

교육과정평가연구
The Journal of Curriculum and Evaluation
2025, Vol. 28, No. 1, pp.313~345
DOI: <https://doi.org/10.29221/jce.2025.28.1.313>

청소년기 디지털 활용 및 역량 유형의 전이 양상과 예측변인 분석: XGBoost와 잠재전이분석의 적용¹⁾

장은아 (충남대학교 박사과정)*
정혜원 (충남대학교 교수)**

요약

본 연구는 초·중 학교급별로 디지털기기 활용 및 디지털 역량 잠재프로파일과 전이 양상을 살펴보고, 전이 패턴의 분류를 예측하는 핵심 변인을 탐색하기 위하여 수행되었다. 이를 위해 서울특별시교육청 교육연구정보원에서 조사한 서울학생종단연구2020의 초교 패널($n=3,464$) 및 중학교 패널($n=3,556$)의 1, 2차년도 자료에 잠재전이분석, XGBoost와 더불어 설명 가능한 인공지능 알고리즘 중 하나인 SHAP을 순차적으로 적용하였다. 분석 결과, 첫째, 초교 패널과 중학교 패널의 디지털기기 활용 및 역량 잠재프로파일은 모두 3개 집단으로 분류되었으나, 구체적인 프로파일 특성과 할당 비율에는 다소 차이를 보였다. 둘째, 초교 패널과 중학교 패널에 공통적으로 동일한 잠재프로파일을 유지하는 비율이 가장 높았으나, 디지털기기 활용 및 역량 프로파일의 하향 전이보다는 상향 전이하는 비율이 상대적으로 더 높은 것으로 나타났다. 셋째, 초교 패널의 잠재프로파일 전이 패턴 분류에 미치는 예측변인을 살펴본 결과, 성별, 보호자와의 관계, 독서량, 학습방법, 목표의식, 학업소진, 인지역량, 정서역량이 상위 중요도 변인으로 도출되었다. 넷째, 중학교 패널에서는 국어 학교 교사의 수업, 교외 문화활동 참여경험, 여가 시간, 긍정정서, 학습방법, 인지역량과 사회역량, 정서역량이 높은 중요도를 보이는 변인으로 확인되었다. 주요 연구결과를 토대로 초등학생의 디지털기기 활용 빈도 감소 및 중학생의 디지털 역량 함양을 위한 지원 방안을 모색하였다.

주제어: 디지털기기 활용, 디지털 역량, 청소년, XGBoost, 잠재전이분석, SHAP

1) 본 논문은 제7회 서울교육종단연구 학술대회 발표 내용을 수정·보완하여 작성하였음

* 제1저자, rhcsla456@naver.com

** 교신저자, chw7@cnu.ac.kr

I. 서론

최근 디지털 기술의 급격한 발전과 더불어 현대 사회는 디지털 환경이 일상화된 정보화 사회로 빠르게 전환되고 있다. 디지털 전환이 가속화됨에 따라 교육 현장에서도 교육의 도구이자 내용으로서의 디지털 기술 이해와 활용의 필요성이 부각되고 있다(교육부, 2022a). 이 가운데 디지털 기반 교육은 코로나19 이후 급격히 현재화되었으며, 특히 학교 교실 공간 안에서 대면 소통을 바탕으로 하던 학교 수업이 디지털 매체를 통한 원격 수업의 활성화로 이어졌다(권순정 외, 2022). 이러한 환경에서 디지털 역량은 학습의 성패를 좌우하는 중요한 요소로 자리 잡고 있다. 2022 개정 교육과정에서는 인공지능 기술 발전에 따른 디지털 전환 등의 시대적 요구를 반영하고, 미래 사회의 변화에 능동적으로 대응할 수 있도록 모든 학생이 언어 및 수리 소양과 더불어 ‘디지털 기초소양’을 갖출 수 있도록 교육과정을 구성하고자 하였다(교육부, 2022b). 또한 교육부(2022a)는 「디지털 인재양성 종합방안」을 발표하여 모든 국민이 디지털 역량을 충분히 갖출 수 있도록 초·중등교육부터 평생교육까지 AI 디지털교과서 도입, 디지털 교육 역량 강화를 위한 교원 연수, 디지털 인프라 확충 등 다양한 에듀테크 기술을 적용한 교육 혁신 추진을 발표한 바 있다.

디지털 역량은 디지털 미디어를 사용하여 문제를 해결하고, 소통하며, 정보를 관리하는 등의 목적을 위해 지식과 내용을 창조하고 공유하는 데 필요한 기술, 태도, 능력의 집합을 의미하며(Ferrari et al., 2012), 디지털 리터러시와 ICT 리터러시, 미디어 리터러시 등의 용어와 혼용되어 사용되고 있다(길혜지, 이상훈, 황정미, 2022; 장지현, 2023). 디지털기기를 얼마나 효과적으로 능숙하게 활용할 수 있는지에 관한 디지털 역량과 더불어 청소년들이 디지털기기를 어떠한 목적으로 얼마나 자주 활용하는지에 대한 디지털 기기 활용 유형 역시 주요한 연구 주제로 다루어지고 있다(길혜지, 박지민, 김훈, 2023). 즉, 청소년들의 일상생활 속 디지털기기 활용은 학습이나 과제 수행을 위한 학업적 활용부터, 의사소통 및 관계 맺기와 같은 사회적 상호작용, 게임과 음악 감상 등의 오락까지 그 유형이 상이할 수 있다는 것이다(이쌍철 외, 2024; 정수정, 최고은, 2022). 그리고 최근에는 청소년들의 디지털기기 활용 관련 유형을 보다 면밀하고 종합적으로 살펴보고자 하는 목적으로 이러한 디지털기기 활용 목적과 더불어 디지털 역량을 함께 고려하여 잠재계층을 유형화하는 연구가 일부 시도되고 있다(길혜지, 박지민, 김훈, 2023; 노연경, 2021; 정수정, 최고은, 2022). 상기 선행연구를 종합하면, 디지털기기 활용 목적과 그에 대한 빈도, 그리고 디지털 역량을 함께 고려함으로써 전체 집단 내 질적으로 상이한 잠재계층으로 세분화될 수 있음을 알 수 있다.

한편 디지털기기 활용 및 디지털 역량과 관련된 잠재집단을 분류한 연구들이 대부분 중학생을 대상으로 수행되었으며, 이에 따라 이미 디지털기기를 활발히 사용하는 초등학생을 대상으로 한 디지털기기 활용 및 역량 관련 잠재유형 연구는 상대적으로 부족한 실정이다. 또한 기존 연구들은 모두 단일 시점의 자료를 활용하여 잠재계층의 변화양상을 분석하지 못한 아쉬움이 있다. 그러나 잠재계층의 변화양상, 즉 잠재전이에 대한 연구는 단순히 정적으로 계층을 분류하는 데 그치지 않고 디지털기기 활용 및 역량의 동적 변화를 추적할 수 있다는 점에서 중요하다. 특히 특정 변수가 시간의 흐름에 따라 잠재전이에 미치는 영향을 분석함으로써 시간적 선행성을 확보할 수 있으며, 이에 따라 잠재전이에 관한 영향요인을 종합적으로 이해할 수 있다는 점에서 강점이 있다(Zyberaj, Bakaç, & Seibel, 2022). 또한 청소년을 대상으로 한 하루 평균 스마트폰 이용 시간에 대한 조사(배상률, 이창호, 김남두, 2021)에서 학년이 올라갈수록 하루 4시간 이상 스마트폰을 사용하는 학생의 비율이 증가하는 경향이 확인되었으며, 미디어 비판적 사고 능력 집단의 종단적 변화를 분석한 연구(장재홍, 고유정, 2021)에서도 1년 사이 상위 집단의 구성비율이 크게 증가하는 것으로 나타났다. 이러한 통계조사 및 연구결과를 종합하여 보면, 청소년기 디지털기기 활용 및 역량은 비교적 짧은 기간에도 의미 있는 변화를 보일 수 있음을 유추해볼 수 있다. 이를 바탕으로 본 연구는 서울학생종단연구2020 초교 패널 및 중학교 패널의 1, 2차년도 2개 시점 자료에 잠재전이분석(latent transition analysis)을 적용함으로써 청소년의 디지털기기 활용 및 역량과 관련된 개인차와 그 변화양상에 대한 연구를 보다 폭 넓게 수행하고자 한다.

그간 청소년의 디지털기기 활용 및 디지털 역량과 관련하여 수행된 국내 선행연구를 살펴보면, 디지털기기 활용과 디지털 역량에 영향을 미치는 다양한 학생 개인 및 가정, 학교 요인 등 다양한 심리적·사회적 요인이 청소년의 디지털기기 활용 및 디지털 역량에 영향을 미치고 있는 것으로 확인된 바 있다(길혜지, 박지민, 김훈, 2023; 김종민 외, 2018; 박현정, 하여진, 박민호, 2011; 이경숙, 2023). 그러나 상술한 연구는 모두 연구자가 정립한 이론적 배경을 바탕으로 일부 변인만을 분석에 활용하여 청소년의 디지털기기 활용 및 디지털 역량에 대한 요인을 다각적으로 살펴보는 못하였다는 아쉬움이 있다. 따라서 본 연구는 다량의 변수를 투입하여 종속변인에 대한 중요도가 높은 주요 변인을 도출하는 데 유용한 머신러닝 기법 중 하나인 XGBoost(eXtreme Gradient Boosting, Chen & Guestrin, 2016) 기법을 활용하여 디지털기기 활용과 디지털 역량에 영향을 주는 여러 개인 요인뿐만 아니라 가정의 영향력을 광범위하게 탐색해보고자 한다.

즉, 본 연구는 초, 중학교 시기 디지털기기 활용 및 디지털 역량 유형이 어떻게 세분화되고 각 시기의 잠재프로파일이 시간의 경과에 따라 어떻게 전이되는지 살펴본 후, 학교급별로 잠재프로파일의 전이 패턴 분류에 미치는 주요 예측변인을 탐색하고자 한다.

더불어 설명 가능한 인공지능(eXplainable Artificial Intelligence, XAI) 알고리즘 중 하나인 SHAP(SHapley Addictive exPlanations, Lundberg & Lee, 2017)을 활용하여, XGBoost를 통하여 도출한 핵심 변인의 응답수준에 따라 참조집단과 관심집단 중 어느 집단에 속할 확률이 높아지는지에 대한 시각적인 정보를 함께 제시하였다. 주요 연구문제는 다음과 같다.

연구문제 1. 초등학교 4학년에서 초등학교 5학년, 중학교 1학년에서 중학교 2학년 시점에서의 디지털기기 활용 및 디지털 역량 잠재프로파일의 전이 양상은 어떠한가?

연구문제 2. 청소년의 디지털기기 활용 및 디지털 역량 잠재프로파일 전이 패턴 분류를 예측하는 주요 변인은 무엇이며, 주요 예측변인이 미치는 영향력은 어떠한가?

II. 선행연구 고찰

1. 디지털기기 활용 및 디지털 역량 유형

알파세대와 Z세대에 속하는 현재의 초등학생과 중학생은 태어날 때부터 디지털기기와 밀접한 환경에서 성장해온 디지털 네이티브(digital natives)로 불린다. 이들은 스마트폰, 태블릿, 컴퓨터 등 다양한 디지털기기를 자연스럽게 접하며, 일상생활뿐만 아니라 학습 환경에서도 디지털 기술을 적극적으로 활용하는 특징을 보인다(박인우, 2024; 배상률, 이창호, 김남두, 2021). 관련하여 여성가족부(2023)에서 실시한 「청소년 인터넷·스마트폰 이용습관 진단조사」에 따르면, 2023년을 기준으로 초등학교 4학년의 스마트폰 보유율은 96.4%, 중학교 1학년의 스마트폰 보유율은 99.2%에 이르는 것으로 나타났다.²⁾ 이는 대부분의 초·중학생들이 디지털기기를 일상적이고 필수적인 도구로 인식하고 있음을 시사하며, 학년이 높을수록 디지털기기의 이용률이 상대적으로 더 높음을 알 수 있다.

청소년들의 일상생활 속 디지털기기 활용은 다양한 목적과 형태로 나타나며, 크게 학업적 활용 및 비학업적 활용으로 구분할 수 있다. 예컨대 학업적 활용은 학습, 과제 수행의 목적으로 온라인 학습플랫폼과 강의를 수강하거나 디지털기기를 사용하여 학업과 관련된 정보를 검색하는 등의 경우를 의미한다. 한편 비학업적 활용은 소셜 네트워크 등의 상호작용, 게임과 쇼핑, 영상 시청 등의 오락 및 여가 등이 이에 해당한다(정수정, 최고은, 2022). 관련하여 디지털기기 활용 목적 및 빈도에 따른 유형을 실증적으로 분석한

2) 청소년 인터넷·스마트폰 이용습관 진단조사 참여 학생 수(초4 399,129명, 중1 439,655명) 대비 스마트폰 미사용 결과 값을 제외한 학생 수(초4 384,842명, 중1 436,223명)의 비율을 계산하여 산출하였음(여성가족부, 2023).

국내 선행연구를 살펴보면, 먼저 박현정, 하여진과 박민호(2011)의 연구에서는 PISA (Programme for International Student Assessment) 2009 자료를 통하여 한국 학생들의 ICT 활용(게임, 엔터테인먼트, 커뮤니케이션)에 관한 잠재프로파일을 분류한 결과, 고험용 집단, 게임 고험용 집단, 게임 저활용 집단, 저활용 집단의 4개 집단으로 유형화되었다. 이쌍철 등(2024)은 부산교육중단연구 자료를 활용하여 중1 학생의 디지털기기 활용 유형을 살펴보았는데, 공부 이외의 목적 과사용 집단, 공부 목적 사용 집단, 전반적 과사용 집단, 전반적 저사용 집단, 스마트기기 중심 사용 집단의 5개 집단으로 구분한 바 있다 즉, 상술한 연구를 종합하면 디지털기기 사용 빈도에 대한 전반적인 상, 하위 수준 유형뿐만 아니라, 게임, 학습 등 디지털기기 활용 목적에 따라 개별화된 디지털기기 사용 패턴이 상이하게 존재함을 확인할 수 있다.

디지털 역량은 디지털 리터러시와 함께 미디어 리터러시, 인터넷 리터러시, ICT 리터러시 등의 유사 개념들과 다양한 관점에서 혼용되고 있다(길혜지, 이상훈, 황정미, 2022; 장지현, 2023). 디지털 역량은 문제 해결, 의사소통, 지식 생성과 공유 등을 목적으로 ICT와 디지털 미디어를 사용할 때 필요한 지식, 태도, 능력, 기술 및 인식의 집합으로 정의된다(Ferrari et al., 2012), 디지털 역량에 따른 잠재프로파일 유형화를 수행한 선행연구를 살펴보면, 이은지 등(2021)의 연구에서는 한국교육중단연구2013 자료를 사용하여 중학교 2학년 학생의 디지털 역량에 따른 잠재프로파일을 분류하였으며, 디지털 역량의 수준에 따라 부진 집단, 양호 집단, 우수 집단의 3개 집단으로 유형화하였다. ICILS(International Computer and Information Literacy Study) 2018 데이터를 활용하여 우리나라 학생들의 컴퓨터·정보 소양과 컴퓨팅 사고력에 대한 잠재프로파일을 도출한 김경희, 노언경과 상경아(2021)의 연구에서도 마찬가지로 컴퓨터·정보 소양 및 컴퓨팅 사고력 수준에 따라 상, 중, 하의 3계층으로 분류한 바 있다. 성은모와 최효선(2016)은 ICCS(International Civic and Citizenship Study) 2016 자료를 바탕으로 중학생의 ICT 리터러시 유형을 세분화한 결과, ICT 활용의 즐거움과 유용성 인식, ICT 보편적 활용, ICT 비활용 집단, ICT 우수 활용의 4개 집단으로 분류하여 역량 지표변인별로 상이한 패턴이 나타남을 보고한 바 있다.

최근 디지털기기 활용과 디지털 역량, 디지털기기 관련 인식 및 태도를 함께 고려함으로써 보다 종합적인 유형을 도출하는 연구가 일부 시도되고 있다. 노언경(2021)은 ICILS 2018 자료를 기반으로 중학생의 ICT 활용(수업에서 ICT 활용, 학교에서 ICT 학습 등) 및 인식(ICT 진로기대, 긍·부정 인식, 효능감 등) 유형을 분류하였다. 분석 결과, 전반적 저수준 집단, 활용 고수준 집단, 태도 고수준 집단의 3개 집단으로 유형화되었으며, 활용과 태도 수준이 교차되는 집단이 있음을 보고하였다. 정수정과 최고은(2022)은 서울교육중단연구2010 자료를 사용하여 중학교 3학년의 ICT 활용 유형과 리터러시에

따른 잠재프로파일을 구분하였으며, 전반적 고수준 집단, 저빈도·고리터러시 집단, 중빈도·고리터러시 집단의 3개 계층이 존재함을 보고한 바 있다. 길혜지, 박지민과 김훈(2023)은 PISA 2018 데이터를 바탕으로 우리나라 만15세 청소년의 디지털 친숙도, 효능감, 기기활용, 정보검색, 의사소통, 사회참여 등의 활용과 인식 및 태도를 포괄한 디지털 리터러시 유형을 분류하였다. 분석 결과, 저수준, 고수준 집단과 더불어 학교 안 디지털 기기 활용의 활성화 정도에 따른 2개 집단까지 총 4개 집단으로 분류됨을 보고하였다.

이상의 선행연구 결과를 종합하면, 디지털기기 활용과 디지털 역량에 대한 잠재프로파일은 전반적인 사용 빈도 수준에 따라 상, 중, 하위 수준으로 구분될 뿐만 아니라, 활용 목적과 역량 수준에 따라 각 수준이 교차하는 독특한 패턴 역시 존재하는 것으로 나타났다. 그러나 기존 연구는 하나의 학교급에 국한된 연구를 수행하여 학교급간 비교 연구는 미비하였으며, 두 개 시점 이상의 자료를 사용하여 잠재프로파일의 변화 양상을 분석한 연구 역시 부족한 실정이다. 이를 바탕으로 본 연구는 서울학생종단연구2020 자료의 초교 패널과 중학교 패널 자료를 모두 사용하여 초, 중학생의 디지털기기 활용 및 역량에 따른 잠재프로파일을 세분화한 후, 잠재프로파일의 전이 양상을 살펴보고자 한다.

2. 디지털기기 활용 및 역량에 영향을 미치는 요인

본 연구는 디지털기기 활용과 디지털 역량이라는 질적으로 상이한 지표를 함께 사용하는 연구를 수행함에 따라, 디지털기기 활용 빈도에 관한 영향변인과 디지털 역량에 대한 영향변인을 분석한 연구를 별도로 분리하여 살펴보고자 한다.

먼저 디지털기기 활용에 미치는 영향변인을 검증한 최신의 연구가 양적으로 많지는 않았으나, 대부분의 선행문헌에서 성별과 가정 특성에 대한 영향력을 보고하고 있는 것으로 나타났다. 남창우와 신수영(2014)은 남학생일수록 오락용으로 컴퓨터를 활용하는 빈도가 높았으나, 과제용으로는 여학생이 컴퓨터를 활용하는 빈도가 유의하게 더 높은 것으로 보고하였다. 홍원준(2020)은 오락적 사용을 제외하고 남학생보다는 여학생의 스마트폰 사용 빈도가 통계적으로 유의한 수준에서 더 높은 것으로 보고하였다. 박현정, 하여진과 박민호(2011)의 연구에서는 부모와의 독서활동 및 학생에게 사용되는 교육비용이 많을수록, 그리고 어머니의 학력이 높을수록 ICT 고활용 집단보다는 ICT 저활용 집단에 속할 확률이 증가하는 것으로 나타났다. 김혜숙(2010)은 아버지 학력과 가구수입이 높을수록 컴퓨터 사용시간이 감소하는 것으로 보고하였다. 이우현과 임상호(2013)의 연구에서는 가족과의 커뮤니케이션 유형(협의, 다원, 권위, 방임)에 따라 디지털기기(TV, PC, 휴대전화) 이용시간에 유의한 차이가 발생하는 것으로 나타났다.

다음으로 디지털 역량에 대한 영향변인을 분석한 선행연구를 정리하면, 이경숙(2023)

은 독서활동, 학업성적, 공부시간, 자기주도학습, 수업태도, 학생중심수업이 높을수록 디지털 리터러시 태도가 높아지는 것으로 보고하였다. 김종민 등(2018)은 ICT 리터러시에 미치는 영향변인을 초, 중학생으로 구분하여 살펴보았는데, 두 학교급에 공통적으로 여학생이고 학교에서 정보·컴퓨터 수업을 한다고 응답할수록 ICT 리터러시가 높아지는 것으로 나타났다. 학교급별 특징적인 변인에서는 대도시일수록 초등학생의 ICT 리터러시가 높아진 반면 중학생 모형에서는 지역 규모가 유의하지 않았으며, 초등학생 모형에서는 교사 변수가 유의하지 않았지만 교사의 컴퓨터 활용 능력이 높을수록 중학생의 ICT 리터러시가 높아지는 것으로 나타났다. 이은지 등(2021)의 연구에서는 학생의 학업성취도와 월평균가구소득, 부모의 교육적 지원 정도가 낮을수록 디지털 역량 양호 집단 대비 부진 집단에 속할 확률이 높아지는 것으로 나타났으며, 사이버가정학습 프로그램을 많이 이용할수록 디지털 역량 양호 집단 대비 우수 집단에 속할 가능성이 높아졌다.

이상의 선행연구를 정리하면, 청소년의 디지털기기 활용과 디지털 역량은 청소년 개인 요인뿐만 아니라, 가정, 학교 등 다양한 배경 변인에 종합적으로 영향을 받는 것으로 확인할 수 있었다. 그러나 상술한 선행연구들은 대부분 전통적인 분석 기법의 특성상 자유도의 제약이나 다중공선성의 문제로 인하여 모형에 투입 가능한 영향변인의 수가 제한적이었다. 이로 인해 연구자가 설정한 이론적 근거를 기반으로 한 영향변인을 선택적으로 투입하였고, 결국 디지털기기 활용과 역량에 미치는 영향변인을 부분적으로 탐색하였다는 아쉬움이 있었다. 이러한 문제의식을 바탕으로 본 연구는 머신러닝 알고리즘 중 하나인 XGBoost와 설명 가능한 인공지능 알고리즘 중 하나인 SHAP을 활용하여 디지털기기 활용 및 역량의 잠재프로파일 전이 패턴을 분류하는 데 미치는 예측변인을 광범위하게 탐색함으로써 기존의 연구에서 비교적 주목이 부족하였던 새로운 변인을 함께 도출하고자 하였다.

III. 연구방법

1. 연구대상

본 연구는 초등학교와 중학교의 각 학교급별로 디지털기기 활용 및 디지털 역량 잠재프로파일을 분류하고 전이 패턴을 살펴보기 위하여 서울특별시교육청교육연구정보원에서 조사한 서울학생종단연구2020(Seoul Educational Longitudinal Study of Students 2020)의 초교 패널 및 중학교 패널 자료를 사용하였다. 구체적으로 서울학생종단연구

2020 데이터의 1차년도(초교 패널 초4, 중학교 패널 중1)와 2차년도(초교 패널 초5, 중학교 패널 중2) 자료를 활용하였으며, 두 시점의 디지털기기 활용 빈도 및 디지털 역량 문항에 모두 응답한 학생을 연구대상으로 선정하였다. 최종적으로 초교 패널 3,464명(남학생 1,710명, 여학생 1,754명), 중학교 패널 3,556명(남학생 1,788명, 여학생 1,768명)을 분석에 활용하였다.

2. 측정도구

가. 종속변인

본 연구는 초등학교 4학년과 5학년, 중학교 1학년과 2학년 시점에 동일하게 측정된 디지털기기 활용 빈도 및 디지털 역량³⁾과 관련된 문항을 잠재프로파일 분류의 종속변인으로 투입하였으며, 종속변인으로 활용한 구체적인 문항 및 기술통계 분석 결과는 <표 1>과 같다. 디지털기기 활용 빈도는 총 10문항으로 인터넷 검색, 게임, 소셜 네트워크 활동과 더불어 온라인 강의 및 학교 숙제 등의 목적에 대한 디지털기기에 대한 사용 빈도를 5점 척도(1=전혀 안 함, 2=한 달에 1~2번, 3=한 달에 3~4번, 4=일주일에 2~4번, 5=거의 매일)로 측정하였다. 그리고 디지털 역량은 디지털기기 사용 및 문제 해결에 대한 동의 정도를 5점 Likert 척도(1=전혀 그렇지 않다, 2=그렇지 않다, 3=보통이다, 4=그렇다, 5=매우 그렇다)로 측정하였다.

<표 1> 디지털기기 활용 및 디지털 역량 문항 및 기술통계

구 분 번호	문항	초교 패널		중학교 패널	
		초4	초5	중1	중2
		M(SD)	M(SD)	M(SD)	M(SD)
디지털기기를 사용할 때 다음과 같은 활동을 얼마나 자주하는지 선택해 주십시오.					
1	인터넷 검색	3.589(1.279)	3.932(1.127)	4.285(0.983)	4.319(0.982)
2	게임	3.650(1.374)	3.723(1.380)	3.683(1.453)	3.599(1.508)
3	동영상 검색 및 감상(예: 유튜브)	3.876(1.301)	4.179(1.141)	4.577(0.851)	4.629(0.796)
4	이메일 및 메신저 서비스(예: 카카오톡)로 채팅	3.236(1.664)	3.951(1.406)	4.593(0.876)	4.639(0.822)
5	소셜 네트워크 활동(예: 인스타그램, 페이스북, 트위터, 블로그)	1.431(1.060)	1.835(1.453)	3.230(1.792)	3.715(1.640)
6	쇼핑	1.408(0.881)	1.497(0.978)	1.958(1.278)	2.283(1.358)

3) 본 연구는 디지털 역량에 관하여 디지털 미디어를 통하여 소통, 정보 관리, 문제 해결 등의 목적을 위하여 지식 및 내용을 창조하고 공유하는 데 필요한 기술, 태도, 능력의 집합으로 개념화한 Ferrari 등(2012)의 연구를 바탕으로, 서울학생종단연구 2020 자료에서 'ICT 역량'으로 측정된 문항을 디지털 역량 변인으로 활용하였음.

구 분 번호	문항	초교 패널		중학교 패널	
		초4	초5	중1	중2
		<i>M(SD)</i>	<i>M(SD)</i>	<i>M(SD)</i>	<i>M(SD)</i>
7	온라인 강의(EBS, 에듀넷, 사이버가정학습 등) 듣기	2.554(1.571)	2.125(1.428)	2.690(1.005)	2.569(1.452)
8	학교 과제를 위한 인터넷 검색 및 문서 작성	2.717(1.326)	2.783(1.240)	3.335(1.237)	3.089(1.182)
9	학교 과제를 위한 이메일, 메신저(예: 카카오톡), 소셜 네트워크 활동	2.208(1.441)	2.574(1.439)	3.642(1.298)	3.367(1.263)
10	학습 어플(또는 앱) 다운로드하기	2.111(1.290)	2.047(1.263)	2.476(1.332)	2.381(1.272)
디지털기기를 사용한 경험을 기준으로 다음 내용에 어느 정도 동의하는지 선택해 주십시오.					
역 량	1 새로운 프로그램이 필요할 때 스스로 설치할 수 있다	3.669(1.395)	3.892(1.332)	4.315(0.967)	4.361(0.904)
	2 디지털기기에 대한 정보를 익혀 스스로 활용한다	3.872(1.156)	4.129(1.033)	4.307(0.886)	4.344(0.879)
	3 디지털기기 사용을 어려움 없이 할 수 있다	3.905(1.128)	4.143(1.044)	4.238(0.936)	4.287(0.937)
	4 디지털기기에 문제가 발생했을 때 스스로 해결할 수 있다	3.153(1.285)	3.428(1.224)	3.882(1.101)	3.917(1.050)
	5 새로운 어플(또는 앱)이 필요할 때 스스로 선택하고 사용할 수 있다	3.706(1.324)	3.902(1.209)	4.306(0.899)	4.284(0.884)

나. 예측변인

본 연구는 학교급별 디지털기기 활용 및 역량 잠재프로파일의 전이에 영향을 주는 핵심적인 예측변인을 탐색하기 위하여 서울학생종단연구2020의 초교 패널 및 중학교 패널 1차년도 학생 설문 응답 변인을 활용하였다. 이 과정에서 학생 및 학교 ID, 가중치, 생년월일, 주관식 문항 등 디지털기기 활용 및 역량과 직접적으로 관련되지 않은 문항은 제외하였고, 중복응답을 허용하여 순위를 지정하는 문항은 1순위 변인만을 사용하였다(예. 원격수업 형식, 원격 수업 시 피드백 유형 등). 또한 ‘해당사항 없음’, ‘잘 모르겠다’로 응답한 사례는 결측으로 재코딩하였고, 범주형 변인은 더미코딩하여 분석에 사용하였다. 더불어 일부 부정문항(예. 사회역량(갈등해결))은 역코딩하였으며, 개별 문항의 의미가 명료한 문항은 별도 전처리 없이 개별 문항을 그대로 활용하였고(예. 학교만족도, 동아리활동 인식도, 독서량 등), 다수의 유사한 문항으로 측정된 구인은 평균을 계산하여 분석에 투입하였다(예. 교우관계, 학교수업 태도, 목표의식 등). 분석의 정확성을 위하여 결측률이 20%를 초과하는 변인⁴⁾은 삭제하였고, 그 외 결측 데이터는 R 4.2.2의 mice(Ver, 3.15.0) 패키지(van Buuren & Groothuis-Oudshoorn, 2011)를 통해 연속형 변인에는 예측 평균 매칭(predictive mean matching) 기법을 적용하여 예측된 값과 가

4) 대부분 활동 경험 및 참여 여부에 ‘예(참여)’라고 선택한 경우에만 응답하는 연계문항에서 높은 결측률이 확인됨

장 가까운 실제 관측값으로 대체하였고, 이분형 변인에는 로지스틱 회귀분석(logistic regression)을 통해 예측된 확률을 계산하여 0 또는 1로 대체하였다(van Buuren & Groothuis-Oudshoorn, 2011). 마지막으로 학생 설문 변인 중 디지털기기 활용 및 역량과 관련이 높을 것으로 판단한 ICT 흥미 및 디지털기기 의존도 변인 역시 분석에서 제외하였다. 이상의 전처리를 거친 후 최종적으로 초교 패널은 학생 변인 62개를, 중학교 패널은 학생 변인 120개를 모형에 투입하였으며, 상세한 예측변인 목록은 <표 2>와 같다.

<표 2> 예측변인 설명

영역	초4	중1
기본 인적사항	성별	
학교 및 사회생활	학교만족도, 교우관계, 학교교사에 대한 인식, 동아리 활동(인식도, 자기보고식 평가, 프로그램 평가), 문화활동 참여경험(교내, 교외)	
학교 정규수업	학교 교육목표 달성 정도, 학교에 대한 평가, 학교풍토, 자유학기(년)제(인식도, 만족도, 자기보고식 평가, 프로그램 평가, 프로그램 시설 및 환경 평가, 봉사활동 참여 경험, 노동인권교육 이수여부, 아르바이트 경험유무	
방과후학교	학교 수업태도(국어, 영어, 수학), 학교 교사의 수업(국어, 영어, 수학), 성취가치(국어, 영어, 수학), 학교 내 평가활동	
사교육	방과후학교 참여 여부	
가정생활	과목별 사교육 참여 여부 보호자와의 관계, 독서량	
ICT 활용	여가시간, 여가생활 동반자, 여가시간 중 스마트기기 활용 시간	
학습 및 심리 특성	디지털기기 최초 사용 시점	
진로	학습방법(인지조절, 상위인지조절, 행동조절), 목표의식, 성장신념, 고정신념, 우울, 불안(과도한 걱정, 예민함, 신체수면상의 문제, 부정적 정서), 학업소진(정서적 고갈, 냉소), 삶의 만족도 자살사고, 긍정정서, 부정정서	
민주시민 교육	진로성숙도 진학진로활동 프로그램 경험 정도, 고등학교 진학 계획, 자기주도적 학습전형 인지 정도	
비행 및 일탈	민주시민의식(민주주의, 인권), 세계시민의식(평화, 생태, 다문화), 평화통일의식	
학생 특성	현실비행 피해경험, 사이버비행 피해경험, 일탈 및 현실비행 경험, 사이버비행 경험	
교육환경	자습시간, 운동시간, 수면시간, 아침식사 횟수, 건강 학교시설평가(문제발생정도)	
코로나19	원격수업 형식, 원격수업 피드백 유형	

영역	초4	중1
역량검사	인지역량(국어, 영어, 수학), 사회역량(관계성, 협력, 갈등해결), 정서역량(정서조절, 정서인식, 정서표현)	
	참여자치역량	

주. 영역 구분 및 변인명은 서울특별시교육청교육연구정보원에서 제공하는 서울학생종단연구 2020 사용자 매뉴얼을 참조하여 작성하였음

3. 분석방법

본 연구는 초, 중학생의 디지털기기 활용 및 디지털 역량 유형과 그 전이 양상을 분석하고, 전이 패턴을 분류하는 데 주요한 영향을 미치는 변인을 탐색하며, 주요 예측변인의 응답수준에 따라 관심집단에 속할 확률이 어떻게 달라지는지 확인하기 위하여 잠재전이분석과 XGBoost, SHAP을 순차적으로 적용하였다. 분석 절차는 다음과 같다.

첫째, 두 학교급 패널의 디지털기기 활용 및 역량 유형을 분류하고, 시간의 경과에 따라 변화하는 잠재프로파일의 전이 양상을 살펴보고자 3단계 접근법(Asparouhov & Muthén, 2014)을 적용한 잠재전이분석을 활용하였다. 잠재전이분석은 잠재프로파일분석(latent profile analysis)을 종단적 맥락으로 확장한 방법론으로, 연구대상이 시간의 흐름에 따라 서로 다른 잠재프로파일에 속할 가능성을 고려하여 t 시점의 잠재프로파일이 $t+1$ 시점의 잠재프로파일로 전이할 확률을 추정한다(Collins & Lanza, 2013). 3단계 접근법은 각 시점 공변인을 투입하지 않은 무조건 모형을 통하여 적합한 잠재프로파일 수를 상정한 후, 각 시점 잠재프로파일 추정 시의 분류 오차를 확인하여 고정한 다음 잠재전이분석을 실시함으로써 잠재프로파일 분류 시 발생하는 측정오차를 고려하여 보다 정확한 전이 확률 추정을 가능케 하는 접근 방식을 의미한다(Asparouhov & Muthén, 2014).

본 연구는 가장 적합한 잠재프로파일 수를 결정하기 위하여 값이 작을수록 양호한 모형인 것으로 해석하는 정보지수 AIC(Akaike Information Criteria), BIC(Baysian Information Criteria), SABIC(Sample size Adjusted BIC)와 더불어, 1에 가까울수록 분류의 질이 좋은 것으로 판단하는 Entropy, 모형 비교 검증 방법 중 하나로 p -value가 유의하지 않으면 더 간명한 모형을 선택하는 LMRT(Lo- Mendell-Rubin adjusted Likelihood Ratio Test)를 종합적으로 살펴보았다. 상술한 모형 적합도 지수와 각 잠재프로파일별 최소 분류 비율 및 해석의 용이성을 함께 고려하여 가장 적합한 잠재프로파일 수를 선택하였다. 잠재전이분석에는 Mplus 8.11(Muthén & Muthén, 1998-2017) 프로그램을 사용하였고, 잠재프로파일별 지표변인으로 활용한 디지털기기 활용 빈도 및 역량 문항 점수에 유의한 차이가 발생하는지 검증하기 위하여 수행한 일원분산분석에는

SPSS 29.0을 활용하였다.

둘째, 각 학교급별 잠재프로파일의 전이 패턴을 살펴본 후 참조집단과 관심집단을 설정하여 XGBoost 분류 모델을 구축하였다. XGBoost는 약한 예측기인 결정 트리의 오차를 줄이기 위해 잔차를 예측하는 방식으로 순차적으로 예측성능을 향상시키는 부스팅을 확장한 알고리즘을 의미한다(Chen & Guestrin, 2016). 본 연구에서는 잠재전이분석과 머신러닝 기법을 병행하여 전이 패턴의 예측 변인을 탐색했던 선행연구(강유림, 김소희, 2024; 김미림, 홍민주, 2022)를 바탕으로, 분류 모델 수행에 앞서 참조집단을 설정할 때 관심 있는 전이 양상을 고려함과 동시에 사례 수를 중요한 기준으로 활용하였다는 점을 반영하여 참조집단과 관심집단을 선정하였다. 즉, 참조집단은 사례수가 가장 많아 분석의 안정성을 확보할 수 있는 집단을 기준으로 설정하였고, 관심집단은 참조집단과 비교하여 의미있는 해석이 가능하도록 전이 패턴의 특성을 고려하여 선정하였다. 또한 최적의 하이퍼파라미터 조합 탐색을 위한 훈련 모델을 구축하기에 앞서 데이터 불균형 문제에 대응하기 위하여 소수 클래스의 데이터를 무선 복제 추출하는 랜덤 오버샘플링(random oversampling)을 실시하였다. 오버샘플링한 데이터를 추출하여 훈련용 자료 70%, 검증용 자료 30%로 구분하였고, 훈련용 자료로 하이퍼파라미터 튜닝 및 자료 적합을 수행한 후 검증용 자료로 모형의 예측 성능을 확인하였다(신은진, 손원숙, 2024). 하이퍼파라미터 튜닝에 사용된 하이퍼파라미터 조건 및 최고 성능을 보인 하이퍼파라미터 조합을 정리하면 <표 3>과 같다. 본 연구는 오버샘플링을 위하여 Python(Ver, 3.12.5) imbalanced-learn(Ver, 0.10.1) 라이브러리의 RandomOverSampler 모듈을, XGBoost 분석에는 xgboost(Ver, 2.1.1) 라이브러리에 포함된 XGBRegressor 모듈을 활용하였다.

<표 3> 하이퍼파라미터 설명 및 조합

구분	하이퍼파라미터 설명	하이퍼파라미터 조합	최적의 하이퍼파라미터	
			초교	중학교
n_estimators	결정트리의 수	[100, 300, 500, 700, 1,000, 1,500]	700	500
learning_rate	개별 결정트리의 기여도를 조절하는 학습률	[0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5]	0.05	0.05

셋째, 잠재프로파일의 전이 패턴 분류에 영향을 미치는 주요 예측변인의 중요도 지수를 산출하고 분석결과를 시각적으로 해석하기 위하여 설명 가능한 인공지능 알고리즘 중 하나인 SHAP(Lundberg & Lee, 2017)을 적용하였다. SHAP은 복잡한 머신러닝 모델의 결과를 해석할 수 있도록 설명 가능한 인공지능 알고리즘 중 하나로, Shapley value에 기반하여 모델의 예측 결과를 설명하는 기법 중 하나이다(Lundberg et al.,

2020). 여기서 Shapley value란 모델의 예측 결과에서 각 예측변인이 차지하는 기여도를 정량적으로 평가하기 위하여 해당 변인이 포함된 모든 변인 조합에서의 기여도를 평균적으로 계산한 값을 의미한다. 본 연구는 Shapley value의 절댓값에 평균을 취한 SHAP 지수를 바탕으로 상위 중요도 10위 변인을 이직의도에 관한 핵심 예측변인으로 산출하였다.

더불어 주요 예측변인의 응답수준에 따라 관심집단에 속할 확률이 어떻게 달라지는지에 관한 시각적인 정보를 확인하고자 SHAP 의존성 도표(SHAP dependence plot)를 제시하였다. SHAP 의존성 도표에서 각 점은 개별 관측 사례를, X축은 각 예측변인의 응답 수준을, Y축은 Shapley value를 의미한다(Lundberg et al., 2020). 즉, 개별 관측치의 Shapley value가 양수 영역에 포진되어 있다면 해당 변인이 참조집단(0) 대비 관심집단(1)에 속하도록 기여하는 것으로 해석 가능하다. 본 연구는 SHAP 분석을 위하여 Python (Ver, 3.12.5) shap(Ver, 0.44.1) 라이브러리의 TreeExplainer 모듈을 사용하였다.

IV. 연구결과

1. 디지털기기 활용 및 역량 잠재프로파일 분석

가. 초교 패널

초교 패널의 각 시점별 디지털기기 활용 및 역량 잠재프로파일의 수를 결정하고자 잠재프로파일의 수를 2개부터 4개까지 순차적으로 1개씩 증가시키면서 모형적합도를 비교하였으며, 분석결과는 <표 4>와 같다. 초등학교 4학년과 5학년 시기에 공통적으로 정보지수 AIC, BIC, SABIC는 계층의 수를 늘릴수록 지속적으로 감소하는 것으로 나타났고, Entropy는 잠재프로파일 수가 3개일 때 가장 높았다. 한편 초5 시기 잠재프로파일 수가 4개일 때 LMRT 결과가 유의하지 않아, 더 간명한 모형인 3개 계층이 보다 적합한 것을 알 수 있다. 그 외 계층별 소속 비율, 해석 용이성 등을 고려하여 본 연구는 두 시기에 공통적으로 3계층 모형을 최종 모형으로 결정하였다.

<표 4> 초교 패널 잠재프로파일 수에 따른 모형 적합도 지수

분류기준	초등학교 4학년			초등학교 5학년			
	2개	3개	4개	2개	3개	4개	
AIC	166270.139	160970.170	159527.267	161680.968	156450.028	154924.011	
BIC	166553.047	161351.481	160006.981	161963.877	156831.339	155403.725	
SABIC	166406.883	161154.478	159759.138	161817.713	156634.335	155155.881	
Entropy	0.855	0.913	0.864	0.875	0.930	0.905	
LMRT	<.001	<.001	<.001	<.001	<.001	>.05	
분류 (%)	집단1	39.6	12.4	15.0	33.6	22.9	31.7
	집단2	60.4	52.5	37.2	66.4	50.9	7.7
	집단3		35.1	35.4		26.2	38.2
	집단4			12.4			22.5
	집단5						

각 시점 3개 계층 유사 범위 내의 잠재프로파일 수를 포함하여 최적의 잠재프로파일 전이 조합을 결정하고자 하였으며, 초4와 초5 시기의 잠재프로파일 수를 2~4개까지 조합한 9개 모형의 적합도를 산출한 결과는 <표 5>와 같다. 분석 결과, 초4와 초5 시기 각각 잠재프로파일의 수가 3개일 때 Entropy가 가장 높은 것으로 확인되었다. 이를 바탕으로 잠재전이분석을 위한 각 시점 잠재프로파일의 수를 3개로 최종 결정하였다.

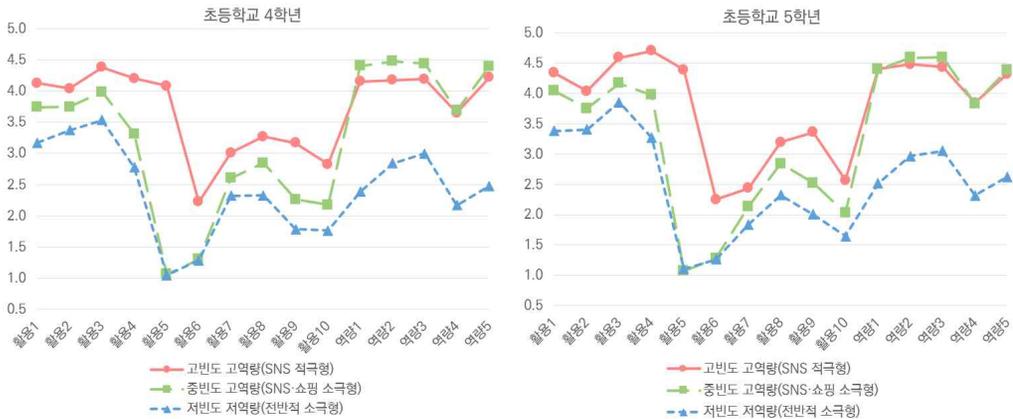
<표 5> 초교 패널 잠재프로파일 전이 조합

잠재프로파일의 수		AIC	BIC	SABIC	Entropy
초4	초5				
2	2	8585.734	8604.184	8594.652	0.786
2	3	11347.052	11377.803	11361.916	0.839
2	4	12899.940	12942.992	12920.749	0.821
3	2	10645.167	10675.918	10660.030	0.838
3	3	13169.027	13218.228	13192.808	0.872
3	4	14722.863	14790.515	14755.563	0.854
4	2	12737.364	12780.416	12758.173	0.799
4	3	15267.592	15335.244	15300.292	0.835
4	4	16756.768	16849.020	16801.358	0.825

초4와 초5 시기 디지털기기 활용 및 역량 프로파일의 양상을 확인하기 위하여 지표변인별 평균 수준을 시각화한 결과는 [그림 1]과 같다. 두 시기에서 도출된 잠재프로파일은 유사한 양상을 보였는데, 실선으로 표기된 집단1은 디지털기기 활용 빈도와 디지털역량이 모두 높은 수준인 것으로 나타났으며, 특히 소셜 네트워크 활동(활용5) 문항에서 여타 집단 대비 평균이 3점 이상 높은 것으로 나타났다. 이를 바탕으로 집단1을 ‘고빈도

고역량(SNS 적극형) 집단으로 명명하였으며, 초4 시기는 12.4%, 초5 시기는 22.9%의 학생이 고빈도 고역량(SNS 적극형) 집단에 속하는 것으로 확인되었다. 파선으로 표기된 집단2는 디지털기기 활용 빈도는 중간 수준이지만 소셜 네트워크 활동(활용5)과 쇼핑(활용6) 문항에서 유난히 낮은 빈도를 보였고, 디지털 역량은 고빈도 고역량(SNS 적극형) 집단과 유사한 것으로 나타나 ‘중빈도 고역량(SNS·쇼핑 소극형)’ 집단으로 명명하였다. 초4 시기는 52.5%, 초5 시기는 50.9%의 학생이 이 집단에 소속되는 것으로 나타났다. 점선으로 표기된 집단3은 디지털기기 활용 빈도와 역량이 세 집단 중 가장 낮은 것으로 나타나 ‘저빈도 저역량(전반적 소극형)’ 집단으로 명명하였고, 초4 시기의 35.1%, 초5 시기의 26.2%의 학생이 저빈도 저역량(전반적 소극형) 집단에 속하는 것으로 나타났다.

한편 잠재프로파일별 디지털기기 활용 빈도와 역량 지표 수준에 통계적으로 유의한 차이가 있는지 일원 분산분석을 통하여 사후검증을 수행한 결과, 디지털기기 활용 빈도를 측정된 10개 문항 중 1~2개 문항에서 중빈도 고역량(SNS·쇼핑 소극형) 집단과 저빈도 저역량(전반적 소극형) 집단 간 유의한 차이가 없는 것으로 나타났다. 또한 두 시점에 공통적으로 디지털 역량 문항 5개 중 일부 문항에서 고빈도 고역량(SNS 적극형) 집단과 중빈도 고역량(SNS·쇼핑 소극형) 집단 간의 유의한 차이가 없는 것으로 나타났고, 특히 초5 시기의 경우 디지털 역량 총점에 대해서도 고빈도 고역량(SNS 적극형) 집단과 중빈도 고역량(SNS·쇼핑 소극형) 간의 유의한 점수 차이는 없는 것으로 확인되었다(부록 <표 14>, <표 15> 참조).



[그림 1] 초교 패널 잠재프로파일별 디지털기기 활용 및 역량 수준⁵⁾

5) 초교 패널(그림 1) 및 중학교 패널(그림 2) 잠재프로파일 분류에 활용된 지표변인은 다음과 같음. 활용1: 인터넷 검색/ 활용2: 게임/ 활용3: 동영상 검색 및 감상/ 활용4: 이메일 및 메신저 서비스/ 활용5: 소셜 네트워크 활동/ 활용6: 쇼핑/ 활용7: 온라인 강의 시청/ 활용8: 학교 숙제 목적 인터넷 검색 및 문서 작성/ 활용9: 학교 숙제 목적 메신저 및 소셜 네트워크/ 활용10: 학습 어플 다운로드/ 역량1: 새로운 프로그램 스스로 설치 가능/ 역량2: 기기 정보 활용 가능/ 역량3: 어려움 없이 기기 사용 가능/ 역량4: 기기에 문제 발생 시 해결 가능/ 역량5: 새로운 어플 필요 시 선택 및 사용 가능

나. 중학교 패널

중학교 패널의 각 시점별 디지털기기 활용 및 역량 잠재프로파일의 수를 선택하기 위하여 잠재프로파일의 수 모형별 적합도를 비교하였고, 분석결과는 <표 6>에 제시하였다. 중학교 1학년 및 2학년 시기에 공통적으로 AIC, BIC, SABIC는 계층의 수가 증가됨에 따라 지속적으로 감소하였으나 계층 수가 2개에서 3개로 높아질 때 감소폭이 가장 큰 것으로 나타났으며, Entropy는 초교 패널과 마찬가지로 잠재프로파일 수가 3개일 때 가장 높았다. 모형 적합도 외에도 계층별 소속 비율과 해석 용이성을 함께 반영하여 본 연구에서는 중1과 중2 두 시기에 공통적으로 3계층 모형을 최종 모형으로 결정하였다.

<표 6> 중학교 패널 잠재프로파일 수에 따른 모형 적합도 지수

분류기준	중학교 1학년			중학교 2학년			
	2개	3개	4개	2개	3개	4개	
AIC	155528.592	151307.402	149872.685	151628.054	146671.079	144906.108	
BIC	155812.706	151690.338	150354.443	151912.168	147054.015	145387.866	
SABIC	155666.541	151493.334	150106.599	151766.003	146857.011	145140.022	
Entropy	0.863	0.943	0.870	0.881	0.962	0.889	
LMRT	<.001	<.001	<.001	<.001	<.001	<.001	
분류 (%)	집단1	31.9	50.3	23.5	30.1	54.6	14.2
	집단2	68.1	34.8	14.8	69.9	31.0	32.5
	집단3		14.9	34.4		14.4	30.7
	집단4			27.3			22.6
	집단5						

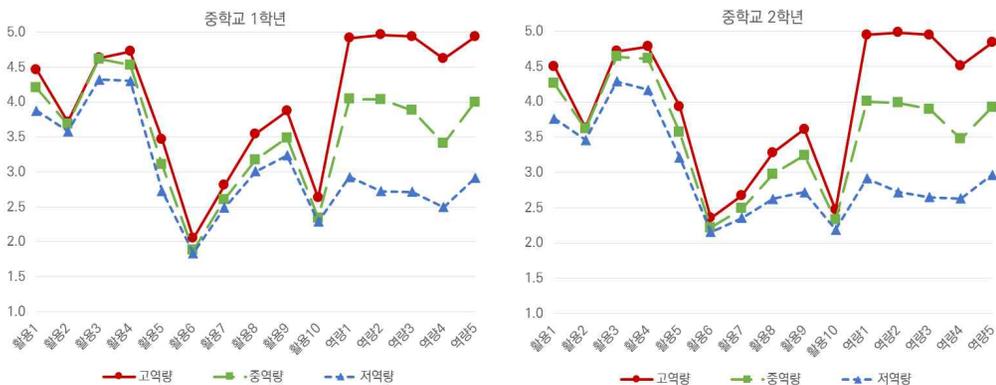
중1과 중2 시기의 잠재프로파일 수를 각 시점 2~4개까지 조합하여 9개 모형의 적합도를 산출하였으며, 분석결과는 <표 7>에 제시하였다. 초교 패널과 마찬가지로 중1과 중2 시기 잠재프로파일의 수가 각각 3개일 때 Entropy가 가장 높은 것으로 나타나 각 시점 잠재프로파일의 수를 분류의 질이 가장 높은 3개로 최종 결정하였다.

<표 7> 중학교 패널 잠재프로파일 전이 조합

잠재프로파일의 수		AIC	BIC	SABIC	Entropy
중1	중2				
2	2	8334.167	8352.696	8343.163	0.789
2	3	10889.938	10920.820	10904.933	0.865
2	4	13494.910	13538.145	13515.902	0.805
3	2	10897.019	10927.901	10912.014	0.857
3	3	13295.840	13345.251	13319.831	0.907
3	4	15921.148	15989.088	15954.136	0.851
4	2	13377.147	13420.381	13398.139	0.796
4	3	15759.494	15827.434	15792.482	0.851
4	4	18215.333	18307.979	18260.316	0.819

중1과 중2 시기 잠재프로파일에 따른 디지털기기 활용과 디지털 역량의 양상을 검토하기 위하여 지표변인별 평균 수준을 시각화한 결과는 [그림 2]와 같다. 후술할 계층 간 디지털기기 활용 빈도 및 역량 문항의 평균 차이 검증을 수행한 부록 <표 16>, <표 17>에서 나타나는 것처럼, 일부 문항에서 일원분산분석 결과가 유의하지 않을 뿐만 아니라 사후검증에서도 통계적으로 유의한 평균 차이가 없는 것으로 확인된 바, 중학교 패널 결과에 대하여 본 연구에서는 디지털 역량에 의한 차이를 중심으로 다음과 같이 집단 명칭을 부여하였다. 먼저 실선으로 나타낸 집단1은 디지털기기 활용 빈도 및 역량이 가장 높은 것으로 나타나 ‘고역량’ 집단으로 명명하였으며, 중1 시기는 50.3%, 중2 시기는 54.6%의 과반수 학생이 고역량 집단에 포함되는 것으로 나타났다. 파선으로 표기한 집단2는 디지털기기 활용 빈도와 역량이 모두 중간 수준인 것으로 드러나 ‘중역량’ 집단으로 명명하였고, 전체 학생 중 중1 시기는 34.8%, 중2 시기는 31.0%의 학생이 중역량 집단에 속하는 것으로 확인되었다. 점선으로 나타낸 집단3은 디지털기기 활용 빈도와 디지털 역량이 가장 저조한 것으로 나타나 ‘저역량’ 집단으로 명명하였고, 중1 시기의 14.9%, 중2 시기의 14.4%의 학생이 저역량 집단에 속하는 것으로 나타났다.

잠재프로파일별 디지털기기 활용 빈도 및 디지털 역량 수준에 유의한 차이가 발생하는지 검증한 결과, 중1과 중2 시기에 공통적으로 디지털기기 활용 빈도를 측정하는 10개 문항 중 게임 목적의 디지털기기 활용 문항의 경우 일원분산분석 결과가 유의하지 않았고, 그 외 시점별로 2~4개 문항에서 중역량 집단과 저역량 집단 간 유의한 평균 차이가 없는 것으로 확인되었다. 한편 디지털기기 활용 빈도 총점과 더불어 디지털 역량 5개 문항, 디지털 역량 총점의 경우 고역량, 중역량, 저역량 순으로 유의한 평균 차이를 확인할 수 있었다(부록 <표 16>, <표 17> 참조).



[그림 2] 중학교 패널 잠재프로파일별 디지털기기 활용 및 역량 수준

2. 디지털기기 활용 및 역량 잠재프로파일 중단적 전이 확률

가. 초교 패널

각 시기별 잠재프로파일을 분류한 후, 초4와 초5 시기 잠재프로파일 수를 각각 3개로 지정하여 잠재전이분석을 수행하였으며, 초4에서 초5 시기로의 잠재프로파일 전이 확률을 산출한 결과는 <표 8>과 같다. 먼저 초5의 고빈도 고역량(SNS 적극형) 집단은 고빈도 고역량(SNS 적극형) 집단을 초5 시기에도 유지할 확률이 57.5%로 가장 높았고, 이어서 중빈도 고역량(SNS·쇼핑 소극형) 집단으로 30.6%, 저빈도 저역량(전반적 소극형) 집단으로 11.9%가 하향 전이하는 것으로 나타났다. 초4 시기 중빈도 고역량(SNS·쇼핑 소극형) 집단 역시 초5 시기에도 동일 집단을 유지할 확률이 65.8%로 가장 높았으며, 이후 상대적으로 디지털기기 활용 빈도가 높은 고빈도 고역량(SNS 적극형) 집단으로 전이할 확률은 21.4%, 저빈도 저역량(전반적 소극형) 집단으로 하향 전이할 확률은 12.7%로 나타났다. 마지막으로 초4 시기 저빈도 저역량(전반적 소극형) 집단의 경우에서도 마찬가지로 초5 시기에도 저빈도 저역량(전반적 소극형) 집단을 유지할 확률이 53.4%로 가장 높았고, 다음으로 중빈도 고역량(SNS·쇼핑 소극형)(34.3%), 고빈도 고역량(SNS 적극형)(12.3%) 집단으로 상향 전이하는 것으로 확인되었다.

<표 8> 초교 패널 잠재프로파일 간 중단적 전이확률 (단위: %)

		초5			합계
		고빈도 고역량 (SNS 적극형) (22.9)	중빈도 고역량 (SNS·쇼핑 소극형) (50.9)	저빈도 저역량 (전반적 소극형) (26.2)	
초4	고빈도 고역량 (SNS 적극형) (12.4)	57.5	30.6	11.9	100.0
	중빈도 고역량 (SNS·쇼핑 소극형) (52.5)	21.4	65.8	12.7	100.0
	저빈도 저역량 (전반적 소극형) (35.1)	12.3	34.3	53.4	100.0

주. 괄호 안의 값은 각 시기에서 해당 잠재프로파일에 속하는 대상의 비율을 나타냄

나. 중학교 패널

중1에서 중2로의 디지털기기 활용 및 역량 프로파일의 전이 확률을 분석한 결과는 <표 9>와 같다. 중1 시기 고역량 집단의 경우 중2 시기에도 고역량 집단을 유지할 확률

이 74.8%로 매우 높았으며, 이어서 중역량(19.3%), 저역량(6.0%)으로 하향 전이할 확률은 상대적으로 낮았다. 중1 시기 중역량 집단도 마찬가지로 중2 시기에도 중역량 집단을 유지할 확률이 46.4%로 높았고, 다음으로 고역량 집단으로 상향 전이할 확률이 39.3%, 저역량 집단으로 하향 전이할 확률이 14.3%인 것으로 나타났다. 중1 시기 저역량 집단 역시 저역량 집단을 유지할 확률이 42.9%로 가장 높았으며, 중역량(35.8%), 고역량(21.4%) 순으로 상향 전이할 확률이 높았다.

<표 9> 중학교 패널 잠재프로파일 간 종단적 전이확률 (단위: %)

	중2			합계	
	고역량 (54.6)	중역량 (31.0)	저역량 (14.4)		
중1	고역량 (50.3)	74.8	19.3	6.0	100.0
	중역량 (34.8)	39.3	46.4	14.3	100.0
	저역량 (14.9)	21.4	35.8	42.9	100.0

주. 괄호 안의 값은 각 시기에서 해당 잠재프로파일에 속하는 대상의 비율을 나타냄

3. 디지털기기 활용 및 역량 잠재전이 패턴의 예측변인 탐색

본 연구는 각 학교급 패널별로 디지털기기 활용 및 역량 잠재프로파일의 전이 패턴 분류를 예측하는 핵심 변인을 탐색하기에 앞서, 잠재프로파일의 사례수 및 전이 양상을 바탕으로 참조집단과 관심집단을 설정하였다. 초교 패널과 중학교 패널 간에 질적으로 상이한 디지털기기 활용 및 역량 잠재프로파일이 확인된 바, 각 학교급 패널별로 상이한 참조집단 및 관심집단을 상정하였다. 이후 참조집단과 관심집단을 분류하는 데 영향을 미치는 핵심 변인 탐색을 위하여 XGBoost 모델을 수행하였고, 각 패널에서 SHAP 지수 상위 10개의 주요 변인을 도출하였다. 더하여 예측변인의 응답수준에 따라 참조집단 대비 관심집단에 속할 확률이 어떻게 달라지는지 시각화하여 확인하고자 하였다.

가. 참조 및 관심집단 설정

(1) 초교 패널

초교 패널 전체 사례수 대비 디지털기기 활용 및 디지털 역량 잠재프로파일의 전이 패턴별 사례수와 비율 산출 결과는 <표 10>과 같다. 본 연구에서는 초교 패널 전이 패턴에서 가장 많은 학생에 속하는 것으로 나타난 중빈도 고역량(SNS·쇼핑 소극형) 집단을 유지하는 경우를 참조집단으로 설정하였다. 그리고 초4 시기 중빈도 고역량(SNS·쇼핑 소극형) 집단에서 초5 시기 고빈도 고역량(SNS 적극형) 집단으로 전이하는 경우를

관심집단으로 설정하여, 디지털 역량은 유사하게 유지하되 디지털기기 활용 빈도가 높아지는 집단에 속할 확률을 높이는 예측변인을 탐색하고자 하였다.

<표 10> 초교 패널 잠재프로파일 전이 패턴별 사례수와 비율 (단위: 명, %)

		초5			총계
		고빈도 고역량 (SNS 적극형)	중빈도 고역량 (SNS·쇼핑 소극형)	저빈도 저역량 (전반적 소극형)	
초4	고빈도 고역량 (SNS 적극형)	247(7.1)	132(3.8)	52(1.5)	431(12.4)
	중빈도 고역량 (SNS·쇼핑 소극형)	386(11.1)	1,163(33.6)	271(7.8)	1,820(52.5)
	저빈도 저역량 (전반적 소극형)	159(4.6)	467(13.5)	587(16.9)	1,213(35.1)
	총계	792(22.9)	1,762(50.9)	910(26.2)	3,464(100.0)

(2) 중학교 패널

중학교 패널 전체 사례수 대비 잠재프로파일 전이 패턴별 사례수와 비율을 산출한 결과는 <표 11>에 제시하였다. 본 연구에서는 중학교 패널 전이 패턴에서 가장 많은 사례수가 포함되는 것으로 나타난 고역량 집단을 유지하는 경우를 참조집단으로 설정하였다. 그리고 중1 시기 고역량 집단에서 중2 시기 중역량 집단으로 하향 전이하는 경우를 관심집단으로 상정하여, 디지털기기 활용 빈도는 유사하게 유지하되 디지털 역량이 감소하는 집단에 소속할 가능성을 높이는 예측변인을 탐색하고자 하였다.

<표 11> 중학교 패널 잠재프로파일 전이 패턴별 사례수와 비율 (단위: 명, %)

		중2			총계
		고역량	중역량	저역량	
중1	고역량	1,322(37.2)	356(10.0)	111(3.1)	1,789(50.3)
	중역량	500(14.1)	555(15.6)	181(5.1)	1,236(34.8)
	저역량	119(3.3)	192(5.4)	220(6.2)	531(14.9)
	총계	1,941(54.6)	1,103(31.0)	512(14.4)	3,556(100.0)

나. 전이 패턴 분류 정확도

초교 패널의 경우 중빈도 고역량(SNS·쇼핑 소극형) 집단을 유지하는 패턴 대비 중빈도 고역량(SNS·쇼핑 소극형) 집단에서 고빈도 고역량(SNS 적극형) 집단으로의 전이 패턴 사례가 상대적으로 부족하였고, 중학교 패널 역시 고역량 집단을 유지하는 패턴보다는 고역량에서 중역량으로 전이하는 패턴의 사례수가 더 적게 나타남에 따라, 데이터 불

균형 문제를 해결하기 위하여 랜덤 오버샘플링을 수행하였다. 오버샘플링한 자료를 훈련용 및 시험용 자료로 구분하여 XGBoost 분류 모델을 적합한 후 각 패널별 예측성능을 분석한 결과는 <표 12>와 같다. 분석 결과, 초교 패널은 모두 정확도(accuracy), 정밀도(precision), 재현율(recall) 모두 0.8~0.9 수준으로 높은 성능을 보이며, 정밀도와 재현율의 조화평균인 F1 score 역시 0.9에 가깝게 나타났다. 중학교 패널 모형에서는 모든 성능지표가 0.9를 초과하는 것으로 나타났다. 즉, 두 패널 모두 참조집단과 관심집단을 분류함에 있어 과적합이 발생하지 않고 분류 예측이 비교적 정확하게 이루어진 것으로 판단할 수 있다.

<표 12> 각 패널별 전이 패턴 분류 예측성능

패널	Accuracy	Precision	Recall	F1 Score
초교	0.865	0.838	0.905	0.871
중학교	0.927	0.914	0.942	0.928

다. 주요 예측변인 도출 및 응답수준에 따른 경향성

초교 패널 및 중학교 패널의 전이 패턴 분류를 예측하는 주요 변인을 탐색하여 SHAP 지수를 바탕으로 상위 중요도 10위 변인을 패널별로 각각 도출한 결과는 <표 13>과 같다.

먼저 초교 패널의 인적사항 영역에서 ‘성별’이 최상위 중요도 변인으로 도출되었으며, 가정생활 영역에서는 ‘보호자와의 관계’, ‘독서량’이 도출되었다. 다음으로 학습 및 심리 특성 영역에서는 ‘학습방법(행동조절)’과 ‘목표의식’, ‘학업소진(정서적 고갈, 냉소)’이 중요도가 높은 변인으로 도출되었고, 역량검사를 통한 조사영역에서는 ‘인지역량(국어, 수학)’, ‘정서역량(정서조절)’이 상위 중요도 변인으로 확인되었다.

중학교 패널 결과를 살펴보면, 학교 및 사회생활 영역에서 ‘국어 학교 교사의 수업’ 및 ‘교외 문화활동 참여경험’의 중요도가 높게 나타났으며, 가정생활 영역에서는 ‘여가 시간’이 주요 변인으로 도출되었다. 이어서 학습 및 심리 특성 영역에서는 ‘긍정정서’와 ‘학습방법(상위인지조절)’이 중요도가 높은 핵심 변인으로 도출되었고, 역량검사에 따른 조사영역에서는 ‘인지역량(국어, 수학)’과 ‘사회역량(관계성, 갈등해결)’, ‘정서역량(정서인식)’이 상위 중요도 변인으로 나타났다.

<표 13> 각 패널별 전이 패턴 분류 주요 예측변인

초교					중학교				
영역	변인	척도	SHAP	순위	영역	변인	척도	SHAP	순위
인적사항	성별	a	0.843	1	학교 및 사회생활	국어 학교 교사의 수업	b	0.227	2
가정생활	보호자와의 관계	b	0.276	9	사회생활	교외 문화활동 참여경험	f	0.167	9
	독서량	c	0.436	2	가정생활	여가시간	g	0.185	7
학습 및 심리 특성	학습방법(행동조절)		0.238	10	학습 및 심리 특성	긍정정서	h	0.211	5
	목표의식	b	0.308	4	역량검사	학습방법(상위인지조절)	b	0.168	8
	학업소진(정서적 고갈)		0.352	3		인지역량(국어)	i	0.216	4
학업소진(냉소)		0.291	6	인지역량(수학)	j	0.167	10		
역량검사	인지역량(국어)	d	0.295	5	역량검사	사회역량(관계성)		0.276	1
	인지역량(수학)	e	0.286	7	사회역량(갈등해결)	b	0.203	6	
	정서역량(정서조절)	b	0.280	8	정서역량(정서인식)		0.227	3	

a) 0=남자, 1=여자

b) 1=전혀 그렇지 않다 ~ 5=매우 그렇다

c) 1=읽지 않음 ~ 8=7권 이상

d) 국어 역량검사 점수(총점 26점)

e) 수학 역량검사 점수(총점 23점)

f) 1=전혀 안 함 ~ 5=1주에 1~2번

g) 1=1시간 미만 ~ 8=7시간 이상

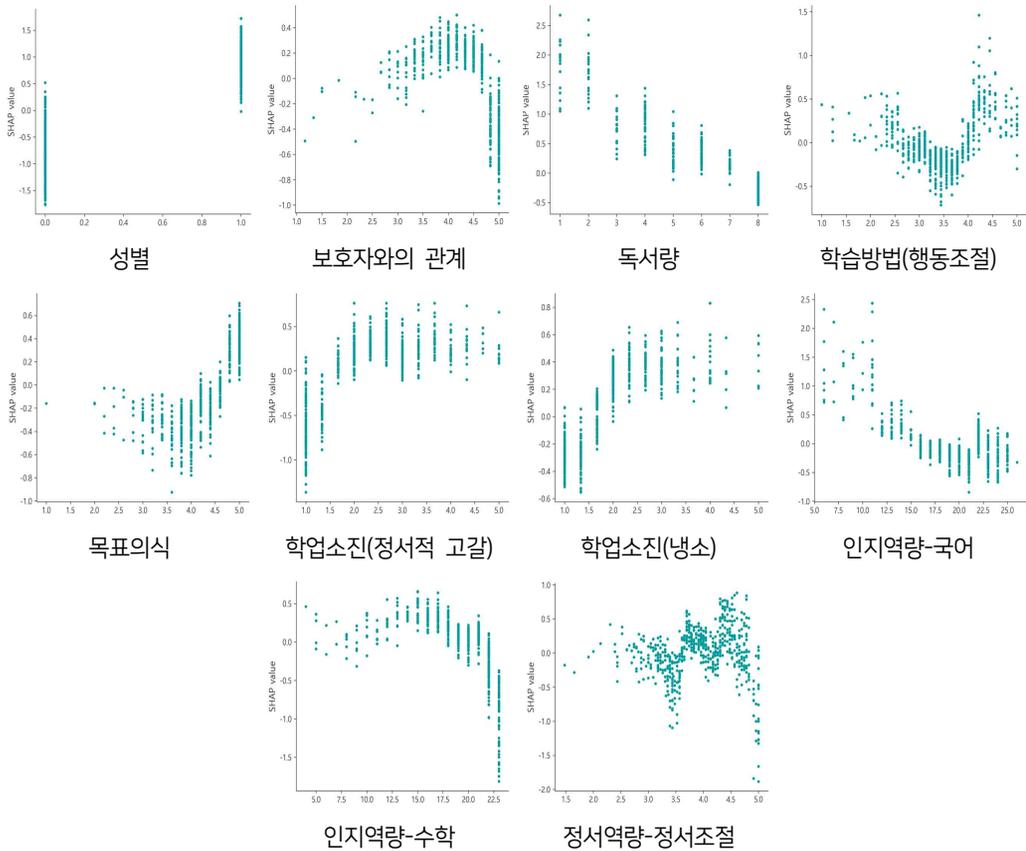
h) 1=전혀 아니다 ~ 5=매우 그렇다

i) 국어 역량검사 점수(총점 29점)

j) 수학 역량검사 점수(총점 25점)

(1) 초교 패널

먼저 [그림 3]에 제시한 바와 같이 초교 패널의 상위 중요도 10위 변인과 응답수준에 따른 경향성을 살펴보면, 먼저 남성보다는 여성인 경우 참조집단인 중빈도 고역량(SNS·쇼핑 소극형) 유지 패턴보다는 관심집단인 고빈도 고역량(SNS 적극형) 집단으로의 전이 패턴에 속할 확률이 높은 것으로 나타났다. 그리고 보호자와의 관계, 독서량, 인지역량(국어, 수학)의 경우 응답수준이 높아질수록 Shapley value가 점차 음수로 전환되었으며, 이는 이들 변인의 응답수준이 높을수록 관심집단보다는 참조집단인 중빈도 고역량(SNS·쇼핑 소극형) 집단을 유지할 가능성이 높음을 의미한다. 반면 목표의식과 학업소진(정서적 고갈, 냉소)의 경우 응답수준이 높아질수록 Shapley value가 증가하는 것으로 나타나, 초4 시기 목표의식과 학업소진을 높게 인식할수록 중빈도 고역량(SNS·쇼핑 소극형) 집단을 유지하기 보다는 초5 시기 고빈도 고역량(SNS 적극형) 집단으로의 전이 패턴에 속할 확률이 높아짐을 알 수 있다. 한편 학습방법(행동조절)과 정서역량(정서조절) 역시 상위 중요도 변인으로 도출되었으나, SHAP 의존성 도표에서 명확한 경향성을 나타내지는 않았다.

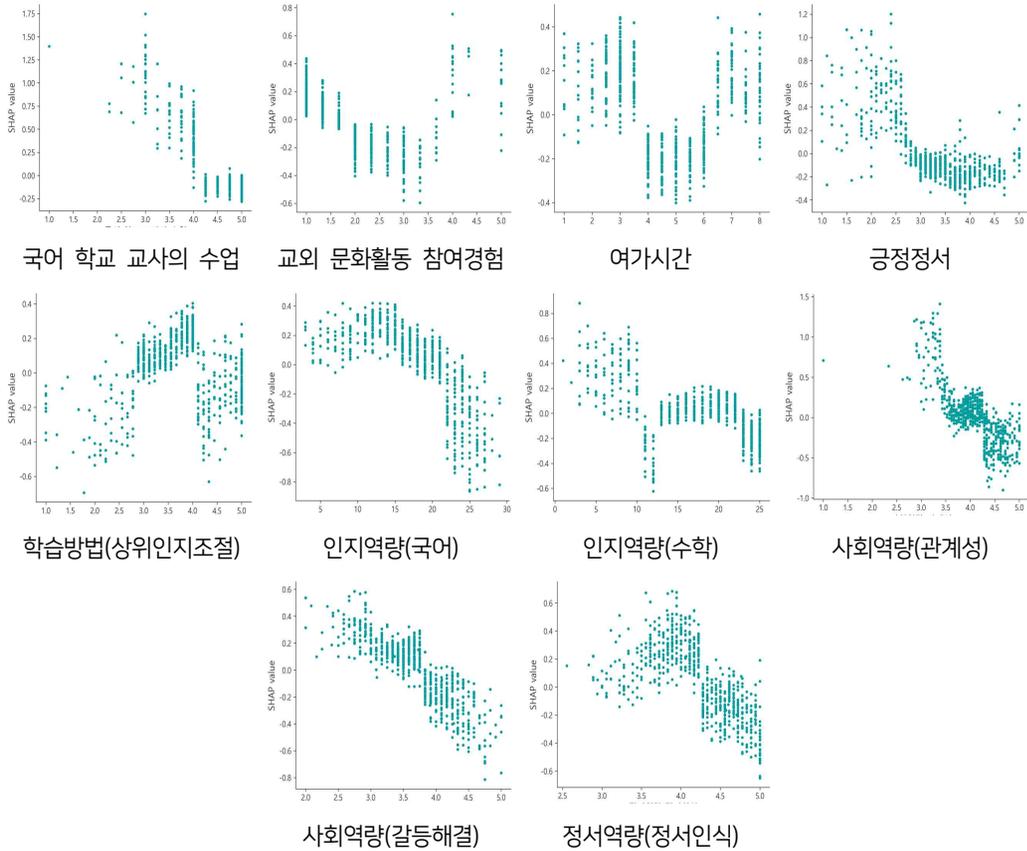


[그림 3] 초교 패널 전이 패턴 분류 SHAP 의존성 도표(제시 순서는 <표 13> 초교 패널 결과 참조)

(2) 중학교 패널

[그림 4]에 제시한 바와 같이 주요 예측변인과 응답수준에 따른 경향성을 살펴본 결과, 국어 학교 교사의 수업, 긍정정서, 인지역량(국어, 수학), 사회역량(관계성, 갈등해결), 정서역량(정서인식)은 응답수준이 높아짐에 따라 Shapley value가 감소하는 것으로 나타났으며, 이러한 결과는 이들 변인에 대한 응답수준이 높아질수록 고역량에서 상대적으로 디지털 역량이 낮은 중역량 집단으로 전이하기 보다는 고역량 집단을 유지하는 패턴에 속할 확률이 증가하는 것을 의미한다. 이어서 교외문화활동 참여경험, 여가시간, 학습방법(상위인지조절) 역시 상위 중요도 변인으로 도출되었지만 명료한 경향성을 보이지는 않았다. 교외문화활동 참여경험은 응답수준이 1학기에 1~2번(3점)에 이르기까지는 Shapley value가 감소하여 관심집단 대비 참조집단인 고역량 유지 패턴에 속할 확률이 높아졌다. 그러나 그 이후부터는 데이터 포인트가 드물게 나타나면서 뚜렷한 경향을 보이지는 않았다. 여가시간의 경우 한 주의 평균 여가시간이 4시간 이상~5시간 미만

(5점)에 이르기까지는 Shapley value가 감소하여 참조집단에 속할 가능성이 높아졌지만, 5시간 이상~6시간 미만(6점) 이후에는 Shapley value가 증가하여 관심집단인 중역량 집단으로의 전이 패턴에 소속할 확률이 높아지는 것으로 나타났다.



[그림 4] 중학교 패널 전이 패턴 분류 SHAP 의존성 도표(제시 순서는 <표 13> 중학교 패널 결과 참조)

V. 논의 및 결론

본 연구는 디지털기기 활용 및 디지털 역량 수준에 따른 잠재프로파일을 학교급(초등학교, 중학교)별로 구분하고, 잠재프로파일의 전이 양상을 살펴본 후, 전이 패턴의 분류를 예측하는 변인을 광범위하게 탐색하였다. 이어서 주요 예측변인의 응답수준에 따라 관심집단에 속할 확률이 어떻게 달라지는지에 대한 시각적인 정보를 함께 제시하고자

하였다. 분석결과를 바탕으로 도출한 시사점은 다음과 같다.

첫째, 잠재프로파일 분석을 통하여 초등학교 4학년과 5학년, 중학교 1학년과 2학년 시점 디지털기기 활용 및 역량 잠재프로파일을 유형화한 결과, 두 학교급에 공통적으로 3개 잠재프로파일이 도출되었으나 구체적인 양상과 각 계층의 할당 비율에는 다소 차이가 나타났다. 구체적으로, 초교 패널은 고빈도 고역량(SNS 적극형)(초4 12.4%, 초5 22.9%), 중빈도 고역량(SNS·쇼핑 소극형)(초4 52.5%, 초5 50.9%), 저빈도 저역량(전반적 소극형)(초4 35.1%, 초5 26.2%)의 3개 계층으로 유형화되었으며, 중학교 패널은 고역량(중1 50.3%, 중2 54.6%), 중역량(중1 34.8%, 중2 31.0%), 저역량(중1 14.9%, 중2 14.4%)의 3개 집단으로 분류되는 것으로 나타났다. 즉, 디지털기기 활용 빈도와 디지털 역량 수준에 따라 상·하위 수준으로 구분되는 것은 두 패널이 동일하나, 초교 패널에서는 디지털기기 활용 빈도는 중간 수준이지만 디지털 역량은 고역량 집단과 유사한 중빈도 고역량(SNS·쇼핑 소극형) 집단이 도출된 것이 특징적이다. 관련해서 서울시교육청(2021)이 초등학교 정보교육 시수가 17시간에 불과한 문제를 인식하고, AI 기반 융합 교육과 디지털 역량 강화를 목표로 'AI 기반 융합 혁신미래교육 중장기 발전 계획('21~'25)'을 발표한 점을 고려할 때, 본 연구에서 과반수의 초등학생이 중빈도 고역량(SNS·쇼핑 소극형) 집단에 속한 결과는 학교와 가정에서 제한된 시간 동안 디지털기기를 사용하더라도 학교 교육을 통해 효과적으로 디지털 역량을 함양하고 있음을 시사한다.

또한 초교 패널은 고빈도 고역량(SNS 적극형) 집단 비율이 초4와 초5 시점에 공통적으로 가장 적은 것으로 확인된 반면, 중학교 패널의 경우 고역량 집단의 비율이 중1과 중2 시기 공통적으로 과반수의 학생이 속하는 것으로 나타난 것 역시 두 패널의 차이점이었다. 이는 초등학생 대비 중학생의 스마트폰 보유율이 상대적으로 높은 것으로 나타난 조사결과(여성가족부, 2023)를 고려하면, 중학교로 진학하면서 학생들의 디지털기기 사용 빈도와 역량이 함께 증가하였음을 알 수 있다. 이와 같이 디지털기기 활용과 역량에 있어 초등학생과 중학생에 상이한 유형이 도출된 결과는 디지털기기 활용 및 역량의 다차원적인 속성을 고려하여 학교급별로 차별적이고 개별화된 개입 방안이 필요함을 함의한다.

둘째, 초등학교 4학년에서 5학년으로, 중학교 1학년에서 2학년으로 시간의 경과에 따라 디지털기기 활용 및 역량 잠재프로파일의 전이 패턴을 분석한 결과, 두 학교급 패널에 전반적으로 시간이 경과해도 동일한 잠재프로파일을 유지하는 비율이 가장 높은 것을 확인할 수 있었다. 그러나 두 학교급 패널에서 공통적으로 디지털기기 활용 빈도와 역량이 감소하는 하향 전이보다 증가하는 상향 전이의 비율이 더 높게 나타난 점은 주목할 만하며, 이러한 결과는 학생들의 디지털 역량을 강화할 수 있는 교육적 가능성을 시사한다. 따라서 학교와 교육청 차원에서는 상향 전이의 긍정적 가능성을 바탕으로 디

지텔 역량을 향상시킬 수 있는 디지털 교육 프로그램을 보다 강화하는 것이 중요하다. 그러한 한편으로는 지속적인 모니터링과 체계적인 지원을 통해 학생들의 디지털 역량 격차를 줄이는 것 역시 병행되어야 한다.

셋째, 초교 패널과 중학교 패널의 디지털기기 활용 및 역량 프로파일에서 질적으로 상이한 계층이 분류됨에 따라, XGBoost 분류 모델 구축에 앞서 참조집단과 관심집단을 학교급에 따라 차별적으로 설정하였다. 이 과정에서, 전이 패턴뿐만 아니라 사례수를 함께 고려하여 참조집단과 관심집단을 선정하였으며, 초교 패널에서는 디지털기기 활용 빈도의 차이가 보다 명확하게 나타났기 때문에 빈도 변화에 초점을 맞추어 전이 패턴을 설정하였고, 중학교 패널에서는 디지털 역량 수준의 차이가 두드러져 디지털 역량 변화를 중심으로 전이 패턴을 정하였다.

먼저 초교 패널에서는 중빈도 고역량(SNS·쇼핑 소극형) 집단을 유지하는 전이 패턴을 참조집단으로, 중빈도 고역량(SNS·쇼핑 소극형) 집단에서 디지털 역량은 유사하되 디지털기기 활용 빈도가 높아지는 고빈도 고역량(SNS 적극형)으로의 빈도 상향 전이 패턴을 관심집단으로 설정하여 두 전이 패턴 분류에 영향을 미치는 핵심 변인을 도출하였다. 상대적으로 명료한 경향성을 보인 변인을 중심으로 살펴보면, 먼저 여학생인 경우, 목표의식, 학업소진(정서적 고갈, 냉소)이 높을수록 상대적으로 디지털기기 활용 빈도가 높은 고빈도 고역량(SNS 적극형) 집단으로의 전이 패턴에 속할 확률이 높아졌다. 이러한 결과는 여성인 경우 디지털기기 활용 빈도가 더 높은 것으로 보고한 홍원준(2020)의 연구와 맥을 함께한다. 또한 고빈도 고역량(SNS 적극형) 집단과 중빈도 고역량(SNS·쇼핑 소극형) 집단을 비교함에 있어 고빈도 고역량(SNS 적극형) 집단의 ‘소셜 네트워크 활동’ 활용 빈도가 약 3점이 더 높은 결과로 미루어 보아 여학생의 SNS활동 및 채팅이 남학생보다 더 활발한 것으로 보고된 보도자료(세계일보, 2022.3.31.)와도 일치하는 것으로 이해 가능하다. 더불어 학업소진과 스마트폰 과이용 간 정적 상관이 있는 것으로 나타난 정인경과 김정현(2017)과 유사한 결과이다. 한편 목표의식의 경우 디지털기기 활용 및 역량 연구에 있어 비교적 관심이 부족하였던 새로운 변인이라는 점에서 주목할 만하다. ‘나는 목표를 이루기 위해 열심히 노력하고 있다’ 등의 문항으로 측정된 목표의식은 응답수준이 높아질수록 중빈도 고역량(SNS·쇼핑 소극형)에서 고빈도 고역량(SNS 적극형)으로의 전이 패턴에 속할 확률이 높아졌다. 중빈도 고역량(SNS·쇼핑 소극형) 집단보다 고빈도 고역량(SNS 적극형) 집단에서 학교 숙제를 목적으로 하는 디지털기기 활용 빈도 역시 더 높게 나타난 점을 고려하면, 목표를 위한 정보 탐색과 학습 도구로서의 디지털기기 활용이 점차 활성화되는 것으로 해석할 수 있다. 따라서 목표 지향적 학습 환경이 학생들의 디지털기기 활용을 긍정적으로 유도할 수 있음을 인지하여 학교에서는 명확한 목표를 설정하고 달성할 수 있도록 돕는 프로그램을 제공할 필요가

있다. 또한 학업에 대한 심리적 피로로 인하여 디지털기기 사용으로의 회피 행동으로 이어지지 않도록 대처 전략이 함께 요구된다.

한편 보호자와의 관계, 독서량, 인지역량이 높을수록 참조집단인 중빈도 고역량(SNS·쇼핑 소극형) 집단의 유지 패턴에 속할 가능성을 높이는 것으로 나타났다. 이는 부모님과 과의 관계가 스마트폰 과사용을 조절하는 것으로 나타난 김경미와 염유식(2014)의 연구와 더불어 초등학생의 학업성취 만족도가 스마트폰 의존에 영향을 미치는 것으로 보고된 박소영과 정혜원(2022)의 연구와 유사한 맥락이다. 더불어 초교 패널 상위 중요도 예측변인 간의 관계성 분석을 위하여 상관분석을 수행한 결과(부록 <표 18> 참조), 보호자와의 관계 변인의 경우 학습방법(행동조절), 목표의식 등과는 정적 상관을, 학업소진과는 부적인 상관을 보이는 것으로 나타났다. 이를 미루어 보면 보호자와의 원활한 관계 형성은 이후 시점의 디지털기기 활용 빈도를 조절하는 중요한 보호요인으로 작용할 뿐만 아니라, 보호자의 정서적 지지 및 학업에 대한 관심이 자녀의 학업과 관련된 정서적 경험과도 관련이 있을 가능성을 시사한다. 따라서 학교에서는 부모 교육 프로그램을 통해 자녀와의 긍정적인 관계 형성을 지원하는 한편으로, 가정에서는 보호자가 자녀의 건강한 학습 습관을 길러줄 수 있도록 돕는 환경을 조성한다면 자녀가 디지털기기를 균형있게 활용할 수 있도록 유도하는 데 중요한 역할을 할 것으로 사료된다. 한편 독서량은 선행연구에서 비교적 다루어지지 않았던 새로운 변인으로 도출됨에 따라 디지털기기 활용과 역량 연구에 있어 주목할 필요가 있다. 관련해서 서울시교육청에서는 학생의 성장단계에 적합한 독서교육을 목적으로 「독서·토론·인문소양교육 기본 계획」을 매해 발표해오고 있다(서울특별시교육청, 2024). 이를 근거로 학교에서는 디지털기기 활용 빈도가 높아지지 않도록 예방하기 위한 방안 중 하나로 독서에 대한 흥미를 높여 독서 습관을 형성하도록 독려하는 것으로 고려해본다면 디지털 역량은 높은 수준으로 유지하면서 자연스럽게 디지털기기 사용은 줄일 수 있을 것으로 기대한다.

넷째, 중학교 패널에서 고역량 집단을 유지하는 전이 패턴을 참조집단으로, 고역량에서 중역량으로의 하향 전이 패턴을 관심집단으로 설정하여 두 전이 패턴의 분류 모델을 수행하였다. 먼저 사회정서역량과 긍정정서, 그리고 인지역량이 높을수록 디지털 역량 수준이 가장 높은 고역량 집단을 유지할 확률이 높아지는 것으로 나타났으며, 이는 학업성취 수준이 높을수록 디지털 역량이 증진되는 것으로 나타난 선행연구(이경숙, 2023; 이은지 외, 2021)와 맥락을 함께한다. 이처럼 인지, 사회, 정서역량 모두가 디지털 역량 잠재프로파일의 전이 분류를 예측하는 핵심 변인으로 도출된 것은 교육적으로 매우 고무적이며, 단순히 특정 역량만을 강화하는 접근이 아니라 학생들의 전반적인 균형 잡힌 역량 발달이 디지털 역량 강화에 기여할 수 있음을 보여준다. 중학교 패널 상위 중요도 예측변인 간 상관분석 결과를 살펴보면(부록 <표 19> 참조) 인지역량과 사회역량, 정서

역량 간의 정적인 상관관계를 확인할 수 있으며, 이는 학생들의 전반적인 역량이 상호유기적으로 연결되어 있음을 시사한다. 따라서 학교 현장에서는 디지털 역량 함양을 위한 디지털 활용 교육은 기기 조작법이나 정보검색과 같은 단순한 기술 습득에 국한되지 않고, 문제 해결력, 협력적 의사소통, 감정 조절과 같은 다양한 역량과 연계될 수 있도록 설계되어야 한다. 이를 위해 학교 교육에서는 디지털 학습 환경에서 협력적 탐구 활동을 강화하고, 정서적 자기조절 및 사회적 상호작용을 포함한 프로그램을 개발하여 학생들이 인지·사회·정서 역량을 균형 있게 성장시키면서 궁극적으로 디지털 역량을 보다 효과적으로 함양할 수 있도록 지원할 필요가 있다.

한편 국어 교사가 수업을 열심히 하고 학생들의 이해 정도를 확인한다고 높게 인식할수록 고역량 집단의 유지 패턴에 속할 확률이 높아졌다. 즉, 학생들의 디지털 역량의 함양에 있어서 교사의 수업의 질이 중요한 영향을 미칠 수 있음을 고려하여, 교육청 차원에서는 디지털기기 활용 우수 수업 사례를 보급하고 교사 대상의 디지털 역량 교육 연수 등의 노력을 통하여 교사의 수업 역량을 높이는 것이 중요하다. 더불어, 기존 연구에서 상대적으로 주목받지 못했던 교외 문화활동 참여 경험 변인이 일정 수준까지 디지털 역량이 높은 집단을 유지하는 데 긍정적인 영향을 미치는 것으로 나타난 점은 주목할 만한 결과이다. 이는 학교 차원에서 교과 학습과 디지털 기술 교육을 보완하는 방식으로 교외 문화활동 참여를 독려하고, 이를 정규 교육과정과 연계하는 방안을 모색할 필요성을 시사한다. 다만, 일정 수준 이상에서는 데이터 포인트가 희박해지면서도 Shapley value가 미약하게 증가하는 경향을 보였기 때문에, 단순히 교외 문화활동의 빈도를 증가시키는 데 초점을 맞추기보다는, 디지털 기술의 실제 활용 사례를 관찰하거나 교외 활동에서 얻은 경험을 디지털 환경에서의 활용 역량 개발로 연결할 수 있는 통합적이고 체계적인 교육 설계가 병행되어야 할 것이다.

본 연구는 디지털기기 활용 및 디지털 역량 잠재프로파일을 학교급별로 분류하고, 이에 대한 전이 양상을 함께 분석하였다는 점에서 횡단적인 관점에서 잠재프로파일 유형화를 시도한 선행연구와는 차별성이 있다. 또한 머신러닝 기반의 XGBoost와 설명 가능한 인공지능 알고리즘 중 하나인 SHAP을 적용하여 전이 패턴의 분류에 미치는 주요 예측변인을 탐색할 뿐만 아니라, 예측변인의 응답값에 따른 경향성을 시각적으로 확인함으로써 초, 중학생의 디지털기기 활용 및 역량 교육을 위한 기초자료를 제공하였다는 점에서 의의가 있다. 특히 초등학교와 중학교에서의 디지털기기 활용 및 역량 프로파일에 있어 질적으로 상이한 계층이 분류됨에 따라 참조집단과 관심집단을 학교급에 따라 차별적으로 설정하여, 초등학교에서는 디지털기기 활용 빈도가 높아지지 않도록 하는, 중학교에서는 디지털 역량이 감소하지 않도록 하는 교육적 시사점을 별도로 제시하였다는 점에서 연구의 의의가 있다.

한편, 본 연구는 서울학생종단연구2020 자료를 기반으로 수행되었으며, 서울 지역에 국한된 표본을 사용하였다는 점에서 연구 결과의 해석과 일반화에 유의할 필요가 있다. 이에 향후 연구에서는 보다 다양한 지역과 표본을 포함한 데이터를 활용하여 연구 결과의 외적 타당성을 강화할 필요가 있다. 또한 본 연구에서는 분석의 안정성 제고를 위하여 전이 패턴 분류 모델 수행에 필요한 참조집단으로 사례수가 가장 많은 중빈도 고역량(SNS·쇼핑 소극형) 유지 패턴(초교 패널), 고역량 유지 패턴(중학교 패널)을 설정하였다. 특히 교육적으로 바람직한 전이 패턴(예. 초4 시기 고빈도 고역량(SNS 적극형) 집단에서 초5 시기 중빈도 고역량(SNS·쇼핑 소극형) 집단으로 전이하여, 역량은 유지하되 디지털기기 사용 빈도가 낮아지는 경우)에 대한 분류를 선행하였으나 사례수가 적어 분류 모델에 과적합 문제가 발생하는 것을 확인함에 따라, 전이 패턴에 따른 분석이 제한되었다는 아쉬움이 있다. 따라서 후속 연구에서는 디지털 역량이 높아지는 상향 전이 패턴(예. 초4 시기 저빈도 저역량(전반적 소극형) 집단에서 초5 시기 중빈도 고역량(SNS·쇼핑 소극형) 집단으로의 전이, 중1 시기 저역량에서 중역량, 고역량 집단으로의 전이) 등 다양한 전이 패턴을 분석할 수 있도록 데이터 축적과 함께 데이터 불균형 문제를 완화할 수 있는 보다 발전된 방법론적 접근 활용이 요구된다.

참고문헌

- 강유림, 김소희(2024). 전환기 교육소의 청소년의 긍정심리 역량 유형 및 영향 요인 탐색: 잠재전이분석과 랜덤 포레스트 분석 활용. *열린교육연구*, 32(1), 73-93.
- 교육부(2022a). 디지털 시대의 주인공이 될 100만 인재를 양성합니다. - 22일 국무회의서 「디지털 인재양성 종합방안」 보고 -. 교육부 보도자료(2022.8.22.).
- 교육부(2022b). 초·중등학교 교육과정 총론(교육부 고시 제2022-33호[별책 1]). 세종: 교육부.
- 권순정, 김아미, 정해진, 안영은(2022). 미래교육을 위한 학교 운영 방안: 혁신미래학교의 디지털 기반 교육 사례를 중심으로(서교연 2022-64). 서울특별시교육청교육연구정보원.
- 길혜지, 이상훈, 황정미(2022). 디지털 역량 교육을 위한 프레임워크 개발 연구. *핵심역량 교육연구*, 7(2), 1-24.
- 길혜지, 박지민, 김훈(2023). 잠재프로파일 분석을 활용한 학생의 디지털 리터러시 유형 분류와 영향요인 분석: PISA 2018을 중심으로. *교육문화연구*, 29(3), 351-374.
- 김경미, 염유식(2014). 청소년의 스마트폰 이용과 주관적 행복: 부모관계의 조절효과에 대한 성별 접근. *정보사회와 미디어*, 15(3), 31-56.
- 김경희, 노언경, 상경아(2021). ICILS 2018 상위 5개국의 컴퓨터·정보 소양과 컴퓨팅 사고력 수준에 따른 잠재 프로파일 도출 및 영향요인 비교. *교육평가연구*, 34(4), 649-678.
- 김미림, 홍민주(2022). 잠재전이분석과 랜덤 포레스트를 활용한 전환기 청소년의 자기결정성 동기 발달 유형과 영향 요인 탐색. *교육과정평가연구*, 25(4), 89-116.
- 김종민, 김한성, 임현정, 안성훈(2018). 다층모형을 활용한 초·중학생 ICT 리터러시 수준 영향요인의 공통점과 차이점 분석. *교육정보미디어연구*, 24(2), 415-440.
- 김혜숙(2010). 청소년의 컴퓨터 사용에 관한 종단적 추이 분석. *아시아교육연구*, 11(2), 93-112.
- 남창우, 신수영(2014). 학생의 ICT 관련 변인이 ICT 활용에 대한 태도 및 문제해결력에 미치는 영향. *교육평가연구*, 27(5), 1265-1286.
- 노언경(2021). 중학생의 ICT 활용과 태도에 따른 유형별 컴퓨터·정보 소양 및 컴퓨팅 사고력의 차이. *교육문화연구*, 27(4), 381-400.
- 박소영, 정혜원(2022). 초등학생의 스마트폰 의존도 관련 주요 설명변수 탐색: 별점화 회귀모형 적용. *아시아교육연구*, 23(1), 121-143.

- 박인우(2024). 디지털 네이티브, 교육. KERIS 디지털교육 동향, 201, 10-13.
- 박현정, 하여진, 박민호(2011). ICT 활용 유형에 따른 학습자 특성 및 성취도에 대한 혼합 모형 분석. *교육평가연구*, 24(2), 733-754.
- 배상률, 이창호, 김남두(2021). 청소년 미디어 이용 실태 및 대상별 정책 대응방안 연구 II: 10대 청소년(연구보고 21-R17). 세종: 한국청소년정책연구원.
- 서울특별시교육청(2021). AI 기반 융합 혁신미래교육 중장기 발전 계획 발표. 서울특별시교육청 보도자료(2021.2.9.).
- 서울특별시교육청(2024). 협력적 독서·인문 교육 활성화 2024 독서·토론·인문소양교육 기본 계획. 서울특별시교육청.
- 성은모, 최효선(2016). 잠재계층분석(Laten Class Analysis)을 활용한 중학생의 ICT 리터러시 유형 및 특성 탐색. *교육공학연구*, 32(4), 987-1013.
- 세계일보(2022.3.31.). 청소년 중 여학생의 스마트폰 과의존, 남학생보다 높아. <https://www.segye.com/newsView/20220331509465>(검색일: 2025.1.2.)
- 신은진, 손원숙(2024). 중·고등 학교급 전환기 자기조절학습 잠재프로파일 전이양상과 영향요인 탐색. *교육과정평가연구*, 27(2), 299-322.
- 여성가족부(2023). 인터넷·스마트폰 과의존 청소년 5년 만에 감소. 여성가족부 보도자료(2023.5.29.).
- 이경숙(2023). 중학생의 디지털 리터러시 태도에 영향을 미치는 학생 요인에 대한 연구. *교육과정평가연구*, 26(1), 103-131.
- 이쌍철, 김혜영, 김혜진, 홍창남(2024). 중학생의 디지털 기기 활용 유형에 따른 학업, 심리 및 행동 특성: 부산교육중단연구(BELS)를 활용하여. *학습자중심교과교육연구*, 24(4), 939-955.
- 이우현, 임상호(2013). 가족의 커뮤니케이션 유형이 아동의 미디어 이용에 미치는 영향에 관한 연구. *디지털융복합연구*, 11(1), 173-179.
- 이은지, 이윤희, 김은영, 추지운(2021). 중학생의 디지털 역량에 따른 잠재계층 분류 및 영향요인 분석. *미래교육학연구*, 34(1), 155-176.
- 장재홍, 고유정(2021). 청소년의 미디어 비판적 사고 능력의 종단적 변화 및 영향요인 탐색. *교육정보미디어연구*, 27(4), 1429-1456.
- 장지현(2023). 대학생 디지털역량 분석: A대학 사례를 중심으로. *핵심역량교육연구*, 8(2), 27-45.
- 정수정, 최고은(2022). 중학생의 ICT 활용유형과 ICT 리터러시에 따른 잠재집단 분류 및

- 창의성, 학업성취와의 관계. *교육정보미디어연구*, 28(4), 1099-1120.
- 정인경, 김정현(2017). 중학생들의 학업스트레스와 학업소진이 스마트폰 중독에 미치는 영향. *한국지역사회생활과학회지*, 28(2), 289-300.
- 홍원준(2020). 스마트폰 사용 목적에 따른 잠재집단과 학업태도와의 관계 성별을 중심으로. *교육정보미디어연구*, 26(3), 567-590.
- Asparouhov, T., & Muthen, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 329-341.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (pp.785-794).
- Collins, L. M., & Lanza, S. T. (2013). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Hoboken, NJ: John Wiley & Sons.
- Ferrari, A., Punie, Y., & Redecker, C. (2012). Understanding digital competence in the 21st century: An analysis of current frameworks. *In European Conference on Technology Enhanced Learning*, (pp.79-92).
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56-67.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 4765-4774.
- Muthen, L. K., & Muthen, B. O. (1998-2017). *Mplus users guide*(8th ed.). Los Angeles, CA: Muthen & Muthen.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1-67.
- Zyberaj J, Bakac C & Seibel, S. (2022). Latent transition analysis in organizational psychology: A simplified “how to” guide by using an applied example. *Frontiers in Psychology*, 13, 977378.

논문접수 : 2025. 1. 3. / 수정본접수 : 2025. 2. 4. / 게재승인 : 2025. 2. 11.

ABSTRACT

Exploring Transition Patterns and Key Predictors of adolescents' Digital Usage and Competence: Applying XGBoost and Latent Transition Analysis

Jang, Eunah

Doctoral Student, Chungnam National University

Chung, Hyewon

Professor, Chungnam National University

The purpose of this study is to classify the latent profiles and transition patterns of digital device usage and digital competence across elementary and middle school students, as well as to explore key predictors influencing these transitions. To this end, latent transition analysis, XGBoost, and SHAP(Shapley Additive Explanations), an explainable artificial intelligence algorithm, was applied to 1st and 2nd wave data of the elementary school students' panel($n=3,464$) and middle school students' panel($n=3,556$) of Seoul Educational Longitudinal Study of Students 2020(SELS 2020). The main findings are as follows. First, latent profiles were classified into three distinct groups across both educational levels, with differences observed in profile characteristics and proportions between panels. Second, while most students remained stable in their profiles, upward transitions in digital device usage and competence were more prevalent than downward transitions. Third, in the elementary school students' panel, gender, guardians relationships, reading volume, academic burnout, among others were identified as key predictors. Fourth, in the middle school students' panel, instructional quality in Korean language education, participation in extracurricular cultural activities, competence (cognitive, social, emotional), among others were derived as key predictors. Based on the findings, implications and recommendations for managing digital device usage and fostering digital competence in students are presented.

Key Words: Digital Device Usage, Digital Competence, Adolescent, XGBoost, Latent Transition Analysis, SHAP