

현상학의 관점에서 본 심리철학의 타당성

사범대학 물리교육과
2002-13620 이정규

I. 문제의 지평

헤겔철학의 붕괴와 함께 20세기에 접어들면서 현대인은 정신적으로 병들게 되었고 이처럼 병든 20세기를 후설은 위기의 시대라고 진단한다. 그리고 이러한 현대인의 위기의 일차적인 책임은 현대과학이 병들어서 그 본래적인 기능을 상실하고 위기상황에 빠졌다는 데 있다. 시대가 지남에 따라 비약적으로 발전했던 과학이 그 본래적인 의미를 상실하며 그 기능을 제대로 수행할 수 없게 된 것이다.

후설에 따르면 이런 위기의 근본적인 원인은 바로 철학의 위기에 있다. 철학은 20세기에 접어들면서 자신의 고유한 기능을 상실하고 병들게 됨으로 일대 위기에 처하게 되었고, 그 필연적인 결과로 과학의 위기를 초래하게 되었다. 이러한 철학의 정체성의 위기가 계속되면서 19세기 중후반에서 20세기 초에 이르는 사이에 실증적인 개별과학의 비약적 발전에 ‘실증주의’라는 사조가 등장하게 되는데 이들은 모두 실증적인 개별과학을 그 철학적 토대로 삼는다는 점에서 공통점을 지닌다.¹⁾

그리고 실증주의 중에서도 논리 실증주의와 자연주의의 영향을 받아 탄생한 철학의 한 분야가 바로 ‘심리철학’이다. 심리철학은 ‘마음을 가진 생물’이라는 개념을 명확하게 하는 문제, 특정한 심적 속성이나 심적 상태와 사건의 종류 및 상호 관계에 관한 문제들, 그리고 심적 속성과 물리적 속성간의 관계에 관한 문제 등을 다루는 철학의 분야를 말한다.

이와 같은 심리철학의 주제에 대해서 물론 다양한 입장들이 있다. 그렇지만 실제로 많은 경우 심리철학자들은 그것이 강하든지 약하든지 간에 어느 정도의 물리주의적인 입장을 취하고 있으며 주로 그를 토대로 논의가 진행 되고 있는 것이 사실이다.

이 글에서는 물리주의에 토대를 두고 있는 심리철학의 세 가지 주된 사조인 행동주의, 심신 동일론, 기계 기능주의에 대하여 후설의 현상학적 태도에서 그것들을 비판적으로 살펴 보겠다. 사실 후설은 『엄밀학으로서의 철학』에서 이미 자연주의 철학과 역사주의 철학을 비판한 바 있으며 심리철학 역시 많은 입장들이 자연주의 철학의 입장과 상통한다. 그리고 이 글은 그런 비판의 연장선에 서있다고 볼 수 있다.

물론 이 글의 목적이 단순히 후설의 입장에서 심리철학이 어떻게 보일지에 대해서 서술하고 끝내는 것은 아니다. 그것은 타당성에 대해서 주장하는 글이 아니라 관점 차이가 무엇 인지를 나타내는 글에 불과할 것이기 때문이다. 때문에 명확히 해야 할 것은 필자가 이런 글을 쓰는 이유는 그런 단순한 관점 차이를 보이기 위해서가 아니라 타당한 논변을 펼치기 위해서이며 그 목적을 위해 후설의 주장들을 잠시 빌려 쓰는 것이다. 물론 그와 같은 주장은 후설이 했던 다른 사람이 했던 전혀 상관없으며 그것이 타당하기만 하면 되는 것이다.

17세기의 누군가가 그랬던 것처럼²⁾ 필자 역시 심리철학의 모든 문제들을 면밀히 살펴

1) 이남인, 『현상학과 해석학』, 서울대학교출판부, 2004, pp. 22-24.

2) “그러나 이를 위해 모든 의견이 거짓임을 증명해 보일 필요는 없다. 이것은 내가 도저히 해 낼 수 없기 때문이

하나하나 옳고 그름을 판단하기는 힘든 일이며 또한 지금의 필자에겐 능력을 벗어나는 일이기 때문에 이들이 토대로 삼고 있는 건물의 가장 아랫부분을 빼내기만 하면 그와 같은 건물들은 자연스럽게 무너지게 될 것이다. 그렇지만 아랫부분을 빼내기에 앞서 과연 그 건물이 무엇인지부터 알아보도록 하자. 그리고 경우에 따라서는 그 건물에 적합하지 않은 부분들을 제거해 버리기로 하자. 그리고 마지막으로 그것들이 토대로 삼고 있는 전체를 반박하여 건물 전체를 무너뜨리기로 하자.

비판에 앞서 일단 명시하고 가야할 것이 있다. 심적인 것은 과연 어떤 특징을 갖는가 하는 점이다. 여기서 주의해야 할 점은 선입견을 가지고 심적인 것의 특징을 규정하려하는 것이 아니라 가능한 모든 선입견을 제거하고 그 사태자체로서 우리에게 나타나는 심적인 것의 특징을 살펴보아야 한다는 점이다. 그렇다면 이와 같은 태도에서 본 심적인 것의 특징에는 어떤 것들이 있을까?

II. 심적인 것의 특징에는 무엇이 있는가?

심적인 것의 특징을 밝히기 위해 우리가 할 일은 바로 심리현상의 본질구조를 해명하는 일이다. 왜냐하면 우리가 여기에서 문제로 삼고 있는 것은 바로 심리학이며 이와 같은 경험과학은 그것이 과학인 한에서 나름대로 어떤 본질적인 가정들에 입각해서 전개되는데, 경험과학이 일종의 학으로서 정립될 수 있기 위해서는 꼭 필요한 본질적인 가정들이 무엇인지를 알아야 하기 때문이다. 후설은 이런 본질적인 가정들이 무엇인지를 다루는 학문을 “내용적 존재론(material Ontologie)”이라 부른다. 따라서 우리가 여기서 수행하는 것은 바로 심리학의 내용적 존재론이며 이를 다른 말로 현상학적 심리학이라고 부를 수 있다.

심리현상의 본질구조 해명을 목표로 하는 현상학적 심리학의 과제는 심리현상 일반의 본질구조를 해명하는 일과 다양한 개별적인 유형의 심리현상의 본질구조를 해명하는 일로 나눌 수 있다. 여기에서는 논의를 진행하기 위해서 꼭 필요한 본질구조만을 중점적으로 살펴 보도록 하겠다.

후설은 심리현상 일반의 본질구조를 지향성, 즉 “대상을 향한 자아의 의식적 관계”에서 찾고 있다. 심리현상 일반의 본질구조가 지향성이라는 말은 모든 심리현상을 심리현상에게 끄 해 주는 핵심적인 요소가 바로 지향성이라는 것이다. 따라서 심리현상 탐구를 목표로 하는 경험적 심리학 역시 그것이 심리현상에 관한 참된 경험과학으로 정립되기 위해서는 “지향성이라는 관점에서” 심리현상의 다양한 성질을 해명해야 한다는 사실을 함축한다.³⁾

이제 개별적인 유형의 심리현상의 본질구조를 해명하는 작업을 하도록 하자. 여기에서는 논의를 진행시키는 데에 핵심이 되는 심리현상인 내적지각과 외부지각, 타인지각만을 살펴 보도록 하겠다. 우선 내적 지각은 나의 현재의 심리상태를 지각 하는 것으로써 지각 대상, 즉 체험 자체는 필증적 명증의 양상에서 의심할 수 없이 지각이 가능하다. 예를 들어 ‘책상이 갈색을 지닌 것으로서 나에게 경험되고 있다.’는 사실이나 강의 내용에 대해서 내가 이해하고 있는지 못하는지, 혹은 그것조차 아리송한지에 대한 사실은 나에게 필증적 명증의

다.[...] 이보다는 오히려 토대가 무너지면 그 위에 세워진 것도 저절로 무너질 것이기에, 기존의 의견이 의존하고 있는 원리들 자체(ipsa principia)를 바로 검토해 보자.”(René Descartes, 이현복 옮김, 『성찰』, 문예출판사, 1997, pp. 34-35.)

3) 이남인, 『현상학과 해석학』, 서울대학교출판부, 2004, p. 62.

양상에서 경험이 되는 것이다. 그러나 외부 지각은 이와는 사정이 조금 다르다. 물론, 지각 대상을 생생하게 떠올린다는 점에서는 내부 지각과 동일한 지향적 구조를 가지고 있으나 결코 필증적 명증의 양상에서 지각될 수 없다는 점에서 외부지각은 내부 지각과는 다른 본질 구조를 가지고 있는 것이다. 예를 들어 우리가 책상이 갈색이라는 사실을 지각할 때 그것이 실제로 갈색이 아닐 수 있는 가능성은 얼마든지 존재하며, 혹은 태양이 지구 주위를 돈다고 지각할 때, 우리는 우리의 외부지각이 종종 잘 못 될 수 있음을 쉽게 알 수 있다. 마지막으로 내가 타인의 심리 상태를 다양한 방식으로 파악하면서 수행하는 타인 지각-타인의 신체적 표현에 대한 외부지각이라는 지향체험과 그를 토대로 형성된, 타인의 심리상태에 대한 새로운 지향체험이 연결된 특수한 형태의 지향체험-의 경우는 외부지각 보다도 더 낮은 단계의 명증도를 띠게 될 수밖에 없다. 왜냐하면 타인 지각은 바로 외부 지각대상의 한 유형인 타인의 신체적 표현을 지각하고 그것을 나름대로 새롭게 해석하면서 수행되기 때문이다.⁴⁾ 사실 우리가 타인의 심리 상태에 대해서 수많은 오해를 한다는 점을 상기시켜보기만 하더라도 타인지각의 명증도가 외부지각보다 더 낮은 이유에 대해서 쉽게 파악할 수 있을 것이다.(사실 엄밀히 말해서 여기에서 ‘오해’라는 말은 의미가 없다. 왜냐하면 타인지각에서는 ‘진실’이 무엇인지 역시 원천적으로 찾을 수 없게 되어 있기 때문이다. 따라서 ‘타인의 심리 상태에 대해서 의심 가능하다는 점을 상기시켜보기만 하더라도’와 같이 의심 가능성으로 표현 하는 것이 더 적절하다.)

여기에서 우리는 내적지각의 경우, 그것이 나에게 직접적으로 알려지는 반면에 타인지각의 경우에는 그렇지 않다는 사실을 쉽게 알 수 있다. 우리는 타인의 심리상태를 해석하듯이 우리의 심리상태를 해석하고 판단하는 것이 아니라 그것을 직접적인 양상에서 알게 되는 것이다. 예를 들어 타인의 경우에는 그가 몸부림치는 것을 보며 고통이라는 심적 상태에 있는 것이라 판단하지만 우리 자신의 경우에는 내가 몸부림을 치고 있는 것을 보며 고통이라는 심적상태에 있다고 해석하는 것이 아니라 나의 몸부림에 대한 경험이 일어나기 이전에 이미 직접적으로 그런 상태를 알게 되는 것이다.

그런데 심리현상은 실제 주어진 것보다 “더 많이 사념함”이라는 본질적 속성을 가지고 있다. 예를 들어 우리가 책상을 볼 때 그 주어진 단면만을 보고 그 대상을 지각하는 것은 아니며, 우리는 그것을 3차원적인 대상으로서 우리에게 보이지 않는 부분까지 생각하며 직접적으로 주어진 의미보다 더 많은 의미를 지니고 있는 대상으로 지각하고 있음을 알 수 있다. 지향체험으로서의 외부지각은 이처럼 실제 주어진 것보다 “더 많이 사념함”이라는 특징을 가지고 있으며 이는 지각의 지향성이 과거에서 이미 주어진 의미와 현재 주어진 의미를 종합하면서 더 높은 단계의 새로운 의미를 지향하면서 파악하기 때문에 가능하다. 그리고 후설은 이런 새로운 의미 지향의 작용을 구성작용(Konstitution)이라고 부른다.⁵⁾

그런데 앞서 살펴본 현상학적 심리학이 이와 같은 구성작용에 초점을 맞추고 진행될 경우 현상학은 현상학적 심리학의 단계를 넘어 “구성적 현상학”의 단계로 넘어간다. 구성작용은 우리에게 주어진 의식이 앞서 주어진 낮은 단계의 대상적 의미로부터 더 높은 단계의 대상적 의미를 향해 초월해 가는 과정이라 할 수 있으며, 이러한 이유에서 구성적 현상학은 “초월론적 현상학”이라 불린다. 초월론적 현상학은 초월론적 주관의 구조를 탐구함을 목표로 하는데 여기서 초월론적 주관이란 구성작용의 주체로서 다양한 의미로서의 대상을 구성하는 주관이다. 그리고 이는 세계의 한 부분으로서 세계 속에 존재하는 심리학적 주관이 아

4) 이남인, 『후설, 엄밀학으로서의 철학』, 서울 관악사 p. 67.

5) 이남인, 『현상학과 해석학』, 서울대학교출판부, 2004, p. 75.

나라, 세계 및 세계내 대상을 구성하는 주관인 것이다.⁶⁾

그리고 이 같은 초월론적 주관은 초월론적 현상학적 판단중지를 통해서 우리에게 주어질 수 있는데, 여기에서 초월론적 현상학적 판단중지만 세계가 가능한 모든 존재자의 총체라고 믿고 있는 “자연적 태도의 일반정립”(die Generalthesis der natürlichen Einstellung)에 대해 그 믿음이 틀렸거나 의심스럽다고 주장하는 것이 아니라 다만 그에 대해 판단을 유보하는 것을 말한다. 세계에 대한 일반정립이 지속되는 한 우리에게 의식의 지향성은 바로 세계 속에 있는 대상 중의 하나로 파악되지만 초월론적 현상학적 판단중지가 수행되면, 의식의 지향성은 이제 더 이상 세계 속에 존재하는 수많은 대상 중의 하나로 파악되는 것이 아니라, 세계 및 세계내 대상을 구성하는 초월론적 기능을 지닌 것으로서 파악될 수 있다. 즉 초월론적 주관은 더 이상 세계의 한 부분으로서 세계 속에 존재하는 심리학적 주관이 아니라, 세계 및 세계내 대상을 구성하는 초월론적 주관으로 자신의 모습을 드러내는 것이다.⁷⁾

지금까지 심리현상의 본질 구조와 초월론적 주관에 대해 간략하게나마 살펴보았다. 압축적으로 말해서 심리철학의 문제들은 전부 지향성이라는 측면을 배제하고 논의를 진행하고 있으며, 경험과학으로서의 심리학에 대해 그 전제를 넘어서는 것까지 설명하려고 하는 수행적 모순을 저지르고 주체로서의 초월론적 주관을 말살시킨다. 그리고 내적지각과 타인지각의 구분이 전혀 되어 있지 않아 필요에 따라 그것을 혼용하여 쓰고 심리현상이 모두 밝혀진 것처럼 말한다. 예를 들어 어떤 입장의 경우, 타인지각이 무엇인지만을 분석해 놓고 그것이 마치 모든 심리현상의 본질인 것 마냥 주장하는 식이다. 좌우간 이제 심리철학을 비판하기 위해 필요한 기본적인 무기는 모두 갖추어졌다. 구체적인 목표물을 향해 조금씩 전진해 보자.

III. 행동주의 비판

20세기 초에 심리학의 성격과 방법론에 관한 이론으로서 등장한 사조가 바로 행동주의이다. 행동주의는 내성주의적인 심리학의 주관적이고 비과학적인 성격에 대해서 반기를 든 심리학자들이 주도하였다. 철학적 행동주의는 행동이 심성을 구성하는 것으로 본다. 이러한 입장에 따르면, 어떤 마음을 갖는다는 것은 적합한 패턴의 관찰 가능한 행동을 표출하거나 표출하는 성향을 갖는 것의 문제라고 본다.

행동주의자들은 우리가 TV에 나오는 사람들이나 만화를 볼 때도 종종 그것들에게 심적인 것을 귀속시킨다는 사실을 주의 깊게 관찰해보면 우리가 지금까지 생각해 오던 심적인 것이라는 것은 존재하지 않는다고 말한다. 길버트 라일(Gilbert Ryle)은 이에 대해 “범주의 오류”(Category mistake)라는 표현을 쓰는 데, 간단히 말해 이것은 우리가 인간에게 무언가 내적인 원리를 귀속시키는 것은 마치 서울대학교가 도서관, 본부, 강의실, 학생, 교수와 같은 구성요소 이외에 '+α'로 이루어져 있다는 착각을 하는 것과 같다는 뜻이다. 이에 따르면 팀워크라는 단어가 단순히 선수들의 행동 방식에서 찾아 볼 수 있는 것이듯이 마음이라는 것도 역시 다른 또 하나의 실체가 아니라 몸이 움직이는 어떤 경향성에 불과한 것이라 주장한다.

우선 논리적 행동주의에 대한 험펠의 논변을 살펴보기로 하자. 이러한 형태의 주장은 심

6) Ibid., pp. 77-79.

7) Ibid., pp. 85-87

리 문장의 번역 가능성에 관한 주장으로서 다음과 같이 요약될 수 있다.

[논리적 행동주의 I] 심적 현상들을 서술하는 유의미한 심리 진술은 어떠한 것이든지 오직 행동적인 현상과 물리적인 현상만을 서술하는 진술로 번역될 수 있다.

이 진술은 모든 유의미한 심리적인 표현들을 행동주의적으로 정의할 수 있다는 논제로 더 넓게 공식화 할 수 있다고 한다.

[논리적 행동주의 II] 모든 유의미한 심리적인 표현은 행동적이고 물리적인 표현들, 즉 행동적인 현상과 물리적인 현상을 가리키는 표현들에 의해서만 정의될 수 있다.⁸⁾

우리가 타인에 대한 심리현상을, 즉 타인지각을 수행할 때 타인의 행동이나 언어 등을 통하여 그의 심리상태를 해석한다는 점을 볼 때 이는 행동주의의 논변과 어느 정도 일치하는 것처럼 보인다. 그러나 과연 모든 유의미한 심적인 표현이 물리적인 표현들에 의해서만 정의되는 것이 가능할까?

행동주의가 심리철학 내에서 받는 비판을 우선 살펴보도록 하자. 행동주의를 반대하는 심리철학자들은 심적 상태가 필합하는 행동이라는 것은 있을 수 없다고 말한다. 예를 들어 우리가 인사하고 싶어 하는 욕구를 필합하는 행동이 있다고 말하고 싶다면 그것은 인사하는 행동이라고 말할지도 모르지만 그것이 손을 흔드는 것인지, 아니면 고개를 숙이는 것인지, 혹은 우리가 손을 흔든다면 다른 사람이 불쾌해 할 것이라고 생각했을 때도 손을 흔들 것인지, 인사하고 싶어 하는 욕구가 있다고 해서 꼭 인사를 하게 되는 것인지, 인사를 할 때 발화되는 언어들에 언제나 같은 것인지 등등 여러 가지 의문이 생기며, 현재의 문화나 관습, 습관, 규약, 상호간의 기대, 행위자의 믿음과 의도들, 그리고 다른 많은 사회적 및 심리적인 요소들로 이루어진 거대한 네트워크가 우리의 행위를 결정짓기 때문에 심적인 표현을 순전히 물리적인 행동(운동과 소리)만을 가리키는 행동주의적인 표현으로 옮겨서 정의하는 시도가 성공할 가망성은 거의 없을 것 같다고 한다.⁹⁾

그런데 이는 현상학적인 관점에서 보았을 때 너무나도 당연한 사실이다. 왜냐하면 행동주의가 분석하고 있는 것은 단순히 타인지각일 뿐이며, 앞서 살펴보았듯이 타인지각이란 필증적 명증의 양상을 떨 수 없기 때문이다. 그런데 행동주의의 경우 그 점을 포착하지 못하고 타인에게서 나타나는 행동(운동과 소리)에서 필증적 명증성을 찾으려 하기 때문에 이런 문제점이 나타나는 것이다. 따라서 행동주의를 통한 분석은 난관에 부딪칠 수밖에 없다.

한편, 내적지각의 경우는 어떠한가? 우리는 우리 자신이 고통스럽다는 심적 상태에 있을 때 그것은 단지 우리의 행동으로서만 정의 될 수 있는 것일까? 내가 결박당한 상태에서 고통을 겪고 있는 경우에도 나는 일반적인 경우와 같은 행동을 하는가? 혹은 반대로 내가 고통스러울 때 하는 것과 똑같은 행동을 하고 있지만 정작 '감각질로서의 고통 그 자체'를 느끼고 있지 않더라도 나는 고통이라는 심적 상태에 있는 것인가? 혹은 내가 머릿속으로

$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ 를 생각할 때(혹은 지향할 때) 나는 무슨 행동을 하는가? 그리고 내가 무

8) Jaegwon Kim, 하종호·김선희 옮김, 『심리철학』, 철학과현실사, 1997, p. 59.

9) Ibid., pp. 69-71.

엠티가를 지향할 때 나는 객관적으로 드러나는 어떤 행동을 꼭 하게 되는가?

지금까지의 비판은 행동주의가 실패 할 수밖에 없는 이유를 주로 심리적 상태가 일관된 행동으로 표출되지 못한다는 점에 초점을 맞춘 것이었다. 그런데 만약 우리가 고통을 느끼는 일반적인 경우에 우리는 제자리에서 한바퀴 회전을 하는 경향이 있으며 많은 실험적 사실들에서 그런 결과가 관찰이 되었다고 가정을 해본다면, 그런 경우에 우리는 ‘고통은 제자리에서 한바퀴 회전하는 행동이다.’라는 정의를 내릴 수 있을까? 사실 여기서 제기한 가정에 대한 문제는 심신 동일론의 문제와 다르지 않다. 필자는 이 문제에 대한 논의를 다음 장인 심신동일론 비판으로 미루도록 하겠다.

그럼에도 불구하고 어쨌든 행동을 통한 심리상태에 대한 분석은 타인의 심리상태를 아는 데 상당한 도움을 줄 수 있다. 그리고 필자는 그런 유효한 점까지 부정하려는 것이 아니다. 필자가 여기서 부정하려는 것은 행동을 통해 타인의 심리상태를 파악하려고 하는 방법이 아니라 바로 심리적 표현이 행동으로써만 정의 될 수 있다는 주장이다. 행동을 통해 타인의 심리상태를 파악하려고 하는 방법은 본질 직관의 방법을 따라야 하고, 만약 거기에서 물리학과 같이 정밀한 수준의 명증성을 찾으려고 한다면 이와 같은 방법은 아무 의미 없이 걸들게 될 것이다.

마지막으로 행동주의의 데카르트적인 심성 개념에 대한 반동이 무엇인지를 비판적으로 살펴보고 이장을 마치기로 하자. 데카르트로부터 유래한 심성에 관한 전통적인 개념에 따르면 심성이 발생한 주체는 그 현상에 직접적으로 인지적인 접근을 할 수 있는 반면에 타인의 경우 그것을 언어적인 보고나 다른 관찰 가능한 행동에 근거하여 그것을 추론하여야 한다고 본다. 그런데 많은 철학자들과 심리학자들은 심성에 대한 이러한 견해로부터 받아들이기 어려운 결과들이 나오게 된다고 생각해 왔다고 한다. 이런 사람들은, 전통적인 개념에 따르면 우리는 다른 사람의 마음에 관해서 전혀 알 수 없다는 결론이 나오게 된다고 한다. 그리고 이런 심성 개념에 근거해서 봤을 때, 어떻게 우리가 ‘아픔’이라는 말의 의미를 배울 수 있고 다른 사람에게 정보를 알려 주기 위해서 “나는 팔꿈치에서 격심한 아픔을 느낀다.”와 같은 문장을 사용하여 의사소통을 할 수 있는지는 불가사의하다고 한다. 그러한 발언들이 사람들 사이에서 정보를 전달하는 데 사용될 수 있으며 “아픔”과 “비가 내릴 것 같다는 생각”과 같은 표현들이 간(間)주관적인 의미를 가진다는 것은 분명한 사실이기 때문에 마음이라는 것을 단일한 주체만이 들여다볼 수 있는 내면의 사적 극장으로 이해하는 데카르트적 모델이 거짓임이 입증된다고 한다.¹⁰⁾

필자의 견해로는 이 문제가 그리 쉽지 않은 많은 많은 문제이며 해결하기 어려운 난제라고 생각한다. 우선 이 같은 문제에 대해 현상학적으로 접근해 보기로 하자. 우리가 데카르트적 심성 개념에 따랐을 때 다른 사람의 마음에 관해서는 전혀 알 수 없는가? 아니다. 우리는 그것을 전혀 모르는 것이 아니라 단지 필증적 명증의 양상으로서 우리의 내적지각처럼 알 수 없을 뿐이다. 그리고 필증적 명증의 양상에서 아는 것이 아니라고 해서 그것을 전혀 모르는 것이라고는 말할 수 없다. 또한, 앞서 심적인 것의 특징에 대해서 살펴보았듯이 내적 지각과 타인지각은 명증의 양상에서 분명히 다르다. 그런데 왜 행동주의자들은 ‘우리는 다른 사람의 심리상태에 대해서 전혀 모른다는 결론이 나온다.’와 같은 주장을 하는 것일까? 필자의 견해로는 이 문제를 해결하기 위해서는 무엇보다도 용어를 명확히 하는 게 가장 중요하다고 생각한다. 그리고 명확히 해야 할 용어는 바로 “의사소통”과 “간(間)주관적인 의미”이다. 그리고 이 문제는 또 하나의 큰 주제라고 생각되므로 여기에서 다루지는 않도록

10) Ibid., pp. 53-55.

하겠다.

IV. 심신 동일론 비판

심신 동일론은 마음과 두뇌가 동일하다는 이론, 즉 심성을 갖는다는 것은 적절한 구조와 복잡성을 지니고서 작동하는 두뇌를 가진다는 것과 같다는 이론이다. 번개란 대기상의 전기 방전 이외의 어떠한 것도 아니듯이, 심적 사건도 두뇌에서 일어나는 신경(궁극적으로는 물리, 화학적인) 과정 이외의 어떠한 것도 아니라는 것이다. 옛날 사람들이 번개에 관해서는 뭔가 알았어도 전기 방전에 관해서는 아무것도 몰랐음에도 불구하고 그 두 표현들은 동일한 현상을 가리키듯이, “고통”과 “C-신경 섬유에의 작동”이라는 표현 역시, 비록 우리 조상들이 전자에 관해서는 잘 알고 후자에 관해서는 전혀 몰랐더라도 그 둘은 동일한 현상을 가리키는 것이며 고통이란 다름 아닌 C-신경 섬유의 작동이라는 것이 판명된다고 동일론자들은 주장한다.¹¹⁾ 그들은 심적 사건과 두뇌 과정 사이에는 포괄적이고 넓은 상관관계가 체계적으로 성립하기 때문에 그 둘 사이에 경험과학적인 이론적 동일성을 얻을 수 있다고 말한다.

사실 필자가 보기에 심신 동일론은 행동주의와 다른 점이 별로 없다. 앞서 행동주의를 비판할 때 필자가 가정했던 것처럼, 만약 사람들이 고통을 느낄 때 마다 제자리에서 한바퀴를 회전하는 행동이 많은 실험적 사실을 통해 입증된다면 ‘고통은 제자리에서 한바퀴 회전하는 행동이다.’라는 정의가 가능할까? 심신 동일론은 이와 같은 경험과학적인 동일성을 전제로 ‘고통은 C-신경 섬유의 작동이다.’와 같은 주장을 하며 심적 사건과 두뇌 과정이 동일한 것이라 말한다. 필자가 보기에 행동주의와 심신 동일론의 차이는 단순히 그것의 명증성 차이 뿐이다.

심신 동일론자들은 여기서 말하는 동일성이란 ‘ $2+2=4$ ’와 같은 명제에서 찾을 수 있는 선험적인 동일성이 아니라 우리가 경험을 통해서 알 수 있는 경험적인 동일성을 말하는 것이라 주장한다. 따라서 ‘물은 H_2O 이다.’에서 찾을 수 있는 이론적 동일성처럼 ‘고통은 C-신경 섬유의 작동이다.’에서도 그와 같은 이론적 동일성을 찾을 수 있다고 한다. 그런데 여기에서 심신 동일론자들은 물이 H_2O 외에 다름 아니듯이 고통 또한 C-신경 섬유의 작동 외에 다름 아니라고 말한다. 그리고 심성을 물리적 및 생물학적인 속성이나 과정으로 환원시킬 수 있다고 주장하며, 물리적 사실들 이외의 심적 사실이란 없다고 한다. 곧 심적인 표현들이 유용하고 실제로는 필요 불가결한 것처럼 보이더라도 그러한 표현들을 배제한다고 해서 우리의 언어가 지니는 서술력이 상실되지는 않을 것이며 물리 언어가 원리상 모든 사실들을 서술하기에 충분하다고 보는 것이다. 심신 동일론은 강력한 유물론이다.¹²⁾

물론 물은 H_2O 와 다르지 않다. 그러나 고통 역시 C-신경 섬유의 작동 외에 다름 아닐까? 우리는 C-신경 섬유의 작동을 면밀히 살핌으로써 거기에서 우리가 ‘아픈 느낌’이라는 의미가 귀속되어 있다고 생각하는 ‘고통’이란 개념을 찾을 수 있을까? 물론 그렇지 않다. 우리는 분명 ‘아픈 느낌’이라는 하나의 감각질을 가지고 있기 때문이다. 그리고 이 감각질은 바로 내적 지각에 속하며, 주체가 고통을 일으키는 내적 지각을 직관할 때에만 이런 아픈 느낌이라는 의미가 귀속되어 있는 ‘고통’이라는 개념이 찾아질 수 있다. 누누이 말했듯이 내적지각과 타인지각이 다른 것은 분명한 사실인데, 심신 동일론은 모든 C-신경 섬유의 작동

11) Ibid., p. 95.

12) Ibid., p. 113.

이 고통과 동일하다고 봄으로서, 만일 그들의 주장에 일관성이 있도록 하려면 내적지각과 타인지각이 동일하다고 보아야 하는 결과를 낳게 되고 그렇게 되면 당연히 이는 타당하지 못한 주장이 되고 만다.

물리학의 경우 연구대상이 되는 모든 것들은 우리에게 객체로 나타나는 것들이다. 따라서 여기에서는 어떤 상황에서도 동일하게 나타나는 이론을 찾는 것이 가능하다. 질량과 모양만 똑같은 물체라면 우리는 그것을 같은 물체라고 간주하고 물리학을 탐구하여도 전혀 하자가 없다. 그러나 신경생리학적인 연구의 경우는 그렇지 않다. 물론 그것이 단순히 객체로서의 사람들만을 연구한다면 고통이 C-신경 섬유에 의해 작동 외에 다름이 아니라는 결론을 내릴 수도 있다. 그러나 이를 ‘나 자신’에까지 적용시켜 버린다면 어떻게 되는가? 과연 나 또한 단순한 물리적인 객체 외에 다름 아닌 것인가? 심신 동일론은 우리가 무슨 생각을 하든지 무엇을 느끼고 있는지 간에 머릿속에서 일어나는 것은 모두 두뇌현상외에 다름 아니라고 주장하면서 세계에서 주체라는 것은 도무지 찾아 볼 수 없게 만들어 버린다.

심신 동일론의 문제점을 살펴보기 위해 한 가지 예를 들어 설명해 보기로 하자. 세 사람 A, B, C가 있다. 이제 이 세 사람에게 C-신경 섬유를 자극 했더니 모두 고통을 느끼는 것처럼 보였다. 우리는 여기에서 A, B, C의 심리상태를 비슷한 양상으로 추측할 수 있다. 그런데 이제 상황을 바꾸어서, A대신에 나 자신이 들어간다고 생각해 보자. 그리고 나, B, C에 또 다시 C-신경 섬유를 자극한다. 그리고 이 경우에 비로소 나는 ‘아픈 느낌’으로서의 고통을 갖기 시작한다. 자, 여기서 나는 존재가 평생 고통이라는 것을 못 느껴보았다고 가정해 보자. 첫 번째 실험을 통해서도 아무리 A, B, C가 고통스럽다고 말하며 몸부림을 친다고 해도 그 ‘고통’이라는 것이 무엇인지 나는 알 수가 없다. 그러나 두 번째 실험이 진행되는 순간 내 C-신경 섬유가 자극되면서 나는 ‘아픈 느낌’으로서의 고통을 갖기 시작하고 비로소 ‘아, 고통이 이런 것이구나!’하는 것을 알게 된다.

물론 심신 동일론에 따르면 C-신경 섬유의 자극과 고통이 같은 것이기 때문에 첫 번째 실험만을 통해서도 나는 고통이 무엇인지를 알아야 한다. 그런데 우리가 아무리 C-신경 섬유의 자극을 관찰한다고 해도 나의 C-신경 섬유가 자극되기 전까지는 고통이라는 것의 진짜 의미를 알 수가 없다. 그리고 여기에는 고통이라는 감각질이 단독으로 존재하는 것이 아니라 주체의 초월론적 주관과 지향적인 상관관계를 맺고 있기 때문에 이 같은 결과가 나타나는 것이다. 그럼에도 불구하고 아직 필자의 의도를 제대로 파악하지 못한 혹자들은 아마도 이렇게 주장할 것이다. ‘그와 같이 ‘아픈 느낌’으로서의 고통에 대해 알게 되는 것도 결국엔 두뇌작용이다. 우리가 수학적 명제를 생각하든 동일론을 생각하든 모든 것은 두뇌작용이며 그것들은 물리적인 언어로 서술 될 수 있다.’ 아마도 이런 물리적인 태도에 간혀 있는 사람들은 지향성, 구성작용, 초월론적 주관 역시 두뇌작용이거나 두뇌의 일부분이라고 말할 것이다.

정말 그런지 살펴보자. 초월론적 주관이란 앞서 말했듯이 세계 및 세계내 대상을 구성하는 주관이다. 이는 세계에 대한 일반 정립에 대한 초월론적 현상학적 판단중지를 통해 우리에게 나타날 수 있다. 그런데 이를 동일론의 방식으로 설명하자면 두뇌에서 초월론적 주관에 해당하는 부분이 있어서 그것들이 세계 내대상을 구성하게 된다고 말해야 한다. 그런데 그렇게 설명하면 초월론적 주관이 세계 내대상으로서만 존재하게 되는 것이 아닌가? 그렇다면 동일론자들은 아마 물리계에서 벗어날 수 있는 초월론적 주관이라는 것을 부정할 것이며 또다시 물리적인 태도에 간혀서, 세계를 구성하는 주관조차도 물리적인 것에서 자유로울 수는 없다고 말할 것이다.

이를 반박하기 위해 또 다른 예를 들어보겠다. 나 자신이 복제가 되어 A라는 개체가 생겼다고 가정해 보자. 이제 나와 A라는 두 개체가 두뇌에 있어서 물리적으로 완전히 똑같은 작용이 일어나고 있다고 생각해 보자. 이때 나와 A의 세계 내 대상을 구성하고 있는 초월론적 주관도 역시 동일한 것인가? 전혀 그렇지 않다. 아무리 물리적인 것에서 우리가 동일함을 찾을 수 있다고 하더라도 그 둘은 엄연히 다른 개체이다. 다시 말해 물리적인 것이 복제가 된다고 하더라도 초월론적 주관으로서의 나의 주체까지 복제가 되는 것은 아닌 것이다. 즉, 내가 물리적으로 복제가 된다고 해서 동일한 나의 주체가 두개가 되는 것은 아니며 복제된 인간은 ‘초월론적 주관으로서의 나의 주체’가 아니라 그것은 하나의 객체로써 나에게 나타나는 것이다. 또한, 나 자신은 이런 초월론적 주관이 있기 때문에 자신의 초월론적 주관에서 각종 지각과 같은 구성작용이나 감각질이 타인과는 전혀 다른 직접적인 양상에서 나타나며 모든 심적인 것은 언제나 초월론적 주관과 지향적인 관계를 맺고 있기 때문에 단순히 물리적인 것으로 환원될 수 없는 것이다.

심신 동일론자들은 자신이 전제로 삼고 있는 물리적인 태도로부터 심적인 것의 모든 측면을 설명하려 하는데 그것은 현상학적 입장에서 보았을 때 그것들이 전제로 삼고 있는 형식적 존재론과 영역적 존재론, 나아가서는 초월론적 현상학 까지 넘어서는 주장을 하고 있는 것이다. 이것이 바로 수행적 모순이다. 그리고 지금까지의 비판은 역시 물리적인 태도를 전제로 삼고 있는 행동주의에도 똑같이 적용 될 수 있다.

그런데 혹시 필자가 ‘동일론을 주장하는 사람들도 초월론적 주관을 갖고 있다.’라는 선입견이 있는 것은 아닐까? 심신 동일론자들은 사실은 초월론적 주관이 없는 자동기계들이라서 동일론의 주장이 그들에게 있어서는 타당하며 일관성이 있을 수도 있진 않을까? 만약에 심신 동일론자가 ‘나는 초월론적 주관이 없고 의식도 없는 자동기계이므로 나의 주장은 정당화 될 수 있다.’고 주장한다면 필자는 거기에 대해서 뭐라 할말이 없을 것이다. 물론 주관인 없는 기계의 말이 주장이 될 수 있는지 또 그것이 정당화 되는지도 문제이지만, 그러나 그럼에도 불구하고 물론 필자의 믿음은 그들이 정말 주관인 없는 것이 아니라 단지 잘못된 태도에 빠져서 허우적거리고 있다는 것일 것이다. 만일 이 글을 읽는 독자들이 의식이 없는 자동기계라면 필자의 의견은 받아들여질 수 없을 것이다. 그러나 독자 자신이 분명히 초월론적 주관이나 의식이 있는 존재라면 그가 자신에게 심신 동일론을 적용하는 것이 정당하지 못하다는 사실은 두 말할 필요도 없다.

끝으로 심적인과정으로부터의 동일성에 대한 비판을 하고 이장을 마치자. 심적 인과란 심적인 것과 물리적인 것들 간의 인과관계를 말한다. 예를 들어 고통을 느껴서 우리가 몸을 움직이면 바로 심적인 것으로부터 물리적인 것으로의 인과가 이루어진 것이라 보는 것이다. 그런데 이 세상은 우리의 신체까지 포함해 물리적인 것으로 구성되어 있으며 물리적인 것은 인과적인 폐쇄성을 갖기 때문에 심적인 것을 인과의 어느 위치로 놓아야 하는 것이 문제가 된다. 부수 현상론은 물리적인 인과에서 부수 적으로 심적인 것이 결과로써만 발생하고 다른 것의 원인이 될 수 없다고 말한다. 그러나 동일론자들은 이는 심적인 것의 인과력을 박탈하는 것이며 그렇다면 그것의 실재성도 박탈되는 것이라며 심적인 것의 인과적 효력을 남기기 위해서는 심적인 것과 물리적인 것을 동일하게 보아야 한다고 주장한다. 예를 들어 바늘에 찔림이라는 물리적 상태가 통증과 동일한 어떤 신경상태 N의 원인이 되고 그것의 결과로서 물리적인 움츠림이 일어난다는 것이다.¹³⁾

필자가 보기에 이런 주장은 시작부터 잘못되었다. 심적인 것을 인과관계로만 파악하려는

13) Ibid., pp. 263-264.

선입견을 가지고 들어왔기 때문이다. 그 때문에 심적인 것과 물리적인 것이 같아야 한다는 엉뚱한 결과가 나오게 되는 것이다. 세계가 물리세계 외에 다른 아나라는 것을 동일론자들은 도대체 어떻게 안 것일까? 그것은 근거가 불충분한 선입견 외에 다른 아니다. 솔직히 필자는 이 주장과 근대철학자인 로크의 기계론의 차이를 별로 느끼지 못하겠으며 도대체 이런 주장이 현대에 와서 왜 제기가 되는 것인지 사실 좀 이해할 수가 없다. 어찌됐든 물리적인 태도에 대해서 판단중지를 하고 심적인 것의 본질을 의미 지향성으로서 파악한다면 이런 문제점은 말끔히 사라지게 된다. 물론, 물리세계의 인과적 폐쇄성이라는 것 역시 논의가 필요한 주제이다. 필자는 이 글의 마지막 장에서 이에 대해 다시 논의할 것이다.

지금까지 심신 동일론에 대한 비판을 살펴보았다. 여기에서 필자는 고통이 C-신경 섬유 자극과 아무 관련이 없다는 것을 주장하는 것은 아니다. 그것은 경험과학의 지식으로서 유의미 하게 활용될 수 있다. 그러나 이런 단순히 경험적인 근거를 토대로 모든 심적인 것이 물리적인 것으로 환원될 수 있다는 주장을 해서는 안 될 것이다. 몇 백 년 전만해도 물리학이 발전하여 초기조건만 알게 된다면 모든 것을 예측할 수 있을 것이라 믿었지만 현대 물리학에서는 인간의 인식 능력으로는 모든 초기 조건을 알지 못한다는 것이 밝혀졌듯이, 생물학 역시 그것이 발전함에 따라 단순히 실험 과학적 방법으로는 밝혀낼 수 없는 많은 부분들이 나타나게 될 것이라고 필자는 조심스레 추측해 본다.

V. 기계 기능주의 비판

기계 기능주의는 심적 속성들의 다수 실현 가능성이라는 아이디어로부터 비롯되었다. 여기서 다수 실현 논변이 무엇인지를 간단히 살펴보기 위해 앞서 심신 동일론에서 살펴본 고통과 C-신경 섬유 자극의 예를 살펴보도록 하자. 동일론에서는 C-신경 섬유 자극과 고통이 동일하다고 말한다. 그렇다면 한 생물체가 C-신경 섬유를 가지고 있지 않거나 적절한 생물학적인 구조로 된 두뇌를 가지고 있지 않는 한, 그 생물체는 고통을 느낄 수 없다는 결론이 나온다. 그렇지만 인간의 두뇌와 매우 다른 두뇌를 가지고 있으면서 고통을 느낄 수 있는 생물체들(예를 들면, 파충류나 연체 동물 등)이 있을 수도 있으며, 혹은 우리 인간처럼 풍요로운 정신생활을 영위하고 지능을 갖추었지만 생체는 탄소로 되어 있는 외계인이 있을 수도 있다. 고통이 물리적 상태 C와 동일하다면 그것은 모든 가능한 생물체와 조직 체계 내에서 C상태와 동일 하여야 하기 때문에 이런 다수 실현 논변은 심신 동일론의 몰락-물론 아직도 그것을 주장하는 사람이 있긴 하지만-을 초래하였다.¹⁴⁾

따라서 기계 기능주의자들은 마음이 어떤 생리학적 물리 상태와 동일한 것이 아니라 일종의 튜링 기계¹⁵⁾(혹은 확률적 자동기계)로 간주할 수 있다고 주장한다. 그들의 중심적인 아이디어는 어떤 것이 심성을 갖는다는 것은 그 것의 심적 상태들-즉 심적 상태 유형들-을 기계표의 내적 상태들과 동일시함으로써, 그것이 물리적으로 실현되는 적절히 복잡한 튜링 기계가 된다는 것이다.

이 생각을 설명하는 다른 방식은, 기계 기술의 개념을 사용하는 것이다. 한 유기체의 기

14) Ibid., pp. 126-128.

15) 필자는 여기에서 튜링 기계에 대한 자세한 설명은 생략하였다. 튜링 기계에 대한 설명은 다음 문헌에서 찾아볼 수 있다. (Jean Lassègue, 임기대 옮김, 튜링, 東文選, 2003, pp. 83-94), (Jaegwon Kim, 하종호·김선희 옮김, 『심리철학』, 철학과현실사, 1997, pp. 144-151.)

계 기술-적절한 복잡성을 지닌-인 튜링 기계가 존재하고 그것의 심적 상태 유형들이 그 튜링 기계의 내적 상태들과 동일시될 경우에만, 한 유기체는 심적 특성을 갖는다. 이 모든 것들은 물론 적절하게 선택된 입력/출력 명세서에 의존한다. 따라서 한 유기체나 체계가 심리적 특성을 갖는다는 것은 그것이 적절한 하나의 튜링 기계를 실현한다는 것이다. 그것은 단지 심성을 가진 어떤 것도 하나의 적절한 기계 기술을 갖는다는 것만을 말하지 않는다. 기계 기능주의는, 그것이 적절한 종류의 기계 기술을 갖는다는 것이 그것의 심성을 구성한다는 보다 강한 주장을 한다.

그들에 따르면 심성, 혹은 마음을 갖는다는 것은 적절한 튜링 기계를 실현하는 데 있다. 그러므로 우리를 심성을 가진 피조물들로 만드는 것은, 바로 우리가 튜링 기계들이라는 사실이다. 기능주의는 어떤 구조적 복잡성을 지닌 두뇌를 갖는 것이 심성을 위해 중요하다는 것을 인정할 것이다. 그러나 그 두뇌의 중요성은 정확히 그것이 하나의 튜링 기계라는 점에 놓여 있다. 우리의 심성을 구성하는 것은 우리 두뇌의 계산적 속성들이지, 그것의 생물학적 속성들이 아니다. 즉, 우리의 두뇌가 우리의 마음인 이유는, 그것이 컴퓨터이기 때문이지 그것이 유기적인 생물학적 구조이기 때문이 아니라는 것이다.¹⁶⁾

현재 많은 이슈가 되고 있는 인공지능이 바로 기계 기능주의로부터 비롯된 것이라고 볼 수 있다. 인공지능을 연구하는 다수의 사람들은 그것이 인간과 같은 마음을 갖게끔 만들 수 있을 것이라고 생각한다. 그런데 여기에서 가장 처음 부딪치는 문제가 바로 목적론적인 문제이다. 인간은 어떤 목적에 적합한 것으로 보이게끔 입-출력이 일어나는데, 기계들은 맹목적으로 입-출력이 나타난다는 것이다.

지금까지 살펴본 기계의 구조에 초점을 맞추고 있는 기능주의 역시 물리주의적 입장에 기반을 두고 있으며, 이에 대해서는 앞서 심신 동일론에서 했던 비판을 그대로 적용할 수 있다. 이는 물리적인 태도에서는 주어질 수 없는 의미지향성과 구성작용, 초월론적 주관의 예시만 들어 보아도 충분하다.

그럼에도 불구하고 동일론에서 했던 비판을 한 가지만 다시 해보도록 하자. 인간의 경우 C-섬유의 자극이 고통을 불러일으킨다는 사실이 경험적으로 관찰 되었으며, 어떤 생물의 종 A의 경우 D-섬유가, B의 경우 E-섬유가, C의 경우 F-섬유의 자극이 고통을 불러일으키는 것 같다는 사실이 관찰되었다고 가정하자. 그리고 이 고통의 상태에 있는 경우 모든 생물체가 A라는 유사한 내적 구조에 있는 것이라 파악했다고 역시 가정해보자.(물론, 이런 일은 불가능 하다고 생각하지만) 자 그렇다면 우리는 ‘고통은 A라는 내적구조에 있는 상태이다.’라고 결론 내릴 수 있을까? 물론, 아니다. 우리가 아무리 실험에 근거를 두고 이와 같은 사실들을 파악했다고 해도 나는 아직 ‘아픈 느낌’으로서의 고통이 무엇인지는 모른다. 그리고 이는 초월론적 주관의 의미 지향성을 통해서만 주체에게 직접적으로 알려지는 것이다. 이런 점에서 볼 때 기계 기능주의 역시 심신 동일론과 마찬가지로 주체를 말살시켜버리는 결과를 낳게 된다.

이제 튜링 테스트¹⁷⁾에 관해서 살펴보자. 그의 제안은, 모든 면으로 볼 때 충분히 심리적인 평균 인간들과 비교하여 계산 기계들(컴퓨터)의 수행 능력을 관찰하는 테스트를 하는 것이다. 그것들이 어떤 수준의 지적 과제들을 인간만큼 잘 수행할 수 있다면, 그것들은 인간만큼 심리적인 존재라고 판단되어야 한다는 것이다. 여기에서 튜링이 제안한 게임은 간단히

16) Jaegwon Kim, 하중호 · 김선희 옮김, 『심리철학』, 철학과현실사, 1997, pp. 155-162.

17) 팔자는 튜링 테스트의 핵심만 압축하여 설명하였다. 이에 대한 좀 더 자세한 설명은 다음을 참조하기 바란다. (Jean Lassègue, 임기대 옮김, 튜링, 東文選, 2003, pp. 167-187)

말해서 질문자가 자판의 터미널을 통하여 대화를 함으로써 우리가 인간과 말하고 있는지 계산 기계와 말하고 있는지 파악하는 것이다. 그리고 우리가 그것의 정체를 파악하는 정도가 평균이하라면 그 기계도 인간의 심성을 갖는 존재로 인정해야 한다는 것이다. 그리고 여기에서의 튜링의 논제는 다음과 같다.

[튜링의 논제] 만일 두 체계들이 입-출력에서 동일하다면, 그것들은 동일한 심리적 지위를 갖는다. 특히, 그중 하나가 심적 존재일 경우에만 다른 하나도 심적이다.¹⁸⁾

심리철학 내에서는 아마 이 논제가 단순한 양상에서 비판 받고 있는 것 같다. 예를 들어 5+7을 계산기에 입력했을 경우, 논리적 계산 과정을 거쳐서 12를 출력하는 기계와 단순히 모든 덧셈 연산에 대한 결과가 저장되어 있는 계산기에서 출력되는 12라는 결과는 비록 입-출력이 동일할지라도 그 내적구조가 다를 수 있다는 것이다.

그런데 필자가 보기에는 이 문제가 그리 쉬운 문제 같지 않은 않다. 물론 우리가 그것의 내적 구조를 밝힐 수 있는 경우에는 그것이 다르다는 주장이 가능하다. 그런데 다음과 같은 반 사실적 가정은 어떠한가? 외계인이 인간과 거의 비슷한 기계를 만들어서 우리 안에 심어 놓았다. 물론 그것은 기계이기 때문에 의식이 발현되지 않지만 인간 사회에서 하나의 인간의 역할을 하고 있으며 우리는 겉으로는 그것이 진짜 인간과 구분이 가지 않기 때문에 아마도 그를 인간으로 취급할 것이고, 우리가 다른 인간에게서 느끼는 각종 감정을 그에게서도 느낄 것이다.

물론, 그럼에도 불구하고 그는 주관과 의식이 있는 인간은 아님에 분명하다. 그러나 우리가 내 주체로서의 초월론적 주관이 아니라 타인에 대해서 생각해 볼 경우, 입-출력에서 다른 인간들과 똑같아 보이는 그 기계는 나에게 있어서 타인들과 동등한 지위를 갖게 된다는 것을 부인할 수 없다. 이와 비슷하게 다른 사람들 중의 몇몇이 사실은 자동기계일 수도 있다는 반 사실적 가정도 가능하며 이 경우 우리는 우리에게 주어지는 사태에서 그 차이를 찾을 수 없다면 그것들에게 동일한 지위를 부여해야 한다는 결론을 얻게 된다. 물론 그럼에도 불구하고 그들에게 부여되는 지위는 내가 초월론적 주관을 가지고 있다고 믿고 있는 타인의 지위일 뿐이며 초월론적 주관을 필연적으로 갖고 있는 내 주체의 지위는 아니다.

이제 마지막으로 이 문제의 연장선상에 있는 서얼의 “중국어 방” 논증을 살펴보고 기계-기능주의에 대한 논의를 마치도록 하자. 이 논증은 중국어를 전혀 이해하지 못하는 사람이 일련의 기호들과 그와 다른 기호들을 변형하는 규칙들을 가지고 있는 어떤 방에 갇혀 있다는 가정으로 시작한다. 이 방에 있는 사람은 일련의 중국어 문자들을 받을 때마다 신속히 작업을 하고 그 즉시 적절한 일련의 중국어 문자들을 내보낼 수 있을 정도로, 중국어 표현들을 조작하는데 매우 숙련되어 있어 방 밖에 있는 사람의 관점에서 보면 정말 중국어를 이해하는 사람이 그 방 안에 갇혀 있는 경우와 완전히 일치한다. 그러나 사실 이 사람은 중국어를 이해하는 것이 아니며 이 기호들의 형태나 “통사론”에 기초해서 그것들을 조작하는 것 뿐이다. 그러나 진정한 이해란 “의미론”, 즉 그 기호들이 표상하거나 의미하는 바를 아는 것을 포함하는 것이기 때문에 입-출력이 동일하다 하더라도 그것을 이해했다고는 볼 수 없다는 것이다. 여기서 바로 중국어 방에 있는 사람은 튜링 기계(컴퓨터)에 비유되고 있는 것이다.

만약 우리가 컴퓨터의 경우처럼 그 것의 내부구조를 알 수 있다면 이런 주장은 가능할

18) Jaegwon Kim, 하종호·김선희 옮김, 『심리철학』, 철학과현실사, 1997, p. 173.

것이다. 다시 말해 우리와 같이 있던 사람이 수상해서 잡아 조사해보았더니 자동기계임이 밝혀졌다면, 그 자동기계는 초월론적 주관이나 의식을 가지고 있지 않는 것이라 “기대해 볼 수” 있다. 그러나 문제는 우리가 그것을 구분하지 못한다는 데에 있다. 다시 말해 나와 함께 사회를 구성하는 타인들이 “실제로” 누군가는 주관이 없고 누군가는 주관이 있다는 반사 실적 가정이 가능하지만 우리가 나 외에 타인들이 주관이나 의식이 있다거나 없다고 증명하는 것이란 불가능 하다. 그러나 그럼에도 불구하고 나의 주관에게 주어지는 사태가 같다면 그것은 나의 주관에 같은 의미로써 다가올 것이다. 다시 말해 정말 타인과 같이 행동하고, 말하고, 먹는 기계가 있다면 그 기계는 비록 의식이 없을지라도 우리가 그것이 기계라는 것을 알기 전까지는 나에게 타인으로서 나타날 것이며, 내 주관에서 그 기계는 타인과 동등한 지위를 가질 수 있을 것이다.

그런데 필자는 그런 기계가 만들어 질 수 있다는 것에 회의적이다. 아마 기술이 발전해서 원초적인 수준에서의 인간과 비슷한 기계들은 만들어 질지 모르나 인간 사이에 섞여 놓아도 구분이 안갈 정도의 수준은 기술적인 문제가 아니라 인간의 초월론적 주관의 성격을 보았을 때, 애당초 불가능 할 것 같다. 혹자들은 복제인간이 탄생한다면 우리와 입-출력의 방식이 똑같지 않느냐는 주장을 하는 사람도 있을지 모르겠다. 물론 같을 것이다. 그리고 필자는 필자의 앞에 복제인간이 나타나더라도 그 또한 엄연한 인격체로 대할 것이다. 그것은 기계가 아니기 때문이다.

VI. 세 가지 근본 전제에 대한 비판

지금까지 세 사조에 대한 비판을 살펴보았다. 그런데 이 세 사조는 세 가지 측면에서 동일한 전제를 가지고 있다. 이제 최소 물리주의로 표현되는 이 세 가지 전제에 대한 비판을 토대로 건물의 밑바닥을 뒤흔들어 보기로 하자.

[심신 수반] 모든 물리적 속성들이 똑같은 두 사물(대상, 사건, 유기체, 사람 등)들이 심적 속성에서도 다를 수 없다면 심적인 것은 물리적인 것에 수반한다. 즉 물리적 식별 불가능성은 심리적 식별 불가능성을 필함(entail)한다.¹⁹⁾

심신 수반이론은 물리적인 것이 같다면 심적인 것도 같다는 것이다. 이는 물론 심적인 것과 물리적인 것이 완전 동일하다는 심신 동일론과는 다르다. 우리는 심신 동일론을 살펴보면서 비록 물리적인 속성들이 완전히 동일하더라도 주체의 초월론적 주관과 의미 지향성, 감각질과 같은 측면에서 그것이 같지 않을 수 있음을 살펴보았다.

[반(反) 데카르트 원리] 순전히 심적인 존재(예컨대 데카르트가 말한 영혼)란 있을 수 없다. 즉 물리적 속성을 갖지 않고서는, 따라서 물체가 되지 않고서는 어떠한 것도 심적 속성을 가질 수 없다.²⁰⁾

앞서 말한 수반 원리는 순전히 심적인 존재를 부정하고 있지는 않기 때문에 이 원리가 반 데카르트 원리를 함축하고 있는 것은 아니다. 그러나 수반원리에 따르면 만일 순전히 심

19) Ibid., p. 26.

20) Ibid., p. 27.

적인 존재가 있다고 하더라도 그것은 기껏해야 하나뿐이라는 결론이 나오게 된다. 반 데카르트 원리는 오직 심적 속성들만 갖는 것은 애초에 존재할 수 없다는 입장으로 수반 논제를 받아들이는 사람이라면 누구나 받아들여지게 되는 입장이다. 현상학적인 관점에서 역시 반 데카르트 원리는 수용될 수 있다. 후설은 『데카르트적 성찰』에서 다음과 같이 말한다.

그 전회는 자아를 사유실체(substantia cogitans), 즉 분리된 인간의 ‘정신 즉 영혼’(mens sive animus)으로 만들고, 인과원리에 따른 추론의 출발점으로 만든다. 간략히 말하자면, 이 전회를 통해 데카르트는(여기서는 아직 명백히 될 수 없는) 불합리한 초월론적 실재론(transzendentes Realismus)의 시조가 되었다.[...]

[...]나는 이미 자연적 자기 경험 속에서 스스로를 인간으로서 발견하게 되는 ‘나’는 아니며, 내적 자기 경험, 즉 순수하게 심리학적인 자기 경험의 순수한 요소들로 추상적으로 제한함에서 자기 자신의 순수한 ‘마음=영혼=지성’(mens sive animus sive intellectus)으로 발견되는 인간 또는 그 자체만으로 파악되는 영혼 자체도 아니다.²¹⁾

그러나 여기서 주의해야 할 점은 그것이 심리철학에서처럼 최소 물리주의를 표방하기 위해 수용되는 것이 아니라는 점이다. 물리주의자들은 심적 속성들이 물리적 속성들에 의존한다고 보는 반면 초월론적 현상학의 경우에는(물론 이 경우에도 역시 순전히 심적인 존재로서의 데카르트적 영혼을 상정하는 것은 아니지만) 단순히 심성이 물리적인 것에 의존한다고 보는 것이 아니라 존재자와 그를 향한 의식이 본질 구조에 있어 독특한 상관관계를 가진다고 보는 것이다.

이 양자 사이의 상관관계에 의하면 의식의 대상이 없는 의식은 없으며, 의식작용이 없는 의식의 대상도 우리에게서 경험되지 않는다. 이 것이 바로 노에시스-노에마 상관관계이다. 여기에서 “노에시스”(Noesis)란 후설이 “무엇에 관한 의식”으로서의 지향성을 ‘사유’라는 의미를 지닌 희랍어 명사를 빌어 부르는 것이며 “노에마”(Noema)란 의식이 향하고 있는 “그 무엇”, 즉 지향성의 대상적 상관자를 ‘사유된 것’이라는 의미를 지닌 희랍어 명사를 빌어 부르는 것이다. 그리고 여기에서 우리는 노에시스-노에마 상관관계가 보여 주듯이 지향성을 탐구주제로 삼는 현상학이 대상과의 관련성을 도외시한 채 의식작용만을 분석하는 것이 아니며, 어디까지나 노에시스로서의 의식작용과 더불어 노에마적 상관자인 대상의 구조도 함께 분석하는 것임을 알 수 있다.²²⁾

[심신 의존] 임의의 대상이 어떠한 심적 속성들을 갖는지는 그것이 어떠한 물리적 속성들을 갖느냐에 의존하고 그것에 의해서 결정된다. 즉 한 대상의 심적 성격은 그것의 물리적 성격에 의해서 전적으로 결정된다.²³⁾

심신 의존 원리는 심신 수반 원리를 필함하고 있기 때문에 추가로 반박할 필요가 없다. 이 원리 역시 물리적인 태도에서 벗어나지 못하고 있다.

다음으로 심신인과에서 전제로 삼고 있었던 물리계의 인과적 폐쇄성에 대해서 살펴보자. 물리계의 인과적 폐쇄성은 말 그대로 물리적 사건을 포함하는 어떠한 인과연쇄도 물리계의 경계를 넘어서 비물리계로 넘어가지 못한다는 원리이다. 따라서 물리적인 것으로 구성되어

21) Edmund Husserl · Eugen Fink, 이종훈 옮김, 『데카르트적 성찰』, 한길사, 2002, pp. 68-69.

22) 이남인, 『현상학과 해석학』, 서울대학교출판부, 2004, pp. 91-92

23) Jaegwon Kim, 하중호 · 김선희 옮김, 『심리철학』, 철학과현실사, 1997, p. 28.

있는 인간의 몸 역시 물리계안에서의 인과연쇄에 귀속되어 있다는 주장이 나온다. 이는 로크의 기계론이나 흄의 결정론과 비슷하며 여기에서 우리는 칸트의 세 번째 이율배반을 쉽게 떠올릴 수 있다. 인과율은 감성계에서만 적용될 수 있다는 것이 그 요지이다. 어찌되었든 그들의 선입견은 물리계의 인과적 폐쇄성을 이용해서 세상의 모든 것을 설명하려 하는 것이다. 즉, 세상이 물리적으로만 이루어져 있다는 선입견이 깔려 있기 때문에 다른 측면들을 보지 못하는 것이며 자기 자신의 존재마저 단순한 C, H, O, N, S의 집합으로 환원시켜 버리고 경험적 사실을 넘어서는 이념의 영역에 까지도 왈가왈부 하는 수행적 모순을 저지른다. 또한, 필자는 물리세계의 인과적 폐쇄성 자체가 올바른 테제인지조차 의심스럽다. 그것이 물리학에서 쓰이는 일반적인 경우를 제외하고 과연 모든 경우에 적용될 수 있는 것인지 우리는 어떻게 아는가? 우리는 아무런 영향을 받지 않고 자유로이 움직이고 있는 입자를 분명히 상상할 수 있으며, 물리적인 세계가 인과적으로 폐쇄되어 있다는 사실은 단순히 경험적으로 얻어진 것으로 보인다. 그리고 아마도 이 문제는 현대물리학에서 현재 활발하게 논의되고 있을 것이다.

이제 예상되는 반론에 대한 몇 가지 답변을 한 후 이 글을 마치기로 하겠다. 우선 첫 번째는 내적지각이 타인지각에 비해서 필증적 명증의 양상을 언제나 띠는 것인지에 대한 반론이다. 혹자들은 고통과 같은 심적 사건에 대한 지식은 타인지각에 비해 내적지각만이 필증적 명증의 양상을 띠 수 있지만 물리적인 사건에서도 과연 그런 것인지 의문을 내비친다. 예를 들어 내가 벽에 붉고 둥근 색점이 있다는 것을 “직접”알 수 있다면, 그 벽과 적당한 거리에 서 있는 나나 다른 사람들 모두가 그렇게 “직접”알 수 있다는 것이다. 이런 반론은 내적지각과 타인지각을 전혀 이해하지 못한 반론이다. 나는 둥근 색점이 있다는 것을 물론 “직접”아는 것이다. 그러나 타인의 경우도 그가 둥근 색점을 보고 있다는 것을 내가 “직접”아는 것인가? 전혀 그렇지 않다. 우리는 타인의 눈동자나 그의 시선을 보고 그가 둥근 색점을 보고 있다고 해석하는 것이다. 따라서 타인지각은 해석의 과정이 들어가기 때문에 내적지각보다 명증성이 떨어지는 것은 분명하며 이런 반론이 정당하지 못하다는 것은 쉽게 알 수 있다.

또 하나의 반론은 과연 심리현상 일반의 본질구조가 지향성이나 하는 것이다. 이 사람들은 통증이라든지, 시각, 촉각, 배고픔과 같은 감각적 상태는 심적 상태임에도 불구하고 아무것도 지향하고 있지 않느냐는 주장을 할 수 있다. 초중기 후설의 경우에는 아마 이와 같은 견해를 인정할 것이다. 그러나 그의 후기 현상학에서는 이에 대한 지향성 개념의 수정이 일어난다. 이에 대해 간단히 살펴보자면 다음과 같다. 비록 감각경험이 객관적인 시간과 공간에 존재하는 대상과 아무런 관련도 맺고 있지 않지만, 그럼에도 불구하고 그것은 나름대로 이러한 경험을 수행하는 자아에게 “자극”을 주는 대상적인 것과 관련을 맺고 있다는 점에서 넓은 의미의 지향성을 띠고 있다. 그리고 그것의 수동적 종합 구조를 분석할 때 핵심적인 역할을 담당하는 것이 “촉발”(Affektion)이라는 개념이다. 촉발은 수동적 종합의 영역에서 자아로부터 감각자료로 “향하는”수동적 의식이며 촉발이 이처럼 “주관으로부터 ~로 향함”이라는 계기를 지니고 있기 때문에, 후설은 “대상에 대해서 우리는 촉발을 대상을 향한 지향성이 일깨워짐이라고 표현할 수도 있을 것이다.”라고 말한다. 즉 이 지향성은 “자기동일적 대상을 향한 자아의 의식적 관계”로 규정된 지향성과는 달리 수동적 양태에 있어서의 지향성인 “수동적 지향성”이다.²⁴⁾

우리는 지금까지 심리철학에서 논의 되는 주요 주장들이 어떤 점에서 문제점을 가지고

24) 이남인, 『현상학과 해석학』, 서울대학교출판부, 2004, pp. 291-297

있는지 살펴보았으며 심리현상을 제대로 파악하기 위해서는 어떤 태도를 가지고 어떤 방법을 써야 하는지도 알아보았다. 물리주의적인 태도를 전제로 하여 나 자신의 주체마저도 말살시켜 버리는 이런 사조는 분명 지양되어야 할 것이다. 그리고 현상학의 기본 모토인 모든 선입견을 제거-모든 선입견을 제거할 수 있다는 전제도 선입견일 수 있지만, 어쨌든 배제 가능한 모든 선입견을 제거하고-하고 ‘사태 자체로’ 돌아갈 때에만 비로소 심리현상에 대한 올바른 이해와 파악이 이루어질 수 있을 것이다.

참고문헌

이남인, 「발생적 현상학과 상호주관성의 문제」, 서울: 관악사, 2001

이남인, 「후설, 엄밀학으로서의 철학」, 서울: 관악사, 2005

랄프 루드비히, 박중목 옮김, 『쉽게 읽는 칸트-순수이성 비판』, 이학사, 1999

이남인, 『현상학과 해석학』, 서울대학교출판부, 2004

Descartes R., *Meditationes de Prima Philosophia*. (이현복 옮김, 『성찰』, 문예출판사, 1997)

Husserl E., *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*. (이종훈 옮김, 『유럽학문의 위기와 선형적 현상학』, 한길사, 1997)

Husserl E., *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*. (최경호 옮김, 『순수 현상학과 현상학적 철학의 이념들 : 순수 현상학의 입문 일반』, 문학과 지성사, 1997)

Husserl E. · Fink E., *Cartesianische Meditation und Pariser Vorträge*. (이종훈 옮김, 『데카르트적 성찰』, 한길사, 2002)

Kim Jaekwon, *Mind in a physical world : an essay on the mind-body problem and mental causation*. (하중호 옮김, 『물리계 안에서의 마음』, 철학과현실사, 1999)

Kim Jaekwon, *Philosophy of Mind*. (하중호·김선희 옮김, 『심리철학』, 철학과현실사, 1997)

Lassègue, J., *Turing*. (임기대 옮김, 『튜링: 인공지능 창시자』, 東文選, 2003)

Wittgenstein L., *Philosophische Untersuchungen*. (이영철 옮김, 『철학적 탐구』, 서광사, 2002)