creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

교육학박사학위논문

사회적 직관주의 이론의 비판적 해석 및 도덕 교육적 함의 연구

2019년 2월

서울대학교 대학원
윤리교육과
양 해 성

국문초록

사회적 직관주의 이론의 비판적 해석 및 도덕 교육적 함의 연구

양 해 성

도덕적 판단의 기제를 사실적으로 규명하려한 사회적 직관주의 이론 즉, 그린(J. Greene)의 이중 과정 모델(Dual Process Model)과 하이트(J. Haidt)의 사회적 직관주의 모델(Social Intuitionist Model)이 타당성을 담보하기 위해서는, 두 이론이 도덕적 판단의 범주에 속하는 다양한 유형의 현상들을 포괄할 수 있어야 하며, 도덕적 판단의 기제에 관한 일반적이고 통일된 설명을 제공할 수 있어야 한다. 만일, 두 사회적 직관주의 이론이 실제 규명한 현상보다 더 넓은 범주의 현상을 포괄하는 것으로 확대 해석되고 있거나, 이론 내부에 모순된 해석의 위험성이 내재한다면, 두 이론은 타당성을 담보하기에 어려움이 있다. 그렇기에 두 사회적 직관주의 이론의 타당성을 담보하기 위해서는 첫째, 이론이 확대 해석되지 않도록, 이론이 규명한 도덕적 판단의 현상 범주를 구체화하는 작업 즉, 규명 범주(verification category) 구체화 작업을 수행하여야 하며 둘째, 이론에 내재할 수 있는 해석상 모순의 위험성을 경계하기 위해, 이론이 활용한 개념 범주(conceptual category)가 연구자의 의도와 다르게 해석될 여지가 없는지 검토하는 작업 즉, 개념 범주 명료화 작업을 수행하여야 한다.

그러나 두 사회적 직관주의 이론에 관한 논의들은 이러한 규명 범주

구체화 및 개념 범주 명료화 작업이 충분히 선행되지 않은 상태에서 도출된 측면이 있다. 이와 같이 규명 범주 구체화 작업이 수행되지 않은 상태로, 하이트의 사회적 직관주의 모델은 도덕적 판단의 현상 범주 중 일부 범주만을 포괄하는 편향된 이론이라는 비판을 받았으며, 그린의 이중 과정 모델은 더욱 넓은 범주의 도덕적 판단 현상을 포괄하는 타당한 이론으로 지지되었다. 또한 개념 범주 명료화 작업이 수행되지 않은 상태로, 의무론적 판단은 감정적인 직관적 도덕 판단과 이에 대한 합리화에 불과하다는 비판이 이루어졌고, 이로 인해 의무론의 규범적 지위가 의심을 받게 되었다. 또한 두 사회적 직관주의 이론에 관한 논의들이 도덕적 판단의 현상을 대상으로 인지 심리학의 발견을 재확인 하는 차원에서 수행됨에 따라, 사회적 직관주의 이론 고유의 발견이 충분히 도출되지 않았으며 이를 교육에 적용하기 위한 방안도 충분히 모색되지 않았다.

본 연구는 두 사회적 직관주의 이론의 규명 범주 구체화 작업 및 개념 범주 명료화 작업이 선행되지 않을 경우, 관련 논의의 타당성을 담보할 수 없다는 문제의식을 가지고 두 이론의 규명 범주 구체화 및 개념 범주 명료화 작업을 수행하였다. 그리고 이에 기반하여 두 이론에서 파생된 논의의 타당성을 검토하였으며, 인지 심리학과 구별되는 사회적 직관주의 이론 고유의 발견을 도출할 수 있는 가능성을 모색하였다. 또한 이에 기반하여 직관적 인지의 기제를 반영한 대표적인 도덕교육 이론인 나바에츠(D. Narvaez)의 통합적 윤리교육 모델(Integrative Ethical Education)에 사회적 직관주의 이론 고유의 발견을 적용할 수 있는 방안을 모색하였다. 이러한 작업들을 통해 규명한 세부 사항은 다음과 같다.

첫째, 규명 범주 구체화 작업에 기반할 때, 두 사회적 직관주의 이론은 상충하는 관계가 아니며, 오히려 보완적으로 기능한다. 두 이론에서 주장의 근거로 제시하는 인터뷰 문항을 비교할 때, 하이트의 사회적 직관주의 모델의 핵심 주장은 '피해와 공정성의 기준이 적용될 수 없는

상황’을 규명 범주로 하여 도출되었으며, 그린의 이중과정 모델의 핵심 주장은 ‘피해와 공정성의 기준이 적용되는 상황’을 규명 범주로 하여 도출되었다. 즉, 두 이론은 도덕적 판단이라는 현상을 구성하는 상이한 하위 범주의 현상을 각각 규명한 것으로서, 상호 보완적인 관계로 파악하는 것이 타당하며, 두 이론 중 어떠한 이론이 더욱 타당한지를 비교하는 논의는 성립하기 어렵다.

둘째, 직관과 추론이라는 개념 범주를 차용하고 감정적 활성화 정도를 그 구분 기준으로 삼을 경우, 두 이론 내부에 모순된 해석의 위험성이 발생한다. 그린은 감정적 활성화 정도를 기준으로 도덕과 무관한 딜레마(Non-Moral Dilemma)에서의 판단을 추론적 판단으로 분류하였으나, 직관과 추론을 구분하는 인지 심리학의 다양한 기준들을 고려할 때, 도덕과 무관한 딜레마 상황에서의 판단을 직관적 판단과 추론적 판단 중 한 종류의 판단으로 명확하게 분류하기에 어려움이 있다. 즉, 감정의 활성화 정도만을 가지고 직관과 추론을 구분하는 것은 모순된 해석의 위험성을 내포한다.

또한, 딜레마 유형별 응답 소요 시간 그래프를 기준으로 할 때, 그린이 추론적 판단으로 분류한 비인신적 딜레마(Moral Impersonal Dilemma) 상황에서의 긍정적 응답 도출에 소요된 시간이 그가 직관적 판단으로 분류한 인신적 딜레마(Moral Personal Dilemma) 상황에서의 부정적 응답 도출에 소요된 시간보다 짧다는 점도, 감정의 활성화 정도만을 가지고 특정한 판단을 직관 또는 추론으로 분류하는 것이 모순된 해석의 위험성을 내포하는 성급한 분류라는 것을 보여준다. 수단을 제외하고 동일한 조건으로 구성된 두 상황을 전제로 도출된 이러한 실험 결과는, 직관적 판단이 신속한 판단으로서의 특성을 갖고 추론적 판단이 느린 판단으로서의 특성을 갖는다는 통설과 정면으로 충돌한다. 그린이 추론으로 분류한 비인신적 딜레마 상황에서의 판단은 덜 감정적이고 더 의식적이며, 판단의 과정을 파악할 수 있다는 추론적 판단의 특성들을 가지지만, 동시에 신속하다는 직관적 판단의 특성도 갖는다. 그렇기에

이는 추론이라 단정하기에도 어려움이 있으며 직관이라 단정하기에도 어려움이 있다.

하이트의 사회적 직관주의 모델에 있어서도 이러한 모순된 해석의 위험성이 존재한다. 하이트가 추론으로 분류한 인지 또한 상기한 바와 같이, 숙달 정도와 자동성 등에 따라 직관으로도 분류할 수 있고 추론으로도 분류될 수 있다. 그렇기에 모순된 해석의 위험성을 제거하고 두 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 보다 타당하게 해석하기 위하여, 개념 범주의 보완 또는 대체가 필요하다. 또한 이러한 보완적 또는 대체적 개념 범주는 하이트와 그린이 직관 및 추론으로 범주화한 대상이 어떠한 실제를 지칭하는지 명료화 하는 작업을 통해 도출할 수 있다.

하이트와 그린이 직관으로 분류한 개념 범주는, 나바에츠(D. Narvaez)가 내용적 차원의 제한 없이 수월성과 자동성이라는 형식적 차원의 특성에 기준하여 직관이라는 개념을 사용하는 것과 달리, 불승인의 감정 반응을 필요조건으로 하며, 그렇기에 내용적으로 불승인에 특정된 '불승인 직관' 으로 지칭하는 것이 타당하다. 또한 상기한 바와 같이, 두 사회적 직관주의 이론이 추론으로 분류한 인지 과정 즉, 불승인 직관 외의 인지는 추론으로 분류될 수 있는 가능성과 직관으로 분류될 수 있는 가능성을 모두 갖는다. 그렇기에 두 사회적 직관주의 이론이 추론으로 분류한 심리 과정은 형식적으로 특정되지 않는 내용적 차원의 개념일 필요가 있으며, 이는 '목적 지향적 인지' 로 지칭할 수 있다.

셋째, 이러한 규명 범주 구체화와 개념 범주 명료화 작업에 기반 할 때, 두 사회적 직관주의 이론의 발견은 다음과 같이 정리할 수 있다. (1)불승인 직관이 강하게 도출 되는 상황에 있어, 도덕적 판단은 피해와 공정성의 기준 적용 가능 여부와 무관하게 불승인의 직관적 판단으로 도출된다. (2)피해와 공정성의 기준이 적용되지 않는 상황에서는 불승인 직관이 존재하지 않는 경우, 어떠한 도덕적 판단도 도출되지 않는다. 그러나 피해와 공정성의 기준이 적용되는 상황에서는 불승인 직관이 존재

하지 않더라도 목적 지향적 판단의 형태로 도덕적 판단이 도출될 수 있다. (3)불승인의 직관적 판단이 도출된다 하더라도 목적 지향적 인지에 의해 수정이 이루어질 수 있다. 단, 피해와 공정성의 기준이 적용되지 않는 상황에서 이러한 수정은 매우 드물게 일어나며, 피해와 공정성의 기준이 적용되는 상황에서는 강한 결과론적 근거에 의해 이러한 수정이 더욱 용이하게 이루어지기도 한다.

넷째, 새로운 인지 유형 분류를 적용할 경우, 의무론적 판단과 결과론적 판단은 상호 보완적인 관계를 맺는다. 의무론적 판단이란 '불승인 직관'에 의해 추동된 판단으로, 오직 특정 행위의 허용 불가능성 여부를 판단하는 것만을 내용으로 한다. 반면, 결과론적 판단은 목적 지향적인 판단으로서, 목적의 달성에 더욱 적합한 행위를 선택하는 것을 내용으로 한다. 즉, 의무론적 판단은 목적과 독립적으로 특정 행위가 수용 가능한 범주를 벗어나 용인될 수 없는 행위인지 여부를 판단하는 데 기능하며, 결과론적 판단은 행위 선택지 중 목적의 달성에 부합하는 행위를 선택하는 데 기능한다. 즉, 두 유형의 판단은 도덕 기능에 있어 상이한 역할을 담당함으로써 보완적으로 작용하며, 이에 따라 의무론적 판단이 윤리적 가치를 갖기 어렵다는 그린의 주장은 타당성을 갖기 어렵다.

다섯째, 동일한 맥락에서 도덕적 전문성은 목적 지향적 인지의 전문성과 불승인 직관의 전문성으로 구분된다. 이러한 전문성은 앞서 도출한 세 가지 유형의 도덕적 판단에 있어, 판단의 수월성을 향상시키는 '적극적 차원의 전문성'과 각각의 인지에 내재한 위험성을 경계하는 '소극적 차원의 전문성'을 포괄하는 개념이어야 하며 이를 바탕으로 도덕적 전문성 유형을 보다 체계적으로 설정하면 다음과 같다. (1)적극적 차원의 불승인 직관의 전문성이란, 피해와 공정성의 기준이 적용되지 않는 상황에서, 충분한 문화화가 이루어진 불승인 직관이 내재 및 조율되어 적절하게 도출되는 것이다. (2)소극적 차원의 불승인 직관의 전문성은, 피해와 공정성의 기준이 적용되는 상황에서, 허용될 수 없는 다양한 행위들에 대한 발달된 민감성을 갖추는 것이다. (3)적극적 차원의 목적 지

향적 인지의 전문성이란, 피해와 공정성의 기준이 적용되는 상황에서, 상황에 내재한 도덕적 요소를 인지하고 해당 상황을 도덕의 상황으로 파악하며 가장 도덕적인 판단을 효과적으로 도출하는 것이다. (4)소극적 차원의 목적 지향적 인지의 전문성이란 불승인 직관과 적극적 차원의 목적 지향적 인지에 내재하는 왜곡과 편향의 위험성을 효과적으로 경계하는 것이다.

여섯째, 도제식 교육이 효과적으로 이루어지기에 많은 제약이 있는 학교 현장에서, 윤리적 기술을 기르기 위한 인지적 도제 모델을 보다 효과적으로 시행하기 위해서는 윤리적 기술의 습득이 효과적으로 이루어질 수 있는 상황을 구성할 수 있어야 한다. 특히, 직관적 기술의 습득 및 발현이 상황 맥락성을 갖는다는 점을 고려할 때, 상황 구성의 원칙을 마련하는 것은 효과적인 교육 전략을 모색하는 데 가장 핵심적인 요소 중 하나이다. 이러한 상황은 '피해와 공정성의 기준이 적용되며 불승인의 직관적 판단이 도출되는 상황'과 '피해와 공정성의 기준이 적용되지 않으며 불승인의 직관적 판단이 도출되는 상황' 그리고 '피해와 공정성의 기준이 적용되며 목적 지향적 판단이 도출되는 상황'의 조합을 통해 구성할 수 있다.

일곱째, 구성된 상황별로 관련 내용을 모두 교육하기에는 현실적 제약이 많기에, 효과적인 교육 전략을 마련해야 한다. 이를 위해서는 (1)당면 상황을 기준으로 학생들의 스키마를 점검하고, (2)그러한 점검에 기반하여 단위 수업에서 중점적으로 교육할 내용을 결정하며, (3)교육을 통해 발달시키고자 하는 인지의 특성에 맞춘 교수 방법을 정해야 한다.

주요어: 사회적 직관주의 모델, 이중과정 모델, 통합적 윤리교육 모델, 불승인 직관

학 번: 2015-30416

목 차

Ⅰ. 서론 ····· 1
1. 연구 목적 및 필요성 ····· 1
2. 연구 방법 및 범위 ····· 10

Ⅱ. 사회적 직관주의 이론의 토대 ····· 13
1. 사회적 직관주의 이론의 등장 배경 ····· 14
1) 인지발달 이론의 한계와 극복 노력 ····· 14
2) 직관적 인지 처리 과정의 강조 ····· 18
2. 사회적 직관주의의 이론적 기초 ····· 23
1) 신경과학적 기초 ····· 23
2) 심리학적 기초 ····· 27

Ⅲ. 사회적 직관주의 이론의 구조 ····· 31
1. 하이트의 사회적 직관주의 모델의 구조 ····· 32
1) 도덕 범주 인식의 문화 종속성 ····· 32
2) 직관적 도덕 판단과 사후 추론의 기제 ····· 37
3) 직관적 도덕 판단의 수정 기제 ····· 41
2. 그린의 이중 과정 모델의 구조 ····· 48
1) 직관적 도덕 판단의 기제 ····· 48
2) 직관적 도덕 판단에 대한 인지적 조절의 기제 ····· 53

Ⅳ. 사회적 직관주의 이론 비판 ····· 59
1. 도덕적 판단의 범주 설정의 문제점과 대안 ····· 60
1) 규명 범주 구체화 필요성 ····· 60
2) 규명 범주 구체화 가능성 ····· 62

3) 규명 범주 구체화의 적용 ·····66
2. 도덕적 인지의 유형 분류의 문제점과 대안 ·····75
1) 개념 범주 명료화 필요성 ·····75
2) 개념 범주 명료화 가능성 ·····80
3) 개념 범주 명료화의 적용 ·····87

Ⅴ. 사회적 직관주의 이론 비판 및 재해석의 도덕 심리학적 적용 91
1. 사회적 직관주의 이론의 재해석 ·····92
1) 사회적 직관주의 모델의 재해석 ·····92
2) 이중과정 모델의 재해석 ·····95
2. 유형별 인지의 상호 보완 가능성 분석 ·····100
1) 직관과 추론의 상호 보완 가능성 ·····100
2) 불승인 직관과 목적 지향적 인지의 상호 보완 가능성 ·····104

Ⅵ. 사회적 직관주의 이론 비판 및 재해석의 도덕 교육적 함의 109
1. 교육 목표 설정에의 적용 가능성 ·····110
1) 도덕적 전문성의 개념과 보완 필요성 ·····110
2) 도덕적 전문성 개념의 보완 가능성 ·····112
2. 교육 방법 구성에의 적용 가능성 ·····119
1) 통합적 윤리교육 모델의 특성 ·····119
2) 통합적 윤리교육 모델의 보완 가능성 ·····126

Ⅶ. 결론 ·····136

참고문헌 ·····141

ABSTRACT ·····152

[표목차]

<표 1> 추론 체제와 직관 체제의 비교 ·····19
<표 2> 도덕적 판단 관련 뇌 영역별 기능 개요 ·····25
<표 3> 인신적 딜레마와 비인신적 딜레마 예시 ·····49
<표 4> 기술 습득의 5단계 ·····81
<표 5> 인지 심리학과 도덕 심리학의 인지 유형 분류 비교 ·····88
<표 6> 도덕적 행동의 네 가지 요소와 기술 및 각각의 하위기술 ·····122
<표 7> 인지적 도제 모델의 수업 원리 ·····124
<표 8> 전문성 교육을 위한 상황 구성 원리 ·····129
<표 9> 도덕적 전문성 향상을 위한 교육 방안의 구성 원리 ·····133

[그림목차]

<그림 1> 도덕적 판단 관련 뇌 영역 ·····26
<그림 2> 웨이슨의 네 장의 카드 과제 ·····28
<그림 3> 도덕적 판단에 관한 사회적 직관주의 모델 ·····36
<그림 4> 도덕적 판단에 관한 이성주의 모델 ·····46
<그림 5> 인도교 딜레마의 수단적 변경 ·····52
<그림 6> 딜레마 유형별 응답 소요 시간 ·····55
<그림 7> 도덕적 판단에 관한 이중과정 모델 ·····57
<그림 8> 직관적 판단 및 사후 추론의 기제 ·····67
<그림 9> 추론적 도덕 판단 기제 ·····70
<그림 10> 인지적 조절에 의한 직관적 판단의 수정 기제 1 ·····71
<그림 11> 인지적 조절에 의한 직관적 판단의 수정 기제 2 ·····72
<그림 12> 두 사회적 직관주의 이론의 통합 가능성 ·····74
<그림 13> 딜레마 유형에 따른 대뇌 부위별 활성화 정도 ·····76
<그림 14> 사회적 직관주의 모델 재해석 ·····92
<그림 15> 이중과정 모델 재해석 ·····96

Ⅰ. 서론

1. 연구 목적 및 필요성

도덕성에 대한 관심은 크게 '도덕이란 무엇인가'에 대한 관심과 '도덕 기능은 어떠한 방식으로 발현 되는가'에 대한 관심으로 나눌 수 있다. 전자에 대한 논의는 주로 도덕 철학 분야의 담론을 통해 전개되어 왔으며 후자에 대한 논의는 주로 도덕 심리학 분야의 담론을 통해 전개되어 왔다. 즉, 도덕 철학의 목적이 정당화 가능한 도덕의 실체를 규명하는 것이라면 도덕 심리학의 주된 목적은 도덕 기능의 기제(mechanism of moral functioning)를 사실적 차원에서 규명하는 것이다. 정당화 가능한 도덕의 실체를 규명하는 도덕 철학 분야의 연구와 도덕 기능의 사실적 기제를 규명하는 도덕 심리학 분야의 연구는, 도덕을 어떠한 방식으로 가르쳐야 하는가에 대한 교육학 분야의 연구와 함께 성공적인 도덕 교육 방안을 모색함에 있어 필수적인 연구이다(정창우, 2013, pp. 175-177).

도덕 철학 분야와 도덕 심리학 분야의 결합을 시도한 최초의 연구는 인지주의를 대표하는 콜버그(L. Kohlberg)에 의해 수행되었다. 콜버그는 도덕의 본질에 대한 관심 없이 도덕 기능의 기제를 고려하는 것이 불가능하다고 보았기에 칸트(I. Kant)로 대표되는 이성주의 윤리에 기반하여 도덕의 정당화 작업을 선행한 후, 보편타당한 도덕 법칙에 근거한 판단과 행위라는 도덕성 발달의 지향점을 설정하고, 이성에 의한 도덕적 추론 능력의 발달을 강조하였다. 그의 연구는 도덕 철학과 도덕 심리학의 접목을 이루었다는 점에서 긍정적인 평가를 받았지만, 도덕성 발달 및 도덕 기능의 기제 규명에 있어 이성적 추론 이외의 요소들을 간과하였다는 비판을 받았다. 또한 수업에서의 도덕적 판단이 일상에서의 도덕적 행위로 이어지지 않는 경우가 빈번하다는 비판이 거세짐에 따라 인지

주의 이론과 그에 기반한 도덕 교육의 실효성에 대한 회의감이 공유되었고, 그에 따라 도덕성의 구성 요소와 각 요소들의 복합적인 작용과 같은 도덕 기능의 기제를 사실적으로 규명하려는 다양한 연구가 수행되었다.

특히, 심리학 및 신경과학 분야의 다양한 연구들이 도덕적 판단의 도출에 있어 감정 반응 및 신체 반응이 강력한 영향을 미친다는 점과 인지의 많은 부분이 암묵적이고 자동적인 특성을 갖는다는 점을 규명함에 따라, 도덕 기능 기제의 직관적 특성들을 사실적 차원에서 규명하기 위한 다양한 연구가 수행되었다. 그러한 대표적인 연구로 하이트(J. Haidt)의 사회적 직관주의 모델(Social Intuitionist Model) 및 그린(J. Greene)의 이중과정 모델(Dual Process Model)과 같은 사회적 직관주의 이론이 있다. 이들은 도덕의 정당화에 대하여 관심을 갖기보다는 다양한 신경과학 및 심리학 분야의 연구에 근거하여, 도덕 기능의 기제를 보다 사실적으로 규명하고자 하였다.

하이트는 사람들이 어떠한 것을 도덕으로 간주하며 어떠한 기제를 통해 도덕적 판단이 도출되는지를 사회적 직관주의 모델을 통해 설명하고자 하였다. 그는 동일한 상황을 대상으로 인도 오리사 지역 주민의 도덕적 평가와 미국인의 도덕적 평가가 상이한 형태로 도출된다는 슈웨더(R. Shweder)의 연구에 영향을 받아, 특정한 상황에 대하여 도덕적 판단을 내리는 일련의 과정들이 사회 구성원들이 공유하고 있는 문화적 기준에 크게 영향을 받는다고 주장하였다. 그는 특히, 서구 전통의 인지주의 패러다임에 의해 개인은 피해와 공정성의 기준에 따라 특정 상황을 도덕의 영역으로 인식하고 그에 대해 추론적 판단을 내리는 존재로 간주 되어 왔지만, 현실에서 개인들은 피해와 공정성의 기준을 적용할 수 없는 상황에 대해서도 옳고 그름에 대한 판단을 내리고 이를 도덕적 영역의 문제로 인식한다고 주장하였으며 도덕적 판단에 있어 문화의 영향을 강조하였다(Haidt, 2012/2014, pp. 56-63).

하이트는 또한 도덕적 판단은 직관적 형태로 도출되며 추론은 직관적 도덕 판단을 합리화 하는 역할을 한다고 주장하였다. 그는 도덕적 판단

의 직관적 특성을 증명하기 위하여, 서구 전통의 도덕적 기준인 피해와 공정성의 기준을 침해하지는 않지만 사회적으로 금기시 되는 행위들을 제시하고 개인들에게 평가를 내리도록 요구하였다. 개인들은 논리적인 근거를 제시하지 못하더라도 자신의 판단을 합리화하며 최초의 판단을 수정하지 않으려는 모습을 보였고, 논리적으로 설득하여도 최초의 판단을 수정하지 않으려는 모습을 보였다(Haidt, Björklund, & Murphy, 2000; Haidt, 2001).

그린(J. Greene)도 개인의 도덕적 판단이 어떠한 기제를 통해 도출되는지를 이중과정 모델을 통해 규명하고자 하였다. 그 또한 하이트와 마찬가지로 도덕적 판단이 감정 반응과 밀접한 관련을 가지며, 많은 경우에 있어 직관적 형태로 도출된다고 보았다. 그러나 추론의 기능을 주로 합리화에 한정한 하이트와 달리, 그린은 추론이 도덕적 판단의 과정에 직접 작용하며 인지적 조절에도 관여한다고 주장하였다.

그린은 도덕적 판단이 직관적 형태와 추론적 형태 모두에 의해 도출된다는 것을 증명하기 위하여, 인도교 딜레마(Footbridge Dilemma)상황과 트롤리 딜레마(Trolly Dilemma) 상황에서 도출되는 도덕적 판단의 특성을 비교 분석하였다. 비교적 강한 감정적 반응이 수반되는 인신적 도덕 딜레마(Moral Personal Dilemma)의 특성을 갖는 인도교 딜레마 상황에 있어, 피험자들은 감정 반응 및 사회적 인지와 관련된 두뇌 영역의 신경 활동이 증가하였으며 높은 비율로 사람을 밀지 않을 것을 선택하였다. 반면 감정적 반응이 비교적 약한 비인신적 도덕 딜레마(Moral Impersonal Dilemma)의 특성을 갖는 트롤리 딜레마 상황에 있어, 피험자들은 고차적 인지와 관련된 두뇌 영역이 상대적으로 더 활성화 되었으며 대부분이 스위치를 조작하는 것을 선택하였다. 이를 통해 그린은 감정적인 반응의 활성화 정도에 따라 도덕적 판단의 형태(직관과 추론)와 내용이 크게 뒤바뀔 수 있다고 주장하였다(Greene, 2001, p. 2106).

덧붙여, 그린은 직관적 판단이 도출된 경우에도 추론을 통해 최초의

판단을 변경하는 것이 가능하지만, 직관적 판단을 변경하기 위해서는 강한 감정적 반응을 극복해야하기 때문에 추론적 판단을 변경하는 것보다 더 많은 시간이 걸릴 것이라고 주장하였다. 이를 증명하기 위해 그린은 대뇌의 활성화 부위가 상이한 판단 유형별로, 긍정적인 답변과 부정적인 답변에 소요되는 시간을 비교하였다. 실험 결과, 직관과 관련된 뇌 영역이 활성화되고 강한 감정 반응이 수반되는 인도교 딜레마 상황에서는 사람을 미는 것에 대한 긍정적인 답변에 소요되는 시간이 부정적인 답변에 소요되는 시간보다 길었다. 반면, 추론과 관련된 뇌 영역이 활성화 되고 상대적으로 약한 감정 반응이 수반되는 트롤리 딜레마 상황에서는 스위치를 조작하는 것에 대한 긍정적인 답변에 소요되는 시간과 부정적인 답변에 소요되는 시간 사이에 별 차이가 없었다(Greene, 2001, p. 2107).

이렇듯 두 이론은 도덕적 판단에 있어 직관과 추론의 역할 및 관계에 대해 상충하는 주장을 제시하였고 이는 후속된 다양한 논의로 이어졌다. 가장 우선적으로 이루어졌던 논의는 과연 두 이론 중 어떠한 이론이 더욱 타당한가에 대한 것이었다. 이론의 타당성에 대한 논의는 추론적 도덕 판단의 가능성 또는 도덕적 판단에 있어 추론의 역할에 관한 논의를 중심으로 진행되었다. 이와 관련하여 나바에츠(D. Narvaez)는 일상적으로 일어나는 다양한 도덕적 판단에 있어 직관적인 과정이 절대적으로 높은 비중을 차지하지 않으며 심사숙고에 의한 추론적 과정도 높은 비중을 차지한다고 지적하며(Narvaez, 2008, p. 235) 그린의 이론을 옹호하였고, 다른 많은 연구자들도 다양한 근거를 제시하며 직관과 의식적인 추론이 우선순위를 번갈아가며 상호작용한다는 그린의 주장을 지지하였다(정창우, 2011, pp. 111-120).

추론의 역할과 관련하여 그린의 이론이 지지됨에 따라, 추론적 인지와 직관적 인지의 관계 방식에 대한 논의 또한 그린의 프레임 속에서 이루어졌다. 이러한 두 인지의 관계 방식에 관한 관점은 그린의 카메라 비유를 통해 이해할 수 있다. 그린은 인간의 뇌를 자동 및 수동의 이중적 방식

에 의해 구동되는 디자인의 카메라로 묘사하였다(Greene, 2014/2017, pp. 205-212). 그에 따르면 직관적 도덕 판단은 자동 방식으로서, 많은 경우에 있어 감정적인 특성을 띄며, 개인은 이러한 직관적인 반응이 유발되는 과정에 대하여 의식하지 못한다. 반면, 추론적 도덕 판단은 수동 방식으로서, 현재의 환경적 자극이나 신체 반응에 의해 자동적으로 활성화 되지 않는 장기적인 목표를 지향하는 등 이성적인 특성을 강하게 갖는다(Greene, 2015, pp. 142-143). 이와 같이, 직관은 효율성을 담보하지만 왜곡과 편향을 수정하기 어렵다는 위험성을 갖는 반면, 추론은 왜곡과 편향의 위험을 경계할 수 있는 가능성은 더 높지만 효율성이 떨어지기에 두 인지 유형은 상호 보완적 특성을 갖는다는 관점이 통설로 받아들여지고 있다.

이렇듯 직관과 추론의 역할 및 관계에 대해 두 사회적 직관주의 이론은 충돌하는 관계로 상정되어 왔으며, 합의점을 도출하지 못한 채 그린의 이론을 더욱 포괄적이고도 타당한 이론으로 받아들이게 되었다. 또한 도덕적 판단의 현상을 포괄적이고 일반적으로 규명하려는 노력은 도덕적 판단의 현상에 있어 인지 심리학적 규명들을 재확인하는 결과로 이어지게 되었다. 또한 이로 인하여 사회적 직관주의 이론과 관련 논의는 도덕 심리학 고유의 발견으로 긴밀히 연결되지 못하였으며, 이로 인하여 다양한 연구들은 실효성 있는 교육 방안의 마련으로 충분히 이어지지 못하였다.

그런데 과연 이러한 논의들은 타당성을 담보하고 있을까. 상기한 바와 같이, 두 사회적 직관주의 이론이 실제로 규명한 현상 범주보다 더 넓은 범위의 현상들을 포괄하는 것으로 확대 해석될 위험을 경계하기 위해서는 이론이 규명한 도덕적 판단의 현상 범주를 구체화하는 작업이 선행되어야 한다. 또한 사회적 직관주의 이론에 내재할 수 있는 해석상 모순의 위험성을 경계하기 위해서는 이론이 활용한 개념 범주[1]가 연구자의 의

1) 개념 범주란 개별적인 개념들을 포괄하는 추상적인 상위 차원의 개념을 의미한다. 이러한 개념 범주가 있어야지만 연구의 명제나 가설이 존재할 수 있으며, 개념 범주를 공유하는 연구자 사이에 의사소통도 가능해진다(Strauss & Corbin, 1990/1995,

도와 다르게 해석될 여지가 없는지 검토하는 작업 즉, 개념 범주 명료화 작업이 선행되어야 한다. 그러나 사회적 직관주의 이론에 관한 논의들은 이러한 비판적 검토가 선행되지 않은 상태로 진행된 측면이 있다.

본 연구는 두 사회적 직관주의 이론의 규명 범주 구체화 및 개념 범주 명료화 작업이 선행되지 않을 경우, 관련 논의의 타당성을 담보할 수 없다는 문제의식을 가지고 두 이론의 규명 범주 구체화 및 개념 범주 명료화 작업을 수행하고자 한다. 또한 이에 기반하여 두 이론에서 파생된 논의의 타당성을 검토하고, 인지 심리학 분야의 발견들과 구별되는 사회적 직관주의 이론 고유의 발견을 도출할 수 있는 가능성을 모색하고자 한다. 또한 이러한 과정을 통해 도출된 도덕적 판단의 사실적 기제에 관한 발견을 도덕 교육에 적용할 수 있는 방안을 모색하고자 하며, 이는 직관적 인지의 기제를 반영한 대표적인 교육 이론인 나바에츠의 통합적 윤리 교육 모델(Integrative Ethical Education)을 중심으로 수행하고자 한다.[2)]

이에 대하여 보다 상세하게 살펴보면 다음과 같다. 두 이론에 관한 논의는 두 연구자의 표현에 전적으로 의존하고 있다. 즉, 두 이론은 두 연구자가 표기한 '도덕적 판단'이라는 동일한 현상 범주에 대해 상이한 규명을 하고 있는 것으로 이해되었으며, 두 이론이 규명한 도덕적 판단의 현상 범주를 구체화하는 작업은 수행되지 않았다. 그러나 두 이론이 실제 규명한 현상보다 광범위한 현상 범주를 규명한 것으로 확대 해석하는 오류를 경계하고 두 이론을 타당하게 평가하며 이를 통해 두 이론의 합의점을 모색하기 위해서는 두 이론이 어떠한 범주의 도덕적 판단 현상을 대상으로 하는지에 대하여 보다 구체적으로 특정할 필요가 있다. 예를 들어, 만일 두 이론이 도덕적 판단이라는 현상의 상이한 범주를 규명 대상으로 한다면, 두 이론 사이에는 충돌이 성립되기 어려우며, 오히려

---

pp.75-79).

2) 하이트와 그린의 경우, 도덕 교육에 관한 직접적인 연구를 별로 수행한 바가 없기 때문에 직관에 관한 인지 심리학 분야의 연구 성과를 적극 반영한 나바에츠의 도덕교육 이론을 중심으로 이론의 보완 가능성을 모색하고자 한다.

두 이론은 도덕적 판단이라는 현상을 구성하는 하위의 현상 범주들을 보완적으로 규명한 관계로 해석될 여지도 있다.

스트라우스(A. Strauss)와 코빈(J. Corbin)이 지적하였듯 자료에 대한 모든 이론적 설명과 범주 설정 등은 언제나 특수한 상황을 전제로 한다(Strauss & Corbin, 1990/1995, p.55). 그럼에도 불구하고 뷰러워이(Burawoy, 1991)가 지적하였듯, 이론은 특정한 연구에서 얻어낸 경험적 일반화를 보다 거대한 실재에 대한 일반화된 진술로 받아들이게 하는 위험성을 내포한다(Charmaz, 2006/2013, pp. 266-271). 그렇기에 규명 범주를 구체화하는 작업은 두 이론의 규명을 명료하게 이해하고 이론 간 합의점을 도출하기 위해 필수적으로 요청된다.

이러한 작업과 함께, 본 연구는 두 이론이 차용한 직관과 추론이라는 인지 심리학적 개념 범주를 활용하여 도덕적 판단의 현상을 규명하는 것이 타당한지 검토할 것이다. 또한 만일 인지 심리학적 개념 범주의 차용으로 인하여 도덕적 판단의 현상 규명에 오류가 발생한다면, 이를 보완하기 위하여 보다 타당한 개념 범주 설정의 가능성을 모색할 것이다. 이를 통해 도덕적 판단의 현상을 대상으로 인지 심리학적 규명들을 재확인하는 것에 그치지 않고, 도덕 심리학적으로 보다 의미 있는 사회적 직관주의 이론 고유의 발견을 도출할 수 있는 가능성을 모색할 것이다.

이러한 두 번째 작업에 대하여 보다 상세하게 살펴보면 다음과 같다. 상기한 바와 같이 사회적 직관주의 이론은 인지 심리학적 개념 범주인 직관과 추론을 차용하여 도덕적 판단의 현상을 설명하고자 하였다. 또한 연구자들이 '도덕적 판단'의 현상에 대한 포괄적이고 일반적인 규명을 보다 타당한 것으로 간주하였기에, 사회적 직관주의 이론에 관한 논의들은 도덕 심리학의 현상을 대상으로 인지 심리학의 발견을 재확인하는 결과로 이어졌다. 그러나 이러한 과정에서, 과연 도덕적 판단의 현상을 해석하기 위해 인지 심리학적 개념 범주를 차용할 경우, 모순이 발생할 위험성이 없는지에 대한 검토가 이루어지지 않았고, 도덕적 판단의 현상을 더욱 타당하게 해석하기 위해 인지 심리학적 개념 범주를 보완할 수 있

는 방안이 없는지에 대한 검토 또한 이루어지지 않았다.

차마즈(K. Charmaz)에 따르면 복합적이고 다의적인 해석의 가능성을 갖는 용어의 엄밀하지 못한 사용은 이론의 의의를 손상시키며(Charmaz, 2006/2013, p. 292), 이론에서 얻은 이해는 현상에 대한 이론가의 해석에 의존하기에(Charmaz, 2006/2013, p. 254) 연구자의 가정과 전제에 대해 인지하고 그러한 가정과 전제가 연구에 영향을 미치는 방식을 분석하는 것은 이론을 보다 명료하게 이해함에 있어 매우 중요하다(Charmaz, 2006/2013, p. 263). 그렇기에 만일 도덕적 판단의 현상을 설명함에 있어 인지 심리학적 개념 범주가 연구자의 의도와 다르게 해석될 위험성을 내포한다면, 현상을 보다 타당하게 해석하기 위해 개념 범주의 수정 또는 보완이 불가피해진다.

덧붙이자면, 이러한 개념 범주 설정의 타당성에 대한 검토는 개방 코딩[3] 단계에 대한 검토라고 볼 수 있다. 개방코딩은 면밀한 자료 검토를 통해 현상에 이름을 붙이고 범주화하는 일종의 분석 작업으로, 이러한 초기 분석이 전제되어야지만 이후 단계의 분석이 가능하다(Strauss & Corbin, 1990/1995, p. 76). 그렇기에 이론이 도덕적 판단의 현상을 해석하기 위해 활용한 개념 범주 즉, 직관과 추론이라는 인지 유형 분류의 적합성을 검토하는 것은 유형별 인지의 상호 보완 가능성 등 사회적 직관주의 이론 고유의 발견을 도출하고 이를 도덕 교육에 반영함에 있어 기초가 되는 작업이다.

최종적으로, 본 연구는 이러한 선행 작업에 기반하여 두 사회적 직관주의 이론을 재해석하고, 사회적 직관주의 이론 고유의 발견을 도출하며, 이를 도덕 교육 이론에 적용하기 위한 방안을 도출하고자 한다. 또한 상기한 바와 같이 이러한 교육적 적용은 직관적 인지의 기제를 반영

3) 스트라우스와 코빈에 따르면 코딩이란 자료를 분석하고 개념화시켜 새로운 방식으로 재조합하는 과정으로, 자료에서 이론이 정립되는 핵심과정을 의미한다. 이론의 형성에 있어 코딩은 개방 코딩(open coding), 축 코딩(axial coding), 선택 코딩(selective coding)으로 나뉜다. 개방 코딩은 현상을 개념화하고 범주를 발견하는 과정이다(Strauss & Corbin, 1990/1995, pp. 71-90).

한 대표적인 교육 이론인 나바에츠의 통합적 윤리교육 모델(Integrative Ethical Education)을 중심으로 수행하고자 한다.

이를 위해 본 연구는 다음의 연구 문제들을 탐구하고자 한다. 첫째, 사회적 직관주의 모델과 이중과정 모델이 어떠한 현상 범주를 규명 대상으로 하고 있으며, 그러한 규명 범주 구체화 작업에 기반 할 때 두 이론의 주장은 합의점을 도출할 수 있는가. 둘째, 두 이론이 차용하고 있는 인지 심리학 분야의 인지 유형 분류 방식은 도덕적 판단의 현상을 모순없이 설명하고 있는가. 또한 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 보다 타당하게 해석하기 위해 그러한 인지 유형 분류를 보완할 수 있는 방안으로는 어떠한 것이 있는가. 셋째, 이러한 비판적 검토 작업에 기반 할 때, 두 이론은 어떠한 방식으로 재해석 될 수 있으며, 유형별 인지는 어떠한 방식으로 관계를 맺는가. 넷째, 이러한 사회적 직관주의 고유의 규명들은 통합적 윤리교육 모델에 어떠한 방식으로 적용될 수 있는가. 각 연구 문제의 답을 찾는 일련의 과정이 도덕적 판단의 기제를 보다 명료하게 해석하고 이를 통해 보다 실효성 있는 도덕 교육의 방안을 모색하는데 기여할 수 있기를 기대한다.

2. 연구 방법 및 범위

본 연구는 사회적 직관주의 모델과 이중과정 모델에 관한 논의가 전제하고 있는 '도덕적 판단'이라는 규명 범주를 구체화하고, '직관과 추론'이라는 개념 범주의 의미를 명료화함으로써, 두 이론의 합의점을 모색하고 인지 심리학적 규명과 차별화되는 사회적 직관주의 이론 고유의 규명을 도출하고자 한다. 그리고 이에 기반하여 사회적 직관주의 이론과 관련된 도덕 심리학 분야의 논의들을 분석하고, 이를 통해 도덕 교육학에의 적용 방안을 도출하고자 한다. 이에 본 연구자는 다음의 사항들을 중심으로 문헌 분석을 통해 연구를 수행하고자 한다.

첫째, 하이트의 사회적 직관주의 모델과 그린의 이중과정 모델이 기반하고 있는 제반 연구들을 검토하고 이에 기반하여 두 이론의 구조와 특성을 상세하게 검토할 것이다. 이는 두 이론에 대한 비판적 검토 작업을 수행하기 위한 기초 작업으로서의 의미를 갖는다. 이는 하이트와 그린의 연구에 영향을 미친 인지 심리학 및 신경 과학 분야의 연구 문헌을 검토하는 작업과 하이트와 그린의 이론이 직접 도출된 연구 문헌을 검토하는 작업을 통해 수행할 것이다.

Ⅱ장에서는 사회적 직관주의 이론들의 형성에 영향을 준 연구들을 검토함으로써 사회적 직관주의 이론의 학적 기반을 검토할 것이다. 먼저 인지주의를 대표하는 콜버그의 인지발달론이 갖는 특징을 살펴보고 인지주의 이론의 한계를 극복하기 위한 연구의 등장 과정 및 직관주의가 조명을 받게 된 배경을 살펴볼 것이다. 또한 직관주의 연구들이 공유하는 인지 심리학적 개념 범주로서의 직관과 추론의 특성을 살펴보고, 사회적 직관주의 이론에 직접적으로 영향을 준 신경과학 및 인지 심리학 연구를 검토함으로써, 도덕 기능의 기제 규명에 있어 직관이 어떠한 방식으로 강조되어 왔는지 보다 상세하게 검토할 것이다.

신경과학적 연구로는 복내측전전두피질(ventromedial prefrontal cortex, 이하 vmPFC)의 손상이 도덕적 기능의 발현에 지장을 준다는

다마지오(A. Damosio)의 연구를 비롯하여, 대뇌 부위와 도덕 기능 간의 관련성을 도출한 연구들을 살필 것이며, 심리학적 연구로는 확증편향(confirmation bias)에 관한 쿤(D. Kuhn)의 연구와 인지 과정의 분류에 관한 마골리스(H. Margolis)의 연구를 살필 것이다. 또한 Ⅲ장에서는 하이트의 사회적 직관주의 모델과 그린의 이중과정 모델의 구조와 특성을 세부적으로 검토함으로써 두 이론에 대한 비판적 검토의 기초 작업을 수행할 것이다.

둘째, 하이트의 사회적 직관주의 모델과 그린의 이중과정 모델에 관한 비판적 검토를 통해 관련 논의들의 타당성을 검토하는 작업을 수행할 것이다. 이는 도덕적 판단으로 범주화 된 두 연구가 실제 규명한 현상의 범주를 구체화 하는 작업과 도덕적 판단이라는 현상을 규명하기 위하여 활용된 직관과 추론이라는 개념 범주 활용의 타당성에 대한 검토 작업 및 보완 가능성 모색 작업으로 나누어 수행할 것이다. 상기한 바와 같이, 전자는 두 이론의 합의점 도출 가능성을 모색하기 위하여 요청되며, 후자는 인지 심리학의 발견을 재확인하는 것을 넘어 사회적 직관주의 이론 고유의 규명을 도출하기 위하여 요청된다.

먼저, Ⅳ장 1절에서는 두 사회적 직관주의 이론의 규명 범주를 구체화하는 작업을 통해 두 이론의 합의점 도출 가능성을 모색할 것이다. 스트라우스와 코빈의 연구(Strauss & Corbin, 1990) 및 차마즈의 연구(Charmaz, 2006)에 기반하여 두 사회적 직관주의 이론의 규명 범주를 구체화함으로써 두 이론의 관계 방식을 명료화 하고자 한다. 규명 범주의 구체화 작업을 수행하기 위해 두 연구가 상정하는 도덕적 판단의 현상에 대한 언급을 검토하고 두 연구가 도덕적 판단의 기제를 규명하기 위해 활용한 문항의 특성을 비교할 것이다. 이를 통해 이론간 관계 설정의 적절성을 검토하고 두 이론이 합의점을 도출할 수 있을지 여부에 대하여 살펴볼 것이다.

Ⅳ장 2절에서는 도덕적 판단이라는 현상을 규명하기 위하여 두 사회적 직관주의 이론이 차용한 직관과 추론이라는 인지 심리학적 개념 범주

설정이 갖는 타당성을 검토하고 그 보완 가능성을 모색할 것이다. 개념 범주 설정의 타당성을 검토하는 작업은, 현상의 규명에 있어 인지 심리학적 개념 범주 차용으로 인하여 오류가 발생하지 않는지 여부를 검토하는 방식으로 수행할 것이다. 만일 인지 심리학적 개념 범주가 연구자의 의도와 다르게 해석될 위험성을 내포하거나 이로 인하여 도덕적 판단의 현상을 규명하는데 있어 오류가 발생한다면, 이를 해결하기 위한 개념 범주의 보완 방안을 모색할 것이다.

셋째, 사회적 직관주의 모델과 이중과정 모델에 대한 비판적 검토 작업이 갖는 도덕 심리학적 의의를 도출할 것이다. Ⅴ장 1절에서는 두 이론의 규명 범주 구체화 작업 및 개념 범주 명료화 작업에 기반하여, 두 사회적 직관주의 이론의 발견들을 보다 타당하게 해석할 것이다. Ⅴ장 2절에서는 유형별 인지의 상호 보완 가능성을 모색함에 있어 사회적 직관주의 이론에 대한 비판적 검토 작업이 갖는 의의를 도출할 것이다.

끝으로, 선행한 작업들을 통해 도출한 사회적 직관주의 이론 고유의 발견을 도덕 교육에 적용하기 위한 방안을 모색할 것이다. 이를 위해 Ⅵ장에서는 직관주의 연구의 성과를 적극 반영한 대표적인 도덕 교육 이론인 나바에츠의 도덕적 전문가 모델의 보완 가능성을 모색할 것이다. Ⅵ장 1절에서는 사회적 직관주의 이론 고유의 발견에 기반하여, 도덕적 전문가 모델의 교육 목표인 도덕적 전문성 개념의 보완 가능성을 모색하는 작업을 수행할 것이다. Ⅵ장 2절에서는 도덕적 전문가 모델이 제시한 교육 방법의 세부적인 구조와 특성을 검토하고, 사회적 직관주의 이론 고유의 발견에 기반하여 도덕적 전문가 모델이 제시한 교육 방법을 보완할 수 있는 방안을 모색할 것이다.

Ⅱ. 사회적 직관주의 이론의 토대

본 장에서는 사회적 직관주의 이론들이 등장하게 된 배경에 대하여 살펴볼 것이다. 먼저 그간 도덕 심리학 분야와 도덕 교육학 분야를 이끌어 온 인지발달 이론의 한계를 검토하고, 인지발달 이론의 한계를 극복하고자 수행된 연구들을 살펴보고자 한다. 도덕적 판단에 있어 이성적 추론의 역할을 강조한 인지발달 이론의 한계를 극복하고자 수행된 연구들은 도덕적 판단에 있어 감정의 중요성을 강조하고 보다 통합적인 관점에서 도덕적 판단의 기제를 규명하고자 하였다. 도덕 심리학계에서 수행된 다양한 연구의 흐름을 거시적 맥락 속에서 살피는 작업은 사회적 직관주의 이론들을 보다 명료하게 이해하기 위한 수단이 될 것이다. 또한 본 장에서는, 사회적 직관주의 이론들이 도덕적 판단의 기제 규명을 위해 차용한 인지 심리학적 개념 범주의 특성을 검토할 것이며, 두 사회적 직관주의 이론에 영향을 준 신경과학 및 인지 심리학 연구들을 검토할 것이다. 두 사회적 직관주의 이론에 직접적으로 영향을 준 신경과학 및 심리학적 기반들을 세부적으로 살피는 작업은 사회적 직관주의 이론을 보다 명료하게 이해하기 위한 토대가 될 것이다.

1. 사회적 직관주의 이론의 등장 배경

1) 인지발달 이론의 한계와 극복 노력

과학적 연구 방법의 등장과 함께 도덕성에 대한 연구는 정신분석학 이론, 사회학습 이론, 인지발달 이론을 통해 진행되어 왔다. 프로이드(S. Freud)에 의해 시작된 정신분석학 이론은 도덕성 발달에 있어 이성적이고 인지적인 요인보다는 감정적 또는 비인지적인 충동들에 초점을 맞추었다. 반두라(A. Bandura)로 대표되는 사회학습 이론은 도덕성 발달에 있어 행동적 측면에 초점을 맞추었으며 도덕성 발달을 학습과 조건화 및 사회화의 산물로 보았다. 피아제(J. Piaget)와 콜버그(L. Kohlberg)를 위시한 인지발달 이론은 도덕성 발달의 핵심을 인지적 추론 능력의 발달로 보았으며 개인의 도덕성 발달이 지적 발달과 동일하게 인지 구조의 평형화 과정을 통해 이루어진다고 보았다.

이러한 이론들 중 인지발달 이론은 콜버그에 의해 1960년대 인지혁명(cognitive revolution)이 일어난 후, 개인의 도덕성을 설명하고 도덕 교육의 방향을 설정함에 있어 최근까지 지배적인 위치를 점해왔다. 콜버그는 도덕성에 관한 논의를 철학적이고 사변적인 범주에 한정시키거나 심리학적이고 경험적인 범주에만 한정하는 것을 지양하였고, 이성주의 윤리와 인지발달 이론의 결합을 통해 각각의 영역에서 별개로 논의되어 온 철학과 심리학이 갖는 한계를 보완하고자 하였다. 그는 도덕성의 개념을 철학자들처럼 철학적 사변의 성격으로 규정하거나 정신분석 이론과 행동주의 이론에서와 같이 심리학적 경험의 대상으로만 여기는 이분법에서 탈피하여, 경험적 자료의 측면과 도덕 철학적 논리의 두 측면을 동시적으로 고려할 필요가 있다는 점을 강조했다(정창우, 2006, p. 35).

콜버그는 피아제의 구성주의적 인지발달론을 받아들여 추론 능력의 발달이 도덕성 발달을 가장 잘 설명해 줄 수 있다고 믿었다. 그는 도덕적 인지가 도덕적 행동에 있어 가장 중요한 요소이자, 행위를 도덕적으로

정당화 할 수 있는 유일한 요소라고 보았다. 콜버그에게 있어 도덕 발달 단계란 도덕 원리의 발달 단계를 의미하며, 도덕 원리란 도덕적 선택의 내용이 아닌 선택의 근거를 나타내는 형식 즉, 인지구조를 의미한다(정창우, 2013, p. 184).

이러한 관점을 전제로, 콜버그는 피아제의 임상적 면담법을 활용하여 중산층 백인 청소년들을 대상으로 도덕성 발달을 연구하였으며, 이를 통해 도덕적 성숙이 3수준 6단계의 불변의 계열성을 통하여 이루어진다고 주장하였다(추병완, 2004, p. 200). 콜버그는 이러한 도덕 발달 단계와 관련하여, '구체적인 개인주의적 관점'으로 규정된 인습이전 수준에 1, 2단계를, '사회구성원적 관점'인 인습 수준에 3, 4단계를, '사회에 선행하는 관점'으로 규정된 인습 이후 수준에 원리화된 추론 단계로서의 5, 6단계를 포함시켰다. 덧붙여, 그는 피아제와 같이, 후속 단계는 선행 단계로부터 나오는 것이며 동시에 선행 단계를 대체하는 것이기 때문에 높은 도덕적 추론 단계에 도달한 사람은 결코 이전의 단계로 퇴행하지 않는다고 주장하였다(정창우, 2006, p. 42). 덧붙여, 그는 도덕적 판단이 강한 정서적 요소를 포함하기도 하지만, 그러한 경우에도 인지적 요소에 의한 정서적 요소의 해석에 따라 판단이 이루어진다고 지적하며, 인지적 요소가 행위 결정에 있어 우선성을 갖는다고 주장하였다(Kohlberg, 1981/2000, pp. 186-191).

또한 콜버그는 이러한 순차적이고 계열적인 도덕 발달이 문화적 특수성에 구애받지 않는 보편적 현상임을 주장하였다(추병완, 2004, p. 200). 그는 도덕성에는 보편적인 도덕적 사고의 원리가 존재하고, 그렇기에 도덕성이 이상적인 수준으로 향할수록 모든 문화에서 동일한 형식을 지니며, 도덕적 사고의 원리는 동일한 발달 단계를 밟는다고 주장하였다. 즉, 콜버그는 다양한 문화권의 모든 개인에게 있어 도덕적 사고의 원리가 발달하는 정도는 다양하지만, 그러한 발달의 단계는 보편적이고 동일한 불변적 순서를 따른다고 보았다(Kohlberg, 1981/2000, p. 60).

콜버그는 피아제의 구성주의적 인지발달론을 받아들였기에, 도덕적 추

론 능력의 발달이 환경과의 능동적인 상호작용을 통해 이루어지며 동화-조절-평형화의 과정 즉, 연속적인 조절(accommodation)의 결과라고 주장하였다. 이러한 관점에 근거하여 그는 상위 단계의 도덕적 추론에 개인을 노출시키는 방식을 통해 인지적 불균형을 유발하고, 이러한 인지적 비평형 상태를 통해 개인의 도덕 발달 수준 즉, 도덕적 사고의 형식 차원의 추론 능력을 향상시키고자 하였다. 또한 이러한 블래트 효과(Blatt Effect)를 달성하기 위한 방안으로 그는 도덕 딜레마 토론 및 정의 공동체 접근(Just Community Approach)을 제시하였다(정창우, 2006, pp. 48-51).

콜버그의 이러한 접근은 도덕철학과 도덕 심리학의 결합에 큰 역할을 하였으나 다양한 심리학 연구의 이론적이고 경험적인 성과들을 반영하지 못함으로써 고립되었으며, 도덕 기능의 사실적 기제를 설명하는데 실패하였다는 비판을 받았다. 그의 이론은 이성에 의한 도덕 추론만을 강조하였고, 그로 인하여 도덕적 정서와 같이 개인의 도덕 기능에 관여하는 다양한 요소의 역할과 중요성에 대하여 간과하였다(정창우, 2013, p. 200). 이는 콜버그가 도덕의 영역을 주로 상이한 개인들의 주장, 권리, 이해관계에서의 갈등을 공평하고 합리적으로 해결하는 방식과 절차에 국한하였다는 점에 기인한다. 이로 인하여 콜버그의 인지발달 이론은 일상적인 사회적 상호 작용에 내재한 도덕적 특성들을 간과하였으며 도덕적 정서와 같은 도덕성의 비인지적 요소가 갖는 중요성 또한 간과하였다(추병완, 2004, p. 206).

이에 대한 극복 노력으로, 나딩스(N. Noddings)는 배려윤리를 제시하였다. 인지발달 이론이 공정성, 권리, 자율성, 독립성, 보편성 등을 중요한 도덕발달의 척도로 삼았던데 비해, 배려윤리는 특수성, 맥락성, 책임감, 인간관계, 배려 등을 중요한 도덕 발달의 척도로 제시하였다(박병춘, 2005, pp. 241-242). 나딩스에 따르면 배려의 핵심 요소는 전념(engrossment)과 동기 전환(motivational displacement)으로, 전념이란 자신을 향해 맞춰져 있던 의식의 초점을 타인을 향해 맞추는 것을 의

미하며, 동기 전환이란 이러한 전념을 통해 자아에서 타자로 동기가 전환되는 것을 의미한다. 또한 그녀는 배려를 자연적 배려와 윤리적 배려로 구분하며, 자연적 배려는 보다 근원적이고 상위의 배려이고 윤리적 배려는 이러한 자연적 배려에 기초하며 동시에 자연적 배려를 목표로 한다고 주장하였다. 또한 그녀는 자연적 배려와 달리 윤리적 배려는 다른 사람에 대해서 느끼는 의무감에 대한 응답에서 도출되며, 이러한 윤리적 배려는 규칙 혹은 원리가 아닌 윤리적 자아의 발달에 의존한다고 강조하였다(박병춘, 1998, pp. 383-388).

호프만(M. Hoffman) 또한 개인이 일상에서 마주하는 전형적인 도덕적 상황은 곤경에 빠진 사람에게 도움을 줄 수 있는 방관자적 상황과 누군가에게 해를 끼치는 위반자적 상황이며, 도덕적 딜레마 상황은 전형적인 도덕적 상황이 아니라고 지적하였다. 그는 합리적 추론이 도덕적 행위를 동기화 하지 못한다고 비판하며, 도덕적 행위의 동기화에 있어 감정적 요소의 작용 기제에 관심을 가졌고 도덕적 삶에 있어 공감(empathy)의 중요성을 강조하였으며, 이러한 공감적 감정의 네 가지 주요 발달단계를 제시하였다(박병기, 추병완, 2007, pp. 338-341). 그는 도덕적 인지가 공감이라는 정서적 요소에 의해 '뜨거운 인지(hot cognition)'가 될 수 있다고 주장하였다(정창우, 2013, p. 186).

이렇듯, 인지발달 이론의 한계를 극복하기 위한 노력은 도덕적 판단의 맥락성에 대한 강조 및 도덕적 판단에 있어 감정의 역할에 대한 강조로 도출되었다. 이와 함께 일상의 행위 결정에 있어 이성과 감정의 분리된 기능이 불가능하다는 관점을 전제로, 인지의 형태를 이성과 감정이 아닌 직관과 추론으로 분류하는 인지 심리학 연구들이 수용되었다. 이에 따라, 맥락 의존적이고 감정적인 특성을 갖는 직관에 대한 관심이 커지게 되었고, 도덕적 직관에 관한 다양한 연구와 논의가 이어졌다. 이러한 도덕적 직관 개념에 대한 강조는 도덕적 행동이 반드시 이성에 의한 숙고를 통해 발현되는 것이 아니라는 점을 보여줌으로써 도덕적 행동의 기준과 동기화에 대한 새로운 관점을 제공하였다(정창우, 2013, p. 201).

이러한 관점의 전환은, ‘우리의 인지 활동이 콜버그가 주장하는 것처럼 의식적이고 명시적이지 않다’는 랩슬리(D. Lapsley)와 나바에츠(D. Narvaez)의 지적에서 드러난다. 이들은 사회 인지적 기능의 상당 부분이 함축적이고 암묵적이거나 혹은 자동적이며, 대개 도덕적인 행동을 하는 개인들이 자신들의 동기가 무엇인지에 대해 분명하게 표현할 수 없고, 어떤 판단들을 통해 그 행위를 하게 되었는지에 대해서도 말하지 못한다고 강조하였다. 이들은 또한 콜버그의 주장과 같이 도덕적 행동이 의식과 명백한 숙고에 달려 있다고 한다면, 사회적 인지의 이러한 자동적이고 암묵적인 특성에 비추어 볼 때, 많은 인간 행동들은 절대로 도덕적 행동이라고 할 수 없을 것이라고 지적하였다(Lapsley & Power, 2005/2008, p. 255). 이러한 관점의 전환은 도덕적 직관이 무엇이고, 도덕적 직관이 어떻게 발현되며, 도덕 판단에 있어 도덕적 직관의 타당성을 얼마나 신뢰할 수 있는가 등과 같은 다양한 논의들로 이어졌으며, 사회적 직관주의 이론의 형성에 영향을 미치게 되었다.

2) 직관적 인지 처리 과정의 강조

도덕 기능의 사실적 기제를 규명하기 위한 다양한 노력의 일환으로 직관에 대한 연구가 수행되었다. 직관주의 분야의 연구자들은 이성과 감정이 분리되어 작용하기 어렵다는 점을 강조하였으며, 인지주의 접근이 현실을 제대로 설명하지 못하는 이유를 ‘감정과 분리된 이성의 작용에 대한 강조’에서 찾았다. 도덕적 판단에 관한 직관주의 접근의 대표적인 연구자인 하이트(J. Haidt)는 다양한 연구를 근거로, 다음과 같이 이성의 강조에 대한 의구심을 갖게 되는 네 가지 이유를 제시하였다.

첫째, 직관적 판단의 존재는 감정과 분리된 이성의 중요성에 대한 의구심으로 이어진다. 다양한 연구들은 직관적 판단이 감정의 영향 하에 놓인다는 것을 밝힌다. 둘째, 추론이 작용하기 위해서는 감정적 동기화가 필요하고 셋째, 동기화 된 추론은 직관적 판단에 대한 사후 합리화의 역할을

한다. 즉, 감정적 반응은 최초의 직관적 판단을 도출하고 이에 대한 합리화를 담당하는 추론을 동기화 하는데 영향을 미친다. 넷째, 도덕적 행동의 도출은 도덕적 추론보다는 도덕적 감정에 의해 이루어진다(Haidt, 2001, pp. 819-825).

이렇듯 직관주의 연구자들은 감정과 분리된 이성이 독립적으로 작용할 수 없다는 관점을 전제로, 이성적 요소와 감정적 요소가 복합적으로 작용하는 인지(cognition)의 형태로서 추론(reasoning)과 직관(intuition)을 구분하고 그 형태별 특성을 도출하는데 관심을 기울였다. 하이트(J. Haidt)는 이와 관련하여, 여러 학자들의 견해를 종합하여 <표-1>과 같이 추론적 인지 과정과 직관적 인지 과정의 특성을 비교하였다.

<표-1> 추론 체제와 직관 체제의 비교 (Haidt, 2001, p. 818)

추론 체제(reasoning system)	직관 체제(intuitive system)
느림	자동적이고 반사적이고 신속함
의도적인 인지적 노력이 필요	인지적 노력이 필요 없음
의식적으로 과정을 파악할 수 있음	과정을 파악하기 어렵고 단지 결과만을 파악할 수 있음
주의 깊은 자료원(resources)이 요구됨	주의 깊은 자료원(resources)이 요구되지 않음
연속적인 과정	평행적인 과정
분석적, 진실 보존적 사고	은유적, 총체적 사고
인간은 2세 이후에 가능하며 훈련받은 약간의 원숭이도 가능	모든 포유동물에 공통적으로 있음
맥락(context) 독립적	맥락(context) 의존적
기반(platform) 독립적	기반(platform) 의존적(직관을 저장하고 있는 두뇌와 신체에 의존)

직관주의 분야의 연구자들은 일상적 판단이 높은 비중으로 직관적 형

태로 도출된다는 인지 심리학적 연구 결과를 공유하였고 도덕적 판단에 있어 직관의 역할에 관심을 가졌다. 인지 심리학이 전제하는 직관적 인지의 특성은 대표적인 인지 심리학자인 바흐(J. Bargh)의 연구를 통해 살펴볼 수 있다. 그는 직관적 인지의 특성과 관련하여 비의도성(의지를 가지고 있지 않더라도 처리가 이루어짐), 비노력성(처리에 있어 노력을 필요로 하지 않음), 비자각성(개인이 정보 처리 과정을 의식하지 못함), 비통제성(개인이 처리 과정을 멈출 수 없음), 효율성(최소한의 인지적 용량으로 처리가 이루어짐)을 제시하였다(Bargh, 1989, pp. 4-5).

인지 심리학 분야의 직관주의 연구자들은 이러한 직관적 인지가 일상의 사회적 정보 수용 및 처리 과정 전반에 있어 다양한 방식으로 영향을 미친다고 강조하였다. 이와 관련하여, 호거스(R. Hogarth)는 자동적 정보 처리의 범주를 기본적 무의식(basic unconscious), 원초적 무의식(primitive unconscious), 정교한 무의식(sophisticated unconscious)의 세 가지 차원으로 분류하였다. 기본적인 무의식은 삶을 유지하고 조절하기 위한 본능적 처리를 담당한다. 예를 들어, 혈당이 떨어지면서 배고픔을 느끼게 되어 먹을 것을 찾게 되는 것을 들 수 있다. 원초적 무의식은 해석이 거의 없이 일어나는 기본적인 정보 처리과정과 관련되는데, 기계적으로 등록되는 사건의 빈도와 변동에서부터 체계의 암묵적 규칙을 추론하는 것에 이르기까지 환경 자극의 상징적인 처리과정을 포함한다. 정교한 무의식은 의미, 정서, 환경 등 지각 과정에 영향을 미치는 정보들에 주의를 기울이는 것과 관련된다. 이는 우리가 종종 이해(understanding)라고 부르는 것에 해당하며 암묵적이지만 단순히 변수들을 연결하는 것을 넘어선 인지적 상태를 의미한다(Lapsley & Power, 2005/ 2008, pp. 262-263).

바흐(J. Bargh) 또한 사회적 정보의 수용 및 처리 과정에 있어 직관적 인지가 영향을 미치는 방식과 관련하여, 사회적 자극에 대한 자동적 인지 처리 과정을 전의식적 자동처리 과정과 후의식적 자동처리 과정 그리고 목표 의존적인 자동처리 과정으로 분류하였다. 전의식적 자동처리

과정이란 지각자 스스로가 전혀 의식하지 못하여도 관련 정보에 만성적으로 접근 가능한 경로가 있어서 주의를 기울이지 않아도 활성화 이루어지는 과정이다(Bargh, 1989, pp. 11). 이는 의식적인 자각 밖의 스키마(schema), 스크립트(script), 계획, 고정관념, 원형(prototype)과 같은 사회적 구성요인(social constructs)이 무의식적으로 활성화되어 나타나는 것이다(Lapsley & Power, 2005/ 2008, p. 257). 사회적 정보에 대한 만성적 접근이 용이한 정보들이 활성화가 되는 예로는 피부색이나 성별에 따라 만성적으로 활성화되는 다양한 종류의 고정관념 등이 있다. 또한 우리의 많은 해석과 평가는 의도적인 인지적 노력 없이 무의식적으로 이루어지며 이는 판단과 결정의 중요한 요소로 작용하기 때문에, 이러한 사회적 구성요인들의 전의식적 활성화는 사회적 정보 처리 과정 전반에 광범위한 영향을 미칠 수 있으며, 사회적 판단에 대한 강한 확신 또는 신념을 유발할 수 있다(Lapsley & Power, 2005/ 2008, p. 257).

후의식적 자동처리 과정이란 최근 의식한 어떠한 사건에 의해 후속적인 처리과정이 영향을 받게 되는 것으로(Bargh, 1989, pp. 14), 의식적인 자각 또는 주의가 일어난 후 의식 밖에서 자동적 인지작용이 일어나는 것이다. 피부색이나 성별에 따른 편견에 대한 교육을 받은 후 마주한 후속 상황에서 이러한 교육 내용이 점화 자극(priming effect)으로 이어져 마주한 상황에서의 정보 처리 과정에 확산되는 현상이 그 예이다. 전의식적 자동처리 과정과 후의식적 자동처리 과정은 만성적으로 접근이 용이한 정보의 틀을 지니고 있느냐 여부에 의해 구분이 가능하다.

목표의존적 자동처리 과정은 의도되지 않은 과정과 의도 된 과정으로 나눌 수 있다. 의도되지 않은 자동 처리란 의도적 행위의 부산물로 획득되는 자동처리 과정으로(Bargh, 1989, pp. 19-20), 피험자들이 문자열에 대한 암기를 요청받았을 때 문자열 뿐 아니라 자극의 반복되는 패턴까지 학습을 하게 되는 암묵적인 학습현상이 그 예이다. 의도 된 자동처리란 능숙하고 전문적인 숙련된 행위의 결과로서, 의식적으로 주의를

기울이거나 통제하는 노력 없이 자동적으로 수행되는 과정이다. 이는 오랜 시간동안의 의식적인 주의를 수반하는 빈번한 수행과 적용과정을 필요로 하며, 도덕적 행동에 관련된 통합적 통찰력의 원천이다. 도식화된 대인관계 관련 절차의 숙지 등이 그 예이다(Narvaez & Lapsley, 2005, p. 147).

이렇듯 인지 심리학적 개념 범주로서의 직관과 추론 즉, 인지 현상을 설명하기 위한 인지 유형에 대한 연구는 사회적 직관주의 연구자들이 도덕적 판단의 현상을 규명하는데 있어 큰 영향을 미쳤다. 도덕적 판단의 현상을 규명함에 있어 인지의 내용이 아닌 형식에 초점을 맞추게 되었으며, 직관적 도덕 판단에 대한 강조가 이루어지게 되었다. 도덕적 판단의 기제를 규명하기 위한 도덕 심리학 분야의 직관주의 연구인 사회적 직관주의 이론들에 대해 본격적으로 살펴보기에 앞서, 다음 절에서는 사회적 직관주의 이론의 형성에 직접적으로 영향을 준 신경과학 및 인지 심리학 분야의 연구들을 검토하고자 한다.

2. 사회적 직관주의의 이론적 기초

상기한 바와 같이 직관주의 이론들은 인지 과정에 감정적 요소가 많은 영향을 미치며, 이성과 감정이 결합된 인지의 형태로서 직관적 판단이 일상의 판단에서 높은 비중을 차지한다고 주장한다. 도덕적 판단에 관한 사회적 직관주의 이론들을 검토하고 분석하기에 앞서, 본 절에서는 사회적 직관주의 이론의 형성에 영향을 준 신경과학 및 심리학 연구들을 검토하고자 한다.

1) 신경과학적 기초

도덕 기능에 관한 신경과학 연구들은 신체 반응이 개인의 도덕 기능에 미치는 영향을 다양한 방식으로 해석하였다. 도덕적 판단에 관한 신경과학 연구는 접근 방식에 따라, 뇌병변 환자의 손상된 뇌 부위와 상실된 사회적 기능의 인과관계를 규명하고자 한 연구와 정상인의 뇌 부위와 도덕 기능의 관련성을 규명하고자 한 연구로 나눌 수 있다.

도덕성과 도덕적 행동이 신경생물학적 기반과 관련이 있다는 최초의 연구로는 전자의 범주 즉, 손상된 뇌 부위와 상실된 사회적 기능의 인과관계 규명 연구에 해당하는 피니어스 게이지(P. Gage)의 두개골 이미지 복원을 통한 다마지오의 연구(Damosio, Grabowski, Frank, Galaburda, & Damasio, 1994)가 있다. 게이지는 사고로 인하여 뇌 손상을 입었으나 회복 후 논리적 추론 능력은 온전하였고 IQ에도 결함이 없었고 기초적인 인지 능력에 장애를 보이지도 않았으며 사회적 지식의 수준 또한 뇌 손상 전에 비하여 떨어지지 않았다. 그러나 그는 정상적인 감정 반응을 느끼지 못하였으며, 실제 생활에 들어가 무엇인가 결정해야 하는 상황이 발생하면 결정을 내리지 못하였고 무책임하고 부적합한 사회적 행동들을 보였다.

다마지오는 사후 부검이 이루어지지 않았던 게이지의 뇌를 현대 신경

영상법을 통해 복원하였고 게이지가 복내측전전두피질(ventromedial prefrontal cortex, 이하 vmPFC)에 손상을 입었으며 이러한 경우, 감정을 처리하거나 적합한 사회적 행동을 하는데 어려움이 있다고 지적하였다. 이를 통해 다마지오는 일상의 판단에는 반드시 감정 및 신체의 반응이 필요하다고 주장하였다.

보다 최근에 수행된 연구(Dimitrov, Phipps, Zahn, & Grafman, 1999) 또한 복내측전전두피질(vmPFC)과 안와전두피질(orbitofrontal cortex, 이하 OFC)의 손상이 사회적, 도덕적 행동의 결핍을 초래한다고 보고한바 있다. 카밀리에(Camille et al., 2004)는 도박 실험을 통해, 안와전두피질(OFC) 병변 환자들이 부정적 결과에 대한 예측 능력이 결핍되고 후회를 느끼지 못한다는 사실을 관찰하였으며, 쾨닉스(Koenigs, 2007)는 복내측전전두피질(vmPFC)이 손상된 환자들이 일반인에 비하여 감정적 사진에 대해 축소된 반응을 보이며 공감과 죄책감을 경험하는 정도가 감소하였고 딜레마 상황에 있어 보다 공리주의적인 판단을 내리는 경향을 보인다고 보고하였다.

정상인의 뇌 부위와 도덕 기능 간의 관련성을 규명한 연구들은 기능성 자기공명영상(functional magnetic resonance imaging, 이하 fMRI)을 활용하였다. 이러한 연구들은 간단한 도덕적 결정을 활용한 연구(Moll, Eslinger, & Oliveira-Souza, 2001; Moll, Oliveira-Souza, Bramati, & Grafman, 2002; Moll, Oliveira-Souza, Bramati et al., 2002; Heekeren, Wartenburger, Schmidt, Schwintowski, & Villringer, 2003; Heekeren et al., 2005; Luo et al., 2006)와 복잡한 딜레마 상황에서의 도덕적 추론을 활용한 연구(Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Borg, Hynes, Van, Grafton, & Sinnott-Armstrong, 2006)로 나눌 수 있다(Prehn, Heekeren, 2009, p. 137).

대표적인 예를 살펴보면, 몰과 동료들(Moll, Oliveira-Souza, Bramati, et al., 2002)은 피험자들에게 도덕적 위반이 나타난 불쾌한

장면을 보여주었고 그들에게서 안와전두피질(OFC)과 내측전두피질(medial prefrontal cortex, 이하 MPC) 및 후측상측두구(posterior superior temporal sulcus, 이하 PSTS)의 활성화를 관찰하였다. 또한 그린은 사람들의 감정적 반응의 활성화 정도가 도덕적 판단에 미치는 영향을 연구하였다. 그린은 직접적으로 사람을 대상으로 하는 '인신적 도덕(Moral Personal) 상황'과 직접적으로 사람을 대상으로 하지 않는 '비인신적 도덕(Moral Impersonal) 상황'에서 감정의 활성화 정도에 차이가 나타날 것이라고 가정하였다. 실험 결과 실제로 인신적 딜레마 상황에서는 정서 반응과 관련된 후위대상회피질(posterior cingulate cortex), 내측전전두피질(medial prefrontal cortex) 등이 활성화 되었고, 비인신적 딜레마 상황에서는 고차적 인지와 연관된 두뇌영역인 배외측전전두엽피질(dorsolateral prefrontal cortex, 이하 DLPFC)과 하측두정엽(inferior parietal lobe)이 활성화 되었다(Greene, 2001, p. 2106; Greene, 2015, p. 147).

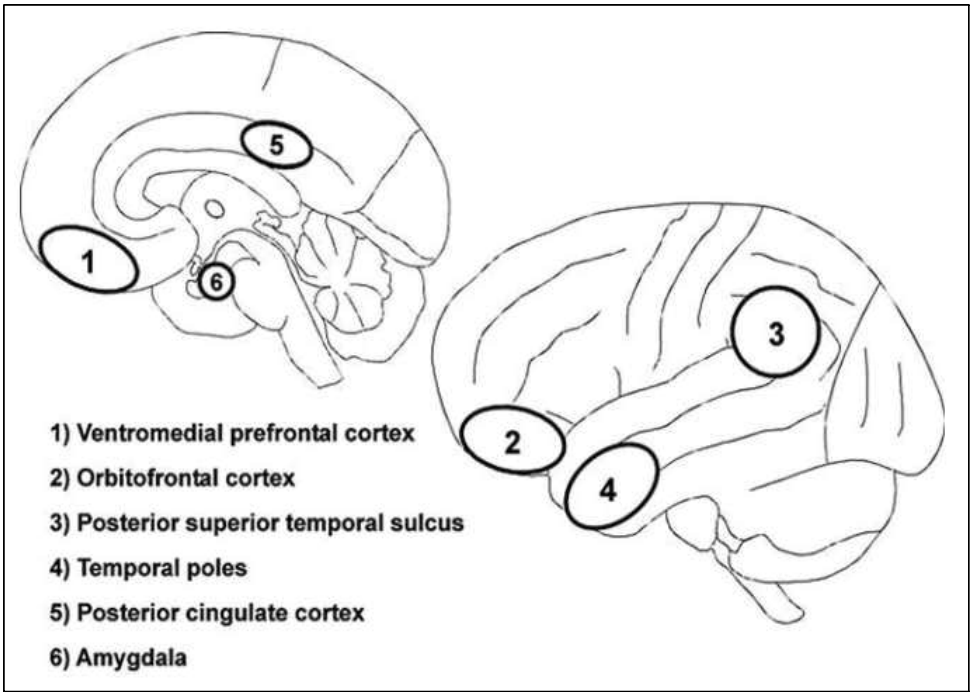
프렌(K. Prehn)과 히케렌(H. Heekeren)은 다양한 연구들에 대한 분석을 통해 뇌 부위와 도덕 기능 간의 관련성을 <표-2> 및 <그림-1>과 같이 정리하였다.

<표-2> 도덕적 판단 관련 뇌 영역별 기능 개요 (Prehn & Heekeren, 2009, p. 139)

뇌 부위	기능
복내측전전두피질 (vmPFC)	- 타인 행위에 내재된 의도 및 심리 상태 파악 - 도덕적 가치 및 사회 규범 위반 등과 관련되는 연민, 수치심, 죄책감과 같은 사회적 정서 생성
안와전두피질 (OFC)	- 행동에 수반될 보상 또는 처벌의 기대 값 감지

후측상측두구 (PSTS)	- 다양한 감각적 신호를 통합하거나 생물학적 운동 신호를 처리하거나 타인 행위의 의도를 감지해내는 등 다양한 양상의 정보들로부터 사회적으로 의미 있는 정보를 감지
측두극 (Temporal poles)	- 삽화적 기억과 자전적 기억
후위대상회피질 (Posterior cingulate cortex)	- 어휘나 물건 및 이미지의 감정적 의미 처리
편도체(Amygdala)	- 공포에 질린 표정이나 화난 표정과 같은 부정적 자극과 위험 신호 차원의 감정 처리

<그림-1> 도덕적 판단 관련 뇌 영역 (Prehn & Heekeren, 2009, p. 134)



덧붙여 프렌(K. Prehn)과 히케렌(H. Heekeren)은 도덕 기능에 관한 신경과학 연구들은 상호연결 네트워크 모델을 지지하고 있기에 도덕 기능을 하나의 뇌 부위와 직접적으로 관련시키는 것은 무리가 있지만, 감정적 반응이 도덕적 판단에 미치는 영향을 보다 사실적으로 규명하기 위해서는 도덕적 판단의 과정에서 도출되는 신경 활동의 변화를 규명하는 작업이 필수적으로 요청된다고 지적하였다(Prehn & Heekeren, 2009, p. 140).

즉, 신경과학적 관찰만을 통해 도덕 기능의 사실적 기제를 규명할 수는 없지만, 그러한 연구들은 도덕 기능의 기제 규명을 목적으로 하는 다양한 이론이 갖는 규명상의 한계와 개선 방향을 명확하게 도출하는데 있어 중요한 역할을 수행한다. 실제, 다양한 신경과학적 연구들은 도덕 기능에 있어, 이성 또는 감정이 독립적으로 작용할 수 없다는 것을 보여줌으로써 인지발달 이론이 갖는 규명상의 한계를 명확하게 도출하는데 기여하였으며, 사회적 직관주의 이론들의 개념 범주 설정 등에 있어 중요한 자료로 기능하며 이론들의 형성에 기여하였다.

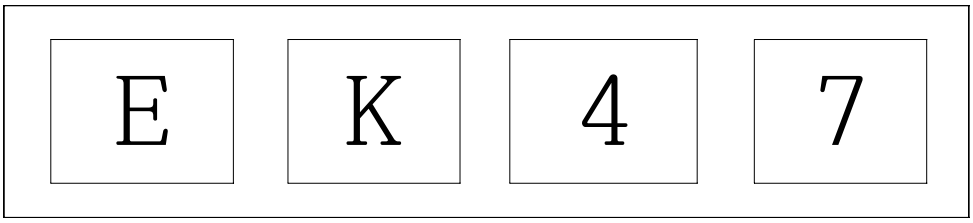
2) 심리학적 기초

직관적 인지의 기제를 규명하고자 한 다양한 인지 심리학 연구들 또한 사회적 직관주의 이론의 형성에 많은 영향을 주었다. 특히, 확증편향(confirmation bias)에 관한 연구들은 사회적 직관주의 이론의 형성에 직접적으로 영향을 미쳤다. 확증편향에 관한 대표적인 연구로는 마골리스(H. Margolis)의 연구가 있다. 마골리스는 '웨이슨(P. Wason)의 네 장의 카드 과제' 실험을 통해 개인의 판단이 객관적 사실에 근거한 추론과 동떨어져 있는 이유를 밝히고자 하였다.

'웨이슨의 네 장의 카드 과제'에서 피험자는 탁자에 놓인 네 장의 카드를 보게 된다. 네 장의 카드는 모두 한쪽에는 알파벳이 쓰여 있고 한쪽에는 숫자가 쓰여 있다. 피험자는 <그림-2>처럼 놓여있는 카드 중 최소한의 카드

를 뒤집어서 ‘카드의 한 면에 모음이 적혀있으면 그 뒷면에는 짝수가 적혀있다.’는 규칙이 지켜지고 있는지를 확인하도록 요청받는다. 논리적으로 추론한다면 E와 7을 뒤집어 보는 것이 맞음에도 많은 사람들은 E와 4를 뒤집어 보았다.[4] 또한 피험자들은 정답을 제시하는 경우와 오답을 제시하는 경우 모두 자신의 답을 얼마든지 설명하며, 추론에 자신감 있는 태도를 보였다(Haidt, 2012/2014, pp. 95-96).

<그림-2> 웨이슨의 네 장의 카드 과제(Haidt, 2012/2014, p. 96)



이러한 연구 결과를 토대로 웨이슨은 판단과 정당화가 별개의 과정이라는 결론에 이르렀다. 마골리스 또한 이러한 상황에 대하여, ‘인간은 판단이 내려지면, 그 근거를 만들어내고 그러한 근거들이 자신이 내린 판단을 도출한 논리적 과정에 대한 설명이라고 믿는다. 그러나 그 근거라는 것들은 사실 자신의 판단에 대한 사후 합리화에 지나지 않는다.’고 해석하였다(Haidt, 2012/2014, p. 97). 즉, 마골리스에 따르면 개인이 판단을 내릴 때 머릿속에는 두 가지의 전혀 다른 인지과정이 작동하며, 이는 ‘보이는 그대로의 인지과정’과 ‘이유를 찾아내는 인지과정’으로 구분된다.

쿤(D. Kuhn) 또한 실험을 통해 이러한 확증편향(confirmation bias)의 모습을 보여주었다. 그는 여덟 장짜리 카드를 여러 벌 만들었으며, 각 카

4) 숫자 4 카드를 뒤집을 때 뒷면에 무엇이 나오든 규칙이 틀리다는 것을 증명하지 못하지만, 숫자 7 카드를 뒤집을 때 뒷면에 모음이 적혀있으면 규칙이 틀리다는 것이 증명된다.

드에는 어떤 아이가 초콜릿 케이크 또는 당근 케이크를 먹고 난 후 어떤 일을 겪는지를 보여주는 만화(예를 들어, 음식을 먹은 후 아이가 웃고 있거나 또는 어디가 아픈 듯 인상을 쓰는 장면)가 그려져 있다. 쿤은 아동과 어른을 대상으로 한 번에 한 벌씩 카드를 보여주었고, 이 '증거(여덟 장의 카드)'를 보고 어떤 음식이 아이를 탈나게 만들었는지를 물어보았다. 여덟 장의 카드 중에는 당근 케이크가 아이를 탈나게 만들었다는 것을 보여주는 카드가 초콜릿 케이크가 아이를 탈나게 만들었다고 보여주는 카드보다 더 많았다. 그럼에도 피험자로 참가한 어른과 아이 모두 직감적으로 초콜릿 케이크를 배탈의 원인으로 지목하고 한두 장의 카드를 증거로 내세우며 자신의 판단이 증명되었다고 결론을 내렸다. 또한 당근 케이크가 배탈의 원인이라는 것을 보여주는 카드가 더 많았음에도 이러한 증거를 무시하는 경향을 보였다(Haidt, 2012/2014, pp. 161-162).

사회적 직관주의 이론들에 영향을 미친 다른 심리학 연구들로 사회적 판단과 행동 및 태도가 암묵적 인지와 고정관념 및 상황 요소에 크게 영향을 받기에 그러한 판단과 행동 및 태도가 인지 왜곡과 편견의 위험에 노출되어 있다는 연구들(Dasgupta, 2009; Gilovich, Griffin, & Kahneman, 2002; Greenwald & Banaji, 1995; Hassin, Ulemanm, & Bargh, 2006)이 있다[5]. 이러한 연구들은, 많은 경우에 있어 도덕적

5) 다스굽타(N. Dasgupta)는 외집단에 대한 암묵적 차원의 편견과 고정관념이 대상 집단에 대한 개인의 태도와 신념 및 행위의 형성과 발현에 영향을 미친다는 것을 강조하며, 인지적 조절 과정에 의해 이러한 암묵적 편견과 고정관념의 영향을 일시적 또는 만성적으로 변화시킬 수 있다고 주장하였다. 그는 인지적 조절에 영향을 미치는 요인으로 내집단의 성원권에 대한 강조 정도, 대상 집단에 대한 정보에의 노출 정도, 판단자의 감정 상태, 내집단에 대한 존중 정도 등을 제시하였다. 길로비치(T. Gilovich)와 그리핀(D. Griffin) 그리고 카네만(D. Kahneman)은 합리적인 도덕적 판단자의 추론 능력 또한 암묵적 정보처리에 종속된다는 것을 보여주는 다양한 연구를 제시하였다. 그린월드(G. Greenwald)와 바나지(R. Banaji)는 판단자가 자신의 판단에 대해 스스로 검토하고 살펴봄으로써 도덕 판단에 대한 암묵적인 편견의 영향력을 감소시킬 수 있으며, 이러한 인지 과정을 통해 의도하지 않은 차별을 예방할 수 있다고 주장하였다. 하신(R. Hassin)과 울레만(J. Uleman) 그리고 바흐(J. Bargh) 또한 도덕적 판단과 사회적 행위에 있어 사소한 암묵적 자동처리 과정이 미치는 다양한 영향에 관한 연구들을 제시하였다.

판단이 감정적 요소와 암묵적 지식 및 상황 요인에 영향을 받고 추론이 그를 확증하는 형태로 기능한다는 것을 지지함으로써 사회적 직관주의 이론들의 형성에 많은 영향을 주었다.

Ⅲ. 사회적 직관주의 이론의 구조

대표적인 사회적 직관주의 이론으로는 하이트(J. Haidt)의 사회적 직관주의 모델(Social Intuitionist Model)과 그린(J. Greene)의 이중과정 모델(Dual Process Model)이 있다. 두 이론은 모두 감정적 특성을 강하게 갖는 도덕적 직관이 도덕적 판단에 강력한 영향을 미친다고 주장하였다. 다만, 하이트가 도덕적 판단은 주로 직관적 형태로 도출되며 추론의 역할은 직관적 도덕 판단에 대한 합리화에 국한된다고 주장한 것과 달리, 그린은 추론이 도덕적 판단에 직접 기능하며 직관과 경쟁할 수 있다고 주장하였다. 이렇듯 두 이론은 핵심적인 부분에 있어 충돌하였고, 이로 인하여 두 사회적 직관주의 이론은 합의점을 도출하는데 실패하였다.

본 장에서는, 도덕적 판단에 관한 두 사회적 직관주의 이론을 비판적으로 검토하기에 앞서, 사회적 직관주의 이론의 구조와 특성을 세부적으로 검토하고자 한다. 두 사회적 직관주의 이론이 규명하는 도덕적 판단의 기제를 검토하기에 앞서, 인지 심리학 분야의 용어 사용으로 인한 혼선을 줄이기 위해 재차 강조하여야 하는 것은 추론의 개념에 대한 정의이다. 상기한 바와 같이, 인지발달 이론은 이성과 감정이 분리되어 작용한다는 관점을 전제로, ‘이성에 의한’ 추론을 강조하였다. 그러나 이와 달리 사회적 직관주의 이론은 도덕적 판단에 있어 이성과 정서가 분리되어 작용하지 않는다는 관점을 전제하며, 인지(cognition)의 형태로서 추론(reasoning)과 직관(intuition)을 구분한다. 즉, 사회적 직관주의 이론들은 추론 또한 감정의 영향을 받는다는 관점을 전제한다.

1. 하이트의 사회적 직관주의 모델의 구조

1) 도덕 범주 인식의 문화 종속성

사회적 직관주의 모델을 검토함에 있어 한 가지 주의해야 할 것은, 인지주의 접근이 도덕에 대한 정당화를 전제로 피해와 공정성이라는 도구를 통해 인습적 영역과 구분되는 도덕의 영역을 정의하고 논의를 진행한 것과 달리 사회적 직관주의 모델은 도덕에 대한 정당화 과정 없이 도덕의 영역을 '개인들이 도덕의 영역으로 받아들이는 무언가'로 열어놓고 논의를 진행하였다는 점이다. 즉, 사회적 직관주의 모델이 다루는 도덕적 판단이란 서구적 문화의 관점에서 개인적 영역 및 인습적 영역과 구분되는 도덕적 영역에 있어서의 판단만을 의미하는 것이 아니다. 하이트가 도덕적 직관, 도덕적 추론 등에 있어 사용하는 '도덕적' 이라는 용어는 '사람들이 도덕으로 인지하는 무언가'를 의미한다.

이러한 관점은 슈웨더(R. Shweder)의 연구(Shweder, Mahapatra, Miller, 1987)에 영향을 받은 바가 크다. 슈웨더는 개인 중심적 문화를 공유하는 미국인과 사회 중심적 문화를 공유하는 인도 오리사(Orisa) 지역 주민의 사고방식에 큰 차이가 있음을 발견하였다. 그는 콜버그(L. Kohlberg)의 이론과 그에 기반한 튜리엘(E. Turiel)의 이론이 개인주의 문화 속에 사는 사람들에 의해 그리고 그들을 위해 만들어진 것이라고 지적하며, 문화적 차이가 자아 관념에 대한 차이를 유발하고 이는 도덕에 대한 인식의 차이를 유발할 것이라 가정하였다.

이러한 가정 하에 슈웨더는 사회 중심적 문화를 공유하는 오리사 주민들이 개인주의적 문화를 공유하는 미국인들처럼 도덕적 규칙(타인에게 해가 가지 않도록 규제하거나 공평성을 지키는 것)과 사회적 규약(피해와 직접 관련이 없는 행동을 규제하는 것)을 구분하는지 여부를 알아보기 위한 인터뷰를 진행한다. 인터뷰 질문은 미국 또는 오리사에서 규칙 위반으로 비칠 수 있는 39개의 문항으로 구성되었으며, 인터뷰는 미국의

아동과 오리사의 브라만 계층의 아동 각 180명, 미국의 성인과 오리사의 브라만 계층의 성인 각 60명, 오리사의 하층민 120명을 대상으로 이루어졌다. 질문은 누군가에게 분명하게 피해가 가는 상황과 타인에게 해가 되거나 불공평할 것이 없는 질문으로 구성되었다. 누군가에게 피해가 가는 상황은 피험자의 반응을 기준으로 두 범주로 나눌 수 있다. 한 범주는 미국인과 오리사 주민 모두가 비난하는 상황들이며, 다른 한 범주는 미국인은 비난하지만 오리사 주민은 비난하지 않는 상황들이다. 누군가에게 해가 되거나 불공평할 것이 없는 상황에 대해서는 미국인들이 충분히 용인할 수 있다는 태도를 보인 반면, 오리사 주민들은 비난적인 태도를 보였다(Haidt, 2012/2014, pp. 49-55).

[미국인과 오리사 주민 모두가 잘못이라 말한 행동]

· 한 남자가 길을 걷고 있다. 남자는 길 위에 개 한 마리가 누워 자는 것을 보고는 개에게 다가가 발로 걷어찬다.
· 아버지가 아들에게 말했다. “이번 시험을 잘 보면 만년필을 사 주마.” 아들은 시험에서 좋은 성적을 거뒀지만, 아버지는 아들에게 아무것도 해주지 않았다.

[미국인은 잘못이라고 했지만, 오리사 주민은 괜찮다고 한 행동]

· 젊은 아내가 남편에게 이야기하지 않은 채 혼자 영화를 보러 갔다. 집에 돌아온 아내를 보고 남편이 말했다. “또 한 번 그랬다간 흠씬 두들겨 맞을 줄 알아.” 아내는 또 혼자서 영화를 보러 갔다. 남편은 아내를 흠씬 두들겨 팼다. (남편의 행동에 대한 평가를 요청함)
· 슬하에 결혼한 아들과 딸을 하나씩 둔 남자가 있었다. 남자가 세상을 떠나자 아들이 아버지의 재산을 대

부분 차지했다. 딸은 거의 아무것도 받지 못했다. (아들의 행동에 대한 평가를 요청함)

[오리사 주민은 잘못이라고 했지만, 미국인은 괜찮다고 한 행동]

· 스물다섯 살 된 아들이 집 안에서 아버지를 이름으로 부른다.
· 어떤 여자가 밥을 지었다. 그녀는 남편 및 오빠와 한 자리에서 밥을 먹고 싶었다. 그래서 그들과 한 상에 앉아 밥을 먹었다. (여자의 행동에 대한 평가를 요청함)
· 나와 한동네에 사는 과부가 일주일에 물고기를 두세 번은 먹는다.
· 한 여자가 대변을 보고 난 후 옷을 갈아입지 않은 채로 요리를 했다.

이러한 행동들 중 오리사 주민은 잘못이라고 했지만 미국인은 괜찮다고 한 행동들은 누군가에게 피해를 주지도 않고 공평성에 어긋나지도 않는다. 즉, 튜리엘의 영역 구분 방식에 따랐을 때 이러한 행동들에 대한 오리사 주민들의 비난은 도덕적 위반에 대한 비난이 아닌 사회적 규약을 어긴 것에 대한 비난이 되어야 한다. 그러나 오리사 주민들은 그러한 행동들이 언제 어디에서나 도덕적으로 잘못된 행동이라고 하였으며, 그들에게 있어 음식과 의복, 성생활과 이성간 관계에 관련된 인도의 관습들은 대부분 사회적 규약이 아닌 도덕적 규칙을 적용받는 쟁점으로 인지되었다. 슈웨더는 이러한 현상에 대하여, 인도인들은 서구의 개인 중심적 문화의 사회 구성원들처럼 피해와 공정성의 차원에서 도덕의 영역을 인지하고 도덕적 판단을 내리지 않으며, 사회 중심적 문화를 받아들임으로써 그에 근거하여 도덕의 범주와 기준에 대한 관점을 정립한다고 주장하였다. 즉, 슈웨더에 따르면 개인 중심 혹은 사회 중심과 같은 문화적 특

성의 차이는 도덕의 범주나 기준에 대한 인식의 차이를 유발하며, ‘피해와 공정성에 기반하여 도덕적 판단의 원리를 도출하는 것’이나 ‘인습적 영역과 도덕적 영역으로 사회적 영역을 구분하는 것’은 인류 보편적인 것이 아니라 개인 중심적인 서구 문화의 산물이다(Haidt, 2012/2014, pp. 49-55).

그러나 미국인이 문제가 없다고 했지만 오리사 지역 주민들은 도덕적으로 옳지 않다고 판단한 문항들과 관련하여, 튜리엘은 이러한 문항들이 미국인의 관점에서 파악하기 힘든 숨겨진 피해와 공정성의 기준을 내포하고 있다고 비판하며, 피해와 공정성의 기준을 적용하여 도덕의 영역을 구분하는 현상은 보편적인 현상이라고 주장하였다(Haidt, 2012/2014, p. 55). 이에 하이트는 과연 튜리엘의 주장처럼 모든 문화에 있어 피해와 공정성의 기준에 의해 도덕의 영역과 인습의 영역이 구분되는 것인지, 아니면 슈웨더의 주장처럼 그러한 영역 구분이 문화적 특성에 종속되는 것인지에 대하여 규명하고자 하였다(Haidt, 2012/2014, pp. 59-60).

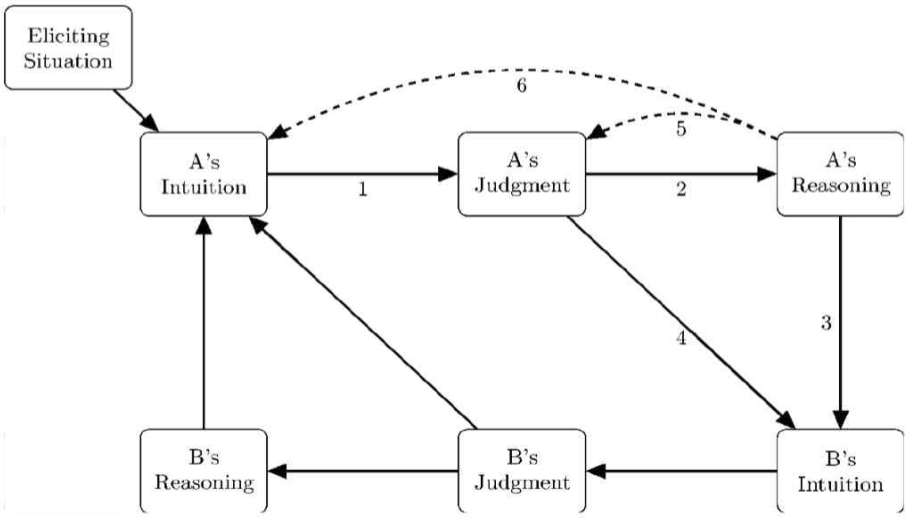
그는 미국 필라델피아와 브라질 지역의 주민들을 대상으로 추가적인 연구를 진행하였으며, 브라질의 경우, 비교적 산업화되고 개발된 남부 도시인 포르투 알그레 지역과 비교적 덜 개발된 북부 도시인 헤시피 지역을 나누어 연구를 진행하였다. 또한 그는 브라질의 빈부격차가 심각하다는 점에 착안하여 각 도시별로 사회 계층 기준을 추가하였고 연령 기준 또한 추가하였다. 즉, 하이트는 세 개의 도시에서 각 도시마다 두 개의 사회계층을 구분하고 각 사회계층마다 두 개의 연령 집단을 구성하였다(Haidt, 2012/2014, p. 61).

그는 각각의 연구 집단이 주어진 상황을 어떠한 영역의 문제로 인식하는지에 대하여 살펴보고자, 튜리엘이 도덕의 영역과 인습의 영역을 구분하기 위해 사용한 두 문항을 활용하였다. 한 이야기는 여자아이가 그네에서 남자아이를 밀어 떨어뜨리는 이야기이며, 다른 하나의 이야기는 교복을 입지 않고 학교에 가는 남자아이 이야기이다. 연구 결과, 필라델피

아의 피험자 네 그룹은 모두 튜리엘의 주장과 같이 도덕 위반과 인습 위반을 확실하게 구분하였다. 그러나 브라질에서는 교복 미착용과 같은 무해한 금기 위반과 관련하여, 남부의 포르투 알그레 지역에서 더 많은 피험자가 이를 도덕적 위반으로 판단하였고, 사회계층이 낮을수록 그리고 연령이 어릴수록 더 많은 피험자가 이를 도덕적 위반으로 판단하였다. 또한 이러한 변수들 중 가장 큰 영향을 미치는 것은 사회 계층이었다. 이에 기반하여 하이트는 도덕 영역에 대한 구분과 도덕적 판단의 기준 설정이 문화집단에 영향을 받는다고 주장하였다(Haidt, 2012/2014, pp. 61-63).

이에 하이트는 피해와 공정성의 기준이 적용되는 서구 전통의 도덕적 영역에 논의의 범주를 한정하지 않았으며 '개인들이 도덕의 언어로 표현하는 무언가'를 도덕의 영역으로 설정하고 열린 논의를 진행하고자 하였다. 이를 전제로 하이트는 <그림-3>과 같이 사회적 직관주의 모델을 도출하였다.

<그림-3> 도덕적 판단에 관한 사회적 직관주의 모델(Haidt, 2001, p. 815)



사회적 직관주의 모델이 제시하는 도덕적 판단의 기제는 다음과 같다. (1)은 '직관적 판단(the intuitive judgment link)'으로서 아무런 노력 없이 즉각적이고 자동적으로 판단이 일어나는 과정이다. (2)는 '사후 추론(the post-hoc reasoning link)'으로서 직관적 판단을 합리화하고 지지하기 위한 근거를 모색하는 추론의 과정이다. 하이트는 추론의 기능을 주로 이러한 직관적 판단에 대한 합리화에 한정하였다. (3)은 '추론된 설득(the reasoned persuasion link)'으로서 직관적 도덕 판단에 대한 사후 추론을 언어적으로 전달함으로써 타인에게 영향을 주는 과정이다. (4)는 '사회적 설득(the social persuasion link)'으로 판단 도출의 과정이나 근거 등에 대한 추론적인 제시 없이 직관적 도덕 판단에 대한 제시만으로 타인의 직관에 영향을 주는 과정이다. (5)는 '추론된 판단(the reasoned judgment link)'으로서 스스로의 추론을 통해 초기 직관에 의해 도출된 판단을 직접적으로 수정하는 것이다. (6)은 '사적인 반성적 사고(the private reflection link)'로서 추론을 통해 초기 직관 그 자체가 영향을 받아 수정되는 것이다.

2) 직관적 도덕 판단과 사후 추론의 기제

하이트는 도덕적 판단은 신속하고 자동적인 직관의 결과물이며, 직관적 판단이 이루어진 후에 느리고 의식적인 도덕적 추론이 발생 한다고 주장한다(Haidt, 2001, p. 818). 즉, 사회적 직관주의 모델의 핵심 아이디어는 첫째, 개인의 도덕적 판단이 감정적 특성을 강하게 갖는 즉각적이고 자동적인 직관의 형태로 도출된다는 것이고 둘째, 의식적인 추론은 판단의 과정에 직접 기능하지 않으며 직관적 판단이 이루어진 후 판단의 합리화에 기능한다는 것이다.

앞서 살펴본 마골리스와 쿤의 연구에 영향을 받아, 하이트는 개인의 추론 능력이 도덕적 판단의 옳고 그름을 논리적으로 검토하고 판단을 정정하는데 사용되기 보다는 최초의 도덕적 판단을 지지하는 그럴싸한 근

거를 만들어내는데 기능한다는 가설을 설정한다. 이를 증명하기 위해 그는 스콧 머피(S. Murphy)와 함께 실험을 진행한다. 그들은 먼저 추론적 도덕 판단의 전형으로 삼을 수 있는 콜버그의 하인츠 딜레마 상황에서의 판단 양식과 직감적인 반응을 유도하는 상황에서의 판단 양식을 살펴보고, 그가 피해와 공정성의 기준이 적용되지는 않지만 개인들이 도덕의 영역으로 받아들이는 판단이 두 판단 중 어떠한 판단의 양식과 닮아 있는지 규명하고자 하였다(Haidt, Björklund, & Murphy, 2000).

그는 우선 전형적인 추론적 도덕 판단 양식의 특징을 살펴보고자, 하인츠 딜레마[6] 상황을 제시하고 인터뷰를 진행하였다. 하이트는 하인츠 딜레마가 피해 요소를 포함하고 있으며, 생명과 재산 또는 생명과 법제도와 같은 가치의 대립 지점이 명확하기 때문에 합리적인 추론에 의한 판단을 도출하기 위한 상황에 충분히 적합하다고 판단하였다. 예상대로 대다수의 피험자들은 자신의 도덕적 판단에 대하여 훌륭한 근거를 제시하였고, 판단의 근거를 찾지 못한 채 판단만을 유지하는 도덕적 당혹감(moral dumbfounding)이 나타나는 비율도 낮았다. 또한 이러한 판단에 대하여 여러 합리적인 근거를 제시하며 피험자의 판단을 수정하려 해보아도 피험자들은 자신이 최초에 선택한 판단을 포기하지 않는 모습을 보였다.

다음으로 그는 직감적인 반응을 유도하는 상황을 피험자에게 제시하고 인터뷰를 진행하였다. 하나는 '바퀴벌레 주스' 시나리오이며 하나는 '영혼 매매 계약' 시나리오이다.

[바퀴벌레 주스 시나리오]

6) 유럽에 거주하고 있는 한 여성은 희귀성 암에 걸려 치료가 어려운 상태이다. 의사의 말에 따르면 그 암을 치료할 수 있는 약은 세상에 단 한 가지뿐이다. 그 약을 개발한 약제사는 자신이 개발에 들인 돈의 10배에 해당하는 금액으로 약의 가격을 책정하였다. 그녀의 남편인 하인츠는 여기저기에서 돈을 빌려보았지만 약 값의 절반 밖에 구하지 못하였다. 하인츠는 약제사에게 할인을 해주거나 나중에 잔액을 내는 방식으로 약을 판매해 줄 것을 요청하였지만, 약제사는 자신이 이 약을 개발하였고 이제 이 약으로 돈을 벌 것이라며 거절하였다. 절박한 마음에 하인츠는 약제사의 가게에 침입하여 그 약을 훔쳤다. 하인츠의 행동은 잘못된 행동인가.

실험자: 당신은 사과 주스를 좋아합니까?
피험자가 예라고 대답한 경우: 좋습니다.
피험자가 아니라고 대답한 경우: 네 그렇다면 여기에 물이 있습니다.
실험자는 피험자의 응답에 맞게 사과주스 또는 물과 냅킨, 컵, 바퀴벌레 상자와 티볼(tea ball)을 가져온다.
실험자: 자, 여기 있는 사과주스(또는 물)를 컵에 따르도록 하겠습니다. 한 모금 드셔보시겠어요?(잠시 기다린다) 자, 여기 상자 안에 살균된 바퀴벌레가 있습니다. 실험실에 납품하는 업체로부터 가져온 바퀴벌레입니다. 이 바퀴벌레는 깨끗한 환경에서 길러졌으며, 우리는 이 바퀴벌레를 다시 멸균기에서 살균하여 어떠한 균도 남아있지 않도록 처리하였습니다. 이제 이 바퀴벌레를 사과주스(또는 물)에 담글 것입니다. 이렇게 말이죠. 자, 이 사과주스(또는 물)를 한 모금 더 드시겠어요?

[영혼 매매 계약 시나리오]
문서에는 '나 ○○○은 총 2달러를 받고 죽은 후 내 영혼을 스콧 머피에게 매도함을 이 문서로써 합의한다 ... 이 문서는 심리학 실험의 일부입니다. 이 계약은 어떤 식으로든 법적 효력이나 구속력을 갖지 않습니다.' 라고 쓰여 있다. 서명 후에 곧바로 종이를 찢어버려도 되며, 그렇게 하더라도 여전히 2달러를 받게 된다. 문서에 서명 하겠는가.

바퀴벌레를 담갔던 주스를 마시겠다고 한 피험자는 전체의 37퍼센트였고 자발적으로 영혼 매매 문서에 서명한 피험자는 전체의 23퍼센트였다(Haidt, Björklund, & Murphy, 2000, p.14). 피험자 중 다수가 요구

에 응하지 않았고, 응하지 않은 이유를 설명하도록 요구 받았을 때 변변한 이유를 제시하지는 못하였다. 스콧 머피의 이의 제기와 설득에 의해 바퀴 주스를 마시려는 피험자는 10퍼센트가 늘어난 47퍼센트가 되었고, 영혼 매매 문서에 서명한 사람은 17퍼센트가 늘어난 40퍼센트가 되었다. 그러나 대부분의 피험자는 두 시나리오 모두에서 최초의 입장을 끝까지 고수하였고 여전히 변변한 이유를 제시하지 못하였다(Haidt, Björklund, & Murphy, 2000, p. 10).

하이트는 이어서 불온한 느낌을 주지만 피해자는 없는 상황을 제시하며, 이에 대한 도덕적 판단이 하인츠 딜레마 상황에서의 '추론적 방식'과 바퀴 주스 과제와 영혼 매매 계약 과제에서의 '직감적 방식' 중 어느 방식과 더 유사한 방식으로 도출되는지를 규명하고자 하였다.

> [불온하지만 피해자는 없는 상황 시나리오 1]
> 남매지간인 줄리와 마크는 함께 프랑스를 여행하는 중이다. 둘은 모두 대학생이다. 그러던 어느 날 이들은 바닷가의 한 오두막집에서 둘이서만 밤을 보내게 되었다. 둘은 이참에 잠자리를 함께해보면 신기하기도 하고 재미있을 거라고 결정을 내린다. 최소한 서로에게 새로운 경험은 될 것이다. 줄리는 이미 피임약을 복용한 상태였지만, 만일의 사태에 대비해 마크도 콘돔을 사용한다. 두 사람은 모두 즐겁게 잠자리를 가졌지만, 다시는 그러지 않기로 결심한다. 그리고 그날 밤 일은 둘만 아는 특별한 비밀로 지키기로 한다. 이제 둘은 서로가 훨씬 가깝게 느껴진다. 마크와 줄리가 섹스를 한 것은 잘못인가.

> [불온하지만 피해자는 없는 상황 시나리오 2]
> 제니퍼는 병원의 병리 검사 센터에서 일하고 있다. 제니퍼는 채식주의자이다. 고기를 먹기 위해 일부러 동물을

죽이는 것이 도덕적으로 잘못된 일이라고 생각하기 때문이다. 그러던 어느 날 밤 제니퍼는 시체를 소각하는 일을 맡게 된다. 제니퍼는 식용으로 아무런 문제가 없는 고기를 버리자니 아깝다는 생각이 든다. 그래서 제니퍼는 시체의 살을 한 조각 잘라 집으로 가져온다. 그리고 그것을 요리해서 먹는다. 제니퍼의 행동은 잘못된 행동인가.

이러한 불온한 감정을 유발하지만 피해자가 존재하지는 않는 상황들을 마주한 피험자들은 직감을 통한 판단과 유사한 방식으로 판단을 도출하였다. 줄리와 마크가 섹스를 해도 좋다고 응답한 피험자는 전체의 20%였고 제니퍼가 시체의 일부를 먹어도 좋다고 응답한 피험자는 전체의 13%였다(Haidt, Björklund, & Murphy, 2000, p. 14). 이러한 행동이 잘못된 행동이라고 응답한 피험자들은 근거 제시를 요구받자 여러 가지 근거들을 찾아내려 노력하는 모습을 보였지만, 그럴싸한 근거를 제시하지 못하고 도덕적 당혹감(moral dumbfounding)을 보이는 경우가 많았다(Haidt, Björklund, & Murphy, 2000, p. 10). 또한 실험자가 피험자의 근거들이 타당하지 못하다는 것을 증명해 내어도 피험자가 판단을 바꾸는 일은 10퍼센트 대의 낮은 비율에 머물렀다(Haidt, Björklund, & Murphy, 2000, p. 14).

이러한 실험을 통해 하이트는 첫째, 개인의 도덕적 판단이 직관적 형태로 도출되며 둘째, 추론은 최초의 직관적 판단에 대한 반성적 검토와 수정에 기여하기보다는 판단의 합리화에 기여한다고 주장하였다. 이러한 특징은 사회적 직관주의 모델(<그림-3>)의 (1) '직관적 판단(the intuitive judgment link)'과 (2) '사후 추론(the post-hoc reasoning link)'에 해당한다.

3) 직관적 도덕 판단의 수정 기제

하이트에 따르면 직관적 도덕 판단의 수정은 크게 타인의 설득에 의한 수정과 스스로의 추론에 의한 수정으로 구분할 수 있다. 먼저 타인의 설득에 의한 직관적 도덕 판단의 수정에 대해 먼저 살펴보면, 이는 사회적 직관주의 모델의 (3) ‘추론된 설득(the reasoned persuasion link)’과 (4) ‘사회적 설득(the social persuasion link)’에 해당한다.

추론된 설득이란 도덕적 판단을 정당화하기 위해 언어적으로 표현된 추론에 의해 타인의 직관에 영향을 미치는 것을 의미한다. 하이트는 추론의 주 역할을 직관적 판단의 합리화로 보았기에, 개인이 스스로의 자발적인 추론에 의해 초기 직관에 기반한 판단을 변경하거나 포기하는 일은 매우 드물다고 보았으며, ‘추론된 설득’이 도덕적 판단의 과정에 있어 추론이 직접적으로 기능하는 유일한 경우에 해당한다고 지적하였다.

또한 그는 상대의 정서적 태도를 변화시키기 위해 인지적 설득보다 정서적 설득이 더욱 효과적이라는 에드워드(K. Edwards)와 폰하이펠(W. Von Hippel)의 연구(Edward, Von Hippel, 1995)를 근거로, ‘추론된 설득’이 성공하기 위해서는 논리적인 논쟁이 아닌 정서적인 설득의 특성을 가져야 한다고 강조하였다(Haidt, 2001, p. 819). 그에 따르면 논리적인 추론을 통해 상대와 논쟁을 벌이는 것은 직관적 판단의 변화를 가져오는데 도움이 되지 않는다(Haidt, Björklund, & Murphy, 2000, p. 9). 즉, 하이트에 따르면 추론에 의해 최초의 판단을 수정하도록 설득함에 있어 필요한 것은 추론의 논리적 완결성이 아닌 추론의 감정적(affective) 영향력이다.

하이트는 판단자의 감정적 반응의 정도가 약화된 경우에 추론적 설득의 가능성이 더욱 커진다는 것을 보여주는 팩스턴(J. Paxton)과 그린(J. Greene)의 연구(Paxton, Greene, 2012) 또한 근거로 삼았다(Haidt, 2012/2014, pp. 142-144). 팩스턴은 피험자들에게 하이트가 사용했던 줄리와 마크 남매 이야기를 들려준 후 하이트와 마찬가지로 해당 행위에

대하여 판단을 내리도록 하였다. 그리고 난 후 피험자의 절반에게는 허술하기 짝이 없는 논거(예를 들어, 줄리와 마크가 사랑을 나누면 이 세상에는 더 많은 사랑이 존재하게 된다)를 제시하고 나머지 절반에게는 그럴듯한 논거(예를 들어, 근친상간에 대한 반감은 사실 고대에 일어난 일종의 진화적 반응으로, 피임법이 따로 없던 세계에서 선천적 기형을 막기 위해 사람들이 불러일으킨 정서였다. 그러나 줄리와 마크의 경우에는 둘 다 피임법을 활용했기 때문에 크게 걱정할 것이 없다)를 제시하며 그들의 판단이 수정될 수 있는지를 알아보았다. 언뜻 보기에 피험자들이 논리적으로 허술한 이유보다 논리적으로 그럴싸한 이유에 더 많이 설득되었을 것으로 생각되지만, 사실 두 설득은 효과 측면에 있어 차이가 전혀 나타나지 않았다. 피험자들은 설득의 논거가 그럴듯한가 여부와 관련 없이 설득자의 논거를 반박할 방법을 모색하며 판단을 바꾸지 않았다.

팩스턴은 방식을 조금 바꾸어 방금 자신의 직관적 판단을 유지하던 피험자들을 추려 2분간 응답을 하지 못하도록 하였다. 허술한 논거를 접한 피험자들은 2분의 시간이 흐른 후에도 여전히 줄리와 마크를 비난하였지만 그럴듯한 논거를 접한 피험자들은 2분이 지난 후에는 줄리와 마크의 행동을 훨씬 더 너그러운 태도로 받아들이는 모습을 보였다. 이러한 현상에 대하여 팩스턴은 피험자의 강한 정서적 반응이 가라앉음으로써 확증편향(confirmation bias)이 조금씩 사라졌고 이에 추론이 판단에 영향력을 행사할 수 있게 되어 이러한 변화가 발생한 것이라고 해석하였다.

다음으로, 사회적 설득(the social persuasion link)은 판단을 정당화하기 위한 언어적 추론을 제시하지 않고 직관적 판단의 결과만을 전달함으로써 타인의 직관에 변화를 주는 과정이다. 하이트에 따르면 사람들은 집단 규준에 적응하기 때문에 친구, 동료, 아는 사람들이 어떤 방식으로 사회적 상황에서 판단한다는 사실을 알게 되면 그들의 판단에 영향을 받는다. 이는 타인의 판단을 가치롭게 여기거나 판단의 주체인 타인을 정서적으로 수용함으로써 직관이 수정되는 과정이다(Haidt, 2001, p.

819).

주목할 점은 판단을 직접 전달함으로써 이루어지는 사회적 설득은 근거의 제시를 필요로 하지 않는다는 것이다. 하이트는 이와 관련하여, 사람들이 특정 개인을 가치롭게 여기거나 정서적으로 수용하는 경우, 그 개인의 판단에 영향을 받는다는 하딘과 히긴스(Hardin & Higgins, 1996)의 연구 등을 근거로 제시한다(Haidt, 2001, p. 819). 즉, 타인에 의한 설득의 과정을 종합하면, 타인의 설득에 의한 직관적 도덕 판단의 수정에 있어, 정서적 영향을 주지 못한다면 잘 구조화된 논리적 추론으로도 상대를 설득할 수 없으며, 정서적인 영향을 줄 수만 있다면 단순히 판단을 공유하는 것만으로도 상대를 설득할 수 있다.

다음으로 스스로의 추론에 의해 직관적 판단을 수정하는 경우에 대해 살펴보자면, 이는 사회적 직관주의 모델의 (5) '추론된 판단(the reasoned judgment link)' 과 (6) '사적인 반성적 사고(the private reflection link)'에 해당한다. 먼저 전자에 대하여 살펴보면, 추론된 판단은 스스로의 추론을 통해 초기 직관에 의해 도출된 판단을 직접적으로 수정하는 것이다. 하이트에 따르면, 드물긴 하지만, 사람들은 종종 순수한 논리적 추론에 의해서 그들의 초기 직관에 모순되는 판단을 내릴 수 있다. 즉, 매우 드물게 추론은 직관적 판단을 합리화 하는데 그치지 않고 그 타당성을 검토하고 대안을 모색하며 판단을 이끌어 내는 역할을 한다. 하이트에 따르면 이러한 추론된 판단은 초기 직관이 약하며 처리 능력이 높은 경우에 발생한다. 주의할 점은 추론된 판단은 추론이 직관에 영향을 미쳐 판단에 변화를 가져오는 과정이 아니라, 추론이 직접적으로 판단의 변화를 가져오는 과정이라는 것이다(Haidt, 2001, p. 819).

덧붙여 하이트는 윌슨과 슐러의 연구(Wilson & Schooler, 2000)를 근거로, 추론에 의한 판단이 초기 직관과 강한 갈등을 보이며 상존하는 경우, 사람들은 추론에 의한 판단을 언어적으로 표현하면서 직관적인 판단이 여전히 표면 아래 함께 존재하는 이중적 태도(dual attitude)를 취

할 수 있다고 지적한다(Haidt, 2001, p. 819).

이러한 이중적 태도를 수반하는 추론은 특정 상황에서 도출된 직관적 판단에 반응적으로 도출된 것일 수도 있지만, 상황을 당면하기 이전에 해당 상황과 관련하여 학습된 추론일 수도 있다. 전자의 예를 들자면, 추론 과정을 통해 스스로가 미신적인 금기에 기반하여 직관적 판단을 도출하였다는 것을 인지하고 판단을 수정하지만 여전히 미신에 대한 금기가 마음속에 존재하는 경우 등이 있다. 후자의 예는 2장에서 살펴본 바흐의 자동적 정보처리 과정 분류 틀을 활용함으로써 보다 명료하게 이해할 수 있으며 이는 다음과 같다.

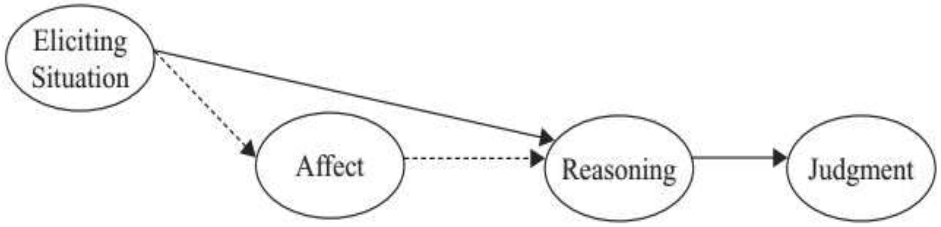
상기한 바와 같이, 바흐는 인지의 자동 처리 과정을 전의식적 자동 처리 과정과 후의식적 자동처리 과정 및 목표의존적 자동처리 과정으로 분류하였다. 전의식적 자동처리 과정이란 지각자 스스로가 전혀 의식하지 못하여도 관련 정보에 만성적으로 접근 가능한 경로가 있어서 주의를 기울이지 않아도 활성화 이루어지는 과정이며, 의식적인 자각 밖의 스키마(schema)와 고정관념 등과 같은 사회적 인지의 구성요인이 무의식적으로 활성화되어 나타나는 것이다. 또한 후의식적 자동처리 과정이란 최근 의식한 어떠한 사건에 의해 후속적인 처리과정이 영향을 받게 되는 것으로, 의식적인 자각 또는 주의가 일어난 후 의식 밖에서 자동적 인지작용이 일어나는 것이다.

후자 즉, 학습된 추론의 도출 및 그에 수반된 이중적 태도란 전의식적 자동처리 과정에 의해 도출된 직관적 판단과 그에 반대되는 후의식적 자동처리 과정에 의해 활성화 된 새로운 직관적 판단이 충돌하는 것으로 볼 수 있다. 예를 들어, 개인이 피부색이나 성별에 따른 편견을 가지고 있으나 그러한 편견에 대한 교육이나 경험을 통해 그러한 편견이 잘못된 것이라고 숙지하는 경우, 마주한 상황에서 전의식적 자동처리 과정으로서 편견에 기반한 직관적 판단이 도출되면서 동시에 후의식적 자동처리 과정의 점화(priming effect)가 이루어지며 반편견적인 추론적 판단을 도출할 수 있다. 즉, 전의식적 자동처리 과정으로서의 스키마가 완전히

수정되지 않은 상태에서 그와 상충하는 후의식적 자동처리 과정의 점화에 의해 감정 반응과 상충하는 추론적 판단을 내리게 되는 것이다.

다음으로 사적인 반성적 사고(the private reflection link)는 초기 직관에 모순되는 새로운 직관이 자발적으로 활성화 되어 초기 직관에 의한 도덕적 판단에 대하여 수정이 이루어지는 과정을 의미한다. 새로운 직관을 촉진시키는 방법으로서 가장 광범위하게 논의되고 있는 것은 역할 채택(Role Taking)이다. 이러한 역할 채택은 인지발달 이론가들이 언급하고 있는 반성적 사고의 경로 중 하나로서, 개인은 하나 이상의 관점에서 이슈나 딜레마를 접하는 과정을 통해 복합적이고 경쟁적인 직관들을 경험한다. 최종 판단은 가장 강한 직관에 의해 도출되거나 대안들 중에서 추론을 통하여 원리를 선택함으로써 도출된다(Haidt, 2001, p. 819).

<그림-4> 도덕적 판단에 관한 이성주의 모델(Haidt J, 2001, p. 815)



주목할 점은 이러한 추론적 판단과 인지발달 이론이 강조하는 이성주의 모델(<그림-4>)은 다르다는 것이다. 이성주의 모델(<그림-4>)은 판단에 있어 감정이 간접적으로 영향을 미치기는 하지만, 감정의 영향이 없다 하더라도 이성에 의해 판단이 도출될 수 있다고 본다. 그러나 사회적 직관주의 모델은 도덕적 판단이 직관적으로 도출된다고 보기에, 추론적 판단은 그에 선행하는 직관적 판단의 도출을 필요로 하며, 그 역할도 직관적 판단에 대한 합리화 또는 수정에 한정된다. 즉, 사회적 직관주의 모델은 상황을 마주하여 처음부터 추론적 형태의 도덕적 판단이 도출될

수 없다고 보고 있다.

2. 그린의 이중과정 모델의 구조

또 하나의 대표적인 사회적 직관주의 이론으로는 그린(J. Greene)의 이중과정 모델(Dual Process Model)이 있다. 그린 또한 하이트와 마찬가지로 감정의 영향을 크게 받는 직관적 도덕 판단이 높은 비중으로 도출된다고 보았다. 그러나 상기한 바와 같이, 하이트가 추론의 역할을 직관적 도덕 판단에 대한 합리화에 국한되는 것으로 본 것과 달리, 그린은 추론이 도덕적 판단에 직접 기능하며 직관과 경쟁할 수 있다고 주장함으로써 두 이론은 충돌하였다.

1) 직관적 도덕 판단의 기제

그린은 피해와 공정성 차원에서 동일한 조건을 갖는 다음의 두 딜레마 상황에서 상이한 판단이 도출되는 이유를 규명하고자 하였다.

> [트롤리(trolley) 딜레마]
> 당신은 철로 옆에 서서 멈출 수 없는 트롤리가 철로를 따라 달려오고 있는 광경을 보고 있다. 트롤리가 이대로 계속 달린다면 터널에 있어서 트롤리를 피할 수 없는 다섯 명의 인부가 치어 죽을 것이다. 그러나 당신이 선로 전환 레버를 당겨 트롤리의 진행 경로를 옆 선로로 바꾼다면, 트롤리는 선로를 바꾸어 달리게 되고 옆 선로에 있던 인부 한 명이 트롤리에 치어 죽을 것이다. 선로 전환 레버를 당겨야 하는가?
>
> [인도교(footbridge) 딜레마]
> 당신은 선로 위 육교에 서서 멈출 수 없는 트롤리가 다섯 명의 인부를 향해 돌진하는 광경을 보고 있다. 그 인

부들을 죽음에서 구하려면, 당신이 당신 옆에 서 있는 덩치 큰 사람을 선로로 밀어 떨어뜨려야 한다. 덩치 큰 사람을 떨어뜨릴 경우 그 사람은 죽겠지만 트롤리는 멈추게 되어 다섯 명의 인부는 목숨을 건질 수 있다. 덩치 큰 사람을 밀어서 철로로 떨어뜨려야 하는가?

트롤리 딜레마 상황에 있어 피험자의 대부분은 레버를 당기는 것이 옳다고 판단한 반면, 인도교 딜레마 상황에서는 소수의 피험자만이 덩치 큰 사람을 밀어서 떨어뜨려도 좋다고 판단하였다. 그린은 수단만 다를 뿐 피해와 공정성 차원에서 동일한 결과로 이어지는 두 딜레마 상황에서 개인의 도덕적 판단이 상이하게 도출되는 현상은 감정의 활성화 정도에 있어서의 차이에 기인할 것이라고 가정하였고, 직접적으로 사람을 대상으로 하는 '인신적 도덕 딜레마(Moral Personal Dilemma) 상황'과 직접적으로 사람을 대상으로 하지 않는 '비인신적 도덕 딜레마(Moral Impersonal Dilemma) 상황'에서 감정의 활성화 정도에 차이가 발생하여 도덕적 판단의 차이로 이어진 것으로 예상하였다(Greene et al, 2001, p. 2106).

<표-3> 인신적 도덕 딜레마와 비인신적 도덕 딜레마 예시(Greene, 2001 재구성)[7]

인신적	[장기 이식 딜레마] 당신은 의사이다. 각기 다른 장기 손상으로 죽음에 직면한 5명의 환자와 건강한 환자 한명이 있다. 5명을 살리기 위한 유일한 방법은 장기가 건강한 한 명의 환자를 희생시켜 그의 장기를 이식하는 방법이 있다. 이러한 방법을 통할 경우 장기가 건강한

7) http://science.sciencemag.org/content/suppl/2001/09/13/293.5537.2105.DC1 에서 그린이 연구에 활용한 인신적 도덕 딜레마(Moral Personal Dilemma)와 비인신적 도덕 딜레마(Moral Impersonal Dilemma) 및 도덕과 무관한 딜레마(Non-Moral Dilemma)의 전체 문항을 확인할 수 있다.

도덕 딜레마	환자 한명은 사망한다. 5명을 살리기 위해 한명을 희생시키는 것은 적절한가.
	[사파리 딜레마] 당신과 환경연구자들은 외딴 정글을 연구하는 과정에서 테러집단에 억류되었다. 며칠이 지나며 한 테러리스트가 당신을 좋아하게 되었고, 당신에게 다가와 내일 아침이면 두목이 당신과 연구자들을 모두 죽이려 한다고 알려준다. 테러리스트는 당신이 도망칠 수 있게 도와주고 싶지만 당신이 경찰에 신고할 위험이 있다고 걱정한다. 그는 당신이 동료들 중 한명을 죽이고 자신이 그 장면을 녹화하여 보증으로 가지고 있는다면, 당신이 도망갈 수 있게 도와주겠다고 약속한다. 동료를 죽이고 도망가는 것은 적절한가.
	[구명보트 딜레마] 당신이 탄 배가 조난 당함에 따라 당신은 구명보트에 타게 되었다. 정해진 인원보다 많은 인원이 배에 올라타는 바람에 배는 가라앉기 직전의 상태이다. 파도가 점점 거세지고 있는 상황에서 구명보트가 가라앉지 않도록 하기 위해서는 힘이 약한 노인들을 밀어서 바다로 떨어뜨리는 방법 밖에는 없다. 당신과 나머지 탑승객들을 살리기 위해 노인들을 바다로 밀어넣는 것은 적절한가.
비인신적 도덕	[유독 가스 딜레마] 당신은 병원의 야간 감시원이다. 어느 날 옆 건물의 사고로 인하여 유독가스가 건물의 환기 시스템으로 유입되었다. 병원에는 3명의 환자가 있는 병실과 1명의 환자가 있는 병실이 있다. 당신이 아무런 조치를 취하지 않으면 유독가스는 3명의 환자가 있는 병실로 흘러들어가고 3명은 사망하게 된다. 당신이 환기 통로 변경 스위치를 누르면 유독가스는 1명의 환자가 있는 병실로 흘러들어가고 1명의 환자가 사망하는 대신 3명의 환자는 살 수 있다. 스위치를 누르는 것은 적절한가.
	[환경 정책 딜레마] 당신은 정부의 정책 결정 위원이며, 환경 재해를 해결하기 위한 두 정책 중 한 가지를 선택해야 한다. 첫 번째 정책을

딜레마	선택할 경우 90%의 확률로 아무도 희생되지 않지만 10%의 확률로 1000명의 사상자가 발생할 수 있다. 두 번째 정책을 선택할 경우 92%의 확률로 아무도 희생되지 않지만 8%의 확률로 10,000명의 사상자가 발생할 수 있다. 어떠한 정책을 선택하는 것이 적절한가.
	[이력서 딜레마] 당신에게는 최근 직장을 갖기 위해 무던히 노력하였음에도 연이어 취업에 실패한 친구가 있다. 그는 자신의 이력서에 조금만 더 인상적인 내용들이 들어간다면 채용이 될 것이라고 생각하여, 당신에게 허위로 인상적인 내용들을 추가해달라고 부탁한다. 당신이 도와줄 경우, 친구는 자신보다 자격 조건을 갖춘 사람들을 경쟁에서 이기고 무조건 합격할 것이 분명하다. 친구가 취업을 할 수 있도록 이력서에 허위 정보를 기재해주는 것은 적절한가.

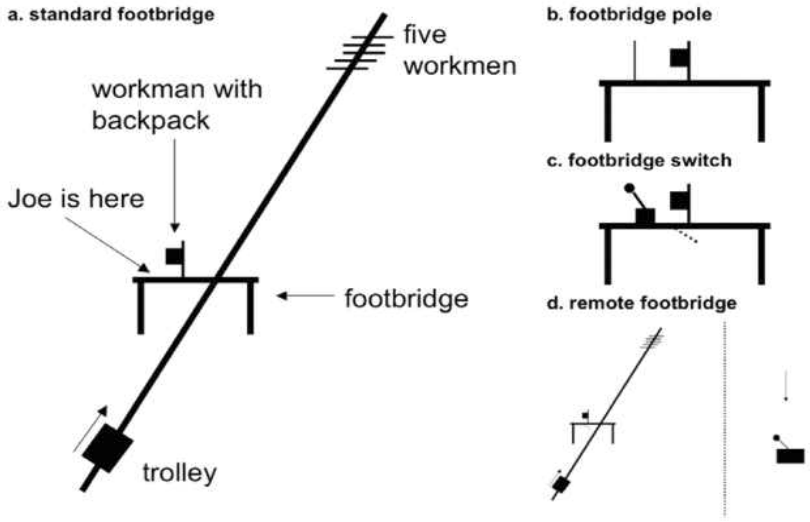
그린은 인도교 딜레마 상황과 같은 인신적 도덕 딜레마 상황에서는 정서 반응 및 사회적 인지와 연관된 두뇌 영역의 신경 활동이 증가할 것이며, 트롤리 딜레마 상황과 같은 비인신적 도덕 딜레마 상황에서는 고차적 인지와 관련된 두뇌 영역이 상대적으로 더 활성화 될 것이라고 예측하였다. fMRI 촬영 결과, 인신적 도덕 딜레마 상황에서는 정서 반응과 관련된 복내측전전두피질(vmPFC)을 포함한 내측전전두피질(medial prefrontal cortex) 등이 활성화 되었으며, 비인신적 도덕 딜레마 상황에서는 고차적 인지와 연관된 두뇌영역인 배외측전전두피질(DLFPC)과 하측두정엽(inferior parietal lobe)이 활성화 되었다(Greene, 2015, p. 147; Greene et al, 2001, p. 2106).

이에 대하여 그린은 '덩치 큰 사람을 직접 밀어서 떨어뜨리는 인신적 행위는 트롤리 딜레마에서 레버를 당기는 것과 같은 결과로 귀결됨에도, 트롤리 딜레마에서의 행위보다 훨씬 강한 부정적 정서를 갖게 만든다. 판단의 정당화 가능성 여부와 관련 없이 이러한 부정적 정서는 사람을

인도 밑으로 밀면 안 된다는 판단으로 도출된다.' 라고 해석하며 도덕적 판단에 있어 감정의 영향력을 강조하였다(Greene, 2008, p. 63).

나아가 그린은 <그림-5>과 같이 인도교 딜레마 상황에서 사람을 떨어뜨리는 방식 즉, 감정 반응의 정도를 좌우하는 수단에 있어 차이를 두고 그에 따라 긍정적 응답의 정도에 변화가 있는지 조사하였다.

<그림-5> 인도교 딜레마의 수단적 변형(Greene, 2009, p. 366)



그린은 상황 A에서 손으로 사람을 밀어야 하는 전형적인 인도교 딜레마 상황을 제시하였고 상황 B에서는 손으로 사람을 미는 것은 아니지만 신체적 힘을 이용하여 레버를 당겨 낙하문을 개방하여 사람을 떨어뜨려야 하는 상황을 제시하였으며, 상황 C에서는 근거리에서 스위치를 누름으로써 낙하문을 개방하여 사람을 떨어뜨려야 하는 상황을 제시하였고, 상황 D에서는 원거리에서 스위치를 누름으로써 낙하문을 개방하여 사람을 떨어뜨려야 하는 상황을 제시하였다. 즉, 상황 A와 상황 B, C, D 사이에는 피해자와의 신체적 접촉 여부에 있어 차이를 두었으며 상황 A, B와 상황 C, D 사이에는 신체적 힘의 사용 여부에 있어 차이를 두었고,

상황 A, B, C와 상황 D 사이에는 거리의 가까운 정도에 차이를 두었다(Greene, 2009, p. 366).

실험 결과, 상황 B에서는 상황 A와 유사한 정도의 긍정적인 응답이 도출되었지만 상황 C와 D에서는 상황 A에 비하여 두 배에 가까운 비율로 긍정적인 응답이 도출되었다. 그린은 이에 대하여, 신체적 접촉 여부와 피해자와의 거리 정도는 불승인의 감정 반응 정도에 큰 영향을 주지 않았지만 신체적인 힘을 사용하는가 여부 즉, 자신의 신체적 움직임을 통해 직접 인신적 피해를 끼치는 것은 불승인 감정의 반응 정도에 큰 영향을 주었으며, 이러한 감정 반응의 차이가 판단의 차이로 이어졌다고 지적하였다(Greene, 2009, p. 367).

이와 관련하여, 발데솔로(P. Valdesolo)와 디스테노(D. DeSteno)는 감정적 민감성이 맥락적 특성을 갖는다고 지적하였다(Valdesolo & Desteno, 2006). 그들은, 감정적 상태란 환경 요인에 의해 좌우되는 일시적인 성격을 가지며 그러한 감정 상태에는 여러 요인들이 다중적으로 영향을 미친다는 슈바르츠(N. Schwarz)와 클로어(G. Clore)의 연구(Schwarz & Clore, 1996)에 기반하여, 피험자들에게 코미디 프로그램을 보여주고, 그들이 긍정적 감정 상태에 있는 동안, 트롤리 딜레마 상황과 인도교 딜레마 상황에서의 응답 비율에 차이가 목격되는지 여부를 조사하였다. 그들에 따르면 인지적 영역이 활성화 되었던 트롤리 딜레마 상황에 있어서의 판단은 감정 상태의 변화에 영향을 받지 않았지만 감정적 영역이 활성화 되었던 인도교 딜레마 상황에 있어서의 판단은 감정 상태의 변화에 영향을 받았고, 그로 인하여 긍정적 응답의 비율이 높게 관측되었다(Valdesolo & Desteno, 2006, pp. 476-477). 즉, 개인의 도덕적 판단은 감정의 활성화 정도에 크게 영향을 받는다.

2) 직관적 도덕 판단에 대한 인지적 조절의 기제

하이트가 도덕적 판단이 직관적 형태로 도출된다고 보고 추론의 역할

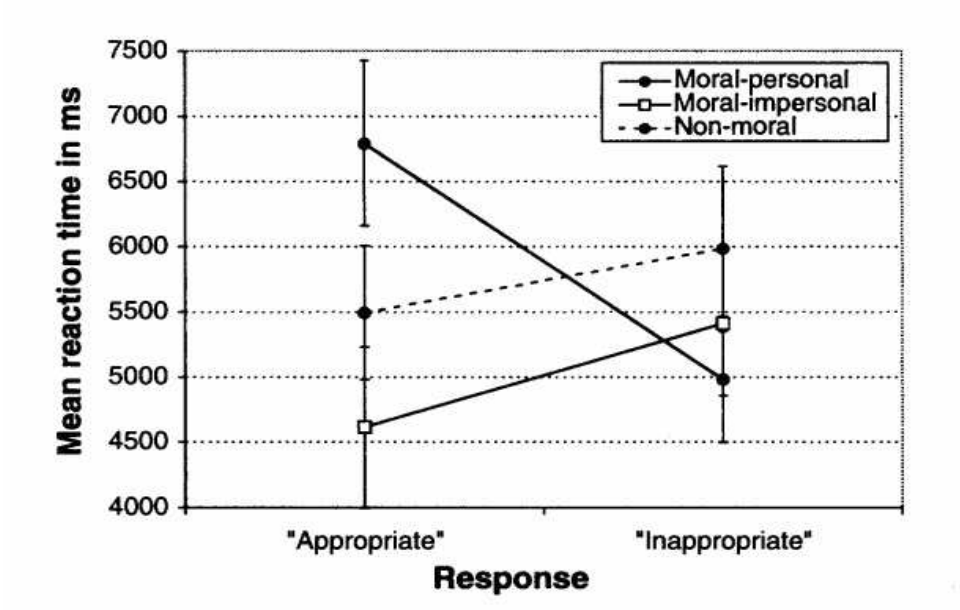
을 직관적 도덕 판단에 대한 합리화에 한정한 것과 달리, 그린은 추론이 도덕 판단의 도출과 직관적 도덕 판단의 수정에 직접적으로 기능한다고 보았다. 즉, 그린에 따르면 앞서 살펴본 비인신적 도덕 딜레마 상황에서와 같이 추론적 형태의 도덕 판단이 도출되기도 하며, 직관적 도덕 판단이 도출된 경우에 있어서도 추론에 의한 인지적 조절이 이루어지기도 한다. 그린은 특히, 직관적 도덕 판단에 대한 인지적 조절이 수월하게 이루어지기 위한 조건으로 감정 반응의 정도가 약할 것과 결과론적으로 강한 근거가 존재할 것을 지적하였다. 먼저 감정적 반응의 정도와 인지적 조절의 관계에 대해 살펴보면 다음과 같다.

그린은 각 딜레마 상황에 있어 긍정적 응답과 부정적 응답 사이에 목격되는 시간의 차이를 조사함으로써 감정적 반응의 강한 정도가 도덕적 판단의 수정에 미치는 영향을 분석하였다. 그린은 인도교 딜레마에서 덩치 큰 사람을 떨어뜨리기로 결정하기 위해서는 정서적 반응을 눌러 이기는 것이 필요하기 때문에, 인신적 침해에 대하여 긍정적 응답을 제시한 피험자의 판단이 도출되는 데 걸리는 시간은 부정적 응답을 제시한 피험자의 판단이 도출되는 시간보다 길 것이라고 예상하였으며, 트롤리 딜레마와 같은 비인신적 도덕 딜레마 상황에서는 눌러 이겨야 할 정서적 반응이 없거나 훨씬 약하기 때문에 긍정적으로 응답하는 사람과 부정적으로 응답하는 사람 사이에 반응 시간의 차이가 없을 것이라 예상하였다(Greene et al, 2001, p. 2107).

실험 결과, <그림-6>와 같이, 비인신적 도덕 딜레마(Moral Impersonal Dilemma) 상황에서 긍정적 응답과 부정적 응답 사이에 발생한 시간 차이는 도덕과 무관한(Non-Moral) 딜레마 상황(시간적 제약 조건 하에 버스와 기차 중에 어떠한 것을 선택하는 것이 옳은가와 같은 문항)에서 긍정적 응답과 부정적 응답 사이에 발생한 시간 차이와 별 차이가 없을 정도로 짧았다. 그러나 인신적 도덕 딜레마(Moral Personal Dilemma) 상황에서의 긍정적 응답과 부정적 응답 사이에 소요되는 시간은 비도덕적 딜레마의 세 배, 비인신적 도덕 딜레마의 두 배

이상에 해당하였다(Greene et al, 2001, p. 2107).

<그림-6> 딜레마 유형별 응답 소요 시간(Greene, 2001, p. 2107)



즉, 그린의 실험에 따르면 인지적 조절을 통해 감정적 반응에 반하는 판단이 도출되기도 하며, 강한 감정 반응에 기인한 판단일수록 그를 인지적으로 조절하기 위해서 더 많은 노력과 시간이 소요된다(Greene & Haidt, 2002, p. 522). 이는 하이트가 추론된 판단(the reasoned judgment link)이 성공적으로 도출되기 위해 인지적 처리 능력이 높으면서 동시에 초기 직관이 약해야 한다고 주장한 것과 일맥상통한다.

다음으로 결과론적으로 강한 근거의 존재 여부와 인지적 조절의 관계에 대하여 살펴보면 다음과 같다. 앞서 살펴본 바와 같이 그린은 인도교 딜레마 상황과 같은 인신적 도덕 딜레마 상황에 있어서의 두뇌 영상과 반응소요시간 자료를 모두 놓고 볼 때, 인신적으로 유해한 행동에 대해 불승인하도록 유도하는 강한 부정적인 감정적 반응들이 존재하며, 그러

한 감정적 반응이 강하게 도출될수록 그에 반하는 추론적 판단을 도출하기 위해서는 더 많은 노력과 시간이 소요된다고 지적하였다. 그에 덧붙여, 그린은 강한 결과론적 근거가 있는 경우, 인간의 인지적 심리 과정이 이러한 불승인의 감정적 심리 과정과 경쟁하는 것이 더욱 용이해진다고도 지적하였다(Greene, 2008, p. 46).

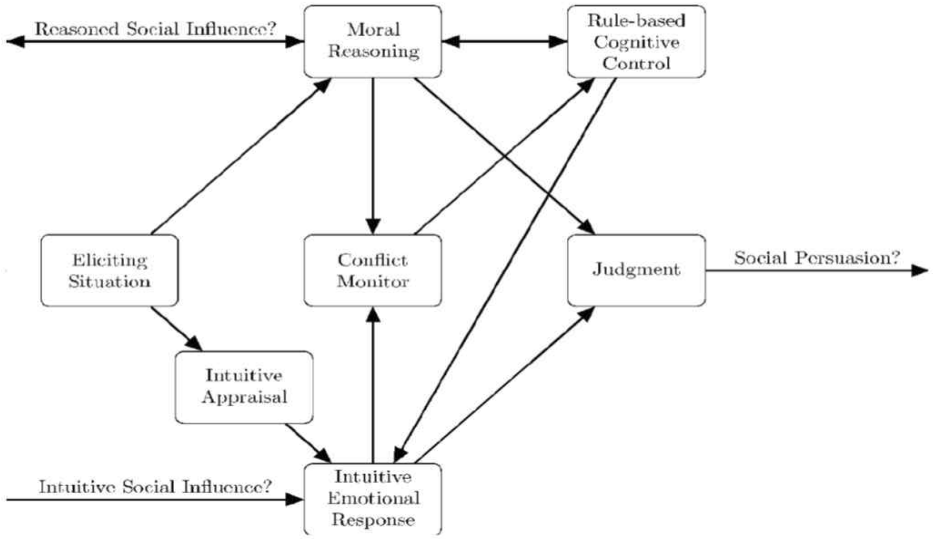
> [우는 아이 딜레마]
> 전쟁이 진행 중인 어느 날, 당신은 아기를 데리고 이웃들과 함께 적군의 눈을 피해 지하실에 숨어 있다. 아기가 울기 시작하자 당신은 소리가 새어나가지 않도록 아기의 입을 막는다. 만약 손을 치우면 아기는 더 크게 울 것이고 그 소리를 들은 적군은 당신과 아기 및 마을 사람들을 찾아내 모두 죽일 것이다. 그러나 만약 손을 치우지 않으면 아기는 질식사 하게 된다. 이러한 상황에서 손을 치우지 않고 아기를 죽게 하는 것은 도덕적으로 올바른 선택인가?

그린은 우는 아이 딜레마 상황에서 피험자들의 응답 양상과 영아 살해의 상황에서 피험자들의 응답 양상을 비교하며 강한 결과론적 근거가 있는 경우, 인간의 인지적 심리 과정이 불승인의 직관적 심리 과정을 더욱 쉽게 극복할 수 있다고 지적하였다. 십대 소녀가 원치 않은 임신에 의해 낳은 아이를 죽이는 것이 적절한지 여부를 판단하는 영아 살해 상황에 대하여 대부분의 피험자들이 순간적으로 잘못된 행위라는 판단을 제시하였고 감정과 관련된 부위들이 활성화 되었다. 반면, 우는 아이 딜레마 상황에서 다양한 응답을 제시한 피험자들은 답변을 도출하는데 있어 모두 오랜 시간이 소요되었으며, 그 과정 동안 뇌의 인지적 작용과 연관이 깊은 부위들이 활성화 되었다. 이에 근거하여 그린은 강한 결과론적 근거와 직관적인 불승인의 감정이 충돌하는 경우, 응답까지의 반응시간이

늘어나며 고차적 인지기능과 연관된 두뇌 영역이 활성화된다고 지적하였다(Greene, 2008, p. 44-46). 즉, 결과론적 근거가 강하고 인지적 정보 처리 능력이 우수하며 초기 직관이 약할수록 추론적 심리 과정은 최초의 직관적 판단과 수월하게 경쟁할 수 있게 된다.

이러한 발견에 기반하여 그린은 도덕적 판단에 관한 이중 과정 모델(<그림-7>)을 도출하였다. 이중과정 모델(Dual Process Model)에 따르면 도덕적 판단은 보다 감정적인 직관적 판단과 보다 인지적인 추론적 판단 모두에 의해 도출되며, 인지적 조절을 통해 직관적 판단을 극복하고 추론적 판단을 도출하는 것 또한 가능하다. 초기 직관에 영향을 미치는 감정 반응은 맥락 종속적인 성격을 가지며, 반응의 강도가 강할수록 직관적 판단이 도출될 가능성이 높아진다. 추론적 판단이 도출되기 위해서는 감정 반응의 정도가 약하거나, 강한 결과론적 근거가 존재하여야 한다.

<그림-7> 도덕적 판단에 관한 이중 과정 모델 (Paxton & Greene, 2010, p. 514)



나아가 그린은 의무론적 판단이란 감정 반응에 추동된 판단이며, 이러한 의무론적 판단에 있어 추론은 합리화 기능만을 담당할 뿐 판단의 과정에 직접적으로 기능하지 못하기에, 의무론적 판단의 신뢰성은 높이 평가하기 어렵다고 지적하였다. 덧붙여 그는 의무론이란 인간이 자신의 행위를 설명하기 위하여 만들어낸 일종의 도덕적 작화(Moral Confabulation)라고 비판하였다. 또한 그린은 의무론적 판단과 대조적으로 결과론적 판단은 비교적 인지적이며 그러한 결과론적 판단의 과정에 있어 추론은 사후 합리화가 아닌 판단의 산출에 작용하는 진정한 의미의 도덕적 추론으로 작용하기에 결과론적 판단은 비교적 신뢰할 수 있다고 주장하였다(Greene, 2008, pp. 59-61).

이후에도 그린은 다양한 연구들에 근거하여 이러한 관점을 점차 공고히 하였다(Greene, 2014, p. 701). 그가 근거로 삼은 연구로는 감정 반응과 관련된 전측두엽에 치매를 앓고 있는 환자들이 인도교 딜레마 상황에 있어 일반인에 비하여 세 배 가량 높은 비율로 결과론적 판단을 내린다는 연구(Mendez, Anderson, & Shapira, 2005, pp. 193-197), 감정 반응과 관련이 깊은 복내측 전전두엽 병변 환자들이 결과론적 고려를 중시한다는 연구(Thoma, Croft, & Tranel, 2011, pp. 2186-2196), 보다 공감적인 성향의 사람들이 의무론적 판단을 내리는 경향이 더 강하다는 연구(Conway, & Gawronski, 2013, pp. 216-235) 등이 있다.

Ⅳ. 사회적 직관주의 이론 비판

도덕적 판단의 기제 규명을 목표로 하는 두 사회적 직관주의 이론은 직관과 추론의 역할 및 관계 방식에 대하여 충돌하는 주장을 제시하는 관계로 상정되어 왔다. 이로 인하여 두 이론은 합의점을 도출하지 못하였고, 그린의 이론이 상대적으로 더 넓은 범주의 현상들을 포괄할 수 있는 타당한 이론으로 간주되었다. 또한 의무론적 판단은 감정적인 직관적 판단과 이에 대한 합리화에 불과하기에 규범적 지위를 인정받기 어렵다는 비판을 받게 되었다.

그러나 서두에 언급한 바와 같이, 이러한 논의들이 타당성을 담보하기 위해서는 두 이론이 규명한 도덕적 판단의 현상 범주를 구체화하는 규명 범주 구체화 작업을 선행하고, 이를 통해 각 이론이 실제 규명한 도덕적 판단의 현상 범주보다 더 넓은 범위의 현상들을 포괄하는 것으로 확대 해석된 부분이 없는지 검토할 필요성이 있다. 또한 두 이론이 활용한 직관과 추론이라는 인지 심리학적 개념 범주가 실제 지칭하는 바를 도출하는 개념 범주 명료화 작업을 선행하고, 이를 통해 인지 심리학적 개념 범주를 차용하여 도덕적 판단의 현상을 해석함에 있어 모순이 발생할 위험성이 없는지 검토할 필요성이 있다.

1. 도덕적 판단의 범주 설정의 문제점과 대안

1) 규명 범주 구체화 필요성

두 사회적 직관주의 이론의 타당성에 관한 다양한 평가와 논의는 추론적 판단의 가능성에 관한 사회적 직관주의 모델과 이중과정 모델의 견해 차이에서 시작된다. 하이트(J. Haidt)의 사회적 직관주의 모델과 그린(J. Greene)의 이중과정 모델은 도덕적 판단에 있어 추론의 역할 즉, 추론이 도덕적 판단의 과정에 직접 기능 하는지 아니면 직관적 판단의 합리화에 한정되는지에 대하여 충돌하는 주장을 제시하였다.

결론부터 살펴보면, 다수의 연구자들은 하이트의 이론이 추론의 역할을 지나치게 축소하여 파악하고 있다고 비판하며, 추론에 의한 도덕 판단의 가능성을 긍정하는 그린의 이론이 더욱 설득력을 갖는다고 지지하였다(정창우, 2011, p.. 111-120). 하이트에 대한 대표적인 비판으로 튜리엘(E. Turiel)과 나바에츠(D. Narvaez)의 비판을 살펴볼 수 있다. 튜리엘은 하이트가 다양한 발달심리학 및 인지심리학 영역의 연구들을 외면하고 있다고 비판하였다. 그는 하이트가 도덕적 판단이 감정적이고 직관적이라는 것을 지지하는 여러 사회심리학 연구를 인용하고 있지만, 반성적이고 신중한 추론과 관련된 연구들은 반영하지 않고 있다고 지적한 바 있다(Turiel, 2006/2010, pp. 49-53).

나바에츠는 일상적으로 일어나는 다양한 도덕적 판단에 있어 직관적인 과정이 절대적으로 높은 비중을 차지하지 않으며 심사숙고에 의한 추론적 과정 또한 높은 비중을 차지한다고 비판하였다(Narvaez, 2008a, p. 235). 나바에츠는 다양한 사례를 근거로, 개인의 도덕적 판단 과정에 있어 직관과 의식적인 추론은 우선순위를 번갈아가며 상호작용한다고 주장하였다. 그녀는 도덕적 판단에 있어 추론이 사용되는 상황으로, ‘그가 화난 것으로 보여. 내가 무슨 말을 해야 할까?’, ‘그녀의 마음을 상하게 하고 싶지 않아. 어떻게 해야 할까?’, ‘이 미팅은 시간낭비에 불과

해. 모두에게 도움이 될 수 있도록 하기 위해 어떻게 해야 할까?', '어떻게 하지? 약속시간을 맞추려면 빨리 가야하는데, 내 도움을 필요로 하는 사람이 있네.' 등의 예를 제시하였다(Narvaez, 2008a, p. 235). 나아가 그녀는 도덕적 의사결정에는 개인적인 목표와 계획을 설정하는 것(B. Williams), 자신의 책임이 무엇인가를 결정하는 것(H. Frankfurt), 대안들 중에서 어느 행동 선택이 최선인가를 심사숙고 하는 것(J. Rawls), 다양한 고려들을 조화시키는 것(J. Wallace), 선택된 의사결정과 행동을 평가하는 것(L. Blum)등이 포함되며 이는 추론의 도움이 없이는 이루어질 수 없다고 지적하였다(Narvaez, 2008a, p. 234).

이렇듯 사회적 직관주의 모델은 도덕적 판단에 있어 직관의 역할에 대한 관심을 불러일으켰지만, 이론의 핵심적 주장에 대한 이러한 비판들은 사회적 직관주의 모델의 타당성에 대한 의구심으로 이어지게 되었다. 이에 따라, 상기한 바와 같이, 하이트의 사회적 직관주의 모델에 비하여 그린의 이중과정 모델이 도덕적 판단의 기제를 보다 포괄적이고 타당하게 설명하고 있는 이론으로 평가받게 되었다. 즉, 직관과 추론의 역할 및 관계에 대한 두 사회적 직관주의 이론의 주장은 충돌하는 것으로 상정되어 왔으며, 그로 인하여 합의점을 도출하지 못한 채 그린의 이론이 더욱 포괄적이고도 타당한 이론으로 받아들여졌다.

그런데 과연 두 사회적 직관주의 이론에 관한 이러한 논의가 타당성을 담보하고 있을까. 이러한 논의가 타당성을 담보하기 위해서는, 직관과 추론의 역할 및 관계에 대한 두 이론의 주장이 충돌하여야 하며, 이러한 충돌이 존재하기 위해서는 두 이론이 동일한 범주의 도덕적 판단 현상을 대상으로 상충하는 주장을 펼치고 있어야 한다. 그러나 두 이론이 규명한 도덕적 판단의 현상 범주를 구체화하는 작업이 선행되지 않은 채, 두 이론에 대한 논의가 진행된 측면이 있다. 이러한 논의들은 전적으로 두 연구자의 표현에 의존하였고, 이에 두 사회적 직관주의 이론이 '도덕적 판단'이라는 동일한 현상 범주에 대해 상이한 규명을 하고 있는 것으로

전제하였다. 그러나 만일, 예를 들어, 두 이론의 규명이 상이한 범주의 도덕적 판단의 현상을 전제로 한다면, 두 이론은 충돌하는 관계가 아닌, 도덕적 판단이라는 현상의 하위 현상 범주들을 보완적으로 규명하는 관계로 해석하는 것이 타당하다.

그렇기에 두 이론에 관한 논의의 타당성을 담보하기 위해서는 두 이론이 실제로 규명한 현상 범주를 구체화 하는 작업을 수행할 필요가 있다. 스트라우스(A. Strauss)와 코빈(J. Corbin)이 지적하였듯 이론 연구에 있어 자료에 대한 모든 이론적 설명과 범주 설정 등은 언제나 특수한 상황을 전제로 한다(Strauss & Corbin, 1990/1995, p.55). 그럼에도 불구하고, 뷰러워이(Burawoy, 1991)가 지적하였듯 이론은 특정한 범주에 대한 연구에서 얻어낸 경험적 일반화를 보다 거대한 실재에 대한 일반화된 진술로 받아들이게 하는 위험성을 내포한다(Charmaz, 2006/2013, pp. 266-271). 그렇기에 두 이론의 규명 범주를 구체화하는 작업은 각 이론이 규명한 도덕적 판단의 현상 범주를 잘못 파악함으로써 발생하는 불필요한 논쟁의 가능성을 없애고 두 이론의 규명들을 타당하게 관계 맺도록 하기 위해 반드시 선행되어야 할 필요성이 있다.

2) 규명 범주 구체화 가능성

하이트는 '사회적 직관주의 모델'을 통해 도덕적 판단이 직관적 형태로 도출된다고 주장하며, 그린은 '이중과정 모델'을 통해 도덕적 판단이 직관적 형태 뿐 아니라 추론적 형태로도 도출된다고 주장한다. 이러한 상이한 주장을 내세우는 두 연구의 규명 범주 즉, 연구의 규명 대상이 되는 현상의 범주를 구체화하기 위해서는, 규명 범주에 대한 하이트와 그린의 언급 및 각 이론이 주장의 근거로 제시한 연구 문항을 분석함으로써 두 연구자가 상정하고 있는 '도덕적 판단'의 범주를 비교할 필요가 있다.

먼저 하이트의 연구에 대하여 살펴보면, 하이트의 사회적 직관주의 모

델은 도덕에 대한 정당화를 전제로 논의를 진행하지도 않았으며, 피해와 공정성이라는 기준을 적용하여 인습적 영역과 구분되는 도덕의 영역을 설정하고 논의를 진행하지도 않았다. 하이트는 도덕에 대한 정당화 과정 없이, 도덕을 '개인들이 도덕으로 받아들이는 무언가'라는 열린 차원의 개념으로 설정하고 논의를 진행하였다. 즉, 사회적 직관주의 모델에서 다루는 도덕적 판단의 범주는 서구적 문화의 관점에서 개인적 영역 및 인습적 영역과 구분하는 도덕적 영역에 있어서의 판단을 지칭하는 것이 아니며, '사람들이 도덕으로 인지하는 무언가'를 의미한다.

그렇다면 하이트의 규명은 '사람들이 도덕으로 인지하는 무언가'에 해당하는 모든 현상을 포괄하고 있을까. 하이트는 인지주의 접근이 상정한 전통적인 도덕의 상황 외에도 사람들이 도덕의 문제로 인지하는 상황들이 있다고 주장하였으며, 그러한 상황에서의 도덕적 판단의 기제를 규명하기 위해, '피해와 공정성의 기준이 적용되지는 않지만 불온한 감정 반응을 유발하는 상황'에 초점을 맞추었다. 그는 이러한 전통적인 도덕 상황에 해당하지 않지만 불온한 감정 반응을 일으키는 상황이 개인들에게 도덕적 상황으로 인식되고 도덕적 평가의 대상이 된다는 것에 관심을 가졌고, 이러한 상황을 전제로 도덕적 판단의 기제를 규명하고자 하였다. 그는 이러한 상황에서 도출되는 도덕적 판단의 양식을 '피해와 공정성의 기준이 적용되어 추론적 판단의 전형으로 삼을 수 있는 콜버그의 하인츠 딜레마 상황에서의 판단 양식' 및 '도덕과 무관하지만 직감적인 반응을 유도하는 상황에서의 판단 양식'과 비교하였으며, 이를 통해 도덕적 판단이 직감적인 반응을 유도하는 상황에서의 판단과 닮아있으며 추론은 그러한 직관적 도덕 판단을 합리화하는데 주로 기능한다고 주장하였다(Haidt, Bjorkulund, & Murphy, 2000, pp. 5-11; Haidt & Bjorklund, 2008, pp. 197-198).

즉, 하이트의 사회적 직관주의 모델의 핵심 주장은 모든 도덕적 판단의 상황을 대상으로 하지 않으며, '피해와 공정성의 기준이 적용되는 전통적 도덕 상황은 아니지만 불승인의 감정 반응이 도출되는 불온한 상

황’에서의 도덕적 판단을 그 규명 대상으로 한다. 이러한 규명 범주는 앞서 살펴본 ‘근친상간 시나리오’와 ‘식인 풍습 시나리오’와 같이 주요 근거로 제시된 인터뷰 문항을 통해서도 확인할 수 있다. 그렇기에 하이트의 사회적 직관주의 모델은 ‘피해와 공정성의 기준이 적용되지 않지만 불온한 상황에서의 도덕적 판단이 직감적인 반응이 유발되는 상황에서의 판단과 닮아있으며, 그러한 경우에 있어 추론은 해당 도덕 판단을 합리화하는데 주로 기능한다’는 것을 규명하였다고 해석하는 것이 타당하다.

전통적 도덕 상황 이외의 상황을 대상으로 도덕적 판단의 기제를 규명하고자 한 하이트와 달리 그린은 ‘피해와 공정성의 기준이 적용되는 인지주의 전통의 도덕적 상황’에서, 감정의 역할과 직관적 도덕 판단의 기제를 규명하고자 하였다. 그린의 이중과정 모델이 대상으로 하는 ‘피해와 공정성의 기준이 적용되는 도덕적 상황’이라는 규명 범주는 그가 근거로 제시한 인터뷰 문항을 통해서도 확인할 수 있다. 그린은 피해와 공정성의 기준이 적용되는 도덕 상황에 있어 감정의 역할을 규명하기 위하여, 결과론적으로 동일한 피해와 공정성의 기준이 적용되지만 수단에 있어 차이를 갖는 트롤리 딜레마 상황과 인도교 딜레마 상황에서의 응답 양상을 비교하였다.

또한 상기한 바와 같이, 그는 이러한 두 도덕적 딜레마 상황에 있어 활성화 되는 뇌의 영역을 비교하고 긍정적 응답과 부정적 응답의 도출에 소요되는 시간을 비교함으로써, 도덕적 판단이 감정적 반응의 활성화 정도에 따라 직관적 형태와 추론적 형태 모두로 도출된다고 주장하였다. 또한 직관적 형태의 도덕 판단이 도출된다 하더라도 결과론적 근거가 있다면 인지적인 심리과정이 감정적이고 직관적인 심리과정과 경쟁할 수 있으며, 감정적 반응이 강할수록 그와 상충하는 판단을 도출하기 위해서 더 많은 인지적 노력이 필요하다고 지적하였다.

즉, 규명 범주 구체화 작업에 따를 때, 이중과정 모델의 핵심 주장은 모든 도덕적 상황이 아닌, ‘피해와 공정성의 기준이 적용되는 도덕적

상황’을 전제로 한다. 그렇기에 그린의 이중과정 모델은 ‘피해와 공정성의 기준이 적용되는 인지주의 전통의 도덕적 딜레마 상황’에서의 도덕적 판단이 감정적 반응의 활성화 정도에 따라 직관적 형태와 추론적 형태 모두로 도출되며, 결과론적 근거가 있다면 인지적인 심리과정이 감정적이고 직관적인 심리과정과 경쟁할 수 있으나, 감정적 반응이 강할수록 직관에 상충하는 판단을 도출하기 위해서 더 많은 인지적 노력이 필요하다는 것을 규명하였다고 보는 것이 타당하다.

종합하면 다음과 같다. 사회적 직관주의 모델과 이중과정 모델은 ‘도덕적 판단’이라는 동일한 현상 범주에 대하여 충돌하는 주장을 펼치는 것으로 간주되어 왔으며, 그로 인하여 두 이론의 발견은 합의점을 도출하지 못하였다. 그러나 규명 범주 구체화 작업에 기반할 때, 두 이론은 동일한 현상 범주를 대상으로 하지 않았다. 하이트는 ‘피해와 공정성의 기준이 적용되지 않지만 불온한 감정 반응을 수반하는 상황’을 대상으로 도덕적 판단의 기제를 규명하였으며, 그린은 ‘피해와 공정성의 기준이 적용되는 인지주의 전통의 도덕 상황’을 대상으로 도덕적 판단의 기제를 규명하였다.

즉, 두 이론은 도덕적 판단이라는 현상 범주를 구성하는 상이한 하위의 현상 범주를 대상으로 각각의 규명을 제시하고 있는 관계로 파악하는 것이 타당하며, 도덕적 판단이라는 동일한 현상 범주를 대상으로 상충하는 주장을 펼친 것으로 파악하여서는 안 된다. 두 이론이 도덕적 판단이라는 동일한 현상 범주를 규명 대상으로 한다고 주장하는 것은, 상기한 바와 같이, 특정한 연구에서 얻어낸 경험적 일반화를 보다 거대한 실재에 대한 일반화된 진술로 받아들이게 하는 오류에 해당한다. 즉, 두 이론은 ‘도덕적 판단’이라는 상위의 현상 범주를 규명함에 있어 충돌하는 해석을 제시한 것이 아니며, 도덕적 판단이라는 상위의 현상 범주를 구성하는 상이한 하위의 현상 범주를 대상으로 각각의 규명을 제시한 보완적인 관계로 파악하는 것이 타당하다.

3) 규명 범주 구체화의 적용

사회적 직관주의 모델과 이중과정 모델은 상이한 도덕적 판단의 현상 범주 즉, 피해와 공정성이라는 인지주의 전통의 도덕적 기준이 적용될 수 있는 상황과 적용될 수 없는 상황을 전제로 각각 도덕적 판단의 기제를 규명하였다. 그렇기에 두 이론은 도덕적 판단의 기제를 규명함에 있어 보완적인 관계로 파악하는 것이 타당하며, 상이한 현상 범주를 대상으로 도출된 두 사회적 직관주의 이론은 보완적으로 하나의 통합 모델을 구성할 수 있다. 두 이론은 피해와 공정성의 기준이 적용되지 않는 상황과 피해와 공정성의 기준이 적용되는 상황을 대상으로 각 유형별 인지 즉, 직관적 판단의 기제와 추론적 판단의 기제 및 추론을 통한 직관적 판단의 수정 기제를 규명하였으며 다음과 같이 통합될 수 있다.

(1) 직관적 도덕 판단의 기제

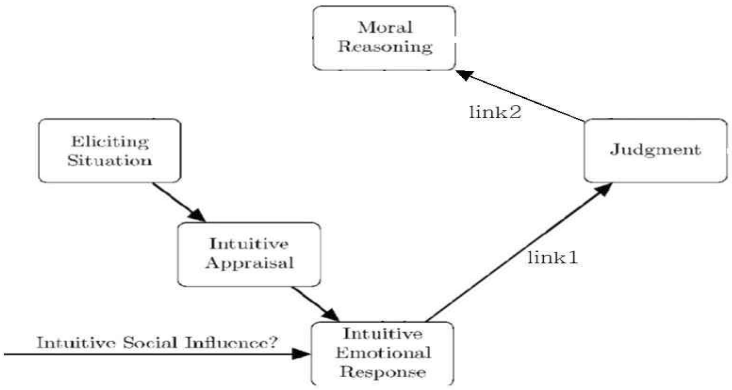
직관적 도덕 판단의 기제는 그린의 이중 과정 모델의 직관적 판단의 기제 부분과 하이트의 사회적 직관주의 모델의 Link 1 '직관적 판단(the intuitive judgment link)' 부분에 해당한다. 상기한 바와 같이 두 이론의 차이점은 그린의 이중 과정 모델이 피해와 공정성의 기준이 적용되는 상황에 있어 행위에 대한 직관적 판단의 기제를 규명하고 있는 반면, 하이트의 사회적 직관주의 모델은 피해와 공정성의 기준을 적용할 수 없는 상황에 있어 직관적 판단의 기제를 규명하고 있다는 점이다. 즉, 두 이론은 피해와 공정성 기준이 적용되는 상황에 해당하는지 여부와 무관하게 직관적 반응이 강하게 발현되는 경우에 있어 도덕적 판단이 직관적 형태로 도출된다는 결론으로 통합된다.

두 이론은 직관적 도덕 판단이 도출되는 경우에 있어, 추론의 역할에 대해서도 동일한 입장을 견지한다. 그린은 추론을 통한 도덕적 판단의 기제와 추론을 통한 직관적 판단의 수정 기제를 규명함으로써 추론의 역

할을 다양하게 인정하고 있지만, 강한 직관적 반응에 의해 도덕 판단이 도출되는 상황에 있어 추론의 역할은 합리화에 한정하고 있다. 그린은 인신적 도덕 딜레마 상황에서 부정적 응답을 보인 피험자들에게 목격되는 의무론적 판단이란 감정 반응에 추동되며, 이러한 의무론적 판단에 있어 추론이란 합리화에 기능할 뿐이라고 지적한다(Greene, 2008, p. 60). 하이트 또한 '인도 오리사 지역 주민과 미국인의 도덕적 판단의 차이를 비교한 슈웨더(R. Shweder)의 문화 비교 연구' 및 '피해와 공정성의 기준이 적용되지 않지만 불온한 감정을 수반하는 상황에 대한 도덕적 판단의 기제 연구' 등을 통해, 강한 직관적 반응에 의해 판단이 도출된 경우 피험자들은 근거를 제시하지 못하거나 의무론적 논리에 의해 판단을 합리화하려 하였다고 지적하며, 이러한 현상을 Link 2 '사후 추론(the post-hoc reasoning link)' 부분을 통해 설명하고자 하였다(Haidt, 2012/2014, pp. 53-55, 90-91).

종합하면, 피해와 공정성 기준의 적용 가능 여부와 무관하게, 직관적 반응이 강하게 도출되는 경우에 있어 도덕 판단은 직관적 형태로 도출되며, 추론은 그러한 직관적 판단을 합리화 하는데 기능한다. 또한 판단에 대한 근거를 제시하기 어려운 경우, 이러한 직관적 판단은 의무론의 형태로 지지된다. 이는 아래의 <그림-8>과 같이 표현할 수 있다.

<그림-8> 직관적 판단 및 사후 추론의 기제 (양해성, 2017, p. 133)



그린의 이중과정 모델(<그림-7>) 중 직관적 도덕 판단의 기제는 <그림-8>에 해당한다. 사회적 직관주의 모델(<그림-3>)의 직관적 도덕 판단의 기제 부분인 Link 1 '직관적 판단(the intuitive judgment link)' 과 Link 2 '사후추론(the post-hoc reasoning link) 또한 <그림-8>과 같이 나타낼 수 있다. 주목할 점은, <그림-8>의 Link 2 부분의 화살표 방향이 그린의 이중과정 모델(<그림-7>)에서 해당 부분의 화살표 방향과 반대라는 점이다.

<그림-8>의 Link 2 부분의 화살표는 하이트의 사회적 직관주의 모델의 사후 추론 기제에 해당한다. 그린의 이중과정 모델에 있어서도 직관적 판단에 이어 도출된 추론이 합리화에 기능하였기에, 직관적 도덕 판단의 기제를 누락 없이 표현하기 위해서는 Link 2 부분의 화살표 방향이 표시되어야 할 필요성이 있다. 즉, 이중과정 모델의 도식은 확증적 추론[8]의 기제를 반영하고 있지 않으며 이에 수정이 필요하다는 것을 알 수 있다.

또한 한 가지 주의할 점은 <그림-8>의 직관적인 사회적 영향(Intuitive Social Influence) 부분은 사회적 직관주의 모델(<그림-3>)의 Link 3 '추론된 설득(the reasoned persuasion link)' 과 Link 4 '사회적 설득(the social persuasion link)' 을 포괄한다는 점이다. 사회적 직관주의 모델의 Link 3과 Link 4는 모두 판단을 도출시키는 직관에 대한 수정 기제를 규명하고 있기 때문이다. 이를 고려할 때, 이중과정 모델(<그림-7>)은 직관적인 사회적 영향(Intuitive Social Influence)과 추론적인 사회적 영향(Reasoned Social Influence)을 구분하여 표기하고 있으나, 이는 직관에 대한 사회적 영향(Social Influence to Intuition)과 추론에 대한 사회적 영향(Social Influence

8) 확증적 추론의 개념을 이해하기 위해서는, 추론을 탐구적 추론(exploratory thought)과 확증적 추론(confirmatory thought)로 분류한 러너(J. Lerner)와 테틀록(P.Tetlock)의 연구(Lerner, Tetlock, 2003)을 참고할 필요가 있다. 탐구적 추론은 대안이 될 수 있는 여러 관점을 공평하게 헤아려보는 역할을 하며, 확증적 추론은 특정 관점을 합리화하기 위하여 기울이는 일방적인 노력으로서 기능한다.

to Reason)으로 구분하여 표기하는 것이 더욱 타당하다.

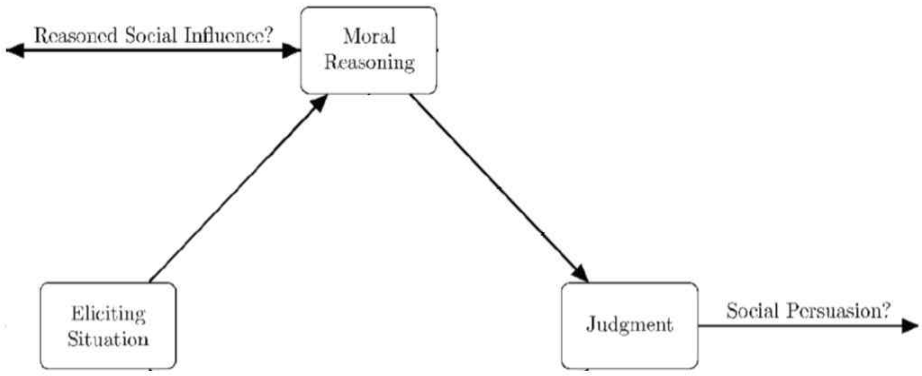
(2) 추론적 도덕 판단의 기제

그린은 추론이 도덕적 판단의 과정에 직접적으로 참여한다고 주장하였으나, 하이트는 추론은 판단 과정에 직접적으로 참여하지 않으며 직관적 판단의 합리화에 기능할 뿐이라고 주장하였다. 그러나 앞서 수행한 도덕적 판단의 범주 구체화에 기반할 때, 하이트는 모든 도덕적 상황에 있어 추론적 도덕 판단의 가능성을 부정한 것이 아니며, 피해와 공정성의 기준이 적용되지 않는 상황에 한정하여, 추론적 도덕 판단의 가능성을 부정한 것으로 보아야 한다.

하이트는 그의 연구(Haidt, Björklund, & Murphy, 2000; Haidt, 2001)에서 피해와 공정성의 기준이 적용되는 전통적인 딜레마 상황에서의 판단의 경우, 추론적 도덕 판단의 특성을 갖는다는 것을 긍정하였다. 또한 이와 대비하여, 피해와 공정성의 기준이 적용되지 않지만 불온한 느낌을 주는 상황에서의 도덕적 판단의 기제가 직감을 자극하는 상황에서의 판단 기제와 닮아있고, 그러한 경우에 있어 추론의 기능은 직관적 판단을 합리화 하는데 국한된다고 주장하였다.

그린 또한 피해와 공정성의 기준이 적용되는 상황을 전제로 직관적 반응이 강하지 않을 때, 추론에 의해 도덕적 판단이 직접 도출된다고 주장한 것이며, 피해와 공정성의 기준이 적용되지 않는 상황에 있어서까지 추론적 도덕 판단이 도출된다고 규명한 것은 아니다. 그렇기에, 두 이론은 추론적 도덕 판단의 가능성에 대하여 충돌하지 않으며, 피해와 공정성의 기준이 적용되는 상황이면서 동시에 직관적 반응이 강하게 도출되지 않을 때, 추론적 형태의 도덕 판단이 도출될 수 있다는 결론으로 수렴 가능하다. 이는 이중과정 모델(<그림-7>)의 인지적 판단 부분에 해당하며 <그림-9>와 같이 표현할 수 있다.

<그림-9> 추론적 도덕 판단 기제 (양해성, 2017, p. 134)



주목할 점은, 앞서 언급한 바와 같이, 이중과정 모델의 도식(<그림-7>)을 기준으로 살펴볼 때, <그림-8>의 Link 2와 <그림-9>의 추론적 판단의 화살표가 반대 방향을 향하고 있다는 것이다. 즉, <그림-8>에 있어서의 추론은 직관적 판단에 대한 확증적 추론을 의미하며 <그림-9>에 있어서의 추론은 피해와 공정성의 기준이 적용되는 상황에서 강한 직관적 반응이 도출되지 않는 경우에 도출되는 탐구적 추론을 의미한다.

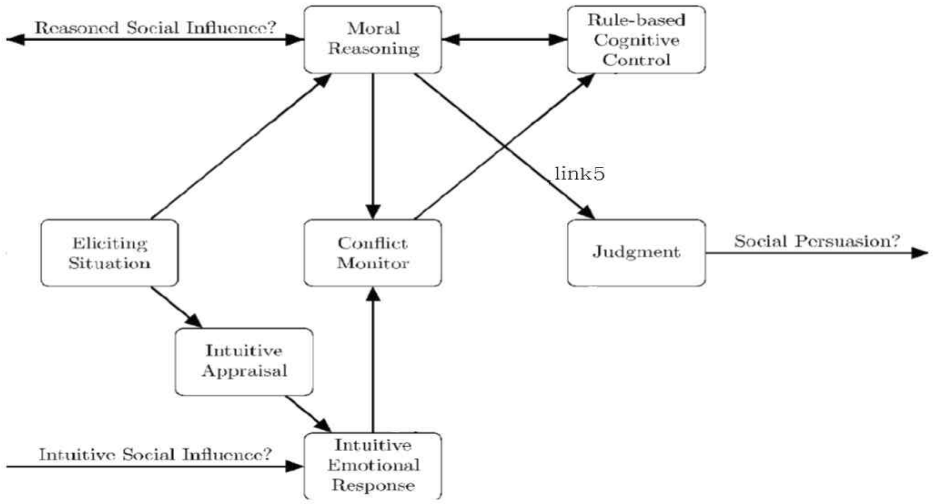
(3) 추론에 의한 직관적 판단의 수정 기제

앞서 살펴본 바와 같이, 하이트에 따르면 개인은 '추론된 설득(사회적 직관주의 모델의 Link 3)' 또는 '사회적 설득(사회적 직관주의 모델의 Link 4)을 통해 타인의 직관에 영향을 미칠 수 있다. 이 외에도 두 사회적 직관주의 이론은 타인에 의한 설득이 아닌 스스로의 추론적 심리과정에 의한 직관적 판단의 수정 가능성을 열어두었다. 하이트는 피해와 공정성의 기준이 적용되지 않는 상황에서, 매우 드물지만, 최초의 직관을 유지한 상태로 추론을 통해 수정된 판단을 도출하거나 추론의 과

정에서 새로이 활성화된 직관에 의해 초기 직관에 모순되는 새로운 판단을 도출할 수 있다고 주장하였다. 그린 또한 피해와 공정성의 기준을 적용할 수 있는 상황에 있어, 직관적 판단에 대한 인지적 조절을 통해 직접 수정된 판단을 도출하거나 인지적 조절 과정에서 촉발된 직관에 의해 수정된 판단을 도출할 수 있다고 주장하였다.

즉, 두 이론은 각각 피해와 공정성의 기준이 적용되거나 적용되지 않는 상황을 전제로, 추론에 의해 직접적으로 판단이 수정되기 위한 조건 및 새로이 도출된 직관에 의해 판단이 수정되기 위한 조건을 규명하였다고 볼 수 있다. 먼저, 직관은 유지된 채로 추론에 의해 직접 수정된 판단이 도출되는 기제는 아래 <그림-10>과 같이 표현할 수 있다.

<그림-10> 인지적 조절에 의한 직관적 판단의 수정 기제 1 (양해성, 2017, p. 135)

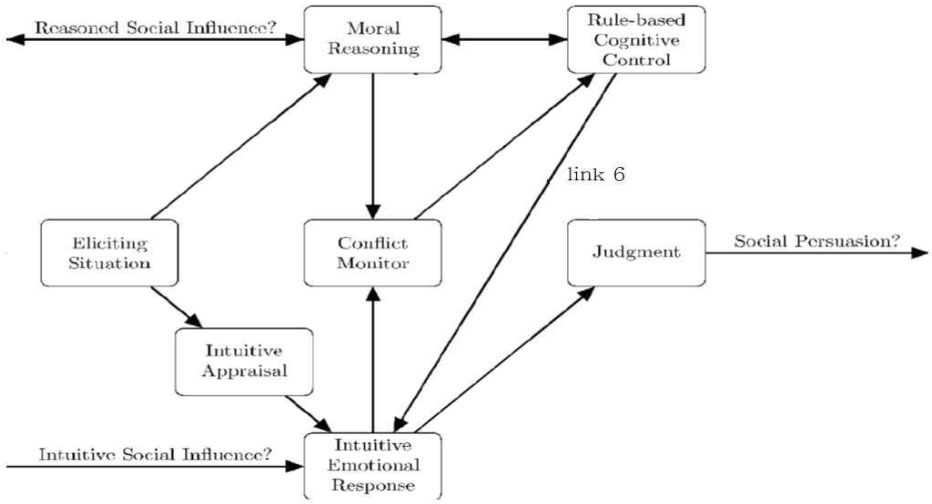


그린은 피해와 공정성의 기준이 적용되는 전통적 도덕 상황을 전제로, 강한 결과론적 근거가 존재하는 경우, 직관적 반응이 그대로 유지된 상태로 인지적 조절 과정을 거쳐 추론에 의한 직관적 판단의 수정이 이루

어질 수 있다고 주장하였다. 하이트 또한 사회적 직관주의 모델의 Link 5 '추론된 판단(the reasoned judgment link)'을 통해, 피해와 공정성의 기준을 적용할 수 없지만 직관적 판단이 도출된 상황에 있어, 초기 직관이 약하고 처리 능력이 뛰어난 경우 추론적 판단이 도출 될 수 있으며, 초기 직관과 추론이 강하게 갈등하는 경우에도 이중적 태도(dual attitude)가 수반된 채 추론에 의해 직관적 판단을 수정할 수 있다고 주장하였다.

다음으로 추론적 심리 과정을 통해 새로이 활성화된 직관에 의해 최초의 직관적 판단이 수정되는 경우에 대하여 살펴보면 아래의 <그림-11>과 같다.

<그림-11> 인지적 조절에 의한 직관적 판단의 수정 기제 2 (양해성, 2017, p. 136)



그린은 피해와 공정성의 기준이 적용되는 전통적인 도덕 상황을 전제로, 인지적 조절 과정에서 새로이 활성화된 직관에 의해 수정된 판단이 도출될 수 있다고 보았다. 또한 하이트는 피해와 공정성의 기준을 적용

할 수 없지만 직관적 반응이 도출되는 상황을 전제로, 직관적 판단에 이어진 추론의 과정에서 새로이 활성화 된 직관에 의해 수정된 판단이 도출될 수 있다고 보았다. 이는 사회적 직관주의 모델의 Link 6 '사적인 반성적 사고(the private reflection link)'에 해당하며, 하나 이상의 관점에서 상황을 접하는 과정을 통해 복합적이고 경쟁적인 직관들을 경험하는 경우 등이 그 예에 해당한다. 이 경우, 최종 판단은 가장 강한 직관에 의해 도출되거나 새로이 활성화 되어 경합하는 직관적 판단들을 대상으로 원리에 기반한 추론적 과정에 의해 도출된다. 그러나 하이트는 개인에게 존재하는 확증편향(confirmation bias)으로 인하여 이러한 관점의 전환이 흔하게 도출되지는 않는다고 덧붙였다(Haidt, 2012/2014, pp. 161-162).

살펴본 바와 같이 사회적 직관주의 모델과 이중과정 모델은 상이한 규명 범주를 전제로 하며 보완적으로 직관적 판단의 기제와 추론적 판단의 기제 및 추론에 의한 직관적 판단의 수정 기제를 규명하였다. 그렇기에 두 이론은 도덕적 판단의 기제 규명에 있어 변증법적으로 통합될 수 있고, <그림-12>와 같이 나타낼 수 있으며, 그 기제의 세부 내용은 다음과 같이 정리할 수 있다.

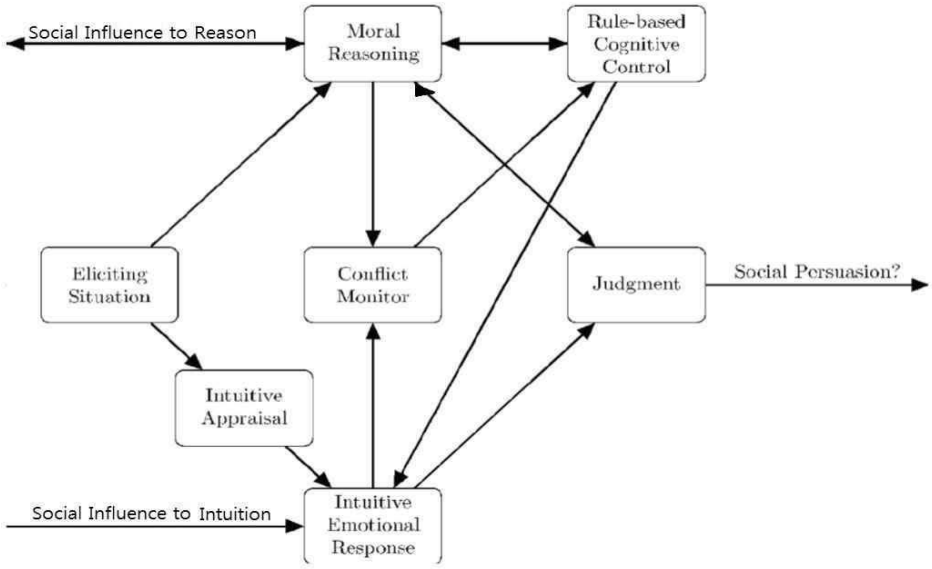
첫째, 피해와 공정성 기준의 적용 가능 여부와 무관하게, 행위에 대한 직관적 반응이 강하게 도출되는 경우에 있어, 도덕 판단은 직관적 형태로 도출되며, 추론은 직관적 판단을 합리화 하는데 기능한다. 또한 판단에 대한 근거를 제시하기 어려운 경우에도, 이러한 직관적 판단은 근거 제시 없이 유지되기도 하며 의무론적 근거에 의해 지지되기도 한다.

둘째, 피해와 공정성의 기준이 적용되는 전통적인 도덕 상황에 있어서는 직관적 판단이 도출되지 않더라도 추론에 의해 최초의 판단이 도출되지만, 피해와 공정성의 기준이 적용되지 않는 상황에 있어서는 직관적 판단 없이 추론에 의해서만 판단이 도출되지 않는다. 즉, 피해와 공정성의 기준이 적용되지 않는 상황에 있어서는, 직관적 판단이 도출되는 경우에 있어서만 도덕적 판단의 과정이 시작될 수 있으며 추론적 심리과정

도 기능을 할 수 있다.

셋째, 직관적 판단이 도출된다 하더라도, 추론에 의해 직관적 판단에 대한 수정이 가능하다. 이러한 수정은 직관이 유지된 상태로 추론을 통해 직접적으로 도출되기도 하며, 추론에 의한 인지적 조절 과정에서 새로이 형성된 직관에 의해 기존의 직관이 대체되며 도출되기도 한다. 단, 피해와 공정성의 기준이 적용되지 않는 상황에 있어 추론에 의한 직관의 수정은 매우 드물게 일어난다. 또한 피해와 공정성의 기준이 적용되는 상황에 있어서는 강한 결과론적 근거가 존재하는 경우, 추론에 의한 직관의 수정이 더욱 용이하게 이루어지기도 한다.

<그림-12> 두 사회적 직관주의 이론의 통합 가능성



2. 도덕적 인지의 유형 분류의 문제점과 대안

1) 개념 범주 명료화 필요성

상기한 바와 같이, 두 사회적 직관주의 이론은 직관과 추론이라는 인지 심리학적 개념 범주를 차용하여 도덕적 판단의 기제를 규명하였다. 또한 도덕적 판단의 현상에 대한 포괄적이고 일반적인 규명을 제시하고자 하는 시도로 인하여, 두 이론에 관한 논의들은 도덕 심리학의 현상을 대상으로 인지 심리학의 규명을 재확인하는 결과로 이어졌다. 그러나 이러한 과정에서, 과연 도덕적 판단의 현상을 타당하게 해석하기 위해, 인지 심리학적 개념 범주를 차용하는 것이 적절한지 여부에 대한 검토가 충분히 이루어지지 않았다. 즉, 인지 심리학적 개념 범주를 차용할 경우 모순이 발생할 위험성이 없는지에 대한 검토가 이루어지지 않았고, 도덕적 판단의 현상을 더욱 타당하게 해석하기 위해 인지 심리학적 개념 범주를 보완할 수 있는 방안이 없는지에 대한 검토 또한 이루어지지 않았다.

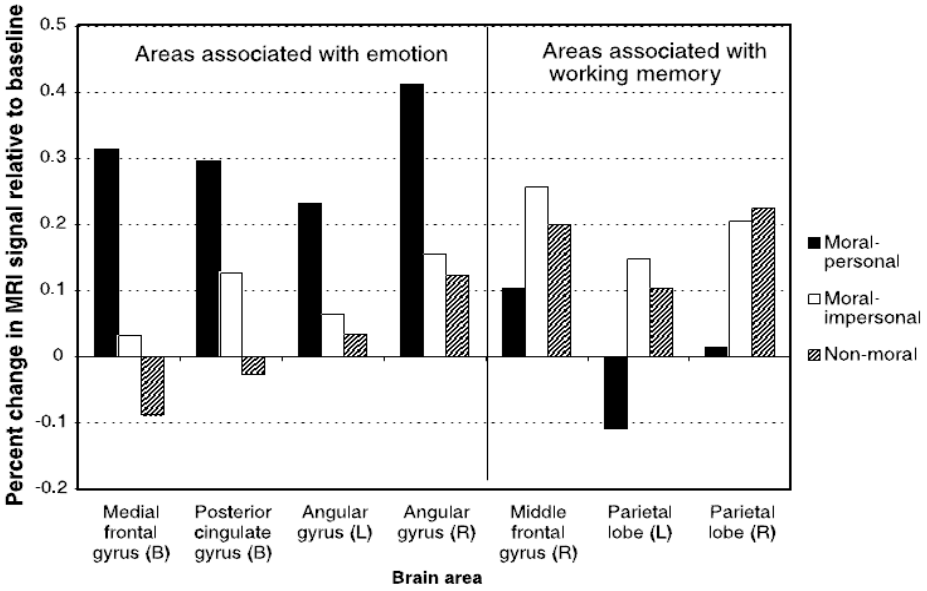
상기한 바 있듯, 차마즈(K. Charmaz)는 복합적이고 다의적인 해석의 가능성을 갖는 용어의 엄밀하지 못한 사용이 이론의 의의를 손상시킨다(Charmaz, 2006/2013, p. 292)고 지적하였다. 즉, 두 사회적 직관주의 이론이 인지 심리학적 개념 범주 즉, 직관과 추론이라는 개념 범주를 활용할 경우, 도덕적 판단이라는 현상을 해석함에 있어 오류가 발생할 위험성이 없는지 살펴보아야 한다. 여러 특성을 동시에 갖는 하나의 현상이 직관과 추론의 특성을 모두 가짐으로 인하여 다양하게 해석될 여지가 있다면, 이러한 개념 범주의 활용은 이론 내부의 모순으로 이어질 위험성을 내재하게 된다.

즉, 만일 두 사회적 직관주의 이론이 현상을 규명함에 있어, 추론적 판단으로 분류한 현상이 인지 심리학의 분류 기준을 따를 때 직관적 인지로 분류될 여지가 있다면 또는 그와 반대로 직관적 판단으로 분류한 현상이 인지 심리학의 분류 기준을 따를 때 추론적 인지로 분류될 여지

가 있다면, 해당 현상을 설명하기 위해 직관과 추론이라는 인지 심리학적 해석 틀을 차용하는 것은 규명상의 한계를 가질 수밖에 없다.

이러한 한계가 그린의 연구(Greene, 2001)에서 목격된다. 그린은 대뇌의 부위별 활성화 정도(<그림-13>; Greene, 2001, p. 2106) 및 딜레마 유형별 응답 소요 시간 그래프(<그림-5>; Greene, 2001, p. 2107)를 근거로, 비인신적 도덕 딜레마 상황에서의 판단이 덜 감정적이고 더 인지적인 특성을 갖는 추론적 판단에 해당하며 인신적 도덕 딜레마 상황에서의 판단이 더 감정적이고 덜 인지적인 특성을 갖는 직관적 판단에 해당한다고 분류하였다. 즉, 그린의 연구에 있어 직관과 추론을 구분하는 기준은 감정 반응의 강한 정도이다.

<그림-13> 딜레마 유형에 따른 대뇌 부위별 활성화 정도(Greene, 2001, p. 2106)



또한 대뇌의 부위별 활성화 정도 또는 유형별 딜레마에 있어 긍정적

응답에 걸리는 시간의 차이를 기준으로 할 때, 도덕과 무관한 딜레마(Non-Moral Dilemma) 상황에서의 반응 양상은 직관으로 분류된 인신적 도덕 딜레마 상황에서의 반응보다는 추론으로 분류된 비인신적 도덕 딜레마 상황에서의 반응과 유사성을 갖는다. 즉, 그린의 분류에 따를 때, 도덕과 무관한 딜레마 상황에서의 판단은 덜 감정적이고 더 인지적인 추론적 판단에 가까운 것으로 분류할 수 있다(Greene, 2001, p. 2107).

그러나 도덕과 무관한 딜레마 상황의 예로 제시된 문항들을 살펴보면, 인지 심리학의 기준을 적용할 때, 해당 상황에 대한 판단을 직관적 판단이 아닌 추론적 판단으로 분류하는 것이 타당하다고 단정하기에는 어려움이 있어 보인다. 다음은 도덕과 무관한 딜레마 상황으로 제시된 문항의 예시이다.

> [제네릭 약품 딜레마]
> 당신은 두통이 있어서 약국을 방문하였다. 오랜 기간 동안 알아와 당신과 신뢰가 두터운 약사는 당신이 구매하려고 했던 상품명의 약품의 재고가 없기에 성분과 효과가 완전히 동일한 제네릭 약품을 복용하라고 권하였다. 원래 사려던 상품명의 약품이 아닌 제네릭 약품을 구입하는 것은 적절한가.

> [브라우니 딜레마]
> 당신은 자신이 먹을 브라우니를 만들려고 하고 있다. 요리책을 열어보니 브라우니에 호두를 넣으라고 나와 있다. 당신은 호두를 싫어하기에 호두 대신 마카다미아 너트를 넣으려고 한다. 마카다미아 너트를 넣는 것은 적절한가.

> [튤립 딜레마]

당신은 튤립 농장에서 수확 기계를 운전하고 있다. 당신 앞에는 두 갈래의 갈림길이 있다. 왼쪽으로 갈 경우, 수확 기계는 10부셸의 튤립을 수확한다. 오른쪽으로 갈 경우, 수확 기계는 20부셸의 튤립을 수확한다. 만일 당신이 아무런 조작을 하지 않는다면 수확 기계는 왼쪽으로 가게 된다. 당신이 수확 기계를 왼쪽으로 가게 놔두어 10부셸을 수확하는 대신 오른쪽으로 가게 하여 20부셸을 수확하게 하는 것은 적절한가.

[고장난 VCR 딜레마]
당신은 고장난 VCR을 가지고 수리점에 왔다. 점원은 수리비로 100달러가 나올 것이라고 하였다. 당신은 수리점에 오는 길에 가전용품점에서 당신의 VCR보다 더 최신 모델의 VCR을 할인하여 100달러에 판매하고 있는 것을 보았다. 새로운 VCR을 구입하는 대신 당신의 VCR을 수리하는 것은 적절한가.

직관과 추론의 특성을 비교한 <표-1>을 참고할 때, 감정 반응이 강하게 도출되었다는 것을 이유로 도덕과 무관한 딜레마 상황에서의 판단을 직관적 판단이 아닌 추론적 판단으로 분류해야 한다고 단언할 수 있을까. 또한 이러한 상황에서의 판단이 바흐(J. Bargh)가 직관적 판단의 특징으로 제시한 비노력성(처리에 있어 노력을 필요로 하지 않음) 또는 효율성(최소한의 인지적 용량으로 처리가 이루어짐)과 같은 특성을 보이지 않는다고 단언할 수 있을까. 특정 판단이 도출되는 과정에 있어, 감정과 연관된 대뇌 부위에 비하여 인지와 연관된 대뇌의 부위가 활성화 되었다는 이유로 추론적 판단으로 분류하고 감정과 관련된 대뇌의 부위가 활성화 되었다는 이유로 직관으로 분류하는 것은 직관과 추론의 여러 특성을 충분히 고려하지 않은 성급한 분류의 위험성을 내포한다.

또한, 딜레마 유형별 응답 소요 시간 그래프를 기준으로 할 때, 그린이 추론적 판단으로 분류한 비인신적 딜레마(Moral Impersonal Dilemma) 상황에서의 긍정적 응답 도출에 소요된 시간이 그가 직관적 판단으로 분류한 인신적 딜레마(Moral Personal Dilemma) 상황에서의 부정적 응답 도출에 소요된 시간보다 짧다는 점도, 감정의 활성화 정도만을 가지고 특정한 판단을 직관 또는 추론으로 분류하는 것이 모순된 해석의 위험성을 내포하는 성급한 분류라는 것을 보여준다. 수단을 제외하고 동일한 조건으로 구성된 두 상황을 전제로 도출된 이러한 실험 결과는, 직관적 판단이 신속한 판단으로서의 특성을 갖고 추론적 판단이 느린 판단으로서의 특성을 갖는다는 통설과 정면으로 충돌한다. 그린이 추론으로 분류한 비인신적 딜레마 상황에서의 판단은 덜 감정적이고 더 의식적이며, 판단의 과정을 파악할 수 있다는 추론적 판단의 특성들을 가지지만, 동시에 신속하다는 직관적 판단의 특성도 갖는다. 그렇기에 이는 추론이라 단정하기에도 어려움이 있으며 직관이라 단정하기에도 어려움이 있다.

하이트의 사회적 직관주의 모델에 있어서도 이러한 모순된 해석의 위험성이 존재한다. 하이트가 추론으로 분류한 인지 또한, 상기한 바와 같이, 숙달 정도와 자동성 등에 따라 직관으로도 분류할 수 있고 추론으로도 분류될 수 있다. 이렇듯, 직관과 추론이라는 개념 범주의 활용으로 인하여 모순된 해석의 위험성이 내재하기에 이를 보완하기 위한 개념 범주를 모색할 필요가 있으며, 이를 위해 연구자가 직관과 추론이라는 개념 범주를 통해 어떠한 실제를 지칭하고자 하였는지 도출하는 작업이 수행되어야 한다. 이론에서 얻은 이해는 현상에 대한 이론가의 해석에 의존하기에(Charmaz, 2006/2013, p. 254) 연구자의 가정과 전제에 대해 인지하고 그러한 가정과 전제가 연구에 영향을 미치는 방식을 분석하는 것은 이론을 보다 명료하게 이해함에 있어 매우 중요하다(Charmaz, 2006/2013, p. 263). 즉, 그린의 이중과정 모델이 규명한 도덕적 판단의 기제를 보다 타당하게 해석하기 위해서는 그가 설정한 개념 범주가

실제 무엇을 지칭하는지 규명하고 그에 기반하여 개념 범주를 보완 또는 대체할 필요성이 있다.

주목할 점은 두 이론의 개념 범주 명료화 작업을 각각 진행하기보다는 두 이론 모두에 적용 가능한 개념 범주를 도출해야 한다는 것이다. 앞서 살펴본 바와 같이, 두 이론은 도덕적 판단이라는 상위의 현상 범주를 구성하는 상이한 하위의 현상 범주를 규명 대상으로 하는 보완적 관계로 파악하는 것이 타당하다. 그렇기에 두 이론이 보완적으로 기능하며 도덕적 판단의 기제를 통일성 있게 규명하기 위해서는 두 이론에 모두 적용 가능한 개념 범주를 마련해야 할 필요성이 있다.

2) 개념 범주 명료화 가능성

앞선 논의를 통해, 사회적 직관주의 이론이 규명한 현상을 보다 타당하게 해석하기 위해 인지 심리학적 개념 범주로서의 직관 및 추론과 구분되는 고유의 개념 범주를 활용해야 할 필요성이 있음을 도출하였다. 그렇다면 두 사회적 직관주의 이론이 사용하고 있는 직관과 추론이라는 개념 범주는 실제 무엇을 지칭하는 것일까. 먼저, 이는 두 이론이 도덕적 판단의 기제 규명에 있어 강조하고 있는 '직관' 개념에 대한 비교를 통해 살펴볼 수 있다.

하이트와 그린이 사용하고 있는 직관 개념은 나바에츠(D. Narvaez)가 사용하고 있는 직관 개념과의 비교를 통해 보다 명료하게 이해할 수 있다. 나바에츠는 드레이퍼스(H. Dreyfus)와 드레이퍼스(S. Dreyfus)가 제시한 기술 습득의 다섯 단계(Dreyfus & Dreyfus, 1990)에 기준하여, 모든 분야에서 학습자들이 초보적 이해에서 전문가적 이해로 이행해 간다고 보았고(Narvaez, 2004, p. 93), 통합적 윤리교육 모델(Integrative Ethical Education)을 통해 도덕적으로 가장 적합한 행위를 직관적으로 판단할 수 있는 도덕적 전문가(Moral Expert)를 양성하는 것을 도덕 교육의 목표로 삼았다. 그녀가 사용한 직관 개념은 학습

및 정보 처리에 관한 인지 심리학의 직관 개념과 유사성을 갖는다.

드레이퍼스와 드레이퍼스(Dreyfus & Dreyfus, 1990, pp. 239-241)에 따르면, 첫 번째 단계인 초보자(Novice)는 선행 경험이 없기 때문에 상황에 대한 전체적인 인지 없이, 맥락 독립적으로 상황을 분석하고, 교육을 통해 제공받은 새로운 기술과 규칙에 기초하여 행동을 결정한다. 두 번째 단계인 준 초보자(Advanced Beginner)는 실제 상황을 대처하는 경험에서 교육자의 지도나 스스로의 지각을 통해 유의미한 추가적인 구성요소를 인지하기 시작한다. 맥락 독립적으로 규칙을 적용하는 초보자와 달리 준 초보자는 상황 맥락적인 규칙에 의존한다. 세 번째 단계인 미숙련자(Competence)는 경험의 증가에 따라 고려 요소가 많아지고 이로 인하여 의사 결정의 위계적 우선순위를 설정한다. 네 번째 단계인 숙련자(Proficiency)는 자신의 행동을 안내해줄 원리를 찾거나 문제 상황을 자신과 분리된 것으로 여기지 않으며, 수많은 경험에 기반하여 계획을 선택하고 선택의 적절성을 감정적으로 경험함으로써 특정한 계획과 목표 및 관점을 선택한다. 숙련자는 당면한 상황 안에서 자신이 해야 할 일을 직관적으로 이해하지만, 무엇을 해야 하는지를 결정하는 과정에서는 여전히 분석적으로 접근한다. 다섯 번째 단계인 전문가(Expertise)는 충분한 연습과 경험을 통한 이해를 토대로 자신을 둘러싼 상황 안에서 목표와 계획, 전략을 세우고 실행에 옮긴다. 이러한 과정은 의식적인 노력 없이 즉시적이고 직관적인 암묵적 과정을 통해 이루어진다.

<표-4> 기술 습득의 5단계 (Dreyfus & Dreyfus, 1986, p. 50)

기술 단계	구성 요소	관점	의사 결정	숙고
초보자 (Novice)	맥락독립적	없음	분석적	분리된 숙고 (원리에 호소)
준초보자	상황적	없음	분석적	분리된 숙고

(Advanced Beginner)				(원리에 호소)
미숙련자 (Competence)	상황적	선택	분석적	분리된 이해 수반된 결과
숙련자 (Proficiency)	상황적	경험	분석적	수반된 이해 분리된 결정
전문가 (Expertise)	상황적	경험	직관적	수반된 숙고

즉, 나바에츠가 통합적 윤리교육 모델에서 제시하고 있는 직관 개념은 가능한 선택지 중 가장 도덕적인 판단을 내리는 정보 처리 과정에 있어서의 자동성 향상 즉, 도덕적 판단의 수월성을 의미하는 개념이다. 이러한 수월성으로서의 직관 개념은 그녀가 제시한 통합적 윤리교육 모델(Integrative Ethical Education)에서도 찾아볼 수 있다. 나바에츠는 도덕적 전문가 양성을 목표로 레스트(J. Rest)의 4구성요소 모형(Rest, 1983)에 근거하여 도덕성을 구성하는 네 가지 요소들(도덕적 민감성, 도덕적 판단, 도덕적 동기화 혹은 초점, 도덕적 행동)을 제시한 후 각 요소별로 도덕적 전문성을 신장시키기 위한 일련의 기술들(skills)을 제시한다. 도덕성의 네 가지 요소를 구성하는 윤리적 기술들에 대한 교수 과정은 상기한 기술 습득의 단계를 활용한 네 단계 수준에 따라 전개된다(Narvaez & Bock, 2014, p. 151).

즉, 나바에츠의 직관 개념은 도덕적 판단과 행위를 구성하는 요소와 요소별 기술에 대한 수월성을 의미하는 형식적 차원의 개념이다. 나바에츠와 인지 심리학의 직관 개념은 불승인의 감정 반응과 같은 내용 요소를 필요로 하지 않는다. 반면, 하이트와 그린이 사용하고 있는 직관 개념은 '불승인의 강한 감정 반응'이라는 내용 요소를 특성으로 한다. 하이트는 피해와 공정성의 기준이 적용되지 않지만 금기(taboo)를 위반한 행위에 대한 불승인의 감정 즉, 부정적 직감(Gut Feeling)을 직관으로 표현하고 있으며, 그린은 피해와 공정성의 기준을 적용할 때 동일한

결과를 가져오는 행위 중 특정 선택지에 대한 불승인의 감정을 직관으로 표현하고 있다. 즉, 인지 심리학 및 나바에츠의 직관 개념이 내용적으로 열려있는 형식적 차원의 개념으로서의 수월성에 해당하는 것과 달리, 두 사회적 직관주의 이론의 직관 개념은 자동성과 같은 형식적 차원의 특성을 가지면서도 동시에 행위에 대한 불승인이라는 내용적 차원의 개념으로서의 특성 또한 갖는다. 그렇기에 두 사회적 직관주의 이론은 도덕적 판단에 있어 '직관'과 '추론'의 기제를 규명한 이론으로 해석하기보다는 '불승인 직관'과 '추론으로 표현된 인지'의 기제를 규명한 이론으로 해석할 필요성이 있다.

그렇다면 두 사회적 직관주의 이론의 개념 범주를 명료화함에 있어, 이러한 '추론으로 표현된 인지' 즉, '불승인 직관 외의 인지'는 어떠한 것을 의미할까. 우선, 그린의 이중과정 모델을 대상으로 살펴보면 다음과 같다. 살펴본 바와 같이 그린의 이중과정 모델이 현상 규명에 있어 인지심리학적 용어인 직관과 추론이라는 형식적 차원의 개념 범주를 사용하는 것은 내부적인 모순으로 이어질 위험성을 내포한다. 또한 이중과정 모델에 있어 직관으로 분류되었던 '불승인 직관'의 경우 내용과 형식을 특정할 수 있지만, 그 외의 인지는 숙달 정도와 자동성 등에 따라 직관으로도 분류할 수 있으며 추론으로도 분류할 수 있다. 그렇기에 이중과정 모델의 규명을 명료하게 해석하기 위한 개념 범주로서의 추론 즉, '불승인 직관 외의 인지'는 인지 심리학적 개념 범주인 직관이나 추론과 같이 형식적 차원의 특성을 갖는 개념 범주가 아니면서도, 불승인 직관과 구분되는 내용적인 특성을 갖는 개념 범주일 필요성이 있다.

이중과정 모델의 '불승인 직관 외의 인지'에 대하여 자세히 살펴보면 다음과 같다. 먼저, 그린은 직관적 판단의 예로 인도교 딜레마 상황에서의 판단을 제시하며, 직관은 판단에 기능하고 추론은 합리화에 기능한다고 보았다. 즉, 피해와 공정성의 기준이 적용되는 전통적인 도덕 상황에 있어 '직관적 판단'은 '직관'과 '확증적 추론'에 의해 구성된다고 볼 수 있다. 또한 상기한 명료화 작업에 기반 할 때, 이러한 직

관은 형식적으로 그리고 내용적으로 특정된 '불승인 직관'으로 보완하는 것이 타당하며, 확증적 추론은 형식적 차원에서 특정되지 않고 내용적 차원의 특성만을 갖는 개념 범주로서 '확증적 인지'[9]로 대체하는 것이 타당하다. 즉, 이중과정 모델이 규명한 직관적 판단의 기제에 있어 '불승인 직관 외의 인지'는 불승인 직관을 합리화 하는 '확증적 인지'라고 할 수 있다.

그렇다면 직관적 판단의 상황과 대비되는 추론적 판단의 상황으로 제시된 비인신적 도덕 딜레마 상황에서의 판단에 있어 '불승인 직관 외의 인지'는 무엇을 의미할까. 직관적 판단과 대조되는 트롤리 딜레마 상황에서의 추론적 판단은 가장 도덕적인 행위 방안을 모색하는 '탐구적 추론'과 그에 대한 '확증적 추론'에 의해 구성된다고 볼 수 있다. 또한 선행한 명료화 작업에 기반할 때, 이러한 탐구적 추론과 확증적 추론은 형식적 차원이 특정되지 않은 내용적 차원의 개념 범주로서 '탐구적 인지'와 '확증적 인지'로 대체하는 것이 타당하다. 즉, 이중과정 모델이 규명한 추론적 판단의 기제에 있어 '불승인 직관 외의 인지'는 '탐구적 인지'와 '확증적 인지'라고 할 수 있다.

덧붙여, 직관적 판단과 추론적 판단이 경합하는 경우 즉, 강한 결과론적 근거가 존재할 때 직관적 판단에 대한 인지적 조절이 이루어지는 경우로 제시된 우는 아이 딜레마(crying baby dilemma) 상황에서의 판단은 '불승인 직관'과 그에 대한 '확증적 인지' 그리고 가장 도덕적인 판단의 도출을 목적으로 하는 '탐구적 인지'와 그에 대한 '확증적 인지'가 경합하며 조율되는 과정으로 볼 수 있다.

종합하면, 그린이 추론으로 분류한 '불승인 직관 외의 인지'는 형식적 차원에서 특정되지 않으면서도 '탐구적 인지'와 '확증적 인지'의

9) 확증적 인지와 탐구적 인지는 앞서 살펴본 러너(J. Lerner)와 테틀록(P. Tetlock)의 추론에 대한 분류를 응용한 개념이다. 계속 언급하고 있듯, 감정의 활성화 정도만을 가지고 직관과 추론을 구분하는 것은 이론 내부의 모순을 유발할 수 있기에, 충분한 검토와 근거 제시 없이 직관 또는 추론이라는 용어를 사용하는 것을 지양하고, 형식적 차원의 두 개념을 모두 포괄하는 '인지'라는 개념을 탐구와 확증이라는 내용적 분류에 접목시켰다.

내용적 특성을 포괄하는 개념 범주여야 한다. 또한 그린에게 있어 추론적 판단은 '가장 도덕적인 행위를 모색하는 것을 목적으로 하는' 탐구적 인지와 '탐구적 인지에 의한 판단을 합리화하는 것을 목적으로 하는' 확증적 인지로 구성되며, 직관적 판단은 불승인 직관과 '불승인 직관에 의한 판단을 합리화하는 것을 목적으로 하는' 확증적 인지로 구성된다. 그렇기에, 탐구적 인지와 확증적 인지의 내용적 특성을 아우르는 '불승인 직관 외의 인지'는 '목적 지향적 인지'로 범주화 가능하다. 즉, 도덕적 판단의 기제 규명에 관한 이중과정 모델을 보다 타당하게 해석하기 위해서는 직관과 추론이라는 개념 범주 대신, '불승인 직관'과 '목적 지향적 인지'라는 개념 범주를 활용할 필요성이 있다. 즉, 그린의 이중과정 모델은 피해와 공정성의 기준이 적용되는 인지주의 전통의 도덕적 판단 상황에 있어, '불승인 직관'과 '목적 지향적 인지'의 기제를 규명한 것이라고 보아야 한다.

그렇다면 하이트의 사회적 직관주의 모델에 있어 '불승인 직관 외의 인지'는 어떠한 것을 의미할까. 앞서 언급한 바와 같이, 사회적 직관주의 모델에 있어서도 추론으로 분류된 개념 범주가 그 수월성 또는 자동성 정도에 따라 다의적으로 해석될 위험성이 존재하기에 이에 대한 대체 개념의 마련이 필수적으로 요청된다. 이에 하이트의 사회적 직관주의 모델에 있어 '불승인 직관 외의 인지'가 의미하는 바에 대하여 살펴보면 다음과 같다.

우선, 하이트의 사회적 직관주의 모델(<그림-3>)에 있어서도 Link 1 '직관적 판단(the intuitive judgment link)'은 '불승인 직관'으로 형식적 차원과 내용적 차원을 특정할 수 있다. 또한 불승인 직관인 Link 1과 짝을 이루는 Link 2 '사후 추론(the post-hoc reasoning link)'은 형식적 차원에서 특정되지 않고, 불승인 직관에 대한 합리화를 그 내용으로 하기에 '확증적 인지'로 표현할 수 있다. 또한 Link 3 '추론된 설득(the reasoned persuasion link)'과 Link 4 '사회적 설득(the social persuasion link)'은 이러한 '불승인 직관' 또는 '확증

적 인지’를 상대방에게 전달함으로써 타인의 판단에 영향을 미치는 기제라고 볼 수 있다.

또한 Link 5 ‘추론된 판단(the reasoned judgment link)’은 ‘불승인 직관’이 유지된 상태로, ‘불승인 직관 외의 인지’가 활성화 되어 수정된 판단이 도출되는 심리 과정이며, Link 6 ‘사적인 반성적 사고(the private reflection link)’는 ‘불승인 직관 외의 인지’에 의해 최초의 불승인 직관과 충돌하는 새로운 직관이 활성화됨으로써, 최초의 ‘불승인의 직관적 판단’이 수정되는 과정이다.

하이트의 사회적 직관주의 모델에 있어 ‘불승인 직관 외의 인지’는 Link 2의 ‘확증적 인지’와 함께 Link 5 ‘추론된 판단(the reasoned judgment link)’과 Link 6 ‘사적인 반성적 사고(the private reflection link)’에 기능하는 심리 과정을 모두 포괄하는 것이어야 한다. Link 2의 ‘확증적 인지’는 Link 1의 불승인 직관을 합리화 하는 것을 목적으로 하기에 그린의 경우와 마찬가지로 ‘목적 지향적 인지’로 범주화 할 수 있다. Link 5와 Link 6의 ‘불승인 직관 외의 인지’ 또한 비도덕적인 ‘불승인 직관’의 경계와 수정을 목적으로 하기에 ‘목적 지향적 인지’로 범주화 가능하다. 즉, 사회적 직관주의 모델에 있어 추론 즉, ‘불승인 직관 외의 인지’는 이중과정 모델의 경우와 마찬가지로 ‘목적 지향적 인지’로 대체할 수 있다. 또한 이러한 목적에는 도덕적으로 더욱 타당한 판단을 모색하는 것[10]이나 편견에 근거한 판단이나 왜곡된 판단을 내리지 않으려는 것[11]과 같은 도덕적 목적이 포함될 수 있으며 타인에게 좋은 평판을 받는 등의 도덕 외적인 목적[12]도 포함될 수 있다.

---

10) 하이트의 언급을 활용할 때, 직관적 판단의 합리화가 아닌 판단의 타당성을 검토하는 등에 기능하는 추론은 더욱 타당한 판단을 모색하는 것을 목적으로 한다.

11) 앞서 언급한 후의식적 자동처리 과정에 의해 도출된 추론 또는 역할 소유와 같은 인지 과정 등은 편견에 근거한 판단이나 왜곡된 판단을 피하는 것을 목적으로 할 수 있다.

12) 이중적 태도를 수반하는 추론의 경우, 판단의 내용보다는 판단의 발화에 따른 상대의 반응과 같은 도덕 외의 목적을 전제할 수 있다.

즉, 사회적 직관주의 모델은 피해와 공정성의 기준이 적용되지 않는 도덕적 판단의 상황에 있어, ‘불승인의 직관적 판단’과 ‘목적 지향적 판단’의 기제를 규명한 것이라고 해석해야 한다. 주의할 점은 사회적 직관주의 모델에서의 ‘목적 지향적 판단’은 이중과정 모델에서의 ‘목적 지향적 판단’과 달리, ‘불승인의 직관적 판단’이 도출된 후에만 후속적으로 발현될 수 있다는 것이다.

종합하면, 피해와 공정성의 기준이 적용되는 상황과 피해와 공정성의 기준이 적용되지 않는 상황을 각각의 규명 범주로 하는 두 사회적 직관주의 이론의 규명을 타당하게 해석하기 위해서는, 형식적 차원과 내용적 차원이 특정된 개념 범주로서 ‘불승인 직관’과 형식적 차원에서 열려 있고 내용적 차원에서 특정된 개념 범주로서 ‘목적 지향적 인지’를 활용할 필요성이 있다.

3) 개념 범주 명료화의 적용

살펴본 바와 같이 도덕 심리학이 현상을 해석함에 있어 인지 심리학의 개념 범주를 그대로 차용하는 것은 한계를 가질 수밖에 없다. 이에 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 타당하게 해석하기 위해, 인지 심리학적 개념 범주로서 직관과 추론을 대체할 수 있는 ‘불승인 직관’과 ‘목적 지향적 인지’라는 개념 범주를 도출하였다. 이를 반영할 경우, 선행한 규명 범주 구체화 작업에 따른 사회적 직관주의 이론의 발견은 다음과 같이 수정되어야 한다.

첫째, 피해와 공정성 기준의 적용 가능 여부와 무관하게, 행위에 대한 ‘불승인 직관’이 강하게 도출되는 경우에 있어, 도덕 판단은 불승인의 직관적 판단으로 도출된다. 또한 이러한 불승인의 직관적 판단은 근거 제시 없이 유지되기도 하며 의무론적 근거에 의해 지지되기도 한다.

둘째, 피해와 공정성의 기준이 적용되는 전통적인 도덕 상황에 있어서는 ‘불승인 직관’이 도출되지 않는다 하더라도 ‘목적 지향적 인지’

에 의한 목적 지향적 판단의 도출이 가능하지만, 피해와 공정성의 기준이 적용되지 않는 상황에서는 '불승인 직관'이 도출되지 않는다면 어떠한 도덕적 판단도 도출되지 않는다. 피해와 공정성의 기준이 적용되지 않는 상황에서는 '불승인 직관'이 존재하는 경우에만 도덕적 판단의 심리 과정이 시작될 수 있다.

셋째, 불승인의 직관적 판단이 도출된다 하더라도 '목적 지향적 인지'에 의해 수정이 이루어질 수 있다. 이러한 판단의 수정은 불승인 직관이 유지된 상태로 이루어지기도 하며, 목적 지향적 인지에 의해 점화된 인지적 조절 과정에서 새로이 형성된 '불승인 직관'이 기존의 '불승인 직관'을 대체하며 도출되기도 한다. 단, 피해와 공정성의 기준이 적용되지 않는 상황에 있어, 목적 지향적 인지에 의한 불승인의 직관적 판단에 대한 수정은 매우 드물게 일어난다. 또한 피해와 공정성의 기준이 적용되는 상황에 있어서는 강한 결과론적 근거가 존재하는 경우, 목적 지향적 인지에 의한 불승인 직관의 수정이 더욱 용이하게 이루어지기도 한다.

또한 선행한 작업에 기반할 때, 두 사회적 직관주의 이론의 개념 범주로서 '불승인 직관'과 '목적 지향적 인지' 및 인지 심리학적 개념 범주로서 '직관'과 '추론'의 관계는 <표-5>와 같이 나타낼 수 있다.

<표-5> 인지 심리학과 사회적 직관주의 이론의 인지 유형 분류 비교

인지 심리학	추론	직관
사회적 직관주의 이론	목적 지향적 인지	불승인 직관

유형별 인지의 특성	느림	자동적이고 반사적이고 신속함
	의도적인 인지적 노력을 필요로 함	의도적인 인지적 노력이 필요하지 않음
	의식적으로 과정을 파악할 수 있음	과정을 파악하기 어렵고 단지 결과만을 파악할 수 있음
	주의 깊은 자료원 (resources)이 요구됨	주의 깊은 자료원(resources)이 요구되지 않음
	맥락(context) 독립적	맥락(context) 의존적
	목적 종속적	목적 독립적
	지향적 형태, 적극적, 탐색적	지양적 형태, 소극적, 통제적
	상대적으로 불승인 감정 반응에 덜 영향 받음	불승인 감정 반응에 크게 영향 받음
	상대적, 결과론적	절대적, 의무론적

이에 대하여 살펴보면 첫째, ‘목적 지향적 인지’는 설정된 목적 즉, 가장 도덕적인 판단의 도출과 같은 목적에 종속되는 특성을 갖는 반면, ‘불승인 직관’은 설정된 목적에 종속되지 않는 독립적인 특성을 갖는다. 즉, 이중과정 모델이 규명한 도덕적 판단의 현상을 전제로 살펴보면, 피해와 공정성의 기준이 적용되는 인지주의 전통의 도덕적 상황에 있어 ‘목적 지향적 인지’가 가장 도덕적인 행위의 선택이라는 설정된 목적

에 종속되어 목적 달성을 위한 최선의 방안을 도출하는 역할을 담당하는 것과 달리, ‘불승인 직관’은 설정된 목적과 독립적으로 특정 행위의 허용 불가능성 여부에 대한 판단만을 담당한다. 피해와 공정성의 기준이 적용되지 않는 상황 즉, 사회적 직관주의 모델이 규명한 도덕적 판단의 현상에 있어서도, ‘불승인 직관’은 행위의 허용 불가능성을 판단하며 ‘목적 지향적 인지’는 불승인의 직관적 판단에 대한 합리화 또는 수정이라는 설정된 목적에 종속되어 기능한다.

이러한 맥락에서 둘째, ‘목적 지향적 인지’는 지향적인 형태를 가지며 적극적이고 탐색적인 특성을 갖는다. 이는 더 도덕적인 행위 방식에 대한 탐색일 수도 있으며 불승인의 직관적 판단의 합리화 근거에 대한 탐색일 수도 있다. 반면, ‘불승인 직관’은 지양적인 형태를 가지며 소극적이고 통제적인 특성을 갖는다. 셋째, ‘불승인 직관’은 불승인의 감정 반응에 크게 영향을 받지만 ‘목적 지향적 인지’는 이러한 불승인의 감정에 크게 영향을 받지 않거나 경쟁할 수 있다. 넷째, ‘불승인 직관’은 의무론적 근거에 의해 지지되거나 근거를 제시할 수 없는 경우에도 유지되지만, ‘목적 지향적 인지’는 결과론적 근거에 의해 지지된다.

Ⅴ. 사회적 직관주의 이론 비판 및 재해석의 도덕 심리학적 적용

앞선 작업을 통해 첫째, 두 사회적 직관주의 이론의 규명 범주를 구체화함으로써 두 이론이 도덕적 판단을 구성하는 상이한 하위의 범주 즉, 피해와 공정성의 기준이 적용되는 상황과 적용되지 않는 상황을 보완적으로 설명하고 있다는 결론을 도출하였다. 둘째, 두 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 모순 없이 타당하게 해석하기 위해서는 인지 심리학적 개념 범주인 직관과 추론이 아닌, 사회적 직관주의 이론 고유의 개념 범주로서 불승인 직관과 목적 지향적 인지를 활용해야 할 필요성이 있음을 도출하였다.

본 장에서는 이러한 비판적 검토 작업의 도덕 심리학적 적용 가능성을 도출하고자 한다. 이에 첫 번째 절에서는 선행한 작업에 기반하여, 두 사회적 직관주의 이론의 규명을 비판적으로 재해석 하는 작업을 수행할 것이다. 두 번째 절에서는 선행한 작업에 기반하여 유형별 인지의 상호 보완 방식을 명료화 할 것이다. 이는 인지 심리학적 인지 유형인 직관과 추론의 상호 보완 방식에 대한 규명 작업과 사회적 직관주의 이론의 인지 유형인 불승인 직관과 목적 지향적 인지의 상호 보완 방식에 대한 규명 작업으로 나누어 수행할 것이다.

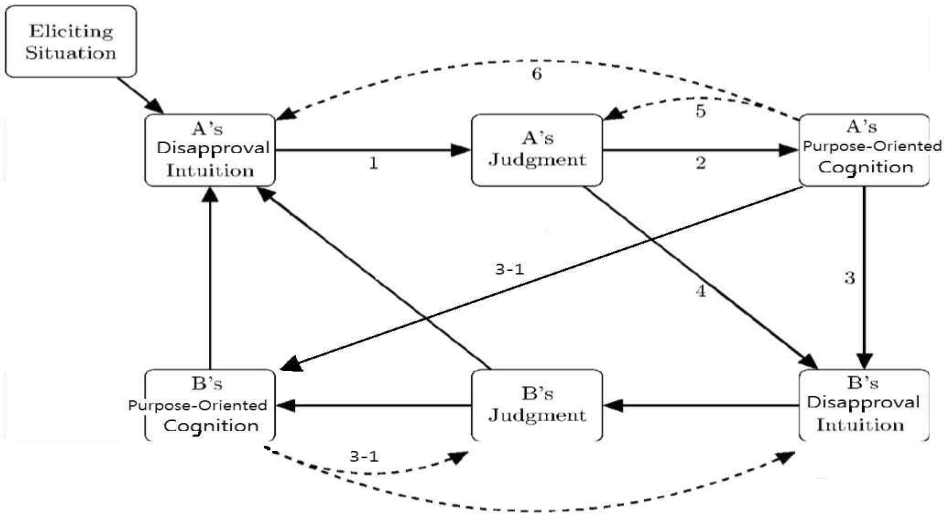
1. 사회적 직관주의 이론의 재해석

1) 사회적 직관주의 모델의 재해석

하이트의 사회적 직관주의 모델은 피해와 공정성의 기준이 적용되지 않는 상황을 전제로 도덕적 판단의 기제를 규명하고 있다. 사회적 직관주의 모델의 핵심 주장은 다음과 같다. 첫째, 피해와 공정성의 기준이 적용되지 않는 상황에 있어, 도덕적 판단은 '불승인 직관'에 의해 도출된다. 둘째, 피해와 공정성의 기준이 적용되지 않는 상황의 경우, '목적 지향적 인지'는 대부분 '불승인의 직관적 판단'을 합리화하는데 기능한다. 이러한 '목적 지향적 인지'는 드물게 불승인의 직관적 판단을 수정하는데 기능하기도 하지만, 개인의 확증 편향(confirmation bias)으로 인하여 그러한 경우는 매우 드물게 도출된다.

선행한 비판적 검토 작업에 기반하여 사회적 직관주의 모델을 재해석하면 아래의 <그림-14>와 같다.

<그림-14> 사회적 직관주의 모델 재해석



<그림-14>의 Link 1은 '불승인의 직관적 판단(Intuitive Disapproval Judgment Link)'으로서, 피해와 공정성의 기준이 적용되지 않지만 불온한 감정 반응을 유발하는 상황에 있어 불승인 직관에 의해 즉각적이고 자동적인 직관적 형태의 도덕적 판단이 도출되는 과정이다. 하이트는 피해와 공정성의 기준이 적용되지 않지만 불온한 '근친상간 시나리오'와 '식인 시나리오' 에서의 판단이 갖는 특성을 통해, 피해와 공정성의 기준을 적용할 수 있는 인지주의 전통의 도덕 상황이 아닌 상황에 있어, 불승인의 직관에 의해 도덕적 판단이 도출된다는 것을 규명하였다. 또한 하이트에 따르면 이러한 불승인 직관은 지역이나 계층 집단이 공유하는 문화적 기준에 영향을 받는다.

Link 2는 '확증적 인지(Confirmative Cognition Link)'로서, 불승인의 직관적 판단을 합리화하고 지지하기 위한 근거를 모색하는 인지 과정이다. 하이트는 이러한 확증적 인지를 추론으로 구분하며 추론의 역할을 합리화에 한정하였으나, 앞서 수행한 개념 범주 명료화 작업에 기반할 때, 확증적 인지는 수월성 정도에 따라 추론적 형태 또는 직관적 형태로 도출될 수도 있기에 형식적 차원에서 특정하지 않는 것이 타당하다. 이와 관련하여, <그림-14> 우측 상단의 박스를 확증적 인지로 표기하지 않고 '목적 지향적 인지(Purpose-Oriented Cognition)'로 표기한 이유는, 해당 인지는 Link2 외에도 Link 3과 Link 5 및 Link 6에 관여하는 인지의 내용적 특성을 모두 포괄할 수 있어야 하며, 확증적 인지 또한 '불승인의 직관적 판단에 대한 합리화'라는 목적에 종속된 목적 지향적 인지의 일종 즉, 목적 지향적 인지라는 개념 범주에 포함되는 하위 개념 범주이기 때문이다.

Link 3은 '추론된 설득(Reasoned Persuasion Link)'으로서, 확증적 인지와 같은 '목적 지향적 인지'를 언어적으로 전달함으로써, 상대방의 불승인 직관에 대한 수정을 이끌어 내는 과정이다. 이러한 '추론된 설득'은 감정적인 영향을 미칠 수 있는 정서적 설득의 특성을 가질 때 더욱 설득력을 갖는다. Link 3-1은 설득이 상대의 '불승인 직관'

에는 영향을 미치지 못한 채, ‘목적 지향적 인지’에만 영향을 주어, 이중적 태도(Dual Attitude)가 수반된 판단의 변화가 도출되는 과정을 가리킨다.

Link 4는 ‘사회적 설득(Social Persuasion Link)’으로서, 목적 지향적 인지의 전달 없이, ‘불승인의 직관적 판단’만을 전달함으로써 상대방의 불승인 직관의 수정을 가져오는 과정이다. 사회적 설득은 근거의 제시를 필요로 하지 않으며, 가치 있게 여기는 개인의 판단이거나 정서적으로 수용하는 개인의 판단일수록 더 큰 영향력을 갖는다.

Link 5는 ‘목적 지향적 판단(Purpose-Oriented Judgment Link)’으로서, ‘목적 지향적 인지’를 통해 ‘불승인 직관’에 대한 변화 없이, ‘불승인의 직관적 판단’을 직접적으로 수정하는 과정이다. 목적 지향적 판단에 있어서의 ‘목적 지향적 인지’는 불승인의 직관적 판단에 반응적으로 도출된 것일 수도 있지만, 해당 상황에 당면하기 이전에 학습된 것일 수도 있다. 이는 수월성 정도에 따라 직관적 형태일 수도 있으며 추론적 형태일 수도 있다. 또한 해당 목적에는 도덕적으로 더욱 타당한 판단을 모색하는 것과 편견에 근거한 판단이나 왜곡된 판단을 내리지 않으려는 것이 포함될 수 있으며 타인에게 좋은 평판을 받는 등의 도덕 외적인 목적 또한 포함될 수 있다. 목적 지향적 판단의 과정에서 ‘불승인 직관’은 계속 유지되기에, 불승인 직관과 목적 지향적 인지가 갈등하며 이중적 태도가 수반된다.

Link 6 ‘사적인 반성적 사고(the Private Reflection Link)’는 개인의 사고과정 속에서 불승인 직관과 충돌하는 ‘감정 반응’이 자발적으로 활성화되고 이를 통해 ‘불승인의 직관적 판단’에 수정이 이루어지는 과정이다. 즉, Link 5는 ‘불승인 직관’이 유지된 채 ‘목적 지향적 인지’를 통해 직접적으로 판단을 변경하는 과정이며, Link 6은 불승인 직관이 수정됨으로써 새로운 판단이 도출되는 과정이다.

각 과정에 대한 재해석 작업과 함께, 도덕적 판단에 대한 사회적 직관주의 모델의 정의 또한 수정되어야 할 필요성이 있다. 하이트는 ‘도덕적

판단이란, 문화 혹은 하위문화에 의해 의무적인 것으로 지지되는 일련의 덕에 대한 존경심으로 형성되는, 행동이나 사람의 인격에 대한 평가(옳고 그름)’라고 정의하였다(Haidt, 2001, p. 817). 그러나 앞선 규명 범주 구체화 작업에 기반할 때, 사회적 직관주의 모델은 피해와 공정성의 기준이 적용되는 전통적인 도덕 상황에 대한 규명을 포괄하지 않는다. 그렇기에 사회적 직관주의 모델의 도덕적 판단에 대한 정의는 ‘피해와 공정성의 기준이 적용되지 않는 상황에 대한 도덕적 판단이란, 문화 혹은 하위문화에 의해 의무적인 것으로 지지되는 일련의 덕에 대한 존경심으로 형성되는 행동이나 사람의 인격에 대한 평가(옳고 그름)’로 수정되어야 한다. 동일한 맥락에서, ‘충분히 문화화 된 인간이 덕 있는 인간이다’라는 그의 주장(Haidt & Bjorkulund, 2008, pp. 215-216) 또한, ‘피해와 공정성의 기준이 적용되지 않는 상황에 대한 도덕적 판단에 있어, 충분히 문화화 된 인간이 덕 있는 판단을 내린다’는 주장으로 구체화되어야 한다.

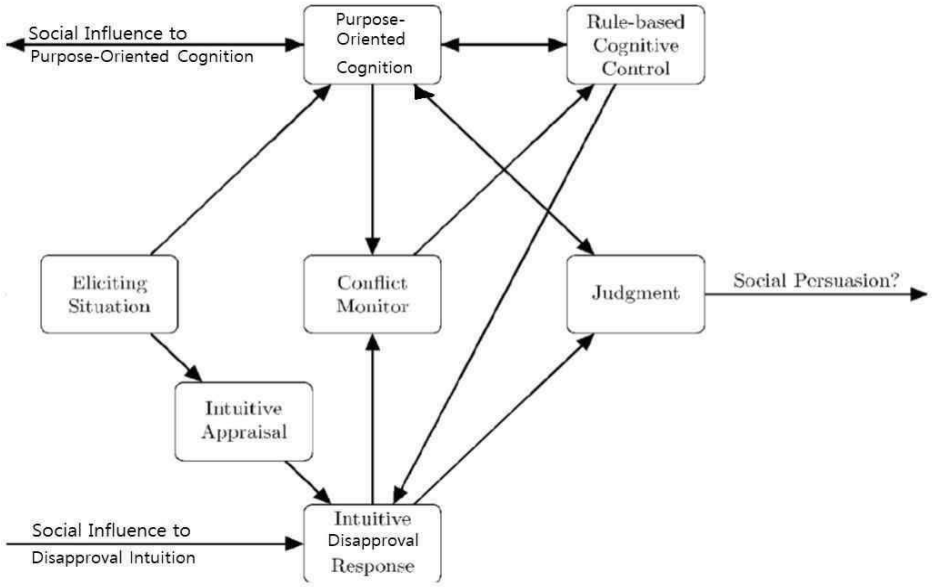
이와 관련하여 주의할 점은 사회적 직관주의 모델에 있어 ‘덕이 있다’는 표현은 실제로는 ‘부덕하지 않음’을 의미한다는 것이다. 사회적 직관주의 모델이 규명한 도덕적 판단의 기제는, 행위나 인격에 대한 불승인의 판단을 그 내용으로 하기에, 사회적 직관주의 모델에 있어 덕이 있다는 표현은 승인될 수 없는 행위가 무엇인지 잘 알고 그러한 행동을 지양한다는 의미로 받아들이는 것이 타당하다. 덧붙여, ‘충분한’ 문화화란, 피해와 공정성의 기준이 적용되지 않는 특정 상황에 있어, 고려되어야 할 다양한 문화적 금기(taboo)의 기준들이 내재 및 조율되어 직관적으로 도출되는 상태를 의미한다고 볼 수 있다.

2) 이중과정 모델의 재해석

그린의 이중과정 모델은 피해와 공정성의 기준이 적용되는 전통적인 도덕적 판단의 상황을 전제로 불승인 직관이 어떠한 방식으로 판단에 영

향을 미치는지를 규명하였다. 선행한 이론적 비판적 검토 작업에 기반하여 이중과정 모델을 재해석하면 <그림-15>과 같다.

<그림-15> 이중과정 모델 재해석



이중과정 모델의 대표적인 두 가지 특성은 다음과 같다. 첫째, 피해와 공정성의 기준이 적용되는 전통적인 도덕적 상황에 있어 도출되는 도덕적 판단의 양태는 '불승인 직관'의 도출 여부에 종속된다. 둘째, 결과론적 근거가 명확한 경우, '목적 지향적 인지'를 통해 '불승인의 직관적 판단'을 수정할 수 있다.

먼저, 첫 번째 특성과 관련하여 살펴보면 다음과 같다. 그린은 전통적인 도덕적 딜레마 상황에서 도출되는 도덕적 판단의 양태가 상황에 대한 감정 반응과 밀접한 관련을 가진다고 보았다. 그는 그 이유를 감정 반응의 활성화 정도에서 찾았고 이를 증명하기 위하여 인도교 딜레마와 트롤리 딜레마를 활용하였다. 실험 결과, 강한 감정 반응을 수반하는 인신적

도덕 딜레마인 인도교 딜레마 경우, 피험자들에게서 감정 반응 및 사회적 인지와 관련된 두뇌 영역의 신경 활동의 증가가 목격되었으며, 대다수의 피험자가 인도교 위의 사람을 밀지 않는 것을 선택하였다. 반면 비교적 약한 감정 반응을 수반하는 비인신적 도덕 딜레마인 트롤리 딜레마의 경우, 고차적 인지와 관련된 두뇌 영역이 상대적으로 더 활성화 되었으며 대부분의 피험자가 스위치를 조작하여 선로를 변경하는 것을 선택하였다.

이러한 결과에 기반하여, 그린은 감정적인 반응의 활성화 정도가 도덕적 판단의 도출 양태 즉, 직관적 판단 또는 추론적 판단의 도출 방식을 결정한다고 보았다(Greene, 2001, pp. 2105-2106). 그러나 앞서 언급한 바와 같이, 감정의 활성화 정도를 기준으로 하여 형식적 차원의 인지 심리학적 개념 범주인 '추론적 판단'과 '직관적 판단'을 구분하는 것은, 해당 개념 범주가 내포하는 다양한 특성으로 인하여, 해석자의 초점에 따라 모순되고 충돌하는 해석의 위험성을 내포한다. 그렇기에 이중과정 모델을 해석함에 있어, '직관적 또는 추론적 도덕 판단의 도출 양태는 상황에 대한 감정 반응의 활성화 정도에 종속된다.'고 표현하는 것은 잘못된 표현이며, '불승인의 직관적 판단 또는 목적 지향적 판단의 도출 양태는 불승인 직관의 도출 여부에 종속된다.'고 표현하는 것이 보다 적절하다.

두 번째 특성과 관련하여 살펴보면 다음과 같다. 그린은 트롤리 딜레마 상황과 인도교 딜레마 상황에서 각각 상이한 응답에 걸리는 시간을 비교함으로써, 인신적 도덕 딜레마 상황에서는 부정적인 답변과 긍정적인 답변 사이의 반응 시간에 차이가 나타나지만, 그와 비교했을 때 비인신적 도덕 딜레마 상황에서는 부정적인 답변과 긍정적인 답변 사이의 반응 시간에 뚜렷한 차이가 없다고 지적하였다. 그는 이를 통해, 결과론적 근거가 명확한 경우, 딜레마 상황에서 감정적 반응과 일치하지 않는 추론적 판단을 내리는 것이 가능하긴 하지만 인지적 조절 과정을 통해 이를 극복해야하기 때문에 판단에 간섭하는 감정적 반응이 강한 경우, 더

많은 시간이 소요된다고 지적하였다(Greene, 2001, pp. 2106-2107; Greene, 2008, pp. 45-46).

그러나 앞서 수행한 작업에 기반할 때, 이러한 '감정 반응'은 '불승인 직관'으로 한정할 필요가 있으며, '추론적 판단'은 '목적 지향적 판단'으로 대체해야 할 필요성이 있다. 그렇기에 반응 시간 차이 비교 실험의 결과는, '결과론적 근거가 명확한 경우, 딜레마 상황에서 불승인 직관과 일치하지 않는 목적 지향적 판단을 내리는 것이 가능하지만 인지적 조절 과정을 통해 불승인 직관을 극복해야 하기 때문에, 판단에 간섭하는 불승인의 감정 반응이 강한 경우 더 많은 시간이 소요된다.' 와 같이 수정될 필요성이 있다.

덧붙여, 그린은 직관적 판단이 의무론적 근거에 의해 지지되고 추론적 판단은 결과론적 근거에 의해 지지된다고 지적하며, 의무론적 판단이란 감정 반응에 추동된 판단일 뿐이며 이러한 의무론적 판단에 있어 추론이란 합리화에 기능할 뿐 판단의 과정에 직접 기능하지 못한다고 주장하였다. 이어서 그는 의무론이란 인간이 자신의 행위를 설명하기 위하여 만들어낸 일종의 도덕적 작화(Moral Confabulation)라고 비판하며, 의무론적 판단과 대조적인 결과론적 판단은 비교적 인지적이며 그러한 결과론적 판단의 과정에 있어 추론은 사후 합리화가 아닌 판단의 산출에 작용하는 진정한 의미의 도덕적 추론으로 작용한다고 주장하였다. 또한 이에 기반하여 그린은 의무론적 판단은 의식적으로 추론에 의해 도출된 것이 아니라 감정적으로 직관의 형태로 도출된 것이기에 규범 윤리로서 신뢰하기 어렵다고 지적하였다(Greene, 2008, pp. 59-61).

이러한 심리학적 근거에 기반한 그린의 의무론 비판이 타당성을 담보하기 위해서는, 직관과 추론이라는 개념 범주를 활용한 그린의 이중과정 모델이 도덕적 판단의 현상을 타당하게 규명하고 있어야 한다. 그러나 상기한 바와 같이 그린이 직관으로 분류한 것은 내용적으로 그리고 형식적으로 특정되는 '불승인 직관'으로 표기하는 것이 타당하며, 그린이 추론으로 분류한 인지는 형식적 차원에서 열려있는 내용적 개념 범주인

‘목적 지향적 판단’으로 대체하는 것이 적합하다. 즉, 의무론의 규범적 지위에 대한 그린의 비판은 잘못 해석된 심리학적 근거에 기반한 비판이라고 할 수 있다.

앞서 살펴본 바와 같이, 불승인의 직관적 판단과 목적 지향적 판단은 도덕적 판단에 있어 상이한 역할을 담당한다. 의무론적 근거에 의해 지지되는 불승인의 직관적 판단이란 설정된 목적과 독립적으로 특정 행위의 허용 불가능성 여부를 판단하는 기능을 담당하는 반면, 결과론적 근거에 의해 지지되는 목적 지향적 판단은 설정된 목적의 달성에 더욱 부합하는 행위를 선택하는 것을 내용으로 한다. 즉, 두 유형의 판단은 도덕적 판단에 있어 상이한 역할을 담당하며 보완적으로 기능하고, 그렇기에 의무론적 판단과 결과론적 판단은 상이한 방식으로 고유의 규범적 가치를 갖는다고 평가할 수 있다.

또한 만일 도덕적 딜레마 상황에서의 판단과 같이 가장 도덕적인 행위의 선택이라는 목적을 전제한 상황이 아닌 비도덕적인 목적이 전제될 수 있는 일상의 많은 상황들을 상정한다면, 그에 종속된 목적 지향적 인지 과정은 오히려 비도덕적인 판단의 도출로 이어질 위험성을 내재한다. 그렇기에 ‘비교적 덜 감정적이고 인지적’이라는 특성만을 통해 인지 유형의 신뢰도를 평가하는 것은 타당하지 않은 접근이다. 즉, 상이한 기제를 통해 보완적으로 기능하는 두 유형의 인지가 갖는 신뢰도는 상이한 기준과 방식에 의해 평가되어야 하며, 그렇기에 어떠한 유형의 인지가 더욱 신뢰도가 있는가를 논하는 것은 우문이 아닐 수 없다.

이와 관련하여, 상이한 기제의 두 유형의 인지가 도덕적 판단에 있어 어떠한 방식으로 보완되며 어떠한 위험성을 내재하는지에 대한 세부적 논의는 다음절에서 진행하도록 하겠다.

2. 유형별 인지의 상호 보완 가능성 분석

시노트 암스트롱(W. Sinnot-Armstrong)의 지적과 같이, 그간 철학자와 심리학자들은 도덕적 직관에 대하여 많은 관심을 가져왔지만 그와 관련하여 동의할만한 정의조차 합의되지 않았다(Sinnot-Armstrong, 2008, pp. 246-247). 이렇듯 개인 도덕 기능의 기제를 규명하기 위한 기본적인 도구인 개념 범주에 대한 기본적 합의조차 이루어지지 않았기에 유형별 인지의 관계 방식 또한 명료하게 규명되지 못하였다.

도덕 기능의 기제 규명을 위한 사회적 직관주의 이론 고유의 개념 범주에 대한 협의가 이루어지지 않음에 따라, 사회적 직관주의 이론에 관한 논의에 있어 연구자들은 형식적 차원의 특성을 갖는 인지 심리학적 개념 범주인 직관과 추론을 활용하였다. 이로 인하여, 유형별 인지의 상호 보완 가능성에 대한 논의는, 도덕적 판단의 현상을 대상으로 인지 심리학적 인지 유형인 직관과 추론의 상호 보완 방식을 재확인하는 결과로 이어졌으며, 사회적 직관주의 이론 고유의 인지 유형별 상호 보완 방식의 도출로 이어지지는 못하였다. 이에, 본 절에서는, 인지 심리학 고유의 개념 범주인 직관과 추론의 상호 보완 가능성에 대한 논의를 먼저 살펴본 후, 사회적 직관주의 이론이 규명한 바를 타당하게 해석하기 위해 도출한 인지 유형인 '불승인 직관'과 '목적 지향적 인지'의 상호 보완 가능성을 도출하고자 한다.

1) 직관과 추론의 상호 보완 가능성

인지 심리학적 개념 범주인 직관과 추론의 상호 보완 가능성을 명료화하기에 앞서, 인지 심리학적 인지 유형으로서의 직관과 추론에 대하여 살펴보면 다음과 같다. 직관과 추론이라는 두 유형의 인지 과정이 갖는 특징을 구분하는데 있어 가장 널리 사용되어 온 분류로는 휴리스틱(heuristic)과 애널리틱(analytic)의 분류가 있다. 카네만(D.

Kahneman)은 인지적 무의식 상태에서 작동하는 휴리스틱으로서의 시스템 1과 의식 수준의 합리적 추론에 해당하는 시스템 2를 분류하였으며(Kahneman, 2011/2012), 호거스(R. Hogarth)는 이를 암묵적(tacit) 시스템과 숙고(deliberate) 시스템으로 분류하였다(Hogarth, 2001). 또한 앞서 2장에서 살펴본 바와 같이, 직관적 인지는 무의식적이고, 자동적이며, 암묵적인 특성 등을 가지며, 추론적 인지는 의식적이고, 노력을 필요로 하며, 명시적인 특성 등을 갖는다.

이러한 추론과 직관의 특징은 그린의 카메라 비유를 통해서도 이해할 수 있다. 그린은 인간의 뇌를 자동 및 수동의 이중적 방식에 의해 구동되는 디자인의 카메라로 묘사하였다. 그린에 따르면 자동 방식은 특정한 판단을 도출하는 직관적 반응으로 많은 경우에 있어 감정적인 특성을 갖는다. 개인은 이러한 감정적이고 직관적인 반응이 유발되는 과정에 대하여 의식하지 못한다. 수동 방식은 현재의 환경적 자극이나 신체 반응에 의해 자동적으로 활성화 되지 않는 장기적인 목표를 지향하는 추론적 판단으로 비교적 이성적인 특성을 강하게 갖는다(Greene, 2015, pp. 142-143). 즉, 적응의 산물인 직관은 효율성을 담보하지만 암묵적이고 무의식적인 특성으로 인하여 왜곡과 편향을 수정하기 어렵다는 위험성을 가지며, 추론은 왜곡과 편향의 위험을 경계할 수 있는 가능성은 더 높지만 효율성이 떨어진다.

카네만은 직관적 인지의 오류 가능성을 경계하기 위해, 추론적인 시스템 2를 통해 무의식적 시스템 1에서 소홀히 다루어졌을 법한 요인들을 검토할 필요가 있다고 지적하면서도, 직관적 인지의 유용성을 인정할 필요가 있음을 강조하였다(Kahneman, 2011/2012). 스턴버그(R. Sternberg) 또한 자동적 인지 처리는 의식적 통제 처리를 필요로 하지 않기 때문에 여러 개를 동시에 수행할 수 있지만, 노력이 많이 소요되는 통제된 인지 처리 과정의 경우 둘 이상을 동시에 수행하기가 어렵다고 지적하며 직관적 인지의 유용성을 강조하였다(Sternberg, 2003/2005, p. 67). 브루너(J. Bruner) 또한 직관적 방법으로 얻은 지식은 분석적

으로 검증될 필요가 있긴 하지만, 직관적 인지 과정은 분석적 인지과정을 통해 충족하기 어려운 전체적 사고를 가능하게 하기 때문에 직관적 인지 과정을 통해 도출된 통찰을 존중할 필요가 있다고 강조하였다(Bruner, 1960, pp. 56-68). 즉, 두 인지 유형이 효율성과 안전성의 차원에서 상호 보완적으로 기능하는 것이 가장 이상적인 인지의 형태라고 할 수 있다.

우선, 효율성 차원의 지향점으로서 직관의 발달과 관련하여, 나바에츠(D. Narvaez)와 랩슬리(D. Lapsley)의 연구를 참고할 수 있다. 나바에츠와 랩슬리는 전문가와 초보자를 비교하며 전문가는 초보자에 비하여 빠르고 자동적인 직관적 형태의 의사 결정에 의존하며, 효과적으로 해결책을 찾는 경향이 있다고 지적하였다(Lapsley, Power, 2005/2008, pp. 266-269). 나바에츠와 랩슬리에 따르면 전문가의 직관적 인지는 다음과 같은 특징을 갖는다.

첫째, 전문가는 자신의 전문 영역에서 특별한 표상(representations)들을 가지고 있기에, 지식을 더 잘 조직한다. 전문가는 선언적(declarative) 지식과 절차적(procedural) 지식 및 조건적(conditional) 지식이 잘 조직화되어 원활하게 상호작용 하는 스키마를 가지고 있다. 둘째, 전문가는 초보자들이 놓치는 세부적인 사항에 주목하고, 그에 기반하여 초보자와 상이한 방식으로 세계를 바라본다. 전문가들은 초보자들에 비해 더 풍부하게 조직화된 스키마를 가지고 있기 때문에, 상황에 대한 인식과 수행 수준에 있어 초보자와 차이를 보인다. 상황을 대면하였을 때, 초보자의 경우 어떤 지식을 선별하여 적용해야 하는지에 대하여 빠르게 파악하지 못하고 정보에 압도당하는 경향이 있지만, 전문가의 경우 상황을 빠르게 인식하고 그에 적합한 정보를 빠르게 즉각적으로 파악하는 경향이 있다. 셋째, 전문가는 독특한 기술들을 가지고 있다. 초보자가 문제 상황을 명확히 하지 않고 바로 문제를 풀려고 뛰어드는 것과 달리, 전문가는 의사결정을 할 때 당면한 상황의 문제를 명확히 밝히고 그 문제를 해결하기 위해 자신이 가지고 있는 기존의 지식망을 활용한

다. 이러한 과정에서 전문가들은 과거의 유사 경험과 현재의 당면 상황을 연결함으로써, 적절한 지식을 어떠한 절차에 따라 언제 어떻게 적용해야 하는지를 신속하고 자동적으로 처리한다.

그렇다면 직관적 판단에 내재한 오류의 위험성에는 어떠한 것이 있으며 추론은 이를 경계하는데 어떻게 기능할 수 있을까. 자동적인 정보 처리 과정과 그에 기반한 판단에 있어서의 오류와 관련하여서는 자이언스(R. Zajonc)와 보우어(G. Bower)의 연구를 참고할 수 있다. 자이언스와 보우어에 따르면 추론 또는 직관적 인지에 영향을 미치는 감정적 반응을 처리하는 과정에 있어 자동처리 과정으로서의 특성이 목격된다. 자이언스에 따르면 의도적이고 의식적인 처리과정이 굳이 개입되지 않은 상태 즉, 자신도 전혀 자각하지 못하는 상태에서도 개인들은 감정을 유발하는 자극에 대해 반응한다(Zajonc, 1980). 또한 보우어에 따르면 감정과 관련된 사건이나 경험 또는 심장 박동이나 피부전도반사와 같은 특정한 생리적 패턴들은 밀접하게 연결되어 있어 하나의 연결고리가 각성하는 순간 그러한 자극이 퍼져나가 후속적으로 입력되는 정보의 처리에 영향을 미치게 된다(Bower, 1991).

그러나 이러한 직관적 인지의 오류 가능성은 추론적 인지에 의해 경계될 수 있다. 예를 들어, 머피(S. Murphy)와 자이언스(R. Zajonc)는 감정적 자극의 점화(priming)가 의식적 통제를 통해 조절될 수 있다는 것을 규명하였다. 머피와 자이언스는 의식할 수 없는 짧은 시간동안 감정적 자극(얼굴 표정)이 제시된 경우, 이러한 자극이 뒤이어 제시된 중립적인 단어에 대한 호오도에 강력한 영향을 미치지만, 그러한 감정적 자극이 의식할 수 있는 수준에서 제시된 경우 점화 효과[13]가 발견되지 않았다는 연구 결과를 제시하였다(Murphy, 1993, p. 723).

자커비(L. Jacoby)의 연구(Jacoby, 1989) 또한 개인이 자신의 인지 상태를 유발시킨 원인 즉, 자신의 인지에 영향을 미친 암묵적 정보에 대

13) 점화 효과란 시간적으로 먼저 제시된 자극이 나중에 제시된 자극의 처리에 긍정적 혹은 부정적 영향을 주는 현상을 의미한다.

하여 인식하는 경우 보다 객관적 관점에서 인지를 통제할 수 있음을 보여주었다. 자커비는 피험자를 두 그룹으로 나누어 이름을 외우게 하는 실험을 진행하였으며, 그는 피험자를 모아놓고 명단의 이름들(절반은 유명 인사, 절반은 유명 인사가 아닌 일반인의 이름)을 외우게 하였고 피험자들에게 그 이름들이 전화번호부에서 선정한 이름들이라고 이야기함으로써 그 이름들이 유명 인사의 이름이 아니라는 사실을 인식시켰다.

이후, 한 그룹에게는 일련의 이름들에 대해 즉시 유명 정도를 판단하게 하였고 다른 그룹은 일단 돌려보낸 후 다음날 다시 방문하여 일련의 이름들의 유명 정도를 판단하게 하였다. 암기 후 즉시 과제를 수행한 그룹은 모든 이름에 대한 친숙도가 높은 것이 암기의 영향이라는 것을 의식하여 각 이름의 유명 정도를 제대로 판단하였으나 다음날 과제를 수행한 그룹은 이름들에 대해 친숙한 이유가 실제 그 이름이 유명 인사의 이름이어서인지 전날의 훈련으로 인해 친숙해 진 것인지를 구별하지 못하여 학습된 이름들도 유명하다고 판단하였다.

즉, 이름에 대한 친숙성의 원인을 명확하게 인지하지 못하는 경우에는 그로 인한 영향을 받았지만, 친숙성의 원인을 명확하게 인지하고 있는 경우에는 그로 인한 영향을 받지 않았다. 베게너(D. Wegener)와 페티(R. Petty)는 이러한 현상이 바로 지각자의 교정 욕구 때문에 일어날 수 있다고 설명한다(Wegener & Petty, 1997, p. 189). 즉, 자신의 판단이 왜곡될 수 있다고 생각하는 경우에는 기존의 자동화된 처리과정에 대해 부가적인 인지적 노력을 기울여 좀 더 점화자극으로부터 독립적이며 객관적인 판단과정을 취하고자 노력하게 된다는 것이다.

그러나 이러한 추론과 직관의 상호 보완 가능성에 대한 논의만으로는 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 온전히 규명하기에 어려움이 있다. 이에 사회적 직관주의 이론 고유의 인지 유형 분류에 기반한 유형별 인지의 상호 보완 가능성에 대한 모색이 필요하다.

2) 불승인 직관과 목적 지향적 인지의 상호 보완 가능성

인지 심리학적 인지 유형인 직관과 추론에 대한 연구들은, 도덕 심리학이 이성과 감정이라는 개념 범주를 활용하여 현실을 설명함에 있어 발생하는 규명상의 한계를 극복하는데 도움을 주었다. 그러나 사회적 직관주의 이론이 규명한 현상을 해석함에 있어 이러한 인지 심리학적 개념 범주의 활용은 모순된 다의적 해석의 위험성을 내포하였다. 그렇다면 사회적 직관주의 이론을 보다 타당하게 해석하기 위한 인지 유형으로서의 '불승인 직관' 과 '목적 지향적 인지' 는 어떠한 방식으로 상호 보완적으로 기능할 수 있을까.

우선, 피해와 공정성의 기준이 적용되지 않는 상황을 대상으로 살펴보면 다음과 같다. 피해와 공정성의 기준이 적용되지 않는 상황에서는 불승인 직관에 의해서만 판단이 도출되며, 목적 지향적 인지는 이러한 불승인 여부에 대한 판단을 합리화 하는데 기능한다. 이 경우에 있어 목적 지향적 인지는 불승인 직관에 종속되기에 두 인지는 인지의 한계나 위험성을 상호 보완하는데 기능할 수 없다.

다만, 드물게 도출되기는 하지만, 목적 지향적 인지에 의한 불승인 직관의 수정 기제 또는 왜곡된 불승인 직관에의 경계와 같이, 불승인의 직관적 판단에 대한 목적 지향적 인지의 '일방적인 보완' 은 가능하다. 이는 구체적으로 사회적 직관주의 모델에서 제시된 '추론된 판단' 의 가능성과 '사적인 반성적 사고' 의 가능성 및 앞선 인지 심리학적 인지 유형의 보완 가능성에서 살펴본 '불승인 직관의 왜곡에 대한 추론적 보완' 가능성이 있다. 또한 충분한 문화화의 구성 요건으로서 직관 형성의 과정에 있어서 '추론에 의한 불승인 기준들 간의 조율' 가능성 등으로 나누어 볼 수 있다.

이러한 보완 과정에 있어 목적 지향적 인지는 다음과 같은 역할을 담당할 수 있다. 목적 지향적 인지는 추론된 판단의 과정에서 후의식적으로 도출되어 전의식적 자동 처리 과정으로서의 불승인 직관에 의해 도출된 편견적 판단을 통제하는 역할을 수행할 수 있다. 또한 사적인 반성적

사고 과정과 같이, 다양한 관점에서 상황을 바라보게 도움으로써 불승인의 직관적 판단이 편향 및 왜곡되는 것을 경계할 수 있다. 그리고 불승인의 직관적 판단이 도출되는 과정에 영향을 미친 암묵적 정보에 대하여 인식할 수 있게 함으로써 판단자가 보다 객관적 관점에서 불승인의 직관적 판단을 통제할 수 있게 도울 수도 있다. 또한 불승인 직관의 형성 및 발달 과정에 있어 다양한 불승인의 기준들을 조율하는 방법을 모색함으로써 현실과 괴리되지 않고 충분한 문화화를 거친 성숙한 불승인 직관이 형성될 수 있도록 도울 수 있다.

이러한 보완 가능성은 그 구체적인 방식에 있어 다양하지만, 모두 암묵적이고 자동적으로 도출된 불승인 직관의 왜곡과 편견을 경계하는 것을 목적으로 한다. 도덕적 판단은 행위 자체에 대한 불승인의 감정 반응뿐 아니라, 행위와 독립적으로 주어진 감정 반응에 의해서도 영향을 받는다는 것을 규명한 하이트의 연구(Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005)는, 불승인의 직관적 판단에 내재한 암묵적 지식과 감정적 정보의 영향을 경계할 필요성을 더욱 여실히 보여준다. 하이트는 정서적으로 영향을 미치는 단어 또는 악취와 지저분한 환경 등 개인에게 부정적인 감정 반응을 일으키지만 평가의 대상이 되는 행위와는 관련이 없는 요인들이 행위에 대한 평가에 영향을 미칠 수 있는 가능성을 규명하였다.

피해와 공정성의 기준이 적용되지 않는 상황에 있어 불승인 직관에 대한 목적 지향적 인지의 일방적인 보완만이 가능했던 것과 달리, 피해와 공정성의 기준의 적용되는 전통적인 도덕 상황에 있어서는 '불승인 직관'과 '목적 지향적 인지'가 상호 보완적으로 기능할 수 있다.

불승인 직관에 대한 목적 지향적 인지의 보완 가능성은 피해와 공정성의 기준이 적용되지 않는 상황에 있어서의 보완 방식과 유사하다. 목적 지향적 인지 과정은 후의식적으로 도출되어 전의식적 자동 처리 과정으로서의 불승인 직관에 의해 도출된 편견적 판단을 통제하는 역할을 수행할 수 있으며, 다양한 관점에서 상황을 바라보게 도움으로써 불승인의

직관적 판단이 편향되거나 왜곡되는 것을 경계할 수 있다. 또한 불승인 직관의 도출에 영향을 미친 암묵적 정보에 대하여 인식할 수 있게 함으로써 판단자가 보다 객관적 관점을 가질 수 있게 할 수 있으며, 불승인 직관의 형성 및 발달에 있어 다양한 금기의 기준들을 조율하는 방법을 모색하게 함으로써 수준 높은 불승인 직관이 도출될 수 있도록 도울 수 있다.

그렇다면 불승인 직관은 목적 지향적 인지를 어떠한 방식으로 보완할 수 있을까. 앞서 언급한 바와 같이, 그린은 합리주의 전통의 도덕적 딜레마 상황에 있어서의 판단 즉, 도덕이라는 목적이 전제된 상황에서의 판단을 대상으로 그의 연구를 진행하고 주장을 제시하였다. 그렇기에 그의 연구에 있어 '목적 지향적 판단'은 도덕적 목적을 달성하기 위한 최적의 수단을 선택하는 것과 관련된다. 그러나 일상의 다양한 상황 중, 도덕이 목적이 아닌 상황(이익 추구를 목적으로 하는 상황, 경쟁에서의 승리를 목적으로 하는 상황 등)에 있어 결과론적 인지 즉, 목적 지향적 판단은 '도덕 외의 목적'에 종속되는 경우가 많으며, 도덕이라는 '목적의 설정'에까지 기능하기에는 어려움이 있다. 이러한 경우에 있어, 도덕적 위반에 대한 불승인 직관은 개인이 비도덕적인 목적을 전제로 한 목적 지향적 인지에 의해 비도덕적 행위를 자행하는 것을 방지하는 역할을 담당할 수 있다.

나아가 개인이 도덕 외의 목적을 추구하는 상황 뿐 아니라, 도덕을 추구하는 상황에 있어서도, 불승인 직관은 목적 지향적 판단을 보완하는 역할을 담당한다. 일상의 상황에 있어, 개인은 왜곡된 정보 수용에 기반하거나 편중된 가치 선택에 기반하여 상황을 오해석하고 그에 의해 도덕적 목표 설정에 오류를 범할 수도 있다. 예를 들어, 정보의 수용에 있어 특정 개인 또는 집단 등에 대한 편견이 작용하는 경우, 정의와 같은 도덕적 가치를 내세우며 실제로 죄가 없는 사람에게 처벌을 가해야 한다고 주장하거나 공정함을 내세우며 혜택을 받아야 하는 사람에게 부담을 더 지우는 일 등이 발생할 수 있다. 즉, 일상의 상황에서 개인이 판단의 근

거로 삼는 정보들은 한계가 존재할 수밖에 없으며, 목적 지향적 인지도 불승인 직관과 같이 왜곡된 감정적 정보 또는 암묵적 정보에 기반하는 측면이 있을 수밖에 없다. 그렇기에 개인이 도덕적 가치를 추구한다 하더라도, 구체적인 상황에 있어 개인이 도덕적이라 믿는 선택이 실제로는 비도덕적인 선택일 가능성이 있으며, 일부 도덕적 요소를 고려하지 못함으로써 편향된 의사 결정을 내릴 가능성이 있다. 이러한 상황에 있어서도 불승인 직관은, 개인이 편향된 목적 지향적 인지 과정에 기반하여 비도덕적 행위를 선택할 위험성 등을 경계한다.

이러한 위험성에 대한 예시로 말루프(A. Maalouf)가 제시한 '살상하는 정체성(identities that kill)'의 발현 상황을 참고할 수 있다. 그는 엘살바도르의 엘 모조테(El Mozote)라는 지방에서 일어난 대량학살을 예로 든다. 그 곳에서 군인들은 어린이와 여성을 가리지 않고 사람들을 무참하게 살해하였다. 이러한 대량학살에 대해 우리는 그 군인들이 도덕적 정체성이 결핍되어 있거나 적어도 도덕적 자아 개념보다 강한 다른 자아 개념을 가지고 있다고 생각할 수 있을 것이다. 그러나 그 군인들은 자신들의 행동이 도덕적 임무라고 생각하였고 그러한 행동이 정의를 실현하는 일이라고 생각하였다(이정렬, 2009, p. 152). 이러한 예는 목적 지향적 인지의 편향성을 불승인 직관을 통해 보완해야 할 필요성을 절실하게 보여준다.

Ⅵ. 사회적 직관주의 이론 비판 및 재해석의 도덕 교육적 함의

직관에 관한 심리학적 연구 성과들을 적극 반영한 대표적인 도덕 교육 이론으로는 나바에츠(D. Narvaez)의 통합적 윤리교육 모델(Integrative Ethical Education)이 있다. 앞선 작업을 통해, 도덕 심리학의 현상을 보다 타당하게 해석하기 위해서는 인지 심리학적 개념 범주를 차용하는 것에 머무르지 않고 사회적 직관주의 이론 고유의 인지 유형 분류를 활용할 필요성이 있음을 지적하였다. 또한 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 보다 타당하게 해석하기 위해서는 도덕적 인지의 유형을 '불승인 직관'과 '목적 지향적 인지'로 구분할 필요성이 있음을 도출하였다. 본 장에서는 선행한 사회적 직관주의 이론에 대한 비판적 검토 작업과 그에 기반한 도덕 심리학적 발견에 기초하여, 나바에츠의 통합적 윤리교육 모델의 개선 가능성을 모색하고자 한다.

1. 교육 목표 설정에의 적용 가능성

직관주의 연구의 성과들을 도덕 교육에 접목한 대표적인 연구자인 나바에츠(D. Narvaez)는 인간의 도덕 기능에 있어 직관적 인지 또한 추론적 인지 만큼이나 중요한 역할을 수행함을 강조하며, 두 유형의 인지가 상호 보완적으로 기능하는 도덕적 전문가를 양성해야 한다고 주장하였다. 또한 그녀는 도덕적 전문가를 양성하기 위한 교육방법 즉, 성숙한 도덕 기능 개발을 위한 교육 방법으로서 통합적 윤리교육(Integrative Ethical Education) 모델을 제시하였다. 본 절에서는 이러한 접근이 갖는 한계를 극복하기 위해 불승인 직관과 목적 지향적 인지라는 도덕 심리학적 인지 유형 분류를 적용하여 도덕적 전문성 개념의 보완 가능성을 모색하고자 한다. 이는 통합적 윤리교육 모델의 보완 가능성을 모색하기 위한 기초 작업으로서 의미를 갖는다.

1) 도덕적 전문성의 개념과 보완 필요성

나바에츠는 사회 인지적 스키마의 발달을 통한 전문성 향상의 개념을 차용하여, 도덕적 스키마의 발달을 통한 도덕적 전문성의 향상을 강조하였다. 그녀는 도덕적 스키마를 발달시키는 교육을 통해 학생들에게 도덕적 목적을 향한 지향성을 심어주고 도덕적 목적을 달성하기 위한 수월성으로서의 전문성을 길러주고자 하였다.

먼저, 전문성(expertise)의 개념에 대하여 살펴보면 다음과 같다. 앞서 언급한 바와 같이, 나바에츠에 따르면, 전문가는 자신의 전문 영역에서 특별한 표상(representations)들을 가지고 있기에 초보자에 비하여 지식을 더 잘 조직한다. 전문가는 해당 영역의 사실 및 정보에 관한 선언적 지식과 이를 적용할 수 있는 절차적 지식 및 조건적 지식이 잘 조직된 풍부한 스키마를 가지고 있다. 반면 초보자는 이러한 지식이 풍부하지 않기에 덜 조직화된 얕은 수준의 스키마를 가지고 있다.

그녀에 따르면, 이러한 스키마의 차이는 상황을 이해하고 대응하는 인지 과정 전반의 차이로 연결된다. 초보자는 상황을 직면하여 정보에 압도당하여 상황을 제대로 파악하지 못함으로써 적합한 지식을 선별하고 상황에 적용하지 못하는 경향이 있으나, 전문가는 상황을 빠르게 인식하고 그에 적합한 정보를 즉각적으로 선별하고 적용한다. 즉, 전문가는 초보자에 비해 풍부하게 조직화된 스키마를 가지고 있음으로써 상황에 대한 인식과 수행에 있어, 초보자가 인지하지 못하는 세부적인 정보를 알아차리고 이에 기반하여 세계를 바라보고 반응한다(Narvaez, 2010, p.171).

이러한 스키마의 발달 단계를 설명하기 위하여, 나바에츠는 드레이퍼스(H. Dreyfus)와 드레이퍼스(S. Dreyfus)가 제시한 기술 습득의 다섯 단계(<표-4>)를 차용한다. 그녀는 모든 분야에서 학습자들이 초보적 이해에서 전문가적 이해로 이행해 간다고 보고(Narvaez, 2004, p. 93) 도덕적으로 가장 적합한 행위를 직관적으로 판단할 수 있는 도덕적 전문가(Moral Expert)를 양성하는 것을 도덕 교육의 목표로 설정하였다.

사회 인지적 스키마 발달을 통한 전문성 향상에 대한 나바에츠의 설명은 도덕적 스키마 향상을 통한 도덕적 전문성 발달의 개념을 이해하는데 도움을 준다. 그러나 그녀가 강조하는 전문성의 핵심 요소로서의 직관 개념은 형식적인 차원의 개념 범주인 인지 심리학적 인지 유형에 의존한 것이기에, 도덕적 전문성의 목표 설정에 있어 내용적 기준을 제시하지 못한다는 한계를 갖는다. 그렇기에 도덕 기능의 사실적 기제를 반영하여 도덕적 전문성 교육의 목표를 보다 구체적으로 설정하기 위해서는, 선행한 도덕 심리학적 인지 유형 분류 작업과 그에 기초한 유형별 인지의 보완 방식에 대한 규명에 기반하여 도덕적 전문성의 개념을 보완해야 할 필요성이 있다.

실제 나바에츠의 통합적 윤리교육 모델에 있어 드레이퍼스와 드레이퍼스의 기술 습득의 다섯 단계는 추론적 판단에서 직관적 판단으로의 발전 방향을 제시하였을 뿐, 도덕 기능에 있어서의 전문성이 고려해야 할 구

체적인 내용에 있어서는 기준을 제시하지 못하였다. 이러한 한계는 전문성 교육이 갖는 특징에 대한 나바에츠의 언급에 있어서도 목격된다. 실제 그녀는 전문성 교육이 갖는 특징과 관련하여 '첫째, 전문가들은 상호 작용 또는 교육을 통해 배우며, 이는 멘토링과 적절한 피드백 및 일대일의 지도를 포함한다. 둘째, 전문가들은 암묵적 지식을 형성함에 있어 명시적인 이론을 배운다. 즉, 전문가는 상황 몰입의 과정 속에서 맥락에 맞는 명시적 이론의 적용 방법을 숙지한다. 셋째, 전문가들은 광범위하고 집중적인 연습을 경험한다(Lapsley & Power, 2005/2008, pp. 270-273).' 라고 언급하고 있을 뿐이며, 전문성 개념에 대한 구체적인 내용적 기준을 충분히 제시하지 못하고 있다.

물론 그녀는, 이어지는 논의에서 살펴보겠지만, 레스트(J. Rest)의 4구성요소 모형에 근거하여 도덕성을 구성하는 네 가지 요소들(도덕적 민감성, 도덕적 판단, 도덕적 동기화 혹은 초점, 도덕적 행동)을 제시한 후 각 요소별로 도덕적 전문성을 신장시키기 위한 일련의 기술들(skills)을 제시하고 있기는 하다. 그러나 이러한 네 가지 요소들은 도덕적 판단의 사실적 기제를 규명한 직관주의 연구들에서 필연적으로 도출된 내용이 아니며, 직관주의 연구와 별개로 존재하는 도덕 기능에 대한 통합적 접근의 연구 성과를 차용한 것이라는 한계를 갖는다.

2) 도덕적 전문성 개념의 보완 가능성

선행한 사회적 직관주의 이론 고유의 인지 유형 분류 방식을 적용할 경우, 도덕적 전문성은 목적 지향적 인지에 대한 전문성과 불승인 직관에 대한 전문성으로 구분될 수 있다. 이러한 전문성은 앞서 도출한 세 가지 유형의 도덕적 판단에 있어, 판단의 수월성을 향상시키는 적극적 차원의 전문성과 각각의 인지에 내재한 위험성을 경계하는 소극적 차원의 전문성을 포괄하는 개념이어야 한다. 앞선 비판적 검토 작업에서 도출된 도덕적 판단의 세 가지 발현 양태는 다음과 같다.

첫째, 불승인 직관이 강하게 도출 되는 경우에 있어 도덕적 판단은 피해와 공정성의 기준 적용 가능 여부와 무관하게 불승인의 직관적 판단으로 도출된다. 둘째, 피해와 공정성의 기준이 적용되지 않는 상황에 있어서는 불승인 직관이 존재하지 않는 경우 어떠한 도덕적 판단도 도출되지 않지만 피해와 공정성의 기준이 적용되는 상황에 있어서는 불승인 직관이 존재하지 않더라도 목적 지향적 판단의 형태로 도덕적 판단이 도출될 수 있다. 셋째, 불승인의 직관적 판단이 도출된다 하더라도 목적 지향적 인지에 의해 수정이 이루어질 수 있다. 단, 피해와 공정성의 기준이 적용되지 않는 상황에 있어 이러한 수정은 매우 드물게 일어나며, 피해와 공정성의 기준이 적용되는 상황에 있어서는 강한 결과론적 근거에 의해 이러한 수정이 더욱 용이하게 이루어지기도 한다.

각각의 경우에 있어 도덕 심리학적 인지 유형으로서의 불승인 직관과 목적 지향적 인지의 발달 지향점으로서 전문성에 대하여 살펴보면 다음과 같다. 도덕적 판단의 첫 번째 유형은, 불승인 직관이 강하게 도출됨으로써 불승인의 직관적 판단이 도출되는 경우이다. 우선, 피해와 공정성의 기준이 적용되지 않는 상황에 있어 불승인 직관의 전문성은 도덕성 발달에 관한 하이트의 정의를 통해 보다 구체적으로 이해할 수 있다.

앞서 살펴본 바와 같이, 하이트는 '충분히 문화화 된 인간이 덕 있는 인간이다' 라는 주장(Haidt & Bjorkulund, 2008, pp. 215-216)을 통해 불승인 직관의 발달 지향점을 제시하였다. 또한 상기한 바와 같이 하이트의 규명에 있어, 덕이 있다는 것은 실제로는 '부도덕하지 않음' 을 의미한다. 즉, 불승인 직관의 발달 지향점은 승인될 수 없는 행위가 무엇인지 잘 알고 그러한 행동을 지양하는 것이어야 한다. 또한 그러한 덕의 조건인 '충분한' 문화화란, 고려되어야 할 다양한 금기(taboo)의 기준들이 내재 및 조율되어 직관적으로 도출되는 상태를 의미한다고 볼 수 있다. 즉, 불승인 직관의 전문성은, 고려해야 하는 다양한 불승인의 기준들이 내재화되고 조율되어 직관적으로 도출됨으로써, 승인될 수 없는 행위가 무엇인지 잘 인지하고 그러한 행위를 지양하는 것을 의미한

다.

이러한 전문성으로서 불승인 직관의 발달 과정은 다음과 같은 하이트의 주장을 통해 보다 구체적으로 이해할 수 있다. 하이트는 햄린(K. Hamlin), 윈(K. Wynn), 블룸(P. Bloom)의 실험을 인용하며 아이들이 사회적 세계에 대하여 일정 수준의 이해 능력을 갖추고 태어난다고 주장한다. 실험을 인용하기에 앞서 하이트는 생후 2개월 된 영아들이 무언가가 자신의 예상대로 일어날 때보다 어딘가 놀라운 부분이 있을 때 더 오랫동안 쳐다보는 경향이 있다는 발달 심리학의 발견을 언급한다.

햄린과 윈 그리고 블룸은 6에서 10개월의 영아들을 데려다 놓고 진행한 '오르기쟁이' 인형극 실험을 진행하였다. '오르기쟁이' 라는 인형이 끙끙대며 언덕을 올라가는 동안 두 개의 인형을 더 등장시켜 한 인형은 오르기쟁이 인형을 뒤에서 밀어주는 역할을 하고 한 인형은 오르기쟁이 인형을 비탈로 밀어 언덕 아래로 구르게 하는 역할을 한다. 몇 분이 지난 후 실험자들은 아이들에게 새로운 인형극을 보여주었다. 이는 오르기쟁이 인형이 도우미 인형과 방해꾼 인형을 번갈아 쳐다보다 결국 방해꾼 인형과 놀기로 한다는 내용이다. 그것은 말도 안 되는 일이었고, 아이들은 그 광경을 오르기쟁이 인형이 도우미 인형과 놀기로 했을 때 보다 더 오랫동안 쳐다보았다. 실험의 막바지에 연구진은 도우미 인형과 방해꾼 인형을 쟁반에 담아 아이들의 앞에 놓아두었고 아이들은 도우미 인형에게 훨씬 더 많이 손을 뻗는 경향을 보였다. 이를 통해 연구진은 '사회적 상호작용을 바탕으로 개개인을 평가하는 능력은 누구에게나 보편적으로 나타나며 이러한 능력은 학습만의 결과가 아니다' 라고 결론을 내렸다(Haidt, 2012/2014, pp. 132-134).

덧붙여, 하이트는 불승인 직관이 일부 선천적이라 하더라도 아동은 결국 문화 혹은 집단의 독특한 기준을 접하고 상호작용을 함으로써 이를 발전시켜 나간다고 지적하였다. 하이트는 이러한 과정을, 도덕성을 형성하는 도덕적 직관 즉, 불승인 직관을 문화적으로 수정하고, 강화하며, 억제하는 것으로 본다. 즉, 하이트에게 있어 불승인 직관의 전문성 향상이란

문화적 조형화와 직관 성숙의 문제이다. 이는 직관의 선택적인 상실(selective loss of intuition)과 관습 복합체에의 몰두(immersion in custom complexes) 그리고 또래집단 사회화(peer socialization)에 의해 이루어진다.

첫째, 직관의 선택적인 상실이란 개인이 소속된 문화에서 강조하는 도덕적 요소에 집중함으로써, 강조되지 않는 요소에 대한 예민한 반응이 상실되어 가는 기제를 의미한다. 하이트는 히긴스의 연구(E. Higgins, 1996)를 인용하며 문화적으로 강조되는 윤리와 관련된 직관은 더 예민하고 만성적으로 나타나며 강조되지 않는 윤리와 관련된 직관은 더 약해지고 나타나지 않는다고 주장한다(Haidt, 2001, p. 827).

둘째, 관습 복합체에의 몰두에 대해 살펴보면, 관습 복합체란 '관습적인 실제, 신념, 가치, 제재, 규칙, 동인 및 이와 관계된 충족'을 의미한다(Whiting & Child, 1953). 이러한 관습 복합체에 대한 몰두는 아동의 도덕적 사회화에서 발견된다. 예를 들어, 인도 오리사 지역의 아동들은 순수성으로 구조화된 신체와 공간을 지속적으로 접하게 되며, 그들은 순결과 불순에 대한 민감한 감각이 발달한다. 이러한 직관은 아동들에 내재화되며, 추후 아동들이 신성함과 관련된 윤리에 대한 인지적인 내용을 배우게 되면 이러한 직관 즉, 불순에 대한 불승인 직관은 합리화의 과정을 거쳐 더욱 강화된다(Haidt, 2001, p. 827).

셋째, 또래집단 사회화에 대해 살펴보면, 하이트는 인간을 주변의 도덕적 판단에 영향을 받으며 자신의 도덕적 판단을 형성해가는 사회적 존재로 보았다. 그는 해리스(Harris, 1995)를 인용하며 아동의 도덕 판단 형성에 가장 강력한 영향을 미치는 주변의 도덕 판단은 또래 집단의 도덕 판단임을 강조한다. 해리스에 따르면 아동들에게는 또래 집단에서 결속을 다지고 신망을 쌓는 것이 중요하기 때문에, 아동기 후기의 아동과 청소년의 과업은 부모님처럼 되어가는 것이 아닌 또래 집단에 적응해 가

는 것이 된다(Haidt, 2001, pp. 827-828).

종합하면, 피해와 공정성의 기준이 적용되지 않는 상황에 있어 불승인 직관의 발달이란, 직관의 선택적인 상실과 관습 복합체에의 몰두 그리고 또래집단 사회화에 의한 성숙과 문화적 조형화를 의미한다. 또한 피해와 공정성의 기준이 적용되지 않는 상황에 있어, 불승인 직관의 전문성이란 이러한 발달의 과정을 통해 충분한 문화화가 이루어진 불승인 직관이 도출되는 상태 즉, 고려되어야 할 다양한 문화 및 하위문화의 불승인의 기준들이 내재 및 조율되어 직관적으로 도출되는 상태라고 할 수 있다.

다음으로 피해와 공정성의 기준이 적용되는 전통적 도덕 상황에 있어 불승인 직관의 전문성 개념은 앞선 사회적 직관주의 이론 고유의 인지 유형간 보완 가능성 도출 작업을 통해 도출할 수 있다. 피해와 공정성의 기준이 적용되는 상황에 있어 불승인 직관은 설정된 목적과 별개로, 허용될 수 없는 행위를 불승인 하는데 기능하기에, 이는 허용될 수 없는 다양한 행위들에 대한 민감성을 갖추는 소극적 차원의 전문성이라 할 수 있다.

앞서 기술한 바와 같이, 도덕이 목적이 아닌 일상의 다양한 상황들에 있어 목적 지향적 판단은 '도덕 외의 목적'에 종속되는 경우가 많다. 이러한 경우에 있어, 불승인 직관은 개인이 비도덕적인 목적에 종속된 목적 지향적 인지에 허용할 수 없는 비도덕적 행위를 자행하는 것을 방지하는데 기능한다. 또한 개인이 도덕적 목적을 추구하는 경우에 있어서도, 개인은 왜곡된 정보 수용에 기반하거나 편중된 가치 선택에 기반하여 상황을 오해석하고 그에 기반하여 편향된 도덕적 목표를 설정하는 오류를 범할 수 있다. 즉, 일상의 상황에서 개인이 판단의 근거로 삼는 정보들은 한계가 존재할 수밖에 없으며, 목적 지향적 인지도 불승인 직관과 같이 왜곡된 감정적 정보 또는 암묵적 정보에 기반하는 측면이 있을 수밖에 없기에 개인이 도덕적 가치를 추구한다 하더라도, 구체적인 상황

에 있어 개인이 도덕적이라 믿는 선택이 실제로는 비도덕적인 선택일 가능성이 있다.

불승인 직관에 의한 도덕적 판단의 전문성 개념에 대한 논의에 이어, 목적 지향적 인지에 의해 도출된 도덕적 판단의 전문성에 대해 살펴보자면 다음과 같다. 이러한 목적 지향적 인지에 기반하는 도덕적 판단은 피해와 공정성의 기준이 적용되는 상황에 있어서만 도출된다. 이러한 목적 지향적 판단에 있어서의 전문성이란 상황에 내재한 도덕적 요소를 인지하고 해당 상황을 도덕의 상황으로 파악하며 가장 도덕적인 판단을 도출하는 것이다. 앞서 언급한 바 있는, 통합적 윤리교육 모델에서 강조되는 전문가로서의 윤리적 기술 즉, 레스트(J. Rest)의 4구성요소 모형에 근거하여 각 요소별로 도덕적 전문성을 신장시키기 위해 필요한 일련의 기술들(skills)이 바로 이러한 목적 지향적 인지에 의해 도출된 도덕적 판단의 전문성에 해당한다.

마지막으로 불승인의 직관적 판단에 대한 목적 지향적 인지의 수정에 의해 도출된 도덕적 판단의 전문성에 대해 살펴보면 다음과 같다. 방금 살펴본 목적 지향적 인지에 의해 도출된 도덕적 판단의 전문성이 도덕적 목적의 달성에 가장 적합한 행위 방안을 선택하는 적극적 차원의 전문성이라면, 불승인의 직관적 판단에 대한 목적 지향적 인지의 수정에 의해 도출된 도덕적 판단의 전문성은 불승인 직관의 위험성을 경계하는 것과 관련되는 소극적 차원의 전문성이라고 볼 수 있다.

이러한 목적 지향적 인지의 소극적 차원의 전문성은 첫째, 사회적 직관주의 모델의 추론된 판단의 과정과 같이, 자동 처리 과정으로서의 불승인 직관에 의해 도출된 편견적 판단을 통제하는 등의 역할을 성공적으로 수행하는 것이다. 둘째, 사회적 직관주의 모델의 사적인 반성적 사고 과정과 같이, 다양한 관점에서 상황을 바라보게 도움으로써 편향 및 왜곡된 불승인의 직관적 판단을 수정하도록 돕는 것이다. 특히, 피해와 공정성의 기준이 적용되는 상황에 있어서는 행위의 결과를 예상하고 결과

론적 근거를 제시함으로써 '불승인의 직관적 판단'을 수정하도록 돕는 것도 포함된다. 셋째, 불승인의 직관적 판단이 도출되는 과정에 영향을 미친 암묵적 정보에 대하여 인식할 수 있게 함으로써 판단자가 보다 객관적 관점에서 불승인의 직관적 판단을 통제할 수 있게 돕는 것이다. 넷째, 불승인 직관의 형성 및 발달 과정에 있어 다양한 불승인의 기준 및 감정들을 조율하는 방법을 모색함으로써 현실과 괴리되지 않은 성숙한 불승인 직관이 도출될 수 있도록 돕는 것이다.

종합하면, 도덕적 전문성의 향상을 위한 도덕적 스키마의 발달은 불승인 직관에 있어서의 전문성 발달과 목적 지향적 인지에 있어서의 전문성 발달을 포괄하여야 한다. 보다 세부적으로 살펴보면 첫째, 불승인 직관의 적극적 차원의 전문성이란, 피해와 공정성의 기준이 적용되지 않는 상황에 있어, 충분한 문화화가 이루어진 불승인 직관이 내재 및 조율되어 적절하게 도출되는 것이다. 둘째, 불승인 직관의 소극적 차원의 전문성은, 피해와 공정성의 기준이 적용되는 상황에 있어, 허용될 수 없는 다양한 행위들에 대한 발달된 민감성을 갖추는 것이다. 셋째, 목적 지향적 인지의 적극적 차원의 전문성이란, 피해와 공정성의 기준이 적용되는 상황에 있어, 상황에 내재한 도덕적 요소를 인지하고 해당 상황을 도덕의 상황으로 파악하며 가장 도덕적인 판단을 효과적으로 도출하는 것이다. 넷째, 목적 지향적 인지의 소극적 차원의 전문성이란 불승인 직관과 적극적 차원의 목적 지향적 인지에 내재하는 왜곡과 편향의 위험성을 효과적으로 경계하는 것이다.

2. 교육 방법 구성에의 적용 가능성

상기한 바와 같이 나바에츠는 도덕적 전문가를 양성하기 위한 교육방법 즉, 성숙한 도덕 기능 개발을 위한 교육 방법으로서 통합적 윤리교육(Integrative Ethical Education) 모델을 제시하였다. 그러나 앞서 언급한 바와 같이, 통합적 윤리교육 모델의 목표가 되는 도덕적 전문성에 대한 정의는 보완될 필요가 있다. 교육의 목표에 대한 세분화는 교육 내용에 대한 세분화로 이어져야 하기에, 본 절에서는 앞서 수행한 전문성 범주의 세분화 작업에 기반하여, 통합적 윤리교육 모델의 보완 가능성을 모색하고자 한다.

1) 통합적 윤리교육 모델의 특성

나바에츠는 전문가의 특성에 착안하여, 학습이란 해당 분야의 초보적 단계에서 전문가적 단계로 나아가는 과정이라고 지적하며 도덕적 전문가를 길러내기 위한 교육 방안을 모색하였다. 나바에츠에 따르면, 앞서 언급한 바와 같이, 전문가는 전문가로부터 멘토링을 제공받으며 잘 구조화된 환경에서 상호 작용을 통해 기술을 배우고, 전문 분야에서 개발된 명시적 이론을 사용하는 법과 메타 인지 및 문제 해결의 전략을 배우며, 자신의 영역에 많은 시간을 투여하며 집중적으로 연습한다(Lapsley & Power, 2005/2008, pp. 270-273).

초보자에서 전문가로 나아가는 과정에 기반하여, 나바에츠는 학생들을 도덕적 전문가로 이끌기 위한 다섯 단계의 지도 절차로서 도덕적 전문성 모델 즉, 통합적 윤리교육 모델을 제시한다. 각각의 단계들은 순차적으로 진행할 수도 있으며, 각 단계들을 결합하여 동시에 진행할 수도 있다. 이는 첫째, 각각의 학생들과 배려적 관계 형성하기이다. 둘째, 학생의 도덕적 인격과 도덕적 성취를 지지하는 교실 풍토 형성하기이다. 셋째, 교육과정 및 교과 외 활동 전범위에 걸쳐 '초보자에서 전문가로'

접근에 기반한 윤리적 기술 교육하기이다. 넷째, 자신을 객관적으로 바라볼 수 있는 메타인지 능력과 자기 규제 능력 배양하기이다. 다섯째, 도덕적 전문성 형성에 도움이 되는 공동체의 지지구조 회복하기 이다(Narvaez, Bock, 2014, pp. 148-153). 이에 대해 보다 세부적으로 살펴보면 다음과 같다.

(1) 배려적 관계 형성

나바에츠는 교사의 교수 태도가 학생의 학업과 학교에 대한 태도에 많은 영향을 미친다는 벤첼(K. R. Wentzel)의 연구를 인용하며, 성공적인 도덕 교육을 위해 교사가 학생과 배려적인 관계를 형성하는 것이 중요하다고 강조한다(Narvaez & Lapsley, 2008, p. 2). 나바에츠는 이를 위해 교사가 다음과 같은 역할을 수행해야 한다고 주장한다.

첫째, 교사는 학생 개개인과 전인적으로 관계를 맺음으로써 신뢰를 형성하여야 한다. 이는 심리상담에 있어 상담자가 내담자와 래포(rapport)를 형성하는 것과 유사하다. 이러한 과정은 인간적이고 배려적인 교실 분위기를 형성하기 위한 출발점이다. 둘째, 교사는 학생들과 건강한 감정적 신호(signal)을 주고받음으로써, 학생들의 감정이 올바르게 흘러갈 수 있도록 영향을 미칠 수 있어야 한다. 셋째, 교사는 학생 개개인에게 개별적인 배려를 제공하며, 그 과정에서 배려 제공자로서의 도덕적 모범을 보여주어야 한다. 나딩스에 의하면, 배려의 태도는 언어적 설명보다는 본보기를 통해 발달되므로 교사는 학생들과 배려적인 관계를 형성하고 직접적으로 배려를 제공하는 방식을 통해 모범을 보여주어야 한다. 넷째, 교사는 학생의 학업성취 및 도덕적 행동에 대한 높은 기대를 가지고 학생들에 대해 긍정적인 태도를 가져야 한다.

(2) 지지적 교실 풍토 형성

교실의 지지적인 풍토(supportive climate)란 학생들이 도덕적 인격을 형성하고 수행하는 것을 지지하는 구조화된 환경을 조성하는 것을 의미한다. 풍토는 드러나거나 숨겨진 상벌체계, 집단의 목적 및 목적에 대한 담론과 같은 사회 환경의 구조 및 사람들이 다른 사람을 대하는 태도, 함께 일하고 의사 결정 하는 방식, 집단 내의 기대 등을 포함한다(Narvaez, 2007, p. 137). 이러한 풍토 조성이 중요한 이유는 인간의 대부분의 행동이 암묵적 지식에 토대를 두고 있으며, 이러한 암묵적 지식은 상당 부분 사회문화적 환경에 의해 형성 및 조율되기 때문이다. 따라서 교사들은 학생들이 도덕적인 암묵적 지식을 가질 수 있는 풍토를 조성하는데 힘써야 한다.

솔로몬(D. Solomon)에 따르면 배려하는 학교와 교실 공동체는 다음과 같은 특징을 갖는다. 첫째, 교사는 존경스러운 행위 모델이며, 학생들에게 따뜻하고 수용적이며 지지적(supportive)인 태도를 갖는다. 둘째, 학생들은 교실의 중요한 결정에 영향력을 갖는다. 셋째, 학생들은 타인과 상호 작용하고 협력하며 중요한 이슈에 대한 열린 논의에 참여하는 기회를 갖는다. 넷째, 학생들은 사회적 기술을 연습하고 다른 사람들을 도울 기회를 갖는다(Narvaez, 2007, p. 137).

(3) 전문가로서의 윤리적 기술 발달

나바에츠는 레스트(J. Rest)의 4구성요소 모형에 근거하여 도덕성을 구성하는 네 가지 요소들(도덕적 민감성, 도덕적 판단, 도덕적 동기화 혹은 초점, 도덕적 행동)을 제시한 후 각 요소별로 도덕적 전문성을 신장시키기 위한 일련의 기술들(skills)을 제시한다(정창우, 2012, p. 346). 나바에츠가 제시한 도덕성의 네 가지 요소와 각 요소를 구성하는 일곱 가지 윤리적 기술 및 각 기술을 구성하는 세 가지 하위 기술들은 <표-6>과 같다.

<표-6> 도덕적 행동의 네 가지 요소와 기술 및 각각의 하위 기술(정창우, 2008, p. 277)

감수성 sensitivity	판단 judgment	중점 focus	행동 action
ES-1 : **감정표현의 이해** 감정의 확인과 표현 자신의 감정 조정하기 노여움과 공격성 다스리기	EJ-1 : **윤리적 문제의 이해** 정보 수집하기 문제 분류하기 윤리적 문제 분석하기	EF-1 : **타인 존중하기** 공손하고 예의 갖추기 공격적이지 않기 존경 나타내기	EA-1 : **갈등과 문제 해결** 상호 관계의 문제 해결하기 협상하기 개선하기
ES-2 : **다른 사람의 관점 채택** 대안적 관점 채택하기 문화적 관점 채택하기 정의(justice) 관점 채택하기	EJ-2 : **준칙(codes)의 사용과 도덕적 기준확인** 준칙 특징화하기 준칙을 분별하여 적용하기 준칙의 다양성 판단하기	EF-2 : **양심의 함양** 자제(극기)하기 영향력 관리하기 명예롭게 되기	EA-2 : **공손하게 주장하기** 인간의 욕구에 주목하기 주장기술 형성하기 공손한 수사법 사용하기
ES-3 : **다른 사람과 관계 형성** 다른 사람과 관계 맺기 배려 나타내기 친구 되기	EJ-3 : **일반적으로 추론하기** 객관적으로 추론하기 건전한 추론 사용하기 추론의 함정 피하기	EF-3 : **책임감 있게 행동하기** 의무 이행하기 훌륭한 담당자 되기 훌륭한 세계시민 되기	EA-3 : **지도자로서 솔선수범하기** 리더 되기 다른 사람을 위해, 그리고 다른 사람과 함께 솔선수범하기 다른 사람의 조언자 되기
ES-4 : **다양성에 반응하기** 그룹과 개인차 속에서 일하기 다양성 인식하기 다문화적으로 되기	EJ-4 : **윤리적으로 추론하기** 관점들 판단하기 표준과 이상에 대해 추론하기 행동과 결과에 대해 추론하기	EF-4 : **공동체의 일원되기** 협동하기 자원 공유하기 지혜 함양하기	EA-4 : **결정의 실행 계획하기** 전략적으로 사고하기 성공적으로 수행하기 자원의 사용 결정하기
ES-5 : **사회적 편견 통제하기** 편견 진단하기	EJ-5 : **결과 이해하기** 자신의 환경 선택하	EF-5 : **삶의 의미 찾기** 자신에게 집중하기	EA-5 : **용기 함양하기** 두려움 다스리기

편견 극복하기 관용 가르치기	기 결과 예측하기 결과에 대응하기	헌신적 태도 기르기 경이로움 함양하기	압력 견뎌내기 변화와 불확실성 다스리기
ES-6 : **상황을 해석하기** 일어나고 있는 일이 어떤 일인지 규정하기 도덕적으로 인식하기 창의적으로 대응하기	EJ-6 : **과정과 결과에 대해 깊이 생각해보기** 목표와 수단에 대해 추론하기 올바르게 선택하기 과정을 재구성하기	EF-6 : **전통과 제도에 가치 부여하기** 전통을 명확히 알고 가치 있게 여기기 사회구조를 이해하기 민주주의를 실천하기	EA-6 : **참을성 기르기** 확고부동 장애를 극복하기 능력을 형성하기
ES-7 : **의사소통 잘하기** 말하고 듣기 말하지 않고 다른 방법으로 의사소통하기 의사소통을 모니터하기	EJ-7 : **맞서서 극복하기** 긍정적인 추론을 적용하기 실망과 실패를 다루기 회복방법을 개발하기	EF-7 : **윤리적 정체성과 자기 통합성 개발하기** 좋은 가치를 선택하기 자신의 정체성 형성하기 자신의 잠재력에 도달하기	EA-7 : **열심히 일하기** 달성가능한 목표를 세우기 시간을 관리하기 자신의 삶을 책임지기

도덕성의 네 가지 요소와 각 요소를 구성하는 일곱 가지 윤리적 기술 및 각 기술을 구성하는 세 가지 하위 기술들에 대한 교수 과정은 네 단계로 이루어진 단계적 수준에 따라 전개된다(Narvaez & Bock, 2014, p. 151). 이는 기술에 대한 전문가의 지식을 식별하는 단계, 그러한 지식을 정교화 하는 단계, 기술을 특정한 맥락에서 실천하는 절차적 측면을 연마하는 단계 그리고 지식과 절차를 온전히 통합하는 단계로 전개된다. 즉, 첫 번째와 두 번째 단계는 윤리적 기술의 내용적 측면에 대한 함양을 목표로 하며, 세 번째 단계는 윤리적 기술의 절차적 측면에 대한 함양, 네 번째 단계는 윤리적 기술의 내용적 측면과 절차적 측면에 대한 유기적 통합을 목표로 한다(석자춘, 정창우, 2017, p. 36).

나바에츠는 이러한 전문가 양성의 교수방법에 도제식 교육이 효과적임을 주장하며, 지식과 기술의 단순 습득이 아닌 '인지적' 지식과 기술 습득을 목표로 하는 콜린스(A. Collins), 브라운(J. Brown), 뉴먼(S. Newmann)의 '인지적 도제 모델(cognitive apprenticeship model)'

을 차용한다. ‘도제’란 학습이나 지식 습득이 체험을 통해 이루어지는 교육 방식으로, 인지적 도제 모델은 특정 상황과 맥락 안에서 실제 과제 해결 과정을 체험함으로써 인지적 지식과 기술을 내면화하고 높은 수준의 전문가적 수행능력을 키우는 것을 목적으로 한다(추병완, 2000, p. 40). 이 모델이 제시하고 있는 몇 가지 수업 원리들을 제시하면 <표-7>과 같다.

<표-7> 인지적 도제 모델의 수업 원리(Wilson, Jonassen, & Cole, 1993, pp. 8-9; **추병완**, 2000, pp. 40-41)

내용	암묵적이고 발견적인 지식을 가르쳐야 한다.
상황학습	실생활에서 사용되어질 방식을 반영하고 있는 맥락 속에서 지식과 기능을 가르쳐야 한다.
시범과 설명	과정이 어떻게 전개되는지 학생들에게 시범을 보여주고, 그러한 방식으로 전개되는 이유를 설명해주어야 한다.
지도와 피드백	적절한 힌트, 도움, 고무적인 피드백을 사용하는 가운데 학습자의 수행에 대한 개별적인 관심을 보여주어야 한다.
비계설정과 제거	학생들이 수행할 수 없는 과제의 부분들을 교사가 수행함으로써 학생들을 지원해 주어야 한다. 점차적으로 비계설정의 양을 줄여나감으로써, 학습에 대한 통제권을 학생들이 가질 수 있게 해 주어야 한다.
정교화와 성찰	학생들로 하여금 자신들의 행동에 대하여 사고하고 그렇게 하게 된 이유를 설명하게 함으로써 학생들이 암묵적인 지식을 더욱 분명하게 만들 수 있도록 해 주어

	야 한다.
탐색	학생들로 하여금 상이한 전략들을 시험해보고, 그 결과들을 관찰할 수 있게 해야 한다. 이것은 학생들이 이전 지식을 활용하여 연습할 수 있는 기회를 부여함과 동시에 그러한 지식들을 해결이 필요한 문제들과 결합할 수 있도록 해 준다. 만약 학생들이 오개념을 만들어내면, 그들이 예외나 반대사례를 접할 수 있게 해 주어야 한다.
계열성	다양성을 늘려나가는 가운데, 간단한 것에서부터 복잡한 것으로 나아가야 한다. 기저의 원리를 먼저 가르친 후에 그 원리를 특정한 수행 상황에 적용하는 것을 세부 조정해 주어야 한다.

(4) 메타인지 능력 및 자기 규제 능력 발달

전문성이란 해당 영역에서의 기술적 노하우로서의 전문성뿐 아니라 자기규제능력에 있어서의 전문성을 포괄한다. 따라서 전문가 양성을 위한 교육 프로그램은 자기 자신을 돌아보는 메타인지 능력의 발달을 위한 교육 즉, 자기규제 기술(self-regulation skills)에 대한 교육을 포함하여야 한다. 짐머만(B. Zimmerman)에 따르면, 자기규제 기술은 세 가지 메타인지 기술로 정리할 수 있다. 첫째, 스스로가 유혹에 흔들리지 않기 위해 자신의 주의를 유혹으로부터 분리하는 자기 점검(self-monitoring) 기술이다. 둘째, 목표 달성 과정에 있어 스스로를 독려하며 지속적으로 나아가게 하는 자기옹호(self-cheerleading) 기술이다. 셋째, 목표달성에 효과적인 환경을 선택하고 조성하는 기술이다(Narvaez & Lapsley, 2008, p. 5).

나바에츠에 따르면, 이러한 자기규제 기술은 직관과 추론이 동시에 작

동되는 복합적인 과정이라 할 수 있다. 자신의 직관적 판단 및 충동적 감정들에 대해 숙고함으로써 행동을 규제하고, 직관이 올바른 방향으로 발달할 수 있도록 환경을 선택하는 등의 자기 규제는 추론과 직관의 상호 작용을 필요로 한다. 이에, 나바에츠는 자기규제능력을 신장하기 위해서는 직관적 정신(intuitive mind) 을 올바르게 기를 수 있는 숙고적 정신(deliberative mind)을 발달시켜야 한다고 강조한다(Narvaez, 2007, pp. 139-140).

(5) 공동체의 지지구조 회복

나바에츠는 개인의 도덕적 인격발달에 영향을 미칠 수 있는 모든 사회 구성원을 교육의 주체에 포함시킨다. 이는 부모와 교사, 교직원 등과 같이 아이들에게 유의미한 성인 및 공동체의 구성원을 포괄한다. 아이들은 주변인들로부터 쉽게 영향을 받고, 모방하기 때문에 성인들은 도덕 교육자로서의 역할을 수행하게 된다. 따라서 공동체의 성인들은 학생들의 친사회적 정서 및 도덕적 행동을 발달시키고 회복탄력성을 키워주는 등 도덕적 인격을 발달시키는데 도움이 되는 역할을 담당하는데 관심을 가져야 한다.

2) 통합적 윤리교육 모델의 보완 가능성

나바에츠의 통합적 윤리교육 모델에 있어, 전문성 교육의 특징이 직접적으로 가장 잘 나타나는 부분은 세 번째 단계인 '윤리적 기술 발달을 위한 교육'이라 할 수 있다. 이를 통해 나바에츠는 도덕이라는 목표 설정을 전제로, 숙련 즉, 수월성으로서의 직관이 강조된 전문가 개념을 제시하고 구체적인 기술을 나열(<표-6>)하며, 이를 효과적으로 가르치기 위한 수업 모델로서 인지적 도제모델(<표-7>)을 제시하고 있다. 선행한 전문성 개념의 보완 작업에 따를 경우, 세 번째 단계인 '윤리적 기술

발달을 위한 교육’은 주로 전문성 개념의 세 번째 특성 즉, 상황에 내재한 도덕적 요소를 인지하고 해당 상황을 도덕적 상황으로 파악하며 가장 도덕적인 판단을 효과적으로 도출하는 목적 지향적 인지의 적극적 차원의 전문성을 전제로 하며, 다른 차원의 전문성까지는 충분히 포괄하고 있지 못하다고 평가할 수 있다.

특히, 모든 교사가 도제식 교육을 시행할 수 있는 역량을 갖춘 도덕적 전문가인 것은 아니며, 입시 등의 다양한 목표가 강조되는 학교 현장에서 개인별로 장시간의 집중적인 도제식 교육을 제공하는 것에는 다양한 제약이 있을 수밖에 없다. 그렇기에, 도덕 기능의 사실적 기제를 반영하여 학생들의 윤리적 기술을 효과적으로 발달시킬 수 있는 교육을 제공하기 위해서는, 보완된 전문성 개념을 반영하여 통합적 윤리교육 모델을 보완할 필요성 있다. 이를 위해서는 도덕적 전문성 교육이 이루어지는 상황 유형별로, 각 유형의 상황에 있어 기능하는 인지의 기제와 해당 인지를 보완하기 위한 인지의 기제에 있어 전문성 신장의 목표를 설정하고 구체적인 교육 방안을 모색하여야 한다. 이에 대하여 보다 자세히 살펴보면 다음과 같다.

(1) 교육 내용의 구성

교육 내용을 구성함에 있어 주목해야 할 점은, 도덕적 전문성을 향상시키기 위한 교육을 효과적으로 적용하기 위해서는 도덕적 기술의 습득이 효과적으로 이루어질 수 있는 상황을 구성할 필요성이 있다는 것이다. 직관의 습득 및 발현이 상황 맥락성을 갖는다는 점을 고려할 때, 효과적이고 전략적인 상황 맥락의 구성은 효과적인 교육 내용 구성 방안을 모색함에 있어 가장 핵심적인 요소 중 하나이다.

도덕적 전문성이란 도덕 영역의 사실 및 정보에 관한 선언적 지식(declarative knowledge)과 이를 적용할 수 있는 절차적 지식(procedural knowledge) 및 조건적 지식(conditional knowledge)이

잘 조직된 풍부한 스키마를 전제로 하며(Narvaez, 2010, p. 171), 이러한 다양한 범주의 지식들이 잘 조직화되기 위해서는 다양한 지식들의 복합적인 작용이 이루어질 수 있는 상황 즉, 도덕적 판단의 도출 과정에 기능하는 다양한 인지 요인들을 자극할 수 있는 상황이 제공되어야 한다. 선행한 작업에 기반할 때, 도덕적 전문성 교육을 위해 구성된 상황은 목적 지향적 인지와 불승인의 직관적 인지가 적절하게 보완적으로 작용하며 발달될 수 있는 요소를 다양하게 포함하는 상황이어야 한다.

한 가지 짚고 넘어가야 할 점은, 이러한 상황 구성은 기술의 목록을 정해놓고 이를 교육하기 위한 상황들을 제시하는 방식으로 수행되어서는 안 된다는 것이다. 기술 목록을 중심으로 상황을 제시할 경우 첫째, 더욱 집중적인 교육이 필요한 미발달된 기술과 이미 충분히 발달된 기술 사이에 차등적인 비중을 두어 교육을 진행하는 것이 어려워지며 둘째, 기술 간 협응 방법을 배우는 것이 아닌, 특정 기술과 관련된 상황의 목록을 암기하는 수준의 피상적인 교육에 머무를 위험이 있다. 즉, 기술을 중심으로 상황을 구성하는 방식은 '상황에 대한 학습'이 아닌 '학습을 위한 상황의 구성'에 머무를 위험성을 갖는다.

이성과 감정이 분리되어 작용하지 않고 밀접하게 상호작용하며 추론과 직관이라는 인지 유형으로 도출되는 것처럼, 다양한 종류의 직관과 추론 또한 하나의 판단이 도출되는 과정에 있어 분리되어 작용하지 않으며 밀접하게 상호작용한다. 신뢰할만한 직관은, 환경에 대한 예측 가능성(Predictability of the Environment)과 오랜 시간의 경험을 통한 연습을 조건으로 하며(Kahneman & Klein, 2009), 전문가들이 많은 경험들을 통해 직관을 발달시킬 수 있었던 것은 경험들을 패턴화하여 즉각적으로 활용할 수 있는 직관으로 형성했기 때문(이정렬, 2017, pp. 123-124)이라는 주장을 반영할 때, 다양한 유형의 인지 간 상호 작용을 경험할 수 있는 '상황에 대한 학습'이 전문가로서의 수월성을 발달시킴에 있어, 보다 적합한 방식이라고 볼 수 있다.

덧붙여, 이러한 상황의 구성은 학생의 발달 수준에 따라, 단일한 기술

또는 지식의 대입을 통해 해결되는 간단한 상황에서 복잡한 고려와 기술이 필요한 일상의 상황 및 복잡한 딜레마 상황 등으로 다층화 되어야 한다. 즉, 학교는 학생들에게 점차 더 많은 상황 구성 요소들을 제시함으로써, 학생들의 도덕적 스키마가 전문가 수준으로 발달해 갈 수 있도록 하여야 한다.

선행한 작업에 기반할 때, 도덕적 판단이 도출되는 상황은 크게 '불승인 직관이 수반되는 상황'과 '불승인 직관이 수반되지 않는 상황'으로 구분할 수 있다. 이를 더 세부적으로 구분하면 '피해와 공정성의 기준이 적용되며 불승인의 직관적 판단이 도출되는 상황'과 '피해와 공정성의 기준이 적용되지 않으며 불승인의 직관적 판단이 도출되는 상황' 그리고 '피해와 공정성의 기준이 적용되며 목적 지향적 판단이 도출되는 상황'과 같이 한 가지 유형의 인지에 의해 판단이 도출되는 상황과 '피해와 공정성의 기준이 적용되며 불승인 직관과 목적 지향적 인지가 경쟁하는 상황'과 같이 여러 유형의 인지가 개입되어 판단이 도출되는 상황으로 구분할 수 있다. 이에 기반하여, 간단한 상황에서 복잡한 상황으로 심화되는, 전문성 교육을 위한 상황 구성 원리를 도출하면 <표-8>과 같다.

<표-8> 전문성 교육을 위한 상황 구성 원리

학습 단계	학습 단위가 포함해야 할 상황 유형		
	피해와 공정성의 기준이 적용되지 않으며, 불승인 직관이 수반되는 상황	피해와 공정성의 기준이 적용되며, 불승인 직관이 수반되는 상황	피해와 공정성의 기준이 적용되며, 불승인 직관이 수반되지 않는 상황

1단계	세 가지 상황 유형 중 한 가지 유형만 포함 (단, 목적 지향적 인지는 도덕을 목표로 제시)
2단계	세 가지 상황 유형 중 두 가지 유형 포함 (단, 목적 지향적 인지는 도덕을 목표로 제시)
3단계	세 가지 상황 유형 모두 포함 (단, 목적 지향적 인지는 도덕을 목표로 제시)
4단계	세 가지 상황 유형 모두 포함 (도덕 외의 목표가 도출될 수 있는 상황 제시)

1단계는 단일 유형의 인지 전문성을 신장시키고 인지의 한계를 보완할 수 있도록 상황 유형을 단순하게 제시하는 초급 단계이며, 이후의 단계들은 복잡한 상황 속에서 각종 유형의 인지들을 조율함으로서 도덕적 전문성을 신장시키는 연습을 제공하는 것을 목표로 한다. 가장 기본이 되는 세 가지 유형의 상황에 있어 유형별 인지의 전문성 신장을 위한 교육은 다음과 같은 내용을 포함하여야 한다.

세 가지 상황 유형 중 첫 번째 상황 유형은 불승인 직관에 의해 판단이 도출되는 상황이다. 상기한 작업에 기반 할 때, 이러한 불승인 직관은 충분한 문화화의 과정에서 조율 및 내재된 기준에 민감하게 반응하거나 적절하게 설정된 피해와 공정성의 기준에 민감하게 반응할 때 그 전문성이 발현된 것으로 볼 수 있다. 그러나 이러한 불승인 직관은 왜곡된 정보에 기반 할 수도 있으며, 비도덕적인 문화적 기준에 근거할 수도 있다.

그렇기에 불승인 직관에 의해 판단이 도출되는 첫 번째 상황 유형에 있어서의 전문성 교육은 타인의 입장에서 감정을 이입해보는 등의 기회를 제공함으로써 민감성의 향상을 도모하여야 하고, 자신의 인지를 객관

적으로 점검할 수 있는 기회를 제공함으로써 불승인 직관의 왜곡과 편향을 경계할 수 있도록 하여야 한다. 즉, 이는 '불승인 직관의 적극적 차원의 차원의 전문성'과 그 한계를 보완하기 위한 '목적 지향적 인지의 소극적 차원의 전문성'을 향상시키는 내용으로 구성되어야 한다.

두 번째 상황 유형은 피해와 공정성의 기준이 적용되며, 불승인 직관에 의해 판단이 도출되는 상황이다. 상기한 작업에 기반 할 때, 이러한 상황에서의 불승인 직관의 전문성은 허용될 수 없는 다양한 행위에 대한 민감성을 갖추는 것이다. 그러나 이러한 불승인 직관은 왜곡된 정보에 기반 하거나 이기적인 관점에 근거할 수도 있으며, 피해와 공정성의 기준에서 고려해야 할 다양한 요소나 관점 중 일부 요소나 관점에만 편향되어 도출될 수도 있다.

그렇기에 두 번째 상황 유형에 있어서의 전문성 교육은 다양한 관점에서 상황을 바라보는 기회를 제공하여야 하고, 자신의 인지를 객관적으로 점검할 수 있는 기회를 제공함으로써 불승인 직관의 왜곡과 편향을 경계할 수 있도록 하여야 한다. 즉, 이는 '불승인 직관의 소극적 차원의 차원의 전문성'과 그 한계를 보완하기 위한 '목적 지향적 인지의 적극적 및 소극적 차원의 전문성'을 향상시키는 내용으로 구성되어야 한다.

세 번째 상황 유형은 피해와 공정성의 기준이 적용되며, 불승인 직관의 개입 없이, 목적 지향적 인지에 의해 판단이 도출되는 상황이다. 상기한 작업에 기반 할 때, 이러한 목적 지향적 인지의 적극적 차원의 전문성이란, 상황에 내재한 도덕적 요소를 인지하고 해당 상황을 도덕적 상황으로 파악하며 가장 도덕적인 판단을 효과적으로 도출하는 것이다. 그러나 이러한 목적 지향적 인지의 적극적 차원의 전문성은 도덕 외의 목적 또는 비도덕적 목적을 전제로 도출되거나 왜곡된 정보에 기반 하여 도출될 수도 있으며, 고려해야 할 다양한 관점 중 일부에만 편향되어 도출되거나 허용될 수 없는 수단을 활용할 수도 있다.

그렇기에 목적 지향적 인지에 의해 판단이 도출되는 상황에 있어서의 전문성 교육은, 상황에 포함된 도덕적 요소를 파악하고 가장 도덕적인

행위를 선택하는 훈련의 기회를 제공하여야 하며, 타인의 입장이 되어보거나 다른 관점에서 상황을 바라보는 기회를 제공함으로써 허용할 수 없는 행위를 경계하는 훈련의 기회를 제공하여야 하고, 자신의 인지를 점검할 수 있는 기회를 제공함으로써 목적 지향적 인지와 불승인 직관의 왜곡과 편향을 경계할 수 있도록 하여야 한다. 즉, 목적 지향적 인지에 의해 판단이 도출되는 상황에 있어서의 전문성 교육은 '목적 지향적 인지의 적극적 차원의 전문성'과 그 한계를 보완하기 위한 '불승인 직관의 소극적 전문성' 및 '목적 지향적 인지의 소극적 전문성'을 향상시키는 내용으로 구성되어야 한다.

(2) 교수 전략의 설정

상황 학습은 교사가 전달하고자 의도한 순서와 절차에 따라 이루어지는 것이 아니며, 학습자와 환경의 상호작용 가운데 즉각적이고 즉흥적으로 때론 비합리적이라고 여겨질 만큼 비순차적이고 비논리적인 순서에 의해 비선형적으로 의미를 창출하며 이루어진다. 또한 교사의 교수행위는 학습의 방향을 구성하고 행위에 의미를 부여해주는 학습의 설계로서 역할을 할 수는 있지만, 실제로 그 탐구 활동을 경험하는 것은 학생 자신이고, 이러한 과정에서 개별 학습자의 상황 인지는 차이를 나타낼 수밖에 없으며, 이러한 차이는 학습을 위한 지식을 계열화 순서에 차이를 유발한다(조현영, 2016, p. 116).

그렇기에 상기한 바와 같이 선별적으로 구성된 상황을 전제로 구체적인 전문성 교육의 내용을 설정하였다면, 도덕적 스키마 발달의 효율적인 전략을 구성하기 위한 기초 작업으로서, 학생들의 스키마가 갖는 특성에 대한 점검이 필요하다. 특정한 상황을 마주한 개인이 해당 상황을 어떻게 받아들이고 어떠한 요소에 초점을 맞추어 어떠한 판단을 내리는지를 살피는 작업이 선행되지 않는다면, 집단을 대상으로 이루어지는 학교 교육을 전략적으로 구성하기에 어려움이 수반될 수밖에 없다. 예를 들어,

‘주어진 상황이 도덕적 상황이 아니며 그렇기에 도덕적 기준을 적용할 필요가 없다고 판단하는 학생’에 대한 수업의 내용과 ‘주어진 상황을 도덕적 상황이라고 판단하지만 기준을 잘못 적용하는 학생’에 대한 수업의 내용은 달라질 수밖에 없다. 또한 ‘특정 행위나 대상에 대한 왜곡 또는 편향된 불승인 직관에 의해 판단을 내리는 학생’에 대한 교수 방법과 ‘목적 지향적 판단 과정에서 고려해야 할 사항을 누락한 학생’에 대한 교수방법은 달라질 수밖에 없다.

스키마 점검을 위한 체크 리스트는, 유형별 상황에 있어 도덕적 판단의 도출에 기능하는 인지의 전문성 신장 및 해당 인지의 한계를 보완하는 인지의 전문성을 신장시키기 위한 교육 내용에 기반하여 도출되어야 한다. 또한 수업에서 중점적으로 다루어야 할 구체적인 교육적 초점은 이러한 스키마 점검 결과를 반영하여 선정되어야 한다. 또한 선정된 단위 수업의 교육 내용을 효과적으로 전달하기 위한 교수 방법 즉, 해당 인지 유형의 전문성 향상을 위한 효과적인 교수 방법의 설정은 해당 인지의 발달 및 수정의 기제를 반영하여야 한다. 예를 들어, 불승인의 직관적 판단에 있어, 근거의 제시는 판단 이후에 도출되는 것이고 근거가 유지될 수 없는 상황에서도 판단은 유지되기에, 불승인의 직관적 판단을 뒷받침하는 근거를 반박하는 방식으로 학생을 설득하려고 하여서는 안 된다. 상기한 작업을 종합하여 교육 방안의 구성 순서를 정리하면 <표-9>과 같다.

<표-9> 도덕적 전문성 향상을 위한 교육 방안의 구성 순서

상황 유형	피해와 공정성의 기준 미적용	피해와 공정성의 기준 적용	
판단의 도출 기제	불승인의 직관적 판단	불승인의 직관적 판단	목적 지향적 판단

판단의 도출 유형별 전문성 교육의 내용	1. 불승인 직관의 전문성 향상을 위한 내용 2. 불승인 직관의 한계를 보완하기 위한 목적 지향적 인지의 소극적 차원의 전문성 향상을 위한 내용	1. 불승인 직관의 전문성 향상을 위한 내용 2. 불승인 직관의 한계를 보완하기 위한 목적 지향적 인지의 적극적 및 소극적 차원의 전문성 향상을 위한 내용	1. 목적 지향적 인지의 적극적 차원의 전문성 향상을 위한 내용 2. 적극적 차원의 목적 지향적 인지의 한계를 보완하기 위한, 불승인 직관의 전문성과 목적 지향적 인지의 소극적 차원의 전문성 향상을 위한 내용
스키마 점검 기준 설정	판단의 도출 또는 판단의 한계 보완에 기능하는 유형별 인지 중, 어떠한 인지의 전문성이 어떠한 방식으로 발현되거나 발현되지 못하고 있는가.		
단위 수업의 교육 내용 선정	판단의 도출 유형별 전문성 교육의 내용 중 어떠한 것에 중점을 맞추어 가르칠 것인가.		
교수 방법의 설정	단위 수업의 교육 내용으로 선정된 유형별 인지를 강화 또는 수정하기 위해 고려해야 하는 심리 과정의 기제에는 무엇이 있는가.		

도덕적 전문성 교육이 성공하기 위해 도제식 교육의 주체로서 교사는 스스로의 도덕 기능에 대한 전문성을 갖추어야 함은 물론이고, 유형별 인지의 특성 등에 관한 심리학적 전문성과 이를 반영한 수업 전략의 구성 등과 관련된 교육학적 전문성 또한 갖추어야 한다. 교사가 스스로의 도덕 기능에 관한 전문성 없이 심리적 전문성과 교육적 전문성만 갖출 경우, 학생들에게 잘못된 전문성을 효과적으로 전달할 위험이 있으며, 도덕 기능에 관한 전문성을 갖추었지만 심리학적 전문성과 교육학적 전

문성을 갖추지 못할 경우, 학생들에게 전문가로서의 노하우를 제대로 전달하지 못할 수 있다.

Ⅶ. 결론

사회적 직관주의 이론은 인지주의 접근의 도덕 심리학이 이성과 감정을 대립적인 관계로 규정함으로써 충분하게 설명하지 못했던 다양한 도덕적 판단의 현상을 더욱 타당하게 규명하는 데 기여하였다. 그러나 하이트의 사회적 직관주의 모델과 그린의 이중과정 모델은 직관과 추론의 역할 및 관계 방식과 관련하여 대립하는 주장을 펼치는 관계로 상정되었으며, 그로 인하여 두 이론의 발견은 합의점을 도출할 수 없는 것으로 간주되었다. 이에 본 연구는 과연 두 이론이 상충하는 관계이며 합의점을 도출할 수 없는지 여부에 대해 살펴보았다.

덧붙여, 두 사회적 직관주의 이론은 인지 심리학적 개념 범주를 차용하여 현상을 해석하였고, 도덕적 판단의 현상을 대상으로 인지 심리학적 규명을 재확인하였다. 이와 관련하여 본 연구는 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 보다 타당하게 해석하기 위해, 추론과 직관이라는 인지 심리학적 개념 범주를 차용하는 것이 타당한지에 대하여 검토하고 이를 보완할 수 있는 사회적 직관주의 이론 고유의 개념 범주 마련 가능성을 모색하였다.

이러한 비판적 검토 작업을 통해 본 연구는 다음과 같은 결론을 얻었다. 첫째, 규명 범주 구체화 작업에 기반할 때, 두 사회적 직관주의 이론은 상충하는 관계가 아니며, 오히려 보완적으로 기능한다. 두 이론이 주장의 근거로 제시하는 인터뷰 문항을 비교할 때, 하이트의 사회적 직관주의 모델의 핵심 주장은 '피해와 공정성의 기준이 적용될 수 없는 상황'을 규명 범주로 하여 도출되었으며, 그린의 이중과정 모델의 핵심 주장은 '피해와 공정성의 기준이 적용되는 상황'을 규명 범주로 하여 도출되었다. 즉, 두 이론은 도덕적 판단이라는 현상을 구성하는 상이한 하위 범주의 현상을 각각 규명한 보완적인 관계로 파악하는 것이 타당하며, 두 이론 중 어떠한 이론이 더욱 타당한지를 비교하는 논의는 성립하

기 어렵다.

둘째, 직관과 추론이라는 개념 범주를 차용하며 감정적 활성화 정도만을 기준으로 삼을 경우, 두 이론 내부에 모순된 해석의 위험성이 발생한다. 그린은 감정적 활성화 정도를 기준으로 도덕과 무관한 딜레마(Non-Moral Dilemma)에서의 판단을 추론적 판단으로 분류하였으나, 직관과 추론을 구분하는 인지 심리학의 다양한 기준들을 고려할 때, 도덕과 무관한 딜레마 상황에서의 판단을 직관적 판단과 추론적 판단 중 한 종류의 판단으로 명확하게 분류하기에 어려움이 있다. 즉, 감정의 활성화 정도만을 가지고 직관과 추론을 구분하는 것은 모순된 해석의 위험성을 내포한다.

또한, 딜레마 유형별 응답 소요 시간 그래프를 기준으로 할 때, 그린이 추론적 판단으로 분류한 비인신적 딜레마(Moral Impersonal Dilemma) 상황에서의 긍정적 응답 도출에 소요된 시간이 그가 직관적 판단으로 분류한 인신적 딜레마(Moral Personal Dilemma) 상황에서의 부정적 응답 도출에 소요된 시간보다 더 짧다. 이러한 실험 결과는 직관적 판단이 신속한 판단으로서의 특성을 갖고 추론적 판단이 느린 판단으로서의 특성을 갖는다는 통설과 정면으로 충돌한다. 즉, 그린이 추론으로 분류한 비인신적 딜레마 상황에서의 판단은 덜 감정적이고 더 의식적이며, 판단의 과정을 파악할 수 있다는 추론적 판단의 특성들을 가지지만, 동시에 신속하다는 직관적 판단의 특성도 갖는다. 이 또한 감정의 활성화 정도만을 가지고 직관과 추론을 구분하는 것이 모순된 해석의 위험성을 내포하는 성급한 분류라는 것을 보여준다.

하이트의 사회적 직관주의 모델에 있어서도 이러한 모순된 해석의 위험성이 존재한다. 하이트가 추론으로 분류한 인지 또한 상기한 바와 같이, 숙달 정도와 자동성 등에 따라 직관으로도 분류할 수 있고 추론으로도 분류될 수 있다. 그렇기에 모순된 해석의 위험성을 제거하고 두 사회적 직관주의 이론이 규명한 도덕적 판단의 현상을 보다 타당하게 해석하기 위한 개념 범주의 보완 또는 대체가 필요하다. 또한 이러한 보완적

또는 대체적 개념 범주는 하이트와 그린이 직관 및 추론으로 범주화 한 대상이 어떠한 실제를 지칭하는지 명료화 하는 작업을 통해 도출할 수 있다.

하이트와 그린이 직관으로 분류한 개념 범주는, 나바에츠(D. Narvaez)가 내용적 차원의 제한 없이 수월성과 자동성이라는 형식적 차원의 특성에 기준하여 직관이라는 개념을 사용하는 것과 달리, 불승인의 감정 반응을 필요조건으로 하며, 그렇기에 내용적으로 불승인에 특정된 '불승인 직관'으로 지칭하는 것이 타당하다. 또한 상기한 바와 같이, 두 사회적 직관주의 이론이 추론으로 분류한 인지 과정 즉, 불승인 직관 외의 인지는 추론으로 분류될 수 있는 가능성과 직관으로 분류될 수 있는 가능성을 모두 갖는다. 그렇기에 두 사회적 직관주의 이론이 추론으로 분류한 심리 과정은 형식적으로 특정되지 않는 내용적 차원의 개념일 필요가 있으며, 이는 '목적 지향적 인지'로 지칭할 수 있다.

셋째, 이러한 규명 범주 구체화와 개념 범주 명료화 작업에 기반 할 때, 두 사회적 직관주의 이론의 발견은 다음과 같이 정리할 수 있다. (1)불승인 직관이 강하게 도출 되는 상황에 있어, 도덕적 판단은 피해와 공정성의 기준 적용 가능 여부와 무관하게 불승인의 직관적 판단으로 도출된다. (2)피해와 공정성의 기준이 적용되지 않는 상황에서는 불승인 직관이 존재하지 않는 경우, 어떠한 도덕적 판단도 도출되지 않는다. 그러나 피해와 공정성의 기준이 적용되는 상황에서는 불승인 직관이 존재하지 않더라도 목적 지향적 판단의 형태로 도덕적 판단이 도출될 수 있다. (3)불승인의 직관적 판단이 도출된다 하더라도 목적 지향적 인지에 의해 수정이 이루어질 수 있다. 단, 피해와 공정성의 기준이 적용되지 않는 상황에서 이러한 수정은 매우 드물게 일어나며, 피해와 공정성의 기준이 적용되는 상황에서는 강한 결과론적 근거에 의해 이러한 수정이 더욱 용이하게 이루어지기도 한다.

넷째, 새로운 인지 유형 분류를 적용할 경우, 의무론적 판단과 결과론적 판단은 상호 보완적인 관계를 맺는다. 의무론적 판단이란 '불승인

직관’에 의해 추동된 판단으로, 오직 특정 행위의 허용 불가능성 여부를 판단하는 것만을 내용으로 한다. 반면, 결과론적 판단은 목적 지향적인 판단으로서, 목적의 달성에 더욱 적합한 행위를 선택하는 것을 내용으로 한다. 즉, 의무론적 판단은 목적과 독립적으로 특정 행위가 수용 가능한 범주를 벗어난 용인될 수 없는 행위인지 여부를 판단하는 데 기능하며, 결과론적 판단은 행위 선택지 중 목적의 달성에 부합하는 행위를 선택하는 데 기능한다. 즉, 두 유형의 판단은 도덕 기능에 있어 상이한 역할을 담당함으로써 보완적으로 작용하며, 그렇기에 의무론적 판단이 윤리적 가치를 갖기 어렵다는 그린의 주장은 타당성을 갖기 어렵다.

다섯째, 동일한 맥락에서 도덕적 전문성은 목적 지향적 인지의 전문성과 불승인 직관의 전문성으로 구분된다. 이러한 전문성은 앞서 도출한 세 가지 유형의 도덕적 판단에 있어, 판단의 수월성을 향상시키는 적극적 차원의 전문성과 각각의 인지에 내재한 위험성을 경계하는 소극적 차원의 전문성을 포괄하는 개념이어야 하며 이는 다음과 같다. (1)불승인 직관의 적극적 차원의 전문성이란, 피해와 공정성의 기준이 적용되지 않는 상황에서, 충분한 문화화가 이루어진 불승인 직관이 내재 및 조율되어 적절하게 도출되는 것이다. (2)불승인 직관의 소극적 차원의 전문성은, 피해와 공정성의 기준이 적용되는 상황에서, 허용될 수 없는 다양한 행위들에 대한 발달된 민감성을 갖추는 것이다. (3)목적 지향적 인지의 적극적 차원의 전문성이란, 피해와 공정성의 기준이 적용되는 상황에서, 상황에 내재한 도덕적 요소를 인지하고 해당 상황을 도덕의 상황으로 파악하며 가장 도덕적인 판단을 효과적으로 도출하는 것이다. (4)목적 지향적 인지의 소극적 차원의 전문성이란 불승인 직관과 적극적 차원의 목적 지향적 인지에 내재하는 왜곡과 편향의 위험성을 효과적으로 경계하는 것이다.

여섯째, 도제식 교육이 효과적으로 이루어지기에 많은 제약이 있는 학교 현장에서, 윤리적 기술을 기르기 위한 인지적 도제 모델을 보다 효과적으로 시행하기 위해서는 윤리적 기술의 습득이 효과적으로 이루어질

수 있는 상황을 구성할 수 있어야 한다. 특히, 직관적 기술의 습득 및 발현이 상황 맥락성을 갖는다는 점을 고려할 때, 상황 구성의 원칙을 마련하는 것은 효과적인 교육 전략을 모색하는 데 가장 핵심적인 요소 중 하나이다. 이러한 상황은 '피해와 공정성의 기준이 적용되며 불승인의 직관적 판단이 도출되는 상황'과 '피해와 공정성의 기준이 적용되지 않으며 불승인의 직관적 판단이 도출되는 상황' 그리고 '피해와 공정성의 기준이 적용되며 목적 지향적 판단이 도출되는 상황'의 조합을 통해 구성할 수 있다.

일곱째, 구성된 상황별로 관련 내용을 모두 교육하기에는 현실적 제약이 많기에, 효과적인 교육 전략을 마련해야 한다. 이를 위해서는 (1)당면 상황을 기준으로 학생들의 스키마를 점검하고, (2)그러한 점검에 기반하여 단위 수업에서 중점적으로 교육할 내용을 결정하며, (3)교육을 통해 발달시키고자 하는 인지의 특성에 맞춘 교수 방법을 정해야 한다.

사회적 직관주의 이론의 가장 큰 발견은 도덕적 판단의 기제에 있어 '불승인 직관'의 역할에 대한 발견이라고 할 수 있다. 그간 비교적 덜 중요하게 취급되었던 소극적 차원의 도덕적 판단은 적극적 차원의 도덕적 판단과 상이한 기제로 작동하며, 적극적 차원의 도덕적 판단이 갖는 한계와 위험성을 경계하고 보완하는 데 중요한 역할을 담당한다. 그렇기에 인간 도덕 기능의 두 가지 축 중 하나인 '불승인 직관'이 왜곡되거나 편향되지 않고 바르게 내면화되도록 교육하는 것은 도덕성을 바르게 형성시키는 데 매우 중요하며, 그에 대한 구체적인 교육 방법의 마련이 시급하다. 특히, 앞서 살펴본 바와 같이, 강한 불승인 직관에 기반하는 도덕적 판단은 주장의 근거를 제시하기 어려운 경우에도 여타의 심리과정 및 판단을 압도하는 특성을 갖는다. 그렇기에 학생의 불승인 직관이 편향된 방향으로 형성되지 않도록 지도해야 할 필요성은 더욱 강조되어야 한다.

참 고 문 헌

1. 국내 문헌

1) 단행본

박병기, 추병완. (2007). 윤리학과 도덕교육1. 일산: 인간사랑.

정창우. (2006). 도덕교육의 새로운 해법. 서울: 교육과학사.

정창우. (2013). 도덕과 교육의 이론과 쟁점. 서울: 울력.

추병완. (2000). 열린 도덕과 교육론. 서울: 하우

추병완. (2004). 도덕 교육의 이해. 서울: 백의.

2) 논문

박병춘. (1998). 나딩스의 배려 윤리와 도덕 교육. 도덕윤리과교육, 9, 381-398.

박병춘. (2005). 배려윤리의 초등 도덕과 교육에의 적용방안. 도덕윤리과교육, 20, 241-264.

박장호. (2012). 신경과학의 도덕적 직관 이해와 도덕 교육. 윤리교육연구, 29, 137-164.

석자춘, 정창우. (2017). 나바에츠 통합적 윤리교육모형을 활용한 도덕교육 개선 방향. 도덕윤리과교육, 54, 23-52.

양해성. (2017). 사회적 직관주의 모델과 이중과정 모델의 통합 가능성 연구. 도덕윤리과교육, 55, 123-146.

이정렬. (2009). 도덕적 인격 형성을 위한 도덕적 정체성의 역할과 도덕교육적 함의. 박사학위논문, 서울대학교, 서울.

이정렬. (2017). 도덕적 직관과 도덕교육. 도덕윤리과교육, 56, 109-138.

정창우. (2011). 도덕 심리학 연구의 최근 동향과 도덕교육적 함의-헤이트(J. Haidt)의 뇌과학 연구를 중심으로. 초등도덕교육, 37,

95-130.
정창우. (2012). 나바에츠(D. Narvaez) 도덕발달 이론의 특성과 도덕교육적 함의. 도덕윤리과교육, 37, 329-361.
조현영. (2016). ‘학습을 위한 상황’에서 ‘상황에 대한 학습’으로, 학습 디자인은 어떻게 달라질 수 있는가. 교육문화연구, 22, 99-120.

2. 국외 문헌

1) 단행본 및 번역서

Bruner, J. S. (2009). *The process of education*. Harvard University Press.
Charmaz, K. (2013). 근거이론의 구성: 질적 분석의 실천 지침. (박현선, 이상균, 이채원 역). 서울: 학지사. (원서출판 2006)
Colby, A., & Damon, W. (2010). *Some do care.* Simon and Schuster.
Creswell, J. W. (2015). 질적 연구방법론: 다섯 가지 접근. (조흥식, 정선욱, 김진숙, 권지성 역). 서울: 학지사. (원서출판 2013)
Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain.* Putnam.
Dasgupta, N. (2009). 13 Mechanisms Underlying the Malleability of Implicit Prejudice and Stereotypes. *Handbook of prejudice, stereotyping, and discrimination.* SAGE.
Dreyfus, H. L., Dreyfus, S. E. (1986). *Mind over Machine: the power of human intuition and expertise in the era of the computer.* Free Press.
Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive*

judgment. Cambridge University Press.

Greene, J. D. (2008). The Secret Joke of Kant' s Soul. *Moral Psychology(V. 3): The neuroscience of Morality: Emotion, Brain disorders, and Development,* MIT Press

Greene, J. D. (2010). 넥스트(Next) 천재과학자 18인이 그리는 10년 후 미래. (한세정 역). 파주: 21세기북스. (원서출판 2009)

Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them.* the Penguin Press.

Greene, J. (2017). 옳고 그름. (최호영 역). 서울: 시공사. (원서출판 2014)

Haidt, J. (2014). 바른마음. (왕수민 역). 서울: 웅진지식하우스. (원서출판 2012)

Haidt, J., Björklund, F. (2008). Social intuitionist answer six questions about moral psychology. *Moral Psychology(V. 2): The cognitive science of morality: Intuition and diversity,* MIT Press

Hakoda, Y., Tsuzuki, T., Kawabata, H., Hagiwara, S. (2011). 인지심리학. (강윤봉 역). 서울: 교육을 바꾸는 책. (원서출판 2010)

Hassin, R., Uleman, J., & Bargh, J. (2006). *The New Unconscious,* Oxford University Press.

Hogarth, R. M. (2001). *Educating intuition.* University of Chicago Press.

Kahneman, D. (2012). 생각에 관한 생각. (이진원 역). 파주: 김영사. (원서출판 2011)

Kohlberg, L. (2000). 도덕발달의 철학. (김민남 역). 서울: 교육과학사. (원서출판 1981)

Lapsley, D, K. & Power, F, C. (2008). 도덕심리학과 도덕교육. (정창우 역). 일산: 인간사랑. (원서출판 2005)

Lerner, J. S., & Tetlock, P. E. (2003). 13 Bridging Individual, Interpersonal, and Institutional Approaches to Judgment and Decision Making: The Impact of Accountability on Cognitive Bias. *Emerging perspectives on judgment and decision research,* Cambridge University Press.

Margolis, H. (1987). *Patterns, thinking, and cognition: A theory of judgment.* University of Chicago Press.

Narvaez, D. (2007). How Cognitive and Neurobiological Science Inform Values Education for Creatures like US. *Values Education and Lifelong Learning,* Springer.

Narvaez, D. (2008a). The social intuitionist model: Some counter intuitions. *Moral Psychology(V. 2): The cognitive science of morality: Intuition and diversity*, MIT Press.

Narvaez, D., & Bock, T. (2014). Developing ethical expertise and moral personalities. *Handbook of moral and character education,* Routledge.

Prehn, K., & Heekeren, H. R. (2009). Moral judgment and the brain: a functional approach to the question of emotion and cognition in moral judgment integrating psychology, neuroscience and evolutionary biology. *The moral brain*, Springer.

Rich, J. M., & DeVitis, J. L. (1999). 도덕 발달 이론. (추병완 역). 서울: 백의. (원서출판 1994)

Sternberg, R, J. (2005). 인지심리학. (김민식, 손영숙, 안서원 역). 서울: 박학사. (원서출판 2003)

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. *Social psychology: Handbook of basic principles,* the Guilford Press.

Shweder, R. A., Much, N. C. (1991). *Thinking through cultures: Expeditions in cultural psychology*, Harvard University Press.

Sinnott-Armstrong, W. (2008). Framing Moral Intuitions, *Moral Psychology, Vol. 2: The cognitive science of morality: Intuition and diversity*, MIT Press.

Strauss, A., Corbin, J. (1995). 근거이론의 이해. (김수지, 신경림 역). 서울: 한울. (원서출판 1990)

Turiel, E. (2010). 도덕성 발달 핸드북. (김태훈 역). 경기: 인간사랑. (원서출판 2006)

Vozzola, E. C. (2018). 도덕 발달의 이론과 적용. (정창우 역). 서울: 울력. (원서출판 2014)

2) 논문

Bebeau, M. J., Rest, J. R., & Narvaez, D. (1999). Beyond the promise: *A perspective on research in moral education. Educational Researcher*, 28(4), 18-26.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. *Unintended thought*, 3, 3-51.

Bargh. J. A., Morsella. E. (2008). The Unconscious Mind. *Perspectives on Psychological Science*, 3(1), 73-79.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674), 1167-1170.

Collins, A., Brown, J. S., & Newman, S. E. (1988). Cognitive apprenticeship: Teaching the craft of reading, writing and

mathematics. Thinking: *The Journal of Philosophy for Children,* 8(1), 2-10.

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. *Journal of personality and social psychology,* 104(2), 216-235.

Damasio, A. R., Everitt, B. J., & Bishop, D. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions: Biological Sciences*, 351(1346), 1413-1420.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264(5162), 1102-1105.

Davis. J. L., Rusbult. C. E. (2001). Attitude Alignment in Close Relationships. *Journal of Personality and Social Psychology*, 81, 65-84.

Dimitrov, M., Phipps, M., Zahn, T. P., & Grafman, J. (1999). A thoroughly modern Gage. *Neurocase*, 5(4), 345-354.

Dreyfus, H. L., & Dreyfus, S. E. (1990). What is moral maturity? A phenomenological account of the development of ethical expertise. *Universalism vs. communitarianism,* 237-264.

Edwards. K., Hippel. W. (1995). Hearts and Minds: The Priority of Affective Versus Cognitive Factors in Person Perception. *Personality and Social Psychology Bulletin*, 21, 996-1011.

Graham. J., et al. (2013). Moral Foundation Theory: The

pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55-130.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3), 364-371.

Greene, J. D. (2015). Beyond point-and-shoot morality: Why cognitive (neuro) science matters for ethics. *The Law & Ethics of Human Rights,* 9(2), 141-172.

Greene. J. D., Haidt. J. (2002), How (and where) does moral judgment work?. *Trends in Cognitive Science,* 6(12), 517-523.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107(3), 1144-1154.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105-2108.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, 102(1), 4-27.

Haidt, J. (2001). The Emotional Dog and Its Rational Tail : A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108(4), 814-834.

Haidt. J. (2010). Moral psychology must not be based on faith and hope: Commentary on Narvaez. *Perspectives on*

Psychological Science, 5(2), 182-184.

Haidt, J., Björklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason, Unpublished manuscript, University of Virginia

Jacoby. L. L., Kelley. C., Brown. J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influence of the past. *Journal of Personality and Social Psychology,* 56(3), 326-338.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446(7138), 908-911.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological review*, 96(4), 674.

Lai, C. K., Haidt, J., & Nosek, B. A. (2014). Moral elevation reduces prejudice against gay men. *Cognition & emotion*, 28(5), 781-794.

Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An investigation of moral judgement in frontotemporal dementia. *Cognitive and behavioral neurology*, 18(4), 193-197.

Murphy. S. T., Zajonc. R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of personality and social psychology,* 64(5), 723-739.

Musschenga, A. W. (2009). Moral Intuitions, Moral Expertise and Moral Reasoning. *Journal of Philosophy of Education,* 43(4), 597-613.

Narvaez, D. (2008b). Triune ethics: The neurobiological roots of our multiple moralities. *New Ideas in Psychology,* 26(1), 95-119.

Narvaez. D. (2010). Moral Complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163-181.

Narvaez, D. (2012). Moral neuroeducation from early life through the lifespan. *Neuroethics,* 5(2), 145-157.

Narvaez, D., Bock, T., Endicott, L., & Lies, J. M. (2004). Minnesota's community voices and character education project. *Journal of research in Character Education,* 2(2), 89-112.

Narvaez, D., & Lapsley, D. K. (2005). The psychological foundations of everyday morality and moral expertise. *Character psychology and character education,* 140-165.

Narvaez, D., & Lapsley, D. K. (2008). Teaching for Moral Character: Two strategies for teacher education. *Teacher educator,* pp. 1-27.

Narvaez. D., Vaydich, J. L. (2008). Moral development and bebaviour under the spotlight of the neurobiological science. *Journal of Moral Education*, 37(3), 289-313.

Parkinson. C., Sinnott-Armstrong. W. (2011). Is Morality Unified? Evidence that Distinct Neural System Underlie Moral Judgment of Harm, Dishonesty, and Disgust. *Journal of Cognitive Science*, 23(10), 3162-3180.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in cognitive science,* 2(3), 511-527.

Rest, J. R. (1983). Morality. *Handbook of child psychology*, 3, 556-629.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and social psychology bulletin,* 34(8), 1096-1109.

Shavitt, S. (1990). The role of attitude objects in attitude functions. *Journal of Experimental Social Psychology*, 26(2), 124-148.

Sherman, G. D., & Haidt, J. (2011). Cuteness and disgust: the humanizing and dehumanizing effects of emotion. *Emotion Review,* 3(3), 245-251.

Shweder, R., Haidt, J. (1993). The future of moral psychology: truth, intuition, and the pluralist way. *Psychological Science,* 4(6), 360-366.

Shweder, R. A., Mahapatra, M., Miller, J. G. (1987). Culture and Moral Development. *The emergence of morality in young children,* 1-83.

Thomas, B. C., Croft, K. E., & Tranel, D. (2011). Harming kin to save strangers: further evidence for abnormally utilitarian moral judgments after ventromedial prefrontal damage. *Journal of cognitive neuroscience*, 23(9), 2186-2196.

Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological science,* 17(6), 476-477.

Wegener. D. T., Petty. R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. *Advances in experimental social psychology*, 29, 141-208.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological science*, 16(10), 780-784.

Wilson, B. G., Jonassen, D. H., & Cole, P. (1993). Cognitive approaches to instructional design. *The ASTD handbook of instructional technology,* 4, 21-21.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological review,* 107(1), 101-126.

Abstract

A Study on Critical Interpretation of Social Intuitionism and Its Implication for Moral Education

Yang Hae Sung
Department of Ethics Education
The Graduate School
Seoul National University

Theories of social intuitionism, Dual Process Model of J. Greene and Social Intuitionist Model of J. Haidt, which attempts to clarify the mechanism of moral judgment, are accepted as having validity, on the assumption that each theory encompasses various types of phenomena of moral judgment and provides general and unified explanation of the mechanisms of moral judgment. If both theories are extensively interpreted to encompass a wider category of phenomena than the actual phenomena that they have actually clarified, or if the risks of contradictory interpretations are inherent within the theories, the two theories are difficult to assure validity. In order to guarantee

the validity of the both theories, first, it is necessary to specify the verification category of the moral judgment which is the subject of each theory, in order to be aware of the risk of contradictory interpretations. And second, it is also necessary to examine whether the conceptual category used in the theory can be interpreted differently from the intention of the researcher.

However, discussions about the theory of two social intuitionism have been derived from the conditions that the specification of the verification category and the clarification of the conceptual category are not sufficiently preceded. However, in the absence of such refinement, Haidt's Social Intuitionism Model has been criticized as a biased theory that encompasses only a subset of the category of phenomena of moral judgment, and Green's Dual Process Model has been conceded as a reasonable theory covering a broader category of moral judgment. In addition, without carrying out of clarification of the concept category, deontological judgement was criticized merely emotional intuitive moral judgment and rationalization after it, and the normative status of deontology became suspected. In addition, as the discussions about the two theories were conducted to reaffirm the discovery of cognitive psychology on the phenomenon of moral judgment, the intrinsic discovery of the social intuitionism has not been sufficiently applied to moral education.

This study was conducted to clarify the verification category and the conceptual category of two theories with the concern that the validity of the related debate can not be guaranteed if these clarification is not preceded. And based on this clarification, this study examined the validity of the arguments

derived from the two theories and searched for the possibility of deriving the intrinsic findings of theses theories distinct from cognitive psychology. And on the basis of these works, this study sought for a method to apply the inherent discoveries to Integrative Ethical Education of D. Narvaez, which is a representative moral education theory reflecting the intuitive mechanism of cognition. Details of these works are as follows.

First, based on the clarification of verification categories, the two social intuitionism theories are not conflicts, but rather complementary. In comparing the interview questions presented on the basis of the two theories, the key claim of Social Intuitionism Model was derived on the category of 'situations in which damage and fairness criteria can not be applied' and the main argument of Dual Process Model was derived on the category of 'the situations where the criteria of damage and fairness are applied'. In other words, both theories are each a different subcategory of phenomena that constitute a phenomenon of moral judgment. It is proper to grasp them as mutually complementary relations, and it does not make sense to make a debate on which theory is more valid.

Second, if borrowing conceptual category of intuition and reasoning and using level of emotional activation as the criterion to classify intuition and reasoning, there remains a danger of contradictory interpretation within these two theories. Although Greene classified the judgments in the Non Moral Dilemma on the basis of the degree of emotional activation, cognitive psychological studies consider various criteria to distinguish intuition and reasoning. So it is difficult to clearly classify the judgment in the situation into one kind of judgment among the

intuitive judgment and the reasoning judgment from a non-moral dilemma. In other words, the distinction between intuition and reasoning with only the activation of emotion implies the danger of contradictory interpretation.

In addition, the time required to elicit a positive response in the Moral Impersonal Dilemma situation which Greene classified as a reasoning takes less time than the time required to elicit a negative response in the Moral Personal Dilemma situation which Greene classified as a intuition. These experimental results, which are derived on the assumption of two situations with the same conditions except for the means, collide in front of the assumption that the intuitive judgment has characteristics as a quick judgment and the reasoning judgment has a characteristic as a slow judgment. That shows classification of a certain judgment with only the degree of activation of emotion is a hasty classification that implies the risk of contradictory interpretation. Judgement in Moral Impersonal Dilemma has the characteristics of reasoning judgment that it is less emotional, more conscious, and able to recognize the process of judgment, but also has the characteristic of intuitive judgment that it is quick at the same time. Therefore, it is difficult to classify it as reasoning or intuition.

There is also the danger of this contradictory interpretation in Haidt's Social Intuitionist Model. As mentioned above, Haidt's classification as reasoning can also be classified as intuition according to degree of mastery and automatability, and can also be classified as reasoning. Therefore, it is necessary to supplement or replace the concept category to eliminate the risk of contradictory interpretation and more appropriately interpret

the phenomenon of moral judgment that the two social intuitionism theories have clarified. This complementary or alternative conceptual category can also be derived by clarifying what Haidt and Greene categorize objects really means.

Contrary to D. Narvaez's use of the concept of intuition of excellence and autonomy without any concerns of contents, the category of concept classified by Haidt and Greene as intuition is indispensible from the specific contents as 'disapproval'. So we can define it as 'Disapproval Intuition'. In addition, the cognitive processes that both theories classify as reasoning have both the possibility to be classified as reasoning or intuition. Therefore, the psychological processes that are classified as reasoning by these two theories need to be contentual concept but not formally, which can be referred to as 'Goal Oriented Cognition'.

Third, based on the clarification of verification categories and conceptual categories, the findings of two theories can be summarized as follows. (1) In the context of disapproval intuition, moral judgment is derived from the intuitive judgment of disapproval regardless of whether or not the criteria of damage and fairness can be applied. (2) In situations where the criteria for damage and fairness do not apply, no moral judgment is made if there is no dissapproval intuition. However, in situations where the criteria of damage and fairness are applied, moral judgment can be derived by the goal oriented judgment, even if there is no intuition. (3) Even if intuitive judgment of disapproval is derived, correction by reasoning can be made by object oriented. But such corrections are very rare in situations where the criteria for damage and fairness do not apply, and in situations where the criteria for damage and fairness are applied,

these modifications may be made easier in case strong consequences are presented.

Fourth, when a new cognitive type classification is applied, complementarity between deontological judgment and consequential judgement is established. The deontological judgment is only a judgment judged by the "disapproval intuition" to judge the unacceptability of a specific act. On the other hand, the consequential judgment is a “goal oriented judgment”, which is to select an action more suitable for achieving the objective. In other words, deontological judgment functions to judge whether or not a certain act can not be tolerated beyond the acceptable category independently of the purpose, and the consequential judgment is to select an action suitable for achieving the purpose. In other words, the two types of judgment play a complementary role by playing a different role in moral functioning and Greene's argument that deontological judgment is difficult to have ethical value is hardly valid.

The most important discovery of social intuitionism is the discovering the role of 'disapproval intuition' in the mechanism of moral judgment. The passive dimension of moral judgment, which has been regarded as relatively less important in the past, operates with a moral judgment and a different mechanism of active dimension. The passive dimension of moral judgment by the disapproval intuition plays an important role in guarding the limits and risks of active moral judgment. Therefore, it is very important to educate moral function to be properly internalized without being distorted or biased, which is one of the two axes of human moral function, and it is urgent to establish concrete education methods for it. In particular, moral judgment based on

strong disapproval intuition has the characteristic of overwhelming other psychological processes and judgments even when it is difficult to provide a basis for the judgement. Therefore, the need to instruct students' disapproval intuition not to be formed in the direction of bias need to be more emphasized.

keywords : Social Intuitionist Model, Dual Process Model, Integrative Ethical Education, Disapproval Intuition

Student Number : 2015-30416