

【논문】

칸트의 인간성 정식과 도덕적 자율성

오근창

【주제분류】 윤리학, 서양근대철학, 칸트

【주요어】 인간성 정식, 목적 그 자체, 목적 수립 능력, 선의지, 자율성

【요약문】 직관적으로 호소력이 있지만 상당히 모호하게 보이는 칸트의 인간성 정식(FH)을 어떻게 이해해야 하는지에 대한 많은 논의들이 존재한다. 이 글에서 나는 현존하는 견해들과 다른 칸트의 인간성 정식에 대한 대안적 독해를 제안하고자 한다. 이를 위해 나는 먼저 『윤리형이상학 정초』를 중심으로 칸트의 인간성 정식을 재구성한다(2장). 아래에서 나는 센센, 코스가드, 단에 의해 제안된 인간성 정식에 대한 여러 영향력 있는 해석들이 갖는 몇몇 결점들을 비판적으로 검토한다. 센센과 달리, 내가 보기에 인간성 정식은 단지 보편법 정식을 보다 직관적이게 만드는 소극적인 역할만을 하고 있지 않다(3장). 또한 코스가드의 인간성 정식에 대한 독해는 너무 얇은 인간성 관념을 가짐으로써 칸트의 원래 의도와 달리 목적 수립 능력 일반이 비도덕적인 행위마저 허용하게 한다는 점에서 문제가 있다(4장). 최근에 단은 인간성이 선의지와 등가적이라고 제안했지만, 나는 선의지 독해가 너무 두껍고 도덕주의적인 인간성 개념을 가정함으로써 인간성의 내용을 과도하게 제한하고 있다고 주장한다(5장). 마지막으로 이 글은 몇몇 유관한 문헌적, 철학적 이유들을 조명함으로써 인간성 정식의 핵심이 도덕적 자율성 개념에 있음을 보인다.

I. 들어가며

많은 이들이 인정하듯이 그간의 칸트 윤리학에 관한 해석들이 주로 초점을 맞춰왔던 것은 이른바 ‘형식주의적’인 칸트였던 것처럼 보인다. 단적으로 이 형식주의적 면모는 정언명법의 보편법 정식, “마치 네 행위의 준칙이 네 의지에 의해 보편적인 자연법칙이 되어야 할 것처럼 그렇게 행위하라”(G 421)¹⁾는 정식에서 잘 드러난다. 반면, 지금까지 정언명법을 나타내는 또 다른 정식인 인간성 정식에 대한 주목은 상대적으로 많은 주목을 받지 못했으며, 칸트 자신도 『실천이성비판』에서 이를 제대로 다루지 않았다. 형식주의와는 대비되는 인본주의적이고 질료적, 실질적인 칸트 윤리학의 면모는 인간성 정식에서 잘 드러난다. 인간을 존중하라, 인간을 단지 수단으로서뿐 아니라 목적으로서도 대하라는 칸트의 주장은 반론의 여지가 적은 보편타당한 도덕 원리로서 상당히 강한 호소력을 갖는다. 한 가지 특기할 만한 사실은 칸트가 인간성 정식을 논의할 때 단순히 인간이 아니라 인간성이라고 말한다는 점이다(G 429). 그렇다면 여기서 인간성을 존중한다는 것은 무슨 뜻인가? 또는 모든 인간은 인간이라는 점 하나만으로 존중할 만한 가치를 갖는가?

결국 칸트가 목적 그 자체로서의 인간성을 이야기할 때 우리에게 문제되는 것은 첫째로, 인간성이 가리키는 것이 무엇인지에 대한 정체 규명과 둘째로, 인간성 정식이 다른 정식들과 맺고 있는 관계이다. 이 글이 주로 다룰 것은 첫째 문제이지만, 간접적으로 두 번째 문제도 다룬다. 이하에서는 칸트 주석가들의 여러 유력한 해석을 살펴봄으로써 각 해석들이 갖는 장점과 문제점을 드러내고 이에 대해 평가함으로써 나 자신이 지지하는 해석을 드러내도록 하겠다. 먼저 위에서 제시한 질문들에 대답하기 위해서 나는 우선

1) 이하에서 칸트에 대한 인용은 괄호 안에 저작 약호와 더불어 나토르프(Paul Natorp)가 편집한 베를린 학술원판 전집(Immanuel Kant, *Kants gesammelte Schriften*, ed. Königlich Preußischen Akademie der Wissenschaften, Bd. I-XXII, Berlin, 1910ff)의 쪽수를 표기할 것이다. 각 저작의 약호는 다음과 같다. G: 『윤리형이상학 정초』, MM: 『윤리형이상학』, R: 『이성의 한계 내에서의 종교』, C2: 『실천이성비판』, C3: 『판단력 비판』.

『정초』에서 나타난 칸트의 인간성 논증을 재구성하며 소급 논증의 타당성을 검토할 것이다(2장). 그리고 인간성 정식에 대한 여러 가지 유력한 해석들, 셴센의 해석(3장), 코스가드의 해석(4장)과 딘의 해석(5장)을 중심으로 살펴보고, 이 해석들이 갖는 문제들을 제시할 것이다.²⁾ 나의 해석에 따르면 인간성 정식이 말하는 인간성은 단지 보편법 정식을 직관에 가깝도록 하는 부차적인 역할을 하는 것(셴센)이 아니며, 또한 단순히 목적 수립의 능력(코스가드)이나 선의지(딘)와 동일한 것이 아니다. 결국 인간성 정식이 말하는 인간성의 핵심은 도덕적 자율성 개념에 있음(6장)을 드러내려는 것이 이 글의 목적이다.

II. 인간성 정식 논증

칸트가 제시하는 인간성 정식은 크게 두 부분으로 나뉘는데 첫째는 무조건적인 목적이 있어야 한다는 점과 둘째로 그 목적은 인간성일 수밖에 없다는 점이다. 만약 정언명령이 있다면 그것은 질료로서 필연적 목적을 가질 수밖에 없고 이는 인간성에 해당한다. 먼저 인간성 정식을 칸트가 어떤 방식으로 도입하는지 살펴보자. 앞서 보편법 정식에 대한 검토를 마친 후 칸트는 “어떤 법칙의 표상에 맞게 행위하도록 자신을 규정하는 능력”(G 427)이라는 의지 개념에 관한 정의를 제시한 후, 이러한 의지의 자기규정의 객관적 근거가 되는 것이 목적이라고 정의한다. 칸트에게 모든 행위는 항상 어떤 목적을 가지기에(MM 385) 의지가 목적을 수반하는 것은 불가피하다. 단적으로 말해 의지는 원칙에 따라 행동하는 능력을 가리키기도 하지만, 또한 목적을 추구하는 것이다(C2 34, R 4). 칸트가 이어서 제기하는 물음은 과연 참된 도덕적 행위가 있다면 그것이 가져야 할 목적이 어떤 목적인가 하는 점이다. 칸트는 만약 정언명령의 근거가 될 수 있는 어떤 것이 있

2) 이 해석들의 대강은 이전에 페이튼에 의해 각각 보편법 정식으로부터의 논증, 이성적 행위자의 본성으로부터의 논증, 성격으로서 선의지로부터의 논증으로 분류된 바 있다 (Paton 176-78).

다면, 그것은 목적 그 자체 또는 최종목적이라고 주장한다. 어떤 목적이 특수한 개인의 경향성에만 의존하여 상대적인 것이라면 이는 단지 가언 명법의 목적에 불과하다. 따라서 도덕적 행위의 목적은 이런 것이 아닌 자립적인 것, 목적 자체이어야 한다. 칸트는 목적 자체가 될 수 있는 여러 가지 후보들을 검토한 후 인간성이 바로 그런 목적 자체라고 말한다. 그는 여기서 단지 조건적, 상대적, 수단의 가치만을 갖는 물건들과 무조건적, 절대적, 목적의 가치를 갖는 이성적인 존재인 인격들을 구별하며, 또 ‘목적의 나라’에 대한 짧은 논의 이후에 이러한 구별을 다시금 대체가능한 것의 가격과 가치를 초월한 존재가 갖는 존엄성으로 나눔으로써 되풀이한다(또한 MM 434-35). 이렇게 도출된 인간성 정식은 “**네 인격 안의 인간성(Menschheit) 뿐 아니라 모든 사람의 인격 안의 인간성까지 결코 단지 수단으로만 사용하지 말고, 언제나 동시에 목적으로도 사용하도록 그렇게 행위하라**”(G 429)이다. 오직 가언명령만 제공할 뿐인 주관적 목적과 달리 정언명령의 근거가 되는 객관적 목적은 목적 그 자체, 인간성에 해당한다. 즉 칸트가 말하는 인간성은 다른 어떤 것을 위해 좋은 도구적 선이 아니라 그 자체로 목적인 내생적(intrinsic)이고 최종적(final) 선이다.

인간성 정식에서 목적 자체를 인간성이라고 규정하는 칸트의 논증은 코스가드와 우드를 비롯한 여러 해석자들에 의해 어떤 소급(regress)의 방식 내지 소거의 방식으로 이루어진다고 여겨져 왔다(Korsgaard 1996a, Wood). 사람에 따라 상이한 주관적이고 상대적 목적에서 객관적인 목적, 목적 그 자체의 방향으로의 소급. 또한 주관적 목적, 경향성의 대상들, 경향성들, 비이성적 존재들(동물, 식물)의 소거. 이는 일단 수단 같은 것에서 보다 목적 같은 것으로, 조건적인 것에서 무조건적인 것으로의 이행을 드러낸다는 점에서, 큰 그림에서는 대강 옳은 지적이다. 그러나 이것이 과연 세밀한 측면에서도 옳은 것인지, 그러니까 경향성의 대상에서 시작하여 각각 경향성, 자연 존재나 사물, 우리 자신, 우리 자신 속의 인격으로 나아가는 논증의 방향이 정말 자의적이지 않은 소급의 방식으로 이루어지고 있는지에 대해서는 의문을 제기할 수 있다. 실상 소급이라고 부를 수 있는 것은 경향성의 대상들에서 경향성으로 이동할 때에만 일어났으며, 칸트는 곧바로 목

적 자체가 될 수 있는 한정된 후보들로서 비이성적 존재와 이성적 존재만을 차례로 검토하고 있을 뿐이기 때문이다. 여기서 인간성이 칸트의 논증에서 수행하는 역할은 무엇인가? 칸트는 인간성이 가치의 최종원천이라고 이미 전제하고 있는가? 칸트는 마치 소급의 무한퇴행을 목적 자체로서 인간성의 단계에서 임의적으로 중단시키는 듯하다.

뒤에서 더욱 자세히 살펴보겠지만, 코스가드는 칸트가 말하는 인간성을 목적 수립 능력이라고 부른다. 그런데 그녀가 재구성하는 칸트의 인간성 정식의 도출 과정이 과연 정당한 소급 논증인가? 가령 사물의 가치가 조건적이라는 점을 인정한다고 할지라도 여기에서 인간의 무조건적인 가치가 논리적으로 따라나오지는 않는다는 점을 지적할 수 있다. 사물의 가치가 상대적이거나 도구적 가치로 생각된다는 점으로부터 인간의 가치가 절대적이거나 내생적이라는 점은 필연적으로 나오지 않기 때문이다. 예컨대 칼이 유용한 것은 오직 그것이 인간에게 유용할 때라는 주장으로부터 인간이 절대적 가치를 갖는다는 점을 도출할 수는 없다(Sensen 105). 코스가드가 재구성한 칸트의 소급 논증은 가치의 원천을 추적하다보면 이성적 능력 자체가 객관적 가치를 가질 수밖에 없다고 가정하지만, 우리가 어떤 목적을 가치있는 것으로 여긴다는 사실로부터 목적의 원천을 추론하는 것은 문제가 있어 보인다. 마틴이 지적하듯이, 이를테면 y 에 의해 만들어진 x 가 속성 p 를 가졌다고 y 도 p 를 가져야 할 것으로 가정해야 할 것 같지 않다(Martin 111, Kerstein 59). 다시 말해서 목적 수립 능력이 목적의 조건임을 인정한다고 해도 왜 목적 수립 능력 자체를 객관적이고 무조건적으로 가치있는 것으로 간주해야 하는지 불확실하다. 이 점에서 모든 이성적 존재를 목적 자체로 만들 뿐 아니라 자연의 최종 목적을 형성하고 심지어 자연 전체도 정당화하는 것이 우리의 목적 수립 능력이라는 코스가드의 주장(Korsgaard 1996a: 130)이 얼마나 충분히 정당화된 것인지는 의문스럽다.

그러나 이러한 불확실성은 코스가드의 해석 자체의 문제라기보다는 그 이전에 자연의 궁극 목적이 인간성에 있다고 요청하는 칸트 자신의 목적론적 세계관(C3 431, 442-43)에서 기인하는 것이라고 볼 수 있기 때문에, 센센이나 마틴이 하듯이 인간성 정식을 이끌어내는 소급 논증 자체를 문제시

하는 것은 다소 자비롭지 못한 비판처럼 보인다. 코스가드는 이들이 생각하는 것처럼 목적 수립 능력이 목적을 생산한다거나 이전에 없던 속성을 부여한다고 주장하는 것이 아니라, 우리가 목적 수립 능력에 대해 취해야 하는 일종의 실천적 태도에 관한 초월론적 논증에 대해 이야기하는 것으로, 다시 말해 우리가 어떤 것을 가치있는 것으로 여긴다면 그것에 가치를 부여하는 자신의 목적 수립 능력 역시도 가치있는 것처럼 간주해야 한다는 것이다(특히 Korsgaard 1996b: 62). 여기서서는 일단 칸트의 인간성 정식 논증을 타당한 것으로 받아들이고 논의를 계속하도록 하자.

Ⅲ. 인간성 정식에 대한 해석에 앞선 예비적 고찰: 형식주의적 해석

칸트가 인간성을 목적 자체로 이끌어내는 논증이 가질 수 있는 잠재적인 목적론적인 함의를 벗겨내려는 해석들도 존재한다. 이를테면 인간성을 존중해야 하는 이유는 모든 인간이 갖는 절대적인 내적 가치이거나, 인간이 그 자체로 존엄성을 갖고 있기 때문이 아니라는 것이다(Paton 189). 여기서서는 일단 보편법 정식으로 인간성 정식이 환원될 수 있다고 보거나 인간성 정식이 실질적으로 정언명법에 관한 칸트의 논의에서 기여하는 바가 없다고 주장하는 해석을 형식주의라고 칭할 것이다. 이러한 형식주의적 해석에 따르면, 인간성 정식은 단지 정언명법을 사람들의 직관에 가깝게 하도록 도입된 것일 뿐이지 그 자체로 인간성의 실질적 가치를 주장하려는 것은 칸트의 애초의 의도가 아니다. 요컨대 이 해석이 지지하는 바는 인간성은 결코 가치의 최종토대, 최종목적 같은 것이 될 수 없다는 점이다. 만약 우리가 타인을 존중해야 한다면 이는 그들이 자체로 갖고 있다고 여겨지는 어떤 비관계적(non-relational) 가치 속성 때문이 아니라 그것이 단지 정언명법에 의해 명령되는 것이기 때문이다(Guyer 2006: 194, Sensen 103). 여기서 중요하게 부각되는 구절 중 하나는 타인에 대해 느끼는 존중심은 실상 도덕법에 대한 존경이지 사람이 갖는다고 가정되는 어떤 절대적인 속성을 향하는

것이 아니라는 주장이다. 가령 타인에 대해 느끼는 동정심과 같은 감정에 근거한 것은 엄밀한 의미에서 도덕이라고 볼 수 없고, “오직 **법칙에 대한 표상** 그 자체만이 우리가 도덕적이라 부르는 탁월한 선을 이룬다.”(G 401) 칸트는 쓰기를 “내가 본성상 그에 책무가 있는 모든 존경은 법칙 일반에 대한 존경”(MM 467)이고, “법칙에 대한 존경은 윤리를 위한 동기가 아니라, 오히려 윤리 자체”(C2 76)이다.

예컨대 셴센이 보기에, 칸트 윤리학에서 “존재론적으로, 존재하는 것은 오직 의지와 정언명법뿐이다(Sensen 107).” 타인을 존중해야 하는 의무는 그것이 정언명법에 의해 명령되기 때문이고, 정언명법은 예외를 허용할 수 없는(G 424) 보편성을 가진 것으로서 셴센은 정언명법을 보편법 정식과 같은 것으로 사용한다(Sensen 103). 그런 한에서 인간성 정식은 근본에 있어 보편법 정식과 다르지 않다. 이렇게 되면 인간성 정식은 어떤 새로운 주장을 하는 것이 아니라, 정언명령과 같은 도덕적 요구 사항을 보다 직관에 가깝게 하는 역할만을 할 뿐이다(G 437). 인간성은 절대적 가치이거나 도덕의 토대가 아니다. 요컨대 인간성 정식에 대한 형식주의적 해석은 타인에 대한 존중에 대한 요구는 모든 인간이 소유한다고 가정되는 어떤 절대적 가치 때문이 아니라 다른 아닌 정언명법으로부터 오는 것일 뿐이라고 주장한다. 이 점에서 가치, 목적 그 자체, 존엄성과 같은 개념은 칸트에게서는 이차적인 개념이며 이중 어떤 것도 정초의 역할을 하지 않는다(Sensen 118). 셴센에 따르면 칸트의 형이상학 체계에게 존재론적으로 실재하는 것은 자연의 법칙과 자유의 법칙, 즉 정언명법뿐이다.

이러한 해석은 칸트의 보편법 정식에서 명시적으로 드러났던 형식주의의 틀에서 인간성 정식을 일관되게 해석하고 정언명령이 칸트의 윤리학 체계에서 갖는 중요성을 강조한다는 점에서 장점을 갖는다. 그런데 이는 반대로 칸트가 목적 그 자체로서 인간성 정식에 부여하는 중요성을 너무 과소평가한다는 점에서 문제적인데, 각각의 정식은 정언명령의 고유한 특징들을 드러내도록 의도되었기 때문이다. 우선 칸트는 정언명법의 정식들 사이의 관계에 대해 논의하면서 모든 준칙은, 특히 일인칭 관점에서, 보편적인 형식과 마찬가지로 질료, 즉 목적을 가져야 한다고 명시적으로 주장한다(G

436, MM 385). “윤리학은 **질료**, (...) 순수 이성의 **목적**을 제시”하며, 여기서 목적이란 “한 대상의 표상에 의해 의사가 이 대상을 산출하는 행위를 하도록 규정되는, (이성적 존재자의) 의사의 대상이다.”(MM 381) “목적들에 관해 아무런 이해관심도 갖지 않는 것은 하나의 모순”(MM 395)이며, 이 점에서 보편법 정식이 형식을 다룰 때 질료에 관계하는 것은 바로 인간성 정식으로, 인간성 정식은 단순히 정언명법을 서술하는 부차적인 역할만을 하지 않는다. 오히려 보편법 정식과 인간성 정식은 단지 동일한 기본 원리의 상이하지만 동등한 진술 방식들일 뿐이며, 이는 인간성 정식의 역할이 없다는 뜻이 아니다. 칸트가 보기에 모든 실천적 입법의 근거는 객관적으로는 보편성의 형식에 있으나 주관적으로는 목적, 즉 목적 그 자체인 이성적 존재자에 있다(G 431)는 점에서, 인간성 정식의 중요성은 재차 강화된다.³⁾ 둘째로, 칸트가 도덕 판단의 엄밀성을 위해서는 정언명령의 보편적 정식을 따르는 것이 좋다고 하면서 덧붙이는 말은 도덕법칙을 널리 알리기 위한 목적을 위해서는 세 가지 개념으로 검토함으로써 직관에 가깝게 해야 한다는 것(G 437)이지, 유독 인간성 정식‘만’이 단지 직관을 위해서만 도입된 것이라고 말하고 있지는 않다는 점에 주목하라. 마지막으로 『윤리형이상학』의 덕론이나 『정초』에서 다루었던 여러 의무, 예컨대 자기계발의 의무나 자선행위의 의무(G 424) 또는 자기사랑과 이웃사랑의 의무(MM 410)처럼 더 넓은(칭찬할만한) 의무는 보편법 정식으로 검토할 경우는 자살이나 거짓약속의 경우처럼 자체로 내적 모순을 보이지는 않는 반면, 인간성이 항상 무조건적으로 가치 있는 목적 자체로 다뤄져야 한다는 점으로부터는 더욱 설득력 있게 설명될 수 있다. 이렇게 볼 경우 우리는 형식주의적 해석이 하듯이 인간성 정식을 단지 정언명법에 대해 잉여적인 재서술의 역할을 할 뿐이라고 간주하기 어려워지며, 이로써 우리는 인간성 정식을 진지하게 고려해볼 만한 이유를 확보하게 된다.

3) 칸트의 인간성 정식에 대한 주목 및 이 정식과 보편법 정식과의 관련성에 대해서는 강병호 2013b를 보라.

IV. 인간성 정식에 대한 해석: 목적 수립 능력으로서 인간성

앞에서의 검토를 통해 칸트가 인간성 정식이 그렇게 사소한 것이 아니라 는 점이 드러났다면, 이제 과연 칸트가 “인격 안의 인간성”(G 429)이라는 말로써 뜻하는 것이 어떤 것인지를 규명하는 것이 인간성 정식에서 해석의 관건이 된다. 이를 본격적으로 살피기에 앞서, 칸트가 말하는 인간성을 단순히 인간(mankind/human being)으로 볼 수 있지 않느냐는 질문에 대해 살펴보자.⁴⁾ 이는 인간성 정식만을 따로 접했을 때 지극히 상식적으로 가능한 반응이다. 인간성 정식은 우리가 일상적으로 만나는 사람들 하나하나를 단지 수단으로만 대하지 말고 또한 목적으로 대하라는 행위 정식처럼 들리기 때문이며, 이 점에서 인간성을 인간으로 읽는 것은 칸트가 무조건적이고 객관적으로 가치를 갖는 목적 자체로서 인간성을 논의한다는 점을 직관적으로 드러내준다는 장점을 갖는다. 이 점에서 팀머만 같은 논자들은 단순히 이성적 ‘본성’(Natur)을 이성적 존재자(Wesen)로 고쳐 읽자고 제안하기도 한다.(Timmermann 2007: 95-96)⁵⁾

그러나 유념해야 할 것은 칸트가 인간을 경험적 인간보다는 이성적 존재와 거의 등치시키는 경우가 많다는 점이다. 따라서 칸트가 인간성을 존중하라고 할 때 이를 단순히 생물학적인 종의 하나로서 사람에 대한 존중으로 읽는 것은 오도적이다. 칸트는 분명하게도 일종의 공간적 은유, 즉 인간성이 우리 모두가 지니는 어떤 것이라는 것을 이용하고 있으며, “인격 속에 인간성”(G 429, 나의 강조)이라는 표현이나 “인간성 일반의 이념 (...) 인간이

4) 주지하다시피, 이 독일어 원어 Menschheit는 영어에서 humanity와 mankind로 번역될 수 있다.

5) 팀머만은 다른 곳에서도(Timmermann 2006: 71-72) 본성(Natur)이 오늘날의 용법과는 달리 어떤 속성이나 능력을 가리키기 보다는 단순히 피조물(creature), 존재자나 개별적인 대상 등을 가리키고 있다고 지적한다. 확실히 여러 군데에서(G 408, 430, 436, 437 등) 이성적 본성을 이성적 존재자로 읽는 것이 자연스러우나 이러한 전거만으로는 칸트가 인간성이라는 말로 뜻하고자 하는 바에 대한 여러 해석들에 확실한 관정을 주기는 어려워 보인다. 이는 아래에서 자세히 볼 것처럼 칸트가 인간성으로 경험적인 개별 인간들을 가리킨다고 보기 어렵기 때문이다.

아니라”(MM 451)⁶⁾이라는 구절 또한 칸트가 인간성이라는 이념과 경험적인 인간(human being)을 구별하고 있음을 보여준다. 『실천이성비판』에서 칸트는 감성 세계의 일부로서 인간이 자신을 넘어서게 하는 것, 오직 지성만이 생각할 수 있는 인격성이 자유 내지 자연의 기계성으로부터 독립성이라고 정의한다. “감성 세계에 속하는 것으로서 인격은 그러므로 그것이 동시에 예지의 세계에 속하는 한에서 그 자신의 인격성에 복종해 있는 것”(C2 87)이다.

따라서 우선 주목해야 할 것은 칸트가 인간, 특히 인간성을 여러 차례 이성적 본성(Natur) 내지 이성적 존재(Wesen)라는 말과 혼용한다는 점이다(특히 G 429-439).⁷⁾ 이에 대해 유력하게 제기되었던 해석은 이성적 존재로서 인간성이란 곧 목적 수립의 능력을 지시한다는 견해이다. 여기서는 최근의 칸트 해석 중 가장 표준적인 견해로 간주되는 것으로서(Timmermann 2006: 69, Wood 127) 코스가드가 『목적의 왕국』 4장에서 전개한 논증을 중심으로 살펴보기로 하자.⁸⁾ 그녀에 따르면 인간이 절대적 가치를 가질 수 있는 것은 그가 사물의 가치에 대해서 항상 무조건적인 조건이기 때문이다.

6) “인간이 아니라”라는 표현은 초판에만 등장한다.

7) 칸트는 특히 자신이 말하는 도덕성이란 모든 이성적 존재에게 공히 타당한 것이라고 주장하는데, 이 또한 이성적 존재로 뜻하는 말이 단지 인간만이 아님을(다시 말해 만약 그런 것이 있다면 신이나 천사 같은 이성적 존재에게도 도덕성은 마찬가지로 타당하다는 것을) 시사한다. 다만 정언·명법’은 여타의 병리적 경향성들을 갖고 있고 도덕법에 복종하지 않을 수도 있는 인간에게만 해당된다. 의무 같은 도덕 명령, 자기강제는 신성한 존재를 포함한 이성적 존재자들 일반이 아니라 이성적 자연존재자인 인간에게만 관련되기 때문이다(MM 379).

8) 여기서 자세히 다루지는 않지만, 몇 가지 차이에도 불구하고 우드의 경우도 인간성을 이성적 행위자, 특히 목적 수립의 능력과 관련하여 사고한다는 점에서 코스가드의 생각과 궤를 같이 한다(Wood 127). 그에 따르면 “이성을 통해 목적을 수립하는 능력은 우리의 인간성을 구성하는 능력들의 집합을 결합시킨다.”(Wood 119) 또한 다음 구절도 참조하라. “칸트는 정언명법을 인간성, 즉 그 의지가 선하거나 악하거나와 무관하게, 이성으로부터 목적을 수립하는 능력을 갖는 어떤 존재의 가치 위에 정초할 것을 제안한다.”(Wood 120-21) 우드는 인간성 정식이 칸트가 도덕법에 관해 언급한 것들 중 가장 중요한 정식이라고 생각하며, 정언명법을 따라야 할 이유 역시 목적 자체이자 최고의 가치를 갖는 인간성 또는 이성적 본성에서 온다고 생각한다. 코스가드처럼 우리가 목적 수립을 통해 가치를 ‘부여’한다고 생각하지 않더라도, 우드는 이성적 본성이 객관적 가치의 유일한 원천일 수밖에 없다고 주장한다.

사물의 가치가 항상 조건적인 것에 그치는 반면에, 사물에 가치를 부여하는 것은 언제나 인간이다. 코스가드가 보기에 칸트는 인간성과 이성적 본성을 혼용하는데, 그녀가 특히 주목하는 것은 인간만이 유일하게 “스스로 목적을 세운다는 점”(G 437)이다. “일반적으로 어떤 목적을 세울 수 있는 능력이 (동물성과 구별되는) 인간성의 특징”(MM 392)이며 인간에게 의무는 “오로지 스스로 목적을 세울 수 있는 능력이 있는 인간성으로 점점 더 높이 솟아오르는 것”(MM 387)이다. 이처럼 목적을 자유롭게 선택한다는 점에서 인간은 동물과 구별된다. 그런데 코스가드가 『정초』의 텍스트를 넘어서, 「추측해 본 인류 역사의 기원」 및 『이성의 한계 내에서의 종교』와 같은 전거를 인용하며 지지하는 해석은 인간성 정식이 무조건적으로 가치 있게 여기도록 명령하는 것이 “**도덕적으로 의무적인 목적들을 채택하는 능력 뿐 아니라 목적 일반의 이성적 규정 능력**”(capacity for the rational determination of ends in general)이라고 주장하는 대목이다(Korsgaard 1996a: 111, 인용자의 강조). 코스가드는 칸트가 인간성의 특징을 목적 수립의 능력이라고 말할 때 이것이 단지 도덕적 목적을 선택할 수 있는 인격성을 뜻한 것만이 아니고 목적을 선택하거나 의욕하고, 가치 있게 여기는 보다 일반적인 능력에 있다고 본다. 인간성 정식에서 수단으로서만이 아니라 항상 목적 자체로 대우되어야 하는 까닭은 바로 이러한 이성의 능력에 있다. 코스가드는 칸트가 인간성 정식에서 전제하는 것이 우리가 행위할 때 스스로를 이성적으로 행위하는 것으로 간주하며, 또 우리의 목적이 객관적으로 선한 것으로 가정한다는 점이라고 주장한다(Korsgaard 1996a: 116). 목적이 이성에 의해 규정되고 또 그 목적이 선한 것으로 여겨진다면 이는 모든 이성적 존재 모두에게 타당한 이유를 갖는다. 만약 목적의 선함이 단지 상대적이면 모든 이성적 존재에게 적용될 수 없고 정언명령의 기초도 될 수 없을 것이다.

코스가드는 정언명령에 요구되는 목적 자체가 도출되는 배경이 충분한 이유가 있는 목적에 있다고 생각하며, 정당화를 추구하는 이성의 무조건적인 것에 대한 추구에서 온다고 본다. 이 점에서 인간성 정식에 대한 논증은 무조건적/조건적인 것의 구별을 선 개념에 적용한 것이다. 즉 이성은 조건적인 것에서 무조건적인 조건이 발견될 때까지 소급한다(Korsgaard

1996a: 117). 이성은 그 목적을 완전히 정당화할 수 있는 것, 무조건적으로 선한 것을 추구한다. 이 점에서 실천이성은 이론이성과 마찬가지로이지만, 무조건적인 것의 추구에 있어 필연적으로 이율배반을 일으키지 않는다는 점에서 이론이성과 다르다. 앞에서 보았듯이 이러한 소급의 방식은 또한 소거의 방식이기도 한데, 경향성의 대상, 경향성, 우리 자신, 우리의 이성적 본성으로 소급되면서 이성적 본성이 최종적으로는 객관적 목적으로 선택되기 때문이다. 그런데 여기서 코스가드 본인의 적극적인 해석이 게재되는 것처럼 보인다. 이성적인 선택의 대상을 선하게 만드는 것은 그것이 이성적 선택의 대상이라는 점으로, 이때 이성적인 선택은 가치 부여적 지위를 갖는다(Korsgaard 1996a: 122-23). 만약 이성적 선택 능력이 가치 부여적 지위를 갖는다면, 이성적 선택 능력을 가진 자는 누구나 그 능력에 의해 가치 부여적 지위를 가진 것으로 보아야 한다. 코스가드는 이 점에서 이성적 본성이 행위의 합리적 제한 조건이기 때문에(G 437-8), 만약 이처럼 자기의 목적이 갖는 선함의 원천인 이성적 본성을 부정하는 것은 자신의 목적뿐 아니라, 그것을 목적으로 하는 행위 또한 선하거나 이성적일 수 없게 만든다고 주장한다. 요컨대 어떤 사람을 목적 자체로 대하는 것은 그 인격을 그녀가 선택한 대상에 가치를 부여하는 자로 간주하는 것이다. 그러므로 목적 수립 능력으로서 인간성을 무시하는 강제와 기만 같은 것은 타인을 단지 수단으로만 대한다는 점(G 429)에서 타인에게 행할 수 있는 가장 나쁜 행위에 속한다.

코스가드는 비본능적인(non-instinctual)한 욕구에서 비롯한 ‘목적 일반’의 선택에서 이성적 본성을 보지만(Korsgaard 1996a: 112-13), 과연 ‘모든’ 목적 수립의 능력이 가치 부여하는 능력이라고 볼 수 있는지, 또는 칸트의 입장에서 볼 때 존엄하다고 여겨질 수 있는 것인지에 대한 의문을 불러 일으킨다. 과연 목적 수립 능력 ‘일반’을 인간성으로 해석할 경우, 그러니까 인간성을 너무 최소주의적으로 해석하는 것이 아닌가? 목적 수립 능력이 수립한 목적이 늘 가치있다고 볼 수는 없기 때문이다. 코스가드는 어떤 사람을 목적 자체로 대하는 것은 그 인격을 가치를 부여하는 자로 간주하는 것이라고 하면서 강제와 기만 같은 경우만을 비판하지만, 우리는 비도덕적

이유로 세워진 목적 수립 역시 가치있는 것으로 허용할 수 있다. 그러나 목적 수립에 있어 비도덕적 목적 수립을 배제할 어떤 특정한 제약 같은 것을 코스가드가 인간성 정식에 대한 해석에서 제시하는 것 같지는 않다. 주목해야 할 것은 칸트가 『윤리형이상학』 등의 저작에서 목적 수립의 능력이 인간성을 특징짓는 것이라 쓸 때 이는 이성적 존재자로서 인간이 선택한 것이라면 어떤 목적이든 상관없다는 것이 아니라는 점이다. 라틴어 선(bonum)이라는 말이 지닌 애매성 때문에 칸트는 일반적인 의미의 선(das Wohl)과 도덕적 의미의 선(das Gute) 사이의 중요한 구별을 도입한다(C2 59-62). 이때 코스가드나 우드가 생각하는 목적을 수립하는 능력 자체는 넓은 의미의 선에 관련되는 것으로 보이며, 이는 칸트가, 그리고 후일 헤겔이 비판하는 바 한갓 선택의 자유를 가리키는 자의(Willkür)에 해당할 것이다.⁹⁾ 오히려 칸트가 인간성 정식에서 의미하는 것은 도덕적 의미의 선을 택하는 능력이다. 그가 명백히 하듯이 “비록 인간이 (...) 자기 자신에게 목적들을 세울 수 있기는 하지만, 그럼에도 이것은 단지 그의 유용성의 **외생적인(extrinsic)** 가치를 줄 따름”이다(MM 434).

이에 대해 코스가드는 자신도 모든 목적이 가치있다고 주장하는 것이 아니며, 비도덕적이거나 반도덕적인 목적이 수립되는 이유는 목적의 타당성의 범위(scope)가 이성적 존재자 전체를 포괄하지 못하기 때문이라고 응수할 수도 있을 것이다. 그러나 설령 그렇다고 해도 목적 수립 능력이 모두에

9) 코스가드가 인간성 개념의 의미를 ‘도덕적’ 목적의 수립이 아니라 목적 ‘일반’ 수립 능력에서 찾을 때, 그녀는 『정초』보다는 「추측해 본 인류 역사의 기원」 및 『이성의 한계 내에서의 종교』와 같은 전거에 의존하여 그렇게 한다. 나는 이때 코스가드가 『정초』와 인간성을 동물성과 대비시키면서 인간성을 보다 폭넓게 정의하는 다른 저작들(C2, MM, R)간의 미묘한 차이를 고려하지 않음으로써 칸트가 『정초』에서 의미하고자 하는 바를 공정하게 다루지 못하고 있다고 생각한다. 우드 역시 『이성의 한계 내에서의 종교』를 끌어다가 이를 해석함으로써, 코스가드와 유사한 오류를 범한다. 우드는 인간성의 경향이 “어떤 자유, 즉 욕망과 충동의 즉각적인 강제에 저항할 수 있는 능력”을 전제한다고 지적한다(Wood 119). 이는 물론 도덕성을 위한 필요조건일 수 있지만 충분조건이 아니며, 적어도 칸트가 『정초』에서 생각하는 인간성 개념을 너무 넓게 해석하고 만다. 여기서 자세히 다룰 수는 없지만, 이와 같은 완전가능성으로부터의 자유 개념이나 칸트의 *Eigenliebe*나 *Eigendünkel* 사이의 구별은 루소의 『인간불평등기원론』에서 나타나는 자기애(amour de soi)와 자기편애(amour propre)의 구별로부터 가져온 것이다.

게 가치있는 것이 될 수 있는가? 코스가드는 목적에 객관적 가치를 부여하는 목적 수립 능력을 모든 목적에 해당되는 것으로 보지만, 목적 수립 일반이 과연 이성적 존재 ‘모두’에게 타당한 어떤 객관적인 가치의 원천이 될 수 있는가? 목적 수립 능력 일반이 과연 객관적으로 타당할 수 있는지에 대한 몇 가지 반례들을 생각해볼 수 있다. 먼저 목적이 가치있는 것은 오직 제한된 집단의 사람들에게만 해당될 수 있다. 가령 죄수와 간수가 각자의 상황에서 갖는 목적 수립은 각자에게 상대적일 것이며, 서로 상충될 것이다. 또한 나치 군인처럼 그의 목적에 긍정적인 가치를 전혀 부여하지 않고 부정적인 가치를 갖는다고 믿는 경우, 가정 폭력 등 어떤 트라우마로 인해 스스로가 가치있다고 여기지 못하는 경우도 있을 것이다. 마지막으로 아무런 목적 없이 시간낭비를 하는 것처럼 상당히 사소하고도 도구적으로 보이는 목적 수립의 여러 경우들, 긍정적이기도 부정적이기도 않아 보이는 목적의 경우 등이 있을 수 있다. (그 밖의 예들에 대한 자세한 설명은 Martin 99, 106). 만약 이러한 반례들이 설득력 있다면 코스가드가 생각하는 ‘모든’ 이성적 존재자에게 타당한 목적 수립 능력 ‘일반’이라는 생각은 그 자체로 비일관적인 이상일 수 있다. 설령 목적 수립 능력과 그 목적 수립에서의 오용가능성을 분리해서 생각한다고 하더라도, 오용가능성을 허용함으로써 코스가드의 논의는 최소한 인간성에 관한 칸트의 텍스트로부터 벗어나게 되는 듯하다. 우리는 목적 수립을 통해서 어떤 목적을 모두에게 좋은 것으로 간주하고 있다는 주장은 칸트의 텍스트에서 찾기 어려우며, 각자가 수립하는 주관적이고 상대적인 목적이 모두에게 타당할 수 있는지, 그렇다고 하더라도 왜 그렇게 간주해야만 하는지 여부 역시 분명하지 않기 때문이다. 팀머만도 주목하였듯이(Timmerman 2007), 칸트가 항상 강조하는 것은 의사 결정 절차로서 보편법 정식이나 도덕성의 근거로서 정언명법과 같은 어려운 철학적 정당화 작업 없이도, 평범한 인간 이성을 갖춘 자는 누구나 도덕적인 것이 무엇인지를 즉각 안다는 점이다(G 391, 404). 코스가드는 한편으로는 실정적인 목적에 가치를 부여하는 것으로서 목적 수립 능력을 최소주의적으로 정의하는 듯 하면서도, 다른 한편으로는 객관적 타당성을 주장할 수 있는 소극적 의미를 가진 제한 조건으로서 목적 자체에 어떤 규범적 속성을

부여하는 듯 하다(Korsgaard 1996a 108, 110). 코스가드는 전자에 대한 설명에 많은 부분을 할애하면서 전자와 후자의 설명 사이의 필연적 연관을 수립하려고 시도하는데, 이 연관이 위의 비판들을 견뎌내기에 얼마나 튼튼한지는 의문스럽다.

나는 코스가드가 주장하는 ‘인간성 = 목적 수립 능력 = 모든 가치의 최종 근거 또는 필요충분조건’이라는 등식은 칸트의 본래 주장이라기보다는 코스가드 자신이 부가한 해석이라고 생각한다.¹⁰⁾ 오히려 칸트가 인간성 정식을 다룰 때의 요점은 코스가드나 우드가 관심을 가지는 것처럼 가치의 원천이 어디에 있느냐 하는 것이라기보다는 도덕성이 어떻게 인간성과 연관되는지 그리고 우리가 이성적 존재자를 어떻게 도덕적으로 대우해야 하느냐는 문제에 관련된 것으로 보인다. 내가 보기에 칸트는 인간성 정식을 통해 타산적 합리성을 포함한 이성적 행위자의 이성적 목적 수립 능력 ‘일반’에 대해 이야기하는 것이 아니라, ‘도덕적’ 능력에 관련되는 것에 국한된 논의를 펼친다.¹¹⁾ 일종의 자유 선택의 능력으로서 목적 수립 능력 일반이 정언명법과 같은 것의 필요조건이 될 수 있을지는 몰라도, 이것이 인간성이 목적 그 자체로 대우받도록 하는 근거로서 충분조건인지는 상당히 불명확하다. 이 점에서 고려해보아야 할 것은 코스가드의 해석이 인간성과 더불어

10) 코스가드는 인간이 목적 자체가 무조건적인 가치를 가지고, 다른 사물들에 가치를 부여한다고 주장하며, 더 나아가 이후 『규범성의 원천』 같은 저서에서는 인간이 스스로에게도 가치를 부여한다고 주장한다(Korsgaard 1996b: 407, 122). 그러나 가치의 ‘원천’에 관한 이러한 주장은 칸트보다는 코스가드의 관심사에 속한다고 보아야 할 것이다. 코스가드는 분명히 규범성의 원천을 자율성으로 꼽고 있지만, 이는 도구론이나 실재론 등의 메타윤리적 논쟁의 맥락에서 제기된 것이다. 목적 수립 능력은 자율성과 분명히 관련된다고 볼 수 있지만, 내가 말하고자 하는 것은 이는 너무 광범위한 목적을 연루시키게 된다는 점에서 칸트의 본의와 멀어진다는 점이며, 이런 한에서 ‘도덕적’ 자율성이라는 제한 조건이 필요하다고 생각한다. 목적 수립 능력으로서 인간성에 대한 코스가드의 독해와 구별되는 필자의 독해는 6장에서 제시될 것이다.

11) 파핏도 칸트가 말하는 인간성을 단순히 가치의 원천으로서 합리성 내지는 이성적 본성으로 이해되는 것에 대해 우려하면서 존엄성을 갖는 존재로서 인간성이 단순히 비도덕적 합리성(non-moral rationality, 예컨대 체스를 잘 두거나 십자말풀이를 잘 푸는 등의 도구적, 타산적 합리성)이 아니라 도덕적으로 행하고 선의지를 가질 수 있는 능력에 해당함을 지적한다(Parfitt Vol. 1: 241-42, Vol. 2: 162, 165). 유사한 맥락에서 팀머만 역시 코스가드의 인간성 정식 해석에서는 도덕적 실천이성과 비도덕적 실천이성 사이의 선명한 구별을 하기가 어렵다는 점을 지적한다(Timmermann 2006: 80).

이야기되는 선의지나 존엄성과 같은 인접 개념들을 얼마나 잘 다루고 있는가 하는 점이다.

V. 인간성 정식에 대한 해석: 선의지로서 인간성

앞에서 보았듯 과연 목적 수립의 능력이 인간성일 뿐이라면 그러한 능력 일체를 가지 있는 것으로 간주해야 하는지에 대한 의문은, 코스가드와 같은 해석에 제기될 수 있는 가장 강력한 도전이다. 이와 관련하여 인간성이 그 자체로 절대적으로, 비교불가능한 가치를 갖는다고 전제하고 이를 칸트에서 일종의 인본주의적 가치설정으로 간주하는 입장에 대한 비판 역시 존재한다. 가령 딘의 경우는, 목적 수립의 능력이 선의지에 부합할 때 비로소 가치 있는 것이라고 본다. 칸트가 말하는 인간성은 오직 도덕적 원리를 행위를 위한 충분한 이유로 실제로 받아들이고 자기이익에 대해 도덕적 고려를 우선시하는 이들에게 나타나는 보다 완전한 인간성이라는 것이다(Dean 2006: 6, 20-21). 요컨대 딘의 논제는 인간성은 목적 자체로서 선의지라는 것으로 요약된다.

이러한 해석은 칸트가 모든 인간은 자체로 절대적인 가치를 갖는다(G 428)고 쓰기 이전에 『윤리형이상학 정초』 서두에서 이미 오직 도덕적인 선의지만이 세계 속에서도 심지어 세계 밖에서도 절대적인 가치를 갖는다고(G 393) 쓴 사실에 주목한다. 이런 선의지는 “유일한 선이고 완전한 선일 필요는 없지만, 그래도 최고선이고 나머지 모든 선의 조건이며, 행복에 대한 모든 열망 자체의 조건”(G 396)이기에, 인간성 정식에서 목적 자체로 제시되는 인간성과 선의지는 모종의 관계를 맺는다. 딘은 이처럼 칸트의 선의지에 관한 설명과 목적 그 자체에 대한 설명 사이의 어떤 유사성을 중시한다. 칸트가 『정초』 서두에서 말하는 바에 따르면 오로지 선의지만이 비교불가능한 높은 가치, 존엄성을 갖는다. 또한 『정초』 2장에서 말하듯 인간성은 목적 자체이고 존엄성을 갖는다. 그는 코스가드가 생각하는 것처럼 모든 목적 수립 능력 일반이 가치 부여 능력이라고 보는 것이 아니라, 선의지야말로

모든 가치의 근원이고, 이는 또 다시 선의지와 인간성은 동일하다는 주장으로 나아간다. 딘은 그가 재구성한 다음의 두 가지 전제들 즉 P1) ‘만약 x가 경향성으로부터 독립적으로 가치있다면, 그 경우 x는 목적 그 자체임에 틀림없다’와 P2) ‘만약 x가 모든 가능한 상황에서 가치있다면, 그 경우 x는 조건 없이 선함에 틀림없다’라는 두 가지 전제를, ‘인간성이 모든 가능한 상황에서 가치있다’는 칸트의 주장 및 ‘선의지는 경향성으로부터 독립적으로 가치있다’는 또 다른 주장과 결합함으로써 자신의 선의지 해석을 정당화시킨다(Dean 2006: 38-40). 이미 『정초』 1장에서 칸트가 제기한바 조건 없이 선한 유일한 것은 오직 선의지뿐이라는 주장과 인간성이 조건 없이 선하다는 주장을 일관적있게 만들기 위해서는 인간성과 선의지를 동일시하는 것이 필수적이다.

딘의 해석은 또한 칸트가 목적 그 자체는 감각적 경향성에서 비롯한 주관적 목적, 상대적인 목적처럼 어떤 야기되어야 할 사태나 대상 같은 것이 아니라 “모든 인간 행위의 자유를 제한하는 최상의 제한 조건”(G 430)과 같은 소극적인 역할을 한다고 쓴 점(G 437)을 잘 부각시킨다. 이처럼 객관적 목적 또는 목적 그 자체와 상대적인 목적의 구분, 자유로운 의사의 대상인 도덕적 목적과 감성적 충동의 목적의 구분(MM 385)을 염두에 둘 때, 칸트가 실천 원리로부터 배제되어야 한다고 생각하는 질료라는 말의 뜻을 일관적으로 이해할 수 있다. 경험적이고 상대적이며 주관적인 목적, 질료는 “그 실현이 욕구되는 대상”(C2 21)에 지나지 않기 때문이다. 반면 인간성은 야기되어야 할 이런저런 구성적 목적이 아니라 행위를 규제하는 자립적인 목적, 제한 조건으로 기능하는 소극적인 의미의 목적 자체에 해당한다. 이 점에서 딘은 인간성을 선의지에 동일시하는 해석이 칸트가 인간성을 그것을 실현하기 위해 노력해야 하는 하나의 이상(ideal), 도덕적 원형 같은 것으로 간주한다는 점과도 잘 조화된다고 주장한다(Dean 2006: 47-49).

이렇게 보면 인간성 정식이 뜻하는 바는 모든 인간은 자신의 선의지를 보존해야 하고, 다른 사람의 선의지를 단순히 수단으로 대해서 희생시켜서는 안 된다는 것으로 읽힐 수 있다(Dean 2006: 130). 인간성 정식이 타당한 이유는 선의지를 지닌 한에서 모두는 보편적으로 평등한 가치, 존엄성을 가

지기 때문이다. 딘의 해석은 우연적인 자유선택인 자의와 의지, 즉 이성적 존재에게 보편적으로 타당할 “어떤 법칙에 대한 표상에 맞춰 행동하도록 자신을 결정하는 능력”(G 427) 사이의 구별을 도입하고, 절대적으로 선한 선의지와 인간성이 동일하다고 주장함으로써 코스가드의 최소주의적 해석을 교정하는 것으로 보인다. 어떻게 단순히 목적 수립의 능력이 그것이 어떻게 쓰이느냐와 무관하게 인간에게 존엄성을 줄 수 있는가라는 의문은 인간성을 선의지와 결부시킬 경우에는 발생하지 않는다. 인간성의 지위는 도덕법과 무관하지 않다.

그런데 이러한 딘의 견해는 선의지를 곧 인간성과 등치시킴으로써 인간성의 영역을 너무 협소하게 만들어 놓았다는 점에서 문제적이다. 선의지만이 존엄성을 가진 대상이라고 보는 것은 너무 강한 도덕주의적인 주장이 아닌가? 이것이 정말 칸트가 말한 인간성의 존중에 해당하는가? 물론 인간성은 인격성(personality)의 실현 속에서만 완성되고 완벽해지며, 그러한 인격성의 실현은 선의지이다. 그러나 선의지를 ‘가질’ 능력을 가지고 있음은 그것이 현실적으로 실현되는가의 여부와 별개로 무조건적인 목적으로 대해져야 한다는 조건으로 충분할 것으로 보인다. 우리가 흔하게 접하는 자제력 없음(akrasia)의 상황, 도덕적으로 올바른 동기가 무엇인지 알면서도 행하지 못하는 경우, 즉 선의지를 실현하지 못하는 경우를 인간성의 범위에서 제한해야 할 것인가? 이에 대해 딘은 선의지가 어떤 이의 특정 행위의 문제라기보다는 행위의 원리에 관한 문제일 것이라는 이유로 자제력 없음으로 선의지를 상실할 수는 없다고 반론할 수도 있을 것이다. 특수한 행위들과 상대적으로 더 영속적인 성격(character)을 구별하면서 그는 선의지를 특수한 행위들의 변화하는 특성이라기보다는 일종의 “고유하고 완성된 개인적 성격”(Dean 2006: 22, 102, Ameriks도 참조)이라고 정의한다. 다음의 경우를 생각해보자. 칸트의 ‘의무로부터의(aus Pflicht) 행위’와 ‘의무에 맞는(pflichtmäßig) 행위’ 사이의 유명한 구별(G 397, C2 71)에서 선의지는 전자에 해당할텐데, 만약 인간성을 이렇게 좁게 설정할 경우 우리는 인간성의 외연으로부터 너무나 많은 것들을 배제해야만 한다. 예컨대 어수룩한 손님에게 바가지를 씌우지 않는 주인이 단지 평판이나 명성을 위한 의도에서

‘의무에 맞는 행위’에 그칠 경우 이를 인간성에 미달한다고 보는 것은 너무 지나친 것으로 보인다. 이 주인이 칸트적인 의미에서 참된 윤리적 행위를 하지 못했다고 말할 수는 있어도, 인간성을 결여한 존재라고 보기는 힘들다. 우리는 특정 경우에서 주인이 어떤 행위 이유를 가지고 있는지, 그리고 행위들의 기저에 놓인 성격이 무엇인지에 대해 확신할 수 없다. 성격으로서 선의지는 단도 인정하다시피 어떤 행위들에 대한 경험적 관찰들을 통해서 는 확실하게 식별하기가 어렵기 때문에, 어떤 이의 선한 성격에 대한 정확한 평가 이전에 왜 우리가 모든 이성적 행위자들을 존경해야만 하는지는 불분명하게 된다. 내가 보기에 칸트가 인간성에 대해 견지하는 것은 어떤 “혁명”(R 47)을 요구하는 도덕적으로 바람직한 성격으로서 선의지보다는 약한 개념이다.

결국 딘의 선의지 해석은 선의지로서 인간성이라는 ‘이상’의 ‘현실적 실현’ 여부에 초점을 맞추으로써 칸트 윤리학에 대한 과도한 엄격주의 내지 엘리트주의적인 독해가 되고 마는 것 같다. 이에 대해 딘 자신은 선의지가 그렇게 드문 것이 아니기 때문에 선의지 해석이 도덕적 엘리트주의로 흐르지 않을 것이라고 반론할 수도 있을 것이다. 그러나 설령 딘의 설명이 우리가 그들에 대해 의무를 갖는 이들인 도덕성의 대상에 관해서는 도덕적 엘리트주의의 위험을 피할 수 있을지 몰라도, 의무를 갖는 것으로 가정되는 도덕성의 주체에 관해서는 그럴 수 없어 보인다. 만약 선의지를 갖는 것이 도덕성의 주체라는 지위를 얻는 데 있어 필연적인 조건이라면, 선의지를 결여하는 이들은 도덕적으로 의무를 갖지 않는다는 것이 따라 나오기 때문이다 (Allison 214). 만약 던처럼 인간성과 선의지를 동일시할 경우 몇몇 약한 인간은 선의지를 결여하므로 인간성은 오직 일부의(some) 인간에게만 귀속 가능한 것이 될텐데, 이는 인간성이 모든 인간에게 속하는 특징이고 존엄성을 갖는다는 칸트의 명시적인 설명(MM 463)과 상충되고 말 것이다(Denis 122-23도 참고). 요컨대 앞서 본 코스가드의 해석이 칸트가 말하는 인간성을 목적 수립 능력 일반으로 설정함으로써 인간성의 외연을 너무 넓게 잡았다는 점에서 문제였다면, 딘의 해석은 인간성을 선의지로 봄으로써 인간성의 외연을 너무 좁게 설정하는 듯 보인다.

이처럼 인간성을 선의지와 등치시킬 때 야기되는 문제들을 고려할 때, 나는 선의지를 위한 능력의 소지와 그 능력의 현실적인 실현으로서 선의지를 구분할 필요가 있다고 생각한다. 딘은 실현되지 않은 능력에 실현보다 더 높은 가치를 할당하는 것이 개념적으로 괴이하고도 비정합적인 것이라고 비판하지만(Dean 2006: 87), 나는 여기서 선의지가 선의지를 위한 능력보다 더 높은 가치를 지녔다거나 또는 그 역에 관해 주장하고자 하는 것이 아니다. 선의지는 바람직한 이상일 수는 있으나 인간이 지닌 복잡하고도 미규정적인 의지의 한 양상이며, 이를 전제한다. 핵심은 인간의 의지는 선하게 될 수 있을 뿐 아니라 악하게도 될 수 있다는 점이며, 이 점에서 나는 딘의 논의가 더 포괄적이고 설득력 있기 위해서는 칸트의 근본악에 대한 종교철학적 논의를 고려할 필요가 있다고 생각한다. 『이성의 한계 내에서 종교』와 같은 저작에서는 아무리 악한 인간도 선의 씨앗을 가진 것으로 간주되고, 또 반대로 어떤 선한 인간도 근본악, 다시 말해 악으로의 본성적인 성벽을 갖고 있다(R 36-7)는 점을 감안한다면, 인간성과 선의지를 곧바로 동일시하기는 더욱 어려워질 것이기 때문이다. 악으로의 성벽은 인간의 힘으로 절멸할 수는 없지만 적어도 극복 가능한 것이다. “만약 인간이 ‘선하게 창조되었다’고 말한다면, 그것은, 인간은 **선으로** 향하도록 창작되었고, 인간 안의 근원적 **소질**은 선하다는 것을 의미할 수 있을 뿐이다. 인간은 이 소질만으로는 아직 선한 것이 아니고, 그가 이 소질이 함유하고 있는 동기들을 그의 준칙 안에 채용하느냐 않느냐 - 이 일은 그의 자유로운 선택에 전적으로 맡겨져 있음이 틀림없다 - 에 따라서 그는 그를 선하게도 악하게도 만드는 것이다.”(R 44) 명백히, 이는 딘이 생각하는 그런 의미의 선의지가 아니다.

이렇게 보면 설령 그 인간이 선의지를 당장 발휘하지는 못할지라도 선의지를 가질 수 있는 능력을 가지고 있다는 사실만으로도 일단 목적 그 자체로 간주될 수 있을 것으로 생각된다. 딘은 선의지를 가진 존재만이 존엄하다고 주장하지만, 칸트의 입장은 이와 반대로 선의지를 갖는지 여부와는 독립적으로 인간성이 존엄하다는 것이다. 도덕적 행위에 있어 인간성으로의

성향(predisposition to humanity)이 도덕법 아래 있음을 의식하는 것과 관련된다고 할 때에도, 여기서 도덕법에 복종하느냐 그렇지 않느냐 하는 문제는 인간성의 개념 자체와는 다른 것으로 여겨져야 한다. 여기서 내가 선의지를 가질 수 있는 능력이라고 부르는 것은 바로 칸트의 맥락에서는 자율성의 능력인데, 자율성은 “자기입법”(G 431) 또는 “의지가 자기에게 하나의 법칙이 되는 성질”(G 440)로 여기서 선의지는 자율성 자체가 아니라 자율성이 낳는 하나의 상태이다.¹²⁾ 인간성이 어떤 특정한 선에 결부되느냐와 무관하게, 일종의 도덕적 지위로서 인간성은 도덕적 의무를 발생시키는 역할을 한다. 딘이 선의지라고 부르는 것은 여기서 특정한 도덕적인 선에 상응하고, 칸트가 인간성으로 뜻하는 바는 선의지를 가능하게 하는 도덕적 지위이기 때문에, 선의지와 인간성은 분리되어야 한다. 결국 자율성을 가진 이성적 존재는 단순히 수단으로서뿐 아니라 목적으로 대해야 한다는 것이 인간성 정식이 궁극적으로 전하고자 하는 것이다. 선의지를 결여한 존재를 인간성을 결여한 존재로 보는 것은 지나치

12) 딘은 최근의 논문에서도 인간성을 하나의 순수한 도덕적 이상, 합리적 이념(idea) 같은 것으로 분석하면서 기존의 입장을 고수하는 것처럼 보이는데(Dean 2013), 바로 이 이상 또는 이념의 지위가 어떤 것인지가 여전히 문젯거리이다. 여기서 딘의 해석은 차라리 인간성을 칸트가 말하는 덕, “하나의 이상ियो 도달할 수 없는 것이나, 그럼에도 끊임없이 그것에 가까이 다가감이 의무인 그런 것”(MM 409)에 등치시키는 것 같다. 다시 말해 딘은 마치 무조건적으로 가치있는 선의지라는 것이 마치 객관적 실재론의 입장에서 말하는 좋음처럼 이미 실존하고 있는 것처럼 가정하는 것으로 보인다. 만약 딘처럼 선의지나 도덕적 이상으로서 인간성을 존중의 대상으로 바라볼 경우 코스가드의 해석 이상으로 반직관적인 결과가 따르는 것처럼 보인다. 이러한 딘의 독해는 칸트 윤리학을 의무론이나 형식주의와 동일시하는 기존의 주류 견해에 대한 비교적 최근의 이의 제기와 관련된다고 볼 수 있는데, 최근 몇몇 논자는 칸트 윤리학의 목적론적 해석의 측면, 덕윤리와의 관련성 등에 관심을 가지면서 실현하고자 노력해야 하는 도덕적 이상 내지 원형으로서 선의지라는 개념에 주목한다(Herman, Guyer: 2005). 그러나 칸트가 주장하는 ‘좋은데 대한 옳음의 우선성’ 논제를 상기해보자. 칸트의 널리 알려진 보편법 정식의 형식적 규정이 보여주듯이, 가치란 보편화 시험 등의 통과 이전에 그 자체로 절대적인 것으로 간주될 수 없다. **“선악의 개념은 도덕 법칙에 앞서서가 아니라 (일핏 보면 심지어 이 개념이 도덕 법칙의 기초에 놓여야 할 법하지만), 오히려 (여기서 보이는 바대로) 도덕 법칙에 따라서 그리고 도덕 법칙에 의해서 규정될 수밖에 없다 (C2 63).”** 딘이 선의지를 추구해야 할 이상이나 이념으로 간주하건 하지 않건 간에 그는 선의지라는 절대적 가치를 마치 도덕법 이전에 규정되는 선처럼 바라보는 것 같은데, 이렇게 해석된 인간성은 칸트 윤리학의 근본적 규정과 상충될 것이다.

다. “인간성 자체는 존엄하다. (...) 사람은 다른 사람 누구에게나 인간성의 존엄을 실천적으로 인정해야 할 책무가 있다. 그러니까 다른 사람 누구에게나 반드시 보여야 할 존경과 관련한 의무는 인간에 의거한다.”(MM 462) 따라서 우리는 선의지를 갖지 않는 이에 대해서 도덕적 의무를 지니지 않는 것이 아니라 오히려 의무란 모두에게 해당되는 것이다. 우리가 누군가를 비인간적(inhuman)이라고 비난할 수 있다 하더라도, 선의지를 결여하는 이런 비인간적인 이들에게도 역시 인간성이 귀속될 수 있을 것이다.

VI. 인간성 정식에 대한 대안적 해석: 도덕적 자율성으로서 인간성

위에서 살펴본 이유들로 나는 인간성을 너무 광범위한 목적 수립 능력 일반이나 너무 협소한 선의지로 정의하기보다는, 도덕적 자율성이라는 능력으로 해석하는 것이 칸트의 본의에 가장 가깝다고 생각한다. 이러한 해석은 칸트 텍스트 자체에 놓여있는 긴장, 즉 한편으로는 모든 인간이 있는 자체만으로 목적 자체라는 주장(G 428, MM 463)와 다른 한편으로는 도덕성이나 선의지가 목적 자체임의 조건이라는 주장(G 435, 393) 모두와 일관적이다.¹³⁾ 목적 수립 능력 해석이 전자의 관점만을 채택함으로써 자율성을 너무 넓은 것으로 간주하고 선의지 해석이 후자의 관점을 도입함으로써 자율성을 너무 좁은 것으로 생각한다는 점에서 문제적인 반면, 여기서 내가 제안하고 싶은 것은 ‘도덕적’이라는 조건이 붙는다는 점에서는 목적 수립 능력보다는 제한적이고, 항상 선의지로 완전히 실현될 필요는 없다는 점에서는 선의지보다는 불완전한 자율성으로서 인간성 해석이다.¹⁴⁾

13) 이러한 이중성은 또한 롤즈에 의해서 인간성 정식에 대한 정의의 의무(허용불가능한 목적)에 관한 부정적 해석과 덕의 의무(허용가능한 목적)에 관한 긍정적 해석 사이의 긴장으로도 해석될 수 있다(Rawls 190-95).

14) 오해를 피하기 위해서 몇 가지 구별에 대해 말해두자. 코스가드는 이성적 자율성(Korsgaard 1996a: 240)이나 이성적 선택 능력(Korsgaard 1996a: 262)을 가치의 원천

나는 먼저 보편법 정식과 인간성 정식의 설명 이후 제시되는 자율성 정식은 단순히 정언명법의 한 정식에 그치는 것이 아니라 도덕성의 최고 원리이기도 하다는 점에 주목한다. 이 점에서 보편법 정식과 인간성 정식을 통합하는 것으로서 자율성의 구성적인 역할이 중요해진다. 이와 더불어 칸트는 기존의 도덕철학은 모두 어떤 법칙에 종속됨만을 주목했지만 자신의 도덕철학이 혁신적인 까닭은 자율성의 원리를 조명한 것에 있다고 말한다(G 432-3). 지금까지 윤리성의 원리를 발견하려는 노력들이 실패한 까닭은 인간이 법에 매여 있음만을 보았지 인간이 자기 자신의 보편적인 입법에 종속되어 있다는 사실을 보지 못했기 때문이라는 것이다(G 432). 따라서 인간성 정식만을 고립시키지 않고, 칸트의 도덕철학 전체를 관통하며 정언명법의 연역 경로의 축이 되는 도덕적 자율성을 시야에서 놓치지 않는 해석이야말로 관건이 될 것이다. 이 해석을 지지하기 위한 몇 가지 이유들을 살펴보자.

첫째로 이와 관련한 칸트의 여러 문헌적 근거들이 존재한다. 이를테면 『정초』에서 제시되는 결정적인 구절은 다음과 같다. “**자율성은** 인간과 모든 이성적 존재의 존엄성의 근거”(G 436)이고, “의지의 자율성이란 의지가 그 자신에게 (의욕의 대상들의 모든 성질로부터 독립적으로) 법칙인 그런 의지의 성질이다.”(G 440)¹⁵⁾ 이는 『정초』의 다른 여러 구절에서도 확인되

이라고 부르면서 목적 수립 능력과 유사한 것으로 생각하지만, 자율성 개념의 도덕적 함축, 그리고 인간성 정식과 다른 정식들 간의 관계를 잘 규명하지는 못한 것으로 보인다. 또한 그녀는 “완전히 이성적인 의지”를 선의지라고 부르면서 목적 수립 능력과 선의지를 구별하면서도(Korsgaard 1996a: 17, 240) 다시 선의지가 가치의 원천이라고 주장하는 등 목적 수립 능력, 자율성, 선의지를 다소 혼동되는 방식으로 사용하지만, 가능한 오해를 피하기 위해서 이 개념들 각각은 보다 명확히 구분될 필요가 있다. 이 점에서 나는 칸트가 자율성이 인간 존엄성의 근거라고 이야기할 때(G 436) 선의지와 구별되는 어떤 것을 이야기한다고 생각한다.

- 15) 자율성 개념의 철학사적 연원과 관련하여 주목할 것은 칸트가 1765년에 쓴 다음의 글이다. “나는 본래 탐구자다. 나는 인식에 대한 전적인 갈망과 거기서 전진하고자 하는 탐욕스러운 조바심을 느끼며, 또한 매번의 획득에서 만족을 느낀다. 한때 나는 이것만이 인류의 영예를 구성한다고 믿은 나머지 아무 것도 모르는 저속한 사람들을 경멸했다. 그러나 루소가 나에게 올바른 길을 알려주었다. 내가 지니고 있던 맹목적인 편견이 사라지게 되었다. 나는 인간을 존중하는 것을 배웠다(인용자의 강조).”(Kant 1910ff: 44, Kant 2005: 7) 칸트가 존중의 대상으로서 말하는 인간성이란 단순한 목적 수립 능

는 것이지만, 특히 『실천이성비판』의 변증론 같은 곳을 보더라도 잘 드러난다. 이에 따르면 우리가 객관적 목적일 수 있는 것의 바탕에는 자율성이 있다. “목적들의 순서에 있어서 인간은 (그와 함께 모든 이성적 존재자는) **목적 그 자체**라는 것, (...) 그러므로 우리 인격의 **인간성**은 우리 자신에게 신성하지 않을 수 없다는 것, 이것은 이제 당연한 결론이다. 왜냐하면, 인간은 **도덕법칙의 주체요**, 그러니까 그 자체로 신성한 것의 주체며 (...) 이 도덕법칙이 자유 의지인 인간 의지의 자율에 기초해 있고, 자유의지는 인간의 보편적 법칙들에 따라 반드시 그가 **복종해야만** 할 것에 동시에 **일치할** 수 있어야만 하기 때문이다.”(C2 131-2)

더 나아가 칸트는 ‘도덕적’ 자율성이 바로 인격성을 특징짓는 것이라고 여러 차례 언급한다.¹⁶⁾ “인간만이 자신에게 스스로 줄 수 있는 그런 가치의 소유할 할 수 없는 조건은 어떤 뿌리에서 유래할 수 있는가? (...) 그것은 **인격성**으로, 다시 말해 자유 내지 전 자연의 기계성으로부터의 독립성으로, 그러면서도 동시에 고유한, 곧 자기 자신에 의해 주어진 순수한 실천 법칙들에 복종하고 있는 존재자의 한 능력으로 보여진다. 감성세계에 속하는 것으로서 인격은 그러므로 그것이 동시에 예지의 세계에 속하는 한에서 그 자신의 인격성에 복종해 있는 것이다.”(C2 86-7) 그런데 감성계와 예지계에 모두 속하는 이중적 존재로서 인간은 바로 그 이유 때문에 인간임 그 자체로, 또는 목적 수립을 할 수 있다는 이유 하나만으로는 목적 그 자체가 될 수 없다. 왜냐하면 칸트의 다음과 같은 구절은 인간성은 단순히 인간도 아니고 목적 수립의 능력 일반도 아님을 시사하기 때

력 일반도 선의지도 아니며, 도덕적인 능력으로서 자율성이다. 이 점이 바로 칸트가 1765년에 겪은 전회의 핵심으로, 이는 루소로부터 이어받은 자율성의 정의를 환기시킨다. 루소는 『사회계약론』 1부 8장에서 인간을 인간답게 하는 성질로서 도덕적 자유를 제시한다. “오직 그것에 의해서만 인간이 자신의 참된 주인이 될 수 있는 도덕적 자유(liberté morale) (...) 왜냐하면 단순히 욕구의 충동만을 따르는 것은 노예이며 **스스로 부과한 법에 복종하는 것은 자유**이기 때문이다.” (Rousseau 365, 인용자의 강조).

- 16) 칸트의 맥락에서 이야기되는 자율성은 현대의 중립적 용법, 가령 생명윤리에서 논의되는 자율성과 온전히 같지 않다(Dean 2006: 10장 참조). 이 점에서 나는 ‘도덕적’이라는 술어가 단순히 잉여적이라고 생각하지 않는데, 이는 특히 칸트의 자율성이 이미 도덕성을 함축한다는 독해를 공유하지 않는 해석자들(가령 마틴 같은 경우), 심지어 비칸트주의자들에게는 더욱 그렇다.

문이다. “도덕 법칙은 **신성하다**(불가침이다). 인간은 비록 충분히 신성하지는 못하지만, 그러나 그의 인격에서 **인간성**은 그에게 신성하지 않을 수 없다. (...) 오로지 인간만은, 그리고 그와 더불어 모든 이성적 피조물은 **목적 그 자체**이다. 인간은 곧 그의 자유의 자율의 힘에 의해, 신성한 도덕 법칙의 주체다.”(C2 87)

둘째로 인간성 정식의 인간성이 자율성이라고 볼 경우, 칸트가 말하는 존엄성의 의미에 대해서도 보다 좋은 해석을 제공할 수 있다. 칸트가 서술 하듯이 “내적 가치, 즉 **존엄성**”(G 435)이라거나 “무조건적이고, 비교불가능한 가치”(G 436)라는 표현은 인간성이 갖는 존엄성은 그 자체로 어떤 절대적인 가치를 가짐을 보여주는 것 같다. 그러나 이러한 존엄성은 실은 키케로와 로마적 전통의 *diagnitas*에 빚진 것으로 일종의 “고양된 위치나 등급”(elevated position or rank)로 읽을 수 있다(Sensen 115). 이때 칸트는 존엄성을 일종의 숭고함(Erhabenheit) 내지 다른 것보다 더 높은 것과 연관하여 생각한다. 이 점에서 존경/존중의 구별을 선의지/인간성에 대응할 수 있을까? 만약 그런 유비가 가능하다면, 우리는 선의지에 대한 존경과 인간(성)에 대한 존중을 구별할 수 있다. 우리는 선의지를 가진 이를 ‘존경’하지만 이것으로부터 선의지를 결여하는 사람을 ‘존중’하지 않아야 한다는 것이 따라 나오지는 않는다. 또한 이를 칸트가 『윤리형이상학』의 「덕론」(MM 448)에서 하듯이 “인간 일반에게 마땅히 빚지고 있는 존중”(dem Menschen überhaupt schuldige Achtung, MM 462-63)과 업적으로 얻은 존경(verdiente Achtung)을 구별할 수 있고,¹⁷⁾ 이를 통해 우리는 딴 식의 선의지 해석을 논박할 만한 또 다른 근거를 갖게 되는 셈이다. 칸트에 따르면 인간성의 존엄성은 인간이 자유를 가진다는 점에서 자연에 속한 다른 어떤 것보다도 특별하다는 점으로부터 온다. 이 점에서 존엄성은 자율성 정식, “의지가 자기의 준칙에 의해 스스로를 동시에 보편적으로 법칙수립적인 것으로 생각할 수 있도록 행위하라”(G 434)와 관계된다. 다시 말해 이성

17) Achtung의 번역에 있어 존중과 존경의 구별에 대해서는 강병호 2013a: 133를 및 이에 앞서 업적에 대한 평가 존경(appraisal respect)와 사람에게 관련되는 인정 존경(recognition respect) 사이의 유사한 구별을 도입에 대해서는 Darwall 36-49 참조.

이 보편적으로 법칙수립적인 의지의 모든 준칙을 자기 자신뿐 아니라 모든 의지에 적용하는 것은 다른 어떤 실천적 동인이나 이익을 위해서가 아니라 자신이 다른 이성적 존재 뿐 아니라 자신에게 준 법칙 외에 따르지 않는 이성적 존재의 존엄성(G 434) 때문이다. 인간은 무엇보다도 이성적인 존재이고, 이 이성적인 존재가 목적 자체가 될 수 있게 하는 유일한 조건은 도덕성이다. 결정적인 대목에서 칸트는 “도덕성과 도덕적일 수 있는 한에서 인간성이 유일하게 존엄성을 갖는다”(G 435)고 지적한다. 인간은 자체로 어떤 자생적인 가치를 갖거나 하는 것이 아니고 도덕성이 인간에 존엄성을 부여한다. 『윤리형이상학』의 「덕론」 또한 “인격으로서, 다시 말해 도덕적-실천적 이성의 주체로 여겨지는 인간은 모든 가격을 뛰어넘는다. (...) 다시 말해 그는 존엄성(절대적인 내적 가치)를 가지며, 이에 의해 그는 다른 모든 이성적 세계존재자들에게 그에 대한 존경을 강요하고, 이런 종류의 모든 타자들과 자신을 견주어 평등의 입장에서 평가할 수 있기 때문”(MM 434-35)이라고 지적한다. “인간은, 의무의 주체인 한, 그러니까 그가 살고 있는 한, 인격성을 포기할 수가 없다(MM 422).”

셋째로 자율성으로 해석된 인간성은 앞서 본 선의지 해석과는 달리 목적들의 다양한 종류들, 불완전하거나 완전하게 이성적인 존재들 모두를 포괄할 수도 있다는 장점을 갖는다. 여러 사람들 사이에, 또는 심지어 한 사람 조차도 다른 시점에서 다양한 등급의 도덕성을 가질 수 있을 것이기 때문이다. 그래서 “우리가 언제나 존중해야 할 기본적인 존엄성을 인간에게 주는 것은 그들의 자율성, 그들 자신의 법에 따라 행위할 능력(성공적인 실현이 아니라)이다.”(Timmerman 2013: 218) 자율성을 이루는 요소가 스스로가 세운 도덕법 하에 있다는 의식과 그 법에 따르는 것이라고 할 때 후자가 현실적으로 일어나지 않다는 것, 도덕법을 따르지 않는다는 사실이 인간성 개념을 무효화하는 것은 아니다. 칸트가 인간성 정식에서 말하고자 한 인간성, 우리가 단지 수단으로서뿐 아니라 목적으로도 대해야 할 인간성이란 바로 ‘자율성’을 가진 존재로서 예지적 인간이다.

마지막으로, 타인을 자율적인 존재로 간주한다는 것이 자기입법적이고 보편적으로 간주될 수 있다는 사실은 인간성 정식과 보편법 정식, 자율성 정식,

목적의 나라 정식 사이의 밀접한 등가관계(G 436)를 잘 보여준다. 보편법 정식은 분명 어떤 목적에 대한 준거를 포함하며, 목적 자체로서 인간성이란 실상 자율성이다. 우리가 단지 스스로의 인간성을 목적 자체로 간주할 뿐 아니라 타인의 인간성 역시도 목적 자체로 간주해야 하는 까닭은, 인간성 정식이 이성적 존재자들 사이의 상호성에 찬동하고 있기 때문이다. 이성적 존재자로서 우리가 자율적인 존재라는 것은 인간 존중의 이유이며, 이 점에서 공통의 법 아래에 있는 목적의 나라의 구성원(G 433)은 마치 민주주의적 질서의 규범적 이상을 대표하는 자유롭고 평등한 시민들인 것처럼 그려진다. 모든 인격들 또는 목적의 나라의 시민들은 각자 자유롭게 가치수립의 능력을 갖는다는 점에서 자유롭고 평등하다. 다시 말해 우리는 자율적 존재로서 서로를 불편부당하게, 또는 같은 말이지만 공정하게 존중해야 한다. 우리는 자율적인 존재로서 각인이 평등한 외적 자유의 행사를 할 자격이 있음을 인정한다. 칸트가 **“이성적 존재가 목적 자체로 실존한다”**라는 보편적 실천법칙이 인간이 필연적으로 자신의 현존을 그렇게 표상한다는 점에서 **“인간 행위들의 주관적 원리”**이면서도 또한 **“다른 모든 이성적 존재자도, 나에게도 타당한 바로 그 동일한 이성 근거를 좇아, 그의 현존재를 그러한 것으로 표상한다. 그러므로 그것은 동시에 객관적 원리”**(G 429)라고 하는 것도 이 때문이다. 우리가 스스로의 인간성을 목적 자체로 간주할 이유를 가질 때, 마찬가지로 이유로 우리는 타인의 인간성도 목적 자체로 간주할 이유를 갖는다.

이는 자연스럽게 칸트의 윤리학과 정치철학과의 상관관계에 대한 논의로 이어진다. 칸트는 인간 권리의 위배가 타인을 단지 수단으로만 이용하는 것(G 430)이라고 하며, 이 점에서 인간성을 목적 자체로 대우하는 것은 권리의 보호 문제를 함축한다. **“자유(즉 타인의 강요하는 의사로부터의 독립성)는, 모든 타인의 자유와 보편적 법칙에 따라서 공존할 수 있는 한에서, 모든 인간에게 그의 인간성의 힘으로 귀속하는 유일하고 근원적인 권리이다.”**(MM 237) 다시 말해 인간성 정식에서 자율성의 의미를 중요하게 취급할 경우 불가피하게 자율적인 존재들 간의 목적 추구 간의 조정이나 평가의 문제 같은 정치철학적 문제들이 불거져 나온다. 이는 앞에서 살펴본 것처럼, 만민에게 보편적으로 타당한 목적 수립이라는 것이

난망한 일이기 때문이다. 칸트가 유아론이나 자아중심주의에 입각하지 않고 여럿의 행위자들의 존재를 인정하는 한에서, 각자의 자율성의 행사는 타인의 결정에 의존해 있고, 각각의 목적 수립의 와중에 상충 등이 발생할 수 있기 때문이다.¹⁸⁾ 개인들은 선의지를 항상 발휘하지 않을 수도 있고 자신만의 고유한 가치를 부여하며 살아간다. 이때 자율적인 존재로서 각자는 목적 그 자체로 존중받아야 하며 단순히 수단으로서만 간주되어서는 안 된다는 것, 그리고 다수의 자율적인 존재가 있다는 것은 타인의 목적 선택을 일정 정도 제한해야 할 수 있음을 함축한다. 앞서 본 목적 수립 능력 일반으로서 인간성 해석은 너무 많은 목적을 도입하게 되며, 선의지로서 인간성 해석은 너무나 적은 목적만을 도입하게 된다는 점에서 문제적이었다. 칸트가 말하는 목적은 추구되어야 할 적극적인 목적일 수도 있지만 그것에 반하여 행위하면 안 되는 “모든 한갓 상대적이고 자의적인 목적들을 제한하는 조건”(G 436)으로서 소극적인 역할을 할 수도 있다. 딘이 목적 수립 능력 일반을 인간성과 동치시키는 시도에 대해 우려했던 것과 달리, 우리가 도덕적으로 자율적인 존재로서 인간성을 목적 자체로 간주한다는 것이 모든 종류의 목적의 촉진을 허용함을 뜻할 필요는 없다. 누군가의 ‘준칙’이 타인에 대한 기본적인 존중과 상충될 때 이 준칙은 허용될 수 없다. 우리가 각자의 행복을 좇는다는 준칙을 따를 때 이는 각종의 불화와 갈등을 유발할 수 있지만, 인간성 정식은 보편법 정식(“마치 네 행위의 준칙이 네 의지에 의해 보편적인 자연법칙이 되어야 할 것처럼 그렇게 행위하라”)과 더불어 보충될 필요가 있다. 그리고 이 “보편적 자연 법칙은 모든 것을 일치/조화시킨다”(C2 28).

결론적으로, 지금까지 살펴본 것과 관련하여 칸트가 제기하는 난제는 이

18) 가령 칸트의 악명높은 예, 거짓말을 하지 말아야 할 의무에 대해서 칸트는 두 가지 답변을 할 수 있다. 『윤리형이상학』의 「덕론」에서 칸트는 거짓말을 절대적인 금지로서, 심지어 친구가 내 집에 숨어있고 살인자가 그를 찾아왔을 때조차 해서는 안 되는 것으로 제시하며 이는 매우 반직관적이다. 그러나 이는 단지 스스로에 대한 의무인 한에서 만 그렇다. 타인에 대한 의무를 동시에 고려하는 「법론」에서 그는 살인자의 비도덕적인 목적을 고려하여 문 앞의 살인자에게 거짓말을 할 수 있다고 주장한다. 타인을 단지 수단으로 대하지 말 것을 요구하는 인간성 정식은 이러한 방식으로 정치철학과 연계될 수 있다(Japa Pallikkathayil 144-46).

런 것이다. 인간성 정식이 보여주는 목적에 대한 실질적 가치 부여를 강조하면서도 어떻게 정식에 관한 형식적 설명을 할 수 있을까? 문제는 각자의 자유를 확보하면서도, 그러니까 의무적인 선행의 의무 목록들을 지나치게 강제하지 않으면서도, 비도덕적이거나 타인을 단순히 수단으로만 간주하는 그런 목적을 배제할 수 있는 적절한 방식을 확보하는 것이다. 우리 각자는 자율적인 존재로서 단지 수단으로서만 대우되어서는 안 되지만, 자율성의 오용이나 잘못된 목적의 등장 같은 문제의 돌발 가능성은 항존한다. 그러므로 인간성 정식의 자율성 해석을 끝까지 밀고 나갈 경우, 마치 루소의 사회계약론과 유사한 문제의식, 즉 타인의 동등한 자유를 보호하기 위한 적절한 규제를 통해서 어떻게 사회 상태에서도 자연-상태에서와 마찬가지로 여전히 자유로울 수 있는 제도를 창출할 수 있을지에 대한 문제의식에 이르게 된다. 모두에게 합리적으로 수용가능한, 정당화가능한 보편적 원리를 따른다는 상호성의 원리는 사회계약론의 핵심이기 때문이다.

VII. 나가며

지금까지 이 글에서 나는 인간성 정식에서 인간성이 도덕적 자율성을 의미한다고 주장했다. 이를 위해 나는 우선 인간성 정식이 단순히 정언명법을 직관에 가깝도록 하는 소극적인 역할만을 하는 것이 아님을 형식주의적 독해에 대한 비판을 통해 제시했다. 그리고 목적 수립 능력 독해가 인간성에 대한 최소주의적인 관점을 취하는 반면 선의지 독해는 최대주의적 관점을 취한다는 점을 비판적으로 살펴보았다. 내가 제안한 도덕적 자율성 독해는 이 두 가지 독해들이 가진 여러 단점들을 피할 수 있지만, 그렇다고 앞서 살펴본 독해들과 전혀 관계가 없는 것은 물론 아니다. 한편으로 내가 코스가드의 관점에 대해 몇몇 수정과 구별의 필요를 제안하긴 했지만, 도덕적 자율성 독해는 그것이 도덕적 구성주의와 자율성의 가치에 찬동하는 한에서 선의지 독해보다는 목적 수립 능력 독해에 가깝다고 간주될 수도 있다. 다른 한편으로 도덕적 자율성 독해가 목적 수립 능력과 모순적이지는 않더라

도, 선의지 독해가 도덕성과 존엄성을 강조하는 점에 있어서는 선의지 독해에 동의한다. 그러나 선의지 독해가 너무나 협소한 인간성에 대한 관점 및 이상이나 덕으로서 인간성이라는 목적론적 해석을 고수하는 한에서, 도덕적 자율성 독해는 이러한 선의지 독해와는 구별되어야 한다.¹⁹⁾

투 고 일: 2014. 06. 30.
심사완료일: 2014. 07. 21.
게재확정일: 2014. 07. 22.

오근창
서울대학교 철학과

19) 이 글의 초고에 관해 여러 좋은 질문들과 비판적 논평들을 전해준 익명의 심사위원들 및 강병호, 강진호, 김현섭 선생님께 감사드린다. 그럼에도 불구하고 이 글이 가질 잠재적 문제들과 오류들은 전적으로 필자의 것임을 밝혀두면서, 이 글에 반영하지 못한 많은 지적들에 관해서는 향후 다른 기회를 통해 보완할 것을 약속드린다.

참고문헌

칸트의 저작

Kant, Immanuel. *Kants gesammelte Schriften*. Ed. Königlich Preußischen Akademie der Wissenschaften. Bd. I-XXII, Berlin, 1910ff.

_____. *Practical Philosophy*. Trans. Mary J. Gregor. Cambridge: Cambridge University Press, 1999.

_____. *Remarks on the Observations on the Feeling of the Beautiful and the Sublime, Notes and Fragments. The Cambridge Edition of the Works of Immanuel Kant*. Eds. Paul Guyer & Curtis Bowman. Trans. Paul Guyer & Frederick Rauscher. Cambridge: Harvard University Press, 2005.

_____. *Religion within the Limits of Reason Alone*. Trans. Theodore M. Greene & Hoyt H. Hudson. New York: Harper Torchbooks, 1960.

임마누엘 칸트, 『실천이성비판』, 백종현 옮김, 아카넷, 2009.

_____, 『윤리형이상학』, 백종현 옮김, 아카넷, 2012.

_____, 『윤리형이상학 정초』, 백종현 옮김, 아카넷, 2005.

_____, 『이성의 한계 안에서의 종교』, 백종현 옮김, 아카넷, 2011.

_____, 『판단력비판』, 백종현 옮김, 아카넷, 2009.

이차문헌

강병호, 「칸트에게서 사람에 대한 존중과 법에 대한 존경 - 관계 규명을 위한 시도」, 『칸트 연구』 31집, 2013a.

_____, 「정언명령의 보편법 정식과 목적 그 자체의 정식. 두 정식의 등가성에 관하여」, 『칸트 연구』 32집, 2013b.

- Allison, Henry. *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford: Oxford University Press, 2011.
- Ameriks, Karl. "Kant on the Good Will." in *Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar*. ed. Otfried Höffe. Frankfurt am Main: Vittorio Klostermann, 1989.
- Cureton, Adam. "A Contractualist Reading of Kant's Proof of the Formula of Humanity." *Kantian Review* Vol. 18, No. 3 (2013): 363-386.
- Darwall, Stephen. "Two Kinds of Respect." *Ethics* Vol. 88, No. 1 (1977): 36-49.
- Dean, Richard. *The Value of Humanity in Kant's Moral Theory*. Clarendon: Oxford University Press, 2006.
- _____. "Humanity as an Idea, as an Ideal, and as an End in Itself." *Kantian Review*, Vol. 18, No. 2 (2013): 171-195.
- Denis, Lara. "Humanity, Obligation, and the Good Will: An Argument against Dean's Interpretation of Humanity." *Kantian Review*, Vol. 15, No. 1 (2010): 118-141.
- Glasgow, Joshua. "Kant's Conception of Humanity." *Journal of the History of Philosophy*. Vol. 45, No. 2 (2007): 291-308.
- Guyer, Paul. "Ends of Reason and Ends of Nature: the Place of Teleology in Kant's Ethics." in *Kant's System of Nature and Freedom: Selected Essays*. Oxford: Clarendon Press, 2005.
- _____. *Kant*. London: Routledge, 2006.
- Herman, Babara. "Leaving deontology behind." in *The Practice of Moral Judgment*. Cambridge: Harvard University Press, 1993.
- Hill, Thomas E Jr. *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca, NY: Cornell University Press, 1992.
- Kerstein, Samuel. *Kant's Search for the Supreme Principle of Morality*. Cambridge: Cambridge University Press, 2002.

- Korsgaard, Christine. *Creating Kingdom of Ends*. Cambridge: Cambridge University Press, 1996a.
- _____. *The Sources of Normativity*. New York: Cambridge University Press, 1996b.
- Martin, Adrienne. “How to Argue for the Value of Humanity.” *Pacific Philosophical Quarterly*. Vol. 87, No. 1 (2006): 96–125.
- Pallikkathayil, Japa. “Deriving Morality From Politics: Rethinking the Formula of Humanity.” *Ethics*. Vol. 121, No. 1 (2010): 116–147.
- Parfit, Derek. *On What Matters*. ed. Samuel Scheffler. Oxford: Oxford University Press, 2011.
- Paton, H. J. *The Categorical Imperative*. Philadelphia: University of Pennsylvania Press, 1971.
- Rawls, John. *Lectures on the History of Moral Philosophy*. ed. Barbara Herman. Cambridge: Cambridge University Press, 2000.
- Rousseau, Jean-Jacques. *Oeuvres Complètes*, Vol. 3. Eds. Bernard Gagnebin & Marcel Raymond. Paris: Gallimard, 1964.
- Sensen, Oliver. “Dignity and the Formula of Humanity.” *Kant’s Groundwork of the Metaphysics of Morals: A Critical Guide*. ed. Jens Timmermann. Cambridge: Cambridge University Press, 2009.
- Timmermann, Jens. “Autonomy and Moral Regard for Ends.” *Kant on Moral Autonomy*. ed. Oliver Sensen. Cambridge: Cambridge University, 2013.
- _____. “Value Without Regress: Kant’s ‘Formula of Humanity’ Revisited.” *European Journal of Philosophy*. Vol. 14, No. 1 (2006): 69–93.
- _____. *Kant’s Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge: Cambridge University Press, 2007.
- Wood, Allen. *Kant’s Ethical Thought*. Cambridge: Cambridge University Press, 1999.

ABSTRACT

Kant's Formula of Humanity and Moral Autonomy

Oh, Keun-Chang

There have been many debates regarding how to make sense of Kant's formula of humanity (hereafter, "FH"), which seems to be intuitively appealing but perplexingly obscure. In this paper, I aim to propose an alternative reading of FH against the existing views and hope to show what FH as an end itself refers to and how it could be coherently connected to other formulas. To this end, I first reconstruct Kant's argument for FH by focusing on *Groundwork of the Metaphysics of Morals* (hereafter, "*Groundwork*") and briefly evaluate the validity of his regress argument (Section 2). In the following, I critically examine the flaws of several influential interpretations of FH proposed by Sensen, Korsgaard, and Dean. Unlike Sensen, my interpretation shows that FH plays a role that goes beyond that of merely bringing the formula of universal law (hereafter, "FU") closer to intuition (Section 3). In addition, Korsgaard's reading of FH is flawed since it has such a thin conception of humanity that the end-setting capacity in general even permits non-moral acts, contrary to Kant's intention (Section 4). Although Dean has suggested that humanity is equivalent to a good will, I argue that a good will reading supposes a too thick and moralistic notion of humanity and thereby excessively limits the contents of humanity (Section 5). Finally, I conclude that FH should be interpreted as moral autonomy by elucidating some relevant textual support and philosophical reasons (Section 6).

Keywords: Formula of humanity, end in itself, end-setting capacity, good will, autonomy

