

한국 학령기 학생들의 관계추론능력: 한국판 관계추론능력검사 활용*

채 수 은¹⁾

강릉원주대학교 교직과정부

본 연구에서는 최근 개발된 관계추론능력 검사(Test of Relational Reasoning: TORR)를 한국 초, 중, 고 학생들에게 적용하여 관계추론능력의 하위 유형과 학년에 따라 추론 수행결과가 어떻게 다른지 살펴보고자 하였다. 이를 위해 반복측정분산분석과 차별문항기능분석을 실시하였다. 강원 지역 청소년 약 749명[초4 196명, 초6 200명, 중2 206명, 고1 146명, 초4 남학생 109명(55.33%), 초6 남학생 105명(52.50%), 중2 남학생 81명(39.32%), 고1 남학생 63명(43.15%)]을 대상으로 관계추론능력검사를 실시한 결과, 학년이 올라감에 따라 점수가 높아지는 경향을 보였다. 그러나 네 개 관계추론 영역별로 학년에 따른 점수 변화 패턴은 차이가 있었다. 초6과 중2 사이에 모순(antinomy)과 대조(antithesis) 점수가 유의하게 다른 반면, 초4와 초6 사이에는 유사성(analogy) 유추와 변칙(anomaly) 점수가 유의하게 달랐다. 본 연구의 결과를 통해서 단순히 유사성 유추만을 중심으로 한 기존의 추론 연구의 범위가 넓혀져야 함을 알 수 있다.

주요어 : 관계추론, 유사, 변칙, 모순, 대조

* 본 연구는 2019년 강릉원주대학교 장기해외파견연구 지원을 받아 수행되었음.

1) 강릉원주대학교 교직과정부, 교수 (E-mail: schae@gwnu.ac.kr)

I. 서 론

1. 연구의 필요성 및 목적

관계 추론은 “관련이 없는 것으로 보이는 정보들 사이에서 의미 있는 관계를 인식하거나 파생시킬 수 있는 능력”(Alexander & the Disciplined Reading and Learning Research Laboratory, 2012)으로 정의할 수 있다. 다양한 학문 영역에서 얻은 경험 연구 결과에 따르면 관계추론능력은 학령기 학습자로부터 전문가에 이르기까지 전 영역과 발달 단계에서 필수적이다(Jablansky, Alexander, Dumas, & Compton, 2016). 관계추론에 대한 기존 연구들을 통해서 관계추론이 아동으로부터(Goswami, 1989, 2013; Richland, Morrison, & Holyoak, 2006) 여러 연령대의 성인에 이르기까지(Alexander, 2019; Dumontheil, Houlton, Christoff, & Blakemore., 2010; Goswami, 1989, 2013; Jablansky et al., 2016) 여러 발달단계에서 일어나는 매우 기초적이며 핵심적인 인지 능력임을 알 수 있다. 따라서 인간발달의 두 가지 쟁점인 “언제 발현되느냐”와 “어떻게 발현되느냐”의 문제가 관계추론 발달과 관련해서도 연구자들 사이에서 오랜 기간 논쟁의 대상이었다.

우선 관계추론능력이 “언제 발현되느냐”와 관련하여 과거 고전적 연구자들은 개념화 과정에서 나타나는 “진정한(true)” 범주화 작업이 어렵기 때문에 아주 어린 나이에 관계추론이 거의 불가능하다고 주장하였다(Inhelder & Piaget, 1964). 그러나 반세기도 지나지 않아 아이들의 추리 능력에 대한 비관적 견해는 뒤집혔다. 최근 연구에 따르면 관계추론능력이

아동기까지 미쳐 형성되지 않는 것처럼 보이는 이유는 관계추론에 대한 정의와 개념화가 잘못된 채 그에 기반한 측정도구를 활용하였기 때문이다(Buehl & Alexander, 2004; Goswami, 1989, 2013). 대개 기존 연구에서 성인을 위해 고안된 관계추론 측정도구(예: 고전 관계추론 검사의 경우 $a : b :: c : d$ 와 같은 유사성 추론을 측정함)를 아동에게 적용해 왔기 때문에 어린 아이들의 관계추론능력이 잘 포착되지 않았다고 할 수 있다. 이러한 문제를 극복하기 위해 아동에게 익숙한 이야기를 사용하여 관계추론능력검사를 하는 대안적 방식이 제안되었다(Chen, 1996; Goswami, 2013). 아동에게 적절한 도구를 사용하고 아동의 일상생활과 관련된 활동을 활용하여 관계추론능력을 측정한 결과, 빠르면 3세에도 아동의 관계추론능력이 포착될 수 있음을 발견하였다(Goswami, 2013). 학령기 학생을 대상으로 한 또 다른 연구에서는 5~6세 아동 중 절반 가량이 삼단논법을 이해하는 능력을 습득했음을 밝혔다(Andrews, 1996). 즉, 초등학교 초기(7세)의 학습자는 일반적으로 세 가지 이상의 요인과 관련된 매우 복잡한 추론을 할 준비가 이미 되어 있다. 뿐만 아니라 아이들에게 적절한 개입을 제공함으로써 추론 능력을 발달시킬 수 있음도 밝혀졌다(Alexander, 1984; Alexander, White, Haensly, & Crimmins-Jeanes, 1987).

관계추론의 발달 패턴에 관한 연구들은 대부분 단순히 선형적인 성장패턴이 아닌 곡선형 성장패턴을 보인다고 주장하고 있다. 그러나 어느 지점에서 어떠한 곡선적 성장을 보이는지에 관한 선행연구는 두 갈래로 나눌

수 있다. 그 중 지금껏 대세를 이룬 주장은 1) 사춘기정도에 발달의 “일시적 감소”을 보인다는 것이다(Dumontheil et al., 2010). 그러나 보다 최근에 2) 청소년기까지 꾸준한 성장패턴을 보이다가 사춘기 이후 성장세가 둔화된다는 주장이 힘을 얻고 있다(Jablansky et al., 2016).

사춘기경 추론 능력 발달의 “일시적 감소”를 주장한 대표적인 연구로 Dumontheil 등 (2010)의 연구가 있다. 이 연구에서는 청소년에서 성인을 대상으로 관계추론능력과 관련된 발달 패턴을 보기 위하여 관계추론 활동을 실시하면서 연구대상자의 두뇌를 영상 촬영하였다. 관계추론검사 결과와 fMRI 촬영 결과를 비교하였는데 14세 이하 청소년시기 학생들에 비해 15세 학생들의 관계추론 수행 검사 점수가 낮았다. Dumontheil 등은 이러한 현상을 대략 15세 경에 나타나는 뇌의 구조적 형태 변화와 연관 지어 설명하였다. 15세 경 전전두엽피질(rostromedial prefrontal cortex, RLPFC)의 구조적 변화가 급격히 진행되면서 관계추론 수행능력이 일시적으로 감소하는 현상이 발생하였다는 것이다. 즉 15세경에 관계추론 인지 능력과 관련된 두뇌 영역의 구조적 변화에 따라 관계추론능력은 비선형적 발달 패턴을 보인다. 그 밖에 많은 연구들이 인간의 인지적 수행 능력의 발달 패턴은 15세경 또는 11세에서 12세 사이에 변곡점을 보이며 곡선적(curvilinear)이라고 보고하였다(Carey, Diamond, & Woods, 1980; Diamond, Carey, & Back, 1983; McGivern, Andersen, Byrd, Mutter, & Reilly, 2002).

반면 Jablansky와 동료들(2016)은 관계추론

발달에 대해 사춘기 이후 성장세가 둔화되지만 성장은 지속된다고 주장하였다. 이들은 관계추론능력의 발달 패턴을 확인하기 위해 5세에서 17세 아동과 청소년을 대상으로 연구를 진행하였는데 이때 연령에 따라 유치원~2학년, 4학년~8학년, 10~11학년 등 크게 세 개 연령 그룹 연구 대상을 묶었다. 각 그룹 청소년들에게 평소에 친밀하거나(예: 주스 박스) 혹은 친밀하지 않은 대상물(예: 야채커터기)을 주고 이 대상물의 사용에 관해 토론을 유도하였다. 이 토론 중에 학생들이 사용한 언어를 기록하고 의미단위(unit)로 분석하였는데 이들이 사용한 관계추론과 관련된 총 의미단위(unit) 수는 11~14세 그룹에서 가장 높은 빈도를 보였다. 즉 현상적으로 14세까지 관계추론 관련 토론 빈도가 높아지다가 이후 다시 꺾이는 모양새를 보인 것이다. 그러나 관계추론의 하위 구인별로 토론 내용을 다시 분석한 결과, 흥미롭게도 관계추론 전체 점수 패턴과 다른 양상을 보였다. 즉 15세 이후 학생들은 자신의 기초 지식과 관련된 단순 유사성을 찾는 식의 추론보다는 상이성을 찾아내는 추론에 더 많은 시간을 보내는 것으로 나타났다. 이를 토대로 Jablansky와 동료들은 관계추론능력이 사춘기 이후 성장세가 다소 둔화되는 것으로 보이는 현상은 관계추론능력 하위 구성 요인을 중심으로 본다면 단순히 줄어든다고 볼 수 없다고 결론지었다. 즉 15세 이후 청소년들은 상이성에 입각한 추론활동에 힘을 쏟느라 단순한 유사성을 찾는 추론활동에 덜 집중하는 경향을 보인다는 것이다.

이처럼 관계추론능력의 발달에 관한 다소

상이한 주장은 발달의 형태가 15세경 일시적 감소로 볼 것인지, 15세 이후 둔화된 성장세가 지속된다고 볼지 두 가지 발달학적 쟁점을 남겼다. 이러한 상충점이 발생하는 원인으로 연구 대상, 연구 방법 등이 문제가 될 수 있을 것이다. 그러나 본 연구에서는 기존 관계추론 연구들이 기반으로 삼은 관계추론능력의 개념화, 구성요인 변별 문제가 이러한 상충된 의견을 이해할 수 있는 열쇠라고 보았다. 즉 관계추론능력이 언제 어떤 양상으로 발달하는지 명확히 보여 주기 위해 핵심 구인에 대한 명확한 학술적 정의와 분석을 바탕으로 관계추론의 연령 간 수행 차이를 제시할 필요가 있어 보인다.

관계추론능력과 관계추론능력의 하위 구인에 대해 보다 세밀한 학술적 정의를 수립하기 위해 최근에 여러 연구자들의 노력이 있었다(Alexander, Dumas, Grossnickle, List, & Firetto, 2016; Dumas & Alexander, 2016, 2018; Dumas, Alexander, & Grossnickle, 2013; Richland & Burchinal, 2013; Richland et al., 2006; Richland, Zur, & Holyoak, 2007). 이러한 노력의 결과, 관계추론능력척도(Test of Relational Reasoning: TORR)가 개발되었다(Alexander et al., 2016; Murphy, Firetto, & Greene, 2017). Alexander와 동료 연구진(2016)이 개발한 초기 관계추론능력척도는 성인을 대상으로 하였으며 관계추론의 하위 구인으로 유사(analogy), 변칙(anomaly), 모순(antinomy), 대조(antithesis) 등 네 가지 척도를 포함하였다. 이후 8~15세 아동을 위한 아동용 TORR(TORR junior: TORRjr.)가 제작되어 지금까지 사용되고 있다(Zhao & Sun, 2019). TORR와 TORRjr.는 개발 이후 대부분 서구 문

화권에서 연구가 진행되어 오다가 최근에 중국 학생을 대상으로 한 연구가 진행되었다(Zhao & Sun, 2019). 이 연구에서는 초등학교 3, 4, 5학년과 중학교 1, 2학년 학생들에게 TORRjr.(아동용 관계추론능력 검사)를 실시하였는데, 초등학교 3학년 학생들의 유사, 변칙, 모순 유형의 관계추론능력 검사 점수가 가장 낮고 그 후 증가함을 발견하였다. 반면에 대조 척도 점수는 초등학생과 중학생 그룹 간에 차이가 뚜렷하였다. Zhao의 연구에서 초등학교 3학년은 대략 9세이고 중학교 2학년은 13세에 해당한다. 따라서 기존의 선행 연구에서 지적했던 15세 전후 변곡점을 보인다는 주장(Dumontheil et al., 2010)은 Zhao의 연구와 상충됨을 알 수 있다. 오히려 보다 넓은 인지능력 연구 맥락에서 12~13세 경 변곡점을 주장한 McGivern 등의 연구(2002)가 Zhao와 Sun의 연구결과(2019)에 근접한 것으로 나타났다. 그러나 관계추론능력의 전체적 발달 패턴은 하위 유형과 상관없이 전체적으로 높아지는 경향을 보여 기존 McGivern 등(2002)이나 Dumontheil 등(2010)의 뇌 연구와는 괴리가 있어 보인다. 완전히 동일하지는 않으나 꾸준한 성장세를 보여주었다는 점에서 Jablansky 등의 연구(2016)에 근접한 결과를 보여주었다.

위에 살펴본 선행 연구 결과를 토대로 본 연구에서 설정한 연구 가설은 크게 두 가지이다. 먼저 본 연구에서는 Jablansky 등(2016)을 필두로 하여 Alexander 연구 그룹(Alexander et al., 2016)과 Richland 등(2007)이 주장한대로, 관계추론능력이 네 개의 하위 구인(유사, 변칙, 모순, 대조)으로 구성된 이중 모형으로 정의될 수 있다고 보았다. 관계추론능력의 이중요인

구조라는 주장은 이미 Spearman을 위시한 고전 인지심리학자들의 일반 지능이론과 맥락을 같이 한다. 최근에는 Dumas와 Alexander(2016)와 Zhao와 Sun(2019)의 연구에서 밝혀진 바 있다. 예컨대, 이들은 이중요인모형, 하위 요인 간 상관만을 허용한 4요인 모형, 관계추론 일반 요인모형 중 어떤 모형이 TORR 수행 데이터에 가장 적합한지 비교하였다. 그 결과 이중요인모형이 가장 설명력이 높은 모형임을 입증하였다. 그러나 이들의 연구는 중학생까지만을 대상으로 하거나 대학생 단일 연령대에 대해서만 요인분석을 실시하여 초등학교부터 고등학교를 망라하는 광범위한 표집에서 적절한 타당성이 있는지에 대한 연구는 진행된 바 없다.

관계추론의 이중 요인 구조를 기초로 한다는 점에서 본 연구는 Jablansky 등의 연구(2016)나 Zhao와 Sun의 발달 연구(2019)와 궤를 같이 한다. 즉 관계추론능력은 15세 이후까지 점진적 발달을 지속할 것이라고 가정하였다. 다만 14세 경 나타나는 두뇌구조의 변화로 말미암아 다소 성장세가 둔화되는 패턴을 보일 것이라 가정하였다. 또한 하위 요인 별로 발달 패턴을 다를 것이라 가정하였다. 이는 Dumontheil 등의 연구(2010)나 기존 신경상관 연구에서 흔히 활용되어 온 레이븐 측정 도구와 같이 유사성에만 집중한 기존 연구와 구별된다.

본 연구는 또한 차별기능분항 분석을 토대로 연구 결과를 정리하고자 하였다는 점에서 기존연구들과도 구분된다. 문항반응 이론(Item Response Theory: IRT), 차별기능분항(Differential Item Functioning: DIF) 분석은 일반 문항분석에

비해 피험자의 실제 잠재능력점수에 가까운 값을 구할 수 있다는 장점이 있다. 따라서 본 연구와 같이 연구 여건상 하나의 그룹을 추적 조사하지 않고 각기 다른 연령 그룹에서 동시에 연구대상을 모집하여 발달 연구를 할 경우 이들 간 차이 검증을 통해 발달 추이를 검토할 수 있다(Wilson, Zheng, & McGuire, 2012). 또한 문항 수준에서 값을 얻기 때문에 전체 점수를 이용함으로써 피험자 집단 간에 정확히 어떤 문항에서 어떻게 차이를 보이는지에 대해 명확한 결과를 얻기 위해 유리한 방법이다. 본 연구에서는 특히 각기 다른 발달단계에서 추출된 여러 집단 간 차이를 구분하기 위하여 Multigroup IRT-DIF를 활용하였다.

유사성 유추를 중심으로 한 선행 연구(Dumontheil et al., 2010; McGivern et al., 2002)에서 특정 연령 이후 보인 발달의 변곡점은 관계추론능력의 기본 구성요인, 내재된 요인 구조에 대한 가설의 왜곡 때문이 아닐지 의심하게 된다. 이에 본 연구에서는 유사, 변칙, 모순, 대조의 네 개 하위 요인을 포함하는 이중요인구조를 바탕으로 한 관계추론능력 검사를 활용하여 아동 청소년의 발달 양상을 탐색해 보고자 하였다.

2. 연구문제

연구문제 1. 한국 학령기 학생들의 관계추론능력은 네 가지 형태(유사, 변칙, 모순, 대조)와 학년에 따라 다른 수행을 보이는가?

연구문제 2. 관계추론능력 검사 한국어 버전에 나타난 관계추론능력에서 학년 간 차별기능분항은 무엇이며 어떠한 특성을 갖는가?

II. 연구방법

1. 연구대상

한국어 버전 관계추론능력검사를 적용하기 위하여 강원도 소재의 초등학교, 중학교, 고등학교 각 1개교에 재학 중인 초4(196명), 초6(200명), 중2(206명), 고1(146명) 학생을 표집하였다. 성별 구성 비율은 초4 남학생 109명(55.33%), 초6 남학생 105명(52.50%), 중2 남학생 81명(39.32%), 고1 남학생 63명(43.15%)으로 나타났다. 이들의 평균 연령은 초4 10세, 초6 12세, 중2 14세, 고1 16세로서 모두 한국에서 정상적인 학교교육을 받아 온 학생들이다. 연구대상은 모두 같은 지역의 학교에 재학 중이었기 때문에 지역에 따른 편차는 없을 것으로 예상되었다. 또한 연구대상이 속한 고등학교는 평준화 지역에 속해 있어 보편적인 한국 고등학생의 인지 능력에 가깝다고 판단할 수 있다. 본 연구표집은 코로나19 발발 이전 학생들의 학교 등교 중에 학교를 통해 실시되었으며 학생들은 출석일에 지필식 검사에 응하였다.

2. 측정도구

아동용 관계추론능력 검사

아동용 관계추론능력 검사(TORRjr)를 한국 학생들에게 적용하기 위해 한국말 버전을 작성하였다. TORRjr는 유사(analogy), 변칙(anomaly), 모순(antinomy), 대조(antithesis) 등 관계추론 네 개 유형별로 네 세트로 구성되어 있으며 각 세트별로 해당 관계추론능력을 측

정한다. 이를 위해 그림으로 제시된 문항에서 네 개의 응답지 중 정답을 찾는 방식으로 진행된다. 각 세트에는 2개의 예시 문항과 8개 문항이 포함되어 있어 피험자는 총 8개의 연습용 예시 문항과 32개의 검사문항을 접하도록 설계되었다. 학생들은 예시문항을 통해 문항을 푸는 방법을 습득한 후 이어지는 8개 문항에 실제로 응답하였다. 그림 1에 각 유형별 예시문항을 하나씩 제시하였다. 보는 바와 같이 유사성(analogy) 유추를 위한 문항은 기존 레이븐 검사(Raven, 2003)의 비언어성 유추문항과 유사한 형태를 띠고 있다. 그림에 제시된 바와 같이 여러 가지 도형과 도형 안에 찍힌 점의 숫자를 보고 이 행렬의 패턴 규칙에 따라 물음표 표시된 빈칸에 들어가야 할 그림은 무엇인지 찾아야 한다. 그림1-a의 예에서 세로 행렬의 규칙은 삼각형이고 밑으로 갈수록 점의 수가 하나씩 늘고 있으므로 정답은 A이다.

변칙(anomaly)은 주어진 예시 그림 집합의 규칙을 따르지 않는 상황을 선택하는 문항이다. 보기에서 A, B, C 모두 직각삼각형을 나타내지만 D만 평행사변형을 나타내기 때문에 다른 도형과는 다른, 변칙에 해당한다.

모순(antinomy) 문항은 두 개의 세트에 속하는 규칙을 알아차리고 한쪽의 규칙에는 맞고 다른 한쪽에는 위배되는 모순적 상황을 찾는 문항이다. 예시1의 세트1은 모두 검은색 동그라미라는 규칙, 세트2는 모두 진한 테두리로 된 다양한 도형의 집합이라는 규칙을 따른다. 따라서 세트1의 규칙에만 따르고 세트2의 규칙에는 따르지 않는 상호 모순적 상황 A가 답이 될 수 있다.

대조(antithesis)는 주어진 보기의 규칙과 정반대의 절차를 따르는 선택지를 고르는 문항이다. 그림1-d에 제시된 보기는 직선이 구불구불한 선으로 바뀌는 과정을 보여준다. 이 과정과 정반대의 과정은 구불구불한 선이 직선으로 퍼지는 것이다. 따라서 구불구불한 선으로 된 삼각형이 직선으로 된 삼각형으로 바뀐 과정을 보여준 A가 정답이다.

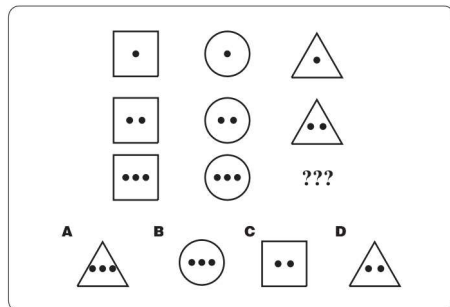
이 관계추론능력검사를 한국어 버전으로 제작하는데 두 명의 연구자가 참여하였다. 연구자1은 교육심리학 전공자로서 해당 분야의

전문성을 보유하고 이 검사지 활용에 능통한 국내 대학 교수이다. 연구자2는 언어심리학 전공자로서 국내 모 대학의 영문과에서 신경심리 및 언어학을 담당하고 있는 교수이다. 두 연구자 모두 한국어와 영어 사용에 능통한 심리학자로서 각자 관계추론능력 검사의 영문판을 읽고 번역하였다. 독립적으로 번역한 전문은 상호 비교 후 상이한 내용에 대해 합의에 이를 때까지 반복적으로 번역과 수정 작업을 거쳤다.

a. 유사성(Analogy)

지침: 아래에 보이는 것은 미완성된 패턴입니다. 이 패턴을 완성시키는 모양을 아래에서 고르세요.

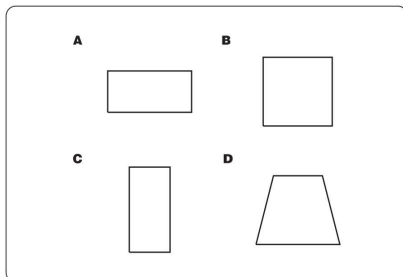
예시 1



b. 변칙(anomaly)

지침: 아래 모양에 해당하지 않는 것을 고르세요.

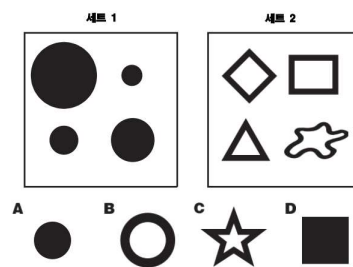
예시 1.



c. 모순(antinomy)

지침: 세트 1에 **있** 해당하는 되고 세트 2에는 해당이 되지 **않**는 모양을 아래에서 고르세요.

예시 1



d. 대조(antithesis)

지침: <보기>와 정반대 과정을 아래에서 고르세요.

예시 1.

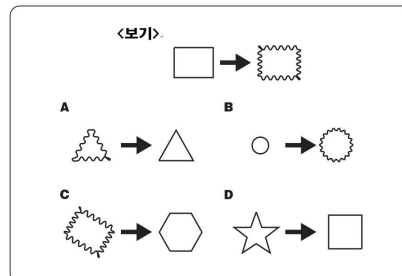


그림 1. 한국판 관계추론능력검사 (아동용) 예시 문항

3. 연구절차

검사에 앞서 2019년 4월 기관생명윤리위원회(IRB)에 연구 참여자용 설명서와 동의서 심의를 신청하여 승인(GWNUIRB-2019-7) 후 연구를 실시하였다. 학생들은 연구동의서에 동의여부를 체크한 후 인쇄물로 제작된 TORRjr 한국어 버전에 응답하였다. TORR와 TORRjr는 당초 온라인 버전과 오프라인 버전 두가지로 제작되었으나 여러 가지 여건상 온라인 버전 검사는 실시하기 어려워 모든 피험자들에게 오프라인 버전 검사(지필 검사)를 실시하였다. 피험자 모두 유사, 변칙, 모순, 대조의 순으로 두 개의 연습문항과 여덟 개의 검사 문항을 풀었다.

4. 자료분석

관계추론능력검사의 네 개 유형에 따른 학

년별 수행 점수 패턴을 분석하기 위하여 반복 측정분산분석(repeated measure ANOVA)을 실시하였다. 이때 네 개 관계추론 유형은 집단 내 요인이며 학년은 집단 간 요인으로 간주되었다. 이 과정은 SPSS 21.0을 이용하여 수행되었다.

다음으로, 학년과 관계추론 유형별 발달 패턴을 파악하기 위해 MIRT-DIF 분석을 실시하였다. 차별문항 검사는 기존 TORRjr와 TORR 관련문헌(Dumas & Alexander, 2016; Zhao & Sun, 2019)에서 제시한 이중요인(bi-factor) 모형을 기반으로 하였다. 해당 문헌에서 제시한 이중요인(bi-factor) 모형을 도식화 하면 그림 2와 같다. 학생 반응의 학년 간 차이가 검사 문항의 어디에서 발생하는지 보다 명확히 알아보기 위하여 세 개 근접 학년 쌍(4학년과 6학년; 6학년과 중2; 중2와 고1)으로 이루어진 세 개의 데이터를 만들어 다집단 문항반응이론-차별문항기능 검사(Multigroup IRT-Differential Item Functioning: MIRT-DIF)를 실시하였다. 그

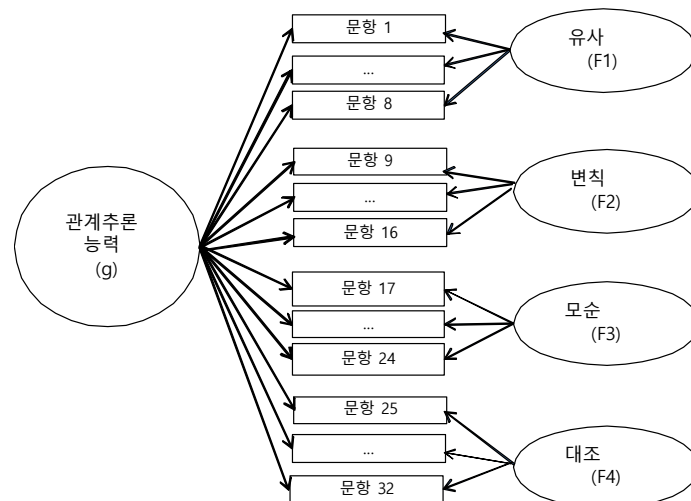


그림 2. 관계추론능력 이중요인 모형

표 1. 관계추론능력검사 유형 및 학년에 대한 반복측정분산분석 결과

	소스	제 III 유형 제곱합	자유도	평균제곱	F	유의확률	부분 에타 제곱
집단내	유형	166055.96	2.00	83113.39	6941.36	0.00	0.90
	유형*학년	1323.97	5.99	220.89	18.45	0.00	0.07
집단간	학년	1987.69	3.00	662.56	24.77	0.00	0.09

후, 학년 모두를 포함한 전체 데이터를 활용하여 위의 두 학년 쌍 별 MIRT-DIF 분석에서 나타난 차별기능 문항을 4개 학년 급 모두에 허용하는 최종 모형 기반 MIRT-DIF 분석을 실시하였다. 이 최종 모형(그룹 간 DIF를 허용한 경우)은 영가설모형(그룹 간 DIF를 허용하지 않는 경우)과 비교하여 모형 적합성을 평가하였다. 또한 이 과정을 통해 학년 및 문항 특성과 상관없이 존재하는 고유한(Kranzler, Benson, & Floyd, 2015) 잠재 변수에 대한 요인값을 추출하였다. 최종 모형에서 얻은 요인값을 활용하여 학년과 관계추론능력의 하위 유형별 차이를 알아보았다. 문항 수준 분석을 위해 R(R Core Team, 2016)의 MIRT-DIF 패러다임과 psych(Revelle, 2019), mirt(Chalmers, 2012), psy(Falissard, 2012) 등 공개되어 있는 R 패키지를 사용하였다.

III. 결과 및 해석

1. 관계추론 유형과 학년 간 차이 분석

그림 3과 같이 전체 TORR_{jr} 점수 평균은 저학년 학생보다 고학년 학생이 높은 점수를 보였다. 하위 영역별로도 고학년으로 갈수록 높은 점수를 보였다. 그러나 각 영역별로 학년 간 평균 차이 패턴은 다소 상이 하였다. 유사와 변칙 척도는 초4와 초6 사이에 유의한 변화를 보였고 그 이후 학년에는 학년 간 차이가 적었다. 즉 초6, 중2, 고1의 유사성 점수는 비슷하였다. 유사와 변칙 척도의 학년 간 차이와는 다르게 모순과 대조 척도 점수는 초6과 중2 사이에 뚜렷한 차이를 보였다.

관계추론능력의 유형과 학년에 따른 편차를 보다 명확히 파악하기 위하여 반복측정분

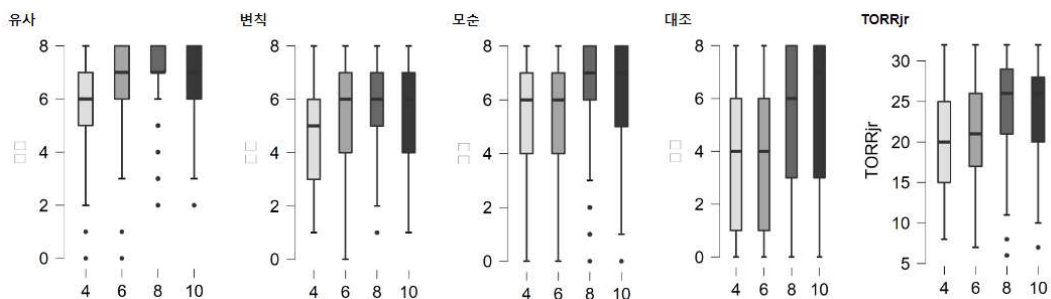


그림 3. 유형 및 학년에 따른 관계추론능력검사 결과 박스플롯(평균 및 표준편차)

산분석을 실시하였다. 먼저 분석을 위한 조건을 만족시키는지 검사한 결과, 데이터의 구형성을 만족하지 않았다($p < .05$). 또한 오차분산의 동질성은 유사성과 모순 유형에 대해서 만족시키지 않았으나($p < .05$) 변칙과 대조 유형은 오차분산 동질성을 보이는 것으로 나타났다($p > .05$). 이러한 검증결과를 토대로 이후 분석에서는 구형성과 오차분산 동질성 가정에서 자유로운 Greenhouse-Geisser 분석법을 적용하였다. 표 1에서 보는 바와 같이 집단 내 효과(관계추론 유형) $F(2) = 6941.36, p < .001$ 와 집단 외 효과(학년) $F(3) = 24.77, p < .001$ 에 따라 관계추론능력의 점수 차이가 유의하였다. 또한 관계추론 유형과 학년 간 상호작용효과도 유의하였다($F(5.99) = 18.45, p < .001$). 다시 말해 학년이 올라가면서 보이는 검사 점수의 차이가 뚜렷했으며 이러한 차이 패턴은 관계추론 유형별로 달랐다. 어느 학년 간에 그 차이가 뚜렷해지는지 Tukey 사후 검사를 통해 살펴 본 결과, 전체 관계추론 능력검사의 점수(TORRjr)는 초6과 중2 사이에 점수 차이가 커서 초4와 초6이 비슷한 그룹으로, 중2와 고1이 유사한 다른 그룹으로 묶이는 것을 알 수 있었다. 하위 유형별로 살펴 본 결과, 유사성 유추 점수는 학년별로 큰 차이가 없었다. 그러나 변칙은 다른 학년에 비해 중2의 점수가 눈에 띄게 높았다. 모순과 대조 척도의 점수는 초6과 중2 간의 차이가 커서 초4초6 대 중2-고1 그룹으로 양분됨을 알 수 있었다. 다소 어려운 유형인 모순과 대조 추론 능력 점수가 초등학교생들에 비해 중2와 고1 학생들에게서 높게 나타남을 확인하였다.

2. 차별문항검사 결과

문항 수준의 발달 패턴을 알아보기 위하여 다집단 문항반응이론에 기초한 차별문항검사(MIRT-DIF)를 실시하였다. 그 결과 세 개 문항에서 초4-초6 학년 간 차이를, 네 개 문항에서 초6-중2 학년 간의 차이를 발견하였다. 이 가운데 한 개의 문항은 초4-초6과 초6-중2 학년 쌍에서 모두 차별문항인 것으로 나타났다(표 2 참조).

앞서 척도수준분석(분산분석)에서 본 것과 마찬가지로 문항수준분석(차별문항 검사)에서도 유사와 변칙문항은 초4와 초6 사이에, 모순과 대조는 초6과 중2 사이에 눈에 띄는 차이를 보였다. 초4-초6 간 DIF를 보인 세 개의 문항은 3번과 4번 유사문항과 2번 변칙문항이었다(표 3 참조). 즉 초4-초6 간에는 유사와 변칙문항의 차이가 뚜렷하며(표 3 참조), 이러한 차이는 그림 3에서 척도수준의 유사, 변칙 점수 평균값 차이에서도 포착된 바와 비슷하다.

초6-중2 간에 DIF를 보인 두 개의 문항은 5번 모순문항과 8번 대조문항이었다. 위의 다섯 개 DIF문항 외에 두 개의 변칙문항(1번과 2번 변칙문항)이 초6과 중2 간에 유의한 DIF를 보이는 것으로 나타났다. 즉 초6-중2 간에는 모순과 대조 문항의 차이가 뚜렷하였으며(표 3 참조), 이러한 차이 역시 그림 3의 척도 수준 평균값 차이에서도 포착된 바와 같다. 반면에 중2와 고1 간에는 DIF가 포착되지 않아 이 두 집단은 문항에 대해 유사하게 반응하는 것을 알 수 있었다.

요약하면 차별문항검사 결과는 반복측정분

표 2. 학년 쌍 별 차별기능 문항 - 영가설 모형과 학년 간 차이 허용 모형 비교

비교집단	모형	문항번호	AIC	AICc	SABIC	HQ	BIC	logLik	$\Delta\chi^2$	Δdf	p-value
초4초6	영가설		13584.03	13662.25	13669.99	13751.32	14006.33	-6686.02			
초4초6	차이 허용	유사3	13580.49	13664.05	13668.88	13752.51	14014.74	-6681.25	9.54	3	0.023
초4초6	차이 허용	유사4	13581.92	13665.47	13670.31	13753.94	14016.17	-6681.96	8.12	3	0.044
초4초6	차이 허용	변칙2	13547.13	13621.93	13631.47	13711.26	13961.46	-6669.57	46.73	3	0.000
초6중2	영가설		12346.68	12422.54	12435.00	12514.75	12771.35	-6067.34			
초6중2	차이 허용	변칙1	12338.14	12419.16	12428.96	12510.98	12774.83	-6060.07	14.54	3	0.002
초6중2	차이 허용	변칙2	12340.20	12421.21	12431.02	12513.03	12776.89	-6061.10	12.48	3	0.006
초6중2	차이 허용	모순5	12340.39	12421.41	12431.21	12513.23	12777.09	-6061.20	12.28	3	0.006
초6중2	차이 허용	대조8	12342.92	12423.93	12433.74	12515.75	12779.61	-6062.46	9.76	3	0.021

주. AIC = Akaike Information Criterion; AICc = sample size adjusted AIC; SABIC = Sample-Size Adjusted BIC; HQ = Hannan-Quinn; BIC = Bayesian Information Criterion

산분석 결과와 유사하게 초4-초6과 중2-고1 그룹으로 나뉘었다. 유사나 변칙과 같은 다소 단순한 척도는 어린 나이인 초4와 초6 사이에 뚜렷한 차이를 보인 반면 모순과 대조는 보다 늦은 나이인 초6와 중2 간에 뚜렷한 차이를 보였다.

위의 DIF 문항들의 변별력과 난이도는 학

년 간 차이를 보였으며(표 3 참조), 이 역시 그림 3에서 이미 본 것과 비슷하다. 다만 마지막 문항인 대조 8번 문항은 관계추론 일반 요인(g)에 대해 변별력이 높아지고(초4 = 0.74, 초6 = 1.02, 중2 = 1.05, 고1 = 1.53), 난이도는 점점 낮아지는(초4 = 0.89, 초6 = 0.97, 중2 = 1.4, 고1 = 1.58) 특징을 보였다.

표 3. 다집단 문항반응 이론 차별기능문항 지수(문항 변별도와 난이도)

문항번호	초4			초6			중2			고1		
	a	a total	d	a	a total	d	a	a total	d	a	a total	d
1. 유사 3	0.82	-0	0.29	0.14	0.35	1.11	0.49	0.3	1	0.54	0.23	0.63
4. 유사 4	0.61	0.6	2.02	0.84	1.21	3.94	1.04	0.8	3.36	-0.4	0.39	2.74
9. 변칙 1	0.79	0.33	-0.10	0.84	0.47	0.06	1.01	0.42	-0.80	0.46	0.03	-0.60
10. 모순 2	0.97	0.51	-0.6	0.84	0.12	0.73	0.85	0.18	0.02	0.69	-0.1	-0.20
21. 모순 5	0.90	0.88	1.14	0.77	0.91	1.36	-0.20	0.27	0.83	0.89	0.57	0.97
32. 대조 8	1.53	0.74	0.89	1.56	1.02	0.97	1.03	1.05	1.4	1.48	1.53	1.58

주. a = 각 유형에 대한 문항변별도; a total = 전체 관계추론능력(일반요인)에 대한 문항 변별도; d = 난이도 (쉬운 정도)

난이도 추세를 보았을 때 대조 8번 문항에 대한 고1 학생의 정답 비율이 상대적으로 높고 중2, 초6, 초4 학년 순으로 높다는 것을 알 수 있다. 대조 8번 문항은 애초에 관계추론능력 검사에서도 최상위 수준에 속하는 문항으로 설계되었기 때문에 이러한 추세는 주목할 필요가 있다. 즉 상위의 관계추론 문항이 고학년 학생들의 관계추론능력을 보다 잘 구별하는 것으로 보인다. 그러나 대조 8번 외의 다른 대조 문항들은 차별기능문항으로 포착되지 않았으며 이 같은 추세는 다른 관계추론 유형(유사, 변칙, 모순)에서도 비슷하게 유지되었다. 다시 말해 관계추론능력 검사 마지막 문항을 제외한 다른 문항들은 모든 학년에게 비슷한 수준으로 기능하고 비슷한 수준의 난이도를 제공하는 것으로 보였다.

IV. 논의 및 결론

특정 연령이 지나면 관계추론능력이 감소한다는 주장(Carey et al., 1980; Diamond et al., 1983; Dumontheil et al., 2010; Jablansky et al., 2016; McGivern et al., 2002)은 본 연구를 통해 확인하기 어려웠다. 본 연구에 따르면 학령이 높은 학생들일수록 전반적인 관계추론 수행력이 높은 것으로 나타났다. 그러나 이러한 발달은 초등학교 시절을 지나 중고등학교로 올라가면서 완만하게 진행되는 것으로 보인다. 전체 관계추론능력 검사 점수의 평균값은 4학년과 6학년 사이에 약간 차이가 있었고 6학년과 중2 사이에서는 유의하게 달랐다. 그러나 이러한 성장세는 중2 이후에 뚜렷하게

줄었다.

선행연구 결과가 본 연구의 결과와 이렇게 상이한 이유는 본 연구의 기초가 된 관계추론능력의 요인구조 때문인 것으로 보인다. 유사성 유추에만 기반한 기존 선행 연구와 달리 본 연구에서는 관계추론능력을 네 개의 하위 요인으로 구성된 일반 요인으로 보았다. 청소년의 관계추론능력을 연령에 따라 추적한 Dumontheil 등의 연구(2010)에서 참가자들은 주어진 그림을 보고 유사점을 식별하는 과제를 수행하였다. 본 연구에서는 관계추론의 네 개 유형(유사, 변칙, 모순, 대조)을 측정하는 32개 구조화된 문항에 응답하였다. 그 결과 Dumontheil의 참가자들은 15세 경 점수가 일시적으로 줄었지만 본 연구에서는 고학년으로 갈수록 점수가 높아지다가 중2 이후 점수가 비슷한 수준을 보였다.

중학생과 고등학생들이 보인 관계추론능력 성장세의 정체기는 Jablansky 등의 연구(2016)와 Zhao와 Sun의 연구결과(2019)와 비슷하다. 이러한 현상은 이 두 개의 기존 연구가 본 연구와 마찬가지로 관계추론능력을 네 개 하위 구인으로 구성된 일반 요인으로 보았다는 바탕 이론이 일치하기 때문에 당연한 결과일 것이다. 관계추론에 유사성 유추만을 포함시켰을 경우 학령 간 차이를 곡해할 수 있음을 확인하였다.

네 개 관계추론능력 하위 영역 중 모순과 대조 척도 점수는 유사, 변칙보다 고학년 사이에 뚜렷한 차이를 보이는 것으로 나타났다. 반복측정분산분석 결과와 MIRT-DIF 분석 결과, 유사와 변칙과 같은 다소 단순한 관계추론수행 능력은 초4-초6 학년에서 뚜렷한 차이

가 포착되었다. 따라서 이 두 개 영역 점수만을 관계추론능력의 하위 영역으로 본다면 관계추론능력은 초기 생애에서 현저하게 발달하는 것처럼 보일 수 있다. 반면 모순, 대조 척도와 같은 보다 난해한 추론 능력에 대한 수행검사 결과는 초6과 중2 사이에 더 현격했다.

유사와 변칙 수행에서 다소 큰 차이를 보이는 초4에서 초6 학년은 만 10세에서 12세로서 Dumontheil과 같은 연구자가 보여준 관계추론능력이 감소하는 시기 혹은 Jablansky가 보여준 관계추론능력 발달의 변곡점으로 보이는 약 14세보다는 어렸다. 또한 본 연구에서 확인한 바, 중2나 고1 학생은 초6학년 학생에 비해 관계추론 수행 점수는 크게 다르지는 않지만 약간 높은 경향을 보였는데 이러한 결과는 사춘기 초기에 관계추론수행능력이 “일시적 감소(dip)”한다고 주장한 기존의 연구(Dumontheil et al., 2010; McGivern et al., 2002)와는 다소 차이가 있다. 그러나 관계추론능력에 기여하는 두뇌 변화가 14세 전후(Dumontheil et al., 2010) 또는 11세에서 12세 사이(McGivern et al., 2002)에 뚜렷하게 나타난다는 신경학적 연구와 주장에 대해서는 어느 정도 지지할 수 있는 경험 자료가 될 수 있을 것이다.

그러나 모순과 대조 추론은 초6과 중2 사이에서 뚜렷한 수행능력 차이를 보였다. 관계추론능력을 기존 연구에서 주로 활용한 유사성 유추를 중심으로 개념화할 경우, 초4 경을 성장의 변곡점으로 볼 수 있겠지만 보다 포괄적으로 모순과 대조까지 포함시킨다면 초6(12세)과 중2(14세) 간 차이가 뚜렷하다고 할 수

있다. 즉 복잡하고 어려운 작업에 해당하는 모순과 대조 문항을 해결하는 데 필요한 관계추론능력은 다소 늦은 시기, 즉 14세 이후에 발달하는 것으로 보인다. 초4와 초6 학년 학생들의 모순과 대조 관계추론 수행 능력 점수는 중고등학생의 점수보다 낮은 편이었다. 결론적으로 관계추론능력 발달의 변곡점으로 지적되었던 12~13세 혹은 14세는 관계추론의 어느 하위 영역을 중심으로 관찰하느냐에 따라 달라질 수 있을 것이다.

관계추론능력 네 개 하위 구인별 학년 간 수행 능력 차이는 과거 TORR연구들과 약간 상이한 측면이 있다. Zhao와 Sun의 연구(2019)에서 3학년에서 6학년(한국 나이 대략 만 12세)까지 관계추론 점수가 높아지다가 7학년(13세)에 미세하게 낮아지는 경향(3학년 = 15.78, 4학년 = 16.99, 5학년 = 18.35, 6학년 = 21.63, 7학년 = 20.84)을 보였는데 이는 본 연구의 전체 TORR 검사 점수의 학년 간 차이 패턴(초4 = 19.89, 초6 = 21.29, 중2 = 24.33, 고1 = 23.7)과 비슷했다. 구체적인 점수를 통해 확인하였듯 본 연구에 참여한 한국 학생들의 점수가 Zhao와 Sun(2019)의 중국 학생들에 비해 약간 더 높은 경향을 보였으나 전체적인 패턴은 비슷하다. 그러나 하위 구인 점수는 본 연구의 한국 학생들의 모순과 대조 추론의 학년 간 수행능력 차이가 뚜렷한 것에 비해 Zhao와 Sun의 연구(2019)에서는 유사, 변칙, 모순, 대조 모든 영역의 추론 점수가 비슷하게 고학년으로 갈수록 높아지는 양상을 보였다.

본 연구에서 흥미롭게 포착된 초6-중2 간 모순과 대조 추론 능력의 현격한 차이는 신경 과학 문헌을 통해서도 어느 정도 설명이

가능하다. 이 연령대의 행동 변화는 뇌의 구조적 변화와 관련이 있어 보인다(Horska et al., 2002; Tseng & O'Donnell, 2005). 전전두엽 피질(rostralateral prefrontal cortex: RLPFC)은 청소년기에 상당한 변화를 겪는다고 알려져 있다(Dumontheil, Burgess, & Blakemore, 2008). 전전두엽피질은 특히 관계추론능력을 관장하는 주요 기관이다(Hobeika, Diard-Detoeuf, Garcin, Levy, & Volle, 2016; Magis-Weinberg, Blakemore, & Dumontheil, 2017; Otero, 2017; Westphal, Reggente, Ito, & Rissman, 2016; Whitaker, Vendetti, Wendelken, & Bunge, 2018). 즉 본 연구결과 나타난 모순과 대조 관계추론능력의 변화는 전전두엽피질의 발달과 관련된 뇌의 구조적 변화에 기인했을 것으로 보인다. 또한 Jablansky 등의 연구(2016)에서 보여주었듯이 청소년 후기로 갈수록 유사나 변칙과 같은 다소 쉬운 관계추론 사고가 덜 두드러지는 경향이 있는데 이는 이 시기 학생들이 보다 고차적 사고(모순과 대조)를 하는데 집중함으로 인한 결과일 수 있다. 다만 Jablansky의 연구에서는 이러한 현상이 실험 환경에서 주어진 시간의 제약으로 인해 모순과 대조와 관련된 대화가 줄고 그 대신 유사나 변칙과 관련된 대화를 더 많이 하게 되는 현상으로 나타났다. 본 연구에서는 이처럼 유사-변칙 관계추론 사고 비율이 줄고 대신 모순-대조 관계추론 비율이 늘어나는 현상이 실제 중고등학생의 관계추론능력 검사 수행 점수를 통해 입증됨을 확인하였다.

본 연구에서 나타난 TORRjr의 차별기능문항 분석 결과에 따르면 관계추론능력은 학생들에게 익숙한 시험 형태가 유리하게 작용할

수 있어 보인다. MIRT-DIF 분석 결과 변칙 2번 문항은 초4와 초6 사이에서도, 초6과 중2 사이에서도 차별문항인 것으로 나타났으며 흥미롭게도 변칙 2번 문항의 정답률은 본 연구 참여자 가운데 가장 어린 학년인 초등학교 4학년 학생들에게서 높게 나타났다. 본 연구를 통해 이를 설명할 수 있는 명확한 단서를 찾을 수는 없었지만 학생들의 학습 경험을 통해 유추해볼 때 이 측정이 4학년에게 특히 유리했을 수 있을 것으로 보인다. 왜냐하면 이 측정 문항의 그림이 저학년 학생들이 자주 하는 활동인 종이접기와 유사했기 때문이다.

전체적 추세를 보면, 초4~중2 사이에서 차별기능 문항들이 상당히 많이 포착된 것과 대조적으로, 중2와 고1 사이에서는 차별기능문항이 보이지 않았다. 이는 초4, 초6, 중2 시기까지는 문항에 대한 학생들의 반응이 학년 간에 뚜렷하게 달랐다는 점을 보여 준다. 즉 같은 문항이 학년에 따라 불리하게, 혹은 유리하게 기능했을 수 있다. 반면에 중2와 고1에서는 차별기능문항이 나타나지 않았는데 이는 관계추론능력 발달이 일정 연령(대략 중2) 이후에 안정화된다는 앞서의 주장과 어느 정도 관련이 되어 보인다.

본 연구에서는 기존에 서구와 중국, 러시아 등지에서만 활용되었던 관계추론능력 검사의 한국판을 한국학생들에게 실시하였다. 그 결과, 관계추론 하위 영역과 학년에 따라 수행 점수 차이를 나타냄을 알 수 있었다. 특히 다소 난해한 문항인 모순과 대조 문항의 정답률은 유추와 변칙에 비해 늦은 시기, 즉 초6과 중2 사이에 차이를 보였다. 이는 기존에 유사성 유추만을 다루어 오던 관계추론 연구의 범

위가 넓어져야 함을 의미한다. 즉 유사, 변칙, 모순, 대조의 네 가지 하위 구인을 모두 고려하여 학령 간 차이를 보아야 비로소 언제 유추 능력이 높거나 낮은지 알 수 있을 것이다. 이러한 정교한 작업을 통해 발달연구의 주요 쟁점인 언제 특정 심리적 구인이 눈에 띄게 성장하는지에 대한 이해를 얻을 수 있을 것이다. 이러한 이해를 토대로 각 학령에 적합한 추론능력 배양 프로그램을 설계하기 위한 시사점을 얻을 수 있다. 유사와 변칙은 초4-초6 사이에 뚜렷한 변화를, 모순과 대조와 같은 보다 고등적 추론 능력은 초6-중2 사이에 뚜렷이 변하는 것으로 나타났다. 따라서 초등학교 6학년으로 넘어 가는 시기는 유사, 변칙 추론, 중2로 넘어 가는 시기에는 모순, 대조 추론 능력 배양에 초점을 둔 발달 중재를 고려하는 것이 타당한 것으로 나타났다. 또한 명확한 관계추론능력 파악을 위해 유사, 변칙, 모순, 대조의 네 개 구인이 모두 포함된 측정 도구를 활용함으로써 관계추론능력의 발달 정도를 보다 명확하게 이해하는데 도움이 되리라 예상된다. 본 연구를 통하여 관계추론능력에 대한 포괄적 접근이 필요함을 지지하는 근거를 얻을 수 있었다.

본 연구는 기존에 충분히 다루지 못했던 한국의 청소년들의 관계추론 발달 추이를 살펴보고려는 시도였다는 점에서 의의가 있다. 이러한 노력에도 불구하고 몇 가지 제한점이 있어 추후 연구에 반영할 필요가 있다. 우선 연구에서 활용한 표집 방식, 즉 동시에 각기 다른 연령 그룹에서 학생을 추출하여 발달연구를 진행한 점은 향후 보다 정교한 자료수집 설계를 통해 보완해야 할 사항이다. 물론 본

연구에서 활용한 특수한 분석 방법, 즉 MIRT-DIF를 통해 어느 정도 이러한 자료수집의 문제는 보완하고자 하였으나 발달과 관련하여 제기될 수 있는 여러 가지 연구 타당성, 즉 성장에 따른 생리적 변화나 인지발달 이외의 요인에 의한 성숙의 문제와 자료 오염의 문제를 피하려면 종단자료수집에 근거한 분석을 하는 것이 적절할 것이다. 표집설계의 한계로 인한 집단 간 차이를 보여주는 것인지 집단 내 차이를 보이고 있는지에 대한 결과의 혼재(confounding)의 문제는 추후 연구에서 보완되어야 할 사항이다.

본 연구에서 우리는 각기 다른 학교 급의 학생들이 문항에 반응하는 방식이 다르다는 점에 대해 양적 분석만을 제시하였는데 이에 대한 질적 분석이 추가되어야 할 것이다. 예컨대 본 연구에서 관계추론의 활동 중에 특정 문항은 특정 학년에서 눈에 띄게 우수한 수행력을 보이는 것을 발견하였는데, 이는 학령 간 반응의 질적 차이도 있을 것임을 암시한다. 관계추론능력의 발달 양상이 단순히 점수의 상향이 아닌, 각 하위 구인 능력별로 어떠한 질적 차이를 보이는지 향후 연구에서 다루어 주어야 할 것이다.

참고문헌

- Alexander, P. A. (1984). Training analogical reasoning skills in the gifted. *Roeper Review*, 6(4), 191-193.
<https://doi.org/10.1080/02783198409552807>
 Alexander, P. A. (2019). Individual differences

- in college-age learners: The importance of relational reasoning for learning and assessment in higher education. *British Journal of Educational Psychology*, 89(3), 416-428.
<https://doi.org/10.1111/bjep.12264>
- Alexander, P. A., Dumas, D., Grossnickle, E. M., List, A., & Firetto, C. M. (2016). Measuring relational reasoning. *Journal of Experimental Education*, 84(1), 119-151.
<https://doi.org/10.1080/00220973.2014.963216>
- Alexander, P. A., & the Disciplined Reading and Learning Research Laboratory (2012). *Test of relational reasoning*. College Park, MD: University of Maryland.
- Alexander, P. A., White, C. S., Haensly, P. A., & Crimmins-Jeanes, M. (1987). Training in analogical reasoning. *American Educational Research Journal*, 24(3), 387-404.
<https://doi.org/10.3102/00028312024003387>
- Andrews, G. (1996). Assessment of relational reasoning in children aged 4 to 8 years. *Society for the Study of Behavioural Development*, 26. Retrieved August 3, 2021, from <https://www.eric.ed.gov/ERICWebPortal/detail?accno=ED402073>
- Buehl, M. M., & Alexander, P. A. (2004). Longitudinal and cross-cultural trends in young childrens' analogical and mathematical reasoning abilities. In L. D. English (Ed.), *Mathematical and analogical reasoning of young learners* (pp. 47-73). New York: Routledge.
- Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: A maturational component? *Developmental Psychology*, 16(4), 257-269.
<https://doi.org/10.1037/0012-1649.16.4.257>
- Chalmers, R. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 32480.
<https://doi.org/10.18637/jss.v048.i06>
- Chen, Z. (1996). Children's analogical problem solving: The effects of superficial, structural, and procedural similarity. *Journal of Experimental Child Psychology*, 62(3), 410-431.
<https://doi.org/10.1006/jecp.1996.0037>
- Diamond, R., Carey, S., & Back, K. J. (1983). Genetic influences on the development of spatial skills during early adolescence. *Cognition*, 13(2), 167-185.
[https://doi.org/10.1016/0010-0277\(83\)90021-5](https://doi.org/10.1016/0010-0277(83)90021-5)
- Dumas, D., & Alexander, P. A. (2016). Calibration of the test of relational reasoning. *Psychological Assessment*, 28(10), 1303-1318.
<https://doi.org/10.1037/pas0000267>
- Dumas, D., & Alexander, P. A. (2018). Assessing differential item functioning on the test of relational reasoning. *Frontiers in Education*, 3(14). <https://doi.org/10.3389/feduc.2018.00014>
- Dumas, D., Alexander, P. A., & Grossnickle, E. M. (2013). Relational reasoning and its manifestations in the educational context: A systematic review of the literature. *Educational Psychology Review*, 25(3), 391-427.
<https://doi.org/10.1007/s10648-013-9224-4>
- Dumontheil, I., Burgess, P. W., & Blakemore, S.

- J. (2008). Development of rostral prefrontal cortex and cognitive and behavioural disorders. *Developmental Medicine & Child Neurology*, 50(3), 168-181.
<https://doi.org/10.1111/j.1469-8749.2008.02026.x>
- Dumontheil, I., Houlton, R., Christoff, K., & Blakemore, S. J. (2010). Development of relational reasoning during adolescence. *Developmental Science*, 13(6), 15-25.
<https://doi.org/10.1111/j.1467-7687.2010.01014.x>
- Falissard, B. (2012). Software package ‘psy’. Various procedures used in psychometry version. *R Package Version 1.1*. Retrieved August 30, 2021, from
<https://cran.r-project.org/web/packages/psy/index.html>
- Goswami, U. (1989). Relational complexity and the development of analogical reasoning. *Cognitive Development*, 4(3), 251-268.
[https://doi.org/10.1016/0885-2014\(89\)90008-7](https://doi.org/10.1016/0885-2014(89)90008-7)
- Goswami, U. (2013). Analogical reasoning in children. In J. Campione & K. Metz (Eds.), *Children’s learning in laboratory and classroom contexts: Essays in honor of Ann Brown* (pp. 55-70). New York: Taylor & Francis.
<https://doi.org/10.4324/9780203826966>
- Hobeika, L., Diard-Detoeuf, C., Garcin, B., Levy, R., & Volle, E. (2016). General and specialized brain correlates for analogical reasoning: A meta-analysis of functional imaging studies. *Human Brain Mapping*, 37(5), 1953-1969. <https://doi.org/10.1002/hbm.23149>
- Horska, A., Kaufmann, W. E., Brant, L. J., Naidu, S., Harris, J. C., & Barker, P. B. (2002). In vivo quantitative proton MRSI study of brain development from childhood to adolescence. *Journal of Magnetic Resonance Imaging*, 15(2), 137-143.
<https://doi.org/10.1002/jmri.10057>
- Inhelder, B., & Piaget, J. (1964). Class inclusion and hierarchical classification. In B. Inhelder & J. Piaget (Eds.), *The early growth of logic in the child: Classification and seriation* (pp. 100-150). Oxon: Routledge.
- Jablansky, S., Alexander, P. A., Dumas, D., & Compton, V. (2016). Developmental differences in relational reasoning among primary and secondary school students. *Journal of Educational Psychology*, 108(4), 592-608.
<https://doi.org/10.1037/edu0000070>
- Kranzler, J. H., Benson, N., & Floyd, R. G. (2015). Using estimated factor scores from a bifactor analysis to examine the unique effects of the latent variables measured by the WAIS-IV on academic achievement. *Psychological Assessment*, 27(4), 1402-1416.
<https://doi.org/10.1037/pas0000119>
- Magis-Weinberg, L., Blakemore, S.-J., & Dumontheil, I. (2017). Social and nonsocial relational reasoning in adolescence and adulthood. *Journal of Cognitive Neuroscience*, 29(10), 1739-1754.
https://doi.org/10.1162/jocn_a_01153
- McGivern, R. F., Andersen, J., Byrd, D., Mutter, K. L., & Reilly, J. (2002). Cognitive efficiency on a match to sample task decreases at the

- onset of puberty in children. *Brain and Cognition*, 50(1), 73-89.
[https://doi.org/10.1016/S0278-2626\(02\)00012-X](https://doi.org/10.1016/S0278-2626(02)00012-X)
- Murphy, P. K., Firetto, C. M., & Greene, J. A. (2017). Enriching students' scientific thinking through relational reasoning: Seeking evidence in texts, tasks, and talk. *Educational Psychology Review*, 29(1), 105-117.
<https://doi.org/10.1007/s10648-016-9387-x>
- Otero, T. M. (2017). Brief review of fluid reasoning: Conceptualization, neurobasis, and applications. *Applied Neuropsychology: Child*, 6(3), 204-211.
<https://doi.org/10.1080/21622965.2017.1317484>
- R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
- Raven, J. (2003). Raven progressive matrices. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 223-237). Boston: Springer.
https://doi.org/10.1007/978-1-4615-0153-4_11
- Revelle, W. (2019). Software package 'psych.' Retrieved August 3, 2021, from <https://personality-project.org/r/psych-manual.pdf>
- Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological Science*, 24(1), 87-92.
<https://doi.org/10.1177/0956797612450883>
- Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94(3), 249-273.
<https://doi.org/10.1016/j.jecp.2006.02.002>
- Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, 316(5828), 1128-1129.
<https://doi.org/10.1126/science.1142103>
- Tseng, K. Y., & O'Donnell, P. (2005). Post-pubertal emergence of prefrontal cortical up states induced by D₁-NMDA co-activation. *Cerebral Cortex*, 15(1), 49-57.
<https://doi.org/10.1093/cercor/bhh107>
- Westphal, A. J., Reggente, N., Ito, K. L., & Rissman, J. (2016). Shared and distinct contributions of rostrolateral prefrontal cortex to analogical reasoning and episodic memory retrieval. *Human Brain Mapping*, 37(3), 896-912. <https://doi.org/10.1002/hbm.23074>
- Whitaker, K. J., Vendetti, M. S., Wendelken, C., & Bunge, S. A. (2018). Neuroscientific insights into the development of analogical reasoning. *Developmental Science*, 21(2), e12531. <https://doi.org/10.1111/desc.12531>
- Wilson, M., Zheng, X., & McGuire, L. (2012). Formulating latent growth using an explanatory item response model approach, *Journal of Applied Measurement*, 13(1), 1-22. Retrieved August 30, 2021, from <https://bearcenter.berkeley.edu/sites/default/files/Wilson%20%237.pdf>
- Zhao, H., & Sun, Y. (2019, March). *Development of relational reasoning ability among chinese*

elementary and middle school students [Paper presentation]. The Annual Meeting of American Educational Research Association, San Antonio, TX, United States.

1 차원고접수 : 2021. 05. 13.

수정원고접수 : 2021. 08. 16.

최종게재결정 : 2021. 09. 27.

Relational Reasoning Ability of South Korean School-Age Students: Use of the Test of Relational Reasoning (TORR)

Chae, Soo Eun

Division of Teacher Education, Gangneung-Wonju National University

This study explores how the recently developed Test of Relational Reasoning (TORR) performs in different age groups of South Korean elementary, middle, and high school students. TORR was applied to South Korean adolescents and analyzed using repeated measure ANOVAs and differential item functioning techniques. As a result of conducting TORR on 749 adolescents enrolled in elementary school grade 4 ($N = 196$) and grade 6 ($N = 200$), middle school grade 2 ($N = 206$), and high school grade 1 ($N = 146$) in Gangwon-do (6th grade: 105 boys (52.5%), middle school 2 grade: 81 boys (39.32%), and high school 1: 63 boys (43.15%)), we found that relational reasoning test performances tended to increase as the students' grade level increased. However, there were differences in the patterns of change in scores according to grade levels for each of the four relational reasoning forms. While the antinomy and antithesis scores were significantly different between elementary grade 6 and middle grade 2, analogy and anomaly scale scores were significantly different between elementary grades 4 and 6. As a result of this study, the existing reasoning research centering on analogical reasoning should be expanded to incorporate various formats.

Keywords: relational reasoning, analogy, anomaly, antinomy, antithesis