

【논문】

인과와 도덕적 책임

—행위와 부재 간의 비대칭성 주장에 대한 비판—*,**

김 성 수

【주제분류】 현대영미철학, 형이상학

【주요어】 인과, 도덕적 책임, 행위, 부재

【요약문】 행위와 행위의 부재에 대한 ‘도덕적 비대칭성’ 주장에 따르면, 행위의 결과가 실제 결과와 다를 수 없는 경우에도 행위자는 자신의 행위에 대해 책임이 있을 수 있는 반면, 부재의 결과가 실제와 다를 수 없는 경우 행위자는 행위의 부재에 대해 책임이 있을 수 없다. 이 논문은 ‘스위치’라고 불리는 인과 구조의 성격에 근거하여 행위와 부재 간의 인과적 비대칭성을 보임으로써 도덕적 비대칭성을 지지하고자 하는 논증을 비판적으로 살펴본다. 특히 ‘스위치’를 특징짓는 두 가지 인과 구조의 차이점을 밝히고 이 논증이 성립하기 위해 추가적으로 요구되는 주장의 문제점을 보인다. 또한 ‘실제 인과’ 이론의 입장에서 이 논증이 주장하는 행위와 부재 간의 인과적 비대칭성이 성립하지 않음을 주장한다.

* 이 논문은 2012년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 연구되었음 (NRF-2012S1A5A8023935).

** 이 논문의 심사를 맡아 유익한 비판과 도움말을 해주신 익명의 심사위원님들께 감사드립니다.

I. 들어가는 말

자신이 한 일로 인해 생긴 결과에 대해 우리는 종종 그것이 자기 탓이라거나 혹은 자기 덕분이라고 한다. 이러한 ‘탓’이나 ‘덕분’이라는 것을 행위에 대한 ‘도덕적 책임’(moral responsibility)이라고 하자.¹⁾ 이러한 행위와 책임의 관계는 신체의 움직임을 동반하는 ‘긍정적인’(positive) 행위(이하 ‘긍정적 행위’ 대신 ‘행위’로 표기함)뿐만 아니라 행위의 자제나 누락과 같은 ‘행위의 부재’(이하 ‘행위의 부재’ 혹은 ‘부재’로 표기함)의 경우에도 성립한다. 예를 들어 물에 빠진 아이를 구한 것에 대해서뿐만 아니라 그러지 않은 것에 대해서도 행위자(agent)에게 책임을 물을 수 있다. 그리고 행위자에게 그 책임을 물을 수 있는 이유는 행위자의 행위나 행위의 부재가 원인이기 때문이다.

인과 논제: 행위자는 어떤 일이 자신의 행위의 결과이거나 행위의 부재의 결과인 한에서만 그 일에 대해 도덕적 책임이 있다.

인과 논제는 행위와 그 결과 또는 부재와 그 결과 간에 성립하는 인과 관계가 도덕적 책임에 대한 필요조건임을 주장한다.²⁾ 따라서 이 조건이 만족되는 경우 도덕적 책임이 성립할 수 있지만, 반면 이 조건이 만족되지 않으면 도덕적 책임이 성립할 수 없게 된다. 이 글에서는 인과 논제를 받아들이는 입장에서 행위에 대한 도덕적 책임과 행위의 부재에 대한 도덕적 책임 간의 관계에 대해 살펴보고자 한다. 특히 행위에 대한 도덕적 책임과 행위의 부재에 대한 도덕적 책임의 관계가 항상 대칭적인지에 대한 논쟁을 인과

-
- 1) 칭찬할만한 행위에 대해 그 행위의 ‘책임’을 행위자에게 부여하는 것은 일상적 용법에 비추어볼 때 부자연스러울 수 있으나, 이 논문에서는 관련된 철학적 논의의 관례에 따르기로 한다.
 - 2) 행위의 결과에 대한 도덕적 책임을 부여하기 위해서는 그 행위가 저항하기 힘든 강요나 강압에 의한 것이 아니라는 의미에서 ‘자유로운’ 행위일 것과 행위자가 행위의 결과에 대해 예견하는 것이 불가능하지는 않았을 것 등과 같은 추가조건이 요구된다. 익명의 심사자가 지적하듯 이들 추가조건에 대한 논의 및 도덕적 책임과 법적 책임 간의 관계에 대한 논의는 매우 중요하기는 하지만 이 논문의 범위를 벗어나 있음을 밝혀둔다.

적 관점에서 살펴볼 것이다.

행위와 부재 간의 ‘도덕적 비대칭성’ 주장에 따르면, 행위와 부재는 도덕적 책임의 부여에 있어 어떤 특정한 조건 하에서는 서로 차이를 보인다.

도덕적 비대칭성 논제: 행위의 결과가 실제 결과와 다를 수 없는 경우에도 행위자는 자신의 행위에 대해 책임이 있을 수 있는 반면, 부재의 결과가 실제와 다를 수 없는 경우 행위자는 행위의 부재에 대해 책임이 있을 수 없다.

도덕적 비대칭성 논제가 다루고 있는 것은 행위의 결과나 부재의 결과가 실제 결과와 다를 수 없는 경우이다. 잘 알려진 것처럼 프랑크푸르트(Frankfurt 1969)는 행위자가 달리 선택할 수 없어서 그 결과를 피할 수 없는 사례를 (이런 사례를 ‘프랑크푸르트 유형’의 사례라고 한다) 사용하여 도덕적 책임의 논의에서 중요한 역할을 해왔던 대안 가능성(alternative possibilities) 이론을 비판하였다. 더 나아가 그는 위 주장과는 달리 행위와 부재 간에 도덕적 대칭성이 성립한다고 주장하였다.³⁾ 하지만 도덕적 비대칭성 논제를 지지하는 철학자들은 일반적으로 행위와 부재 간에는 인과적 비대칭성이 성립하기 때문에 도덕적 비대칭성 논제 역시 성립한다고 주장한다. 예를 들어, 피셔와 라비자(Fisher & Ravizza 1991)는 “가능한 신체 움직임과 행위자 간의 관계와 관련하여 행위와 부재는 서로 차이가 있다; 행위의 경우 이 관계는 일종의 인과 관계이지만 부재의 경우에는 그런 종류의 인과가 결여되어 있다”⁴⁾고 주장한다. 그리고 사토리오(Sartorio 2005)는 “도덕적 비대칭성은 행위와 부재가 서로 다른 인과력을 갖는다는 인과 원칙에 기대고 있다”⁵⁾고 주장한다. 흥미 있는 점은, 이들 철학자들은 행위의 부재가 어떤 일의 원인이 되는 경우가 있다는 점을 인정하거나 가정함에도 불구하고, 행위 또는 부재로 발생한 결과를 피할 수 없는 경우에는 인과

3) “내가 보기에, 완전히 대칭적인 이론을 선호할 충분한 이유가 있다. 행위의 경우 도덕적 책임이 요구하는 것이 단지 관련된 움직임의 ‘실제 인과적 제어’(actual causal control)라면, 부재의 경우에 도덕적 책임이 요구하는 것은 관련된 움직임의 부재에 대한 ‘실제 인과적 제어’이다.” (Frankfurt 1994, 621-22쪽, 강조는 원문을 따름)

4) Fisher & Ravizza 1991, 270쪽.

5) Sartorio 2005, 461쪽, 강조는 원문을 따름.

적·도덕적 비대칭성이 성립한다고 주장한다는 것이다.

사토리오는 (곧 보게 되겠지만) 행위를 포함하는 프랑크푸르트 유형의 사례와 부재를 포함하는 프랑크푸르트 유형의 사례에 인과적 비대칭성이 성립한다는 주장을 통해 도덕적 비대칭성 논제를 지지한다. 이를 위해 사토리오는 ‘스위치’라고 알려진 인과 구조가 갖는 특성에 기대어, 부재를 포함하는 프랑크푸르트 유형의 사례는 도덕적 책임에 대해 인과 논제가 요구하는 조건을 만족하지 않는다고 주장한다. 다음 절에서는 이 논변을 소개하고 그 문제점을 비판적으로 살펴볼 것이다. 특히 스위치와 연관된 두 가지 서로 다른 인과 구조를 구분하고, 이를 통해 스위치의 인과적 성격에 근거한 논증이 성공하기 위해 요구되는 것이 무엇인지 밝히고 그것이 갖는 문제점을 논의할 것이다. 그 다음으로, 인과 구조 파악에 실제 상황을 반영하여 원인과 결과 간의 의존 관계를 이해하는 ‘실제 인과’ 개념을 사용하여 행위를 포함하는 프랑크푸르트 유형의 사례와 부재를 포함하는 프랑크푸르트 유형의 사례를 살펴볼 것이다. 특히 이 두 가지 유형의 사례 간에 인과적 대칭성이 성립함을 보임으로써 도덕적 책임의 문제와 독립적으로 전개된 인과 이론 역시 도덕적 비대칭성 논제에 반하는 결론을 내놓는다는 점을 주장할 것이다.

II. ‘스위치’와 행위와 부재 간의 인과적 비대칭성

언뜻 보기에, 행위와 책임 간의 관계에 대해 ‘우리는 (자신이 실제 한 것과) 달리 행동할 수 있었을 경우에만 자신의 행위에 대해 책임이 있다’는 대안 가능성의 원리가 성립하는 것 같다. 실제로 행한 행위 이외의 다른 행위를 할 수 있는 가능성 자체가 봉쇄된 상황에서는 행위자에게 그 행위에 대한 책임을 따지는 것이 어려울 것 같아 보이므로, 대안 가능성은 행위자의 책임이 성립하기 위한 필요조건인 것으로 여겨져 왔다. 하지만 다음과 같은 프랑크푸르트 유형의 사례는 이 원리의 문제점을 설득력 있게 보여준다.

(F-행위) 철수는 (강압이 아니라 자율적으로) 희생자를 죽이기로 결심했다. 그는 방아쇠를 당겼고 그 결과 희생자가 죽었다. 그런데 철수 자신은 모르고 있었지만, 악한 과학자가 철수의 뇌를 모니터링하고 있었다. 그 과학자는, 만약 철수가 희생자를 죽이지 않을 이유를 심각하게 고려하여 방아쇠를 당기겠다는 그의 결심이 흔들리게 되었다라면 철수의 뇌에 신호를 보내 방아쇠를 당기기로 결심하도록 만들었을 것이다. 실제로 철수의 결심은 흔들리지 않았고 과학자는 개입하지 않았다.

이 경우 철수는 방아쇠를 당기는 것 이외의 다른 결심이나 선택을 할 수 없었고 따라서 희생자를 쏘는 일과 그 결과로 발생하는 희생자의 죽음은 피할 수 없는 일이었다. 그럼에도 불구하고, 실제로 희생자의 죽음을 초래한 것은 철수 자신의 결정에 따른 그 자신의 행위이며 철수는 이에 대해 도덕적 책임이 있다.

이제 (F-행위)와 유사하지만 행위의 부재를 포함하는 프랑크푸르트 유형의 사례를 생각해보자.

(F-부재) 철수는 어린이가 물에 빠진 것을 보았고 자신이 물에 뛰어들어 그 아이를 구할 수 있다고 생각했다. 하지만 그는 물에 뛰어들지 않기로 결정했다. 그런데 철수는 모르고 있었지만, 악한 과학자가 그의 뇌를 모니터링하고 있었다. 그 과학자는, 만약 철수가 물에 뛰어들 이유를 심각하게 고려하여 다른 결정을 내리려고 했었다라면 그의 뇌를 조종하여 물에 뛰어들지 않겠다는 결정을 하도록 했을 것이다. 하지만 악한 과학자는 개입하지 않았고 아이는 익사하였다.

이 경우 철수는 물에 뛰어들지 않기로 결심하는 것 이외의 다른 결심이나 선택을 할 수 없었고 또한 아이가 익사하는 것은 피할 수 없는 일이었다. 이 사례는 행위 대신 행위의 부재를 포함한다는 차이가 있긴 하지만 (F-행위)와 매우 유사해 보인다. 이러한 유사점에도 불구하고 도덕적 비대칭성 논제의 근거를 두 사례 간의 인과적 차이에서 발견하려는 시도는 놀랍기도 하고 새롭기도 하다.

사토리오(Sartorio 2005)는 (F-행위)와 (F-부재) 간에 인과적 차이가 있다는 이유를 들어, 전자의 경우 철수는 희생자의 죽음에 대해 도덕적 책임이 있지만 후자의 경우 철수는 아이의 죽음에 대해 도덕적 책임이 없다고

주장한다. 이 주장을 지지하기 위해 사토리오스는 ‘스위치’라고 불리는 인과 구조에 주목하는데, 그 주장은 다음과 같이 요약될 수 있다.

- (1) (F-부재)에서 물에 뛰지 않기로 한 철수의 결정은 스위치 역할을 한다.
- (2) ‘스위치’라는 인과 구조에서, 스위치 역할을 하는 것은 원인이 아니다.
- (3) 따라서 (F-부재)의 경우 철수의 행위 부재는 아이 죽음의 원인이 아니며 철수는 아이의 죽음에 도덕적 책임이 없다.

이 주장을 이해하기 위해, 먼저 전형적인 스위치의 사례에 대해 생각해 보자.

(기차) 선로를 따라 내달리는 기차가 있고 그 선로 앞쪽에 희생자가 묶여있다. 스위치 옆에 서 있던 철수는 스위치를 당겼고 그 결과 오른쪽 선로를 달리던 기차는 왼쪽 선로로 들어섰다. 그러나 두 선로는 희생자가 묶여있는 지점 앞에서 다시 합쳐졌고, 희생자는 기차에 치여 죽었다.⁶⁾

직관적으로 철수가 스위치를 당긴 것은 희생자 죽음의 원인이 아니다. 그렇다면 그 이유는 어디에서 찾을 수 있는가? 사토리오스는 스위치가 다음과 같은 인과 구조를 갖는다고 주장한다.

스위치의 인과 구조: 스위치는 ① 행위자가 기존의 위협(threat)이 그대로 진행되도록 내버려두거나 ② 그것의 경로(course)를 단지 살짝 바꾸기만 하는 그런 인과 구조이다.⁷⁾

(기차)에서 철수는 ‘기차가 희생자를 향해 달려오고 있다’는 기존 위협의 경로를 살짝 바꾸기만 하였다. 반면 ‘비-스위치’에서 행위자는 기존의 위협을 증진시키거나 새로운 위협을 창출하거나 한다. 예를 들어, 암살자가 희생자를 향해 총을 겨누고 있는 상태에서 철수가 먼저 총을 쏘았다면 철수는 새로운 위협을 창출한 것이다. 한편, 암살자가 총을 쏜 직후 희생자 앞을 막고 있는 방탄장치를 제거하여 희생자가 총에 맞아 죽었다면 철수의 제거 행

6) 이 예는 Rowe (1989), Sartorio (2005), Hall (2007)에서 (유사한 형태로) 제시된 예이다.
 7) Sartorio 2005, 472-75쪽.

위는 기존의 위협을 증진한 것이다.⁸⁾

이제, 스위치의 인과적 구조에 대한 분석을 토대로 앞서 논의한 (F-행위)와 (F-부재)에 대해 다시 생각해보자. (F-행위)에서 철수의 행위는 그전에 없던 새로운 위협을 창출했고 그런 점에서 그의 행위는 희생자 죽음의 원인이다. 반면 (F-부재)에서 물에 뛰어들지 않았다는 행위 부재는 이미 물에 빠져 허우적거리는 아이의 죽음과 연관된 새로운 위협을 창출하지 않았고 기존의 위협을 증진시키지도 않았으며, 대신 기존의 위협이 진행되도록 내버려두었을 뿐이다. 따라서 스위치와 비-스위치 간의 인과적 비대칭성이 성립한다면 이는 행위와 부재 간의 도덕적 비대칭성을 지지하게 된다.

이 주장을 평가하기 위해 먼저 스위치는 두 종류의 인과 구조로 구성된다라는 점에 주목할 필요가 있다. 스위치의 전형적인 사례인 (기차)는 ②를 만족한다. (기차)의 경우 실제로 기차는 선로를 따라 내달리고 있었고, 이러한 상황에서 철수가 스위치를 조작하는 것은 주어진 결과가 발생하는 물리적 경로만 바꿀 뿐 결과에 대해 어떤 인과적 차이도 만들지 못한다. 따라서 스위치를 조작하는 것이 희생자 죽음의 원인일 수 없다는 것이 분명해 보인다. 그러나 (F-부재)는 ②의 경우가 아니라 사토리오가 스위치를 구성한다고 주장하는 또 다른 인과 구조인 ①의 경우이다. 그렇다면 어떻게 ②의 경우로부터 ①의 경우에도 동일한 결론을 이끌어 낼 수 있을까? 사토리오는 이 점에 대해 전혀 언급하고 있지 않지만, 이런 결론을 이끌어 낼 수 있는 방법이 있어 보인다. 즉, 만약 ②의 경우인 기존 위협의 경로를 단지 살짝 바꾸기만 하는 것이 원인이 될 수 없다는 점에 동의한다면, ①과 같이 기존의 위협을 그 경로조차 바꾸지 않고 그대로 진행되도록 내버려두는 경우 역시 원인이 될 수 없다고 생각할 수 있다. 그렇다면 스위치의 인과 구조에 근거한 논증에 문제가 있음을 보이는 방법은 이러한 생각의 문제점을 보이거나 ②의 인과적 성격에 대해 동의한다고 해서 ①의 인과적 성격에 대해 동의해야만 하는 것은 아님을 보이는 것이다.

①에서처럼 기존의 위협이 그대로 진행되도록 내버려두는 것은 행위의 부재나 사건의 부재가 갖는 전형적인 특징이다. 예를 들어, 악한 과학자와

8) Sartorio 2005, 472-73쪽.

같은 존재가 없는 상황에서, 철수가 물에 뛰어들지 않은 경우 이러한 행위의 부재는 기존의 위협이 그대로 진행되도록 내버려두는 역할을 한다. 하지만 이 경우 행위의 부재는 아이 죽음의 원인이며 철수는 이에 책임이 있다. 사토리오 역시 이 점에 동의하지만, 그 결과를 피할 수 없는 경우에는 인과적 성격이 달라진다고 주장한다.

부재는 기존의 위협을 증진시키지 않거나 새로운 위협을 만들어내지 않는다고 할 때, 부재는 결코 어떤 것의 원인도 될 수 없다고 제안하고 있는 것은 *아니라*는 점을 강조하는 바이다. 많은 경우 어떤 과정이 특정한 방식으로 진행되도록 내버려두는 것은 어떤 결과의 원인이 되기에 충분하다. [...] 내가 제안하는 것은, *결과가 발생할 수밖에 없을 때* 부재는, 첫째 결과가 발생할 수밖에 없으며 둘째 그 부재가 위협을 증진시키지도 창출하지도 않는다는 것이 주어지면, *결과*의 원인이 아니라는 것이다. 이것이 바로 [기차]에서 일어나고 있는 것이다.⁹⁾

(기차)는 물론 결과가 발생할 수밖에 없는 경우이고 기존의 위협을 증진시키거나 창출하지 않는 경우이다. 하지만 앞서 보았듯 (기차)는 ②의 경우이지 ①의 경우가 아니다. 물론 ②의 경우 부재는 그 ‘결과’가 발생할 수밖에 없다는 특징을 갖지만, 단지 ②의 경우로부터 ①의 경우를 이끌어낼 수는 없다. 특히 기존의 위협이 그대로 진행되도록 내버려두는 것이 행위나 사건의 부재가 갖는 전형적인 특징이라는 점에 Sartorio가 동의하고 있다는 점을 생각해보면, 단지 그 결과가 발생할 수밖에 없다는 조건을 추가함으로써 어떻게 ①에 대한 주장이 성립하는지 알기 어렵다.

이상의 논의는 스위치가 인과적으로 통합된 단일한 구조를 갖지 않음을 제시할 뿐만 아니라, ①을 만족시키는 경우에도 역시 인과 관계가 성립하지 않음을 주장하기 위해서는 별도의 논증이 요구된다는 것을 보여준다. 사토리오는 이에 대해 직접적인 논증을 제시하고 있지 않지만 부재에 관한 사토리오의 논의에서 관련 논증을 재구성할 수 있다. 아래에서는 관련된 두 가지 논증을 재구성하고 이를 통해 ①에 관한 사토리오의 주장을 비판적으로 살펴보고자 한다.

9) Sartorio 2005, 474쪽, 강조는 원문을 따름.

사토리오스는 먼저 (F-부재)의 경우 물에 떨어지지 않겠다는 철수의 결심은 아이의 죽음의 원인이 아니라고 주장한다.

(2') 물에 떨어지지 않기로 한 [철수의] 결정은 아이의 죽음의 원인이다. [...] 그러나 (2')는 거짓이다. 아마도(arguably) 아이의 죽음의 원인은 물에 떨어지지 않겠다는 [철수의] 결정이 아니라 물에 떨어졌다는 결정을 하는 데 실패한 것이다. [...] 물에 떨어지지 않겠다는 결정과 물에 떨어졌다는 결정의 부재 중 어느 것이 아이 죽음에 대한 더 나은 원인인가? 아마도 후자이다. [...] [철수]가 물에 떨어지지 않기로 결정했다는 사실은 그 사실이 [철수]가 물에 떨어지기로 결정하지 않았음을 함축하는 한에서만 아이 죽음과 상관있다. (반약 [철수]가 물에 떨어지지 않기로 결정했다면 그는 물에 떨어지기로 결정할 수가 없고 따라서 그는 아이를 구할 수 없었다.) 이는 아이의 죽음과 상관있는 것은 물에 떨어지지 않기로 결정했다는 사실이 아니라 물에 떨어지기로 결정하지 않았다는 사실이라는 것을 제시한다.¹⁰⁾

사토리오스가 주장하는 것과 반대로, 철수가 물에 떨어지지 않기로 결정한 것이 아이 죽음의 원인이라고 해보자. 물에 떨어지지 않기로 결심한 것은 일종의 정신적 행위이다. 도덕적 비대칭성 논제와 인과 논제에 따르면, 이 결정은 정신적 행위이므로 피할 수 없는 결과의 원인이 될 수 있고 행위자는 그 결과에 대해 책임이 있을 수 있다. 또한 물에 떨어지지 않기로 결심한 것은 철수가 물에 떨어지지 않은 행위 부재의 원인이며 이 결정은 이러한 행위의 부재를 인과적 매개로 하여 아이 죽음의 원인이 된다. 그렇다면 이 인과 고리의 중간에 있는 철수의 행위 부재 역시 아이 죽음의 원인이며 철수는 아이의 죽음에 대해 도덕적 책임이 있을 수 있다. 하지만 이는 도덕적 비대칭성 논제에 반하는 주장이다. 이 점을 고려해보면 도덕적 비대칭성 논제를 지지하기 위해서는 적어도 아이 죽음의 원인이 철수가 물에 떨어지지 않기로 결정한 것이 아니라는 점을 보일 필요가 있다.

그러나 (F-부재)의 경우 철수는 아이가 물에 빠져 허우적거리는 것을 보았고 자신이 그 아이를 구할 수 있다고 믿었지만 물에 떨어지지 않기로 결정했고, 그 결정에 따라 물에 떨어지지 않았다. 이때 물에 뛰어드는 행위 부

10) Sartorio 2005, 464-65쪽, 강조는 원문을 따름.

재의 원인은 단순히 물에 떨어졌다는 결정의 부재가 아니라, 물에 떨어지지 않겠다는 결정이라는 점이 분명해 보인다. 사토리오는 이 기술 자체를 부인하지는 않지만, 위의 인용문에 나타난 함축 관계에 근거한 논증을 통해 (F-부재)의 경우 아이 죽음의 ‘진짜’ 원인은 이러한 결정이라기보다는 결정의 부재임을 주장한다. 그렇다면 어떤 근거에서 이런 주장을 할 수 있는 것인가? 사토리오는 다음과 같은 인과에 대한 함축 관계를 고려하고 있는 것으로 보인다.

만약 $D(c_1)$ 이 $D(c_2)$ 를 함축하고 c_2 가 e 의 원인이면 c_1 은 e 의 원인이 아니다
(여기에서 $D(c_1)$ 와 $D(c_2)$ 는 각각 c_1 과 c_2 의 기술이다).

이를 (F-부재)에 적용해보면, 물에 떨어지지 않기로 결심한 것은 물에 떨어지기로 결심하지 않았음을 함축하며 물에 떨어지기로 결심하지 않았음이 아이 죽음의 원인이라면 이를 함축하는 사실은 원인이 아니다. 함축에 관한 이 주장은 일견 설득력이 있어 보인다. 예를 들어 영희가 유리창을 향해 녹색 공을 던졌음은 영희가 유리창을 향해 공을 던졌음을 함축한다. 그리고 영희가 공을 던지지 않았더라면 유리창은 깨지지 않았을 것이지만, 영희가 던진 공이 녹색이 아니었다라도 유리창은 여전히 깨졌을 것이다. 유리창이 깨진 것의 원인은 영희가 녹색 공을 던졌다는 것이 아니라 그녀가 공을 던졌다는 것이다.

하지만 사건의 기술 간에 성립하는 함축과 인과의 관계에 관한 이 주장이 일반적으로 성립하는 것은 아니다. 예컨대 영희가 유리창을 향해 무게가 0.2 Kg인 공을 2 Kg m/s의 운동량으로 던졌다고 해보자. 그리고 유리창이 깨지기 위해서는 적어도 2 Kg m/s의 운동량이 필요하다고 해보자. 이 경우, 영희가 0.2 Kg인 공을 2 Kg m/s의 운동량으로 공을 던졌다는 사실은 공의 속도가 10 m/s임을 함축한다. 그리고 공이 10 m/s 속도를 갖도록 던진 것은 분명 유리창이 깨진 것의 원인이다.¹¹⁾ 사건의 기술 간의 함축과 인과

11) 이 경우 함축의 관계는, 운동량은 질량과 운동 속도를 곱한 것이라는 정의를 전제로 요구한다. 반면 ‘녹색 공을 던졌다’가 ‘공을 던졌다’를 함축하는 것은 별도의 전제를 요구하지 않는다. 이 점 때문에 위 반례에 문제가 있다고 생각할 수 있다. 그렇다면 다음

에 대한 앞서의 주장이 옳다면, 영희가 2 Kg m/s의 운동량을 갖도록 공을 던졌다는 것은 유리창이 깨진 것의 원인이 될 수 없다. 하지만 영희가 2 Kg m/s의 운동량을 갖도록 공을 던졌다는 것이 유리창이 깨진 것의 원인이라는 점을 부인할 수 없다. 따라서 단지 함축 관계에 근거하여 아이 죽음의 원인은 철수가 물에 뛰어들기로 결정하지 않은 것이지 철수가 물에 뛰어들지 않기로 결정한 것이 아니라는 주장을 정당화할 수 없다. 물론 함축과 관련된 인과 주장이 일반적으로 성립하지 않는다고 해서 이 경우에도 반드시 성립하지 않으리라는 보장은 없다. 하지만 이는 별도의 논증을 요구한다.

논의를 더 진행하기에 앞서, 설령 아이 죽음의 원인이 물에 뛰어들기로 결정하는 것에 실패한 것, 즉 물에 뛰어들기로 결정하지 않은 것이라는 주장을 받아들인다고 해도 그것만으로는 도덕적 비대칭성 논제를 지지할 수 없다는 점에 주목할 필요가 있다. 만약 물에 뛰어들기로 결정하지 않았다는 정신적 행위의 부재가 아이 죽음의 원인이라면 이 부재는 아이 죽음이 피할 수 없는 것일지라도 그것의 원인이 된다. 그리고 철수의 정신적 행위의 부재는 물에 뛰어드는 행위 부재의 원인이며 이를 통해 아이 죽음의 원인이 된다. 그렇다면 물에 뛰어드는 행위의 부재는 아이 죽음이 피할 수 없는 것이라고 해도 여전히 그것의 원인이 된다. 따라서 이것만으로는 도덕적 비대칭성 논제를 옹호하는 인과적 근거가 되지 못한다. 그렇다면 행위와 부재의 인과적 비대칭성을 통해 도덕적 비대칭성을 주장하기 위해 사토리오의 추가적인 근거를 제공할 필요가 있다.

(F-부재)와 같은 사례를 논의하면서 사토리오(Sartorio 2009)는, 철수가 물에 뛰어들지 않은 것은 “내가 보기에, 주어진 시간에 방 안에 코끼리가 없음이 그 시간 바로 전에 방 안에 코끼리가 없었음에 의해 설명되는 것과 유사한 방식으로—행위자가 그 일을 할 것을 누락한(omit) 다른 것들로부터

의 경우에 대해 생각해보자. 영희는 탈진한 사람을 담요로 덮어주고 물을 먹였고 그 결과 그 사람은 회복하였다. 이 경우 탈진한 사람을 담요로 덮어주고 물을 먹인 것은 그 사람이 회복한 것의 원인이다. 그리고 ‘담요로 덮어주고 물을 먹였다’는 것은 ‘물을 먹였다’는 것을 함축한다. 그런데 물을 먹였음이 회복의 원인임이 분명하다. 그리고 이 경우 함축 관계는 별도의 전제를 요구하지 않으므로 함축과 인과에 대한 위의 주장에 문제가 있음을 다시 확인할 수 있다. 녹색 공의 사례와 운동량의 사례 간의 차이점을 지적해준 익명의 심사자에게 감사드린다.

‘흘러나오는’(flow from) 것 같다”¹²⁾고 주장한다. 철수가 물에 뛰어들지 않은 것의 원인이 물에 뛰어들기로 결정하는 것의 부재라고 해보자. 이것과 철수의 행위의 부재 즉 그가 물에 뛰어들지 않은 것 간의 관계는, 사토리오에 따르면 단순히 ‘흘러나오는’ 것이다. 사토리오는 아마도 이렇듯 ‘단순히 흘러나오는’ 경우는 인과의 경우가 아니며 (F-부재) 역시 그런 경우임에 분명하다고 가정하는 것 같다. 그렇다면 철수가 물에 뛰어들지 않은 것의 원인이 물에 뛰어들기로 결정하는 것의 부재라는 주장과 부재와 부재 간에는 인과 관계가 성립하지 않는다는 주장으로부터 철수의 정신적 행위 부재는 아이 죽음의 원인이 아니라는 결론이 따라 나온다.

사토리오의 두 번째 주장 즉 부재와 부재 간에는 인과 관계가 성립하지 않는다는 주장과 관련하여 먼저 다음의 경우를 생각해보자. 영희는 점심을 먹기로 결정했다. 이는 그녀가 조깅하려 결정하지 않았음, 연못에 뛰어들기로 결정하지 않았음, 의자를 수리하기로 결정하지 않았음 등을 의미한다. 이러한 결정의 부재는 ‘단순한’ 정신적 부재이다. 이러한 단순한 부재는 다른 상태가 성립하였기 때문에 그리고 단지 그 이유로 ‘성립하는’ 부재이다. 만약 단순한 부재가 원인이라면, 우리가 받아들일 준비가 되어있는 것보다 훨씬 더 많은 원인이 있게 될 것이다. 예컨대 철수는 영희의 화초에 물을 주기로 약속했지만 그러지 않았고 그 결과 화초가 고사했다고 해보자. 실제로 아무도 화초에 물을 주지 않은 경우, 만약 철수가 물을 주었다라면 화초는 살았을 것이다. 그러나 만약 영국 여왕이 물을 주었다라면 화초는 살았을 것이라는 주장 역시 성립한다.¹³⁾ 하지만 영국 여왕이 화초에 물을 주지 않았다는 것은 화초가 고사한 것의 원인이 아니며, 이 점은 단순한 부재는 어떤 사건의 원인이 아니라는 것을 시사한다.¹⁴⁾ 단순한 정신적 부재로부터 ‘흘러나오는’ 것이 무엇이든 그것과 단순 부재 간에는 인과 관계가 성립하지 않는다. 또한 행위자는 단순한 정신적 부재와 그것의 ‘결과’에 대해 도덕적 책임이 없다는 것 역시 분명해 보인다.

12) Sartorio 2009, 520쪽.

13) Beebe (2004). 반면 적어도 어떤 부재는 원인이라는 주장에 대한 옹호로는 Schaffer (2004)를 참조하라.

14) 이른바 ‘영국 여왕의 문제’를 해결하려는 시도로는 McGrath (2005)를 참조하라.

물론 위의 논의는 부재와 부재 간의 인과 관계를 직접적으로 논의하지도 않고 부재 인과 일반에 대한 어떤 증명을 제공하지도 않는다. 하지만 부재 인과 자체를 부인하지 않는다면, 적어도 단순한 부재와 원인이 될 수 있는 부재 간의 구분을 받아들일 필요가 있다는 점은 보여준다. 이제 앞서 살펴본 예를 조금 변경하여, 철수가 영희의 화분에 물을 주지 않은 것과 화초가 세포 내에 수분을 유지하지 못하는 것 간의 관계에 대해 생각해보자. 철수가 화분에 물을 주지 않은 것은 앞서 말한 것처럼 단순한 부재가 아니다. 그리고 철수가 화분에 물을 주지 않은 것과 화초가 세포 내에 수분을 유지하지 못하는 것 간에는 의존 관계가 성립한다. 즉 만약 철수가 화분에 물을 주었다면 화초는 세포 내에 수분을 유지할 수 있었을 것이다. 앞서 언급했듯 단순한 부재가 보여주는 의존 관계는 인과에 관한 좋은 근거가 될 수 없지만, 이 경우에 성립하는 의존 관계는, 부재 인과 자체를 부인하지 않는 한, 인과 관계에 관한 좋은 근거가 된다. 일반적으로 원인은 어떤 사건을 발생시키거나 혹은 어떤 사건의 발생을 방지함으로써 원인이 된다. 그리고 모든 부재가 그렇지는 않지만 적어도 어떤 부재는 다른 것의 원인이 될 수 있다면 그런 부재 역시 어떤 사건을 발생시키거나 혹은 어떤 사건의 발생을 방지함으로써 원인이 된다는 것을 부정할 이유를 찾기 힘들다. 이러한 고려가 사토리오의 주장에 대해 제시하는 바는 명료하다. 즉 사토리오가 언급하고 있는 방안의 코끼리의 부재는 단순한 부재에 불과하며 이에 근거한 논의는 부재가 다른 부재의 원인이 될 수 없음을 확립할 수 없다.

이상을 종합해보면, (F-행위)와 (F-부재) 간의 인과적 비대칭성을 통해 도덕적 비대칭성을 옹호하고자 하는 사토리오의 논변을 받아들일 좋은 이유가 없다. 먼저 ②와 같은 스위치 구조에 대한 주장이 옳다고 해서 (F-부재)의 경우인 ①과 같은 스위치 구조에 대한 주장이 자동적으로 성립하지 않는다. 그리고 ①과 같은 스위치 구조가 ②와 같은 인과적 성격을 갖는 것을 보이기 위한 시도 역시 성립하지 않는다. 이 시도의 첫 번째 단계를 확립하고자 하는 함축 관계에 근거한 인과적 주장이 성립하지 않고, 설령 그 주장이 성립한다고 해도 부재와 부재 간에는 인과적 관계가 성립할 수 없다는 주장에 문제가 있다. 물론 (F-부재)에서 아이의 죽음은 피할 수 없는

일이다. 그리고 이 점이 부재의 인과적 지위에 대해 함축하는 바가 있을 수도 있다. 하지만 사토리오의 주장을 통해서도 이를 보여주지 못한다. 그렇지만 사토리오의 주장을 반박한다고 해서 모든 문제가 해결되는 것은 아니다. (F-행위)와 (F-부재) 간에 인과적 대칭성이 성립한다고 주장하기 위해서는 적어도 어떤 근거에서 그들 사례 간에 대칭성이 성립한다고 볼 수 있는지에 대해 논의할 필요가 있다. 다음 절에서는 이 문제를 조금 더 살펴보고자 한다.

III. ‘실제 인과’와 (F-행위)와 (F-부재) 간의 인과적 대칭성

(F-행위)는 피할 수 없는 일에 대해서도 행위자가 자신의 행위에 대해 책임이 있을 수 있음을 보여주는데, 이를 인과적 측면에서 지지하기 위해서는 행위자의 행위가 피할 수 없는 어떤 것의 원인이 될 수 있음을 보일 필요가 있다. 그리고 인과 논제를 받아들이는 입장에서 (F-행위)와 (F-부재) 간의 도덕적 대칭성을 인과적으로 지지하기 위해서는 (F-부재) 역시 마찬가지로 원인이 될 수 있음을 보일 필요가 있다. 그런데 인과 개념 자체가 논쟁의 대상이라는 점을 고려해볼 때, (F-행위)와 (F-부재)의 인과적 성격을 규명하는 것과 관련하여 인과 관계를 어떻게 이해해야 하는가 하는 문제가 대두된다.

일반적으로 인과 관계를 이해하는 입장은 크게 두 가지로 나눌 수 있다. 즉, 인과 관계를 일종의 물리적 ‘연결’로 보는 입장과 일종의 ‘의존’ 관계로 보는 입장이 있다.¹⁵⁾ 그런데 행위의 부재나 누락과 같은 것은 발생하지 않은 것이며, 발생하지 않은 것은 어떤 다른 사건과 물리적으로 연결될 수 없다. 따라서 인과 관계를 일종의 물리적 연결로 보는 입장을 받아들이면 부재는 다른 어떤 것의 원인이 될 수 없다. 하지만 이 글의 서론에서 언급했듯이 논문이 다루고 있는 논쟁은 부재가 원인이 될 수 있다는 것을 받아들이

15) Hall (2004).

는 철학자들 간의 논쟁이며 특히 사토리오 역시 부재가 원인이 될 수 있다는 것 자체를 부인하지 않으므로, 인과 관계를 원인과 결과 간에 성립하는 일종의 의존 관계로 보고 논의를 진행시켜야 한다.

인과 관계를 일종의 의존 관계로 보는 대표적인 이론은 반사실적 인과 이론이다. 이 이론의 핵심은 원인과 결과 간에 반사실적 의존 관계가 성립한다는 것이다. 즉, 만약 원인이 발생하지 않았더라면 결과 역시 발생하지 않았을 것이다. 하지만 인과 관계를 단순한 반사실적 의존 관계로만 보는 것에 문제가 있음을 이미 잘 알려진 사실이다.¹⁶⁾ 예를 들어, 철수와 영희 둘 다 유리창을 향해 돌을 던진다고 해보자. 그리고 영희가 철수보다 먼저 혹은 좀 더 세계 돌을 던져 영희가 던진 돌이 유리창과 접촉하여 유리창이 깨지고, 철수가 던진 돌은 이미 유리창이 깨지고 없는 빈 공간을 통과한다고 해보자. 이 경우 유리창이 깨진 것의 원인은 영희가 돌을 던진 것이다. 하지만 원인과 결과 간에서는 반사실적 의존 관계가 성립하지 않는다. 즉, 만약 영희가 돌을 던지지 않았더라도 (철수가 던진 돌로 인해) 유리창은 여전히 깨졌을 것이다.

그렇다면 인과 관계를 의존 관계로 보면서도 이러한 문제에 당면하지 않는 방식으로 인과 관계를 파악할 필요가 있다. 이와 관련하여 필자는 ‘실제 인과’(actual causation)라고 알려진 인과에 관한 최근의 이론(Halpern and Pearl 2001, Hitchcock 2001, 2007)을 염두에 두고 도덕적 책임의 문제와 관련된 인과의 구조에 대한 논의를 진행하고자 한다. 여기서 ‘실제 인과’라는 것은, 기본적으로, 실제 상황에 대한 고려를 어떤 식으로든 반영하면서 원인과 결과 간의 반사실적 의존 관계를 살펴봄으로써 인과 관계를 이해하고자 하는 접근법을 뜻한다. 이 접근법의 핵심은 실제 상황에서 원인과 결과 간의 의존 관계를 훼손하는 것이 있다면 그것을 일정한 방식으로 제어함으로써 원인과 결과 간의 의존 관계를 회복시키고자 하는 데 있다. 예를 들어, 영희와 철수가 돌을 던진 위의 경우에서 철수가 돌을 던진 것은 실제 원인인 영희가 돌을 던진 것과 그 결과인 유리창이 깨진 것 간의 의존 관계를 훼손하는 요인이다. 그런데 이 경우 철수가 던진 돌은 실제로 유리창과 접

16) Lewis (1973, 1986, 2000).

촉하지 않았다. 그리고 이러한 사실 즉 철수가 던진 돌이 유리창과 접촉하지 않았다는 사실을 ‘고정한’ 상태에서 영희가 돌을 던지지 않았더라면 어떤 일이 생겼을지 생각해보면 유리창은 깨지지 않았을 것이라는 것을 쉽게 알 수 있다. 달리 말하면, 반사실적 인과론의 기본적 아이디어는 의존 관계를 살펴보기 위해 원인이 ‘제거된’ 반사실적 상황을 상정하고 그러한 상황에서 결과 역시 ‘제거’될 것인지를 살펴보는 것인데, 실제 인과는 이를 받아들이면서도 이에 더해 특정 사실을 ‘고정’할 것을 요구한다. 여기서 특정 사실을 ‘고정’한다는 것은 단순히 원인이 ‘제거’된 반사실적 상황 대신 원인이 ‘제거’되고 *그리고*, 예를 들어, 철수가 던진 돌이 (실제로 그랬던 것처럼) 유리창과 접촉하지 않은 그런 반사실적 상황에서, 결과 역시 제거될 것인지를 살펴보는 것을 뜻한다.

이렇듯 ‘실제 인과’는 실제 상황을 일정한 방식으로 ‘고정’한 상태에서 원인과 결과 간의 반사실적 의존 관계를 파악하고자 한다. 이를 조금 다른 방식으로 부연 설명해보자. 앞서의 사례에는 유리창이 깨지는 것과 관련된 두 가지 가능한 ‘경로’가 있다. 한 가지는 ‘영희가 돌을 던짐-영희의 돌이 날아감-영희의 돌이 유리창과 접촉함-유리창이 깨짐’의 경로이고 또 다른 것은 ‘철수가 돌을 던짐-철수의 돌이 날아감-철수의 돌이 유리창과 접촉함-유리창이 깨짐’이다. 그런데 실제로 발생한 것은 영희와 관련된 경로의 모든 것이고 철수와 관련된 경로에서는 ‘철수가 돌을 던짐-철수의 돌이 날아감’은 실제로 발생하였으나 ‘철수의 돌이 유리창과 접촉함’은 실제로 발생하지 않았다. 이제 영희가 돌을 던진 것과 유리창이 깨진 것 간의 인과 관계를 밝히기 위해, 원인-후보와 결과-후보를 제외하고 영희와 관련된 경로에 속하지 않는 것들을 실제와 동일한 상태로 고정한다고 해보자. 즉 ‘철수가 돌을 던짐-철수의 돌이 날아감’을 실제 상황대로 ‘발생’으로 고정하고 ‘철수의 돌이 유리창과 접촉함’ 역시 실제 상황대로 ‘발생하지 않음’으로 고정한다고 해보자. 이런 상태에서 영희가 돌을 던진 것과 유리창이 깨진 것 간의 관계를 살펴보면 이 둘 사이에 반사실적 의존 관계가 성립함을 알 수 있다. 즉 이렇게 고정된 상태에서, 만약 영희가 돌을 던지지 않았더라면 유리창은 깨지지 않았을 것이다. 반면 철수가 돌을 던진 것과 유리창이 깨진 것 간의 의존 관

계를 살펴보기 위해서는 영희가 던진 돌이 유리창과 접촉하였다는 실제 상황을 고정하여야 한다. 그리고 이렇게 하면, 만약 철수가 돌을 던지지 않았더라도 유리창은 깨졌을 것이라는 반사실적 조건문이 성립하므로 철수가 돌을 던진 것은 유리창이 깨진 것의 원인이 아니게 되며, 이는 정확한 답이다.

그런데 실제 상황에 대한 고려가 도덕적 책임과 관련된 논의에서 중요하다는 점은 이미 프랑크푸르트의 글에서도 찾아볼 수 있다.

어떤 일을 하는 것을 피할 수 없었다는 사실은 그 일을 한 것에 대한 충분조건이다. 그러나 나의 예들이 보여주듯, 이 사실은 행위자가 그 일을 왜 했는지에 대한 설명에서 어떤 역할도 하지 않을 수 있다. 그 사실은 행위자가 자신이 한 일을 실제로 하게끔 한 상황의 일부가 전혀 아닌 것으로 밝혀질 수 있으며 그 경우 그의 행위는 완전히 다른 근거에 기대어 설명되어야 한다.¹⁷⁾

여기서 ‘행위자가 자신이 한 일을 실제로 하게끔 하는 것을 행위의 원인으로 이해하는 것이 자연스러워 보이는데, (F-행위)와 같은 상황과 관련된 인과 구조를 해명하는 방식으로 프랑크푸르트는 다음을 제안한다.

[철수]가 달리 하는 것이 불가능하게 만드는 그 상황을, 무엇이 발생했는지 혹은 어째서 발생했는지에 영향을 주지 않으면서, 주어진 상황에서 빼버릴(subtract) 수 있다. 그 사람이 자신이 한 일을 하게끔 이끈 것이 혹은 그렇게 만든 것이 실제로 무엇이든 그것은, 설령 그가 그 일 외의 무엇인가를 하는 것이 가능했었다고 하더라도, 그로 하여금 그 일을 하도록 이끌거나 혹은 그 일을 하도록 만들었을 것이다.¹⁸⁾

하지만 단순히 ‘주어진 상황에서 빼버릴(subtract) 수’ 있는지 여부에 의해서는 인과 관계를 정확하게 가려낼 수 없다. 철수가 돌을 던진 것과 관련된 경로를 빼버리고 단지 영희가 돌을 던진 것과 유리창이 깨진 것의 관계를 보면 물론 반사실적 의존 관계가 성립한다. 그러나 영희가 돌을 던진 것과 관련된 경로를 빼버려도 철수가 돌을 던진 것과 유리창이 깨진 것 간에 반사실적 의존 관계가 성립한다. 물론 위 주장에는 ‘무엇이 발생했는지 혹

17) Frankfurt 1969, 836-37쪽, 강조 첨가.

18) Frankfurt 1969, 837쪽.

은 어째서 발생했는지에 영향을 주지 않으면서'라는 단서가 있긴 하지만, 이 단서를 어떻게 이해해야 할지에 대해서는 명확한 해명을 찾을 수 없다. 반면 '실제 인과'는 관련된 아이디어를 체계적으로 생각할 수 있는 방법을 제공한다. 그리고 만약 도덕적 책임의 문제와는 독립적으로 제시된 인과 이론인 실제 인과가 이 문제에 대한 이해를 도울 수 있다면, 이는 (F-행위)와 (F-부재)의 인과적 차이에 근거해 도덕적 비대칭성 논제를 지지하고자 하는 시도를 도덕적 책임에 관한 논의와 독립적으로 평가할 수 있게 해 준다.

이제 실제 인과의 틀을 사용하여 (F-행위)의 경우에 대해 생각해보자. 먼저 희생자의 죽음과 관련하여 두 가지 경로가 있다. 즉 희생자를 쏘아야 할 이유를 긍정적으로 고려하는 철수의 결정-이전 상태에서부터 철수의 결심과 행동을 거쳐 희생자의 죽음에 이르는 경로(이를 '철수-경로'라고 하자)와 희생자를 쏘아야 할 이유에 대해 부정적으로 고려하는 결정-이전 상태에서부터 과학자의 개입, 철수의 결심, 행동을 거쳐 희생자의 죽음에 이르는 경로(이를 '과학자-경로'라고 하자)가 있다. 이제 철수의 결정과 희생자의 죽음 간에 의존 관계가 있는지 살펴보기 위해 철수-경로에 속하지 않는 요인인 과학자의 개입 여부를 실제와 동일한 상태인 '개입하지 않음'으로 고정해보자. 이 경우, 만약 철수가 희생자를 쏘고자 결정하지 않았더라면 철수는 방아쇠를 당기지 않았을 것이고 희생자는 죽지 않았을 것이다. 따라서 철수의 결정은 희생자 죽음의 원인이며, 철수의 행동 역시 희생자 죽음의 원인이 된다. 마찬가지로 (F-부재)의 경우 역시 과학자가 개입하지 않음을 고정하면, 만약 철수가 물에 뛰어들지 않을 결정을 하지 않았더라면 그는 물에 뛰어들었을 것이며 아이는 살아남았을 것이므로, 철수의 결심과 그의 행위 부재는 아이 죽음의 원인이다.

이미 언급했듯이 (F-행위)와 같은 사례들은, 도덕적 책임은 행위자가 달리 할 수 있었을 것을 요구한다는 대안 가능성(alternative possibility) 이론을 비판하기 위해 고안된 것이다. (F-행위)에서 철수의 결정, 행동 및 행동의 결과는 과학자의 존재로 인해 피할 수 없는 것이며, 여기서 과학자의 존재는, 프랑크푸르트의 표현을 빌리면, 이런 것들을 피할 수 없게 만드는 '충분조건'이다. 하지만 위의 논의는 그것을 피할 수 없는 충분조건이 성립한

경우에도 행위자는 여전히 책임이 있을 수 있음을 보여준다. 이것이 함축하는 바를 좀 더 살펴보기 위해 다음의 경우를 상상해보자. 철수의 결정-이전 상태가 방아쇠를 당기지 않는 쪽으로 기울고 있고 이를 알아챈 과학자가 개입하여 철수로 하여금 방아쇠를 당기는 결정을 하도록 만든다. 철수는 방아쇠를 당기고 희생자는 죽게 된다. 이 경우는 철수-경로와 과학자-경로 중 과학자-경로가 실제 경로인 경우이다. 과학자의 개입이 그 다음 발생한 일들의 원인인지 살펴보기 위해 ‘철수의 결정-이전 상태’를 실제와 동일하게 ‘방아쇠를 당기지 않는 쪽으로 기울고 있었음’으로 고정하면¹⁹⁾ 다음이 성립함을 확인할 수 있다. 즉, 만약 과학자가 개입하지 않았더라면 철수의 결정도 행동도 발생하지 않았을 것이고 아이는 죽지 않았을 것이다. 따라서 과학자의 개입은 철수의 결정, 행동 및 아이 죽음의 원인이다. 반면 철수의 결정-이전 상태가 방아쇠를 당기는 쪽으로 기울었더라면(과학자는 개입하지 않았을 것이지만) 철수의 결정과 행동은 여전히 발생하였을 것이고 아이는 죽었을 것이다. 따라서 이 경우 철수의 결정-이전 상태는(과학자 개입 여부를 결정한다는 점에서 과학자 개입의 원인이지만) 철수의 결정과 행동 및 아이 죽음의 원인이 아니다.

위의 경우는, 철수 결정-이전의 상태가 과학자 개입의 원인이고 과학자 개입이 그 다음에 발생하는 일련의 사건들의 원인이라고 해도, 결정-이전 상태는 이러한 일련의 사건들의 원인이 아님을 보여준다. 달리 말하면 ‘실제 인과’에 따르면 인과의 전이성(transitivity)이 성립하지 않는다. 그리고 철수의 결정과 행동 그리고 아이의 죽음은 철수의 결정-이전 상태가 방아쇠를 당기지 않는 쪽으로 기울고 있었음에도 불구하고 과학자의 개입 때문에 이런 일들이 발생했다는 주장이 직관적으로 설득력이 있다는 점을 고려할 때, 인과의 전이성을 부인하는 것에 설득력이 있다. 더 나아가 인과의 전이성이 성립하지 않는다는 것은 도덕적 책임의 문제에서 인과에 대한 고려

19) 이 경우 철수의 결정-이전 상태를 실제와 동일하게 고정하는 것은 실제 인과 이론에만 국한되는 요구 사항이 아니다. 반사실적 의존 관계를 통해 인과 관계를 규명할 때 원인-후보 발생 이전의 것들을 실제와 동일하게 고정하는 것은 반사실적 인과 이론에서 ‘표준’으로 받아드려진다. 원인-후보 발생 이전의 것들도 변화시키는 이른바 후-추적(backtracking)의 문제점과 관련하여서는 Lewis (1973)를 참조하라.

가 하는 역할에 대해서도 의미하는 바가 있다. 만일 인과가 전이적이어야만 한다면 일련의 사건들이 발생할 수밖에 없는 상황에서 철수의 결정-이전의 상태, 결정 및 행동은 철수-경로를 통해 발생하든 아니면 과학자-경로를 통해 발생하든 상관없이 아이의 죽음에 대해 동일한 성격의 인과적 기여를 한다. 반면 인과의 전이성이 성립하지 않으면, 아이의 죽음이 어떤 경로를 따라 발생했는지에 따라 인과적 기여의 성격이 달라진다. 철수-경로를 따라 발생한 아이 죽음의 원인은 철수 자신에게 의해 즉 그의 결정-이전 상태에서부터 촉발된 일련의 인과 고리를 따라 추적할 수 있지만, 과학자-경로를 따라 발생한 아이 죽음의 원인은 과학자가의 개입이 일련의 인과 고리를 촉발한 것으로 밝혀진다. 따라서 아이 죽음이 설령 피할 수 없는 것이었다고 하더라도, 철수-경로를 통해 발생한 경우에는 철수에게 그 도덕적 책임이 있을 수 있는 반면, 과학자-경로를 통해 발생한 경우에는 그 책임이 철수가 아닌 과학자에게 있을 수 있음을 보여준다.

마지막으로 스위치를 구성하는 ②의 경우인 (기차)에서 스위치를 조작하는 것이 원인이 아니라는 것을 실제 인과 이론을 통해 보여줄 수 있는지에 대해 생각해보기로 하자. 홀(Hall 2007)은 (기차)와 같은 경우는 실제 인과 이론에 문제가 되는 경우라고 주장한다. 실제 상황에서 “오른쪽 선로는 *실제로*(in fact) 비어있는데, 만약 스위치를 조작하지 않는 반사실적 상황에서 이 선로가 실제와 마찬가지로 남아있었다라면 기차는 그것의 종착지에 도달하지 않는다.”²⁰⁾ 만약 이 주장이 옳다면 실제 인과는 (기차)에 대해 스위치의 조작이 희생자 죽음의 원인이라는 잘못된 답을 내놓음으로써 그 이론 자체에 문제가 있는 것이 된다. 또한 (기차)와 (F-행위) 및 (F-부재)를 통합적으로 다루지 못함으로써 사토리오의 주장을 반박했던 우리의 논의를 뒷받침해 줄 만한 인과 이론이 부재할 수도 있다는 문제가 생기게 된다.

홀이 제기한 문제는 기차가 오른쪽 선로를 달리는 것과 왼쪽 선로를 달리는 것을 인과적 ‘요인’으로 볼 것인지 아니면 인과적 요인이 갖는 다양한 ‘값 혹은 상태’로 볼 것인지를 문제와 연관되어 있다. 여기서 어느 선로를 따라 달리는지가 인과적 요인이라는 것은 그것을 예를 들어 철수의 결정,

20) Hall 2007, 119쪽, 강조는 원문을 따름.

행동 혹은 과학자의 개입과 같은 것으로 본다는 것이다. 이 경우 ‘오른쪽 선로를 따라 달림’의 인과적 요인은 그 선로에 ‘기차가 있음’이나 ‘기차가 없음’의 값을 가질 수 있다. 반면 어느 선로를 따라 달리는지가 인과적 요인이 아니라 인과적 요인의 값 혹은 상태라는 것은 예를 들어 ‘과학자의 개입’의 요인이 ‘개입함’ ‘개입하지 않음’의 상태를 갖는 것처럼 ‘기차가 선로를 따라 달림’이라는 인과적 요인은 ‘오른쪽 선로를 따라 달림’이나 ‘왼쪽 선로를 따라 달림’의 값이나 상태를 갖는다는 것을 뜻한다.

이 점을 고려할 때 홀의 주장은 각각의 선로를 따라 달리는 것이 서로 다른 인과적 요인이라는 것과 또한 실제 인과 이론은 이들을 서로 다른 인과적 요인으로 보아야만 한다는 것을 전제하고 있다. 이 전제를 수용하게 되면, 선로 스위치를 조작한 것이 기차가 희생자를 친 것의 원인인지를 밝히기 위해서는 실제 경로에 속하지 않은 다른 경로에 속한 것들을 그것이 실제로 발생한 상태로 고정된 후에 원인-후보와 결과-후보 간에 반사실적 의존 관계가 성립하는지 살펴봐야 한다. 그런데 기차는 실제로 왼쪽 경로를 따라 달리고 있었으므로 다른 경로에 속하는 ‘오른쪽 경로를 따라 달림’이라는 인과적 요인의 값을 ‘기차가 없음’ 혹은 ‘비어 있음’으로 고정해야 한다. 그리고 이를 고정하게 되면, 만약 철수가 스위치를 조작하지 않았더라면 기차는 희생자는 치지 않았을 것이라는 주장이 성립하는 것 같아 보인다. 그 이유를 이해하기 위해 철수가 스위치를 조작하지 않는 반사실적 상황을 생각해보자. 그 상황에서는 기차는 오던 대로 계속해서 오른쪽 선로를 따라 달릴 것이라고 생각하는 것이 자연스럽지만, 이미 오른쪽 선로가 비어있었다는 사실을 고정된 상태에서 철수가 스위치를 조작하지 않는 반사실적 상황에서는 오른쪽 선로를 따라 달리던 기차는 스위치가 조작되는 교차점을 지나자마자 돌연 선로에서 사라져버리게 되고 그 결과 희생자는 죽음을 면하게 된다.

한편 기차가 오른쪽 혹은 왼쪽 선로를 따라 달리는 것을 기차가 선로를 따라 달린다는 인과적 요인의 값으로 보게 되면 홀이 제기한 문제가 발생하지 않는다. 기차가 선로를 따라 달리는 것은 기차가 선로를 따라 달리고 있음-스위치를 조작함-희생자의 죽음 경로에 속하는 한 요인이며 별도의 다른 인과 경로가 없으므로 달리 그 값을 고정해야 할 요인이 없다. 그리고 만약

스위치를 조작하지 않았더라면 기차는 기존에 달리고 있던 선로를 따라 즉 오른쪽 선로를 따라 달렸을 것이고 희생자는 여전히 죽게 되었을 것이므로 스위치 조작은 희생자 죽음의 원인이 아니다. 즉, 오른쪽 혹은 왼쪽 선로를 따라 달리는 것을 인과적 요인의 값이나 상태로 봄으로써 실제 인과 이론은 정확한 답을 내놓는다. 앞서 논의한 실제 인과에 대한 논의를 점검해보면 실제 인과는 (기차)와 같은 경우를 반드시 홀이 생각하는 방식으로 다루는 것에 동의해야(committed) 한다고 볼 근거가 없다. 또한 홀이 제시하는 방식은 오른쪽 선로를 따라 달리던 기차가 교차점을 지나자마자 돌연 선로에서 사라져버리게 되는 실제와 많이 동떨어진 상황을 생각할 것을 요구하는데, 이 점은 홀의 논점을 지지하기 보다는 그의 제안에 문제가 있음을 보여주는 것 같다. 그렇다면 (기차)와 같은 상황에서 각각의 선로를 달리는 것을 인과적 요인이 아닌 주어진 인과적 요인이 갖는 값이나 상태로 보는 것에 문제가 없을 뿐만 아니라 오히려 그렇게 볼 좋은 이유가 있다고 주장할 수 있다.

홀의 주장은 만약 두 선로의 물리적 상이함이 인과적 요인의 상이함을 함축한다면 여전히 유효한 주장일 수 있다. 하지만 물리적 상이성은 인과적 상이성을 함축하지 않는다. 예컨대 양손잡이 철수가 자신의 오른손으로 공을 던져 유리창을 깬다고 해보자. 오른손으로 던진 공의 물리적 궤적과 왼손으로 던진 공의 물리적 궤적은 서로 상이하지만, 유리창이 깨진 것과 관련해 인과적으로 중요한 것은 유리창을 향해 돌을 던졌다는 것이지 그 공을 오른손으로 던졌는지 아니면 왼손으로 던졌는지가 아니다. 철수는 왼손으로 공을 던질 수도 있었지만 유리창이 깨진 것에 대해 이 점이 특별히 인과적으로 중요하지 않다. 그렇다면 (기차)에 실제 인과 이론을 적용하는 데 있어 기차가 오른쪽 선로를 따라 달리는지와 왼쪽 선로를 따라 달리는지와 관련하여, 그 둘이 물리적으로 상이하다고 해서 인과적으로 상이한 인과 경로로 다루어야만 할 필요가 없다.

이상의 논의는 실제 인과를 사용하여 (F-행위)와 (F-부재)에서 행위자의 행위나 행위의 부재가 주어진 결과의 원인일 수 있음을 보일 수 있을 뿐만 아니라 스위치를 구성하는 한 구조인 ②의 경우에도 직관과 일치하는 답을 내놓을 있다는 것을 보여준다. 그렇다면 스위치의 인과 구조에 기대어 (F-

행위)와 (F-부재) 간의 인과적 비대칭성을 옹호하는 사토리오의 논변에 문제가 있을 뿐만 아니라, 이에 더해 실제 인과 이론을 통해 사토리오 논변의 핵심이 되는 스위치의 인과적 구조를 수용하면서도 (F-행위)와 (F-부재) 간의 인과적 대칭성을 지지할 수 있게 된다.

IV. 맺음말

인과 관계에 관한 고려가 도덕적 책임의 논의에서 중요한 역할을 함에도 불구하고 도덕적 책임에 관한 논의에서 인과 관계 자체에 관한 논의는 상대적으로 소홀히 다루어져 온 경향이 있다. 사실 도덕적 책임에 대한 논의만큼이나 인과에 대한 논의가 복잡하고 논쟁의 여지가 많다는 점을 생각해보면 이는 충분히 이해할 수 있는 일이기도 하다. 이런 점에 비추어볼 때 사토리오의 논문은 도덕적 책임의 문제와 관련하여 명시적으로 인과의 문제를 논의하고 그를 통해 논쟁의 대상이 되어온 도덕적 책임의 비대칭성을 지지하고자 했다는 점에서 중요한 기여를 한다. 이 글에서 전개한 필자의 주장이 옳다면 (F-행위)와 (F-부재) 사례를 통한 사토리오의 논변에는 많은 문제점이 있고 실제 인과에 대한 고려 역시 이 사례의 인과적 대칭성을 지지한다. 하지만 이로부터 도덕적 비대칭성 논제 자체가 성립하지 않는다는 결론이 도출되는 것은 물론 아니다. 따라서 도덕적 비대칭성 논제를 지지할 수 있는 또 다른 인과적 사례가 있는지 그렇다면 그것이 도덕적 책임의 성격에 관해 의미하는 바가 무엇인지는 여전히 주목할 필요가 있는 문제이다.

투 고 일: 2014. 10. 17.
 심사완료일: 2014. 11. 16.
 게재확정일: 2014. 11. 17.

김성수
 서울시립대학교 철학과 부교수

참고문헌

- Beebe, Helen (2004), 'Causing and Nothingness', in L. A. Paul, E. J. Hall & J. Collins (eds.), *Causation and Counterfactuals*: 291-308, The MIT Press.
- Clarke, Randolph (2011), 'Omissions, Responsibility, and Symmetry', *Philosophy and Phenomenological Research* 82: 594-624.
- Fischer, John and Ravizza, Mark (1991), 'Responsibility and Inevitability', *Ethics* 101: 258-78.
- Fischer, John and Ravizza, Mark (1998), *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge University Press.
- Frankfurt, Harry (1969), 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy* 66: 829-39.
- Frankfurt, Harry (1994), 'An Alleged Asymmetry between Actions and Omissions', *Ethics* 104: 620-23.
- Hall, Ned (2004), 'Two Concepts of Causation', in J. Collins, N. Hall & L. A. Paul (Eds.), *Causation and Counterfactuals*: 225-76, The Mit Press.
- Hall, Ned (2007), 'Structural Equations and Causation', *Philosophical Studies* 132: 109-36.
- Halpern, J. and Pearl, J. (2001), 'Causes and Explanations: A Structural-Model Approach—Part 1: Causes', in *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*: 194-202.
- Hitchcock, Christopher (2001), 'The Intransitivity of Causation Revealed in Equations and Graphs', *Journal of Philosophy* 98: 273-99.
- Hitchcock, Christopher (2007), 'Prevention, Preemption, and the Principle of Sufficient Reason', *Philosophical Review* 116: 495-532.
- Hitchcock, Christopher (2009), 'Structural Equations and Causation: Six Counterexamples', *Philosophical Studies* 144:391-401.

- Lewis, David (1973), 'Causation', *Journal of Philosophy* 70: 556-67.
- Lewis, David (1986), 'Postscripts to 'Causation'', in D. Lewis, *Philosophical Papers*, Vol. 2: 172-213, Oxford University Press.
- Lewis, David (2000), 'Causation as Influence', *Journal of Philosophy* 97:182-97.
- McGrath, Sarah (2005), 'Causation by Omission: A Dilemma', *Philosophical Studies* 123: 125-48.
- Rowe, William (1989), 'Causing and Being Responsible for What Is Inevitable', *American Philosophical Quarterly* 26: 153-59.
- Sartorio, Carolina (2005), 'A New Asymmetry Between Actions and Omissions', *Noûs* 39: 460-82.
- Sartorio, Carolina (2009), 'Omissions and Causalism', *Noûs* 43: 513-30.
- Schaffer, Jonathan (2004), 'Causes Need Not Be Physically Connected to Their Effects: The Case for Negative Causation', in C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science*: 197-216, Blackwell.

ABSTRACT

Causation and Moral Responsibility:

Moral Asymmetry and Its Problems

Kim, Sungsu

It has been claimed that when the outcome of an action or an omission would not have been different, an agent can be morally responsible for an action but an agent cannot be responsible for an omission. In order to defend this moral asymmetry claim, philosophers often appeal to some sort of causal asymmetry between actions and omissions. In this paper, I critically examine this asymmetry claim. More specifically, I examine Sartorio's argument for the moral asymmetry thesis that appeals to certain causal characteristics of "switches." I argue that the argument fails. I then analyze in terms of "actual causation" the causal scenarios that play a crucial role for the argument. It turns out that these scenarios do not support the argument either.

Keywords: Causation, moral responsibility, action, omission