

선언적 설명에 대해서 어떻게 생각할 것인가?

이 재 호[†]

이 논문에서 필자는 우선, 우리가 선언적 인과를 받아들이지 않는다고 하더라도 선언적 설명의 존재를 받아들이 필요가 있다는 것을 논증한다. 다음으로 필자는 기존의 주요 설명이론들, 즉 D-N 모델, 인과적 설명 이론, 설명에 대한 통일 이론, 반사실적 의존성 설명 이론 등은 모두 선언적 설명에 대한 우리의 직관을 제대로 설명할 수 없다는 것을 논증한다. 마지막으로 필자는 필자의 설명이론은 기존의 이론들과는 달리 선언적 설명에 대한 우리의 직관에 대해 원리지원된 방식으로 체계적 설명을 제공해줄 수 있다는 것을 논증한다.

주요어 : 설명, 선언, 평형설명, 조작, 레시피

[†] 중앙대학교 철학과 교수 (jaeho.jaeho@gmail.com)

1. 선언과 선언적 설명

선언(disjunction)은 형이상학, 과학철학, 인식론 등 철학의 주요 분과에서 다양한 이슈를 만들어 내는 주제로 알려져 있다. 특히 어떤 것이 “선언적”이라는 수식을 받을 때 그것은 종종 의심스러운 것으로 여겨진다. 예를 들어, “선언적 속성”이라는 명칭은 “진정한 속성은 아닌” 또는 적어도 “자연적 속성은 아닌”과 쉽게 연상된다. 이것은 충분히 이해 가능하다. 우리는 일반적으로 상이한 대상이 속성을 공유하는 것을 통해서 서로 객관적으로 유사해질 수 있다고 생각한다. 그러나 선언적 속성은 이런 역할을 잘 수행할 수 없는 것으로 보인다. 예를 들어 우리는 고양이와 개가 ‘포유류임’이라는 속성을 공유하는 것을 통해서 유사해진다고 생각하지만 이들이 ‘고양이거나 개임’을 공유하는 것을 통해서 유사해진다고 생각하지는 않는다. 마찬가지로 잘 알려진 굿맨의 초랑(grue) 역설에서 ‘초랑’이라는 속성이 투사 불가능한 속성이며 이것이 과학적으로 존중받을 수 없는 속성이라고 간주될 때 우리는 흔히 ‘초랑’이 갖는 선언적 성격에 주목한다.

선언이 갖는 이런 부정적인 인상은 인과에 관한 논의에서도 나타난다. 우리는 흔히 인과 관계의 관계항이 사건이라고 생각한다.¹⁾ 그리고 어떤 것이 진정한 사건이기 위한 조건으로 우리는 종종 그것이 선언적이지 않을 것을 요구한다. 예를 들어, D. 루이스가 강조하듯이 우리는 사건이 개별적인 것이며 따라서 특정한 시공간적 장소를 가져야 한다고 생각한다. 사건은 특정한 장소에서 발생하는 것이다. 그러나 선언적 사건은 그 선언적 성격 때문에 특정한 시공간적 장소를 갖는다고 말하기는 어렵다(Lewis 1986b). 철수가 집에서 밥을 먹거나 학교에서 공부하는 “사건”은 집에서 일어난 “사건”이라고 보기도 어렵고 학교에서 일어난 “사건”이라고 보기도 어렵다. 만약 우리가 선언적 사건을 쉽게 받아들일 수 없다면 마찬가지로

1) 소위 행위자 인과(agent causation)을 받아들이는 사람들은 인과의 관계항이 사건이 아니라 실체일 수도 있다고 생각한다(Clark 2003; O'Connor 1995). 그러나 행위자 인과를 받아들이는 사람이 오늘날 다수라고 보기는 어렵다.

우리는 선언적 인과 역시 쉽게 받아들일 수 없다.

인과와 설명 사이의 관계가 정확히 어떤 것인지에 대해서는 다양한 의견이 있을 수 있다. 어떤 “인과적 설명 이론가”들은 설명적 관계는 사실상 인과 관계로 분석될 수 있다고 생각한다(Salmon 1984). 설명에 관한 통일 이론을 옹호하는 P. 키처의 경우 반대로 인과 개념은 설명 개념으로부터 파생되어 나온 개념이라고 생각한다(Kitcher 1989). 그러나 오늘날 인과와 설명 개념이 매우 밀접하게 연결되어 있다는 것 자체를 부정하는 설명 이론가는 없다. 그렇다면 우리는 자연스럽게 다음과 같이 물을 수 있다: 선언적 인과에 대한 우리의 부정적 생각은 선언적 설명에 대한 부정적 생각으로 이어져야 하는가?

이 논문에서 필자는 두 가지 것을 보이고자 하는데, 그 첫 번째 것은 이 자연스런 질문에 대해 ‘아니오’라고 대답하는 것이다. 필자는 우리가 선언적 인과를 받아들이지 않는다고 하더라도 선언적 설명을 받아들일 필요가 있다고 주장할 것이다. 필자가 이 논문에서 보이고자 하는 두 번째 것은 우리가 선언적 설명에 대해서 갖는 직관을 가장 만족스럽게 설명하는 이론이 필자가 다른 곳에서 제안한 바 있는 새로운 조작주의적 설명이론이라는 것이다. 과학적 설명에 대한 철학적 논의를 사실상 출발시킨 햄펠의 D-N 모델, 새먼, 루이스 등에 의해 옹호된 인과적 설명 이론, 키처의 통일 접근, 선우환 교수 등이 옹호하는 반사실적 의존성 접근 등 기존의 주요 설명 이론들은, 앞으로 논증되겠지만, 모두 선언적 설명에 관한 우리의 직관을 제대로 수용하지 못한다. 만약 우리에게 거부할 수 없는 선언적 설명에 대한 직관이 존재하고 그 직관을 필자의 새로운 조작주의적 설명 이론이 가장 잘 설명할 수 있다면 자연스럽게 우리는 새로운 조작주의적 설명이론을 받아들일 좋은 동기를 갖게 될 것이다.

2. 선언적 설명의 존재와 종류

먼저 왜 선언적 설명이 존재한다는 직관이 무시될 수 없는지에 대해

서 생각해 보자. 우선, 인과에 대한 논의에서 많이 논의되는 다음의 선취(preemption) 사례를 생각해 보자.

(돌던지기 사례) 철수와 영수는 꽃병을 향해 동시에 돌을 던졌다. 이들의 돌은 모두 꽃병을 향해 정확하게 날아갔지만 보다 강한 어깨를 갖고 있었던 철수가 던진 돌이 꽃병에 먼저 도달해 꽃병을 깨고, 영수가 던진 돌은 허공을 갈랐다.

이 사례에서 우리의 **인과**에 대한 직관은 명확하다. 철수가 돌을 던지는 사건은 꽃병이 깨지는 사건의 원인이 되지만 영수가 돌을 던지는 사건은 원인이 되지 않는다. 그러나 이 사례에서 우리의 **설명적** 직관은 그렇게 명확하지 않다. 물론 우리는 한편으로 그 꽃병이 깨지는 사건이 철수가 돌을 던지는 사건 때문에 발생했다는 무시할 수 없는 직관을 갖고 있다. 그러나 동시에 우리는 철수가 돌을 던지지 않았더라도 그 꽃병은 어차피 (영수가 던진 돌에 의해서) 깨졌을 것이기 때문에 그 꽃병이 철수가 돌을 던졌기 때문에 깨진 것은 아니라는 직관 역시 무시할 수 없을 정도로 갖고 있다. 이 후자의 직관을 따라가 보자. 만약 그 꽃병이 깨지는 사건이 발생한 것이 철수가 돌을 던진 사건 때문에 발생한 것이 아니라면 그 꽃병은 왜 깨진 것인가? 아마도 이런 직관을 강하게 갖는 사람들은 다음의 선언적 설명이 이 상황에서 가장 적절한 설명이라고 생각할 것이다.²⁾

선언적 설명1(표준적 선취, 과잉결정 상황): 철수 또는 영수가 돌을 던졌기 때문에 그 꽃병이 깨졌다.

만약 철수 또는 영수가 돌을 던지는 것이 성립하지 않았다면, 즉 철수도 돌을 던지지 않았고 영수도 돌을 던지지 않았다면 그 꽃병은 깨지지 않았을 것이며 이런 반사실적 의존성의 성립은 선언적 설명1을 직관적으로 그럴듯한 설명으로 보이게 만든다.

2) 이 길이 정확히 선우환이 가는 길이다(선우환 2020, p. 141).

선취 상황에서의 우리의 설명적 직관은 과잉결정 상황에서의 우리의 설명적 직관에도 확장된다. 이번에는 철수와 영수가 던진 돌이 정확히 같은 시간에 꽃병에 충돌했다고 가정해 보자. 이 경우 우리의 인과적 직관은 철수가 돌을 던지는 사건과 영수가 돌을 던지는 사건 모두가 꽃병이 깨지는 사건의 원인이라고 말한다.³⁾ 그러나 선취 사례에서와 마찬가지로 과잉결정 사례에서도 우리의 설명적 직관은 아주 명확하지는 않다. 우리는 이 경우에도 ‘선언적 설명1’이 적절한 설명이라는 무시할 수 없는 직관을 가지고 있다.

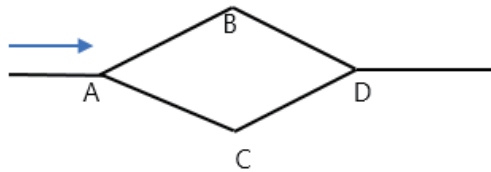
선취나 과잉결정 상황에서 우리가 선언적 설명에 대한 무시할 수 없는 직관을 갖는다는 것은 진정한 선언적 설명이 존재한다는 것을 보여주는가? 필자는 그렇다고 생각한다. 그러나 필자는 필자의 생각을 독자들에게 강요할 생각은 없다. 설명적 관계와 인과적 관계 사이의 거리가 매우 가깝다고 생각하는 사람들 가운데는 우리의 설명적 직관은 철저하게 우리의 인과적 직관을 따라가야 한다고 생각하는 사람이 있을 수 있다. 이런 사람들은 위의 선취 사례에서 철수가 돌을 던지는 사건 때문에 그 꽃병이 깨지는 사건이 발생했다는 무시할 수 없는 설명적 직관을 최대한 존중하고 철수 또는 영수가 돌을 던졌기 때문에 그 꽃병이 깨졌다는 무시할 수 없는 설명적 직관을 최대한 억제 또는 무시하는 것을 통해 진정한 선언적 설명이 존재한다는 것을 부정하는 길을 갈 수도 있다.

그러나 이런 식으로 선언적 설명에 대한 직관에 저항하는 전략이 잘 작동하지 않는 선언적 설명의 사례가 존재한다. 이를 확인하기 위해 우선 다음의 스위치 구조를 갖는 사례를 생각해 보도록 하자.

(철길 사례) 기술자 한 명이 철도 교차로(A지점) 스위치에 서 있다. 이 스위치는 왼쪽으로 가 있거나 오른쪽으로 가 있어야 하는데 현재 오른쪽에 있다. 만약 이 기술자가 그 스위치를 왼쪽으로 밀면 다

3) D. 루이스는 이런 인과적 직관을 받아들이지 않는다(Lewis 1986c). 필자가 생각하기에 루이스의 과잉 결정 사례에 대한 직관은 일반적으로 받아들여지는 직관이라고 보기는 힘들며, 이 논문에서 루이스적 직관은 무시될 것이다.

가오고 있는 열차는 위쪽 철도로 진행하게 되며 B지점을 거쳐 시간 t 에 D에 도달하게 된다. 반면 이 기술자가 스위치를 오른쪽으로 놔 둔다면 다가오고 있는 열차는 아래쪽 철도로 진행하게 되고, C지점을 거쳐 t 에 D에 도달하게 된다. 실제로 그 기술자는 스위치를 왼쪽으로 밀었고 열차는 B지점을 지나 D에 도착했다(Hall 2004, p. 187; 이재호 2023, p. 11).



이 상황에서 우리는 기술자가 스위치를 왼쪽으로 미는 사건과 열차가 D에 도착하는 사건 사이에 인과적 연결이 존재한다는 무시할 수 없는 직관을 갖는다. 열차가 B를 지나 D에 도착하는 과정은 분명 인과적인 과정으로 보이며 기술자가 스위치를 왼쪽으로 미는 사건은 이 인과적 과정을 발생시킨 것으로 보인다. 그러나 이 경우 전형적인 선취 사례와 마찬가지로 원인 사건과 결과 사건 사이에 반사실적 의존성은 존재하지 않는다. 스위치를 왼쪽으로 밀지 않았어도 열차는 C를 지나 D에 도달했을 것이기 때문이다. 스위치 사례는 실제의 원인이 대안적 백업 원인이 되는 것을 막는다는 점에서 선취 사례와 유사한 특징을 보인다. 그러나 스위치 사례에서는 실제의 원인 사건을 대체하는 사건이, 즉 실제의 원인이 발생하지 않는 최근접 세계에서 실제의 원인을 대체하는 사건이 자체로 백업 원인이 되며 따라서 실제의 원인의 발생이 논리적으로 백업 원인의 발생을 차단한다는 점에서 전형적인 선취 사례와는 구분된다. 스위치 사례가 전형적인 선취 사례와 구별되는 또 다른 특징은 스위치 사례의 경우 실제 원인이 설명적 힘을 거의 갖지 않는다는 것이다. 위의 철도 사례에서 우리는 기술자가 스위치를 왼쪽으로 밀었기 때문에 열차가 D에 도달했다는 직관을 거의 갖지 않는다. 철도가 연결되어 있는 구조가 주어질 경우 스위치가 어떤 위치에 있건 열차는 D에 도달하게 되어 있으며 따라서 스위치의 위치는 열차의 D

지점에서의 도착에 대해서 설명적으로 무관하다. 현재의 스위치 사례가 앞서 언급된 선취 사례와 구분되는 또 하나의 중요한 특징은 이 스위치 사례의 경우 선언지뿐만 아니라 단순한 선언 역시 설명적 힘을 갖고 있지 않은 것으로 보인다는 것이다. 사례 속의 스위치가 왼쪽 또는 오른쪽에 있었기 때문에 열차는 D 지점에 도착했다고 말하는 것은 자연스럽게 들리지 않는다. 주어진 상황 속에서 스위치는 왼쪽 또는 오른쪽에 있을 수밖에 없으며 이런 사실상 필연적인 사실이 피설명항 사건을 설명할 수 있을 것으로 보이지는 않는다. 이 사례에서 열차가 D 지점에 도착한 것에 대한 가장 적절한 설명은 아마도 철도의 전반적인 연결 구조를 지적하는 설명일 것이다.

이런 점에서 위의 스위치 상황은 흔히 인과적 설명이론의 반례로 제시되는 “총체적 평형 설명(equilibrium explanation) 상황”이라고 불리는 상황과 닮아 있다.⁴⁾ 예를 들어 모퉁한 접시에 쇠구슬을 놓았다고 가정해 보자. 그 쇠구슬은 굴러가다 결국 그 접시의 가장 낮은 부분에 멈출 것이다. 왜 그 구슬이 접시의 가장 낮은 부분에 멈추었는지를 설명하기 위해서 그 쇠구슬이 지나온 실제의 인과적 경로를 지적하는 것은 적절하지 않다. 주어진 상황에서 그 쇠구슬이 어떤 위치에 어떤 속력으로 놓였든 그 쇠구슬은 결국 접시의 가장 낮은 부분에서 멈추었을 것이기 때문이다. 위에서의 열차 사례와 마찬가지로 이 쇠구슬 사례는 스위치 구조를 갖는다. 실제 그 쇠구슬이 접시의 특정 위치에 특정한 속도로 놓이는 사건이 그 쇠구슬이 최종적으로 접시의 가장 낮은 부분에서 멈추는 사건의 원인 사건이기는 하지만 그 구슬이 그 특정 위치에 그 특정 속도로 놓이지 않고 다른 위치에 (다른 속도로) 놓였다고 하더라도 그 다른 위치에 놓이는 사건이 동일한 결과를 야기하였을 것이다. 위의 열차 사례와 마찬가지로 쇠구슬 사례에서도 단순한 선언 역시 설명적 힘을 갖지 않는다. “위치1 또는 위치2 또는 위치3, ..., 위치n에 쇠구슬이 놓였기 때문에 그 쇠구슬은 접시의 가장 낮은 부분에서

4) 평형 설명에 관한 논의는 E. 소버의 고전적인 논문(Sober 1983)에서 시작되었다. 소버는 평형 설명이 선언적 설명이라는 것은 명시적으로 이해했지만 평형 설명이 갖는 주요 특징들이 그 스위치 구조 때문에 발생한다는 것은 정확하게 이해하지 못한 것으로 보인다. 보다 자세한 내용은 다음을 참고할 것. 이재호, 박재희 (2024).

멈추게 되었다”는 전혀 설명적이지 않다. 이 경우 아마도 가장 적절한 설명은 접시의 전반적인 형태적 특징과 중력의 작용을 언급하는 설명이 될 것이다.

그러나 스위치 구조를 갖는 평형 설명 가운데는 국소적인 형태의 평형 설명도 있다. 여기서 평형 설명이 국소적이라는 것은 평형 설명이 갖는 스위치 구조에서 스위치를 어디에 두건 동일한 결과가 야기되는 것은 아니지만 스위치의 **현실적인 위치 주변의 위치는 모두 동일한 결과를 야기한다**는 것을 의미한다. 예를 들어, 철수가 유리컵에 든 음료를 마시고 사망했다고 가정해 보자. 그 음료 안에는 청산가리(시안화 칼륨) 0.8g이 들어 있었는데 청산가리의 치사량은 0.2g이다. 철수는 분명 청산가리 0.8g을 먹은 것이 원인이 되어 사망했다. 그러나 우리는 이 경우 철수가 정확히 청산가리 0.8g을 먹었기 때문에 사망했다고 말할 수는 없다. 왜냐하면 그가 청산가리 0.81g을 먹거나 0.79g을 먹었어도 그는 여전히 사망했을 것이기 때문이다. 이 사례에서 실제 원인 사건이 발생하지 않는 최근접 세계에서 그 실제 원인 사건을 대체하는 사건 역시 동일한 결과를 야기한다. 스위치 구조를 갖는 사례에서 실제 원인이 되는 선언지가 설명적 힘을 갖지 않듯이 이 사건에서도 실제 원인은 설명적 힘을 갖지 않는다. 그렇다면 이 상황에서 철수의 죽음은 어떻게 설명되어야 하는가? 이 상황에서 가장 적절한 설명은 다음의 선언적 설명2이다.

선언적 설명2(국소적 평형 설명 상황): (1) 철수는 0.2g 이상의 청산가리를 먹었기 때문에 사망했다. ((2) 즉, 그는 0.2g 또는 0.21g 또는 0.22g 또는 ... 의 청산가리를 먹었기 때문에 사망했다.)

이 설명의 설명항은, 0.2g이상의 특정 양의 청산가리를 먹는 사건들을 그 선언지로 갖는 선언이며 따라서 이 설명은 선언적 설명이 된다. 이 국소적 형태의 평형 설명 상황에서는 일반적인 스위치 구조를 갖는 상황에서의 설명이 그러하듯이 선언지는 설명적 힘을 갖지 않는다. 그러나 이 경우에는 총체적 형태의 평형 설명에서와는 달리 선언은 설명적 힘을 갖는다. 이

런 특성은 평형 설명에 관한 논의의 출발점을 이루는 E. 소비의 예, 즉 R. A. 피셔의 1:1 성비에 관한 평형 설명에서도 그대로 나타난다. 피셔가 제시하는 평형 설명은 다음과 같다. 만약 어떤 집단의 성비가 1:1로부터 크게 벗어날 경우 더 적은 수의 성을 더 많이 생산하는 부모가 재생산적으로 더 적합하게 된다. 왜냐하면 그들이 생산하는 자식들이 재생산에 성공할 확률이 높아질 것이기 때문이다. 반면에 이 과정을 통해서 성비가 다시 1:1이 될 경우 특정한 성의 자식들을 더 많이 생산하는 부모가 가졌던 재생산적 이점은 사라진다. 이런 메커니즘을 통해서 인구집단의 성비는 1:1이라는 평형상태에 도달하고 이 상태는 지속된다.⁵⁾

소비가 지적하듯이 피셔의 설명에서 어떤 인구집단의 초기 성비는 설명적인 힘을 갖지 않는다. 왜냐하면 어떤 성비로 출발했든 피셔가 지적하는 메커니즘에 따라서 중국에는 1:1의 성비로 귀결될 것이기 때문이다. 그러면 특정 인구집단의 성비가 1:1이라는 것은 어떻게 설명되어야 하는가? B. 스코우에 따르면, 이 경우 가장 적절한 설명은 그 인구 집단의 부모들이 아들 또는 딸만 낳는 성향을 갖지 않았다는 것이다.⁶⁾ 모두가 아들만 낳거나 모두가 딸만 낳는 극단적인 상황에서는 피셔가 제시한 메커니즘은 작동하지 않으며 그 인구 집단은 절멸된다. 그러나 이런 극단적인 상황을 배제할 경우 피셔의 메커니즘이 작동해 그 인구 집단은 결국 1:1 성비라는 평형

5) 이재호, 박재희 (2024), p. 7.

6) Skow (2016), p. 69. 스코우는 해당 인구집단의 초기 성비는 최종적인 1:1 성비의 원인이 될 수 없으며 ‘아들 또는 딸만 낳지는 않음’이라는 것만이 진정한 원인이 된다고 주장하며, 이에 기초해 피셔의 평형 설명이 인과적 설명이론에 대한 반례가 될 수 없다고 주장한다. 이 점에서 스코우는 사실상 선언적 인과의 가능성을 받아들이고 있는 것으로 보인다. 필자는 이 주장을 받아들이지 않는다. 필자는 소비가 주장하듯이 초기 성비가 진정한 원인이라는 것을 받아들이며, ‘아들 또는 딸만 낳지는 않음’은 그것이 갖는 선언적 성격 때문에 원인이 될 수 없다고 생각한다. 그러나 이 논문의 논증 구조 하에서 필자의 이런 생각이 옳은지는 그렇게 중요하지 않을 수 있다. 이 논문의 핵심 주장은 우리가 선언적 인과를 받아들일 수 없다고 하더라도 선언적 설명은 받아들여야 한다는 것이다. 이 주장을 논증하기 위해서 필자는 이 논문에서 선언적 인과는 존재하지 않는다는 것을 전제하고 논의를 진행할 것이다.

상태에 도달한다. 특정 인구집단의 부모들이 아들 또는 딸만 낳는 성향을 갖지 않았다는 것은 1:1, 2:1, 3:1 등의 다양한 비율을 그 선언지로 갖는 선언적 사실이며 따라서 스코우의 해석 하에서 피셔의 설명은 국소적 평형 설명이 갖는 모든 특징을 갖는다.

이제 국소적 평형 설명에 대한 우리의 직관, 즉 선언적 설명2에 대한 우리의 직관을 받아들일 경우 우리는 선언적 인과의 존재를 부정한다고 하더라도 선언적 설명의 존재는 받아들여야 한다. 왜냐하면 이 경우 스위치 구조의 존재 때문에 선언지는 설명적 힘을 아예 잃어버리고 따라서 선언적 설명1에 대한 직관에 저항하기 위해 우리가 위에서 검토해 봤던 전략, 즉 선언지 설명에 대한 직관을 최대한 존중하고 선언 설명에 대한 직관을 최대한 무시하는 전략은 여기서는 작동하지 않게 되기 때문이다. 그리고 이 사실은 역으로 선언적 설명1에 대한 이런 방식의 저항은 아무런 동기가 없다는 것을 보여준다. 우리가 일단 선언적 설명2에 대한 직관을 받아들인다면 우리는 이미 선언적 설명의 존재를 받아들인 것이 되는데 그런 상황에서 선언적 설명1에 대한 우리의 직관에 굳이 저항하는 것은 별로 동기 부여되지 않기 때문이다.

지금까지 우리의 논의를 정리하면 다음과 같다. (1) 우리는 선언적 설명의 존재를 받아들여야 한다. (2) 선언적 설명에는 다양한 종류가 있을 수 있다. 이 절의 나머지 부분에서는 선언적 설명의 종류를 보다 정교화하는 작업을 수행하고자 한다. 이미 위에서의 논의에서 분명하게 드러났듯이 선언적 설명에는 스위치 구조를 포함하는 선언적 설명과 스위치 구조를 포함하지 않는 선언적 설명이 있다. 스위치 구조를 포함하는 선언적 설명의 경우 선언지는 설명력을 갖지 않으며 (종종) 선언만이 설명력을 갖는다. 반면에 스위치 구조를 포함하지 않는 선언적 설명에서는 선언지와 선언이 모두 무시할 수 없는 설명력을 갖는 것으로 보이며, 적어도 선언적 설명1의 경우에는, 선언이 선언지보다 더 나은 설명적 힘을 갖는다. 더 나아가 어떤 선언적 설명은 명시적으로 선언적이다. 위의 선언적 설명1이 그런 경우이다. 그에 반해서 어떤 선언적 설명은 명시적으로 선언적이지는 않지만 내용적으로 선언적이다. 예를 들어 “0.2g이상의 청산가리를 먹음”은 명시적

으로 선언적인 문장은 아니다. 그러나 내용적으로는 선언적이다. 이런 특징은 존재 양화된 사실을 통한 설명의 경우 두드러진다. 예를 들어 육사 동기인 철수, 영수, 명수, 준수가 군사 반란을 모의했다고 가정해 보자. 이들의 쿠데타 계획은 실행 전에 정보가 새어 나가 발각이 되었고 실패로 끝나게 되었다. 왜 이들의 쿠데타는 실패했는가? 이 질문에 대한 가능한 하나의 대답은 다음과 같은 선언적 설명³이다.

선언적 설명3(존재양화 설명): 누군가가 쿠데타 계획을 밀고했기 때문에 쿠데타는 실패했다. (쿠데타 모의 세력 안에 밀고자가 존재했기 때문에 쿠데타가 실패한 것이다.)

이 설명은 직관적으로 수용 가능한 설명이며 이 설명에서 설명항은 존재 양화되어 있다. 그리고 이 존재양화된 설명항은 내용적으로 선언적이다. 즉 쿠데타 모의 세력 안에 밀고자가 존재했다는 것은 철수 또는 영수 또는 명수 또는 준수가 밀고했다는 것을 의미한다. 만약 이 예에서 실제 쿠데타 계획을 밀고한 사람이 영수였다고 하자. 그 경우 우리는 “그 쿠데타 모의는 영수의 밀고 때문에 실패했다”라는 선언지를 통한 설명 역시 갖게 되며 이 경우 (일반적인 문맥 하에서) 선언지를 통한 설명(선언지 설명)은 선언을 통한 설명(선언 설명)보다 더 만족스럽다.⁷⁾ 최소한 이 경우 우리는 쿠데타의 실패에 대한 보다 많은 정보를 얻게 된다.

따라서 우리는 선언적 설명을 스위치 구조의 유무와 추상화된 개념의 사용 여부에 따라서 다음과 같이 4가지로 구분할 수 있다.

7) 존재 양화 설명에서 선언지 설명이 선언설명보다 항상 더 만족스러운 것은 아니다. 예를 들어 위의 사례에서 영수가 밀고하지 않았을 경우 준수가 밀고할 예정이었다고 가정해 보자. 이 경우 영수가 밀고했기 때문에 쿠데타가 실패했다는 설명은 오히려 선언을 통한 설명보다도 불만족스럽다. 또 쿠데타의 역사와 실패 양태를 논의하는 수업 시간에서는 추상화된 개념을 사용하는 존재양화 설명이 특정한 밀고자를 적시하는 설명보다 더 만족스러울 수도 있다. 이런 선언적 설명이 갖는 문맥 상대성에 대한 논의는 후에 다시 다뤄질 것이다.

		스위치 구조 O	스위치 구조 X
추상화된 개념 사용	사례	선언적 설명2(1) (국소적 평형 설명)	선언적 설명3 (존재양화 설명)
	설명력	선언지 X 선언 O	(종종) 선언지 > 선언
명시적인 선언 사용	사례	선언적 설명2(2) (국소적 평형 설명)	선언적 설명1 (선취, 과잉결정 사례)
	설명력	선언지 X 선언 O	(종종) 선언지 < 선언

〈표1〉

이제 우리가 해야 할 것은 이런 다양한 선언적 설명에서 왜 우리는 상이한 설명적 직관을 갖게 되는지를 체계적으로 설명하는 일이다. 필자는 이 논문의 나머지 부분에서 지금까지 과학적 설명에 관한 철학적 논의들을 지배해온 주요 이론들이 이 기대되는 체계적 설명을 제시하는 데 실패한다는 것을 보이고 필자가 다른 곳에서 제시한 바 있는 새로운 조작주의적 설명은 바로 그것을 제시할 수 있다는 것을 보이고자 한다. 필자는 앞으로 햐펠의 D-N 모델, 인과적 설명 모델, 통일 이론, 단순 반사실적 의존성 모델 등을 검토할 것이다. 필자는 잘 알려진 설명에 관한 이론들 가운데 반 프라센의 화용론적 설명 이론과 우드워드의 조작주의적 설명 이론은 검토하지 않을 것이다. 반 프라센의 이론은 새먼과 키처에 의해서 이 이론이 완전한 설명이론이라고 보기 어렵다는 점이 잘 밝혀졌기 때문에 생략되었다(P. Kitcher and Salmon 1987). 후자의 경우 그 이론이 갖는 이중성 때문에 생략되었다. 우드워드의 이론은 한편으로는 인과적 설명 모델로서의 성격을 가지며 다른 한편으로는 반사실적 의존성 모델로서의 성격을 갖는다(Woodward 2003). 따라서 필자는 이 이론의 한계가 다른 이론들을 검토하는 과정에서 자연스럽게 드러나게 될 것으로 믿는다.

3. D-N 모델과 선언적 설명

D-N 모델은 비록 오늘날 이 모델의 원래 형태를 그대로 수용하는 철학

자는 없지만 이후의 모든 설명 이론의 출발점이 되었다는 점에서 여전히 중요하게 다루어질 필요가 있다. D-N 모델은 기본적으로 2 개의 반례 그룹에 의해 붕괴되었다. 첫 번째 반례 그룹은 “인과적 반례”라고 불릴 법한 반례들의 모임인데, 인과적 비대칭성 사례(브롬버거의 깃대 사례), 인과적 무관성 사례(주문걸린 소금 사례), 인과적이지 않은 공변 사례(기압계 눈금 하강과 폭풍우 사례) 등과 같은 반례들이다. 이런 반례들이 공통적으로 공격하는 D-N 모델의 특징은 이 모델이 설명항 사건과 피설명항 사건 사이의 인과적 연결을 요구하지 않으며 단순한 연역적 도출 관계만을 요구한다는 것이다. 비대칭성 사례들은 이런 연역적 도출 관계가 원인에서 결과로의 도출뿐만 아니라 결과에서 원인으로의 도출에서도 성립할 수 있다는 것을 활용한다. 인과적 무관성 사례는 연역 추론의 단조성(monotonicity) 때문에 인과적으로 유관한 것에 인과적으로 유관하지 않은 것을 덧붙여도 여전히 연역적 도출 관계가 성립한다는 사실을 활용한다. 인과적이지 않은 공변 사례는 공변 관계만 성립하면 인과 관계가 성립하지 않아도 연역적 도출 관계가 성립한다는 사실을 활용한다.

D-N 모델이 이런 문제를 갖는다는 사실은 자체로는 D-N 모델이 선언적 설명에 대한 직관을 포섭하지 못하게 만들지 않는다. 오히려 D-N 모델은 설명항 사건과 피설명항 사건 사이의 인과적 연결을 요구하지 않기 때문에 보다 자유롭게 선언적 설명을 허용할 수 있다. D-N 모델이 선언적 설명과 관련해서 갖는 문제는 이 모델이 선언적 설명을 수용할 수 없다는 것이 아니라 선언적 설명을 너무 “무절제하게” 수용한다는 사실이다. D-N 모델이 갖는 이런 특징은 D-N 모델을 붕괴로 이끌어간 두 번째 반례들의 그룹에서 잘 나타난다. 이 두 번째 그룹의 반례들은 “선언을 활용한 기술적 반례들”이라고 불릴 법한 반례들이며 카플란 등에 의해서 제기된 바 있다(Eberle, Kaplan and Montague 1961). 필자는 비슷한 특징을 갖는 새먼의 반례를 활용해 이 문제를 설명할 것이다.⁸⁾

우선 T를 “모든 사람은 불완전하다($(\forall x)Ix$)”라고 해 보자. 그럴 경우 다

8) Salmon (1989), pp. 21-2.

음의 T'은 T의 하나의 논리적 귀결이다.

T': $(\forall x)(\forall y)(Ix \vee (Py \supset My))$ (모든 사람은 불완전하거나 모든 철학자(P)는 남자(M)다)

이제, C를 “만약 콰인(q)이 불완전하거나 햄펠(h)이 철학자가 아니라면 햄펠은 남자이다 $((Iq \vee \sim Ph) \supset Mh)$ ”라고 해보자. 마지막으로 E를 “햄펠이 남자이다(Mh)”라고 하자. 그럴 경우 우리는 다음의 D-N 추론을 만들 수 있다.

T': $(\forall x)(\forall y)(Ix \vee (Py \supset My))$
 C: $((Iq \vee \sim Ph) \supset Mh)$
 E: Mh

모든 사람이 불완전하다는 것은 틀림없는 사실이고 그것은 법칙일 수도 있지만 그 사실이 햄펠이 남자라는 것에 설명적으로 유관하다는 것은 불합리하다. 그러나 위의 추론은 D-N 모델이 제시하는 모든 조건을 만족시킨다.⁹⁾ 이 반례는 분명 C가 갖고 있는 선언적 성격을 활용한다. 따라서 이런 문제를 막기 위해서는 C에 선언이 무제한적으로 들어오는 것을 막는 장치가 필요하다. 실제로 이런 종류의 반례가 한창 논의될 당시 가장 영향력 있었던 대응 방식은 카플란이 제안한 다음의 조건을 삽입하는 것이었다: “단칭 전제 C는 참인 기초 문장들의 연언이어야 한다.”¹⁰⁾ 김재권은 이후에 카플란의 순수하게 논리적인 제안에 형이상학적인 의미를 부여한다. 그는 카플란의 이 제한은 사실상 단칭 전제 C가 하나의 사건을 기술해야 한다는

9) T'은 (T가 참이고 T'은 T의 귀결이므로) 참인 본질적으로 일반화된 문장이며, C는 (Iq와 Mh가 참이므로) 참인 단칭 문장이다. 이 둘로부터 E는 도출 가능하다. 그리고 $\sim Iq$ 와 Ph로 이루어진 T'과 양립 가능한 기초 문장의 집합은 (C의 전제를 거짓으로 만드는 것을 통해) C를 귀결로 갖지만 Mh를 귀결로 갖지는 않는다.

10) Kaplan (1961), p. 433. 강조는 필자. 여기서 기초문장은 원자 문장이거나 그것의 부정을 의미함.

것을 의미한다고 이해한다.¹¹⁾ 우리가 선언적 사건을 받아들일 수 없다면, 그리고 단칭 설명항은 피설명항의 원인이 되는 사건을 기술해야 한다면 C 자리에 선언적인 문장은 금지되어야 한다는 것이 김재권의 생각이다. 카플란-김재권의 이런 제안은 분명 선언을 활용한 기술적 반례들에 대응하는데 있어 효과적이다. 그러나 우리는 여기서 D-N 모델이 선언적 설명과 관련해 갖는 근원적인 딜레마를 발견하게 된다. 만약 우리가 원래의 D-N 모델을 고수한다면 우리는 D-N 모델이 너무나 쉽게 선언적 설명을 허용하는 문제를 갖게 된다. 이 경우 설명적이지 않은 선언적 사실이 설명적인 것으로 판명되는 문제를 갖는다. 반면에 우리가 이 문제를 막기 위해서 카플란-김재권 식의 제한을 가한다면 우리는 선언적 설명 모두를 포기하게 되어 선언적 설명에 대한 우리의 직관을 수용할 수 없게 된다.

김재권은 분명 이 문제를 인식했으며 따라서 그는 우리의 선언적 설명에 대한 직관을 해소하기 위한 논증을 제시한다. 그는 다음과 같은 선언적 D-N 설명을 고려해 볼 것을 권유한다.

메리는 낭창 또는 류마티스 관절염을 갖고 있다.

낭창 환자는 관절에 통증을 갖는다.

류마티스 환자는 관절에 통증을 갖는다.

따라서, 메리는 관절에 통증을 갖는다.

그는 이 상황에서 우리는 하나의 **선언적 설명**을 갖는 것이 아니라 (잠재적인) 두 개의 **설명**의 선언을 가질 뿐이라고 주장한다. 그는 “메리의 상황에 대한 설명을 갖기 위해, 즉 **그것의 원인을 알기 위해**, 우리는 추가로 해야 할 일을 갖고 있다. 우리는 메리가 낭창을 갖는지 류마티스 관절염을 갖고 있는지를 알아 내야 한다”라고 말한다.¹²⁾ 필자는 김재권의 직관에 동의하지 않는다. 김재권의 주장은 위의 인용문이 잘 드러내듯이 강한 형태의 인과적 설명 이론, 즉 설명이 성립하기 위해서는 원인을 지적할 수 있어야

11) Kim (1999), p. 13, n. 27.

12) Kim (1999), p. 13. 강조는 필자.

한다는 것을 전제하고 있다.¹³⁾ 그러나 우리가 선언적 설명에 대한 직관을 논의할 때 이런 강한 인과적 설명이론을 전제할 이유는 없다. 뿐만 아니라 우리는 김재권의 직관을 거부할 좋은 이유를 갖고 있다. 예를 들어 다음의 문장을 생각해 보자.

메리는 낭창 또는 류마티스를 가졌기 때문에 관절에 통증을 가졌다.

이 문장은 상정된 상황에서 분명 참이다. 이는 우리가 관련된 선언적 문장과 피설명항 문장 사이에 “때문에” 관계가 성립한다는 것을 인정한다는 것이다. 때문에 관계의 성립과 설명적 관계의 성립이 동일한 것은 아니라고 하더라도 이들 밀접하게 연결되어 있다는 것을 부정할 수는 없다. 그렇다면 우리는 이 경우 하나의 선언적 설명을 갖는다는 직관을 부정할 좋은 이유는 없다. 물론 이 경우는 앞서 선언적 설명³의 경우와 마찬가지로 선언지 설명이 선언 설명보다 더 만족스러운 경우이다. 이 점에서 김재권이 지적한 대로 우리는 해야 할 추가적인 작업이 있다. 그러나 선언지 설명이 선언 설명보다 만족스럽다는 것이 선언 설명이 설명이 아니라는 것을 의미하는 것은 아니다. 만약 우리가 낭창과 류마티스를 구분할 진단 방법을 갖고 있지 않다면 현재의 증거 상황에서 이 선언 설명은 우리가 가질 수 있는 최선의 설명이 될 것이다.

더 나아가 선언적 설명에 대한 직관을 김재권식으로 해소하는 것이 필자가 강조한 바 있는 스위치 구조가 존재하는 선언 설명의 경우에 작동할 가능성은 별로 없다. 그렇다면 D-N 모델이 갖는 두 번째 유형의 반례들을 카플란-김재권 식의 처방을 통해서 극복하려는 시도는 그리 바람직한 시도라고 볼 수는 없다.

13) 강한 형태의 인과적 설명 모델과 약한 형태의 인과적 설명 모델의 차이는 다음 절에서 설명될 것이다.

4. 인과적 설명이론과 선언적 설명

인과적 설명이론은 태생적으로 선언적 설명에 대한 직관을 수용하기에 별로 적절하지 않은 설명 이론이다. 선언적 인과에 대한 부정적인 생각이 선언적 설명에 대한 부정적인 생각으로 이어지는 것을 차단하기 가장 어려운 이론이기 때문이다. 그러나 이런 인상에 기초해 인과적 설명이론이 선언적 설명에 대한 직관을 수용할 수 없는 이론이라고 쉽게 단정하는 것은 성급할 수 있다. 인과적 설명 이론은 단일한 하나의 이론이라기보다는 다양한 이론들의 집합이며 아주 약한 형태의 인과적 설명 이론은 어떤 형태의 선언적 설명을 수용할 수 있기 때문이다. 모든 형태의 인과적 설명 이론이 공통적으로 주장하는 것은 설명적 관계의 분석에 있어서 인과적 관계가 큰 역할을 한다는 것이다. 그러나 설명적 관계와 인과적 관계가 관계를 맺는 데는 다양한 방식이 있을 수 있으며, 필자의 견해로는 다음의 두 개의 기준이 특정한 인과적 설명 이론의 성격(강도)을 규정하는 데 도움을 준다.

- (1) 구체적인 원인의 제시가 설명과 연결되는지 아니면 (구체적인 원인의 제시 없는) 적절한 인과적 정보 제시만으로도 설명과 연결되는지 여부
- (2) 원인의 제시 또는 인과적 정보의 제시가 설명관계의 성립에 단지 필요하기만 한 것인지(즉, '설명→인과' 논제만 받아들일 것인지) 아니면 충분하기까지 한 것인지(즉, '인과→설명' 논제까지 받아들일 것인지) 여부

이 기준에 따르면 원리적으로 다음의 4개의 인과적 설명 이론의 종류가 있을 수 있으며 그 각각은 모두 실제 사례를 갖고 있다.

	'설명→인과'만 요구	'설명→인과'와 '인과→설명' 모두 요구
인과적 정보의 제시만으로 충분	스트레븐스의 인과적 설명이론(Strevens 2004)	루이스의 인과적 설명이론(Lewis 1986a)

실제 원인의 제시가 필요	새먼의 수정된 인과적 설명이론(Salmon 1997)	새먼의 수정 전 인과적 설명이론(Salmon 1984), 스코우의 인과적 설명 이론(Skow 2016)
------------------	----------------------------------	--

〈표2〉

인과적 설명이론의 가장 기본적인 아이디어는 어떤 것을 설명하는 것은 그것의 (구체적인) 원인이 무엇인지를 알려 주는 것이라는 것이며 새먼의 초기 인과적 설명이론과 스코우의 인과적 설명 이론이 이런 **가장 강한 형태의** 인과적 설명이론에 해당된다. 이 이론이 선언적 설명에 대한 우리의 직관과 양립 불가능하다는 것은 명백하다. 평형 설명의 사례에서 잘 드러나듯이 어떤 경우 실제의 원인을 제시하는 것은 전혀 설명적이지 않다. 즉 우리가 평형 설명에 관한 직관을 받아들이면 우리는 ‘인과→설명’ 논제, 즉 피설명항의 실제 원인이 무엇인지를 제공하면 설명을 제공할 수 있다는 논제를 부정해야 한다.

평형 설명의 사례는 구체적 원인의 제시를 요구하는 설명→인과 논제 역시 받아들이 수 없다는 것을 보여준다. 위에서 언급된 피서의 평형 설명의 경우 인구 집단의 초기 성비, 즉 실제 원인을 전혀 언급하지 않으면서도 충분히 만족스런 설명이 된다. 마찬가지로 오목한 접시에 놓여진 쇠구슬 사례에서도 가장 만족스런 설명은 구체적인 실제 원인을 전혀 언급하지 않는다. 그 설명은 접시의 전체적인 형태적 특징과 중력의 작용에 대해서 언급할 뿐이다. 따라서 평형 설명에서의 우리의 직관과 양립 가능할 가능성이 있는 유일한 인과적 설명 이론은 ‘설명→인과’ 논제만 받아들이면서 ‘인과’라는 조건을 실제 원인을 제시하는 것으로 이해하지 않고 인과적 정보를 제공하는 것으로 약하게 이해하는 이론, 즉 **가장 약한 형태의** 인과적 설명이론이다. 이런 형태의 인과적 설명이론은 ‘설명→인과적 정보 제공’ 논제만을 받아들이는데 스트레븐스의 카이레틱(Kairetic) 이론이 그 대표적인 사례이다. 이 이론에 따르면 설명적 관계는 “인과 관계 + 차이 만들기”로 이해될 수 있다. 즉, 어떤 것이 피설명항 사건의 설명항이 되기 위해서

는 그것이 차이를 만들어 내는 원인이어야 한다. 문제는 “차이를 만들어 낸다”를 어떻게 이해할 것이냐이다. 스트레븐스는 어떤 것이 차이를 만들어 내는 원인이기 위해서는 가장 추상화된 방식으로 그것으로부터 피설명항을 군더더기 없이 도출할 수 있어야 한다고 주장한다. 이 점에서 스트레븐스의 이론은 인과적 설명 이론의 장점과 키처의 통일 이론의 장점을 교묘하게 종합하려고 한다.¹⁴⁾ 스트레븐스는 설명은 단순한 연역적 도출이 아니라 **인과적인** 연역적 도출이어야 한다고 생각하며, 단순히 인과적인 연역적 도출이 아니라 가장 (추상화된) **통일적인** 연역적 도출이어야 한다고 생각한다. 스트레븐스의 이론에서 전자의 조건은 (인과 개념을 도입하지 않는) D-N 모델의 문제를 해결하는 역할을 하며, 후자의 조건은 평형 설명이 제기하는 인과적 설명 모델의 문제를 해결하는 역할을 수행한다. 앞서 언급되었던 청산가리 사례를 생각해 보면 철수가 0.8g의 청산가리를 먹었다는 것으로부터 철수의 죽음을 인과적으로 연역적으로 도출하는 것은 가능하다. 그러나 이 인과적 연역적 도출은 스트레븐스에 따르면 설명적이지 않은데, 그것은 가장 추상화된(통일적인) 도출이 아니기 때문이다. 철수가 0.8g의 청산가리를 먹은 사건은 가장 통일적인 방식으로 추론을 만들기 위해 철수가 0.2g 이상의 청산가리를 먹은 사건으로 추상화되어 기술되어야 하며, 우리가 원하는 만족스런 설명은 (기대되는 대로) 철수는 0.2g 이상의 청산가리를 먹었기 때문에 죽었다는 것이 된다. 스트레븐스에 따르면 이런 추상화 과정이 극단적으로 이루어진 경우가 총체적 평형 설명이 되며 이 경우 우리는 전체적인 인과적 구조에 대한 정보를 제공할 뿐 구체적인 원인을 제공하지는 않는다.

이렇게 보면 스트레븐스의 인과적 설명이론은 키처의 통일이론이 갖는 특징을 빌려 오는 것을 통해서 평형 설명에서의 우리의 설명적 직관을 설명하려고 한다고 볼 수 있으며, 따라서 키처의 통일이론이 얼마나 선언적 설명에 대한 우리의 직관에 잘 대처하는지에 따라서 스트레븐스의 인과적

14) 그래서 스트레븐스는 자신의 이론이 인과적 설명이론과 통일 이론을 인과적으로 통일한 이론이라고 부른다(Strevens 2004, 2008).

설명이론이 얼마나 선언적 설명에 대한 직관을 잘 설명할 수 있는지가 결정된다. 이런 점에서 스트레븐스의 인과적 설명이론은 별로 인과적 설명 이론답지 않은 방식으로 선언적 설명에 대한 직관에 대처한다. 이는 인과적 설명 이론이 태생적으로 선언적 설명과 양립 가능하기 쉽지 않은 이론이라는 점에서 충분히 이해 가능하다.

지금까지의 논의를 종합해 보면 다음과 같다. 일반적으로 인과적 설명이론은 선언적 설명에 대한 직관을 수용하기에 적합한 이론이 아니며 인과적 설명 이론과 선언적 설명에 대한 직관을 화해시키기 위해서는, 설명적 관계를 '인과적 정보 제공+ α '로 생각하는 가장 약한 형태의 인과적 설명 이론을 받아들일 필요가 있다. 그리고 이런 형태의 인과적 설명이론에서 선언적 설명에 대한 직관의 수용을 담당하는 부분은 '인과적 정보 제공' 이라기 보다는 ' α '이며 일반적으로 이 α 를 뒷받침하는 직관은 인과적 설명이론에 고유한 직관이 아니라 어떤 다른 설명이론이 호소하는 직관에 기반한다. 스트레븐스의 경우 이 α 의 역할을 하는 것이 통일 이론이 호소하는 직관에 기반하고 있으며 따라서 스트레븐스의 이론이 얼마나 선언적 설명에 대한 직관을 잘 설명하는지는 통일 이론이 얼마나 선언적 설명에 대한 직관을 잘 설명하는지에 따라 크게 좌우된다.

5. 통일 이론과 선언적 설명

키처의 통일 이론(Philip Kitcher 1989)은 크게 봐서 햄펠의 D-N모델의 후예라고 볼 수 있다. 이 이론들은 모두 설명이 기본적으로 일반화된 규칙성으로부터 피설명항을 연역적으로 도출하는 것이라고 생각한다. 물론 모든 일반화된 규칙성으로부터의 도출이 설명적이지는 않다. 예를 들어, 이 파티에 참석한 모든 사람이 키가 크다는 것과 철수가 이 파티에 참석했다는 것으로부터 철수가 키가 크다는 것을 도출할 수 있지만 이 도출은 철수가 왜 키가 큰 지를 전혀 설명해주지 않는다. 따라서 우리는 설명적인 도출과 설명적이지 않은 도출을 구분하는 장치를 갖고 있어야 하는데, 햄펠과

키쳐는 이 장치가 어떤 성격을 가져야 하는지에 대해서 다른 생각을 갖는다. 햄펠은 기본적으로 규칙성이 법칙적이게 되기 위한 다양한 논리적 조건과 법칙으로부터의 도출이 가져야 할 추가적인 논리적 조건을 제시하는 것을 통해서 설명적인 도출과 설명적이지 않은 도출을 구분하려고 시도한다. 이런 점에서 햄펠의 구분장치는 **구문론적이다**. 이에 반해서 키쳐는 구문론적인 장치에 호소하지 않는다. 키쳐에 따르면 어떤 도출이 설명적인지는 그 도출의 논리적인 특징에 따라 결정되는 것이 아니라 그 도출이 현재 우리가 갖고 있는 지식체계 안에서 가장 통일적인 도출 패턴의 사례이냐에 따라서 결정된다. 위의 파티 사례에서 “이 파티에 참석한 모든 사람이 키가 크다”는 규칙성은 기껏해야 그 파티에 참석한 사람들에게 대해서만 적용되는 따라서 통일적 힘이 매우 약한 규칙성이다. 따라서 이런 제한된 규칙성으로부터의 도출은 가장 통일적인 설명적 도출 패턴들의 집합 $E(K)$ 에 포함될 수 없으며 따라서 이런 제한된 규칙성에 호소하는 도출은 설명적이지 않게 된다. 이런 통일 이론의 정신은 위의 선언적 설명2의 경우에 잘 적용된다. 청산가리 0.8g을 먹은 모든 사람은 죽는다는 것과 철수가 청산가리 0.8g을 먹었다는 것으로부터 철수가 죽는다는 것을 도출하는 것은 설명적이지 않는데, 그것은 이 도출이 어떤 문제 있는 논리적 특징을 갖고 있기 때문이 아니라 이 도출이 가장 통일적인 도출이 아니기 때문이다. 이 형식의 도출은 철수의 죽음을 도출하는데 사용될 수는 있지만 0.6g의 청산가리를 먹고 죽은 영수의 죽음을 도출하는 데는 사용될 수 없다. 반면에 0.2g이상의 청산가리를 먹은 모든 사람은 죽는다는 것과 철수가 0.2g 이상의 청산가리를 먹었다는 것으로부터 철수가 죽었다는 것을 도출하는 것은 훨씬 통일적인 설명이다. 이 도출 형식은 철수의 죽음뿐만 아니라 영수의 죽음도 설명할 수 있기 때문이다.

이런 점에서 키쳐의 통일이론은 어떤 종류의 선언적 설명에 대한 직관을 설명할 수 있다. 가장 대표적으로 이 이론은 표1에서 선언적 설명2(1), 즉 추상화된 개념을 사용하는 스위치 구조를 갖는 선언 설명에 대한 우리의 직관

을 잘 설명할 수 있다.¹⁵⁾ 그러나 통일 이론이 모든 종류의 선언적 설명에 대한 직관을 만족스럽게 설명하는 것은 아니다. 통일 이론에는 선언적 설명에 대한 수용을 어렵게 만드는 두 개의 특성이 존재하기 때문이다. 그 첫번째 특성은 통일 이론이 투사 가능한 술어와 그렇지 않은 술어의 엄격한 구분을 요구한다는 것이다. 통일 이론은 가능한 한 최소의 도출 패턴을 사용할 것을 요구한다. 그러나 우리가 게리맨더링 된 패턴을 허용한다면 도출 패턴의 수를 줄이는 것은 사소한 작업이 된다. 키처는 이런 문제를 막기 위해서 다음과 같이 말한다. “우리는 패턴의 게리맨더링을 막을 수 있게 해주는 패턴 개별화에 대한 어떤 요구조건을 필요로 한다. … 이 도전을 해결하는 명백한 방식은 도식적 문장에 나타나는 술어, 채우기 교본을 정식화하는 데 사용되는 술어들, 그리고 부류화에 등장하는 술어들이 모두 투사 가능한 술어들일 것을 요구하는 것이다.”¹⁶⁾ 이런 식으로 투사 가능한 술어만을 설명적 도출의 도식에 사용할 경우 많은 선언적 설명은 통일 이론에서 허용할 수 없는 설명으로 전락한다. 예를 들어 위에서 언급된 바 있는 김재권의 예에서 “낭창 또는 류마티스를 가짐”이라는 술어는 분명 투사 가능한 술어가 아니며 따라서 키처의 통일 이론에서 허용할 수 있는 선언적 설명에서 사용되기 어려운 술어가 된다. 일반적으로 통일 이론은 추상화된 개념을 사용하는 선언적 설명보다는 추상화된 개념을 사용하지 않는 선언적 설명, 위의 선언적 설명1과

15) 앞서 살펴본 스트레븐스의 인과적 설명 이론은 통일 이론이 갖는 이런 장점을 승계하는 이론이다.

16) P. Kitcher, (1989), p. 482. 이런 게리맨더링을 막는 것은 키처의 통일 이론과 비슷한 아이디어를 사용하는 루이스의 법칙 이론에서도 중요하다. 루이스에 따르면 법칙이라는 것은 세계에 발생하는 기초적인 사실들을 체계화하는 최선의 공리화 시스템에서 공리와 상위 정리에 대응하는 것이 법칙이 된다(Lewis 1973). 키처는 우리가 특정 시점에 갖고 있는 지식의 최선의 체계화에 대해서 말하는 반면 루이스는 지식이 아니라 사실의 체계화에 대해서 말하는 점에서 둘의 이론은 구분되어야 하지만 그 체계화의 핵심이 가장 통일적인 도출 시스템을 만들어 내는 것이라는 점에서 이들은 매우 유사한 아이디어에 기반하고 있다. 루이스는 이 최선의 체계화가 게리맨더링에 의해서 사소화되는 것을 막기 위해 키처가 ‘투사 가능한 술어’라는 개념을 도입하는 지점에서 ‘자연적 속성’이라는 개념을 도입한다(Lewis 1983).

같은 선언적 설명에 대한 직관을 설명하는데 난점을 갖는다.

선언적 설명에 관한 우리의 직관을 설명하는데 있어 키처의 통일 이론이 갖는 또 하나의 중요한 문제적인 특징은 이 이론이, J. 우드워드¹⁷⁾가 지적하듯이, 승자독식적(the winner-take-all) 성격을 갖는다는 것이다. 즉 이 이론에 따르면 가장 통일적인 도출만이 설명적이며 그렇지 않은 모든 도출은 설명적이지 않게 된다.¹⁷⁾ 이 이론이 갖는 승자독식적 성격은 앞서 지적했듯이 선언적 설명 2(1)에 대한 우리의 직관을 설명하는 데는 장점이 된다. 왜냐하면 이 경우 우리는 보다 추상적인 개념을 사용하는 선언은 설명력을 갖지만 선언지는 설명력을 갖지 않는다고 생각하기 때문이다. 그러나 선언적 설명 1과 3의 경우 사정이 다르다. 이 경우에 우리는 선언과 선언지 모두가 설명력을 갖는다고 생각한다. 특히 선언적 설명 3의 경우에는 문제가 더 심각하다. 왜냐하면 이 경우에는 보다 추상화된 개념을 사용하는 따라서 보다 통일적인 선언적 설명보다 선언지 설명이 더 강한 설명력을 갖기 때문이다. 우드워드가 지적하듯이, 보다 덜 통일적인 설명도 종종 충분히 설명적일 수 있다. 따라서 통일의 정도는 설명의 정도를 설명하기에 보다 적합한 척도이지 설명력의 유무를 판단하기에 적합한 척도는 아니다. 즉, 통일 이론은 더 좋은 설명과 덜 좋은 설명의 구분을 받아들여야 하는 곳에서 잘 작동하는 설명 이론이 아니다. 그리고 선언 설명에 대한 직관은, 우리가 살펴본 것처럼, 더 좋은 설명과 덜 좋은 설명의 구분을 요구하는 것으로 보인다.

6. 반사실적 의존 이론과 선언적 설명

우리의 설명적 직관이 반사실적 의존성과 매우 깊은 관계를 갖는다는 것은 분명하다. 앞서 지적된 바 있듯이 우리는 전형적인 선취 사례에서 선취하는 원인이 설명력을 갖는다는 무시할 수 없는 직관도 갖지만 선취적 원인이 설명력을 갖지 않는다는 직관 역시 무시할 수 없는 정도로 갖는다.

17) Woodward (2003), p. 367.

그리고 이 후자의 직관 기저에는 선취하는 원인과 결과 사이에는 반사실적 의존성 관계가 성립하지 않는다는 사실이 있는 것은 분명해 보인다. 마찬가지로 선언적 설명 1이 진정한 설명이라는 우리의 직관의 기저에는 여기서 선언적 사실과 피설명항 사실에 명백한 반사실적 의존성 관계가 성립한다는 사실이 있는 것으로 보인다.

만약 반사실적 의존 관계가 설명적 관계에서 중요한 역할을 한다는 것을 받아들이지만 인과적 설명 이론은 받아들일 생각이 없는 사람이라면 설명과 반사실적 의존성을 직접적으로 연결하는 것은 매우 매력적인 대안이 될 수 있으며, 선우환의 이론은 바로 이런 길을 가는 이론이다. 선우환에 따르면 ‘P이기 때문에 Q이다’는 P와 Q가 참이고 $\sim P$ 였더라면 $\sim Q$ 였을 것이다가 참일 경우 오직 그 경우에만 참이다.¹⁸⁾ 선우환의 이론은 선언적 설명과 관련해서 ‘인과에 대한 반사실적 의존성 분석 + 인과적 설명이론’ 조합에 분명한 이점을 갖고 있다. 왜냐하면 이 이론은 선언적 설명1에 대한 우리의 직관을 만족스럽게 설명할 수 있기 때문이다. 선언적 설명1에서 철수 또는 영수가 돌을 던지는 것은 꽃병이 깨지는 것의 원인이 될 수는 없지만 철수 또는 영수가 돌을 던지지 않았다면, 즉 둘이 모두 돌을 던지지 않았다면 꽃병이 깨지지 않았을 것이라는 것은 여전히 참이다. 선우환의 이론은 선언적 설명2에 대한 직관도 잘 설명할 수 있다. 철수가 정확하게 청산가리 0.8g을 먹었다는 것은 그가 죽었다는 사실을 설명할 수 없는데, 그것은 철수가 정확하게 청산가리 0.8g를 먹지 않았다고 하더라도, 즉 그가 0.79g이나 0.81g을 먹었다고 하더라도 그는 여전히 죽었을 것이기 때문이다. 반면 철수가 청산가리 0.2g 이상을 먹었다는 사실은 그가 죽었다는 사실을 설명할 수 있는데, 그것은 그가 청산가리 0.2g 이상을 먹지 않았다면 그는 죽지 않았을 것이기 때문이다. 선우환의 이론은 선언적 설명 3에 대한 직관도 설명한다. 그 쿠데타가 실패한 것은 누군가가 쿠데타 계획을 밀고했기 때문인데, 그것은 밀고자가 없었다면 그 쿠데타는 성공했을 것이기 때문이다. 이 사례에서 선우환의 이론은 선언적 설명이 왜 설명이 되는지에 대해

18) 선우환 (2020), p. 78.

서도 설명할 수 있다. 영수의 밀고가 없었다면 그 쿠데타는 성공했을 것이기 때문이다.

이 모든 것을 고려해 봤을 때, 선우환의 반사실적 의존성 설명이론은 분명 지금까지 우리가 고려해 본 모든 이론들 가운데 선언적 설명에 대한 우리의 직관을 설명하는 데 있어 가장 우위에 있는 이론으로 보인다. 그러나 반사실적 의존성 이론은 선언적 설명과 관련해서 두 개의 치명적인 문제를 갖는다. 그 첫번째 문제는 이 이론이 D-N 모델과 마찬가지로 너무 “무절제하게” 선언적 설명을 용인한다는 것이다. 만약 E가 C에 대해서 반사실적으로 의존한다면, 사소하게 E는 C 또는 D에 반사실적으로 의존한다. 예를 들어 쿠데타의 실패가 영수의 밀고에 반사실적으로 의존한다면 쿠데타의 실패는 영수의 밀고 또는 (쿠데타와는 아무런 관계가 없는) 동건의 잘생김에도 반사실적으로 의존한다. 영수가 밀고하지 않고 동시에 동건이 잘생기지 않았다면 그 쿠데타는 성공했을 것이기 때문이다. D-N 모델이 무관한 **연언지**를 설명함에 무제한적으로 허용하기 때문에 (주문결린 소금 사례와 같은) 인과적 무관성 반례에 직면했던 것처럼 반사실적 의존성 이론은 무관한 **선언지**를 설명함에 무제한적으로 허용하기 때문에 문제가 된다.

다음으로, 반사실적 의존성 이론은 반사실적 의존성 관계가 없음에도 불구하고 설명적 관계가 존재하는 경우를 설명할 수 없다. 선언적 설명 1이 바로 그런 경우다. 선취나 과잉결정 사례에서 우리는 결과가 실제의 원인에 반사실적으로 의존하지 없음에도 불구하고 실제의 원인의 발생이 결과의 발생을 설명할 수 있다는 무시할 수 없는 직관을 갖는데, 왜 우리가 이 직관을 갖는지 반사실적 의존성 이론은 설명하기 어렵다. 인과에 대한 철학적 논의의 역사는 선취 사례가 인과에 대한 반사실적 의존성 분석을 옹호하는 철학자들에게 매우 해결하기 어려운 문제라는 것을 드러냈다. 그리고 반사실적 의존성 설명 이론은 반사실적 의존성 인과 이론이 갖는 이 문제를 계승한다.¹⁹⁾

19) 선우환의 이런 문제들에 대한, 화용론적 고려에 기반한, 답변과 그 답변에 대한 필자의 비판은 다음을 볼 것. 선우환 (2022); 이재호 (2022). 형이상학적 설명과 관련한

7. 새로운 조작주의 이론과 선언적 설명

3절에서 6절까지 우리는 기존의 주요 설명 이론들이 선언적 설명에 대한 우리의 다양한 직관을 제대로 수용하지 못한다는 것을 논증했다. 이 절에서 필자는 필자의 이론이 이런 문제를 갖지 않는다는 것을 논증한다.

필자의 새로운 조작주의적 이론의 기본적인 아이디어는 다음과 같다.

비유적으로 표현해서, X를 설명하고자 하는 욕구는 신이 X를 발생시키기 위해서 실제로 사용했을 것으로 생각될 수 있는 레시피를 훔치려는 욕구이다. 이 욕구의 배후에는 X와 유사한 것들을 유사한 상황에서 제어하려는 욕구가 존재한다(이재호 2023, p. 15).

필자는 이 아이디어를 “확장된 루이스적 다이어그램”을 활용해서 엄밀하게 정식화한 바 있다. 그러나 그런 복잡한 정식화가 이 논문에서의 논의를 위해 꼭 필요한 것은 아니다. 여기서는 다음의 단순화된 그러나 그다지 엄밀하지는 않은 정식화만으로도 충분할 것이다.

- A, 또는 A의 진리 확정적 부분이 실제로 B를 생성하고, A를 발생시키는 것이 자체로 B를 발생시키는데 충분할 경우 A는 B에 대한 강한 긍정적 레시피가 된다.
- A가 자체로 B에 대한 강한 긍정적 레시피는 아니지만 B에 대한 어떤 강한 긍정적 레시피의 군더더기적이지 않은(non-redundant) 부분이라면 A는 B에 대한 약한 긍정적 레시피가 된다.
- A가 발생하는 것을 막는 것을 통해 B가 발생하는 것을 막을 수 있다면 A는 B에 대한 강한 부정적 레시피가 된다.
- A가 자체로 B에 대한 강한 부정적 레시피는 아니지만 이들을 둘러싸고 있는 현재의 인과적 네트워크와 동형적인 어떤 인과적 네트워크 하에서 B에 대한 강한 부정적 레시피가 될 수 있다면 A는 B에 대한 약한 부정적 레시피가 된다.
- A는 B에 대해서 어떤 강도의 긍정적 레시피이면서 동시에 어떤

유사한 비판을 위해서는 다음을 참고할 것. 한성일 (2022).

강도의 부정적 레시피일 경우 설명적이게 되며, 이때 설명의 정도는 긍정적 레시피와 부정적 레시피의 강도에 의존한다.

설명적 관계가 기본적으로 레시피 관계라는 것을 받아들인다면 ‘선언적 설명’이라는 개념은 매우 자연스런 개념이 된다. 왜냐하면 ‘선언적 레시피’라는 개념은 ‘선언적 인과’와는 달리 매우 자연스런 개념이기 때문이다. 예를 들어, 삼국지연의에는 수경선생, 즉 사마휘가 유비에게 “와룡과 봉추 중에서 하나만 얻어도 천하를 손에 넣을 수 있다”라고 조언하는 장면이 나온다. 여기서 수경선생은 분명 유비에게 천하를 손에 넣기 위한 선언적 레시피를 제공하고 있다고 볼 수 있다. 비록 이 레시피는 잘못된 것으로 드러났지만²⁰⁾, 이 레시피가 갖는 선언적 성격 자체에는 아무런 문제가 없다. 수경선생의 레시피는 기본적으로 선언적 설명1의 구조와 유사하다. 선언적 설명1의 상황에서 철수 또는 영수가 돌을 던지는 것은 그 꽃병이 깨지게 만드는데 있어 강한 긍정적 레시피이다. 동시에 철수나 영수가 돌을 던지는 것은 그것을 막았을 경우 그 꽃병이 깨지지 않았을 것이기 때문에 강한 부정적 레시피이다. 이 점에서 이 선언적 설명은 진정한 설명으로서의 자격을 갖는다. 동시에 선언적 설명, 즉 “그 꽃병은 철수가 돌을 던졌기 때문에 깨졌다” 역시 정당화된다. 철수가 돌을 던지는 것은 그 꽃병이 깨지는데 자체로 충분하며 따라서 강한 긍정적 레시피이다. 반면에 관련된 상황에서 철수가 돌을 던지는 것을 막는 것을 통해서 그 꽃병이 깨지지 않게 만들 수는 없다. 철수가 돌을 던지지 않았더라도 영수가 던진 돌에 의해서 그 꽃병은 깨졌을 것이기 때문이다. 그러나 이 상황과 충분히 유사한 상황²¹⁾, 즉 영수가 돌을 던지는 것을 포기하는 상황에서 철수가 돌을 던지는 것을 막는 것은 그 꽃병이 깨지지 않게 만들며 따라서 이것은 강한 부정적 레시피가 될 수 있다. 따라서 철수가 돌을 던지는 것은 이 상황에서 꽃병이 깨지는 것에 대한 강한 긍정적 레시피이자 **약한** 부정적 레시피가 된다. 요약하

20) 유비는 와룡(제갈량)과 봉추(방통)를 모두 얻고도 천하를 얻지 못했다.

21) 필자는 “유사한 상황”을 “x에 대한 어떤 법칙적으로 일관적인 동형적 ELD”라는 개념을 통해서 정식화한 바 있다(이재호, 2023, p. 26).

면 선언적 설명1의 상황에서 선언 설명은 강한 긍정적 레시피와 강한 부정적 레시피를 제공하는 강한 설명이 되고 선언지 설명은 강한 긍정적 레시피와 약한 부정적 레시피를 제공하는 강하지는 않은, 그러나 진정한 설명이 된다. 즉, 필자의 이론은 이 상황에서 왜 우리는 선언 설명과 선언지 설명이 모두 설명이 되는지 뿐만 아니라 왜 이 경우 선언 설명이 선언지 설명보다 더 강한 설명력을 갖는지도 설명할 수 있다.

필자의 이론은 왜 국소적 평형 설명 상황에서 선언은 설명력을 갖지만 선언지는 설명력을 갖지 않는지 역시 설명할 수 있다. 0.2g 이상의 청산가리를 먹이는 것은 철수를 죽이는데 있어 강한 긍정적 레시피가 될 수 있다. 청산가리의 치사량이 0.2g이기 때문이다. 그리고 철수가 0.2g 이상의 청산가리를 먹지 않게 만드는 것은 철수가 죽지 않게 만드는 좋은 방법이므로 이것은 동시에 강한 부정적 레시피가 될 수 있다. 반면 0.8g의 청산가리를 먹는 것은 철수를 죽이는데 있어 강한 긍정적 레시피가 될 수는 있지만 강한 부정적 레시피는 될 수 없다. 그가 0.8g 대신 0.79g이나 0.81g의 청산가리를 먹었어도 그는 죽었을 것이기 때문이다. 더 나아가 이 상황과 인과적으로 유사한 상황 안에서 0.8g의 청산가리를 먹지 못하게 하는 것이 강한 부정적 레시피가 되는 경우는 없다. 모든 유사한 상황에서 철수가 정확하게 0.8g의 청산가리를 먹지 않는 최근접 세계는 그가 0.79g이나 0.81g의 청산가리를 먹는 세계가 될 것이며 그런 세계에서 철수는 여전히 죽었을 것이기 때문이다. 스위치 상황과 선취 상황의 결정적인 차이가 바로 이것이다. 표준적인 선취 상황에서는 선취되는 (대안적) 원인이 발생하지 않는 유사한 상황에서 선취하는 원인과 결과 사이에 반사실적 의존성 관계가 성립한다. 즉, **영수가 돌을 던지지 않는 상황에서** 철수가 돌을 던지는 사건과 꽃병이 깨지는 사건 사이에는 반사실적 의존성 관계가 성립한다. 그러나 스위치 상황에서는 실제 원인에 의해서 대안적 원인 자체가 논리적으로 차단되기 때문에 현실 세계에서 대안적 원인은 발생하지 않는다. 따라서 (사소하게) 대안적 원인이 발생하지 않는 상황에서 피설명항은 실제 원인에 반사실적으로 의존하지 않게 된다. 즉, **철수가 0.79g이나 0.81g의 청산가리를 먹지 않는 상황에서, 즉 현실세계에서의 상황에서** 철수가 0.8g의

청산가리를 먹는 사건과 철수가 죽는 사건 사이에는 반사실적 의존성 관계가 성립하지 않는다. 따라서 스위치 상황에서의 실제의 원인은 단지 강한 부정적 레시피가 될 수 없는 것이 아니라 약한 부정적 레시피조차 될 수 없게 된다. 따라서 필자의 이론은 왜 선언적 설명2에서 선언적 설명은 성립하지만 선언지 설명은 성립하지 않는지를 잘 설명할 수 있다.

필자의 이론은 선언적 설명3에서의 설명적 직관 역시 설명할 수 있다. 쿠데타 모의 세력 가운데 누군가가 계획을 밀고하는 것은 그 쿠데타가 실패하게 만드는 데 있어 강한 긍정적 레시피가 된다. 반면에 철수, 영수, 명수, 준수 가운데 누구도 밀고하지 않았다면 그 쿠데타는 실패하지 않았을 것이라는 점에서 누군가가 밀고하는 것은 쿠데타가 실패하는데 있어 강한 부정적 레시피가 된다. 동시에 영수가 밀고하는 것은 쿠데타의 실패에 강한 긍정적 레시피인 동시에 강한 부정적 레시피가 된다. 따라서 이 경우 선언 설명(존재양화 설명)과 선언지 설명이 모두 좋은 설명이 된다. 앞서 각주7에서 밝힌 바 있듯이 선언적 설명3과 같은 존재양화 설명에서 선언 설명과 선언지 설명 가운데 어떤 설명이 더 만족스러운지는 문맥에 따라서 변할 수 있다. 예를 들어 쿠데타의 역사와 실패 또는 성공 방식에 대한 이론을 배우는 수업의 경우 선언 설명은 선언지 설명보다 더 만족스러울 수 있다. 왜냐하면 이 경우 선언 설명은 존재양화된 보다 추상화된 개념을 사용하기 때문에 보다 통일적인 설명을 제공해 줄 수 있기 때문이다. 보다 통일적인 설명은 보다 광범위하게 유사한 상황에서 레시피를 제공해 줄 수 있으며 따라서 그런 레시피가 필요한 상황에서 선호된다. 그러나 선언지 설명은 선언 설명보다 일반적으로 더 쉬운 레시피를 제공해 준다. 철수, 영수, 명수, 준수 가운데 누구도 밀고하지 않게 만드는 것은 영수가 밀고하지 않게 만드는 것보다 어렵다. 4명 모두를 감시하는 것보다는 영수만 감시하는 것이 더 쉽기 때문이다. 누가 그 쿠데타가 실패하는 것을 어떻게 막을 수 있었냐고 질문했을 때 우리가 제공할 수 있는 가장 간단한 레시피는 영수가 밀고하지 못하게 했어야 한다는 것일 것이다. 따라서 통일적인 레시피보다 쉬운 레시피가 더 필요한 맥락에서 존재양화된 선언적 설명보다는 선언지 설명이 더 만족스러울 수 있다. 레시피 개념에 기반한 필자의 이론

은 설명적 힘의 문맥 의존성에 매우 유연한 방식으로 대응할 수 있다.

8. 결론

2절에서의 논의에서 드러났듯이 우리의 선언적 설명에 대한 직관을 설명하는데 있어서 다음의 두가지가 핵심적이다.

- 왜 스위치 구조가 있는 경우 선언 설명은 성립하는데 선언지 설명은 결코 성립하지 않는가?
- 왜 종종 선언 설명과 선언지 설명은 모두 성립하는가? 특히 왜 선취나 과잉결정의 상황에서 반사실적 의존성 없는 선언지가 설명력을 갖는가?

현대 설명이론의 양대 라이벌인 인과적 설명이론과 통일 이론은 모두 이 둘을 동시에 설명하지 못한다. 인과적 설명이론은, 그것이 순수하고 기본적인 형태로 제시되었을 때, 전자를 설명하는데 난점을 갖는다. 인과적 설명 이론의 정신을 충실히 따랐을 경우 우리는 항상 선언적 설명보다 선언지 설명을 선호해야 하며 그런 점에서 인과적 설명 이론은 기본적으로 선언적 설명과 양립하기 어려운 이론이다.

반면에 설명에 대한 통일 접근은 후자의 직관을 설명하는데 난점을 갖는다. 통일 접근의 경우 추상화된 개념을 사용하는 스위치 구조를 갖는 선언 설명을 정당화하는 데는 문제를 갖지 않지만 추상화되지 않고 투사 가능하지 않은 술어를 사용하는 선언 설명을 정당화하는데 문제를 갖는다. 또 통일 접근은 그것이 갖는 승자독식적 성격 때문에 (선언적 설명3의 경우처럼) 추상화된 개념을 사용하는 선언 설명과 선언지 설명이 모두 성립하는 경우를 정당화하는데 문제를 갖는다.

D-N 모델과 반사실적 의존성 이론은 모두 선언적 설명에 대해 친화적이다. 그러나 이들은 허용되지 말아야 할 선언적 설명마저 포용하는 문제를 갖는다. 그리고 이 문제를 해결하기 위해 제안된 D-N 모델의 수정된 형태

들은 목록물과 함께 아이를 버려 버리는 우를 범한다. 선언적 설명 자체를 불가능하게 만들어 버리기 때문이다. 반사실적 의존성 이론의 경우 무관한 선언지를 매우 쉽게 허용하는 문제를 가지며, (선언적 설명1의 경우처럼) 선언과 함께 선언지도 설명력을 갖는 경우에서 우리의 직관을 제대로 설명 못하는 문제 역시 갖는다.

이에 반해 필자가 옹호하는 새로운 조작주의적 설명 이론은 선언적 설명에 대해서 우리가 갖는 다양한 직관을 체계적으로 설명할 수 있는 힘을 갖는다. 이 이론은 왜 스위치 구조를 갖는 선언적 설명의 경우 선언은 설명력을 갖는데 선언지는 설명력을 갖지 않는지를 설명하는 동시에 왜 선취나 과잉결정 상황에서 선언과 선언지가 모두 설명력을 갖는지를 설명할 수 있다. 이 이론은 왜 무관한 선언지가 설명에서 용인될 수 없는지도 설명할 수 있다. 예를 들어, 철수에게 돌을 던지게 하는 것은 꽃병을 깨는데 적합한 긍정적 레시피가 될 수 있지만 철수가 돌을 던지거나 영수가 제자리 걷기를 하는 것은 꽃병을 깨는데 적합한 긍정적 레시피가 될 수 없다. 영수가 제자리 걷기를 하도록 시키는 것을 통해서 ‘철수가 돌을 던지거나 영수가 제자리 걷기를 하는’ 선언적 사태를 발생시킬 수는 있지만 그렇다고 해서 꽃병이 깨지는 것은 아니기 때문이다. 선언이 긍정적인 레시피가 되기 위해서는 모든 선언지가 피설명항에 대해 유관해야 한다.

이 모두를 고려할 경우 필자의 이론은 지금까지 제시된 어떤 주요 설명 이론보다 선언적 설명에 대한 우리의 직관을 잘 설명할 수 있는 이론이다. 레시피 개념은 인과 개념과 밀접한 관계를 가지지만 인과 개념과는 달리 애초부터 선언에 친화적인 개념이며, 그것이 이론적으로 정교화 되었을 때 우리의 선언적 설명에 관한 직관을 원리 지워지고 체계적인 방식으로 설명할 수 있다. 따라서 우리가 선언적 설명을 부정할 수 없다면 우리는 필자의 이론을 받아들일 좋은 동기를 갖게 된다.

참고문헌

- Clark, R. (2003), *Libertarian Accounts of Free Will*, Oxford University Press.
- Clarke, R. and Clarke, R. K. (2005), *Libertarian Accounts of Free Will*, Oxford University Press.
- Eberle, R., Kaplan, D. and Montague, R. (1961), “Hempel and Oppenheim on Explanation”, *Philosophy of Science* 28(4): pp. 418-28.
- Hall, E. J. (2004), “Two Concepts of Causation”, in Collins, J. D., Hall, E. J. and Paul, L. A. (eds.), *Causation and Counterfactuals*, Cambridge, Mass: MIT Press, pp. 225-76.
- Kaplan, D. (1961), “Explanation Revisited”, *Philosophy of Science* 28, pp. 429-36.
- Kim, J. (1999), “Hempel, Explanation, Metaphysics”, *Philosophical Studies* 94(1-2): pp. 1-20.
- Kitcher, P. (1989), “Explanatory Unification and the Causal Structure of the World”, in Kitcher, P. and Salmon, W. C. (eds.), *Scientific Explanation*, Minneapolis: University of Minnesota Press, pp. 410-505.
- Kitcher, P. and Salmon, W. (1987), “Van Fraassen on Explanation”, *Journal of Philosophy* 84(6): pp. 315-30.
- Lewis, D. (1973), *Counterfactuals*, Oxford: Blackwell.
- _____ (1983), “New Work for a Theory of Universals”, *Australasian Journal of Philosophy* 61(4): pp. 343-77.
- _____ (1986a), “Causal Explanation”, in *Philosophical Papers* Vol. 2, Oxford: Oxford University Press, pp. 214-40.
- _____ (1986b), “Event”, in *Philosophical Papers* Vol. 2, Oxford: Oxford University Press, pp. 241-69.
- _____ (1986c), “Postscripts to ‘Causation’”, in *Philosophical Papers* Vol. 2, Oxford: Oxford University Press, pp. 172-213.
- O’Connor, T. (1995), “Agent Causation”, in O’Connor, T. (ed.), *Agents*,

- Causes, and Events: Essays on Indeterminism and Free Will*, New York: Oxford University Press, pp. 173-200.
- Salmon, W. C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton, N. J.: Princeton University Press.
- _____ (1989), “Four Decades of Scientific Explanation”, in Kitcher, P. and Salmon, W. C. (eds.), *Scientific Explanation*, Minneapolis: University of Minnesota Press, pp. xiv-528.
- _____ (1997), “Causality and Explanation: A Reply to Two Critiques”, *Philosophy of Science* 64(3): pp. 461-77.
- Skow, B. (2016), *Reasons Why*, Oxford University Press.
- Sober, E. (1983), “Equilibrium Explanation”, *Philosophical Studies* 43: pp. 201-10.
- Strevens, M. (2004), “The Causal and Unification Approaches to Explanation Unified—Causally”, *Nous* 38(1): pp. 154-76.
- _____ (2008), *Depth: An Account of Scientific Explanation*, Harvard University Press.
- Woodward, J. (2003), *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.
- 선우환 (2020), 『때문에—‘때문에’의 의미에 대한 철학적 연구』, 아카넷.
- _____ (2022), 「반사실적 조건문적 〈때문에〉 이론에 대한 옹호」, 『과학철학』 25권 3호, pp. 1-37.
- 이재호 (2022), 「완전한 설명 이론과 반사실적 의존성—선우환 교수의 반론에 대한 답변」, 『과학철학』 25권 3호, pp. 89-114.
- _____ (2023), 「신의 레시피로서의 설명—새로운 조작주의적 설명 이론」, 『과학철학』 26권 3호, pp. 1-37.
- 이재호, 박재희 (2024), 「인과적 설명 이론의 딜레마와 평형 설명」, 『과학철학』 27권 1호, pp. 1-37.
- 한성일 (2022), 「비술어화 사실에 대한 형이상학적 설명과 반사실적 의존성」, 『철학적 분석』, 8호, pp. 29-60.

논문 투고일	2024. 08. 08.
심사 완료일	2024. 10. 02.
게재 확정일	2024. 10. 02.

What Should We Think about Disjunctive Explanations?

Jaeho Lee

In this paper, I first argue that we need to accept the existence of disjunctive explanation even if we do not accept the existence of disjunctive causation. Next, I argue that none of the major existing theories of explanation, namely the D-N model, the causal theories explanation, the unification account of explanation, and the counterfactual dependence theory of explanation, can properly explain our intuitions about disjunctive explanations. Finally, I argue that my theory of explanation, unlike the existing theories, can provide a systematic explanation of our intuitions about disjunctive explanations in a principled way.

Keywords: Explanation, Disjunction, Equilibrium explanation,
Manipulation, Recipe