結果の望ましさによって選択される方略が異なることについては、課題において課せられる認知的負荷に加えて、重要な要素を保持することの効用という付加的な文脈の影響が指摘された。例えば、上記の面接の状況においては、結末が面接に失敗したという否定的な事象であったため、VOTAT のコストは低いといえる (仮説が正しければ今より望ましい状況になり、誤っていたとしても現状より悪くはない)。逆に、結末が面接に成功したという肯定的な事象である場合は、VOTAT のコストは高い (仮説が正しければ今より状況は悪化し、仮説が誤っていた場合は現状維持)。しかし、あくまで全体的な傾向としては反証型検証が選択される傾向にあり、結果の望ましさの効果は二次的なものであると考えられた。

しかし、Farris & Revlin (1989) の研究には重大な問題がある。まず第一に、彼らが仮説評価課題として、規則発見状況ではなく重要な変数を分離する能力に焦点を当てた因果推論課題を用いていること、第二に、被験者が反証型検証を選択できるという主張を支持する証拠が実験条件と交絡していることが挙げられる。第一の問題点については、Farris & Revlin (1989) は、通常の仮説検証課題としては伝統的な規則発見課題である 2 - 4 - 6 問題を使う一方で、仮説

評価課題としては、Tschirgi (1980) の因果推論状況下で重要な条件を操作するかどうかを選択させる課題を用い、構造の異なる両課題での反応を比較したことが問題であるといえる。そのため、通常の仮説検証過程と仮説評価の間での反応傾向の違いは、彼らが言うような課題における認知的負荷の違いではなく、むしろ課題の構造そのものが直接的に影響を与えている可能性が高い。

例えば、演繹的推論課題として知られる 4 枚カード問題 (Wason, 1966) では、抽象的な規則で記述されるオリジナルのカード問題と比べて、課題の構造は等価だが、より具体的な内容で記述された飲酒問題において正解率が高くなることが知られている (Cheng & Holyoak, 1985; Grrigs & Cox, 1982; Johnson-Laird, Legrenzi, & Legrenzi, 1972; Manktelow & Evans, 1979)。4 枚カード問題とは、表にアルファベット、裏に数字が書かれたカードが存在し、このカードが「表のアルファベットが母音であれば、裏の数字は偶数である」という規則にのっとって作成されているという状況において、この規則が正しく守られているかどうかを確認する方法を問うものであり、単一試行での仮説検証方法を選択する課題であるといえる。被験者は、例えば、A, K, 4, 7 と書かれた 4 枚のカードの中から、規則の真偽を確認するためには、最低限どのカードの反対側を確認する必要があるかを尋ねられる。ここで、カードの表に母音が書かれていることを p 、子音が書かれていることを $not-p$ 、裏に偶数が書かれていることを q 、奇数が書かれていることを $not-q$ であらわすと、前述の規則は「 $p \Rightarrow q$ 」と表現することができる。このような命題は条件文 (conditional) と呼ばれる。論理的には、 $p \Rightarrow q$ と等価なのは、 $not-q \Rightarrow not-p$ であり、これを対偶

(contraposition) という。4 枚カード問題における正解は p カード (A) と $not-q$ カード (7) を選択することであるが、多くの被験者は $not-q$ カードを選択しない、すなわち対偶の確認を怠るという誤りを示すことが知られてきた (e.g. Wason, 1966; 1968)。一方、この 4 枚カード問題では、問題の構造が同じでありながら、表面的な表現が変えられただけで正解率が劇的に変化することが知られている。例えば、Griggs & Cox (1982) は、「酒を飲んでいるならば 20 歳以上でなければならない」という規則が守られているかどうかを確かめるという状況で、ビールを飲んでいる人物、コーラを飲んでいる人物、16 歳の人物、22 歳の人物のうち、誰を確認する必要があるかを問うという課題を用いた。彼らは、このような課題では、被験者が、論理的に正しい「ビールを飲んでいる人物」「16 歳の人物」を選択できることを示した。具体的な状況、特に被験者の有している知識、経験と密接に結びついた状況において論理的に正しい解を選択できることについては、Cheng & Holyoak (1985) による実用的推論スキーマ説 (pragmatic reasoning schema theory) によって一応の説明がされている。実用的推論スキーマ説は、人が日常的によく遭遇する状況に関する知識体系 (スキーマ) を持っており、この知識体系に合致する状況では、特に論理的操作を意識しなくとも、外見上は規範的な解を導き出すことができるという主張である。彼らは特に、飲酒問題では「許可」、「義務」というスキーマが発動されると指摘している。例えば、飲酒問題における条件文は、「20 歳以上であれば酒を飲んでも良い」という許可スキーマに合致する内容に置き換えることができるため、スキーマが発動することで、正しい推論ができるといえる。実用

的推論スキーマの種類には、4 枚カード問題の飲酒バージョン課題で発動されるような「義務」ないし「許可」の他に、「因果」「共変性」といったものが考えられている (Cheng & Holyoak, 1985; 高橋, 服部, 1996)。

ここで Farris & Revlin (1989), Tschirgi (1980) が用いた仮説評価課題は、因果的推論場面を扱った課題であることを考えると、仮説評価課題における反証型検証の選択頻度の増大は、仮説評価課題において要求される認知的負荷の低さではなく、使用した題材が規範的に正しい反応を導きやすい具体的な状況であったことに起因するのではないかという疑問が生じてくる。すなわち、この日常的因果推論場面を扱った仮説評価課題では、「因果」の実用的推論スキーマが働いたために、規範的に正しい選択肢が選択されやすかったのではないかと考えられる。

1.4.2.3 仮説検証における情報の解釈・評価

Farris & Revlin (1989) に対する批判としてさらに重要なのは、前述の第二の問題、すなわち被験者が反証型検証を選択できるという主張を支持する証拠が実験条件と交絡しているという問題である。彼らの研究の被験者は、望ましくない状況が生じている場合において、反証型方略に相当する VOTAT が用いられる傾向が強固になるとされていた。しかし、Kareev & Halberstadt (1993) は、この傾向を、被験者が仮説を確証しようとしたか反証しようとしたかという分類で見るべきではないと指摘した。彼らは、被験者が、物語中の人物がどのように振舞えば否定的な出来事を避けられるかに基づいた選択をしているのであ

って、否定的な事象が記述された状況において、それを避ける方法と反証型検証に相当する VOTAT とがたまたま一致したに過ぎず、Farris & Revlin (1989) が示した仮説評価課題における反証型検証、すなわち仮説否定 (refutation) の試みが多く見られたという傾向は、方法論上のアーティファクトに過ぎないと述べた。Kareev & Halberstadt (1993) は、この主張に基づき、規則発見課題における反証型検証の有効性、および仮説を否定する情報の有効性がどのように評価されるかを実験的に検討した。彼らは、先行する試行で得られた情報の違いによって、後に行われる検証の種類に対する有効性の評価が異なるかどうかを調べるため、2 - 4 - 6 問題に参加した架空の被験者の記述を与えた。先行情報としては、(1) 仮説を支持する確証型検証 (hypothesis-confirming positive test), (2) 仮説を否定する確証型検証 (hypothesis-refuting positive test), (3) 仮説を支持する反証型検証 (hypothesis-confirming negative test), (4) 仮説を否定する反証型検証 (hypothesis-refuting negative test), (5) 先行情報なし (初期事例「2 4 6」のみ), の5種類が用意された。Kareev & Halberstadt (1993) は、実験1において、上記の5種類の先行情報を呈示した上で、確証事例と反証事例のそれぞれがどの程度有効なテストであるかを被験者に評価させた。先行情報は被験者間要因であり、検証事例の種類は被験者内要因であった。その結果、全体としては確証型検証の評価が高いものの、先行情報として確証型検証が行われた場合、あるいはターゲットに属する正事例の情報が与えられる数が多ければ多いほど、反証型検証に対する有効性の評価が高くなることが示された。また、実験2では、被験者は5種類の先行情報を与えられた上で、確証型仮説

支持，反証型仮説支持，確証型仮説否定，反証型仮説否定の4種類のそれぞれが，規則発見に対してどの程度有効な情報を提供するかを評価するよう求められた．この結果の有効性評価では，先行情報とは無関係に，反証型仮説否定 > 確証型仮説支持 > 確証型仮説否定 > 反証型仮説支持という順序で評価されることが示された．この結果から，Kareev & Halberstadt (1993) は，確証バイアスというものが人間の認知システムに根ざした欠陥ではなく，仮説を否定する結果の情動的価値が正しく認識されうると述べた．

1.4.3 受動的情報収集手続きを含む課題における推論特徴の検討

本研究の第2の目的は，前述した通り，受容方略課題，仮説評価課題という受動的情報収集手続きを含む課題状況下での推論過程の特徴を調べることである．受容方略課題は，日常的な場面での観察学習に相当し，仮説検証の文脈でいうと，被験者が検証事例を自由に選択できない課題であるといえる．一方，仮説評価課題は，被験者は検証事例のみならず仮説も自由に選択できない課題である．これまでの研究において，受容方略課題を仮説検証研究に適用した研究はほとんどなく，また，仮説評価研究では，仮説評価課題で反証型検証の選択率が高くなるという結果が示されているものの，比較の対象となる課題として，規則発見課題については2 - 4 - 6問題，仮説評価課題について日常的因果推論課題が用いられており，両者の間で課題の背景状況が大きく異なるため，一般性を欠くという問題がある．本研究では，同じ背景状況で規則発見，受容方略，仮説評価の各課題を比較するために，まず規則発見課題を実施し，その反

応を元に受容方略課題，仮説評価課題を構成し，それぞれにおける推論過程の特徴を分析する。

1.5 第1章のまとめ

本章では，従来の仮説検証に関する研究を概観し，そこでは仮説検証過程のダイナミズムという重要な側面に関する考察が不十分であることを指摘した。そのため，本研究では，動態的変容分析パラダイムを適用し，いくつかの段階を経て変化するであろう仮説検証過程の特徴を分析するという目的を設定した。また，時系列に沿った段階的变化に加えて，受動的情報獲得手続きという要因が検証方略の選択に与える影響を指摘し，そのような手続きを含む課題として受容方略課題，および仮説評価課題について議論した。以下の章では，これら2つの要因が検証方略の選択に与える影響を検討した実証的研究を報告する。特に次章では，時系列に沿った分析を行うための動態的変容分析パラダイムを仮説検証課題における被験者の反応の分析に適用し，その有効性について議論する。

第2章 仮説検証過程の動態的変容分析

2.1 第1実験 — 規則発見課題における検証方略の段階的変化

本章では、従来の仮説検証研究ではほとんど問題とされてこなかった、検証方略の時系列に沿った段階的変化について検討する。これまでの2-4-6問題をはじめとする規則発見課題を用いた研究では、検証過程の全体を通じた事例選択傾向のみに注目し、全体的な傾向として、被験者が仮説の確証事例を多用する確証型検証方略を用いることを示してきた (Gorman et al., 1987; Klayman & Ha, 1987, 1989; Mynatt et al., 1977, 1978; Tukey, 1986; Tweney et al., 1980; Wason, 1960)。しかし、仮説検証は、仮説生成、検証事例の選択、フィードバックの解釈、仮説の更新が繰り返されるダイナミックなプロセスであり、常に同じ方略が用いられているとは限らず、むしろ時系列に沿ってダイナミックに変化していると考えた方が自然である。そこで、本実験では、検証過程をいくつかの段階に分割して、それぞれの段階における推論の特徴、また時系列に沿った変化に注目した分析を行う。課題としては、従来の研究と同様に、規則発見課題を用いる。

2.2 方法

課題

被験者の課題は、架空の生物が好む物理的環境の条件に関する規則を発見することであった。被験者には、以下のカバーストーリーが与えられた。

あなたは A 国の国立研究所に勤める生物学者です。最近、この国では、2 種類の未知の生物が発見されました。それぞれ、「イエール」「スウ」と名付けられましたが、まだその習性 などについてはよくわかっていません。ところで、「イエール」という生物は、特殊な角(つの)を持っており、その角が戦略的に重要な新世代素材の原料として期待されています。また、この角は一度切りとっても、また生え代わってくるという性質があり、資源の枯渇についてもそれほど心配がないとされています。それに対して、「スウ」の方は、特に害を及ぼすといったことがないものの、これといった利用価値もないと考えられています。

そこで、国をあげて「イエール」を捕獲しようという気運が高まってきました。しかし、前に述べたように、いったいどのようにしたらこの「イエール」をおびき寄せられるのかがわかりません。そこで、あなたに対して、この「イエール」が好む環境を発見して頂きたいという依頼がありました。

規則を構成する次元には、(1) 色光 (赤, 緑, 青のそれぞれについて, 0~2 の 3 段階), (2) 温度 (0~40 度の間で 1℃間隔), (3) 湿度 (40% ~ 80%の間で 1%間隔), (4) 音 (なし, 1000Hz, 2000Hz, 3000Hz, 6000Hz, 7000Hz, 8000Hz の 7 段階), (5) 匂い (甘い, すっぱいの 2 種類のそれぞれについて, 0~2 の 3 段階) の 5 次元を設定した。色光と匂いの強さについては、「0」は「なし」, 「1」は「弱」, 「2」は「強」というラベルが付与されていた。被験者が発見すべき規則は、「『イエール』を捕獲するための条件は、最低でもいずれか 1 つの色光の強度が『2』, かつ甘い (すっぱい) 匂いがある (『1』以上)」というもので

あった。半分の被験者は甘い匂いがあるという規則 (規則 1)、残り半分の被験者はすっぱい匂いがあるという規則 (規則 2) に割り当てられた。フィードバックはイエス・ノーの形ではなく、生物の名前で与えられた (c.f. 2-4-6 問題の DAX, MED バージョン; Tweney et al., 1980)。また、『イエール』が捕まらない時は、常に『スウ』が捕まる、すなわち『スウ』を捕獲するための条件は、『イエール』を捕まえるための条件の補集合になっていた。

被験者は、規則を発見するために、自ら仮説を生成し、それを検証するための事例を自ら作成した。仮説は、別途配布された質問紙に、被験者自身の手によって毎回記入され、検証事例の選択、実際の検証、及び結果のフィードバックは全てコンピューター上で行われ、試行毎に記録された。被験者は望むときに、過去に行った検証時の設定、すなわち検証事例とその結果に関する履歴を参照することができた。仮説を生成し、それに対する検証 1 回をもって 1 試行とし、途中で正しい規則を発見したか否かに関わりなく、全ての被験者が 30 試行を行なった。

手続き

被験者は最初に、カバーストーリー、および規則を構成する次元に関する教示を与えられた。被験者は、30 回の試行の毎回において、まず反应用紙に仮説を書いてから、その仮説に対する検証をコンピューター上で行うという手続きに関する説明を受けた後で、反应用紙の書き方に対する説明を与えられた。反应用紙は、30 枚の仮説記録用紙と、全試行終了後に記入する 1 枚の質問紙か

ら構成されていた。30 枚の仮説記録用紙では、それぞれの試行を実施する時点で考えられる可能性を全て列挙した上で、それぞれに対して、どの程度自信が持てるかを 0 (絶対に間違っていると思う) ~ 100 (絶対に正しいと思う) の間で評価することが求められた。またその試行においてどの可能性を検証するか、すなわちどの仮説に注目するかを 1 つ選択することも併せて指示された。さらに、可能性が思い付かない場合には「わからない」と記入し、決して飛ばさずに進めることが強調された。続いて被験者は、コンピューターの操作に関する説明を受けた後に、操作に慣れるために 6 回の練習試行を実施した。練習試行においては、被験者の設定に関わりなく、前半では、「イエール」という生物が、また後半では「スウ」という生物が捕獲されるようになっており、被験者はその旨通知され、コンピューターの操作の習熟に専念するよう求められた。練習試行終了後、被験者は各自のペースで本試行を行った。30 回の本試行終了後は、最終的な仮説、実験中に気をつけた点をたずねられる質問紙に回答し、報酬を受け取った時点で実験は終了した。質問紙では、「正しい規則を見つけること」「多くのボーナスを獲得すること」「自分が正しいと思う設定をテストすること」「見落としていることがないかどうか確かめること」「得点を得ること」「その他」の中から複数を選択し、特に重要視した項目に印をつけることが求められた。

被験者

被験者は、札幌市在住の大学生、大学院生、社会人 40 名 (年齢は 18 歳~32

歳)であった。実験は個別に行なわれ、1人当たり概ね1時間半程度で終了した。

2.3 結果

正解到達率

40名の被験者のうち、正解に達したのは22名であった。これら22名の被験者は、平均して23.7試行目で正解に達した。また、不正解であった18名の内訳は、正しい規則の下位集合となる規則を報告した者が13名、正しい規則と部分的に重なる規則を報告した者が5名であった。このうち、下位集合型の誤りは、2-4-6問題でよく見られるタイプの誤りであるが、部分的重なり型の誤りは、2-4-6問題においてはほとんど見られなかったタイプの誤りである。これは、2-4-6問題が、数字列という1次元のみで構成される問題であったのに対し、今回使用した問題が5つの次元から構成されていたという理由によると考えられる。

検証事例選択

次に被験者の検証事例の選択傾向を分析する。まず、被験者の選択した事例を、当該の仮説に対する確証事例 (+H)、反証事例 (-H)、その他のいずれかに分類した。また、第1試行のデータは、被験者が仮説を生成していないため、以降の分析の対象からは除外した。その他に該当する事例は、仮説が存在しない、あるいは複数の仮説が立てられている場合に、どれを検証の対象とするかを

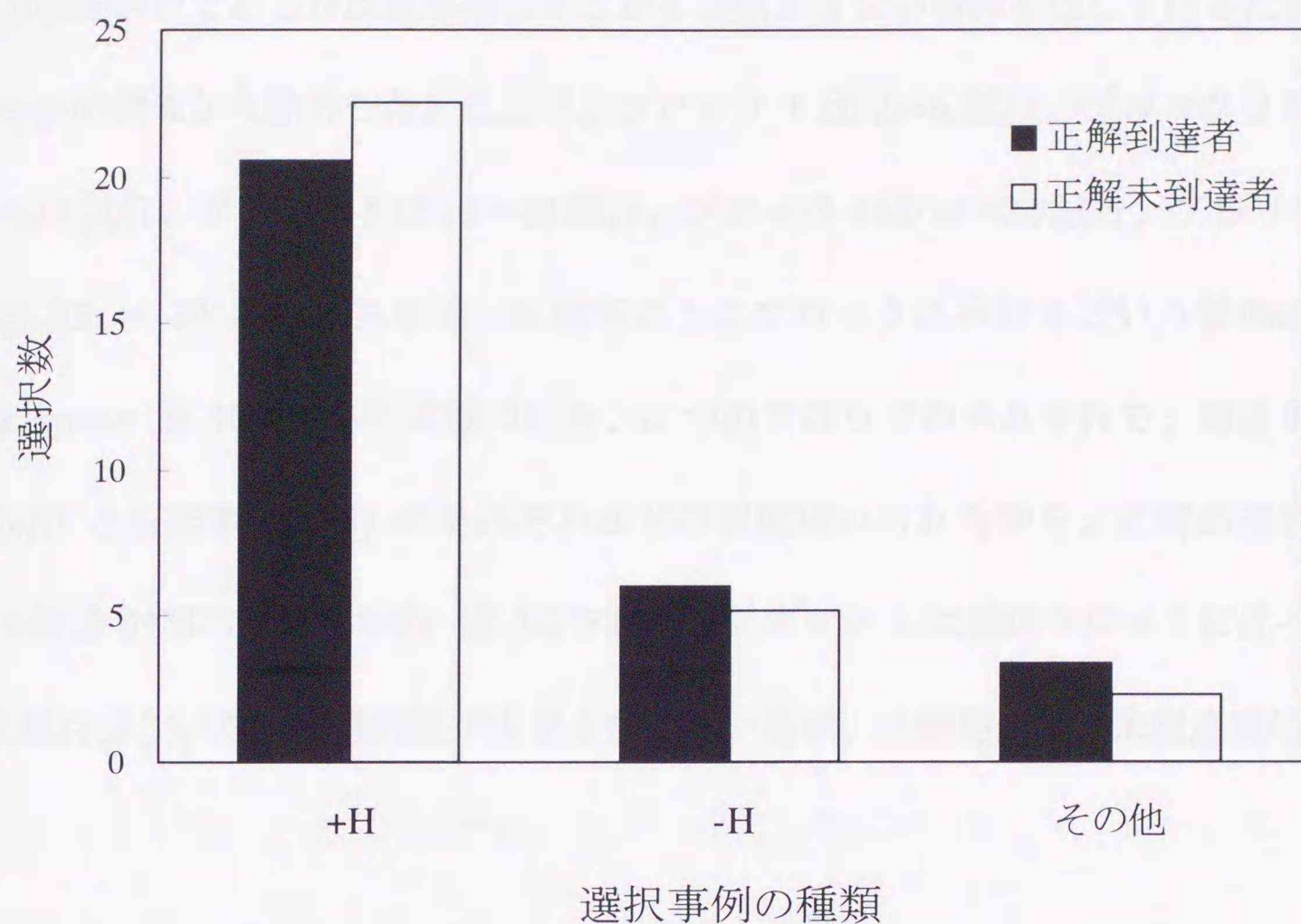


図2-1 選択事例の種類と頻度

選択しなかったため、いずれにも分類し得ない事例を指す。

全体的な傾向としては、確証型検証が多く選択されるという結果が見られた。第1試行を除く全29試行のうち、確証事例は平均して21.5試行で選択されており、反証事例は5.6試行、その他は2.8試行であった。正解到達者と未到達者の間にも差は認められなかった(図2-1)。

次に29試行を、6試行ずつ5つのブロックに分け、それぞれのブロックでの事例選択傾向がどのように変化するかを分析する。ここで「その他」に属する事例は、被験者が仮説を生成していない、あるいは複数の仮説を生成している場合に、どの仮説に注目しているかを指摘し忘れたために、選択事例が仮説

の確証事例であるか反証事例であるかを分類できない事例を指しているため、分析の対象から除外した。ここで、ブロック 1 は 2～6 試行、ブロック 2 は 7～12 試行、ブロック 3 は 13～18 試行、ブロック 4 は 19～24 試行、ブロック 5 は 25～30 試行からなる。6 試行ごとにブロックに分けるという基準は、Klayman & Ha (1989) に基づいた。6 つのブロックのそれぞれで、確証事例 (+H) と反証事例 (-H) のそれぞれがどの程度用いられたかを、正解到達者、未到達者別に示したのが、図 2-2 である。ブロック 1 は他のブロックに比べて 1 試行少ないため、選択数が少なくなっているが、正解到達者、未到達者に共

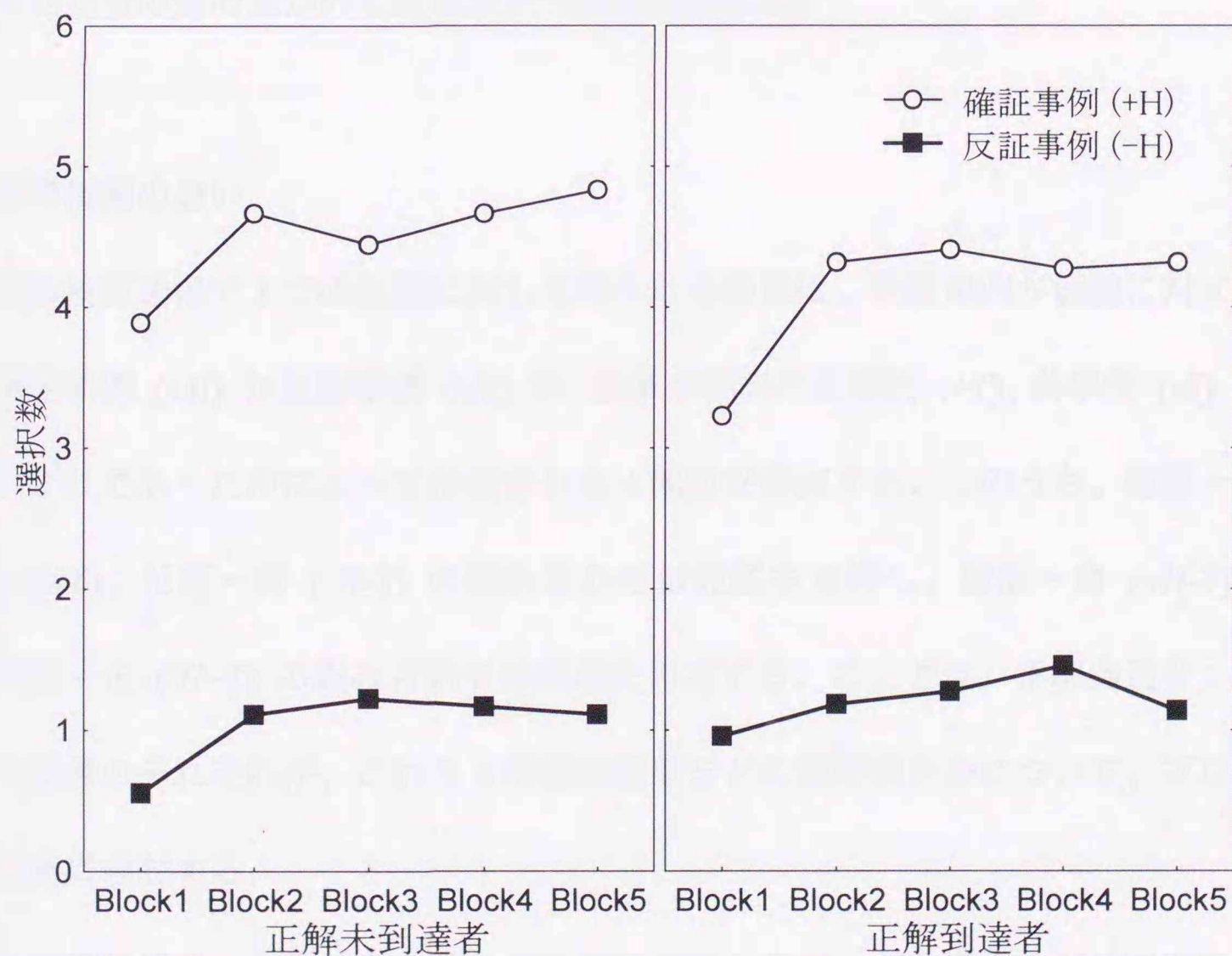


図2-2 ブロック別に見た事例選択傾向

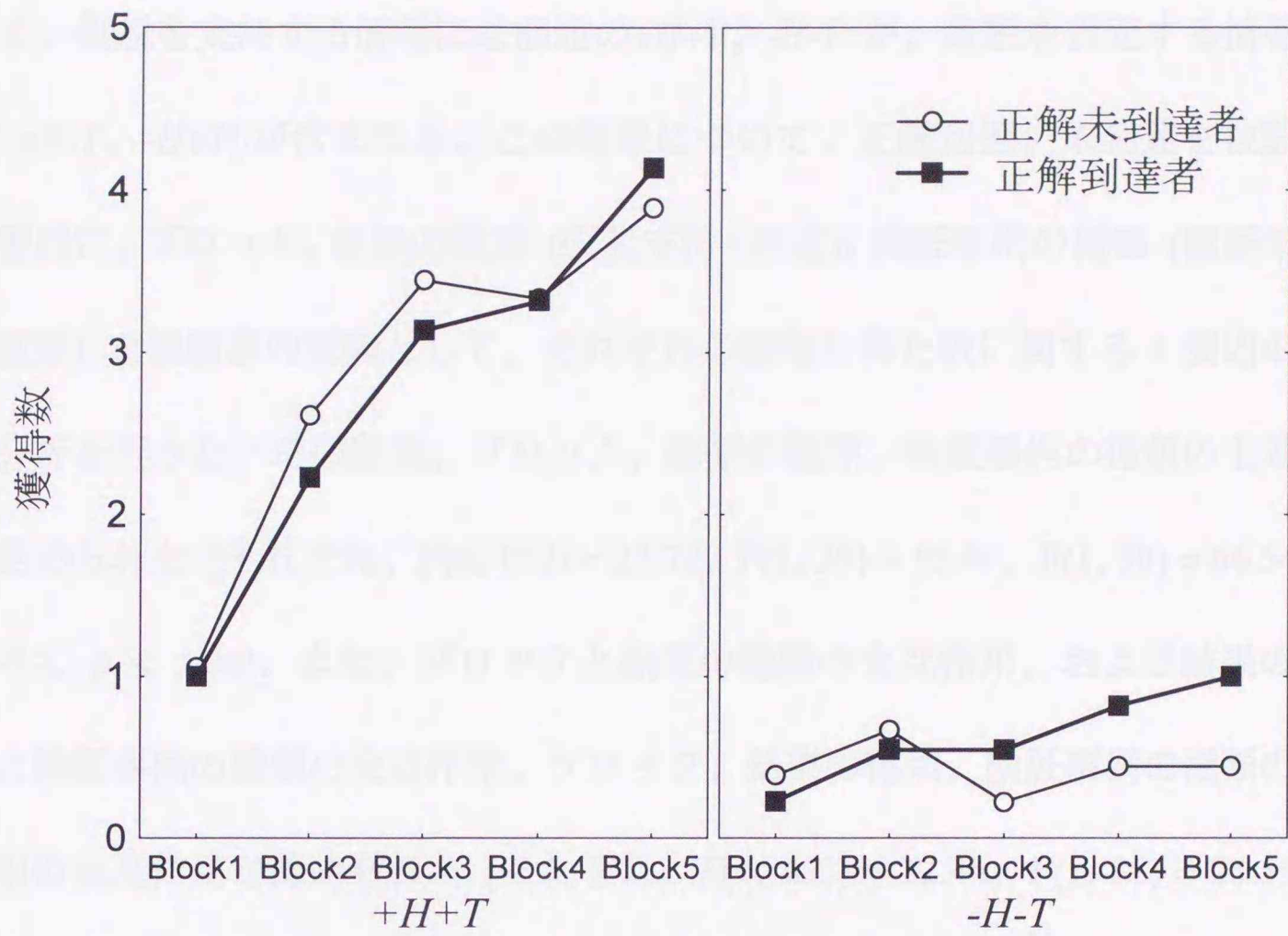
通して全ブロックを通じて確証型検証の選択比率が一貫して高いことがわかる。

正解到達・未到達を被験者間要因、ブロック、選択事例の種類を被験者内要因として、事例選択数に関する分散分析を行った結果、ブロックと事例の種類の主効果が認められた [それぞれ、 $F(4,152) = 23.72$, $p < .001$, $F(1,38) = 66.54$, $p < .001$] が、到達・未到達の主効果、交互作用は認められなかった。ブロックの主効果については、第1ブロックが、その他のブロックに比べて試行数が少ないためであり、事例の種類の効果については、確証事例が反証事例よりも多いことによる。このように、正解到達者・未到達者の間で、統計上の有意差は認められなかった。以下では、確証・反証事例の選択数以外で正解到達者と未到達者の間の差がみられるかどうかを検討する。

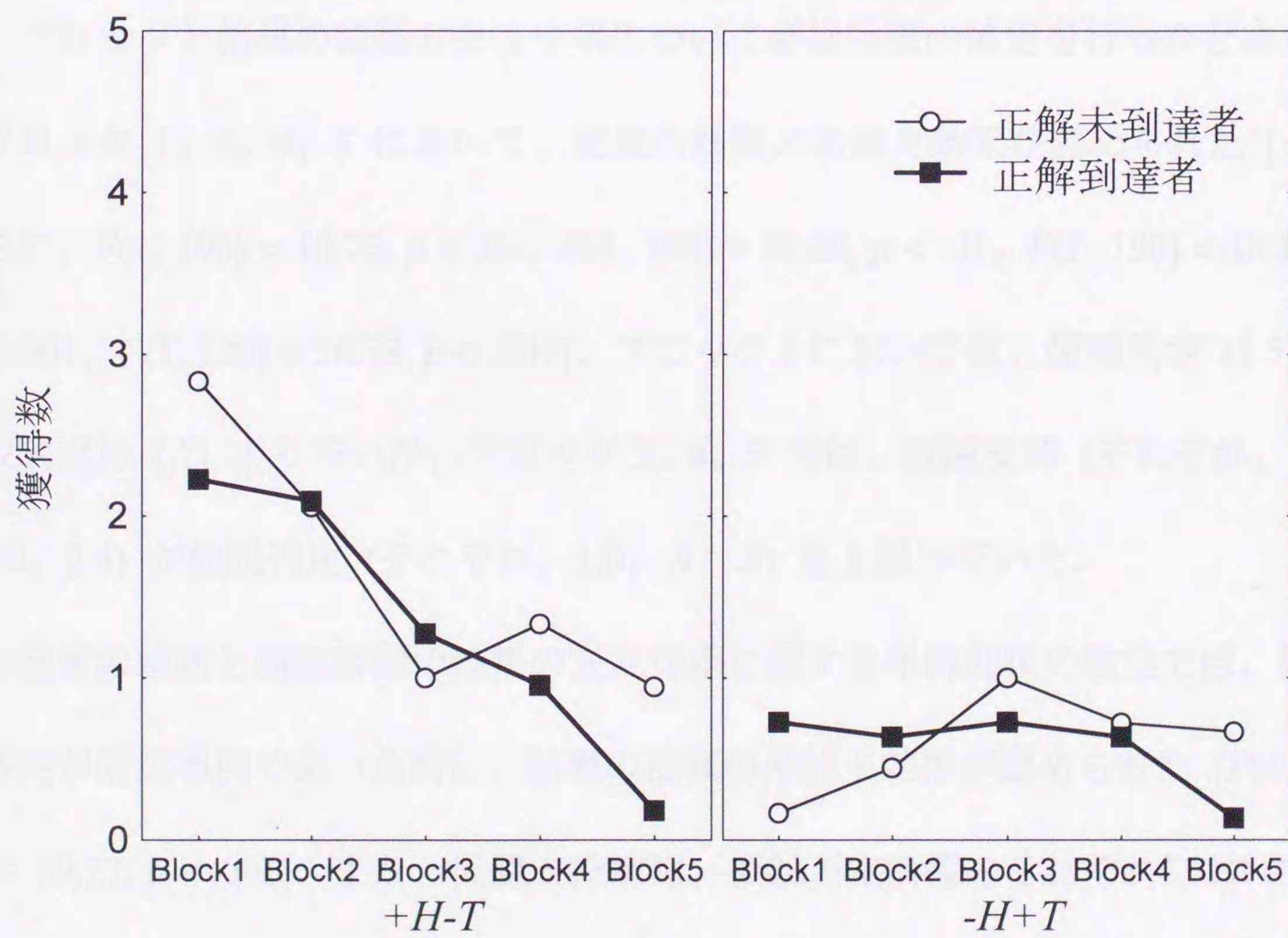
獲得情報の違い

仮説検証過程で1つの仮説に対して得られる情報は、検証事例が仮説に対する確証事例 (+H) か反証事例 (-H) か、結果が規則の正事例 (+T), 負事例 (-T) のいずれであったかによって決定される4種類が存在する。このうち、確証－正 (+H+T), 反証－負 (-H-T) の組み合わせは仮説を支持し、確証－負 (+H-T), 反証－正 (-H+T) の組み合わせは仮説を否定する。ここでは、正解到達者、未到達者のそれぞれが、これら4種類の結果をどの程度得たかについて、ブロック別に検討する。

正解到達者、未到達者のそれぞれが各ブロックで、どの情報をどの程度得たかを、仮説を支持する結果、仮説を否定する結果に分けて示したのが 図 2-3 で



仮説を支持する結果



仮説を否定する結果

図2-3 各ブロックにおける獲得情報数

ある。仮説を支持する情報には前述の $+H+T$ 、 $-H-T$ が、仮説を否定する情報には $+H-T$ 、 $-H+T$ が含まれる。この結果について、正解到達、未到達を被験者間要因に、ブロック、結果の種類 (仮説支持・否定)、検証事例の種類 (確証型・反証型) を被験者内要因として、それぞれの情報を得た数に関する 4 要因の分散分析を行った。その結果、ブロック、結果の種類、検証事例の種類の主効果が認められた [それぞれ、 $F(4, 152) = 23.72$, $F(1, 38) = 15.44$, $F(1, 38) = 66.54$, すべて $p < .001$]。また、ブロックと結果の種類の交互作用、および結果の種類と検証事例の種類の交互作用、ブロック、結果の種類、検証事例の種類の 3 要因の交互作用が認められた [それぞれ、 $F(4, 152) = 22.44$, $F(1, 38) = 25.59$, $F(4, 152) = 21.20$, すべて $p < .001$]。

ブロックと結果の種類の交互作用について単純効果の検定を行ったところ、ブロック 1, 3, 4, 5 において、結果の種類の単純主効果が認められた [それぞれ、 $F(1, 190) = 10.76$, $p < .01$, $F(1, 190) = 11.20$, $p < .01$, $F(1, 190) = 18.19$, $p < .001$, $F(1, 190) = 56.78$, $p < .001$]。ブロック 1 においては、仮説否定 (1.5) が仮説支持 (.7) より多いが、ブロック 3, 4, 5 では、仮説支持 (それぞれ、1.8, 2.0, 2.4) が仮説否定 (それぞれ、1.0, .9, .5) を上回っていた。

結果の種類と検証事例の種類の交互作用に関する単純効果の検定では、検証事例が確証事例であった時に、結果の種類の単純主効果が認められた [$F(1, 76) = 39.27$, $p < .001$]。また、仮説支持情報、仮説否定情報のそれぞれにおいて、検証事例の種類の単純主効果が認められた [それぞれ、 $F(1, 76) = 91.85$, $F(1, 76) = 14.35$, 共に $p < .001$]。検証事例および結果の種類の組み合わせからなる 4 種

類の情報のうち、最も多く得られたのが確証型仮説支持 (+H+T: ブロックあたり 2.8), 次が確証型仮説否定 (+H-T: 1.5)であった。反証型仮説支持 (-H-T: .5), 反証型仮説否定 (-H-T: .6) は、元々反証型検証の数が少ないために、それぞれの頻度も少なくなっている。

また、3 要因の交互作用については、主な結果を挙げると、ブロック 1, 3, 4, 5 において、確証型仮説支持が確証型仮説否定より多く得られたといえる [それぞれ, $F(1, 380) = 19.72, F(1, 380) = 40.53, F(1, 380) = 41.59, F(1, 380) = 103.94$, 全て $p < .001$].

正解到達者、未到達者という比較では、仮説否定情報に関して、ブロック 3 までは、正解到達者が未到達者に比べて、やや多くの仮説否定情報を獲得しているが、ブロック 4 以降では、未到達者の方がより多くの仮説否定情報を獲得しているようにも思われる。これは、本実験課題が正解に到達したかしまなかったかに関わらず、30 試行で実験を終了したため、被験者にとって十分な検証を行う前に課題が終了してしまったためではないかと推測される。しかし、正解到達・未到達という要因を含む交互作用は見られなかった。

単一仮説検証・複数仮説検証

仮説検証においては、複数の仮説を同時に検証する方が、1 つの仮説について考慮するよりも効率的な課題解決が可能となることが知られている (Gorman et al. 1987; Kareev & Halberstadt, 1993; Tweney et al., 1980; Wharton et al., 1993)。

ここでは、前述した 5 つのブロックのそれぞれにおいて、正解到達者、未到達

者が同時に複数の仮説を考慮したか、複数の仮説が存在する時に、どのような仮説を検証したかについて検討する。図2-4に、5つのブロックのそれぞれで、仮説が1つだけ挙げられている場合(単一)、複数の仮説が挙げられているが、検証の対象とされているのがもっとも可能性が高いと考えられる仮説、すなわち、主仮説である場合(複数-主)、複数の仮説が挙げられており、検証の対象が主仮説ではなく、対立仮説である場合(複数-対立)のそれぞれが、どの程度の頻度で見られるかを、正解到達者と未到達者に分けて示した。

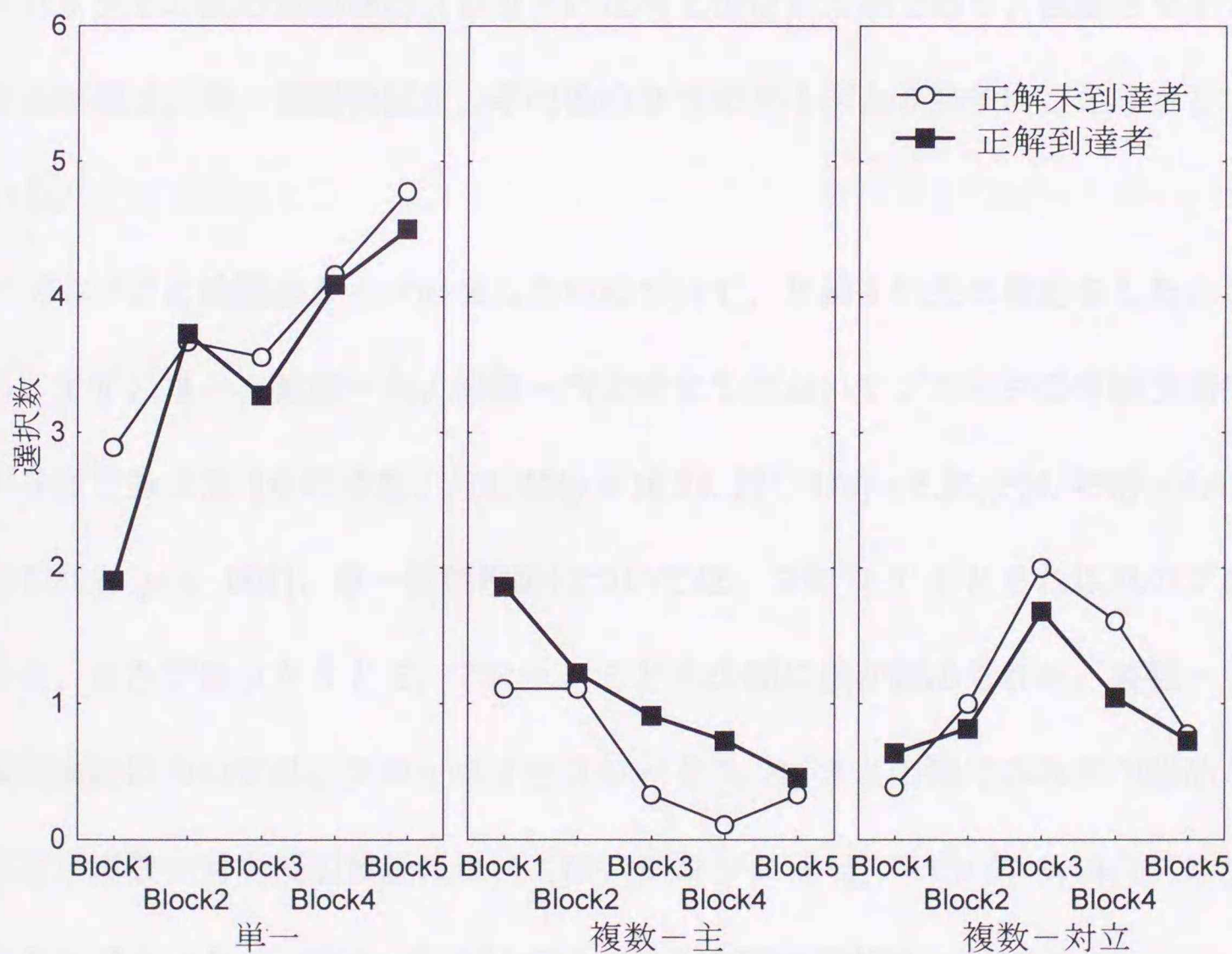


図2-4 検証対象となる仮説の選択

検証の対象が主仮説であるか対立仮説であるかは、仮説に対する確信度の評価に基づいた。すなわち、最も確信度が高い仮説が選択されている場合は主仮説検証とし、それ以外の仮説が選択されている場合は対立仮説検証として分類した。この焦点仮説の選択数に関して、最終的な規則報告（正解到達・未到達）、ブロック、検証のタイプ（単一、複数－主、複数－対立）を独立変数とした分散分析を行った結果、ブロックの主効果、検証のタイプ的主効果、およびブロックと検証のタイプの交互作用が認められた [それぞれ、 $F(4, 152) = 42.43$, $F(2, 76) = 42.34$, $F(8, 304) = 8.65$, いずれも $p < .001$]。ブロックの主効果については、ブロック 1 の試行数が他のブロックに比べて少ないためであり、検証のタイプ的主効果は、単一仮説検証が、その他の 2 つのタイプよりも多いことを示している。

ブロックと検証のタイプの交互作用について、単純主効果の検定をしたところ、まず、単一、複数－主、複数－対立の全てにおいてブロックの単純主効果が有意であった [それぞれ、 $F(4, 456) = 16.54$, $F(4, 456) = 5.55$, $F(4, 456) = 6.42$, いずれも $p < .001$]。単一仮説検証については、ブロック 1 とそれ以外のブロック、またブロック 5 と 2、ブロック 5 と 3 の間に差が認められた。複数－主仮説検証については、ブロック 1 とブロック 3, 4, 5 との間でのみ差が認められた。複数－対立仮説検証については、ブロック 1 とブロック 3, 4、ブロック 3 とブロック 5、ブロック 3 とブロック 2 の間で差が見られた。

また、全てのブロックで 3 タイプの検証の選択数に差が見られた [それぞれ、 $F(2, 380) = 8.21$, 21.90 , 18.10 , 34.51 , 51.17 , 全て $p < .001$]。ブロック 1 にお

いては、単一と複数一対立の間に差が見られた。また、ブロック 2, 4, 5 では、単一とそれ以外の間で差が認められ、ブロック 3 では、全てのタイプの間に差が認められた。

正解到達・未到達という比較では、両者の間で比較的差が見られるのが、ブロック 1 である。ブロック 1 において、正解到達者は平均して 2.5 試行で複数の仮説を考慮しているのに対して、未到達者のそれは 1.6 試行であった。逆に、1 つの仮説のみを挙げている試行は、正解到達者が 2.1 試行であるのに対し、未到達者は 3.2 試行みられた。また、図 2-3 にみられるように、ブロック 1 で仮説否定情報が得られた数については、正解到達者の方が、未到達者に比べて反証型仮説否定、すなわち負事例だと思われたものが正事例であるというタイプの誤りに気づくことが多かったため、正解到達者は比較的早い段階で正解に近い仮説に到達し、誤った規則報告でよく見られた下位集合型のエラーに陥りにくかったのではないかと推測される。しかし、正解到達・未到達という要因を含んだ交互作用は統計的には有意には見られず、あくまで推測の域を出ない。

2.4 考察 — 仮説検証過程の動態的変容プロセス

本実験の主要な目的は、仮説検証方略を時系列に沿って分析し、仮説検証過程がダイナミックに変化するかどうか、変化する場合はどのように変化するかを調べることであった。眞嶋・瀧川 (1997; 1999; Majima & Takigawa, 1999) は、Wason (1960) 以来の仮説検証研究の大きな問題点として、全体的な事例選択

傾向によって被験者の検証方略をスタティックに記述しているため、検証過程がどのように進行するかを分析できなかったことを挙げている。仮説検証過程は、仮説の生成、検証事例の選択、フィードバックの獲得と解釈、仮説の更新などから構成される、極めてダイナミックなプロセスであり、全体的な事例選択傾向のみに基づいた分析では、このプロセスの持つ決定的に重要な側面を見落としてしまうこととなる。そのため、課題における被験者の反応を全体的に分析するのではなく、時系列に沿った段階的变化をとらえることが必要である。本実験は、以上の考え方に基づいて、全試行をいくつかのブロックに分けて、被験者の反応、特に事例選択、獲得情報、仮説の取捨選択方法の3点について注目した。

事例選択傾向については、一貫して仮説に沿った確証事例の選択が多いという結果が得られており、過去の研究とも対応した傾向が見られる (Klayman & Ha, 1987; Mynatt et al., 1977, 1978; Wason, 1960)。この結果からすると、少なくとも本実験課題における被験者の仮説検証方略は、時系列に沿って、ダイナミックに変化したとは言えないようである。しかし、本実験で正解に達した被験者でも、比較的終盤の状況においてようやく正解に達した被験者が多く、22名中16名が20試行以上、さらにその内の9名が25試行以上を要していた。このことから考えて、本実験で設定された30試行という試行数は、被験者が正しい規則を発見するのに必要最低限の試行数に極めて近い値であると言えるであろう。すなわち、本実験では試行数が限られていたために、被験者が十分に検証し尽くすことができないまま終了してしまったことが考えら

れる。科学的探求では、その初期において仮説の蓋然性を高めるために確証型検証を用いることは、直感的にも納得ができ、また合理的な方略であることが指摘されており (e.g. 内井, 1995), その意味では、本実験の被験者も極めて常識的な反応を示していると言える。また、確証型検証が多いものの、Wason (1960) で見られたように、常に確証型検証のみが用いられるというわけではなく、全体の $1/6$ 程度は、反証型検証が用いられることが示された。その意味では、被験者は必ずしも確証型検証のみに頼った、いわゆる確証バイアス傾向をもつのではなく、ある一定の割合で反証型検証を選択する確証・反証併用方略を持っていると考えることもできる。第1章で示したように、2 - 4 - 6 問題では、確証型検証がもっとも多く(66%), ついで診断性検証 (21%) が比較的多く行われ、反証型検証は極めて少ない (7%) ことが確かめられている (Klayman & Ha, 1989)。この結果と比較すると、本実験における反証型検証の選択率はやや高め (17%) であるといえる。

検証の結果得られた情報については、前述の事例選択傾向とも対応して、全体的には確証型仮説支持情報が最も多く得られるという傾向が見られた。また、正解到達者と未到達者の間で、得られた情報の種類およびその数には差が見られなかった。ただし、正解未到達者は、正解到達者に比べると終盤の試行になっても仮説否定情報を得ることがやや多く、また最終的な規則の報告では、下位集合型の誤りを犯したタイプが多かった。このことから考えると正解に到達できなかった被験者は、十分に検証しつくす前に課題が終了してしまったのではないかと思われる。一方、未到達者と獲得した情報の種類や数に差は

ないものの、正解到達者は、得られた情報をうまく利用し、このことが正しい規則を発見するのに貢献したと推測される。特に、仮説からは非ターゲット集合であると予測されるものが実際にはターゲット集合であることが判明するタイプの誤りであることを示す、反証型仮説否定情報をうまく利用することで、下位集合型の誤りに陥ることがなかったのではないかと考えられる。

仮説の取捨選択方法については、全体的に単一仮説検証が多く、被験者が同時に複数の仮説を立てることはそれほど多くはないといえる。また、単一仮説検証の数は、試行を経る毎に増えており、1つの仮説へと絞り込まれていくプロセスが観察されたと言える。さらに、複数一対立仮説検証については、ブロック3, 4でやや数が多くなっており、中盤において可能性が低そうな仮説が注目されていることを示している。これは、可能性が低そうな仮説の棄却を目的とした消去法的な検証過程の表れだと解釈することができるであろう。しかし、正解到達者と未到達者の間では、仮説の取捨選択方法に差は見られなかったことから、両者の間で解決成績の差を生じたのは、仮説の取捨選択方法というよりは、前述したような獲得した情報の利用法の違いではないかと推測される。

2.5 第2章のまとめ

本章の主要な目的は、仮説検証課題で得られた被験者の反応を、時系列に沿っていくつかの段階に分けて吟味する動態的変容分析パラダイムによって分析し、そのダイナミックな変化をとらえることにあった。本実験で採用した動態的変

容分析パラダイムによって、これまでの研究では捉えきれなかった仮説検証方略の時間を追った変化をある程度は示すことができたといっただいであろう。検証事例の選択については、全体的には確証型検証の使用が多く、反証型検証の使用は比較的少ないといえる。また両者の選択比率も課題を通してほぼ一定であり、さらに正解到達者と未到達者の間でも差は見られなかった。しかし、仮説の取捨選択方法という観点から見ると、より終盤に近づくに従って1つの仮説のみを考慮する試行が増え、それに伴って複数の仮説を考慮する試行が減っていく過程が観察された。また、複数の仮説を立てた場合でも、比較的初期の段階では、その中でももっとも可能性が高そうな主仮説に注目するが、中盤では可能性が低い対立仮説に注目した消去法的方略を用いている可能性が示唆された。この意味で、一部限定はつくものの、動態的変容分析パラダイムによる被験者の反応の分析はある程度の有効性を持っていると考えられる。次章では、この分析手法を、受容方略課題および仮説評価課題における被験者の反応の分析に適用し、これらの課題が持つ受動的情報獲得手続きが、被験者の推論過程にどのように影響を与えるか、またその影響は時系列に沿ってどのように変化するかを検討する。

第3章 受動的情報収集手続きを含む課題における 方略の変化

本章では、第2章の第1実験で用いられた動態的変容分析パラダイムを他の課題状況へと適用し、時系列に沿った段階的な検証方略の変化について更なる検討を加える。また、本研究の第2の目的である、受動的な情報獲得手続きが推論過程に与える影響を検討するために、第1実験で用いられた課題を元にして、被験者に呈示する情報を作成する。

3.1 仮説検証における受動的情報収集

仮説検証研究に用いられる規則発見課題は、科学的探求過程をシミュレートしたものであり、(1) 仮説の生成、(2) 検証事例の選択、(3) フィードバックの獲得と解釈、(4) 必要に応じた仮説の更新という手続きによって構成されている。このうち、結果のフィードバックは常に実験者など外部の情報源から与えられる情報であるが、仮説生成および検証事例の選択は、検証者自らが能動的に行なうのが普通である。このときの事例選択の傾向として確証バイアス、あるいは確証型検証方略が観察されることが知られている (Klayman & Ha, 1987, 1989; Mynatt et al., 1977, 1978; Wason, 1960)。

一方、検証者が検証事例や仮説を選択しない課題として、受容方略課題と仮説評価課題が知られている。かつて Bruner, Goodnow, & Austin (1956) によって始められた受容方略課題の研究は、その後ほとんど行われていないが、こ

の研究は、概念達成において被験者が事例情報を自分で選択しない、もしくは選択できない、他者から情報が与えられる観察学習状況を対象とした研究であるといえる。受容方略と類似の状況と言える仮説評価課題は、他者が行なった仮説検証過程を観察・評価する課題である。仮説評価に関する研究としては、物語中の登場人物が行なう仮説検証過程を評価させる課題を用いた Tschirgi (1980) がよく知られている。

これらの受容方略課題と仮説評価課題は、情報獲得における受動性という特徴を共有している。受容方略課題は、検証事例は自由に選択できないものの、仮説は被験者自身が生成する課題であるのに対し、仮説評価課題は他者の検証過程の観察・評価が求められ、仮説さえもが情報として外部から与えられる課題である。そのため、仮説評価課題では、仮説生成に認知的資源を分配する必要がなく、事例選択や結果の解釈により多くの注意を振り向けることができるとされている (Farris & Revlin, 1989; Tschirgi, 1980)。もちろん、仮説評価課題において、被験者が全く仮説を立てないというわけではなく、評価を行なっている最中にも、被験者が別個の仮説を生成し、検証している可能性がある。この場合、仮説評価課題は、外部から対立仮説が与えられた課題であり、被験者は自らが保持している仮説と呈示された仮説という2つの競合する仮説に対して検証を行なうことになる。一般に仮説検証課題では、複数の仮説を保持している方がより正しい規則を発見しやすいことが知られており (Gorman et al., 1987; Tweney et al., 1980; Wharton et al., 1993)、その意味で仮説評価課題は受容方略課題や通常の仮説検証課題と比べても、より正しい推論が可能となる場合

が多いと考えられる。

このように、これら 2 つの課題については、仮説検証における確証バイアスの解消手続きとも関係した、共通の構造的特徴を有していると言えるが、これまで 2 つの課題を直接的に比較検討した研究はなされていない。本章では、これら 2 つの課題で同一の事例系列を用いることによって、両課題が共有する受動的情報収集という構造的特徴が検証事例の選択に対して与える効果について検討する。また、第 1 実験において用いられた動態的変容分析パラダイムを適用し、2 つの課題状況下でどのように推論が進行するかを調べる。

3.2 第 2 実験 — 受容方略型課題における仮説検証過程

受動的情報収集手続きが仮説検証過程に与える影響を調べるため、まず本実験では、受容方略型課題を用いて被験者の推論過程を分析する。Bruner et al. (1956) による、受容方略および選択方略課題という課題手続きの分類は、仮説検証過程の研究においても極めて有効だと考えられる (c.f. Kareev & Halberstadt, 1993; Halberstadt & Kareev, 1995) が、近年、両手続き下における仮説検証過程に注目した研究は見当たらない。また、Bruner et al. (1956) が用いた課題には、事例空間の小ささという重大な問題がある。彼らの課題における事例は、4 次元が各 3 属性を持つという極めて狭い空間に限定されており、日常的に人が遭遇するであろう、より複雑な状況との対応を著しく欠いているといえる。以上の観点から、本実験では、概念を構成する属性空間が広く、概念獲得に必要な事例の数が極めて多い課題を構成するため、第 1 実験 (第 2 章) の規則発見課題

で仮説検証を行なった別の被験者のデータを基にして、呈示する事例系列を作成した。

3.2.1 方法

課題と事例情報

課題は、第1実験で使用された架空の生物が好む環境を発見する課題に従事した被験者による事例選択系列を呈示する、という受容方略型仮説検証課題を採用した。本実験で用意された事例情報は、第1実験での被験者のうち、正解に到達した者と達しなかった者各1名から得られたデータを基に構成された。第1実験で、各被験者は全部で30試行の課題に従事したが、この正解到達者は17、18試行目で初めて連続して正しい規則を唯一の可能性として挙げ、後続の試行で仮説を変更していなかったため、18試行目までを呈示情報として採用した。この正解到達者の試行系列は、15試行の正のフィードバック（正事例）、3試行の負のフィードバック（負事例）から構成されており、以降この事象系列情報を与えられる条件を正事例系列（positive instance series: *P series*）条件と呼ぶ。それに対して、正解に達しなかった被験者の事象系列は30試行中28試行目までが負事例であるが、試行回数を揃えるために、正解未到達者のデータも18試行目までを呈示することとした。以降この事象系列情報を与えられる条件を負事例系列（negative instance series: *N series*）条件と呼ぶ。

受容方略型課題においては、実際に利用可能な検証事例は実験者からの呈示情報であり、呈示される情報が被験者の仮説に対する確証事例となるか反証事

例となるかは被験者が生成した仮説に依存する。本実験では、この実際に呈示される事例情報とは別に、被験者自身が検証事例を自由に選択できると仮定した場合にどのような選択を希望するかを尋ねた。これにより、受容方略型課題における事例選択傾向を検討し、通常の規則発見課題における事例選択傾向との比較を行う。また、第1実験において採用された動態的変容分析パラダイムを、結果の分析に適用する。

手続き

被験者は回答用の冊子を渡され、表紙に記載される注意を読み、名前を記入した後、課題のカバーストーリーを読んだ。カバーストーリーは、概ね、第1実験において呈示されたものと同一であったが、受容方略型課題手続きが不自然と感じられることがないように、「大きな研究プロジェクトに参加しており、実験は他の研究機関が行う」という趣旨の記述を追加した。各試行において、被験者は、まず第一に、それまでに得られた架空実験の結果をもとに、仮説を生成するよう求められた。その時、複数の仮説が考えられる場合は全て列挙すること、また、その試行において特に注目する仮説、すなわち検証の対象とする仮説を1つ選択することを求めた。また、生成した仮説の全てについて、どの程度正しいと思うかを0～100点で評価するよう求めた。第二に、「もし上記の仮説について自分で自由に検証事例を選択できるとしたら、どのようなテストをするべきだと思うか」を尋ねた。ただし、「この選択は、実際に行なわれる保証は全くなく、あくま

で希望として回答するよう」求めた。これは、質問紙上のチェックリストに記入するようになっていた。第三に、予め用意された実験条件の組み合わせとその結果の情報が1組呈示された。また、被験者には、後にこの情報を参照できるように、メモをとるよう求めた。第四に、呈示された事例が、初めに記入した仮説の真偽を確かめるために適切なものであったかどうかを、5件法（-2: 全く不適切 ～ 2: 非常に適切）で評価するよう求めた。被験者は以上の手続きを、用意された18試行分繰り返し、その後に正しいと思う規則を1つ報告した。また、最初に呈示される、元データの第1試行に当たる情報は、初期情報として与えられるので反応の対象とならないことが伝えられた。

被験者

被験者は、北海道大学の学部、大学院学生20名（男子8名、女子12名）で、年齢は19～25歳であった。被験者は、無作為に10名ずつ前述の正事例系列条件と負事例系列条件に割り当てられた。実験は個別、あるいは数名の集団で実施され、40～60分程度で終了した。

3.2.2 結果

正解到達率

まず、最終的に報告された規則が正しいものであったかどうかにより被験者を二分した。正事例系列条件では、10名の被験者のうち、正しい規則を発見で

きたものは3名、負事例系列条件では0名であった。第1実験においては、全体としては40名中22名が正解しており、本実験における正解到達率は極めて低いといえる。ただし、本実験で負事例系列条件に与えられた情報は、正しい規則を発見するには必要十分ではないため、この結果は当然であるといえる。また、正事例系列条件では、最終的に報告された規則と呈示された事例情報との間の関係を分析したところ、正解に到達しなかった7名は全て少なくとも1つ以上の試行で、報告した規則と論理的に矛盾する結果を見落したり、フィードバックの誤認をしていた。明白にフィードバックの誤認が確認された被験者は1名のみであり、残り6名は自らの仮説と矛盾する情報を見落としたか無視したと考えられる。

選択希望事例

反応の対象とならなかった第1試行を除く全17試行において、被験者が希望した事例を、仮説に沿う確証事例 (+H) か、沿わない反証事例 (-H) か、その他の事例かに分類した。仮説が複数存在する場合は、被験者がその試行において注目すると報告した仮説と事例との関係から分類した。正事例系列条件では正解到達者と未到達者が存在したが、正解到達者の数が少ないため、以後の分析では正解到達者と未到達者の区別は行なわないこととした。また、「その他」に該当する事例の数も極めて少ないため、以降の分析の対象から除外した。本実験における選択は実際に検証に使用されたものではなく、あくまで自分が自由に検証できると仮定した場合にどのような選択を希望するか、という意味で

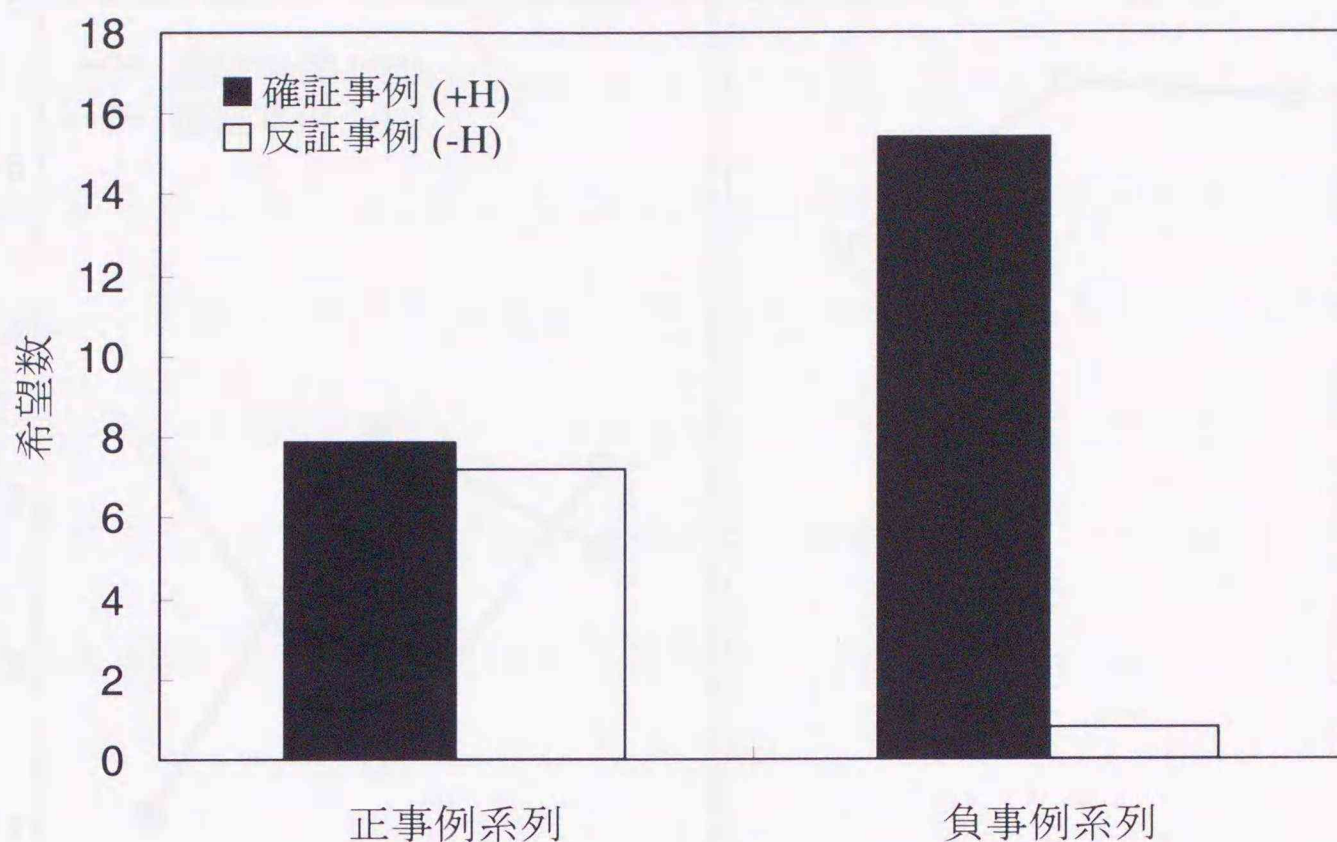


図3-1 事例選択の頻度 (受容方略課題)

の選択であるが、正事例系列条件と負事例系列条件では選択のパターンが大きく異なっていた (図 3-1)。正事例系列条件においては、全 17 試行中確証事例 (7.9 試行) と反証事例 (7.2) がほぼ同程度選択されたが、負事例系列条件においては、確証事例 (15.4) が反証事例 (0.8) よりも遥かに多く選択された。

次に第 1 実験と同様に、この事例選択傾向が時系列に沿ってどのように変化したかを分析する。ここで、第 1 試行を除いた全 17 試行を試行順に 3 つのブロックに分け、それぞれのブロックで確証事例および反証事例の使用が希望された頻度を示したのが 図 3-2 である。この図で、正事例系列条件においては、第 1 ブロック (2~6 試行) では確証事例の選択が多く (5 試行中 3.3 試行)、反証

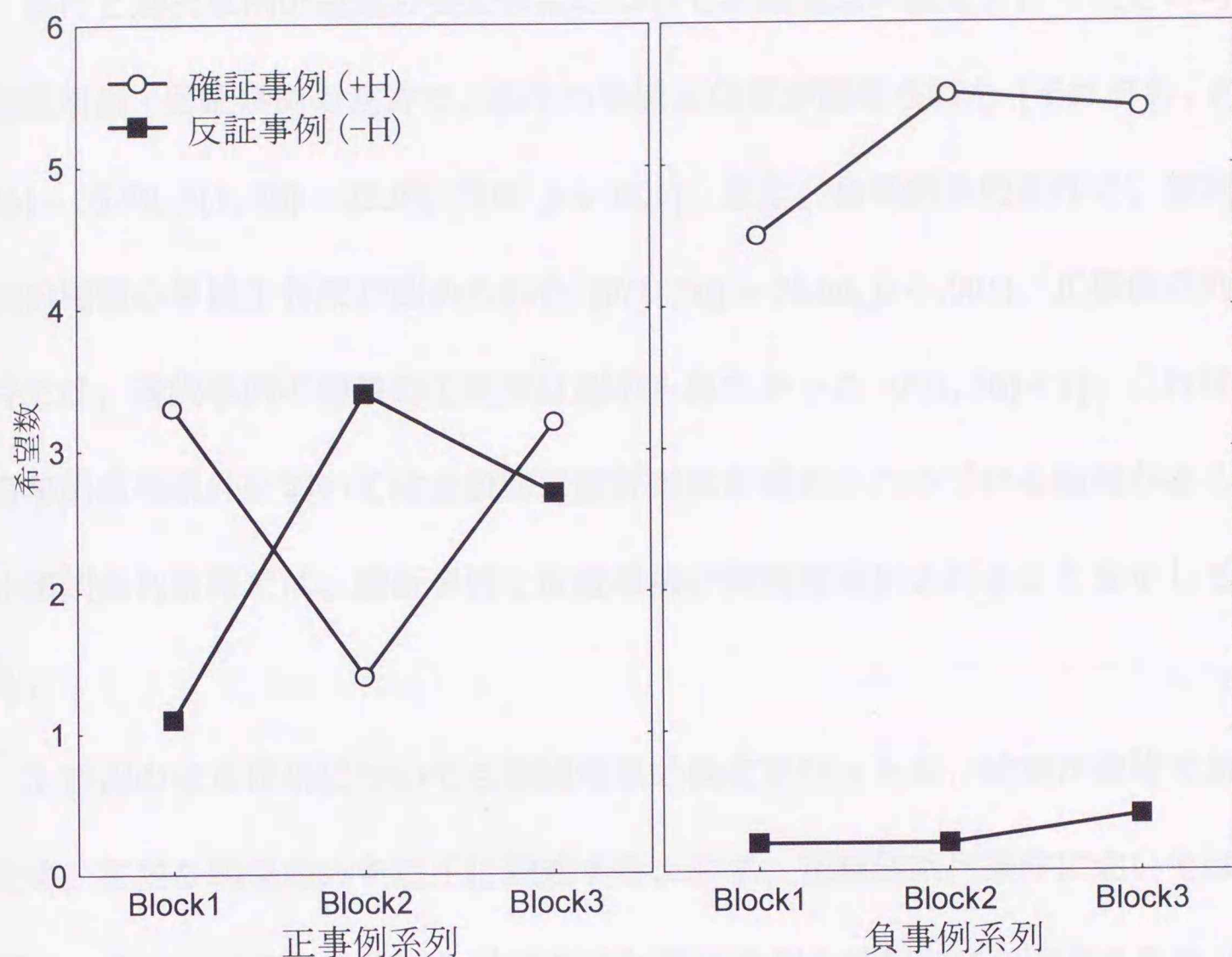


図3-2 選択希望事例の変遷 (受容方略型課題)

事例の選択が少ない (1.1 試行) のに対し、第 2 ブロック (7~12 試行) では、両者の比率が逆転している。一方、負事例系列条件においては、一貫して確認事例の選択頻度が高くなっている。ここで、事例の選択数について、条件、ブロック、選択事例の種類を独立変数とした分散分析を行ったところ、ブロックおよび選択事例の種類の主効果 [それぞれ $F(2, 36) = 15.34$, $F(1, 18) = 42.10$, 共に $p < .001$]、条件と選択事例の種類の交互作用 [$F(1, 18) = 34.74$, $p < .001$]、3 要因の交互作用が認められた [$F(2, 36) = 6.24$, $p < .01$]。

条件と選択事例の種類の交互作用について単純効果の検定を行ったところ、
確証事例・反証事例の双方で、条件の単純主効果が認められた [それぞれ、 $F(1, 36) = 35.59$, $F(1, 36) = 25.92$, 共に $p < .001$]。また、負事例系列条件で、選択事例の種類の単純主効果が認められた [$F(1, 36) = 76.66$, $p < .001$]。正事例系列条件では、選択事例の種類の主効果は認められなかった [$F(1, 36) < 1$]。これは、負事例系列条件においては全般的に確証事例が選択されつづける傾向があるが、正事例系列条件では、確証事例と反証事例が同程度選択されることを示している。

3 要因の交互作用についても単純効果の検定を行ったが、結果が複雑であるため、主要な結果のみを以下に記述する。まず、正事例系列条件においては、ブロック 1 とブロック 2 で、確証事例と反証事例の選択数に差が見られた [それぞれ、 $F(1, 54) = 7.21$, $p < .01$, $F(1, 54) = 5.96$, $p < .05$]。また、負事例系列条件においては、ブロック 1~3 のすべてにおいて、両事例の選択数に差が認められた [それぞれ、 $F(1, 54) = 27.56$, $F(1, 54) = 41.87$, $F(1, 54) = 37.26$, すべて $p < .001$]。正事例系列条件に関しては、確証事例と反証事例の総選択数に差はないが、ブロック 1 では、確証事例が多く (3.3), 反証事例が少ない (1.1) のに対し、ブロック 2 では反証事例が多く (3.4), 確証事例が少ない (1.4) というように、ブロック毎に傾向が変化している。一方の負事例系列条件では、ブロック 1~3 のすべてにおいて、一貫して確証事例が多く選択されているといえる (4.5, 5.5, 5.4)。両条件において、このような選択傾向の違いが見られた原因について、以下で検討する。

反証型検証の選択に影響を与える要因 (1) - 呈示情報の違い

なぜ正事例系列条件と負事例系列条件で、事例選択傾向がこのように大きく異なるのだろうか。これまでの仮説検証研究で得られた知見 (e.g. Klayman & Ha, 1989; Mynatt et al. 1977, 1978; Wason, 1960) や、第2章の第1実験で用いられた通常の規則発見課題でも全試行を通じて確証事例による検証の頻度が高いという結果から考えて、被験者が元々反証事例を選択しやすい傾向があったとは考えにくく、受容方略型課題の手続き自体、あるいは本実験で呈示された情報の性質によるものであるという可能性が考えられる。特に、正事例系列条件では、序盤に反証型検証が行なわれた結果、仮説が否定されることで、被験者が反証型検証の有効性を認識し、後の試行で反証型検証を選択される頻度が高くなったと推測される。本実験における事例選択は、被験者があくまで希望として挙げたものであり、実際に検証に用いられたのは、実験者が呈示した事例である。ここで、各条件において、実験者によって実際に呈示された事例が被験者の立てた仮説に対する確証、あるいは反証事例であったかを、ブロック別に示したのが図3-3である。この結果について、条件、ブロック、事例の種類を独立変数として、それぞれの出現数に関する分析を行ったところ、ブロックの主効果、条件と事例の種類の交互作用が認められた [それぞれ、 $F(2, 36) = 12.13, p < .001$, $F(1, 18) = 10.20, p < .01$]。このうち、ブロックの主効果については、ブロック1に該当する試行数が少ないためである。

条件と事例の種類の交互作用について、単純効果の検定を行ったところ、確

証事例，反証事例のそれぞれに関して，条件の差が認められた [それぞれ， $F(1, 36) = 7.03, p < .05$ ， $F(1, 36) = 11.92, p < .01$]。また，負事例系列条件に関して，事例の種類の差が認められた [$F(1, 18) = 12.03, p < .01$]。このことから，正事例系列条件では，確証事例 (8.6) と反証事例 (6.3) の頻度に差がない，すなわち両者が比較的バランス良く呈示されたが，負事例系列条件では，反証事例 (11.9) の方が確証事例 (4.3) より多かったということがわかる。

さらに，呈示された事例が結果として仮説を支持したか否定したかに関して分類を行うと，表 3-1 のようになる。ただし，負事例系列条件では得られる

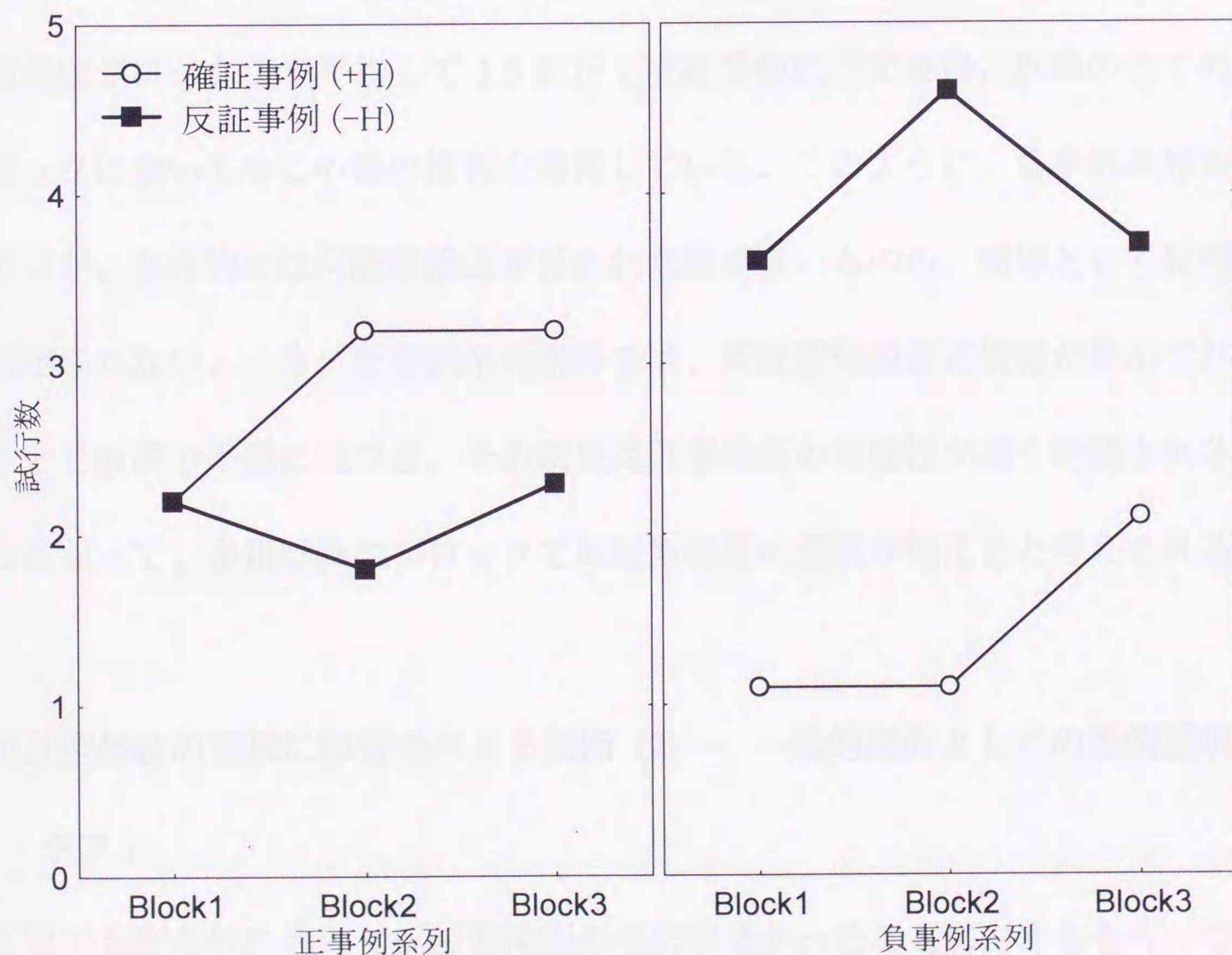


図3-3 呈示事例と仮説との関係

表3-1 獲得した結果の分布

		正事例系列 (<i>P series</i>)			負事例系列 (<i>N series</i>)		
		Block1	Block2	Block3	Block1	Block2	Block3
仮説支持	$+H+T$	2	2.8	2.8	—	—	—
	$-H-T$	0.7	0.4	0.4	3.5	4.6	3.7
仮説否定	$+H-T$	0.2	0.4	0.5	1.1	1.1	2.1
	$-H+T$	1.5	1.4	1.8	—	—	—

結果が全て負事例であるため、確証型仮説支持 ($+H+T$)、および反証型仮説否定 ($-H+T$) はありえない。表に示される通り、正事例系列条件においては、被験者はブロック 1 で平均して 1.5 試行で反証型仮説否定を得、以降の全てのブロックにおいてもこの種の情報を獲得している。このように、負事例系列条件の方が、全体的には反証型検証が行われた数が多いものの、結果として仮説は否定されない。一方、正事例系列条件では、反証型仮説否定情報が呈示されることで仮説の不備に気づき、その結果反証型検証の有効性が高く評価されるようになって、中盤以降のブロックで反証型検証の選択が増えたと考えられる。

反証型検証の選択に影響を与える要因 (2) — 一般的傾向としての事例選択の

シフト

正事例系列条件において反証型検証の選択が多かったことに対するもう一つの説明として、仮説の構築 → 仮説の洗練という検証段階のシフトに伴って、事

例選択傾向が変化したというものが挙げられる。この方略は、まず仮説に対する確信度がある一定の水準以上になるまで、仮説を支持する可能性の高い検証を行い、その後に仮説に不備が存在しないかどうかを確かめるために仮説を否定する可能性のある検証にも注目するというものである。この方略が採用された時の事例選択傾向として、特に中盤から終盤にかけて反証型検証の選択比率が増加するという、事例選択のシフトが見られることが予測される。これは直感的にも十分納得できるものであり、また適応的な意味で合理的な方略といえる (c.f. Friedrich, 1993)。正事例系列条件においては、確証型仮説支持情報を得る場合が多く、仮説に対する確信度は比較的上昇しやすいと考えられる。したがって、被験者は見落としを確認するための事例として反証事例を希望することが多かったと考えられる。一方、負事例系列条件では、与えられる事例が確証事例である場合には仮説が否定され、反証事例のときのみ仮説が間接的に支持されるに過ぎない。そのため、仮説に対する確信度が高まらず、確証事例による仮説支持情報を求めた結果、確証事例の選択頻度が高くなったとも考えられる。

ここで、被験者が各試行において最も可能性が高いと考えている仮説に対する確信度が、ブロックを通じてどのように変化したかを見てみる (図 3-4)。図に示される通り、正事例系列条件では、ブロック 1 から 2 へと確信度が上昇し、ブロック 3 では頭打ち状態になっているのに対し、負事例系列条件では、仮説に対する確信度がブロック 2 で一時低下し、その後やや上昇している。条件とブロックを独立変数として分散分析を行った結果、ブロックの主効果、および

2 要因の交互作用が認められた [それぞれ, $F(2, 36) = 7.24$, $F(2, 36) = 5.57$, 共に $p < .01$]. ブロックの主効果については, ブロック 1 とブロック 3 の間に差が認められた. また, 交互作用について単純効果の検定を行ったところ, 正事例系列条件におけるブロックの差が認められ [$F(2, 36) = 10.51$, $p < .001$], 特にブロック 1 と 2, 1 と 3 の間に差が認められた. このことから, 少なくとも正事例系列条件に関しては, 仮説に対する確信度がブロック 1 から 2 で上昇し, その後は頭打ちになったといえる. また, 負事例系列条件におけるブロックの

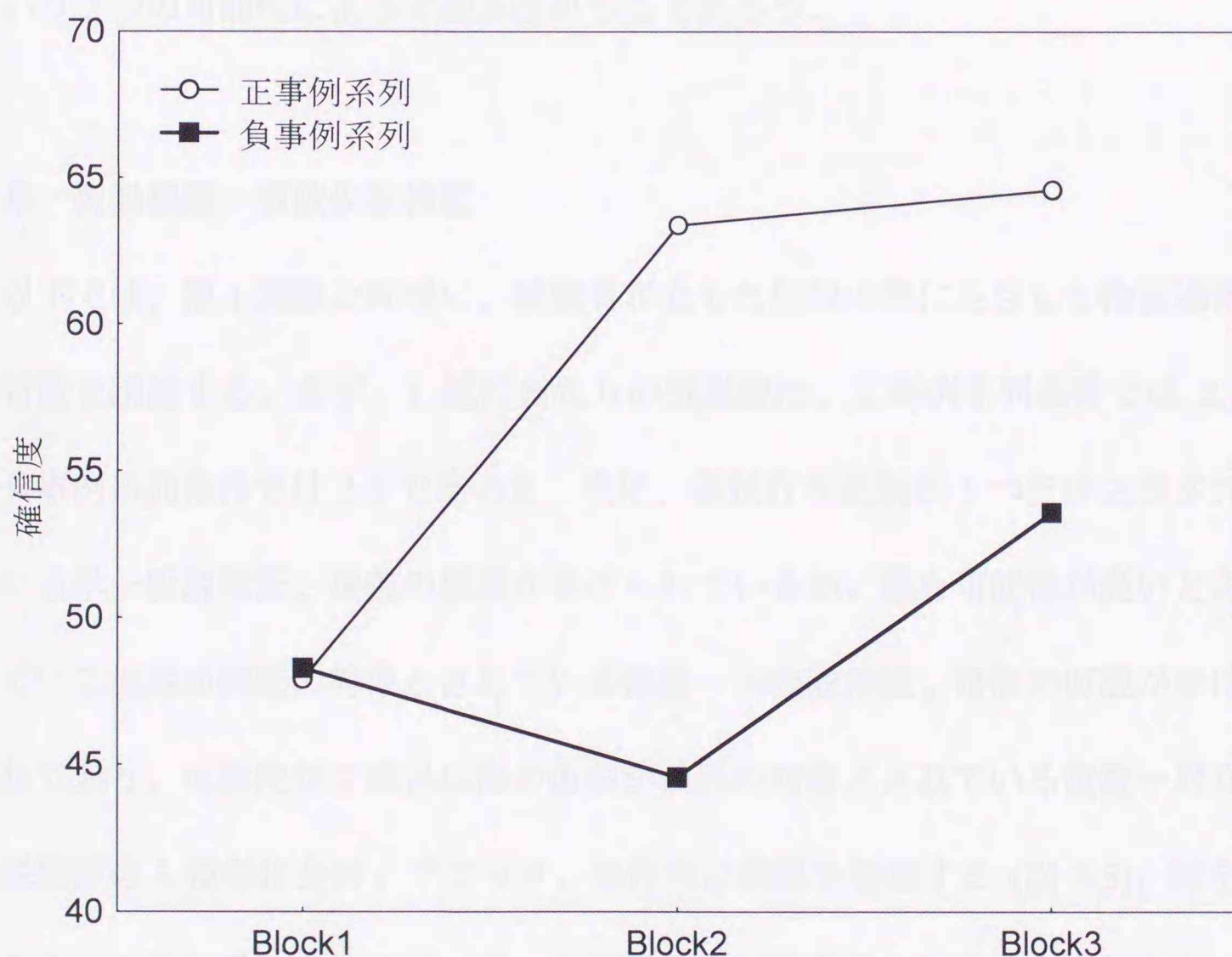


図3-4 仮説に対する確信度の変化

差は有意となる傾向にあった [$F(2, 36) = 2.46, p = .10$]. 正事例系列条件において確信度の上昇が見られるが、負事例系列条件では確信度の上昇が見られず、この確信度評定の結果は、まず仮説の蓋然性を高めてから、より洗練されたものに修正しようとするという検証段階の移行に伴った事例選択傾向のシフトという説明を支持していると考えられる。

以上に示した通り、本実験で見られた、特に正事例系列条件における反証型検証の選択比率の増大は、(1) 反証型仮説否定情報を獲得することによる反証型検証の有効性評価の上昇、(2) 一般的な傾向としての検証段階のシフトという2つの可能性によって説明されうるであろう。

単一仮説検証・複数仮説検証

以下では、第1実験と同様に、被験者が立てた仮説の数に注目して検証過程の特徴を議論する。まず、1試行当たりの仮説数は、正事例系列条件では2.1、負事例系列条件では3.1であった。次に、各試行を仮説が1つだけ生成されている単一仮説検証、複数の仮説が挙げられているが、最も可能性が高いとされている仮説が検証の対象とされている複数－主仮説検証、複数の仮説が挙げられており、可能性が2番目以降の仮説が検証の対象とされている複数－対立仮説検証の3種類に分け、ブロック、条件別に結果を整理する(図3-5)。図をみると、正事例系列条件では、単一仮説試行の頻度が徐々に高くなっており、被験者が、複数の候補から徐々に一つの候補へと絞り込みを行っている過程の現れであると考えられる。一方、負事例系列条件では、一貫して複数仮説検証が

多く、また試行が進むにつれて、特に複数－対立仮説検証の数が増えており、正事例系列条件のように絞込みができなかったことを示していると考えられる。これは、負事例系列条件では負事例のみが与えられるため、規則の発見に困難をきたしていることの現れであると思われる。ここで、条件、ブロック、3つの検証のタイプ(単一、複数－主仮説、複数－対立仮説)を独立変数として、それぞれの選択試行数に関する分析を行った結果、ブロックの主効果 [$F(2, 36) = 28.05, p < .001$], 条件と検証のタイプの交互作用 [$F(2, 36) = 6.23, p < .01$], プ

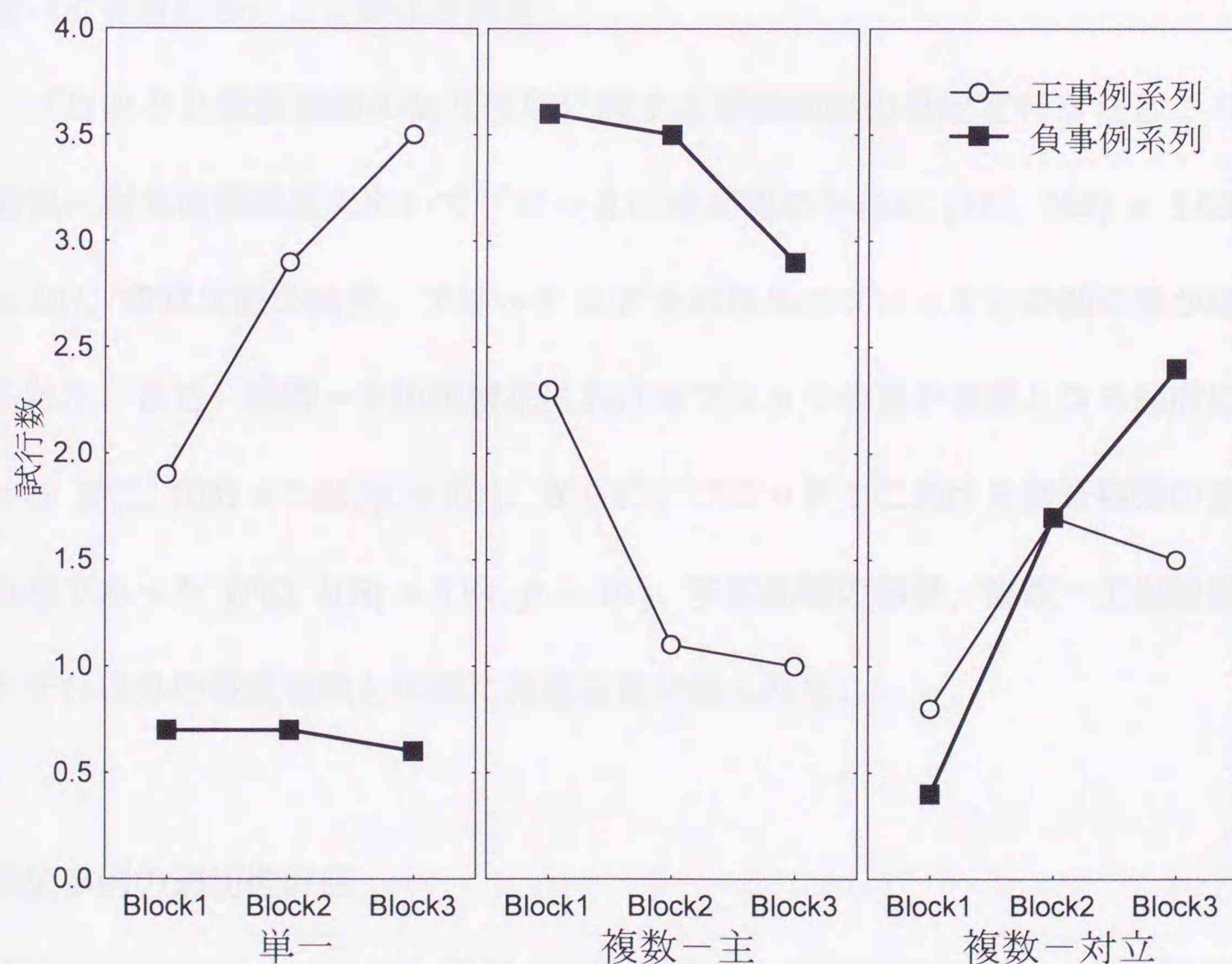


図3-5 単一・複数仮説試行の分布

ブロックと検証の種類の変異作用 [$F(4, 72) = 2.96, p < .05$] が認められた。

条件と検証の種類の変異作用について単純主効果の検討をしたところ、単一仮説検証の数に関する条件差が認められ [$F(1, 54) = 10.35, p < .01$]、正事例系列条件の方が単一仮説検証の数が多いことを示している。また、複数主仮説検証の数に関する条件差が認められ [$F(1, 54) = 8.18, p < .01$]、単一仮説検証とは逆に、この種の検証については負事例系列条件の方が多いことを示している。さらに、負事例系列条件において、3つの検証タイプの差が認められ [$F(2, 36) = 5.85, p < .01$]、多重比較の結果、複数主仮説検証がその他の2種類のものに比べて有意に多いことが示された。

ブロックと検証種類の変異作用に関する単純効果の検定を行ったところ、複数対立仮説検証においてブロックの差が認められた [$F(2, 108) = 5.63, p < .01$]。多重比較の結果、ブロック1とそれ以外のブロックとの間に差が認められた。また、複数主仮説検証におけるブロックの差が有意となる傾向にあった [$F(2, 108) = 2.81, p < .07$]。さらに、ブロック1における検証種類の差が有意であった [$F(2, 108) = 5.84, p < .01$]。多重比較の結果、複数主仮説検証とそれ以外の検証種類との間に有意な差が見られた。

検証事例の適切性評価

次に、各試行で与えられた事例の、検証事例としての適切性（-2: 全く不適当 ～2: 非常に適切）の評価を、条件、ブロック、事例の種類（確証・反証）別に示したのが表 3-2である。この適切性の評価について、3 要因の分散分析を

行った結果、事例の主効果、および条件と事例の種類の交互作用が認められた [それぞれ、 $F(1, 18) = 8.98, p < .01$, $F(1, 18) = 7.28, p < .05$]。事例の主効果については、確証事例の方が反証事例よりも高く評価されていたが、これは主として負事例系列条件における反証事例に対する評価の低さが影響を与えていると考えられる (表 3-2 下段右)。条件と事例の交互作用について、単純効果の検定を行ったところ、反証事例の評価に関する条件の差、および負事例系列条件における 2 種類の事例に対する評価の差が見られた [それぞれ、 $F(1, 36) = 7.28, p < .05$, $F(1, 18) = 16.22, p < .001$]。この結果から、特に負事例系列条件において、反証事例を用いた検証が適切ではないと評価されているといえる。この結果は、前述の事例選択傾向と対応している。正事例系列条件では、ブロック 1 では確証型検証の方が多く、ブロック 2 では反証型検証の方が多い。また、ブロック 3 では確証型検証と反証型検証の比率に差はなく、この結果は表 3-1 左側の正事例系列条件における適切性評価の傾向と対応している。また、負事例系列条件の評定も、確証型検証に対する評価が一貫して高いが、反証型検証に

表3-2 適切性評価の変遷 (受容方略課題)

	正事例系列			負事例系列		
	Block1	Block2	Block3	Block1	Block2	Block3
確証事例 (+H)	.50	-.07	.57	.50	.75	.83
反証事例 (-H)	.18	.06	.57	-.68	-.62	-.42

対する評価は一貫して低く、図 3-2 の事例選択傾向と対応した結果となっている。

3.2.3 考察

受容方略型課題における反応の特徴

本実験では、受容方略型課題における推論過程の特徴について吟味した。まず、事例の選択希望について、与えられる系列によって異なった傾向が見られた。正事例を多く含む系列を与えられた正事例系列条件では、被験者は反証事例の使用を希望する頻度が高く、特に中盤以降でその傾向が強く見られた。一方、負事例のみから構成される系列を与えられた負事例系列条件においては、一貫して確証事例が希望された。この結果について、現時点で考えられる説明は 2 通りある。1 つは、被験者が課題中に獲得した情報、特に反証型仮説否定情報 (false negative: $-H+T$) の数の違いに原因を求める説明である。正事例系列条件では、反証事例が呈示された場合に、それが結果として仮説を否定する情報となりうる可能性があるが、負事例系列条件では、反証事例が呈示されても、それが仮説を否定する情報とはなり得ない。そのため、正事例系列条件では、反証型仮説否定情報が呈示されることで、反証型検証の有効性が強く認識され、後の段階で要求されるようになったと考えられる。

もう 1 つの説明は、一般的な科学的探求の手続きとして、仮説の蓋然性を高める段階から、仮説をより洗練されたものにする段階へのシフトが存在するという考え方である。このときの事例選択傾向として、確証事例を多用する段階

から確証・反証事例を併用する段階への移行が観察されると予想される。すなわち、仮説を思いついた後は、まず初めに仮説に対する確信度を高めるための確証型検証を中心的に行い、ある程度自信が持てるようになったところで、見落としを確認するために反証型検証と確証型検証を併用すると考えられ、現実的な場面での仮説検証方略としては直感的にも理解しやすく、また理に適っているといえる。本実験の正事例系列条件では、確証型仮説支持情報が全てのブロックで2つ以上得られているが、負事例系列条件ではこの種の情報は全く得られない。そのため、正事例系列条件では仮説に対する確信度は試行を重ねる毎に高くなった結果、反証事例による検証が要求される頻度が多くなったが、負事例系列条件では確信度が高くなり、引き続き確証事例による検証が希望されたと考えられる。

仮説を複数立てたか、立てた場合どのような仮説に注目したかについて分析した結果は、以下のように要約される。まず正事例系列条件では、比較的初期の段階では、複数の仮説を立てた上で最も可能性の高い主仮説に注目した試行が多いが、試行が進むにつれて仮説を一つだけ考える場合が多くなる。また中盤では、複数の仮説を立てながら対立仮説に注目した検証を行う場合がやや多く、仮説の絞り込み → 可能性の低そうな対立仮説の排除 → 仮説の洗練という段階で仮説検証が進行していることをうかがわせる。一方、負事例系列条件では、仮説を一つだけ考える試行は少なく、複数の仮説を同時に考えることが多かった。この条件の被験者は、負事例しか与えられなかったため、可能性が高そうな仮説を一つに絞り込むことができず、課題解決に失敗したと考えられる。

3.3 第3実験 — 仮説評価課題における推論の特徴

仮説評価課題は、他者の行った仮説検証過程について、その適切性等を評価する課題であり、第2実験で用いた受容方略課題とは、受動的な情報収集という課題手続き上の共通点を有している。Tschirgi (1980) は、日常的因果推論を題材とした仮説評価課題を用いて、評価状況では反証型検証が好まれることを示した。また、Farris & Revlin (1989) は、反証型検証の選択には、結末の望ましさという要因が媒介変数となっており、反証型検証を用いることで得られるであろう結果が持つ潜在的なコストが高い場合は、むしろ確証型検証が用いられやすいことを示した。しかし、Kareev & Halberstadt (1993) は、Farris & Revlin の課題が、結果の望ましさと反証・確証の要因が交絡しているという問題点を抱えていることを指摘し、仮説検証研究で伝統的に用いられてきた2-4-6問題を評価課題に転用した実験課題によって、確証型検証の方が反証型検証に比べて有効なテストであると評価されることを示した。また、一方で仮説を否定する結果の情動的価値は高く評価されることも示された。しかし、これらの仮説評価研究でも、Wason (1960) 以降の仮説検証研究と同様、評価過程を全体的、かつスタティックに記述する傾向があることを指摘しなくてはならない。本実験では、仮説検証プロセスの変化と対応して、評価プロセスも段階的に変化するという可能性を検討するために、第1、第2実験で用いられた動態的変容分析パラダイムを適用して評価プロセスが段階を経て変化するかどうかを検討する。また、本実験で得られた結果と、第1、第2実験で得られた結果とを

比較し、受動的情報収集手続きが推論過程に与える影響について考察する。

本実験では、第2実験と同様に、第2章の第1実験で用いられた規則発見課題に基づいて、仮説評価課題を作成した。また、評価すべき仮説、検証事例は、第2実験の受容方略課題で用いられたのと同じ被験者が行った検証過程のデータを元に作成した。

3.3.1 方法

課題及び手続き

実際に行われた仮説検証過程を評価する課題を、質問紙を用いて行った。被験者は、第2実験の受容方略課題と同様に、第1実験の規則発見課題における正解到達者および未到達者各1名の試行系列を与えられた。以降それぞれの系列を与えられる条件を正解到達条件 (*acquired*)、未到達条件 (*not-acquired*) と呼ぶ。正解到達条件は第2実験の正事例系列条件に対応し、未到達条件は負事例系列条件に対応した。課題の性質上、第2実験では先行実験の被験者の仮説が呈示されなかったが、本実験では仮説も併せて呈示された。被験者には第1実験と同様のカバーストーリーが与えられたが、「同じ研究に従事している同僚の考え方について評価を求める」という趣旨の教示が追加された。

各試行において被験者は、まず第一に「同僚」の仮説を教えられ、それを書きとめた。続いて、被験者は、呈示された仮説がどの程度正しいと思うかを0～100点で評価した。第二に「もし自分が上記のような仮説をテストするにあたって、自由に条件の組み合わせを選択できるとしたら、どのようなテスト

をすべきだと思うか」を尋ねた。これは、質問紙中のチェックリストに記入させた。第三に、「同僚」が選択した実験条件の組み合わせ、すなわち検証事例が与えられ、被験者はこの情報を後で参照できるように質問紙中に書き留めた。第四に、この条件選択について、(a) 仮説を検証するのにどの程度適切な選択であるか (-2: 全く不適切 ~ 2: 非常に適切), (b) この選択についてどの程度予想していたか (-2: 全く予想外 ~ 2: 全く予想通り), (c) 結果の予測を尋ねた。第五に、実験結果が伝えられた。第六に、(a) この結果についてどの程度意外だと思うか (-2: 全く予想外 ~ 2: 全く予想通り), (b) この情報は最終的な規則の発見にどの程度有効であると思うか (-2: 全く役に立たない ~ 2: 非常に役に立つ) を評価するよう求めた。結果についても、検証事例と同様にメモをとるよう求めた。被験者は以上の手続きを 18 試行繰り返し、その後「同僚」の検証過程について気づいたこと、もし自分が「同僚」の立場に立ったときに取ったであろう方法について自由に記述させた。また、1 試行目にあたる情報は、初期情報として与えられるものであり、反応の対象とはならないことが伝えられた。

被験者

被験者は、大学の学部・大学院学生 20 名 (男子 8 名, 女子 12 名) で、年齢は 19~26 歳であった。被験者は、10 名ずつ無作為に前述した 2 種類の情報条件のいずれかに割り当てられた。実験は個別、あるいは数名の集団で実施され、40~60 分程度で終了した。

3.3.2 結果

選択希望事例

被験者に呈示された試行系列のうち、両条件の第 1 試行、および正解到達条件の第 2 試行では、「同僚」が立てた仮説が「わからない」というものであったため、両条件とも第 1, 第 2 試行のデータは以降の分析の対象から除外した。反応の対象となった全 16 試行において、呈示された仮説を検証するにはどのような事例を用いるべきか、すなわち被験者が使用を希望した事例を、呈示された仮説に対する確証・反証事例に分類した結果、正解到達条件と未到達条件の間に大きな違いが見られた。正解到達条件においては、確証事例 (5.4 試行) より反証事例 (10.5) の方が多いのに対し、未到達条件ではほとんどの試行 (15.6) で確証事例が選択されており、第 2 実験における正事例系列条件と負事例系列条件との間の差と同じ傾向を示している。

一般に仮説評価課題では、反証型検証が採用されやすい傾向にあることが示されている (例えば, Farris & Revlin, 1989; Tschirgi, 1980) が、本結果は呈示される情報によって選択される事例に違いがあることを示している。ここで、第 2 実験と同様に、全試行を 3 つのブロック (3-6, 7-12, 13-18) に分け、それぞれのブロックで確証型検証、反証型検証がどの程度希望されたかを 図 3-6 に示した。図に示される通り、正解到達条件においては、徐々に反証事例の使用希望頻度が高くなり、13 試行以降の終盤部で特にその傾向が顕著となる (6 試行中 4.6 試行) 一方で、未到達条件においては、一貫して確証事例の

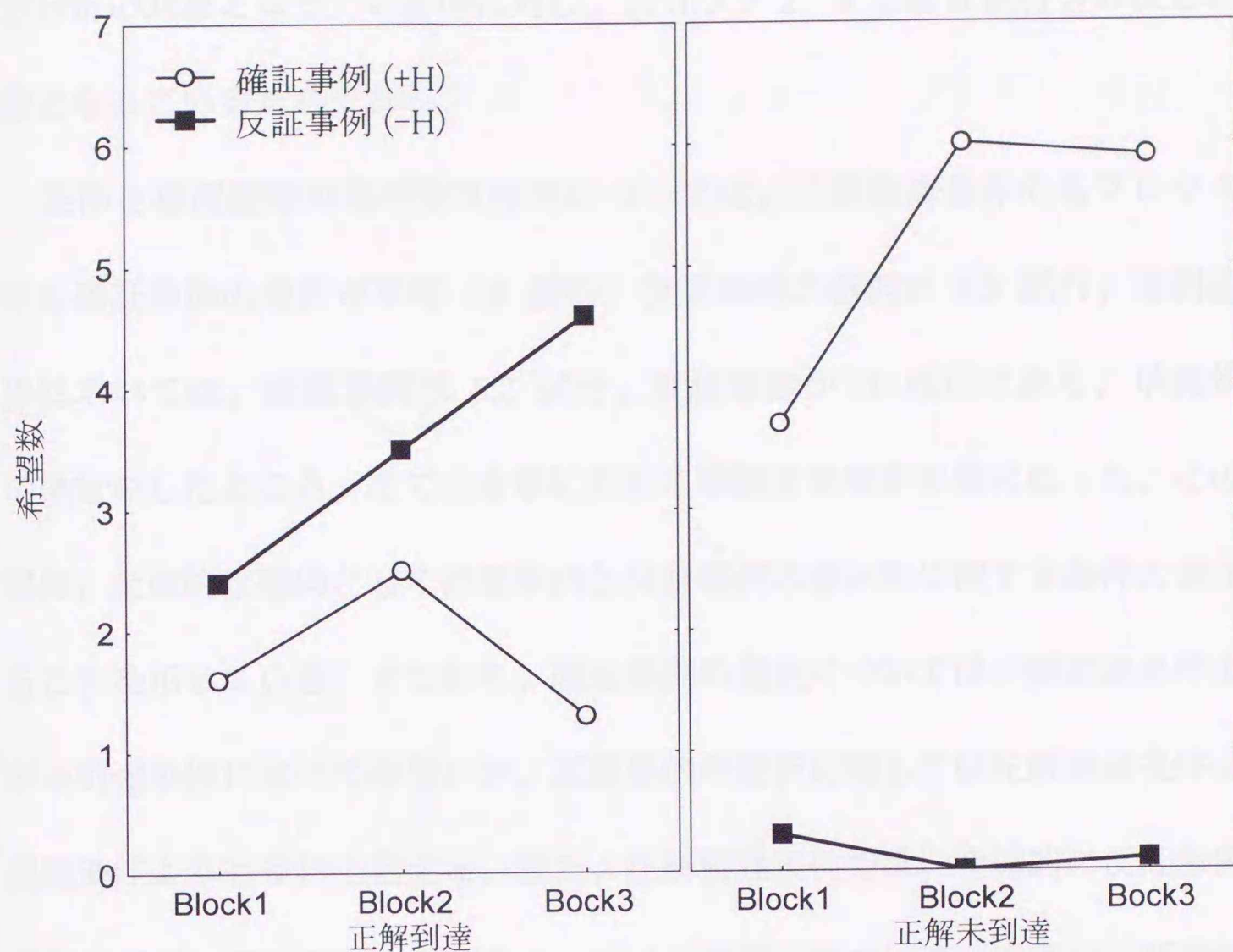


図3-6 選択希望事例の変遷 (仮説評価課題)

使用を希望する頻度が極めて高い。この傾向も、第2実験の受容方略課題の結果と対応している。この結果について、条件 (正解到達・未到達)、ブロック、事例の種類 (確認・反証) を独立変数とした分析を行ったところ、ブロックの主効果 [$F(2, 36) = 1561, p < .001$], 事例の主効果 [$F(1, 18) = 15.24, p < .001$], 条件と事例の種類の交互作用 [$F(1, 18) = 61.56, p < .001$], ブロックと事例の種類の交互作用 [$F(2, 36) = 3.51, p < .05$], 3要因の交互作用 [$F(2, 36) = 10.58, p < .001$] が認められた。ブロックの主効果については、ブロック1では4試行分の反応

が分析の対象となっているのに対し、ブロック 2, 3 では 6 試行分が反応の対象となっているためである。

条件と事例種類の間の交互作用については、正解到達条件の各ブロックにおける確証事例の選択が平均 1.8 試行、反証事例の選択が 3.5 試行、未到達条件においては、確証事例が 5.2 試行、反証事例が .1 試行であり、単純効果の検定をしたところ、全ての水準における単純主効果が有意となった。この結果は、全体的な傾向として確証事例と反証事例の選択数に関する条件の差があることを示している。すなわち、確証事例の選択については正解到達条件の方が未到達条件に比べて少ないが、反証事例の選択に関しては正解到達条件は未到達条件よりも多いと言える。また、正解到達条件では、全体的に反証事例の選択の方が、確証事例よりも多く、逆に未到達条件では確証事例が反証事例を上回っているといえる。ブロックと事例の間の交互作用についても、全ての水準で単純主効果が有意であった。

3 要因の交互作用についても、同様に単純効果の検定を行ったが、ここでは主要な結果のみを記述することとする。まず、正解到達条件のブロック 1、ブロック 2 では、確証事例と反証事例の選択数に差は見られないが [それぞれ、 $F(1, 54) = 1.14$, $F(1, 54) = 1.18$, 共に $p > .10$]、ブロック 3 においては反証事例が確証事例より多く選択された [$F(1, 54) = 19.36$, $p < .001$]。一方、未到達条件では、ブロック 1~3 の全てにおいて確証事例の方が反証事例よりも多く選択されていた [それぞれ、 $F(1, 54) = 20.55$, $F(1, 54) = 64.01$, $F(1, 54) = 59.91$, 全て $p < .001$]。

以上の通り，仮説評価課題においても，呈示される情報の系列によって，事例の選択傾向に違いがあることが示された．特に，第2実験の正事例系列条件と同様に，本実験の正解到達条件では反証事例の選択が多いという結果が得られた．しかし，前者では，17 試行中確証事例と反証事例が概ね同程度選択されていた（それぞれ，7.9，7.2 試行）のに対し，同じ事例と結果の組み合わせを与えられた後者では，反証事例が確証事例の2倍程度選択されたという違いがあり，特に終盤のブロック3における事例選択の差が顕著であるといえる．正事例系列条件（実験2）では，ブロック3における確証事例（3.2）と反証事例（2.7）の選択数はほぼ等しかったが，本実験の正解到達条件では，ブロック3における反証事例の選択数（4.6）は，確証事例の選択数（1.3）比べて多いといえる．正事例系列条件と正解到達条件における事例選択傾向について，条件，ブロック，事例の種類を独立変数とした分散分析を行った結果，ブロックと事例の種類の交互作用，および3要因の交互作用が認められた [それぞれ， $F(2, 36) = 4.47$ ， $F(2, 36) = 4.79$ ，共に $p < .05$]．特に3要因の交互作用については，ブロック3における確証事例，反証事例それぞれの選択数について条件の差が認められた [共に， $F(1, 108) = 6.59$ ， $p < .05$]．

ここで，被験者に与えられた情報について注目する（表3-3）．受容方略課題においては，呈示された検証事例が仮説の確証事例となるか反証事例となるかは被験者が生成した仮説の内容に依存するが，仮説評価課題においては仮説と呈示事例との関係は予め決められている．受容方略課題の正事例系列条件においては，ブロック1のうち，1.7 試行が反証事例であったが，そのうち結果と

して仮説が否定された ($-H+T$) のは 1.0 試行であった。一方で、同じ事例を与えられた仮説評価課題の正解到達条件では、ブロック 1 のうち、2.0 試行が反証事例で、そのすべてが仮説を否定する結果となった。また、続くブロック 2, 3 においても、一貫して正解到達条件の方が、反証事例による仮説の否定を多く得ている。このように、両条件では、反証型仮説否定を得る頻度が異なっており、このことが仮説評価課題においてより多く反証事例を選択させる方向に影響したと考えられる。特に、本実験の正解到達条件において、反証型検証の有効性をより強く認識した結果、このような選択傾向が見られたと推測される。

もう一つの可能性として、検証段階のシフトに伴う事例選択傾向の変化という側面からの検討を試みる。正事例系列条件 (第 2 実験)、正解到達条件 (第 3 実験) において、最も可能性が高いと考えられている仮説に対する確信度の

表3-3 正事例系列条件、正解到達条件における情報獲得数の違い

		正事例系列条件			正解到達条件		
		Block1	Block2	Block3	Block1	Block2	Block3
仮説支持	$+H+T$	1.8	2.8	2.8	1.0	3.0	3.0
	$-H-T$	.7	.4	.4	0	0	1.0
仮説否定	$+H-T$	.2	.2	.4	1.0	1.0	0
	$-H+T$	1.0	1.4	1.9	2.0	2.0	2.0

評定の変動を示したのが、図 3-7 である。図に示される通り、一貫して正解到達 (*acquired*) 条件の方が、仮説に対する確信度が高いといえる。条件、およびブロックを独立変数として分散分析を行ったところ、ブロックの主効果が認められた [$F(2, 36) = 7.93, p < .01$]。また、条件の主効果が有意となる傾向にあった [$F(1, 18) = 3.20, p < .10$]。交互作用は有意ではなかった。ブロックの主効果については、ブロック 1 とそれ以外のブロックの間に差が認められた (いずれも、 $p < .01$)。概ね全体的な傾向として正解到達条件の方が正事例系列条件よりも高いようではあるが、条件の主効果はそれほど強くなく、正事例系列

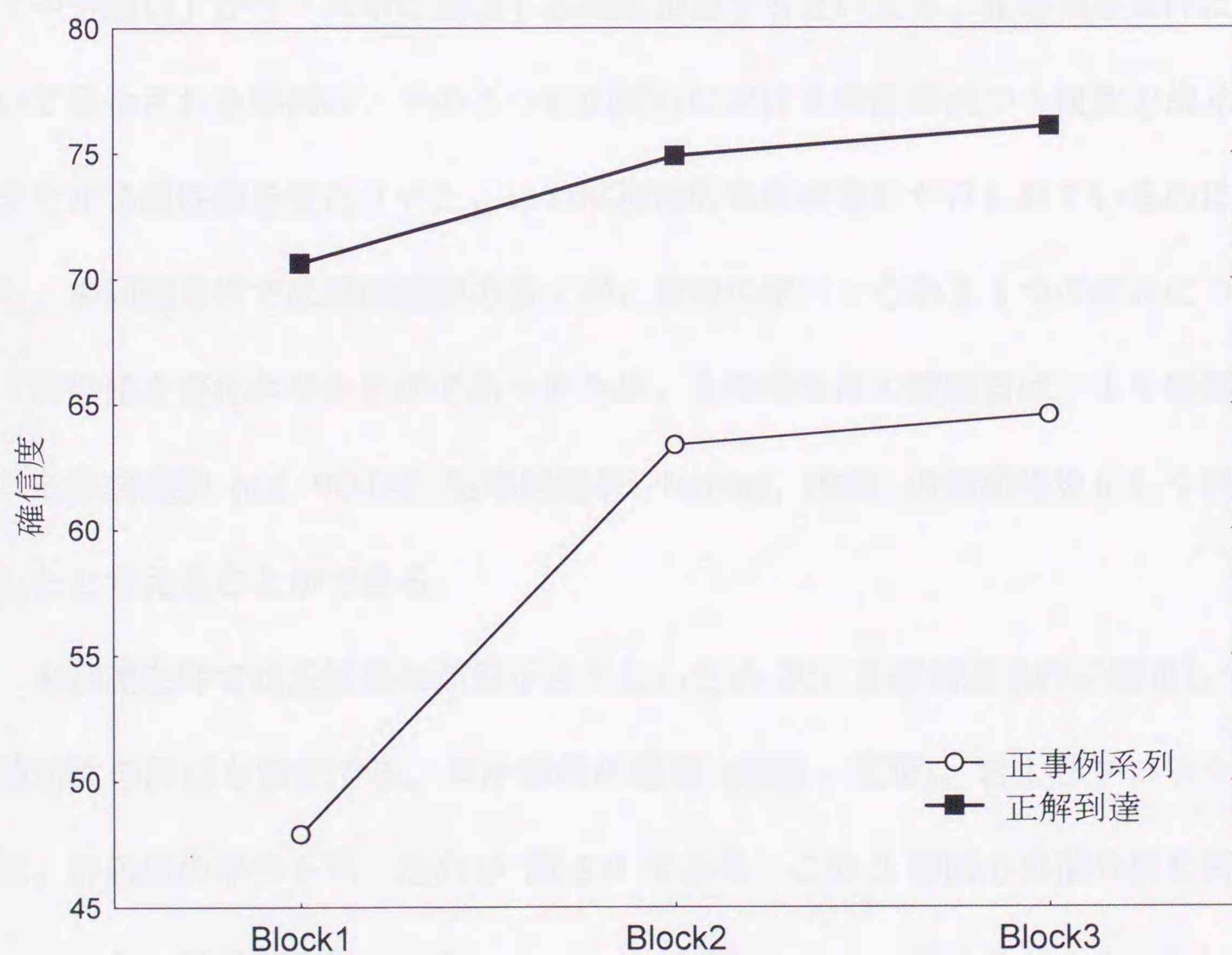


図3-7 仮説に対する確信度の変動

条件に比べて、正解到達条件の方で反証型検証の選択率が高いのは、特に後者の条件で、多くの反証型仮説否定情報を得たためであると推測される。

検証事例の適切性評価

検証事例選択の適切性は、呈示された検証事例が仮説の真偽を検討するのに適しているかどうかの評価として、 $-2 \sim 2$ までの 5 件法で求めた。被験者の評価の平均値は、正解到達条件で .53、未到達条件では 1.18 であり、未到達条件の方が全体的に高い評価であった。正解到達条件の評価平均値は、「どちらともいえない」から「やや適切」の間に位置し、未到達条件の評価平均値は「やや適切」から「非常に適切」の間に位置するといえる。正解到達条件において呈示された事例は、その 1 つ前の試行における検証事例から複数の次元にまたがる属性値を変化させた、いわば投機的な事例選択が行われているのに対し、未到達条件では事例選択の多くが、直前の試行からある 1 つの次元について属性値を変化させたものであったため、未到達条件の被験者は、より統制された事例選択 (c.f. VOTAT 型事例選択: Tschirgi, 1980) の適切性を正しく評価したと考えることができる。

未到達条件では反証事例が呈示されないため、次に正解到達条件に限定して、適切性の評価を検討する。呈示事例の種類 (確証・反証)、およびブロック別に、評価値の平均を示したのが 図 3-8 である。この 2 要因の分散分析を実施した結果、呈示事例の主効果 ($F(1, 9) = 10.94, p < .01$) が認められた。これは、反証事例が呈示された場合 (.75) の方が、確証事例が呈示された場合 (.33) よ

りも、選択が適切であると評価される傾向があった。この結果は、前述した被験者の事例選択の希望を反映していると考えられる。すなわち、正解到達条件では反証事例が希望される頻度が高く、被験者の希望通りに反証型検証が実施されたときに、より適切性の評価が高くなったと考えられる。また、2 要因の交互作用が有意となる傾向にあった [$F(2, 18) = 2.70, p < .10$]。反証事例の適切性評価は、試行段階が進む毎にやや低くなっているが、確証事例の適切性評価は、逆に高くなっているといえる。この結果は、第2実験の受容方略型課題と

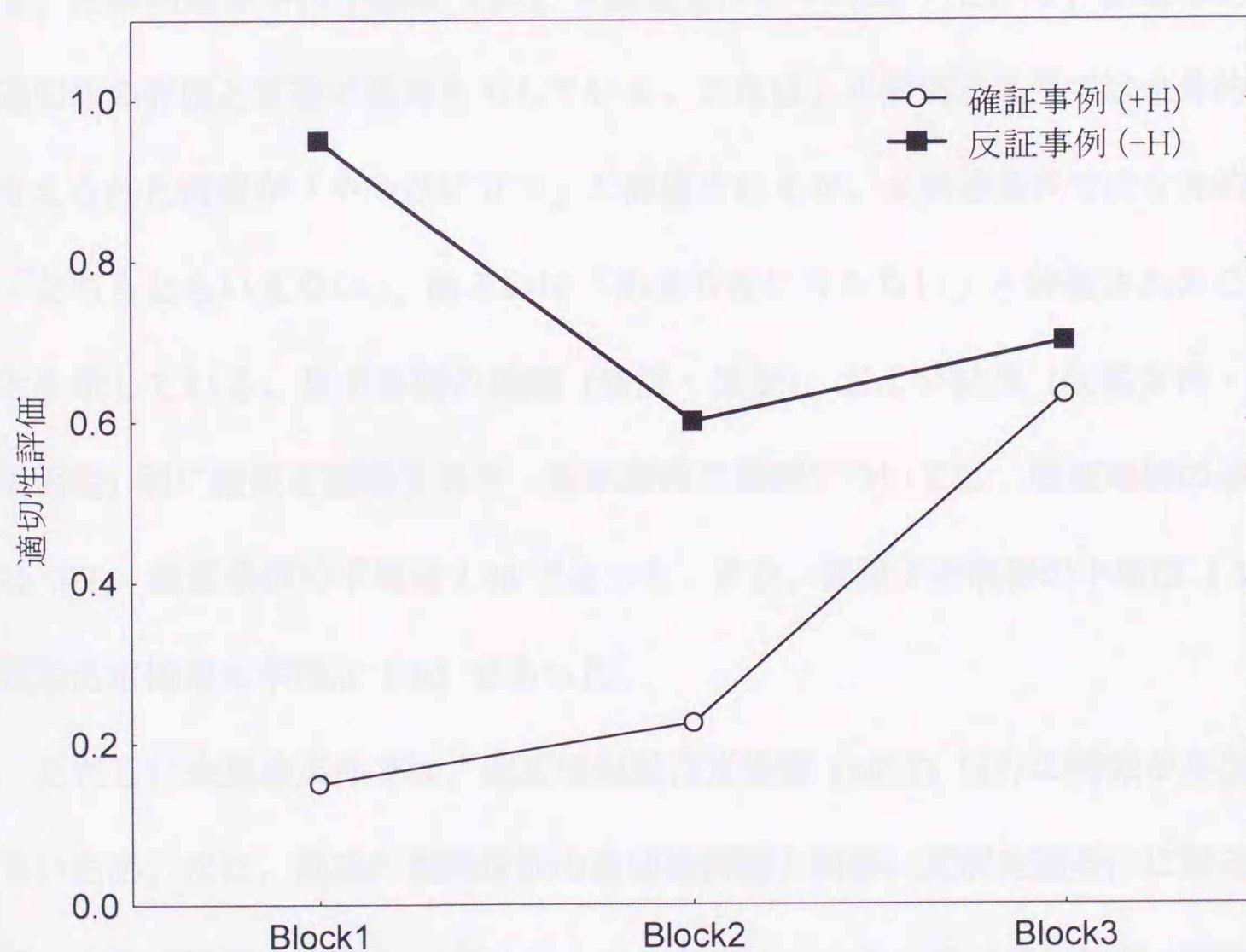


図3-8 正解到達 (acquired) 条件における呈示事例の適切性評価

異なり、事例選択傾向とは必ずしも対応していない。事例選択傾向では、ブロックが進行するにつれて反証型検証の選択率が上昇したが、適切性評価では、全体的に反証型検証の適切性は高く評価されているものの、課題の進行に伴って反証型検証の適切性が上昇するという傾向は見られなかった。

検証事例の有効性評価

被験者は、仮説・検証事例・結果の情報をすべて呈示された後に、この一連の検証過程の有効性を $-2 \sim 2$ の値で評価した。この評価の全体的傾向としては、正解到達条件の平均は 1.25 、未到達条件の平均は $-.31$ で、検証事例の適切性の評価とは逆の傾向を示している。これは、正解到達条件では全体的に与えられた情報が「やや役に立つ」と評価されるが、未到達条件では全体的に「どちらともいえない」、あるいは「あまり役に立たない」と評価されたことを示している。呈示事例の種類（確証・反証）、および結果（仮説支持・仮説否定）別に結果を整理すると、呈示事例の種類については、確証事例の平均は $.89$ 、反証事例の平均は 1.46 であった。また、仮説支持情報の平均は 1.14 、仮説否定情報の平均は 1.03 であった。

ただし、未到達条件では、確証型仮説否定情報（ $+H-T$ ）以外の情報が存在しないため、次に、前述の事例選択の適切性評価と同様に正解到達条件に限定して、有効性評価を検討する（図 3-9）。この評定について、呈示事例種類（確証・反証）、および結果（仮説支持・否定）の 2 要因による分散分析を実施した結果、事例の種類の主効果のみが認められた [$F(1, 9) = 7.74, p < .05$]。交互作用は

有意ではなかったが，反証型仮説支持情報 ($-H-T$) の評価が高く，全体的な評価の傾向と同様，過去の研究の知見とは相反する結果となっていた (c.f. Kareev & Halberstadt, 1993).

しかし，ここで注意しなくてはならないことがある．それは，呈示事例と結果に対する予測の関係，すなわち，被験者が呈示された仮説をどう考えているかである．呈示事例，結果の予測，実験結果の関係は，表 3-4 のように被験者自身の考えが否定されるかどうか，および呈示された仮説が否定されるか

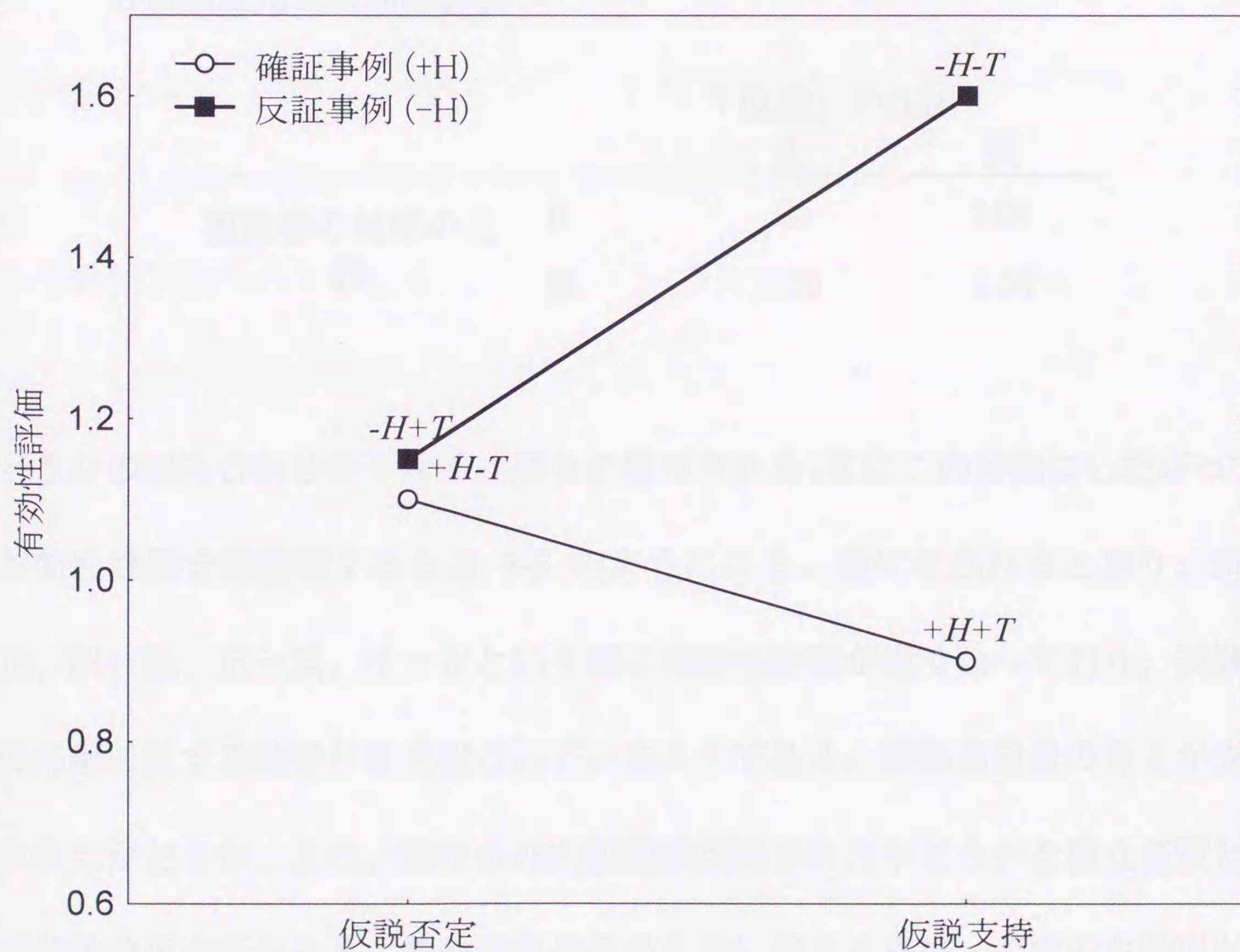


図3-9 正解到達 (acquired) 条件における情報の有効性評価

表3-4 呈示事例，予測，結果の関係とその意味

事例	予測	結果	状態	意味
確証事例	正	正	正－正	自分も「仮説」も正しい
		負	誤－誤	自分も「仮説」も間違っている
	負	正	誤－正	自分は間違っている，「仮説」は正しい
		負	正－誤	自分は正しい，「仮説」は間違っている
反証事例	正	正	正－誤	自分は正しい，「仮説」は間違っている
		負	誤－正	自分は間違っている，「仮説」は正しい
	負	正	誤－誤	自分も「仮説」も間違っている
		負	正－正	自分も「仮説」も正しい

表3-5 被験者が呈示された仮説に対して持つ見解，および仮説の正誤に基づく

分類別に見た有効性評価

		「仮説」の正誤	
		正	誤
被験者の見解の正誤	正	.86	1.06
	誤	1.29	1.26

どうかの組み合わせからなる4通りに整理される。次にこの分類にしたがって、有効性評価を再整理すると表3-5のようになる。表に示されるとおり、誤－正，誤－誤，正－誤，正－正という順に有効性評価が低くなっており、被験者の見解に反する結果が重要視されているようである。被験者自身の考えが支持されたかどうか，また，呈示された仮説が支持されたかどうかを独立変数として分散分析を行なうと，被験者自身の考え方に対する支持・否定の主効果が有意に認められた [$F(1, 156) = 5.34, p < .05$]。呈示された仮説に対する支持・否定の主効果，および両者の交互作用は有意ではなかった [共に， $F(1, 156) < 1$]。

被験者自身の考えに対する支持・否定については、否定情報を重視しており、この結果は、「最終的な目的は正しい規則を発見することであるが、実際に毎回実施するのは同じ研究に従事している『同僚』の考え方を評価することである」という教示を与えたことにより、被験者自身の考え方の正しさに対する認知が情動的価値の評価に大きく影響したという解釈が成り立つであろう。

検証事例の意外性評価

被験者は、「同僚」による事例選択、すなわち呈示事例と、実験結果の意外性を、それぞれ-2 ~ 2 の値で評価した。まず、呈示事例の意外性について、被験者の希望（確証・反証）と呈示事例の種類（確証・反証）別に示したのが表 3-6 である。未到達条件では反証型検証が選択されていないため、この表は正解到達条件のみの結果を示している。2 要因の分散分析を行なったところ、いずれの効果も認められなかった [それぞれ、 $F(1, 155) = 1.73, 2.17, .36$, 全て $p > .10$]。被験者の希望と、呈示事例の種類が異なる場合は意外だと評価される、すなわち評価値が低くなることが期待されたが、希望が確証事例 (+H) であるときに反証型検証 (-H) が行われた場合の評価がそれほど低くはなく (.20)、反証事例が希望された時に確証型検証が行われた時は評価が低くなっていた (-.30)。これは反証型検証の適切性に関する高い評価によって、意外性の評価が歪められたのではないかと推測される。

表3-6 呈示事例の意外性評価

		呈示事例	
		+H	-H
希望事例	+H	.03	.20
	-H	-.35	.06

表3-7 結果の意外性に対する評価

		結果	
		正	負
予測	正	1.32	-1.42
	負	-1.04	1.39

次に、実験結果の意外性について、被験者の予測（正・負）、および実際の結果（正・負）別に評価値の平均を示したのが表 3-7 である。実験結果についても、未到達条件では負事例以外はありえないために、正解到達条件に限定して検討した。この結果について予測と結果を独立変数とした分散分析を行なったところ、2 要因の交互作用のみ有意に認められた [$F(1, 156) = 341.26, p < .001$]。これは、予測と結果が一致する場合意外でないと評価し、一致しない場合意外であると評価したことを示し、ごく自然な結果であるといえる。

「同僚」の仮説検証過程に対する総合評価

全ての試行を終えた後で、被験者は「同僚」の仮説検証過程について気づいたこと、自分が同じ立場に立ったとしたらどのような方法をとったかについて自由に記述した。その内容を分類した結果、正解到達条件においては、複数の次元にまたがって属性値を変化させるよりも、一つ一つ属性値を変化させる、より統制された事例選択、すなわち VOTAT 型事例選択を採用すべきとした被験者が最も多く、10 名中 5 名であった。また、2 名は反証型検証の有効性について言及しており、残り 3 名は無記入であった。一方、未到達条件においては、VOTAT を評価するものが 3 名、VOTAT を用いるべきでないとするものが 3 名と評価が分かれていた。その他では、複数の次元を組み合わせた仮説をたてるべきという、仮説の立て方について言及したものが 3 名いた。

3.3.3 考察

仮説評価課題における事例選択希望傾向

まず第一に、事例選択希望の傾向については、与えられた試行系列によって異なったパターンが観察されたといえる。正解に到達した被験者の試行系列を与えられた正解到達条件（第 2 実験の正事例系列条件と対応）では反証型検証が希望される頻度が高く、特に中盤以降でその傾向が強く見られた。一方、正解に到達しなかった被験者の試行系列を与えられた未到達条件（第 2 実験の負事例系列条件と対応）では、一貫して確証型検証が希望された。この傾向は、第 2 実験の受容方略型課題における事例選択希望傾向と極めてよく類似していた。

これまでの仮説評価研究においては、評価課題において反証型検証の選択が多くなることの原因として、規則発見課題と比べて要求される認知的負荷が低いことによるという説明が考えられてきたが、本実験の結果は、認知的負荷という要因では説明できないと考えられる。むしろ、第2実験の考察で述べた2つの説明、すなわち反証型仮説否定情報の獲得数の違いに原因を求める説明、あるいは、検証過程の段階的变化に原因を求める説明の方が妥当な解釈であると言えるであろう。この点については、後に詳しく考察する。

仮説評価課題における情報評価

情報の評価は2つの側面から検討された。まず第1に、呈示された検証事例が適切な事例選択であったかどうかという問題、第2に、呈示された結果が規則発見に対してどの程度有効（適切）か、すなわち結果の情報的価値の問題の2点である。仮説評価に関する先行研究では、2つの問題が混同されている（Farris & Revlin, 1989; Kareev & Halberstadt, 1993）が、本来、これら2つは独立に議論されるべきものであろう。本実験では、このような観点に基づき、検証事例としての適切性、結果の情報的価値を個別に検討した。

まず、検証事例としての適切性評価については、全体的傾向として、正解到達条件の方が、未到達条件よりも評価が低かった。これは、正解到達条件における事例選択が、直前の試行に比べて複数の次元にまたがる属性値を変更した投機的、あるいは非統制的選択であったのに対し、未到達条件では、直前の試行から、特定の次元のみに限定した属性値の変更が行われた統制的事例選択が

多く、この違いが適切性評価に影響を与えたと考えられる。正解到達条件で見られた非統制的事例選択、および未到達条件で見られた統制的事例選択は、Tschirgi (1980) による CA (Change All), HOTAT (Hold One Thing At a Time), VOTAT (Vary One Thing At a Time) の分類と類似している。CA 型の事例選択は、直前の試行から全ての次元に関して属性値の変更を行う選択であり、HOTAT 型の事例選択とは、直前の試行から、焦点次元以外の全ての次元に関して属性値の変更を行う選択である。CA にせよ、HOTAT にせよ、一度に多くの次元の属性値を変化させるため、統制的な事例選択であるとはいえない。一方の VOTAT 型の選択では、焦点次元のみにおいて属性値の変更が行われる統制的な事例選択であるだけでなく、いわゆる反証型検証に相当するものであり、規範的に適切な行動であるとされている (Farris & Revlin, 1989; Tschirgi, 1980)。その意味では、本実験の被験者の反応は、規範的に正しい選択を適切に評価できたと考えることもできるであろう。

また、正解到達条件に限定して、確証型検証と反証型検証のそれぞれがどのように評価されたかを分析したところ、全体的には反証型検証の評価が高いが、終盤に向かうにつれて両者の間の差がなくなるという傾向がみられた。反証型検証の適切性が高く評価されるということについては、仮説評価研究で得られた知見と一致している (Farris & Revlin, 1989; Tschirgi, 1980)。しかし、一方で、前述の通り、反証型検証が希望される頻度はブロック 1 からブロック 3 に移行するに従って高くなっており、この選択傾向と、実際に呈示された事例に対する適切性評価はかならずしも一致した傾向を示していない。

次に、結果の情動的価値について検討する。仮説検証において得られる情報は、確認型仮説支持、確認型仮説否定、反証型仮説支持、反証型仮説否定の4種類に分けられる。これら4種類については、反証型仮説否定 > 確認型仮説支持 > 確認型仮説否定 > 反証型仮説支持、という順序でそれぞれの有効性が評価されるということがこれまでの研究で知られている (Kareev & Halberstadt, 1993)。しかし、本実験の結果からいうと、反証型仮説支持に対する評価が最も高く、ついで反証型仮説否定、確認型仮説否定、確認型仮説支持という順序での評価がなされており、これまでの知見とは反する結果を得ている。

ここで観点を変えて、被験者が呈示された仮説をどのように考えているかに注目してみると、被験者の考え方が否定されるかどうか、結果の情動的価値に対する評価の主要な規定因である可能性が示された。これは、呈示された情報の価値の判断に、被験者の主観的評価が大きな影響を与えている可能性を示唆しており、被験者すなわち評価者が、呈示されている情報を基に独自の推論過程を進行させる、いわば参照学習 (reference learning) とでも呼ぶべき過程が存在していることを推測させる。評価者が、仮説評価課題において、呈示された情報から独自の推論を進行させているかどうかについては、第5章の第5実験で改めて検討する。

3.4 受動的情報収集手続きによる方略選択の変化

本章では、受容方略型課題、仮説評価課題という、受動的情報収集手続きを含

む課題を用いて、それぞれの課題における推論過程の特徴を検討した。科学的探求のプロセスは、仮説の生成、検証事例の選択、フィードバックの獲得と解釈、仮説の更新が繰り返される過程であり、従来の仮説検証研究で用いられてきた規則発見課題は、これら全てのプロセスを含んでいる。一方、受容方略課題は、検証事例の選択、すなわち仮説に対するテストが検証者の自由選択によって行われるのではなく、外部から情報が与えられる課題であり、仮説評価課題は仮説までもが外部から与えられる課題である。このように本章の第2、第3実験で用いた受容方略課題、および仮説評価課題には、受動的情報収集手続きという構造的な共通点を持っている。

これらの実験から、受容方略課題および仮説評価課題では、共に正事例を多く含む情報系列を与えられた場合に、反証型検証の選択が促進されることが示された。特に、仮説評価課題において、その傾向が顕著であり、受容方略型課題では確証型検証と反証型検証がほぼ同数選択されていたのに対し、仮説評価課題では反証型検証が確証型検証の2倍程度選択された。一方、負事例のみから構成される系列を与えられた場合は、受容方略課題でも仮説評価課題でも一貫して確証型検証が選択されるという傾向が見られた。この結果に対しては、(1) 正事例を多く含む系列で与えられた反証型仮説否定情報 (false negative) が、反証型検証の有効性を認識させるよう働いた結果、より後の段階で反証型検証の選択が増加したという可能性、(2) 科学的探求の一般的方略として、まず確証型検証による仮説に対する確信度を高めた後、反証型検証を利用した仮説の見落としの確認を目指すという方法論に基づいた反応の結果、終盤に向かうに

従って反証事例の選択が増加したという可能性の2つが考えられた。第2、第3実験の基となった、第1実験における反応では、一貫して確証型検証の選択数が多いということからは、受容方略課題、仮説評価課題の持つ受動的情報収集手続きによる影響、すなわち(1)の反証型仮説否定情報の獲得という条件の影響が強いようにも思われる。しかし、第1実験の問題点として、正解に到達したかしなかったかに関わりなく、30試行で実験が終了してしまったということが挙げられる。そのために、特に正解に到達しなかった被験者については、十分な検証を行う前に強制的に実験が終了させられてしまったという可能性は十分に考えられる。すなわち、反証型検証を用いて仮説の不備を調べるという段階に至る前に実験が終了させられたため、反証型検証の選択が少ないのではないかと推測される。さらに、最初に仮説のもっともらしさを高めた後に、見落としを確認するためのテストを行うという検証段階の移行に伴う事例選択傾向の変化は、直感的には非常に納得しやすく、また適応的な観点から考えて合理的な方法であると考えられる。これら2つの説明のうち、いずれがより強力な規定因であるかという問題は、章を改めてより詳細に分析する必要があると思われる。

さらに、受動的情報収集手続きに原因を求める反証型仮説否定の獲得数の違いという説明と、一般的方略としての事例選択傾向のシフトという説明は、必ずしも相互排他的な説明であるとはいえないということを指摘しておくべきであろう。例えば、第1実験における正解到達者の反応を見ても、反証型検証の選択数はそれほど多いとはいえない。一方、受容方略課題(第2実験)、およ

び仮説評価課題 (第 3 実験) の正事例を多く含む系列を与えられた被験者の反応では、反証型検証の選択数が極めて大きいといえる。特に、仮説評価課題での反証型検証の選択は極めて多く、これは受容方略課題で獲得する以上の反証型仮説否定情報を得たことによる促進効果であると考えられる。すなわち、人は一般的に反証型検証を、より後半の段階において仮説の見落としを確認する手段として使用するが、その傾向が受容方略課題、仮説評価課題において呈示された反証型仮説否定情報によってさらに促進されたと考えることもできないわけではない。

3.5 第 3 章のまとめ

本章において報告された第 2, 第 3 実験の結果から、少なくとも受容方略課題、仮説評価課題における被験者の仮説検証方略は、時系列に沿って大きく変容していることが示された。特に、受容方略課題において正事例を多く含む系列を与えられたり、仮説評価課題において正解に到達した系列を与えられた被験者は、中盤以降で反証型検証を多く希望する一方で、受容方略課題で負事例のみからなる系列を与えられたり、仮説評価課題において正解に到達しなかった系列を与えられた被験者は、確証型検証を希望する回数が極めて多いことが示された。この反応傾向の違いは、確証型検証重視→確証・反証型検証併用というシフトを伴う一般的な仮説検証方略のあらわれ、あるいは反証型仮説否定情報の呈示による反証型検証選択の促進という 2 つの考え方によって説明されるであろう。次章においては、これら 2 つの要因のうち、どちらが検証方略

第4章 反証型検証の選択を促進する要因

4.1 第4実験 — 反証型仮説否定の獲得か、一般的方略としての選択傾向のシフトか

本実験は、第3章の第2、第3実験で観察された、反証型検証の選択頻度の増加に影響を与える要因について検討するためにデザインされた。第2実験の正事例系列条件、および第3実験の正解到達条件では、特に中盤以降における反証型検証の選択頻度の増大が見られたが、第2実験の負事例系列条件、第3実験の正解未到達条件では、一貫して確証型検証が多く選択されることが観察された。この傾向に対しては、(1) 正事例系列条件、正解到達条件で被験者が獲得した反証型仮説否定情報 (false negative: $-H+T$) が反証型検証の有効性を認識させるよう働いたとする説明、あるいは (2) 科学的探求場面における一般的方略としての、時系列に沿った事例選択傾向のダイナミックな変容によるものとする説明の2つが考えられた。本実験では、いずれの要因が反証型検証の選択に対してより強い影響を与えているかを分析する。

以上の目的のため、本実験では一般的な規則発見課題とやや手続きの異なる課題を用いる。通常、規則発見課題では、被験者は自由に仮説の生成、および検証事例の選択を行うため、結果として得られる情報は被験者毎にまちまちになるのが普通である。本実験の目的である2つの説明に関する分析を行うためには、先行的に与えられる情報の種類を統制することによって、後の検証過程が変化するかどうか、変化する場合はどのように変化するかを観察する必要

がある。そこで、規則発見課題を受容方略段階 (reception strategy phase) と選択方略段階 (selection strategy phase) に分割し、前者の段階で呈示する情報を統制する。受容方略、選択方略は、Bruner et al. (1956) による概念学習研究において用いられた課題手続きであり、両者の違いは、概念を達成するための事例を、前者では被験者が自由に選択できず実験者が順次呈示するが、後者では被験者が自由に選択するところにある。これを規則発見課題に当てはめて考えると、前者は、仮説は自由に生成できるが検証事例を選択できない手続き、後者は、仮説の生成、事例の選択共に被験者の自由選択が可能な手続きであるといえる。これらの手続きを併用することで、先行的呈示情報を統制しつつ、被験者の自由な仮説検証過程を検討することが可能となる。本実験では、まず 3 試行からなる受容方略段階を行い、この段階で呈示する情報の種類を統制した。その後は、通常の規則発見課題と同様に、全ての事例選択を被験者が自由に行えるようにした。

また、本実験では、被験者の用いる仮説検証方略を多面的な側面から考察する試みとして、従来の確証事例および反証事例という、仮説との関係で決定される事例の選択傾向の分類に加えて、関連次元・無関連次元、焦点次元内・次元間変動という 2 つの側面を取り上げる。

関連次元 (relevant dimension)、および無関連次元 (irrelevant dimension) とは、元々は、Bruner et al. (1956) を始めとする概念学習研究によって提唱された用語であり、前者は概念の規定に参加する次元、後者は概念の規定に参加しない次元を指す (東, 1963; Bruner et al., 1956; 寺岡, 1966)。本実験課題における事例

空間は3つの次元から構成されており，ここでは，被験者が仮説において言及している次元を関連次元，言及していない次元を無関連次元として扱い，被験者がそれぞれの次元にどの程度注目するかを分析する．この時，被験者が関連次元のみ属性値を変化させた事例選択を行っている場合は「関連次元」，無関連次元の属性値も変化させている場合は「無関連次元」というラベルを与えて分類する．

一方の，焦点次元内，および次元間変動とは，被験者がある試行で選択した検証事例が，直前の試行で選択した事例と比べて，いくつかの次元で属性値を変化させたかで表現される．例えば，ある被験者が第4試行から第5試行に進んだ時に，3つの次元のうち1つ，例えば次元Aに関して属性値を変更したとする．さらに，この被験者が第5試行から第6試行に進んだ時に，次元A，次元Bという2つの次元の属性値を変更したとする．この場合，第4試行→第5試行で変更された次元Aを，被験者がこの系列で注目している次元（焦点次元）として考え，第5試行→第6試行で同じく行われた次元Aに関する属性値の変更に対して焦点次元内変動（within-dimension change）というラベルを付与する．一方，第5試行→第6試行で新たに行われた次元Bに関する属性値の変更は，焦点次元以外の次元へと注目がシフトしており，このような事例選択には次元間変動（between-dimension change）というラベルを付与する．

前述の関連次元・無関連次元という分類は，仮説と検証事例との関係に基づいて決定されていたのに対し，この焦点次元内・次元間変動という分類は，被験者の選択した検証事例のみに注目して，その選択の仕方を記述するものである

と言える。被験者が検証事例の選択に際して、直前の試行で用いた検証事例からどのように変化させた事例を選択するかについては、Halberstadt & Kareev (1995) が、属性値が変更された次元の総数を統制された実験 (controlled experiments) の指標として分析している。彼らは、変更の数が少ないほど統制された実験が行われたことになり、この場合は仮説の改良 (fine-tuning) が意図されていると考え、属性値が変更された次元数が多い場合は、ターゲット事例を獲得するための冒険的試み (looking for target items) としての性格が強いものとして扱った。しかし、彼らは、あくまで属性値が変更された次元の総数を問題としており、その変更が焦点次元内のみのものか、焦点次元以外の属性値を含むものかによっては分類していない。しかし、焦点次元内のみに限定した属性値の変更は、特定の次元において属性値のとりうる範囲を決定するために行われる一方で、焦点次元以外での属性値の変更は、それまでの試行で考慮していなかった別の次元へと視点が移動していることを示していると考えられ、焦点次元内変動と次元間変動は異なった目標を持っているといえる。また、次元間変動は、焦点次元が移動していることを示すものであり、この数が多いということは、あまり統制的ではない検証事例選択を行っていることを意味する。従って、これら2種類の属性値の変更の仕方は区別して論じられるべきであろう。本実験では、属性値が変更された次元の総数のみではなく、その変更の仕方によって焦点次元内変動、次元間変動の2つに分類し、それぞれがどの程度の頻度で行われるかを分析する。

4.2. 方法

課題

架空の細菌が繁殖する物理的条件を発見するという規則発見課題が用いられた。物理的条件は、(1) 温度 ($-10 \sim 30^{\circ}\text{C}$)、(2) 湿度 ($40 \sim 80\%$)、(3) 紫外線強度 ($60 \sim 100\text{ k J/m}^2$)、という 3 つの次元から構成されていた。また、全ての被験者が、初期事例として、温度 10°C 、湿度 60% 、紫外線強度 80 k J/m^2 (キロジュール毎平方メートル) という条件で細菌が繁殖するという情報を与えられた。被験者は、細菌が繁殖する条件に関する規則性を発見するために、仮説を生成し、それに対する検証を行った。ただし、最初の 3 試行に限り、検証事例を自ら選択せず、実験者から指定された事例を用いた検証を行う受容方略型手続きを用い、4 試行目以降から通常の実験方略型手続きを用いるという、ハイブリッド課題を用いた。検証事例の選択は、3 つの次元のそれぞれについて属性値を 1 つ選択することで行われた。フィードバックは細菌が繁殖したかどうかという形で与えられた。また、被験者は望むときに、それまでに選択した検証事例とそれに対応する結果からなる履歴を参照することができた。本実験では、被験者の事例選択を、従来の研究と同様の仮説に対する確証事例か反証事例かという分類だけではなく、関連次元のみで属性値を変更したか、あるいは無関連次元での属性値の変更を行ったかどうか、さらに、どれだけ多くの次元に関する属性値の変更を行ったかという点に注目して分類を行った。Halberstadt & Kareev (1995) は属性値の変更について、変更の総数のみを考えていたが、本実験課題は複数の次元から構成される事例空間を持つため、焦点

次元内変動,次元間変動という2つの側面から属性値の変更について検討する。

条件

本実験の目的は、反証型検証の選択が、先行的に呈示された情報の影響を受けるかどうかを検討することである。そのため、課題の最初の3試行において、被験者が事例を自由に選択できない受容方略型手続きを採用し、この段階で与えられる情報の種類を操作した。仮説検証過程において獲得可能な情報は、大別すると仮説を支持する情報と仮説を否定する情報の2種類があり、さらにそれぞれが確証型検証の結果によるものと反証型検証の結果によるものから構成される、計4種類の情報が存在する。そこで、これら4種類の情報のどれか1つが与えられる4つの条件、すなわち確証型仮説支持 (+H+T) 呈示条件、反証型仮説支持 (-H-T) 呈示条件、確証型仮説否定 (+H-T) 呈示条件、反証型仮説否定 (-H+T) 呈示条件が設定された。また、この操作は、受容方略型手続きで行われる3試行のうちの第2試行において実施され、残りの第1、第3試行では、全ての被験者に確証型仮説支持情報が与えられるようにした。第2試行において与えられる事例は、被験者が第1試行において生成した仮説に基づいて、その都度各条件に対応するものが1つ決定され、呈示された。

また、第1実験、第2実験のように試行回数を限定するのではなく、被験者自身が十分に検討し尽くしたと感じるようになるまで課題を続行することとし、さらに初期仮説の性質によって検証過程が変化しないように、被験者が報告した仮説を正しい規則として扱うこと、すなわち仮説支持情報が呈示される

条件 (+H+T, -H-T) では第 1 試行で挙げられた仮説，仮説否定情報が呈示される条件 (+H-T, -H+T) では，第 1 試行で挙げられた仮説が第 2 試行で否定されるため，第 3 試行で挙げられた仮説を，正解の規則とすることとした。

手続き

実験はコンピュータによって制御され，被験者の反応は全てコンピュータ上でなされた。被験者は，実験のカバーストーリー，および課題の目的に関する教示を読んだ後で，操作に慣れるため，4 回の練習試行を実施した。続く本試行において，被験者は毎回，(1) 仮説生成，(2) 検証事例選択ないし受容，(3) 結果の予測，(4) 結果の呈示および確認を行った。(1) の仮説生成については，被験者は毎試行において，その時点で最も可能性が高いと思われる仮説を 1 つ挙げた。以上の (1) ～ (4) までの反応は全てコンピュータに対して入力された。(1) の仮説については，被験者からの口頭報告を基に，実験者が入力した。(2) の事例選択ないし受容に関しては，実験開始から第 3 試行までは，検証事例が実験者によって指定される受容方略型手続きが採用された。受容方略型手続きの段階では，第 2 試行で指定される事例の種類は条件によって異なっていたが，第 1，第 3 試行で指定される事例は，各被験者の仮説を支持する確証型の検証事例であった。第 4 試行以降の選択方略型手続きの段階では，被験者は，自由に検証事例を選択することができた。被験者は，以上の (1) ～ (4) までの手続きを，現在考慮中の仮説以外に，細菌が繁殖する条件の規則性に関する説明がありえないことに確実な自信が持てるまで続け，被験者から仮説に絶対

の自信があるという報告がなされた時点で実験は終了した。

被験者

大学の学部学生、および大学院学生 40 名が、前述した 4 つの条件に各 10 名ずつ無作為に割り当てられた。

4.3 結果

総試行数

本実験は、規則発見課題ではあるが、従来の課題のように実験者側が想定する正しい規則というものは存在せず、被験者が第 1 試行で考えた仮説 (確証型仮説支持提示条件、反証型仮説支持提示条件)、あるいは第 3 試行で考えた仮説 (確証型仮説否定提示条件、反証型仮説否定提示条件) を正しい規則と見なした。さらに、実験は、被験者が十分な検証をし尽くしたと納得がいくまで続けられた。ここでは、まずはじめに、被験者が実験終了に至るまでに要した総試行数について分析する。

最も少ない試行数で実験を終了した被験者が要した試行数は 9 試行であった。また、最も多い試行数を要した被験者は 63 試行で終了した ($SD = 11.19$)。表 4-1 に、条件別にみた試行数の平均を示した。傾向を見ると、反証型仮説否定 < 確証型仮説否定 < 確証型仮説支持 < 反証型仮説支持、といったように、仮説否定情報を得た 2 条件の方が、やや早い段階で規則に到達したと判断する傾向がみられる。しかし、第 2 試行で与えられる事例の種類 (確証型・反証型)、

表4-1 呈示された情報による総試行数の違い

括弧内は SD

	確証型	反証型
仮説支持	24.6 (10.34)	25.2 (16.69)
仮説否定	21.3 (9.07)	18.7 (6.67)

表4-2 呈示された情報による短試行・長試行解決者の分布

		短試行解決	長試行解決
確証型	仮説支持	3	7
	仮説否定	4	6
反証型	仮説支持	6	4
	仮説否定	5	5

および結果 (仮説支持・仮説否定) を独立変数とした分散分析の結果は、いずれの主効果、交互作用も有意ではなかった [結果の主効果: $F(1, 36) = 1.88, p > .10$, 事例の種類的主効果, および交互作用: $F(1, 36) < 1$].

次に、全被験者の試行数の中央値 (19 試行) によって被験者を、短試行解決者 (19 試行未満で解決) と長試行解決者 (19 試行以上で解決) の 2 つに分け、4 つの条件におけるこれらの比率を比較した (表 4-2)。この結果について、比率の逆正弦変換値を用いて、呈示事例の種類、結果の 2 要因の効果を検討したところ、いずれの主効果、交互作用も有意ではなかった [事例の主効果: $\chi^2(1) = 2.33, p > .10$, 結果の主効果, および交互作用: $\chi^2(1) < 1$]。これらの結果から、解決に要した総試行数については、受容方略手続きの段階で仮説否定情報が呈示される条件 (+H-T, -H+T) の方がやや少ないようではあるが、統計的

な差は認められないということがいえるであろう。

事例選択傾向

次に、被験者が選択方略段階において仮説を検証するために選択した事例を、仮説に対する確証事例と反証事例に分類し、両者の選択頻度を比較する。本実験では、被験者毎に総試行数が異なっているため、これまでの実験のように特定の試行数によってブロックに分割することは不可能であること、また、総試行数が9試行しかない被験者がおり、この被験者は受容方略型段階を除いた選択方略型段階では総試行数が6試行と少なくなってしまうことから、多数のブロックへ分割することは適切ではないと思われる。そこで、それぞれの被験者の事例選択系列を前半（ブロック1）と後半（ブロック2）の2つのブロックに分割し、それぞれのブロックにおける事例選択傾向を分析する。

本実験においては、被験者は毎回、仮説を1つだけ挙げるよう教示されたため、被験者が選択した事例は、この仮説との関係においてのみ決定される。ブロック1とブロック2のそれぞれにおいて、確証事例（+H）・反証事例（-H）がどの程度用いられたかを条件別に示したものが図4-1である。この結果について、条件（2×2）、ブロック、事例の種類（確証・反証）を独立変数とした分散分析を実施した結果、ブロックの主効果、及びブロックと事例の種類の交互作用が認められた [それぞれ、 $F(1, 36) = 36.73, p < .001, F(1, 36) = 4.67, p < .05$]。このうち、ブロックの主効果については、総試行数が奇数である被験者の反応を2つのブロックに分割する際に、ブロック2の方が試行数が多くなるため

ある。

ブロックと事例の種類の交互作用については、ブロック 1 における確証事例の選択数は 5.2、反証事例は 4.3、ブロック 2 における確証事例の選択数は 4.2、反証事例は 5.8 であった。全体的な傾向として、確証事例はブロック 1 の方が多く、ブロック 2 で減少するが、反証事例はブロック 1 よりもブロック 2 の方が多くなっており、事例選択の時系列に沿った動態的変容という説明に対して支持を与えているようである。この交互作用について、単純効果の検定

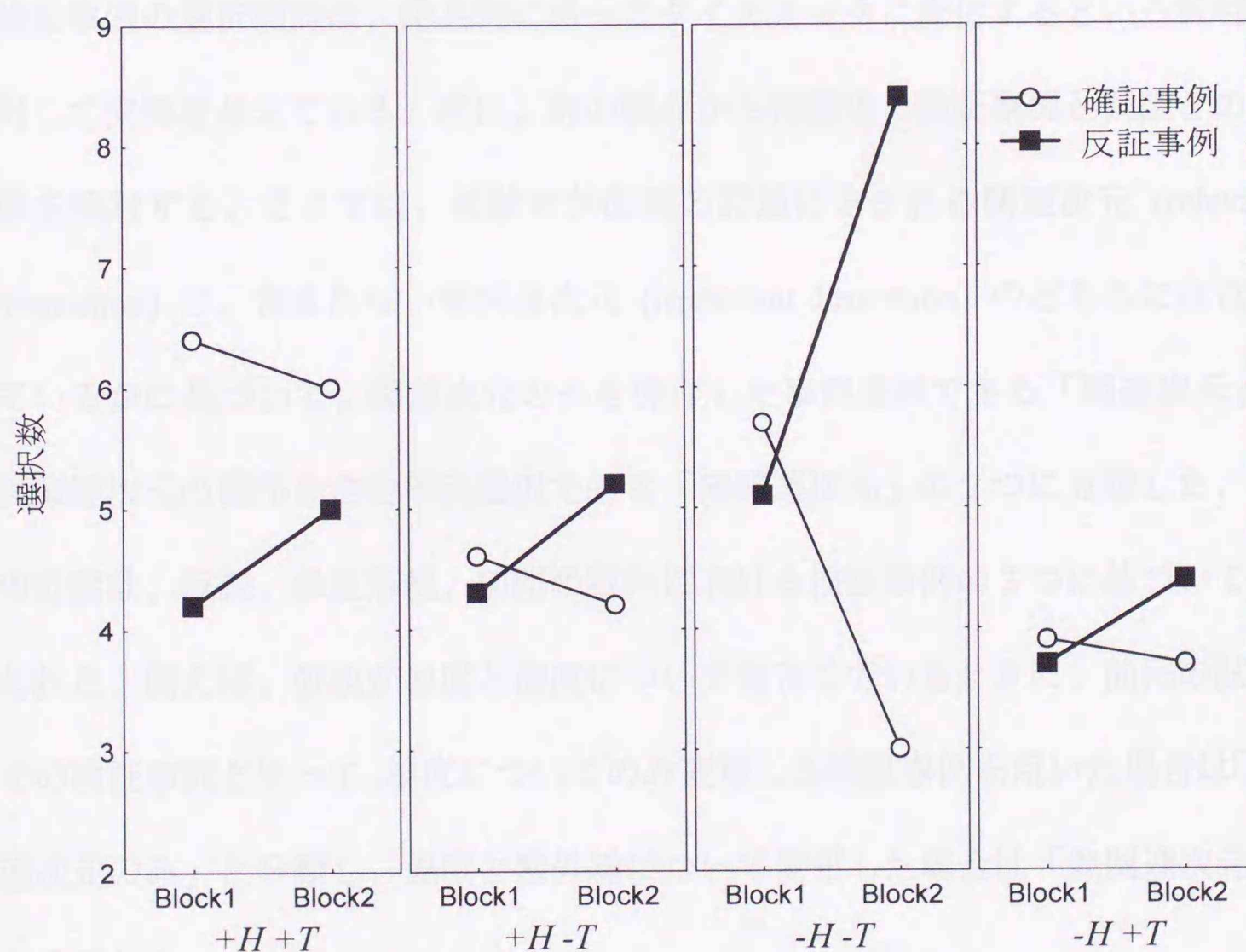


図4-1 検証事例の分布

を実施したところ、ブロック 1 における反証事例の選択数とブロック 2 における反証事例の選択数に有意な差が認められた [$F(1, 72) = 6.82, p < .05$]. これ以外の交互作用は認められなかった。以上の結果を総合すると、条件に関わりなく、ブロック 1 よりもブロック 2 において反証事例の選択数が多くなっており、一般的傾向としての事例選択のシフトという説明を支持する結果となっているといえるであろう。

関連次元・無関連次元

検証事例の選択傾向は、時系列に沿ってダイナミックに変化するという説明に対して支持を与えている。次に、別の観点から被験者の検証事例と仮説との関係を検討する。ここでは、被験者が仮説の記述に含まれる関連次元 (relevant dimension) と、含まれない無関連次元 (irrelevant dimension) のどちらに注目しているかに基づいて、関連次元のみを操作した事例選択である「関連次元」、無関連次元の操作を含む事例選択である「無関連次元」の 2 つに分類した。この分類は、仮説、検証事例、前回の試行における検証事例の 3 つに基づいて行われた。例えば、仮説が温度と湿度について言及しているときに、前回の試行での検証事例と比べて、温度についてのみ変更した検証事例を用いた場合は「関連次元のみ」と分類し、温度と紫外線について変更した場合は「無関連次元」と分類した。

以上の分類について、条件、ブロック別に示したのが 図 4-2 である。全体的な傾向として、関連次元のみを操作した事例選択が多いということがいえる。

ここで、条件 (2×2)、ブロック、事例選択型 (関連・無関連) を独立変数とした4要因の分散分析を行ったところ、ブロックの主効果、事例選択型の主効果、および両者の交互作用が認められた [それぞれ、 $F(1, 36) = 9.48, p < .01$, $F(1, 36) = 16.82, p < .001$, $F(2, 72) = 5.07, p < .05$]。交互作用について単純効果の検定をしたところ、ブロック 1, 2 共に事例選択型の単純主効果が見られた [それぞれ、 $F(1, 72) = 21.32, p < .001$, $F(1, 72) = 9.74, p < .01$]。これは、両方のブロックで「関連次元のみ」の数が「無関連」を上回っているためである。また、「無関連次元」についてブロックの単純主効果が認められた [$F(1, 72) = 8.35, p < .01$]

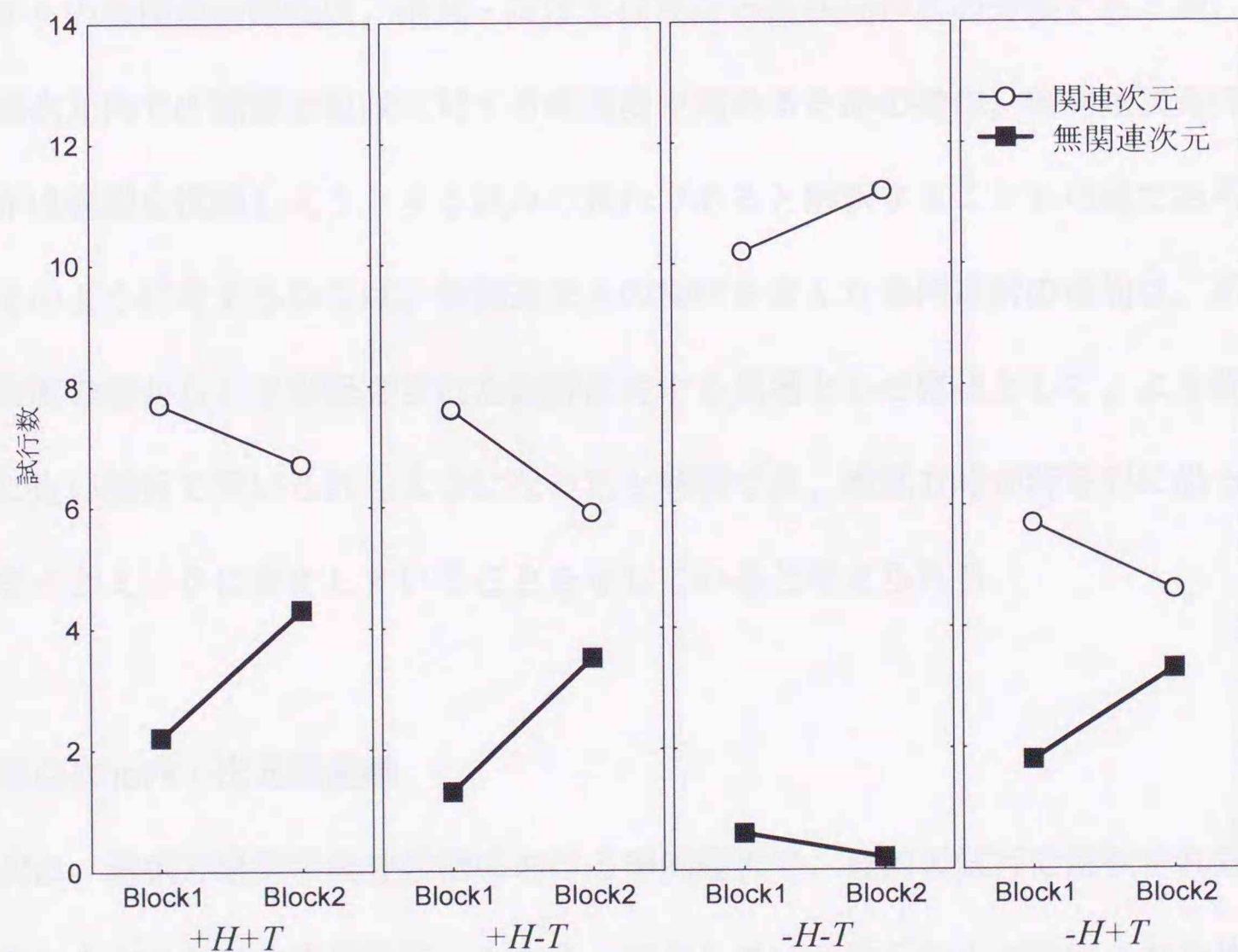


図4-2 関連次元・無関連次元のそれぞれに注目した事例選択数

J. 無関連次元の属性値を操作した検証事例の選択は、ブロック 1 (1.48) よりもブロック 2 (2.83) の方が頻度が高いといえる。また、これに伴って、関連次元のみを操作した検証事例の選択が、試行が進むにつれてやや少なくなっていくという傾向がみられる。この結果も、一般的傾向としての時系列に沿った方略の変容を支持するものであるといえる。このシフトは、まず仮説に対する確信度を高めるための検証が行われる段階が先行し、続いて見落としがないかどうかを確認するための検証が行われるという段階へと移行することを主張するものであり、このときの事例選択傾向として、確証型重視から確証型・反証型併用へのシフトが見られる。ここで検討した関連次元、無関連次元という観点からの事例選択傾向は、確証・反証とは異なった次元からの分類であるが、関連次元内での操作は仮説に対する確信度を高めるための操作、無関連次元の操作は仮説を洗練しようとする試みの表れであると解釈することも可能であろう。そのように考えるならば、無関連次元の操作を含んだ事例選択の増加は、ある程度の確からしさが保証された仮説に対する見落としの確認として、より終盤に近い試行で用いられるようになったと解釈でき、検証方略が時系列に沿ってダイナミックに変化していることを示していると考えられる。

焦点次元内・次元間変動

次に、選択方略型手続き段階における事例選択で、直前の試行で選択された事例からどれだけの変更を行ったかを、注目している次元のみに限定された焦点次元内変動、それ以外の次元に関する属性値の変動を行った次元間変動に分類

し、両者の数を集計した。本課題の事例空間は 3 次元であるため、1 試行あたりでの両者の総数は最高で 3 となる。例えば、温度という次元に注目しているときに、直前の試行から温度と湿度について属性値の変更を行った場合は、焦点次元内変動が 1、次元間変動が 1、合計で 2 としてカウントされた。ここで、注目している次元とは、仮説において言及されている次元ではなく、事例選択系列から推測される注目次元を指す。4 つの条件別に、各ブロックにおいて焦点次元内・次元間変動が、1 試行あたりそれぞれどの程度行われたかを示したのが、図 4-3 である。

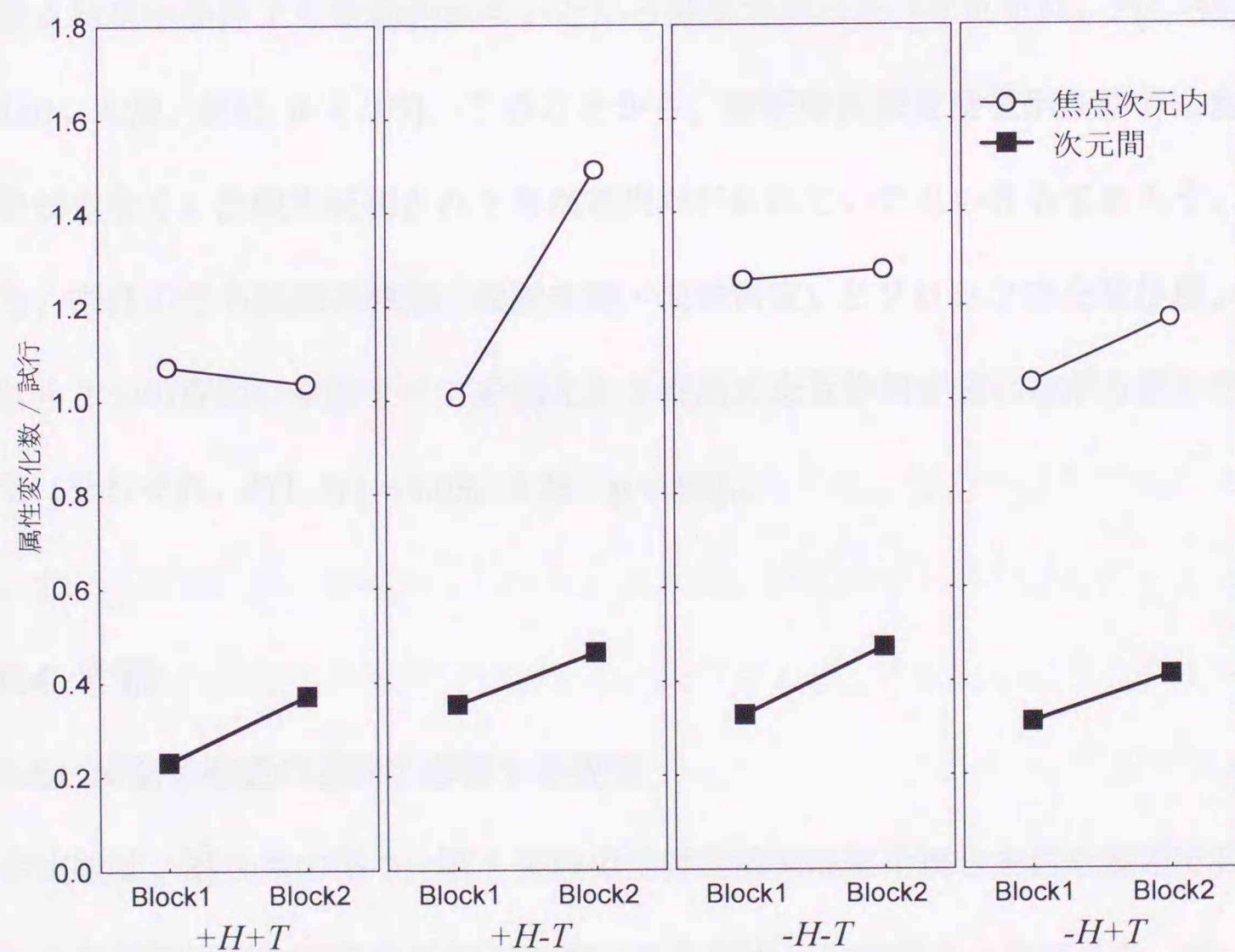


図4-3 各ブロックにおける 1 試行あたりの焦点次元内・次元間変動数

この結果について、条件 (2×2)、ブロック、変動のタイプ (次元内・次元間) を独立変数とした分散分析を実施したところ、ブロックの主効果、変動タイプの主効果、条件、すなわち呈示事例および呈示結果の交互作用が認められた [それぞれ、 $F(1, 36) = 15.72, p < .001$, $F(1, 36) = 278.87, p < .001$, $F(1, 36) = 6.52, p < .05$]。このうち、変動タイプ的主効果については、焦点次元内変動 (1.17) の方が、次元間変動 (.37) よりも有意に多いことが示された。条件に関する交互作用については、仮説支持情報呈示条件のうち、反証型仮説支持呈示条件 (-H-T) の方が確証型仮説支持呈示条件 (+H+T) より変動数が多く、確証型検証が呈示された条件のうち、確証型仮説否定呈示条件 (+H-T) の方が確証型仮説支持呈示条件より変動数が多いという結果であった [それぞれ、 $F(1, 36) = 5.01, 4.73$, 共に $p < .05$]。このことから、確証型仮説支持呈示条件では変動数が少なく、比較的統制された事例選択が行われていたといえるであろう。また、条件のうち結果の種類 (仮説支持・仮説否定) とブロックの交互作用、これら2つの要因に変動タイプを加えた3要因の交互作用が弱いながら認められた [それぞれ、 $F(1, 36) = 4.08, 3.78, p < .06$]。

4.4 考察

4.4.1 反証型検証の選択に影響する要因

本実験は、第3章の第2、第3実験の受動的情報収集手続きを含む課題で見られた反証型検証の選択数の増大に寄与する要因として考えられてきた、(1) 反証型仮説否定情報 (false negative) 獲得による反証型検証の有効性認識の増大、

および、(2) 一般的傾向としての時系列に沿った検証方略のダイナミックなシフト，という2つの説明のどちらがより強力な規定因であるかを検討するために実施された。本実験の結果は、(2) の検証方略のシフトという説明に対して、より強力な証拠を提示しているといえるであろう。

本実験では、規則発見課題の初期段階において、検証事例を実験者側から指定することで、被験者が先行的に獲得する情報の種類を操作した。被験者の半分は、第2試行において、仮説に対する確証型検証を行うよう指示され、残りの被験者は、同試行において、反証型検証を行うよう指示された。それぞれのグループは、さらに、検証の結果として、仮説が支持されるグループと、仮説が否定されるグループに分けられた。もし、反証型仮説否定情報の獲得によって、反証型検証の有効性が認識され、後にその選択頻度が増大するというのであれば、少なくとも反証型仮説否定情報を先行的に呈示された条件では、他の3条件に比べて後の反証型検証の選択頻度が高くなっていなければならない。しかし、本実験の結果では、4条件の間で反証型検証の選択頻度の違いは認められなかった。むしろ、全条件に共通した傾向として、試行が進むにつれて反証型検証の選択頻度が増大しており、これは、確信度の上昇の後見落としを確認するための反証型検証が行われることを予測する、時系列に沿った検証方略の段階的変化という主張に沿った結果であるといえる。このように全体的傾向としては、時系列に沿った検証方略の段階的変化を示していると考えられるが、条件によってやや選択パターンが異なっていることも指摘しておかなくてはならないだろう。確証型仮説支持情報 (+H+T) を先行的に呈示された条件では、

常に確証型検証 (+H) の数が反証型検証 (-H) を上回っているが、その他の条件では前半の試行では両者の数がほぼ等しいか、確証型検証の方がやや上回っているものの、後半の試行では両者の比率が逆転し、反証型検証の方が多くなるという傾向が見られた。この結果から、全体的な傾向としては、確証型を重視した検証→確証型・反証型検証の併用というシフトが存在するものの、2 次的な効果として反証型検証の呈示による有効性認識の強化などの別の要因も影響を与えているということが考えられる。

4.4.2. 仮説検証のダイナミックな変容過程

本実験、第2実験、第3実験で示された結果から、受容方略課題・仮説評価課題において特に顕著に観察された反証型検証の選択数の増大は、まず特定の仮説に対する確信度を高めることで多数の候補の中から一番もっともらしいものを選び出す仮説抽出段階が先行し、後に絞り込まれた1つの仮説ををより詳細に検討し、特に仮説が見落としている条件が存在しないかどうかを確認する仮説洗練段階に移行するという、検証方略のシフトを示していると言えるであろう。このときの事例選択傾向として、前半では確証型検証が多く、後半になるに従って反証型検証の選択比率が高まるという事例選択パターンが観察されると考えられる。

確証型・反証型検証という観点からみた事例選択傾向の他にも、本実験では、時系列に沿った仮説検証方略のシフトを支持するデータが得られている。本実験では、被験者の事例選択が、仮説中で言及されている関連次元と、言及

されていない無関連次元のどちらに注目したものであるかについて、関連次元のみで属性値の変更が行われた事例選択と、無関連次元での属性値の変更も行われた事例選択とに分類した。関連次元は概念の規定に参与する次元であり、本実験においては仮説中で言及されている次元を指す。従って、関連次元のみに注目した事例選択は、被験者がターゲット規則を構成する可能性が高いと考える次元に対する確証型検証の試みであるといえ、これは、仮説の蓋然性を高めるために行われているのではないかと推測される。一方、無関連次元は概念の規定に参与しない次元を指し、本実験においては仮説において言及されていない次元である。従って、無関連次元に注目した事例選択は、仮説に見落としがないかどうかを確認し、仮説を洗練されたものにしようとする試みであると解釈することができる。このように考えるならば、関連次元のみに注目した事例選択は、試行の序盤で多く行われ、無関連次元を含む事例選択は、試行の終盤に向かうにしたがって頻度が多くなることが予想される。本実験の結果から、全般的には関連次元のみに関する事例選択が多く、特に前半の試行でその傾向が強いものの、無関連次元の属性値の変更を含む事例選択が後半の試行でより多くなることが示された。本実験の結果は、この予測を裏付けているといえる。

また、本実験では、試行系列で注目されている焦点次元内での属性値の変動、および焦点次元以外の次元での属性値の変動を、それぞれ、焦点次元内変動・次元間変動とよび、各試行における変動の数を集計した。その結果、全体的傾向として焦点次元内変動が多く、被験者は比較的統制された実験を行っていたといえる。また、確証型仮説支持情報を呈示される条件 (+H+T) では、

他の条件に比べて、全体的な属性値の変動数が少ないという結果が得られた。さらに、非ターゲット事例が呈示される 2 条件 (+H-T, -H-T) では、焦点次元内変動が多いという傾向があるようにも見える。これらの条件においては、より統制された実験が行われやすかったと考えることもできる。しかし、本実験課題は 3 次元からなるものであり、被験者の仮説も 1 次元のみに関するものから 3 次元全てに言及したものまでさまざまであった。従って、初期仮説の違いにより被験者が注目する次元の数に違いが生じる可能性があり、単純に次元変動の数のみに基づいて議論するのは適当ではないかもしれない。

次元の変動数という測度をどのように解釈すべきであるかについては、疑問の余地を残すものの、それ以外の結果、特に、検証過程の後半になって反証型検証の選択数が増大するという傾向は、仮説検証が時系列に沿って変容するダイナミックなプロセスであることを示していると言える。特に、一般的な仮説検証方略として、まず特定の仮説に対する蓋然性を高める段階から、続いてその仮説をより洗練されたものにするために見落としの確認などを行う段階へと向かうシフトが存在することが示唆される。

一方で、第 3 章で指摘した通り、第 2 実験の正事例系列条件と第 3 実験の正解到達条件との間で反証型検証の選択率を比較したところ、全般的に後者の方が反証型検証の選択率が高いという結果が見られたことから考えると、反証型仮説否定情報の呈示による反証型検証に対する有効性認識の強化という要因が事例選択に与える影響も無視し得ない。従って、被験者が検証事例を選択する場合、本実験で示されたような一般的傾向としての選択傾向のシフトに加え

て、反証型検証に対する認識が強化されるかどうかという要因がオーバーライドする形で影響を与えるという可能性が考えられる。すなわち、検証事例の選択とは、単一の要因の影響によってのみ説明されるのではなく、複数の要因が複雑に絡み合って働くプロセスであると言えるであろう。

4.5 第4章のまとめ

本章で報告した第4実験の結果から、第3章で報告した第2、第3実験において、課題の中盤から終盤にかけて反証型検証の選択数が増加したのは、仮説に対する確信度がある程度高くなった後で、見落としの確認を行うという検証方略のシフトによるところが大きいと言えるであろう。すなわち、人は仮説に対する確信度が低い初期段階と、ある程度仮説に自信が持てるようになった中期ないし後期段階とで、選択する検証事例の種類を変化させ、前者では確証型検証を重視し、後者では反証型検証も併用する傾向を持っていると考えられる。

また、第3章で述べたように、第2実験の正事例系列条件と第3実験の正解到達条件とを比較したところ、特に後者において反証型検証の選択頻度が高いという結果が得られていることから考えると、反証型仮説否定が呈示されることによる反証型検証に対する有効性認識の強化という要因も、検証方略の選択に影響を与えていると考えられる。すなわち、検証方略の選択は、一般的傾向として初期段階と後期段階で変化するだけでなく、複数の要因が絡み合った結果として決定されるということができるとであろう。

次章では、検証方略の選択が時系列に沿ってダイナミックに変化するという

第5章 仮説評価課題における評価者独自の推論過程

本章では、これまでの実験結果から示された、時系列に沿った検証方略のダイナミックな変容が、より一般的な状況においても当てはまるかどうかについて検討する。特に、これまでの仮説検証研究で用いられてきた2-4-6問題(Wason, 1960)という状況の下で、時系列に沿った段階的な方略の変化が見られるかどうかを調べることは有意義であろう。また、第3章の第3実験で得られた結果から示唆された、仮説評価課題における評価者独自の推論過程の可能性についても検討する。以上のような目的のため、本章の第5実験では、2-4-6問題を仮説評価課題に転用した課題を用いる。

5.1 第5実験 - 参照学習過程としての仮説評価課題

第3章の第3実験で用いられた仮説評価課題では、獲得された結果が規則発見に対してどれほど有効かの評価に関して、被験者、すなわち評価者が呈示された仮説に対してもつ見解が影響を与えているという可能性が示された。これは、評価者が、呈示された情報を基に独自の推論過程を進行させていることを示唆している。これまでの仮説評価研究においては、評価者が、呈示された仮説、および検証方略を客観的に評価している、あるいは評価することができるということが前提とされてきた。その上で、評価者は反証型検証、あるいは仮説否定情報の有効性を正しく認識できることが示されてきた (Farris & Revlin, 1989; Kareev & Halberstadt, 1993; Tschirgi, 1980)。一方、第3章の第3実験では、特に

得られた結果の情動的価値の評価に関して、被験者が呈示された仮説を正しいと考えているかどうかという要因の効果が見られた。特に、被験者は、呈示された仮説を正しいと思っている時にその仮説が否定される場合、及び呈示された仮説を誤っていると思っているときにその仮説が支持される場合に、それぞれの情報を有効で価値がある情報であると考えていることが示された。この結果は、被験者が呈示された仮説に対して持つ見解が否定されるかどうか、情報の価値評価に決定的な影響を及ぼしていると言いかえることができ、被験者、すなわち評価者が、課題遂行中に与えられた情報に基づいて自ら仮説を立てるなどの、独自の推論過程を進行させている可能性を示唆するものである。本実験では、この可能性をを検討することを第一の目的とする。また、課題には、伝統的な規則発見課題として用いられてきた2-4-6問題を仮説評価課題に修正した課題を用い、検証方略の段階的変化傾向の一般性についても併せて検討する。

5.2 方法

課題および条件

2-4-6問題 (Wason, 1960) を題材とした仮説評価課題を作成した。第3章の第3実験とは異なり、実際の被験者の試行系列ではなく、架空の試行系列が用意されたが、被験者に対しては、ある特定の人物が実際に実施した試行系列であるという教示を与えた。

用意された系列は2つで、一つは正しい規則である「上昇系列」に到達す

表5-1 呈示情報系列

試行	規則発見 (<i>correct</i>)		下位集合型誤り (<i>embedded</i>)	
	仮説	事例	仮説	事例
1	連続する偶数	$+H+T$	連続する偶数	$+H+T$
2	同上	$+H+T$	同上	$+H+T$
3	同上	$-H+T$	同上	$+H+T$
4	2 ずつ増える数	$+H+T$	同上	$+H+T$
5	同上	$-H+T$	同上	$+H+T$
6	等差数列	$+H+T$	同上	$+H+T$
7	同上	$+H-T$	同上	$+H+T$
8	公差が正の等差数列	$-H-T$	同上	$+H+T$
9	同上	$-H+T$	同上	$+H+T$
10	上昇系列	$+H+T$	同上	$+H+T$

る過程をシミュレートした系列，もう一つはいわば確証バイアスに陥り，下位集合型の誤りを犯した過程をシミュレートした系列である．それぞれの系列情報を与えられる条件を規則発見 (*correct*) 条件，下位集合型誤り (*embedded*) 条件と呼ぶ．それぞれの条件で呈示された情報系列を表 5-1に示す．表に示される通り，規則発見条件においては，確証型仮説支持情報が 5，確証型仮説否定情報が 1，反証型仮説支持情報が 1，反証型仮説否定情報が 3 というように，数にばらつきはあるものの，全ての種類の情報が呈示される，確証・反証併用型の方略を用いた反応系列である．一方，下位集合型誤り条件では，全ての事例が確証型仮説支持情報であり，典型的な確証バイアスを示す反応系列になっている．

実験は質問紙形式で実施され，被験者は各試行において，呈示された仮説が正しいと思うかどうか，呈示された仮説よりも可能性が高そうな別の仮説は考えられるかどうか，どのような検証事例が適切だと思うか，呈示された検証事

例がどの程度適切だと思うか、呈示された結果がどの程度有効なデータだと思うかについて評価するよう求められた。また、それまでの試行で呈示された情報については自由に参照することができた。

手続き

被験者は、割り当てられた条件に応じた質問紙を受け取った。2つの条件は、呈示される情報系列が異なるが、カバーストーリーや、その他の手続きについては、全く同じ記述が与えられた。被験者はまず、2-4-6問題に関する簡単な記述を与えられた後、他者が課題を解決した過程について評価するという趣旨の教示を読んだ。被験者は各試行に先立ち、他者が考えている仮説と、その時点でわかっている事例と結果の組み合わせを呈示された。事例と結果の組み合わせについては、第1試行では初期事例である「2 4 6」のみであるが、以降の試行では、事例と結果の組み合わせが随時追加、呈示された。例えば、第5試行では初期事例の「2 4 6」に加え、第1～第4試行で呈示された事例と結果の組み合わせが4つ、併せて5つの情報が呈示された。これら2つの情報が呈示された後、(1) 被験者は、仮説が正しいと思うかどうかを「正しい」、ないし「間違っている」のいずれかを選択することで評価し、また、呈示された仮説よりも可能性が高いと思われる仮説がある場合は、それを記述した。続いて被験者は、(2) 検証事例として適切であると思われる数字列、すなわち検証事例を1つ挙げるよう求められた。この事例は、第3実験の仮説評価課題と同様に、実際に用いられる検証事例としてではなく、あくまで希望とし

て選択するよう求められた。次に、(3) 予め用意された検証事例が呈示され、この検証事例の選択がどの程度適切なものであるかを 0 ～ 10 点で評価した。最後に、(4) 呈示された事例がターゲット集合に属するかどうか被告知られ、この結果が最終的に正しい規則を発見するためにどの程度役に立つものであるかを 0 ～ 10 点で評価した。以上の (1) ～ (4) までの手続きが、用意された 10 試行分繰り返された。

被験者

被験者は、札幌市内の専門学校 3 年次学生 83 名であり、無作為に規則発見条件と下位集合型誤り条件のいずれかに割り当てられた。規則発見条件には 42 名、下位集合型誤り条件には 41 名の被験者が参加した。

5.3 結果

仮説に対する評価

仮説に対しては、まず各試行において提示された仮説が正しいと思うかどうかの評価がなされた。規則発見条件においては、10 試行中 9.4 試行で呈示された仮説が正しいと評価されており、仮説が誤っていると評価されたのは .6 試行にとどまっていた。また、下位集合型誤り条件においても、10 試行中 9.1 試行で仮説が正しいと評価され、誤っていると評価されたのは .9 試行であった。以上の結果から、呈示された仮説は、ほとんどの場合において正しいと考えられているといえるであろう。

次に、各試行において、被験者が別の可能性を考えたかどうかについて検討する。呈示された仮説以外に別の可能性があるとして被験者が指摘した試行数を、それぞれの被験者毎に計算したところ、規則発見条件では平均して .6 試行、下位集合型誤り条件では、1.2 試行という結果であった。被験者は、特に別の可能性を考えていないようである。両条件間の差も有意ではなかった [$t(81) = 1.18, p > .10$]。

事例選択傾向

次に、これまでの仮説評価課題と同様に、検証事例の選択傾向について分析する。ここでの選択傾向とは、第 3 章の第 2, 第 3 実験と同様に、実際に検証に用いられた事例ではなく、あくまで希望としての選択である。被験者が毎試行において選択を希望した事例を、呈示された仮説に対する確証事例 (+H) か反証事例 (-H) かによって分類し、さらに前半 5 試行のブロック 1, 後半 5 試行のブロック 2 に分けて、それぞれの事例の選択数を表したのが図 5-1 である。条件、ブロック、事例の種類 (+H, -H) を独立変数とした分散分析を行ったところ、事例の種類の主効果と 3 要因の交互作用が認められた [それぞれ, $F(1, 75) = 174.3, p < .001$, $F(1, 75) = 5.9, p < .05$]。事例の主効果は、両条件の全ブロックを通じて、確証事例の選択が反証事例より多いためであるといえる。3 要因の交互作用について単純効果の検定を行ったところ、規則発見条件の確証事例、および反証事例のそれぞれについて、ブロックの差が認められた [それぞれ, $F(1, 150) = 8.47, p < .01$, $F(1, 150) = 4.32, p < .05$]。確証事例については、

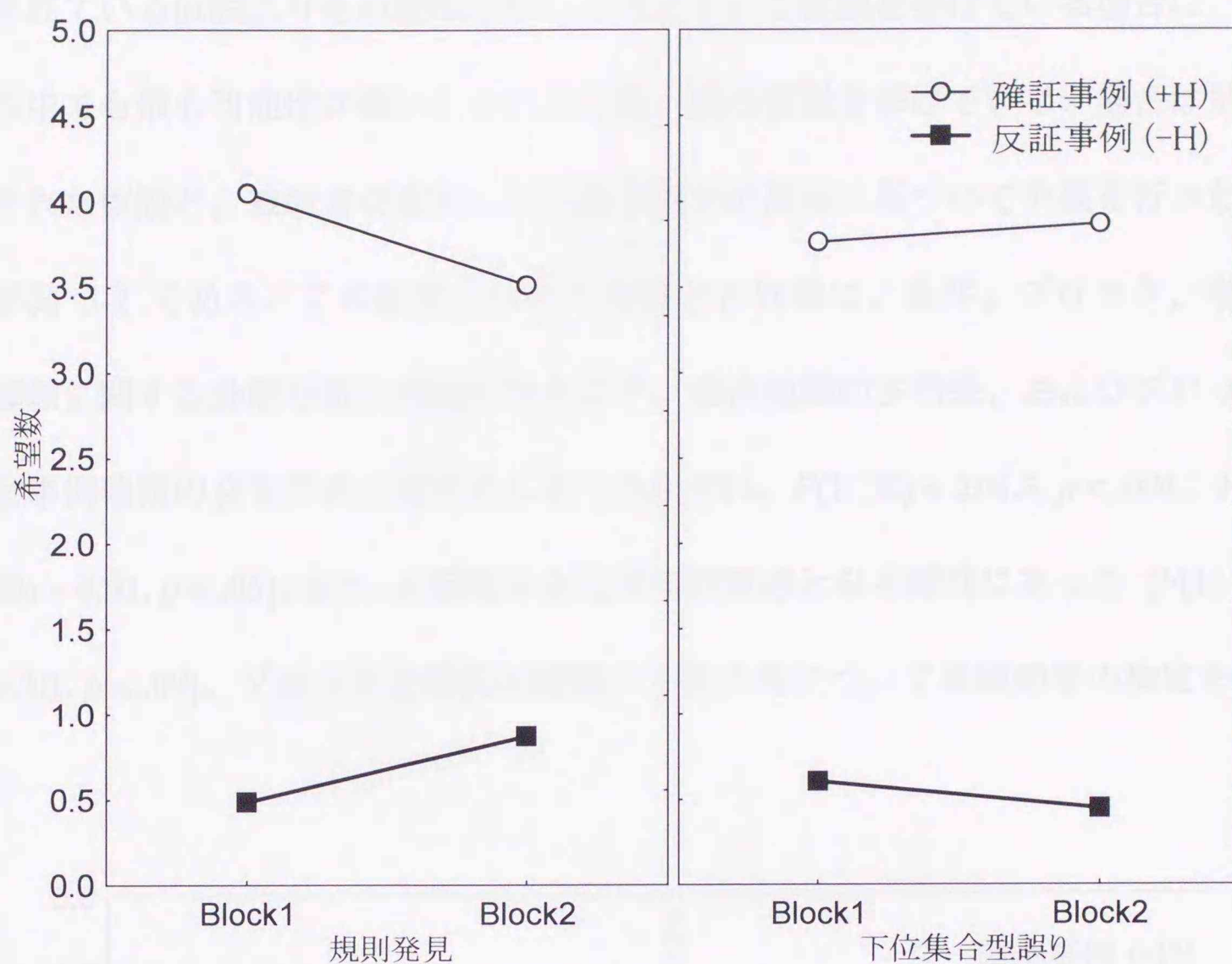


図5-1 条件・ブロック別に見た事例選択傾向

ブロック 1 (4.1) の方がブロック 2 (3.5) よりも多いが、反証事例はブロック 1 (.5) よりもブロック 2 (.9) の方が多く選択された。また、両条件の全てのブロックにおいて、確認事例の方が反証事例よりも多く選択されていた (全て $p < .001$)。

次に、呈示された仮説とは異なる別の可能性を挙げた試行が少数ながら存在することを考慮に入れ、被験者の選択した事例と、その時点で被験者がもっとも可能性が高いと考えている仮説との関係に注目する。もし被験者が、呈示

されている仮説よりも可能性が高いと考えている仮説を挙げている場合は、その中でも最も可能性が高いとされる仮説、別の仮説を挙げていない場合は呈示された仮説と、被験者が選択した検証事例との関係に基づいて分類を行ったのが図 5-2 である。この結果について先ほどと同様に、条件、ブロック、事例種類に関する分散分析を実施したところ、事例種類の主効果、およびブロックと事例種類の交互作用が認められた [それぞれ、 $F(1, 70) = 294.5, p < .001$, $F(1, 70) = 4.91, p < .05$]。また、3 要因の交互作用が有意となる傾向にあった [$F(1, 70) = 3.1, p < .09$]。ブロックと事例の種類の交互作用について単純効果の検定を行

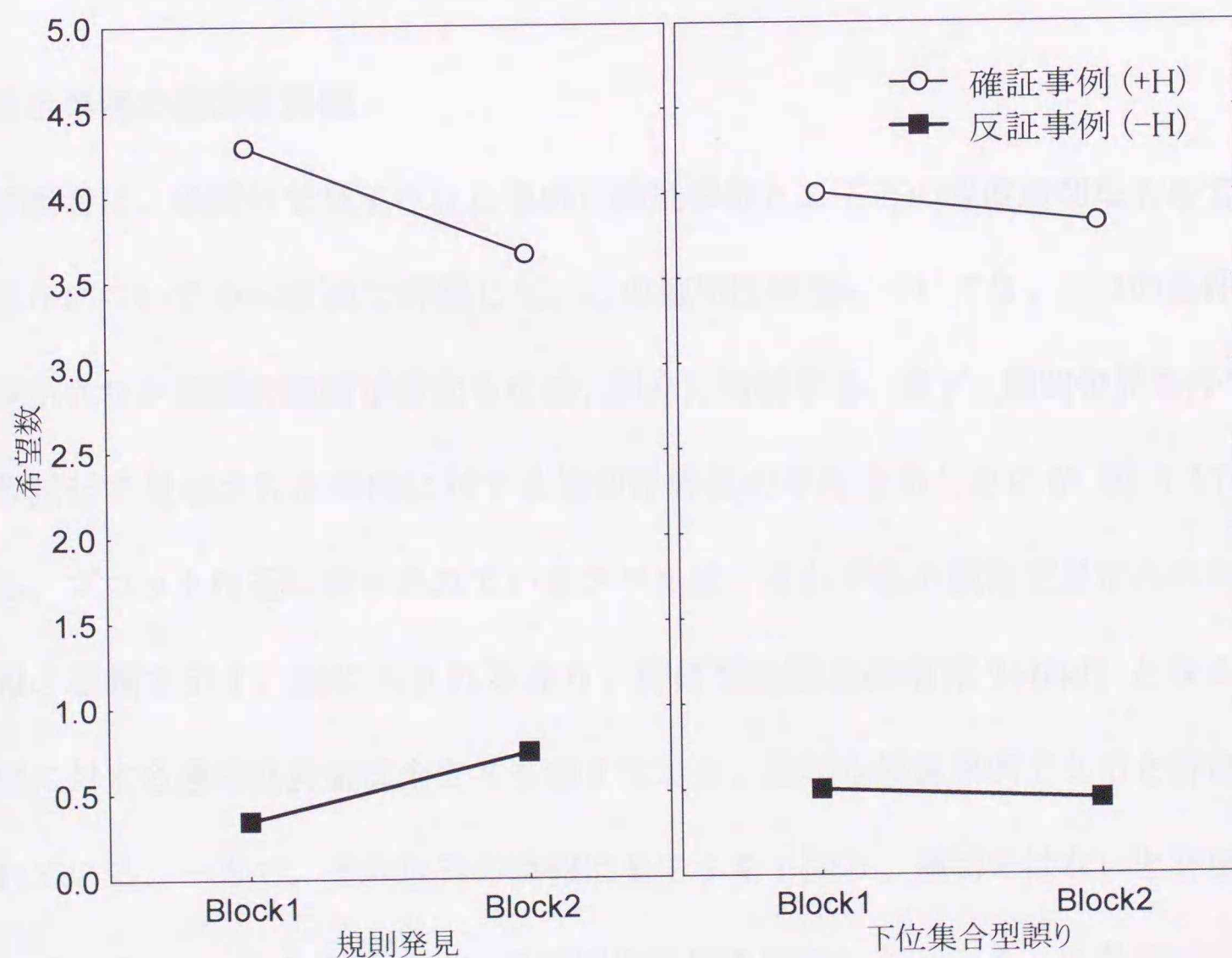


図5-2 被験者が最も可能性が高いと考える仮説と選択事例の関係

ったところ、確証事例の選択数についてブロックの単純主効果が見られた [$F(1, 140) = 8.34, p < .01$]. また、ブロック 1, 2 共に確証事例の方が多く選択されていた [それぞれ、 $F(1, 140) = 244.8, 175.4$, 共に $p < .001$]. 3 要因の交互作用については、5%水準では有意であるとはいえないものの、規則発見条件の確証事例、反証事例のそれぞれにおいてブロックの差が認められた [それぞれ、 $F(1, 140) = 10.19, p < .01$, $F(1, 140) = 3.98, p < .05$]. この結果は、仮想生物課題を使った第 3 実験の結果とは異なり、反証型仮説否定の獲得による反証型検証の有効性の認識という説明に支持を与えているようである。この点については、後に考察する。

呈示事例の適切性評価

被験者は、毎試行で呈示された事例が検証事例としてどの程度適切なものであるかについて 0~10 点で評価した。この適切性評価については、2 つの条件で呈示された事例の種類が異なるため、別々に検討する。まず、規則発見条件で、各試行で呈示された事例に対する適切性評価の平均を示したのが 図 5-3 である。プロット付近に表示されているラベルは、それぞれの試行で呈示された情報の種類を示す。図に示される通り、確証型仮説支持情報 (+H+T) となる事例に対する適切性評価は全て 5 を越えており、適切な検証事例であると評価されている。一方で、それ以外の情報は全て 5 を下回り、適切ではないと評価されている。10 個の事例について適切性評価を比較したところ、主効果が認められた [$F(9, 342) = 13.11, p < .001$]. 多重比較の結果、第 1, 第 2, 第 4, 第

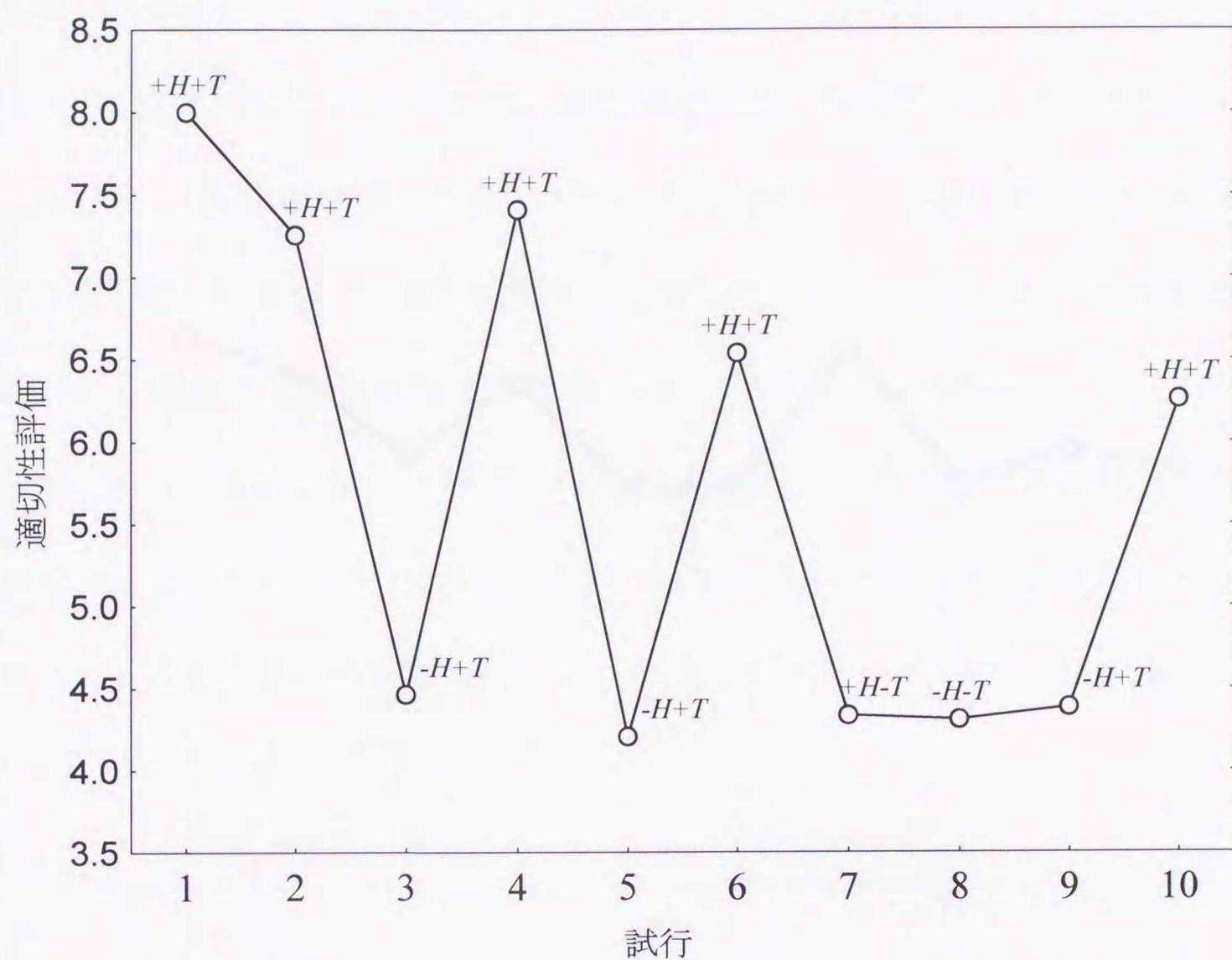


図5-3 規則発見条件における適切性評価

6, 第 10 試行で呈示された確証型仮説支持情報と, 第 3, 5, 7, 8, 9 試行で呈示されたそれ以外の種類の情報との間で評価に差が認められた. 一方, 下位集合型誤り条件における評定値の平均をあらわしているのが図 5-4である. この条件で呈示されたのは全て確証型仮説支持情報となった事例であるが, 事例間で評定の差が認められた [$F(9, 315) = 2.27, p < .05$]. 特に, 第 1, 第 7 試行で呈示された事例と, 第 5, 第 6, 第 8 試行で呈示された事例との間に差があるようである. しかし, 全般的に言って, 全ての呈示事例に対する適切性評価が,

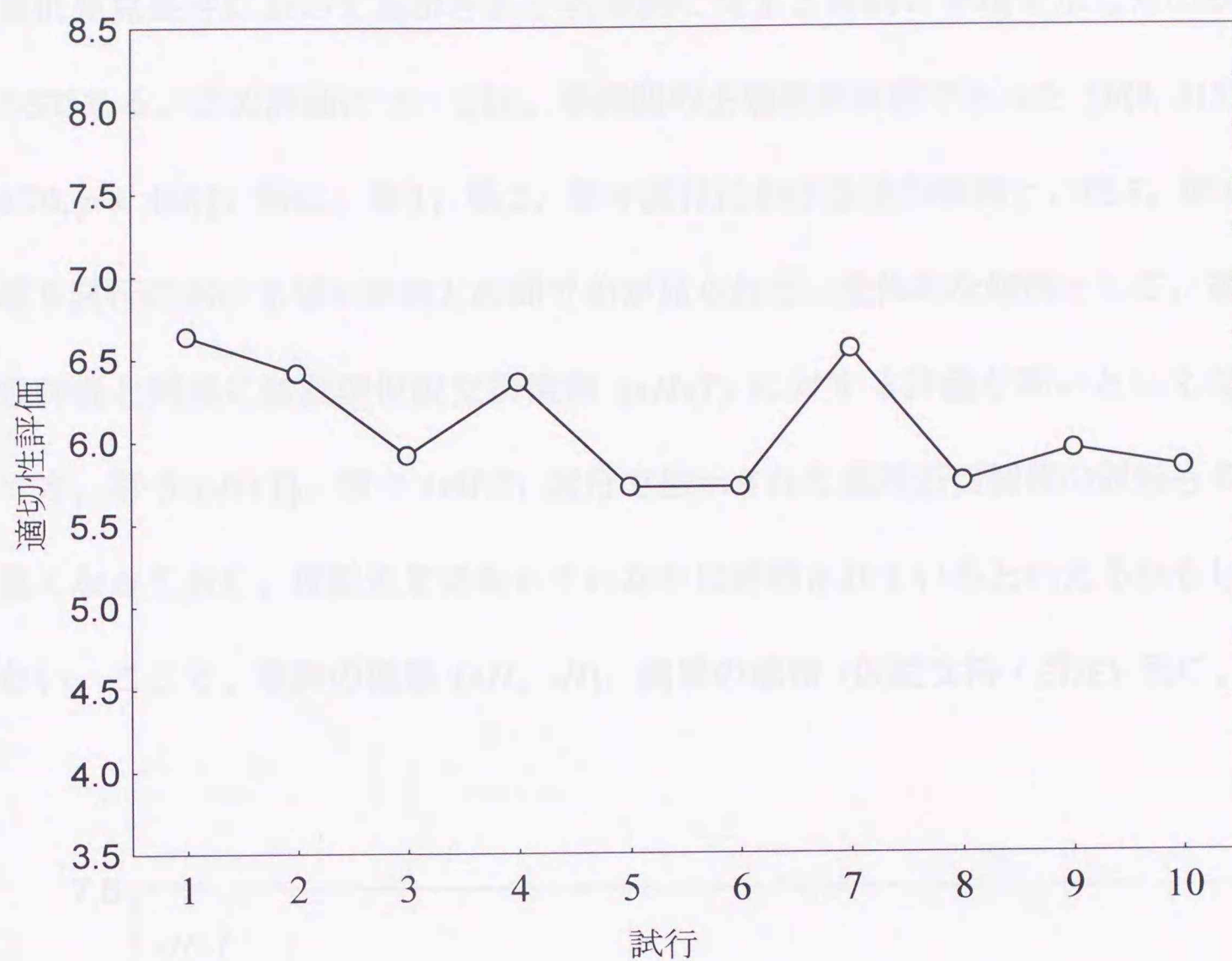


図5-4 下位集合型誤り条件における適切性評価

平均として5を越えており、概ね適切な事例選択であると評価されているといえる。また、緩やかではあるが右下がりの直線的傾向を示しており、試行が進むにつれて適切性評価が低くなっているようである。

結果の有効性評価

前述の適切性評価と同様、被験者は、呈示された結果の有効性評価についても0～10点で評価しており、ここでもまた2つの条件について別々に分析する。

規則発見条件において呈示された各事例に対する評価の平均を示したのが図5-5である。この評価については、事例間の主効果が有意であった [$F(9, 315) = 4.70, p < .001$]。特に、第1, 第2, 第4試行における呈示事例と、第5, 第8, 第9試行における呈示事例との間で差が見られた。全体的な傾向として、適切性評価と同様に確証型仮説支持情報 (+H+T) に対する評価が高いといえる。一方、第3 (-H+T), 第7 (+H-T) 試行で呈示された仮説否定情報の評価もやや高くなっており、仮説否定情報もそれなりに評価されているといえるかもしれない。ここで、事例の種類 (+H, -H), 結果の種類 (仮説支持・否定) 別に,

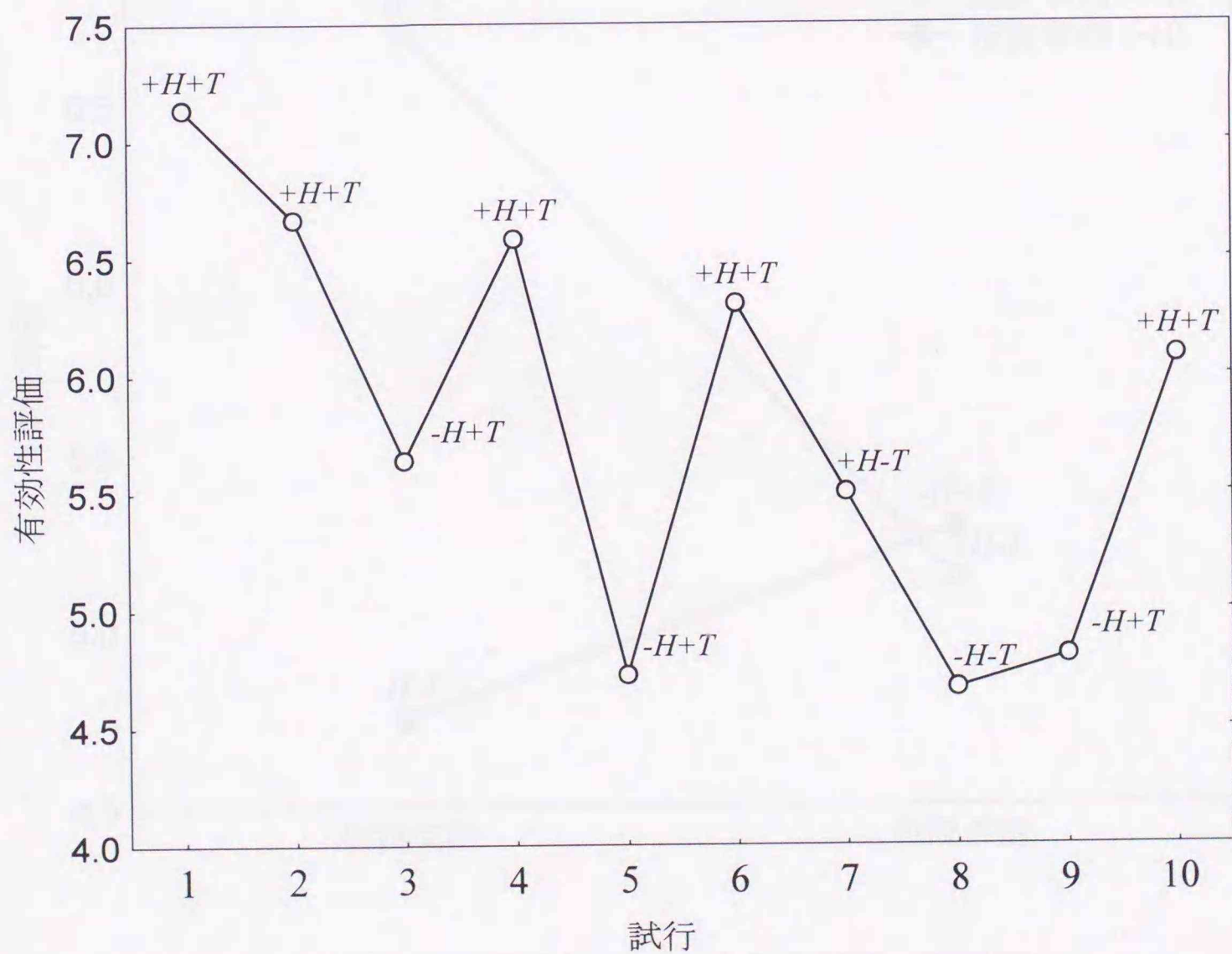


図5-5 規則発見条件における有効性評価

それぞれの種類の情報が、どの程度有効だと評価されたかを示したのが図 5-6 である。2 要因の分散分析の結果、事例の主効果、および 2 要因の交互作用が認められた [それぞれ、 $F(1, 41) = 10.77, 10.73$, 共に、 $p < .01$]。交互作用について単純効果の検定をしたところ、仮説支持情報のうち、確証型仮説支持 (+H+T) の方が、反証型仮説支持 (-H-T) よりも高く評価されていた [$F(1, 82) = 21.37, p < .001$]。また、確証型仮説支持情報の方が、確証型仮説否定情報(+H-T) よりも高く評価されていた [$F(1, 82) = 8.93, p < .01$]。以上のように、確証型

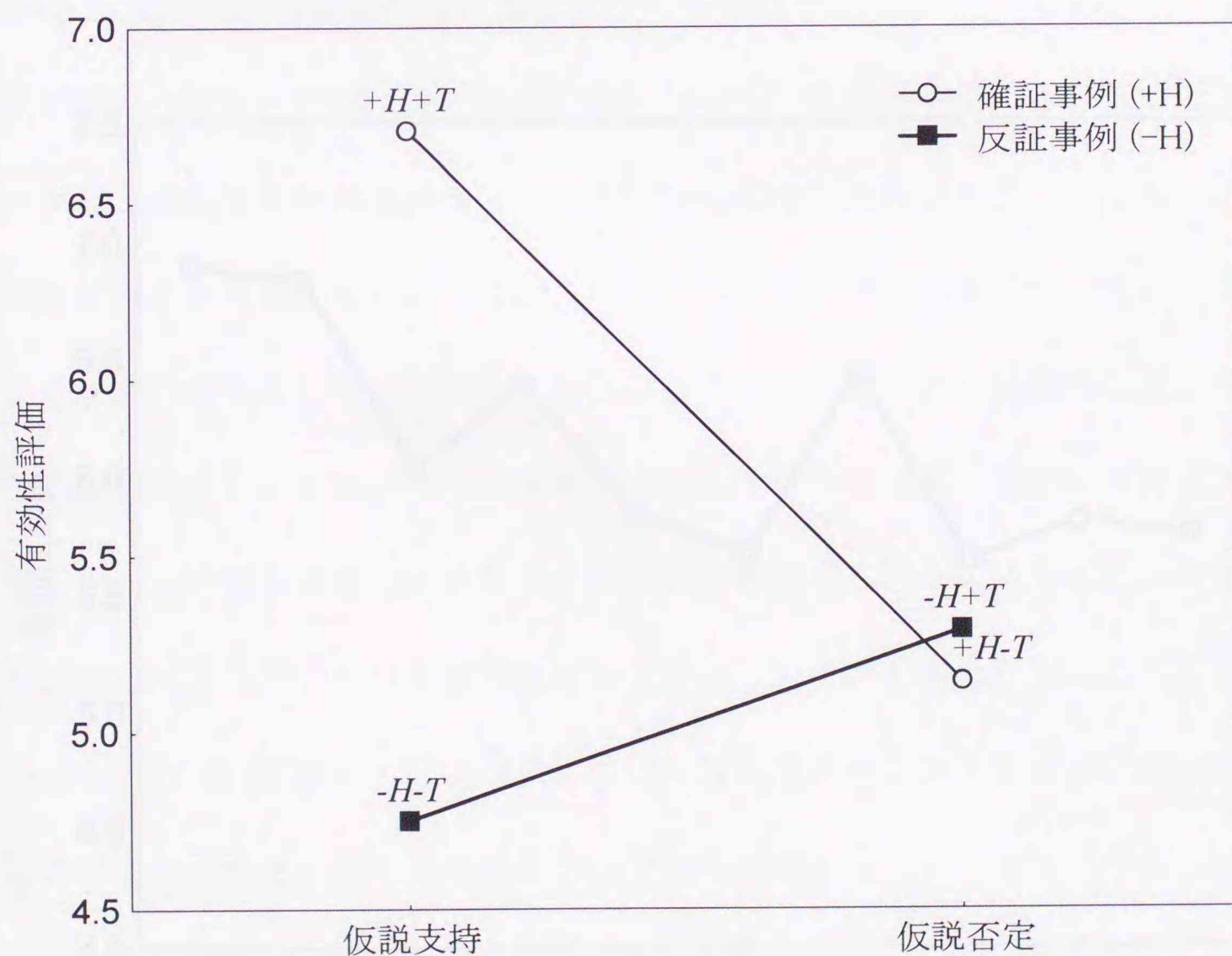


図5-6 規則発見条件における有効性評価 (種類別)

仮説支持情報が特に規則発見に対して有効であると評価されていた。また、結果の種類について差が見られないことから、仮説否定情報もそれなりに評価されているといえ、第3実験や従来の研究からの知見と一致している (c.f. Kareev & Halberstadt, 1993)。

次に、下位集合型誤り条件について、同様に有効性評価の平均を算出して示したのが図 5-7である。この条件で呈示されたのは全て確証型仮説支持情報 (+H+T) であるが、事例間の主効果が有意となった [$F(9, 315) = 3.57, p < .001$]。特に第1, 第2試行における呈示事例と、第6, 第8, 第10試行における呈示

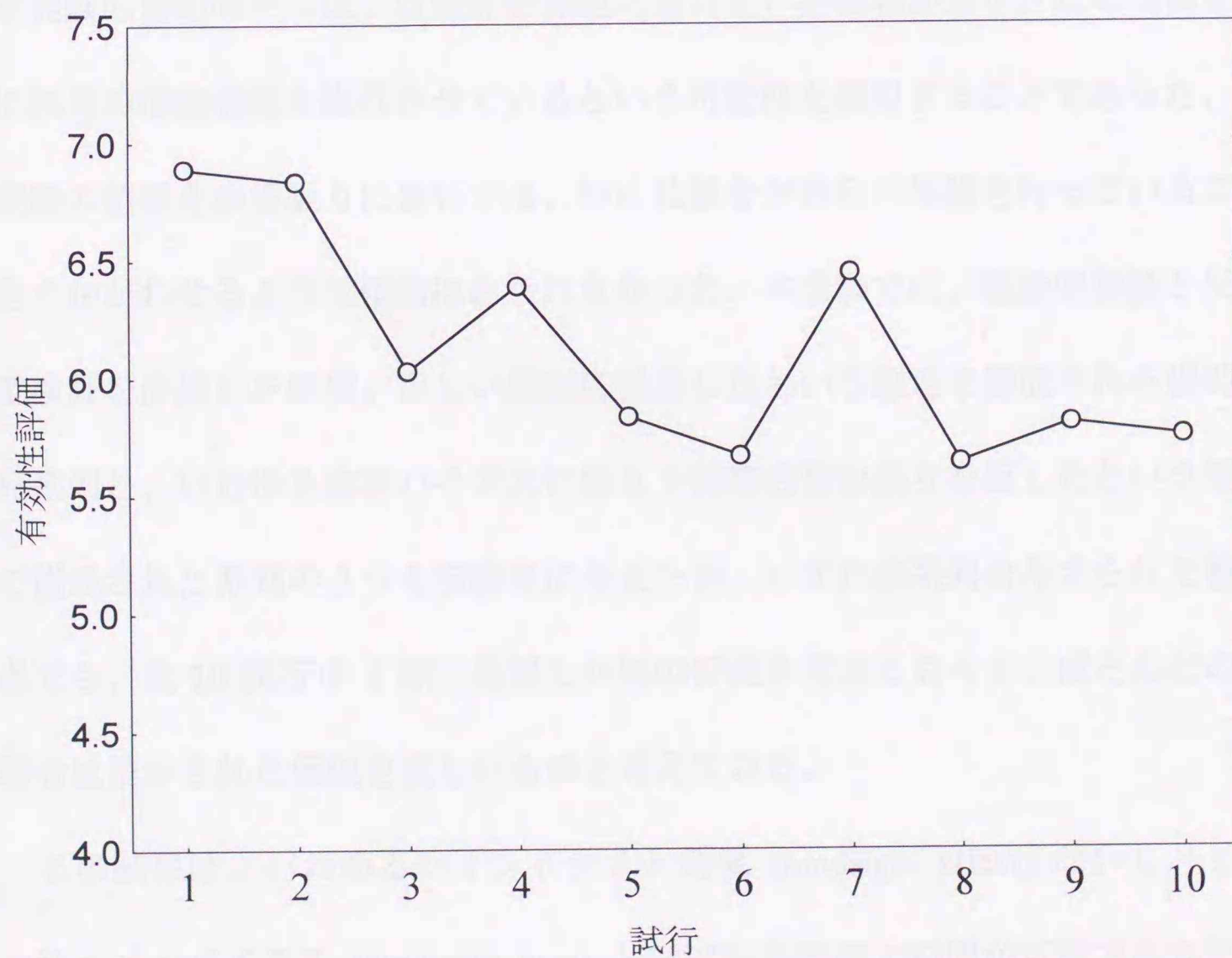


図5-7 下位集合型誤り条件における有効性評価

事例との間で評価値に差が認められた。また、全体的な傾向として、右下がりの直線的傾向を示しており、確証型仮説支持情報は終盤に向かうほど有効ではないと評価されていることがわかる。同様の傾向は、規則発見条件においても観察される。規則発見条件において確証型仮説支持情報が与えられたのは、第 1、第 2、第 4、第 6、第 10 試行であるが、徐々にその有効性評価は低くなっているといえる。

5.4 考察

5.4.1 仮説評価課題における評価者独自の推論過程

本実験の目的の一つは、仮説評価課題において、評価者が呈示される情報を基に独自の推論過程を進行させているという可能性を検討することであった。本実験の結果をみる限りにおいては、特に被験者が独自の推論を行っていることをうかがわせるような傾向はみられなかった。本実験では、確証型検証と反証型検証を併用した結果、正しい規則に到達したという想定で構成された規則発見系列と、いわゆる確証バイアスに陥り下位集合型の誤りを犯したという想定で構成された系列の 2 つを被験者に与えたが、いずれの系列を与えられた被験者でも、全 10 試行中 1 試行程度しか別の仮説を考えておらず、ほとんどの被験者は呈示された仮説を正しいものと考えていた。

この結果は、いわゆるハインドサイト効果 (hindsight effect) ないしハインドサイト・バイアス (hindsight bias) と呼ばれる現象で説明が可能であると思われる。ハインドサイト効果とは、ある情報を与えられることによって、その

情報を得る前の心的状態に戻ることができなくなることを指し、因果推論や自己の判断に対する確信度評価などにおいて観察されることが知られている (Fischhoff, 1975, 1977; Hawkins & Hastie, 1990; Slovic & Fischhoff, 1977; Wasserman, Lempart, & Hastie, 1991). 特に、因果推論状況においては、結末を知らされた後で、それを知らなかったと仮定して、どのような結末となるかを考えるといったような場合にハインドサイト効果が働き、教えられた結末以外の選択肢を結末として考えることが困難になるという傾向、すなわちハインドサイト・バイアスとして観察される。ハインドサイト・バイアスについては、漸進的決定論 (creeping determinism; Fischhoff, 1975) という説明が提唱されており、その基本的主張は、結末に関する情報が与えられることにより、自動的かつ無意識的に、その選択肢とそれを支持する証拠とを強く結びつけるような知識の再構成が行われ、情報が与えられる前の知識構造に基づいた判断・推論ができなくなるということにある。本実験の場合は、仮説が呈示されることにより、この仮説とそれを支持する証拠とを結びつけるような知識の再構成が行われ、それ以外の可能性を考慮することが困難になったと考えられる。以上のことから、被験者による情報の評価は、呈示された仮説との関係で考えられるべきものであって、第3実験で述べたような、被験者自身の見解が支持されるかどうかという要因は無関係であるといえるであろう。

しかし、一方で、本実験と第3実験では異なった課題状況を用いていることにも注意する必要があると思われる。第3実験で用いられた仮説評価課題は、仮想生物課題という現実的な背景をもつ課題における被験者の反応に基づいて

作成されている一方で、本実験課題は、2-4-6問題という人工的な問題空間によって構成される課題を背景としている。前者の課題では、実際には架空の生物が好む環境という恣意的な規則を発見する課題であったものの、生物学に関して評価者が持つ常識的な知識が発動されやすく、評価者自身が独自の推論を行っていたかも知れない。一方、後者の課題は、そのような常識的知識が発動される余地のない完全に人工的な課題状況であったため、評価者にとって独自に推論を行うのが困難であったという可能性は考えられる。このように、被験者が当該の課題状況に関して持っている知識、すなわち領域固有的知識の影響も無視し得ないものである。

5.4.2 評価課題における事例選択傾向 — 時系列に沿った検証方略のダイナミックな変化

本実験の第2の目的は、第2、第3、第4実験に見られたように、仮説検証方略が段階をふんで変化するという現象が一般的な傾向であるかどうかを、これまでの研究で伝統的に用いられてきた2-4-6問題の評価版課題において検討することであった。

全体的な傾向としては、確証型検証が反証型検証よりも多く希望されるという傾向が見られ、2-4-6問題の評価版課題を用いた先行研究と対応した結果を示している (Kareev & Halberstadt, 1993)。また、規則発見条件においては、前半5試行に比べて、後半5試行の方で反証型検証が希望される頻度がやや高くなっているが、下位集合型誤り条件では、そのような傾向は見られなか

った。この結果は、反証型仮説否定情報の呈示による反証型検証の有効性に対する認識の強化という主張に沿っているといえる。ここで、この傾向についてさらに詳しく分析するために、各試行においてどの程度の被験者が確証型検証を希望したかに関する比率を算出した (図 5-8)。図に示される通り、規則発見条件においては、第 8、第 9 試行における確証型検証の選択頻度が、その他の試行に比べて低くなっているといえる。これは、反証型検証の選択頻度が高く

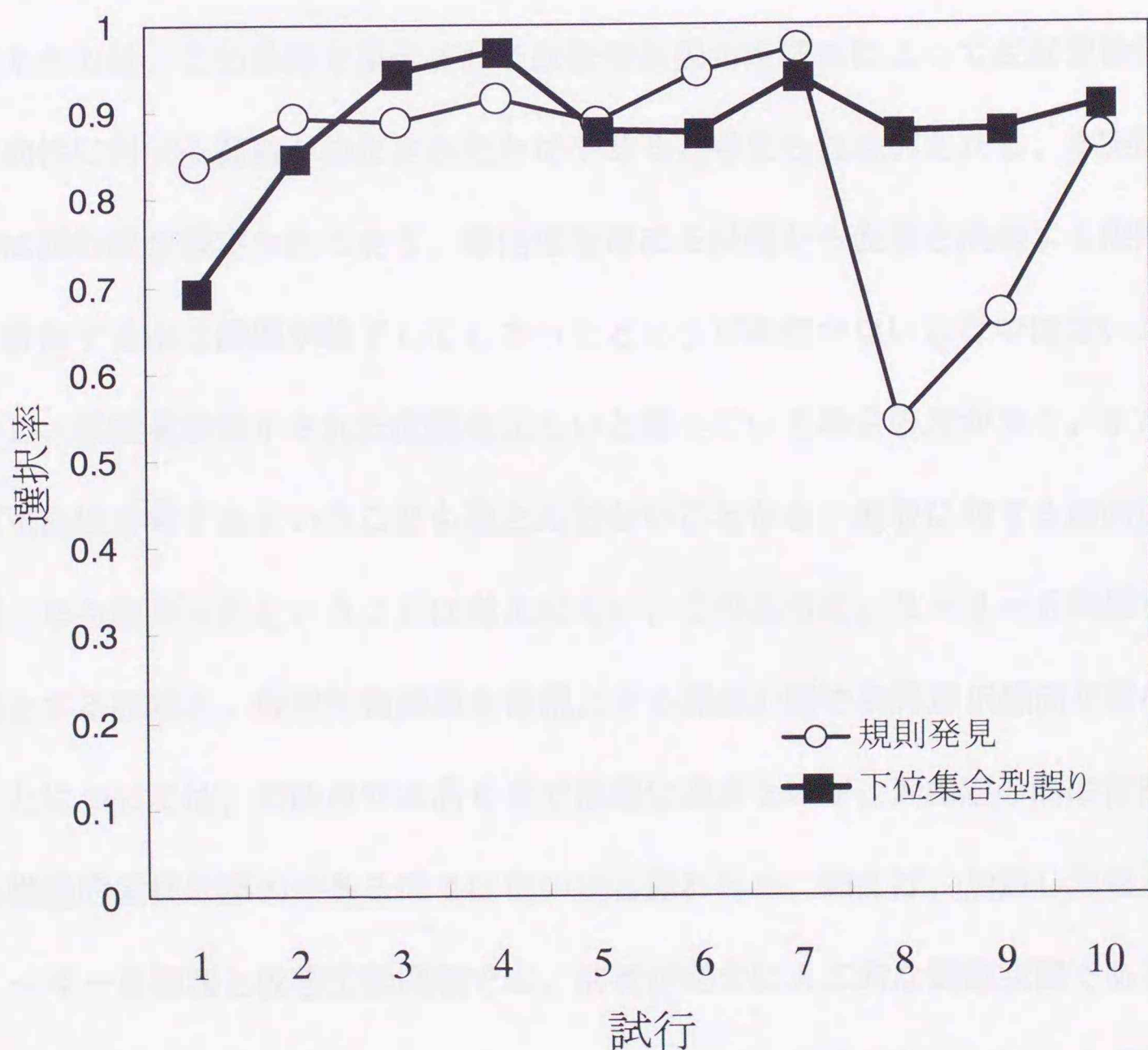


図5-8 各試行における確証型検証の選択率

なっていることを意味する。特に第 8 試行においては、確証型検証を希望した被験者と反証型検証を希望した被験者がほぼ同数であった。一方、下位集合型誤り条件では、第 1 試行において 3 分の 1 ほどの被験者が反証型検証の選択を希望していたが、それ以外の試行では、大半の被験者が確証型検証を希望していた。第 5, 第 6, 第 8, 第 9 試行で、やや反証型検証を希望する被験者の数が多くなっているようであるが、その比率は 15% 程度であった。

このように、本実験の被験者は、概ね確証型検証を多用する傾向があり、規則発見条件の被験者が後半のブロック 2 において反証型検証を選択する頻度が増えたのは、この条件で呈示される反証型仮説否定情報によって反証型検証の有効性に対する認識が強化されたためであると考えられる。ただし、本実験課題は試行数が限定されており、確信度を高める段階から仮説を洗練する段階へと移行する前に課題が終了してしまったという可能性がないわけではない。しかし、被験者が呈示された仮説を正しいと思っている場合の方が多く、また別の可能性を考えたということもほとんどないことから、仮説に対する確信度が高くならなかったということは考えにくい。このように、2-4-6 問題を背景とする課題と、仮想生物課題を背景とする課題の間で事例選択傾向が異なることについては、現時点ではあくまで推測に過ぎないが、両課題の間に存在する構造的差異に原因があるのではないかと思われる。例えば、前述した通り、2-4-6 問題と仮想生物課題では、前者が完全に人工的な問題空間であるのに対して、後者では被験者が現実を持っているであろう生物学に関する常識的知識が発動する可能性があるという違いがある。後者のように、既存の知識が

利用できる領域では、その知識によって推論の方向性が大きく影響されることが考えられる。例えば、仮想生物課題は3つの次元からなる事例空間を持っていたが、被験者は既に持っている知識に基づいて、このうちの特定の次元のみに注目し、他の次元を重視しないといったことや、次元内の属性値の一部しか注目しないなどといったように、自発的に推論を制約していたかもしれない。一方、2-4-6問題は、仮想生物課題に対応する生物学的知識のような領域固有の知識体系がないために、確証型検証方略 (positive test strategy; Klayman & Ha, 1987) というヒューリスティックスが用いられ、結果として確証事例に偏った選択を行ったということは考えられる。

また、2つの課題では、事例集合に大きな違いがあることも指摘しておかなくてはならないであろう。2-4-6問題は、1次元のみから構成される空間ではあるが、事例の数は無限大に存在しうるのに対して、仮想生物課題の事例空間は、5次元という複雑な空間であり、潜在的な事例数は非常に大きいものの無限に広がる集合ではない。そのため、2-4-6問題と生物学課題では、反証事例の想起しやすさ、あるいは利用可能性 (availability: Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1973, 1974) が異なっており、特に2-4-6問題の方で反証事例の利用可能性が低かったのではないかと推測される。仮想生物課題と2-4-6問題の間の差については、第7章で改めて議論する。

5.4.3 呈示事例の適切性、および情報的価値の評価

事例選択傾向については、第1～第4実験で用いられた仮想生物課題とは異

なった結果がみられるものの、呈示された事例の適切性、および結果の情動的価値の評価については、部分的にはあるが、検証方略のシフトを支持する結果が得られた。

まず、呈示事例の適切性の評価については、少なくとも確認型仮説支持 (+H+T) となるタイプの検証事例は、規則発見条件、下位集合型誤り条件の双方で試行を追う毎に評価が徐々に下がっており、これは検証方略のシフトから導かれる予測と一致する。しかし、規則発見条件における反証型検証の適切性の評価については、全体的に評価スケールの中央値である 5 を下回っており、適切な検証ではないと評価されていることを示している。また、反証型検証の有効性が、試行を追う毎に高くなるという傾向も見られない。この結果は、検証方略のシフトとは異なった傾向を示している。確認型重視から確認型・反証型併用へと時系列に沿った事例選択傾向の変化が存在するとすれば、少なくとも試行を追う毎に、反証型検証の適切性評価が高くなることが予測されるからである。

情動的価値の評価についても以下のことが言える。規則発見条件、下位集合型誤り条件共に、少なくとも確認型仮説支持情報 (+H+T) については、試行が進むにつれて情動的価値の評価が徐々に下がっており、検証方略のシフトに基づく予測に沿っていると考えられる。しかし、反証型仮説否定情報 (-H+T) の評価については、規則発見条件の第 3 試行で呈示されたものについては比較的评价が高いものの、第 5、第 8 試行で呈示されたものについては評価が低く、両者共に評価スケールの中央値である 5 を下回っていた。従って、第 5、第 8

試行で呈示された反証型仮説否定情報は、どちらかという有効な情報ではないと評価されていることになる。この傾向も前述の適切性評価と同様、検証方略のシフトに基づく予測とは異なる。さらに、従来の仮説評価研究の知見とも反する (c.f. Farris & Revlin, 1989; Kareev & Halberstadt, 1993; Tschirgi, 1980)。

このように、検証事例の適切性評価、および獲得した情報の情報的価値の評価に関して言うと、本実験の結果は、検証方略の段階をふんだ変化という説明に基づく予測に反しているといえる。この点については、2 - 4 - 6 問題と仮想生物課題の間の構造的差異によるものであるという可能性は考えられる。一方、本研究では、これまで第 2, 第 3, 第 5 実験において、検証過程で得られた情報の価値の評価に関する検討を行ってきたものの、実験の主眼点が検証事例の選択傾向の検討という点におかれてきたため、呈示される情報の種類が未統制で、バランスのとれていない系列となっていたという問題点があることも事実である。そのような実験デザインのために、4 種の情報に対する評価を直接比較できず、検証の初期段階と、より後期の段階とで情報評価がどのように異なるかを分析することができなかった。次章の第 6 実験では、この問題を解決し、実験の目的を検証過程で得られる情報に対する評価の系統的な検討に置く。また、本研究で提唱された動態的変容分析パラダイムの考え方に基づいて、検証の初期段階と、より後期の段階とで各種の情報に対する評価を比較する。

5.5 第5章のまとめ

本章において報告した第5実験の目的は、仮説評価課題において、評価者が呈示される情報を基にした独自の推論を行っている、すなわち参照学習(reference learning)が行われているかどうかを検討することであった。第5実験の結果を見る限り、評価者、すなわち被験者は、呈示された仮説が誤っていると考えることは少なく、また、呈示された仮説以外の可能性を考えることも少なかった。以上のことから、本実験においては、参照学習の存在を示唆する結果は得られなかったといえる。

また、時系列に沿った仮説検証方略のシフトの一般性に関しては、全体的に確証型検証が多く、第2～第4実験で見られたような中盤以降で反証型検証の選択率が高くなるという傾向は観察されなかった。むしろ、反証型仮説否定情報を呈示された規則発見条件で反証型検証の選択が多くなっており、この情報を呈示されたことによる反証型検証に対する有効性認識の強化という要因が働いていると思われる。このように、一般的傾向としての段階的な方略の変化が本実験で見られなかったことについては、第2～第4実験で用いられた仮想生物課題と、本実験で用いられた2-4-6問題の間の構造的差異、特に領域固有の知識体系の有無や両課題での事例空間の大きさの違いに原因があることが推測された。この問題は第7章で改めて論じることとし、次の第6章では、仮説評価課題で与えられる4種類の情報の評価について、より統制された課題を用いて比較検討する。

第6章 仮説検証過程で得られた情報の評価

6.1 第6実験 - 4種類の情報に対する評価の比較

本章では、仮説検証過程において得られる情報に対する評価が、段階をふんで変化するかどうかを検討する。この情報の解釈については、Kareev & Halberstadt (1993) が、仮説支持情報と仮説否定情報とを比較し、被験者が仮説否定情報の効用を正しく認識できるということを確かめている。しかし、本研究の第2, 第3, 第4実験において示されたように、被験者の検証方略は時系列を通じてダイナミックに変化することが示されており、この方略変化に伴い、情報に対する評価が変動することも考えられる。ただし、第2～第4実験で用いられた課題では、主たる注目点はダイナミックなプロセスの中での事例選択傾向の変化を記述することであり、呈示される情報の種類の統制という点では問題があった。そこで、本実験では、仮説検証研究において伝統的に用いられており、かつ単純な課題状況である2-4-6問題を題材に用い、情報評価を系統的に分析する。特に、検証過程の比較的初期の状況と、終盤に差し掛かった状況とで、情報評価の違いを比較する。

仮説検証過程において得られる情報は、仮説に対する関係から、仮説を支持する情報と仮説を否定する情報の2つに分けられる。また、それぞれが、確証型検証の結果であるか、反証型検証の結果であるかによってさらに分けられ、合わせて4種類の組み合わせが生じる。これらは、確証型仮説支持 (*positive hit: +H+T*)、反証型仮説支持 (*negative hit: -H-T*)、確証型仮説否定 (*false positive:*

+H-T), 反証型仮説否定 (*false negative*: -H+T) である。仮説否定情報の1つである確証型仮説否定は、仮説に沿う確証事例 (+H) が非ターゲット事例 (-T) である場合で、ターゲットに含まれない事例集合が仮説に含まれていることを示す情報である。一方、反証型仮説否定は、仮説に沿わない反証事例 (-H) がターゲット事例 (+T) である場合で、ターゲットに含まれる事例集合が仮説集合に含まれず、仮説に見落としがあることを示す情報であるといえる。

Kareev & Halberstadt (1993) は、2-4-6問題を元にした仮説評価状況を設定し、被験者に上記の4種類の情報に関する情報的価値を評価させるという実験を行った。彼らは、先行する事例で呈示された情報の種類によって、後続の試行で行われる確証型検証および反証型検証の有効性、また、その結果得られる情報の有効性に関する評価が変化するかどうかを検討した。先行する事例で呈示される情報は、(1) 確証型仮説支持、(2) 確証型仮説否定、(3) 反証型仮説支持、(4) 反証型仮説否定、(5) なし (初期事例のみ) の5条件が用意された。Kareev & Halberstadt (1993) の実験1では、被験者は上記の5種類のうちどれか1つの情報を呈示された後、確証型検証、反証型検証のそれぞれについて、どの程度有効なテストとなりうるかを評価した。ここでの要因計画は、5つの先行情報 × 2種類の検証事例 (確証型・反証型) となり、前者は被験者間要因、後者は被験者内要因であった。この結果、被験者は全体的傾向として確証型検証を反証型検証より有効性が高いものとして評価すること、さらに、この結果が先行する試行で行われた確証型検証の数、および先行試行で得られたターゲット事例の数と負の相関関係にあることが示された。すなわち、先行

情報中に確証型検証，あるいはターゲット事例が含まれている場合は，そうでない場合と比べて反証型検証の選択率が高くなることが示された．一方，Kareev & Halberstadt (1993) の実験 2 では，被験者は前述した 5 種類の先行情報のうちどれか 1 つを呈示された後，確証型仮説支持，反証型仮説支持，確証型仮説否定，反証型仮説否定のそれぞれが，規則発見に対してどの程度有効であるかの評価を行った．従って，彼らの実験 2 は，5 つの先行情報 × 2 種類の検証事例（確証型・反証型）× 2 種類の結果（仮説支持・仮説否定）という 3 要因計画の実験を行ったことになる．このうち，結果の要因，すなわち呈示事例が仮説を支持するか，あるいは否定するかについては，被験者に直接的に明示されたのではなく，被験者は発見されるべき真の規則を告げられ，その規則と事例との関係から，事例が正負いずれのフィードバックを生じるかを推論するようになっていた．これら 3 要因のうち，先行情報は被験者間要因であったが，残り 2 つの条件は被験者内要因であった．この実験の結果，有効性評価に対する結果の主効果，および結果と検証事例の種類の交互作用が認められた．結果の主効果については，仮説否定情報の方が，仮説支持情報よりも高く評価されており，2 要因の交互作用については，反証型仮説支持 < 確証型仮説否定 < 確証型仮説支持 < 反証型仮説否定という順番で評価が高くなるという結果が見られた．

しかし，この Kareev & Halberstadt の実験には，被験者，すなわち評価者が発見すべき規則を知っているという問題がある．現実場面における科学的探求や，日常的な信念形成においては，発見すべき法則や真実が予めわかっている

ということは、通常ありえない。唯一そのような状況がありうるとすれば、過去において既に発見された真実に関して、回想的に思考を行う場合であろう。既に生じた事象に対する判断は、ハインドサイト判断 (hindsight judgment) と呼ばれるものであり、このような判断では、既に得られた結果の情報による判断の歪み、すなわちハインドサイト・バイアス (hindsight bias), ないしハインドサイト効果が生じることが知られている (Fischhoff, 1975, 1977; Hawkins & Hastie, 1990; Slovic & Fischhoff, 1977; Wasserman, et al., 1991)。既に指摘した通り、ハインドサイト・バイアスとは、ある知識をもつことによって、それを知らなかった時点での心的状態に戻った判断ができなくなるという現象をさす。Fischhoff (1975) は、この現象の説明として、漸進的決定論 (creeping determinism) というモデルを主張した。これは、結末に関する情報が与えられることによって、自動的に、かつ無意識的に、仮説とそれを支持する証拠とを強く結び付ける知識構造の再構成が生じ、情報が与えられる前の知識構造に基づいた判断ができなくなるというモデルである。このようなハインドサイト・バイアスを回避しつつ、現実場面で起こりうる仮説評価状況で、各種の情報がどのように評価されるかを検討するためには、被験者に発見すべき真の規則を教えるという手続きを採用するべきではないだろう。むしろ、検証の結果がターゲット事例 (正事例) であったか、非ターゲット事例 (負事例) であったかというフィードバックのみを与えることで評価を求める手続きの方が、現実的に人が遭遇する場面との対応がとれているといえる。

Kareev & Halberstadt (1993) のもう一つの問題としては、事例を有効性

(usefulness) という基準のみに基づいて評価するよう求めていることが挙げられる。4 種類の情報は、検証事例の種類、および結果から構成されているが、この情報の「有効性」とは何を指すのかが曖昧であると思われる。例えば、検証事例としては適切であると思われたが、結果として得られた情報は、それほど役に立たないといったこともありうるだろう。従って、検証事例の適切性と得られた結果の有効性は、別々に検討する必要があると考えられる。

本実験では、このような問題を回避するために、(1) 検証事例としての適切性、(2) 得られた結果も含めた形での情報の有効性、という 2 つの観点から情報を評価するよう求めた。

6.2 方法

課題および条件

Wason (1960) による 2 - 4 - 6 問題を基にして、架空の被験者が規則発見を試みている状況を設定し、そこで得られた情報の価値を評価させるという課題を用いた。本実験で設定された条件は、状況、検証事例の種類、結果の 3 種類であり、すべての条件が被験者内要因であった。

まず、状況要因については、「2 4 6」がターゲット事例であるという情報のみが与えられた状態である初期状況 (initial phase) と、ターゲット事例が 10 個、非ターゲット事例が 7 個提示される終盤状況 (latter phase) の 2 種類が用意された。終盤状況で呈示される 17 個の情報には、初期状況において評価の対象となる 8 つの事例が含まれていた。

検証事例，結果の要因については，2種類の検証事例（確証型・反証型事例），および2種類の結果（仮説支持・仮説否定）を組み合わせることにより，確証型仮説支持（ $+H+T$ ），確証型仮説否定（ $+H-T$ ），反証型仮説支持（ $-H-T$ ），反証型仮説否定（ $-H+T$ ）の4種類の呈示情報が用意された．初期・終盤の2状況のそれぞれにおいて，4種類の情報に該当する事例が2つずつ用意された．すなわち，被験者は，それぞれの状況において8つの事例に対する評価，全体では16の事例に対する評価を行ったことになる．また，終盤状況において，確証型仮説否定を可能にするため，発見すべき規則をオリジナルの「上昇系列（ascending numbers）」から「正の上昇系列（ascending positive numbers）」，すなわち，「第1項より第2項が，第2項より第3項が大きく，かつ第1，第2，第3項すべてが正の数」という想定の下で呈示事例を作成した．それぞれの状況において呈示された仮説，および事例を表6-1に示す．

情報の評価は，(1) 呈示された事例の検証事例としての適切さ，(2) 結果までを含めた情報の有効性（情動的価値）という2つの側面から，それぞれ0～

表6-1 各状況における仮説および呈示事例

	初期状況 「連続する偶数」	終盤状況 「公差が正の等差数列」
$+H+T$	“10 12 14” “102 104 106”	“2 8 14” “100 200 300”
$-H-T$	“2 2 4” “5 1 4”	“19 10 1” “10 2 6”
$+H-T$	“8 6 4” “100 98 96”	“-2 0 2” “-30 -20 -10”
$-H+T$	“1 3 5” “2 6 10”	“2 8 10” “3 6 12”

10 点で評価するよう求めた。

手続き

実験は質問紙形式で行われた。被験者は、まずはじめに、2 - 4 - 6 問題の概略の説明、および架空の人物が問題解決を行う過程について評価するという趣旨の教示を読んだ。次に、これまでにわかっていることとして、初期事例「2 4 6」がターゲット事例、すなわち、発見すべき規則に沿う事例であることが告げられた。続いて、初期状況における仮説、すなわち「連続する偶数」が呈示された後、確証型仮説支持 (+H+T)、確証型仮説否定 (+H-T)、反証型仮説支持 (-H-T)、反証型仮説否定 (-H+T) の 4 種類に対応した情報がそれぞれ 2 つずつ、事例と結果の組として呈示された。このとき、被験者は、呈示された 8 つの情報のそれぞれに対して、検証事例としての適切さ (0: 全く不適切 ~ 10: 非常に適切)、得られた結果の情報的価値 (0: 全く役に立たない ~ 10: 非常に役に立つ) という観点から評価した。初期状況における評価が終わった後で、終盤状況の評価が行われた。終盤状況においては、これまでにわかっていることとして、初期事例の「2 4 6」、および初期状況で明らかになった 8 つの結果に加え、終盤状況での仮説である「公差が正の等差数列」に矛盾しない 8 つの事例が与えられたが、それ以外の手続き等は初期状況と同一であった。

被験者

被験者は、大学生および専門学校生 38 名であった。実験は質問紙形式により、

講義中に集団で、あるいは個別に実施された。1人あたりの所要時間は概ね 15 ～ 30 分程度であった。

6.3 結果

検証事例の適切性

事例は、2つの状況のそれぞれにおいて、確証型仮説支持、確証型仮説否定、反証型仮説支持、反証型仮説否定の4種類の事例が2つずつ呈示され、被験者はそのそれぞれについて、仮説の真偽を確かめるために適切な事例であったかどうかを0～10点で評価した。被験者は、それぞれの状況において、1種類あたり2つの事例に対して適切性を評価しており、その2つの事例に対する評価の平均値をそれぞれの種類の事例に対する被験者の評価値とした。図6-1は、検証事例としての適切性評価を、4種類の事例、および状況別に示したものである。図に示される通り、仮説を支持する事例の適切性評価は終盤状況で上昇しているが、仮説を否定する事例の適切性評価は逆に下降している。特に終盤状況での仮説否定事例に対する適切性評価は、確証型仮説否定が4.72、反証型仮説否定が4.99と、評価スケールの中心である5を下回っており、どちらかというと適切でないと評価されていることがわかる。

検証事例の適切性について、状況（初期・終盤）、事例種類（確証・反証）、結果（仮説支持・否定）を独立変数とした3要因の分散分析を行った結果、結果の主効果 [$F(1,37) = 5.16, p < .05$]、状況と結果の交互作用 [$F(1,37) = 19.38, p < .001$]、事例種類と結果の交互作用 [$F(1,37) = 7.80, p < .01$] がそれぞれ認めら

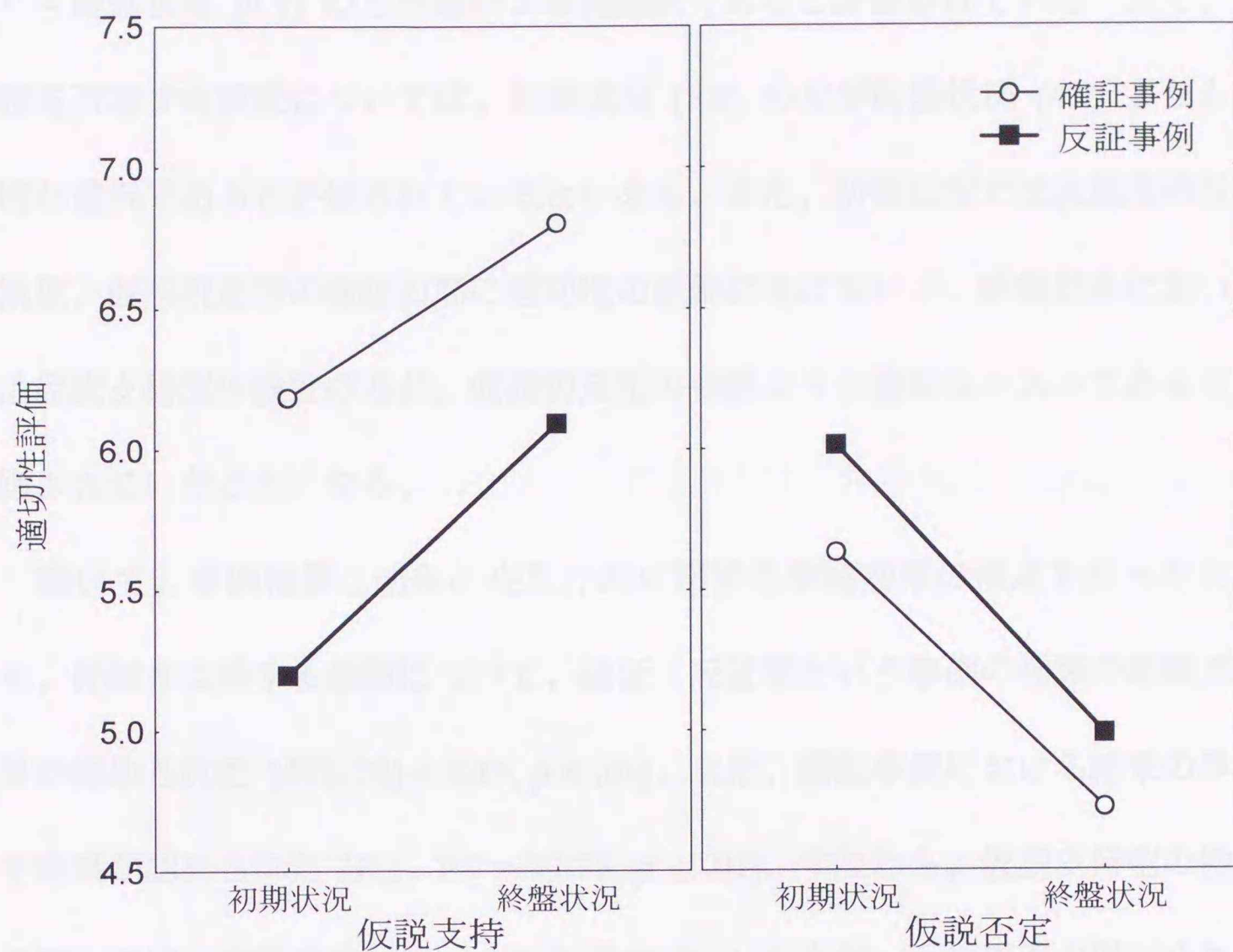


図6-1 各状況における検証事例の適切性

れた。結果の主効果については、仮説を支持する検証事例 (6.01) の方が仮説を否定する検証事例 (5.34) よりも適切であると評価された。

状況と結果の交互作用について、単純効果の検定を行ったところ、仮説を支持する検証事例、仮説を否定する検証事例のそれぞれにおける状況の単純主効果が有意であり [それぞれ $F(1, 74) = 4.18$, $F(1, 74) = 6.83$, 共に $p < .05$]、また、終盤状況における結果の種類の単純主効果が有意であった [$F(1, 74) = 17.89$, $p < .001$]。すなわち、仮説を支持する検証については、初期状況 (5.7) よ

りも終盤状況 (6.4) の方が適切な事例選択であると評価されている一方で、仮説を否定する検証については、初期状況 (5.8) の方が終盤状況 (4.9) よりも適切な選択であると評価されているといえる。また、初期状況では仮説支持型の検証、仮説否定型の検証の間に適切性の評価に差はないが、終盤状況においては仮説支持型の検証の方が、仮説否定型の検証よりも適切なテストであると評価されていたことになる。

続いて、事例種類と結果の交互作用に関する単純効果の検定を行ったところ、仮説を支持する事例について、確証・反証型という事例の種類の単純主効果が認められた [$F(1, 74) = 8.15, p < .01$]。また、確証事例における結果の単純主効果が認められた [$F(1, 74) = 11.76, p < .01$]。すなわち、仮説支持型の検証に関しては、確証事例が用いられた場合 (6.5) の方が、反証事例が用いられた場合 (5.6) よりも適切なテストであると評価される一方で、仮説否定型の検証に関しては、用いられた事例が確証事例 (5.2) であれ、反証事例 (5.5) であれ、適切性の評価に差は認められないという結果であった。また、確証型仮説支持と確証型仮説否定の間では適切性評価に差が見られるが、反証型仮説支持と反証型仮説否定の間では差が認められなかった。

事例種類の主効果、事例種類と状況の交互作用、3 要因の交互作用は認められなかった [全て、 $F(1, 37) < 1$]。確証事例の適切性に対する全体的な平均は 5.84、反証事例は 5.57 であった。また初期状況における確証事例に対する適切性評価の平均は 5.91、反証事例では 5.61 であり、終盤状況における確証事例の平均は 5.76、反証事例の平均は 5.54 であった。

情報の有効性

呈示事例の有効性、すなわち情報的価値についても、適切性評価と同様に、それぞれの種類について2つずつ呈示された事例の平均値を、被験者の評定として扱う。図 6-2 は、4 種類の事例に対する有効性の評価を、状況別に示したものである。有効性の評価については、反証型仮説否定情報が、初期状況より終盤状況で低くなっているのに対し、それ以外の3種類の情報は、逆に初期状況より終盤状況のほうが高くなっている。また、初期状況における反証型仮説支持情報の有効性が評定尺度の中央値である5を下回っており、どちらかというと役に立たない情報として評価されていた。

事例の有効性について、状況（序盤・終盤）、事例種類（確証・反証）、結果（仮説支持・否定）を独立変数とした3要因の分散分析を行った結果、状況と結果の交互作用 [$F(1,37) = 9.09, p < .01$]、および事例種類と結果の交互作用 [$F(1,37) = 5.80, p < .05$] が認められた。また、事例種類の主効果が有意となる傾向にあった [$F(1,37) = 3.26, p < .08$]。

状況と結果の交互作用について単純効果の検定を行ったところ、仮説支持情報、および仮説否定情報における状況の単純主効果はいずれも有意ではなかった。一方、初期状況および、終盤状況における結果の単純主効果は、共に有意となる傾向があった [それぞれ、 $F(1, 74) = 3.05, p < .10$, $F(1, 74) = 3.23, p < .10$]。この結果については、以下のことがいえるだろう。まず、初期状況においては、仮説否定情報 (5.9) の方が仮説支持情報 (5.4) に比べてやや評価が

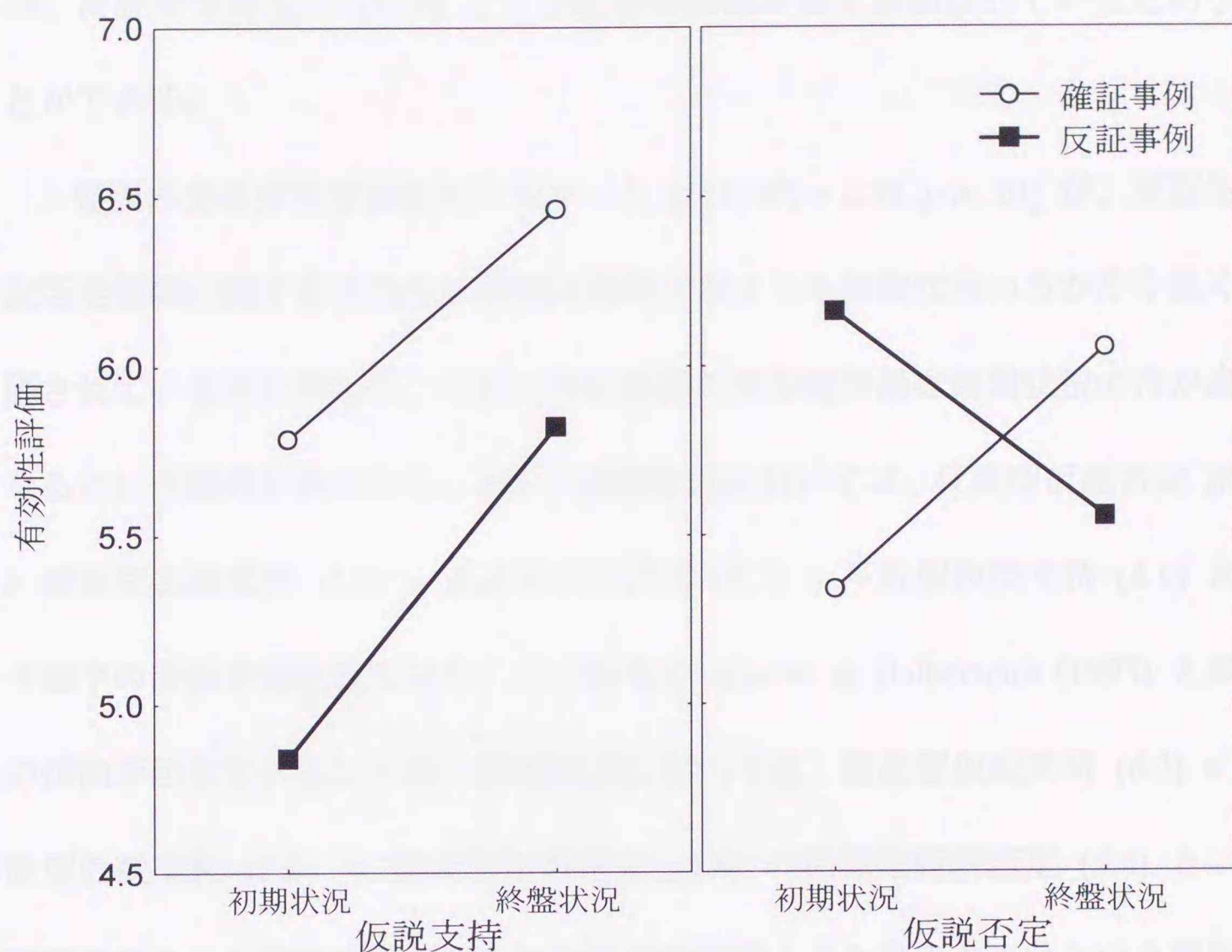


図6-2 事例の情報的価値の評価

高く、終盤状況では、逆に仮説支持情報 (6.0) の方が仮説否定情報 (5.5) よりも評価が高くなっている。

一方、事例種類と結果の交互作用について、同様に単純効果の検定を行ったところ、仮説を支持する結果において事例種類の単純主効果が認められた [$F(1, 74) = 8.89, p < .01$] が、仮説を否定する結果における事例種類の単純主効果、確証、反証それぞれの事例種類における結果の単純主効果が認められなかった。この結果については、仮説支持情報のうち、確証型仮説支持 (6.1) の方

が、反証型仮説支持 (5.29) よりも情報的価値が高く評価されていたということが出来る。

3 要因の交互作用は認められなかった [$F(1, 37) = 1.55, p > .10$] が、反証型仮説否定情報に関する有効性の評価は初期状況よりも終盤状況の方がやや低く評価されているのに対して、それ以外の情報の有効性評価は終盤状況の方が高くなるという傾向が見られる。また、初期状況においては、反証型仮説否定 (6.1) > 確証型仮説支持 (5.9) > 確証型仮説否定 (5.7) > 反証型仮説支持 (4.9) という順序の評価がなされており、この結果は Kareev & Halberstadt (1993) と同様の傾向を示している。一方、終盤状況においては、確証型仮説支持 (6.3) > 反証型仮説支持 (5.6) $\div$ 確証型仮説否定 (5.6) > 反証型仮説否定 (5.4) という順序であり、仮説支持情報の方が仮説否定情報よりも評価が高いという傾向が見られる。

6.4 考察

6.4.1 本実験で得られた結果

本実験では、確証事例・反証事例、および仮説支持情報・仮説否定情報の組み合わせからなる 4 種類の情報が、検証初期の状況と、終盤の状況のそれぞれにおいてどのように評価されるかについて、検証事例としての適切さ、および得られた情報が規則発見に対して持つ情報的価値の二点に分けて検討した。

呈示された事例が、検証事例としてどの程度適切なものであったかについては、その事例が確証型か反証型かというよりは、むしろ最終的な結果として

仮説を支持したか、否定したかに基づいて評価がなされたという傾向が見られた。全体的傾向としては、仮説を支持する検証事例がより適切であると評価され、仮説を否定する検証事例の評価を上回っていた。また、この傾向は、状況の種類との交互作用が認められ、情報が少ない初期状況においては仮説支持情報と仮説否定情報の間に差は認められないが、情報が多く集まった終盤状況においては、仮説支持情報は適切な検証であると評価される一方で、仮説否定情報はどちらかというと適切ではないという評価がなされていた。また、全体的傾向として、確証型仮説支持情報 (*positive hit: +H+T*) をもたらす検証事例が、その他の3種類の情報よりも、適切な検証事例であるという評価がなされていた。この結果は、仮想生物課題を用いた、第3章の第2～第3実験で得られた適切性評価の結果とはやや異なっている。特に第3実験においては、後半の試行になるほど反証型検証の適切性評価が高くなる傾向が見られたが、本実験では、確証型検証と反証型検証の間に差は見られない。このように、2-4-6問題を背景とする本実験と、仮想生物課題を背景とする第2、第3実験との間で適切性の評価が違うことについては、あくまで推測ではあるが、第5章で指摘したような両課題の構造的差異、特に領域固有の知識体系の有無や事例空間の大きさの違いといった要因が関与している可能性は考えられる。

しかし、全体的な傾向としては、どちらの場合も評価スケールの中央値である5を越えており、どちらかというと適切な検証であると評価されているといえるだろう。この結果は、先行呈示された確証型検証、あるいはターゲット事例が多いほど確証型検証と反証型検証との間で評価に差がなくなることを

示した Kareev & Halberstadt (1993) の結果とよく類似しており、被験者が反証型検証を必ずしも役に立たない検証であるとは考えていないことを示している。

一方の規則発見に対する情報的価値については、反証型仮説否定情報 (*false negative*) が初期状況から終盤状況に移行することで評価が低下するのに対し、その他の3つの情報は評価が高くなるという傾向が見られた。統計的分析の結果からは、仮説を支持する情報に関しては、確証型仮説支持 (*positive hit*) の方が反証型仮説支持 (*negative hit*) よりも評価が高いということがいえる。しかし、仮説支持情報と仮説否定情報との間では、評価に差はなく、また仮説否定情報に関していえば、確証型と反証型では差が見られなかった。

また、初期状況では仮説支持情報よりも仮説否定情報の方が有効な情報であると評価される一方で、終盤状況では仮説支持情報の方が仮説否定情報よりも有効であると評価されるというように、一般的傾向としての段階的な検証方略の変化という考え方から導かれる予測とは逆の傾向があると言える。

6.4.2 検証方略の時系列に沿った動態的変容という考え方の一般性

検証方略の時系列に沿った動態的な方略変容という考え方では、仮説検証において、まず仮説の蓋然性を高めるための試みがなされ、後に仮説に見落としがないかどうかを確認するなど仮説を洗練する試みがなされるというように検証のフェーズが移行することを予測する。このときの事例選択傾向としては、本研究のこれまでの実験が示してきた通り、まず確証型検証によって仮説に対する支持を獲得し、仮説の蓋然性を高めようとし、次に見落としを確認するため

に反証型検証と確証型検証を併用するようになると考えられる。本研究における一連の実験で示されたように、少なくとも第1～第4実験で用いられた架空の生物を題材にした規則発見課題、受容方略型課題、仮説評価課題では、このモデルを支持する結果が得られている。また、この事例選択傾向に伴って、仮説評価状況において呈示される事例の適切性や、結果の情報的価値の評価が変化すると考えられる。特に検証事例の適切性の評価については、反証型検証の適切性評価が、少なくとも検証の終盤状況において高くなることが推測される。しかし、本実験の結果を見る限り、初期・終盤状況において反証型検証の適切性評価に差はなく、結果の情報的価値の評価についても、反証型仮説否定情報に対する評価が初期状況から終盤状況に移行するに従って低下するという結果が得られた。

このように本実験の第1～第4実験で用いられた仮想生物課題と、第5、第6実験で用いられた2-4-6問題とでは、被験者の反応傾向に差が見られる。この違いを生じた原因として、両課題のもつ構造的差異が考えられる。第5章においても指摘した通り、仮想生物課題については、生物学における常識的知識という領域固有の知識体系が推論時に用いられる可能性がある一方で、人工的な問題空間である2-4-6問題においては、そのような知識体系が存在しない。この領域固有知識の有無が推論に影響を与えたということは十分に考えられる。また、仮想生物課題における事例空間は、大きいものではあるが有限の集合である一方で、2-4-6問題における事例空間は無限の広がりを持っている。そのために、特に後者では反証事例の想起しやすさが低いために、

そのような事例が選択されにくいこと、またそれに伴って反証型検証の有効性がそれほど高く評価されなかったのではないかと考えられる。

6.5 第6章のまとめ

本章の目的は、仮説評価課題において呈示される情報に対する評価が、方略選択傾向に伴って、時系列に沿って変化するかどうかを検討することであった。

その結果、初期状況では仮説支持情報よりも仮説否定情報の方が有効な情報であると評価される一方で、終盤状況では仮説支持情報の方が仮説否定情報よりも有効な情報として評価されるというように、一般的傾向としての検証方略のシフトという考え方から導かれる予測とは逆の傾向があることが示された。また、全体的な傾向としても、確証型検証と反証型検証の間で適切性評価に差が見られないといったように、第3章の第3実験における仮想生物課題をベースにした仮説評価課題の結果と異なっていた。以上のような結果を生じた原因として、本実験のベースとなった2-4-6問題と、第1～第4実験で用いられた仮想生物課題との間の構造的差異、特に領域固有的知識の有無や両課題における事例空間の大きさの違いというものが考えられる。第7章では、2-4-6問題と仮想生物課題の間で事例選択、および得られた情報に対する評価が異なることについて、両課題間の構造的差異という観点から考察する。

第7章 課題の種類と反応傾向についての考察

7.1 仮想生物課題と2-4-6問題の構造的差異

本章では、第1～第3、ならびに第4実験で用いられた仮想生物課題と、第5、第6実験で用いられた2-4-6問題で観察された反応傾向の違いについて検討する。まず、両課題における反応傾向をみる前に、これらの課題の構造的差異を指摘しておかなくてはならないだろう。本研究で用いられた仮想生物課題と、仮説検証研究で伝統的に用いられてきた2-4-6問題は、共に被験者が、実験者の想定しているルールを特定するために仮説生成・テストを繰り返す規則発見課題である。しかし、これらの課題には、領域固有の知識体系の有無、事例空間の大きさという違いがあることに注意しなくてはならない。

領域固有の知識体系に関しては、第1～第4実験で用いられた仮想生物課題では、架空の生物や細菌が好む環境という規則は恣意的ではあるものの、被験者が持つ生物学的常識という知識体系が推論に影響を与えた可能性があることを指摘しておかなくてはならないだろう。一方で、2-4-6問題とは、完全に人工的な問題空間によって構成されており、仮想生物課題における生物学的常識のような領域固有の知識体系に対応するものが存在しない、いわば具体性・日常性を欠く課題であるといえる。

また、事例空間という点から言うと、前者は5つ(第1～第3実験)、あるいは3つ(第4実験)というように複数の次元から構成される事例空間を持つが、後者は数字という1つの次元のみから構成される事例空間を持つ点が異なる。

る。また、仮想生物課題は複数の次元を持つものの、それぞれの次元における属性値の範囲は決まっており、事例空間は有限であるが、2-4-6問題における数字列という事例空間は1次元でありながら、無限の広がりを持っているという点においても2つの課題は異なっている。ここで特に重要なのは、両課題における事例空間の大きさの違いである。この違いが、反応パターンに与える影響としては、反証事例の想起しやすさ、ないし利用可能性 (availability) が挙げられるであろう。仮想生物課題では、各次元における属性値が有限であるために、それぞれの次元に対する仮説中の言及の仕方は、「次元 X は p から q の範囲」であるという形をとりやすい。一方、2-4-6問題では、事例空間が無限に広がっており、かつ規則の形式が、「第1項と第2項の差 d_{12} と第2項と第3項の差 d_{23} はしかじかである」というように、仮想生物課題とは全く異なった形式の仮説となる。そのために、仮想生物課題では反証事例を想起しやすく (p から q の範囲に含まれない任意の属性値を選べば良い)、2-4-6問題では反証事例が想起しにくかったのではないと思われる。

7.2 仮想生物課題と2-4-6問題における事例選択傾向の違い

それでは、両課題において実際に得られた反応を比較する試みとして、事例選択傾向の違いをみてみよう。第1～第3実験で用いられた仮想生物課題、第4実験で用いられた仮想生物課題、第5実験で用いられた2-4-6問題の評価バージョン課題では、それぞれ試行数、またブロック数が異なるため、全ての課題における反応を、前半部と後半部の2つのブロックに分け直し、それぞ

れにおける確証型検証の選択率を計算し、表 7-1 に示した。表中の第 1 列は、本研究においてそれぞれの課題が用いられた実験を示す。第 2 列は、課題の種類を示す。第 1 ～ 第 3 実験で用いられたものは、5 次元の事例空間からなる仮想生物課題であるため「5 次元」と表記される。同様に、第 4 実験で用いられた仮想生物課題は「3 次元」、第 5 実験で用いられた 2-4-6 問題の評価バージョンは「2-4-6」と表記される。

第 3 列は、それぞれの実験で分析の単位となった条件を示す。第 1 実験では実験開始時点での条件分けは行われていなかったが、正解到達者と未到達者を分けて検討していたため、ここでもその分類を適用する。第 2 実験では、正事例系列条件と負事例系列条件に分けられ、それぞれの条件で別の情報系列が呈示された。第 3 実験では、正解到達条件と未到達条件に分けられ、それぞれで

表 7-1 各課題における選択率のシフト

実験	課題	条件	Block1		Block2		N	P
			+H	-H	+H	-H		
第 1 実験	5 次元	正解到達	.79	.21	.75	.25	22	.58
		正解未到達	.82	.18	.79	.21	18	.68
第 2 実験	5 次元	正事例系列	.63	.37	.47	.53	10	.03
		負事例系列	.94	.06	.95	.05	10	.91
第 3 実験	5 次元	正解到達	.40	.60	.29	.71	10	.14
		正解未到達	.96	.04	.99	.01	10	.66
第 4 実験	3 次元	+H+T	.59	.41	.56	.44	10	.67
		-H-T	.45	.55	.28	.72	10	.02
		+H-T	.52	.48	.50	.50	10	.77
		-H+T	.52	.48	.47	.53	10	.49
第 5 実験	2-4-6	規則発見	.93	.07	.83	.17	39	.17
		下位集合型誤り	.88	.12	.91	.09	38	.72

異なる情報系列が呈示された。第4実験は、先行的に呈示される情報の種類によって4つの条件に分けられ、それぞれの条件は確証型仮説支持 (+H+T) 呈示条件、反証型仮説支持 (-H-T) 呈示条件、確証型仮説否定 (+H-T) 呈示条件、反証型仮説否定 (-H+T) 呈示条件であった。第5実験では、規則発見条件と下位集合型誤り条件に分けられ、それぞれで異なった情報系列が呈示された。

表の第4列から第7列は、それぞれの実験・条件における確証型検証および反証型検証の選択比率を、前半の試行 (Block1) と後半の試行 (Block2) 別に示している。第8列は、各行に対応する実験・条件に参加した被験者数を示す。また第9列は、それぞれの実験・条件において、前半と後半で選択のパターンに変化がみられたかどうか、すなわちブロックと選択事例の種類の交互作用が有意であったかどうかの検定結果を示した。

前半から後半へと移行するにあたって選択パターンが変化したかどうかを検討するために、交互作用に関する p 値を用いたメタ分析 (meta-analysis) を行った。ここでの分析は、Stouffer-method (Rosenthal, 1984) に基づいた。第1実験から第5実験の結果を総合した結果、ブロックおよび選択事例の種類の交互作用は、全体としては有意ではなかった ($p = .31$)。

一方、表を見ると、これらの実験結果はいくつかのパターンに分けることができるようである。まず、全体的な傾向として反証型検証の選択率が比較的高いグループとそうでないグループによって分類すると、第2実験の正事例系列条件、第3実験の正解到達条件、第4実験の4つの条件は前者に相当し、第2実験の負事例系列条件、第3実験の正解未到達条件、第5実験の2条件は後者

に相当するといえる。第1実験の2条件は、どちらかというと後者よりではあるが、やや反証型検証の選択率が高いといえる。また、ブロック1から2へと試行段階が進むに連れて、反証型検証の選択率が上がっているかどうかによって分類すると、第2実験の正事例系列条件、第3実験の正解到達条件、第4実験の $-H-T$ 条件、第5実験の規則発見条件は後半に向かうにしたがって、反証型検証の占める割合が10%ほど増加しているのに対し、残りの条件は反証型検証の選択は増加していない。また、ブロック2で反証型検証の選択率が増加した4つのグループのうち、第2実験の正事例系列条件と第4実験の $-H-T$ 条件では交互作用が有意であったが (共に $p < .05$)、第3実験の正解到達条件と第5実験の規則発見条件では有意ではなかった (共に $p > .10$)。しかし、第3実験の正解到達条件では、元々反証型検証の選択率が高いが、第5実験の規則発見条件では、逆に確証型検証の選択率が高いという違いは存在する。

ここで、12のグループの分類を試みる。まず、それぞれのグループにおける全体を通した反証型検証の選択比率、およびブロック1からブロック2の移行に際して生じた反証型検証の選択比率の変化を元にクラスター分析を行った (図7-1)。ここでの距離測度としては平方ユークリッド距離を、凝集ルールとしてはワード法を用いた。図に示されるように、グループは大きく2つのクラスターに分けられるといえる。1つめのクラスターは、実験1の2条件、実験2の負事例系列条件、実験3の正解未到達条件、実験5の2条件というグループから構成され、全体的に反証型検証の選択率が低い (図7-1 上方のクラスター)。このクラスターは、確証型検証偏向群とでも呼ぶべきものである。

それに対して、2 つ目のクラスターは、反証型検証の選択率が比較的高いグループが集まって構成されている。このクラスターは、確証・反証型検証併用群と呼ぶのが適当であろう。

この違いを生じた要因としては、まず第1に本研究で指摘してきた時系列に沿った段階的な方略の変化が挙げられるであろう。少なくとも、第1実験の2条件、第2実験の負事例系列条件、第3実験の正解未到達条件に関しては、試行回数が足りなかったため十分に検証し尽くされていない、あるいは仮説に対

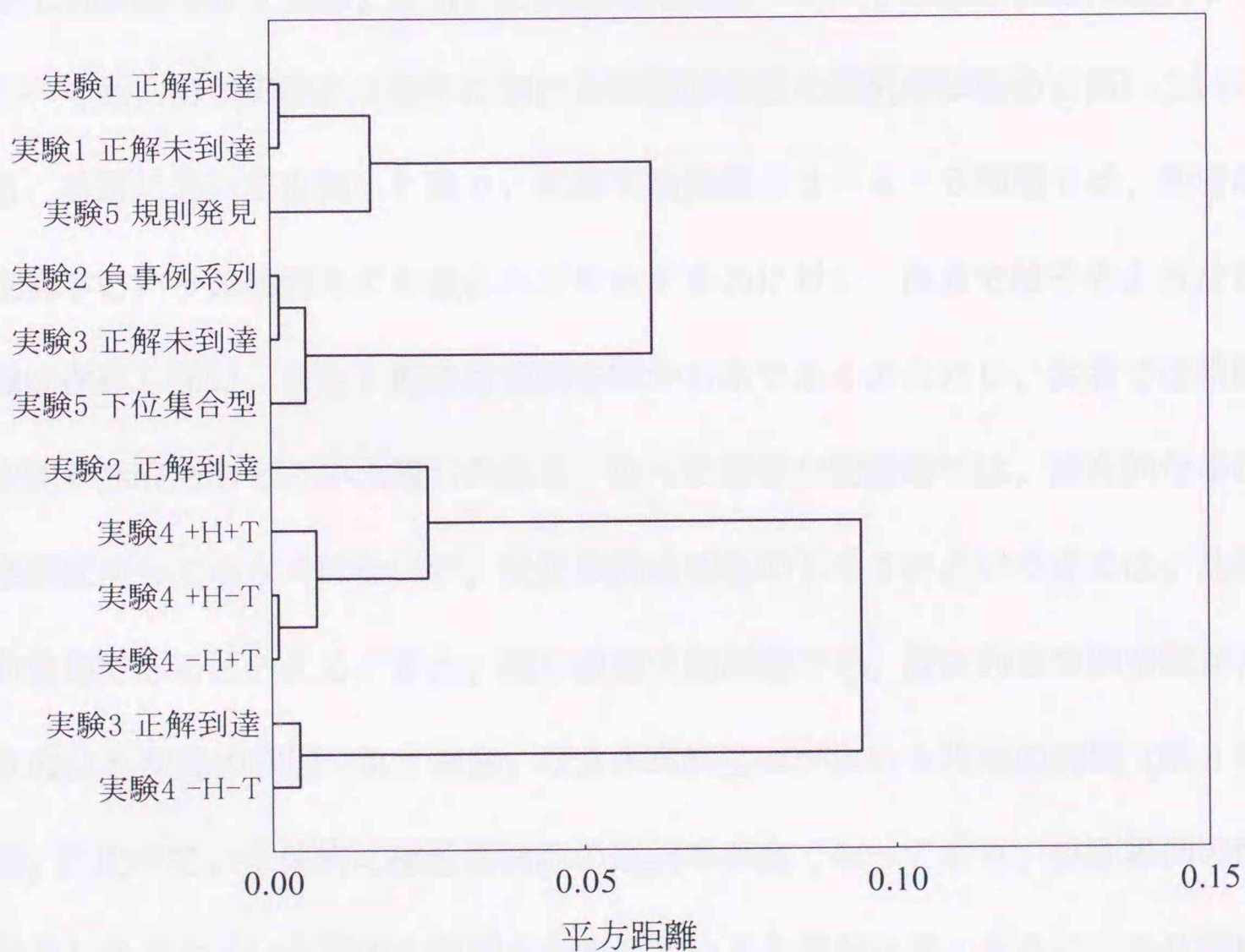


図7-1 反証型検証の選択率およびその変化に基づく階層的クラスター分析によるグループ間の類似構造

する確信度が高まっていないため反証型検証への移行が行われないという解釈が可能であろうと思われる。一方、第5実験で用いられた2-4-6問題を背景とする仮説評価課題で確証型検証の選択比率が極めて高いのは、第1～第4実験で用いられた仮想生物課題と第5実験で用いられた2-4-6問題の構造的な差異、特に領域固有の知識体系の有無や、反証事例の想起のしやすさの違いに起因するものではないかと考えられる。一般に規則発見課題は仮説評価課題に比べて確証型検証の選択率が高くなることが知られている (e.g. Farris & Revlin, 1989; Tschirgi, 1980) が、5次元の仮想生物課題 (第1実験)、3次元の仮想生物課題 (第4実験) を用いた規則発見課題における確証型検証の選択率と比べても、第5実験の2条件における確証型検証の選択率は極めて高いといえる。前節において指摘した通り、仮想生物課題と2-4-6問題では、前者は生物学という領域固有の知識体系が存在するのに対し、後者ではそのような知識は存在しない。また、前者は事例空間が有限であるのに対し、後者では事例空間が無限であるという違いがある。従って仮想生物課題では、潜在的な事例空間は決して小さくはないが、反証事例の想起のしやすさという点では、比較的容易であるといえる。また、同じ仮想生物課題でも、潜在的な事例空間がより広い5次元の課題 (第1実験) の方が事例空間が狭い3次元の課題 (第4実験) に比べて、全体的に確証型検証の選択率が高くなっており、反証事例の想起のしやすさという要因の影響を示唆していると思われる。さらに、全体的には確証型検証の選択が多い第5実験の2条件のうち、規則発見条件の方が、下位集合型誤り条件よりも反証型検証の選択比率が高いということについては、

全体的な傾向としては反証事例が想起しにくかったものの、規則発見条件では反証型仮説否定情報が呈示されることで反証型検証に対する有効性が高く評価され、下位集合型誤り条件よりは選択されやすかったのではないかとと思われる。

以上のことから、仮説検証における事例選択は、仮説の蓋然性を高める段階から仮説を洗練させる段階への移行に伴う、確証型重視 → 確証型・反証型併用という事例選択傾向のシフトに加えて、領域固有の知識体系の有無、反証型検証の想起のしやすさ、反証型仮説否定情報の獲得による反証型検証に対する有効性認識の強化などの要因が複雑に絡み合ったプロセスであると考えられる。

7.3 仮想生物課題と2-4-6問題における検証事例の適切性評価

次に、2つの課題において呈示された事例に対する適切性の評価がどの程度異なるかについてみてみよう。呈示事例の適切性評価が求められたのは、第2実験の受容方略型5次元仮想生物課題、第3実験の仮説評価型5次元仮想生物課題、および第5実験の評価型2-4-6問題である。このうち、第3実験の正解未到達条件と、第5実験の下位集合型誤り条件では、反証事例が呈示されないため、第2実験の正事例系列条件、負事例系列条件、第3実験の正解到達条件、第5実験の規則発見条件で、検証事例の適切性評価を比較する。また、第2、第3実験では-2～2の5件法で、第5実験では0～10の11件法での評価がなされていたので、第2、第3実験における評価値を線形変換し、第5実験における評価値と揃えて比較しよう。比較の対象となった4つのグループで、前半の試行(Block1)、および後半の試行(Block2)のそれぞれにおける確

証型検証 (+H) と反証型検証 (-H) の適切性評価を示したのが 図 7-2 である。

ここで、前節において行われたクラスター分析の結果を基にすると、上記の 4 つのグループは、第 2 実験正事例系列条件・第 3 実験正解到達条件、第 2 実験負事例系列条件・第 5 実験規則発見条件というカテゴリに分けることが可能である。まず第 1 のカテゴリでは、確証型検証と反証型検証の双方に対する有効性評価が高く、逆に第 2 のカテゴリでは反証型検証に対する評価は低く、かつ評価スケールの中央値である 5 を下回っており、どちらかという適切な検証

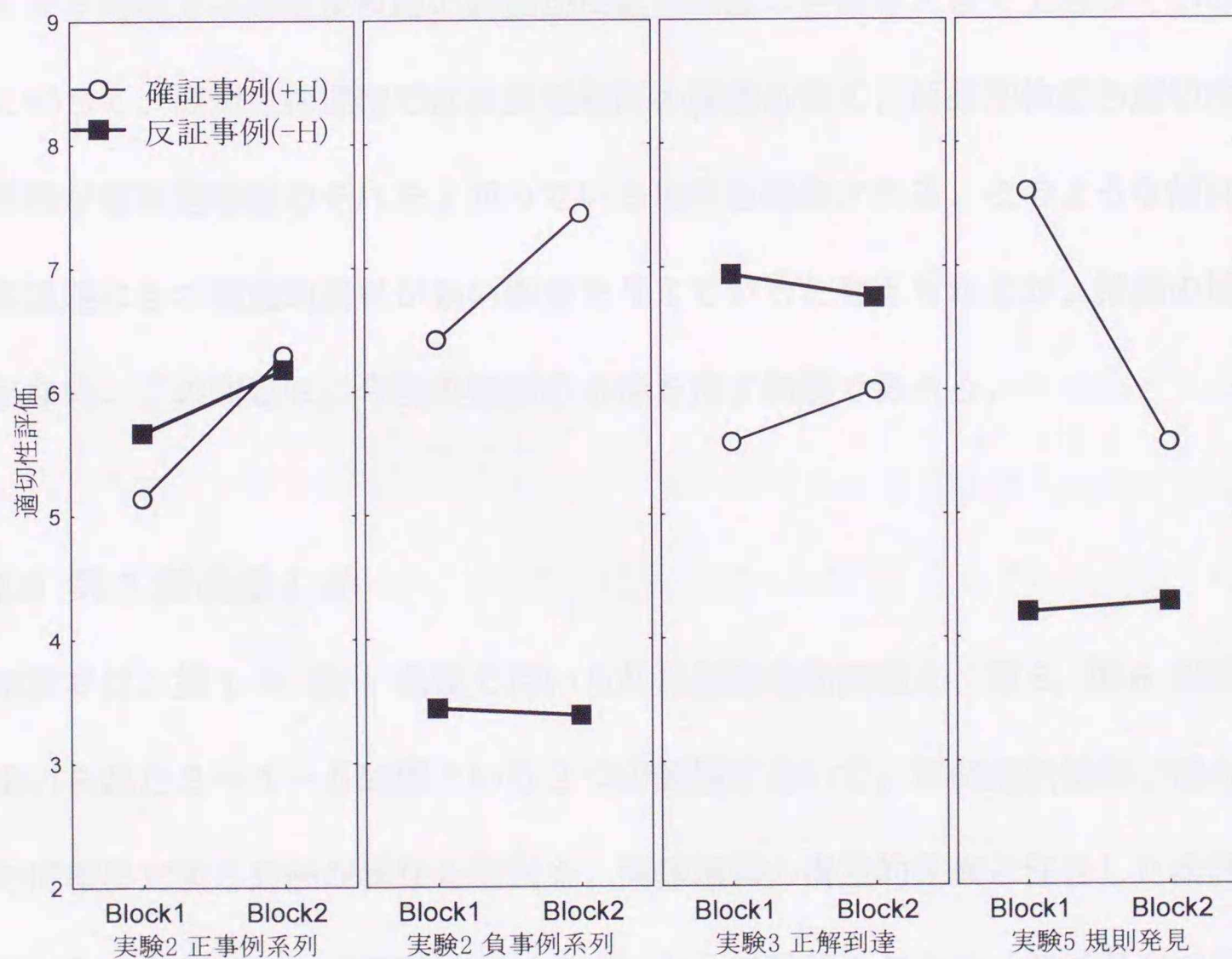


図7-2 各グループにおける呈示事例の適切性評価

ではないと評価されていることがわかる。この傾向は事例選択傾向とも対応しており、前者では確証型検証と反証型検証がほぼ同じ程度か、あるいは反証型検証の方が選択率が高いのに対し、後者では確証型検証の選択率が、反証型検証の選択率を常に上回っている。

このように、2-4-6問題における事例選択傾向と、仮想生物課題における事例選択傾向は大きく異なっている。特に、仮想生物課題においては反証型検証の選択率が高い。また、この事例選択傾向と対応して、受容方略課題、仮説評価課題において呈示された事例の検証事例としての適切性の評価は、2-4-6問題では確証型検証の評価が反証型検証の評価を大きく上回っているのに対して、仮想生物課題では反証型検証の評価も高く、反証型検証の適切性の評価が確証型検証のそれを上回っている場合も観察される。このような傾向は両課題のもつ構造的差異が強い影響を与えていると考えられるが、推測の域を出ない。この問題は、今後の検討の余地を残す問題であろう。

7.4 第7章のまとめ

本章では、第1～第4実験で用いられた仮想生物課題と、第5、第6実験で用いられた2-4-6問題という2つの課題において、事例選択傾向、得られた情報に対する評価が異なる原因を、両課題間の構造的差異に注目した検討を行った。上記の2つの課題における違いとして重要なことは、領域固有の知識体系の有無、および事例空間の大きさである。仮想生物課題は、生物学という領域固有の知識体系が背景にあると共に、事例空間が有限の集合となっている。

一方、2-4-6問題には、背景となる知識体系が存在せず、事例空間も無限の広がりを持っている。そのため、特に後者では、反証事例の想起しやすさという点では極めて困難な課題であり、また、背景知識がないためにヒューリスティック的な方略が使われやすいと思われる。そのために、前者をベースとした第1～第4実験では、比較的反証型検証が選択されやすく、後者をベースとした第5、第6実験では、確証事例が選択されやすかったのではないかと推測される。さらに、この事例選択傾向に対応する形で、得られた情報に対する評価が決定され、前者では特に中盤以降で反証型検証に対する適切性評価が高くなるが、後者では確証型検証が適切であると評価されたのではないかと考えられる。従って、課題構造という要因も、検証方略の選択に大きな影響を与えていると考えられる。一方、第2～第4実験の結果が示すように、少なくとも仮想生物課題においては、課題が進行するにつれて検証事例の選択パターンが変化しており、一般的な傾向としての事例選択のシフトという要因も大きな影響を持っているといえる。また、第2実験の正事例系列条件に比べて、第3実験の正解到達条件の方が、反証型検証の選択率が高いという結果が得られており、反証型仮説否定情報の呈示による反証型検証の有効性認識の強化という要因も無視し得ない。このように、仮説検証方略の選択においては、複数の要因が複雑に絡み合った形で作用していると考えられる。

第8章 総合的考察

8.1 本研究で得られた知見とその意義

人の仮説検証は、概ね以下のような過程をたどるといえる。まず、(1) 既存の知識体系からでは説明不可能なデータの存在が認識されることで、仮説が生成される。この仮説の生成に際して、さらなるデータ収集が行われる場合もありうる。続いて、(2) 仮説の真偽を決定するためのテストが行われ、具体的には、仮説から演繹された事例が仮説の予測どおりの結果を導くかどうかの確認がなされる。この検討は、科学的探求における実験のように検証者が積極的、かつ自発的に情報収集を試みる場合もあるが、単に外部から与えられる情報を観察するだけの場合もある。(3) 次に、この検討の結果から、必要に応じた仮説の更新がなされ、以降このテストと更新のフェーズが繰り返される。従来の仮説検証研究の多くが、前述した過程のうちの(2)、すなわち仮説の真偽を決定するためのテストの実行段階に注目した検討を行ってきた (Klayman & Ha, 1987; Wason, 1960)。ここで特に問題とされてきたのは、検証事例の選択の仕方であり、人が現在保持している仮説の確証事例による検証を多く行う傾向があることが示されてきた。

8.1.1 仮説検証過程の動的側面

本研究において指摘してきた通り、仮説検証に関する従来の研究では、本来ダイナミックに進行する認知過程を全体的、スタティックに記述するという問題

点を持っていた。その問題点を解決するために、本研究では動態的変容分析パラダイムという考え方にに基づき、検証過程が時系列に沿ってどのように進行するかを記述すること、すなわち仮説検証過程の動的側面をとらえることを第1の目的としてきた。本研究の第1実験においては、生物学の話題を題材にした架空の規則発見状況を設定した仮想生物課題を用い、被験者の反応パターンを検討した。その結果、全体的には確証型検証が多用されるものの、1/6程度の試行では反証型検証が用いられることが示された。しかし、両者の選択比率は時系列に沿って変化するというよりも、常に一定の割合で選択されること、正解到達者と未到達者の間でも特に差はないことが示された。一方、続く第2・第3実験で用いられた受容方略型課題、仮説評価課題では事例選択傾向が大きく変化し、特に中盤以降で、反証型検証の方が多く選択される場面が見うけられた。ただし、反証型検証の選択は呈示される情報の系列に依存し、非ターゲット事例（負事例）のみから構成される系列では、むしろほとんどの選択が確証型検証であった。

さらに、先行的に呈示される情報を操作した第4実験では、呈示情報の種類に関わらず、反証型検証の選択が比較的多く、また終盤に向かうに従って反証型検証の選択がわずかに増加するという傾向が見られた。これらの結果は、まず仮説の蓋然性を高める段階から、次に見落としがないかなどを確認するための洗練の段階へと移行する検証段階のシフトに伴い、事例選択がダイナミックに変化したものとして解釈することができるであろう。仮説の蓋然性を高めるには、仮説を支持する情報を獲得することが必要であり、そのためには、確証

型検証が適しているといえる(c.f. Klayman & Ha, 1987). 一方、仮説を洗練されたものにするためには、確証型検証に加えて反証型検証を併用することで、仮説が必要かつ十分であることを確かめなくてはならない。そのために、特に検証過程の中盤以降において反証型検証の選択が増加することが予測され、第4実験の結果もこの予測に沿った傾向を示している。このことから、本研究の主要な目的である仮説検証過程の動的側面をとらえる目的は達成されたのではないかと考えられる。

一方、検証方略選択における段階的变化という傾向の普遍性を確かめるために、2-4-6問題の評価版を用いて検討した第5実験では、第2～第4実験の結果と一転して確証型検証の選択比率が高いという結果が見られた。また、この結果は2-4-6問題の評価版課題を用いた Kareev & Halberstadt (1993)の結果と極めて良く似た傾向を示しているといえる。このように2-4-6問題を題材にした課題とそれ以外の課題で見られた違いは、2つの課題、すなわち第1実験～第4実験で用いられた仮想生物課題と、第5実験で用いられた2-4-6問題の間に存在する構造的な差異によると思われる。これら2つの課題の間には、大きく分けると2つの違いが存在すると考えられる。まず第一に、仮想生物課題は、発見すべき規則は恣意的なものではあるが、生物学的常識という領域固有の知識体系が背景に存在する課題であるのに対し、2-4-6問題では、そのような領域固有の知識が通常は存在しないと考えられる。言いかえるなら、前者はより現実的、具体的な領域での仮説検証課題であるが、後者は極めて現実性に乏しい人工的な課題であるといえる。

第二に、仮想生物課題と2-4-6問題では、潜在的な検証事例の集合の大きさに極めて重大な違いがあることを指摘しておく必要があるであろう。前者では、事例空間は複数の次元から構成されるものの、特定の次元がとりうる属性値の範囲は極めて限定的である。一方、2-4-6問題では、属性値の範囲は無限大の広がりをもっている。そのため、2-4-6問題では反証事例の想起のしやすさ、あるいは利用可能性 (availability: Kahneman, et al., 1982; Tversky & Kahneman, 1973, 1974) に差が生じ、事例選択傾向が異なったのではないかと思われる。以上指摘したような領域固有的知識体系の有無、事例集合の大きさの違いという要因は、現時点では推測の域を出ないが、今後の重要な検討項目となるであろう。

また、上記の確証事例・反証事例という基準による事例選択の分類とは異なった観点から、検証方略の段階的变化を記述することも重要な問題であろう。本研究では、第4実験においてそのような分析を試みた。第4実験では、3次元で構成される事例空間を持つ規則発見課題が用いられていたため、事例選択の基準として、仮説中に言及されている次元 (関連次元) のみに注目したものであるか、仮説中に言及されていない次元 (無関連次元) にも注目したものであるかという分類を採用した。関連次元のみに注目した事例選択は、被験者がターゲット規則を構成する可能性が高いと考える次元に対して確証型検証を試みていると考えられ、これは、仮説の蓋然性を高めるために行われていると推測される。一方、無関連次元を含む事例選択は、仮説に見落としがないかどうかを確認し、仮説を洗練されたものにしようとする試みとして解釈することが

できる。第4実験の結果から、全体的には関連次元のみに注目した事例選択が多いものの、無関連次元に注目した事例選択が後半の試行でより多くなること
が示されており、蓋然性を高める仮説構築の段階から見落としを確認する仮説
洗練の段階へと向かう、検証過程の段階的移行が存在していることを示唆して
いると言えるであろう。

8.1.2 受動的情報収集手続きの検証方略選択に対する影響

前述した通り、少なくとも第1～第4実験で用いた仮想生物課題においては、反証型検証の選択が終盤に向かうにつれて増加するという傾向が見られた。
特にこの傾向が顕著であったのは、第2、第3実験で用いられた受容方略型課題、
仮説評価課題である。5次元の仮想生物課題を用いて、通常の規則発見課題
を行った第1実験では、全体の7割強が確証型検証であったのに対し、第1
実験のデータを下に構成された第2実験の受容方略型課題では、確証型検証と
反証型検証の選択数がほぼ同数であり（正事例系列条件）、同じく第3実験の
仮説評価課題では、反証型検証の選択数が確証型検証の選択数を上回っていた
（正解到達条件）。受容方略型課題と仮説評価課題の手続き上の共通点は、両課
題が共に受動的な情報収集手続きを含み、特に検証事例が被験者による自発的
選択が行われないという点にある。また、2つの課題には、前者では被験者自
身が仮説を生成しなくてはならないが、後者では仮説も外部から呈示されると
いう相違点もある。そのため、結果として呈示された事例が、その時点での仮
説とどのような関係になるかがやや異なっていた。そのため、前者の受容方略

型課題よりも、後者の仮説評価課題の方が、結果として多くの反証型仮説否定 (false negative) を得ることとなり、仮説評価課題において反証型検証の有効性がより強く認識されたと考えられる。これにより、第2実験の正事例系列条件よりも、第3実験の正解到達者条件において、さらに反証型検証を選択する傾向が強くなったと考えられる。これは、検証方略選択において、検証段階の移行に伴う方略の変化から予測される確証事例・反証事例の選択傾向にオーバーライドする形で、反証型仮説否定情報による反証型検証の有効性認識の強化という要因が働いたと考えられる。

本研究の結果から、ダイナミックに進行するプロセスである仮説検証の理解においては、本研究で提唱した動態的変容分析パラダイムの考え方が適しているといえるであろう。ただし、仮説検証方略の選択は、確証型検証を重視する仮説構築段階→反証型検証も併用する見落とし確認段階という極めて単純な形式のシフトのみで説明されるものではなく、反証型仮説否定呈示による反証型検証の有効性認識の強化、領域固有の知識体系の有無、反証事例の想起のしやすさ等の複数の要因が複雑に絡み合ったプロセスであると考えられる。今後の発展的研究において、このモデルをさらに洗練させることによって、仮説検証過程の統一的な理解が進むことであろう。

8.1.3 事例選択方略モデルの精緻化

本研究の結果、検証事例の選択は、検証過程を通じて段階的に変化する可能性が示唆された。この変化は、1 つには、確信度がある程度高まった仮説に関す

る見落としを確認するという形で生じると考えられる。これは、検証の初期段階と後期段階では、情報の認識的効用 (epistemic utility) が異なることによると言いかえることもできる。

検証事例の選択が認識的効用によって決定されたと考えたのは、Kirby (1994)をはじめとする最近の 4 枚カード問題を用いた一連の研究である (Evans & Over, 1996; Hattori, 1999; Kirby, 1994; Oaksford & Chater, 1994b, 1996; Yama, 1999)。これらの研究に共通する点は、4 枚カード問題におけるカード選択傾向は、カードを裏返したときに得られることが期待される情報の価値に対する評価によって説明可能であると考えることにある。例えば、Kirby (1994)、あるいはそれを発展させた Evans & Over (1996) は、4 枚カード問題において *not-q* カードが選択されないのは、通常 *p* 集合の大きさが *not-p* 集合より小さいため、*not-q* カードを裏返して仮説を否定する情報を獲得できる確率の主観的評価値が低いからであると考え、その主張を裏付ける実験を行い、認識的効用説 (epistemic utility theory) に対する支持を得ている。また、Oaksford & Chater (1994b) は、カードの選択が、不確実性の低減 (uncertainty reduction) を目的に行われると主張した。彼らは、Shannon-Weiner の情報量に基づいて仮説の不確実さを推定し、またベイズの定理を導入することで、あるデータを得た時に仮説の不確実さがどの程度減じられるかを定式化した。そして、4 枚カードにおける *p*, *not-p*, *q*, *not-q* のカード選択のそれぞれについて、この式に基づいて期待される不確実性の低減の度合いを計算した。また、このモデルの予測と実際の被験者の反応とを比較し、カード選択の頻度が期待情報獲得量と単調的

関数関係にあることを見出した。モデルと反応との対応は、オリジナルの4枚カード問題に限らず、Kirby (1994) が用いた p カードの確率を変化させた課題、さらには、「義務」スキーマが活性化されるような具体的状況を用いた課題においても確かめられた。

Oaksford & Chater (1994b) が提唱したこの最適データ選択 (Optimal Data Selection: ODS) モデルは、2 - 4 - 6 問題をはじめとする規則発見課題のパフォーマンスに対しても適用可能であるが、実際に適用して検討を行った研究は、これまでではない。また、ODS モデルでは、情報の認識的効用が、検証過程の途中で変化するということを予測していないが、モデルのパラメータとして、情報の必要度や優先度といったようなものを付け加えることによって、本研究で得られた結果についても説明が可能であるかもしれない。これは、今後の検討の余地を残す極めて重要な問題であると考えられる。

また、検証方略選択モデルの問題について、従来の心理学的研究が規範的モデルとして想定していた反証主義に対する批判が、科学哲学の領域からも提唱されている。反証主義の主要な問題点としては、補助仮説の取り扱いの不備 (Putnam, 1974; 高野, 1994; 内井, 1995)、確率・統計的命題の取り扱いの不備 (内井, 1995) などが挙げられている。特に重要なのは、確率・統計的命題の取り扱いの不備であると考えられている。近代科学の発展を促した大きな要因の一つに、確率論の導入があることは論を待たない。実際に、本研究における実験や、その他の心理学研究、社会学や経済学などの社会科学、物理学や化学をはじめとする自然科学、またそれらを応用した工学など、非常に多くの領域で確

率論的な考え方が利用されている。

この問題を解決するために、科学哲学の領域からは確率論的帰納法というモデルが提唱されている (e.g. 内井, 1995)。確率論的帰納法は、科学的知識が本質的な不確実性をまぬがれないことをはっきりと認識した上で、不確実性の計量を目指すモデルであり、帰納法および仮説演繹法を確率論の枠組みから再定義し、手持ちのデータに照らして与えられた仮説の確からしさを求めるといふ、ベイズの定理に基づいた仮説検証である。確率的帰納法においては、仮説に対する確証・反証は程度の問題であるが、ある程度以上の事後確率を得た時に検証手続きを終了し、仮説を受容するという規則を導入することは可能である。また、論理的命題についても、確率的命題の特殊事例として扱うことが可能である。また、確率論的帰納法は、確率評価が人の持つ主観的基準を下回る時に、仮説生成過程が始動することも予測する (眞嶋・瀧川, 1997)。以上のことから、現実的な場面での科学的探求において反証主義を規範的なモデルと置くことには大きな問題があり、むしろ確率論的帰納法の方が多くの点でより実用的なモデルであるといえる。

前述したような ODS モデルを元にした検証方略選択のモデル化は、基本的な考え方を確率論的帰納法と共有したものであるといえる。今後の発展的研究においては、ODS モデルをはじめとする認識的効用や、確率論的帰納法の概念を導入して、より妥当性の高い記述的モデルを構築することが重要となると考えられる。その際には、本研究において指摘された、反証型仮説否定情報の獲得に伴う反証型検証の有効性 (認識的効用) の強化や、領域固有知識の有無、

反証事例の想起のしやすさといった要因による方略の変化を含めた統合的なモデルの構築が求められるであろう。

8.1.4 人の仮説検証方略の選択における適応的合理性

仮説検証に関する近年の研究は、それまで規範的モデルから逸脱した非合理的な推論として扱われてきたこの認知的プロセスを、適応性を重視した営みとして積極的かつ肯定的に評価しようとしてきた (e.g. Friedrich, 1993; Klayman & Ha, 1987)。本研究も、この適応論的アプローチの立場に立ち、方略選択を規定する要因の探求を試みた。本研究の結果が示すのは、仮説検証方略の選択が、検証段階の移行に伴って、また獲得情報の種類による確証・反証型検証の有効性に対する認識の変化や課題構造の違いに応じて変化する、きわめて柔軟性に富んだプロセスであるということである。現実場面で仮説検証が行われる状況としては、まず第一に科学的探求が挙げられる。近年、科学的探求過程に関する研究では、従来通りの実験室的な *in vitro* アプローチをとる研究だけに頼るのではなく、科学者が現実に行われる科学的探求場面でどのように推論するかを調査・観察する、ある種フィールドワーク的な *in vivo* アプローチをとる研究の重要性が指摘されている (e.g. Dunbar, 1995)。本研究は、上記の2つのアプローチのうち、実験室的な *in vitro* アプローチによって、方略選択の段階的变化という動的側面をとらえることを試みた。仮説検証の動的側面に対する注目は、*in vivo* アプローチにおいても、今後重要な問題となっていくであろう。例えば、前述の Dunbar (1995) では、新たな発見を導く要因として、研究グル

ープのメンバーのそれぞれが、部分的に重なり合う研究背景を持つこと、類推をうまく使用すること、研究の成果について議論する場を設けることなどを挙げているが、これらは最終的な結果として新しい理論や現象の発見に至った場合と、そうでない場合とを全体的、スタティックに比較した結論として挙げられたものである。その意味では、科学的探求に対する *in vivo* アプローチにおいても、プロセスの動的側面に対する考察が十分であるとは言えない。現実場面における科学者（あるいは科学者集団）が、どのように仮説を立て、それをどのように洗練させていくかというダイナミズムを記述することは、仮説検証過程、科学的探求に対するより良い理解において重要な役割を果たすことであろう。本研究は、今後の研究の方向性に対して有意義な知見を提供することができたのではないかと思う。

また、仮説検証は上記に挙げたような科学的探求状況に限定されるものではなく、日常的な信念形成においても決定的な役割を果たすと考えられる。人が日常的に行うさまざまな判断、推論、意思決定では、人がさまざまな誤りを犯すことが指摘されており (e.g. Kahneman, Slovic, & Tversky, 1982)、思考において陥りやすい誤りを回避する脱バイアス技法の開発は、思考心理学の重要な研究対象である。本研究や、今後の発展的研究において得られた知見を、判断、推論、意思決定の分析に応用することで、これらの心的活動に対するより良い理解が進み、結果として脱バイアス技法の開発にも大きな貢献となるであろう。

本研究では、既に述べた通り、仮説検証方略のダイナミックな変化や、情報を獲得する手続きの違いによる方略選択の変化について議論した。本研究にお

いて、検証方略の選択に影響を与える全ての要因について十分に検討し尽くしたとは言えないであろうが、今後のこの領域における発展的研究の方向性について重要な示唆を与えることができたと考えられる。本論文の結びとして、本研究においては検討の対象とはならなかった仮説検証研究の諸問題を概観し、今後の研究発展の可能性を模索したい。

8.2 未解決の問題と今後の課題

8.2.1 協調的解決過程の検討

本研究の発展として位置付けられる問題として第1に指摘すべきは、協調的解決、ないし協調的発見過程 (collaborative discovery process) の検討である。協調的解決、あるいは協調的発見とは、複数の人間が問題の解決に従事する、特にここでは、複数の人間が仮説検証過程を共に行う状況である。協調的発見過程は、ここ近年、世界的に注目されてきている問題である (c.f. Dunbar, 1995; 亀田, 1996; Laughlin & Futoran, 1985; Laughlin & Hollingshead, 1995; Miwa, 1999; Okada & Simon, 1996)。協調的発見過程の研究において問題となるのは、複数のメンバーが課題に従事した場合、単独で行う場合と比較して、課題解決が促進されるかどうか、またどのような条件が揃えば、この促進が見られるのかということである。メンバー間が協調的に働く場合に、課題解決が促進される現象は、創発 (emergence) と呼ばれている。

例えば、Laughlin とその同僚による研究では、トランプを用いた規則発見課題を用いて、グループと個人の課題遂行を比較する試みがなされている

(Laughlin, et al., 1995; Laughlin & Futoran, 1985; Laughlin & Hollingshead, 1995).

彼らは、トランプの属性 (色, 種類, 数) や, その組み合わせに関する規則 (例えば, 「赤で 6 以上または黒で 4 以下のカード」) を発見させる規則発見課題を用い, グループによる解決は個人による解決と比べて成績が良くなる (創発性), いわば「三人寄らば文殊の智慧」という現象が生じるかどうかを検討している. しかし, 彼らの研究の結果からは, グループが個人に比べて成績が良いとは言えず, グループによる創発性は観察されなかった.

一方, Okada & Simon (1996) は, Dunbar (1993) による分子遺伝学実験を, 2 人で協同して解決させるという実験を行い, グループの創発性について詳細に検討した結果, グループによる創発性現象が見られることを確認している. この実験課題は, Jacob & Monod (1961) によって実際に発見された, 複数の遺伝子によるラクトーゼ分解酵素の制御メカニズムを, コンピュータ上の実験を通して発見させるというものであった. Okada & Simon (1996) は, この課題解決の過程を, 個人で解決する場合と, 友人とペアで解決する場合とに分け, 両者を比較した. この実験の結果は, ペアの方が個人よりも成績が良く, 協調的解決による創発性現象が認められた. Laughlin らによる一連の研究と, Okada & Simon (1996) による研究とで創発性の有無に違いが見られたことについては, 集団の大きさの違い, 集団成員の性質の違いという 2 つの説明が考えられている (亀田, 1996). 集団の大きさという点については, Laughlin らの研究では 3 人以上の比較的大きな集団であったのに対し, Okada & Simon (1997) では 2 人ペアという違いが存在する. 3 人以上の集団では, 多数派对少数派といったよ

うな集団力学的要因が働くが、2人ペアではそのような要因は存在しない。一方、集団成員の質については、Laughlinらの研究で集団を形成したメンバーが、実験開始時点では全く見ず知らずの者であったのに対し、Okada & Simon (1996)における2人ペアは、友人同士という関係にあり、お互いの専門性や知識に関するメタ知識が存在したと考えられる。見ず知らずのメンバーで構成された Laughlinらの研究での集団では、このようなメタ知識はほとんど存在しないと考える良い。メタ知識は、効率的な認知的分業体制を可能にし、グループの遂行を高めると考えられる。

また、Miwa (1999) は、2-4-6問題を題材にして、2人の人間が課題解決にあたった場合に、両者がどのような関わり方をした時に創発性が生じるかをコンピュータシミュレーションによって検討した。Miwa (1999) は、特に、ペアで問題解決にあたるメンバー間でどの程度の認知的空間 (cognitive space) の共有の度合いによって、創発性の程度に差が見られるかどうか注目した。彼は、(1) ペアが実験空間 (experimental space)、すなわち検証事例とその結果を共有する場合、(2) ペアが実験空間に加え仮説空間 (hypothesis space) を共有する場合、(3) 実験空間と仮説空間が共有されるだけでなく、両者が統合される場合とを比較し、認知的空間の共有度が高くなるに従って課題解決が促進されることを見出した。例えば、実験空間のみを共有する (1) の状況では創発性はほとんど生じなかったが、仮説空間も共有する (2) では、一部条件付ながら創発性は生じ、認知的空間の共有度が最も高い (3) では、高い創発性現象が確認された。

本研究での検討内容と、この集団による協調的解決との関連で考えると、以下のような発展的研究が考えられる。まず第一に、仮説評価課題と規則発見課題とを組み合わせ、両者を同時に進行させるという実験手続きが考えられる。例えば、2人ペアを組み、1人の被験者はこれまで通りの規則発見課題に従事させる。そして、もう1人の被験者には、実際に進行している規則発見課題における反応に対して評価をさせる。この時、規則発見課題に従事する被験者に対して、評価課題に従事した被験者の反応をフィードバックするかどうかを操作し、単独で解決する場合と比較して成績がよくなるかどうかを検討する。この手続きによって、認知的分業体制が確立された状況と、単独で課題解決にあたる状況との比較が可能となり、集団による創発性の有無の検討が可能となる。また、評価課題に従事する被験者についても、実際に評価対象者と相対して評価を下す場合と、匿名性が確保された状態で評価を下す場合とで反応パターンが変化するかどうかを検討するということも、社会心理学的に非常に興味深い問題である。

8.2.2 多面的な観点からの検証過程の検討

仮説検証研究における最大の注目点は、人がどのような方法をもって仮説を検証するか、すなわち仮説検証における方略使用である。この問題については Wason (1960) 以来さまざまな研究が行われてきたが、今日に至るまでのほとんどの研究では、どのような事例選択を行うかによって、仮説検証方略を記述しようとする試みが行われてきた (Klayman & Ha, 1987, 1989; Mynatt et al.,

1977, 1978; Oaksford & Chater, 1994b; Tukey, 1986; Tweney et al., 1980; Wason, 1960, 1966). しかし、広義の仮説検証過程は、仮説の生成、検証事例の選択、フィードバックの獲得と解釈、必要に応じた仮説の修正などを含む、極めてダイナミックなプロセスであり、事例選択傾向のみから検証方略を定義することは、非常に限られた一側面にしか注目していないことになる。

近年、これまでとは異なったパラダイムによって検証過程を記述しようとする試みが始められている。例えば、Klahr & Dunbar (1988) は、科学的発見状況における問題空間を仮説空間 (hypothesis space: H-space) および実験空間 (experimental space: E-space) に分け、それぞれの空間の探索を重視する型として、理論家 (theorists) と実証家 (experimenters) を挙げている。理論家とは、データを説明できる仮説がなくなるまで実験空間を探索しない、仮説空間内の探索に重点を置く、すなわち理論的考察を重視するタイプである。一方の実証家とは、初期仮説は立てるものの、それが失敗に終わると特定の仮説を立てずに事例収集を行なう、実験空間内の探索を重視する、すなわち実験による多数の事例情報収集を重視するタイプである。また、Halberstadt & Kareev (1995) は、受動的情報収集を多用する受容型 (getters) と、能動的事例生成を多用する生成型 (builders) の2つのタイプを挙げている。Halberstadt & Kareev (1995) は、仮説検証におけるデータ収集のプロセスとして、現実場面における実験に相当する生成パラダイム (generation paradigm) と、調査・観察などに相当する受容パラダイム (reception paradigm) の2つがあることを指摘し、従来の研究で用いられてきた実験課題 (e.g. 2 - 4 - 6 問題: Wason, 1960, 人工空間課題:

Mynatt et al., 1977, 1978, 分子遺伝学課題: Dunbar, 1993) が、課題手続き上、生成パラダイムでの反応を被験者に強いていたという問題があると述べた。彼らは、仮説検証研究において指摘されている認知的制約、例えば Evans (1989) の指摘した肯定的情報への偏向 (positivity bias) といったものも、このパラダイムの固定に原因があると考え、生成・受容の両パラダイムを被験者が自発的に選択できる実験課題を用いた。その結果、受容パラダイムは生成パラダイムより先に選択されること、また受容 → 生成へのシフトは正反応 (positive instance) を得た後に生じやすいことなどを見出した。また、このような全体的傾向とは別に、受容パラダイムを多用する受容型 (getters) と、生成パラダイムを多用する生成型 (builders) という 2 つのタイプの被験者が存在することを示した。

また、本研究では、第 1, 第 2 実験において、被験者が複数の仮説を生成したかどうか、複数の仮説を生成した場合はどの仮説に注目したかによって、Single, Multiple-Main, Multiple-Alternative という 3 つのタイプに分類し、各被験者がこの 3 つのタイプをどのように使い分けたかを検討した。そして、規則を発見できた被験者では、Single, すなわち仮説を一つだけ挙げる試行が、後半になるほど多くなり、うまく仮説を絞り込めていることが示された。

事例選択傾向に注目した従来の研究は、特定の仮説の真偽を決定するための方法に焦点を絞った研究であると考えられる。一方で, Klahr & Dunbar (1988), Halberstadt & Kareev (1995) や、本研究などは、仮説の生成過程や、生成された仮説の取捨選択の仕方といったプロセスを視野に入れた方略の分類を行なっ

ているといえる。仮説検証方略を分類、整理する規準は様々なものがあり、現実には用いられる検証方略はある特定の規準のみから記述されるべきではない。例えば、Halberstadt & Kareev (1995) における生成型でも、確証事例を多用する者や確証・反証事例を併用する者がいるだろうし、受容型、生成型の両方において消去法な検証を行なうことも可能と思われる。今後の発展的研究においては、検証方略を特徴づける規準を整理した上で、様々な検証方略をその規準に基づいて分類すること、それぞれの方略がどのような場面で用いられ、またどのようなきっかけで方略が変化するかの境界条件を探求することが重要となるであろう。

8.2.3 仮説生成過程の探求 — 科学における発見の文脈

前述の問題とも関連するが、今後の発展的研究における重要な問題として、仮説生成過程の検討を挙げておく必要があるだろう。仮説生成は、仮説検証、あるいは帰納的推論過程を始動する契機となるプロセスであり、当然のことながら、仮説が生成されない限り、いかなる検証プロセスも生じ得ない。また、この生成段階でどのような仮説を立てるかによって、後続の検証過程が大きく影響を受けることは明らかである。例えば、発見すべき規則（現実世界における真実）に近い仮説を立てることができたならば、後の検証過程は非常に少ないステップで済む事が期待される。逆に、真実とはかけ離れた仮説を立ててしまった場合、真実に到達するには極めて多くのステップを要するであろう。仮説の生成には、既存の知識体系が利用され、特に類推的思考が重要な役割を果た

すと考えられる。例えば、本研究の第1～第4実験で用いた仮想生物課題では、架空の状況であるとは言え生物学という現実的な背景が存在するのに対し、第5、第6実験で用いた2-4-6問題にはそのような背景が存在しない。従って、2-4-6問題では既存の知識体系が利用しづらい、あるいは全く利用できないのに対し、仮想生物課題では既存の知識体系から、生物学に関する知識を動員し、仮説を生成することが可能であるという違いが生じる。すなわち、仮想生物課題の解決にあたっては、良くも悪くも既存の知識体系が影響を与えるのである。例えば、発見すべき規則が既存の知識体系と矛盾しないようなものである場合は、仮説の生成段階において、より正解に近い仮説を構築することが可能となり、少ないステップで正解に到達することが可能になるであろう。一方、既存の知識体系と矛盾する規則を発見しなければならない場合は、逆に知識が課題解決を阻害するかも知れない。

このように仮説生成プロセスは、極めて重要な役割を果たすものであるが、科学哲学においては、このプロセスに関する検討はなされてこなかった。科学哲学では、科学的探求過程を理論や、仮説を生成・構築する発見の文脈と、それらを検証する正当化の文脈を区別し、後者のみが科学哲学の対象であるとする(戸田山, 1999; 内井, 1995)。仮説を検証し、その真偽を決定する方法は論理的であるが、その仮説を考え出すプロセスは、論理的飛躍を含み、必ずしも正しいという方法が存在しない。また、前述の通り、確率論的帰納法の主張に従えば、どのような初期仮説から初めても、最終的には十分な検証を重ねることによって、正しい規則へと収束することは十分可能である。むしろ、理論や

仮説を思いつく生成のプロセスは心理学的問題であると考えられている (戸田山, 1999)。しかし、これまでの思考心理学の研究では、人が既存の情報からどのようにして有意義なパターンを抽出し、仮説の構築に利用するかについては全く研究がなされていないといってよい。既に述べた通り、仮説の生成には、既存の知識体系が利用され、特に類推的思考が重要な役割を果たすと考えられる。時には、科学革命 (Kuhn, 1962) のように、全く新しい理論ないし説明体系がドラスティックに構築されることもないわけではない。しかし、日常的に多く行われる仮説の生成とテストは、既存の知識体系と少々沿わない結果を説明するために、知識体系に多少の変更を加えるという目的の下で行われる場合が多いであろう。その目的を達成するために、個人は、既に経験によって得た知識を基に推論を進め、実際の観察結果との矛盾を解消するように仮説を検証し、知識を更新していくであろう。いずれにせよ、仮説の生成過程に関連する問題は、現在のところほとんどが手つかずの状態であり、今後の発展的研究が期待される領域である。

8.2.4 検証方略選択における個人差

最後に指摘しておくべきは、個人差に関連する問題である。本研究に限らず、過去の研究 (Klayman & Ha, 1987, 1989; Tweney, et al., 1980; Wason, 1960) のほとんどにおいて取り上げられてこなかった問題が、この個人差である。例外的に、Klahr & Dunbar (1988) や、Halberstadt & Kareev (1995) などが、被験者の反応をいくつかのカテゴリーに分類する試みを行っているものの、その検討は

未だ不十分であると言わざるをえない。検証方略の選択においては、各個人が持つ領域的知識体系のみならず、個人の認知的能力 (cognitive ability) や、性格特徴などが影響を与えていると推測されるが、Klahr & Dunbar や Halberstadt & Kareev においても、個人差を生じた要因については検討されていない。

最近、思考における個人差の問題について、Stanovich & West (1998) が、4枚カード問題、三段論法、ベースレートを用いた確率評価など思考心理学で伝統的に用いられてきた課題において観察される、規範からの逸脱傾向について検討し、個人の持つ認知的能力と、思考向性 (thinking dispositions) が、個人差と相関関係にあることを見出している。今後は、このアプローチを、仮説検証過程の分析にも応用し、検証方略選択における個人差を検討することも重要な問題であるといえよう。

8.3 結語

本研究における最も重要な目的は、仮説検証における方略使用の変化を時系列に沿って分析し、この認知的過程が課題状況や獲得された情報に対応する形でダイナミックに変化していくプロセスを記述することであった。これまでの研究においては、全体として仮説に対する確証事例と反証事例がどの程度の頻度で用いられるかを分析し、検証過程をスタティックにとらえることがほとんどであった。しかし、仮説検証とはその過程で得られた情報を基に仮説を更新する極めて柔軟性に富んだプロセスであり、全体的かつスタティックな分析を行うだけでは、この認知的過程の重要な側面を検討の対象から外してしまうこと

になる。本研究の結果、仮説検証過程は少なくとも、一番もっともらしい仮説を一つ構築する段階と、構築された仮説をさらに洗練させる段階の2つに分けて考える必要があり、それぞれの段階では検証事例の選択の仕方が異なることが示された。

本研究における第2の目的は、仮説検証研究に伝統的に用いられてきた規則発見課題とは異なった状況である、受容方略課題および仮説評価課題における被験者の反応を分析することであった。これらの課題は、外部から与えられた情報を観察するという、受動的情報収集手続きを共有している。特に両者に共通した検証事例の受動的獲得という手続きが、後の検証において反証型検証の選択率を増加させるように働いていることが示された。この傾向は、方略の段階的变化にオーバーライドする形で影響を与えていると考えられる。すなわち、人は基本的に確証型検証重視から確証型・反証型検証併用へと、事例選択がシフトする傾向を持っており、このシフトは確証型・反証型検証の認知的効用の変化によって引き起こされることが考えられるが、認知的効用の変化に対して、受動的な反証型仮説否定情報の獲得が関わっていると考えられる。また、本実験で用いられた実験課題間の比較によって、領域固有の知識体系の有無、あるいは反証事例の想起のしやすさという要因が、反証型検証の選択にさらに影響を与えていることが推測された。このような課題間の構造的差異が方略選択に与える影響については、今後の詳細な検討が期待される。

本研究の結果得られた以上のような知見は、人の仮説検証方略の選択が様々な要因に影響され、かつ当該の状況に対して適応的に対処するという方向性を

持ってなされるということを示しているといえる。この、「適応的に変化するプロセスとしての仮説検証過程」という観点は、今後の発展的研究においても重要なものとなるであろう。

仮説検証過程の統一的理解のためには、解決しなければならない問題は多くあるが、この認知的プロセスを時系列的に分析する必要性を指摘した本研究の成果は、今後の思考心理学の研究に対して有意義な知見をもたらしたといえる。今後は、思考心理学における研究をさらに発展させることは言うまでもなく、科学哲学、論理学、人工知能、教育学、社会心理学などの、思考心理学以外の領域の研究者の方々とともに研究を進めることが重要である。本論文を契機として、著者がこれらの方々との交流をもてるようになれば、これほど嬉しいことはない。

文献

- 東 洋 (1963). 概念学習におけるストラテジーの問題. 心理学評論, 7, 126-134.
- Baron, J. (1985). *Rationality and intelligence*. Cambridge: Cambridge University Press.
- Bourne, L. E., & Guy, D. E. (1968). Learning conceptual rules II: The role of positive and negative instances. *Journal of Experimental Psychology*, 77, 488-494.
- Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: Wiley
- Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-406.
- Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, 17, 397-434.
- Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg and J. E. Davidson (Eds.) *The Nature of Insight*. Cambridge: MIT Press.
- Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99, 3-19.
- Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Brighton: Lawrence Erlbaum.
- Evans, J. St. B. T. & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction. *Psychological Review*, 103, 356-363.
- Farris, H. H. & Revlin, R. (1989). Sensible reasoning in two tasks: Rule discovery and hypothesis evaluation. *Memory and Cognition*, 17, 221-232.
- Fischhoff, B. (1975). Hindsight $\neq$ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology, Human Perception and Performance*, 1, 288-299.

- Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology, Human Perception and Performance*, **3**, 349-358.
- Freedman (1992). Scientific induction: Individual versus group processes and multiple hypothesis. *Proceedings of the 14th annual meeting of cognitive science society*, 183-188.
- Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review*, **100**, 298-319.
- Gorman M. E., Stafford, A., & Gorman, M. E. (1987). Disconfirmation and dual hypothesis on a more difficult version of Wason's 2-4-6 task. *The Quarterly Journal of Experimental Psychology*, **39A**, 1-28.
- Griggs, R. A. & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, **73**, 407-420.
- Halberstadt, N. & Kareev, Y. (1995). Transition between modes of inquiry in a rule discovery task. *Quarterly Journal of Experimental Psychology*, **48A**, 280-295.
- Hattori, M. (1999). The effects of probabilistic information in Wason's selection task: An analysis of strategy based on the ODS model. *Proceedings of the Second International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference* (pp. 623-626). Waseda University.
- Hawkins, S. & Hastie, R. (1990). Hindsight: Biased judgments of past events after outcome are known. *Psychological Bulletin*, **107**, 311-327.
- Jacob, F. & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, **3**, 318-356.
- Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972) Reasoning and a sense of reality. *British Journal of Psychology*, **63**, 395-400.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics*

and biases. Cambridge: Cambridge University Press.

亀田 達也 (1996). 合議の知を求めて. グループの意思決定. 認知科学モノグラフ, 第3巻, 東京: 共立出版.

Kareev, Y. & Halberstadt, N. (1993). Evaluating negative tests and refutations in a rule discovery task. *Quarterly Journal of Experimental Psychology*, **46A**, 715-727.

Kirby, K. N. (1994). Probabilities and utilities of fictional outcomes in Wason's four card selection task. *Cognition*, **51**, 1-28.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation and information in hypothesis testing. *Psychological Review*, **94**, 211-228.

Klayman, J., & Ha, Y.-W. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 596-604.

Laughlin, P. R., Chandler, J.S., Shupe, E. I., Magley, V. J., & Hulbert, L. G. (1995). Generality of theory collective induction: Face-to-face and computer-mediated interaction, amount of potential information, and group versus member choice of evidence. *Organizational Behavior and Human Decision Processes*, **63**, 98-111.

Laughlin, P. R. & Futoran, C. G. (1985). Collective induction: Social combination and sequential transition. *Journal of Personality and Social Psychology*, **48**, 608-613.

Laughlin, P. R. & Hollingshead, A. B. (1995). A theory of collective induction. *Organizational Behavior and Human Decision Processes*, **61**, 94-107.

Lord, C. G., Ross, L., & Lepper M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, **37**, 2098-2110.

眞嶋 良全・瀧川 哲夫 (1997). 帰納的推論及び仮説検証過程に関する研究概観. *Technical Report*, No. 7, January, Department of Psychology, Hokkaido University.

Majima, Y. & Takigawa, T. (1999) Differences in strategy selection in hypothesis testing

- and evaluation. *Proceedings of the 2nd International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference*, (pp. 343-348). Waseda University.
- 眞嶋 良全・瀧川 哲夫 (1999) 受容方略型課題における仮説検証過程. 日本心理学会第 63 回大会発表論文集. 中京大学.
- Manktelow, K. I. & Evans, J. St. B. T. (1979) Facilitation of reasoning by realism: Effect or non-effect? *British Journal of Psychology*, **70**, 477-488
- McDonald, J. (1992). Is strong inference really superior to simple inference? *Synthese*, **92**, 261-282.
- 三輪 和久 (1999). 科学的発見における有効な仮説検証方略 — 計算機シミュレーションに基づく検討. 岡田, 田村, 戸田山, 三輪(編著). 科学を考える 人工知能からカルチュラル・スタディーズまで 14 の視点, (pp. 96-119), 京都: 北大路書房.
- Miwa, K. (1999). Collaborative discovery process by interactive production systems: Degree of sharing cognitive space and possibility of emergence. *Proceedings of the 2nd International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference*, (pp. 343-348). Waseda University.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, **29**, 85-95.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, **30**, 395-406.
- Oaksford, M. & Chater, N. (1994a). Another look at eliminative and enumerative behaviour in a conceptual task. *European Journal of Cognitive Psychology*, **6**, 149-

- Oaksford, M. & Chater, N. (1994b). A rational analysis of the selection task as optimal data selection. *Psychological Review*, **101**, 608-631.
- Oaksford, M. & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review*, **103**, 381-391.
- Platt, J. R. (1964). Strong inference. *Science*, **146**, 347-353.
- Poletiek, F. H. (1995). Testing in a rule discovery task: Strategies of test choice and test result interpretation. In J.-P. Caverall, M. Bar-Hillel, F. H. Baron, & H. Jungermann (Eds.), *Contributions to Decision Making I*, 335-349, Elsevier Science, B. V.
- Politzer (1986). Laws of language use and formal logic. *Journal of Psycholinguistic Research*, **15**, 47-92.
- Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books.
- Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills CA: Sage.
- Skov, R. B. & Sherman S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, **22**, 93-121.
- Sloviaczek, L. M., Klayman, J., Sherman, S. J., & Skov, R. B. (1992). Information selection and use in hypothesis testing: What is a good question, and what is a good answer? *Memory & Cognition*, **20**, 392-405.
- Slovic, P. & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology, Human Perception and Performance*, **3**, 544-551.
- Snyder, M., & Swann, W.B.J. (1978). Hypothesis testing processes in social interaction. *Journal of Personality and Social Psychology*, **36**, 1202-1212.
- Stanovich, K. E. & West, R. F. (1998). Individual Differences in Rational Thought. *Journal of Experimental Psychology: General*, **127**, 161-188

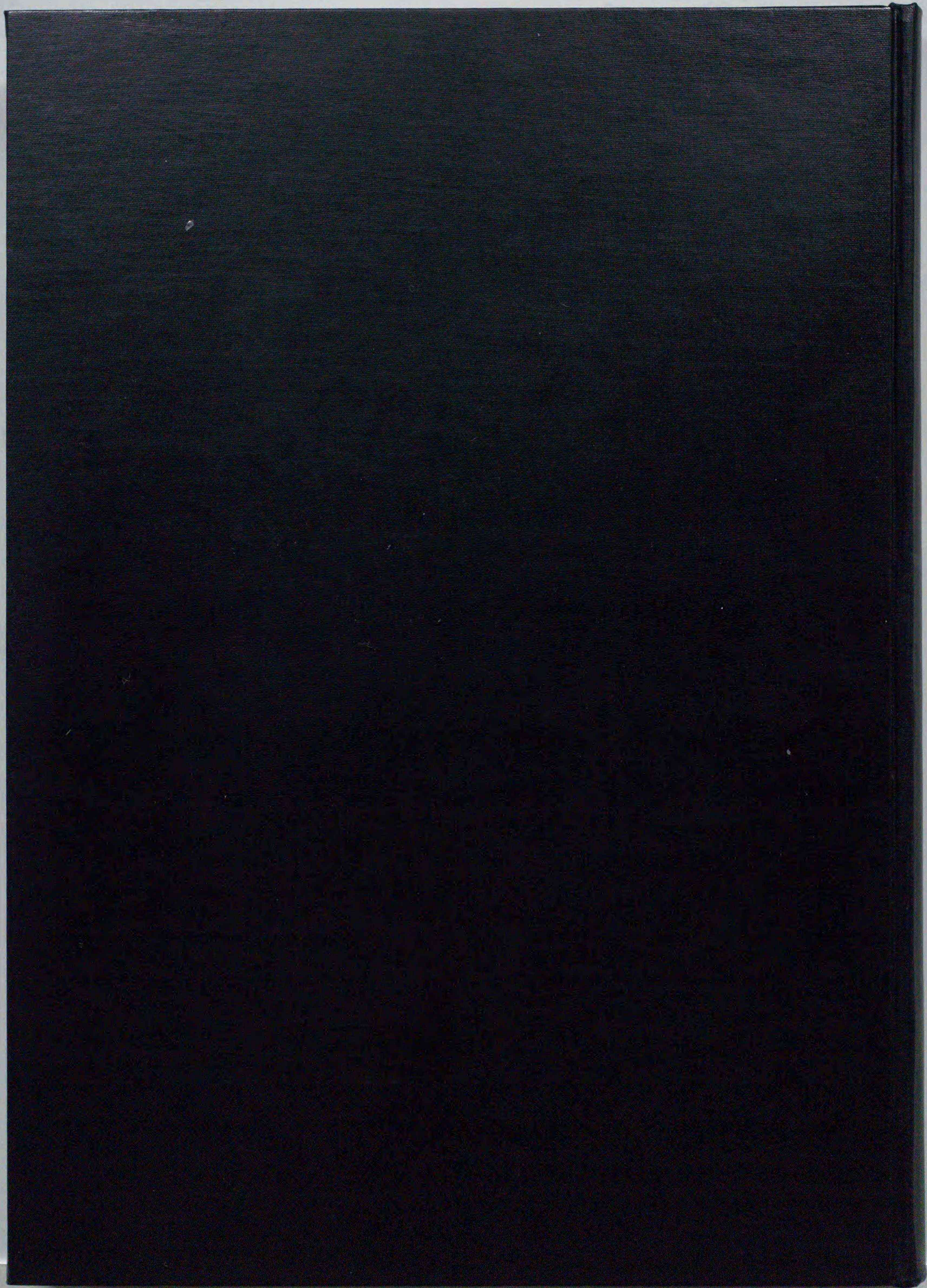
- 高橋 和弘・服部 雅史 (1996). 演繹的推論. 市川 伸一(編), 認知心理学, 第 4 卷, 第 2 章, 思考, (pp. 15-35), 東京: 東京大学出版会
- 寺岡 隆 (1966). 概念学習. 梅岡 義貴, 大山 正(編著), 学習心理学, 第 11 章、東京: 誠信書房.
- Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology*, **43**, 22-34.
- Trope, Y. & Bassok, M. (1983). Information gathering strategies in hypothesis testing. *Journal of Experimental Social Psychology*, **19**, 560-576.
- Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypothesis. *Child Development*, **51**, 1-10.
- Tukey, D. D. (1986). A philosophical and empirical analysis of subjects' modes of inquiry in Wason's 2-4-6 task. *The Quarterly Journal of Experimental Psychology*, **38A**, 5-33.
- Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **5**, 207-232.
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124-1131.
- Tweney, R. D., Doherty, M. E., Warner, W. J., Pliske, D. B., Mynatt, C. R., Gross, A., & Arkkelin, D. (1980). Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology*, **32**, 109-124.
- 内井 惣七 (1995). 科学哲学入門. 科学の方法・科学の目的. 京都: 世界思想社
- Yama, H. (1999). Probabilities and explicitness in the Wason selection task. *Proceedings of the Second International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference* (pp. 153-158). Waseda University.
- Wason, P. C. (1960). On the failure to eliminate hypothesis in a conceptual task. *Quarterly Journal of Experimental Psychology*, **12**, 129-140.

- Wason, P. C. (1966) Reasoning. In B. M. Foss (Ed.), *New horizons in psychology*. Harmondsworth, England: Penguin.
- Wason, P. C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, **20**, 273-281.
- Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of reasoning. Structure and content*. London: Batsford.
- Wasserman, D., Lempart, R. O., & Hastie, R. (1991). Hindsight and causality, *Personality and Social Psychological Bulletin*, **17**, 30-35.
- Weimer, W. (1977). Scientific inquiry, assessment, and logic: Comments on Bowers and Mahoney-DeMonbreun. *Cognitive Therapy and Research*, **1**, 247-255.
- Wetherick, N. E. (1962). Eliminative and enumerative behavior in a conceptual task. *Quarterly Journal of Experimental Psychology*, **14**, 246-249.
- Wharton, C., Cheng, P., & Wickens, T. (1993). Hypothesis-testing strategies: Why two goals are better than one? *Quarterly Journal of Experimental Psychology*, **46A**, 743-758.

謝辞

博士論文をまとめるにあたっては、諸先生方、および北海道大学文学部行動科学科基礎行動学専攻比較行動学講座（現心理システム科学講座）のメンバーに大変お世話になりました。特に、瀧川哲夫先生からは、ご多忙の中、本論文の着想から完成まで多くのご指導を賜り、公私共々さまざまな助言を頂きました。また、菱谷晋介先生には、本論文の草稿に目を通していただき、貴重なコメントを頂きました。著者が北海道大学大学院文学研究科の在学中に、同じく在籍されていた、鈴木修司さん（現、北海学園大学講師）、平真木夫さん（現、宮城教育大学講師）、増地あゆみさん（現、北海道大学文学部助手）、黒澤勝士さんとは、多くの議論を交わし、それが研究面において非常に有意義な刺激となりました。ここに記して感謝の意を表します。

最後に、実験実施にあたり、被験者の募集に協力した頂いた方々、実験に参加していただいた多くの方々にも御礼申し上げます。

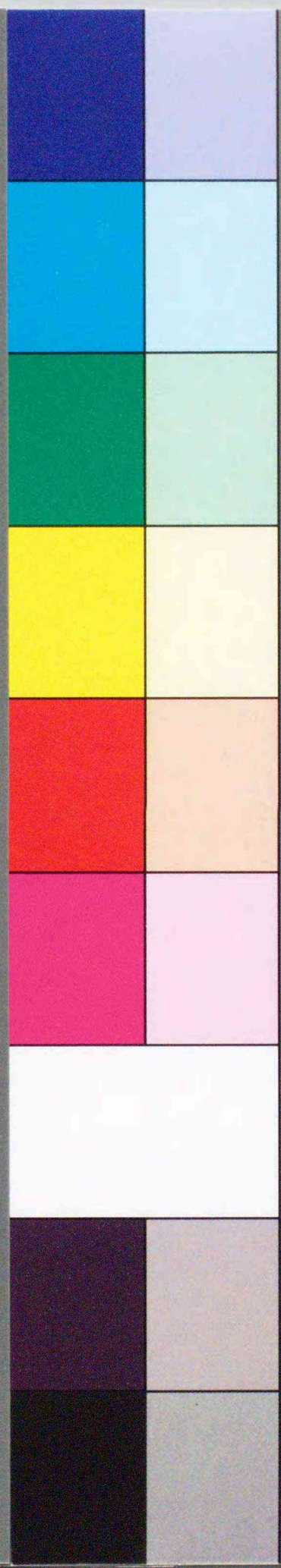


inches 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

© Kodak, 2007 TM: Kodak

Blue Cyan Green Yellow Red Magenta White 3/Color Black



Kodak Gray Scale



© Kodak, 2007 TM: Kodak

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19

