

RESEARCH ARTICLE

A Study on The Scientific Reasoning Ability of Students Enrolled in The Gifted Education Center

Minsu Ha
Kangwon National University

과학 영재원 재학생의 과학적 추론 능력에 대한 연구

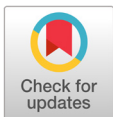
하민수
강원대학교

First Author: msha@kangwon.ac.kr

ABSTRACT

The purpose of this study was to assess the scientific reasoning abilities of students enrolled in a gifted education center. Students' scientific reasoning abilities were assessed using a translated version of the FORT (Formal Reasoning Test) developed by Kalinowski and Willoughby (2019), and data from two years of elementary and middle school students were analyzed. As a result of the study, there was a limit to analyzing 20 questions in a single dimension, but it was confirmed that the overall validity of the questions was good. The ability of elementary and middle school students to reason scientifically was found to be superior to that of college students in the United States. It was established that the ability to read and comprehend the question items had no significant effect, and some meta-thinking about confirming the correct answer occurred as well. Middle school students demonstrated a greater capacity for scientific reasoning than elementary school students, and the gap between elementary and middle school students was greater for female students.

Key words : Scientific reasoning, gifted student, gender, academic year, formal operational stage



OPEN ACCESS

Brain, Digital, & Learning
2022, Vol. 12, No. 2, 243-251.

<https://doi.org/10.31216/BDL.20220016>

Received: March 26, 2022

Revised: June 07, 2022

Accepted: June 07, 2022

© 2022. 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

과학지식을 생성하거나 알려진 과학지식을 판단하기 위해서 필수적인 사고가 과학적 추론이다. 학생들의 과학적 추론 능력은 과학 개념에 대한 이해(Alchin & Zemlén, 2020; Saad et al., 2017; Wilhelm et al., 2018)와 과학적 의사결정(Čavojová et al., 2020) 능력을 신장시키고, 일상 상황에서 학생들이 과학에 참여할 수 있도록 하는데 중요한 기능을 한다(Kind & Osborne, 2017; Van der Graaf et al., 2019). Kuhn et al. (1988, 1992), Kuhn & Dean (2004)의 연구에서 과학적 추론은 개인이 자연세계를 개념화 하고 추상적 사고와 논리적 추론을 통해 결론을 도출하는 능력을 의미한다고 하였다. Lawson (1978, 2000), Kalinowski & Willoughby (2019)의 연구에서는 과학적 추론이 형식적 조작 추론과 동일하다고 하였으며, 크게 가설 검정, 조합사고, 상관추리, 비례사고, 확률사고, 및 변인 통제에 관한 과학적 사고 기술을 과학적 추론 능력으로 정의하였다.

과학적 추론 능력은 과학 탐구를 얼마나 잘 수행할 수 있는지를 보여주는 중요한 지표이기 때문에 오래전부터 과학적 추론 능력을 측정하는 도구를 개발하고자 하였다(Burnester, 1952; Lawson, 1995). Lawson (1978)은 학생들의 교실에서 과학적 추론 능력을 진단하는 측정 도구로 CTSR(classroom test of formal reasoning)을 개발했다. Tobin & Capie (1981)도 비슷한 방법으로 Test of Logical Thinking (TOLT)를 개발하였다. 그 이후 Roadrangka et al. (1983)는 TOLT를 활용하여 우리나라에서 널리 활용된 GALT(group assessment of logical thinking)를 개발했다. 그 이후 다시 Lawson (2000)에 의하여 그의 이전 과학적 추론 검사도구에서 상관추론과 가설추론을 정교화하여 LCTSR(Lawson classroom test of scientific reasoning)을 개발하였다. 이와 같은 과정으로 과학적 추론 검사도구에 대한 개발이 이어져오다, 최근 Kalinowski & Willoughby (2019)가 FORT(Formal Reasoning Test)를 개발하였다. 그 동안 개발되어 왔던 과학적 추론 검사 도구의 타당도와 신뢰도를 높이기 위하여 개발, 수행, 분석의 과정을 반복하여 최종적으로 5개 과학적 추론 요인 측정하는 20개 문항을 개발하였다. 이 검사도구는 미국 대학생들을 대상으로 비교적 짧은 시간인 20분 정도의 시간에 5가지 유형의 추론 능력(변인 통제, 가설 설정, 상관 추리, 확률 사고, 비례 사고)을 측정할 수 있다.

이 연구에서는 이 검사도구를 번역하여 우리나라 한 대학의 영재원에 다니고 있는 초등학생과 중학생의 과학적 추론 능력을 확인하고자 하였다. 특히, 과학적 추론 능력이 발달 과정과 관련이 높고, 일부 성차에 관한 연구도 관심을 가지고 있기 때문에 초등학생과 중학생, 남학생과 여학생의 점수 차이를 확인하고 관련하여 의미를 발견하는 것을 목적으로 한다. 이 연구는 철저하게 통제된 상황에서 수집된 자료를 바탕으로 특정한 가설을 검증하는 것이 목적이 아닌 영재원에서 생성한 자료를 바탕으로 여러 경향성을 확인하고 새로운 연구에 대한 아이디어를 발견하는 것이 주된 연구의 목적이다.

Methods

Research tool

이 연구에서 Kalinowski & Willoughby (2019)가 개발한 FORT (Formal Reasoning Test) 문항을 번역하였다. 번역은 과학 교육 전문가 3인에 의하여 이루어졌다. 이 검사도구는 문항에 정답과 오답을 포함한 3~5개의 보기가 포함된 선택형 평가이다. 각 문항에는 학생이 문항을 읽고 이해되었는지, 자신의 답에 대한 확신은 어느 정도인지를 응답한다. 문항의 형태는 부록에 문항의 예시에서 확인할 수 있다.

학생들이 정답을 몰라도 우연히 맞출 수 있는 가능성이 있기 때문에 측정 오차가 발생할 수 있다. 더욱이 5가지 하부 요인으로 구성된 과학적 추론 검사가 하나의 능력을 측정하는지 아니면 여러 요소로 구성되어 있는지에 대한 면밀한 검토가 필요하다. 그래서 Winsteps을 활용하여 차원 분석을 실시하였다. 먼저 20개 문항 전체를 활용할 경우 검사도구의 측정으로 설명되는 분산의 Eigenvalue는 9.02로 31.1%가 설명되었다. 설명되지 않는 1번째 요인의 Eigenvalue는 2.08이며, 전체 분산의 7.2%이었다. 내적 일관성 신뢰도 Cronbach alpha는 0.614, 라쉬 분석을 통해 확인된 Person reliability는 0.61, Item reliability는 0.97이었다. 설명력이 40%가 넘지 않으며, 설명되지 않는 1번째 요인의 Eigenvalue가 2 이상이라는 점에서 단일 차원으로 분석하기에는 한계가 있다. 차원 분석 결과 문항 10문항 12문항 8문항 15문항 2문항 4문항 18문항 16을 제외한 12문항으로 분석할 경우 측정으로 설명되는 분산의 Eigenvalue는 8.76, 설명력은 42.2%이었으며, 설명되지 않는 1번째 요인의 Eigenvalue는 1.63이었다. Cronbach alpha는 0.741, Person reliability는 0.68, Item reliability는 0.98로 측정 도구로서 타당한기능을 갖출 수 있다(Boone et al., 2014). 따라서 이 연구에서는 20개 문항을 개별적으로 분석함과 동시에 12개 문항으로 생성된 점수를 전체 과학적 추론 점수로 하여 분석하였다. 특히 제거된 문항 10, 문항 12, 문항 8문항 15문항 2문항 4문항 18문항 16에서는 의미 있는 추가 요인이 나타나지 않아 12개 문항으로 단일 차원으로 가정하고 분석하였다. Table 1은 12개 문항의 문항 적합도(MNSQ)와 해당 문항을 삭제하였을 때 Cronbach alpha (AID, alpha if item deleted)를 보여준다. 전체 자료에서 응답을 하지 않은 비율은 1.4%이다.

Table 1. Item fitness and AID values for 12 items

Item	MEASURE	MNSQ		PTMA	AID
		Infit	Outfit		
Item1	0.55	1.165	1.413	0.456	0.735
Item3	-2.10	1.075	1.324	0.357	0.737
Item5	-0.31	1.152	1.072	0.472	0.731
Item6	-0.96	0.967	1.369	0.500	0.722
Item7	-0.21	1.049	1.069	0.516	0.728
Item9	2.52	1.129	1.422	0.405	0.742
Item11	0.00	0.776	0.640	0.668	0.694
Item13	2.51	1.034	1.177	0.439	0.735
Item14	1.20	0.898	0.767	0.608	0.714
Item17	-1.69	0.872	0.869	0.507	0.720
Item19	-0.93	0.844	0.773	0.584	0.710
Item20	-0.57	0.954	0.851	0.557	0.717

Participants

이 연구에서 사용한 자료는 한 대학의 영재원에 다니는 초등학생과 중학생을 대상으로 2년간 수집된 자료이다. 해당 대학의 영재원은 입원 경쟁률이 높지 않으며, 영재원에 입학하는 많은 학생들이 사는 지역은 전국적으로 학업 성취도가 낮은 편에 속한다. 학생들이 영재원에 입학한 이후 영재원은 학부모에게 학생의 수준을 이해할 수 있도록 창의력, 과학적 추론 능력, 과학 동기 등 다양한 능력들을 측정하여 개별적으로 안내하는 프로그램을 제공한다. 이 자료는 해당 프로그램의 한 부분으로 진행되어 수집된 자료이다. 2020년도, 2021년도에 입학한 학생들을 대상으로 수집한 자료이며, 169명의 학생의 학교급과 성별은 Table 2에 제시되어 있다.

Table 2. School level and gender of participating students

School	Male		Female		Total	
	Frequency	(%)	Frequency	(%)	Frequency	(%)
Elementary	65	(38.5)	25	(14.8)	90	(53.3)
Middle	57	(33.7)	22	(13.0)	79	(46.7)
Total	122	(72.2)	47	(27.8)	169	(100.0)

Analysis methods

20개의 과학적 추론 검사도구의 자료는 문항별 분석과 함께 단일 차원으로 판단되는 12개 문항에서 추출한 점수로 구분하여 분석하였다. 문항별로는 정답유무와 집단별 카이제곱검정을 통해 확인하였다. 정답 유무별 이해 정도도 카이제곱검정으로, 확신 점수에 대한 비교는 Mann-Whitney 검정으로 확인하였다. 라쉬 분석을 통해 문항 적합도 등의 타당도 정보, 단일 차원 분석 등을 수행하였다. 라쉬 분석은 Winsteps 4.1.0으로, 기타 통계분석은 SPSS 23으로 분석하였다.

Results

본 연구에 참여한 초등학생의 평균 점수는 20점 만점에 8.2점, 중학생은 10.5점이었다. 남학생의 평균 점수는 9.3, 여학생은 9.1점이었다. Kalinowski & Willoughby (2019)의 연구에 참여한 미국 대학생의 평균 점수는 11점인 것으로 보았을 때, 중학생은 거의 근접하고 초등학생은 약 20% 정도 부족하다. 20개 문항의 난이도를 미국 대학생들의 결과에서 확인된 난이도와 상관관계를 보면 r 은 0.80로 높은 상관관계를 보인다. 상관관계는 한국 학생의 난이도와 미국 학생의 난이도를 산점도로 나타낸 Fig. 1을 보면 분명히 확인할 수 있다. 난이도가 두 나라에서 비슷하다는 것은 과학적 추론 검사 문항의 정답률이 교육과 문화(언어)가 달라도 큰 차이가 없다는 것을 의미한다.

Table 3과 Fig. 1을 보면 본 연구의 참여자의 문항 난이도(정답률)과 Kalinowski & Willoughby (2019) 연구의 난이도의 상관관계를 대략적으로 확인할 수 있다. 또한 회색 선을 경계로 본 연구의 참여자들이 더 높은 정답률을 보이는 문항과 더 낮은 정답률을 보이는 문항도 확인할 수 있다. 비례 사고를 확인하는 문항의 경우에는 본 연구의 참여자들이 더 높은 점수를 보이며, 가설과 변인통제의 경우 전반적으로 미국의 대학생들이 더 높은 사고를 보이고 있다.

Table 4는 오답과 정답별 문항 이해 비율과 확신 점수의 평균을 보여준다. 예를 들어서 가설 능력을 평가하는 2번 문항의 경우 오답을 선택한 학생 중 94%, 정답을 선택한 학생 중 89%가 문항이 무엇을 묻고 있는지 이해된다고 응답하였다. 그리고 이 비율의 차이가 의미가 있는지를 확인하기 위해서 카이제곱검정을 실시하였고, 유의도 p 값이 제시되어 있다. 확신 점수의 경우에도 1~5점 척도인데, 오답을 선택한 학생과 정답을 선택한 학생별 확신점수의 평균이 제시되어 있다. 2번 문항의 경우 오답을 선택한 학생은 확신이 5점 척도에 3.32점, 정답을 선택한 학생은 3.56점이었다. 두 값의 차이가 유의미한지는 비모수 검정인 Mann-Whitney 검정으로 확인하였고, p 값이 제시되어 있다.

앞서 연구 방법에서 설명한 같이 20개 문항에서 단일차원으로 추정되는 12개 문항의 라쉬 점수를 바탕으로 학교급과 성별로 이원분산분석을 실시하였다. 학교급은 유의미한 차이를 보였으며, 성별은 유의미한 차이가 없었다(학교급 분석: $F[1, 165]=32.802, p=0.000, PES=0.166$; 성별 분석: $F[1, 165]=0.916, p=0.340$,

PES=0.006). 하지만 학교급과 성별의 교호작용은 유의미하였다(교호작용: $F[1, 165]=5.566, p=0.019$, PES=0.033). 교호작용이 유의미한 이유는 Fig. 2를 보면 초등학생의 경우 여학생의 점수가 약간 낮는데, 중학생은 그 반대이기 때문이다.

Table 3. The correct answer percentange and chi-square test results for each question (* Questions excluded from Rasch analysis)

Domain	Item	Correct answer rate(%)				<i>p</i>	Korea	USA	Difference
		Elementary		Middle					
		Male	Female	Male	Female				
Hypothesis testing	2*	30	30	26	19	0.790	27	58	31
	5	49	74	66	90	0.004	64	78	14
	9	20	12	16	36	0.153	20	21	1
	12*	25	13	39	14	0.031	27	47	20
Control of variables	1	40	30	60	71	0.007	49	71	22
	7	51	60	67	82	0.050	62	73	11
	16*	28	18	27	27	0.827	26	49	23
	20	63	44	77	86	0.005	68	74	6
Proportions	6	69	52	79	95	0.006	73	66	-7
	11	40	32	82	82	0.000	58	48	-10
	15*	22	16	23	14	0.759	20	15	-5
	17	82	64	84	100	0.014	82	82	0
Correlations	4*	18	13	25	29	0.465	21	44	23
	10*	45	56	31	18	0.024	38	43	5
	13	17	13	20	32	0.374	19	23	4
	18*	42	29	44	55	0.373	42	48	6
Probability	3	85	84	86	95	0.606	86	81	-5
	8*	28	20	45	30	0.082	33	68	35
	14	32	29	45	45	0.357	38	48	10
	19	68	52	82	86	0.012	73	61	-12

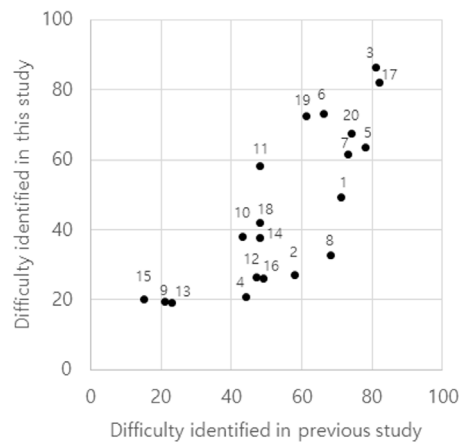
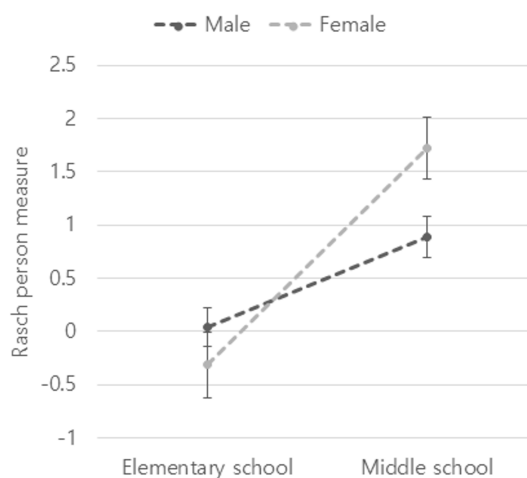


Fig. 1. Scatterplot of the difficulty of eacy items in this study and previous study

Table 4. Percentage of students who understood the item and the average score of confidence by wrong answers and correct answers.

Domain	Item	Understanding rate		<i>p</i>	Confidence score		<i>p</i>
		Incorrect	Correct		Incorrect	Correct	
Hypothesis testing	2*	94	89	0.251	3.32	3.56	0.132
	5	85	97	0.004	2.97	3.59	0.001
	9	92	94	0.680	3.41	3.70	0.149
	12*	87	95	0.117	3.15	3.34	0.234
Control of variables	1	95	96	0.724	3.94	3.99	0.667
	7	84	96	0.008	3.20	4.20	0.000
	16*	89	95	0.224	3.30	3.68	0.058
	20	86	100	0.002	3.31	4.25	0.000
Proportions	6	96	99	0.118	3.69	4.36	0.000
	11	93	99	0.035	3.81	4.53	0.000
	15*	97	97	0.995	3.75	3.53	0.199
	17	93	98	0.158	3.29	4.35	0.000
Correlations	4*	95	100	0.195	3.71	3.80	0.488
	10*	91	92	0.834	3.32	3.52	0.244
	13	89	91	0.737	3.33	3.72	0.093
	18*	92	96	0.311	3.59	3.60	0.967
Probability	3	91	95	0.439	3.82	4.44	0.001
	8*	92	94	0.544	3.39	3.83	0.028
	14	86	98	0.009	3.33	3.97	0.001
	19	98	98	0.908	3.46	4.42	0.000

**Fig. 2.** Average of Rasch person measure by school-level and gender

Discussions and Implications

이 연구는 과학 영재원에 재원 중인 학생의 과학적 추론 능력을 조사하기 위하여 Kalinowski & Willoughby (2019)가 개발한 FORT(Formal Reasoning Test) 문항을 번역하여 타당도를 확인하고, 나타난 점수를 활용하여 학교급과 성별로 분석하였다. 이 연구의 결과를 바탕으로 논의하면 다음과 같다. 먼저 FORT (Formal Reasoning Test)의 20개 문항은 1개 요인으로 분석하기에는 한계가 있다. 과학적 추론이 복잡한 사고 과정일 뿐만 아니라, 문항의 특징을 살펴보면 쉽게 이해하기 어려울 정도로 높은 이해력을 요구하는 문항이 상당수 있기 때문에 다양한 요인이 복잡하게 나타날 수 있을 것이다. 그럼에도 불구하고 12개 문항은 단일 차원의 성격을 지니고 있었으며, 12개의 문항의 문항 적합도 등 문항 타당도에 관한 통계 지표는 양호한 편이었다. 학생의 과학적 추론 능력을 진단하기 위해서 각 문항의 점수를 활용해도 되며, 과학적 추론 능력의 전체 점수를 생성하기 위해서 12개 문항만을 활용해도 될 것이다.

두 번째는 본 연구의 초등학생의 평균 점수는 20점 만점에 8.2점, 중학생은 10.5점으로 Kalinowski & Willoughby (2019)의 연구에 참여한 미국 대학생의 평균 점수(11점)를 근거로 살폈을 때 학생들의 과학적 추론 능력이 높은 것으로 확인된다. 본 연구의 중학생은 주로 중학교 1학년 학생으로 아직 저학년임에도 불구하고 미국 대학생의 점수에 근접하는 평균치를 보이고 있다. 일부 문항의 경우는 미국의 대학생 보다 더 높은 성취를 보이고 있다. Kalinowski & Willoughby (2019)의 연구가 미국의 중하위권 대학에서 진행되었기 때문에 직접적으로 비교하여 논의하기는 어려우나, 과학적 추론이 교육 수준과 발달적 측면에서 모두 영향을 받는 점을 고려하였을 때 본 연구에 참여한 초등학생과 중학생의 수준은 상당히 높다고 평가된다. 과학 영재원에서 학생들은 대부분 과학적 추론 능력이 요구되는 가설-연역적 실험 기반의 활동을 많이 수행하기 때문에 변인통제, 가설설정 등 형식적 조작기 학생들이 가능한 통합탐구능력이 많이 요구된다. 본 검사도구로 학생들의 통합탐구능력의 평가도 가능할 것으로 판단된다.

눈여겨 볼 점은 본 연구에서 확인한 FORT (Formal Reasoning Test) 문항의 난이도가 미국 대학생들의 결과에서 확인된 난이도와 상관계수가 매우 높다는 것이다. 과학적 추론 문항에는 일부 내용이 포함되어 있으며, 두 국가의 교육과정과 참여자들의 맥락이 상당히 다름에도 불구하고 상당히 높은 상관계수는 매우 의미 있는 결과로 이해될 수 있다. 다시 말하면, 과학적 추론 능력이 문화와 교육 수준 등 사회적 맥락보다는 인간 뇌 활동이라는 생물학적 요인에 더 큰 영향을 받는 것이 아닌가 하는 추론을 매우 조심스럽게 제시할 수 있다. 예를 들어 Ha & Nehm (2014)은 진화 추론에서 교육 수준과 문화가 다양한 집단을 비교하면서 교육 수준과 다른 두 국가와 상관없이 특정한 상황에 대한 진화 추론에 어려움과 쉬움이 비슷하게 나타나는 것을 관찰하였다. 그들은 결과를 바탕으로 아마도 진화 추론이 교육이나 문화적인 맥락보다는 인간의 인지적 능력에 더 큰 영향을 받을 것이라 논의하였다. 변인통제, 가설 설정, 확률적 사고 등은 맥락보다는 인지 능력을 통해서 해결해야 되기 때문에 교육 내용과 문화가 차이나더라도 비슷한 어려움을 보이는 것으로 이해될 수 있다. 물론 이 논의에 대해서는 분명히 추가적인 연구가 필요할 것으로 보이며, 상당히 흥미로운 주제가 될 것으로 판단한다.

학생들은 매우 복잡한 과학적 추론 문항을 이해하는데에는 큰 어려움이 없었던 것으로 판단된다. 정답을 선택한 학생의 이해 비율이 일부 문항에서 높기는 하지만 큰 차이가 없었다. 문항을 읽고 이해하는 능력이 정답을 선택하는데 영향을 주긴 하지만 큰 효과는 없음을 확인할 수 있다. 또한 정답을 선택한 학생의 확신 점수가 유의미하게 높은 경우가 많은 것은 자기 평가, 즉 메타적 사고까지 일부 가능한 것으로 이해된다. 학생

들은 문제를 읽고 해결한 뒤 자신의 선택한 답이 정답일 가능성이 높은지를 판단할 수 있는 능력까지 갖춘 경우가 있음을 의미한다. 확신 점수를 통해 과학적 추론에 대한 메타 능력까지 측정이 가능함을 보여주며, 이 능력에 대한 다양한 추가 연구도 요구된다.

마지막으로 초등학생에 비하여 중학생의 점수가 높으며, 특히 중학교에서는 여학생의 점수가 상당히 높을 통해 과학적 추론 능력의 발달 과정을 일부 유추할 수 있다. 중학생이 초등학생에 비하여 교육을 더 많이 받았기 때문에 과학적 추론 점수가 높아진 것이 발달로 인하여 발생한 효과인지 아니면 교육을 통해 이루어진 효과인지 단정하기는 어렵다. 그럼에도 불구하고 과학적 추론 능력의 발달이 성별에 따라 다르게 나타난다는 것은 의미있게 생각할 결과이다. 또한 이 결과를 통해 추론해 보건데, 초등학생과 중학생이 해결하기에는 20개 문항이 상당히 길며, 고도의 집중력과 끈기가 요구되는 문항이기 때문에 과학적 추론 능력 이외에 정의적인 역량의 영향도 많았을 가능성이 있다. Oh (2013)는 여학생이 남학생보다 인내력 기질이 유의미하게 높았다는 성차 연구를 발표한 바 있다. 과학적 추론과 같은 고도의 정신 활동 상황에서의 인내력 기질은 더 복잡한 양상을 보일 수 있을 것이며, FORT (Formal Reasoning Test) 검사도구를 활용한 과학적 추론에 관한 성차 연구도 흥미로운 결과를 많이 생성할 것으로 기대한다.

서론에서 밝힌 바와 같이 이 연구는 과학적 추론 능력에 대한 학교급별, 성별 차이를 확인하고자 하는 연구가 아니라 과학 영재원에서 생성된 자료를 바탕으로 의미있고 연구할 가치가 있을만한 주제들을 탐색한 연구이다. 따라서 본 연구 결과들은 매우 제한된 수준에서 이해되어야 할 것이며, 과도하게 일반화를 하지 않아야 될 것이다. 과학적 추론에 관한 타당하고 의미있는 검사도구가 개발되었고 그것을 활용하여 우리나라 학생들의 과학적 추론 능력을 확인한 자료를 분석한 결과 여러 흥미로운 결과들을 확인할 수 있었다. 체계적인 선행연구의 분석과 자료수집을 통해 과학적 추론의 발달과 그와 관련된 변인에 관한 많은 연구가 진행될 수 있기를 바라며, FORT (Formal Reasoning Test) 번역본과 이 연구의 결과들이 활용될 수 있기를 바란다.

Acknowledgement

This paper was conducted with support from college accounting of Kangwon National University in 2021.

References

- Allchin, D., & Zemplén, G. Á. (2020). Finding the place of argumentation in science education: Epistemics and whole science. *Science Education*, 104, 907-933.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Dordrecht, the Netherlands: Springer.
- Burmester, M. A. (1952). Behavior involved in the critical aspects of scientific thinking. *Science Education*, 36, 259-263.
- Čavojová, V., Šrol, J., & Jurkovič, M. (2020). Why should we try to think like scientists? Scientific reasoning and susceptibility to epistemically suspect beliefs and cognitive biases. *Applied Cognitive Psychology*, 34, 85-95.
- Ha, M., & Nehm, R. H. (2014). Darwin's difficulties and students' struggles with trait loss: Cognitive-historical parallelisms in evolutionary explanation. *Science & Education*, 23, 1051-1074.

- Kalinowski, S. T., & Willoughby, S. (2019). Development and validation of a scientific (formal) reasoning test for college students. *Journal of Research in Science Teaching*, 56, 1269-1284.
- Kind, P. M., & Osborne, J. (2017). Styles of scientific reasoning: A cultural rationale for science education? *Science Education*, 101, 8-31.
- Kuhn, D., & Dean, Jr, D. (2004). Connecting scientific reasoning and causal inference. *Journal of Cognition and Development*, 5, 261-288.
- Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., & Yotive, W. (1988). *The Development of Scientific Thinking Skills*. San Diego : Academic Press.
- Kuhn, D., Schauble, L., & Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. *Cognition and Instruction*, 9, 285-327.
- Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15, 11-24.
- Lawson, A. E. (1995). *Science Teaching and the Development of Thinking*. Belmont, CA: Wadsworth.
- Lawson, A. E. (2000). Classroom test of scientific reasoning. Revised edition of Lawson(1978) version. Unpublished Manuscript.
- Oh, H. S. (2013). Gender differences in relations between attention performance and temperament of adolescents. *Korean Journal of Clinical Psychology*, 32, 589-610.
- Roadrangka, V., Yeany, R. H., & Padilla, M. J. (1983). The construction and validation of group assessment of logical thinking (GALT). In Paper Presented at the Annual Meeting of the National Association for Research in Science Teaching, Dallas, TX.
- Saad, M. I. M., Baharom, S., & Mokhsein, S. E. (2017). Scientific reasoning skills based on socio-scientific issues in the biology subject. *International Journal of Advanced and Applied Sciences*, 4, 13-18. <https://doi.org/10.21833/ijaas.2017.03.003>
- Tobin, K. G., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement*, 41, 413-423.
- Van der Graaf, J., Van de Sande, E. Gijssels, M., & Segers, E. (2019). A combined approach to strengthen children's scientific thinking: Direct instruction on scientific reasoning and training of teacher's verbal support. *International Journal of Science Education*, 41, 1119-1138.
- Wilhelm, J., Cole, M., Cohen, C., & Lindell, R. (2018). How middle level science teachers visualize and translate motion, scale, and geometric space of the Earth-Moon-Sun system with their students. *Physical Review Physics Education Research*, 14, Article 010150.

Author Information

Ha, Minsu: Kangwon National University Professor, First Author