

조작논증과 어니, 다이애나, 번개*

김 세 화

【국문요약】 “자유의지에 대한 조작논증과 근원-양립가능주의의 대응”에서 김성수 교수는 근원-자유와 결정론이 양립가능하지 않다는 근원-양립불가능주의에서 제시한 논증인 조작논증에 대해 다룬다. 김성수는 조작논증에 대한 양립가능주의의 두 종류의 대응 중 강경한 대응보다는 온건한 대응이 더 설득력이 있다고 주장하는데, 특히 그는 번개 사례를 통해 강경한 대응을 펼치는 사토리오에 집중하여 반론을 펼침으로써 강경한 대응이 설득력이 부족하다고 주장한다. 본 논문에서 필자는 사토리오에 대한 김성수의 반론이 성공적이지 않다고 논증한다. 이를 위해서 필자는 사토리오의 책임회석 효과에 대한 설명에 문제가 있다는 김성수의 반론에 대해서 이러한 문제점을 피할 수 있는 원격 집행부대 예를 새롭게 제시한다. 또한 필자는 번개 사례에서 보이는 직관이 조작논증의 전제를 지지하는 직관만큼 강해보이지 않는다는 김성수의 반론에 대해 사토리오의 논증이 성공하기 위해서 번개 사례에서 보이는 직관이 다이애나 사례에서 보이는 직관만큼 강할 필요는 없다고 논증한다.

【주요어】 조작논증, 자유행위, 책임, 강경한 대응, 사토리오, 다이애나 사례, 번개 사례, 김성수

투고일: 2019. 1. 28 심사 및 수정 완료일: 2019. 2. 24 게재확정일: 2019. 2. 24

* 이 논문의 심사를 맡아 유익한 논평을 해주신 익명의 심사위원들께 감사드린다.

1

“자유의지에 대한 조작논증과 근원-양립가능주의의 대응”¹⁾에서 김성수 교수(이하 ‘김성수’)는 조작논증(the manipulation argument)에 대해 다룬다. 조작논증은 근원-자유²⁾(이하 ‘자유’)와 결정론이 양립 가능하지 않다는 근원-양립불가능주의(source incompatibilism, 이하 ‘양립불가능주의’)에서 제시한 논증이다. 조작논증에 대한 양립가능주의의 대응은 강경한 대응과 온건한 대응으로 나뉜다. 김성수는 각각의 대응을 살펴본 후 강경한 대응보다는 온건한 대응이 더 설득력이 있다고 주장한다. 강경한 대응이 설득력이 부족하다는 것을 보이는데 있어 김성수는 특히 번개 사례를 통해 조작논증에 반론을 가하는 사토리오에 집중을 하여 이에 대해 반론을 펼친다. 본 논문에서 필자는 사토리오에 대한 김성수의 반론이 성공적이지 않다는 점을 밝히고자 한다.

2

조작논증은 다음과 같다.

조작논증:

(1) 일정한 방식으로 조작된 행위는 도덕적 책임을 물을 수 있는

1) 김성수 (2018).

2) 전통적인 의미에서의 자유는 달리 할 수 있음으로서의 자유인 것에 반해 최근 자유의지 논의에서는 근원-자유를 중심으로 논의가 진행되는 경우가 많다. 근원-자유란 “합리적 이유에 일정한 방식으로 반응하거나 행위의 원인인 일차적 욕구가 행위자의 이차적 욕구와 잘 맞는 것”이다. 김성수 (2018), p. 374. 이하 본문에서 ‘자유’는 ‘근원-자유’를, ‘양립가능주의’는 ‘근원-양립주의’로, ‘양립불가능주의’는 ‘근원-양립불가능주의’를 의미한다.

자유행위가 아니다.

- (2) 일정한 방식으로 조작된 행위와 결정론적으로 발생한 일상적 행위 사이에 차이가 없다.
- (3) 따라서, 결정론적으로 발생한 일상적 행위 역시 자유행위가 아니다.³⁾

조작논증의 전제 (1)을 뒷받침하는 것으로서 김성수가 고려하는 사례는 밀(Mele)의 수정란 예(‘zygote’ example)⁴⁾의 한 버전인 다음과 같은 사례이다.

다이애나 사례:

메리는 임신을 위해 시험관 시술을 하려한다. 다이애나 여신이 개입하여 시험관에서 형성되고 있는 수정란에 유전자 변형이 생기게 하는데, 다이애나는 그 결과로 선천적인 유전적 성향을 가진 아이 어머니(Ermie)가 잉태될 것을 안다. 다이애나는, 부분적으로 이 성향으로 인해 삼십년 후 어머니가 유산을 물려받기 위해 삼촌을 살해할 것 역시 안다. 세계의 상태와 결정론적 법칙을 완전히 알고 자신이 개입한 이후 세계가 정확히 어떤 식으로 전개될지 예측할 수 있는 다이애나는 어머니가 삼십년 후 자기 삼촌을 살해하길 원하기 때문에 그에 맞는 유전자 변형이 생기게 한다. 그럼에도 삼십년 후 자기 삼촌을 살해할 때 어머니는 (어떤 강제나 강요 없이, 이유에 반응하여, 자신의 행동을 이끄는 욕구를 자신의 것으로 여기고, 관련 메커니즘에 대한 자신의 책임을 받아들이는 등) 자유에 대한 표준적인 양립가능주의의 모든 조건을 만족한다.⁵⁾

3) 김성수 (2018), p. 375.

4) Mele (2006), p. 188.

5) Sartorio (2016), p. 159. 김성수 (2018), pp. 377-378에서 재인용.

이 사례에서 어머니의 삼촌 살해 행위는 일정한 방식으로 조작된 행위로서, “양립가능주의자들이 제시하는 근원-자유의 조건을 만족”⁶⁾ 함에도 불구하고 “직관적으로 자유행위가 아니”⁷⁾다. 이와 같은 사례를 통해 양립불가능주의는 일정한 방식으로 조작된 행위는 자유행위가 아니며, 따라서 이와 근본적인 차이가 없는, 결정론적으로 발생한 일상적 행위 역시 자유행위가 아니라고 주장한다.

이 다이애나 사례에 적용하여 조작논증을 다시 정식화하면 다음과 같이 된다.

다이애나 사례를 통해 본 조작논증:

- (1)* 어머니의 행위는 도덕적 책임을 물을 수 있는 자유행위가 아니다.
- (2)* 어머니의 행위와 결정론적으로 발생한 일상적 행위 사이에 차이가 없다.
- (3) 따라서, 결정론적으로 발생한 일상적 행위 역시 자유행위가 아니다.

3

조작논증에 대한 양립가능주의자의 대응은 크게 두 가지로 나뉜다. 먼저 강경한 대응은 조작논증의 전제 (2)는 받아들이지만 전제 (1)과 (1)*을 거부함으로써 조작논증을 반박한다. 김성수에 의하면 강경한 대응은 “결정론적으로 발생한 일상적 행위가 도덕적 책임을 물을 수 있는 자유행위라는 전제로부터 출발한다.”⁸⁾ 이와 같은 전제에서 출발하여 조작논증의 전제 (2)를 받아들인 후 이로부터 다이애나 사례에서 어머니의 행위 역시 자유행위라고 주장한다는 것이다. 김성수는 강경한 대응 방식을 취하는 양립가능주의자의 논증을

⁶⁾ 김성수 (2018), p. 377.

⁷⁾ Ibid.

⁸⁾ Ibid., p. 375.

보다 구체적으로 다음과 같이 제시한다.

강경한 대응 논증:

- (4) (양립가능주의자들이 내놓은 자유의 조건을 만족하는) 결정론적으로 발생한 일상적 행위는 도덕적 책임을 물을 수 있는 자유행위이다.
- (5) 결정론적으로 발생한 일상적 행위와 어머니의 행위 사이에 차이가 없다.
- (6) 따라서, 어머니의 행위 역시 도덕적 책임을 물을 수 있는 자유행위이다.⁹⁾

이에 반해 온건한 대응은 전제 (1)과 (1)*은 받아들이지만 전제 (2)를 거부함으로써 조작논증을 반박한다.¹⁰⁾ 온건한 대응은 일정한 방식으로 조작된 어머니의 행위와 결정론적으로 발생한 일상적 행위 사이에는 중대한 차이가 있다는 주장에서 출발한다. 즉, 조작논증의 전제 (1)과 이를 다이애나 사례에 적용한 전제 (1)*이 참이라고 하더라도 “자유와 도덕적 책임의 측면에서 볼 때 결정론적으로 발생한 일상적 행위와 일정한 방식으로 조작된 행위 사이에 차이가 있”¹¹⁾으므로, 어머니의 행위가 자유행위가 아니라는 사실이 결정론적으로 발생한 일상적 행위 역시 자유행위가 아니라는 것을 보이지는 않는다고 주장한다.

본 논문에서 필자가 목표로 하는 바는 사토리오가 강경한 대응을 위해 제시하는 논증에 대한 김성수의 반론이 성공적이지 않다는 것이다. 따라서 앞으로 전개되는 논의에서는 강경한 대응에만 집중하여 진행하기로 하겠다.

⁹⁾ Ibid., p. 379.

¹⁰⁾ Ibid., p. 375.

¹¹⁾ Ibid., p. 376.

4

다음을 다시 보자.

다이애나 사례를 통해 본 조작논증:

- (1)* 어니의 행위는 도덕적 책임을 물을 수 있는 자유행위가 아니다.
- (2)* 어니의 행위와 결정론적으로 발생한 일상적 행위 사이에 차이가 없다.
- (3) 따라서, 결정론적으로 발생한 일상적 행위 역시 자유행위가 아니다.

다이애나 사례를 통해 본 강경한 대응 논증:

- (4) (양립가능주의자들이 내놓은 자유의 조건을 만족하는) 결정론적으로 발생한 일상적 행위는 도덕적 책임을 물을 수 있는 자유행위이다.
- (5) 결정론적으로 발생한 일상적 행위와 어니의 행위 사이에 차이가 없다.
- (6) 따라서, 어니의 행위 역시 도덕적 책임을 물을 수 있는 자유행위이다.¹²⁾

위의 두 논증을 보면 전제 (2)*와 (5)는 동일한 명제이며, 전제 (1)*은 강경한 대응 논증의 결론에 대한 부정이고, 전제 (4)는 조작 논증의 결론 (3)에 대한 부정이라는 것을 알 수 있다. 사토리오에 따라서 이러한 강경한 대응 논증이 취하는 전략을 ‘후건 부정 전략 (*modus-tollens strategy*)’라고 부른다.¹³⁾ 즉, (1)*이 참이기 때문에 (3)이 참이라는 조작논증에 대하여, 강경한 대응은 (3)이 거짓이기 때문에 (1)*도 거짓이라고 논증한다.

¹²⁾ Ibid., p. 379.

¹³⁾ Sartorio (2016), pp. 163-164.

여기에서 김성수는 페리붐(Pereboom)의 주장을 받아들여 전제 (4)를 지지하는 직관이 전제 (1)*을 지지하는 직관만큼 강하지 않다고 주장한다.¹⁴⁾ 사토리오 역시 (1)*이 직관적으로는 참으로 보인다는 점을 인정한다.¹⁵⁾ 사토리오는 후건 부정(강경한 대응)보다 전건 긍정(양립불가능주의)을 선호할 좋은 이유가 없을 경우에만 후건 부정 전략이 통할 수 있다고 말한다.¹⁶⁾ 즉 만약 양립가능주의자의 직관보다는 양립불가능주의자의 직관이 보다 신뢰할만하다고 생각될 수 있는 이유가 있다면 이는 후건 부정보다는 전건 긍정을 지지하는 근거가 될 것이라고 한다.

사토리오는 그렇지만 양립불가능주의자의 직관, 즉 다이애나 사례에서 어니의 행위가 자유행위가 아니라는 직관의 신뢰성에 대해 의심할 수 있게 만드는 근거가 있다고 주장하며 다음과 같이 다이애나 사례를 수정한 번개 사례를 제시한다.

번개 사례:

이 경우에 다이애나 여신은 없다. 대신 시험관에서 수정란 형성이 진행되는 중에 번개가 치고 그 결과 다이애나 사례에서 다이애나가 개입한 것과 동일한 유전자 변형이 생겨난다. 다른 모든 것은 다이애나 사례와 동일하다.¹⁷⁾

사토리오는 번개 사례에서 어니의 행위에 대한 자신의 직관이 다이애나 사례에서 어니의 행위에 대한 직관과는 매우 달라진다고 주장한다. 즉, 다이애나 사례에서는 어니의 도덕적 책임과 자유가 침해된다고(undermined) 보는 반면, 번개 사례에서는 어니의 행위

14) 김성수 (2018), pp. 381-382. Pereboom (2008), p. 162.

15) Sartorio (2016), p. 160. Sartorio (2018), p. 1550.

16) Sartorio (2016), p. 164.

17) Ibid., pp. 160-161. 김성수 (2018), p. 383에서 재인용.

가 자유행위가 아닌 것으로 여겨지지 않는다는 것이다. 더 나아가 사토리오의 다른 많은 이들도 자신과 비슷한 반응을 보일 것이라고 말한다. 그런데 다이애나 사례와 번개 사례의 차이점은 단 하나이다. 다이애나 사례에는 다이애나의 의도적 개입이 존재하는 반면 번개 사례에는 이러한 의도적 개입이 존재하지 않는다. 사토리오의 유전자 변형이 어디에서 기인했는지가 어니의 행위가 자유행위인지 아닌지 여부의 차이를 결정할 수는 없다고 주장한다. 어니의 행위는 두 사례 모두에서 도덕적 책임이 있는 자유행위이거나 두 사례 모두에서 아니거나 둘 중 하나이다.¹⁸⁾ 번개 사례에 직면하여 어니의 행위에 대한 우리의 직관이 달라진다는 사실은 다이애나의 사례에 대한 우리의 직관을 의심할 수 있게 만들며, 따라서 조작논증의 전제 (1)*에 대하여 심각한 의심을 불러올 수 있다는 것이 사토리오의 논증이다.

5

김성수는 다이애나 사례와 번개 사례에서 대한 우리의 직관에 차이가 있다는 것을 인정하면서도,¹⁹⁾ 사토리오의 반론이 별로 설득력이 없다고 주장한다.²⁰⁾ 그 근거로서 김성수는 이러한 직관 차이에 대한 사토리오의 설명이 충분치 않다고 주장한다.

사토리오의 두 사례에 대해 우리가 가지는 직관의 차이를 심리적인 것에서 찾는다. 사토리오의 다음과 같이 말한다.

어째서 우리의 직관은 이런 식으로 작동하는가? …… 내가 보기에는 ‘책임의 회석’ 같은 어떤 심리적 효과가 작동하는 것 같다.

¹⁸⁾ Sartorio (2016), p. 161.

¹⁹⁾ 김성수 (2018), p. 383.

²⁰⁾ Ibid., pp. 385-386.

…… [싱어와 영거]는 다른 사람들 역시 돕지 않는다는 (따라서 도덕적 의무를 다하지 못하고 있다는) 사실이 우리가 스스로를 덜 책임이 있거나 실제로 그런 것보다 덜 잘못하고 있는 것으로 보게 하는 것이라고 함으로써 이 현상을 해명한다. …… 이 현상의 또 다른 가능한 예로는 총살형 집행부대 같은 과잉인과 (overdetermination)의 사례가 있다. 일반적으로 부대의 개별 대원은 자기 혼자 집행할 때보다 사형수의 죽음에 대해 책임을 덜 느끼리라는 것이 집행부대가 존재하는 이유의 일부이다. 그러나 많은 이들이 이는 단지 그래 보일 뿐이라 생각하는데 …… 다른 대원들의 존재가 실제로 개별 대원이 그 죽음에 덜 책임이 있게 만들지 않기 때문이다. …… 밖에서 볼 때에도 우리는, 실제로는 그렇지 않다고 할 만한데도, 각각의 대원이 덜 책임이 있는 것으로 보는 경향이 있다.²¹⁾

바로 이러한 심리적 이유 때문에 우리는 어머니의 도덕적 책임과 자유행위 여부에 대해 두 사례에서 서로 다른 직관을 갖게 된다는 것이다. 즉, 다이애나 사례에서는 어머니의 책임을 희석시켜줄 수 있는 다른 의도적 행위자가 존재하는 반면 번개 사례에서는 그런 것이 존재하지 않기 때문에 직관이 달라진다는 것이다.

김성수는 사토리오의 주장에 설득력이 부족하다고 주장한다. 김성수는 두 가지 이유를 들고 있다. 첫째, “강경한 대응의 문제를 보완하기 위해서는 책임분산 효과가 발생하지 않는 상황에서 어머니 [가] 책임[이 있다는] 직관이 조작논증의 전제 [(1)*]을 지지하는 직관만큼 분명해야”²²⁾ 하는데 다른 의도적 행위자가 없어 책임분산 효과가 없는 번개 사례에서 어머니의 책임에 대한 직관이 그 정도로 강해보이지 않는다는 것이다. 둘째, 사토리오가 든 책임희석 효과의 경우들과 다이애나의 사례 사이에 여러 가지 중대한 비유사성이 존재한다는 것이다.

21) Sartorio (2016), pp. 167-168. 김성수 (2018), pp. 383-384에서 재인용.

22) 김성수 (2018), p. 385.

6

필자는 김성수가 드는 두 가지 이유 중 먼저 두 번째 이유에 대해 반론을 제기하고자 한다. 책임회석이라는 심리적 효과를 보여주는 동시에 김성수가 지적한 비유사성이 존재하지 않는 예는 다음과 같이 주어질 수 있다.

원격 집행부대:

총살형은 다음과 같이 집행된다. 사형수 앞에 총이 놓여있다. 이 총은 원격으로 조정된다. 원격 집행부대는 반드시 두 명 이상으로 구성되는데 이들은 각자 서로 다른 공간에 있다. 그들의 앞에는 다이얼이 있다. 첫 번째 집행부대원이 먼저 다이얼을 스스로 선택하여 다이얼을 왼쪽 혹은 오른쪽으로 돌리면 다음 집행부대원들이 차례대로 다이얼을 돌려야 한다. 그런데 첫 번째 다이얼 방향이 결정되면 그 다음 다이얼도 모두 같은 방향으로만 움직일 수 있게 고정되고 그 이후의 집행부대원은 그 방향으로 다이얼을 돌려야 한다. 모든 집행부대원은 앞의 사람이 돌린 이후에야 다이얼을 돌릴 수 있다. 그리고 모든 집행부대원은 한 명도 빠짐없이 다이얼을 돌려야 한다. 마지막 집행부대원이 다이얼을 돌리면 원격으로 조정되는 총이 발사된다.

집행부대원들은 자기 혼자만 다이얼을 돌린 것이 아니기에 본인들에게 실제보다 덜 책임이 있는 것으로 느낄 것이다. 심지어 두 번째로 다이얼을 돌린 집행부대원부터는 앞 사람 때문에 그 방향으로 돌릴 수밖에 없었으므로 앞 사람보다 훨씬 덜 책임이 있다고 느끼

기도 할 것이다. 그러나 이런 책임회석 효과는 단지 심리적인 효과이지 실제로 책임을 덜어주는 효과는 없다. 첫 번째 사람이 더 책임이 있는 것도 물론 아니다.

이 새로운 예에는 김성수가 지적인 비유사성이 존재하지 않는다. 김성수가 지적인 첫 번째 비유사성은 “사토리오가 든 책임회석 효과의 경우들에서는, 비슷한 정도의 책임이 있을 수 있는 행위자들이 대략 동시에 그리고 서로 독립성을 유지하면서 행동”²³⁾하는데 다이애나 사례의 경우 그렇지 않다는 것이다. 그러나 원격 집행부대 예에서도 행위자들은 동시에 서로 독립성을 유지하며 행동하지 않는다. 모든 집행부대원의 행위는 마지막 집행부대원의 행위에 도달하는 인과사슬 내에 있으며 뒷사람은 앞사람이 돌리는 방향대로만 다이얼을 돌려야 한다.

두 번째 비유사성은 “집행부대의 경우와는 달리 다이애나 사례에서는 어니를 제거하고 다이애나만 있는 상황에서 (어니의 행위에 대한) 다이애나의 책임에 대해 생각할 수 없다”²⁴⁾는 것이다. 그러나 원격 집행부대에서도 마지막 집행부대원을 제거하고 첫 번째 집행부대원만 있는 상황에서 총살에 대한 첫 번째 집행부대원의 책임에 대해 생각할 수 없다.

세 번째 비유사성은 “다이애나 사례에 책임회석 효과가 적용된다면 다이애나의 책임의 정도에 대한 우리의 직관 역시 영향을 받아야 될 것”²⁵⁾ 같은데 그렇지 않다는 것이다. 우리가 집행부대에서 모든 대원에 대해 각각 덜 책임이 있는 것으로 보는 경향이 있는데 반해 다이애나에 대해서는 어니처럼 덜 책임이 있는 것으로 보게 되지 않는다는 것이다. 그러나 필자가 생각하기에는, 우리가 처음에 다이애나 사례만 보았을 때는 다이애나에게 어니의 행위에 대

23) Ibid.

24) Ibid., p. 385-386.

25) Ibid., p. 386.

한 모든 책임이 있고 어니에게는 책임이 하나도 없다고 생각할 수 있지만 그 이후 번개 사례에 대한 우리의 직관이 변한다는 사실을 본 이후에는 이러한 생각이 흔들리는 것 같다. 그리고 원격 집행부대 예에서 첫 번째 사람에게 보다 많은 책임이 있는 것으로 잘못 생각할 수 있지만 사실은 그렇지 않다는 것을 깨달은 후에 다시 다이애나 사례를 보았을 때는 우리의 직관이 충분히 변할 수 있는 것으로 보인다.

이렇듯 원격 집행부대 예에는 김성수가 말한 세 가지 비유사성이 존재하지 않는다는 점에서 사토리오가 든 예보다 더 발전된 예라고 할 수 있다. 그리고 이 새로운 예를 보았을 때 김성수의 두 번째 반론 이유는 성립하지 않는다.

7

이제 김성수가 든 첫 번째 이유에 대해 살펴보자. 그것은 “강경한 대응의 문제를 보완하기 위해서는 책임분산 효과가 발생하지 않는 상황에서 어니[가] 책임[이 있다는] 직관이 조작논증의 전제 [(1)*]을 지지하는 직관만큼 분명해야”²⁶⁾ 하는데 다른 의도적 행위자가 없어 책임분산 효과가 없는 번개 사례에서 어니의 책임에 대한 직관이 그 정도로 강해보이지 않는다는 것이다. 그러나 과연 정말로 강경한 대응의 문제를 보완하기 위해서는 번개 사례에서 어니가 책임이 있다는 직관이 조작논증의 전제 (1)*, 즉 다이애나 사례에서 어니가 책임이 없다는 직관만큼 분명해야 하는가? 사토리오의 논증이 성공하기 위해서는 정말로 이 조건이 만족되어야 하는가? 필자가 보기에는 그렇지 않다.

사실 사토리오가 염두에 두고 있는 조작논증은 위에서 본 조작

²⁶⁾ Ibid., p. 385.

논증이 아니라 그녀가 단순한 버전(simple version)이라고 부른 다음의 조작논증이다.

다이애나 사례를 통해 본 단순한 버전의 조작논증:

- (1)* 어머니의 행위는 도덕적 책임을 물을 수 있는 자유행위가 아니다.
- (2)** 어머니의 행위는 표준적인(standard) 양립가능주의자들이 내놓은 자유의 조건을 모두 만족한다.
- (3)* 따라서, 모든 표준적인 형태의 양립가능주의는 실패이다.

사토리오의 위에서 제시된 조작논증이 아닌 단순한 버전의 조작논증에 집중하는 이유에 대해 단순한 버전이 더 약한 결론을 확립하고자 하는 만큼 더 강한 논증이기 때문이라고 말한다.²⁷⁾ 그러나 필자가 보기에 사토리오가 염두에 두고 있는 것이 (1)*-(3)의 논증이 아니라 (1)*-(3)*의 논증이라는 점은 이보다 더 중요한 함의를 가진다.

양립불가능주의자들과 강경한 대응이 서로 어떻게 맞서 싸웠는지를 다시 상기해보자.

다이애나 사례를 통해 본 조작논증:

- (1)* 어머니의 행위는 도덕적 책임을 물을 수 있는 자유행위가 아니다.
- (2)* 어머니의 행위와 결정론적으로 발생한 일상적 행위 사이에 차이가 없다.
- (3) 따라서, 결정론적으로 발생한 일상적 행위 역시 자유행위가 아니다.

다이애나 사례를 통해 본 강경한 대응 논증:

- (4) (양립가능주의자들이 내놓은 자유의 조건을 만족하는) 결정론적

²⁷⁾ Sartorio (2016), p. 158.

으로 발생한 일상적 행위는 도덕적 책임을 물을 수 있는 자유행위이다.

- (5) 결정론적으로 발생한 일상적 행위와 어니의 행위 사이에 차이가 없다.
- (6) 따라서, 어니의 행위 역시 도덕적 책임을 물을 수 있는 자유행위이다.²⁸⁾

(1)*이 참이기 때문에 (3)이 참이라는 조작논증이 ‘전건 긍정 전략(modus ponens strategy)’을 취한다면 (3)이 거짓이기 때문에 (1)*도 거짓이라는 강경한 대응은 ‘후건 부정 전략(modus-tollens strategy)’을 취한다. 필자가 보기에 사토리오의 전건 긍정과 후건 부정을 가지고 서로 선결문제요구의 오류를 저지른다고 항변하면서 끝없이 빙빙 돌면서 싸우는 지형으로부터 벗어나 다른 각도에서 우회적으로 접근하는 방식을 택하고자 한 것으로 보인다. 사토리오 자신도 다른 방식으로 접근할 것이라고 말하고 있다.²⁹⁾ 그 다른 식의 전략이란 (3)이 거짓이라는, 즉 (4)가 참이라는 주장으로부터 시작하여 (1)*이 거짓이라는, 즉 (6)이 참이라는 주장으로 향하는 것이 아니라, 번개 사례를 가지고 다른 방식으로 (1)*을 공격하는 것이다. 필자가 보기에는, 바로 그렇기 때문에 사토리오가 가장 먼저 취한 행동이 바로 (1)*-(3)의 논증을 직접적으로 다루지 않겠다고 한 것이었다.

사토리오가 번개 사례를 통해 보이고자 하는 것은 번개 사례에서 어니의 책임에 대한 직관이 조작논증의 전제 (1)*을 지지하는 직관만큼 강하다는 것이 아니다. 사토리오의 훨씬 약한, 그럼에도 훨씬 효과적인 것을 목표로 한다. 사토리오의 다음과 같이 말한다.

²⁸⁾ Ibid., p. 164.

²⁹⁾ Ibid., p. 166.

나의 전략은 [전제 (1)*] 뒤에 깔려있는 직관의 신뢰성을 공격함으로써 [다이애나 사례를 통해 본 단순한 버전의 조작논증]에 대응하는 것이었다. [전제 (1)*]이 직관적으로 참으로 보인다는 것은 인정하면서도 이 직관을 *신뢰하지 말아야* 한다는 논증을 나는 제시했다. 만약 우리가 그 직관을 신뢰할 수 없다면 [다이애나 사례를 통해 본 단순한 버전의 조작논증]은 출발조차 할 수 없게 된다.³⁰⁾

다이애나 사례에 대한 우리의 직관을 신뢰하지 말아야 한다는 주장을 뒷받침하는 사토리오의 논증은 다음과 같은 두 주장이 참이라는 전제에서 시작한다.

- (a) 번개 사례에 대한 우리의 반응과 다이애나의 사례에 대한 우리의 반응 사이에 유의미한 차이가 있다.
- (b) 어니의 도덕적 책임은 다이애나 사례와 번개 사례에서 동일하다.³¹⁾

사토리오가 말하는 유의미한 차이에는 다이애나 사례에서 어니에게 도덕적 책임이 없지만 번개 사례에서 어니에게 도덕적 책임이 있다고 보는 것도 포함하지만 또한 다이애나 사례에서 어니에게 도덕적 책임이 없지만 번개 사례에서 어니에게 도덕적 책임이 있는지 없는지에 대해서는 잘 모르겠다고 보는 것 역시 포함된다.³²⁾ 사토리오는 (b)가 참인데도 우리가 두 사례에 대해 서로 다른 직관적 판단을 하게 된다는 사실은 처음에 다이애나 사례에 대해 가졌던 직관적 판단에 대해 의심을 하게 만든다고 주장한다. 그리고 이 경우 우리가 합리적으로 취해야 하는 태도는 두 사례 모두에 대한 우리

³⁰⁾ Sartorio (2018), p. 1550.

³¹⁾ Ibid.

³²⁾ Ibid.

의 판단을 유보(suspend)해야 하는 것이라고 사토리오의 주장한다.³³⁾ 즉, 이 경우 두 사례에 대한 우리 자신의 판단 그 어느 것도 신뢰할 수 없다는 것이다.³⁴⁾

다이애나 사례를 통해 본 단순한 버전의 조작논증에 대해 반론을 가하는 사토리오의 논증은 다음과 같이 정식화될 수 있다.

사토리오의 대응:

- (7) 번개 사례에 대한 우리의 반응과 다이애나의 사례에 대한 우리의 반응 사이에 유의미한 차이가 있다. ((a)에 해당)
- (8) 어니의 도덕적 책임은 다이애나 사례와 번개 사례에서 동일하다. ((b)에 해당)
- (9) 어니의 도덕적 책임이 다이애나 사례와 번개 사례에서 동일함에도 불구하고, 번개 사례에 대한 우리의 반응과 다이애나의 사례에 대한 우리의 반응 사이에 유의미한 차이가 있다는 사실은 다이애나 사례에 대한 우리의 반응을 의심해야 한다는 것을 보여준다.
- (10) 이 경우 우리가 가져야 하는 합리적인 태도는 두 사례에 모두 대한 우리의 판단을 유보하는 것이다.
- (11) 따라서, (1)*이 참인지 거짓인지에 대해 우리는 판단을 유보해야 한다.

이 논증의 결론은 다이애나 사례를 통해 본 단순한 버전의 조작논증의 전제 (1)*, 즉 “어니의 행위는 도덕적 책임을 물을 수 있는 자유행위가 아니다”라는 명제가 참인지 거짓인지에 대해 판단을 유

³³⁾ Ibid., p. 1550-1551.

³⁴⁾ Ibid., p. 1551.

보해야 한다는 것이다. 이는 다이애나 사례를 통해 본 단순한 버전의 조작논증의 첫 번째 전제를 참으로서 받아들이지 말아야 한다는 것이며, 이렇게 되면 이 단순한 버전의 조작논증은 출발조차 할 수 없게 된다.

여기에서 우리가 주목해야 할 것은 사토리오가 확립하고자 한 결론이 번개 사례에서 어니의 책임에 대한 직관이 조작논증의 전제 (1)*을 지지하는 직관만큼 강하다는 것이 아니라는 점, 그리고 이것이 결론 (11)을 확립하기 위해 필요하지도 않다는 점이다.³⁵⁾ 사토리오가 확립하고자 한 결론은 단순한 버전의 조작논증의 첫 번째 전제 (1)*을 참이라고 받아들이지 않을 이유가 있다는 것이며 이를 위해서는 (1)*이 거짓이라는 것을 반드시 보일 필요가 없다. 더 나아가 우리가 반드시 번개 사례에서 어니가 책임이 있다는 직관을 가져야 할 필요도 없다. 단지 두 사례에 대한 우리의 직관이 서로 다르다는 사실만이 필요하며 이것은 “잘 모르겠다”는 불가지론적 태도로도 충분하다.

김성수의 주장과는 달리, 사토리오의 논증이 성공하기 위해서는 번개 사례에서 어니가 책임이 있다는 직관이 다이애나 사례에서 어니가 책임이 없다는 직관만큼 분명해야 할 필요가 없다. 따라서 이 조건이 만족되지 않았다는 이유로 사토리오의 논증에 문제점을 제기하는 것은 적절치 않다.

8

다이애나 사례를 통해 양립불가능주의를 지지하는 조작논증에 대해

³⁵⁾ 우리가 왜 두 사례에 대해 서로 다른 직관을 가지게 되는가에 대한 설명도 사토리오의 논증을 위해서는 필요하지 않다는 점에도 주목하자. 사토리오의 논증에서 이 설명은 아무런 역할도 하지 않기 때문이다. 그래서 사토리오는 자신을 비롯한 양립가능주의자가 이에 대한 설명을 제시하는 것이 반드시 필요하지는 않다고 말하고 있다. Sartorio (2016), p. 170.

사토리오의 번개 사례를 통해 이에 반론을 가한다. 김성수는 강경한 대응을 보장하기 위해 제시된 사토리오의 논증이 설득력이 부족하다고 주장한다. 사토리오의 책임희석 효과에 대한 설명에 문제가 있다는 김성수의 반론에 대해서 필자는 이러한 문제점을 피할 수 있는 원격 집행부대 예를 새롭게 제시한다. 번개 사례에서 어니가 책임이 있다는 직관이 조작논증의 전제 (1)*을 지지하는 직관만큼 강해보이지 않는다는 김성수의 반론에 대해 필자는 사토리오의 논증이 성공하기 위해서 번개 사례에서 어니가 책임이 있다는 직관이 다이애나 사례에서 어니가 책임이 없다는 직관만큼 분명해야 할 필요는 없었다고 논증한다. 따라서 사토리오에 대한 김성수의 반론은 성공적이지 않다.

참고문헌

- 김성수 (2016), “자유의지에 대한 조작논증과 근원-양립가능주의의 대응”, 『논리연구』, 21집 3호, pp. 373-393.
- Mele, A. (2006), *Free Will and Luck*, Oxford University Press.
- Pereboom, D. (2008), “A Hard-Line Reply to the Multiple-Case Manipulation Argument”, *Philosophy and Phenomenological Research*, 77, pp. 160-170.
- Sartorio, C. (2016), *Causation and Free Will*, Oxford University Press.
- Sartorio, C. (2018), “Replies to Critics”, *Philosophical Studies*, 175, pp. 1545-1556.

이화여자대학교 스크랜튼학부

Scranton Honors Program, Ewha Womans University
seahwak@ewha.ac.kr

The Manipulation Argument: Ernie, Diana, and Lightning Strike

Seahwa Kim

In this paper, I raise objections to Sungsu Kim's argument that Sartorio's hard-line reply to the manipulation argument fails. In attacking Sartorio's argument, Sungsu Kim claims that there are two problems with Sartorio's. I argue that Sungsu Kim's argument fails by responding to these two problems. With respect to the first problem, I provide a new example of dilution of responsibility. With respect to the second problem, I argue that, contrary to what Sungsu Kim assumes, for Sartorio's argument to succeed, our intuition that Ernie is responsible in the Lightning Strike Scenario does not have to be as strong as our intuition that Ernie is not responsible in the Diana scenario.

Key Words: The manipulation argument, Free action, Moral responsibility, Hard-line reply, Sartorio, Diana scenario, Lightning strike scenario, Sungsu Kim