

유전자와 도덕 사이: 이타성을 중심으로

정 상 모
(신라대학교 철학과)

1. 들어가는 말

일찍이 다윈은 인간의 도덕성의 본질이 협동과 이타성(“to aid one another and to sacrifice themselves for the good of common good”)에 있으며(Darwin, 1871: 166), 진화의 긴 여정에서 발생한 협동과 이타적 본성의 맹아에 인간의 높은 지적 능력과 언어 능력이 더해짐으로써 마침내 인간에게만 특유한 문화의 한 부분으로서 도덕이 발생했다고 보았다. 주지하듯이, 이런 다윈의 주장을 더욱 구체적이고 체계적으로 발전시킨 이가 사회생물학의 창시자 에드워드 윌슨(Wilson, 1978)이다.

윌슨은 도덕적 본능의 기본적 기원을 “협동과 배신 간의 역동적 관계”에서 찾는다. 그는 수천 세대를 거쳐 협동을 가능케 하는 유전자들이 전체 인

주 제 어: 사회생물학, 진화론적 이타성, 심리적 이타성, 도덕성, 발생계 이론
sociobiology, evolutionary altruism, psychological altruism, morality,
developmental system theory

류에서 우세하게 되면서 도덕 감정이라는 본능이 발생했다고 본다. 그런 본능이 각 개인들에게 양심, 자존심, 자책감, 공감, 수치심, 겸손, 도덕적 분노 등의 다양한 형태로 경험되고 나아가 명예심, 애국심, 이타성, 정의, 동정심, 자비, 구원 등의 보편적 도덕적 코드들을 표현하는 관습의 형성을 통해 문화적 진화를 몰고 간다고 한다(Wilson, 2005: 435-437).

그는 협동의 두 가지 기본 유형으로 “맹목적 이타성”(hardcore altruism)과 “목적성 이타성”(softcore altruism)을 든다(Wilson, 2000: 217-8). 전자는 타인을 향한 비합리적이고 일방적인 이타적 충동으로, 사회적 보상이나 처벌에 비교적 영향을 받지 않는 환경에서 발휘되는 반응 집합이다. 후자는 자신에게 돌아올 보답을 염두에 두는 궁극적으로는 이기적인 조건부 이타적 욕구이다. 윌슨은 맹목적 이타성이 혈연 선택을 통해 진화해 온데 비해, 목적성 이타성은 주로 개체 선택을 통해 진화해 왔으며 문화적 진화에 크게 영향을 받는 것으로 추측한다. 맹목적 이타성은 사회성 곤충에서 완벽하게 구현되고 사회생활을 하는 고등한 동물에서 흔히 볼 수 있다. 목적성 이타성은 혈연적으로 무관한 개인들 사이에서 이루어지는 호혜적 이타성으로 발전해 왔으며, 그 극단적인 발전의 형태를 인간 사회에서 볼 수 있다고 한다. 윌슨은 인간의 이타성도 혈연관계에 있는 타 개체들에게 작용하는 맹목적 이타성과 비혈연관계에 있는 타 개체들에게 작용하는 목적성 이타성으로 구성되어 있다고 한다. 그리고 비록 목적성 이타성이 주류를 차지하지만, 그 역시 원래는 맹목적인 형태를 갖는 강력한 감정적 통제 장치를 기초로 해서 성립한다고 본다. 혈연 집단에서는 희생의 대가가 자동적으로 혈연을 통해 자신에게 돌아오지만 비혈연 집단에서는 그렇지 않기 때문에 인지 능력의 도움을 받아 대가가 돌아올 수 있는 방식으로 이타성이 작동되도록 진화했다는 것이다. 또한 윌슨은 혈연 이타성과 호혜성 이타성은 개체 수준의 선택압을 통해 진화한다고 보았다.¹⁾

1) 윌슨이 비혈연 개체들에게 작용하는 순수 이타성을 따로 설정하지 않는데 비해 이어는 순수 이타성을 다른 두 종류의 이타성보다 훨씬 더 중요한 인간 윤리의 핵심 요소로 간주한다. 그리고 후자는 문화 집단에 작용하는 선택압에 의해 진화한다고 본다(Mayr, 2002: 400).

기나긴 진화의 과정에서 인간의 마음은 것처럼 특정한 방식으로 편향되었으며, 그 결과 인간의 도덕적 본성, 즉 이타성이 성립하게 된 것이라고 한다. 윌슨은 그런 마음의 편향이 기본적으로 유전자가 인지 발달의 신경 회로를 만들고 그 속에 인간에 특유한 종적인 행동을 위한 어림법을 새겨놓음으로서 가능해졌다고 보고 그 어림법을 “후성 규칙”(epigenetic rule)이라 부른다(Wilson, 2005: 232; Ruse, 1998: 221-3).²⁾ 인간이 갖는 도덕 감정과 인지 능력도 후성 규칙의 산물이다. 도덕은 바로 그런 능력들이 만들어낸 문화인 것이다. 유전자는 정신 발달의 유전적 규칙성, 즉 후성 규칙을 통해 (도덕) 문화의 산출과 전달로 연결된다.³⁾ 그렇게 산출된 도덕이나 다른 문화가 새로운 환경을 낳고, 그 환경은 다시 새로운 선택을 낳는 과정을 윌슨은

-
- 2) ‘후성 규칙’이라는 말은 원래 ‘후생학’(epi + genetics = epigenetics)에서 파생되었다. 후생학의 의미는 왜딩턴(Conrad Waddington)이 1942년에 그것을 처음 만든 이래 부단히 변해왔다. 그 만큼 다중의미(polysemic)를 갖는 단어이다. 모렌지(Morange)에 따르면 그것은 원래 “유전학을 넘어서”(over genetics)를 의미하지만, “유전학 외”(besides genetics)의 뜻으로 더 많이 사용되어 왔다고 한다. ‘후생학’은 원래 유전학에 의해 적절하게 설명되지 못한다고 간주되는 문제들에 답을 제시하고자 사용되었다. 오늘날 후생학이라는 말은 “DNA 사슬 상의 변화(modification)에서 유래하지 않는 유전자 활동 상의 되물림가능한 변화(inheritable change in gene activity)를 유발하는 메커니즘”을 의미한다(Morange, 2006).
- 3) 윌슨의 후성 규칙은 일차과 이차 후성 규칙으로 나뉜다. 일차 후성 규칙은 감각기관에서 작동하는데, 자극을 거르고 압호화하고 두뇌가 그 자극들을 지각하도록 하는 자동적 과정 전체의 규칙성을 말한다. 이차 후성 규칙은 많은 양의 정보를 통합하는 과정에서 기능하는 규칙성으로 지각, 기억, 감정의 파편들을 끌어올려 우리의 마음이 특정 문화 단위에 대한 편향된 선택을 하도록 만든다. 일차 후성 규칙의 가장 기본적 속성들중에 대표적인 것으로 연속적 감각을 구분된 단위로 끊어 주는 속성이 있다. 이차 후성 규칙의 쉬운 예는 표정, 특히 미소다. 미소는 시각적인 비언어적 의사소통과 심리 발달의 보편적 편향이다. 그러나 정신 활동의 최고 수준에서 나타나는 복잡한 이차 후성 규칙들은 “구상화”(reification) 과정에서 따라 나온다고 한다. 구상화란 복잡한 현상이나 아이디어를 친숙한 대상과 활동에 비유하여 좀 더 단순한 개념으로 압축하는 절차 또는 세상에 질서를 창조하는 빠르고 쉬운 정신 알고리즘 내지 어림법(rules of thumb)이다. 대표적 예로 양분 본성(dyadic instinct)—사회적으로 중요한 것들을 양분하고자 하는 성향—이 있다(Wilson, 2005: 270-274, Lumsden and Wilson, 1981: 370-371).

문화와 유전자의 공진화라고 부른다. 도덕의 진화 도식은 다음과 같다.

유전자 → 후성 규칙(도덕 감정 → 도덕 의식) → 도덕 규범의 산출과 전달 → 다음 세대 유전자 → 새로운(변화된) 후성 규칙 → 새로운 도덕 규범의 산출과 전달 → …

윌슨은 콜버그의 도덕성 발달 단계에서 넷째 단계를 비비와 침팬지의 도덕성 수준으로, 다섯째 단계를 홍적세의 수렵 채취인 무리 단계의 도덕성 수준으로, 그리고 그 이후의 문명사회에서의 도덕성을 여섯째 단계 수준으로 간주했다. 그리고 여섯째 단계는 감정에 바탕을 둔 이전 단계와 구분되는 비생물학적 단계로 본다. 그러면서도 승고한 도덕 가치들의 문화적 진화가 스스로 방향을 설정하고 자체 추진력을 획득하여 유전적 진화를 대체할 수 있는 것으로는 보지 않는다. 그는 “인간 행동은—그것을 유도하고 지도하는 가장 깊은 감정적 반응 능력들처럼—인간의 유전 물질이 자신을 고스란히 보존해 가는 우회적인 방법”이며, 도덕은 유전 물질 외에 스스로를 설명할 수 있는 다른 어떠한 궁극적 기능도 갖고 있지 않다고 주장한다 (Wilson, 2000: 232).⁴⁾

윌슨 주장의 핵심은 도덕성의 중심에 이타성이 있으며 그것은 혈연 이타성과 호혜 이타성과 같은 진화적 이타성에 새겨진 유전자에서 유래한다는 점이다. 이 주장이 도덕성은 생물학적 진화의 산물인 인간의 뇌가 유발하는 행동 성향에서 나오며 인간 뇌의 존재는 기본적으로 유전자의 유전 정보에서 유래한다는 것을 뜻한다면, 그것을 부인할 사람은 없을 것이다. 그러나 만약 그 주장이 뇌가 유전자에 새겨진 정보의 틀 안에서 선택과 진화에 부응하는 행동 성향만을 산출한다는 것을 뜻한다면, 그것은 논리적 비약이다. 인간의 뇌는 비록 진화의 산물이지만, 진화는 뇌에게 생물학적 선택압을 상쇄할 수 있는 마음의 창발적 자유를 허용했다. 그 결과 인간의 마음은 그

4) 콜버그의 도덕성 발달 6단계: 처벌-복종 단계, 시장 교환 단계(이상 전인습 윤리), 착한 소년/소녀 단계, 법질서 단계(인습 윤리), 사회적 계약 단계, 보편적 원리 단계(후인습 윤리)(Kohlberg, 1969).

자체의 고유한 진화 논리에 따라 도덕과 문화를 만들어 냈다. 그렇기 때문에 사회생물학에서 제시하는 이정표를 따라가도 유전자에서 마음을 거쳐 도덕으로 나아가는 길을 찾을 수는 없다.

이 논문에서 필자는 바로 그 점을 진화적 이타성 개념을 중심으로 보여주려 한다. 작업은 세 단계로 이루어지는데, 먼저 논의의 예비적 단계로 필자가 최근(2008) 제시한 새로운 정의(definition)를 통해 기존의 진화적 이타성 개념을 둘러싼 의미론적 혼란을 걷어내는 작업을 한다. 그 다음으로 이타성에 관한 소버와 윌슨(Sober, 2000; Sober and Wilson, 1998)의 분석을 활용하여 유전자에서 마음으로 가는 길 및 마음에서 도덕으로 가는 길이 사회생물학자들의 견해와 달리 잘 연결되어 있지 않다는 점을 부각시킨다. 마지막으로 최근의 발생계 이론이 제시하는 비환원주의적 입장이 그와 같은 점을 뒷받침하고 있음을 주장할 것이다. 이런 작업을 통해 사회생물학의 환원주의적 입장에 대한 다소 과다한 세간의 환호에 경계심을 환기하고자 한다.

2. 이타성 개념의 혼란

진화생물학에서 이타성 개념은 매우 혼란스럽게 사용되어 왔다.⁵⁾ 그 원죄는 ‘이타성’이 원래 그리고 지금도 심리적 또는 윤리적 용어라는 점에서 유래한다. 그것을 진화론의 전문용어로 전환하면서 적절한 새 용어를 찾거나 만드는 대신 그 의미만 진화론적으로 바꾸었을 뿐 아니라 바꾸는 작업 역시 불완전했기 때문에, 지금과 같은 혼란을 보이고 있다. 혼란은 진화와 도덕을 연결하려는 사회생물학에서 특히 심하다.⁶⁾ 사회생물학에서 이타성 개념의 혼란스런 사용은 대체로 다음과 같은 양태를 보인다. 우선 유전자에

5) 소버(Sober 1998)와 정상모(2008)는 이타성의 일상적 용법과 진화적 용법 간에 관해 상세한 대비를 보여준다.

6) 사회생물학의 몇몇 개념들(근친상간, 성, 강간 등)의 모호한 사용에 관해서는 Sober 2000: ch. 7 참조.

서 도덕성에 이르는 전체 과정에서 서로 다른 종류의 이타성 개념을 각기 다른 의도로 사용하거나 일상적 의미의 이타성을 진화적 의미와 구분 없이 사용하곤 한다. 또 협동과 이타성은 서로 매우 다른 종류의 행동 경향인데도 종종 동일한 범주로 간주되곤 한다. 게다가 도덕의 기원이 협동에 있는지 이타성에 있는지를 분명하게 밝히지 못하고 있다. 이타성 개념을 중심으로 유전자와 도덕 간의 간격을 가늠해보려는 작업의 예비 단계로 진화적 이타성의 엄밀한 의미와 심리적 이타성과 이타주의의 구분에 관한 논의부터 시작한다.

2.1. 진화적 이타성

근래 진화 생물학에서의 이타성 논의는 세 가지 중요한 생물학적 이타성 개념을 확립하는 성과를 거두었다. 그 세 가지 개념은 혈연관계에 있는 집단 성원들 간에 성립하는 혈연 이타성, 비혈연 집단의 성원들 간에 성립하는 순수 이타성, 그리고 비혈연 집단의 성원들 간에 성립하는 조건적 이타성인 호혜 이타성이다.

혈연 이타성이란 혈연 선택(kin selection)을 가능하게 하는 형질이다. 혈연 선택이란 자신의 직접재생산 기회를 희생하고 대신 유전자의 일부를 공유한 혈연의 자손증식을 도움으로써 간접적으로 자신을 재생산하는 진화 방식을 말한다. 혈연 선택의 핵심은 개체의 진화적 생존 전략이 개체적응도만을 고려하는 단선적 방식에 의존하는 것이 아니라, 때로는 포괄적응도도 고려하는 이원적 방식으로 이루어진다는 점이다. 즉 자신의 개체적응도가 포괄적응도보다 높을 경우에는 통상적인 다윈적 이기적 개체로 행동하지만, 포괄적응도가 개체적응도를 초과할 경우에는 이타적으로 행동하기도 한다는 것이다(Hamilton, 1963). 여기서 ‘이타적으로 행동한다’는 말은 ‘개체적응도 원리를 포기하고 포괄적응도 원리에 따라 행동한다’는 것을 뜻한다. 한편 ‘포괄적응도 원리에 따라 행동한다’는 말을 관련되는 형질의 입장에서 해석하면 ‘직접적으로 자신을 재생산하는 대신 혈연관계에 있는 개체를 통

해 간접적으로 자신을 재생산하려는 형질에 의해 유발되는 행동을 한다'가 된다. 진화적으로 이기성과 이타성을 구분하는 기준은 바로 이 재생산의 직·간접성이다.⁷⁾

이제 간접성을 통해 혈연 이타성을 재정의 하면 다음과 같다:

혈연 이타성: 혈연관계에 있는 타 개체들을 통해 간접적으로 자신을 재생산하는 형질에 의해 발현되는 표현형(또는 행동 성향)

혈연 이타성은 각 성원들이 다른 성원들의 혈연관계를 확인할 수 있는 집단 내에서의 이타성을 규정하는 데 편리한 개념이다. 모든 인류가 혈연관계에 있고 단지 근연도 상의 정도의 차이만 있을 뿐이라는 의미로 이 개념은 혈연관계를 넘어서 모든 집단에 보편적으로 적용될 수도 있다. 그래서 혈연 선택론은 이타성에 대한 개체주의적 접근법으로 인정받고 있다. 그렇지만, 근연도가 낮거나 혈연성을 식별할 수 없는 집단에 적용하기에는 개념적으로 또 인과적 설명으로 부자연스럽다. 즉 실제로 선택이 일어나는 인과적 과정을 제대로 반영하지 못한다.⁸⁾

소버와 윌슨의 집단 선택론(다수준 또는 형질집단 선택론이라 부른다)은 그러한 부자연스러움을 피하는 이타성 개념을 제시하고 있다. 순수 이타성 진화의 핵심은 충분한 수의 다른 이타적 개체들이 어떤 집단에 주기적으로 집중되어야 한다는 점이다. 그런 경우 이기적 개체들이 개별적으로는 이타적 개체보다 더 많은 이득을 보지만, 이타적 개체들의 수가 충분히 많기 때문에 개체군 전체적으로 볼 때 총 이득에 있어서는 이타적 개체들의 수혜 총량이 이기적 개체들의 그것보다 더 많아질 수 있다. 그렇게 되면, 비록 집단 내에서는 이기성이 선택되지만, 개체군 전체를 통해서는 이타성이 진화할 수 있는 것이다. 혈연관계일 필요도 호혜관계일 필요도 없는 집단의 성

7) 이 기준은 필자가 최근(2008) 제시한 입장으로, 상세한 내용은 해당 논문을 참조할 수 있다.

8) 소버와 윌슨(Sober and Wilson 1998: 31)은 그것을 “평균화의 오류”(averaging fallacy)라고 부른다.

원들 사이에서 대가 없이 이루어진다는 점에서, 형질집단 선택론에서의 이타성은 순수 이타성이라고 할 수 있겠다.

순수 이타성을 순전히 진화론적 어휘로 표현하면 다음과 같다:

순수 이타성: 집단 내 다른 개체들을 통해 간접적으로 자신을 재생산하는 형질에 의해 발현되는 표현형(또는 행동 성향)

한편으로 혈연 이타성은 순수 이타성의 특별한 경우, 즉 혈연관계에 있는 집단에서의 이타성을 나타낸다고 볼 수 있고, 다른 편으로 순수 이타성은 근연도가 0인 혈연들 간의 혈연 이타성으로 볼 수 있다. 따라서 양자는 수학적으로 서로 동일하고 상호 번역도 가능하다. 그러나 ‘혈연’이라는 개념의 의미론적 제약과 집단 간 선택의 중요성을 감안하면, 소비와 월슨의 주장처럼 순수 이타성은 좀 더 일반적인 종류의 이타성이라고 볼 수 있다 (Sober and Wilson, 1998: 110).

트리버스(Trivers, 1971)의 호혜 이타성(reciprocal altruism)도 형질의 입장에서 엄밀히 진화론적 어휘로 정의될 수 있다. 호혜 이타성은 ‘동일한 형질을 가진 상대를 통한 간접재생산을 대가로만 직접재생산이 가능한 형질에 의해 발현되는 표현형’으로 정의된다. 여기서도 간접재생산이 관여되지만, 상대의 그것과 맞교환한다는 조건에서만 그렇다. 따라서 호혜 이타성은 순수한 의미의 이타성이 아니라 ‘합리적 이기성’이나 ‘협동’(cooperation)을 의미한다.⁹⁾ 협동을 유발하는 형질은 동일한 형질을 가진 상대의 재생산을

9) 월슨과 두갯킨(Wilson and Dugatkin, 1992: 30)의 다음 구분법 참고(괄호 속은 필자의 새로운 정의):

- 1) 이타성: 자신의 적응도 감소/타자의 적응도 증가(타자를 통한 간접재생산 추구함)
- 2) 협동: 자신과 타자 모두의 적응도 증가(타자를 통한 간접재생산을 대가로 그 이상의 직접재생산 추구함)
- 3) 이기성: 자신의 적응도 증가/타자의 적응도 감소(직접재생산만 추구함)
- 4) 인색(Spite): 자신과 타자 모두의 적응도 감소(직접재생산도 간접재생산도 추구하지 않음)

돕게 만듦으로써 서로서로가 간접적으로 자신을 재생산하되 단 그것을 나의 직접재생산 효율을 높이기 위한 수단으로만 사용한다. 즉 재생산의 간접성은 오직 직접성을 최적화시키기 위한 수단으로만 활용된다.

어떤 경우든 진화적으로 볼 때, 모든 종류의 이타성에 공통적 요소는 재생산의 간접성 뿐이다. 그래서 우리는 진화적 이타성을 다음과 같이 정의한다:

진화적 이타성: 다른 개체들을 통해 간접적으로 자신을 재생산하는 형질에 의해 발현되는 표현형(또는 행동 성향)

이렇게 하여, 적응도 상으로 ‘남에게 이익을 준다’든지 ‘자신을 희생한다’는 식의 일상어적 표현이 없는 이타성의 엄밀한 생물학적 정의를 얻게 되었다. 진화적 입장에서 볼 때, 이타적 종이란 이제 ‘간접재생산 시스템을 갖고 있는 종’이라는 더 적절한 표현으로 바뀌어야 한다. 그렇게 함으로써 “진화적 이타성이 유기체의 이타적 행동의 계통발생적 원인을 구성한다.” (Sober and Wilson, 1998: 199)는, 의미론적으로 모호하고 혼란스러운 표현 대신 ‘간접재생산 시스템이 유기체의 이타적 행동의 계통발생적(또는 자연선택적) 원인을 구성한다.’는 식의 명료한 표현을 얻을 수 있게 된다.¹⁰⁾

세 가지의 이타성 개념을 통해 드러나듯이, 생물학적 이타성의 핵심 내포는 재생산의 간접성이다. 호혜 이타성에서 재생산의 간접성은 단지 직접성의 증대를 위한 수단적인 것에 불과하고 혈연 이타성은 그 간접성이 혈연관계에 의해 매개 또는 확인되기 때문에 둘 다 이타성의 일반적 정의로는 부족하다. 순수 이타성만이 이타성 일반의 정의로 손색이 없기 때문에 이하에서의 이타성 관련 논의는 순수 이타성을 중심으로 전개된다.

여기서 제시된 새로운 진화적 이타성 개념은 하나의 중요한 특징을 갖는

10) 마찬가지로 “만약 한 유기체가 자신의 적응도를 낮추고 다른 유기체의 적응도를 높인다면, 진화론적 의미로 그것은 이타적으로 행동한다.”(Sober and Wilson, 1998: 199)는 역설적 표현 또한 다음과 같이 바뀌어야 한다: “만약 한 유기체가 직접재생산 대신 간접재생산 전략을 따르는 행동을 한다면 그는 진화론적으로 이타적으로 행동한다.”

다. 그것은 바로 진화적 이타성이 일상적 이타성에서 우리가 흔히 떠올리는 통속 심리적 의미들과 의미론적으로는 무관하다는 점이다. 재생산의 간접성은 여러 가지 방식으로 실현 가능하기 때문이다. 그럼에도 불구하고 인류에게 공통적으로 관찰되는 이타적 심리 또는 행동 성향이 존재한다는 것도 사실이다. 양자 간의 그와 같은 무관성과 유관성이 다음 논의들에서 보게 될 유전자와 도덕 사이의 간격을 암시한다.

2.2. 심리적 이타성과 이타주의

앞 소절에서 논의했던 진화적 이타성이 특정 행동 성향(property of behavior)의 유형에 관한 용어인데 비해 여기서 다루는 심리적 이타성은 특정 동기 성향(property of motives)의 유형에 관한 용어이다. 사소하게 보이는 이 차이와 함께, 선택이 고려하는 것은 전자이지 후자가 아니라는 점도 염두에 두고 다음 논의를 보자. ‘A가 심리적으로 이타주의자이다.’는 표현이 무엇을 의미하는가? 일반적으로 우리는 ‘A가 타인들에게 이익을 주기 위해 자신의 손해를 감수하는 심리적 성향을 갖는다.’는 정도로 이해한다. 과연 그럴까? 소버와 윌슨(Sober and Wilson, 1998)의 분석은 이런 표현이 갖는 모호함을 잘 보여준다.

누군가가 굶주리는 아이들에 관한 방송을 보고 ARS로 천원을 기부하기로 한 경우를 생각해 보자. 기부하는 행동은 그 동기에 측면에서 세 가지로 구분할 수 있다. 첫째, 자신의 이익을 위한 수단이 아니라 순수하게 그 자체를 목적으로 굶주리는 아이들의 복지를 위한다는 이타적 동기가 있다. 물론 그러한 기부로 심리적 만족감을 얻겠지만, 만족 그 자체는 진정한 동기가 아니고 부수효과에 불과하다. 그 다음으로 죄책감의 회피나 만족감을 얻기 위한 이기적 동기가 있다. 마지막으로 두 가지 동기가 동시에 작용한 동기가 있는데, 이를 “복수동기”(pluralism)라고 하자. 세 번째의 경우는 또 이타심이 우세한 경우와 이기심이 우세한 경우의 두 가지로 나눌 수 있다. 각 동기에 충실한 사람들이 어떤 행동 양태를 보일지 알아보기 위해 나와 타인

들 사이의 손익관계 입장에서 행동을 분류하면 다음 네 경우가 나온다 (242-248):

- ① 내계도 타인에게도 이익인 행동
- ② 내계는 이익이고 타인에게는 손해인 행동
- ③ 내계는 손해고 타인에게는 이익인 행동
- ④ 내계도 타인에게도 손해인 행동

이들 네 경우의 행동에 대해 위의 네 가지 동기를 가진 사람들이 각기 취할 행동 선택의 선호도는 다음과 같다:

- 가) 순수 이기주의자: ①② → ③④
- 나) 순수 이타주의자: ①③ → ②④
- 다) 이기주의가 강한 복수동기주의자: ① → ② → ③ → ④
- 라) 이타주의가 강한 복수동기주의자: ① → ③ → ② → ④

‘가)’와 ‘나)’는 순수 이기주의자는 타자의 상태는 개의치 않고 오직 자신의 이익만 챙기고 순수 이타주의자는 자신의 상태는 개의치 않고 오직 타인의 이익만 챙기는 욕구를 가진다는 점을 나타낸다. 그것에 비해 ‘다)’와 ‘라)’는 나와 타인의 이익을 모두 챙길 때의 선호 순서를 나타낸다. 잘 알려져 있듯이, 심리적 이기주의는 “우리의 궁극적 욕구는 **모두** 자기-지향적이다(self-directed)”는 사고방식이다(202, 필자의 강조). 따라서 누군가가 드물게라도 “타인-지향적”(other-directed)인 궁극 욕구를 갖는다면, 그는 이타주의자에 속해야 한다. 이기주의만이 단일동기적 이론인데 비해 이타주의는 복수동기적 이론과 양립한다는 점을 알 수 있다. 일견 이기주의자인 것처럼 보이는 ‘다)’가 이타주의자인 까닭은 비록 자기-지향적 욕구가 타인-지향적 욕구보다 강하긴 하지만, 어차피 내가 손해인 경우 남에게라도 이익이 돌아가는 것을 바라는 궁극 욕구가 있기 때문이다. ‘다)’처럼 ‘타인들에게 이익

을 주기 위해 자신의 손해를 감수하는 심리적 성향'을 갖지 않아도 이타주의자로 간주될 수 있기 때문에, 이타주의가 단순히 '타인들에게 이익을 주기 위해 자신의 손해를 감수하는 심리적 성향'을 가짐을 의미한다고 주장하는 것은 맞지 않다. 이 말은 이타주의와 이기주의의 논의를 이 둘중 하나의 심리적 성향(나아가 행동 성향)을 고정적으로 갖는 것으로 몰고 가면 안 된다는 것을 함축한다. 따라서 A가 이타주의자라는 말은 그가 어떤 특정 상황에서는 순전히 타인의 복지를 위해 자신의 이익을 기꺼이 희생하겠지만, 다른 상황에서는 그렇지 않을 수도 있다는 것을 의미한다.

요약하자면, '심리적 이타성'은 우리 욕구의 궁극적 타인-지향성을 의미하지만, '심리적 이타주의'는 "우리는 [다른 때에는 모르지만] **때때로** 타인-지향적(other-directed)인 궁극적 욕구를 갖는다."는 사고방식이나 이론을 의미한다(Sober and Wilson, 1998: 228, 필자의 강조). 여기서 '궁극적'이란 말은 '자기-지향성의 타인-지향성으로의 환원불가능성'(irreducibility)을 의미한다. 또한 거기서 타인-지향적 궁극 욕구가 자기-지향적 욕구보다 더 강한지 약한지는 알 수 없다.¹¹⁾

그뿐만이 아니다. 타인-지향적인 욕구(선의의 의도를 갖는 것)도 두 가지가 있다. 하나는 타인이 자신이 원하는 바를 갖기를 원하는 것이고 다른 하나는 타인의 선호와 무관하지만 타인에게 좋다고 내가 생각하는 바를 그가 갖기를 원하는 것이다. "개인이 갖고 있는 유일한 궁극적 목적은 자기-지향적이다."⁽²²⁴⁾고 주장하는 이기주의도 마찬가지다. 자기-지향적 행동에는 자기 자신만 관련될 수도 있고 다른 사람의 욕구가 함께 관련될 수도, 즉 타인과 무관할 수도 있고 유관할 수도 있다. 따라서 자신의 복지에 대한 궁

11) 위의 도식에 따르면, 자기 이익과 타인의 복지가 일치하면 네 가지 선호도 구조는 모두 동일한 행동을 예측한다. 양자가 상치하면, 이기주의와 이기주의가 강한 다원주의가 동일한 예측을, 이타주의와 이타주의가 강한 다원주의가 다른 동일한 예측을 한다. 이 경우에도 만약에 행위자가 자기-희생을 피한다면, 이기주의와 이기주의가 강한 다원주의를 구분하지 못한다. 사람들의 행동만으로는 심리적 동기를 구분해 내기는 그만큼 곤란하다. 이기주의와 이타주의는 어떤 특정 욕구를 사람들이 갖고 있는가에 대해서는 말하지 않는다. 그들은 단지 사람들이 궁극적 목적으로 어떤 유형의 욕구를 갖고 있는지를 말할 뿐이다(Sober and Wilson, 1998: 248).

극적 욕구와 타인에 대한 궁극적 욕구 외에 고려해야 할 추가적인 욕구, 즉 자신과 특정 타인들 간에 어떤 관계의 획득에 관한 욕구가 있어야 한다. 소버와 윌슨은 그것을 “관계주의”(relationism)라 칭한다(Sober and Wilson, 1998: 231).

이처럼 심리적 이타성은 진화적 이타성은 물론 심리적 이타주의와도 분명한 의미상의 차이가 있다. 이러한 구분을 무시하고 이타성이 도덕의 본질이라 간주하고 그것의 진화론적 근거를 찾으려는 기도는 선결문제 요구의 오류를 범하는 일이라 할 수 있다. 도덕 원리를 가지려는 궁극 욕구를 오직 진화적 이타성에서만 찾으려는 태도의 문제점들은 다음 두 절에서의 본격적 논의를 통해 더 분명히 드러난다.

3. 유전자와 심리 사이

윤리학의 생물학화 가능성과 관련해서 윌슨은 다음과 같이 주장한다: “도덕 감정의 유전학이 윤리적 행동의 심리학적·생리학적 과정들의 유전성을 측정함으로써 가장 쉽게 접근할 수 있을 것이며, 좀 어렵더라도 마침내는 규정적 유전자(prescribing gene)를 확인함으로써 접근 가능할 것이다”(Wilson, 2005: 441). 사회생물학의 출발점이 도덕 감정을 성립시키는 유전자의 존재라는 가정을 확인시켜주는 대목이다. 그는 도덕이 후성 규칙의 산물인 도덕 감정을 통해 유전자와 연결된다고 주장한다. 간접재생산과 같은 계통발생적인 진화 전략이 인간의 이타적 행동과 같은 개체발생적 행동 양식으로 직접 연결될 수는 없기 때문이다. 마이어(Mayr 1961)가 행동 전략의 진화는 “궁극인”(ultimate cause)과 “근인”(proximate cause)의 두 수준에서 접근할 수 있다고 한 것과 상응한다. 하나의 행동 전략이 진화하려면 선택을 위한 최종 기제뿐 아니라 유기체로 하여금 최종 기제를 행동 수준에서 수행하게 하는 근인(近因)적 기제도 진화해야 한다. 담쟁이의 광합성이 진화하려면 담쟁이로 하여금 다른 방향이 아닌 한 방향으로 자라게 유발하는 어

편 기제가 담쟁이 내부에 있어야 하듯이 말이다(Sober and Wilson, 1998: 200):

광합성을 위한 선택 → 내적 기제 → 빛을 향해 성장함

이것을 이타적 행동 경향에 적용한다면 다음과 같이 될 것이다:

간접재생산을 위한 선택 → 내적 기제 → 이타적 행동

이타적 행동에 대한 이해를 위해서는 근인적 기제인 내적 기제에 대한 이해가 필수적이다. 내적 기제가 무엇인지를 알기 위해서 우선 고려해야 하는 중요한 구분이 있다. 그것은 의도적 행동을 유발하는 마음의 존재 여부이다. 마음에 대한 제대로 된 정의는 그 자체가 쉽지 않은 문제이지만, 여기서는 행동을 유발하는 욕구 능력 정도로만 이해한다. 인간의 이타적 행동은 욕구의 산물이다. 그러나 뇌충(Brain worm), 저독성 점액종 바이러스(Myxoma virus), 암컷 새끼를 더 많이 낳는 장수말벌(wasp) 등과 같은 유기체의 이타적 행동은 욕구의 산물이 아니다. 마음이 없어도 이타적인 유기체가 있을 수 있는 것이다. 따라서 이타적 행동도 마음의 산물인 경우와 마음 이전의 유기체의 어떤 적응방식인 경우로 나눌 수 있다. 이런 구분에 따르면 근인적 기제도 비심적 근인과 심적 근인으로 나눌 수 있다.

인간의 마음을 중심으로 진화적 이타성, 즉 간접재생산 형질에 의해 유발되는 행동 경향과 심리적 이타성 간의 관계를 알아보자. 마음은 인류로 하여금 상이한 상황에서 상이한 행동을 산출하게끔 하는 근사적 기제이다. 인간의 이기성이나 이타성은 심적 근인에 따라 작동한다. 궁극인과 근인의 두 원인 및 이타적 행동과 이기적 행동 사이에는 다음과 같은 네 가지 조합의 관계가 성립한다(Sober and Wilson, 1998: 202-204):

- ① 심리적으로 그리고 진화적으로 이기적인 행동

- ② 심리적으로 그리고 진화적으로 이타적인 행동
- ③ 심리적으로 이기적이지만 진화적으로 이타적인 행동
- ④ 심리적으로 이타적이지만 진화적으로 이기적인 행동

만약 ③이나 ④에 해당하는 표준적 사례들이 있다면, 진화와 심리 사이에는 일대일 대응관계가 성립하지 않는다는 결론이 나온다. 먼저, 제한된 음식이 있는 가운데 자신만이 최대한의 음식을 챙기려는 사람이 있다고 치자. 그의 욕구는 자기-지향적이기에 그는 심리적으로 이기적이고, 음식을 적게 가지려는 사람보다 개체의 직접재생산 효율(적응도)이 더 높기에 진화적으로도 이기적이다. 이런 사람의 행위는 ①의 예이다. ②의 예는 궁극적인 목적으로 자신이 속한 집단 내 다른 사람들의 복지를 위해 음식을 최소한으로 챙기려는 사람에서 볼 수 있다. 그의 욕구는 심리적 이타성이고, 그런 행위의 결과 그 자신의 직접재생산 효율은 낮아지고 집단 내 타 개체들을 통한 간접재생산 효율은 그 이상으로 높아지기 때문에 그것은 또한 진화적 이타성이기도 하다.

어떤 사람이 자신의 집단 내 모든 이들의 복지를 위해 항상 최소한의 음식만 취하는데 그 이유가 그렇게 하는 것이 자신을 기쁘게 하기 때문이라면 어떻게 되는가? 이 경우 행위자의 행위는 자신의 직접재생산 효율은 낮추고 집단 내 타 개체들을 통한 간접재생산 효율을 그 이상으로 높이기 때문에 진화적으로 이타적이지만 심리적으로는 이기적이다. 집단의 방어를 위해 요새에 말뚝을 박는 일을 돕고자 하는 사람의 경우도 마찬가지이다. 그가 일을 돕고자 하는 궁극 목적은 자신의 안전이다. 그렇기 때문에 그의 행동은 심리적 이기성의 결과이다. 그렇지만, 일을 돕는 행위는 무임승차에 비하면 자신의 직접재생산 효율을 낮추는 행위이다. 요새의 안전으로 모두가 다 동일한 이익을 볼 수 있지만, 일을 하는 자들은 더 많은 비용을 지불해야 하기 때문이다. 따라서 일을 돕는 행위는 진화적으로는 이타적이다. 이들 행위는 ③의 예이다. ④와 같은 행동은 아이들의 복지를 위해 헌신적으로 배려하는 부모에서 볼 수 있다. 그런 행동은 자신 외에 누군가에 대한 순수한 배려이

기 때문에 이타적인 심리적 욕구의 산물이다. 그렇지만 그런 부모가 그렇지 않은 부모들에 비해 더 성공적인 직접재생산 효율을 보인다면, 그런 행동은 진화적으로 이기적인 행동이다.¹²⁾

이상 네 가지 행동의 예를 정리하면 다음과 같다(204):

		진화적	
		이기성	이타성
심리적	이기주의	① S는 가능한 한 많은 음식을 챙기기를 원한다.	③ S는 요새 건설을 돕기 원한다.
	이타주의	④ S는 자신의 아이들의 복지를 위해 헌신한다.	② S는 집단 내 다른 개체들의 복지를 배려한다.

표 안의 ③과 ④에서 볼 수 있듯이, 심리적 이기성·이타성과 진화적 이기성·이타성은 서로 어긋날 수 있다. 그래서 소비와 월슨은 이 대비를 통해 다음 사실을 지적한다.

모든 동기들은 두 개의 서로 아주 다른 각도에서 접근될 수 있다. 한 동기가 진화적으로 이기적 또는 이타적 행동을 산출했다는 사실이 그 동기가 심리적으로 이기적인지 이타적인지 하는 문제를 해결해 주지는 않는다. 또한 한 행동이 주어진 한 심리적 동기에 의해 산출되었다는 사실은 그 행동이 왜 진화했는가 하는 물음을 해결해 주지도 않는다(Sober and Wilson, 1998: 205).

그뿐만이 아니다. ‘진화적으로 이타적’이라는 말이 ‘관련 형질의 진화가 간접재생산으로만 가능함’을 뜻하기 때문에 ‘③’, 즉 ‘심리적으로 이기적이지만 진화적으로 이타적인 행동’은 해당 행동을 통해 얻는 포괄적응도가 개체적응도를 능가한다든가 집단 내 선택압 손실분을 집단 간 선택압 이익분이 초과할 경우에만 성립한다. 또한 ‘④’, 즉 ‘심리적으로 이타적이지만 진

12) 이타성을 재생산의 간접성으로 정의하면 직계자손에 대한 이타성은 모두 진화적 이기성으로 간주된다. 직계 자손의 복지에 헌신하는 것은 결과적으로 자신의 직접 재생산 효율을 높이는 것이기 때문이다.

화적으로 이기적인 행동'도 그런 행동을 통해 행위자의 적응도(재생산효율)를 비교 대상의 그것 이상으로 높일 수 있어야 성립한다. 결국 ③과 ④에 해당하는 행동을 결정하는 것은 심리와 무관한 요소, 즉 집단 내 또는 집단 간에 성립하는 상황이라는 말이다. 이 점은 진화적 이기성과 이타성이 행위자의 심리와 무관하게 성립할 수 있다는 앞의 결론을 더욱 강화한다.

심리와 진화 사이의 그러한 관계는 행동을 유발하는 원인(=근사적 기작)과 그 원인을 유발하는 원인(선택되는 유전 형질)이 작용하는 기작이 서로 다른 데서 유래한다. 기작만 다를 뿐 아니라, 위계적으로도 동등하지 않다. 즉 최종 산물인 행동은 근사적 기작에 대해서는 궁극적이지만, 선택되는 형질에 대해서는 수단적이다. 따라서 그 행동은 관련 형질을 달성하는 여러 행동중 하나에 불과하다. 예컨대 “이타적으로 행동하려는 심적 욕구는 심리적으로는 궁극적이나 진화적으로는 수단적”(201)이고 그런 욕구 외에도 동일한 진화적 목적, 즉 관련 형질을 간접재생산하려는 목적에 봉사하는 다양한 수단이 있을 수 있다는 말이다. 결론적으로 말해, 인간의 심리적 이타성과 이타주의의 전모가 진화적 이타성만으로 밝혀질 수는 없다. 월슨은 ‘진화적’(evolutionary)을 ‘유전자적’(genetic)을 포함한 의미로 사용하고 있기 때문에, 이 결론은 유전자와 심리 사이에는 창발적 비약을 허용하는 위계적 차이가 있음을 함축한다.¹³⁾

4. 심리와 도덕 사이

사회생물학은 도덕성과 이타성을 행동의 수준과 동기의 수준 모두에서 종종 같은 것으로 간주한다. 첫째 등식은 도덕성이 항상 우리로 하여금 타

13) 대부분의 표현형 진화 모델은 관련되는 유전자의 존재에 대해 어떤 증거도 주지 않는 것으로 알려져 있다. 피셔(Fisher, 1938)의 성비 연구나 파커(Parker, 1978)의 톱파리 교미시간 연구가 보여주듯이, 관련 유전자를 전혀 상정하지 않고도 특정 형질의 유전에 관해 충분히 설명할 수 있다고 한다(Sober, 2000). 이에 관한 논의는 여기서 다루지는 않는다.

인들을 위해 자기-이익을 희생하도록 요구한다는 것이고, 둘째 등식은 이타적 욕구에 의해 유발된다는 것은 도덕 원리에 의해 유발된다는 것과 동일하다는 것이다. 그런데 우선 만약 도덕 원리가 그 동기가 되지 않고서도 이타적 욕구를 가질 수 있다면, 즉 한 개인이 도덕 원리에 의해 추동되지 않고서도 이타주의자가 될 수 있다면, 이타성이 도덕성의 본질이라는 주장 또는 ‘도덕성=이타성’이라는 등식은 약화될 것이다. 이것이 소버와 윌슨(Sober and Wilson, 1998: 238-240)이 취하는 입장이다.

도덕 원리에 의해 추동되지 않는 이타성의 예로 부모의 자식에 대한 헌신을 들 수 있다. 부모가 이기적인 이유에서가 아니라 아이들의 행복 그 자체를 위해서 헌신하고파 한다면 그것은 일종의 이타성이다. 부모들은 그와 같은 심리적인 이타적 욕구를 어떤 도덕 체계에 의거하지 않고서도 가질 수 있다. 그리고 그러한 심리적 이타성은 도덕성에 핵심적인 일반성이나 보편화가가능성(universalizability)을 결하고 있다는 점에서 도덕성과 본질적으로 구분된다.¹⁴⁾ 자식에 대한 순수한 헌신이 진화적으로는 이타성이 아닌 이기성의 일종이라는 점에 입각하여 이 주장을 반박할 수도 있을 것이다.¹⁵⁾ 그러나 그것은 논점을 일탈한다. 여기서 문제가 되는 것은 심리적 이타성이며, 그것은 정의상 자기 아닌 다른 사람에 대해 헌신하는 것이기 때문이다.

앞의 경우와 반대로 만약 이타성 없이도 도덕성을 실천할 수 있다면, 즉 이타적 동기 없이도 도덕 원리에 따를 수가 있다면, (심리적) 이타성이 도덕의 본질이라는 주장은 더욱 심각한 도전에 처하게 된다. 소버와 윌슨은 이타적 심적 동기 없이도 도덕성을 실천하는 경우의 좋은 예로 의무주의적 도덕성과 종교적 신념, 특히 유신론적 확신에 근거한 도덕성을 들고 있다. 의무주의는 행동의 동기가 자신이든 타인이든 사람들의 행복과 무관하다고 간주한다. 따라서 이타적 욕구(심적 동기)와 무관하게 도덕적으로 올바른 행동을 할 수 있다. “타자-지향 욕구(other-directed desires)는 감정적 통로의

14) 정상모(1998)는 자식에 대한 배려의 도덕성이 보편화될 때 자기 “자기-파괴적으로”(self-defeating) 되어버리기 때문에 보편적 도덕 원리가 될 수 없음을 논증하고 있다.

15) 주‘12)에서 밝히고 있듯이, 이 심리적 이타성은 그러나 진화적으로는 이기성이다.

중재 없이도 발생할 수 있다.”(237)는 말이다.¹⁶⁾ 단지 신이 그렇게 하도록 명하기에 어떤 행동을 해야 한다고 확신하는 사람들의 경우도 마찬가지다. 의무주의적 도덕 원리나 유신론적 도덕 원리가 잘못된 원리이기 때문에 이런 주장은 설득력이 없지 않느냐는 비판에 대해 소버와 윌슨은 자신들의 입장은 그런 원리의 옳고 그름 이전에 지금까지 많은 의무주의자들과 유신론 신봉자들이 있어 왔다는 엄연한 사실만으로도 자신들의 주장은 충분히 지지된다고 응답한다.

이타성과 도덕성 사이가 더욱 멀어지게 되는 것은 이타적 동기를 갖는 행위가 도덕적으로 나쁠 수도 있을 뿐 아니라, 반대로 이기적 동기로 유발된 행위가 도덕적일 수도 있다는 점 때문이다. 누군가를 돕는 일이 제3자에게 해가 되는 경우가 전자의 대표적인 경우다. 병든 친구를 돕고 싶은 마음이 아주 커서 다른 친구를 속여서 돈을 얻어 병든 친구를 돕는 경우, 그 행동은 이타적 욕구나 관심의 산물이지만, 도덕적으로 정당화되기 어렵다. 나치 권력에 선의로 헌신적으로 협조했던 사람들, 국가를 위한다는 나름대로의 선의로 유신 독재의 꼭두각시 노릇을 한 사람들 등은 심리적 이타성이 도덕적 악을 보증하는 데 기여할 수 있음을 보여주는 좋은 사례들이다.

이기적 동기로 유발된 행위가 도덕적일 수도 있는 경우로 소버와 윌슨은 다음과 같은 예를 든다. 갑과 을 두 사람이 목숨이 위중한 상태에 있는데, 특효약이 딱 일인분 밖에 없고 또 공교롭게도 약을 복용할 사람을 정할 결정권이 갑에게 있다고 하자. 만약 갑이 자신의 이기적 욕구에 따라 그 약을 자신이 복용하기로 결정했다면, 그는 도덕적으로 비난받을 것이다. 그러나 만약 공리주의자가 갑의 행동을 판단한다면 다른 결과가 나올 수도 있다. 공리주의에 따르면 약은 전체적으로 더 큰 효과를 발휘하는 사람에게 가야 한다. 그런데 만약 갑이 바로 그런 사람이고, 그렇기 때문에 갑이 그 약을 자신이 사용하기로 결정했다면, 그의 결정은 공리주의 원리에 충실하다. 갑의 행동을 그런 식으로 계산하는 사람에게는 갑의 이기적 행동도 도덕적으

16) 주지하듯이, 칸트에 따르면 선한 것을 단지 그것이 선하다는 이유만으로 행하려는 의지에 따라 행동하는 것이 도덕적으로 옳은 행위이다.

로 옳은 행동으로 둔갑하는 것이다.

도덕성이 심리적 이타성이나 이기성과 무관할 수도 있다는 소비와 월슨의 주장은 황금률을 비롯한 주요 도덕 이론들이 취하는 근본 입장과도 일치한다. 예컨대 황금률은 자신과 다른 사람을 공평하게 대하라는 입장을 취하고 있지 이기적이거나 이타적이기를 요구하지는 않는다. 모든 사람을 각각한 몫으로 해서 최대다수의 최대 행복을 주장하는 공리주의 도덕 원리나 네 의지의 준칙이 항상 동시에 보편적 입법의 원리에 타당하도록 행동하라(보우이의 말로 표현하면, “자기 자신을 예외로 만들지 말라.”[Bowie, 1999: 17])고 주장하는 의무주의 도덕 원리도 정확하게 그렇다. 그렇다면 도덕성은 주로 호혜 이타성적 욕구와 관련된다는 말인가 하고 의문을 제기할 수 있다. 그러나 그런 의문은 2절에서 봤던 호혜 이타성의 정의를 떠올리면 해소된다. 호혜 이타성은 나와 타자를 똑 같이 대하려는 욕구가 아니다. 그것은 자신을 특별히 대하는 데 필요한 만큼만 타자를 이용하려는 욕구와 관련되기 때문이다.

이상과 같은 논의의 결론은 이렇다: 도덕성은 이기적 욕구의 명령과 일치하는 (따라서 이타적 욕구와 충돌하는) 함축도 갖고 이타적 욕구의 명령과 일치하는 (따라서 이기적 욕구와 충돌하는) 함축도 갖는다. 그것은 도덕 원리가 비개인적(impersonal) 특성을 갖는 데 비해 이타적·이기적 욕구는 개인적 특성과 관련되기 때문이며, 따라서 일반적 도덕 원리를 지지하는 궁극적 욕구는 이타적으로 간주되어서도 이기적으로 간주되어서도 안 된다(Sober & Wilson, 1998: 231). 이 비판은 인간은 선천적으로 쾌를 선호하고 고통을 회피한다는 심리적 동기를 도덕 원리 도출을 위한 출발점으로 삼는 자연주의 윤리설(대표적으로는 공리주의)에도 물론 동일하게 적용된다.¹⁷⁾

이 점을 앞 두 절의 결론과 종합하면 이렇게 된다. 심리적 이타성의 본질을 진화적 이타성에서 찾을 수 없을 뿐 아니라 도덕성의 본질을 심리적 이타성에서 찾을 수도 없다. 그렇기 때문에 진화적 이타성에서 도덕의 본질을

17) 프란츠 부케티츠는 이러한 입장을 “<선>과 <악>: 자연은 이를 정하지 않았다.”고 요약하고 있다(Wuketits, 1999: 164).

찾는다는 것은 이중으로 곤란하다. 만약 진화적 이타성이 전적으로 유전자의 산물이라면, 심리적 이타성과 진화적 이타성은 정확히 일치해야 할 것이다. 꿀벌 사회에서처럼 말이다. 심리적(일상적) 이타성과 진화적 이타성 간의 일치와 불일치의 이러한 복합적 관계 그 자체는 이타성이 유전자와 문화를 포함하는 복합적 실재의 산물일 가능성을 함축한다. 다음 절에서는 보게 될 유전자와 도덕 사이의 위계적 특성이 갖는 구조적인 문제점과 최근의 발달계 이론이 보여주는 형질 발현의 창발적 특성은 이런 결론을 지지하는 좋은 근거가 되고 있다.

5. 창발성과 발달계 이론

에드워드 윌슨에 따르면, 유전자와 문화로서의 도덕을 매개하는 인간의 마음은 후성 규칙에 따라 작동한다. 맹목적 이타성과 목적성 이타성 같은 심리 작용이 바로 그런 후성 규칙의 산물이다. 후성 규칙은 한편으로 도덕으로 하여금 자율적 문화로 발전하게 허용하면서 다른 편으로 “유전자로부터 발흥하여 유전자의 검인을 영원히 간직”해야 하는 상반된 역할을 해야 한다(Wilson, 2005: 289). 그렇게 하려면 후성 규칙은 일군의 DNA 분절로 된 유전자의 구문론적(syntactical) 디지털 신호 체계를 마음의 의미론적(semantic) 신호 체계로 바꿔주는 일종의 교환기 역할을 해야 한다. 후성 규칙이 제공하는 의미론적 신호 체계가 있어야 도덕의 화용론적 신호 체계가 가능해 진다. 과연 그런 교환기가 가능한가?, 또 가능하다면 일군의 DNA 분절의 구문론적 신호 체계는 도덕의 화용론적 신호 체계와 어떤 연결을 가질 수 있는가?

디지털 컴퓨터를 예로 들어 보자. 프로그램, 이를테면 한글워드 프로그램이 깔아지기 전의 기초 상태의 컴퓨터의 인지 체계가 ATGC의 신호 체계로 된 DNA 분절에 해당한다면, 컴퓨터에 심어진 한글워드 프로그램의 인지 체계는 후성 규칙에 해당할 것이고 워드로 쓰고 있는 이 논문의 인지 체계

는 도덕에 해당할 것이다. 유명한 존 설(John Searle, 1984)의 “중국어 방”(Chinese room) 논증에서 볼 수 있듯이, 컴퓨터의 경우 기본적 작동 프로그램은 물론 한글워드 프로그램까지도 의미를 발생시키지 않는다. 의미는 내가 자판을 두드리려 글자들을 유의미하게 조합했을 때 비로소 발생한다. 후성 규칙이 한글워드 프로그램에 해당한다면, 도덕을 유의미하게 편향되게 해 줄 수 있는 교환기, 즉 후성 규칙은 불가능하다. 한글워드 프로그램이 타이핑 기법에 어떤 편향을 줄 수는 있어도 논문의 내용에는 어떤 편향도 초래할 수 없듯이 말이다. 한글워드 프로그램이 후성 규칙이 아니고 논문 작성법이 그것이라고 반론을 제기할 지 모르겠다. 그러나 논문 작성법은 교육을 통해 전달되고 발전되는 것이지 컴퓨터 디지털 신호 체계나 한글워드를 통해 전달되는 것이 아니기 때문에, 그것은 반론이 될 수 없다.

한글워드 프로그램이 없는 기초 상태의 컴퓨터 인지 체계와 내가 쓰고 있는 논문의 인지 체계를 연결하는 것은 양자 모두 동일한 디지털 신호 체계를 바탕으로 한다는 구문론적 사실뿐이다. 논문의 내용에 관한 어떤 요소도 기초 상태의 컴퓨터가 말해 줄 수 있는 것은 없다. 그 점은 한글워드 프로그램도 마찬가지이다. 결과적으로 유전자와 도덕 사이에 후성 규칙이라는 우회로를 추가해도 도덕의 내용에 관해 유전자가 구체적으로 할 수 있는 말은 없다.¹⁸⁾

유전자에서 후성 규칙을 거쳐 도덕 사이에 성립하는 의존 관계—만약 그런 것이 있다면—는 구문론적 토대 관계이지 의미론적 또는 화용론적 상호관계는 아니다. 후성 규칙은 유전자에 그리고 도덕은 (부분적으로) 후성 규칙에 구문론적으로 의존하지만, 의미론적으로 도덕과 문화는 자율적이다.¹⁹⁾ 원추세포에서 뇌의 최후부 신경 세포들로 연결되는 과정에서 작동하는 색 인지의 후성 규칙은 인간의 타고난 색 지각 체계를 구문론적으로 결정해 주지만 미술 문화에 대해 우리가 주로 알고 싶어 하는 내용들에 관해 이야기

18) 정상모(2009)에 제시된 논증.

19) 이것은 환원주의에 대한 표준적인 비판과 유사하다. 세 가지 종류의 환원에 관한 아이얼러(Ayala, 1974)의 구분 참조. 좀 더 본격적 논의는 Sarkar(1998) 참조.

해줄 수 있는 바는 별로 없다. 유전자에서 후성 규칙을 통해 도덕으로 가는 길에서 도덕은 이중으로 자유(창발성)를 얻는다. 문화가 가능한 것은 분명 수정란 속의 유전자가 발현하여 성장한 개체들이 사회를 이루어 살기 때문이다. 그런 의미로 도덕이 “유전자로부터 발흥하며 유전자의 검인을 영원히 간직한다.”는 윌슨의 말에는 전혀 특별할 것이 없다. 게다가 만약 그의 말처럼 도덕을 포함한 다양한 문화 사이의 거의 모든 차이들이 대개 역사와 환경의 산물일 개연성이 높다면(Wilson, 2005: 258), 그 사실만으로도 도덕과 문화에 대한 이해를 위해서 생물학이 해야 할 역할은 생각보다 크지 않음을 알 수 있다.²⁰⁾

최근에 유전자 환원주의의 대안으로 급부상하고 있는 발생계 이론(Developmental system theory)은 이런 결론을 뒷받침한다. 발생계 이론은 기존의 유전자 중심 진화론을 발생계 중심으로 다시 정식화하려는 시도이다. 발생계는 특정한 진화 계통의 생활사를 산출하도록 상호작용하는 물리적 자원들의 전체 체계를 의미하는데, 거기서는 기존의 중요 개념들이 새로운 의미를 갖는다. 이를테면 계통은 유사한 계별 생활사가 인과적으로 연결된 서열을, 유전은 발생 자원이 후속 계통들로 믿을만하게 재생산 되는 것을, 자연선택은 발생계의 대물림가능한 변이들의 차별적 재생산을, 그리고 진화는 발생계의 개체군 구성물이 시간이 지남에 따라 변화하는 것을 의미한다. 간단히 말해 발생계 이론은 기존에 유전자가 있던 자리에 발생계를 대입하여 성립하는 유전학이다(Oyama, 2000).

이 이론의 핵심 특징은 발생계 “등가성 논제”(parity thesis)로 잘 알려져 있는데, 그것에 따르면 유전자를 포함한 여러 자원들이 발생 과정에서 인과

20) 사회생물학은 근친상간 회피가 이계교배를 위한 선택을 가능하게 하는 유전자에서 발흥하여 유아기 때 친숙해진 이성에 대한 성적 매력 상실을 초래하는 후성 규칙을 거쳐 형성된 문화라고 주장한다. 그러나 그런 후성 규칙은 교배기회 늘리기를 위한 선택을 가능하게 하는 유전자에서 발흥할 수도 있으며, 유아기 때 친숙해진 이성에 대한 매력 상실을 초래하는 후성 규칙은 근친상간 회피 문화가 아니라 다른 문화를 가능하게 할 수도 있다. 근친상간 회피 문화도 다른 후성 규칙의 산물일 가능성도 있는 것은 물론이고. 사정이 이렇다면 근친상간 회피 문화의 유전적 근거를 찾는 일의 의미는 무엇인가?(Sober 2000, ch. 7 참조).

적으로 동등한 역할을 담당하고 있으며 또한 각 자원은 각기 독자적으로 발휘되지 않고 다른 많은 자원들과의 상호작용 속에 발휘된다.²¹⁾ 예컨대 DNA 정보가 정상적인 세포로 발달하려면 미세소관(microtubule)으로 구성된 기저체와 중심립(centriole), 세포질의 적절한 화학 농도, DNA 메틸화(DNA의 구성 염기에 메틸기를 첨가하는 작용) 양상, 세포막, 그리고 세포 내 소기관 등과 같은 자원들의 공동 작업이 필요하다. 당연히 유전자는 다른 발달 요소들과 다른 종류의 특별한 부류의 마스트 분자를 구성하지 않으며, 진화도 단지 유전자 빈도의 변화가 아니라 발달 자원 전체 스펙트럼의 빈도 변화로 측정된다. 또한 각 자원은 발생 과정에서 세포별·시간별로 다르게 번역된다. 발생계를 구성하는 모든 다른 요소들이 그러하듯이, 유전자 코드는 단백질의 기본 구조 외에 어떤 다른 결과를 상술하는 표상적 자원을 전혀 갖고 있지 않다. 한 마디로 일련의 DNA 분절(유전자) 내에 미리 잘 짜인 프로그램이 있고 발생 과정에서 유전자가 그 프로그램을 작동시키는 식의 발생은 없다는 말이다(Keller, 2005).²²⁾ 이 분야의 선두 주자중 한 사람인 오이야마는 발생이 유전자에 미리 존재하는 프로그램의 실행이 아니고 “구성적 상호 작용, 변형(transformations), 그리고 창발(emergences)의 우연한 시리즈”라고 규정한다(Oyama, 2000: 136).

발생계 이론의 이러한 전체론적 특성은 유전자와 표현형 간의 전통적인

21) 이를 “상호작용적 구성”(interactive construction)이라 한다(Griffiths and Stotz, 2000: 36).

22) 이른바 이보디보(진화발달생물학)도 비슷한 입장이다. 양자는 발달을 많은 요인들 간의 상호작용을 통해 적응적 표현형을 산출하는 하나의 체계로 이해해야 한다는 견해를 공유한다. 따라서 세부 표현형적 특징을 하나 혹은 단순하고 부가적인(additive) 방식으로 상호작용하는 적은 수의 유전자 책임으로 돌리는 데 반대한다. 그러나 발달계 이론은 행동 발달의 경험적 요인들에 관한 심리생물학적 연구에 중점을 두고 발달 체계의 포괄적 정의를 추구하다 보니 후성 유전의 진화적 잠재성과 유기체-환경 공진화를 탐구하는 다른 연구 전통에 수렴하는 데 비해, 이보디보는 발달 유전학 상의 발견에 밀착하고 발달 체계를 유전체의 창발적 특징(emergent feature of the genome)으로 이해한다(발달계 이론은 유전체 밖의 요소들도 진화하는 발달계의 부분으로 보는데 비해 이보디보는 유전체 내에서의 발달만 인정한다.). 그런 점에서 양자는 상보적일 수 있다(Griffiths and Gray, 2005: 421).

이분법적 사고방식을 거부하게 한다. 천성과 양육, 유전자와 환경, 유전자와 문화와 같은 식의 이분법이 틀렸을 뿐 아니라, 그런 이분법에 대한 대안적 설명으로 제시된 양자 간 상호작용론마저도 거부한다. 모든 형질은 많은 발생 자원들 간의 상호작용으로 인해 산출되며, 이분법적 사고는 이런 상호작용자들을 나누는 많은 방식중 하나에 불과하다고 본다. 유전자도 수많은 발생 자원들중 하나일 뿐이며, 형질도 부모로부터 자손에게 그저 전달되는 것이 아니라 발생 과정에서 재구성된다. 그 과정에서 수많은 발생 자원들이 함께 작용한다. 또한 그 각각은 부모로부터 자손으로 가는 수많은 인과적 통로중 어느 하나에서 특정 시간과 장소에서 고유한 방식으로 작동하며 발생의 각 단계와 시점에 따라 새로운 의미를 갖는다. 발생 과정의 그러한 역동적 연속성 속에서 창발성이 생성된다(Oyama, Griffiths & Gray, 2001).

한편 발생계 이론은 전체 발생 과정을 증시하기 때문에 되물림(inheritance)도 전통적인 표현형과 후성 규칙을 넘어서 “각 세대에 현존하면서 삶의 순환을 재구성하도록 도와주는 모든 것”(whatever is present in each generation and helps reconstruct the life-cycle)으로 획기적으로 확장된다(Griffiths and Stotz, 2000: 35). 인간의 문화도 당연히 포함된다. 그렇기 때문에 발달계 이론은 인간성(human nature)도 유전자에서 문화에 이르는 수많은 자원들로 이루어진 전체 발달계의 산물로 본다. 인간성의 발달은 유전적 소인 뿐 아니라 문화라는 외적 골격에도 의존하는데, 그 의존도가 여느 생물 종과도 비교할 수 없을 정도로 심대하다. 대표적인 발생계 이론가인 그립피스와 스토츠는 인지 상의 종-특유의 규칙성에 대한 논의를 통해 발생계의 그런 특성을 보여주고 있다. 그들에 따르면, 종-특유의 규칙성은 발달계의 다양한 요소들 간의 호혜적 상호작용의 결과로 창발한다. 인간성이 진화의 산물인 것은 틀림없지만, 그렇다고 해서 그것이 어떤 제한된 방식으로 발달하도록 미리 약정되어 있는 것은 아니다. 우리가 조상으로부터 물려받는 것은 마음이 아니라 특정한 방식으로 마음을 발달시킬 능력이다. “수정란은 언어 획득 장치도 **통속 심리학의 기본 교의(tenets)**에 관한 지식도 갖고 있지 않다. 이들 특성들은 마음이 성장하면서 생기게 된다” (Griffiths and

Stotz, 2000: 31, 필자의 강조). 이 말은 인간의 이타적 본성과 같은 것도 유전자의 산물이 아니라 유전자에서 문화에 이르는 발생계 전체의 창발적 생산물로서, 유전자나 문화 등과 같은 계의 특정 요소로 환원될 수 없는 독자적인 실재라는 점을 함축한다.

그립피스와 스토츠는 어린이의 행동에 대한 텔른의 연구 결과(Thelen, 1995; Griffiths and Stotz, 2000: 38에서 재인용)에서 작은 증거를 찾는다. 딸랑이를 갖고 처음으로 소리를 내며 놀게 된 어린아이의 예를 보자. ‘빨기-위한-대상’이라는 도식만을 갖고 있던 아이에게 이 새로운 행위를 통해서 ‘소리내기-위한-대상’이라는 새로운 도식이 창발한다. 아이의 행위는 행위하는 시점의 신경계의 구조에 의해 제약되지만, 동시에 그 구조를 수정하고 이전 인지 동력에서는 불가능한 행동의 출현으로 나아간다. 3, 4세 아이들의 손닿기와 물건 쥐기 실험도 동일한 결론을 보여준다. 개체 발생 중의 행동과 인지 및 그것들의 변화는 정해진 두뇌 구조나 기호 또는 유전자 암호로 계 내에 미리 재현되어 있지 않다. 생각과 행동은 개체가 지금 하려는 일과 개체의 내적 동력의 함수로 생기는 역동적 과정의 산물이다.

그들은 헨드릭스-잔슨의 연구도 좋은 예증으로 제시한다. 실험에 따르면 4개월 정도의 아기가, 그전에는 늘 엄마의 얼굴만 강하게 응시했는데, 어느 날 더 먼 다른 곳을 보기 시작했다. 비록 아이의 응시는 사실상 특별한 대상을 향하지 않았지만, 엄마에 의해 그것은 다양한 대상에 대한 흥미로 해석된다. 바로 이 순간에,

아마도 아기는 어떤 대상에 대해 초보적인 내밀어서 쥐는 자세를 취했을지 모른다. 엄마는 그것을 어떤 대상을 취려는 시도로 해석하고는 아기의 행동을 완성하려고 개입한다. 그렇게 함으로써 엄마는 아기의 원초적 행동 양식을 대상에 대한 행위로 구체화시킨다. [이런 방식으로] 엄마는 행동 양식에 지향적 의미를 부여하는 틀 또는 골격(scaffolding)을 제공한다(Hendriks-Janson, 1996: 273, Griffiths and Stotz, 2000: 39에서 재인용).

이처럼 우리의 인지 과정은 언어와 다른 표상 체계를 포함한 고도로 조직

화된 문화적 환경의 존재를 가정하고 이용한다. 두뇌가 화학적 전기 신호 밖의 능력, 즉 마음을 갖게 되는 것은 바로 그런 외부의 의미론적 골격의 도움에서 비롯된다. 이런 식으로 발달의 통로는 그 계의 창발적 동력에 의해 지배된다. 그러한 동력의 작용 방식은 시간에 따라 변하는데, 부분적으로는 발달하는 계가 그 자체의 매개변수에 미치는 영향이 그 원인이다.

발생계 이론이 인간의 이타성을 설명할 수 있으려면 발생 과정에 대한 생물학적 연구 상의 진전이 한참 더 진행되어야 하겠지만 지금으로서도 앞의 두 절에서의 논의에 대한 좋은 뒷받침이 된다. 인간성이 그리고 인간의 마음이 광의의 발생계를 구성하는 많은 요소들의 협동 작업의 결과라는 점만으로도 인간성의 중요한 부분을 차지하는 이타성 또한 그러한 결과라는 점은 지지되기 때문이다. 순수 이타성이 독특한 집단 구조 및 변화 과정에 의해서만 선택되고 진화될 수 있다는 점도 발생계 이론의 유용성을 시사한다. 소비와 월슨에 따르면, 이타성은 유유상종하고 주기적으로 해제되고 재구성되는 집단에서 선택되고 진화한다. 유유상종을 위해서는 사기꾼이나 무임승차자를 구분하는 능력, 즉 인지 능력이 필수다. 발생계 이론에 따르면 인지 능력은 전체 발생계의 산물이다. 이타성을 가능하게 하는 필수 요소의 하나가 발생계 이론으로 잘 설명된다면, 이타성 그 자체도 그럴 가능성을 기대할 수 있을 것이다.

6. 나오는 말

논란이 없는 것은 아니지만 진화생물학계에 대체적인 경향은 선택의 단위가 유전자에서 집단에 이르기까지 다양한 수준에서 이루어진다는 점을 인정한다. 유전자, 개체, 혈연 집단, 사회 집단 등은 대표적인 선택의 단위이다. 개체를 통해서는 직접재생산 형질, 즉 진화적으로 이기적인 형질만이 선택되고 진화한다. 이타성은 집단이 선택의 단위가 되어야 성립한다. 혈연 집단이 선택의 단위가 되는 경우가 혈연 이타성이고 협동적 집단이 선택 단위

가 되는 경우가 호혜 이타성이다. 마이어는 혈연 이타성에 근거를 두는 혈연 이타주의는 주로 자식에 대한 부모의 사랑을 중심으로 인간 윤리의 작은 부분을 차지하고, 호혜 이타성에 근거하는 호혜 이타주의는 다양한 방식으로 이루어지는 서로 주고받는 행위 또는 신체의 상호 교환의 윤리로 인간 윤리의 상당 부분을 차지하지만, 그것들보다 훨씬 더 중요한 인간 윤리의 요소는 문화 집단을 통해 선택되고 진화하는 순수 이타성에 근거를 두고 있다고 본다(Mayr 2002). 혈연 이타성과 호혜 이타성은 근본적으로 개체 수준의 선택압을 통해 진화한다. 그것에 비해 순수 이타성은 집단 수준에 작용하는 선택압에 의해 진화한다.

혈연 이타성에서 순수 이타성으로의 진화는 본능적 이타성에서 의식적 이타성으로의 진화이다. 간접재생산 전략으로서의 진화적 이타성은 인간을 포함한 많은 동물들에서 확인된 형질이다. 그리고 진화적 이타성을 실현하기 위해 인간의 종적 마음에 새겨진 것(본능)으로서의 어떤 심리적 기제들이 있다는 주장도 옳을 것이다. 가족에 대한 이타성(혈연 이타성)은 주로 그리고 비혈연 개체들 간의 상호 교환적 이타성은 상당 정도 그러한 본능적 기제의 영향을 받은 것으로 추측된다. 그러나 문화 집단 간에 선택되는 형질로 가면 그런 본능은 도덕 진화 상의 주도권을 상실한다. 순수 이타주의는 그런 문화 집단에 작용하는 선택압에 의해 진화한다. 그 단적인 증거는 문화의 진화가 유전자의 진화에 비교할 수 없이 빠르다는 점이다. 오늘날의 인간의 도덕 문화는 그 대부분이 문명의 발생 이후의 산물이다. 그에 비해, 인간의 본성은 인간 진화사의 대부분을 차지하는 문명 이전의 수렵채취 기간 동안 진화한 것이다.

유전자와 도덕 간의 논란의 핵심은 연결의 유무가 아니라 도덕 형성에서 유전자가 맡는 역할의 강도 또는 주도권 문제이다. 사회생물학은 유전자가 도덕 규범 형성의 골격을 제공한다는 강한 입장을 선호한다. 그러나 우리가 앞에서 봤던 주장들을 감안해 보면 그런 입장은 성립되기 힘들다. 본능에 새겨진 진화적 이타성의 후성 규칙은 도덕적 본성의 내용을 결정짓는다고 볼 수 없다. 인간의 도덕적 본성은 인간이 갖고 있는 다양한 지적 정서적

상호작용의 결과물이다. 지적 정서적 능력이 유전자의 산물이라고 해서 그 능력의 산물까지 유전자의 산물일 필요는 없다. 마이어는 도덕적 행동을 가능하게 한 인간의 능력이 “포괄적응도를 기본으로 하는 본능적 이타성으로부터 판단에 따른 집단 윤리로 넘어가는 인간화 과정”에서 형성되었다고 본다. 그리고 그런 능력은 유아기와 아동기가 엄청나게 길어진 것, 그래서 자식을 돌보는 기간 역시 길어진 것, 인간 사회가 혈연을 넘는 대규모로 늘어난 것, 종족 특유의 전통과 문화가 발달한 것 등과 같은 다른 특성들의 진화와 밀접하게 연관되어 있으며, 전체적으로 볼 때, 이처럼 서로 연관된 발달 과정에서 무엇이 원인이고 무엇이 결과인지를 밝힌다는 것은 불가능하다고 주장한다(Mayr, 2002: 401).

그런 점에서도 발달계 이론은 설득력을 갖는다. 발달계 이론은 이타성을 유전자적 요소와 문화적 요소의 두 요소로 분리될 수 있는 것이 아니라 유전자 및 후성유전자에서부터 문화에 이르는 하나의 계 전체의 협업으로 이루어지는 발생 과정의 산물로 본다. 그 과정에서 어느 특정 요소가 주도권을 갖는다고 말할 수도 없다. 당연히 이타성을 계의 특정 요소로 환원하려는 시도도 실패할 수밖에 없다. 인간의 도덕적 행동은 도덕적 믿음에 기초해서 수행된다. 그런 믿음을 습득할 수 있는 능력 자체는 선천적으로 타고났을 것이다. 그러나 그것은 특정 도덕적 믿음에 대한 선천적 능력은 아니다.²³⁾ 도덕의 형성에서 유전자에 근거하는 윤리 성향이 주도권을 갖고 있지 못하다면, 그리고 인간성을 산출하는 발생계가 엄청난 양의 문화적 골격들을 포함하고 있으며 동시에 문화가 인간의 유전적 속박을 벗어난 독자적 실재의 성격을 갖는다면,²⁴⁾ 사회생물학의 역할은 기대보다 훨씬 작을 것이고,

23) “인간은 아마 선천적으로 윤리적 믿음을 습득할 수 있는 능력을 타고났지만, 그것은 특정한 믿음에 대한 능력은 아니다”(Waddington, 1960: 126, Mayr, 1997: 410에서 재인용).

24) 소버와 월슨은 인간의 마음의 자율성을 다음과 같이 표현하고 있다:
생물학적 진화가 두뇌를 낳았고 두뇌가 지금과 같은 우리의 행동을 유발한다는 것은 분명히 참이다. 그러나 그 사실로부터 두뇌가 단지 수동적인 근인적 기제의 역할만을 하며 오직 우연히 다윈적 이점을 부여하는 행동들만 실행한다는 결론은 나오지 않는다. 생물학적 선택은 두뇌를 낳았다. 그러나 두뇌는 생물학적 선택압

문화적 진화론 모델 같은 융통성 있는 이론의 역할은 기대 이상일 것이다.²⁵⁾ 그래서 필자는 유전자의 관점에서 도덕을 설명하려는 윌슨의 사회생물학이 도덕에 관한 “**일부** 사실들에 대한 설명의 **일부**”(part of the explanation of *some* facts about)라는 소버와 윌슨(Sober and Wilson, 1998: 298, 저자의 강조)의 평가가 틀릴 가능성은 없을 것이라고 예단한다.

을 상쇄할 수 있는 강력한 하나의 과정을 만들어 냈다. 정신은 생물학적 선택이 선호하는 행동들을 산출하는 장치 이상의 것이다. 그것은 자기 나름의 적응도와 되물림가능성(heritability)의 척도에 의해 정의되는 고유한 선택 과정의 기초이다. 자연 선택은 자유롭게 떠다니는 하나의 선택 과정을 탄생시켰던 것이다(Sober 2000: 215).

25) 보이드와 리처슨(Boyd & Richerson, 1985) 및 카발리-스포르자와 펠드만(Cavali-Sforza and Feldman, 1981)의 문화적 진화론 모델 참조.

참고문헌

- 정상모 (1998), 「파핏의 상식 도덕성」, 『철학』 54: 219-257.
- 정상모(2008), “A new definition of evolutionary altruism”, 『과학철학』 11권 2호: 223-255.
- 정상모(2009), 「유전자와 문화 사이: 후성 규칙의 빛」, 『철학논총』 56집: 107-126.
- Ayala, F. J.(1987), “The biological root of morality”, *Biol. and Phil.* 2: 235-252.
- Bowie, N.(1999), *Business Ethics: A Kantian Perspective*, Oxford: Blackwell Publishers.
- Boyd, R. & Richerson, P.(1985), *Culture and The Evolution Process*, Chicago: University of Chicago Press.
- Cavali-Sforza, L and Feldman, M.(1981), *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton: Princeton University Press.
- Darwin, C.(1871), *The Descent of Man and Selection in Relation to Sex*, London: John Murray.
- Fisher, R.(1930): *The Genetical Theory of Natural Selection*, Dover Book, 2nd ed. 1957.
- Griffiths, P. E. and K. Stotz(2000), “How the mind grows: a developmental perspective on the biology of cognition”, *Synthese* 122: 29-51.
- Griffiths, P. E. and R. D. Gray(2005), “Discussion: Three ways to misunderstand developmental systems theory”, *Biology and Philosophy* 20: 417-425.
- Hamilton W. D.(1963), “The evolution of altruistic behavior”, *American Naturalist* 97: 354-356.
- Hendriks-Janson, H.(1996), *Catching Ourselves in the Act: Situated Activity, Interactive Emergence, Evolution and Human Thought*, Cambridge, MA.: MIT Press.
- Keller, E. F. (2005), “DDS: Dynamics of developmental system”, *Biology and Philosophy* 20: 409-416.
- Lumsden, J and E. O. Wilson(1981), *Genes, Mind and Culture-The Coevolutionary Process*, Massachusetts: Harvard University Press.
- Mayr. E.(1997), *This Is Biology*, Cambridge, MA.: Harvard University Press; 최재천 외 옮김, 『이것이 생물학이다』, 몸과 마음, 2002

- Oyama, S.(2000), *Evolution's Eyes: A Systems View of the Biology-Culture Divide*, Duke University Press, Durham, NC.
- Parker, G.(1978): "Search for mate", in J. Krebs and N. Davies (eds.), *Behavioral Ecology: An Evolutionary Approach*. Blackwell.
- Ruse, M.(1998), *Taking Darwin Seriously*, Amherst, NY: Prometheus Books.
- Sarkar S.(1998), *Genetics and Reductionism*, Cambridge: Cambridge University Press.
- Sober, E.(1998), "What is Evolutionary Altruism?" in D. Hull and M. Ruse(eds.), *The Philosophy of Biology*, Oxford University Press, 459-478.
- _____ (2000), *Philosophy of Biology*, 2nd, Westview Press.
- _____ (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge: Harvard University Press.
- Thelen, E.(1995), "Time-Scale Dynamics and the Development of an Embodied Cognition", in R. F. Port and T. van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*", Cambridge, MA: MIT Press. pp. 68-100.
- Trivers, R. L.(1971), "The Evolution of Reciprocal Altruism", *Quarterly Review of Biology* 46: 35-57.
- Waddington, C. H.(1960), *The Ethical Animal*, London: Allen and Unwin.
- Wilson, E. O.(1975), *Sociobiology: The New Synthesis*, Cambridge, Mass.: Harvard University Press, 이병훈·박시룡 옮김(1992), 『사회생물학』, 서울: 민음사.
- Wilson, E.(1978), *On Human Nature*, The President and Fellows of Harvard College, 이한음 옮김(2000), 『인간 본성에 대하여』, 사이언스북스.
- Wilson, D. S. and Dugatkin, L. A.(1992), "Altruism: contemporary debates", in Keller E. F. and Lloyd E. A.(eds.), *Keywords in Evolutionary Biology*, Cambridge: Harvard University Press, 29-33.
- Wuketits, F. M.(1990), *Gene, Kultur and Moral: Soziologie-Pro und Contra*, 김영철 옮김 (1999), 『사회생물학 논쟁』, 서울: 사이언스북스.

원고 접수일: 2009년 3월 30일

심사 완료일: 2009년 5월 27일

계재 확정일: 2009년 5월 28일

ABSTRACT

Between Genes and Morality: Centering around Altruism

Jung, Sang Mo

Edward Wilson argues that the core of morality be altruism and that altruism derive from the genes inscribed in the evolutionary altruism such as kin and reciprocal altruism. If this argument means that morality comes out from the disposition of behavior caused by the human brain and that the human brain comes out from genetic information of genes, then nobody would deny it. However, if the argument means that the brain produces only the disposition of behavior that meets selection and evolution within the frame work of information inscribed in genes, it would be a jump of logic. Although the human brain is the result of evolution, evolution has allowed it to have the emergent freedom of mind. In consequence, if one departs from genes and goes toward morality through the pathway of sociobiology, he cannot reach the destination. In this paper, I will show this point of view with several arguments concerning evolutionary altruism. As a preliminary discussion, I will get rid of the semantical obscurity around the concept of altruism, and then I will bring into relief the fact that both the way from genes to mind and the way from mind to morality do not exist. Finally, I will argue that the developmental system theory support the fact well.