

인과적 사슬구조에서의 범주기반 속성추론

Category-based Feature Inference in Causal Chain

최인범¹ · 이형철² · 김신우^{3*}

InBeom Choi¹ · Hyung-Chul O. Li² · ShinWoo Kim^{3*}

Abstract

Concepts and categories offer the basis for inference pertaining to unobserved features. Prior research on category-based induction that used blank properties has suggested that similarity between categories and features explains feature inference (Rips, 1975; Osherson et al., 1990). However, it was shown by later research that prior knowledge had a large influence on category-based inference and cases were reported where similarity effects completely disappeared. Thus, this study tested category-based feature inference when features are connected in a causal chain and proposed a feature inference model that predicts participants' inference ratings. Each participant learned a category with four features connected in a causal chain and then performed feature inference tasks for an unobserved feature in various exemplars of the category. The results revealed nonindependence, that is, the features not only linked directly to the target feature but also to those screened-off by other feature nodes and affected feature inference (a violation of the causal Markov condition). Feature inference model of causal model theory (Sloman, 2005) explained nonindependence by predicting the effects of directly linked features and indirectly related features. Indirect features equally affected participants' inference regardless of causal distance, and the model predicted smaller effects regarding causally distant features.

Key words: Causal Reasoning, Category-Based Feature Inference, Causal Markov Condition, Causal Model Theory, Causal Chain

요약

개념과 범주는 관찰하지 못한 속성을 추론할 수 있는 기반을 제공한다. 무의미 속성을 사용한 범주기반 속성추론 연구들은 범주 및 속성의 유사성이 추론을 설명하는 핵심 요인이라는 것을 제안했다(Rips, 1975; Osherson et al., 1990). 이후 연구들은 사람들의 사전지식이 범주기반 추론에 막대한 영향을 미치며 심지어 유사성 효과가 완전히 사라지는 경우도 있음을 보고했다. 본 연구는 범주 속성들이 사전지식의 한 종류인 인과적 지식에 의해 사슬구조로 연결되었을 때의 범주기반 속성추론을 검증했으며 그 결과를 예측하는 속성추론모형을 제안했다. 참가자들은 네 개의 속성들이 사슬구조를 이루는 인과적 범주를 학습한 뒤 해당 범주의 다양한 범주 예시들의 숨겨진 속성에 대한 추론을 실시했다. 그 결과 인과적으로 직접 연결된 속성뿐만 아니라 다른 속성 노드에 의해 차폐된 속성들도 추론에 영향을 미치는 비독립성이 나타났다(인과적 마코프 조건의 위배). 인과모형이론(Sloman, 2005)에 기반한 속성추론모형을 적용하여 참가자들의 추론을 모델링한 결과 인과적 연결의 직접 효과뿐만 아니라 간접 효과 즉 인과추론의 비독립성도 예측하는 것으로 나타났다. 다만 간접적으로 연결된 속성들은 인과적 거리와 무관하게 참가자들의 추론

* 이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2018S1A5A2A01037248).

¹ 최인범: 광운대학교 산업심리학과 석사졸업

² 이형철: 광운대학교 산업심리학과 교수

^{3*} (교신저자) 김신우: 광운대학교 산업심리학과 부교수 / E-mail: shinwoo.kim@kw.ac.kr / TEL: 02-940-5421

평정에 동일하게 영향을 미쳤지만 모형은 거리가 멀어짐에 따라 추론에 미치는 영향이 작아짐을 예측했다.

주제어: 인과적 추론, 범주기반 속성추론, 인과적 마코프 조건, 인과모형 이론, 인과적 사슬

1. 서론

개념과 범주는 인간 사고의 기본 단위이다. 선언적 지식의 이해와 학습, 절차적 지식의 초기표상, 기억, 귀납 및 연역 추론, 일상적 의사결정, 법률적 판단, 분류체계의 확립 등을 포함하는 모든 고차 인지과정에서 개념과 범주는 중추적인 역할을 담당한다(e.g., Murphy, 2002). 인지과정에서 범주의 중요성은 특히 관찰하지 못한 속성을 추론할 수 있게 해준다는 점에서 두드러진다. 처음 보는 눈앞의 대상이 동물 혹은 식물이라는 것을 아는 것은 그 대상의 속성에 대해 막대한 정보를 제공한다. 그 대상이 포유류라는 것을 아는 것은 추가적인 정보를 제공해주며, 박쥐라는 것을 알게 되면 더 많은 것을 직접 경험하지 않고도 추측할 수 있게 된다. 손목 관절이 아플 때 피부과가 아니라 정형외과를 찾는 것도 범주 기반 사고에 따른 선택이다. 이러한 추론은 범주의 구조적 특성에 의해 제한되기도 하고 어떤 경우에는 추론의 강한 근거가 되기도 한다(Patalano & Ross, 2007; Thagard & Nisbett, 1982). 강릉에서 측정된 금속 K의 녹는 점이 250°C라면 뉴욕에서 발견한 금속 K도 같은 녹는점을 가질 것이라고 추론할 수 있다. 반면 유학시절 만났던 영국인 룸메이트가 김치 볶음밥을 좋아했다고 해서 내일 만날 영국인 바이어도 같은 식성을 가질 것이라 장담하기는 어렵다.

범주기반 속성추론은 다양한 방식으로 발생할 수 있다. 가령 독일산 무선 키보드가 튼튼했던 경험을 통해 독일산 마우스도 튼튼할 것이라 추론할 수 있다. 나아가 독일산 전자제품 전체의 내구성이 높을 것이라 결론 내리기도 하고, 독일을 포함하는 서유럽 제품은 역시 믿을만 하다고 일반화할 수 있게 될지도 모른다. 혹은 어떤 공산품이 독일산이라는 것을 알면 그 제품은 튼튼할 것이라고 선형적으로 추측하게 될 수도 있다. 즉 특정 범주에 대한 경험은 그 범주뿐만 아니라 인접 범주 혹은 그 범주를 포함하는 상위 범주에 대한 추론도 가능하게 해준다. 본 연구에서는 특정 범주에 대한 경험이나 지식이 그 범주에 속하는 다양한 대상의 속성

에 대한 추론을 검증하고 모델링하고자 한다.

범주기반 속성추론에 대한 기존 연구들은 다양한 흥미로운 결과들을 보고했다. 초기 연구들은 추론에서 사전 지식의 효과를 배제하기 위해 무의미 속성(blank properties)을 사용하여 속성추론에 영향을 미치는 범주의 영향을 검증했다. 대표적인 결과로 전형성 효과(typicality effect)를 들 수 있는데, 가령 전형적인 새인 참새가 sesamoid 뼈를 가진다고 할 때 다른 새들도 sesamoid 뼈를 가진다고 추론할 개연성이 높지만, 비전형적인 타조가 sesamoid 뼈를 가진다고 해서 다른 새들도 같은 뼈를 가진다고 생각하기는 어렵다는 것이다. 즉, 전형적인 범주일수록 그 특성이 다른 범주 혹은 상위범주로 전이될 가능성이 높다는 것이다. 이때 사용한 sesamoid 뼈는 무의미 속성으로 사람들이 무엇인지 모르기 때문에 사전 지식에 기반한 추론이 거의 불가능하다. 아울러 다양성 효과(diversity effect)도 발견되었는데 이는 추론의 기반이 되는 둘 이상의 범주들이 다양할수록 결론에 대한 확신이 높아진다는 것이다(Feeney & Heit, 2011; Heit & Hahn, 2001). 예를 들어, 참새와 박새가 sesamoid 뼈를 가질 때 보다 참새와 타조가 sesamoid 뼈를 가지는 경우에 다른 새 혹은 새 범주 전체에 대해 sesamoid 뼈를 일반화하는 경향이 더 강하다는 것이다. 다양성 효과는 규범적 추론의 한 예로 표본검사를 할 때 가능한 다른 시간, 위치 등에서 추출된 샘플일수록 전체를 대표할 가능성이 높을 것이라는 일반적인 믿음과 일치한다(Kim & Keil, 2003; Heit et al., 2005; Nagel, 1939).

범주기반 속성추론의 실험결과들을 설명하기 위한 주요 이론들이 제안되었는데 이들은 모두 Rips(1975)가 제안한 유사성에 근거하여 추론을 설명한다. 먼저 유사성-범위 모형(similarity coverage model; Osherson et al., 1990)에 따르면 전제 범주의 다양한 범주 예시들을 장기기억의 가용성에 따라 인출하고 이들을 표적 범주와의 최대 유사성에 근거하여 속성추론이 이루어진다고 제안한다. 즉, 전제 범주의 속성들이 표적 범주의 속성들을 포괄하는 정도에 따라 추론 강도가 결정

된다는 것이다. 이는 전제 범주의 전형성이 높거나 다양할수록 표적 범주를 더 많이 포괄할 수 있기 때문에 전형성 및 다양성 효과를 설명하는 근거가 된다. 또한 위계적인 관련성이 없는 범주들간의 속성추론은 단순 유사성에 의해 결정된다고 제안한다. 범주와 범주 속성 간 유사성이 속성추론을 결정한다고 제안한 유사성-범위 모형은 무의미 속성에 대한 범주기반 추론에 대해 높은 설명력을 가진 이론으로 여겨지고 있다. 또 다른 이론인 속성-기반 추론모형(feature-based induction model; Sloman, 1993)도 마찬가지로 유사성에 기반하여 설명을 전개하는데 속성수준에서 연결주의 모형을 사용했다는 점에서 Osherson et al.(1990)과 차이가 있을 뿐이다. 인지적, 계산적 처리의 관점에서 두 모형은 서로 차이를 보이지만 결국 유사성 변인을 통해 속성추론을 설명했다는 점에서 동일한 종류의 모형으로 분류할 수 있다. 최근 제안된 Osta-Vélez & Gärdenforsd(2020)의 모형도 마찬가지로 서로 다른 범주들의 여러 속성간 유사성과 이들의 가중치를 토대로 계산한 거리를 추론의 근거로 가정한다는 점에서 유사성 기반 접근에 속한다.

유사성을 설명변수로 둔 속성추론 모형들은 꽤 성공적이었지만 무의미 속성을 대상으로 했다는 점에서 제한을 가진다. 무의미 속성은 사전지식의 영향을 통제하고 범주 자체의 영향을 검증할 수 있다는 장점은 존재한다. 그러나 이후 연구들은 범주와 범주 속성에 대한 사람들의 지식이 다양한 방식으로 속성 추론에 영향을 미친다는 것을 보고했다(Bailenson et al., 2002; Coley et al., 2005; Coley & Vasilyeva, 2010; Doh & Lee, 2020; Heit & Rubinstein, 1994; Lee et al., 2014; Proffitt et al., 2000; Rehder, 2006; Shafto et al., 2007). 예를 들어, Heit & Rubinstein(1994)는 “두개의 공실을 가진 간(liver with two chambers)”과 같은 해부학적 속성은 전제범주가 참치일 때보다 곰일 때 고래에게 더 잘 전이됐는데 이는 곰과 고래가 모두 포유류이기 때문이다. 반면 “지그재그로 움직인다(travels in a zigzag path)”와 같은 행동적인 속성은 오히려 동일한 서식지를 공유하는 참치와 고래에서 전이가 잘 된다는 것을 보고했다. 즉, 범주와 속성에 대한 사전 지식은 속성추론의 패턴을 뒤집을 만큼 강했던 것이다. 또 다른 예로 Sloman(1994, 1997)은 “많은 전과자들은 보디가드로

채용된다”라는 전제를 보고 참가자들은 “많은 참전군인들은 보디가드로 채용된다”라는 결론이 타당하다고 반응하지만 “많은 전과자들은 무직이다”라는 전제를 제시했을 때 “많은 참전군인들은 무직이다”라는 결론에는 크게 동의하지 않는다는 것을 보고했다.

위의 예시들은 범주와 속성에 대한 참가자들의 인과적 지식에 대한 적용이 범주기반 추론에 중요한 영향을 미친다는 것을 보여준다. 곰과 고래가 포유류라는 것은 유사한 해부학적 특징을 설명하는 원인이 되며, 참치와 고래가 서식지와 먹이사슬을 공유한다는 점은 이들이 헤엄치는 방식을 인과적으로 설명하게 된다. 마찬가지로 전과자와 참전군인이 공유하는 거친 특성은 이들이 보디가드로 적합함을 설명하지만, 전과자가 무직인 이유는 참전 군인이 무직이라는 것을 설명하지 못한다. 왜냐하면 무직이라는 결과를 설명하는 인과적 원인(전과경력)이 후자에는 부재하기 때문이다.

속성추론에서 인과적 설명이 유사성 요인 외에도 중요한 역할을 한다는 것은 직관적으로도 명확해보인다. 앞선 예시에서 유학시절 만났던 영국인 룸메이트가 김치 볶음밥을 좋아했다는 전제에 그 룸메이트의 배우자가 한국인이라는 지식을 덧붙이면, 내일 만날 영국인 바이어도 같은 식성을 가질 것이라는 생각은 거의 할 수 없을 것이다. 한국인 배우자를 둔 것이 원인이 되어 김치 볶음밥을 좋아했다는 결과를 추론할 수 있기 때문이다. 사전 지식 혹은 인과적 설명이라는 요인이 속성추론에 강한 영향을 미친다는 것은 명확한 사실이다. 더 중요한 것은 특정한 인과적 지식이 범주기반 속성추론에 어떤 방식으로 영향을 미치는지 검증하는 것이 될 것이다.

기존 연구에서 범주 속성들이 다양한 방식의 인과관계로 연결되었을 때 속성추론에 미치는 영향을 검증한 바 있다(Kim & Li, 2017; Lee & Park, 2016; Rehder & Burnett, 2005; Rehder & Kim, 2009). 예를 들어 범주 속성들이 공통원인(common cause)의 인과관계를 가질 때 인과적 마코프 조건(causal Markov condition)에 대한 위배가 나타나며 공통효과(common effect)의 인과구조에서는 결과속성에 대한 인과적 절감(causal discounting)이 관찰되었다. 특히 인과모형이론(causal model theory; Sloman, 2005; Waldmann & Holyoak, 1992)을 적용한 속성추론 모형은 인과적 범주의 속성

추론에서 전형성 효과, 비독립성, 인과적 절감을 모두 설명한다는 것을 보여주었다(Kim & Li, 2017).

본 연구에서는 범주내 속성들이 인과적 사슬(causal chain)로 연결되었 때의 범주기반 속성추론을 실험하고 이를 설명하는 모형을 제안한다. Fig. 1은 F1, F2, F3, F4의 네 속성으로 이루어진 범주와 범주속성 간의 인과관계를 보여준다. F1은 F2의 원인이며, F2는 F3의 원인이며, F3는 F4의 원인이 된다. 가령 전과자 범주에서 충동성은 범죄행위의 원인이며, 범죄행위에 의해 전과기록이 생기고, 전과기록은 미취업을 낳게된다는 속성의 인과적 연쇄가 가능하다. 인과적 연쇄는 그 자체로써 연구자들의 흥미를 유발했으나 분석 방법에 대해서는 논란이 있어왔다(Ahn et al., 2000; Hadjichristidis et al., 2004; Marsh & Ahn, 2006; Mayrhofer & Rothe, 2012; Rehder & Kim, 2010). 가령 인과위계가설(causal status hypothesis; Sloman et al., 1998)에서는 인과적 사슬에서 원인이 되는 속성이 가장 핵심적인 지위를 차지하며 인과적으로 멀어져 말단으로 갈수록 그 중요성이 낮아진다고 제안한다. 이에 따르면 원인의 결과 속성들은 원인에 종속되는 종속성(dependency)이 있으며 종속성에 따라 개별 속성들의 중심성(centrality)이 결정된다고 제안한다(Ahn et al., 2017). 이러한 해석은 결국 인과적 사슬구조에서 속성들의 중심성 분석이 해석의 기초를 이룬다는 결론을 낳게된다. 반면 인과모형이론(causal model theory)에 따르면 범주속성들은 인과관계에 따라 서로 다른 가중치를 가지며 동시에 속성들 간의 인과성에 따른 상호작용 효과도 독립적으로 존재한다고 제안한다(Fenker et al., 2005; Rehder, 2003; Rehder & Kim, 2010; Waldmann & Holyoak, 1992).

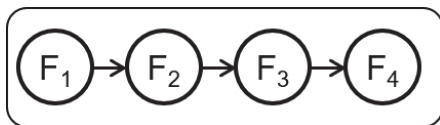


Fig. 1. Four-feature causal chain

본 연구에서는 인과모형이론에 근거한 속성추론 모형이 인과적 사슬구조에서의 추론패턴을 타당하게 예측하는지 확인하고자 한다. 예를 들어, Fig. 1과 같은 인과적 사슬구조에서 F3에 대한 추론은 F1의 존재여부와 관련이 없다. 왜냐하면 F1과 F3를 매개하는 F2가

차폐(screen-off)하기 때문에 규범적(rational) 추론자는 F3를 추론할 때 F1은 고려하지 않아야 한다. 실험 결과 이러한 비독립성 즉 인과적 마코프 조건에 대한 위배가 나타났으며 인과모형이론에 근거한 속성추론 모형이 범주기반 속성추론에서 이러한 비독립성을 설명하는 것으로 나타났다.

본 연구와 같이 인과적 사슬구조 혹은 그에 상응하는 인과구조에 대한 기존 연구들이 존재한다(Ahn et al., 2017; Kim et al., 2009; Rehder, 2006; Rehder, 2009; Rehder & Burnett, 2005; Rehder & Hastie, 2004). 이 연구들이 사용한 범주와 인과적 지식이 유사한 경우도 있었으나 검증 내용에는 명확한 차이가 존재한다. Ahn et al.(2017)과 Kim et al.(2009)은 정신질환 범주의 진단, 치료에서 인과위계이론이 제안하는 중심성(centrality)이 핵심역할을 한다는 것을 보여주었다. Rehder(2006)은 유사성과 인과관계가 속성추론에서 배타적으로 사용된다는 것을 보여준 것이며, Rehder & Burnett(2005)은 본 연구와 유사한 사슬구조를 포함하는 인과적 범주에서 속성추론을 검증했으나 추론의 비독립성을 심리적 본질주의(psychological essentialism)에 근거하여 개념적 모형으로 설명하였으며 계산적 모델링은 시도하지 않았다. Rehder & Hastie(2004)는 범주 응집성(coherence)이 속성추론에 미치는 영향을 분석했으며, Rehder(2009)는 특정 범주예시가 아닌 전체 범주에 대한 일반화를 검증한 결과를 보고하였다. 본 연구에서는 범주 속성들이 인과적 사슬로 연결되었을 때 개별 범주 예시의 속성에 대한 추론과 이에 대한 모델링 결과를 보고할 것이다.

2. 실험

2.1. 방법

2.1.1. 실험자극 및 설계

실험을 위해 키호 개미(생물 범주), 미야 별(자연물 범주), 넷툰 컴퓨터(인공물 범주)의 세 범주를 사용했다(Rehder & Kim, 2006; Rehder & Kim 2009). 각 범주에는 네 개의 속성차원이 있었으며 각각은 전형적이거나 비전형적인 두 개의 속성값을 가질 수 있었다. 표

1은 미야 별을 예시로 한 속성차원과 속성값이다. 전형적인 미야 별은 헬륨이 이온화되었으며, 온도가 매우 높고, 밀도가 높으며, 행성의 수가 매우 많다. 전형적인 속성과 비전형적인 속성의 기저율은 각각 ‘대부분’과 ‘일부’로 제시했다. 전형적인 속성과 비전형적인 속성은 각각 F와 ~F로 표시할 것이다. 따라서, F2는 두번째 속성이 전형적이라는 것을 의미하며, ~F4는 네번째 속성이 비전형적임을 나타낸다. 각 범주들의 전형적인 속성들은 F1→F2→F3→F4의 인과적 사슬구조(causal chain)을 가지고 있었다. 따라서 인과적 사슬을 형성하는데 필요한 세 개의 인과관계와 그에 대한 간략한 설명도 추가로 구성했다. 각 참가자들은 키호 개미, 미야 별, 넷톤 컴퓨터 중 하나의 범주를 학습한 후 속성추론 과제를 수행했기 때문에 범주는 참가자간 요인이었다.

Table 1. Example of feature dimensions and their associated values (Myastar)

Feature dimension	Typical value (most)	Atypical value (some)
Helium	Ionized	Non-ionized
Temperature	Very hot	Less hot
Density	High	Low
Number of planets	Larage	Small

2.1.2. 참가자

총 30명의 학부생들이 실험에 참여했으며, 각 범주에 동일한 인원을 무선적으로 할당했다.

2.1.3. 실험절차

참가자들은 할당된 범주에 대한 지식을 습득한 후 속성추론과제를 수행했다. 범주학습 단계에서 참가자들은 화면의 여러 페이지에 제시된 범주관련 내용을 자율적으로 학습했다. 첫 화면에는 범주명칭과 범주에 대한 한두 문장의 소개글을 제시했으며, 두 번째 화면에는 범주의 속성차원과 전형적인 및 비전형적인 속성값 및 기저율(대부분, 일부)을 제시했다. 세 번째 화면에는 범주 속성간 인과관계(F1→F2, F2→F3, F3→F4)를 서술했으며, 마지막 화면에는 이를 Fig. 1과 같이 그림으로 요약하여 제시했다. 범주 학습이 완료되면 참가자들은 범주의 속성차원, 속성값과 기저율, 인과관계에 대해 서술했으며 모두 정확하게 기억한 경우 속성

추론과제를 수행했다.

속성추론 과제에서 참가자들은 주어진 범주예시에서 숨겨진 한 속성의 속성값을 추론하는 과제를 수행했다. 속성차원이 총 네 개였으며 이중 세 개는 두 가지 속성값을 가질 수 있으므로, 총 32회의 속성추론 시행이 제시되었으며(4 x 2 x 2 x 2 = 32) 각 시행은 무선적으로 제시되었다. 전형적인 속성과 비전형적인 속성을 1과 0으로 나타낼 때, x111은 F2, F3, F4가 전형적일 때 F1의 속성값을 추론하는 시행이 된다. 유사하게, 1X00은 F1이 전형적이고 F3, F4가 비전형적일 때 F2를 추론하는 시행이다.

Fig. 2는 속성추론의 한 시행이다. 맨 위에 범주명칭을 표시했으며, 그 아래에 각 줄에 속성차원 및 속성값을 순서대로 제시했다. 참가자들은 표적속성 “????”의 속성값에 대해 좌우 방향키를 눌러 세로막대를 이동한 뒤 엔터키를 눌러 반응했다. 가로막대의 왼쪽에는 비전형적인 속성값, 오른쪽에는 전형적인 속성값을 표시했다. 참가자들은 11점 척도로 반응했으며 0-100 범위로 변환하여 분석했다.

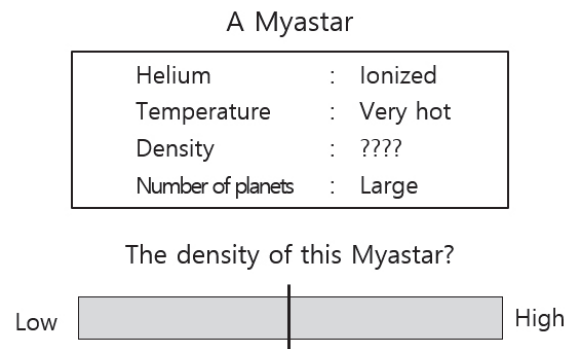


Fig. 2. Example of inference rating trial

2.2. 결과 및 논의

참가자간 요인인 범주(키호 개미, 미야 별, 넷톤 컴퓨터)에 따른 속성추론의 유의미한 차이가 없어 결과를 통합하여 분석했다.

2.2.1. 분석방법

그림 1과 같은 인과적 사슬구조에서의 속성과 속성관 인과관계가 추론에 미치는 영향에 대한 분석에는 약간의 모호함이 존재한다(e.g., Mayrhofer & Rothe,

2012). 이러한 모호함이 발생하는 이유 중 하나는 방향성을 가진 비대칭적 관계가 연쇄적으로 발생하기 때문이다. 가령, F2에 대한 추론은 선행속성인 F1의 존재여부에 의해 영향을 받지만 후행속성인 F3의 존재여부도 F2에 대한 추론에 영향을 미칠 수밖에 없다. 본 연구에서는 Fig. 3과 같이 속성간 거리(distance)를 통해 실험 결과를 분석하여 속성추론에서의 비독립성을 검증하고자 한다. 구체적으로 살펴보면, 네 개의 속성으로 이루어진 사슬구조에서는 세 가지 속성간 거리가 존재한다(dist 1, 2, 3). F1에 대한 추론의 경우 dist 1에 존재하는 F2는 추론에 영향을 미쳐야 하지만, dist 2, 3의 F3, F4는 ‘인과적 추론’에는 영향을 미치지 않아야 한다. 마찬가지로 Fig. 3에서 F1, F2, F3, F4의 각 속성에 대한 추론에서 공통적으로 dist 1에 있는 속성은 추론에 영향을 미쳐야 하지만 나머지 dist 2, 3의 속성값은 인과추론에 영향을 미치지 않아야 한다. 01x1의 x에 대해 dist 1, 2, 3에는 각각 2, -1, 0개의 속성이 존재하며, 110x에는 각 거리에 -1, 1, 1개의 속성이 존재하게 된다. 이때 양수는 x의 전형적 속성값에 대한 긍정단서이며, 0은 중립, 음수는 부정단서이다.

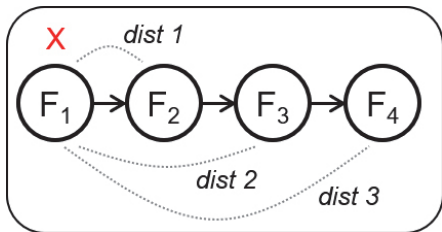


Fig. 3. Distances 1, 2, 3 from F1

2.2.2. 결과 및 논의

Fig. 4는 각 거리에서 속성개수에 따른 속성추론의 평정결과를 보여준다. 이때 dist 1에 대해서는 dist 2, 3의 결과를 통합했으며, dist 2는 dist 1, 3를 통합한 결과이며, dist 3는 dist 1, 2를 통합했다. 표적속성 x에 대해 dist 1에는 -2, -1, 0, 1, 2 개의 속성이 존재하며, dist 2, 3 각각에서 속성 개수의 합은 0으로 상쇄된다. 마찬가지로 dist 2에는 -1, 1개의 속성이 존재하며, dist 1, 3 각각에서 속성개수의 합은 0이다. Dist 3에는 -1, 0, 1개의 속성이 존재하며 dist 1, 2에서 속성개수의 합은 0으로 상쇄된다. 또한 표적속성 x에 대해 다른 속성들이 인과적으로 선행 혹은 후행하는지 여부도 완전

역균형화되어 상쇄된다. 규범적(rational) 추론자의 경우, 인과적 사슬구조에서 속성 x에 대한 추론은 dist 1에 있는 직접 연결된 속성의 개수에 따라 달라져야 하며 dist 2, 3인 경우에는 다른 속성에 의해 차폐(screen-off)되기 때문에 속성 개수에 영향을 받지 않아야 한다.

Fig. 4는 dist 1, 2, 3에서 모두 속성의 개수 증가에 따라 표적속성 x에 대한 평정이 높아졌음을 보여준다(cf. Rehder & Burnett, 2005, p. 297). Dist 1에서는 직접적인 원인 혹은 결과의 존재가 표적속성 x에 대한 추론에 강한 영향을 미쳤으며 이는 당연한 결과이다. 그런데 dist 2, 3에서도 속성의 개수 증가에 따른 추론평정이 높아졌으며 이는 인과적 마코프 조건에 대한 위배를 보여준다. 즉, 범주기반 추론에서 인과적으로 무관한 속성 개수의 증가가 다른 속성 x의 존재 가능성을 높게 평가하게 만든다는 것이다. 특히, dist 2와 dist 3에서 기울기가 거의 동일하다는 것은 참가자들이 좀 더 가까운 dist 2의 속성들을 dist 3의 속성들 보다 더 유용한 증거로 여기지는 않음을 보여준다. 이는 참가자들이 암묵적인 인과관계를 가정했다기 보다는 범주 예시의 전형성을 토대로 속성 x의 존재여부를 추론했다는 증거가 된다. 예를 들어, x000과 x011에 대해 규범적 인과 모형이론은 x에 대해 동일한 추론을 예측하지만 후자에서 전형성 효과로 인해 속성추론이 강하게 발생했다는 것이다.

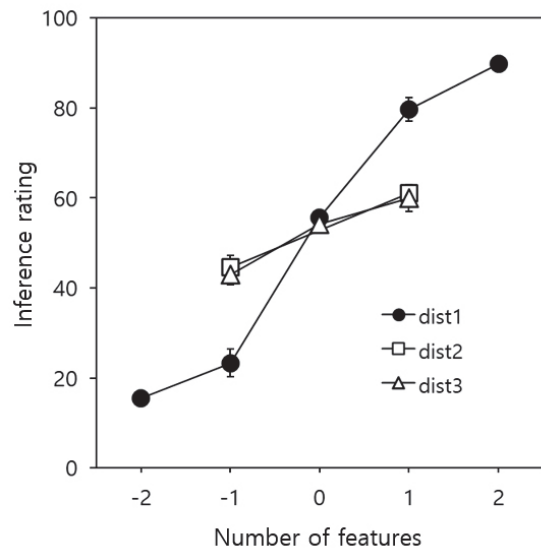


Fig. 4. Mean inference ratings at distances 1, 2, 3 (Error bars indicate SE)

통계적 검증을 위해 각 참가자들의 평정값을 dist 1, 2, 3로 예측하는 다중회귀분석을 실시했다. 그 결과 dist 1에 대한 평균 가중치는 21.8로 나타났으며 이는 dist 2와 3에 대한 평균 가중치 8.3과 8.5 보다 모두 유의미하게 높았다, $ps < .001$. 이는 직접 연결된 속성의 존재여부가 추론에 가장 강한 영향을 미친다는 것을 보여준다. 나아가 dist 2와 3의 가중치 차이는 유의미하지 않았으나 모두 0보다는 유의미하게 높은 것으로 나타났다, $ps < .001$. 이는 인과적 연결관계가 없는 속성도 추론에 영향을 미치며 그 강도는 인과적 거리와 무관함을 보여준다.

3. 속성추론의 모델링

본 연구의 목적 중 하나는 인과적 사슬구조에서의 속성추론을 검증하고 이를 모델링하는 것이다. 기존 연구에서 제안한 속성추론 모형이 공통원인(common cause)과 공통효과(common effect) 인과구조에서의 전형성 효과, 비독립성, 인과적 절감효과들을 재현해낸다는 것을 보여주었다(Kim & Li, 2017). 이 모형이 인과적 사슬구조를 가지는 범주에서 거리분석에 따른 속성추론의 질적인 특성들을 보여주는지 검증할 것이다.

3.1. 모형의 작동방식과 계산절차

Fig. 5는 인과적 사슬에서 표적속성 x가 존재할 가능

성이 높은 경우(1x11)와 낮은 경우(0x01)를 도식적으로 보여준다. 1x11의 경우 x의 원인과 결과가 모두 존재하기 때문에 F2인 x의 존재가능성이 높다. 0x01은 원인과 결과가 모두 부재하므로 F2의 존재가능성이 낮다.

속성추론모형은 x에 대한 추론이 두 단계에 걸쳐 발생한다고 제안한다. 먼저 추론단계(단계 1)에서는 속성 x가 존재하거나 하지 않는 두 가지 가능성에 대한 확률 계산이 발생한다고 제안한다. 직관적으로 볼 때 x가 있는지 혹은 없는지를 판단해야하므로 두 가능성 모두에 대해 가늠해볼 가능성이 높다. Fig. 5A를 예로 들면 1x11에서 F2가 존재하는 경우와 하지 않는 경우는 각각 p(1111)과 p(1011)로 표현할 수 있다. 판단단계(단계 2)에서는 두 확률간의 차이에 대한 계산이 이루어지고 이를 토대로 속성추론 평정이 이루어진다고 가정한다.

Fig. 5B의 F2에 대한 추론은 p(1111) - p(1011)에 근거한다고 예측하는데 차이값이 양수가 될 것으로 추측할 수 있다. 반면 Fig. 5B의 0x01은 속성의 존재가능성이 낮은 경우이다. 실제로 p(0101)과 p(0001)를 비교하면 좌향의 경우 인과관계가 모두 어긋나기 때문에 확률이 낮게 추정될 것이며 따라서 p(0101) - p(0001)은 음수가 되어 참가자들이 척도의 왼쪽에 반응할 것으로 예측한다.

속성추론모형의 계산에서 핵심을 이루는 것은 속성이 존재하거나 존재하지 않는 범주에서의 확률값을 계산하는 것이다. 이를 정형화하여 서술하면 아래의 같은 식 (1)을 도출할 수 있다.

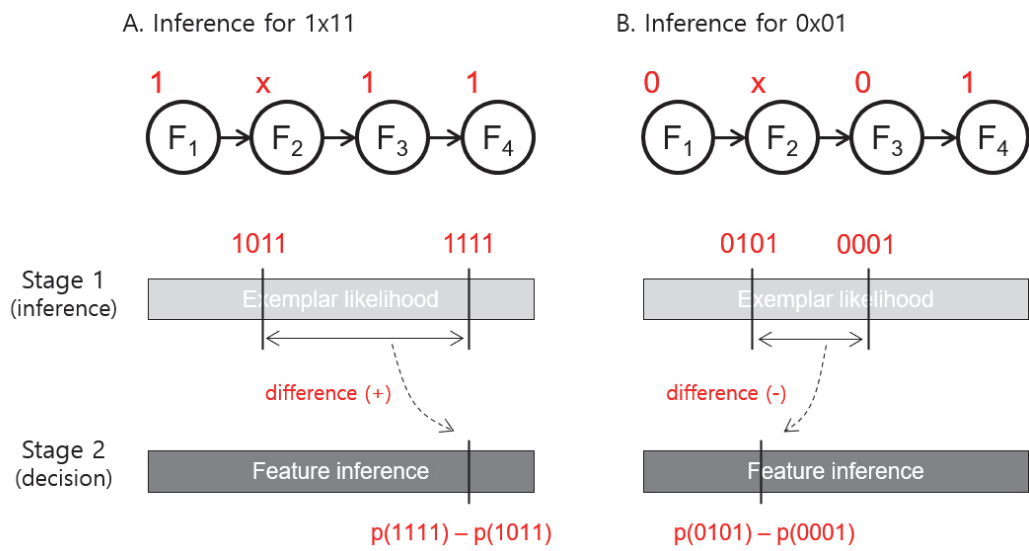


Fig. 5. Worked-out examples of feature inference in 1x11 (A) and 0x01 (B)

$$\text{Rating}(F_X) = f[p(E_{F(X)}|\text{Cat}) - p(E_{F(\sim X)}|\text{Cat})] \quad (1)$$

이때, F_X 는 추론대상이 되는 표적속성이며, $p(E_{F(X)}|\text{Cat})$ 은 표적속성 F_X 가 존재하는 범주예시(exemplar) E 의 확률이며, $p(E_{F(\sim X)}|\text{Cat})$ 은 표적속성 F_X 가 존재하지 않는 범주예시 E 의 확률이다. 실험결과와 비교는 다음 식 (2)를 통해 이루어질 수 있다.

$$\text{Rating}(F_X) = K_1*[p(E_{F(X)}|\text{Cat}) - p(E_{F(\sim X)}|\text{Cat})]+K_2 \quad (2)$$

이때, K_1 과 K_2 는 모형의 예측을 참가자들의 [0 100] 척도상 반응으로 변환하는 역할을 한다. 표적 속성이 존재하거나 $p(E_{F(X)}|\text{Cat})$ 존재하지 않는 경우의 확률은 $p(E_{F(\sim X)}|\text{Cat})$ 속성간 인과관계를 확률적으로 표상하는 인과모형이론(Rehder, 2003; Sloman, 2005; Waldmann & Holyoak, 1992)으로 계산할 수 있다. 가령 F_1, F_2 의 두 속성으로 이루어진 범주에서 $F_1 \rightarrow F_2$ 의 인과관계가 있다고 할 때, 속성조합에 따라 네 가지 범주예시가 존재할 수 있다(11, 10, 01, 00). 원인속성(F_1), 결과속성(F_2), 인과관계의 확률을 각각 ‘c’, ‘b’, ‘m’으로 놓으면 확률규칙에 따라 $p(11) = c*(m+b-m*b)$, $p(10) = c*(1-m)*(1-b)$, $p(01) = (1-c)*b$, $p(00) = (1-c)*(1-b)$ 으로 쉽게 계산이 가능하다. 가령, x_1 이라는 추론시행이 주어진 경우 속성추론모형은 식 (1)에 따라 $p(11) - p(01)$ 의 강도로 F_1 의 존재를 예측하게 된다. 식 (2)를 적용하면 $\text{Rating}(F_1) = K_1*[p(11)-p(01)]+K_2 = K_1*[c*(m+b-m*b) - (1-c)*b]+K_2$ 이라는 계산식이 도출된다. 속성의 개수가 늘어나거나 인과관계가 복잡해지더라도 동일한 확률

규칙을 통해 특정 속성조합을 가지는 범주예시의 개연성을 계산할 수 있다.

인과모형이론은 베이저언 그래프 모형(Glymour, 2001; Pearl, 2000)의 한 형태로서 속성간 인과관계를 확률적으로 계산하여 개별 범주예시의 개연성을 예측하게 된다. 모형의 식 (1), (2)가 나타내는 것은 속성추론 과정에서 참가자는 표적속성 x 가 존재하는 범주예시와 존재하지 않는 범주예시의 개연성을 암묵적으로 계산하여 그 차이를 근거로 속성 x 의 존재를 판단한다는 것이다. 인과모형이론은 규범적 확률계산을 기반으로 하기 때문에 직접적인 인과관계가 없는 속성들의 독립성을 가정한다. 속성추론이론은 이를 처리모형으로 변형하여 두 가능성의 직관적인 차이가 속성추론의 근거가 된다고 제안한다. 즉, 참가자들은 속성이 존재할 가능성과 존재하지 않을 가능성 간의 차이를 직관적으로 인식하여, 이를 토대로 반응한다는 것이다.

모형의 예측에 대한 직관적 이해를 위해 $F_1 \rightarrow F_2$ 의 두 속성으로 이루어진 범주의 범주예시와 속성추론에 대한 결과값을 제시하였다(Fig. 6; $c=m=b=.5$). Fig. 6A는 개별 범주예시의 존재확률을 보여주는데 인과관계가 존재함에 따라 속성의 개수뿐만 아니라 속성간 상호작용에 의해 확률값이 달라짐을 알 수 있다. Fig. 6B는 네 가지 속성추론 항목에 대한 속성추론 강도의 예측을 보여준다. 원인 혹은 결과속성이 존재하지 않는 경우(0x, x0)보다 존재하는 경우($x_1, 1x$) 속성의 존재를 더 강하게 예측하였으며 이는 직관적인 이해와 일치하는 결과이다. 이때 범주예시 확률과 속성추론 강도는 절대적인 값이 아니라 상대적인 차이에 따른 패턴을

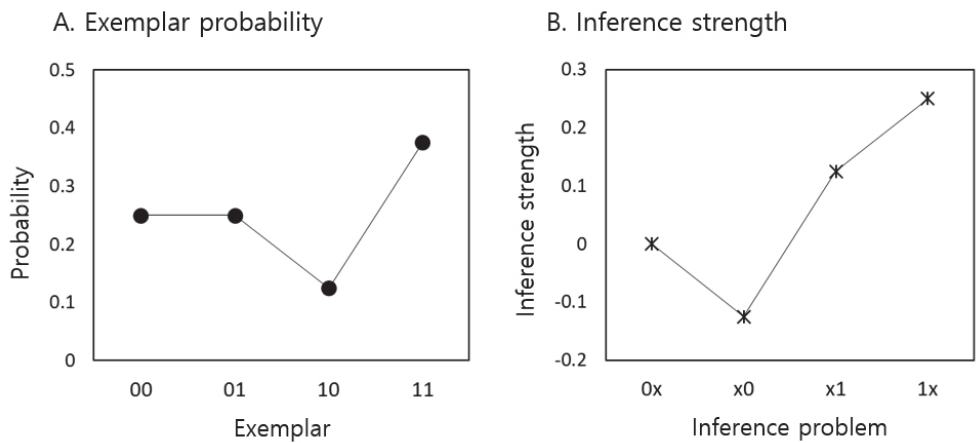


Fig. 6. Exemplar probabilities (A) and inference strength (B) in a two-feature causal category

확인해야 한다. 인과모형이론에서 모든 범주예시에 대한 확률값의 합은 규칙에 따라 1.0이 되며 따라서 속성 노드가 증가할수록 개별 범주예시의 확률은 점점 작아진다. 속성추론의 강도는 두 범주예시의 확률 차이에 대한 직관적인 인상을 반영하는 것이기 때문에 대부분의 심리학 모형과 마찬가지로 절대적인 값을 예측하는 것은 아니다.

3.2. 속성추론모형의 검증

속성추론모형이 인과적 사슬구조의 범주에서 인과적 마코프 조건의 위배를 예측할 수 있는지가 본 연구의 핵심적인 검증사항이었다. 모형 검증을 위해 각 참가자의 평정값에 대해 모형의 예측을 dist 1, 2, 3에 따라 Fig. 7에 제시했다(avg. RMSD= 15.3). 그 결과 dist 1(Fig. 7A)에서 속성의 개수에 따라 모형의 예측값도 증가하는 것을 확인할 수 있다. 물론 실제 평정값은 S 형태를 나타내지만 이는 바닥효과와 천장효과에 의한 것으로 보인다. 따라서 dist 1에서 모형은 속성추론 패턴을 타당하게 예측하는 것으로 해석할 수 있다. 마찬가지로 dist 2, 3(Fig. 7B와 7C)에서도 속성 개수의 증가에 따라 모형의 예측값도 증가하며 이는 모형이 실제 데이터의 형태를 질적으로 타당하게 설명한다고 해석할 수 있다. 특히 속성 개수가 증가함에 따라 예측값도 증가하는 것은 모형이 인과추론의 비독립성 즉 인과적 마코프 조건의 위배를 올바르게 설명함을 보여준다. 다만 dist 2, 3에서는 모형이 변화량을 과소추정하는 것으

로 나타났으며 이는 dist 3에서 더욱 명확하게 나타난다.

통계적 검증을 위해 각 참가자에 대한 모형의 평정값을 dist 1, 2, 3로 예측하는 다중회귀분석을 실시했다. 그 결과 dist 1에 대한 평균 가중치는 17.6으로 나타났으며 이는 dist 2와 3에 대한 평균 가중치 5.0과 2.3보다 모두 유의미하게 높았다, $ps < .001$. 이는 직접 연결된 속성이 추론에 더 강한 영향을 미친다는 실험결과와 일치하는 예측이다. 나아가 dist 2와 3의 가중치 모두 0보다 유의미하게 크게 나타났으며 이는 직접 연결되지 않은 속성도 인과추론에 영향을 미친다는 비독립성을 모형이 타당하게 예측함을 보여준다. 그러나 Fig. 7B와 7C에서 보여지듯이 dist 2의 가중치는 dist 3의 가중치에 보다 더 컸으며, $t(29) = 15.2, p < .001$, 이는 속성추론모형이 인과적으로 먼 속성의 영향을 과소추정한다는 것을 보여준다.

종합하면, 제안한 속성추론모형은 속성들이 인과적 사슬구조로 연결된 범주에서의 속성추론의 패턴을 전반적으로 타당하게 예측한다고 해석할 수 있다. 특히 속성추론모형은 dist 2, 3에서 직접적인 인과관계가 없는 속성의 존재여부도 표적 속성의 추론에 영향을 미치는 인과추론의 비독립성을 적합하게 예측하였다. 다만 인과적으로 직접 연결되지 않은 속성들은 거리와 상관없이 동일하게 평정값에 영향을 미친 반면 모형은 인과적인 거리에 따라 추론에 미치는 영향은 차별적임을 예측했다는 점에서는 차이가 있다. 이에 대해 본 실험의 특수성과 모형의 특성에 비추어 결론에서 논의하고자 한다.

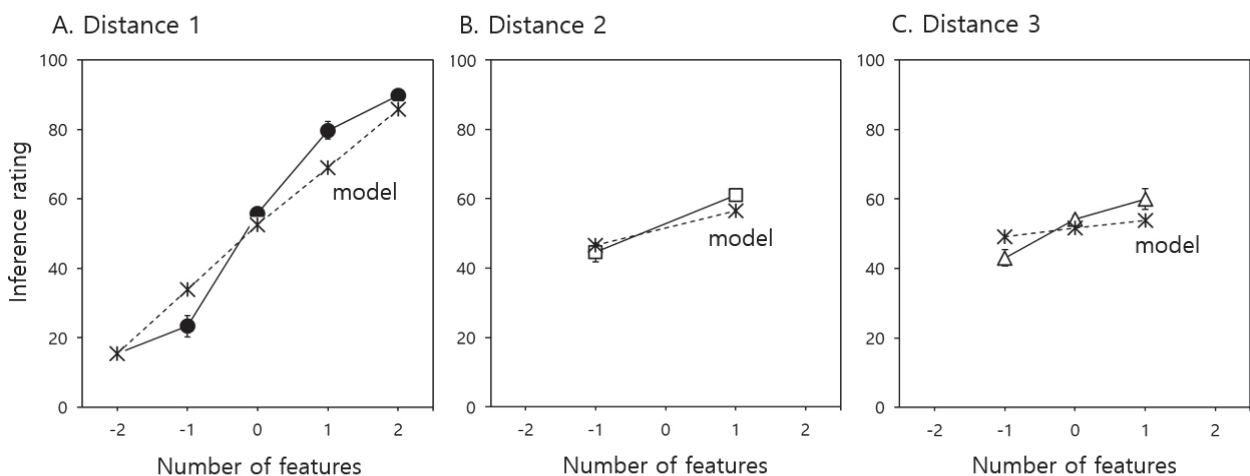


Fig. 7. Subjects' ratings (Error bars indicate SE) and model predictions at distances 1, 2, 3 (panels A, B, C, respectively)

4. 결론

본 논문에서는 속성들이 인과적 사슬구조로 연결된 범주에서의 속성추론을 실험하고 이를 모델링하였다. 이를 위해 네 개의 속성 차원으로 구성된 세 가지 범주를 사용했으며 참가자들은 할당된 조건의 범주를 학습한 뒤 다양한 속성값 조합으로 이루어진 32개의 범주 예시에 대한 속성추론을 실시했다. 인과관계가 연쇄적으로 발생하는 사슬구조에서 속성들의 직접 영향과 간접 영향을 분리하기 위해 거리분석을 통해 평정값을 코딩하여 결과를 통합했다. 그 결과 인과적으로 직접 연결된 속성이 추론에 가장 강한 영향을 미쳤으며, 간접적으로 연결된 dist 2, dist 3의 속성들은 추론에 동등한 영향을 미친 것을 확인할 수 있었다. 이 결과는 기존 연구에서 밝혀진 인과적 마코프 조건의 위배가 인과적 사슬구조에서 나타난다는 것을 보여준다(cf. Rehder & Burnett, 2005).

실험 결과를 모델링하기 위해 인과모형이론(causal model theory)에 기반한 속성추론 모형의 예측값과 참가자들의 평정값을 비교했다. 그 결과 속성추론 모형은 참가자들의 속성추론 패턴을 전반적으로 유사하게 묘사하는 것으로 나타났다. 특히 다른 속성 노드에 의해 차폐(screen-off)된 dist 2, dist 3의 속성들이 추론에 영향을 미치는 인과적 마코프 조건의 위배를 예측하는 것으로 나타났다.

제안한 속성추론 모형은 몇 가지 중요한 가정에 기초한다. 표적 속성값에 대한 추론과제가 주어질 때 모형은 먼저 해당 속성이 존재하거나 존재하지 않은 두 가능성의 개연성을 내적으로 추정한다고 가정한다. 이어서 두 가능성의 주관적인 차이를 계산한 뒤 이에 근거하여 의사결정을 한다고 제안한다. 이 과정은 Fig. 5에서 예를 들어 설명했으며 식 (1), (2)의 계산과정에 구체화되어 표현되어있다. 어떤 속성의 존재여부를 추론하는 것은 그 자체로써 존재하거나 하지 않거나 하는 가능성들을 가늠할 수 밖에 없기 때문에 이 가정은 직관적으로 타당하다고 생각된다. 특정 속성이 존재할 가능성이 매우 높은 경우에는 그 속성이 존재하지 않을 가능성은 낮으므로 차이값은 (+)가 될 것이고 반대의 경우 차이값은 (-)가 될 것이다. 두 가능성의 개연성이 유사하다면 차이값은 0이 될 것이다.

본 연구에서는 속성추론모형을 참가자들의 반응과 직접 비교하기 위해 참가자들에게 Fig. 2와 같은 척도에 반응하도록 요청하였다. 대안적인 종속변인 측정방안으로 속성의 존재 혹은 부재의 두 가능성 중 하나를 선택하게 하는 2AFC (two-alternative force choice)를 사용할 수 있다. 이 방법을 선택할 경우 식 (1)에서 차이값 대신 선택공리(the axiom of choice; Luce, 1959, 1977)를 적용하여 모델링할 수 있으며 dist 1에서 나타난 천정 및 바닥효과를 피할 수 있다. 또한 선택 반응은 매우 민감한 측정치이므로 작은 차이도 종속변인에 반영된다는 장점이 있다. 다만 각 참가자에게 충분히 신뢰로운 선택 확률을 획득하기 위해 본 연구에서 사용한 32개의 추론문항을 여러 번 반복해 제시해야한다는 단점이 존재하기 때문에 실제 실험에서 적용하기에는 현실적으로 어렵다.

본 논문에서 제안한 속성추론모형은 기존의 인과모형이론(CMT; causal model theory)에 기반한다. 인과모형이론은 속성들 간 인과관계에 대한 명시적인 지식(explicit knowledge)가 존재할 때의 범주화 패턴을 매우 잘 설명하는 검증된 이론이다(Rehder, 2003; Rehder & Kim, 2010). 본 연구에서도 참가자들에게 속성과 그들의 인과관계에 대한 직접적인 지식을 제공한 뒤 속성추론과제를 수행하도록 요청했다. 그런데 인과관계에 대한 지식은 선언적으로만 존재하는 것이 아니라 경험과 관찰에 의해서도 형성되며 이러한 경우 방향성을 가진 인과관계가 아닌 속성간 관련성을 단순 연합으로 표상하여 추론을 수행할 가능성도 있다(e.g., Danks, 2009; Hanus, 2016; Lagnado et al., 2007). 따라서 학습방법과 표상의 종류에 따른 속성추론 양상을 확인하고 이에 대한 모형검증도 필요할 것이다.

본 연구에서는 범주 속성들이 인과적 사슬구조를 가질 때의 속성추론을 검증하고 이를 모델링하였다. 기존 연구(Kim & Li, 2017)에서는 범주 속성이 공통원인과 공통효과 구조로 연결되었을 때 동일한 모형으로 속성추론을 검증한 바 있다. 식 (1)에 기반한 속성추론 모형이 서로 다른 세 가지 인과 네트워크에서 추론의 비독립성, 인과적 절감(causal discounting), 전형성 효과를 모두 설명했다는 것은 본 모형이 인과적 범주에서의 속성추론의 핵심을 잘 반영한다고 볼 수 있다.

Fig. 7B와 7C에 대한 분석에서 드러난 것과 같이 속

성추론 모형은 인과적 마코프 조건의 위배를 설명함에도 불구하고 거리가 먼 속성의 경우 추론에 미치는 영향이 작아짐을 예측했다. 이는 아마도 본 모형이 네 개의 속성으로 이루어진 작은 범주보다 훨씬 더 일반화된 상황을 가정하기 때문인 것으로 보인다. 가령 n 개의 속성으로 이루어진 인과적 범주에서 n 의 값이 커짐에 따라 먼 거리의 속성값이 표적속성의 추론에 미칠 영향은 작아질 수 밖에 없다. 즉, 속성의 인과 네트워크의 크기가 큰 경우(직관적으로) 거리에 따른 영향력의 차이가 발생할 수 밖에 없는데 본 실험에서는 단지 네 개의 속성만을 사용했기 때문에 참가자들의 평정값에서 dist 2와 3의 차이가 발생하지 않았을 수 있다. 또한 인과 네트워크가 커지는 경우 특히 작업기억이나 주의 용량의 제약으로 거리가 먼 속성의 영향이 작아질 것으로 추측해볼 수 있을 것으로 보인다.

본 연구에서는 인지적 추론과정을 검증하기 위해 사전지식을 통제하고 정서 및 감성의 영향을 배제하기 위해 인공적인 범주를 사용했다. 그러나 현실세계에서의 속성추론은 지각된 위험과 감성의 영향에 의해 크게 달라질 수 있다. 예를 들어 화려한 색을 보이는 야생 버섯의 독성 여부를 판단해야하는 상황에서 사람들은 독성을 과대 추정할 수도 있다. 마른 기침을 하는 사람에게 독감이나 폐렴 바이러스가 있을 수 있다고 지각하여 회피하는 것은 정서적 공포가 속성의 확률을 과대추정할 수 있다는 것을 보여주는 직관적인 예이다. 혹은 위험지각이 특정 속성의 확률적 추론에는 영향을 미치지 않지만 단순히 행동이나 대처방식에 대한 의사결정에만 영향을 미칠 수도 있다. 즉 속성추론은 비교적 합리적으로 발생하지만 그 결과를 활용하는 측면에서 정서나 감성이 영향을 미칠 수 있는 것이다. 실제로 위험한 범주를 포함하는 범주기반 속성추론에서 참가자들이 반드시 지각된 위험의 영향을 받지 않는다는 연구 결과가 존재한다(Zhu & Murphy, 2013). 아울러 Dunsmoor & Murphy(2014)는 특정한 범주의 범주예시(exemplar)에 대해 공포조건형성(fear conditioning)이 되었을 때 공포의 일반화 정도는 조건형성 당시에 경험한 범주예시의 전형성(typicality)에 따라 달라진다는 것을 보고했다. 이는 범주기반 속성추론에 대한 기존 연구 결과들이 범주나 속성의 정서적 특성에 의해서는 크게 영향을 받지 않는 일반적 원리라는 것을 암시한다.

마지막으로 본 연구에서는 특정 범주 내에 명확하게 속하는 개별 범주 예시들의 속성에 대한 추론을 검증하였다. 이 때문에 식 (1), (2)에서의 범주 예시에 대한 확률값은 해당 예시가 그 범주에 속한다는 것을 가정하고 계산하였다. 그런데 속성추론은 범주가 모호한 경우에도 빈번하게 발생하고 또 하나의 대상이 복수의 범주에 속하기도 한다(Chen et al., 2016; Murphy & Ross, 2010; Ross & Murphy, 1999). 교차 범주화 혹은 범주가 모호한 경우에 나타난 속성추론의 양상을 이론적 모형으로 구체화하고 이를 모델링하는 것도 향후 중요한 과제중 하나일 것이다. 아울러 독립변인 특히 인과강도의 조작에 따른 모형의 파라미터 값의 변화를 확인함으로써 모형의 질적 타당성을 추가적으로 검증하는 것이 향후 연구에서 필요할 것이다.

REFERENCES

- Ahn, W. K., Kim, N. S., & Lebowitz, M. S. (2017). The role of causal knowledge in reasoning about mental disorders. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 603-617). Oxford University Press.
- Ahn, W. K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, *41*(4), 361-416. DOI: 10.1006/cogp.2000.0741
- Bailenson, J. N., Shum, M. S., Atran, S., Medin, D. L., & Coley, J. D. (2002). A bird's eye view: Biological categorization and reasoning within and across cultures. *Cognition*, *84*(1), 1-53. DOI: 10.1016/S0010-0277(02)00011-2
- Coley, J. D., & Vasilyeva, N. Y. (2010). Generating inductive inferences: Premise relations and property effects. In B. H. Ross (Ed.), *The psychology of learning and motivation*, *53* (pp. 183-226). Burlington, VT: Academic Press. DOI: 10.1016/S0079-7421(10)53005-6
- Danks, D. (2009). The psychology of causal perception and reasoning. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *Oxford handbook of causation* (pp. 447-470). Oxford: Oxford University Press.

- Doh, E. Y., & Lee, G. H. (2020). Effect of interaction between category coherence and base rate on presumption of reasons for preference. *Korean Journal of Cognitive Science*, 31(3), 77-102. DOI: 10.19066/cogsci.2020.31.3.002
- Dunsmoor, J. E., & Murphy, G. L. (2014). Stimulus typicality determines how broadly fear is generalized. *Psychological Science*, 25(9), 1816-1821. DOI: 10.1177/0956797614535401
- Feeney, A., & Heit, E. (2011). Properties of the diversity effect in category-based inductive reasoning. *Thinking & Reasoning*, 17(2), 156-181. DOI: 10.1080/13546783.2011.566703
- Fenker, D. B., Waldmann, M. R., & Holyoak, K. J. (2005). Accessing causal relations in semantic memory. *Memory & Cognition*, 33(6), 1036-1046. DOI: 10.3758/BF03193211
- Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.
- Hadjichristidis, C., Sloman, S., Stevenson, R., & Over, D. (2004). Feature centrality and property induction. *Cognitive Science*, 28(1), 45-74. DOI: 10.1016/j.cogsci.2003.09.001
- Hanus, D. (2016). Causal reasoning versus associative learning: A useful dichotomy or a strawman battle in comparative psychology?. *Journal of Comparative Psychology*, 130(3), 241-248. DOI: 10.1037/a0040235
- Heit, E., & Hahn, U. (2001). Diversity-based reasoning in children. *Cognitive Psychology*, 43(4), 243-273. DOI: 10.1006/cogp.2001.0757
- Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(2), 411-422. DOI: 10.1037/0278-7393.20.2.411
- Heit, E., Hahn, U., & Feeney, A. (2005). Defending diversity. In W. K. Ahn, R. L. Goldstone, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside of the laboratory: Essays in honor of Douglas L. Medin* (pp. 87-99). Washington, DC: American Psychological Association.
- Kim, N. S., & Keil, F. C. (2003). From symptoms to causes: Diversity effects in diagnostic reasoning. *Memory & Cognition* 31(1), 155-165. DOI: 10.3758/BF03196090
- Kim, N. S., Luhmann, C. C., Pierce, M. L., & Ryan, M. M. (2009). The conceptual centrality of causal cycles. *Memory & Cognition*, 37(6), 744-758. DOI: 10.3758/MC.37.6.744
- Kim, S., & Li, H. C. O. (2017). Modeling feature inference in causal categories. *Korean Journal of Cognitive Science*, 28(4), 329-347. DOI: 10.19066/cogsci.2017.28.4.007
- Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (p. 154-172). Oxford University Press. DOI: 10.1093/acprof:oso/9780195176803.003.0011
- Lee, G. H., & Park, S. Y. (2016). The effect of manufacturing method preferences for different product types on purchase intent and product quality perception. *Science of Emotion & Sensibility*, 19(4), 21-32. DOI: 10.14695/KJSOS.2016.19.4.21
- Lee, G. H., Choi, J. H., Ahn, C. H., Li, H. C. O., & Kim, S. (2014). Suggestion of similarity-based representative odor for video reality. *Science of Emotion & Sensibility*, 17(1), 39-52. DOI: 10.14695/KJSOS.2014.17.1.39
- Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.
- Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, 15(3), 215-233. DOI: 10.1016/0022-2496(77)90032-3
- Marsh, J., & Ahn, W. (2006). The role of causal status versus inter-feature links in feature weighting. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 561-566). Mahwah, NJ: Cognitive Science Society.
- Mayrhofer, R., & Rothe, A. (2012). Causal status meets coherence: The explanatory role of causal models in categorization. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 743-748). Austin, TX: Cognitive Science Society.

- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Murphy, G. L., & Ross, B. H. (2010). Category vs. object knowledge in category-based induction. *Journal of Memory and Language*, 63(1), 1-17. DOI: 10.1016/j.jml.2009.12.002
- Murphy, G. L., Chen, S. Y., & Ross, B. H. (2012). Reasoning with uncertain categories. *Thinking and Reasoning*, 18(1), 81-117. DOI: 10.1080/13546783.2011.650506
- Nagel, E. (1939). Probability and the theory of knowledge. *Philosophy of Science*, 6(2), 212-253.
- Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, 97(2), 185-200. DOI: 10.1037/0033-295X.97.2.185
- Osta-Vélez, M., & Gärdenfors, P. (2020). Category-based induction in conceptual spaces. *Journal of Mathematical Psychology*, 96, 102357. DOI: 10.1016/j.jmp.2020.102357
- Patalano, A. L., & Ross, B. H. (2007). The role of category coherence in experience-based prediction. *Psychonomic Bulletin & Review*, 14(4), 629-634. DOI: 10.3758/BF03196812
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.
- Proffitt, J. B., Coley, J. D., & Medin, D. L. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(4), 811-828. DOI: 10.1037/0278-7393.26.4.811
- Rehder, B. (2003). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1141-1159. DOI: 10.1037/0278-7393.29.6.1141
- Rehder, B. (2006). When similarity and causality compete in category-based property generalization. *Memory & Cognition* 34(1), 3-16. DOI: 10.3758/BF03193382
- Rehder, B. (2009). Causal-based property generalization. *Cognitive Science*, 33(3), 301-344. DOI: 10.1111/j.1551-6709.2009.01015.x
- Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology*, 50(3), 264-314. DOI: 10.1016/j.cogpsych.2004.09.002
- Rehder, B., & Hastie, R. (2004). Category coherence and category-based property induction. *Cognition*, 91(2), 113-153. DOI: 10.1016/S0010-0277(03)00167-7
- Rehder, B., & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 659-683. DOI: 10.1037/0278-7393.32.4.659
- Rehder, B., & Kim, S. (2009). Classification as diagnostic reasoning. *Memory and Cognition*, 37(6), 715-729. DOI: 10.3758/MC.37.6.715
- Rehder, B., & Kim, S. (2010). Causal status and coherence in causal-based categorization. *Journal of Experimental Psychology: Learning Memory and Cognition*, 36(5), 1171-1206. DOI: 10.1037/a0019765
- Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning & Verbal Behavior*, 14(6), 665-681. DOI: 10.1016/S0022-5371(75)80055-7
- Ross, B. H., & Murphy, G. L. (1999) Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology*, 38(4), 495-553. DOI: 10.1006/cogp.1998.0712
- Shafto, P., Coley, J. D., & Baldwin, D. (2007). Effects of time pressure on context-sensitive property induction. *Psychonomic Bulletin & Review*, 14(5), 890-894. DOI: 10.3758/BF03194117
- Shafto, P., Coley, J. D., & Vitkin, A. (2007). Availability in category-based induction. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Experimental, developmental, and computational approaches* (pp. 114-136). Cambridge University Press. DOI: 10.1017/CBO9780511619304.006
- Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, 25(2), 231-280. DOI: 10.1006/cogp.1993.1006
- Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition*, 52(1), 1-21. DOI: 10.1016/

0010-0277(94)90002-7

Sloman, S. A. (1997). Explanatory coherence and the induction of properties. *Thinking & Reasoning*, 3(2), 81-110. DOI: 10.1080/135467897394374

Sloman, S. A. (2005). *Causal models: How we think about the world and its alternatives*. New York: Oxford University Press.

Sloman, S. A., Love, B. C., & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, 22(2), 189-228. DOI: 10.1016/S0364-0213(99)80039-1

Thagard, P., & Nisbett R. E. (1982). Variability and confirmation. *Philosophical Studies*, 42, 379-394. DOI: 10.1007/BF00714369

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive

and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, 121(2), 222-236. DOI: 10.1037/0096-3445.121.2.222

Zhu, J., & Murphy, G. L. (2013). Influence of emotionally charged information on category-based induction. *PLoS One*, 8(1), e54286. DOI: 10.1371/journal.pone.0054286

원고접수: 2021.02.17

수정접수: 2021.02.18

게재확정: 2021.03.01