報告番号　甲　㊀　第　　号　　氏名 安西　祐一郎

論文題目 The Epistemology of Learning and Interaction: A Goal-Directed Adaptive Agent is an Epistemic Agent
(学習と相互作用の認識論：目標指向の適応的行為主体は認識的行為主体である)

論文審査担当者

主査	慶應義塾大学文学部教授 文学研究科委員	岡田 光弘
副査	慶應義塾大学文学部教授 文学研究科委員	柏端 達也
副査	名古屋大学大学院情報学研究科教授	戸田山 和久
副査	名古屋大学大学院情報学研究科教授	三輪　和久
学識確認	慶應義塾大学文学部教授 文学研究科委員	岡田 光弘

論文概要

　認知科学的「目標指向の適応的行為主体」が哲学的認識論の意味での「認識主体」であるというテーゼを、いくつかの付帯条件を明確にしながら提示している。特に、認識についてのプロセス指向構成主義の立場を導入し、「学習」(learning) と「相互作用」(interaction) という人間の２つの根源的な活動をもとに、このことを示している。また、哲学的認識論の文脈におけるこのテーゼの重要性について議論を展開している。

論文の構成は次の通りである。

各章の概要

まず序章において、本論文の目的、論文作成の動機、および第１章から第３章の概要を述べ、さらに認識論への本論文の貢献について予めまとめている。

第１章では、認知科学分野における学習の諸研究について、本論文の著者が関与

してきた多くの研究を含め、特に理論とモデルに関わる代表的な研究について議論している。

1.1 節で学習の基本概念について述べたのち、1.2 節では、認知科学における学習の研究を広く概観している。また、1.3 節で、近年の学習の研究に情報科学が大きな影響を与えてきたことを指摘している。

次に、1.4 節において、著者が長年にわたり主導してきた「することによる学習の理論 (The theory of learning by doing)」に関する研究成果の概要とその意味について述べ、1.5 節ではさらに、著者が同理論を拡張して提起した「することによる学習に基づく問題理解」のプロセスに関する研究成果を述べている。これら２つの節を通して、著者は、問題解決の行為主体による問題解決方略の学習プロセスが新たな知識を発見するプロセスであること、および、行為主体によって学習される方略の構造が行為の系列の構造として特徴づけられることを指摘している。

1.7 節では、認知科学における最近の学習研究の概要を述べるとともに、1.4 節と 1.5 節に述べた研究成果と近年の学習研究の間の多様な関係について、総括的な提示を行っている。1.8 節は第１章全体の結言である。

第２章では、認知科学分野における相互作用の諸研究とその意味について、特に理論とモデルに関わる研究を中心として、著者自身による研究を含め、概要を述べている。

2.1 節で相互作用の基本概念を提示したのち、2.2 節で、相互作用に関する認知科学的研究が理解とデザインを軸とする新しい方法論に変化してきたことを指摘している。2.3 節では、著者が長年主導してきた人間・ロボット間相互作用の研究成果について述べ、2.4 節では、2.3 節で得た結果を踏まえ、人間・ロボット間相互作用の本質が「情報の共有」にあることを論じている。

そのうえで、2.5 節において、相互作用の理論として著者が提唱してきた「情報共有による相互作用の理論 (The theory of interaction by information sharing)」の研究成果を説明するとともに、第３章の認識論的議論の基礎となる、「行為主体」の概念をはじめとする多くの基本概念を定義している。さらに 2.6 節では、同理論が相互作用の認知科学的諸研究にもたらす意味を明らかにし、特に、知識、信念、意図、行為、欲求、目標等の諸概念を再検討して、認知科学的な意味での知識の４条件を提示している。

また、2.7 節で、2.3 節に述べた人間・ロボット間相互作用の研究が 2.5 節に述べた情報共有による相互作用の理論を産み出した経緯を説明し、2.8 節では、同理論を踏まえて相互作用に関する最近の認知科学的研究を概観している。続く 2.9 節では、相互作用の認知神経科学的基盤の機能的ネットワークモデルを提示する。2.10 節では、本章の記述を基に、相互作用を情報の交換ではなく共有とみなす立場をあらためて示している。2.11 節は本章の結言である。

第３章では、第１章に述べた学習の理論とその意味、また第２章に述べた相互作用の理論とその意味についての議論を受けて、実在論に立脚しつつ内的表象の構成プロセスに注目した「プロセス指向構成主義 (process-oriented constructivism)」の立場を導入し、「目標指向の適応的行為主体は認識的行為主体である」というテーゼについて肯定的な議論を提示する。また、哲学的認識論の文脈のもとで、このテーゼの重要性をめぐる議論を展開する。

まず 3.1 節で、表象の概念およびプロセス指向構成主義の基本的考え方を述べるとともに、６つの認識論的規範を導入する。3.2 節では、表象、表象される対象、表象主体の三項関係を表象の構造として導入し、特に表象主体の概念について、Dretske, Giere, Millikan らの議論に言及しつつ、その必要性を主張している。3.3 節では、内的表象の構造を検討し、表象の記述形式として「構造形式」と「表示形式」の概念を導入して、内的表象の構造と内容の関係を議論している。3.4 節ではさらに、知識の構造化に関する制約、表象主体による表象の構成方略、表象主体にとっての世界と表象の構造的関係等について論じている。3.5 節では特に、表示形式の選択方略が経験科学の研究において重要であることを指摘し、3.6 節において、著者らによるいくつかの認知科学的研究を引用し、科学者にとって表示形式の選択

方略が理論やモデルを構成する鍵となることを例示している。
次に、3.7 節において、プロセス指向構成主義と従来からの認識論の議論とを対比し、実在論に立脚しつつ内的表象の構成プロセスに注目したプロセス指向構成主義の特徴を挙げている。3.8 節では、プロセス指向構成主義と伝統的な知識の理論の関係を検討し、いくつかの条件のもとで「(認知科学的な) 知識条件を満たす信念は (認識論における標準的な) 真理条件を満たす」ことを示している。3.9 節では、第１章の学習の理論と第２章の相互作用の理論がともに「手続き的な科学理論」であることを示している。また、多様な構造を持つ内的表象について、特定の領域に熟達した行為主体が構成する表象の対象が実在することを、収束的実在論 (convergent realism) 等を援用して主張している。そのうえで、行為主体の内部に起源を持つ情報と外部に持つ情報の両方に基づくプロセス指向構成主義によって、実在論と反実在論の二項対立的議論を超えて実在論的議論が可能になることを示唆している。3.10 節では、プロセス指向構成主義の立場のもとで、2.5 節で与えた知識条件の意味でいわゆる KK 仮説が肯定されることを示している。また、内的表象の構成プロセスにおける相似性の推論について議論するとともに、機能主義や計算主義を実在論の立場から批判している。
3.11 節では、「いかにしての信念 (belief-how)」が「いかにしての知識 (knowledge-how)」に変換されるプロセスについて、Stanley と Williamson による知性主義的な議論を引用しつつ、プロセス指向構成主義の立場に立てばその議論がより明確に可能になることを示している。3.12 節では、行為や因果性の認識論において Anscombe や Davidson をはじめとして長く議論されてきた、原因、理由、行為、およびそれらの間の関係について、同一の行為主体が、内的状態と行為の関係を、因果関係、帰属関係、目的論的関係、合理的関係等の関係として多様に解釈し得ることを指摘している。3.13 節では、「社会的行為」の概念について Bratman による shared agency の議論 (集合的概念を導入しない) と Gilbert による joint commitment の議論 (集合的概念を導入する) に言及しつつ、集合的概念を導入することなく社会的相互作用の認識論的な説明が可能なことを主張している。3.14 節では、Quine らを嚆矢とする認識論の自然化をめぐる議論に関し、プロセス指向構成主義を媒介として規範的認識論と記述的認知科学が共存する、中庸な自然化認識論を主張している。
3.15 節では、それまでの議論を踏まえ、プロセス指向構成主義の立場に立てば、(1)「目標指向の適応的行為主体は認識的行為主体である」こと、(2)行為主体の内部に起源を持つ情報と外部に起源を持つ情報の両方を用いて内的表象を構成する機能は、目標の達成を指向し環境の変化に適応する行為主体にとっても、知識と真理を探究する認識的行為主体にとっても必要な前提条件であること、(3)行為主体に独立な世界の存在を認めるべきであり、また、多様な構造を持つ内的表象の対象がその世界に実在すると考えるべきであること、したがって、(4)プロセス指向構成主義が行為主体による内的表象の構成プロセスに焦点を当てた認識論的実在論として位置づけられることを主張している。
最後に、「結言と認識論への貢献」として、第１～３章の要約とともに、本論文による認識論への 10 項目の貢献を列挙している。

審査要旨

「知識とは何か」、「真理とは何か」、「人が知識を得るとはどういうことか」等、知識や真理に関わる根本問題を問う認識論は、古代から哲学の中核分野の一つであった。一方で、２０世紀半ばに誕生した認知科学は、情報の概念を基に心のはたらきを科学的に解明することを目的として、「環境の変化に適応しつつ複雑な目標の達成に向けて心が働くように動機づけられた人間」の探究を行ってきた。伝統的認識論と新しい認知科学との間の適切な関係と架橋可能性については、機能主義などに関する一部の話題を除き、本格的に取り上げて検討されることはこれまであまり

なされてこなかった。本論文で著者は、認知科学の経験科学的アプローチが伝統的認識論の規範的「認識」の根本問題にどのように答え得るかという難問に正面から挑戦している。そして、人間の根源的な活動である「学習」と「相互作用」の観点から、「(認知科学的な意味での) 目標指向の適応的行為主体は (哲学的認識論の言う意味での) 認識主体である」というテーゼを立て、付帯条件を明確にすることを通じて、そのテーゼの妥当性を論証している。これは現代認識論への新たな貢献である。すなわち、伝統的認識論の内部で自己完結するこれまでの標準的認識論の枠を超え、学習や相互作用の経験科学的観点を取り入れた新しい規範的認識論が本論文においては体系的に提案されており、その学術的価値はきわめて高いものであると評価することができる。

本論文では、まず、内的表象が認知科学的な意味で「知識」であるための条件として活用可能性などの４条件が独自の観点から提示され、さらに、「目標指向の適応的行為主体」はその意味での「知識」を持つことが要請されると論じられる。他方で、内的表象が (哲学的) 規範的認識論の意味での「知識」であるための条件 (正当化条件、信頼性条件など) が示され、認知科学的「知識」の条件と認識論的「知識」の条件との関係が詳細に検討される。そして、それらの議論を通じて、上記テーゼの成立が導出されている。これらの議論展開の仕方はきわめて独創的である。その展開の要となっているのは、「知識」としての「内的表象」構成プロセスに対する著者独自の立場の提案である。そこにおいては、実在論に立脚しつつ、内的表象を構成する「プロセス指向構成主義」の立場が詳細に提案され、また、環境変化に適応しつつ内的表象を構成・制御・調整して認識を成立させる構造とその諸付帯条件をめぐる議論が展開されている。論理的に構築された、以上の議論は十分な説得力を持つものと言える。

さらに本論文では、「プロセス指向構成主義」の表象構成論を支える一般的な諸規範が、驚くほど簡潔な仕方で明示されている。そして、認知科学的「知識」の条件を満たす表象が科学的「理論」の条件を満たすこと、およびそのことから、それが認識論的「知識」の条件をも満たすことが、科学哲学の諸知見を援用して明らかにされている。これら主要テーゼの妥当性を示す各段階の議論もまた、緻密かつ明解であり、独創的である。

本論文は、上記テーゼの妥当性だけでなく、伝統的認識論における諸論争に対して新しい知見を与えることにも成功している。例を挙げれば、内的表象の対象についての実在論-反実在論論争に対し、外部情報と内的表象の関係性に準拠した著者の「プロセス指向構成主義」は、これまでの論争点を超える新たな実在論の立場の提案であるとみなすことができる。また本論文の立場は、相互作用を通じた共有表象に関し、構成的経験主義などの反実在論を批判し、一人称推論と類似性推論のみに準拠する非集合的表象構成プロセスを主張するものであるが、ここにも高い独創性が見られ、著者の新たな実在論的見方に拡がりを与えたものとなっている。そしてまた、初心者が熟達者になっていく学習の過程を、信念が知識に変換される内的プロセスとして、科学哲学における実在論の議論に依拠して主張している点にも著者の独創性が見られる。さらに、認識論の「自然化」に対し、その方向の有効性を認めつつ、「自然化」によっては捉えられない認識論の「規範的」側面があり、それこそが認識論にとっては重要であるということを、認知科学を出発点として指摘するくだりには、大きな啓発性と説得力がある。

一方で審査委員会は以下のような点を、今後の研究課題として指摘しておきたい。

経験諸科学や認知科学に隣接する経験科学諸領域と比較すると、認知科学は単に経験科学的・記述的だけではなく既に規範的特徴を内に含んでいる場合がある。そのような観点は、認知科学の方法論自体や認知科学と規範的学との関係をさらに検討するうえで重要であろう。

内的表象が認知科学的知識と呼び得る十分条件として活用可能性 (utilizability)、頑健性 (robustness)、適応性(adaptability)、許可性(admittability)が議論されているが、それらの間の関係について、例えば、頑健性と適応性はどのような意味で両立するのか、あるいは、許可性は独立の条件とし

て加える必要があるのかといった理論的な疑問が提起された。これらの点についてはさらなる考察の余地が残されていると思われる。また、これらの条件の一部は進化論的適応の過程で満たされた条件と考えることが可能かもしれない。このような進化論的観点を取り入れた知識の成立条件という課題も検討に値すると思われる。さらに、ここで提起された学習や相互作用の理論が、創発や創造性などのより複雑な認知プロセスの理解にどのように寄与するのかも、興味深い課題である。

本論文では、科学理論の知識の実在性の議論とは別に、特定領域に熟達した行為主体が構成する表象の対象の実在性が、収束的実在論を援用して主張されている。これに対しては、科学哲学において多くの批判にさらされてきた、奇跡論法を援用した収束的実在論に依拠するのではなく、「共有ビュー」等のプロセス指向的構成主義特有の概念装置を用いた、これまでにないタイプの実在論擁護の可能性があるのではないか、との指摘があった。収束的実在論よりもむしろ、Ronald Giere 的な構成的実在論に近い実在論擁護のためのアーギュメントを展開することが、今後の研究課題として期待される。

このような研究課題を残してはいるものの、この論文の成果は認知科学的観点を含めて哲学的認識論に新しい知見を与えるものであり、その学術的価値は国際的にみても第一線級であると言える。審査委員会は、本論文が慶應義塾大学大学院文学研究科博士学位授与に充分に値するものであると判定する。

平成３０年 10 月 4 日

審査委員会一同