

【특집】

이론적 합리성과 실천적 합리성

김기현*

【주요어】 합리성, 효용, 진리, 도구주의, 목적(rationality, utility, truth, instrumentalism, purpose)

I. 들어가며

우리는 세상에 대하여 알고자 하여 이러저러한 믿음을 구성한다. 우리는 또 이들 믿음이 합리적으로 구성되었는가를 묻고, 합리적인 믿음을 위한 기준을 찾으려 하며, 이를 통하여 진리를 향한 우리의 인식을 향상시키고자 한다. 한편 우리는 세상에 대한 진리 탐구만 할 수는 없으며, 구체적인 행위를 통하여 실천하며 살아가야 한다. 현실 상황에서의 우리의 구체적인 행위가 합리적으로 이루어졌는가라는 질문이 제기된다. 현명한 행위 선택을 위한 질문은 행위 지침을 유도하며 우리는 이런 과정을 통하여 합리적인 행위자가 되어 간다. 전자와 관련된 합리성의 문제는 흔히 인식적 또는 이론적 합리성이라는 제목 하에 인식론과 과학철학의 탐구 대상이 되어 왔다. 반면에 후자는 실천적 합리성이라는 제목 하에 사회과학자, 철학자, 윤리학자들의 관심이 되어 독자적인 영역으로 발전하고 있다. 이 글은 이론적

* 서울대학교 철학과

합리성과 실천적 합리성의 논의 구조를 비교하여 양자 사이의 유사점과 차이점을 제시하고, 양자의 가능한 연관성에 대한 몇 가지 시론적인 주장을 제시한다.

이 글은 크게 보면 이론-실천의 구분을 다루는데, 이 글에서 다루는 구분과 다른 이론-실천의 구분이 있다. 이론과 실천이 관심사로 제기될 때 흔히 양자 사이의 괴리가 주목을 받는다. 이러한 이론과 실천 사이의 괴리는 철학자들 뿐 아니라 일상인에게도 친근한 주제다. 어떤 주장에 대하여 그것은 이론적으로는 맞는 이야기이지만 현실에 적용되지 않는다고 대꾸하는 경우를 볼 수 있다. 이론적 차원에서 맞는 것이 현실 생활에서 올바른 지침이 되지 못할 수 있음을 지적하는 말이다. 여기서의 구분은 인간의 실천에 관한 이론적 지침이 현실적 맥락에서 적용되기 어려운 경우가 있음을 말하고 있기 때문이다. 따라서 정확히 말하면 위의 구분은 실천에 관한 이론과 실제적 실행 사이의 괴리를 지적하고 있는 것이다. 이 문제는 철학에서 널리 다루어진 주제이기는 하지만, 이 글의 관심사는 아니다.

이 글에서 관심을 갖는 이론-실천의 구분은 인간 행위가 행하여지는 목적에 있어서의 차이 또는 인간의 행위를 평가하는 기준의 차이와 관련된다. 우리는 천체 물리학을 연구하여 우주의 생성과 관련된 진리를 탐구하는 행위에 가담하기도 하고, 주어진 돈을 가장 효과적으로 쓰는 방법은 극장을 가는 것인지 맛있는 음식을 사먹는 것인지를 숙고하여 어떤 행동을 할 것인지를 결정하기도 한다. 전자를 이론적 작업이라 한다면 후자는 실천적 작업이라 할 수 있다. 이들은 최종적 산물에 있어서도 차이가 난다. 이론적 작업의 산물은 사태에 대한 믿음인 반면, 실천적 작업의 산물은 행위다. 양자는 목적에 있어서도 차이가 난다.

이론적 작업과 실천적 작업은 목표도 다르고 최종적 산물도 달라 서로 독립적인 영역으로 간주되어 각기 다른 분야에서 발전하여 왔다. 각 영역의 발전에 비하여 볼 때 양자 사이의 관계에 대한 논의는 상대적으로 빈약하다. 이러한 상황은 다소 의아한 일이다. 우선 양자를 구분하는 믿음과 행위의 구분이 그렇게 확고한 것이 아니다. 진리

를 탐구하는 작업은 주어진 증거에 비추어 가장 참일 것으로 보이는 명제를 선택하는 의사 결정 과정으로 볼 수 있다. 마찬가지로 실천적 행위도 비록 진리를 목표로 하지는 않을지라도 주어진 상황에 비추어 일정한 기준에 따라 가장 바람직한 행위를 결정하는 의사 결정 과정을 포함한다. 따라서, 양자는 의사 결정이라는 유사한 논리적 구조를 갖고 있는 것으로 볼 수 있을 뿐 아니라, 명제를 믿는 것도 넓은 의미의 행위로 볼 수 있으므로 이론적 합리성의 구조와 실천적 합리성의 구조가 유사할 근거가 충분하다. 양자의 구조 사이의 유사점과 차이점을 살펴보기로 하자.

II. 합리성 논의의 일반적 구조

결과주의와 본래주의

우리의 많은 행동들은 목표를 지향한다. 나는 일정한 장소에 가기 위하여 버스를 타고 배고픔을 달래기 위하여 밥을 식사를 한다. 이러한 목적 지향적인 행위를 행위의 전형으로 보면 합리성에 대한 특정한 견해가 따른다. 행위가 목적을 지향하므로 그 행동의 합리성이 목적 도달에 기여하는가에 의하여 결정된다는 견해가 그것이다. 이렇게 한 행동의 합리성이 목적 성취 여부에 미치는 결과에 의하여 결정된다는 입장이 결과주의consequentialism이다. 합리성에 대한 비결과주의non-consequentialism는 모든 행위들이 목표에 기여하는가에 의하여 결정되지는 않는다는 입장이다. 즉, 어떤 행위들은 목적과의 연관성에 의해서가 아니라 그 행위 자체가 갖는 본래적인inherent 속성 때문에 합리적이라고 주장한다. 따라서 결과주의에 반대하는 입장을 본래주의라고도 부를 수 있다.

도구주의와 비도구주의

합리성에 관한 다음 구분은 행위 자체가 아니라 행위의 목적에 관한 상반된 견해와 관련된다. 한 행위가 일정한 목적을 지향하고 있다고 할 때, 그 행동의 합리성은 주어진 목적과의 관계 하에서 평가된다. 목적 자체는 어떠한가? 일정한 목적을 취하는 것, 일정한 대상을 다른 대상보다 더 선호하는 것 등의 태도는 합리성의 관점에서 평가될 수 있는가? 아니면, 목적 설정은 합리성 논의를 위한 전제의 역할을 하므로 그 자체는 합리성 평가의 대상이 될 수 없는 것인가? 흄David Hume은 행위의 궁극적인 목적은 합리적 평가의 대상이 되지 않는다고 주장하는 대표적인 철학자다. 목적들은 기본적인 목적unmotivated end과 그 목적을 달성하기 위한 도상의 과정에서 설정되는 부차적 목적motivated end이 구분될 수 있다. 이 경우에 다른 목적을 달성하기 위한 부차적 목적의 설정은 그 합리성이 평가될 수 있다. 그 부차적 목적이 기본 목적을 달성하기 위한 효과적인 길인가가 합리성의 관점에서 평가될 수 있을 것이기 때문이다. 그러나 기본적인 목적의 설정과 관련하여서는 합리성의 평가가 개입할 여지가 없다고 흄은 주장한다. 이러한 입장이 합리성에 관한 도구주의instrumentalism이다. 반면에 도구주의에 반대하는 비도구주의non-instrumentalism는 목적들의 적절성이 이성적으로 인식되고 평가될 수 있다고 주장한다.

도구주의가 모든 합리성의 평가가 목적과 수단과 관련된 평가라고 주장하고 있지 **않다**는 점에 주의할 필요가 있다. 이들은 여러 상반된 목적이 주어졌을 때 이들을 어떻게 조화할 것이며, 이들에게 어떤 비중을 부여하고, 또 이들을 달성하기 위하여 시간상의 계획을 세우는 것들이 모두 합리성과 관련된 논의의 영역에 속한다는 것을 부정하지 않는다. 다만 도구주의는 이러한 모든 논의를 촉발시키는 기본적인 목적이 합리성 논의의 영역 밖에 있다고 주장한다.

일인칭적 관점과 삼인칭적 관점

한 학생의 학업 성취도를 평가하는 경우를 보자. 일등에게 주는 최고상의 예에서 보여지듯이 때로는 학생의 성적이 우수할 때 상장을 준다. 학생이 공부를 열심히 하였는가는 그러한 평가에 있어 고려되지 않는다. 반면에, 장려상의 경우에서 보여지듯이, 다른 학생들과 비교해 볼 때 우수하지 않을지라도 학생이 상당한 노력의 결과로 발전을 하였을 때 상을 주기도 한다. 여기서 우리는 평가를 위한 두 가지 상반된 견해를 볼 수 있다. 전자는 행위 주체의 동기와 무관한 결과에 의한 평가이며 후자는 결과보다는 동기를 중시하는 평가다. 위에서 나타나는 차이는 행위의 도덕성을 주체의 동기를 통하여 평가하여야 하는가, 아니면 행위 자체가 초래하는 결과를 통하여 평가할 것인가와 관련한 윤리학의 오랜 대립을 반영하고 있다. 윤리학에서 이 대립은 결과주의와 동기주의 사이의 대립으로 잘 알려져 있다.

합리성도 윤리적 개념과 마찬가지로 평가적 개념이기에, 결과주의적 관점과 동기주의적 관점 사이의 대립이 나타난다. 한 행동이 일정한 목적을 향하고 있을 때, 그 행동 주체가 그 목적을 달성하기 위하여 최선의 노력을 하였을 경우 그 행동을 합리적이라고 평가할 수 있다. 반면에 주관적 의지와 관계없이 그 행동이 객관적 상황에 비추어 볼 때 그 목적을 달성하기 위한 가장 효율적인 방법일 경우에 그 행동을 합리적이라고 평가할 수 있다. 전자는 행위자 자신의 관점에서 벌어지는 일을 합리성 평가의 핵심으로 삼으므로 합리성에 관한 일인칭적 관점이며, 후자는 그보다는 모든 관찰자가 객관적으로 판단할 수 있는 근거를 통하여 합리성을 평가하므로 합리성에 관한 삼인칭적 관점이다.

여기서 윤리학에서와 같이 결과주의/동기주의의 구분을 사용하지 않는 이유는 앞서 결과주의란 표현이 이미 사용되었기 때문이다. 앞서 설명한 바에 따르면 결과주의는 행위가 본래적인 합리성의 가치

를 가질 수 없으며, 오직 목적과의 연관성 하에 나타나는 결과에 의하여 합리성이 평가되어야 한다는 입장이다. 따라서, 결과주의 틀 내에서도 그 결과를 끌어내기 위하여 그 행위를 하였을 경우에 그 행위를 합리적인 것으로 볼 것인가, 아니면 실제 결과가 그러한 결과를 가져올 경우에 합리적인 행위로 볼 것인가의 구분이 가능하다. 즉, 행위가 갖는 본래적 합리성을 부정하는 결과주의의 틀 내에서도 일인칭적 견해와 삼인칭적 견해의 구분이 있을 수 있다.

Ⅲ. 이론적 합리성의 특성

아리스토텔레스적인 전통 하에서 합리성 개념은 목표 지향적인 개념으로 간주된다. 이에 따르면, 한 행위가 합리적인가 아닌가는 그 행위가 주어진 목표에 도달하기 위한 효과적인 수단인가에 의하여 결정된다. 그렇다면, 한 믿음이 이론적으로 합리적인가를 결정하기 위해서는 우선 이론의 목표가 무엇인가가 밝혀져야 한다. 인식론자들은 이론적 행위의 목표는 참에 도달하고 거짓을 피하는 것이라고 간주한다. 따라서, 이론적인 관점에서 볼 때 합리적인 믿음 또는 인식적으로 정당한 믿음은) 참에 도달하고 거짓을 피하는 목표에 비추어 바람직한 믿음이며, 믿음의 이론적 합리성은 그러한 목표에 기여하는 바에 의하여 평가된다. 이런 점에서 이론적 합리성은 명백히 결과주의적이다. 즉, 한 믿음의 인식적 합리성은 참의 추구, 거짓의 피함이라는 목표와의 상관 관계에 의하여 평가되는 것이지, 이론적인 관점에서 볼 때 본래적으로 합리적인 믿음이란 있을 수 없다.

1) “이론적”이라는 표현과 “인식적”이라는 표현은 상호 교환하여 사용될 수 있다. 그리고, 이론적 맥락에서 “합리성”은 “정당성”과 상호 교환하여 사용될 수 있다. 따라서, 이론적 합리성, 이론적 정당성, 인식적 합리성, 인식적 정당성은 모두 같은 의미를 갖는다. 이 글에서도 이들이 서로 같은 의미로 사용될 것이다.

다음으로 이론적 합리성은 도구주의적 합리성인가, 비도구주의적 합리성인가?2) 우선 이론과 인식이라는 개념에 참의 추구하고 거짓의 회피라는 것이 포함되어 있음에 주목하자. 참을 추구하고 거짓을 피하는 것은 인식 행위와 관련된 기본적 목표이며, 앞에서 보았듯이 인식 행위의 목표가 이러한 특성을 갖는다는 것은 순수히 개념적 분석으로부터 도출된다. 개념 분석은 이성적 작업이다. 따라서 참의 추구하고 거짓의 회피라는 이론적 행위의 목표가 개념 분석을 통하여 정당화되고 인식되므로 이론적 합리성은 비도구주의적이라고 할 수 있다. 그러나, 인식 행위의 목표가 이성적 평가의 대상이기는 하지만, 그 목표가 선험적 개념 분석으로부터 도출되기 때문에 개인차가 나타나지 않는다. 이런 이유 때문에 비록 이론적 행위의 목표가 이성적 합리화의 대상이면서도 논란의 대상이 되지는 않았으며, 이론적 합리성이 도구주의적인가 하는 문제가 인식론에서 논의 주제로 떠오르지 않았다고 할 수 있다.

다음으로 일인칭적 견해와 삼인칭적 견해의 구분이 이론적 합리성에서 어떻게 나타나는가를 보자. 이론적 합리성을 다루는 인식론에서는, 윤리학에서의 경우와 마찬가지로, 일인칭적 견해와 삼인칭적 견해를 모두 볼 수 있다.3) 윤리학에서의 동기주의의 경우에서와 같은 믿음의 이론적 합리성과 관련된 평가는 인식 주관을 평가의 대상으로 삼을 수 있다. 이 경우에 이 인식 주관이 참을 추구하고 거짓을 피하는 인식론적 목표에 도달하기 위하여 최선을 다하고 있는가가

2) 이 문제와 관련하여 이론적 합리성과 실천적 합리성의 중요한 차이가 드러난다. 그리고 이 차이 때문에 이론적 합리성에 대한 기존의 논의와 실천적 합리성에 관한 기존의 논의가 상당히 다른 양상을 띠고 나타나게 된다. 이것은 뒤에 실천적 합리성을 설명하는 부분에서 제시될 것이므로, 여기서는 이론적 합리성이 이 문제의 맥락에서 갖는 특성을 우선 살펴보기로 하자. 도구주의/비도구주의라는 관점에서 볼 때 이론적 합리성은 특이한 위상을 갖는다.

3) 이 주제에 대한 상세한 논의는 줄고 김기현(1998)을 참조하라. 이 글은 두 견해를 상세히 소개함과 동시에 양 개념을 모두 비판한다.

고려 사항이 된다. 이 사람이 거짓을 피하고 참에 도달하려는 주체적 노력을 통하여 한 믿음을 형성하였을 경우에 그 사람은 칭찬을 받고 그렇지 못한 경우에 비난을 받는다. 이러한 의미에서 칭찬을 받으려면, 인식 주관은 주어진 근거를 고찰하여 그것이 문제의 명제를 참이 게끔 한다는 것을 인식하고 그에 의하여 이 명제를 믿어야 한다. 윤리학에 있어서의 동기주의와 유사한 이러한 견해를 인식적 합리성에 대한 전통적인 견해라고 부르자. 이러한 견해가 전통적 인식론의 주류를 이루고 있기 때문이다. 이제 인식적 합리성에 대한 전통적 견해의 핵심을 이루는 주장은 다음과 같이 요약된다:

S가 R이라는 근거에 의하여 p라고 믿는 것이 인식론적으로 합리적 이기 위해서는, S는 R이 주어졌을 때 p가 참일 가능성이 높다는 상위 의식을 가져야 한다.

여기서 상위의식이라는 표현이 쓰이는 이유는 이 의식이 단순한 외적인 사태에 대한 의식이 아니라 인식 주관 내에 존재하는 심리적 상태인 믿음과 그 근거들에 관한 의식이기 때문이다.⁴⁾

이론적 합리성에 대한 전통적 견해는 사람들이 반성을 통하여 한 믿음이 과연 참일 수 있는가를 평가하는 능력이 있다는 사실에 근거하고 있다. 상위의식의 요청은 인간이 갖는 이러한 비판적 고찰 능력에 주안점을 두고서 인식적 합리성을 분석한 결과로 보인다. 만약 사람들이 그러한 능력을 갖고 있지 않다면, 사람들이 믿음을 받아들임에 있어 참을 구하고 거짓을 피하는 목적에 비추어 주어진 믿음을 평가할 것을 요구할 수 없을 것이다.

한편 인간의 인식체계에는 문제의 믿음을 비판적으로 고찰하는 능력과는 다른 그리고 아마도 더 본질적인 인식적 능력이 있다. 이 능력은 외부 세계로부터 입력을 받아들이고 그로부터 믿음들을 산출하

4) 전통적 견해를대표하는 이론들로는 다음이 있다: Chisholm(1968, 1977, 1982), Ginet(1975), Moser(1985), Naylor(1988), Wolterstorff(1983), Pollock(1986), BonJour(1986).

는 능력이다. 인식 체계의 이러한 측면에 주목하게 되면, 인간의 인식 체계는 믿음을 비판적으로 고찰하는 체계라기보다는 믿음을 형성하는 체계로 보인다. 인식 체계를 믿음 형성자로 보는 견해를 진리 추구를 인식적 목표로 보는 견해와 결합하면, 인식적으로 합리적인 믿음은 진리 추구와 거짓을 피하는 목표에 기여하는 방식으로 잘 형성된 well-formed 믿음 또는 잘 유지된 well-sustained 믿음으로 파악된다. 예를 들어, 발생적 견해를 대표하는 골드만은 인식적 합리성을 다음과 같이 정의한다.

만약 S가 t의 시점에 p라고 믿는 것이 믿음을 형성하는 신빙성 있는 인지과정의 결과라면, t에 S가 p를 믿음은 정당화된다.⁵⁾

골드만은 발생적 견해의 근간을 이루는 잘 형성됨을 신빙성 있는 인지과정에 의하여 산출됨으로 분석하고 있으며,⁶⁾ 신빙성을 거짓 믿음보다 많은 참된 믿음을 산출하는 성향으로 정의하고 있다.

한편 다른 발생적 인식론자들은 인식적 합리성을 증거와의 연관에 의하여 정의하는 일반적 틀을 받아들이면서 한 믿음의 인식적 합리성을 발생론적으로 정의한다. 이들에 따르면, 한 믿음과 그 증거 사이에 적절한 인과적 관계가 성립할 때, 그 믿음이 잘 형성된다. 알스톤 William Alston,⁷⁾ 스웨인 Marshall Swain,⁸⁾ 웰드만 Richard Feldman 과 코니 Earl Conee⁹⁾가 이러한 이론을 옹호한다. 이제 “근거”라는 표현을 단지 증거에 해당하는 심리적 상태 뿐 아니라 심리적 과정까지도 포함하는 넓은 의미로 사용하면, 인식적 합리성에 대한 발생론적 견해의 핵심은 다음과 같이 요약된다:

5) Goldman(1992), 116.

6) 골드만 스스로 “잘 형성됨”, “발생적 이론” 등의 표현을 사용하고 있다. 위의 논문, 117.

7) Alston(1986).

8) Swain(1981).

9) Feldman and Conee(1985).

R이라는 근거에 의하여 S가 p라고 믿는 것이 합리적이기 위해서는, S가 p라고 믿는 것이 R에 의하여 야기되어야 한다.¹⁰⁾

이론적 합리성에 대한 전통적 견해가 일인칭적 관점을, 발생적 견해가 삼인칭적 관점을 대변하고 있음은 분명하다. 전통적 견해는 일인칭적인 관점에서 합리성을 고찰하면서 진리를 추구한다는 이론적 합리성의 목적에 도달하기 위하여 최선을 다하였는가에 의하여 이론적 합리성을 평가한다. 반면에 발생적 견해는 한 믿음을 형성한 구체적 과정이 실제로 진리 추구라는 인식 목표에 합당한가에 의하여 일인칭적 관점보다는 제3자의 관점에서 인식적 합리성을 평가한다. 이 두 관점은 인식적 합리성에 기본적으로 상이한 태도를 대변하면서 현대 인식론에서 주요한 두 주류를 형성하고 있다. 그리고 인식적 합리성에 대한 기본적 관점의 차이는 이론의 구체적 내용에 있어서의 다양한 차이를 낳게 된다.

우선 전통적 견해와 발생적 견해는 인식적 합리성과 진리 추구라는 인식적 목표와의 관계에 대한 상이한 입장을 보인다. 인식 주관을 평가하는 전통적 입장에 따르면, 한 믿음이 인식 주관에게 합리적이기 위해서는 그 믿음이 근거를 성찰하는 인식 주관의 반성적 의식에 참인 것으로 나타나야 한다. 따라서, 전통적 견해에 따르면, 한 믿음이 합리적이기 위해서는 그 믿음의 근거가 내재적으로 포착되고 그 근거의 참됨이 (또는 참일 개연성이 높음이) 의식에 내재적으로 포착되어야 한다. 반면에, 이미 보았듯이, 발생적 견해는 이러한 내재적 포착의 요구를 부정하고, 한 믿음이 사실상 참일 개연성이 높은 것으로 족하다고 주장한다. 이러한 대비 때문에 전통적 견해와 발생적 견해의 대립이 때로는 내재주의와 외재주의의 구분으로 나타난다.¹¹⁾

10) 여기서 야기한다함은 발생을 야기하는 것일 수도 유지를 야기하는 것일 수도 있다.

11) 위의 내재주의와 외재주의의 구분은 거친 구분이다. 내재주의와 외재주의의 구분은 현대 인식론에서 널리 쓰이고 있음에도 불구하고, 이들은 다른 맥락에서 다른 의미로 사용되고 있어 많은 혼란을 야기하고 있다. 이러한

전통적 견해에 따르면, 한 믿음이 인식 주관의 반성적 의식에 비추어 참인 것으로 나타나면, 그 믿음이 객관적으로 참일 개연성이 높지 않다 하더라도 인식적으로 합리적인 것이 된다. 인식 주관이 최대한의 반성적 노력을 기울인 결과 그 믿음이 자신이 성찰한 모든 증거에 비추어 참인 것으로 보인다면, 그의 믿음은 나무랄 수 없는 것이고 따라서 합리적이라고 하여야 하기 때문이다. 한편, 인식 주관이 아니라 믿음 자체를 평가하는 발생적 관점에서 볼 때, 한 믿음이 인식적으로 합리적이기 위해서는 그 믿음은 객관적으로 참일 확률이 높아야 한다. 한 믿음이 반성적 의식에는 정당한 것으로 보인다 하더라도, 그 믿음이 정당한 방식으로 발생하지 않았을 경우에는 그 믿음 자체로서는 참을 추구하는 목표에 기여하지 못하는 바람직하지 못한 믿음이기 때문이다. 이 사실이 왜 전통적 견해와 발생적 견해의 대비가 때로 주관적 견해와 객관적 견해의 대립으로, 때로 개인적 견해와 진리적 견해의 대립으로 나타나고 있는가를 설명해준다.

실천적 합리성에 대한 논의로 들어가기에 앞서 지금까지의 이론적 합리성에 대한 논의를 요약해보자. 합리성의 문제를 세 가지 구분의 축을 통하여 살펴본 결과, 이론적 합리성은 결과주의/본래주의의 축과 도구주의/비도구주의의 축에서 볼 때 별 논란이 없는 것으로 나타났다. 인식적 합리성에 관한 한 참을 추구하고 거짓을 피한다는 이론적 목표에는 이론의 여지가 없어, 모든 개별적인 믿음이 이 목표와 관련하여 결과주의적으로 평가되고, 또한 이론적 목표가 모든 사람들에게서 이견이 없는 이유로 그 목표 자체가 합리적 평가의 대상이 되는가의 문제는 사람들의 흥미를 끌지 않았기 때문이다. 이러한 사실이 이론적 합리성의 논의가 일인칭적 견해/삼인칭적 견해 사이의 대립에 집중된 것에 부분적으로 기여하였다고 볼 수 있다. 물론 합리성이 논의될 수 있는 세 문제 지평 중에서 두 지평에서 논란이 없다는 단순한 그 이유 때문에 인식적 합리성의 논의가 셋째 차원에 집중되었다는 것은 아니다. 잘 알다시피 인간이 외부 세계를 있는 모습

혼란과 그 해소를 위해서는 Kim(1993)을 보라.

그래도 인식할 수 있는가라는 문제는 철학의 역사를 통하여 가장 큰 주제라고 할 수 있다. 이 문제는 있는 그대로의 세계를 말하는 것이 합리적인가, 세계는 실재하는가 구성되는가, 진리의 본성은 무엇인가 등의 철학의 영원한 주제와 긴밀히 연관되어 나타난다. 이론적 합리성을 일인칭적 관점에서 파악하여야 하는가, 삼인칭적 관점에서 파악하여야 하는가의 문제는 이러한 더 큰 철학적 문제맥락과 닿아 있다. 인식이 궁극적으로 세계의 있는 모습 그대로를 반영할 수 있다고 생각하면, 믿음의 정당성을 세계의 모습과 비교하여 평가하는 것이 설득력을 갖는다. 반면에 그러한 실재론적 세계관과 그에 따르는 인식관을 받아들이지 않을 경우에는 한 믿음이 합리적인가의 평가는 인식 주관이 그 믿음과 관련하여 최선을 다하였는가에 의하여 평가될 수밖에 없을 것이다. 이것이 바로 인식론에서의 외재론과 내재론의 구분이며, 일인칭적 관점과 삼인칭적 관점의 구분과 대응하는 부분이다. 이 사실이 이론적 합리성 논의가 앞에서 제시한 문제의 세 축 중에서 일인칭적 견해와 삼인칭적 견해의 대립 부분에 집중된 주된 이유이며, 다른 문제 차원에서 논란이 없었던 것이 이런 집중화를 가속화하였다고 볼 수 있다.

IV. 실천적 합리성

실천적 합리성에 대한 현대적 논의를 고찰하는 가장 효과적인 방법은 실천적 합리성에 관한 기대효용 이론(expected utility theory)을 소개하고, 이를 둘러싼 논의를 고찰하는 것이다. 기대효용 이론이 실천적 합리성에 대한 지배적인 이론으로서 실제로 사회과학 분야에서 한 동안 패러다임의 역할을 하여 왔고, 실천적 합리성에 대한 논의는 이에 대한 비판과 수정의 방식으로 전개되어 기대효용 이론이 실천적 합리성과 관련한 담론의 주축을 이루고 있기 때문이다.

기대효용 이론은 효용 추구가 실천적 행위의 목표라는 공리주의적

인 정신에서 출발한 것으로 볼 수 있으며, 현대에 들어와 폰 노이만 John von Neumann 모겐스텐 Oskar Morgenstern, 세비지 Leonrad Savage에 의하여 이론적 체계로 완성되어 기대효용 이론이라는 이름을 얻는다.¹²⁾ 이 이론의 핵심은 행위자에게 주어진 가능한 행위의 선택지들 중에서 효용성이 가장 큰 행위를 하는 것이 합리적이라는 것이다. 다음의 예를 보자.

당신이 지금 10,000원을 주고 당첨되면 10,000,000원을 받는 복권을 살 것인가를 결정하고자 한다고 하자. 이때 당신에게 주어진 선택지는 복권을 사는 행위와 사지 않는 행위의 두 가지로 압축된다. 복권을 살 경우에 당신은 우선 10,000원의 순 지출을 하게 된다. 이러한 지출을 함으로서 당신은 복권에 당첨될 기회를 산다. 복권을 사게 되면, 그 결과는 당첨되든가 당첨되지 않든가 둘 중의 하나다. 당첨될 확률이 p 라고 하자. 그러면 당첨되지 않을 확률이 $(1 - p)$ 가 된다. 당첨될 경우에는 천만원을 받게 되고, 당첨되지 않을 경우에는 아무 것도 받지 못하므로, 복권을 삼으로써 기대할 수 있는 가치는 $10,000,000원 \times p + 0원 \times (1 - p)$ 라고 할 수 있다. 한편, 복권을 살 경우에 드는 비용 10,000원은 복권 사는 행위의 초기 조건에 해당하므로 복권 사는 행위와 관련된 기대할 수 있는 가치는 $10,000,000원 \times p + 0원 \times (1 - p) - 10,000원$ 이다. 반면, 복권을 사지 않는 행위가 갖는 기대치는 0이다. 복권을 사지 않을 경우 지출도 없고, 들어올 수입의 가능성도 전혀 없기 때문이다. 기대효용 이론에 따르면, 이 상황에서 어떤 행위를 하는 것이 실천적으로 합리적인가는 위에서 나타난 수식의 값에 의하여 결정된다. 만약 복권을 사는 행위에 동반하는 기대치가 0원보다 크다면 복권을 사는 것이 합리적이고, 0보다 작다면 사지 않는 것이 합리적이라는 것이다. 복권을 살 것인가 말 것인가를 결정하는 상황에서 선택지와 동반된 기대치를 비교하고, 이 중에서 기대치가 가장 큰 행위를 선택하는 것이 합리적이라는 것이다.

기대효용 이론은 우리의 모든 행위 결정 과정이 본질적으로 위와

12) Neumann, J. von and Morgenstern, O.(1944); Savage(1954).

같은 도박의 과정과 동일하다고 생각한다. 요지를 일반화하면 다음과 같다: 현재 당신에게 주어진 행위의 선택지가 A_1, A_2, A_3 라고 하자. 한 행위를 하였을 때 나타날 수 있는 결과가 확정적인 경우는 드물다. 따라서 한 행위에 상이한 결과들이 따르는 것을 생각해 볼 수 있다. A_1 에 동반하는 가능한 결과들을 $O_{11}, O_{12}, O_{13} \dots O_{1n}$ 라고 하자. A_2 에 동반하는 가능한 결과들은 $O_{21}, O_{22}, \dots O_{2n}$, A_3 에 동반하는 결과들은 $O_{31}, O_{32}, \dots O_{3n}$ 이라고 하자. 이 경우에 각 O 에 대하여 당신은 스스로의 선호에 따라 일정한 값을 부여할 수 있을 것이다. 좋아하는 결과라면 양수의 값을, 피하고자 하는 결과라면 음수의 값을 부여할 수 있을 것이다. 이 값이 각 결과가 갖는 효용성utility이다. 그렇다면, A_1 에 대하여 우리가 기대할 수 있는 효용성은 가능한 각 결과들의 효용성을 그 결과의 확률과 곱한 다음에 이들을 모두 합함으로써 얻어진다. 이를 도식화하면, A_1 과 관련된 임의의 결과 O_{1i} 의 효용성을 $U(O_{1i})$ 이라 하고, 그 확률을 $P(O_{1i})$ 라고 하면, A_1 이 갖는 기대효용은 $\sum U(O_{1i}) \times P(O_{1i})$ 가 된다. 그리고 A_2 의 기대효용은 $\sum U(O_{2i}) \times P(O_{2i})$, A_3 의 기대효용은 $\sum U(O_{3i}) \times P(O_{3i})$ 가 된다. 이제 당신은 이들 각 값을 계산하여 그 중 값이 제일 큰 행위를 선택하여야 한다. 그 경우에만 당신의 행위는 실천적 합리성을 갖는다.

이제 위에서 서술한 기대효용 이론이 2절에서 제시한 합리성 이론의 여러 부류 중에서 어디에 속하는가를 살펴보기로 하자. 우선 행위의 실천적 합리성에 대한 평가와 관련하여 기대효용 이론은 결과주의적 입장을 취하고 있음이 명백한 듯하다. 한 행위가 실천적으로 합리적인 것은 주어진 선택지들 중에서 기대효용이 가장 큰을 통하여 정의되고 있으므로 한 행위의 실천적 합리성은 효용성이라는 목표에 기여하는 방식에 의하여 정의되고 있다고 할 수 있기 때문이다.

다음으로 기대효용 이론은 도구주의적인가를 살펴보자. 우선 한 이론이 결과주의적이라는 것만으로 그 이론이 도구주의적이라는 결론이 필연적으로 따라 나오지 않는다는 것에 주목하자. 행위의 실천적 합리성을 설정된 목표와의 관계 하에서 평가할 것인가의 문제와 목

표 자체가 합리적 평가의 대상이 될 수 있는가의 문제는 상호독립적이다. 따라서, 목적의 설정이 합리적 평가의 대상이 될 수 있다고 하면서 동시에 행위의 실천적 합리성은 이렇게 설정된 목적에 기여하는 바에 의하여 규정된다고 할 수 있다. 이런 주장에는 아무 모순이 없으며, 이런 주장을 하는 이론은 결과주의적이지만 도구주의적이지는 않은 이론이다.¹³⁾ 그럼 기대효용 이론은 도구주의적인가?

많은 사람들은 기대효용 이론이 전형적인 도구주의적 이론이라고 생각한다. 많은 사람들은 실천적 합리성에 대한 도구주의적 이론을 거론하면 흔히 기대효용 이론을 연상한다. 이렇게 생각하는 이유는 기대효용 이론에서의 목표는 효용으로 대변되며, 이 효용은 공리주의에서의 고통과 쾌락에 따른 계산을 연상시키기 때문이다. 고통과 쾌락과 같은 경험은 합리적 평가의 대상이 될 수 없으며, 효용이 이러한 고통과 쾌락에 의하여 계산되는 것이라면 기대효용 이론은 도구주의적 이론이라고 할 수 있을 것이다. 그러나 기대효용 이론을 대변하는 현대의 이론가들은 효용성과 그와 관련된 선호preference를 고통, 쾌락 등의 경험을 통하여 정의할 수 있는가에 대하여 침묵한다. 이들은 효용성과 선호도를 행위자의 선택 행위를 통하여 정의한다. 예를 들어, 당첨 확률이 동일하고 단지 당첨품이 a와 b로 다른 두 복권에서 당신의 선택이 아무 차이가 나지 않는다면, 이는 a와 b의 효용성이 당신에게서 동일함을 보여준다. A가 당첨품인 복권을 선택한다면 a를, b가 당첨품인 복권을 선택한다면 b를 선호한다는 것을 보여준다. a의 당첨 확률이 b의 당첨 확률보다 두 배가 높은 두 복권 사이에서 당신의 선택이 차이가 나지 않는다면, 이는 b에 대한 당신의 선호도가 a에 대한 선호도의 두 배임을 보여준다. 이렇게 동일한 확률 사이에서의 선택 행위, 상이한 확률 사이에서의 선택 들을 통하

13) 그리고 도구주의적 이론이 결과주의적일 필요도 없다. 우리의 행위들을 목적 지향적인 행위와 그렇지 않은 행위로 구분하여, 목적에 관하여 도구주의적인 입장을 취하고 목적을 지향하는 행위에 대하여는 결과주의적인 입장을 취하면서 목적과 무관한 행위에 대하여는 본래적 합리성을 부여할 수 있기 때문이다.

여 행위자가 갖는 선호도, 효용성은 결정된다. 이들이 이렇게 선호도를 선택 행위를 통하여 정의하는 이유는 논리실증주의의 영향 하에 순수히 행동을 예측하는 이론을 구성하는 데에 관심을 갖고 있었기 때문이다. 어떤 과정을 통하여 일정한 선호도를 갖게 되었는가를 연구하는 것은 객관적으로 검증하기 어려운 심리의 영역으로 들어가는 것이므로 주저하고 있으며, 선호도를 행동의 선택 방식에 맞추어 정의하는 것은 효용을 극대화하려는 행위자의 미래 행위를 예측하는 데에 오히려 더 효과적일 수 있기 때문이다.

그럼에도 불구하고 기대효용 이론을 도구주의적으로 해석하는 것은 일리가 있는 일이다. 이들은 개별적 행위의 실천적 합리성에 대하여는 적극적인 이론을 제시하면서도 효용성과 선호도 자체가 어떻게 형성되는가, 그리고 이들은 합리적 평가의 대상이 될 수 있는가라는 문제에 침묵하고 있다. 이들이 효용성의 문제를 합리적 논의의 영역을 넘어서는 것으로 보고 있다고 해석할 때 이 침묵이 가장 잘 설명된다. 또 공리주의적 토양에서 기대효용 이론이 성장했다는 사실도 이러한 해석에 설득력을 더해준다. 이러한 측면들을 고려할 때 상반된 해석이 논리적으로 가능함에도 불구하고 기대효용 이론은 대체로 실천적 합리성에 대한 도구주의적 이론이라고 보아 무리가 없을 듯하다.

이러한 해석에 반하여 기대효용 이론을 비도구주의적으로 해석하는 입장이 있을 수 있다. 위에서 보았듯이 기대효용의 주장자들은 효용성을 선호 행위를 통하여 정의한다. 그런데 이들이 모든 선호 행위가 효용성을 정의한다고 생각하지는 않는다. 예를 들어, 한 행위자가 a를 b보다 선호하고, b를 c보다 선호하면서 c를 a보다 선호한다고 해보자. 이런 형태의 선호 행위는 효용성의 측정에 사용될 수 없으며 미래의 행위 예측에 도움이 되지 않는다. 이런 이유로 이들은 선호 행위가 만족해야 할 몇 가지 논리적 조건(일관성 조건)을 제시하고 이를 만족하는 선호를 통하여 효용성을 정의한다. 그렇다면 기대효용 이론에 있어 효용성의 정의에 논리적 고려, 따라서 합리성에 대한 고

려가 포함된다고 할 수 있을 것이다. 그러나 이 정도만으로 기대효용 이론이 비도구주의적이라고 결론내리기는 무리가 있는 듯하다. 왜냐하면 이 조건은 일단 형성된 선호도의 집합이 일정한 논리적 조건을 만족할 것을 요구하고 있을 뿐이며, 애초에 한 대상을 다른 대상보다 더 선호한다는 사실 자체에 대하여는 합리성에 따른 평가를 인정하고 있지 않기 때문이다. 우리가 비도구주의적 합리성 이론이라고 할 때는 이 정도까지를 기대한다.

더 나아가 선호도의 집합이 일관성이라는 논리적 조건을 만족시켜야 한다는 요구 자체가 도구주의적으로 정당화될 수 있다. 예를 들어 한 사람의 선호도가 $a > b$, $b > c$, $c > a$ 와 같이 일관성을 갖지 못한다고 하자. 이제 내가 c 를 갖고 있고, 그 사람은 a 를 갖고 있다고 하자. 그 사람의 첫째 선호도 때문에 이 사람은 나로부터 a 를 받고 b 와 덤을 주는 거래에 참여할 것이다.¹⁴⁾ 다시 두 번째 선호도 때문에 내가 그에게 b 를 줄 때 나에게 c 와 덤을 주는 거래에 그는 참여할 것이다. 이제 세 번째 선호도 때문에 내가 그에게 c 를 줄 때 나에게 a 와 덤을 주는 거래에 그가 참여할 것이다. 그렇게 되면, 결국 나는 처음에 갖고 있던 a 를 그대로 지니고 세 번의 거래를 통한 덤들을 부가로 갖는다. 반면, 그는 처음에 갖고 있던 c 와 덤들을 잃는 순 손실로 거래를 끝마치게 된다. 이러한 논증은 선호도가 일관성을 지녀야 하는 이유가 효용성의 증가라는 도구주의적 관점에서 정당화될 수 있음을 보여준다.¹⁵⁾

지금까지 우리가 살펴본 바에 따르면, 기대 효용 이론은 상이한 해석이 논리적으로 가능함에도 불구하고 실제적으로는 도구주의적이면서 결과주의적인 이론으로 보는 것이 합당하며, 이는 기대효용이론에 대한 일상적인 관점과 부합한다. 이제 실천적 합리성에 대한 기대효용 이론이 이론적 합리성에 대한 이론과 마찬가지로 일인칭적 관점

14) 이 덤은 극히 작은 것이라도 상관없다.

15) 이 논증은 확률의 일관성에 대한 요구를 정당화하는 네덜란드 책 논증 Dutchbook Argument과 유사하다.

과 삼인칭적 관점에 따른 상이한 해석이 있는가 하는 문제를 살펴보자. 흥미 있는 점은 이러한 두 관점 사이의 차이에 의한 접근법이 기대효용 이론의 초기에 지적이 되기는 하였지만, 그것이 이후의 발전 과정을 통하여 커다란 논쟁점으로 자리잡고 있지는 않다는 것이다. 기대효용 이론의 창시자라고 할 수 있는 폰 노이만과 모겐스톤은 기대효용 최대화에 사용되는 확률을 객관적 확률로 삼은 반면, 세비지는 주관적 확률을 계산의 근거로 제시하고 있다. 이는 각기 실천적 합리성에 대한 삼인칭적 접근과 일인칭적 접근에 대응하는 것으로 볼 수 있다. 행위자의 주관적 확률 판단에 의거한 효용성의 계산에 비추어 기대 효용치가 가장 높은 행위를 선택하는 것이 합리적이라고 주장하는 것은 실천적 합리성의 판단 기준을 일인칭적 관점에서 찾는 것이다. 반면에 주관적 확률과 달리 나타날 수 있는 객관적 확률을 효용치 계산을 위한 근거로 삼는 것은 삼인칭적 관점을 채택한 결과로 볼 수 있다.

기대효용 이론과 관련된 논쟁은 일인칭/삼인칭 사이의 대립보다는 효용의 극대화를 합리성의 기준으로 삼을 수 있는가라는 보다 근본적인 문제를 둘러싸고 제기된다. 이는 인식적 합리성의 경우와는 상반된 현상이다. 이론적 합리성의 논의에서는 참을 추구하고 거짓을 피한다는 이론적 목표 자체와 개별적 믿음이 그와의 관련 속에서 결과주의적으로 결정된다는 것에 대하여는 이론이 없었다. 그러나 실천적 합리성의 경우에는 효용의 극대화를 통하여 합리성을 정의하는 기대효용 이론이 막강한 형태도 제기되고, 그 초기에서부터 이 근본적 개념에 대한 논쟁이 시작되어 오히려 일인칭/삼인칭 사이의 논쟁은 상대적으로 주목을 덜 받은 것으로 보인다.¹⁶⁾

우선 목적 또는 선호 자체가 합리적 평가의 대상이 될 수 있다는 주장이 있다. 이런 주장의 대표적인 형태는 도덕적 추론과 실천적 합리성을 연결시키는 입장들에서 두드러지게 나타난다. 예를 들어, 나

16) 앞에서 설명하였듯이, 인식적 합리성의 논의가 여기에 집중된 것이 단지 다른 영역에서 논란이 많지 않았기 때문인 것은 아니다.

의 의지의 준칙이 보편적 입법의 원리가 되게 하라는 칸트의 주장에 따르면, 내가 일정한 목적을 세우고 그를 추구하기 위하여 하는 행위가 만약 다른 사람에게 해를 미치며 내가 피해자의 입장에 서있을 경우에 동의할 수 없는 그런 행위라면 그런 일련의 실천 행위는 용납될 수 없다는 결론이 따른다. 목적과 그에 따른 행위가 도덕과 관련된 이성적 추론에 의하여 평가받고 제약될 수 있는 셈이 된다. 계약론적 입장에서도 마찬가지로 결과가 따른다. 계약론적인 입장에서 볼 때 나의 행위에 의하여 영향을 받을 사람이 동의하지 않을 일련의 행위 과정(목적 설정을 포함한)은 실천적 합리성을 가질 수 없다.¹⁷⁾ 이런 입장들이 비도구주의적임은 자명하다.¹⁸⁾

행위 자체가 갖는 가치, 또는 그 행위를 통하여 만들어지는 상황이 좋아서 행위를 선택하는 경우들도 기대효용 이론에 대한 반론으로 제기된다. 한 사람이 효용성에 대한 고려 없이 음악회의 분위기가 좋다는 단순한 이유에서 음악회를 가는 경우가 그 예가 될 수 있다. 즉 이들은 우리의 행동 선택이 항상 목적 지향적인 것임은 아니라는 것을 지적하고 이것으로부터 실천적 행위의 규범을 설명하는 이론은 이러한 행동 선택의 원리를 규범 속에 포함하여야 한다고 주장한다.¹⁹⁾ 또 어떤 이들은 자애심에서 출발하여 또는 절제를 미덕으로 생각하여 적당한 만족에서 더 이상의 만족의 추구를 중단하는 것도 실천적 관점에서 볼 때 이성적 행위일 수 있음을 주장한다.²⁰⁾ 이들의 주장은 행위의 실천적 합리성이 단지 효용성 추구에 기여하는 바에 의하여 결정될 수 없다는 것을 주장함으로써 결과주의적 기대효

17) Rawls가 이런 입장을 갖는 철학자라고 할 수 있다.

18) 윤리적 고려가 이성적 논의의 영역 안에 있다는 것을 인정하고, 또한 목적을 설정하고 그를 성취하기 위하여 행위를 수행하는 일련의 과정이 도덕적 고려에 의하여 제약될 수 있다고 하는 입장은, 아마도 협소한 공리주의적 윤리설을 옹호하는 경우를 제외하고는, 실천적 합리성에 관한 비도구주의적 입장을 옹호하게 되리라 생각된다.

19) Kahneman and Tversky(1990).

20) Slotte(1989).

용 이론에 반대함과 동시에 실천적 합리성에 관한 비결과주의적 또는 본래주의적 대안을 제시하고 있다.

기대효용 이론을 둘러싼 논란은 기대효용 이론이 포함하는 도구주의와 결과주의에 관한 논란에 제한되지 않는다. 행위의 목표로서의 효용성이 합리적 논의의 대상이 되지 않는다는 것(도구주의 이성관)을 받아들이고, 또 행위의 비판이 효용성 성취와 관련된다는 결과주의를 받아들이면서도, 그 구체적 내용에 들어가 실천적 합리성을 기대효용의 극대화라고 정의하는 것에 문제를 제기하는 사람들도 있다. 때로는 우리의 정보력과 계산 능력의 한계 때문에 가능한 선택지들 사이의 정확한 기대 효용을 계산하고 그 중에 최고 기대 효용을 갖는 행위를 결정하는 것이 불가능하다고 주장하는 사람들이 있는가 하면,²¹⁾ 상이한 가치들이 충돌할 때 이들에 대하여 일괄적인 값을 부여하여 비교하는 것이 불가능하다는 이유에서 같은 결론을 주장하는 사람도 있다.²²⁾ 기대효용 이론은 이러한 비판들에 직면하여 여러 변수들을 도입하고 초기의 공리들을 변화시켜 가면서 계속 수정되어 가고 있다.

V. 이론적 합리성과 실천적 합리성

지금까지 우리는 이론적 합리성과 실천적 합리성을 개별적으로 살펴 펴보았다. 양자 모두 합리성이라는 공통 주제를 갖고 있으며, 비록 논의의 핵심이 서로 다른 측면에 맞추어져 있기는 하지만 양자는 유사한 문제 구조를 갖고 있으며, 아울러 이론적 합리성이 다루는 믿음이라는 것도 의사 결정의 결과로 나타나는 행위로 볼 수 있다. 이러한 사실에 비추어 볼 때, 만약 실천적 합리성 논의와 이론적 합리성 논의가 서로 교류점 없이 진행된다면 이는 놀라운 일이다.

21) Herbert Simon(1982).

22) Levi(1986).

이미 보았듯이 실천적 합리성의 논의에서 가장 중요한 쟁점은 도구주의적 이성이 실천 이성의 모든 것인가 하는 문제다. 어느 누구도 도구적 이성이 실천적 이성의 한 부분을 이루고 있음은 부정하지 않는다. 다시 말하면, 일정한 행위의 목표가 주어졌을 경우에 이 목표를 구현하기 위한 효율적 방법을 모색하는 것이 실천적 합리성의 중요한 부분임을 인정하되, 다만 이것이 실천적 이성이 하는 모든 일인가가 논란의 대상이 되고 있다. 이론적 합리성이 실천적 합리성의 바로 이 공통 부분에 관여하고 있음은 자명하다. 한 목표를 이루기 위하여 중간 단계의 잠정적 목표를 설정하는 행위, 그리고 주어진 행위가 초래하는 기대효용을 계산하는 행위 등은 모두 이론적 합리성의 통제를 받는다. 엉터리로 추정된 행위 결과에 엉터리로 확률을 부여하여 한 행위의 기대효용을 계산하고, 이 결과를 비교하여 최고의 기대효용을 갖는 행위를 선택하였다고 하자. 우리는 이러한 행위에 실천적 합리성을 부여하지 않는다. 이러한 사실은 실천적 합리성의 결정에 이론적 합리성이 개입하여 있음을 보여준다.

이제 반대 방향의 개입 여부를 살펴보자. 이론적 합리성의 결정에 실천적 고려가 개입할 수 있을까? 진리라는 숭고한 이념을 추구하는 이론적 이성은 실천적 가치로부터 진정으로 자유로운가? 어떤 명제를 알고 싶어 하며, 따라서 어떤 명제의 이론적 합리성이 문제가 되는가에 있어 인식자의 이해, 관심, 효용 등의 실천적 요소가 개입한다는 것을 주목하자. 그 뿐 아니라, 인식적 합리성의 관심이 한 명제 또는 믿음에 한정된 경우에조차도 실천적 요소가 인식적 합리성의 결정에 변수로 작용한다. 예를 들어, ‘영수가 정희를 사랑한다’는 믿음의 이론적 합리성을 검증하려 한다고 하자. 이 명제와 상반되는 명제 — 회의적 가설 — 들로는 다음과 같은 것들이 있다.

1. 영수는 정희를 사랑하는 것이 아니라 미워한다.
2. 영수는 정희를 사랑하는 것이 아니라, 주연이를 사랑한다.
3. 영수가 정희를 사랑하는 것이 아니라, 철호가 정희를 사랑한다.

누군가가 자신은 영수가 정희를 사랑한다는 것을 알고 있다고 주장한다고 하자. 그런데 나는 영수라는 사람은 지극히 이기적인 사람이어서 다른 사람을 사랑할 수 있는 마음의 여유가 없는 사람이며, 또한 대체로 주변 사람들을 미워하는 부정적인 성격을 갖고 있다고 생각한다고 하자. 이 경우에 나는 “당신은 정말 영수가 정희를 사랑한다는 것을 아는가?”라고 묻게 된다. 이 질문을 통하여 내가 알고 싶어하는 것은 영수가 사랑하는 대상이 구체적으로 어떤 사람인가 하는 것이 아니다. 나는 또한 정희를 사랑하는 사람이 누구인가에 관심을 갖고 있지도 않다. 나는 정희가 누구인지도 모를 수 있으며, 알고 있다 하더라도 누가 그녀를 사랑하는가에 관심을 갖고 있지 않다. 내가 관심을 갖는 것은 과연 영수가 다른 사람을 미워하지 않고 사랑할 마음의 여유가 있는가 하는 것이다. 따라서, 이 경우에 누군가가 이러한 사실을 입증한다면, 나는 그가 영수가 정희를 사랑한다고 믿는 것은 정당하다고 받아들일 것이다.

위의 1, 2, 3 모두는 영수가 정희를 사랑한다는 명제에 대한 회의적 가설에 해당한다. 위의 예가 시사하는 것은 어떤 사람이 영수가 정희를 사랑한다고 믿는 것이 합리적이기 위해서는, 그가 1, 2, 3의 가설 모두에 응전해야 할 필요는 없다는 것이다. 영수의 정서적 성향에 대하여 의심을 갖고 있는 나와 같은 사람이 개입한 맥락에서, “영수가 정희를 사랑한다”는 믿음이 정당하기 위해서는 1과 같은 회의적 가설이 응전되어야 하지만, 2나 3과 같은 회의적 가설은 대응되어야 할 필요가 없다. 마찬가지로 2의 가설만이 대응될 필요가 있는 경우를 생각해 볼 수 있으며, 3의 가설만이 대응될 필요가 있는 경우를 생각해 볼 수 있다. 이러한 고찰은 한 명제가 이론적으로 합리적인가는 주어진 맥락에서 인식자가 무엇을 알고 싶어하는가에 의하여 결정되며, 그 명제의 이론적 합리성을 확보하기 위하여 배격되어야 할 회의적 가설이 달라지고, 따라서 그 명제의 이론적 합리성을 위한 증거의 집합도 달라진다는 것을 보여준다.

위에서 우리는 행위자의 관심과 이해가 어떤 명제를 이론적 합리성 평가의 대상으로 삼을 것인가에 영향을 미칠 뿐 아니라, 더 나아가 선택된 명제의 이론적 합리성을 결정하기 위한 증거의 집합을 결정하는 데에도 영향을 미침을 보았다. 사실 이 두 가지가 서로 독립적이지 않음은 쉽게 알 수 있다. 이론적 합리성의 대상이 되는 명제의 선택에 작용하는 관심과 이해가 바로 그 명제의 합리성 논의를 위한 맥락을 구성하는 것이며, 그 명제의 이론적 합리성을 결정하는 증거의 집합이 바로 이 맥락에 의하여 결정되기 때문이다. 물론 이러한 사실로부터 이론적 합리성의 영역에 실천적 합리성이 개입한다는 결론이 바로 따르는 것은 아니다. 이론적 합리성의 논의를 촉발하는 맥락, 관심, 이해 등이 순수히 이론적인 것일 수 있기 때문이다.²³⁾ 다만 여기서 주목하고자 하는 것은 (구체적 실천적 문제를 해결하기 위하여 이론에 관심을 갖게 되는 흔한 경우에서 볼 수 있듯이) 이들 맥락, 관심, 이해가 실천적인 차원에 있을 수 있다는 것이다. 지구 밖으로 가고 싶어 로켓 이론이 발전하였고, 좀더 따듯하고 편하게 지내고 싶어 섬유에 관한 이론이 발전하였을 것이며, 편리하고 안전한 건물을 짓고 싶어 토목 공학이 발전한 것 아니겠는가? 이러한 추론을 진행하면 궁극적으로는 모든 이론적 합리성은 실천적 관심에서 출발하였다는 과감한 주장을 할 수 있을 것으로 보인다.

지금까지 우리는 실천적 합리성의 논의에 이론적 차원이 개입하며, 이론적 합리성의 논의에 실천적 차원이 개입함을 보았다. 그런데 위에서 살펴본 예들을 종합하면 실천적 차원이 이론적 차원의 상위에 위치하는 듯이 보인다. 실천적 합리성을 결정함에 있어 이론적 합리성이 영향을 미치기는 하되 그것은 다만 행위의 효용을 계산하는 데에 수동적으로 봉사하는 정도의 역할이고, 반면에 실천적 차원은 이

23) 우주의 출발에 관심을 갖는 과학자가 일정한 가설을 구성하고 이를 실험을 통하여 검증하는 경우에서와 같이 순수한 이론적 관심에 의하여 일정한 명제가 이론적 합리성 평가의 대상으로 선택될 수 있으며, 이런 이론적 관심이 다시 선택된 명제의 이론적 합리성을 결정하는 증거의 집합을 결정할 수 있다.

론적 합리성의 문제가 되는 명제를 선택하고 그와 관련된 증거의 범위를 결정하는 방식으로 상위에서 적극적인 규정적 역할을 하고 있기 때문이다. 그러나, 반대의 경우, 즉 이론적 차원이 실천적 차원을 제약하는 경우도 있다. 우리의 속담에 오르지 못할 나무는 쳐다보지도 말라는 말이 있다. 이를 수 없는 목표는 애초부터 바라지도 않는 것이 낫다는 주장이다. 자신이 원하는 포도를 따먹을 수 없는 여우가 그 포도는 어차피 시어서 맛이 없다고 판단하는 이솝 우화의 여우는 이러한 교훈을 잘 실천하고 있다.²⁴⁾ 그런데 이 교훈은 설정된 목표가 도달할 수 없다는 것을 알고 있는 사람에게만 적용될 수 있다. 그 목표가 도달할 수 있는 것이라고 믿는 사람에게에는 위의 실천적 교훈이 아무 효력을 갖지 못한다. 이로부터 우리는 이론적 영역이 실천적 합리성에 영향을 미치는 한 방식을 찾아낼 수 있다. 여기서 제시되는 교훈의 핵심은 다음과 같이 요약된다: 당신이 만약 현재의 목표가 달성될 수 없는 것인 줄을 **안다면**, 목표를 변경하는 것이 **합리적**이다. 이 조건문의 전건은 지식과 관련된 이론적 합리성의 영역이며, 후건은 명백히 목표의 설정과 관련된 실천적 합리성의 영역이다.

위의 교훈은 우리가 이루지 못할 목표를 갖고 있을 경우에 이를 수 있는 목표로 방향 전환을 하라는 실천적 교훈을 담고 있다. 위의 논의를 한발 더 발전시키면, 실제로 우리가 지향하는 목표들은 이미 이론적 차원의 여과를 거쳤다고 주장할 수 있을 듯하다. 목표가 이를 수 없는 것이라는 것을 알 때 우리가 상당한 갈등을 겪는다는 사실은 거꾸로 우리가 갈등 없이 추구하는 목표는 이미 이론적 차원에서의 실현가능성이라는 판단을 통과하였음을 시사한다. 이 경우에 이론적 차원의 검증이란 물론 무의식적인 차원에서 일어나는 것이라고 보아야 할 것이다.

지금까지의 논의를 다시 요약하자. 이론적 차원의 합리성 논의는 실천적 차원의 영향을 받는다. 실천적 차원에서 어떤 목표를 설정하

24) 현실적 고려에서부터 목표가 바뀌는 것에 관한 상세한 논의를 위하여는 Elster(1983)을 보라.

는 것이 합리적인가는 이론적 차원의 영향을 받는다. 물론 모든 이론적 문제가 궁극적으로 실천적 차원의 영향을 받는지, 모든 실천적 목표 설정이 이론적 차원의 영향을 받는지의 미래의 연구를 필요로 하는 부분이다. 그러나 이상의 논의는 이론적 합리성의 문제와 실천적 합리성의 문제는 유사한 논리적 구조를 가지고 있을 뿐 아니라 서로 간 문제와 목표의 설정에 영향을 미치고 있음을 보여준다. 실천적 차원의 논의가 결국 이론적 차원의 논의의 합리성에 영향을 미치고, 이론적 합리성의 논의가 다시 실천적 차원에서의 목표 설정의 합리성에 영향을 미친다. 이렇듯 이론적 차원과 실천적 차원은 상호간 영향을 미치면서 변증법적으로(?) 진행되는 것으로 보인다.

그렇다면 이론적 합리성의 영역과 실천적 합리성의 영역은 상하의 차원에서 서로 앞서거나 뒤서거나 하면서 조화롭게 진행되기만 하는 것인가? 이들이 서로 충돌하는 경우는 없는가? 이런 경우가 혹시 있다면, 어떤 것에 우선성을 부여하여야 하는가? 다음과 같은 경우를 보자. 영수는 심각한 질병을 앓고 있으며, 그는 그렇게 믿을 만한 충분히 많은 정당한 이유들을 갖고 있다고 하자. 그러나, 영수는 심각한 사람이어서 자신이 증병을 앓고 있다는 것을 받아들일 경우에 심각한 정신적 타격을 받을 것이고 이는 결국 병으로부터 회복되는데 막대한 장애가 된다고 하자. 이러한 상황에서 영수는 자신이 앓고 있는 병은 심각하지 않으며 그는 그 병에서 쉽게 회복될 것이라고 믿는다고 하자. 이 믿음은 믿음의 소유자에게 유익한가라는 관점에서 볼 때 바람직한 믿음이다. 왜냐하면, 그렇게 믿음으로 하여 영수는 마음의 평화를 얻을 뿐 아니라 그러한 평화는 영수가 현재 앓고 있는 질병으로부터 회복하는데 실질적으로 도움이 되기 때문이다. 따라서, 영수의 믿음은 효용성의 관점에서 합리적이라 할 수 있다. 그러나 영수의 믿음이 이론적으로 합리적이지는 않다. 그 이유는 그 믿음을 찬이게 하는 근거를 영수는 갖고 있지 않기 때문이다. 이론적 합리성과 실천적 합리성이 상충하는 진단을 내린다. 이 경우에 이론적 합리성과 실천적 합리성을 포괄하는 상위의 차원에서, 즉 합리성 자

체의 차원에서 한 진단의 손을 올려주기는 쉬운 일이 아니다. 이에 대한 최종적인 결론은 상당한 논의를 필요로 하는 것으로 보인다. 이 글의 나머지에서는 이러한 해결로 가기 위한 첫걸음을 놓는 의미에서 이론적 합리성에 대한 실천적 합리성의 상대적 우위를 시사하는 한 논증을 제시하고자 한다.

여기서 보이고자 하는 것은 실천적 차원의 고려가 이론적 합리성에 영향을 미치는 방식과 이론적 고려가 실천적 합리성에 영향을 미치는 방식 사이에 존재하는 비대칭성이다. 다음과 같은 경우를 보자. 한 사람이 어느 날 갑자기 한 컵에 들어 있는 물 분자의 수를 결정하는 일에 관심을 가졌다고 하자. 이 문제가 해결 불가능한 것은 아니지만 상당한 시간과 노력을 요한다고 하자. 그리고 그 시간과 노력을 기울이면, 더욱 많은 참인 명제를 찾아 낼 수 있다고 하자. 이 경우에 모래알의 수를 찾고자 하는 이론적 행위는 합리적인가? 이 행위는 실천적으로 합리적이지 않은 듯하다. 같은 시간과 노력으로 자신에게 더욱 유용한 많은 진리 명제를 찾아 낼 수 있기 때문이다. 효율성을 고려한 실천적 관점에서 이러한 인식 행위가 합리적이지 않은 것으로 판단되는 이유는 단지 그 행위를 통하여 얻을 수 있는 참인 명제의 수가 작기 때문은 아니다. 비록 많은 시간과 노력을 기울여 오직 하나의 참인 명제만을 찾아낼 수 있다 하더라도 그 명제가 매우 유용한 것이라면, 그 명제의 참을 확인하는 작업이 실천적으로 불합리하다는 판단을 내리지 않는다는 것이다. 위의 경우에 있어서도, 만약 물의 부자 수를 확인하는 것이 주어진 상황에서 실천적으로 중요한 의미를 지닌다면 그러한 행위를 실천적으로 불합리하다고 판단하지 않을 것이다.

이제 반대의 방향에서 유사한 문제가 발생한다고 하자. 즉, 한 사람이 효율성을 목표로 하여 하는 행위를 이론적인 관점에서 비판하는 경우를 살펴보자. 이러한 맥락에서의 A 행위에 대한 다음과 같은 모습을 지닐 것이다: A 대신에 B의 행위를 하였더라면, 인식적으로 더욱 우월할 것이기 때문에 A의 행위는 인식적으로 합리적이지 않

다. 앞서 이론적 행위를 실천적 관점에서 비판하는 경우와는 달리, 이 경우에는 특이한 문제가 발생한다. 첫째로, 이 비판은 인식적 차원의 우월성을 비교하기 위하여 순수히 계량적인 수에 의존하여야 한다. 순수히 수적인 차이에 의존하지 않고 B의 행위에 동반하는 진리 명제의 중요성을 언급함으로써 B가 갖는 인식적 우월성을 이야기하는 것은 순수히 인식적인 평가가 아니다. 이는 효용성과 관련된 실천적인 평가를 뒷문으로 들여오는 것이다. 둘째로, 순수히 참인 명제의 수에 의존하여 B 명제의 인식적 우월성을 주장하려 한다고 하자. 이 경우에는 설득력이 없는 주장이 되고 만다. B를 하였을 때 비록 많은 수의 참인 명제들이 획득될 수 있다 하더라도 그 명제들이 모두 하찮은 것들인 반면 A를 하였을 때 비록 적은 수의 참인 명제들 일지언정 중요한 명제들이 얻어진다면, B가 A에 비하여 이론적으로 우월하다는 판단은 아무런 직관적 호소력을 갖지 못한다. 결국 실천적 행위를 이론적 관점에서 비판하는 논의는 이론적 행위를 실천적 관점에서 비판하는 경우와 달리 다시 실천적 차원으로 논의가 되돌아가게 된다.

위 논증은 이론적으로 합리적인 명제를 실천적인 차원에서 비판하는 경우와 실천적으로 합리적인 명제를 이론적인 차원에서 비판하는 경우 사이에 비대칭성이 존재함을 보여준다. 이론적 차원의 행위는 실천적 차원에서 일관되게 비판하는 것이 가능하지만, 실천적 차원의 행위를 이론적 차원에서 비판하는 것은 딜레마에 빠지게 된다. 즉, 진리의 중요성을 언급하면 인식적 차원의 비판이 설득력을 갖지만, 진리의 중요성은 실천적으로 결정되기에 이는 순수히 이론적 비판이 될 수 없다. 반면에 진리의 양만을 언급하면, 그 행위 자체가 이론적으로 비판 받을만한가가 불명확해진다. 위 논증이 어떤 의미를 갖는가에 대하여 다양한 해석이 있을 수 있다. 가능한 한 해석은 이론적 평가의 기준 속에 효용성이 한 요소로 자리를 잡고 있다고 보는 것이다. 이러한 고찰이 이론적 합리성과 실천적 합리성이 정면으로 충돌하는 경우를 포함하여, 양자의 전반적 역학 관계를 해명하는 과정

에 중요한 역할을 하리라고 기대한다. 그것에 대한 구체적인 논의는 앞으로의 숙제로 남겨두자.

참고문헌

- 김기현, 『인식적 합리성의 두 개념』, 한국 분석철학회편, 『합리성』(서울: 철학과 현실사, 1998): 102-132.
- Alston, William 1986, “Internalism and Externalism in Epistemology”, *Philosophical Topics* 14: 179-221.
- Chisholm, Roderick 1968, “Lewis’ Ethics of Belief”, in *The Philosophy of C. I. Lewis*, ed. P. A. Schlipp (La Salle, Illinois: Open Court).
- _____, 1977, *Theory of Knowledge*, 2d ed. (Englewood Cliffs, N.J.: Prentice-Hall).
- _____, 1982, “A Version of Foundationalism”, in *The Foundations of Knowing* (Minneapolis: University of Minnesota Press).
- _____, 1986, *Contemporary Theories of Knowledge* (Totowa, N.J.: Rowman and Littlefield).
- Elster, J. 1983, *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge, MA: Cambridge University Press).
- Feldman, Richard and Conee, Earl 1985, “Evidentialism”, *Philosophical Studies* 48: 15-34.
- Ginet, Carl 1975, *Knowledge, Perception and Memory* (Dordrecht: D. Redidel).
- Goldman, Alvin 1992, “What Is Justified Belief?”, in *Liasons* (Cambridge, MA: MIT press).
- Kahneman, D. and Tversky, A. 1990, “Rational Choice and the Framing of Decisions”, in K. Cook and M. Levi (eds.) *The Limits of Rationality* (Chicago, IL: Chicago University Press): 60-89.

- Kim, Kihyeon 1993, "Internalism and Externalism in Epistemology." *American Philosophical Quarterly* 30: 303-316.
- Levi, I. 1986, *Hard Choices: Decision-Making under Reresolved Conflict* (New York: Cambridge University Press).
- Moser, Paul 1985, *Empirical Justification* (Dordrecht: D. Redidel)
- Naylor, Margery 1988, "Epistemic Justification", *American Philosophical Quarterly* 25: 49-58.
- Neumann, J. von and Morgenstern, O. 1944, *Theory of Games and Economic Behavior* (Princeton, NJ: Princeton University Press).
- Pollock, John 1979, "A Plethora of Epistemological Theories." In G. Pappas ed., *Justification and Knowledge* (Dordrecht: D. Redidel).
- Savage I. 1954, *The Foundations of Statistics* (New York: Wiley & Sons).
- Simon, Herbert 1982, *Models of Bounded Rationality* (Cambridge, MA: Cambridge University Press).
- Slote, M., 1989, *Beyond Optimizing: A Study of Rational Choice* (Cambridge, MA: Cambridge University Press).
- Swain, Marshall 1981, *Justification and Knowledge* (Ithaca, New York: Cornell University Press).
- Wolterstorff, Nicholas 1983, "Can Belief in God be Rational If It Has No Foundations?", in *Faith and Rationality*, ed. Alvin Plantinga and Nicholas Wolterstorff (Notre Dame, Indiana: University of Notre Dame press, 1983).