

RESEARCH ARTICLE

Validation of the Revised Systems Thinking Measuring Instrument(Re_STMI) for High School Students

Hyundong Lee¹, Byung-Yeol Park², Jaedon Jeon³, Hyonyong Lee^{3*}¹Daegu National University of Education²Dankook University³Kyungpook National University

고등학생들을 대상으로 한 개정 시스템 사고 검사 도구(Re_STMI)의 타당도 검증

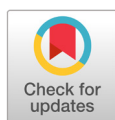
이현동¹, 박병열², 전재돈³, 이효녕^{3*}¹대구교육대학교, ²단국대학교, ³경북대학교

*Corresponding Author: Hyonyong Lee, hlee@knu.ac.kr(Corresponding Author's e-mail)

ABSTRACT

The purpose of this study is to verify the validity of the revised STMI, which is a revised and supplemented version of the Systems Thinking Measuring Instrument (STMI) developed by Lee et al. (2013) for high school students. For this purpose, we used data from 475 high school students' responses to 20 Revised STMI questions and 20 STS scale questions. Validity verification includes item response analysis (Item Reliability, Item Map, Infit and Outfit MNSQ, and DIF of male and female groups), exploratory factor analysis (main axis factor analysis using Promax), and confirmatory factor analysis using structural equation model. And central/cross validity verification and correlation analysis between the two scales were used. The research results are as follows. First, as a result of the item response analysis of the Revised STMI 20 questions, Item Reliability was .98 and Infit MNSQ was .84~1.39, and valid results were also analyzed in Item Map and DIF analysis. Second, in the exploratory factor analysis, all sub-questions of the five sub-factors showed factor loading values of .3 or higher, and the reliability of each factor was .631~.800 and the overall reliability was .863. Third, in the confirmatory factor analysis, all model fit values were good (χ^2/df : 2.406, CFI: .923, TLI: .908, SRMR: .046, RMSEA: .054), and the convergent validity and cross-validity were also established. The results showed that the model was suitable. Lastly, in the correlation analysis with the 20 items of the STS scale, the correlation between all items was .754, and the correlation with STS items for each sub-factor was high, ranging from .426 to .658. Therefore, in the case of the Revised STMI, significant results were obtained not only in construct validity but also in official validity, and the results of this study showed improved validity results compared to research results using existing scales. Continuously, additional validation research should be conducted so that the Revised STMI can be applied to various groups, including not only high school students, but also elementary and middle school students, and college students. In addition, ongoing research on the relationship between system thinking and other learning-related measures should be actively conducted using validated instrument.

Key words: Re_STMI(Revised Systems Thinking Measuring Instrument), STS(Systems Thinking Scale), item response analysis, factor analysis, validation



OPEN ACCESS

Brain, Digital, & Learning

2024, Vol. 14, No. 1, 1-19

<https://doi.org/10.31216/BDL.20240001>**Received:** March 08, 2024**Revised:** March 19, 2024**Accepted:** March 20, 2024© 2024. 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

시스템 사고(Systems thinking)는 고등 사고 능력 중 하나로 시스템과 그 시스템을 이루고 있는 하위 시스템(하위 요소)에 대한 인식, 그리고 시스템 내에서의 상호작용, 시스템 내 역동성이나 숨겨진 차원 등에 대한 이해를 바탕으로 상위 시스템의 현재 모습과 함께 앞으로의 변화를 예측할 수 있는 사고 방식을 의미한다(Kowch, 2019; Lee et al., 2018; Maani & Maharaj, 2004; O'Connor & McDermmot, 1997). 이러한 시스템 사고는 국내외의 과학 교육에도 반영되어, 우리나라의 경우 7차 교육과정의 '지구계'에 그 사고 과정이 반영된 것을 시작으로 2015 및 2022 개정 과학과 교육과정에서도 시스템과 상호작용이 하나의 단원으로 제시되고 있으며 하위 성취 기준에 지구 시스템, 역학 시스템, 생명 시스템 등을 이해하는데 시스템 사고가 활용된다(MOE, 2015, 2022; Park et al, 2019).

국외의 사례에서는 미국의 경우 국가과학교육표준(NGSS)에 포함된 Dimension 2의 Crosscutting Concepts에서 과학과 기술, 공학을 활용하는데 시스템과 시스템 사고를 포함한 모델을 제시하고 있으며(NRC, 2012), 뉴욕주 등 여러 주에서 이루어지는 과학 교육에 반영된 Learning Standards for Math, Science, and Technology (NYSED, 2017)의 Standard 6에서는 다양한 주제 내에서 상호 작용을 이해하는 사고 방식으로 시스템 사고를 제시하며 전체와 하위 요소들 간의 상호 작용을 강조하고 있다.

이와 같이 국내외 과학 교육에서 전체와 부분들 사이의 상호 작용 및 이를 바탕으로 한 시스템 전체를 이해하는 사고 능력으로 시스템 사고가 강조되면서 시스템 사고를 향상시키기 위한 다양한 교육 프로그램 개발 및 적용 연구(Ben-zvi Assaraf & Orion, 2009; Kali et al., 2003; Kang et al., 2008; Kwon et al., 2011; Lee et al., 2011; Lee & Lee, 2017; Momsen et al., 2022; Moon & Kim, 2007; Park et al., 2019)가 이루어져 왔으며, 최근에는 학생들의 시스템 사고를 측정하기 위한 평가 도구와 측정을 위한 평가 척도에 대한 연구도 이루어져 왔다(Lee et al., 2013; Lee et al., 2018; Nguyen, & Santagata, 2021).

학생들의 시스템 사고의 변화를 살펴보는 연구 중 국내에서 가장 잘 알려지고, 인용되는 연구로는 Ben-zvi Assaraf & Orion (2005a, 2005b, 2009)가 있다. 물 순환에 대한 주제를 활용하여 인과 지도, 단어 간 상관, 그림 그리기 등 다양한 정성적 분석 도구들을 활용하여 물 순환 주제 활동 전과 후 학생들의 시스템 사고 변화를 제시하였으며, 이 연구에서 활용한 도구들은 학생들의 정성적 변화를 판단하는데 있어 최근 연구에서도 활용되고 있다(Lee, 2020; Jeon et al, 2022; Ratinen et al., 2023).

정성적 도구로 이루어지던 시스템 사고 연구에서, Lee et al. (2013)에서는 학생들의 시스템 사고 변화를 해석하는 데 있어 연구자의 주관 개입이나 학생들의 인지적인 변화가 미치는 영향 등이 정성적 도구만으로 분석하는데 한계점을 제시하고, 정성적 도구와 함께 정량적인 해석이 가능한 시스템 사고 검사 도구(Systems Thinking Measuring Instrument; STMI)를 개발하였다. STMI는 Senge (2006, 2012)가 제시한 Fifth discipline을 이론적 배경으로 하여 5가지 요인(시스템 분석, 정신 모델, 개인 숙련, 팀 학습, 공유 비전)에서 각 요인 별 4문항씩 20문항으로 이루어진 검사지로 요인 분석, 구조 방정식 모형을 통한 검증을 통해 타당도를 검증하였다.

검사 도구가 개발된 이후, Lee & Lee (2016)에서는 과학적 자기 효능감 척도와의 경로 분석 등을 통해 다른 척도와의 인과 관계 검증을 통한 타당도 및 신뢰도를 검증, Han & Jo (2015), Jeon & Lee (2015), Cho & Hwang (2016), Park & Choi (2018) 등 과학 교육뿐 아니라 간호학, 교육학 등 다양한 분야에서 고등학생 외 여러 집단을 대상으로 시스템 사고 검사 도구가 활용되었다. 그러나 시스템 사고 검사 도구와 비교 가능한 다른 도구

가 개발되지 않아 공인 타당도를 측정할 수 없다는 부분, 그리고 고등학생 외 다른 집단에서는 요인 구조가 5 요인으로 분명하게 드러나지 않는다는 한계점이 지속적으로 제기되었다(Lee & Lee, 2016; Lee et al., 2019).

최근에 들어와 시스템 사고를 정량적으로 측정하고자 하는 방향은 국외 연구에서도 이루어지고 있다. Davis & Stroink (2016), Roslan et al. (2021)에서는 교육학 및 심리학 분야에서 시스템 사고를 측정 가능한 구인으로 설정하고, 이를 측정하는 문항을 개발하기도 하였으며, Dolansky et al. (2020)의 연구에서 STS (Systems Thinking Scale)라 명명된 시스템 사고를 측정하는 20문항의 검사지를 보고하였다. 이 연구에서 주목할만한 점은 STS의 문항을 개발하는 이론적 배경으로 Senge (2012)의 Fifth discipline을 활용하였다는 것인데, 이는 국내 Lee et al. (2013)의 STMI 검사지와 동일한 이론적 배경을 가지고 개발한 것이기 때문에 시스템 사고 검사 도구 간 공인 타당도 검증을 실시할 수 있는 척도로 확인할 수 있다. 최근 실시된 Lee et al. (2023)의 연구에서는 구조 방정식 모형에서 기존에 개발된 STMI와 STS 검사 도구 간 상관 분석을 실시한 결과, 두 척도 간 상관이 .903으로 매우 높은 상관을 나눴다는 결과가 보고되기도 하였다.

나아가 최근에 수행된 다양한 집단에서 시스템 사고를 분석한 여러 선행 연구(Lee et al., 2019; Park et al., 2019; Lee, 2020)의 결과에서 STMI를 구성하는 일부 문항의 타당도가 낮게 도출되었다는 것을 확인할 수 있었다. 이로 인하여 초등학생이나 대학생을 대상으로 한 연구에서 20문항 전체에서의 구인 타당도 검증이 이루어지지 않았다는 한계점을 확인할 수 있었으며, 나아가 STMI가 개발되고 약 10년이 지나가면서 문항 개정의 필요성이 제시되었다(Lee et al., 2023).

STMI를 활용한 최근 연구들에서 검사지의 신뢰도는 적절한 값이 보고되고 있지만, 탐색적 요인 분석 과정에서 일부 문항이 의도한 요인이 아닌 다른 요인에 적재되거나, 2개의 요인이 하나의 요인으로 합쳐져 분석되는 등의 결과가 보고되며 시스템 사고 검사 도구의 개정 필요성이 제기되었다(Lee et al., 2019; Lee, 2020; Jeon et al., 2022).

따라서 Lee et al. (2023) 등의 연구 결과를 바탕으로 본 연구에서는 시스템 사고 검사 도구의 개발자 및 관련 전문가들과 함께 그 이유를 파악하고, 문항의 개정을 제안하고자 하였다. 학생들이 응답함에 있어서 문항의 표현 방식을 수정하여 전달되는 의도를 명확히 반영되지 않을 경우, 검사 참여자들이 문항의 의도를 다르게 해석할 수 있기 때문에(Bak & Chung, 2012; Park & Jo, 2017), 개발된 검사 도구 내 문항의 의도가 달라지지 않는 부분에서 사회·문화적 맥락에 적합한 표현을 반영한 개정은 필요하다고 볼 수 있다.

또한 이러한 검사 도구의 개정을 통해 후속 연구들이 지속적으로 진행되는 과정에서 시스템 사고 혹은 고등 능력 사고 측정의 신뢰도를 높이는 데 도움이 될 수 있다. 따라서 이 연구에서는 Lee et al. (2013)에서 개발한 시스템 사고 검사 도구의 문항을 수정·보완한 개정 시스템 사고 검사 도구(Revised STMI)에 대한 타당도 검증을 실시하였다. 이 연구 목표를 위해 설정한 연구 문제는 다음과 같다.

첫째, 개정된 시스템 사고 검사 도구의 문항 반응 분석 결과는 타당한가?

둘째, 개정된 시스템 사고 검사 도구의 요인분석 및 수렴·교차 타당도 결과는 타당한가?

셋째, 개정된 시스템 사고 검사 도구와 STS(Systems Thinking Scale) 도구 간 상관관계는 유의미한가?

Method

Process of Study

이 연구의 과정은 Fig. 1과 같다. Lee et al. (2013)의 시스템 사고 검사 도구의 개정을 위하여 과학교육과 시스템 사고를 전공한 전문가 집단 3인이 기존의 시스템 사고 검사 도구 문항을 검토하고, 수정·보완 과정을 거쳐 시스템 사고 검사 도구의 개정 문항을 도출하였다. 그리고 이 검사 도구의 공인 타당도 검증을 위하여 Dolansky et al.(2020)에서 개발한 시스템 사고 척도(STS) 문항을 활용하였다.

검사 도구의 타당도 분석을 위한 연구 대상으로는 국내 고등학생 475명으로부터 검사 결과를 수집하였다. 수집된 데이터는 Winsteps, SPSS 26.0과 Amos 26.0을 활용하여 문항 반응 분석, 신뢰도 분석, 탐색적 요인 분석, 확인적 요인 분석을 통해 시스템 사고 검사 도구 개정 문항의 신뢰도와 구인 타당도를 검증하였으며, STS 검사 도구와의 상관 분석을 통해 공인 타당도를 검증함으로써 시스템 사고 검사 도구 개정 문항의 재타당화 결과를 도출하였다.

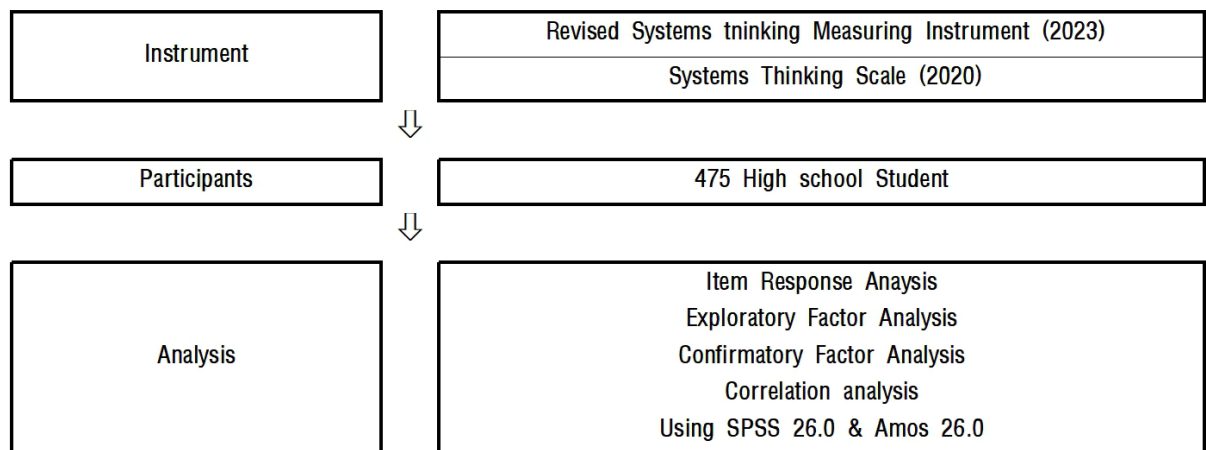


Fig. 1. Process of study

Participants

고등학생들을 대상으로 시스템 사고 검사 도구 개정 문항의 타당화를 위하여 광역시 및 중소도시 소재 일반계 고등학생 1~2학년을 대상으로 총 523명을 표집하였다. 고등학교 집단에 대하여 광역시 소재 고등학교 2곳과 중소도시 소재 1곳의 군집을 추출한 뒤, 각 학교에서 무선 표집을 통해 학생을 선정하고 구글 설문을 이용하여 응답하도록 하였다. 응답 전 학생들에게 데이터 수집의 목적과 활용 여부에 대한 안내를 하였으며, 학부모와 해당 학교 교사 및 학생의 동의 후 설문에 응답하도록 안내하였다. 그리고 응답한 결과로부터 무성 의하게 응답한 결과나 목종 경향성을 보이는 응답을 제외한 475명의 데이터로부터 결과를 도출하였다. 475명의 데이터 중 남학생은 204명으로 42.9%, 여학생은 271명으로 57.1%에 해당하였다.

Instruments

Revised Systems Thinking Measuring Instrument

이 연구에서 타당화하고자 한 검사 도구는 Lee et al. (2013)에서 개발한 시스템 사고 검사 도구의 문항을 수정·보완한 시스템 사고 검사 도구 개정 문항이다. 이 도구는 Senge (2012)의 학습 조직의 5가지 수련을 바탕으로 5개 요인 20문항(정신 모델 4문항, 개인 숙련 4문항, 시스템 분석 4문항, 공유 비전 4문항, 팀 학습 4문항)으로 구성되었으며, 문항에 대한 응답은 리커트 5단계로 이루어진 검사 도구이다. 각 요인에 대한 정의, 기존 문항과 개정 문항의 예는 Table 1과 같다.

문항의 개정에는 과학교육 및 시스템 사고를 전공한 전문가 3인이 여러 차례 회의를 거쳐 문항을 수정·보완하였다. 문항 개정과 관련된 전문가 회의에서, 문항의 일부 개정 또는 전면 개정에 대하여 논의한 결과, 문항의 내용 타당도나 이론적 배경이 검증되었으며, 신뢰도가 검증되었기 때문에 전면 개정 보다는 일부 개정으로 논의가 이루어졌다. 따라서 기존 선행 연구에서 나타난 문항의 중복이나 요인이 중복되어 나타나는 문제점 등 검사 도구에서 수정이 필요한 부분을 공유한 뒤, 먼저 전문가 각자가 수정이 부분을 찾아 의견을 제시하였다. 이후, 여러 번의 대면 협의에서 학습 조직의 5가지 수련에 대한 정의와 이론적 배경을 토대로 문항 간 의미의 중복성, 요인 간 간섭을 줄이는 방향으로 문항 내 단어나 문항의 표현 내용을 수정하였다.

특히 문항의 표현을 국내외 여러 가지 척도의 표현 방식을 참고하여, 도입 문장을 요인별 특징이 드러나는 용어로 통일하여 시작하도록 수정하였다. 그리고 ‘비판적인 시각’, ‘목표 달성의 결과’ 등 고등학생뿐 아니라 선행 연구의 다른 학교급 학생들에게서 어렵게 느껴졌던 표현을 ‘다양한 시각’, ‘목표의 성취가’ 등의 쉽게 읽히는 표현으로 수정하는 등 검사 도구를 읽고 해석하는데 있어 쉽게 읽고 해석할 수 있는 방향으로 문항을 개정하였다. 개정된 20문항의 전체 신뢰도는 본 연구에서 .863, 하위 요인별로 .631~.800의 Cronbach- α 값이 도출되었다.

Table 1. Example of Re_STMI items

Sub Factor	Definition of Sub factor	STMI(Lee et al., 2013)	Re_STMI
Mental Model	Changes in thinking that occur through constant interaction with perspective (TV, Internet, etc.) outside	I try to view newspapers or news from a critical perspective (TV, Internet, etc.).	When I usually think to myself, I try to view newspapers or news from a critical perspective (TV, Internet, etc.).
Personal Mastery	Change in the process of setting your own vision and moving toward achieving	When I set a vision, I consider how the results of vision achievement will affect me.	When I set goals or plans, I think about how achieving my goals will affect me.
Team Learning	Interaction between team members in a group and changes in the opinion. process	When I discuss topics, I actively announce my opinion.	When I do group learning (activity), I actively express my opinions.
Systems Analysis	Thinking ability to take into account different perspectives	When put in a difficult situation, I consider that this happens from the background.	When I face a problem situation, I consider as many different perspectives as possible.
Shared Vision	Formation of a group vision and leadership that a leader must possess	When I engage in team learning, I am able to accept the opinions of other team members.	When I am the leader of a group, I induce opinions from other members in group activities.

Systems Thinking Scale

Dolansky et al. (2020)의 STS 검사 도구 또한 학습 조직의 5가지 수련을 이론적 배경으로 하여 시스템 사고를 측정하고자 개발된 검사지이다. 이 검사지의 경우 단일 요인의 20문항으로 개발되었으며 STMI와 같이 리커트 5단계 척도로 이루어진 검사지이다(Table 2). STS 20문항의 전체 신뢰도는 본 연구에서 .905로 Cronbach- α 값이 도출되었다.

Table 2. Example of STS Items

STS	Items
Item 1	When I want to make an improvement, I seek everyone's view of the situation.
Item 2	When I want to make an improvement, I look beyond a specific event to determine the cause of the problem.
Item 3	When I want to make an improvement, I think understanding how the chain of events occurs is crucial.
Item 4	When I want to make an improvement, I include people in my work unit to find a solution.
Item 5	When I want to make an improvement, I think recurring patterns are more important than any one specific event.

Analysis

시스템 사고 검사 도구 개정 문항의 타당도 검증을 위하여 475명의 응답 결과를 Winsteps, SPSS 26.0과 Amos 21.0을 이용하여 분석하였다. 먼저 문항 반응 분석에서는 Rating Scale Model 방법을 적용하여 Item Reliability, Item Map, Infit and Outfit MNSQ 및 남녀 집단 간의 DIF를 분석하여 문항의 타당도를 살펴보았다. 그리고 SPSS 26.0을 이용하여 탐색 요인 분석을 실시하여 의도한 하위 요인에 4문항씩 적재되는지 확인하였다. 탐색적 요인 분석에는 4개의 하위 요인이 독립적이지 않다고 가정되었기 때문에 프로맥스(사교회전)를 이용한 주축 요인 분석을 실시하였으며, 문항의 타당도를 검증하기 위하여 Factor loading 값이 .3이상이거나 다른 요인의 Factor loading 값이 .3이상인지를 확인하였다.

또한 검사 도구를 이루는 요인 구조 모형과 모형 적합도 확인을 위하여 확인적 요인 분석을 실시하여 모형 적합도 지수(χ^2/df , CFI, TLI, SRMR, RMSEA)와 평균 분산 추출을 이용한 모형의 집중 타당도와 표본을 둘로 나누어 교차 타당도를 검증하였다. 그리고 개정된 검사 도구의 공인 타당도 검증을 위해 STS 검사 도구 간 상관 분석을 통한 수렴 타당도를 분석함으로써 개정된 시스템 사고 검사 도구(Re_STMI)의 타당도를 분석하였다.

Results

Item Response Analysis

개정된 시스템 사고 검사 도구의 문항 반응 분석은 Item Measure 값과 Item map, 남녀 학생의 집단 간 Differential Item Functioning(이하 DIF)를 확인하여 진행하였다. 문항의 Infit과 Outfit의 MNSQ 값의 기준은 Rating Scale model에서는 .6 ~ 1.4의 범위로 설정하였다(Ha, 2016; Ha et al., 2018; Wright & Linacre, 1994). 전체 20문항과 각 문항 별 분석 결과는 Table 3, 4와 같다.

Table 3. Result of rating scale model

Item	Mean	S.D.	Infit MNSQ	Outfit MNSQ	Item Reliability
20	1774.7	148.9	.98	1.01	.98

Table 4. Result of rating scale model _ item #

Item #	Measure	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
MM1	0.49	1.15	2.31	1.19	2.85
MM2	0.35	1.2	3.08	1.24	3.67
MM3	0.58	1.39	5.9	1.46	6.51
MM4	0.32	0.97	-0.4	1.01	0.15
PM1	-0.14	0.94	-0.67	0.94	-0.69
PM2	-0.35	0.91	-1.37	0.9	-1.51
PM3	-0.38	1.00	0.00	1.01	0.15
PM4	-0.23	0.92	-1.14	0.95	-0.78
TL1	-0.78	0.92	-1.18	0.89	-1.66
TL2	-0.17	0.95	-0.77	0.97	-0.51
TL3	1.08	1.21	3.22	1.27	4.08
TL4	0.59	1.36	5.36	1.53	7.42
SA1	-0.18	0.84	-2.48	0.85	-2.43
SA2	-0.18	0.86	-2.19	0.87	-2.08
SA3	0.25	0.98	-0.24	0.99	-0.07
SA4	-0.18	0.85	-2.39	0.86	-2.26
SV1	0.03	0.85	-2.35	0.85	-2.37
SV2	0.13	0.93	-1.04	0.98	-0.36
SV3	-0.43	0.85	-2.12	0.85	-2.23
SV4	-0.81	0.89	-1.71	0.86	-2.14

전체 20문항에 대한 Infit과 Outfit의 MNSQ의 값은 .1.00, 1.02로 나타났으며 Item Reliability는 .98로 높은 신뢰도와 타당도를 보여주었다. 각 문항에 대한 Infit과 Outfit의 MNSQ를 살펴본 결과, .84 ~ 1.53으로 나타났으며 Infit의 경우 0.84 ~ 1.39로 나타났다. Rasch 모델에서는 보수적으로 산출되는 Infit 값에 가중치를 두고 해석하는 것으로 볼 때(Boone et al., 2011; Boone et al., 2014; Lee, 2019), fit 값은 기준치에 적합하여 타당한 문항으로 볼 수 있다.

문항 반응 분석에서 도출된 Item Map은 Fig. 2와 같다. 20개 문항의 경우 평균을 중심으로 정규분포의 형태로 분포하는 것을 확인할 수 있었으며, 고등학생 응답자의 평균인 1.00을 중심으로 높은 수준의 학생과 낮은 수준의 학생들이 대체로 정규 분포하는 것을 확인할 수 있다. 시스템 사고 검사 도구의 경우 5점 리커트식 척도로 구성된 자기 보고식 설문 문항이므로 문항의 평균을 기준으로 정규 분포의 형태로 분석되는 것이 타당하다고 볼 수 있다(Boone et al., 2011; Boone et al., 2014). 따라서 20문항의 시스템 사고 검사 도구는 고등학생의 시스템 사고를 적절하게 측정하면서도 각 문항들의 곤란도는 유사하도록 구성된 것이라 볼 수 있다.

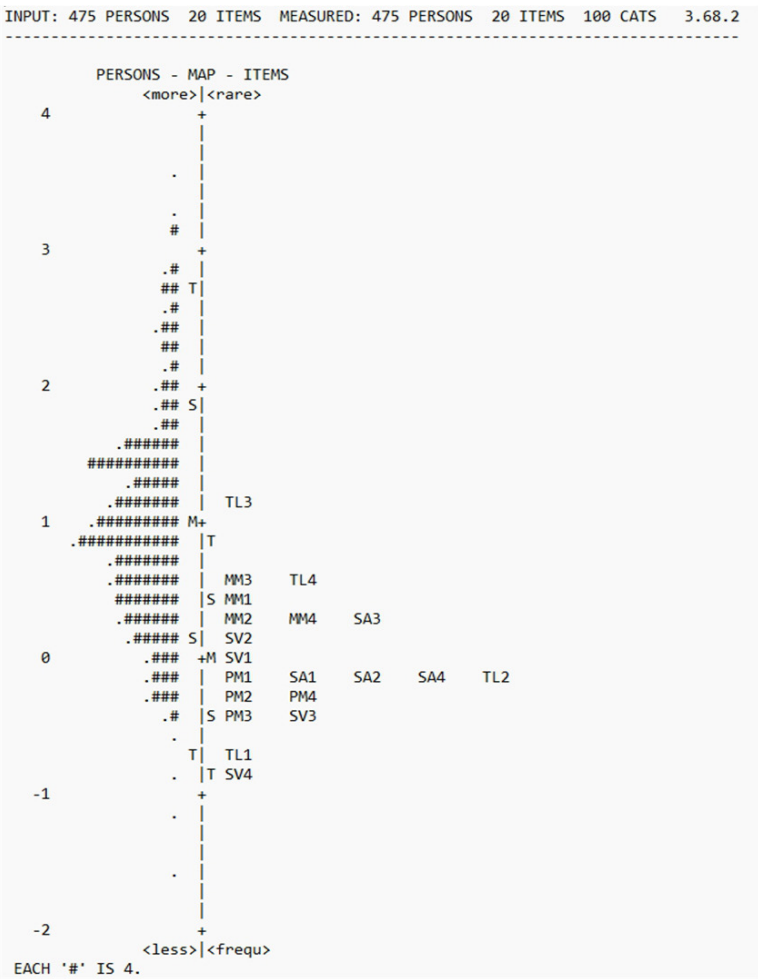


Fig. 2. Item map

문항 반응 분석에서 추가적인 타당도 검증을 위하여 고등학생 내 서로 다른 두 집단인 남학생과 여학생 집단에서 문항이 동일한 기능을 하였는지를 살펴보고자 DIF 값을 비교하였으며, 그 결과는 Table 5, Fig. 3과 같다.

Table 5. Result of DIF

Gender	MM1	MM2	MM3	MM4	PM1	PM2	PM3	PM4	TL1	TL2
male	0.42	0.46	0.55	0.20	-0.14	-0.40	-0.31	-0.17	-0.81	-0.17
female	0.55	0.26	0.58	0.39	-0.14	-0.31	-0.43	-0.27	-0.78	-0.17
Gender	TL3	TL4	SA1	SA2	SA3	SA4	SV1	SV2	SV3	SV4
male	1.32	0.53	-0.34	-0.18	0.13	-0.42	0.07	0.13	-0.38	-0.64
female	0.9	0.62	-0.07	-0.18	0.34	-0.01	0.00	0.13	-0.46	-0.94

고등학생 내 남학생과 여학생 집단에서 DIF를 분석한 결과, TL3에서 .42로 그 차이가 크게 나타났으나 대부분의 문항에서는 차이가 크게 나타나지 않았다. DIF의 경우 두 집단의 차이가 .68을 기준으로 판단하므로 개정된 시스템 사고 검사 도구 20문항에서는 남학생과 여학생 두 집단이 문항을 같은 의미로 해석하고 응답하였다는 것을 알 수 있다(Boone & Scantlebury, 2006; Boone et al., 2011; Boone et al., 2014).

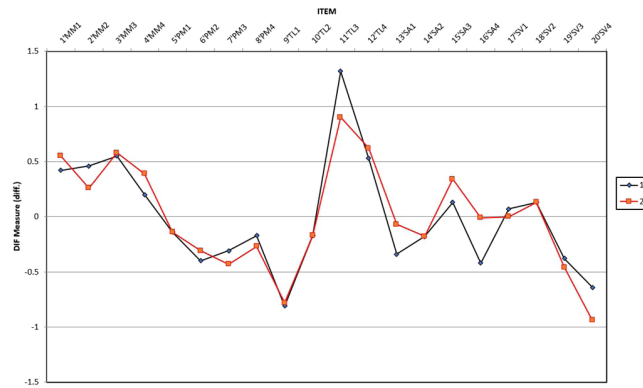


Fig. 3. DIF('1'= male, '2'=female)

Exploratory Factor Analysis

탐색적 요인 분석 결과 및 신뢰도 분석 결과는 Table 6과 같다. 5개 요인 구조의 확인을 위하여 SPSS 26.0에 서 사교회전(promax)를 이용한 공통 요인 분석(주축 요인 분석)을 실시하였다. 이는 직교회전, 주성분 분석 보다 오차를 정교하게 통제하는 분석 방법으로 검사지의 문항을 단순히 요인별로 묶어주는 효과에서 나아가 문항의 타당도를 정확하게 분석해주는 것으로 이 경우, Factor loading 값이 .30이상이면 타당한 문항으로 분석한다(Jung & Tak, 2023; Lim & Lee, 2020; Song, 2010). 초기 고유값 1을 기준으로 한 요인의 수는 5개로 도출되었으며, 각 요인별 회전 제공합 적재값은 2.661~4.060으로 나타났다. 그리고 패턴 행렬과 구조 행렬에서 Factor loading 값이 .3이하로 나타나거나, 다른 요인에 .3이상의 Factor loading 값을 보인 문항은 없었다.

Table 6. Result: exploratory factor analysis: Revised STMI

STMI	Item (Factor #)	Result of EFA		
		Factor loading	Cronbach's- α	
Team Learning	1	.826	.776	
	2	.773		
	3	.606		
	4	.384		
Systems Analysis	1	.775	.800	
	2	.757		
	3	.629		
	4	.606		
Personal Mastery	1	.705	.715	.863
	2	.630		
	3	.618		
	4	.488		
Shared Vision	1	.777	.737	
	2	.739		
	3	.451		
	4	.418		
Mental Model	1	.694	.631	
	2	.556		
	3	.465		
	4	.334		

Unweighted Least Square: Factor 1 – 3.494, Factor 2 – 4.060, Factor 3 – 3.696, Factor 4 – 3.556, Factor 5 – 2.661

요인 분석 결과를 살펴보면, 5개 요인의 하위 문항들이 모두 .3이상의 Factor loading 값이 도출됨으로써 보수적인 기준에서 부하량을 만족하고 있으며 하위 요인 간 상관은 .241~.516으로 나타났다. 따라서 탐색적 요인 분석의 요인 간 회전을 promax로 설정한 부분이 적절하였다고 볼 수 있다. 하위 요인별 신뢰도는 .631~.800으로 .6이상의 값이 도출되었으며, 20문항의 전체 신뢰도는 .863으로 나타났다. 사회과학이나 교육학 영역에서 측정 도구의 신뢰도는 Cronbach- α 가 .60 이상일 경우 신뢰롭다고 인정하고 있다(Mun & Kim, 2024; Song, 2010). 따라서 탐색적 요인 분석 결과에서는 신뢰도와 타당도가 검증되었다고 볼 수 있다.

Confirmatory Factor Analysis

개정된 시스템 사고 검사 도구의 요인 모형 및 모형 적합도 수치의 검증을 위하여 구조 방정식 모형을 이용한 확인적 요인 분석을 실시하였다. 분석을 위한 구조 방정식 모형은 Fig. 4와 같다.

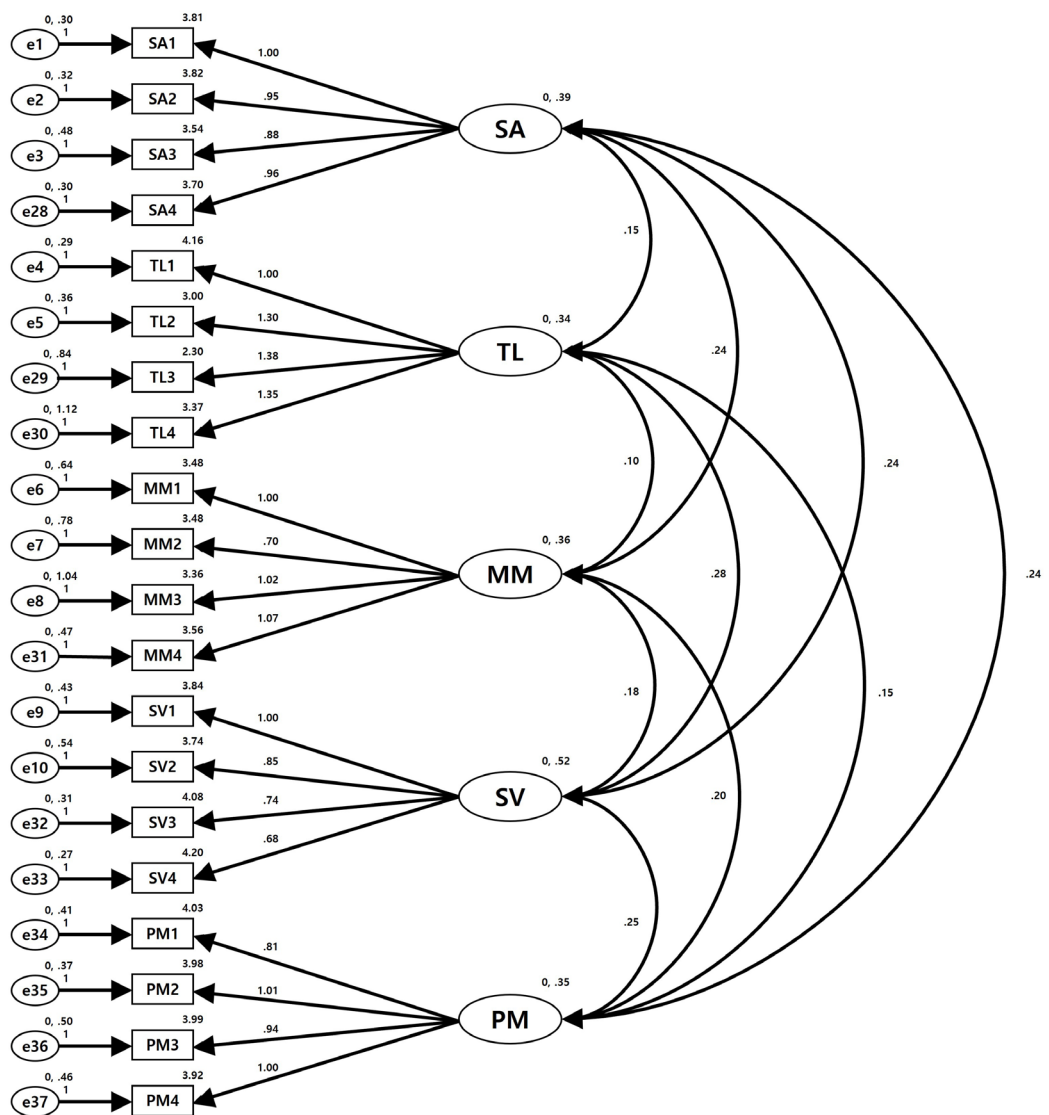


Fig. 4. Structural equation models for confirmatory factor analysis

구조 방정식 모형의 집중타당도와 모형 적합도 수치를 분석한 결과는 Table 7, 8과 같다.

Table 7. Results: convergent validity of the model

	Estimate	S.E.	C.R.	S.R.W.	AVE	Concept Reliability
MM → MM1	1.000			.599		
MM → MM2	.704	.099	7.123	.431	.915	.977
MM → MM3	1.017	.125	8.119	.513		
MM → MM4	1.067	.113	9.444	.680		
TL → TL1	1.000			.732		
TL → TL2	1.297	.093	13.942	.782	.945	.985
TL → TL3	1.382	.111	12.463	.660		
TL → TL4	.852	.104	8.230	.425		
SA → SA1	1.000			.752		
SA → SA2	.954	.066	14.431	.726	.946	.986
SA → SA3	.876	.071	12.412	.620		
SA → SA4	.956	.066	14.599	.736		
PM → PM1	.811	.078	10.447	.601		
PM → PM2	1.010	.087	11.629	.700	.933	.982
PM → PM3	.938	.088	10.672	.618		
PM → PM4	1.000			.656		
SV → SV1	1.000			.740		
SV → SV2	.855	.068	12.545	.645	.940	.984
SV → SV3	.742	.055	13.369	.692		
SV → SV4	.684	.051	13.306	.688		

구조 방정식 모형의 확인적 요인 분석에서 estimates, Standardized regression weight(S.R.W.), average variance extraction(AVE), concept reliability를 확인하였다. Standardized regression weight과 average variance extraction에서 .5이상 그리고 concept reliability가 .7 이상이면 집중타당도가 검증되는 것으로 볼 수 있다. Table 7를 보면, 여러 적합도 수치가 모두 기준값 이상의 값이 도출되었으므로 집중타당도로부터 추정된 결과 측정 모형이 적합하다고 판단할 수 있다.

Table 8. Result of model fit

	χ^2	df	χ^2/df (1~3)	CFI (>.9)	TLI (>.9)	SRMR (<.08)	RMSEA (less .08)
Model	385.01	160	2.406	.923	.908	.046	.054

Table 8에는 확인적 요인 분석 결과를 확인하기 위한 모형 적합도 수치를 제시하였다. 모형 적합도의 경우 χ^2/df 의 경우 1~3, CFI와 TLI는 .9이상, SRMR과 RMSEA는 .08이하일 경우 모형이 수용된다고 본다. 이 연구에서 제시한 모형의 경우 χ^2/df 는 2.406, CFI는 .923, TLI는 .908, SRMR은 .046, RMSEA는 .054으로 분석되어 모형 적합도 수치가 모두 양호한 것으로 나타났다. 따라서 개정된 시스템 사고 검사 도구의 모형 적합도는 적절한 것으로 도출되었으며, 탐색적 요인 분석 결과와 확인적 요인 분석 결과 모두에서 타당한 검사 도구라는 것을 확인할 수 있다.

Cross-validation Analysis

교차 타당도는 연구 모형이 다른 표본에서도 일반화될 수 있는지 정도를 검증하는 것으로(Bae, 2018), 이

연구에서는 이미 수집된 자료를 둘로 나누어 하나는 추정 표본, 다른 하나는 검증 표본으로 사용하여 분석하였다. 2개의 표본으로 나누는 방법은 무작위 추출법을 활용하였으며 연구 대상을 Group 1, 추출한 집단을 Group 2로 설정하고 다중집단분석을 실시하였다. 두 집단 간 동일성을 검증한 결과는 Table 9와 같다.

Table 9. Result of model fit on cross-validation analysis

	χ^2	df	χ^2/df	CFI	TLI	RMSEA
Model 1 (Measurement weights)	549.908	335	1.642	.928	.918	.036
Model 2 (Measurement intercepts)	573.778	355	1.616	.927	.921	.035
Model 3 (Structural covariances)	580.701	370	1.569	.929	.927	.034
Model 4 (Measurement residuals)	588.338	390	1.509	.933	.935	.032

Model 1은 측정 가중치(Measurement weight)가 두 집단에 걸쳐 동일하도록 설정한 모델로, 요인 계수가 동일한 모델이며, Model 2는 Model 1의 가정 외에 측정 절편(Measurement intercept)이 동일하도록 설정한 모델로 측정된 변수를 예측하기 위한 방정식의 절편을 동일하게 제한한 모델이다. Model 3은 Model 2의 가정 외에 구조 공분산(Structural covariances)이 동일하도록 가정한 것으로 요인분산 및 공분산이 동일하도록 설정한 것이며, Model 4는 Model 3의 가정 외에 측정 잔차(Measurement residuals)가 동일하다는 가정으로 측정 오차가 동일하도록 설정한 것이다.

Model 1~4까지 도출된 결과를 확인하였을 때 모두 안정적인 것으로 나타났다.

Table 10에서와 같이 모형 간 차이를 통해서 비제약모델과 제약을 가한 모델 간 모델 적합도 지수를 비교한 결과, 각 모형 간 적합도 지수에 차이가 없는 것으로 판단하였다(Cheung & Rensvold, 2002). χ^2/df 차이 검증 결과 Model 2와 Model 1의 $\Delta\chi^2/df$ 는 -.026($p > .05$)으로, Model 3와 Model 2의 $\Delta\chi^2/df$ 는 -.047($p > .05$)으로, Model 4와 Model 3의 $\Delta\chi^2/df$ 는 -.060($p > .05$)으로 모두 통계적으로 유의하지 않게 나타나 최종모델이 잘 재현된 것으로 나타났다.

또한 ΔCFI , ΔTLI , $\Delta RMSEA$ 모두 모형을 제약한 후에도 적합도가 크게 낮아지지 않은 것으로 나타나 집단 간의 교차 타당도가 확보되는 것을 확인할 수 있었다.

Table 10. Result of comparison between models

	$\Delta\chi^2$	Δdf	$\Delta\chi^2/df$	p	ΔCFI	ΔTLI	$\Delta RMSEA$
Model 2 vs. Model 1	23.87	20	-.026	.807	-.001	.003	-.001
Model 3 vs. Model 2	6.923	15	-.047	.952	.002	.006	-.001
Model 4 vs. Model 3	7.637	20	-.060	.997	.004	.008	-.002

Correlation Analysis

본 연구에서는 개정된 시스템 사고 검사지의 공인타당도 검증을 위하여 동일한 이론적 배경을 가지고 개발된 STS 검사지와 상관 분석을 실시하였다. 두 검사지 모두 시스템 사고를 측정하는 검사지로 학습 조직의 5가지 수련을 토대로 개발되었기에 상관 분석에서 높은 상관이 도출될 경우 공인 타당도가 검증되는 것으로 판단할 수 있다. 상관 분석을 위해 SPSS 26.0에서 상관 분석을 실시하였으며 그 결과는 Table 11과 같다.

상관 분석 결과, STMI와 STS 간 상관은 .754로 유의수준 95%에서 유의미한 값이 도출되었으며, 상관 계수도 .7이상 매우 높은 정적 상관이 나타났다. 그리고 STS와 STMI의 하위 요인 간 상관에서도 .426~.658까지 높은 정적 상관 계수가 도출되었으며 모든 값에서 유의수준 95%에서 유의미한 결과가 나타난 것을 확인할 수 있다. 따라서 고등학생들을 대상으로 시스템 사고를 측정하는 개정 시스템 사고 검사 도구(Re_STMI)의 구인 타당도뿐 아니라 공인 타당도에서도 유의미한 결과가 도출된 것을 확인할 수 있다.

Table 11. Result: correlation analysis between Revised STMI and STS

	STS	TL	SA	MM	PM	SV	Re_STMI
STS	1						
TL	.456**	1					
SA	.624**	.309**	1				
MM	.426**	.184**	.455**	1			
PM	.542**	.309**	.502**	.405**	1		
SV	.658**	.546**	.297**	.297**	.444**	1	
Re_STMI	.754**	.679**	.744**	.655**	.732**	.760**	1

Discussion and Conclusions

이 연구에서는 Lee et al. (2013)에서 개발한 시스템 사고 검사 도구의 문항을 수정하고, 고등학생을 대상으로 수정한 시스템 사고 검사 도구의 타당도를 분석하였다. 국내에서 개발된 Lee et al. (2013)에서 개발한 시스템 사고 검사 도구는 지난 10년 동안 여러 연구에서 연구 대상자들의 시스템 사고나 고등 사고 능력을 분석하는데 활용되어 왔다(Chung & Choi, 2018; Cho & Hwang, 2016; Han & Jo, 2015; Kim & Mun, 2017; Kim et al., 2017; Lee et al, 2019; Park et al., 2019). 그러나 이 검사 도구를 개발하는 과정에서 기존에 개발된 시스템 사고 측정 검사 도구나 척도가 없어 공인 타당도 검증을 실시할 수 없었으며, Lee et al. (2020)과 Jeon et al. (2022) 등 최근의 연구 결과에서 시스템 사고 검사 도구의 문항이 최근 사회·문화적 맥락 등을 반영하여 수정할 필요성이 점차 제기되어 왔다. 뿐만 아니라 국외의 Dolansky et al. (2020) 연구에서 동일한 이론적 배경을 토대로 시스템 사고 검사 도구가 개발되어 공인타당도를 분석할 수 있게 되어 시스템 사고 검사 도구 문항의 개정과 함께 재타당화를 진행하였다. 연구의 토의와 결론을 정리하면 다음과 같다.

Discussion

2000년대 이후부터 국내에서 시스템 사고를 적용한 교육 및 연구가 이루어진 이후 학생들의 시스템 사고를 알아보는 연구가 활발하게 이루어져 왔다(Im & Lee, 2014; Jeon & Lee, 2015; Lee et al., 2011; Lee et al., 2017). 이러한 연구들을 살펴보면 공통적으로 Ben-Zvi Assaraf & Orion (2005a, 2005b, 2009)에서 활용한 단어 간 상관, 인과 지도, 그림 그리기 등을 활용한 정성적 분석을 통해 학생들의 시스템 사고를 분석하였다. 이후, 정성적 검사 도구의 분석 결과를 보완하면서 학습자들의 특성을 정량적으로 살펴볼 수 있는 검사지 개발의 필요성이 제기된 후, 국내에서는 Lee et al. (2013)에서 시스템 사고 검사 도구(STMI)가 개발되었다. Senge (1996, 2006, 2012)가 제시한 학습 조직의 5가지 수련을 바탕으로 개발된 검사지는 학생들의 시스템 사고를 살펴보는 여러 연구에 활용되어 왔으며, 최근에는 Dolansky et al. (2020)에서도 Senge (2012)의 이론을 구인으로 설정한 시스템 사고 척도가 소개되며, 국내 시스템 사고 척도의 이론적 타당성을 확인할 수 있었다. 최근 국내 연구(Cho & Hwang, 2016; Han & Jo, 2015; Jeon & Lee, 2015)에서 STMI의 신뢰도는 양호하게 분석되었으나, 타

당도 검증 과정에서 일부 문항이 제외되거나 수정되어 사용되며, 추후 개정에 대한 필요성이 제기되어 왔다 (Lee et al., 2019).

따라서 본 연구에서는 이러한 선행 연구 결과 및 이론적 배경을 바탕으로 문항의 요인별 정의를 다시 확인하고, 문항을 최근의 사회맥락에 맞춰 수정·보완하였다. 이를 통해 학생들의 지속가능한 시스템 사고 연구가 가능하는 기반을 제시하였다는 것에 의의가 있다. 과학교육 분야에서 시스템 사고를 적용하려는 흐름은 국내뿐 아니라 국외에서도 지속적으로 연구가 되는 분야이다. 최근 지속가능한 발전과 관련하여 Momsen et al. (2022)과 Ratinen et al. (2023)에서는 생물교육에서의 시스템 사고 수준을 제시하거나 시스템 사고를 기반한 탄소 순환 교육에서 시스템 사고 능력을 측정하기도 하였다. 즉, 학생들의 지구시스템 내 순환이나 시스템 내 상호작용(생태계, 역학 등)을 학습하는 분야에서는 Ben-Zvi Assaraf & Orion (2009)이나 Aguilar-Cisneros et al. (2022)에 따른 시스템 사고 수준을 측정하고자 하는 연구가 수행되고 있으며, 이를 위해 적절한 검사 도구를 필요로 하는 것이다.

본 연구에서는 STMI 검사 도구 문항 개정과 함께 문항 타당도 분석 연구 방법으로 문항 반응 분석, 집중 및 교차 타당도 검증을 추가하여 Lee et al. (2013)의 연구보다 검사지의 타당도를 엄격하게 분석하였으며, 이와 함께 국외에서 개발된 STS 척도와와의 상관 분석을 통해 공인타당도를 검증하였다. 이를 통해 학생들의 시스템 사고를 분석하는 연구에 활용 가능한 정량적 검사 도구로 이 연구가 활용될 수 있을 것이라 기대한다.

게다가 2022 개정 교육과정에서도 과학 과목에서 시스템적 요소가 반영되어 있기에, 학생들은 시스템과 시스템 내 상호작용에 대한 이해에 바탕을 둔 학습을 할 필요성이 있다(MOE, 2022). 이러한 학습의 결과, 학생들의 고등 사고 능력 함양과 관련하여 타당화된 Re_STMI를 활용하면 학생들의 시스템 사고 역량과 다른 핵심 역량 및 기초 소양 함양과 관련성 연구도 진행될 수 있을 것으로 판단된다(Batt et al., 2021; Kowch, 2019; Ulku & Gaye, 2024).

끝으로 국외의 과학 교육 중 순환에 대한 개념 학습, 디지털 기반 교육, STEM 교육 등에서도 고등 사고 능력 함양을 강조하고 있으며, 특히 미국에서는 과학 교육 표준과 함께 제시된 STEM 교육에 대한 표준에서 시스템과 시스템을 고려하는 사고 능력인 시스템 사고 함양을 강조하고 있는 것을 확인할 수 있다(Clark et al., 2017; de Sousa et al., 2019; Emihovich, 2024; Ignacio et al., 2023; NRC, 2012; NYSED, 2017; Ratinen et al., 2023; Stavrianeas et al., 2022). 우리나라에서도 융합교육(STEM/STEAM)이 지속적으로 이루어지고 있기 때문에, 이러한 융합교육에서 함양하고자 하는 기초 소양이나 핵심 역량 함양(Kwon et al., 2021)과 시스템 사고와의 관련성을 함께 분석하는 연구도 진행될 필요성이 있을 것이다.

Conclusions

첫째, 시스템 사고 검사 도구 문항의 수정은 검사 도구가 개발될 당시 설정한 5요인 구조는 그대로 유지하며 각 하위 요인 별 정의와 이론적 배경을 토대로 국내 여러 척도의 표현 방식, 사회·문화적 맥락의 변화 등을 반영하여 수정·보완을 진행하였다. 이 과정에서 과학 교육 및 시스템 사고를 전공한 전문가 3인이 회의를 거쳤으며, 수정된 20 문항의 신뢰도는 Cronbach- α 값이 .863(각 하위 요인별 .631~.800)으로 도출되었다. 이 결과는 Lee et al. (2013)에서 이 검사 도구를 개발할 당시 신뢰도 .840(각 하위 요인별 .604~.723)과 유사하거나 조금 높은 수준으로 나타났다.

둘째, 본 연구에서 개정한 시스템 사고 검사 도구의 타당도 검증 결과, 문항 반응 분석의 Item Measure, Item

map, DIF 분석에서 모두 타당한 결과가 도출되었다. 각 문항의 Factor loading 값은 .3 이상으로 나타났으며 요인별 회전 제곱합 적재값은 2.661~4.060으로 도출되어 측정하고 하는 시스템 사고의 각 하위 요인들을 충분히 설명하는 검사지로 나타났으며, 검사지의 설명력도 Lee et al.(2013)보다 향상된 수치를 확인할 수 있었다.

셋째, 구조 방정식 모형을 이용한 확인적 요인분석, 수렴 타당도 교차 타당도 분석에서 모형 간 차이를 검증한 결과 모형 간 차이는 통계적으로 유의미하지 않아 연구에서 설정한 최종 모형의 타당도가 확보되는 것을 확인할 수 있었다.

넷째, 개정한 시스템 사고 검사 도구와 Dolansky et al. (2020)의 STS 20문항과의 상관 분석을 실시하여 공인 타당도를 검증한 결과 두 검사지 간 상관은 .754로 매우 높은 상관을 보여주었다. 그리고 각 하위 요인 간 상관에서도 .426~.658의 정적 상관 계수가 도출되었다.

이 결과는 Lee et al. (2023)에서 기존 시스템 사고 검사 도구 문항과 STS 간 상관을 분석한 연구 결과이며, 이를 토대로 본 연구에서 타당화한 개정 시스템 사고 검사 도구는 기존의 시스템 사고 검사 도구 및 STS 척도와 같이 고등학생들의 시스템 사고를 일관되고 타당하게 측정하는 검사지로 분석되었다.

References

- Aguilar-Cisneros, J. R., Valerdi, R., & Sullivan, B. P. (2022). Students' systems thinking competency level detection through software cost estimation concept modeling. *Programming & Computer Software*, 48, 499-512.
- Bae, B. (2018). *Amos 24 Advanced Structural Equation Modeling*. Seoul: Cheongram.
- Bak, B., & Chung, Y. (2012). Scale revalidation of epistemological beliefs and exploration of their domainality. *The Korean Journal of Educational Psychology*, 26, 739-767.
- Batt, A. M., Williams, B., Brydges, M., Leyenaar, M., & Tavares, W. (2021). New ways of seeing: Supplementing existing competency framework development guidelines with systems thinking. *Advances in Health Sciences Education*, 26, 1355-1371.
- Ben-zvi-Assaraf, O., & Orion, N. (2005a). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, 42, 518-560.
- Ben-zvi-Assaraf, O., & Orion, N. (2005b). A study of junior high students' perceptions of the water cycle. *Journal of Geoscience Education*, 53, 366-373.
- Ben-Zvi-Assaraf, O., & Orion, N. (2009). A design based research of an earth systems based environmental curriculum. *Eurasia Journal of Mathematics, Science & Technology Education*, 5, 47-62.
- Boone, W., & 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., & Yale, M. (2014). *Rasch Analysis in the Human Sciences*. Springer: Switzerland.
- Boone, W., Townsend, J., & 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.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255.
- Cho, O., & 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, 147-154.

- Chung, S., & Choi, N. (2018). The validity and norms of reading motivation scale for elementary school students. *The Korean Journal Child Education*, 27, 215-240.
- Clark, S., Petersen, J. E., Frantz, C.M., Roose, D., Ginn, J., & Rosenberg Daneri, D. (2017). Teaching systems thinking to 4th and 5th graders using environmental dashboard display technology. *PLoS ONE* 12, 1-11. <https://doi.org/10.1371/journal.pone.0176322>.
- Davis A., & Stroink, M. (2016). The relationship between systems thinking and the new ecological paradigm. *Systems Research and Behavioral Science*, 33, 575-586.
- de Sousa, L. O., Hay, E. A., & Liebenberg, D. (2019). Teachers' understanding of the interconnectedness of soil and climate change when developing a systems thinking concept map for teaching and learning. *International Research in Geographical & Environmental Education*, 28, 324-342.
- Dolansky, M. A., Moore, S. M., Palmieri, P. A., & Singh, M. K. (2020). Development and validation of the systems thinking scale. *Journal of General Internal Medicine*, 35, 2314-2320.
- Emihovich, B. (2024). Game-based learning and systems thinking: An innovative instructional approach for the 21st century. *TechTrends: Linking Research & Practice to Improve Learning*, 68, 174-185.
- 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, 509-517.
- Ha, M., Park, H., Kim, Y., Knag, N., Oh, P., Kim, M., ... Son, M. (2018). Developing and applying the questionnaire to measure science core competencies based on the 2015 revised national science curriculum. *Journal of the Korean Association for Science Education*, 38, 495-504.
- Han, D., & 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.
- Ignacio, J. T., Gonzales, C. K., & Lee-Chua, Q. N. (2023). Development of an e-learning module integrating systems thinking for climate change. *Ubiquitous Learning: An International Journal*, 16, 1-16. [doi:10.18848/1835-9795/CGP/v16i02/1-16](https://doi.org/10.18848/1835-9795/CGP/v16i02/1-16).
- Im, Y., & 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., & 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, 1007-1018.
- Jeon, J., Lee, H., & Lee, H. (2022). Development and application of instrument for level scale of the systems thinking ability about carbon cycle. *Journal of the Korean Association for Science Education*, 42, 397-415.
- Jung, J., & Tak, J. (2023). Development and validation of the Kkondae tendency scale. *Korean Journal of Coaching Psychology*, 7, 153-196.
- Kali, Y., Orion, N., & 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, 545-565.
- Kang, C., Lee, H., Yoon, I., & 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, 61-72.
- Kim, H., & Mun, S. (2017). The effect of a system thinking-based steam-type ESD program on environmental literacy and system thinking ability of elementary students. *Journal of Korean Journal of Environmental Education*, 30, 85-102.
- Kim, H., Jeong, S., Jeong, S., & Mun, S. (2017). Development of a systems thinking-based steam education program for elementary school students and the effect of its application. *School Science Journal*, 11, 288-301.

- 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.
- Kwon, H., Kim, E., Park, H., Bae, Y., Lee, D., Lee, H., Lee, H., Choi, S., & Ham, H. (2021). Current status of the implementation of convergence education in primary and secondary schools. *Journal of Science Education*, 45, 336-348.
- Kwon, Y., Kim, W., Lee, H., Byun, J., & Lee, I. (2011). Analysis of biology teachers' systems thinking about ecosystem. *Biology Education*, 39, 529-543.
- 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, 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, 21-38.
- Lee, H., & Lee, H. (2016). Effects of systems thinking on high school students' science self-efficacy. *The Journal of The Korean Earth Science Society*, 37, 133-145.
- Lee, H., & 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, 611-624.
- Lee, H., Jeon, J., & Lee, H. (2018). Development of framework and rubric for measuring students' level of systems thinking. *Journal of the Korean Association for Science Education*, 38, 355-367.
- Lee, H., Jeon, J., & 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, 161-171.
- Lee, H., Kwon, H., Park, K., & Lee, H. (2013). An instrument development and validation for measuring high school students' systems thinking. *Journal of the Korean Association for Science Education*, 33, 995-1006.
- Lee, H., Kwon, Y., Oh, H., & 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., Kwon, Y., Oh, H., & 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., & 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, 318-330.
- Lim, A., & Lee, H. (2020). Development and validation of the acceptance for differences in couple relationships scale based on IBCT. *Korean Journal of Psychology: General*, 39, 599-632.
- Maani, K. E., & Maharaj, V. (2004). Links between systems thinking and complex decision making. *System Dynamics Review*, 20, 21-48.
- Ministry of Education [MOE]. (2015). *Science Curriculum*. Notification No. 2015-74 [issue 9]. Sejong: Ministry of Education.
- Ministry of Education [MOE]. (2022). *2022 Revised Curriculum: Science*. Sejong: Ministry of Education.
- Momsen, J., Speth, E. B., Wyse, S., & Long, T. (2022). Using systems and systems thinking to unify biology education. *CBE—Life Sciences Education*, 21, es3.
- Moon, B., & Kim, H. (2007). A study on the abilities and characteristics of the systems thinking for pre-service elementary teachers. *Korean System Dynamics Review*, 8, 235-252.

- Mun, B., & Kim, S. (2024). Development and validation of a scale to measure of early childhood teachers climate literacy. *Journal of Eco Early Childhood Education & Care*, 23, 71-94.
- National Research Council [NRC] (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press.
- Next Generation Learning Standards, New York State Education Department [NYSED], (2017). <https://www.nysed.gov/curriculum-instruction>, 2023. 12. 12)
- Nguyen, H., & Santagata, R. (2021). Impact of computer modeling on learning and teaching systems thinking. *Journal of Research in Science Teaching*, 58, 661-688.
- O'Connor, J., & McDermmot, I. (1997). *The Art of Systems Thinking: Essential Skills for Creativity and Problem Solving*. London, UK: Thorsons Publishers.
- Park, K., Lee, H., Lee, H., & 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, 599-611.
- Park, M., & 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, 221-231.
- Park, S., & Jo, Y. (2017). Item analysis and modification of a parental empathy inventory perceived by adolescent children scale under multiple item response theory. *Korean Journal of Youth Studies*, 24, 1-21.
- Ratinen, I., Linnanen, L., Claudelin, A., & Halonen, V. (2023). Toward sustainable development: connecting systems thinking competency and carbon footprint knowledge. *Sustainable Development*, 31, 1593-1605.
- Roslan, S., Hasan, S., Zaremohzzabieh, Z., & Arsad, N. M. (2021). Big five personality traits as predictors of systems thinking ability of upper secondary school students. *Pertanika Journal of Social Sciences & Humanities*, 29, 251-269.
- Senge, P. M. (1996). *The Fifth Discipline: Fieldbook*. New York: Broadway Business.
- Senge, P. M. (2006). *The Fifth Discipline: The Art & Practice of the Learning Organization*. New York: Crown 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. (2010). *Statistical Analysis Methods Required for Writing SPSS/AMOS*. Seoul; 21st century
- Stavrianeas, S., Bangera, G., Bronson, C., Byers, S., Davis, W., DeMarais, A., ... Offerdahl, E. G. (2022). Empowering faculty to initiate STEM education transformation: Efficacy of a systems thinking approach. *PLoS ONE* 17: 1-17. <https://doi.org/10.1371/journal.pone>.
- Ulku, S. B., & Gaye, D. C. (2024). Research trends on systems thinking approach in science education. *International Journal of Science Education*, 46, 485-502.
- Wright, B., & Linacre, J.(1994). Reasonable mean-square-fit values. *Rasch Measurement Transactions*, 8, 370.

Authors' Information

Lee Hyundong: Daegu National University of Education, Associate Professor, 1st Author. <https://orcid.org/0009-0000-8028-4135>

Park, Byung-Yeol: Dankook University, Research Professor, Co-author <https://orcid.org/0000-0002-7474-9693>

Jeon, Jaedon: Kyungpook National University, Research Professor, Co-author <https://orcid.org/0000-0002-4248-2779>

Lee, Hyonyong: Kyungpook National University, Professor, Corresponding Author <https://orcid.org/0000-0001-5104-1847>