
우리의 오류추리는 얼마나 합리적인가?: <웨이슨 선택 과제>와 오류추리의 합리성

박미영 (한신대학교)

I. 서론

사람들은 종종 논리 규칙을 위반하는 판단과 행동을 하는데, 이 현상은 대체로 비합리적인 판단과 행동으로 이해되어 왔다. 1966년에 웨이슨 (Peter. C. Wason)이 고안한 카드 선택 과제(Wason's selection task)는 이러한 오류 판단의 편재성을 드러내는 대표적인 실험이다. 이 실험의 결과는 사람들이 학습 부족이나 의지박약 등의 이유로 자주 오류 판단을 저지른다고 해석되었다. 그러나 이같은 설명은 비합리적 판단과 행동에 대한 가설은 제시할 수 있으나, 그러한 현상이 왜 편재하며 또한 그 편재성의 패턴이 왜 규칙적인지에 대한 만족스러운 설명은 제공하지 못한다. 이에 대한 대안으로 비합리적인 판단과 행동을 오류가 아닌 일종의 추론규칙인 휴리스틱스(heuristics)로 이해하는 이론은 설명력과 예측력 면에서 기존의 해석보다 더 나은 성과를 거두고 있으나, 여전히 비합리성의 제약에 간혀 일관되고 설득력있는 이론을 제시하지 못한다.

논리 규칙을 위반하는 판단과 행동을 오류가 아닌 휴리스틱스로 이해하고 이를 통해 마음의 작동을 보다 정확하게 설명하고 예측하기 위해서는, 기존의 이상적인 합리성에 기반 한 해석이 아니라 맥락-의존적이고 영역특수적인 합리성에 기반 한 해석이 필요하다. 이 논문은 (1) <웨이슨 선택 과제>의 답변은 오류 판단이 아니라 휴리스틱스라는 추론규칙을 따르는 판단이라는 점을 드러내고, (2) 휴리스틱스는 논리 규칙과 다르지만 판단에

대한 규범의 역할을 할 수 있으므로, (3) 〈웨이슨 선택 과제〉에서 작동하는 휴리스틱스에 대한 최선의 해석은 논리 규칙과 확률 규칙만 합리적 규범으로 상정하는 해석이 아니라, 맥락-의존적이고 영역-특수적인 제한된 합리성 규범에 기반 한 해석임을 주장하고자 한다. 이 논의가 옳다면 웨이슨의 실험에서 소위 ‘전형적인 오답’을 제시한 사람들은 비합리적인 오류추리를 한 것이 아니라, 추론자가 처한 상황과 맥락에 알맞은 합리적인 추리를 하였다고 볼 수 있다. 더 나아가 우리가 일상에서 저지르는 오류추리 중 상당수가 제한된 의미의 합리성을 갖춘 휴리스틱스라는 주장은 더 높히 정당화될 것이다.

II. 〈웨이슨 선택 과제〉

전건 긍정(Modus ponens)이나 후건 부정(Modus Tollens) 규칙은 기본적인 연역규칙으로, 우리는 이러한 규칙들을 의도적으로 따르는 판단, 또는 설령 그런 규칙을 모른다고 하더라도 추론의 과정이 실제로 그러한 규칙에 부합하는 방식으로 이루어진 판단을 합리적 판단이라고 배운다. 그렇다면 사람들은 실제로 이 규칙에 부합하는 추론을 하는가? 웨이슨은 이와 관련하여 사람들에게 주어진 조건문에 부합하는 카드를 고르게 하는 실험을 고안한다. 이 실험을 [웨이슨1]이라고 부른다.

[웨이슨1]



실험 참여자들에게 카드 4장과 조건법칙을 함께 제시한다. 각 카드의 펼쳐진 면에는 A, D, 4, 7이 적혀있다. 조건법칙은 ‘카드 앞면이 모음이면, 그 카드 뒷면은 짝수의 수이다’이다. 연구자는 참여자들에게 각 카드의 한 면에는 알파벳이, 다른 한 면에는 숫자가 적혀있다고 말해주고, 제시한 조건법칙의 참/거짓을 판별하기 위하여 반드시 뒤집어 봐야만 하는 카드가 무엇인지 묻는다. (Wason, 1966)

위 실험의 조건법칙인 ‘카드 앞면이 모음이면, 그 카드 뒷면은 짝수의 수이다’라는 문장은 $p \supset q$ 의 형식을 가지고 있다. 따라서 모음인 카드를 살펴보거나 짝수가 아닌 카드를 살펴보아야 한다. 이 조건문의 참/거짓과 관련된 연역규칙은 전건 긍정과 후건 부정이다. 물론 전건 긍정과 후건 부정이라는 말은 논리학을 따로 공부하지 않은 사람들은 모를 수 있는 용어이다. 그러나 전건 긍정과 후건 부정은 올바른 추론 형태이고, 전건 부정과 후건 긍정은 올바르지 않은 추론 형태임은 논리학을 공부하지 않은 사람들도 깨달을 수 있다. 다시 말해 $p \supset q$ 의 형식을 갖는 자연어 문장은 p 가 참이거나 q 가 거짓인 사례가 $p \supset q$ 의 거짓을 보일 수 있는 사례가 된다는 것은 대부분의 사람들이 조금 생각해봄으로써 알 수 있는 것이다. 마찬가지로 p 가 거짓인 경우나 q 가 참인 경우는 $p \supset q$ 의 거짓을 보일 수 없는 사례라는 것도 알 수 있다. 그렇다면 제시된 조건문의 참/거짓을 판별하기 위해서는 ‘A’와 ‘7’을 뒤집어야 한다는 답이 도출된다.

그러나 위 실험에서 참여자들 중 10%만이 ‘A’와 ‘7’을 선택하였고, 90%는 ‘A’와 ‘4’를 선택하였다(Johnson-Laird & Wason, 1970). 대부분이 선택한 ‘4’는 사실상 제시한 규칙 확인에 도움을 주지 못한다. 제시한 규칙은 ‘모든 모음 카드는 뒷면에 짝수가 있다’이지, ‘모든 짝수 카드는 뒷면에 모음이 있다’가 아니기 때문이다. ‘4’의 앞면에 어떠한 알파벳이 있더라도 제시한 규칙을 확인할 수 없다.

웨이슨과 존슨-라이드는 이 결과를 일종의 오류/편향으로 설명하고 이를 확증 편향(confirmation/verification biases)이라고 불렀다. 실험 참여자들은 과제에서 제시한 규칙을 반증할 수 있는 증거를 찾기보다는 오히려 그 규칙을 참으로 만드는 사례만 찾기 때문에 ‘p라면 q’라는 조건문에서 ‘-q’ 보다는 ‘q’를 선택한다는 것이다. 과제를 수행한 대부분의 사람들은 소위 ‘후건 긍정 오류’로 불리는 추론을 하였다고 설명된다(Wason & Johnson-Laird, 1972).

그러나 위 과제의 형식에 구체적인 문제 상황을 등장시키면 정답률이 매우 높아진다고 한다. 1971년 웨이슨과 사피로(Wason & Shapiro)가

[웨이슨1]에 내용(구체적인 도시이름과 이동수단)을 넣어 참여자들에게 과제를 수행하였더니, 기존 선택 과제에서 10%에 불과하던 정답률이 63%로 증가하였다고 한다(Reich & Ruth 396). 이후 [웨이슨1]은 여러 방식으로 조작되어 연구되었는데, 크게 두 종류의 조작으로 구분된다. 하나는 전건과 후건의 내용이 특정한 유형에 속하게 내용을 조작(내용 조작)하는 것이고, 다른 하나는 조건명제에서 후건을 부정문의 형태로 표현하는 조작(후건 조작)이다(Evans 1984, 1989; Oaksford & Stenning 1992)(도경수 78 참고).

[웨이슨1]에 대한 대표적인 내용 조작 과제를 살펴보자. 이 실험을 [웨이슨 2]라고 부르자.

[웨이슨2]

맥주	콜라	19	17
----	----	----	----

당신은 나이트클럽 관리인이다. 클럽 고객은 입장 시 카드를 받는데, 앞면에는 기본 메뉴가, 뒷면에는 고객의 나이가 적혀있다. 4장의 카드(맥주, 콜라, 19, 17)가 있다. 규칙은 '고객이 맥주를 마시려면, 그는 19세 이상이어야 한다.'이다. 이 규칙의 T/F를 판별하기 위해서는 어떤 카드를 뒤집어 보아야 할까? (Griggs & Cox, 1982)

위 실험의 조건법칙인 '고객이 맥주를 마시려면, 그는 19세 이상이어야 한다'라는 문장은 [웨이슨1]과 동일하게 $p \supset q$ 의 형식을 갖고 있다. 따라서 '맥주' 카드를 살펴보거나 '19세'가 아닌 카드를 살펴보아야 한다. 따라서 제시된 조건문의 참/거짓을 판별하기 위해서는 '맥주'와 '17'을 뒤집어야 한다는 답이 도출된다.

놀랍게도 [웨이슨1]과 달리 [웨이슨2]에서 실험 참여자의 74%가 '맥주'와 '17'을 골랐다. 실험 연구자인 그릭스와 콕스는 이를 기억-단서 가설(memory-cuing hypothesis)로 설명한다. 이들은 참여자인 "대학생들이 이 조건 법칙의 내용과 음주 연령을 잘 알고 있으며, 많은 실험 참여자들이 이 규칙을 위반해 본 경험이 있기 때문에, 이 문제의 수행 수준이 향상되었

다”(박권생 71)고 설명한다.

다른 종류의 내용 조작 실험도 이와 유사하다. 참여자가 자기 스스로를 편지를 분류하는 우체국 직원이라고 생각하게 하고 ‘봉투가 봉해져 있으면, 5d 우표가 붙어 있다’는 규칙의 위반 여부를 확인하려면 어떠한 카드를 선택해야하는지에 대한 과제에서 참여자의 81%가 조건문의 진리값을 판별하기 위한 조건에 부합하는 카드를 선택하였다고 한다(이 참여자들은 내용 조작 이전의 [웨이슨1]에서는 15%만이 해당 카드를 선택하였다.)(홍승렬 4).

III. <웨이슨 선택 과제> 대한 해석 및 연관된 실험

그렇다면 [웨이슨1]과 [웨이슨2]에 대하여 어떠한 해석이 가능할까? [웨이슨1]에 대한 가장 일반적인 해석은 웨이슨과 존슨-라이드가 설명하였듯이 사람들이 오류/편향을 저지른다는 것이다. 그러나 이러한 설명은 특정 과제에 대하여 대부분의 사람들이 공통적이며 반복적으로 하게 되는 판단을 실수, 오판, 잘못으로 규정해 버린다. 즉 사람들의 일상적인 추론의 형식을 ‘오류’로 만들어버리는 문제점을 안고 있다. 또한 [웨이슨2]에서 증가한 정답률도 설명하기 어렵다.

이러한 해석의 문제점을 인식하고 <웨이슨 선택 과제>의 결과를 오류가 아닌 것으로 설명하는 다양한 입장들이 있는데, 대표적인 입장들을 다섯 가지로 분류하여 보았다.

(1) 구체적 경험 및 기억-단서설: 먼저 [웨이슨2]를 설계한 그릭스와 콕스는 참여자들이 자신의 기억을 단서로 과제를 수행하였다고 설명한다. 영역-특수적인 경험의 기억을 사용하여 과제를 해낸 것이다(Griggs & Cox, 1982; Rieth & Ruth 1982)(홍승렬 4 참고). 일상생활에서 접하는 많은 정보를 장기 기억에 저장하고 그 정보를 인출하여 추론 과제를 수행해야

하는 어려움은 논리 형식보다는 전제의 내용과 사전 지식간의 관계에 더 의존한다. (Holyoak & Spellman, 1993) 이 입장에 의하면, [웨이슨2]에서 내용의 수준이 매칭과 확증(verification) 비율에는 영향을 끼치지지만, 반증(falsification)에는 영향을 끼치지 않는다고 한다. 즉 웨이슨 과제에 내용을 첨가하여 매칭비가 감소하고 확증비가 증가했다고 하더라도, 반증의 비율은 이와 상관없이 낮다는 것이다(Reich & Ruth 401 참고).

(2) 매칭 편향설: 에반스는 매칭 편향(matching biases) 가설을 제시하였다. 참여자들이 과제에서 제시한 규칙에서 언급된 항목과 판단해야할 사례의 항목 간의 매칭이 없을 경우 그러한 매칭이 없는 진술은 부적절한 것으로 간주하기 때문에, 규칙에 제시된 항목에만 집중하였다는 것이다.(Evans & Ball 435 참고) [웨이슨1]에서 제시된 규칙에는 ‘모음, 짝수’가 언급되어 있으므로, 사람들은 직관적으로 빠르게 언급된 사항과 어울리는 ‘A’와 ‘4’를 선택하게 된다. 그러나 전건 긍정 규칙을 적용하였을 경우 선택했어야 하는 ‘A’와 ‘7’에서 ‘7’은 제시된 규칙에서 언급되지 않았으므로 매칭 역시 이루어지지 못하였다.

에반스는 두 단계 이론(two-stage theory)으로 이를 설명한다. 단계1에서 인지지는 휴리스틱 과정으로 적절한 과제 정보 목록을 선택하고, 단계2에서는 분석적 과정에서는 선택한 정보에서 추론의 형식을 일반화한다. 에반스는 [웨이슨1]에서 논리적으로 틀린 선택이 많은 것은 추론이 단계1 과정만 거쳐서 매칭 편향을 보이기 때문이라고 주장한다(Evans & Ball, 2010).

(3) 쌍조건문 이해설: 사람들이 [웨이슨1]에서 오류를 저지르는 이유는 그들이 전건-후건 관계를 조건문이 아니라 쌍조건문으로 이해하기 때문이라는 해석도 있다. 참여자들은 어떤 규칙은 강하게 논리적인 조건절로, 반면 다른 것들은 쌍조건을 표현한 것으로 간주한다고 한다. 원래의 추상 규칙은 이 점이 매우 애매하고 의무 규칙과는 다르다는 것이다. 즉 대부분의 사람들의 유형적 오류는 그들이 추상 규칙을 쌍조건문으로 고려하기 때문에 만들어진다는 사실에 의해 가장 잘 설명된다(Wagner-Egger, 2007 참고). 소위 논리적 오류나 편향이라는 것은 사실은 규칙에 대한 상이한 해석의 결과이지

잘못된 추론의 결과가 아니라는 것이다(Wagner-Egger 491). 사람들은 [웨이슨1]은 주로 쌍조건문으로 해석하고, 반면 [웨이슨2]와 같은 종류의 과제들은 조건문이나 쌍조건문, 혹은 이 두 형식 모두로 해석한다고 한다 (Wagner-Egger 500).

(4) 실용적 추리-도식설: 쉹과 홀요크(Cheng & Holyoak, 1985)는 사람들은 허가, 의무, 인과관계와 같이 일상적인 경험에서 나온 지식 구조를 사용해 추리를 한다고 제안하면서, 이러한 구조를 실용적 추리 도식(pragmatic reasoning schemas)이라고 하였다.¹ 이 입장에 의하면 사람들은 구문론이나 구체적 경험에 기초하여 추리를 하는 것이 아니라, '전제 조건이 충족되면, 수행을 해도 좋다'는 허가(permissions), '전제 조건이 충족되어야 수행을 할 수 있다'는 의무(obligation), 그리고 인과 관계(causation)와 같은 삶의 경험을 바탕으로 유도한 추상 지식 구조를 사용해서 추리를 한다는 것이다. 그런데 [웨이슨1]은 일상생활과 관련이 없으므로 이러한 추리 도식이 쉽게 활성화되지 못한 것이며, <맥주-나이 선택 과제>와 <봉투-우표 선택 과제>는 일상생활과 관련 있으므로 그러한 과제를 수행하는 맥락에서는 관련 추리 도식이 원활하게 활성화된 것이다.

유사하게 [웨이슨2]의 조건문을 이익을 얻기 위해 비용을 치러야하는 사회교환적인 것으로 이해하게 하게 함으로써 허용-도식이나 사회교환-도식이 활성화되어 그러한 도식이 요구하는 방식에 따라 선택과제가 수행되는 것으로 해석할 수도 있다(Cosmides 1989; Gigerenzer & Hug 1992).

(5) 심적 모델 이론: 참여자들은 오직 조건 규칙에 대한 그들의 심적 모델에서 정확하게 표현되는 카드만을 고려하며, 따라서 그런 규칙의 참고 관련 있는 것들을 선택한다 (Johnson-Laird & Byrne, 2002)(Wagner-Egger 487). 참여자들은 과제의 전제에 드러난 구조와 내용을 표상하는 내적

1. 이에 대하여 추상적 허가 도식 문제는 분명한 부정이 사용과 위반 확인 맥락의 조합에 의해 수행이 촉진된 것이지, 허가 도식에 기초한 결과가 아니라는 비판이 있다.(Jackson & Griggs 1990) 쉹과 홀요크는 이 비판에 대하여 실험을 제시하면서 분명한 부정이 없이도 연역 추리에서 촉진 효과가 있었다고 답변한다.(Cheng & Holyoak, 1993)

모델을 구성하며, 이 모델에 분명하게 표상된 카드만을 고른다는 것이다.

한편, [웨이슨1]과 [웨이슨2]의 결과의 보편성은 단지 상술한 외국의 사례에만 국한되는 것은 아니다. <웨이슨 선택 과제>는 다양하게 변형된 형태로 계속 진행되고 있는데, 참여자의 성별, 나이, 국가와 상관없이 대체로 유사한 결과를 얻고 있다. 우리나라의 경우도 다르지 않다. 다음은 한국에서 실행된 유사한 종류의 실험 과제이다.

1. 카드문제: 네 장의 카드를 보고 철수가 '이 카드의 한쪽 면에 B가 인쇄되어 있으면, 다른 면에는 3이 인쇄되어 있어'라고 말했다. 철수의 말이 어겨졌는지 알아보려면 반드시 뒤집어 보아야 하는 카드는 어느 것인가?

2. 회원명부문제: 민지가 동아리실 회원카드를 보고 '남자 카드 윗면에는 파랑 꽃이 인쇄되어 있어'라고 말했다. 민지의 말이 어겨졌는지 알아보려면 반드시 뒤집어 보아야 하는 카드는 어느 것인가?

3-1. 주말근무문제 (근로자 관점): 주5일 근무하는 회사에서 주말에 근로자가 근무하는 것은 사장에게 득이 되지만, 근로자에게는 불편한 일이다. 이전에 한 번도 주말근무를 하지 않았던 근로자 A가 이번 주말에 근무하는 대신 다음 주중에 하루를 쉬는 것이 좋을 것 같아 주말 근무를 할 것인지를 따져보고 있다. 그런데 '주말 근무를 하면 주중에 4일을 근무해야한다'는 규정이 지켜지지 않은 적이 있다는 소문이 있다. 이에 근로자 A는 근로자 네 명의 근무기록을 조사하기로 했다. 개인별 작업기록부 한 면에는 주중 근무상황이, 다른 면에는 주말 근무상황이 기재되어 있다. 근무 규정이 어겨진 적이 있는지 알아보려면 반드시 뒤집어 보아야 하는 카드는 어느 것인가?

3-2. 주말근무문제 (사장 관점): 주5일 근무하는 회사에서 주말에 근로자가 근무하는 것은 사장에게 득이 되지만, 근로자에게는 불편한 일이다. 그래서 주말 근무를 하면 대신 다음 주중에 4일만 근무하는 체도를 만들었다. 그런데 '주말 근무를 하면 주중에 4일을 근무해야한다'는

규정이 지켜지지 않은 적이 있다는 소문이 있다. 이에 사장은 근로자 네 명의 근무기록을 조사하기로 했다. 개인별 작업기록부 한 면에는 주중 근무상황이, 다른 면에는 주말 근무상황이 기재되어 있다. 근무 규정이 어겨진 적이 있는지 알아보려면 반드시 뒤집어 보아야 하는 카드는 어느 것인가?

4. 보조금문제: 공해를 줄이기 위해 정부는 중유 보일러를 도시가스 보일러로 개량하는 가정에 보조금을 준다. 그런데 ‘지원금을 받았으면 도시가스 보일러를 사용해야한다.’는 규정이 지켜지지 않은 적이 있다는 소문이 있다. 그래서 환경청 직원이 네 가구의 주택대상을 보았는데, 다른 서류에 가려 두 개의 주택대상에서는 개량 공사여부만, 다른 두 주택대상에서는 보조금 지급 여부만 보였다. 지원금 규칙이 어겨진 적이 있는지 알아보려면 반드시 뒤집어 보아야 하는 주택대상은 어느 것인가?

5. 청소문제: 한 아이가 어머니가 직장 일을 하는 동안 방청소를 지시 받았다. 어머니는 ‘내 방을 정리하면, 나가 놀아도 된다’라는 규칙을 말했다. 그리고 가사도우미에게 며칠간의 아이의 행적을 기록하여 한 면에는 방청소 여부를, 다른 면에는 나가 놀았는지 여부를 적게 했다. 이 규칙이 어겨진 적이 있는지 알아보려면 반드시 뒤집어 보아야 하는 기록카드는 어느 것인가?

6. 연금문제: 회사에 사원이 오래 근무하는 것은 회사에는 득이 된다. 반면 사원은 회사가 마음에 안들 경우 고역일 수 있다. 반대로 장기 근무 퇴직자에게 주는 연금은 회사로서는 비용이지만 사원에게는 이득이다. 그런데 ‘연금을 받으려면 20년 이상 근무해야한다’는 규정이 지켜지지 않은 적이 있다는 소문이 있다. 그래서 노조의 한 간부가 최근 퇴직한 사원 네 명의 기록을 조사하기로 하였다. 개인 기록부의 한 면에는 연금 수령 여부가, 다른 한 면에는 근무연수가 기재되어 있다. 이 연금 규정이 어겨진 적이 있는지 알아보려면 반드시 뒤집어 보아야

하는 주택대장은 어느 것인가?(도경수 92-93)

위 연구는 질문3-1을 포함하여 질문1~6으로 구성된 과제를 제시하는 <실험1>과, 질문3-2를 포함하여 질문1~6으로 구성된 과제를 제시하는 <실험2>를 사용하였다. 실험자는 “<실험1>에서는 후건을 부정문으로 표현하면 후건 부정을 선택하는 비율이 증가하는 것과, 카드의 선택비율과 탐색 순서 간의 상관성이 높은 것을 관찰하였다. <실험2>에서는 후건의 표현양식과 조건문의 내용이 후건 부정을 고르게 하는 데 독립적으로 영향을 미칠 수 있음을 시사하는 결과를 얻었다. 아울러 판단자의 관점에 따라 선택의 양상이 달라질 수도 있음을 관찰하였다”라고 결론을 내린다(도경수 77).

이 실험은 [웨이슨2] 결과의 원인이 선행 연구들에서 제시한 가설들 중 하나가 될 수 있음을 보여준다. 즉 과제의 형식, 내용 등이 모두 과제 수행에 영향을 끼치며, 실험 참여자의 관점도 역시 과제 수행에 영향을 미친다는 것이다.

<웨이슨 선택 과제> 결과의 보편성은 실험 참여자의 학력과 직업에 있어서도 큰 차이를 보이지 않는다. 쉹 등의 연구에 의하면, 대학에서 논리학을 수강한 학생들이 형식 논리 훈련을 전혀 받지 않은 학생들보다 [웨이슨1]에서 단지 3% 높게 논리적으로 맞는 카드를 선택하였다고 한다 (Cheng, Holyoak, Nisbett, & Oliver, 1986). 또한 박사학위를 가진 과학자 집단에서도 정반응의 참여자가 전체의 10%도 미치지 못하였다고 한다 (Jackson & Griggs, 1988)(박권생 61).

다음은 국내의 현직 판사, 사법연수원생, 대학생을 대상으로 한 내용-조작의 카드 선택 과제이다.

귀하는 재판을 진행 중에 있습니다. 현재 다루고 있는 사건은 조직폭력배 사건으로 재판을 종결해야 할 시한이 다가오고 있습니다. 그런데, 최종적으로 판결을 내리는 데에는 다음의 사실을 확인하는 것이 아주

중요합니다.

조직폭력배 집단인 도끼파는 모두 태백 출신이다.

도끼파와 갈고리파의 집단 폭력현장에서 4명이 현행법으로 체포되어 구속되었습니다. 이들은 다음과 같습니다.

- (1) 도끼파
- (2) 갈고리파
- (3) 태백 출신
- (4) 춘천 출신

사실의 진위를 판정하기 위해 구속된 네 명 중 누구(들)를 증인으로 소환하시겠습니까? 시간적 제약 때문에 가능하면 적은 수의 증인을 소환해야 합니다.(김청택, 최인철 2000)

제시한 과제의 진위를 알기 위해서는 (1)과 (4)를 확인하여야 한다. 실험 참여자 세 집단이 (1)을 선택한 비율은 유사하였다. 그러나 (4)를 확인하는 비율에서는 통계적으로 유의한 차이가 났다. 판사는 21%, 연수생은 43%, 대학생은 39%가 (4)을 선택하였다. 법관은 연수생과 대학생보다 반증 사례를 찾으려는 시도를 적게 한 것이다. 즉 법관들은 도리어 대학생보다 높은 편향을 보였다. 이 실험과 함께 행한 [웨이슨1]류의 실험에서는 판사의 69.4%, 연수생의 71.1%, 대학생의 48.5%가 논리 규칙에 맞지 않는 카드를 선택하였다. 또한 [웨이슨1]를 재판 시나리오에 맞추어 반증질문(후견 부정)을 하는지를 살펴보았는데, 판사의 78.6%, 연수생의 86.8%, 대학생의 57.5%가 논리 규칙에 맞는 질문을 선택하였다(김청택, 최인철 331-332).

참여자들 중 판사와 연수생 부류는 대학생 부류나 그 외 일반인 부류에 비하여 논리 규칙을 상대적으로 잘 알고 있다고 가정되지만, [웨이슨1]에서는

(일반적인 참여자들보다는 약간 낮은 수치이지만) 여전히 높은 편향률을 보인다. “흥미로운 것은 법학공부를 시작하는 사람들이 일반인보다 확증편향적 사고를 하는 성향이 더 높은 것으로 나왔다”는 점이다.(김청택, 최인철 333) 또한 [웨이슨1]과 [웨이슨2]의 결과와 유사하게 과제에 내용이 첨가될 경우 오답률이 줄었고, 그러한 내용이 참여자의 직업과 직접 관련 있는 내용일 경우 오답률은 더 많이 줄어든다고 한다.

청 등의 실험과 김청택 등의 실험의 참여자들은 일반적인 참여자들에 비해 비교적 논리 규칙을 잘 알고 있다. 그러나 전건 긍정과 후건 부정 규칙에 대한 이들의 지식은 과제 수행에 잘 활용되지 못하였다.

대학생들을 대상으로 한 다른 실험에서도 [웨이슨1]에서는 정반응률이 5% 이하로 나타났고, [웨이슨2]에서는 53%가 넘는 매우 큰 정반응률이 나타났다고 한다. 흥미롭게도 미리 전건 긍정 연습문제를 풀게 하고 답안을 알려준 점은 참여자들은 정반응률 12%로 그렇지 않은 참여자들의 정반응률 자체의 차이는 있었으나, 오반응률에는 차이가 없었다고 한다(박권생 77). 실험자는 이러한 결과를 토대로 웨이슨 과제 수행의 결과는 인종과 문화를 초월하여 나타나는 현상이라고 해석한다(박권생 80).

이상의 실험 결과와 해석들은 추론 구조 및 과정에 관한 이론에 어떠한 시사점을 주는가? 위 실험들과 그에 대한 해석은 다음의 세 진술을 함의하는 것으로 볼 수 있다.

첫째, 가장 기본적인 보편적인 연역추론 규칙으로 알려져 있는 ‘전건 긍정의 위반 및 후건 부정 오류’가 추론자의 상황에 상관없이 보편적으로 나타나고 있다는 점이다. 1966년 이후 다양한 방식으로 변형된 외국의 실험뿐 아니라, 국내의 여러 직업군을 대상으로 했던 실험에서도 이 점이 관찰된다.

둘째, 편향이 유의미하게 사라지는 결정적인 요인에 대한 여러 해석이 있는데, 이 해석들은 공통적으로 내용 조작이 편향 감소와 매우 깊은 상관관계를 보이고 있음에 주목한다. 특히 사회계약 및 의무와 관련된 내용은

대체적으로 편향을 감소시킨다고 분석한다. 즉 동일한 형식의 과제라고 하더라도 그것의 기술 방식에 따라 이를 수행하는 사람들의 추론 과정이 달라지는 것으로 보인다.

셋째, 편향 감소의 일차적 요인은 내용 조작이겠으나, 다른 여러 요소들도 편향 감소에 영향을 끼치는 것으로 보인다. 즉 내용과 형식, 판단자의 상황이라는 변수들이 상호작용을 하면서 총체적으로 과제 수행에 영향을 끼치고 있다고 볼 수 있다. 예컨대 〈웨이슨 선택 과제〉를 주말근무 문제로 변형시킨 실험에서는 동일한 문제를 근로자의 관점과 시장의 관점으로 기술하였을 때 각 관점에 대한 평가와 답변이 상이하였으며(도경수, 2000), 〈웨이슨 선택 과제〉를 폭력사건의 증인 소환 문제로 변형시킨 실험에서는 참여자들이 직업군별(현직 판사, 사법연수생, 대학생)로 규칙적으로 상이한 답변을 하였다고 한다(김정택, 최인철, 2010).

이러한 점에 비추어 볼 때, 인간의 추론과정을 이해하기 위해서는 판단의 형식뿐 아니라 내용, 즉 문맥이 중요하다고 주장할 수 있다. 추상적 기호로 주어진 과제의 문맥에서는 연역규칙의 도식에 따라 추론을 하는 것이 쉽지 않다. 그러나 구체적 내용, 특히 그 중에서도 의무 불이행 등의 사회계약과 관련된 내용으로 주어진 과제의 문맥에서는 상당수가 외견상 연역규칙의 도식에 따르는 추론과 동일한 결과를 산출하는 추론을 한다.

그렇다면 〈웨이슨 선택 과제〉의 결과를 단지 ‘사람들은 자주 오류추리를 한다’고 해석하는 것보다는, 그 오류추리의 전형성과 보편성을 특정 휴리스틱스로 담아내어 ‘사람들은 특정한 과제에 대하여 특정한 휴리스틱스 규범을 따른다’고 해석하는 것이 더욱 적절하지 않을까? 본 논문은 〈웨이슨 선택 과제〉에 대한 후자의 해석을 지지하고자 하며, 이를 위하여 먼저 휴리스틱스가 무엇인지 설명하고, 이어서 어떻게 휴리스틱스가 추론의 규범이 될 수 있는지를 논의하여 보겠다.

IV. 휴리스틱스는 오류가 아니라 추론 규칙이다?

마음과 세계를 서로 독립된 존재로 상정한 데카르트의 심신이론은 마음을 몸이 속한 이 세계의 특성과는 별개의 원리에 의해 통제되는 방식으로 규정한다. 이에 기반을 둔 이성 개념은 여타의 동물과는 질적으로 다른 인간의 뛰어난 사유체계에 대한 확신과 이 체계의 작동 과정 전반에 대하여 통제와 조절이 가능한 능력의 소유, 논리적 타당성, 무모순성, 그리고 사유과정의 결과로서 타당하게 도출되는 행위의 규범적 완결성 등으로 정의되어 왔다. 이러한 전통에서 합리적이라는 것은 보편적이며 일관적인 규칙을 항상 따르는 것으로 이해되었다. 현대의 명제논리학은 보편적 규범으로 여겨지는 추론의 형식을 제시하는데, 이러한 논리 규칙에 속하는 대표적인 연역 규칙과 확률 규칙은 다음과 같다.

전건 긍정 (Modus ponens): $p \supset q, p \therefore q$

후건 부정 (Modus tollens): $p \supset q, \neg q \therefore \neg p$

선언지 제거 (Modus tollendo ponens): $p \vee q, \neg p \therefore q$

가언삼단논법 (Hypothetical Syllogism): $p \supset q, q \supset r \therefore p \supset r$

귀류법 (Reductio ad absurdum): $p \supset (q \cdot \neg q) \therefore \neg p$

선언지 첨가 (Addition): $p \therefore p \vee q$

연연화 (Adjunction): $p, q \therefore p \cdot q$

단순화 (Simplification): $p \cdot q \therefore p$

결합법칙 (associative law): $P(A \cap B) \leq P(A)$

조건부확률 (conditional probability): $P(B|A) = P(A \cap B) / P(A)^2$

-
2. 연역규칙과 확률 규칙은 형식이 다르지만, 후자는 좌항과 우항의 관계가 변항의 내용에 상관없이 보편적 규범으로 여겨진다는 점에서 추론 규칙에 확률 규칙도 추가하였다. 이에 대한 지적을 해주신 익명의 심사위원께 감사드립니다.

위 규칙들은 전통적으로 그것의 내용과 상관없이 형식적으로 언제나 타당한 것으로 간주되어 왔다. 대표적으로 전건 긍정(Modus ponens, 이하 MP)은 예외 없는 보편적 규범으로, p와 q의 내용이 무엇이 되든지 상관없이 언제나 참인 결과를 도출하는 추론 형식이다. 그러나 추론자가 과제를 수행하고자 할 때, 그 과제의 내용과 관련된 사회적 맥락에 따라 MP를 사용하기도 하지만, 다른 경우에는 MP를 위반하는 추론을 사용하기도 하여, 오히려 MP의 위반이 더 적절한 판단이나 행위가 되기도 한다. 예를 들어, '나에게 날개가 있다면, 나는 즐거울 것이다. 나는 지금 즐겁다. 따라서 나는 날개가 있다'라는 추리는 형식적으로는 후건 긍정의 오류이며 내용상으로 거짓이 도출되는 추리이다. 그런데 어떤 형사가 용의자를 취조하면서 '당신이 어제 그 사람을 죽이지 않았다면 당신은 나흘 전에 발생한 사건의 범인이 아니다. 그런데 당신은 어제 그 사람을 죽였다. 따라서 당신은 나흘 전에 발생한 사건의 범인이다'라고 추리하는 상황을 가정해 보자. 이 추리는 MP를 위반하여 전제의 참이 결론의 참을 보장하지 못함에도 불구하고 실제로 유용하게 자주 사용되는 추론의 형식이다. 즉 이러한 종류의 추리는 MP를 위반하고 형식적 오류를 저지르는 것으로 보이지만, 그것의 추리가 발생한 문제 상황과 p와 q의 내용에 따라 유용하고 실제로 참을 도출하는 추론의 형식이 될 수 있다.

한편, 사람들은 자주 위의 규칙이 아닌 다른 종류의 규칙에 따라 판단과 행동을 하는데, 그 규칙이 '모든 경우를 고려하지 않고 나름대로의 편리한 기준에 따라 그 중 일부만을 고려하여 문제를 해결하는 신속한 방법'인 경우에 이를 휴리스틱스(heuristics) 또는 휴리스틱(heuristic)이라고 부른다. 여러 휴리스틱스 중에서도, 정보 처리 부담을 줄여주면서 동시에 적절하고 효과적인 결론을 이끌어낼 수 있는 휴리스틱스는 다음의 10가지로 제시될 수 있다.

- (1) 인지 휴리스틱 (Recognition heuristic): 두 대안 중 하나가 인지된다면, 인지된 그것이 더 높은 가치를 가진다고 추론하라.

- (2) 거침없음 휴리스틱 (Fluency heuristic): 두 대안 모두 인지되는데, 하나가 더 빠르게 인지된다면, 그것이 더 높은 가치를 가진다고 추론하라.
- (3) 최선-취하기 (Take-the-best): 두 대안 중 어느 것이 더 가치 있는지 추론하기 위해서 (a) 타당한 단서를 통해서 탐색하고, (b) 단서가 구별될 때 탐색을 멈추고, (c) 이 단서를 지지하는 대안을 선택하라.
- (4) 단순계산하기 (Tallying): 기준 평가를 위해서, 비중을 평가하지 말고, 단순하게 긍정적 단서의 수를 고려하라.
- (5) 만족화 (Satisficing): 대안들을 통해 탐지하고, 당신의 기대 수준을 능가하는 첫 번째 대안을 선택하라.
- (6) 동등성 휴리스틱 (1/N; equality heuristic): 자원을 동등하게 1/N로 분배하라.
- (7) 불이행 휴리스틱 (Default heuristic): 만약 불이행이 존재한다면, 아무것도 행하지 말라.
- (8) 맞받아치기 (Tit-for-tat): 처음에는 협력하고, 다음에는 당신의 파트너의 마지막 행동을 모방하라.
- (9) 다수 모방하기 (Imitate the majority): 당신이 속한 그룹의 다수를 고려하고, 그들의 행동을 모방하라.
- (10) 성공한 것 모방하기 (Imitate the successful): 가장 성공적인 사람을 고려하고, 그의 행동을 모방하라.

(Gigerenzer & Brighton 130-131)

위 목록에서 보듯이 논리 규칙과 휴리스틱스는 매우 다른 형식을 가진다. 그리고 휴리스틱스 분류는 논리 규칙의 분류만큼 명확하지도 않다. 그러나 그러한 형식적 문제가 휴리스틱스를 논리 규칙과 다른 진술, 즉 규범이 아닌 것으로 만드는 것은 아니다. 물론 하나의 휴리스틱이 기술되는 방식은 시대와 문화마다 다를 수 있으나, 축적된 다양한 연구를 통해 특정 휴리스틱에 대한 대표적인 정의와 설명이 제시되어 왔고 앞의 목록은 휴리스틱스 연구자들에게 승인된 것들이다. 이렇게 승인된 휴리스틱스의 내용은 사실판

단이 아닌 가치판단의 형식으로 제시된다. 따라서 휴리스틱스는 규범의 역할을 할 수 있고, 우리는 그것이 어떠한 규범이 될 수 있는지 물을 수 있다.

그렇다면 휴리스틱스는 합리성의 규범이 될 수 있는가? ‘원칙적으로 논리 규칙과 확률 규칙을 언제나 준수하는 판단과 행위’만을 합리적인 것으로 보는 입장(무제한적 합리성 unbounded rationality, 이하 unBR)에서는 휴리스틱스가 합리성이 규범이 될 수 없다. 그러나 ‘원칙적으로 논리 규칙, 확률 규칙, 그리고 휴리스틱스를 과제의 문맥에 의존적이며 영역 제한적인 방식으로 준수하는 판단과 행위’를 합리적인 것으로 보는 입장(제한된 합리성 bounded rationality, 이하 BR)에서는 휴리스틱스가 합리성의 규범들 중 하나가 될 수 있다.³

휴리스틱스는 제한된 의미의 합리성의 규범이 될 수 있으며, 따라서 휴리스틱스를 따르는 추론 규칙은 합리적 판단과 행동을 이끈다. 그러나 휴리스틱스 선택과 적용이 합리적이기 위해서는 적절한 문맥에서 영역 제한적인 규범으로 제시되어야 한다는 것이 중요하다. 어떠한 상황에서도 무조건 특정 휴리스틱을 따르는 것이 합리적이라고 주장하는 것은, 어떠한 상황에서도 무조건 특정 논리 규칙을 따르는 것이 합리적이라고 주장하는 것만큼이나 틀렸다. 휴리스틱스를 항상 따르면 당연히 문제가 생긴다. 예를 들어 긍정적 단서의 수만 고려하는 ‘(4) 단순계산하기 휴리스틱’은

3. 무제한적 합리성 개념은 플라톤에서 시작하여 데카르트, 데이빗슨, 로반 등에 이르기까지 이상적인 이성 개념을 상징하는 많은 철학자들이 지지한다. 그러나 이러한 이성과 합리성 개념은 꾸준한 비판을 받고 있다. 특히 1970년대 이후 추론, 판단, 결정에 대한 합리적 규범과 인간의 실제적 판단 사이에 많은 불일치가 있다는 관찰이 보고되고 다양한 심리학적 연구들이 인간의 판단과 행동이 비합리적이라는 것을 지속적으로 보여주면서, 이에 기반하여 이상적인 이성 개념과 함께 무제한적 합리성도 비판을 받는다. (Nisbet & Ross 1980; K&T 1982; Evans 1983, 1989; Simon 1983; Cherniak 1986; de Sousa 1987; Stich 1990, 1994; Baron 1994; Gigerenzer & Goldstein 1996; Gigerenzer 2000, 2008) 그 중 일부는 합리성 개념 자체를 거부하고(Hume, Nietzsche, Freud 등), 일부는 합리성을 제한된 것으로 이해한다(Simon 1976, 2000; Cherniak 1986; Gigerenzer 2000, 2001 등). 이에 관한 보다 자세한 연구는 필자의 다른 논문에서 진행 중이다.

어떠한 특정 맥락에서는 적절한 추론 도구이지만, 로르샤흐 검사(Rorschach test)⁴에서는 매우 적은 수의 단서를 임의로 긍정적 단서로 ‘인정하는 방식’을 통해 부적절한 추론 도구의 역할을 하며, 따라서 실제의 환경과 부합하지 않는 진짜 오류를 산출한다. 이처럼 논리 규칙을 위반하였기 때문에 오류로 불렸지만 사실은 과제 맥락에서 가장 적절한 추론 기제인 휴리스틱스(가짜-오류)와 해결해야 할 문제의 맥락을 부적절하게 이해한 잘못된 추론 방법(진짜-오류)을 명확히 구분하는 것이 중요하다.

인간의 판단과 선택은 이론적 측면과 실천적 측면 모두에서 수많은 휴리스틱스로 이루어진다. 편견이 전혀 없다는 것은 가능하지도 않을뿐더러, 실제 인간의 마음이 그것을 항상 욕구하지도 않는다(Gigerenzer & Henry 136 참고). 어떤 편견은 매우 쓸모 있고 그 맥락에 따른 적절한 결과를 산출한다. 즉 연역규칙이나 확률 규칙만을 언제나 일관적으로 따르는 것은 어떤 경우 매우 쓸모없고, 따라서 그 맥락에서 부적절한 결과를 산출하게 된다. 행위자가 합리적이기 위하여 어떠한 규칙을 어떠한 방식으로 따라야하는지의 문제는, 그 규칙의 내적 완결성과는 별도로 그것의 맥락과 실제로 적절하게 잘 어울리는지 여부를 고려해야 한다. 유사하게 휴리스틱스 역시 그 자체로는 좋은 추론기제인지의 여부를 알 수 없고 이는 오직 환경과의 적절한 상보적 통합을 통해서만 평가 가능하다. 인지과정은 당면 과제의 외적 요소들과 그 과제를 풀어야하는 행위자가 속한 환경적 요소에 의존하여 작동하기 때문이다.

이러한 이유로 휴리스틱스가 논리 규칙에 비하여 규범이 요구하는 형식적

4. Rorschach test: 특정 질문에 대한 특정 반응이 정신적으로 문제가 있는 사람의 특성을 드러내준다고 알려진 투사 검사(projective test)의 일종. 좌우 대칭인 다양한 잉크 반점을 보고 이것이 무엇으로 보이는지에 대한 답을 통해 그 사람이 동성애자인지, 편집증 환자인지, 자살할 가능성이 있는지 등을 밝혀낼 수 있다고 주장하는 검사이다. 예를 들어, 어떤 잉크 반점에 대해 영당이, 여성의 의복, 성기를 언급하는 답변이나, 성별이 모호하다거나 혹은 성별이 혼재되어 있다는 답변은 모두 다 동성애를 의심할만한 지표라고 한다. 이 검사의 무용성을 보여주는 연구에도 불구하고, 심리학자들은 수십 년 간 이 검사를 사용해왔고, 많은 이들이 여전히 이것을 사용하고 있다고 한다.

엄밀성과 보편성의 측면에서 덜 완벽할 수는 있으나, 합리성이 요구하는 최선의 이유 제시와 믿음-욕구에 기반 한 목표 달성의 측면에서는 오히려 더 나을 수 있다. 이처럼 휴리스틱스를 추론의 규범으로 보는 관점은 <웨이슨 선택 과제>의 결과를 오류로 보는 기존의 해석이 아닌, 휴리스틱스라는 추론 규칙에 따른 것으로 보는 새로운 해석을 가능하게 해준다. 이 해석들을 비교하기 위하여, 먼저 <웨이슨 선택 과제>에 대한 기존의 해석, 즉 ‘많은 사람들은 과제 풀이에서 오류를 범하였다’는 해석(unBR에 근거한 해석)이 왜 높이 정당화될 수 없는지 논하고, 이어서 새로운 해석인 ‘많은 사람들은 과제 풀이에서 특정 휴리스틱을 따르는 추리를 하였다’는 해석(BR에 근거한 해석)이 왜 더 적절하고 높이 정당화되는지를 논하겠다.⁵

V. <웨이슨 선택 과제>에 대한 논리 규칙-규범 해석 및 이에 대한 비판

<웨이슨 선택 과제>를 오류 추론 혹은 오류 추론의 하나인 휴리스틱스로 설명하는 해석이 왜 문제인가? [웨이슨1]와 [웨이슨2] 모두에서 unBR이 작동한다고 가정한다면, [웨이슨1]의 결과는 설명되지만, [웨이슨2]의 결과, 즉 [웨이슨1]에서 확증 카드를 선택한 참여자들이 [웨이슨2]에서는 반증 카드를 선택한 결과를 제대로 설명할 수 없다. 두 설명 모두에서 unBR은 MP 또는 후건 부정과 같은 연역규칙이 판단의 보편 규범임을 전제하고 있는데, 이러한 전제는 실험의 결과와 충돌한다.

unBR에 의하면, [웨이슨1]은 MP, 정확히 말하자면 ‘MP가 제시하는 조건문의 참/거짓 판별법’(이하, MP식 판별법)을 위반하고 확증 오류를 저지르는 사례이다. 즉 비합리적 추론의 사례가 된다. 이런 현상의 원인은

5. 이 새로운 해석은 기거렌저(Gigerenzer 2009; 2010)의 해석에 기반하여 발전시킨 것으로, 기거렌저식 해석의 취약점을 보완하고 철학적 근거를 추가하여 더 견고한 해석으로 재구성한 것이다.

참여자들의 추론력과 학습력의 차이, 관련 지식의 유무 등으로 인해 올바른 규범인 MP에 대한 의식적 이해(MP식 판별법을 인지하고 이를 의도적으로 과제 풀이에 적용한 경우) 또는 MP에 대한 무의식적 이해(MP식 판별법을 인지하지는 못하였지만 추론 과정이 그러한 판별법을 드러내거나 그 판별법에 부합하는 방식으로 과제 풀이가 진행된 경우)가 결합되어 있는 것이라고 주장할 수 있다. 이러한 설명은 어느 정도 설득력이 있다.

그러나 이 설명 방식은 [웨이슨2]에 가서는 설득력을 잃게 된다. unBR을 전제하면서 [웨이슨2]에 대해 제시할 수 있는 가능한 해석은 두 가지이다. 첫 번째 해석은 (1) '사람들은 [웨이슨1]에서는 의식하였든 혹은 의식하지 못하였든 MP식 판별법을 위반하고, [웨이슨2]에서는 의식하였든 혹은 의식하지 못하였든 MP식 판별법을 이행한다'는 것이다. 이러한 unBR식 해석은 [웨이슨2]에서 MP식 판별법이 사용되었다면, 왜 [웨이슨1]에서는 MP식 판별법이 사용되지 않았는지도 함께 설명해야 한다. 두 과제의 차이는 내용뿐이며, 실험 참여자들의 조건과 과제 형식은 동일하다. 따라서 unBR은 MP식 판별법의 사용 유무를 과제 내용과 관련 있다고 설명해야 하는데, 그러한 설명은 논리 규칙의 적용이 맥락에 의존하지 않는다고 가정하는 unBR의 입장과 충돌한다. 그러므로 unBR 지지자는 이 해석을 수용할 수 없다.

가능한 또 다른 unBR식 해석은 (2) '두 과제 모두에서 사람들이 의미론적으로 과제를 해석하여 추론하였다'는 것이다. 즉 웨이슨 실험 결과는 사람들이 MP식 판별법에 따르는 '논리적 형식에 근거한' 추론의 사례가 아니라, 과제의 구체적인 내용에 따르는 '의미론적' 추론의 사례라고 주장하는 것이다. 그런데 상술하였듯이 웨이슨 실험 결과는 상당히 규칙적이고 일반적이므로, 이러한 unBR식 해석은 MP식 판별법이 실제로 사람들이 잘 사용하지 않는 규범이라는 점을 스스로 인정하는 셈이다. 해석(2)는 MP식 판별법이 추론의 규범이 되어야 하는 사례에서조차도 그러한 규범을 따르기 보다는, 의미론적 추론이 꾸준히 수행되고 있음을 보여준다. 이 해석은 보편적 규범으로 가정했던 MP식 판별법에 대한 부정적 견해를 함축하게 되고

결국 MP식 판별법을 거부하는 가능성도 인정하므로, unBR의 관점과 양립불가능하게 된다. 그러므로 unBR 지지자는 이 해석 역시 수용할 수 없다.

이처럼 unBR식 해석이 웨이슨 사례에 대한 좋은 해석이 될 수 없는 이유는 모든 과제에서 공통적으로 MP가 보편적 규범으로 제시되어야 한다는 전제 때문이다. 이러한 전제는 합리적인 판단과 행위라는 것이 MP와 같은 영역 일반적이며 맥락-초월적인 규범을 선택하고 따르는 판단과 행위라는 주장에서 파생된다. 따라서 unBR식 해석의 부적절성은 합리성의 그러한 주장을 거부하고, 새로운 종류의 합리성 규범을 제시함으로써 극복될 수 있을 것이다.

VI. 〈웨이슨 선택 과제〉에 대한 휴리스틱스 - 규범 해석

이어서 〈웨이슨 선택 과제〉에 대한 BR식 해석을 살펴보자. 이 해석에 의하면 [웨이슨1]은 추상적 기호로 표기된 과제로, 어떠한 사회적 맥락도 들어오지 않은 사례이다. 그러한 사례는 과제 형식의 내적 구조상 MP식 판별법이 적절한 규범이지만, 실제로 사람들은 이러한 규범을 이미 알고 있다고 해도 거의 인식하지 못한다. 과제는 단지 기호로만 표기되고 내용상 사회적 맥락이 결여되어 있으므로, 과제를 부여받은 이는 가장 적절한 추론을 가능하게 하는 단서를 충분히 제공받지 못한다. 이 과제에서 적절한 추론 규범을 찾게 해주는 단서는 거의 없다. 주어진 단서만으로는 ‘확증 휴리스틱’이 가장 그럴듯해 보인다. 따라서 대부분의 인지자들은 ‘확증 휴리스틱’을 사용한다. 또는 과제를 쌍조건문을 따르는 사례로 읽어내기도 한다. 이런 상황에서 MP를 이미 지식으로 가지고 있는 사람들은 MP식 판별법을 부적절한 규범으로 인지하였다기보다는, 선택 가능한 규범들 중 하나로도 거의 인지하지 못하였다. MP를 지식으로 가지고 있지 않는 사람들은 평소 다른 과제에서는 MP식 판별법에 부합하는 방식으로 추론하는 것이 가능하였다고 하더라도, 이 맥락에서는 그러한 판별법에 부합하는

추론을 진행하지 못하였다. 사람들은 [웨이슨1]의 기술 방식으로 인하여 부적절한 규범인 ‘확증 휴리스틱’을 선택하는 판단을 저지르게 된다.

반면 [웨이슨2]는 그것의 내용에 비추어 보았을 때 ‘인지 휴리스틱’이나 ‘최선 취하기 휴리스틱’이 적절한 규범이 되고, 오히려 ‘확증 휴리스틱’은 부적절한 규범으로 분명하게 인지된다. 이 과제는 의무의 이행/불이행의 내용이 들어간 맥락인데, 이 과제가 제시하는 적절한 규범에 대한 단서는 매우 명확하므로 대부분의 인지자들은 그 단서를 사회적 맥락에서 해석하여 최선의 규범을 선택하게 된다. 이 맥락에서 ‘확증 휴리스틱’을 사용하면 ‘19’를 선택하게 되는데, ‘19’는 ‘19세 미만’이면서 맥주를 마시는 의무 불이행자’를 골라내는데 어떠한 역할도 하지 못한다. 반면 ‘인지 휴리스틱’이나 ‘최선 취하기 휴리스틱’은 규칙을 어기는 사람을 골라내는데 반드시 필요한 추론규칙이다. [웨이슨2]가 제시한 규칙은 의무 불이행자를 선별하는 기준에 대한 것인데, 그 기준에 부합하는 의무 불이행자를 골라내기 위해서는 맥주를 마시는 자 중 19세 미만이 있는지 살펴보고, 동시에 19세 미만이면 맥주를 마시는 자가 있는지 살펴봐야 한다. 따라서 ‘맥주’와 ‘19’보다는 ‘맥주’와 ‘17’이라는 카드가 의무 불이행자를 판별하는 과제에 직면하여 즉각적으로 인지되는 선택지이다. 이러한 추론 과정에서 ‘콜라’와 ‘19’는 선택지의 후보로서 쉽게 제거될 수 있다.

이때 ‘인지 휴리스틱’은 전건 긍정과 외견상 동일한 형식을 가지고 있으므로, ‘인지 휴리스틱’이나 ‘최선 취하기 휴리스틱’을 따르는 판단과 MP식 판별법을 따르는 판단은 동일한 결과를 산출하게 된다. 그렇다면 [웨이슨2]에서 대부분의 사람들은 MP를 합리적 규범으로 인식하여 판단하거나 혹은 인식하지는 못하였어도 MP식 판별법과 정합적인 방식으로 판단하였는가? 상술한 실험 결과와 이에 대한 해석들에 의하면 그렇지 않다. 사람들은 ‘확증 휴리스틱’이라는 부적절한 규범을 제외한 다른 적절한 휴리스틱스 중 하나 혹은 그 이상을 사용하여 판단하였다.⁶ 그러한 휴리스틱스의 내적

6. 실험 참여자들이 정확히 어떤 종류의 휴리스틱을 사용하는지는 각 실험 내용의 세팅에

구조가 MP와 동일한 논리적 형식을 가진 것이다. 따라서 [웨이슨2]에서 적절한 규범이 될 수 있는 휴리스틱스의 논리적 형식이 MP와 동일하다고 해서 이를 MP로 환원할 수는 없다. 인지자가 실제의 인지과정에서 자신이 따른다고 인지한 규범은 특정 휴리스틱이지 MP가 아니기 때문이다. 이를 MP로 환원하는 것은 이행된 규범에 대한 일종의 사후 조작이 될 것이다.

이처럼 BR식 해석이 웨이슨 사례에 대한 좋은 해석이 되는 이유는 모든 과제에서 각 과제의 맥락에 따라 그에 적절한 다양한 규칙들이 영역-특수적 규범으로 제시되어야 한다는 전제 때문이다. 이러한 전제는 합리적인 판단과 행위라는 것이 논리 규칙, 확률 규칙, 휴리스틱스 등의 규범들 중 하나를 선택하고 따르는 판단과 행위라는 주장에서 파생된다. 이들 규범들은 영역-일반적이며 맥락-초월적 규범이 아니라, 영역-특수적이며 맥락-의존적인 규범으로 간주된다.⁷

한편 unBR을 수용하더라도 특정 상황에 따라 합리성은 제한적으로 발휘될 수 있으며 따라서 <웨이슨 선택 과제> 과제 역시 unBR을 가정하여 설명할 수도 있다고 하는 반론이 가능하다. 그러나 그 경우 '무제한적 합리성의 제한적 발휘'는 '제한된 합리성의 발휘'와는 다르다. 전자는 '원칙적

따라 상이하다. 다양한 휴리스틱스 중 어느 것을 선택할지의 여부는 참여자 개인의 성향 및 위치(처한 상황)와 무관하지 않기 때문이다. 물론 이 과제에서 사람들이 대체로 사용하는 효과적인 휴리스틱스의 목록을 제시할 수 있다.(예컨대, '인지 휴리스틱', '최선 취하기 휴리스틱', '고정 및 조정 휴리스틱', '후광효과 휴리스틱' 등) 그러나 어떠한 휴리스틱을 사용하였던지 공통적으로 그것이 효과적인 휴리스틱이라면 이는 MP의 형식으로 재구성 이 가능한 논리적 구조를 갖고 있으며, 그러한 이유로 그러한 휴리스틱을 사용하면 MP를 사용하였을 때와 동일한 결과를 산출할 수 있었다. 이는 명백하게 MP를 위반하고 있는 '확증 휴리스틱'과는 대조적이다.

7. 오류 추리의 발생을 특정 맥락에서 설명하는 것과 그 추리의 정당화 문제는 별개라는 심사의견이 있었다. 본 논문은 추리 형식 그 자체의 정당화 여부가 아니라, 추리가 그것이 작동하는 상황에서 얼마나 합리적으로 실행되는지 여부를 문제 삼고 있다. 후자에서 일부 오류추리는 비합리적인 추리이지만, 다른 오류추리는 합리적일 수 있다. 따라서 이 지적은 본 논문의 논지에 대한 직접적이고 심각한 비판으로 보이지는 않는다. 중요한 지적을 해주신 익명의 심사위원께 감사드린다.

으로 무제한적인 합리성이 특정 이유로 인하여 그 능력을 발휘하지 못하는 경우'이며, 후자는 '원칙적으로 제한된 합리성이 그것의 능력을 적절하게 발휘하는 경우'이다. 선언지첨가법을 생각해보자. unBR을 수용하면 선언지첨가법은 맥락에 상관없이 언제나 따라야할 규범이 된다. 그러나 실생활에서 선언지첨가법은 무의미하거나 오히려 인지자의 의도에 반하는 결과를 산출하는 추론이 되기도 한다. 즉 선언지첨가법은 '언제나 참'이 아니라, '맥락에 따라 참이거나 거짓이거나 또는 무의미한' 규범이다.

또한 BR을 수용한다는 것이 비합리성과 진짜 오류까지도 모두 수용해야한다는 주장을 함의하는 것은 아니다. BR이 제시하는 규범을 이행하지 않는 추론의 결과는 명백한 비합리성이나 진짜 오류를 야기한다. BR을 지지하는 관점은 그러한 비합리성과 진짜 오류의 존재를 부정하지 않는다. BR이 합리적인 것으로 규정한 것은 실제 맥락에서도 합리적인 것을 의미하지만, unBR이 비합리적인 것으로 규정한 것에는 실제 맥락에서 합리적인 것이 포함된다. 예컨대 베이지안 규칙을 위반하는 여러 사례는 실제로 과제 수행에서 인지자의 의도에 부합하는 유용한 결과를 산출하는 경우가 매우 많다. <린다 직업 맞추기>⁸, <암 발병률 예측하기>⁹ 등의 과제에서 실험 참여자 대다수는 기저율(base rate)을 무시하였으나 관련 내용의 사회적 의미를 비중 있는 정보로 간주하여 과제를 수행하였다고 한다. 그러한

8. ●과제: 린다는 31세의 독신이며, 뛰어나고 매우 총명하다. 대학에서는 철학을 전공했다. 학생 때 린다는 차별과 사회정의에 매우 관심이 많았고 반핵데모에도 참여했다. 그렇다면 다음 중 어느 쪽이 린다일 확률이 높을까? (1) 린다는 은행원이다. (2) 린다는 은행원 이면서 페미니즘 운동에 적극적으로 참여한다.

●결과: 실험 참여자들 중 85%가 (2)를 선택하였다. - Kahneman & Tversky, 299(1983)

9. ●과제: 어떤 사람이 병원에서 특정 질병에 걸렸는지 알아보기 위한 검사를 하였다. 이 질병은 발병률이 0.001%이며, 이 질병을 찾기 위한 검사는 5%의 오류가능성이 있다. 그 사람이 이 검사에 양성 반응을 보였고, 그의 증상 등에 대해 아무 것도 모를 경우, 실제로 병에 걸렸을 확률은 어느 정도일까? (참여자: 하버드 의과대학의 학생과 스태프 60명)

●결과: 약 절반이 0.95(죄빈값)라고 답했고, 평균은 0.56이었다. 18%만이 정답으로 간주되는 0.02라고 답하였다.(Kahneman & Tversky, 1882, 154)

추론에 의한 결과는 실험실이라는 맥락에서는 소위 오류 추론으로 해석할 수 있지만, 실제 의사 결정 상황에서는 유의미하고 유용하게 사용될 가능성이 높은 추론이라고 해석할 수 있다.

VII. 결론

이상의 논의를 통하여 본 논문은 다음을 주장하였다. (1) 휴리스틱스는 전통적인 논리규범과는 다르지만 적절한 문맥에서 영역 제한적인 규범으로 제시될 때, 충분히 합리적인 규칙이 될 수 있다. (2) 전통적인 논리규범에 기반 한 〈웨이슨 선택 과제〉 해석은 두 종류의 웨이슨 과제인 [웨이슨1]과 [웨이슨2] 결과 중 특히 중요한 [웨이슨2]의 결과를 제대로 설명하지 못한다. 이 해석은 MP와 같은 논리 규칙은 보편적으로 따라야하는 합리성의 규범으로 이해하기 때문에, MP식 판별법의 보편성과 상충하는 [웨이슨2]의 결과를 제대로 설명하지 못한다. 반면, 휴리스틱스-규범에 기반 한 〈웨이슨 선택 과제〉 해석은 [웨이슨1]과 [웨이슨2] 결과 모두를 잘 설명한다. 사람들이 과제에서 수행하였던 ‘확증 휴리스틱’, ‘인지 휴리스틱’ 등은 합리성의 규범들 중 하나이며, 휴리스틱스는 보편적으로 따라야하는 규범이 아니라 과제 영역에 따라 선택적으로 따를 수 있는 규범이기 때문에, 이를 통해 [웨이슨2]의 결과가 잘 설명된다. (3) 그렇다면 〈웨이슨 선택 과제〉에서 작동하는 휴리스틱스에 대한 최선의 해석은 논리 규칙과 확률 규칙만을 합리적 규범으로 상정하는 해석이 아니라, 논리 규칙과 확률 규칙, 그리고 휴리스틱스도 합리적 규범에 포함시키는 해석일 것이다.

〈웨이슨 선택 과제〉는 일반인에게는 물론이고 관련 연구자들에게도 오류추리의 편향을 보여주는 대표적인 실험으로 인식되어 왔다. 그러나 사람들의 오류추리가 단지 논리규범을 어기는 추리가 아니라, 휴리스틱스라는 규범을 따르는 추리일 수도 있다는 가설과 이를 지지하는 다수의 심리실험이 이어지고 있다. 이 새로운 해석은 우리 인간이 자주 틀리고 오류를

저지르는 결점투성이가 아니라, 다양한 추론의 규범들 중 맥락에 맞는 규범을 잘 선별하여 사용하는 현명한 추론자라는 자부심을 부여한다. 아울러 철학사에서 오랜 논쟁의 대상이던 합리성 이론에 ‘전통적인 의미에서는) 비합리적이지만 사실은 (제한된 의미에서는) 합리적인 추론과 행동’이라는 현상에 대한 새롭고 유의미한 시각을 제시할 수 있을 것이다.

참고문헌

- 김청택·최인철(2010), 「법정의사결정에서의 판사들의 인지편향」, 『서울대학교법학』 51(4), 317-345쪽.
- 도경수(2000), 「후건의 표현양식과 조건문의 내용이 선택과제의 수행에 미치는 영향」, 『한국심리학회지: 인지 및 생물』 12(1), 77-94쪽.
- 박권생(1992), 「조건추리의 과제내용과 수행수준」, 『사고개발연구』 2(2) 61-86쪽.
- 홍승렬(1993), 『카드선택문제에서의 실용적 추리도식의 적용』, 이화여자대학교 대학원 교육심리학과 석사논문.
- Cheng, P. W. & Holyoak, K. J. (1985). "Pragmatic reasoning schemas", *Cognitive Psychology* 17(4), 391-416.
- Evans, J. St. B. & Ball, L. J. (2010). "Do people reason on the Wason selection task? A new look at the data of Ball et al.(2003)", *The Quarterly Journal of Experimental Psychology* 63(3), 434-441.
- Gigerenzer, G. (2010). "Moral Satisficing: Rethinking Moral Behavior as Bounded Rationality". *Topics in Cognitive Science* 2(3), 528-554.
- Gigerenzer, G. & Brighton, H. (2009). "Homo Heuristicus: Why Biased Minds Make Better Inferences". *Topics in Cognitive Science* 1(1), 107-150.
- Griggs, R. A. & Cox, J. R. (1982). "The elusive thematic-materials effect in Wason's selection task". *British Journal of Psychology*, 73, 407-420.
- Johnson-Laird, P. N. & Wason, P. C. (1970). "A theoretical analysis of insight into a reasoning task". *Cognitive Psychology*, 1, 134-148.
- Kahneman, D. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University press.
- _____ (1983). "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgement". *Psychological Review* 90(4), 293-315.
- Reich, S. S. & Ruth, P. (1982). "Wason's selection task: Verification, falsification and matching". *British Journal of Psychology* 73, 395-405.
- Wagner-Egger, P. (2007). "Conditional reasoning and the Wason selection task: Biconditional interpretation instead of reasoning bias". *Thinking & Reasoning* 13(4), 484-505.
- Wason, P. C. (1966). "Reasoning", In Foss, B. M. (ed.) *New Horizons*

in Psychology I. Harmondsworth, UK: Penguin press.
Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of reasoning:
Structure and content*. Cambridge: Harvard University press.

국문초록

사람들은 종종 논리 규칙을 위반하는 판단과 행동을 하는데, 이러한 현상은 대체로 비합리적인 판단과 행동으로 이해되어 왔다. 1966년에 웨이슨이 고안한 카드 선택 과제는 이러한 오류 판단의 편재성을 드러내는 대표적인 실험이다. 이 실험의 결과는 사람들이 학습 부족이나 의지박약 등의 이유로 자주 오류 판단을 저지른다고 해석되었다. 그러나 이같은 설명은 비합리적 판단과 행동에 대한 가설은 제시할 수 있으나, 그러한 현상이 왜 편재하며 또한 그 편재성의 패턴이 왜 규칙적인지에 대한 만족스러운 설명을 제공하지 못한다. 이에 대한 대안으로 비합리적인 판단과 행동을 오류가 아닌 일종의 추론규칙인 휴리스틱스(heuristics)로 이해하는 이론은 설명력과 예측력 면에서 기존의 해석보다 더 나은 성과를 거두고 있으나, 여전히 비합리성의 제약에 갇혀 일관되고 설득력있는 이론을 제시하지 못한다. 논리 규칙을 위반하는 판단과 행동을 오류가 아닌 휴리스틱스로 이해하고 이를 통해 마음의 작동을 보다 정확하게 설명하고 예측하기 위해서는 기존의 이상적인 합리성에 기반 한 해석보다는, 맥락-의존적이고 영역-특수적인 합리성에 기반 한 해석이 필요하다. 이 논문은 (1) 웨이슨의 선택 과제의 답변은 오류 판단이 아니라 휴리스틱스라는 추론규칙을 따르는 판단이라는 점을 드러내고, (2) 휴리스틱스는 논리 규칙과는 다르지만 일종의 판단의 규범의 역할을 할 수 있으므로, (3) 웨이슨 선택 과제에서 작동하는 휴리스틱스에 대한 최선의 해석은 논리 규칙과 확률 규칙만을 합리적 규범으로 상정하는 해석이 아니라, 맥락-의존적이고 영역-특수적인 제한된 합리성 규범에 기반 한 해석임을 주장하고자 한다. 이 논의가 옳다면 웨이슨의 실험에서 소위 '전형적인 오답'을 제시한 사람들은 비합리적인 오류추리를 한 것이 아니라, 추론자가 처한 상황과 맥락에 알맞은 합리적인 추리를 하였다고 볼 수 있다. 더 나아가 우리가 일상에서 저지르는 오류추리 중 상당수가 제한된 의미의 합리성을 갖춘 휴리스틱스라는 주장의 정당성은 더 높아질 것이다.

키워드: 웨이슨 선택 과제, 오류추리, 휴리스틱스, 인지심리학, 오류추리의 합리성, 제한된 합리성

Abstract

How Rational are Our Fallacies?: Philosophical Examination on Wason's Selection Task

Miyoung Park (Hanshin University)

People often make judgements and actions in violation of the logical rules, and those judgements and actions have been generally regarded as irrational. The Wason's selection task designed by Wason, P. C. in 1966 is the representative psychological experiment revealing the ubiquity of these fallacy-judgements. The result of the experiment was interpreted as 'People often make an error by the lack of learning or will.' This explanation offers a hypothesis for irrational judgements and actions, but does not offer any satisfying explanation about why the phenomenon is widespread and why it shows some regular pattern. An alternative is a theory that "irrational" judgements and actions are heuristics as an inference rule, not a fallacy. This theory provides more compelling explanation and predictability than the existing interpretation. Unfortunately, however, it still cannot offer any consistent and convincing argument due to its limitations with respect to irrationality. Understanding judgements and actions violating logical rules not as fallacy but as heuristics and accurately explaining and predicting how the mind works requires an explanation based on context-dependent, domain-specific rationality rather than one based on ideal rationality. This paper will show the followings: (1) Answers to the Wason's selection task are not fallacy but judgements following heuristics as an inference rule. (2) Heuristics, different to logical rules, can function as norms of judgements. (3) Therefore, the best interpretation of heuristics working on the Wason's selection task is not an interpretation postulating only logical rules and probabilistic rules as rational norms, but an interpretation based on context-dependent, domain-specific

rationality norms. If the above arguments are proper, we can say that people who make a typical wrong answer to the Wason's selection task do not commit an irrational fallacy but make a rational inference adequate to situations and contexts that they are in. Furthermore, it might be highly justified that a number of common fallacies in a daily life are heuristics based on bounded rationality.

Keywords: Wason's selection task, Fallacy, Heuristics, Cognitive Psychology, Rationality of Fallacy, Bounded Rationality

Received : 20 January 2019
Reviewed: 17 February 2019
Accepted : 20 February 2019