

博士学位申請論文

因果性の学習と推論に関する

実験心理学的研究

—因果ベイズネットの心理学的妥当性の検討—

関西学院大学大学院文学研究科  
博士課程後期課程  
総合心理学専攻心理学領域  
斎藤 元幸

## 要旨

因果の知識をどのように獲得して使用するかという問題は数多くの哲学者の論考を背景に、心理学においても様々なアプローチから解決が試みられてきた。近年、新たなアプローチとして因果ベイズネットを用いた説明がなされている。因果ベイズネットとは、既存の情報から因果関係を適切に導き出し、与えられた情報から推測を行うために統計学や計算機科学といった分野で開発された数理的手法である。因果ベイズネットを用いることで因果性の学習と推論に関する一部の側面はうまく説明されるものの、まだ十分に検討されていない側面も数多く存在している。本研究の目的は、因果構造の学習と因果強度の学習と因果推論の3つの側面から因果ベイズネットの心理学的妥当性を検討することであった。

因果構造の学習では事象間の因果の方向性が未知であり、学習者は様々な手がかりに基づいて因果構造を推測するように求められる。研究1では因果構造の学習における共変動情報の役割を検討した、2 事象で構成される最も単純な因果構造の判断課題を用いて、共変動手がかりと時間順序手がかりのどちらが使用されるか検討した（実験1）。その結果、共変動手がかりが重視されることが示された。また、共変動手がかりと時間順序手がかりが学習過程でどのように影響を及ぼしているか検討したところ、学習初期では時間順序手がかりが優先され、試行の経過に伴って共変動手がかりが重視されることが明らかとなった（実験2）。さらに、共変動情報を体系的に操作して共変動手がかりから因果の方向性がどのように導かれるかについて検討を行った結果、実験課題に応じて、因果の必要性に基づいた解釈を行うか、因果の十分性に基づいた解釈を行うか変化することが確認された（実験3、実験4）。研究1の結果は、時間順序を重視するヒューリスティックモデルよりも、共変動情報を重視する因果ベイズネットを支持していた。研究2では因果構造の学習において共変動手がかりと時間順序手がかりの使用が課題の複雑性によってどのように変化するか検討した。実験の結果、事象の数が少ない

単純な課題では共変動手がかりが、事象の数が多い複雑な課題では時間順序手がかりが重視されることが示された（実験 1, 実験 2）。また、介入による学習ではこのような判断方略の変化は確認されなかった（実験 3）。研究 2 の結果から、観察と介入では異なるプロセスが働いていることが示唆された。

因果強度の学習では、原因と結果が予め決められており、学習者はそれらの因果関係の強さを判断することが求められる。研究 3 では因果強度の学習において共変動情報を操作した 9 つの実験から 114 条件をまとめ、規則ベースアプローチや連合的アプローチや因果的アプローチの代表的なモデルの比較検討を行った。シミュレーションの結果、因果ベイズネットを採用している SS パワーモデルがデータに対する最も高い適合度を示し、その妥当性が示唆された。研究 4 では観察による学習と介入による学習のどちらが因果強度の正確な推定をもたらすか検討した。実験の結果、介入による学習の促進効果が明らかとなった（実験 1, 実験 2）。研究 4 の結果は観察と介入では異なるプロセスが働いていることを示唆するものであった。

因果推論では、獲得した因果の知識に基づいてある事象から別の事象に対して推測を行うことが求められる。研究 5 では因果推論において観察に基づく推論と介入に基づく推論が区別されるか検討したところ、それぞれで因果ベイズネットの予測と一致した推測がなされていることが明らかとなり、その妥当性が示唆された（実験 1, 実験 2）。研究 6 では介入に基づく因果推論を意思決定場面にもまで拡張した選択の因果モデル理論に焦点を当て、選択が介入として機能するか、因果モデルが考慮されるか、因果モデルのパラメータが考慮されるか検討を行った。実験の結果、因果モデルは考慮されるが（実験 1）、基準率や因果強度などのパラメータは考慮されないことが示され（実験 2, 実験 3）、選択の因果モデル理論の妥当性が示唆された。

以上の 15 の実験を含む 6 つの研究を通して、因果性の学習と推論における因果ベイズネットの心理学的妥当性が示された。これらの結果は、従来の連合的アプローチや規則ベースアプローチでは説明することが困難であった。

	頁
要旨	
<b>第1部 序論</b>	
1. はじめに	1
1. 1. 因果性の学習と推論	1
1. 2. 因果ベイズネット	1
1. 3. 博士論文の目的および構成	3
2. 哲学的背景	4
3. 因果性の学習と推論におけるモデル	9
3. 1. 序論	9
3. 2. 規則ベースアプローチ	9
3. 2. 1. $\Delta P$ ルール	10
3. 2. 2. 重み付き $\Delta P$ ルール	11
3. 2. 3. $\Delta D$ ルール	12
3. 2. 4. $pCI$ ルール	12
3. 2. 5. $EI$ ルール	12
3. 2. 6. 二要因ヒューリスティックモデル	13
3. 2. 7. 条件付き $\Delta P$ ルール	13
3. 2. 8. 確率対比モデル	14
3. 3. 連合的アプローチ	14
3. 3. 1. Rescorla-Wagnerモデル	15
3. 3. 2. 改訂版Rescorla-Wagnerモデル	16
3. 3. 3. Pearceモデル	16
3. 4. 規則ベースアプローチと連合的アプローチの問題点	17
4. 因果性の学習と推論における因果ベイズネット	19
4. 1. 序論	19
4. 1. 1. はじめに	19
4. 1. 2. ベイズの定理	19
4. 1. 3. ベイズの定理の応用	20
4. 2. 因果ベイズネット	21
4. 2. 1. 因果ベイズネットとは	21
4. 2. 2. 因果マルコフ条件	23
4. 2. 3. 介入	23
4. 3. 因果ベイズネットの哲学的背景	24
4. 3. 1. 確率	24
4. 3. 2. 因果性	25

4. 4. 因果ベイズネットによる因果構造の学習	27
4. 4. 1. 因果構造の学習とは	27
4. 4. 2. 制約ベース法	28
4. 4. 3. ベイズ法	29
4. 4. 4. 因果構造の学習における観察と介入	30
4. 5. 因果ベイズネットによる因果強度の学習	31
4. 5. 1. 因果強度の学習とは	31
4. 5. 2. パワーPC理論	32
4. 5. 3. 因果サポートモデル	34
4. 5. 4. SSパワーモデル	36
4. 5. 5. 原因事象が複数存在する場合の因果強度の学習	37
4. 6. 因果ベイズネットによる因果推論	38
4. 6. 1. 因果推論とは	38
4. 6. 2. 直接原因モデル ( $C \rightarrow E$ ) における因果推論	38
4. 6. 3. 因果連鎖モデル ( $C \rightarrow M \rightarrow E$ ) における因果推論	39
4. 6. 4. 共通結果モデル ( $C_1 \rightarrow E \leftarrow C_2$ ) における因果推論	40
4. 6. 5. 共通原因モデル ( $E_1 \leftarrow C \rightarrow E_2$ ) における因果推論	40
4. 6. 6. 因果推論における観察と介入	41
5. 博士論文研究の目的	43
5. 1. 序論	43
5. 2. 因果構造の学習	44
5. 3. 因果強度の学習	45
5. 4. 因果推論	46
<b>第2部 因果構造の学習</b>	
6. 研究1 因果構造の学習における共変動情報の役割	48
6. 1. 序論	48
6. 1. 1. はじめに	48
6. 1. 2. 共変動手がかりと時間順序手がかり	48
6. 1. 3. 本研究の目的	50
6. 2. 実験1	51
6. 2. 1. 方法	51
6. 2. 2. 結果および考察	53
6. 3. 実験2	56
6. 3. 1. 方法	57
6. 3. 2. 結果および考察	58
6. 4. 実験3	60
6. 4. 1. 方法	62
6. 4. 2. 結果および考察	63
6. 5. 実験4	65

6.5.1. 方法	65
6.5.2. 結果および考察	65
6.6. 総合論議	67
7. 研究2 因果構造の学習における課題の複雑性と判断方略の変化	69
7.1. 序論	69
7.2. 実験1	70
7.2.1. 方法	70
7.2.2. 結果および考察	74
7.3. 実験2	76
7.3.1. 方法	76
7.3.2. 結果および考察	78
7.4. 実験3	80
7.4.1. 方法	80
7.4.2. 結果および考察	81
7.5. 総合論議	83
<b>第3部 因果強度の学習</b>	
8. 研究3 因果強度の学習に関する数理モデルの比較検討	87
8.1. 序論	87
8.1.1. はじめに	87
8.1.2. 因果強度の学習に関する数理モデル	88
8.1.3. 本研究の目的	97
8.2. 方法	97
8.3. 結果	98
8.4. 考察	101
9. 研究4 因果強度の学習における介入の促進効果	103
9.1. 序論	103
9.1.1. はじめに	103
9.1.2. 観察と介入	103
9.1.3. 本研究の目的	106
9.2. 実験1	107
9.2.1. 方法	108
9.2.2. 結果および考察	112
9.3. 実験2	114
9.3.1. 方法	114
9.3.2. 結果および考察	117
9.4. 総合論議	118

## 第4部 因果推論

10. 研究5 因果推論における観察と介入	124
10. 1. 序論	124
10. 1. 1. はじめに	124
10. 1. 2. 連合的アプローチと因果的アプローチ	124
10. 1. 3. 本研究の目的	126
10. 2. 実験1	127
10. 2. 1. 方法	127
10. 2. 2. 結果および考察	130
10. 3. 実験2	132
10. 3. 1. 方法	132
10. 3. 2. 結果および考察	133
10. 4. 総合論議	135
11. 研究6 意思決定における因果推論—因果モデルとパラメータの役割—	139
11. 1. 序論	139
11. 1. 1. はじめに	139
11. 1. 2. 因果推論と意思決定	140
11. 1. 3. 選択の因果モデル理論	141
11. 1. 4. 質的推定説と量的推定説	148
11. 1. 5. 本研究の目的	148
11. 2. 実験1	149
11. 2. 1. 方法	151
11. 2. 2. 結果および考察	153
11. 3. 実験2	155
11. 3. 1. 方法	155
11. 3. 2. 結果および考察	161
11. 4. 実験3	164
11. 4. 1. 方法	164
11. 4. 2. 結果および考察	167
11. 5. 総合論議	170

## 第5部 総合論議

12. 総合論議	174
12. 1. 博士論文研究のまとめ	174
12. 2. 理論的含意	175
12. 3. 今後の展望	177
12. 4. おわりに	178
References	180

# 第 1 部

## 序論

# 1. はじめに

## 1. 1. 因果性の学習と推論

因果の知識を獲得して使用することは、複雑な環境において適応的な行動を取るために必要不可欠である。事象間の因果関係を知ることによって、過去の説明や未来の予測が可能になる。また、行為とその結果の因果関係を知ることによって適切な選択を行うことができるようになり、望ましい状態に近づけようと現在の環境を制御することが可能になる。では、因果関係はどのようにして学習されるのか？また、因果の知識からどのような推論が行われるのか？数多くの哲学者の論考を背景に、これらの心的過程を解明しようと様々なアプローチから研究が進められてきた (see Allan, 1993; De Houwer & Beckers, 2002a; Holyoak & Cheng, 2011; Penn & Povinelli, 2007; Sawa, 2009; Shanks, 1993, 2007 for reviews). 成人だけでなく、幼児やヒト以外の動物においても因果の表象が形成されることは多くの心理学者によって認められている (Gopnik & Schulz, 2007; McCormack, Hoerl, & Butterfill, 2011; Shanks, Holyoak, & Medin, 1996; Sloman, 2005). 近年では、因果性の学習と推論を因果ベイズネットによって説明する試みが盛んに行われている (Glymour, 2001; Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Griffiths & Tenenbaum, 2005, 2009; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008; Sloman & Lagnado, 2005; Waldmann & Hagmayer, 2005).

## 1. 2. 因果ベイズネット

因果ベイズネットとは、既存の情報から因果関係を適切に導き出し、与えられた情報から推測を行うために統計学や計算機科学といった分野で開発された数理的手法である (Pearl, 2000). 因果ベイズネットの観点から因果性の学習と推論を捉えた場合、「因果構造の学習」と「因果強度の学習」と「因果推論」という 3 つの側面に分けて考え

ることができる (Figure 1-1). 因果構造の学習では, 何が原因で何が結果となっているか考えることが求められる. 例えば, ある話題を Xさんと Yさんから別々に聞いた場合, Xさんが Yさんに話したのか (i.e.,  $X \rightarrow Y$ ), Yさんが Xさんに話したのか (i.e.,  $X \leftarrow Y$ ), あるいは第三者が 2人に先に話したのか (i.e.,  $X \leftarrow Z \rightarrow Y$ ), どのような経路か判断することは難しい. 因果の順序を判断する際には, これらの確率関係や時間関係は有益な手がかりとなる. 因果強度の学習では, 原因が結果に対してどのような影響を及ぼすか考えることが求められる. 例えば, 投薬と健康状態の因果関係を考える場合, 薬によって体調が改善する可能性もあるが, 副作用によって悪化する可能性もある. 薬の効果を適切に判断するには, 薬を飲んだ時だけでなく, 飲まなかった時にどのような結果が生じたかも考える必要がある. 因果推論では, ある事象の状態から他の事象の状

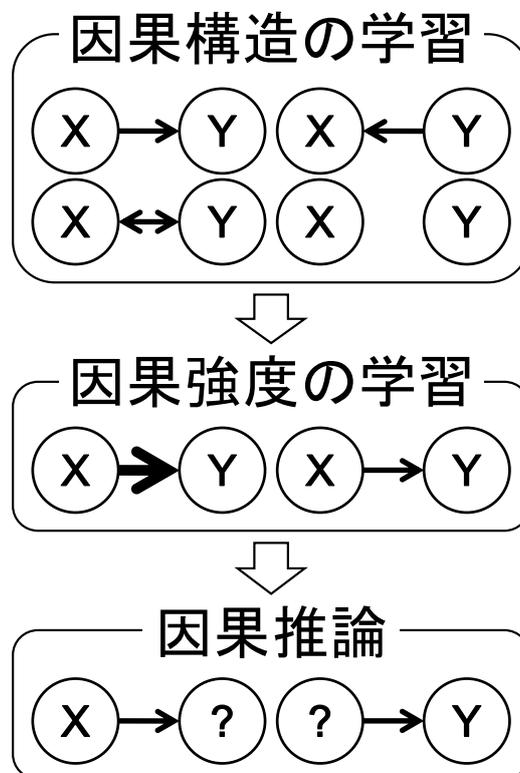


Figure 1-1. 因果性の学習と推論に関する3つの側面.

態について考えることが求められる<sup>1</sup>。例えば、空が曇ってくれば雨が降るのではないかと原因から結果を予測したり、外に出た際に地面が濡れていれば雨が降ったのではないかと結果から原因を推測したりすることができる。このような因果推論が可能になるには、事象間にどのような因果関係があるか事前に知っておく必要がある。因果ベイズネットを用いることでこれらの側面の一部はうまく説明されるものの、まだ十分に検討されていない側面も多数存在している。

### 1. 3. 博士論文の目的および構成

本博士論文の目的は因果ベイズネットの心理学的妥当性を検討することである。第 1 部では、因果性に関する哲学的背景を述べた後、因果性の学習と推論における規則ベースアプローチと連合的アプローチの数理モデルについて説明する。その後、因果的アプローチの核となる因果ベイズネットについて詳しく解説し、博士論文研究の目的について述べる。第 2 部では「因果構造の学習」について行った研究、第 3 部では「因果強度の学習」について行った研究、第 4 部では「因果推論」について行った研究をそれぞれ紹介する。第 5 部ではこれらの研究を踏まえ、因果ベイズネットの心理学的妥当性や今後の展望について議論を行った。

---

<sup>1</sup> ここでは獲得した因果の知識の利用という意味で因果推論と記述しているが、因果構造の学習と因果強度の学習を因果帰納と捉えて、3つの側面すべてを対象として因果推論と呼ぶこともある。

## 2. 哲学的背景

因果関係をどのように認識するかという問題は、心理学が実験的検討を始めるはるか以前から哲学のなかで重要な問題の 1 つとして扱われてきた。ここでは、因果性に関する哲学的議論を歴史の流れに沿って整理し、因果性がどのように定義されるかについて解説を行う (see Psillos, 2005; White, 1990 for reviews)。心理学では因果の知識がどのように獲得および使用されるかという認識論的側面に焦点が当てられているのに対して、哲学では認識論的側面だけでなく、原因はどのようにして結果を引き起こすかといった形而上学的側面についても検討がなされてきた。心理学と哲学では扱う問題が若干異なるものの、哲学的論考は因果性の学習と推論に関する心理学モデルの多くに影響を与えており、ここでは心理学モデルとの関連についても述べる。

Aristotle (trans. 1984) は原因をその意味から 4 つに分類している。それらは、そこから事物が生じる質料的なものである質料因 (material cause) と事物が完成する時、その事物が最後まで持っている形相である形相因 (formal cause) と事物を完成させる作用因である動力因 (efficient cause) と事物の目的である目的因 (final cause) である。例えば、彫像の質料因は材料であり、その形相因は形や輪郭であり、その動力因は作者であり、その目的因は彫像が作られた目的である。完全な因果的説明には、これら 4 つの原因を挙げる必要がある。Aristotle は因果性について 2 つの考えを述べている。1 点目は、あらゆる変化には原因が必要であり、もし何の干渉もなければ、そのような因果現象は規則的に起きるということである。ここで動力因の連鎖を遡っていくと、それ自身は決して動かされることはないが他のものを動かす「不動の動者 (unmoved mover)」に辿りつく。スコラ哲学者はこのような究極原因は神であると考えた。この必要性は、論証的議論における論理的必然性と同等であるというのが 2 点目の主張である。Aristotle によると、前提から結論を導く論理的必然性というのは、原因が結果を発生させる物理的必然性を反映したものである。

Descartes (1644/1991) は、世界が物理的実体と心的実体から成り立っていると考え、物心二元論を提唱した。物心二元論によると、物体は力を持たないものと考えられ、物体を持った時に感じる重さは心的実体である我々の知覚に基づくものであり、物理的実体が有する性質とは捉えられない。また、因果性に関する Aristotle との相違点として、運動や静止が続いた状態にあるのを説明するのに原因を仮定する必要はないという慣性 (inertia) を導入したことが挙げられる。物心二元論においては、物体のある性質が他の物体に伝達されるという因果性の伝達モデル (transference model of causality) を用いて物体同士の相互作用を説明する。しかしながら、どのようにして物体と心の相互作用が可能になるのかという問いに対してこのモデルは答えることができなかった。Descartes の後継者は機械原因論者のグループと物体に力能を再導入するグループに分けられる。前者の代表者である Malebranche (1674-1675/1997) によると、真の原因とは、原因と結果の間に必然的結合を知覚するような原因であり、精神は無限に完全な存在 (神) とその結果の間には必然的結合を知覚する。したがって、真の原因は神のみであることが導かれる。また、必然的結合のような関係はどこにも見当たらないため、因果性は存在しないかもしれないと考えられていた。一方、後者の代表者である Leibniz (1765/1896) は物質界に力や効力を再導入し、物質はモナド (monad) によって維持され、モナドによって後の状態が決定すると考えた。因果的結合の代わりに、モナド間における神の予定調和が想定されており、モナドは相互作用しないという主張は因果性の排除と解釈されている。

Hume (1748/2003) は近世哲学において因果性の問題に対して最も重要な寄与をなした。Hume は著書『人間本性論』において、原因と結果の間にある必然的結合のようなものは存在せず、そこにあるのは蓋然性だけであると主張している。例えば、ボール *A* がボール *B* にぶつかり、ボール *B* が動き出すとき、一般的には *A* によって *B* が動いたと捉えられる。しかし、これら 2 つの出来事をどれだけ細かく突き詰めたとしても、この原因「によって」が見えることはない。それまでの哲学者が物事の生起を神の力に依拠して

いたのに対して、Hume は恒常的連接 (constant conjunction), 時間的継起 (temporal succession), 時空間的接近 (spatiotemporal contiguity) によって因果性の観念が形成されるとした。因果推論は、これまで経験した恒常的連接がこれからも続くという「自然の斉一性 (uniformity of nature)」を前提としている。しかし、自然の斉一性は理性的に基礎づけることができないため、因果推論を支えているのは「理性」ではなく、「習慣」であると考えられた。このような Hume の思想は、すべての因果的知識は経験から生じるという連合的アプローチに強く影響を及ぼしている。

Hume に対する批判を行ったのは Kant (1781/2003) であった。Kant は著書『純粋理性批判』において、ある出来事と他の出来事との因果関係を認識するためには、どちらが先行しているかが規定されなければならないと述べている。この規定は時間の前後における構想力の連結だけでは不十分であるため、「原因と結果の概念」が必要となると主張している。因果性を心の働きによって構成されるものと捉える点で Hume と Kant の見解は一致していたが、Hume がア・ポステリオリな判断と考えているのに対して、Kant はア・プリオリな判断を想定していた。このような Kant の思想は、経験だけでなく先験的知識も重要視する因果的アプローチや、事象間の共変関係よりもメカニズムを重視する力学モデル (Wolff, 2007) などに反映されている。

Mill (1843) は、原因は結果に対する必要十分条件だと考え、Hume の思想に洗練を加えた。すべての知識が経験に由来するという考えは Hume と同様であったが、恒常的連接が因果的連接となるのは、それが他の要因の有無に依存しない無条件のときだけであると考えた。例えば、2 つの事象に恒常的連接が生じていたとしても、それが共通の原因によってもたらされている場合は原因事象なしに連接は生じないため、因果的連接と呼ぶことはできない。また、Mill は何が原因であるかということよりも、我々がどのように原因を発見するかということに関心を持っており、因果関係を解明するいくつかの方法を提唱している。因果の必要性によると結果の生起には原因の生起が必要であり、同一の結果が生じるという以外は互いに似ていない事例が多く存在する場合はそれら

の事例に共通する要因が原因である。因果の十分性によると原因の生起は結果の生起に十分であり、ある要因が存在しない場合に限定して結果が生じなければその要因が原因である。

一方、Mackie (1974) は結果の生起には複数の原因が必要であると考えた。Mackie によると、原因は結果の生起に「不十分だが余分ではなく (Insufficient but Non-redundant)」, 結果の発生に「必要ではないが十分 (Unnecessary but Sufficient)」な条件の一部である。例えば、火災の原因は漏電だったという場合、漏電はそれだけでは火災を生じさせるのには不十分だが、火災を生じさせるのに必要ではないが十分な条件セット(e.g., 可燃材料や酸素の存在や漏電)の一部である。このことは、それぞれの頭文字から INUS 条件と呼ばれている。

Mill や Mackie が必要性や十分性の論理を用いて恒常的接続に修正を加えていたのに対して、Suppes (1970) は確率を用いて恒常的接続の修正を試みた。Suppes は恒常的接続の問題点として、不完全な接続を扱うことができないことを指摘している。例えば、喫煙が肺がんを引き起こすといっても全ての喫煙者が肺がんを患うわけではない。Hume の定義では両者の間に因果関係は存在しないとなるが、原因  $C$  が結果  $E$  の生起確率を高める (i.e.,  $P(e|c) > P(e|\neg c)$ ) と考えることにより、上記の主張は可能になる。しかしながら、この定義には擬似相関と因果関係を区別できないという問題点があった。例えば、気圧計の変化と降雨の有無には相関関係が認められるが、気圧計が天候の変化を引き起こしているわけではなく、気圧という共通原因によって両者が変化しているだけである。そこで、Eells (1991) は状況要因  $F$  を統制した上で原因  $C$  が結果  $E$  の生起確率を上昇させること (i.e.,  $P(e|c, f) > P(e|\neg c, f)$ ) を因果性の定義に採用し、この問題点を解決した。状況要因の統制によってこの問題は解決したように見えたが、どの要因を統制すればよいかという新たな問題に直面することとなった。因果性を 2 つの確率の差分とみなす考え方は  $\Delta P$  ルール (Jenkins & Ward, 1965) や確率対比モデル (Cheng & Novick, 1992) といった心理学モデルにも現れている。

Mill や Mackie や Suppes が因果性を規則的な継起と捉えていたのに対して、Lewis (1973) は、「もし原因が生起しなかったならば結果も生起しなかった」という反実仮想に基づいて因果性を定義している。例えば、漏電が火災の原因となるのは、もし漏電が生じていなかったならば火災は生じなかったという場合に限定される。Lewis は「もし事象  $C$  が生じていれば事象  $E$  も生じていた」・「もし事象  $C$  が生じていなければ事象  $E$  も生じていなかった」という反事実的条件文の真偽を判定するために、現実世界に類似した世界が別に存在するという可能世界論を利用した。しかしながら、可能世界論の適用には多くの困難が伴い、否定的な見解が多い (cf. 佐金, 2008)。

Woodward (2003) は介入 (intervention) の概念を導入して因果性の定義を行っている。介入とは事象の操作であり、事象  $C$  への介入によって事象  $E$  の状態が変化する場合  $C$  は  $E$  の原因である。例えば、気圧計の変化を観察して天候を予測することは可能であるが、両者の関係は気圧という共通原因による擬似相関であって因果関係ではないため、気圧計の針を操作して天候を変化させることはできない。Pearl (2000) はグラフ理論と確率論を基盤として観察と介入の違いを定量的に扱う因果ベイズネット (causal Bayes nets) と呼ばれる数理的手法を開発しており、既存の情報から因果関係を適切に導き出し、与えられた情報から推測を行うことを可能にした。近年では、因果ベイズネットを認知モデルとして採用し、どのように因果的知識が獲得・使用されるのかを説明する試みが盛んに行われている (Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Griffiths & Tenenbaum, 2005, 2009; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008)。

### 3. 因果性の学習と推論に関するモデル

#### 3. 1. 序論

数多くの哲学者の論考を背景に、因果性の学習と推論のプロセスを説明しようと様々な理論やモデルがこれまでに考案されてきた (see Pineño & Miller, 2007 for a review). それらはアプローチの違いから大きく3つに分類することができる. 1つ目は規則ベースアプローチ (rule-based approach) であり, この立場では事象間の関係性を表す規範的な統計量あるいはヒューリスティックスといった簡便な方略を用いた説明が提供されている. 2つ目の連合的アプローチ (associative approach) では, 条件づけの理論を援用して因果学習の過程を説明する試みがなされている. 3つ目は因果的アプローチ (causal approach) と呼ばれ, 因果ベイズネットを用いて因果性の学習と推論の記述を行う. 因果的アプローチの説明は次章で詳述するため, 本章では規則ベースアプローチと連合的アプローチのモデルをそれぞれ概説し, それらの問題点について述べる. なお, 関連する他のアプローチとして, De Houwer (2009) による命題的アプローチ (propositional approach) や Wolff (2007) による力学モデル (dynamics model) なども存在する. 前者は数量的予測を行わないため, 後者は条件文推論課題を主な対象としており博士論文研究の実験課題とは直接関連しないため, 解説は省略した.

#### 3. 2. 規則ベースアプローチ

規則ベースアプローチでは, ヒトが直観的統計学者のように振る舞い, 事象の生起情報から因果関係が引き出されることが想定されている. 事象の生起情報は Figure 3-1 に示した  $2 \times 2$  の分割表で表現される. 4つのセルは各々の事象の共生起の頻度を表している. 例えば, セル  $a$  は原因と結果が共生起した事例を, セル  $d$  は原因と結果が生起しなかった事例を表している. この分割表で表現される事象の生起情報に対して何らか

		Event <i>E</i>	
		e	¬e
Event <i>C</i>	c	a	b
	¬c	c	d

Figure 3-1. 原因事象(*C*)と結果事象(*E*)の分割表. アルファベット a, b, c, d はそれぞれの共生起の頻度を表す.

の計算が行われる. このアプローチの中には, 事象間の関係性を表す規範的な統計量に基づく判断を仮定する規範的モデルだけでなく, ヒューリスティックスといった簡便な方略に基づく判断を想定する記述的モデルも含まれる (cf. Hattori & Oaksford, 2007).

### 3. 2. 1. $\Delta P$ ルール

規則ベースアプローチにおける最も代表的なモデルは $\Delta P$ ルール (Jenkins & Ward, 1965; Ward & Jenkins, 1965) であり, 以下のように定義される.

$$\Delta P = P(e|c) - P(e|\neg c) = \frac{a}{a+b} - \frac{c}{c+d} \quad (3-1)$$

ここで  $P(e|c)$  は原因が生起した時に結果が生起する条件付き確率を,  $P(e|\neg c)$  は原因が生起しなかった時に結果が生起する条件付き確率を表す.  $\Delta P$  はこれら 2 つの条件付き確率の差分をとることで算出され, -1 から +1 までの値になる. 原因の生起によって結果

の生起確率が上昇する場合 (i.e.,  $\Delta P > 0$ ), 原因と結果の間には発生的因果関係が想定される. 一方, 原因の生起によって結果の生起確率が低下する場合 (i.e.,  $\Delta P < 0$ ), 原因と結果の間には抑制的因果関係が想定される. 原因の有無によって結果の生起確率が変化しない場合 (i.e.,  $\Delta P = 0$ ), 原因と結果は無関係となる.

$\Delta P$  ルールについては支持する結果と支持しない結果の両方が報告されている. いくつかの研究は因果強度の評定値と  $\Delta P$  におおまかな対応関係がみられることを示している (e.g., Allan & Jenkins, 1980; Alloy & Abramson, 1979; Dickinson, Shanks, & Evenden, 1984; Wasserman, Elek, Chatlosh, & Baker, 1993; Ward & Jenkins, 1965). しかしながら, 原因事象の生起確率が上昇すると評定値も上昇する密度バイアスなど  $\Delta P$  から逸脱した結果も報告されている (e.g., Shanks, 1985a).

### 3. 2. 2. 重み付き $\Delta P$ ルール

$\Delta P$  ルールでは各セルの情報は等しく重み付けられているが, 実際にはそうならないことが示されている. 例えば Kao and Wasserman (1993) は, 4 種類のセル頻度は  $a > b > c > d$  の順で優先されることを示している. また, 抑制的因果関係を判断する場合は  $b > a > d > c$  の順に重視されることが報告されている (Mandel & Vartanian, 2009). これらの知見を反映したより記述的なモデルとして, それぞれの生起頻度に重みを付けた重み付き  $\Delta P$  ルールが提唱されている. 重み付けの方法はいくつか考案されているが (e.g., Einhorn & Hogarth, 1986; Schustack & Sternberg, 1981), Anderson and Sheu (1995) による重み付き  $\Delta P$  ルールは以下の通りである.

$$\Delta P_w = \beta_0 + \beta_1 P(e|c) - \beta_2 P(e|\neg c) \quad (3-2)$$

ここで  $\beta_1$  と  $\beta_2$  は条件付き確率の重みを表すパラメータであるが,  $\beta_0$  を切片,  $\beta_1$  と  $\beta_2$  を回帰係数とする回帰モデルとしても捉えられる.

### 3. 2. 3. $\Delta D$ ルール

いくつかの研究では  $\Delta P$  ルールとは異なる判断方略が使用されることが報告されている (e.g., Shimazaki, Tsuda, & Imada, 1991). その中でも頻繁に見られるのは  $\Delta D$  ルール (Inhelder & Piaget, 1958) と呼ばれるものであり, これは確信事例 (i.e., セル  $a$  とセル  $d$ ) の数から非確信事例 (i.e., セル  $b$  とセル  $c$ ) の数を引くことで計算される.

$$\Delta D = (a + d) - (b + c) \quad (3-3)$$

### 3. 2. 4. $pCI$ ルール

$\Delta D$  ルールは頻度に基づいた予測を行うため, 確信事例が増え続けた場合に上昇し続けてしまうという欠点を持つ. White (2003) は  $\Delta D$  ルールを比率に基づく形に変更し,  $pCI$  ルールと命名した.  $pCI$  ルールは  $\Delta P$  ルールと同様に  $-1$  から  $+1$  までの値をとる.

$$pCI = \frac{a + d - b - c}{a + b + c + d} \quad (3-4)$$

### 3. 2. 5. $EI$ ルール

Perales and Shanks (2007) は  $\Delta D$  ルールを比率に基づく形に変更し, それぞれの生起頻度に重みを付け,  $EI$  ルールと命名している.  $w$  はセルの重みを表すパラメータであり, 全ての情報の重みが等しい場合 (i.e.,  $w_a = w_b = w_c = w_d = 1$ ),  $pCI$  ルールと一致する.

$$EI = \frac{w_a a + w_d d}{w_a a + w_b b + w_c c + w_d d} - \frac{w_b b + w_c c}{w_a a + w_b b + w_c c + w_d d} \quad (3-5)$$

Perales and Shanks (2007) では共変動情報を操作した実験の中から 114 条件を選出してメタ分析を行ったところ,  $EI$  ルールが最も高い適合度を示すことが明らかになった.

### 3. 2. 6. 二要因ヒューリスティックモデル

Hattori and Oaksford (2007) は連関係数の指標である  $\varphi$  係数に変更を加えて独自のモデルを提唱している. 原因と結果が非生起の情報 (i.e., セル  $d$ ) は何も起きていないことを表しており, 現実的にはこの生起頻度を定義することは困難であり, また便宜的に定義しても無数に存在しているので数え上げることは難しい. そこで,  $\varphi$  係数に含まれる  $d$  を無限大に発散させた二要因ヒューリスティックモデルが提唱されている.

$$H_G = \lim_{d \rightarrow \infty} \varphi = \sqrt{P(e|c)P(c|e)} = \frac{a}{\sqrt{(a+b)(a+c)}} \quad (3-6)$$

ここで言う二要因とは, 原因から結果への予測可能性  $P(e|c)$  と原因の結果に対する適合性  $P(c|e)$  という 2 つの要因のことである. Hattori and Oaksford (2007) は共変動情報を体系的に操作した実験を対象としてメタ分析を行い, 41 個のモデルを比較した中で二要因ヒューリスティックモデルが最も高い適合度を示したことを報告している.

### 3. 2. 7. 条件付き $\Delta P$ ルール

$\Delta P$  ルールでは, 原因事象が存在する時と存在しない時の比較に基づいて因果関係が判断される. しかしながら, 複数の原因事象が存在する場合には, 他の要因 ( $f$ ) の影

響を統制したうえで比較を行う必要がある。条件付き  $\Delta P$  ルールは以下のように算出される (e.g, Spellman, 1996; Waldmann & Holyoak, 1992)。

$$\Delta P_{cond} = P(e|c, f) - P(e|\neg c, f) \quad (3-7)$$

他の要因が存在しない事態において条件付き  $\Delta P$  を算出した場合は、交絡のない因果強度の推定値となる。

### 3. 2. 8. 確率対比モデル

Cheng and Novick (1992) によって提唱された確率対比モデルでは、与えられた情報全体で条件付き  $\Delta P$  が計算されるのではなく、その部分集合である焦点集合 (focal set) に対して条件付き  $\Delta P$  が計算されることを想定している。焦点集合を用いることによって、原因と可能化条件 (enabling condition) の区別が可能になる。なお、確率対比モデルでは原因事象の相互作用を交互作用対比として条件付き  $\Delta P$  の差分で計算しているが、交互作用の計算としては不適切なことが後に指摘されている (Novick & Cheng, 2004)。

### 3. 3. 連合的アプローチ<sup>2</sup>

因果学習の実験事態で原因事象と結果事象を観察させることは、古典的条件づけの条件刺激 (conditioned stimulus: CS) と無条件刺激 (unconditioned stimulus: US) の対呈示と構造的に類似していることから、条件づけの理論を用いて因果性の学習過程を説明する試みがなされてきた。連合的アプローチでは、因果関係の学習を事象間の連合強度の変化として捉える。

---

<sup>2</sup> ここで紹介しきれなかったモデルとして、修正 SOP モデル (Dickinson & Burke, 1996) やコンパレータ仮説 (Denniston, Savastano, & Miller, 2001; Miller & Matzel, 1988; Stout & Miller, 2007) が挙げられる。

連合的アプローチの強みは、手がかり競合 (cue competition) と呼ばれる、原因事象が複数存在する事態における様々な現象を説明できることである。動物の条件づけ研究と同様に、因果学習においても隠蔽 (overshadowing) やブロッキング (blocking) といった現象が確認されている (e.g., Baker, Mercier, Vallée-Tourangeau, Frank, & Pan, 1993; Chapman & Robbins, 1990; Price & Yates, 1993; Shanks, 1985b)。これらの現象は連合的アプローチのモデルでは容易に説明可能であるが、規則ベースアプローチの多くのモデルでは説明に困難を伴う。

### 3. 3. 1. Rescorla-Wagner モデル

連合的アプローチにおける最も代表的なモデルは Rescorla-Wagner モデル (Rescorla & Wagner, 1972) であり、連合強度の変化は次のように計算される。

$$\Delta V_i = \alpha\beta \left( \lambda - \sum V_T \right) \quad (3-8)$$

ここで  $\Delta V_i$  は原因と結果の連合強度の変化量を、 $\sum V_T$  は既に形成されている連合強度を表す。結果事象が生起する場合には  $\lambda = 1$  となり、連合強度は増加する。一方、結果事象が生起しない場合には  $\lambda = 0$  となり、連合強度は減少する。学習の初期などで  $\sum V_T$  の値が小さいほど、連合強度の変化量は大きくなる。 $\alpha$  は原因事象の明瞭度、 $\beta$  は結果事象の明瞭度を表すパラメータであり、0 から 1 の値をとる。これら 2 つのパラメータも連合強度の変化量に影響を及ぼす。

一般的に結果事象の明瞭度  $\beta$  は、結果が生起する時と生起しない時で異なる値になると想定されているが、両者を一定の値に設定した場合は Rescorla-Wagner モデルの漸近値と(3-1)式に示した  $\Delta P$  の値が一致することが知られている (Chapman & Robbins, 1990; Wasserman et al., 1993; see Danks, 2003 for detailed

analysis). Rescorla-Wagner モデルの漸近値はセルの生起頻度を用いて以下のように表現することも可能である (Wasserman et al., 1993).

$$V_{asympt} = \frac{\beta_e a}{\beta_e a + \beta_{-e} b} - \frac{\beta_e c}{\beta_e c + \beta_{-e} d} \quad (3-9)$$

### 3. 3. 2. 改訂版 Rescorla-Wagner モデル

Rescorla-Wagner モデルでは, 原因事象が存在しない時に連合強度が変化しないことを想定しているため, 逆行ブロッキング (backward blocking) などの回顧的再評価 (retrospective revaluation) を説明することはできない. この問題に対しては, 原因事象の明瞭度  $\alpha$  を負の値に設定することで原因事象が存在しない時でも連合強度が変化するように修正が提案されている (Van-Hamme & Wasserman, 1994). しかし, De Houwer and Beckers (2002b) によって明らかにされた高次の回顧的再評価は改訂版 Rescorla-Wagner モデルでも説明困難である.

### 3. 3. 3. Pearce モデル

Rescorla-Wagner モデルと同様の学習規則に基づくものの, Pearce モデル (Pearce, 1987) では異なる予測値が導かれる. Pearce モデルの特徴は, 複数の手がかり刺激を 1 つの複合刺激として捉えることであり, 2 つの刺激の間には類似性に基づいた般化が生じることが想定されている. 原因事象と結果事象が 1 つしか存在しない事象では, 文脈刺激 ( $X$ ) と原因事象と文脈刺激の複合刺激 ( $CX$ ) という 2 つの刺激に対して学習が生じる. 一方の刺激が提示された場合には, 他方の刺激表象が類似性に応じて活性化されるため, 連合強度の変化量は以下のように計算される.

$$\Delta V_{CX} = \beta(\lambda - V_{CX} - x_1 V_X) \quad (3-10)$$

$$\Delta V_X = \beta(\lambda - V_X - x_2 V_{CX}) \quad (3-11)$$

原因事象が生起している場合は(3-10)式が、生起していない場合は(3-11)式が適用される。この式に含まれる多くのパラメータは Rescorla-Wagner モデルと同じ意味を持つ。新たなパラメータ  $x$  は刺激間の類似性によって規定され、般化の程度を表す。原因事象に対する判断は原因事象( $C$ )と原因事象と文脈刺激の複合刺激( $CX$ )の類似性によって決まる。

$$J_C = x_3 V_{CX} \quad (3-12)$$

それぞれのセルの生起頻度を用いて表現した場合は以下のようになる (Perales & Shanks, 2003)。

$$J_C = \frac{x_3[a(c+d) - c(ax_1 + bx_1)]}{a(c+d) + b(c+d) - cx_2(ax_1 + bx_1) - dx_2(ax_1 + bx_1)} \quad (3-13)$$

### 3. 4. 規則ベースアプローチと連合的アプローチの問題点

規則ベースアプローチは因果性を確率に還元しており、連合的アプローチは因果性を連合強度に還元しているため、これらのアプローチは因果関係を記述する術を持たない。そのため、2つの事象の共変動が直接の因果関係によるものか、共通の原因による擬似相関なのか区別して扱うことができない。因果関係と擬似相関の差異は事象に介入を行う際に顕著に表れる。因果関係がある場合は原因事象への介入が結果事象を発生させるのに対して、擬似相関しかない場合は介入を行ったとしても何も変化しない。例えば、気圧計の値と天候の状態は共変関係にあるが、両者の関係は共通原因 (i.e., 低気圧) による擬似相関のため、気圧計の針を操作しても天候が変わることはない。実

験参加者が因果構造と介入の論理を理解していることは、多くの研究によって示されてきた (e.g., Meder, Hagmayer, & Waldmann, 2008, 2009; 斎藤・嶋崎, 2012; Sloman & Lagnado, 2005; Waldmann & Hagmayer, 2005).

規則ベースアプローチや連合的アプローチでは、原因は結果を発生させるが、結果が原因を発生させることはないという因果の非対称性が無視されている。原因事象から結果事象を推測する予測的推論では、原因事象の因果強度と他の原因の影響を考慮する必要がある。一方、結果事象から原因事象を推測する診断的推論では、複数の原因事象の因果強度だけでなく、それらがどのくらいの確率で生起するかという基準率も考慮する必要がある。それぞれの推論では異なる判断が見られることが報告されており (e.g., Fernbach, Darlow, & Sloman, 2011), 規則ベースアプローチや連合的アプローチではこの差異を説明することはできない。また、因果の非対称性を考慮しないことは、2つの原因事象と1つの結果事象が存在する事態 (i.e., 共通結果モデル) と、1つの原因事象と2つの結果事象が存在する事態 (i.e., 共通原因モデル) について同じ予測を行うことを意味する。しかしながら、前者の事態でのみ個々の因果強度に対する判断が互いに影響しあうことが多くの研究で報告されている (e.g., Booth & Buehner, 2007; Waldmann, 2000, 2001; Waldmann & Holyoak, 1992; Waldmann & Walker, 2005)<sup>3</sup>。原因事象同士は競合するが、結果事象同士は競合しないことは、手がかり競合の非対称性と言われている。

因果関係と相関関係の区別や因果の非対称性といった因果性に起因する問題は、因果性を確率に還元する規則ベースアプローチや連合強度に還元する連合的アプローチでは対処することが困難である。しかし、因果ベイズネットを用いる因果的アプローチではこれらの現象をうまく説明することが可能である。

---

<sup>3</sup> 手がかり競合の非対称性に関する知見には否定的な見解も存在するが (e.g., Arcediano, Matute, Escobar, & Miller, 2005), 細かな実験手続きの影響を受けやすいだけで現象の存在は確かであることが示されている (Waldmann & Walker, 2005).

## 4. 因果性の学習と推論における因果ベイズネット

### 4. 1. 序論

#### 4. 1. 1. はじめに

連合的アプローチ (associative approach) では因果性を連合に還元し、ヒトの因果学習や因果推論を動物の連合学習と同じ枠組みで捉えることで節約的かつ包括的な説明が試みられている。一方、規則ベースアプローチ (rule-based approach) では因果性を確率に還元し、事象間の関係性を表す規範的な統計量あるいはヒューリスティックスといった簡便な方略を用いた説明が提供されている。しかしながら、これらのアプローチに対しては因果関係と相関関係が区別されていない、原因と結果の非対称性が考慮されていないといった問題点が指摘されている。

近年、これらの問題点を克服する新たなアプローチとして因果ベイズネット (causal Bayes nets) を用いるベイズ的アプローチによる説明が注目されている。因果ベイズネットとは、既存の情報から因果関係を適切に導き出し、与えられた情報から推測を行うために哲学や統計学や計算機科学といった分野で開発された統計的因果推論の手法である (Pearl, 2000)。ベイズ的アプローチの根幹には、Thomas Bayes が発見したベイズの定理 (Bayes, 1763) に基づく計算がある。

#### 4. 1. 2. ベイズの定理

認知主体がデータ  $d$  を生成する仮説  $h$  について考える事態において、データを得る前の仮説についての確信の度合いを  $P(h)$  という主観確率で表す。認知主体はデータを得た後に仮説に対する確信の度合いをどのように変化させるべきか？ベイズの定理はこのような確率の更新の度合いを定めるものであり、以下の式で表現される。

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)} \quad (4-1)$$

ここで  $P(d|h)$  は仮説  $h$  のもとでデータ  $d$  が発生する確率を表すが、データが発生した後には確率でなくなるため、仮説  $h$  についての尤度と呼ばれる。また、分母の  $P(d)$  は考えられるすべての仮説のもとでデータが得られる確率を表す。この数値は分子の計算可能なすべての組合せについて、合計が 1 になるように基準化する項であると考えられることから、基準化定数や正規化定数と呼ばれる。

ベイズの定理はデータを得る前の仮説の確からしさである事前確率  $P(h)$  から、データを得た後の仮説の確からしさである事後確率  $P(h|d)$  への確信の度合いの変化を記述する。結果であるデータに基づき、それを生成する原因である仮説を推測していることから、逆確率の定理と言われることもある。

#### 4. 1. 3. ベイズの定理の応用

ベイズの定理は、事前の知識を新たに獲得したデータに基づいて更新するという学習過程を表現したものと解釈可能であり、その汎用性の高さから様々な場面に応用されている。母数の推定や仮説の比較にベイズの定理を用いることはベイズ統計 (Bayesian statistics) と呼ばれ、従来の伝統的な統計学では扱うことができなかった問題に解決手段を提供している (cf. 繁樹, 1985; 渡部, 1999)。行動に対して損失関数を設定し、不確実状況下における最適な意思決定をベイズの定理に基づいて考えるのがベイズ決定理論 (Bayesian decision theory) である (cf. 松原, 2008)。

ベイジアンネットワーク (Bayesian networks) とは、変数間の関係性をグラフと確率で表現し、グラフ構造の学習や確率推論を行う数理的な手法であり、機械学習などの領域で盛んに研究されている (Pearl, 1985; cf. 植野, 2013)。因果関係に基づいて

ベイジアンネットワークを構築した場合は、因果ベイズネット<sup>4</sup> (causal Bayes nets) と  
言われる (Glymour, 2001). ベイズの定理や因果ベイズネットを駆使して認知モデル  
を作成することはベイジアンモデリング (Bayesian modeling) と呼ばれ、条件づけ  
(Courville, Daw, & Touretzky, 2006), 視知覚 (Kersten, Mamassian, & Yuille,  
2004; Knill & Richards, 1996), 感覚運動学習 (Körding & Wolpert, 2004), 語  
彙学習 (Xu & Tenenbaum, 2007), 言語処理 (Chater & Manning, 2006), 意味  
表象 (Griffiths, Steyvers, & Tenenbaum, 2007), 帰納推論 (Kemp &  
Tenenbaum, 2009), 因果推論 (Gopnik, Glymour, Sobel, Schulz, Kushnir, &  
Danks, 2004; Griffiths & Tenenbaum, 2005, 2009) など様々な領域に適用され  
ている<sup>5</sup> (see Tenenbaum, Kemp, Griffiths, & Goodman, 2011 for a review).

## 4. 2. 因果ベイズネット

### 4. 2. 1. 因果ベイズネットとは

因果ベイズネットは事象間の因果関係を非循環有向グラフで、その強さを条件付き  
確率で表現する。グラフ内のノードは変数を、エッジ (矢印) は変数間の因果関係を表  
しており、非循環とはノードからエッジを辿っても元のノードに戻らないことを意味している。  
変数は二値変数のような離散変数の場合もあれば、連続変数の場合もあり、また、顕在  
変数や潜在変数のどちらでも組み込むことができる。因果関係は、決定的あるいは確率  
的、発生的あるいは抑制的、線形もしくは非線形といったように様々な形態をとる。

因果ベイズネットのグラフ構造は変数の同時確率分布を定義するために使用される。  
同時確率分布に基づき、グラフ構造が明らかな場合は変数の予測を行い、グラフ構造  
が明らかでない場合はその学習を行う。例えば、喫煙 ( $S$ ) が歯の黄ばみ ( $Y$ ) と肺がん ( $L$ ) を

---

<sup>4</sup> 因果ベイジアンネットワークと言われることもある。

<sup>5</sup> ベイジアンモデリングについては、*Trends in Cognitive Science* 誌 (Chater, Tenenbaum, & Yuille, 2006; Kousta, 2010) や *Developmental Science* 誌 (Gopnik & Tenenbaum, 2007) や *Cognition* 誌 (Xu & Griffiths, 2011) の特集号なども参照されたい。

引き起こしている関係は Figure 4-1(左図)のように表現される。これら 3 つの変数の同時確率分布は連鎖公式<sup>6</sup>を用いて以下のように表現される。

$$P(S, Y, L) = P(S)P(Y|S)P(L|Y, S) \quad (4-2)$$

また、連鎖公式においてこれらの変数は任意に並び替えることが可能であり、以下のように記述することも可能である。

$$P(Y, L, S) = P(Y)P(L|Y)P(S|L, Y) \quad (4-3)$$

しかしながら、これら(4-2)式や(4-3)式は実際のグラフ構造を反映した形にはなっておらず、変数の同時確率分布を一意に表現するために、因果ベイズネットの核となる仮定が設けられている。

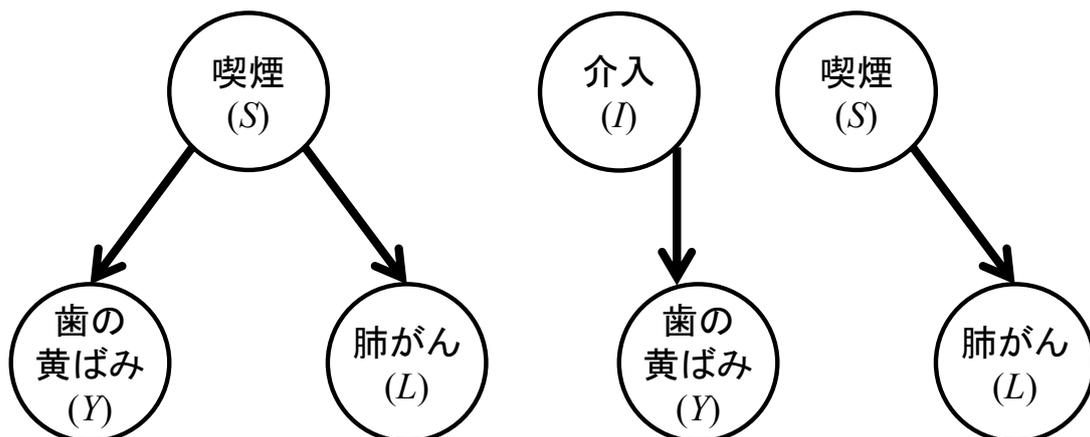


Figure 4-1. 喫煙(S)と歯の黄ばみ(Y)と肺がん(L)の因果モデル(左図)。歯の黄ばみに対する介入によって喫煙から歯の黄ばみへのエッジが取り除かれた因果モデル(右図)。

<sup>6</sup> 連鎖公式によると、n 個の離散型確率変数の分布は n 個の条件付き確率の積として分解することができる (e.g.,  $P(X_1, X_2, \dots, X_n) = P(X_n|X_{n-1}, \dots, X_2, X_1) \cdots P(X_2|X_1)P(X_1)$ ).

#### 4. 2. 2. 因果マルコフ条件

因果マルコフ条件 (causal Markov condition) によると, 変数はその直接の原因事象を固定した場合, その結果事象以外の変数と独立になる. 上述の例では, ある人が喫煙していると知ってしまえば, その人の歯が黄色いか否かは肺がんの予測とは無関係となる. このように, ある変数を所与として変数間に独立関係が生じることは条件付き独立と言われる. 因果マルコフ条件を設けることによって(4-2)式や(4-3)式は以下のように書き換えられる.

$$P(S, Y, L) = P(S)P(Y|S)P(L|S) \quad (4-4)$$

(4-2)式における  $P(L|Y, S)$  が(4-4)式では  $P(L|S)$  に単純化されており, 同時確率分布が変数の順序に沿った形で分解されていることが見て取れる. この例における因果マルコフ条件は喫煙と歯の黄ばみに基づく肺がんの予測と喫煙のみに基づく肺がんの予測が等しい (i.e.,  $P(L|Y, S) = P(L|S)$ ) という仮定に他ならない.

因果マルコフ条件の背景には, ある変数の原因となるすべての変数がネットワークに含まれていることが仮定されている. また, 因果のない相関はないという Reichenbach (1956) の共通原因に関する仮定を満たしていることも想定されている. つまり, 変数間に相関関係が生じている時は, どちらか一方の変数が他方の変数の原因となっているか, 両方の変数を引き起こす第三の変数が存在するというわけである. これら 2 つの仮定が満たされるとき, 因果マルコフ条件は正当化される.

#### 4. 2. 3. 介入

因果ベイズネットの特徴の 1 つとして変数に操作を加えた時の結果を予測できることが挙げられる. 変数を強制的にある値に設定することは介入 (intervention) と呼ばれ, 介入によって制御された変数は本来の原因とは独立になる. このことは, 外生変数とし

て介入の変数を加え、介入された事象に対するエッジを取り除くことで表現され、グラフ切除 (graph surgery) や無効化 (undoing) とされる (Pearl, 2000; Sloman, 2005)<sup>7</sup>. 上記の例においてホワイトニングによって歯の黄ばみを除去した場合、喫煙から歯の黄ばみへのエッジが除去され (Figure 4-1 右図), 歯の状態から喫煙の有無や肺がんの状態について予測することはできなくなる. このことは当然のように思われるかもしれないが、観察と介入の違いを表現し、因果関係を適切に記述できるようになったのは Pearl (2000) によって因果ベイズネットが開発されてからのことである.

### 4. 3. 因果ベイズネットの哲学的背景

#### 4. 3. 1. 確率

不確実性を表現する方法として確率を用いることは多いが、その解釈には様々な立場が存在する (Hájek, 2003). 最も一般的なものは頻度主義と呼ばれる立場であり、この立場では同一条件下における無限回の試行の結果、相対頻度が収束したものとして確率を定義する. 例えば、コインを投げた際に表が出る確率は、コインを繰り返し投げて、何回表が出たか数えることで求められる.

一方、ベイズ主義の立場では、ある事象についての個人的な信念の度合いとして確率が定義される<sup>8</sup>. このような確率は主観確率と呼ばれ、測定には様々な方法が存在する (e.g., 戸田, 1982). 確信の程度として確率を捉えることによって、明日雨が降る確率といった未知の事象や火星に生命が存在する確率といった無知に起因する不確定

---

<sup>7</sup> 対象となる変数の値を固定する介入が強い介入 (strong intervention) とも呼ばれるのに対して、変数に影響を及ぼすが、その値を固定しないような操作は弱い介入 (weak intervention) と呼ばれ、この場合は他の変数からのエッジが取り除かれることはなく、依然として本来の原因からの影響も受ける (Eberhardt & Scheines, 2007; Meder, Gerstenberg, Hagmayer, & Waldmann, 2010).

<sup>8</sup> ベイズ主義の中には相対頻度に関する信念の度合いを確率と定義し、客観的な確率解釈に留まろうとする考えもある (e.g., Neapolitan, 2004). なお、ベイズ統計における主観ベイズと客観ベイズの違いは、確率解釈に基づくものではなく、事前確率分布の設定方法の違いに基づくものである (Berger, 2006; Dey & Rao, 2005 繁榘他訳 2011).

な事象に対しても確率を定義することができる。ベイズ統計の領域では主観確率の主観性が問題になることもあるが、認知モデルにおいて主観確率を使用する場合、この問題点はむしろ個人差を扱いやすくなる利点と考えられる。

主観確率が確率法則に従うことは様々な形で示されてきた (Fishburn, 1986)。例えば, Cox (1946) は信念の度合いが満たすべきいくつかの公理を提唱し, 確率がそれらの公理を満たすことを示している。また, de Finetti (1937) は信念の度合いに対する合理的な制限として確率の公理を提唱しており, 確率の公理が満たされない場合には不公平な賭けが生じることを指摘している (cf. 園, 2014)。例えば, コイントスの結果を予測し, 予測が的中すれば 1000 円貰える賭けに参加したとする。コインの表が出ることに對してある人が 0.6 の信念を抱いている場合, その人は 600 円までならお金を払って賭けを行うと考えられる。また同時に, コインの裏が出ることに對しても 0.6 の信念を抱いていたとする。この場合, 賭けの主催者はこれら 2 つの賭けを合わせて 1200 円で売れば, 結果の如何に関わらず必ず 200 円儲けることができる。確率の公理が満たされていない事態では, 相手が必ず負ける賭け (ダッチブック) を仕掛けることが可能になる。主観確率がダッチブックを生み出さない場合には確率の公理が満たされることが知られており, このような正当化はダッチブック論証 (Dutch book argument) とされる。また, 確率の公理からダッチブックが導かれることはない。

#### 4. 3. 2. 因果性

確率の概念は不確実性を表現するためだけではなく, 因果性の定義<sup>9</sup>にも使用されてきた (e.g., Eells, 1991; 大塚, 2010)。Suppes (1970) は原因事象( $C$ )が結果事象( $E$ )の生起確率を上昇させるものとして因果関係を捉えた。

$$P(E|C) > P(E) \quad (4-5)$$

---

<sup>9</sup>因果性の定義に関する様々な議論については Psillos (2005) を参照されたい。

しかしながら、この定義だけでは因果関係と相関関係を区別することはできない。例えば、歯の黄ばみが存在している場合は、そうでない場合よりも肺がんの確率が高くなるが、歯の黄ばみが肺がんを引き起こしているわけではない。両者の関係は喫煙という共通原因によってもたらされる擬似相関だからである (Figure 4-1 参照)。真の原因と判断するには、他の要因 ( $F$ ) を統制した上で(4-5)式の関係が維持されるか確認する必要がある。

$$P(E|C, F) > P(E|F) \quad (4-6)$$

ただし、この式には時間情報が含まれておらず、因果の方向性を決定することはできない。例えば、喫煙が肺がんを引き起こすか ( $S \rightarrow L$ )、肺がんになりやすい人が喫煙習慣を持ちやすいか ( $S \leftarrow L$ ) 判断することはできない。また、真の原因であるかどうかを決定するためには他の事象との因果的関連を考慮に入れなければならないが、それを判断するにはまた別の事象について考慮する必要が生じ、無限後退に陥ってしまう。

因果ベイズネットは変数間の確率的関係を要約しているが、確率のみによって因果性を定義することは稀であり、変数の操作である介入と組み合わせて因果性を定義することが多い。このような立場は介入主義 (interventionism) と呼ばれている。介入主義によれば、変数  $X$  への介入が変数  $Y$  の値に影響を及ぼす時、変数  $X$  は変数  $Y$  の原因である。さらに、他の変数を一定の値に統制した時に介入が影響する場合、変数  $X$  は変数  $Y$  の直接の原因であると言える (Woodward, 2003)。

$$P(e|do(c)) > P(e|do(\neg c)) \quad (4-7)$$

事象に対して何らかの介入を行うことは  $do(\cdot)$  と記述され、 $do(c)$  は変数  $C$  を値  $c$  に固定することを表している。Pearl (2000) によって導入されたこのような表記法は  $do$  オペレ

一タと呼ばれ、観察や介入に基づく確率推論が定式化された。このような介入による因果性の定義は、「もし原因事象が存在しなければ結果事象は存在しなかった」といった反事実による因果性の定義 (e.g., Lewis, 1973) との整合性が認められている。

#### 4. 4. 因果ベイズネットによる因果構造の学習

##### 4. 4. 1. 因果構造の学習とは

事象間の因果関係は、その有無と方向性を表す因果構造と、関係の強さを表す因果強度に二分される。因果構造の学習では変数間の因果の方向性が未知であり、学習者は様々な手がかりに基づいて因果構造を推測するように求められる。一般的な実験事態では3つ以上の2値変数が繰り返し提示され、実験参加者はそれらの生起・非生起といった共変情報や事象の提示順序といった時間情報に基づいて因果構造を推測する。例えば、Steyvers, Tenenbaum, Wagenmakers, and Blum (2003) の実験では、心を読む3人の宇宙人の思考伝播を観察して、どのような方向性で思考伝播が生じているか判断することが実験課題として用いられている。ヒトが因果の方向性をどのように判断しているかという問題は、因果ベイズネットが認知モデルとして適用されてから検討され出した問題であるため、どのような手がかりがどのように利用されているかについてはまだ分かっていないことが多い (Lagnado, Waldmann, Hagmayer, & Sloman, 2007)。

因果ベイズネットでは構造学習について様々なアルゴリズムが考案されている (Neapolitan, 2004; Pearl, 2000)。ここでは制約ベース法 (constraint-based method) とベイズ法 (Bayesian method) という2つの代表的な手法について述べる。これら2つの方法は異なるモデルで採択されており、因果マップ (Gopnik et al., 2004) は制約ベース法を、因果モデル理論 (Waldmann, 1996; Waldmann, Hagmayer, & Blaisdell, 2006; Waldmann & Holyoak, 1992) はベイズ法を用いて因果構造の学習を説明している。

制約ベース法とベイズ法は、共変関係から因果関係を導くために定常性 (stability) もしくは忠実性 (faithfulness) と呼ばれる仮定を設けている。定常性とは、任意のパラメータによって変数間の独立・従属関係が変化しないことである (Pearl, 2000; Spirtes, Glymour, & Scheines, 1993)。このことは、変数間の独立関係は因果マルコフ条件によってのみ生じ、特定のパラメータから生じることはないことを意味している。例えば、変数  $X$  が変数  $Y$  に直接的だけでなく、媒介変数  $Z$  を通して間接的にも影響を及ぼしているとする。このような関係は、 $Z = aX$  と  $Y = bX + cZ$  という2つの構造方程式で記述できる。このような因果モデルでは、ほとんど全てのパラメータにおいて変数  $X$  と変数  $Z$  は従属関係になる。しかしながら、 $a = -bc$  という条件を満たすパラメータでは、直接効果と間接効果が互いを完全に相殺するため、変数  $X$  と変数  $Z$  は独立関係になってしまい、共変関係から正しい因果関係が導かれることは無い。このような問題は理論的には可能であっても、現実的にはほとんど生じないと考えられている。

#### 4. 4. 2. 制約ベース法

制約ベース法は事象間の共変情報を基にボトムアップで因果構造を学習するアルゴリズムであり、2変数間の独立性の検定と因果関係の方向付けという2段階の処理で構成されている (e.g., PC アルゴリズム, Spirtes et al., 1993)。PC アルゴリズムは IC アルゴリズム (Verma & Pearl, 1990) を改良して開発されたものであり、最も代表的な手法である。

PC アルゴリズムによると、最初のステップでは全ての変数を無向辺で結び、無向グラフを構築する。次のステップでは2つの変数を取り出して  $X^2$  検定などを用いて独立性を分析し、独立関係にある2変数間のエッジを取り除く。続いて、他の変数を固定した上での条件付き独立性を調べる。因果構造が異なれば独立・従属関係も変化するため、これらの検定結果は因果構造に制約をもたらす。

方向付けを行うステップにおいても、独立・従属関係は重要となる。変数  $X$  と変数  $Y$

がそれぞれ変数  $Z$  と結びついており (i.e.,  $X-Z-Y$ ),  $Z$  を所与として  $X$  と  $Y$  に従属関係が生じる場合,  $Z$  が共通結果となるようにエッジに向きを付ける (i.e.,  $X \rightarrow Z \leftarrow Y$ ). 上記以外の無向辺については, 共通結果を形成しないようにエッジの方向付けを行う. 例えば, 変数  $X$  が変数  $Z$  に影響しており, 変数  $Y$  が変数  $Z$  と結びついている場合 (i.e.,  $X \rightarrow Z - Y$ ),  $Z$  から  $Y$  へエッジを付ける (i.e.,  $X \rightarrow Z \rightarrow Y$ ). 制約ベース法は事象間の共変情報から因果構造を推測するので領域固有の知識などは必要としないが, それらを用いて更なる制約を掛けることも可能である.

#### 4. 4. 3. ベイズ法

ベイズ法 (e.g., Cooper & Herskovitz, 1992; Heckerman, Geiger, & Chickering, 1995) は共変動や他の手がかり (e.g., 時間順序, 先行知識) を利用してトップダウンで因果構造を学習するアルゴリズムであり, 候補となる因果構造に事前確率を割り当て, ベイズの定理を使用して因果構造の仮説を改訂していく. 事前確率は一様に与えられることもあれば, 先行知識に基づいて特定の因果構造の事前確率が高くなることもある.

事前確率の設定後, それぞれの因果構造において特定の生起パターンに対する尤度が算出される. 例えば, 共通結果の因果モデル (e.g.,  $X \rightarrow Z \leftarrow Y$ ) では原因が独立に生起するため, 3 つの事象が共生起する確率は共通原因の因果モデル (e.g.,  $X \leftarrow Z \rightarrow Y$ ) よりも低くなる. 利用可能なデータにベイズの定理を適用して事後確率が計算され, 事後確率の最も高い因果構造がデータを生成した因果モデルとして選択される. モデルの選択基準には事後確率だけでなくベイズ情報量基準 (Bayesian Information Criterion: BIC) や最少記述長 (Minimum Description Length: MDL) が用いられることもある.

#### 4. 4. 4. 因果構造の学習における観察と介入

制約ベース法やベイズ法では定常性を仮定しても因果構造を特定できないことがある。例えば、変数  $X$  が変数  $Y$  と変数  $Z$  に影響している因果モデル (i.e.,  $Y \leftarrow X \rightarrow Z$ ) の確率分布を因果マルコフ条件に基づいて分解した場合、 $P(X, Y, Z) = P(X)P(Y|X)P(Z|X)$  のように記述される。また、変数  $Y$  が変数  $X$  に影響し、変数  $X$  が変数  $Z$  に影響している因果モデル (i.e.,  $Y \rightarrow X \rightarrow Z$ ) の場合、その確率分布は  $P(X, Y, Z) = P(Y)P(X|Y)P(Z|X)$  のように記述される。ここで  $P(X)P(Y|X) = P(Y)P(X|Y) = P(X, Y)$  となるため、これら 2 つの因果モデルの確率分布は同一となる。このような因果モデルはマルコフ等価モデル (Markov equivalent model) と呼ばれており、観察によって得られた共変情報からは因果構造を一意に定めることができない。

介入によって変数の状態を操作することは、マルコフ等価モデルの区別を可能にする。変数  $X$  が変数  $Y$  と変数  $Z$  に影響する共通原因モデル (i.e.,  $Y \leftarrow X \rightarrow Z$ ) において  $X$  を発生させる場合、その結果である  $Y$  と  $Z$  も生起することが予測される。一方、変数  $Y$  が変数  $X$  を媒介して変数  $Z$  に影響する因果連鎖モデル (i.e.,  $Y \rightarrow X \rightarrow Z$ ) において  $X$  を発生させる場合、 $X$  の結果である  $Z$  は生起するが、 $X$  の原因である  $Y$  が生起することは無い。介入を行うことによって共通原因モデルと因果連鎖モデルでは異なる生起パターンが生じるようになり、マルコフ等価モデルを区別することが可能になる。このことは、調査観察データから因果関係を実証することは困難であるが、実験研究では可能になるという研究法の話と対応する。

介入による因果構造の学習の促進効果は多くの研究で認められている (e.g., Steyvers et al., 2003; Lagnado & Sloman, 2004; Sobel & Kushnir, 2006)。Steyvers et al. (2003) は、観察による学習に比べて介入による学習では因果構造の判断の正答率が上昇することを示した。Lagnado & Sloman (2004) は介入による因果構造の学習の促進効果に時間順序手がかりが含まれていることを明らかにした。また、介入と時間順序手がかりは加算的に促進効果をもたらすことが明らかにされている

(Lagnado & Sloman, 2006). さらに, 介入によって因果構造を学習する際にはどの事象に介入するかという意思決定も重要であることが Sobel & Kushnir (2006) によって示されている. 介入による促進効果は就学前児を対象とした研究においても確認されており (Burns & McCormack, 2009; Schulz, Gopnik, & Glymour, 2007; but see also Kushnir, Wellman, & Gelman, 2009), 関連する様々な要因について実験的検討がなされている (e.g., Sobel & Sommerville 2009, 2010).

#### 4. 5. 因果ベイズネットによる因果強度の学習

##### 4. 5. 1. 因果強度の学習とは

因果構造が因果の方向性を表すのに対して, 因果強度は原因と結果の関係の強さを表す. 因果強度の学習では, 原因候補と結果が予め決められており, 学習者はそれらの因果関係の強さを判断することが求められる. なお, 原因候補と表現するのは, 実際には無関係である可能性も含まれているからである. 一般的な実験事態では原因候補と結果の 2 値 2 変数が使用され, これらの共変情報は Figure 4-2 に示すような 2×2 の分割表で表現することができる. 実験参加者はこれらの情報を継時的もしくは要約的

		Event $E$	
		$e$	$\neg e$
Event $C$	$c$	$a$	$b$
	$\neg c$	$c$	$d$

Figure 4-2. 原因事象( $C$ )と結果事象( $E$ )の分割表. アルファベット  $a$ ,  $b$ ,  $c$ ,  $d$  はそれぞれの共生起の頻度を表す.

に取得した後、0 から 100 の間で因果強度を評価する。例えば、Buehner, Cheng, and Clifford (2003) の実験では、放射線被曝がウイルスの突然変異を引き起こすか調べるというカバーストーリーの下で、被曝の有無と突然変異の有無に関する情報が与えられ、放射線がどのくらい変異を引き起こすか判断することが求められた。

原因候補と結果の共変関係の強さは、以下に示す  $\Delta P$  で記述されることが多い (e.g., Allan, 1980; Jenkins & Ward, 1965; Ward & Jenkins, 1965)。

$$\Delta P = P(e|c) - P(e|\neg c) \quad (4-8)$$

ここで  $P(e|c)$  は原因候補が生起した時に結果が生起する条件付き確率を、 $P(e|\neg c)$  は原因候補が生起しなかった時に結果が生起する条件付き確率を表す。 $\Delta P$  は -1 から +1 までの値をとり、その値によって異なった意味を持つ。原因候補の生起によって結果の生起確率が上昇する場合 (i.e.,  $\Delta P > 0$ )、原因と結果の間には発生的因果関係が想定される。一方、原因候補の生起によって結果の生起確率が低下する場合 (i.e.,  $\Delta P < 0$ )、原因と結果の間には抑制的因果関係が想定される。原因候補の有無によって結果の生起確率が変化しない場合 (i.e.,  $\Delta P = 0$ )、原因候補と結果は無関係となる。

因果強度の評定値と  $\Delta P$  にはおおまかな対応関係がみられたことから (e.g., Allan & Jenkins, 1980; Alloy & Abramson, 1979; Dickinson, Shanks, & Evenden, 1984; Wasserman, Elek, Chatlosh, & Baker, 1993; Ward & Jenkins, 1965),  $\Delta P$  ルールとして説明モデルに使用されるようになった。しかし、 $\Delta P$  から逸脱した結果も報告されており (e.g., Shanks, 1985), その後、数々の説明モデルが提唱された (see Hattori & Oaksford, 2007; Perales & Shanks, 2007)。

#### 4. 5. 2. パワーPC 理論

Cheng (1997) は  $\Delta P$  ルールの主要な問題点として、以下の 2 点を指摘している。原

因の有無に関わらず結果が常に生起する天井効果の事態では (i.e.,  $P(e|c) = P(e|\neg c) = 1$ ), 実際には因果効果を判断することはできないが,  $\Delta P$  ルールは効果がないと判断してしまう. また,  $\Delta P$  は他の原因事象を考慮していないため, 因果効果の適切な推定値とはならない.

彼女はこれらの問題点を克服する新たな理論として, パワーPC 理論 (Cheng, 1997; Novick & Cheng, 2004) を提唱し, 共変動と因果を区別する因果パワー (causal power) を定式化した (cf. 斎藤・嶋崎, 2009). 因果パワーとは, 他の原因が存在しない事態において原因事象が結果事象を発生させる確率である. パワーPC 理論は因果ベイズネットとは独立に提唱されたが, 特定の状況下では両者が一致することが後に指摘されている (Glymour, 2001, 2003).

パワーPC 理論では, ヒトが因果関係を学習する際に以下の 4 つの信念を抱いていると仮定する. (1)背景原因( $B$ )<sup>10</sup>と原因候補( $C$ )は独立に結果( $E$ )に影響する. (2)背景原因によって結果が発生することはあるが, 抑制されることはない. (3)背景原因と原因候補の因果パワーはそれらの生起頻度とは無関係である. (4)原因が存在しない時に結果が生じることはない. これらの仮定が満たされる場合, 発生的因果パワー $q_G$  は以下の式から算出される.

$$q_G = \frac{\Delta P}{1 - P(e|\neg c)} \quad (4-9)$$

$q_G = 1$  は原因が結果を必ず発生させることを意味し,  $q_G = 0$  は影響を及ぼさないことを意味する. 背景原因によって結果が常に生起する, つまり天井効果が生じている場合 (i.e.,  $P(e|c) = P(e|\neg c) = 1$ ), 因果パワー $q_G$  は定義されない. また, 背景原因が存在しな

---

<sup>10</sup> 厳密には背景原因( $B$ )は原因候補( $C$ )以外の全ての原因を表し, その中には観測不能な原因も含む.

い場合 (i.e.,  $P(e|\neg c) = 0$ ),  $q_G = \Delta P$  となり, 共変動と因果は一致する. Buehner et al. (2003) は  $\Delta P$  を一定に保ちつつ,  $P(e|\neg c)$  を体系的に操作し,  $\Delta P = 0$  の事態を除いて実験参加者の評定値がパワーPC理論の予測値と一致することを報告している.

#### 4. 5. 3. 因果サポートモデル

Griffiths and Tenenbaum (2005, 2009) は, これまでパラメータ推定の問題として扱われていた因果強度の判断をモデル比較の問題として捉え直し, 原因候補と結果の間に因果関係が存在するかという構造の判断から説明を試みた. ここで言うモデルとは, Figure 4-3 に示すような背景原因 ( $B$ ) のみが結果 ( $E$ ) に影響を及ぼす因果モデル (Graph 0) と背景原因 ( $B$ ) と原因候補 ( $C$ ) が結果 ( $E$ ) に影響する因果モデル (Graph 1) のことである. 背景原因と原因候補が独立に結果を生起させる場合, 結果の生起確率は noisy-OR 関数を用いて以下のように算出される.

$$P(e|b, c; w_0, w_1) = w_0b + w_1c - w_0w_1bc \quad (4-10)$$

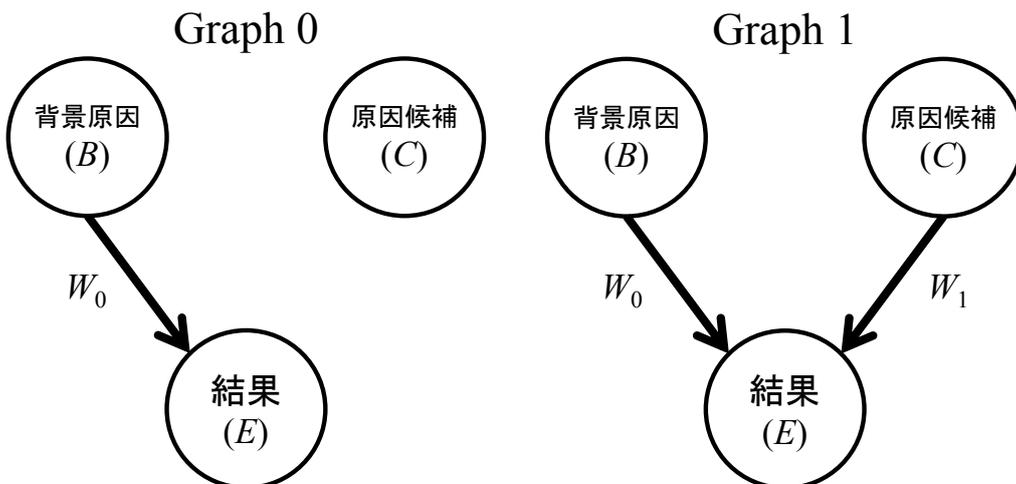


Figure 4-3. 背景原因 ( $B$ ) と原因候補 ( $C$ ) と結果 ( $E$ ) の因果モデル.  $W_0$  と  $W_1$  は背景原因と原因候補の因果強度をそれぞれ表す.

ここで  $w_0$  は背景原因の因果強度を,  $w_1$  は原因候補の因果強度を表す. また,  $b$  と  $c$  はそれぞれ背景原因と原因候補の有無を表し, 存在するときは 1, 存在しないときは 0 になる. なお, 発生的因果パワー  $q_G$  は noisy-OR 関数の下での  $w_1$  の最尤推定量となることが示されている.

ベイズの定理を適用することで, 得られたデータが Graph 0 によって生じたか, Graph 1 によって生じたかは対数オッズの形で以下のように表現できる.

$$\log \frac{P(\text{Graph 1}|D)}{P(\text{Graph 0}|D)} = \log \frac{P(D|\text{Graph 1})}{P(D|\text{Graph 0})} + \log \frac{P(\text{Graph 1})}{P(\text{Graph 0})} \quad (4-11)$$

この式の右辺の第一項である対数尤度比はどちらの仮説が尤もらしいかを表す指標であり, ベイズファクター (Kass & Raftery, 1995) と呼ばれている. Griffiths and Tenenbaum (2005) はこの対数尤度比を因果サポートと名づけ, データが Graph 0 よりも Graph 1 を支持する指標とした.

$$\text{causal support} = \log \frac{P(D|\text{Graph 1})}{P(D|\text{Graph 0})} \quad (4-12)$$

なお, それぞれの尤度は以下の式から算出される.

$$P(D|\text{Graph 0}) = \int_0^1 P(D|w_0, \text{Graph 0})P(w_0|\text{Graph 0})dw_0 \quad (4-13)$$

$$P(D|\text{Graph 1}) = \int_0^1 \int_0^1 P(D|w_0, w_1, \text{Graph 1})P(w_0, w_1|\text{Graph 1})dw_0dw_1 \quad (4-14)$$

ここで  $P(w_0|\text{Graph}0)$ と  $P(w_0, w_1|\text{Graph}1)$  は因果強度の事前分布となっており, 事前分布には 0 から 1 までの一様分布が想定されている.

因果サポートモデルは,  $\Delta P$  ルールやパワーPC 理論では説明困難であった  $\Delta P = 0$  の事態における実験参加者の評定値に適合するだけでなく, サンプルサイズの効果や不完全な分割表における判断や比率データからの学習を上手く説明することができていた (Griffiths & Tenenbaum, 2005).

#### 4. 5. 4. SS パワーモデル

Lu, Liljeholm, Cheng, and Holyoak (2008) は, 因果強度の事前分布に偏りが無い一様分布を想定するよりも, 因果の単純性 (Chater & Vitányi, 2003) を反映した分布を想定する方が適切であると主張している. 因果関係における単純性とは, 原因事象の数を最少にしつつ因果強度を最大にする因果モデルへの選好である.

Lu et al. (2008) は(4-14)式における因果強度の事前分布に必要な十分な因果モデルへの選好を組み込み, この事前分布に基づくベイズモデルを SS パワーモデルと名づけた. 背景原因と原因候補が結果を発生させる場合, 事前分布は以下のように表現される.

$$P(w_0, w_1|\text{gen}, \text{Graph} 1) \propto (e^{-\alpha w_0 - \alpha(1-w_1)} + e^{-\alpha(1-w_0) - \alpha w_1}) \quad (4-15)$$

ここで  $\alpha$  は必要十分な原因への選好を示すパラメータであり,  $\alpha = 0$  となる時, 事前分布は一様分布となる. この事前分布は原因候補の因果強度と背景原因の因果強度が両立しないこと, つまり, 一方の値が高く想定される場合は他方の値が低く想定されることを示している.

SS パワーモデルは, 複数の実験のデータセット (Perales & Shanks, 2007) に対して因果サポートモデルよりも良い当てはまりを示すだけでなく, 発生的因果関係と抑制

的因果関係の差異をうまく説明できていた。また、「原因は結果をどの程度発生させるか？」という強度の判断と「因果関係はどのくらいありそうか？」という構造の判断の違いも記述することができていた。

#### 4. 5. 5. 原因事象が複数存在する場合の因果強度の学習

単一の原因事象に対する因果強度の判断では、背景原因のみが影響するモデルと、原因事象と背景原因の両方が影響するモデルの2つの因果モデルを比較していたが、複数の原因事象が想定される場合はより多くの因果モデルを考慮する必要がある。例えば、原因事象 A と B が結果事象に影響を及ぼしているか考える場合、A のみが影響するモデル、B のみが影響するモデル、A と B の両方が影響するモデル、どちらも影響しないモデルという4つの因果モデルの比較が行われる。

Carroll, Cheng, and Lu (2013) は、回顧的再評価 (retrospective reevaluation)<sup>11</sup>と呼ばれる現象に関して、連合学習の代表的なモデルと因果ベイズネットの比較を行っている。回顧的再評価の代表的な例である逆行ブロッキングの手続きでは、原因事象 A と B が結果事象 E と対呈示された後、片方の原因 A と結果 E が対呈示される。例えば、リンゴ(A)とバナナ(B)を食べた後にアレルギー症状(E)が出たとする。この段階ではどちらが原因であるか分からないため、どちらも危険だと判断される。しかし、その後にリンゴ(A)だけを食べてアレルギー症状(E)が出た場合、リンゴがアレルギー症状の原因だと判断され、バナナは安全だと判断される。当初はどちらも危険だと判断していたものが、その後の経験により、過去を顧みて評価され直すわけである。Carroll et al., (2013) は、連合学習モデルと因果ベイズネットのシミュレーションの結果と行動実験の結果を照らし合わせ、因果ベイズネットの方がデータの当てはまりがよいことを示している。また、Griffiths, Sobel, Tenenbaum, and Gopnik (2011) は、因果関係が

---

<sup>11</sup> この現象はヒトおよび動物の高次認知過程を理解するための試金石として様々な領域で研究されている (see 川合・久保(川合), 2008)。

存在する確率はどのくらいか、因果関係は決定的か確率的かといった事前知識の影響を検討し、それらが因果ベイズネットでうまく説明されることを報告している。

#### 4. 6. 因果ベイズネットによる因果推論

##### 4. 6. 1. 因果推論とは

因果構造の学習や因果強度の学習が因果の知識の獲得を扱うのに対して、因果推論では獲得した因果の知識をどのように使用するかということに焦点が当てられている。一般的な実験事態では、学習者に因果構造を教示して事象間の確率関係を学ばせた後で、ある事象の状態から別の事象に対して推測を行うことが求められる。使用される因果構造には、複数の原因が単独の結果に影響する共通結果モデル、単独の原因が複数の結果を生じさせる共通原因モデル、原因が媒介事象を通して結果に影響を及ぼす因果連鎖モデルなど様々なものが含まれる (see Rottman & Hastie, 2014 for a review). 例えば, Waldmann and Hagmayer (2005) の実験では, 1つの原因が2つの結果をもたらす共通原因の因果モデルにおいて, 一方の結果事象の状態から他方の結果事象の状態を推測することが求められた。ここでは, 様々な因果モデルにおいて因果ベイズネットがどのような因果推論を行うか説明する。

##### 4. 6. 2. 直接原因モデル ( $C \rightarrow E$ ) における因果推論

原因事象から結果事象を推測することは予測的推論と言われる。原因事象  $C$  が生起している時に結果事象  $E$  が生起する確率はその条件付き確率と一致する (i.e.,  $P(e|c)$ ). ただし, 原因事象が生起している時に背景原因も生起していて, その背景原因によって結果が生起する可能性もあるため,  $P(e|c)$ を原因  $C$  の結果  $E$  に対する因果強度と見なすことはできない。Fernbach, Darlow, and Sloman (2011) は, 予測的推論において背景原因の影響が考慮されないことを報告している。

結果事象から原因事象を推測することは診断的推論と言われる。結果事象  $E$  が生起している時に原因事象  $C$  が生起している確率はベイズの定理によって以下のように算出される。

$$P(c|e) = \frac{P(e|c)P(c)}{P(e)} = \frac{P(e|c)P(c)}{P(e|c)P(c) + P(e|\neg c)P(\neg c)} \quad (4-16)$$

より因果的な表現として、原因が生起している時に結果が生起する確率  $P(e|c)$ ではなく、原因が結果を発生させる確率  $w_1$ と背景原因が結果を発生させる確率  $w_0$ を用いて診断的推論を記述することもできる (see Meder, Mayrhofer, & Waldmann, 2014)。

(4-16)式において原因事象が生起する確率  $P(c)$ は基準率 (base rate) と呼ばれ、診断的推論に基準率が正しく反映されないことは基準率錯誤 (base rate fallacy) と呼ばれている (e.g., Bar-Hillel, 1980; see Barbey & Sloman, 2007 for a review)。基準率錯誤については、基準率は無視されるものであるという説 (Tversky & Kahneman, 1980) や、自然頻度の形で提示されれば使用されるという説 (Gigerenzer & Hoffrage, 1995) や、心的表象次第であるという説 (Hattori & Nishida, 2009; Krynski & Tenenbaum, 2007) など様々な立場が存在する。基準率錯誤に関する実験の多くは統計的關係を対象としたものであるが、因果關係に基づく推論を対象とした実験では基準率が考慮されることが示されている (Waldmann & Hagmayer, 2005; Meder, Hagmayer, & Waldmann, 2009)。

#### 4. 6. 3. 因果連鎖モデル ( $C \rightarrow M \rightarrow E$ ) における因果推論

原因事象  $C$  が媒介変数  $M$  を通して結果事象  $E$  に影響する因果連鎖モデルでは、原因から結果への推移的推論が検討されている。Baetu and Baker (2009) は原因  $C$  と媒介変数  $M$  の関係や媒介変数  $M$  と結果  $E$  の関係をそれぞれ学習させた後に、原

因と結果の関係について判断を求めた。規範的な推測と一致するように、どちらも正の相関か負の相関の場合には結果の生起を予測し、一方が正の相関で他方が負の相関の場合は結果が生起しないことを予測することが示された。類似した結果は二次条件づけの手続きを用いた Jara, Vila, and Maldonado (2006) でも報告されている。

#### 4. 6. 4. 共通結果モデル ( $C_1 \rightarrow E \leftarrow C_2$ ) における因果推論

因果マルコフ条件によると、共通結果モデルでは結果事象の状態が不明な場合、2つの原因事象は独立関係にあるが、結果の状態が明らかになると両者には従属関係が生じる。そのため、結果  $E$  が生じたときに原因  $C_1$  が生起している確率  $P(c_1|e)$  と結果  $E$  と片方の原因  $C_2$  が生じたときに残りの原因  $C_1$  が生起している確率  $P(c_1|e, c_2)$  は異なるものになる。結果の生起は片方の原因によって説明されるため、前者の確率  $P(c_1|e)$  が後者の確率  $P(c_1|e, c_2)$  よりも高くなることが知られている (see Khemlani & Oppenheimer, 2010)。一方の原因の生起によって他方の原因の生起確率が減少することは、心理学の領域では割引効果 (discounting effect) として (e.g., Kelley, 1973)、人工知能の領域では言い逃れの効果 (explaining away effect) として知られている (Pearl, 2000)。Morris and Larrick (1995) では、実験参加者にエッセイを読ませてそれぞれの確率を尋ねたところ、割引効果を反映した判断を行うことが実証されている。

#### 4. 6. 5. 共通原因モデル ( $E_1 \leftarrow C \rightarrow E_2$ ) における因果推論

因果マルコフ条件によると、共通原因モデルでは原因事象  $C$  の状態が明らかになると、それぞれの結果事象  $E_1 \cdot E_2$  は独立関係になることが想定されている。つまり、結果  $E_1$  が生起するかは原因  $C$  によって決まるため、結果  $E_2$  の状態は影響しないわけである。しかしながら、他方の結果の状態によって判断が変化することが知られており (e.g., Rehder & Burnett, 2005)、このことはマルコフ条件の逸脱 (Markov violation) と

言われている。この逸脱に対しては因果関係のメカニズムに関する知識がそれぞれの事象に隠れた原因として影響しているといった説明がなされており (e.g., Rehder, 2014; Rehder & Burnett, 2005), メカニズムに関する知識を操作した実験ではこの仮説を支持する結果が得られている (Park & Sloman, 2013)。

#### 4. 6. 6. 因果推論における観察と介入

ここまで説明してきた因果推論では、ある事象の状態が判明した時に他の事象の状態についてどのような推測が成立するかということに焦点が当てられていた。このような推論は観察に基づく推論であり、因果モデルとそのパラメータによってその内容が決まる。一方、事象に対して操作を行った場合の推論は介入に基づく推論であり、両者は明確に区別される。例えば、外に出たときに地面が濡れていれば雨が降ったかもしれないとその原因を推測することができるが、庭に水を撒いた後に地面の濡れから降雨に対して推測を行うことはない。介入によって制御された事象 (e.g., 地面の濡れ) は本来の原因 (e.g., 降雨) とは独立になるからであり、このことはグラフ切除や無効化と言われる (Pearl, 2000; Sloman, 2005)。グラフ切除は介入された事象に対するエッジを取り除くことで表現され、この場合の原因事象の生起確率は基準率に従う。

$$P(c|do(e)) = P(c|do(\neg e)) = P(c) \quad (4-17)$$

因果推論における観察と介入の違いは多くの研究によって示されてきた。Sloman and Lagnado (2005) は複数の事象が関連している事態で特定の事象に対して介入を行った場合に他の事象についてどのような推測がなされるか検討した結果、介入された事象が本来の原因と独立に扱われていることを示した。Waldmann and Hagmayer (2005) は、事象の生起情報を表に記載した要約提示の因果学習課題を用いて、ヒトが観察による学習から介入の結果を正しく予測できることを報告している。同様の結果

は、事象の生起情報を 1 試行毎に提示する事例提示の因果学習課題においても得られている (Meder, Hagmayer, & Waldmann, 2008, 2009). 近年では, ラットでさえも観察と介入を区別して因果推論を行っているということを示唆する研究も報告されている (Blaisdell, Sawa, Leising, & Waldmann, 2006; Leising, Wong, Waldmann, & Blaisdell, 2008; but see also Dwyer & Honey, 2009).

## 5. 博士論文研究の目的

### 5. 1. 序論

環境内の因果関係に関する情報は、過去の説明および現在の制御ならびに未来の予測を可能にする手段を我々に提供する。因果関係をどのように認識するかという問題は Hume や Kant を初めとする数多くの哲学者の論考を背景に、心理学においても様々なアプローチから解決が試みられてきた。連合的アプローチでは因果学習を原因と結果の連合強度の変化と捉え、規則ベースアプローチでは因果学習は何らかの単純な計算規則に基づいて行われると想定されている。しかし、これらのアプローチに対しては、因果関係と相関関係が区別されていない、原因と結果の非対称性が考慮されていないといった問題点が指摘されている (see Cheng & Holyoak, 2011 for a review).

近年、これらの問題点を克服する新たなアプローチとして因果ベイズネット (causal Bayes nets) を用いたベイズモデルによる説明がなされている (e.g., Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Griffiths & Tenenbaum, 2005, 2009; Lu, Liljeholm, Cheng, & Holyoak, 2008). 因果ベイズネットとは、既存の情報から因果関係を適切に導き出し、与えられた情報から推測を行うために統計学や計算機科学といった分野で開発された数理的手法である (Glymour, 2001; Pearl, 2000). 因果の知識をどのように獲得して使用するかという問題を因果ベイズネットの観点から考えた場合、何が原因で何が結果であるかという「因果構造の学習」と、原因はどのくらい強く結果に影響するかという「因果強度の学習」と、原因や結果の状態からどのような推測が可能かという「因果推論」の3つの側面に分類することができる。因果ベイズネットを用いることで因果性の学習と推論に関する一部の側面はうまく説明されるものの、まだ十分に検討されていない側面も数多く存在している。さらに、研究者ごとに異なるベイズモデルが提唱されており、それぞれのモデルに関する比較検討もほとんど行われていない。

本博士論文研究の目的は、因果構造の学習と因果強度の学習と因果推論の 3 つの側面から因果ベイズネットの心理学的妥当性を検討することである。因果性の学習と推論の認知プロセスを探求する際に因果ベイズネットを使用する利点の 1 つは、因果ベイズネットが合理性を定義する手段として機能することである。ベイズの定理を使用することによって、「新たに情報が得られた場合に信念をどのように改訂すべきか」という問題に対する規範的な答えを計算することができる。因果ベイズネットによる規範的な振る舞いとヒトの行動を比較検討することは、それらが一致する場合にはその行動に対する説明を、一致しない場合には計算を阻害する要因についての示唆をもたらす。これらの検討により、ヒトがどのように振る舞うべきかという規範的モデルから、ヒトが実際にどのように振る舞うかという記述的モデルを構築することが可能になる。

## 5. 2. 因果構造の学習

事象間の因果関係は、その有無と方向性を表す因果構造と、関係の強さを表す因果強度に二分される。因果構造の学習では事象間の因果の方向性が未知であり、学習者は事象の生起・非生起といった共変動手がかりや事象の提示順序といった時間順序手がかりに基づいて因果構造を推測する。因果ベイズネットでは共変動手がかりに基づいて因果構造を学習することが想定されているのに対し、先行研究では共変動手がかりよりも時間順序手がかりが重視されることが報告されている (Lagnado & Sloman, 2006)。

研究 1 では、先行研究の結果が複雑な実験課題に起因するものと考え、2 事象から構成される最も単純な因果構造の判断課題を用い、共変動手がかりと時間順序手がかりのどちらが使用されるか検討した (実験 1)。また、課題中に繰り返し判断を求めることによって、共変動手がかりと時間順序手がかりが学習過程でどのように影響を及ぼしているか検討した (実験 2)。さらに、共変動情報を体系的に操作し、共変動手がかりがどのように使用されるかについても検討を行った (実験 3)。続けて、実験課題のカバー

ストーリーを変更して、同様の検討を行った（実験 4）。

研究 2 では、共変動手がかりと時間順序手がかりの使用が課題の複雑性によってどのように変化するか検討した（実験 1）。また、課題の複雑性に伴う判断方略の変化が異なる因果モデルでも確認されるか調べた（実験 2）。事象の状態を見て因果関係を学習する場合（i.e., 観察）と事象を操作して因果関係を学習する場合（i.e., 介入）とでは、課題の複雑性の影響がどのように異なるかについても検討した（実験 3）。

### 5. 3. 因果強度の学習

因果強度の学習では、原因事象と結果事象があらかじめ決められており、学習者は 2×2 の分割表で表現可能な共変動情報から原因と結果の関係の強さを評価する。因果強度の学習に関しては、因果バイズネットだけでなく、規則ベースアプローチや連合的アプローチからも各種モデルが提唱されている（Hattori & Oaksford, 2007; Perales & Shanks, 2007）。

研究 3 では、共変動情報を操作した各種実験の結果をまとめ、各種モデルの比較検討を行った。Perales & Shanks (2007) で集められた研究結果をデータセットとして、規則ベースアプローチと連合的アプローチとバイズ的アプローチの代表的なモデルのシミュレーションを実施した。また、複数のバイズモデルについてデータとの適合度を検討した。

研究 4 では、因果強度の学習における学習方法に焦点を当てた。観察による学習では事象の状態を見て因果関係を学習し、介入による学習では事象を操作して因果関係を学習する。観察と介入のどちらがより正確な学習をもたらすか検討した（実験 1）。また、実験課題の因果モデルを変更し、より詳細な検討を行った（実験 2）。

#### 5. 4. 因果推論

因果推論では、学習者は獲得した因果関係の知識を基に、ある事象の状態から別の事象の状態を推測することが求められる。事象の状態が観察によって得られたものか、介入によって得られたものかによって、推測結果は大きく異なる (Meder, Hagmayer, & Waldmann, 2008, 2009; Waldmann & Hagmayer, 2005)。因果ベイズネットは因果推論における観察と介入の違いを説明するだけでなく、観察による学習から介入を行った際の結果が推測可能なことを予測する。なお、規則ベースアプローチや連合的アプローチのモデルでは、これらを説明することは困難である。

研究 5 では、Meder et al. (2008) と同様の手続きを用いて観察による学習から観察および介入の結果を予測できるか確認した (実験 1)。また、学習方法を介入による学習に変更し、介入による学習から観察および介入の結果を予測できるか検討した (実験 2)。

研究 6 では、介入に基づく因果推論を意思決定場面にまで拡張した選択の因果モデル理論 (Hagmayer & Sloman, 2009; Sloman & Hagmayer, 2006) に焦点を当てた。選択に基づく推論や介入に基づく推論や観察に基づく推論において、因果構造が考慮されるか否か検討した (実験 1)。また、これらの推論に因果モデルのパラメータの 1 つである基準率が反映されるか否か検討を行った (実験 2)。さらに、これらの推論において原因と結果間の因果強度が考慮されるか否かも検討した (実験 3)。

## 第 2 部

# 因果構造の学習

## 6. 研究 1 因果構造の学習における共変動情報の役割<sup>12</sup>

### 6. 1. 序論

#### 6. 1. 1. はじめに

複雑に絡み合った事象の中から何が原因で何が結果となっているか判断することは難しい。因果構造を推測する際の基本的な手がかりとして、共変動手がかり (covariation cue) と時間順序手がかり (temporal order cue) が挙げられる。研究 1 では、因果構造の学習における共変動手がかりに焦点を当てる。具体的には、共変動手がかりと時間順序手がかりが一致しない時どちらの手がかりが重視されるか、また共変動手がかりから因果の方向性はどのように判断されるかについて検討を行った。

#### 6. 1. 2. 共変動手がかりと時間順序手がかり

共変動手がかりとは、変数の同時確率分布であり、2 値変数の場合には  $2 \times 2$  の分割表を用いて表現することができる (Figure 6-1)。因果関係が存在する場合には、2 つの事象に強い共変関係が期待される。一方、共変関係がないことは、2 つの事象が無関係であることを示唆する。しかしながら、共変動手がかりの使用にはいくつかの限界が指摘できる。第一に、対象となる事象の数が増加すると、共変動情報はより複雑になる。例えば、2 値 (i.e., 生起, 非生起) 2 事象の場合生起パターンは 4 種類であるが、2 値 5 事象の場合は 32 もの生起パターンが生じ得る。第二に、共変動手がかりだけでは区別できない因果モデルが存在する。例えば、事象 X と事象 Y が繰り返し共生起した場合には両者の間に何らかの因果関係があると想定されるが、事象 X が事象 Y を引き起こしているのか (i.e.,  $X \rightarrow Y$ )、事象 Y が事象 X を引き起こしているのか (i.e.,  $X \leftarrow Y$ )、あるいは第三の原因 Z によって両者が引き起こされているのか (i.e.,  $X \leftarrow Z \rightarrow Y$ )

---

<sup>12</sup>本研究の一部は 2009 年度関西学院大学大学院文学研究科修士論文として提出されている。

判断することはできない。これらの例は共変動手がかかり単独で因果構造を推測することが困難なことを示している。ただし、他の原因は存在しないということが分かっているならば共変動手がかかりはより有益なものとなる。このことは、第三の原因による擬似相関という可能性を排除する。さらに一方の事象  $X$  が単独生起していれば、他方の事象  $Y$  が結果であると推測できる。原因が存在しない時に結果が生起することはないという因果の必要性から、単独生起する事象は結果事象ではなく、原因事象と考えることができる。

因果構造を推測する際のもう 1 つの基本的な手がかりは、事象の状態を観察する時間順序である。原因は結果に先行して観察されやすいことから、事象  $X$  が事象  $Y$  に先立って観察された場合には事象  $X$  が事象  $Y$  に影響するという因果関係が示唆される。しかしながら、観察する時間順序が因果の順序と一致するとは限らない。原因よりも結果を先に観察する事態では、誤った因果の方向性が導かれる。また、時間順序手がかりによって擬似相関を因果関係と捉えてしまうこともある。事象  $X$  を観察した後に事象  $Y$  を観察し、 $X$  から  $Y$  への因果関係を推測しても、両者が共通の原因  $Z$  によって生じた擬似相関の可能性もある。時間順序手がかりは因果構造の学習を促進することもあれば、誤った推論を導くこともある。

		<b>Event <math>E</math></b>	
		$e$	$\neg e$
<b>Event <math>C</math></b>	$c$	$a$	$b$
	$\neg c$	$c$	$d$

Figure 6-1. 原因事象( $C$ )と結果事象( $E$ )の分割表。アルファベット  $a$ ,  $b$ ,  $c$ ,  $d$  はそれぞれの共生起の頻度を表す。

### 6. 1. 3. 本研究の目的

因果構造の学習に関するこれまでの研究から共変動手がかりに基づいて因果の方向性を判断するという証拠は得られていない。Lagnado and Sloman (2006) は、事象 A が事象 B を発生させ、事象 B が事象 C と事象 D を発生させている事態 (i.e.,  $A \rightarrow B \rightarrow C \ \& \ D$ ) において、実験参加者に実際の因果関係とは異なる順序 (e.g.,  $A \rightarrow D \rightarrow C \rightarrow B$ ) でそれらの状態を観察させ、共変動手がかりと時間順序手がかりのどちらに基づいて因果構造を判断するか検討した。実験の結果、実験参加者は時間順序手がかりが当てにならないと教示されていたにも関わらず、因果構造を時間順序手がかりに基づいて判断していた。同様の傾向は共変動手がかりから因果構造を推測する方法を教示に明示した White (2006) においても確認されている。これらの知見に基づき、Fernbach and Sloman (2009) は因果構造の学習において共変動手がかりから因果の方向性が引き出されることはない述べている。

しかし、共変動の使用に関するこれまでの研究では事象の数が見落とされてきた。事象の数が増えるほど共変動は複雑になり、当然手がかりとして使用しにくくなる。実際、Lagnado & Sloman (2006) では 4 つの事象、White (2006) では 5 つの事象からなる因果構造を共変動手がかりに基づいて判断することが求められていた。共変動手がかりに基づいて因果構造を判断できるか否か調べるにはより単純な因果構造を用いた事態での検討が必要である。

そこで本研究では、2 つの事象から構成される実験事態において因果構造の学習に共変動の手がかりが使用されるか否か検討することを目的とした。実験 1 では共変動手がかりと時間順序手がかりと事前知識の相互作用を検討した。続く実験 2 では、因果構造の学習過程に焦点を当てた。実験 3 と実験 4 では共変動情報からどのようにして因果の方向性が導き出されるかについて検討を行った。因果ベイズネットでは共変動手がかりに基づいて因果構造の学習を行うため、共変動手がかりが使用されるかどうかは因果ベイズネットの心理学的妥当性を考えるうえで重要な問いとなる。

## 6. 2. 実験 1

実験 1 の目的は、共変動情報から因果の方向性を引き出すことが可能かどうか検討することであった。実験課題は 2 種類の架空の生物の発生を観察し、それらの因果関係を判断することであった。共変動手がかりと時間順序手がかりの不一致を作るために、実験参加者が 2 つの事象を観察する時間順序を操作した。実験参加者が時間順序手がかりよりも共変動手がかりを優先する場合、時間順序に関わらず一貫した判断が見られるはずである。一方、共変動手がかりよりも時間順序手がかりを重視する場合、時間順序によって異なる判断が生じるはずである。また、他の原因が存在しないことや時間順序が当てにならないことの教示を操作し、これら 2 つの判断方略の境界条件を検討した。

### 6. 2. 1. 方法

**実験参加者および実験計画** 大学生 96 名が実験に参加し、成績への加点となるコースクレジットを受け取った。実験参加者は、時間順序・時間順序の信頼性の教示・他の原因の不在の教示の 3 つの要因の操作から生じる 8 群のいずれかに無作為に割り当てられた。

**教示** 実験参加者に印刷された教示を手渡し、実験者がその内容を読み上げた。教示の概要は次の通りであった。

以下に記されているストーリーの登場人物になることを想像してください。あなたはとある研究所で微生物の研究を行っています。現在の研究テーマは「新たに発見された 2 種類の生物の因果関係」を調べることです。生物の関係を調べるために、あなたは生物の様子を観察することにしました。

この図は赤い生物と青い生物の関係を表しています。赤い生物が青い生物を発生させているか、青い生物が赤い生物を発生させているか、あるいは、両者は無関係であるか分かりません。一

方の生物がもう一方の生物を発生させる場合でも必ず発生させるとは限りません。

あなたにはこれら 2 種類の生物を観察して頂き、『赤い生物と青い生物の因果関係』を判断して頂きたいと思います。なお、この課題では生物学に関する知識は不要です。(残りの教示は学習フェイズとテストフェイズの進め方に関するものであった)

上記の教示に加えて、完全教示群や時間順序教示群には生物の状態を観察する順序が当てにならないことが伝えられた。この教示の概要は次の通りであった。

生物が見つかる順序は必ずしも生物の因果関係を表しているわけではありません。つまり、赤い生物が先に観察されたからといって赤い生物が原因であるとは限りません。これは、生物は非常に小さいのでその存在の有無が判明するのに時間が掛かったりするためです。

時間順序に関する教示とは別に、完全情報群と他の原因教示群にはこれらの生物に影響を与えるような他の原因は存在しないことが伝えられた。なお、この情報は時間順序教示群や非教示群には伝えられなかった。この教示の概要は以下の通りであった。

また、これらの生物は厳重に管理された実験室で観察されているので、何らかの他の原因が生物に影響を与えることはありません。

実験参加者は教示の操作から生じる 4 種類の教示のいずれかを伝えられた。残りの教示は学習フェイズやテストフェイズの進め方に関するものであった。教示を読み終えた後、実験参加者が内容をきちんと理解しているかどうか確認した<sup>13</sup>。

**学習フェイズ** 学習フェイズは、2 種類の生物の有無を観察するのを 1 試行とし、全

---

<sup>13</sup>実験では赤い生物と青い生物と表現しているが、本文では説明を簡潔にするため生物 X、生物 Y という表記を用いる。

40 試行で構成されていた。試行はディスプレイ上に表示された「NEXT」というボタンをクリックすることで開始され、2 種類の生物に「？」と書かれた状態が表示され、1 秒ごとに生物の状態が表示された。生物が生起する場合には生物の絵が、非生起の場合には生物の絵に×が書かれたものが表示された。一致条件では時間順序と因果関係が一致するように生物の状態が表示されたのに対して (i.e., 原因の提示後、結果の提示)、逆方向条件では因果関係とは逆の順序で生物の状態が表示された (i.e., 結果の提示後、原因の提示)。2 種類の生物の状態が判明してから 2 秒経過後、最初の画面に移行した。実験参加者には前の試行の提示内容が現在の試行の提示内容に影響しないことが教示されていた。

実験参加者は 40 試行の観察を通して共変動手がかりを受け取っていた。共変動情報はすべての群に共通していた。40 試行のうち、20 試行では生物 X と生物 Y がともに存在していたが、10 試行ではどちらの生物も存在していなかった。残りの 10 試行では、生物 X のみが存在していて、生物 Y が存在していなかった。この実験課題では、因果の必要性の観点から単独生起によって原因が特定されるため、この情報によって生物 X が生物 Y に影響するという因果の方向性が決定された。原因事象の提示位置は実験参加者間でカウンターバランスをとった。また、課題に馴れさせるために、学習フェイズの前に練習試行を設け、実験参加者には練習試行の内容と学習フェイズの内容は無関係であることを伝えた。

**テストフェイズ** テストフェイズでは、生物 X が生物 Y を発生させているかという問いに対して 2 件法 (はい・いいえ) で回答するように指示した。生物 Y が生物 X を発生させているか否かについても同様に尋ねた。

## 6. 2. 2. 結果および考察

テストフェイズにおける 2 つの質問への回答を組み合わせると、(1)X が Y を発生させている、(2)Y が X を発生させている、(3)相互に因果関係が存在する、(4)無関係であると

いう4つのカテゴリーに回答を分類した。これら4つの因果モデルのうち、正答は生物 X が生物 Y を発生させる因果モデルであった。それぞれの群における回答の割合を Table 6-1 に示した。時間順序手がかりが因果関係の順序と一致している場合、各群の正答率はチャンスレベルの 25% を大きく上回っていた (83.3% in the full-instruction group,  $p < .001$ ; 75.0% in the unreliable-temporal-order group,  $p < .001$ ; 91.7% in the no-hidden-cause group,  $p < .001$ ; 83.3% in the non-instruction group,  $p < .001$  by binomial tests)。しかしながら、時間順序手がかりも同じ判断を導くため、これらの回答が共変動手がかりに基づく判断とは限らない。したがって、時間順序が因果の順序と逆向きになっている事態での回答が重要となる。二項検定の結果、追加の教示を受け取った場合には正しく判断できていることが明らかとなった (75.0% in the full-instruction group,  $p < .001$ ; 50.0% in the unreliable-temporal-order group,  $p = .054$ ; 66.7% in the no-hidden-cause group,  $p = .003$ )。一方、追加の教示がない場合には因果関係を見つけることができていなかった (25.0% in the non-instruction group,  $p = .609$ )。時間順序が因果の順序と逆向きになっている事態において多くの実験参加者が因果構造を正しく推測できたことは、共変動手がかりから因果の方向性を学習していることを示唆している。また、各群での異なる正答率は、時間順序手がかりが因果構造の判断に影響を及ぼすことだ

#### 実験 1 における各群の回答パターン

	Full-Instruction	Unreliable-Temporal-Order	No-Hidden-Cause	Non-Instruction
Consistent Order				
X→Y	83.3	75.0	91.7	83.3
X←Y	0.0	0.0	0.0	0.0
X→Y & X←Y	8.3	0.0	0.0	0.0
X Y	8.3	25.0	8.3	16.7
Reverse Order				
X→Y	75.0	50.0	66.7	25.0
X←Y	8.3	33.3	0.0	16.7
X→Y & X←Y	8.3	0.0	0.0	0.0
X Y	8.3	16.7	33.3	58.3

けでなく、時間順序や他の原因に関する事前知識が因果構造の学習を促進することを示唆している。

Table 6-2 にそれぞれの質問項目に対する実験参加者の回答を示した。時間順序手がかりや事前知識の効果を検討するため、時間順序(2)×時間順序の教示(2)×他の原因の教示(2)のロジスティック回帰分析を実施した。主効果や交互作用は Wald 検定を用いて査定した。生物 X が生物 Y を発生させるかという質問への回答では、時間順序の主効果 ( $\chi^2(1, N = 96) = 9.77, p = .002$ ) や他の原因の教示の主効果 ( $\chi^2(1, N = 96) = 7.58, p = .006$ ) が有意であった。このことは、他の原因に関する教示が共変動手がかりの使用を促進していることを示唆する。一方、生物 Y が生物 X を発生させているかという質問への回答では、時間順序の主効果が有意であり ( $\chi^2(1, N = 96) = 4.29, p = .038$ )、時間順序の教示の効果は有意でなかった ( $\chi^2(1, N = 96) = 3.27, p = .071$ )。因果の順序が事象の提示順序と矛盾する場合、時間順序手がかりは X が Y を発生させるという判断を減少させ、Y が X に影響するという判断を増加させることが示された。実験参加者は共変動手がかりに基づいて因果構造を正しく推測できていたが、これらの結果は実験参加者が時間順序手がかりや事前知識に大きく影響されることをはっきりと示している。

実験 1 の結果、共変動手がかりと時間順序手がかりが一致しないことを観察した場合に、共変動手がかりから因果構造を正しく判断できることが示された。何名かの実験参加者は共変動手がかりよりも時間順序手がかりに基づいた判断を行うなど不完全な部

Table 6-2

実験 1 のそれぞれの質問項目に対する各群の Yes 反応率

	Full-Instruction	Unreliable-Temporal-Order	No-Hidden-Cause	Non-Instruction
Question about X→Y				
Consistent Order	91.7	75.0	91.7	83.3
Reverse Order	83.3	50.0	66.7	25.0
Question about X←Y				
Consistent Order	8.3	0.0	0.0	0.0
Reverse Order	16.7	33.3	0.0	16.7

分も見られたが、共変動情報から因果の方向性を導く認知能力を備えていることは明確に反映されていた。しかしながら、これらの能力は時間順序が当てにならないこと、あるいは他の原因が存在しないことが予め分かっている事態に限定されていた。具体的には、他の原因が存在しないことは共変動手がかりに基づく判断を促進していた。これらの結果は共変動情報を利用する際の事前知識の重要性を示唆するものである。

実験 1 の結果に関する重要な問いは、いつ実験参加者が共変動情報から因果構造を導き出したのかということである。共変動手がかりと時間順序手がかりはその入手可能性において大きく異なる。時間順序手がかりが常にある因果構造を示唆するのに対して、共変動手がかりは繰り返しの観察を必要とする。そのため、学習フェイズの初期には時間順序手がかりを好み、学習フェイズの最後には共変動手がかりを重視したかもしれない。実験 2 では、このような切り替えが学習過程で生じているか検討する。

### 6. 3. 実験 2

実験 1 では、実験参加者は学習フェイズにおいて 2 つの事象の共変動情報を取得し、テストフェイズにおいて因果構造を判断することが求められた。実験の結果、時間順序手がかりよりも共変動手がかりを重視して因果構造を推測していることが示された。実験 1 の手続きの限界として、共変動や時間順序の情報が増えていく際の学習過程を見落としていることが挙げられる。共変動情報と時間順序情報では利用し易さが大きく異なる。時間順序手がかりは常に使用可能なのに対して、共変動手がかりはいくつかの観察後にはじめて利用できるようになる。したがって、初期の判断では時間順序手がかりの影響が強く、最終的な判断では共変動手がかりの影響が強くなることが予想される。実験 2 の目的は学習フェイズの初期にどのような因果モデルが形成され、試行の経過に伴ってどのように修正されていくか検討することである。実験 2 では学習フェイズの途中でテストフェイズを挿入することによって因果構造の学習過程を検討した。実験 1 では時間順序手がかりの信頼性を教示によって操作していたが、実験 2 ではランダムな時間順序を提

示する群を設けることによって時間順序手がかりの信頼性を操作した。これらの実験操作によって時間順序手がかりと共変動手がかりの相互作用を検討することが可能になる。

### 6. 3. 1. 方法

**実験参加者および実験計画** 大学生 48 名が実験に参加し、成績への加点となるコースクレジットを受け取った。この中に実験 1 の参加者は含まれていなかった。実験参加者は時間順序の操作（一致、逆方向、ランダム）によって生じる 3 群のいずれかに無作為に割り当てられた。

**実験手続き** 基本的な手続きは実験 1 と同様であった。教示では、科学者になったつもりで微生物の因果関係を推測するように伝えた。また、因果関係が確率的であることや他の原因は存在しないことが教示された。これらの教示は共変動情報から因果の方向性を引き出すのに必要なものであった。実験 1 とは異なり、時間順序手がかりが当てにならないことは知らせなかった。その代わりに学習フェイズ中の時間順序を操作した。

学習フェイズでは、実験参加者は 2 種類の微生物の状態に関する情報を受け取った。生物の生起・非生起を観察する順序は各群で異なっていた。一致群の実験参加者は、原因事象の提示後に結果事象が提示され、因果の方向性と同じ順序で共変動情報を獲得した。一方、逆方向群では結果事象の提示後に原因事象が提示され、実際の因果関係とは逆向きの順序で観察することが求められた。これらの 2 群に加えて、実験 2 ではランダム群が追加された。ランダム群では半分の試行は因果の方向と一致する順序で、残りの半分の試行は因果の方向とは逆の順序で事象の状態が提示された。つまり、ランダム群では時間順序手がかりが実際に役に立たないわけである。共変動情報は実験 1 と同様であり、X と Y の共生起 20 試行、X と Y の非生起 10 試行、X の単独生起 10 試行となっていた。因果の必要性によって、共変動情報から生物 X が生物 Y を発生させるという因果関係が示唆されていた。

テストフェイズは学習フェイズに繰り返し挿入された。因果構造の学習過程を検討するために、学習フェイズの4試行目・16試行目・28試行目・40試行目が終了した段階で因果構造を判断させた。テストフェイズが等間隔に挿入されていないのは、因果の方向を引き出すために最低限必要な観察をした後の初期の因果モデルを調べるためであった。実験1と同様に、生物Xが生物Yを発生させているかという問いや、生物Yが生物Xを発生させているかという問いに対して2件法（はい・いいえ）で回答するように指示した。構造の判断に加えて、判断に対する確信度を0（まったく自信がない）から100（非常に自信がある）の範囲で評定するように伝えた。

### 6. 3. 2. 結果および考察

実験参加者の回答は実験1と同様の方法で4つのカテゴリーに分類した。Figure 6-2に各テストフェイズにおける回答の割合を示した。いずれの群においても大半の実験参加者が最初のテストフェイズで正しい因果構造を回答しており（consistent group,  $p < .001$ ; reverse group,  $p = .027$ ; random group,  $p < .001$ , by binomial tests), この傾向は最後のテストフェイズでも見られた（consistent group,  $p = .007$ ; reverse group,  $p = .002$ ; random group,  $p = .027$ , by binomial tests). これらの結果は共変動手がかりから因果構造を判断するという実験1の結果を再現した物であり、単純な因果構造を学習するには時間順序手がかりと同じくらい早く共変動手がかりを使用

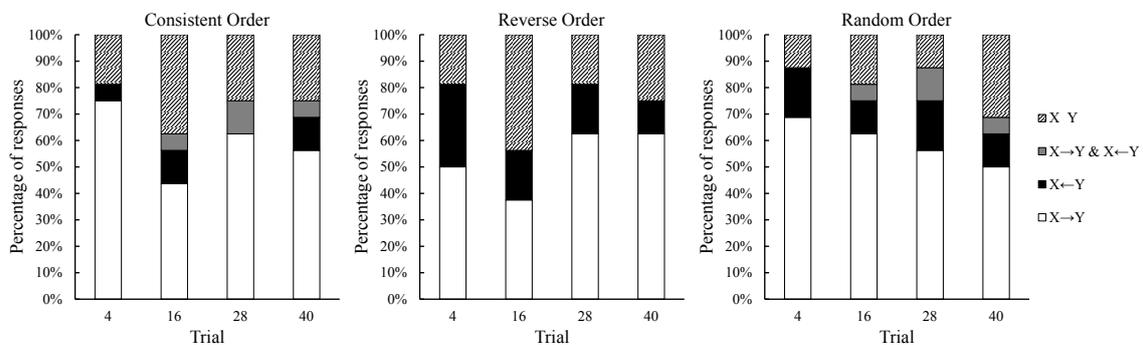


Figure 6-2. 実験2のテストフェイズにおける各群の回答の割合。

することが示唆された。テストフェイズを通して正答率はおおまかに一定であったが、試行の経過に伴って共変動手がかりの使用は若干減少していた。2つの事象間の因果関係がそれほど強くないために (i.e.,  $P(\text{effect} \mid \text{cause}) = .67$ ,  $P(\text{effect} \mid \neg\text{cause}) = 0$ ), 何名かの実験参加者は相関の検出に失敗して無関係であると考えたかもしれない。予測とは異なり、共変動手がかりの使用が徐々に増え、時間順序手がかりの使用が徐々に減るといった傾向は見られなかった。この結果は、質問項目が2件法での回答を求めており、項目の感受性が低かったことに起因すると考えられる。実験参加者の判断を詳細に検討するため、各群の確信度に関する分析を行った。

Figure 6-3に各群の平均確信度得点を示した。因果関係があるという回答に+1を、因果関係がないという判断に-1を割り当て、確信度の評定値をかけ、異なる回答をまとめて加算した。これらの確信度得点は実験参加者がどの程度その因果関係の存在を信じているかを示していると考えられる。確信度得点は正規分布しないため、ノンパラメトリック検定を使用した。生物 X が生物 Y を発生させるという判断では、一致群とランダム群では試行の経過に伴う変化が見られなかったのに対して (consistent group,  $p = .879$ ; random group,  $p = .584$ , by Friedman tests), 逆方向群では徐々に増加

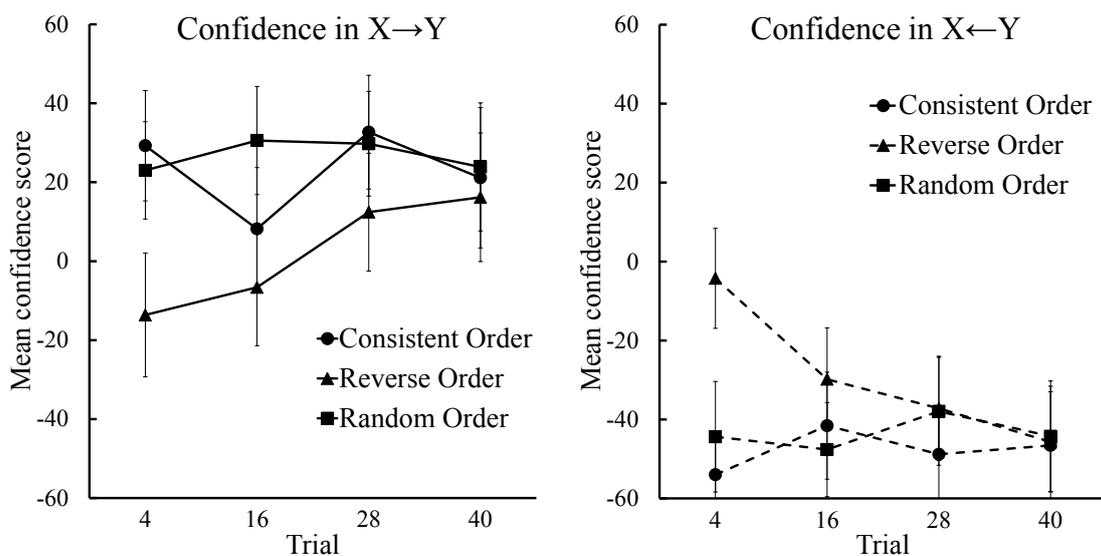


Figure 6-3. 実験 2 における各群の確信度得点。

する傾向が示された ( $p = .026$  by Friedman test). 多重比較の結果, 1 回目のテストフェイズと 4 回目のテストフェイズ間で有意差が確認された ( $p = .030$ ). 逆方向群における確信度スコアの漸次的な変化は試行の経過に伴って共変動手がかりを重視するようになっていていることを示唆する. 同様の分析を生物 Y が生物 X を発生させるという判断に対しても行った. その結果, 逆方向群では試行の経過に伴って確信度スコアが低下していたのに対して ( $p = .021$  by Friedman test), 一致群とランダム群ではそのような変化は生じていなかった (consistent group,  $p = .804$ ; random group,  $p = .750$ , by Friedman tests). 逆方向群の結果について多重比較を行ったところ, 1 回目のテストフェイズと 4 回目のテストフェイズ間で有意差が見られた ( $p = .019$ ). 確信度スコアの漸次的な減少は試行の経過に伴って時間順序手がかりを使用しなくなったことを反映していると考えられる.

実験 2 の結果, 多くの実験参加者が共変動手がかりや他の原因が無いという情報に基づいて因果構造を正しく判断できるという実験 1 の結果が再現された. さらに, これらの判断は最小限の観察でも形成されることが示された. また, 逆方向群の確信度スコアからは学習初期では時間順序手がかりを優先するが, 繰り返しの観察によって共変動手がかりが重視されるようになることが示唆された.

#### 6. 4. 実験 3

実験 1 や実験 2 の結果から共変動手がかりに基づいて因果構造を判断できることが示された. しかしながら, 共変動情報から因果の方向性をどのようにして引き出しているかは依然として不明なままである. これまでの実験では因果の必要性を所与のものとして, 必要性で定義された因果関係を見出せるかということに焦点が当てられていた. しかし, 因果性には十分性や必要十分性など異なる側面もあり, これらによって共変動情報の解釈は大きく異なる. 例えば, 事象 X と事象 Y が頻繁に共生起しており, 時々事象 X のみが単独で生起していたとする. 原因事象が生起していない時に結果事象は生起し

ないとする因果の必要性 (i.e.,  $P(\text{effect}|\neg\text{cause}) = 0$ ) の観点からは事象 X が事象 Y を発生させるという因果関係 (i.e.,  $X \rightarrow Y$ ) が示唆される. 一方, 原因の生起は必ず結果の生起を伴うとする因果の十分性 (i.e.,  $P(\text{effect}|\text{cause}) = 1$ ) という観点では事象 Y が事象 X を発生させるという因果関係 (i.e.,  $X \leftarrow Y$ ) として解釈できる.

Pearl (2000) は因果の必要性・十分性・必要十分性を評価する指標として, 必要性の確率 ( $PN$ , probability of necessity), 十分性の確率 ( $PS$ , probability of sufficiency), 必要十分性の確率 ( $PNS$ , probability necessity and sufficiency) を提唱している. これらの指標は, 共変動情報に原因が生起していて結果が生起していない事例 (i.e., セル b) と原因が生起していなくて結果が生起している事例 (i.e., セル c) が同時に含まれていない場合には容易に計算することができる. 必要性の確率  $PN$  は次の式によって求められる.

$$PN = \frac{P(y|x) - P(y|\neg x)}{P(y|x)} \quad (6-1)$$

事象 X が事象 Y を発生させる事態 (i.e.,  $\Delta P > 0$ ) において,  $PN$  の値は事象 X が無い時に事象 Y が生起する確率  $P(y|\neg x)$  が減少するにつれて大きくなる. このことは, 原因が無い時に結果が発生することはないという因果の必要性の側面を反映している. 一方, 因果の十分性は, 原因の生起が結果の生起を必ず伴うことを想定しており, 十分性の確率  $PS$  は以下の式によって算出される.

$$PS = \frac{P(y|x) - P(y|\neg x)}{1 - P(y|\neg x)} \quad (6-2)$$

十分性の確率はパワーPC 理論 (Cheng, 1997) の因果パワーと一致する. また, 必

要十分性の確率  $PNS$  は、因果関係における必要性の側面と十分性の側面の両方を考慮に入れた指標であり、 $\Delta P$  と一致することが知られている。

$$PNS = P(y|x) - P(y|\neg x) \quad (6-3)$$

これらの指標は事象  $X$  から事象  $Y$  という方向性に基づいて算出されており、 $2 \times 2$  の分割表の行と列を入れ替えることによって事象  $Y$  から事象  $X$  という方向性に基づいた指標を算出することもできる。

実験 3 の目的は因果構造の判断を行う際に共変動情報がどのように解釈されているか検討することである。2 事象間の共変動情報を体系的に操作し、因果構造の判断を求めた。因果の必要性に関する教示や他の原因が存在しないことを伝えないため、因果の方向性を一意に決定することは不可能であるが、この実験事態では Pearl (2000) による 3 つの指標を用いて実験参加者の解釈を分類することが可能になる。

#### 6. 4. 1. 方法

**実験参加者および実験計画** 大学生 43 名が実験に参加した。この中に実験 1 や実験 2 の参加者は含まれていなかった。共変動情報を体系的に操作し、15 条件を設定した (see Table 6-3)。具体的には、事象  $X$  が生起している時に事象  $Y$  が生起している確率  $[P(y|x) = .00, .25, .50, .75, 1]$  と事象  $X$  が生起していない時に事象  $Y$  が生起している確率  $[P(y|\neg x) = .00, .25, .50, .75, 1]$  を  $\Delta P \geq 0$  となるように組み合わせた。実験は実験参加者内計画で行われ、実験参加者は共変動情報の異なる複数の課題を遂行することが求められた。

**実験手続き** 教示では、微生物の研究者になったつもりで新たに発見された 2 種類の微生物 ( $X, Y$ ) の状態を観察して、生物の間に因果関係があるかどうか、因果関係がある場合にはどちらが原因であるかを判断するように伝えた。本課題における因果関

係とは一方の生物が他方の生物を発生させる関係であると説明した。学習フェイズは、1つの採集地における生物の有無を観察するのを1試行として全16試行で構成されていた。ディスプレイ上に表示された「NEXT」というボタンをクリックした後、2種類の微生物の状態が同時に表示された。生物が生起している場合には生物の絵が、非生起の場合には生物の絵に×が書かれたものが表示された。生物の状態が判明してから1秒経過後、最初の画面に移行した。テストフェイズでは、生物Xが生物Yを発生させているかという問いに対して2件法（はい・いいえ）で回答するように指示した。生物Yが生物Xを発生させているか否かについても同様に尋ねた。それぞれの条件は無関係であることを教示した上で次の条件に移行し、同様の手続きで共変動情報の異なる15条件における因果構造の判断を記録した。

#### 6. 4. 2. 結果および考察

テストフェイズにおける2つの質問への回答を組み合わせることにより、(1)XがYに影響している、(2)YがXに影響している、(3)相互の因果関係が存在する、(4)無関係である

Table 6-3

実験3の各条件における回答の割合と予測

Conditions			Causal models				Predictions		
$P(Y X)$	$P(Y \neg X)$	$\Delta P$	$X \rightarrow Y$	$X \leftarrow Y$	$X \rightarrow Y \& X \leftarrow Y$	$X \perp Y$	$PN$	$PS$	$PMS$
1.00	.00	1.00	0.00	2.33	53.49	44.19			
1.00	.25	.75	20.93	48.84	6.98	23.26	$X \leftarrow Y$	$X \rightarrow Y$	$X \leftarrow Y$
.75	.00	.75	53.49	18.60	6.98	20.93	$X \rightarrow Y$	$X \leftarrow Y$	$X \leftarrow Y$
1.00	.50	.50	27.91	53.49	0.00	18.60	$X \leftarrow Y$	$X \rightarrow Y$	$X \leftarrow Y$
.75	.25	.50	11.63	13.95	18.60	55.81			
.50	.00	.50	51.16	23.26	2.33	23.26	$X \rightarrow Y$	$X \leftarrow Y$	$X \leftarrow Y$
1.00	.75	.25	27.91	60.47	2.33	9.30	$X \leftarrow Y$	$X \rightarrow Y$	$X \leftarrow Y$
.75	.50	.25	9.30	20.93	11.63	58.14			
.50	.25	.25	13.95	13.95	16.28	55.81			
.25	.00	.25	46.51	23.26	2.33	27.91	$X \rightarrow Y$	$X \leftarrow Y$	$X \leftarrow Y$
1.00	1.00	.00	20.93	51.16	0.00	27.91			
.75	.75	.00	11.63	27.91	13.95	46.51			
.50	.50	.00	11.63	4.65	11.63	72.09			
.25	.25	.00	16.28	13.95	6.98	62.79			
.00	.00	.00	23.26	16.28	0.00	60.47			

るといふ 4 つのカテゴリーに実験参加者の回答を分類した。Table 6-3 に各条件における回答の割合と 3 種類の指標に基づく予測を示す。実験の結果、いずれの条件においても因果関係を特定することはできないにも関わらず、条件間で回答パターンは大きく異なっていることが明らかとなった ( $\chi^2(42) = 291.37, p < .001$ )。実験参加者の回答は因果の必要性に基づく予測と概ね対応しており、十分性や必要十分性に基づく予測とは異なっていた。また、これらの指標が定義されない条件では、事象 X と事象 Y が無関係であるという判断が多く見られた。

実験参加者がどのような観点から共変動情報を解釈していたか検討するために、各指標が定義可能な 6 つの条件における予測と回答の対応に基づいて、実験参加者を (1)必要性、(2)十分性、(3)必要十分性、(4)分類不可、のいずれかのグループに分類した。予測と一致した回答に 1、そうでない回答に 0 と得点を付け、実験参加者を最もスコアの高くなるグループに割り当てた。(4)分類不可にはチャンスレベルと同程度の一致を示した実験参加者と複数のグループに同程度の一致を示した実験参加者が含まれた。その結果、過半数の実験参加者 (55.81%) が必要性のグループに分類され、少数の参加者 (27.91%) が十分性のグループに分類され、他のグループにはほとんど分類されていないことが示された (必要十分性 6.98%; 分類不可 9.30%)。

実験 3 の結果、共変動情報が必要性の観点から解釈されていることが明らかとなった。しかしながら、この知見は Mayrhofer and Waldmann (2011) の知見とは一致していない。Mayrhofer and Waldmann (2011) は、宇宙人のコミュニケーションを観察して思考伝播の方向性を判断するという実験事態で、共変動情報が因果の十分性の観点から解釈されることを報告している。これら 2 つの研究の主な差異として、実験に用いたカバーストーリーや共変動情報の設定方法などが挙げられる。実験 4 では、実験 3 からカバーストーリーのみを変更して同様の実験を行い、両者の差異がカバーストーリーによるものか、共変動情報の設定方法によるものか検討した。

## 6. 5. 実験 4

実験 3 では、二種類の微生物の状態を観察した後に因果構造の判断を求めたところ、因果の必要に基づいて共変動情報が解釈されていることが示された。実験 4 では、実験に用いるカバーストーリーを Mayrhofer and Waldmann (2011) と同一のものに変更し、Mayrhofer and Waldmann (2011) と同様に因果の十分性に基づく判断が見られるか検討した。

### 6. 5. 1. 方法

**実験参加者および実験計画** 大学生 28 名が実験に参加した。この中にこれまでの実験の参加者は含まれていなかった。実験 3 と同様に共変動情報を体系的に操作し、15 条件を設定した (see Table 6-4)。

**実験手続き** 実験手続きは使用するカバーストーリーを除いて実験 3 と同様であった。教示では、宇宙人の中には他者の思考を読み取る能力によって他の宇宙人と同じことを考えることが可能な宇宙人が存在しており、科学者になったつもりで新たに発見された 2 種類の宇宙人 (X, Y) の因果関係を判断するように伝えた。本課題における因果関係とは一方の宇宙人の思考が他方の宇宙人に伝染する関係であると説明した。学習フェイズは、2 種類の宇宙人が考えている内容を観察するのを 1 試行として全 16 試行で構成されていた。思考内容はアルファベットの母音と子音で構成された無意味綴りであり、宇宙人が何も考えていない場合は何も表示されなかった。テストフェイズでは、宇宙人 X の思考が宇宙人 Y に伝染しているか否かについて 2 件法 (はい・いいえ) で回答させた。宇宙人 Y の思考が宇宙人 X に伝染しているか否かについても同様に尋ねた。

### 6. 5. 2. 結果および考察

実験 3 と同様の方法で分析を行った。Table 6-4 に各条件における回答の割合と 3 種類の指標に基づく予測を示す。実験の結果、いずれの条件においても因果関係を特

定することはできないにも関わらず、条件間で回答パターンは大きく異なっていることが明らかとなった ( $\chi^2(27) = 315.35, p < .001$ ). 回答パターンに基づいて実験参加者を分類したところ、過半数の実験参加者 (60.71%) が十分性のグループに分類され、少数の参加者 (17.86%) が必要性グループに分類され、他のグループにはほとんど分類されていないことが示された (必要十分性 7.14%; 分類不可 14.3%). これらの結果は共変動情報が因果の十分性に基づいて解釈されていることを示している.

実験 4 の結果, Mayrhofer & Waldmann (2011) と同様に, 共変動情報が因果の十分性の観点から解釈されていることが示された. 実験 3 との違いが実験に使用するカバーストーリーのみであったことから, カバーストーリーによって解釈が変化することが明らかとなった. カバーストーリー内のどのような要因が影響しているか本実験から断定することはできないが, 他の原因事象の有無が可能性として挙げられる. 微生物の発生に関するカバーストーリーでは, 片方の生物がもう一方の生物を産出する場合, 原因となる生物がいない状態で結果となる生物が発生するとは考えにくい. 一方, 宇宙人の思考伝

Table 6-4

実験 4 の各条件における回答の割合と予測

Conditions			Causal models				Predictions		
$P(Y X)$	$P(Y \neg X)$	$\Delta P$	$X \rightarrow Y$	$X \leftarrow Y$	$X \rightarrow Y \& X \leftarrow Y$	$X \quad Y$	$PN$	$PS$	$PNS$
1.00	.00	1.00	3.57	3.57	82.14	10.71			
1.00	.25	.75	57.14	14.29	14.29	14.29	$X \leftarrow Y$	$X \rightarrow Y$	$X \leftarrow Y$
.75	.00	.75	25.00	57.14	7.14	10.71	$X \rightarrow Y$	$X \leftarrow Y$	$X \leftarrow Y$
1.00	.50	.50	60.71	25.00	0.00	14.29	$X \leftarrow Y$	$X \rightarrow Y$	$X \leftarrow Y$
.75	.25	.50	25.00	14.29	10.71	50.00			
.50	.00	.50	21.43	46.43	0.00	32.14	$X \rightarrow Y$	$X \leftarrow Y$	$X \leftarrow Y$
1.00	.75	.25	57.14	21.43	7.14	14.29	$X \leftarrow Y$	$X \rightarrow Y$	$X \leftarrow Y$
.75	.50	.25	32.14	17.86	7.14	42.86			
.50	.25	.25	7.14	28.57	3.57	60.71			
.25	.00	.25	14.29	64.29	3.57	17.86	$X \rightarrow Y$	$X \leftarrow Y$	$X \leftarrow Y$
1.00	1.00	.00	60.71	21.43	3.57	14.29			
.75	.75	.00	39.29	7.14	7.14	46.43			
.50	.50	.00	7.14	14.29	10.71	67.86			
.25	.25	.00	10.71	21.43	3.57	64.29			
.00	.00	.00	7.14	28.57	0.00	64.29			

播に関するカバーストーリー、一方の宇宙人からの影響を受けない場合でも、自発的に考えた結果として思考内容が生じると考えられる。実際、他の原因に関する教示を操作した実験 1 では、他の原因が無いことを教示された場合には必要性に基づく共変動手がかりの使用が増加していた。それぞれのカバーストーリーで事前の知識を操作することによって、この可能性は検討できる。現在進行中の研究ではこの仮説を支持する結果が得られている。

## 6. 6. 総合論議

本研究では 2 事象で構成される因果構造を判断する際に、共変動手がかりから因果の方向性が引き出されるか否か検討を行った。実験 1 の結果、単純な因果構造を学習する場合には時間順序手がかりよりも共変動手がかりが重視されることが明らかとなった。学習過程の検討を行った実験 2 の結果からは、共変動手がかりに基づく判断が非常に早い段階からなされていることが明らかとなった。また、共変動情報からどのように因果の方向性が引き出されるか検討したところ、微生物の発生関係を題材とした課題では必要性に基づく解釈が（実験 3）、宇宙人の思考伝播の関係を題材とした課題では充分性に基づく解釈が見られた（実験 4）。

本研究の結果は、共変動手がかりから因果の方向性は導き出されないとする先行研究（Lagnado & Sloman, 2006; White, 2006）の結果とは異なるものであった。これらの相反する知見は課題の複雑性の観点から解釈できる。単純な因果構造の場合には共変動情報の種類も少なく、手がかりとして利用し易いが、複雑な因果構造の場合には共変動情報の種類が多く、そこから因果の方向性を判断することは困難になる。課題の複雑性による判断方略の変化については研究 2 で検討を行った。

共変動手がかりの使用に関する本研究の知見は、共変動情報から因果構造を推測する因果ベイズネットの説明と整合的である。しかし、宇宙人のコミュニケーションに関するカバーストーリーで充分性に基づく解釈が見られ、微生物の発生に関するカバースト

一リーで必要性に基づく解釈が見られたことは、同じ情報から全く異なる結論が導かれることを意味しており、直ちに説明することはできない。この結果は領域一般的な共変動情報と領域固有的な知識の相互作用として考えることが可能であり、このような相互作用を説明するには階層ベイズモデルが必要になる。階層ベイズモデルによって本研究の結果が説明できるかどうかは今後の研究課題である。

## 7. 研究 2 因果構造の学習における課題の複雑性と判断方略の変化

### 7. 1. 序論

研究 1 の結果, 単純な因果構造を判断する課題では時間順序手がかりに惑わされず, 共変動手がかりから因果の方向性を正しく判断できることが示された. しかしながら, この結果は共変動手がかりよりも時間順序手がかりが優先されるという先行研究 (Lagnado & Sloman, 2006; White, 2006) の結果とは異なっていた. これらの相反する知見は課題の複雑性の観点から解釈することができる. 因果構造を構成する事象の数やそれらを結びつけるエッジの数の増加に伴って共変動情報は複雑になり, そこから因果関係を見出すことは困難になる. Lagnado and Sloman (2006) では 4 事象の因果構造が, White (2006) では 5 事象の因果構造が使用されていたのに対して, 研究 1 では 2 事象の因果構造が用いられており, 事象の数によって判断方略が変化したと考えられる.

いくつかの研究によって因果学習における課題の複雑性の役割が明らかにされてきた (e.g., Marsh and Ahn, 2006; Reips & Waldmann, 2008). Marsh and Ahn (2006) は, 因果関係の強さを判断する実験課題において課題の複雑性が初頭効果と新近効果の決定因として機能することを実証している. 観察する事象の数が少ない場合, 学習フェイズの前半に与えられた情報は後半に与えられた情報よりも重視されていた (i.e., 初頭効果). 一方, 観察する事象の数が多い場合には学習フェイズの後半に提示された情報に重点を置いた判断が見られた (i.e., 新近効果). また, Reips and Waldmann (2008) は診断的推論課題における事象の数の効果を検討している. 実験課題で提示する事象の数が少ない場合は基準率を考慮した正しい判断が見られたのに対して, 事象の数が増えて課題が複雑になった場合には基準率が考慮されなくなることを報告している. これらの研究は因果性の学習と推論において課題の複雑性が大きな影響力を持つこ

とを示唆している。

本研究の目的は因果構造の学習における課題の複雑性と判断方略の関係を検討することである。課題の複雑性の操作として、因果構造を構成する事象の数を操作した（実験 1）。事象の数を操作は因果構造の形態の変化を伴うため、判断の変化が因果構造を構成する事象の数の変化だけによってもたらされたものか、因果構造の形態の影響も含んでいるか区別することはできない。実験 2 では、実験 1 と異なる因果構造を使用し、実験間の比較によって因果構造の形態の影響を調べた。また、実験 3 では介入による学習の際に課題の複雑性によって判断方略が変化するか検討した。提示される事象の数が少ない単純な課題では共変動手がかりに基づく判断が、事象の数が多し複雑な課題では時間順序手がかりに基づく判断が多くなると仮説を立てた。

## 7. 2. 実験 1

実験 1 の目的は課題の複雑性の変化に伴って判断方略がどのように変化するか検討することであった。実験課題は、複数の事象の状態を観察し、それらの因果構造を推測することであった。課題の複雑性の操作として、因果構造を構成する事象の数を操作した。事象の数が少ない場合は共変動手がかりに基づく判断が、事象の数が多し場合には時間順序手がかりに基づく判断が多くなると予測した。

### 7. 2. 1. 方法

**実験参加者および実験計画** 大学生 24 名が実験に参加し、成績への加点となるコースクレジットを受け取った。実験は実験参加者内計画で行われ、実験参加者は出現する事象の数が異なる 3 つの課題（i.e., 3 事象, 4 事象, 5 事象）を遂行した。

**教示** 実験参加者に印刷された教示を手渡し、実験者がその内容を読み上げた。教示の概要は次の通りであった。

以下に記されているストーリーの登場人物になることを想像してください。あなたはとある研究所で微生物の研究を行っています。現在の研究テーマは「新たに発見された数種類の微生物の因果関係」を明らかにすることです。下記の図は赤い生物と青い生物と黄色い生物の関係を表しています。どの生物がどの生物を発生させているか、あるいは、生物同士が無関係であるか分かりません。

微生物の関係を調べるために、一種類の微生物だけを培養液で満たされた容器 40 個に入れて、他の微生物が発生するか否かを顕微鏡で観察することにしました。どの微生物が入れられるかはそれぞれの課題の前にお知らせします。どの微生物が容器に入れられるかということに加えて、以下の情報はあなたが生物の関係を考える際に役に立ちます。(1)ある生物が他の生物を発生させる場合でも必ず発生させるとは限りません。(2)これらの生物は厳重に管理された実験室で観察されているため、生物以外の他の原因によって生物が発生することはありません。つまり、最初に入れられた生物以外の生物が発生した場合、その生物を発生させる生物が存在するという事です。(3)生物が発生しているか否かは順番に判明します。これは、ある生物が他の生物を発生させるのに時間が掛かったり、あるいは生物は非常に小さいのでその存在の有無が判明するのに時間が掛かったりするためです。生物が存在すると判明する前にその生物が他の生物を発生させることもあり得ます。

あなたにはこれらの生物の有無を観察していただき、『どの生物がどの生物を発生させているか』判断して頂きたいと思います。なお、この課題では生物学に関する知識は不要です。(残りの教示は学習フェイズとテストフェイズの進め方に関するものであった)

教示を読み終えた後、実験参加者が内容をきちんと理解しているかどうか確認した。このカバーストーリーにおいて、事象の数 (3 事象 vs. 4 事象 vs. 5 事象)

は実験参加者が課題中に観察する微生物の種類と対応していた。また、教示内の(1)は因果関係が確率的であることを、(2)は他に隠れた原因は存在しないことを、(3)は時間順序手がかりが当てにならないことをそれぞれ意味しており、これらは先行研究の Lagnado and Sloman (2006) でも伝えられたものであった。

**学習フェイズ** 学習フェイズの開始時に、実験参加者には何種類の微生物が現れるか (i.e., 3 事象, 4 事象, 5 事象), 培養液で満たされた容器に初めから投入されている微生物はどれか (i.e., 最初の原因) についての情報が与えられた。1つの容器における微生物の有無を観察するのを1試行として全40試行で構成されていた。実験参加者には生物の状態を観察し、それらの因果関係を考えることが求められた。試行はディスプレイ上に表示された「NEXT」というボタンをクリックすることで開始され、複数の生物に「?」と書かれた状態が表示され、1秒ごとに微生物の状態が表示された。生物が生起する場合には微生物の絵が、非生起の際には微生物の絵に×が書かれたものが表示された。全ての生物の状態が判明してから1秒経過後、最初の画面に移行した。それぞれの課題では出現する微生物の種類が異なっており、3事象条件では3種類の生物が、4事象条件では4種類の生物が、5事象条件では5種類の生物が提示された。これらの生物は同じ形をしていたが、それぞれの色は異なっていた。

Figure 7-1 に事象間の因果関係と時間関係を示した。共変動情報はこれらの因果構造と因果強度を基に作成された。教示で因果関係が確率的であると述べたように、原因が結果を生起する確率は80%に設定されていた。例えば、3事象条件では、40試行のうち26試行で生物 A・B・C が存在していた。また、生物 A・B が存在していて生物 C が存在していないのが6試行、生物 A・C が存在していて生物 B が存在していないのが6試行含まれていた。残りの2試行は生物 A のみ存在し、生物 B・C は存在していなかった。生物 A (i.e., 最初の

原因) が常に存在しているのに対して、それ以外の生物はその直接の原因が存在しない時に発生することはなかった。事象の数の増加に伴って、共変動情報はより複雑になっていた。Figure 7-1 で明らかなように時間関係と実際の因果関係は一致していなかった。生物の状態を観察する時間順序と実際の因果関係を異なったものにするによって、共変動手がかりと時間順序手がかりのどちらを重視しているか査定することが可能になった。

**テストフェイズ** 40 試行の観察の後、実験参加者には因果構造を推測することが求められた。学習フェイズと同じ配置で複数の微生物が描かれたシートを渡し、因果関係があると思う場合は原因から結果に対して矢印を描くように指示した。生物の間には薄いグレーの線が引かれており、それぞれの線に対して因果関係があると思うか、ある場合にはどちらが原因でどちらが結果であるか考えるように教示した。3 事象条件では A-B 間・A-C 間・B-C 間という 3 種類の関係について、4 事象条件では 6 種類の関係について、5 事象条件では 10 種類の関係について回答を求めた。なお、それぞれの実験課題は独立であること

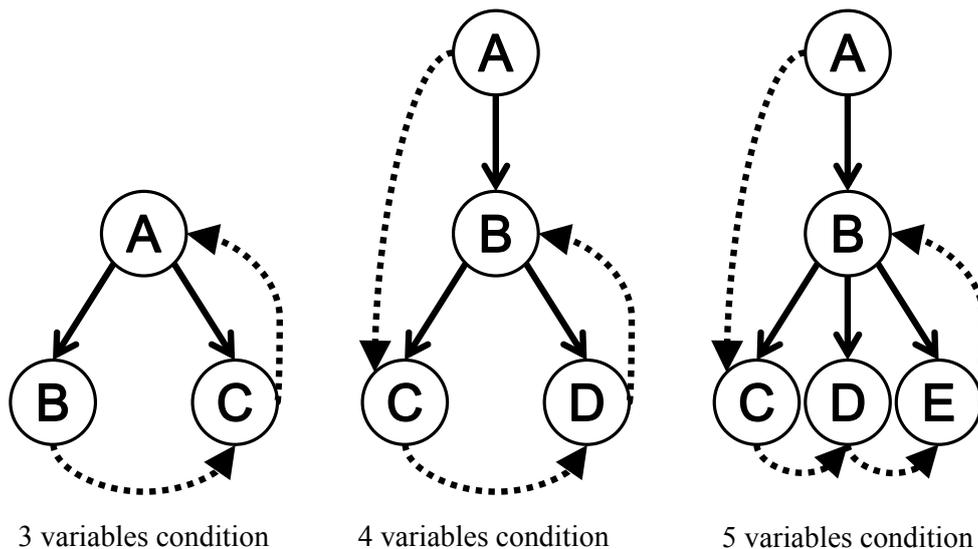


Figure 7-1. 実験 1 における因果構造。実線は実際の因果関係を、点線は実験参加者が観察する時間順序をそれぞれ表す。

を教示し、各条件の順序ならびに原因事象の配置については実験参加者間でカウンターバランスをとった。

## 7. 2. 2. 結果および考察

因果構造の学習における判断方略を同定するために、実験参加者がどの程度共変動手がかりと時間順序手がかりを使用していたかを表す手がかり使用率を定義した。共変動手がかりの使用率は、実験参加者が共変動手がかりと一致する形で描いたエッジの数を、共変動情報によって示唆されるエッジの総数で割ることで算出された。例えば、3事象条件ではAからBへのエッジ (i.e., A→B) とAからCへのエッジ (i.e., A→C) が共変動手がかりによって支持される。実験参加者がこれらのエッジを描いた場合、手がかり使用率は100%となる。一方、これらのエッジを回答しなかった場合、手がかり使用率は0%となる。同様の手順で時間順序手がかりの使用率も計算した。3事象条件では、BからCへのエッジ (i.e., B→C) とCからAへのエッジ (i.e., C→A) が時間順序手がかりによって支持されていた。実験参加者がこれらのどちらかを回答していた場合、手がかり使用率は50%となる。これらの指標は、実験参加者がどの程度その手がかりを使用していたかを表している。

Figure 7-2に各条件における共変動手がかりと時間順序手がかりの使用率を示す。3事象条件や4事象条件では共変動手がかりがよく使用されているのに対して、5事象条件では手がかり使用率にほとんど差は見られなかった。手がかり（共変動 vs. 時間順序）×事象の数（3事象 vs. 4事象 vs. 5事象）の二要因分散分析の結果、手がかりの主効果 ( $F(1, 23) = 14.01, MSE = 0.23, p = .001, \eta^2_G = .17$ ) および事象の数の主効果 ( $F(2, 46) = 4.43, MSE = 0.04, p = .017, \eta^2_G = .02$ ) ならびに手がかりと事象の数の交互作用 ( $F(2, 46) = 6.76, MSE = 0.12, p = .003, \eta^2_G = .09$ ) が有意であった。単純主効果の検定を行ったところ、

3 事象条件と 4 事象条件では共変動手がかりが優先される傾向が見られたのに対して ( $F(1, 23) = 16.21$ ,  $MSE = 0.20$ ,  $p < .001$ ,  $\eta_G^2 = .33$ ,  $F(1, 23) = 12.50$ ,  $MSE = 0.13$ ,  $p = .002$ ,  $\eta_G^2 = .23$ ), 5 事象条件では手がかり使用率に差は見られなかった ( $F < 1$ ). これらの結果は課題の複雑性によって判断方略が変化することを示唆している.

事象の数の変化に伴って判断方略も変化していたが, 4 事象条件の結果は時間順序手がかりが共変動手がかりよりも優先されるという先行研究 (Lagnado & Sloman, 2006) の結果とは一致しなかった. Lagnado and Sloman (2006) の実験において共変動手がかり (i.e.,  $A \rightarrow B \rightarrow C \& D$ ) と時間順序手がかり (i.e.,  $A \rightarrow D \rightarrow C \rightarrow B$ ) が完全に異なる条件では, 時間順序手がかりの使用率が 58%となっており, 共変動手がかりの使用率 39%を上回っていた. これらの相反する結

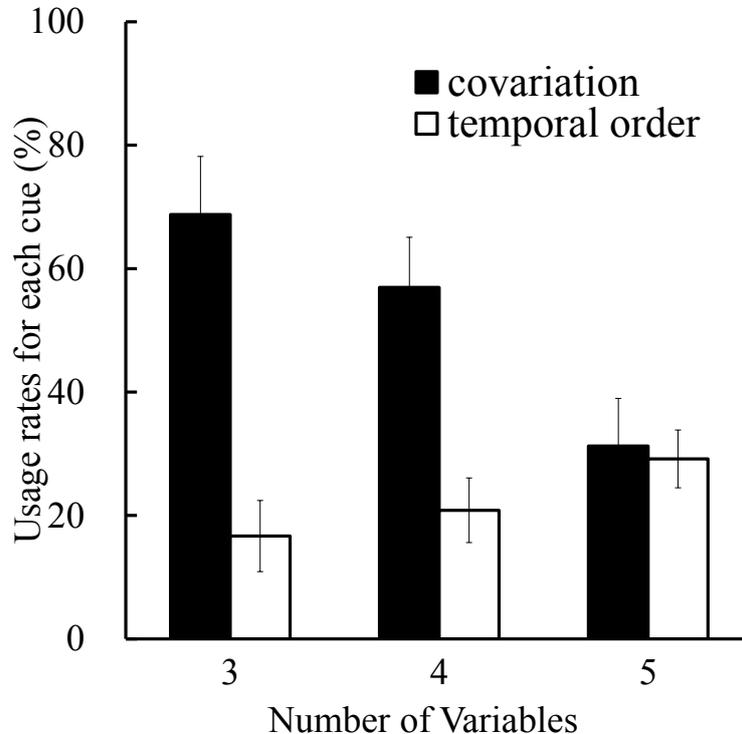


Figure 7-2. 実験 1 における共変動手がかりと時間順序手がかりの使用率. エラーバーは標準誤差を表す.

果は実験に使用するカバーストーリーに起因するかもしれない。本研究では微生物間の因果関係を取り扱ったのに対して、先行研究ではメールのやり取りを行うコンピューターネットワークが使用されていた。コンピューターの通信に関する事前の知識や経験が時間順序手がかりの重要性を高めたと考えられる。実際、研究1の結果は時間順序の信頼性が時間順序手がかりに基づく判断に影響することを示している。

要約すると、3事象や4事象の因果構造を学習する場合には共変動手がかりが重視されることが実験1の結果から示された。しかしながら、5事象の因果構造では共変動手がかりに対する選好が消失していた。このような結果の傾向は課題の複雑性によって判断方略が変化するという仮説を支持している。

### 7. 3. 実験2

実験1の結果、課題の複雑性によって因果構造の判断が変化することが明らかとなった。しかし、時間順序手がかりが共変動手がかりよりも重視されるという結果は見られなかった。実験2の目的は課題の複雑性と判断方略の関係について更なる検討を行うことであった。複雑な因果構造において時間順序手がかりが共変動手がかりよりも影響力を持つか確かめるため、異なる形態の因果構造が用いられた。具体的には、複数の因果関係を含む因果構造を採用した。因果構造に含まれるエッジの数を増加することによって、共変動情報はより複雑になるからである。因果構造が比較的単純な場合には共変動手がかりに基づく判断が、複雑な場合には時間順序手がかりに基づく判断が多くなると予測した。

#### 7. 3. 1. 方法

**実験参加者および実験計画** 大学生 24 名が実験に参加し、成績への加点と

なるコースクレジットを受け取った。この中に実験 1 の参加者は含まれていなかった。実験 1 と同様に、実験は実験参加者内計画で行われ、実験参加者は出現する事象の数が異なる 3 つの課題 (i.e., 3 事象, 4 事象, 5 事象) を遂行した。

**実験手続き** 実験参加者には微生物の状態を観察した後に、それらの因果関係を推測することが求められた。実験課題に使用する因果構造を除いて、手続きは実験 1 と同じであった。因果構造を構成する事象の数は 2 つの実験で等しくなっていたが、エッジの数は実験 2 の方が実験 1 よりも多くなっていた (Figure 7-1 および Figure 7-3 参照)。因果関係を表すエッジを増加した結果、共変動情報はより複雑になっていた。例えば、実験 1 の 4 事象条件では 5 種類の共変動情報が与えられたのに対して、実験 2 の 4 事象条件では 7 種類の共変動情報が与えられた。

教示では、カバーストーリーを通して微生物の因果関係を判断することが課題であると説明した。実験 1 と同様に、実験参加者には因果関係が確率的であることや、他に隠れた原因が存在しないことや、時間順序が当てにならないことが伝えられた。学習フェイズでは、実験参加者は 40 試行の観察を通して微

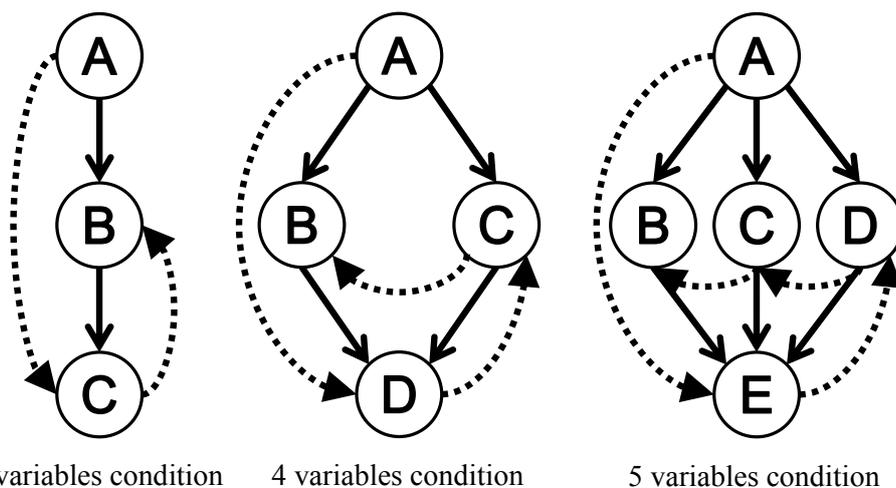


Figure 7-3. 実験 2 における因果構造。実線は実際の因果関係を、点線は実験参加者が観察する時間順序をそれぞれ表す。

生物の状態に関する共変動情報を受け取った。各条件では、観察する微生物の種類が異なっていた。Figure 7-3 に示したように、3 事象条件では 3 個の事象が 2 個のエッジで結ばれており、4 事象条件では 4 個の事象が 4 個のエッジで結ばれており、5 事象条件では 5 個の事象が 6 個のエッジで結ばれていた。また、各条件において共変動手がかりによって示唆される因果関係は実験参加者が観察する時間関係 (i.e., 時間順序手がかり) と異なっており、それぞれの手がかりがどの程度使用されているか査定することが可能になっていた。テストフェイズでは、実験 1 と同様の方法で因果構造を推測させた。

### 7. 3. 2. 結果および考察

実験参加者の回答は実験 1 と同様の方法で分析した。Figure 7-4 に各条件における共変動手がかりと時間順序手がかりの使用率を示す。3 事象条件では共変動手がかりが優先されており、実験 1 の結果が再現されていた。対照的に、4 事象条件や 5 事象条件では共変動手がかりよりも時間順序手がかりが重視されていた。手がかり (共変動 vs. 時間順序) × 事象の数 (3 事象 vs. 4 事象 vs. 5 事象) の二要因分散分析の結果、事象の数の主効果 ( $F(2, 46) = 9.31, MSE = 0.04, p < .001, \eta^2_G = .05$ ) および手がかりと事象の数の交互作用 ( $F(2, 46) = 9.66, MSE = 0.07, p < .001, \eta^2_G = .10$ ) が有意であった。単純主効果の検定を行ったところ、5 事象条件では時間順序手がかりが優先されることが明らかとなった ( $F(1, 23) = 10.97, MSE = 0.09, p = .003, \eta^2_G = .22$ )。3 事象条件と 4 事象条件では 2 つの手がかりの使用率に有意な差は見られなかった ( $F(1, 23) = 2.06, MSE = 0.20, ns, \eta^2_G = .06, F(1, 23) = 2.71, MSE = 0.08, ns, \eta^2_G = .03$ )。これらの結果は課題の複雑性によって判断方略が異なることを示唆している。

因果構造の形態が判断方略に及ぼす影響を検討するために、因果構造 (実験 1 vs. 実験 2) × 手がかり (共変動 vs. 時間順序) × 事象の数 (3 事象 vs. 4 事

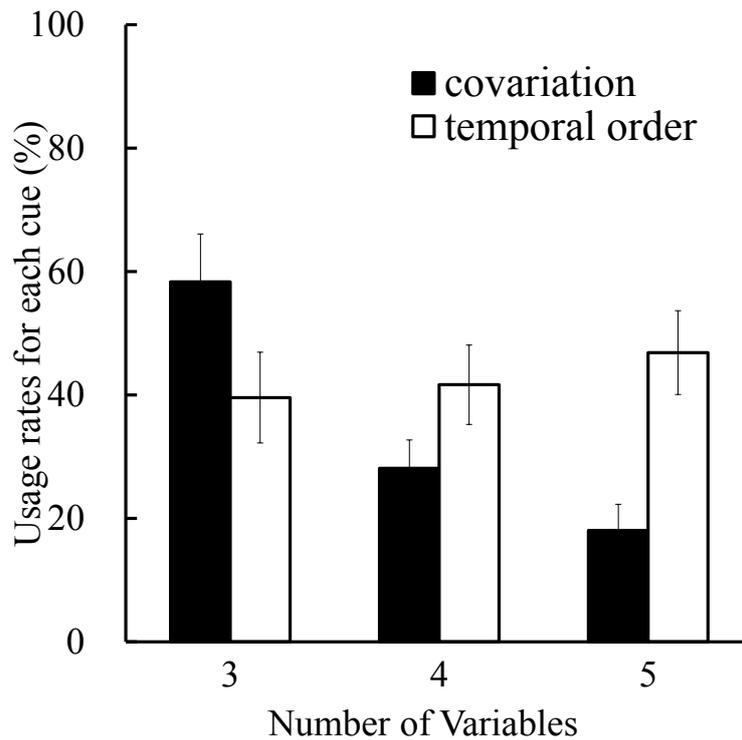


Figure 7-4. 実験 2 における共変動手がかりと時間順序手がかりの使用率。エラーバーは標準誤差を表す。

象 vs. 5 事象) の混合三要因分散分析を実施した。分析の結果、事象の数の主効果 ( $F(2, 92) = 12.03, MSE = 0.04, p < .001, \eta_G^2 = .03$ ) および因果構造と手がかりの交互作用 ( $F(1, 46) = 11.21, MSE = 0.23, p = .002, \eta_G^2 = .08$ ) ならびに手がかりと事象の数の交互作用 ( $F(2, 92) = 15.11, MSE = 0.06, p < .001, \eta_G^2 = .09$ ) が有意であった。因果構造と手がかりの交互作用について単純主効果の検定を行ったところ、実験 1 の因果構造の方が実験 2 の因果構造よりも共変動手がかりの使用率が高かった ( $F(1, 46) = 5.35, MSE = 0.21, p = .025, \eta_G^2 = .06$ )。一方、時間順序手がかりの使用率は実験 2 の因果構造の方が実験 1 の因果構造よりも高かった ( $F(1, 46) = 9.17, MSE = 0.16, p = .004, \eta_G^2 = .11$ )。これらの結果は、因果構造の形態もまた判断方略に影響することを示唆している。

実験 1 の結果と合わせると、実験 2 の結果は課題の複雑性による判断方略の変化について更なる証拠を提供している。実験参加者は 3 事象条件において共変動手がかりを重視し、5 事象条件では時間順序手がかりに重きを置いていた。これらの結果は、共変動手がかりへの選好を示す研究 1 の結果と時間順序手がかりへの選好を示す先行研究の結果 (Lagnado & Sloman, 2006) に統一的な見解をもたらすものであった。また、実験 1 と実験 2 の比較は、事象の数だけでなく、因果構造の形態も判断方略に影響を及ぼすことを示していた。

#### 7. 4. 実験 3

因果構造を学習する主要な方法として、観察 (observation) と介入 (intervention) という 2 つの方法が存在する。観察とは事象の自動的な振る舞いを観測することであり、介入とは事象に対して何らかの操作を加えることである (cf. Pearl, 2000)。実験 1 や実験 2 の結果、観察による学習において課題の複雑性によって判断方略が変化することが明らかとなったが、同様の傾向が介入による学習においても見られるかどうかは分からない。介入を行うことによって観察では区別できない因果構造を区別できるようになるため、介入による学習の方が観察による学習よりも因果構造を正しく学習できることは多くの研究によって示されており (e.g., Lagnado & Sloman, 2004; Sobel & Kushnir, 2006; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003)、両者には異なる過程が働いている可能性を示唆する研究もある (e.g., Fernbach & Sloman, 2009; Rottman & Keil, 2012)。実験 3 の目的は介入による学習において課題の複雑性が判断方略にどのような影響を及ぼすか検討することである。

##### 7. 4. 1. 方法

**実験参加者および実験計画** 大学生 24 名が実験に参加し、成績への加点と

なるコースクレジットを受け取った。この中に実験 1 や実験 2 の参加者は含まれていなかった。実験 1 や実験 2 と同様に、実験は実験参加者内計画で行われ、実験参加者は出現する事象の数が異なる 3 つの課題 (i.e., 3 事象, 4 事象, 5 事象) を遂行した。

**実験手続き** 実験課題は実験 2 とほぼ同様であったが、因果構造の学習方法が異なっていた。学習フェイズにおいて、実験参加者には容器の中に入れる微生物 (i.e., 最初の原因) を毎試行選択することが求められた。選択後に「NEXT」と書かれたボタンが画面上に表示され、そのボタンをクリックすると実際の因果関係とは異なる順序で生物の状態が表示された。各条件における因果構造や事象の提示順序は実験 2 と同様であった (Figure 7-3 参照)。テストフェイズでは、これまでの実験と同じ方法で因果構造を推測させた。

#### 7. 4. 2. 結果および考察

実験参加者の回答は実験 1 や実験 2 と同様の方法で分析した。Figure 7-5 に各条件における共変動手がかりと時間順序手がかりの使用率を示す。いずれの条件においても共変動手がかりが時間順序手がかりよりも重視されており、共変動手がかりの使用率が非常に高くなっていることが明らかとなった。手がかり (共変動 vs. 時間順序) × 事象の数 (3 事象 vs. 4 事象 vs. 5 事象) の二要因分散分析の結果、手がかりの主効果 ( $F(1, 23) = 61.10, MSE = 0.18, p < .001, \eta^2_G = .55$ ) および事象の数の主効果 ( $F(2, 46) = 7.39, MSE = 0.02, p = .002, \eta^2_G = .04$ ) が有意であり、手がかりと事象の数の交互作用が有意傾向であった ( $F(2, 46) = 3.11, MSE = 0.02, p = .005, \eta^2_G = .01$ )。観察による学習を扱った実験 2 とは異なり、介入による学習では課題の複雑性によって判断方略が共変動手がかりから時間順序手がかりへと変化することはなかった。なお、事象の数の増加に伴って時間順序手がかりの使用率が低下している傾向が見られたが、

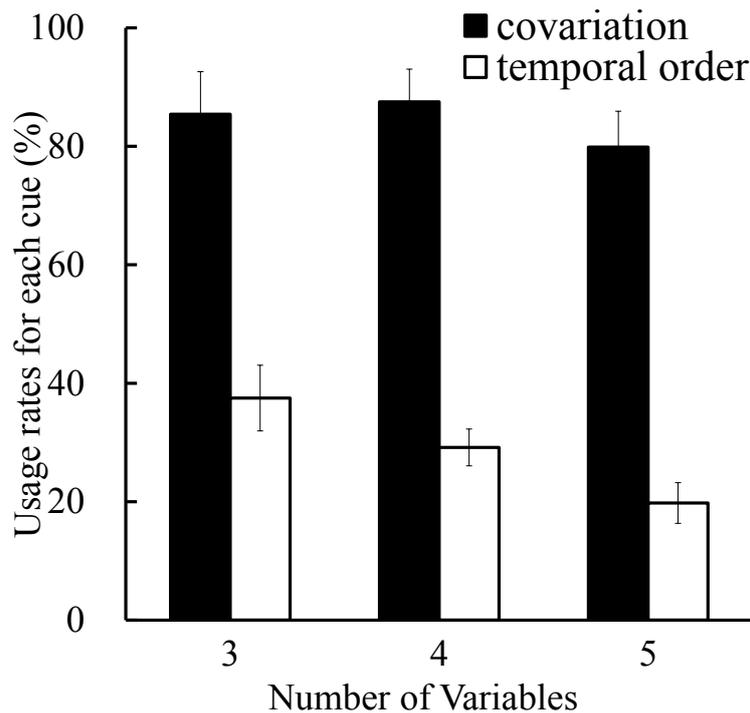


Figure 7-5. 実験 3 における共変動手がかりと時間順序手がかりの使用率. エラーバーは標準誤差を表す.

これは多くの実験参加者が最初の原因事象 (e.g., 5 事象条件における事象 E) に対してエッジを引いており, 時間順序手がかりが示唆する総エッジ数が条件間で異なることに由来するものであった.

学習方法の効果を詳細に検討するために, 観察による学習を扱った実験 2 の結果と介入による学習を扱った実験 3 の結果の比較を行った. 学習方法 (観察 vs. 介入) ×手がかり (共変動 vs. 時間順序) ×事象の数 (3 事象 vs. 4 事象 vs. 5 事象) の混合三要因分散分析を実施した. 分析の結果, 学習方法の主効果 ( $F(1, 46) = 17.98, MSE = 0.13, p < .001, \eta_G^2 = .09$ ) および手がかりの主効果 ( $F(1, 46) = 19.81, MSE = 0.21, p < .001, \eta_G^2 = .15$ ) ならびに事象の数の主効果 ( $F(2, 92) = 14.98, MSE = 0.03, p < .001, \eta_G^2 = .04$ ) が有意であった. また, 学習方法と手がかりの交互作用 ( $F(1, 46) = 35.10, MSE = 0.21, p < .001, \eta_G^2 = .24$ ) や手がかりと事象の数の交互作用 ( $F(2, 92) = 4.26, MSE = 0.04, p$

= .016,  $\eta_G^2 = .02$ ) も有意であった。さらに、学習方法と手がかりと事象の数の二次交互作用も有意であった ( $F(2, 92) = 12.64, MSE = 0.04, p < .001, \eta_G^2 = .05$ )。二次交互作用について単純交互作用の検定を行ったところ、観察による学習では手がかりの効果が事象の数によって大きく変化していたのに対して ( $F(2, 46) = 9.66, MSE = 0.07, p < .001, \eta_G^2 = .10$ )、介入による学習ではあまり変化していなかった ( $F(2, 46) = 3.11, MSE = 0.02, p = .005, \eta_G^2 = .01$ )。また、学習方法と手がかりの交互作用に関して単純主効果の検定を実施した結果、介入群の方が観察群よりも共変動手がかりを多く、時間順序手がかりを少なく利用していることが明らかとなった ( $F(1, 46) = 47.54, MSE = 0.18, p < .001, \eta_G^2 = .43, F(1, 46) = 4.72, MSE = 0.15, p = .035, \eta_G^2 = .06, respectively$ )。この結果は、介入を行った際には共変動情報がより確かな手がかりとして重視されることを示している。共変動手がかりに基づく正しい判断が増加していたことは、介入による因果構造の学習の促進効果を報告している先行研究 (e.g., Steyvers et al., 2003) と一致していた。

実験 3 の結果、介入による学習では課題の複雑性に伴う判断方略の変化は見られなかった。このことは、観察による学習と介入による学習では異なる過程が機能していることを示唆している。また、共変動手がかりは介入による学習においてより重視されることが示された。

## 7. 5. 総合論議

本研究では因果構造の学習における課題の複雑性と判断方略の関係について検討を行った。実験の結果、事象の数が少ない単純な課題では共変動手がかりを、事象の数が多し複雑な課題では時間順序手がかりを重視することが明らかとなった (実験 1)。このような判断方略の変化は異なる因果構造でも確認された (実験 2)。これらの結果は、共変動手がかりに基づく判断を報告している研究 1 だけでなく、時間順序手がかり

りに基づく判断を報告している Lagnado and Sloman (2006) や White (2006) とも一致しており、相反する 2 つの知見に対して統一的な見解をもたらしている。また、介入による学習では課題の複雑性の変化に伴う判断方略の変化は見られなかった (実験 3)。

本研究の結果はカテゴリー学習における明瞭度と妥当性の観点から解釈することもできる。Kruschke and Johansen (1999)によると、無関係な手がかりの存在は妥当な手がかりの使用を阻害し、この効果は無関係な手がかりの明瞭度が増すほどに大きくなる。本実験では、共変動情報が妥当な手がかりであり、時間順序は明瞭ではあるが無関係な手がかりであった。因果構造を構成する事象の数の増加によって、共変動手がかりの明瞭度が下がり、時間順序手がかりの明瞭度が上昇した結果、共変動手がかりが使用されなくなったと考えられる。これらの類似した現象は、因果学習とカテゴリー学習の間に密接な関係があることを示唆している。

観察による学習 (実験 1 と実験 2) と介入による学習 (実験 3) で違いが見られたことは、両者のメカニズムが異なっていることを示唆している。因果構造の学習における観察と介入の相違点に関して以下の 2 点が考えられる。第一に、共変動手がかりの変化が挙げられる。観察とは異なり、介入を行った際には共変動手がかりはより有益なものとなる (e.g., Steyvers et al., 2003)。例えば、事象 X と事象 Y の共変関係によって 3 種類の因果構造 (i.e.,  $X \rightarrow Y$ ,  $X \leftarrow Y$ , or  $X \leftarrow Z \rightarrow Y$ ) が示唆される事態において、事象 X を操作して事象 Y に影響がなければ、事象 X が事象 Y に影響するという因果関係 (i.e.,  $X \rightarrow Y$ ) は排除される。介入によって共変動手がかりがより有益になり、その使用が促進されたと考えられる。第二に、時間順序手がかりの信頼性が挙げられる。事象の状態を観察して因果構造を学習している場合には時間順序手がかりが実際の因果関係と一致しているか否か判断することは困難であるが、介入を行う場合には時間順序手がかりが有効であるか否か検証することができる。時間順序手がかりが有効でないことに気付いた結果、時間順序手がかりの使用率が低くなったと思われる。介入による因果構

造の学習の促進はこれらの差異に由来すると考えられる。

因果構造の学習に関して、因果ベイズネットでは観察と介入で学習メカニズムが異なることは想定されていないが、本研究の結果はこれらを区別して考える必要性を示している。観察による因果構造の学習と介入による因果構造の学習の両者を包括的に説明するモデルの構築が求められる。また、モデリングを通してどのように事象を操作することが最適であるか、実験参加者の介入対象の選択は最適であるか検討することも望まれる。

## 第 3 部

# 因果強度の学習

## 8. 研究 3 因果強度の学習に関する数理モデルの比較検討

### 8. 1. 序論

#### 8. 1. 1. はじめに

事象間の因果関係を学習することによって、その事象を制御することが可能になる。ただし、対象を正確に制御するためには、どの変数がどのような影響を持つかきちんと査定されていなければならない。研究 1 と研究 2 では何が原因で何が結果となるかという因果構造の学習に焦点を当てたが、研究 3 では原因が結果に対してどのような影響を及ぼすかという因果強度の学習に焦点を当てる。この問題は因果性の学習と推論の中でも最も検討されてきたテーマであり、異なるアプローチから様々なモデルが提唱されている。ここでは、これまでに蓄積された知見を用いてそれぞれの数理モデルの比較検討を行う。

因果強度の学習では、原因候補と結果が予め決められており、学習者はそれらの因果関係の強さを判断することが求められる。なお、原因候補と表現するのは、実際には無関係である可能性も含まれているからである。一般的な実験事態では原因候補と結

		Event $E$	
		$e$	$\neg e$
Event $C$	$c$	a	b
	$\neg c$	c	d

Figure 8-1. 原因事象( $C$ )と結果事象( $E$ )の分割表。アルファベット  $a$ ,  $b$ ,  $c$ ,  $d$  はそれぞれの共生起の頻度を表す。

果の 2 値 2 変数が使用され、これらの共変動情報は Figure 8-1 に示すような 2×2 の分割表で表現することができる。実験参加者には原因事象と結果事象の状態が 1 つずつ継時的に、あるいは何らかの形でまとめられて要約的に提示される。これらの情報は「投薬の有無と症状の改善の有無」などのように、架空のカバーストーリーを通して与えられる。その後、「症状の改善に対する投薬の効果」といった質問に対して、0（まったく効果がない）から 100（必ず改善する）の尺度を用いてカバーストーリーに応じた形で因果強度を評価することが求められる。抑制的因果関係の可能性も含む場合には、-100 から+100 の間で評定が求められる。例えば、Buehner, Cheng, and Clifford (2003) の実験では、放射線被曝がウイルスの突然変異を引き起こすか調べるというカバーストーリーの下で、被曝の有無と突然変異の有無に関する情報が与えられ、放射線がどのくらい変異を引き起こすか判断することが求められた。

#### 8. 1. 2. 因果強度の学習に関する数理モデル

**規則ベースアプローチ** 規則ベースアプローチでは、事象間の関係性を表す規範的な統計量あるいはヒューリスティックスといった簡便な方略によって事象の生起情報から因果関係が引き出されると考えられている。このアプローチにおける最も代表的なモデルは  $\Delta P$  ルール (Jenkins & Ward, 1965; Ward & Jenkins, 1965) であり、以下のように定義される。

$$\Delta P = P(e|c) - P(e|\neg c) = \frac{a}{a+b} - \frac{c}{c+d} \quad (8-1)$$

ここで  $P(e|c)$  は原因が生起した時に結果が生起する条件付き確率を、 $P(e|\neg c)$  は原因が生起しなかった時に結果が生起する条件付き確率を表す。 $\Delta P$  はこれら 2 つの条件付き確率の差分をとることで算出され、-1 から+1 までの値になる。原因の生起によって結果

の生起確率が上昇する場合 (i.e.,  $\Delta P > 0$ ), 原因と結果の間には発生的因果関係が想定される. 一方, 原因の生起によって結果の生起確率が低下する場合 (i.e.,  $\Delta P < 0$ ), 原因と結果の間には抑制的因果関係が想定される. 原因の有無によって結果の生起確率が変化しない場合 (i.e.,  $\Delta P = 0$ ), 原因と結果は無関係となる.

$\Delta P$  モデルでは各セルの情報は等しく重み付けられているが, 実際にはそうならないことが示されている. 例えば Kao and Wasserman (1993) は, 4 種類のセル頻度は  $a > b > c > d$  の順で優先されることを示している. また, 抑制的因果関係を判断する場合は  $b > a > d > c$  の順に重視されることが報告されている (Mandel & Vartanian, 2009). これらの知見を反映したより記述的なモデルとして, それぞれの生起頻度に重みを付けた重み付き  $\Delta P$  ルールが提唱されている. 重み付けの方法はいくつか考案されている (e.g., Einhorn & Hogarth, 1986; Schustack & Sternberg, 1981), 本研究では Anderson and Sheu (1995) の重み付き  $\Delta P$  ルールを使用する.

$$\Delta P_w = \beta_0 + \beta_1 P(e|c) - \beta_2 P(e|\neg c) \quad (8-2)$$

ここで  $\beta_1$  と  $\beta_2$  は条件付き確率の重みを表すパラメータであるが,  $\beta_0$  を切片,  $\beta_1$  と  $\beta_2$  を回帰係数とする回帰モデルとしても捉えられる.

いくつかの研究では  $\Delta P$  ルールとは異なる判断方略が使用されることが報告されている (e.g., Shimazaki, Tsuda, & Imada, 1991). その中でも頻繁に見られるのは  $\Delta D$  ルール (Inhelder & Piaget, 1958) と呼ばれるものであり, これは確信事例 (i.e., セル  $a$  とセル  $d$ ) の数から非確信事例 (i.e., セル  $b$  とセル  $c$ ) の数を引くことで計算される.

$$\Delta D = (a + d) - (b + c) \quad (8-3)$$

$\Delta D$  ルールは頻度に基づいた予測を行うため、確信事例が増え続けた場合に上昇し続けてしまうという欠点を持つ。White (2003) は  $\Delta D$  ルールを比率に基づく形に変更し、 $pCI$  ルールと命名した。 $pCI$  ルールは  $\Delta P$  ルールと同様に  $-1$  から  $+1$  までの値をとる。

$$pCI = \frac{a + d - b - c}{a + b + c + d} \quad (8-4)$$

Perales and Shanks (2007) は  $\Delta D$  ルールを比率に基づく形に変更し、それぞれの生起頻度に重みを付け、 $EI$  ルールと命名している。 $w$  はセルの重みを表すパラメータであり、全ての情報の重みが等しい場合 (i.e.,  $w_a = w_b = w_c = w_d = 1$ )、 $pCI$  ルールと一致する。

$$EI = \frac{w_a a + w_d d}{w_a a + w_b b + w_c c + w_d d} - \frac{w_b b + w_c c}{w_a a + w_b b + w_c c + w_d d} \quad (8-5)$$

Hattori and Oaksford (2007) は連関係数の指標である  $\varphi$  係数に変更を加えて独自のモデルを提唱している。原因と結果が非生起の情報 (i.e., セル  $d$ ) は何も起きていないことを表しており、現実的にはこの生起頻度を定義することは困難であり、また便宜的に定義しても無数に存在しているので数え上げることは難しい。そこで、 $\varphi$  係数に含まれる  $d$  を無限大に発散させた二要因ヒューリスティックモデルが提唱されている。

$$H_G = \lim_{d \rightarrow \infty} \varphi = \sqrt{P(e|c)P(c|e)} = \frac{a}{\sqrt{(a+b)(a+c)}} \quad (8-6)$$

ここで言う二要因とは、原因から結果への予測可能性  $P(e|c)$  と原因の結果に対する適

合性  $P(c|e)$  という 2 つの要因のことである。また、抑制的因果関係の場合  $H$  は以下のよう  
に算出される。

$$H_p = \frac{b}{\sqrt{(a+b)(b+d)}} \quad (8-7)$$

**連合的アプローチ** 動物の条件づけ研究に端を発する連合的アプローチでは、因果関係の学習を事象間の連合強度の変化として捉える。このアプローチにおける最も代表的なモデルは Rescorla-Wagner モデル (Rescorla & Wagner, 1972) であり、連合強度の変化は次のように計算される。

$$\Delta V_i = \alpha\beta \left( \lambda - \sum V_T \right) \quad (8-8)$$

ここで  $\Delta V_i$  は原因と結果の連合強度の変化量を、 $\sum V_T$  は既に形成されている連合強度を表す。結果事象が生起する場合には  $\lambda = 1$  となり、連合強度は増加する。一方、結果事象が生起しない場合には  $\lambda = 0$  となり、連合強度は減少する。学習の初期などで  $\sum V_T$  の値が小さいほど、連合強度の変化量は大きくなる。 $\alpha$  は原因事象の明瞭度、 $\beta$  は結果事象の明瞭度を表すパラメータであり、0 から 1 の値をとる。これら 2 つのパラメータも連合強度の変化量に影響を及ぼす。

一般的に結果事象の明瞭度  $\beta$  は、結果が生起する時と生起しない時で異なる値になると想定されているが、両者を一定の値に設定した場合は Rescorla-Wagner モデルの漸近値と(8-1)式に示した  $\Delta P$  の値が一致することが知られている (Chapman & Robbins, 1990; Wassrman, Elek, Chatlosh, & Baker, 1993; see Danks, 2003 for detailed analysis)。Rescorla-Wagner モデルの漸近値はセルの生起頻度を用

いて以下のように表現することも可能である (Wasserman et al., 1993).

$$V_{asympt} = \frac{\beta_e a}{\beta_e a + \beta_{-e} b} - \frac{\beta_e c}{\beta_e c + \beta_{-e} d} \quad (8-9)$$

Rescorla-Wagner モデルに対してはいくつかの修正が提案されているが (e.g., Van-Hamme & Wasserman, 1994), 単一の原因事象と結果事象の事象において両者は同じ予測を行う (see Perales & Shanks, 2003). また, 修正 SOP モデル (Dickinson & Burke, 1996) も単一の原因事象と結果事象の事象では類似した予測値となるため (see Perales & Shanks, 2003), Rescorla-Wagner モデルにまとめて検討する.

Rescorla-Wagner モデルと同様の学習規則に基づくものの, Pearce モデル (Pearce, 1987) では異なる予測値が導かれる. Pearce モデルの特徴は, 複数の手がかり刺激を 1 つの複合刺激として捉えることであり, 2 つの刺激の間には類似性に基づいた般化が生じることが想定されている. 原因事象と結果事象が 1 つしか存在しない事象では, 文脈刺激 ( $X$ ) と原因事象と文脈刺激の複合刺激 ( $CX$ ) という 2 つの刺激に対して学習が生じる. 一方の刺激が提示された場合には, 他方の刺激表象が類似性に応じて活性化されるため, 連合強度の変化量は以下のように計算される.

$$\Delta V_{CX} = \beta(\lambda - V_{CX} - x_1 V_X) \quad (8-10)$$

$$\Delta V_X = \beta(\lambda - V_X - x_2 V_{CX}) \quad (8-11)$$

原因事象が生起している場合は(8-10)式が, 生起していない場合は(8-11)式が適用される. この式に含まれる多くのパラメータは Rescorla-Wagner モデルと同じ意味を持つ. 新たなパラメータ  $x$  は刺激間の類似性によって規定され, 般化の程度を表す. 原因事

象に対する判断は原因事象( $C$ )と原因事象と文脈刺激の複合刺激( $CX$ )の類似性によって決まる.

$$J_C = x_3 V_{CX} \quad (8-12)$$

それぞれのセルの生起頻度を用いて表現した場合は以下ようになる (Perales & Shanks, 2003).

$$J_C = \frac{x_3[a(c+d) - c(ax_1 + bx_1)]}{a(c+d) + b(c+d) - cx_2(ax_1 + bx_1) - dx_2(ax_1 + bx_1)} \quad (8-13)$$

**因果的アプローチ** ヒトは与えられた共変動情報を表面的に処理するのではなく、その背後に潜む因果関係を推測する (Waldmann, Hagmayer, & Blaisdell, 2006). 因果的アプローチでは、因果モデルの構築およびパラメータの推定の問題として因果学習が扱われる. Cheng (1997) は  $\Delta P$  ルールの主要な問題点として、以下の2点を指摘している. 原因の有無に関わらず結果が常に生起する天井効果の事態では (i.e.,  $P(e|c) = P(e|\neg c) = 1$ ), 実際には因果効果を判断することはできないが、 $\Delta P$  ルールは効果がないと判断してしまう. また、 $\Delta P$  は背景原因の影響を考慮していないため、因果効果の適切な推定値とはならない. 共変動と因果を区別する指標として、因果パワーが提唱されている.

$$q_G = \frac{\Delta P}{1 - P(e|\neg c)} \quad (8-14)$$

因果パワーとは、他の原因が存在しない事態において原因事象が結果事象を発生させ

る確率である。  $q_G = 1$  は原因が結果を必ず発生させることを意味し、  $q_G = 0$  は影響を及ぼさないことを意味する。背景原因によって結果が常に生起する、つまり天井効果が生じている場合 (i.e.,  $P(e|c) = P(e|\neg c) = 1$ ), 因果パワー  $q_G$  は定義されない。また、背景原因が存在しない場合 (i.e.,  $P(e|\neg c) = 0$ ),  $q_G = \Delta P$  となり、共変動と因果は一致する。また、抑制的因果関係における因果パワーは以下のように算出される。

$$q_P = \frac{\Delta P}{P(e|\neg c)} \quad (8-15)$$

Griffiths and Tenenbaum (2005, 2009) は、これまでパラメータ推定の問題として扱われていた因果強度の学習をモデル比較の問題として捉え直し、原因事象と結果事象の間に因果関係が存在するかという構造の判断から説明を試みた。ここで言うモデルとは、背景原因 ( $B$ ) のみが結果事象 ( $E$ ) に影響を及ぼす因果モデル (Graph 0:  $B \rightarrow E$ ) と背景原因 ( $B$ ) と原因事象 ( $C$ ) が結果事象 ( $E$ ) に影響する因果モデル (Graph 1:  $B \rightarrow E \leftarrow C$ ) のことである。背景原因と原因事象が独立に結果を生起させる場合、結果事象の生起確率は noisy-OR 関数を用いて以下のように算出される。

$$P(e|b, c; w_0, w_1) = w_0 b + w_1 c - w_0 w_1 bc \quad (8-16)$$

ここで  $w_0$  は背景原因の因果強度を、  $w_1$  は原因事象の因果強度を表す。また、  $b$  と  $c$  はそれぞれ背景原因と原因事象の有無を表し、存在するときは 1、存在しないときは 0 になる。抑制的因果関係の場合、背景原因には発生的影響を、原因事象には抑制的影響を想定する。このときの結果事象の生起確率は noisy-AND-NOT 関数を用いて以下のように算出される。

$$P(e|b, c; w_0, w_1) = w_0b - w_0w_1bc \quad (8-17)$$

なお、発生的因果パワー $q_G$ は noisy-OR 関数の下での、抑制的因果パワー $q_P$ は noisy-AND-NOT 関数の下での  $w_1$  の最尤推定量となることが示されている。

ベイズの定理を適用することにより、得られたデータが Graph 0 によって生じたか、Graph 1 によって生じたかは対数オッズの形で以下のように表現できる。

$$\log \frac{P(\text{Graph 1}|D)}{P(\text{Graph 0}|D)} = \log \frac{P(D|\text{Graph 1})}{P(D|\text{Graph 0})} + \log \frac{P(\text{Graph 1})}{P(\text{Graph 0})} \quad (8-18)$$

この式の右辺の第一項である対数尤度比はデータが 2 つの因果モデルのどちらが尤もらしいかを表す指標であり、Griffiths and Tenenbaum (2005) はこの対数尤度比を因果サポートと名づけた。

$$\text{causal support} = \log \frac{P(D|\text{Graph 1})}{P(D|\text{Graph 0})} \quad (8-19)$$

なお、それぞれの尤度は以下の式から算出される。

$$P(D|\text{Graph 0}) = \int_0^1 P(D|w_0, \text{Graph 0})P(w_0|\text{Graph 0})dw_0 \quad (8-20)$$

$$P(D|\text{Graph 1}) = \int_0^1 \int_0^1 P(D|w_0, w_1, \text{Graph 1})P(w_0, w_1|\text{Graph 1})dw_0dw_1 \quad (8-21)$$

ここで  $P(w_0|\text{Graph 0})$ と  $P(w_0, w_1|\text{Graph 1})$  は因果強度の事前分布となっており、事前分

布には 0 から 1 までの一様分布が想定されている。

Lu, Liljeholm, Cheng, and Holyoak (2008) は, 因果強度の事前分布に偏りが  
ない一様分布を想定するよりも, 原因事象の数を最少にしつつ因果強度を最大にする  
因果モデルへの選好を反映した分布を想定する方が適切であると主張している. Lu et  
al. (2008) は(8-21)式における因果強度の事前分布に必要な十分な因果モデルへの  
選好を組み込み, この事前分布に基づくベイズモデルを SS パワーモデルと名づけた.  
背景原因と原因事象が結果事象を発生させる場合, 事前分布は以下のように表現さ  
れる.

$$P(w_0, w_1 | gen, Graph 1) \propto (e^{-\alpha w_0 - \alpha(1-w_1)} + e^{-\alpha(1-w_0) - \alpha w_1}) \quad (8-22)$$

ここで  $\alpha$  は必要十分な原因への選好を示すパラメータであり,  $\alpha = 0$  となる時, 事前分布  
は一様分布となる. この事前分布は原因候補の因果強度と背景原因の因果強度が両  
立しないこと, つまり, 一方の値が高く想定される場合は他方の値が低く想定されること  
を示している. また, 抑制的因果関係に対する因果強度の事前分布は以下の形で計  
算される.

$$P(w_0, w_1 | prev, Graph 1) \propto (e^{-\alpha(1-w_0) - \alpha(1-w_1)} + e^{-\alpha(1-w_0) - \alpha w_1}) \quad (8-23)$$

この事前分布では, 背景原因の因果強度が高くなること, および原因事象の因果強度  
は高いか低いかの両極端な値になることが想定されている. なお, Lu et al. (2008) に  
従って,  $\alpha = 5$  としてシミュレーションを行い, 因果強度の判断の予測値には  $w_1$  の推定平  
均値を使用した.

### 8. 1. 3. 本研究の目的

因果強度の学習に関して、異なるアプローチから様々なモデルが提唱されており、共変動情報を操作した実験によってその比較検討が行われてきた。Perales and Shanks (2007) は共変動情報のみを独立変数とした研究を対象としてメタ分析を行い、主要なモデルの比較検討を行った。9つの実験から114条件が選出され、それぞれのモデルのシミュレーションを行った結果、EIルールが最も高い適合度を示したことを報告している。一方、Hattori and Oaksford (2007) は同じようなメタ分析の結果、二要因ヒューリスティックモデルの適合度が最も高くなることを示している。これら2つの研究はほぼ同時期に発表されたため、EIルールと二要因ヒューリスティックモデルの比較はまだ行われていない。また、この後に提唱されたSSパワーモデルとの比較もまだ行われていない。そこで本研究では、Hattori and Oaksford (2007) より多くの研究を対象としている Perales and Shanks (2007) と同様のメタ分析を行い、新たに提唱されたモデルを加えた上で代表的な数理モデルの比較検討を行った。

### 8. 2. 方法

Perales and Shanks (2007) と同様に、以下の5つの基準に基づいて研究を選定した。(1)単一の原因事象と単一の結果事象の因果関係を対象としている。(2)離散試行手続きを用いて一試行ずつ共変動情報を与えている。(3)因果関係を尋ねる一般的な質問方法を採用している。(4)試行ごとではなく、全体での評定を求めている。(5)共変動情報以外の要因を操作していない。これらの5つの基準を満たすものとして、Wasserman, Kao, Van-Hamme, Katagiri, and Young (1996), Vallée-Tourangeau, Murphy, Drew, and Baker (1998), Lober and Shanks (2000), Shanks (2002), Buehner, Cheng, and Clifford (2003), Perales and Shanks (2003), White (2003), Perales and Shanks (2004), Collins and Shanks (2006) の9つ実験から114条件が選ばれた。なお、Anderson and Sheu

(1995) の実験は上記の基準を満たすものの、一部の条件の試行数が極めて少なく、連合学習モデルが漸近値に達しないと考えられるため、分析には含めなかった。

本研究で比較する数理モデルの中には、事象の生起頻度のみから計算できるモデルもあれば、いくつかのフリーパラメータを含むモデルもある。モデルが複雑になればなるほど、データへの当てはまりは高まるため、これらを直接比較することは適切ではない。Perales and Shanks (2007) と同様に、交差検証を用いてパラメータの推定を行った。この方法では、データを二分した後、その半分をパラメータの推定に使用し、残りの半分を用いてデータへの適合度を検討する。この手続きを採用することにより、同じ指標を用いた比較が可能になる。モデルの予測値と評定値の決定係数を算出してモデル比較を行った。

### 8. 3. 結果

Table 8-1 にメタ分析に使用した 114 条件の共変動情報および各種数理モデルの予測値ならびに実際の評定値を示した。シミュレーションの結果、最も高い適合度を示したのは SS パワーモデルであった ( $r^2 = .91$ )。フリーパラメータを含む他のモデルも SS パワーモデルに及ばないものの、高い適合度を示していた。(EI ルール,  $r^2 = .89$ ; 重み付き  $\Delta P$  ルール,  $r^2 = .85$ ; Pearce モデル,  $r^2 = .85$ , Rescorla-Wagner モデル,  $r^2 = .80$ )。また、パワー PC 理論と二要因ヒューリスティックモデルはフリーパラメータを含まないにも関わらず、それぞれ良い当てはまりを示していた ( $r^2 = .86$ ,  $r^2 = .85$ , respectively)。 $\Delta P$  ルールと  $pCI$  ルールはパワー PC 理論や二要因ヒューリスティックモデルを若干下回る程度であったが ( $r^2 = .80$ ,  $r^2 = .81$ , respectively)、 $\Delta D$  ルールの適合度はこれらを大きく下回っていた ( $r^2 = .55$ )。また、因果的アプローチの 1 つである因果サポートモデルも適合度はあまり高くはなかった ( $r^2 = .43$ )。

$\Delta D$  ルールと  $pCI$  ルールと EI ルールは、確信事例と非確信事例の比較を行うという点で共通しているが、比率の形を取らない  $\Delta D$  ルールの適合度は低かった。また、 $pCI$

Table 8-1

メタ分析に使用したデータおよび数理モデルの予測値

Study	Experiment	Covariation information				Non-parameterized models						Parameterized models				Rating	
		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	$\Delta P$	$\Delta D$	<i>pCI</i>	<i>H</i>	causal power	causal support	SS power	$w\Delta P$	EI	RW		Pearce
Perales & Shanks (2004)	1a	36	4	8	32	0.70	56.00	0.70	0.86	0.88	11.87	0.87	0.74	0.76	0.71	0.84	72.00
		22	18	4	36	0.45	36.00	0.45	0.68	0.50	5.50	0.53	0.47	0.43	0.43	0.54	46.00
		36	4	32	8	0.10	8.00	0.10	0.69	0.50	1.75	0.18	0.25	0.38	0.11	0.28	40.00
Perales & Shanks (2003)	1	10	0	6	4	0.40	8.00	0.40	0.79	1.00	2.38	0.68	0.51	0.61	0.44	0.58	87.00
		4	6	0	10	0.40	8.00	0.40	0.63	0.40	1.36	0.46	0.40	0.31	0.37	0.46	35.00
		6	4	10	0	-0.40	-8.00	-0.40	-0.63	-0.40	-1.36	-0.30	-0.21	-0.10	-0.44	-0.25	3.00
		0	10	4	6	-0.40	-8.00	-0.40	-0.79	-1.00	-2.38	-0.87	-0.32	-0.58	-0.37	-0.38	-64.00
	2	10	0	1	9	0.90	18.00	0.90	0.95	1.00	5.30	0.93	0.92	0.93	0.91	1.05	92.00
		10	0	7	3	0.30	6.00	0.30	0.77	1.00	2.05	0.46	0.43	0.55	0.33	0.49	61.00
		8	2	8	2	0.00	0.00	0.00	0.63	0.00	0.54	0.05	0.15	0.27	0.00	0.17	25.00
		2	8	2	8	0.00	0.00	0.00	0.32	0.00	-0.54	-0.15	0.04	-0.11	0.00	0.04	-15.00
	3	0	10	3	7	-0.30	-6.00	-0.30	-0.77	-1.00	-2.05	-0.83	-0.24	-0.50	-0.27	-0.28	-61.00
		0	10	9	1	-0.90	-18.00	-0.90	-0.95	-1.00	-5.28	-0.93	-0.73	-0.93	-0.89	-0.85	-78.00
		10	0	2	8	0.80	16.00	0.80	0.91	1.00	4.46	0.93	0.84	0.87	0.82	0.96	83.00
		10	0	8	2	0.20	4.00	0.20	0.75	1.00	1.78	0.29	0.35	0.49	0.22	0.39	53.00
Shanks (2002)	1	24	6	6	24	0.60	36.00	0.60	0.80	0.75	6.59	0.76	0.64	0.66	0.60	0.73	68.00
		0	30	6	24	-0.20	-12.00	-0.20	-0.75	-1.00	-3.37	-0.90	-0.16	-0.43	-0.18	-0.19	-35.00
		0	30	30	0	-1.00	-60.00	-1.00	-1.00	-1.00	-20.03	-0.97	-0.81	-1.00	-1.00	-0.94	-99.00
	2	24	6	6	24	0.60	36.00	0.60	0.80	0.75	6.59	0.76	0.64	0.66	0.60	0.73	67.00
		3	27	10	20	-0.23	-14.00	-0.23	-0.72	-0.70	-2.51	-0.67	-0.17	-0.36	-0.21	-0.20	-48.00
		9	21	30	0	-0.70	-42.00	-0.70	-0.84	-0.70	-10.14	-0.70	-0.51	-0.52	-0.73	-0.60	-70.00
Lober & Shanks (2000)	1	21	9	0	30	0.70	42.00	0.70	0.84	0.70	17.41	0.71	0.70	0.68	0.67	0.80	68.00
		24	6	10	20	0.47	28.00	0.47	0.75	0.70	5.95	0.72	0.53	0.57	0.47	0.60	63.00
		27	3	20	10	0.23	14.00	0.23	0.72	0.70	2.09	0.44	0.36	0.46	0.25	0.40	55.00
		28	0	0	28	1.00	56.00	1.00	1.00	1.00	33.94	0.97	1.00	1.00	1.00	1.15	96.00
	2	28	0	7	21	0.75	42.00	0.75	0.89	1.00	18.47	0.96	0.80	0.84	0.78	0.91	87.00
		28	0	14	14	0.50	28.00	0.50	0.82	1.00	9.92	0.95	0.60	0.68	0.54	0.68	78.00
		28	0	21	7	0.25	14.00	0.25	0.76	1.00	3.99	0.77	0.39	0.52	0.28	0.44	65.00
		24	6	12	18	0.40	24.00	0.40	0.73	0.67	4.35	0.67	0.48	0.52	0.41	0.54	54.00
	3	12	18	0	30	0.40	24.00	0.40	0.63	0.40	6.98	0.43	0.40	0.31	0.37	0.46	38.00
		27	3	25	5	0.07	4.00	0.07	0.68	0.40	0.15	0.21	0.23	0.36	0.07	0.25	34.00
		30	15	0	45	0.67	60.00	0.67	0.82	0.67	25.17	0.68	0.67	0.64	0.63	0.76	63.00
		36	9	18	27	0.40	36.00	0.40	0.73	0.67	6.73	0.68	0.48	0.52	0.41	0.54	55.00
Collins & Shanks (2006)	1	18	27	0	45	0.40	36.00	0.40	0.63	0.40	11.48	0.42	0.40	0.31	0.37	0.46	41.00
		32	8	4	36	0.70	56.00	0.70	0.84	0.78	20.14	0.78	0.72	0.72	0.69	0.82	82.00
		36	4	22	18	0.35	28.00	0.35	0.75	0.78	5.66	0.75	0.45	0.54	0.37	0.51	63.00
	2	14	26	0	40	0.35	28.00	0.35	0.59	0.35	8.08	0.37	0.35	0.25	0.32	0.40	49.00
		20	20	20	20	0.00	0.00	0.00	0.50	0.00	-1.37	0.10	0.10	0.09	0.00	0.10	32.00
		20	0	10	10	0.50	20.00	0.50	0.82	1.00	4.92	0.93	0.60	0.68	0.54	0.68	94.00
Vallée-Tourangeau et al. (1998)	1	5	15	5	15	0.00	0.00	0.00	0.35	0.00	-0.47	-0.09	0.05	-0.08	0.00	0.05	-20.00
		10	10	10	10	0.00	0.00	0.00	0.50	0.00	0.00	-0.03	0.10	0.09	0.00	0.10	1.00
		15	5	15	5	0.00	0.00	0.00	0.61	0.00	0.47	0.04	0.14	0.24	0.00	0.15	30.00
		10	10	0	20	0.50	20.00	0.50	0.71	0.50	4.07	0.54	0.50	0.44	0.46	0.57	19.00
	2	15	5	5	15	0.50	20.00	0.50	0.75	0.67	3.37	0.69	0.55	0.57	0.50	0.62	55.00
		20	0	10	10	0.50	20.00	0.50	0.82	1.00	4.92	0.93	0.60	0.68	0.54	0.68	94.00
		5	15	5	15	0.00	0.00	0.00	0.35	0.00	-0.47	-0.09	0.05	-0.08	0.00	0.05	-6.00
		10	10	10	10	0.00	0.00	0.00	0.50	0.00	0.00	-0.03	0.10	0.09	0.00	0.10	16.00
	3	15	5	15	5	0.00	0.00	0.00	0.61	0.00	0.47	0.04	0.14	0.24	0.00	0.15	39.00
		0	20	10	10	-0.50	-20.00	-0.50	-0.82	-1.00	-4.92	-0.94	-0.41	-0.65	-0.46	-0.47	-73.00
		5	15	15	5	-0.50	-20.00	-0.50	-0.75	-0.67	-3.37	-0.67	-0.36	-0.43	-0.50	-0.42	-28.00
		10	10	20	0	-0.50	-20.00	-0.50	-0.71	-0.50	-4.07	-0.47	-0.31	-0.24	-0.54	-0.37	-15.00
White (2003a)	2	4	0	10	10	0.50	4.00	0.17	0.53	1.00	1.61	0.74	0.60	0.26	0.54	0.68	40.00
		10	10	0	4	0.50	4.00	0.17	0.71	0.50	1.07	0.47	0.50	0.25	0.46	0.57	47.00
		8	0	8	8	0.50	8.00	0.33	0.71	1.00	2.98	0.84	0.60	0.50	0.54	0.68	60.00
		8	8	0	8	0.50	8.00	0.33	0.71	0.50	2.34	0.53	0.50	0.34	0.46	0.57	47.00
	6	12	0	6	6	0.50	12.00	0.50	0.82	1.00	3.84	0.88	0.60	0.68	0.54	0.68	74.00
		6	6	0	12	0.50	12.00	0.50	0.71	0.50	3.14	0.56	0.50	0.44	0.46	0.57	55.00
		12	4	4	4	0.25	8.00	0.33	0.75	0.50	0.54	0.35	0.35	0.49	0.26	0.39	61.00
		4	4	4	12	0.25	8.00	0.33	0.50	0.33	0.14	0.42	0.30	0.31	0.24	0.34	43.00
	6	6	12	2	4	0.00	-4.00	-0.17	0.50	0.00	-1.16	0.16	0.06	-0.10	0.00	0.07	31.00
		4	2	12	6	0.00	-4.00	-0.17	0.41	0.00	-0.48	0.19	0.13	-0.04	0.00	0.14	35.00
		2	4	6	12	0.00	4.00	0.17	0.29	0.00	-0.93	0.23	0.06	0.10	0.00	0.07	29.00
		12	6	4	2	0.00	4.00	0.17	0.71	0.00	-0.50	0.17	0.13	0.36	0.00	0.14	54.00

(次ページに続く)

Table 8-1 (続き)

メタ分析に使用したデータおよび数理モデルの予測値

Study	Experiment	Covariation information				Non-parameterized models						Parameterized models				Rating			
		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	$\Delta P$	$\Delta D$	<i>pCI</i>	<i>H</i>	causal power	causal support	SS power	w $\Delta P$	EI	RW		Pearce		
Buehner et al. (2003) (Generative component)	1	8	0	0	8	1.00	16.00	1.00	1.00	1.00	8.02	0.93	1.00	1.00	1.00	1.15	89.00		
		8	0	2	6	0.75	12.00	0.75	0.89	1.00	4.50	0.91	0.80	0.84	0.78	0.91	77.00		
		6	2	0	8	0.75	12.00	0.75	0.87	0.75	4.21	0.77	0.75	0.74	0.72	0.86	69.00		
		8	0	4	4	0.50	8.00	0.50	0.82	1.00	2.48	0.80	0.60	0.68	0.54	0.68	71.00		
		6	2	2	6	0.50	8.00	0.50	0.75	0.67	1.59	0.68	0.55	0.57	0.50	0.62	54.00		
		4	4	0	8	0.50	8.00	0.50	0.71	0.50	1.79	0.58	0.50	0.44	0.46	0.57	57.00		
		8	0	6	2	0.25	4.00	0.25	0.76	1.00	1.17	0.40	0.39	0.52	0.28	0.44	58.00		
		6	2	4	4	0.25	4.00	0.25	0.67	0.50	0.38	0.38	0.35	0.40	0.26	0.39	47.00		
		4	4	2	6	0.25	4.00	0.25	0.58	0.33	0.00	0.40	0.30	0.27	0.24	0.34	46.00		
		2	6	0	8	0.25	4.00	0.25	0.50	0.25	-0.12	0.34	0.25	0.11	0.22	0.29	34.00		
		8	0	8	0	0.00	0.00	0.00	0.71	NA	0.31	0.22	0.19	0.37	0.00	0.21	41.00		
		6	2	6	2	0.00	0.00	0.00	0.61	0.00	-0.31	0.19	0.14	0.24	0.00	0.15	43.00		
		4	4	4	4	0.00	0.00	0.00	0.50	0.00	-0.70	0.20	0.10	0.09	0.00	0.10	37.00		
		2	6	2	6	0.00	0.00	0.00	0.35	0.00	-1.15	0.20	0.05	-0.08	0.00	0.05	29.00		
		0	8	0	8	0.00	0.00	0.00	NA	0.00	-2.19	0.12	0.00	-0.27	0.00	0.00	9.00		
		(Preventive component)	1	0	8	8	0	-1.00	-16.00	-1.00	-1.00	-1.00	-8.02	-0.92	-0.81	-1.00	-1.00	-0.94	-94.00
				0	8	6	2	-0.75	-12.00	-0.75	-0.89	-1.00	-4.49	-0.91	-0.61	-0.83	-0.72	-0.70	-85.00
				2	6	8	0	-0.75	-12.00	-0.75	-0.87	-0.75	-4.20	-0.73	-0.56	-0.59	-0.78	-0.65	-72.00
				0	8	4	4	-0.50	-8.00	-0.50	-0.82	-1.00	-2.47	-0.88	-0.41	-0.65	-0.46	-0.47	-79.00
				2	6	6	2	-0.50	-8.00	-0.50	-0.75	-0.67	-1.60	-0.63	-0.36	-0.43	-0.50	-0.42	-65.00
				4	4	8	0	-0.50	-8.00	-0.50	-0.71	-0.50	-1.79	-0.43	-0.31	-0.24	-0.54	-0.37	-46.00
				0	8	2	6	-0.25	-4.00	-0.25	-0.76	-1.00	-1.17	-0.80	-0.20	-0.46	-0.22	-0.23	-72.00
				2	6	4	4	-0.25	-4.00	-0.25	-0.67	-0.50	-0.38	-0.48	-0.16	-0.25	-0.24	-0.18	-59.00
				4	4	6	2	-0.25	-4.00	-0.25	-0.58	-0.33	-0.01	-0.29	-0.11	-0.07	-0.26	-0.13	-43.00
6	2			8	0	-0.25	-4.00	-0.25	-0.50	-0.25	0.13	-0.19	-0.06	0.08	-0.28	-0.08	-21.00		
0	8			0	8	0.00	0.00	0.00	-0.71	NA	-0.31	-0.62	0.00	-0.27	0.00	0.00	-45.00		
2	6			2	6	0.00	0.00	0.00	-0.61	0.00	0.31	-0.34	0.05	-0.08	0.00	0.05	-48.00		
4	4			4	4	0.00	0.00	0.00	-0.50	0.00	0.71	-0.21	0.10	0.09	0.00	0.10	-34.00		
6	2			6	2	0.00	0.00	0.00	-0.35	0.00	1.14	-0.14	0.14	0.24	0.00	0.15	-23.00		
8	0			8	0	0.00	0.00	0.00	NA	0.00	2.21	-0.07	0.19	0.37	0.00	0.21	-9.00		
Wasserman et al. (1996)	1			6	6	6	6	0.00	0.00	0.00	0.50	0.00	0.00	-0.03	0.10	0.09	0.00	3.00	
				3	3	9	9	0.00	0.00	0.00	0.35	0.00	0.00	-0.02	0.10	0.02	0.00	0.10	-1.00
				3	9	3	9	0.00	0.00	0.00	0.35	0.00	-0.44	-0.11	0.05	-0.08	0.00	0.05	-14.00
				3	3	9	3	0.00	0.00	0.00	0.61	0.00	0.45	0.03	0.14	0.24	0.00	0.15	27.00
				9	9	3	3	0.00	0.00	0.00	0.61	0.00	0.00	-0.06	0.10	0.14	0.00	0.10	8.00
				3	5	6	10	0.00	2.00	0.08	0.35	0.00	-0.18	-0.04	0.07	0.05	0.00	0.08	-4.00
				3	6	5	10	0.00	2.00	0.08	0.35	0.00	-0.26	-0.06	0.06	0.03	0.00	0.07	-7.00
				5	3	10	6	0.00	-2.00	-0.08	0.46	0.00	0.19	0.00	0.12	0.05	0.00	0.13	1.00
				5	10	3	6	0.00	-2.00	-0.08	0.46	0.00	-0.31	-0.10	0.06	-0.06	0.00	0.07	-17.00
		6	3	10	5	0.00	-2.00	-0.08	0.50	0.00	0.26	0.01	0.13	0.09	0.00	0.14	-1.00		
		6	10	3	5	0.00	-2.00	-0.08	0.50	0.00	-0.23	-0.09	0.07	-0.02	0.00	0.08	-11.00		
		10	5	6	3	0.00	2.00	0.08	0.65	0.00	0.31	0.00	0.13	0.28	0.00	0.14	30.00		
		10	6	5	3	0.00	2.00	0.08	0.65	0.00	0.23	-0.01	0.12	0.27	0.00	0.13	29.00		
		2	6	6	6	-0.25	-4.00	-0.20	-0.61	-0.50	-0.98	-0.38	-0.16	-0.23	-0.24	-0.18	-27.00		
		6	2	6	6	0.25	4.00	0.20	0.61	0.50	0.98	0.25	0.35	0.33	0.26	0.39	30.00		
		6	6	2	6	0.25	4.00	0.20	0.61	0.33	0.64	0.23	0.30	0.24	0.24	0.34	18.00		
		6	6	6	2	-0.25	-4.00	-0.20	-0.61	-0.33	-0.64	-0.21	-0.11	-0.01	-0.26	-0.13	-6.00		
		18	6	6	6	0.25	12.00	0.33	0.75	0.50	1.47	0.26	0.35	0.49	0.26	0.39	60.00		
		6	18	6	6	-0.25	-12.00	-0.33	-0.75	-0.50	-1.46	-0.44	-0.16	-0.29	-0.24	-0.18	-52.00		
		6	6	18	6	-0.25	-12.00	-0.33	-0.50	-0.33	-0.93	-0.22	-0.11	-0.20	-0.26	-0.13	-22.00		
		6	6	6	18	0.25	12.00	0.33	0.50	0.33	0.94	0.32	0.30	0.31	0.24	0.34	23.00		
		Proportion variance accounted for ( $R^2$ )					0.80	0.55	0.81	0.85	0.86	0.43	0.91	0.85	0.89	0.80	0.85		

ルールと EI ルールの違いとしてフリーパラメータの有無が挙げられるが、フリーパラメータを入れることによって適合度が上昇していることが読み取れる。同様の傾向は  $\Delta P$  ルールと Rescorla-Wagner モデルの比較や、 $\Delta P$  ルールと重み付き  $\Delta P$  ルールの比較においても示されている。

#### 8. 4. 考察

本研究の目的は、これまでに実施された研究に対してメタ分析を行い、規則ベースアプローチ・連合的アプローチ・因果的アプローチのそれぞれの数理モデルについて比較検討を行うことであった。シミュレーション実験の結果、因果ベイズネットを使用したモデルの1つであるSSパワーモデルが最もデータとの適合度が高いことが示され、その妥当性が示唆された。SSパワーモデルは事前分布の形状に関する定数を含むが、複数のフリーパラメータを含むモデルよりも当てはまりが良く、優れたモデルであることが示された。また、3つのセルの生起頻度のみから計算可能な二要因ヒューリスティックモデルは最も単純なモデルであるが、他のモデルに引けを取らない適合度を示していた。合理的なベイズモデルと単純なヒューリスティクスがともに高い適合度となっていたことは、ヒューリスティクスがベイズ推論の近似として機能することを示唆している。

ベイズモデルの1つであるSSパワーモデルが評定値をうまく説明していたのに対して、同様のアプローチに属する因果サポートモデルの適合度は低かった。このことは、因果サポートモデルがモデル比較を行うことに起因していると考えられる。因果サポートは因果関係を想定するモデルと無関係を想定するモデルの比較によって計算されるが、本研究で収集した実験の中には発生的因果関係と抑制的因果関係を同時に含む条件もあった。そのため、モデル比較が上手く機能せず、適合度の低下を招いたと考えられる。

本研究ではSSパワーモデルが一番高い適合度を示していたが、EIルールや二要因ヒューリスティックモデルといった他のモデルでも十分高い適合度が得られており、この結果だけでモデルの優劣の判断を下すことは十分ではない。今後はこれらのモデルの予測値が大きく異なる事態での実験的検討が望まれる。また、本実験では単一の原因事象と単一の結果事象によって構成される実験事態に限定して比較検討を行ったが、複数の原因事象で構成される実験事態などにおいても、これらのモデルの比較検討を行うべきである。

SS パワーモデルでは必要十分な原因への選好が事前分布に組み込まれている。これは、原因事象は単体で結果事象に強く影響するという因果の必要十分性に対する思い込みと言い換えることもできる。しかしながら、研究 1 では因果の必要性に基づく判断や因果の十分性に基づく判断が多く見られており、必要十分性に基づく判断はほとんど見られていなかった。因果構造の学習と因果強度の学習で実験事態が大きく異なるものの、因果構造の学習と同様に、どのような因果関係が想定されるかは状況によって変化することも考えられる。因果の必要性や十分性に関する情報を与えることがどのように影響するか検討することによって、SS パワーモデルが想定する事前分布が妥当かどうか明らかになるだろう。

## 9. 研究 4 因果強度の学習における介入の促進効果

### 9. 1. 序論

#### 9. 1. 1. はじめに

因果の知識は、自身の経験や他者からの教育など様々な方法によって獲得される。科学において多様な研究法が存在するように、自分自身で経験して因果関係を学習する場合でも、対象を繰り返し観察する、あるいは対象を直接操作してその効果を調べるなど様々な学習方法がある。研究 3 では因果強度の学習に関する様々なモデルの比較検討を行ったが、研究 4 では因果強度の学習における学習方法に焦点を当てる。具体的には、事象の状態を見て因果関係を学習する場合 (i.e., 観察) と事象を操作して因果関係を学習する場合 (i.e., 介入) とでは、どちらがより正確な学習をもたらすか検討する。

#### 9. 1. 2. 観察と介入

事象間の因果関係は観察や介入によって学習される。観察は変数の自動的な振る舞いを観測することに対して、介入は変数に何らかの操作を加えることである (Pearl, 2000; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003)。例えば、薬の投与と症状の改善の因果関係を考える場合、観察による学習では患者が薬を投与されるか否かとその結果を見て因果関係を判断することになる。一方、介入による学習では、患者に薬を投与するか否か意思決定を行い、選択の結果を見ることで因果関係を判断する。これら 2 つの学習方法は原因と結果の状態 (i.e., 生起もしくは非生起) に関する情報を得るという点で同じであるが、両者はその情報がどのようにして得られたかという点において大きく異なる。

観察と介入の差異については、実験参加者に事象間の因果の方向性を判断させる因果構造の学習課題で盛んに検討されてきた。Steyvers et al. (2003) は観

察による学習と介入による学習ではどちらがより正確に因果構造を推測できるか検討したところ、介入によって変数を操作した場合に判断がより正確になることが明らかになった。同様の結果は Lagnado and Sloman (2004) でも得られている。介入によって因果構造の判断が正確になることは、観察では区別できない因果モデルを区別できるようになることから説明されている。例えば、変数 X と変数 Y の共生起を観察していても、X が Y に影響しているか、Y が X に影響しているか判断することはできない。ここで変数 X を操作して何も生じなかった場合、X が Y に影響を及ぼしているという可能性は排除される。変数 Y を操作して変数 X の状態が変化することは、Y が X に影響しているという因果関係の証拠となる。観察では区別できない因果モデルはマルコフ等価モデルと呼ばれており、介入によってこれらのモデルの区別が可能になることから正確な判断が導かれる。また、Sobel and Kushnir (2006) は他者の介入を見て因果関係を学習するよりも、自分自身で介入を行って因果関係を学習する方が正確に因果構造を推測できることを示している。この結果は、自分自身で介入を行う際には自分の好きな順序で情報を獲得できるからだと解釈され、介入による学習における意思決定の重要性を示唆している。

観察と介入の差異は因果強度の学習課題ではあまり検討されていない。このことは 2 つの学習方法が離散試行手続きとフリーオペラント手続きという異なる実験手続きで研究されてきたことに起因する。離散試行手続きでは、試行は原因事象の提示期と結果事象の提示期に明確に区分される。例えば、Dickinson, Shanks, and Evenden (1984) はミサイルが戦車を破壊する効果を判断させる実験課題を作成している。各試行において実験参加者は戦車に対してミサイルを発射するか否かを選択し、戦車が爆発するかどうかを観察した。離散試行手続きの初期の研究 (e.g., Dickinson et al., 1984) では実験参加者が原因事象の状態を決めることができたが、原因事象と結果事象の生起頻度を統制するため、

後の研究 (e.g., DeHouwer & Beckers, 2002b) では実験参加者が原因の有無を選択できないようになっていた。フリーオペラント手続きは連続試行手続きとも呼ばれ (Anderson & Sheu, 1995), この手続きでは試行は定義されない。実験参加者は好きなときに反応すること (e.g., キー押し) が可能であり, 結果 (e.g., ランプの点灯) が一定の確率で提示される (e.g., Reed, 2001; Wasserman, Elek, Chatlosh, & Baker, 1993)。離散試行手続きでは観察による学習に用いられており, フリーオペラント手続きでは介入による学習が採用されているため, 両者を比較する試みはほとんどなされなかった。実験参加者の判断は離散試行手続きやフリーオペラント手続きといった課題にあまり影響されないと考えられていた (Allan, 1993)。

介入が因果強度の学習を促進するか否かについては一貫した知見が得られていない。Jenkins and Ward (1965) は, 自己関与の効果として自分でボタンを押してライトが点灯するか調べる場合 (i.e., 介入による学習) とただ単にそれらの情報が提示される場合 (i.e., 観察による学習) で判断に違いが生じるか検討したところ, 両者に統計的に有意な差異は認められなかった。同様の結果は Kelly and Athy (2009) でも報告されている。一方, Hattori and Oaksford (2007) は, 肥料の投与と植物の開花の因果関係を判断させる実験課題を用いて観察と介入の違いを検討し, 観察による学習では原因と結果が共に非生起の情報 (i.e.,  $2 \times 2$  の分割表におけるセル *d*) が無視されやすいことを示している。Hattori and Oaksford (2007) の結果は因果強度の学習における観察と介入の違いを示唆する貴重な知見であるものの, 彼らの知見は以下の 2 点で制限されている。第一に, 実験の目的がメタ分析のための基準を設けることであったため, 観察群の評定値と介入群の評定値との比較や, 評定値と因果パワー (Cheng, 1997) などの規範値との比較が行われていない。また, 各群の共変情報が別々にプログラムされているため, 観察群に提示される共変情報と介入群に提示さ

れる共変情報が異なっていた。そのため、観察による学習と介入による学習の違いが、原因の生起確率の上昇に伴って評定値が上昇する原因の密度効果 (e.g., Perales & Shanks, 2007) や結果の生起確率の上昇に伴って評定値が上昇する結果の密度効果 (e.g., Shanks, 1985a) によって生じた可能性も否定しきれない。提示情報を統制する手段として、ヨークトコントロール手続きが挙げられる。この手続きでは、介入群によって生成された共変情報が後に観察群に提示されるため、両群が受け取る情報は等しくなる。因果強度の学習における介入の役割について、ヨークトコントロール手続きを用いた検討はまだなされていない。

### 9. 1. 3. 本研究の目的

介入による学習を扱った研究から、人々が因果強度を査定する際に情報価の高い状況を選択可能なことが示唆されている (e.g., Barberia, Baetu, Sansa, & Baker, 2010; Green & Over, 2009)。Barberia et al., (2010) は、発生的原因の因果強度を調べる際には他の原因の影響の小さい、つまり結果の基準率の低い状況が選択されることを示している。このような状況では、標的原因の効果と他の原因の効果が混同されにくい。介入による学習には情報価の高い選択の機会が含まれているのに対して、観察による学習には選択の機会は含まれていない。ヨークトコントロール手続きによって両方で提示される情報は等しくなるが、共変動情報を望ましい順序で受け取れることは学習の促進効果として機能するだろう。また、介入による学習において実験参加者が変数を操作することによってより分析的になり、判断が正確になるかもしれない。これらのことから、介入によって因果強度の判断がより正確になると仮説を立てた。

本研究の目的は、介入が因果強度の推定を促進させるか否か検討することであった。実験 1 では、標的原因と文脈が独立に結果に影響する事態において実

験参加者にそれぞれの因果強度の判断を求めた。実験 2 では、観察と介入の違いを強調するため、標的原因と文脈に加えて別の原因事象が追加された。介入によって因果強度の判断が改善するならば、介入群の評定値は観察群よりも規範値に近い値をとることが予想される。

## 9. 2. 実験 1

因果学習における観察と介入の差異を検討するため、2 つの原因が共通の結果を独立に生起させる共通結果の因果モデルを使用した。一方の原因事象は実験参加者が操作可能な標的原因であり、他方の原因事象は標的原因の有無に関わらず常に存在している文脈であった (i.e.,  $P(\text{context}) = 1$ )。したがって、実験参加者が結果事象の生起を観察した時、それが標的原因によるものか、文脈によるものか判断できないようになっていた。実験のカバーストーリーでは、標的原因はミサイルに、文脈は地雷に、結果は戦車の爆発にそれぞれ対応していた。実験参加者はミサイルがどのくらい戦車を爆発させるか判断するように求められた。共通結果の因果モデルにおいては、標的原因の状態によって異なる情報がもたらされる。標的原因が存在している場合、実験参加者は標的原因と文脈が合わさった影響に関する情報を受け取る。一方、標的原因が存在しない場合、実験参加者は文脈の影響に関する情報だけを受け取ることになる。標的原因の因果強度を正確に判断するためには、文脈の影響を正しく把握しておく必要がある。介入群は知りたい情報を自分が望む順序で獲得することが可能であった。

因果パワー (causal power) とは、他の原因が存在しない事態において原因事象が結果事象を発生させる確率である。標的原因と文脈がそれぞれ独立に結果の生起に影響を及ぼすとき、標的原因の因果パワー  $q_{(\text{target cause})}$  は以下の式によって算出される。

$$q_{(target\ cause)} = \frac{P(effect | cause, context) - P(effect | \neg cause, context)}{1 - P(effect | \neg cause, context)} \quad (9-1)$$

この式において、 $P(effect|cause, context)$ は標的原因と文脈が存在する時に結果が生起する条件付き確率を、 $P(effect|\neg cause, context)$ は標的原因が存在せず、文脈のみが存在している時に結果が生起する条件付き確率を表している。一方、文脈の因果パワーが以下の式によって求められる。

$$q_{(context)} = P(effect | \neg cause, context ) \quad (9-2)$$

これらの指標は提示される情報から容易に算出可能であり、実験参加者には標的原因と文脈の因果パワーを推定することが求められた。観察と介入の違いを検討するために、ヨークトコントロール手続きを使用したことや共通結果の因果モデルを採用したことが本研究の特色であり、介入を行ったときの方が観察を行うときよりも正確な推定が可能になると予測した。

## 9. 2. 1. 方法

**実験参加者および実験計画** 大学生 24 名が実験に参加し、成績への加点となるコースクレジットを受け取った。実験参加者は介入群か観察群のいずれかに無作為に割り当てられた。

**教示** 実験参加者に印刷された教示を手渡し、実験者がその内容を読み上げた。教示の概要は次の通りであった。

以下に記されているストーリーの登場人物になることを想像してください。あなたはとある軍事施設で対戦車ミサイルの研究を行っています。現在の研究テーマは「新たに開発された対戦車ミサ

イルの性能」を検証することです。ここで言うミサイルの性能とは、ミサイルが戦車を破壊できるか否かということを表しています。ミサイルが戦車に命中すると戦車が爆発します。ただし、ミサイルは必ずしも戦車に命中するというわけではなく、ミサイルが発射されても戦車が爆発しないこともあります。また、狙撃対象の戦車は地雷原を走行しており、地雷によって爆発することもあります。したがって、ある戦車が爆発したとしても、それがミサイルによるものか、地雷によるものか断定することはできません。ミサイルが発射されずに戦車が爆発した場合には、それが地雷によるものだと判断することができます。ミサイルの効果調べるために、あなたはミサイルが戦車を爆発させるかどうか何度も観察することにしました。なお、戦車は毎回異なる場所を走行するので、一度地雷で爆発したからといって、その後、地雷で爆発することがなくなるというわけではありません。あなたの課題は、ミサイルを発射するかどうかを選択し（介入群）、ミサイルの有無と戦車の爆発の有無を観察し（観察群）、ミサイルと戦車の爆発の因果関係を判断することです。なお、この課題ではミサイルや戦車に関する知識は不要です。（残りの教示は学習フェイズとテストフェイズの進め方に関するものであった）

教示の大部分は介入群と観察群で共通していたが、下線部分に関しては両者で異なっていた。また、介入群は標的原因の操作を行うため、学習フェイズの説明に関しても観察群と若干異なったものになっていた。教示を読み終えた後、実験参加者が内容をきちんと理解しているかどうか確認した。

**学習フェイズ** 学習フェイズは 40 試行から成り、各試行ではミサイルと戦車の爆発についての情報が与えられた。実験参加者は観察群か介入群のいずれかに割り当てられ、ミサイルと戦車の状態を見て、それらの因果関係を学ぶように伝えられた。介入群の実験参加者はミサイルを発射するか否か選択することが求められた。学習フェイズ開始時には画面上に戦車と 2 つのボタンが表示されていた。一方のボタンはミサイルを発射することと、他方のボタンはミサイルを発射しないことと対応していた。実験参加者がどちらかのボタンを選択

した後、ミサイルが画面上に表示された。ミサイルが発射される場合にはミサイルのイラストが表示され、ミサイルが発射されない場合にはミサイルのイラスト上にバツマークが描かれた。ミサイルの表示と同時に“NEXT”と書かれたボタンが画面上に提示され、ボタンをクリックすると戦車の爆発の有無が伝えられた。戦車が爆発する場合には爆発のイラストが表示され、戦車が爆発しない場合には戦車のイラストが表示されたままであった。爆発の有無が提示されてから 1.5 秒後に最初の画面に移行した。

ほぼ同様の手続きが観察群の実験参加者にも適用されたが、ミサイル発射の選択を行わないという点で両者は異なっていた。最初にミサイルの状態と“NEXT”と書かれたボタンが表示され、ボタンをクリックすると爆発の有無が表示された。ミサイル発射の有無とその結果について観察群の各実験参加者が受け取る情報は、介入群の各実験参加者と対応するようになっていた (i.e., ヨークトコントロール手続き)。したがって、両群に提示される共変動情報は同一のものであった。実験参加者は .25-.25 条件・.75-.75 条件・.75-0 条件・.75-.25 条件という随伴性の異なる 4 つの条件を遂行した (Table 9-1 参照)。条件名の前半部分の数字はミサイルが発射された時に戦車が爆発する確率

Table 9-1

実験 1 の各条件における実際の確率と因果パワー

	Contingency conditions							
	.25-.25		.75-.75		.75-0		.75-.25	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
$P(\text{target cause})$	.56	.09	.49	.08	.54	.12	.54	.06
$P(\neg\text{target cause})$	.44	.09	.51	.08	.46	.12	.46	.06
$P(\text{effect} \text{target cause, context})$	.27	.02	.73	.03	.72	.02	.71	.03
$P(\text{effect} \neg\text{target cause, context})$	.29	.02	.74	.02	.00	.00	.29	.02
$q(\text{target cause})$	.00	.00	.03	.04	.72	.02	.60	.04
$q(\text{context})$	.29	.02	.74	.02	.00	.00	.29	.02

注) 条件名は  $P(\text{effect}|\text{target cause, context})$  と  $P(\text{effect}|\neg\text{target cause, context})$  をそれぞれ表す。因果パワーが負の値となる場合は 0 として計算した。

$P(\text{effect}|\text{target cause, context})$ を表しており、後半の部分の数字はミサイルが発射されない時に戦車が爆発する確率  $P(\text{effect}|\neg\text{target cause, context})$ を表していた。例えば.25-.25 条件では、ミサイルの発射の有無に関わらず 25%の割合で戦車が爆発した。このことは、ミサイルの効果がないこと [ $q(\text{target cause}) = 0$ ]および地雷に弱い効果があること [ $q(\text{context}) = .25$ ]を意味していた。同様に.75-.75 条件でもミサイルは戦車の爆発に影響を及ぼさなかった。ただし、地雷が強い効果を持っていたため [ $q(\text{context}) = .75$ ]、結果の密度は.25-.25 条件よりも高くなっていた。これらの条件とは対照的に.75-0 条件ではミサイルが強い効果を持っており、地雷には効果がなかった [i.e.,  $q(\text{target cause}) = .75, q(\text{context}) = 0$ ]。また.75-.25 条件では、ミサイルと地雷がともに戦車の爆発を引き起こしていた [i.e.,  $q(\text{target cause}) = .75, q(\text{context}) = .25$ ]。上記のように、それぞれの随伴性条件ではミサイルの因果強度と地雷の因果強度が異なっていた。実験参加者が行う選択は実験者が完全に制御することはできないため、実際の随伴性は実験参加者ごとに僅かに異なるものになっていた。実際の随伴性のバラつきを抑えるため、結果の状態の 90%はプログラムで設定した値に収束するように決定され、残りの 10%はランダムに決定された。随伴性条件の順序は実験参加者間でカウンターバランスを行った。

**テストフェイズ** 学習フェイズの 40 試行終了後、実験参加者には標的原因であるミサイルの因果強度と文脈として提示された地雷の因果強度を推定することが求められた。「ミサイルはどのくらい戦車を爆発させますか？」という文章とともに、Visual Analog Scale (0 : 全く爆発させない～100 : 必ず爆発させる)を画面上に表示し、反応を記録した。同様の方法で文脈である地雷についての評定も求めた。テストフェイズ終了後、同様の手順で次の随伴性条件が開始された。実験参加者にはそれぞれの条件がまったくの別物であることを伝えた。

## 9. 2. 2. 結果および考察

Figure 9-1 にそれぞれの随伴性条件における標的原因と文脈の平均評定値を示した。標的原因の評定値と文脈の評定値のそれぞれに対して、学習方法(2)×随伴性条件(4)の混合二要因分散分析を実施した。標的原因の評定値に対する分散分析の結果、随伴性条件の主効果は見られたが ( $F(3, 66) = 52.12, MSE = 252.31, p < .001, \eta^2_G = .58$ ), 学習方法の主効果は見られなかった ( $F(1, 22) = 0.67, MSE = 516.81, p = .420, \eta^2_G = .01$ )。学習方法と随伴性条件の交互作用は有意でなかった ( $F(3, 66) = 0.28, MSE = 252.31, p = .839, \eta^2_G = .01$ )。随伴性条件の主効果について多重比較を行ったところ、すべての条件間で有意差が確認された ( $ps < .05$ )。このことは、観察による学習と介入による学習のいずれの場合でも各随伴性条件における標的原因の因果強度を弁別していたことを示唆している。 .25-.25 条件や.75-.75 条件では標的原因は結果に対して効果がないにも関わらず [i.e.,  $q(\text{target cause}) = 0$ ], .75-.75 条件の評定値は.25-.25 条件の評定値よりも高くなっていた。このような傾向は、結果の生起確率が上昇すると因果強度の評定値も上昇するという結果の密度バイアス (e.g., Shanks, 1985)

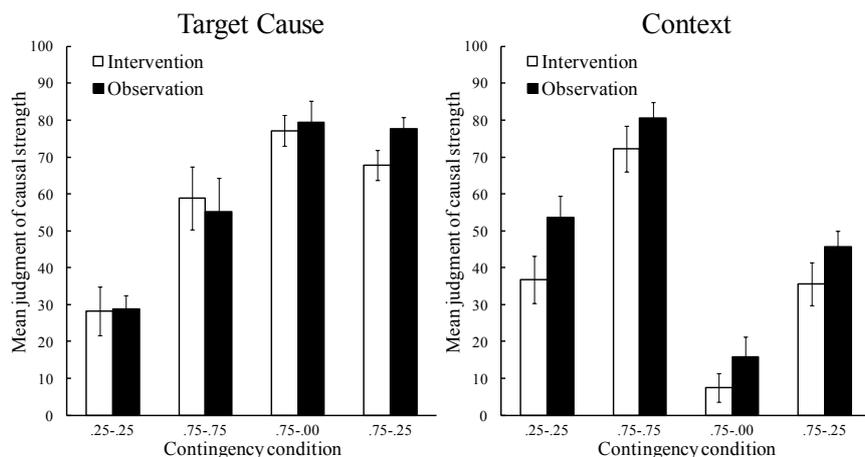


Figure 9-1. 実験 1 における標的原因(左図)と文脈(右図)の因果強度の平均評定値。エラーバーは平均の標準誤差を表す。

が生じていることを示している。

文脈の評定値に対して同様の分散分析を行ったところ、学習方法の主効果 ( $F(1, 22) = 5.00, MSE = 292.88, p = .035, \eta_G^2 = .07$ ) および随伴性の主効果 ( $F(3, 66) = 91.88, MSE = 220.24, p < .001, \eta_G^2 = .74$ ) が有意であった。学習方法と随伴性条件の交互作用は見られなかった ( $F(3, 66) = 0.92, MSE = 220.24, p = .435, \eta_G^2 = .03$ )。随伴性条件の主効果に関する多重比較の結果、.25-.25 条件と .75-.25 条件間を除く ( $p = .258$ )、すべての水準間で有意差が見られた ( $ps < .001$ )。文脈の因果パワーは標的原因が生起していない時に結果が生起する確率  $P(\text{effect}|\neg\text{target cause, context})$  から計算されるため (Table 9-1 参照)、介入による学習行った場合は観察による学習を行うよりも文脈の因果強度をより正確に判断していることが分かる。

介入による促進効果を詳細に検討するため、評定値と規範値である因果パワーの平均二乗誤差 (RMSE) を算出した。この値は評定値と因果パワーの差を表したものであり、値が小さいほど規範値に近い判断を行っていることを意味する。標的原因に関する判断では、介入群と観察群の間に大きな差は見られなかった (介入群:  $RMSE_{(\text{target cause})} = 31.51$ , 観察群:  $RMSE_{(\text{target cause})} = 33.09$ )。一方、文脈に関する判断においては、介入群の方が観察群よりも正確な推定を行っていることが明らかとなった (介入群:  $RMSE_{(\text{context})} = 13.19$ , 観察群:  $RMSE_{(\text{context})} = 18.90$ )。これらの結果は、介入によって文脈の因果強度がより正確に判断されることを示している。

実験 1 の結果は、介入が因果強度のより正確な推定を導くという仮説と一致するものであった。しかしながら、文脈の因果強度を低く見積もっていたことに対しては、介入によって他の原因の評価が下がったという別の解釈も可能である。実際、Kushnir and Gopnik (2005) は共変動情報が介入の結果と矛盾する場合に幼児が介入の結果を好むことを発見している。また、Kushnir,

Wellman, and Gelman (2009) は、他者の行為よりも自分の行為の影響を高く見積もる自己エージェントバイアスを報告している。実験 1 の結果に対しては介入群の実験参加者が自己エージェントバイアスによって文脈の因果強度を低く評価した可能性が残されている。実験 2 はこの可能性を排除する目的で行われた。

### 9. 3. 実験 2

実験 1 では、実験参加者が学習方法に関わらず標的原因に同様の判断を行うのに対して、文脈の因果強度については異なる判断を行っていることが示された。この結果は介入によって因果強度の正確な推定が促進されたものとして解釈された。しかしながら、介入の際に自己エージェントバイアスが生じ、介入対象以外の原因が過小評価されたという別の解釈も成立する。実験 2 では、これら 2 つの解釈について検討を行った。実験手続きは実験 1 と類似していたが、標的原因と文脈に加えて別原因が追加された。別の原因事象を追加することによって、介入は因果強度の推定を促進するのか、あるいは自己エージェントバイアスを生じさせて他の因果強度を低く見積もるのか、これら 2 つの解釈を分離することが可能になる。介入によって因果強度の判断が促進される場合、介入群における規範値からの逸脱は観察群よりも小さくなるはずである。一方、介入が他の原因を軽視させるだけならば、別原因と文脈に対して因果強度を低く見積もる結果となるはずである。

#### 9. 3. 1. 方法

**実験参加者および実験計画** 大学生 18 名が実験に参加し、成績への加点となるコースクレジットを受け取った。実験は 3 名まで同時に実施され、実験参加者はパーティションによって区切られたスペースにそれぞれ着席し、実験課題を遂行した。なお、この中に実験 1 の参加者は含まれていなかった。

**実験手続き** 手続きは標的原因と文脈に加えて別原因が追加されるという点を除いて実験 1 と同様であった。教示では、新たに開発されたミサイル (i.e., 標的原因) が戦車に及ぼす効果を調べるように伝えた。戦車の爆発は実験参加者が操作するミサイル (i.e., 標的原因) だけでなく、地雷 (i.e., 文脈) や別のミサイル (i.e., 別原因) によっても引き起こされることが教示された。

学習フェイズでは、40 試行を通して標的原因や別原因や文脈に関する情報が与えられた。文脈は常に生起しているのに対して (i.e.,  $P(\text{context}) = 1$ ), 別原因は半分の試行において生起していた (i.e.,  $P(\text{another cause}) = .5$ )。また、標的原因の生起確率は介入群の選択によって変化していた。介入群の実験参加者は別原因の状態を見てから標的原因の状態を選択することができた。一方、観察群の実験参加者は標的原因と別原因の状態が提示されるだけであった。それぞれの実験参加者は .75-.75-.75 条件・.25-.75-.25 条件・.75-.25-.25 条件・.75-.75-.25 条件という 4 つの随伴性条件を遂行した (Table 9-2)。条件名の一番目の数値は標的原因が存在するときに戦車が爆発する確率 (i.e.,  $P(\text{effect}|\text{target cause}, \neg\text{another cause}, \text{context})$ ) を、二番目の数値は別原因が存在するときに戦車が爆発する確率 (i.e.,  $P(\text{effect}|\neg\text{target cause}, \text{another cause}, \text{context})$ ) を、三番目の数値は標的原因

Table 9-2

実験 2 の各条件における因果パワー

Causal power	Contingency conditions			
	.75-.75-.75	.25-.75-.25	.75-.25-.25	.75-.75-.25
$q(\text{target cause})$	0	0	.67	.67
$q(\text{another cause})$	0	.67	0	.67
$q(\text{context})$	.75	.25	.25	.25

注) 条件名は  $P(\text{effect}|\text{target cause}, \neg\text{another cause}, \text{context})$  と  $P(\text{effect}|\neg\text{target cause}, \text{another cause}, \text{context})$  と  $P(\text{effect}|\neg\text{target cause}, \neg\text{another cause}, \text{context})$  をそれぞれ表す。

と別原因が両方存在するときに戦車が爆発する確率 (i.e.,  $P(\text{effect}|\neg\text{target cause}, \neg\text{another cause}, \text{context})$ ) をそれぞれ表している. 各随伴性条件では, 標的原因と別原因と文脈の因果強度が異なっていた. .75-.75-.75 条件では, 標的原因と別原因のどちらのミサイルも戦車の爆発に効果はなく [ $q_{(\text{target cause})} = 0, q_{(\text{another cause})} = 0$ ], 地雷が強い因果強度を有していた [ $q_{(\text{context})} = .75$ ]. .25-.75-.25 条件では別原因の因果強度 [ $q_{(\text{another cause})} = .67$ ] は標的原因の因果強度 [ $q_{(\text{target cause})} = 0$ ] よりも高く設定されていたが, .75-.25-.25 条件では標的原因の因果強度 [ $q_{(\text{target cause})} = .67$ ] の方が別原因の因果強度 [ $q_{(\text{another cause})} = 0$ ] よりも高くなっていた. これらの条件では, 文脈の因果強度は低く設定されていた [ $q_{(\text{context})} = .25$ ]. また, .75-.75-.25 条件では標的原因と別原因ともに高い因果強度となっていた [ $q_{(\text{target cause})} = .67, q_{(\text{another cause})} = .67, q_{(\text{context})} = .25$ ]. 標的原因と別原因のミサイルが同時に存在する場合に戦車が爆発する確率は noisy-OR 関数 (Pearl, 1988) を用いて以下のように計算した.

$$P(\text{effect} | \text{target cause}, \text{another cause}, \text{context}) = 1 - (1 - q_{(\text{target cause})})(1 - q_{(\text{another cause})})(1 - q_{(\text{context})}) \quad (9-3)$$

例えば, 標的原因と別原因がそれぞれ戦車の爆発に影響する.75-.75-.25 条件では, 両者が存在する時に結果が生起する確率は  $P(\text{effect}|\text{target cause}, \text{another cause}, \text{context}) = [1 - (1 - 2/3)(1 - 2/3)(1 - 1/4)] \approx .92$  として計算された.

テストフェイズでは, 実験 1 と同様に標的原因と別原因と文脈の因果強度を判断するように求めた. 続いて, 実験参加者は次の随伴性条件における学習フェイズとテストフェイズを遂行した. 随伴性条件の順序は Graeco-Latin 方格デザインを用いてカウンターバランスをとった.

### 9. 3. 2 結果および考察

Figure 9-2 にそれぞれの随伴性条件における標的原因と別原因と文脈の平均評定値を示した。標的原因の評定値と別原因の評定値と文脈の評定値のそれぞれに対して、学習方法(2)×随伴性条件(4)の混合二要因分散分析を実施した。標的原因の評定値に対する分散分析の結果、随伴性条件の主効果のみ有意であり ( $F(3, 48) = 13.26, MSE = 344.45, p < .001, \eta^2_G = .38$ ), 学習方法の主効果や学習方法と随伴性条件の交互作用は有意でなかった ( $F_s < 1$ )。実験 1 と同様に、介入群と観察群ともに標的原因の因果パワーが 0 である .75-.75-.75 条件において密度バイアスを示していた。実験参加者の評定値と因果パワーを比較した場合、.25-.75-.25 条件では介入群の方が規範値の 0 に近い判断を行っていたが、統計的な支持は得られなかった。

同様の分散分析を別原因の評定値に対しても行ったところ、随伴性条件の主効果が有意であった ( $F(3, 48) = 25.60, MSE = 240.38, p < .001, \eta^2_G = .51$ )。学習方法の主効果 ( $F(1, 16) = 1.17, MSE = 377.59, p = .296, \eta^2_G = .02$ ) や学習方法と随伴性条件の交互作用 ( $F < 1$ ) は見られなかった。実験参加者が自己エージェントバイアスを示していた場合、介入群の評定値は観察群よりも低くなるはずであるが、

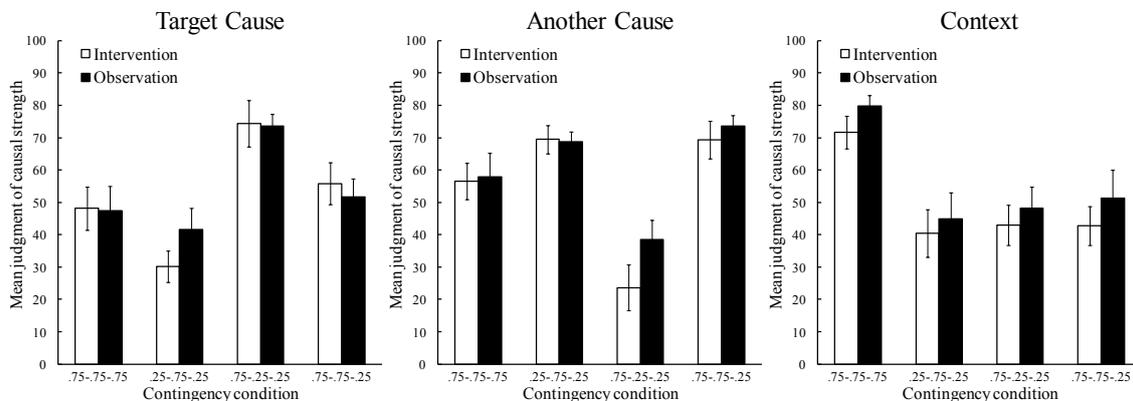


Figure 9-2. 実験 2 における標的原因(左図)と別原因(中央図)と文脈(右図)の因果強度の平均評定値。エラーバーは平均の標準誤差を表す。

.75-.25-.25 条件を除いて両群の評定値はほぼ等しかった。 .75-.25-.25 条件では別原因の因果パワーは 0 であるため、介入群の方が正確な推定を行っていたことが分かる。

文脈の評定値について学習方法(2)×随伴性条件(4)の混合二要因分散分析の結果、随伴性条件の主効果のみ有意であり ( $F(3, 48) = 13.56, MSE = 315.13, p < .001, \eta^2_G = .34$ ), 学習方法の主効果 ( $F(1, 16) = 1.27, MSE = 613.26, p = .276, \eta^2_G = .03$ ) や学習方法と随伴性条件の交互作用 ( $F < 1$ ) は有意でなかった。この結果は実験参加者が文脈の因果強度を適切に弁別できていたことを示している。

学習方法の差異を詳細に検討するために評定値と規範値である因果パワーの平均二乗誤差 (RMSE) を算出した。その結果、介入群における平均二乗誤差 ( $RMSE_{(target\ cause)} = 33.4, RMSE_{(another\ cause)} = 33.3, RMSE_{(context)} = 20.7$ ) は、いずれの判断においても観察群の平均二乗誤差 ( $RMSE_{(target\ cause)} = 35.4, RMSE_{(another\ cause)} = 36.7, RMSE_{(context)} = 26.4$ ) を下回っていた。これらの結果は、介入が因果強度のより正確な判断を導くという仮説を支持していた。

#### 9. 4. 総合論議

因果構造の学習では観察による学習よりも介入による学習の方が因果の方向性を正確に判断できることが数多く示されているのに対して (e.g., Steyvers et al., 2003; Sobel & Kushnir, 2006), 因果強度の学習において同様の促進効果がみられるか否かはきちんと確認されていない。本研究では 2 つの実験を通して介入が因果強度の正確な推定を導くという仮説を検証した。実験 1 の結果、標的原因については学習方法に関わらず類似した判断が見られたのに対して、文脈については介入を行った場合により正確な判断が見られた。実験 2 では、弱い因果強度に対して促進効果が生じやすいことが示された。これらの結果は、自身の行動を他の原因よりも重視する自己エージェントバイアスなどでは説明

困難であった。実験 1 の結果と実験 2 の結果はともに、介入によって因果強度の推定が促進されるという結論を支持している。

介入による促進効果については、いくつかの説明が可能である。第一に、どの情報を得ようとするかという意思決定が判断を促進したと考えられる。観察による学習とは異なり、介入による学習ではどのような介入を行うか決定しなければならず、この意思決定によって情報を好きな順序で受け取ることが可能になる。因果強度の判断において情報の提示順序が影響を及ぼすことは多くの研究によって示されている (Chapman, 1991; Dennis & Ahn, 2001; López, Shanks, Almaraz, & Fernández, 1998; Marsh & Ahn, 2006; Rottman & Ahn, 2011)。例えば、因果構造の学習においてランダムな順序で情報を受け取るよりも、構造化された順序で情報を受け取る方が正確な判断になることが知られている (Sobel & Kushnir, 2006)。毎試行において実験参加者にどのような介入を行うか指示するような手続き (e.g., Hannah & Beneteau, 2009) を用いれば、因果学習における意思決定の役割を検討することができるだろう。第二に、介入による学習が熟慮を促した可能性が挙げられる。観察による学習が事象の状態を見るだけなのに対して、介入による学習では介入の結果を見るだけでなく、その結果に基づいてどのような介入を行うか考えなければならない。適切な介入を行うためには、実験参加者はより分析的になり、これまでの情報を覚えておかなければならない。このような認知スタイルの変化が正確な判断を導いたかもしれない。科学的に振る舞うように教示された実験参加者は、より正確な判断を行うことが報告されている (Matute, 1996)。第三に、介入を行った場合には獲得した情報がより明瞭になり、個々のデータにより注意が向けられることも考えられる。情報の明瞭度や注意における差異が介入による学習を観察による学習と異なるものにしてしているかもしれない。介入による促進効果に対しては、単一のメカニズムを想定するよりも、意思決定や認知スタイルや明瞭

度や注意といった複数の要因の組み合わせを想定する方がより現実的である。

介入による促進効果に関して、いくつかの限界も確認された。第一に、実験参加者は学習方法に関わらず、結果の生起頻度が上昇すると因果強度も高く判断するという密度バイアスを示していた。具体的には、標的原因や別原因が結果の生起と無関係であるのに、その因果強度を過大評価していた。これは、結果の生起頻度が高い場合 (i.e., 天井効果) は原因と結果間の正の相関関係を否定する情報が得られにくいことに起因すると考えられる。第二に、介入による促進効果は因果強度が弱い場合においてのみ顕著であった。弱い因果パワーを正確に査定するためには、負の相関関係を示す情報 (i.e.,  $2 \times 2$  の分割表におけるセル *b* やセル *c*) に注目しなければならない。Maldonado, Jiménez, Herrera, Perales, and Catena (2006) は、偶発的な状況においてこれらの情報が無視されることを報告している。介入による学習ではこれらの情報がより考慮された結果、正確な判断が行われたかもしれない。介入による促進効果がなぜ生じたのか、また促進効果はどのような事態で生じるかについては更なる検討が必要である。

介入による因果強度の学習の促進効果は、因果性の学習と推論研究に理論的示唆を提供する。観察と介入の差異は、実験参加者に事象間の因果の方向性を推測させる因果構造の学習課題で検討されてきた。これらの実験では学習方法によって全く異なる判断が見られたことから、観察による学習 (i.e., Rottman & Keil, 2012) と介入による学習 (i.e., Fernbach & Sloman, 2009) には異なるモデルが提唱されてきた。対照的に、因果強度の学習事態では観察による学習と介入による学習に大きな違いはないと考えられてきた (e.g., Allan, 1993)。そのため、因果強度の学習に関する多くのモデルでは両者は明確に区別されていない (e.g., Cheng, 1997; Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004)。これらのモデルでは本研究の結果をそのまま説明することは難しい。いくつかの連合学習モデル (e.g., Rescorla & Wagner, 1972) では観察と介入の差異を刺激の明瞭度の差異として捉えることも可能

であるが、そのような試みでは過大な判断を説明することはできても、正確な判断が増えることを説明することは困難である。Hattori and Oaksford (2007) は因果学習にはヒューリスティックスと分析的思考の二段階の過程が存在することを提唱している。ヒューリスティックス段階では共変関係の検出に焦点が当てられ、将来に対する予測が主な目的となる。一方、分析的思考段階では因果関係の検出に焦点が当てられ、環境の制御が主な目的となる。これら2つの段階は観察と介入という2つの学習方法に対応することが想定されている。観察と介入の差異を示した本研究の結果はこのような二段階過程や推論の二重過程理論と一致するように思われる (e.g, Sloman, 1996, 2014)。

理論的含意に加えて、本研究は判断と意思決定の関連領域への示唆も含む。観察と介入の差異は因果学習事態だけでなく、関連領域でも検討されてきた。例えば、Yarritu, Matute, and Vadillo (2014) は制御幻想における介入の役割を検討している。制御幻想とは、自分の行動が結果を生じさせる確率を客観的確率よりも高く見積もる現象である (Langer, 1975)。彼らの結果は制御不可能な結果に対するコントロール感は学習方法によって変化しないことを示していた。Enkvist, Newell, Juslin, and Olsson (2006) は複数手がかり判断課題における介入の役割について検討している。実験の結果、連続変数を用いた事態では介入の促進効果が、二値変数を用いた事態では介入によってパフォーマンスの低下が生じていた。Osman and Speekenbrink (2012) は動的環境における手がかりと結果の関係を学習する課題において、観察による学習と介入による学習を比較した。その結果、学習方法に関わらず、手がかりと結果間の関係を学習できていたことが明らかとなった。これらの研究では互いに異なる結論が得られているが、本研究の結果も含め、判断と意思決定における介入の役割に関する累積的な知見となると考えられる。

本研究の結論として、能動的に介入を行って因果関係を学習した場合は受動的に観察して学習するよりも因果強度を正確に判断できることが示された。しかしながら、介入の促進効果は弱い因果関係に限定して生じていた。これらの知見は判断と意思決定

における観察と介入について新たな洞察をもたらしている。

第 4 部  
因果推論

## 10. 研究 5 因果推論における観察と介入<sup>14</sup>

### 10. 1. 序論

#### 10. 1. 1. はじめに

事象間の因果関係を学習することによって、過去の説明や将来の予測が可能になり、自身の行動とその結果の因果関係を学習することによって、適切な選択が可能になる。研究 4 では因果強度の学習事態における観察と介入の差異に焦点を当てたが、研究 5 では因果推論事態における両者の差異について検討を行う。具体的には、ある出来事が起こった場合 (i.e., 観察) とある出来事を生じさせた場合 (i.e., 介入) に、その原因や結果についてどのような推測がなされるか検討した。因果推論における観察と介入について、Rescorla-Wagner モデル (Rescorla & Wagner, 1972) などの連合学習理論を用いて説明を試みる連合的アプローチと、因果ベイズネット<sup>15</sup> (causal Bayes nets, Glymour, 2001; Pearl, 2000) を援用して説明を試みる因果的アプローチでは、その見解が大きく異なっている。

#### 10. 1. 2. 連合的アプローチと因果的アプローチ

連合的アプローチでは因果学習を事象間の連合強度の変化として捉えるのに対して、因果的アプローチでは因果モデルの構築およびパラメータの推定の問題として因果学習を扱う。事象の観察と介入に関して、これら 2 つのアプローチは大きく異なる。連合的アプローチでは観察による学習は古典的条件づけに、介入による学習はオペラント条件

---

<sup>14</sup>本研究の一部は 2007 年度関西学院大学文学部卒業論文として提出されており、本研究は以下の学術論文を修正・加筆したものである。

斎藤元幸・嶋崎恒雄 (2012). 因果推論に対する因果ベイズネットの適用の妥当性について—観察と介入の違いに関する実験的検討—. 認知科学, 19, 380-390.

<sup>15</sup> 因果ベイジアンネットワーク (causal Bayesian networks) と呼ばれることもある。

づけに対応する。両者の間には古典的条件づけがオペラント行動に影響するなどの相互作用 (e.g., Pavlovian-instrumental transfer, Estes & Skinner, 1941; Holland, 2004) が生じるとされており、ヒトを対象とした実験においてもその確認が報告されている (Talmi, Seymour, Dayan, & Dolan, 2008)。しかしながら、連合学習理論では観察から推測できることと介入から推測できることの違いを十分に説明することはできない。例えば、気圧計の値を観察することで雨の到来を予測できるが、気圧計の針を操作しても天候が変化することはない。なぜなら低気圧という共通原因によって擬似相関が生じているためである。連合学習理論によると、介入を行った経験がなければ介入の結果を予測することはできず、観察の経験から転移が生じると仮定した場合には気圧計の操作によって天候が変わるといった誤った予測が導かれてしまう。

一方、因果的アプローチでは操作を行った経験がなくても観察による学習だけで介入の結果を予測できると考える。元来、因果ベイズネットは因果関係を適切に扱うために統計学や計算機科学といった分野で開発された数理的手法であり、観察と介入の違いを的確に表現することが可能である。因果ベイズネットによると、観察 (observation) とは変数の自動的な振る舞いを観測することであり、介入 (intervention) とは変数に対して何らかの操作を加えることである (cf. Pearl, 2000; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003)。介入によって制御された変数は本来の原因からの影響を受けなくなる。すなわち、気圧計の針を操作している場合には、低気圧が到来しても気圧計の示す値が変化することはない。したがって、雨が降るか否かは気圧計の値とは無関係になり、低気圧が到来する確率や低気圧が雨をもたらす確率から推測される。

因果推論における観察と介入の違いは多くの研究によって示されてきた。Sloman and Lagnado (2005) は複数の事象が関連している事態で特定の事象に対して介入を行った場合に他の事象についてどのような推測がなされるか検討した結果、介入された事象が本来の原因と独立に扱われていることを示した。Waldmann and Hagmayer

(2005) は、事象の生起情報を表に記載した要約提示の因果学習課題を用いて、ヒトが観察による学習から介入の結果を正しく予測できることを報告している。同様の結果は、事象の生起情報を 1 試行毎に提示する事例提示の因果学習課題においても得られている (Meder, Hagmayer, & Waldmann, 2008, 2009)。近年では、ラットでさえも観察と介入を区別して因果推論を行っているということを示唆する研究も報告されている (Blaisdell, Sawa, Leising, & Waldmann, 2006; Leising, Wong, Waldmann, & Blaisdell, 2008; but see also Bonawitz, Ferranti, Saxe, Gopnik, Meltzoff, Woodward, & Schulz, 2010; Dwyer & Honey, 2009)。これらの結果はいずれも連合学習理論では説明不可能なものであり、因果ベイズネットの心理学的妥当性を示している。

### 10. 1. 3. 本研究の目的

連合的アプローチでは事象の生起情報は連合強度に変換されることが想定されており、観察の判断と介入の判断は形成された連合に基づいてなされるため、観察による学習を行った際にはその内容を介入の判断にも適用し、介入による学習を行った際にはその内容を観察の判断にも適用することが予想される (cf. Reips & Waldmann, 2008; Vadillo & Matute, 2007)。一方、因果的アプローチでは事象の生起情報は因果モデルに集約されることが想定されており、学習方法によって判断内容が変化することはない。観察による学習から介入の結果を予測できることは多くの研究によって示されているが (e.g., Meder et al., 2008)、介入による学習から観察の結果を予測できるか否かについてはまだ研究されておらず、因果ベイズネットの心理学的妥当性を検討する上では重要な問題と考えられる。本研究の実験 1 では Meder et al. (2008) と同様の手続きを用いて観察による学習から観察および介入の結果を予測できるか確認する。続く実験 2 では学習方法を介入による学習に変更し、介入による学習から観察および介入の結果を予測できるか検討する。

また、観察による学習を行う実験 1 の結果と介入による学習を行う実験 2 の結果を比較することにより、因果関係の学習方法の違いが因果推論の精度に及ぼす影響についても検討を行う。因果の方向性を判断する因果構造の学習事態では介入による学習の促進効果が多くの研究によって示されている(e.g., Lagnado & Sloman, 2004, 2006; Sobel & Kushnir, 2006; Steyvers et al., 2003)。ただし、2 値変数を対象とした複数手がかり判断課題においては介入によって不正確な判断が増大することが報告されており (Enkvist, Newell, Juslin & Olsson, 2006), 介入が必ずしも学習を促進させるとは限らない。実験間の比較を行うことにより、観察や介入の結果を推測する場面で介入による学習の効果が検討可能になる。

## 10. 2. 実験 1

### 10. 2. 1. 方法

**実験参加者および実験計画** 大学生 24 名が実験に参加した。実験は実験参加者内計画で行われ、推論の種類 (2; 観察, 介入) および推論の方向性 (2; 予測, 診断) ならびに事象の有無 (2; 生起, 非生起) の三要因が操作された。

**手続き** 実験課題は教示、学習フェイズ、テストフェイズによって構成されていた。教示では、架空のカバーストーリーが記された用紙が実験参加者に渡され、実験者によってその内容が読み上げられた。カバーストーリーはワインの樽内に存在する化学物質に関するものであり、カバーストーリーを通して、因果関係を学習することおよびそれらの関係について判断することを求めた。Figure 10-1 から確率の表記を省いたものを実験参加者に提示し、化学物質 A が化学物質 B と C を発生させ、それらが独立に化学物質 D を発生させることを伝えた。これらの因果関係は確率的であり、それぞれの樽における化学物質の有無を観察して因果関係の強さを学習するよう指示した。また、ワインの樽の中には上記以外の化学物質が含まれているかもしれず、それによって上記の化学物質が発生する可能性があることも伝えた。

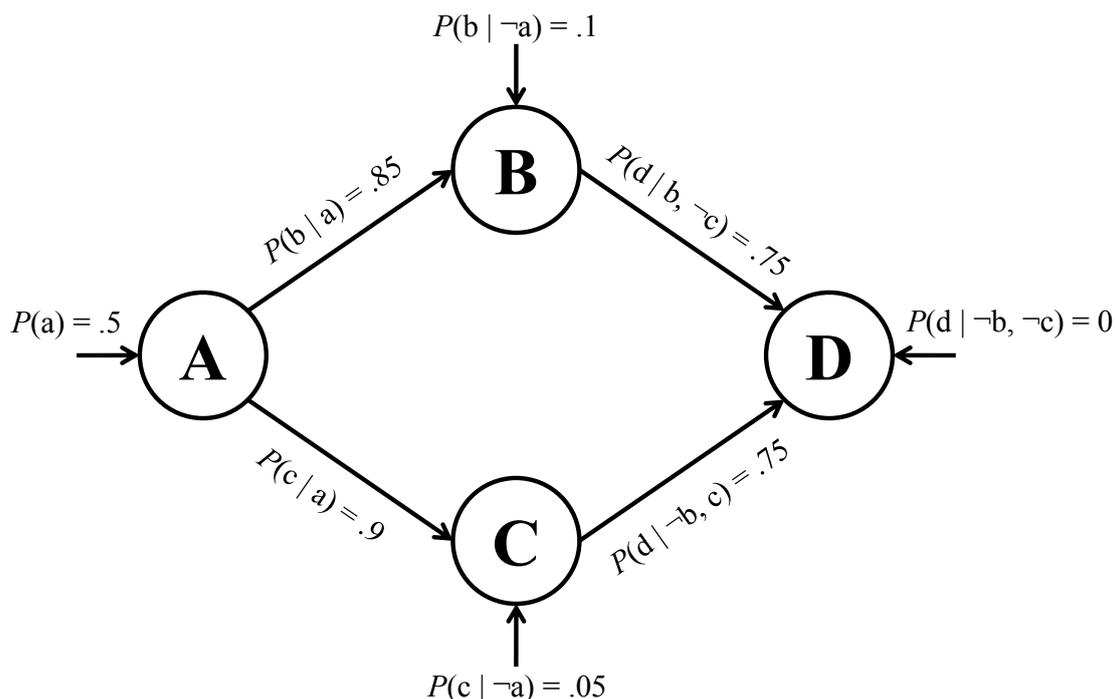


Figure 10-1. 実験 1 と実験 2 で用いた因果モデル.

学習フェイズは 40 試行から成り、各試行では 4 種類の化学物質の生起および非生起を観察することが求められた。試行開始時には NEXT と書かれたボタンが画面上に提示されていた。ボタンをクリックするとそれぞれの化学物質を表す異なる色の円にクエスチョンマークが付随して表示され、これらは化学物質の有無がまだ不明であることを示していた。その後、化学物質の状態が順番に明らかとなった。化学物質が生起している場合はクエスチョンマークが化学物質名 (e.g., A) に変わり、非生起の場合には円上の化学物質名に大きくバツマークが描かれた。画面上での化学物質の配置は Figure 10-1 と同様であり、因果関係と一致する順序で化学物質の状態が判明した。これらの順序は 1s の遅延を挿入することによって作られ、全ての化学物質の状態が示されてから 2s 後に最初の画面に移行した。Table 10-1 に学習フェイズにおける事象の生起情報を示した。また、学習フェイズにおける試行の順序は実験参加者ごとにランダム化した。

テストフェイズは観察や介入に関する 8 つの質問項目から構成されていた。観察に関

Table 10-1

実験 1 と実験 2 における事象の生起情報

	Experiment 1	Experiment 2
a, b, c, d	14	12.33 (2.61)
a, b, c, ¬d	1	2.17 (1.07)
a, ¬b, c, d	2	2.67 (1.25)
a, ¬b, c, ¬d	1	0.67 (0.80)
a, b, ¬c, d	2	2.42 (1.38)
a, b, ¬c, ¬d	0	0.54 (0.71)
a, ¬b, ¬c, d	0	0.00
a, ¬b, ¬c, ¬d	0	0.00
¬a, b, c, d	0	0.58 (0.64)
¬a, b, c, ¬d	0	0.00
¬a, ¬b, c, d	1	2.38 (1.38)
¬a, ¬b, c, ¬d	0	0.79 (0.87)
¬a, b, ¬c, d	1	3.13 (1.45)
¬a, b, ¬c, ¬d	1	1.13 (0.88)
¬a, ¬b, ¬c, d	0	1.63 (1.11)
¬a, ¬b, ¬c, ¬d	17	9.58 (2.84)
Σ	40	40

注) 数値は生起頻度の平均値と標準偏差をそれぞれ表す。

する質問では、化学物質 C の状態を確認した上で他の化学物質がどのくらい生起しているか推測することが求められた。状態は生起もしくは非生起の 2 通りであり、推測対象は化学物質 C の原因である化学物質 A (i.e., 結果から原因への診断的推論) と化学物質 C の結果である化学物質 D (i.e., 原因から結果への予測的推論) の 2 つであった。介入に関する質問でも同様に、化学物質 C の有無と推論の方向性によって 4 種類の質問が設けられた。例えば、介入によって化学物質 C が生起した時の化学物質 D について予測する場合、「新たなワインの樽 40 個に化学物質 C を加えました。この時化学物質 D はどのくらいの数の樽に存在していると思いますか？」という質問項目への評定が求められた。評定は「0:全ての樽において化学物質 D が存在しない」から「40:全ての樽において化学物質 D が存在する」の尺度上で行われた。観察に関する質問と介入に関する質問をそれぞれまとめ、どちらを先に行うかは実験参加者間でカウンターバランス

をとった。

## 10. 2. 2. 結果および考察

**結果の予測** 気圧計の例で示したように、観察に基づく予測と介入に基づく予測が一致するとは限らない。因果ベイズネットでは、ノードとリンクによって事象間の因果関係を表現する因果モデル (cf. Figure 10-1) を用いて両者の違いを説明する。事象に対して介入を行った場合、本来の原因がその事象に影響することはなくなり、本来の原因とその事象は独立になる。このことは、介入された事象に対するリンクを取り除くことで表現され、グラフ切除 (graph surgery) や無効化 (undoing) と呼ばれる (Pearl, 2000; Sloman, 2005)。例えば、気圧計の針を操作した場合、本来の原因である大気圧から気圧計に対するリンクが取り除かれる。事象に対して何らかの介入を行うことは  $do(\cdot)$  と記述される。Pearl (2000) によって導入されたこのような表記法は do オペレータと呼ばれ、観察や介入に基づく確率推論が定式化された。

診断的推論において、結果の有無は原因の生起確率を推測するための有力な証拠となり、その計算にはベイズの定理が用いられる。C の生起を観察することは A の生起確率を上昇させ、非生起である場合には A の生起確率は低下する [ $P(a|c) > P(a|\neg c)$ ]。対照的に、C に対して介入を行った場合には同様の推論は成り立たない。つまり、C の状態は介入によって規定されているため、その状態によって A の生起確率が変わることはない [ $P(a|do(c)) = P(a|do(\neg c))$ ]。このことは因果モデルにおける A から C へのリンクを取り除くことによって表現される。また、予測的推論では C の状態だけでなく、B の状態も考慮しなければならない。例えば、C の非生起を観察した場合には A の生起確率が低下し、それに伴って B の生起確率も低下する。一方、介入によって C が非生起となる場合、A は元々の基準率で生起して化学物質 B を発生させる。したがって、C の非生起を観察した場合の D の生起確率は、介入により C を非生起にする場合よりも小さくなる [ $P(d|\neg c) < P(d|do(\neg c))$ ]。

**診断的推論** 診断的推論と予測的推論における平均評定値と因果ベイズネットの予測値を Table 10-2 に示す. Holm の方法による事前比較の結果, 因果ベイズネットで予測されているように C の生起を観察した時の A の生起確率  $P(a|c)$  は C の非生起を観察した時の A の生起確率  $P(a|\neg c)$  よりも高く評定されているが ( $t(23) = 6.58, p < .001, d = 1.34$ ), 介入に基づく推論では C の生起に関わらず一貫した評定がみられた ( $t(23) = 0.61, p = .546, d = 0.13$ ). さらに,  $P(a|c)$  は  $P(a|do(c))$  よりも高く,  $P(a|\neg c)$  は  $P(a|do(\neg c))$  よりも低く評定されていることが示された ( $t(23) = 2.50, p = .020, d = 0.51$ ;  $t(23) = 3.91, p = .001, d = 0.80$ , respectively).

**予測的推論** 診断的推論の分析と同様に, Holm の方法による事前比較を用いて分析を行った. 因果ベイズネットの予測と一致して C の非生起を観察した場合の D の生起確率  $P(d|\neg c)$  は, 介入によって C を非生起にした場合の D の生起確率  $P(d|do(\neg c))$  よりも低く評定されており, その差は有意傾向であった ( $t(23) = 2.00, p = .058, d = 0.41$ ). C が生起している事態では観察と介入による評定値の違いはみられなかった ( $t(23) = 0.25, p = .806, d = 0.05$ ). また, 観察に基づく推論と介入に基づく推論のどちらにおいても C が生起している場合の方が, C の非生起時よりも D の生起確率が高く評定されていた ( $t(23) = 6.93, p < .001, d = 1.42$ ;  $t(23) = 6.35, p < .001, d = 1.30$ , respectively).

実験の結果, Meder et al. (2008) による先行研究と同様に, 観察による学習から

Table 10-2

実験 1 における評定値の平均と標準偏差および因果ベイズネットの予測値

Inference	Observation		Intervention	
	Presence	Absence	Presence	Absence
Diagnostic inference	$P(a c)$	$P(a \neg c)$	$P(a do(c))$	$P(a do(\neg c))$
<i>Predicted values</i>	38	4	20	20
<i>M</i>	28.82	16.30	22.30	23.70
<i>SD</i>	9.37	8.52	7.74	10.30
Predictive inference	$P(d c)$	$P(d \neg c)$	$P(d do(c))$	$P(d do(\neg c))$
<i>Predicted values</i>	36	5	33	14
<i>M</i>	28.80	12.13	29.23	15.43
<i>SD</i>	8.16	8.25	8.02	8.69

介入の結果を予測できることが示された。これらの結果は連合学習理論では説明困難なものであり、因果ベイズネットによる予測と一致するものである。因果ベイズネットの心理学的妥当性の更なる検討のために、実験 2 では介入による学習から観察の結果を予測できるか否かを検討する。連合学習理論では介入による学習から観察の結果を予測することはできないと考えるのに対して、因果ベイズネットでは観察による学習からの介入の予測だけでなく、介入による学習からの観察の予測も可能であると考ええる。

### 10. 3. 実験 2

実験 2 では、実験 1 とほぼ同様の手続きを用いて、介入による学習から観察の結果が予測できるか否かを検討する。また、因果ベイズネットの予測値との差異に基づき、観察による学習（実験 1）と介入による学習（実験 2）のどちらがより正確な推測を導くかに関しても検討を行う。

#### 10. 3. 1. 方法

**実験参加者および実験計画** 大学生 24 名が実験に参加した。実験 1 と同様に、実験は実験参加者内計画で行われ、推論の種類（2；観察，介入）および推論の方向性（2；予測，診断）ならびに事象の有無（2；生起，非生起）の三要因が操作された。

**手続き** 実験課題は実験 1 とほぼ同様であり、教示ではワインの樽内に存在する化学物質の因果関係を学ぶように指示した。実験 1 から変更が加えられた学習フェイズでは、実験参加者は化学物質の状態を操作することが求められた。操作は 4 種類の化学物質のいずれかを発生または抑制させることであり、各試行において画面上に提示された 8 種類のボタンの中から選択させた。発生させる場合にはその化学物質がワインの樽内に加えられ、抑制させる場合にはその化学物質が発生しなくなる薬品が投入されることを実験参加者に伝えた。操作の対象となった化学物質の状態は選択後すぐに表示さ

れ、同時に提示された NEXT と書かれたボタンを押すことによって他の化学物質の状態が表示された。他の化学物質の状態は実験 1 で用いた因果モデルのパラメータに基づいて決定され (cf. Figure 10-1), 全 40 試行の選択内容 (Table 10-3) および表示内容 (Table 10-1) が記録された。テストフェイズでは、観察や介入についての 8 つの質問に対する回答を求めた。

### 10. 3. 2. 結果および考察

**事象の操作** Table 10-3 に実験参加者の選択内容を示す。事象の種類(4)×操作の種類(2)の二要因分散分析の結果、事象の種類の主効果 ( $F(3, 69) = 2.88, MSE = 2.18, p = .042, \eta^2_G = .04$ ) および操作の種類の主効果 ( $F(1, 23) = 7.63, MSE = 6.82, p = .011, \eta^2_G = .11$ ) が有意であり、交互作用は有意でなかった ( $F(3, 69) = 2.24, MSE = 1.74, p = .091, \eta^2_G = .03$ )。事象の種類の効果について下位検定を行ったところ、原因事象 A と結果事象 D の操作回数の中に有意差が見られた ( $t(23) = 2.80, p = .010, d = 0.57$ )。結果事象よりも原因事象に多く介入するという結果は Steyvers et al. (2003) や Lagnado and Sloman (2006) と合致するものであり、推定可能になるパラメータ数の違いから説明できる。原因事象 A の操作とは異なり、結果事象 D を操作した場合には B から D への因果関係と C から D への因果関係についての情報は得られない。また、事象を抑制させた場合にはその結果事象に対する影響が

Table 10-3

実験 2 において各事象を発生または抑制した回数

Event	Generation	Prevention
A	5.46(1.47)	5.25(1.61)
B	5.63(1.47)	4.42(1.55)
C	5.75(1.53)	4.50(1.50)
D	5.25(1.51)	3.75(1.27)

注) 括弧内の数値は標準偏差を表す。

判断できなくなることから、抑制させるよりも発生させる回数の方が多いことも同様の理由で説明される。

**診断的推論** 評定値の平均および標準偏差と因果ベイズネットの予測値を Table 10-4 に示す。Holm の方法による事前比較を行ったところ、C の生起を観察している場合の方が C の非生起を観察している場合よりも A の生起確率を高く評定することが示された ( $t(23) = 5.22, p < .001, d = 1.07$ )。しかし、同様の傾向は介入に基づく推論においても見られており ( $t(23) = 2.60, p = .016, d = 0.53$ )、因果ベイズネットの予測と一致していない。また、C の生起に関しては観察と介入で評定値に差が見られたものの、非生起時にそのような差は見られなかった ( $t(23) = 3.16, p = .004, d = 0.65; t(23) = 1.39, p = .177, d = 0.28, respectively$ )。これらの結果から因果ベイズネットを支持する明確な証拠は得られなかったが、推論の種類と事象の有無による二要因の分散分析では交互作用が見られていることから ( $F(1, 23) = 9.15, MSE = 193.19, p = .025, \eta^2_G = .03$ )、何らかの形で観察と介入を区別していることが示唆される。

**予測的推論** Table 10-4 に示されている予測的推論における評定値について Holm の方法による事前比較を行った結果、因果ベイズネットの予測と異なり、C が非生起の事態において観察と介入の間に差は見られなかった ( $t(23) = 0.10, p = .919, d = 0.02$ )。実験 1 と同様に、C が生起している事態では観察と介入による評定値の違いはみられなかった ( $t(23) = 1.34, p = .194, d = 0.27$ )。また、観察と介入のどちらに

Table 10-4

実験 2 における評定値の平均と標準偏差および因果ベイズネットの予測値

Inference	Observation		Intervention	
	Presence	Absence	Presence	Absence
Diagnostic inference	$P(a   c)$	$P(a   \neg c)$	$P(a   do(c))$	$P(a   do(\neg c))$
Predicted values	37.27	1.52	21.00	21.00
M	29.25	16.12	24.72	18.45
SD	8.71	8.20	10.63	9.72
Predictive inference	$P(d   c)$	$P(d   \neg c)$	$P(d   do(c))$	$P(d   do(\neg c))$
Predicted values	37.56	6.05	35.19	18.39
M	28.27	18.17	27.10	18.02
SD	8.09	8.91	7.95	7.61

においても C が生起している場合の方が、C の非生起時よりも D の生起確率が高く評定されていた ( $t(23) = 4.55, p < .001, d = 0.93$ ;  $t(23) = 4.44, p < .001, d = 1.30$ , respectively). 予測的推論においては観察と介入を区別して推論を行っているという傾向はほとんど見られず、因果ベイズネットを支持するような証拠は得られなかった。

**実験 1 との比較** 実験 1 と実験 2 の結果に対する因果ベイズネットモデルの適合度を Table 10-5 に示す。数値は評定値と予測値の平均二乗誤差 (root-mean-squared-error) を表しており、この値が大きいほどモデルの予測から逸脱していることを意味している。分析の結果、観察に基づいて因果関係を学習した方が、介入による学習を行うよりも観察に基づく推論と介入に基づく推論の両者を正確に行っていることが明らかとなった。また、観察と介入どちらの学習方法においても観察に基づく推論よりも介入に基づく推論において誤差が小さいことが示された。

実験の結果、介入による学習から観察の結果を推測できることが示されたが、介入に関する推論において因果ベイズネットの予測と異なる評定も見られた。これらの結果は連合的アプローチでは説明不可能であり、因果的アプローチでも説明困難である。また、観察による学習を行った実験 1 と介入による学習を行った実験 2 の比較によって、観察に基づく学習を行った場合により正確な推測がなされることが明らかとなった。

#### 10. 4. 総合論議

本研究では、因果ベイズネットの心理学的妥当性を検討することを目的として、因果推論における観察と介入について実験を行った。実験 1 では Meder et al. (2008) と

Table 10-5

実験 1 および実験 2 における評定値と因果ベイズネットの予測値との平均二乗誤差

	Observational inference	Interventional inference	Total
Observational Learning(Exp.1)	12.58	9.23	11.04
Interventional Learning(Exp.2)	14.47	10.80	12.77

同様の手続きを用いて観察による学習から介入の結果を予測できるか否かを検討し、因果ベイズネットの予測と一致して、観察と介入を区別した予測が正しく行われていることが示された。この結果は連合学習理論では説明不可能なものであり、因果ベイズネットの心理学的妥当性を示唆するものであった。続く実験 2 においては、因果関係の学習方法を観察から介入に変更し、実験 1 と同様に観察と介入に関する推測を求めたところ、観察と介入を区別しているという十分な証拠は得られなかった。これらの結果から因果ベイズネットには一定の妥当性が認められるものの、その適用には慎重な議論が必要であることが示唆された。介入による学習を行った後に適切な推論が行われなかった理由として以下の 2 点が挙げられる。

介入によって因果関係を学習した場合には観察による学習の時とは異なる因果推論がなされた可能性が考えられる。具体的には、介入による学習後の判断では実際に経験した内容をそのまま推論に用いるということであり、これは連合学習理論から導かれる予測である。例えば、事象 C を発生させた時に事象 A が生起しており、事象 C を抑制した時に事象 A が非生起であることを経験している場合、C に対する操作によって A の生起確率の評定値が変化することは連合学習の立場から考えると尤もらしい結果である。しかしながら、実験参加者が実際に経験した確率と因果ベイズネットが予測する確率の間に大きな差は見られない ( $P(a | do(c))_{\text{experienced}} = .58$ ;  $P(a | do(\neg c))_{\text{experienced}} = .54$ ;  $P(a | do(c))_{\text{predicted}} = P(a | do(\neg c))_{\text{predicted}} = .53$ )。さらに、介入に基づく推論の評定値に関して実際に経験した頻度情報から算出した予測値と因果ベイズネットの予測値の適合度をそれぞれ算出したところ、因果ベイズネットのほうがよい当てはまりを示していた ( $RMSE = 13.53$ ;  $RMSE = 10.80$ , respectively)。これらのことを踏まえると、介入による学習後の推論が連合学習理論によって想定されているものと一致しているとは考えにくい。

因果推論に違いがないとすれば、因果関係を学習する段階で何らかの違いが生じたと考えられ、適切な推論が行われなかったという結果は介入による学習において因果関

係が十分に学習されなかったことに起因するかもしれない。つまり、不適切な因果推論の原因は不適切な因果学習にあるということである。本研究では、何が原因で何が結果であるかという因果構造は予め教示されているため、それぞれの因果関係の強さを学習する際に観察と介入で違いが生じたと思われる。介入では因果関係が正しく学習されないとすると、因果ベイズネットの予測値に対する適合度において観察を行った後の評定値のほうが介入を行った後の評定値よりも当てはまりがよかったことの説明がつく。ただし、研究 4 では、介入による学習の促進効果が示されており、それぞれの知見が矛盾しているように見える。実験で採用している因果モデルやテストフェイズの判断内容や細かな実験手続きが異なっているため、この不一致の原因を断定することはできないが、本研究の結果を考察するうえで研究 4 の知見は有用であると考えられる。以下、研究 4 の結果を交えながら、本研究における観察と介入の相違点について考察を進める。

第一に、学習フェイズにおける提示情報の差異が考えられる。Table 10-1 に示したように、介入による学習では事象の生起および非生起を操作できるため、観察による学習を行った場合とでは異なった情報が提示されることになる。介入される事象以外は特定の確率に従って生起することから提示情報が大きくことなることはないものの、提示されるパターンの増加によって課題が困難になった可能性がある<sup>16</sup>。研究 4 ではヨークトコントロール手続きを用いて観察による学習と介入による学習の提示情報を揃えており、そのような事態でも両者に違いが見られたことから、提示情報の違いだけで本研究の結果を説明することは難しい。

第二の相違点として事象の提示順序が挙げられる。観察によって学習する際には事象は因果の方向性と同じ順序で事象が提示されていたが、介入によって学習する際には介入された事象の生起または非生起が先に提示されていた。Meder et al. (2008) では事象の提示順序が因果の方向性と逆である場合には誤った判断が増加することが

---

<sup>16</sup> 観察による学習と介入による学習の提示情報を揃える方法として、ヨークトコントロール手続きが挙げられる。本実験の場合、介入による学習を行った際の提示情報を記録し、その提示情報を観察させることに相当する。

明らかにされており、介入する対象によって提示順序の一部が変更されたことが影響したかもしれない。

第三に、観察と介入の機能の違いが考えられる。観察とは異なり、介入が行われる時には操作された事象とその本来の原因は独立状態になることが想定されている (i.e., グラフ切除)。例えば、介入によって事象 C を発生させた場合に事象 A が生起していたとしても、このとき事象 A が事象 C に影響を及ぼしていたかは判断できない。介入の対象として選択した事象とその結果事象にのみ注目した結果、原因事象との因果関係の判断が不適切になったと考えられる。このような事態は 2 値事象に起因するものであり (cf. Beckers, De Houwer, Pineño, & Miller, 2005), 2 値事象から多値事象へと結果の最大性 (outcome maximality) を操作することによって、上記の場面においてどのような判断がなされているか検討可能になる。また、介入による学習においては実験参加者が選択した事象は介入の結果として、それ以外の事象は観察の結果として学習することが求められるため、観察と介入を混同してしまった可能性も考えられる。

観察と介入には様々な相違点が存在するが、どのような差異が介入後の因果推論を歪めたかについて本研究の結果から断定することはできない。本研究よりも単純な因果モデルを用いた上で上記の事柄を検討することによって因果学習における観察と介入の違いが明らかになるだろう。実際、研究 4 では 2 つの原因事象と 1 つの結果事象から成る共通結果の因果モデルを採用し、介入によって因果強度の判断が促進されることが明らかにされている。また、観察による学習を行った実験 1 の結果は因果ベイズネットを支持するものであるが、モデルの当てはまりが十分であるとは言えない評価項目も存在する。このような予測値からの逸脱は記憶容量の限度によって説明されることが多いが (e.g., Buehner, Cheng, & Clifford, 2003), どのような要因が適切な因果推論を妨害するかについては十分に明らかにされておらず、今後の重要な研究課題である。

## 11. 研究 6 意思決定における因果推論

### —因果モデルとパラメータの役割—<sup>17</sup>

#### 11. 1. 序論

##### 11. 1. 1. はじめに

因果の知識は適応的意思決定に必要不可欠である。因果性は質的側面と量的側面から構成されており、前者は因果関係の有無やその方向性を表す因果構造である。行為が望ましい結果と無関係な場合、その行為に従事することは無意味である。一方、後者は因果関係の強さを表す因果強度である。2つの選択肢が望ましい結果と結びついている場合、どちらの選択肢がより高い確率で結果をもたらすか判断しなければならない。因果の知識の欠如は迷信行動 (superstitious behavior) や制御幻想 (illusion of control) などの不合理な行為をもたらす。

研究 5 では、事象の状態が観察の結果から得られたものか、介入の結果から得られたものかによって、因果推論がどのように変化するか検討した。研究 6 では、これらの推論を意思決定場面に拡張して検討を行う。具体的には、行為の状態が観察の結果から得られた場合、介入の結果から得られた場合、選択の結果から得られた場合、それぞれについてどのような推論がなされるか検討した。

---

<sup>17</sup>本研究の一部は 2009 年度関西学院大学大学院文学研究科修士論文として提出されており、本研究は以下の学術論文を修正・加筆したものである。  
Saito, M., & Shimazaki, T. (2011). Causal reasoning in decision making: The role of causal models and their parameters. In B. Kokinov, A. Karmiloff-Smith, & N. J. Nersessian (Eds.), *European Perspectives on Cognitive Science*, New Bulgarian University Press.

## 11. 1. 2. 因果推論と意思決定

Johnson-Laird and Shafir (1993) によって推論と意思決定の密接な関係が指摘されたにも関わらず、両者の関係を扱った研究は少ない (Sloman & Hagmayer, 2006). 因果推論と意思決定の関係を扱った研究は大きく 2 つのアプローチに分けられる. 一方のアプローチでは、因果推論が意思決定にどのような影響を及ぼすかが検討されている. ヒトは時間や知性や計算能力に限界があるという限定合理性の観点から、Gigerenzer, Todd, and ABC Research Group (1999) は最低限の情報のみを使用して意思決定を行う高速・儉約ヒューリスティクス (fast and frugal heuristics) を提唱している. Garcia-Retamero and Hoffrage (2006) は、高速・儉約ヒューリスティクスの 1 つである最善選択ヒューリスティクス (take-the-best heuristics) における因果の知識の役割について論じた. Garcia-Retamero, Wallin, and Dieckmann (2007) は、手がかり妥当性を学習する課題において、因果的知識が妥当な手がかりの発見を促進させることを明らかにした. Garcia-Retamero, Müller, Catena, and Maldonado (2009) は、因果的信念と経験的証拠の両方が意思決定に影響を及ぼしており、どちらを重視するかは事前の訓練や手がかり妥当性に依存することを示した. 事前の訓練がない場合、毎試行のフィードバックが与えられたにもかかわらず、実験参加者は妥当性の高い手がかりよりも因果的信念と合致する妥当性の低い手がかりを選好していた. これらの結果は、因果の知識が適切な意思決定を促進するだけでなく、阻害することもあることを示唆している (see Garcia-Retamero, Hoffrage, Müller, & Maldonado, 2010). ヒューリスティックによる意思決定ではなく、熟慮に基づく意思決定を対象とした研究も存在する. Hagmayer and Sloman (2005) は因果関係に基づいて意思決定が行われることを示した. Hagmayer and Sloman (2006) は Newcomb のパラドックス (Nozick, 1969) を題材とした意思決定課題においても因果関係に基づく意思決定が同様にみられることを報告している. Nichols and Danks (2007) は因果強度を学習した上で適切な意思決定が行われることを明らかにした.

意思決定に焦点を当てた上記の研究とは異なり、他方のアプローチでは意思決定の際にどのような因果推論がなされているかが検討されている。Hagmayer and Meder (2008) および Meder and Hagmayer (2009) は、繰り返しの意思決定課題において因果の知識を基に意思決定が行われ、意思決定の結果から因果構造を学習していることを実証した。また、Hagmayer, Meder, Osman, Mangold, and Lagnado (2010) は、動的システムをコントロールする課題において、実験参加者が自発的にシステム内の因果関係を推測していたことを示した。Sloman and Hagmayer (2006) および Hagmayer and Sloman (2009) は因果推論と意思決定を包括的に説明する理論として選択の因果モデル理論 (causal model theory of choice) を提唱している。

### 11. 1. 3. 選択の因果モデル理論

選択の因果モデル理論 (Sloman & Hagmayer, 2006; Hagmayer & Sloman, 2009) は因果推論と意思決定を 1 つの枠組みにまとめたものであり、因果ベイズネットがその核となっている。また、意思決定による選択を特別視しており、意思決定者とその観察者では抱いている因果モデルが異なることを想定している。ここでは、因果ベイズネットの簡単な説明を行った後、観察者および意思決定者の観点から意思決定を考察し、意思決定の過程について解説を行う。

**因果ベイズネット** 因果ベイズネット (e.g., Glymour, 2001; Pearl, 2000; Sloman, 2005) は事象間の因果関係を因果モデル (causal model) によって表現して推論を行うための数理的な手法である。因果モデルにおいて事象間の因果関係は非循環有向グラフで、因果関係の強さは条件付き確率を用いて表現される。それぞれの事象はノードと言われ、エッジと呼ばれる矢印によって結びつけられている。非循環とはノードからエッジを辿って元のノードに戻らないことであり、循環論に陥るようなことはない。因果ベイズネットでは因果関係と相関関係が明確に区別される。前者は原因事象から結果事象に対するエッジであり、後者はそれらのエッジによって生じるものである。

Figure 11-1 左上の因果モデルは運動靴の購入とスリムな体型がトレーニングによって生じていることを示している。因果関係が無い場合，運動靴を買ったところでスリムな体型になることはない。Figure 11-1 右上の因果モデルは階段の使用とスリムな体型の因果関係を示している。この場合，階段を使用することでスリムになることが予想される。

Figure 11-1 の例において運動靴を買うことや階段を使用することは事象の操作として捉えられ，因果ベイズネットでは新たなノードとエッジを挿入して表現する (Figure 11-1 左下および右下参照)。このような事象の操作は介入 (intervention) といわれ，事象の状態を確認する観察 (observation) とは明確に区別される。介入は外生的 (i.e., 介入対象の変数を除く全ての変数と独立) であり，介入された事象は本来の原因事象と独立になる。介入によってエッジが取り除かれることはグラフ切除 (graph surgery) や無効化 (undoing) と呼ばれる (Pearl, 2000; Sloman, 2005)。例えば，運動靴を買うように命令された場合，運動靴の購入とトレーニング間のエッジが取り除か

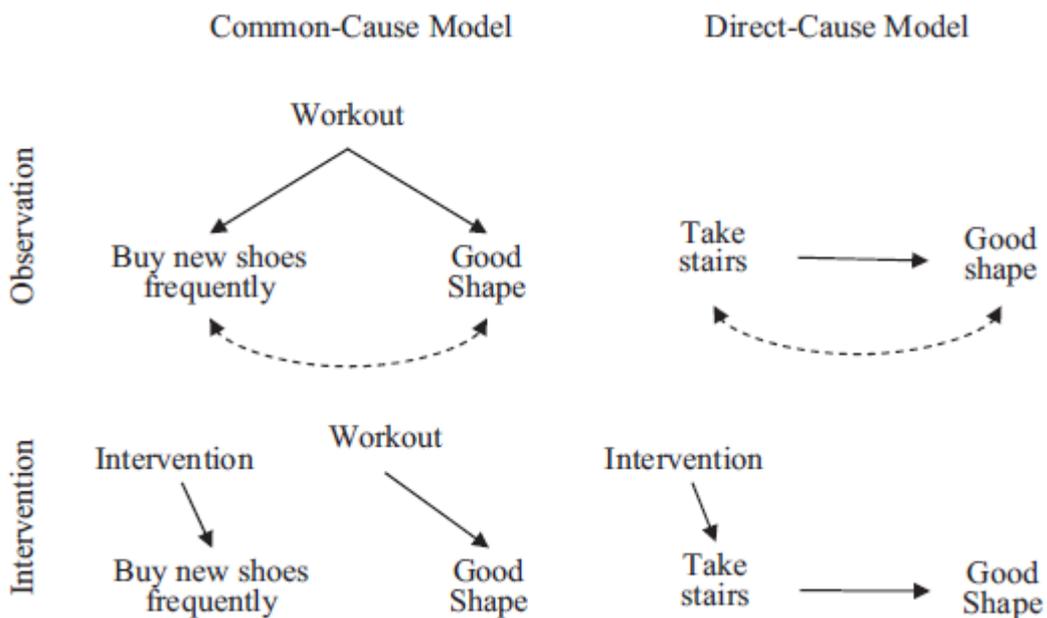


Figure 11-1. 行動と結果の関係を示す因果モデルにおける観察と介入。単方向の矢印は因果関係を，双方向の矢印は相関関係を表す。(Hagmayer & Sloman, 2009 より)

れ、その人がトレーニングを行っているか否かは判断できない (Figure 11-1 左下参照)。一方、階段を使うように命令された場合、エッジが取り除かれることはないため、スリムな体型になると予測できる (Figure 11-1 右下参照)。因果ベイズネットは介入の概念を利用することで、相関関係に基づく確率と事象の操作や行動による因果関係に基づく確率の区別を行っている。

**観察者の観点での意思決定** 因果ベイズネットと介入の概念を用いることで、意思決定の因果モデルを作成することが可能になる。体調を良くするために仕事に自転車で行くべきか否か熟慮している人を例として考えた場合、Figure 11-2 左側のような因果モデルが記述される。この因果モデルにおいて意思決定者の選択 (e.g., 仕事に自転車で行くかどうか) はその人自身の意思決定の過程 (e.g., 健康 vs. 早い帰宅)

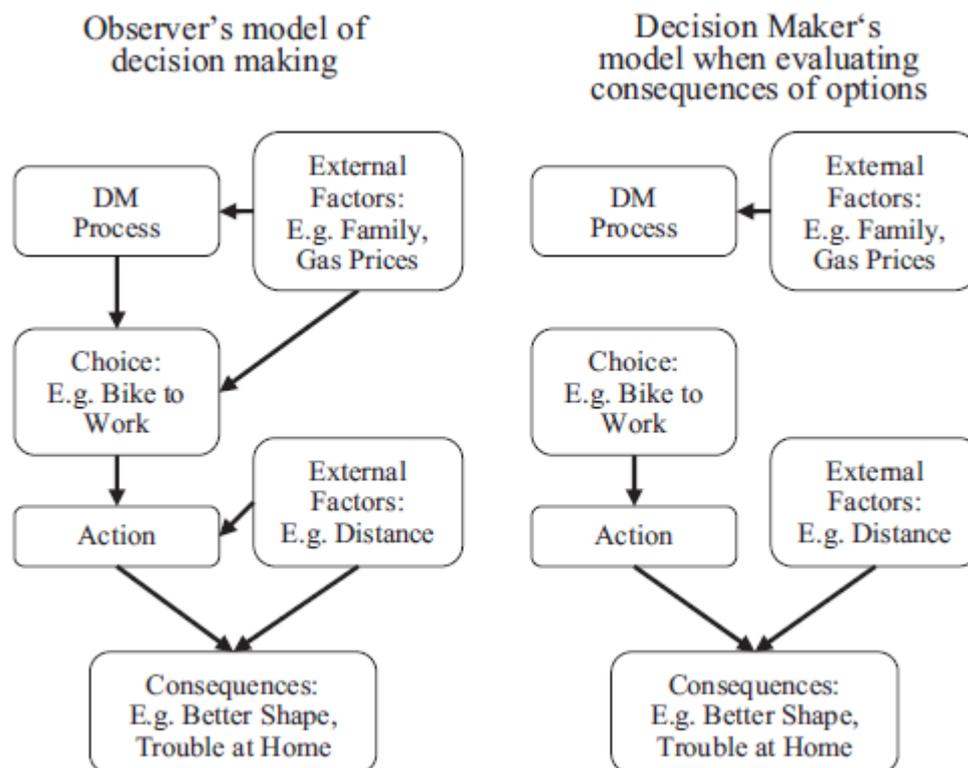


Figure 11-2. 観察者 (左側) および意思決定者 (右側) の観点からモデル化した熟慮に基づく選択。DM は意思決定 (Decision Making) の略。(Hagmayer & Sloman, 2009 より)

に影響される。意思決定者が気付いているか否かに関わらず、いくつかの外的な要因（e.g., 重要な他者からの圧力、ガソリン代）が意思決定の過程や選択に影響を及ぼすかもしれない。

意思決定過程の結果は選択である。Figure 11-2 の因果モデルにおいて選択は意思決定や行動とは区別して表現される。意思決定の結果が選択で終わらない場合が存在することから（e.g., いろいろ悩んだけど、仕事に自転車で行くかまだ選べない）、意思決定と選択は区別される。また、選択をしたにも関わらず選択された行動が行われない場合が存在することから（e.g., 自転車で行こうと決めただけど、いつも通車を使っている）、選択と行動は区別される。選択（e.g., 自転車に乗るかの決定）は行動に影響し、行動は他の環境要因（e.g., タイヤが扁平か否か）とともに結果に帰着する（e.g., 運動量の増加と到着時間の遅延）。ただし、Figure 11-2 左側に示したような因果モデルでは、選択から生じる行動や選択は因果ベイズネットの意味する介入になりえない。それらが介入であるならば、他の原因と独立でなければならないからである。

**意思決定者の観点での意思決定** これまで観察者の観点から意思決定について説明してきたが、以下では意思決定者自身の観点から意思決定について述べる。仕事に自転車で行くかどうか決定する際には、習慣（e.g., 夏の間はいつも自転車を使う）や外的な圧力（e.g., 放課後子どもたちを迎えに行かなければ、夕飯抜きになるかもしれない）や直感（e.g., ただなんとなく、自転車に乗る気がしない）など様々な過程の意思決定が考えられるが、ここでは選択の結果を考慮して行われる意思決定について焦点を当てる。Hagmayer and Slovic (2009) の仮説では、意思決定者は選択肢から生じる結果を Figure 11-2 右側のような因果モデルによって予測する。意思決定者の因果モデルは観察者のそれと類似しているが、意思決定者は自身の選択を行動に対する介入として捉えている点で両者は異なる。意思決定者が自身の選択を独立に生成し、行動を決定することが選択の因果モデル理論では仮定されている。選択肢から生じる結果を考慮する際、選択はあらゆる事象と独立であり、行動は選択からのみ影響を

受ける。したがって、選択によってグラフ切除が実行されると考えられ、Figure 11-2 右側の因果モデルにおいて行動に対するエッジが取り除かれる。選択が介入として捉えられるという仮定に対して、Hagmayer and Sloman (2009) は以下の 3 つの理由を述べている。

第一に、選択が自由意志に基づいて行われていることが挙げられる。自由意志 (free will) と決定論 (determinism) の間には様々な議論があるが、多くの人が自由意志の存在を信じていることは確かである。自由意志に基づく選択はあらゆる事象から影響を受けることはない。このことから選択を表すノードを因果モデルにおける外生変数として表現している。

ヒトは選択に基づく行動が何らかの結果を起こした場合、結果の原因を行動と考えることが第二の理由である。つまり、選択や行動は結果に影響を及ぼす他の要因と混同されないということである。もし行動に何らかの要因が常に混同していたら、因果関係を正しく学習できず、行動と結果の関係が擬似相関によるものか悩まされ続けるだろう。このことは、選択のノードのみが行動に影響し、行動に影響を及ぼす外的要因からのエッジが取り除くことで表現されている。

第三の理由として、選択から選択に影響する外的要因についての推論が行われないことが考えられる。これらの要因は前もって知られていて意思決定の過程で考慮されるか、未知のままであるかのどちらかである。例えば、自転車で仕事に行くことを決めた時、ガソリン代の高さが新たに分かることはなく、ガソリン代は意思決定の過程で既に考慮されている。このように、選択からその原因を推測できないことは選択のノードに対するエッジが取り除かれることで表されている。

**意思決定の過程** 選択の因果モデル理論における中心的な考えは、熟慮に基づく意思決定が単純化された因果モデルと仮想的な選択に基づくということである。そして、意思決定者は結果に対する相関関係よりも因果関係を考慮し、自身の選択を介入と捉えることがこの理論では仮定されている。選択の因果モデル理論によると、意思決定

の過程には因果モデルの構築と因果モデルにおける選択結果の予測と選択肢の決定という3つの段階 (Figure 11-3 参照) が含まれている。

第一段階では、意思決定問題について因果モデルの表象が形成される。まず初めに達成すべき目標が設定され、その目標に対して因果関係がある要因が選定される。ヒトは結果に影響を及ぼす要因を容易に思い浮かべることができるが (e.g., Ahn, Kalish, Medin, & Gelman, 1995), 意識的に考慮される要因の数は比較的少数である (Fischhoff, Slovic, & Lichtenstein, 1978)。次に、これらの要因がどのように影響しあうかについて因果モデルが構築される。そして、利用可能な情報を基に因果モデルが

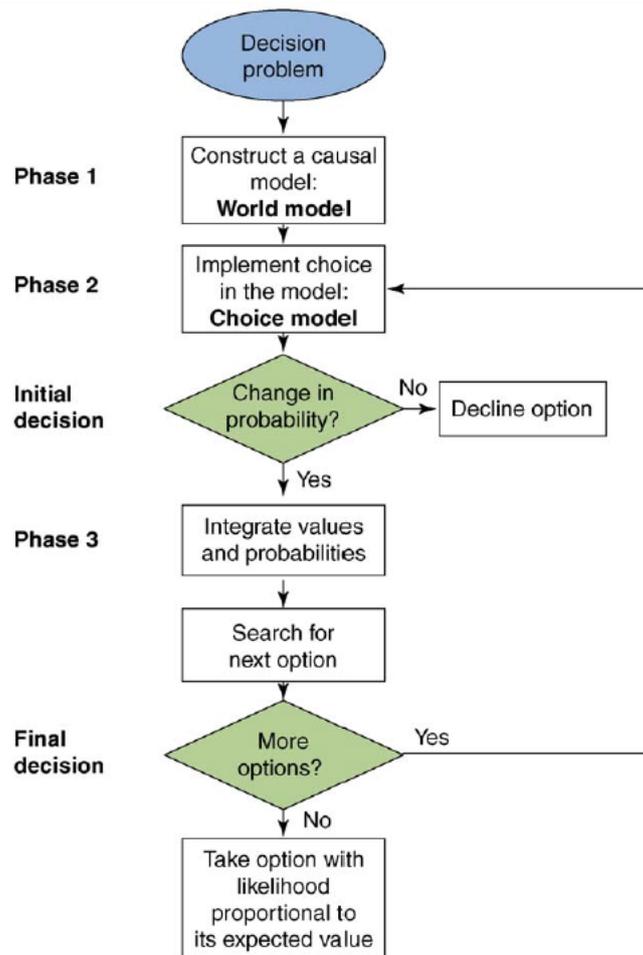


Figure 11-3. 選択の因果モデル理論における意思決定の過程 (Sloman & Haggmayer, 2006 より)

更新される。

第二段階では、意思決定問題について選択肢の表象が形成される。意思決定者が  
選択肢の結果を評価する前に、各々の選択肢から生ずる結果を表す新しい因果モデル  
が構成されなければならない。選択肢を含む新たな因果モデルは選択モデル  
(choice model) と言われ、因果モデルとは以下の点で異なる。選択を表すノードが追  
加され、変数の値が指定される。選択は介入と捉えられるため、グラフ切除 (i.e., 介入  
される事象は他の原因と独立になる) が実行される (Figure 11-1 参照)。ヒトは推論を  
行う際に、介入が他の原因との独立関係を生成することを理解している (Meder,  
Hagmayer, & Waldmann, 2008, 2009; 斎藤・嶋崎, 2012; Sloman & Lagnado,  
2005; Waldman & Hagmayer, 2005)。そして、選択モデルにおいて仮想的な選択と  
その結果の予測が為される。つまり、介入に基づく結果の生起確率が計算される。  
Hagmayer and Sloman (2009) は、ヒトが介入に基づく確率を正確に推測することは  
ないと主張しているが、いくつかの研究では発生的な介入の結果の確率を極めて的確  
に推測することが示されている (Meder, Hagmayer, & Waldmann, 2009; Waldman,  
Hagmayer, 2005)。

第三段階では、選択モデルの結果に基づいて意思決定が行われる。あらゆる行為は  
多かれ少なかれ時間や労力などを必要とするため、初期状態の選択は何もしないこと  
である。共通の原因によって行動と何らかの事象が関連している場合にこのような選択が  
行われる (Figure 11-1 左側参照)。対照的に、選択が結果の生起確率を高める場合、  
コストが利得に勝らない限り、選択が受理される。行動が何らかの事象の原因となってい  
る場合にこのような選択が行われる。もし、選択が結果の生起確率を十分に高めるなら  
ば、その選択肢は即座に受け入れられる (i.e., initial decision)。このような戦略は満  
足化 (satisfying) と言われる (Simon, 1957)。緊迫している事態において、ヒトは最  
初に思いついた選択肢をとる傾向がある (Klein, 1997; Zsombok & Klein, 1997)。  
結果の生起確率を上昇させる選択肢が複数存在する場合、それぞれの選択肢に対し

て介入に基づく確率が計算される。最終的な意思決定 (i.e., final decision) のために、介入に基づく確率とその効用が統合される (Koehler & Harvey, 2004)。

#### 11. 1. 4. 質的推定説と量的推定説

選択の因果モデル理論では、強制された行動や意図的に選択された行動は介入として機能することが想定されている。介入に基づく確率計算には質的推定と量的推定の2つの立場が存在する。質的推定説 (e.g., Sloman, 2005) によると、介入における確率は因果モデルの質的側面 (i.e., 因果構造) によってのみ判断され、因果モデルの量的側面 (i.e., 基準率, 因果強度) は無視される。選択の因果モデル理論は質的推定説の立場をとっており、“人々は介入における確率を正確に推測することはできないが、判断は無効化 (undoing) と一致するだろう” (Hagmayer & Sloman, 2009, p.26) と述べている。一方、量的推定説 (e.g., Waldmann, Hagmayer, Blaisdell, 2006) では、ヒトが因果モデルとそのパラメータの両方を考慮することが主張されている。介入における確率は因果モデルの量的側面と質的側面から推定されるため、基準率や因果強度などのパラメータが反映されることが想定されている。

これまでの研究では量的推定説を支持する結果が報告されている。Waldmann and Hagmayer (2005) は、観察によって獲得した知識から介入の結果を予測できるか、また、この推論に基準率や因果強度が反映されるか否か検討を行った。実験参加者はリスト形式で与えられた共変動情報を基に因果モデルの質的側面と量的側面を考慮し、介入における確率を正確に推測することができていた。さらに、このような推論能力は毎試行ごとに事象の状態を提示する離散試行手続きでも確認されている (Meder, Hagmayer, & Waldmann, 2009)。

#### 11. 1. 5. 本研究の目的

Hagmayer and Sloman (2009) は、ヒトが意思決定場面の因果モデルを用いて意

思決定を行うことや、意図的に選択された行動が介入として機能することを実証し、選択の因果モデル理論の妥当性が示唆された。しかしながら、選択された行動や強制された行動に関する推測において、基準率や因果強度といった因果モデルの量的側面が考慮されるか否かはまだ不明のままである。意図的に選択された行動や強制された行動における介入は、変数を操作した場合の介入と異なる可能性もあるため、意思決定における因果推論においてどのような判断がなされるか検討することは非常に重要である。いずれの先行研究もこの事態に関する直接的な知見をまだ提供していない。

本研究の目的は、選択された行動や強制された行動に基づいて推測を行う際に因果モデルの量的側面が反映されるか否か検討することである。量的情報の使用の検討に先立ち、実験 1 では意図的に選択された行動が介入として機能するかどうか確認した。続く実験 2 では、因果強度を一定に保ちつつ基準率を操作し、基準率が考慮されるか検討した。実験参加者が因果モデルの基準率を使用しない場合、判断は基準率の操作による影響を受けないだろう。また、実験 3 では基準率を一定に保ちつつ因果強度を操作し、因果強度が考慮されるか検討した。実験参加者が因果モデルの因果強度を考慮に入れない場合、判断は因果強度の操作による影響を受けないだろう。

## 11. 2. 実験 1

実験 1 の目的は、意図的に選択された行動が介入として機能するという Hagmayer and Sloman (2009, Experiment 4) の知見を確認することであった。実験手続きは先行研究とほぼ同様であったが、行為が生起していた場合だけでなく、生起していなかった場合の推測も求めた点で異なっていた。実験参加者の課題はシナリオを読んで、行為の原因や結果について推測を行うことであった。実験課題には 2 種類のシナリオを使用した (Figure 11-4 参照)。一方のシナリオは行為と結果が共通原因によって引き起こされるものであり (i.e., 共通原因モデル)、他方のシナリオは行為が媒介事象に影響し、媒介事象が結果に影響するというものであった (i.e., 因果連鎖モデル)。

行為は選択・介入・観察の 3 通りに記述された。具体的には、選択の場合は意図的に選択された行動として、介入の場合は強制された行動として、観察の場合は単に観察された行動として説明されていた。これらはすべて行為を記述したものであるが、行為の原因に関して互いに異なるものとなっていた。因果連鎖モデルでは行為の理由に関わらず、行動の生起は非生起の場合と比べて媒介事象の生起確率を上昇させる [i.e.,  $P(\text{intermediate}|\text{action}) > P(\text{intermediate}|\neg\text{action})$ ]. しかし、同様の推論は共通原因モデルでは成立しない。選択の因果モデル理論 (Hagmayer & Sloman, 2009) によると、意図的に選択された行動や強制された行動は介入として機能するため、行動と原因事象の関係を独立にする (i.e., グラフ切除, 無効化). そのため、選択された行為や強制された行為の状態は、その原因事象の推測に影響しない [i.e.,  $P(\text{cause}|\text{action}) = P(\text{cause}|\neg\text{action})$ ]. 対照的に、単に観察された行為の状態は、その原因事象に対して診断的証拠をもたらすため、原因事象の推測に影響する [i.e.,  $P(\text{cause}|\text{action}) > P(\text{cause}|\neg\text{action})$ ]. 意図的に選択された行動が介入として機能する場合、選択に基づく推論は介入に基づく推論と同様の結果になると予測される。つまり、行為の有無によって原因事象に対する評定値は変化せず [i.e.,  $P(\text{cause}|\text{action}) = P(\text{cause}|\neg\text{action})$ ], 選択に基づく推論と介入に基づく推論の評定値は等しくなるはずである。

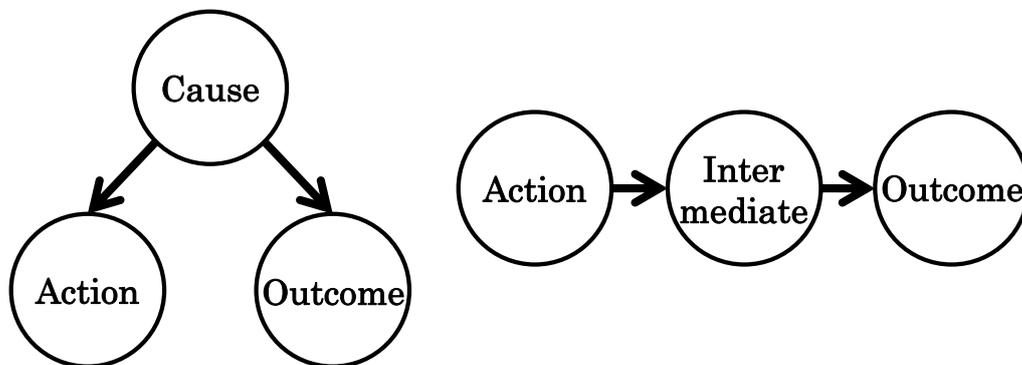


Figure 11-4. 共通原因モデル(左図)と因果連鎖モデル(右図)

## 11. 2. 1. 方法

**実験参加者および実験計画** 大学生 36 名を実験参加者とした。実験参加者には、実験参加のお礼として成績の加点となる実験参加証を渡した。実験参加者は推論の種類によって生じる 3 つの群（選択、介入、観察）のいずれかに無作為に割り当てられた。また、2 種類の因果モデル（共通原因モデル、因果連鎖モデル）を実験参加者内で操作した。

**実験手続き** 実験課題はシナリオを読んで推測を行うことであり、3 ページで構成される小冊子を 2 種類用意した。1 ページ目にはシナリオに関する情報が、2 ページ目と 3 ページ目には確率判断を尋ねる項目が記載されていた。2 種類のシナリオは共通原因モデルか因果連鎖モデルのいずれかで記されていた（Figure 11-4 参照）。これらのシナリオの大部分は Hagmayer and Sloman (2009, Experiment 4) と同様であったが、因果モデルのパラメータを操作するという一連の研究目的のため、シナリオに記載されていた数値は取り除かれた。

1 ページ目では、行動と結果の間に強い統計的関係がみられることが伝えられ、その背後にある因果構造について説明がなされた（Figure 11-4 参照）。2 つのシナリオは家事の手伝いや運動に関するものであった。家事の手伝いに関するシナリオでは、家事の手伝いが行為（action）、健康が結果（outcome）、人生の見通しが共通原因（common cause）または媒介事象（intermediate）であり、具体的な内容は以下の通りであった。

最近の研究によって以下のことが明らかにされました。家事を手伝う男性は、そうでない男性よりも健康状態が良好である。研究によると、この結果は、男性の人生に対する見通しに起因することがわかっている。人生に対して明確な見通しを持つことによって、家事を手伝うようになる。また、人生に対して明確な見通しを持つことによって、より健康になる（共通原因モデル）。研究によると、家事を手伝うことで達成感が得られ、達成感は人生に対する明確な見通しを持つことに

つながることがわかっている。人生に対して明確な見通しを持つことによって、より健康になる（因果連鎖モデル）。この図は研究報告によって明らかにされた因果関係を表しています。人生に対する見通しを持つと、家事を手伝うようになります。また、人生に対する見通しを持つと、健康状態が改善します（共通原因モデル）。家事を手伝うと、人生に対する見通しを持つようになります。そして、人生に対する見通しを持つと、健康状態が改善します（因果連鎖モデル）。

実験参加者は自分のペースでシナリオを読み、実験者に要約を伝えた。実験者は実験参加者がシナリオの因果モデルを理解しているか否か確認し、理解していないと判断した場合にはシナリオを再読させた。シナリオの順序やシナリオと因果モデルの組み合わせは実験参加者間でカウンターバランスをとった。

実験参加者は3群に分けられ、各群では2ページ目と3ページ目で求められる確率判断の問題文における行動の理由が異なっていた。選択群では、知り合いの男性がこの研究報告を読み、意図的に家事を手伝うように決断して家事を手伝っていることを想像するよう求めた。介入群では知り合いの男性が家事を手伝わされていることを、観察群では知り合いの男性が家事を手伝っていることを想像するよう指示した。その後、共通原因モデルにおける原因事象、あるいは因果連鎖モデルにおける媒介事象の確率判断と判断に対する確信度の評定が求められた。家事の手伝いに関するシナリオでは、「知り合いの男性の人生に対する明確な見通しについて、どのような推測ができますか？」という質問にVAS（0:存在しない, 50:推測できない, 100:存在する）で回答するよう求めた。「この回答にどのくらい自信がありますか？」という確信度の評定も同様にVAS（0:全く自信がない, 100:完全に自信がある）を用いた。同様に、行動を行っていない場合についても推測することが求められた。2つのシナリオを通して  $P(\text{cause}|\text{action})$  と  $P(\text{cause}|\neg\text{action})$  と  $P(\text{intermediate}|\text{action})$  と  $P(\text{intermediate}|\neg\text{action})$  という4種類の確率を推測させた。

## 11. 2. 2. 結果および考察

Table 11-1 に評定値の平均と標準偏差を示した。因果連鎖モデルでは、推論の種類に関わらず、行動の生起時に媒介事象の生起確率が高く、行動の非生起時に媒介事象の生起確率が低くなっている。一方、共通原因モデルでは、選択群と介入群は行動の有無に関わらず同様の評定を行っていることが示された。しかしながら、観察群では行動生起時の評定値が行動非生起時よりも高くなっていた。

推論の種類(3)×因果モデル(2)×行為の有無(2)の混合三要因分散分析の結果、行為の有無の主効果 ( $F(1, 33) = 26.61, MSE = 707.12, p < .001, \eta^2_G = .16$ ) および因果モデルと行為の有無の交互作用 ( $F(1, 33) = 12.10, MSE = 456.75, p = .001, \eta^2_G = .05$ ) が有意であった。結果の更なる分析のために、各群のデータに対して二要因の分散分析を実施した。選択群の評定値について因果モデル(2)×行為の有無(3)の分散分析を行ったところ、行為の有無の主効果 ( $F(1, 11) = 8.93, MSE = 497.89, p = .012, \eta^2_G = .12$ ) と因果モデルと行為の有無の交互作用 ( $F(1, 11) = 5.80, MSE = 465.23, p = .034, \eta^2_G = .08$ ) が確認された。選択の因果モデル理論の予測と一致して、共通原因モデルでは行為の有無によって評定値が変化することはなく ( $F < 1$ )、因果連鎖モデルでは行為の有無によって評定値が変化していた ( $F(1, 11) = 12.41, MSE = 567.28, p = .004, \eta^2_G = .37$ )。同様の分析を介入群の評定値に対し

Table 11-1

実験 1 における確率判断の平均値と標準偏差

Inference	Common-cause		Causal-chain	
	$P(\text{cause}   \text{action})$	$P(\text{cause}   \neg\text{action})$	$P(\text{intermediate}   \text{action})$	$P(\text{intermediate}   \neg\text{action})$
Choice				
<i>M</i>	48.50	44.25	71.25	37.00
<i>SD</i>	32.54	27.49	15.02	29.33
Intervention				
<i>M</i>	50.50	54.17	76.75	42.42
<i>SD</i>	35.74	30.70	30.27	32.51
Observation				
<i>M</i>	68.50	37.67	74.33	37.17
<i>SD</i>	17.32	25.42	15.44	21.60

でも行ったところ、因果モデルと行為の有無の交互作用が有意であり ( $F(1, 11) = 5.18$ ,  $MSE = 836.77$ ,  $p = .043$ ,  $\eta^2_G = .09$ ), 行動が生起している場合の媒介事象の生起確率は行動が生起していない場合よりも高く評定されていた ( $F(1, 11) = 14.48$ ,  $MSE = 488.39$ ,  $p = .002$ ,  $\eta^2_G = .25$ ). 選択群や介入群とは対照的に、観察群の評定値に対する分散分析では行為の有無の主効果のみが有意であり ( $F(1, 11) = 21.04$ ,  $MSE = 659.27$ ,  $p < .001$ ,  $\eta^2_G = .43$ ), 交互作用は見られなかった ( $F(1, 11) = 1.76$ ,  $MSE = 68.24$ ,  $p = .211$ ,  $\eta^2_G = .01$ ).

Table 11-2 に各群における確信度の平均値と標準偏差を示した。推論の種類(3)×因果モデル(2)×行為の有無(2)の混合三要因分散分析の結果、行為の有無の主効果のみ有意であり ( $F(1, 33) = 7.10$ ,  $MSE = 174.27$ ,  $p = .011$ ,  $\eta^2_G = .02$ ), その他の主効果や交互作用は有意でなかった。

要約すると、実験1の結果は意図的に選択された行動は介入として機能するという先行研究の知見を再現するものであった。因果連鎖モデルにおける実験参加者の評定値は0や100といった極端な値よりも50といった中央の値に近かったが、これは確信度がそこまで高くないために中心化傾向が生じたものと解釈できる。しかしながら、実験1の結果は選択や介入に基づく推論と観察に基づく推論が機能的に異なることを明確に示している。

Table 11-2

実験1における確信度の平均値と標準偏差

Inference	Common-cause		Causal-chain	
	$P(\text{cause}   \text{action})$	$P(\text{cause}   \neg\text{action})$	$P(\text{intermediate}   \text{action})$	$P(\text{intermediate}   \neg\text{action})$
Choice				
<i>M</i>	56.25	45.58	52.83	48.17
<i>SD</i>	31.04	20.80	25.57	25.65
Intervention				
<i>M</i>	65.58	55.67	69.33	59.58
<i>SD</i>	19.94	24.09	21.63	21.93
Observation				
<i>M</i>	53.75	55.50	57.75	55.83
<i>SD</i>	25.91	23.39	21.30	18.99

### 11. 3. 実験 2

実験 1 の結果、意図的に選択された行動は介入と同様に原因事象との独立関係を形成し、行動から本来の原因事象の状態を予測できなくすることが示された。これらの結果は選択の因果モデル理論の予測と一致し、その妥当性が示唆された。しかし、選択に基づく推論や介入に基づく推論が因果モデルの構造にのみ従って行われるか、因果モデルのパラメータ（e.g., 因果強度、事象の基準率）まで考慮に入れるかは明らかにされていない。実験 2 の目的は選択および介入に基づく推論が因果モデルの構造だけでなく、因果モデルのパラメータを考慮して行われているか否か検討することである。実験 2 では原因事象がどのくらい発生するかを表す事象の基準率を操作し、シナリオを読んだから行われる実験参加者の推測に影響を及ぼすか検討した。実験 1 と異なり、因果モデルのパラメータを示すために実験 2 ではリスト提示課題を用いて事象間の関係を伝えた。Hagmayer and Sloman (2009) が主張するように選択や介入に基づく推論が因果構造のみを使用して行われるならば、事象の基準率に関わらず一貫した評定値が得られるはずである。一方、これらの推論が因果モデルのパラメータも考慮に入れているなら、事象の基準率に応じて評定値が変化するはずである。

#### 11. 3. 1. 方法

**実験参加者および実験計画** 大学生 48 名を実験参加者とした。この中に実験 1 に参加した者はおらず、実験参加者には実験参加のお礼として成績の加点となる実験参加証を渡した。実験参加者は選択群、介入群、観察群のいずれかに無作為に割り当てられた。また、原因事象の基準率（高基準率、低基準率）を実験参加者内で操作した。

**実験手続き** 実験参加者の課題は 2 つのシナリオを読み、推測を行うことであった。2 つのシナリオは実験 1 で使用した家事のシナリオと運動のシナリオと基本的に同様であったが、実験 1 とは異なり、どちらのシナリオも共通原因モデルで記述されていた

(Figure 11-5 参照). 1 つのシナリオは 4 枚の用紙で構成されており, 1 枚目は因果モデルの教示, 2 枚目はリストによる事象の生起情報の提示, 3・4 枚目は判断課題となっていた. 1 枚目の用紙には, シナリオの内容とその内容を図で表現した因果モデル (Figure 11-5 からパラメータを省いたもの) が記載されていた. 運動に関するシナリオの具体的な内容は以下の通りであった.

最近の研究によって以下のことが明らかにされました. 運動をしている人は, そうでない人よりも摂取カロリーが平均より少ない. この結果は神経伝達物質 CNA が原因となって運動する量が多くなったり, 食べる量が少なくなったりすることで生じると研究では述べられている. 上記の図は研究報告によって明らかにされた因果関係を表しています. 脳内における神経伝達物質 CNA の量が増加すると, より運動するようになります. また, 脳内における神経伝達物質 CNA の量が増加すると, 食欲が減って摂取カロリーが減少します.

2 枚目の用紙には, シナリオで示した事象の生起情報とその読み取り方が記載されていた (Table 11-3 参照). 事象の生起情報は 2 種類設けられ, 高基準率条件では原因事象の基準率が 0.8 であり (i.e.,  $P(\text{cause}) = 0.8$ ), 低基準率条件では原因事象の基準率が 0.2 であった (i.e.,  $P(\text{cause}) = 0.2$ ). 基準率の要因は実験参加者内で操作され,

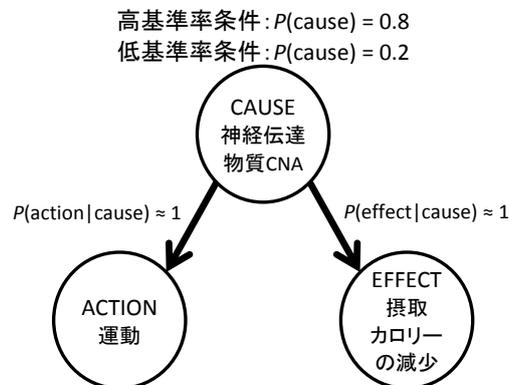


Figure 11-5. 実験 2 で使用した因果モデル.

一方のシナリオが高基準率条件となっていれば、もう一方のシナリオは低基準率条件に設定された。因果モデルにおける全てのパラメータを定義するために、事象の生起情報はシナリオにおける研究報告の詳細としてリストで提示され、各行は個々人の状態を表し、各列は3つの事象の状態（例えば、神経伝達物質 CNA の増加・運動・摂取カロリーの減少など）を表していた。事象が生起している場合は○、生起していない場合は×と書くことで各事象の状態を示した。実験参加者にはこのリストに一行ずつ目を通すことが求められた。リストの各行の順序は実験参加者ごとにランダムイズし、リストの左の列には原因事象が必ず表示され<sup>18</sup>、真ん中と右の列に表示する事象は実験参加者間でカウンターバランスを取った。

3・4 枚目の用紙では、判断課題と判断に対する確信度の評定を行うように伝えた。判断課題は、行動している場合に原因事象が存在する確率  $P(\text{cause}|\text{action})$  を判断する課題と行動していない場合に原因事象が存在する確率  $P(\text{cause}|\neg\text{action})$  を判断する課題の2種類があった。どちらの課題を先に行うかは実験参加者間でカウンター balan

Table 11-3

実験 2 における事象の生起情報

Data pattern	base-rate	
	high	low
cause, action, effect	14	4
cause, action, $\neg$ effect	1	0
cause, $\neg$ action, effect	1	0
cause, $\neg$ action, $\neg$ effect	0	0
$\neg$ cause, action, effect	0	0
$\neg$ cause, action, $\neg$ effect	0	1
$\neg$ cause, $\neg$ action, effect	0	1
$\neg$ cause, $\neg$ action, $\neg$ effect	4	14
$\Sigma$	20	20

注) 数値は生起頻度を表す

<sup>18</sup> リストの行を見る順序と因果の方向性を一致させるために、このような提示方法を用いた。

スを取った。実験参加者は選択群，介入群，観察群のいずれかに無作為に割り当てられ，2つの判断課題に回答した。運動のシナリオで選択群において行動が生起する際は，知り合いの男性がこの研究報告を読んで意図的に運動を始めようと決断し，運動していることを想像するよう求めた。介入群では知り合いの男性が運動させられていることを，観察群では知り合いの男性が運動していることを想像するよう指示した。その後，原因事象の確率判断と判断に対する確信度の評定が求められた。運動に関するシナリオでは，「知り合いの脳内における神経伝達物質 CNA の有無について，どのような推測ができますか？」という質問に VAS (0:存在しない, 100:存在する) で回答させた。また，「この回答にどのくらい自信がありますか？」という確信度の評定も同様に VAS (0:全く自信がない, 100:非常に自信がある) を用いた。なお，判断の際に1枚目の用紙 (i.e., 因果モデルを示した用紙) や2枚目の用紙 (i.e., 事象の生起情報を示した用紙) を見ることは許可されていた。

実験参加者は4枚の用紙で構成されている1つのシナリオを読んで判断を行い，もう一方のシナリオにおいても同様の手順で判断を行うことが求められた。シナリオを十分に読ませるために，3・4枚目の判断課題を除いて用紙は一枚ずつ手渡された。事象の基準率やシナリオの順序や判断課題の順序は実験参加者間でカウンターバランスをとった。

**結果の予測** ここでは，まず始めに因果ベイズネットを用いて量的推定説の予測を導出し，その後，因果モデルのパラメータを考慮しない質的推定説の予測について説明を行う。因果ベイズネットにおける主要な仮定の1つは因果マルコフ条件である。因果マルコフ条件によると，因果モデルにおける事象はその直接の原因事象を固定するとその結果事象以外の事象と独立になる。例えば，Figure 11-5 に示されているような共通原因モデルでは，原因事象の状態 (e.g., 神経伝達物質 CNA の有無) が分かると，行動 (e.g., 運動) と結果事象 (e.g., 摂取カロリーの減少) は独立になる。つまり，結果事象の生起確率の推測には原因事象の状態さえ分かっただけで，行動の状態は不

必要になり、結果事象の生起確率に影響しない。因果マルコフ条件を適用すると、共通原因モデルの同時確率分布は次のように分解される。

$$\begin{aligned} P(CAUSE, ACTION, EFFECT) \\ = P(CAUSE) \times P(ACTION | CAUSE) \times P(EFFECT | CAUSE) \end{aligned} \quad (11-1)$$

観察条件における推論は事象の観察された状態に基づいて行われる。行動の生起を観察した場合に原因事象が生起する確率はベイズの定理を用いて以下のように算出される。

$$\begin{aligned} P(cause | action) &= \frac{P(action | cause) \times P(cause)}{P(action | cause) \times P(cause) + P(action | \neg cause) \times P(\neg cause)} \\ &= \frac{P(action | cause) \times P(cause)}{P(action)} \end{aligned} \quad (11-2)$$

(11-2)式に因果モデルのパラメータを代入すると以下の予測値が得られる。

$$\begin{aligned} P(cause | action)_{high \ base-rate} &= \frac{0.9375 \times 0.8}{0.75} = 1 \\ P(cause | action)_{low \ base-rate} &= \frac{1 \times 0.2}{0.25} = 0.8 \end{aligned} \quad (11-3)$$

行動の非生起を観察した場合の原因事象の生起確率は次のように計算される。

$$\begin{aligned} P(cause | \neg action) &= \frac{P(\neg action | cause) \times P(cause)}{P(\neg action | cause) \times P(cause) + P(\neg action | \neg cause) \times P(\neg cause)} \\ &= \frac{P(\neg action | cause) \times P(cause)}{P(\neg action)} \end{aligned} \quad (11-4)$$

(11-4)式に因果モデルのパラメータを代入すると以下の予測値が得られる.

$$\begin{aligned}
 P(\text{cause} | \neg \text{action})_{\text{high base-rate}} &= \frac{(1-0.9375) \times 0.8}{(1-0.75)} = 0.2 \\
 P(\text{cause} | \neg \text{action})_{\text{low base-rate}} &= \frac{(1-1) \times 0.2}{(1-0.25)} = 0
 \end{aligned}
 \tag{11-5}$$

観察の推論と異なって介入条件や選択条件における推論は、行動と原因事象間のリンクが取り除かれた因果モデルに基づいて行われる。これは、グラフ切除（無効化）によって行動と原因事象が独立になるからである。介入によって事象がある状態になることは  $\text{do}[\ ]$  と記述され、このような記述方法は  $\text{do}$  オペレータと言われる (Pearl, 2000)。行動と原因事象は独立になるので、行動の状態に関わらず原因事象は基準率で生起することが予測される。

$$P(\text{cause} | \text{do} [\text{action}]) = P(\text{cause} | \text{do} [\neg \text{action}]) = P(\text{cause})
 \tag{11-6}$$

(11-6)式に因果モデルのパラメータを代入すると、量的推定説の予測値が求められる。

$$\begin{aligned}
 P(\text{cause} | \text{do} [\text{action}])_{\text{high base-rate}} &= P(\text{cause} | \text{do} [\neg \text{action}])_{\text{high base-rate}} = 0.8 \\
 P(\text{cause} | \text{do} [\text{action}])_{\text{low base-rate}} &= P(\text{cause} | \text{do} [\neg \text{action}])_{\text{low base-rate}} = 0.2
 \end{aligned}
 \tag{11-7}$$

続いて、選択の因果モデル理論、すなわち質的推定説の予測について説明を行う。観察群では選択の因果モデル理論と量的推定説の予測は同じになるが、介入群および選択群では異なる。質的推定説では介入に基づく推論は因果モデルの構造にのみ従い、量的推定説では介入に基づく推論は因果モデルの構造とパラメータに従うと考え

るからである。介入群や選択群では、評定値に事象の基準率が反映されないことを質的推定説は予測する。つまり、(11-7)式における各条件の評定値が同じになる。

$$\begin{aligned}
 P(\text{cause} | \text{do} [\text{action}])_{\text{high base-rate}} &= P(\text{cause} | \text{do} [\text{action}])_{\text{low base-rate}} \\
 P(\text{cause} | \text{do} [\neg\text{action}])_{\text{high base-rate}} &= P(\text{cause} | \text{do} [\neg\text{action}])_{\text{low base-rate}}
 \end{aligned}
 \tag{11-8}$$

質的推定説では介入や選択の推論は因果モデルの構造にのみ従うため、予測として数値を算出することを想定していないが、数量的な予測を行う場合は基準率の初期値である 0.5 になることが想定される (cf. Waldmann & Hagmayer, 2005)。この仮定の下で(11-8)式は次のようになる。

$$\begin{aligned}
 P(\text{cause} | \text{do} [\text{action}])_{\text{high base-rate}} &= P(\text{cause} | \text{do} [\text{action}])_{\text{low base-rate}} = 0.5 \\
 P(\text{cause} | \text{do} [\neg\text{action}])_{\text{high base-rate}} &= P(\text{cause} | \text{do} [\neg\text{action}])_{\text{low base-rate}} = 0.5
 \end{aligned}
 \tag{11-9}$$

### 11. 3. 2. 結果および考察

Table 11-4 に評定値の平均と標準偏差を示した。高基準率条件と低基準率条件と

Table 11-4

実験 2 における確率判断の平均値と標準偏差

Inference	High base rate		Low base rate	
	$P(\text{cause}   \text{action})$	$P(\text{cause}   \neg\text{action})$	$P(\text{cause}   \text{action})$	$P(\text{cause}   \neg\text{action})$
Choice				
<i>Predicted values</i>	50	50	50	50
<i>M</i>	53.06	48.69	52.25	44.38
<i>SD</i>	32.25	27.49	28.71	26.12
Intervention				
<i>Predicted values</i>	50	50	50	50
<i>M</i>	49.63	52.81	50.06	50.13
<i>SD</i>	35.20	31.65	29.48	29.74
Observation				
<i>Predicted values</i>	100	20	80	0
<i>M</i>	83.19	29.63	72.31	33.75
<i>SD</i>	18.48	22.13	24.34	27.64

もに、選択に基づく推論では行動の生起・非生起で原因事象の生起確率の評定値がほぼ同様であり、介入に基づく推論でも行動が生起している時と非生起の時で類似した評定値を示しており、観察に基づく推論では行動が生起している時の方が生起していない時よりも原因事象の生起確率が高く評定されている。これらの結果は、ヒトが観察に基づく推論と介入に基づく推論を区別しており、選択を介入として捉えることを示している。実験 2 の目的である因果モデルのパラメータが考慮されているか否かは高基準率条件と低基準率条件の比較によって明らかになる。選択に基づく推論では、高基準率条件における評定値と低基準率条件における評定値に大きな差はみられない。介入に基づく推論でも同様に、2 つの条件間に明らかな差はみられない。観察に基づく推論では、高基準率条件において行動が生起している時の原因事象の生起確率が低基準率条件よりもやや高く評定されている。行動が非生起の場合、高基準率条件と低基準率条件の評定値はほぼ等しい。これらの結果は、選択に基づく推論や介入に基づく推論において事象の基準率が考慮されていないことを意味し、質的推定説を採用している選択の因果モデル理論を支持している。

これらの結果は以下の統計解析によって確認されている。評定値について推論の種類(3)×基準率(2)×行動の有無(2)の混合三要因分散分析を行ったところ、行動の有無の主効果 ( $F(1, 45) = 9.25, MSE = 1473.88, p = .003, \eta^2_G = .09$ ) および推論の種類と行動の有無の交互作用が有意であった ( $F(2, 45) = 7.11, MSE = 1473.88, p = .002, \eta^2_G = .13$ )。結果を詳細に調べるために、それぞれの推論ごとに評定値のデータを別々にしてから分析を行った。選択群における評定値について基準率(2)×行動の有無(2)の二要因分散分析を行ったところ、いずれの主効果および交互作用も有意でなかった ( $F_s < 1$ )。また、介入群における評定値について基準率(2)×行動の有無(2)の二要因分散分析を同様に行ったところ、いずれの主効果および交互作用も有意でなかった ( $F_s < 1$ )。しかし、観察群における評定値について基準率(2)×行動の有無(2)の二要因分散分析を行ったところ、行動の有無の主効果が有意であり ( $F(1, 15) =$

36.32,  $MSE = 934.70$ ,  $p < .001$ ,  $\eta^2_G = .51$ ), 基準率と行動の有無の交互作用が有意傾向であった ( $F(1, 15) = 3.76$ ,  $MSE = 239.50$ ,  $p = .071$ ,  $\eta^2_G = .03$ ). 基準率と行動の有無の交互作用について下位検定を行ったところ,  $P(\text{cause}|\text{action})$ の評定において基準率の効果が有意傾向であった ( $F(1, 15) = 4.04$ ,  $MSE = 234.46$ ,  $p = .062$ ,  $\eta^2_G = .06$ ). 選択や介入に基づく推論では事象の基準率が考慮されず, 観察に基づく推論では事象の基準率が考慮されることが明らかとなった.

Table 11-5 に確信度の平均値を示した. 論の種類(3)×基準率(2)×行為の有無(2)の混合三要因分散分析の結果, 推論の種類の主効果および行為の有無の主効果が有意であり ( $F(2, 45) = 3.79$ ,  $MSE = 1443.48$ ,  $p = .030$ ,  $\eta^2_G = .10$ ,  $F(1, 45) = 6.30$ ,  $MSE = 190.50$ ,  $p = .016$ ,  $\eta^2_G = .13$ , respectively), その他の主効果や交互作用は有意でなかった. また, 選択群や介入群の確信度は観察群よりも若干低いものの, 中程度に留まっていた. このことは, 選択群や介入群における確率判断が確信度の低下による中心化傾向で生じたものではないことを示唆している.

これらの結果をまとめると, 実験 2 の結果は選択の因果モデル理論の予測と一致しており, 選択や介入に基づく推論では基準率が使用されず, 観察に基づく推論では基準率が考慮されていた. また, これらの判断は確信度の低下によって引き起こされていたわけではなかった.

Table 11-5

実験 2 における確信度の平均値と標準偏差 Table 6-1

Inference	High base rate		Low base rate	
	$P(\text{cause} \text{action})$	$P(\text{cause} \text{-action})$	$P(\text{cause} \text{action})$	$P(\text{cause} \text{-action})$
Choice				
<i>M</i>	47.06	49.13	55.25	45.88
<i>SD</i>	22.97	28.86	25.70	29.56
Intervention				
<i>M</i>	63.38	54.69	59.38	58.38
<i>SD</i>	21.15	23.02	17.18	21.29
Observation				
<i>M</i>	70.75	63.63	71.38	65.50
<i>SD</i>	21.34	21.95	14.38	23.22

### 11. 4. 実験 3

実験 2 の結果，意図的に選択された行動は介入として捉えられ，そのような行為に基づく推論では因果モデルにおける事象の基準率は無視されていることが示された．これらの結果は，選択や介入に基づく推論において因果モデルのパラメータ（e.g., 事象の基準率，因果強度）は考慮されないとする質的推定説の予測と一致している．実験 2 の結果から事象の基準率が考慮されていないことが明らかとなったが，事象間の因果強度が考慮されているか否かは明らかにされていない．実験 3 の目的は選択および介入に基づく推論が因果強度を考慮して行われているか否か検討することである．実験 2 と同様の課題を用いて，事象間の因果関係の強さを表す因果強度を操作し，実験参加者の推測に影響を及ぼすか検討した．選択や介入に基づく推論が因果構造のみを使用して行われるならば，因果強度に関わらず一貫した評定値が得られるはずである．一方，これらの推論が因果モデルにおける因果強度を考慮に入れているなら，因果強度に応じて評定値が変化するはずである．

#### 11. 4. 1. 方法

**実験参加者および実験計画** 大学生 42 名を実験参加者とした．この中に実験 1 ならびに実験 2 に参加した者はおらず，実験参加者には実験参加のお礼として成績の加点となる実験参加証を渡した．実験参加者は選択群，介入群，観察群のいずれかに無作為に割り当てられた．また，媒介事象に対する行動の因果強度（高因果強度，低因果強度）を実験参加者内で操作した．

**実験手続き** 手続きの大部分は実験 2 と同様であったが，使用する因果モデルとそのパラメータが異なっていた．家事のシナリオと運動のシナリオを因果連鎖モデルで記述し，基準率を一定に保ちながら因果強度を操作した（Figure 11-6 参照）．1 つのシナリオは 4 枚の用紙で構成されており，1 枚目は因果モデルの教示，2 枚目はリストによる事象の生起情報の提示，3・4 枚目は判断課題となっていた．

1 枚目の用紙には、シナリオの内容とその内容を図で表現した因果モデル (Figure 11-6 からパラメータを省いたもの) が記載されていた。2 枚目の用紙には、シナリオで示した事象の生起情報とその読み取り方が記載されていた (Table 11-6 参照)。事象の生起情報は 2 種類設けられ、高因果強度条件では行動が媒介事象を生起させる確率が高く (i.e.,  $P(\text{intermediate}|\text{action}) = 1$ )、媒介事象が結果事象を生起させる確率が低かった (i.e.,  $P(\text{effect}|\text{intermediate}) = 0.5$ )。対照的に、低因果強度条件では行動が媒介事象を生起させる確率が低く (i.e.,  $P(\text{intermediate}|\text{action}) = 0.5$ )、媒介事象が結果事象を生起させる確率が高く設定されていた (i.e.,  $P(\text{effect}|\text{intermediate}) = 1$ )。因果強度の要因 (高因果強度 vs. 低因果強度) は実験参加者内で操作され、一方のシナリオが高因果強度条件となっていれば、もう一方のシナリオは低因果強度条件に設定された。

3・4 枚目の用紙では、判断課題と判断に対する確信度の評定を行うことが求められた。判断課題は、行動している場合に媒介事象が存在する確率  $P(\text{intermediate}|\text{action})$  を判断する課題と行動していない場合に媒介事象が存在する確率  $P(\text{intermediate}|\neg\text{action})$  を判断する課題の 2 種類があった。実験参加者は、知り合いが意図的に行動を選択していることを想像する選択群、知り合いが行動させられていることを想像する介入群、知り合いが行動していることを想像する観察群のいずれかに無作為に割り当てられ、2 つの判断課題に回答した。

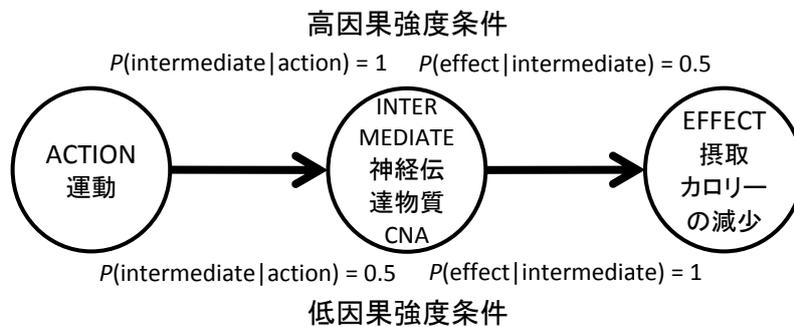


Figure 11-6. 実験 3 で使用した因果モデル。

**結果の予測** Figure 11-6 に示した因果モデルが与えられとき、因果ベイズネットを用いた量的推定説の予測は以下の通りである。因果連鎖モデルでは行為の理由に関わらず、行動が生起している場合はそうでない場合よりも媒介事象の生起確率が高くなる [i.e.,  $P(\text{intermediate}|\text{action}) > P(\text{intermediate}|\neg\text{action})$ ]. Table 11-6 に示した生起情報から、行動が生起している時の条件付き確率を計算すると、以下の予測値が得られる。

$$\begin{aligned} P(\text{intermediate} | \text{action})_{\text{strong}} &= 1 \\ P(\text{intermediate} | \text{action})_{\text{weak}} &= 0.5 \end{aligned} \tag{11-10}$$

量的推定説では因果強度が反映されるため、2つの因果強度条件で異なる予測値が導かれる。ただし、行動が生起していない場合は媒介事象が生起することはないため、2つの条件で同じ予測値が得られる。

$$P(\text{intermediate} | \neg\text{action})_{\text{strong}} = P(\text{intermediate} | \neg\text{action})_{\text{weak}} = 0 \tag{11-11}$$

Table 11-6

実験 3 における事象の生起情報

Data pattern	causal strength	
	strong	weak
action, intermediate, effect	5	5
action, intermediate, $\neg$ effect	5	0
action, $\neg$ intermediate, effect	0	0
action, $\neg$ intermediate, $\neg$ effect	0	5
$\neg$ action, intermediate, effect	0	0
$\neg$ action, intermediate, $\neg$ effect	0	0
$\neg$ action, $\neg$ intermediate, effect	0	0
$\neg$ action, $\neg$ intermediate, $\neg$ effect	10	10
$\Sigma$	20	20

注) 数値は生起頻度を表す

続いて、選択の因果モデル理論、すなわち質的推定説の予測について説明を行う。観察群では選択の因果モデル理論と量的推定説の予測は同じになるが、介入群および選択群では異なる。質的推定説では介入に基づく推論は因果モデルの構造にのみ従い、量的推定説では介入に基づく推論は因果モデルの構造とパラメータに従うと考えられるからである。介入群や選択群では、評定値に因果強度が反映されないことを質的推定説は予測する。つまり、(11-10)式における各条件の評定値が同じになる。

$$P(\text{intermediate} | \text{do} [\text{action}])_{\text{strong}} = P(\text{intermediate} | \text{do} [\text{action}])_{\text{weak}} = 1 \quad (11-12)$$

ここでは因果強度の初期値として 1 を設定しており、この値は因果関係が決定的なものとして考えられることを意味している。行動が生起していない場合は媒介事象が生起しないため、予測値は量的推定説と同様になる。

$$P(\text{intermediate} | \text{do} [-\text{action}])_{\text{strong}} = P(\text{intermediate} | \text{do} [-\text{action}])_{\text{weak}} = 0 \quad (11-13)$$

#### 11. 4. 2. 結果および考察

Table 11-7 に評定値の平均と標準偏差を示した。いずれの推論においても行動生起時の媒介事象の確率は行動非生起時よりも高く判断されている。選択に基づく推論では高因果強度条件と低因果強度条件における評定値にほとんど差は見られない。一方、介入に基づく推論では、行動が生起している場合において高因果強度条件の評定値が低因果強度条件よりも若干高くなっていた。また、観察に基づく推論では行動の有無に関わらず、高因果強度条件の方が低因果強度条件よりも高く評定されていた。選択群や観察群の結果は質的推定説を採用している選択の因果モデル理論を支持しているものの、介入群の結果は量的推定説の予測に近いものであった。

評定値について推論の種類(3)×因果強度(2)×行動の有無(2)の混合三要因分散

分析を行ったところ、因果強度の主効果 ( $F(1, 39) = 22.09, MSE = 271.10, p < .001, \eta_G^2 = .06$ ) と行動の有無の主効果 ( $F(1, 39) = 90.12, MSE = 578.93, p < .001, \eta_G^2 = .48$ ) が有意であり、推論の種類と因果強度の交互作用 ( $F(2, 39) = 4.78, MSE = 271.10, p = .013, \eta_G^2 = .03$ ) や因果強度と行動の有無の交互作用 ( $F(1, 39) = 4.30, MSE = 296.67, p = .044, \eta_G^2 = .01$ ) が見られた。結果を詳細に調べるために、それぞれの推論ごとに評定値のデータを別々にしてから分析を行った。選択群における評定値について因果強度(2)×行動の有無(2)の二要因分散分析を行ったところ、行動の有無の主効果のみ有意であった ( $F(1, 13) = 13.10, MSE = 1590.49, p = .003, \eta_G^2 = .35$ )。また、介入群における評定値について因果強度(2)×行動の有無(2)の二要因分散分析を同様に行ったところ、因果強度の主効果 ( $F(1, 13) = 12.17, MSE = 243.09, p = .004, \eta_G^2 = .12$ ) および行動の有無の主効果 ( $F(1, 13) = 44.13, MSE = 837.93, p < .001, \eta_G^2 = .63$ ) ならびに因果強度と行動の有無の交互作用 ( $F(1, 13) = 5.19, MSE = 84.85, p = .040, \eta_G^2 = .02$ ) が有意であった。下位検定として単純主効果の検定を行った結果、行動が生起している場合に因果強度の効果が有意であり ( $F(1, 13) = 17.18, MSE = 165.30, p = .001, \eta_G^2 = .34$ )、行動が生起していない場合は因果強度の効果は有意傾向であった。

Table 11-7

実験 3 における確率判断の平均値と標準偏差

Inference	Strong causal strength		Weak causal strength	
	$P(\text{intermediate}   \text{action})$	$P(\text{intermediate}   \neg \text{action})$	$P(\text{intermediate}   \text{action})$	$P(\text{intermediate}   \neg \text{action})$
Choice				
<i>Predicted values</i>	100	0	100	0
<i>M</i>	71.93	26.93	63.27	29.93
<i>SD</i>	23.17	24.51	23.29	33.03
Intervention				
<i>Predicted values</i>	100	0	100	0
<i>M</i>	83.79	26.79	63.64	17.86
<i>SD</i>	11.18	30.06	17.41	17.76
Observation				
<i>Predicted values</i>	100	0	50	0
<i>M</i>	81.07	29.93	55.79	15.21
<i>SD</i>	16.35	33.18	25.85	20.39

( $F(1, 13) = 3.43, MSE = 162.65, p = .087, \eta_G^2 = .03$ ). 同様に, 観察群における評  
定値について因果強度(2)×行動の有無(2)の二要因分散分析を同様に行ったところ,  
因果強度の主効果 ( $F(1, 13) = 15.96, MSE = 350.77, p = .002, \eta_G^2 = .15$ ) およ  
び行動の有無の主効果 ( $F(1, 13) = 67.34, MSE = 437.21, p < .001, \eta_G^2 = .48$ )  
が有意であり, 交互作用は有意でなかった ( $F(1, 13) = 1.20, MSE = 327.22, p$   
 $= .294, \eta_G^2 = .01$ ). 選択に基づく推論では因果強度が考慮されず, 介入に基づく推  
論や観察に基づく推論には因果強度が反映されることが示された.

Table 11-8 に確信度の平均値を示した. 推論の種類(3)×因果強度(2)×行為の有  
無(2)の混合三要因分散分析の結果, 因果強度と行動の有無の交互作用のみが有意  
であった ( $F(1, 39) = 7.45, MSE = 187.16, p = .009, \eta_G^2 = .01$ ). 単純主効果の検  
定の結果, 低因果強度条件では行動の有無によって確信度が異なることが明らかとな  
った ( $F(1, 39) = 5.35, MSE = 185.86, p = .021, \eta_G^2 = .02$ ).

これらの結果を要約すると, 実験 3 の結果は選択の因果モデル理論の予測と部分的  
一致していた. 具体的には, 選択に基づく推論では因果強度が反映されていなかった  
が, 介入に基づく推論では因果強度がある程度考慮されていた. また, 観察に基づく推  
論で因果強度が考慮されることは予測と一致していた.

Table 11-8

実験 3 における確信度の平均値と標準偏差

Inference	Strong causal strength		Weak causal strength	
	$P(\text{intermediate}   \text{action})$	$P(\text{intermediate}   \neg \text{action})$	$P(\text{intermediate}   \text{action})$	$P(\text{intermediate}   \neg \text{action})$
Choice				
<i>M</i>	59.93	55.80	53.13	59.07
<i>SD</i>	25.16	28.67	26.55	26.51
Intervention				
<i>M</i>	67.79	62.14	60.57	63.71
<i>SD</i>	20.58	21.60	19.28	19.15
Observation				
<i>M</i>	73.79	69.14	69.07	80.36
<i>SD</i>	26.74	26.39	24.21	19.92

## 11. 5. 総合論議

本研究は、意思決定における因果推論について選択が介入として捉えられているか否か、選択や介入に基づく推論は因果モデルのパラメータを無視して因果モデルの構造にのみ従って行われているか否か検討することを目的として実施された。実験参加者にシナリオを読ませて共通原因モデルにおける原因事象や因果連鎖モデルにおける媒介事象を推測させたところ、意図的に選択された行動や強制された行動と単に観察された行動を区別し、選択を介入として捉えていることが示された（実験 1）。リスト提示課題を用いて因果モデルのパラメータを考慮した推論が行われているか否か検討した結果、選択や介入に基づく推論では原因事象の基準率が考慮されることはなかった（実験 2）。また、選択に基づく推論では因果強度が反映されず、介入に基づく推論では因果強度が反映されることが示された（実験 3）。選択に基づく推論において基準率や因果強度が考慮されなかったことは質的推定説を支持するものであった。ただし、介入に基づく推論において因果強度が考慮されているなど質的推定説と一致しない結果も一部では見られた。

因果推論において基準率や因果強度が考慮されることを示し、量的推定説を支持する先行研究（Meder et al., 2009; Waldmann & Hagmayer, 2005）とは対照的に、本研究の結果はそれらのパラメータが考慮されないことを示し、質的推定説を支持するものであった。選択に基づく推論に基準率や因果強度が反映されなかったことの原因として、実験事態の差異が考えられる。量的推定説を支持する研究は事象の操作としての介入を扱っているのに対して、本研究における介入は選択された行動や強制された行動であった。これらの行動の原因は明確に示されておらず、どのようにして介入が生じているかという実験事態の差異が因果モデルのパラメータの使用に影響したかもしれない。また、推論に基準率が反映されないことは基準率錯誤として捉えることもできるが、基準率錯誤の実験事態と本研究の実験事態も異なっている。基準率の使用に関しては Tversky & Kahneman (1974) を始めとして数多くの研究が行われており

(Barbey & Sloman, 2007 for a review), 基準率錯誤では尤度に基づく判断が広く観察されている。しかしながら、本研究ではそのような傾向は見られず、本研究の結果を一般的な基準率錯誤と捉えてよいかについては議論の余地が残されている。

選択の因果モデル理論では、意図的に選択された行動と強制された行動はともに介入として機能し、両者では同じ推論が行われることが想定されている。しかしながら、本研究の実験 3 では、選択に基づく推論で因果強度が考慮されず、介入に基づく推論では因果強度が考慮されていた。介入に基づく推論で因果強度が反映されていたことについては、意図的な行動に対しては自発的に取り組むこと自体の効果が考えられたために、因果強度などの他の情報があまり考慮されなかったかもしれない。このような影響が存在しているならば、両者が完全には同じ機能を果たしているとは言い切れない。

本研究では、選択が介入として機能することを示したが、選択は必ずしも介入となるわけではなく、介入の論理から逸脱することもある。Hagmayer & Sloman (2009) は自身の選択が介入として捉えられていない現象として自己欺瞞 (self-deception) とセルフハンディキャッピング (self-handicapping) を挙げている。前者は自分で行動を選択しているにも関わらず、何らかの他の要因によって行動がなされていると考えることである。Quattron & Tversky (1984) によると健康的な心臓を持つ人は冷水の中に長い時間手を浸けていられることが教示された場合、実験参加者が手を浸す時間が長くなる。手を浸す時間は自身の選択によって決まるにも関わらず、健康的な心臓のために手を浸す時間が長くなると実験参加者は回答していた。一方、後者は悪い結果を導くような障害を自分自身で作り出すことである。障害を作ることによって悪い結果をその障害に帰属させることが可能になり、自尊感情が保たれる。自己欺瞞やセルフハンディキャッピングはどちらも意図的な選択にも関わらず、そのように捉えていないために結果は外的要因に帰属される。これらと対照的に、実際は外的要因によって結果が生じているにも関わらず結果を内的要因つまり自身の行動に帰属するのが制御幻想 (illusion of control) である。

原因帰属の研究では行為者は外的要因に帰属する傾向があり、観察者は行為者の内的要因に帰属する傾向があることが知られており、これは行為者-観察者バイアス (actor - observer bias) と呼ばれている (Jones & Nisbett, 1972). 例えば、試験の結果が悪かった場合テストを受けなかった観察者は行為者の勉強不足と内的に帰属し、テストを受けた行為者は問題が難しかったと外的に帰属することがある. 行為者と観察者における推論の違いは4枚カード問題でも指摘されている (Gigerenzer & Hug, 1992). これらの現象が選択の因果モデル理論で説明可能かについては今後の重要な課題である.

# 第 5 部

## 総合論議

## 12. 総合論議

### 12. 1. 博士論文研究のまとめ

本博士論文研究の目的は、因果構造の学習と因果強度の学習と因果推論の 3 つの側面から因果ベイズネットの心理学的妥当性を検討することであった。研究 1 では因果構造の学習における共変動情報の役割を検討した、2 事象で構成される最も単純な因果構造の判断課題を用い、共変動手がかりと時間順序手がかりのどちらが使用されるか検討した（実験 1）。その結果、共変動手がかりが重視されることが示された。また、課題中に繰り返し構造の判断を求めることによって、共変動手がかりと時間順序手がかりが学習過程でどのように影響を及ぼしているか検討したところ、学習初期では時間順序手がかりが優先され、試行の経過に伴って共変動手がかりが重視されるようになることが明らかとなった（実験 2）。さらに、共変動情報を体系的に操作して共変動手がかりから因果の方向性がどのように導かれるかについて検討を行った結果、微生物の発生に関するカバーストーリーでは因果の必要性に基づく解釈が（実験 3）。宇宙人のコミュニケーションに関するカバーストーリーでは因果の十分性に基づく解釈が確認された（実験 4）。研究 1 の結果から、共変動情報がどのように使用されるかは状況に応じて変化することが示唆された。研究 2 では因果構造の学習において共変動手がかりと時間順序手がかりの使用が課題の複雑性によってどのように変化するか検討した。実験の結果、事象の数が少ない単純な課題では共変動手がかりが、事象の数が多し複雑な課題では時間順序手がかりが重視されることが示された（実験 1 および実験 2）。また、介入による学習ではこのような判断方略の変化は確認されなかった（実験 3）。研究 2 の結果から、共変動手がかりの使用に限界があることや、観察と介入では異なるプロセスが働いていることが示唆された。

研究 3 では因果強度の学習において共変動情報を操作した 9 つの実験から 114 条件をまとめ、規則ベースアプローチや連合的アプローチや因果的アプローチの代表的な

モデルの比較検討を行った。シミュレーションの結果、因果ベイズネットモデルの 1 つである SS パワーモデルがデータに対する最も高い適合度を示しており、その妥当性が示唆された。研究 4 では因果強度の学習において観察による学習と介入による学習のどちらが因果強度の正確な推定をもたらすか検討した。実験の結果、介入による学習の促進効果がみられた（実験 1 および実験 2）。この結果は観察と介入では異なるプロセスが働いていることを示唆するものであった。

研究 5 では因果推論において観察に基づく推論と介入に基づく推論が区別されるか検討した。観察によって因果関係を学習した場合は両者を明確に区別することが示されたのに対して（実験 1）、介入によって因果関係を学習した場合には区別していることを示す十分な証拠は得られなかった（実験 2）。これらの結果から因果ベイズネットには一定の妥当性が認められるものの、その適用には慎重な議論が必要であることが示唆された。研究 6 では介入に基づく因果推論を意思決定場面にまで拡張した選択の因果モデル理論に焦点を当て、選択が介入として機能するか、因果モデルが考慮されるか、因果モデルのパラメータが考慮されるか検討した。その結果、因果モデルは考慮されるが（実験 1）、基準率や因果強度などのパラメータは考慮されないことが示され（実験 2 および実験 3）、選択の因果モデル理論の妥当性が示唆された。

## 12. 2. 理論的含意

ここでは本研究の結果と因果ベイズネットの関係について、それぞれの側面から詳しく述べる。因果構造の学習に関しては、因果ベイズネットによる説明だけでなく、ヒューリスティックスを用いた説明も提唱されている（e.g., Fernbach & Sloman, 2009; Lagnado & Sloman, 2004）。因果ベイズネットが共変動手がかりに重きを置くのに対して、これらのヒューリスティックスでは時間順序手がかりを重視している。時間順序手がかりの使用を示す先行研究（e.g., Lagnado & Sloman, 2006; White, 2006）ではこれらのヒューリスティックスが支持されていたが、本研究では共変動手がかりに基づく因

因果構造の判断が見られており、因果ベイズネットを支持する結果となっている。しかしながら、研究 2 では課題の複雑性によって判断方略が変化することが示されており、どちらか一方のモデルが正しいというより、二重過程として両方機能していると考えの方が妥当である。また、Mayrhofer and Waldmann (2011) は因果の十分性に基づくヒューリスティックスを提唱しているが、研究 1 では必要性和十分性の両方に基づく解釈が見られており、これだけでは説明できない。ベイズモデルを拡張してこれらの説明が可能かどうかは今後の検討課題である。

因果強度の学習では因果ベイズネットの SS パワーモデルがデータに対して最もよく当てはまっており、因果ベイズネットが支持された。また、因果推論においても因果ベイズネット固有の予測と一致する結果が得られ、その妥当性が示唆された。しかしながら、観察と介入については一貫してベイズネットが想定する以上の違いが見られた。ベイズネットでは両者の違いをグラフ切除や無効化として捉えるが、研究 4 では介入によって因果強度の推定が促進されていた。また、意思決定における推論を扱った研究 6 では観察に基づく推論において因果モデルとパラメータが考慮されたのに対して、介入に基づく推論では因果モデルのみが考慮され、そのパラメータは考慮されなかった。このことは選択の因果モデル理論の予測とは一致するが、数理的手法としての因果ベイズネットからは予測されないものであった。これらの結果は観察と介入にはグラフ切除以上の心理的効果を含んだ違いがあることを示唆している。

規則ベースアプローチや連合的アプローチでは因果性を確率や連合に還元しており、因果関係を記述する術を持たないため、因果構造の学習や因果推論にそのモデルを適用することはできない。また、因果強度の学習では因果ベイズネットの SS パワーモデルよりも高い適合度を示すことができなかった。これらの結果は、因果性の学習と推論において規則ベースアプローチや連合的アプローチが因果ベイズネットを採用する因果的アプローチよりも優れている部分が見られないことを意味している。

### 12. 3. 今後の展望

本研究では様々な実験を通して因果ベイズネットの妥当性を検討してきたが、まだ検討できていない部分も多数残されている。第一に、それぞれの過程に関する相互作用が検討されていないことである。本研究では、因果構造の学習と因果強度の学習と因果推論の3つの側面に分けてそれぞれに関して実験を行った。しかしながら、これらの過程は決して独立なものではなく、互いに影響しあうと考えられる。実際、因果構造の知識が因果強度の学習に影響することは手がかり競合の非対称性をはじめとして多数報告されている (e.g., Booth & Buehner, 2007; Waldmann, 2000, 2001; Waldmann & Holyoak, 1992; Waldmann & Walker, 2005)。手続き的に見た場合、これらの研究では因果構造を操作して因果強度の判断を求めているのに対して、研究1では因果強度を体系的に操作して因果構造の判断を求めている。研究1では構造の判断のみを尋ねていたが強度の判断を尋ねることで両者の関係について更なる知見が得られるだろう。また、学習と推論に関しても同様に相互作用の検討が望まれる。

第二に、実験課題に使用する変数が挙げられる。この領域の他の多くの研究と同様に、本研究では二値変数のみ使用していた。現実的には連続変数や多値変数が多く存在するにもかかわらず、連続変数を用いた実験はほとんど行われていない。因果ベイズネットは二値変数だけでなく、連続変数も扱うことが可能であり、今後はこのような変数を用いた実験も必要である。

第三に、因果関係の性質が挙げられる。因果強度の学習におけるメタ分析を除いて、本研究では発生的因果関係が対象とされていた。しかし、因果関係には発生的関係だけでなく、抑制的關係もあり、両者が非対称であるという知見も存在している (Baetu & Baker, 2012)。特に因果構造の学習に関して抑制的因果関係を扱った研究は管見の限り存在しておらず、知見の一般化可能性のためにも、抑制的因果構造の学習の実験が期待される。

第四に、サンプルサイズが挙げられる。実験によって多少の違いはあるものの、本研

究では学習フェイズが 16 試行から 40 試行で構成されていた。現実的には特定の対象に対して何十回も観察を行うことは珍しく、より早い段階で知識を獲得して使用しなければならない。一部の研究ではサンプルサイズが小さいときに相関の検出が容易になることが報告されている (e.g., Kareev, 1995, 2000)。ただし、これらの知見は相関係数や連関係数を基にしたシミュレーションの結果であり、ヒトがどのように関係性を検出するかということを考慮に入れていない。本研究で比較した様々な数理モデルを用いて同様の結果が得られるか確かめる必要がある。

第五に、命題的アプローチ (De Houwer, 2009) との比較が挙げられる。本研究の実験事態では因果ベイズネットと命題推論モデルの比較することはできなかったが、ブロッキングにおける加法性と最大性と呼ばれる現象は命題推論モデルでのみ説明されており、因果ベイズネットで説明可能かどうかははっきりしていない。ブロッキング現象では、ある原因事象  $C_1$  が結果事象  $E$  を発生させているのを観察した後に、 $C_1$  と他の原因事象  $C_2$  が  $E$  を引き起こしているのを観察した場合、 $E$  の原因は  $C_1$  であって  $C_2$  ではないと判断される。このように先行呈示された事象によって後続事象の因果性が低く判断されることは順行ブロッキングと呼ばれている。この現象が生起するためには、それぞれの原因が加算的に作用すること(加法性)、および先行呈示された原因が結果を最大まで引き起こしていること(最大性)が必要であることが知られている。因果ベイズネットのシミュレーションで加法性と最大性の効果が説明可能か否か、因果ベイズネットでも説明可能な場合、命題モデルと因果ベイズネットのどちらのモデルがより妥当であるか検討が望まれる。

#### 12. 4. 終わりに

本博士論文研究では、因果ベイズネットの心理学的妥当性を検討することを目的として、因果構造の学習と因果強度の学習と因果推論の 3 つの側面から実験的検討を行った。その結果、因果ベイズネットの予測と概ね一致する結果が得られ、その心理学

的妥当性が示唆された。ただし、一部の実験では予測と異なる結果も得られており、万能ではないことも同時に示された。しかしながら、因果ベイズネットが因果性の学習と推論の研究を進めていく上で有用な手段となることに間違いないだろう。

## References

- Ahn, W.-K., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition, 54*, 299–352.
- Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15*, 147–149.
- Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin, 114*, 435–448.
- Allan, L. G., & Jenkins, H. M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology, 34*, 1–11.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108*, 441-485.
- Anderson, J. R., & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory & Cognition, 23*, 510-524.
- Aristotle (trans. 1984). *Physics* (Vol.2). In J. Barnes (Ed.), *The Complete Works of Aristotle* (Vol.1). Princeton, NJ: Princeton University Press.
- Arcediano, F., Matute, H., Escobar, M., & Miller, R. R. (2005). Competition between antecedent and between subsequent stimuli in causal judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 228–237.
- Baetu, I., & Baker, A. G. (2009). Human judgments of positive and negative causal chains. *Journal of Experimental Psychology. Animal Behavior*

*Processes*, 35, 153–168.

- Baker, A. G., Mercier, P., Vallée-Tourangeau, F., Frank, R., & Pan, M. (1993). Selective associations and causality judgments: Presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 414–432.
- Barberia, I., Baetu, I., Sansa, J., & Baker, A. G. (2010). Choosing optimal causal backgrounds for causal discovery. *Quarterly Journal of Experimental Psychology*, 63, 2413–2431.
- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, 30, 241–254.
- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44, 211–233.
- Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53, 370–418.
- Beckers, T., De Houwer, J., Pineño, O., & Miller, R. R. (2005). Outcome additivity and outcome maximality influence cue competition in human causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 238–249.
- Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, 3, 385–402.
- Blaisdell, A. P., Sawa, K., Leising, K. J., & Waldmann, M. R. (2006). Causal reasoning in rats. *Science*, 311, 1020–1022.
- Bonawitz, E. B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A. N., Woodward, J., & Schulz, L. E. (2010). Just do it? Investigating the gap

- between prediction and action in toddlers' causal inferences. *Cognition*, *115*, 104-117.
- Booth, S. L., & Buehner, M. J. (2007). Asymmetries in cue competition in forward and backward blocking designs: Further evidence for causal model theory. *Quarterly Journal of Experimental Psychology*, *60*, 387-399.
- Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1119-1140.
- Burns, P., & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning*, *15*, 167-196.
- Carroll, C. D., Cheng, P. W., & Lu, H. (2013). Inferential dependencies in causal inference: a comparison of belief-distribution and associative approaches. *Journal of Experimental Psychology: General*, *142*, 845-863.
- Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 837-854.
- Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, *18*, 537-545.
- Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*, 335-344.
- Chater, N., Tenenbaum, J. B., & Yuille, A. (Eds.). (2006). Probabilistic models of cognition [Special issue]. *Trends in Cognitive Sciences*, *10*.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405.

- Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365-382.
- Collins, D. J., & Shanks, D. R. (2006). Conformity to the power PC theory of causal induction depends on type of probe question. *Quarterly Journal of Experimental Psychology*, *59*, 225-232.
- Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine learning*, *9*, 309-347.
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*, 294-300.
- Cox, R. T. (1946). Probability, frequency and reasonable expectation. *American Journal of Physics*, *14*, 1-13.
- Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology*, *47*, 109-121.
- de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives [Foresights: Its logical laws, its subjective source]. *Annales de l'institut Henri Poincaré*, *7*, 1-68.
- De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1-20.
- De Houwer, J., & Beckers, T. (2002a). A review of recent developments in research and theories on human contingency learning. *The Quarterly Journal of Experimental Psychology*, *55B*, 289-310.
- De Houwer, J., & Beckers, T. (2002b). Higher-order retrospective reevaluation in human causal learning. *The Quarterly Journal of Experimental Psychology*, *55B*, 137-151.
- Dennis, M. J., & Ahn, W.-K. (2001). Primacy in causal strength judgments:

- The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition*, *29*, 152–164.
- Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65–117). Mahwah, NJ: Erlbaum.
- Descartes, R. (1991). *Principles of philosophy* (V. R. Miller & R. P. Miller, Trans.). Netherlands: Kluwer Academic. (Original work published 1644)
- Dey, D. K., & Rao, C. R. (2005). *Handbook of statistics 25: Bayesian thinking, modeling, and computation*. Amsterdam, Netherlands: Elsevier. (繁榘算男・岸野洋久・大森裕浩(監訳)(2011). ベイズ統計分析ハンドブック. 東京: 朝倉書店.)
- Dickinson, A. & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *The Quarterly Journal of Experimental Psychology*, *49B*, 60–80.
- Dickinson, A., Shanks, D. R., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *36A*, 29–50.
- Dwyer, D. M., Starns, J., & Honey, R. C. (2009). “Causal reasoning” in rats: A reappraisal. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 578-586.
- Eberhardt, F., & Scheines, R. (2007). Interventions and causal inference. *Philosophy of Science*, *74*, 981-995.
- Eells, E. (1991). *Probabilistic causality*. Cambridge, United Kingdom: Cambridge University Press.

- Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99*, 3–19.
- Estes, W. K., & Skinner, B. F. (1941) Some quantitative properties of anxiety. *Journal of Experimental Psychology, 29*, 390–400.
- Enkvist, T., Newell, B., Juslin, P., & Olsson, H. (2006). On the role of causal intervention in multiple-cue judgment: Positive and negative effects on learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 163–179.
- Fernbach, P. M., Darlow, A., & Sloman, S. a. (2011). Asymmetries in predictive and diagnostic reasoning. *Journal of Experimental Psychology: General, 140*, 168–185.
- Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 678–693.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representations. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 330–344.
- Fishburn, P. (1986). The axioms of subjective probability. *Statistical Science, 1*, 335–345.
- Garcia-Retamero, R., & Hoffrage, U. (2006). How causal knowledge simplifies decision-making. *Minds and Machines, 16*, 365-380.
- Garcia-Retamero, R., Hoffrage, U., Müller, S. M., & Maldonado, A. (2010). The influence of causal knowledge in two-alternative forced-choice tasks. *Open Psychology Journal, 3*, 136-144.
- Garcia-Retamero, R., Müller, S. M., Catena, A., & Maldonado, A. (2009). The power of causal beliefs and conflicting evidence on causal judgments

- and decision making. *Learning and Motivation*, *40*, 284-297.
- Garcia-Retamero, R., Wallin, A., & Dieckmann, A. (2007). Does causal knowledge help us be faster and more frugal in our decisions? *Memory & Cognition*, *35*, 1399-1409.
- Chater, N., & Vita'nyi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Science*, *7*, 19–22.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*, 127-171.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.
- Glymour, C. (2003). Learning, prediction, and causal Bayes nets. *Trends in Cognitive Sciences*, *7*, 43-48.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3–32.
- Gopnik, A., & Schulz, L. E. (2007). *Causal learning: Psychology, philosophy, and computation*. New York, NY: Oxford University Press.
- Gopnik, A., & Tenenbaum, J. B. (Eds.). (2007). Bayesian networks, Bayesian learning and cognitive development [Special section]. *Developmental Science*, *10*, 281–364.
- Green, D. W., & Over, D. E. (2009). Valuing intervention and observation. *Quarterly Journal of Experimental Psychology*, *62*, 1010–1022.

- Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011). Bayes and blickets: effects of knowledge on causal induction in children and adults. *Cognitive Science*, *35*, 1407–1455.
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*, 211–244.
- Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*, 334–384.
- Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*, 661–716.
- Hagmayer, Y., & Meder, B. (2008). Causal Learning Through Repeated Decision Making. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 179-184). Austin TX: Cognitive Science Society.
- Hagmayer, Y., Meder, B., Osman, M., Mangold, S., & Lagnado, D. (2010). Spontaneous causal learning while controlling a dynamic system. *The Open Psychology Journal*, *3*, 145-162.
- Hagmayer, Y., & Sloman, S. A. (2005). A causal model theory of choice. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 881-886). Mahwah, NJ: Erlbaum.
- Hagmayer, Y., & Sloman, S. A. (2006). Causal vs. evidential decision making in Newcomb's paradox. In R. Sun, & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Hagmayer, Y., & Sloman, S. A. (2009). Decision makers conceive of their choices as interventions. *Journal of Experimental Psychology: General*, *138*, 22-38.

- Hájek, A. (2003). Interpretations of probability. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2012 ed.). Retrieved from <http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>
- Hannah, S. D., & Beneteau, J. L. (2009). Just tell me what to do: Bringing back experimenter control in active contingency tasks with the command-performance procedure and finding cue density effects along the way. *Canadian Journal of Experimental Psychology*, *63*, 59–73.
- Hattori, M., & Nishida, Y. (2009). Why does the base rate appear to be ignored? The equiprobability hypothesis. *Psychonomic Bulletin & Review*, *16*, 1065–1070.
- Hattori, M., & Oaksford, M. (2007). Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science*, *31*, 765–814.
- Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine learning*, *20*, 197-243.
- Holland P. C. (2004) Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 104 –117.
- Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology*, *62*, 135-163.
- Hume, D. (2003). *A treatise of human nature*. New York, NY: Dover Publications. (Original work published 1748)
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal*

- operational structures* (A. Parsons & S. Milgram, Trans.). London: Routledge & Kegan Paul.
- Jara, E., Vila, J., & Maldonado, A. (2006). Second-order conditioning of human causal learning. *Learning and Motivation*, *37*, 230–246.
- Jenkins, H., & Ward, W. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs*, *7*, 1-17.
- Johnson-Laird, P. N., & Shafir, E. (1993). The interaction between reasoning and decision making: An introduction. In P. N. Johnson-Laird, & E. Shafir (Eds.), *Reasoning and decision making* (pp.1-9). Oxford: Blackwell.
- Jones, E. E. and Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of the behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Morristown, NJ: General Learning Press.
- Kant, I. (2003). *Critique of pure reason* (J. M. D. Meiklejohn, Trans.). New York, NY: Dover Publications. (Original work published 1781)
- Kao, S.-F., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1363–1386.
- Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, *56*, 263–269.
- Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, *107*, 397–402.
- Kass, R. E., & Rafferty, A. E. (1995). Bayes factors. *Journal of the American*

*Statistical Association, 90, 773-795.*

川合伸幸・久保(川合)南海子 (2008). ヒトと動物の回顧的推論について. 認知科学, *15, 378-391.*

Kelley, A. M., & Athy, J. R. (2009). *The effects of observation and intervention on the judgment of causal and correlational relationships* (USAARL Report No. 2009-16). Fort Rucker, AL: U. S. Army Aeromedical Research Laboratory.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28, 107-128.*

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review, 116, 20-58.*

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55, 271-304.*

Khemlani, S. S., & Oppenheimer, D. M. (2010). When one model casts doubt on another: A levels-of-analysis approach to causal discounting. *Psychological Bulletin, 137, 195-210.*

Klein, G. A. (1999). *Sources of power: How people make decisions.* Cambridge, MA: MIT Press.

Knill D, C., & Richards, W. (1996). *Perception as Bayesian Inference.* Cambridge, United Kingdom: Cambridge University Press.

Körding, K., & Wolpert, D. (2004). Bayesian integration in sensorimotor learning. *Nature, 427, 244-247.*

Kousta, S. (Ed.), (2010). Approaches to cognitive modeling [Special section]. *Trends in Cognitive Sciences, 14, 339-364.*

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and*

- Cognition*, 25, 1083–1119.
- Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgement under uncertainty. *Journal of Experimental Psychology: General*, 136, 430–450.
- Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological Science*, 16, 678–683.
- Kushnir, T., Wellman, H. M., & Gelman, S. A. (2009). A self-agency bias in preschoolers' causal inferences. *Developmental Psychology*, 45, 597–603.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32, 311–328.
- Lagnado, D. A., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 856–876.
- Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 451–460.
- Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). Oxford: Oxford University Press.
- Leibniz, G. W. (1896). *New essays on human understanding* (A. G. Langley, Trans.). London: Macmillan. (Original work published 1765)
- Leising, K. J., Wong, J., Waldmann, M. R., & Blaisdell, A. P. (2008). The special status of actions in causal reasoning in rats. *Journal of Experimental Psychology: General*, 137, 514–527.
- Lewis, D. (1973). Causation. *The Journal of Philosophy*, 70, 556–567.

- Lober, K., & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, *107*, 195–212.
- López, F. J., Shanks, D. R., Almaraz, J., & Fernández, P. (1998). Effects of trial order on contingency judgments: A comparison of associative and probabilistic contrast accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 672–694.
- Lu, H., Yuille, A., Liljeholm, M., Cheng, P.W., & Holyoak, K.J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*, 955-984.
- Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford, United Kingdom: Clarendon Press.
- Maldonado, A., Jiménez, G., Herrera, A., Perales, J. C., & Catena, A. (2006). Inattentional blindness for negative relationships in human causal learning. *Quarterly Journal of Experimental Psychology*, *59*, 457–470.
- Malebranche, N. (1997). *The search after truth* (T. M. Lennon & P. J. Olscamp, Trans). Cambridge, United Kingdom: Cambridge University Press. (Original work published 1674-1675)
- Mandel, D. R., & Vartanian, O. (2009). Weighting of contingency information in causal judgement: evidence of hypothesis dependence and use of a positive-test strategy. *Quarterly Journal of Experimental Psychology*, *62*, 2388–2408.
- Marsh, J. K., & Ahn, W.-K. (2006). Order effects in contingency learning: The role of task complexity. *Memory & Cognition*, *34*, 568–576.
- 松原望 (2008). 入門ベイズ統計—意思決定の理論と発展—. 東京: 東京図書.
- Matute, H. (1996). Illusion of control: Detecting response-outcome independence in analytic but not in naturalistic conditions. *Psychological Science*, *7*, 289–293.

- Mayrhofer, R., & Waldmann, M. R. (2011). Heuristics in covariation-based induction of causal models: Sufficiency and necessity priors. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3110-3115). Austin, TX: Cognitive Science Society.
- McCormack, T., Hoerl, C., & Butterfill, S. (Eds.). (2011). *Tool use and causal cognition*. Oxford: Oxford University Press.
- Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal*, *3*, 119–135.
- Meder, B., & Hagmayer, Y. (2009). Causal induction enables adaptive decision making. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 1651-1656). Austin, TX: Cognitive Science Society.
- Meder, B., Hagmayer, Y., & Waldmann, M. R. (2008). Inferring interventional predictions from observational learning data. *Psychonomic Bulletin & Review*, *15*, 75-80.
- Meder, B., Hagmayer, Y., & Waldmann, M. R. (2009). The role of learning data in causal reasoning about observations and interventions. *Memory & Cognition*, *37*, 249-264.
- Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*, 277–301.
- Mill, J. S. (1843). *A system of logic: Ratiocinative and inductive* (8th ed.). New York, NY: Harper & Brothers.
- Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 51–92). San Diego,

CA: Academic Press.

Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, *102*, 331–355.

Neapolitan, R. E. (2004). *Learning Bayesian networks*. Upper Saddle River, NJ: Prentice Hall.

Nichols, W., & Danks, D. (2007). Decision making using learned causal structures. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual meeting of the cognitive science society* (pp. 1343-1348). Austin, TX: Cognitive Science Society.

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review*, *111*, 455–485.

Nozick, R. (1969). Newcomb's Problem and two principles of choice. In N. Rescher (Ed.), *Essays in honor of Carl G. Hempel* (pp. 114-146). Dordrecht, The Netherlands: Reidel.

大塚淳 (2010). ベイズネットから見た因果と確率. 科学基礎論研究, *38*, 39-47.

Osman, M., & Speekenbrink, M. (2012). Prediction and control in a dynamic environment. *Frontiers in Psychology*, *3*: 68.

Park, J., & Sloman, S. a. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology*, *67*, 186–216.

Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, *94*, 61–73.

Pearl, J. (1985). Bayesian networks: A model of self-activated memory for evidential reasoning. *Proceedings of the 7th Annual Conference of the Cognitive Science Society* (pp.329-334).

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of*

- plausible inference*. San Francisco, CA: Morgan Kaufmann Publishers.
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, United Kingdom: Cambridge University Press.
- Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology, 58*, 97–118.
- Perales, J. C., & Shanks, D. R. (2003). Normative and descriptive accounts of the influence of power and contingency on causal judgment. *Quarterly Journal of Experimental Psychology, 56A*, 977–1007.
- Perales, J. C., & Shanks, D. R. (2004, May). *The cause–density effect as a tool to discriminate between causal learning models*. Paper presented at the Special Interest Meeting on Human Contingency Learning, Lignely, Belgium.
- Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review, 14*, 577–596.
- Pineño, O., & Miller, R. R. (2007). Comparing associative, statistical, and inferential reasoning accounts of human contingency learning. *Quarterly Journal of Experimental Psychology, 60*, 310–329.
- Price, P. C., & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition, 21*, 561–572.
- Psillos, S. (2005). Causality. In M. Horowitz (Ed.), *New dictionary of the history of ideas*. New York, NY: Charles Scribner's Sons.
- Quattrone, G., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology, 46*, 237–248.

- Reed, P. (2001). Schedules of reinforcement as determinants of human causality judgments and response rates. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 187–195.
- Rehder, B. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology*, *72*, 54–107.
- Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology*, *50*, 264–314.
- Reichenbach, H. (1956). *The direction of time*. Berkeley, CA: University of California Press.
- Reips, U.-D., & Waldmann, M. R. (2008). When learning order affects sensitivity to base rates: Challenges for theories of causal learning. *Experimental Psychology*, *55*, 9–22.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp.64-99). New York: Appleton-Century-Crofts.
- Rottman, B. M., & Ahn, W.-K. (2011). Effect of grouping of evidence types on learning about interactions between observed and unobserved causes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1432–1448.
- Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*, 109–139.
- Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*, 93–125.
- 佐金武 (2008). 因果とは何かをめぐる哲学的論争(1)—D.ルイスの反事実的条件法

- による分析とその批判一. 哲学論叢, 35, 82-93.
- 斎藤元幸・嶋崎恒雄 (2009). 因果推論におけるパワーPC理論について. 人文論究, 59(3), 1-21.
- Saito, M., & Shimazaki, T. (2011). Causal reasoning in decision making: The role of causal models and their parameters. In B. Kokinov, A. Karmiloff-Smith, & N. J. Nersessian (Eds.), *European Perspectives on Cognitive Science* (280, pp.1-6). Sofia, Bulgaria: New Bulgarian University Press.
- 斎藤元幸・嶋崎恒雄 (2012). 因果推論に対する因果ベイズネットの適用の妥当性について—観察と介入の違いに関する実験的検討—. 認知科学, 19, 380-390.
- Sawa, K. (2009). Predictive behavior and causal learning in animals and humans. *Japanese Psychological Research*, 51, 222-233.
- Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, 10, 322-332.
- Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, 110, 101-120.
- Shanks, D. R. (1985a). Continuous monitoring of human contingency judgment across trials. *Memory & Cognition*, 13, 158-167.
- Shanks, D. R. (1985b). Forward and backward blocking in human contingency judgement. *The Quarterly Journal of Experimental Psychology*, 37, 1-21.
- Shanks, D. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology*, 84, 319-354.
- Shanks, D. R. (2002). Tests of the power PC theory of causal induction with

- negative contingencies. *Experimental Psychology*, *49*, 81-88.
- Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology*, *60*, 291-309.
- Shanks, D. R., Holyoak, K. J., & Medin, D. L. (Eds.). (1996). The psychology of learning and motivation: Causal learning (Vol. 34). San Diego, CA: Academic Press.
- 繁榘算男 (1985). ベイズ統計入門. 東京: 東京大学出版会.
- Shimazaki, T., Tsuda, Y., & Imada, H. (1991). Strategy changes in human contingency judgments as a function of contingency tables. *The Journal of general psychology*, *118*, 349-360.
- Simon, H. A. (1957). *Models of man: Social and rational*. New York, NY: Wiley.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22.
- Sloman, S. A. (2014). Two systems of reasoning, an update. In Sherman, J., Gawronski, B., & Trope, Y. (Eds.). *Dual process theories of the social mind*. Guilford Press.
- Sloman, S. A. (2005). *Causal models: how people think about the world and its alternatives*. New York, NY: Oxford University Press.
- Sloman, S. A., & Hagmayer, Y. (2006). The causal psychologic of choice. *Trends in Cognitive Sciences*, *10*, 407-412.
- Sloman, S. A., & Lagnado, D. A. (2005). Do we 'do'? *Cognitive Science*, *29*, 5-39.
- Sobel, D. M. & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34*, 411-419.
- Sobel, D. M., & Sommerville, J. A. (2009). Rationales in children's causal

- learning from others' actions. *Cognitive Development*, 24, 70–79.
- Sobel, D. M., & Sommerville, J. A. (2010). The importance of discovery in children's causal learning from interventions. *Frontiers in psychology*, 1: 176.
- 園信太郎 (2014). 確率概念の近傍—ベイズ統計学の基礎をなす確率概念—. 東京: 内田鶴圃.
- Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, 7, 337-342.
- Spirtes, P., Glymour, C., & Scheines, P. (1993). *Causation, prediction, and search*. New York, NY: Springer-Verlag.
- Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, 27, 453-489.
- Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, 114, 759-783.
- Suppes, P. (1970). *A probabilistic theory of causation*. Amsterdam, Netherlands: North-Holland.
- Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human Pavlovian-instrumental transfer. *The Journal of Neuroscience*, 28, 360–368.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331, 1279–1285.
- 戸田正直 (1982). 主観確率の測定法について. 行動計量学, 9, 1-11.

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.
- Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social Psychology* (Vol. 1, 49-72). Hillsdale, NJ: Lawrence Erlbaum.
- 植野真臣 (2013). ベイジアンネットワーク. 東京: コロナ社.
- Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *Quarterly Journal of Experimental Psychology, 60*, 433-447.
- Vallée-Tourangeau, F., Murphy, R. A., Drew, S., & Baker, A. G. (1998). Judging the importance of constant and variable candidate causes: A test of the power PC theory. *Quarterly Journal of Experimental Psychology, 51A*, 65-84.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and motivation, 25*, 127-151.
- Verma, T. S., & Pearl, J. (1990). Equivalence and synthesis of causal models [Technical report R-150]. *Department of Computer Science, University of California, Los Angeles*.
- Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 34, pp. 47-88). San Diego, CA: Academic Press.
- Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 53-76.
- Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin &*

*Review, 8*, 600-608.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 216-227.

Waldmann, M. R., Hagmayer, Y., & Blaisdell, A. P. (2006). Beyond the information given: Causal models in learning and reasoning. *Current Directions in Psychological Science, 15*, 307-311.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222-236.

Waldmann, M. R., & Walker, J. M. (2005). Competence and performance in causal learning. *Learning & Behavior, 33*, 211-229.

Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 174-188.

Wasserman, E. A., Kao, S.-F., Van Hamme, L. J., Katagiri, M., & Young, M. E. (1996). Causation and association. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 34, pp. 207-264). San Diego, CA: Academic Press.

Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology, 19*, 231-241.

渡部洋 (1999). ベイズ統計学入門. 東京: 福村出版.

White, P. A. (1990). Ideas about causation in philosophy and psychology. *Psychological Bulletin, 108*, 3-18.

White, P. A. (2003). Making causal judgments from the proportion of confirming instances: The *pCI* rule. *Journal of Experimental*

*Psychology: Learning, Memory, and Cognition*, 29, 710–727.

White, P. A. (2006). How well is causal structure inferred from cooccurrence information? *European Journal of Cognitive Psychology*, 18, 454-480.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General*, 136, 82-111.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford, UK: Oxford University Press.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114, 245–272.

Xu, F., & Griffiths, T. L. (2011). Probabilistic models of cognitive development [Special issue]. *Cognition*, 120, 299–412.

Yarritu, I., Matute, H., & Vadillo, M. A. (2014). Illusion of control: The role of personal involvement. *Experimental Psychology*, 61, 38–47.

Zsombok, C. E., & Klein, G. (1997). *Naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum.

## 謝辞

本博士論文の作成にあたり、多くの方からご指導ご鞭撻を賜りました。嶋崎恒雄先生には、学部生の頃から今日まで卒業論文や修士論文をはじめとして多くの叱咤激励を頂きました。分かりやすい説明を大切にする嶋崎先生からは研究発表の仕方についてお叱りを受けることも多々ありましたが、研究内容についてはいつも面白いとおっしゃってくださり、そのことは研究活動の何よりの励みとなりました。中島定彦先生には、大学院入学時よりゼミに出席させて頂き、研究者として知っておくべき様々なことについてご教授頂きました。学生にきめ細やかな指導をしつつも、多くの論文を出版し続ける中島先生の姿勢は私の目指す理想の研究者像となりました。立命館大学の服部雅史先生には、ご多忙中にも関わらず博士論文の審査を快諾して頂き、今後検討すべき課題について多くの示唆を与えて頂きました。楽しく読ませてもらいましたという先生のお言葉に感激しつつも、明確な回答を用意できなかった質疑にはより一層精進しなければと身の引き締まる思いが致しました。哲学倫理学専修の久米暁先生には因果性の哲学的背景について様々な文献を紹介して頂きました。

研究室の先輩、同期の友人、後輩には、何度も研究の話聞いてもらって様々なコメントを頂きました。話が逸れて雑談となることも多々ありましたが、それはそれで大変楽しいものでした。関西学院からは私の研究に対して様々な形でご支援を頂きました。実験好きな私がほぼ毎日実験することができたのは心理科学研究室の充実した設備のおかげです。また、多くの方々に実験に参加して頂いたおかげで研究活動を進めることができました。皆さまのご支援がなければ、博士論文が完成することはありませんでした。ここに心からの感謝を記します。

斎藤 元幸