



저작자표시-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

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

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



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

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

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

인지과학석사학위논문

다른 맥락에서의 의미 활성화가
가설 생성에 미치는 영향

2014년 8월

서울대학교 대학원

협동과정 인지과학 인지심리학 전공

이 윤 하

초 록

진단적 추론이란 결과로부터 가능한 원인을 추론하는 인지활동을 가리킨다. 불확실한 상황에서 추론은 가설에 의존하기 때문에, 가설생성은 진단적 추론의 가장 기본이자 핵심이 되는 인지 과정이다. 판단과 다른 맥락에서의 의미의 활성화가 이후 문제에 영향을 미친다는 연구가 있기는 하지만, 진단적 추론 상황에서 이러한 활성화의 영향에 대한 연구는 찾아보기 어렵다. 더욱이 기존의 가설 생성 연구에서는 가설의 유형 차이를 고려하지 않았다. 가설 유형은 가설들이 서로 배타적인지, 혹은 양립 가능한지로 구분할 수 있다.

본 연구에서는 진단적 추론과 다른 맥락에서 활성화된 의미가 가설의 유형에 따라 가설생성에 어떠한 영향을 미치는지 알아보려고 하였다. 구체적으로 활성화가 가설들이 서로 배타적인지 아니면 양립가능한지에 따라, 최종 가설 선택 패턴, 가설의 생성 수 그리고 추론의 확신에 다른 영향을 미칠 가능성을 탐색하였다. 이와 함께 주어진 단서의 개수에 따라서 추가적으로 가설생성에 어떤 변화를 야기하는지도 검증하였다.

이상의 목적을 위해 세 개의 실험이 수행되었다. 첫 번째 실험에서는 진단적 추론을 하게 하기 전에, 단순한 문제에 답을 하게 하여 한 가설을 직접적으로 활성화시켰다. 실험 참여자들은 짧은 지문을 읽은 후 추론 문제를 풀었는데, 실험집단은 추론과 다른 맥락의 문제를 사전에 풀게 하여 특정 의미를 활성화시켰다. 그 결과 배타적인 가설의 경우, 사전 활성화가 최종 가설 생성 패턴에 영향을 미쳤지만, 가설의 생성 개수에는 영향을 미치지 않았다. 반면 양립 가능한 가설의 경우, 활성화가 최종 가설 생성 패턴엔 영향을 미치지 않았지만, 가설의 생성 수에는 영향을 미쳤다. 가설 유형에 상관없이 사전 활성화는 추론의 확신에 유의미한 영향을 미치지 않았다.

실험 2와 실험 3에서는 다른 맥락에서 활성화된 의미가 추론에 미치는 영향의 경계조건(boundary condition)을 탐색하고자 하였다. 실험 2에서 진단적 추론에 앞서 실험 1에서 보다 더 어려운 문제에 답을 하도록 하여, 인지적 부담을 가중시켰다. 만일 실험 1의 결과가 피상적 사고로 인해 얻어진 결과라면, 더 느리고 분석적인 사고 (System 2)를 유도하는 상황에서 활성화의 영향이 감소할 것으로 예상할 수 있다. 하지만 더 깊은 사고를 하도록 유도했음에도 불구하고, 실험 1의 결과가 반복 검증 되었다.

실험 3에서는 다른 맥락의 활성화와 더불어 추론 맥락에서 고려되어야 하는 단서를 제공하였다. 단서는 여러 수준에서 조작되었는데, 결과적으로 추론에 고려해야 하는 관련 단서가 하나라도 제시된다면, 사전 활성화의 영향이 사라진다는 것을 확인하였다. 단서의 개수가 증가함에 따라 양립 가능한 가설의 경우에는 더 많은 수의 가설이 생성되었고, 배타적인 가설의 경우에는 추론의 확신이 높아졌다.

이상의 결과를 종합하면, 본 연구를 통해 추론과 다른 맥락에서 발생한 활성화가 진단적 추론에 영향을 미친다는 것과, 두 가설이 서로 양립가능한지 혹은 서로 배타적인지에 따라서 가설 생성 과정이 다르다는 것이 확인됐다. 추론과 다른 맥락에서 발생한 의미 활성화는 직관적인 사고를 할 때뿐 아니라 분석적인 사고를 할 때조차 두 가설의 유형에 따라서 최종 가설의 선택 패턴, 가설의 생성 수에 다른 영향을 미친다. 하지만 추론에 고려해야 하는 진단 단서가 포함되었을 때 활성화의 영향은 사라지며, 단서 개수의 증가는 두 가설의 유형에 따라 추론 확신과 가설의 생성 수에 다른 영향을 미친다. 본 연구의 결과는 일상 상황에서 발생할 수 있는 진단적 추론의 오류를 경고하며, 가설 생성/가설 평가를 다루는 연구들에서 가설 유형에 따른 차이를 고려해야 할 필요성을 제기한다.

.....
키워드: 진단적 추론, 가설 생성, 가설 유형, 프레이밍, 단서

학번: 2012-22916

목 차

서론	7
선행 연구 및 이론적 배경	10
실험 1	17
실험 방법	17
실험 결과	21
논의	24
실험 2	26
실험 방법	27
실험 결과	29
논의	31
실험 1과 실험 2에 대한 논의	32
실험 3	34
실험 방법	35
실험 결과	38
논의	41
종합논의	43
참고문헌	51
부록	57

표 목 차

[표 1] <과제 1>에서 사용된 지문	-----	17
[표 2] <과제 2>에서 사용된 지문	-----	18
[표 3] <과제 1>과 <과제 2>에서 사용된 문제들	-----	19
[표 4] 실험 1에서 사용된 진단적 추론과 다른 맥락의 문제들	-----	19
[표 5] 실험 1 조건 간 차이 정리	-----	20
[표 6] 실험 2에서 사용된 진단적 추론과 다른 맥락의 문제들	-----	28
[표 7] 실험 1과 실험 2의 결과 요약	-----	32
[표 8] 실험 3에서 사용된 단서가 포함된 과제 1의 지문	-----	35
[표 9] 실험 3에서 사용된 단서가 포함된 과제 2의 지문	-----	36
[표 10] 실험 3에서 각 조건마다 주어진 자극의 차이	-----	37
[표 11] 전체 실험 결과 정리	-----	43

그림 목 차

[그림 1] HyGene모델에서 외부 사건(external events)과 심적 표상(one' s mental representation of the external events)의 관계 -----	11
[그림 2] 과제 1(배타적인 가설)에서의 최종 가설 선택 패턴 -----	21
[그림 3] 과제 2(양립 가능한 가설)에서의 최종 가설 선택 패턴 -----	21
[그림 4] 가설 유형에 따른 가설 생성 수의 차이 -----	22
[그림 5] 과제 1(배타적인 가설)에서의 최종 가설 선택 패턴 -----	29
[그림 6] 과제 2(양립 가능한 가설)에서의 최종 가설 선택 패턴 -----	29
[그림 7] 가설 유형에 따른 가설 생성 수의 차이 -----	30
[그림 8] 단서의 개수에 따른 실험집단 1의 배타적인 가설 선택 패턴 ----	39
[그림 9] 단서의 개수에 따른 실험집단 2의 배타적인 가설 선택 패턴 ----	39
[그림 10] 단서의 개수에 따라 총 생성된 양립 가능한 가설의 수 -----	39
[그림 11] 단서의 개수에 따른 배타적인 가설의 추론 확신 정도의 차이 ----	40
[그림 12] 제 3가설의 생성 현황 -----	46

서론

어제까지 잘 달리던 자동차가 오늘 아침에 시동이 걸리지 않는다면, 운전자는 무슨 생각을 할까? 한 가지 가능한 추론은 자동차 배터리에 문제가 있을지 모른다는 것이다. 이와 비슷하게, 출근길에 젖은 도로를 본다면, 우리는 어제 저녁이나 새벽에 비가 내렸을 것이라고 추론할 것이다. 이러한 추론을 진단적 추론 (diagnostic inference)이라고 부르는데, 진단적 추론은 일상적 추리 (everyday reasoning)에서 자주 활용된다. 어떤 이는 진단적 추론을 귀추적 추론 (abduction inference)이라고 부르기도 한다.

철학자이자 논리학자인 Charles Sanders Peirce (1839-1914)은 연역적 추론 (deductive inference)과 귀납적 추론 (inductive inference)이외에 귀추적 추론 (abductive inference, retroductive inference)이 있음을 주장하였다. 19세기 철학자이자 생리학자인 Herman von Helmholtz (1821-1894) 역시 사람들이 지각하는 방식을 귀추적인 추론 과정으로 보았으며, 이 과정에서 적절한 세트의 가설집합은 이미 우리 내부에 가지고 있음을 가정한 바 있다 (이 부분은 다음 장에서 조금 더 논의될 것이다). 컴퓨터 과학자인 Charniak와 McDermont(1985)는 귀추적 추론을 거꾸로 가는 전건긍정법칙으로 (as modus ponens turned backward) 설명하였다. 그들은 이 추론 과정이 1)어떤 현상에 대한 원인 추론, 2)자료에 대한 기술, 3)가장 좋은 설명의 추론 (inference to the best explanation)으로 일어난다고 주장하였다 (Josephon & Josephon, 1996). 요컨대 귀추적 추론은 어떤 결과 (effect, data, cue)로부터, 현상을 가장 잘 설명할 수 있는 원인 (cause)을 찾아내는 인지 활동을 가리킨다.

귀추적 추론은 가장 그럴 듯한, 현재 현상을 가장 잘 설명하는 가설을 추론하는 방식이지만, 이 추론은 항상 참을 보장하지는 못한다. 자동차에 시동이 걸리지 않을 때 배터리가 없다고 추론하는 것, 바닥이 젖어있을 때 새벽에 비가 왔을 것이라고 추론하는 것은 맞을 가능성이 높지만, 틀릴 가능성 역시 존재한다. 하지만 우리가 잠정적으로 이러한 원인을 생각하는 이유는 현재로서 이 보다 더 좋은 설명을 아직 찾아내지 못했기 때문이다. 설명은 가설의 형태로 표상되며, 가설 생성은 진단적 추론의 핵심이다.

인간의 가설 생성 과정을 설명하는 모델 중 하나는 기억 연구에서의 발견을 확장한 HyGene모델이다. HyGene모델에 의하면, 가설생성은 기억에 의존한다. 진단 상황에 주어진 단서들은 관련된 가설들을 기억에서 활성화시키며, 가장 크게 활성화된 가설은 최종 가설로 선택된다. 많은 선행 연구들에서 기억에서의 가용성

(availability)과 가설생성의 관계에 대해서 다루어 왔다. 판단의 오류는 직관적으로 떠오르기 쉬운 가설들을 택함으로써 발생할 수 있다. 이 밖에도 단서들의 순서나 생생함 (vividness) 등을 조작하여 한 가설의 활성화를 일시적으로 바꿀 수 있다는 것이 보고되었다. 최근의 제시되거나, 눈에 띄는 단서와 관련된 가설들은 최종 가설로 더 많이 선택되고, 더 잘 떠올려지는 가설을 선택할수록 추론의 확신은 증가한다.

하지만 진단적 추론에서 다른 맥락 (context)이 한 가설의 활성화에 미치는 영향을 검증한 연구는 거의 없었다. 사전에 제시된 문제로 인해 고차적 사고 (high-order thinking)를 요하는 과제에서도 점화 효과 (priming effect)가 관찰되었음을 고려해 볼 때, 진단적 추론에서도 다른 맥락에서 활성화된 의미가 이후 가설생성에 영향을 미칠 수 있다. 본 연구는 일차적으로 이 가능성을 탐색하기 위해 수행되었다.

가설의 중요한 기능 중 하나는 설명이며, 가설 생성은 가장 좋은 설명을 추론 (inference to the best explanation)하는 심리적 활동이다. 설명은 모두 같은 구조와 기능을 가지고 있지 않으며, 맥락에 따라서 선호하는 설명이 달라질 수도 있다. 예를 들어, Lombrozo(2007)의 연구에서 하나의 질병 (simple explanation)과 두 가지 질병의 조합 (complex explanation)이 있을 때, 사람들은 기저율 정보의 유무에 따라서 다른 설명을 선택하였다. 뿐만 아니라 설명은 유형에 따라서 서로 경쟁하거나 서로 양립될 수 있기도 하다. Sloman (2004)에 따르면, 설명이 서로 경쟁할 때 한 설명의 가능성이 높아지면 다른 설명의 가능성은 줄어들거나 혹은 한 설명이 다른 설명의 생성을 억제할 수 있다. 이를 고려해 보면 가설의 유형에 따라서 가설 생성 과정이 다를 가능성이 있지만, 가설 생성을 다룬 선행 연구들에서는 가설의 유형 차이를 고려하지 않았다.

종합하자면 본 연구는 크게 두 가지 목적을 가지고 있다. 하나는 진단적 추론과 다른 맥락에서 활성화된 의미가 가설 생성에 영향을 미칠 수 있는지 알아보는 것이며, 다른 하나는 가설이 서로 배타적인지/양립가능한지에 따라 가설 생성 과정이 다르다는 것을 밝히는 것이다. 이를 검증하기 위해 본 연구의 가설은 다음과 같다.

가설: 추론과 다른 맥락에서 발생한 의미 활성화는 두 가설이 서로 배타적인지 (mutually exclusive) 혹은 서로 양립될 수 있는 지에 따라 가설 생성 과정에 다른 영향을 미친다.

두 가설이 서로 배타적일 때 활성화를 통해 한 가설의 가능성이 높아지면 다

른 가설의 가능성은 낮아질 것이다. 반면 두 가설이 서로 양립할 수 있으며, 경우에 따라서 수렴할 수도 있다면 한 가설의 가능성의 증가는 다른 가설의 가능성을 낮추지 않을 것이다. 이러한 차이로 인해, 최종 생성된 가설의 수에서는 물론 최종 선택된 가설에 대한 확신의 정도에서도 차이가 나타날 것이다. 총 세 번의 실험은 이 가능성을 검증하기 위해 실행되었다.

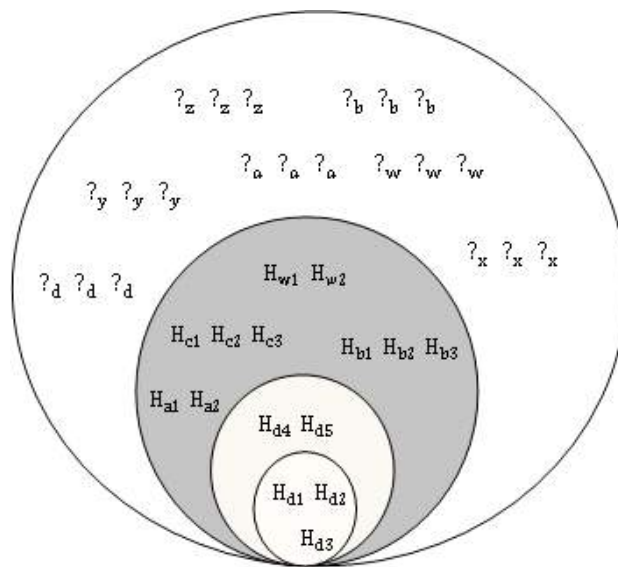
선행 연구 및 이론적 배경

기존의 추론과 판단 연구는 가설 생성 (hypothesis generation)보다는 연구가 상대적으로 용이한 가설 평가 (hypothesis evaluation)와 정보 선택 (information selection)에 초점을 맞추어 왔다 (Weber, Bockenholt, Hilton & Wallace, 1993; Ho & Keller, 1994; Dougherty & Sprenger, 2006; Tversky & Koehler, 1994; Baron & Hershey, 1988; Wedell, 2011). 하지만 가설 생성은 고려되어야 할 가설의 수를 포함하여, 이후 가설의 확률 평가, 가설 검증을 위한 정보 탐색 과정에 핵심이 되는 과정이다 (Thomas, Dougherty, Sprenger & Harbison, 2008; Mehle, Gettys, Manning, Baca & Fisher, 1981; barrows, Norman, Neufeld & Feightner, 1982; Lange, Thomas & Davelaar, 2013; Gettys & Fisher, 1979). 적어도 한 개 이상의 가설이 생성되었을 때 비로소 평가와 판단이 가능해지고, 처음 생성된 가설이 무엇인지에 따라 정보 수집, 최종 판단이 달라지기 때문이다 (Baron, Beattie & Hershey, 1988; Slowiaczek, Klayman, Sherman & Skov, 1992; Klayman & Ha, 1987; Zuckerman, Knee, Hodgins, & Miyake, 1995; Wason, 1960; Weber, Bockenholt, Hilton, & Wallace, 1993).

가설 생성은 데이터에 기반 하여 설명을 생성하는 영역에서 모두 적용될 수 있다. 예컨대 환자의 증상을 관찰하여 진단을 생성하는 의학 분야 (Weber, Bockenholt, Hilton, & Wallace, 1993; Kassirer & Kopelman, 1989; Baron & Hershey, 1988), 신경 정보에 기반을 두어 진단을 내리는 신경 과학 분야 (Barrows & Bennett, 1972), 회계 장부에 기초하여 문제를 발견하는 회계 분야 (Libby, 1985; Ho & Keller, 1994; Waller & Felix, 1989; Libby & Fredrick, 1990), 자동차 이상의 원인을 진단하는 공학 분야 (Mehle, 1982), 행동 패턴에 기반을 두어 성격 특성을 설명하는 사회 분야 (Zuckerman, Knee, Hodgins & Miyake, 1995; Strack, Schwarz, Bless, Kubler & Wanke, 1993; Higgins & Bargh, 1987) 주어진 증거를 기반으로 유/무죄를 판결하는 법학 분야 (Pennington & Hastie, 1988; Fein & McCloskey, 1997; Reyes, Thomson & Bower, 1980) 가설 생성/평가 과정을 바탕으로 진단 전문가 시스템 (diagnostic expert system)을 구현하는 컴퓨터 과학 분야 (Reggia, Nau & Wang, 1985a; Reggia, Nau, Wang & Peng, 1985b) 등에서 모두 가설생성 과정을 다루어 왔다 (Thomas, Dougherty, Sprenger & Harbison, 2008).

1. 기억과 가설생성

다양한 영역에서 이루어진 가설생성에 대한 연구들로 인해, 가설 생성 과정이 우리의 기억과 밀접한 관련이 있다는 것은 잘 알려졌다. 기억과 가설 생성의 계산론적 모형의 하나는 HyGene모형이다 (Thomas, Dougherty, Sprenger & Harbison, 2008). 이 모형에 따르면 가설 생성은 우리의 기억에 많이 의존하는데 그 세부 처리 과정을 소개하면 다음과 같다 (그림 1). 그림 1에서, 큰 원안에 있는 모든 외부 사건은 가능한 상태의 전체 집합 (the universe of possible states)으로, 우리는 학습을 통하여 일부의 사건을 배워 내재화 한다. ‘가설’은 모든 외부 사건들 중에서 의미 지식 (semantic knowledge)으로 표상된 사건들을 지칭한다. 가장 바깥에 위치하고 있는 첫 번째 원 안에 있는 ‘?’은 외부 환경에 있는 지식들을 의미하고, 우리가 아직 배우거나 경험하지 않아서 내재화되지 못한 사건들이기 때문에 이것의 정체를 아직은 알 수 없다. 이들 중에서 학습과 경험을 통해 알게 된 지식들은 비로소 가설 ‘H’로 기억 속에 표상되며 (바깥으로부터 두 번째 원), 아래 첨자 w, c 등은 데이터w, 데이터c을 나타낸다. 예를 들어, H_c은 데이터c와 관련된 가설임을 의미한다. 만약 d라는 데이터가 주어지면 d와 관련된 가설인 H_{d1},H_{d2},H_{d3},H_{d4},H_{d5}가 활성화 되고 (세 번째 원), 이 중 몇 개의 가설 (H_{d1},H_{d2},H_{d3})만이 제한된 용량을 지닌 우리의 작업 기억 (working memory)안에 들어와 (네 번째 가장 작은 원) 의식적으로 처리된다.



[그림 1] HyGene모형에서 외부 사건(external events)과 심적 표상(one's mental representation of the external events)의 관계

HyGene모델이 가정하는 가설생성 과정을 살펴보면, 1)어떤 사건, 단서(d)가 주어졌을 때¹⁾ 2) 이와 관련된 지식 ($H_{d1}, H_{d2}, H_{d3}, H_{d4}, H_{d5}$)이 활성화 되고, 3) 이들 중 가장 크게 활성화 된 가설은 최종가설로 선택된다. HyGene모델에서는 관찰된 단서를 기반으로 장기기억으로부터 관련된 가설이 활성화 된다는 것을 전제한다. 한 가설을 지지하는 단서가 많을수록 그 가설이 최종적으로 선택될 가능성은 높아지며, 이와 더불어 사람들은 높은 기저율을 (경험을 통해서 획득한 일화 기억들) 가진 가설을 더 많이 생성한다 (Dougherty & Hunter, 2003a; Weber, Böckenholt, Hilton, & Wallace, 1993, Dougherty, Gettys & Thomas, 1997; Getty & Fisher, 1987; Schwarz, 2004; Libby, 1985).

이와 같이 가설들은 이미 기억 속에 있는 지식들로부터 생성되며, 19세기 철학자인 헬름홀츠(Herman von Helmholtz)도 이와 비슷한 주장을 하였다. 그는 외부의 물체가 무엇인지 ‘알기(perception)’ 위해서는 물리적 상태를 보고 추론하는 과정이 필요하며 (Lombardo, 1987; Pastore, 1971), 이러한 귀추적 추론에 필요한 적절한 가설 세트는 이미 우리 내부(기억)에 있어야 한다는 것을 주장하였다. Weber, Bockenholt, Hilton 그리고 Wallace (1993)도 가설은 의미 지식 (semantic knowledge) 과 일화기억 (episodic memory)을 활용하는 사고 과정을 통해 생성된다고 주장하였으며, Libby (1985) 역시 진단적 모델에서 가설생성의 핵심은 환경에서 주어진 단서와 지식 구조와의 상호작용이라고 주장하였다. 이 밖에 다른 연구들에서 가설 생성에서의 개인차는 작업기억의 용량, 일반 지식의 양, 경험에 의해서 발생한다고 밝힌 것은 가설이 기억에 의존한다는 사실을 더욱 공고히 해준다 (Dougherty & Hunter, 2003a; Dougherty & Hunter, 2003b; Sprenger & Dougherty, 2006; Weber, Bockenholt, Hilton & Wallace, 1993; Thomas, Dougherty, Sprenger & Harbison, 2008; Dougherty & Sprenger, 2006; Ho & Keller, 1994; Libby & Fredrick, 1990; Schwarz, 2004; Gettys & Fisher, 1979; Lange, Thomas, Buttacio, Illingworth & Davelaar, 2013).

주어진 단서를 기반으로 기억으로부터 가설을 생성한다는 HyGene 모델의 설명에 부합되게, 기존 연구들에서 단서의 빈도수를 조절하거나, 생생함 (vividness), 주어지는 순서를 바꾸어 한 가설의 활성화를 조작하였다 (Koehler, 1991; Tversky & Kahneman, 1973; Reyes, Thompson & Bower, 1980; Lange, Thomas & Davelaar, 2013; Sprenger, & Dougherty, 2012; Mehlhorn, Taatgen & Lebiere, 2011; Hogarth & Einhorn, 1992; Pennington & Hastie, 1988; Castellan & Edgell, 1973; Schwarz, 2004; Lange, Thomas, Buttacio, Illingworth & Davelaar, 2013). 더 빈번하게, 최근에

1) 여기서의 단서는 추론에 직접적으로 관련이 있어 사용되는 모든 정보들을 의미한다(Thomas, Dougherty, Sprenger & Harbison, 2008).

제시되거나 이해하기 쉬운 단서와 관련된 가설의 접근성은 증가하여, 이 가설이 최종적으로 선택될 확률은 높아진다.

하지만 활성화는 단서에 기반 하지 않아도 맥락에 의해서 발생할 수 있다 (Gilad & Kilger, 2008; Bioy & Hilton, 2002; Kusev, Schaik & Aldrovandi, 2012; Schunn & Dunbar, 1996; Kahneman, 2011; Jacoby, Kelley, Brown & Jasechko, 1989; Ho & Keller, 1994; Green, Fugelsang & Dunbar, 2006; Wheeler & Berger, 2007; Berger & Fitzsimons, 2008; Carroll, 1978). 선행 연구들을 고려해 보았을 때, HyGene모델이 가정하는 것처럼 단서에 기반을 두어 한 가설의 활성화를 변화시키는 것 이외에도, 맥락을 조작하여 일시적으로 가설의 가용성을 바꿀 수 있는 가능성이 존재한다. 예를 들어, 몇몇 연구들에서 사전에 풀었던 다른 맥락에서의 문제가 활성화를 통해 판단에 영향을 미치는 것을 보여주었다. 예를 들어, Gilad 와 Kilger의 연구에서 (2008), 사전에 보여준 글은 이후 위험 태도에 (risk attitude) 영향을 미쳐 투자 결정을 바꾸게 한다. 연구자들은 실험 참여자들에게 가상의 인물의 특징을 담은 짧은 글을 보여주고, 인물을 묘사하는 형용사를 최대한 많이 회상하여 쓰게 하였다. 이후 참여자들은 주식을 선택하였는데, 사전 위험추구 성향이 있는 인물에 대한 설명을 받은 집단은, 위험 회피 성향이 있는 인물에 대한 설명을 받은 집단보다 주식에 더 높은 금액으로 투자했으며, 투자에 대한 위험을 더 작게 인식하였다. 두 번째 실험에서는 단순히 사전에 광고 문구를 보여주는 것 또한 이후 투자 결정에 영향을 미친다는 것을 확인하였다 (Gilad & Kilger, 2008). 이와 유사하게, Erb, Bioy 그리고 Hilton의 연구에서도(2002) 위험 회피와 관련된 형용사를 받은 집단이 더 확실한 해결책을 선호했다 (Erb, Bioy & Hilton, 2002).

Kusev, Schaik 그리고 Aldrovandi의 연구에서 (2012)는 보험을 선택하기 이전에 다른 맥락의 문제를 먼저 풀게 하면, 의미 활성화로 인해 이후 보험 가입 유무가 달라질 수 있음을 보여주었다. 이들의 연구에서 실험 집단은 ‘차 도둑’이 더 많은 도시를 고르는 문제를 먼저 풀고, 이후 ‘1%의 가능성으로 차 안에서 30파운드 상당을 도둑맞을 수 있는데, 30파운드의 비용으로 보험에 가입할 것인지’에 대해 답하였다. 결과는 사전에 ‘차 도둑’ 의미 활성화를 받은 실험 집단이 그렇지 않은 집단보다 더 위험 회피 (risk aversion) 성향을 보이게 돼, 보험에 더 많이 가입하였다 (Kusev, Schaik, & Aldrovandi, 2012). 이 밖에도 사전에 복잡한 문제를 풀 때 사용된 개념이, 활성화를 통해 이후 다른 복잡한 문제를 풀 때 영향을 미친다는 결과를 고려해 보면 (Schunn & Dunbar, 1996), 사전 활성화는 심지어 고차적인 사고 (high order thinking)할 때에도 영향을 미친다는 것을 알 수 있다.

이러한 현상이 나타나는 이유로 크게 두 가지를 생각해 볼 수 있다. 먼저,

사전 문제를 통해 의미 접화가 발생하면 (semantic priming) 이전 경험 접근성 (pre-experimental accessibility)과 독립적으로 ‘추가적인’ 활성화가 유발될 수 있다 (Kusev, Schaik & Aldrovandi, 2012). Kusev와 동료들은 접근성을 장기기억에 저장되어 있는 ‘경험의 가용성’과 결정 맥락에서 ‘즉시 떠올리기 쉬운 것’으로 나누었는데, 의미 활성화를 통해 경험적 가용성과 독립적으로 특정 단서를 인출하기 쉽게 조작할 수 있음을 검증하였다.

둘째로, Kahneman과 Tversky(1973)가 주장한 가용성 휴리스틱 (availability heuristic)에서 볼 수 있듯이, 사람들은 기저율을 쉽게 무시하고 사건이 더 쉽게 회상될수록 그것이 더 높은 빈도를 지닌다고 추론 한다 (Kahneman & Tversky, 1973; Jacoby, Kelley, Brown & Jasechko, 1989; Reyes, Thomson & Bower, 1980; Carroll, 1978; Ross, Lepper, Strack & Steinmetz, 1977). 따라서 선행 연구들의 결과는 사전에 활성화된 의미가 한 가설의 가용성을 일시적으로 높여서, 사건의 실제 발생빈도를 왜곡시켰기 때문으로 설명된다.

이 밖에도 사람들은 한 사건이 기억에서 쉽게 다루어질수록 더욱 진실에 가깝다고 여기며 (Schwarz, Sanna, Skurnik & Yoon, 2007; Connel & Keane, 2004; Alter & Oppenheimer, 2009; Reber & Schwarz, 1999), 결정에 더욱 확신을 가진다는 연구 결과도 존재 한다 (Haddock et al., 1999; Tormala, Petty & Brinol, 2002; Schwarz, Sanna, Skurnik & Yoon, 2007; Pennington & Hastie, 1988; Alter & Oppenheimer, 2009; Koehler, 1991; Mehle, Gettys, Manning, Beca & Fisher, 1981).

이상을 종합해보면, 진단적 추론과 다른 맥락에서 발생한 의미의 활성화는 가설 생성과 추론 확신에 영향을 미칠 것임을 예상해 볼 수 있다. 선행 연구들에서 (Gilad & Kilger, 2008; Erb, Biov & Hilton, 2002; Kusev, Schaik, & Aldrovandi, 2012; Schunn & Dunbar, 1996) 다른 맥락에서의 활성화가 판단과 문제해결에 영향을 미친다는 것을 보여주었다. 판단과 문제해결 그리고 진단적 추론은 모두 고차적 사고를 요한다는 특징이 있고, 과제 수행이 기억에 의존한다는 공통점이 있다. 따라서 진단적 추론에서도 사전 맥락에 따라 한 가설의 가용성이 달라질 것이라 예상된다.

2. 가설의 유형과 가설 생성

가설의 중요한 기능 중에 하나는 설명이다 (Lombroze, 2012; Einhorn & Hogarth, 1986; Thagard, 2000). 가설 생성이란 어떠한 현상이 관찰되었을 때 그 원

인에 대한 설명들을 만들고, 이 중 가장 좋은 설명을 최종적으로 선택하는 일련의 과정을 의미한다 (Josephson & Josephson, 1994).

모든 설명은 같은 구조와 기능을 가지고 있지 않으며, 설명의 특성에 따라서 추론의 결과가 달라질 수 있다 (Lombrozo, 2006; Sloman, 2004; Lombroze, 2007). 예를 들어, Sloman의 연구에서 (2004) 만약 설명들이 양립 가능하면, 설명이 더 많은 추론 결과를 진실로 여기지만, 설명들이 서로 배타적인 경우엔 반대의 현상이 나타난다 (explanation discounting) (Sloman, 2004). 배타적인 설명의 경우, 사람들은 적은 수의 설명이 있는 추론 결과를 더 진실로 여긴다. 또한 맥락에 따라서 단순한 설명과 복잡한 설명의 선호에서도 차이가 나타나는데, Lombroze (2007)는 주어진 정보에 따라서 사람들의 설명 선호도가 달라진다는 것을 보여주었다. 일반적으로 사람들은 단순한 설명을 더 선호하고, 더 진실이라고 받아들이지만, 기저율 정보가 주어진 경우에는 복잡한 설명을 더 선호한다 (Lombroze, 2007).

이처럼 설명의 특성에 따라 추론이 달라질 수 있음에도 불구하고, 대부분의 가설 생성/평가 연구들에서는 오직 배타적인 가설 세트만을 다루어 왔다 (McKenzie, 1998; Ross & Murphy, 1996; Malt, Ross & Murphy, 1995; Wallendael & Hastie, 1990; Teigen, 1983; Tversky & Koehler, 1994; Castellan & Edgell, 1973; Gettys & Fisher, 1979). 배타적인 가설 세트에서 (mutually exclusive and exhaustive) 한 가설의 가능성은 다른 가설과의 비교를 통해서 이루어지며, 이 경우 오직 하나의 가설만이 참이 된다. 비록 배타적인 세트에서 한 가설의 가능성의 증가가 다른 가설의 가능성을 반드시 감소시키지는 않는다는 연구 결과도 존재하지만 (Wallendael & Hastie, 1990), 다른 연구에서 가설의 수가 적을 경우에는 (예, 2개) 상호 보완적 (complementary)으로 평가한다는 것이 밝혀졌다 (Payne, 1976).

McKenzie (1998)는 기존 연구 결과들이 주어진 가설 중에 반드시 하나만 답이 되고 (mutually exclusive), 주어진 가설 중에 답이 있다는 (exhaustive) 것에 기반하는 것을 지적하였다. 일상의 진단적 추론상황에서 가설들은 동시에 답이 되기도 하고 (no mutually exclusive), 충분히 근거가 쌓이지 않으면 현재 고려되고 있는 가설이 선택되지 않기도 한다 (no exhaustive). 또한 두 개의 가설이 모두 답이 가능한 경우에는 두 개의 가설이 수렴하여 제 3의 다른 가설 하나가 생성될 가능성도 존재한다 (McKenzie, 1998).

이상을 종합하면, 주어진 맥락과 가설의 유형에 따라 가설의 최종 선택 패턴과, 생성된 가설의 수, 추론의 확신이 달라질 수 있음을 예상해 볼 수 있다. 사람들은 다른 맥락으로부터 활성화된 가설을 더 쉽게 떠올리며 실제 발생 빈도를 왜곡할 것이다. 하지만 이 현상은 가설의 유형에 따라서 달라질 가능성이 있는데, 만약 두

가설이 배타적이라면 활성화로 인해 한 가설의 가능성이 높아지면 반드시 다른 가설의 가능성은 낮아질 것이다 (complementary evaluation). 반면 두 가설이 서로 양립할 수 있으며, 경우에 따라서 수렴하여 제 3의 가설이 생성될 수도 있다면, 한 가설의 가능성의 증가가 다른 가설의 가능성을 낮추지 않을 것이다. 이 과정에서 전체 가설세트의 크기와 추론의 확신 또한 달라질 것이다.

총 3번의 실험을 통해, 진단적 추론과 다른 맥락의 문제에 의한 의미 활성화는 가설들이 서로 배타적인지, 양립가능한지에 따라서, 1) 최종 가설 선택 패턴, 2) 가설의 생성 수 3) 추론 확신에 다른 영향을 미칠 것임을 검증하였다.

실 험 1

실 험 방 법

실험 1은 서울 소재 S대학교 학부 생 120명을 대상으로 진행되었다 (남 66명, 여 54명, 평균 나이 23세). 실험에 사용된 과제는 기존 연구에서 주로 쓰이던 범인 추론 문제와 더불어 (Malt & Murphy 1995; Ross & Murphy, 1996), 원인 추론 문제 (Fenton & Neil, 2012)가 사용되었다²⁾. 구체적으로 <과제 1>에서는 짧은 글을 읽고 (표 1) ‘초인종을 누른 사람은 누구인지’ 를 추론하게 하였고, ‘도둑’ 과 ‘택배 기사’ 가설이 제시되었다. 이 과제에서 문 앞에 남자는 단 1명이기 때문에 가설들은 상호 배타적이다 (mutually exclusive).

반면 <과제 2>는 (표 2) ‘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 복수응답 또한 가능하다는 특성을 지닌다.

[표 1] <과제 1>에서 사용된 지문

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 정

2) <과제 1>과 <과제 2>에서 사용된 지문들은 기존 연구들 (과제 1: Malt & Murphy, 1995; Ross & Murphy, 1996; 과제 2: Fenton & Neil, 2012)에서 사용된 것들이다. 이 지문들을 한국 문화에 맞게, 그리고 진단적 추론의 형식으로 수정하였다.

리하고 있었는데 초인종이 눌렸다. 확인해 보니 한 사람이 있었으나 얼굴이 보이지 않았다. 팔 부분의 옷자락만 보일 뿐이었다. 지금 시간에 올 사람이 누굴까 생각하다가, 불현듯 며칠 전 부모님들께서 포도를 보내주신다고 말씀하셨던 것이 생각났다. 아마도 택배 직원이 포도를 가지고 왔을 것 같았다. 박 모양은 응답을 하기 위해 서둘러 인터폰을 들었다.

[표 2] <과제 2>에서 사용된 지문

Tomas Sowell 박사는 1987년 워싱턴 DC에 사는 산모들을 조사하다가 흥미로운 사실을 발견하였다. ‘출산 전 건강관리’와 새로 태어난 ‘아이 몸무게’ 간의 상관관계가 있음을 발견한 것이다. 출산 전 낮은 수준의 건강관리를 받은 여자들은 대부분 흑인과 가난한 사람이었다. 정부의 잘못된 정책으로 인해 가난한 사람들이 건강관리 수혜를 입지 못해, 저 체중 아이를 낳게 되고 결과적으로 유아의 건강에 심각한 영향을 미치는 것이다. 하지만 이 연구가 발표되자 다른 관점으로 문제를 바라보는 사람들도 나타났다. 이러한 결과는 인종과 소득과는 무관하게, 산모의 흡연과 알코올 섭취와 관련이 있다는 것이다. 출산 전 건강관리 후원을 받지 못했던 산모들의 흡연과 알코올 섭취의 비율이 각각 2배, 6배가 더 높다는 통계가 있기 때문이다. 하지만 이러한 주장이 제기되자, 많은 사람들이 통계수치의 신빙성과, 과연 저 체중 아이의 출산을 전적으로 가난한 산모의 개인 책임으로 물을 수 있느냐고 반박하였다.

실험 참여자들은 ‘글 이해와 추론’이라는 연구로 알고 참여하였으며, 지문을 읽은 후 이어지는 문제에 답을 하였다 (표 3). 문제를 풀 때 다시 지문의 내용을 보지 않도록 하였으며, 지문에는 가설이 두 개 제시되어 있으나, 글 속에 제시되지 않았던 가설을 생성할 가능성도 있기 때문에 추론 결과는 주관식으로 자유롭게 답하게 하였다. 또한 추론의 답(최종 가설)이외에 사전에 고려했던 가설이 있었는지 알아보기 위해서, 다음 문제에서 이 밖에 추가로 고려한 대안이 있었는지, 만약 있었다면 무엇이 있었는지 쓰도록 하였다.

[표 3] <과제 1>과 <과제 2>에서 사용된 문제들

	문 밖에 있는 사람은 누구라고 생각하는가?
과제 1	위에 응답한 답 이외에 생각한 대안(들)은 있었는가? 있다면 무엇이고, 그 답이 맞을 확률은 어느 정도라고 생각하는가?
	지금까지의 추론에 얼마나 확신이 있는가?
과제 2	흑인과 저 소득층 산모가 저체중아를 낳는 원인은 무엇이었다고 생각하는가?
	위에 응답한 답 이외에 생각한 대안(들)은 있었는가? 있다면 무엇이고, 그 답이 맞을 확률은 어느 정도라고 생각하는가?
	앞서 제시한 답들에 얼마나 확신이 있는가?

집단은 네 집단으로 나누었으며 각 집단마다 약 30명의 실험 참여자가 무작위로 배정되었다. 세 집단은 실험집단으로 추론을 하기 전 진단적 추론과 다른 문제가 들어간 설문지를 받았으며 (표 4), 통제집단은 진단적 추론과 다른 문제가 제시되지 않았다. 실험집단 1, 실험집단 2에서는 진단적 추론과 다른 문제 두 개로 인해 한 가설의 의미가 사전에 활성화 되었고, 실험집단 3에서는 실험집단 1과 실험집단 2에서 사용되었던 진단적 추론과 다른 문제 네 개로 인해 두 가설의 의미가 사전에 동시에 활성화 되었다 (표 5). 연구 참여자들이 실험의 목적을 눈치 채지 못하도록 하기 위해서 이 외에 다른 문제들도 일부 포함되었다 (부록 1).

[표 4] 실험 1에서 사용된 진단적 추론과 다른 맥락의 문제들

과제 1	실험 집단1	집에 도둑이 들었던 경험이 있는가? 2013년 현재를 기준으로 절도범죄 검거율은 얼마쯤이라고 생각하는가? 그 이유는? (%)
	실험 집단2	택배를 받아 본 경험이 있는가? 100명중 몇 명이 택배서비스를 이용하고 있을 것 같은가? 그 이유는? (%)
과제 2	실험 집단1	영국 매체 미러(Mirror)에 따르면 알코올중독에 빠진 24세 여성이 출산이 임박했음에도 음주 욕구를 못 이겨 한 주류 판매점에서 술을 마시다가 만삭인 상태로 쓰러졌다. 이 후 이 여성에게 영국 정부는 어떠한 제재를 가했는지 추측해 보시오. 개인 스스로 고치기 힘든 나쁜 습관을 고치기 위해 집단 상담 프로그램에 많이 참여한다. 보통 프로그램의 효과가 어느 정도 나타날지 예상해 본다

	면? (%)
실험 집단2	무상보육정책의 장단점은 무엇이라고 생각하는가? 당신이 생각하기에 가장 이상적인 보육정책은 무엇인가?

[표 5] 실험 1 조건 간 차이 정리

실험집단				통제집단
집단1	집단2	집단3		
첫 번째 가설과 같은 어휘가 있는 사전 문제 2개 포함	두 번째 가설과 같은 어휘가 있는 사전 문제 2개 포함	두 가설 모두 활성화시키기 위해 문제 4개 포함		활성화 없음
활성화된 가설	과제1: 도둑 과제2: 음주	과제1: 택배기사 과제2: 정부정책	과제1: 도둑, 택배기사 과제2: 음주, 정부정책	

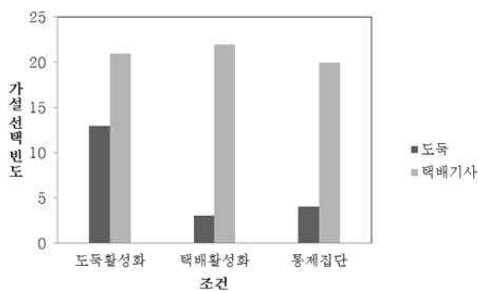
사람들은 일반적으로 시간제한이 있을 때 더 적은 수의 가설을 생성한다는 것을 고려하여 (Thomas, Dougherty, Sprenger & Harbison, 2008), 시간제한을 두지 않고 자유롭게 글을 읽고 문제를 풀도록 하였다. 평균적으로 한 사람당 약 15~20분 정도의 시간이 소요되었다.

실험 결과

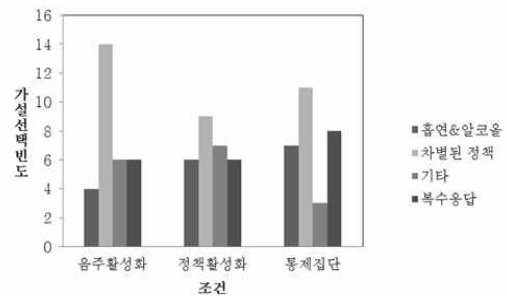
결과 1) 최종가설 선택패턴

사전 진단적 추론과 다른 맥락의 문제에 의해 활성화된 의미가 ‘추론 결과 (최종 가설 선택 패턴)’에 미치는 영향을 검증하기 위하여, 교차분석을 통해 실험 집단 1과 실험집단 2, 그리고 통제집단 간에 가설 선택 빈도를 비교해 보았다. 그 결과 배타적인 가설의 추론에서는 (과제1), 의미 활성화에 따라 특정 가설의 선호가 더 높아졌다 ($\chi^2 (df=2) = 6.875, p < .05$). 그림 2에서 보는 것과 같이, ‘도둑’ 활성화를 받은 실험집단 1은 다른 조건보다 더 많이 ‘도둑’ 가설을 생성하였으며, ‘택배’ 활성화를 받은 실험집단 2는 다른 조건들보다 더 많이 ‘택배기사’ 가설을 생성하였다. 모든 집단에서 택배기사의 가설이 도둑 가설보다 더 많이 생성되었음을 고려해볼 때, 사람들은 기저율을 어느 정도 고려하고 있음을 알 수 있다³⁾. 또한 그림 2를 보면, 택배기사 가설의 선택 빈도는 조건들 간에 별다른 차이가 없지만, 도둑 가설의 경우엔 확실하게 ‘도둑’ 의미를 사전에 활성화 받았던 집단에서 선택 빈도가 더 높다는 것을 알 수 있다. 활성화의 영향이 ‘택배기사’ 가설보다 ‘도둑’ 가설에서 더 크다는 결과에 대해서는 마지막 <종합 논의>에서 자세히 설명할 것이다 (p. 43).

배타적인 가설과는 달리, 양립 가능한 가설의 추론에서는 (과제 2), 사전 활성화가 최종 가설 선택에 영향을 미치지 않았다 ($\chi^2 (df=6) = 3.894, p > .05$). 그림 3에서 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.



[그림 2] 과제 1(배타적인 가설)에서의 최종 가설 선택 패턴



[그림 3] 과제 2(양립 가능한 가설)에서의 최종 가설 선택 패턴

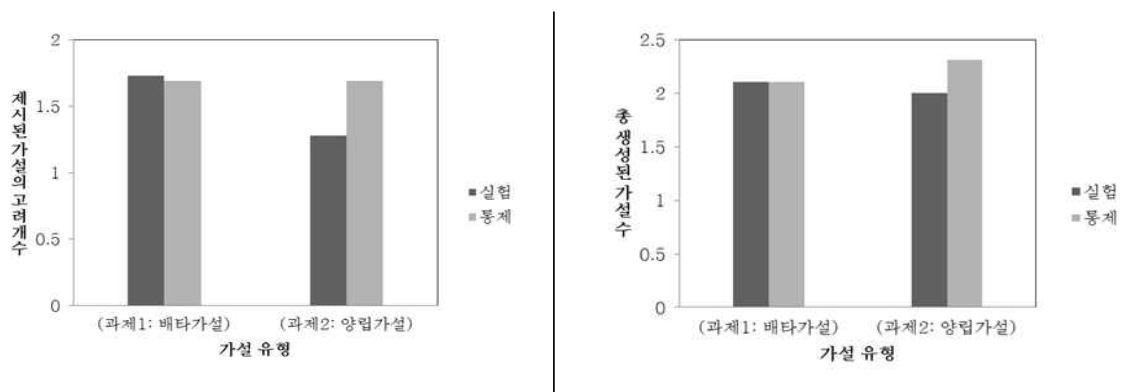
3) 조건에 상관없이 각 집단에서 ‘택배’ 가설보다 ‘도둑’ 가설의 선택 빈도가 높다는 사실은 기저율의 차이뿐 아니라 최신 효과(recent effect)에 의해서 발생했을 가능성도 있다. 두 가설이 지문에 제시된 순서를 보면 도둑 가설은 지문의 앞 단락에서 제시되었고, 택배기사 가설은 지문의 마지막 단락에 제시되었기 때문이다. 하지만 <과제 2>의 경우를 살펴보면, 조건에 상관없이 ‘흡연&알코올’ 보다 ‘차별된 정책’ 가설이 더 많이 선택되었는데, <과제 2>의 지문에서는 차별된 정책가설이 지문의 앞 단락에 제시되었다. 이를 통해 가설의 선택에 최신 효과가 영향을 미쳤을 가능성은 제외된다.

결과 2) 생성된 가설의 수

진단적 추론과 다른 맥락에서 발생한 활성화가 ‘가설의 생성 개수’에 미치는 영향을 알아보기 위해서, 통제집단과 실험집단 간에 독립표본 T 검정을 실시하였다. 활성화 여부에 따라, 지문에 제시된 2개의 가설들 중 몇 개의 가설을 고려했는지를 비교해 보았고, 이를 포함하여 총 생성된 가설의 개수 역시 비교해 보았다.

결과 2에서는 결과 1과는 반대의 효과가 관찰되었는데, <과제 1>의 경우 실험집단과 통제집단 간에 지문에 등장한 가설의 고려 개수는 차이가 없었으며 ($t(87) = 0.385, p > .05$), 전체 고려된 가설의 개수에서도 차이가 발견되지 않았다 ($t(86) = 0.011, p > .05$). 특정 가설의 활성화 여부와 상관없이, 통제집단과 실험집단은 지문에 등장한 두 개의 가설 중 평균 1.5개의 가설을 고려하였으며, 총 생성한 가설의 개수는 평균 2개였다 (그림 4).

반면 <과제 2>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 ($t(84) = 2.072, p < .05$). 다른 맥락에서의 활성화가 없었던 통제집단은 활성화 집단보다 제시된 가설을 더 많이 고려하였을 뿐만 아니라, 최종적으로 더 많은 수의 가설을 생성했다 (그림 4).



[그림 4] 가설 유형에 따른 가설 생성 수의 차이

결과 3) 추론의 확신

다른 맥락에서 유발된 활성화가 ‘추론의 확신’에 영향을 미치는지 확인하기 위해 실험집단과 통제집단 간에 추론의 확신 정도를 비교하였다. 독립표본 T검정 결과, 사전 예상과는 다르게 가설의 유형에 상관없이, 실험 집단과 통제집단 간에 추론의 확신은 차이가 없었다 (과제1: $t(88) = 0.303$, $p > .05$; 과제2: $t(88) = 0.629$, $p > .05$).

논 의

실험 1을 통하여, 진단적 추론과 다른 맥락에서 유발된 활성화는 두 가설의 관계에 따라서 가설생성에 다른 영향을 미친다는 것을 확인하였다.

<다른 맥락에서의 활성화가 최종 가설 선택에 미치는 영향>

실험 결과 1을 보면, 다른 맥락에서의 활성화는 배타적인 가설의 최종가설 선택에만 유의미한 영향을 미쳤다. 배타적인 가설의 경우 (과제 1) 다른 맥락의 활성화에 따라서 편향된 가설이 생성됐지만, 양립 가능한 가설의 경우 (과제 2) 이러한 현상이 나타나지 않았다. 이 결과는 먼저 부채효과 (fan effect) (Anderson & Redar, 1999; Anderson, 2011)로 설명해 볼 수 있다. 부채효과에 의하면 마디의 수가 많을수록 활성화의 양은 분산된다. <과제 1>의 경우 원인 마디는 2개지만 (도둑 vs 택배기사), <과제 2>의 경우 원인 마디는 3개 혹은 그 이상이다 (흡연&알코올, 차별된 정부 정책, 기타, 복수응답). 각 과제의 지문에는 같은 수의 가설이 (2개) 제시되었지만, 양립 가능한 가설의 경우 만들어질 수 있는 가설의 개수가 더 많다. 배타적인 가설의 경우는 두 가설 중 단 하나만이 답이 된다는 특징이 있지만, 양립 가능한 가설의 경우는 두 가설이 동시에 답이 될 수 있다는 특징이 있다 (복수 응답). 또한 McKenzie (1998)에 따르면, 두 가설이 동시에 답이 될 수 있는 경우, 수렴하여 새로운 제 3의 가설이 만들어질 수 있다. <과제 2>에서 ‘흡연&알코올’ 과 ‘차별된 정부정책’ 의 두 가설은 수렴하여 ‘산모의 건강 악화’ 라는 제 3의 가설이 생성되기도 하였다⁴⁾. 따라서 각 원인마디에 미치는 다른 맥락의 활성화의 영향은 <과제 2>보다 <과제 1>에서 더 크다.

또 다른 해석으로, 배타적인 가설의 경우 (과제 1) 각각의 가설이 서로 독립적이므로, 다른 맥락에서 유발된 한 가설의 의미 활성화는 다른 가설에 영향을 끼칠 필요가 없다. 하지만 양립 가능한 가설의 경우 (과제 2) 서로 상호작용이 가능하기 때문에, 다른 맥락의 활성화는 한 가설에만 영향을 주는 것이 아니라 가설 세트 전체 집합에 영향을 미치게 된다 (McKenzie, 1998). 따라서 배타적인 가설의 경우 다른 맥락에서의 활성화가 한 가설에 직접적으로 영향을 미쳐서 활성화에 따라 편향되게 가설이 생성되지만, 양립 가능한 가설의 경우엔 이 현상은 나타나지 않는다.

4) 양립 가능한 가설의 제 3 가설의 생성 패턴은 <종합 논의>에서 보다 자세히 다루어 졌다 (p. 45).

<다른 맥락에서의 활성화가 가설의 생성 수에 미치는 영향>

실험 결과 1과는 반대로, 실험 결과 2에서 다른 맥락에서의 활성화는 양립 가능한 가설의 생성 수에만 유의미한 영향을 미친다는 것이 확인됐다. <과제 2>에서 다른 맥락의 활성화를 받은 집단은 통제집단보다 더 적은 수의 가설을 생성하였다. Sloman(1994)은 설명이 서로 배타적일 때, 한 가설의 가능성의 증가가 다른 가설의 생성을 억제하여 더 적은 수의 설명이 만들어질 수 있다고 생각하였다. 그러나 본 연구에서 배타적인 가설의 경우 한 가설의 가능성의 증가는 다른 가설의 생성에 영향을 미치지 못하지만, 오히려 배타적이지 않은 가설의 경우 한 가설의 가능성의 증가가 다른 가설의 생성을 방해한다는 것을 밝혔다. 다른 맥락에서의 활성화를 받은 집단이 그렇지 않은 집단보다 더 적은 수의 가설을 생성하는 현상은 ‘차단block’ 또는 ‘고착fixation’ 현상으로 설명될 수 있다. 차단 및 고착 현상이란, 이전에 생성된 가설로 인해 다른 가설을 생성하지 못하는 것을 의미 한다 (Tseng, Cagan & Kotovsky, 2008; Bassok & Novick, 2012; Smith & Blankship, 1991). 다른 맥락에서의 활성화는 한 가설을 먼저 작업 기억으로 불러들이고, 이 가설에 고착되게 하여 추가로 다른 가설을 생성하지 못하게 된다. 이는 가설생성이 복잡할 때, 한 가설의 활성화가 다양한 가설의 생성을 억제시킬 수 있다는 것을 함축한다.

<다른 맥락에서의 활성화가 추론의 확신에 미치는 영향>

실험 결과 3에서 다른 맥락에서의 활성화는 가설의 유형에 상관없이 추론의 확신에 영향을 미치지 못한다는 것이 확인됐다. 이는 가설이 생생할수록 확신이 높다는 선행 연구들과 대치되는 듯하지만, 가설의 활성화를 기존 연구는 추론에 필요한 단서를 이용해 조작한 반면, 본 연구는 추론과 다른 맥락을 통해 유발시켰다는 차이가 있다. 일반적으로 사람들은 자신의 태도에 대한 근거가 더 많을수록 더 높은 확신을 가진다 (Haddock, Rothman, Reber, & Schwarz, 1999; Tormala, Petty & Brinol, 2002; Schwarz, Sanna, Skurnik & Yoon, 2007). 하지만 본 연구의 경우 추론의 단서가 없이 단순히 다른 맥락을 통해 활성화를 조작했기 때문에, 기존 연구와는 달리 실험집단과 통제집단 간에 추론 확신에 차이가 없었을 것이다. 이는 단순히 활성화가 높은 가설을 선택하는 것은 추론의 확신과 어떠한 관련이 없고, 한 가설을 지지하는 단서를 통한 활성화만이 추론의 확신에 영향을 미친다는 것을 시사한다.

실 험 2

실험 1을 통하여, 사람들은 진단적인 추론을 할 때 사전에 제시된 다른 맥락의 문제에 영향을 받는다는 것을 알게 되었다. 진단적 추론은 우리가 매일 하는 추론 방식이기 때문에, 실험 1의 결과는 일상생활에서 우리가 잘못된 판단을 하고 있을 가능성을 보여준다. 예를 들어, 단순히 사전에 친구가 핸드폰 배터리를 바꾸었다는 이유로 인해 ‘자동차의 시동이 걸리지 않는 것은 배터리가 없기 때문’으로 추론할 수 있다. 합리적인 근거에 기반 하지 않고 ‘현재 배터리가 쉽게 떠오르는 것을 보면, 자동차의 배터리 부족은 매우 빈번하게 발생하는 문제이다’라는 엉터리 추론하고 있는 셈이다. 실험 2는 이러한 판단의 오류를 줄이기 위한 방안을 탐색하기 위하여 실시되었다.

기존 연구에서 활성화의 효과가 직관적이고 덜 분석적인 (less analytic) 사고를 할 때 더 크며 (Gilad & Kliger, 2008), 정신적인 노력을 더 들일수록 더 깊은 사고를 할 수 있다는 것을 보여주었다 (Kahneman, 2011; Stanovich, 2011). 만약 실험 1에서 다른 맥락에서 의미 활성화가 가설 생성에 영향을 미친 이유가 피상적인 사고에서 비롯된 것이라면, 인지적인 자극을 제공하여 보다 깊이 있는 사고를 할 때 활성화의 영향은 줄어들 수 있을 것이다. 사람들은 다른 맥락에서의 활성화를 추론과 관련 없는 원천 (resource)으로 분류하여, 이를 배제하고 추론할 가능성이 더 커질 것이기 때문이다 (Schwarz, Sanna, Skurnik & Yoon, 2007; Schwarz, Bless, Strack, Klumpp & Rittenauer-Schatka, 1991; Strack, Schwarz, Bless, Kubler & Wanke, 1993).

실험 2는 더 깊은 사고를 하는 것이 다른 맥락에서의 활성화가 가설생성에 미치는 영향의 경계조건 (boundary condition)인지 확인하기 위해 수행되었다.

실험 방법

실험 2는 서울 소재 S대학교 학부 생 86명을 대상으로 진행되었다 (남 46명, 여 39명, 평균나이 21.5세). 실험에 사용된 과제의 지문들은 실험 1에서 사용된 것과 같지만, 실험집단에 제시되는 진단적 추론과 다른 맥락의 문제의 난이도를 높여 인지적 노력을 증가시키고자 하였다 (인지시스템 2의 활성화 (Kahneman, 2011; Stanovich, 2011)). 예컨대, 실험 1에서 추론에 앞서 택배를 받아본 경험 유무나 현재 도둑 검거율을 대략적으로 묻는 문제가 제시됐다면, 실험 2에서는 통계 계산문제, 주어진 주장을 비판하는 문제를 제시하였다 (표 6). 실험에 성실히 참여했는지를 점검하기 위하여 지문 내용 확인 문제를 추가로 3개 제시하였다 (특정 가설을 떠올리지 않도록 중립을 유지하는 문제를 선정하였다) (부록 2).

실험 참여자들은 ‘비판적 사고력과 추론’이라는 연구로 알고 참여하였으며, 실험 1과 마찬가지로 지문에는 가설이 2개 제시되어 있으나, 추론은 자유롭게 할 수 있도록 주관식으로 답을 쓰게 하였으며, 바로 이어서 대안 가설의 고려 여부를 묻는 문제가 제시됐다.

실험 1에서 두 가설의 의미를 모두 활성화 시켰던 실험집단 3은 결과에 미치는 영향이 없다고 판단되어, 실험 2에서 실험집단 3을 제외하고 집단을 세 집단으로 나누었다. 실험집단 1과 실험집단 2에서 각각 활성화된 가설은 실험 1과 동일하였으며, 각 집단에 약 30명의 실험 참여자가 무작위로 배정되었다. 실험 집단은 추론을 하기 전 특정 가설을 활성화시키기 위한 진단적 추론과 다른 어려운 문제가 2개 제시되었으며, 통제집단은 추론과 다른 맥락의 문제가 제시되지 않았다. 시간제한은 없었으며, 개인차에 따라 평균적으로 약 15~30분 정도의 시간이 소요되었다.

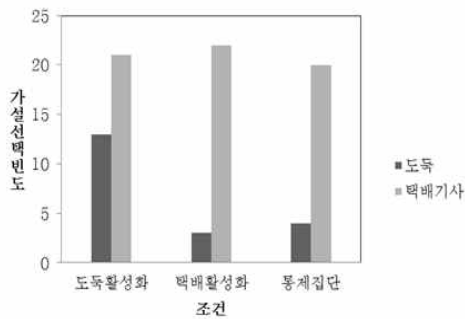
[표 6] 실험 2에서 사용된 진단적 추론과 다른 맥락의 문제들

과제 1	실험 집단1	<p>도둑들 중에서 남자의 확률은 92%이라고 하였을 때, 남자 중에서 도둑인 확률은 어느 정도라고 생각하는가? (한 사람이 남/여일 확률은 각각 50%라고 가정) (%)</p> <p>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.</p>
	실험 집단2	<p>택배기사들 중에서 남자의 확률은 92%이라고 하였을 때, 남자 중에서 택배기사인 확률은 어느 정도라고 생각하는가? (한 사람이 남/여일 확률은 각각 50%라고 가정) (%)</p> <p>2013년 조사 결과 운송업체 직원들의 평균 나이는 22세였다. 이 글을 본 어떤 사람이 대학생들을 대상으로 운전교육을 실시해야 한다고 주장했다. 이 주장에 대해서 간단하게 평가해 보시오.</p>
과제 2	실험 집단1	<p>영국 매체 미러(Mirror)에 따르면 알코올중독에 빠진 24세 여성이 출산이 임박했음에도 음주 욕구를 못 이겨 한 주류 판매점에서 술을 마시다가 만삭인 상태로 쓰러졌다. 임신 중 음주한 여성이 그렇지 못한 여성보다 ‘태아 알코올 증후군’에 시달릴 확률은 어느 정도 차이가 날 것이라고 생각하는가? (%)</p> <p>5. 다음 조건에서 캘리포니아에 거주하고 있는 산모들이 알코올을 섭취했을 확률을 계산해 본다면? (조건: 캘리포니아 주에서 거주하고 있는 산모들은 미국의 전체 산모들의 30%이다. 미국의 전체 산모들의 알코올 섭취율은 20%이다.) (%) 위 답에 얼마나 확신하는가? (%)</p>
	실험 집단2	<p>한 매체의 조사결과 서울 시민 100명은 우리사회의 차별적 요인으로 다음과 같은 점을 언급하였다. 괄호 안에 있는 요소를 각 파이에 대응시켜 보고, 이렇게 생각한 이유를 간단하게 설명해 보시오. (예시. 1. 직업(60%) > 2. 소득수준(30%) > 3. 교육수준(10%)에 따라서 차별하는 정도가 다르다.)</p> <p>2013년 조사 결과, 암 치료를 받았던 사람들 중에서 환자의 경제적 조건이나 사회적 지위에 따라 의료서비스 이용에 있어서 차별을 받는 것 같다는 응답을 한 사람들은 94.2%, ‘그렇지 않다’ 라는 응답이 5.8%로 나왔다. 어떤 사람이 이 결과를 보고 현재 대한민국 사람들은 의료서비스에 대단히 불만을 품고 있으므로, 의료 개혁이 시급하다고 주장하였다. 이 사람의 주장에 대해서 간단하게 평가해 본다면?</p>

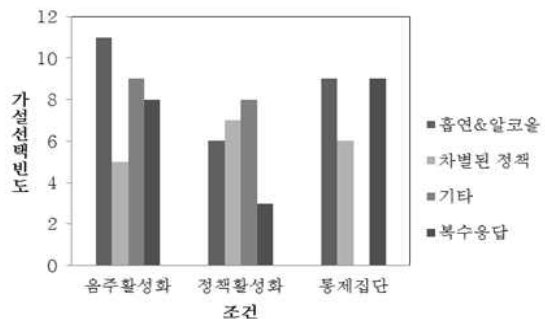
실험 결과

결과 1) 최종가설 선택패턴

사전 추론 맥락과 다른 문제에 의해 활성화된 의미가 최종 가설 선택에 미치는 영향을 검증하기 위하여 교차분석을 실시하였다. 결과는 실험 1과 마찬가지로 <과제 1>에서 다른 맥락에서의 활성화가 이후 특정 가설 선택 빈도에 유의미한 영향을 미쳤다 ($\chi^2 (df=2)=6.441, p<.05$). 그림 5에서 볼 수 있듯이, 배타적인 가설의 경우 ‘도둑’ 활성화를 받은 실험집단1은 ‘도둑’ 가설을 더 많이 생성하였으며, ‘택배’ 활성화를 받은 실험집단 2는 ‘택배기사’ 가설을 더 많이 생성하였다. 반면 <과제 2>에서는 다른 맥락에서의 활성화는 특정 가설의 선택 빈도에 미치지 않았다($\chi^2 (df=6)=12.40, p>.05$) (그림 6).



[그림 5] 과제 1(배타적인 가설)에서의 최종 가설 선택 패턴



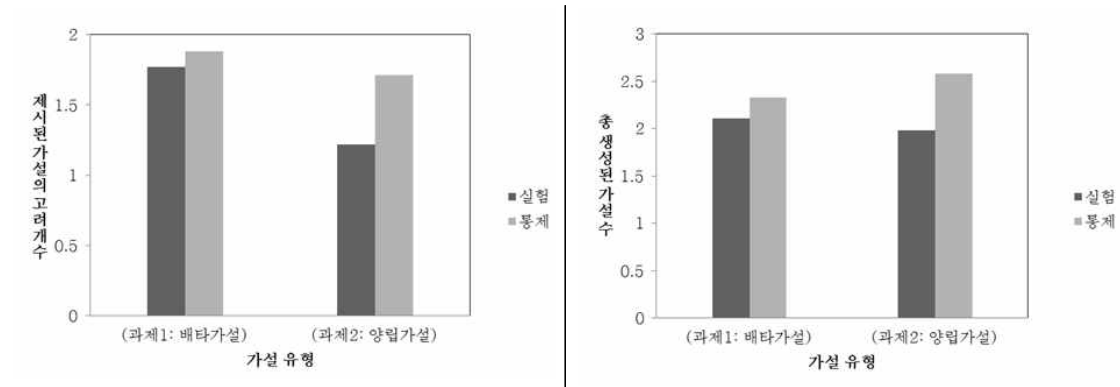
[그림 6] 과제 2(양립 가능한 가설)에서의 최종 가설 선택 패턴

결과 2) 생성된 가설의 수

진단적 추론과 다른 맥락에서의 활성화가 ‘가설의 생성 개수’에 미치는 영향을 보기 위해서 통제집단과 실험집단 간에 독립표본 T 검정을 실시하였다. 활성화 여부에 따라, 지문에 제시된 2개의 가설들 중 몇 개의 가설을 고려했는지를 비교해 보았고, 이를 포함하여 총 생성된 가설의 개수 역시 비교해 보았다.

실험 1의 결과와 마찬가지로, 양립 가능한 가설의 경우에만 활성화를 받은 실험집단이 통제집단보다 더 적은 수의 가설을 생성했다 (그림 7). <과제 1>의 경우 실험 집단과 통제 집단 간에 제시된 가설의 고려개수는 차이가 없었고 ($t(52.599)=1.191, p>.05$), 총 생성된 가설 수에서도 차이가 없었다 ($t(83)=1.472, p>.05$). 반면 <과제 2>의 경우 실험집단 (평균=1.22, 표준편차=0.68)보다 통제집단 (평균=1.71, 표

준편차=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 ($t(32.603)= 3.237, p<.01$).



[그림 7] 가설 유형에 따른 가설 생성 수의 차이

결과 3) 추론의 확산

실험 1과 마찬가지로, 가설의 유형에 관계없이 다른 맥락에서의 의미 활성화는 추론의 확산에 영향을 미치지 못하였다 (과제 1: $t(32.179)= 0.424, p>.05$; 과제 2: $t(74)= 1.428, p>.05$).

논 의

실험 2는 다른 맥락에서의 활성화가 가설생성에 미치는 영향의 경계 조건을 알아보기 위하여 실시되었지만, 실험 1의 결과와 마찬가지로 다른 맥락에서의 활성화는 가설의 유형에 따라서 최종 가설의 패턴, 생성된 가설의 수에 다른 영향을 미쳤다. 더 깊은 사고를 할 수 있도록 어려운 문제를 사전에 제시하였을 때조차 여전히 추론과 직접적으로 관련이 없는 활성화는 가설 생성에 영향을 미친다.

기존 연구에서 사전에 인지적인 노력을 더 하도록 유도한다면 사람들이 더욱 깊이 있는 사고를 하게 되어 인지적 편향이 감소한다는 것을 주장하였다 (Stanovich, 2011; Kahneman, 2011; Gilad & Kilger, 2008). 하지만 실험 2에서 더 분석적인 사고를 하더라도 사전에 추론과 다른 맥락의 문제가 진단적 추론에 영향을 미친다는 것이 반복 검증 되었다. 이 현상은 기존에 Gilad 와 Kilger의 연구에서 (2008) 느리고 분석적인 판단을 하는 집단일수록 활성화의 영향을 더 적게 받는다는 결과와 상반된다. 이러한 결과가 발생한 가장 큰 원인 중 하나는 추론에 필요한 단서의 존재 여부라고 생각된다. Gilad와 Kilger의 연구에서 판단에 필요한 단서가 제시되어 있었지만, 본 연구에서는 추론에 필요한 단서는 주어지지 않고, 추론과 다른 맥락에서 의미 활성화만 발생하였다. 예를 들어, 문 앞에 있는 남자를 추론할 때 (과제 1) 남자를 ‘도둑’ 혹은 ‘택배기사’로 특정할 수 있는 단서들은 (인상착의, 행동, 말투 등) 없었다. 대부분의 진단적 추론연구에서 추론에 필요한 단서가 제시되지 않은 경우는 거의 없으며, HyGene모델에서도 단서를 기반으로 가설이 생성된다는 것을 전제한다. 본 실험에서는 순수하게 다른 맥락에서의 활성화가 진단적 추론에 영향을 미칠 수 있는지를 검증하기 위해서 고려해야 할 단서를 제시하지 않았다. 이러한 차이로 인해, 기존 연구들과 다른 결과를 초래하였을 가능성이 있다.

실험 1과 실험 2에 대한 논의

본 연구는 진단적 추론과 다른 맥락에서 유발된 의미 활성화가 가설 생성에 영향을 미치는지, 그리고 가설의 유형에 따라 각각 다른 영향을 미치는지를 알아보 고자 하였다. 두 개의 실험을 통해 다른 맥락에서의 활성화는 가설 생성에 영향을 미치는데, 가설 유형이 배타적인 경우 최종 가설 패턴에 영향을 주고, 가설의 유형 이 양립 가능한 경우 가설의 생성 수에 영향을 미친다는 것이 확인되었다. 하지만 다른 맥락에서의 활성화는 가설의 유형에 관계없이 추론의 확신에 영향을 미치지 않았다. 이상의 결과는 다음과 같이 정리될 수 있다 (표 7).

[표 7] 실험 1과 실험 2의 결과 요약

	실험 1 (쉬운 문제를 통해 의미 활성화 발생)			실험 2 (어려운 문제를 통해 의미 활성화 발생)		
	최종가설 선택패턴	가설의 생성 개수	추론 확신	최종가설 선택패턴	가설의 생성 개수	추론 확신
배타 가설	활성화에 따라 최종 가설 선택	N/S ⁵⁾	N/S	활성화에 따라 최종 가설 선택	N/S	N/S
양립 가설	N/S	활성화에 따라 가설 개수 감소	N/S	N/S	활성화에 따라 가설 개수 감소	N/S

Kahneman의 주장에 따르면(2011) 인지 시스템 2의 활성화를 위한 도움이 없는 경우, 사람들은 일반적으로 직관적으로 판단 한다 (System 1). 결국 직관적으로 판단하였던 실험 1과 인지적인 부담을 가중시켜 더 분석적인 사고를 유도하였던 실험 2에서 모두 다른 맥락에서의 활성화의 영향이 나타난 것은, 사고방식의 차이는 다른 맥락에서의 활성화의 영향과 어떠한 관계가 없음을 시사한다.

하지만 분명 분석적인 사고를 하게 되면 활성화의 영향이 감소한다는 연구 결과들은 존재 한다 (Stanovich, 2011; Kahneman, 2011; Gilad & Kilger, 2008). Gilad 와 Kilger의 연구에서 (2008) 주식 선택을 할 때 느리고 분석적인 판단을 하는 초심자 집단이 빠르고 직관적인 판단을 하는 전문가 집단보다 사전 활성화의 영향을 더

5) N/S=Not Significant

적게 받았다. 본 연구의 결과와 사전 연구의 결과에 일관성이 없는 가장 큰 이유 중 하나는 추론에 필요한 단서의 제시 여부이다. 기존 연구들에서 판단 및 추론에 필요한 단서가 제시되지 않은 경우는 거의 없으며, 발표된 결과들은 단서에 기초한 결과들이다. 하지만 본 연구는 제 3의 영향력을 배제하고 다른 맥락에서의 활성화가 진단적 추론에 영향을 미칠 수 있는지를 검증한다는 이유로 추론에 필요한 단서는 하나도 제시하지 않은 채, 맥락과 직접적으로 관련이 없는 의미 활성화만 제공하였다.

일상생활에서 우리가 진단적인 추론을 할 때도 단서 없이 추론을 하는 상황은 드물다. 예컨대, 자동차에 시동이 걸리지 않을 때, 우리는 보조등에 불이 들어오는지 먼저 확인한 후 배터리가 없다고 가설을 세운다. 만약 보조등을 켜고 불이 들어온다면 배터리가 없다고 추론하기보다, 추운 날씨로 인해 핸들과 자동차 열쇠를 넣는 이음새 부분이 매끄럽지 않기 때문이라는 가설로 바꾸게 될 것이다. 검사가 용의자를 지목할 때도 항상 현장에서 관찰된 데이터들을 기반으로 가설을 생성하는 것도 이와 마찬가지로이다.

이와 같이 단서가 투입되면 다른 맥락에서의 활성화가 최종 가설의 선택에 미치는 영향이 감소할 것임을 예상해 볼 수 있다. 또한 Tseng와 동료 (2008)들의 연구를 고려해보면, 단서가 투입되면 가설 생성의 고착현상도 극복할 수 있다. 실험 1과 실험 2에서 양립 가능한 가설의 경우 다른 맥락에서의 활성화의 영향으로 인해 더 적은 수의 가설이 생성되었다. 이는 사전에 활성화된 가설이 다른 가설의 생성을 간섭하기 때문인데, 새로운 단서가 투입되면 이러한 고착현상은 사라진다 (Tseng, Cagan & Kotovsky, 2008; Bassok & Novick, 2012).

이 밖에도 몇몇 연구들에서 사람들은 선택한 가설에 대한 단서가 더 많을수록 추론에 더 높은 확신을 가진다는 결과가 발표한 바 있다 (Schwarz, Sanna, Skurnik & Yoon, 2007; Haddock, Rothman, Reber, & Schwarz, 1999; Tsai, Klayman & Hastie, 2008).

종합해 본다면, 추론에 필요한 단서가 주어졌을 때 다음과 같은 점을 예상해 볼 수 있다. 1) 다른 맥락에서의 활성화가 배타적인 가설의 최종 선택패턴에 미치는 영향은 사라질 것이며, 2) 양립 가능한 가설의 생성 개수에서 나타나는 고착현상도 사라질 것이다. 3) 추론에 필요한 단서의 개수가 증가하면 추론의 확신은 높아질 것이다. 이와 같은 가능성을 확인하기 위해 실험 3을 수행하였다

실 험 3

실험 1과 실험 2를 통해 지금까지 발견한 사실은 진단적 추론을 할 때 사고 방식의 변화, 즉 직관적이거나 분석적인 사고방식만으로는 다른 맥락에서 유발된 활성화를 배제하지 못한다는 것이다. 더 합리적으로 진단적 추론을 하기 위해서 이러한 활성화 영향의 경계조건 (boundary condition)을 찾을 필요가 있다.

실험 1과 실험 2에서 추론에 필요한 단서가 제시되지 않았다는 점에 주목하여, 실험 3은 다른 맥락에서의 활성화와 더불어 추론에 직접적으로 고려해야 하는 단서가 제시되었다.

실험 3의 목적은 크게 두 가지이다. 첫째는 활성화 영향의 경계조건이 단서의 존재 여부인지를 검증하기 위함이고, 둘째는 단서 개수의 증가에 따른 가설 생성의 추가적인 변화를 확인하는 것이다.

실험 방법

실험 3은 서울 소재 S대학교 학부 생 155명을 대상으로 진행되었다. 단서 1개 조건에는 80명의 실험 참여자들이 (남 43, 여 36, 평균 연령 21.2세), 단서 3개 조건에는 총 75명의 실험 참여자들이 (남 34, 여 41, 평균 연령 21.43세) 참여하였다. 참여자들은 세 집단에 (실험집단1, 실험집단2, 통제집단) 무작위로 배정되었다.

실험에 사용된 과제의 짧은 글들은 실험 2에서 사용된 것과 같으며, 차이점은 짧은 글 속에 추론과 관련된 단서가 집단 마다 각각 1개, 3개가 포함되었다는 것이다. 예를 들어, 단서가 1개 제시된 집단의 경우, 초인종을 누른 남자는 모자를 쓰고 있는데, 도둑은 모자를 쓰고 다니고, 현재 택배 차량이 아파트에 주차되어 있다는 단서가 포함되었다 (표 8, 표 9).

[표 8] 실험 3에서 사용된 단서가 포함된 과제 1의 지문

* 실제 사용된 지문에는 단서에 굵은 표시가 제시되지 않았다.

<단서 1개>

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 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 눌렸다. **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.

<단서 3개>

박 양(32)은 요즘 집을 나서기가 무섭다. 최근 며칠 동안 아파트 단지 내에 도둑이 출몰하고 있으며, **이 도둑은 초인종을 눌러 먼저 인기척을 살피고, 주로 모자를 쓰고 다니며,**

주부가 혼자 있을 때 주로 활동하니 조심하라는 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 눌렸다. 확인해 보니 한 사람이 있었으나 모자를 쓰고 있어서 얼굴이 보이지 않았다. 지금 시간에 올 사람이 누굴까 생각하다가, 불현듯 며칠 전 부모님들께서 ‘Quick회사’ 택배로 포도를 보내주신다고 말씀하셨던 것이 생각났다. 박 씨는 아주머니들의 주의도 있고 해서 확인하기 위해 아파트 입구에 ‘Quick회사’ 택배차량이 있고, 몇 개의 택배상자들이 입구 앞에 놓여있는 것을 확인한 후, 문을 열어주기 위해 서둘러 현관문으로 향했다.

[표 9] 실험 3에서 사용된 단서가 포함된 과제 2의 지문

* 실제 사용된 지문에는 단서에 굵은 표시가 제시되지 않았다.

<단서 1개>

Tomas Sowell 박사는 1987년 워싱턴 DC에 사는 산모들을 조사하다가 흥미로운 사실을 발견하였다. ‘출산 전 건강관리’와 새로 태어난 ‘아이 몸무게’ 간의 상관관계가 있음을 발견한 것이다. 출산 전 낮은 수준의 건강관리를 받은 여자들은 대부분 흑인과 가난한 사람이었다. 정부의 잘못된 정책으로 인해 가난한 사람들이 건강관리 수혜를 입지 못해, 저 체중 아이를 낳게 되고 결과적으로 유아의 건강에 심각한 영향을 미치는 것이다. **조사 결과 정부가 후원하는 건강관리 센터는 가난한 사람들이 이용하기 힘든 지역에 위치한 곳들뿐이었다.** 하지만 이 연구가 발표되자 다른 관점으로 문제를 바라보는 사람들도 나타났다. 이러한 결과는 인종과 소득과는 무관하게, 산모의 흡연과 알코올 섭취와 관련이 있다는 것이다. 출산 전 건강관리 후원을 받지 못했던 산모들의 흡연과 알코올 섭취의 비율이 각각 2배, 4배가 더 높다는 통계가 **여러 학자들에 의해 잇따라 발표되었기 때문이다.** 하지만 이러한 주장이 제기되자, 몇몇 연구자들이 통계수치의 신빙성에 의문을 제기하였으며, 과연 저 체중 아이의 출산을 전적으로 가난한 산모의 개인 책임으로 물을 수 있느냐고 반박하였다.

<단서 3개>

수 많은 업적으로 세계적으로 존경 받는 Tomas Sowell 박사는 1987년 워싱턴 DC에 사는 산모들을 조사하다가 흥미로운 사실을 발견하였다. ‘출산 전 건강관리’와 새로

태어난 ‘아이 몸무게’ 간의 상관관계가 있음을 발견한 것이다. 출산 전 낮은 수준의 건강관리를 받은 여자들은 대부분 흑인과 가난한 사람이었다. 정부의 잘못된 정책으로 인해 가난한 사람들이 건강관리 수혜를 입지 못해, 저 체중 아이를 낳게 되고 결과적으로 유아의 건강에 심각한 영향을 미치는 것이다. **조사 결과 정부가 후원하는 건강관리 센터는 가난한 사람들이 이용하기 힘든 지역에 위치한 곳들뿐이었고, 정부의 지원 없이 개인적으로 관리를 받기엔 너무 비싸서 개인이 부담하기 힘든 금액이었다.** 하지만 이 연구가 발표되자 다른 관점으로 문제를 바라보는 사람들도 나타났다. 이러한 결과는 인종과 소득과는 무관하게, 산모의 흡연과 알코올 섭취와 관련이 있다는 것이다. 출산 전 건강관리 후원을 받지 못했던 산모들의 흡연과 알코올 섭취의 비율이 각각 2배, 4배가 더 높다는 통계가 발표되었고, **잇따른 다른 조사들에서도 이러한 통계수치가 꽤 신뢰롭다는 것을 증명하였다.** 또한 임신 중 알코올 흡연을 복용한 산모는 그렇지 않은 산모보다 2배 이상 저 체중아를 낳을 가능성이 높을 뿐 아니라 흑인과 가난한 산모들의 비타민 섭취량도 부족하다는 주장도 제기되었다. 하지만 Sowell 박사와 몇몇 학자들은 이러한 결과 조사 방식에 문제가 있다는 것을 주장했으며, 과연 저 체중 아이의 출산을 전적으로 가난한 산모의 개인 책임으로 물을 수 있느냐고 반박했다.

지문을 읽은 후 풀어야 할 문제는 실험 2에서 사용된 것과 같으며, 역시 주관식으로 자유롭게 서술하도록 하였다. 실험 2와 마찬가지로 실험 참여자들은 ‘비판적 사고력과 추론’이라는 연구로 알고 참여하였으며, 시간제한 없이 짧은 글을 읽고 이후 이어지는 문제에 답을 하였다. [표 10]을 보면 알 수 있듯이, 실험 3이 기존 실험 2와 다른 점은 단서가 들어간 지문이 사용되었다는 점이다.

[표 10] 실험 3에서 각 조건마다 주어진 자극의 차이

	단서 1개 집단			단서 3개 집단		
	실험집단1	실험집단2	통제집단	실험집단1	실험집단2	통제집단
지문	과제1: ‘도둑’, ‘택배기사’와 관련된 단서가 각각 1개씩 포함			과제1: ‘도둑’, ‘택배기사’와 관련된 단서가 각각 3개씩 포함		
	과제2: ‘흡연&알코올’, ‘정부정책’과 관련된 단서가 각각 1개씩 포함			과제2: ‘흡연&알코올’, ‘정부정책’과 관련된 단서가 각각 3개씩 포함		
문제	실험 2에서 사용된 문제와 같다.			실험 2에서 사용된 문제와 같다.		
	다른 맥락 문제 2개 포함	다른 맥락 문제 2개 포함	다른 맥락 문제 없음	다른 맥락 문제 2개 포함	다른 맥락 문제 2개 포함	다른 맥락 문제 없음

실험 결과

결과 1) 다른 맥락에서 유발된 활성화의 경계조건

예상했던 대로 실험 1과 실험 2와는 다른 결과가 관찰되었다. 단서가 한 개 주어졌을 때 추론과 다른 문제로 인한 활성화는 최종가설 선택패턴 (과제1: χ^2 (df=6)=8.768, $p>.05$; 과제2: χ^2 (df=6)=3.205, $p>.05$), 가설의 총 생성 개수 (과제1: $t(77)= 0.427$, $p>.05$; 과제2: $t(76)= 0.052$, $p>.05$), 추론의 확신에 모두 영향을 미치지 못하였다 (과제1: $t(78)= 1.039$, $p>.05$; 과제2: $t(77)= 0.844$, $p>.05$).

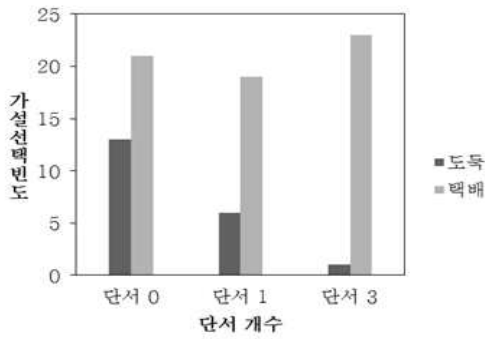
단서가 세 개 있을 때도 역시 다른 맥락에서의 활성화의 영향은 최종가설 선택패턴 (과제1: χ^2 (df=4)=3.741, $p>.05$; 과제2: χ^2 (df=6)=7.062, $p>.05$) 가설의 총 생성 수 (과제1: $t(73)= 0.163$, $p>.05$; 과제2: $t(71)= 0.386$, $p>.05$), 확신에 영향을 미치지 않았다 (과제1: $t(37)= 0.332$, $p>.05$; 과제2: $t(70)= 0.745$, $p>.05$).

결과 2) 단서의 개수의 따른 가설 생성의 변화

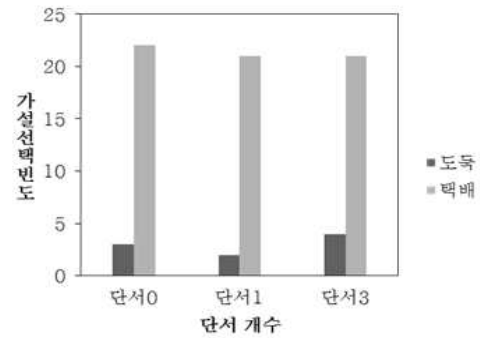
단서의 개수 변화에 따라서 가설 생성에 추가적으로 어떠한 변화가 야기됐는지 확인해 보았다. 단서 개수의 변화는 가설의 유형에 따라서 다음과 같이 최종가설 선택패턴, 가설의 총 생성 수, 추론의 확신에 다른 영향을 미쳤다.

2-1) 최종가설 선택패턴

단서 개수에 따라서 실험집단1의 배타적인 가설의 최종 선택패턴이 달라졌다 (χ^2 (df=2)=8.928, $p=.012$). 기존에 단서가 하나도 없을 때 배타적인 가설의 경우 ‘도둑’ 활성화의 영향을 받아 ‘도둑’ 가설이 편향되게 선택되었지만, 단서가 추가될수록 이러한 현상은 점차 감소하였다 (그림 8). 하지만 그림 9를 보면 알 수 있듯이 ‘택배’ 활성화를 받았던 실험집단 2에서는 단서의 투입과 상관없이 일정한 수의 택배기사 가설이 최종적으로 선택됐다 (χ^2 (df=2)=0.595, $p>.05$).



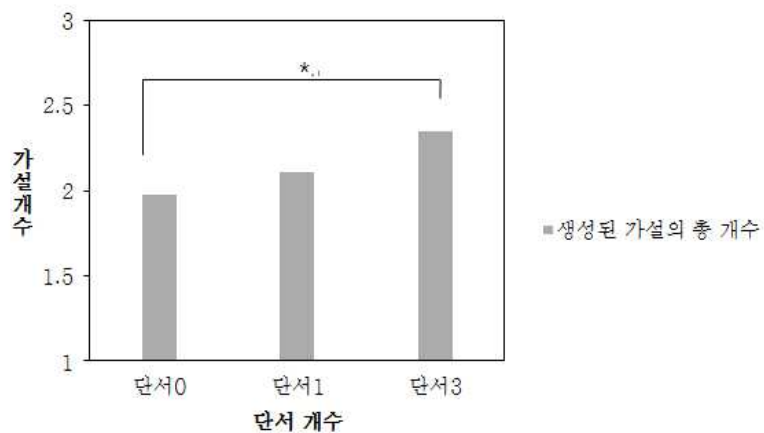
[그림 8] 단서의 개수에 따른 실험집단 1의 배타적인 가설 선택 패턴 (실험집단1은 ‘도둑’ 활성화를 받았다)



[그림 9] 단서의 개수에 따른 실험집단 2의 배타적인 가설 선택 패턴 (실험집단2는 ‘택배’ 활성화를 받았다)

2-2) 가설의 총 생성 수

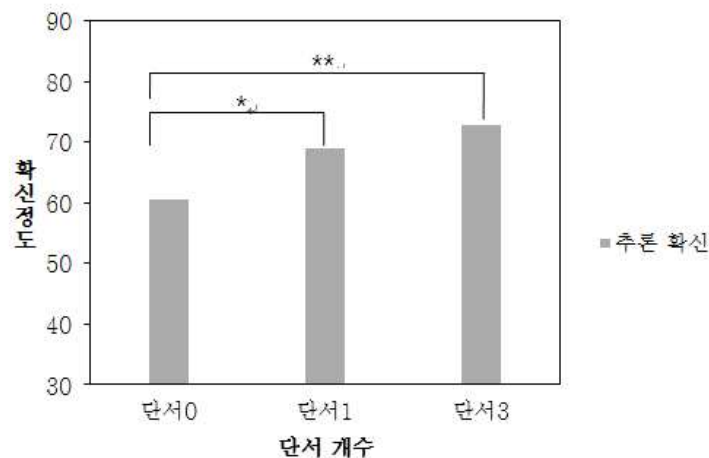
아래 그림 10이 보여주듯이, 단서의 개수가 증가할수록 양립 가능한 가설의 총 생성 수는 증가한다. 전체 실험집단에서 <과제 2>의 총 가설 생성 개수는 단서의 개수에 따라서 달라졌으며 ($F(2,154)=3.39, p<.05$) (단서0: 평균=1.98, 표준편차=0.58; 단서1: 평균=2.08, 표준편차=0.75; 단서 3: 평균=2.35, 표준편차=0.88), Scheffe와 Tukey의 사후 검정 결과 단서 0개 조건과 단서가 3개 조건 사이에 총 생성된 가설의 수에 유의미한 차이가 있음을 확인하였다 ($p<.05$). 하지만 배타적인 가설의 경우 가설의 총 생성 수는 단서 개수의 변화에 따라서 달라지지 않았다 ($F(2,159)=0.104, p>.05$)



[그림 10] 단서의 개수에 따라 총 생성된 양립 가능한 가설의 수

2-3) 추론의 확신

단서의 개수가 증가함에 따라서 배타적인 가설의 추론 확신은 더 높아진다 (그림 11). 전체 실험집단은 단서의 개수가 증가할 때 배타적인 가설의 추론에 더 확신을 가졌다 ($F(2,158)=7.557, p<.001$) (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 없었을 때와 단서가 3개 있었을 때 간에 ($p<.01$) 추론의 확신에 유의미한 차이가 발견되었다. 하지만 양립 가능한 가설의 경우 단서의 개수에 따라서 추론의 확신은 달라지지 않았다 ($F(2,152)=1.009, p>.05$).



[그림 11] 단서의 개수에 따른 배타적인 가설의 추론 확신 정도의 차이

논 의

<다른 맥락에서 유발된 활성화의 경계조건>

실험 3을 통하여, 사전에 추론 맥락과 다른 문제로 인한 의미 활성화가 진단적 추론에 미치는 영향의 경계 조건 (boundary condition)을 찾아낼 수 있었다. 추론에 고려해야할 단서가 존재할 때 다른 맥락에서의 활성화는 진단적 추론에 영향을 미치지 못한다.

<단서 개수에 따른 최종가설 선택패턴의 변화>

배타적인 가설의 경우, 단서의 개수가 증가할수록 다른 맥락에서 ‘도둑’ 활성화가 최종가설 선택에 미치는 영향은 비례해서 감소하였지만, 단서 개수의 변화에 따라 다른 맥락에서 ‘택배기사’ 활성화가 최종가설 선택에 미치는 영향에는 차이가 없었다. 이를 고려해 보았을 때, 이전 실험에서 집단 간 나타났던 다른 맥락에서의 활성화에 따른 배타적인 가설의 최종 선택의 차이는 ‘도둑’ 가설의 활성화 (실험집단 1)에 기반 했음을 알 수 있다. 이는 추론과 다른 맥락에서의 활성화는 ‘택배기사’ 가설 보다, ‘도둑’ 가설에 더 큰 영향을 미치는 것을 의미한다.

‘택배기사’ 가설과 ‘도둑’ 가설의 가장 큰 차이는 바로 기저율 (base rate)로, 일반적으로 집에 택배기사의 방문율이, 도둑의 방문율 보다 더 높다. 통제 집단의 응답을 분석해본 결과 도둑의 기저율은 약 22%, 택배기사의 기저율은 약 76%임을 확인할 수 있었다. 따라서 다른 맥락에서 유발된 활성화는 기저율이 높은 가설의 경우엔 크게 영향을 미치지 못하지만, 기저율이 낮은 가설엔 영향을 더 많이 미친다는 것을 생각해 볼 수 있으며, 이 영향은 단서의 개수가 증가할수록 급격하게 감소한다는 것을 알 수 있다. 이와 관련한 보다 상세한 논의는 <종합 논의>에서 다루도록 하겠다.

<단서 개수에 따른 가설 생성 수와 추론 확신의 변화>

사전 연구들의 결과와 마찬가지로 단서의 개수가 증가할수록 가설 생성의 고착 효과가 감소되었고, 추론의 확신이 높아진다는 것이 확인됐다. 사전 연구에서는 밝히지 못했지만 본 연구에서 추가적으로 발견한 사실은, 단서의 개수가 증가할

수록 가설의 생성 수가 증가하는 경우는 가설들이 서로 양립 가능할 때뿐이라는 것과 (과제 2), 단서가 증가할수록 확신이 증가하는 경우는 가설들이 서로 배타적일 때뿐이라는 것이다 (과제 1). 이는 기존 연구 결과들은 가설의 유형에 따라 적용여부가 달라질 수 있음을 시사한다.

종합 논의

본 연구를 통하여, 진단적 추론과 직접적으로 관련이 없는 맥락에서 유발된 의미 활성화가 가설의 유형에 따라 가설 생성에 다른 영향을 미친다는 것을 확인하였다. 이러한 활성화의 영향은 직관적인 사고를 할 때뿐 아니라 분석적인 사고를 할 때조차 유효하지만, 추론에 고려해야 할 단서가 주어지면 사라진다. 또한 실험 3을 통하여 단서의 개수의 변화는 가설 유형에 따라서 가설 생성에 다른 변화를 야기한다는 것도 확인하였다. 전체 실험의 결과를 간략하게 요약하면 다음과 같다 (표 11).

[표 11] 전체 실험 결과 정리

	실험 1		실험 2 (더 분석적인 사고를 유도)		실험 3 (추론에 필요한 단서를 제시)		
	배타 가설	양립 가설	배타 가설	양립 가설	배타 가설	양립 가설	
의미 활성화	1) 최종 가설 선택 패턴	활성화 에 따른 가설 N/S	활성화 에 따른 가설 N/S	N/S	N/S	N/S	
	2) 생성된 가설의 개수	N/S	가설 개수 감소	N/S	가설 개수 감소	N/S	
	3) 확신 정도	N/S	N/S	N/S	N/S	N/S	
단서 개수의 증가	1) 최종 가설 선택 패턴					실험집단 1의 최종 선택변화	N/S
	2) 생성된 가설의 개수					N/S	총가설 개수 증가
	3) 확신 정도					추론 확신 증가	N/S

실험 1과 실험 2를 통하여, 배타적인 가설의 경우 (과제 1) 다른 맥락에서 유발된 활성화에 따라서 편향되게 가설이 생성되지만, 양립 가능한 가설의 경우 (과제 2) 이러한 현상이 나타나지 않음을 확인하였다. 사전에 ‘도둑’ 활성화를 받은 집단은 ‘도둑’ 가설을 더 많이 선택하였다. 이 결과는 원인 마디의 수의 차이로 인해 활성화의 영향이 양립 가능한 가설보다 배타적인 가설에서 더 컸을 것으로 해석된다 (부채 효과). 또한 이는 배타적인 가설의 경우 각각의 가설이 서로 독립적이기 때문에 한 가설이 활성화를 직접적으로 받을 수 있지만, 공존 가능한 가설의 경우 활성화는 전체 가설 집합으로 퍼지기 때문으로도 설명될 수 있다.

생성된 가설의 개수의 경우, 활성화에 따라 배타적인 가설에서는 차이가 없지만, 양립 가능한 가설에서는 가설이 더 적게 생성되었다. 이 현상은 사전에 활성화를 받은 가설이 이후 새로운 가설의 생성을 간섭하기 때문으로 설명된다. 답이 하나 뿐인 단순한 배타적인 가설 생성에서는 이러한 간섭이 나타나지 않지만, 양립 가능한 가설과 같이 복잡한 가설 생성의 경우에는 한 가설의 사전 활성화가 이후 다양한 가설의 생성을 억제한다.

다른 맥락에서 유발된 활성화는 가설의 유형에 상관없이 추론의 확신에 영향을 미치지 못한다. 본 연구를 통해 단순히 활성화가 높은 가설을 선택하는 것은 추론의 확신과 어떠한 관련이 없다는 것이 밝혀졌다. 추론의 활성화를 높이는 요인은 단서에 기반 하여 활성화가 높은 가설을 선택하는 것과 관련된다.

실험 3을 통하여, 추론에 고려해야할 단서가 존재할 때 다른 맥락에서의 활성화는 진단적 추론에 영향을 미치지 못함을 밝혔으며, 단서의 개수가 증가할 때 가설의 유형에 따라 가설생성 과정의 추가적인 변화를 확인하였다.

<다른 맥락에서의 활성화가 기저율이 낮은 가설에 미치는 영향>

실험 1과 실험 2에서 배타적인 가설의 최종 선택 패턴을 보면, 다른 맥락에서 유발된 활성화에 따라서 ‘도둑’ 가설의 선택은 크게 달라졌지만, 활성화 여부와 상관없이 ‘택배기사’ 가설의 선택 빈도는 일정하였다. 또한 실험 3에서 단서의 개수가 증가할수록 활성화가 최종가설에 미치는 영향은 비례해서 감소한다는 것이 확인됐는데, 이 효과는 ‘도둑’ 활성화를 받은 집단에서만 유의미하였다. ‘택배’ 활성화를 받은 집단은 단서의 개수에 상관없이 일정하게 ‘택배’ 가설을 생성하였다. 이 현상들은 다른 맥락에서의 활성화가 최종 가설 선택 패턴에 미치는 영향이 기저율이 높은 ‘택배기사’ 가설 보다 기저율이 낮은 ‘도둑’ 가설에서 더 크다는 해석을 가능하게 한다. 왜 다른 맥락에서 유발된 활성화는 기저율이 낮은 가설에서

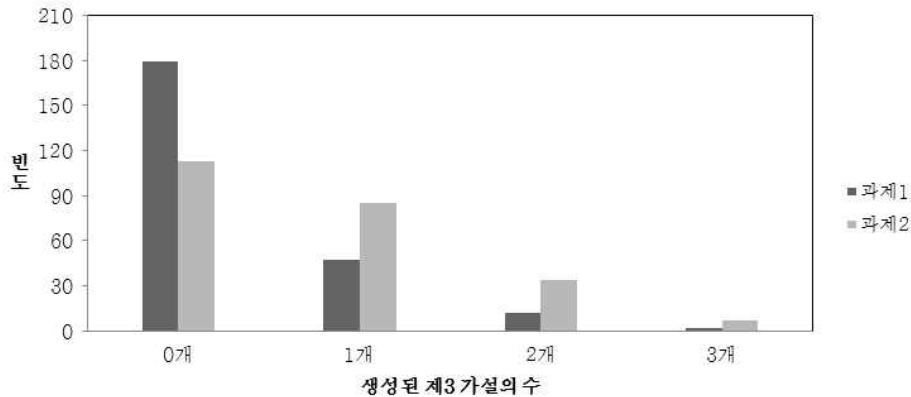
효과가 더 큰 것일까?

Wheeler 와 Berger 의 연구에서 (2007) 활성화가 조건에 의존한다는 것이 검증되었는데, 모호한 대상일수록 더 많은 해석이 필요하기 때문에, 활성화는 분명한 대상보다 모호한 대상 (ambiguous targets)에 더 큰 영향을 미친다 (Wheeler & Berger, 2007). 일반적으로 기저율이 낮은 가설은 기저율이 높은 가설보다 선택에 근거가 더 많이 필요하다. 기저율이 낮은 가설은 참이 될 확률이 낮으며, 어떠한 근거가 있기 전까지 이 가설을 참으로 생각하기 어렵다. 반면 기저율이 높은 가설은 근거가 있지 않아도 그 자체로 참이 될 확률이 크다 (반대되는 증거가 계속 쌓이기 전까지). 결국 사전 다른 맥락에서의 활성화는 기저율이 낮은 가설이 선택되어야 하는 ‘근거’를 마련해 준 셈이다. 비록 활성화가 추론과 관련 없는 맥락에서 유발되었음에도 불구하고, 활성화의 영향으로 사람들은 ‘지금 즉시 생각나기 쉽기 때문’이라는 근거가 생긴 도둑 가설을 최종적으로 선택하였다. 반면 기저율이 높은 ‘택배기사’의 가설의 경우 다른 맥락에서의 활성화가 추가적인 근거를 마련해 주지 않아도, 이미 ‘일반적으로 도둑 보다는 출현 가능성이 더 높다’는 근거를 가지고 있기 때문에 다른 맥락에서 유발된 활성화의 영향을 거의 받지 않는다.

<제 3 가설의 생성 패턴>

본 연구는 가설 생성과정을 다루고 있다. 기존 연구들에서는 주로 가설들이 이미 제시된 상태에서 가설들의 확률을 평가하거나, 객관식으로 주어진 가설들 중 하나를 선택하도록 하였다. 본 연구에서는 사람들이 가설을 어떻게 생성하는지에 관심이 있기 때문에 주관식으로 추론에 답을 하게 하여 자유롭게 가설을 생성하도록 하였으며, 추론에 고려했던 가설들이 어떤 것들인지 알아보고자 하였다. 총 3번의 실험에서는 이렇게 생성된 가설들의 개수만 파악하였지만, 사람들이 어떻게 가설을 생성하는지 그 패턴을 확인하는 것도 필요하다. 지문에 등장했던 두 개의 가설 이외에 고려했던 다른 가설들은 무엇이 있었는지, 이 가설들은 어떠한 방식으로 생성했는지 추가로 확인하였다 (그림 12).

제시된 가설 이외에 다른 가설을 생성한 수를 비교해보면, 배타적인 가설의 경우 추가로 하나도 생성하지 않은 사람의 비율은 전체 참여자 중에서 74.6%으로 가장 많았으며, 1개 생성한 사람은 19.6%, 2개 생성한 사람은 5%, 3개 생성한 사람은 0.8%였다. 양립 가능한 가설의 경우 제 3의 가설을 하나도 생성하지 않은 사람의 비율은 전체에서 47.3%, 1개는 35.6%, 2개는 14.2%, 3개는 2.9%였다. 조건 간 (다른 맥락의 활성화 여부 및 단서 개수) 제 3가설의 생성 수에는 차이가 없었다.



[그림 12] 제 3가설의 생성 현황

배타적인 가설⁶⁾의 경우 가장 많이 생성된 제 3가설은 ‘경비아저씨’와 ‘이웃주민’으로 전체 생성된 273개의 제 3의 가설에서 각각 6.6% (18명)를 차지하였다. 이 밖에 ‘방문객’이 2.9% (8명)를 차지하였으며, ‘가스 검침원’, ‘부모님’, ‘다른 택배’가 각각 1.8% (5명)를, ‘친척’, ‘친구’가 각각 1.5% (4명)를 차지하였다. 배타적인 가설의 경우 지문에 등장했던 인물들이 (경비아저씨, 이웃주민, 부모님 등) 주로 제 3가설로 생성되었으며, 이 밖에 개인의 경험 (가스 검침원, 방문객, 세탁소 아저씨, 친척 등)이 반영되어 새로운 가설을 생성했음을 알 수 있다.

양립 가능한 가설⁷⁾의 경우 ‘영양결핍 (낮은 섭취량)’이 전체 생성된 262개의 제 3가설들 중 19.1% (54명)로 가장 많이 생성되었으며, 뒤를 이어서 ‘가난 (경제력, 소득)’, ‘열악한 환경 (비위생적 환경, 낮은 생활수준)’이 각각 7.4% (21명), 6.7% (19명)를 차지하였다. 이 밖에 ‘산모의 선천적 건강 (유전)’이 5.3% (15명)였고, ‘산모의 스트레스’가 3.9% (11명), ‘충분한 휴식 부족 (과도한 노동)’이 전체에서 2.8% (8명)를 차지하였다. 양립 가능한 가설의 경우 지문에 등장하지 않았지만 근본적인 원인 (낮은 경제력과 지위, 사회 구조, 낮은 교육수준 등)이 제 3가설로 생성되었으며, 또는 ‘저 체중’과 관련되거나 (산모의 낮은 섭취량으로 저 체중의 아이 탄생), ‘임신’과 관련된 가설들(유전, 산모의 스트레스, 입덧 등)이 생성되었음을 확인하였다.

6) 배타적인 가설을 다루었던 <과제 1>에서 사람들은 초인종을 누른 사람이 ‘도둑’인지 ‘택배기사’인지 추론하였다.

7) 양립 가능한 가설을 다루었던 <과제 2>에서 사람들은 가난한 산모들이 저체중아를 낳는 원인이 ‘산모의 흡연&알코올 섭취량’ 때문인지, ‘정부의 차별된 정책’ 때문인지 추론하였다.

<단서의 진단적 역할과 추론의 확신>

실험 3에서 또한 1) 단서의 개수가 증가함에 따라 가설의 수가 증가하는 경우는 양립 가능한 가설과 같이 복잡한 가설 생성을 할 때뿐이라는 것과, 2) 개수 증가에 따라 추론의 확신이 증가하는 현상은 배타적인 가설을 생성을 할 때만 나타난다는 것을 확인하였다.

하지만 단서의 개수가 늘어남에 따라서 추론의 확신이 증가한다는 것을 밝힌 기존 연구들에서는 다른 맥락에서의 활성화가 존재하지 않았기 때문에, 본 연구에서도 단서에 따른 통제집단의 추론 확신 정도를 확인해볼 필요가 있다. 추가 분석을 실시한 결과, 본 연구의 통제집단에서 단서가 증가함에 따라서 추론의 확신은 달라지지 않았다 ($F(2,75)=1.413, p>.05$). 왜 이러한 결과가 발생한 것일까?

기존의 연구는 단서를 통해 올바른 가설을 찾아낼 수 있었다 (단서가 진단적이다). 반면 본 연구에서는 단서가 두 개의 가설에 모두 제시되었기 때문에, 단서가 주어졌다고 하더라도 단서를 통해 올바른 가설을 찾아낼 수 없었다 (단서가 비진단적이다). 이를 고려해 보았을 때, 사전 활성화가 발생하였던 실험집단에서만 단서의 개수가 증가함에 따라서 확신이 증가한 결과는 단서와 더불어 활성화라는 ‘추가적인 근거’가 함께 제시되었기 때문으로 해석해 볼 수 있다. 따라서 단서의 개수가 증가할수록 추론의 확신이 커진다는 기존의 연구 결과는 단서를 통해 올바른 가설을 찾아낼 수 있을 때만 나타나는 현상으로 한정될 수 있다.

<다른 맥락에서의 활성화와 추론의 확신>

이 밖에 추가적으로 더 고려해 보아야 할 사항은 사전 활성화와 반대로 답을 한 사람들의 판단 확신이 더 낮다는 기존의 연구 결과이다 (Erb, Bioy & Hilton, 2002). Erb와 동료들의 실험에서 ‘위험 추구’ 단어를 사전에 보았던 실험 참여자들 중에서, ‘안정적인 옵션’을 선택한 실험 참여자들의 확신은 ‘위험추구 옵션’을 선택한 실험 참여자들보다 더 낮았다. 본 연구에서도 이러한 결과가 반복되는지 알아보기 위해 추가 분석을 실시하였다.

기존 연구와는 달리 본 연구에서는 (가설의 유형, 단서의 개수에 상관없이) 이러한 현상이 나타나지 않았다. 예컨대 ‘도둑’ 활성화가 있는 집단에서 ‘도둑’ 가설을 최종적으로 선택한 사람의 확신과, ‘택배기사’를 최종적으로 선택한 사람들의 확신 간에 유의미한 차이가 없었다.

기존의 연구 결과와 본 연구의 결과가 일치하지 않는 것은, 활성화의 대상

에 차이가 있기 때문으로 생각된다. 기존의 연구 (Erb, Bioy & Hilton, 2002)는 활성화가 (위험 추구 의미) 감정에 (위험 추구) 영향을 미친 것이었지만, 본 연구에서 활성화는 (도둑 의미) 가설에 (도둑) 직접적으로 영향을 미쳤다. 결국 기존 연구에서 사전 활성화와 반대로 판단한 집단의 확신이 떨어진 것은, 활성화의 의미와 일관된 판단을 내리지 않았기 때문이 아니라, 감정에 반대되는 행동을 했기 때문으로 재해석될 필요가 있다.

<비판과 논의: 과제 난이도의 차이와 다루어진 가설들의 기저율>

본 연구에서 추론과 관련 없는 맥락에서의 활성화가 가설의 유형에 따라 다른 영향을 미친다는 것을 주장하였다. 배타적인 가설을 다루었던 <과제 1>과 양립 가능한 가설을 다루었던 <과제 2>에서 다른 맥락에서의 활성화가 가설 생성에 미친 영향이 각각 달랐기 때문이다. 하지만 두 과제 간 난이도에 차이에 의해서 활성화의 영향이 달라졌을 수 있으며, 각 과제에서 다루어진 가설들의 기저율 차이에 의해 활성화의 영향이 달라졌을 가능성도 있다.

먼저 Erb, Bioy 와 Hilton (2002)의 연구에서 더 복잡하고 이해하기 어려운 과제일수록, 사람들이 활성화보다 과제의 내용에 더 초점을 맞추기 때문에, 활성화가 미치는 영향이 감소할 수 있다는 것을 언급한 바 있다 (Erb, Bioy & Hilton, 2002). 이에 본 연구의 결과가 가설들의 유형 차이 때문인지 과제 이해도의 차이에서 비롯된 것인지 확인해 볼 필요가 있다.

본 연구의 과제 간 이해도의 차이를 비교해 본 결과, 실험 2에서 <과제 1>의 평균 이해도는 4.16 (표준편차=0.83)정도이며, <과제 2>의 평균 이해도는 3.13 (표준편차=1.08)로, 이 둘의 차이는 유의미 하였다 ($t(151.67)=6.771, p<.001$). 실험 3에서 단서가 추가 되었을 때도 이 차이는 유효 하였다 (단서 1개: 과제 1(평균=4.1, 표준편차=0.86), 과제 2 (평균=2.86, 표준편차=1.11), $t(122.573)=7.192, p<.001$); 단서 3개: 과제1 (평균=4.2, 표준편차=0.68), 과제 2 (평균=2.9, 표준편차=0.91, $t(112.63)=9.1, p<.001$). 이처럼 본 연구에서 확실히 <과제 1>과 <과제 2>간의 난이도의 차이는 존재 했다. 이에 Erb와 동료들의 연구 결과대로 단순히 활성화의 영향이 과제의 난이도에 따라 달라졌을 가능성이 있다. 하지만 이해 정도에 따라서 다른 맥락에서 유발된 활성화의 영향을 다르게 받을 수 있는지 추가적으로 검증을 실시한 결과, <과제 1>, <과제 2>에서 이해를 더 잘한 사람과 못 한 사람 간에 최종가설의 선택 패턴, 가설의 생성 수, 확신에 유의미한 차이가 없음을 확인하였다. 이를 통해 본 연구의 결과가 과제의 이해 정도에 따라서 발생했을 수 있다는 가능성은 배제된다.

또한 본 연구의 결과가 가설의 유형 차이로 인해 기인된 것이 아니라 두 과제가 다른 가설들의 기저율의 차이에 근거한다는 해석의 여지가 있다. 앞서 언급하였던 Berger의 연구 (2007)와 본 연구 실험 결과들에서, 기저율이 낮은 가설에 다른 맥락에서의 활성화가 미치는 영향이 더 크다는 것이 확인됐기 때문이다. 이를 확인하기 위해 본 연구의 통제집단을 분석하여 각 과제가 다른 가설들의 기저율을 비교해 보았다.

단서가 하나도 없을 때 배타적인 가설 (과제 1)의 경우 ‘도둑’의 기저율은 약 22%, ‘택배기사’의 기저율은 76%정도였지만, 공존 가능한 가설 (과제 2)의 경우 ‘흡연 및 알코올’의 기저율은 30%, ‘차별된 정부 정책’ 34% 정도였다. 만약 본 연구의 결과가 가설간의 유형 차이로 기인하는 것이 아니라 과제에 사용된 가설들의 기저율의 차이에서 비롯된 것이라면, 오히려 <과제 2>에서 주어진 가설들 모두 기저율이 낮기 때문에 각각의 활성화의 영향이 더 커야 한다 (과제 2의 최종 가설의 선택패턴에서 유의미한 차이가 나타났어야 한다). 하지만 결과는 배타적인 가설의 경우 (과제 1)에만 활성화의 영향으로 최종 가설 선택 패턴이 유의미하게 달라졌고, 각각의 가설의 기저율이 낮았던 <과제 2>의 경우에는 이러한 현상이 관찰되지 않았다. 또한 Kusev와 동료들의 연구 (2012)에서 의미 활성화의 영향은 기저율과는 독립적임을 밝힌 바 있다 (Kusev, Schaik & Aldrovandi, 2012).

따라서 본 연구의 결과가 과제의 난이도 차이나, 각 과제에서 다른 가설들간의 기저율의 차이에 기인할 수 있다는 주장은 타당하지 않다. 본 연구에서 밝힌 다른 맥락에서 유발된 활성화가 각 과제의 추론에 다른 영향을 미친 결과는 가설들의 유형 차이에서 비롯된 것으로 결론내릴 수 있다.

<연구의 의의 및 후속 연구 제안>

먼저 본 연구는 실제 상황에서 발생할 수 있는 가설생성과정을 다루었다. 본 연구의 결과는 이상적인 상황에서의 가설생성만 다루었던 HyGene모델이 설명하지 못하는 부분이며, 일상생활에서 사람들이 비합리적으로 진단적 추론을 할 수 있음을 경고한다. 단서에 기초하지 않고 내리는 결정은 사고방식에 관계없이 (직관적인 사고 또는 분석적인 사고) 추론과 다른 맥락에서 비롯된 활성화의 영향에 취약하기 때문에, 가설을 생성할 때 반드시 진단적 데이터에 기초해야 한다. 가설 생성은 인간 판단에 가장 기초가 되는 부분이며, 진단적 추론은 실제 거의 모든 영역에서 자주 사용되는 추론 방식이기 때문에 본 연구가 가지는 의의는 매우 크다고 하겠다.

이 밖에도 본 연구는 앞으로 가설생성, 가설평가 연구에서 가설들의 특성을 고려해야 할 필요성을 제기한다. 기존의 많은 선행 연구들에서 밝힌 결론들은 배타적인 가설만을 가지고 내린 결론이기 때문에, 이 결과가 다른 특성을 지닌 가설에도 역시 적용될 수 있을지는 의문이다. 일상생활에서는 배타적인 가설뿐 아니라 다양한 특성을 지닌 가설들이 존재한다. 이를 고려하지 않고 배타적인 가설만을 다룬다면, 연구 결과를 일반화시키기 어려울 것으로 예상된다.

본 연구에서 가설의 유형에 따라 가설 생성이 달라질 수 있음을 보여주었다. 하지만 본 연구에서도 단지 2개의 가설 유형만을 고려했을 뿐이다. 후속 연구에서 배타적이거나 양립 가능한 특성을 지닌 가설 이외에 어떤 다른 특성을 지닌 가설이 있을지 고려하여, 각각의 가설 생성 과정의 차이를 더 연구해볼 필요가 있으며, HyGene모형을 확장하여 가설의 유형까지 고려한 통합 모델을 만들기 위한 노력이 요구된다.

참 고 문 헌

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*.
- Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*(2), 186.
- Anderson, J. R., (2011). *인지심리학과 그 응용 (제 7 판)*. Ewha Womans University Press.
- Baron, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: I. Priors, error costs, and test accuracy. *Organizational behavior and human decision processes*, *41*(2), 259-279.
- Barrows, H. S., & Bennett, K. (1972). The diagnostic (problem solving) skill of the neurologist: experimental studies and their implications for neurological training. *Archives of Neurology*, *28*(3), 273.
- Berger, J., & Fitzsimons, G. (2008). Dogs on the street, pumas on your feet: How cues in the environment influence product evaluation and choice. *Journal of Marketing Research*, *45*(1), 1-14.
- Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of experimental social psychology*, *14*(1), 88-96.
- Castellan Jr, N. J., & Edgell, S. E. (1973). An hypothesis generation model for judgment in nonmetric multiple-cue probability learning. *Journal of Mathematical Psychology*, *10*(2), 204-222.
- Connell, L., & Keane, M. T. (2004). What plausibly affects plausibility? Concept coherence and distributional word coherence as factors influencing plausibility judgments. *Memory & cognition*, *32*(2), 185-197.
- 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.

- Dougherty, M. R., & Hunter, J. E. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta psychologica, 113*(3), 263-282.
- Dougherty, M. R., & Sprenger, A. (2006). The influence of improper sets of information on judgment: how irrelevant information can bias judged probability. *Journal of Experimental Psychology: General, 135*(2), 262.
- Dougherty, M. R., Gettys, C. F., & Thomas, R. P. (1997). The role of mental simulation in judgments of likelihood. *Organizational Behavior and Human Decision Processes, 70*(2), 135-148.
- Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99*(1), 3.
- Fein, S., McCloskey, A. L., & Tomlinson, T. M. (1997). Can the jury disregard that information? The use of suspicion to reduce the prejudicial effects of pretrial publicity and inadmissible testimony. *Personality and Social Psychology Bulletin, 23*(11), 1215-1226.
- Gettys, C. F., & Fisher, S. D. (1979). Hypothesis plausibility and hypothesis generation. *Organizational behavior and human performance, 24*(1), 93-110.
- Gettys, C. F., & Fisher, S. D. (1979). Hypothesis plausibility and hypothesis generation. *Organizational behavior and human performance, 24*(1), 93-110.
- Gilad, D., & Kliger, D. (2008). Priming the Risk Attitudes of Professionals in Financial Decision Making. *Review of Finance, 12*(3), 567-586.
- Green, A. E., Fugelsang, J. A., & Dunbar, K. N. (2006). Automatic activation of categorical and abstract analogical relations in analogical reasoning. *Memory & cognition, 34*(7), 1414-1421.
- Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual review of psychology, 38*(1), 369-425.
- 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.

- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive psychology*, 24(1), 1-55.
- Holyoak, K. J., & Morrison, R. G. (Eds.). (2013). *The Oxford handbook of thinking and reasoning*. Oxford University Press.
- Josephson, J. R., & Josephson, S. G. (Eds.). (1996). *Abductive inference: Computation, philosophy, technology*. Cambridge University Press.
- Kahneman, D. (2012). *생각에 관한 생각*. 김영사.
- Kassirer, J. P., & Kopelman, R. I. (1989). Cognitive errors in diagnosis: instantiation, classification, and consequences. *The American journal of medicine*, 86(4), 433-441.
- Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological bulletin*, 110(3), 499.
- 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., Davelaar, E. J., & Thomas, R. P. (2013). Data acquisition dynamics and hypothesis generation. *Cognitive Systems Research*, 24, 9-17.
- 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.
- Libby, R. (1985). Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research*, 648-667.
- Libby, R., & Frederick, D. M. (1990). Experience and the ability to explain audit findings. *Journal of Accounting Research*, 348-367.
- Lombardo, T. J. (1987). *The reciprocity of perceiver and environment: The evolution of James J. Gibson's ecological psychology*. Lawrence Erlbaum Associates, Inc.
- 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.
- Mehle, T. (1982). Hypothesis generation in an automobile malfunction inference task. *Acta Psychologica*, 52(1), 87-106.
- Mehle, T., Gettys, C. F., Manning, C., Baca, S., & Fisher, S. (1981). The availability explanation of excessive plausibility assessments. *Acta Psychologica*, 49(2), 127-140.
- Mehlhorn, K., Taatgen, N. A., Lebiere, C., & Krams, 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.
- Pastore, N. (1971). Selective history of theories of visual perception, 1650-1950.
- 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.
- Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and cognition*, 8(3), 338-342.
- Reggia, J. A., Nau, D. S., & Wang, P. Y. (1985a). 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, B. H., & Murphy, G. L. (1996). Category-based predictions: influence of

- uncertainty and feature associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 736.
- 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.
- Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *The American journal of psychology*, 61-87.
- Sprenger, A., & Dougherty, M. R. (2006). Differences between probability and frequency judgments: The role of individual differences in working memory capacity. *Organizational Behavior and Human Decision Processes*, 99(2), 202-211.
- 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.
- Stanovich, K. (2011). *Rationality and the reflective mind*. Oxford University Press.
- 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.
- 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., & Koehler, D. J. (1994). Support theory: a non extensional representation of subjective probability. *Psychological Review, 101*(4), 547.
- Waller, W. S., & Felix Jr, W. L. (1989). Auditors' causal judgments: Effects of forward vs backward inference on information processing. *Accounting, Organizations and Society, 14*(1), 179-200.
- 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.
- Wedell, D. (2010). Probabilistic reasoning in prediction and diagnosis: Effects of problem type, response mode, and individual differences. *Journal of Behavioral Decision Making, 24*: 157-179
- Wheeler, S. C., & Berger, J. (2007). When the same prime leads to different effects. *Journal of Consumer Research, 34*(3), 357-368.
- Fenton, N., & Neil, M. (2012). Risk assessment and decision analysis with Bayesian networks. CRC Press.

부 록

목 차

[부록 1] 실험 1에서 사용된 지문과 문제 세트	-----	58
[부록 2] 실험 2에서 사용된 지문과 문제 세트	-----	62
[부록 3] 실험 3에서 사용된 지문과 문제 세트	-----	68

() (%)

4. 100명중 몇 명이 택배서비스를 이용하고 있을 것 같은가? 그 이유는?

() (명)

통제집단 해당 없음

3. 지문에서 제시된 문 밖에 있는 사람이 집 안에 사람이 있는지를 신경 쓸 확률은 어느 정도 될 것 같은가?

() (%)

4. 문 밖에 있는 사람이 전과가 있을 확률을 추측해 본다면?

() (%)

5. 문 밖에 있는 사람을 추론하기 위해 필요한 정보들을 최대 2개 고르면?

수상한 가방을 들고 있었는가?	CCTV를 두려워했는가?	비가 오는 날 범죄가 일어나는 확률은?	택배차량이 밖에 있었는가?
이웃 주민들과 안면이 있는 사람인가?	택배직원들이 가장 스트레스 받는 일은 무엇인가?	경찰들이 가장 스트레스 받는 일은 무엇인가?	날씨와 범죄와의 상관관계가 있는가?

6. 문 밖에 있는 사람은 누구라고 생각하는가?

7. 6번에 응답한 답 이외에 생각한 대안(들)은 있었는가? 있다면 무엇이고, 그 답이 맞을 확률은 어느 정도라고 생각 하는가?

8. 지문 다음에 이어질 가장 그럴듯한 스토리를 추론해 보아라.

9. 지금까지의 추론에 얼마큼 확신이 있는가?

() (%)

<과제 2 지문>

[다음 지문을 읽고 질문에 답하세요]

Tomas Sowell 박사는 1987년 워싱턴 DC에 사는 산모들을 조사하다가 흥미로운 사실

3. 위 글을 읽은 전체 산모들 중 몇 퍼센트의 산모들이 분노를 느꼈을 것 같은가?

() (%)

4. '산모 알콜 금지법'을 만들면 저 체중 아이 문제가 얼마큼 해결될 것 같은가?

() (%)

5. 저 체중 아이를 낳는 원인을 추론하기 위해 필요한 정보를 최대 2개까지 고르세요.

음주가 태아에 미치는 영향은 어느 정도였는가?	여자들의 평균 주량과 산모들의 평균 주량에는 차이가 있었는가?	보통 정책의 혜택을 가장 많이 받았던 사람들은 누구인가?	저소득층과 고소득층의 출산 전 건강 관리에 어떤 차이점이 있었는가?
건강관리 정책은 누가 어떤 목적으로 만들었는가?	흡연이 태아에 미치는 영향은 어느 정도였는가?	산모의 사치품 소비율은 어떠했는가?	산모들이 스트레스 받는 주된 원인은 무엇인가?

6. 흑인과 저 소득층 산모가 저체중아를 낳는 원인은 무엇이었다고 생각하는가?

7. 6번에 응답한 답 이외에 생각한 대안(들)은 있었는가? 있다면 무엇이고, 그 답이 맞을 확률은 어느 정도라고 생각하는가?

8. 당시 위 글을 읽은 사람들이 했을 행동을 추측해 보시오.

9. 앞서 제시한 답들에 얼마큼 확신이 있는가?

() (%)

[부록 2] 실험 2에서 사용된 지문과 문제 세트

<과제 1 지문>

[다음 글을 읽고 질문에 답하세요]

박 양(32)은 요즘 집을 나서기가 무섭다. 최근 며칠 동안 아파트 단지 내에 도둑이 출몰하고 있으니 조심하라는 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 모자를 쓰고 있어서 얼굴이 보이지 않았다. 지금 시간에 올 사람이 누굴까 생각하다가, 불현듯 며칠 전 부모님들께서 포도를 보내주신다고 말씀하셨던 것이 생각났다. 박 씨는 문을 열어주기 위해 서둘러 현관문으로 향했다.

<과제 1 문제>

참여자의 성별: (남 / 여)

참여자의 나이(age): ()

다음 이어지는 질문들에 답을 하시오.

1. Ipsos InnoQuest에 따르면 글로벌 소비자의 45%는 타 지역 및 문화권에서 유래된 음식/식품을 Home dining으로 선호하는 반면, 한국 소비자의 55%는 간단히 조리할 수 있는 자작 (DIY) 포장 식품을 선호하는 것으로 나타났다. 이러한 결과가 나타난 이유는 무엇이라고 생각하는가?

2. 박 양의 부모님이 좋아하시는 음식은 ()이다.

3. 도둑들 중에서 남자의 확률은 92%이라고 하였을 때, 남자 중에서 도
실험집단1 득인 확률은 어느 정도라고 생각하는가? (한 사람이 남/여일 확률은 각각
50%라고 가정)

	() (%)
실험집단2	3. 택배기사들 중에서 남자의 확률은 92%이라고 하였을 때, 남자 중에서 택배기사인 확률은 어느 정도라고 생각하는가? (한 사람이 남/여일 확률은 각각 50%라고 가정)
	() (%)
통제집단	해당 없음

4. 박 양이 전업주부일 확률은 어느 정도라고 생각하는가?

() (%)

실험집단1	5. 2013년 조사 결과 절도 범죄자들의 평균 나이는 23세였다. 이 글을 본 어떤 사람이 범죄를 예방하기 위해서 대학생들을 대상으로 운리교육을 실시해야 한다고 주장했다. 이 주장에 대해서 간단하게 평가해 보시오.
실험집단2	5. 2013년 조사 결과 운송업체 직원들의 평균 나이는 22세였다. 이 글을 본 어떤 사람이 대학생들을 대상으로 운전교육을 실시해야 한다고 주장했다. 이 주장에 대해서 간단하게 평가해 보시오.
통제집단	해당 없음

6. 박 양의 나이는 ()이다.

7. 당신은 지문을 어느 정도 이해했다고 생각하는가? (_____ %)

8. 박 양의 집 문 밖에 있는 사람은 누구인가? 이를 어느 정도 확신하는가?

(_____)

(_____ %)

8번 답 이외에 다른 가능성을 생각해 본 적 있는가? 만약 있다면 누구(들) 이었는가?

(_____)

전체 산모들의 알코올 섭취율은 20%이다.)

() (%)

6. 5번 답에 얼마나 확신하는가? () (%)

7. 당시 조사받은 산모들은 ()에 거주하고 있었다.

8. 당신은 지문을 얼마나 이해했다고 생각하는가?

- ① 전혀 이해하지 못했다.
- ② 20%이상 이해했다.
- ③ 40%이상 이해했다.
- ④ 60%이상 이해했다.
- ⑤ 80%이상 이해했다.
- ⑥ 완벽하게 이해했다.

9. 지문의 내용을 기억하며 다음 질문에 대답 하시오.

8-1 당시 흑인과 가난한 산모들이 저체중아를 낳은 원인은 무엇이라고 생각하는가?

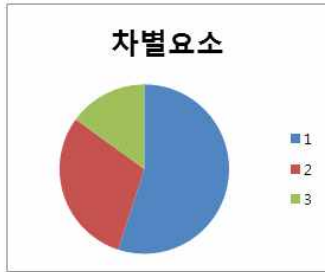
8-2 이 외에 다른 원인이 있을 수 있는가? 있다면 무엇(들)인가?

8-3. 8번 추론에 얼마나 확신하는가? (_____ %)

<실험집단2>

1. 한 매체의 조사결과 서울 시민 100명은 우리사회의 차별적 요인으로 다음과 같은 점을 언급하였다. 괄호 안에 있는 요소를 각 파이에 대응시켜 보고, 이렇게 생각한 이유를 간단하게 설명해 보시오. (예시. 1. 직업(60%) > 2. 소득수준(30%) > 3. 교육수준(10%)에 따라서 차별하는 정도가 다르다.)

('소득수준' / '교육수준' / '직업')



2. 당시 조사받은 산모들은 ()에 거주하고 있었다.
3. 2013년 조사 결과, 암 치료를 받았던 사람들 중에서 환자의 경제적 조건이나 사회적 지위에 따라 의료서비스 이용에 있어서 차별을 받는 것 같다는 응답을 한 사람들은 94.2%, ‘그렇지 않다’ 라는 응답이 5.8%로 나왔다. 어떤 사람이 이 결과를 보고 현재 대한민국 사람들은 의료서비스에 대단히 불만을 품고 있으므로, 의료 개혁이 시급하다고 주장하였다. 이 사람의 주장에 대해서 간단하게 평가해 본다면?
4. 위 글이 쓰인 시기는 ()년도) 이다.
5. Tomas Sowell박사가 태어난 연도는 ()이다.
6. 당신은 지문을 얼마나 이해했다고 생각하는가?
- ① 전혀 이해하지 못했다.
 - ② 20%이상 이해했다.
 - ③ 40%이상 이해했다.
 - ④ 60%이상 이해했다.
 - ⑤ 80%이상 이해했다.
 - ⑥ 완벽하게 이해했다.
7. 지문의 내용을 기억하며 다음 질문에 대답 하시오.
- 7-1 당시 흑인과 가난한 산모들이 저체중아를 낳은 원인은 무엇이라고 생각하는가?
- 7-2 이 외에 다른 원인이 있을 수 있는가? 있다면 무엇(들)인가?
- 7-3. 6번 추론에 얼마나 확신하는가?(_____ %)

<통제집단>

다음 이어지는 질문들에 답을 하시오

1. 위 글이 쓰인 시기는 () 년도) 이다.
2. 당시 조사받은 산모들은 ()에 거주하고 있었다.
3. Tomas Sowell박사가 태어난 연도는 ()이다.
4. 당신은 지문을 얼마나 이해했다고 생각하는가?
 - ① 전혀 이해하지 못했다.
 - ② 20%이상 이해했다.
 - ③ 40%이상 이해했다.
 - ④ 60%이상 이해했다.
 - ⑤ 80%이상 이해했다.
 - ⑥ 완벽하게 이해했다.

지문의 내용을 기억하며 다음 질문에 대답 하시오.

5. 당시 흑인과 가난한 산모들이 저체중아를 낳은 원인은 무엇이라고 생각하는가?
이 외에 다른 원인이 있을 수 있는가? 있다면 무엇(들)인가?
5번 추론에 얼마나 확신하는가?

[부록 3] 실험 3에서 사용된 지문과 문제⁸⁾ 세트

<단서 1개 지문- 과제1>

[다음 글을 읽고 질문에 답하세요]

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 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 열린 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 눌렀다. 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.

<단서 1개 지문- 과제2>

[다음 글을 읽고 질문에 답하세요]

Tomas Sowell 박사는 1987년 워싱턴 DC에 사는 산모들을 조사하다가 흥미로운 사실을 발견하였다. ‘출산 전 건강관리’와 새로 태어난 ‘아이 몸무게’ 간의 상관관계가 있음을 발견한 것이다. 출산 전 낮은 수준의 건강관리를 받은 여자들은 대부분 흑인과 가난한 사람이었다. 정부의 잘못된 정책으로 인해 가난한 사람들이 건강관리 수혜를 입지 못해, 저 체중 아이를 낳게 되고 결과적으로 유아의 건강에 심각한 영향을 미치는 것이다. 조사 결과 정부가 후원하는 건강관리 센터는 가난한 사람들이 이용하기 힘든 지역에 위치한 곳들뿐이었다. 하지만 이 연구가 발표되자 다른 관점으로 문제를 바라보는 사람들도 나타났다. 이러한 결과는 인종과 소득과는 무관하게, 산모의 흡연과 알코올 섭취와 관련이 있다는 것이다. 출산 전 건강관리 후원을 받지 못했던 산모들의 흡연과 알코올 섭취의 비율이 각각 2배, 4배가 더 높다는 통계가 여러 학자들에 의해 잇따라 발표되

8) 실험3에서 사용된 문제세트는 실험2에서 사용된 문제세트와 완전히 동일하다.

있기 때문이다. 하지만 이러한 주장이 제기되자, 몇몇 연구자들이 통계수치의 신빙성에 의문을 제기하였으며, 과연 저 체중 아이의 출산을 전적으로 가난한 산모의 개인 책임으로 물을 수 있느냐고 반박하였다.

<단서 3개 지문- 과제1>

[다음 글을 읽고 질문에 답하세요]

박 양(32)은 요즘 집을 나서기가 무섭다. 최근 며칠 동안 아파트 단지 내에 도둑이 출몰하고 있으며, 이 도둑은 초인종을 눌러 먼저 인기척을 살피고, 주로 모자를 쓰고 다니며, 주부가 혼자 있을 때 주로 활동하니 조심하라는 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 눌렸다. 확인해 보니 한 사람이 있었으나 모자를 쓰고 있어서 얼굴이 보이지 않았다. 지금 시간에 올 사람이 누굴까 생각하다가, 불현듯 며칠 전 부모님들께서 ‘Quick회사’ 택배로 포도를 보내주신다고 말씀하셨던 것이 생각났다. 박 씨는 아주머니들의 주의도 있고 해서 확인하기 위해 아파트 입구에 ‘Quick회사’ 택배차량이 있고, 몇 개의 택배상자들이 입구 앞에 놓여있는 것을 확인한 후, 문을 열어주기 위해 서둘러 현관문으로 향했다.

<단서 3개 지문- 과제2>

[다음 글을 읽고 질문에 답하세요]

수 많은 업적으로 세계적으로 존경 받는 Tomas Sowell 박사는 1987년 워싱턴 DC에 사는 산모들을 조사하다가 흥미로운 사실을 발견하였다. ‘출산 전 건강관리’와 새로 태어난 ‘아이 몸무게’ 간의 상관관계가 있음을 발견한 것이다. 출산 전 낮은 수준의 건강관리를 받은 여자들은 대부분 흑인과 가난한 사람이었다. 정부의 잘못된 정책으로 인해 가난한 사람들이 건강관리 수혜를 입지 못해, 저 체중 아이를 낳게 되고 결과적으로 유아의 건강에 심각한 영향을 미치는 것이다. 조사 결과 정부가 후원하는 건강관리 센터는 가

난한 사람들이 이용하기 힘든 지역에 위치한 곳들뿐이었고, 정부의 지원 없이 개인적으로 관리를 받기엔 너무 비싸서 개인이 부담하기 힘든 금액이었다. 하지만 이 연구가 발표되자 다른 관점으로 문제를 바라보는 사람들도 나타났다. 이러한 결과는 인종과 소득과는 무관하게, 산모의 흡연과 알코올 섭취와 관련이 있다는 것이다. 출산 전 건강관리 후원을 받지 못했던 산모들의 흡연과 알코올 섭취의 비율이 각각 2배, 4배가 더 높다는 통계가 발표되었고, 잇따른 다른 조사들에서도 이러한 통계수치가 꽤 신뢰롭다는 것을 증명하였다. 또한 임신 중 알코올 흡연을 복용한 산모는 그렇지 않은 산모보다 2배 이상 저체중아를 낳을 가능성이 높을 뿐 아니라 흑인과 가난한 산모들의 비타민 섭취양도 부족하다는 주장도 제기되었다. 하지만 Sowell 박사와 몇몇 학자들은 이러한 결과 조사 방식에 문제가 있다는 것을 주장했으며, 과연 저 체중 아이의 출산을 전적으로 가난한 산모의 개인 책임으로 물을 수 있느냐고 반박했다.

Abstract

The Effect of Semantic Activation in the Different Context on Hypothesis Generation

Younha Lee

Interdisciplinary Program in Cognitive Science

The Graduate School of

Seoul National University

People make diagnostic inferences every day. Imagine that your car suddenly does not start. You generate several hypotheses to come up with the reason, such as “the battery might be dead,” or “the car’s engine might have failed.” Likewise diagnostic inference is the cognitive process of inferring possible causes from the effect. Because we depend on hypotheses in uncertain situations, hypothesis generation is the fundamental and essential cognitive process for diagnostic inference.

Although there have been some studies on the effect of semantic priming in the judgment context, little research has been done about the effect of the semantic activation on the diagnostic inference. Moreover, previous researches about hypothesis generation have dealt with only mutually exclusive hypotheses. In the mutually exclusive hypotheses, just one hypothesis can be true.

This research explores how the semantic activation affects the process of the hypothesis generation differently by the type of hypotheses. Hypotheses fall into two categories: mutually exclusive hypothesis or not. Specifically this research anticipated that the semantic activation affects differences in 1) the selection of the final hypothesis, 2) the number of hypotheses and 3) the inference confidence, according to whether two

hypotheses are mutually exclusive or not. Also, how cue(s) plays a role in this context was studied.

Three experiments were conducted. Subjects read two short stories and made diagnostic inferences. In the first experiment, subjects in the experimental groups were given simple questions before making diagnostic inference. The simple questions were unrelated to the diagnostic inference but contained the same concept of one of the hypotheses which was given in the diagnostic context. Subjects in the control group did not answer these questions in advance, but made only diagnostic inferences.

As a result, semantic activation affected the selection of the final hypothesis in a case of the mutually exclusive hypotheses, but not the number of hypotheses. Conversely, in a case of the non-mutually exclusive hypotheses (i.e., both hypotheses can be true), the semantic activation affected generating hypotheses (i.e., the number of generated hypotheses varied according to whether the activation existed or not), but not the selection of the final hypothesis. The semantic activation did not have any influence on confidence in the diagnostic inference.

The experiment 2 and experiment 3 were conducted to find the boundary condition of the semantic effect on the diagnostic inference. In experiment 2, more difficult problems were given before taking the diagnostic problem, instead of the easy problems given in experiment 1. By doing so, subjects were induced to think harder. If the results of experiment 1 were caused by shallow thinking, slow and analytical thinking (i.e., Cognitive System 2 (Kahneman, 2011; Stanovich, 2011)) will prevent the semantic activation in the different context from affecting diagnostic inference. However, experiment 2 repeated the results of experiment 1.

In the experiment 3, cue was added. In the previous experiments, any cue to be considered was not given when subjects were making diagnostic inference. As a result, when at least one cue existed in the diagnostic context, the effect of semantic activation on the hypothesis generation disappeared. As the number of cues increased, more hypotheses generated in a case of the non-mutually exclusive hypotheses, and inference confidence increased in a case of mutually exclusive hypotheses.

In sum, this research revealed not only the semantic activation caused in the different context from diagnostic inference can affect the process of hypothesis generation in diagnostic inference, but this effect is not same for different types of hypotheses (mutually exclusive or not). Regardless of how people think (i.e., shallow thinking or deep thinking), semantic activation in different context from diagnostic inference affects the selection of final hypothesis and the number of hypotheses differently according to the type of hypotheses. However this effect disappeared when a cue was involved. Moreover when the number of cues increased, the number of hypotheses and the inference confidence varied according to the type of hypotheses.

These results caution people about making irrational diagnostic inferences due to previous context in daily life, and make suggestion that researches studying hypothesis generation/evaluation should take the type of hypotheses into account.

.....

Keywords: diagnostic inference, hypothesis generation, type of hypotheses, priming, cue

Student Number: 2012-22916