

RESEARCH ARTICLE

Development of a Survey Tool to Diagnose Cognitive Biases in Science Researchers' Research Environment

Hyojeong Hwang · Eunju Park · Minsu Ha*
Kangwon National University

과학연구자의 연구환경에서 인지편향적 요소를 진단하는 설문도구 개발

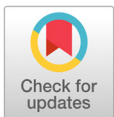
황효정 · 박은주 · 하민수*
강원대학교

*Corresponding Author: msha@kangwon.ac.kr

ABSTRACT

The purpose of this study is to develop a survey tool that diagnoses cognitive biases in a scientist's research environment. The survey questions were developed based on 40 cognitive biases that are believed to be present in the scientists' research environment. A survey of 83 scientist was conducted to verify the statistical validity and reliability of the developed questions. The questions developed can be used as a single-dimensional measurement tool. The survey tool demonstrates an appropriate level of statistical validity and reliability. This survey tool can be used to measure science researchers' sensitivity to cognitive bias. It can also be useful for checking the research environment.

Key words: Cognitive bias, scientist, question development, validity, science research



OPEN ACCESS

Brain, Digital, & Learning
2021, Vol. 11, No. 2, 393-403.

<https://doi.org/10.31216/BDL20210025>

Received: March 21, 2021

Revised: May 14, 2021

Accepted: May 31, 2021

© 2021. Institute of Brain based Education,
Korea National University of Education



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

과학 지식은 객관적이고 보편타당한 근거를 바탕으로 지식을 생성하여야 하지만, 과학 연구의 결과들을 분석한 여러 연구들은 과학자의 논증이 항상 객관적이고 보편타당하지 않을 수 있음을 보여준다(Ioannidis, 2005). 예를 들어 Ioannidis (2005)는 출판된 많은 연구들이 결과가 거짓이거나 또는 매우 과장될 가능성이 있다고 보고하였다. 이와 같이 연구 결과들이 잘못되었거나, 또는 과장되는 과정이 연구자의 의도도 아주 일부 있겠지만, 대부분의 과정은 연구자의 편향에 기인한다(Topol, 2004). 일반적으로 인간은 자신의 믿음을 뒷받침할 수 있는 정보와 증거를 찾으려고 노력하고, 그렇지 않은 정보와 증거를 무시하는 경향이 있다(Klayman, 1995). 의도하지 않더라도 일부 과학자는 무의식적으로 자신의 주장을 뒷받침하는 정보들을 선택적으로 받아들일 수 있다. 이와 같은 심리적 속성에 의하여 자신도 모르게 잘못된 과학적 주장을 할 수 있다(Nuzzo, 2015). 그리고 그런 심리적 속성 중에 대표적인 것이 인지편향이다.

인지편향은 1972년 Amos Tversky와 Daniel Kahneman에 의해 대중화된 이론이다. 인지편향은 의사결정, 기억, 행동, 판단 등과 같이 인간의 사고 과정에 큰 영향을 주는 심리적 속성이다(Haselton et al., 2016). 인지편향에 관련된 많은 심리학 이론은 인간의 인지적 편견이 직관에 의존하며 분석적 추론을 수행하는 데 어려움이 있다고 강조한다. 생각하는 과정에서 많은 오류가 발생하는데, 제한된 기억력과 주의력, 인과관계에 대한 오류, 위험에 대한 회피 등 빠르고 안전한 방향으로 탐색하려는 인간의 의식에 근거한다(Kahneman & Tversky, 1972; Tversky & Kahneman, 1973, 1974). 인간은 빠르고, 직관적이고, 무의식적인 사고를 담당하는 시스템1과, 느리고 통제되며 의식적으로 작동하는 시스템2, 두 가지의 사고 시스템을 가지고 있다. 시스템1이 작동하는 과정에서 다양한 인지편향이 영향을 줄 수 있는데, 이 과정을 통제하는 시스템이 시스템2이다. 하지만 높은 수준의 동기나 기대와 같은 감정적 요인, 빠르게 판단해야 하는 환경적 요인, 잠재적 위험이 있는 불확실성 등은 시스템2가 작동하는데 충분한 시간적 여유를 주지 못하여 결국 인지편향에 영향을 받게 된다.

인지편향은 현재까지 매우 넓은 범위에서 연구되고 있다. 인지편향 리스트를 정리한 위키피디아 정보¹를 보면 대략 180개 정도의 인지편향이 연구되어 왔다. 이와 함께 인지편향에 영향을 주는 다양한 요인도 함께 연구되어져 왔다. 인지편향은 사회적 압력, 문화, 감정, 개인의 동기 등 다양한 요인들에 의해 영향을 받는다(Haselton et al., 2016). 인지편향은 복잡하고, 합리적인 사고가 요구되는 경우에는 부정적인 영향을 미치나 빠르고 직관적인 정보처리가 유리한 상황에서는 긍정적일 수 있다. 다만 이 연구에서 주목하는 과학적 활동의 경우에는 빠르고 직관적인 사고는 다양한 오류를 생성할 수 있고, 잘못된 판단을 이끌어 낼 수 있기 때문에 인지편향을 경계하고 조절해야 되는 것이다. 과학자에게서도 다양한 인지편향이 영향을 준다고 알려졌다. Kahan (2012)은 과학자의 의사결정에서 예상되는 편향으로 동기기반추론을 지적하였다. 간혹 과학자가 객관적 근거에 기반하여 타당한 추론보다는 자신의 생각을 정당화하는데 더 많은 관심을 둔다. 결국, 답을 정해 놓고 역으로 답에 맞춰진 추론을 하는 동기기반추론이 진행되는 것이다(Kunda, 1990). 동기기반추론과 밀접한 인지편향은 확증편향이다. 확증편향은 자신의 주장을 지지하는 정보를 선택적으로 고려하는 편향을 의미한다(Klayman, 1995). 그 외에도 과학자의 연구활동과 관련이 높은 인지편향은 다양할 수 있다. 예를 들어 더 많은 정보가 유용할 것으로 믿는 정보편향(Baron et al., 1988), 자신의 지식이나 능력을 과대 평가하는 과신편향(Scott et al., 2003), 시간 계획에 영향을 주는 계획오류(Buehler et al., 1994) 등 다양한 인지편향이 있다. 이 연구에서는 그동안 연구된 180여 가지의 인지편향 중에서 과학자의 연구 활동과 관련이 높을 것으로 판단

되는 인지편향을 선택하여, 문항으로 개발하는 것이 목적이다. 개발된 문항을 바탕으로 실제 과학자들에게 설문문을 통해 문항의 타당도와 신뢰도, 그리고 어느 정도 수준에서 과학자들이 해당 인지편향을 경험하고 있는지 확인할 것이다.

Methodology

Research Method

과학연구자의 연구환경에서 인지편향적 요소를 진단하는 설문 도구를 개발하기 위하여 먼저 현재까지 연구된 모든 인지편향을 수집하여 정리하였다. 현재까지 연구된 인지편향은 약 180개 정도로, 그 수가 상당하다. 인지편향은 주로 의사결정, 사회적 관계, 기억 등과 관련된 내용으로 연구되어져 왔다. 이 연구에서 개발한 문항은 지난 50여 년 동안 연구되어온 인지편향의 내용을 바탕으로 과학연구자의 연구활동에 맞게 수정하여 개발되었다. 이 연구에서는 과학연구자의 연구상황에 맞는 인지편향을 재정의하거나, 새로운 구인을 설정하는 것을 최대한 피하여 검사내용의 타당도를 높이고자 하였다. 기존에 알려진 인지편향의 내용을 최대한 활용하고, 상황적 맥락만 변화시킴으로써 내용 타당도를 확보하고자 하였다(Messick, 1995). 수집된 인지편향 중에서 과학연구자의 연구환경과 밀접한 관련이 있을 것으로 확인되는 인지편향을 선별하였다. 과학연구자의 연구 상황 및 환경과 관련된 인지편향은 과학적 의사결정(e.g., 정보편향, 손실혐오), 과학적 탐구(e.g., 선택적 지각), 반성적 사고(e.g., 과신 편향), 협력 및 집단지성(e.g., 집단 사고), 권위나 집단에 대한 의존(e.g., 권위 편향, 편승 효과) 등으로 구분된다. 개발한 설문은 진단의 목적이 있기 때문에 특별한 항목으로 구분되지 않더라도 과학연구자가 자신의 연구활동에서 경험할 수 있다고 판단되는 문항은 포함시켰다. 이 과정은 과학교육전문가 4인이 함께 수행하였다. 개발된 문항과 각 문항과 관련된 인지편향은 Table 3에 제시되어 있다.

인지편향 요소를 진단하는 문항은 ‘여러분과 여러분 주위의 동료들의 연구 문화를 생각했을 때 아래의 상황에 어느 정도 동의하는지 5단계 척도에 응답해 주시기 바랍니다.’의 문구로 진단한다. 이와 같은 질문은 과학연구자 자신을 둘러싼 환경에서 전반적으로 인지편향이 어느 정도 나타나는지를 진단한다. 높은 동의를 하는 것은 이와 같은 요소를 자주 접해보았음을 의미하고 낮은 동의는 흔하게 접하지 않았음을 의미한다.

Participants

개발한 설문지는 과학 활동을 수행하는 총 83명의 석사과정 이상의 연구자에게서 수집하였다. 연구자의 연구 분야는 매우 다양하다. 반도체, 발광소자, 암세포, 영상처리, 생물 유전, 해안분지, 퇴적암, 면역, 신경, 신약, 우주 진화, 유기화학, 나노물질 합성 등 매우 다양한 분야의 연구분야의 연구원들로부터 자료를 수집하였다. 인구통계학적 정보는 현재 신분, 성별, 연령을 물었다. 해당 문항에 응답한 참여자를 기준으로 비율을 보면, 신분의 경우 석사과정이 25명(30.9%), 박사과정이 26명(32.1%), 박사후연구원 8명(9.9%), 대학 전임교원 22명(27.2%)이다. 남자는 52명(65.0%), 여자는 28명(35.0%)이었다. 연령은 평균 33.0세(표준편차=9.2)이었다(Table 1).

Table 1. Participants' demographic information

Gender	Position	Frequency	Percentage	Age
Male	Master course	14	16.9	26.4
	PhD course	13	15.7	29.9
	Post Doctoral	5	6.0	35.2
	Professor	19	22.9	45.2
	Not reported	1	1.2	37.0
Female	Master course	10	12.0	25.3
	PhD course	12	14.5	27.3
	Post Doctoral	3	3.6	34.3
	Professor	1	1.2	40.0
	Not reported	1	1.2	23.0
Not reported	Professor	1	1.2	59.0
	Not reported	3	3.6	
Total		83	100.0	

Data Analysis

과학연구환경에서 인지편향적 요소를 진단하는 설문 도구를 개발하고, 설문 도구의 점수를 활용하여 정량적 정보를 취득하기 위해서는 설문 도구가 통계적 타당도를 보여주어야 한다. 개발된 설문 문항에 대한 응답을 활용하여 통계적 타당도와 신뢰도를 확인하기 위하여 라쉬 분석을 활용하였다. 특히 이 연구에서는 Messick (1995)의 타당도 관점에서 실제로 근거한 타당도와 내적 구조에 관한 타당도의 두 가지를 확인하였다. 실제로 근거한 타당도는 문항적합도로, 내적 구조에 관한 타당도는 차원분석 등으로 확인하였다.

먼저 라쉬분석에서 자료의 수는 중요한데, 이 연구에서는 83개의 자료를 활용하여 분석하고 있다. 라쉬분석에서 자료의 수는 측정도구의 안정성과 관련이 있다(Boone et al., 2014). 50개의 자료일 경우 99%의 신뢰수준(confidence)에서 ± 1 로짓(logit)의 측정 오차가 발생할 수 있고, 100명이면 ± 0.5 로짓의 측정 오차가 발생할 수 있다. 이 연구에서는 83명 수준이기 때문에 $\pm 0.7 \sim 0.8$ 수준의 측정오차가 발생할 것으로 보인다. 따라서 이 부분에 대해서는 해석의 주의가 요구된다. 먼저 40개 문항을 다차원으로 분석하는 것이 타당한지, 단일차원으로 분석하는 것이 타당한지 차원 분석을 실시하였다. 개발된 설문 도구의 점수분포가 문항의 타당도를 확인하기 위해 적합한 수준인지 확인하기 위하여 Item Reliability를 확인하였고, 과학연구자의 인지편향적 요소를 정량화하는데 개발된 문항들이 적합하게 활용될 수 있는지 확인하기 위하여 Person Reliability를 확인하였다. 라쉬 분석의 문항 적합도(MNSQ, mean-square)를 통해 응답 반응의 적합도를 확인하였다. MNSQ 점수는 1을 기준으로 낮거나 높아지는데, 적합도의 기준은 문항의 특성에 따라 다르다. 이 연구에서 개발한 문항은 평정척도이고, 고부담 시험이 아닌 단순 설문이므로 Wright & Linacre (1994)가 제시한 0.6-1.4가 기준으로 활용될 수 있다. 또는 라쉬분석의 권위자인 Linacre 박사가 제안한 기준인 0.5-1.5도 활용될 수 있다(Boone et al., 2014). 이 연구에서는 두 기준을 모두 활용하여 문항의 적합도를 판단할 것이다. 설문 자료의 내적 일관성을 확인하기 위하여 Cronbach's alpha를 확인하였다. 인지편향적 요소의 점수에 관련된 변인을 확인하기 위하여 선형 회귀분석과 분산분석을 실시하였다. 라쉬 분석은 WINSTEPS 버전 4.1.0을 사용하였으며, 기타 통계 분석은 SPSS 23버전을 사용하였다.

Results and Discussions

개발된 문항의 통계적 타당도를 확인하기 전에 40개 문항이 단일차원으로 활용되는지, 아니면 다양한 차원의 문항으로 구분되는지 통계적 방법으로 확인하였다. 40개 문항을 바탕으로 측정용 도구로 설명이 되는 분산의 eigenvalue는 28.19(41.3%)이었다. 설명이 되지 않는 분산의 eigenvalue는 40.00(58.7%)이었다. 설명이 되지 않는 분산이 상당하긴 하지만, 이 분산에서 의미 있게 활용될 수 있는 차원이 존재하는지 확인해야 된다. 먼저 첫 번째로 구분되는 차원의 경우 eigenvalue는 3.80(5.6%), 두 번째는 3.11(4.6%), 세 번째는 2.73(4.0%), 네 번째는 2.39(3.5%), 마지막으로 다섯 번째는 2.27(3.3%)이다. 설명이 되는 분산에 비해서 의미있게 활용될 만큼 큰 차원은 존재하지 않는 것으로 확인된다. 두 번째 방법은 Pearson measure와 1~5개의 의미있는 차원에서 문항 집단에서 생성된 점수와의 상관관계(Disattenuated Correlation)를 확인하는 방법이다². Disattenuated Correlation의 값이 0.3 이하일 경우 다른 차원으로 이해할 수 있다. 이 기준으로 분석한 결과 설명되지 않는 분산에서 확인된 3번째 차원의 첫 번째 문항 집단에서 생성된 점수와의 Disattenuated Correlation이 0.21로 0.3보다 작았으며, 다른 14개 Disattenuated Correlation는 만족하였다. 40개 문항에서 생성된 점수끼리의 Spearman's rank correlation coefficient도 확인하였는데, 유의미한 부적상관관계를 보이는 문항도 없었다. 이상의 분석 내용을 고려하였을 때 40개 문항을 단일 차원으로 활용할 수 있다고 결론 내릴 수 있다.

라쉬 분석은 라쉬 모형을 근거로 예측된 점수와 실제 점수의 차이를 바탕으로 문항 적합도(mean-square, MNSQ)로 나타낸다. 적합도의 기준은 Wright & Linacre (1994)가 제시한 0.6~1.4, Boone et al.(2014)에 제시된 0.5~1.5를 활용한다. 40문항을 분석한 결과 Infit MNSQ의 경우 0.730~1.384, Outfit MNSQ의 경우에도 0.695~1.375로 적합한 것으로 확인되었다(Table 2). Person Reliability는 0.87, Item Reliability는 0.97로 아주 좋은 신뢰도 수준을 보이고 있다(DeVellis, 2003). 40문항의 내적 일관성 신뢰도(Cronbach α 계수)는 0.883으로 적합한 수준의 내적 일관성 신뢰도를 보여주고 있다(Fauth et al., 2014).

Table 2. Statistical validity and reliability index of developed questions

Item Number	Measure	Infit MNSQ	Outfit MNSQ	AID	Item Number	Measure	Infit MNSQ	Outfit MNSQ	AID
1	1.000	1.129	1.089	0.881	21	1.370	0.943	0.938	0.881
2	0.270	1.171	1.175	0.878	22	0.420	0.960	0.944	0.877
3	-0.690	1.085	1.142	0.883	23	-0.080	0.892	0.903	0.878
4	0.750	1.384	1.375	0.882	24	-0.190	0.884	0.886	0.879
5	-0.270	0.844	0.834	0.878	25	0.710	0.862	0.851	0.874
6	-0.690	1.220	1.166	0.883	26	0.710	0.783	0.752	0.878
7	0.530	0.989	0.974	0.885	27	0.050	0.884	0.894	0.873
8	0.140	1.029	1.025	0.882	28	0.120	1.194	1.248	0.887
9	-1.020	1.145	1.115	0.882	29	-1.390	1.068	1.123	0.885
10	-1.020	0.938	1.008	0.884	30	1.690	1.016	0.929	0.879
11	-1.870	0.747	0.772	0.882	31	1.210	1.061	1.009	0.881
12	-0.690	1.142	1.123	0.881	32	0.080	0.895	0.881	0.878
13	0.140	0.804	0.805	0.875	33	0.750	1.165	1.222	0.887
14	0.960	0.899	0.889	0.882	34	-0.140	1.089	1.104	0.880
15	0.530	0.854	0.841	0.878	35	0.030	0.951	0.977	0.885
16	-0.570	1.022	1.047	0.882	36	-0.470	0.838	0.841	0.878
17	-1.050	0.867	0.839	0.878	37	1.280	0.730	0.695	0.876
18	0.270	0.802	0.795	0.881	38	1.150	1.229	1.310	0.880
19	-0.060	1.028	1.048	0.880	39	-0.920	1.068	1.111	0.886
20	-0.570	1.144	1.130	0.879	40	-2.490	0.886	0.908	0.884

Fig. 1은 40개 문항과 83명의 참여자의 분포를 보여주는 Wright Map이다. 왼쪽 참여자의 분포와 오른쪽의 문항 분포를 보면 비슷한 위치에 분포한다. 오른쪽에 위치한 문항은 왼쪽에 위치한 사람의 인지편향을 구분하는 기준점이 될 수 있는데, 문항의 위치는 자의 눈금에 해당한다. 자의 눈금이 규칙적으로 배열되어 있으며 참여자의 분포와 비슷하기 때문에 측정도구가 참여자의 인지편향을 정량화하는데 적합하다고 판단할 수 있다. 가장 점수가 높은 문항(동의하기 어려운 문항)은 CB_30번(자신이 속한 연구 집단은 좋게 평가하고, 다른 연구 집단은 나쁘게 평가한다.)이며, 가장 점수가 낮은 문항(동의하기 쉬운 문항)은 CB_40번(연구 결과를 해석할 때 더 많은 정보가 있으면 좋을 것이라 생각한다.)이다.

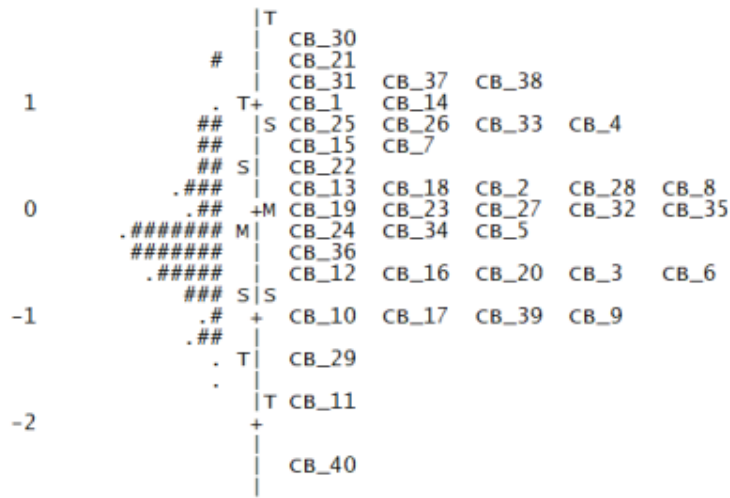


Fig. 1. Wright Map of developed 40 items

Table 3은 40개 문항에 대한 응답별로 비율을 보여주고 있다. 응답은 5단계 Likert척도로 매우 아니다(SD), 아니다(D), 보통이다(N), 그렇다(A), 매우 그렇다(SA)로 구성되어 있다. Likert척도의 점수를 1~5점으로 변환하여 통계적 타당도와 신뢰도를 확인한 것이 앞에 소개한 분석이라면, Table 2에서는 각 문항의 응답 빈도를 보여주고 있다. 문항은 Appendix에 제시되어 있다. Likert척도를 흔히 양적 정보로 변환하기 위하여 활용하지만, 질문이 긍정과 부정의 의미로 구성되어 있기 때문에 동의하는 수준의 정도로도 이해할 수 있다. Table 3는 ‘그렇다’와 ‘매우 그렇다’의 두 긍정의 응답의 비율의 합을 기준으로 정렬하여 가장 많은 동의를 보여주는 문항을 첫 번째에 나타내었다. 가장 높은 수준의 동의를 받은 문항은 ‘연구 결과를 해석할 때 더 많은 정보가 있으면 좋을 것이라 생각한다.’이다. 이 문항은 정보편향(information bias)을 의미한다. 정보편향은 의사결정을 위해서 더 많은 정보가 있으면 좋을 것이라 믿는 것으로, 실제 의사결정에 도움이 되지 않더라도 더 많으면 좋다고 믿는 편향된 사고를 의미한다(Baron et al., 1988). 통계를 위한 자료의 수나 판단을 위해 필요한 정보의 수가 일정 수준 이상일 경우 특별한 도움이 되지 않음에도 불구하고 많은 자료가 무조건 좋은 의사결정에 도움이 된다고 믿게 되면 연구의 효율성은 매우 낮아지게 된다. 물론 더 많은 정보가 도움을 줄 수도 있다. 따라서 효율적인 연구를 위해서는 연구자가 더 많은 정보를 추구할 때 정보편향에 의한 편향된 의사결정인지, 더 많은 정보가 진정으로 필요해서 구하는지에 대한 반성적 사고가 필요하다.

Table 3. The percentage of responses to the developed items(sorted by the sum of percentage of 'agree' and 'strongly agree' responses).

Item #	Cognitive bias	SD	D	N	A	SA	A/SA
40	Information Bias	0.0	0.0	7.2	66.3	26.5	92.8
11	Dunning-Kruger effect	0.0	1.2	15.7	69.9	13.3	83.1
29	Planing Fallacy	0.0	6.0	24.1	60.2	9.6	69.9
17	Bandwagon Effect	1.2	11.0	24.4	57.3	6.1	63.4
9	Stereotype	0.0	15.9	23.2	52.4	8.5	61.0
12	Hot-hand Phenomenon	3.6	14.5	28.9	49.4	3.6	53.0
39	Cognitive dissonance	0.0	8.4	38.6	49.4	3.6	53.0
6	Backfire Effect	4.8	10.8	32.5	48.2	3.6	51.8
10	Attentional Bias	0.0	3.6	45.8	43.4	7.2	50.6
20	Sunk Cost Fallacy	3.7	19.5	29.3	40.2	7.3	47.6
16	Ambiguity Effect	1.2	19.3	32.5	44.6	2.4	47.0
3	Loss Aversion	0.0	18.1	36.1	38.6	7.2	45.8
36	Halo Effect	1.2	21.7	32.5	43.4	1.2	44.6
5	Outcome Bias	2.4	26.5	32.5	37.3	1.2	38.6
24	Availability heuristic	2.4	28.9	32.5	34.9	1.2	36.1
27	Selective Perception	6.1	36.6	24.4	29.3	3.7	32.9
34	Hyperbolic Discounting	2.4	32.5	33.7	25.3	6.0	31.3
23	Appeal from Authority	3.6	30.1	36.1	26.5	3.6	30.1
19	Blind-spot Bias	3.6	30.1	37.3	25.3	3.6	28.9
13	Ad hoc Hypothesis	6.0	33.7	33.7	25.3	1.2	26.5
32	Conformity Effect	4.8	32.5	37.3	22.9	2.4	25.3
2	Social Desirability Bias	12.0	27.7	38.6	18.1	3.6	21.7
8	Elimination by Aspects	4.8	31.3	42.2	20.5	1.2	21.7
35	choice-supportive Bias	2.4	26.5	50.6	20.5	0.0	20.5
22	Groupthink	7.2	44.6	28.9	15.7	3.6	19.3
28	Wishful Thinking Bias	3.6	30.1	47.0	18.1	1.2	19.3
4	Hindsight Bias	15.7	43.4	22.9	18.1	0.0	18.1
15	Belief Bias	7.2	47.0	27.7	18.1	0.0	18.1
26	Spiral of Silence	6.0	59.0	18.1	16.9	0.0	16.9
18	Overconfidence Bias	2.4	38.6	43.4	14.5	1.2	15.7
25	Confirmation bias	12.0	49.4	22.9	12.0	3.6	15.7
7	Endowment Effect	4.9	45.1	39.0	11.0	0.0	11.0
1	Fundamental Attribution Error	15.7	51.8	22.9	7.2	2.4	9.6
33	Psychological Reactance Bias	6.0	54.2	31.3	7.2	1.2	8.4
38	Ostrich Effect	22.9	43.4	25.3	7.2	1.2	8.4
37	Self-serving Bias	19.3	55.4	18.1	7.2	0.0	7.2
14	Optimism Bias	10.8	53.0	30.1	6.0	0.0	6.0
21	Emotional bias	19.3	57.8	18.1	4.8	0.0	4.8
30	Ingroup Outgroup Bias	28.9	55.4	10.8	3.6	1.2	4.8
31	Negativity Effect	20.5	48.2	26.5	3.6	1.2	4.8

두 번째로 많은 사람이 동의한 ‘숙련되지 않은 연구자들은 무슨 문제가 있는지 몰라 더 많은 실수를 저지른다.’는 Dunning-Kruger 효과를 의미한다(Schlösser et al., 2013). 일반적으로 숙달이 되지 않은 초보자는 메타인지의 부족으로 더 많은 실수를 저지른다. 이와 같은 상황이 연구자의 주변에서 흔하게 발견이 된다면 많은 문제가 발생할 수 있다. Dunning-Kruger 효과를 방지하기 위해서는 초보자에게 상세히 안내된 매뉴얼을 제공하고, 함부로 의사결정하지 않고, 연구실에 경험이 많은 동료에게 도움을 받으면서 전문성을 높여야 된다. 세 번째로 많은 동의를 받은 문항은 ‘연구를 진행할 때 계획한 시간보다 더 오래 걸리는 편이다.’이다. 이 편향은 계획오류(planning fallacy)로 불린다. 계획오류는 작업을 완료하는 데 얼마나 많은 시간이 필요한지에 대한 예측이 낙관적인 데에서 발생한다(Buehler et al., 1994). 인지편향은 중요한 의사결정이 이루어지고, 합리적인 추론과정이 보장되어야 하는 과학연구환경에서는 흔하게 나타나서는 안 된다. 일반적인 이론에 맞게 연구 결과를 해석하는 것은 편승효과와 과학기술연구자들은 특별한 집단으로 이해하는 고정관념 모두 일반적인 인지편향이다. Table 2에서 제시된 인지편향들이 연구자의 주변에서 흔하게 나타나고 있다면 연구의 효율을 크게 저하시키거나, 잠재적인 위험 요소를 가지고 있다고 볼 수 있다.

이와 같은 요소들이 많다고 느낀다면 인지편향을 조절하는 방안에 대해서 논의해 볼 필요가 있다. Rusmana et al. (2020)은 인지편향을 조절하는 방법으로 인지편향을 알기, 인지편향을 인식하기, 인지편향 조절을 위해 행동하기, 반성하기의 네 단계를 제시하였다. Table 2에 제시된 빈도수가 높은 항목들을 바탕으로 과학연구환경을 점검할 필요가 있으며, 해당 인지편향 요소를 조절하기 위한 전략을 수립할 수 있다.

마지막으로 이 연구에서 수집한 인구통계학적 정보에 따라 인지편향 요소에 대한 인식의 점수 차이가 있는지 확인하였다. 먼저 신분으로 구분된 4개 집단의 차이를 확인한 결과 분산분석결과와 사후분석결과 모두 통계적으로 유의미하지 않았다($F(3, 77)=0.598, p=0.618$). 성별과 연령의 효과를 확인하고자 선행회귀분석을 실시하였다. 종속 변인을 인지편향에 대한 인식 점수로 하고, 독립변인으로 성별과 연령을 설정하였다. 분석 결과 R제곱은 0.084이었으며, 분산분석결과는 0.05 수준에서 유의미하였다($F(2, 76)=3.467, p=0.036$). 성별은 유의미한 효과가 없었으나($t=-1.059, p=0.293$), 연령은 유의미한 효과가 있었다($t=2.633, p=0.010$). 연령이 높아질수록 인지편향에 대한 인식 점수가 낮아진다. 이 결과에 대한 해석은 다양할 수 있는데, 먼저 연령이 높아질수록 인지편향에 대한 민감도가 낮아진다고 이해할 수 있다. 인지편향적 요소가 주변에 많이 있음에도 불구하고 인지하지 못하는 것이다. 두 번째는 연령이 높은 연구자일수록 주변에 그런 인지편향적 요소가 나타나지 않을 수도 있다. 전문성을 바탕으로 연구환경을 잘 구축할 경우 인지편향적 요소가 덜 할 수 있다. 이 연구의 자료를 바탕으로 이 두 가지 가능한 해석 중에서 어느 설명이 타당한지는 확인할 수 없다. 아마도 이 부분은 후속 연구의 과제가 되어야 할 것으로 판단된다.

Conclusions and Educational Implications

이 연구는 180여 가지의 인지편향 중에서 과학연구자의 연구 활동과 관련이 높을 것으로 판단되는 인지편향을 선택하여, 문항으로 개발하는 것을 목적으로 하였다. 개발된 문항을 바탕으로 실제 과학연구자들에게 설문문을 통해 문항의 타당도와 신뢰도, 그리고 어느 정도 수준에서 과학연구자들이 해당 인지편향을 경험하고 있는지 확인하였다. 연구 결과, 과학연구자의 연구환경에서 나타날 수 있을 것으로 판단되는 40개의 인지편향이 추출되었으며, 과학연구자의 연구상황에 맞게 40개 문항으로 개발되었다. 개발된 문항은 통계적으

로 타당하고, 신뢰할만한 것으로 확인되었다. 특히 40개 문항은 단일차원으로 활용할 수 있으며, 이처럼 분석할 경우 인지편향에 대한 인식 수준을 측정하는 점수가 획득될 수 있다.

설문조사 결과 일부 문항에서는 많은 과학연구자들이 동의하고 있었다. 과학연구자의 연구환경에서는 정보편향, Dunning-Kruger 효과, 계획오류, 편승효과, 고정관념 등이 흔하게 나타날 수 있다. 이와 같은 편향된 인식은 과학연구자의 연구 활동에 부정적인 영향을 미칠 수 있다. 이 연구에서 개발된 도구를 활용하여 과학연구자의 연구환경을 진단하면, 인지편향적 요소를 조절하는 데 큰 도움이 될 수 있다. 또한, 이 연구에서 개발한 문항들은 연구를 배우는 석사과정생이나 이공계 대학생, 또는 과학고등학교 학생 등의 교육자료로도 활용될 수 있다. 또한 문제 중심이나 프로젝트 중심의 학생 주도형 탐구 학습에도 활용될 수 있다(Hong et al., 2020; Jeong & Kim, 2020; Oh & Sohn, 2020). 인지편향을 이해해야 인지편향을 인지하고, 조절할 수 있는 역량이 길러질 수 있다. 이 연구에서 개발된 내용이 과학연구자들이 역량 신장과 효율적인 연구환경을 유지하는데 활용될 수 있을 것이다.

Acknowledgements

This work was supported by the National Research Foundation of Korea (No. NRF-2020R1C1C1006167) funded by the Ministry of Science and ICT (MSIT, Korea) and by 2018 Research Grant (PoINT) from Kangwon National University.

Endnotes

1 https://en.wikipedia.org/wiki/List_of_cognitive_biases

2 https://www.winsteps.com/winman/table23_1.htm

References

- Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, 42, 88-110.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. New York, NY: Springer.
- Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, 366.
- DeVellis, R.F. (2003). *Scale Development: Theory and Applications*. California: Sage Publications.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1-9.
- Haselton, M. G., Nettle, D., & Murray, D. R. (2016). The evolution of cognitive bias. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology: Integrations* (pp. 968-987). John Wiley & Sons, Inc.
- Hong, N. I., Lee, K., & Choi, H. (2020). Development and application of science and technology convergence general education program based on PBL for pre-service teacher. *Brain, Digital, & Learning*, 10, 431-442.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2, e124.
- Jeong, H. R., & Kim, H. (2020). The effect of climate change teaching and learning using Havruta on the high school student's attitude toward science. *Brain, Digital, & Learning*, 10, 443-452.

- Kahan, D. M. (2012). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision making*, 8, 407-424.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation*, 32, 385-418.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480-498.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Oh, H. J., & Sohn, J. (2020). Development and application of STEM education program applying problem-based learning model: Focusing on 'space exploration' subject in middle school. *Brain, Digital, & Learning*, 10, 523-533.
- Rusmana, A. N., Roshayanti, F., & Ha, M. (2020). Debiasing overconfidence among Indonesian undergraduate students in the biology classroom: An intervention study of the KAAR model. *Asia-Pacific Science Education*, 6, 228-254.
- Schlösser, T., Dunning, D., Johnson, K. L., & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the "signal extraction" counterexplanation for the Dunning-Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, 39, 85-100.
- Scott, J., Stumpp, M., & Xu, P. (2003). Overconfidence bias in international stock prices. *The Journal of Portfolio Management*, 29, 80-89.
- Topol, E. J. (2004). Failing the public health-rofecoxib, Merck, and the FDA. *New England Journal of Medicine*, 351, 1707-1709.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square-fit values. *Rasch Measurement Transactions*, 8, 370.

Authors Information

Hwang, Hyojeong: Kangwon National University, Doctoral student, First Author

ORCID: <https://orcid.org/0000-0001-8006-8950>

Park, Eunju: Kangwon National University, Postdoctoral researcher, Co-author

ORCID: <https://orcid.org/0000-0003-4675-4919>

Ha, Minsu: Kangwon National University, Professor, Corresponding Author

ORCID: <http://orcid.org/0000-0003-3087-3833>

Appendix

번호	문항
1	다른 연구자들의 실수는 그 사람이 문제가 있다고 말하면서, 자신의 실수에 대해서는 주변 상황 때문이라고 말한다.
2	연구 결과를 해석할 때 사회적으로 바람직한 가치를 지지하는 방향으로 해석한다.
3	미래에는 유망하지만 위험성이 있는 새로운 연구 주제보다 현재 잘 진행되는 자신의 연구 주제에 더 집중한다.
4	연구 결과 해석이 잘못되었다는 것을 사후에 알았을 때 처음부터 잘못된 해석임을 알았다고 설명한다.
5	실험 결과가 원하는 방향으로 나오면 연구 과정은 옳았을 것으로 판단한다.
6	연구 가설을 지지하지 않는 증거를 가질 때 가설을 포기하기보다 가설을 증명하려고 더 노력한다.
7	자신이 연구하는 주제들은 다른 사람의 연구 주제보다 더 대단하다고 생각한다.
8	여러 조건들을 고려해야 되는 복잡한 문제에 직면할 때 전체 조건을 고려하기보다 특정 변인에 집중하여 해석한다.
9	일반 대중들은 과학기술연구자들에 대해 특별한 고정관념이 있는 것 같다.
10	연구 주제에 몰두하여 다른 중요한 관찰 사실들을 놓친 적이 있다.
11	숙련되지 않은 연구자들은 무슨 문제가 있는지 몰라 더 많은 실수를 저지른다.
12	실험 과정에서 계속 성공한 경험이 있을 때 계속 잘 될 것이라 믿는다.
13	자신의 생각이나 이론에 반하는 증거를 얻었을 때에도 수긍하기보다 임시방편적이지만 그럴듯한 설명으로 방어하려고 한다.
14	특별한 이유 없이 진행 중인 연구들은 전부 잘 될 것이라고 믿는다.
15	연구 결과를 해석할 때 연구자의 신념과 맞으면 좋은 해석이라고 판단한다.
16	연구 결과를 해석할 때 의미가 있으나 불확실한 해석보다 새롭지는 않지만 명확하고 당연한 해석을 선호한다.
17	학문 분야에서 일반적으로 알려진 이론에 맞게 연구 결과를 해석한다.
18	주변의 과학기술연구자들은 자신의 능력을 과대평가한다.
19	다른 사람의 편향과 실수는 잘 지적하면서 자신의 편향이나 실수는 잘 모르는 것 같다.
20	연구에 투자한 시간과 노력이 아까워 지속적인 실패에도 포기하지 못한다.
21	연구 결과를 해석할 때 순간적인 감정에 영향을 받는 편이다.
22	연구실 내에서 갈등을 줄이려고 비판적인 생각을 하지 않고 집단의 의견에 따른다.
23	주변의 석학이나 유명한 연구자들의 생각은 옳을 것이라고 생각한다.
24	연구 결과를 해석할 때 가장 먼저 떠오르는 정보를 중요하게 고려한다.
25	연구 가설이나 생각을 지지하는 증거만 관찰하고 수집하려 한다.
26	특이한 연구 결과 해석이 떠올라도 주변의 나쁜 평가를 받을까봐 제시하지 못한다.
27	자신의 생각을 지지하거나 일치하는 것 중심으로 실험 결과를 관찰한다.
28	진행하는 연구들이 잘 될 것이라는 믿음이 항상 있다.
29	연구를 진행할 때 계획한 시간보다 더 오래 걸리는 편이다.
30	자신이 속한 연구 집단은 좋게 평가하고, 다른 연구 집단은 나쁘게 평가한다.
31	다른 연구원을 평가할 때 긍정적인 면보다 부정적인 면을 더 고려한다.
32	연구 결과를 해석할 때 가급적 사회적(학문적)으로 바람직한 방향으로 유도한다.
33	주변 동료들이 하지 말라고 하는 연구 주제들에 대해 더 관심이 생긴다.
34	미래에 수행할 수 있는 좋은 연구 주제보다 지금 당장 논문으로 출판할 수 있는 연구에 집중한다.
35	연구에 대한 의사결정을 한 이후에는 최고의 선택이라고 믿는다.
36	연구 능력이 매우 뛰어난 대가들의 연구에는 덜 비판적이다.
37	연구에 성공하면 자신의 능력 덕이라고 하면서 잘못되면 주변의 상황 때문이라고 한다.
38	연구 결과와 대치되는 이론이나 선행연구를 발견했을 때 의도적으로 보지 않으려 노력한다.
39	자신의 가설이나 생각과 다른 연구 결과를 얻었을 때 초기 가설이나 생각을 바꾼다.
40	연구 결과를 해석할 때 더 많은 정보가 있으면 좋을 것이라 생각한다.