

## 베트남 고등학생을 대상으로 한 개정 시스템 사고 검사 도구 재타당화 및 한국과 베트남 고등학생의 잠재 평균 비교

이호녕<sup>1</sup> · Nguyen Thi Thuy<sup>2</sup> · 박병열<sup>3</sup> · 전재돈<sup>4</sup> · 이현동<sup>5,\*</sup>

<sup>1</sup>경북대학교 지구과학교육과, 41566, 대구광역시 북구 대학로 80

<sup>2</sup>Thuan Hoa High School, Hue 54000, Vietnam

<sup>3</sup>단국대학교부설 통합과학교육연구소, 16890, 경기도 용인시 수지구 죽전로 152

<sup>4</sup>경북대학교 과학교육연구소, 41566, 대구광역시 북구 대학로 80

<sup>5</sup>대구교육대학교 과학교육과, 42411, 대구광역시 남구 중앙대로 219

## Re-validation of the Revised Systems Thinking Measuring Instrument for Vietnamese High School Students and Comparison of Latent Means between Korean and Vietnamese High School Students

Hyonyong Lee<sup>1</sup>, Nguyen Thi Thuy<sup>2</sup>, Byung-Yeol Park<sup>3</sup>,  
Jaedon Jeon<sup>4</sup>, and Hyundong Lee<sup>5,\*</sup>

<sup>1</sup>Department of Earth Science Education, Kyungpook National University, Daegu 41566,  
Republic of Korea

<sup>2</sup>Thuan Hoa High School, Hue 54000, Vietnam

<sup>3</sup>Institute of Integrated Science Education, Dankook University, Yongin 16890, Republic of Korea

<sup>4</sup>Science Education Research Institute, Kyungpook National University, Daegu 41566, Republic of Korea

<sup>5</sup>Department of Science Education, Daegu National University of Education, Daegu 42411,  
Republic of Korea

**Abstract:** The purposes of this study were: (1) to revalidate the revised Systems Thinking Measuring Instrument (Re\_STMI) reported by Lee et al. (2024) among Vietnamese high school students and (2) to investigate the differences in systems thinking abilities between Korean and Vietnamese high school students. To achieve this, data from 234 Vietnamese high school students who responded to translated Re\_STMI consisting of 20 items and an Scale consisting of 20 items were used. Validity analysis was conducted through item response analysis (Item Reliability, Item Map, Infit and Outfit MNSQ, DIF between male and female) and exploratory factor analysis (principal axis factor analysis using Promax). Furthermore, structural equation modeling was employed with data from 475 Korean high school students to verify the latent mean analysis. The results were as follows: First, in the item response analysis of the 20 translated Re\_STMI items in Vietnamese, the Item Reliability was .97, and the Infit MNSQ ranged from .67 to 1.38. The results from the Item Map and DIF analysis align with previous findings. In the exploratory factor analysis, all items were loaded onto intended sub-factors, with sub-factor reliabilities ranging from .662 to .833 and total reliability at .876. Confirmatory factor analysis for latent mean analysis between Korean and Vietnamese students yielded acceptable model fit indices ( $\chi^2/df$ : 2.830, CFI: .931, TLI: .918, SRMR: .043, RMSEA: .051). Lastly, the latent mean analysis between

\*Corresponding author: leehd@dnue.ac.kr  
Tel: +82-53-620-1349

Korean and Vietnamese students revealed a small effect size in systems analysis, mental models, team learning, and shared vision factors, whereas a medium effect size was observed in personal mastery factors, with Vietnamese high school students showing significantly higher results in systems thinking. This study confirmed the reliability and validity of the Re\_STMI items. Furthermore, international comparative studies on systems thinking using Re\_STMI translated into Vietnamese, English, and other languages are warranted in the context of students' systems thinking analysis.

Keywords: Revised Systems Thinking Measuring Instrument, Systems Thinking Scale, Reliability, Validity, Vietnamese High School Student

**요약:** 이 연구의 목적은 1) 베트남 고등학생들을 대상으로 Lee et al. (2024)에서 보고한 개정 시스템 사고 검사 도구의 타당도를 재검증하고, 2) 연구에 참여한 우리나라 고등학생과 베트남 고등학생 간 시스템 사고 능력에 대한 차이를 알아보는 것이다. 이를 위하여 베트남 고등학생 234명이 베트남어로 번역된 개정 시스템 사고 검사 도구 20문항과 STS 척도 20문항에 응답한 자료를 활용하였다. 타당도 분석은 문항 반응 분석(Item Reliability, Item Map, Infit and Outfit MNSQ, 남녀 집단의 DIF)과 탐색적 요인 분석(프로맥스를 활용한 주축 요인 분석)을 통해 검증하고, 나아가 우리나라 고등학생 475명의 데이터를 함께 활용하여 구조 방정식 모형을 이용한 잠재평균비교를 통해 검증하였다. 연구 결과는 다음과 같다. 첫째, 베트남어로 번역된 Re\_STMI 20문항의 문항 반응 분석 결과 Item Reliability는 .97, Infit MNSQ는 .67-1.38으로 나타났으며 Item Map과 DIF 분석에서도 선행 연구에서 보고된 결과와 일치하는 결과가 도출되었다. 탐색적 요인 분석에서는 모든 문항들이 의도한 하위 요인에 적재되었으며, 요인별 신뢰도는 .662-.833, 전체 신뢰도는 .876으로 분석되었다. 우리나라 고등학생과 잠재평균비교를 위한 확인적 요인 분석에서 도출된 모형 적합도 수치는 모두 수용 가능한 값으로 분석되었다( $\chi^2/df$ : 2.830, CFI: .931, TLI: .918, SRMR: .043, RMSEA: .051). 마지막으로 연구에 참여한 우리나라 고등학생과 베트남 고등학생 간 잠재평균비교에서는 시스템 분석, 정신 모델, 팀 학습, 공유 비전 요인에서 작은 효과 크기가, 개인 숙련 요인에선 중간 이상의 효과 크기를 보이며, 베트남 고등학생들이 시스템 사고 능력에서 유의미하게 높은 결과를 보여주었다. 이를 통해 개정 시스템 사고 검사 도구 문항은 안정적인 신뢰도와 타당도를 가지고 있음을 확인할 수 있었다. 앞으로 학생들의 시스템 사고 연구와 관련하여 베트남어 및 영어 등으로 번역한 문항을 활용하여 시스템 사고의 국제 비교 연구도 진행할 필요성이 있을 것이다.

주요어: 개정 시스템 사고 검사 도구, 시스템 사고 척도, 신뢰도, 타당도, 베트남 고등학생

## I. 서론

1990년대부터 국내의 학교 교육에서 학생들의 고등 사고 능력을 향상시키기 위해 시스템 사고(Systems Thinking)가 적용되었다. 이러한 시스템 사고는 학생들의 복잡한 문제해결을 위한 사고력 증진, 지적 성장, 자기 효능감이나 동기 유발 등에 유의미한 영향을 준다(Ben-zvi Assaraf & Orion, 2009; Kali et al., 2003; Kang et al., 2008; Kwon et al., 2011; Lee et al., 2011; Lee & Lee, 2017; Moon et al., 2004). 더욱이 2022 개정 과학과 교육과정에서도 시스템과 상호 작용은 주요 핵심 아이디어로 제시되어 있으며, 시스템적인 지식(예, 물질의 순환, 에너지의 순환 등)의 학습이나 체계적인 문제해결을 위해 시스템 사고의 적용을 강조하고 있다(MOE, 2022). 선행 연구를 살펴보면, 시스템 사고를 통해 학생들의 고등 사고 능력, 핵심 역량 및 기초 소양 등의 함양이 향상될 수 있음을 제시하고 있다(Batt et al., 2021; Kowch, 2019; Ulku & Gaye, 2024).

국내외에서 시스템 사고 연구가 진행되면서 학생들의 시스템 사고 능력을 측정하는 다양한 방법이 개발 및 시도 되었다. 이러한 방법과 시도 중 우리나라 과학 교육에서 많이 인용되고 활용되었던 것은 Ben-zvi Assaraf & Orion(2005a, 2005b)이 제시한 방법이다. 이 연구에서는 물 순환과 관련하여 학생들의 시스템 사고 능력의 변화를 살펴보고자 시스템 사고 검사지, 단어 간 관계 분석, 인과 지도 그리기, 그림 그리기, 인터뷰 등에서 도출된 다양한 정성적 데이터를 활용하여 학생들의 시스템 사고 변화를 보고하였다. 이 연구에서 활용된 방법은 이후 국내 논문에도 다수 인용되었다. 예를 들어 Kang et al.(2008)에서는 인과 지도를 이용한 시스템 사고 분석을, Lee et al.(2011)에서는 단어 간 관계 분석, 인과 지도, 그림 그리기, 면담 등의 방법을 활용하여 학생들의 시스템 사고 능력을 분석하고, 교수-학습 프로그램 전과 후의 시스템 사고 변화를 살펴보는데 활용되었다. 또한 이 도구들은 시스템 사고의 변화를 살펴보는데 유용한 도구로 활용되었으며, 교육학적으로도 시스템 사

고의 변화가 학생들의 성장에 유의미한 영향을 준다는 결과를 나타내었다(Kwon et al., 2011; Im & Lee, 2014; Song et al., 2015; Lee, 2020).

이러한 시스템 사고를 측정하는 연구가 진행되는 과정에서 시스템 사고의 변화를 정성적 데이터에 의존하여 그 변화를 살펴본다는 문제점을 극복하고자 정량적 도구 개발의 필요성이 제기되었다. 가장 대표적인 정량적 시스템 사고 검사 도구(STMI)는 Senge(1996, 2012)의 학습 조직의 5가지 수련에 제시된 5개 하위 요인을 토대로 Lee et al.(2013)에 의해 개발되었다. 이 검사 도구는 5개 하위 요인별로 4문항씩, 총 20문항으로 구성된 자기 보고식 검사 도구로, 정성적 검사 도구의 결과를 정량적으로 상호 보완할 수 있는 검사지로 개발되었으며 이후 다양한 집단, 다른 검사 도구와의 관련성을 통해 그 신뢰도와 타당도가 검증되었다(Han & Jo, 2015; Jeon & Lee, 2015; Cho & Hwang, 2016; Lee & Lee, 2016; Park & Choi, 2018; Park et al., 2019; Lee, 2020).

시스템 사고에 관련된 연구가 지속적으로 수행되면서 연구 대상이 확대되고, 측정 도구나 검사 결과에 대한 기법이 다양하게 발전하였다. 이 과정에서 주목할 만한 변화는 연구 대상자의 시스템 사고 능력을 수준으로 정의하고, 지속적인 교육을 통해 도달 가능한 상위점을 설정한 연구가 수행되었다는 것이다(Lee & Lee, 2017; Lee et al., 2018; Park et al., 2019). Lee et al.(2018)에서는 시스템 사고와 관련된 교육학, 사회학, 정치학 등 다양한 선행 연구 분석을 통해 시스템 사고 수준을 4단계로 설정하였으며, 각 단계별 학습자들이 수행할 수 있는 시스템 사고 능력을 정의하고 제시하였다. 그리고 Hung(2008)의 연구를 기반으로 하여 정성적 데이터를 통해 시스템 사고 수준에 도달 여부를 정량적으로 측정할 수 있는 루브릭을 개발하였다. 즉, 앞에서 제시한 흐름처럼 국내에서 교육과 관련된 시스템 사고 연구는 시스템 사고를 향상시키는 다양한 교육 프로그램의 개발뿐만 아니라 학생들의 변화를 신뢰롭고 타당하게 측정하고자 하는 부분에서 활발한 연구가 진행되어 왔다(Im & Lee, 2014; Jeon & Lee, 2015; Lee et al., 2011; Moon et al., 2004). 즉, 교육학이나 경제학, 정치학 등 다양한 분야에서 연구 대상이 되는 학생들이나 성인을 대상으로 시스템 사고의 수준이나 역량을 측정하는 연구는 여전히 활발하게 이루어지고 있다(Aguilar-Cisneros et al., 2022; Momsen et al., 2022;

Ratinen et al., 2023).

교육학이나 과학 교육 분야에서 적용되었던 시스템 사고는 간호학, 의학 계열 대학생들에게도 필요한 사고 능력으로 제기되면서 다양한 영역에서 대학생을 대상으로 시스템 사고 능력을 분석하거나, 다른 요인들과의 관련성을 파악하는 연구도 수행되었다(Cho & Hwang, 2016; Cho & Kim, 2017; Lee et al., 2020). 대학생에 대한 시스템 사고 능력을 살펴보는 연구는 국내뿐 아니라 국외에서도 진행되었는데, Davis & Stroink(2016)는 심리학 전공 분야와 관련하여 시스템 사고 척도 문항을 개발하고 대학생들을 대상으로 시스템 사고를 측정하였으며, Dolansky et al.(2020)는 의학 계열에서 필요한 역량으로 시스템 사고를 제시하면서 대학 교육에서 시스템 사고 능력을 측정하고자 시스템 사고 측정 도구(STS; Systems Thinking Scale)을 개발하고 타당화하는 연구가 진행되었다. 즉, 시스템 사고 능력에서 요구되는 시스템 내 패턴과 상호작용을 인식하고 종합하는 능력이나 상호작용에 따른 행동이 가져오는 강화적 또는 균형적 피드백 등이 교육학뿐만 아니라 다양한 영역에서도 필요한 사고방식이라는 점을 이 연구에서 확인할 수 있다. 게다가 이 연구에서 개발한 STS는 시스템 사고와 관련된 다양한 문헌 연구를 통해 최종적으로 Senge(1996, 2012)가 제시한 시스템 사고 향상을 위한 학습 조직의 5가지 요소(Fifth Discipline)를 근거로 문항을 구성하였다. 즉, 국외에서도 국내에서 개발한 시스템 사고 검사 도구(STMI)와 동일한 이론적 배경과 정의를 가진 문항들을 구성하여 시스템 사고 측정을 위한 문항을 제작하였다는 것을 확인할 수 있다.

선행 연구로부터 STMI를 구성하는 문항의 경우 시스템 사고 측정을 위한 타당한 이론적 배경, 정의, 요인을 가지고 있다는 것을 국외 연구 사례로부터 확인할 수 있으며, 국내와 국외 검사지 간 상관 분석을 통해 신뢰도와 타당도를 확인할 수 있었다. 최근 Lee et al.(2023)에서는 개발도상국을 대상으로 한 교육 영역의 개발 원조 사업(ODA; Official Development Assistance)의 일환으로 베트남을 방문하여 베트남 고등학생들을 대상으로 STMI와 STS 두 척도에 대하여 구조 방정식 모형을 활용한 모형 적합도 분석을 실시하였다. 베트남 고등학생들에게 STEM 교육을 실시하면서 학생들과 학교장의 동의를 받아 연구를 진행하였으며 해당 연구의 결과에서 두 척도 간 상

관 계수가 .90 이상의 높은 상관관이 도출되었다. 그리고 Lee et al.(2024)에서 국내 고등학생들을 대상으로 문항을 개정한 시스템 사고 검사 도구(Re\_STMI)와 STS 검사지 간 상관 분석을 실시하였으며, 선행 연구에서도 .70 이상의 상관 계수가 도출되어 지금까지 개발된 국·내외의 시스템 사고 측정 도구의 높은 신뢰도와 타당도를 확인할 수 있었다. 이와 같이 검사 도구 또는 척도를 개발하거나 수정 및 보완하는 경우 척도의 타당도를 지속적으로 검증하여 해당 척도가 의도하는 요인을 측정하고 있는지에 대한 추가적인 검증은 척도 개발 연구에서 필요한 절차 중 하나이다(Rust & Golombok, 2014). 그러므로 국내에서 개발되고 타당화한 개정 시스템 사고 검사 도구(Re\_STMI)의 경우, 국내에서 추가적인 타당도를 검증할 필요가 있고, 국외에서도 검사지를 제대로 활용하기 위해 타당도를 검증할 필요성이 있다.

이 연구는 베트남 교육 봉사 활동 과정에서 베트남 학생의 시스템 사고를 분석한 Lee et al.(2023)의 후속 연구이다. Lee et al.(2023)의 연구에서는 베트남 고등학생들에게 2013년에 보고된 시스템 사고 검사 도구를 활용하여 해당 도구의 타당도를 확인하고 나아가 STS 검사지 간 상관 분석을 통해 공인 타당도를 검증하였다. 시스템 사고 검사 도구의 신뢰도와 타당도를 확인한 결과 일부 문항의 타당도가 낮게 도출되어 검사 도구 문항의 일부 수정의 필요성이 제기되었다. 이를 토대로 Lee et al.(2024)에서는 시스템 사고 검사 도구를 수정하고 국내 고등학생들에게 투입하여 타당도를 검증한 개정 시스템 사고 검사 도구를 보고하였다. 이 연구에서는 추가적인 타당도 검증 및 국내에서 개발된 검사지가 국외 학생들에게도 타당한지에 대한 여부를 살펴보기 위하여 2024년 1월 4주 동안 K대학에서 실시한 베트남 단기 해외교육봉사 교육 프로그램에 참여한 베트남 고등학생들을 대상으로 개정된 시스템 사고 검사 도구에 대한 타당도를 확인하는 연구를 수행하였다. 이를 위해 개정 시스템 사고 검사 도구 문항을 베트남어로 번역하였으며, 번역된 문항에 대하여 베트남 학생들이 신뢰롭고 타당하게 응답하는지를 확인하고, Re\_STMI를 국외 연구에 활용할 수 있는지에 대한 타당도를 검증하고자 하였다.

이 연구 목적을 달성하기 위해, Lee et al.(2024)에 의해 개발되고 우리나라 고등학생을 대상으로 타당도

를 검증한 개정 시스템 사고 검사 도구(Re\_STMI)를 베트남어로 번역하였다. 그리고 베트남 고등학생들을 대상으로 Re\_STMI를 투입하여 개정된 검사지의 신뢰도와 타당도(문항 반응 분석, 탐색적 요인 분석)를 재검증하고자 하였다. 또한 개정된 검사지로부터 얻은 우리나라 고등학생 데이터를 함께 활용하여 구조 방정식 모형을 통해 추가적인 타당도 검증을 실시하면서 베트남 고등학생들의 시스템 사고 능력을 잠재 평균비교를 통해 분석하였다. 이 연구 목적을 위해 설정한 연구 문제는 다음과 같다.

첫째, 베트남 학생들을 대상으로 개정 시스템 사고 검사 도구의 문항 반응 분석 결과는 타당한가?

둘째, 베트남 학생들을 대상으로 개정 시스템 사고 검사 도구의 탐색적 요인 분석 결과 및 공인 타당도는 타당한가?

셋째, 구조방정식 모형을 이용한 잠재평균비교로부터 도출된 한국과 베트남 고등학생의 시스템 사고 능력에는 차이가 있는가?

## II. 연구 방법

### 1. Process of study

이 연구의 과정은 Fig. 1과 같다. 베트남 학생들을 대상으로 개정된 시스템 사고 검사 도구의 타당도를 검증하기 위하여 신뢰도 분석 및 문항 반응 분석, 탐색적 요인 분석을 실시하였다. 그리고 공인 타당도 검증을 위하여 Dolansky et al.(2020)에서 개발한 시스템 사고 척도(STS) 문항을 활용하여 상관 분석을 실시하였다.

우리나라와 베트남 고등학생과의 시스템 사고 차이를 알아보기 위한 방법으로는 구조 방정식 모형을 활용한 잠재평균분석을 실시하였다. 우리나라 고등학생의 데이터는 Lee et al.(2024)의 개정 시스템 사고 검사 도구 재타당화에서 활용한 우리나라 고등학생 데이터를 이용하였다. 잠재평균분석을 위하여 구조 방정식 모형을 활용한 확인적 요인 분석을 실시하였다. 그리고 잠재평균분석을 위하여 모형에 제약을 가하여 형태 동일성, 매트릭 동일성, 절편 동일성, 요인 분산 동일성 검증을 통하여 모형이 지지되는지를 확인한 후, 우리나라 고등학생과 베트남 고등학생 간 시스템 사고 검사 도구 결과의 효과 크기를 살펴보고자 하였다.

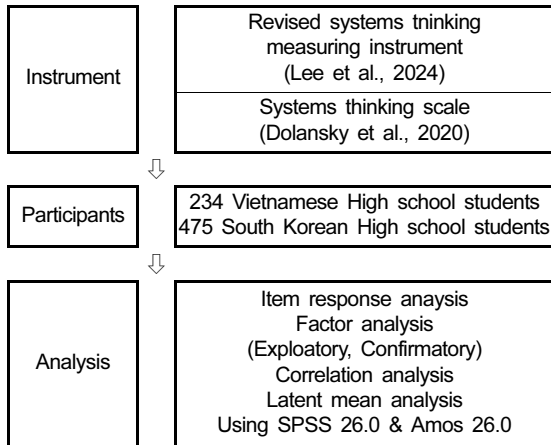


Fig. 1. Process of Study

## 2. 연구 대상

베트남 고등학생을 대상으로 개정 시스템 사고 검사 도구의 재타당화를 위하여 베트남 H시에 위치한 T 고등학교의 학생을 대상으로 검사를 실시하였다. 연구에 참여한 학교는 베트남에서 대학과 연계하여 STEM 교육을 실시하는 학교로 베트남 내에서도 우수한 학생들로 구성된 고등학교이다. 베트남의 H시 교육청에 교육 봉사를 진행하면서 Lee et al.(2023)에서 고등학생들의 시스템 사고를 측정하는 부분이 있었으며, 이와 관련한 후속 연구의 성격이 있어 교육 봉사에 참여한 베트남 고등학생들을 대상으로 연구가 수행되었다. 검사와 관련하여 고등학교 전체 학생들에게 검사의 목적과 데이터 활용 여부에 대한 안내를 실시하였으며 학교장의 사전 동의를 받고 조사를 시작하였으며 학부모 및 학생의 동의를 받았다. 설문은 <https://forms.gle/du1Rba8wjBa4FyNb8>을 통해 구글로 응답하도록 안내하였으며, 총 247명의 베트남 고등학생이 설문에 응답하였다. 응답한 데이터 중 일관된 번호를 응답하거나 무성의한 응답을 한 학생들의 데이터를 제외한 234명의 데이터로부터 결과를 도출하였다. 234명의 데이터 중 남학생은 101명으로 43.2%, 여학생은 133명으로 56.8%에 해당하였다. 잠재평균비교를 위한 우리나라 고등학생의 데이터는 무선 표집된 475명의 고등학생들이 응답한 Lee et al.(2024)에서 사용한 데이터를 활용하였다.

## 3. 검사 도구

1) 개정 시스템 사고 검사 도구(Revised Systems Thinking Measuring Instrument; Re\_STMI, 2024)

이 연구에서 타당도를 재검증하고 나아가 우리나라와 베트남 고등학생의 시스템 사고 능력을 비교하기 위한 검사 도구로는 Lee et al.(2024)에서 타당화한 개정 시스템 사고 검사 도구를 이용하였다. 이 검사 도구는 시스템 사고가 5개의 하위 요인으로 구성되었다는 이론적 배경을 바탕으로 20개 문항(정신 모델 4문항, 개인 숙련 4문항, 시스템 분석 4문항, 공유 비전 4문항, 팀 학습 4문항)으로 개발되었다. 검사지 개발 이후 10년이 지나면서 사회·문화적 맥락이 맞지 않는 문항 내용 등이 나타나면서 최근 문항의 개정을 실시하였다. 예를 들어 기존의 문항에서는 ‘신문 기사나 뉴스 등의 매체를 TV로부터 접한다’라는 내용을 유튜브나 스마트 기기를 활용하는 부분으로 수정하는 등의 수정 보완과 함께 문항 도입 시 ‘나는 평소 스스로 생각할 때’와 같이 도입 문구를 제시하고, 자기 보고식 리커트 5단계 척도로 응답하도록 구성하였다. 그리고 베트남 학생들에게 검사 도구를 적용하기 위하여 국내 베트남 유학생 중 한국어 및 영어에 능숙한 이중 언어자가 한국어로 된 시스템 사고 검사 도구 문항을 베트남어로 번역하였다(Table 1). 베트남어로 번역된 내용과 한국어 버전과의 내용 타당도 확인을 위해 한국어로 재번역 후, 재번역된 내용에 대하여 시스템 사고 및 과학교육 전문가 2인이 내용 타당도를 확인하였다. 응답한 결과를 바탕으로 분석된 20문항 전체의 Cronbach- $\alpha$ 값은 .868로 나타났다, 하위 요인에서는 .678-.814로 분석되었다.

2) 시스템 사고 척도(STS; Systems Thinking Scale, 2020)

Re\_STMI의 공인 타당도를 검증하기 위하여 Dolansky et al.(2020)에서 개발한 STS 검사 도구를 활용하여 상관 분석을 실시하였다. 이 검사지는 Lee et al.(2023), Lee et al.(2024)의 연구에서도 동일한 척도 간 공인 타당도 검증을 위해 활용되었으며, 각 연구에서 모두 두 척도 간 .7 이상의 높은 상관을 나타내었다. 이 연구에서도 두 선행 연구에서의 결과와 유사한 결과가 도출될 경우 두 척도 간 공인 타당도가 검증될 뿐만 아니라, 두 검사지 모두 시스템 사고를 측정하는 검증된 검사 도구로 적용할 수 있다는 결론을 도출할 수 있다. 이를 위하여 Lee et al.(2023)에서 베트남어로 번역된 20문항을 활용하였으며, 베트남 학생들이 응답한 결과로부터 도출된 신뢰도(Cronbach- $\alpha$ )는 .914로 나타났다(Table 2).

**Table 1.** Example of Revised Systems Thinking Measuring Instrument Items (Korean-Vietnamese)

Re_STMI	Korean	Vietnamese
Item 1	나는 평소 스스로 생각할 때, 내가 원하는 결과가 나오지 않으면, 반성의 시간을 가진다.	Khi tôi nghĩ về bản thân mình, Thời gian tôi dành ra để tự suy nghĩ về bản thân nhiều hơn 1 lần một tuần.
Item 2	나는 목표나 계획을 세울 때, 항상 내가 처해진 현재 상황을 고려한다.	Khi tôi lên một kế hoạch hoặc mục tiêu nào đó, Tôi luôn nghĩ về tình hình hiện tại của mình.
Item 3	나는 모둠 학습(활동)을 할 때, 나의 의견을 적극적으로 이야기한다.	Khi tôi hoạt động nhóm, Tôi đóng góp ý kiến của mình một cách tích cực.
Item 4	나는 어떤 문제 상황을 마주하였을 때, 가능한 다양한 관점을 고려한다.	Khi tôi đối mặt với vấn đề nào đó, Tôi cân nhắc tất cả các quan điểm một cách đa dạng.
Item 5	내가 모둠의 리더라면, 활동에서 다른 구성원들 사이의 관계를 고려한다.	Khi tôi là đội trưởng của một nhóm, Tôi luôn để ý đến mối quan hệ giữa các thành viên trong hoạt động.

**Table 2.** Example of STS Items

STS	Korean	Vietnamese
Item 1	나는 상황에 대한 다른 사람들의 관점을 찾아본다.	Tôi sẽ tìm kiếm quan điểm của người khác về tình hình đó.
Item 2	나는 문제의 원인을 결정하기 위해, 하나의 특정한 사건을 넘어 살펴본다.	Tôi sẽ nhìn lại một sự kiện nhất định nào đó để quyết định nguyên nhân của vấn đề.
Item 3	나는 일련의 사건들이 어떻게 발생하는 지에 대해 이해하는 것이 매우 중요하다고 생각한다.	Tôi nghĩ rằng việc hiểu nguyên nhân và lý do khiến các sự kiện xảy ra là rất quan trọng.
Item 4	나는 팀원들과 함께 해결책을 찾는다.	Tôi luôn tìm cách giải quyết vấn đề cùng với các thành viên khác.
Item 5	나는 특정한 사건보다 반복되는 패턴(양상)이 더 중요하다고 생각한다.	Tôi nghĩ rằng các khuôn mẫu lặp lại xảy ra một cách lặp lại quan trọng hơn là một sự kiện nào đó.

**4. 분석 방법**

베트남 학생들을 대상으로 Re\_STMI의 타당도 검증에 위하여 234명의 응답 결과를 Winsteps, SPSS 26.0과 Amos 21.0을 이용하여 분석하였다. Lee(2019)와 Lee et al.(2024)의 연구 방법과 같이 문항 반응 분석은 Rating Scale model을 활용하여 Item Reliability, Item Map, Infit and Outfit MNSQ (Mean Square; 제곱 평균) 및 남녀 집단 간의 DIF (Differential Item Function; 차별 문항 기능)를 분석하여 문항의 타당도를 분석하였다. 그리고 SPSS 26.0을 이용하여 탐색 요인 분석(프로맥스, 주축 요인 분석)을 실시하여 의도한 하위 요인에 4문항씩 적재되는지, 그리고 Factor loading 값이 .3 이상으로 나오는지 확인하였다. 그리고 시스템 사고 검사 도구의 공인 타당도 검증을 위해 STS 척도와와의 상관 분석을 실시하여 상관 계수를 확인하였다. 마지막으로 검사 도구를 이루는 요인 구조 모형과 모형 적합도 확인을 위하여 확인적 요인 분석을 실시하였으며, 두 집단 간 평균비교에서 t-test보다 오차를 엄격하게 통제하여 통계적으로 유의미한 효과 크기를 산출하기 위하여 우리나라 고등학생들의 데이터를 함께 활용한 잠재평균분석

을 실시하였다. 또한 잠재평균분석을 위하여 모형에 제약을 가하면서 도출되는 모형 적합도 지수( $\chi^2/df$ , CFI, TLI, SRMR, RMSEA)를 통해 베트남 고등학생과 우리나라 고등학생 두 집단에 대한 동일성을 검증할 수 있다(형태, 메트릭, 절편, 요인 분산 동일성). 그리고 우리나라 고등학생의 평균을 0으로 한 잠재평균 도출 및 효과 크기 비교를 통해 두 집단 간 차이를 분석함으로써 Re\_STMI의 재타당화와 우리나라와 베트남 고등학생 간 시스템 사고 비교 분석을 실시하였다.

**III. 연구 결과**

**1. 문항 반응 분석**

베트남 고등학생이 응답한 Re\_STMI 문항에 대한 문항 반응 분석 결과(Item Measure, Item map, 남녀 학생의 집단 간 DIF)는 Table 3-4와 같다. Rating Scale model에서 문항의 내적합도 평균지수인 Infit MNSQ, 외적합도 평균 지수인 Outfit MNSQ은 응답자가 얼마나 일관되게 응답하였는지를 살펴보는 지수로, 기준은 선행 연구에서 제시된 .6-1.4의 범위로 설

**Table 3.** Result of Rating Scale Model

Item	Mean	S.D.	Infit MNSQ	Outfit MNSQ	Item Reliability
20	917.5	81.6	1.00	1.05	.97

**Table 4.** Result of Rating Scale Model \_ Item #

Item #	Measure	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
MM1	0.53	1.12	1.36	1.17	1.72
MM2	-0.06	0.86	-1.60	0.91	-0.82
MM3	0.30	1.12	1.31	1.10	1.02
MM4	0.26	1.06	0.68	1.21	1.97
PM1	-0.61	1.02	0.28	1.07	0.59
PM2	-0.36	1.05	0.50	1.37	2.87
PM3	-0.88	1.05	0.49	1.20	1.35
PM4	-0.71	1.00	0.07	1.01	0.15
TL1	-0.29	0.85	-1.61	0.78	-2.08
TL2	-0.19	0.79	-2.42	0.73	-2.60
TL3	1.35	1.38	4.03	1.37	3.84
TL4	-0.40	1.24	2.32	1.21	1.72
SA1	0.04	0.69	-3.74	0.67	-3.51
SA2	0.11	0.76	-2.88	0.95	-0.43
SA3	0.25	0.78	-2.64	0.74	-2.86
SA4	0.10	0.67	-4.08	0.65	-3.85
SV1	0.25	1.16	1.74	1.24	2.26
SV2	-0.01	1.05	0.54	1.08	0.77
SV3	-0.18	1.17	1.72	1.19	1.62
SV4	0.49	1.17	1.83	1.25	2.40

정하였다(Ha, 2016; Lee et al., 2024; Wright & Linacre, 1994).

Re\_STMI 20문항에 대한 Infit과 Outfit의 MNSQ의 값은 .1.00, 1.05로 나타났으며 Item Reliability는 .97로 높은 신뢰도와 타당도 값이 도출되었다. 각 문항에 대한 Infit, Outfit MNSQ를 살펴본 결과, .67-1.38로 나타났으며 Infit의 경우 0.69-1.38로 나타났다. 여러 선행 연구에서 제시된 바와 같이 Rasch 모델에서는 보수적으로 산출되는 Infit 값에 가중치를 두고 해석하는 것으로 볼 때(Boone et al., 2011; Boone et al., 2014; Lee, 2019; Lee et al., 2024), fit 값은 기준치에 적합한 것으로 분석되었다.

문항 반응 분석에서 도출된 Item Map은 Fig. 2와 같다. Re\_STMI의 경우 평균 값을 중심으로 정규 분포로 문항들이 분포하는 것을 확인할 수 있으며, 고등학생 응답자의 평균인 1.00을 중심으로 Measure 값이 높은 학생과 낮은 학생들이 대체로 정규 분포

로 도출된 것을 확인할 수 있다. 개정된 시스템 사고 검사 도구는 5점 리커트식 척도의 자기 보고식 설문 문항이므로 선행 연구에서와 같이 문항들이 정규 분포의 형태로 도출되는 것이 타당하다고 볼 수 있다(Boone et al., 2011; Boone et al., 2014; Lee et al., 2024). 따라서 베트남 고등학생들을 대상으로도 개정된 시스템 사고 검사 도구는 문항 반응 분석에서도 선행 연구에서 도출된 결과와 유사한 결과가 도출된 것을 확인할 수 있다. 그리고 선행 연구에서 실시한 분석을 참고하여 추가적인 타당도 검증을 위하여 베트남 고등학생 내 남학생과 여학생 집단에서 문항이 동일한 기능을 하였는지를 살펴보고자 DIF 값을 비교하였으며(Lee, 2019; Lee et al., 2024), 그 결과는 Table 5, Fig. 3과 같다.

DIF는 서로 다른 두 집단에서 동일한 평가 문항을 얼마나 편파적으로 응답하는지를 살펴보는 값으로, 베트남 고등학생 중 남학생과 여학생 집단에서 DIF

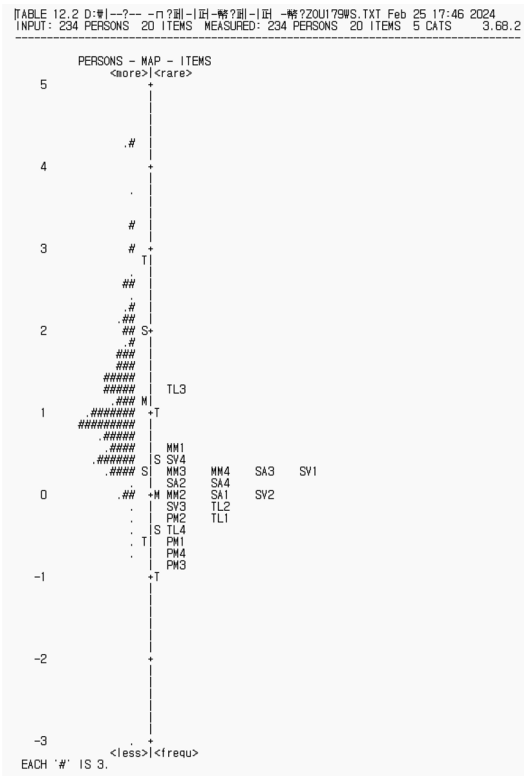


Fig. 2. Item map

를 분석하였다. 분석 결과, 20개 문항 중 두 집단 간 가장 큰 차이를 보여준 문항이 PM3로 .45의 차이가 있었으며, 그 외 문항에서도 두 집단 간 차이가 .00-.43으로 기준치인 .68 이상의 차이가 나타나지 않았다(Lee, 2019; Lee et al., 2024). 따라서 Re\_STMI 20문항에서는 베트남 고등학생 중 남학생과 여학생 두 집단이 문항을 같은 의미로 해석하고 응답하였다는 것을 알 수 있다(Boone & Scantlebury, 2006; Boone et al., 2011; Boone et al., 2014).

Table 5. Result of DIF

Gender	MM1	MM2	MM3	MM4	PM1	PM2	PM3	PM4	TL1	TL2
male	0.33	-0.24	0.3	0.14	-0.61	-0.31	-0.62	-0.48	-0.13	-0.07
female	0.67	0.07	0.3	0.34	-0.61	-0.39	-1.07	-0.88	-0.4	-0.28
Gender	TL3	TL4	SA1	SA2	SA3	SA4	SV1	SV2	SV3	SV4
male	1.3	-0.48	-0.22	-0.02	0.28	0.10	0.39	0.11	-0.01	0.45
female	1.39	-0.34	0.21	0.21	0.25	0.10	0.14	-0.09	-0.3	0.52

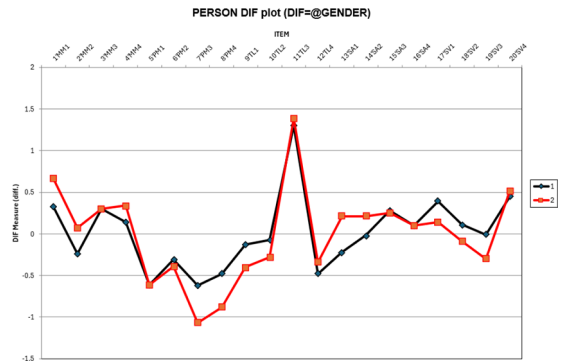


Fig. 3. DIF ('1'= male, '2'=female)

## 2. 탐색적 요인 분석

베트남 고등학생들을 대상으로 실시한 탐색적 요인 분석 결과 및 신뢰도 분석 결과는 Table 6과 같다. 우리나라 고등학생들을 대상으로 타당도를 재검증한 연구에서와 같이 탐색적 요인 분석은 SPSS 26.0에서 프로맥스를 이용한 주축 요인 분석을 실시하였다(Lee et al., 2024). 초기 고유값 1을 기준으로 한 요인의 수는 5개로 도출되었으며, 각 요인별 회전 제공합 적재값은 3.047-4.297로 나타났다. 그리고 패턴 행렬과 구조 행렬에서 Factor loading 값이 타당하게 도출되었다.

탐색적 요인 분석 결과를 살펴보면, 각 요인별 하위 문항에서 모두 .30 이상의 Factor loading 값이 도출된 것을 확인할 수 있었다. 또한 하위 요인 간 상관관은 .275-.550으로 도출되었기에 요인 분석에서 요인 간 회전을 사교회전(promax)으로 설정한 부분은 적절하였다고 볼 수 있다. 하위 요인별 신뢰도는 .662 - .833으로 .6 이상의 값이 도출되었으며, 20문항의 전체 신뢰도는 .876으로 분석되어 탐색적 요인 분석에서는 신뢰도와 타당도가 검증되었다고 볼 수 있다.



**Table 6.** Result: exploratory factor analysis: Revised STMI

STMI	Item (Factor #)	Result of EFA	
		Factor loading	Cronbach's- $\alpha$
Team Learning	1	.748	.737
	2	.656	
	3	.554	
	4	.540	
Systems Analysis	1	.822	.750
	2	.772	
	3	.683	
	4	.567	
Personal Mastery	1	.659	.727
	2	.643	
	3	.627	
	4	.587	
Shared Vision	1	.737	.833
	2	.590	
	3	.403	
	4	.318	
Mental Model	1	.970	.662
	2	.630	
	3	.444	
	4	.365	

Unweighted Least Square: Factor 1, 4.297; Factor 2, 3.290; Factor 3, 4.028; Factor 4, 3.047; Factor 5, 3.322

### 3. STS 검사지와와의 상관 분석 결과

Re\_STMI와 STS 검사지 모두 Senge(1996, 2012)가 제시한 학습 조직의 5가지 수련(팀 학습, 개인 숙련, 정신 모델, 시스템 분석, 공유 비전)을 이론적 배경으로 하여 학생들의 시스템 사고를 측정하는 검사지로 개발되었기 때문에, 상관 분석에서 높은 상관이 도출될 것을 기대할 수 있었으며, Lee et al.(2023), Lee et al.(2024)에서는 이와 일치하는 결과가 도출되

었다. 이 연구에서도 선행 연구와 같이 SPSS 26.0에서 두 검사지 간 상관 분석을 실시하였으며 그 결과는 Table 7와 같다.

상관 분석 결과, Re\_STMI와 STS 간 상관은 .725로 유의수준 95%에서 유의미한 값이 도출되었으며, 상관 계수도 .7 이상 매우 높은 정적 상관이 나타났다. 그리고 STS와 STMI의 하위 요인 간 상관에서도 .430-.657로 높은 정적 상관 계수가 도출되었다. 따라

**Table 7.** Result: Correlation analysis between Revised STMI and STS

	MM	PM	TL	SA	SV	Re_STMI	STS
MM	1						
PM	.298**	1					
TL	.428**	.332**	1				
SA	.457**	.427**	.523**	1			
SV	.361**	.353**	.465**	.446**	1		
Re_STMI	.712**	.636**	.767**	.791**	.723**	1	
STS	.450**	.430**	.486**	.610**	.657**	.725**	1

\*\* $p < .01$

서 베트남 고등학생들을 대상으로 개정 시스템 사고 검사 도구(Revised STMI)의 공인 타당도를 분석한 결과 선행 연구에서와 같이 유의미한 상관이 도출된 것을 확인할 수 있다(Lee et al., 2023; Lee et al., 2024).

4. 확인적 요인 분석 및 잠재평균비교

1) 확인적 요인 분석 결과

베트남 고등학생과 우리나라 고등학생들의 잠재평균비교를 위하여 Lee et al.(2013), Lee et al.(2024)에서 설정한 Fig. 4의 구조 방정식 모형을 활용하였다. Re\_STMI의 확인적 요인 분석 결과는 Table 8과 같다.

Table 8에는 Re\_STMI에서 베트남 고등학생과 우리나라 고등학생들을 대상으로 확인적 요인 분석 결과를 확인한 모형 적합도를 제시하였다. 각 모형 적합도 수치를 확인해 보면,  $\chi^2/df$ 는 2.830(1-3), CFI는 .931(> .90), TLI는 .918(> .90), SRMR은 .043(< .080), RMSEA는 .051(< .080)으로 나타나 모형 적합도 수치가 양호한 것으로 나타났다. 따라서 베트남 고등학생과 우리나라 고등학생들을 대상으로 한 검사에서도 개정된 시스템 사고 검사 도구의 모형 적합도는 안정적으로 도출되었으며, 확인적 요인 분석 결과를 토대로 적절한 문항과 요인으로 구성되었다는 것을 확인할 수 있다.

2) 베트남 고등학생과 우리나라 고등학생 간 잠재평균비교

구조 방정식 모형에 제약을 가하여 오차를 통제하면서 추가적인 타당도를 검증하면서 나아가 베트남 고등학생과 우리나라 고등학생 간 시스템 사고 능력의 차이를 알아보기 위하여 잠재평균비교를 실시하였다. 잠재평균비교를 위하여 구조 방정식 모형에 제약

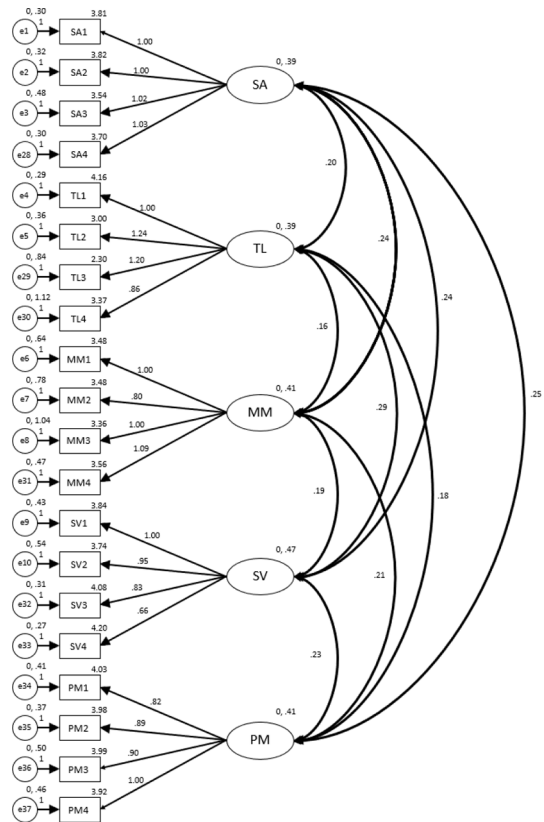


Fig. 4. Structural equation models for confirmatory factor analysis

을 가하여 형태 동일성, 측정 동일성, 절편 동일성, 요인 분산 동일성에 대한 검증을 실시하였으며 그 결과는 Table 9과 같다.

베트남 고등학생과 우리나라 고등학생에 대하여 모든 잠재변인 간 상관관계를 허용하면서 모수 측정을 자유롭게 추정한 기저 모형(형태 동일성)과 베트남과 우리나라 고등학생 각 집단에 동일하게 제약을 가한

Table 8. Result of model fit

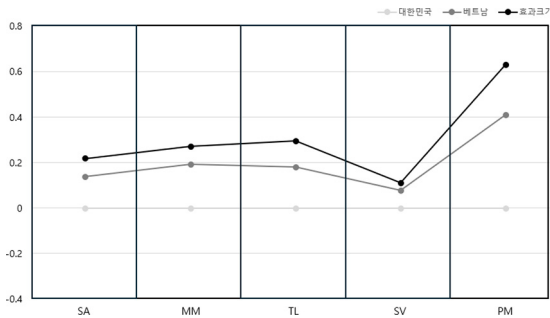
	$\chi^2$	df	$\chi^2/df$ (1-3)	CFI (> .9)	TLI (> .9)	SRMR (< .08)	RMSEA (less .08)
Model	452.80	160	2.830	.931	.918	.043	.051

Table 9. Result of model fit on identify verification

	$\chi^2$	df	$\chi^2/df$	CFI	TLI	RMSEA
Configural Invariance	478.88	250	1.916	.941	.928	.036
Metric Invariance	508.13	263	1.932	.937	.927	.036
Structural Invariance	719.66	281	2.561	.898	.898	.047
Equal factor variances/covariances	730.59	286	2.555	.896	.898	.047

**Table 10.** Result of latent mean analysis

		Korean	Vietnamese	S.D.	Cohen's <i>d</i>
Systems Thinking	Systems Analysis	0	.138	.63	.22
	Mental Model	0	.193	.71	.27
	Team Learning	0	.180	.60	.29
	Shared Vision	0	.076	.68	.11
	Personal Mastery	0	.411	.65	.63

**Fig. 5.** Result of latent mean analysis and Cohen's *d*

측정 동일성 모형에서는 적합도 수치가 모두 양호하게 나타났다. 다만 절편 동일성과 요인 분산 동일성에서 CFI와 TLI의 수치가 다소 낮게 나타났으나 잠재평균비교와 관련된 여러 선행 연구(Park & Bae, 2016; Jo, 2019; Lee, 2022)의 연구 결과를 근거로 하였을 때 RMSEA나  $\chi^2/df$  등 다른 모형 적합도 수치와의 관련성을 토대로 잠재평균비교를 실시하여 그 결과를 비교하는 것이 타당하다고 판단하여 분석을 진행하였다.

이 연구에서는 우리나라 고등학생의 잠재 평균을 0으로 가정하고 측정된 잠재평균과 공통의 표준편차를 활용하여 Cohen's *d* 값을 산출하였으며, 그 결과는 Table 10, Fig. 5와 같다.

베트남 고등학생과 우리나라 고등학생 간 잠재평균 비교를 실시한 결과, 우리나라 고등학생들에 비하여 이 연구에 참여한 베트남 고등학생들의 잠재평균이 5개 하위 요인에서 모두 높게 나타났다. 두 집단 간 차이를 살펴보는 Cohen's *d* 값을 하위 요인별로 살펴본 결과 시스템 분석, 공유 비전 요인에서는 작은 값(비중첩 비율; 15% 이하)으로 나타났으며, 정신 모델, 팀 학습에서는 중간 정도(비중첩 비율; 15-21%)의 차이가 있음을 확인할 수 있었다. 개인 숙련에서의 차이가 다른 요인들보다 크게 나타났는데, Cohen's *d*가 .63으로 .8보다는 작은 값으로 나타났지만, 비중

첩 비율이 38% 이상으로 다른 하위 요인에 비해서는 큰 차이가 도출되었다(Cohen, 2013).

#### IV. 결론 및 제언

이 연구는 Lee et al.(2023)에서 우리나라 고등학생들을 대상으로 타당화한 '개정 시스템 사고 검사 도구(Re\_STMI)'를 베트남어로 번역한 후 베트남 고등학생에게 투입하여 타당도를 재검증하였다. 우리나라와 베트남 고등학생 간 시스템 사고 결과를 잠재평균비교를 통해 분석하였다. 검사 도구의 타당도 재검증은 문항 반응 분석, 탐색적 요인 분석, 확인적 요인 분석, STS 검사와의 상관 분석을 통해 다양한 적합도 수치를 살펴보았다. 그리고 우리나라와 베트남 고등학생 간 시스템 사고 능력 비교를 위해, 잠재평균비교를 실시하는 과정에서 구조 방정식 모형에 제약을 가함으로써 모형 적합도 수치의 변화를 통해 검사지의 타당도 확인 및 두 집단 간 차이를 검증하였다. 연구의 결과를 요약하면 다음과 같다.

첫째, 문항 반응 분석을 통해 타당도를 살펴본 결과, Item Measure, Item map, 성별에 따른 DIF 등에서 타당한 결과가 도출되었다. 문항 반응 분석은 Rating Scale model로 이루어졌으며, 전체 20문항에 대한 Infit MNSQ는 1.00, Outfit MNSQ는 1.05, Item Reliability는 .97로 분석되었다. 그리고 각 문항의 Infit과 Outfit 값을 살펴본 결과 20개 문항 모두 .6-1.5 사이의 값이 도출된 것을 확인할 수 있었다(Boone et al., 2011; Boone et al., 2014; Lee, 2019; Lee et al., 2023). 그리고 Item map에서 나타난 응답한 학생들의 분포와 문항 분포는 Lee(2019), Lee et al.(2024)의 선행 연구에서 나타난 분포와 유사한 분포를 보여주었으며, 평균을 중심으로 학생과 문항 분포가 정규 분포를 보여주고 있음을 확인할 수 있었다. 성별에 따라 문항에 응답하는 차이를 보는 DIF 분석에서도 두 집단 간 차이가 .00-.43으로 나타나

다른 두 집단에서도 같은 의미로 문항을 해석하고 있다는 것을 확인하였다. 시스템 사고 검사 도구를 문항 반응 분석으로 분석한 선행 연구와의 결과와 비교해 볼 때(Lee, 2019; Lee et al. 2024), 문항 반응 분석에서 타당도를 검증하는 적합도 수치나, 베트남 고등학생들이 반응한 문항 분포 등이 기존 연구 결과와 유사하게 도출되었다. 베트남 고등학생들에게 2013년에 개발된 검사 도구를 활용한 Lee et al. (2023)의 연구에서는 문항 반응 분석을 실시하지 않아 기존 검사 도구와 개정된 검사 도구 간 문항 분석 결과가를 직접적으로 비교하기는 어렵지만, Lee et al.(2024)에서 국내 고등학생들에게 개정 검사 도구를 활용한 결과와 비교하면 국내 고등학생과 베트남 고등학생들에게서 나타나는 문항 반응 분석 결과가 일관된 것을 확인할 수 있다. 그리고 분석된 적합도 수치들 또한 허용 가능한 범위 내에서 결과가 도출되었기에 문항 반응 분포를 통해 살펴본 베트남어 버전의 Re\_STMI 문항은 타당하다고 판단할 수 있다.

둘째, 베트남 고등학생들의 자료로 사교회전(promax)을 이용한 주축 요인 분석을 실시한 결과, 20문항의 개정된 시스템 사고 검사 도구가 4문항씩 5개 요인으로 분석되었다. 신뢰도는 각 요인별 Cronbach's- $\alpha$ 가 .662-.833, 20개 문항의 Cronbach's- $\alpha$ 는 .876으로 분석되었다. 각 문항의 Factor loading 값은 모두 .3 이상으로 도출되었다. 이 연구를 포함하여 Re\_STMI의 탐색적 요인 분석을 실시한 결과는 개정 전 시스템 사고 검사 도구에 대한 타당도를 검증한 기존의 선행 연구(Lee et al., 2013; Lee & Lee, 2013; Lee & Lee, 2016; Lee et al., 2019) 등의 연구 결과와 비교해 볼 때, 각 문항들이 의도한 요인에 높은 Factor loading 값과 회전 제곱합 적재값을 보여주며 안정적으로 적재되는 것을 확인할 수 있으며, 높은 신뢰도를 보여주는 것으로 보아 개정된 문항의 내용이 의도한 요인의 정의를 잘 반영하고 있고 나아가 설문에 응답하는 학생들도 문항의 의도를 잘 이해하도록 번역되었음을 확인할 수 있다. 그리고 2013년에 개발된 검사지를 활용한 Lee et al.(2023)의 결과에서는 탐색적 요인 분석 결과 일부 문항의 타당도가 낮거나 해당 요인에 적재되지 못하여 확인적 요인 분석에서도 20문항 모두 활용하지 못하였다는 한계점이 제기되었다. 본 연구에서는 Re\_STMI를 활용하여 베트남 고등학생들에게 다시 타당도를 검증하였으며, 탐색적 요인 분석 결과 20문항 모두 타당한 결과가

도출되었으며, 이를 토대로 국내 고등학생과 비교할 수 있도록 신뢰도와 타당도를 검증하였다는 것에 의의가 있다.

셋째, Re\_STMI의 공인 타당도 검증을 위하여 STS 검사지와 상관 분석을 실시한 결과, 두 척도 간 상관은 .725로 도출되었으며, 각 하위 요인과 STS 척도 간 상관도 .4 이상으로 높은 상관을 보여주었다. 시스템 사고 검사 도구와 STS 검사지 간 공인 타당도 검증을 실시한 선행 연구에서의 결과(Lee et al., 2023; Lee et al., 2024)와 비교해 볼 때, 이 연구의 결과에서도 .7 이상의 상관이 분석된 것으로 볼 때, Re\_STMI 문항들도 시스템 사고를 잘 반영한 문항들도 구성되어 있다는 것을 확인할 수 있다.

넷째, 구조 방정식 모형을 활용하여 베트남 고등학생과 우리나라 고등학생 간 잠재평균비교를 실시하며 검사도구의 타당도를 살펴보았다. 구조 방정식 모형의 확인적 요인 분석과 잠재평균비교를 위해 구조 방정식 모형에 제약을 가한 모형에서도 적합도 수치가 수용 가능한 수치로 도출되었으며, 여러 모형에서 적합도 수치의 변화가 크지 않고 안정적으로 도출되는 것을 확인할 수 있었다(Park & Bae, 2016; Jo, 2019; Lee, 2022). 마지막으로 베트남 고등학생과 우리나라 고등학생 간 잠재평균비교 결과 5개 하위 요인 모두에서 베트남 고등학생들이 우리나라 고등학생보다 높은 평균 값을 보여주었고, Cohen's  $d$ 로 그 차이를 살펴본 결과 개인 숙련 요인을 제외하고는 작은 차이가 나타났으며, 개인 숙련에서 중간 이상의 차이가 나타났음을 알 수 있었다(Cohen, 2013).

본 연구의 결론에서 다음과 같은 한계점과 함께 제언을 제시할 수 있다. 첫째, 베트남 학생과 우리나라 학생들 간 성취도 비교 연구나 다른 인지적 검사 도구의 비교 결과가 충분하지 않아 이 연구의 결과가 도출된 원인을 구체적으로 설명하는 데 한계가 있다. 즉, 본 연구에서는 베트남 고등학생을 표집하는 과정에서 교육 봉사를 실시한 1개 고등학교 학생들을 대상으로 목적 표집을 실시하였기에 연구 대상이 베트남 학생 모집단을 대표하는 것으로 보기에 어려움이 있다. 그러므로 이 연구의 결과로부터 베트남 학생이 우수하다는 일반적인 결론을 도출하기에는 연구의 제한점이 따르기에, 이 연구는 개정된 시스템 사고 검사 도구의 재검증에 초점을 맞추어 해석할 필요가 있다.

둘째, 타당화가 이루어진 시스템 사고 검사 도구의

경우 Re\_STMI와 STS가 있으며, 2개의 검사 도구 모두 국어, 영어, 베트남어의 3개 국어로 번안되어 활용되고 있다. 따라서 시스템 사고를 측정하는 연구를 진행하기 위한 다국어의 검사 도구가 확보되었기에, 고등학생들의 고등 사고 능력을 살펴보는 국제 연구에서 시스템 사고 능력도 측정할 수 있을 것으로 사료된다.

셋째, 국가 간 성취도 비교가 가능한 경우 성취도 검사와 함께 시스템 사고 검사 문항도 함께 검사가 이루어진다면 성취도와 시스템 사고 능력 간 관련성을 파악할 수 있으며, 나아가 국가 간 성취도와 시스템 사고 능력 간 관련성뿐만 아니라 성취도와 시스템 사고에 영향을 미치는 요인을 분석하는 연구가 진행될 수 있을 것으로 판단된다. 학생들의 국제 비교와 관련된 선행 연구를 살펴볼 때, PISA 자료로부터 성취도와 성취와 관련된 예측 변수들을 탐색하는 연구 등이 수행되고 있음을 확인할 수 있다(Kim & Yoo, 2022; Jo, 2011). 이러한 예측 변수들은 학생들의 학업 문제나 성취도 향상과 관련하여 학교 현장에서 활용되는 중요한 교수-학습 전략에 반영되는 요소들이다. 따라서 시스템 사고 능력도 학업 성취도와 관련성, 나아가 다른 학습 요인들 간의 관련성을 분석하는 연구를 통해 학생들의 효과적인 학습 전략을 발달시키는데 기여할 수 있을 것으로 판단된다.

마지막으로 이 연구의 결론에서 도출된 제한점을 보완하기 위한 후속 연구도 이루어질 필요성이 있다. 최근 ODA 사업 등 다양한 루트를 통해 베트남 고등학교 혹은 대학교와의 교류가 이루어지고 있다(Lee et al., 2023). 이 연구에서는 참여한 베트남 고등학교 학생들의 경우 우리나라 일반계 고등학생들보다 우수한 시스템 사고를 하고 있다는 결과가 도출되었으나, 두 모집단에서 같은 결과가 도출될 것이라는 결론을 도출하기는 어렵다. 그러므로 추후 베트남 고등학생들의 성취도와 우리나라 고등학생들의 성취도, 시스템 사고 등 다양한 학습적 요소나 사고 능력에 대한 데이터를 통한 비교 연구가 진행될 필요가 있을 것이다.

## 참고문헌

- Aguilar-Cisneros, J. R., Valerdi, R., and Sullivan, B. P., 2022, Students' Systems Thinking Competency Level Detection through Software Cost Estimation Concept Modeling. *Programming & Computer Software*, 48(8), 499-512.
- Batt, A. M., Williams, B., Brydges, M., Leyenaar, M., and Tavares, W., 2021, New ways of seeing: supplementing existing competency framework development guidelines with systems thinking. *Advances in Health Sciences Education*, 26(4), 1355-1371.
- Ben-zvi-Assaraf, O., and Orion, N., 2005a, Development of System Thinking Skills in the Context of Earth System Education. *Journal of Research in Science Teaching*, 42(5), 518-560.
- Ben-zvi-Assaraf, O., and Orion, N., 2005b, A Study of Junior High Students' Perceptions of the water cycle. *Journal of Geoscience Education*, 53(4), 366-373.
- Ben-Zvi-Assaraf, O., and Orion, N., 2009, A Design Based Research of an Earth Systems Based Environmental Curriculum. *Eurasia Journal of Mathematics, Science & Technology Education*, 5(1), 47-62.
- Boone, W., and Scantlebury, K., 2006, The Role of Rasch Analysis When Conducting Science Education Research Utilizing Multiple-Choice Tests. *Science Education*, 90, 253-269.
- Boone, W., Staver, J. and Yale, M., 2014, *Rasch Analysis in the Human Sciences*. Springer: Switzerland.
- Boone, W., Townsend, J., and Staver, J., 2011, Using Rasch Theory to Guide the Practice of Survey Development and Survey Data Analysis in Science Education and to Inform Science Reform Efforts: An Exemplar Utilizing STEBI Self-Efficacy Data. *Science Education*, 95, 258-280.
- Cho, K., and Kim, H., 2017, Effect of Systems Thinking Based STEAM Education Program on Climate Change Topics. *Journal of Digital Contents Society*, 17(7), 113-123.
- Cho, O., and Hwang, K., 2016, The effects of simulation-based education on nursing students' presence in education, systems thinking and proactivity in problem solving. *Journal of Korean Academic Society of Home Health Care Nursing*, 23(2), 147-154.
- Cohen, J., 2013, *Statistical power analysis for the behavioral sciences*. Academic press.
- Davis A., and Stroink, M., 2016, The Relationship between Systems Thinking and the New Ecological Paradigm. *Systems Research and Behavioral Science*, 33, 575-586.
- Dolansky, M. A., Moore, S. M., Palmieri, P. A., and Singh, M. K., 2020, Development and validation of the systems thinking scale. *Journal of General Internal Medicine*, 35, 2314-2320.
- Ha, M., 2016, Examining the Validity of History-of-Science-Based Evolution Concept Assessment and Exploring Conceptual Progressions by Contexts. *Journal of the Korean Association for Science Education*, 36(3), 509-517.
- Han, D., and Jo, E., 2015, A study on the exploration of

- relationship between environmental literacy and systems thinking for sustainability education in social studies. *Social Studies Education*, 54, 65-83.
- Hung, W., 2008, Enhancing systems thinking skills with modelling. *British Journal of Educational Technology*, 39(6), 1099-1120.
- Im, Y., and Lee, H., 2014, Development and analysis of effects of writing educational program for improving system thinking ability. *Journal of Learner-Centered Curriculum and Instruction*, 14, 407-427.
- Jeon, J., and Lee, H., 2015, The development and application of STEAM education program based on systems thinking for high school students. *Journal of the Korean Association for Science Education*, 35(6), 1007-1018.
- Jo, J., 2019, Test of Construct Equivalence of the Career Indecision Inventory for Korean University Students and the Comparison of Latent Means between Male and Female Students. *Journal of Education Culture*, 25(2), 339-357.
- Jo, S., 2011, A Study of Factors Effecting on Gifted Students' Achievement: Self-determination, Learning Goal-orientation, Self-efficacy, Implicit Theory of Intelligence, and Self-regulated Learning Strategy. *Journal of Gifted/Talented Education*, 21(3), 611-630.
- Kali, Y., Orion, N., and Eylon, B., 2003, Effect of Knowledge Integration Activities on Students' Perception of the Earth's Crust as a Cyclic System. *Journal of Research in Science Teaching*, 40(6), 545-565.
- Kang, C., Lee, H., Yoon, I., and Kim, E., 2008, Analysis of conceptions related to Earth system and systems-thinking of high school student about water cycle. *Journal of Science Education*, 32(1), 61-72.
- Kim, M., and Yoo, J., 2022, Exploration of Predictors for Male and Female Students' Reading Literacy: Evidence from PISA 2018 via the Elastic Net. *Asian Journal of Education*, 23(4), 787-812.
- Kowch, E. G., 2019, Introduction to Systems Thinking and Change. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology* (pp. 1-14). Springer International Publishing. [https://doi.org/10.1007/978-3-319-17727-4\\_133-1](https://doi.org/10.1007/978-3-319-17727-4_133-1).
- Kwon, Y., Kim, W., Lee, H., Byun, J., and Lee, I., 2011, Analysis of Biology Teachers' Systems Thinking about Ecosystem. *Biology Education*, 39(4), 529-543.
- Lee, H., Lee, H., and Oh, H., 2020, Path Analysis Study among Science Motivation, Science Self-Efficacy and System Thinking of Student related Medical Major. *The Journal of Learner-Centered Curriculum and Instruction*, 20(18), 1193-1209.
- Lee, H., and Lee, H., 2013, Revalidation of measuring instrument systems thinking and comparison of systems thinking between science and general high school students. *Journal of the Korean Association for Science Education*, 33(4), 1237-1247.
- Lee, H., 2019, Validity Verification of Systems Thinking Measuring Instrument of Elementary School Students' and Middle School Students': Using Item Response Analysis. *Secondary Education Research*, 67(2), 249-277.
- Lee, H., 2020, An Analysis of Systems Thinking Levels of the Water Cycle in Earth System of Primary School Pre-Service Teachers. *Teacher Education Research*, 59(1), 21-38.
- Lee, H., 2022, Validation of Science Self-Efficacy Scale for Pre-Service Teachers and Latent Mean Analysis According to Background Variable. *Journal of Korean Elementary Science Education*, 41(1), 65-78.
- Lee, H., and Lee, H., 2016, Effects of Systems Thinking on High School Students' Science Self-Efficacy. *The Journal of The Korean Earth Science Society*, 37(3), 133-145.
- Lee, H., and Lee, H., 2017, Analysis and Effects of High School Students' Systems Thinking Using Iceberg (IB) Model. *Journal of the Korean Association for Science Education*, 37(4), 611-624.
- Lee, H., Jeon, J., and Lee, H., 2018, Development of Framework and Rubric for Measuring Students' Level of Systems Thinking. *Journal of the Korean Association for Science Education*, 38(3), 355-367.
- Lee, H., Jeon, J., and Lee, H., 2019, Verification the Systems Thinking Factor Structure and Comparison of Systems Thinking Based on Preferred Subjects about Elementary School Students'. *Journal of the Korean Association for Science Education*, 39(2), 161-171.
- Lee, H., Kwon, H., Park, K., and Lee, H., 2013, An instrument development and validation for measuring high school students' systems thinking. *Journal of the Korean Association for Science Education*, 33(3), 995-1006.
- Lee, H., Kwon, Y., Oh, H., and Lee, H., 2011, Development and application of the educational program to increase high school students' systems thinking skills: Focus on global warming. *Journal of the Korean Earth Science Society*, 32, 784-797.
- Lee, H., Nguyen, T., Lee, H., Jeon, J., and Park, B., 2023, The Validation of the Systems Thinking Assessment Tool for Measuring the Higher-order Thinking Ability of Vietnamese High School Students. *The Journal of The Korean Earth Science Society*, 44(4), 318-330.
- Lee, H., Park, B., Jeon, J., and Lee, H., 2024, Validation of the Revised Systems Thinking Measuring Instrument (Re\_STMI) for High School Students. *Brain, Digital, & Learning*, 14(1), 1-19.
- Ministry of Education [MOE], 2022, 2022 revised curriculum: Science. Sejong: Ministry of Education.

- Momsen, J., Speth, E. B., Wyse, S., and Long, T., 2022, Using Systems and Systems Thinking to Unify Biology Education. *CBE—Life Sciences Education*, 21(2), es3.
- Moon, B., Jeong, J., Kyung, J., Koh, Y., Youn, S., Kim, H., and Oh, K., 2004, Related conceptions to earth system and applying of systems thinking about carbon cycle of the preservice teachers. *Journal of the Korean Earth Science Society*, 25, 684-696.
- Park, B., and Bae, S., 2016, Testing Measurement Invariance and Latent Mean Analysis of the Internet Game Addiction Scale of Adolescents Depending on Gender and School stage. *Journal of Youth Welfare*, 18(2), 65-84.
- Park, K., Lee, H., Lee, H., and Jeon, J., 2019, Analysis of Systems Thinking Level of Pre-service Teachers about Carbon Cycle in Earth Systems using Rubrics of Evaluating Systems Thinking. *Journal of the Korean Association for Science Education*, 39(5), 599-611.
- Park, M., and Choi, D., 2018, The Effect of Simulation Integrated with Problem Based Learning on System Thinking, Learning Flow, Proactivity in Problem Solving and Performance Ability for Medication in Nursing Students. *Journal of Digital Convergence*, 16(8), 221-231.
- Ratinen, I., Linnanen, L., Claudelin, A., and Halonen, V., 2023, Toward Sustainable Development: Connecting Systems Thinking Competency and Carbon Footprint Knowledge. *Sustainable Development*, 31(3), 1593-1605.
- Rust, J., and Golombok, S., 2014, *Modern psychometrics: The science of psychological assessment*. Routledge.
- Senge, P. M., 1996, *The fifth discipline: Fieldbook*. New York: Broadway Business.
- Senge, P. M., 2012, *Schools that learn (Updated and Revised): A fifth discipline fieldbook for educators, parents, and everyone who cares about education*. New York: Doubleday.
- Song, J., Moon, B., and Kim, J., 2015, The Development and Application of the Teaching-Learning Program for Systems Thinking Learning in Elementary Science Classes. *Journal of Korean Society of Earth Science Education*, 8(3), 318-331.
- Ulku, S. B., and Gaye, D. C., 2024, Research Trends on Systems Thinking Approach in Science Education. *International Journal of Science Education*, 46(5), 485-502.
- Wright, B., and Linacre, J., 1994, Reasonable mean-square-fit values. *Rasch Measurement Transactions*, 8(3), 370.

---

Manuscript received: April 4, 2024

Revised manuscript received: April 18, 2024

Manuscript accepted: April 23, 2024