

【논문】

모든 인간은 본성상 악하다*

— 악의 보편성에 관한 칸트의 증명 —

김 화 성

【주제분류】 윤리학, 근대독일철학

【주요어】 악, 보편성, 도덕성, 자유, 선택의지, 칸트

【요약문】 칸트는 『단순한 이성 한계 내에서의 종교』 제1부에서 악의 문제를 본격적으로 다루었다. 하지만 이 논문은 이미 철학자로서의 명성이 높던 그에게 찬사보다는 거부와 비판을 안겨 주었다. 특히 인간이 본성상 악하다는 그의 주장은 경험적 사실들에 의존하고 있거나 전혀 증명하지 않았다는 해석이 오늘날에도 여전히 팽배하다. 이는 보편적 악에 대한 증명이 텍스트 속에서 그리 명확하게 제시되어 있지 않을 뿐만 아니라 그와 관련한 칸트의 언급들이 서로 모순적이라는 것에 기인한다. 본 논문은 문제가 되는 텍스트의 면밀한 재검토를 통해 보편적 악의 증명을 재구성하여 기존의 해석이 타당하지 않을 뿐만 아니라 왜 모든 인간은 예외 없이 악한지를 칸트가 설득력 있게 제시하고 있음을 보이고자 한다. 또한 이를 통해 흔히 근본악의 개념과 관련해서 지적되곤 하는 모순성, 즉 악은 자유로부터 초래된다는 주장과 악은 인간에게 보편적이고 생득적이므로 필연적이라는 주장 사이의 모순도 잘못된 해석에 기인한 오해에 불과함을 밝히고자 한다.

* 필자는 이 논문을 지난 5월 8일에 중앙대학교 중앙철학연구소가 ‘악의 문제’라는 주제로 개최한 춘계학술발표회에서도 발표할 기회를 가졌다. 이 자리를 빌려 맹주만 연구소장님과 더불어 열띤 토론과 값진 지적을 해 주신 참석자 여러분께 깊이 감사드린다.

칸트는 『베를린 월간지(Berlinische Monatsschrift)』 1792년 4월 호에 처음으로 발표한 “Von der Einwohnung des bösen Prinzips neben dem Guten: Oder über das radiakle Böse in der menschlichen Natur (선 이외에 악한 원리의 내재성에 대하여: 또는 인간본성의 근본악에 관하여)”라는 논문에서 악의 문제를 본격적으로 다루었다. 그리고 그는 이 논문을 다시 1794년에 『단순한 이성 한계 내에서의 종교(Die Religion innerhalb der Grenzen der bloßen Vernunft)』(이하에서는 『종교』로 약칭함)의 제1부로 수록하여 출판하였다.¹⁾

이 논문은 이미 철학자로서의 명성이 높던 그에게 찬사보다는 거부와 비판을 안겨 주었다. 인간의 자기규정성이라는 주체성의 이념을 핵심으로 한 계몽주의의 절정에서 당시 지성인들에게 이 짧은 논문은 전통적인 기독교의 핵심교리 중 하나인 원죄설을 새로이 옷단장시키는 복구적 시도로 비추어졌기 때문이었다. 널리 알려진 바대로 괴테는 헤르더에게 보낸 한 편지에서 이 논문을 칸트가 자신의 이론을 대중화시키려는 의도에서 기독교인들로부터 환심을 사기위한 시도라고 비꼬면서 그의 철학의투를 근본악의 불명예스러운 얼룩으로 더럽혔다고 혹평했다.²⁾ 나아가서 이 논문은 칸트의 도덕철학 체계와 관련해서도 논쟁을 불러일으켰다. 과연 그의 도덕철학 체계에서 악에 관한 논의는 어떤 의미를 지니는가? 특히 칸트가 선과 자유에 대비되는 악과 자유라는 새로운 관계를 논의함으로써 그의 자유에 입장변화가 발생하고 있지는 않은가?³⁾

1) 칸트는 이 과정에서 사상을 검열하려는 프로이센 당국과 갈등을 겪게 되었다. 자세한 내용은 Ch. Schulte(²1991), 13-21쪽과 B. Stangneth(2003), IX-LXI쪽 참조.

본 논문에서 모든 칸트의 인용은 아카데미판의 『칸트 전집(Kants gesammelte Schriften)』을 이용하며, 인용한 책의 권수는 로마숫자로 그리고 쪽수는 아라비아 숫자로 각각 본문이나 주석의 괄호 안에 표기한다.

2) Ch. Schulte(²1991), 26쪽 참조. 칸트의 악에 대한 주장은 기독교적 원죄설의 잔존물이거나 낙관주의적 계몽주의와 기독교적 사유전통 사이의 갈등이 표출한 것이라는 해석은 오늘날의 칸트연구에서도 여전히 발견된다. D. O'Conner(1985), 299-300쪽과 G. E. Michalson(1990), 6-9쪽 참조. 이 같은 해석에 비판적 입장으로는 Ch. Schulte(²1991), 21-32쪽 참조.

이러한 논쟁에 대한 적절한 평가를 위해서는 우선 칸트가 제시하고 있는 악의 개념에 대한 적절한 이해가 선행되어야 한다. 하지만 이와 관련하여 그의 논문은 독자들에게 적지 않은 어려움을 안겨 주고 있다. 그 중 하나가 칸트의 지지자조차도 당혹시켰던 주장, 즉 모든 인간은 본성상 악하다는 주장에 관한 증명문제일 것이다. 일견 기독교의 원죄설을 그대로 수용한 듯이 보이는 이 주장에 대한 일반적 평가는 매우 비판적이다. 칸트가 주장하는 악의 보편성은 전적으로 경험적 일반화에 의존하거나⁴⁾ 또는 단순히 경험적 가정에 지나지 않으며⁵⁾ 나아가서는 애초부터 증명 불가능한 것으로서⁶⁾ 그 타당성이 전혀 입증되지 않은 것에 불과해 보이기 때문이다. 한마디로 칸트는 어떠한 납득할 만한 증명도 없이 단지 경험적 사실들을 토대로 악의 보편성을 주장하였다는 것이다. 물론 이 같은 비판적 해석에 대해 칸트의 주장을 옹호하는 입장들도 있다. 하지만 이들 또한 칸트가 보편적 악의 증명을 『종교』에서는 제시하고 있지 못하다는 점에는 전적으로 동의하고 있다.⁷⁾ 과연 칸트는, 비판자와 옹호자 모두가 한결같이 주장하

-
- 3) 이 질문에 G. Prauss(1983)는 긍정적으로 답변한(특히 83-100쪽 참조) 반면에, J. Bojanowski(2006)는 자유개념의 일관성을 주장한다(229-264쪽 참조). 이처럼 칸트에게 악의 문제는 그의 실천철학을 평가하고 음미할 때 결코 무시할 수 없는 중대한 사안임에도 불구하고 많은 영향력 있는 칸트문헌들에서 이 문제는 매우 소홀히 다루어지고 있다.
- 4) 이 같은 주장으로는 P. Quinn(1988), 111쪽과 Th. Buchheim(2001), 663쪽 그리고 P. R. Frierson(2003), 104쪽 참조. 또한 P. Guyer(1998)는 칸트의 이 주장을 경험적 증거와 원죄설의 혼합물이라고 비판한다(192쪽).
- 5) M. Willaschek(1992), 153쪽과 D. O'Conner(1985), 298쪽 참조.
- 6) 이 같은 주장은 R. Wimmer(1990), 122쪽 및 129쪽과 Ch. Schulte(1991), 99-101쪽 참조. 그 외에도 G. E. Michalson은 악의 보편성 주장의 근거로서 인간의 육체성을 제시한다(1990, 69쪽). 하지만 이는 감각적인 것을 악의 원인으로 본 스토아주의에 비판적이었던 칸트의 입장과 정합적이지 않다.
- 7) 보편적 악에 관한 칸트의 주장은 대부분 비판적이거나 회의적으로 평가된다. 이에 비해 그의 주장을 옹호하는 연구는 극히 소수에 불과하다. 그럼에도 그 옹호전략은 다양각색이다. H. E. Allison(1990)은 칸트가 『종교』에서 증명을 제시하는 데에 성공적이지 못했음을 인정하면서 다음과 같은 대안적 증명을 제시한다. 즉, 그에 따르면 인간이 '선への 성향'대신에 '악への 성향'을 가지는 것은 필연적이다. 왜냐하면 인간은 감각적으로 축

듯이, 『종교』에서 보편적 악에 관한 증명을 제시하지 못했는가?

본 논문은 모든 인간은 본성상 악하다는 칸트의 주장에 관한 증명문제를 『종교』 제1부에 대한 보다 면밀한 텍스트분석을 통해 재검토해 보고자 한다. 이를 위해 먼저 악의 문제를 다루고 있는 텍스트의 구성을 개괄한 다음, 악 개념에 대한 칸트의 이해를 살펴보고자 한다. 이를 바탕으로 특히 악의 보편성 주장에 대한 증명의 단초를 논의하는 텍스트를 분석한 후, 악의 보편성에 대한 칸트의 증명을 재구성해보고자 한다.

발되는 욕망과 경향성을 넘어서지 못하는 유한한 또는 신성하지 못한 존재여서 결코 자발적으로 도덕법칙에 복종하지 않기 때문이다(154-157쪽 참조). 그러나 이러한 접근은 악에 대한 성향이 인간 자유와 결부된 문제가 아니라 인간현존의 유한성으로부터 도출되는 자연스럽고 불가피한 결과임을 의미할 뿐이다. 그 결과 악은 자유로부터 나온다는 칸트의 핵심 테제는 사실상 설 자리를 잃게 될 것이다. A. Wood(1999)는 『종교』에서 드러나고 있지 않은 보편적 악의 원리를 칸트가 인간의 사회적 본성으로 언급한 “반사회적 사회성(unsociable sociability)”에서 찾아야 한다고 주장한다(288쪽). 즉, 악의 원천은 인간들이 상호비교와 적대관계라는 조건을 통해 사회를 형성한다는 데에 있다는 것이다. Wood의 인간학적 접근에 S. R. Grimm(2002)은 원칙적으로 동의하지만, 인간은 원래 선했지만 사회형성과 더불어 악이 발생하게 되었다는 루소식의 설명에는 반대한다. 오히려 그는 보편적 악의 원천을 자연적인 감정과 그 이후에 발생하게 되는 이성 간의 갈등 내지는 투쟁에서 찾고자 한다(170-171쪽 참조). 그러나 이러한 인간학적 내지는 발생학적 해석은 칸트가 악의 문제를 인간의 자유문제와 결부시키고 있다는 중요한 사실을 간과하면서 악의 문제를 Wood처럼 “자긍심(self-worth)” 사이의 길항관계의 문제(290쪽 참조)로 또는 Grimm처럼 감성과 이성 사이의 진부한 갈등 문제로 파악하고 있다. 이에 반해 칸트는 『종교』에서 보편적 악의 문제를 발생학적 관점에서가 아니라 도덕적 관점에서 자유와 관련시켜 논구하고 있다. S. Morgan(2005)은 칸트가 악의 문제를 자유의 문제와 결부시키고 있음을 직시하면서 보편적 악의 증명을 재구성하고자 한다. 그에 따르면 악은 『도덕 형이상학』의 「법론」에서 제시된 “외적 자유(outer freedom)”, 즉 자신의 무제한적인 자유에 대한 맹목적 추구의 결과이다(81-85쪽 참조). 그러나 이러한 시도는 도덕적 의미에서의 악을 사법적 차원에서의 (외적) 자유로부터 규정짓는 우를 범하고 있다. 끝으로 비록 소수이지만 국외의 칸트연구에서는 보편적 악의 증명을 다루고 있는 반면에, 이 문제와 관련한 국내연구는 거의 찾아보기 힘든 실정이다.

I. 텍스트 구성과 증명문제

칸트는 『종교』 제1부에서 모든 인간이 그 본성에 있어 악하다고 주장한다. 어떤 보편성의 주장에서나 그러하듯이, 이 주장 또한 그것을 뒷받침하는 증명이 요구된다. 따라서 악의 보편성을 증명하는 것은 제1부의 주요한 과제라고 할 수 있다. 칸트는 이와 관련한 논의를 여섯 부분 또는 단계로 나누어 진행한다.

- 도입부(VI 19-25)에서 칸트는 자신의 도덕철학적 틀에서 악을 규정한다. 특히 그는 악을 선택의지가 자유로이 채택한 준칙과 더불어 준칙의 자유로운 채택에 있어 최초의 주관적 근거와 관련시키면서 인간의 본성상 악함을 주장한다. 이에 대한 증명을 위해서는 “인간학적 탐구”(anthropologische Nachforschung, VI 25)가 필요함을 언급한다.

- 제1절(VI 26-28)은 도덕성의 관점에서 살펴본 인간학적 탐구의 첫 부분으로서 인간 본성에 있는 ‘선에 대한 소질’(Anlage zum Guten)을 소개하고,

- 제2절(VI 28-32)은 인간학적 탐구의 두 번째 부분으로서 인간 본성에 있는 ‘악에의 성향’(Hang zum Bösen)을 소개한다.

- 제3절(VI 32-39)은 “인간은 본성상 악하다”는 제목 아래 여러 경험적 예시들을 통해 악에의 성향이 현존(Dasein)함을 언급한 후 그 성향의 본래적 성질에 대하여 논의한다.

- 제4절(VI 39-40)은 인간 본성에 있는 악의 근원(Ursprung)이 시간적 근원이 아닌 이성적 근원임을 강조하고,

- 끝으로 일반주석(VI 44-53)에서는 인간에게 보편적인 악함과 관련해서 어떻게 선에의 소질을 다시 작동시킬 수 있는가를 제시한다.

이상과 같은 텍스트 구성에 비추어 볼 때, 악의 보편성에 대한 증명은 주로 제2절과 제3절에서 제시되어 있는 것으로 보인다. 그러나 막상 이로부터 증명을 재구성하는 작업은 일단 악의 보편성에 대한 증명이 텍스트에서 그리 명확하게 드러나 있지 않다는 사실 때문에 어려움에 봉착하게 된다. 게다가, 곧 자세히 언급하겠지만, 증명의 방

향과 지침에 관한 칸트의 주요 언급들은 서로 모순하고 있어 보인다. 이러한 점들로 인해 과연 그가 악의 보편성을 증명했을까 하는 의구심과 더불어 악의 보편성에 대한 증명 자체가 애초부터 잘못 채워진 단추가 아닌가하는 비판은 피하기 어려워 보이는 것이 사실이다.

먼저 도입부의 말미에서 증명의 방향을 제시하고 있는 칸트의 첫 언급에 주목해 보자.

“우리가 인간은 본성상 선하거나 악하다고 말할 때 그 인간을 개개의 인간이 아니라 [...] 인류 전체로 이해할 권한이 우리에게 있다는 것은 오로지 인간학적 탐구를 통해 다음과 같은 것을 밝힌다면 장차 증명될 수 있다. 즉, 우리가 양자의 특성들[=선하거나 악하거나] 중 하나를 타고난 것으로 인간에게 부여하는 것을 정당화시켜주는 근거들에 따르면 그 누구도 그것들로부터 제외시킬 근거가 없으며, 따라서 그 특성은 모든 인간에게 타당하다는 것을.”(VI 25)

여기서 칸트는 인간의 본성상 선함 또는 악함을 개개의 인간과 관련된 것이 아니라 모든 인간에게 예외 없이 해당하는 것임을 강조하고, 이에 대한 증명의 실마리로서 “인간학적 탐구”를 제시한다. 그리고 이를 각각 제1절과 제2절에서 소개한 후, 제2절의 중간쯤에서 증명의 지침에 대한 두 번째 언급을 한다.

“우리는 여기서 인간에게, 심지어 (행위에 관한 한) 가장 선한 인간에게조차도 악에의 성향이 있음을 발견하게 될 것인데, 이는 만약 인간에게서 악에의 성향의 보편성이 또는 그와 동일한 의미인 악에의 성향이 인간적 본성과 밀접하게 짜여져(verwebt) 있다는 것이 **증명된다면** 분명해 질 수밖에 없다.”(VI 30, 강조는 필자)

이로부터 분명한 것은 앞서 제시된 인간학적 탐구가 증명의 실마리를 제공할지언정 그것만으로 악의 보편성이 증명되었다고 말할 수 없다는 점과 본격적인 증명은 뒤이은 텍스트에서 제시될 것이라는 점이다. 그러나 뒤이은 텍스트를 살펴보면, 과연 이곳에 제시되어야 할 증명이 담겨 있는가 하는 의심을 떨쳐 버릴 수가 없다. 덧붙여 마치

방금 언급한 문제점을 전혀 의식하지 못하는 듯, 칸트는 이어지는 제3절의 첫 단락(VI 32)에서 이미 제2절에서 악의 보편성에 대한 증명이 이루어졌다는 인상을 강하게 주고 있다. 나아가서 이어지는 다음 단락에서는 “인간의 행위들에 관한 우리의 경험이 눈앞에 제시하는 수많은 생생한 사례들”에 비추어 볼 때, 인간이 본성상 악하다는 것에 대한 “형식적인 증명”(den förmlichen Beweis, VI 33)이 불필요하다고 말함으로써 상황을 더욱 혼란스럽게 만들고 있다. 만약 그가 단지 많은 생생한 경험적 사례들에 의거해서 보편적 악을 주장하고 있다면, 이는 전혀 근거 없는 주장에 불과함을 적나라하게 드러내는 꼴이 될 것이다. 왜냐하면 보편성이란 결코 경험적으로 증명될 수 없기 때문이다.

과연 칸트는 경험적 근거를 바탕으로 악의 보편성을 주장을 했을까? 아니면 애초부터 악의 보편성을 증명하려고 의도나 했을까? 또는 은근슬쩍 증명의 부담을 회피하고 있는 것은 아닐까? 이와 관련해서 제3절의 끝에 붙여진 각주는 악의 보편성에 대한 증명에 얽힌 어려움과 혼란으로부터 벗어날 수 있는 중요한 단서를 제공하는 듯이 보인다.

“도덕적 판결을 내리는 이성의 이러한 유죄판결에 관한 본래적 증명은 이 절[=제3절]에서가 아니라 앞선 절[=제2절]에 담겨져 있다. 이 절에서는 단지 경험에 의한 그 판결의 확정을 제시하고 있을 뿐이다. 그러나 경험은 [...] 악의 뿌리를 결코 들추어 낼 수 없다.”(VI 39 각주)

우선 위의 각주는 a) 기대되었던 텍스트에서 보편적 악의 증명이 전혀 명확하게 제시되고 있지 않다는 인상과 b) 많은 생생한 사례들로 형식적이고 딱딱한 증명을 대신할 수 있다고 한 제3절의 언급과 상치된다. 오히려 위의 각주는 위에서 언급한 혼란을 불식시키는 다음의 두 가지 분명한 주장을 담고 있다. 즉,

- 악의 보편성은 결코 경험적으로 정당화될 수 없으며,
- 악의 보편성에 대한 증명 또는 그것을 증명하는 결정적인 논증은 이미 제2절에서 제시되었다.

특히 후자는 이미 제2절에서 악의 보편성에 대한 증명이 이루어졌

다는 제3절 첫째 단락의 인상과 일치할 뿐만 아니라 증명의 재구성과 관련해서도 결정적인 단초를 제공하고 있다. 따라서 증명의 재구성은 제2절에 대한 면밀한 내용분석을 통해 비록 칸트 스스로는 제시했다고 말하지만 그리 명확하게 드러나 있지 않은 증명을 명료하게 밝혀내는 데에 초점을 맞추어져야 할 것으로 보인다.⁸⁾ 이에 대한 본격적인 논의에 앞서서 먼저 칸트가 말하는 악에 대한 개념적 이해를 살펴보는 것은 상당히 유익할 듯하다.

II. 악에 대한 개념적 이해

1. 도덕적 악과 책임성

칸트에 따르면 악은 도덕적 위법성(Gesetzwidrigkeit)을 뜻한다. 위법성은 적법성을 전제할 때에만 가능하므로, 악은 개념적으로 도덕적 적법성을 함축하고 있다. 칸트에게 도덕적 적법성의 기준이란 그 자체로 선함을 확정짓는 도덕법칙인 정언명법이다. 따라서 도덕적 위법성으로서 악은 도덕법칙을 부정하면서 선에 대척하는 절대적 개념으로 이해될 수 있다.⁹⁾

그렇다면 악하다는 도덕적 평가의 근거는 무엇인가? 칸트는 우리가 어떤 사람을 악하다고 하는 평가의 초점을 단지 그가 위법적인 행위를 했다는 결과보다는 그러한 행위를 “의식적으로”(mit Bewußtsein, VI 20) 했다는 동기에 맞춘다. 즉, 선악의 평가는 단순히 행위자의 위법적 행위와 관련된 것이 아니라 그 행위로부터 추론되는 행위자의 주관적 원칙인 준칙과 관련된다. 이러한 맥락에서 악이란 행위자가 악한 준칙에 따라 고의적으로 위법을 행했다는 것을 의미하며 그것의 근거가 되

8) 이와는 대조적으로 Th. Buchheim(2001, 663쪽)과 J. Bojanowski(2006, 270-271쪽)은 증명이 포함된 곳을 제3절로 본다. 그러나 이럴 경우에 증명은 제3절에 담긴 내용으로 인해 다분히 경험적인 특성을 지닐 수밖에 없다.

9) H.-J. Ehm(2006), 32-49쪽 참조.

는 행위자의 주관적 근거인 준칙과 결부된 가치개념임을 알 수 있다.

그럼 칸트는 왜 악의 개념을 행위가 아닌 준칙이라는 행위자의 주관적 근거와 연결시키려 하는가? 이는 도덕성에 대한 그의 이해와 불가분의 관계에 있다.¹⁰⁾ 그에 따르면 인간은 이성적 존재이며 또한 자신의 이성에 따라 스스로에게 행위의 보편타당한 법칙(도덕법칙)을 제정할 수 있는 자율적 존재이다. 이처럼 인간은 자신의 자율에 근거해서 행위할 수 있으므로 그에 대한 전적인 책임을 가지게 된다. 하지만 선한 행위의 충분한 동기로서 도덕법칙을 채택하느냐 마느냐는 또한 각자의 자유로운 결정에 따른다. 다시 말해 어떤 사람이 행위의 충분한 동기로 도덕법칙을 채택하지 않는 이유는 그에게 이 법칙이 애초부터 부재하기 때문이 아니다. 오히려 인간이면 누구나가 선함 그 자체를 확정짓는 도덕법칙을 행위의 동기로 언제나 의식하고 있다. 행위의 충분한 동기로서 도덕법칙의 채택여부는 오로지 각자의 선택의 지(Willkür)의 자유에 달려있다. 결국 선악의 문제에서 칸트가 행위보다도 준칙에 초점을 맞추는 근본적 이유는 행위의 주관적 근거가 되는 준칙의 채택이 각자의 자유로운 선택의지에 따른 것이며 바로 그렇기 때문에 각자에게 행위에 대한 **책임귀속**도 가능해지기 때문이다.

이와 같은 이해는 인간이 선하면서 동시에 악하다는 양립론이나 선하지도 악하지도 않다는 무차별론을 애초부터 봉쇄시키는 효과를 가진다. 칸트에 따르면 도덕적 행위의 의욕에는 동기가 필요한데, 이는 행위자가 준칙을 채택하는 것을 전제로 한다. 이때 행위자는 다수의 특수한 준칙들을 가지지만, 이것들은 결국 보편타당한 최상의 준칙에 종속된다. 인간이 도덕적 행위와 관련해서 가지는 최상의 준칙은 유일한 것이므로 그것은 이성이 명하는 도덕법칙으로서 선하거나 또는 그 반대로서 악한 것 중 하나일 수밖에 없다. 따라서 인간은 선하지도 악하지도 않다는 주장은 성립하지 않는다. 이는 선하면서도 동시에 악하다는 주장에도 똑같이 해당된다. 왜냐하면 최상의 준칙으로서 도덕법칙인 선한 준칙과 그것의 부정인 악한 준칙은 각기 보편

10) VI 3 참조.

타당성을 가지는데, 두 개의 보편타당한 준칙들을 동시에 최상의 준칙으로 삼는다면, 결국 그 어느 것도 보편타당하지 못하다는 모순에 빠지기 때문이다. 달리 말해 최상의 준칙으로서 도덕법칙은 그 자체 보편적이지만 또 다른 최상의 준칙으로서 악한 준칙과의 관계에서 보면 단지 특수한 준칙에 불과하기 때문이다. 따라서 선한 준칙과 악한 준칙 사이에서의 채택에서는 원칙적으로 양자택일의 가능성 이외에 다른 어떤 대안도 있을 수 없게 된다. 이런 이유에서 칸트는 인간은 선하기도 하고 악하기도 하다는 ‘절충주의’(Synkretismus)나 선하지도 악하지도 않다는 도덕적 ‘무차별주의’(Indifferentismus)를 거부하고 인간은 선하거나 아니면 악하다는 ‘엄격주의’(Rigorismus)를 견지한다.¹¹⁾

정리해 보면, 칸트는 인간이 자유로운 의지와 그에 따르는 도덕적 책임능력을 가졌다는 근본적인 이해에서 출발하면서 악하다는 가치판단의 준거를 행위 그 자체보다는 그로부터 추론 가능한 “행위자의 주관 안에 모든 도덕적으로 악한 특수준칙들의 보편적 근거”(VI 20)가 되는 최상의 준칙에서 찾는다.

2. 인간의 본성과 자유

행위자의 의욕을 규정하는 것은 동기이고, 이는 “행위자의 주관적 원칙”(IV 421 각주)인 준칙으로 채택되는 한에서만 비로소 작용할 수 있다. 이때 준칙의 채택은 인과법칙에 지배하는 자연의 합법칙성에 따른 것이 아니며 인간의 육체적 또는 정신적인 유전과도 무관하다.¹²⁾ 이러 저러한 동기들 중 어떤 것을 최상의 준칙으로 채택하는 것은 오로지 선택의지의 자유에 기초한다. 물론 이미 언급했다시피 최상의 준칙을 자유로이 채택함에 있어서 행위자는 그가 이성적 존재인 한에서 도덕법칙을 동기로서 분명히 의식하지만,¹³⁾ 그에 대한 분

11) VI 22 참조.

12) VI 40 각주 참조. 이는 유전적 결정론과 마찬가지로 자유 및 도덕적 책임성과 결코 양립할 수 없는 기독교적 원죄설에 대한 칸트의 거부를 강력히 시사하고 있다.

명한 의식이 곧 그것의 필연적 채택을 보장하지는 않는다. 무엇을 최상의 준칙으로 채택하느냐 하는 것은 선택의지의 자유에 달려있기 때문이다. 다시 말해 모든 준칙의 채택과 관련한 최초의 주관적 근거는 선택의지의 자유로운 결정행위, 즉 “절대적 자발성(자유)”(VI 24)이며, 바로 이것이 모든 다른 특수한 준칙의 채택에 바탕이 되는 자유사용의 주관적 근거이기도 하다. 칸트는 이 자유행위를 “인간의 본성”(Natur des Menschen, VI 21)으로 이해한다.

“여기서 인간의 본성은 (객관적 도덕법칙 하에서) 단지 감각적으로 경험 가능한 모든 행위에 앞서는 자유사용 일반의 주관적 근거(der subjektive Grund des Gebrauchs der Freiheit überhaupt)를 의미한다. 그런데 이 근거가 어디에 있건 그것은 상관이 없다. 그렇지만 주관적 근거는 항상 그 자체 자유행위(ein Aktus der Freiheit)여야 한다.”(같은 곳)

여기서 ‘본성’으로 번역된 ‘Natur’는 칸트철학에서 대체로 ‘자유’(Freiheit)와 상반되는 개념인 ‘자연’을 뜻한다. 그런데 만일 인간의 본성이 자연의 영역에 귀속하는 것이라면, 준칙채택과 관련한 최초의 주관적 근거는 자연적 충동과 연결된다. 이럴 경우 우리는 사실상 어떠한 행위에 대해서도 책임을 물을 수 없게 되며, 선과 악이라는 평가 자체도 무의미하게 될 것이다. 왜냐하면 자연적 충동이란 인과적 사슬에 묶여 있으므로, 준칙채택의 자유란 사실상 자연적 원인에 의한 결과에 불과하기 때문이다. 따라서 여기서의 ‘본성’(Natur)은 경험세계인 자연이 아니라 오히려 그것을 초월한 세계로서 도덕세계와 연결시켜 이해되어야 한다. 그리고 이러한 의미에서 칸트는 인간을 본질적으로 “자유로이 행위하는 존재”(VI 37)로 이해하고 선택의지의 자유를 그의 본성으로 파악하고자 한다.

칸트가 선택의지의 절대적 자발성, 즉 자유행위를 인간의 본성으로 이해하는 데에는 몇 가지 중요한 이유가 있다. 첫 번째 이유는 무엇

13) 이는 『도덕형이상학의 정초』나 『실천이성비판』에서 다룬 내용을 이미 증명된 것으로 간주할 때 가능하다. 실제로 칸트는 바로 이러한 전제에서 악의 문제를 논의하고 있다.

보다도 인간의 본성과 도덕성의 연결 가능성이다. 칸트가 여기서 말하는 인간의 본성은 필연적 인과법칙이 지배하는 경험적 차원이 아닌 **예지적 차원**에 속하는 것이므로 선택의지의 자유도 예지적 행위에 속한다. 이러한 의미에서 선택의지가 자유로이 준칙을 채택하기 때문에 각 자에게 행위에 대한 책임귀속이 가능해진다.

두 번째 이유는 자유가 어떻게 가능한가는 물음과 관련이 있다. 칸트는 이미 『도덕형이상학의 정초』나 『실천이성비판』에서 자유 그 자체의 가능성에 대한 설명은 불가능하다고 말한 바 있다.¹⁴⁾ 그에 따르면 비록 ‘초월적 자유’(transzendente Freiheit)의 가능성에 대해서는 모순 없이 사유할 수 있지만, 자유 그 자체의 가능성을 설명한다는 것은 인간의 능력을 넘어선다. 그래서 우리는 준칙선택과 관련한 최초의 주관적 근거에 대하여 더 이상 캐물 수 없으며, 그것은 “우리에게 불가해한”(uns unerforschlich, VI 21) 것으로 남는다. 하지만 선택의지의 자유를 인간의 본성으로 이해한다면, 적어도 그 자유의 근거에 대해 끝없이 캐물은 악무한은 피할 수 있게 된다.

세 번째 이유는 선택의지의 절대적 자발성(자유)을 예지적 차원에서 인간본성으로 이해함으로써 이 자유가 단지 특정 개인에게만 해당하는 것이 아니라 인류 전체에게 예외 없이 속하는 것임을 밝힐 수 있다는 점이다. 즉,

“우리가 인간이 본성상 선하거나 악하다고 말하면, 그것은 그가 선한 또는 악한 (위법적) 준칙을 채택하는 (우리에게 불가해한) 최초의 근거를 가지고 있음을 의미한다. 더군다나 이는 인간에게 보편적이어서 인간은 선하거나 악한 준칙의 채택을 통해 인류 전체의 특성(den Charakter seiner Gattung)을 표현한다.”(VI 21)

끝으로 준칙선택과 관련한 최초의 주관적 근거인 선택의지의 자유를 인간본성으로 이해하는 것은 악의 문제가 인간본성과 밀접하게 관련되어 있음을 암시한다. 이미 언급했듯이, 악의 근거는 행위에 관한

14) IV 468-469쪽과 V 46쪽 참조.

행위자의 주관적 원칙인 준칙에 있다. 특수한 준칙들뿐만 아니라 그것들이 최종적으로 근거하는 최상의 준칙은 선택의지에 의해 자유로이 채택되므로 악의 근거는 준칙채택의 주관적 근거가 되는 선택의지의 자유에서 찾아야 할 것이다.¹⁵⁾ 이때 자유는 인간의 본성에 해당하므로 악은 결국 인간의 본성과 깊은 관계에 있다고 볼 수 있다.

지금까지의 논의를 간략히 정리해 보면, 칸트는 악의 문제를 도덕의 문제로 이해하면서 이를 인간에 대한 근본적인 이해, 즉 인간의 본성인 자유로부터 접근한다. 그에 따르면 도덕적 악이란 행위에 대한 최상의 준칙과 불가분의 관계가 있는데, 이 준칙의 채택은 선택의지의 자유에 기초하므로 악은 오로지 인간 자신에게 책임이 있다. 또한 이 자유행위는 인간의 본성으로서 모든 인간에게 예외 없이 속하므로, 만일 인간이 악하다면, 그것은 모든 인간에게 예외 없이 해당하는 보편진술로 이해되어야 한다.

그러나 도덕적 악이 최상의 행위준칙과 관계있다는 점, 준칙채택은 선택의지의 자유에 기초한다는 점 그리고 자유는 인간의 본성이라는 점으로부터 모든 인간이 본성상 선한지 또는 악한지에 대한 어떠한 결론도 도출되지는 않는다. 이에 대한 논의의 실마리로서 칸트가 제시한 ‘인간학적 탐구’의 내용을 살펴보도록 하자.

III. 도덕성의 관점에서 살펴 본 인간 본성의 구조

1. 선의의 소질

칸트는 인간의 마음 또는 영혼능력을 크게 인식능력, 쾌와 불쾌의 감정 그리고 의욕능력의 자발적인 능력으로 이루어져 있다고 보고¹⁶⁾ “인간학적 탐구”도 이러한 능력들에 상응하여 진행시키고자 한다. 다만 여기서는 탐구범위를 도덕적 행위의 근거를 이루는 “의욕능력과

15) VI 35와 37을 참조.

16) V 196쪽 이하 참조.

선택의지의 사용”(VI 28)과 관련된 것으로 한정시켜 선예의 소질에 관한 탐구를 진행한다. 이러한 틀 속에서 그는 본격적인 논의를 소질의 정의로부터 시작한다.

“우리는 어떤 존재자의 소질을 그러한 것으로 존재하기 위해 필히 요구되는 구성요소일 뿐 아니라 그들 사이의 결합형식들로 이해한다. 소질은 그것이 어떠한 존재자의 가능성에 필연적으로 속해 있는 한 **근원적(ursprünglich)**이다.” (VI 28)

이에 따르면 선예의 소질은 인간을 규정하는 근원적인 요소들로서 인간에 관한 본질적인 규정을 담고 있다. 다시 말해 인간 각자가 비로소 인간의 보편성을 지닌 존재자일 수 있는 규정적 근거가 바로 소질에 있다. 특히 인간의 의욕능력과 관련해서 소질은 보편적 인간의 가능성을 실현하기 위한 필연적인 욕구를 지시한다고 할 수 있다. 그리고 이 욕구의 충족이 바로 보편적 인간 가능성의 실현이기 때문에, 소질은 다름 아닌 인간행위의 실천적 목적을 지시할 뿐만 아니라 그 목적의 실현을 위한 행위의 요청도 함의하고 있다고 할 수 있다. 이러한 맥락에서 소질은 보편적 인간의 실현이라는 실천적 목적을 위해 필히 요청되는 행위의 동기वाद 밀접하게 결부됨을 알 수 있다.¹⁷⁾ 그리고 칸트에 따르면 선예의 소질은 크게 “세 종류”(Klassen, VI 26)로 이루어진다.

—첫 번째 선예의 소질은 생명체로서 가지는 인간의 “동물성”(Tierheit)이다. 이것은 이성능력이 요구되지 않는 “기계적인 자기애”(Selbstliebe, 같은 곳)인데, 자기보존, 종족번식과 보존 그리고 다른 인간들과의 공동생활을 실천적 목적으로 지시한다. 이러한 목적으로부터 벗어난 것은 “자연의 악덕”(VI 27)으로 불린다.

—두 번째 선예의 소질에는 인간이 생명체이면서 동시에 이성적 존재로서 지니는 “인간성”(Menschheit)이 해당된다. 이것은 이성이 요구되는 “비교하는 자기애”인데, 타자와의 비교를 통해 자신에게 가치를

17) 이에 관한 논의는 Hans-Jörg Ehn(2006), 70-71쪽 참조.

부여할 때 평등성의 가치를 지향하도록 하는 실천적 목적을 지시한다. 이로부터 벗어나는 것은 “문화의 악덕”(같은 곳)이 된다.

– 마지막 소질은 이성적이면서도 동시에 책임능력이 있는 존재로서 가지는 인간의 “인격성”(Persönlichkeit)이다. 이는 인간에게 “그 자체가 선택의지의 충분한 동기가 되는 도덕법칙에 대한 존중심의 수용”(같은 곳)을 실천적 목적으로 부여한다.

이처럼 소질은 각각 인간의 필연적이고 보편적 욕구에 상응하는 행위의 실천적 목적을 지시할 뿐만 아니라 그 목적의 달성을 의욕하는 동기와 결부되어 있으므로 각 소질은 의지의 동기로 환치가 가능하다. 이에 따르면 모든 인간은 기본적으로 ‘자기애’와 ‘도덕법칙에의 존중’이라는 두 동기를 가진다고 할 수 있다. 그리고 보편적 인간의 실현가능성과 관련해서 각 소질이 부여하는 실천적 목적들을 고려할 때, 동물성, 인간성 그리고 인격성이라는 선예의 소질들이 모두 근원적이고 필연적이어서 그 어떤 것도 배제되어서는 안 되는 것과 마찬가지로 자기애와 도덕법칙에의 존중이라는 두 동기들도 모두 근원적이고 필연적인 것으로 간주되어야 한다. 다시 말해 도덕법칙에의 존중만을 의지의 유일한 동기로 내세우면서 자기애를 전적으로 배제한다거나 도덕법칙을 무시한 채 자기애만을 유일한 의지의 동기로 간주되어서는 안 된다. 오히려 인간은 자기애와 도덕법칙에의 존중을 모두 의지의 동기로 가질 수밖에 없다. 따라서 문제는 어떤 동기를 배제하느냐가 아니라 두 동기들 사이의 관계를 어떤 식으로 규정하는가에 있다.¹⁸⁾ 이 관계의 규정은, 앞서 살펴보았듯이, 전적으로 행위자의 자유로운 선택의지에 달려 있다.

2. 악예의 성향

칸트에 따르면 성향(Hang)이란 “인간성에 대하여 전적으로 우연적인 경향성(습관적 욕망)을 가능하게 하는 주관적 근거”(VI 28)를 말한

18) VI 36 참조.

다. 다시 말해 성향이란 대상을 표상하거나 경험하기 이전에 “선행하는 어떤 일정한 욕망 생성에 대한 주관적 가능성”을 말하는데, 이는 주관이 대상을 경험하면서 습관적으로 그것에 대한 욕망을 좇는 경향성(Neigung)의 발생근거가 된다.¹⁹⁾

성향은 모든 인간에게 “타고난(angeboren)”(VI 29) 것이라는 점에서 일견 소질과 별다른 구분이 없어 보인다.²⁰⁾ 하지만 성향과 소질은 모두 인간의 본성에 속한다는 공통점 이외에 양자 간에 분명한 차이가 있다. 소질은 각 개인이 비로소 인간일 수 있는 가능성을 규정하고 있다는 점에서 그것이 지시하는 실천적 목적 또는 행위의 동기 그 자체에 대해서 인간은 어떠한 자유로운 선택권이나 결정권도 갖지 못한다. 이에 반해 성향 그 자체에는 앞으로 생성될 경향성의 방향과 관련해서 전혀 결정된 바가 없다. 오히려 어떤 경향성이 생성될 것인지 전적으로 각자가 그 주관적 근거가 되는 성향을 어떻게 규정해 나가는가에 달려 있다. 이를 칸트는 다음과 같이 표현하고 있다.

“성향은 다음과 같은 점에서 소질과 구별된다. 즉 그것은 비록 타고난 것이지만 그 자체로 표상되어 질 수 없고, 오히려 그것은 (만약 선하다면) 획득된(erworben) 것이거나 (만약 악하다면) 초래된(zugezogen) 것으로 여겨져야 한다.”(같은 곳)

이처럼 선한 성향의 획득이나 악한 성향의 초래는 “자유로운 선택 의지의 규정으로서만 가능”(같은 쪽)하다. 이는 무엇보다도 소질과는 달리 성향이 선택의지의 자유와 보다 직접적인 관련성을 가진다는 것을 암시한다. 달리 표현하자면, 모든 인간에게 보편적이고 타고난 성향은 **열린 구조**를 가진다고 할 수 있다. 왜냐하면 인간의 가능성을 규정하는 소질이 지시하고 있는 실천적 목적의 실현여부는 각자의 자유로운 의지에 달려있기 때문이다. 반면에 인간에게 행위의 실천적 목적과 동기를 지시하는 소질은 근원적이고 보편적이지만, 그 자체

19) 성향에 대한 정의와 관련해서는 VII 266을 참조.

20) VI 22 참조.

인간의 자유로운 결정이나 선택에 의존적이지 않다. 이러한 의미에서 성향과는 달리 소질은 선택의지의 자유와는 보다 덜 직접적인 관련성을 갖는 **단한 구조**를 가진다고 할 수 있다. 결국 성향이 선택의지의 자유와 보다 직접적인 관계에 놓여 있다면, 선에의 소질은 오로지 성향의 매개를 통해서 자유와 관계한다고 할 수 있다.

같은 맥락에서 ‘성향은 그 자체로 표상될 수 없다’는 위 인용문의 구절도 이해될 수 있다. 소질이 주어진 인간의 의욕능력에 상응하는 행위의 실천적 목적이라는 보편적 구조로 표상될 수 있는 것과는 달리, 성향은 자유행위와 결부되어 있으므로 그것의 구조는 결국 자유로운 선택에 따른 우연적 결과의 모습만을 지닐 뿐이다. 따라서 소질에서처럼 성향 그 자체를 표상한다는 것은 가능하지 않다.

다만 한 가지 주의해야 할 점이 있다. 선택의지의 자유와 보다 직접적인 관계를 맺음으로써 갖는 성향의 열린 구조는 결코 임의적이지 않다는 사실이다. 앞서 선에의 소질에 관한 논의에서 보았듯이, 인간은 누구나가 예외 없이 행위와 관련한 의지의 동기로서 자기애와 도덕법칙에의 존중을 가지게 되며 양자 중 어느 것도 배제할 수 없다. 문제는 선택의지가 양자 사이의 관계를 어떤 질서 속에 규정하는가에 있다. 양자의 관계에서 도덕법칙의 존중이라는 동기가 언제나 우선성을 확보한다면 도덕적 선이 획득될 것이지만, 그렇지 못할 경우에는 도덕적 악이 초래될 것이다. 이러한 의미에서 성향의 열린 구조란 바로 양자 사이의 결정되지 않은 관계 틀을 의미한다. 여기에서 칸트는 도덕적 악을 초래하는 양자 사이의 관계규정을 ‘악에의 성향’과 관련 짓고 이를 “세 단계”(Stufen, VI 29)로 구분한다.

— 그 첫 단계에는 “인간본성의 유약성(Gebrechlichkeit)”(같은 곳)이 관계된다. 이는 인간이 비록 도덕법칙을 객관적 동기로 인식하여 준칙으로 채택하기는 하지만, 그것을 따르기에는 충분히 강하지 못할 뿐만 아니라 경향성과 비교해서 그것을 너무 약한 동기라고 여기는 경우이다.

— 두 번째 단계는 인간 심정²¹⁾의 “불순성”(Unlauterkeit, VI 30)이

다. 이는 인간이 도덕법칙을 준칙으로 채택하고 그에 따라 실천하지만 그것 자체만을 충분한 동기로 삼지 않고 언제나 다른 비도덕적 동기를 필요로 하는 경우이다. 다시 말해 “의무적 행위가 순수하게 의무로서 행해지지 않는”(같은 곳) 경우이다.

-마지막 단계는 인간 심정의 “부패성”(Verderbtheit) 또는 “사악성”(Bösartigkeit, 같은 곳)이다. 이는 인간이 그 자체만으로 충분한 동기인 도덕법칙보다는 다른 비도덕적 동기를 최상의 준칙으로 삼는 경우이다. 즉, 다수의 동기들 사이에 마땅히 세워져야 할 도덕적 질서가 전도된 경우이다.

이와 같은 논의에서 우리는 악에의 성향이 자유로운 선택의지와 깊은 관계가 있음을 알 수 있다. 왜냐하면 보편적 인간의 가능성을 구성하는 선에의 소질에 따르면, 인간은 행위의 기본적인 동기로서 자기애와 도덕법칙에의 존중을 가지는데, 선악의 도덕적 평가는 이들 중 어떤 것을 최상의 준칙으로 채택하는가 하는 자유로운 선택에 달려 있기 때문이다.

한편 칸트는 악에의 성향에 대한 소개에 곧이어 악의 보편성에 대한 증명과 관련한 중요한 언급한다. 그에 따르면 “악에의 성향이 인간적 본성과 하나로 엮여져 있다는 것”을 증명한다면, 인간은 예외 없이 본성상 악하다는 것을 밝힐 수 있다(VI 30). 이로부터 우리는 칸트가 바로 뒤이은 텍스트부분(VI 30 이하)에서 보편적 악에 관한 증명을 제시할 것이라고 기대할 수 있다. 특히 이 기대는, 이미 서두에서 말한 것처럼, 제3절의 첫 단락이 주는 인상과 함께 같은 절의 마지막 각주에서 명시적으로 언급된 바를 통해 충분히 뒷받침된다. 따라서 문제는 칸트가 문제의 텍스트부분에서 정확하게 어떤 식으로 악의 보편성을 증명하고 있는가를 밝혀내는 것이다. 이제 이를 본격적으로 살펴보도록 하자.

21) 여기서 심정이란 “자연적인 성향에서 발생하는 것으로서 도덕법칙을 준칙으로 채택하는 선택의지의 능력 또는 무능력”(VI 29)을 일컫는다.

IV. 악의 보편성에 대한 칸트의 증명

1. 텍스트 분석과 증명의 재구성을 위한 단초

제2절에서 증명을 담고 있다고 여겨지는 텍스트부분은 두 단락으로 구성되어 있다. 첫 번째 단락에서 칸트는 진정 선한 사람이란 도덕법칙을 “글자대로”(den Buchstaben nach) 따르는 사람이 아니라 그 법칙을 “정신에 따라”(dem Geist nach, VI 30) 행하는 자임을 강조한다.²²⁾ 양자의 본질적인 차이는, 전자가 도덕법칙 이외에 다른 동기들을 필요로 하는 반면에, 후자는 항상 도덕법칙만을 “유일하고도 최상인 동기”(zur alleinigen und obersten Triebfeder, 같은 곳)²³⁾ 삼는다는 점에 있다. 달리 말해 선악의 평가기준은 결과로서의 행위가 아니라 오로지 그것의 동기인 준칙의 성격에 달려있으며, 행위의 동기로서 오직 도덕법칙만을 ‘유일하고 그 자체만으로 충분한 최상의 준칙’으로 삼는 자만이 선한 사람일 수 있다. 반면에 어떤 사람이 아무리 합법적이고 선한 행위를 할지라도 그것이 준하는 동기 자체가 도덕적 질서에 어긋나 있다면, 그는 악한 것이다. 왜냐하면 도덕법칙 이외에도 다른 동기에 따른 행위는 이미 그 자체 위법성의 가능성을 담고 있으며, 그것이 도덕법칙에 일치하는 것은 순전한 우연에 불과하기 때문이다. 또한 도덕법칙만을 최상의 유일한 준칙으로 선택할 수 있는 것만큼이나 그렇지 못한 선택을 하는 것은 순전히 행위자 자신의 자유로운 선택에 따른 것이므로 그의 책임에 속하기 때문이다.

이어지는 단락에서 칸트는 행위자의 자유로운 준칙선택이 성향과 어떤 연관성을 가지는지를 밝히고자 한다. 이를 위해 그는 먼저 성향을 크게 “자연적 존재로서 인간의 선택의지에 속하는” 물리적 성향과 “도덕적 존재로서 인간의 선택의지에 속하는” 도덕적 성향으로 구분

22) 이는 도덕적으로 선한 행위란 “의무에 맞게”(pflichtmäßig) 한 행위가 아니라 오로지 “의무로부터”(aus Pflicht) 한 행위라고 한 『도덕형이상학 정초』의 주장과 맥을 같이한다. IV 397-399을 참조.

23) VI 37을 참조.

한다(VI 31). 전자는 감각적 충동에 기초하여 인과적 자연법칙, 즉 자연적 필연성에 종속되므로 자유로운 선택의지와는 전혀 상관이 없다.²⁴⁾ 이에 반해 도덕적 성향은 행위준칙의 자유로운 선택의지와 불가분의 관계에 있으므로 인간의 책임영역에 귀속한다. 이런 이유로 도덕적 성향에만 “선택의지의 도덕적 능력에 부착될 수 있다.”²⁵⁾

그런데 책임귀속성은 실상 우리 자신의 행위(Tat)²⁶⁾와 관련되어 있는 반면, 성향은 “선택의지를 규정하는 주관적 근거”(같은 곳)로서 모든 행위에 선행하는 것으로 행위 그 자체는 아니다. 그렇다면 성향의 개념에는 준칙선택에 있어 자유로운 선택의지와 결부되어 있으므로 책임귀속성과 관련되면서도 동시에 행위 그 자체는 아니므로 책임귀속성과 무관하다는 모순이 발생하는 것처럼 보인다.

이러한 이유에서 칸트는 행위(Tat)라는 표현에 관한 좀 더 세밀한 의미구분을 덧붙인다. 그에 따르면 행위(Tat)는 최상의 준칙을 선택의지로 수용하는 자유의 사용뿐만 아니라 이 준칙에 따라 실제로 행하는 자유의 사용을 함께 뜻한다. 전자의 경우에서 행위는 시간적 제약을 넘어서 오직 이성을 통해서만 인식될 수 있는 “예지적 행위”(intelligibele Tat, 같은 곳)를 뜻하고, 후자의 경우에서 행위는 시간 속에 행해지는 ‘경험적이고 감각적임’을 뜻한다. 이러한 구분에 따르면 선택의지를 규정하는 주관적 근거인 성향이 결부되어 있는 선택의지의 자유행위는 바로 예지적 행위에 해당하며, 바로 그렇기 때문에 악의 성향은 경험적 차원에서의 모든 위법적 행위에 대한 “형식적 근거”(den formalen Grund, 같은 곳)가 된다. 이로써 앞서 성향의 개념에 잠재된 듯이 보였던 모순은 일단 해소된다.

이상의 텍스트 분석에서 사실 악의 보편성에 대한 증명이 뚜렷하

24) 칸트 이전까지만 하더라도 물리적 성향과 관련된 악은 자연적 악에 해당하였다. 이에 반해 칸트는 자연적 악을 악의 개념적 영역으로부터 철저하게 배제시키고 오로지 인간의 도덕적 책임영역 속에 한정시켜 규정한다.

25) “부착되어 있다”는 “ankleben”의 번역이다. 같은 의미로 칸트는 “깊이 뿌리내림”(tiefe Einwurzelung, VI 35)이라는 표현을 사용하고 있다.

26) Handlung이 결과적 의미에서의 행위를 뜻한다면, Tat는 Handlung의 원조(Urheber)가 되는 차원에서의 행위를 뜻한다. VI 227 참조.

게 제시되고 있다고 단언하기는 어려워 보인다. 그럼에도 여기에는 증명의 재구성을 위한 중요한 단초가 제시되고 있다. 먼저 악의 보편성은 “악에의 성향이 인간의 본성과 하나로 엮어져 있다는 것”을 증명하는 것에 있다는 칸트의 언급과 함께 위의 텍스트 분석을 통해 드러나는 다음의 사실에 주목해보자. 즉, 여기에서 모든 위법적 행위의 형식적 근거인 성향으로 이해된 ‘예지적 행위’와 앞서 모든 자유사용의 주관적 근거이면서 그 자체 자유행위로 이해된 ‘인간의 본성’이 동질적 영역인 초경험적 영역에 속한다는 점이다. 성향의 예지적 행위와 인간의 본성인 자유행위가 동일한 영역에 속한다는 사실로부터 우리는 **성향이 자유행위가 본질인 ‘인간의 본성’과 하나로 엮어져 있음**을 추론할 수 있다. 이로부터 성향은 시간 속에서, 즉 경험적으로 획득될 수 있는 것이 아니라 인간에게는 결코 근절될 수 없는 “생득적인”(angeboren) 것으로도 자연스럽게 드러난다.²⁷⁾ 물론 생득적이란 표현은 상당히 애매모호한 것이 사실이다. 하지만 분명한 것은 이 표현이 인간의 출생이라는 생물학적 차원을 의미하기보다는 자유행위가 고유한 본질인 인간의 본성, 즉 실천적·도덕적 차원과 관련되어 있다는 사실이다. 바로 이러한 맥락에서 인간이면 누구나가 예외 없이 모든 경험적인 자유사용의 주관적 근거인 악에의 성향을 가지며, 따라서 악을 초래하는 것은 준칙선택과 관련한 자유의 사용을 본성으로 하는 인간 자신에게 있으므로 인간은 결코 악에 대한 책임성으로부터 자유로울 수 없게 된다.

이제 우리는 칸트가 말하는 보편적 악을 보다 정확히 이해할 수 있게 되었다. 악은 언제나 자유로부터 초래되므로 그 자체 우연적인 것이다. 하지만 그 자유사용의 주관적 근거가 되는 성향은 “인간성 그 자체와 하나로 엮어져 있으며 또한 그 곳에 뿌리박고”(mit der Menschheit selbst verwebt und darin gleichsam verwurzelt, VI 32) 있으므로 모든 인간에게 보편적이고 생득적이다. 바로 성향이 자유행위라는 인간의 고유한 본성과 불가분의 관계 속에 얽혀 있으므로 해

27) 이러한 의미에서 칸트는 악에의 성향을 “악에의 자연적 성향”(einen natürlichen Hang zum Bösen, VI 32)이라고 부르기도 한다. VII 29와 37도 참조.

서 인간이 본성상 악하다는 주장은 자연스럽게 도출될 수 있다.

2. 증명의 재구성

제2절의 텍스트 분석을 통해 살펴본 증명의 단초를 토대로 악의 보편성에 대한 칸트의 증명을 재구성해보도록 하자.

a) 인간에 대한 도덕적 평가는 그가 행위와 관련해서 어떤 동기를 충분하면서도 유일한 최상의 준칙으로 채택하는가와 전적으로 결부되는데, 오직 도덕법칙만을 그러한 준칙으로 삼고 행위할 때에만 선한 자가 된다.

b) 인간학적 탐구에 따르면 선의의 소질은 인간의 가능성과 관련한 행위의 실천적 목적을 지시하고 그 목적의 실현에 상응하는 행위의 동기로서 자기애와 도덕법칙에의 존중을 부여한다. 이러한 의미에서 행위의 실천적 목적을 지시하는 선의의 소질은 닫힌 구조를 가진다.

c) 인간 가능성의 실현여부는 위의 두 동기들 중 어느 것을 그 자체만으로 충분하고 유일한 최상의 준칙으로 채택할 것인가에 직결되는데, 이는 전적으로 선택의지의 자유에 달려있다.

d) 이때 자유로운 선택의지를 규정하는 주관적 근거는 악에의 성향이다. 이것은 최상의 준칙을 선택의지에로 수용하는 예지적 행위일 뿐만 아니라 모든 경험적 자유행위의 형식적 근거가 된다. 이러한 의미에서 악에의 성향은 선의의 소질과는 반대로 자유행위와 보다 직접적인 관계에 놓여 있는 열린 구조를 가진다.

e) 그런데 자유행위와 직접적인 관계 속에서 그 자체 예지적 행위로 이해된 성향은 그 개념적 규정에 있어서 모든 자유사용의 주관적 근거이자 절대적 자유행위로 이해된 인간의 본성과 동질적이다. 이처럼 악에의 성향은 생물학적 차원이 아니라 실천적·도덕적 차원에서 이해되는 인간의 본성과 하나로 엮어져 있으며, 그러한 한에서 성향은 모든 인간에게 생득적이고 보편적이다.

f) 그러므로 인간은 본성상, 즉 누구도 예외 없이 악하다.

이상과 같은 논증을 통해 칸트는 인간에게 악이 보편적이라는 자

신의 주장을 증명한다. 물론 악이 우리 자신의 자유로운 행위에 의해 초래된다고 할지라도, 우리는 “왜 그것이 우리의 내부에서 최고의 준칙을 부패시켰는가에 관한 원인(Ursache)을 제시할 수는 없다”(VI 32). 이는 우리가 왜 악에의 성향이 우리 본성의 근본 성격인가는 물론 왜 인간이 본성상 자유로이 선택하는 존재인지에 대한 원인에 대해 더 이상 캐물을 수 없는 것과 같다.

하지만 이상과 같은 칸트의 증명에도 불구하고 한 가지 의문은 여전히 해소되지 않고 남는다. 칸트의 언급을 좇아 악이 우리의 내부에서 최고의 준칙을 부패시킨 원인은 더 이상 탐구될 수 없다는 것을 일단 받아들인다 하더라도 ‘악이 우리의 내부에서 최고의 준칙을 부패시켰다’는²⁸⁾ 주장에 대해서는 해명이 필요한 듯하다. 왜냐하면 그와 관련해서 칸트가 ‘근본악’(radikales Böse)이라는 개념을 도입하는 데, 이 개념은 다음과 같은 심각한 문제점을 드러내기 때문이다. 즉, 한편으로는 예지적 행위로서 이해된 악에의 성향은 자유행위를 본질로 하는 인간의 본성과 결부되어 있으므로 모든 인간에게 예외 없이 생득적이다. 하지만 다른 한편으로 악이 이미 모든 준칙의 근거를 부패시켰다면, 사실 악의 초래는 필연적일 수밖에 없어 보인다. 따라서 악이 선택의지의 자유로부터 초래되는 한에서 인간 자신에게 그 책임이 전적으로 귀속되어야 하지만, 이미 모든 준칙의 근거가 그 뿌리에서부터 부패되어 있어서 악의 초래는 필연적일 수밖에 없다면, 인간에게 악의 책임을 물을 수 없게 된다. 간단히 말해 자유로워야 할 선택의지는 그 출발점에서부터 이미 악을 필연적으로 선택하는 것 이외에 다른 방도가 없어 보인다.²⁹⁾ 바로 이러한 이유에서 ‘악이 우리의 내부에서 최상의 준칙을 부패시켰다’라는 칸트의 언급은 반드시 해명되어야 할 필요가 있다.

28) 칸트는 같은 언급을 VI 37과 38에서 반복적으로 하고 있다.

29) Ch. Schulte는 이러한 모순성 때문에 칸트의 근본악 개념은 유지될 수 없다고 본다(1991, 83쪽). 이와 관련해서 M. Willaschek(1992), 157-1578쪽도 참조. 특히 R. Bernstein(2002)은 이러한 모순성과 관련해서 칸트가 “자기 자신과 전쟁을 벌이고 있다(at war with himself)”는 표현으로 꼬집는다(33쪽).

3. ‘근본악’ 개념의 모순성에 대한 해명

방금 서술한 근본악 개념의 모순성은 결코 벗어날 수 없는 딜레마에 빠진 것으로 보인다. 그러나 이미 앞서 언급한 몇 가지 사실들을 종합해서 살펴본다면, 이에 관한 해명이 전혀 불가능하다고만 단정 내릴 수 없어 보인다.

먼저 모든 인간은 예외 없이 본성상 선의의 소질을 가진다. 특히 인간의 의욕능력과 선택의지의 사용과 관련해서 소질은 보편적 인간의 구현이라는 실천적 목적을 지시하고 또한 그 실현을 위한 행위의 동기로서 자기애와 도덕법칙에의 존중을 부여한다. 이때 도덕적으로 선택할 수 있는 유일한 조건은 도덕법칙만을 충분하고 유일한 행위의 동기로 채택하는 것이다. 왜냐하면 도덕법칙만이 인간성의 이념, 즉 자유를 담보해줄 수 있기 때문이다. 또한 도덕적 동기인 도덕법칙에의 존중과는 달리 자기애는 비도덕적 동기이기 때문이다.

하지만 후자는 전자만큼이나 모든 인간에게 필연적이고 근원적인 것으로 그 자체 선하다. 그러므로 인간의 가능성을 실현하기 위한 동기를 선택할 때, 위의 두 동기들 중 어느 것도 배제되어서는 안 된다. 그렇다고 두 동기들을 모두 동시에 최상의 준칙으로 삼을 수도 없다. 소질이 지시하는 목적의 실현을 위해서는 두 동기들 중 하나를 최상의 준칙으로 채택하여 양자 사이의 우선성을 규정해야만 한다. 즉, 인간성의 이념 또는 인격성의 실현은 도덕법칙만을 이미 충분하고 유일한 행위의 동기로서 그리고 최상의 준칙으로 채택할 것을 요구한다. 하지만 보편적 인간의 가능성 전체에서 볼 때, 이 요구는 자기애라는 또 다른 동기와 필연적으로 공존할 수밖에 없다. 왜냐하면 인간은 도덕법칙을 행위의 보편적이고 필연적 동기로 언제나 의식하고 있지만 그것 못지않게 자기애도 보편적 인간의 가능성을 규정하는 근원적 동기로 의식하기 때문이다.³⁰⁾

30) 바로 이러한 배경에서 인간은 비록 도덕법칙을 객관적 동기로 의식하고 준칙으로도 채택하지만, 그것만으로는 결코 충분하지 못하다고 여기거나 그에 따른 실천을 위해서는 또 다른 동기가 필요하다고 여긴다. 이 보다

이러한 인간학적 조건이 시사하는 바는 인간성 이념의 실현을 위해 요청되는 동기 또는 최상의 준칙의 **순수성과 유일성**은 이미 훼손되어 있다는 사실이다.³¹⁾ 왜냐하면 인간성 이념의 실현을 위한 동기의 순수성과 유일성은 보편적 인간의 가능성 전체라는 관점에서 주어지는 보편적이고 근원적인 동기의 **복수성**과 필연적으로 공존하기 때문이다. 바로 이 점에서 칸트는 최상의 준칙이 그 뿌리에서 부패되어 있다고 말한다. 이를 좀 더 도식적으로 정리해 보자면,

a) 악에의 성향은 인간에게 보편적이고 생득적이다. 물론 여기서 생득적이란 자연적 필연성을 의미하는 것이 아니라 선택의지를 규정하는 주관적 근거로서의 보편성을 뜻한다.

b) 이 성향은 인간의 본성으로 이해된 자유, 즉 선택의지의 자유와 하나로 엮여져 있다. 이때 자유는 소질에 의해 부여된 보편적 동기들인 자기애와 도덕법칙에의 존중 중 하나를 최고의 준칙으로 선택하는 열린 구조를 가진다.

c) 인간성 이념의 실현은 선택의지가 도덕법칙만을 유일하고 충분한 행위의 동기로 채택할 때이다. 그런데 자기애도 도덕법칙에의 존중과 마찬가지로 동등한 존재적 정당성을 가진다. 이런 이유로 인해 인간성 이념의 실현을 위해 요청되는 동기 또는 최상 준칙의 순수성과 유일성은 이미 훼손되어 버린다.

d) 그러나 비록 도덕성의 관점에서 요구되는 동기 또는 최상 준칙의 순수성과 유일성은 이미 부패되었지만, 이로 인해 악의 초래가 필연적이지는 않다. 왜냐하면 악에의 성향은 열린 구조를 가지며, 그

더 진진된 단계에서는 심지어 도덕법칙보다는 자기애를 최상의 준칙으로 삼는 경우가 발생하게 된다.

31) 이 점은 칸트가 인간의 본성상 악함을 극복하는 유일한 길로서 내세운 ‘심성의 혁명’에서 충분히 읽을 수 있다. 그에 따르면 우리가 자신의 악함을 극복하는 길은 도덕법칙에의 존중을 다른 동기와 결부되어야 하거나 또는 그것에 종속되어야 하는 것이 아니라 오로지 그 자체만을 유일하고 충분한 의지규정의 동기로 채택하는 데에 있다. VI 46 참조. 즉 칸트는 “선에의 근원적인 소질의 복구”(die Wiederherstellung der ursprünglichen Anlage zum Guten)를 모든 준칙의 근거로서 선을 정화시키는 것(본문에는 “die Herstellung der **Reinigkeit** [des Guten]”)으로 이해한다(VI 46).

구조 속에서 동기를 선택하는 것은 결국 다름 아닌 선택의지의 자유에 있기 때문이다.

이러한 의미에서 악에의 성향은 보편적이고 생득적이지만, 그로 인해서 악의 초래는 결코 필연적이지 않다. 최상 준칙의 부패성에도 불구하고 인간은 전적으로 자신의 자유로운 의지에 따라 최상 준칙을 선택하므로 도덕적 책임성을 가지며 나아가서 도덕법칙만을 충분하고 유일한 행위의 동기로 선택할 수 있는 가능성을 여전히 가진다.

V. 나가는 말

칸트는 인간이 본성상 악하다는 주장을 전혀 증명하지 않았거나 경험적 사실들에 의존하고 있다는 해석은 오래 전부터 널리 퍼져 있다. 이러한 해석은 보편적 악에 대한 증거가 텍스트 속에서 그리 명확하게 제시되어 있지 않을 뿐만 아니라 그와 관련한 칸트의 언급들이 서로 모순적이라는 것과 무관하지 않다. 이에 대해 본 논문은 텍스트 분석을 통해 보편적 악의 증거를 재구성해냄으로써 기존의 해석이 타당하지 못할 뿐만 아니라 왜 모든 인간은 예외 없이 악한지를 칸트가 설득력 있게 제시하고 있음을 보였다. 또한 이를 통해 흔히 근본악의 개념과 관련해서 줄곧 지적되어 온 모순성, 악은 자유로부터 초래된다는 주장과 악은 인간에게 보편적이고 생득적이므로 필연적이라는 주장 사이의 모순도 칸트의 주장에 대한 증거가 불가능하다거나 또는 실패했다는 잘못된 해석에 기인한 오해에 불과함을 밝혔다.

투고일: 2009. 04. 16.
 심사완료일: 2009. 05. 18.
 게재확정일: 2009. 05. 20.

김화성
 건국대학교

참고문헌

- Allison, Henry E., *Kant's Theory of Freedom*, Cambridge University Press, 1990.
- Bernstein, Richard, *Radical Evil: A Philosophical Interrogation*, Polity, 2002.
- Bojanowski, Jochen, *Kants Theorie der Freiheit: Rekonstruktion und Rehabilitierung*, de Gruyter, 2006.
- Buchheim, Thomas, “Die Universität des Bösen nach Kants Religionschrift”, in: *Kant und die Berliner Aufklärung. Akten des IX. Internationalen Kant-Kongresses*, vol.3, de Gruyter, 2001.
- Ehni, Hans-Jörg, *Das moralische Böse: Überlegungen nach Kant und Ricoeur*, Alber, 2006.
- Frierson, Patrick R., *Freedom and Anthropology in Kant's moral Philosophy*, Cambridge University Press, 2003.
- Grimm, Stephen. R., “Kant's Argument for Radical Evil”, in: *European Journal of Philosophy*, 10(2002).
- Guyer, Paul, “Immanuel Kant”, in: Edward Craig (ed.), *The Routledge Encyclopedia of Philosophy*, Routledge, 1998.
- Kant, Immanuel, *Grundlegung der Metaphysik der Sitten*, Meiner, 1999.
- _____, *Kritik der praktischen Vernunft*, Meiner, 1990.
- _____, *Die Religion innerhalb der Grenzen der bloßen Vernunft*, Meiner, 2003.
- _____, *Die Metaphysik der Sitten*, Werkausgabe, Bd.8, Suhrkamp, 2005.
- _____, *Anthropologie in praktischer Hinsicht*, Meiner, 2003.

- Michalson, Gordon. E., *Fallen Freedom: Kant on Radical Evil and Moral Regeneration*, Cambridge University Press, 1990.
- Morgan, Seiriol, “The Missing Formal Proof of Humanity’s Radical Evil”, in: *The Philosophical Review* 114(2005).
- O’Conner, Daniel, “Good and Evil Disposition”, *Kant-Studien* 76(1985).
- Prauss, Gerold, *Kant über Freiheit als Autonomie*, Klostermann, 1983.
- Quinn, Phillip, “In Adam’s Fall, We Sinned All”, in: *Philosophical Topics* 16(1988).
- Schulte, Christoph, *radikal böse. Die Karriere des Bösen von Kant bis Nietzsche*. Fink, 21991.
- Stengeth, Bettina, “Kants schädliche Schriften. Eine Einleitung”, in: *Immanuel Kant, Religion innderhalb der Grenzen der bloßen Vernunft*, Meiner, 2003.
- Willaschek, Marcus, *Praktische Vernunft: Handlungstheorie und Moralbegründung bei Kant*, Metzler, 1992.
- Wimmer, Reiner, *Kants kritische Religionsphilosophie*, de Gruyter, 1990.
- Wood, Allen, *Kant’s Ethical Thought*, Cambridge University Press, 1999.

ABSTRACT

The Human Being Is Evil by Nature

—Kant's Argument for the Claim of Universal Evil—

Kim, Hwa-Sung

In *Religion within the Boundaries of Mere Reason* Kant held that human beings, as a species, are evil. But the text of *Religion* is convoluted and difficult, with inadequate exposition of central concepts and confusing articulation of important claims. For this reason it is generally said that Kant did not or could not justify his claim of universal evil at all. Furthermore the concept of radical evil outlined there is inconsistent. Unlike the critics, however, I will be attempting to provide a reconstruction of Kant's argument for the claim of universal evil using the text as a basis and propose a way explaining how the contradictory positions on the concept of the radical evil could be understood properly.

Keywords: Evil, Radical Evil, Universality, Morality, Free Will, Kant