

결과를 보고 가능한 원인을 생각하는 추론을 진단 추론(diagnostic inference)이라고 한다. 초인종이 울렸을 때 누가 왔는지를 추론하는 것, 감기에 걸린 원인을 추론하는 것이 이에 해당한다. 이와 같은 추론은 우리가 일반적으로 자주하는 추론이기에 일상 추리(everyday reasoning)로 불리거나, 혹은 귀추 추론(abductive inference)으로 불리기도 한다(Josephson & Josephson, 1996).

이러한 진단 추론을 우리는 얼마나 합리적으로 해낼까? 혹시 추론 직전에 있었던 사건에 일시적이거나 영향을 받아 편향된 추론을 하지는 않을까? 예를 들어, 한가로운 오후에 낯선 사람이 초인종을 눌렀을 때, 세 시간 전 ‘도둑들’이라는 영화를 본 사람들은, 그렇지 않은 사람들보다 문을 더 열어주지 않을지도 모른다. 사전 맥락이 이후 고차 인지 활동에도 영향을 미친다는 연구 결과들을 생각해 보면, 그 가능성은 높아 보인다. 그렇다면 과연 이 영향의 적용 범위는 어디까지일까? 가능한 여러 대안 중 오직 하나의 대안만이 답이 되는 추론의 경우와, 대안들 간 상호작용이 가능한 복합적인 원인을 추론하는 경우에, 사전 맥락이 동일하게 영향을 미칠 수 있는 것일까?

본 연구는 이상의 질문들에 대한 답을 찾아보기 위해 수행되었다. 구체적으로 사전에 우연적으로 활성화된 의미가, 진단 추론의 유형에 따라 어떠한 영향을 미치는지를 탐색하고자 한다.

가설생성과 기억

진단 추론은 결과를 가장 잘 설명할 수 있는 원인을 추구하지만, 이 추론이 항상 참을 보장하지는 않는다. 우리 집 초인종을 누른 사람이 옆집 이웃일 것이라고 추론하는 것, 감기에 걸린 원인을 추운 날씨 때문으로 추론하는 것은 맞을 가능성이 높지만, 동시에 틀릴 가능성이 존재한다. 이 때문에 진단 추론의 결론은 하나의 가설로 볼 수 있고, 따라서 가설 생성(hypothesis generation)이 진단 추론에서 핵심 역할을 담당한다고 할 수 있다.

실제로 가설생성은 판단의 기초가 되는 중요한 인지과정이기 때문에(Thomas et al., 2008; Lange et al., 2013; Zuckerman et al., 1995; Weber et al., 1993), 진단 추론 뿐 아니라 다양한 분야에서 연구가 진행되어왔다. 예컨대, 환자의 증상을 관찰하여 진단을 생성하는 의학 분야(Weber et al., 1993), 행동 패턴에 기반을 두어 성격 특성

을 설명하는 사회 분야(Zuckerman et al., 1995; Strack et al., 1993) 주어진 증거를 기반으로 유/무죄를 판결하는 법학 분야(Pennington & Hastie, 1988; Reyes et al., 1980), 가설 생성/평가 과정을 바탕으로 진단 전문가 시스템(diagnostic expert system)을 구현하는 컴퓨터 과학 분야(Reggia et al., 1985) 등이다.

이상의 연구들을 통해 밝혀진 사실 중 하나는, 가설 생성이 기억과 밀접한 관련이 있다는 것이다. Weber와 동료들(1993)은, 가설은 의미 지식(semantic knowledge)과 일화기억(episodic memory)을 활용하는 사고 과정을 통해 생성된다는 것을 주장하였으며, Libby(1985) 역시 진단 모델에서 가설생성의 핵심은 주어진 단서와 지식 구조와의 상호작용임을 역설했다. 이 밖에 가설 생성에서 개인차는 작업기억의 용량, 일반 지식의 양, 경험에서 비롯된다는 결과들은 가설이 기억에 의존한다는 사실을 수렴적으로 보여준다(Dougherty & Hunter, 2003; Weber et al., 1993; Ho & Keller, 1994; Libby & Fredrick, 1990; Schwarz, 2004).

현재 기억과 가설 생성의 과정을 가장 잘 설명하는 모형 중 하나는 *HyGene* 모형이다(Thomas, Dougherty, Sprenger & Harbison, 2008). 이 모형에 따르면 가설이란 모든 외부 사건의 전체 집합(the universe of possible states) 중에서 학습과 경험을 통해 의미 지식으로 표상된 사건들을 지칭한다. *HyGene* 모형이 가정하는 가설생성 과정을 간단히 살펴보면, 1) 어떤 사건과 단서가 주어졌을 때 2) 주어진 단서와 관련된 가설들 중 일부가 작업기억에서 활성화 되고, 3) 가장 높은 활성화를 지닌 가설은 최종가설로 선택된다. 예를 들어, 1) 누군가 초인종을 눌렀을 때, 이와 관련된 단서 - 택배 상자, 유니폼, 남자 - 를 기초로 하여, 2) 단서와 관련된 가설들이 - 택배기사, 경비원, 남편 등 - 생성된다. 그리고 3) 단서에 의해 가장 크게 활성화 되는 가설인 택배기사는(택배기사 (3), 경비원 (2), 남편 (1)), 최종 가설로 선택된다. 이와 같은 가정은 여러 선행 연구들에 근거한다. 단서의 빈도수, 생생함, 주어지는 순서를 변경하여 한 가설의 활성화 정도를 조작하였을 때(예, Reyes et al., 1980; Sprenger & Dougherty, 2012; Mehlhorn et al., 2011; Pennington & Hastie, 1988; Schwarz, 2004; Lange et al., 2013), 그 단서와 관련된 가설이 최종 선택될 확률은 높아졌다.

지금까지 가설 생성과 관련된 연구들과 모형은 가설 생성에 직접적으로 관련된 단서들의 효과에 초점을 두고 있다. 즉, 단서를 통한 가설들의 활성화 변화만을 다루어왔다. 하지만 이렇게 직접적으로 제시된 단서가 없더라도, 가설의 활성화는 사

전 맥락을 통해 달라질 가능성이 있다. 예를 들어, Kusev 와 동료들의 연구(2012)에서, 사전 의미 활성화의 유무에 따라 사람들은 다른 판단을 내렸다. 연구에 참여했던 모든 참가자들은 도난 보험에 가입할 것인지 결정해야 했는데(표 1의 2번 문제), 이 판단에 앞서, 아래 1번 문제(의미 활성화 유발)에 먼저 답을 한 실험집단은 다른 집단들과 다른 선택패턴을 보여주었다: ‘차 도둑’의 의미 활성화가 사전에 발생했던 실험 집단은, 통제집단이나, 관련 없는 의미 활성화(‘차 도둑’ 대신에 ‘차 사고’문제가 제시되었다)를 받은 집단보다, 도난 보험에 더 많이 가입하였다. 사전에 다른 맥락을 통해 발생한 의미 활성화로 인해, 실험집단은 ‘차 도둑’에 대해 높은 위험 회피 성향을 보인 것이다.

표 1. Kusev연구(2012)에서 활용되었던 문제 일부

1. 12개월 동안 차 안에 소지품을 훔쳐가는 도둑이 더 많았던 도시는 어디일까? a) 런던 b) 뉴욕
2. 1%의 확률로 당신의 차 안에 300파운드 가량의 물건이 도둑맞을 수 있다. 당신은 30파운드의 비용을 들여 보험에 가입할 것인가? a) 예 b) 아니오

이 연구를 통해 Kusev와 동료들은, 사전 다른 맥락에서 발생한 의미 점화가 이후 판단에 영향을 미친다는 것을 주장하였다. 사전 우연적 의미 점화는 이후 같은 의미를 가진 사건의 활성화를 추가로 증가시킨다(Kusev, Schaik & Aldrovandi, 2012; Ross et al., 1977; Tversky & Kahneman, 1973).

이와 유사하게 Erb, Bioy 그리고 Hilton의 연구(2002)에서도 사전에 위험 회피와 관련된 의미가 활성화된 집단은, 문제해결 상황에서 더 확실한 해결책을 선호했다. 이들의 연구에서 실험 집단은 본 문제를 풀기 전에 기억문제에 먼저 답을 하였다. ‘위험 추구’ 집단의 경우, 진취적인(enterprising), 모험심있는(adventurous)과 같은 위험 추구를 긍정적으로 보는 형용사들이 기억 문제로 제시된 데 반해, ‘위험 회피’ 집단에게는 주의깊은(careful), 조심스러운(conscientious)과 같은 형용사들이 제시되었다. 그 다음에 두 집단은 같은 글을 읽고 해결책을 선택하였는데¹⁾, 예상했던 대로, 위

험 추구 집단은 위험 추구 해결책을, 위험 회피 집단은 반대의 해결책을 더 선호하였다.

이상의 연구들은 사전에 다른 맥락에서 우연적으로 발생한 의미 활성화가, 이후 문제를 풀 때 특정 대안의 선택을 증가시킬 수 있음을 보여준다. 예를 들어 ‘도둑’의 의미가 사전에 활성화 된 경우, 이후 ‘도둑’과 관련된 가설의 선택이 증가하는 것이다. 뿐만 아니라, 이 과정에서 우연적 활성화는 추론에 대한 확신에도 영향을 미칠 수 있다. Kelly와 Linsay(1993)에 따르면, 사람들은 더 빨리 기억나는 답을 정답으로 확신하는 경향이 있다. 이들의 연구에서 참가자들은 사람, 장소, 물건에 대한 목록을 먼저 읽었다. 이 목록에는 다음에 풀어야 할 일반 상식 문제에 대한 답, 혹은 답과 관련은 있지만 오답, 혹은 상관없는 답들이 섞여있었다. 목록을 모두 읽은 사람들은 일반 상식 문제에 답을 하였는데, 사전에 각 문제와 관련된 단어를 읽었던 사람들은, 그 목록에 제시된 단어들이 정답이든 오답이든 상관없이, 자신의 답에 더욱 확신을 갖는 경향을 보였다(Kelly & Linsay, 1993; Alter & Oppenheimer, 2009).

이상의 논의를 종합해 보면, 사전 우연적 의미 활성화는 한 가설의 활성화를 일시적으로 높이고, 이로 인해 추론에 대한 확신을 증가시킬 것을 예상해 볼 수 있다.

기존 연구들은 주로 선택, 판단, 태도에 미치는 의미 활성화의 영향을 다루어 왔지만, 진단 추론에 미치는 영향을 직접적으로 검증한 연구는 거의 없었다. 판단과 진단 추론은 모두 고차적 사고를 요한다는 특징이 있고, 과제 수행이 기억에 상당히 의존하고 있는 점을 고려해 보았을 때, 동일한 현상이 진단 추론에서도 발생할 수 있다. 본 연구에서는 일차적으로 이 가능성을 탐색하고자 하였다.

가설의 유형

이와 더불어 본 연구의 또 다른 관심은, 우연적 의미 활성화의 영향이 가설의 유형에 따라서 달라지는지를 알아보는 것이다. 우연적 활성화를 다루었던 기존 연

1) 예를 들어, 현재 휴가를 마치고 집을 가야하는 상황에서 참여자들은 a) 계획했던 것보다 14시간이 늦어지는 기차(위험 회피), 또는 b) 계획했던 시간에 도착 가능, 하지만 파업 시 계획보다 이틀이 늦어질 수 있는 기차(위험 추구) 중에서 하나를 선택해야 했다.

구들은 오직 배타가설(mutually exclusive and exhaustive hypothesis)에만 초점을 맞추어 왔다. 하지만, 가설의 중요한 기능 중 하나가 설명이라는 것을 고려해 보았을 때 (Lombrozo, 2012; Thagard, 1997), 다음과 같이 가설의 유형에 따라 우연적 활성화의 영향이 달라질 가능성이 있다.

Josephson과 Josephson(1994)에 따르면, 가설 생성이란 어떠한 현상에 대한 설명들을 만들고, 이 중 가장 좋은 설명을 최종 선택하는 일련의 과정을 의미한다. 주목해야 할 점은, 설명은 모두 같은 구조와 기능을 가지고 있지 않으며, 설명의 특성에 따라서 추론 결과가 달라질 수 있다는 것이다(Lombrozo, 2006, 2007; Sloman, 1994). 예를 들어, Sloman(1994)은 설명들이 서로 양립 가능할 때와, 배타적인 경우에 따라 추론의 타당성이 달라질 것을 주장하였다(Sloman, 1994). 설명들이 서로 배타적인 경우 한 설명의 타당성은 다른 설명과의 비교를 통해서 이루어지며, 이 경우 한 설명의 타당성이 증가하면, 다른 설명의 타당성은 감소한다. 반면 양립 가능한 설명들은 서로 공존할 수 있으며, 두 설명이 동시에 높은 타당성을 가질 수 있다.

기존에 우연적 의미 활성화의 영향을 다루었던 대부분의 연구들은, 배타가설 세트에만 초점을 맞추어 왔다(McKenzie, 1998). 또한 연구 참가자들은 ‘주어진’ 가설의 가능성을 각각 평가해야 했다(가설 평가). 하지만 일상생활에서 많은 문제들은 배타가설 뿐 아니라 양립할 수 있는 가설을 다루고 있으며, 더욱이 모든 사람들이 같은 수의 가설을 고려하는 것도 아니다.

이에 본 연구는, 기존 연구들을 확장하여, 우연적 활성화가 배타가설뿐 아니라 양립가설에 미치는 영향을 탐색하고, 주어진 가설 세트에서 ‘선택’하는 것이 아닌, 스스로 ‘생성’하는 과정을 다루고자 하였다. 두 가설이 서로 배타적인 경우, 우연적 의미 활성화로 한 가설의 가능성이 높아지면 다른 가설의 가능성은 낮아질 것이지만(complementary evaluation)²⁾, 두 가설이 서로 양립할 수 있다면, 한 가설의 가능성의 증가로 인해 다른 가설의 가능성이 낮아질 필요는 없을 것이다. 이 과정에서 1) 최종 선택되는 가설과, 2) 고려되는 가설 세트의 크기, 3) 추론에 대한 확신이 가설 유형에 따라서 달라질 것을 예상할 수 있다. 실험 1에서는 이 세 예상을

2) 가설의 수가 적을 경우에는(예, 2개) 상호 보완적으로 평가 한다(Van Wallendael & Hastie, 1990).

검증한 다음, 이 결과가 인지적 노력이 가중되었을 때 어떻게 달라지는지를 탐색하였다. 실험 2에서는 단서가 주어진 경우 발생하는 변화에 대해 알아보았다.

실 험 1

실험 방법

대상

실험 1은 서울 소재 S대학교 학부 생 96명을 대상으로 진행되었다(남 57명, 여 39명, 평균 나이 23세). 실험에 참여한 모든 참가자들은 각 수업에서 참여점수를 획득하였다.

과제

총 두 개의 과제가 사용되었다. 첫 번째 과제(과제1)는 기존 연구에서(Malt, Ross & Murphy, 1996) 사용된 범인 추론 문제를 번역 및 수정한 것이고, 두 번째 과제(과제2)는 Fenton과 Neil(2012)에서 사용된 일반 추론 문제를 번역 및 수정한 것이다. 과제1은 ‘초인종을 누른 사람’을 추론하는 문제(표 2 참조), 지문에는 ‘도둑’과 ‘택배기사’가설이 모두 제시되었다. 이 과제에서 문 앞에 남자는 단 한 명이기 때문에 가설들은 상호 배타적인 특징을 갖는다(배타가설). 반면 과제2는 ‘가난한 산모들이 저체중아를 더 많이 출산하는 원인’에 대해서 추론하는 것으로, 제시된 가설로는 ‘정부의 차별된 건강관리 정책’과 ‘가난한 산모들의 흡연과 알코올 섭취가 높다는 것’이었다. 이 과제의 가설들은 서로 상호작용이 가능하며, 동시에 모두 답이 될 수 있다는 특징이 있다(양립가설).

절차

모든 실험 참가자들은 ‘글 이해와 추론’이라는 연구로 알고 참여하였으며, 지문을 읽고 각 문제에 대한 추론(초인종을 누른 사람/ 저 체중아 출산 원인)을 하였다. 문제를 풀 때 다시 지문의 내용은 볼 수 없었으며, 추론은 주관식으로 자유롭게

답하였다. 실험집단은 총 두 집단으로 이들은 다른 맥락의 문제를 먼저 푼 뒤(표 2의 1, 2번), 진단 추론을 하였으며(표 2의 3, 4, 5번), 통제집단은 다른 맥락의 문제가 제시되지 않았다(표 3).

표 2. <과제1>에서 사용된 지문과 일부 문제 예시(실험에 사용된 과제에는 밑줄이 생략되었다)

주부 박 모양(32)은 요즘 집을 나서기가 무섭다. 최근 며칠 동안 아파트 단지 내에 도둑이 출몰하고 있으며, 이 도둑은 주부 혼자 있거나 외출했을 때 집중적으로 활동하고 있으니 더욱 조심해야 한다는 동네 아주머니들의 주의가 계속 이어졌기 때문이다. 하지만 오늘은 부모님들께서 오후 6시경 모처럼 집에 방문한다고 하셨기에 집안 청소를 하고 장을 보러 나가야 한다. 두 명의 아이들이 현재 유치원에 가서 한가할 때 얼른 청소를 마치고, 부모님들께서 좋아하시는 닭백숙을 하기 위해 재료를 사러 집에서 20분 거리에 있는 마트에 가서 장을 봐서 2시쯤 돌아올 계획이다. 또한 아침에 아파트 경비아저씨가 재활용품을 내놓는 날이 오늘로 바뀌었다고 알려주셔서, 재활용품도 밖에 내다 놓아야 한다. 박 양은 먼저 빨래를 세탁기에 돌리고, 간단히 설거지를 한 뒤 분리수거 할 것을 정리하고 있었는데 초인종이 울렸다. 확인해 보니 한 사람이 있었으나 얼굴이 보이지 않았다. 팔 부분의 옷자락만 보일 뿐이었다. 지금 시간에 올 사람이 누구일까 생각하다가, 불현듯 며칠 전 부모님들께서 포도를 보내주신다고 말씀하셨던 것이 생각났다. 아마도 택배직원이 포도를 가지고 왔을 것 같았다. 박 모양은 응답을 하기 위해 서둘러 인터폰을 들었다.		
다른 맥락의 문제(통제집단의 문제지에는 생략)	실험 집단 1	1. 집에 도둑이 들었던 경험이 있는가? 2. 2013년 현재를 기준으로 절도범죄 검거율은 얼마큼이라고 생각하는가? (%)
	실험 집단 2	1. 택배를 받아 본 경험이 있는가? 2. 100명중 몇 명이 택배서비스를 이용하고 있을 것 같는가? 그 이유는? (명)
진단 추론과 확신(실험집단 1, 실험집단 2, 통제집단)		3. 문 밖에 있는 사람은 누구라고 생각하는가? 4. 위에 응답한 답 이외에 생각한 대안(들)은 있었는가? 있다면 무엇이고, 그 답이 맞을 확률은 어느 정도라고 생각하는가? 5. 지금까지의 추론에 얼마큼 확신이 있는가?

표 3. 실험 1 조건 간 차이 정리

		실험집단		통제집단
		집단1	집단2	
		한 가설의 의미를 활성화시 키는 사전 문제 2개 포함	한 가설의 의미를 활성화시 키는 사전 문제 2개 포함	활성화
활성화된 가설	과제1: 도둑 과제2: 음주	과제1: 택배기사 과제2: 정부정책		없음

다른 맥락의 문제인 표 2의 1번과 2번은 실험참가자들에게 특정 가설의 의미를 일시적으로 활성화시키는 목적으로 삽입되었다(실험집단1에는 도둑 의미가, 실험집단2에는 택배기사 의미가 진단 추론 이전에 활성화되었다). 예를 들어, 표 2를 보았을 때, 실험 집단은 1번과 2번의 문제를 푼 뒤 초인종을 누른 사람을 추론하였고(3번), 통제집단은 1번과 2번의 문제가 없이 바로 초인종을 누른 사람을 추론하였다. 참여자들이 연구의 목적을 사전에 알아채지 못하도록 표 2에 제시된 문제 이외에, 무의미 문제 4개가 포함되었다. 요약하면, 모든 참가자들은 같은 지문을 읽고 같은 추론 문제에 답을 하였다. 다만 실험집단은 추론 문제를 풀기에 앞서 다른 맥락의 문제 두 개에 추가로 답을 했다는 차이가 있다.

실험이 모두 끝나고 실험 참가자들은 연구의 목적에 대하여 자유롭게 서술한 뒤 퇴실하였다. 시간제한은 없었으나 평균 15~20분 정도의 시간이 소요되었다.

실험 결과

실험의 목적을 알아차린 6명의 피험자들을 제외하고, 총 90개의 데이터를 분석하였다.

전반적으로 확인된 사실은, 예상했던 대로 우연적 의미 활성화는 1) 진단 추론에 영향을 미치며, 2) 이 영향은 가설의 유형에 따라서 다음과 같이 다르게 영향을 미친다는 것이다.

결과1) 최종가설 선택패턴

진단 추론과 다른 맥락의 문제에 의해 활성화된 의미가 ‘추론 결과(최종 가설 선택 패턴)’에 미치는 영향을 검증하기 위하여, 교차분석을 실시하였다. 실험집단1과 실험집단2, 그리고 통제집단 간에 가설 선택 빈도를 비교하였다. 그 결과 배타가설의 추론에서는(과제1), 우연적 의미 활성화에 따라 특정 가설의 선택 빈도가 더 높아졌다($\chi^2(2) = 6.875, p < .05$). 그림 1에서 보는 것과 같이, ‘도둑’활성화를 받은 실험집단1은 다른 조건보다 더 많이 ‘도둑’가설을 선택하였다. 모든 집단에서 택배기사 가설이 도둑 가설보다 더 많이 선택되었음을 고려해볼 때, 사람들은 기

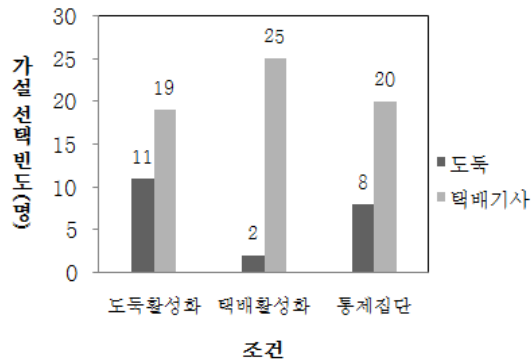


그림 1. 과제1(배타가설)의 최종 가설 선택

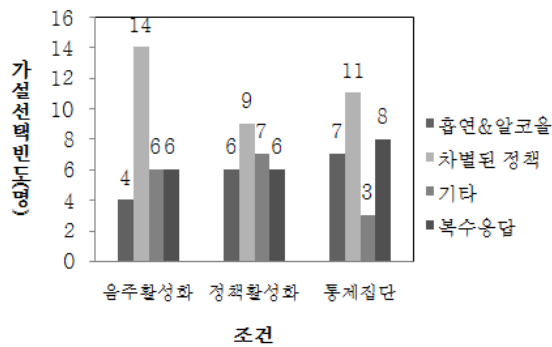


그림 2. 과제2(양립가설)의 최종 가설 선택

저율(base rate)을 어느 정도 고려하고 있음을 알 수 있다.

반면 양립가설의 추론에서는(과제2), 사전 활성화가 최종 가설 선택에 영향을 미치지 않았다($\chi^2(6) = 3.894, p > .05$). 그림 2에서 보는 것과 같이, 특정 가설의 활성화 여부와 상관없이 사람들은 ‘흡연&알코올’가설보다 ‘차별된 정부 정책’ 가설을 더 많이 선택하였으며, ‘흡연&알코올’가설보다 ‘복수 응답’을 더 선호하였다.

결과2) 생성된 가설의 수

우연적 활성화가 ‘가설의 생성 개수³⁾’에 미치는 영향을 알아보기 위해서, 통제 집단과 실험집단 간에 독립표본 T 검정을 실시하였다. 활성화 여부에 따라, 지문에 제시된 두 개의 가설들 중 몇 개의 가설을 고려했는지를 비교해 보았고, 이를 포함하여 총 생성된 가설의 개수 역시 비교해 보았다.

첫 번째 결과와는 반대의 효과가 관찰되었다. 과제1(배타가설)의 경우 실험집단과 통제집단 간에 지문에 등장한 가설의 고려 개수는 차이가 없었으며($t(87) = 0.385, p > .05$), 전체 생성된 가설의 개수에서도 차이가 발견되지 않았다($t(86) =$

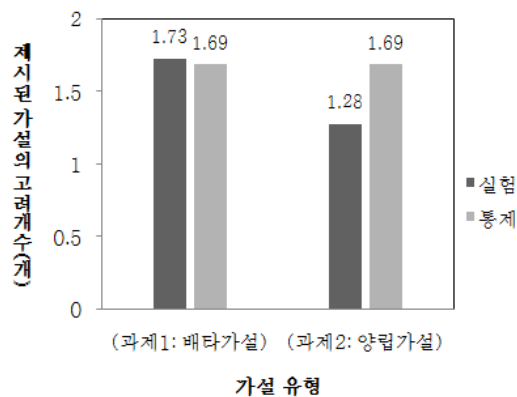


그림 3. 제시된 가설의 생성 개수의 차이

3) 표 2의 3,4번 문제에 작성되었던 가설들의 개수를 세어서 분석하였다. 예를 들어, 3번에 ‘도둑’으로 답하고, 4번에 ‘택배기사, 경비아저씨’를 작성한 사람은 제시된 가설 중 두 개를 생성(그림 3), 총 세 개의 가설을 생성(그림 4)한 것이다.

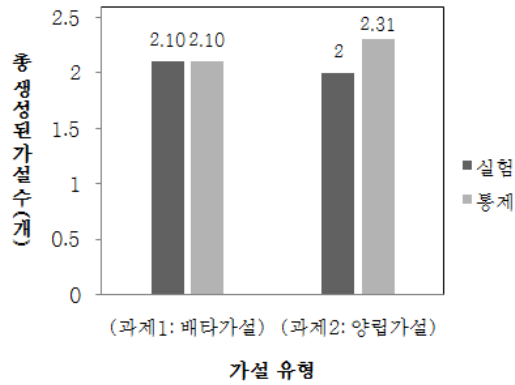


그림 4. 총 생성된 가설 개수의 차이

0.011, $p > .05$). 특정 가설의 활성화 여부와 상관없이, 통제집단과 실험집단은 지문에 등장한 두 개의 가설 중 평균 1.7개의 가설을 고려하였으며, 총 생성한 가설의 개수는 평균 2개였다. 반면 과제2(양립가설)의 경우 실험집단(평균 = 1.28, 표준편차 = 0.7)보다 통제집단(평균 = 1.69, 표준편차 = 0.54)이 제시된 가설을 더 많이 고려하였으며($t(70.484) = 2.989, p < .01$), 총 생성된 개수에서도 실험집단(평균 = 2.0, 표준편차 = 0.7)보다 통제집단(평균 = 2.31, 표준편차 = 0.54)이 더 많은 수의 가설을 생성하였다($t(84) = 2.072, p < .05$) (그림 3, 그림 4).

결과3) 추론에 대한 확신

우연적 활성화가 ‘추론에 대한 확신’에 미치는 영향을 확인하기 위해, 실험집단과 통제집단 간에 추론 확신 정도를 비교하였다. 독립표본 T 검정 결과, 가설의 유형에 상관없이 두 집단의 차이는 없었다(과제1: $t(88) = 0.303, p > .05$; 과제2: $t(88) = 0.629, p > .05$).

논 의

실험 1을 통해, 우연적 의미 활성화가 진단 추론에 미치는 영향은 가설의 유형

에 따라 달라지는 것을 확인하였다(표 4 참조).

표 4. 실험 1 결과 요약

	과제 1 (배타가설추론)	과제 2 (양립가설 추론)
최종 선택 가설	영향O	영향X
생성된 가설 수	영향X	영향O
추론 확신	영향X	영향X

실험 1의 첫 번째 결과로, 우연적 활성화는 배타가설의 최종선택에만 유의미한 영향을 미친다는 것을 확인하였다. 배타가설의 경우 우연적 활성화에 따라서 편향되게 가설이 생성됐지만, 양립가설의 경우 이러한 현상이 나타나지 않았다.

이와 관련하여 가능한 해석은 양립가설 세트에서 점화효과의 확산 가능성이다. 양립가설의 경우 가설들 간에 상호작용이 가능하기 때문에, 의미 점화 효과가 같은 의미를 가진 가설에만 영향을 주지 않고, 가설 세트 전체 집합으로 확산될 수 있다(McKenzie, 1998). 즉, 그림 5에서 보는 것과 같이 배타가설의 경우 서로 경쟁하고 있기 때문에, 우연적 의미 활성화가 같은 의미를 가지고 있는 가설에만 직접 영향을 미친다. 반면 양립가설의 경우, 의미 활성화가 가설들 사이의 상호 연결 마디를 통해, 전체 가설 세트로 확산될 수 있다. 이로 인해 우연적 의미 활성화가 같은 의미를 가진 양립 가설의 최종 선택에 미치는 영향은 작았다.

또 다른 해석으로, Kahneman과 Miller의 규준 이론(Norm Theory, 1986)도 생각해 볼 수 있다. 이 이론에 따르면, 현상에 대한 한 가설은 다른 가설을 활성화 시키지만, 이 둘이 서로 경쟁적일 경우에는 상호 억제 효과가 일어난다고 가정한다. 이를 고려해 보면, 본 연구에서 우연적 의미 활성화가 같은 의미를 가진 배타가설의 활성화를 높이면, 서로 경쟁적인 위치에 놓여있는 다른 배타가설의 활성화는 상대적으로 낮아진다. 이로 인해 사람들은 우연적 의미 활성화와 같은 의미를 가진 가설(예, 도둑)을 최종 선택한 것이다. 한편 양립가설들은 서로 경쟁하지 않으므로, 한 가설의 활성화 증가가 다른 가설의 활성화 억제로 이어지지 않는다. 따라서 양립가설

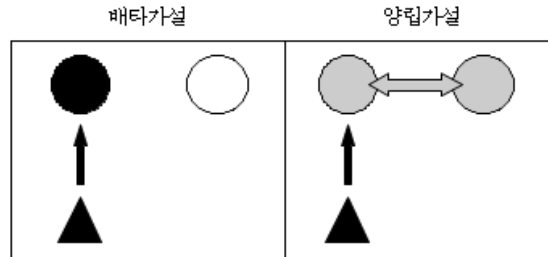


그림 5. 가설 유형에 따른 우연적 의미 활성화의 영향: 배타가설의 경우 다른 맥락의 (세모) 의미 활성화(검정)가 같은 의미를 가진 한 가설(원)에 직접 영향을 미친다(검정 원). 반면 양립가설의 경우 다른 맥락의(세모) 의미 활성화(검정)는 같은 의미를 지닌 가설(원)에만 영향을 미치지 않고, 상호 연결 마디를 통해 전체 가설 세트로 확산된다(회색의 가설세트).

의 최종 선택에 우연적 의미 활성화는 큰 영향을 주지 않았다.

첫 번째 결과와는 반대로, 두 번째 결과에서 우연적 활성화는 양립가설의 생성 수에만 유의미한 영향을 미친다는 것이 확인됐다. 양립가설에서 우연적 의미 활성화를 받은 집단은 통제집단보다 더 적은 수의 가설을 생성하였다. Sloman(1994)은 설명이 서로 배타적일 때, 한 설명의 가능성이 증가하면 다른 설명의 생성이 억제되어, 총 더 적은 수의 설명이 만들어질 것으로 생각하였다. 그러나 본 연구에서 오히려 배타적이지 않은 가설의 경우, 한 가설의 가능성의 증가가 다른 가설의 생성을 방해한다는 것을 발견하였다. 이 결과는 ‘차단(block)’ 또는 ‘고착(fixation)’현상 (Tseng et al., 2008; Bassok & Novick, 2012)으로 이해된다. 차단 및 고착 현상이란, 이전에 생성된 가설로 인해 새로운 가설이 생성되지 못하는 것을 의미 한다. 우연적 활성화는 한 가설을 먼저 사람들의 작업 기억으로 불러들이고, 이 가설에 고착되게 하여, 추가로 다른 가설을 생성하지 못하게 한다. 이는 가설생성이 복잡할 때, 우연적 의미 활성화가 다양한 가설의 생성을 억제할 수 있다는 것을 시사한다.

처음 예상과는 달리, 우연적 활성화는 가설의 유형에 상관없이 추론에 대한 확신에 영향을 미치지 못하였다. 이는 가설이 생생히 떠오를수록 추론에 대한 확신이 높아진다는 선행 연구와(Kelley, 1993) 대치되는 것으로 보일 수도 있다. 그러나 본 연구와 Kelley 등의 연구(1993)에서 문제에 따라 답을 확신하는 정도의 차이가

있음을 고려해야 한다. 구체적으로, Kelley의 연구에서 사람들은 일반 상식 문제에 답을 하였다(예, 한글이 창제된 연도는? 답: 1443년). 일반 상식 문제에서 답은 정답이거나 오답이지, 보통 답의 가능성은 잘 고려하지 않는다(알거나 혹은 모르거나). 이러한 이유로 인해 보통 일반 상식 문제를 풀 때 우리는, 오답이라고 하더라도, 기억에서 빨리 떠오른 답을 정답으로 간주하기 십상이다(Kelley et al., 1993). 하지만 본 연구에서 사람들은 진단 추론에 답을 하였다. 일반 상식 문제와는 달리, 진단 추론의 답들은 모두 ‘가설’로, 정답 가능성을 가지고 있다. 즉, 진단 추론의 결과인 최종 선택된 가설은 유일한 정답이 아닌, 현재 상황에서 ‘가장 그럴 듯한’ 답일 뿐이다(예, 내 시험 점수가 낮은 이유는? 답: 채점에 문제가 있었기 때문). 때문에 현재 하나의 가설이 쉽게 떠오르더라도 다른 가설들을 완전히 배제하기 어렵다. 더욱이 판단을 뒷받침할 근거가 제시되지 않은 상태라면 떠올린 가설에 더욱 확신을 가지기 힘들 것이다. 이러한 차이로 인해 본 연구에서 우연적 활성화는 추론에 대한 확신을 높이는 데 실패한 것으로 판단된다.

이상의 결과들로 우연적 의미 활성화는 진단추론에 영향을 미친다는 것을 알게 되었다. 우연적 의미 활성화는 추론과 다른 맥락으로부터 발생되었기 때문에, 이와 같은 결과는 평소 우리가 비합리적으로 추론할 수 있을 가능성을 보여준다. 예를 들어, 자동차에 시동이 걸리지 않을 때 우리는 현재 가지고 있는 단서, 기저울에 기초하여 합리적으로 그 원인을 추론하지 못할 수 있다. 대신에, 단순히 옆 좌석에 앉아있는 친구의 핸드폰 배터리가 나갔다는 이유로 인해 ‘자동차의 시동이 걸리지 않는 것도 아마 배터리가 없기 때문’으로 추론할 수 있다. 합리적인 근거에 기반하지 않고 ‘현재 배터리가 쉽게 떠오르는 것을 보면, 자동차 시동 문제의 원인은 배터리 때문일 것이다’라는 엉터리 추론을 하고 있는 셈이다.

기존 연구들에서, 의미 활성화의 효과는 직관적이고 덜 분석적인(less analytic) 사고를 할 때 더 크며(Gilad & Kliger, 2008), 더 깊은 사고를 하는 사람일수록 판단과 직접적으로 관련 없는 요소에 덜 민감하게 반응한다는 것을 검증하였다(Smith & Levin, 1996). 본 연구에서도 더 깊은 사고를 하는 것이 우연적 의미 활성화의 경계 조건인지 살펴보기 위해, 추가 실험을 실시해 보았다.

서울 소재 S대학교 학부 생 86명이 추가 실험에 참여하였으며(남 46명, 여 39명, 평균나이 21.5세), 모든 참여자들은 각 수업에서 참여 점수를 획득하였다. 실험 1과

전반적인 실험 절차는 모두 같지만, 추론에 앞서 실험집단이 푸는 다른 맥락의 문제의 난이도를 높였다. 이 방법은 기존 Kahneman(2003), Shafir와 LeBoeuf(2002), Agnoli(1991)의 연구를 참고한 것으로, 확률 문제의 난이도를 높여서 전반적인 인지 자극을 유도한 것이다(Kahneman, 2003; Shafir & Leboeuf, 2002; Agnoli, 1991). 예를 들어, 실험 1에서 ‘1. 집에 도둑이 들었던 경험이 있는가?’라는 문제(표 2 참고) 대신에 ‘1. 2013년 조사 결과 절도 범죄자들의 평균 나이는 23세였다. 이 글을 본 어떤 사람이 범죄를 예방하기 위해서 대학생들을 대상으로 윤리교육을 실시해야 한다고 주장했다. 이 주장에 대해서 간단하게 평가해 보시오.’라는 문제가 제시되었고, ‘2. 2013년 현재를 기준으로 절도범죄 검거율은 얼마쯤이라고 생각하는가?’라는 문제 대신에(표 2), ‘2. 도둑들 중에서 남자의 확률은 92%이라고 하였을 때, 남자 중에서 도둑인 확률은 어느 정도라고 생각하는가?(한 사람이 남/여일 확률은 각각 50%라고 가정)’의 문제가 사전에 제시되었다.

그 결과, 위와 같이 분석적 사고를 유도하였음에도 불구하고, 실험 1의 결과가 정확하게 반복검증 되었다. 즉, 우연적 의미 활성화는 배타가설의 최종선택 패턴에 유의미한 영향을 미쳤지만($\chi^2(2) = 6.441, p < .05$), 양립가설의 최종 가설 선택에는 영향을 미치지 않았다($\chi^2(6) = 12.40, p > .05$). 반면 우연적 의미 활성화는, 배타가설의 생성 개수에는 영향을 미치지 못하였지만(제시된 가설 수: $t(52.599) = 1.191, p > .05$; 총 생성 수: $t(83) = 1.472, p > .05$), 양립가설에는 영향을 미쳤다. 양립가설의 경우 실험집단(평균 = 1.22, 표준편차 = 0.68) 보다 통제집단(평균 = 1.71, 표준편차 = 0.46)이 지문에 제시된 가설을 더 많이 고려하였으며($t(80) = 3.202, p < .01$), 실험집단(평균 = 1.98, 표준편차 = 0.58) 보다 통제집단(평균 = 2.58, 표준편차 = 0.83)이 총 더 많은 수의 가설을 생성하였다($t(32.603) = 3.237, p < .01$)⁴.

4) 지금까지의 결과분석과 관련하여 각 과제 간 차이는, 과제가 각각 다른 가설 유형을 다루고 있기 때문이 아닌(과제1: 배타가설 vs 과제2: 양립가설), 과제의 난이도 차이에서 비롯되었다는 비판이 제기될 수도 있다. Erb, Bioy 와 Hilton(2002)의 연구에서 더 복잡하고 이해하기 어려운 과제일수록, 사람들이 의미 활성화보다 과제의 내용에 더 초점을 맞추기 때문에, 의미 활성화가 미치는 영향이 감소할 수 있다는 것을 추측한 바 있다(Erb, Bioy & Hilton, 2002). 이에 본 연구의 결과가 가설들의 유형 차이 때문인지, 과제 이해도의 차이에서 비롯된 것인지를 확인해 보았다. 두 과제의 난이도 차이를 비교한 결과, 과제1(배타가설)의 참여자들의 평균 이해도는 4.15점 / 5점(표준편차 = 0.83), 과제2(양립가

추가 실험을 통해, 사전 연구와는 달리, 인지자극을 유도했을 때조차 우연적 의미 활성화가 배타가설의 최종 선택, 양립가설의 생성 수에 영향을 미치는 것을 반복해서 확인하였다. 사전 연구와 상반된 이상의 결과는 다음과 같이 사전 연구와 본 연구의 차이에 대해 의문을 야기한다.

Gilad와 Kilger(2008)의 연구에서 우연적 활성화와 더불어 사람들은 단서를 고려하여 판단하였지만, 본 연구에서 추론에 필요한 단서는 제시되지 않았다. 예를 들어, 문 앞에 있는 남자를 추론할 때(과제1) 남자를 ‘도둑’ 혹은 ‘택배기사’로 특정할 수 있는 단서들, 가령 인상착의, 특이 행동 등의 정보는 없었다. 이 차이는 우연적 활성화를 배제하기 위해서 분석적 사고와 더불어 결정적으로 다른 요인 - 단서 - 이 필요하다는 것을 시사한다. 본 연구는, 제 3의 요인은 배제한 채 진단 추론에 우연적 의미 활성화의 영향을 검증하기 위해서, 추론에 필요한 단서를 제시하지 않았다. 이러한 차이로 인해 기존 연구들과 다른 결과를 초래했을 가능성이 있다.

실 험 2

실험 1을 통해, 사람들은 진단 추론 시 우연적 활성화의 영향을 받는다는 것을 알 수 있었다. 우연적 활성화는 가설이 서로 배타적인지 양립가능한지에 따라, 최종가설 선택패턴, 가설의 생성 수에 다른 영향을 미쳤다. 이 결과는 우리의 진단 추론이 때때로 비합리적일 수 있다는 것을 의미하기 때문에, 추가 실험을 통해 인지적 노력을 유도해 보았다. 하지만 더 깊은 사고를 하는 것만으로는 우연적 활성화의 영향을 배제할 수 없었다.

설)의 평균 이해도는 3.13점 / 5점(표준편차 = 1.08)으로 이 둘의 차이는 유의미하였다 ($t(151.66)=6.771, p = 0.00$). 이처럼 두 과제 간 난이도 차이는 존재했지만, 이해 정도에 따라서 우연적 활성화의 영향이 달라지는지의 여부를 검증한 결과, 두 과제 모두 이해도에 따라 최종가설 선택패턴(과제1: $\chi^2(3) = 1.49, p > .05$; 과제2: $\chi^2(12) = 14.19, p > .05$), 생성된 가설의 개수에는 유의미한 차이가 없었다(과제1: $F(3, 57) = 1.02, p > .05$; 과제2: $F(4, 53) = 1.46, p > .05$). 이를 통해 본 연구의 결과가 과제 간 난이도 차이에 의해 발생했을 수 있다는 가능성은 제외된다.

분석적 사고가 우연적 활성화를 배제할 수 있다고 주장한 사전 연구와 본 연구의 중요한 차이는 단서의 제시 여부이다. Gilad와 Kilger(2008)의 연구에서 분석적인/직관적인 참여자들 모두 단서를 고려하여 판단했지만 본 연구에서 추론에 필요한 단서는 주어지지 않았다. 또한 판단에 필요한 단서가 없는 경우 합리적인 판단에 실패할 수 있다는 LeBoeuf와 Shafir의 연구(2003)를 고려할 때, 설사 더 깊은 사고를 하는 사람이라 하더라도 단서가 필요할 수 있다.

이상을 고려하여, 실험 2에서는 인지자극을 유도함과 더불어 단서를 제시하였다. 단서의 제시여부가 우연적 활성화의 경계조건이라면, 선행 연구들과 마찬가지로 우연적 활성화가 최종 가설 선택에 미치는 영향은 감소할 수 있을 것이다.

단서가 제시되는 경우, 또한 다음과 같은 결과를 예상해볼 수 있다: 실험 1과 추가실험에서, 양립가설의 경우 우연적 의미 활성화의 영향으로 인해 더 적은 수의 가설이 생성되었는데, Tseng와 동료(2008)들에 의하면, 단서가 투입되면 가설 생성의 고착현상이 극복될 수 있다(Tseng et al., 2008; Bassok & Novick, 2012). 뿐만 아니라, 선택된 가설에 대한 단서가 많을수록 추론 확신 역시 증가할 수 있다(Schwarz et al., 2007; Tsai et al., 2008; Tormala et al., 2002).

이상을 정리하면, 추론에 필요한 단서가 주어졌을 때 다음 세 결과가 예상된다: 1) 우연적 의미 활성화가 배타가설의 최종 선택에 미치는 영향은 감소한다. 2) 양립가설의 생성에 나타나는 고착현상은 사라질 것이며, 3) 단서의 개수가 증가 할수록 추론에 대한 확신은 높아진다. 이 예상을 검증하기 위해 실험 2를 수행하였다.

실험 방법

대상

서울 소재 S대학교 학부 생 155명이 참여하였다. ‘단서 한 개 조건’에는 80명의 실험 참여자들이(남 43, 여 36, 평균 연령 21.2세), ‘단서 세 개 조건’에는 총 75명의 실험 참여자들이(남 34, 여 41, 평균 연령 21.43세) 참여 하였다. 실험에 참여한 모든 참가자들은 수업 참여점수를 획득하였다.

절차

모든 조건은 실험 1과 동일하였지만, 한 가지 차이점은 짧은 글 속에 추론과 관련된 단서가 집단 마다 각각 한 개, 세 개가 포함되었다는 것이다. 예를 들어, 단서가 한 개 제시된 집단의 경우, 지문에 ‘도둑은 모자를 쓰고 다니고’, ‘현재 택배 차량이 아파트에 주차되어 있다’는 추론에 필요한 단서가 동시에 주어졌다(표 5). 주의해야 할 점은, 각 가설을 지지하는 단서들이 한 지문에 ‘동시에’ 제시되었다는 점이다. 즉, 실험 2은, 실험 1에서 인지자극을 유도했던 추가 실험과 절차와 문제가 모두 같지만, 표 5와 같이 지문에 가설과 관련된 단서들이 주어졌다는 유일한 차이점이 있다.

표 5. 단서 한 개 조건에서 과제1(배타가설)에 사용된 지문(지문에 도둑 관련 단서와 택배기사 관련 단서가 동시에 제시되었다.)

박 양(32)은 요즘 집을 나서기가 무섭다. 최근 며칠 동안 아파트 단지 내에 도둑이 출몰하고 있으며, **이 도둑은 주로 모자를 쓰고 다니니** 모자 쓴 사람을 조심하라는 동네 아주머니들의 주위가 계속 이어졌기 때문이다. 하지만 오늘은 부모님들께서 오후 6시경 모처럼 집에 방문한다고 하셨기에 집안 청소를 하고 장을 보러 나가야 한다. 두 명의 아이들이 현재 유치원에 가서 한가할 때 얼른 청소를 마치고, 부모님들께서 좋아하시는 닭백숙을 하기 위해 재료를 사러 집에서 20분 거리에 있는 마트에 가서 장을 봐서 2시쯤 돌아올 계획이다. 또한 아침에 아파트 경비아저씨가 재활용품을 내놓는 날이 오늘로 바뀌었다고 알려주셔서, 재활용품도 밖에 내다 놓아야 한다. 박 씨는 먼저 빨래를 세탁기에 돌리고, 간단히 설거지를 한 뒤 분리수거 할 것을 정리하고 있었는데 초인종이 울렸다. 확인해 보니 한 사람이 있었으나 **모자를 쓰고 있어서** 얼굴이 보이지 않았다. 지금 시간에 올 사람이 누굴까 생각하다가, 불현듯 며칠 전 부모님들께서 포도를 보내주신다고 말씀하셨던 것이 생각났다. 박 씨는 아주머니들의 주의도 있고 해서 확인하기 위해 **아파트 입구에 택배차량이 세워진 것을 확인한 후**, 문을 열어주기 위해 서둘러 현관문으로 향했다.

실험 결과

실험의 목적을 알아차린 1명의 피험자를 제외하고, 총 154개의 데이터가 분석되

었다.

결과1) 우연적 의미 활성화의 경계조건

사전 예상한 결과가 관찰되었다. 단서가 한 개 주어졌을 때, 우연적 의미 활성화는 최종가설 선택패턴(과제1: $\chi^2(6) = 8.768, p > .05$; 과제2: $\chi^2(6) = 3.205, p > .05$)과 가설의 총 생성 개수(과제1: $t(77) = 0.427, p > .05$; 과제2: $t(76) = 0.052, p > .05$)에 모두 영향을 미치지 못하였다.

단서가 세 개 있을 때도 역시 우연적 활성화의 영향은 최종가설 선택패턴(과제1: $\chi^2(4) = 3.741, p > .05$; 과제2: $\chi^2(6) = 7.062, p > .05$)과 가설의 총 생성 수(과제1: $t(73) = 0.163, p > .05$; 과제2: $t(71) = 0.386, p > .05$)에 영향을 미치지 않았다.

결과2) 단서의 개수의 따른 가설 생성의 변화

단서의 개수 변화에 따라서 가설 생성에 추가적인 변화를 살펴보았다. 단서 개수는 가설의 유형에 따라서 다음과 같이 최종가설 선택패턴, 추론에 대한 확신에 다른 영향을 미쳤다.

2-1) 최종가설 선택패턴

단서 개수에 따라서 실험집단1(도둑 의미 활성화)의 배타가설 추론 패턴이 달라졌다($\chi^2(2) = 8.928, p = .012$). 기존에 단서가 하나도 없을 때 배타가설의 경우 ‘도

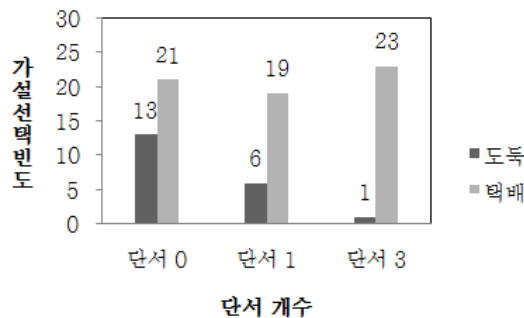


그림 6. 단서의 개수에 따른 실험집단1(도둑 의미 활성화)의 배타가설 선택 패턴

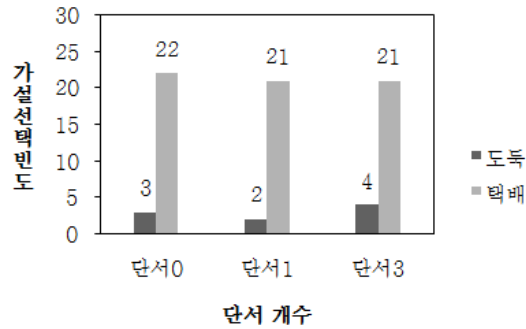


그림 7. 단서의 개수에 따른 실험집단2(택배 의미 활성화)의 배타가설 선택 패턴

‘도둑’의미 활성화의 영향을 받아 ‘도둑’ 가설이 편향되게 선택되었지만, 단서가 추가될수록 이러한 현상은 점차 감소하였다(그림 6). 하지만 그림 7을 보면 알 수 있듯이 ‘택배’ 활성화를 받았던 실험집단2에서는 단서의 투입과 상관없이 일정 수의 택배기사 가설이 최종적으로 선택됐다($\chi^2(2) = 0.595, p > .05$)(그림 7).

2-2) 추론에 대한 확신

단서의 개수가 증가함에 따라서 배타가설의 추론 확신은 높아졌다(그림 8). 실험 집단은 단서의 개수가 증가 할수록 배타가설의 추론에 더 확신을 가졌다($F(2,158)$

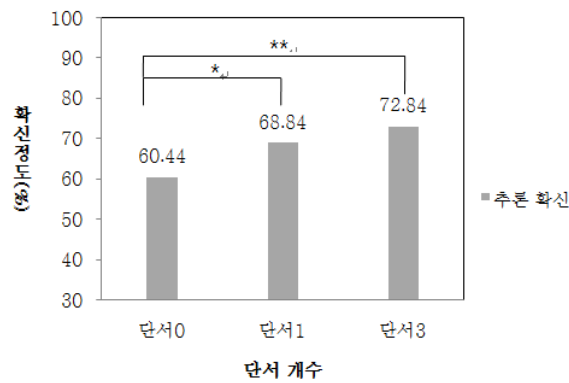


그림 8. 단서의 개수에 따른 배타가설의 추론 확신

= 7.557, $p < .001$ (단서 0: 평균 = 60.44, 표준편차 = 15.78; 단서 1: 평균 = 68.84, 표준편차 = 17.67; 단서 3: 평균 = 72.84, 표준편차 = 18.5). Tukey와 Scheffe의 사후검정 결과 단서가 하나도 없을 때와 단서가 한 개 있을 때($p < .05$), 단서가 하나도 없을 때와 단서가 세 개 있을 때($p < .01$) 추론에 대한 확신에 유의미한 차이를 확인하였다(그림 8). 하지만 양립가설의 경우 이러한 변화는 나타나지 않았다($F(2,152) = 1.009, p > .05$).

논 의

실험 2에서 우연적 의미 활성화가 가설 생성에 미치는 영향의 경계 조건을 발견하였다. 분석적 사고를 유도함과 동시에 추론에 고려해야 할 단서가 있을 때, 우연적 활성화는 배타가설의 최종 선택과, 양립가설의 가설 생성 수에 더 이상 영향을 미치지 못하였다.

그러나 본 연구와 다르게, 단서가 있을 때에도 우연적 의미 활성화가 발생했던 연구도 있다(Kusev, Schaik & Aldrovandi, 2012). 여기서 주목해야 할 점은, 본 연구에 제시된 단서의 역할이다. 기존 연구들에서는 한 가설을 지지하는 단서만 제시되었지만, 본 연구에서는 각 가설들을 지지하는 단서들이 동시에 주어졌다. 즉, 도둑 가설을 지지하는 ‘도둑은 모자를 쓰고 다니는데, 문 밖에 있는 남자는 모자를 쓰고 있다’라는 단서와 함께, 택배기사 가설을 지지하는 ‘아파트 앞에 택배차량이 있다’라는 단서가 동시에 제시되었다. 따라서 사람들은 단서를 통해 양쪽 가설을 모두 고려할 수 있었다. 이 결과는, 보다 합리적인 진단 추론을 하기 위해서, 각 가설과 관련된 단서를 모두 고려하여 추론해야 함을 시사한다. 예를 들어 법정에서, 검사는 유력한 용의자와 관련된 단서만 고려해서는 안 되고, 각 용의자들과 관련된 단서들을 모두 고려해야 할 필요가 있다. 관련된 단서들을 동시에 고려할 때, 비로소 추론과 관련 없는 요소(예, 우연적 활성화)를 무시할 수 있기 때문이다.

또한 단서 개수의 변화는 가설 유형에 따라 다른 영향을 미친다. 실험 결과 2-1를 보면 배타가설의 경우, 단서의 개수가 증가할수록 다른 맥락에서 ‘도둑’의미 활성화가 최종가설 선택에 미치는 영향은 비례해서 감소하였지만, 단서 개수의 변화

에 따라 다른 맥락에서 ‘택배기사’의 의미 활성화가 최종가설 선택에 미치는 영향은 차이가 없었다. ‘택배기사’가설과 ‘도둑’가설의 가장 큰 차이는 바로 기저율로, 일반적으로 집에 택배기사의 방문율이, 도둑의 방문율 보다 더 높다. 통제집단의 응답을 분석해본 결과 도둑의 기저율은 약 22%, 택배기사의 기저율은 약 76%임을 확인할 수 있었다. 이와 같은 결과는, 기저율이 높은 가설과 낮은 가설을 각각 지지하는 단서가 있을 때, 기저율이 높은 가설을 지지하는 단서가 더 비중 있게 다루어진다는 해석을 가능하게 한다. 이와 관련하여 보다 상세한 논의는 종합논의에서 다루도록 하겠다.

결과 2-2에서 단서의 개수가 증가할수록 추론에 대한 확신이 높아진다는 것이 확인됐다. 사전 연구에서는 밝히지 못했지만 본 연구에서 추가로 발견한 사실은, 단서가 증가할수록 확신이 증가하는 경우는 가설들이 서로 배타적일 때뿐이라는 것이다(과제1). 이 결과는 가설의 유형에 따라 특정 요인이 미치는 영향이 달라질 수 있다는 본 연구의 주장을 더욱 공고히 해준다.

종합논의

본 연구는, 가설의 유형에 따라, 우연적 의미 활성화가 가설 생성에 다른 영향을 미치는 것을 검증하였다(실험 1). 우연적 활성화는 더욱 깊은 사고를 할 때조차 추론에 영향을 미치지 않지만(실험 1), 추론에 고려해야 할 단서들이 동시에 주어지면 사라진다(실험 2). 이상의 연구 결과에 대한 추가 해석과 본 연구의 의의, 그리고 후속연구 주제에 대한 논의는 다음과 같다.

가설의 기저율에 따른 단서의 영향

실험 2에서, 배타가설의 최종선택에 우연적 의미 활성화가 미치는 영향은 단서들의 개수가 증가할수록 감소했는데, 이 현상은 상대적으로 기저율이 낮았던 ‘도둑’의 의미 활성화를 받은 집단에서만 유의미하였다(그림 6 참고). 본 연구는 한 가설을 지지하는 단서만을 제시했던 선행 연구들과는 다르게, 지문에 등장했던 두 가설들을 각각 지지하는 단서들을 동시에 제시하였다(표 5 참고). 그 결과, 단서가 없

을 때 베타 가설의 최종 선택에 작용했던 우연적 활성화의 영향은 사라졌으며, 이와 더불어 단서가 한 개에서 세 개로 증가했을 때 기저율이 낮은 가설(도둑)의 선택은 더욱 감소하였다. 이 결과는 두 가설을 각각 지지하는 단서가 동시에 주어졌을 때, 사람들은 기저율이 낮은 가설(도둑)을 지지하는 단서보다, 기저율이 높은 가설(택배기사)을 지지하는 단서에 더 큰 비중을 두고 판단한다는 것을 보여준다. 왜 이러한 현상이 관찰된 것일까?

한 가지 가능한 해석으로, ‘확증편향(confirmation bias)⁵⁾성’을 생각해 볼 수 있다: 사람들은 일차적으로 떠올린 택배기사 가설을 지지하는 단서만 고려하고, 도둑 가설을 지지하는 단서는 무시한 것이다. 일반적으로, 기저율이 높은 가설은 기저율이 낮은 가설보다 참일 가능성이 더 높다. 즉, 본 연구에서 기저율이 높은 가설이었던 택배기사는 기저율이 낮은 도둑 가설 보다 가설 선택에 있어서 우위성을 가진다. 하지만 본 연구의 실험 1에서 택배기사 가설의 참/거짓 여부를 검증할 수 있는 단서는 없었다. 이로 인해 사람들은 이 기저율이 높은 가설에 대한 가설 검증을 할 수 없었고, 추론 맥락과 상관없는 우연적 활성화에 따라 더 쉽게 생각나는 가설을 선택하였다. 그러나 실험 2에서 각 가설과 관련된 단서가 주어졌을 때, 사람들은 우선순위에 놓여있는, 기저율이 높은 가설을 확증할 수 있는 단서에 더 비중을 두었다. 처음 떠올린 택배기사 가설을 지지하는 단서만 고려한 채, 이를 반증할 수 있는 도둑 가설을 지지하는 단서는 무시한 것이다(확증 편향). 이 과정에서 우연적 활성화는 더 이상 의미 있는 영향을 미치지 못하였다.

이 결과는, 단서가 우연적 활성화를 배제시킬 수 있음에도 불구하고, 우리는 여전히 합리적으로 진단 추론을 하지 못할 수 있음을 경고한다. 각 가설과 관련된 단서가 확보되었을 경우라도 사람들은 단서를 활용하는 데 편향적이다.

우연적 활성화와 추론에 대한 확신

추가로 더 고려해 보아야 할 사항은, 사전 의미 활성화와 반대로 답을 한 사람들의 확신이 더 낮다는 연구 결과이다(Erb, Bioy & Hilton, 2002). Erb와 동료들의 실험에서 ‘위험 추구 단어’를 사전에 보았던 실험 참여자들 중, ‘안정된 옵션’을 선

5) 확증편향이란 처음에 떠올린 가설의 가능성을 시험할 때, 그 가설을 지지하는 단서만을 고려하는 인지 편향성을 의미한다(Synder & Swann, Jr, 1978).

택한 실험 참여자들의 확신은 ‘불확실한 옵션’을 선택한 실험 참여자들보다 더 낮았다. 본 연구에서도 이러한 결과가 반복되는지 알아보기 위해 추가 분석을 실시해 보았다. 그러나 본 연구에서 이상의 결과는 관찰되지 않았다. 예를 들어, ‘도둑’ 의미 활성화가 있는 집단에서 도둑 가설을 최종 선택한 사람의 확신과, 택배기사를 최종 선택한 사람들의 확신 간에 유의미한 차이는 없었다($t(32) = 0.282, p > .05$).

기존 연구와 본 연구의 결과가 일치하지 않는 이유는, 활성화 대상에 차이가 있기 때문으로 생각된다. Erb와 동료들의 연구에서 의미 활성화는(위험 추구 의미), 해결 옵션에 직접 영향을 미친 것이 아니라, 감정에(위험 추구) 영향을 미친 것이었다. 즉, ‘진취적인’이라는 의미 활성화는, 사람들이 위험을 추구하려는 감정에 영향을 주어, 불확실한 해결책(예, 이틀 연속 파업할 수 있는 기차)의 선택을 높였다. 반면 본 연구에서 ‘도둑’ 의미 활성화는 도둑 가설에 직접적인 영향을 주었다. 결국 기존 Erb의 연구에서 사전 의미 활성화와 반대로 판단한 집단의 확신이 떨어진 것은, 활성화의 의미와 일관된 옵션을 선택하지 않았기 때문이 아니라, 감정에 반대되는 행동을 했기 때문일 가능성이 있다.

연구의 의의

본 연구는 실제 상황에서 발생할 수 있는 가설생성과정을 다루었다. 본 연구의 결과는 이상적인 상황에서 가설생성을 다루었던 HyGene모델이 다루지 않은 부분으로, 배타가설만을 두고 의미 활성화의 영향을 다루었던 선행 연구들만으로 예측되지 않는다.

후속 연구를 통해 보다 세부적인 과정이 밝혀져야겠지만, 본 연구는 추론 직전에 발생했던 의미 활성화로 인해 사람들이 비합리적으로 진단추론을 할 수 있음을 경고한다. 단서에 기초하지 않고 내리는 결정은, 더 깊은 사고를 하더라도 추론과 다른 맥락에서 비롯된 의미 활성화의 영향에 취약하기 때문에, 가설을 생성할 때는 반드시 각 가설들과 관련된 단서들에 주의를 기울여야 한다.

이 밖에도 본 연구는, 앞으로 가설생성과 평가를 다루는 연구에서 가설들의 유형 차이를 고려해야 할 필요성을 제기한다. 많은 선행 연구들은 배타가설만을 다루어왔기 때문에, 이 결과가 다른 유형의 가설에도 역시 적용될 수 있을지는 의문

이다. 일상생활에서는 베타가설뿐 아니라 다양한 특성을 지닌 가설들이 존재한다. 이를 고려하지 않고 계속 베타 가설만을 다룬다면, 연구 결과를 일반화시키는 데 어려움이 있을 것이다.

후속 연구 제안

본 연구는 가설 유형에 따른 의미 활성화 영향의 차이를 주장하였다. 그러나 아직 두 개의 가설 유형 (베타가설 vs 양립가설) 차이만 검증했을 뿐이다. 후속 연구를 통해 베타적이거나 양립 가능한 특성을 지닌 가설 이외에, 다양한 특성을 가진 가설들을 고려하여(예를 들어, 세 개 이상의 답이 동시에 가능하거나, 최종 선택된 가설이 더 이상 유효하게 되지 않을 경우), 각각의 가설 생성 과정을 더 연구해 볼 필요가 있다. 아울러 더 다양한 상황에서 본 결과를 반복 검증할 필요가 있다.

마지막으로, 본 연구에서 가설 유형에 따른 우연적 활성화의 영향을 확인했지만, 이러한 현상을 야기한 메커니즘을 다루지는 못했다. 앞서 논의한 부분을 바탕으로 후속 연구를 통해 이에 대한 연구를 진행할 예정이다.

참고문헌

- Agnoli, F. (1991). Development of judgmental heuristics and logical reasoning: Training counteracts the representativeness heuristic. *Cognitive Development*, 6, 195-217
- Alter, A. L. & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*.
- Bassok, M. & Novick, L. R. (2012). Problem solving. *Oxford handbook of thinking and reasoning*, 413-432.
- Dougherty, M. R. & Hunter, J. (2003). Probability judgment and subadditivity: The role of working memory capacity and constraining retrieval. *Memory & Cognition*, 31(6), 968-982.
- Erb, H. P., Bioy, A., & Hilton, D. J. (2002). Choice preferences without inferences: Subconscious priming of risk attitudes. *Journal of Behavioral Decision Making*, 15(3),

251-262.

- Fenton, N. & Neil, M. (2012). Risk assessment and decision analysis with Bayesian networks. CRC Press.
- Gilad, D. & Kliger, D. (2008). Priming the Risk Attitudes of Professionals in Financial Decision Making. *Review of Finance*, 12(3), 567-586.
- Ho, J. L. & Keller, L. R. (1994). The effect of inference order and experience-related knowledge on diagnostic conjunction probabilities. *Organizational Behavior and Human Decision Processes*, 59(1), 51-74.
- Josephson, J. R. & Josephson, S. G. (Eds.). (1996). *Abductive inference: Computation, philosophy, technology*. Cambridge University Press.
- Kahneman, D. & Miller, D. T. (1986). Norm theory: comparing reality to its alternatives. *Psychological Review*, 93, 136-153.
- Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9), 697.
- Kelley, C. M. & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, 32(1), 1-24.
- Kusev, P., van Schaik, P., & Aldrovandi, S. (2012). Preferences induced by accessibility: Evidence from priming. *Journal of Neuroscience, Psychology, and Economics*, 5(4), 250.
- Lange, N. D., Thomas, R. P., Buttaccio, D. R., Illingworth, D. A., & Davelaar, E. J., (2013). Working memory dynamics bias the generation of beliefs: The influence of data presentation rate on hypothesis generation. *Psychonomic bulletin and review*, 20:171-176.
- LeBoeuf, R. A. & Shafir, E. (2003). Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making*, 16(2), 77-92.
- Libby, R. & Frederick, D. M. (1990). Experience and the ability to explain audit findings. *Journal of Accounting Research*, 348-367.
- Libby, R. (1985). Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research*, 648-667.

- Lombrozo, T. (2006). The structure and function of explanations. *Trends in cognitive sciences*, 10(10), 464-470.
- Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive psychology*, 55(3), 232-257.
- Lombrozo, T. (2012). Explanation and abductive inference. *Oxford handbook of thinking and reasoning*, 260-276.
- Malt, B. C., Ross, B. H., & Murphy, G. L. (1995). Predicting features for members of natural categories when categorization is uncertain. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(3), 646.
- McKenzie, C. R. (1998). Taking into account the strength of an alternative hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(3), 771.
- Mehlhorn, K., Taatgen, N. A., Lebiere, C., & Krems, J. F. (2011). Memory activation and the availability of explanations in sequential diagnostic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1391.
- Pennington, N. & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 521.
- Reggia, J. A., Nau, D. S., & Wang, P. Y. (1985). A formal model of diagnostic inference. I. Problem formulation and decomposition. *Information Sciences*, 37(1), 227-256.
- Reyes, R. M., Thompson, W. C., & Bower, G. H. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology*, 39(1), 2.
- Ross, L. D., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, 35(11), 817.
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14(4), 332-348.
- Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public

- information campaigns. *Advances in experimental social psychology*, 39, 127.
- Shafir, E. & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology*, 53, 419-517.
- Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgements of likelihood. *Cognition*, 52(1), 1-21.
- Smith, S. M. & Levin, I. P. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making*, 9(4), 283-290.
- Snyder, M. & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36(11), 1202.
- Sprenger, A. & Dougherty, M. R. (2012). Generating and evaluating options for decision making: the impact of sequentially presented evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 550.
- Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, 23(1), 53-62.
- Thagard, P. (1997). Probabilistic Networks and Explanatory Coherence. *Automated Abduction: Inference to the best explanation*.
- Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review*, 115(1), 155.
- Tormala, Z. L., Petty, R. E., & Briñol, P. (2002). Ease of retrieval effects in persuasion: A self-validation analysis. *Personality and Social Psychology Bulletin*, 28(12), 1700-1712.
- Tsai, C. I., Klayman, J., & Hastie, R. (2008). Effects of amount of information on judgment accuracy and confidence. *Organizational Behavior and Human Decision Processes*, 107(2), 97-105.
- Tseng, I., Moss, J., Cagan, J., & Kotovsky, K. (2008). The role of timing and analogical similarity in the stimulation of idea generation in design. *Design Studies*, 29(3), 203-221.
- Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2), 207-232.

- Van Wallendaal, L. R. & Hastie, R. (1990). Tracing the footsteps of Sherlock Holmes: Cognitive representations of hypothesis testing. *Memory & Cognition*, 18(3), 240-250.
- Weber, E. U., Böckenholt, U., Hilton, D. J., & Wallace, B. (1993). Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(5), 1151.
- Zuckerman, M., Knee, C. R., Hodgins, H. S., & Miyake, K. (1995). Hypothesis confirmation: The joint effect of positive test strategy and acquiescence response set. *Journal of Personality and Social Psychology*, 68(1), 52.

1차원고접수 : 2015. 02. 23

1차심사완료 : 2015. 04. 05

2차원고접수 : 2015. 05. 12

2차심사완료 : 2015. 05. 27

게재확정일 : 2015. 06. 22

(Abstract)

The Effect of Incidental Semantic Activation on Hypothesis Generation: Exclusive vs Compatible Hypotheses

Younha Lee

Jooyong Park

Interdisciplinary Program in Cognitive Science, Psychology, Seoul National University

Previous studies on the effect of incidental semantic priming on judgment, have focused mainly on mutually exclusive hypotheses. However, the present study explored whether incidental semantic activation affects diagnostic inference depending on the type of the hypothesis: mutually exclusive hypotheses vs compatible hypotheses. In Experiment 1, in case of mutually exclusive hypotheses, the final hypothesis was selected according to the incidental semantic priming, but there was no difference in the number of generated hypothesis in comparison with the control. However, for compatible hypotheses (i.e., both hypotheses can be true), the semantic priming affected the number of generated hypotheses, but not the selection of the final hypothesis. The same pattern of results was observed even when the cognitive load was increased. In Experiment 2, we found a boundary condition of incidental semantic activation on diagnostic inference. When cues related to each of the hypotheses were presented simultaneously, the incidental semantic effect disappeared. These results suggest that people consider all possible cues when making diagnostic inference in daily life. In light of these findings, further research on hypothesis generation/evaluation should take the type of hypothesis into account.

Key words : diagnostic inference, hypothesis generation, type of hypothesis, priming, cue