

선의지는 개별 의도인가 성격인가 - 『도덕형이상학 정초』를 중심으로

강은아*

I. 서론

칸트는 『도덕형이상학 정초』 1절 서두에서 “이 세계에서 또는 도대체가 이 세계 밖에서까지라도 아무런 제한 없이 선하다고 생각될 수 있을 것은 오로지 선의지뿐”(IV393)이라고 말한다.¹⁾²⁾ 선의지에 관한 저 언명은, 1절의 제목인 “평범한 윤리적 이성 인식에서 철학적 이성 인식으로의 이행”이 알려주듯이, 정치한 철학적 규정이 아닌 건전한 상식을 갖는 보통 인간의 의견으로서 제시된다. 『정초』 1절의 선의지 논의는 이처럼 도덕의 본질을 보통 사람이

* 서울과학기술대학교 교양대학 시간강사

1) 임마누엘 칸트, 『윤리형이상학 정초』는 Felix Meiner판(Kant, 2016)을 기본으로 백중현의 번역(백중원, 2014)과 김석수, 김중국의 번역(김석수, 김중국, 2019)을 참조하여 필자가 번역하였다. 그리고 칸트 저작들의 저서 번호 및 페이지 표시는 독일 베를린 학술원판 칸트 전집(Kant's gesammelte Schriften)을 따른다.

2) 『윤리형이상학 정초』는 『정초』로 줄여서 표기한다.

견지하는 도덕 및 선에 관한 입장으로부터 찾아내고자 하는 논의 방향에서, 보통 사람이 갖는 견해의 일환으로서 제시되는 것이다.

그렇다면 선의지는 무엇인가? “평범한 윤리적 이성 인식”에서 볼 때, 선의지는 행위 시에 행위자가 갖는 선한 의도를 가리킨다고 하는 것이 가장 자연스러워 보인다. 개별 행위를 어떠한 마음으로 했는지, 선의에 따른 것인지 아닌지에 따라 그 행위의 평가가 달라진다는 것이다. 이러한 생각은 자연스럽게 선의지는 행위자가 개별 행위 시에 갖는 개별 의도라는 해석으로 이끌리기 쉽다. 선의로 한 행위는 무제한적으로 선하지만, 선의 없이 한 행위는 그렇지 않다는 것이다. 하지만 선의지(*guter Wille*)가 의지의 일종임을 고려하면, 선의지는 개별 행위보다는 행위자에 속하는 어느 정도 지속적인 특성으로 보는 것이 더 자연스러운 것일 수 있다. 의지는, 어떤 맥락에서는, 한 행위자의 여러 행위들에서 공통적으로 관철되는 것, 행위자의 일관된 선택의 배후에 있는 것, 어떤 성향으로 이해된다. 이에 따르면 선한 의지는 행위자가 갖는 특정한 종류의 성격이라고 할 수 있을 것이다.

선의지가 개별 행위의 속성이냐 아니면 행위자의 지속적인 성향이냐 하는 물음에 대해, 최근의 논의들은 후자를 지지하는 경향을 보인다.³⁾ 나는 이 문제를 드물게도 주제적으로 다룬 아메릭스의 논의를 통해 이 문제에 접근해 보고자 한다.⁴⁾ 나는 선의지를 성향이나 성격으로 이해해야 한다는 아메릭스의 입장에 동의해서 선의지는 행위자가 견지하는 성격이라고 주장할 것이다. 그렇지만 나는 개별 의도 해석을 비판하는 아메릭스의 논증이 상당한 결함을 가지며, 특히 개별 의도 해석을 배제하려면 아메릭스의 논의를 수정할 필요가

3) 성격 해석을 지지하는 대표적인 학자로는 앨리슨(Allison, 2011), 존슨(Johnson, 2009) 등이 있다.

4) 이 논의를 주제적으로 다룬 논문으로는 칼 아메릭스의 “Kant on the Good Will”이 있다 (Ameriks, 1989). 여기서 아메릭스는 선의지를 개별 의도로 해석하는 입장, 능력으로 해석하는 입장, 성격으로 해석하는 입장으로 구분하고, 세 번째인 성격 해석을 지지한다. 이 글에서 나는 두 번째 입장은 다루지 않고 개별 의도 해석과 성격 해석만을 다룰 것인데, 선의지를 자유롭게 선택하는 능력으로 보는 두 번째 입장은 선의지를 갖는 의지의 범위를 다루는 나머지 두 입장과의 논의 수준이 다르다고 보기 때문이다.

있다고 본다.⁵⁾

이 글의 목표는 일차적으로 선의지가 무엇인지를 탐구하는 것이며, 이차적으로는 이를 통해 『정초』 1절 전반부의 논의를 이해하는 것이다. 이 글의 구성은 이렇다. II에서는 개별 의도 해석을 견지하는 페이튼의 논증과 그에 대한 아메릭스의 비판, 그리고 이에 대한 페이튼의 가능한 반론을 검토한다. III에서는 『정초』 텍스트 독해를 중심으로 선의지에 대한 올바른 규정을 탐구한다. IV에서는 이 해석에 제기될 만한 반론을 다룬다.

II. 개별 의도 해석

1. 아메릭스의 페이튼 비판

개별 의도 해석을 취하는 페이튼에 따르면, 칸트는 행위에 들어있는 도덕과 유관한 요소들을 선의지와 비-선의지적인 것으로 나눈다. 비-선의지적 요소로는 “정신의 재능들”, “기질상의 성질들”(IV 393)이 있다. “정신의 재능들”에는 지적 능력이라고 간주될 만한 지성, 기지, 판단력이 속하고, “기질상의 성질들”에는 심성적인 능력, 덕이라 할 만한 용기, 결단, 초지일관성 등이 속한다. 그렇다면 왜 칸트는 비-선의지인 것들은 무제한적으로 선할 수 없고 선의지만이 무제한적으로 선하다고 여겼는가?⁶⁾

5) 이 글은 『정초』에서의 선의지의 의미에 주로 집중할 것이다. 칸트의 이후 실천철학 저작들에서 선의지 개념이 어떻게 수정, 발전되는지에 대해서는 이 글의 후반에 간략히 다루질 것이지만, 별도의 정치한 연구도 필요하다고 본다.

6) 이 글에서 내가 ‘ 좋음(good)’과 ‘선함(good)’(또한 ‘나쁨’과 ‘악함’) 표현을 번갈아서 사용하는 것은 한국어 서술의 자연스러움을 고려해서일 뿐이지, 두 말에 의미상의 차이가 있다고 생각해서가 아니다. 현 논의 맥락에서 나는 저 두 말이 의미상 구별되지 않으며, 이것이 칸트의 생각과 부합한다고 생각한다. 물론 한국어의 어떤 용법에서 선은 도덕적이고 좋음은 도덕과 무관한 의미로 쓰이는 경우도 있지만, 이것은 나의 의도가 전혀 아니다. 여기서 선의지가 무제한적으로 선하다는 것은 단지 선(좋음)이 어떠한 점에서도 모자라거나 제약되지 않는다는 의미이며, 여기에는 어떠한 도덕적 함의도 들어있지 않다.

페이튼은 이 물음에 대한 칸트의 대답을 다음과 같이 설명한다. 선의지는 어떤 상황에서도 결코 나쁠 수 없지만, 선의지를 제외한 것들은 어떤 상황에서는 나쁠 수 있다. 이를테면 어떤 사람이 뛰어나게 좋은 판단력을 가졌는데 이를 사용해서 속임수로 타인의 재산을 빼앗는 일을 했다고 하자. 이때 그는 좋은 판단력으로 나쁜 행동을 하였다.⁷⁾ 이는 판단력이 제한적으로 선한 것임을 보여주는 것이다. 이처럼 선의지 외의 것들은 모두 나쁘게 사용될 수 있으므로 제한적으로 선하고, 선의지는 나쁘게 사용될 수 없으므로 무제한적으로 선하다는 것이다.

그러나 아메릭스는 페이튼의 위의 추론이 선의지가 무제한적으로 선함을 증명하지 못한다고 본다. 뛰어난 판단력이 나쁜 행위에 사용되는 것과 같이, 선한 의지도 나쁜 행위에 사용되는 것이 가능해 보이기 때문이다. 이를테면, 선의지로 남을 도왔지만 그로써 도움받은 사람이 도리어 큰 위협에 빠지거나 큰 손해를 보는 행위가 그런 경우이다. 이는 판단력이 나쁘게 사용될 수 있으니 무제한적으로 선한 건 아니라는 것과 같은 논리에서, 선의지가 무제한적으로 선하지 않음을 보여주는 것으로 보인다.

이 비판에 대해 페이튼은 선의지로 한 행위가 나쁜 경우가 있다는 것이 선의지가 나쁘다는 것의 근거일 수는 없다고 대응한다. 선의지로 한 행위가 나쁜 것은 선의지의 탓이 아니라 다른 요소의 탓이라는 것이다. 이를테면 선의지로 했지만 나쁜 판단력으로 인해 나쁜 행위를 행하게 되었다면, 행위의

7) 허정훈은 “칸트는 우리가 선이란 말을 사용할 때, 무의식적으로 혼동하고 있는 두 의미 즉 선과 복을 구별해야 하며, 마찬가지로 악이란 말에서도 악과 화 또는 불행이란 의미를 구별해야 한다고 지적한다”(허정훈, 2003 259)고 말하며, 그 전거로 『실천이성비판』의 「순수실천이성의 대상 개념에 대하여」에서 이뤄지는 선(das Gute)과 복(das Wohl)의 구별을 상기시킨다. 하지만 나는 이 구별을 우리가 다루는 「정초」의 선의지 논의 단계에서 활용할 수는 없다고 생각한다. 우리의 논의는 보통 이성의 의견에서 출발하여 도덕적 선이 무엇인가를 밝히는 논의의 시작 단계에 있는 것으로, 이 시작 단계에서 칸트가 저 구별을 미리 전제하고 있다고 보기는 어렵다고 보기 때문이다. 여기서 칸트는 도덕이나 도덕적 가치가 무엇인지를 미리 상정하지 않고 있으며, 오히려 지금 우리가 다루는 논의를 통해서 도덕적 선이 무엇이고, 왜 특별한 가치를 갖는지를 결론으로서 증명하고자 한다. 따라서 나는 현 맥락에서 의지의 선과 다른 선이 의미상으로 구별된다고 전제하지 않을 것이다.

이 나쁨은 선의지가 아니라 판단력에 달려져야 한다는 것이다.⁸⁾

그러나 아메릭스는 이 대응이 설명을 뒤로 미룰 뿐 선의지가 무제한적으로 선함을 보여주는 못한다고 본다. 페이튼은 선의지로 한 나쁜 행위나 나쁨은 선의지의 나쁨이 아니라 다른 요소의 나쁨으로 인한 것이고, 선의지 자체는 선한 것으로 남는다고 주장한다. 하지만 페이튼은 왜 선의지만이 그러한 **예외적** 지위를 갖는지는 설명하지 못한다. 더구나, 오히려 같은 논리에서 다른 요소들의 무제한적 선을 논증할 수도 있는 것으로 보인다. 좋은 판단력과 나쁜 의지에서 나쁜 행위가 일어났다고 할 때, 나쁜 것은 의지이고 **판단력 자체는** 좋은 것으로 남는다고 말하지 못할 이유가 뭐란 말인가. 이로부터, 따라서 좋은 판단력도 마찬가지로 무제한적 가치를 갖는다고 추론하지 못할 이유가 어디에 있는가.

결국, 우리는 아무 소득 없이 처음 질문으로 돌아오게 된다. 왜 오직 선의지만이, 결과가 아무리 나쁘더라도, 좋은 것으로 남는 것인가? (행위의) 다른 요소들과는 다르게 선의지만이 저 특별한 가치를 갖는 이유는 무엇인가? 페이튼의 해석은 이 물음에 대답을 주지 못한다. 결국, 개별 의도 해석은 선의지의 무제한적 가치를 설명하지 못하는 것이다. 따라서 아메릭스는 페이튼의 개별 의도 해석이 선의지에 대한 적절한 해석일 수 없다고 결론내리는 것이다.

그렇다면 개별 의도 해석이 실패하는 지점은 어디인가? 아메릭스는 이 해석이 선의지를 좋은 행위의 **한 요소**로만 보았기 때문에 실패한다고 분석한다. 페이튼을, 재능이나 판단력처럼, 개별 행위를 이루는 요소 중의 하나로 보았다. 하지만 저것들이 모두 행위의 요소일 뿐이라면 한 요소가 다른 요소와 다른 특별한 가치(즉, 무제한적 선)를 갖는다고 할 이유는 없어 보인다. 적어도 그런 이유를 페이튼은 제공하지 못했다. 아메릭스는 페이튼의 개별 의도 해석

8) “선한 사람도 커다란 해악(harm)을 저지러 수 있음은 분명한 사실이다. 그리고 이런 해악은 도덕적 악에 속하는 참견이나 허영심에서가 아니라 단순한 게으름이나 어리석음에서 생겨날 수도 있다. [...] 그러나 나는 칸트가 어리석은 선한 사람이 저지러 해악은 그의 선함 때문이 아니라 어리석음 때문이라고 주장하리라 생각한다. 그는 분명히 선의지 자체는 그런 행위를 결합할 수 없다고 주장하는 듯하다(Paton, 1971, 40).

은 선의지의 선함과 비-선의지의 선함의 유의미한 차이를 보여줄 수 없기 때문에 선의지의 무제한적 선함을 증명하는 데에 실패한다고 주장한다. 그러므로 선의지를 개별 의도로 보는 해석은 견지될 수 없다는 것이다.⁹⁾

2. 최선의 개별 의도 해석

나는 개별 의도 해석이 아메릭스의 반론에 한 번 더 대응할 수 있지만, 최종적으로는 견지될 수 없다고 본다. 하지만 중요한 것은 이 해석이 견지될 수 없는 이유를 아는 것이다. 우리는 이 반론에 대응하는 과정에서 선의지에 관한 중요한 통찰을 얻을 수 있다. 이를 위해 칸트의 냉정한 악한 사례에 주목해 보자.

칸트는 우리가 절제나 성찰 능력을 발휘하는 악한을, 그러한 능력이 없거나 적은 악한보다 더 혐오한다고 말한다. “무릇 선의지의 원칙들이 없이는 이런 것들은 최고로 악해질 수 있고, 악한의 냉혈은, 그가 이런 것[절제, 자제, 냉철한 성찰] 없이 악한으로 여겨졌을 때보다, 그를 훨씬 더 위협하게 만들 뿐만 아니라, 직접 우리의 눈에도 더욱 혐오스럽게 만든다.”(『정초』 IV394). 그렇다면 절제하고 성찰하는 악한을 우리가 더 혐오스럽게 보는 이유는 무엇일까? 이 사례에 대해 칸트는 더 설명하지 않으므로, 아쉽게도 저 물음의 대답을 칸트에게 직접 들을 수는 없다. 하지만 나는 이 설명을 재구성한다면 선의지의 무제한적 가치 문제의 해결에 도움을 주는 통찰을 얻을 수 있다고 본다.

건전한 상식을 가진 보통 인간이 성찰하는 악한을 그렇지 않은 악한보다 더 혐오스럽게 여기는 이유는 무엇일까? 일단, 성찰하는 악한의 의지가 성찰할 줄 모르는 악한의 의지보다 더 악하기 때문은 아닐 것이다. 만약 두 악한이 똑같은 정도의 악의로 똑같이 나쁜 행동을 했다면, 그 경우에도 성찰 능력이 뛰어난 악한을 더 혐오스럽게 여긴다는 것이 보통 사람과 칸트의 생각으로

9) Ameriks(1989, 47-51)

보인다. 그렇지만 왜 전자의 성찰하는 악한을 더 혐오스럽게 여기는 것인가?
이 **추가적인** 혐오는 어디에서 비롯되는 것인가?

1장에서 논의한 아메릭스의 마지막 반론을 염두에 두고서 이 물음에 대해 생각해보자. 만일 악한 의지와 상관없이 성찰 능력 자체는 언제나 좋은 것으로 남는다면, 두 악한은 **같은 정도로** 혐오스러워야 할 것이다. 그러나 칸트도, 보통 사람도 그렇게 여기지 않는다. 우리는 성찰 능력이 좋은 악한을 더 혐오스러워하며, 이 자가 더 악하다고 생각한다. 즉, 우리는 성찰 능력이 좋은 악한이 성찰 능력이 없는 악한보다 나쁨을 더 갖는다고 생각하는 것이다. 그렇다면 성찰 능력을 가진 악한의 이 **추가적인** 나쁨은 어디에서 온 것인가? 두 악한의 의지 및 다른 능력의 정도가 동일하다고 전제하면, 나쁨을 추가해줄 것은 성찰 능력밖에 없다. 성찰하는 악한이 성찰하지 못하는 악한과 비교해 추가적으로 갖는 나쁨은 성찰 능력이 나쁘게 되어서 이 나쁨이 추가적으로 덧붙여진 게 아니라면 설명되지 않는 것이다. 만일 그렇지 않고, 앞의 아메릭스의 반론대로, 성찰 능력이 선의지와 같이 자기의 선함을 내재적으로 유지하는 것이라면, 성찰하는 악한은 성찰하지 못하는 악한과 같은 정도로 악하고 같은 정도로 혐오스러워야 할 것이다. 따라서 보통 사람이 성찰하는 악한을 더 혐오하는 것을 설명하려면, 성찰 능력 **자체**가 나쁘게 변했다고 추론할 수밖에 없다.¹⁰⁾

이 추론에 따르면 성찰 능력은 제한적으로 선하다는 결론이 나온다. 성찰 능력은 어떤 상황에서든 선한 것으로 남지는 않으므로, 그 선함이 제한적이기 때문이다. 또한, 용기, 결단력, 지성 등의 다른 능력들에게도 같은 추론을 적용할

10) 이 재구성된 논증에 대해 후자는 페이튼의 주장은 개별 행위에 대한 것인데, 내가 다루는 것은 악한이라는 한 행위자에 대한 것이므로 페이튼의 입장을 옹호하기에 적절하지 않다고 지적할 수 있을 것이다. 페이튼은 **행위**의 선함에 대한 것인데, 악한 사례는 **행위자**의 선함에 대한 것이라는 것이다. 그러나 나는 내 논증이 행위의 선함으로 확장될 수 있다고 생각한다. 본문의 논증은 다른 능력과 의지가 동일하면서 성찰 능력을 가진 악한의 행위와 갖지 못한 악한의 행위를 비교하면, 우리는 전자의 행위가 더 많은 나쁨을 갖는다는 논증으로 확장될 수 있고, 이것도 마찬가지로 보통 인간의 견해에 부합한다고 할 수 있다고 생각한다.

수 있으므로, 우리는 비-선의지 능력들도 마찬가지로 제한적으로 선하다고 논증할 수 있다.

그렇다면 같은 추론을 선의지에도 적용할 수 있는가? 그럴 수 없어야 저 추론은 최소한의 설득력을 가질 것이다. 이를테면 나쁜 판단력과 선의지를 가지고 나쁜 행동을 한 경우를 생각해보자. 이 경우를 두고, 보통 인간은 나쁜 행위가 일어나게 한 선의지 자체도 문제가 있다고는 말하지 않는다. 오히려 ‘그 사람의 의도는 선했는데...’라고 여기며, (선한) 의지가 원한 바를 관철하지 못한 것을 안타까워한다. 보통 인간은 선의지 자체에 대해서는 언제나 좋다고, 선의지가 적어도 행위의 나쁨에는 아무 기여도 하지 않는다고 여긴다. 이 나쁨은 선의지 외의 요소들에서 온 것이지, 선의지 탓은 아니라고 보는 것이다. 칸트는 이것을 보통 인간의 상식적인 견해라고 본다. 그렇다면 분명히 여기에는 어떤 근본적인 비대칭성이 있다. 보통 인간의 눈으로 볼 때, 선의지와 그 밖의 모든 다른 능력들은 근본적으로 **다른 기준으로** 대우받는다(11)

이 마지막의 논점은 페이튼의 개별 의도 해석이 견지될 수 없는 이유를 알려준다. 페이튼은 선의지와 비-선의지의 능력을 개별 행위를 구성하는 요소로만 본다는 점에서 무차별적인 것으로 대우한다. 하지만 그럴 경우 의지와 다른 능력들이 갖는 가치의 이 근본적인 차이는 설명되지 않기 때문에, 선의지의 무제한적 선이 어떻게 정당화될 수 있는지 알기가 어렵다. 그러나 선의지와 다른 행위 능력들 사이에는 어떤 근본적인 **비대칭성**이 있다. 이 차별성으로 인해, 아무리 나쁜 행위라도 그 나쁨이 선의지까지 물들일 수 없고, 선의지는 어떤 상황에서든 자신의 선함을 유지하는 것이다. 바로 그 덕분에 선의지는 다른 정신 능력들과 달리 무제한적으로 선하다.

개별 의도 해석의 가장 큰 문제는 행위에 영향을 미치는 요소들 자체의

11) 앞 장에서의 아메릭스의 반론에 대한 나의 대답과 페이튼의 대답의 차이는 나의 논의는 선의지와 다른 정신 능력들의 가치를 보통 인간은 다른 잣대로 평가하며, 이 비대칭성의 해명을 이제부터 우리가 해결해야 할 과제라고 본다는 점에 있다. 이에 반해, 페이튼의 논증은 의지가 갖는 예외적 지위를 증명하는 대신에 그저 전제한다는 점에서, 자신이 증명해야 할 것을 전제하는 선결문제의 오류를 범하고 있다.

원천을 구별하지 않는다는 것이다. 행위를 구성하는 것들(의지와 다른 능력들)을 그저 요소로만 다룸으로써 의지를 다른 요소들과 같은 층위로만, 무차별적으로 다루는 것이다. 하지만 선의지의 무제한적인 선을 설명하려면 저 가치를 주는 선의지의 저 특별한 지위가 설명되어야만 한다. 그리고 이 설명을 위해 우리는 칸트를 따라 보통 인간의 건전한 상식을 넘어서 “철학적 이성 인식”으로 이행해야 한다.

III. 선의지란 무엇인가?

1. 페이튼의 선의지 규정

페이튼은 선의지를 “**인간 조건 하에서는** 의무로 말미암아 행위하는 의지”라고 규정한다.¹²⁾ 그가 “인간 조건 하에서는”의 단서를 단 이유는 이 정의가 모든 존재자에게 적용되는 선의지의 일반적인 정의가 아니기 때문이다. 이 정의는 인간이 아닌 이성적 존재자, 즉 완전하게 이성적인 존재자에게는 들어맞지 않는다. 그러한 신성한 존재자는 선의지를 가지지만 의무를 갖진 않기 때문이다. 또한, 이 규정은 인간에게 있어서만큼은 선의지가 **의무 동기**와 밀접한 관계에 있음을 알려준다. 이에 따르면 선의지는 어떤 특정한 동기를 갖는다는 점에서 규정된다. 그런데 동기는 개별 행위마다 갖는 것이므로, 저 규정은 선의지를 행위자보다는 행위의 개별 의도의 성질로 보는 해석을 뒷받침하기 쉽다.

페이튼은 저 규정을 두 단계로 증명한다. 첫째, 행위 일반에서 그것의 한 종류인 의무로부터의(*aus Pflicht*) 행위를 가려낸다. 둘째, 의무로부터의 행위만이 무제한적인 가치를 가짐을 밝힌다. 그런데 오직 선의지만이 무제한적인

12) “A good will *under human condition* is one which acts for the sake of duty.” (Paton 1971, 46)

가치를 가지므로, 의무로부터의 행위는 선의지로 한 행위와 같다는 결론이 나온다. 따라서 의무로부터의 행위와 선의지로 한 행위는 같다는 것이다.¹³⁾

하지만 페이튼의 논증은 의무로부터의 행위가 선의지를 갖는 행위임을 보이기에는 하지만, 선의지를 갖는 행위가 의무로부터의 행위임을 증명하지는 못한다. 다시 말해, 선의지를 가짐이 의무로부터 행위함의 필요조건임은 증명하지만, 충분조건임은 증명하지 못한다. 따라서 의무로부터의 행위를 통해 선의지가 무엇인지를 밝히려는 페이튼의 시도는 불완전하다.

물론 의무로부터의 행위는 선의지가 무엇인지 알기 위해 아주 중요하다는 점을 부정할 수는 없다. 칸트가 의무로부터의 행위 개념을 도입한 것은 선의지 개념을 “발전시키기” 위해서였다. “의무 개념은 비록 어떤 주관적인 제한들과 방해들 하에서이기는 하지만 선의지의 개념을 함유”(『정초』 IV397)한다. 하지만 둘의 관계는 페이튼의 말처럼 “인간 조건 하에서는”이라는 단서만 붙이면 등치되는 그러한 단순한 것이 아니다.

2. 『정초』에서 ‘선의지’의 사용

사실, 『정초』 1절에서 칸트는 여러 번에 걸쳐 선의지가 개별 의도가 아니라 **성격**의 특성임을 명시적으로 밝힌다.

그러나 이런 것들[정신의 재능들, 기질상의 성질들]도, 만약 이런 천부의 자질들을 사용하는, 그 때문에 그것의 특유한 성질을 **성격(Charakter)**이라고 일컫는 의지가 선하지 않다면, 극도로 악하고 해가 될 수도 있다.(『정초』 IV393, 강조는 칸트)

13) “If we are to justify our contention that a good will—under human conditions—is one which acts for the sake of duty, we must first *isolate* actions done for the sake of duty and judge whether they possess the supreme worth which we have ascribed to a good will.”(Paton, 1971, 47)

인용에서 칸트는 (선한) 의지가 “성격”이라고 일컬어진다고 말한다. “일컬える”은 칸트가 여기서 의지에 관한 통상적인 견해를 말하고 있음을 시사한다. 우리는 의지라는 말을 대개 이렇게 이해하고 사용하고 있다는 것이다. 아마도 여기서 의지가 성격이라는 생각은 의지가 개별 행위마다 일회적으로 발휘되는 것이 아니라, 행위자가 어느 정도의 지속성과 일관성을 가지고서 발휘하는 것, 일회적 행위가 아니라 한 인간을 평가할 때 중요하게 참조하는 것이라는 생각과 관련되어 있을 것이다.

또한, 칸트는 (선한) 의지가 원칙을 갖는다고 말한다.

그러나 무릇 선의지의 원칙들이 없이는 이런 것들은 최고로 악해질 수 있고, 악한의 냉혈은, 그가 이런 것 없이 악한으로 여겨졌을 때보다 그를 훨씬 더 위험하게 만들 뿐만 아니라, 직접 우리의 눈에도 더욱 혐오스럽게 만든다.(『정초』 IV394, 강조는 필자)¹⁴⁾

원칙의 가장 중요한 점은 환경이나 외부의 충격 등으로 저절로 생겨나는 것이 아니라, 행위자가 선택하거나 만들어야만 비로소 존재하게 된다는 것이다. 여기서 칸트는 의지가 원칙을 가짐을 지적함으로써, 의지가 맹목적인 힘이나 경향성이 아니라는 것, 어떠한 이성의 결정이 들어간다는 점을 시사한다.

또한, 칸트는 의지가 원칙적 측면과 동기적 측면을 갖는다고 말한다.

무릇, 의지는 형식적인, 그것의 선형적 원칙과 질료적인, 그것의 후험적 동기 사이의 한가운데에, 말하자면 갈림길에 서 있는 것이다. 그럼에도 의지는 무엇인가에 의해 결정되어야만 하는 것이기 때문에, 만약 행위가 의 무로부터 말미암아 일어난다면 의지에서 모든 질료적 원칙은 제거된 것이

14) 선의지가 일련의 원칙들이라는 생각은 다음 인용에서도 발견된다. “예컨대 가게 주인이 어리숙한 고객에게 바가지를 씌우지 않는 것은 물론 의무에 맞는 일이다. (...) 그러나 그렇다고 해서, 그 상인이 의무로부터, 그리고 정직의 원칙들에서 그렇게 처신했다고 믿기에는 아직 불충분하다.”(『정초』 VI397) 논의의 맥락상 여기서 “정직의 원칙들”은 선의지로 대체될 수 있다.

므로, 의지는 의욕 일반의 형식적 원칙에 의해 결정되지 않을 수 없다.(『정초』 IV400, 강조는 필자)

의지는 ‘어떤 원칙에 의해 결정될 것이냐’와 ‘어떤 동기에 좌우될 것이냐’라는 두 가지 면을 갖는다. 전자에서 의지는 자기가 따를 원칙을 정하고, 후자에서 의지는 자신의 욕구, 동기를 갖는다. 전자는 의지의 원리적인 면이고, 후자는 동기적인 면이다. 전자가 행위의 근거가 되는 원칙과 관계된 것이라면, 후자는 실행의 주관적인 면과 관계된다.

그렇다면 선한 의지도 저 두 가지 면을 가질 것이다. 칸트 식으로 말하면, 선의지도 어떠한 형식적인 면과 질료적인 면을 갖는 것이다. 우선, 형식적, 원리적인 면에서 선의지는 선한 원칙을 갖는다. 그리고 선의지의 그러한 원칙은 1절 후반부에 가서 도덕법칙, 보편화 가능한 준칙임이 밝혀질 것이다. 다음으로, 질료적, 동기적인 면에서 선의지는 도덕법칙에 대한 존경을 동기로 갖는다. 결국, 선의지는 보편화가능한 원칙 혹은 법칙을 존경하기 때문에 따르려는 의지의 성격(혹은 성질)이다. 반면, 선하지 않은 의지는 자신의 주관적인 경향성 때문에 편향적인 원칙을 따르려는 의지의 성격일 것이다.¹⁵⁾

그렇다면 의무 동기와 선의지의 관계는 무엇인가? 우리는 앞서 이 둘을 등치함으로써 선의지를 규정하려는 페이튼의 시도가 실패함을 보았다. 이 둘의 관계, 그리고 각각의 가치인 도덕적 가치와 무제한적 가치를 어떻게 이해해야 할까?

둘째 명제는 이렇다: 의무로 한 행위는 그것의 도덕적 가치를, 그 행위를 통해 달성해야 할 의도에서(in) 갖는 것이 아니라, 그에 따라 그 행위가 결의되는 준칙에서(in) 갖는 것으로, 의무로부터의 행위는 그러므로 행위 대상의 현실성에(von) 의존해 있는 것이 아니라, 순전히 욕구 능력의 모든 대상과는 무관하게 행위를 일어나게 한 의욕의 원칙에(von) 의존해(hängt) 있는

15) 이는 『정초』 1절의 첫째 명제와 둘째 명제로부터 추론되는 셋째 명제의 내용과도 같다. “의무는 법칙에 대한 존경으로부터 말미암는 행위의 필연성이다.”(『정초』 IV400)

것이다. (『정초』 IV399-400)

여기서 칸트는 의무 동기로부터의 행위가 갖는 **도덕적 가치의 원천**을 밝히는 데, 의무로부터의 행위의 도덕적 가치는 “그 행위가 결의되는 준칙”, 그 행위를 일어나게 한 “의욕의 원칙”에서 비롯된다고 한다. 도덕적 가치는 그 행위가 일으키거나 일으킬 것으로 기대되는 것과는 무관하다는 것이다.

행위를 도덕적인 것으로 만들어주는 도덕적 가치에 대해서는 세 가지 점을 추가적으로 고려할 필요가 있다. 첫째, 도덕적 가치는 행위의 평가적 술어이다. 도덕적 가치를 갖는 것은 의지나 행위자가 아니라, 적어도 일차적으로는 행위이다.¹⁶⁾ 둘째, 도덕적 가치를 갖는 행위는 의무 동기에서 한 행위이며, 이 동기는 경향성 동기의 대안적인 동기이다. 즉, 행위자는 경향성의 동기row 행위하거나 의무 동기로 행위한다. 이는 앞서 인용한, 의지가 선험적 원칙과 후험적 동기 사이의 갈림길에 있다는 언급에서도 드러난다. 그리고, 셋째, 의무 개념은 선의지 개념을 함유한다.¹⁷⁾ 즉, 의무 동기로 행위하는 행위자는 선의지를 갖는다.

이제, 이상의 (행위의) 도덕적 가치 논의를 오직 선의지만이 무제한적 가치를 갖는다는 생각과 비교해서 생각하면, 의무 동기에서 한 행위의 도덕적 가치의 원천이 무엇인지를 알기는 어렵지 않다. 도덕적 가치의 원천은 행위자의 의지이다. (도덕적) **행위**의 가치는 그 행위를 한 **행위자의 의지**의 가치에서 비롯되는 것이다. 결국, 의무 동기로 한 행위는 자기 가치의 원천을 선의지에 두고 있다. 그렇다면 선의지가 갖는 무제한적 선험의 가치는 어디에서 비롯된 것인가? 앞서 보았듯이, 의지는 형식적, 원칙적인 면에서 볼 때, 원칙들을 승

16) 이 점은 다음 구절로도 뒷받침된다. “우리가 행위들에서 가질 수 있는 의도들과 그리고 의지의 목적들 및 동기들인 그 행위들의 작용들이 행위들에게 무조건적인 도덕적 가치를 부여할 수 없음은 앞서의 설명으로 명백하다.”(『정초』 IV400) 그 외에도 IV397, 398, 401에서 도덕적 가치는 행위를 수식하고 있다.

17) “이 의무 개념은 비록 어떤 주관적인 제한들과 방해들 하에서이기는 하지만 선의지의 개념을 함유하는 바 [...]” (『정초』 IV397).

인하고 채택하는 능력이다. 따라서 선한 의지의 선함의 원천은 선한 원칙, 즉, (결론만 말하면) 보편화가능한 준칙이다.

의무, 도덕적 가치는 개별 행위와 유관한 개념이다. 동기(의무, 경향성)는 행위마다 발휘되며, 의무 동기일 때 그에 따른 행위는 도덕적 가치를 갖는다. 이에 비해, 선의지는 행위자의 지속적인 성격과 유관한 개념이다. 행위자는 자신의 행위 원칙을 결정함으로써 자기 의지의 성격을 결정한다. 선의지를 가진 행위자는 개별 상황에서 개별 의무 동기로 행위하면서 자신의 선의지를 발휘한다. 이러한 의무로부터의 행위는 도덕적 가치를 가지며, 이 가치는 무제한적으로 선한 그의 선의지에서 비롯되는 것이다.¹⁸⁾

요컨대, 칸트는 선의지란 선한 원칙, 즉 보편화가능한 원칙을 자기 준칙의 원리로 삼겠다는 행위자의 의지의 성격이고, 이때 그의 의지는 무제한적인 가치를 갖는다고 본다. 선의지를 가진 행위자가 개별 상황에서 의무 동기를 가지고 행위할 때, 그의 행위는 선의지로부터 비롯된 가치인 도덕적 가치를 갖는 것이다.¹⁹⁾

후기 실천철학으로 가면서 칸트는 의지 개념을 수정한다. 『윤리형이상학』에서는 『정초』의 의지를 의지(Wille)와 의사(Willkür)로 재구분한다.²⁰⁾ 의

18) 앨리슨도 선의지를 성격으로 해석하는 입장을 옹호하면서, 이러한 성격은 어떤 감정적인 상태나 습관이라기보다는 사고방식(Denkungsart)으로 이해해야 한다고 주장한다(Allison, 2011, 78-80). 선의지는 좋은 사고방식으로, “도덕적으로 선한 원칙들을 굳게 승인하는 것”이라는 것이다. 이러한 해석은 의지를 정념보다는 이성적인 원칙으로 이해하는 해석에 무게를 실어준다.

19) 성격 해석을 지지하는 다른 학자들의 선의지 규정들도 필자의 규정과 대동소이하다. 이를테면 우드는 “A good will is such a faculty when it adopts good principles and sets about acting on them.”(Wood, 2008, 32)로, 존슨은 “A good will is thus the disposition to choose courses of action on the basis of certain sort of policies”(Johnson, 2009, 23)로 정의한다.

20) “개념들에 따르는 욕구능력은, 행위를 위한 욕구능력의 규정근거를 그 자신 안에서 마주치고, 대상에서 마주치지 않는 한에서, **임의대로 행동하는** 능력이라 일컫는다. 이 욕구능력이 대상을 만들어내기 위한 자기의 행위 능력에 대한 의식과 결합되어 있는 한에서, 그것은 **의사(Willkür)**라고 일컫는다. 그러나 욕구능력이 그런 것과 결합되어 있지 않으면, 욕구능력의 작용은 **소망**이라고 일컬어진다. 그 내적 규정근거가, 따라서 그 임의가 주체의 이성 안에서 마주쳐지는 그런 욕구능력은 **의지(Wille)**라고 일컫는다. 그러므로 의지는 (의사처럼) 행위와의 관계에서가 아니라, 오히려 행위로의 의사를 규정하는 근거와의 관계에서 고찰되는 욕구능력

사(Willkür)는 어떤 행위를 일으키는 원인이 되는 능력이다. 이는 소망과 상대적인 능력으로, 행위자가 어떤 결과를 바라기만 할 때 그 행위를 의지(Willkür)하는 게 아니라 소망하는 것이다. 자기 행위로서 그 결과의 원인이 되고자 할 때, 비로소 행위자는 그 행위를 의지(Willkür)하는 것이다.²¹⁾ 이에 반해 의지(Wille) 개념은 행위의 원리적 근거와 더 연관된다. 이성적 행위자는 어떤 규칙이나 원칙에 따라서 움직일 수밖에 없으므로, 자기를 움직이는 원칙을 정해야 한다. 이 능력이 의지(Wille)이다. 요컨대, 의사(Willkür)는 실행 능력, 의지(Wille)는 의사 결정의 원칙을 정하는 능력이라고 할 수 있다.

『윤리형이상학』의 이러한 의지(Wille) 개념은 『정초』의 독자에게는 생소해 보인다. 『정초』에서 의지는 어떤 표상을 원인으로 삼아 행위하는 능력으로 주로 실행능력으로 규정되고, 원칙을 정하는 능력으로서의 의지 개념은 상대적으로 덜 다루어지기 때문이다. 그렇지만 선의지는 의지의 원리적 측면과 밀접하게 결부되어 있다.²²⁾ 『정초』의 선의지는 선한 행위를 실행하는 능력이라기 보다는 그러한 실행을 가능하게 하는 원칙을 채택하는 능력에 가깝다. 이는 선의지가 성격이라는 규정과도 잘 부합한다. 선의지는 선한 원칙을 채택하는 성향이다. 행위자의 도덕적 성격은 행위자가 어떤 원칙을 채택하느냐로 결정되며, 이 성격에 따라서 그의 의지가 평가되는 것이다.

IV. 성격 해석에 대한 반론과 함축

지금까지 나는 선의지가 행위자의 의지의 어떤 성격이라는 해석을 옹호하

이며, 그 자신 앞에 본래 어떠한 규정근거도 가지고 있지 않고, 오히려 의지는, 그것이 의사를 규정할 수 있는 것인 한에서, 실천이성 자체이다. 『도덕형이상학』, VI213(강조는 칸트).

21) 의지와 소망의 차이는 이미 칸트가 『정초』에서 가언명령의 분석성을 도출할 때 적극적으로 개진했던 것이기도 하다.

22) 게다가 선의지를 원칙과 연결하는 몇몇 표현들이 발견된다. 앞서 인용된 IV394에서 칸트는 “선의지의 원칙들”이라는 표현을 사용하고 있으며, “의욕의 원칙”(IV400)라고도 한다. 칸트는 계속해서 선의지를 어떤 원칙의 능력으로 보려고 한다.

며, 선의지 개념을 규정하고자 하였다. 이러한 해석에 관해서는 다음과 같은 반론이 가능해 보인다. 만약 선의지가 행위자의 선한 성격이라면, 선의지를 갖는 행위자는 결코 잘못을 범할 수 없을 것처럼 보인다. 그는 언제나 적절한 상황에서는 의무로부터의 행위를 어김없이 할 것이고, 그의 모든 도덕과 유관한 행위는 도덕적 가치를 가질 것이다. 하지만 이는 불가능하다. 제아무리 선의지를 갖는 행위자라고 해도 일탈이나 의무에 반하는 행위를 할 수 있는 것으로 보이기 때문이다.²³⁾

과연 이 반론이 성립하는지 검토하기 위해 반론을 더욱 구체화시켜서 보자. 어떤 사람이 선의지를 가졌음에도 의무에 어긋나는 행위를 하는 것이 가능한가? 선한 원칙을 자기 의지의 법칙으로 승인하고 있으면서도 의무에 어긋나는 행동을 하는 것이 가능한가? 그러한 행동으로 우리는 두 가지 경우를 생각할 수 있다. 첫 번째는 선의지를 가진 행위자가 의무 동기로 행동했지만 잘못된 판단으로 인해 의무에 반하는 행위를 하는 경우다. 두 번째는 행위자는 선의지를 가지고 있고, 즉 도덕법칙에 대한 존경에서 의무를 따르겠다는 원칙을 견지하고 있지만, 일시적인 유혹에 이기지 못하고 의무에 반하는 행위를 하는 경우다.

두 번째 경우의 행위가 도덕적 가치를 가지지 않는다는 것은 명백해 보인다. 확실히 그는 의무 동기에서 행동하지 않았기 때문이다. 그는 선한 원칙을 자신의 행위 원칙으로 정하는 데에는 성공했지만, 이 원칙으로써 자기 행위를 결정하는 데에는 (적어도 이번엔) 실패했다. 그는 의무에서가 아니라 경향성에서 행동한 것이다. 따라서 그의 행위는 도덕적 가치를 가질 수 없다.²⁴⁾ 이는

23) 선의지를 성격으로 해석하는 입장에 대한 표준적인 반론은 선의지를 성격으로 해석할 경우 개별 행위에 도덕적 가치를 귀속시키기가 곤란해진다는 것이다. 필자는 도덕적 가치나 행위의 성격을 논하면서 이 문제가 해소됐다고 여겼기 때문에 본 논문에서 이 반론을 명시적으로 다루지 않았다. 이 반론에 대답하려는 시도로는 Allison(2011)과 Ameriks(1989)의 논의를 참조하라.

24) 이처럼 의무 동기와 선의지가 완전히 맞물리지 않는다는 점이 칸트에게 도덕을 따르는 의지의 순수성/불순성 문제를 안겨준 것일 수 있다. 『이성의 한계 안에서의 종교』에서 칸트는 “법칙 그것만을 충분한 동기로서 그 안에 채용한 것이 아니라, 의무가 요구하는 바로 그 의

설령 선의지를 갖는 행위자라도 의무에 반하는 행위를 할 수 있음을 알려준다. 모든 인간은 의무 동기에서 행동하기 위해서 내적 투쟁을 벌일 수밖에 없는 존재이며, 이 투쟁이 불필요해지는 단계에 적어도 인간이 다다를 수는 없다.

첫 번째 경우는 더 까다로워 보인다. 이 경우 행위자의 행위는 (가정상) 의무에 반하지만 동시에 (가정상) 의무 동기로 이루어진 것이기 때문이다. 하지만 여기엔 중요한 고려사항이 하나 빠져 있다. 바로, 의무에 반하는 행위는 의무 동기로 이뤄질 수 없다는 것이다. 칸트는 도덕적으로 선한 행위는 의무 동기에서 한 것이어야 할 뿐만 아니라 합법적인(의무에 맞는 *pflichtmässig*) 것이어야 한다고 주장한다. 그래서 문제의 행위는 비합법적이므로 (즉, 의무에 맞지 않으므로) 도덕적으로 선한 행위일 수 없다.²⁵⁾

이러한 두 가지 경우에 대한 고려는 선의지를 논하는 영역과 의무 동기를 논하는 영역이 다르다는 것을 알려준다. 하나의 지속하는 인격으로서의 인간을 도덕적으로 평가하고자 할 때 우리는 그의 의지가 선한지를 본다. 반면, 개별 행위를 도덕적으로 평가하고자 할 때는 그 행위 시에 발휘된 동기를 고려한다. 물론 이 동기가 의지의 선과 무관한 것은 아니다. 선의지를 갖지 않는 행위자는 도덕적 가치를 갖는 행위를 할 수도 없다. 하지만 선의지를 갖는다고 해서 언제나 도덕적 가치를 갖는 행위만을 하는 것은 아니다. 선의지를 가진 인간도 의무로부터 행위하기 위해서 고비마다 투쟁해야 하는 것이다.

또한, 행위와 행위자의 선함의 이러한 논의 영역의 구별은 도덕과 무관한 영역을 확보하는 데에도 용이할 수 있다. 선의지를 갖는 행위자의 모든 행위가 도덕과 유관하지는 않다. 사실, 상당수의 행위는 선하지도 악하지도 않음

사를 규정하기 위해서 그 동기 외에 흔히 또 다른 동기들을 필요로 하는'(VI29) 것이 악으로서의 본성적인 성벽이라고 말한다.

25) 사실, 칸트의 실천철학에서 도덕성과 합법성의 관계는 깊이 있는 논의가 필요한 흥미로운 문제거리다. 나는 이 부분에 대해서는, 행위가 도덕적으로 선함 (즉, 도덕적 가치를 가질) 조건으로서, 그 행위가 의무에 맞을(*pflichtmässig*) 뿐만 아니라 의무로부터 나와야(*aus Pflicht*) 한다는 칸트의 생각(『정초』 VI397)을 반복하는 것으로 논의를 유보하고 싶다.

것이다. 행위자는 선의지를 통해 도덕적인 원칙을 채택하고 있지만, 모든 순간에 도덕적으로 가치있는 행위를 하는 것은 아니다. 도덕과 무관한 일을 할 수도 있고, 유혹에 저서 도덕에 반하는 행동을 할 수도 있는 것이다.

V. 결론

이 글에서 나는 선의지에 관한 성격 해석을 옹호함으로써 『정초』 논의에서 선의지의 자리가 어디인지에 답하고자 하였다. 선의지는 보편화가능한 원칙을 자기 준칙의 원리로 삼겠다는 행위자 의지의 성격이다. 이 과정에서 나는 선의지를 행위를 구성하는 한 요소로 보는 개별 의도 해석이 선의지를 무제한적인 가치를 설명하는 일과 양립될 수 없음을 보였고, 의무 동기와 선의지 논의는 밀접한 연관을 갖긴 하지만 행위와 행위자에 관한 것으로서 서로 다른 범위를 가짐을 밝혔다. 의무는 개별 행위의 동기로서 행위가 도덕적 가치를 가질 기준이며, 선의지는 선한 원칙을 채택하려는 행위자가 지속적으로 견지하는 성격이라는 것이다.

지금껏 칸트의 실천철학은 행위 윤리학에 경도되어 해석되어왔고 『정초』 1절의 의무, 도덕적 가치 논의는 종종 이러한 해석을 뒷받침하는 전거로 사용되기도 하였다. 하지만 이러한 해석은 더 이상 견지되기 어렵다. 선의지는 특히 일회적이고 개별적인 행위들을 묶는 공통 원칙을 관철시키는 어떠한 성격으로서, 한 행위자의 자기 표현으로서 보는 것이 칸트의 생각에 더 부합한다.

【주제어】 선의지, 무제한적 가치, 가치의 비대칭성, 의무 동기, 성격 해석

[참고문헌]

- 허정훈 (2003). 칸트의 선의지에 관한 연구. *교육과학연구*, 5(2), 253-268.
- Allison, H. (2011). *Kant's Groundwork for the Metaphysics of Morals — A Commentary*. Oxford: Oxford University Press.
- Ameriks, K. (1989). Kant on the good will. In Höffe, Bittner eds. *Kant's Grundlegung zur Metaphysik der Sitten—Ein Kooperativer Kommentar*. Frankfurt am Mein: Klostermann.
- Johnson, R. (2009). Good will and the moral worth of acting from duty. In Hill ed. *The Blackwell Guide to Kant's Ethics*. [n.p.]: Wiley-Blackwell.
- Kant, I. (1990) *Metaphysische Anfangsgründe der Tugendlehre*. Hamburg: Meiner. 백종현 옮김 (2012). *윤리형이상학*. 서울: 아카넷.
- _____ (2016). *Grundlegung zur Metaphysik der Sitten*. Hamburg: Meiner, 백종현 옮김 (2005). *윤리형이상학 정초* 서울: 아카넷. 김석수, 김종국 옮김 (2019). *도덕형이상학정초 실천이성비판*. 파주: 한길사.
- _____ (2017). *Die Religion innerhalb der Grenzen der bloßen Vernunft*. Hamburg: Meiner. 백종현 옮김 (2011). *이성의 한계 안에서의 종교* 서울: 아카넷. 백종현 옮김.
- Paton, H. (1971). *The Categorical Imperative—A study in Kant's Moral Philosophy*. Philadelphia: University of Pennsylvania Press. 김성호 옮김 (1988). *칸트의 도덕철학*. 서울: 서광사.
- Wood, A. (2008). *Kantian Ethics*. Cambridge University Press.

[국문초록]

본 논문에서 나는 선의지에 관한 개별 의도 해석을 반박하고 성격 해석을 옹호함으로써 『정초』 1절의 논의에서 선의지 개념이 차지하는 자리를 밝히고자 하였다. 결론은 보편화가능한 원칙을 자기 행위의 준칙의 원리로 삼겠다는 행위자의 의지의 성격이 선의지라는 것이다. 선의지가 행위 시에 발휘되는 개별 의도라는 해석은 선의지라는 요소가 왜 다른 정신 능력 요소들과 비교해 특별한 가치를 갖는지를 설명하지 못한다는 점에서 선의지를 무제한적인 가치를 설명하지 못한다. 반면, 선의지를 선한 원칙을 채택하는 행위자의 의지의 성격으로 해석하면, 행위의 도덕적 가치의 원천을 설명하며, 이로부터 개별 행위와 이를 실행하는 행위자가 채택한 원칙을 분리하여 고찰할 여지가 생긴다. 의무로부터의 행위의 도덕적 가치는 선의지의 무제한적 가치로부터 나오지만, 선의지를 가진 행위자가 언제나 의무 동기에 따라 행위하도록 보장하는 것은 아니다. 인간 의지는 언제나 자신이 채택한 선한 의지를 개별 행위에서 관철하기 위한 투쟁 중에 있는 것이다.

[Abstract]

Good Will as Agent's Character in *Groundwork*

Kang, Euna (Seoul National University of Science & Technology)

I would clarify the position of concept of good will in the discussion in Section 1 of *Groundwork* by refuting the particular intention interpretation and advocating for the character interpretation of good will. The conclusion is that good will is the character of an agent's will to adopt the principle that his maxim of action is ruled by a universalizable principle. The interpretation that good will is a particular intention exercised in each action cannot explain the unlimited value of good will in that it does not explain why good will as an element has such a special value compared to other elements of mental abilities. On the other hand, interpreting good will as the character of the agent's will to adopt good principles explains the source of the moral value of an action, and from this, there is room for separate consideration of a particular action and the agent's principle of action. The moral value of an action from duty comes from the unlimited value of good will, but it does not guarantee that a person of good will would always act from duty. Human will is always in the midst of a struggle to realize the good will one has adopted by doing particular actions.

[Keywords] good will, unlimited worth, asymmetry of value, motive of duty, character interpretation

논문투고일: 2024년 02월 24일 / 논문심사일: 2024년 04월 05일 / 게재확정일: 2024년 04월 25일

[저자연락처] chulssu@naver.com