https://doi.org/10.33156/philos.2020.57..005

반응적 태도 이론과 책임의 귀속 문제

박 의 연*

주제분류 윤리학, 서양철학, 영미철학

주 요 어 반응적 태도 이론, 책임, 월리스, 맥케나, 스트로슨

요 약 문

이 논문의 목적은 반응적 태도 이론을 통해 책임 개념이 안정적으로 설명될 수 있음을 주장하는 것이다. 반응적 태도 이론은 행위자의 행위에 대한 동료 행위자의 반응적 태도가 책임을 귀속시키는 근거가 된다고 보는 입장이다. 반응적 태도의 적절성은 나와 상대가 공유한 상호작용적 관계, 즉 관행과 그것을 기반으로 서로에게 가지는 기대에 달려있다. 이런 방식으로 책임의 귀속 여부가 결정된다면, 결정론이 참이라고 해도 책임은 여전히 행위자에게 귀속된다.

필자는 반응적 태도가 우리 책임 귀속의 근거가 된다는 스트로슨의 주장에서 시작해, 그러한 주장을 규범적 개념들을 통해 제약하고 명료화 한 월리스와, 관행이 어떻게 반응적 태도로 하여금 책임을 귀속시키게 하는지 설명한 맥케나의 논증을 두루 살펴본다. 스트로슨은 반응적 태도를 통해 책임을 귀속시키는 시도를 한 최초의 연구자임에도, 개념의 애매모호함과 문제적 관행을 제약할 수 있는 설명을 제시하지 않았다는 점에서 비판받는다. 월리스는 이러한 문제를 해결하기 위해 규범적 개념들을 도입해 반응적 태도 이론을 체계적인 형태로 구성해 냈지만, 책임의 귀속에서 핵심적인 역할을 하는 관행의 역할을 간과한다. 필자는 이러한 문제를 맥케나의 대화이론을 통해 보완할 수 있다고 주장할 것이다. 반응적 태도 이론은 건전하고 새로운 책임 개념의 이해방식이며, 책임의 주체로서의 행위자성에 관하여 더 확장적인 이해를 제공하는 단초가 될 수 있다.

* 전남대학교

www.kci.go.kr

1. 서론

우리는 기존의 방식으로는 해결할 수 없는 도덕 문제를 매일 마주한다. 고도로 발전한 기술과 통제할 수 없는 환경 문제가 공존하는 시기에 무엇을 해야 하는지에 관한 고민은 늘 새로운 양상으로 대두된다. 개인의 옳은 신념만으로는 해결할 수 없는 문제들에 새로운 방식으로 접근하고자 도덕철학은 지금도 다양한 길을 탐색하고 있다. 책임 개념은 그러한 시도들 중 하나다. 책임은 도덕적 대화에 흔히 등장하는 개념임에도 기존의 윤리학 논증에서 쉽게 활용되지 못했다. 그 이유는 자유의지와 결정론 사이의 긴장관계에 있을 것이다. 만약 결정론이 참이라면, 자신의 행위에 대한 궁극적 통제력이 없는 행위자에게 책임을 귀속시키는 것이 합당한지 묻는 물음에 답해야하기 때문이다. 반응적 태도 이론(reactive attitude theory)은 이러한 난문을 해결하여 책임 개념에 대한 설득력 있는 설명을 제공하려는 시도 가운데 하나다.

도덕적 문제 상황에서 책임 개념은 옳음 혹은 그름만큼이나 강력하게 작동하지만, 옳음에서 곧바로 도출되는 개념은 아니다. 그러면 우리는 언제 도덕적 책임을 물을 수 있으며, 그 책임은 어느 경우에 어떤 사람에게 귀속되는가? 그리고 그에게 책임을 귀속시키는 것은 무엇이며, 책임의 귀속은 어떤 함의를 가지는가? 필자는 이 논문을 통해 반응적 태도 이론이 이러한 문제들에 어떻게 답하는지 확인할 것이다.

반응적 태도 이론은 1962년 P. F. 스트로슨(P. F. Strawson)이 반응적 태도를 기반으로 행위자에게 책임을 귀속시킬 수 있음을 주장하며 시작되었다. 이 논문의 두 번째 장은 반응적 태도 이론을 태동하게 한 스트로슨의 논증을 소개하며, 반응적 태도 이론의 전체적인 얼개를 보여줄 것이다. 그 다음에는 규범적 개념들을 통해 스트로슨의 주장을 재해석함

www.kci.go.kr

으로서 우리의 반응적 태도가 어떻게 제약되는지 논증한 제이 월리스(R. Jay Wallace)의 입장을 상세하게 소개한다. 월리스는 기대, 책무, 반성적 자기 통제력이라는 개념을 통해 반응적 태도 이론의 규범적 해석을 제안한다. 그러나 월리스의 해석 속에는 귀책을 가능하게 하는 관행(practice)의 자리가 없다. 필자는 마이클 맥케나(Michael McKenna)의 대화 이론을 통해 이러한 한계를 보완할 수 있다고 주장할 것이다. 맥케나는 우리의 책임 귀속 관행이 흡사 모국어를 구사하는 화자들 간의 대화와 유사하다며, 공동체 안의 행위자들이 서로 책임을 지고, 지워주는 과정을 세 단계로 나누어 설명한다.

반응적 태도 이론가들은 다양한 방식으로 반응적 태도 이론을 전개해 나가지만, 기본적으로 다음 세 가지에 동의한다.[1] 첫째, 도덕적으로 책임이 있다는 것은 특정 행위자에게 도덕적으로 책임이 귀속된다는 것을 의미하며, 책임이 귀속된다는 것은 곧 그 행위자가 비난과 칭찬의 대상이 될 수 있다는 것을 의미한다. 둘째, 반응적 태도는 책임 귀속의 본성에서 핵심적인 역할을 한다. 셋째, 행위자에 대한 비난과 칭찬에서 가장 중요한 것은 그들이 행위하며 가지는 의지의 질(quality of will)이다. 즉 이들

1) Brandon Warmke and Michael McKenna, “Moral Responsibility, Forgiveness, and Conversation” in *Free Will and Moral Responsibility* (Newcastle: Cambridge Scholars Publishing, 2013), ed. Ishtiyaque Haji and Justin Caouette 190쪽 참조. 웜크와 맥케나는 이 세 가지가 스트로슨과 본인들의 입장이 가지는 공통점이라고 말한다. 이 공통점은 다른 연구자들에게도 일반적이다. Gary Watson, “Responsibility and the Limits of Evil: Variations on a Strawsonian Theme,” in *Responsibility Character and the Emotions* (Cambridge: Cambridge University Press, 1987), ed. Ferdinand Schoeman, 256-86쪽, John Martin Fischer and Mark Ravizza, *Responsibility and Control* (Cambridge: Cambridge University Press, 1998), Coleen Macnamara, “Reactive Attitudes as Communicative Entities,” *Philosophy and Phenomenological Research*, Vol. 90, No. 3 (2015), 546-69쪽, Gideon Rosen, “The Alethic Conception of Moral Responsibility,” in *The Nature of Moral Responsibility New Essays* (New York: Oxford University Press, 2015), ed. Randolph Clarke, Michael McKenna, and Angela M. Smith 65-87쪽 참조.

www.kci.go.kr

이 악의를 가지고 의도적으로 해를 끼친 것인지 원치 않았지만 어쩔 수 없이 행위했는지는 중요한 요소다. 이러한 반응적 태도 이론에서는 책임을 귀속시키는 직접적 근거는 나의 완전한 자유가 아니라, 나의 행위와 그에 대한 동료 행위자의 반응적 태도라고 간주한다. 또 나의 행위가 타인에게 반응적 감정을 불러 일으킬만한 행위인지는 나와 상대가 공유한 상호작용적 관계와 그에 따라 서로에게 가지는 기대에 달려있다. 이러한 상호작용적 관계는 관행에 기반해 구성되며, 이런 방식으로 책임의 귀속 여부가 결정된다면 결정론이 참이라고 해도 책임은 여전히 행위자에게 귀속된다.

이 논문을 통해 필자가 보여주고자 하는 것은 반응적 태도 이론의 안정적인 설명력이다. 반응적 태도를 근거로 하는 책임 개념이 건전한 논증을 통해 설득력을 갖출 수 있다는 것이 이 논문에서 주장하고자 하는 바이다. 반응적 태도 이론은 새로운 책임 개념의 이해방식임에도 기존의 도덕직관과 상충하지 않는다. 또한 반응적 태도 이론을 통한 책임 이해는 책임의 주체가 되는 행위자성에 관하여 더 확장적인 이해를 제공하는 단초가 될 수 있을 것이다.

2. 반응적 태도 이론의 기원: P. F. 스트로슨

스트로슨의 1962년 논문 「자유와 분노」(“Freedom and Resentment”)는 책임 논의의 역사에서 가장 중요한 분기점 중 하나다. 이 논문에서 스트로슨은 서로에게 책임을 묻는 관행을 우리가 오랫동안 공유해 왔다는 사실에 주목한다. 그의 성찰은 자유의지와 결정론의 논쟁에 가로막혀 있던 책임 논의를 전면적으로 전환한다. 우리가 진정 가진 것인지 확인할 수 없는 개념적 자유와 통제 가능성 보다는, 서로에게 책임을 묻는 관행과 그에 따른 사람들의 감정이 책임 개념과 더욱 밀접하지 않느냐는 문제제

www.kci.go.kr

기인 것이다. 우리의 반응적 태도가 어떻게 도덕적 요구가 되는지 추적한 이 논의는 후에 월리스, 맥케나, 존 마틴 피셔(John Martine Fischer), 마크 라비짜(Mark Ravizza), 콜린 맥나마라(Coleen Macnamara) 등의 연구자들에 의해 반응적 태도에 기반을 둔 도덕적 책임에 관한 논증으로 발전되었다. 이들의 주장은 '반응적 태도 이론'이라는 이름으로 불리는 것이 일반적이다.[2] 이 장에서는 최초의 반응적 태도 이론이라고 볼 수 있는 스트로슨의 주장이 어떤 형태이고, 어떤 의의와 한계를 가지고 있는지 고찰하고자 한다.

논문「자유와 분노」가 본래 해결하고자 하던 문제는 도덕철학이 아니라 존재론의 것이었다. 이 논문에서 스트로슨은 우리가 실제 삶에서 서로에게 도덕적 기대와 요구를 하는 과정이 결정론의 진위와 무관하게 지속될 수 있다는 점을 보여주며, 책임을 귀속시키는 관행과 결정론이 양립가능하다고 주장한다. 주장의 목표도 물론 중요한 논점을 지적하고 있었지만, 그 과정에서 논의된 책임 개념의 내용은 기존 책임 논의의 방향을 전환시킨다. 실제로 우리에게 책임을 부과하는 것은 우리가 상호작용하며 서로에게 가지게 된 태도라는 스트로슨의 책임 개념은 실천적인 동시에 설득력 있는 논증의 단초를 제공한다.

'반응적 태도'[3]라는 용어 역시 스트로슨에 의해 처음으로 중요하게 해

2) Robert Kane, *A Contemporary Introduction to Free Will* (New York: Oxford University Press, 2005), 10장 참조. Matthew Talbert, "Moral Responsibility," *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/win2019/entries/moral-responsibility/> 참조.

3) 스트로슨의 논문을 번역한 최용철은 reactive attitude를 '반향적 태도'라고 번역한다. 하지만 다른 국내 논문들에서도 반응적 태도라는 번역어를 자주 사용하고, '기대에 대한 반응으로서의 도덕'이라는 중심 의미를 생각해 봤을 때 반응적 태도가 더 적절하다고 필자는 판단했다. 그런 이유로 이 논문에서는 반응적 태도라는 번역어를 차용한다. 피터 스트로슨,「자유와 분노」,『자유의지와 결정론의 철학적 논쟁』최용철 편, 최용철 역 (서울: 간디서원, 2004) 141-186쪽, 김양현 외,「책임 개념에 대한 실천윤리적 해명」,『범한철학』, 39집 (2005), 291-328쪽, 양선이,「중

www.kci.go.kr

석되기 시작한다. 스트로슨의 책임개념을 요약하자면, 책임은 인간이 서로에게 반응하며 기대하게 된, 일종의 태도로서의 사회적 관행이다. 여기서 책임을 질 수 있는 행위자는 분노, 감사, 분개, 비난, 용서와 같은 반응적 태도의 주체이자 대상이다. 타인의 행위에 대한 행위자의 태도는 그에게 책임을 귀속시키는 근거임과 동시에, 자신 역시 타인의 반응적 태도에 따른 책임의 귀속과 비난을 감수해야한다고 강제하는 근거가 된다. 따라서 중요한 것은 우리가 서로에게 하는 기대와 이 기대의 배경이 되는 사회의 관행이지, 완전히 비결정적인 자유의지와 그에 따른 행위의 통제 가능성이 아니다. 간단히 말해서 우리가 상대를 특정한 반응을 기대해도 되는 행위자로 간주하는 경우, 이는 곧 우리가 서로를 도덕 행위자, 즉 도덕적 책임을 지는 존재로 이해하는 것이고, 상대 역시 이런 과정을 통해 나를 규정한다.

스트로슨은 반응적 태도를 세 가지, 즉 인간적(the personal), 대리적(the vicarious), 자기 반응적(the self-reactive) 태도로 구분한다.[4] 예를 들어 어떤 사람이 내 발을 밟는다면 우리는 심한 고통과 함께 즉각적으로 화가 난다. 나와 동등하다 생각되는 행위자에게 우리는 일반적인 범위의 행동을 기대하는데, 그러한 기대의 배반에 따른 분노(resentment)는 우리가 갖게 되는 반응으로서의 감정이다. 이것이 첫 번째 태도인 '인간적 태도'로서의 반응이다.[5] 합당한 판단력이 있는 성인이 나에게 예상하지 못한 손해를 입혔을 때 불쾌한 감정이 솟아난다는 것에 추가적인 논증을 할 필요는 없을 것이다.

독과 도덕적 책임」, 『철학연구』, 109집 (2015), 191-216쪽 참조.

4) 스트로슨은 반응적 태도를 세 종류로 나누지만 그에 따른 이름을 붙이지 않았다. 이 논문에서 필자는 반응적 태도 이론을 주장하는 후대 연구자들의 명명을 따라 스트로슨의 분류를 인간적, 대리적, 자기 반응적 태도라 부를 것이다. Robert Kane, *A Contemporary Introduction to Free Will*, 10장 참조. Deborah P. Tollefsen, *Group as Agents* (Cambridge: Polity Press, 2015), 6장, 특히 120쪽 참조.

5) 피터 스트로슨, 「자유와 분노」, 155쪽 참조.

www.kci.go.kr

‘인간적 태도’로서의 분노는 상대를 확인하거나 의사소통을 한 이후 사그라지는 경우가 많은데, 분노가 사라지며 도덕적 요구 역시 사라지는 방식은 다음 두 가지다. 우선 그 행위자에게 불가피한 사정이 있었다면 그는 책임을 면할 수 있다.[6] 이러한 경우 행위자는 책임에서 일시적으로 면제되지만 그가 여전히 도덕적 행위를 할 능력이 있는 행위자라는 사실은 의심받지 않는다. 두 번째로 상대가 도덕적 판단을 내리기에 합당한 행위자가 아니라고 간주되는 경우에도 도덕적 책임은 사라진다.[7] 내 발을 밟은 사람이 어린아이거나 정신적·육체적 장애를 가지고 있는 경우, 우리는 늘 적용하던 상호작용의 일환으로서의 반응적 태도, 즉 ‘인간적 태도’를 중단한다. 이 때 사람들은 상대가 나와 동등한 관계를 맺을 수 있을 것이라는 기대를 접고 ‘객관적 태도’(the objective attitude)를 취하게 된다.[8] 여기서 객관적이라는 것은 서로에게 책임을 귀속시키는 상호

6) 스트로슨, 「자유와 분노」, 152-54쪽 참조. 스트로슨은 전자를 “인간관계에의 참여 혹은 관여의 태도”, 후자를 “타인에 대한 객관적 태도”라고 부른다. (「자유와 분노」, 155쪽.) 이것을 후대 연구자들은 ‘상호작용적 태도’와 ‘객관적 태도’라고 줄여 부르기도 한다. 그러나 이렇게 ‘-태도’ 라는 명칭을 쓰면, 반응적 태도의 세 범주인 인간적, 대리적, 자기 반응적 태도와 혼동하기 쉽다. 그래서 연구자들은 그 특성을 중심으로 새로운 지칭을 제안하곤 한다. 변명 조건과 면제 조건이라는 명명은 왓슨의 것이고, 국지적 변명과 전면적 변명이라는 명명은 맥케나의 제안이다. 필자는 혼동을 예방하기 위해 왓슨의 용어를 주로 사용할 것이다.
Gary Watson, “Responsibility and the Limits of Evil: Variations on a Strawsonian Theme,” 256-86쪽과 Michael McKenna, “The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism,” in *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's “Freedom and Resentment”* (Farnham: Ashgate Publishing, 2008), ed. Michael McKenna and Paul Russell, 201-18쪽 참조.

7) 스트로슨, 「자유와 분노」, 154쪽 참조.

8) 같은 책, 154-56쪽 참조. 개리 왓슨은 ‘도덕적 지목’(moral address)이라는 개념을 통해 이런 방식의 이해를 부연한다. 왓슨은 도덕적 능력이 없는 상대에게 도덕적 요구를 하면 나의 분노는 갈 곳을 잃는다는 점을 지적하며, 내가 반응적 태도를 보인다는 것은 상대가 도덕적 행위를 할 수 있는 능력을 가짐을 전제하고 있다는 “도덕적 지목의 조건” (Gary Watson, “Responsibility and the Limits of Evil:

www.kci.go.kr

적 관계에서 빠져 나와, 반응적 감정 없이 사건을 객관적으로 조망한다는 의미다. 책임에서 벗어나는 첫 번째 경우를 변명조건, 두 번째를 면제조건이라고 부르며, 이 두 조건이 하나라도 만족되는 경우 책임은 부여되지 않는다.

우리가 타인의 입장에 감정이입 함으로서 반응적 태도를 취하게 되는 경우도 있다. 이것이 두 번째, '대리적 태도'다.[9] 앞의 인간적 태도는 타인이 내게 보이는 선의, 악의, 무심함 등에 대한 직접적 반응이었다면, 대리적 태도는 내가 가진 인간적 태도가 공감을 기반으로 타인의 일에까지 일반화 된 것이다.[10] 내 이익이나 존엄과는 무관한 사건에 타인을 대신해 분노하는 것을 우리는 도덕적 분개(moral indignant)라고 부른다. 누군가 타인에게 악의적으로 해를 가하는 것을 목격한 경우, 저런 행동이 허용되어서는 안 된다고 주장하거나 피해자를 보호하려 나서는 경우가 드물지 않다. 이런 분개는 타인을 대리해서 일어난다는 점에서, 즉 직접 인간을 대면해 발생하는 것이 아니라는 의미에서 무인격적이다. 하지만 우리가 흔히 경험하는 감정이기도 하다.

도덕적 분개는 타인 뿐 아니라 나 자신을 향하기도 한다. 타인에게 해를 입히는 행위를 하지 말아야 하는 것은 나 역시 마찬가지다. 우리는 타인의 잔인함이나 무책임함에 분노하며 자연스럽게 나 역시 잘못하는 것이 없는지 확인한다. 타인을 위해 자신에게 규범을 부여하는 태도가

Variations on a Strawsonian Theme," 266쪽.)을 말한다. 이런 도덕적 지목의 대상이 되기 위해서는, 우선 충분한 도덕적 이해를 할 능력이 있어야 한다. 또 그가 한 행위가 그 자신의 도덕적 자아(moral self)를 반영해야 한다. 이런 도덕적 요구를 받을 능력, 즉 도덕적 지목의 대상이 되고 또 도덕적으로 누군가를 지목할만한 능력을 가지지 못한 자들은 전면적으로(globally) 도덕 행위자에서 제외된다. (같은 책, 266-67쪽 참조.)

9) 스트로슨은 "개인적 반응적 태도의 일반화 된 혹은 대리적인 유비적 상태"(the generalized and vicarious analogue of the personal reactive attitude)라고 부른다. 스트로슨, 「자유와 분노」, 166쪽.

10) 같은 책, 164-65쪽 참조.

www.kci.go.kr

세 번째, “자기 반응적 태도”[11]다. 규범성을 향한 추동과 관련된 다양한 감정들이 이런 자기 반응적 태도를 통해 드러난다. 가책, 죄책감, 후회 등이 이런 반응으로 생겨나는 감정이고, 이것들은 수치심과도 관련되어 있다고 스트로슨은 설명한다.[12]

반응적 태도가 사회 내의 관계와 우리의 감정에서 직접적으로 도출된다는 사실이 보여주고자 하는 지점은, 자유의지와 결정론을 중심으로 책임 개념을 논의하던 과거의 방식에 잘못된 단절이 있다는 것이다. 개인이 완전무결한 자유의지를 가지고 상황을 통제하는 경우에만 그 행위자에게 책임을 부여하고 비난할 수 있다는 과거의 이해방식은, 그 행위나 행위의 대상이 되는 행위자를 실천적 관계 속에서 통합적으로 설명하지 못한다. 다시 말해 이런 방식으로는 가해자와 피해자가 맺고 있는 관계가 들어설 자리가 없다. 기존의 방식에서 개인은 서로 분리되어 있기 때문에 피해자 혹은 같은 공동체의 일원인 타인이 특정 행위자에게 책임을 귀속시키고 그에 따른 도덕적 비난 혹은 칭찬을 할 때 추가적 근거가 요청된다.

반면 스트로슨은 우리가 인간 사회의 일원이라는 이유만으로 서로의 예상 가능한 행동에 대한 공동의 이해를 이미 갖추고 있고, 그에 따라 행위한다는 점을 책임 귀속의 가능 근거로 삼는다.[13] 이러한 방식의 책임 개념이 가지는 특징은 다음 두 가지로 정리될 수 있다. 첫째, 반응적 태도는 그 적절성을 판단함에 있어 사회적 맥락을 반영한다. 우리가 상대에게 어디까지 기대해도 될 것인지, 또 어떤 반응적 태도가 적절한지는 사회의 관행과 우리의 경험에 따라 결정된다. 반응적 태도는 인간의 욕구 혹은 충동이 단독으로 작동한 결과가 아니다. 사회의 관행과 나의 숙고된 욕구가 결합할 때 책임을 귀속시키는 힘을 가진 반응적 감정이

11) 같은 책, 167쪽.
12) 같은 곳.
13) 같은 책, 168쪽.

www.kci.go.kr

생겨난다. 반응적 태도는 인간의 욕구가 단독으로 작동한 결과가 아니다.

둘째, 반응적 태도에 따른 책임의 요구는 내가 상대 행위자를 평등한 행위자로 간주하고 있다는 전제를 함의한다. 우리가 서로에게 책임을 부여하고 그 책임에 따른 행위를 요구하는 것은, 특정 반응을 요구하는 사회의 일원으로 평등하게 상호작용 하는 과정에 일어나는 일이지 처벌을 목표로 일어나는 일이 아니다. 책임의 귀속에 따른 비난 역시 처벌이 아니라 도덕적 요구의 한 양상일 뿐이다. 책임의 귀속은 평등한 공동체의 일원이 되기 위해 우리가 서로에게 부여하고 또 수용하는 것이다. 상대가 나와 같은 지위를 가진 행위자가 아니라면 우리의 반응적 태도는 달라지고 그에게 책임 역시 요구하지 않는다.

그러나 스트로슨의 주장에는 해명되지 않은 지점이 남아있다. 대표적인 비판은 다음 두 가지로 추려질 수 있다. 첫째, 반응적 감정, 태도, 상호적 관계, 대리성 등 주요 개념들이 모호하고 애매하다. 스트로슨은 이 주제에 관하여 「자유와 분노」 한 편을 썼을 뿐이고, 여기서 그는 주장을 할 뿐 그것을 뒷받침할 충분한 논증을 제공하지 못한다. 둘째, 스트로슨의 입장은 기존의 관행에 문제제기 하지 않는다는 난점이 있다. 스트로슨의 논증에서 반응적 태도는 기존의 사회적 관행에 기반을 둔다. 만약 그 관행에 문제가 있을 경우, 반응적 태도는 잘못된 책임 귀속이라는 귀결에서 벗어날 수 없다. 스트로슨의 이론에는 이러한 경우에 대한 적절한 답이 제시되어 있지 않다. 윌리스는 스트로슨의 한계를 보완하고자, 책임의 귀속을 제약할 도구, 즉 도덕적 해석을 가미한 이론을 전개한다.

3. 반응적 태도에 대한 규범적 해석: J. 윌리스

스트로슨 이후 반응적 태도를 중심으로 한 책임의 귀속은 여러 연구자들에 의해 시도되었다. 그 중 윌리스는 적절한 책임의 귀속이 무엇인지

www.kci.go.kr

제약하고 설명하기 위해 규범적 근거들을 도입하였고, 반응적 태도 이론을 체계적인 형태로 구성해 내는데 가장 결정적인 역할을 했다.[14] 월리스는 사람들이 관행을 바탕으로 갖게 되는 '반응적 태도'가 책임의 귀속과 수용에 대한 직접적 근거가 된다는, 그리고 특정한 상황에 처한 사람들은 책임의 부여에서 면제될 수 있다는 스트로슨의 주장에 동의한다. 월리스 역시 책임 판단의 대상이 되는 행위자에게 집중하기보다, 이러한 판단에 도덕적인 성격을 부여하는 맥락이 어떤 의미를 가졌는지 확인해야 책임 귀속의 성격을 이해할 수 있다고 말한다.[15] 그럼에도 책임 판단을 보완하고 뒷받침할 근거가 필요하기 때문에 그는 규범적 근거들을 도입해 추가적 논증을 구성한다. 그런 이유로 월리스는 자신의 입장을 규범적 해석(normative interpretation)이라고 부른다.

이름에서 알 수 있듯이, 월리스가 스트로슨과 결정적인 차이를 보이는 부분은 책임을 귀속시키는 관행을 평가하는 근거로 도덕 개념을 도입했다는 것이다. 그는 책임 있는 도덕 행위자의 조건에서 중요한 것은 행위자가 갖는 규범적 능력(normative competence)라고 말하며, 자유 의지나 대안적 행위를 할 가능성 등은 결정론 논쟁의 과정에서 조명된 지엽적인 문제로 간주한다.[16] 월리스에 따르면 우리의 반응적 태도는 행위자들이 서로를 '반성적 자기 통제력'(the power of reflective self-control)을 가진

14) 이후의 연구자들은 월리스가 구체화 시킨 형태의 반응적 태도 이론을 비판적으로 평가하며 자신의 이론을 전개해 나가는 경우가 많다. John Martin Fischer and Mark Ravizza, *Responsibility and Control*, Coleen Macnamara, "Reactive Attitudes as Communicative Entities," Gideon Rosen, "The Alethic Conception of Moral Responsibility" 참조. 다음 장의 맥케나 역시 여기 포함된다.

15) R. Jay Wallace, *Responsibility and the Moral Sentiments* (camridge: Harvard University Press, 1994). 1쪽 참조.

16) 같은 곳 참조. 월리스는 자신의 주된 관심사가 규범적 능력으로서의 책임 개념이라고 명시하지만, 그 역시 논의의 과정에서 본인이 옹호하는 책임 개념이 결정론과 양립 가능한지 검토하게 된다. 즉 스트로슨과 논증의 목적은 상이함에도, 스트로슨처럼 두 문제 모두를 논의하는 셈이다.

www.kci.go.kr

행위자로 간주하고, 이들이 적절하다고 판단한 행위를 할 것이라 기대하는 데서 기인한다. 이러한 기대가 도덕적 책무(moral oblilgation, 이하 책무)를 기반으로 한 관행의 범주에 속한다고 여겨지는 경우, 기대의 위반에 따른 반응적 감정이 책임을 귀속시킬 수 있는 힘을 갖게 된다. 월리스가 반응적 태도 이론을 정식화하기 위해 사용한 추가적 개념인 '기대', '책무', '반성적 자기 통제력'을 검토하고 월리스의 주장을 구체적으로 진단해보자.

우선 월리스는 분노와 같은 반응적 감정은 본질적으로 '기대'라는 개념과 관련되어 있는 것이지, 도덕적 믿음에서 파생된 것은 아니라고 주장한다. 만약 '분노'가 어떤 부정의나 피해가 발생했다는 생각에서 야기된 믿음의 하나라면, 이 분노는 좋음과 구분되는 옳음 개념에 호소해야 정당성을 갖출 수 있다. 그러면 그 과정에 또 다른 논증이 필요하다. 월리스는 이것이 과도한 도덕화라고 비판한다.[17] 월리스는 반응적 감정이 도덕적 믿음과 전혀 무관한 것은 아니지만, 오직 도덕적 믿음을 통해서만 발생하는 것은 아니라는 점 역시 지적한다.[18] 예를 들어 '죄책감'이라는 반응적 감정은 스스로 부정의에 기여했다는 분명한 믿음 없이도 생겨날 수 있다. 내가 행위해야 했는데 하지 못한 일, 즉 부작위는 요구나 평가가 깊이 개입되어 있어 도덕적 사실에 대한 믿음이라고 단순화하기 어렵다. 반응적 감정을 단순히 도덕적 믿음의 일종이라고 보는 것은 잘못된 분류다.

월리스는 그러므로 믿음이 아니라, '기대'라는 "유사 평가적 입장(stance of quasi evaluation)"[19]을 통해 반응적 감정을 이해해야 한다고

17) 같은 책, 18-19쪽 참조.
18) 같은 곳 참조.
19) 같은 책, 19쪽. 우리가 상대에게 기대 혹은 요구를 하는 입장은 나의 기대가 충족되는지 평가하는 입장일 수 있다는 의미에서 월리스는 기대를 유사 평가적 입장이라고 부른다.

www.kci.go.kr

주장한다. 내가 누군가에게 기대를 한다는 말의 의미는, 이러한 기대가 충족되지 않을 경우 특정 범위의 감정들에 민감해진다는 것이다. 또 내가 이 기대를 배반하는 경우 상대가 그러한 감정을 가지는 것이 적절하다고 나 역시 믿는다는 의미다. 월리스는 이 과정에서 기대 혹은 요구가 충족되지 않을 경우 민감해지는, 기대와 구성적으로 연관된 감정들이 바로 반응적 감정이며 이 감정을 통해 갖게 되는 태도가 반응적 태도라고 정의한다.[20] 기대는 도덕 혹은 반응적 감정과 무관할 수도 있다. 이제 막 미용기술을 배우기 시작한 내 친구가 머리를 잘 자르지 못할 것이라 기대하는 것은 도덕과 무관하다. 그러나 도덕적 감정을 본질적으로 불러일으키는 기대도 있다. 월리스에 따르면 이러한 기대는 "우리가 어떤 대상에게 우리가 기대하는 대로 행위해달라고 요구하는 것"[21]과 같다. 내 친구가 내게 거짓말을 하거나, 물건을 훔치거나, 협박을 하지 않을 것이라고 기대하는 것은 도덕과 관련된 기대다.[22]

그러나 도덕적이라고 분류될 수 있는 요구 혹은 기대는 그 기대를 배반할 경우 도덕적 비난이 수반될 수 있다. 행위자에게 책임과 비난을 귀속시킬 수 있는 특수한 기대는 명확한 적용 범위를 필요로 한다. 그 때문에 월리스는 도덕과 관련된 반응적 감정의 범위를 분노, 도덕적 분개, 죄책감 세 가지로 제한한다. 스트로슨은 인간관계 속에서 타인의 입장에 공감해 느끼는 감정, 즉 "대리적"[23]감정을 모두 도덕적인 범주의 반응적 감정이라고 분류했지만, 그러면 반응적 감정의 범위가 너무 폭넓게 해석

20) 같은 곳 참조.
21) 같은 책, 21쪽.
22) 월리스는 '기대' 혹은 '요구'라는 개념을 다음과 같이 설명한다. 기대는 (1)필수적인 것은 아니라 해도 실천이성에 의해 지지될 수 있어야하고, (2)기대가 언어적 형식을 가질 수 있어야 함에도 불구하고 그런 기대를 자신과 타인에게 부여하는 행위자가 언제나 그런 형식을 생산해야 한다고 요구하지는 않으며, (3)기대들이 상충하는 경우, 요구되는 행위가 상충할 수 있다.(같은 책, 22-23쪽 참조.)
23) 스트로슨, 「자유와 분노」, 164-67쪽 참조.

www.kci.go.kr

되어 개념은 모호해지고 논증의 부담은 커진다. 월리스는 스트로슨처럼 대리성의 유무를 통해 도덕적 성격을 가진 반응적 태도를 구획하기 보다는, 기대와 관련되었다고 확실히 인정되는 소수의 반응적 감정만이 책임을 귀속시킬 수 있다고 주장한다.

월리스가 반응적 태도를 분노, 분개, 죄책감만으로 제한한 이유는 다음 두 가지다.[24] 첫째, 스트로슨의 구분보다 더 명료하다. 좁은 의미의 반응적 태도를 받아들인다면 예컨대 면제조건을 만족시키는 상대에 대한 객관적 태도가 어느 범주에 들어가야 할지 분명해진다. 의사나 상담사, 복지사 등이 도덕적 책임에서 면제되었다 간주되는 환자, 관찰 대상 등과 맺는 관계 역시 인간관계다. 스트로슨은 인간관계에서 갖게 되는 태도가 모두 반응적 태도라고 말하면서도, 객관적 태도는 또 반응적 태도가 아니라고 한다. 월리스는 자신의 정의를 채택한다면 스트로슨의 주장 내부의 이런 애매성을 해결하고, 객관적 태도를 가지는 사람들과 그 대상 사이의 인간관계를 반응적 감정과 명확하게 구별해 낼 수 있다고 주장한다. 즉 기대라는 감정을 통해 반응적 태도와 그 밖의 감정들을 분명히 구분할 수 있다.

둘째, 인간이 사회 속에 살면 자연스럽게 반응적 태도를 갖는다는 스트로슨의 주장은 너무 강하고 선언적이기 때문에, 반응적 태도 이론은 논파하기 쉬운 형태가 된다. 월리스는 반응적 감정의 경계를 명료화함으로서 그 적용범위를 분명히 하는 방식으로 이 문제를 해결하려 한다. 반응적 태도의 범위를 세 개의 감정으로 좁히고 특정한 문화권 안의 태도로 국한시킨다고 해도 반응적 감정을 느끼는 과정이 자연스럽다는 것은 부정하기 어렵다.[25] 자연적 감정과 반응적 감정을 구분하면서도 반응적

24) Wallace, *Responsibility and the Moral Sentiments*, 29-32쪽 참조.

25) 앞서 말했듯이 스트로슨은 인간이라면 당연히 느끼는 감정이 반응적 감정이라고 주장했지만, 월리스는 반응적 태도가 없는 가상의 문화 역시 존재하지 못할 이유는 없다고 말한다. 분노, 분개, 죄책감이 없는 문화는 당연히 드물겠지만, 감정을

www.kci.go.kr

감정은 논증 가능하다.

그래서 월리스가 기대와 관련해 책임을 귀속시킬 수 있는 숙고된 감정이라고 분류한 것은 분노, 분개, 죄책감뿐이다. 이 감정들만이 자신과 타인에게 부여하는 평가, 즉 도덕적 기대와 본질적으로 엮여있는 감정이기 때문이다. 예를 들어 '감사'하는 태도는 스트로슨에게 반응적 태도지만 월리스에게는 아니다. '감사'는 내가 상대에게 기대하거나 요구한 것이 아니기 때문이다.[26] 따라서 감사라는 감정은 상대에게 책임을 귀속시키지 못한다. 반면 반응적 태도로서의 '분노'는 그냥 화가 난 것이 아니라 타인의 행동에 대해 설명 혹은 제재를 요구하는 태도다. 분노가 적절하다면 상대에게는 그 행위에 대한 책임이 귀속되고, 상대의 행위에는 비난 혹은 제재가 뒤따를 수 있다. 책임의 귀속과 비난은 타인의 행위에 대한 중대한 개입이고, 그것의 정당화에는 합당한 근거가 필요하다.

그렇다면 무엇이 적절한 기대일까? 월리스에게 책임을 부과할 수 있는 적절한 기대는 도덕적으로 정당화 된, 나와 상대 모두가 받아들인 기대다. 그는 이것을 '책무'라고 부른다.[27] 반응적 태도의 적절성은 그 태도

절제하는 배경 때문에 반응적 태도가 없는 곳이 존재한다 할지라도 지금 우리 사회가 책임을 귀속시키는 관행을 가지고 있다면 반응적 태도 이론은 충분히 힘이 있다는 것이다. 같은 곳 참조.

26) 모든 반응적 태도 이론 연구자들이 이처럼 부정적인 반응적 태도만 고려하는 것은 아니다. 분노, 분개 등의 부정적인 반응적 태도와 같은 비중으로, 감사함, 찬사 등의 긍정적인 태도 역시 설명할 수 있어야 한다고 주장하는 연구자도 있다. Elisa A. Hurley and Coleen Macnamara, "Beyond Belief: Toward a Theory of the Reactive Attitudes," *Philosophical Papers*, Vol. 39, no. 3 (2010), 373-399쪽 참조.

27) Wallace, *Responsibility and the Moral Sentiments*, 36쪽, 62-63쪽 참조. 월리스는 '책무'를 설명하며 앤스콤이나 윌리엄스를 인용하고, 이 개념을 다소 고전적인 의미에서 사용하겠다고 명시한다. 이 말은 롤즈적 의미에서의 책무, 즉 사회 제도의 기반이 되는 정의의 원칙이 부여하는 의무라는 의미에서 책무의 개념을 정의하지 않겠다는 것이다. 월리스의 표현에 따르면 책무란 "옳음의 요구 전체를 가리키는 말"이다. 롤스적 의미에서의 책무에 '자연적 의무'라는 의미를 더한 것이 월리스의 논증에서 의미하는 책무의 개념과 가깝다.(같은 책, 64쪽, 각주 18

www.kci.go.kr

를 낳은 기대의 내용에 따라 달라진다. 도덕적 근거들의 측면에서 정당화 불가능한 기대는 사람들에게 책임을 부여할 수 없다. 또 월리스는 양자 모두가 받아들인 책무에 대한 기대만이 도덕적 책임을 지울 수 있다고 말한다.[28] 나와 상대방 모두에게 정당화된 책무의 위반에서 분노가 기인했다면, 이 반응적 태도는 상대에게 책임을 귀속시킬 수 있다. 따라서 책무는 도덕적 반응적 태도와 그렇지 않은 것을 구분하는 기준이며, 적절한 기대를 정당화하는 근거다.

책임을 지우는 방식을 떠올려 보자. 모든 '분노'가 도덕적 책임을 타인에게 귀속시킬 수 있는 것은 아니므로, 나 자신부터 나의 분노가 적절한지 묻는다. 정당화되기 어려워 보이거나, 상대가 동의하지 않을만한 기대에서 솟아난 감정이라면, 이 분노는 상대에게 책임을 지울 수 없다. 반응적 태도를 정당화 하려면 이처럼 그 설명 과정에 도덕적 책무가 등장하지 않을 수 없다는 것이 월리스의 주장이다. 이렇게 도덕적 정당화를 통해 적절한 분노, 도덕적 분개, 죄책감이 가려진다.[29]

이렇게 정당화된 반응적 태도는 어떤 행위자가 가질 수 있는 것일까? 월리스의 체계에서 행위자는 결정론이 참이라고 해도 당위를 담보할만한 행위능력을 소유할 수 있는가? 월리스 역시 당위가 가능을 함축해야 한다는 기본 명제를 받아들인다. 그러나 월리스의 체계에서 책임을 질 수 있는 행위자란 강한 의미에서의 자유, 즉 대안적 가능성을 늘 가지고 있는 존재가 아니라, 일반적인 수준의 합리적 능력을 가지고 있는 행위자다.[30] 이러한 합리적 능력을 월리스는 '반성적 자기 통제력'이라고 부른다. 반성적 자기 통제력은 "(1) 도덕적 이유를 이해하고 적용할 수 있는 힘, 그리고 (2) 이러한 이유들의 측면에서 자신의 행위를 통제하고 조절

번 참조.)

28) 같은 책, 63쪽 참조.

29) 같은 책, 36쪽 참조.

30) 같은 책, 7-8쪽과 13-14쪽 참조.

www.kci.go.kr

할 수 있는 힘"[31]이라고 정의된다. 이러한 능력은 강한 의미의 자유의지를 요청하는 것이 아니라, 사회를 구성하는 성인에게 기대할만한 사고력을 소유하는 것이다. 운명이 정해져 있거나 대안가능성이 없다 할지라도, 내가 이해력과 합리적 사고력을 바탕으로 한 행위능력을 가진다는 사실은 사라지지 않는다. 그렇다면 결정론이 참이라고 해도 책임 개념의 존립은 위협받지 않는다.

또한 월리스는 반성적 자기 통제력의 유무를 통해 책임 귀속의 변명조건과 면제조건을 설명한다. 책임을 귀속시킬 수 있는 동시에 책임을 질 수 있는 행위자는 반성적 자기 통제력을 가진다. 그러나 변명조건과 면제조건을 만족시키는 경우, 즉 불가피한 사정이 있었거나 반성적 자기 통제력을 본래 가지지 못한 경우 우리는 행위자에게 책임을 귀속시키지 않는다. 행위자의 행위와 그 행위자가 가졌던 의지의 질이 악의적이지 않았다는 것을 충분히 보여준다면 행위자는 비난 혹은 책임 귀속에서 자유로울 수 있다.[32] 강도의 강요로 돈을 내어준 은행원은 변명조건을 예시할 수 있는 대표적인 경우다. 반면 면제조건은 상당기간동안 반성적 자기 통제력이 결여된 상태에 적용된다. 어린 아이 이거나 심각한 정신적·신체적 장애를 가진 사람의 경우, 행위자는 반성적 자기 통제력을 가지기 어려워진다. 이 경우에는 전자와 달리 책임을 질 수도 또 지울 수도 없다.[33]

월리스의 주장을 요약하자면, 반응적 태도는 '반성적 자기 통제력'을 가진 동료 행위자에 대한 '기대'에서 비롯한다. 이 기대가 자의적인 변덕이나 계산이 아니라 '책무'일 때 책임의 귀속은 도덕적 성격을 가진다. 그래서 어떤 행위자가 도덕적 이유들을 포착하고 적용할 수 있을 때, 또 그 이유에 따라 자신의 행위를 통제할 수 있는 능력을 가진 경우, 적절

31) 같은 책, 157쪽.
32) 같은 책, 128쪽 참조.
33) 같은 책, 155-66쪽 참조.

www.kci.go.kr

한 기대에 따른 우리의 반응적 태도는 그로 하여금 책임의 귀속을 가능하게 한다. 결과적으로 월리스는 스트로슨이 제시한 이론적 이상을 규범적으로 해석함으로서 반응적 태도 이론을 명료화 한다. 그는 스트로슨의 이론에서 모호하게 해석된 반응적 태도를 제약하며, 적절한 기대가 무엇인지 결정해 주는 관행을 평가할 도구로 공정한 책무를 제안한다.

그러나 월리스는 외부의 도덕개념을 통해 반응적 태도의 적절성을 판단함으로서 반응적 태도와 관행의 관계를 약화시킨다. 이렇게 관행과 분리된 도덕 개념들을 근거로 삼으면 반응적 감정이 가지는 불확실성은 줄어들지만, 동시에 책임 개념이 책무 개념에 종속적인 형태가 된다. 책무에 지나친 무게가 실리는 경우, 관행이 책임을 부여한다는 스트로슨의 핵심적인 주장이 무력해지고, 반응적 태도를 통해 책임을 귀속시켜야 한다는 요구는 불필요해진다. 다시 말해서 책임을 귀속시키는 것은 관행과 반응적 태도가 아니라 책무라고 해도 무관해 진다는 것이다. 그러면 우리는 개별적 행위자들이 왜 서로에게 관행에 따른 기대를 하고 그에 따른 반응적 감정을 표현하며 책임을 귀속시키는지에 관한 추가적 설명을 해야만 한다. 필자는 이러한 문제를 맥케나의 '대화 이론'을 통해 보완할 수 있을 것이라고 생각한다. 대화 이론은 관행과 책임 귀속의 과정을 대화와 유비하는 방식으로 논증하기 때문에, 관행이 책임 귀속의 배경으로 작동하는 방식에 관한 적절한 설명을 제공한다.

4. 책임 귀속 관행과 대화적 접근법: M. 맥케나

반응적 태도 이론에서는 반응적 태도가 책임의 귀속에 핵심적인 역할을 하고, 이러한 책임의 귀속에는 비난이 수반될 수 있으며, 그 과정에서 행위자가 가진 의지의 질이 중요한 요소로 고려된다. 또한 이러한 책임의 귀속은 우리의 관행을 배경으로 이루어진다. 이러한 요소들은 스트로

www.kci.go.kr

슨과 이후의 연구자들이 모두 동의하는 부분이다. 그러나 월리스의 규범적 해석은 관행이 갖는 중요성을 간과하고, 그로 인해 책임의 귀속을 규범적 개념들에 종속시키는 경향이 있다.

맥케나 역시 월리스가 책임의 귀속에 관한 서사를 재구성하며 스트로슨의 '상호관계적 관행'(interpersonal practice)이 가지는 중요성을 간과했다고 비판한다.[34] 맥케나에 따르면 반응적 태도를 취하도록 하는 원인이자 배경이 바로 인간이 서로 관계 맺는 방식에 관련된 우리의 관행인데, 이러한 부분이 월리스의 책임 귀속 논증에서는 등한시 된다는 것이다.[35]

맥케나는 기본적으로 스트로슨의 논증에서는 반응적 감정의 범위가 너무 모호하다는 월리스의 주장을 받아들인다. 도덕적인 반응적 감정과 도덕과 무관한 반응적 감정을 대리성 유무로 구별하는 스트로슨의 설명에 문제가 있다는 지점에 동의하기 때문이다.[36] 그래서 맥케나는 반응적 태도를 분노, 분개, 죄책감만으로 제한하는 월리스의 정의를 받아들인다. 그러나 월리스와 달리, 맥케나가 반응적 태도를 중심으로 한 책임 개념 논증에서 가장 중요하게 생각하는 부분은 '관행'의 역할이다. 맥케나는 스트로슨의 책임 개념에서 책임 있는 행위자에 대한 정의가 곧 책임의 귀속에 대한 설명이 된다는 점과, 책임의 귀속에 대한 설명에 반응적 태도와 책임 귀속의 관행이 고려되어 있다는 점을 핵심적인 부분이라고 간주한다.[37]

맥케나와 월리스 사이의 차이는 '관행'이 규범적 개념들에 종속되는 방식에 있다. 맥케나는 책임의 귀속 과정에서 관행이 직접 역할 해야 한다고 본 반면, 월리스는 책임의 주요 속성에 관한 논증에서 관행을 특별

34) McKenna, *Conversation and Responsibility* (New York: Oxford University Press, 2012), 34-39쪽 참조.
35) Warmke and McKenna, "Moral Responsibility, Forgiveness, and Conversation," 191쪽 참조.
36) McKenna, *Conversation and Responsibility*, 65-66쪽 참조.
37) 같은 책, 2-3쪽 참조.

www.kci.go.kr

히 언급하지 않는다. 월리스가 한 줄로 정리한 책임의 조건은 다음과 같다.

> (N) S는 오직 S에게 (행위x에 대한) 도덕적 책임을 지우는 것이 적절할 때만 (행위x에 대한) 도덕적 책임이 있다.[38]

월리스의 정식은 적절함의 내용이 무엇인지에만 주목하기 때문에, 그 내용이 어떻게 채워지는지에 관한 물음은 결론에 이르는 과정에만 자리한다. 월리스의 체계 속에서 책임의 귀속이 이러한 정식화를 통해 일반화되기 때문에, 결국 반응적 태도의 배경으로서의 관행은 직접적 역할을 잃게 된다. 그러면 책임은 공정함과 책무의 제한을 받는 하위규범으로 간주될 수 있다.

반면 맥케나에 따르면 문제 상황에 대한 반응적 태도의 적절성은 책무가 아니라 사람들이 이전에 맺고 있던 관계와 그 관계가 바탕하고 있던 관행에 따라 결정된다.[39] 책무 등의 규범적 근거는 우리의 관행 자체의 옳고 그름을 평가하는 기준으로는 유효하다.[40] 하지만 책임의 귀속 그 자체는 책무가 아니라 우리의 관행에 따른 반응적 태도에 따른다고 보는 것이 적절하다. 책임의 귀속이 하나의 명제 속으로 포섭되는 종류의 것이라고 보는 것 보다, 관행 속에서 유동적으로 변화하는 일종의 맥락으로 보는 것이 책임 개념과 우리의 도덕적 능력이 작동하는 방식을 더 잘 설명한다. 그래서 맥케나는 대화와의 유비를 통해 책임의 귀속 방식을 설명하고자 한다.[41]

맥케나는 우리가 가지는 도덕적 능력을 모국어를 사용하는 화자의 능

38) Wallace, *Responsibility and the Moral Sentiments*, 91쪽.
39) McKenna, *Conversation and Responsibility*, 3-4쪽 참조.
40) 같은 책, 34-35쪽 참조.
41) 맥케나는 자신의 대화적 접근법이 왓슨에게서 왔음을 명시한다. 같은 곳 참조.

www.kci.go.kr

력에 유비한다.[42] 우리는 모국어를 사용하는 환경에서 나고 자란다. 이 언어는 몸에 심어진 것이라고 해도 무방할 만큼 나 자신과 분리불가능하며, 모국어 문법은 그것이 어떤 원리인지 깨닫기 어려울 정도로 체화되어 있다. 누군가 언어의 관습에 어긋나게 말하면 대화에 참여한 사람들은 계속 오해할 뿐 대화를 이해할 수 없으므로, 우리는 언어적 관습을 충실히 따른다. 그래서 풍자나 비꼬기 같은 섬세한 상황에서도 적절한 단어와 표현을 사용할 수 있으며, 이러한 말들을 즉각적으로 이해하고 반응한다. 모국어를 구사하는 능력이 있다는 것의 의미는 이 화자가 자신을 표현할 수 있는 능력과 타인의 말을 이해할 수 있는 능력 두 가지를 다 가지고 있다는 것이다. 이 두 가지 능력은 상호보완적이며, 모두 가지고 있는 경우에만 대화가 가능하다.

책임의 담지자라 간주되는 행위자의 행위도 이 대화와 유사하다. 행위자는 관행 안에서 적절하게 행동해야 하고, 이런 관행에서 어긋나는 행위는 서로에게 오해를 불러일으킨다. 모국어를 구사하는 것처럼, 행위자의 행위 역시 자연스럽게 일어난다. 우리가 상대가 악의를 가지고 있지 않다는 것을, 즉 상대의 의지의 질을 쉽게 알아볼 수 있는 것도 우리가 이러한 행위의 관행 속에 있기 때문이다. 우리가 상대의 행위를 쉽게 해석하고 이해할 수 있는 이유는, 내가 그렇게 행위할 수 있기 때문이다. 우리는 때로 자신이 악의가 없음을 적극적으로 드러내는 행위를 하기도 한다. 맥케나는 환자들을 체크해서 위급한 순서대로 치료실로 보내야 하는 간호사가, 일부러 오가며 환부를 확인하고, 고민하는 표정을 짓고, 이 편 저 편의 병상을 둘러보며 환자들을 지켜보고 있다는 표를 내는 것을

42) 같은 책, 85-86쪽 참조. 이런 이유로 맥케나는 자신의 이론을 '대화 이론'(conversation theory) 혹은 '대화적 접근법'(conversational approach)라고 부르며, 이러한 대화 이론을 주장하는 다른 연구자들로는 맥나마라, 스테판 다웰(Stephen Darwall), 미란다 플리커(Miranda Fricker) 등이 있다. Matthew Talbert, "Moral Responsibility" 참조.

www.kci.go.kr

예로 든다. 맥케나는 타인과 우리의 도덕 관계는 우리의 행위가 상대에게 어떻게 해석될지 파악하며 만들어 진다고 말한다.[43] 타인이 나의 의도를 이해할 수 있게 나의 행위를 조정하는 것은 공동체 속의 행위자들이 숨 쉬듯이 자연스럽게 늘 하고 있는 일이다.

공동체 안의 행위자들이 행위하다 보면 때로 문제적인 상황이 발생한다. 행위자가 분노, 분개, 죄책감이라는 반응적 태도를 드러내는 것이 적절한 상황이라면, 상대 행위자는 책임 질 행위를 한 행위자에게 책임을 귀속시킨다. 맥케나에 따르면 우리가 책임을 귀속시키는 관행은 도덕적 공헌(moral contribution), 도덕적 지목(moral address), 도덕적 논증(moral account)이라는 세 단계를 거친다.[44] 그리고 이 세 단계는 대화를 나누는 것과 유사한 방식으로 이루어진다.

도덕적 공헌은 어떤 행위자가 자신의 의지가 어떤 성질을 가지는지 드러내는 행위를 하는 것이다. 예를 들어 A는 B에게 소수자에 대한 편견을 강화하는 농담을 한다. 최초의 행위가 이루어지는 것이 이 첫 번째 단계다. 두 번째 단계인 도덕적 지목은 관련된 다른 행위자들, 즉 이대로 두지 않아야 할 책임이 있는 행위자들이 적절한 반응적 태도를 보이는 것이다. 분노, 분개 등이 이런 반응적 태도의 예이다. 우리의 예에서 B는 A에게 분개의 태도를 보일 것이다. 마지막 도덕적 논증은 한 쪽이 다른 쪽에게 자신의 행위에 대한 설명을 내놓는 단계다. 행위자는 사과를 할 수도, 거부할 수도, 어쩌면 단지 자신의 잘못을 인지하기만 할 수도 있다. A는 B에게 그런 행위에 대한 문제를 인식했음을 말하고 사과를 함으로서 용서를 구할 수 있다. 하지만 대화를 거부하며 관계를 끊는 것도 가능하다. 도덕적 논증의 단계에서 관계는 언제나 재조정 될 수 있다.

책임 부여의 이러한 세 단계가 스트로슨이 중요시했던 세 주요 요소를 반영하고 있다고 맥케나는 주장한다.[45] 첫 번째 단계는 책임이 있는 행

43) 같은 책, 86-87쪽 참조.
44) 같은 책, 89-90쪽 참조.

www.kci.go.kr

위자의 의지의 질을 보여준다. 이 단계에서 보이는 행위에 행위자의 악의적 의도가 없었다는 것을 설명할 수 있다면 추후 그의 변명은 수용될 것이다. 두 번째 단계에서는 책임 지우는 행위자, 즉 책임을 귀속시키는 행위자의 반응적 태도가 드러난다. 세 번째 단계에는 책임 질 행위자가 내놓는 변명 혹은 또 다른 논증이 자리한다. 책임을 지우고 책임을 지는 쌍방의 행위자들 사이에 오가는 논증은 책임 귀속 관행의 가장 중요한 요소다. 세 번째 단계에서 도덕적 논증은 자신이 가졌던 의지의 질에 관한 해명으로 작용할 수도 있지만, 그 질의 표현 방식에 부적절함이 있었음을 인식하는 계기가 될 수도 있다. 언어가 변화하는 것처럼 관행의 도덕 문법도 늘 변화하기 때문에, 그러한 변화를 학습하는 것도 가능하다.

맥케나의 입장은 월리스의 입장과 상충하기보다 다소 보완적이라고 할 수 있다. 맥케나는 규범적 근거를 도입함으로서 우리의 관행 자체에 문제제기 하는 길을 열어야 한다는 월리스의 방식에 동의한다.[46] 관행 자체에 대한 평가는 규범적인 근거를 통해 이루어 질 수밖에 없다. 맥케나의 대화적 접근법은 월리스가 누락한 책임의 귀속방식에 대한 설명을 보완한다. 대화적 접근법은 다음 세 가지 방식으로 책임 귀속에 대한 이해를 증진시킨다. 첫째, 모국어 화자의 대화와 유비함으로서 우리의 관행이 우리의 책임 귀속에 얼마나 전면적이고 근본적인 영향을 주고 있는지 이해할 수 있다. 둘째, 책임의 귀속과 그 책임을 지는 방식이 얼마나 다양하고 맥락 의존적인지 역시 이해할 수 있다. 마지막으로 맥케나의 논증은 책임을 귀속시키는 '대화'에 참여할 능력이 없다는 것이, 즉 반성적 자기 통제력을 가지지 못하는 경우에 어떻게 그들을 책임의 귀속에서 면제시켜주는지 더 정확하고 직관적으로 설명한다.[47]

맥케나는 월리스의 규범적 해석을 보완해, 책임의 귀속 과정이 관행에

45) 같은 책, 91-92쪽 참조.
46) 같은 책, 34-35쪽 참조.
47) 마지막 요소에 대해서 맥케나 본인도 언급하고 있다. 같은 책, 79-80쪽 참조.

www.kci.go.kr

근거한다는 것을 확인해준다. 대화적 접근법을 통해 관행이 우리의 모국어 사용처럼 근본적인 방식으로 책임 귀속 과정 전체를 이루고, 이를 통해 책임의 귀속과 그를 근거로 타인의 행위에 의견을 개진하는 것이 얼마나 자연스럽게 이루어지고 있는지 설명할 수 있다.

5. 결론

이 논문의 목적은 반응적 태도 이론을 통해 책임 개념이 안정적으로 설명될 수 있음을 주장하는 것이다. 필자는 반응적 태도가 책임 귀속의 근거가 된다는 스트로슨의 주장에서 시작해, 그러한 주장을 규범적 개념들을 통해 제약하고 명료화 한 월리스와 관행이 어떻게 반응적 태도로 하여금 책임을 귀속시키게 하는지 설명한 맥케나의 논증을 두루 살펴보았다. 이를 통해 반응적 태도 이론이 반성적 자기 통제력을 가진 공동체 안의 행위자에게 책임을 귀속시키는 과정을 검토했다. 반응적 태도는 책무에 기반한 적절한 기대에 따라 우리가 평등한 동료 행위자들에게 보이는 태도이며, 이것은 나와 타인에게 모국어 화자가 대화에 참여하듯이 자연스러운 방식으로 책임을 귀속시킨다.

현대의 철학은 기존의 도덕개념들로 해결되지 않는 문제들에 대한 답을 찾고 있다. 책임은 그러한 과정에서 다양한 방식으로 재해석 되고 있는 중이다. 기존의 책임 개념은 자율성, 상호성, 과거지향성의 성격을 고수했지만, 현대의 문제들은 그런 제약에서 벗어나야 해결할 수 있다.[48] 반응적 태도 이론은 책임 개념과 그 귀속에 관한 가장 설득력 있는 주장 중 하나다. 반응적 태도 이론은 자율성을 논증해야 할 부담에서 벗어났고, 상호적이지 않은 관계에도 적용된다는 점에서 앞의 두 문제에 대한

48) 김양현 외, 「책임 개념에 대한 실천윤리적 해명」, 『범한철학』, 39집 (2005년 봄), 304-10쪽 참조.

www.kci.go.kr

답이 될 수 있다. 미래지향적인 책임은 과거의 문제와 다른 방식으로 물어져야 할 것이라고 필자는 생각한다.

반응적 태도 이론의 또 다른 가능성은 인간의 존재론적 지위에 구속받지 않는 책임의 주체를 상상할 수 있게 해준다는 것이다. 그 근거는 행위자가 반성적 자기 통제력을 가진 우리 공동체의 일원이라면 누구라도 행위자로 역할 할 수 있다는 데 있다. 이는 다양한 이성적 존재자들에게 열린 정의이며, 인간 행위자를 넘어서 비 인간 행위자, 즉 기업, 국가, 단체 등의 집단 행위자에까지 책임 부여의 대상을 확장하는 단초가 될 수 있다. 이러한 확장이 성공적이라면, 대규모의 집단적 행위자에게서 기인하는 환경, 전쟁, 착취 등의 문제에 대한 해결책 모색의 시작점이 될 수 있을 것이다.

www.kci.go.kr

참고문헌

김양현 외,「책임 개념에 대한 실천윤리적 해명」,『범한철학』, 39집 (2005), 291-328쪽.

양선이,「중독과 도덕적 책임」,『철학연구』, 109집 (2015), 191-216쪽.

피터 스트로슨,「자유와 분노」,『자유의지와 결정론의 철학적 논쟁』 최용철 편, 최용철 역 (서울: 간디서원, 2004) 141-186쪽.

Fischer, John Martin and Ravizza, Mark. *Responsibility and Control* (Cambridge: Cambridge University Press, 1998).

Kane, Robert. *A Contemporary Introduction to Free Will* (New York: Oxford University Press, 2005).

Macnamara, Coleen and Hurley, Elisa A. "Beyond Belief: Toward a Theory of the Reactive Attitudes," *Philosophical Papers,* Vol. 39, no. 3 (2010), 373-399쪽.

Macnamara, Coleen. "Reactive Attitudes as Communicative Entities," *Philosophy and Phenomenological Research*, Vol. 90, No. 3 (2015), 546-69쪽.

McKenna, Michael. "The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism," in *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's "Freedom and Resentment"* (Farnham: Ashgate Publishing, 2008), ed. Michael McKenna and Paul Russell, 201-18쪽.

_____, *Conversation and Responsibility* (New York: Oxford University Press, 2012).

McKenna, Michael and Warmke, Brandon. "Moral Responsibility, Forgiveness, and Conversation" in *Free Will and Moral Responsibility* (Newcastle: Cambridge Scholars Publishing, 2013), ed. Ishtiyaque

www.kci.go.kr

Haji and Justin Caouette. 189-212쪽.

Rosen, Gideon. “The Alethic Conception of Moral Responsibility,” in *The Nature of Moral Responsibility New Essays* (New York: Oxford University Press, 2015), ed. Randolph Clarke, Michael McKenna, and Angela M. Smith, 65-87쪽.

Talbert, Matthew. “Moral Responsibility,” *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/win2019/entries/moral-responsibility/>.

Tollefsen, Deborah. *Group as Agents* (Cambridge: Polity Press, 2015).

Wallace, R. Jay. *Responsibility and the Moral Sentiments* (camridge: Harvard University Press, 1994).

Watson, Gary. “Responsibility and the Limits of Evil: Variations on a Strawsonian Theme,” in *Responsibility Character and the Emotions* (Cambridge: Cambridge University Press, 1987), ed. Ferdinand Schoeman, 256-86쪽.

www.kci.go.kr

Reactive Attitudes and Responsibility

Park, Ey yeon (Chonnam National Univ.)

The goal of this paper is to show that the concept of moral responsibility can be well explained through the Reactive Attitude Theory. In the Reactive Attitude Theory, reactive attitudes of fellow agents are the basis for attribution of responsibility. The appropriateness of the reactive attitude depends on interpersonal relations, that is, practice, and the expectations that we have for each other. If this theory is valid, responsibility is attributed to the agent, even though determinism is true.

I start with P. F. Strawson's argument that reactive attitudes are the basis for attribution of responsibilities, and I look through the Normative Interpretation by J. Wallace that constrains and clarifies reactive attitudes and practices through normative concepts. And then, I want to show how practices let reactive attitudes hold people responsible in M. McKenna's Conversation Theory. Strawson was the first philosopher to attempt to attribute responsibility through a reactive attitude, but he's account is criticized for not only its ambiguity and vagueness but also unsolved problems of harmful practices. To solve these problems, Wallace organized the theory of reactive attitudes in a systematic form with normative concepts. However, he overlooks the role of practices that play a pivotal role in holding agents responsible. I argue that these problems can be complemented by McKenna's Conversation theory. Reactive Attitude Theory is a sound and original way of understanding responsibility and could provide a more extended

www.kci.go.kr

understanding of the agency as a subject of responsibility.

Key words: reactive attitude theory, responsibility, Wallace, McKenna, Strawson

박의연 e-mail: whale1124@naver.com

투 고 일	2020년 01월 15일
심 사 일	2020년 02월 18일
게재확정	2020년 02월 18일

www.kci.go.kr