

## 무법칙적 일원론과 심성 인과

하 종 호 (송실대)

‘심성 인과(mental causation)’란 심적 사건을 포함하는 인과 관계를 의미한다.<sup>1)</sup> 심적 사건이 원인이 되고 물리적 사건이 결과가 되는 경우, 물리적 사건이 심적 사건의 원인되는 경우, 심적 사건이 심적 사건의 원인이 되는 경우가 바로 심성 인과 관계가 성립하는 경우들이라고 할 수 있다. 물을 마시고 싶다는 생각이 들어서 냉수를 마시러 의자에서 일어나는 경우가 첫 번째의 경우에 해당하고, 뜨거운 냄비에 손을 댔다가 통증을 느끼는 경우가 두 번째의 경우이고, 손에서 통증을 느끼고서 의사한테 가야겠다는 생각이 든 경우가 세 번째 경우에 해당한다.

심성 인과의 문제는 이미 데카르트를 비롯해서 여러 고전적인 철학자들이 해결을 시도하였던 문제로써 심적인 것과 물리적인 것간의 관계를 설명하려는 어떠한 이론도 이 문제에 대한 답변을 제시하여야만 한다. 그런데 근래에 가장 각광을 받아 온 데이빗슨의 무법칙적 일원론<sup>2)</sup>에 대해서 그 이론이 이 문제를 해결하지 못한다는 비판이 제기되어왔다. 본 논문에서는 무법칙적 일원론이 이 비판을 극복할 수 있는지를 검토하고자 한다. 이를 위해서 I절에서는 데이빗슨의 무법칙적 일원론이 심성 인과의 문제를 극복하지 못한다는 비판을 받는 이유를 제시하고, II절에서는 무법칙적 일원론이 이 비판을 극복할 수 있음을 보여주려고 하는 시도들을 논의하였다. 그리고 III절에서는 무법칙적 일원론과 같은 비환원주의적 유물론에 대한 대안으로서 제시된 김재권의 새로운 환원주의적인 전략을 살펴 보고자 한다.

- 
- 1) 이 논문에서는 영어의 ‘mental’을 주로 ‘심적’이라는 말로 옮기되 ‘인과 관계’ 앞에 오는 경우에 한해서 관례에 따라 ‘심성’이라고 옮기겠다.
  - 2) Donald Davidson, “Mental Events,” in *Essays on Actions and Events* (Oxford: Oxford Univ. Press, 1980), pp.207-227; “Psychology as Philosophy,” pp.229-244; “The Material Mind,” pp.245-259.

## I

우리는 매크로플린을 따라서 심물(心物) 관계에 대한 데이빗슨의 입장을 다음 몇 가지의 테제로 요약할 수 있다.<sup>3)</sup>

**인과적 상호 작용의 원리:** (1) 적어도 일부 심적 사건들은 물리적 사건들과 인과적으로 상호 작용한다.

**무법칙적 일원론:** (2) 모든 심적 사건은 물리적 사건이다; (3) 엄밀한 심물 법칙은 있을 수 없다; 그리고 (4) 어떠한 물리적 술어도, 그것이 아무리 복잡한 것이더라도, 심적 술어와 동일한 외연을 가질 수 없다.

**수반 테제:** (5) 심적 특성들은 물리적 특성들에 수반한다.

**심성의 무법칙성:** (6) 심적 사건들을 설명하고 예측할 때 근거가 될 수 있는 엄밀 법칙(strict law)은 없다.

**인과 관계의 법칙적 특성 원리:** (7) 원인과 결과의 관계를 맺는 사건들은 엄밀 법칙에 의해서 포섭된다.

이 테제에 나오는 주요 개념들을 간단히 살펴 보자. 데이빗슨은 사건을 시간 속에서 존재하는 구체적인 개별자로 보고서 이 사건들이 인과 관계를 맺으며 두 사건들이 동일한 원인과 결과를 가질 때 바로 그 때에만 그 사건들은 동일하다고 본다. 어떤 사건이 심적 사건이 되기 위해서는 그 사건이 심적 기술구를 가지거나 그 사건만이 만족시키는 심적인 열린 문장이 있어야 한다. 심적 기술구나 심적 열린 문장이란 심적 술어가 본질적으로 들어있는 기술구나 문장을 말한다. 심적 술어란 심적 동사를 가진 술어로서 이 동사가 술어 안에 들어있음으로 인하여 내포적인 문맥이 만들어진다. 그리고 명제 태도를 표현하는 동사가 심적 동사가 된다. 그러면 어떻게 해서 모든 심적 사건이 물리적 사건이라고 말할 수 있을까? 매크로플린은 이에 대한 데이빗슨의 생각을 단적으로 다음과 같이 압축시켜서 표현한다.

3) 각 테제에 관한 상세한 논의는 다음 논문을 참조하라. Brian P. McLaughlin, "Anomalous Monism and the Irreducibility of the Mental," in *Actions and Events*, ed. Ernest LePore and Brian P. McLaughlin (Oxford: Basil Blackwell, 1985), pp.331-368.

“인과 관계는 엄밀 법칙 하의 포섭을 요구하고, 엄밀 법칙은 물리 법칙이므로 물리적 사건들과 인과적으로 상호작용하는 사건들은 물리적 사건들이다. 그리고 모든 심적 사건은 물리적 사건과 인과적으로 상호 작용하므로 모든 심적 사건은 물리적 사건이다.”<sup>4)</sup>

그런데 데이빗슨의 일원론적인 관점은 심적인 것에 관한 이론이 물리적인 것에 관한 이론으로 환원될 수 있다고 보지는 않는다. 왜냐하면 두 이론을 환원하기 위해서는 교량 법칙(bridge law)이 필요한데 심적인 것과 물리적인 것 사이에는 교량 법칙이 성립할 수 없기 때문이다. 데이빗슨에 따르면 심적인 것의 영역을 지배하는 원리는 합리성의 규범들을 표현하는 원리인 반면에 물리적인 것의 영역은 이 원리의 지배를 받지 않는다고 한다. 합리성의 규범을 표현하는 원리는 규범적인 진리를 표현하는 원리로서 우리가 가지고 있는 기존의 명제 태도에 비추어 볼 때 어떠한 명제 태도를 새로 채택하는 것이 합리적인지를 말해 주는데 이때 정합성, 합리성, 무모순성 등의 조건들이 가해진다. 그러나 물리 이론에서는 이러한 합리성의 규범들이 물리적인 용어로 표현될 수 없으므로 심적인 것에 관한 이론이 물리적인 이론으로 환원될 수 없으며, 그 나름대로의 독자성을 확보할 수 있게 된다.

심성 인과의 문제는 위의 (1), (3), (7)의 테제들이 상충된다는 점에서 비롯된다. 즉 심성 인과성이 성립하기 위해서는 그 인과 관계에 있는 심적 사건들이 법칙을 예화시켜야 하는데 심적 사건들에 관한 법칙은 성립할 수 없다고 하였기 때문이다. 이를 구체적으로 설명하면 다음과 같다. e라는 사건의 원인이 되는 심적 사건 m이 있다고 하자. 데이빗슨에 따르면, 이 사건들간의 인과 관계는 m과 e가 어떤 물리 법칙을 예화함으로 말미암아 성립한다. 즉 m은 특정한 물리적(또는 신경생리학적)인 속성 N을 가지며, e는 P라는 물리적 속성을 가지고 있어서 N류의 사건들과 P류의 사건들을 연결시켜주는 인과 법칙에 의해서 m과 e가 인과 관계를 맺게 된다는 것이다. 그런데 이 경우에 m이 심적 사건이라는 사실은 m이 e와 맺는 인과 관계를 결정함에 있어서 아무 역할도 하지 못하게 된다. m의 인과 관계는 전적으로 m의 물리적 속성들에 의해서 결정될 뿐이며, 그것의 심적 속성(이것을 M이라고 하자)이 인과 관계에 영향을 미칠 수 있는 길은 차단되어 있기 때문이다. 따라서 데이빗슨의 무법칙적 일원론대로 하자면, 심적인 것이 이 세계에서 완전히 제거된다고 하더라도 이 세계의 사건들 간에 성립하는 인과 관계는 하등

4) Ibid., p. 338.

영향을 받지 않을 것이며 이 세계의 인과적인 구조에는 아무 변화도 없게 된다. 이는 곧 심적 속성들이 부수 현상에 지나지 않는다는 결론에 이르게 되는 것이다.

여기서 데이빗슨은  $m$ 의 심적 속성인  $M$ 의 인과적 효력을 부정하든지 아니면 인정하든지 둘 중의 하나를 택하여야 하는데 어느 쪽을 택하더라도 그의 무법칙적 일원론의 기본 전제들과 상충하게 된다. 먼저 전자의 방책을 택하게 되면 심성적인 것의 영역을 제거하거나 그것을 물리적인 것의 영역으로 환원하겠다는 입장에 서게 되는데 이는 제거주의나 환원주의를 받아들이는 결과를 초래한다. 그럴 경우 제거주의나 환원주의를 극복할 수 있는 대안으로서 제시된 비환원적 유물론으로서의 무법칙적 일원론이 갖는 의의는 사라지고 말 것이다.

반면에 후자의 방책을 채택할 경우에는 비환원주의의 입장을 고수할 수 있기는 하지만 “설명적 배제(explanatory exclusion)”의 문제에 직면하게 된다. 설명적 배제의 현상은 동일한 피설명항에 대하여 두 개 이상의 서로 배타적인 설명들이 주어지되 그 설명들의 전제는 상호 무모순적일 경우에 발생한다. 예를 들어, 행위를 설명함에 있어서 기계론적인 설명과 목적론적인 설명이 함께 주어질 때 바로 설명적 배제의 현상이 나타나게 되는 것이다. 위의 예에서는  $e$ 의 발생을  $m$ 의 심적 속성  $M$ 에 의해서 설명하는 것과 물리적 속성  $N$ 에 의해서 설명하는 경우가 이에 해당한다. 그러면 이러한 상황에 처할 때 우리는 어떠한 선택을 하여야 하는가?

김재권에 따르면, 다음의 다섯 가지 선택지가 가능한데 데이빗슨의 무법칙적 일원론 안에서는 어느 것도 받아들일 수 없게 된다고 한다.<sup>5)</sup> 첫째로, 하나의 결과에 대해서 두 사건이 원인으로 주어질 때 그 두 사건이 하나의 원인을 구성하는 것으로 볼 수 있다. 왜 철수의 코가 깨졌느냐고 물었을 때 그가 너무 독한 감기약을 먹은 후 현기증이 나서 길에서 넘어졌기 때문이라고 설명할 경우가 한 예이다. 그런데 심물 관계의 경우에는 이 선택지를 택할 수 없다. 왜냐하면  $M$ 과  $P$ 가 하나의 원인을 구성한다고 할 경우 이는 데이빗슨이 전제한 “물리계의 폐쇄성”의 원칙을 어기는 셈이 되기 때문이다.

둘째로, 두 사건이 한 사건의 원인이 되되 각각은 결과가 되는 사건을 일으키기에 충분하다는 점에서 대등하다고 보는 것이다. 즉 결과가 되는 사건은 그 두 사건들에 의해서 “과잉결정(overdetermination)”되어 있다고 볼 수 있다. 하지

5) Jaegwon Kim, “Explanatory Exclusion and the Problem of Mental Causation,” in *Information, Semantics and Epistemology*, ed. Enrique Villanueva (Oxford: Basil Blackwell, 1990), pp.40-41.

만 이것도 데이빗슨의 이론 안에서는 채택될 수 없다. 이 선택지를 받아들일 경우, 심적 속성 M이 물리적 사건 e의 원인이 된다고 하여야 하는데 이것은 “하향적 인과 관계(downward causation)”를 용인하여 물리계의 폐쇄성의 원칙을 어기게 되기 때문이다.

셋째로, 두 사건이 한 사건의 발생에 이르는 인과 연쇄 상에 놓임으로써 원인의 역할을 한다고 볼 수 있다. 그래서 원인이 되는 두 사건들 중 뒤의 사건은 앞선 사건에 인과적으로 의존하게 된다. 하지만 이 선택지도 데이빗슨은 받아들일 수 없다. M과 P가 e의 원인이 되되 동시에 원인이 된다고 전제할 수 있기 때문이며, 심적 사건과 물리적 사건이 하나의 인과 연쇄를 이루고 상호 의존한다면 이 역시 물리계의 폐쇄성의 원칙을 어기는 셈이 되기 때문이다.

넷째로, 두 원인들을 등 길시하는 방법이 있다. 즉 심적인 원인과 물리적인 원인이 서로 별개로 존재하는 것이 아니라 동일한 원인이라고 보는 것이다. 이것은 심신 문제에 대한 고전적인 동일론이 취하는 해결책이기도 하다. 하지만 데이빗슨이 심성적인 것의 무법칙성을 고수하는 한, 심적인 속성과 물리적인 속성이 동일성이라는 매우 강한 관계에 있다고 주장할 수 없으므로 이 선택지 역시 데이빗슨에게는 애초부터 불가능하다. 심적 속성과 물리적 속성의 동일성은 그 속성들간의 엄밀하고 예외없는 상관 관계를 요구하기 때문이다.

다섯째로, 심성적인 것과 물리적인 것이 동일하지는 않지만 전자가 후자에 수반 내지는 의존한다고 볼 수 있다. 그런데 김재권의 해석에 따르면, 심적 속성이 물리적 속성에 수반한다고 볼 때 그 수반 개념을 어떻게 이해하느냐에 따라 환원 가능성을 받아들일게 된다고 한다. 강한 수반 개념이 한 예이다. 그럴 경우 데이빗슨은 이 선택지 역시 받아들이기 어렵게 된다. 강한 수반은 양자간의 법칙적 관계를 함축하는데 데이빗슨은 그러한 법칙성을 거부하였기 때문이다. 반면에 약한 수반 관계로 볼 경우 심적인 속성과 물리적인 속성 사이의 의존 관계는 너무 약한 것이어서 별로 의미가 없는 의존 관계가 되고 만다.<sup>6)</sup>

결국 데이빗슨은 자신의 기본적인 전제들을 일부 포기하지 않는 한 심성 인과의 문제를 해결하기가 어렵게 된다. 그러나 무법칙적 일원론의 입장도 포기하지 않고 심성 인과의 문제도 해결할 수 있는 길은 없을까? 다음 절에서 그러한 시도

6) 일반적으로 데이빗슨의 수반 개념은 약한 수반 개념으로 이해되고 있다. 위의 (5)에 나오는 수반 개념도 바로 약한 수반으로 이해할 수 있다. 수반 개념의 종류와 이들에 관한 논의는 이 논문의 범위를 넘기 때문에 생략하겠다.

를 검토해 보겠다.

## II

르포와 로어(이하에서는 ‘르-로’로 약칭하겠다)는 데이빗슨의 무법칙적 일원론이 결코 심적인 것의 인과력을 배제하지 않음을 보여주는 논증을 제시하였다.<sup>7)</sup> 그들은 무법칙적 일원론의 주요 테제들을 다음과 같이 해석한다. 첫째로, 심적인 사건과 물리적인 사건은 동일하지만 심적인 속성과 물리적인 속성은 구별된다. 둘째로, 심적인 속성과 물리적인 속성은 상호 작용한다. 셋째로, 물리계의 폐쇄성은 속성의 경우에 해당되지 않는다. 한 사건의 물리적 속성은 다른 사건의 심적 속성에 반사실적으로(counterfactually) 의존할 수 있기 때문이다.

르-로는 인과적 관련성(causal relevance) 개념을 두 가지로 구별하고 무법칙적 일원론을 비판하는 사람들이 이 두 개념을 혼동하고 있다고 지적한다. 제 1 개념은 엄밀 법칙이 전제되는 인과적 관련성의 경우다. c라는 사건이 F라는 속성을 가지고 사건 e는 속성 G를 가지되, F를 가지는 것들은 G를 가지는 것들의 원인이 된다는 것을 필함(entail)하는 엄밀 법칙이 있을 때, F와 G는 c가 e의 원인이 됨에 있어서 관련이 있다. 제 2 개념은 반사실적 의존성만이 전제되는 인과적 관련성의 경우를 나타낸다. 사건 c가 속성 F를 갖는 것이 e가 G를 갖는 것에 인과적으로 관련이 있기 위한 필요 충분 조건으로서 르-로는 다음 네 가지 조건들을 제시한다. “(i) c는 e의 원인이다, (ii) c는 F를 가지며 e는 G를 갖는다, (iii) c가 F를 갖지 않는다면, e는 G를 갖지 않을 것이다, 그리고 (iv) c가 F를 갖는 것과 e가 G를 갖는 것은 논리적으로 그리고 형이상학적으로 독립되어 있다.”<sup>8)</sup> 르-로는 무법칙적 일원론이 인과적 관련성의 제 1 개념의 의미에서는 심적 속성이 인과적으로 무관하다고 보지만 제 2 개념의 의미에서는 인과적으로 관련이 있음을 인정한다고 해석한다. 따라서 무법칙적 일원론 안에서도 심적인 것의 인과적 효력은 보장될 수 있다고 주장한다.

c를 ‘내가 물을 마셔야겠다고 생각하는 (심적) 사건’이라고 하고, e를 ‘내가 물이 든 컵을 잡기 위하여 손을 내뻗는 (물리적) 사건’이라고 하면, 전자는 인과

7) Ernest LePore and Barry Loewer, “Mind Matters,” *Journal of Philosophy* 84 (1987): 630-642.

8) Ibid., p. 635.

적 관련성의 제 2 개념의 의미에서 후자에 대하여 인과적 관련성을 갖는다고 볼 수 있다는 것이다. 이에 대해서 무법칙적 일원론을 비판하는 사람들은 다음과 같은 의문을 제기할 수 있다. “내가 물을 마셔야겠다고 생각함”이라는 속성을  $F$ 라고 하고, 내가 물을 마셔야겠다고 생각할 때 발생하는 두뇌의 특정한 신경생리학적 속성을  $F^*$ 라고 하자. 그러면 사건  $c$ 가  $F$ 를 갖지 않더라도  $F^*$ 를 가지면 여전히  $e$ 는 발생하게 되고, 데이빗슨의 이론 체계 안에서는  $Fc$ 와  $F^*c$ 가 동일하므로, 결국 심적 사건  $Fc$ 는 인과적으로 효력이 없게 되지 않는가? 그리고 이것이 바로 심성 인과의 문제와 관련해서 데이빗슨을 비판하는 사람들이 문제삼는 경우가 아닌가? 따라서 르-로가 제시한 인과적 관련성 개념에 대한 규정은 데이빗슨을 도와주지 못한다고 봐야 한다.”

이에 대한 르-로의 답변을 검토해 보자. 르-로는 앞의  $c$ 를 “신경 사건(neural event)”이라고 하고  $e$ 를 “행동 사건(behavioral event)”이라고 부른다.<sup>9)</sup> 즉 두 사건이 모두 물리적인 사건인 셈이다. 그런데 위의 물음은 사건  $c$ 가 심적 속성  $F$ 를 가지되 물리적 속성  $F^*$ 는 가지지 않더라도 사건  $e$ 가 발생하는 것이 무법칙적 일원론 안에서 허용될 수 있느냐는 물음과 같다. 르-로는 허용될 수 있다고 답한다. 그 이유는 이 때  $c$ 가  $F^*$ 를 갖지 않더라도 이와 “매우 유사한”<sup>10)</sup> 물리적 사건  $F^{**}$ 를 가질 것이기 때문이라고 한다.<sup>11)</sup>

이것이 과연 애초의 물음에 대한 충실한 답변이 될 수 있을까? 먼저  $F^*$ 와  $F^{**}$ 가 매우 유사하다고 할 때 어느 정도 유사한 상태를 가리키는지 분명치 않다. 양자가 어느 정도로 유사해야  $c$ 가 전자의 속성을 갖지 않더라도  $e$ 의 원인이 될 수 있는지에 대해서 르-로는 전혀 언급하지 않고 있다. 만일  $F$ 가 발생할 때 항상  $F^*$ 만이 발생하고 다른 신경생리학적 속성은 발생하지 않는다고 할 때(이러한 가능성을 배제할 수 있는 근거는 전혀 없다)에는 그들은 어떻게 답변할 것인가? 인과적 관련성 제 2 개념에 대한 반론을 봉쇄하기 위해서 제시한 르-로의 답변은 무법칙적 일원론에 대해서 심성 인과 문제가 제기될 때의 문제 상황을 교묘하게 회피할 뿐 적절한 해결책이 되지 못한다고 여겨진다.

이 점은 다음과 같은 물음을 묻게 되면 한층 분명해진다. 제 2 개념의 정의는  $F_c$ 가  $G_e$ 에 인과적으로 관련이 있음을 정의하기 위한 것인데 그렇다면  $F_c$ 가  $G_e$ 의

9) Ibid., p.639.

10) Ibid., p. 640.

11) Ibid., p. 639.

원인이 된다는 것은 어떻게 정의할 것인가? 인과적 관련성의 정의항 (i)에서 인과성 개념을 사용하는 것을 보면 르-로가 인과적 관련성과 인과성을 구분하고 있음을 알 수 있다. 그런데 그들은 (i)에 나오는 인과성 개념에 대해서는 전혀 규정을 하지 않고 있다. 그러면 c가 e의 원인이 된다고 하면서  $F_c$ 는  $G_e$ 의 원인이 된다고 하지 않고 인과적으로 관련이 있다고 말하여야 하는 이유는 무엇인가? 심성 인과의 문제가  $F_c$ 가  $G_e$ 의 원인이 될 수 있느냐는 문제이지 그들간의 (인과성과 구별되는 개념으로서의) 인과적 관련성 여부를 묻는 것은 아니라고 한다면 르-로는 어떻게 답변을 할 것인가? 이 의문들에 대한 답변을 르-로의 논문 안에서는 찾을 길이 없다.

그리고 인과적 관련성 제 2 개념에 대한 정의가 심물 관계는 둘째치고 물물 관계에서도 성립하기가 어렵다는 사실은 시겔과 소버가 제시한 반례를 통해서 분명해진다.<sup>12)</sup> c가 F인 상태를 석탄이 빨강계 된 상태로 보고, e가 G인 상태를 종이 가 타는 상태로 보자. 그러면 정의항의 (i), (ii), (iv)가 만족될 뿐만 아니라, 석탄이 빨강지 않았다면 종이 가 타지 않았을 것이기 때문에 (iii)도 만족된다. 그러나 석탄의 색깔이 종이의 연소에 대해서 인과적으로 관련이 있다고 보기는 어렵다. 이것은 정의항이 인과적 관련성의 필요 충분 조건이 되지 못함을 보여준다. 따라서 제 2 관련성 개념이 심성 인과의 문제를 해결하기 위한 만족스러운 장치가 될 수 없다는 결론이 나온다.

인과적 관련성의 두 개념들 사이에 어떠한 관계가 성립하는지에 대해서도 의문이 제기된다. 르-로는 제 1 개념이 제 2 개념을 필함하는지의 여부에 대해서 전혀 언급하지 않았는데 김재권은 그 두 개념이 필함 관계에 있더라도 문제가 생기고 그렇지 않더라도 문제가 생긴다고 비판한다.<sup>13)</sup> 먼저 제 1 개념이 제 2 개념을 필함한다고 보자. 앞에서 말한 바와 같이 사건 c가 심리적 속성 F와 물리적 속성  $F^*$ 를 함께 가지고 있고 사건 e가 (행동으로 표출된) 물리적 속성 G를 가지고 있다면  $F^*$ 와 G 사이에는 제 1 인과적 관련성 관계가 성립한다. 그리고 제 1 관련성이 제 2 관련성을 필함한다고 전제하였으므로 c가  $F^*$ 를 갖는 것은 e가 G를 갖는 것에 대하여 제 2 관련성도 갖는다고 할 수 있다. 이 경우 데이빗슨의 이론대로

12) Gabriel Segal and Elliott Sober, "The Causal Efficacy of Content," *Philosophical Studies* 63 (1991): 5.

13) Jaegwon Kim, "Explanatory Exclusion and the Problem of Mental Causation," pp.44-45.



하자면,  $e$ 가  $G$ 를 가지는 것은  $c$ 가 심리적 속성  $F$ 를 갖는다는 것에 의해서 합리화되며,  $r$ -로의 정의에 따라  $c$ 가  $F$ 를 갖는 것이  $e$ 가  $G$ 를 갖는 것과 인과적으로 관련(제 2 개념)되어 있다고 볼 수 있다. 그러면 'c가  $F^*$ 를 갖는 것'과 'c가  $F$ 를 갖는 것'이 모두  $e$ 가  $G$ 를 갖는 사태를 야기시키게 되는데 이 두 원인의 관계는 어떠한 관계인가라는 물음에 대해서  $r$ -로는 아무 답변도 제시하지 못한다. 이것은  $r$ -로의 이론이 인과적 설명 배제의 문제에 대해서 속수무책임을 보여준다.

이제 제 1 관련성이 제 2 관련성을 필합하지 않는 경우를 생각해 보자.  $r$ -로의 설명에 따르면,  $c$ 의 물리적 속성  $F^*$ 는 엄밀 법칙에 의거해서  $e$ 의 물리적 속성  $G$ 와 인과적으로 관련(제 1 개념)되어 있고, 동시에 심적 속성  $F$ 는 세테리스 파리부스 법칙 (*ceteris paribus law*)에 의거해서  $G$ 와 인과적으로 관련(제 2 개념)되어 있다고 할 수 있다.<sup>14)</sup> 그러면  $F$ 와  $F^*$ 가 모두  $G$ 에 대해서 인과적으로 관련되어 있을 경우에 다음과 같은 물음이 제기될 수 있다. 'c가 발생하고  $F^*$ 를 가졌고  $e$ 가 발생하고  $G$ 를 가졌다면, c가  $F$ 를 갖지 않았더라도  $e$ 가  $G$ 를 가졌을 것인가?' 이에 대해서 아니라고 답할 경우, 이것은  $e$ 가  $G$ 를 갖는다는 물리적인 상태를 산출하기 위하여 물리적인 속성들에 덧붙여서 심적 속성도 가져야 함을 뜻한다고 볼 수 있는데 이렇게 되면 물리계의 폐쇄성의 원칙을 어기는 결과가 나온다.

반면에 앞의 질문에 대해서 그렇다라고 답하면,  $c$ 가  $F$ 를 갖는 것이  $e$ 가  $G$ 를 갖는 것에 대하여 인과적으로 관련되어 있다는 것이 무슨 뜻이냐는 물음이 제기된다.  $c$ 의 물리적 속성이 특정한 행동을 야기시켰다면,  $c$ 의 심리적 속성은 그 행동의 발생에 대해서 어떠한 인과적인 효력을 갖는지 의심스러워지는 것이다. 이것은 곧  $r$ -로의 이론 체계 안에서는 심리적 속성의 인과적 효력이 인정받을 여지가 없음을 보여주고, 따라서  $r$ -로가 심성 인과의 문제에 대해서 무법칙적 일원론이 당면하는 곤경을 해결해 주지 못함이 입증된다.

14) 세테리스 파리부스 법칙이란 '다른 모든 조건이 같다면'이라든지 '모든 조건이 정상적이라면'과 같은 절이 단서로 붙은 법칙을 말한다. 쉬퍼는 '다른 모든 조건이 같다면 사건  $M$ 이 사건  $B$ 의 원인이 된다'가 의미하는 바는 다음과 같다고 한다. '우리가  $B$ 의 발생을 설명하되  $M$ 에 의한 것으로 설명할 수 있지만  $M$ 의 발생만으로는  $B$ 가 발생하기에 충분치 않다.' 세테리스 파리부스 법칙에 관한 구체적인 논의를 위해서는 다음 논문들을 참조하라. Stephen Schiffer, "Ceteris Paribus Laws," *Mind* 100 (1991): 1-17; Jerry A. Fodor, "You Can Fool Some of The People All of The Time, Everything Else Being Equal: Hedged Laws and Psychological Explanations," *Mind* 100 (1991): 19-34.

이제 다른 사람이 아닌 데이빗슨 본인의 해명을 살펴 보자.<sup>15)</sup> 그는 무법칙적 일원론의 기본적인 테제들이 상충되고 심성적인 것의 인과적 효력을 부정하게 된다고 비판하는 김재권을 비롯한 철학자들이 그 이론을 오해하고 있다고 지적함으로써 자신의 입장을 옹호한다. 심성 인과의 문제에 대한 해명에 앞서서 우선 데이빗슨은 자기가 말하는 인과성이 비추상적인 (즉 구체적인) 개체로서의 사건들간의 관계이고, 그 사건들이 어떻게 기술되느냐는 것은 사건들간의 인과 관계에 아무 영향을 미치지 못함을 주시시킨다.<sup>16)</sup> 예를 들어, 사건 c가 e의 원인이 된다고 할 때 c가 심적 사건으로서 기술되느냐 물리적 사건으로서 기술되느냐는 것은 인과 관계에 하등 영향을 미치지 않는다는 것이다. 데이빗슨은 이 점을 다음과 같이 예시하고 있다.

마젤란이 암초가 앞에 있음을 알게 되었다고 하자. 이것은 그의 배의 항로를 변경하는 원인이 되는 사건이다. 이 때 마젤란이 암초를 알아챈 사건은 심적 사건이며, 인과적인 효력을 가진다. 그런데 이 사건은 마젤란의 신체 안에서 일어나는 사건으로서 물리학의 용어로 서술될 수 있는 물리적 사건이기도 하다.<sup>17)</sup>

따라서 심적 사건으로서 기술되는 사건이 인과력을 갖느냐의 여부를 따지는 것이 심성 인과의 문제라면, 무법칙적 일원론은 심성 인과의 문제에 말려들 이유가 없게 된다. 그렇다고 해서 이것이 심적 속성을 아무 쓸모없는 것으로 배제하는 결과를 낳지는 않는다고 데이빗슨은 주장한다. 동일한 물리적 속성을 가지고 있는 두 사건들 중 하나는 심적 속성을 가지고 있고, 다른 하나는 그렇지 않다면 그 두 사건은 동일한 사건들이 될 수 없으며, 자신의 이론이 이 점을 부정하지는 않는다고 한다. 다만 인과 관계는 사건들간의 관계이므로 “속성이 그 관계에 영향을 미칠 수 없다”<sup>18)</sup>는 점을 말할 뿐이라고 해명한다.

그러면 이것은 한 사건의 심적 속성들이 그 사건의 인과 관계에 대해서 아무 영향을 못미치게 되어 심적 속성의 인과적 효력이 부정됨을 의미하는가? 다시 말

15) Donald Davidson, “Thinking Causes,” in *Mental Causation*, ed. John Heil and Alfred Mele (Oxford: Oxford Univ. Press, 1993), pp.3-17.

16) *Ibid.*, p. 6.

17) *Ibid.*, p. 12.

18) *Ibid.*, p. 8.

해서 심성 인과의 문제를 무법칙적 일원론이 해결할 수 없음을 데이빗슨 자신도 인정하는 것인가? 데이빗슨은 심적 속성이 인과적으로 전혀 영향을 미치지 못한다는 생각이 무법칙적 일원론과 상충되지는 않지만, 무법칙적 일원론이 그 생각을 함의(imply)한다는 결론은 나오지 않으며, 자기를 비판하는 어느 누구도 그것들간의 함의 관계를 입증하지는 못하였다고 주장한다.<sup>19)</sup> 그리고 데이빗슨은 심물 수반 관계가 성립된다면, 심적 속성들은 한 사건의 인과 관계에 대해서 영향을 미친다고 말한다. 왜냐하면 수반 관계에 의해서 심적 속성들은 물리적 속성들과 관계가 있고, 물리적 속성들은 인과 관계에 영향을 미칠 것이기 때문이다.<sup>20)</sup> 따라서 무법칙적 일원론이 심물 수반 관계를 전제하는 한, 심적 속성의 인과적 효력은 보장될 수 있고 심성 인과의 문제는 속성의 차원에서도 문제가 되지 않는다는 결론이 나온다.

데이빗슨이 거론한 또 다른 쟁점은 심물 법칙에 관한 것이다. 그동안 데이빗슨은 다음과 같은 논조의 비판을 받아왔다. ‘심적 속성과 물리적 속성 사이에 상관관계나 의존 관계가 있음을 인정한다면, 심물 법칙이 있음을 인정하여야 한다. 데이빗슨은 전자를 인정하면서도 후자는 부정하는데 이것은 그가 일관성을 결여하고 있음을 보여준다. 반면에 그가 심물 법칙을 인정할 경우 그는 더 이상 환원주의를 배격할 수 없게 된다.’ 이에 대해서 데이빗슨은 자기가 부정한 것은 심물간의 엄밀 법칙일 뿐 비(非)엄밀 법칙은 성립할 수 있다고 답변한다. 그가 말하는 ‘엄밀한 심물 법칙’이란, “물리학의 용어로 환원될 수 없고 또 그 법칙 안에서 제거될 수 없는 심리 용어들을 담고 있으며, 세테리스 파리부스 절을 가지고 있지 않으면서 물리학의 법칙과 같은 폐쇄 체계에 속한 법칙”<sup>21)</sup>을 말한다. 이러한 유의 법칙이 심물 간에 성립하지 못한다는 입장과, 심적 속성과 물리적 속성 사이에 상관관계나 의존, 또는 수반 관계가 성립한다는 입장은 양립가능하다. 그리고 비엄밀법칙을 용인한다고 해서 반드시 환원주의를 받아들여야 한다고 볼 수는 없다. 따라서 위의 비판은 수긍할 수 없다고 데이빗슨은 반박한다.

이상과 같은 데이빗슨의 자기 변호가 받아들여질 수 있는지 살펴 보자. 먼저 심적 속성이 수반 관계를 맺고 있는 물리적 속성에 의해서 그것의 인과적 효력을 인정받을 수 있다고 데이빗슨은 주장하였는데 심적 속성의 인과적 효력을 보장할 만한

19) Ibid., p. 13.

20) Ibid., p. 14.

21) Ibid., p. 11.

강한 수반 관계를 전혀 구체적으로 제시하지 못하고 있다는 비판이 제기될 수 있다.<sup>22)</sup> 김재권은 데이빗슨의 약한 심물 수반 개념이 보장할 수 있는 것이라고는 기껏해야 심적 속성의 “인과적 관련성(causal relevance)”일 뿐 “인과적 효력성(causal efficacy)”은 보장하지 못한다고 지적한다.<sup>23)</sup> 심적 속성이 사건들간의 인과 관계에 관련이 있다고 해서 반드시 결과가 되는 사건에 대해서 인과적인 영향을 미쳤다고는 볼 수 없는데 데이빗슨은 심적 속성의 인과적 효력이 어떻게 성립하는지에 대해서 구체적으로 규명하지 않았다는 것이다. 이것은 심적 속성이 물리적 속성에 약하게 수반하고 이 물리적 속성에 의해서 사건들이 인과적으로 연결된다고 해서 바로 그 심적 속성에 의해서 문제의 사건들이 인과적으로 연결된다고 생각하여야 할 이유는 없다는 맥로플린의 지적과도 일맥상통한다.<sup>24)</sup> 따라서 데이빗슨의 무법칙적 일원론은 여전히 심성 인과의 문제를 해결하지 못한 채로 남게 된다.

데이빗슨이 허용한 비엄밀법칙에 대해서는 우선 그가 말하는 법칙의 비엄밀성을 어떻게 규정할 것이냐는 물음이 제기될 수 있다. 데이빗슨은 법칙의 비엄밀성을 구체적으로 규정하지는 않고 다만 포더의 ‘세테리스 파리부스’ 절이 붙은 비엄밀 법칙을 예로 들었지만 김재권은 포더가 말하는 비엄밀 법칙이 과연 심물간의 무법칙성을 보장해 줄 만큼 느슨한 것인지에 대해서 의구심을 제기한다. 그러나 무엇보다도 비엄밀성이 어떻게 규정되건 간에 비엄밀 법칙도 법칙인 한에 있어서 법칙적인 힘을 지닌다고 보아야 하는데 심물간의 무법칙성을 전제한 데이빗슨의 이론 체계 안에서 비엄밀 법칙이 용인될 여지가 있을지 의심스럽다고 김재권은 지적한다.<sup>25)</sup> 데이빗슨이 심적인 영역과 물리적인 영역을 구분할 때 각 영역은 서로 구별되는 특수한 원리에 의해서 규제를 받는다고 하면서 심물 법칙을 배제하였는데 최소한이나마 법칙적 힘을 가지고 있는 비엄밀 법칙이 허용될 경우 두 영역의

22) 데이빗슨이 “Thinking Causes”에서 제시한 수반 개념은 다음과 같다. “술어들의 집합 S에 의해서 구별될 수 없는 것(entity)들을 술어 p가 구별할 수 없을 때 바로 그 때에만 p는 S에 수반한다. (p.4)” (이 정의에 나오는 술어라는 용어를 개념이나 속성이라는 용어로 대체해도 무방하다.)

23) Jaegwon Kim, “Can Supervenience and ‘Non-Strict Laws’ Save Anomalous Monism?” in *Mental Causation*, pp.23-24.

24) Brian P. McLaughlin, “On Davidson’s Response to the Charge of Epiphenomenalism,” in *Mental Causation*, p. 37.

25) Jaegwon Kim, “Can Supervenience and ‘Non-Strict Laws’ Save Anomalous Monism?”, pp.24-25.

구분이 보존될 수 있는지, 있다면 어떻게 그것이 가능한지를 설명하여야만 함에도 불구하고 데이빗슨은 구체적인 설명을 하지 않고 있다. 결국 이상과 같은 이론상의 문제점이 해결되지 않는 한, 무법칙적 일원론은 심물간의 인과 관계에 있어서 심적인 것의 역할을 보장해 주지 못한다는 비판을 벗어나기가 어려우리라고 본다.

### III

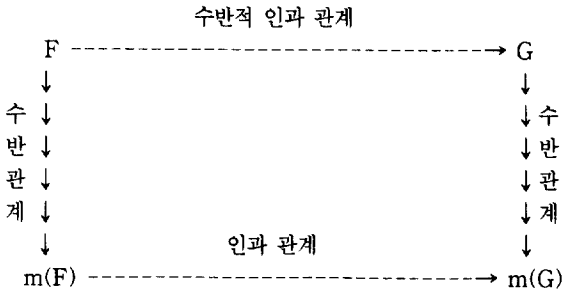
심성 인과에 대한 무법칙적 일원론의 관점을 포기할 경우 남는 대안은 무엇이 있을까? 김재권은 무법칙적 일원론이 심성 인과의 문제를 해결하지 못한 이유는 비환원주의의 입장에 서있기 때문이라고 보고 수반적 인과 관계에 의한 환원주의를 채택할 것을 제안한다.

김재권은 원자나 입자와 같은, 세계의 가장 기본적인 구성 요소를 고정시키고 그것들의 속성과 그것들을 지배하는 법칙들을 고정시키면, 세계의 전체적인 성격이 고정될 것이라고 생각하면서 인과 관계를 “거시적 인과 관계”와 “미시적 인과 관계”로 구분할 것을 제안한다.<sup>26)</sup> 거시적 인과 관계는 거시적 사건들간의 인과 관계로서 거시적 사건은 하나의 사건이 한 시점에서 거시적 속성을 예화하는 것이라고 한다. 거시적인 것과 미시적인 것의 구분은 상대적이라고 할 수 있는데 우리 주위의 물체와 속성을 거시적인 것으로, 그 물체를 구성하는 입자들을 미시적인 것으로 볼 수 있다. 대상의 속성과 행태를 이해할 때 그것의 미시적인 구성 요소들을 특징짓는 속성과 관계에 의해서 이해하는 이론 물리학의 전략을 받아들여서 철학에서도 거시적 속성들이 미시적 속성들에 수반하거나 그것들에 의해서 결정된다고 보는 미시결정론을 채택할 것을 김재권은 제안한다. 이러한 미시결정론을 받아들일 경우 거시적 사건들간의 인과 관계는 그것들보다 더 기본적인 인과 과정으로 환원시킬 수 있다.

김재권이 제시한 두 사건들간의 거시적 인과 관계를 환원시키는 일반적인 도식을 살펴보자. 이를 위해서 F와 G를 거시적 속성이라고 하고 그것들의 미시적 속성을 각각  $m(F)$ ,  $m(G)$ 라고 하자. 그리고 ‘x가 F를 가짐’이라는 사건을  $F_x$ , ‘y가 G를 가짐’이라는 사건을  $G_y$ , ‘x가  $m(F)$ 를 가짐’이라는 사건을  $m(F)_x$ ,

26) 수반적 인과에 근거한 김재권의 환원주의적인 전략은 여러 논문에서 제시되고 있는데 대표적인 논문으로서 “Epiphenomenal and Supervenient Causation,” *Midwest Studies in Philosophy* 9 (1984): 257-270을 들 수 있다.

‘y가 m(G)를 가짐’이라는 사건을 m(G)y라고 하자. 그러면 Fx가 Gy의 원인이 되기 위한 필요 충분 조건은 다음과 같다. “Fx는 m(F)x에 수반하고, Gy는 m(G)y에 수반하며, m(F)x와 m(G)y 사이에는 적합한 인과 연쇄가 성립한다.”<sup>27)</sup> 여기에서 사용된 수반 개념은 다음과 같이 정의되고 있다. “속성 집합 A가 다른 속성 집합 B에 수반함은 다음과 같이 설명된다: ‘A의 원소인 임의의 속성 F에 대해서 임의의 대상 x가 F를 가진다면, B에 속한 임의의 속성 G를 x가 가지며, G를 가지는 어떠한 대상도 필연적으로 F를 가진다’는 것은 필연적이다.”<sup>28)</sup> 김재권의 이러한 생각을 다음의 도식으로 나타낼 수 있다.<sup>29)</sup>



이 이론이 환원주의인 이유는 F와 G의 거시적 인과 관계를 m(F)와 m(G)의 미시적 인과 관계로 환원시켜서 설명하기 때문이다.

김재권은 위의 수반적 인과 관계에 의거해서 심성 인과를 설명할 수 있다고 주장한다. 예를 들어, 내가 뜨거운 냄비에 손을 댔다가 고통을 느끼고서 얼른 손을 떼는 경우를 생각해 보자. 이때 고통을 느끼는 심적 사건(거시적 사건)은 특정한 신경 생리학적인 두뇌 상태(미시적 사건)에 수반하고, 이 미시 물리적인 사건이 내가 손을 떼는 물리적 사건(거시적 사건)이 수반하고 있는 다른 미시 물리적인 사건의 원인이 됨으로써 심적 사건과 물리적 사건간의 인과 관계가 성립하게 된다고 한다. 따라서 심적 속성들의 인과적인 효력은 그것들이 수반하고 있는 물리적인

27) Ibid., p. 262.

28) Ibid.

29) 이 도표는 Gabriel Segal and Elliott Sober, “The Causal Efficacy of Content,” p.7 에서 제시된 것이다.

속성들로부터 나오고, 심물간의 인과 관계는 보다 더 기본적인 물리적 수준에서 발생하는 물물간의 인과 과정으로 환원되는 것이다.

수반 인과론적인 환원주의가 직면하는 가장 위협적인 난제는 심성 내용에 관한 외재주의(Externalism) 이론에 의해서 제기된다.<sup>30)</sup> 이 이론에 따르면, 퍼트남이 제시한 쌍둥이 지구에 의한 사고실험을 통해서 두 사람이 신경 생리학적인 (물리적인) 속성들에 있어서는 동일하지만 상이한 심성 내용을 가질 수 있음이 입증되었다고 한다. 다시 말해서 두 사건이 신경 생리학적인 속성에 관해서는 동일하지만 심적 속성에 관해서는 다를 수 있다는 것이다. 이것은 김재권의 심물수반론 테제와 근본적으로 상충된다. 따라서 심성 내용에 관한 외재주의의 주장을 반박하지 못할 경우 수반 인과론에 근거한 김재권의 환원주의는 성립 근거를 잃고 말 위험에 빠진다.

김재권은 심성 내용에 관한 내재주의와 외재주의의 대립에 대해서 아직 분명한 입장을 표명하지 않고 있으며, 외재주의에 의한 심물수반론의 반박을 어떻게 극복할 것인지에 대해서도 함구하고 있다. 그리고 환원주의적인 전략을 지극히 간명하게 서술하였을 뿐 비환원주의를 대체할 만큼 정교한 이론적인 체계를 제시하지 않았기 때문에 과연 그의 환원주의가 무법칙적 일원론의 대안이 될 수 있을지는 현 단계에서 확인하기가 어렵다. 하지만 데이빗슨이 심성 인과의 문제에 대한 명쾌한 해결책을 제시하지 못하는 한, 김재권의 환원주의는 심리 철학의 새로운 장을 여는 계기가 될 수 있으리라고 보여진다.

30) 외재주의의 대표자로서 버지를 들 수 있는데 그의 이론을 구체적으로 살펴 보기 위해서는 다음 논문들을 참조하라. Tyler Burge, "Individualism and the Mental," in *Midwest Studies in Philosophy* 4, ed. P. French, T. Euhling, and H. Wettstein (Minneapolis: Univ. of Minnesota Press, 1979), pp.73-122; "The Content of Propositional Attitudes," *Nous* 14 (1980): 51-58; "Other Bodies," in *Thought and Object*, ed. A. Woodfield (Oxford: Oxford Univ. Press, 1981), pp.97-120; "Two Thought Experiments Reviewed," *Notre Dame Journal of Formal Logic* 23 (1982): 284-293; "Individualism and Psychology," *Philosophical Review* 95 (1986): 3-45; "Cartesian Error and the Objectivity of Perception," in *Contents of Thought*, ed. R. H. Grimm and D. D. Merrill (Tucson: Univ. of Arizona Press, 1988), pp.62-76.