

인과 구조 작동 방식이 인과 추론의 정확성에 미치는 효과*

박 주 화[†]
연세대학교

도 승 이
성균관대학교

개인과 기업 그리고 국가 정책의 성공은 판단과 예측의 정확성에 달려있다고 해도 과언이 아니다. 하지만 사람들은 판단과 예측을 할 때 주어진 정보 외에 추가적인 정보를 탐색하지 않으며, 심지어 주어진 정보조차 무시한다. 판단과 예측의 정확성을 향상시키기 위해서는 이러한 정보 무시 경향성의 인지적 과정에 대한 이해가 필수적이다. 본 연구는 고등 인지 과정의 새로운 규범적 모델로 떠오르고 있는 인과 베이즈 모형(causal Bayes nets)의 틀에서 정보 무시 현상의 인지적 과정을 연구하였다. 실험 1에서는 인과 베이즈 모형이 적용되기 위한 세 가지 경계 조건인 인과 구조와 파라미터의 충분성 및 인과 구조 작동 방식의 명확성을 만족시키는 상황을 구성한 후 인과 추론을 실시하였다. 선행 연구와는 다르게 실험 1에서는 정보 무시 현상이 관찰되지 않았다. 실험 2는 인과 구조가 작동하는 방식을 변경하였다. 실험 2는 인과 구조가 자발적으로 작동한 실험 1과는 달리 인과 구조가 실험 참여자의 개입에 의해서만 작동하였다. 실험 2에서는 정보 무시 현상이 관찰되었다. 이는 선행 연구에서 보고된 정보 무시 현상이 인과 구조 또는 관련 변인들이 작동하는 방식에 대한 인출이 되지 않아서 발생했을 가능성을 시사한다. 본 연구는 또한 인과 관계 표상에 대한 인과 베이즈 모형의 가정이 실제 사람들의 표상과 다를 가능성이 있음을 시사한다. 본 연구는 인과 구조가 작동하는 방식에 대한 인지적 이해를 바탕으로 판단과 예측의 정확성을 향상시키는 후속 연구가 필요함을 보여 주었다.

주제어 : 정보 무시 현상, 인과 베이즈 모형, 인과 구조 작동 방식, 예측 추론

* 이 논문은 2010년도 정부재원(교육부 인문사회연구역량강화사업비)으로 한국연구재단의 지원을 받아 연구되었음(NRF-2010-32A-H00011).

[†] 교신저자 : 박주화, 연세대 IT 정책전략연구소, (120-740) 서울시 서대문구 연세로 50 새천년관 110호
E-mail : juhwa.park@gmail.com

우리는 불확실한 미래의 일들에 대해 다양한 판단과 예측을 하며 살아간다. 이러한 판단과 예측은 개인적인 문제(예: 레몬차가 생강차보다 감기 예방에 효과적일까?)에서부터 사회/국가적 차원(예: 하반기 경제 성장률은 상승할 것인가?)까지 다양한 수준과 범위에서 우리 생활 전반에 걸쳐 영향을 미치고 있다. 특히 판단과 예측의 실패는 개인적 수준 이상의 경제적 손실뿐만 아니라 인명 사고 등의 치명적인 결과로 이어질 수 있다. 예를 들어 최근 개통한 7개 고속도로는 교통량 예측의 실패로 현재 5657억원의 적자를 기록하고 있을 뿐만 아니라 영업을 할수록 적자가 늘어나고 있다(경향신문, 2014). 또한 2009년 이탈리아 국립 재난 위원회는 대형 지진을 예측하는데 실패해 309명이 생명을 잃었다. 재난 위원회 소속 과학자들은 당시 소규모 진동이 있었음을 보고받았지만 이를 무시하여 지진 경보를 발령하지 않았다(한겨레신문, 2014). 개인과 기업 그리고 국가 정책의 성공은 판단과 예측의 정확성에 달려있다고 해도 과언이 아니며 예측의 정확성을 향상시키기 위한 다양한 연구들이 진행되고 있다. 하지만 판단과 예측의 정확성을 재고하는 규범적/기술적 제안들의 효과는 제한적이다.

판단과 예측의 부정확성은 정보의 부족보다는 정보 활용에 있어 주의(attention)를 받지 못하는 정보는 무시하는 근시안적(myopic) 인지 경향성에 의해 발생할 가능성이 높다(Fernbach, Dalows, & Sloman, 2010, 2011a, 2011b). 사람들은 주어진 정보 외에 추가적인 정보를 탐색하려 하지 않는다(Fischhoff, Slovic, & Lichtenstein, 1977; Koriat, Lichtenstein, & Fischhoff, 1980;

Oskamp, 1965; Shaklee & Fischhoff, 1982). 심지어 이탈리아 국립 재난 위원회 소속 과학자들처럼 주어진 정보조차 무시하는 경향성을 보인다(Kanheman, Slovic, & Tversky, 1982; Hardman, 2009). 따라서 정확한 판단과 예측을 위해서는 근시안적 인지 경향성에 관한 이해가 필수적이다.

주어진 정보조차 무시하는 근시안적 인지 경향성(이하 정보 무시 현상)에 대한 선행 연구는 정보 무시 현상의 인지적 과정 보다는 정보 무시 현상의 견고성(robustness)에 초점을 맞추고 있다(Dawes, 1979; Gilovich & Griffin, 2002; Gigerenzer & Goldstein, 1996; Gigerenzer & Gaissmaier, 2011; Hattori & Oaksford 2007; Klayman & Ha, 1987; Kahneman, et al., 1982; Skowrinski & Goldston 1989; Wason, 1960). 선행 연구들은 사람들이 다양한 정보를 고려하기 보다는 결과를 가장 잘 예측해 줄 수 있는 하나 또는 두 개의 정보만을 이용하여 판단과 예측을 수행한다는 것을 다양한 과제상황에서 보여주고 있다.

정보 무시 현상의 견고성에 대한 대표적 설명은 정보 무시 현상은 인지 과정의 오류로 인해 생기는 편향이라는 것이다(Baron, 2000; Hastie & Dawes, 2010). 반면 Gigerenzer와 그의 동료들은 정보 무시 현상이 인간의 인지적 용량과 계산 능력을 고려하면 최선의 방책이라고 제안하다. 즉 제한된 합리성(bounded rationality) 관점에서는 적응적, 더 나아가 규범적 현상으로 간주한다(Gigerengzer & Goldstein, 1996; Gigerenzer Hoffrage, & Kleinbolting, 1991). 정보 무시 현상을 바라보는 관점은 다르지만 정보 무시 현상이 자발적으로 극복될 수 없는

현상으로 본다는 점에서 두 접근은 공통점을 가지고 있다.

하지만 선행 연구에서 보고된 정보 무시 현상의 견고성은 선행 연구가 의존하고 있는 규범적/기술적 모형의 한계 때문에 발생했을 가능성이 있다. 일반적으로 규범적/기술적 모형은 사람들이 판단과 예측에 사용할 것으로 예상되는 정보(input), 계산과정(computation), 그리고 예상 결과(expected output)를 정의한다(Anderson, 1990; Chater & Oaksford, 1999; Krynski & Tenenbaum, 2007; 이하 경계 조건). 정보 무시 현상에 대한 연구는 정의된 정보, 계산과정이 판단자가 실제 사용하고 있는 인지 과정이라고 가정한다.

그러나 판단과 예측의 인지 과정에 대한 규범적/기술적 모형의 가정과 인간의 실제 인지 과정 사이에 괴리가 있을 수 있다. 더 큰 문제는 선행 연구들이 가정하는 경계 조건들을 경험적으로 검증하기 어렵다는 것이다. 예를 들어 판단과 예측에 대한 선형/통계적 모형은 사람들은 예측과 관련된 모든 관련 정보를 탐색하며, 정보의 유용성을 다중회귀 방식으로 평가한 뒤, 선별된 정보를 선형적으로 조합하여 최종 결론에 이른다고 제안한다. 이러한 가정이 충족될 수 없다는 것은 자명한 사실이다. 따라서 인간의 실제 인지 과정에 기반을 둔 새로운 모형의 관점에서 판단과 예측에 대한 연구가 필요하다.

본 연구는 이러한 필요성을 바탕으로 인간 인지 과정의 새로운 규범적/기술적 모델¹⁾로

각광 받고 있는 인과 베이지 모형(causal Bayes nets)의 틀에서 판단과 예측의 인지 과정을 연구하고자 한다. 인과 베이지 모형의 핵심적 주장은 불확실성을 나타내는 확률적 관계는 선형성이 아닌 인과 구조와 인과 강도에 의해 표상되며, 인과 구조와 인과 강도의 제약을 바탕으로 한 베이지 추론을 통해 판단과 예측의 불확실성을 결정할 수 있다는 것이다.

본 연구의 가장 큰 목표는 사람들이 과연 주어진 정보 이외의 다른 정보를 사용하여 판단과 예측을 할 수 있는지 살펴보는 것이다. 이를 위해 인과 추론에 대한 규범적/기술적 모형인 인과 베이지 모형의 경계 조건을 살펴보고, 경계 조건들이 만족되는 상황에서도 정보 무시 현상이 관찰되는지 검증하는 것이다. 구체적으로 인과 구조의 충분성(sufficiency), 인과 구조가 작동하는 방식(functional form)의 명확성, 그리고 인과 강도(causal strength)를 포함한 파라미터(parameter)의 충분성이 만족되는 상황을 구성하여 인과 추론을 살펴보았다(실험 1). 또한 본 연구는 왜 주어진 정보 이외의 정보를 사용하지 못하는지에 대한 실마리를 탐색하였다. 특히 인과 구조가 작동하는 방식이 인과 추론에서 관찰되었던 정보 무시 현상에 미치는 영향을 살펴보았다(실험 2). 본 연구를 통해 밝혀진 판단과 예측의 인지적 과정을 바탕으로 판단과 예측의 정확성을 향상시키기 위한 책략들이 제안될 수 있을 것이다.

근 기술적 모형으로서도 타당함을 주장하는 연구들이 발표되고 있다(자세한 논의는 Krynski & Tenenbaum(2007) 그리고 Sloman & Lagnado (2015)를 참고).

1) 일반적으로 판단과 의사 결정 분야에서 규범적 모형과 기술적 모형은 구분된다. 인과 베이지 모형은 규범적 모형에서 출발하였지만 최

인과 베이즈 모형과 예측/진단 추론

인과 관계로 인간의 인지 과정을 설명하는 인과 베이즈 모형은 최근 인간의 학습(learning), 추론(reasoning), 지각(perception), 언어 이해(language comprehension) 등 고등 인지 과정의 새로운 패러다임으로 급부상하고 있다(Sloman & Lagnado, 2015). 특히 정보 무시 현상과 관련이 있는 기저율 무시(base rate neglect) 현상과 차단 법칙(screening-off rule) 위반 현상을 포함한 판단과 예측에서 보이는 다양한 인지적 편향을 성공적으로 설명하고 있다(Krynski & Tenenbaum, 2007; Park & Sloman, 2013, 2014).

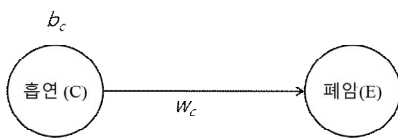
인과 베이즈 모형은 모든 사건들의 의존 관계(dependency 또는 correlation)는 사건들의 인과 관계에 의해 발생한다는 직관에 기반하여, 사건들 사이의 확률/의존적 관계를 인과 구조 그래프로 표상하는 접근이다(Pearl, 2000; Spirtes, Glymour, & Scheines, 2000; Sloman & Lagnado, 2015). 예를 들어 흡연과 폐암은 정적 상관관계가 있다. 이러한 의존 관계는 흡연이 폐암을 일으키기 때문에 발생하는 것이지 폐

암이 흡연의 원인이기 때문에 발생하는 것은 아니라는 것이다. 인과 구조 그래프는 방향적 비순환 그래프(directed acyclic graph)로, 사건들과 인과 관계를 각각 마디와 두 마디들간의 직접적 인과 관계를 표현하는 방향적 화살표로 구성된다. 각 화살표는 인과 강도(Cheng, 1997)에 해당하는 파라미터를 가진다.

인과 베이즈 모형의 틀에서 모든 예측은 두 종류의 추론, (1) 원인이 존재할 때 결과가 발생할 가능성에 대한 추론(predictive inference, 이하, 예측 추론)과 (2) 결과가 존재할 때 원인이 발생했을 가능성에 대한 추론(diagnostic inference, 이하, 진단 추론)으로 수렴된다(Holyoak & Cheng, 2011). 서론에서 언급한 교통량 예측, 지진 예측은 예측 추론, 증상을 통해 질병의 원인을 추론하는 의료 판단은 진단 추론으로 분류할 수 있다.

인과 베이즈 모형은 사람들이 인과 추론을 할 때 (1) 관련 사건들을 탐색하고, (2) 탐색된 사건들의 인과 구조, 작동 방식(functional form) 그리고 인과 강도(causal strength)를 결정하며 (3) 마지막으로 구조적 제약 및 인과 강도에 따라 최종 결론을 내린다고 가정한다(이하 인

A. 단일 모형



B. 완전 모형

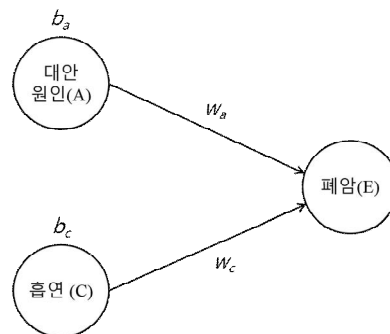


그림 1. 흡연과 폐암에 대한 인과 추론에 사용될 수 있는 두 인과 모형

과 추론의 3단계). 그림 1은 흡연과 폐암에 대한 예측 추론과 진단 추론에 사용될 수 있는 두 종류의 인과 구조를 제시하였다.

그림 1A는 주어진 정보, 즉 흡연과 폐암만으로 이루어진 단일 인과 구조이다. 흡연은 b_c 의 기저율(흡연이 발생할 가능성), 그리고 W_c 의 인과 강도(흡연이 폐암을 일으킬 가능성)를 가진다. 반면 그림 1B는 스트레스와 같이 흡연 이외에 폐암을 일으킬 수 있는 대안적/추가적 정보를 탐색한 후 이를 인과 구조에 반영한 완전 인과 구조이다. 인과 베イズ 모형은 사람들이 흡연과 폐암에 관한 추론을 할 때 흡연과 폐암 외에 추가적 정보가 반영된 완전 모형으로 예측을 한다고 가정한다(Griffiths & Tenenbaum, 2005; 2009; Sloman, 2009).

일반적으로 인과 베イズ 모형에서 대안 원인은 흡연 외에 폐암을 일으키는 사건들의 집합을 의미하며 흡연과 독립적으로 발생한다(Cheng, 1997). 대안 원인 역시 b_a 의 기저율(대안 원인이 발생할 가능성) 그리고 W_a 의 인과 강도(대안 원인이 폐암을 일으킬 가능성)를 가진다. 인과 구조가 작동하는 방식은 흡연과 대안 원인은 단독으로 폐암을 유발할 수 있다는 Noisy-Or(Cheng, 1997; Pearl, 2000) 함수를 따른다. 인과 베イズ 모형은 이러한 가정을 바탕으로 예측 추론($P(E|C)$)과 진단 추론($P(C|E)$)에 대해 아래 수식을 통해 규범적인 예측을 생성한다(Cheng, 1997; Pearl, 2000, 자세

한 수식의 유도는 Fernbach, et al.(2010)을 참고하면 된다).

수식에서 알 수 있듯이 예측 추론 및 진단 추론은 대안 원인의 기저율과 인과 강도에 영향을 받는다. 대안 원인에 의해 결과가 발생할 가능성이 증가하면 예측 추론치($P(E|C)$)는 증가하지만 진단 추론치($P(C|E)$)는 감소한다. 반면 대안 원인에 의해 결과가 발생할 가능성이 감소하면 예측 추론치는 감소하지만 진단 추론치는 증가한다. 예를 들어 철수가 흡연을 할 때 철수가 폐암에 걸릴 가능성은 철수가 흡연 외에 폐암을 일으키는 행위(예: 석면 공장 근무)를 더 많이 할수록 증가한다. 하지만 철수가 폐암에 걸렸을 때 철수가 흡연을 할 가능성은 철수가 흡연 외에 폐암을 일으키는 행위를 더 많이 할수록 감소한다.

예측 추론과 진단 추론의 비대칭성

Fernbach와 그의 동료들(2010, 2011a)은 관련 정보를 탐색하지 않는 정보 무시 현상이 인과 관계를 기반으로 한 예측에서도 관찰된다고 보고하였다. Fernbach와 그의 동료들(2010, 2011a)에 따르면, 사람들은 예측 추론을 할 때 대안적 원인을 고려하지 않았다. 구체적으로 실험 참여자들은 철수가 흡연을 했을 때 철수가 폐암일 가능성(이하 기본 예측 추론, $P(E|C)$)과, 철수가 흡연을 하지만 폐암을 일으킬 다른 이유를 가지고 있지 않을 때

$$P(E|C) = W_c + b_a \cdot W_a - W_c \cdot b_a \cdot W_a$$

$$P(C|E) = 1 - (1 - b_c) \cdot \frac{b_a \cdot W_a}{b_c \cdot W_c + b_a \cdot W_a - b_c \cdot W_c \cdot b_a \cdot W_a}$$

철수가 폐암일 가능성(이하 단독 예측 추론, $P(E|C, -A)$)을 추정하였다. 만일 사람들이 대안적 원인(그림 1B에서 A)을 고려한다면, 기본 예측 추론($P(E|C)$)은 단독 예측 추론($P(E|C, -A)$)보다 높아야 한다. 하지만 실험 결과 기본 예측 추론($P(E|C)$)과 단독 예측 추론($P(E|C, -A)$)은 다르지 않았다.

반면 실험 참여자들이 진단 추론을 할 때는 대안적 정보를 고려하였다. 실험 참여자들은 철수가 폐암일 때 철수가 흡연할 가능성(이하 기본 진단 추론, $P(C|E)$)과, 철수가 폐암이지만 폐암을 일으킬 다른 이유를 가지고 있지 않을 때 철수가 흡연을 할 가능성(이하 단독 진단 추론, $P(C|E, -A)$)을 추정하였다. 만일 사람들이 대안적 정보를 고려한다면 기본 진단 추론($P(C|E)$)이 단독 진단 추론($P(C|E, -A)$)보다 낮아야 하며, 실험 결과는 이러한 예측과 일치하였다. 요약하면 사람들이 예측 추론을 할 때는 관련 정보를 무시하지만, 진단 추론을 할 때는 관련 정보를 고려한다.

예측 추론에서는 대안적 정보, 특히 대안 원인을 고려하지 않지만 진단 추론에서는 대안적 원인을 고려하는 비대칭성은 의사와 같은 전문가들에도 관찰되었으며(Fernbach et al., 2010), 속성 추론(property induction)에서도 관찰되었다(Fernbach et al., 2010). 또한 인과 관계에 대한 표현을 확률이 아닌 빈도로 표현하여 정보의 모호함을 줄였을 때(Hertwig & Gigerenzer, 1999), 제 3자적 관점, 특히 판사의 관점에서 판단을 유도하여 관련 정보의 탐색을 증가시켰을 때(Bonini, Tentori, & Osherson, 2004)에도 두 추론의 비대칭성은 관찰되었다.

대안 무시 현상이 왜 일어나는지에 대해서

는 세 가지 가설이 제안되었다. 첫 번째 가능성은 처리 용이성 가설이다. 예측 추론의 처리 용이성 및 친숙성이 대안 무시 현상의 원인이라는 것이다(Fernbach et al., 2011a; Fernbach & Rehder, 2013). 예측 추론은 원인에서 결과 방향으로 이루어지는 추론이다. 사람들은 일반적으로 인과 관계를 원인에서 결과의 방향으로 생각하기 때문에 예측 추론을 자연스럽게 그리고 쉽게 처리할 수 있는 인지적 과제로 만든다(Tversky & Kahneman, 1980). 이러한 처리의 용이성 및 친숙성은 관련 정보를 탐색하는 등의 인지적 노력을 최소화시키며 다양한 편향을 일으킨다는 것이다(예: Alter, Oppenheimer, Epley, & Eyre, 2007). 반면 결과에서 원인을 판단하는 진단 추론의 경우 인과 모형 학습이 반드시 필요하다(Meder, Mayrhofer, & Waldmann, 2014). 결과적으로 진단 추론은 다양한 정보를 고려할 수 밖에 없다는 것이다.

다양한 선행 연구들이 처리 용이성 가설과 일관된 결과를 보고하고 있다. 선행 연구에 따르면 사람들은 주어진 정보에만 초점을 두고, 자신의 생각이 틀렸을 가능성보다는 참일 가능성만 고려한다(Doherty, Chadwick, Garavan, Barr, & Mynatt, 1996; Doherty, Mynatt, Tweency, & Schiavo, 1979; Hadjichristidis, Sloman, & Over 2009; Klayman & Ha, 1987; Lord, Ross, & Lepper, 1979; Ross & Murphy, 1996). 가용성 및 대표성 휴리스틱으로 대표되는 휴리스틱 처리는 다양한 대안들을 고려하기 보다는 처리의 용이성에 의존한다는 면에서 처리 용이성 가설과 일관된다.

Fernbach와 Rehder(2013)는 대안 원인의 존재

를 명시적으로 언급한 인과 구조를 제시하거나(예: 흡연은 폐암을 유발한다. 스트레스는 폐암을 유발한다), 예측 추론 질문에서 대안 원인의 존재를 언급하였다(예, 철수는 흡연을 한다. 철수의 스트레스 수준은 알 수 없다). 하지만 정보 무시 현상은 사라지지 않았다. Fernbach와 Rehder는 이러한 결과를 바탕으로 정보 무시 현상이 정보 인출의 문제라기보다는 정보를 편향적으로 사용하기 때문에 발생한다고 주장하였다. 하지만 Fernbach와 Rehder는 친숙한 사건(예: 흡연, 폐암)이 아닌 비친숙한 사건(예: 빅토리아 새우의 수면 주기, 호르몬 수준)을 사용하였기 때문에 일반화 가능성에는 제약이 있다.

대안 원인 무시 현상에 대한 두 번째 가설은 화용론적 설명이다. 대안 무시 현상은 사람들이 예측 추론 질문을 인과 강도 질문으로 해석하기 때문이라는 설명이다. 예를 들면 ‘철수가 흡연을 할 때 폐암에 걸릴 가능성’을 묻는 질문을 ‘흡연이 폐암을 일으키는 강도’에 대한 질문으로 이해한다는 것이다. Cummins(2014)는 사람들이 Fernbach 등이 사용한 예측 추론 질문을 인과 강도 질문으로 해석한다는 결과를 보고 하였다(화용론적 설명에 대한 반론은 Fernbach 등(2011b)를 보면 된다).

세 번째 가설은 인과 베이지 모형의 결함에 관한 것이다. Cummins(2014)는 사람들이 예측 추론을 할 때 주어진 원인 외에 대안적 원인을 고려해야 한다는 규범적 모형의 가정이 틀렸다고 주장하였다. 즉 규범적 모형으로 사용된 인과 베이지 모형의 표상(그림 1A)이 사람들이 실제로 사용하는 인과 표상과 다르다는 것이다. Cummins(2014)는 사람들이 예측 추론

시 대안적 원인을 고려하기 보다는 결과 발생을 방해하는 사건들을 떠올린다고 주장하였다. 이를 바탕으로 Cummins(2014)는 그림 1A에서 대안 원인 마디는 억제 변인(preventive 또는 disabling variable)으로 수정되어야 한다고 제안하였다. Mayrhofer와 Waldmann(2014)은 전통적 인과 베이지 모형에 비해 억제 변인을 포함한 수정된 인과 베이지 모형이 사람들의 실제 인과 판단과 부합함을 보였다.

인과 베이지 모형의 경계 조건과 대안 원인 무시 현상

인과 베이지 모형의 규범적 예측과 실제 판단의 불일치는 인과 베이지 모형(causal Bayes nets)의 심리적 타당성(psychological validity)의 문제일 수도 있다. 실제 판단치와 규범적 예측치의 괴리는 인과 베이지 모형이 가정하는 경계 조건과 실제 인지 과정의 차이로 볼 수 있다는 것이다. 다시 말해서 인과 베이지 모형이 사람들이 판단과 예측을 할 때 사용할 것이라고 가정하는 인지 과정과 사람들이 실제로 사용하는 인지 과정이 다를 수 있다는 것이다. 따라서 사람들이 판단과 예측을 할 때 관련정보를 고려할 능력이 있는지를 인과 베이지 모형의 틀에서 검증하기 위해서는 인과 베이지 모형이 가정하고 있는 경계 조건의 만족 여부를 확인하는 것부터 시작해야 한다.

판단과 예측에 대해 인과 베이지 모형을 적용하기 위해서는 먼저 인과적으로 충분한 인과 구조(causally sufficient structure)가 정의되어야 한다. 인과적으로 충분한 구조란 주어진 자료(data)를 설명할 수 있는 인과 구조는 하나

이어야 한다는 것이다. 이는 인과 베이즈 모형에 기본적 가정인 인과 충분성 가정에 기초한 것이다(Pearl, 2000). 또한 모든 파라미터들이 정확하게 인과 모델에 적용되어 free 파라미터가 없어야 한다. 마지막으로 인과 구조가 작동하는 방식이 명확해야 한다(Pearl, 2000; Krynski & Tenenbaum, 2007; Park & Sloman, 2013).

대안 원인 무시 현상이 보고된 연구들은 이러한 경계 조건을 만족시키지 못하고 있다. 먼저 대안 원인 무시 현상 연구에 사용된 시나리오들은 인과적으로 충분하지 않다. 다시 말해 시나리오가 함의하는 인과구조는 공통 결과 구조 외에는 없어야 한다는 인과 충분성 가정을 위반한 것이다. 대안 원인 무시 현상을 보고한 선행 연구들은 공통 결과 인과 구조를 언어적으로 제시하였다(윤민지, 2014; Fernbach et al., 2010, 2011a, 2011b; Fernbach & Rehder, 2013). 흡연과 폐암에 대한 인과 구조는 무엇인가? 인과 베이즈 모형은 그림 1B의 모형을 사용할 것이라고 가정하지만, 실제 사람들은 그림 1A의 인과 구조를 사용했을 가능성이 있다. 이러한 인과 구조의 불충분성은 인과 베이즈 모형에 대한 철학적 논쟁의 핵심이며(Cartwright, 2004; Hausman & Woodward, 1999), 경험적 연구로 확장되고 있다. 인과 추론에서 정보사용의 편향을 보여주는 기저율 무시 현상(Krynski & Tenenbaum, 2007; Barbey & Sloman, 2007)과 차단 규칙 위반 현상(Park & Sloman, 2013, 2014)은 인과적으로 충분한 인과 구조를 제시하면 사라진다.

또한 선행 연구들은 하나의 원인과 하나의 결과에 대한 단일 인과 관계(elementary causal

relation)를 사용하는 경우가 많았다(Cummins, 2014; Fernbach et al., 2010, 2011a). 단일 인과 관계가 확률적으로 작동하게 되면 공통 결과 구조를 내포한다. 하지만 사람들은 모순적인 상황을 만나지 않는 이상 확률적 단일 인과 구조를 인과 베이즈 모형이 가정하는 공통 결과 구조로 확장하지 않는다(Kushnir, Gopnik, Lucas, & Schulz, 2010; Luhmann & Ahn, 2007; Park & Sloman, 2014; Walsh & Sloman, 2004, 2008). 따라서 공통 결과 인과 구조를 내포하는 단일 인과 관계가 아닌 공통 결과 인과 구조가 명백하게 제시한 상황에서 대안 원인 무시 현상이 발생하는지 검증하는 것이 필요하다.

정보 무시 현상을 보고한 Fernbach와 그의 동료들은 대안 원인의 기저율과 대안 원인의 인과 강도를 분리하지 않았다. Fernbach와 그의 동료들은 대안 원인이 결과를 일으키는 영향력을 관심이 되는 원인이 없을 때 결과가 발생하는 확률($P(E| \neg C)$)로 보았다. 인과 베이즈 모형에 따르면 $P(E| \neg C)$ 는 대안 원인의 기저율(b_a)과 대안 원인의 인과 강도(W_a)의 함수이지만 Fernbach와 그의 동료들은 두 파라미터를 구분하지 않고 $P(E| \neg C)$ 를 조작하였다. 결과적으로 대안 원인 무시 현상에 영향을 주는 변인이 기저율인지, 대안 원인의 인과 강도인지, 두 변인의 상호작용인지 알 수 없다.

마지막으로 인과 구조를 언어적으로 제시하는 경우, 인과 구조가 작동하는 방식을 확인할 수 없다. 다양한 원인이 어떻게 결합하여 결과를 생성하는지(Waldman, 2007), 각 변인이 어떤 방식으로 활성화 되는지(Sloman, 2009)에 대한 경계 조건이 만족되지 않은 것이다. 이

러한 분석을 바탕으로, 본 연구는 예측과 판단에 대한 인과 베이스 모형의 경계 조건인 인과 구조 및 파라미터의 충분성을 확보한 상태에서 대안 원인 무시 현상이 발생하는지 살펴보았다.

실험 1

실험 1에서 실험 참여자는 인과 베이스 모형의 경계 조건이 만족되는 상황에서 세 슬라이더로 구성된 공통 결과 구조를 학습한 후 예측 추론과 진단 추론을 실시하였다. 먼저 인과 구조와 파라미터의 충분성을 확보하기 위해 사용한 방법은 50회의 관찰로 이루어진 trial-by-trial 학습이다. 공통 결과 구조를 학습하기 위한 50회의 시행은 부록 1에 제시하였다. 50회 시행은 원인 슬라이더의 기저율이 0.5, 원인 슬라이더가 활성화되었을 때 결과 슬라이더가 활성화될 확률은 0.8, 원인 슬라이더가 활성화되지 않았지만 결과 슬라이더가 활성화될 확률은 0.2인 확률적 공통 결과 구조에서 생성되었다. 50회의 시행은 인과적으로 충분한 시행이다. 다시 말해 50회의 시행을 바탕으로 유도할 수 있는 인과 모형은 공통 결과 구조 외에는 없다(Scheines, Spirtes, Glymour, & Meek, 1994; Fernbach & Sloman, 2009). 또한 50회의 trial-by-trial 학습을 통해 인과 강도, 기저율 등 파라미터에 대한 정보를 충분히 획득할 수 있다(Lagnado & Sloman, 2004).

인과 구조가 작동하는 방식은 Noisy-Or (Cheng, 1997; Pearl, 2000) 함수를 따른다. 인과 구조가 작동되는 방식을 명확하게 하기 위해

실험 1은 세 슬라이드는 인과 관계와 부합하는 시간적 정보를 함께 제시하였다. 시간적 정보는 인과 관계를 암시하는 가장 강력한 요소 중 하나이다: 결과는 원인보다 먼저 발생할 수 없다(Sloman & Lagnado, 2015). 예를 들어 A가 B와 C의 공통 결과라고 가정한다면, 참여자는 B가 활성화된 후 A가 활성화되는 장면, C가 활성화된 후 A가 활성화되는 장면, B와 C가 동시에 활성화된 후 A가 활성화되는 장면을 통해 공통 결과 구조를 학습하게 된다.

학습 단계 이후 참여자들은 자신이 생각한 인과 구조를 보고하였으며 예측 추론과 진단 추론으로 이루어진 인과 추론 과제를 실시하였다. 실험 참여자들이 예측 추론에서 대안 원인(그림 1B에서 A)을 무시한다면 기본 예측 추론($P(E|C)$)과 단독 예측 추론($P(E|C, -A)$)은 다르지 않아야 한다. 하지만 실험 참여자들이 대안 원인을 고려하여 예측 추론을 한다면 기본 예측 추론($P(E|C)$)이 단독 예측 추론($P(E|C, -A)$)보다 높아야 한다. 그러나 실험 참여자들이 진단 추론에서 대안 원인을 무시한다면 기본 진단 추론($P(C|E)$)과 단독 진단 추론($P(C|E, -A)$)이 같아야 한다. 하지만 실험 참여자들이 대안 원인을 고려하여 진단 추론을 한다면 기본 진단 추론($P(C|E)$)과 단독 진단 추론($P(C|E, -A)$)보다 낮아야 한다.

방법

실험 참여자 36명의 브라운 대학교 학부생들이 참여하였다. 실험은 PC가 설치된 부스에서 진행되었으며, 실험 참여자들은 실험 참여 점수를 받았다.

실험 설계 인과 구조(3: 공통 원인 vs. 인과 연쇄 vs. 공통 결과)와 과제(인과 구조 보고 vs. 인과 추론)는 피험자내 변인, 과제 순서(선 인과 추론 vs. 후 인과 추론)는 피험자간 변인인 혼합 설계를 사용하였다. 공통 원인 구조와 인과 연쇄 구조는 대안 원인 무시 현상과 관련이 없기 때문에 결과 분석에서 제외하였다. 인과 구조 학습을 마치고 실험 참여자의 절반은 인과 구조를 보고한 후 인과 추론 과제를 시행하였다. 나머지 절반은 인과 추론 과제를 시행한 후 인과 구조를 보고하였다.

재료 및 절차 실험이 시작되면 실험 참여자는 모니터를 통해 슬라이더를 이용한 실험 형식, 비순환 그래프(acyclic graph), 확률적 인과 관계(probabilistic causal relation)에 대한 설명을 받았다. 또한 참여자들에게 세 슬라이더(그림 2)의 활성화 상태를 50회 동안 관찰한 후 세 슬라이더간 인과 관계를 파악하는 것이 실험의 목적이라고 설명하였다. 그림 2A는 활성화 전 세 슬라이더의 상태를 나타낸다. 3개의 회색 슬라이더가 삼각형 모양으로 화면에 제시되었다. 참여자는 그림 2A 하단에 있는 ‘TEST’

버튼을 눌러 세 슬라이더의 상태를 확인할 수 있다. 세 슬라이더의 활성화 상태는 동시에 나타나는 것이 아니라 인과 관계에 따라 시간적 순서대로 제시되었다. 예를 들어 슬라이더 A가 슬라이더 B와 슬라이더 C의 공통 결과라고 가정해 보자. 그림 2는 실험 참여자가 모든 슬라이더가 활성화되는 상황을 시간적 순서에 따라 제시하였다. 참여자가 화면 하단에 TEST 버튼을 누르면(그림 2A), TEST 버튼은 사라지며 슬라이더 B와 슬라이더 C가 동시에 활성화 되고(그림 2B), 1초 후 슬라이더 A가 활성화된다(그림 2C). 이후 활성화된 슬라이더는 모두 비활성화되고 TEST 버튼은 다시 나타난다. 참여자는 연습 시행에서 실험 방법 이해 정도를 확인한 후, 학습 시행을 실시하였다.

50회의 학습이 끝난 후 참여자는 소책자를 이용해 인과 구조 보고 과제를 실시하였다. 첫 페이지에 방향적 비순환 그래프와 숨겨진 변인에 대한 개념적 설명을 제시하였다. 추가적으로 Glymour(1998)의 예를 통해 숨겨진 변인과 방향적 비순환 그래프의 예를 제공하였다. 지시문을 읽은 후 실험 참여자는 다음 페

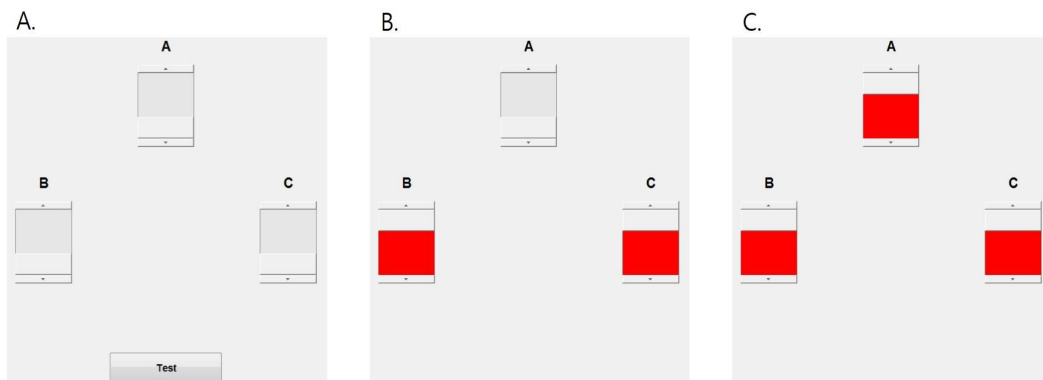


그림 2. 인과 학습 과제에서 참여자들이 관찰하는 화면의 예

이지에 본인이 생각할 때 가장 그럴듯한 인과 구조를 직접 그래프로 나타내었다. 실험 참여자가 그린 그래프에 대해서 피드백은 제공되지 않았다.

인과 추론 과제는 각 슬라이더의 상태를 다른 두 슬라이더들의 상태를 바탕으로 추론하는 24개의 조건 확률 문제들로 구성되었다. 24개의 인과 추론 문제는 슬라이더들의 상태를 보여주면서 언어적으로 질문하였다. 슬라이더 A가 슬라이더 B와 슬라이더 C의 공통 결과일 때 실험 참여자에게 주어진 단독 예측, 기본 예측, 단독 진단, 기본 진단 추론 질문은 다음과 같다.

단독 예측 질문

슬라이더 B는 활성화 되었지만 슬라이더 C는 활성화되지 않았다.

슬라이더 A가 활성화되었을 가능성은 얼마인가?

기본 예측 질문

슬라이더 B는 활성화 되었지만 슬라이더 C의 상태는 알 수 없다.

슬라이더 A가 활성화되었을 가능성은 얼마인가?

단독 진단 질문

슬라이더 A는 활성화 되었지만 슬라이더 B는 활성화되지 않았다.

슬라이더 C가 활성화되었을 가능성은 얼마인가?

기본 진단 질문

슬라이더 A는 활성화 되었지만 슬라이더 B의 상태는 알 수 없다.

슬라이더 C가 활성화되었을 가능성은 얼마인가?

그림 3은 단독 예측 추론을 할 때 실험 참

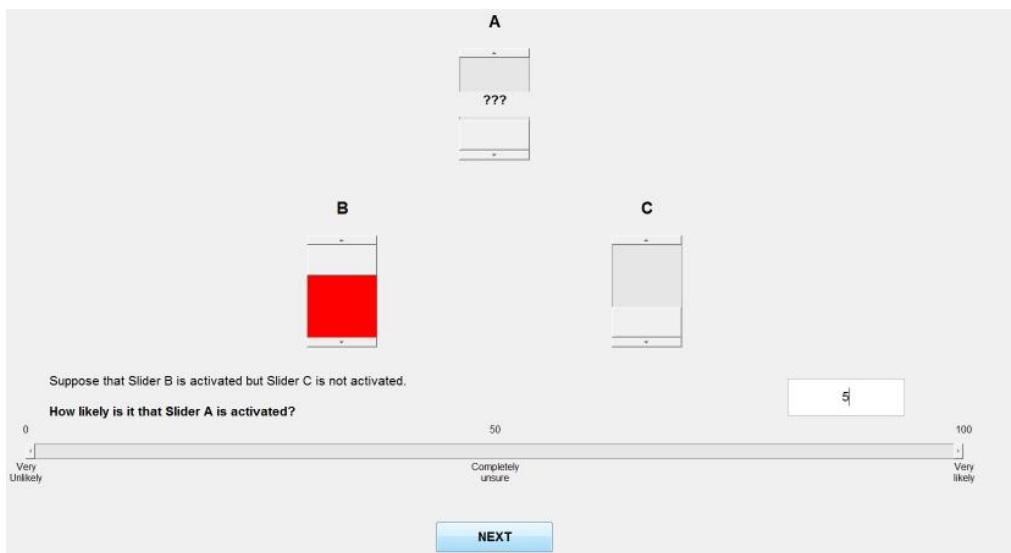


그림 3. 예측 질문의 예

여자에게 주어지는 화면이다.

24개 질문 중 대안 무시 현상과 관련된 여섯 개의 예측 추론 질문과 여섯 개의 진단 추론 질문에 대한 추정치를 분석에 사용하였다. 두 과제들을 모두 마친 실험 참여자는 다음 인과 구조 학습 단계로 진행하였다. 인과 구조의 제시 순서와 공간적 배치는 역균형화(counterbalancing) 하였다. 인과 예측 질문들은 무선적으로 제시되었다.

결 과

구조 그리기 과제 36명중 26명이 정확히 공통 결과 구조를 그렸다. 주어진 슬라이더 외의 다른 변인을 포함하여 인과 관계를 그린 피험자는 없었다. Fisher exact 검증 결과 선 인과 추론 조건(18명중 12명)과 후 인과 추론 조건(18명중 14명)간 유의미한 차이는 없었다, $p = .457$. Binomial test 결과 정확하게 구조를 그린 실험 참여자의 비율이 부정확한 구조를 그린 참여자의 비율보다 유의미하게 높았다(우연수준 = 0.5), $Z = 2.5$, $p = .05$.

인과 추론 공통 결과 구조를 정확히 찾은 26명의 인과 추론 결과에서 대안 무시 현상이 관찰되는지 살펴보았다. 인과 추론 결과의 표기는 E가 C1과 C2의 공통 결과인 인과 구조를 바탕으로 이루어졌다.

예측 추론. 24개의 인과 추론 중 예측 추론과 관련이 있는 질문은 8개이다. 그 중 여섯 개 질문이 대안 원인 무시 현상과 관련이 있다. C1이 단독으로 활성화되었을 때 E가 활

성화되었을 확률에 대한 질문($P(E|C1, -C2)$)과 C2가 단독으로 활성화되었을 때 E가 활성화되었을 확률에 대한 질문($P(E|C2, -C1)$)은 질적으로 동일한 질문이다. 따라서 두 질문의 확률 추정치에 대한 평균을 내었으며, $P(E | C, -A)$ 로 표기하였다. 마찬가지로 $P(E|C1)$ 과 $P(E|C2)$ 역시 질적으로 동일한 질문이기 때문에 평균을 내었으며 $P(E|C)$ 로 표시하였다. 마지막으로 $P(E|C1, C2)$ 는 표기의 일관성을 위해 $P(E|C, A)$ 로 표시하였다.

과제 수행 순서(2: 선 인과 추론 vs. 후 인과 추론)는 피험자간 변인, 확률 추정치(3: $P(E|C, -A)$ vs. $P(E|C)$ vs. $P(E|C, A)$)는 피험자내 변인으로 하는 혼합 변량 분석을 실시하였다. 분석 결과 참여자들은 대안 무시 원인 현상과는 다르게 대안 원인의 상태를 고려하여 예측 추론을 하는 것으로 밝혀졌다. 확률 추정치의 주효과는 유의하였다, $F(2, 48) = 28.62$, $MSE = 76.67$, $p < .001$, $partial \eta^2 = .544$. 과제 순서 및 상호 작용은 유의하지 않았다, $F_s < 1$. 주효과 분석 결과(그림 4) 기본 예측 추론($P(E|C)$)(평균: 82.69 / 표준오차: 2.82)이 단독 예측 추론($P(E|C, -A)$)(평균: 75.00 / 표준오차: 3.82)보다 유의미하게 높았다, $t(25) = 3.9$, $SE = 1.98$, $p < .001$. 이는 인과 베이스 모형의 규범적 예측과 일치하며 실험 참여자들이 예측 추론을 할 때 대안 원인을 고려하였다는 것을 시사한다.

$P(E|C, A)$ (평균: 93.46 / 표준오차: 1.43)는 $P(E|C, -A)$ 보다 유의하게 높았고, $t(25) = 5.9$, $SE = 3.12$, $p < .001$, $P(E|C)$ 보다 유의하게 높았다, $t(25) = 5.8$, $SE = 1.86$, $p < .001$. 이 결과 역시 베이스 모형의 예측과 일치한다.

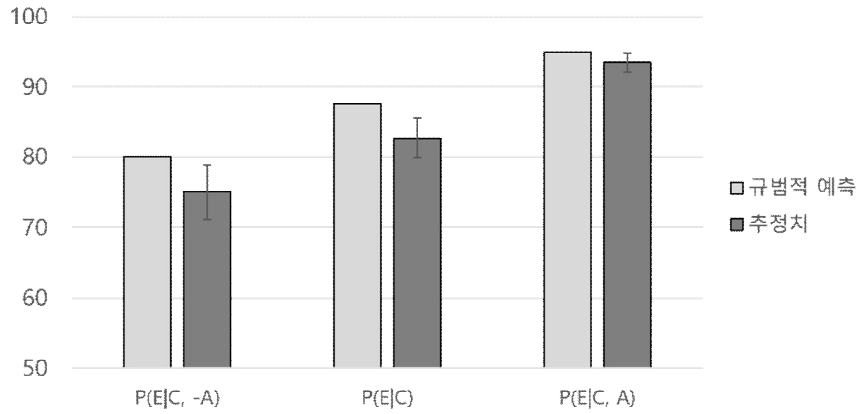


그림 4. 실험 1에서 예측 추론 판단치의 평균과 표준오차

다음으로 각 예측 판단에 대해 규범적 예측치를 계산하고(부록 1), 이를 실제 판단치와 비교하였다. $P(E|C, -A)$ 의 규범적 예측치는 80으로 실제 판단치와 다르지 않았다, $t(25) = 1.30, n.s.$ 다른 두 예측 판단치도 같은 패턴이 관찰되었다. $P(E|C)$ 의 규범적 예측치는 87.5, $P(E|C, A)$ 의 규범적 예측치는 95였으며, 실험 참여자의 실제 판단치와 다르지 않았다, 각각 $t(25) = 1.70, n.s., t(25) = 1.07, n.s.$

진단 추론. 24개의 인과 추론 중 진단 추론과 관련이 있는 질문은 16개이다. 그 중 여섯 개 질문이 대안 원인 무시 현상과 관련이 있다. 이 6개의 질문은 질적으로 동일한 2개의 세트로 구분할 수 있다, $P(C1|E, -C2), P(C1|E), P(C1|E, C2)$ vs. $P(C2|E, -C1), P(C2|E), P(C2|E, C1)$. 예측 추론과 같이 질적으로 동일한 질문에 대한 참여자의 추정치를 평균하였으며, 각각 $P(C|E, -A), P(C|E), P(C|E, A)$ 로 표기하였다.

과제 수행 순서(2: 선 인과 추론 vs. 후 인과 추론)는 피험자간 변인으로 확률 추정치

(3: $P(C|E, -A), P(C|E), P(C|E, A)$)는 피험자내 변인으로 하는 혼합 변량 분석을 실시하였다. 그 결과 실험 참여자들은 대안적 원인을 고려하여 진단 추론을 하는 것으로 드러났다. 확률 추정치의 주효과는 유의하였지만, $F(2, 48) = 10.36, MSE = 195.34, p < .001, partial \eta^2 = .301$. 과제 순서 및 상호 작용은 유의하지 않았다, $F_s < 1$. 주효과 분석 결과(그림 5), 단독 진단 추론($P(C|E, -A)$)(평균: 80.19 / 표준오차: 4.05)은 기본 진단 추론($P(C|E)$)(평균: 67.30 / 표준오차: 3.18)보다 유의하게 높았다, $t(25) = 3.4, SE = 3.74, p < .001$. 실험 참여자는 진단 추론시 대안 원인을 고려한 것이다.

$P(C|E, A)$ (평균: 63.26 / 표준오차: 3.52)보다 유의하게 높았다, $t(25) = 3.8, SE = 1.86, p < .001$. $P(C|E)$ 와 $P(C|E, A)$ 는 유의한 차이가 없었다, $t(25) = 1.4, SE = 3.04, n.s.$

각 진단 판단에 대해 규범적 예측치를 계산하고(부록 1) 이를 실제 판단치와 비교하였다. $P(C|E, -A)$ 의 규범적 예측치는 80으로 실제 판단치와 다르지 않았다, $t(25) < 1$. 또한 $P(C|E)$ 의 규범적 예측치는 63이었으며 실험 참여자

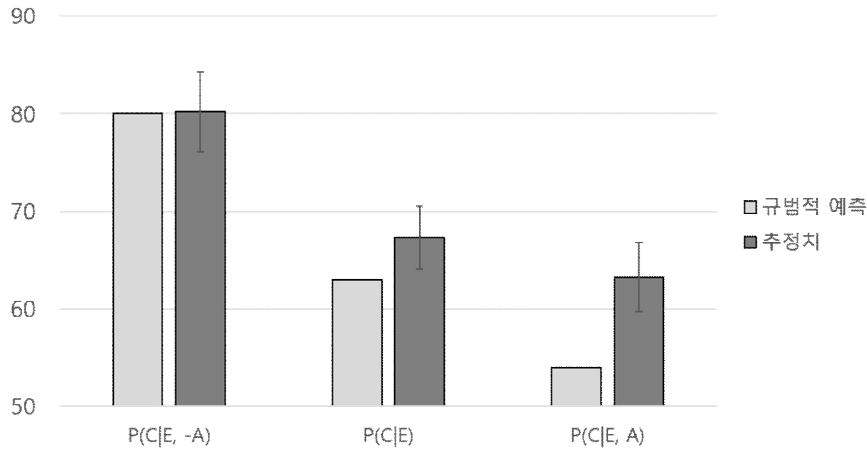


그림 5. 실험 1에서 진단 추론 판단치의 평균과 표준오차

의 실제 판단치와 다르지 않았다, $t(25) = 1.35, n.s.$ 하지만 $P(E|C,A)$ 의 규범적 예측치는 54였지만 실제 판단치는 이보다 유의하게 높았다, $t(25) = 2.63, p < .05$.

논 의

실험 1의 가장 큰 목표는 인과 구조와 파라미터의 충분성, 그리고 인과 구조가 작동하는 방식의 명확성이 확보된 상태에서 대안 원인 무시 현상이 관찰되는지 살펴보는 것이다. 이를 위해 실험 1에서는 인과적으로 충분한 자료를 관찰한 후, 어떤 인과 구조를 파악했는지 알아보는 인과 구조 과제와 예측 추론 그리고 진단 추론 과제를 실시하였다.

인과 구조 과제에서 정확하게 인과 구조를 보고한 참여자에 한해 인과 추론 과제 결과를 분석하였다. 그 결과 대안 원인 무시 현상은 관찰되지 않았다. 기본 예측 추론은 단독 예측 추론보다 유의하게 높았으며, 기본 진단 추론은 단독 예측 추론보다 유의하게 낮았다.

실험 1의 결과는 인과 베이지 모형의 경계 조건이 만족된 상황에서 사람들은 규범적 판단을 내릴 수 있다는 사실을 시사한다.

그렇다면 인과 구조와 파라미터, 그리고 인과 구조가 작동하는 방식을 명확하게 한다면 예측의 정확성은 향상된다고 결론내릴 수 있을까? 다시 말해, (1) 교통량에 영향을 주는 원인(예: 통행료, 기후, 경제 상황 등)과 인과 강도를 파악하고 (2) 각 원인들의 어떤 방식으로 조합되어 결과를 일으키는지 명백히 정의된다면, 교통량을 정확히 예측할 수 있을까?

인과 추론에 대한 선행 연구들은 인과 구조와 파라미터의 충분성은 인과 구조 학습과 인과 추론의 필요충분조건(necessary and sufficient condition)이다(Pearl, 2000; Griffiths & Tenenbaum, 2005). 하지만 대안 원인 무시 현상은 인과 구조가 작동하는 방식에 영향을 받을 가능성이 있다(Lucas & Griffiths, 2010; Waldmann, 2007) Lucas와 Griffiths(2010), 그리고 Waldmann(2007)의 연구에서 참여자들은 공통 결과 구조에 의해 생성된 공변 자료를 학습한 후 예측 추론

과제를 실시하였다. Lucas와 Griffiths(2010), 그리고 Waldmann(2007)은 학습 시행이 시작되기 전 인과 구조가 작동하는 방식을 조작하였다. 참여자의 절반은 학습할 인과 구조가 원인이 모두 활성화되어야 결과가 활성화되는 연결 구조(conjunctive structure)라는 설명을 듣거나 학습을 하였고, 나머지 절반은 하나의 원인이 활성화되어도 결과가 활성화되는 이접 구조(disjunctive structure)라는 설명을 듣거나 학습하였다. 이 후 두 집단은 동일한 공변 정보를 학습하였음에도 예측 추론의 패턴은 달랐다.

만일 인과 구조 작동 방식(functional form)의 학습이 대안 원인 무시 현상과 관련이 있다면, 인과 구조 작동 방식의 변화는 대안 원인을 다루는 방식을 변화시킬 것이다. 실험 2는 실험 1과 인과 구조는 같지만 인과 구조 작동 방식을 변경시켜 대안 원인 무시 현상이 관찰되는지 살펴보았다.

실험 2

실험 2에서 참여자는 세 변수 중 하나 혹은 두 변인의 상태를 사전에 설정하고(이하 상태 개입), 자신의 상태 개입의 결과를 관찰하는 방식으로 인과 구조를 학습하였다. 즉 실험 참여자가 원인 변인을 강제로 활성화시키지 않는 한 결과 변인은 활성화되지 않는다. 비록 인과 구조가 작동하는 방식은 다르지만(원인의 자발적 활성화 vs. 개입에 의한 원인의 활성화), 개별 학습 시행에서 실험 참여자들이 관찰하는 변인들의 활성화 정보는 동일하다. 예를 들어 실험 1과 실험 2의 참여자들은 세 변인이 모두 활성화된 상태를 관찰할

수 있다. 하지만 실험 1에서는 두 원인 변인이 자발적으로 활성화된 후 결과 변인이 활성화된 것이고, 실험 2에서는 참여자가 두 원인 변인을 강제로 활성화 시킨 후 결과 변인이 활성화된 것이다.

공통 결과 구조를 학습하고 인과 추론 과제를 실시한다는 면에서 실험 1과 동일하다. 하지만 상태 개입 학습을 통한 인과 구조 학습을 가능하게 하기 위해 인과 구조의 파라미터를 수정하였다. 실험 2의 개별 시행은 기저율 0, 인과 강도는 0.8, 원인이 없을 때 결과가 발생하는 확률은 0인 공통 결과 구조에서 생성되었다. 실험 2에서는 원인이 자발적 활성화되지 않지만 원인이 존재하였을 때 결과가 존재하는 확률은 0.8로 실험 1과 동일하다.

한가지 주목해야 할 점은 실험 1과는 달리 실험 2에서는 참여자별로 개별 학습 시행의 수가 다르다. 학습 시행의 수가 다른 것은 참여자간 학습 정보의 인과적 충분성이 다를 가능성이 있다. 실험 2에서는 개별 실험 참여자의 학습이 인과적으로 충분한지 Fernbach와 Sloman(2009)이 사용한 인과 충분성 검증을 사용할 것이다. 그 외에는 실험 1과 동일하다.

요약하면 실험 1과 실험 2는 관찰된 인과 구조와 확률 정보는 질적으로 동일하지만 각 인과 구조가 작동하는 방식은 다르다. 실험 1은 자발적으로 작동할 수 있는 인과 구조이고 실험 2는 개입에 의해서만 작동한다. 사람들은 이러한 작동 방식의 차이를 구분하며 질적으로 구분되는 인과 추론을 한다(Lucas & Griffiths, 2010; Mayrhofer, Goodman, Waldmann, & Tenenbaum, 2008; Meder, Hagmayer, & Waldmann, 2008, 2009; Waldmann, 2007).

만일 사람들이 관찰된 정보에 의해서 인과 추론을 한다면 실험 2의 인과 추론 결과는 실험 1과 같을 것이다. 예측 추론과 진단 추론에서 대안 원인을 고려할 것이다. 반면 관찰된 정보가 서로 다른 작동 방식에서 나온 것을 고려한다면, 대안 원인 무시 현상에 대해 예측 추론과 진단 추론의 비대칭성이 관찰될 것이다.

기본 예측 추론에서 대안 원인의 상태가 알 수 없다는 것은 대안 원인이 활성화되지 않은 것으로 해석될 것이다. 원인의 기저율이 0이고 개입에 의해서만 활성화되는 상황에서 참여자의 개입이 없다면, 대안 원인은 활성화되지 않은 것이다. 따라서 기본 예측 추론과 단독 예측 추론이 같아진다.

진단 추론에서 결과 변인이 활성화되었다는 것은 원인 변인들에 대한 개입이 발생했다는 것을 의미한다. 따라서 결과 변인의 활성화가 어떤 원인 변인에 개입에 의해 발생했는지에 대한 예측을 하게 된다. 어느 변인에 개입이 발생했는지에 대한 추론(실험 2)은 어느 변인이 활성화되었는지에 대한 추론(실험 1)과 질적으로 동일하다. 결과적으로 실험 1과 같이 대안 원인을 고려한 추론이 발생할 것이다.

방 법

실험 참여자 및 실험 설계 36명의 브라운 대학교 학부생들이 참여하였다. 실험은 PC가 설치된 부스에서 진행되었으며, 실험 참여자들은 실험 참여 점수를 받았다. 실험 2는 실험 1과 동일한 혼합 설계를 사용하였다. 실험 1과 동일하게 공통 원인 구조, 인과 연쇄 구조의

실험 결과는 본 실험의 목적과 관련이 없기 때문에 분석에 포함시키지 않았다.

재료 및 절차 실험 1은 관찰 학습 패러다임(observation learning paradigm)을 사용한 반면 실험 2는 상태-개입 학습 패러다임(intervention learning paradigm; Lagnado & Sloman, 2004)을 사용하였다. 그 외의 전반적 실험 절차는 실험 1과 동일하다.

그림 6에 실험 참여자들이 학습에 사용한 화면을 제시하였다. 화면 하단에는 각 슬라이더의 상태를 강제-개입하는 세 쌍(활성화-비활성화, on-off)의 회색 버튼이 해당 슬라이더 바로 밑에 제시되었다. 활성화 버튼(A_on, B_on, C_on)은 슬라이더 상태를 강제적으로 활성화 시켰고, 비활성화 버튼(A_off, B_off, C_off)은 슬라이더 상태를 강제적으로 비활성화 시켰다. 활성화 버튼과 비활성화 버튼을 누르면 버튼은 붉은색으로 변한다. 해당 슬라이더, 슬라이더의 이름 역시 붉은색으로 변하며, 슬라이더 바는 위쪽으로 이동한다. 실험 참여자는 한번에 최대 두 개의 슬라이더의 상태를 강제할 수 있었다. 그림 6B는 실험 참여자가 슬라이더 A는 활성화, 슬라이더 B는 비활성화 상태로 강제한 상태를 보여준다. 각 슬라이더의 상태를 강제 한 후 화면 우측 하단에 있는 ‘TEST’ 버튼을 누르면 상태 개입의 결과를 보여준다. 그림 6C와 6D는 그림 6B의 상태-강제에 대한 두 가능한 결과를 보여준다.

실험 참여자는 본 학습 시행 전 연습 시행을 실시하였다. 연습 시행 직전 19개의 모든 가능한 상태-개입의 예를 나열한 후 이를 직접 시행해 보도록 요청하였다. 연습 시행에서

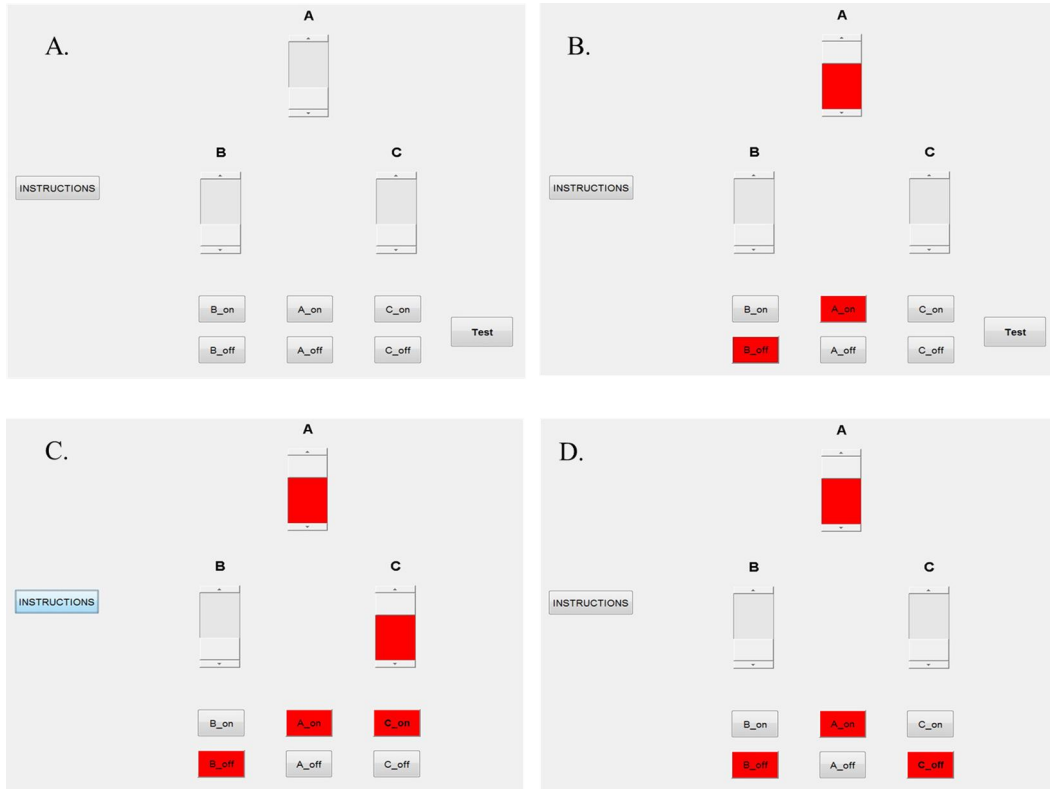


그림 6. 인과 학습 단계에서 피험자들에게 제시된 화면의 예.

A: 학습 시행 전 화면, B: 실험 참여자가 슬라이더 A는 활성화 슬라이더 B는 비활성화시킨 화면, C: 그림 6B의 결과로 슬라이더 C가 활성화된 화면 D: 그림 6B의 결과로 슬라이더 C가 비활성화된 화면

세 슬라이더는 특정 인과 구조에 대한 편향이 발생하지 않도록 서로 독립적으로 설정되었다.

본 학습 시행에서 상태-개입의 결과는 원인 슬라이더에 대한 기저율은 0, 원인 슬라이더가 활성화되었을 경우 결과 슬라이더가 활성화될 확률은 0.8, 원인 슬라이더가 활성화되지 않았을 때 결과 슬라이더가 활성화될 확률은 0인 확률적 공통 결과 구조에 의해 생성되었다. 20회의 시행이 지나면 ‘TEST’ 버튼 우측에 ‘NEXT’ 버튼이 나타나며 실험 참여자는 인과 구조를 파악하였다고 생각했을 때 ‘NEXT’ 버튼을 누르도록 지시받았다. 학습이 끝난 후

참여자는 소책자를 이용해 인과 구조 보고 과제와 24개 인과 추론 과제를 실시하였다.

결 과

인과 학습 시행과 인과 충분성 검증 상태-강제를 통한 학습 단계에서 실험 참여자가 시도한 상태-강제의 종류와 수는 참여자마다 다르다. 공통 결과 구조를 파악하기 위해 충분한 학습 시행이 이루어졌는지 살펴보기 위해 Fernback와 Sloman(2009)의 절차를 사용하였다. 각 실험 참여자의 학습 시행이 공통 결과 구

조에 대해 인과적으로 충분한지 여부를 베이즈 추론 알고리즘(부록 2)을 이용해서 검토하였다. 모든 실험 참여자의 학습 시행이 인과적으로 충분함을 확인하였다. 결과적으로 모든 실험 참여자는 동일한 인과 구조와 인과 강도를 학습하였다.

구조 그리기 과제 36명중 23명이 정확히 공통 결과 구조를 그렸다. 주어진 슬라이더 외의 다른 변인을 포함하여 인과 관계를 그린 피험자는 없었다. Fisher exact 검증 결과 선 인과 추론 조건(18명중 11명) 과 후 인과 추론 조건(18명 중 12)간 유의미한 차이는 없었다, $p = .729$. Binomial test 결과 정확하게 구조를 그린 실험 참여자의 비율이 부정확한 구조를 그린 참여자의 비율과 다르지 않았다(우연수준 = 0.5), $Z = 1.5$, $n.s.$

상태 개입 시행 수 각 참여자에 대해 상태 개입 학습 시행 수를 분석하였다. 먼저, 인과 구조를 정확히 찾은 23명은 평균 35.9회

($SE=3.4$)의 개입 학습을 시행하였으며, 인과 구조를 틀리게 보고한 13명은 평균 27.1($SE=2.6$)의 개입 학습을 시행하였다, $r_{pb}(35) = .30$, $p < .10$.

인과 추론 과제 공통 결과 구조를 정확히 찾은 23명을 대상으로 분석하였다. 실험 1과 동일하게 질적으로 동일한 질문은 평균을 구하여 분석에 사용하였다.

예측 추론. 과제 수행 순서(2: 선 인과 추론 vs. 후 인과 추론)는 피험자간 변인, 확률 추정치(3: $P(E|C, -A)$ vs. $P(E|C)$ vs. $P(E|C, A)$)는 피험자내 변인으로 하는 혼합 변량 분석을 실시하였다. 분석 결과 확률 추정치의 주효과는 유의하였다, $F(2, 42) = 10.12$, $MSE = 118.01$, $p < .001$, $partial \eta^2 = .325$. 과제 순서 및 상호 작용은 유의하지 않았다, $F_s < 1$.

주효과 분석 결과(그림 7) 대안 원인 무시 현상이 관찰되었다. 기본 예측 추론($P(E|C)$)(평균: 86.63 / 표준오차: 2.99)과 단독 예측 추론

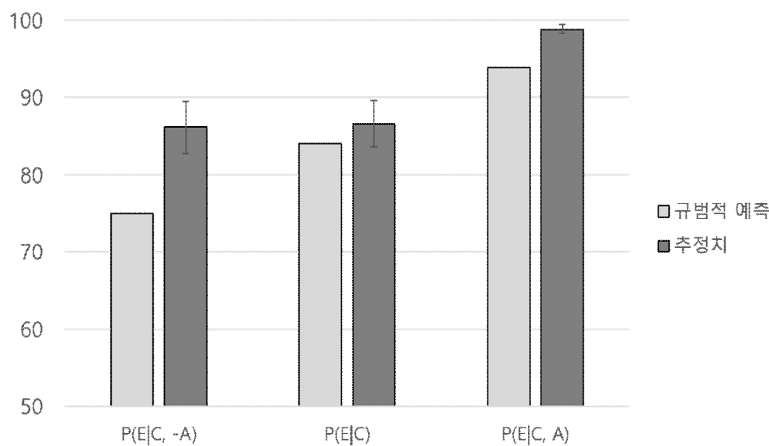


그림 7. 실험 2에서 예측 추론 판단치의 평균과 표준오차

$(P(E|C, -A))$ (평균: 86.19 / 표준오차: 3.34)은 다르지 않았다, $t(22) < 1$, *n.s.*.

또한 $P(E|C, A)$ (평균: 98.91 / 표준오차: .54)는 $P(E|C, -A)$ 보다 유의하게 높았고, $t(22) = 3.99$, $SE = 3.07$, $p < .001$, $P(E|C)$ (평균: 86.63 / 표준오차: 2.99)보다 유의하게 높았다, $t(22) = 5.8$, $SE = 1.86$, $p < .001$.

다음으로 각 예측 판단에 대해 규범적 예측치를 계산하고(부록 1) 이를 실제 판단치와 비교하였다. $P(E|C, -A)$ 의 규범적 예측치는 75로 참여자의 추정치와 달랐다, $t(22) = 3.34$, $p < .01$. $P(E|C)$ 의 규범적 예측치는 84로 참여자의 판단치와 다르지 않았다. 하지만 $P(E|C, A)$ 의 규범적 예측치 94는 실험 참여자의 실제 판단치와 달랐다, $t(22) = 4.9$, $p < .001$.

진단 추론. 과제 수행 순서(2: 선 인과 추론 vs. 후 인과 추론)는 피험자간 변인으로 확률 추정치(3: $P(C|E, -A)$, $P(C|E)$, $P(C|E, A)$)는

피험자내 변인으로 하는 혼합 변량 분석을 실시하였다. 실험 참여자들은 대안적 원인을 고려하여 진단 추론을 하는 것으로 드러났다. 확률 추정치의 주효과는 유의하였지만, $F(2, 42) = 4.73$, $MSE = 483.62$, $p < .05$, *partial* $\eta^2 = .184$, 과제 순서 및 상호 작용은 유의하지 않았다, $F_s < 1$.

주효과 분석 결과(그림 8), 진단 추론에서는 대안 원인 무시 현상이 관찰되지 않았다. 단독 진단 추론($P(C|E, -A)$)(평균: 53.80 / 표준오차: 7.25)은 기본 진단 추론($P(C|E)$)(평균: 40.86 / 표준오차: 4.78)보다 유의하게 높았다, $t(22) = 2.0$, $SE = 6.48$, $p = .06$. $P(C|E, -A)$ 는 $P(C|E, A)$ (평균: 35.21 / 표준오차: 5.04)보다 유의하게 높았으며, $t(22) = 2.4$, $SE = 7.60$, $p < .05$. $P(C|E)$ 와 $P(C|E, A)$ 는 유의한 차이가 없었다, $t(22) = 1.1$, $SE = 4.95$, *n.s.*,

각 진단 판단에 대해 규범적 예측치를 계산하고(부록 1) 이를 실제 판단치와 비교하였다.

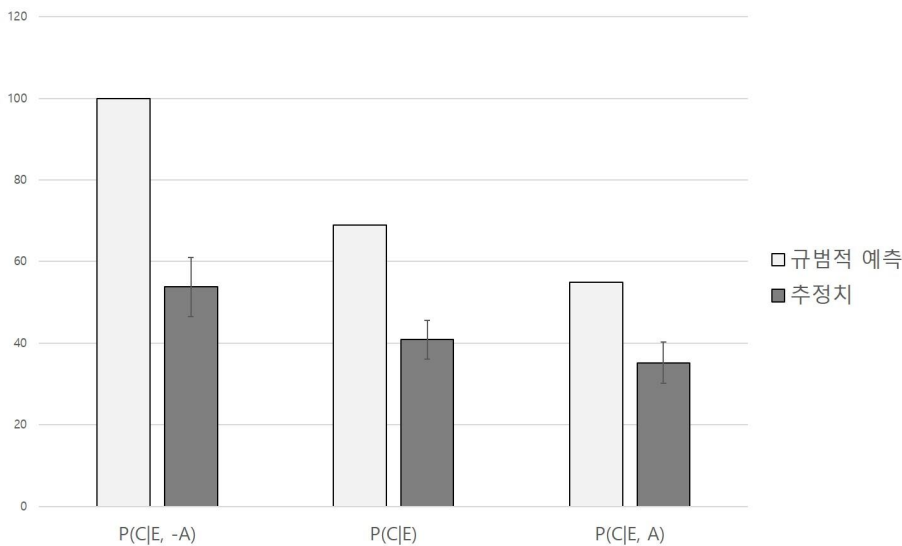


그림 8. 실험 2에서 진단 추론 판단치의 평균과 표준오차

$P(C|E, -A)$ 의 규범적 예측치는 100으로 실제 판단치와 달랐다, $t(22) = 6.37, p < .001$, 또한 $P(C|E)$ 의 규범적 예측치는 69였으며 실험 참여자의 실제 판단치와 달랐다, $t(22) = 5.87, p < .001$. 하지만 $P(E|C, A)$ 의 규범적 예측치는 55이며 실제 판단치는 이보다 유의하게 낮았다, $t(22) = 3.92, p < .01$.

실험 2에서 실제 판단치가 규범적 예측치와 차이가 많이 나는 것은 실제 관찰된 원인 변인들의 기저율이 낮았기 때문일 가능성이 있다. 즉 규범적 예측치는 0.5의 기저율을 사용했지만 실제 관찰된 자료는 이보다 낮았을 것이라는 것이다. 실제 실험 참여자가 관찰한 기저율을 바탕으로 규범적 예측치를 계산하는 것은 본 연구 목적의 달성과 큰 의미가 없다. 본 연구는 대안 원인 무시 현상을 검증하는 것으로 규범적으로 예측된 대안 원인에 대한 고려가 실제 판단에서도 나타났다는 점을 중심으로 결과를 해석해야 할 것이다.

논 의

실험 2의 목표는 인과 정보가 작동하는 방식의 차이가 대안 무시 현상에 미치는 영향을 살펴보는 것이었다. 실험 2에서는 관찰되는 인과 정보의 질은 실험 1과 동일하지만 인과 정보가 작동하는 방식이 차이가 나는 인과 정보를 학습한 후 인과 추론을 실시하였다.

실험 1과 실험 2에 사용된 인과 학습 정보는 공통 결과 구조에서 생성되었으며, 원인 변인이 활성화되었을 때 결과 변인이 활성화되는 확률 역시 동일하였다. 비록 각 원인 변인이 활성화된 상태를 관찰한 빈도는 다르지

만 원인 변인의 활성화를 관찰했다는 점에서 동일하다. 결정적으로 두 실험의 학습 시행의 인과 충분성은 다르지 않았다. 다시 말해서 두 실험 모두에서 공통 결과 구조 외에는 실험 참여자가 학습한 개별 시행을 설명할 인과 구조는 없다. 그러나 두 실험은 개별 학습 시행을 생성한 인과 구조, 즉 공통 결과 구조가 작동하는 방식에서 차이가 있다. 실험 1은 원인 변인이 자발적으로 활성화되지만 실험 2는 원인 변인이 개입(intervention)에 의해서만 활성화되었다.

실험 1과는 다르게 실험 2의 예측 추론에서는 대안 원인 무시 현상이 관찰되었다. 이 결과는 관찰되는 정보가 동일하더라도 그 정보가 어떤 방식으로 생성되었는지, 즉 정보들이 작동하는 방식이 인과 추론, 특히 대안 원인 무시 현상에 영향을 준다는 사실을 시사한다. 실험 2에서 원인 변인은 자발적으로 활성화되지 않는다. 다시 말해 실험 2의 인과 구조는 자발적 활성화가 불가능한 구조이다. 이러한 상황에서 판단자는 각 원인 변인의 기준 상태(default mode)를 비활성화 상태로 표상한다는 것이 우리의 해석이다.

우리는 동기적 측면에서 trial-by-trial 학습의 이점을 설명하였다. 자발적으로 작동하는 인과 구조를 수동적으로 관찰하는 관찰 학습(observation learning, 실험 1)에 비해, 작동을 시키고 싶은 변인을 선택 그 결과를 살펴보는 상태 개입 학습(intervention learning, 실험 2)은 더 많은 동기적 개입이 요구된다(Lagnado & Sloman, 2004, 2006). 실험 1의 결과와 trial-by-trial 학습에 따른 동기적 요소의 이점에서 발생하였다면 더 많은 동기적 개입이 필요

한 실험 2에서도 동일한 결과가 나와야 한다. 이러한 예측은 지지되지 않았다.

요약하면 본 실험의 결과는 인과 추론시 대안 원인이 표상되는 방식, 그리고 그 표상에 영향을 미치는 인과 구조의 작동 방식에 대한 이해가 대안 원인 무시 현상의 원인일 가능성을 강력히 시사한다.

종합 논의

본 연구의 첫 번째 목적은 관심이 되는 정보 이외의 다른 정보를 사용하여 판단과 예측을 할 수 있는지 검증하는 것이었다. 이를 위해 우리는 정보 무시 현상을 인과 베이스 모형의 틀에서 살펴보았다. 구체적으로 선행 연구에서 보고된 대안 원인 무시 현상이 인과 베이스 모형의 규범적 경계 조건이 확립된 상황에서도 관찰되는지 살펴보았다. 실험 1에서는 50회의 관찰학습으로 규범적 경계 조건을 만족시켰다. 50회의 관찰은 배타적으로 공통 결과 구조를 내포하고 있으며 참여자는 50회의 관찰을 통해 세 슬라이더와 관련된 파라미터를 학습할 수 있었다. 참여자들이 인과 구조 과제에서 세 변인만으로 이루어진 정확한 인과 구조를 보고한 점, 그리고 인과 추론이 학습 자료의 파라미터를 바탕으로 한 규범적 추론과 전반적으로 일치한 점은 인과 베이스 모형의 경계 조건이 성공적으로 조작되었음을 의미한다. 선행 연구와는 달리 인과 베이스 모형의 경계 조건이 갖춰진 상황에서 실험 참여자는 대안 원인 무시 현상을 보이지 않았다. 실험 참여자들은 예측 추론 및 진단 추론을 할 때 대안 원인을 고려하였다.

실험 1의 결과는 사람들이 예측 추론을 할 때 대안 원인을 고려할 능력이 있음을 보여주었다. 대안 원인을 고려하여 예측 추론을 실시하라는 지시를 명시적으로 받았을 경우 (Cummins, 2014; 윤민지, 2014)외에는 예측 추론을 할 때 실험 참여자들이 대안 원인을 고려한 연구는 보고되지 않았다. 본 연구는 대안 원인을 고려하라는 지시 없이도 사람들이 대안 원인을 고려하여 예측 추론을 할 수 있음을 보여주는 최초의 연구이다.

본 연구의 두 번째 목적은 왜 사람들은 주어진 정보 이외의 정보를 사용하지 못하는지를 살펴보는 것이었다. 이를 위해 우리는 인과 구조가 작동하는 방식을 조작하였다. 구체적으로 실험 2는 실험 1과 질적으로 동일한 공변 정보(covariation data)를 제공하였지만 그 정보가 작동하는 방식을 변화시켰다. 실험 1은 자발적으로 작동하는 인과 구조에 의해 생성된 공변 정보를 관찰하였다. 실험 2는 실험 1과 질적으로 동일한 공변 정보를 관찰하였지만 공변 정보를 관찰하기 위해서는 실험 참여자가 원인 변인의 상태를 개입하여야 했다. 실험 1과는 달리 상태 개입 학습 후에 이루어진 예측 추론에서는 대안 원인이 무시되었다.

개입 학습에서 대안 원인 무시 현상이 관찰된 실험 2의 결과에 대해 우리는 원인 변인의 기본 상태(default mode)가 비활성화된 상태로 표상되었을 가능성을 제안하였다. Fernbach와 그의 동료들(2011b)역시 비슷한 가능성을 제안하였다. Fernbach와 그의 동료들(2011b)은 대안 원인 무시 현상의 원인으로 기본 예측 질문의 표상에 관심을 가졌다. 사람들이 ‘관심 원인 변인이 존재할 때 결과 변인이 존재할 가능

성'을 묻는 기본 예측 질문을 '관심 원인 변인이 존재하지만 대안 원인 변인은 부재할 때, 결과 변인이 존재할 가능성'을 묻는 질문으로 해석한다는 것이다. Fernbach와 그의 동료들(2011b)은 이러한 가능성을 직접적으로 검증하지 않았다. 본 연구 결과는 Fernbach와 그의 동료들(2011b)의 제안에 대한 경험적 증거를 제공한다는 면에서 의의를 가질 수 있다.

인과 구조가 작동하는 방식에 대한 이해가 대안 원인 무시 현상과 관련이 있다는 본 연구 결과는 예측 추론과 진단 추론의 비대칭성과 일관적이다. Meder 등(2014)에 따르면 대안적 정보를 무시하는 예측 추론은 인과 구조가 작동하는 방식이 현저하지 않는 한 고려되지 않는다, 반면 대안적 정보를 고려하는 진단 추론은 주어진 인과 구조를 넘어 인과 구조가 작동하는 방식에 대한 추론이 필요하다. 대안 원인 무시 현상에 대한 선행 연구는 인과 구조를 언어적으로 제시하였다. 서론에서 언급한 예측 추론의 처리 용이성 가설은 원인에서 결과로의 추론에서 어떠한 인지적 과정이 생략되거나 왜곡되는지에 대한 설명은 제공하지 않는다. 본 연구 결과는 예측 추론에서는 인과 구조가 작동하는 방식에 대한 추론이 이루어지지 않음을 시사한다.

본 연구는 인과 베이스 모형의 기술적 타당성에 대해 긍정적인 함의를 지닌다. 본 연구 결과는 인과 베이스 모형의 경계 조건이 만족된다면, 인과 베이스 모형은 사람들의 인과 추론을 성공적으로 설명할 수 있음을 시사한다. 하지만 본 연구 결과는 부정적인 함의를 동시에 가진다. 특히 원인 사건에 대한 표상 방식에 대해 의문을 제기한다. 일반적으로

인과 베이스 모형은 원인 사건을 기저율을 가진 자동적으로 발생하는 사건으로 표상한다. 우리의 실험 결과는 사람들이 원인 사건을 개입이 되지 않는 한 활성화되지 않는 사건으로 표상할 가능성을 시사한다.

인과 관계가 개입에 의해 발생한다는 가설은 사람들의 일상적 직관과 부합한다. 먼저 우리는 하나의 변인을 조작했을 때 다른 변수는 상수로 가정하는 경우가 많다. 이러한 가정은 사회 과학에서 이루어지는 실험의 기본적인 가정이기도 하다. 또한 일상 생활에서 많은 사건들은 개입에 의해서만 발생한다. 흡연을 0.5의 확률로 우연히 발생하는 사건으로 표상하는 사람은 없을 것이다. 흡연은 흡연자의 선택과 개입으로 이루어지는 사건으로 표상된다. 상태 개입에 의한 추론이 인지적으로 부담스러운 과제임에도 불구하고 정확률이 높은 선행 연구의 결과(Meder, et al., 2008, 2009)는 사람들의 인과관계를 개입에 의해 발생하는 사건으로 표상하기 때문일 가능성도 있다. 원인 변인의 활성화를 어떻게 표상해야 할지, 원인 변인이 활성화 되었을 때 다른 원인 변인이 어떻게 표상될 지에 대한 후속 연구가 필요할 것이다.

또한 본 연구는 예측의 정확성을 향상시킬 수 있는 실질적인 방안을 제시할 수 있다는 점에서 의의가 있다. 먼저 예측 추론에서 관찰되는 대안 원인 무시 현상은 진단 추론을 함께 실시함으로써 약화될 수 있을 것이다. 예를 들어 "새로운 고속도로를 건설한다면 교통량은 분산될 것인가?"와 같은 예측 추론을 할 때, 교통량이 분산이 "새로운 고속도로 건설 때문에 발생할 가능성은 얼마일까?"와 같

은 진단 추론을 함께 한다면 보다 다양한 정보를 고려한 판단이 가능할 것이다. 그리고 각 원인 사건들을 개입에 의해 발생하는 사건(실험 2)으로 표상하기 보다는 자연적으로 발생하는 사건(실험 1)으로 표상하는 것이 예측의 정확성을 높이는 방법이다. 예를 들어 교통량을 줄일 수 있는 원인들을 고려할 때 인간의 개입에 의해서만 발생하는 사건들만을 고려하는 것이 아니라 기후나 경제 상황 등 직접적으로 개입할 수 없는 사건들을 고려한다면 예측의 정확성은 향상될 것이다.

마지막으로 본 연구는 대안 원인 무시 현상에 대한 후속 연구의 방향을 제안했다는 면에서 의의가 있다. 본 연구는 trial-by-trial 학습 그리고 상태 개입 학습을 통해 인과 구조가 작동하는 방식을 현저하게 만들었다. 하지만 이러한 학습 방식이 현실에 적용되기에는 제약이 있다. 따라서 판단자가 인과 구조가 작동하는 방식을 고려할 수 있게 만드는 방안들에 대한 실증적 연구가 필요할 것이다. 예를 들어 Fernbach와 그의 동료들(2011b)은 기저율에 대한 질문이 진단 추론과 같이 인과 구조가 작동하는 방식을 환기시킨다고 제안하였다. 대안 원인 무시 현상에 사용된 실험 패더다임을 적용해 이러한 주장을 검증하는 연구가 가능할 것이다.

본 연구의 제한점으로 인과 추론 질문에서 대안 원인을 언급했다는 것이다. 본 연구에서는 대안 원인의 상태를 ‘알 수 없다’라고 언급하였다. 이러한 언급은 대안 원인을 현저하게 만들었으며 결과적으로 예측 추론에 대안 원인이 반영되었다는 것이다(실험 1). Fernbach와 Rehder(2013)는 예측 추론에서 대안 원인의 상

태를 언급하였지만 대안 원인 무시 현상은 관찰되지 않았다. 반면 윤민지(2014)는 예측 추론 질문에서 대안 원인 상태를 언급하는 것이 대안 원인 무시 현상을 없앨 수 있음을 보여주었다. 하지만 윤민지(2014)의 경우 대안 원인의 상태를 ‘존재할 수 있다’라고 기술하였다. 즉 본 연구와는 다르게 대안 원인이 활성화되었을 가능성을 시사한 것이다. 또한 만일 실험 1에서 관찰된 대안 원인 존중 현상이 질문의 효과라면 동일한 질문을 사용한 실험 2에서도 대안 원인 존중 현상이 관찰되어야 한다. 하지만 실험 2에서는 대안 원인 무시 현상이 관찰되었다. 하지만 일상생활에서 대안 원인을 언급하는 경우는 많지 않다. 대안 원인에 대한 언급이 없는 질문을 사용하여 생태학적 타당성을 높일 필요가 있다.

본 연구는 인과 베이지 모형의 규범적 경계 조건이 만족된 상황에서 대안 원인 무시 현상이 관찰되는지 검증하기 위해 2개의 실험을 실시하였다. 본 실험 결과는 사람들은 대안 원인을 고려할 수 있는 능력이 있음을 보이고 있다. 특히 본 연구는 대안 원인 무시 현상은 인과 구조가 작동하는 방식에 대한 이해와 관련이 있음을 보여주었다. 본 연구는 인과 구조가 작동하는 방식에 대한 인지적 과정에 대한 이해를 기반으로 예측의 정확성을 높이는 구체적인 후속 연구가 필요할 것이다.

참고문헌

- 윤민지 (2014). 예측 추론과 진단 추론의 정보처리 특성. 성균관대학교 석사학위 논문
한겨레신문 (2014. 11. 17). ‘지진’ 예측 실패한

- 과학자는 유죄, 무죄?
경향신문 (2014. 10. 20). 고속도로 교통량 예측 실패... 통행료로 건설비 원금도 못 갚아.
- Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, 136, 569-576.
- Anderson, J. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Barbey, A. K. & Sloman, S. A. (2007). Base-rate respect: From statistical formats to cognitive structures. *Behavioral and Brain Sciences*, 30, 287-298.
- Baron, J. (2000). *Thinking and deciding*. Cambridge University Press.
- Bonini, N., Tentori, K., & Osherson, D. (2004). A different conjunction fallacy. *Mind & Language*, 19(2), 199-210.
- Cartwright, N. (2004). Causation: One word, many things. *Philosophy of Science*, 71, 805-81.
- Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, 3, 57-65.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104, 367-405.
- Cummins, D. D. (2014). The impact of disablers on predictive inference. *Journal of experimental psychology: learning, memory, and cognition*, 40(6), 1638-1655.
- Dawes, R. M. (1979). The robust beauty of improper linear models. *American Psychologist*, 34, 571-582.
- Doherty, M. E., Chadwick, R., Caravan, H., Barr, D. & Mynatt, C. R. (1996). On people's understanding of the diagnostic implications of probabilistic data. *Memory & Cognition*, 24, 644-654.
- Doherty, M. E., Mynatt, C. R., Tweeney, R. D., & Schiavo, M. D. (1979). On pseudodiagnosticity. *Acta Psychologica*, 43, 111-121.
- Fernbach, P. M., & Sloman, S. A. (2009). Causal Learning with Local Computations, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 678-693.
- Fernbach, P. M., & Rehder, B. (2013). Cognitive shortcuts in causal inference. *Argument & Computation*, 4(1), 64-88.
- Fernbach, P. M., Darlow, A., & Sloman, S. A. (2010). Neglect of Alternative Causes in Predictive But not Diagnostic Reasoning, *Psychological Science*, 21, 329-336.
- Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011a). Asymmetries in Predictive and Diagnostic Reasoning, *Journal of Experimental Psychology: General*, 140, 168-185.
- Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011b). When Good Evidence Goes Bad: The Weak Evidence Effect in Judgment and Decision-Making. *Cognition*, 119, 459-467.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental*

- Psychology: Human perception and performance*, 3(4), 552-564.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650-669.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62, 451-482.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. *Psychological review*, 98(4), 506.
- Gilovich, T., & Griffin, D. (2002). Heuristics and biases: Then and now. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 230-249). Cambridge, England: Cambridge University Press.
- Glymour, C. (1998). Learning causes: Psychological explanations of causal explanation. *Minds and Machines*, 8, 39-60.
- Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and Strength in Causal Induction, *Cognitive Psychology*, 51, 334-384.
- Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-Based Causal Induction, *Psychological Review*, 116, 661-716.
- Hadjichristidis, C., Sloman, S. A., & Over, D. E. (2009). *Categorical induction from uncertain premises: Jeffrey's (doesn't) rule*. Manuscript submitted for publication.
- Hattori, M., & Oaksford, M. (2007). Adaptive non interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science*, 31(5), 765-814.
- Hastie, R., & Dawes, R. M. (Eds.). (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Sage.
- Hausman, D. M., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *British Journal for the Philosophy of Science*, 50, 521-583.
- Hertwig, R., & Gigerenzer, G. (1999). The conjunction fallacy revisited: how intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12, 275-306.
- Holyoak, K. J. & Cheng, P. W. (2011). Causal Learning and Inference as a Rational Process, *Annual Review of Psychology*, 62, 135-163.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases. Judgment*. Cambridge University Press.
- Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211-228.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human learning and memory*, 6(2), 107-118.
- Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General*, 136(3), 430-450.
- Kushnir, T., Gopnik, A., Lucas, C., and Schulz,

- L. (2010). Inferring hidden causal structure. *Cognitive Science*, 34(1), 148-160.
- Lagnado, D., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 856-876.
- Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(3), 451-460.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098-2109.
- Lucas, C. G., & Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical Bayesian models. *Cognitive Science*, 34(1), 113-147.
- Luhmann, C. & Ahn, W. (2007). BUCKLE: A model of unobserved cause learning. *Psychological Review*, 92(3), 657-677.
- Mayrhofer, R., & Waldmann, M. R. (2014). Agents and causes: Dispositional intuitions as a guide to causal structure. *Cognitive science*, 1-31.
- Mayrhofer, R., Goodman, N. D., Waldmann, M. R., & Tenenbaum, J. B. (2008). Structured correlation from the causal background. In B. C. Love, K. McRae, V. M. Sloutsky, (Eds.) *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp.303-308). Austin, TX: Cognitive Science Society.
- Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological review*, 121(3), 277-301.
- Meder, B., Hagmayer, Y., & Waldmann, M. R. (2008). Inferring interventional predictions from observational learning data. *Psychonomic Bulletin & Review*, 15(1), 75-80.
- Meder, B., Hagmayer, Y., & Waldmann, M. R. (2009). The role of learning data in causal reasoning about observations and interventions. *Memory & Cognition*, 37, 249-264.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of consulting psychology*, 29(3), 261-265
- Park, J., & Sloman, S. A. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology*, 67, 186-216.
- Park, J., & Sloman, S. A. (2014). Causal explanation in the face of contradiction. *Memory & Cognition*, 42(5), 806-820.
- Pearl, J. (2000). *Causality: models, reasoning and inference*. Cambridge: MIT press.
- Ross, B. H., & Murphy, G. L. (1996). Category-based predictions: Influence of uncertainty and feature associations. *Journal of Experimental Psychology: Learning, Memory*, 22, 736-745.
- Scheines, R., Spirtes, P., Glymour, C., & Meek, C. (1994). *TETRAD II*. Hillsdale, NJ: Erlbaum.
- Shaklee, H., & Fischhoff, B. (1982). Strategies of Information Search in Causal Analysis. *Memory & Cognition*, 10, 520-530.
- Skowronski, J. J., & Carlston, D. E. (1989).

- Negativity and extremity biases in impression formation: A review of explanations. *Psychological bulletin*, 105(1), 131-142.
- Sloman, S. (2009). *Causal models: How people think about the world and its alternatives*. Oxford University Press.
- Sloman, S. A., Lagnado D. (2015). Causality in Thought. *Annual Review of Psychology*, 66(3), 1-25.
- Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search*. New York: Springer-Verlag.
- Tversky, A., & Kahneman, D. (1980). Causal schemata in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (pp.49-72). Hillsdale, NJ: Erlbaum.
- Walsh, C. & Sloman, S. A. (2008). Updating beliefs with causal models: Violations of screening off. Gluck, M. A., Anderson, J. R., & Kosslyn, S. M. (Eds.). *Memory and Mind: A Festschrift for Gordon H. Bower*. New Jersey: Lawrence Erlbaum Associates.
- Walsh, C. R., & Sloman, S. A. (2004). Revising causal beliefs. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly journal of experimental psychology*, 12(3), 129-140.
- Waldmann, M. R. (2007). Combining versus analyzing multiple causes: How domain assumptions and task context affect integration rules. *Cognitive Science*, 31, 233-256.

1 차원고접수 : 2015. 03. 27

수정원고접수 : 2015. 04. 21

최종게재결정 : 2015. 04. 27

The Effect of a Functional form of Causal relation on Accurate Causal Reasoning

Juhwa Park

Yonsei University

Seung-Lee Do

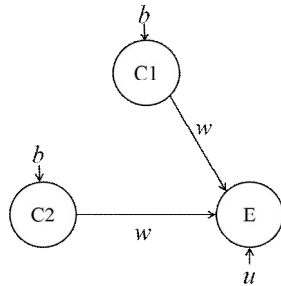
SungKyunkwan university

The success of an individual, a business and even a country depends on the accuracy of judgments and predictions about the future. Reasoners, however, often fail to search for additional pieces of information to make an accurate prediction. Moreover, they tend to neglect a relevant piece of information even if it is available. Thus, it is critical to understand the cognitive process behind the neglect phenomenon in order to increase the accuracy of predictions. The present study aimed to investigate the process that underlies neglect of relevant information based on a newly emerging normative model, the causal Bayes nets. In Experiment 1, we constructed a situation where the three boundary conditions of the causal Bayes nets were satisfied: the sufficiency of a causal structure, parameters, and the clarity of the functional form of causal structure. The results revealed that participants did not neglect a relevant piece of information when they made a predictive inference. Experiment 2 was identical to Experiment 1, except for the functional form of causal structure. The causal structure in Experiment 1 was activated by itself while the structure in Experiment 2 was activated only when it was intervened. With the intervention-based causal structure, participants showed neglect of a relevant piece of information to make a predictive inference. Our results suggest that neglect of a functional form of events in question is responsible for the failure to consider relevant information. Our results also suggest that the assumptions that the causal Bayes nets make about representations of causal relations may not be consistent with the actual representations of reasoners. Our results suggest that a future study is necessary to investigate how understanding of the functional form of causal structures influences prediction accuracy.

Key words : The causal Bayes nets, neglect of relevant information, predictions

부록 1

공통 결과 구조에 대한 규범적 인과 추론치는 아래의 확률적 공통 결과 구조를 통해 결합 확률 (Joint probability)를 계산하고 조건 확률을 구하는 방식으로 생성되었다.



- b : 기저율
- w : 원인 변인이 단독으로 결과를 일으킬 확률
- u : 원인 변인이 없을 때 결과 변인이 발생할 확률

실험 1은 $b=0.5$, $w=0.75$ 그리고 $u=0.2$ 실험 2는 $b=0.5$, $w=0.75$ 그리고 $u=0$ 의 파라미터와 표 1을 사용하여 결합 확률을 계산하였다.

표 1. 실험 1에 사용한 50회의 빈도 및 결합확률 계산에 사용한 공식

슬라이더			빈도	결합확률
C1	C2	E		
1	1	1	11	$b \times b \times (1 - (1 - w) \times (1 - w) \times (1 - u))$
1	1	0	1	$b \times b \times (1 - w) \times (1 - w) \times (1 - u)$
1	0	1	10	$b \times (1 - b) \times (1 - (1 - w) \times (1 - u))$
1	0	0	3	$b \times (1 - b) \times (1 - w) \times (1 - u)$
0	1	1	10	$(1 - b) \times b \times (1 - (1 - w) \times (1 - u))$
0	1	0	3	$(1 - b) \times b \times (1 - w) \times (1 - u)$
0	0	1	3	$(1 - b) \times (1 - b) \times u$
0	0	0	9	$(1 - b) \times (1 - b) \times (1 - u)$

부록 2. 실험 2에서 인과 충분성 검증에 사용한 절차

Bayes rule에 따르면 어떤 데이터 (d)를 관찰한 후 가능한 가설들(h)의 사후 확률(posterior probability)은

$$P(h_i|d) = \frac{P(d|h_i) \cdot P(h_i)}{\sum_{h \in H} P(d|h) \cdot P(h)}$$

로 정의된다.

P(h)는 사전 확률 (prior probability)로 데이터를 보기 전 각 가설이 얼마나 그럴듯한지에 대한 믿음을 반영하며, P(d|h)는 관찰된 데이터가 각 가설에 의해 설명되는 정도를 나타낸다. 분모는 모든 가설에 대한 표준화를 위한 것이다. 세 변인으로 부터 가능한 24개의 인과 구조 (세 변인이 독립적인 경우는 제외) 를 가능한 가설로 설정하였다. 실험 참여자의 각각의 상태-강제 시행을 데이터로 간주하면 각 가설의 사후 확률은

$$P(h_i | d_1, d_2, d_3, \dots, d_n) = \frac{P(h_i) \cdot \prod_{d \in D} P(d|h_i)}{\sum_{h \in H} P(h) \prod_{d \in D} P(d|h)}$$

로 정의된다.

사전 확률은 아래에 제시된 Fernbach 와 Sloman (2009)의 공식을 사용하여 할당하였다.

$$P(h_i) = \frac{\theta^{l_i}}{\sum_{i=1}^{24} \theta^{l_i}}$$

P(h_i)는 i번째 구조의 사전 확률을 뜻하며, l_i는 해당 구조의 화살표의 수를 나타내며, θ는 복잡하거나 간단한 구조에 대한 편향을 반영하는 수치이다. 분모는 표준화를 위한 것이다. θ가 1보다 크면 화살표가 적은 구조에 높은 사전 확률 (단순성 편향)을 할당하고, θ가 1보다 작으면 화살표의 수가 많은 구조에 높은 사전 확률을 할당한다. θ가 1이면 24개 구조의 사전확률은 동일하다.

θ를 1에서 100까지, 모두 100번의 사후 확률을 계산하였다. 24개의 구조들 중 사후 확률이 가장 높은 구조를 데이터가 제안하는 구조로 선택하였다. 모든 실험 참여자의 학습 시행에 대해 모든 100번의 시행에서 베이즈 추론 알고리즘은 공통 인과 구조와 인과 연쇄 구조를 각각의 학습 시행에서 인과적으로 충분한 구조로 선택하였다. 표 2는 실험 2에서 θ가 1과 100일때의 사후 확률의 평균을 제시하였다.

표 2. θ가 1과 100일 때의 사후 확률의 평균

θ	공통 결과 구조
1	.94
100	.99