creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

문학석사 학위논문

자유의 본성에 관한 넓은 원천 양립불가론과 외재적 양방력 이론

2022년 2월

서울대학교 대학원

철학과 서양철학전공

조 민 수

자유의 본성에 관한 넓은 원천
양립불가론과 외재적 양방력 이론

지도 교수 한 성 일

이 논문을 문학석사 학위논문으로 제출함
2021년 10월

서울대학교 대학원
철학과 서양철학전공
조 민 수

조민수의 문학석사 학위논문을 인준함
2022년 1월

위 원 장 _____ 김 기 현 _____ (인)

부위원장 _____ 한 성 일 _____ (인)

위 원 _____ 강 상 진 _____ (인)

초 록

원천 양립불가론에 의하면 행위자는 오직 자신이 수행한 행위의 원천인 경우에 한하여 그것에 대해 도덕적 책임을 지며 문제의 원천-관계는 인과적 결정론이 참인 세계에서 성립할 수 없다. 그중 넓은 원천 양립불가론자들은 인과적 결정론이 참인 세계에서 문제의 원천-관계가 성립할 수 없는 이유를 다음과 같이 설명한다: 행위자는 오직 달리 행할 수 있었던 경우에 한하여 자신이 수행한 행위의 원천일 수 있지만 인과적 결정론이 참인 세계에서는 자신의 어떠한 행위에 대해서도 그것을 수행함에 있어서 달리 행할 수 없다.

어떤 넓은 원천 양립불가론자들은 오직 달리 행할 수 있었던 경우에 한하여 행위자가 자신이 수행한 행위의 원천일 수 있는 이유를 다음과 같이 설명한다: 그것은 문제의 행위가 행위자의 스스로-결정하는 힘의 사례인 경우에 한하여 행위자가 그것의 원천일 수 있으며 스스로-결정하는 힘의 본성 상 그것의 모든 사례는 항상 대안가능성을 동반하기 때문이다. 그중 양방력 이론가들은 특히 이러한 독특한 특징을 갖는 스스로-결정하는 힘을 일종의 양방력, 즉 무엇인가를 할 수도 있지만 하지 않을 수도 있는 힘으로 이해한다. 양방력 이론가들에 의하면 스스로-결정하는 힘의 모든 사례들이 항상 대안가능성을 동반하게 되는 이유는 다름아닌 스스로-결정하는 힘이 무엇인가를 할 수도 있지만 하지 않을 수도 있는 힘, 즉 양방력이기 때문이다. 그렇다면 어떤 힘이 무엇인가를 할 수도 있지만 하지 않을 수도 있는 힘이라는 것이 정확히 무슨 뜻인가? 외재적 양방력 이론가들에 의하면 이는 다음을 의미한다: 문제의 힘은 발현-조건이 만족되었다 하더라도 여전히 발현되지 않을 수 있는 힘이다. 달리 말해 외재적 양방력은 발현의 충분조건으로서의 발현-조건을 결여하는 힘이다.

하지만 넓은 원천 양립불가론자의 입장에서 양방력으로 이해된 스스로-결정하는 힘을 다시금 외재적 양방력의 일종으로 이해하기는

어려워 보인다. 첫째, 만약 넓은 원천 양립불가론자가 양방력으로 이해된 스스로-결정하는 힘을 다시금 외재적 양방력의 일종으로 이해할 경우 피셔가 프랑크푸르트-유형 사례에 근거하여 제기한 탄탄성 비판에 대한 어떤 특정한 종류의 답변을 더이상 이용할 수 없다. 물론 외재적 양방력 이론가들이 이용할 수 있는 답변이 여전히 남아있기는 하지만 이러한 답변은 분명한 한계를 가지고 있다. 따라서 이러한 사실은 최소한 양방력 이론을 받아들이는 몇몇 넓은 원천 양립불가론자들에게 있어서는 외재적 양방력 이론을 거부할 좋은 이유를 제공해줄 것이다. 둘째, 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘은 본성 상 외재적 양방력과 동일시될 수 없는 것처럼 보인다. 왜냐하면 스스로-결정하는 힘이 외재적 양방력의 일종이라는 주장으로부터 다음과 같은 문제적인 귀결이 따라 나오기 때문이다: 행위자가 주어진 행위를 수행할지 말지 여부를 스스로-결정한 경우가 행위자가 주어진 행위를 수행할지 말지 여부를 스스로-결정하는 힘을 행사하지 않은 경우에 해당할 수 있다. 셋째, 들뜬 원자가 가지고 있는 광자를 방출하는 즉자적 힘 또한 발현의 충분조건으로서의 발현-조건을 결여하는 힘인 것처럼 보인다. 따라서 외재적 양방력 이론이 옳다면 이러한 즉자적 힘과 자유 및 도덕적 책임의 근거를 제공하는 스스로-결정하는 힘은 적절히 구분되지 않는다. 이는 자유로운 행위자가 자신의 행위에 대해 가지고 있는 통제력이 기껏해야 들뜬 원자가 광자를 방출할지 말지 여부에 대해 가지고 있었던 통제력과 다를 바 없다는 문제적 귀결을 낳는 것처럼 보인다. 더불어 이러한 문제는 어떤 힘이 외재적 양방력이기 위한 또다른 필요조건을 추가한다 하더라도 해결되기 어려워 보인다. 이러한 사안들을 종합적으로 고려할 때 외재적 양방력 이론은 넓은 원천 양립불가론적 스스로-결정하는 힘의 본성에 대한 만족스러운 이론이기 힘들다.

주요어 : 자유, 도덕적 책임, 스스로-결정하는 힘, 넓은 원천 양립불가론, 양방력 이론, 외재적 양방력 이론

학　번 : 2019-26553

목 차

제 1 장 서 론: 양방력 이론이란 무엇인가? 1

제 2 장 힘과 관련된 용어들 11

제 3 장 외재적 양방력 이론이란 무엇인가? 18

제 4 장 프랑크푸르트-유형 사례 문제 22

제 5 장 스스로-결정하는 힘과의 본성적 긴장 67

제 6 장 즉자적 힘과의 구분 문제 77

제 7 장 결 론: 내재적 양방력 이론을 향하여 93

참고문헌 96

Abstract 101

제 1 장 서론: 양방력 이론이란 무엇인가?

우리에게 일어나는 다양한 일들 중 어떤 것은 행위(action)에 해당하지만 어떤 것은 그렇지 않다. 예를 들어 뇌전증(epilepsy) 환자의 발작(seizure)은 그 환자에게 일어난 일이지만 행위라고 볼 수는 없다. 마찬가지로 하지직거상 검사(straight leg raise test)를 실시하는 정형외과 의사에 의해 나의 다리가 올라갔다 하더라도 이는 나의 행위가 아니다. 더불어 우리의 행위들 중 어떤 것은 자유로운 행위(free action)이지만 어떤 것은 그렇지 않다. 예를 들어 강박장애(obsessive-compulsive disorder)의 증상으로 나타난 강박 행동(compulsion)은 일견 행위로는 간주될 수 있을지 몰라도 자유로운 행위라고 하기는 어려운 것처럼 보인다. 마찬가지로 금단 증상을 견디지 못하고 마약을 사용한 마약 중독자의 행위 또한 일견 자유로운 행위라고 보기는 어려운 것처럼 보인다. 반면 홍차를 우리기 위해 물을 끓이는 행위 혹은 카페에서 커피를 한 잔 사기 위해 지갑을 꺼내는 행위는 분명 자유로운 행위이다.

이렇게 우리에게 일어나는 다양한 일들 가운데 인간적 행위자성(human agency)이 전형적으로 드러나는 사례는 무엇보다도 자유로운 행위인 것처럼 보인다. 또한 자유로운 행위는 다시금 그것을 수행한 행위자에게 도덕적 책임이 귀속된다는 점에서 특별한 것처럼 보인다. 하지만 자유와 도덕적 책임의 본성은 아직 신비스러운 채로 남아 있다. 자유로운 행위란 도대체 무엇이며 우리는 왜 행위자에게 이에 대한 도덕적 책임을 묻는가?

어떤 철학자들은 자유와 도덕적 책임의 본성을 최소한 부분적으로는 다음과 같은 원리를 통해 특징지을 수 있다고 주장한다: 행위자 S는 오직 달리 행할 수 있었던 경우에 한하여 자신이 수행한 행위 A에 대해 도덕적 책임을 진다. 프랑크푸르트(Frankfurt, 1969)를 따라 이러한 논제를 '대안가능성 원리(principle of alternate possibilities)'라고 부르도록 하자.

대안가능성 원리는 일견 다음과 같은 귀결을 갖는 것처럼 보인다: 만약 인과적 결정론(causal determinism)이 참이라면 자유 및 도덕적

책임은 있을 수 없다. 왜냐하면 인과적 결정론이 참인 세계에서는 어떠한 대안가능성도 있을 수 없는 것처럼 보이기 때문이다. 인과적 결정론은 대안가능성의 부재를 함축하는 것처럼 보이며 대안가능성의 부재는 다시금 대안가능성 원리와 함께 도덕적 책임의 부재를 함축하는 것처럼 보인다. 따라서 인과적 결정론이 참인 세계에서 자유 및 도덕적 책임은 있을 수 없는 것처럼 보인다. 이렇게 인과적 결정론이 참인 세계에서 자유 및 도덕적 책임은 있을 수 없다는 입장 혹은 인과적 결정론은 자유 및 도덕적 책임과 양립불가능하다는 입장을 '양립불가론(incompatibilism)'이라고 부르도록 하자.

하지만 어떤 철학자들은 설사 인과적 결정론이 참이라 하더라도 여전히 자유 및 도덕적 책임이 있을 수 있다고 생각한다. 즉 그들은 양립불가론이 틀렸다고 생각한다. 이렇게 인과적 결정론이 참인 세계에서도 여전히 자유 및 도덕적 책임이 있을 수 있다는 입장 혹은 인과적 결정론은 자유 및 도덕적 책임과 양립가능하다는 입장을 '양립가능론(compatibilism)'이라고 부르도록 하자. 양립가능론을 성공적으로 옹호하기 위해서는 앞서 제시된 양립불가론에 대한 논변, 즉 자유 및 도덕적 책임은 대안가능성을 요구하며 (달리 말해, 대안가능성 원리가 참이며) 인과적 결정론이 참인 세계에서 대안가능성은 있을 수 없기 때문에 결국 인과적 결정론이 참인 세계에서 자유 및 도덕적 책임은 있을 수 없다는 논변이 정확히 어떤 문제를 가지고 있는지를 보여주어야 한다.

어떤 양립가능론자들은 개념 〈대안가능성〉에 대한 보다 섬세한 분석을 통해 인과적 결정론이 참인 세계에서도 여전히 어떤 종류의 대안가능성이 있을 수 있음을 보여주려고 노력한다. 이러한 방식으로 양립가능론을 옹호하려는 시도를 오코너와 프랭클린(O'Connor & Franklin, 2020)은 '고전적 양립가능론(classical compatibilism)'이라고 부른다. 하지만 고전적 양립가능론의 기획이 성공적인지 여부는 분명하지 않다. 왜냐하면 설사 인과적 결정론이 참인 세계에서 여전히 "어떤 종류의" 대안가능성이 있을 수 있다 하더라도 과연 문제의 대안가능성이 자유 및 도덕적 책임의 필요조건으로 요구되는 바로

그러한 종류의 대안가능성에 해당하는지 여부는 분명하지 않기 때문이다. 오코너와 프랭클린(O'Connor & Franklin, 2020, pp. 18-26)은 최소한 아직까지 전자와 후자를 모두 보이는데 성공한 고전적 양립가능론자는 없는 것 같다고 평가한다.

한편 어떤 양립가능론자들은 대안가능성 원리 자체를 거부한다. 그들은 달리 행할 수 없었던 행위에 대해서도 여전히 도덕적 책임을 물을 수 있다고 주장한다. 그들에 의하면 설사 인과적 결정론이 참인 세계에서 어떠한 대안가능성도 존재할 수 없다 하더라도 이는 양립불가론을 받아들일 좋은 이유를 제공해주지 못한다. 왜냐하면 애초에 자유 및 도덕적 책임은 대안가능성을 필요로 하지 않기 때문이다. 이러한 입장을 취하는 이들을 일반적으로 '원천 양립가능론자(source compatibilist)'라고 부른다. 왜냐하면 이들은 일반적으로 자유 및 도덕적 책임의 본성을 대안가능성 원리 대신 다음과 같은 원리를 통해 특징짓기 때문이다: 행위자 S는 오직 자신이 수행한 행위 A의 원천인 한에서 A에 대해 도덕적 책임을 진다. 이를 '원천 원리(sourcehood principle)'라고 부르도록 하자. 물론 원천 양립가능론자들이 여기서 요구되는 원천-관계의 본성을 이해하는 방식은 다양하다.[1] 하지만 그들은 모두 다음과 같은 주장을 받아들인다는 점에 있어서는 의견의 일치를 보인다: 행위자 S와 S에 의해 자유롭게 수행된 행위 A 사이에 원천-관계가 성립하기 위해 인과적 결정론이 거짓일 필요는 없다. 즉 원천-관계는 인과적 결정론과 양립가능하다.

하지만 어떤 양립불가론자들은 원천 원리를 받아들인다는 점에서는 원천 양립가능론자들이 옳았지만 원천-관계가 성립하기 위해 인과적 결정론이 거짓일 필요는 없다고 주장한다는 점에 있어서는 원천

[1] 원천-관계의 본성에 대한 대표적인 양립가능론적 이해 방식으로는 동일시 모형(identification model)과 이유-반응성 모형(reason-responsiveness model)이 있다 (O'Connor & Franklin, 2020). 프랑크푸르트(Frankfurt, 1971)는 대표적인 동일시 모형의 옹호자이며 피셔(Fischer, 2006, pp. 63-83)와 맥케나(McKenna, 2013)는 대표적인 이유-반응성 모형의 옹호자이다.

양립가능론자들이 잘못되었다고 주장한다. 그들은 다음과 같이 주장한다: 인과적 결정론이 참인 세계에서 행위자는 자신이 수행한 어떠한 행위에 대해서도 그것의 원천일 수 없다. 즉 원천-관계가 성립하기 위해서는 인과적 결정론이 거짓이여야 한다. 이렇게 인과적 결정론과 원천-관계가 양립불가능하다고 주장하는 입장을 팀페(Timpe, 2007)를 따라 '원천 양립불가론(source incompatibilism)'이라고 부르도록 하자. 원천 양립불가론자들은 원천 원리를 받아들인다는 점에 있어서는 원천 양립가능론자들과 의견을 같이 하지만 원천-관계가 인과적 결정론과 양립불가능하다고 생각한다는 점에 있어서는 원천 양립가능론자들과 의견을 달리 한다.

그렇다면 원천 양립가능론자들과 원천 양립불가론자들 사이에 가장 중요한 쟁점은 과연 인과적 결정론이 참인 세계에서 원천-관계가 성립할 수 있는지 여부일 것이다. 따라서 원천 양립불가론자들은 자신들이 왜 인과적 결정론과 원천-관계가 양립불가능하다고 생각하는지를 설명해줄 필요가 있다. 어떤 원천 양립불가론자들은 원천-관계가 성립하기 위해서는 대안가능성이 존재해야 한다고 생각한다. 즉 그들은 다음과 같은 생각을 받아들인다: 행위자 S는 오직 행위 A를 수행함에 있어서 달리 행할 수 있었던 경우에 한하여 자신이 수행한 행위 A의 원천일 수 있다. 만약 이러한 생각이 옳다면, 원천 양립가능론자들처럼 대안가능성 원리는 거부하면서 원천 원리는 받아들이는 것은 애초에 불가능할 것이다. 왜냐하면 대안가능성 원리는 이제 원천 원리의 귀결로서 따라 나오는 원리일 것이기 때문이다. 팀페(Timpe, 2007)를 따라 이러한 종류의 원천 양립불가론을 '넓은 원천 양립불가론(wide source incompatibilism)'이라 부르도록 하자.[2]

---

[2] 물론 모든 원천 양립불가론자들이 넓은 원천 양립불가론을 받아들이는 것은 아니다. 어떤 원천 양립불가론자들은 행위자가 자신이 수행한 행위의 원천이기 위해 꼭 어떤 대안가능성이 존재해야 하는 것은 아니라고 생각한다. 팀페(Timpe, 2007)는 이러한 입장을 '좁은 원천 양립불가론(narrow source incompatibilism)'이라 부른다. 대표적인 좁은 원천 양립불가론자로는 스텀프(Stump, 2003)가 있다. 물론 맥케나(McKenna, 2001)와 같이 표현 '원천 양립불가론'을 오직 좁은 원천 양립불가론만을 의미하는 용어로 사용하는 철학자들

팀페(Timpe, 2007)가 자신이 이러한 용어법을 채택한 이유에 대해 특별히 명시적으로 설명한 바는 없다. 하지만 우리는 팀페가 문제의 원천 양립불가론을 '넓은 원천 양립불가론'이라 부르는 이유를 다음과 같이 추정해볼 수 있다: 넓은 원천 양립불가론자들이 행위자와 자유롭게 수행된 행위 사이에 성립하는 원천-관계를 이해하는 방식에 따르면, 문제의 원천-관계는, 만약 그것이 성립할 경우에, 어떤 대안가능성의 존재를 함축한다는 점에서 보다 "넓은" 함축을 갖는다.[3]

넓은 원천 양립불가론자들은 앞서 설명된 것처럼 [행위자 S는 자신이 수행한 행위 A에 대해 도덕적 책임을 진다][4]는 [행위자 S는 자신이 수행한 행위 A의 원천이다]에 근거하며 후자의 사실이 성립하기 위해서는 다시금 [행위자 S는 A를 수행함에 있어서 달리 행할 수 있었다]가 성립해야 한다고 주장한다. 하지만 왜 [행위자 S는 자신이 수행한 행위 A의 원천이다]가 성립하기 위해 [행위자 S는 A를 수행함에 있어서 달리 행할 수 있었다]가 성립해야 하는가? 이에 대한 한가지 가능한 답변은 다음과 같다: 그것은 [행위자 S는 자신이 수행한 행위 A의 원천이다]가 무엇보다도 [A는 S가 소유하던 스스로-결정하는 힘(self-determining power)의 사례[5]이다]에 근거하며 후자의 사실은 다시금 [S는 A를 수행함에 있어서 달리 행할 수 있었다]를 함축하기 때문이다.[6] 하지만 이러한 답변이 성공적이기 위해서는 다음과 같은

또한 존재한다. 하지만 본 논문에서는 팀페(Timpe, 2007)의 용법을 따라 '원천 양립불가론'을 일단 넓은 원천 양립불가론과 좁은 원천 양립불가론 모두를 포괄하는 용어로 사용하고 후자가 아니라 전자, 즉 좁은 원천 양립불가론이 아니라 넓은 원천 양립불가론에 집중하여 논의를 전개할 것이다.

[3] 이러한 추정은 한성일 교수님께서 제안해주신 것이다.

[4] 임의의 명제 P에 대해 '[P]'는 P에 의해 표상되는 사실을 의미한다. 예를 들어 '[철수는 도서관에 갔다]'는 철수가 도서관에 갔다는 사실을 의미한다.

[5] 행위자 S에 의해 수행된 행위 A가 S가 소유하던 어떤 힘 P의 사례라는 것은 다음과 같이 이해된다: A는 (i) P의 행사이거나, (ii) P가 행사된 결과 이거나 혹은 (iii) P가 행사된 귀결이다. P의 행사, P가 행사된 결과 그리고 P가 행사된 귀결 사이의 구분과 관련해서는 알바레즈와 하이먼(Alvarez & Hyman, 1998)을 참조하라.

[6] 물론 원천-관계가 왜 대안가능성의 존재를 함축하는지를 묻는 질문에 대해 모든 넓은 원천 양립불가론자가 이러한 방식으로 답변하는 것은 아니다. 페러

질문에 대한 답변 또한 만족스럽게 주어질 수 있어야 한다: 도대체 스스로-결정하는 힘은 어떤 독특한 종류의 힘이길래 문제의 힘의 사례들은 그것이 일어남에 있어서 항상 대안가능성을 동반하는가?

최근 일련의 철학자들에 의해 (원천-관계의 근거를 제공함으로써) 자유 및 도덕적 책임의 근거를 제공하는 스스로-결정하는 힘의 본성은 단방력(one-way power)이 아니라 양방력(two-way power)으로 이해되어야 한다는 주장이 제시된 바 있다. (이하 '양방력 이론(two-way power theory)')[7] 어떤 힘이 단방력이라는 것은 곧 해당 힘이 행사될 수 있는 조건이 만족되었을 때 일어날 수 있는 일이 오직

붐(Pereboom, 2003)은 원천-관계가 성립한다는 사실은 행위자가 문제의 행위를 수행하도록 행위자의 통제를 벗어난 요소들에 의해 인과적으로 결정되어 있지 않았다는 사실에 근거하며 후자의 사실은 다시금 대안가능성의 존재를 함축한다고 주장한다. 하지만 페러붐의 주장 또한 여전히 다음과 같은 함축을 갖는다: 원천-관계는 대안가능성의 존재를 함축한다.

[7] 다음 목록에 해당하는 철학자들의 양방력 이론들은 모두 넓은 원천 양립불가론의 일종으로 발전될 여지가 있는 자유 및 도덕적 책임에 관한 이론들이다: 스튜어드(Steward, 2012a, 2013, 2020, 2021), 알바레즈(Alvarez, 2013), 프로스트(Frost, 2013, 2020), 로우(Lowe, 2013). (핀크(Pink, 2004, 2011, 2016)의 경우 비록 자유 및 도덕적 책임에 대한 그의 이론이 양방력 이론의 일종으로 이해될 수 있을 것인지 여부가 다소 분명하지 않은 측면이 있지만 만약 우리가 그의 이론을 여전히 양방력 이론이라고 부를 수 있다면 그의 양방력 이론 또한 넓은 원천 양립불가론의 일종으로 발전될 여지가 있는 자유 및 도덕적 책임에 관한 이론이라고 말할 수 있다.) 특별히 스튜어드(Steward, 2021)는 명시적으로 자신의 양방력 이론을 자유 및 도덕적 책임에 관한 넓은 원천 양립불가론의 일종으로 발전시키고 있다. 물론 (본 논문에서는 오직 양립불가론의 일종으로 이해된 양방력 이론만이 고려될 예정이지만) 몇몇 양방력 이론가들은 자신의 양방력 이론이 그 자체로 양립불가론을 함축한다고는 생각하지 않는다 (Alvarez, 2013, pp. 111-112; Frost, 2013, 2020, pp. 1141의 각주 2). 하지만 그들 중 양방력 이론이 그 자체로 양립가능론을 함축한다고 생각하는 이 또한 아무도 없는 것처럼 보인다. 따라서 우리는 지금까지 제시되어 온 대부분의 양방력 이론들은 (그리고 최소한 앞서 언급된 목록에 속한 철학자들의 양방력 이론들은 모두) 넓은 원천 양립불가론의 일종으로 충분히 발전될 여지가 있는 자유 및 도덕적 책임에 관한 이론들이라고 무리없이 말할 수 있을 것이다. 양방력 이론이 양립가능론 및 양립불가론과 관련하여 정확히 어떤 관계를 맺고 있는지에 대해 보다 분명히 설명해야 할 필요성은 김기현 교수님께서 지적해주신 것이다.

한가지 뿐임을 의미한다. 예를 들어 물이 가지고 있는 염화나트륨(sodium chloride)을 녹일 수 있는 힘의 경우, 문제의 힘이 행사될 수 있는 조건이 만족되었을 때 일어날 수 있는 일은 오직 한가지, 즉 **염화나트륨이 녹음** 뿐이다.[8] 반면 어떤 힘이 양방력이라는 것은 곧 해당 힘이 행사될 수 있는 조건이 만족되었을 때 일어날 수 있는 일이 다음과 같은 의미에서 한가지가 아니라 두가지임을 의미한다: 어떤 일이 일어날 수도 있었지만 일어나지 않을 수도 있었다. 예를 들어 양방력 이론가들이 주장하는 바에 의하면 방금 전 홍차를 마실 때 내가 행사한 스스로-결정하는 힘은 다음과 같은 의미에서 양방력이었다: 방금 전 홍차를 마실 때 나는 홍차를 마실 수 있었던 것 만큼이나 홍차를 마시지 않을 수도 있었다. 만약 단방력과 양방력을 이러한 방식으로 이해할 경우 넓은 원천 양립불가론자들이 스스로-결정하는 힘의 본성을 단방력이 아니라 양방력으로 이해하는 것은 몹시 자연스러운 것처럼 보인다. 왜냐하면 만약 스스로-결정하는 힘이 양방력이라면 해당 힘이 행사될 수 있는 조건이 만족되었을 경우 일어날 수 있는 일은 한가지가 아니라 두가지일 것이며, 따라서 문제의 힘의 사례들은 항상 대안가능성을 동반할 것이기 때문이다. 예를 들어 나는 방금 전 홍차를 마셨지만 이때 만약 내가 일종의 양방력으로서의 스스로-결정하는 힘을 행사한 것이라면 바로 이 사실에 근거하여 나는 주어진 상황에서 홍차를 마실 수 있었던 것 만큼이나 홍차를 마시지 않을 수도 있었다고 말할 수 있다.

따라서, 양방력 이론을 받아들이는 넓은 원천 양립불가론자들에 의하면, 스스로-결정하는 힘의 사례들이 그것이 일어남에 있어서 항상 대안가능성을 동반하는 이유는 다름아닌 문제의 힘이 어떤 행위 A에 대해 A를 수행하거나 수행하지 않을 수 있는 힘, 즉 양방력이기 때문이다. 하지만 양방력의 본성은 다시금 어떻게 이해되어야 하는가? 이러한 질문이 제기되는 이유는 어떤 행위 A에 대해 A를 수행하거나 수행하지 않을 수 있는 힘 혹은 바로 이러한 의미에서 그것이 행사될

[8] 이는 스튜어드(Steward, 2021, p. 11)에 의해 제시된 사례이다.

수 있는 조건이 만족되었을 때 일어날 수 있는 일이 한가지가 아니라 두가지인 힘의 본성은 (최소한) 서로 다른 두가지 방식으로 이해될 수 있으며, 문제의 힘의 본성을 이해하는 방식에 있어서의 이러한 차이는 다시금 넓은 원천 양립불가론이 겪게 되는 문제들의 종류와 성격, 그리고 해당 문제들의 해결가능성에 있어서 중요한 차이를 낳게 되기 때문이다.[9] 더 나아가 아직까지 (명시적으로든 혹은 암묵적으로든) 스스로를 양방력 이론가로 간주하는 이들 사이에 양방력의 본성을 이해하는 합의된 방식이 존재하지 않는다는 사실[10]은 우리가 아직까지 양방력의 본성에 대해 충분히 만족스러운 이해에 도달하지 못하였음을 시사한다. 이렇게 양방력의 본성에 대한 이해를 결여한 채로 스스로-결정하는 힘의 본성을 개념 〈양방력〉에 호소하여 설명하는 것은 자유 및 도덕적 책임 그리고 그것의 근거를 이루는 원천-관계의 본성을 해명하는데 있어서 별다른 도움이 되지 못할 것이다. 본 논문은 이러한 맥락에서 다음과 같은 질문에 대해 부분적으로 답변하는 것을 목표로 한다: 넓은 원천 양립불가론, 그 중에서도 스스로-결정하는 힘의 본성에 호소하여 원천-관계의 본성을 해명하려 하는 종류의 넓은 원천 양립불가론[11]이 자유 및 도덕적 책임의 본성에 대한 올바른 이론이라고

---

[9] 비록 본 논문에서 주목하려고 하는 문제들과 직접적으로 관련되어 있는 것은 아니지만 프로스트(Frost, 2020) 또한 우리가 만약 그가 제안하는 방식을 따라 양방력의 본성을 이해하지 않을 경우 양방력과 관련하여 발생하는 어떤 특정한 종류의 딜레마로부터 벗어날 수 없게 됨을 논변하면서 이와 (정확히 일치하는 것은 아니지만) 유사한 종류의 우려를 표현한 바 있다. (물론 이때 그가 본 논문에서와 같이 양방력 이론을 "넓은 원천 양립불가론적" 스스로-결정하는 힘의 본성에 관한 이론으로서 고려하고 있는 것은 아니다.)

[10] 예를 들어 스튜어드 (Steward, 2012a, 2013, 2020, 2021)와 알바레즈 (Alvarez, 2013)는 양방력의 본성을 다음과 같이 이해한다: 양방력은 발현의 충분조건으로서의 발현-조건을 결여하는 힘이다. 반면 프로스트(Frost, 2013, 2020)와 로우(Lowe, 2013)는 양방력의 본성을 다음과 같이 이해한다: 양방력은 서로 다른 두가지 발현-유형에 해당하는 방식으로 발현될 수 있는 힘이다. 여기서 사용된 용어 '발현-조건'과 '발현-유형'의 의미는 본 논문의 2장에서 보다 자세히 설명될 것이다.

[11] 본 논문에서는 앞으로 표현 '넓은 원천 양립불가론'을 특별한 별도의 표기가 없다면 오직 스스로-결정하는 힘의 본성에 호소하여 원천-관계의 본성을 해명

가정할 경우, 만약 문제의 스스로-결정하는 힘이 양방력의 일종이라면 이때 양방력의 본성은 어떻게 이해되어야 하는가? 즉 본 논문은 스스로-결정하는 힘의 본성에 호소하여 원천-관계의 본성을 해명하려 하는 넓은 원천 양립불가론자의 입장에서 양방력의 본성을 이해하는 가장 만족스러운 방식은 과연 무엇일지의 문제를 탐구한다. 그리고 이러한 질문에 대해 본 논문에서 궁극적으로 제시하려고 하는 답변은 다음과 같다: 최소한 외재적 양방력 이론[12] 은 넓은 원천 양립불가론자의 입장에서 양방력의 본성을 이해하는 만족스러운 방식이기는 어렵다.

이를 보이기 위해 먼저 나는 본격적인 논의에 앞서 힘의 본성을 특징짓는데 사용되는 몇가지 용어들의 의미를 분명히 한 후 (2장) 양방력의 본성을 개념화하는 한가지 유력한 방식 (본 논문에서는 양방력에 대한 이러한 개념화 방식을 '외재적 양방력 이론'이라 부를 것이다)을 소개할 것이다. (3장) 하지만 내 생각에 외재적 양방력 이론이 스스로-결정하는 힘의 본성에 호소하여 원천-관계의 본성을 해명하려 하는 넓은 원천 양립불가론자의 입장에서 양방력의 본성을 이해하는 만족스러운 방식이기는 어려워 보인다. 일단 넓은 원천 양립불가론자들이 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 채택할 경우 프랑크푸르트-유형 사례[13]로부터 제기되는 문제에 답변함에 있어서 한가지 불필요한 제약을 받아들여야 한다. 물론 이러한 제약 하에서도 어떤 종류의 답변은 여전히 이용가능할 것이다. 하지만 넓은 원천 양립불가론자 일반이 과연 이때 이용가능한 답변만으로 만족할 수 있을 것인지 여부는 분명하지 않으며, 따라서 프랑크푸르트-유형 사례로부터 제기되는 문제는 최소한 몇몇 넓은 원천 양립불가론자들에게 있어서는 스스로-결정하는 힘의 본성과

하려 하는 종류의 넓은 원천 양립불가론만을 의미하는 것으로 사용할 것이다.

[12] 표현 '외재적 양방력 이론'의 정확한 의미와 관련해서는 본 논문의 3장에서 자세히 설명될 것이다.

[13] '프랑크푸르트-유형 사례'란 곧 프랑크푸르트(Frankfurt, 1969)에 의해서 제시되었던 사례 및 이를 대안가능성 원리에 대한 반례로서 더 설득력이 있도록 만들기 위해 적절히 변형하여 구성해낸 사례들을 총칭하는 말이다.

관련하여 외재적 양방력 이론을 채택하는 것을 재고할 한가지 좋은 이유를 제공해줄 수 있을 것이다. (4장) 더불어 우리가 만약 스스로-결정하는 힘과 외재적 양방력을 동일시한다면 이로 인해 스스로-결정하는 힘의 본성과 관련하여 어떤 문제적인 귀결에 개입하는 것을 피하기 어려운 것처럼 보이며 (5장) 또한 우리가 자유로운 행위자로서 자신의 행위에 대해 가지고 있는 통제력과 들뜬 원자가 광자를 방출할지 말지 여부에 대해 가지고 있는 통제력을 구분짓는데 있어서 어려움을 겪게 되는 것처럼 보인다. (6장) 이러한 일련의 사안들을 고려할 때 우리는 다음과 같은 결론에 도달하는 것을 피하기 어려워 보인다: 양방력으로서의 스스로-결정하는 힘의 본성에 호소하여 원천-관계의 본성을 해명하려 하는 넓은 원천 양립불가론자의 입장에서 양방력의 본성과 관련하여 최소한 외재적 양방력 이론을 채택하는 것은 받아들이기 어려운 선택지이다.

제 2 장 힘과 관련된 용어들

양방력의 본성을 본격적으로 논의하기에 앞서 먼저 힘의 본성에 대해 논의할 때 자주 사용되는 용어들의 의미를 분명히 해보자. 먼저 힘이란 무엇인가? 힘은 일반적으로 성향(disposition), 능력(ability), 역량(capability) 등과 동일시되거나 혹은 최소한 이들과 매우 밀접한 관계를 맺고 있는 속성의 일종으로 이해된다 (Jacobs, 2017, p. 1; Molnar, 2003, p. 57). 물론 힘이 정확히 무엇이며 성향, 능력, 역량 등과는 정확히 어떤 관계를 맺고 있는지에 대해서는 논란의 여지가 있을 수 있다. 하지만 본 논문에서 나는 몰나(Molnar, 2003, p. 57)를 따라 해당 문제들과 관련하여 어떠한 이론적 개입도 하지 않은 채 논의를 진행할 것이다. 다만 나는 여기서 힘을 다음과 같은 사례들을 모두 포괄하는 일련의 "성향적 속성들(dispositional properties)"의 전형으로 이해할 것이다: 전자가 가지고 있는 양전하를 띤 물체를 끌어당기는 힘, 자석이 가지고 있는 철을 끌어당기는 힘, 물이 가지고 있는 염화나트륨을 녹이는 힘, 그리고 독미나리즙이 가지고 있는 사람을 죽이는 힘.

우리는 힘에 대해 항상 그것이 무엇을 하는 힘인지를 물을 수 있다. 더불어 이러한 물음에 대한 적합한 답변은 항상 그것은 이러저러한 것을 하는 힘이라는 형식으로 주어진다. 임의의 힘 P에 대해, P가 무엇을 하는 힘인지를 물을 때 우리가 묻고 있는 것, 더불어 이 물음에 대한 적합한 답변으로 P는 이러저러한 것을 하는 힘이라고 답변할 때 우리가 제시하는 것, 그것을 P의 '정체성(identity)을 이루는 발현-유형(manifestation-type)', 혹은 '발현-유형 정체성(manifestation-type identity)'이라고 하자. 예를 들어 항생제는 일반적으로 세균을 죽이는 힘을 가진다. 우리는 항생제가 갖고 있는 문제의 힘에 대해 그것이 무엇을 하는 힘인지를 물어볼 수 있으며 이에 대한 적합한 답변은 그것은 세균을 죽이는 힘이라는 방식으로 주어진다. 따라서 발현-유형

세균을 죽임은 해당 힘의 발현-유형 정체성이라고 할 수 있다.[14]

그런데 자신의 발현-유형 정체성에 해당하지 않는 방식으로 행사되는 힘이 있을 수 있는가? 프로스트(Frost, 2020)는 어떤 힘은 자신의 발현-유형 정체성에 해당하지 않는 방식으로 행사될 수 있으며 양방력 또한 이러한 종류의 힘들 중 하나라고 주장한다. 그는 양방력 외에도 자신의 발현-유형 정체성에 해당하지 않는 방식으로 행사될 수 있는 힘의 사례로 다음을 제시한다. 먼저 어떤 질량을 지닌 물체 O가 뉴턴의 만유인력의 법칙에 의해서 그것의 발현-유형 정체성이 참되게 기술되는 힘 P를 가지고 있다고 가정해보자. 카트라이트(Cartwright, 1980/2012)가 지적했듯이, 이때 O가 가지고 있는 힘 P가 행사되었다고 해서 꼭 O가 P의 발현-유형 정체성에 해당하는 방식으로 운동하게 되는 것은 아니다. 왜냐하면 만유인력의 법칙은 오직 전자기력 등 다른 힘이 전혀 작동하지 않으면서 또한 존재하는 물체가 (O를 포함하여) 오직 두개인 상황에 대해서만 참된 기술을 제공해주기 때문이다. 따라서 O가 가지고 있는 힘 P는 발현-유형 정체성에 해당하지 않는 방식으로도 행사될 수 있는 그러한 힘의 사례라고 할 수 있다.

만약 이러한 프로스트의 주장이 옳다면 발현-유형과 힘 사이에 성립하는 관계는 서로 다른 두가지 종류가 있을 것이다. 첫째, 발현-유형 M과 힘 P 사이에는 관계 〈M은 P의 발현-유형 정체성이다〉가 성립할 수 있다. 둘째, 발현-유형과 힘 사이에는 관계 〈P는 M에 해당하는 방식으로 발현될 수 있다〉가 성립할 수 있다. 편의상 전자의 관계를 '발현-유형 정체성 관계'라 부르고 후자의 관계를 '발현-유형 소유 관계'라고 부르자. 또한 M과 P 사이에 발현-유형 소유 관계가 성립할 경우에는 'P는 M을 소유한다'고 말하도록 하자. 예를 들어

[14] 정체성에 대한 이러한 설명은 아리스토텔레스의 본질 개념에 대한 한성일(2015, p. 83)의 설명 방식을 응용한 것이다. 스튜어드 또한 우리가 힘에 대해 그것이 어떤 발현-유형을 자신의 "정의적 행사(definitive act)"(Steward, 2013, p. 688)로 갖는지를 물을 수 있음을 언급하고 있으며, 그녀가 여기서 염두에 두고 있는 것은 본 논문에서 발현-유형 정체성으로 규정된 바와 유사한 것처럼 보인다.

항생제가 가지고 있는 세균을 죽이는 힘은 발현-유형 **세균을 죽임**과 발현-유형 정체성 관계를 맺으며 또한 동시에 발현-유형 소유 관계를 맺는다. 왜냐하면 문제의 힘은 발현-유형 **세균을 죽임**을 자신의 발현-유형 정체성으로 가지며 또한 발현-유형 **세균을 죽임**에 해당하는 방식으로 발현될 수 있기 때문이다. 우리가 용어 '발현-유형 정체성 관계'와 용어 '발현-유형 소유 관계'를 이와 같은 방식으로 사용한다면 앞서 제시된 프로스트의 주장은 다음과 같이 재진술될 수 있을 것이다: 어떤 발현-유형 M과 어떤 힘 P에 대해 P와 M 사이에는 발현-유형 소유 관계는 성립하지만 발현-유형 정체성 관계는 성립하지 않는다.

그런데 힘은 아무 때나 발현 혹은 행사되지 않는다. 힘은 일반적으로 특정 조건이 만족된 경우에 한하여 행사된다. 예를 들어 진통제가 갖고 있는 힘, 즉 고통을 경감시켜주는 힘이 행사되기 위해서는 먼저 통증을 느끼고 있는 환자에게 투여되어야 한다. 진통제는 병에 담겨있는 채로는 자신의 힘을 행사할 수 없다. 더불어 일반적인 경우 힘은 특정 조건이 만족될 경우 행사되지 않을 수 없다. 예를 들어 음전하를 띠는 입자는 양전하를 띠는 입자를 끌어당기지 않을 수 없는 것처럼 보인다. 모든 발현-조건이 발현의 필요조건으로서의 발현-조건인지 여부 혹은 모든 발현-조건이 발현의 충분조건으로서의 발현-조건인지 여부는 물론 논란의 여지가 있을 수 있다. 하지만 최소한 일반적으로는 분명 발현 혹은 행사와 밀접한 관련을 맺고 있는 어떤 조건들이 존재하는 것처럼 보이며, 이를 '발현-조건(manifestation-condition)'이라고 부르도록 하자.[15][16]

---

[15] 심지어 모든 힘이 이러한 방식으로 규정된 발현-조건을 갖는지 여부조차도 논란의 여지가 있을 수 있다. 로우(Lowe, 2011, p. 22)는 발현-조건을 전적으로 결여하는 힘의 사례로 라듐 원소가 가지고 있는 방사성 붕괴를 겪을 수 있는 힘과 질량을 지닌 물체가 갖는 질량을 지닌 다른 물체를 끌어당기는 힘을 언급하고 있다. 그에 의하면 라듐 원소가 가지고 있는 방사성 붕괴를 겪을 수 있는 힘은 즉자적으로(spontaneously) 행사되는 힘으로 어떠한 발현-조건도 갖지 않는다. 더불어 질량을 지닌 물체가 갖는 질량을 지닌 다른 물체를 끌어당기는 힘의 경우 이는 상시 발현되는 힘으로 마찬가지로 특별한 발현-조건을 갖지 않는다. 하지만 본 논문에서는 일단 모든 힘이 발현의 필요조건으로서의 발현-조건을 갖는다고 가정할 것이다.

더불어 항생제가 가지고 있는 힘과 진통제가 가지고 있는 힘은 분명 서로 다른 힘이다. 왜 그런가? 이에 대한 한가지 자연스러운 답변은 항생제가 갖고 있는 힘의 정체성을 이루는 발현-유형 **세균을 죽임**과 진통제가 갖고 있는 힘의 정체성을 이루는 발현-유형 **고통을 경감시켜줌**은 서로 다른 발현-유형이기 때문이라는 답변이다. 즉 힘의 동일성 기준(identity criterion)에 대한 한가지 자연스러운 생각은, 임의의 힘 P, Q에 대해, P와 Q의 정체성을 이루는 발현-유형이 서로 동일할 경우 오직 그 경우에 P와 Q가 서로 동일하다는 것이다. 또한 힘은 자신의 발현-유형 정체성에 의해 개별화(individuation)되는 것처럼 보인다. 혹은 힘은 자신의 발현-유형 정체성을 개별화 원리(principle of individuation)로 갖는 것처럼 보인다. 즉 발현-유형 M을 자신의 발현-유형 정체성으로 갖는 힘 P에 대해, M은 "[힘] 종에 속하는 존재자들 중 어떤 것이 [P]인지를 결정"(Lowe, 2010, p. 9)해준다. 예를 들어 발현-유형 **세균을 죽임**은 세균을 죽이는 힘, 고통을 경감시켜주는 힘, 혈당을 낮추는 힘, 혈압을 낮추는 힘 등등의 힘들 중에서 세균을 죽이는 힘이 정확히 어떤 것인지를 결정해준다.[17][18]

---

[16] 앞서 소개된 하나의 힘이 복수의 발현-유형과 발현-유형 소유 관계를 맺을 수 있다는 생각은 발현-조건과 관련하여 다음과 같은 질문을 낳을 수 있다. 말하자면, 서로 다른 두 발현-유형 M과 M* 모두와 발현-유형 소유 관계를 맺고 있는 힘 P에 대해 P는 M과 M* 각각에 대해 서로 다른 발현-조건을 갖는가? 아니면, 단 하나의 발현-조건을 가질 뿐인가? 이러한 질문과 관련하여 나는 현재 다음과 같은 생각을 가지고 있다: 최소한 양방력으로 이해된 넓은 원천 양립불가론자들의 스스로-결정하는 힘에 대해서는, 만약 그것이 서로 다른 두 발현-유형 M, M*와 발현-유형 소유 관계를 맺고 있는 힘이라면, M과 결부되어 있는 발현-조건 C와 M*과 결부되어 있는 발현-조건 C* 사이에 동일성 관계가 성립해야 한다. 하지만 복수의 발현-유형과 발현-유형 소유 관계를 맺고 있는 힘 "일반"에 대해 우리가 앞서 제시된 질문에 대해 답변하기 위해서는 자유 및 도덕적 책임의 근거를 이루는 스스로-결정하는 힘을 넘어서 자연 세계에서 발견되는 힘 "일반"을 고려해야 하며, 이러한 고려는 본 논문의 범위를 넘어선다. 따라서 이에 대한 보다 본격적인 논의는 후속 과제로 남겨두도록 하겠다.

[17] 힘의 개별화 원리를 이루는 것이 정확히 무엇인지에 대해서는 여러 견해가 존재한다. 로우(Lowe, 2010, 2012)는 발현-유형 정체성이 힘의 개별화 원리를 이룬다고 주장한 바 있으며 이에 대한 대안적 견해로는 다음을 참조하라:

지금까지 소개된 용어들을 다시 한 번 정리해보자. 첫째, 힘이란 곧 성향, 능력, 역량 등과 동일시되거나 혹은 최소한 이들과 매우 밀접한 관계를 맺고 있는 성향적 속성이다. 둘째, 임의의 힘 P에 대해 P의 발현-유형 정체성이란 곧 P가 무엇을 하는 힘인지를 물을 때 우리가 묻고 있는 것, 더불어 이 물음에 대한 적합한 답변으로 P는 이러저러한

Tugby(2017).

[18] 동일성 기준과 개별화 원리는 구분되어야 한다. 왜냐하면 임의의 종 K에 대해, K에 속하는 존재자들 사이에서 동일성 관계가 언제 성립하고 언제 성립하지 않는지를 말해줄 수 있는 기준, 즉 동일성 기준은 여러가지가 있을 수 있지만 해당 기준들이 모두 K에 속하는 존재자들의 개별화 원리에 해당하는 것은 아닐 수 있기 때문이다. 개별화 원리는 동일성 기준을 제공해주는 것을 넘어 다음과 같은 이론적 역할 또한 수행해줄 수 있다: 개별화 원리는 동일성 사실이 왜 성립하는지를 설명해준다. (혹은 동일성 사실의 형이상학적 근거를 제공한다.) 예를 들어 임의의 두 힘 P와 Q에 대해 P를 소유하는 모든 가능자(possibilia)들의 집합과 Q를 소유하는 모든 가능자들의 집합이 동일할 경우 오직 그 경우에 P와 Q가 동일하다고 해보자 (루이스(Lewis, 1983)는 임의의 속성 F에 대해 F는 곧 F인 가능자들의 집합이라고 주장한 바 있다). 이때 다음의 쌍조건문 또한 마찬가지로 참일 수 있다: P의 발현-유형 정체성과 Q의 발현-유형 정체성이 동일할 경우 오직 그 경우에 P와 Q는 동일하다. 즉 다음 세 진술 'P를 소유하는 모든 가능자들의 집합과 Q를 소유하는 모든 가능자들의 집합이 동일하다', 'P의 발현-유형 정체성과 Q의 발현-유형 정체성이 동일하다', 'P와 Q가 동일하다'는 모두 동치(equivalence) 관계에 있을 수 있다. 그렇다면 우리는 힘을 소유하는 가능자들의 집합과 힘의 발현-유형 정체성이 모두 힘의 개별화 원리를 제공해준다고 말할 수 있을 것인가? 그럴 수는 없다. 왜냐하면 [P와 Q가 동일하다]를 설명해주는 사실은 오직 [P의 발현-유형 정체성과 Q의 발현-유형 정체성이 동일하다] 뿐이기 때문이다. [P를 소유하는 모든 가능자들의 집합과 Q를 소유하는 모든 가능자들의 집합이 동일하다]는 [P와 Q는 동일하다]의 근거를 제공해준다기 보다는 오히려 [P와 Q가 동일하다]를 자신의 근거로 갖는 그러한 종류의 사실에 불과하다. 따라서 힘의 개별화 원리를 제공해주는 것은 오직 힘의 발현-유형 정체성 뿐이며 이로부터 개별화 원리에 해당하지 않는 동일성 기준이 존재함이 예증된다. 또한 로우(Lowe, 2010)는 동일성 기준은 존재하지만 개별화 원리는 존재하지 않는 경우가 있다는 사실에 근거하여 동일성 기준과 개별화 원리가 구분될 필요성이 있음을 논증한다. 로우에 의하면 헬륨 원자를 이루고 있는 서로 다른 두 전자를 각각 a와 b라고 할 때 분명 a와 b는 서로 다른 전자이지만 (따라서 동일성 기준은 존재하지만) 주어진 서로 다른 두 전자들 중 정확히 무엇이 a이고 무엇이 b인지를 결정해줄 수 있는 사실은 존재하지 않는다 (즉 개별화 원리는 존재하지 않는다).

것을 하는 힘이라고 답변할 때 우리가 제시하는 것이다. 셋째, 발현-유형 M과 힘 P 사이에 〈P는 M에 해당하는 방식으로 발현될 수 있다〉가 성립할 경우 그리고 오직 그 경우에 나는 P와 M 사이에는 발현-유형 소유 관계가 성립한다고 말할 것이다. (더불어 P는 M을 소유한다고도 말할 것이다.) 넷째, 본 논문에서 나는 임의의 힘 P에 대해 P의 발현 혹은 행사와 밀접한 관련을 맺고 있는 조건들을 'P의 발현-조건'이라 부를 것이다. 다섯째, 발현-유형 정체성은 일반적으로 힘의 동일성 기준 및 개별화 원리로 이해된다.

이러한 용어들은 앞서 소개된 단방력과 양방력의 차이 및 양방력의 본성을 개념화하는 방식에 대한 우리의 논의를 보다 명료하게 해줄 것이다. 예를 들어, 나는 앞서 어떤 힘이 단방력이라는 것은 곧 해당 힘이 행사될 수 있는 조건이 만족되었을 때 일어날 수 있는 일이 오직 한가지 뿐임을 의미한다고 말한 바 있다. 여기서 사용된 표현 '해당 힘이 행사될 수 있는 조건'은 곧 해당 힘의 발현-조건을 의미하며, 따라서 단방력이란 다시금 발현-조건이 만족되었을 경우 오직 한가지 일만이 일어날 수 있는 그러한 종류의 힘으로 이해될 수 있다. 반대로 어떤 힘이 양방력이라는 것은 곧 해당 힘이 행사될 수 있는 조건이 만족되었을 때 일어날 수 있는 일이 다음과 같은 의미에서 한가지가 아니라 두가지임을 의미했다: 어떤 일이 일어날 수도 있지만 일어나지 않을 수도 있다. 여기서 사용된 표현 '해당 힘이 행사될 수 있는 조건' 또한 마찬가지로 발현-조건을 의미하며, 따라서 양방력이란 다시금 발현-조건이 만족되었을 경우 서로 다른 두가지 일이 일어날 수 있는 그러한 종류의 힘으로 이해될 수 있다. 더불어, 본 논문의 논의를 통해 그 정확한 함의가 드러나겠지만, 우리는 표현 '오직 한가지 일만이 일어날 수 있는'과 표현 '서로 다른 두가지 일이 일어날 수 있는'의 정확한 의미를 여기서 소개된 용어들을 통해 서로 다른 두가지 방식으로 이해할 수 있다. 먼저 우리는 표현 '오직 한가지 일만이 일어날 수 있는'을 발현-조건이 만족되었을 경우 문제의 힘이 발현되지 않을 수 없다는 것을 의미하는 표현으로 이해하고 표현 '서로 다른 두가지 일이 일어날 수 있는'을 발현-조건이 만족되었을 경우 문제의

힘이 발현될 수도 있고 발현되지 않을 수도 있다는 것을 의미하는 표현으로 이해해볼 수 있다. 반면 우리는 표현 '오직 한가지 일만이 일어날 수 있는'을 오직 한가지 발현-유형에 해당하는 방식으로 문제의 힘이 행사될 수 있다는 것을 의미하는 표현으로 이해하고 표현 '서로 다른 두가지 일이 일어날 수 있는'을 서로 다른 두가지 발현-유형에 해당하는 방식으로 문제의 힘이 행사될 수 있다는 것을 의미하는 표현으로 이해해볼 수도 있다.

제 3 장 외재적 양방력 이론이란 무엇인가?

이제 힘의 본성에 대한 논의에 사용되는 용어들의 의미를 분명히 했으니, 양방력의 본성을 개념화하는 한가지 유력한 방식에 대해 살펴보도록 하자. 앞서 나는 어떤 힘이 양방력이라는 것은 곧 문제의 힘의 발현-조건이 만족되었을 때 일어날 수 있는 일이 다음과 같은 의미에서 한가지가 아니라 두가지임을 의미한다고 말한 바 있다: 어떤 일이 일어날 수도 있지만 일어나지 않을 수도 있다. 그리고 바로 이러한 점에서 양방력은 발현-조건이 만족되었을 때 일어날 수 있는 일이 오직 하나뿐인 힘, 즉 단방력과 구분된다고 말한 바 있다.

하지만 (이 또한 앞서 잠시 언급되었듯이) 이렇게 이해된 양방력의 본성 및 양방력과 단방력의 차이를 이해하는 방식에는 서로 다른 두가지가 있을 수 있다. 스튜어드(Steward, 2013, 2020, 2021)와 알바레즈(Alvarez, 2013)는 그중 한가지 방식을 명시적으로 옹호한다. 그녀들은 양방력의 본성을 다음과 같이 개념화한다. 첫째, 양방력은 발현의 충분조건으로서의 발현-조건을 결여하는 그러한 종류의 힘이다. 스튜어드와 알바레즈는 다음과 같이 주장한다: 양방력은 그것의 발현-조건이 만족된다 하더라도 여전히 발현되지 않을 수 있으며 바로 이러한 특징 덕분에 다른 종류의 힘, 즉 단방력들과 구분될 수 있는 그러한 종류의 힘이다. 스튜어드(Steward, 2021, p. 11)는 다음과 같은 예시를 든다. 물은 염화나트륨을 녹일 수 있는 힘을 갖는다. 또한 이 힘은 발현-조건이 만족될 경우 발현되지 않을 수 없는, 즉 염화나트륨과 닿은 경우에는 염화나트륨을 녹이지 않을 수 없는 그러한 종류의 힘이다. 따라서 물이 가지고 있는 염화나트륨을 녹이는 힘의 발현은 발현의 충분조건으로서의 발현-조건을 갖는 그러한 종류의 힘, 즉 단방력이다. 하지만 자유로운 행위자인 우리가 가지고 있는 스스로-결정하는 힘, 즉 양방력은 발현의 충분조건으로서의 발현-조건을 결여한다는 점에서 염화나트륨을 녹이는 힘과 같은 단방력과는 구분된다. 예를 들어 내가 자두를 먹을 때 행사한 힘, 즉 자두를 먹거나 먹지 않을 수 있는 힘은 발현의 충분조건으로서의 발현-조건을

결여하는 힘이다. 나는 방금 전 눈 앞에 자두를 보고 갈증과 허기를 느껴 이를 집어먹을 것인지 여부에 대해 잠시 고민을 한 후에 결국 자두를 먹기로 결정하고 이를 실행에 옮겼다. 하지만 분명 문제의 상황에서 나는 자두를 먹을 수 있었던 것 만큼이나 자두를 먹지 않을 수도 있었다. 즉 문제의 상황에서 내가 행사했던 힘, 즉 자두를 먹거나 먹지 않을 수 있는 힘은 발현-조건이 만족되었음에도 불구하고 여전히 행사되지 않을 수 있었다. (여기서 내가 갈증과 허기를 해소하길 원했고 또한 만약 자두를 먹는다면 이러한 나의 원함을 만족시킬 수 있다고 믿었으며, 더 나아가 실제로 자두를 먹을 것인지 여부에 대해 실제로 고민[19]을 하고 있었던 상황은 특이사항이 없는 경우 자연스럽게 문제의 힘의 발현-조건이 만족된 상황으로 이해될 수 있을 것이다.) 자두를 먹거나 먹지 않을 수 있는 힘과 관련해서 그것이 만족되었을 경우 내가 자두를 먹지 않을 수 없게 되는 그러한 종류의 발현-조건은 존재하지 않는다. 따라서 자두를 먹거나 먹지 않을 수 있는 힘의 경우에는 염화나트륨을 녹이는 힘의 경우와는 달리 설사 그것의 발현-조건이 만족되었다 하더라도 그것의 소유자인 나는 여전히 자두를 먹지 않을 수 있다. 즉 나는 문제의 힘을 행사하지 않을 수 있다. 바로 이러한 점에서 자유로운 행위자인 내가 가지고 있는 양방력, 즉 자두를 먹거나 먹지 않을 수 있는 힘은 물이 가지고 있는 단방력, 즉 염화나트륨을 녹이는 힘과 구분된다.

둘째, 스튜어드와 알바레즈의 개념화 방식에 따르면 양방력은 오직 하나의 발현-유형만을 자신의 발현-유형 정체성으로 가지며 이와 다른 방식으로는 발현될 수 없는 (즉, 자신의 발현-유형 정체성에 해당하지 않는 발현-유형과는 발현-유형 소유 관계를 맺지 않는) 힘이다. 그렇다면 만약 내가 자두를 먹지 않았다면 어떻게 되는 것인가? 나는 자두를 먹을 때 또한 자두를 먹지 않을 수 있었다. 만약 내가 문제의 상황에서 자두를 먹지 않았다면 이때 나는 자두를 먹거나 먹지 않을 수 있는 힘을 행사할 수 있었음에도 불구하고 행사하지 않은 것이다. 즉

[19] 물론 여기서 자두를 먹을지 여부에 대한 고민이 꼭 의식적으로 이루어질 필요는 없을 것이다.

내가 무엇인가를 하거나 하지 않을 수 있을 때 그것을 하지 않는다면 나는 문제의 상황에서 주어진 양방력을 행사하지 않은 것이다. 스튜어드와 알바레즈에 의하면 양방력과 결부되어 있는 대안가능성은 무엇보다 양방력이 행사되지 않을 수 있는 가능성으로, 예를 들어, 발현-유형 **자두를 먹음**을 발현-유형 정체성으로 갖는 양방력과 결부되어 있는 대안가능성, 즉 자두를 먹지 않을 수 있는 가능성은 자두를 먹거나 먹지 않을 수 있는 양방력이 행사되지 않을 수 있는 가능성으로 이해되어야 한다.

나는 본 논문에서 양방력의 본성에 대한 스튜어드와 알바레즈의 견해를 '외재적 양방력 이론'이라 부르고 그녀들을 '외재적 양방력 이론가'라고 부를 것이다. 그녀들이 양방력의 본성을 이해하는 방식에 따르면 양방력이 무엇인가를 하거나 하지 않을 수 있는 힘인 이유는 무엇보다도 문제의 힘이 설사 발현-조건이 만족되었다 하더라도 여전히 행사될 수도 있고 그렇지 않을 수도 있는 힘이기 때문이다. 따라서 스튜어드와 알바레즈의 방식으로 이해된 양방력의 양방향성(two-wayness)은 곧 발현과 비-발현(non-manifestation) 사이에 성립하는 양방향성이며 이는 다시금 양방력의 발현에 대해 "외재적"인 방식으로 성립하는 그러한 종류의 양방향성이다. 이것이 내가 그녀들의 양방력 이론을 '외재적 양방력' 이론이라 부르는 이유이다.[20][21]

---

[20] 스튜어드(Steward, 2020, pp. 348-349)는 양방력의 양방향성에 대한 이러한 방식의 이해 (외재적 양방력 이론)를 본격적으로 옹호한다. 반면 알바레즈의 경우 비록 내가 아는 한 그녀가 외재적 양방력 이론을 스튜어드만큼 본격적으로 옹호한 적은 없지만 알바레즈(Alvarez, 2013)도 스튜어드와 마찬가지로 분명 양방력의 양방향성과 관련하여 외재적 양방력 이론을 받아들이고 있다. 더불어 여기서 사용된 용어 '외재적 양방력 이론' 및 뒤에서 잠시 언급될 용어 '내재적 양방력 이론'은 한성일 교수님께서 처음 제안해주신 것으로 내가 아는 한 기존에 어떤 철학자에 의해서도 사용된 바 없다.

[21] 김율(2008, pp. 222-227)에 의하면 스코투스에 의해 일찍이 의지가 갖는 두 가지 차원의 비결정성이 이미 논의된 바 있다. 김율에 의해 설명된 스코투스의 의지 이론에 따르면 의지는 그것이 작용할 수 있는 복수의 방식들 중에서 정확히 어떤 방식으로 작용할 것인지 여부와 관련하여서도 비결정성을 갖지만

또한 그것이 도대체 어떤 방식으로든 작용하기는 할 것인지 여부와 관련하여서도 마찬가지로 비결정성을 갖는다. 외재적 양방력 이론에 의하면 양방력의 양방향성은 전자의 비결정성에 대응하기 보다는 후자의 비결정성에 대응하는 것으로 이해될 수 있다. 왜냐하면 외재적 양방력의 양방향성은 문제의 힘이 행사될 수 있는 방식들 사이에 성립하는 양방향성이라기 보다는 문제의 힘이 행사될 것인지 아니면 행사되지 않을 것인지 여부와 관련하여 성립하는 양방향성이기 때문이다. 이상 양방력에 대한 현대적 논의와 의지에 대한 스콜라 철학자들의 논의 사이에 성립하는 관계에 대한 고려의 필요성은 강상진 교수님께서 지적해주신 것이다.

제 4 장 프랑크푸르트-유형 사례 문제

하지만 외재적 양방력 이론[22]이 과연 넓은 원천 양립불가론자의 입장에서 스스로-결정하는 힘으로서의 양방력의 본성을 개념화하는 하나의 만족스러운 방식일 수 있을 것인가? 적어도 나에게는 외재적 양방력 이론이 만족스러운 넓은 원천 양립불가론이라고 생각하기 어려운 이유들이 여럿 존재하는 것처럼 보인다. 지금부터 그 이유들을 하나씩 차근차근 살펴보도록 하자.

먼저 4장에서는 프랑크푸르트-유형 사례로부터 발생하는 문제와 관련하여 외재적 양방력 이론이 겪게 되는 어려움을 살펴볼 것이다. 앞서 언급된 바와 같이 넓은 원천 양립불가론은 대안가능성 원리에 개입한다. 즉 넓은 원천 양립불가론에 따르면 행위자 S는 오직 달리 행할 수 있었던 경우에 한하여 자신이 수행한 행위 A에 대해 도덕적 책임을 진다. 하지만 대안가능성 원리에 개입하는 자유 및 도덕적 책임의 본성에 대한 모든 이론은 프랑크푸르트-유형 사례로부터 제기되는 문제에 부딪히게 된다. 왜냐하면 프랑크푸르트-유형 사례는 행위자가 A를 수행함에 있어서 달리 행할 수 없었음에도 불구하고 여전히 A에 대해 도덕적 책임을 지는 것처럼 보이는 사례이기 때문이다. 즉 프랑크푸르트-유형 사례는 대안가능성 원리에 대한 반례인 것처럼 보이기 때문이다. 따라서 넓은 원천 양립불가론자들 또한 그들이 대안가능성 원리에 개입하는 이상 프랑크푸르트-유형 사례로부터 제기되는 문제에 대해 어떤 그럴듯한 답변을 제공해야 한다.

4장에서는 넓은 원천 양립불가론자들이 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 채택할 경우 프랑크푸르트-유형 사례로부터 제기되는 문제에 대해 답변함에 있어서 한가지

---

[22] 이후 별도의 표시가 특별히 주어지지 않을 경우 표현 '외재적 양방력 이론'은 모두 스스로-결정하는 힘의 본성을 외재적 양방력의 일종으로 이해하는 넓은 원천 양립불가론을 의미하는 것으로 사용할 것이며 표현 '외재적 양방력 이론가' 또한 이러한 종류의 넓은 원천 양립불가론을 받아들이는 이들을 의미하는 것으로 사용할 것이다.

불필요한 제약을 받아들여야 함을 논변할 것이다. 앞서 잠시 언급한 바와 같이 이러한 제약 하에서도 물론 어떤 종류의 답변들은 여전히 이용가능할 것이다. 하지만 넓은 원천 양립불가론자 일반이 과연 이때 이용가능한 답변만으로 만족할 수 있을 것인지 여부는 분명치 않으며, 따라서 프랑크푸르트-유형 사례로부터 제기되는 문제는 최소한 몇몇 넓은 원천 양립불가론자들에게 있어서는 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 채택하는 것을 재고할 한가지 좋은 이유를 제공해줄 수 있을 것이다.

이를 보이기 위해 나는 먼저 프랑크푸르트-유형 사례를 한가지 제시한 후 문제의 사례가 왜 대안가능성 원리에 대한 반례인 것처럼 보이는지를 설명할 것이다. 더불어 이에 대해 피셔(Fischer, 2006, pp. 38-62)가 '스쳐 지나가는 자유(flickers of freedom) 전략'이라고 부른 종류의 답변이 어떻게 주어질 수 있는지 살펴볼 것이다. 그 후에는 스쳐 지나가는 자유 전략에 대한 한가지 유력한 비판 논변인 피셔(Fischer, 2006, pp. 38-62)의 탄탄성 비판(robustness objection)을 소개할 것이다. 넓은 원천 양립불가론자가 탄탄성 비판에 대해 제시할 수 있는 답변은 기본적으로 크게 두가지 종류로 구분될 수 있다. 이후 나는 첫번째 종류—본 논문에서는 이러한 종류에 해당하는 답변을 '단호한(hard-line) 답변'이라 부를 것이다—에 해당하는 답변을 한가지 제시한 후 외재적 양방력 이론가는 이러한 종류의 답변, 즉 단호한 답변 일반을 이용할 수 없음을 주장하고 논변할 것이다. 물론 외재적 양방력 이론가들이 단호한 답변을 이용할 수 없다는 사실로부터 외재적 양방력 이론가들이 탄탄성 비판에 대해 도대체 어떠한 답변도 제시할 수 없다는 귀결이 따라 나오는 것은 아니다. 하지만 우리는 최소한 이로부터 외재적 양방력 이론을 받아들이는 넓은 원천 양립불가론자의 입장에서 탄탄성 비판에 대해 제시할 수 있는 답변은 오직 두번째 종류에 해당하는 답변—본 논문에서는 이를 '온건한(soft-line) 답변'이라 부를 것이다—뿐임을 귀결로 얻을 수 있을 것이다. 그런데 외재적 양방력 이론가들에게 이용가능한 탄탄성 비판에 대한 답변이 단지 온건한 답변 뿐이라는 사실은 최소한 몇몇 넓은 원천

양립불가론자들에게는 외재적 양방력 이론을 거부할 충분한 이유에 해당할 수 있다. 왜냐하면 넓은 원천 양립불가론자의 입장에서 단호한 답변을 포기하고 온건한 답변을 제공하는 것에 만족한다는 것은 곧 자유 및 도덕적 책임의 근거를 제공하는 원천-관계와 스스로-결정하는 힘에 대해 보다 "탄탄한(robust)" 이해를 포기하고 이에 대한 상대적으로 "초라한(humble)" 이해 만으로 만족한다는 것을 함의하기 때문이다. 본 논문에서는 먼저 외재적 양방력 이론가들의 입장에서 온건한 답변이 구체적으로 어떻게 주어질 수 있는지를 소개한 후, 이러한 온건한 답변들이 정확히 왜 이러한 함의를 갖게 되는지를 논변할 것이다. 만약 여기서 제시될 논변이 성공적이라면 외재적 양방력 이론가들이 프랑크푸르트-유형 사례로부터 제기되는 문제에 대해 답변하기 위해 채택할 수 있는 답변이 오직 온건한 답변 뿐이라는 사실은 최소한 몇몇 넓은 원천 양립불가론자에게는 외재적 양방력 이론을 거부할 좋은 이유에 해당할 수 있을 것이다.[23]

이제 본격적으로 논의를 시작해보자. 앞서 언급한 바와 같이 많은 철학자들은 자유 및 도덕적 책임과 관련하여 다음과 같은 대안가능성 원리를 받아들인다:

> (대안가능성 원리) 행위자 S는 오직 달리 행할 수 있었던 경우에 한하여 자신이 수행한 행위 A에 대해 도덕적 책임을 진다.

대안가능성 원리는 여러가지 방식으로 옹호될 수 있지만[24] 무엇보다 도덕적 책임 귀속과 관련하여 우리가 일상적으로 내리는 판단들을 통해서 정당화될 수 있는 것처럼 보인다. 롭 (Robb, 2020, pp.

---

[23] 여기서 사용된 다음 용어들의 정확한 의미에 대해서는 뒤에서 보다 상세히 설명될 것이다: '단호한 답변', '온건한 답변', '탄탄한 이해' 그리고 '초라한 이해'.

[24] 여기서 제시된 방식 외에도 대안가능성 원리가 옹호될 수 있는 방식에는 여러가지가 있을 수 있다. 이와 관련하여서는 롭(Robb, 2020, pp. 10-16)을 참조하라.

11-12)에 의하면 우리는 일반적인 경우 대안가능성이 없었던 행위들에 대해서는 바로 그러한 사실 때문에 도덕적 책임을 묻지 않는 것처럼 보인다. 예를 들어 우리는 강요된 행위는 대안가능성을 동반하지 않는다고 생각하며 (혹은 강요된 행위를 수행한 행위자는 다름아닌 강요로 인해 문제의 행위를 수행함에 있어서 달리 행할 수 없었다고 생각하며) 바로 이러한 사실이 우리가 강요된 행위에 대해 도덕적 책임을 묻지 않는 이유라고 생각한다. 자녀를 죽이겠다는 협박으로 인해 상관에게 거짓보고를 한 부하직원의 경우를 생각해보자. 우리는 문제의 상황에서 부하직원에게 (법적 책임을 물을 수는 있을지 몰라도) 도덕적 책임은 묻기 어렵다고 생각한다. 그리고 이러한 우리의 생각은 부하직원에게는 협박으로 인해 상관에게 거짓보고를 하는 것 말고는 별다른 선택지가 없었다는 판단에 기대고 있는 것처럼 보인다. 그렇다면 대안가능성 원리는 사실 이렇게 도덕적 책임 귀속과 관련하여 우리가 일상적으로 내리는 판단들을 일반화한 것에 불과하며, 따라서 어떤 의미에서 대안가능성 원리는 이러한 일상적 판단들에 의해서 충분히 만족스럽게 정당화될 수 있는 것처럼 보인다.

하지만 프랑크푸르트(Frankfurt, 1969)에 의하면 대안가능성 원리는 거짓이다. 왜냐하면 S가 A를 수행함에 있어서 달리 행할 수 없었음에도 불구하고 여전히 A에 대해 도덕적 책임을 지는 경우가 존재하는 것처럼 보이기 때문이다. 즉 대안가능성 원리에 대한 반례가 존재하는 것처럼 보이기 때문이다. 다음 사례를 생각해보자. 존스는 어느 날 자신에게 돈을 빌려달라고 부탁한 친구에게 지금 당장은 빌려줄 돈이 없다고 거짓말을 하기로 결심하고 이를 실행에 옮겼다. 이때 신경과학자 블랙은 몰래 존스의 두뇌에 심어놓은 칩을 통해 존스의 숙고 과정을 감시하고 있었으며 만약 존스가 친구에게 거짓말을 하지 않기로 결심하려는 징후(sign)를 조금이라도 보인다면[25] 즉시 개입하여

[25] 이는 사실 프랑크푸르트(Frankfurt, 1969)에 의해 제시되었던 원래 사례보다는 피셔(Fischer, 2006, pp. 38-62)에 의해 제시된 사례에 가깝다. 프랑크푸르트에 의해 제시되었던 원래 사례에서는 블랙이 존스가 친구에게 거짓말을 하지 않기로 결심하려는 것이 “분명해질 경우”에 개입하려고 준비하고 있었지만

존스의 두뇌에 전기자극을 주어 존스로 하여금 친구에게 거짓말을 하기로 결심하도록 그리고 문제의 결심을 실제로 실행에 옮기도록 만들 준비를 하고 있었다. 하지만 존스는 블랙의 개입 없이 스스로 친구에게 거짓말을 하기로 결심하고 이를 실행에 옮겼으며 블랙은 특별히 존스가 친구에게 거짓말을 하도록 만들기 위해 문제의 상황에 개입할 필요가 없었다.

일단 문제의 상황에서 존스는 일견 친구에게 거짓말을 하기로 결정함에 있어서, 더 나아가 친구에게 거짓말을 함에 있어서 달리 행할 수 없었던 것처럼 보인다. 왜냐하면 존스가 친구에게 거짓말을 하지 않기로 결정하려는 징후를 조금이라도 보일 경우 블랙은 문제의 상황에 개입하여 존스로 하여금 친구에게 거짓말을 하기로 결정하도록 그리고 문제의 결심을 실제로 실행에 옮기도록 만들 예정이었기 때문이다. 존스의 대안가능성은 모두 블랙에 의해 차단되어 있었다. 하지만 그럼에도 불구하고 여전히 존스는 이에 대해 도덕적 책임을 지는 것처럼 보인다. 왜냐하면 실제 상황에서 블랙은 어떠한 개입도 하지 않았으며 모든 결정과 행위는 존스가 "스스로" 내린 것이기 때문이다. 따라서 문제의 사례는 대안가능성 원리에 대한 반례인 것처럼 보인다.

하지만 해당 사례에서 정말로 존스는 달리 행할 수 없었는가? 피셔(Fischer, 2006, pp. 38-62)는 우리가 여전히 문제의 사례에서 다음과 같은 네 가지 종류의 대안가능성을 발견할 수 있음을 지적한다. 첫째, 존스는 최소한 친구에게 거짓말을 하지 않기로 결심하려는 "징후"를 보일 수 있었다. 분명 존스는 문제의 사례에서 친구에게 거짓말을 하지 않기로 결심할 수는 없었다. 왜냐하면 친구에게 거짓말을 하지 않기로 결심하려는 "징후"를 보이는 즉시 블랙은 존스의 의사결정 과정에 개입하여 존스로 하여금 친구에게 거짓말을 하기로

피셔에 의해 제시된 사례에서는 블랙이 존스가 친구에게 거짓말을 하지 않기로 결심하려는 "징후를 조금이라도 보인다면 즉시" 개입하려고 준비하고 있었다. 더불어 본 논문에서 제시된 프랑크푸르트-유형 사례는 물론 프랑크푸르트(Frankfurt, 1969) 및 피셔(Fischer, 2006, pp. 38-62)에 의해 제시된 사례들과 사소한 구체적 사안들에 있어서는 차이가 있다. 하지만 이러한 차이는 바라기로는 무시될 수 있을 것이다.

결심하고 이를 실행에 옮기도록 만들었을 것이기 때문이다. 하지만 최소한 존스는 친구에게 거짓말을 하지 않기로 결심하려는 "징후"는 보일 수 있었다. 둘째, 존스에게는 최소한 실제로 야기되었던 사건-개별자와 수적으로 다른 어떤 사건-개별자가 발생할 수 있었다. 분명 주어진 상황에서 존스에게 사건-유형 **친구에게 거짓말을 하기로 결심함**에 속하는 어떤 사건-개별자가 발생하는 것은 피할 수 없었던 것처럼 보인다. 하지만 존스에게 해당 사건-유형에 속함에도 불구하고 실제로 발생한 사건-개별자와는 수적으로 다른 어떤 사건-개별자가 발생하는 것은 여전히 가능했던 것처럼 보인다.[26] 셋째, 존스는 행위자-인과(agent-causation)를 통해 친구에게 거짓말을 하게 됨을 피할 수 있었다.[27] 존스가 친구에게 거짓말을 하지 않기로 결심하려는 "징후"를 보여서 블랙이 존스로 하여금 친구에게 거짓말을 하기로 결심하고 이를 실행에 옮기도록 만들었다고 해보자. 이때 존스는 분명 친구에게 거짓말을 하기로 결심하고 친구에게 거짓말을 했다고 말할 수 있을 것이다. 하지만 여기서 존스의 결정과 행위는 모두 존스가 행위자-인과를 통해서 야기한 것이 아니다. 이들은 모두 블랙에 의해 야기된 것일 뿐이다. 넷째, 존스는 **친구에게 거짓말을 하기로 "스스로" 결심함**을 피할 수 있었으며 또한 **"스스로" 친구에게 거짓말을 함**을 피할 수도 있었다. 물론 존스는 친구에게 거짓말을 하기로 결심하지 않을 수 없었다. 하지만 친구에게 거짓말을 하기로 "스스로" 결심하지 않을 수는 있었다. 더불어 존스는 친구에게 거짓말을 하지 않을 수

---

[26] 물론 사건-개별자의 수적 동일성 기준에 대해서는 여러 견해가 존재하며 그 중에서 정확히 어떤 견해를 받아들이는지에 따라서 이러한 "스쳐 지나가는 자유"가 존재할 수도 있고 그렇지 않을 수도 있다. 피셔(Fischer, 2006, pp. 38-62)는 만약 모든 사건-개별자가 자신의 원인을 본질적으로 갖는다는 견해를 받아들일 경우 존스에게 실제로 발생했던 사건-개별자와 수적으로 다른 어떤 사건-개별자가 발생할 수 있었음이 귀결로 따라 나온다는 사실을 지적한다.

[27] 행위자-인과를 통해 어떤 행위를 수행한다는 것은 일반적으로는 다음과 같이 이해될 수 있다: 행위자 S가 행위자-인과를 통해 어떤 행위 A를 수행했을 경우 오직 그 경우에 A가 산출된 인과적 역사를 추적했을 때 그 자체로는 다른 원인을 갖지 않는 실체로서의 바로 그 행위자 S를 원인-관계항으로 갖는 최초의 인과관계가 존재한다.

없었다. 하지만 “스스로” 친구에게 거짓말을 하지 않을 수는 있었다. 피셔는 이러한 종류의 대안가능성들을 모두 ‘스쳐 지나가는 자유’라 부르고 프랑크푸르트-유형 사례에서 여전히 이러한 종류의 대안가능성이 존재한다는 사실을 지적함으로써 프랑크푸르트-유형 사례들이 대안가능성 원리에 대한 반례가 되지 못함을 보이려는 시도들을 ‘스쳐 지나가는 자유 전략’이라 부른다.

피셔(Fischer, 2006, pp. 38-62)에 의하면 스쳐 지나가는 자유 전략들은 모두 실패한다. 왜 그런가? 피셔에 의하면, 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 이때 요구되는 대안가능성은 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성, 즉 “탄탄한 대안가능성”일 것이다. 그런데 소위 “스쳐 지나가는 자유”에 해당하는 대안가능성들은 모두 “탄탄한 대안가능성”이 아니다. 예를 들어 첫번째 “스쳐 지나가는 자유”, 즉 존스가 친구에게 거짓말을 하지 않기로 결심하려는 “징후”를 보일 수 있었던 가능성은, 문제의 “징후”가 행위가 아니라는 가정 하에, 존스가 실제로 수행한 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성으로 이해될 수는 없을 것이다. 만약 이러한 피셔의 생각이 옳다면 대안가능성 원리의 옹호자들이 프랑크푸르트-유형 사례로부터 제기되는 문제를 해결하기 위해 찾아내야 하는 대안가능성은 “스쳐 지나가는 자유”에 해당하는 대안가능성이 아니라 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성, 즉 “탄탄한 대안가능성”일 것이다. 하지만 프랑크푸르트-유형 사례에서 “스쳐 지나가는 자유”는 있을지 몰라도 “탄탄한 대안가능성”은 없는 것처럼 보인다. 따라서 스쳐 지나가는 자유 전략은 대안가능성 원리를 구제하는데 실패한다. 이러한 피셔의 비판을 ‘탄탄성 비판’이라 부르도록 하자.

사실 ‘탄탄한 대안가능성’은 다양한 방식으로 정의될 수 있다. 예를 들어 방금 전에 규정한 바에 따르면 ‘탄탄한 대안가능성’은 곧 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성을 의미하는 것으로 이해된다. 하지만 ‘탄탄한 대안가능성’은 다음과 같은 또 다른 방식으로 규정될 수도 있다: 탄탄한

대안가능성이란 곧 도덕적 책임 귀속을 근거짓기에 충분한 대안가능성을 말한다. 만약 '탄탄한 대안가능성'에 대해 후자의 규정을 받아들인다면, 다음과 같은 주장이 굳이 어떤 개념적 혼동을 저지르고 있는 것으로 이해될 필요는 없을 것이다: 어떤 대안가능성은 비록 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성에 해당하는 것은 아니지만 여전히 탄탄한 대안가능성일 수 있다. 하지만 만약 우리가 전자의 규정 (본 논문에서 채택할 규정)을 받아들인다면 우리는 더이상 어떤 개념적 혼동에 개입함 없이 이러한 주장을 하기 어려울 것이다. 이 경우 다만 우리는 다음과 같이 주장할 수 있을 뿐이다: 어떤 대안가능성은 비록 탄탄한 대안가능성에 해당하는 것은 아니지만 여전히 도덕적 책임 귀속을 근거짓기에 충분한 대안가능성일 수 있다. 이러한 차이는 어쩌면 단순히 말하는 방식에 있어서의 차이에 해당하는 것만은 아닐지도 모르겠다. 하지만 '탄탄한 대안가능성'의 의미를 고정하지 않은 채 논의를 진행할 경우 지나치게 논의가 복잡해질 우려가 있기 때문에 일단 본 논문에서는 '탄탄한 대안가능성'에 대한 전자의 규정을 잠정적으로 채택한 후 논의를 진행할 것이다. 이는 '탄탄한 대안가능성'에 대한 전자의 규정이 후자의 규정보다 더 나은 규정이라는 판단에 근거한 것이라기 보다는 단지 논의 상의 편의를 위한 선택에 불과하다.[28]

여하튼, 만약 탄탄성 비판이 성공적이라면 스쳐 지나가는 자유 전략들은 모두 실패할 것이다. (따라서 대안가능성 원리의 옹호자들은 다시금 프랑크푸르트-유형 사례를 처리할 수 있는 또 다른 방책을 고안해 내야 할 것이다.) 하지만 탄탄성 비판이 과연 성공적인가? 앞서 설명된 것과 같이 탄탄성 비판은 다음과 같이 요약될 수 있다:

(1) 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 여기서 요구되는 대안가능성은 탄탄한 대안가능성일 것이다.

(2) 탄탄한 대안가능성이 존재하지 않음에도 불구하고 도덕적

[28] 이에 대한 명시적 언급의 필요성은 한성일 교수님께서 지적해주신 것이다.

책임 귀속이 이루어지는 경우가 존재한다. (프랑크푸르트-유형 사례가 바로 그러한 종류의 사례이다.)

(3) 따라서 도덕적 책임 귀속은 대안가능성을 요구하지 않는다.

이상과 같이 요약된 탄탄성 비판은 분명 타당한 논변이며, 따라서 넓은 원천 양립불가론자가 탄탄성 비판에 대해 답변하기 위해 채택할 수 있는 방식은 (1)을 거부하는 방식과 (2)를 거부하는 방식 뿐이다. 나는 본 논문에서 (1)을 거부하는 방식을 채택하는 답변을 '온건한 답변'이라 부르고 (2)를 거부하는 방식을 채택하는 답변을 '단호한 답변'이라 부를 것이다.[29]

나는 피셔의 탄탄성 비판이 성공적이라고 생각하지 않는다. 왜냐하면 나는 프랑크푸르트-유형 사례가 우리가 "탄탄한 대안가능성"이 없었던 행위에 대해서도 도덕적 책임을 물을 수 있음을 보여주는 그러한 종류의 사례가 아니라고 생각하기 때문이다. 만약 프랑크푸르트-유형 사례가 이러한 종류의 사례가 아니라면, 더이상 우리에게는 (2)가 참이라고 생각할 좋은 이유는 없는 것처럼 보인다. 따라서 탄탄성 비판은 자신의 두번째 전제인 (2)와 함께 설득력을 잃게 된다. 즉 내가 지금부터 탄탄성 비판에 대해 제시할 답변은 단호한 답변에 해당하는 그러한 종류의 답변이다.

그렇다면 왜 프랑크푸르트-유형 사례는 "탄탄한 대안가능성"이 없었던 행위에 대해서도 도덕적 책임을 물을 수 있음을 보여주는 그러한 종류의 사례가 아닌가? 나는 일단 문제의 사례에서 존스의 결심과 행위에 대해 우리가 도덕적 책임을 물을 수 있다는 주장은 거부하기 어렵다고 생각한다. 하지만 동시에 나는 문제의 프랑크푸르트-유형 사례에서 존스에게는 여전히 "탄탄한 대안가능성"이 존재했다고 생각한다. 왜 그런가?

---

[29] 내가 본 논문에서 피셔의 탄탄성 비판에 대해 (1)을 거부하는 방식을 채택하는 답변을 '온건한 답변'이라 부르고 (2)를 거부하는 방식을 채택하는 답변을 '단호한 답변'이라 부르는 이유에 대해서는 뒤에서 보다 상세히 설명될 것이다.

먼저 앞서 제시된 프랑크푸르트-유형 사례와 유사한 상황, 즉 돈을 빌려달라는 말을 들은 존스가 친구에게 거짓말을 할지 말지 여부를 고민한 끝에 결국 거짓말을 하기로 결심하고 이를 실행에 옮겼지만 이때 블랙과 같은 감시자 및 가능적 개입자는 존재하지 않았던 상황을 고려해보도록 하자. (더불어 해당 상황을 '일상적 사례'라고 부르도록 하자.) 문제의 상황은 블랙이 존재하지 않는 상황이기 때문에 분명 존스는 해당 상황에서 친구에게 거짓말을 하기로 결심할 수 있었던 것 만큼이나 친구에게 거짓말을 하지 않기로 결심할 수 있었을 것이다. 행위자-인과 자유론자(agent-causal libertarian)[30]인 오코너(O'Connor, 2002, 2011, 2013)라면 문제의 상황을 다음과 같이 이해할 것이다: 존스는 주어진 상황에서 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**과 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함** 모두에 해당하는 방식으로 행사될 수 있는 행위자-인과력(agent-causal power)를 가지고 있었다; 더불어 문제의 행위자-인과력은 주어진 상황에서 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 발현될 수 있었던 것 만큼이나 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로도 발현될 수 있었다.[31]

이제 문제의 행위자-인과력의 본성에 대해 조금 더 자세히 살펴보자. 먼저 해당 행위자-인과력은 서로 다른 두 발현-유형과 발현-유형 소유 관계를 맺는 그러한 종류의 힘으로 이해된다. 따라서

---

[30] 행위자-인과 자유론이란 거칠게 말해 다음과 같은 입장을 말한다: (i) 인과적 결정론이 참인 세계에서 자유 및 도덕적 책임은 있을 수 없지만 우리 세계는 인과적 결정론이 거짓인 세계이기 때문에 자유 및 도덕적 책임이 있을 수 있으며 (ii) 모든 자유로운 행위는 그것의 인과적 역사가 행위자-인과로부터 시작된 그러한 종류의 행위이다.

[31] 더 정확히 말하자면 오코너는 주어진 상황에서 선택지 **친구에게 거짓말을 함**에 대해 존스가 가지고 있었던 이유를 R이라고 하고 선택지 **친구에게 거짓말을 하지 않음**에 대해 존스가 가지고 있었던 이유를 R*이라 할 때 존스는 발현-유형 **R 때문에 친구에게 거짓말을 하려는 의도를 형성함**과 발현-유형 **R* 때문에 친구에게 거짓말을 하지 않으려는 의도를 형성함** 모두에 해당하는 방식으로 행사될 수 있는 행위자-인과력을 가지고 있다고 말할 것이다.

다음과 같은 질문이 자연스럽게 제기될 수 있다: 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 발현되기 위해 만족되어야 하는 발현-조건 (이하 'C')과 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되기 위해 만족되어야 하는 발현-조건 (이하 'C*') 사이의 관계는 어떻게 되는가? 양립불가론자의 입장에서 이러한 질문에 대해 제시할 수 있는 가장 자연스러운 답변은 아무래도 다음과 같을 것이다: C와 C* 사이에는 동일성 관계가 성립한다. 이를 다음과 같이 표현할 수도 있다: 문제의 행위자-인과력은 자신과 발현-유형 소유 관계를 맺고 있는 발현-유형 각각에 대해 별도의 발현-조건을 갖는 것이 아니다; 문제의 행위자-인과력은 오직 하나의 발현-조건만을 가질 뿐이며 또한 그것의 유일한 발현-조건이 만족될 경우 일단 그것과 발현-유형 소유 관계를 맺고 있는 어떠한 발현-유형 M에 대해서든 M에 해당하는 방식으로 발현될 수 있다.

이번에는 다시 앞서 제시된 프랑크푸르트-유형 사례를 고려해보도록 하자. 문제의 사례에서 존스에게는 분명 일상적 사례의 존스에게 있었던 것과 동일한 종류의 행위자-인과력이 있었을 것이다. 즉 주어진 상황에서 존스는 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**과 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함** 모두에 해당하는 방식으로 행사될 수 있는 행위자-인과력을 여전히 가지고 있었을 것이다. 하지만 일상적 사례와는 달리 프랑크푸르트-유형 사례에서는 감시자이자 가능적 개입자인 블랙이 존재한다. 그렇다면 프랑크푸르트-유형 사례에서 감시자이자 가능적 개입자인 블랙이 해야 하는 일이 정확히 무엇인가? 여기서 블랙의 과제는 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되는 것을 차단하고 오직 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로만 행사되도록 만드는 그러한 일이다.

그렇다면 문제의 행위자 인과력이 발현-유형 **친구에게 거짓말을**

하지 않으려는 의도를 형성함에 해당하는 방식으로 행사되는 것을 차단하기 위해 블랙이 취할 수 있는 방법에는 어떤 것들이 있는가? 첫째, 블랙은 문제의 행위자-인과력의 발현-조건이 애초에 만족되지 못하도록 만들 수 있을 것이다. 만약 발현-조건이 만족되지 않는다면 문제의 행위자-인과력은 어떠한 방식으로든 발현될 수 없을 것이다. 따라서, 한층 더 강력한 이유로, 문제의 행위자-인과력은 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현될 수 없을 것이다. 둘째, 블랙은 만약 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되기 "시작"할 경우 이를 즉시 차단하고 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 발현되도록 만들 수 있을 것이다. 이 경우 존스는 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사하기 "시작"할 수는 있지만 문제의 행사를 성공적으로 "완료"할 수는 없을 것이다. 왜냐하면 문제의 행사가 "시작"될 경우 블랙이 즉시 개입하여 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 사건이 발생하는 것을 차단하고 해당 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 발현되도록 만들 것이기 때문이다.[32]

---

[32] 오코너(O'Connor, 2002, p. 81의 각주 25)는 자신의 행위자-인과 자유론에 의하면 블랙을 포함한 어떤 것에 의해서도 행위자-인과력의 행사 그 자체가 야기되는 것은 불가능하다고 주장한다. 만약 이러한 오코너의 주장이 옳다면 블랙이 개입하여 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 발현되도록 만드는 것은 불가능할 것이다. 하지만 오코너 본인도 인정하고 있듯이 설사 그의 행위자-인과 자유론이 옳다 하더라도 존스가 친구에게 거짓말을 하려는 의도를 갖도록 만드는 것은 여전히 가능하다. 따라서 블랙이 존스의 행위자-인과력이 특정한 방식으로 발현되도록 만들 준비를 하고 있다는 본 논문의 상황 설정에 대해 우려를 가지고 있는 독자가 있다면 문제의 상황을 블랙이 존스로 하여금 친구에게 거짓말을 하려는 의도를 형성하도록 만들 준비를 하고 있는 상황으로 대치하여 이해하라. 물론 프랑크푸르트-유형 사례와 관련된 철학적 논의의 맥락들 중에서

프랑크푸르트-유형 사례에서 블랙이 자신의 과제를 달성하기 위해 첫째 방법을 사용할 수 없다는 사실은 분명한 것처럼 보인다. 왜냐하면 문제의 행위자-인과력의 발현-조건이 만족되지 않을 경우 해당 행위자-인과력은 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현될 수 없을 뿐 아니라 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로도 발현될 수 없을 것이기 때문이다. 프랑크푸르트-유형 사례가 대안가능성 원리에 대한 반례로 작동하기 위해서는 어쨌든 존스가 "스스로" 친구에게 거짓말을 하기로 결심하고 이를 실행에 옮겨야 한다. 달리 말해 존스는 블랙의 개입 없이 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사해야 한다. 이를 위해서는 최소한 문제의 행위자-인과력의 발현-조건이 만족되어야 한다. 왜냐하면 문제의 행위자-인과력은 오직 하나의 발현-조건만을 가질 뿐이며 도대체 어떠한 방식으로든 그것이 행사되기 위해서는 일단 해당 행위자-인과력의 발현-조건이 만족되어야 하기 때문이다. 블랙이 만약 첫째 방법을 사용하여 존스가 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 자신의 행위자-인과력을 행사하는 것을 차단하려 할 경우, 블랙은 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사되는 것 또한 마찬가지로 차단하게 될 것이다.

그렇다면 블랙이 자신의 과제를 달성하기 위해 취할 수 있는 방법은 오직 둘째 방법 뿐인 것처럼 보인다. 즉 블랙이 취할 수 있는 방법은 오직 다음 뿐이다: 만약 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되기 "시작"할 경우 이를 즉시 차단하고 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 발현되도록 만든다. 스텀프(Stump, 2003)에 의해 제시된

이러한 대치가 유의미한 차이를 가져올 수 있는 맥락이 있겠지만, 바라기로는, 현재 논의의 맥락은 다행히도 그러한 맥락은 아닌 것처럼 보인다.

프랑크푸르트-유형 사례를 참조할 경우 우리는 블랙이 바로 이러한 방법을 사용하여 자신의 과제를 달성하려 하고 있는 사례를 고안해낼 수 있는 것처럼 보인다. 앞서 제시된 프랑크푸르트-유형 사례를 다음과 같이 수정해보도록 하자. 일단 우리는 존스가 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사할 경우 문제의 힘이 행사되는 과정, 즉 행위자-인과적 과정(agent-causal process)을 실현하는 신경 상관자(neural correlate)에 해당하는 어떤 신경학적 과정(neurological process)이 또한 동시에 진행될 것이라고 가정해볼 수 있다. 문제의 신경학적 과정을 'N'이라고 부르자. 이때 N은 일련의 사건들, 예를 들면 a, b, c로 이루어져 있을 것이다: a가 발생함에 따라 N은 진행되기 "시작"하고 c가 발생함에 따라 N은 "완료"된다. 이때 블랙은 다음과 같은 방식으로 존스의 두뇌에 개입할 수 있다: a가 발생할 경우 블랙은 즉시 개입하여 b와 c의 발생을 차단하고 대신 존스가 자신의 행위자-인과력을 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사할 경우에 해당 행위자-인과적 과정을 실현하는 신경 상관자로서 진행되었을 신경학적 과정 (말하자면 'N*')을 곧바로 진행시킨다. 이러한 상황에서는 분명 N이 "시작"될 수는 있을지 몰라도 "완료"될 수는 없었을 것이다. 따라서 N에 의해서 실현되는 행위자-인과적 과정, 즉 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되는 과정 또한 "시작"될 수는 있을지 몰라도 "완료"될 수는 없을 것이다. 그렇다면 우리는 여기서 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사될 수 있었던 가능성은 차단되어 있었다고 말해야 하는 것처럼 보인다. 더불어 이때 존스는 (실제로는) 블랙의 개입 없이 "스스로" 친구에게 거짓말을 하기로 결심했으며 따라서 a는 발생하지 않았고 블랙은 개입할 필요가 없었다. 따라서 여전히 우리는 여기서 존스는 자신의 결심과 행위에 대해 도덕적 책임을 진다고 말할 수 있다. 그렇다면 방금 제시된 사례는 "탄탄한 대안가능성"이 없었음에도 불구하고

여전히 도덕적 책임이 귀속될 수 있는 그러한 행위가 존재함을 보여주는 것처럼 보인다.[33]

하지만 만약 블랙이 이러한 방법 (둘째 방법)을 취한다면 우리는 존스에게 여전히 어떤 "탄탄한 대안가능성"이 존재했다고 말할 수 있는 것처럼 보인다. 왜냐하면 존스는 최소한 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사하길 "시작"했을 수 있기 때문이다.[34] 문제의 행위자-인과적 과정을 'P'라고 부르도록 하자. 새로운 사례는 P가 "진행"될 수 없었던 사례가 아니라 P가 "완료"될 수 없었던 사례에 불과하다. 새로운 사례에서 블랙이 개입하기 위해서는 적어도 P가 진행되기 "시작"했어야 한다. 더불어 일단 P가 진행된 이상 비록 P가 완료될 수는 없었다 하더라도 우리는 최소한 다음과 같은 비완료상(imperfective aspect) 문장을 참되게 말할 수 있다: '존스는 행위자-인과를 통해 행위 **친구에게 거짓말을 하지 않기로 결심함**을 수행하고 있었다'. 더불어 스튜어드(Steward, 2012b)가 주장한 것처럼 만약 우리가 이러한 비완료상 문장을 참되게 말할 수 있다면 우리는 또한 존스가 주어진 상황에서 최소한 어떤 활동(activity)에 "참여(engagement)"했던 것이라고 참되게 말할 수 있다. 이러한 "참여"의 가능성은 곧 실제로 수행된 것과 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성으로 우리는 이러한 "참여"의 가능성이 존재했다는 사실로부터 "탄탄한 대안가능성"이 있었다는 사실을 확보할 수 있다.[35] 행위자가 어떤 활동에 참여할 수 있었던

---

[33] 이 사례는 스텀프(Stump, 2003)에 의해 '(G)'라고 불린 사례를 본 논문에서 제시된 프랑크푸르트-유형 사례에 맞추어 약간 단순화하여 재구성한 것이다.

[34] 이와 유사한 논점은 이미 여러 철학자들에 의해서 지적된 바 있다 (Goetz, 2002, p. 135; O'Connor, 2002, p. 82; Widerker, 1995a, p.116, 1995b, pp. 252-253).

[35] 물론 나는 여기서 개념 〈행위〉를 심적 행위(mental action) 뿐 아니라 심적 활동(mental activity)에 대해서도 적용될 수 있는 것으로 이해하고 있다. 이렇게 개념 〈행위〉에 대해 상대적으로 넓은 이해 방식을 취하는 것이, 바라기로는, 특별히 현재 논의가 이루어지고 있는 맥락에서 논점선취와 같은 문제를 저지르고 있는 것으로 간주될 필요는 없다. 왜냐하면 일단 심적 행위 및 심적 활

가능성은 기껏해야 “스쳐 지나가는 자유”에 불과한 것으로 이해될 수 없다. 왜냐하면 해당 가능성 안에서 행위자는 분명 무엇인가를 한 것이기 때문이다. 그렇다면 결국 방금 전 제시된 스텀프의 사례를 참조하여 고안된 사례는 비록 P가 완료될 수 있는 그러한 종류의 사례는 아니지만 그럼에도 불구하고 여전히 존스가 P를 이루는 활동에 참여할 수 있었던 그러한 종류의 사례일 것이며, 따라서 여전히 존스에게 “탄탄한 대안가능성”이 존재했던 그러한 종류의 사례일 것이다. 즉 문제의 사례는 “탄탄한 대안가능성”이 없었음에도 불구하고 여전히 도덕적 책임이 귀속될 수 있는 그러한 행위가 존재함을 보여주는 사례가 아니다.

그런데 우리가 문제의 상황에서 존스에 대해 어떤 비완료상 문장을 참되게 말할 수 있으며 이러한 사실로부터 존스가 주어진 상황에서 최소한 어떤 활동에 참여했음이 귀결로서 따라 나온다는 주장을 왜 받아들여야 하는가? 더 나아가, 이러한 “참여”의 가능성이 곧 “탄탄한 대안가능성”에 해당한다는 주장을 우리는 왜 받아들여야 하는가? 나는 이러한 주장들은 모두 다음과 같은 사례로부터 예증될 수 있다고 생각한다. 도공인 철수가 어느 날 도자기를 만들던 중 작은 실수를 하게 되어 끝내 도자기를 완성하지 못했다고 가정해보자. 이때 완료상(perfective aspect) 문장 ‘철수는 도자기를 만들었다’는 분명 거짓일 것이다. 왜냐하면 철수는 어쨌든 도자기를 완성하지 못했기 때문이다. 하지만 비완료상 문장 ‘철수는 도자기를 만들고 있었다’는 여전히 참이다. 즉 비완료상 문장 ‘철수는 도자기를 만들고 있었다’가

동을 스스로-결정하는 힘의 사례로 이해하지 못할 이유는 찾기 어려우며 우리가 심적 행위 및 심적 활동을 스스로-결정하는 힘의 사례로 이해할 수 있는 이상 심적 행위 및 심적 활동으로서의 대안가능성이 “탄탄한 대안가능성”, 즉 실제로 수행된 것과 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성으로 이해되지 말아야 할 이유 또한 찾기 어렵기 때문이다. 오히려 특별한 논변 없이 심적 행위 및 심적 활동으로서의 대안가능성은 “탄탄한 대안가능성”일 수 없다고 주장하는 것이 대안가능성 원리의 비판자의 입장에서 대안가능성 원리의 옹호자에 대해 논점선취를 저지르는 것일 수 있다. 본 논문에서 개입하고 있는 개념 〈행위〉에 대한 이러한 명시적 설명의 필요성은 한성일 교수님께서 지적해주신 것이다.

참이기 위해 완료상 문장 '철수는 도자기를 만들었다'가 반드시 참일 필요는 없다.[36] 이는 철수가 비록 도자기를 만들지는 못했지만 여전히 무엇인가를 했음을 보여준다. 그렇다면 이때 철수가 한 것은 도대체 무엇인가? 그것은 바로 도자기 만들기-활동에 대한 참여이다. 철수가 도자기 만들기-활동에 참여했을 경우 오직 그 경우에 우리는 비완료상 문장 '철수는 도자기를 만들고 있었다'를 참되게 말할 수 있다.

행위자의 활동 참여는 분명 도덕적 책임 귀속의 적합한 대상일 수 있다. 예를 들어 우리는 어떤 사람이 비록 살인에 성공하진 못했다 하더라도 그가 살인 활동에 참여했다는 사실에 근거하여 그를 비난하고 그에게 살인미수의 책임을 묻는다. 더불어 무엇인가에 대해 도덕적 책임을 묻기 위해서는 그것이 반드시 자유롭게 수행된 것이어야 함을 가정한다면, 행위자의 활동 참여는 분명 자유롭게 수행된 것일 수 있다. 또한 무엇인가에 대해 도덕적 책임을 묻기 위해서는 그것이 반드시 행위자의 스스로-결정하는 힘의 사례여야 함을 가정할 경우, 행위자의 활동 참여는 스스로-결정하는 힘의 사례이기도 할 것이다. 그렇다면 분명 행위자의 활동 참여 가능성으로서의 대안가능성은 "탄탄한 대안가능성"일 수 있다. 즉 방금 전 제시된 프랑크푸르트-유형 사례에서 존스가 P를 이루는 활동에 참여할 수 있었다는 사실은 그가 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었음을 보여준다. 따라서 우리는 다음과 같이 결론지을 수 있을 것이다: 스텀프의 사례를 참조하여 새롭게 고안된 사례는 "탄탄한 대안가능성"이 없었음에도 불구하고 여전히 도덕적 책임이 귀속될 수 있는 그러한 행위가 존재함을 보여주는데 결국 실패한다.

이에 대해 혹자는 또 다른 종류의 프랑크푸르트-유형 사례를 제시함으로써 피셔를 옹호하려 할지도 모르겠다.[37] 다음과 같은

---

[36] 톰슨(Thompson, 2008, p. 124) 또한 어떤 비완료상 문장이 참이기 위해 꼭 대응하는 완료상 문장이 참일 필요는 없음을 지적한 바 있다.

[37] 이는 페러붐(Pereboom, 2001, pp. 18-22)에 의해 제시된 탈세 사례를 적절히 변형하여 만들어낸 사례이다. 이러한 종류의 사례에 대한 고려의 필요성은 한성일 교수님께서 지적해주신 것이다.

사례에서 블랙은 여전히 첫째 방법을 사용하여 자신의 과제를 달성할 수 있는 것처럼 보인다. 먼저 다음과 같은 "징후" S가 존재한다고 가정해보자: (i) S가 관찰되지 않은 채로 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되는 것은 불가능하다; 하지만 (ii) S가 관찰되지 않은 채로 존스가 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하는 것은 충분히 가능하다. 더불어 주어진 상황에서 존스는 S를 보이지 않은 채로 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사했지만 그럼에도 불구하고 여전히 해당 상황에서 존스는 언제든지 S를 보일 수도 있었으며, 따라서 (만약 블랙이 존재하지 않았다면) 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사할 수도 있었을 것이라고 가정해보자. 그렇다면 블랙은 S가 관찰될 경우 오직 그 경우에 존스에게 개입하여 존스로 하여금 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하도록 만듦으로써 자신의 과제를 성공적으로 달성할 수 있을 것이다. 왜냐하면 비록 해당 상황에서 실제로는 S가 관찰되지 않은 채로 블랙의 개입 없이 존스가 "스스로" 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하였지만, 만약 S가 관찰되었다면 블랙은 존스에게 개입하여 존스로 하여금 여전히 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하도록 만들었을 것이기 때문이다.[38] 가정 상 존스가 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지**

[38] 각주 32에서 언급되었듯이 블랙이 존스의 행위자-인과력이 특정한 방식으로 발현되도록 만들 준비를 하고 있다는 본 논문의 상황 설정에 대해 우려를 가지고 있는 독자가 있다면 문제의 상황을 블랙이 존스로 하여금 친구에게 거짓말을 하려는 의도를 형성하도록 만들 준비를 하고 있는 상황으로 대치하여 이해하라.

않으려는 의도를 형성함에 해당하는 방식으로 행사하기 위해서는 "먼저" S가 관찰되어야 한다. 하지만 S가 관찰되는 즉시 블랙은 존스에게 개입하여 문제의 행위자-인과력이 이러한 방식으로 행사되는 것을 차단하려 하고 있었다. 따라서 존스는 해당 상황에서 도저히 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사할 수 없었다. (더 나아가 행사하기 시작할 수조차 없었다.) 그렇다면 여기서 내가 주목하고 있는 "탄탄한 대안가능성"은 존재하지 않았던 것처럼 보인다. 하지만 존스는 여하튼 S를 보이지 않은 채로 블랙의 개입 없이 "스스로" 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하였다. 따라서 존스는 문제의 상황에서 자신의 결정에 대해 여전히 도덕적 책임을 지는 것처럼 보인다. 즉 해당 사례는 "탄탄한 대안가능성"이 없었던 행위에 대해서도 도덕적 책임이 귀속될 수 있음을 보여주는 그러한 사례이다.

나는 이러한 사례는 (i)[39] 와 (ii)[40] 를 동시에 만족하는 S가 존재한다고 가정함에 있어서 논점선취를 저지르고 있다고 생각한다.[41] 가정 상 S는, 만약 그것이 관찰된다면, 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 "이전에" 관찰되는 그러한 종류의 징후이다. 더불어 (i)에 의하면 S는 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 위해서 반드시 관찰되어야 하는 그러한 종류의 징후이다. 그렇다면 S는 다시금 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지**

---

[39] (i)의 내용은 다음과 같다: S가 관찰되지 않은 채로 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되는 것은 불가능하다.

[40] (ii)의 내용은 다음과 같다: S가 관찰되지 않은 채로 존스가 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하는 것은 충분히 가능하다.

[41] 괴츠(Goetz, 2002, p. 140) 또한 페러붐(Pereboom, 2001, pp. 18-22)의 탈세 사례에 대해 (비록 본 논문에서 제시된 것과는 다른 우려에 근거하고 있지만) 유사한 문제를 지적하고 있다.

않으려는 의도를 형성함에 해당하는 방식으로 행사되기 위해서 반드시 만족되어야 하는 발현-조건에 포함될 것이다. 하지만 가정 상 문제의 행위자-인과력은 자신과 발현-유형 소유 관계를 맺고 있는 발현-유형 각각에 대해 별도의 발현-조건을 갖는 힘이라기 보다는 단지 하나의 공통 발현-조건을 가질 뿐인 그러한 종류의 힘이다. 따라서 S는 결국 문제의 행위자-인과력이 도대체 어떠한 방식으로든 발현되기 위해서 반드시 만족되어야 하는 발현-조건에 포함되는 것처럼 보인다. 하지만 그렇다면 S가 관찰되지 않은 상황은 문제의 행위자-인과력의 발현-조건이 애초에 만족되지 못한 상황이다. 따라서 S가 관찰되지 않은 채로 존스가 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하는 것은 사실 불가능하다. 하지만 (ii)는 바로 이러한 일이 가능하다고 말하고 있다. 따라서 (i)과 (ii)를 동시에 만족하는 S가 존재한다고 가정하는 것은 결국 문제의 행위자-인과력이 자신과 발현-유형 소유 관계를 맺고 있는 발현-유형 각각에 대해 별도의 발현-조건들을 갖고 있으며 둘 중 하나가 만족되면서 나머지 하나는 만족되지 않는 것이 가능함을 가정하는 것이다. 이는 해당 상황에서 존스의 행위자-인과력이 오직 하나의 공통 발현-조건만을 가질 뿐이며 또한 그것의 유일한 발현-조건이 만족될 경우 일단 그것과 발현-유형 소유 관계를 맺고 있는 어떠한 발현-유형에 대해서든 그것에 해당하는 방식으로 발현될 수 있다는 앞서 제시된 가정과 충돌한다. 따라서 (i)과 (ii)를 동시에 만족하는 S가 존재한다고 가정하는 것은 행위자-인과력의 본성이 앞서 규정된 것과 같은 방식으로 이해되지 말아야 함을 가정하는 것과 마찬가지며, 이러한 가정은 분명 넓은 원천 양립불가론에 대한 비판으로서는 논점선취적이다.

이러한 나의 답변에 대해 혹자는 S는 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 "이전에" 관찰되는 그러한 종류의 징후이며 또한 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 위해

반드시 관찰되어야 하는 그러한 종류의 징후라는 사실로부터 다음이 따라 나오는 것은 아니라고 비판할지도 모르겠다: S는 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 위해서 반드시 만족되어야 하는 발현-조건에 포함된다. 왜냐하면 어쩌면 (i)[42]과 (ii)[43]를 모두 만족하는 S는 존스의 행위자-인과력의 발현-조건이 아니라, 말하자면, 문제의 힘이 발현되기 위한 배경 조건(background condition)의 일부일지도 모르기 때문이다.[44]

나로서는 일단 (i)과 (ii)를 모두 만족하는 징후 S를 존스의 행위자-인과력의 발현-조건이 아니라 발현의 배경 조건으로 이해해야 할 어떤 특별한 이유를 찾기 어렵다. 넓은 원천 양립불가론자들에게 제기된 피셔의 비판에 답변하는 현재 논의의 맥락에서 S를 발현-조건이 아닌 배경 조건으로 이해해야 할 별도의 이유가 제시되지 않는 이상 이러한 비판은 별로 설득력이 없을 것이다.

더 나아가 설사 S가 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 위해 반드시 만족되어야 하는 발현-조건이 아니라 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 위한 배경 조건이라 하더라도 이러한 사실이 궁극적으로 피셔를 옹호하는데 있어서 도움이 되기 위해서는 다음과 같은 가정이 별도로 추가되어야 한다: 존스의 행위자-인과력은 자신과 발현-유형 소유 관계를 맺고 있는 발현-유형 각각에 대해 별도의 배경 조건들을 따로 갖고 있으며 둘 중 하나가 만족되면서 나머지 하나는 만족되지 않는 것이 충분히 가능하다. 하지만 발현-조건에 대해서도 그랬던 것처럼

---

[42] (i)의 내용은 다음과 같다: S가 관찰되지 않은 채로 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되는 것은 불가능하다.

[43] (ii)의 내용은 다음과 같다: S가 관찰되지 않은 채로 존스가 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하는 것은 충분히 가능하다.

[44] 이러한 답변 가능성은 한성일 교수님께서 지적해주신 것이다.

배경 조건에 대해서도 마찬가지로 넓은 원천 양립불가론자의 입장에서 이러한 가정을 받아들여야 할 어떠한 독립적인 이유도 없는 것처럼 보인다.

오히려 넓은 원천 양립불가론자의 입장에서 존스의 행위자-인과력의 배경 조건에 대해 취할 수 있는 가장 자연스러운 이해 방식은 다음과 같다: 존스의 행위자-인과력은 그것과 발현-유형 소유 관계를 맺고 있는 각각의 발현-유형들에 대해 해당 힘이 그것에 해당하는 방식으로 발현되기 위해서 반드시 만족되어야 하는 고유한 배경 조건을 각각 따로 갖는다기 보다는 오직 하나의 공통적인 배경 조건만을 갖는다. 넓은 원천 양립불가론자들은 블랙과 같은 가능적 개입자가 없는 일상적 상황에서 존스의 행위자-인과력은 정확히 동일한 상황에서 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로도 발현될 수 있고 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로도 발현될 수 있다고 생각한다. 그리고 이러한 생각은 존스의 행위자-인과력의 발현-조건과 배경 조건에 대해 자연스럽게 다음을 시사한다: 존스의 행위자-인과력은 오직 하나의 공통 발현-조건과 배경 조건을 갖는다. 이러한 상황에서 앞서 말한 것처럼 존스의 행위자-인과력은 자신과 발현-유형 소유 관계를 맺고 있는 발현-유형 각각에 대해 별도의 배경 조건들을 따로 갖고 있으며 둘 중 하나가 만족되면서 나머지 하나는 만족되지 않는 것이 충분히 가능함을 별도의 논변 없이 가정하는 것은 논점선취를 저지르는 것이다.

그렇다면 결국 넓은 원천 양립불가론자들은 S가 존스의 행위자-인과력의 발현-조건이 아니라 배경 조건이라는 생각을 받아들인다 하더라도 여전히 존스의 행위자-인과력은 자신과 발현-유형 소유 관계를 맺고 있는 발현-유형 각각에 대해 별도의 배경 조건들을 갖고 있으며 둘 중 하나가 만족되면서 나머지 하나는 만족되지 않는 것이 가능하다는 가정을 거부함으로써 앞서 제시된 비판에 대해 손쉽게 답변할 수 있다. S가 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기

위해 반드시 만족되어야 하는 배경 조건에 포함된다고 가정해보자. 그렇다면 S가 관찰되지 않은 상황은 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사되기 위해 반드시 만족되어야 하는 배경 조건이 애초에 만족되지 못한 상황이다. 그렇다면 S가 관찰되지 않은 상황은 문제의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사되기 위해 반드시 만족되어야 하는 배경 조건 또한 마찬가지로 만족되지 못한 상황에 해당한다. 왜냐하면 문제의 행위자-인과력은 그것과 발현-유형 소유 관계를 맺고 있는 각각의 발현-유형들에 대해 해당 힘이 그것에 해당하는 방식으로 발현되기 위해서 반드시 만족되어야 하는 고유한 배경 조건을 각각 따로 갖는다기 보다는 오직 하나의 공통적인 배경 조건만을 갖기 때문이다. 따라서 S가 관찰되지 않은 채로 존스가 문제의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하는 것은 불가능하다. 하지만 (ii)는 바로 이러한 일이 가능하다고 말하고 있다. 따라서 (i)과 (ii)를 동시에 만족하는 S가 존재한다고 가정하는 것은 마찬가지로 행위자-인과력의 배경 조건에 대한 넓은 원천 양립불가론자의 관점에서 가장 자연스러운 방식의 이해를 특별한 논변 없이 배제하는 것과 마찬가지며, 이러한 가정은 넓은 원천 양립불가론에 대한 비판으로서는 분명 논점선취적이다.

하지만 혹자는 여전히 다음과 같이 불평할지도 모르겠다.[45] 문제의 징후 S는 여전히 존스의 행위자-인과력의 발현-조건의 일부로 이해될 필요가 없다. 먼저 존스의 행위자-인과력의 발현-조건이 모두 만족된 시점을 't_1'라고 하자. 더불어 해당 행위자-인과력이 발현되기 시작하는 시점을 't_2'라고 하자. 이때 앞서 언급한 특징들을 갖는 "징후" S는 다름아닌 t_1과 t_2 사이에 관찰될 수 있는 그러한 종류의 징후일 수 있으며, 따라서 S는 존스의 행위자-인과력의 발현-조건에 포함되는

[45] 이러한 답변 가능성은 한성일 교수님께서 지적해주신 것이다.

것은 아니지만 그럼에도 불구하고 여전히 블랙으로 하여금 성공적으로 자신의 과제를 달성할 수 있게 만들어주는 그러한 징후일 수 있다. 블랙은 S가 관찰될 경우 오직 그 경우에 개입함으로써 존스의 행위자-인과력이 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 발현되는 것을 성공적으로 차단할 수 있다. 왜냐하면 가정 상 S를 보이지 않은 채로 존스가 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사하는 것은 불가능하지만 S가 관찰될 경우 블랙은 존스가 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사하기 시작하기 "전에" (즉 t_1과 t_2 사이에) 개입하여 여전히 존스가 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하지 않으려는 의도를 형성함**에 해당하는 방식으로 행사하기 "시작"할 가능성 조차도 차단하는데 성공할 수 있기 때문이다. 더불어 존스는 실제로는 S를 보이지 않았으며, 따라서 블랙의 개입 없이 스스로 자신의 행위자-인과력을 발현-유형 **친구에게 거짓말을 하려는 의도를 형성함**에 해당하는 방식으로 행사하였다. 또한 S는 문제의 행위자-인과력의 발현-조건이 이미 모두 만족된 이후인 t_1 이후에 비로소 관찰될 수 있었던 그러한 종류의 징후로 해당 행위자-인과력의 발현-조건에 포함되는 것도 아니다. 그러므로 문제의 사례는 앞서 언급한 종류의 논점 선취를 저지르지 않고서도 여전히 "탄탄한 대안가능성" 없는 도덕적 책임 귀속이 가능함을 예증해주는 사례로 이해될 수 있다.

나는 이러한 비판은 t_1과 t_2가 다르다는 가정에 의존하고 있으며 넓은 원천 양립불가론자가 탄탄성 비판에 대해 답변함에 있어서 이러한 가정을 반드시 받아들일 필요는 없다고 생각한다. 물론 이러한 가정을 거부할 경우 존스의 행위자-인과력은 이제 발현-조건이 만족되는 "즉시" 그것과 발현-유형 소유 관계를 맺고 있는 발현-유형들 중 하나에 해당하는 방식으로 발현되는 그러한 종류의 힘으로 이해되어야 할 것이다. 하지만 문제의 행위자-인과력이 무엇보다 스스로-결정하는 힘으로 이해되는 한에 있어서는 그것이 발현-조건이 만족될 경우 즉시

그것과 발현-유형 소유 관계를 맺고 있는 발현-유형들 중 어느 하나에 해당하는 방식으로 발현되기 시작하는 그러한 종류의 힘으로 이해된다 하더라도 특별히 문제될 것은 없는 것처럼 보인다. 특히 우리가 복수의 선택지들 중 하나를 결정하는 과정, 즉 의도를 형성하는 과정과 복수의 선택지들 각각에 대해 그것을 선택해야 할 이유들 및 그것을 선택하지 말아야 할 이유들을 고려하는 숙고적 평가(deliberative appraisal)의 과정을 분명히 구분한다면, 후자의 과정 (숙고적 평가의 과정)이 아닌 전자의 과정 (의도를 형성하는 과정)을 매개하는 힘이 발현-조건이 만족될 경우 즉각적으로 발현되기 시작하는 힘이라고 생각하는 것은 특별히 반직관적으로 보이지 않는다. 따라서 넓은 원천 양립불가론자는 t_1과 t_2가 다르다는 가정을 거부함으로써 위에서 제시된 추가적인 비판에 대해 충분히 만족스럽게 답변할 수 있는 것처럼 보인다.[46]

그렇다면 넓은 원천 양립불가론자들은 결국 (2)[47]를 거부함으로써 탄탄성 비판에 대해 충분히 만족스러운 답변을 제시할 수 있는 것처럼 보인다. 즉 넓은 원천 양립불가론자들은 탄탄성 비판에 대해 단호한 답변을 충분히 만족스러운 방식으로 제시할 수 있는 것처럼 보인다. 왜냐하면, 만약 방금 제시된 나의 논변이 성공적이라면, 프랑크푸르트-유형 사례들은 결국 "탄탄한 대안가능성"이 없었던 행위에 대해서도 도덕적 책임이 귀속될 수 있음을 보여주는데 있어서 실패할 것이기

---

[46] t_1과 t_2가 동일하다고 가정하더라도, 혹자는 존스의 행위자-인과력의 발현-조건이 만족된 바로 그 시점, 그리고 존스의 행위자-인과력이 발현되기 시작하는 바로 그 시점, 즉 t_1에 관찰되는 어떤 징후를 통해 블랙이 자신의 과제를 달성할 수 있다고 생각할지도 모르겠다. (이러한 답변 가능성은 한성일 교수님께서 지적해주신 것이다.) 하지만 블랙이 t_1에 관찰되는 징후를 통해 자신의 개입 여부를 결정한다면 문제의 개입은 t_1 "이후"에야 비로소 가능할 것이다. 왜냐하면 문제의 징후를 알아보고 개입하는 과정이 필요하기 때문이다. 그렇다면 블랙은 존스의 행위자-인과력이 발현되기 시작한 바로 그 시점 "이후"에서야 비로소 개입할 수 있을 것이다. 그리고 앞서 스텀프의 사례를 참조하여 새롭게 고안한 사례를 검토하며 논변한 바와 같이 이러한 방식의 개입은 행위자의 활동 참여 가능성으로서의 탄탄한 대안가능성을 차단하는데 있어서 결국 실패한다.

[47] (2)의 내용은 다음과 같다: 탄탄한 대안가능성이 존재하지 않음에도 불구하고 도덕적 책임 귀속이 이루어지는 경우가 존재한다.

때문이다. 하지만 외재적 양방력 이론을 받아들이는 넓은 원천 양립불가론자가 과연 이러한 종류의 답변을 채택할 수 있을까? 나는 외재적 양방력 이론가는 (2)를 거부하는 방식으로 탄탄성 비판에 대해 답변할 수 없다고 생각한다. 즉 나는 외재적 양방력 이론가가 탄탄성 비판에 대해 단호한 답변을 제시하는 것은 불가능하다고 생각한다. 왜 그런가? 먼저 외재적 양방력 이론가들이 대안가능성 원리에 개입하게 되는 이유는 다음과 같았다: 외재적 양방력 이론에 의하면 [행위자 S는 자신이 수행한 행위 A의 원천이다]는 [S는 A를 수행함에 있어서 자신이 소유하고 있던 어떤 외재적 양방력을 행사했다]에 근거하며 후자의 사실은 다시금 대안가능성의 존재를 함축한다; 따라서 도덕적 책임 귀속은 원천-관계에 근거하며 원천-관계는 다시금 대안가능성의 존재를 함축하기 때문에 도덕적 책임 귀속은 대안가능성의 존재를 함축하게 된다. 이것이 바로 외재적 양방력 이론가들이 대안가능성 원리에 개입하게 되는 이유이다.

하지만 이때 요구되는 대안가능성은 기껏해야 외재적 양방력이 비-행사될 수 있었던 대안가능성에 지나지 않는다. 앞서 살펴보았듯이 외재적 양방력은 발현의 충분조건으로서의 발현-조건을 결여하는 힘이다. 즉 외재적 양방력은 설사 그것의 발현-조건이 만족되었다 하더라도 여전히 비-행사될 수 있는 그러한 종류의 힘이다. 외재적 양방력의 양방향성은 행사와 비-행사 사이에 성립하는 양방향성이며 S에 의해 수행된 행위 A가 어떤 외재적 양방력의 사례라는 사실로부터 따라 나오는 대안가능성 또한 기껏해야 문제의 외재적 양방력이 비-행사될 수 있었던 가능성에 지나지 않는다.

그런데 여기서 함축되는 대안가능성, 즉 외재적 양방력의 비-행사 가능성은 설사 "스쳐 지나가는 자유"일 수는 있을지 몰라도 "탄탄한 대안가능성"일 수는 없는 것처럼 보인다. 왜 그런가? 첫째, 일단 외재적 양방력을 포함하여 자유로운 행위자가 가지고 있는 "모든" 임의의 힘 P에 대해, P의 비-행사 가능성은 도저히 P의 소유자에 의해 자유로운 행위가 수행될 수 있는 그러한 종류의 가능성으로 이해되기 어려워 보인다. 둘째, 만약 외재적 양방력의 비-행사 가능성이 "탄탄한

대안가능성"일 수 있다면 외재적 양방력의 비-행사 가능성은 ("탄탄한 대안가능성"의 정의 상) 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성일 것이다. 하지만 외재적 양방력의 비-행사 가능성 또한 실제로 수행된 것과는 다른 종류의 행위를 자유롭게 수행할 수 있었던 가능성일 수 있다면 이제 외재적 양방력의 "행사"는 자유로운 행위의 필요조건일 수 없다. 왜냐하면 외재적 양방력의 행사 없이도 자유로운 행위가 수행될 수 있을 것이기 때문이다. 하지만 그렇다면 도대체 행위자가 문제의 외재적 양방력을 가지고 있었다는 사실이 그녀가 자유로운 행위를 수행했다는 사실과 무슨 상관이 있을 수 있단 말인가? 더 나아가 외재적 양방력이 (자유 및 도덕적 책임의 근거를 이루는) 원천-관계의 근거를 이루는 것으로 간주되고 있다는 점을 고려할 때, 외재적 양방력의 비-행사 가능성 또한 "탄탄한 대안가능성"일 수 있다는 입장은 자연스럽게 다음과 같은 질문에 부딪힌다: 무엇인가를 하는 경우와 무엇인가를 하지 않는 경우 사이에 문제의 양방력과 관련하여 도대체 어떤 사실이 공통적으로 성립하고 있길래 양방력은 그것이 행사되었는지 여부와는 독립적인 방식으로 무엇인가를 하는 것과 무엇인가를 하지 않는 것을 모두 자유로운 행위로 만들어줄 수 있는 것인가? 나는 이러한 질문에 대해 외재적 양방력 이론가의 입장에서 과연 그럴듯한 답변이 제시될 수 있을지 여부에 대해 회의적이다. 또한 이러한 질문에 대한 그럴듯한 답변이 제시되지 않는 이상 외재적 양방력의 비-행사 가능성 또한 "탄탄한 대안가능성"일 수 있다는 주장은 별로 설득력이 없을 것이다.[48]

그렇다면 외재적 양방력의 비-행사 가능성은 결국 "탄탄한 대안가능성"일 수 없을 것이다. 하지만 외재적 양방력 이론가들이 프랑크푸르트-유형 사례에서 그것의 존재를 주장할 수 있는 대안가능성은 오직 발현-유형 **친구에게 거짓말을 하기로 결심함**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력의 비-행사 가능성

[48] 이상과 같이 외재적 양방력의 비-행사 가능성이 "탄탄한 대안가능성"에 해당한다고 생각하기 어려운 이유에 대해 보다 적극적으로 해명해야 할 필요가 있다는 사실은 강진호 교수님께서 지적해주신 것이다.

뿐이다.[49] 즉, 외재적 양방력 이론가들이 프랑크푸르트-유형 사례에서 그것의 존재를 주장할 수 있는 대안가능성은 오직 탄탄하지 못한 대안가능성 뿐이다. 하지만 외재적 양방력 이론가들이 (2)[50]를 거부하기 위해서는 최소한 주어진 상황에서 존스에게 어떤 "탄탄한 대안가능성"이 존재했다고 주장해야 한다. 따라서 외재적 양방력 이론가들은 (2)를 거부함으로써 탄탄성 비판에 대해 답변할 수 없다. 즉 그녀들은 단호한 답변을 채택할 수 없다.

물론 외재적 양방력 이론을 수용하는 원천 양립불가론자들은 여전히 (1)[51]을 거부함으로써 피셔의 탄탄성 비판에 대해 답변할 수 있을지도 모르겠다. 즉 그녀들은 탄탄성 비판에 대해 (비록 단호한 답변은 제공할 수 없을지 몰라도) 온건한 답변은 충분히 만족스러운 방식으로 제시할 수 있을지도 모르겠다. 실제로 외재적 양방력 이론가인 스튜어드(Steward, 2009, 2012a, 2021)와 알바레즈(Alvarez, 2009)는 모두 (1)이 틀렸기 때문에 피셔의 탄탄성 비판은 별로 설득력이 없다고 생각하는 것처럼 보인다. 즉 스튜어드와 알바레즈는 피셔의 탄탄성 비판 논변에서 다음 전제를 거부한다: 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 여기서 요구되는 대안가능성은 "탄탄한 대안가능성"일 것이다. 그리고 탄탄성 비판에 대한 이러한 방식의 답변은 앞서 규정한 바에 따르면 온건한 답변에 해당하는 그러한 종류의 답변이다.

---

[49] 혹자는 프랑크푸르트-유형 사례의 존스가 주어진 상황에서 발현-유형 **친구에게 거짓말을 하기로 결심함**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력이 아닌 또 다른 외재적 양방력을 행사할 수 있었으며 이러한 대안가능성은 외재적 양방력의 행사 가능성이기 때문에 "탄탄한 대안가능성"의 일종으로 간주되는데 있어서 부족함이 없다고 주장할지도 모르겠다. 하지만 블랙이 존재하는 상황에서 이러한 별도의 외재적 양방력의 행사 가능성이 존재할 수 있는지 여부는 분명하지 않다. 자유로운 행위 및 이와 결부되어 있는 대안가능성을 서로 다른 두 외재적 양방력을 상정함으로써 설명하려 하는 시도에 대해서는 6장 후반부에서 더 자세히 논의될 것이다.

[50] (2)의 내용은 다음과 같다: 탄탄한 대안가능성이 존재하지 않음에도 불구하고 도덕적 책임 귀속이 이루어지는 경우가 존재한다.

[51] (1)의 내용은 다음과 같다: 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 여기서 요구되는 대안가능성은 탄탄한 대안가능성일 것이다.

외재적 양방력 이론가의 입장에서 온건한 답변은 다시금 서로 다른 두가지 방식으로 주어질 수 있다. 첫째, 외재적 양방력 이론가는 먼저 다음과 같이 주장함으로써 (1)을 거부할 수 있다: 도덕적 책임 귀속을 위해 대안가능성의 존재가 요구되는 것은 맞지만 이때 대안가능성의 존재는 도덕적 책임 귀속의 근거가 아니라 기껏해야 도덕적 책임 귀속의 필요조건으로서 요구되는 것에 불과하며 이러한 이유에서 요구되는 대안가능성이 굳이 "탄탄한 대안가능성"일 필요는 없다. 외재적 양방력 이론가들이 이러한 방식으로 탄탄성 비판에 대해 답변할 수 있는 이유는 무엇보다도 피셔가 (1)을 다음과 같이 정당화하고 있기 때문이다:

(4) 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 이는 대안가능성이 도덕적 책임 귀속을 근거짓기 때문일 것이다.
(5) 탄탄하지 못한 대안가능성은 도덕적 책임 귀속을 근거짓기에 충분하지 못하다.
(6) 따라서 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 이때 요구되는 대안가능성은 탄탄한 대안가능성일 것이다. (즉, (1)이 참이다.)

따라서 외재적 양방력 이론가들은 (4)를 거부함을 통해서 (1)에 대한 피셔의 논변을 거부할 수 있으며 이는 탄탄성 비판에 대한 온건한 답변이 주어질 수 있는 한가지 그럴듯한 방식에 해당할 것이다. 실제로 외재적 양방력 이론가인 스튜어드(Steward, 2009, 2012a, 2021)가 탄탄성 비판에 대해 제시하고 있는 답변 또한 온건한 답변의 일종으로 (1)에 대한 피셔의 정당화 논변의 첫번째 전제인 (4)를 거부하는 방식을 취하고 있다. 이러한 방식 ((4)를 거부하는 방식)으로 이루어지는 온건한 답변을 '온건한 답변 [1]'이라 부르도록 하자.

사실 이러한 방식의 답변 (즉 온건한 답변 [1])은 원칙적으로 넓은 원천 양립불가론자 일반이 이용할 수 있는 것처럼 보인다. 왜냐하면 넓은 원천 양립불가론자들에 의하면 도덕적 책임 귀속 사실이

대안가능성의 존재를 요구하는 이유는 무엇보다도 스스로-결정하는 힘의 모든 사례들이 그것의 본성 상 항상 어떤 대안가능성을 동반하게 되기 때문이며 이러한 이유에서 요구되는 대안가능성은 꼭 그 자체로 도덕적 책임 귀속 사실의 근거를 이룰 필요는 없는 것처럼 보이기 때문이다. 하지만 본 논문에서 주목하려고 하는 사실은 넓은 원천 양립불가론자 일반이 이러한 답변을 이용할 수 있다는 사실이 아니라 외재적 양방력 이론가들은 단호한 답변을 취할 수 없다는 사실이며, 이러한 사실은 곧 살펴보게 되겠지만 최소한 몇몇 넓은 원천 양립불가론자들에게는 외재적 양방력 이론을 거부할 좋은 이유를 제공해주는 것처럼 보인다.

둘째, 외재적 양방력 이론가의 입장에서 다음과 같은 방식으로 (1)을 거부하는 것 또한 충분히 가능한 것처럼 보인다: 도덕적 책임 귀속을 위해 대안가능성의 존재가 요구되는 것도 맞고 이는 대안가능성이 도덕적 책임 귀속을 근거짓기 때문인 것도 맞다; 하지만 어떤 대안가능성은 비록 탄탄한 대안가능성이 아님에도 불구하고 여전히 도덕적 책임 귀속을 근거짓기에 충분한 것처럼 보인다. 이러한 방식으로 이루어지는 온건한 답변은 온건한 답변 [1]과는 달리 (4)가 아니라 (5)를 거부함을 통해서 (1)에 대한 피셔의 논변을 거부하는 그러한 종류의 답변에 해당한다. 물론 해석이 분명치 않은 점이 존재하는 것은 사실이지만, 탄탄성 비판에 대한 알바레즈(Alvarez, 2009)의 답변은 실제로 이러한 방식을 취하고 있는 것으로 이해될 수 있다. 이러한 방식 (즉 (5)를 거부하는 방식)으로 이루어지는 온건한 답변을 '온건한 답변 [1]'과 대비하여 '온건한 답변 [2]'라고 부르도록 하자.

알바레즈는(Alvarez, 2009, p. 69)는 도덕적 책임 귀속이 대안가능성을 요구하는 이유에 대해 다음과 같이 설명한다. 먼저 무단횡단을 한 영희가 자신이 수행한 행위 **무단횡단을 함**에 대해 도덕적 책임을 지는 상황을 생각해보자. 이때 알바레즈에 따르면 영희가 자신의 행위에 대해 도덕적 책임을 지는 이유는 (부분적으로) (i) 그녀가 해당 상황에서 문제의 행위를 수행하지 않을 수 있었으며

또한 (ii) 문제의 행위를 수행할지 말지 여부가 그녀에게 달려 있었기 때문이다 (혹은, 영희와 그녀의 행위 사이에 달려 있음-관계가 성립했기 때문이다.) 그런데 영희와 그녀의 행위 사이에 달려 있음-관계가 성립했다는 사실 (즉 (ii)가 성립했다는 사실)은 다시금 어떻게 이해되어야 하는가? 이는 영희가 문제의 행위를 수행할지 말지 여부가 그녀가 해당 행위를 수행하기로 결정할지 말지 여부에 의존하고 있었음을 의미하는 것으로 이해된다. 알바레즈에 의하면 이는 다시금 다음의 쌍조건문이 참이였음을 의미한다: 영희는 행위 **무단횡단을 함**을 수행하기로 결정할 경우 오직 그 경우에 행위 **무단횡단을 함**을 수행할 것이다.

그렇다면 여기서 영희가 자신이 수행한 행위 **무단횡단을 함**에 대해 도덕적 책임을 진다는 사실, 즉 도덕적 책임 귀속 사실을 근거짓는 것으로 간주되고 있는 대안가능성은 정확히 무엇인가? 일단 영희는 문제의 행위를 수행하지 않을 수 있었다. 더불어 영희가 문제의 행위를 수행하지 않을 수 있었다는 사실은 영희와 그녀의 행위 사이에 달려 있음-관계가 성립했다는 사실과 더불어 다시금 다음을 함축한다: 영희는 행위 **무단횡단을 함**을 수행하기로 결정하지 않을 수 있었다. 왜냐하면 영희와 그녀의 행위 **무단횡단을 함** 사이에 달려 있음-관계가 성립했다는 것은 곧 쌍조건문 '영희는 행위 **무단횡단을 함**을 수행하기로 결정했을 경우 오직 그 경우에 행위 **무단횡단을 함**을 수행했을 것이다'가 참임을 의미하기 때문이다. 따라서 알바레즈에게 있어서 문제의 도덕적 책임 귀속 사실을 근거짓는 것으로 간주되고 있는 대안가능성은 결국 다음과 같은 대안가능성들이다: 행위 **무단횡단을 함**을 수행하지 않을 수 있었을 가능성 및 행위 **무단횡단을 함**을 수행하기로 결정하지 않을 수 있었을 가능성. 더 나아가 전자의 가능성은 어떤 의미에서 후자의 가능성에 의존하는 것처럼 보인다. 왜냐하면 달려 있음-관계에 대한 알바레즈의 설명을 고려할 때 영희가 행위 **무단횡단을 함**을 수행하지 않을 수 있었던 것은 무엇보다도 그녀가 행위 **무단횡단을 함**을 수행하기로 결정하지 않을 수 있었다는 사실 덕분인 것처럼 보이기 때문이다. 따라서 여기서 도덕적 책임 귀속

사실을 근거짓는 것으로 간주되고 있는 대안가능성은 무엇보다도 그녀가 행위 **무단횡단을 함**을 수행하기로 결정하지 않을 수 있었던 가능성이라고 할 수 있다.

한편 알바레즈(Alvarez, 2013)는 모든 행위자성의 사례는 외재적 양방력의 행사를 동반하며 또한 외재적 양방력이 행사될지 말지 여부는 항상 해당 외재적 양방력의 소유자에게 달려 있다고 주장한 바 있다. 이러한 알바레즈의 주장은 자연스럽게 다음을 시사하는 것처럼 보인다: 비록 피셔가 말하는 "탄탄한 대안가능성"은 아니지만 여전히 충분히 만족스러운 방식으로 도덕적 책임 귀속 사실의 근거를 이룰 수 있는 그러한 종류의 대안가능성, 즉 "실제로 수행했던 행위를 수행하기로 결정하지 않을 수 있었던 가능성"은 무엇보다도 외재적 양방력의 비-행사 가능성이다. 예를 들어, 무단횡단을 한 영희의 사례에서 영희의 행위 **무단횡단을 함**에 대한 도덕적 책임 귀속 사실을 근거짓는 대안가능성, 즉 행위 **무단횡단을 함**을 수행하기로 결정하지 않을 수 있었던 가능성은 곧 그녀의 외재적 양방력, 즉 행위 **무단횡단을 함**을 수행하기로 결정하거나 결정하지 않을 수 있는 힘이 비-행사될 수 있었던 가능성이다. 그렇다면 알바레즈에 의해 비록 "탄탄한 대안가능성"은 아니지만 그럼에도 불구하고 여전히 도덕적 책임 귀속 사실을 충분히 만족스럽게 근거지을 수 있는 것으로 간주되고 있는 그러한 종류의 대안가능성은 바로 다음과 같은 가능성이다: 행위자 S가 소유하고 있던 A를 하기로 결정하거나 결정하지 않을 수 있는 힘으로서의 외재적 양방력의 비-행사 가능성.

알바레즈(Alvarez, 2009)에 의하면 (만약 그녀에 대한 나의 해석이 옳다면) 이러한 대안가능성, 즉 행위자 S가 소유하고 있던 A를 하기로 결정하거나 결정하지 않을 수 있는 힘으로서의 외재적 양방력의 비-행사 가능성은 비록 "탄탄한 대안가능성"은 아니지만 그럼에도 불구하고 여전히 도덕적 책임 귀속 사실을 충분히 만족스럽게 근거지을 수 있다. 또한, 만약 이러한 알바레즈의 주장이 옳다면, (5), 즉 탄탄하지 못한 대안가능성은 도덕적 책임 귀속을 근거짓기에 충분하지 못하다는 피셔의 주장은 거짓일 것이다. 따라서 이상의 알바레즈의

설명은, 만약 그것이 정말 사실이라면, 외재적 양방력 이론가들로 하여금 온건한 답변 [2]에 해당하는 방식으로 탄탄성 비판에 대해 답변할 수 있는 한가지 그럴듯한 통로를 마련해줄 수 있을 것으로 기대된다.

나는 우리가 온건한 답변 [1] 혹은 온건한 답변 [2]과 관련하여 어떤 결정적인 문제를 발견하기는 어려우며, 따라서 어떤 의미에서 외재적 양방력 이론가들이 온건한 답변을 통해 피셔의 탄탄성 비판에 대해 답변하는 것이 충분히 가능함을 인정한다. 하지만 피셔의 탄탄성 비판의 첫번째 전제인 (1)[52]을 거부하는 이러한 종류의 답변들, 즉 온건한 답변들은 그것이 온건한 답변 [1]에 해당하든 아니면 온건한 답변 [2]에 해당하든 넓은 원천 양립불가론자의 답변으로서 어떤 분명한 한계를 가지고 있는 것처럼 보인다. 혹은 만약 넓은 원천 양립불가론자가 탄탄성 비판에 대해 단호한 답변을 포기한 채 기껏해야 온건한 답변을 제시하는 것으로 만족하게 된다면 이는 곧 넓은 원천 양립불가론자의 입장에서 (어쩌면 이러한 포기의 필요성은 그 자체로 외재적 양방력 이론을 포기할 충분한 이유를 제공해줄지도 모른다는 의미에서) 중요한 무엇인가를 포기하게 됨을 함의한다. 나는 지금부터 외재적 양방력 이론가의 입장에서 탄탄성 비판에 대해 제시해볼 수 있는 온건한 답변의 전형적인 사례로서 스튜어드(Steward, 2009, 2012a, 2021)의 답변을 비판적으로 검토하고, 이를 통해 외재적 양방력 이론가들이 (설사 탄탄성 비판에 대해 성공적으로 답변하게 된다 하더라도) 결국 탄탄성 비판으로 인해 포기해야 하는 것이 무엇인지를 분명히 할 것이다.

탄탄성 비판에 대해 스튜어드(Steward, 2009, 2012a, 2021)가 제시하는 답변은, 기본적으로, 단호한 답변이라기보다는 온건한 답변이라고 할 수 있다.[53] 왜냐하면 그녀는 무엇보다도 탄탄성 비판이

---

[52] (1)의 내용은 다음과 같다: 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 여기서 요구되는 대안가능성은 탄탄한 대안가능성일 것이다.

[53] 더 정확히 말해 탄탄성 비판에 대한 스튜어드의 답변은 온건한 답변 [1]이라고 할 수 있다. 하지만 지금부터 내가 지적하려 하는 (단호한 답변을 포기한

“탄탄한 대안가능성”을 요구하는 대안가능성 원리를 기각하는데 있어서는 성공했음을 인정하고 있기 때문이다.

먼저 그녀(Steward, 2009)는, 만약 도덕적 책임 귀속이 대안가능성을 필요로 하는 이유가 주어진 상황에서 행위자에게 도덕적 책임을 귀속하는 것이 공정(fair)해야 하기 때문이라면, 이때 요구되는 대안가능성은 피셔의 주장처럼 “탄탄한 대안가능성”이어야 할 것임을 받아들인다. 롭(Robb, 2020, p. 14)은 공정함에 대한 고려로부터 대안가능성 원리를 옹호하려는 시도를 다음과 같이 정리한 바 있다: 행위자 S가 자신이 수행한 행위 A에 대해 도덕적 책임을 진다는 것은 최소한 S가 비난이나 칭찬과 같은 “반응적 태도(reactive attitudes)”의 적절한 대상임을 함축한다; 하지만 달리 행할 수 없었던 행위로 인해 누군가를 비난하거나 칭찬하는 것은 “부당한(unfair)” 것처럼 보인다; 더불어 누군가를 비난하거나 칭찬하는 것이 부당하다면 그는 비난이나 칭찬의 적절한 대상일 수 없을 것이다; 따라서 S가 A를 수행함에 있어서 달리 행할 수 없었던 경우에 우리는 S에게 A에 대한 도덕적 책임을 물을 수 없는 것처럼 보인다. 스튜어드는, 만약 이러한 방식으로 대안가능성 원리가 성공적으로 정당화될 수 있다면, 이때 요구되는 대안가능성은 “탄탄한 대안가능성”일 것임을 인정한다. 왜냐하면 설사 S에게 모종의 대안가능성이 존재했다 하더라도 해당 대안가능성이 기껏해야 “스쳐 지나가는 자유”에 불과했다면 여전히 S를 비난하거나 칭찬하는 것은 “부당한” 것처럼 보이며 따라서 S는 비난이나 칭찬의 적절한 대상일 수 없는 것처럼 보이기 때문이다. 예를 들어 기껏해야 친구에게 거짓말을 하지 않기로 결심하려는 징후를 보일 수 있었던 가능성은 문제의 행위에 대한 우리의 비난을 공정한 것으로 만들어주기에는 역부족인 것처럼 보인다.

하지만 스튜어드(Steward, 2012a, pp. 191-192, 2021, p. 5)는 탄탄성 비판이 최소한 다음을 성공적으로 보여주었다고 생각한다:

채 이루어지는) 온건한 답변의 한계는 온건한 답변 [1] 뿐 아니라 온건한 답변 [2] 또한 공유하는 것으로 알바레즈의 답변 또한 마찬가지로 이러한 한계를 갖는다고 할 수 있을 것이다.

우리는 “탄탄한 대안가능성”이 없었던 행위에 대해 충분히 도덕적 책임을 물을 수 있으며, 따라서 공정함에 대한 고려는 도덕적 책임 귀속이 대안가능성을 요구한다는 생각을 정당화해주지 못한다. 왜냐하면, 만약 공정함에 대한 고려가 도덕적 책임 귀속이 대안가능성을 요구한다는 생각을 정당화해줄 수 있다면, 우리는 “탄탄한 대안가능성”이 없었던 행위에 대해서는 아무런 도덕적 책임도 물을 수 없을 것이기 때문이다. 앞서 제시된 프랑크푸르트-유형 사례에서 우리는 분명 “탄탄한 대안가능성”이 없었음에도 불구하고 여전히 존스에게 그의 거짓말에 대한 도덕적 책임을 물을 수 있는 것처럼 보인다.

하지만 그렇다고 스튜어드(Steward, 2009, 2012a, 2021)가 대안가능성 원리 자체를 포기하는 것은 아니다. 그녀는 대안가능성 원리가 행위자성에 대한 고려로부터 여전히 정당화될 수 있다고 주장한다. 스튜어드에 의하면 도덕적 책임 귀속 사실이 대안가능성을 요구하는 이유는 무엇보다도 도덕적 책임 귀속이 행위자성의 행사를 전제하며 행위자성의 행사는 다시금 모종의 대안가능성 (행위자성의 행사를 통해 확정된 사안이 행위자성의 행사 이전에는 확정되어 있지 않았음)을 전제하기 때문이다.

이러한 스튜어드의 주장은 정확히 어떻게 이해되어야 하는가? 일단 도덕적 책임 귀속이 행위자성의 행사를 전제한다는 스튜어드의 생각은 그것이 무엇을 의미하고 있는지가 다소 분명해 보인다. 왜냐하면 행위자 S에게 S의 행위자성의 사례가 아닌 것에 대해 도덕적 책임을 물을 수 있다는 주장은 그 자체로 어색하게 들리기 때문이다. 우리는 철수의 행위에 대해 영희에게 책임을 묻지 않으며 철수의 행위에 대한 책임은 오직 철수에게만 묻는다. 이는 분명 도덕적 책임 귀속이 행위자성의 행사를 전제한다는 스튜어드의 생각을 예증해주는 것처럼 보인다. 하지만 행위자성의 행사가 다시금 그 자체로 대안가능성을 전제한다는 생각은 일견 문제의 생각이 정확히 어떻게 이해되어야 하며 또한 왜 우리가 이러한 생각을 받아들여야 하는지 여부가 분명치 않은 것처럼 보일 수도 있다. 왜 하필 우리는 행위자성이 대안가능성의

존재를 전제한다고 생각해야 하는가? 이러한 질문에 대해 스튜어드(Steward, 2012a)가 제시하는 답변은 다음과 같다: 행위자성은 무엇보다도 행위자가 살아가는 세계에서 사건들이 발생할 수 있는 다양한 방식들 중 정확히 어떤 방식으로 사건들이 발생하게 될지 여부를 확정(settlement)하는 힘으로 이해된다; 더불어 행위자성이 행사되었다는 사실은 다시금 그것의 행사를 통해 확정된 사안이 행위자성의 행사 이전에는 확정되어 있지 않았었음을 전제한다 (왜냐하면 이미 확정된 사안은 다시 확정할 수 없기 때문이다); 달리 말해, 확정하는 힘으로 이해된 행위자성이 행사되었다는 사실은 확정하는 힘의 행사 이전에는 확정된 사안과 관련하여 세계에서 사건들이 발생할 수 있는 다양한 방식들이 존재했었음을, 즉 대안가능성이 존재했었음을 전제한다.

더불어 스튜어드(Steward, 2021)는 이때 행위자성을 이루는 확정하는 힘이 행사되었다는 사실에 의해 전제되는 대안가능성은 다름아닌 행위자성의 본성을 이루는 외재적 양방력이 비-행사될 수 있었던 가능성 혹은 실제로 수행되었던 행위가 수행되지 않을 수 있었던 가능성이라고 주장한다. 따라서 결국 그녀는 여전히 다음과 같은 이유에서 도덕적 책임 귀속을 위해 대안가능성의 존재가 요구된다는 주장에 개입한다: 도덕적 책임 귀속은 확정하는 힘으로서의 행위자성의 행사를 요구하며 확정하는 힘은 다시금 외재적 양방력의 일종이기 때문에 그것의 모든 사례들은 항상 외재적 양방력의 비-행사 가능성으로서의 대안가능성을 동반하게 된다.

스튜어드(Steward, 2021)에 의하면 이러한 이유에서 요구되는 대안가능성은 외재적 양방력의 모든 사례들이 그것의 본성 상 동반하게 되는 그러한 종류의 대안가능성으로 굳이 "탄탄한 대안가능성"일 필요가 없다. 이렇게 행위자성에 대한 고려로부터 정당화되는 대안가능성 원리를 '탄탄하지 못한 대안가능성 원리'라고 부르도록 하자. 반대로 공정함에 대한 고려로부터 정당화되는 (혹은 정당화되는 것처럼 보였던) 대안가능성 원리를 '탄탄한 대안가능성 원리'라고 부르도록 하자. 탄탄하지 못한 대안가능성 원리와 탄탄한 대안가능성

원리에 대한 보다 정확한 정식화는 다음과 같이 제시될 수 있다:

> (탄탄하지 못한 대안가능성 원리) 도덕적 책임 귀속은 대안가능성을 요구하지만 이때 요구되는 대안가능성이 꼭 탄탄한 대안가능성일 필요는 없다.
> (탄탄한 대안가능성 원리) 도덕적 책임 귀속은 탄탄한 대안가능성을 요구한다.

탄탄하지 못한 대안가능성 원리와 탄탄한 대안가능성 원리 사이의 구분을 통해 우리는 앞서 그 정확한 함의가 불분명한 채로 사용되었던 용어 '단호한 답변'과 '온건한 답변'의 의미를 보다 분명히 할 수 있다. 일단 스튜어드는 탄탄성 비판이 탄탄한 대안가능성 원리를 성공적으로 기각했다는 점은 인정한다. (즉 탄탄성 비판의 두번째 전제인 (2)[54]가 참임을 인정한다.) 하지만 그녀에 의하면 탄탄하지 못한 대안가능성 원리는 탄탄성 비판을 통해 기각되지 않는다.[55] 따라서 탄탄성 비판에 대한 그녀의 답변은 무엇보다도 탄탄성 비판의 부분적인 성공 (탄탄한 대안가능성 원리를 기각하는데 있어서 성공함)을 인정한다는 점에서 '온건한 답변'이라 불릴 수 있다.[56] 반대로 앞서 내가 피셔의 탄탄성 비판에 대해 제시했던 답변 (탄탄성 비판의 두번째 전제인 (2)를 거부하는 전략을 채택하는 답변)은 탄탄성 비판이 이러한 부분적인

---

[54] (2)의 내용은 다음과 같다: 탄탄한 대안가능성이 존재하지 않음에도 불구하고 도덕적 책임이 이루어지는 경우가 존재한다.

[55] 스튜어드(Steward, 2009)는 이러한 이유로 탄탄성 비판에 대해 답변하기 위해 자신이 채택하고 있는 전략을 '순수한 대치 전략(pure replacement strategy)'이라고 부른다.

[56] 이를 통해 앞서 제시된 '온건한 답변'에 대한 규정, 즉 〈탄탄성 비판의 첫번째 전제인 (1)을 거부하는 방식으로 이루어지는 답변〉은 다소 부정확했음이 드러난다. 왜냐하면 '온건한 답변'에 대해 방금 이루어진 규정에 의하면 온건한 답변은 곧 〈오직 탄탄성 비판의 첫번째 전제인 (1)을 거부하는 방식으로만 이루어지는 답변〉을 의미하며 이는 앞서 제시된 규정과는 다소 차이가 있기 때문이다. 하지만 앞서 제시된 규정은 엄밀한 규정으로 의도된 것은 아니었다. 따라서 앞에서 사용된 용어 '온건한 답변' 또한 마찬가지로 보다 정확히 말한다면 방금 이루어진 규정에 입각하여 사용된 것으로 이해되어야 한다.

성공조차도 거두지 못했다고 (혹은 탄탄한 대안가능성 원리조차도 기각하는데 성공하지 못했다고) 주장하는 그러한 종류의 답변이며, 따라서 이는 '단호한 답변'이라 불릴 수 있을 것이다.

스튜어드의 답변이 단호한 답변이 아니라 온건한 답변이라는 사실은 생각보다 중요한 함의를 가질 수 있다. 더불어 외재적 양방력 이론가들이 탄탄성 비판에 대해 제시할 수 있는 답변이 오직 온건한 답변 뿐이라는 사실을 고려할 때 (앞서 논변한 바와 같이 외재적 양방력 이론가들은 탄탄성 비판의 두번째 전제인 (2)를 거부하는 전략, 즉 단호한 답변을 제시하는 것을 목표로 하는 전략을 택하지 못하며, 따라서 탄탄성 비판에 대한 외재적 양방력 이론가들의 답변은 단호한 답변일 수 없다) 이러한 함의는 외재적 양방력 이론 일반이 갖게 되는 함의일 것이다. 그렇다면 스튜어드의 답변이 온건한 답변이라는 사실이 갖는 함의는 정확히 무엇인가? 먼저, 롭(Robb, 2020, p. 11)도 지적한 바 있듯이, 프랑크푸르트-유형 사례를 포함하여 대안가능성 원리에 대한 진지한 철학적 비판들이 다수 존재하는 상황에서 대안가능성 원리는 일단 자명한 것으로 간주되기는 어려워 보인다. 따라서 대안가능성 원리를 옹호하는 이들은 최소한 자신이 왜 대안가능성 원리를 참이라고 생각하는지에 대한 그럴듯한 이유를 제시해줄 필요가 있다. 넓은 원천 양립불가론자들이 대안가능성 원리를 옹호하기 위해 제시하는 이유는 일반적으로 다음과 같다. 먼저 넓은 원천 양립불가론에 의하면 사실 [행위자 S는 자신이 수행한 행위 A에 대해 도덕적 책임을 진다]는 사실 [S는 A의 원천이다]에 근거하며 후자의 사실은 다시금 사실 [A는 S가 가지고 있던 스스로-결정하는 힘의 사례이다]에 근거한다. 더불어 모든 스스로-결정하는 힘의 사례들은 대안가능성을 동반한다. 따라서 다음과 같은 귀결이 따라 나온다: 도덕적 책임 귀속 사실은 대안가능성의 존재를 함축한다; 즉, 대안가능성 없는 도덕적 책임 귀속은 불가능하다.

만약 이러한 넓은 원천 양립불가론자들의 생각이 옳다면 탄탄성 비판에 대한 온건한 답변은 자연스럽게 다음과 같은 함축을 갖게 될 것이다: 스스로-결정하는 힘의 사례들이 동반하는 대안가능성은

“탄탄한 대안가능성”일 필요가 없다. 왜냐하면 넓은 원천 양립불가론에 의하면 도덕적 책임 귀속이 요구하는 대안가능성은 곧 스스로-결정하는 힘의 사례들이 스스로-결정하는 힘의 본성 상 동반하게 되는 대안가능성이며, 더 나아가, 온건한 답변에 의하면 이러한 대안가능성은 “탄탄한 대안가능성”일 필요가 없기 때문이다.

그렇다면 이러한 힘이 과연 넓은 원천 양립불가론자들이 일반적으로 원천-관계의 근거를 이룬다고 생각했던 바로 그러한 종류의 스스로-결정하는 힘일 수 있을까? 물론 그렇게 주장하는 것이 어쩌면 논리적으로 불가능한 것은 아닐지도 모르겠다. 하지만 이렇게 이해된 스스로-결정하는 힘 (이하 ‘탄탄하지 못한 스스로-결정하는 힘’)은 최소한 넓은 원천 양립불가론자들이 탄탄성 비판에 대한 단호한 답변을 채택할 경우에 개입하게 되는 스스로-결정하는 힘 (이하 ‘탄탄한 스스로-결정하는 힘’) 보다는 다소 초라한 것처럼 보인다. 일단 탄탄하지 못한 스스로-결정하는 힘과 탄탄한 스스로-결정하는 힘에 대한 보다 더 정확한 규정은 다음과 같이 주어질 수 있다:

> (탄탄하지 못한 스스로-결정하는 힘) 스스로-결정하는 힘 P는 그것의 사례들이 P의 본성 상 동반하게 되는 대안가능성들이 굳이 탄탄한 대안가능성일 필요가 없을 경우 오직 그 경우에 탄탄하지 못한 스스로-결정하는 힘이다.
> (탄탄한 스스로-결정하는 힘) 스스로-결정하는 힘 P는 그것의 사례들이 P의 본성 상 동반하게 되는 대안가능성들이 반드시 탄탄한 대안가능성이어야 할 경우 오직 그 경우에 탄탄한 스스로-결정하는 힘이다.

물론 우리가 자유 및 도덕적 책임의 근거를 이루는 스스로-결정하는 힘을 이렇게 다소 초라한 방식으로 이해할 경우 몇가지 이론적 장점이 있을 수는 있을 것이다. (예를 들어 이러한 입장을 받아들일 경우 우리는 자유로운 행위자의 외연을 다소 확장할 수 있을 것이다. 이는 실제로 스튜어드(Steward, 2012a)의 기획이 가지고 있는

이론적 동기 중 하나인 것처럼 보이기도 한다.) 하지만 우리가 자신이 자유로운 행위자로서 가지고 있는 스스로-결정하는 힘을 이렇게 초라한 방식으로 이해하기 위해서는 상응하는 대가를 지불해야 한다. 넓은 원천 양립불가론에 의하면 자유 및 도덕적 책임 그리고 그것의 근거를 이루는 원천-관계는 모두 다시금 스스로-결정하는 힘에 근거한다. 따라서 탄탄한 스스로-결정하는 힘을 포기하고 이를 상대적으로 초라한 스스로-결정하는 힘, 즉 탄탄하지 못한 스스로-결정하는 힘으로 대치하는 것은, 결국 스스로-결정하는 힘에 근거하고 있는 원천-관계, 더 나아가 자유 및 도덕적 책임의 본성에 대한 상대적으로 초라한 이해를 받아들인다는 것을 함의한다. 이것이 바로 내가 스튜어드와 같이 온건한 답변을 택하는 넓은 원천 양립불가론자들에게 (사실상 모든 외재적 양방력 이론가들에게) 지적하고 싶었던 바로 그 함의이다. 나는 이러한 함의가 최소한 앞서 언급한 자유로운 행위자의 외연을 확장하려는 동기와 같이 스스로-결정하는 힘에 대한 초라한 이해를 추동하는 몇몇 이론적 동기들을 공유하지 않는 넓은 원천 양립불가론자들에게는 가능하다면 단호한 답변을 온건한 답변보다 선호할 좋은 이유를 제공해준다고 생각한다.

물론 어떤 논변에 부딪혔을 때 해당 논변의 귀결을 부분적으로 수용하는 것은 충분히 있을 수 있는 일이며 때로는 오히려 권장되는 일이기까지도 하다. 하지만 앞서 논변한 것과 같이 탄탄성 비판의 두번째 전제 (2)[57]는 별로 설득력이 없는 것처럼 보이며, 따라서 탄탄성 비판이 과연 탄탄한 대안가능성 원리를 기각하는데 성공했는지 여부 또한 분명치 않다. (나는 앞에서 단호한 답변을 한가지 제시하였으며 이를 통해 우리에게 (2)가 참이라고 믿을 좋은 이유는 존재하지 않는다는 주장을 옹호한 바 있다.) 이러한 상황에서 탄탄한 대안가능성 원리 및 탄탄한 스스로-결정하는 힘을 포기하고 자유 및 도덕적 책임에 대한 상대적으로 초라한 이해를 수용하는 것은 최소한 스스로-결정하는 힘에 대한 이렇게 초라한 이해를 추동하는 몇몇 이론적

[57] (2)의 내용은 다음과 같다: 탄탄한 대안가능성이 존재하지 않음에도 불구하고 도덕적 책임이 이루어지는 경우가 존재한다.

동기들을 공유하지 않는 넓은 원천 양립불가론자들로서는 받아들이기 어려운 타협에 해당할 지도 모른다. 넓은 원천 양립불가론자들은 자유롭게 수행된 행위와 행위자 사이에 성립하는 원천-관계 및 이를 근거짓는 스스로-결정하는 힘에 대해 보다 야심찬 동기를 가지고 있을 수 있다.

넓은 원천 양립불가론자로 간주될 수 있는 케인은 비록 인과적 결정론과 양립불가능하지만 그럼에도 불구하고 여전히 "원할 만한 가치가 있는"(Kane, 1998, p. 15)[58] 자유가 가능하며 이는 다름아닌 "자신의 고유한 목표 혹은 목적의 궁극적 창조자 및 지지자일 수 있는 힘"(Kane, 1998, p. 15)이라고 주장한 바 있다. 케인(Kane, 1998)이 문제의 힘을 이해하는 방식에 따르면 자유로운 행위자는 (적어도 몇몇 경우에 한해서는) 바로 그 상황에서 자신이 실제로 수행했던 행위와는 다른 행위를 수행할 수 있었어야 하며 (케인은 이러한 상황에서 수행된 행위를 '자기-형성적 행위(self-forming action)'라고 부른다) 설사 주어진 상황에서 그 행위자가 다른 행위를 수행하게 되었다 하더라도 이는 다시금 그 자체로 해당 행위자의 자유로운 행위로 간주될 수 있는 그러한 종류의 행위여야 한다. 따라서 그가 "자신의 고유한 목표 혹은 목적의 궁극적 창조자 및 지지자일 수 있는 힘"(Kane, 1998, p. 15)이라고 부르는 힘 혹은 행위자를 자신의 행위의 "궁극적 원천(ultimate source)"으로 만들어 줄 수 있는 힘이라고 부르는 힘은 곧 탄탄한 스스로-결정하는 힘에 해당한다.[59]

---

[58] 케인에 의하면 여기서 사용된 표현 "원할 만한 가치가 있는"은 데넷(Dennett, 1984)에 의해서 사용된 것이기도 하다. 한가지 흥미로운 점은 스튜어드(Steward, 2012a, p. 5)가 "원할 만한 가치가 있는" 자유란 무엇인지를 묻는 질문으로부터 자유의 본성에 대한 탐구를 시작하는 것을 거부한다는 사실이다. 그녀는 우리가 이러한 질문으로부터 자유의 본성에 대한 탐구를 시작할 경우 우리의 자유가 인과적 결정론과 양립불가능한 이유는 다름아닌 우리의 자유를 부분적으로 이루며 또한 동물과 공유하는 어떤 능력이 인과적 결정론과 양립불가능하기 때문일 가능성을 고려하기 어려워진다고 지적한다. 왜냐하면 동물과 공유하는 어떤 능력이 우리가 "원할 만한 가치가 있는" 능력이라는 생각은 손쉽게 떠올리기 어렵기 때문이다.

[59] 물론 케인(Kane, 1998)이 문제의 탄탄한 스스로-결정하는 힘의 본성을 일종

하지만 스튜어드가 주목하고 있는 확정하는 힘은 이러한 힘이 아니다. 확정하는 힘은 케인이 말하고 있는 힘, 즉 "자신의 고유한 목표 혹은 목적의 궁극적 창조자 및 지지자일 수 있는 힘"(Kane, 1998, p. 15)과는 전혀 다른 종류의 힘이다. 확정하는 힘은 무엇보다 외재적 양방력이며 외재적 양방력은 다시금 탄탄한 스스로-결정하는 힘과는 달리 "탄탄한 대안가능성"을 동반하지 않는 그러한 종류의 힘이다. 외재적 양방력이 동반하는 대안가능성은 기껏해야 외재적 양방력의 비-행사 가능성 뿐이며 이는 기껏해야 탄탄하지 못한 대안가능성에 지나지 않는다. 따라서 우리는 다음과 같은 귀결을 얻을 수 있다: 외재적 양방력 이론은 혹여 확정하는 힘 혹은 탄탄하지 못한 스스로-결정하는 힘의 본성을 해명해줄 수는 있을지 몰라도 탄탄한 스스로-결정하는 힘의 본성을 해명해줄 수는 없다. 더 나아가 탄탄한 스스로-결정하는 힘은 여전히 넓은 원천 양립불가론자의 입장에서 "원할 만한 가치가 있는"(Kane, 1998, p. 15) 그러한 종류의 힘일 수 있다. 따라서 우리에게 이러한 힘이 없다고 생각할 좋은 이유가 주어지지 않은 이상 우리는 이러한 힘을 섣불리 포기하기보다는 오히려 이러한 힘이 있기를 희망하고 만약 우리에게 이러한 힘이 있다면 이러한 힘의 본성은 어떻게 이해되어야 할 것인지의 문제를 탐구할 수 있다.[60] 하지만 스튜어드의 외재적 양방력 이론은 이러한 탐구, 즉 탄탄한 스스로-결정하는 힘의 본성을 해명하는 것을 목표로 하는 탐구에는 별다른 도움이 되지 못한다.

물론 스튜어드의 행위자성 이론이 도덕적 책임 귀속을 위해 대안가능성이 요구된다고 생각할 한가지 이유를 제공해줄 수 있을지도 모른다. 하지만 그녀의 행위자성 이론을 통해서 제공되는 이유가

의 양방력으로 이해하고 있는 것은 아니다. 하지만 분명한 점은 그가 자유 및 도덕적 책임의 근거로서 제시하고 있는 자유로운 행위자의 스스로-결정하는 힘은 탄탄하지 못한 스스로-결정하는 힘이라기 보다는 탄탄한 스스로-결정하는 힘이라는 사실이다.

[60] 케인(Kane, 1998) 또한 이러한 이론적 동기 하에 "자신의 고유한 목표 혹은 목적의 궁극적 창조자 및 지지자일 수 있는 힘"(Kane, 1998, p. 15)의 본성을 탐구하는 것을 자신의 목표로 삼고 있다.

대안가능성 원리가 참이라고 생각해야 할 유일한 이유인 것은 아니다.[61] 앞서 논변한 바와 같이 우리는 피셔의 탄탄성 비판에 대해 단호한 답변을 채택할 수 있으며 (즉, 탄탄성 비판의 두번째 전제인 (2)[62]를 거부할 수 있으며), 만약 이러한 답변이 성공적이라면 우리에게는 사실 탄탄한 대안가능성 원리가 거짓이라고 생각할 이유가 전혀 없는 것이다. 따라서 탄탄한 대안가능성 원리가 거짓이라고 생각할 이유가 없는 상황에서 보다 탄탄한 원천-관계 및 자유와 도덕적 책임의 근거를 이루는 탄탄한 스스로-결정하는 힘을 포기한 채 (더불어 이와 결부되어 있는 탄탄한 대안가능성 원리를 포기한 채) 기껏해야 탄탄하지 못한 스스로-결정하는 힘을 옹호하는 것만으로 만족하는 것은 케인과 같은 넓은 원천 양립불가론자가 가지고 있던 이론적 동기를 고려할 때 최소한 어떤 넓은 원천 양립불가론자로서는 도저히 받아들이기 어려운 선택지인 것처럼 보인다.

지금까지 4장에서 이루어진 논의를 정리해보자. 먼저 넓은 원천 양립불가론자는 대안가능성 원리에 개입하기 때문에 프랑크푸르트-유형 사례로부터 제기되는 문제에 대해 그럴듯한 답변을 제시해주어야 할 의무를 갖는다. 왜냐하면 프랑크푸르트-유형 사례는 달리 행할 수 없었던 행위에 대해서도 도덕적 책임이 귀속될 수 있음을 보여주는 것처럼 보이는 사례이기 때문이다.

이에 대해 즉각적으로 제시해볼 수 있는 답변은 소위 "스쳐 지나가는 자유"에 호소하는 답변이다. 하지만 피셔는 다음과 같이 주장한다: 만약 도덕적 책임이 대안가능성의 존재를 요구한다면 이때 요구되는 대안가능성은 "탄탄한 대안가능성"일 것이며, 따라서 프랑크푸르트-유형 사례로부터 제기되는 문제를 만족스럽게 해결하기 위해 해당 사례에서 찾아내야 하는 대안가능성은 "스쳐 지나가는 자유"가 아니라 실제로 수행된 것과는 다른 종류의 행위를 자유롭게

---

[61] 스튜어드의 답변이 가지고 있는 이러한 한계는 한성일 교수님께서 지적해주신 것이다.

[62] (2)의 내용은 다음과 같다: 탄탄한 대안가능성이 존재하지 않음에도 불구하고 도덕적 책임이 이루어지는 경우가 존재한다.

수행할 수 있었던 가능성, 즉 “탄탄한 대안가능성”이다.

이러한 피셔의 탄탄성 비판에 대해 성공적으로 답변하기 위해서는 프랑크푸르트-유형 사례의 존스에게는 여전히 “탄탄한 대안가능성”이 존재했음을 설득력 있게 보이거나 혹은 설사 도덕적 책임 귀속이 대안가능성을 요구한다 하더라도 이때 요구되는 대안가능성이 꼭 “탄탄한 대안가능성”일 필요는 없음을 설득력 있게 보여주어야 한다. 본 논문에서는 전자의 방식을 취하는 답변을 ‘단호한 답변’이라 부르고 오직 후자의 방식 만을 취하는 답변을 ‘온건한 답변’이라고 불렀다. 하지만 외재적 양방력 이론가들에게 전자의 과제는 원칙적으로 달성할 수 없는 그러한 종류의 것이다. 왜냐하면 외재적 양방력 이론가들이 스스로-결정하는 힘과 결부되어 있다고 주장하는 대안가능성은 기껏해야 외재적 양방력의 비-행사 가능성에 불과하며 이는 “탄탄한 대안가능성”일 수 없기 때문이다. 따라서 외재적 양방력 이론가가 탄탄성 비판에 대해 답변하기 위해 채택할 수 있는 방식은 오직 후자의 방식, 즉 설사 도덕적 책임 귀속이 대안가능성을 요구한다 하더라도 이때 요구되는 대안가능성이 꼭 “탄탄한 대안가능성”일 필요는 없음을 보이는 방식 뿐이다. 달리 말해 외재적 양방력 이론가는 탄탄성 비판에 대한 단호한 답변을 제공할 수 없으며, 그녀들이 탄탄성 비판에 대해 제시할 수 있는 답변은 오직 온건한 답변 뿐이다. (그리고 본 논문에서는 온건한 답변이 제시될 수 있는 방식을 다시금 두가지로 구분하여 각각 소개하였다.)

하지만 넓은 원천 양립불가론자의 입장에서 탄탄성 비판에 대한 단호한 답변을 포기하고 온건한 답변을 제시하는 것으로 만족하는 것은 (그것이 온건한 답변 [1]에 해당하든 아니면 온건한 답변 [2]에 해당하든) 자유 및 도덕적 책임을 근거짓는 원천-관계 및 스스로-결정하는 힘에 대한 보다 탄탄한 이해를 포기한 채 상대적으로 초라한 이해만으로 만족하는 것을 함의한다. 이러한 함의는 최소한 몇몇 넓은 원천 양립불가론자들에게는 그 자체로 외재적 양방력 이론을 포기할 충분한 이유를 제공해줄 수도 있다. 하지만 이러한 타협은 넓은 원천 양립불가론을 채택하기 위해 반드시 받아들여야 하는 그러한 종류의

타협도 아니고 오직 외재적 양방력을 받아들일 경우에만 넓은 원천 양립불가론자들에게 강제되는 그러한 종류의 타협이다. 따라서 프랑크푸르트-유형 사례로부터 제기되는 문제에 대한 고려는 최소한 이러한 타협을 반기지 않는 넓은 원천 양립불가론자들에게는 자유 및 도덕적 책임의 근거를 이루는 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 채택하지 말아야 할 하나의 좋은 이유를 제공해줄 수 있을 것이다.

제 5 장 스스로-결정하는 힘과의 본성적 긴장

앞서 설명한 바와 같이 넓은 원천 양립불가론자들은 다음과 같은 원천 원리를 받아들인다: 행위자 S는 오직 자신이 수행한 행위 A의 원천인 한에서 A에 대해 도덕적 책임을 진다. 하지만 넓은 원천 양립불가론자들에 의하면 인과적 결정론이 참인 세계에서 행위자는 자신이 수행한 어떠한 행위에 대해서든 그것의 원천일 수 없으며 바로 이러한 이유 때문에 자유 및 도덕적 책임은 인과적 결정론과 양립할 수 없다. 그렇다면 왜 인과적 결정론이 참인 세계에서는 행위자와 수행된 행위 사이에 원천-관계가 성립할 수 없는가? 넓은 원천 양립불가론자들에 의하면 행위자 S는 오직 A를 수행함에 있어서 달리 행할 수 있었던 경우에 한하여 자신이 수행한 행위 A의 원천일 수 있다. 더불어 인과적 결정론이 참인 세계에서는 어떠한 대안가능성도 있을 수 없다. 따라서 행위자는 인과적 결정론이 참인 세계에서 자신이 수행한 어떠한 행위에 대해서도 그것의 원천일 수 없다. 이것이 바로 넓은 원천 양립불가론자들이 인과적 결정론이 참인 세계에서 행위자와 수행된 행위 사이에 원천-관계가 성립할 수 없다고 생각하는 이유이다.

그렇다면 넓은 원천 양립불가론자들이 생각할 때 인과적 결정론이 참인 세계에서 자유 및 도덕적 책임이 있을 수 없는 이유는 결국 원천-관계의 성립을 위해 대안가능성의 존재가 요구되기 때문일 것이다. 하지만 도대체 왜 원천-관계의 성립을 위해 대안가능성의 존재가 요구되는 것인가? 이에 대해 넓은 원천 양립불가론자들은 다음과 같은 설명을 제시한다: [S는 자신이 수행한 행위 A의 원천이다]는 [A는 S의 스스로-결정하는 힘의 사례이다]에 근거하며 후자의 사실로부터 다시금 [S는 A를 수행함에 있어서 달리 행할 수 있었다]가 따라 나오기 때문이다. 즉 스스로-결정하는 힘의 모든 사례들은 그것의 본성 상 항상 어떤 대안가능성을 동반하게 되며 (넓은 원천 양립불가론자들에 의하면) 바로 이러한 사실이 결국 원천-관계의 성립을 위해 대안가능성의 존재가 요구되는 이유인 것이다.

그렇다면 다시금 스스로-결정하는 힘은 도대체 어떤 독특한

종류의 힘이길래 문제의 힘의 사례들이 항상 어떤 대안가능성을 동반하게 되는 것인가? 앞서 살펴보았듯이 외재적 양방력 이론은 이에 대한 한가지 가능한 답변을 제공해주는 것처럼 보였다: 그것은 바로 스스로-결정하는 힘이 외재적 양방력의 일종이기 때문이다. 하지만 나는 최소한 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘은 그것의 본성 상 외재적 양방력과 같은 힘과는 동일시되기 어렵다고 생각한다. 지금부터 나는 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘이 외재적 양방력과 동일시될 경우 이로부터 우리가 손쉽게 받아들이기 어려운 어떤 귀결이 따라 나옴을 보일 것이다. 만약 지금부터 제시될 나의 논변이 성공적이라면 이는 최소한 문제의 귀결을 수용하기를 주저하는 넓은 원천 양립불가론자들에게 있어서는 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 채택하지 않을 하나의 좋은 이유를 제공해줄 수 있을 것이다.

먼저 자유 및 도덕적 책임의 근거를 제공해주는 스스로-결정하는 힘에 대한 넓은 원천 양립불가론적 이해 방식이 정확히 무엇인지를 분명히 해보도록 하자. 앞서 설명한 것처럼 넓은 원천 양립불가론자들은 스스로-결정하는 힘을 무엇보다도 다음과 같은 특징을 갖는 힘으로 이해한다: 스스로-결정하는 힘의 모든 사례들은 문제의 힘의 본성 상 항상 어떤 대안가능성을 동반하게 된다. 예를 들어 무단횡단을 한 영희가 만약 무단횡단을 함에 있어서 그녀의 어떤 스스로-결정하는 힘을 행사했다면 이로부터 다음과 같은 귀결이 따라 나온다: 주어진 상황에서 영희에게는 어떤 대안가능성이 존재했었다. (더불어 이러한 귀결-관계는 다름아닌 스스로-결정하는 힘의 본성 덕분에 성립하는 사실이다.) 이렇게 실제로 성립했던 사실 [영희는 무단횡단을 했다]에 의해 동반되었던 대안가능성에 해당하는 대안적 사실[63] 을 'Q'라고 부르도록 하자.[64] 또한 해당 사례에서 실제로

[63] 물론 대안적 사실은 엄밀히 말해 사실은 아니다. 왜냐하면 대안적 사실은 실제로 성립한 사실이 아니기 때문이다.

[64] 외재적 양방력 이론가들은 아마도 Q를 [영희는 무단횡단을 하지 않았다] 혹

성립했던 사실 [영희는 무단횡단을 했다]는 'P'라고 부르도록 하자. 그렇다면 이제 문제의 사례에서 영희가 가지고 있었던 스스로-결정하는 힘은 "P가 성립할지 아니면 Q가 성립할지 여부"를 스스로-결정하는 힘으로 이해된다. 혹은 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정하는 힘으로 이해된다. 이러한 방식으로 이해된 스스로-결정하는 힘을 '넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘'이라고 부르도록 하자.

그런데 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘은 다음과 같은 의미에서 "대칭적"인 것처럼 보인다: Q가 실현된 경우는 P가 실현된 경우 만큼이나 행위자가 "P가 성립할지 아니면 Q가 성립할지 여부"를 스스로-결정한 경우에 해당한다; 혹은 Q가 실현된 경우는 P가 실현된 경우 만큼이나 행위자가 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정한 경우에 해당한다.[65] 하지만 행위자가 "P가 성립할지 아니면 Q가 성립할지 여부"를 스스로-결정한 경우를 행위자가 "P가 성립할지 아니면 Q가 성립할지 여부"를 스스로-결정하는 힘을 행사하지 않은 경우라고 할 수 있는가? 혹은 행위자가 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정한 경우를 행위자가 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정하는 힘을 행사하지 않은 경우라고 할 수 있는가? 이러한 주장은 항생제가 세균을 죽인 경우가 항생제가 가지고 있던 세균을 죽이는 힘이 행사되지 않은 경우에 해당한다는 주장 만큼이나 받아들이기 어려운 주장이다. 따라서 우리는 다음과 같은 귀결을 얻을 수 있다: Q가 실현된 경우는 P가 실현된 경우 만큼이나 "P가 성립할지 아니면 Q가 성립할지 여부" 혹은 "P와 Q 중에서 무엇이 실현될지 여부"를 스스로-

---

은 [영희는 무단횡단을 하거나 하지 않을 수 있는 외재적 양방력을 비-행사하였다]와 같은 사실들로 이해할 것이다.

[65] 물론 우리는 동일한 내용을 다음과 같은 방식으로 말할 수도 있을 것이다: P가 실현된 경우는 Q가 실현된 경우 만큼이나 행위자가 "P가 성립할지 아니면 Q가 성립할지 말지 여부"를 스스로 결정한 경우에 해당한다; 혹은 P가 실현된 경우는 Q가 실현된 경우 만큼이나 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정한 경우에 해당한다.

결정하는 힘이 “행사”된 경우이다.

그렇다면 넓은 원천 양립불가론자들이 자유 및 도덕적 책임의 근거를 이루는 힘으로 상정하고 있는 스스로-결정하는 힘은 분명 외재적 양방력의 일종으로 이해될 수 없을 것이다. 왜냐하면 문제의 스스로-결정하는 힘은 무엇보다도 그것의 본성 상 그것의 모든 사례들이 반드시 동반하게 되는 대안가능성 또한 그것이 (또 다른 방식으로) 행사될 수 있었던 그러한 종류의 가능성으로 이해되어야 하는 힘인 것처럼 보이기 때문이다. Q가 실현된 경우 또한 P가 실현된 경우와 마찬가지로 “P와 Q 중에서 어떤 것이 성립할지 여부”를 행위자가 스스로-결정한 경우에 해당한다. 더불어 “P와 Q 중에서 어떤 것이 성립할지 여부”를 행위자가 스스로-결정한 경우는 다시금 행위자가 “P와 Q 중에서 어떤 것이 성립할지 여부”를 스스로-결정하는 힘을 행사한 경우로 이해되어야 한다. 왜냐하면 “P와 Q 중에서 어떤 것이 성립할지 여부”를 행위자가 스스로-결정한 경우가 “P와 Q 중에서 어떤 것이 성립할지 여부”를 스스로-결정하는 힘이 행사되지 않은 경우라고 주장하는 것은 항생제가 세균을 죽인 경우가 항생제가 가지고 있던 세균을 죽이는 힘이 행사되지 않은 경우에 해당한다는 주장만큼이나 받아들이기 어려운 주장이기 때문이다. 그렇다면 Q가 실현된 경우 또한 문제의 스스로-결정하는 힘이 행사된 경우로 이해될 수밖에 없다.

하지만 앞서 설명된 바와 같이 외재적 양방력의 양방향성은 행사와 비-행사 사이의 양방향성이다. 즉, 외재적-양방력의 사례들이 동반하는 대안가능성은 그것이 다른 방식으로 행사될 수 있었던 가능성이라기보다는 그것이 비-행사될 수 있었던 가능성이다. 따라서 외재적 양방력은 본성 상, 앞서 설명된 것과 같이 스스로-결정하는 힘이라면 반드시 만족해야 하는 어떤 조건, 즉 그것의 사례들이 동반하는 대안가능성 또한 그것이 (또 다른 방식으로) 행사될 수 있었던 가능성으로 이해되어야 한다는 조건을 만족할 수 없는 종류의 힘이다. 분명 이는 외재적 양방력 이론가들에게 심각한 문제를 불러일으킨다. 왜냐하면 이는 그녀들이 결국 행위자에 의해 “P와 Q 중에서 어떤 것이

성립할지 여부”가 스스로-결정된 경우가 곧 행위자가 소유하던 “P와 Q 중에서 어떤 것이 성립할지 여부”를 스스로-결정하는 힘이 비-행사된 경우일 수 있다는 받아들이기 어려운 주장에 개입하는 것을 피할 수 없음을 함축하기 때문이다.

방금 제시된 나의 논변에 대해 외재적 양방력 이론가는 단지, 말하자면, 영희가 “무단횡단을 할지 말지 여부”를 스스로-결정했다는 사실로부터 영희가 “무단횡단을 할지 말지 여부”를 스스로-결정하는 힘을 행사했음이 따라 나오는 것은 아니라고 답변하려 할지도 모르겠다. 이러한 답변은 다음과 같이 진행될 수 있다. 영희가 “무단횡단을 할지 말지 여부”를 스스로-결정했다는 사실로부터 영희가 “무단횡단을 할지 말지 여부”를 스스로-결정하는 힘을 행사했음이 따라 나오기 위해서는 다음과 같은 원리가 필요하다: 임의의 행위자 S 및 S에 대해 참되게 적용될 수 있는 어떤 넓은 의미에서 행위와 관련된 술어 X에 대해, 만약 S가 X했다면 S는 X하는 힘을 행사한 것이다. 하지만 이러한 원리는 거짓이다. 다음과 같은 사례를 생각해보자. 철수는 어느 날 횡단보도에 서서 빨간 불이 켜진 것을 보았다. 바쁜 일이 었었던 철수는 잠시 무단횡단을 할지 말지 여부에 대해 고민했지만 결국 무단횡단을 하지 않았으며 이를 통해 교통법규를 준수했다. 이때 분명 철수는 교통법규를 준수했지만 그가 교통법규를 준수하는 힘을 행사한 것은 아니다. 철수는 여기서 단지 무단횡단을 하는 힘을 행사하지 않음으로써 교통법규를 준수한 것이지 특별히 교통법규를 준수하는 힘을 행사한 것이 아니다. 만약 문제의 상황에 대한 이러한 해석이 옳다면 분명 철수는 교통법규를 준수했지만 그럼에도 불구하고 이로부터 철수가 교통법규를 준수하는 힘을 행사했다는 귀결이 따라 나오는 것은 아니다. 따라서, 임의의 행위자 S 및 S에 대해 참되게 적용될 수 있는 어떤 넓은 의미에서 행위와 관련된 술어 X에 대해, 만약 S가 X했다면 S는 X하는 힘을 행사한 것이라는 원리는 거짓이다. 더불어 문제의 원리가 거짓이라면 마찬가지로 영희가 “무단횡단을 할지 말지 여부”를 스스로-결정했다는 사실로부터 영희가 “무단횡단을 할지 말지 여부”를 스스로-결정하는 힘을 행사했음이 따라 나오지 않는다.

그리고 앞서 제시된 나의 논변은 실패한다.

하지만 나는 이러한 답변은 실패한다고 생각한다. 임의의 행위자 S 및 S에 대해 참되게 적용될 수 있는 어떤 넓은 의미에서 행위와 관련된 술어 X에 대해, 만약 S가 X했다면 S는 X하는 힘을 행사한 것이라는 원리는 거짓이라는 주장에 대해서는 나도 동의한다. 하지만 내가 영희가 "무단횡단을 할지 말지 여부"를 스스로-결정했다는 사실로부터 영희가 "무단횡단을 할지 말지 여부"를 스스로-결정하는 힘을 행사했음이 따라 나온다고 생각할 때 가정했던 원리는 이보다는 다소 제한된 종류의 원리이다. 내가 가정했던 원리는 다음과 같은 원리이다: 임의의 행위자 S 및 S에 대해 참되게 적용될 수 있는 어떤 넓은 의미에서 행위와 관련된 술어 X에 대해, 만약 S가 X하는 힘을 가지고 있었으며 또한 S가 X를 했으면 S는 X하는 힘을 행사한 것이다. 앞서 제시되었던 철수의 사례에서 철수가 교통법규를 준수하는 힘을 별도로 가지고 있었던 것은 아닌 것처럼 보인다. 따라서 문제의 사례는 내가 실제로 가정했던 다소 제한된 종류의 원리에 대해서는 특별히 반례가 되지 않는다. 하지만 영희가 "무단횡단을 할지 말지 여부"를 스스로-결정한 경우는 무엇보다도 그녀가 "무단횡단을 할지 말지 여부"를 스스로-결정하는 힘을 가지고 있었던 경우로 가정된다. 따라서 이 경우에는 철수의 사례와는 달리 영희가 "무단횡단을 할지 말지 여부"를 스스로-결정했다는 사실로부터 영희가 "무단횡단을 할지 말지 여부"를 스스로-결정하는 힘을 행사했다는 귀결이 따라 나온다.

더 나아가 철수의 사례에서 우리가 철수에게 교통법규를 준수하는 힘이 있었다고 가정한다 하더라도 이는 여전히 내가 실제로 가정했던 다소 제한된 종류의 원리에 대한 반례가 되지 못한다. 왜냐하면 문제의 사례가 철수의 교통법규를 준수하는 힘이 행사되지 않은 경우에 해당한다는 우리의 직관은 애초에 철수에게 교통법규를 준수하는 힘이 없었다는 암묵적인 가정에 의존하고 있는 것처럼 보이기 때문이다. 이번에는 문제의 사례에서 철수에게 교통법규를 준수하는 힘이 있었다고 가정해보자. 그럼에도 여전히 우리에게 여전히 문제의 사례가 철수가 교통법규를 준수하는 힘을 행사하지 않은 경우에 해당하는

사례라는 직관이 남아있는가? 그렇지는 않은 것처럼 보인다. 우리가 문제의 사례에서 철수에게 교통법규를 준수하는 힘이 있었다고 가정할 경우 이제 철수는 교통법규를 준수하는 힘을 행사한 것으로 이해되어야 하는 것처럼 보인다. 따라서 여전히 문제의 사례는 앞서 내가 실제로 가정했던 다소 제한된 종류의 원리에 대한 반례가 되지 못하는 것처럼 보인다.

마지막으로 이러한 나의 답변이 혹시 임시방편적인(ad hoc) 것은 아닌지에 대해 간략히 검토해보자. 최소한 내가 볼 때는 임의의 행위자 S 및 S에 대해 참되게 적용될 수 있는 어떤 넓은 의미에서 행위와 관련된 술어 X에 대해, 만약 S가 X했다면 S는 X하는 힘을 행사한 것이라는 원리 대신 내가 가정했던 다소 제한된 종류의 원리 (즉 S가 X하는 힘을 가져야 한다는 조건을 추가한 원리)를 상정하는 것이 임시방편적인 시도로 간주될 이유는 특별히 없는 것처럼 보인다. 왜냐하면 X하는 힘에 대한 소유는 X하는 힘을 행사하기 위한 최소한의 필요조건이기 때문이다. 그리고 X하는 힘에 대한 소유가 X하는 힘을 행사하기 위한 필요조건에 해당한다는 우리의 생각이 현재 쟁점이 되고 있는 사안과 관련하여 특별히 어떤 논점선취적인 직관에 의존하고 있는 것 또한 아닌 것처럼 보인다. 따라서 여기서 내가 어떤 제한적인 원리에 호소하여 현재 고려되고 있는 외재적 양방력 이론가의 답변이 실패함을 논변한다 하더라도 이러한 나의 논변이 임시방편적인 시도에 불과하다고 생각할 이유는 없는 것처럼 보인다.

그렇다면 결국 우리는 다음과 같은 귀결에 도달한 것처럼 보인다: 스스로-결정하는 힘을 외재적 양방력의 일종으로 이해하는 넓은 원천 양립불가론자들은 모두 “P와 Q 중에 어떤 것이 성립할지 여부”를 행위자가 스스로-결정한 경우가 곧 행위자가 소유하던 “P와 Q 중에 어떤 것이 성립할지 여부”를 스스로-결정하는 힘이 비-행사된 경우라는 주장에 개입하게 된다. 그리고 이러한 주장은 넓은 원천 양립불가론자의 입장에서 손쉽게 받아들일 수 있는 주장이 아니다. 따라서 이는 최소한 문제의 귀결을 받아들이기를 주저하는 넓은 원천 양립불가론자들에게 있어서는 양방력으로 이해된 스스로-결정하는

힘의 본성과 관련하여 외재적 양방력 이론을 거부할 한가지 좋은 이유를 제공해줄 수 있을 것이다.

더불어, 만약 이러한 나의 생각이 옳다면, 앞서 탄탄성 비판에 대해 외재적 양방력 이론가의 입장에서 제시해볼 수 있는 답변으로 소개되었던 (그리고 앞서 지적되었던 것과 같이 탄탄성 비판에 대한 넓은 원천 양립불가론자의 답변으로서 어떤 한계를 가지고 있었던) 온건한 답변 [1]이 심지어 외재적 양방력 이론가들에 의해 제시되었을 때 정말로 탄탄성 비판에 대한 설득력있는 답변일 수 있을 것인지 여부를 다시금 재고해볼 필요가 있는 것처럼 보인다. 앞서 설명했듯이, 온건한 답변 [1]에 의하면 물론 도덕적 책임 귀속을 위해 대안가능성의 존재가 요구되는 것은 맞지만 이때 대안가능성의 존재는 그것이 도덕적 책임 귀속 사실을 근거짓기 때문에 요구되는 것이라기 보다는 기껏해야 도덕적 책임 귀속 사실에 의해 함축되는 사실에 불과했다. 그리고 이는 탄탄성 비판의 첫번째 전제인 (1), 즉 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 이때 요구되는 대안가능성은 "탄탄한 대안가능성"일 것이라는 주장에 대해 피셔가 제시했던 논변을 거부할 수 있는 (하지만 그 자체로는 어떤 한계를 가지고 있는) 한가지 방안 (즉, (4)[66]에 대한 거부 방안)을 마련해주는 것처럼 보였다.

하지만 만약 애초에 여기서 스스로-결정하는 힘의 본성이 외재적 양방력의 일종으로 이해될 수 없는 그러한 종류의 것이라면, 온건한 답변 [1]이 외재적 양방력 이론가의 입장에서 제시되었을 때 과연 여전히 탄탄성 비판에 대한 설득력 있는 답변으로 남아있을 수 있는 것인지 여부 또한 불분명해지는 것처럼 보인다. 왜냐하면 온건한 답변 [1]은 궁극적으로 스스로-결정하는 힘의 모든 사례들은 그것의 본성 상 어떤 대안가능성을 동반하게 된다는 주장에 의존하고 있으며 외재적 양방력 이론 (즉, 스스로-결정하는 힘은 곧 외재적 양방력이라는 주장)은 다시금 바로 이러한 주장이 도대체 왜 성립하는지를 묻는 질문에 대한 답변의 역할을 담당하고 있기 때문이다. 따라서 애초에

[66] (4)의 내용은 다음과 같다: 만약 도덕적 책임 귀속이 대안가능성을 요구한다면 이는 대안가능성이 도덕적 책임 귀속을 근거짓기 때문일 것이다.

스스로-결정하는 힘이 외재적 양방력과 동일시될 수 없다면 외재적 양방력 이론가는 스스로-결정하는 힘의 모든 사례들은 그것의 본성 상 어떤 대안가능성을 동반하게 된다는 주장이 왜 성립하게 되는지를 더이상 설명할 수 없다. 그리고 이러한 주장이 왜 성립하게 되는지를 설명할 수 없게 된다는 것은 곧 온건한 답변 [1]이 더이상 피셔의 탄탄성 비판에 대한 만족스러운 답변일 수 없음을 함축하는 것처럼 보인다. 왜냐하면 결국 이를 설명할 수 없다는 것은 대안가능성의 존재가 도덕적 책임 귀속 사실의 근거를 이루는 것이 아님에도 불구하고 도대체 왜 도덕적 책임 귀속을 위해 요구되는 것인지를 설명할 수 없다는 것을 의미하기 때문이다. 따라서 5장에서 제시된 나의 논변이 성공적이라면, 이로부터 온건한 답변 [1]은 외재적 양방력 이론가의 입장에서 제시되었을 경우 더이상 탄탄성 비판에 대한 설득력 있는 답변이기 어렵다는 귀결이 따라 나오는 것처럼 보인다.

5장의 내용을 정리해보도록 하자. 행위자 S에 의해 자유롭게 수행된 행위 A에 대해 [S가 A를 했다]를 'P'라고 부르자. 더불어 P가 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘의 사례이기 때문에 반드시 동반하게 되는 대안가능성에 대응하는 대안적 사실을 'Q'라고 부르도록 하자. 이때 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘이란 다름아닌 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정하는 힘이다. 그렇다면 대안적 사실 Q가 실현된 경우는 P가 실현된 경우와 마찬가지로 "P와 Q 중에서 어떤 것이 성립할지 여부"가 스스로-결정된 경우에 해당한다. 따라서 대안적 사실 Q가 실현된 경우는 P가 실현된 경우와 마찬가지로 "P와 Q 중에서 어떤 것이 성립할지 여부"를 스스로-결정하는 힘이 행사된 경우에 해당한다. 하지만 만약 문제의 스스로-결정하는 힘이 외재적 양방력이라면 이는 불가능하다. 왜냐하면 외재적 양방력의 모든 사례들이 그것의 본성 상 동반하게 되는 대안가능성은 다름아닌 문제의 외재적 양방력이 행사되지 않을 수 있었던 가능성에 불과하기 때문이다. 더불어 만약 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘이 외재적 양방력과 동일시될 수 없다면, 이로부터 다음과

같은 결론 또한 따라 나온다: 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 채택하는 이들은 더이상 4장에서 언급되었던 피셔의 탄탄성 비판에 대한 온건한 답변 [1]을 이용할 수 없다. 이러한 고려사항들은 결국 넓은 원천 양립불가론자들에게 스스로-결정하는 힘의 본성에 관하여 외재적 양방력 이론을 거부할 좋은 이론적 동기를 제공해줄 수 있을 것이다.

제 6 장 즉자적 힘과의 구분 문제

넓은 원천 양립불가론자들은 일반적으로 스스로-결정하는 힘을 자유 및 도덕적 책임과 유관한 통제력의 원천으로 이해한다. 따라서 스스로-결정하는 힘을 일종의 양방력으로 이해한다는 것은 곧 자유 및 도덕적 책임과 유관한 통제력을 문제의 양방력을 통해 확보되는 통제력과 동일시한다는 것을 의미한다. 이러한 사실은 외재적 양방력 이론을 채택하는 넓은 원천 양립불가론자들에게 다시금 어떤 심각한 문제를 야기한다. 프로스트(Frost, 2020)는 만약 외재적 양방력 이론이 옳다면 이로부터 다음과 같은 귀결이 따라 나옴을 지적한다: 라듐 원자가 가지고 있는 방사선을 방출하는 힘과 같은 즉자적 힘(spontaneous power) 또한 마찬가지로 양방력이다. 라듐 원자가 가지고 있는 방사선을 방출하는 힘은 자두를 먹거나 먹지 않을 수 있는 힘과 마찬가지로 발현의 충분조건으로서의 발현-조건을 결여하고 있는 힘이다. 따라서 라듐 원자가 방사선을 방출한 경우 이는 일어나지 않을 수 없었던 그러한 종류의 일이 아니다. 즉 라듐 원자가 가지고 있는 방사선을 방출하는 힘은 항상 발현되지 않을 수 있었음에도 불구하고 발현되는 그러한 종류의 방식으로만 발현될 수 있는 힘이다. 이는 외재적 양방력 이론가들이 제시하는 외재적 양방력의 정의를 정확히 만족시킨다. 따라서 외재적 양방력 이론을 받아들일 경우 넓은 원천 양립불가론자들에게는 다음과 같은 문제가 발생한다: 양방력은 즉자적 힘과 구분되지 않는다.

물론 이는 일견 그렇게 심각한 문제로 보이지 않을 수도 있다. 왜냐하면 외재적 양방력 이론가는 외재적 양방력을 소유하는 것이 자유로운 행위자성을 갖기 위한 필요조건인 것은 맞지만 충분조건인 것은 아니라고 주장함으로써 라듐 원자가 갖는 방사성 붕괴를 통해 방사선을 방출할 수 있는 힘 또한 자두를 먹거나 먹지 않을 수 있는 힘과 마찬가지로 외재적 양방력이라는 결론을 그냥 끌어안을 수 있는 것처럼 보이기 때문이다. 라듐 원자는 외재적 양방력을 소유하긴

하지만 자유로운 행위자성의 또 다른 어떤 필요조건[67]을 결여하기 때문에 자유로운 행위자가 아닌 것이다. 단지 외재적 양방력이 생물(animate entity) 뿐 아니라 비생물(inanimate entity)에 의해서도 소유된다는 이유로 외재적 양방력 이론이 어떤 어려움을 겪는다고 생각할 이유는 없다.[68]

하지만 외재적 양방력 이론이 라듐 원자와 같은 비생물에 의해 소유되는 힘 또한 양방력일 수 있다는 귀결을 갖는다는 사실은 보다 심각하게 고려될 필요가 있다. 왜 그런가? 외재적 양방력 이론가들은 최소한 외재적 양방력이 자유로운 행위자성의 필요조건이라고 주장한다. 따라서 다음과 같은 질문이 자연스럽게 제기된다: 외재적 양방력은 왜 자유로운 행위자성의 필요조건인가? 달리 말해, 외재적 양방력을 갖지 않으면 왜 자유로운 행위자일 수 없는 것인가? 물론 이에 대해 외재적 양방력 이론가들이 가지고 있는 생각은 조금씩 다를 수 있을 것이다. 하지만 앞서 살펴본 바와 같이 외재적 양방력 이론을 받아들이는 넓은 원천 양립불가론자들에게 외재적 양방력은 결국 자유 및 도덕적 책임과 유관한 통제력을 이루는 힘 혹은 스스로-결정하는 힘으로 가정된다.

---

[67] 예를 들면 자유로운 행위자성은 외재적 양방력과 합리성 모두를 필요조건으로 요구하며 라듐 원자는 외재적 양방력은 갖지만 합리성은 결여하기 때문에 자유로운 행위자가 아니라고 생각하는 것이 충분히 가능해 보인다. (이러한 외재적 양방력 이론가의 답변 가능성은 한성일 교수님께서 지적해주신 것이다.) 더불어 이와 정확히 일치하는 것은 아니지만, 그리고 비록 그가 외재적 양방력 이론가인 것 또한 아니지만, 유사한 생각이 로우에 의해서도 이미 제시된 바 있다. 로우(Lowe, 2013, pp. 176-178)는 인간 행위자의 의지가 라듐 원자의 즉자적 힘과는 달리 다음의 두가지 특징을 갖는 힘이며 바로 이러한 사실에 근거하여 라듐 원자의 즉자적 힘과 구분될 수 있다고 주장한다: (i) 의지는 어떤 행위 X에 대해 **X하기를 의지함**과 **X하지 않기를 의지함** (혹은 **X하지 않으려 함**(refraining from doing X)) 모두에 대해 각각 그것에 해당하는 방식으로 발현될 수 있는 힘이다; (ii) 의지는 또한 "이유-반응적인 방식으로 행사되는 힘"(Lowe, 2013, p. 177)이다. 이중 (ii)는 앞서 언급된 합리성 또한 자유로운 행위자성의 필요조건이라는 생각과 유사한 생각을 반영하고 있는 것처럼 보인다.

[68] 반 밀텐버그와 오메토(van Miltenburg & Ometto, 2020) 또한 여기서 제기되고 있는 문제가 본 논문에서 '외재적 양방력 이론'이라 부르는 종류의 양방력 이론에 대해 어떤 결정적인 문제를 일으킨다고 보지 않는다.

이것이 바로 그들이 외재적 양방력을 자유로운 행위자성의 필요조건이라고 생각하는 이유이다. 예를 들어 철수가 행위 **길을 건넘**을 수행했다고 해보자. 이때 철수는 자신이 소유한 외재적 양방력, 즉 길을 건너거나 건너지 않을 수 있는 힘을 행사했다고 할 수 있으며 이와 관련하여 알바레즈(Alvarez, 2009, 2013)는 다음과 같이 주장한다: 주어진 상황에서 철수가 자신이 소유한 외재적 양방력을 행사했다는 사실 덕분에 [철수는 길을 건넜다]가 성립할지 말지 여부가 철수에게 달려 있게 된다. (즉 철수는 자신의 행위 **길을 건넘**의 원천일 수 있게 된다.) 그렇다면 일반적으로 말해 발현-유형 M을 자신의 발현-유형 정체성으로 갖는 외재적 양방력 P는, 비록 자유로운 행위자성을 갖기 위한 충분조건은 아닐지 몰라도, 다음과 같은 특징을 갖는 힘으로 이해되고 있는 것이다: 다름아닌 P의 소유자가 P를 행사했다는 사실 덕분에 M에 해당하는 사건이 발생할지 말지 여부는 P의 소유자에게 달려 있을 수 있다. (물론 여기서 사용된 개념 〈달려 있다〉는 자유 및 도덕적 책임과 유관한 개념으로 이해되어야 할 것이며 이는 다시금 자유 및 도덕적 책임을 근거짓는 원천-관계를 포착하는 또 다른 개념으로 이해될 수 있을 것이다: A를 수행할지 말지 여부가 S에게 달려 있을 경우 오직 그 경우에 S는 A의 원천이다. 또한 여기서 사용된 개념 〈달려 있다〉는 다시금 자유 및 도덕적 책임과 유관한 통제력을 포착하는 또 다른 개념으로 이해될 수 있을 것이다: A를 수행할지 말지 여부가 S에게 달려 있을 경우 오직 그 경우에 S는 A에 대해 자유 및 도덕적 책임과 유관한 종류의 통제력을 갖는다.)

그렇다면 라듐 원자와 같은 비생물에 의해 소유되는 힘 또한 양방력일 수 있다는 귀결은 생각보다 문제적일 수 있다. 왜냐하면 이는 라듐 원자의 즉자적 힘이 행사된 경우 우리는 발생한 사건 **방사성 붕괴를 통해 방사선이 방출됨**에 대해, 문제의 사건이 발생할지 말지 여부는 즉자적 힘의 소유자, 즉 라듐 원자에게 "자유 및 도덕적 책임과 유관한 의미에서" 달려 있었다고 말해야 함을 함축하기 때문이다. 혹은 라듐 원자는 문제의 사건이 발생할지 말지 여부에 대해 "자유 및 도덕적 책임과 유관한" 통제력을 가지고 있었다고 말해야 함을

함축하기 때문이다. 이는 분명 문제적인 귀결이다. 스튜어드의 표현을 빌리자면 "[방사선] 방출이 일어났을 때 라듐을 (단지 무작위로 시점이 정해진(timed) 하나의 가능성이 실현되는데 있어 수동적으로 참여한 것에 반대되는 것으로서) 시간-특이적 힘(time-specific power)을 행사한 어떤 것으로 생각하는 것에는 매우 이상한 측면이 있다" (Steward, 2020, p. 347). 여기서 라듐은 기껏해야 "무작위로 시간이 정해진 하나의 가능성이 실현되는데 있어 수동적으로 참여한 것"(Steward, 2020, p. 347)에 불과하며, 따라서 사건 **방사성 붕괴를 통해 방사선이 방출됨**이 발생할지 말지 여부가 "자유 및 도덕적 책임과 유관한 의미에서" 라듐 원자에게 달려 있었다든지 혹은 사건 **방사성 붕괴를 통해 방사선이 방출됨**이 발생할지 말지 여부에 대해 라듐 원자가 "자유 및 도덕적 책임과 유관한" 통제력을 가지고 있었다고 말하는 것에는, 스튜어드 또한 인정하듯이, "매우 이상한 측면이 있다" (Steward, 2020, p. 347).

스튜어드(Steward, 2020)에 의해 이러한 문제를 해결할 수 있을지도 모르는 한가지 실마리가 제시된 바 있다. 그녀는 비록 자신이 한때 라듐 원자의 즉자적 힘을 외재적 양방력의 일종으로 이해하고 라듐 원자 또한 외재적 양방력을 소유한다 생각했지만 이러한 생각이 실수였음을 명시적으로 밝힌다 (Steward, 2020, p. 346). 그녀는 이제 라듐 원자가 가지고 있는 즉자적 힘은 외재적 양방력이 아니라고 주장한다. 그렇다면 자연스럽게 다음과 같은 질문이 제기된다: 라듐 원자가 가지고 있는 즉자적 힘과 자유로운 행위자로서의 우리가 가지고 있는 외재적 양방력은 모두 발현의 충분조건으로서의 발현-조건을 결여하는 그러한 종류의 힘인데, 그렇다면 두 힘은 정확히 어떤 차이에 근거하여 구분되는 것인가?

스튜어드는 일반적 힘(general power)과 시간-특이적 힘(time-specific power) 사이의 구분에 호소하여 이러한 질문에 대해 답변한다. 어떤 라듐 원자 r이 어떤 시점 t에 방사선을 방출했다고 해보자. 스튜어드에 의하면 우리는 여기서 r이 t에 일반적 힘 P(방사선을

방출함)[69]을 가지고 있다고는 말할 수 있지만 시간-특이적 힘 P(t에 방사선을 방출함)은 가지고 있다고는 말할 수 없다. 왜 그런가? 스튜어드는 다음과 같이 주장한다. 우리는 왜 하필 시점 t에 r이 일반적 힘 P(방사선을 방출함)을 발현하게 되었는지를 묻는 질문에 대해 어떠한 답변도 제공할 수 없다. 더불어 이는 우리의 인식적인 한계 때문이 아니라 왜 하필 시점 t에 r이 일반적 힘 P(방사선을 방출함)을 발현하게 되었는지를 설명[70]해줄 수 있는 사실 자체가 없기 때문이다. r이 과연 일반적 힘 P(방사선을 방출함)을 (다른 시점이 아닌 바로 그) 시점 t에 발현할 것인지 여부는 "[r]이 소유한 힘의 행사를 촉발할 수 있는 환경과 [r] 사이의 특징적인 상호작용에 의존하지 않으며, 따라서 그 시점에 라듐에 의해 소유된 어떠한 속성에 대해서도 의존하지 않는다" (Steward, 2020, p. 347). 스튜어드는 이로부터 다음과 같은 귀결을 끌어낸다: "그러므로 우리에게는 [다른 시점이 아닌 하필 바로] 그 시점에 [문제의 방사선] 방출이 일어났다는 사실이 [r 그 자체] 혹은 [r]이 그 시점에 다른 어떤 것과 맺게 된 관계와 조금이라도 관련되어 있다고 믿을 어떠한 특정한 이유도 없다" (Steward, 2020, p. 347). r이 갖고 있던 일반적 힘 P(방사선을 방출함)의 발현은 사실상 "무작위(random)이며 설명 불가능(inexplicable)"(Steward, 2020, p. 347)하다. 그녀는 이러한 경우 우리가 r에 대해 그것이 시점 t에 일반적 힘 P(방사선을 방출함)을 가지고 있었다고 말할 수는 있을지는 몰라도 이에 대응하는 시간-특이적 힘 P(t에 방사선을 방출함)을 가지고 있었다고 말하기는 어렵다고 주장한다.

그렇다면 스튜어드는 여기서 어떤 힘이 외재적 양방력이기 위해

---

[69] 본 논문에서는 임의의 발현 유형 M, M*, …에 대해 M, M*, …를 발현-유형 정체성으로 갖는 힘을 표현 'P(M, M*, …)'을 사용하여 가리킬 것이다. 예를 들어 세균을 죽이는 힘은 표현 'P(세균을 죽임)'을 사용하여 가리킬 것이다.

[70] 여기서 제시된 조건이 발현의 충분조건으로서의 발현-조건을 결여하는 외재적 양방력이 만족하는데 무리가 없는 그러한 조건으로 제시되고 있다는 점을 고려할 때, 여기서의 '설명'은 피설명항의 발생을 필연화하는(necessitating) 인과적 조건을 명시하는 방식으로 이루어지는 설명이라기 보다는 피설명항이 발생할 확률에 영향을 미치는 인과적 조건을 명시하는 방식으로 이루어지는 설명으로 이해되어야 할 것이다.

만족해야 하는 조건으로 발현의 충분조건으로서의 발현-조건의 결여와 더불어 다음을 추가적으로 요구하고 있는 것이다: 해당 힘의 소유자는 이를 단지 일반적 힘으로서 가질 뿐 아니라 시간-특이적 힘으로서도 가져야 한다. 이러한 조건을 '시간-특이적 힘 조건'이라고 부르자. 스튜어드는 어떤 힘이 시간-특이적 힘 조건을 만족하기 위해서는 그것이 행사된 시점 t에 대해 우리가 왜 하필 그것이 t에 행사되었는지에 대해 문제의 힘의 소유자와 소유자가 처한 환경 사이의 특징적인 상호작용을 지적함으로써 설명할 수 있어야 한다고 주장한다. 따라서 그녀는 결국 즉자적 힘과의 구분 문제에 대해 다음과 같이 답변하고 있는 것이다: 즉자적 힘은 설사 그것이 시점 t에 행사되었다 하더라도 우리는 왜 하필 그것이 t에 행사되었는지에 대해 문제의 힘의 소유자와 소유자가 처한 환경 사이의 특징적인 상호작용을 지적함으로써 설명할 수 없지만 외재적 양방력에 대해서는 이러한 설명이 가능하다.

하지만 최근 로우, 마모도로 그리고 코리에 의해 이루어진 논의는 이러한 스튜어드의 답변이 실패함을 시사하는 것처럼 보인다. 먼저 로우(Lowe, 2011)는 라듐 원자가 가지고 있는 즉자적 힘, P(방사선을 방출함)과 관련하여 문제의 힘은 "[외부] 조건의 변화가 라듐 원자가 특정한 기간 동안 방사성 붕괴를 겪을 확률에 어떠한 차이도 가져다 주지 못한다는 의미에서 외부 조건에 대해 완전히 무감각"(Lowe, 2011, p. 22)하다고 주장한다. 만약 이러한 로우의 주장이 옳다면 라듐 원자가 가지고 있는 즉자적 힘에 대해 설사 그것이 시점 t에 행사되었다 하더라도 우리는 왜 하필 그것이 t에 행사되었는지에 대해 라듐 원자와 라듐 원자가 처한 환경 사이의 특징적인 상호작용을 지적함으로써 설명할 수 없다는 스튜어드의 주장은 옳을 것이다.

하지만 마모도로(Marmodoro, 2016, p. 205)는 이러한 로우의 주장을 정면으로 거부한다. 그녀는 라듐 원자의 즉자적 힘, P(방사선을 방출함)의 행사는 "진공 요동(vacuum fluctuation)"에 의해 촉발된다는 사실을 지적한다. 만약 이러한 마모도로의 지적이 옳다면, 라듐 원자가 가지고 있는 즉자적 힘, P(방사선을 방출함)이 "[외부] 조건의 변화가

라듐 원자가 특정한 기간 동안 방사성 붕괴를 겪을 확률에 어떠한 차이도 가져다 주지 못한다는 의미에서 외부 조건에 대해 완전히 무감각"(Lowe, 2011, p. 22)하다는 로우의 주장은 명백한 실수일 것이다. 더 나아가 라듐 원자가 가지고 있는 즉자적 힘에 대해 그것이 설사 시점 t에 행사되었다 하더라도 우리는 왜 하필 그것이 t에 행사되었는지에 대해 라듐 원자와 라듐 원자가 처한 환경 사이의 특징적인 상호작용을 지적함으로써 설명할 수 없다는 스튜어드의 주장 또한 별로 설득력이 없을 것이다.

코리는 로우와 마모도로 사이의 논쟁과 관련하여 다음과 같이 논평한 바 있다:

> 하지만 상황은 로우 혹은 마모도로가 제안하는 것보다 더 복잡하다. 첫째, 마모도로가 지적하는 연구(Yokoyama and Ujihara, 1995)는 원자핵의 방사성 붕괴라기 보다는 들뜬 원자에 의한 광자의 '즉자적' 방출에 대한 것이다. 둘째, 즉자적 방출에 대한 이러한 모형들은 진공 요동으로부터의 기여와 전자와 그것의 고유한 장 사이의 상호작용으로 인한 기여 둘 다를 포함한다(Dalibard et al., 1982). (Corry, 2019, p. 49)[71]

만약 이러한 코리의 논평이 옳다면 이는 다음을 시사한다. 최소한 들뜬 원자가 가지고 있는 힘, P(광자를 방출함)에 대해서는 그것이 시점 t에 행사되었을 경우 우리는 왜 하필 그것이 t에 행사되었는지에 대해 들뜬 원자와 들뜬 원자가 처한 환경 사이의 특징적인 상호작용, 즉

---

[71] 원문은 다음과 같다: "But the situation is more complex than Lowe or Marmodoro suggest. First, the research that Marmodoro points to (Yokoyama and Ujihara, 1995) is concerned with the 'spontaneous' emission of photons by excited atoms, rather than radioactive decay of nuclei. Second, these models of spontaneous emission involve both a contribution from vacuum fluctuations and a contribution due to an electron's interaction with its own field (Dalibard et al., 1982)." (Corry, 2019, p. 49) 본 인용문에서 언급되고 있는 요코야마 및 우지하라(Yokoyama and Ujihara, 1995)의 연구와 달리바드 등(Dalibard *et al*, 1982)의 연구에 대한 서지정보는 코리(Corry, 2019)를 참조하라.

“진공 요동”과 “전자와 그것의 고유한 장 사이의 상호작용”을 지적함으로써 설명할 수 있다. 따라서 들뜬 원자가 가지고 있는 힘, P(광자를 방출함)은 발현의 충분조건으로서의 발현-조건을 결여하는 힘이면서 동시에 문제의 힘의 소유자인 들뜬 원자가 그것을 일반적 힘으로도 가지고 있으면서 동시에 시간-특이적 힘으로도 가지고 있는 그러한 종류의 힘이다. 즉 들뜬 원자가 가지고 있는 힘, P(광자를 방출함)은 스튜어드(Steward, 2020)에 의해 제시된 어떤 힘이 외재적 양방력이기 위한 두가지 조건을 모두 만족한다. 이러한 사실은 즉자적 힘과 외재적 양방력을 구분하려는 스튜어드(Steward, 2020)의 시도가 실패함을 보여준다. 왜냐하면 스튜어드가 어떤 힘이 양방력이기 위해 만족해야 하는 조건으로 추가하고 있는 조건, 즉 그것의 소유자가 문제의 힘을 단지 일반적 힘으로 갖는 것을 넘어 시간-특이적 힘으로도 가져야 한다는 조건을 만족하는 비생물의 즉자적 힘이 존재하기 때문이다. 들뜬 원자가 갖는 광자를 방출하는 힘이 바로 그것이다.

그렇다면 결국 외재적 양방력 이론은 들뜬 원자가 갖는 광자를 방출하는 힘 또한 외재적 양방력이라는 귀결을 피하기 어려운 것처럼 보인다. 따라서 외재적 양방력 이론은 다음과 같은 견해에 개입하게 된다: 들뜬 원자가 시점 t에 광자를 방출했다고 가정했을 때, 들뜬 원자는 시점 t에 여전히 광자를 방출하지 않을 수 있었으며 들뜬 원자가 시점 t에 광자를 방출할지 말지 여부는 “자유 및 도덕적 책임과 유관한 의미에서” 들뜬 원자에게 달려 있었던 것이다. (혹은 들뜬 원자는 자신이 시점 t에 광자를 방출할지 말지 여부에 대해 “자유 및 도덕적 책임과 유관한 의미에서” 통제력을 가지고 있었다.) 물론 문제의 달려 있음-관계는 자유 및 도덕적 책임의 충분조건은 아닐 수 있을 것이다. 하지만 “자유 및 도덕적 책임과 유관한” 이러한 관계들이 들뜬 원자와 사건 **시점 t에 광자를 방출함** 사이에 성립한다는 생각은 여전히 그 자체로 문제적인 것처럼 보인다. 왜냐하면 이는 다음과 같은 함축을 갖기 때문이다: 우리가 인간 행위자로서 도덕적 책임 귀속을 동반하는 방식으로 어떤 행위를 자유롭게 수행할 때 문제의 행위가

수행될지 말지 여부는 기껏해야 시점 t에 광자를 방출할지 말지 여부가 들뜬 원자에게 달려 있었다고 말할 때와 정확히 동일한 의미에서 우리에게 달려 있었던 것에 불과하다; 혹은 자유롭게 수행된 행위에 대해 행위자가 가지고 있는 통제력은 기껏해야 들뜬 원자가 광자를 방출할지 말지 여부에 대해 가지고 있었던 통제력에 불과하다.

나는 이러한 귀결은 넓은 원천 양립불가론자들이 외재적 양방력을 통해서 확보할 수 있는 자유로운 행위자의 통제력의 정체를 보다 분명히 드러내준다고 생각한다. 외재적 양방력을 통해서 확보되는 자유로운 행위자의 통제력은 (그것을 여전히 통제력이라 부를 수 있다면) 기껏해야 들뜬 원자가 사건 **시점 t에 광자를 방출함**에 대해서 가지고 있었던 통제력에 불과하다. 물론 우리는 이러한 통제력을 사소하게 가지고 있을 것이다. 하지만 자유 및 도덕적 책임을 근거짓는 통제력은 분명 이러한 통제력보다는, 말하자면, 훨씬 더 탄탄한 통제력인 것처럼 보이며, 따라서 즉자적 힘과의 구분 문제는 외재적 양방력 이론이 자유 및 도덕적 책임을 근거짓는 통제력을 충분히 만족스럽게 포착하는데 실패했음을 보여주는 것처럼 보인다. 이러한 사실은 넓은 원천 양립불가론자로 하여금 스스로-결정하는 힘의 본성에 대한 이론으로 외재적 양방력 이론을 받아들일 수 없게 만드는 또 하나의 좋은 이유이다.

6장의 논의를 마무리하기 전에 여기서 제시된 비판에 대해 외재적 양방력 이론가의 입장에서 제시해볼 수 있는 한가지 답변을 고려해보자. 문제의 답변은 다음과 같은 주장에 기초한 답변이다: 자유로운 행위자는 임의의 행위 A에 대해 발현-유형 **A를 함**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력 P를 가질 때마다 항상 발현-유형 **P를 행사하지 않음**을 자신의 발현-유형 정체성으로 갖는 또 다른 외재적 양방력을 "함께" 소유한다. 이러한 주장은 일견 외재적 양방력 이론가로 하여금 즉자적 힘과의 구분 문제를 해결할 수 있는 유망한 이론적 자원을 제공해주는 것처럼 보인다. 왜냐하면 외재적 양방력 이론가들은 이제 즉자적 힘과의 구분 문제와 관련하여 다음과 같은 답변을 제시할 수 있는 것처럼 보이기 때문이다: 자유로운 행위자는

임의의 행위 A에 대해 발현-유형 **A를 함**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력을 가질 때마다 항상 이와 함께 문제의 힘을 행사하지 않을 수 있는 외재적 양방력을 별도로 갖는다; 반면 들뜬 원자는 오직 하나의 외재적 양방력 즉 발현-유형 **광자를 방출함**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력 만을 가질 뿐 문제의 힘을 행사하지 않을 수 있는 외재적 양방력은 결여한다; 자유 및 도덕적 책임과 유관한 통제력은 하나의 단일한 외재적 양방력 만으로는 확보될 수 없다; 이는 오직 문제의 외재적 양방력과 더불어 이를 행사하지 않을 수 있는 외재적 양방력을 함께 갖는 개체에 대해서만 귀속될 수 있는 그러한 종류의 통제력이다. 어떤 의미에서 이러한 방식의 답변은 즉자적 힘과의 구분 문제를 외재적 양방력과 즉자적 힘과의 구분 없이 새로운 힘 (문제의 외재적 양방력을 행사하지 않을 수 있는 외재적 양방력)을 추가적으로 도입함으로써 해결하려는 종류의 답변이라고 할 수 있다.[72]

하지만 내 생각에 자유 및 도덕적 책임의 근거를 이루는 통제력을 서로 다른 두개의 외재적 양방력을 통해 이해하려는 시도는 일단 그 자체로 지나치게 임시방편적인(ad hoc) 해결책일 뿐더러 한가지 심각한

---

[72] 스튜어드(Steward, 2013, p. 691)는 어떻게 하나의 발현-유형만을 소유하는 힘이 양방향성을 갖는 힘이라고 할 수 있는지를 묻는 질문에 대해 답변하는 맥락에서 이와 유사한 (하지만 정확히 일치하지는 않는) 생각을 잠시 내비친 적이 있다. 그녀는, 예를 들어, 노래를 부르거나 부르지 않을 수 있는 힘으로서의 외재적 양방력은 오직 발현-유형 **노래를 부름**에 해당하는 방식으로만 발현될 수 있는 그러한 종류의 힘이지만, 그럼에도 불구하고 오직 행위자가 문제의 외재적 양방력을 행사하지 않을 수 있는 "수동적(passive)" 힘을 동시에 별도로 가지고 있었던 경우에 한하여 문제의 노래를 부르거나 부르지 않을 수 있는 힘, 즉 외재적 양방력의 행사가 "능동적(active)"인 것일 수 있다고 주장한다. 하지만 내가 아는 한 스튜어드가 이러한 생각 혹은 이와 유사한 생각이 즉자적 힘과의 구분 문제에 대해 어떤 함축을 가질 수 있을 가능성에 대해 논의한 적은 아직 없다. 더불어 적어도 내가 아는 한 이러한 방식 혹은 이와 유사한 방식으로 외재적 양방력 이론가의 입장에서 즉자적 힘과의 구분 문제를 해결하려고 시도했던 철학자 또한 아직 없다. 따라서 이어질 논의는 기본적으로 외재적 양방력 이론가의 입장에서 제시해봄직한 가상의 답변에 대한 재반박이라고 할 수 있다.

문제를 야기할 우려가 있다.

알바레즈(Alvarez, 2013, p. 109)와 스튜어드(Steward, 2020, pp. 345-346)는 왜 우리가 외재적 양방력 이론을 받아들여야 하는지에 대해 다음과 같이 논변한 바 있다. 먼저 지은이가 노트북을 샀으며 이는 외재적 양방력의 사례라고 가정해보자. 그렇다면 이때 지은이는 노트북을 사지 않을 수도 있었을 것이다. 즉 노트북을 사거나 사지 않을 수 있는 힘으로서의 외재적 양방력을 비-행사할 수도 있었을 것이다. 이때 양방력 이론의 반대자들은 다음과 같이 질문할 수 있을 것처럼 보인다: 문제의 상황을 왜 하나의 양방력을 통해 이해해야 하는가? 우리는 문제의 상황을 서로 다른 두개의 단방력 P(노트북을 삼)과 P(노트북을 사지 않음)을 통해서도 충분히 만족스럽게 이해할 수 있다. 하지만 알바레즈와 스튜어드에 의하면 만약 우리가 문제의 상황을 서로 다른 두개의 단방력 P(노트북을 삼)과 P(노트북을 사지 않음)을 통해 이해하게 된다면 다음과 같은 문제가 발생한다. 일단 문제의 상황은 (사소하게) 노트북을 살 수 있었던 상황, 즉 P(노트북을 삼)의 발현-조건이 만족된 상황이다. 더불어 문제의 상황은 노트북을 사지 않을 수도 있었던 상황, 즉 P(노트북을 하지 않음)의 발현-조건이 만족된 상황이다. 하지만 가정 상 P(노트북을 삼)과 P(노트북을 사지 않음)은 둘다 단방력이며, 따라서 문제의 힘들은 모두 발현-조건이 만족될 경우 그것의 발현이 필연화되는 그러한 종류의 힘이다. 그렇다면 해당 상황에서 P(노트북을 삼)과 P(노트북을 사지 않음)의 발현은 모두 필연화된다. 즉 지은이는 해당 상황에서 노트북을 사면서 동시에 사지 않게 된다. 하지만 이는 불가능하다. 따라서 지은이가 노트북을 사거나 사지 않을 수 있었던 상황은 서로 다른 두개의 단방력을 통해 이해될 수 없으며, 이는 오직 발현의 충분조건으로서의 발현-조건을 결여하는 힘, 즉 외재적 양방력을 통해서 이해되어야만 한다.[73]

나는 외재적 양방력 이론가들이 자유 및 도덕적 책임의 근거를

[73] 이는 알바레즈(Alvarez, 2013)의 논변이 구성된 방식을 따른 것이다. 스튜어드(Steward, 2020)의 논변 또한 이와 유사한 문제점을 지적하고 있다.

제공해주는 통제력을 서로 다른 두개의 외재적 양방력을 통해 이해할 경우 알바레즈와 스튜어드가, 말하자면, 지은이가 노트북을 사거나 사지 않을 수 있는 상황을 서로 다른 두 단방력을 통해 이해하려는 입장에 대해 제기했던 문제와 정확히 동일하지는 않더라도 아주 유사한 어떤 문제에 부딪히게 된다고 생각한다. 이번에는 감기에 걸린 영희가 병원에 갈지 말지 여부에 대해 고민하고 있는 상황을 상상해보자. 문제의 상황에서 영희는 병원에 갈 수도 있었고 가지 않을 수도 있었다. 이때 만약 우리가 자유 및 도덕적 책임의 근거를 제공해주는 통제력을 서로 다른 두개의 외재적 양방력을 통해 이해한다면, 문제의 상황은 다시금 외재적 양방력 P(병원에 감)을 행사할 수도 있었고 외재적 양방력 P(외재적 양방력 P(병원에 감)을 행사하지 않음)을 행사할 수도 있었던 상황으로 이해된다. 이제 전자의 힘을 'X'라고 부르고 후자의 힘을 'Y'라고 부르도록 하자. 문제의 상황은 분명 영희가 병원에 갈 수도 있었고 병원에 가지 않을 수도 있었던 상황이다. 즉 X가 행사될 수도 있었고 Y가 행사될 수도 있었던 상황이다. 그런데 주어진 상황에서 X가 행사될 수 있었다는 사실은 X의 발현-조건이 만족되었음을 함축한다. 또한 Y가 행사될 수 있었다는 사실 역시 Y의 발현-조건이 만족되었음을 함축한다. 그렇다면 주어진 상황은 X의 발현-조건과 Y의 발현-조건이 모두 만족된 상황이다. 따라서 다음과 같은 질문이 제기된다: 왜 X와 Y는 동시에 발현될 수 없는가? 영희는 분명 병원에 갈 수도 있었고 가지 않을 수도 있었다. 하지만 병원에 가면서 동시에 가지 않는 것만큼은 불가능하다. 그리고 두 힘 X와 Y의 발현-조건이 모두 만족된 이상 왜 하필 두 힘이 동시에 발현되는 것만큼은 불가능해야 하는 것인지를 이해하기 어렵다.

이에 대해 외재적 양방력 이론가의 입장에서 다음과 같이 답변할 수 있을까? X와 Y의 발현-조건이 동시에 만족되었음에도 불구하고 두 힘이 동시에 발현되는 것이 불가능한 이유는 외재적 양방력 P(병원에 감)을 행사하면서 동시에 행사하지 않는 것이 모순이기 때문이다. 하지만 이러한 답변은 다음과 같은 이유로 불만족스럽다. 첫째, 이는 나의 질문에 대한 답변이 아니다. 물론 양방력 P(병원에 감)을

행사하면서 동시에 행사하지 않는 것은 분명 모순이다. 따라서 이러한 일은 일어날 수 없을 것이다. 하지만 나는 여기서 일반적으로 서로 다른 두 힘 P와 Q의 발현-조건이 동시에 만족된 경우 P와 Q가 동시에 발현되는 것이 최소한 가능해 보인다는 점을 고려할 때 왜 X와 Y의 경우에 한해서는 이러한 일이 일어날 수 없는 지를 묻고 있는 것이다. 이러한 질문은 X와 Y가 동시에 발현되는 것이 가능했다면 애초에 제기되지 않았을 것이다. 따라서 양방력 P(병원에 감)을 행사하면서 동시에 행사하지 않는 것은 모순이라는 사실은 문제를 발생시키는 사실이지 문제에 대한 답변이 아니다.

둘째, 이는 (설사 성공적이라 하더라도) 오직 행위자 앞에 놓인 두 선택지가 무엇인가를 하는 것과 그것을 하지 않는 것인 경우만을 설명해줄 수 있다. 하지만 일단 우리가 무엇인가를 하거나 하지 않을 수 있는 상황을 서로 다른 두개의 힘을 통해 이해한다면, 이제 병원에 갈 수도 있고 가지 않을 수도 있는 상황에서 병원에 가면서 동시에 가지 않는 것만큼은 불가능한 이유와, 말하자면, 삼성 갤럭시를 사기로 결정할 수도 있고 애플 아이폰을 사기로 결정할 수도 있는 상황에서 삼성 갤럭시를 사기로 결정하면서 동시에 애플 아이폰을 사기로 결정하는 것만큼은 불가능한 이유[74]는 다를 수 없는 것처럼 보인다. 하지만 삼성 갤럭시를 사기로 결정하면서 동시에 애플 아이폰을 사기로

[74] 물론 삼성 갤럭시를 사기로 결정하면서 동시에 애플 아이폰을 사기로 결정하는 것이 그 자체로 불가능한 것은 아니다. 여기서 불가능한 것은 바로 삼성 갤럭시를 사는 것과 애플 아이폰을 사는 것 중에서 어느 하나를 선택하기 위해 고민한 끝에 해당 고민의 결과로서 둘 다를 선택하게 되는 것이다. 문제의 상황은 무엇보다 행위자가 고려하고 있는 선택지의 목록이 이미 결정되어 있는 상황으로 이해되어야 한다. 따라서 여기서는 이미 결정되어 있는 선택지 목록에 삼성 갤럭시와 애플 아이폰을 모두 사는 선택지는 포함되어 있지 않은 것으로 가정된다. 그렇다면 여기서 삼성 갤럭시를 사기로 결정하면서 동시에 애플 아이폰을 사기로 결정하는 것은 행위자가 고려하고 있던 선택지에 해당하지 않는 일이 일어나는 것에 해당한다. 물론 앞서 언급한 바와 같이 이러한 일이 일어나는 것이 그 자체로 불가능한 것은 아닐 것이다. 하지만 행위자가 고려하고 있던 선택지에 해당하지 않는 일이 일어났을 때 과연 이를 우리가 여전히 “선택”이라고 부를 수 있는지 여부는 최소한 분명치 않은 것처럼 보인다.

결정하는 것은 모순이 아니다. 따라서 모순에 호소한 답변은 삼성 갤럭시를 사기로 결정할 수도 있고 애플 아이폰을 사기로 결정할 수도 있는 상황에서 삼성 갤럭시를 사기로 결정하면서 동시에 애플 아이폰을 사기로 결정하는 것만큼은 불가능한 이유를 설명해줄 수 없다. 더불어 만약 여기서 모순에 호소한 답변이 실패한다면 앞선 경우, 즉 병원에 갈 수도 있고 가지 않을 수도 있는 상황에서 병원에 가면서 동시에 가지 않는 것만큼은 불가능한 이유를 묻는 질문에 대해서도 또한 마찬가지로 모순에 호소한 답변은 사태를 제대로 설명해내지 못하고 있다고 말할 수 있을 것이다. 따라서 모순에 호소한 답변은 앞서 제시되었던 질문, 즉 왜 X와 Y는 동시에 발현될 수 없는 것인지를 묻는 질문에 대한 만족스러운 답변이 아니다.

그렇다면 혹시 다음과 같이 답변하는 것은 어떤가? X와 Y가 동시에 발현되는 것이 불가능한 이유는 다음과 같이 이해될 수 있다: X는 어떤 의미에서 오직 Y를 극복하는 방식으로만 발현되며 반대로 Y는 오직 X를 극복하는 방식으로만 발현되기 때문이다.[75] 만약 이러한 설명이 옳다면, X가 발현되었다는 사실은 곧 Y가 극복되었음을 함축하며 Y가 발현되었다는 사실은 곧 X가 극복되었음을 함축한다. 따라서 X와 Y는 동시에 발현될 수 없다. 이러한 답변은 또한 삼성 갤럭시를 사기로 결정하면서 동시에 애플 아이폰을 사기로 결정할 수 없는 이유 또한 만족스럽게 설명해주는 것처럼 보인다. 삼성 갤럭시를 사기로 결정하는 힘 (아마도 이는 외재적 양방력의 일종일 것이다)은 항상 애플 아이폰을 사기로 결정하는 힘 (아마도 이 또한 외재적 양방력의 일종일 것이다)을 극복하는 방식으로만 발현되며 후자의 힘 또한 항상 전자의 힘을 극복하는 방식으로만 발현된다. 바로 이것이 주어진 상황에서 두 힘이 동시에 발현되는 것이 불가능한 이유이다.

하지만 내 생각에 이러한 답변 또한 결국 실패한다. 왜 그런가? 그 이유는 다음과 같다. 만약 X는 오직 Y를 극복함을 통해서만 발현되며 Y는 오직 X를 극복함을 통해서만 발현된다면 이로부터 다음과 같은

[75] 이러한 답변 가능성은 한성일 교수님께서 지적해주신 것이다.

함축이 따라 나오게 된다: Y를 극복함은 X의 발현-조건이며 X를 극복함은 Y의 발현-조건이다. 그렇다면 X가 행사될 수도 있고 Y가 행사될 수도 있는 상황은 있을 수 없다. 왜 그런가? 먼저 X가 발현될 수 있으려면 X의 발현-조건이 만족되어야 한다. 따라서 X가 발현되기 위해서는 Y가 극복되어야 한다. 또한 Y가 발현될 수 있으려면 Y의 발현-조건이 만족되어야 한다. 따라서 Y가 발현되기 위해서는, 이번에는 X가 극복되어야 한다. 따라서 X가 행사될 수도 있고 Y가 행사될 수도 있는 상황은 Y도 극복되고 X도 극복된 상황이며 이러한 상황에서는 X와 Y 중 어떠한 힘의 행사도 불가능하다. 왜냐하면 문제의 힘들은 모두 이미 극복되었기 때문이다. 따라서 결국 X가 행사될 수도 있고 Y가 행사될 수도 있는 상황은 존재할 수 없다. 하지만 최소한 양립불가론을 받아들이는 이상 X가 행사될 수도 있고 Y가 행사될 수도 있는 상황이 존재할 가능성은 포기할 수 없으며, 따라서 넓은 원천 양립불가론자는 이러한 답변을 채택할 수 없다.

물론 힘의 형이상학과 관련하여 여러 쟁점들에 대해 보다 구체적인 개입들을 가정할 경우 외재적 양방력 이론가의 입장에서 여기서 제기된 문제를 해결할 수 있는 방법이 아주 없는 것은 아닐지도 모르겠다. 예를 들어 힘의 발현-유형들은 정확히 어떤 방식으로 개별화 될 수 있으며 힘의 발현 혹은 행사가 정확히 어떤 범주의 존재자에 해당하는 것인지에 대한 보다 구체적인 가정들에 개입하거나 혹은 X와 Y 뿐 아니라 또 다른 힘을 다시 한 번 추가적으로 상정하는 방식으로 X와 Y가 왜 동시에 행사될 수 없는 것인지를 묻는 질문에 대해 외재적 양방력 이론가가 어떤 답변을 제시할 수 있는 가능성 일반을 배제하는 것은 몹시 어려운 일이다.

하지만 최소한 나는 앞서 제시한 논변을 통해 다음을 설득력 있게 보이는데 있어서는 성공했다고 생각한다: 외재적 양방력 이론가의 입장에서 왜 X와 Y는 동시에 행사될 수 없는 것인지를 묻는 질문에 대해 손쉽게 제시할 수 있는 답변은 존재하지 않는다. 내가 아는 한 외재적 양방력 이론가의 입장에서 임의의 행위 A에 대해 발현-유형 **A를 함**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력 P에 더해

발현-유형 **P를 행사하지 않음**을 자신의 발현-유형 정체성으로 갖는 외재적 양방력을 추가적으로 상정함으로써 즉자적 힘과의 구분 문제를 해결하려는 어떠한 시도도 아직까지 이루어진 바가 없다. 따라서 나는 외재적 양방력 이론가의 입장에서 왜 X와 Y는 동시에 행사될 수 없는 것인지를 묻는 질문에 대해 손쉽게 제시할 수 있는 답변이 존재하지 않음을 보이는 것만으로도 현재 상황에서는 넓은 원천 양립불가론자들로 하여금 스스로-결정하는 힘의 본성과 관련하여 외재적 양방력 이론을 거부할 한 가지 좋은 이유를 성공적으로 제공할 수 있다고 믿는다.

제 7 장 결론: 내재적 양방력 이론을 향하여

이상의 논의를 정리해보자. 먼저 넓은 원천 양립불가론자들은 원천-관계가 성립하기 위해 모종의 대안가능성이 존재해야 하며 바로 이러한 사실 때문에 인과적 결정론이 참인 세계에서 원천-관계는 성립할 수 없을 것이라고 생각한다. 더불어 원천-관계가 대안가능성의 존재를 요구한다는 그녀들의 생각은 다시금 원천-관계의 근거를 제공하는 자유로운 행위자의 스스로-결정하는 힘이 다음과 같은 특징을 갖는 힘이라는 생각에 근거하고 있다: 스스로-결정하는 힘의 모든 사례들은 그것의 본성 상 항상 어떤 대안가능성을 동반한다. 하지만 이러한 입장은 자연스럽게 다음과 같은 질문을 불러일으킨다: 스스로-결정하는 힘은 도대체 어떤 독특한 종류의 힘이길래 그것의 사례들이 항상 어떤 대안가능성을 동반하게 되는가? 더불어 양방력 이론은 이러한 질문에 대한 하나의 유력한 답변을 제공하는 것처럼 보였다. 왜냐하면 만약 스스로-결정하는 힘이 양방력의 일종이라면 그것의 사례들이 항상 어떤 대안가능성을 동반한다는 사실은 자연스럽게 설명되는 것처럼 보였기 때문이다.

하지만 양방력의 본성을 이해하는 한가지 유력한 방식에 해당하는 외재적 양방력 이론은 앞서 살펴본 바와 같이 여러모로 넓은 원천 양립불가론자들이 스스로-결정하는 힘의 본성과 관련하여 채택하기에는 문제가 많은 이론인 것처럼 보였다. 외재적 양방력 이론은 먼저 (i) 프랑크푸르트-유형 사례로부터 제기되는 문제로 인해 최소한 몇몇 넓은 원천 양립불가론자들로서는 도저히 받아들이기 어려운 타협을 받아들이게끔 강제하며, (ii) 넓은 원천 양립불가론적 방식으로 이해된 스스로-결정하는 힘이 그것의 본성 상 외재적 양방력과 동일시되기 어렵다는 문제를 겪었으며, 마지막으로 (iii) 자유로운 행위자가 자신의 행위에 대해 가지고 있는 통제력이 기껏해야 들뜬 원자가 광자를 방출할지 말지 여부에 대해 가지고 있는 통제력과 다를 바 없는 것이라는 귀결을 가졌다. 이러한 사실들은 분명 넓은 원천 양립불가론자들로 하여금 스스로-결정하는 힘의 본성과 관련하여

외재적 양방력 이론을 채택하기 어렵게 만든다.

하지만 물론 넓은 원천 양립불가론자가 스스로-결정하는 힘을 "외재적 양방력"의 일종으로 이해할 수 없다는 사실로부터 넓은 원천 양립불가론자가 스스로-결정하는 힘을 "양방력"의 일종으로 이해할 수 없다는 사실이 곧바로 따라 나오는 것은 아니다. 즉 넓은 원천 양립불가론자는 여전히 양방력 이론을 외재적 양방력 이론과는 다른 방식으로 받아들일 수 있을지도 모른다. 왜냐하면 외재적 양방력 이론 외에도 양방력의 본성을 이해하는 또 다른 방식이 존재하기 때문이다.

예를 들어, 프로스트(Frost, 2013, 2020)에 의하면 양방력은 서로 다른 두가지 방식으로 발현될 수 있는 그러한 종류의 힘이다. 즉 양방력은 서로 다른 두 발현-유형과 모두 발현-유형 소유-관계를 맺고 있는 그러한 종류의 힘이다. 예를 들어 방금 전 커피를 마실 때 내가 행사한 양방력, 즉 커피를 마시거나 마시지 않을 수 있는 힘은 발현-유형 **커피를 마심**에 해당하는 방식으로도 발현될 수 있고 발현-유형 **커피를 마시지 않음**에 해당하는 방식으로도 발현될 수 있는 그러한 종류의 힘이다. 따라서 양방력과 결부되어 있는 대안가능성은 문제의 양방력이 행사될 수 있는 또 다른 방식으로 이해된다. 나는 커피를 마실 때 또한 커피를 마시지 않을 수도 있었으며 이러한 대안가능성은 커피를 마시거나 마시지 않을 수 있는 양방력이 행사되지 않을 수 있었던 가능성이라기 보다는 또 다른 방식으로 행사될 수도 있었던 가능성으로 이해된다. 우리는 양방력의 본성에 대한 이러한 프로스트의 견해를 외재적 양방력 이론에 대비되는 것으로서 '내재적 양방력 이론'이라 부를 수 있을 것이다. (또한 우리는 그를 '내재적 양방력 이론가'라고 부를 수 있을 것이다.)

하지만 나는 스스로-결정하는 힘의 본성과 관련하여 양방력 이론을 받아들이는 넓은 원천 양립불가론자들의 입장에서 양방력의 본성과 관련하여 내재적 양방력 이론을 받아들이는 것이 과연 만족스러운 선택지일 수 있을 것인지 여부를 검토하는 것은 후속 과제로 남겨두려 한다. 본 논문의 목표는 어디까지나 양방력 이론을 받아들이는 넓은 원천 양립불가론자들이 양방력의 본성과 관련하여

적어도 외재적 양방력 이론을 채택하기는 어려움을 주장하고 논변하는 것에 있다.

더불어 본 논문에서 이루어진 작업은 우리에게 다음과 같은 교훈을 일깨워주는 것처럼 보인다: 철학자들은 자유 및 도덕적 책임과 같이 아직까지 다소 신비스러운 채로 남아있는 현상들을 설명하기 위해 때때로 어떤 특별한 종류의 힘에 호소하곤 한다; 하지만 이러한 방식으로 설명을 시도할 때 문제의 힘의 본성을 개념화하는 방식은 결코 사소한 문제일 수 없다; 우리가 주어진 현상을 설명함에 있어서 가지고 있는 이론적 동기들은 어쩌면 때로 우리가 설명항으로 사용하는 특별한 종류의 힘의 본성을 이해하는 방식과 충돌하게 될지도 모르며, 더 나아가 결국 우리가 문제의 힘의 본성과 관련하여 어떤 방식의 이해를 채택하는지에 따라서 제시될 설명이 궁극적으로 실패하게 될지도 모른다. 나는 자유 및 도덕적 책임의 본성을 설명하려는 넓은 원천 양립불가론의 기획에서 스스로-결정하는 힘의 본성을 외재적 양방력의 일종으로 이해할 경우 발생하게 되는 일련의 문제들이 바로 이러한 교훈을 보여주는 하나의 실례가 될 수 있다고 믿는다.

참고 문헌

김율. (2008). 둔스 스코투스의 의지이론 연구 (I)-'이성적 능력(potential rationalis)'으로서의 의지. *철학사상*, *28*, 213-241.

한성일. (2015). 본질과 정체성. *철학*, *124*, 73-96.

Alvarez, M. (2009). Actions, thought-experiments and the 'principle of alternate possibilities'. *Australasian Journal of Philosophy*, *87*(1), 61-81.

Alvarez, M. (2013, April). Agency and Two-Way Powers. *Proceedings of the Aristotelian Society, 113*(1), 101-121.

Alvarez, M., & Hyman, J. (1998). Agents and their Actions. *Philosophy*, *73*(2), 219-245.

Cartwright, N. (2012). Do the laws of physics state the facts?. In M. Curd, J. A. Cover & C. Pincock (Eds.), *The Philosophy of Science: The Central Issues* (2nd ed., pp. 871-882). W. W. Norton. (Reprinted from *Pacific Philosophical Quarterly*, *61*(1-2), 1980, pp. 75-84)

Corry, R. (2019). *Power and Influence: The Metaphysics of Reductive Explanation*. Oxford University Press.

Dennett, D. (1984). *Elbow Room*. MIT Press.

Fischer, J. M. (2006). *My Way: Essays on Moral Responsibility*. Oxford University Press.

Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *The journal of philosophy*, *66*(23), 829-839.

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The journal of philosophy, 68*(1), 5-20.

Frost, K. (2013). Action as the exercise of a two-way power. *Inquiry*, *56*(6), 611-624.

Frost, K. (2020). What could a two-way power be?. *Topoi*, *39*(5),

1141-1153.

Goetz, S. (2002). Alternative Frankfurt-Style Counterexamples to the Principle of Alternative Possibilities. *Pacific Philosophical Quarterly*, *83*(2), 131-147.

Jacobs, J. D. (2017). Introduction. In J. D. Jacobs (Ed.), *Causal powers* (pp. 1-6). Oxford University Press.

Kane, R. (1998). *The significance of free will*. Oxford University Press.

Lewis, D. (1983). New work for a theory of universals. *Australasian journal of philosophy*, *61*(4), 343-377.

Lowe, E. J. (2003). Individuation. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford handbook of metaphysics* (pp. 75-95). Oxford University Press.

Lowe, E. J. (2010). On the individuation of powers. In A. Marmodoro (Ed.), *The metaphysics of powers: Their grounding and their manifestations* (pp. 8-26). Routledge.

Lowe, E. J. (2011). How not to think of powers: a deconstruction of the 'Dispositions and Conditionals' Debate. *The Monist*, *94*(1), 19-33.

Lowe, E. J. (2012). Asymmetrical dependence in individuation. In F. Correia & B. Schneider (Eds.), *Metaphysical Grounding: Understanding the Structure of Reality* (pp. 214-233). Cambridge University Press.

Lowe, E. J. (2013). The will as a rational free power. In R. Groff & J. Greco (Eds.), *Powers and capacities in philosophy* (pp. 172-185). Routledge.

Marmodoro, A. (2016). Dispositional modality vis-à-vis conditional necessity. *Philosophical Investigations*, *39*(3), 205-214.

McKenna, M. (2001). Source incompatibilism, ultimacy, and the transfer of non-responsibility. *American Philosophical Quarterly*, *38*(1), 37-51.

McKenna, M. (2013). Reasons-responsiveness, agents, and mechanisms. In D. Shoemaker (Ed.), *Oxford studies in agency and responsibility* (Vol. 1, pp. 151-83). Oxford University Press.

Molnar, G. (2003). *Powers: A study in metaphysics*. Oxford University Press.

O'Connor, T. (2002). *Persons and causes: The metaphysics of free will*. Oxford University Press.

O'Connor, T. (2011). Agent-causal theories of freedom. In R. Kane (Ed.), *The Oxford handbook of free will* (2nd ed., pp. 309-328). Oxford University Press.

O'Connor, T. (2013). Agent-causal power. In P. Russell & O. Deery (Eds.), *The Philosophy of Free Will: Essential Readings from the Contemporary Debates* (pp. 229-248). Oxford University Press.

O'Connor, T. & Franklin, C. (2020). Free Will. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition). Retrieved from https://plato.stanford.edu/archives/spr2020/entries/freewill/

Pereboom, D. (2001). *Living without free will*. Cambridge University Press.

Pereboom, D. (2003). Source Incompatibilism and Alternative Possibilities. In D. Widerker & M. McKenna (Eds.), *Moral responsibility and Alternative Possibilities* (pp. 185-199). Ashgate.

Pink, T. (2004). *Free will: A very short introduction*. Oxford University Press.

Pink, T. (2011). Freedom and action without causation: Noncausal theories of freedom and purposive agency. In R. Kane (Ed.), *The Oxford Handbook of Free Will* (2nd ed., pp. 349-365). Oxford University Press.

Pink, T. (2016). *Self-determination: The ethics of action* (Vol. 1).

Oxford University Press.

Robb, D. (2020). Moral responsibility and the Principle of Alternative possibilities. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition). Retrieved from https://plato.stanford.edu/archives/fall2020/entries/alternative-possibilities/

Steward, H. (2009). Fairness, agency and the flicker of freedom. *Noûs*, *43*(1), 64-93.

Steward, H. (2012a). *A metaphysics for freedom*. Oxford University Press.

Steward, H. (2012b). Actions as processes. *Philosophical Perspectives*, *26*, 373-388.

Steward, H. (2013). Responses. *Inquiry*, *56*(6), 681-706.

Steward, H. (2020). Agency as a two-way power: A defence. *The Monist*, *103*(3), 342-355.

Steward, H. (2021). Frankfurt cases, alternative possibilities and agency as a two-way power. *Inquiry*. Advance online publication. https://doi.org/10.1080/0020174X.2021.1904639

Stump, E. (2003). Moral responsibility without alternative possibilities. In D. Widerker & M. McKenna (Eds.), *Moral responsibility and alternative possibilities* (pp. 139-158). Ashgate.

Thompson, M. (2008). *Life and action*. Harvard University Press.

Timpe, K. (2007). Source incompatibilism and its alternatives. *American Philosophical Quarterly*, *44*(2), 143-155.

Tugby, M. (2017). Power worlds and the problem of individuation. *American Philosophical Quarterly*, 54(3), 269-281.

van Miltenburg, N. & Ometto, D. (2020). Free will and mental powers. *Topoi*, *39*, 1155-1165.

Widerker, D. (1995a). Libertarian freedom and the Avoidability of decisions. *Faith and Philosophy*, 12(1), 113-118.

Widerker, D. (1995b). Libertarianism and Frankfurt's attack on the principle of alternative possibilities. *The Philosophical Review*, *104*(2), 247-261.

Abstract

Wide Source Incompatibilism on the Nature of Freedom and External Two-Way Power Theory

Cho, Minsoo
Department of Philosophy
The Graduate School
Seoul National University

According to source incompatibilism, agents are morally responsible for the actions they have performed only if they are the very source of them, but the sourcehood-relationship cannot obtain in a world where causal determinism is true. Particularly, wide source incompatibilists assert that the reason why the sourcehood-relationship at issue cannot obtain in a world where causal determinism is true is the following: agents can be the source of the actions they have performed only if they could have done otherwise, but, in a world where causal determinism is true, they couldn't.

Some wide source incompatibilists appeal to the following reason to explain why agents can be the source of the actions they have performed only if they could have done otherwise: this is because, only when the actions performed by agents are instances of their self-determining powers, they can be the source of the actions they have

performed, and it is in the very nature of self-determining powers that all instances of them always involve an alternative possibility. Especially, two-way power theorists understand self-determining powers, which possess the unique characteristic, as a kind of two-way powers, that is, powers to do something or not. According to two-way power theorists, every instance of self-determining powers always involves an alternative possibility in virtue of their being powers to do something or not, that is, two-way powers. If so, what does it mean exactly that a power is a power to do something or not? The answer given by external two-way power theorists is the following: it means that they can still be not manifested even if all the manifestation-conditions are satisfied. In other words, external two-way powers are powers which lacks any manifestation-conditions which are sufficient for their manifestations.

However, for wide source incompatibilists, it is difficult to regard self-determining powers, which were understood as a sort of two-way powers, as external two-way powers again. First, if wide source incompatibilists understand the nature of self-determining powers, which were identified with two-way powers, as a sort of external two-way powers again, they cannot exploit a specific kind of answer to the robustness objection which Fischer raised based on a Frankfurt-style case. Of course, some other answers to his objection still remain available for external two-way power theorists, but these answers clearly have limitations. Therefore, at least for some wide source incompatibilists who also embrace a two-way power theory, this fact

amounts to a good reason for denying external two-way power theory. Second, it seems that self-determining powers, which were understood in a wide source incompatibilist way, cannot be identified with external two-way powers because of the very nature of them. For the claim that self-determining powers are a sort of external two-way powers has the following problematic consequence: it is possible that the case where agents self-determined whether they perform an action is the case where agents didn't exercise their power to self-determine whether they perform the very action. Third, the spontaneous power of excited atom to emit a photon also lacks any manifestation-conditions which are sufficient for its manifestation. Thus, if external two-way power theory is correct, this spontaneous power cannot be appropriately distinguished from the self-determining power which provides a ground of freedom and moral responsibility. It seems to have a problematic consequence that the control of free agents over their actions is, at best, just like the control of excited atoms over whether they emit a photon or not. In addition, this problem cannot be solved by mere addition of any necessary conditions for a power to be an external two-way power. Given these considerations, external two-way power theory is not satisfactory as a theory concerning the nature of self-determining powers when they are understood in wide source incompatibilist way.

Keywords : freedom, moral responsibility, self-determining power, wide source incompatibilism, two-way power theory, external two-

way power theory

Student Number : 2019-26553