

체화된 인지와 반표상주의

배 문 정[†]

인지과학을 학제적 연구 프로그램으로 존립하게 하는 근간에는 표상과 계산의 개념이 있다. 하지만 최근 체화된 인지논의가 광범위하게 확산되면서 인지과학에서 표상과 계산의 지위에 대한 전면적인 재검토가 요구되고 있다. 이 논문은 1) 체화된 인지의 여러 접근이 표상에 반대하는 이유를 역사적 맥락에 초점을 맞추어 살펴보고, 2) 지난 20여년에 걸쳐 진행된 표상 논쟁이 인지과학의 철학적, 경험적 토대에 가져온 변화를 개괄한다. 그리고 3) 이 변화가 인지과학의 지평을 확장하고 공고히 하는 방향으로 나아가기 위해 필요한 이론적, 실천적 조건들에 대해 논의한다.

【주요어】 표상, 반표상주의, 체화된 인지, 동역학계 이론, 인지과학

[†] 우석대학교 상담심리학과, mjbae64@gmail.com.

지난 반세기 동안 인지과학은 학제적 융합 연구의 가장 훌륭한 모범으로 자리 잡았으며, 앞으로 반세기 동안 인류 문명에 가장 괄목할만한 변화를 가져올 학문 분야로 주목받고 있다. 이러한 놀라운 성과가 가능했던 까닭은 인지과학이 ‘인지’ 또는 ‘마음’이라는 각별히 중요하고 심층적인 실체를 탐구 대상으로 하기 때문이 아니라 그 탐구의 작업이 경험적 검증의 과학적 활동이며, 다양한 학문적 전통과 성과들이 상호 침투하고 수렴하는 과정을 통해 이루어지는 학제적 활동이라는 데 있다. 이 ‘경험 연구에 기반한 학제적 소통’이 가능했던 바탕에는 철학, 언어학, 심리학, 신경과학, 인공지능의 제반 학문 영역들이 기꺼이 함께 사용하는 표상(representation)과 계산(computation)이라는 공통의 언어가 있었으며, 인지과학은 마음의 표상이론과 계산이론이라는 두 반석 위에 성공적으로 세워진 바벨탑이라고 해도 과언이 아니다. 하지만 최근 인지과학 내에서 이 표상과 계산 개념의 지위에 대한 전면적인 재검토를 요구하는 흐름들이 나타났다. 흔히 체화된 인지(embodied cognition)로 통칭되는 이 흐름들은 인지가 개인(또는 개체)의 머릿속에서 일어나는 표상적이고 계산적인 과정이 아니라, 유기체가 환경과 직접적이고 역동적으로 상호작용하는 과정에서 일어나는 적응적이고 창조적인 의미생성의 행위라고 주장한다(Brooks 1991a, Beer 1990, Varela, Thompson and Rosch 1991).

본 논문은 체화된 인지 논의가 인지과학에 던지는 도전이 무엇이며, 이 도전을 넘어 인지과학이 새롭게 열어야 하는 지평이 어디인지를 탐색하는 작업의 일환으로 체화된 인지의 반표상주의를 살펴보고자 한다. 체화된 인지의 여러 새롭고 긍정적인 주장을 뒤로 하고 굳이 반표상주의에 초점을 맞추는 이유는 첫째, 체화된 인지 논의가 하나의 동질적 주장이 아니라 경험 연구 영역에서 부딪힌 다양한 문제들을 해결하는 과정에서 나타난 이질적이고 분산된 흐름이라는 점에서 최소한의 공통적 기반을 확보하는 것이 필요하고, 그 기반이 전통적 표상 개념에 대한 반대라고 생각하기 때문이다. 두 번째는 전통적인 인지과학자들이 반표상주의를 체화된 인지의 가장 위험하고 전복적인 주장으로 받아들이고 있기 때문이다. 많은 이들이 인지가 신체적 활동과 밀접한

연관을 가지며, 개체의 머리 속에 감금된 유아론적(solipsistic) 과정이 아니라는 주장에는 충분히 동의할 수 있지만, 심적 실체로서의 표상의 존재와 설명적 유용성을 부정하는 것은 인지과학의 토대를 흔드는 급진적이고 과격한 주장이라는 비판을 제기했다. 나아가 체화된 인지의 주장에서 반표상주의를 제거하고 온건한 체화된 인지주의를 정립하려는 시도들이 잇따랐다.¹⁾ 따라서 체화된 인지의 반표상주의는 체화된 인지 논의를 이해하는 출발점이며 동시에 체화된 인지가 인지과학의 전반에 가져올 변화를 가능할 눈금자이다. 본 논문은 체화된 인지의 반표상 논쟁을 가능한 철학적 논변이 아니라 경험연구의 맥락에서 살펴보고, 이 논쟁이 촉발한 인지과학의 연구 주제와 방법론의 변화에 초점을 맞추어 논의를 진행하고자 한다.

1. 인지과학의 공통 화폐, 표상

행동주의의 전횡에 맞서 인지과학의 출범을 알린 인지혁명의 성공은 ‘표상’과 ‘계산’의 깃발 없이는 불가능한 일이었다. “인지는 계산”이고 “표상 없는 계산은 없다”(Fodor 1981)라는 슬로건이 인지과학을 정의한다고 해도 좋을 만큼 표상과 계산은 인지과학의 핵심 개념이다.²⁾ 포도어(Fodor 1981)는 적합한 인지과학 연구는 반드시 마음의 표상적 속성을 이해해야 하며, 인지과학적 설명은 생물학이나 물리학, 화학에서의 설명과 달리 반드시 의미론적으로 검증가능하고 행동의 원인이 되는 실체가 개입(etiologically involved)해야 한다고 말한다. 표상

1) 이러한 입장은 클락(Clark 1997)이 반표상주의적 체화된 인지를 급진적 체화된 인지(radical embodied cognition)로 지칭하고, 표상주의적 체화된 인지를 수식어 없는 체화된 인지로 분류한 데서 잘 드러난다.

2) “No representations, no computations. No computations, no model.”(Fodor 1975, p. 31), “The manipulation and use of representations is the primary job of the mind”(Dretske 1995, p. xiv), “Representation seems to me the very essence of mind.”(Rey 2002, p. 403).

의 이 두 가지 조건은 표상에 대한 가장 일반적인 정의에 해당한다. 이런 표상 개념에 근거하여, 쌔가드(Thagard 1996)는 인지과학에서 설명의 기본적인 도식을 다음과 같이 특징짓는다.

1. 설명 대상: 왜 사람들은 특별한 종류의 지능적인 행동을 하는가?
2. 설명 패턴: 사람들은 심적 표상을 가진다. 사람들은 그 표상들에 작동하는 알고리듬의 과정을 가진다. 이 표상들에 적용된 과정들은 행동을 산출한다.

여기서 반드시 짚고 넘어가야 할 점은 심적 표상은 인지를 설명하기 위해 가정된 구성개념(construct) 또는 이론적 실체(theoretical entity)이며, 지능적 행동을 설명하기 위해 인과적, 논리적으로 필요하다면 얼마든지 고안해서 사용할 수 있다는 것이다.³⁾ 이 심적 표상들의 심리적 실재는 그 표상 개념을 요구한 경험 데이터에 의해, 또 때로는 순전히 직관에 의해 논리적으로 보장된다. 이렇게 심리학적으로 가정된 심적 표상과 계산 모형이 물리적 실체로 구현가능한지의 여부를 결정하는 것은 신경과학과 인공지능, 계산론적 시뮬레이션의 끝으로 넘겨진다. 따라서 인지과학에서 심적 표상의 실재는 심리학, 신경과학, 인공지능 등의 수렴적인 협력 연구를 통해 검증되고, 이렇게 검증된 심적 표상은 정합적인 인지모형을 개발하는 데 핵심적인 역할을 수행한다. 이처럼 ‘보이지 않는 실체인 마음’에 대한 과학을 표방하는 인지과학에서 표상과 계산, 그리고 표상과 계산에 의해 조작적으로 정의되는 인지개념은 심리주의의 덫에 걸리지 않고 마음을 과학적으로 탐구하게 하는 이론적·방법론적 기초이다.

³⁾ 인지심리학에서 기억 연구의 데이터를 설명하기 위해 블록(chunk), 도식(schema), 스크립트 등의 심적 표상이 고안되었으며, 삼단추론의 행동적 자료를 설명하기 위해 ‘심성모형(mental model)’이라는 심적 표상이 고안되었다. 이 외에도 마(Marr 1982)의 망막배열, 일차 스케치, 2½차원 스케치, 민스키(Minsky 1974)의 ‘프레임’, 코슬린(Kosslyn 1980)의 ‘심상’ 등은 이론적 실체로서의 심적 표상이다.

하지만 인지과학에서 표상 개념은 그 근본적인 중요성만큼이나 꾸준히 논란의 대상이 되어왔다.⁴⁾ 인지과학 내부에서 진행된 대표적인 표상 논쟁은 표상주의 대 제거론(eliminativism)의 논쟁과 기호주의 대 연결주의의 논쟁이었다. 제거론은 상식적으로 가정되는 심적 상태의 존재를 부정하는 입장이지만, 인지과학적 제거론은 주로 명제태도와 관련되었으며 인지과학의 설명 개념으로서의 표상을 부정하지는 않았다. 오히려 대부분의 제거론적 주장들은 심적 표상에서 통속 심리학적 요소를 제거하고 표상을 자연화하여, 계산론적 인지과학에 정합성과 정당성을 부여하려는 노력에 집중되었다(Dennett 1969, Stich 1983, Churchland 1989). 또 인지과학의 역사에서 가장 급격한 논쟁이었던 기호주의와 연결주의 간의 논쟁도 표상의 실재나 설명적 지위가 아니라 표상의 형식과 계산의 구현방식에 관한 것이었던 만큼, 인지과학의 역사에서 표상과 계산은 보완되거나 확장될 수는 있지만 제거될 수는 없는 핵심 요소로 간주되었다.

반면 1990년대부터 본격화된 체화된 인지의 반표상주의적 주장은 철학자들이 아니라 경험과학자들에 의해 제기되었으며 심적 표상의 존재와 설명적 유용성 모두를 문제 삼았다. 그리고 이들의 반표상주의는 표상에 근거한 기존의 연구 패러다임을 폐기하고 새로운 연구 패러다임을 개척하려는 과정에서 제기되었다는 점에서 이전의 제거론이나 연결주의 표상 논쟁과는 큰 차이를 보였다. 다음 절에서는 체화된 인지 논의를 촉발한 상황지워진 AI와 신경동역학, 진화 로보틱스, 실행적

4) 표상주의에 대한 가장 극렬한 반대는 주로 행동주의 전통으로부터 왔지만, 이외에도 Gibson 1979)의 생태주의 심리학, Rorty(1979)의 비실재론, 드레이퓌스(Dreyfus 1972)의 현상학적 접근, 비고츠키(Vygotsky 1935) 전통의 사회적 구성주의, 후기 비트겐슈타인(Witgenstein 1953)의 일상언어 철학의 전통이 있다. 이들의 주된 문제제기는 표상주의가 심신이 원론과 주객이원론의 철학적 입장에 개입하고 있으며, 이러한 관점이 인지와 세계에 대한 이해를 지나치게 단순화시키고 협약화한다는 것이다. 표상주의 또는 인지주의 패러다임에 반대하는 철학적 입장들과 그에 대한 주류 인지과학자들의 대응에 대한 개관은 Still and Costall (1991), Haugeland (1978) 참조.

인지 프로그램을 중심으로 각각의 연구 프로그램이 표상에 반대한 경험적, 이론적 근거가 무엇인지를 살펴본다.

2. 체화된 인지와 반표상주의

1) 로드니 브룩스와 「표상 없는 지능」(1991): 상황지워진 AI

체화된 인지의 여러 주장을 중 세간의 가장 큰 관심과 주목을 받았던 인물은 「표상 없는 지능」("Intelligence without Representation")이라는 도발적인 제목의 논문을 들고 나타난 MIT의 로보틱스 연구자 브룩스(Rodney Brooks)였다. 그는 논문의 첫머리에서 인간 수준의 완벽한 지능을 구현할 수 있다고 믿었던 GOFAI의 기대는 실질적으로 좌초했으며 AI의 새로운 접근으로 단순하지만 자율적인 시스템을 구축하려는 노력이 시작되었다고 선언한다. 이 새로운 AI는 1) 유기체의 지능이 진화의 과정을 통해 점차적으로 발전한 것과 마찬가지로 단순하고 낮은 단계의 지능적 체계에서 점차적으로 높은 단계로 나아가며, 2) 각각의 단계에서 개발된 지능적 체계는 실세계에서 감각과 운동능력을 가지고 완전하게 작동하는 체계가 되도록 해야 한다는 목표를 가진다. 그리고 이런 자율적인 운동 로봇을 개발하는 과정에서 예상치 못한 두 가지의 결론에 도달했는데, 그것은 첫째, 매우 단순한 수준의 지능을 조사할 때 명시적인 표상과 세계에 대한 모형은 오히려 방해가 될 뿐이며, 지능적 체계는 ‘세계를 그것 자신의 모형으로’(the world as it's own model) 사용하는 것이 낫다는 것이다. 두 번째 결론은 표상은 대부분의 인지적 체계를 개발하는 데 잘못된 추상화의 단위라는 것이다. Brooks의 첫 번째 결론은 그의 논문 중 가장 많이 인용되는 문구로, 인공적 체계 뿐 아니라 자연적 체계(동물이나 사람) 또한 머릿속에 세계에 대한 상세한 정보를 저장하고 인출하는 방식이 아니라 세계 안에 있는 풍부하고 역동적인 정보를 세계에 대한 지식으로 직접 사용할 것이라는 함축을 제공한다. 이런 관점에서 보면, 내적 표상은 불필요하

고 거추장스러운 짐일 뿐이다.

브룩스의 로봇은 이전의 로봇시스템과는 달리, 탐색할 공간에 대한 정보를 미리 프로그램하지 않으며, 내적 모듈과 중앙 통제부를 가지지 않도록 설계되었다. 로봇의 행동은 각각 완전한 감각운동 회로—행동 산출의 층들—의 병렬적이고 중첩된 구조들에 의해 조율되는데, 각각의 층은 다른 층이 시키는 대로 일하는 하위 루틴이 아니라 그 자신들이 행동해야 될 때가 언제인지를 스스로 결정하는 자율적인 회로들이다. 브룩스는 이 층들은 **지각을 행위에 직접 연결**⁵⁾하기 때문에 시스템의 어디에도 표상은 없으며 층들 사이의 국지적인 혼돈(local chaos)에서 창발하는 정합적인 패턴이 관찰자의 눈에 표상으로 해석되는 것일 뿐이라고 강조한다. 즉 체계의 행동에 표상과 통제부를 귀속시키는 것은 체계 자체도, 설계자도 아니며, 바로 관찰자들이다. 이어서 그는 “표상은 그 용어로 말하지 않으면 고립될 수밖에 없었던 논문들과 모듈들 사이의 연결을 원활하게 한다는 단 한 가지 이유로 지난 15년간 AI 분야의 중심적인 논제가 되어왔다.”라고 덧붙인다.

표상에 대한 브룩스의 이런 단호한 반대는 그만의 돌출적인 주장은 아니었다. 이는 당시 AI 영역에서 태동하고 있던 행동기반 로보틱스 (behaviour-based robotics)와 상호작용적 AI(Agre and Chapman 1987), 상황지워진 자동자 이론(situated automata theory, Rosenschein and Kaelbling 1995) 등 이후 상황지워진 AI로 통칭되는 새로운 접근의 관점을 대변하고 있다. 브룩스(Brooks 1991b)는 상황지워진 AI의 주된 아이디어는 상황성과 신체성이라고 명시하고, 1) 상황성 (situatedness)은 “로봇은 세계 안에 상황지워져 있고, 그들은 추상적인 기술(description)을 다루지 않는 대신 체계의 행동에 직접적으로 영향을 주는 환경의 ‘여기(here)’와 ‘지금(now)’을 다룬다.”라는 의미로, 2) 신체성(embodiment)은 “로봇은 신체를 가지고 세계를 직접적으로 경험하며, 그들의 행동은 세계와 연동된 동역학(coupled dynamics)의 부

⁵⁾ 자율적 체계는 ‘지각과 행위가 직접 연결’되고, 환경에는 ‘행위자의 행동을 안내하는 충분한 정보가 있다.’ 따라서 ‘표상은 불필요하다’는 Brooks의 이 주장은 Gibson의 생태주의적 지각이론의 핵심 아이디어다.

분이며, 행동은 로봇 자신의 감각에 대한 즉각적인 피드백”이라는 뜻이라고 설명한다.

2) 프리먼과 스카르다의 자기조직하는 창발적 두뇌: 신경동역학

버클리의 신경과학자인 프리먼(Walt Freeman)과 그의 동료 스카르다(C. A. Skarda)는 1990년 「표상: 누가 그것을 필요로 하는가?」라는 제목의 논문에서 신경과학에서 표상 개념이 사용되어서는 안 되는 두 가지 이유를 제시하는데, 첫째는 표상 개념이 신경계의 활동을 이해하고 설명하는 데 어떤 기여도 하지 않는다는 것이고, 둘째는 표상 개념은 뇌의 실제 활동을 오해하게 만들고 신경과학의 발전을 저해한다는 것이다. 그리고 뇌의 활동을 연구하는 새로운 패러다임으로 신경동역학(neural dynamics)을 제안한다.

1980년대의 신경과학 연구들은 뇌에서 심적 표상에 해당하는 기능적 단위를 찾는 데 집중했고, 프리먼 그룹도 EEG를 사용하여 후각피질에서 특정한 냄새에 반응하는 신경집단의 활성화 패턴을 발견했다. 이 결과는 후각피질에 특정한 냄새에 반응하는 신경적 표상으로 간주될 수 있는 것이었다. 하지만 10년의 연구 끝에 프리먼 그룹이 내린 결론은 1) 후각 망울(olfactor bulb)에서 발견되는 신경적 활동 패턴은 유기체의 동기(motivation)와 행동과 맥락에 의존적이고 2) 신경적 활동 패턴은 동물의 역사와 환경의 상호작용에 의해 정의되며, 신경활동의 이 역동적인 과정에서 고유하게 표상이라고 부를 만한 것은 없다는 것이었다.

이후 프리먼은 신경적 표상의 개념을 완전히 폐기하고, 새로운 연구 질문을 제기했다. 그 질문들은 신경활동의 시간적 동역학(temporal dynamics)과 뇌의 자기조직화 원리와 관련된다. 그들은 이 새로운 질문을 탐색할 방법론으로 수학적 비선형 동역학 이론을 사용하기 시작했고, 이를 통해 신경과학의 전통적인 두 가정에 중대한 전환을 가져왔다. 그것은 첫째, 뇌는 기능적으로 국제화된 모듈들로 이루어진 계산론적 장치가 아니라 비선형적인 혼돈 체계(nonlinear chaotic system)라는 것이다. 여기에는 표상도 통사적 절차도 없다. 둘째, 계산론적 신경

과학에서는 신경활동의 노이즈를 제거 혹은 조정해야 하는 것으로 다루지만, 신경활동의 패턴—질서—은 뇌에 있는 모든 신경세포들의 활동—노이즈—을 배경으로 해서 일어난다. 프리먼 등의 이 결론은 두뇌의 기능을 공간적 구조화와 연결시키는 관점에서 시간적 구조화의 관점에서 바라보게 만들었으며, 뇌를 신체, 환경과 지속적으로 상호작용하며 역동적으로 변하는, 살아있는 구조로 바라볼 것을 요구한다.

프리먼과 스카르다, 그리고 앞서 본 브룩스의 논지는 체화된 인지주의자들이 표상에 대해 가지는 가장 전형적인 태도를 보여준다. 그것은 첫째, 표상에 기초한 연구 패러다임은 경험적으로 실패했으며 새로운 이론과 방법론이 필요하다는 자각과 선언이다. 둘째, 표상 개념이 어려운 문제를 지나치게 단순화시켜 결국 (인공지능이든 신경과학이든) 연구자들이 진정으로 추구해야 하는 질문에 집중하지 못하도록 한다는 평가다. 브룩스는 표상은 지능적 체계의 비용을 증가시키고 비효율적인 체계로 만든다는 점을 강조한 반면, 프리먼은 뇌의 역동성을 포착하기에 표상이 적절한 개념이 아니라는 점을 지적한다. 이들의 강조점은 서로 다르지만 둘 다 표상과 계산은 인지과학의 설명 개념으로 유용하지 않다는 주장이다. 세 번째는 표상은 실험자나 관찰자에 의해 해석되고 삽입되는 것이지 두뇌나 지능적 체계 안에는 표상이 존재하지 않는다는 주장이다. 이는 지능적 체계의 복잡함 때문에 체계를 단순화시켜 해석하고 설명하는 도구로 표상 개념을 사용할 수는 있지만, 이런 설명적 유용성과 무관하게 표상은 그 추상성과 정태성 때문에 물리적인 지능 체계 안에 존재할 수 없다는 주장이다.

이제 우리는 표상이 인지과학의 경험 연구를 안내하고 인지체계를 설명하는 데서 유용성을 가지지 못한다면 무엇으로 인공지능연구, 신경과학, 심리학을 가로지르는 공통의 설명 기반을 삼을 것인지를 묻게 된다. 이 질문은 인지과학 안에서 표상에 반대하는 주장이 그처럼 거센 저항에 직면한 이유이며, 이를 통해 단지 표상을 반대하는 것을 넘어 표상을 대체할 공통의 설명개념을 찾아야 하는 인지과학의 새로운 과제가 제시되었다. 이 질문은 반겔더(van Gelder 1995)의 논문 제목 「만약 계산이 아니라면, 인지는 무엇인가?」(“What Might Cognition

Be, if Not Computation?”)에 상응한다. 만약 표상이 아니라면 우리는 무엇으로 인지과학을 지속할 수 있는가? 우리는 이 질문에 대한 대답을 인공생명학자 비어(Randall Beer)에게서 발견할 수 있다.

3) 랜달 비어와 동역학계 이론: 진화 로보틱스

비어(Beer)는 1990년에 출간한 박사학위 논문 「적응적 행동으로서의 지능」(“Intelligence as Adaptive Behavior”)에서 전통적인 AI가 지능을 ‘신중한 추론’(deliberative reasoning)으로 정의하고 표상을 필수적인 요소로 받아들이지만 표상은 단지 가설일 뿐 참이거나 거짓으로 밝혀진 적이 없다는 점을 강조한다. 그는 표상에 기반한 어떤 시도도 지능을 탈맥락화시키고 근본적으로 유연하지 못한 체계를 만들어 낼 뿐이라는 점을 지적하고, 지능을 탈신체화된 추론이 아니라 행위의 관점에서 정의한다. 그리고 AI의 새로운 과제는 실제 환경에서 자율적으로 행동하는 체계를 개발하는 것이며, 이런 자율적 행위자(autonomous agent)를 개발하는 하나의 방법론으로 곤충의 신경계를 컴퓨터로 시뮬레이션하는 계산론적 신경행동학(computational neuroethology; CNE)을 제안했다. 1992년에는 갤러거(J. Gallagher)와 함께 쓴 논문 「적응적 행동을 위한 역동적 신경망의 진화」(“Evolving Dynamical Neural Networks for Adaptive Behavior”)에서 CNE에 유전자 알고리듬을 접목해 무작위 네트워크를 걸어 다니는 인공 바퀴벌레로 진화시키는 데 성공하고 진화 로보틱스라는 새로운 분야를 개척했다.

나아가 비어(Beer 1995)는 체화된 인지의 관점에서 표상적 계산이론을 대체할 설명 도구로 수학적 동역학계 이론(dynamical system theory; DST)을 제시한다. DST는 여러 개의 자율적인 체계들이 서로 밀접하게 상호작용하면서 만들어내는 역동적 변화를 기술하고 설명하기에 적합한 설명 이론이다. 동역학적 체계는 시간에 걸쳐 변화하는 모든 실체를 포괄하는데, 그것은 움직이는 당구공일 수도, 상태 변화를 겪는 튜링머신일 수도 있으며, 자기조직하는 두뇌일 수도 있다. 행위환경 상호작용은 두 개의 연동된 동역학계(coupled dynamical systems)로 기술된다. DST는 체계들의 변화를 추상적인 상태 공간 안

에서 나타내고 여러 국면을 통해 움직이는 궤적을 추적한다. 따라서 상태 공간(또는 국면 공간)과 위치, 궤적은 동역학계의 기본적이 개념이다. 동역학계가 어떤 국면 공간에 반복적으로 끌려들어간다면 그 지점은 ‘끌개(attractor)’라고 불린다. 반대의 개념은 ‘밀개(repeller)’다, 또 ‘평형상태(equilibrium)’, ‘분할선(separatrices)’, ‘국면 초상(phase portrait)’ 같은 개념들이 사용된다. DST는 기본적으로 미분 방정식의 집합이지만, 그 변화의 궤적을 정성적으로 기술하고 예측할 수 있다. 비어는 이 수학적 도구야 말로 환경과 신체, 신체와 두뇌의 경계를 가지르는 상호작용으로서 인지의 본성을 가장 잘 드러낼 수 있으며, 특히 복잡하고 변화하는 환경 안에 있는 인지 체계의 행동을 예측하는데 가장 적합한 설명 도구라고 제안한다.

4) 바렐라, 톰슨, 로쉬의 실행적(enactive) 인지: 신경현상학과 체화된 실재론

1990년대 초는 반표상적 반계산론적 흐름이 인지언어학(Lakoff and Johnson 1980), 발달심리학(Thelen and Smith 1994), 운동협응 연구 (Kelso 1995) 등의 다양한 모습으로 드러나고 있었다. 바렐라와 톰슨과 로쉬는 『체화된 마음』에서 인지과학의 경험 연구들은 전통적인 인지주의의 기본 가정을 재검토하고, 실행적(enactive)⁶⁾ 인지의 새 패러다임으로의 전환을 요구하고 있다고 역설하고, 자신들의 접근을 다음과 같이 정의한다.

우리는 인지는 ‘이미 주어진 마음에 의한 이미 주어진 세계의 표상’이 아니라, 세계 안에 있는 존재가 수행하는 다양한 행위의 역사에 기초해서 ‘세계와 마음을 실행’하는 것이라는 점증하는 확신을 강조하기 위해 ‘실행적enactive’이라는 용어를 이름으로 제안한다. 그리고 실행적 접근은 마음이 ‘자연의 거울’이라는 생각에 대한 철학적 비판을 심각하게 받아들인다. 하지만 과학의 심장 한

6) ‘enactive’는 ‘발제(發製)’, ‘행화(行化)’, ‘행위화’ 등으로 번역되었으나, 본 논문에서는 ‘실행(實行)’으로 번역한다. ‘enactive’의 번역 문제에 관해서는 배문정(2014) 참조.

가운데서 이 주제를 제기하는 것에 의해 한 발 더 나아간다.⁷⁾

바렐라 등은 자신들의 반표상주의가 경험 연구의 축적된 결과에 기초해 있고 자신들의 접근이 과학적 연구 프로그램으로 지속될 것이라고 밝히고 있지만, 이러한 작업의 근본적인 동기가 인지과학에 깊이 침습해있는 심신 이원론과 주객 이원론의 철학적 난관을 해결하는 데 있다는 점을 분명히 한다. 이는 앞서 살펴본 브룩스, 프리먼, 비어 등의 태도와는 사뭇 다르다. 그들의 관심이 각자의 연구 영역에서 만난 경험적 과제를 해결하기 위해 표상이라는 개념적 결림돌을 치우는 데 있었고 주로 인지적 설명 개념으로서의 표상의 유용성 문제를 지적했다면, 바렐라 등의 관심은 표상에 의해 가로막힌 마음과 세계 사이의 본유적인 관계를 회복하고, 인지의 개념을 확장하는 데 있었다. 여기서 우리는 체화된 인지의 반표상주의가 경험적 동기와 함께 철학적 동기를 포함하고 있다는 것을 발견하게 된다.⁸⁾ 바렐라 등의 이 관심은 철학 일각에서 제기된 반실재론을 넘어 체화된 실재론의 가능성을 탐색하고, 오랫동안 분리되어 있었던 현상학적 전통과 과학의 전통을 발전적으로 접목시키는 작업으로 이어진다. 이 작업의 일환으로 그들은 신경동역학에 일인칭 방법을 접목시킨 신경현상학(neurophenomenology) 연구 프로그램을 개척한다.

인지과학 전반에 대한 깊은 이해와 통찰, 그리고 방대한 경험적 예시에도 불구하고, 그들의 이 제안은 인지과학 주류에서 의도적으로 축소되거나 폄하되었는데,⁹⁾ 그 이유는 무엇보다 그들의 실행적 인지가

7) Varela, Thompson and Rosch (1991), p. 9.

8) Chemero(2009)는 체화된 인지의 반표상주의를 경험적 반표상주의와 형이상학적 반표상주의로 구분하는데, 이 두 관점은 서로 밀접히 연결되어 있지만, 서로가 서로를 반드시 전제해야 하는 것은 아니라고 주장한다. 즉 심적 표상의 존재론적 인식론적 함의에 깊이 개입하지 않는다 하더라도, 인지의 설명 개념으로서 표상의 필요성을 받아들일 수 있으며, 동시에 설명적 유용성과 무관하게 철학적 관점에 개입할 수 있다는 것이다.

9) 바렐라 등의 실행적 인지에 대한 주류 인지과학자들의 평가는 Clark(1997), Boden(2006)의 평가에서 잘 드러난다. “바렐라 등은 그들의

인지과학의 본질적인 질문인 인식론의 영역을 넘어서 있는 것처럼 보였기 때문이다. 인식론의 전통적인 질문은 인지자가 어떻게 객관적인 세계를 정확하게 지각하고, 이해하고 추론할 수 있는가에 놓여 있다. 그리고 이 질문의 틀 안에서는 어떤 존재론적 입장을 취하든 표상은 인식론의 필수불가결한 요소다. 만약 인지가 세계를 정확하게 아는 것 또는 세계에 대한 정합적인 지식을 갖는 것이 아니라면, 인지는 무엇인가? 이 질문에 그들은 ‘인지는 실행(enactment)’이라고 대답한다. ‘실행’은 때로 (세계의) 창출, 생성이라는 의미로 사용되며, 인지의 창조적, 실천적 속성을 강조한다. 그들의 새로운 패러다임은 그 용어에서부터 인지과학이 인식론의 영역을 넘어 실천론과 윤리학의 영역까지 지평을 넓혀야 한다는 생각을 함축한다.

바렐라 등은 반표상주의를 그들의 주장 한 가운데 배치한 것은 표상이 주객 이원론을 떠받치는 견고한 인식론적 장치일 뿐 아니라, 이 개념이 인지과학에서 다루어야 할 본질적인 문제들에 대한 과학적 설명 또한 제거하고 있기 때문이라고 말한다. 예를 들어, 심리학에서 인지주의가 등장하면서 심리학 교과서에서 의식과 창조적 사고와 같은 주제가 사라졌고(Still and Costall 1991 참조), 신체 협응과 운동, 행위 같은 연구 주제는 인지과학의 변방에도 포함되지 않았으며, 인간 경험의 자연적 표현으로서의 통속 심리학은 설명 대상에서 배제되었다. 프리

성찰을 실재론과 세계에 대한 객관주의적 관점에 반대하는 증거로 사용한다. 나는 의도적으로 이 확장을 피하려고 한다. 이러한 확장은 사물은 마음에 독립적이지 않다는 말썽 많은(problematic) 생각을 체화된 인지의 접근에 연결해서, 이 접근의 과학적 가치를 가려버리는 위험을 초래하기 때문이다. 반대로 나의 주장은 단순히 생물학적 뇌가 표상하는 실제 세계의 측면은 종종 구체적인 필요와 감각운동적 능력에 밀접하게 맞물려 있을 것이라고 말하는 정도다(Clark 1997, p. 173).” 또 Boden(2006)은 인지과학의 방대한 역사서 *Mind as Machine*의 곳곳에서 바렐라 등의 작업을 뉴에이지의 반문화적 시도로 폄하한다. 하지만 Clark은 최근 논문(2015)에서 자신의 새로운 이론(Predictive Processing)은 바렐라 등의 (반객관주의) 철학과 완전히 양립 가능하며 반표상적(급진적) 체화된 인지 진영과 표상주적 체화된 인지 진영 사이의 표상 전쟁은 끝났다고 선포한다.

먼 등(Freeman et al. 1990)이 정확히 지적한 것처럼, 전통적인 인지과학은 ‘마음은 무엇을 할 수 있는가?’나 ‘두뇌는 무엇을 할 수 있는가?’를 묻기보다는 ‘컴퓨터는 무엇을 할 수 있는가?’를 묻는다. 이러한 표상주의적 인지 관점에서 마음은 세계 속의 존재가 살아가면서 겪는다 채롭고 독특한 경험의 파노라마가 아니라 보편적이고 객관적이고 논리적인 마음으로 협약화된다.

3. 표상주의 대 반표상주의 논쟁: 길을 내며 가는 길

우리는 앞서 체화된 인지의 여러 접근들이 제기한 반표상주의의 경험적 동기와 철학적 동기들을 살펴보았다. 논문의 이 부분에서는 체화된 인지의 반표상주의가 촉발한 논쟁의 굽은 줄기들을 개괄하고, 이 논쟁들이 인지과학 전반의 철학적 관점과 경험 연구에 어떤 변화를 가져왔는지를 살펴본다.

1) 동역학계 이론과 조속기 논쟁

체화된 인지의 경험적 반표상주의는 표상이 인지과학의 경험연구에 이론적, 설명적 유용성을 가지지 않으며 오히려 연구의 발전을 저해한다고 주장하며, 표상의 대안으로 동역학계 이론을 제안한다. 표상주의 계산이론의 대안으로 동역학계 이론(DST)의 필요성을 역설한 것은 비어가 처음은 아니었다. 비선형 동역학을 인지과학의 실질적인 연구 방법론으로 도입한 최초의 연구자들은 신경동역학을 개척한 프리먼 그룹이었다. 하지만 기계와 동물을 동역학적 체계로 기술하는 기본 아이디어는 인지과학의 전신이라 할 수 있는 초기 사이버네틱스 시절부터 있었다. 애쉬비(Ashby 1953)는 두뇌를 포함하는 모든 적응적인 체계를 연구할 수 있는 방법론으로 변수들의 집합적 장(field)이 시간에 걸쳐 변하는 패턴을 연구하는 동역학계 이론을 제안했다. 또 1980년대 들어서면서 터비 등의 생태주의 심리학자들(Kugler, Kelso and Turvey

1980)도 계산주의의 대안적인 프레임워크로 수학적 동역학계 이론을 도입했다. 그들은 지각을 유기체와 환경 사이의 조율(tuning) 또는 공명(resonance)이라고 표현한 김슨의 개념을 수학적 설명 개념으로 형식화하기 위해, 연동된 진동자(coupled oscillators) 개념을 사용했고, 켈소(Kelso 1995)는 유기체의 동작 협응과 범주적 지각, 두뇌의 활동 등을 ‘자기조직하는 동역학적 패턴’(self-organizing dynamical pattern)으로 설명하는 통합 이론을 제안했다.

체화된 인지의 반표상주의에서 동역학계 이론의 중요성은 반겔더(van Gelder 1995)의 와트 조속기(governor) 비유를 통해 광범위하게 확산된다. 반겔더(van Gelder 1998)는 “디지털 컴퓨터와 달리 동역학계는 본래 표상적이지 않다. 소수의 영향력 있는 동역학계 이론가들은 그들의 목적을 위해 표상 개념이 불필요하거나 심지어는 방해가 된다는 것을 발견했다. 동역학은 표상을 완전히 밀어낸 인지 모형을 발전시키기 위한 강력한 프레임워크이다. 인지가 반드시 표상과 관련되어야 한다는 가정은 부분적으로 비표상적인 체계가 어떻게 인지적 과제를 수행할 수 있는지 상상할 수 없다는 데서 비롯된다. 동역학계적 접근 안에서는 그런 체계는 상상될 수 있을 뿐 아니라 모형화되고 구축될 수 있다.”고 말하고, 그 상상의 한 예시로 와트 조속기를 제시한다. 반겔더(van Gelder 1995)는 기계적 장치인 와트 조속기의 작동을 설명하는 계산론적 모형과 동역학계적 모형을 비교함으로써, 표상을 물리적 실체가 아니라 이론적 실체로 보는 관점을 확고히 한다.

와트 조속기에서 엔진의 회전 속도에 의존하며 벨브의 마개를 열고 닫는 기능을 하는 양 팔의 각도가 과연 표상으로 해석될 수 있는지에 대해 다양한 의견이 제시되었는데, 베텔(Bechtel 1998)은 양 팔의 각도는 엔진의 회전 속도에 인과적으로 연결되어 있다는 점에서 그리고 그 각도가 회전 속도와 동형적(isomorphic)이라는 점에서 표상으로 해석되어야 한다고 주장한다. 반면, 램지(Ramsey 2007)를 비롯한 다른 연구자들은 팔의 각도는 엔진의 회전 속도에 인과적으로 연결되어 있지 만 동형적이지 않으며 인과적 관련성은 표상의 조건이 아니라고 주장한다.

이 논쟁에서 어떤 입장을 취하든 핵심은 기능적이지만 명백히 인지적이라고는 할 수 없는 조속기의 작동을 모형화할 때도 이론가가 표상에 대해 어떤 정의를 택하는가에 따라 그것을 표상적 계산론적 체계로 설명할 수 있다는 점이다. 이로써 표상은 인지의 필요조건도 충분조건도 아니며 이론가의 해석의 둑이라는 관점이 광범위하게 확산되었다. 이 논쟁을 통해 얻은 또 다른 성과는 심적 표상이 불가피하게 개입할 수밖에 없었던 통속심리학적 요소를 배제하고 오직 표상의 설명적 지위에만 집중하는 논의를 확산시켰다는 점이다. 따라서 표상은 더 이상 인지과학이라는 영역의 유일한 게임이 아니라 설명적 개념의 여러 선택지 중 하나로 다뤄지게 되었다.

반면, 비어(Beer 1996)는 표상에 대한 논쟁이 철학적 논변이나 직관, 유비에 머무는 것은 인지과학의 발전을 위해 바람직하지 않으며 이 문제를 극복하기 위해서는 가능한 한 선형적 가정에 의존하지 않는 구체적인 모델에 근거해서 토론이 진행되어야 한다고 강조하면서 이런 토론의 가능한 한 가지 방식으로 진화 시뮬레이션의 사용을 제안했다. 그의 아이디어는 ‘최소로 인지적인 행동’(minimally cognitive behavior)의 이상적인 모형을 사고실험의 형태로 구성하는 것이다. 이러한 시도로 그는 원과 다이아몬드를 범주적으로 지각하는 인공 행위자를 동역학계 이론의 개념과 방법을 사용하여 합성하고 분석하는 실험을 진행했다. 비어의 연구 외에도 엘먼(Elman 1995)의 어휘적, 통사적 구조 연구, 틸렌과 스미스(Thelen and Smith 1994)의 영아들의 ‘A not B’ 오류 연구, 스티븐과 딕슨(Stephen and Dixon 2009)의 기어 문제해결 연구 등은 일반적으로 표상과 계산에 의존해 설명해 왔던 고차적인 인지행동들이 동역학계 이론의 개념과 도구로 훌륭하게 설명되고 예측될 수 있다는 것을 보여주었다.

2) 행위정향적 표상(action oriented representation)

체화된 인지의 주장들이 제기된 후, 인지를 두뇌와 신체, 환경의 광범위한 상호작용으로 이해해야 한다는 전반적인 취지에는 동의하지만, 표상에 대한 명시적인 반대만큼은 수용하기 힘들다는 반론들이 제기되

었다. 그 대표적인 예가 커시(Kirsh 1991)의 주장이다. 그는 “오늘은 집게벌레, 내일은 인간?”(Today the earwig, tomorrow man?)이라는 제목의 논문에서 브룩스(Brooks 1991)의 반표상주의를 반박한다. 그의 주장의 핵심을 요약하면, 1) 하위수준의 인지와 인간의 고차적 인지는 근본적으로 다르며, 2) 언어의 사용과 문화의 전승은 계산적 비용을 효율적으로 줄여 주며, 3) 이런 언어의 사용과 문화적 전승의 모든 이점은 사실을 내적으로 표상하는 능력에서 흘러나온다는 것이다. 그는 언어 능력이 인간의 진화 경로를 얼마나 바꾸었는지를 고려할 때 표상의 필요는 자명한 것이라고 강조한다. 이 반론은 이후 클락에 의해 점진적이고 체계적인 형태로 제시되었고(Clark 1997, 1999, Clark and Toribio 1994, Clark and Grush 1999), 이 주장에 휠러(Wheeler 2005), 롤랜즈(Rowlands 2006), 윌슨(Wilson 2004) 등의 철학자들이 합류하면서 표상주의적 체화된 인지의 흐름을 형성한다. 이 주장의 흐름을 추적하기 위해 클락과 토리비오의 주장부터 살펴보자.

클락과 토리비오(Clark and Toribio 1994)는 브룩스(Brooks 1991a)과 비어(Beer 1992)의 연구를 개괄하면서, 전통적 표상주의와 연결주의 논쟁은 표상의 형태와 속성에 대한 관점의 차이에도 불구하고 마음의 과학은 그 중심에 내적 표상을 전제해야 한다는 공통적인 합의가 있었던 반면 체화된 인지는 표상 그 자체를 문제 삼고 있으며 이런 주장이 철학적 사변이 아니라 경험 연구의 결과로 제기되었다는 점에서 큰 영향력을 가진다고 말한다. 하지만 이 반표상적 주장은 사실 지나치게 과장된 것이고 자신들이 제시하는 경험 자료의 범위를 벗어났다고 지적한다. 그들의 비판은 크게 두 가지로 집약되는데, 1) 체화된 인지주의자들이 표상을 지나치게 좁게 해석한다는 것이다. 즉 그들은 오직 전통적인 기호 표상만을 염두에 두고 반표상주의를 내세우지만, 연결주의의 표상은 암묵적이고, 신경계의 구조에 분산적으로 체화되어 있다. 2) 체화된 인지주의자들이 표상에 반대하는 증거로 제시하는 인지현상들은 모두 하위 수준의 인지이며, 본래부터 표상에 크게 의존하지 않는 문제들이라는 것이다. 반면 고차적 수준의 인지과정—생각, 언어, 기억 등—은 반드시 표상을 필요로 한다. 이런 문제들을 ‘표상에

굶주린 문제’(representation hungry problems)¹⁰⁾라고 부른다. 클락 등은 이런 논지는 체화된 인지의 반표상주의에 대해 이중적인 의미를 지는데, 하나는 전면적인 반표상주의 선언은 지나치게 성급하고 오류일 것이라는 주장, 다른 하나는 최소한 어떤 인지는 표상적 설명이 필요 없음을 인정하는 것이다. 그들은 표상에 대한 이런 관점을 최소표상주의(minimal representationalism)이라고 지칭한다.

최소표상주의자들(Wheeler 2005, Rowlands 2006, Wilson 2004, Clark and Grush 1999)은 표상의 필요성을 역설하면서도 그 표상의 속성이 체화되어 있다고 본다는 점에서, 그리고 인지를 두뇌, 신체, 환경을 포함하는 확장된 과정으로 이해한다는 점에서 스스로를 체화된 인지주의자로 명명한다. 그들은 전통적인 기호 표상과는 구분되는 행위정향적 표상(action-oriented representation)¹¹⁾을 가정하며, 이는 온라인 인지뿐만 아니라 오프라인 인지에서 사용되는 표상의 진화적, 발달적 기초가 될 것이라고 주장한다. 이들에서 중요한 변화는 준언어적이고 객관적이며 행위 중립적인 전통적인 표상 개념의 폐기다. 휠러(Wheeler 2005)는 이 새로운 표상들은 객관적이기 보다는 자아중심적(egocentric)이며, 저장되고 인출될 수 있는 것이 아니라 일시적(transient)이고, 행위의 맥락에 민감하고, 그리고 무엇보다 행위의 통제에 직접적으로 개입한다고 설명한다.

이 관점은 확실히 김순(Gibson 1979)의 ‘어포던스’(affordance)를 표상으로 수용한 것이며, 실제 대부분의 최소표상주의자들은 하이데거의 존재론을 수용한다. 따라서 이들의 표상에 대한 옹호는 경험적이기 보

¹⁰⁾ Clark 등의 표상에 굶주린 문제 영역은 다음 중 하나 또는 둘 다를 만족하는 문제영역들을 의미한다. 1) 눈앞에 없거나(absent), 존재하지 않는 것 (non-existent), 또는 가공적인(count-factual) 상태에 대한 추론과 관련된 문제들. 2) 행위자가 복잡하고 비규칙적인 자극에서 특정 패러미터에 선택적으로 민감하기를 요구하는 문제, 즉 추상적인 처리와 관련된 문제들이다.

¹¹⁾ ‘행위정향적 표상’은 Clark(1997)이 자신의 최소표상을 지칭하기 위해 고안한 개념이다. Wheeler(2005)는 인지를 적응적인 행위를 위한 것으로 이해하고, 전통적인 표상 개념이 가져오는 인지적 부담을 최소화시키려는 관점에서 제안된 모든 종류의 표상을 지칭하는 개념으로 사용한다.

다는 형이상학적이다. 이들은 인지자를 행위자로 이해하면서, 행위자가 객관적 세계의 관찰자이기 보다는 세계의 구성적(constitutional) 요소임을 인정한다. 하지만 하이데거의 존재자(Dasein)가 세계와 맺고 있는 관계가 직접적이고 도구적인 관계(ready to hand)뿐 아니라 대면적 관계(present at hand)를 포함하며 데카르트의 탈신체화된 인지의 문제점은 대면적 관계를 이상화시킨 데 있을 뿐 데카르트적 탈신체화된 인지의 존재 자체를 부정할 수는 없다고 본다. 데카르트의 관념적, 반성적 인지는 세계에 대한 객관적 인식의 가능성을 열어 놓는다는 점에서 인지과학이 반드시 설명하고 이해해야 할 인간 인지의 중요한 속성이며 따라서 그러한 인지를 설명할 최소한의 기반으로 표상 개념은 유지되어야 한다는 것이 그들의 행위정향적 표상주의의 형이상학적 동기다.

행위정향적 표상은 전통적인 표상이 개입했던 철학적 이원론으로부터 벗어나 있지만, 이 표상 개념이 인지의 경험적 설명개념으로 가치를 지닐 수 있는지에 대해서는 논란의 여지가 있다. 갤러거(Gallagher 2008)는 설명개념으로서의 행위정향적 표상이 지니는 가장 심각한 문제로 인과적 확산(causal spread)과 결별가능성(decouplability)을 든다. 롤랜즈(Rowlands 2006)는 “표상의 운반체는 피부에서 멈추지 않는다. 그것들은 세계를 향해 모든 방향으로 확장한다.”고 말한다. 즉 행위정향적 표상은 신체와 환경적 대상을 포함하는 상호작용에 관여하며, 두뇌, 신체, 환경은 모두 행동을 이끄는 동등한 인과적 구성요소이다. 이 때 환경적 상태와 신체의 상태는 미리 예측되거나 고정된 것이 아니기 때문에 두뇌의 신경적 패턴 또한 저장된 고정 패턴일 수 없다. 이처럼 역동적이고 광범위한 상호작용을 설명하기 위해서는 표상은 적절한 설명 개념이 아니라는 것이 초기의 (근본적인) 체화된 인지과학자들이 주장이었으며, 그 인과적 확산과 맞물림을 설명할 수 있는 대안으로 동역학계 이론을 제안한 것이다. 따라서 최소표상주의자들이 이처럼 광범위한 인과적 확산을 전제하면서 어떻게 표상 개념을 유지할 수 있는지는 경험적 검증 이전에 논리적 딜레마이다.

두 번째 결별 가능성 또한 행위 정향적 표상의 큰 문제점이다. 롤랜즈(Rowlands 2006)는 전통적인 표상은 내적이고, 지속하고, 자기 외부

의 의미를 전달하고, 해석을 필요로 하며, 수동적이고, 현재의 맥락으로부터 결별가능(decouplable)하다는 속성들을 가진다고 하며, 이 행위정향적 표상에 포함되는 유일한 속성으로 결별 가능성을 든다. 이는 행위정향적인 표상이 오프라인 인지에서도 사용될 수 있다는 것을 의미한다. 하지만 인과적으로 광범위하게 확산되어 있는 행위정향적 표상이 환경과 신체로부터 결별할 때, 그 표상이 어떻게 기능적으로 작동할 수 있는지, 그 표상이 과연 행위가 진행 중일 때 사용되는 표상과 근본적으로 동일한 속성을 가지는지는 해명되지 않는 채 남겨진다.

3) 설명항으로서의 표상

체화된 인지의 반표상주의는 표상의 전통적인 관점에 큰 변화를 가져왔고 새로운 표상 개념들이 제안되는 계기가 되었다. 하지만 이처럼 다양한 표상 개념이 제안되는 데에 문제를 제기하는 연구자들도 있다.(Haselager, de Groot and van Prppard 2003, Ramsey 2007). 예를 들어, 조속기 논쟁에서 베텔(Bechtel 1998)이 조속기 팔의 각도가 표상으로서의 역할을 있다고 주장했을 때, 이것은 단지 억지나 돌출적인 주장이 아니라 실제 신경과학에서 광범위하게 통용되는 표상의 관점을 대변한 것이다. 신경과학에서는 표상 개념이 충분한 검토 없이 광범위하게 편재해 있으며 더구나 이론적 실체(구성물)인 표상이 문자 그대로 두뇌에 실재하는 것으로 오용된다. 나아가 표상은 이론적 숙고의 결과이기 이전에 이론을 위한 전제로 사용된다(Tonneau 2011). 이러한 난맥상과 관련해 램지(Ramsey 2007)는 인지과학에서 표상은 가장 단순한 무조건 반사를 제외하고 거의 모든 것을 설명할 수 있는 보편적 개념으로 사용되고 있으며, 이 표상적 귀인 앞에서 나타나는 인지과학 자들의 근거 없는 확신과 자신감은 좋은 과학의 징조가 아니라고 지적한다. 차가운 맥주가 음식을 지시하거나, 바위가 어떻게 언덕을 굴러가는지 안다고 말하거나, 등산객의 혈압이 고도를 표상하고, 위가 소화과정을 계산한다고 말하는 것은 표상으로 아무 것도 설명할 수 없는 상황을 초래하며, 인지과학은 표상이 결코 사소하지 않은(non-trivial) 설

명을 제공하는 조건의 집합을 명시해야 한다.

램지는 「재고찰된 표상」(“Representation Reconsidered”)에서 인지과학에서 통용되는 표상의 개념을 입출력 표상, 시뮬레이션 표상, 수용기 표상, 절차적 표상으로 정리하고, 이 중 의미론적 관련성을 지니고 사소하지 않은 설명을 제공하는 적법한 표상은 시뮬레이션(내적 모형)뿐이라고 주장한다. 그리고 단지 인과적 연결을 가지는 것으로서의 수용기나 절차적 표상은 표상으로서의 적법성(legitimacy)을 가지지 않는다고 설명한다. 이러한 인과적 연결은 인지적 체계 이외에 자연적이거나 기계적인 체계에서도 관찰되고, 이러한 연결은 표상이론 보다는 동역학계 이론에 의해서 더 잘 설명될 수 있다. 램지는 표상의 개념을 이처럼 엄격하게 규정하고 사용한다면 인지과학에서 통용되고 있는 대부분의 표상적 설명이 적법하지 않은 것으로 드러날 것이라고 말한다.

4) 피설명항으로서의 표상

인지과학에서 설명 개념으로서 표상의 경험적 가치를 재고하려는 노력과는 달리, 표상은 설명항이기 이전에 피설명항으로 다루어져야 한다는 경험적 반표상주의 주장을 살펴보자. 이 주장은 주로 진화로 보티스 연구자들 사이에서 제기되는데, 하비(Harvey 2008)는 인지과학에서 표상 개념이 가지는 문제는 그것이 경험적으로 검증되고 설명되어야 하는 대상이 아니라 설명 개념으로 전제되는 데 있다고 지적하고, 설명개념으로서 표상이 지니는 완고한 힘은 그것이 자연 언어를 인지의 메타포로 사용하기 때문이라고 말한다. 실제 인간이 사물을 표상하는 능력은 놀라울 정도로 중요하고 흥미로우며, 이 표상 능력의 출현은 인지의 진화에서 극적인 전환을 가져온 하나의 사건이다. 그리고 이 능력은 인지과학의 설명항이 아니라 피설명항이다. 즉 물리학에서 설명의 메타포로 사용되는 당구공이 그 자체로 원자와 분자의 관점에서 설명되어야 하는 대상인 것과 마찬가지로 표상 능력 또한 심적 요소와 과정의 관점에서 설명되어야 하는 대상이다. 이러한 발상의 전환은 명제태도나 통속심리학에 대해서도 동일하게 적용될 수 있다. 인간은 사회적 장면에서 상대방의 믿음, 욕구, 희망을 지각하고 이를 통

해 행동의 이유를 추론한다. 후토(Hutto 2007)는 심리학의 설명에서 암묵적으로 전제되거나 가정되고 있는 통속 심리학적 태도는 단지 제거되어야 하는 것이 아니라 인간의 사회인지 능력의 일부로 설명되어야 하는 것임을 강조하고, 인지발달의 과정에서 일어나는 서사 훈련(narrative practice)에 주목한다.

표상 능력의 출현을 경험적으로 연구할 수 있는 가장 유망한 방법론은 진화 로보틱스의 최소인지 실험이다. 최근 로봇들의 사회적 협응과 집단행동을 조작한 실험 연구들은 로봇들의 행동에서 기호적인 제스처로 이해될 수 있는 행동들을 진화시켰다.(Di Paolo 2000, Quinn et al. 2003). 이 진화적 사회인지 연구들은 인지주의에서 개체 안에 고립되어 있는 것으로 가정되었던 심적 능력들이 계통발생학적으로 그리고 개체발생학적으로 사회적 상호작용을 통해 출현하고 발전한다는 것을 보이는 중요한 증거들을 제공한다. 하비(Harvey 2008)는 이 연구들이 풀어야 할 과제는 인공적 행위자의 체계 안에 내적 표상으로 해석될 수 있는 기제가 존재하는지를 확인하는 것이 아니라 인공적 행위자가 명실상부한 표상의 사용자로 진화하기 위해 요구되는 최소한 환경적, 상호작용적 요건들이 무엇인지를 밝히는 일임을 강조한다. 표상이 더 이상 직관적으로 전제되는 설명 개념이 아니라 그 자체로 엄격하게 탐구되어야 하는 설명 대상으로 다루어지기 시작했다는 것은 체화된 인지의 반표상주의가 가져온 중요한 성과의 하나이다.

5. 성숙한 인지과학을 향하여

우리는 앞선 논의에서 체화된 인지의 반표상주의가 인지과학이 발전하고 성숙하는 과정에서 얻은 경험적 성과로, 다양한 관점과 시도를 양산하면서 인지과학의 새로운 패러다임으로 성공적으로 정착해가고 있음을 보았다. 하지만 체화된 인지가 계산론적 인지과학을 완전히 대체할 수 있을지, 과연 동역학계 이론이 표상과 계산 개념을 통해 형성된 학제적 소통의 대안적 기반이 될 수 있을지는 여전히 의문으로 남

는다. 그 의문의 첫 번째 사소한 이유는 대부분의 인지과학자들에게 동역학계 이론은 지나치게 낯설고 어려운 수학적 도구라는 것이며, 그 보다 더 중요한 두 번째 이유는 표상과 계산 개념이 우리의 상식과 직관에 대해 가지는 강력한 호소력이다. 와트 조속기에서 표상과 계산의 이미지를 제거하는 것은 비교적 쉬운 일이지만 살아있는 유기체의 인지를 설명할 때 통속심리학적 요소를 제거하는 것은 결코 쉬운 일이 아니다. 세 번째로 보다 근본적인 이유는 체화된 인지 패러다임이 아직은 초기 형성기를 벗어나지 못한 발전 중의 패러다임이며, 동역학계 이론 또한 복잡한 인지과정에 적용하기에는 아직 많이 보완되어야 할 수학적 도구라는 점이다. 따라서 이 논문의 결론은, 체화된 인지의 반표상주의의 성공 여부는 인지과학자들이 실천적으로 검증하고 개척해야 하는 경험연구의 뜻이며 이 실천이 인지과학의 지평에 가져올 변화의 정도도 아직 확실하게 가늠할 수 없다는 것이다. 하지만 한 가지 확실한 것은 그 변화가 인지과학을 한층 성숙하고 풍부한 학문으로 만들 것이라는 점이다.

이제 남은 일은 이 자기실현적 예언이 실제로 완수되기 위해서 인지과학자들에게 필요한 실천적 태도가 무엇인지를 물어야 한다. 그 대답은 아마도 본래의 질문을 잊지 않는 것, 그 질문의 답을 향해 가는 과정에서 마주쳐야 하는 감당할 수 없는 모호함을 견뎌내는 것, 그리고 내가 걸어온 길과는 다른 길을 걸어온 사람들의 경험과 지혜에 기꺼이 귀를 기울이고 소통과 협력을 통해 창조적인 발견의 미덕을 터득하는 것일 테다. 하지만 지나온 인지과학의 역사가 그만큼 충분히 진지하고 개방적이고 창조적이었다고 확인할 수는 없다. 인지주의는 다른 극단에서 행동주의만큼이나 강한 실재론적 확신에 개입했으며, 그 확신에서 벗어나는 다양한 의견들을 인지과학의 영역에서 원천적으로 추방해 왔다. 데일 등(Dale, Dietrich and Chemero 2009)은 인지혁명 이후의 심리학과 인지과학의 역사서들은 새로운 이론들이나 비주류의 이론들을 인지주의에 종속적이거나 보완적인 것으로 또 때로는 절대적으로 그려왔다는 점을 지적하고, 지금 인지과학에서 진행되고 있는 기호주의, 연결주의, 동역학계 이론의 ‘프레임 논쟁’은 설명적 다원주의를 받

아들이는 기초 위에서만 인지과학의 발전에 기여할 수 있을 것이라고 강조한다. 설명적 다원주의의 필요성을 역설한 베델과 에이브럼슨의 말을 인용하며 반표상주의 논쟁이 견지해야 할 실천적 태도에 대한 결론을 대신하고자 한다.

우리는 인지과학이 그 많은 논쟁에도 불구하고, 다양한 영향들을 조합하고 재조합하는 지속적인 과정을 통해 발전했다는 것을 강조해왔다. 현상을 특징짓고 설명하는 데서 방정식의 사용 또한 마찬가지다. 다른 영향과 개입들이 조합될 때, 여기서 논의된 결과들은 양적인 조작을 가지는 정보처리모형에서부터 연결주의 네트워크 전역적이고 기계적인 동역학적 설명까지 영역을 넓힌다. 이 접근들은 각각 마음이 언어와 같은 표상을 형식적인 규칙에 따라 처리하는지 아닌지에 대한 상이한 대답을 제공한다. 그리고 우리는 전체적인 대답이 이를 중 하나에 제한될 필요가 없다고 주장한다. 인지과학은 주어진 시대에 다중적인 형태를 취한다. 그리고 시간에 걸쳐 변한다. 나는 거대하다, 나는 다중성을 가진다(*I am large, I contain multitudes*).¹²⁾

¹²⁾ Bechtel and Abrahamsen (2006), p. 187.

참고문헌

- 배문정 (2014), 「Enactivism을 Enact 하기】, 『인지과학』 25권 4호, pp. 303-341.
- Agre, P. E. and Chapman, D. (1987), “Pengi: An Implementation of a Theory of Activity”, *AAAI 87(4)*: pp. 268-272.
- Ashby, W. R. (1953), “Design for a Brain”, *Journal of Consulting Psychology* 17(1): p. 76.
- Bechtel, W. (1998), “Representations and Cognitive Explanations: Assessing the Dynamicist's Challenge in Cognitive Science”, *Cognitive Science* 22(3): pp. 295-318.
- Bechtel, W. and Abrahamsen, A. (2006), “Phenomena and Mechanisms: Putting the Symbolic, Connectionist, and Dynamical Systems Debate in Broader Perspective”, in Stainton, R. J. (ed.), *Contemporary Debates in Cognitive Science*, Oxford: Blackwell, pp. 159-187.
- Beer, R. (1990), *Intelligence as Adaptive Behavior*, New York: Academic press.
- _____ (1995), “A Dynamical Systems Perspective on Agent-Environment Interaction”, *Artificial Intelligence* 72(1): pp. 173-215.
- _____ (1996), “Toward the Evolution of Dynamical Neural Networks for Minimally Cognitive Behavior”, *From Animals to Animats* 4: pp. 421-429.
- _____ (2008), “The Dynamics of Brain-Body-Environment Systems: A Status Report”, *Handbook of Cognitive Science: An Embodied Approach*, pp. 99-120.
- Beer, R. and Gallagher, J. C. (1992), “Evolving Dynamical Neural Networks for Adaptive Behavior”, *Adaptive Behavior* 1(1): pp. 91-122.

- Boden, M. A. (2006), *Mind as Machine: A History of Cognitive Science*, Oxford: Oxford University Press.
- Brooks, R. (1991a), “Intelligence without Representation”, *Artificial Intelligence*, 47(1): pp. 139-159.
- _____, (1991b), “New Approaches to Robotics”, *Science* 253(5025): pp. 1227-1232.
- Chemero, A. (2009), *Radical Embodied Cognition*, Cambridge, MA: MIT press.
- Churchland, P. M. (1989), *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, Cambridge, MA: MIT press.
- Clark, A. (1997), *Being There: Putting Brain, Body, and World Together Again*, Cambridge, MA: MIT press.
- _____, (1999), “An Embodied Cognitive Science?”, *Trends in Cognitive Sciences* 3(9): pp. 345-351.
- _____, (2015), “Predicting Peace: The End of the Representation Wars—A Reply to Michael Madary”, in Metzinger, T. and Windt, J. M. (eds.), *Open MIND*: 7(R), Frankfurt am Main: MIND Group.
- Clark, A. and Grush, R. (1999), “Towards a Cognitive Robotics”, *Adaptive Behavior* 7(1): pp. 5-16.
- Clark, A. and Toribio, J. (1994), “Doing without Representing?”, *Synthese* 101(3): pp. 401-431.
- Dale, R., Dietrich, E. and Chemero, A. (2009), “Explanatory Pluralism in Cognitive Science”, *Cognitive Science* 33(5): pp. 739-742.
- Dennett, D. (1969), “Personal and Sub-Personal Levels of Explanation: Pain”, *Content and Consciousness*, London: Routledge.
- Di Paolo, E. A. (2000), “Behavioral Coordination, Structural Congruence and Entrainment in a Simulation of Acoustically

- Coupled Agents”, *Adaptive Behavior* 8(1): pp. 27-48.
- Dretske, F. I. (1995), “Meaningful Perception”, in Osheron, D. N. (ed.), *An Invitation to Cognitive Science: Visual Cognition* (Vol. 2), Cambridge, MA: MIT press, pp. 331-352.
- Dreyfus, H. L. (1972), *What Computers Can't Do: A Critique of Artificial Reason*, New York: Harper and Row.
- Elman, J. L. (1995), “Language as a Dynamical System”, in Port, R. F. and van Gelder, T. (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge, MA: MIT press, pp. 195-223.
- Fodor, J. A. (1975), *The Language of Thought* (Vol. 5), Cambridge, MA: Harvard University Press.
- _____. (1981), *Representations: Philosophical Essays on the Foundations of Cognitive Science*, Cambridge, MA: MIT press.
- Freeman, W. J. and Skarda, C. A. (1990), “Representations: Who needs them”, in McGaugh, J. L., Weinberger, N. M. and Lynch, G. (eds.), *Brain Organization and Memory: Cells, Systems, and Circuits* (Vol. 3), Oxford: Oxford University Press, pp. 375-380.
- Gallagher, S. (2008), “Are Minimal Representations Still Representations?”, *International Journal of Philosophical Studies* 16(3): pp. 351-369.
- Gibson, J. J. (1979), *The Ecological Approach to Visual Perception*, Boston, MA: Houghton Mifflin.
- Grush, R. (2004), “The Emulation Theory of Representation: Motor Control, Imagery, and Perception”, *Behavioral and Brain Sciences* 27(03): pp. 377-396.
- Harvey, I. (2008), “Misrepresentations”, *Artificial Life* 11: pp. 227-233.
- Haselager, P., de Groot, A. and Van Rappard, H. (2003),

- “Representationalism vs. Anti-Representationalism: a Debate for the Sake of Appearance”, *Philosophical Psychology* 16(1): pp. 5-24.
- Haugeland, J. (1978), “The Nature and Plausibility of Cognitivism”, *Behavioral and Brain Sciences* 1(02): pp. 215-226.
- Hutto, D. D. (2007), “The Narrative Practice Hypothesis: Origins and Applications of Folk Psychology”, *Royal Institute of Philosophy Supplement* 60: pp. 43-68.
- Kelso, J. S. A. (1995), *Dynamic Patterns: The Self-organization of Brain and Behavior*, Cambridge, MA: MIT press.
- Kirsh, D. (1991), “Today the Earwig, Tomorrow Man?”, *Artificial Intelligence* 47(1): pp. 161-184.
- Kosslyn, S. M. (1980), *Image and mind*, Cambridge, MA: Harvard University Press.
- Kugler, P. N. and Turvey, M. T. (1987), *Information, Natural Law, and the Self-assembly of Rhythmic Movement*, Mahwah: Lawrence Erlbaum Associates.
- Kugler, P. N., Kelso, J. S. and Turvey, M. T. (1980), “On the Concept of Coordinative Structures as Dissipative Structures: I. Theoretical Lines of Convergence”, *Tutorials in Motor Behavior* 3: pp. 3-47.
- Lakoff, G. and Johnson, M. (1980), *Metaphors We Live By*, Chicago/London.
- Marr, D. (1982), *Vision: A Computational Approach*, San Francisco: Freeman and Co.
- Minsky, M. (1974), “A Framework for Representing Knowledge”, *Artificial Intelligence Memo* No. 306, M.I.T., Artificial Intelligence Laboratory.
- Quinn, M., Smith, L., Mayley, G. and Husbands, P. (2002), “Evolving Team Behaviour for Real Robots”, *EPSRC/BBSRC International Workshop on Biologically-Inspired Robotics: The*

- Legacy of W. Grey Walter*, *WGW* 2: pp. 14-16.
- Ramsey, W. M. (2007), *Representation Reconsidered*, Cambridge: Cambridge University Press.
- Rorty, R. (1979), *Philosophy and the Mirror of Nature*, Princeton: Princeton University Press.
- Rosenschein, S. J. and Kaelbling, L. P. (1995), “A Situated View of Representation and Control”, *Artificial Intelligence* 73(1): pp. 149-173.
- Rowlands, M. (2006), *Body Language: Representing in Action*, Cambridge, MA: MIT press.
- Stephen, D. G. and Dixon, J. A. (2009), “The Self-organization of Insight: Entropy and Power Laws in Problem Solving”, *The Journal of Problem Solving* 2(1): p. 6.
- Stich, S. (1983), *From Folk Psychology to Cognitive Science: The Case against Belief*, Cambridge, MA: MIT press.
- Still, A. and Costall, A. (eds.) (1991), *Against Cognitivism: Alternative Foundations for Cognitive Psychology*, Harvester Wheatsheaf.
- Thagard, P. (1996), *Mind: Introduction to Cognitive Science* (Vol. 4), Cambridge, MA: MIT press.
- Thelen, E., and Smith, L. B. (1994), *A Dynamic Systems Approach to the Development of Perception and Action*, Cambridge, MA: MIT press.
- Tonneau, F. (2011), “Metaphor and Truth: A Review of Representation Reconsidered by WM Ramsey”, *Behavior and Philosophy* 39: pp. 331-343.
- Van Gelder, T. (1995), “What Might Cognition Be, if Not Computation?”, *The Journal of Philosophy* 92(7): pp. 345-381.
- _____, (1998), “The Dynamical Hypothesis in Cognitive Science”, *Behavioral and Brain Sciences* 21(05): pp. 615-628.

- Varela, F. J. (1996), "Neurophenomenology: A Methodological Remedy for the Hard Problem", *Journal of Consciousness Studies* 3: pp. 330-349.
- Varela, F. J., Thompson, E. and Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: MIT press.
- Vygotsky, L. S. (1934/1986), *Thought and Language* (Kozulin, A., trans.), Cambridge, MA: MIT press.
- Wheeler, M. (2005), *Reconstructing the Cognitive World: The Next Step*, Cambridge, MA: MIT press.
- Wilson, R. (2004), *Boundaries of the Mind*, New York: Cambridge University Press.
- Wittgenstein, L. (1953/2010), *Philosophical Investigations*, New York: John Wiley and Sons.

논문 투고일	2015. 09. 25
심사 완료일	2015. 10. 17
게재 확정일	2015. 11. 25

Embodied Cognition and Anti-Representationalism

Moon-Jung Bae

Cognitive science as an interdisciplinary research program has been based on concepts of representation and computation. However, recent broad support on embodied cognition has challenged these concepts. This paper 1) reviews the debates on anti-representationalism invoked by embodied cognition, 2) discusses the theoretical and empirical change for last twenty years and 3) further discusses theoretical and practical task to utilize these debates in expanding and substantiating philosophical and empirical basis of cognitive science.

[Key Words] representation, anti-representationalism, embodied cognition, dynamical system theory, cognitive science