

Criterion-Related Validity and Temporal Stability of the Affective Domain of Digital Literacy Scale (ADLS): A Three-Wave Longitudinal Study Using Metacognitive Self-Regulation as an External Criterion

Eun-Ji Kim*, Young-Jun Lee**

*Student, Dept. of Computer Education, Korea National University of Education, Cheongju, Korea

**Professor, Dept. of Computer Education, Korea National University of Education, Cheongju, Korea

[Abstract]

This study examined the criterion-related validity and temporal stability of the Affective Domain of Digital Literacy Scale (ADLS) using a three-wave longitudinal design, with metacognitive self-regulation serving as an external criterion. A total of 157 fifth-grade students completed the ADLS and a metacognitive self-regulation measure at three time points: pretest (April), midpoint (September), and posttest (December). Data were analyzed using Pearson correlations, cross-time correlations, and mixed-design repeated-measures analysis of variance (ANOVA). The ADLS total score demonstrated significant positive correlations with metacognitive self-regulation across all three time points ($r = .599 - .641$). The magnitude of these correlations showed no significant differences over time, indicating consistent criterion-related validity. Regarding temporal stability (test-retest reliability), the ADLS total score and the Self-Efficacy subfactor exhibited adequate stability, whereas the Value Perception subfactor showed relatively lower stability. Furthermore, repeated-measures ANOVA revealed that ADLS scores increased significantly from the pretest to the midpoint and remained stable thereafter. In contrast, metacognitive self-regulation showed no significant changes over the same period. These findings provide empirical evidence for the criterion-related validity and temporal stability of the ADLS, supporting its utility as a reliable repeated-measurement tool.

▶ **Key words:** Affective Domain of Digital Literacy Scale (ADLS), metacognitive self-regulation, criterion-related validity, temporal stability, longitudinal study

• First Author: Eun-Ji Kim, Corresponding Author: Young-Jun Lee
*Eun-Ji Kim (153comz@gmail.com), Dept. of Computer Education, Korea National University of Education
**Young-Jun Lee (yjlee@knue.ac.kr), Dept. of Computer Education, Korea National University of Education
• Received: 2026. 02. 20, Revised: 2026. 03. 22, Accepted: 2026. 03. 24.

[요 약]

본 연구는 디지털 소양 정의적 영역 척도(ADLS)의 준거관련 타당도 및 시간적 안정성 근거를 확보하기 위해 메타인지적 자기조절을 외부 준거로 설정하여 3시점 종단 분석을 수행하였다. 초등학교 5학년 157명을 대상으로 사전(4월), 중간(9월), 사후(12월)에 ADLS와 메타인지적 자기조절을 반복측정하고 시점별 상관, 시점 간 상관, 혼합설계 반복측정 분산분석을 실시하였다. 그 결과 ADLS 총점과 메타인지적 자기조절은 모든 시점에서 유의한 정적 상관을 보였으며($r = .599 - .641$) 시점 간 상관 크기의 차이는 유의하지 않아 ADLS 총점의 준거관련 타당도가 시점에 걸쳐 일관되게 유지되었다. 하위요인 중 실천경험과 자기효능감은 시점에 걸쳐 안정적 상관을 보였으나 가치인식은 사후 시점에서 관련성이 약화되는 경향을 보였다. 시간적 안정성(순위안정성)은 ADLS 총점과 자기효능감에서 적절한 수준이었으나($r = .550 - .748$) 가치인식은 상대적으로 낮았다. 반복측정 분산분석 결과 ADLS는 사전에서 중간으로 유의하게 증가한 후 유지된 반면 메타인지적 자기조절은 유의한 변화를 보이지 않아 해당 연령대에서 비교적 안정적인 특성임을 시사하였다. 이상의 결과는 ADLS가 준거관련 타당도와 시간적 안정성 근거를 갖추고 있음을 시사한다.

▶ **주제어:** 디지털 소양 정의적 영역, ADLS, 메타인지적 자기조절, 준거관련 타당도, 검사-재검사 신뢰도, 종단 연구

I. Introduction

1. Research Background

2022 개정 교육과정은 디지털 기초소양을 언어·수리와 더불어 모든 학생의 학습 기초로 명시하고 있으며 이에 따라 디지털 소양의 체계적 진단과 평가에 대한 요구가 높아지고 있다. 디지털 소양은 단순한 기술적 조작 능력을 넘어 정보의 이해·활용·창작·참여를 포함하는 실천적 역량으로 논의되며 이러한 수행이 실제 학습과 문제 해결로 이어지기 위해서는 자기효능감, 가치인식, 실천경험과 같은 정의적 기반이 뒷받침될 필요가 있다[1, 2].

이러한 맥락에서 선행연구[3]는 초등학생의 디지털 소양 정의적 영역을 측정하기 위한 척도(Affective Domain of Digital Literacy Scale; ADLS)를 개발하고 탐색적·확인적 요인분석을 통해 실천경험(E), 자기효능감(Ef), 가치인식(V)의 3요인 20문항 구조를 확인하였다. 이를 통해 ADLS는 내용타당도와 내적 구조에 대한 구인타당도 근거를 갖추게 되었다[3].

그러나 타당도는 고정된 속성이 아닌 점수 해석과 사용을 뒷받침하는 다원적 근거가 누적적으로 축적됨으로써 강화된다[4]. 이 관점에서 구인타당도와 내용 타당도만으로는 도구의 타당도가 충분히 확립되었다고 보기 어렵다. ADLS의 타당도 근거를 보다 폭넓게 확보하기 위해서는 두 가지 추가적인 검토가 필요하다. 첫째, 이론적으로 관련된 외부 구인과의 준거관련 타당도(criterion-related validity) 근거이다. 둘째, 시간의 흐름에 따라 측정 점수

가 일관되게 유지되는지를 확인하는 시간적 안정성(temporal stability) 근거이다. 시간적 안정성은 개인 간 순위가 시점에 걸쳐 유지되는 순위 안정성(rank-order stability; 시점 간 상관으로 추정되는 검사-재검사 신뢰도)과 집단의 평균 수준이 시점에 따라 어떻게 변화하는지를 확인하는 평균 수준 변화(mean-level change)의 두 관점에서 점검될 수 있으며[5] 특히 초등학생의 정의적 특성이 발달 과정에서 상황에 따라 변동할 가능성을 고려할 때 종단 자료를 통한 경험적 검증이 요구된다[5, 6].

2. Research Necessity and Purpose

준거 관련 타당도 근거를 확보하기 위한 외부 준거 변인으로는 학업적 자기효능감, 자기주도적 학습능력, 학업 성취도, 디지털 수행 평가 등 여러 구인이 고려될 수 있다. 본 연구에서는 다음과 같은 근거에서 메타인지적 자기조절을 외부 준거로 선정하였다. 첫째, 디지털 도구 활용은 전략의 선택·점검·수정을 수반하는 자기조절적 성격을 지니므로 메타인지적 자기조절과 이론적으로 관련되면서도 [6-8] ADLS의 하위 구인(실천경험, 자기효능감, 가치인식)과 구인 수준에서 명확히 구별된다. 둘째, 초등학생 대상으로 신뢰도와 타당도가 확보된 측정 도구가 존재하여[9] 실증적 검증이 가능하다. 셋째, 메타인지적 자기조절은 학업 성취와 일관된 관련성이 보고되어 왔으며[10-12] 이를

외부 준거로 설정함으로써 ADLS 점수가 학습 관련 자기 조절 능력과 관련되는지 확인할 수 있다. 다만 본 연구에서는 단일 준거만을 사용하였으므로 후속연구에서 다양한 외부 준거를 포함하여 수렴 및 변별 타당도 근거를 확장할 필요가 있다.

메타인지적 자기조절은 학습 과정에서 자신의 인지 활동을 계획·점검·평가하는 조절 능력으로 효과적인 학습을 가능하게 하는 핵심 기제로 간주된다[6-8]. 디지털 도구의 활용은 전략의 선택·점검·수정을 수반하는 자기조절적 성격을 지니므로 디지털 학습 환경에서 풍부한 실천경험과 높은 자기효능감 및 가치인식을 가진 학습자일수록 자신의 학습 과정을 보다 적극적으로 점검하고 조절할 것으로 예상된다[6, 7, 9]. 이는 ADLS와 메타인지적 자기조절 간에 유의미한 관련성이 존재할 수 있음을 시사한다[10].

아울러 ADLS의 시간적 안정성 근거는 아직 보고된 바 없다. 반복 측정된 점수가 시간에 걸쳐 어느 정도 일관성을 유지하는지를 확인하는 것은 도구의 신뢰도를 판단하는 데 필수적이다. 또한 ADLS 점수의 평균 변화 양상을 확인함으로써 초등학생의 디지털 소양 정의적 영역이 시간에 따라 어떠한 변화를 보이는지에 대한 기초 자료를 제공할 수 있다[10].

이에 본 연구는 초등학교 5학년 학생을 대상으로 3시점에 걸쳐 수집된 종단 자료를 활용하여 ADLS의 준거관련 타당도 및 시간적 안정성 근거를 제시하고자 한다.

3. Research Questions

본 연구의 연구 문제는 다음과 같다.

- (1) 각 시점에서 ADLS(총점 및 하위요인)와 메타인지적 자기조절 간 상관관계는 유의한가, 그리고 그 상관관계의 크기는 시점에 걸쳐 일관되게 유지되는가?
- (2) ADLS(총점 및 하위요인)와 메타인지적 자기조절의 각 변인별 시점 간 상관(순위 안정성)은 어떠한가?
- (3) ADLS(총점 및 하위요인)와 메타인지적 자기조절의 평균은 3시점에 걸쳐 어떠한 변화 양상을 보이는가?

II. Theoretical Background

1. Digital Literacy and the Affective Domain

디지털 소양(Digital Literacy)은 디지털 기술과 정보를 이해하고 활용하여 문제를 해결하고 사회적 참여를 실천하는 역량으로 정의된다[8]. 초기에는 기술적 조작 능력에 초점을 두었으나 이후 정보의 비판적 이해, 윤리적 활용,

협력적 창작 등을 포함하는 다차원적 개념으로 확장되어 왔다[7, 13]. 국내외에서는 UNESCO, DigComp 등 디지털 역량 프레임워크를 통해 디지털 소양의 구성 요소와 수준을 체계적으로 제시해 왔으며[14, 15], 2022 개정 교육과정에서도 디지털 기초소양이 범교과적 학습 기반으로 명시되면서 초등 단계에서의 디지털 소양 교육과 평가에 대한 관심이 높아지고 있다[7, 13].

전통적으로 정의적 영역은 Krathwohl의 위계(감수-반응-가치화-조직화-인격화)와 같이 내면화의 단계적 심화로 개념화되어 왔다[2]. 그러나 이러한 위계적 분류는 특정 교과 맥락에서 학습자의 정의적 상태를 진단하기에는 추상성이 높다는 한계가 있다. ADLS의 실천경험, 자기효능감, 가치인식은 위계적 단계가 아닌 디지털 소양의 정의적 기반을 구성하는 세 가지 병렬적 차원(행동 기반의 경험 인식, 수행에 대한 자기 신념, 활동의 유용성 판단)으로 구조화된 점에서 기존 분류와 구별된다.

디지털 소양의 효과적인 발현은 인지적 능력뿐 아니라 정의적 요인에 의해서도 좌우된다[16, 17]. 학습자가 디지털 도구를 활용할 때 갖는 자기효능감, 관련 경험에 대한 인식, 디지털 활동의 가치에 대한 판단은 실제 수행의 질과 지속성에 영향을 미치는 것으로 보고된다[17-19]. 그러나 기존의 디지털 소양 측정 도구는 주로 인지적·기능적 측면에 초점을 맞추고 있어 특히 초등학생을 대상으로 정의적 하위요인을 체계적으로 분리·측정하는 도구는 제한적이다[17, 20].

이에 선행연구[3]는 초등학생의 디지털 소양 정의적 영역을 측정하기 위한 ADLS를 개발하였다. ADLS는 실천경험(E), 자기효능감(Ef), 가치인식(V)의 3요인 20문항으로 구성되며 탐색적 요인분석과 확인적 요인분석을 통해 구인타당도 근거가 확인되었다. 이 3요인은 경험(행동 기반), 효능감(신념), 가치(평가)의 축에서 디지털 소양의 정의적 기반을 포착하는 구조이다. 실천경험은 디지털 도구를 활용한 학습 및 문제 해결 경험에 대한 인식을, 자기효능감은 디지털 과제를 성공적으로 수행할 수 있다는 자신감을, 가치인식은 디지털 소양 활동의 유용성과 중요성에 대한 판단을 각각 측정한다.

2. Metacognitive Self-Regulation

메타인지(metacognition)는 자신의 인지 과정에 대한 인식과 조절을 포괄하는 개념으로, 일반적으로 인지에 대한 지식(knowledge of cognition)과 인지의 조절(regulation of cognition)의 두 하위 차원으로 구분된다[21-23]. 이 중 인지의 조절에 해당하는 메타인지적 자기

조절(metacognitive self-regulation)은 학습 과정에서 자신의 인지 활동을 계획하고 진행 과정을 점검하며 결과를 평가하는 능동적 조절 활동을 의미한다[22]. 본 연구에서는 이 메타인지적 자기조절에 초점을 둔다[23, 24].

메타인지적 자기조절은 자기조절학습(self-regulated learning)의 핵심 구성 요소로 제시되며 학업 성취와 유의한 관련성을 보이는 것으로 일관되게 보고되어 왔다[10-12]. 특히 초등학생 대상 연구에서 메타인지적 자기조절 수준이 높은 학습자는 과제 수행 시 보다 적절한 전략을 선택하고 학습 과정을 능동적으로 점검하며 오류에 대해 유연하게 대처하는 것으로 나타났다[10, 23].

디지털 학습 환경에서 메타인지적 자기조절의 역할은 더욱 강조된다[24-26]. 디지털 도구를 활용한 학습은 정보의 탐색, 선별, 조직, 표현 등 다층적인 인지적 처리를 요구하며, 이 과정에서 학습자의 자기점검과 전략 조절이 수반된다[24, 25]. 따라서 디지털 소양의 정의적 영역(실천경험, 자기효능감, 가치인식)은 디지털 학습에서 요구되는 자기점검과 전략 조절을 촉진함으로써 메타인지적 자기조절과 유의미하게 관련될 것으로 예상된다[25, 26].

구체적으로 실천경험이 풍부한 학습자는 다양한 디지털 과제 수행을 통해 전략의 선택과 수정에 관한 절차적 지식을 축적함으로써 메타인지적 조절 전략의 활용이 촉진될 수 있다. 자기효능감이 높은 학습자는 도전적인 디지털 과제에서도 자신의 수행을 적극적으로 점검하고 어려움에 직면했을 때 유연하게 전략을 조절할 가능성이 높다[18, 19]. 가치인식이 높은 학습자는 디지털 활동 자체를 의미 있는 학습 기회로 인식하여 과제에 대한 깊은 처리(deep processing)와 자기점검을 수행할 동기를 갖게 된다[10, 25, 26]. 이처럼 ADLS의 세 하위 요인은 각기 다른 기제를 통해 메타인지적 자기조절과 관련될 것으로 예상되며 이는 메타인지적 자기조절이 ADLS의 외부 준거로서 이론적 적합성을 갖추고 있음을 뒷받침한다.

3. Types of Validity Evidence and Temporal Stability

타당도(validity)는 측정 점수에 기반한 해석과 사용을 뒷받침하는 근거의 정도를 의미한다[27]. 현대 측정학에서 타당도는 고정된 속성이 아니라 다양한 유형의 근거가 누적적으로 축적됨으로써 강화되는 것으로 이해된다[28]. 이러한 관점에서 측정 도구의 타당도를 뒷받침하기 위해서는 내적 구조, 다른 변인과의 관계, 반응 과정, 검사 내용, 결과적 측면 등 다원적 근거를 체계적으로 확보할 필요가 있다[29].

본 연구에서 초점을 두는 근거는 다음과 같다. 첫째, 다른 변인과의 관계(relations to other variables)에 해당하는 준거관련 타당도 근거이다[27]. 둘째, 측정의 시간적 일관성을 보여주는 시간적 안정성(신뢰도)이다[28]. 시간적 안정성은 엄밀히는 신뢰도(재현성) 근거에 해당하나 점수 해석의 타당성을 뒷받침하는 중요한 기반으로 타당도 근거와 함께 점검될 필요가 있다[30].

준거관련 타당도는 측정하고자 하는 구인의 점수가 이론적으로 관련된 외부 준거와 체계적인 관련성을 보이는지를 확인하는 것이다[27]. ADLS의 경우 디지털 소양의 정의적 영역이 학습자의 자기조절 능력과 관련된다는 이론적 근거에 기반하여 메타인지적 자기조절을 외부 준거로 설정할 수 있다[25, 26]. 이때 두 구인 간 상관의 방향과 크기가 이론적 예측과 부합하는지를 확인함으로써 다른 변인과의 관계에 대한 타당도 근거를 확보할 수 있다[27]. 또한 이러한 관련성이 특정 시점에 한정되지 않고 시점에 걸쳐 일관되게 유지되는지를 함께 점검하면 단일 시점의 우연적 공변을 넘어 관련성이 반복 관측에서도 재현되는지를 확인할 수 있다[27].

시간적 안정성은 동일한 도구로 반복 측정한 점수가 시간에 걸쳐 어느 정도 일관성을 유지하는지를 나타내며 도구의 신뢰도를 판단하는 데 중요한 근거가 된다[30]. 시간적 안정성은 두 가지 관점에서 점검될 수 있다. 첫째, 순위 안정성(rank-order stability)은 시점 간 상관계수를 통해 개인 간 상대적 순위가 유지되는 정도를 평가하며 검사-재검사 신뢰도의 핵심 지표로 간주된다[31]. 둘째, 평균 수준 변화(mean-level change)는 집단 평균이 시점에 따라 유의하게 변화하는지를 확인하는 것으로 순위 안정성과는 독립적인 정보를 제공한다[31]. 순위 안정성이 높더라도 집단 전체의 평균수준은 변화할 수 있으며 이와 별개로 평균 수준이 유지되더라도 개인 간 순위는 변동할 수 있다[31, 32]. 따라서 두 관점을 함께 점검하는 것이 시간적 안정성에 대한 보다 종합적인 이해를 제공한다.

특히 초등학생의 정의적 특성은 발달적 변화와 학습 경험에 따라 변동 가능성이 있어 어느 정도의 시간적 안정성이 확인되는지를 경험적으로 검증하는 것은 ADLS의 실용적 활용 가능성을 판단하는 데 중요한 기초 자료가 된다.

III. Research Method

1. Participants

본 연구의 참여자는 K도 소재 초등학교의 5학년 학생이

다. 전체 9개 학급 214명 중 사전 시점에 참여하지 않은 13명을 제외한 201명이 연구에 참여하였으며 이중 3시점(사전, 중간, 사후) 모두에 완전 응답 자료가 확보된 157명을 최종 분석 대상으로 선정하였다(완전사례 분석; listwise deletion). 중간 및 사후 시점에서의 추가 결측은 전학, 결석, 미응답 등으로 발생하였으며 본 연구는 동일 표본을 기반으로 시점 간 비교를 수행하기 위해 완전사례만을 분석에 포함하였다. 체계적 탈락 가능성을 점검하기 위해 사전 시점 참여자 중 3시점 완전사례($n = 157$)와 탈락자($n = 43$)의 사전 시점 점수를 Welch t 검정으로 비교한 결과 ADLS 총점($t = -0.45, p = .657, d = -0.08$)과 메타인지적 자기조절($t = -1.10, p = .275, d = -0.20$) 모두에서 유의한 차이가 관찰되지 않았다. 효과크기 역시 무시할 수준(< 0.20)이었으며, 결측 사유가 전학·결석 등 측정 변인과 무관한 요인에 해당하므로 탈락이 무선에 가까운 것으로 판단하였다. 본 자료는 실험집단(마이크로러닝 및 디지털 배지 기반 디지털 소양 교육 프로그램 참여) 86명과 통제집단(전통적 교수법 수업 참여) 71명으로 구성되었다. 본 연구의 초점은 처치 효과 검증이 아니라 ADLS의 측정학적 특성 검토에 있으나 교육 프로그램의 유무가 척도의 측정학적 특성에 영향을 미칠 가능성을 배제할 수 없으므로 집단 구성을 명시하고 분석 시 집단 변인을 통제하거나 통제집단 단독 보조 분석을 병행하였다. 사전 시점에서 실험집단과 통제집단 간 ADLS 총점($t = 0.03, p = .977$) 및 메타인지적 자기조절($t = -1.23, p = .222$)에 유의한 차이가 관찰되지 않아 두 집단의 사전 동질성이 확인되었다. 성별은 설문지의 인구통계 문항을 통해 자기보고로 수집하였으며 남학생 79명(50.3%), 여학생 78명(49.7%)였다.

2. Research Design and Procedure

본 연구는 사전(4월) - 중간(9월) - 사후(12월)의 3시점에 걸친 종단 설계를 적용하였다. 실험집단에는 중간과 사후 시점 사이(11월, 4주간)에 마이크로러닝 및 디지털 배지 기반 디지털 소양 교육 프로그램을 주 2차시씩 총 8차시 실시하였으며, 통제집단에는 통상적 교수학습 방식에 따른 수업이 진행되었다. 사전-중간 구간(4월-9월)에는 실험집단과 통제집단 모두 통상적 교수학습 방식에 따른 수업이 진행되었으며 이 기간에 집단 간 차별적 처치는 없었다. 각 시점에서 ADLS와 메타인지적 자기조절을 동시에 측정하였다. 본 연구의 초점은 처치 효과 검증이 아닌 ADLS의 준거관련 타당도 및 시간적 안정성 근거 확보에 있었다. 따라서 평균 수준 변화 분석에서는 집단 변인을

혼합설계 반복측정 분산분석의 between-subjects 요인으로 포함하여 시점에 따른 변화가 집단에 따라 차이가 있는 지(집단 × 시점 상호작용)를 보조적으로 점검하였다. 만약 이 상호작용이 유의하다면 교육 프로그램의 유무에 따라 ADLS의 시점 간 변화 궤적이 달라지는 것이므로 준거관련 타당도나 순위 안정성 결과가 처치 조건에 의해 교란되었을 가능성을 시사한다. 반대로 상호작용이 유의하지 않았다면 처치 여부와 무관하게 척도 점수의 변화 양상이 일관됨을 의미하며 이는 본 연구에서 보고하는 타당도 및 안정성 근거의 일반화 가능성을 지지한다.

3. Instruments

3.1 Affective Domain of Digital Literacy Scale (ADLS)

ADLS(Affective Domain of Digital Literacy Scale)는 초등학생의 디지털 소양 정의적 영역을 측정하기 위해 개발된 자기보고식 척도이다. 실천경험(E) 6문항, 자기효능감(Ef) 8문항, 가치인식(V) 6문항의 3요인 20문항으로 구성되며 각 문항은 5점 리커트 척도로 응답한다. 역문항은 포함되지 않으며 총점 및 하위요인 점수는 해당 문항의 평균으로 산출하였다(점수범위: 1 - 5점). 본 연구에서의 시점별 내적 일관성 신뢰도는 결과 절에 제시하였다.

3.2 Metacognitive Self-Regulation

메타인지적 자기조절은 박주연(2005)이 개발한 메타인지 검사 도구(총 40문항: 메타인지적 지식 18문항, 메타인지적 자기조절 15문항, 메타인지적 경험 7문항) 중 메타인지적 자기조절 영역 15문항을 사용하여 측정하였다. 해당 영역은 계획하기 5문항, 모니터링 5문항, 자기평가 5문항으로 구성되며 각 문항은 5점 리커트 척도(1 = 전혀 그렇지 않다, 5 = 매우 그렇다)로 응답한다. 모니터링 하위요인의 14번 문항("나는 종종 수업시간 동안 다른 생각을 해서 중요한 핵심을 놓친다")은 부정문으로 구성된 역문항이므로 원 척도의 채점 지침에 따라 역채점 처리하였다[9]. 점수는 15문항 합산(범위: 15 - 75점)으로 산출하였다. 원 도구에서 해당 영역의 일관성 신뢰도는 .87 - .88로 보고되었으며[9], 본 연구의 시점별 내적 일관성 신뢰도는 결과 절에 제시하였다.

4. Data Analysis

본 연구의 자료 분석은 다음과 같이 수행하였다. 준거관련 타당도 및 순위 안정성 분석을 위한 상관계수는 전체 표본($N = 157$)을 대상으로 산출하였다. 첫째, 준거관련 타당도를 검토하기 위해 각 시점에서 ADLS(총점 및 하위요

인)와 메타인지적 자기조절 간 Pearson 적률상관계수를 산출하였다. 이는 관찰 점수 수준의 분석으로 측정 오차가 분리되지 않으므로 잠재 변수 간 관계의 과소추정 가능성이 존재한다. 시점 간 상관 크기의 차이가 유의한지를 탐색적으로 확인하기 위해 동일 표본에서 비중첩 종속 상관을 비교하는 Steiger(1980)의 검정을 각 변인별 3쌍 비교로 실시하였다. 본 연구에서 비교하는 두 상관은 동일 표본에서 산출되었으나 공통 변수를 공유하지 않으므로 비중첩 종속 상관 비교 검정을 적용하였다. 이 비교는 탐색적 성격이므로 1차적으로 보정 없는 개별 검정 결과를 보고하되 12회 비교에 따른 1종 오류 누적 가능성을 고려하여 Benjamin - Hochberg 보정을 보조적으로 적용한 결과를 함께 제시하였다. 둘째, 시간적 안정성(순위 안정성)을 검토하기 위해 각 변인의 시점 간 Pearson 상관계수(사전-중간, 중간-사후, 사전-사후)를 산출하였다. 셋째, 평균 수준 변화를 확인하기 위해 시점(3)을 within-subjects 요인으로, 집단(2: 실험/통제)을 between-subjects 요인으로 설정한 혼합설계 반복측정 분산분석(mixed-design repeated-measures ANOVA)을 실시하였다. 구형성 가정이 충족되지 않을 경우 Greenhouse - Geisser 보정을 적용하였으며 사후 비교는 Bonferroni 보정을 적용하였다. 본 연구의 초점은 시점의 주효과에 있으며, 집단 × 시점 상호작용 효과를 함께 보고하여 시점 간 변화 양상이 집단에 따라 차이가 있는지를 보조적으로 확인하였다. 다만 본 분석은 종단 측정불변성(longitudinal measurement invariance)을 사전에 검증하지 않았으므로 관찰된 평균 변화는 관찰 점수 수준의 변화로 해석되어야 한다. 모든 분석은 R(version 4.5.2)을 사용하여 수행하였다. Cronbach's 알파는 psych 패키지, 혼합설계 반복측정 분산분석은 afex 및 emmeans 패키지를 이용하였다. Steiger(1980) 검정은 cocor 패키지의 cocor.dep.groups.nonoverlap() 함수를 사용하여 동일 표본 내 비중첩 종속 상관을 비교하였다.

IV. Research Result

1. Descriptive Statistics and Reliability

각 변인의 시점별 기술통계와 내적 일관성 신뢰도 Cronbach's 알파를 Table 1에 제시하였다.

Table 1. Descriptive Statistics and Internal Consistency Reliability by Time Point (N = 157)

Variable	Pre			Mid			Post		
	M	SD	α	M	SD	α	M	SD	α
ADLS Total	3.46	0.60	.915	3.63	0.62	.925	3.69	0.60	.914
E	3.18	0.73	.826	3.37	0.68	.790	3.50	0.67	.796
Ef	3.51	0.75	.872	3.68	0.82	.913	3.72	0.78	.888
V	3.66	0.64	.818	3.83	0.61	.825	3.82	0.63	.825
M	48.55	8.27	.850	49.83	9.46	.900	49.55	9.87	.886

Note. E = Practice Experience; Ef = Self-efficacy; V = Value Perception; M = Metacognitive self-regulation.

Note. ADLS scores are item means (range: 1-5). Metacognitive self-regulation scores are 15-item sums (range: 15-75).

ADLS 총점은 Cronbach's α 는 세 시점 모두 .91 이상으로 높은 수준이었으며 하위요인도 .79 - .91 범위로 양호한 내적 일관성을 보였다. 메타인지적 자기조절의 알파는 .85 - .90으로 적절한 수준이었다.

2. Criterion-Related Validity: Correlations Between ADLS and Metacognitive Self-Regulation by Time Point (RQ 1)

각 시점에서 ADLS(총점 및 하위요인)와 메타인지적 자기조절 간 Pearson 상관계수를 Table 2에 제시하였다.

Table 2. Correlations Between ADLS and Metacognitive Self-regulation by Time Point (N = 157)

Variable	Pre(T1)	Mid(T2)	Post(T3)
ADLS Total × M	.641***	.631***	.599***
Practice Experience(E) × M	.562***	.555***	.551***
Self-efficacy(Ef) × M	.587***	.574***	.587***
Value Perception(V) × M	.463***	.481***	.333***

Note. M = Metacognitive self-regulation. ***p < .001.

ADLS 총점과 메타인지적 자기조절은 모든 시점에서 유의한 정적 상관을 보였으며($r = .599 - .641$, $p < .001$), 상관의 크기는 Cohen(1988)의 기준에 따르면 중간에서 큰 수준에 해당하였다. 하위요인 중 실천경험($r = .551 - .562$)과 자기효능감($r = .574 - .587$)은 시점에 걸쳐 비교적 안정적인 수준의 상관을 보였다. 가치인식은 사전($r = .463$)과 중간($r = .481$)에서 중간 수준의 상관을 나타냈으나 사후 시점($r = .333$)에서는 상대적으로 낮아지는 경향을 보였다.

시점 간 상관 크기의 차이를 탐색적으로 확인하기 위해 Steiger(1980) 검정을 실시한 결과를 Table 3에 제시하였다.

Table 3. Comparison of Correlation Magnitudes Across Time Points: Steiger (1980) Test Results

Variable	T1 vs T2		T2 vs T3		T1 vs T3	
	z	p	z	p	z	p
ADLS Total × M	0.110	.912	0.694	.488	0.716	.474
E × M	0.199	.842	-0.032	.974	0.162	.871
Ef × M	0.047	.962	-0.035	.972	0.013	.990
V × M	-0.259	.796	2.011	.044	1.541	.123

When Benjamini-Hochberg correction was applied across all 12 comparisons, no adjusted p-value reached significance (p_{adj} = .528-.990).

ADLS 총점과 메타인지적 자기조절의 상관은 세 시점 간 비교 모두에서 유의한 차이가 나타나지 않았다. 하위요인 중 실천경험과 자기효능감도 모든 시점 간 비교에서 유의한 차이가 없었다. 가치인식의 경우 중간 대 사후 비교에서 상관 크기의 감소가 유의 수준에 도달하였으나(z = 2.011, p = .044), 다중비교 보정이 적용되지 않은 탐색적 결과이며 사전 대 사후 비교에서는 유의하지 않았다(z = 1.541, p = .123). 따라서 가치인식과 메타인지적 자기조절의 관련성이 사후 시점에서 약화되었을 가능성이 제기되나 Benjamini-Hochberg 보정 적용 시 유의하지 않은 탐색적 결과이므로(p_{adj} = .528) 주의 깊은 해석이 필요하다. 이에 대한 심층적인 논의는 V장에서 다룬다.

3. Temporal Stability(RQ 2, 3)

3.1 Rank-Order Stability(RQ 2)

각 변인의 시점 간 Pearson 상관계수를 Table 4에 제시하였다.

인접 시점 간 상관(사전-중간, 중간-사후)은 비인접 시점 간 상관(사전-사후)보다 전반적으로 높았으며 이는 측정 간격이 길어질수록 순위 안정성이 감소하는 일반적 경향과 부합한다.

Table 4. Cross-Time Correlations: Rank-Order Stability (N = 157)

Variable	Pre-Mid	Mid-Post	Pre-Post
ADLS Total	.664***	.686***	.550***
E	.490***	.494***	.456***
Ef	.678***	.748***	.612***
V	.604***	.538***	.340***
M	.730***	.760***	.647***

Note. E = Practice Experience; Ef = Self-efficacy; V = Value Perception; M = Metacognitive self-regulation. ***p < .001.

메타인지적 자기조절은 인접 시점간 r = .730-.760, 사전-사후 r = .647로 가장 높은 순위 안정성을 보였다. ADLS 총점(r = .550-.686)과 자기효능감(r = .612-.748)은 적절한 수준의 안정성을 나타냈다. 실천경험(r = .456-.494)은 중간수준이었으며 가치인식(r = .340-.604)은 인접 시점에서는 중간 수준의 안정성을 보였으나 사전-사후 간(r = .340)에는 상대적으로 낮은 안정성을 나타냈다.

3.2 Mean-Level Change(RQ 3)

각 변인에 대해 시점(3)을 반복측정 요인, 집단(2: 실험/통제)을 피험자 간 요인으로 설정한 혼합설계 반복측정 분산분석을 실시하였다. 결과를 Table 5에 제시하였다.

Table 5. Results of Mixed-Design Repeated-Measures ANOVA (N = 157)

Variable	Time				Group×Time			
	F	df1	df2	p	η _p ²	F	p	η _p ²
ADLS Total ^a	15.96	1.91	296.28	< .001	.093	0.20	.819	.001
E	15.93	2	310	< .001	.093	0.60	.549	.004
Ef ^a	9.22	1.92	298.25	< .001	.056	0.28	.753	.002
V ^a	6.78	1.80	278.53	.002	.042	0.30	.744	.002
M ^a	2.68	1.92	297.00	.073	.017	1.20	.304	.008

Note. ^a Mauchly's test indicated violation of sphericity (p < .05); Greenhouse-Geisser corrected df and p are reported(ADLS Total ε = .956; Ef ε = .962; V ε = .898; Metacognitive self-regulation ε = .958). E met the sphericity assumption and is reported without correction. η_p² = partial eta-squared. For interaction tests, the same (corrected) df as the corresponding time main effect were applied. Pairwise mean differences are based on estimated marginal means; minor discrepancies with Table 1 may occur due to rounding.

집단 주효과는 모든 변인에서 유의하지 않았으며(ps > .05), 집단 × 시점 상호작용도 모든 변인에서 유의하지 않았고 효과크기도 매우 작았다(η_p² = .001 - .008; Table 5). 이는 중간-사후 구간에 실시된 교육 프로그램이 시점에 따른 변화 양상에 차별적 영향을 미치지 않았음을 시사한다.

Bonferroni 사후 비교 결과를 Table 6에 제시하였다.

Table 6. Bonferroni Pairwise Comparisons of Estimated Marginal Means (N = 157)

Variable	Comparison	Diff	SE	95% CI	p
ADLS Total	T1 → T2	0.177	0.040	[0.080, 0.274]	< .001
	T1 → T3	0.222	0.046	[0.111, 0.333]	< .001
	T2 → T3	0.045	0.039	[-0.049, 0.138]	.741
Practice Experience (E)	T1 → T2	0.198	0.057	[0.061, 0.336]	.002
	T1 → T3	0.316	0.059	[0.174, 0.458]	< .001
	T2 → T3	0.118	0.054	[-0.014, 0.249]	.094
Self-efficacy(Ef)	T1 → T2	0.169	0.051	[0.047, 0.291]	.003
	T1 → T3	0.201	0.054	[0.070, 0.332]	.001
	T2 → T3	0.032	0.046	[-0.078, 0.143]	1.000
Value Perception (V)	T1 → T2	0.168	0.045	[0.059, 0.277]	.001
	T1 → T3	0.157	0.059	[0.014, 0.299]	.026
	T2 → T3	-0.012	0.048	[-0.128, 0.105]	1.000
Metacognitive self-regulation	T1 → T2	1.284	0.529	[0.004, 2.564]	.049
	T1 → T3	0.841	0.619	[-0.658, 2.340]	.530
	T2 → T3	-0.443	0.537	[-1.744, 0.858]	1.000

Note. Diff = estimated marginal mean difference (later - earlier; positive values indicate increase). SE = standard error of the difference. 95% CI = Bonferroni-adjusted 95% confidence interval. p values are Bonferroni-adjusted. Estimated marginal means are averaged over the levels of group.

ADLS 총점은 시점의 주효과가 유의하였다($F(1.91, 296.28) = 15.96, p < .001, \eta^2 = .093$). 사후 비교 결과 사전에서 중간 및 사전에서 사후로의 증가가 유의하였으나 중간에서 사후로의 변화는 유의하지 않았다(Table 6). 하위 요인도 동일한 패턴을 보여 실천경험, 자기효능감, 가치인식 모두 사전에서 중간으로 유의하게 증가한 후 사후까지 유지되었다.

메타인지적 자기조절은 시점의 주효과가 유의하지 않았다($F(1.92, 297.00) = 2.68, p = .073, \eta^2 = .017$). Bonferroni 사후 비교에서 사전-중간 비교가 명목상 유의 수준에 도달하였으나($p = .049$), 주효과가 유의하지 않았으므로 해당 결과는 보조적으로 해석하였다.

이상의 결과를 종합하면 ADLS(총점 및 하위요인)는 사전에서 중간으로 유의하게 증가한 후 사후까지 유지되는 양상을 보인 반면 메타인지적 자기조절은 3시점에 걸쳐 유의한 평균 변화를 보이지 않았다.

각 변인의 시점별 추정 평균 변화 양상을 Fig. 1에 제시하였다. 각 점은 집단 수준을 평균한 추정 주변 평균이며 오차 막대는 95% 신뢰구간을 나타낸다. 메타인지적 자기조절은 합산 점수(15 - 75점)를 문항 수(15)로 나누어 1 - 5점 척도로 변환하여 다른 변인과 동일한 축에 표시하였다.

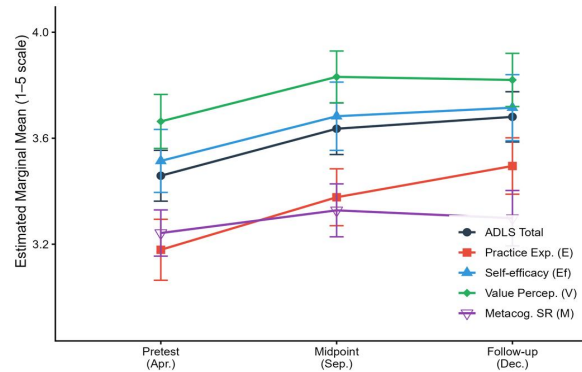


Fig. 1. Mean-Level Changes in ADLS and Metacognitive Self-regulation Across Three Time Points (N = 157)

4. Supplementary Analysis: Control Group Results

Table 7. Supplementary Analysis: Criterion-Related Validity in the Control Group (n = 71) Compared With Full Sample (N = 157)

Variable	Pretest		Mid		Posttest	
	Control	Full	Control	Full	Control	Full
ADLS Total	.669***	.641***	.633***	.631***	.630***	.599***
E	.626***	.562***	.454***	.555***	.524***	.551***
Ef	.650***	.587***	.642***	.574***	.626***	.587***
V	.464***	.463***	.449***	.481***	.437***	.333***

Note. Control = control group only (n = 71); Full = full sample (N = 157). ***p < .001.

본 연구의 자료가 실험집단에 대한 교육 프로그램 처치가 포함된 맥락에서 수집되었으므로 처치가 없었던 통제 집단(n = 71)만을 대상으로 보조분석을 실시하여 주요 결과의 강건성을 점검하였다. 결과를 Table 7과 Table 8에 제시하였다.

Table 8. Supplementary Analysis: Rank-Order Stability in the Control Group (n = 71) Compared With Full Sample(N = 157)

Variable	T1-T2		T1-T3		T2-T3	
	Control	Full	Control	Full	Control	Full
ADLS Total	.627***	.664***	.489***	.550***	.609***	.686***
E	.437***	.490***	.544***	.456***	.473***	.494***
Ef	.631***	.678***	.504***	.612***	.661***	.748***
V	.563***	.604***	.335***	.340***	.501***	.538***
M	.761***	.730***	.655***	.647***	.761***	.760***

Note. Control = control group only (n = 71); Full = full sample (N = 157). ***p < .001.

통제집단의 시점별 ADLS 총점-메타인지적 자기조절 상관($r = .630 - .669$)은 전체 표본($r = .599 - .641$)과 유사한 수준이었으며, 시점 간 상관(순위 안정성) 역시 전체 표본의 패턴과 일관되게 나타났다. 이는 본 연구에서 보고된 준거관련 타당도 및 순위 안정성 결과가 처치 효과에 의해 왜곡되지 않았음을 뒷받침한다.

V. Discussion

본 연구는 초등학생의 디지털 소양 정의적 영역 척도 (ADLS)의 준거관련 타당도 및 시간적 안정성 근거를 메타인지적 자기조절을 외부 준거로 활용한 3시점 종단 자료를 통해 검토하였다. 이하에서는 연구 문제별로 주요 결과를 해석하고 이론적·실천적 시사점과 제한점을 논의한다.

1. Consistency of Criterion-Related Validity(RQ 1)

ADLS 총점과 메타인지적 자기조절은 세 시점 모두에서 유의한 정적 상관을 보였으며($r = .599 - .641$), 탐색적으로 실시한 Steiger(1980) 검정에서 시점 간 상관 크기의 차이는 유의하지 않았다. 이는 ADLS가 측정하는 디지털 소양의 정의적 영역이 메타인지적 자기조절과 이론적으로 예측된 방향의 관련성을 가지며 그 관련성이 특정 시점에 한정되지 않고 반복 관측에서도 재현됨을 의미하여 ADLS의 준거관련 타당도 근거를 보강한다.

하위요인별로 보면 실천경험($r = .551-.562$)과 자기효능감($r = .574-.587$)은 시점에 걸쳐 안정적인 수준의 상관을 보였다. 이는 디지털 도구 활용에 대한 경험적 기반과 수행 자신감이 학습자의 자기조절적 점검·조절 능력과 일관되게 관련됨을 시사한다. 반면 가치인식은 사전($r = .463$)과 중간($r = .481$)에서는 중간 수준의 상관을 보였으나 사후 시점($r = .333$)에서는 상대적으로 약화되었으며 탐색적 Steiger 검정에서 중간 대 사후 비교가 유의 수준에 도달하였으나($z = 2.011, p = .044$), Benjamini-Hochberg 보정 적용 시 유의하지 않았다(Table 3 참조). 이러한 경향은 가치인식이 실천경험이나 자기효능감에 비해 맥락 의존적 속성을 가질 수 있음을 시사한다. 디지털 소양 활동의 유용성과 중요성에 대한 판단은 학습 과제의 내용이나 교실 맥락에 따라 변동할 여지가 있으며 사후 시점에서 두 변인간 관련성이 낮아진 패턴은 가치인식의 상황 민감성과 연결지어 해석해 볼 수 있다. 특히 사후 시점(12월)은 학년말로서 교과 활동이 축소되고 학습 환경이 전환되는 시기임을 고려할 때 디지털 소양 활동에 대한 가

치 판단이 맥락적 요인에 의해 일시적으로 변동하였을 가능성을 배제할 수 없다. 다만 이는 보정 후 유의하지 않은 탐색적 관찰이므로 후속 연구에서 가치인식과 메타인지적 자기조절 간 시점별 상관 변화에 대한 엄격한 가설 검증이 필요하다.

2. Temporal Stability: Rank-Order Stability and Mean-Level Change(RQ 2, 3)

순위 안정성의 측면에서 메타인지적 자기조절은 인접 시점 간 $r = .730 - .760$, 사전-사후 $r = .647$ 로 가장 높은 안정성을 보였다. 이는 메타인지적 자기조절이 비교적 안정적인 개인차 특성으로 기능함을 시사하며 외부준거로서의 적합성을 뒷받침한다. ADLS 총점($r = .550 - .686$)과 자기효능감($r = .612-.748$)은 적절한 수준의 안정성을 나타내어 반복 측정 상황에서 개인 간 상대적 순위가 일정 수준 유지됨을 확인하였다. 실천경험($r = .456-.494$)은 중간 수준이었으며 가치인식($r = .340-.604$)은 인접 시점에서는 중간 수준이었으나 사전-사후 간($r = .340$)에는 상대적으로 낮았다. 전반적으로 인접 시점 간 상관은 비인접 시점 간 상관보다 높게 나타났는데 이는 측정 간격이 길어질수록 순위 안정성이 감소하는 일반적 경향과 일치한다.

가치인식의 상대적으로 낮은 순위 안정성은 준거관련 타당도에서 관찰된 사후 시점의 상관 약화와 맥을 같이 한다. 이는 가치인식이 다른 하위요인에 비해 상황적 요인에 의해 더 민감하게 변동하는 특성을 가질 수 있음을 시사한다. 디지털 소양 활동의 가치에 대한 판단은 과제 경험과 교실 환경에 따라 재구성될 수 있으며 초등학생 시기에는 가치 판단의 변동성이 상대적으로 클 수 있다. 이 점은 ADLS를 활용한 종단 연구에서 가치인식 점수의 해석에 주의가 필요함을 시사한다.

평균 수준 변화와 관련하여 ADLS(총점 및 하위요인)는 사전에서 중간으로 유의하게 증가한 후 사후까지 유지되는 양상을 보였다. 사전-중간 구간(4월-9월)은 실험집단의 교육 프로그램이 실시되기 이전 시기로 이 구간에서 처치 없이도 ADLS 평균의 유의한 증가가 관찰되었다. 따라서 이러한 평균 증가 패턴은 처치 효과가 아닌 정규 교육과정의 학습 경험이나 자연적 발달에 의한 변화로 해석하는 것이 타당하다. 한편 집단 × 시점 상호작용은 모든 변인에서 유의하지 않아 시점에 따른 변화 양상이 실험집단과 통제 집단 간에 차별적이지 않은 것으로 나타났다. 메타인지적 자기조절은 3시점에 걸쳐 유의한 평균 변화를 보이지 않았으며 이는 해당 구인이 약 8개월의 관찰 기간 동안 집단 수준에서 비교적 안정적으로 유지되었음을 의미한다.

ADLS의 평균 증가와 메타인지적 자기조절의 평균 안정성이 동시에 관찰된 것은 두 구인이 관련되면서도 평균 수준 변화 양상에서는 차이가 있음을 시사한다.

3. Theoretical and Practical Implications

본 연구의 결과는 다음과 같은 시사점을 갖는다. 첫째, ADLS가 메타인지적 자기조절이라는 외부 준거와 시점에 걸쳐 일관된 관련성을 보임으로써 구인타당도에 더하여 준거관련 타당도에 대한 초기 근거가 추가적으로 확보되었다. 이는 ADLS 점수의 해석과 활용을 뒷받침하는 다원적 근거의 축적의 일환으로서 의의가 있다. 다만 본 연구의 분석이 관찰 점수 수준에 기반하므로 잠재 변수 수준에서의 타당도 확인은 후속 연구에서 보완될 필요가 있다. 둘째, ADLS의 시간적 안정성에 대한 초기 근거가 관찰됨으로써 종단 연구 및 교육 프로그램 운영 맥락에서 ADLS를 반복 측정 도구로 활용할 수 있는 가능성이 시사되었다. 다만 종단 측정불변성이 검증되지 않았으므로 관찰된 평균 변화가 구인의 실질적 변화를 반영하는지에 대해서는 추가 검증이 필요하다. 특히 ADLS 총점과 자기효능감은 적절한 순위 안정성과 함께 시점에 따른 평균 수준 변화가 관찰되어 변화 측정에 활용할 수 있는 잠재력을 시사한다. 실천적 측면에서 디지털 소양 교육 프로그램의 효과를 장기적으로 평가할 때 ADLS의 자기효능감 점수를 반복 측정함으로써 프로그램이 학습자의 디지털 수행 자신감에 미치는 영향을 모니터링하고 이를 토대로 교육 내용을 조정하는 데 활용할 수 있다. 안정적인 자기효능감은 학습자가 새로운 디지털 과제에 도전하는 데 긍정적으로 기여하므로 ADLS를 통해 자기효능감의 변화 추이를 추적하는 것은 디지털 역량 형성 과정을 이해하는 데 유용한 정보를 제공할 것이다. 셋째, 가치인식 요인이 준거관련 타당도와 순위 안정성 모두에서 상대적으로 낮은 수치를 보인 것은 해당 요인의 이론적 성격에 대한 추가 검토의 필요성을 제기한다. 가치인식이 상황 의존적 특성을 지닌다면 이를 안정적 특질로 보기보다는 맥락에 따라 변동하는 정의적 상태로 이해하는 것이 더 적절할 것이다. 이러한 변동성은 교육적 관점에서 중요한 함의를 갖는다. 디지털 소양 활동의 가치에 대한 인식이 일시적으로 저하될 경우 학습자가 디지털 도구를 활용한 심층적 학습 전략을 선택하는 빈도가 감소하고 이는 장기적인 디지털 역량 형성에 부정적 영향을 미칠 수 있다. 따라서 교육 현장에서는 특정 시점의 맥락적 요인에 의해 가치인식이 저하되지 않도록 디지털 활동의 내재적 유용성을 지속적으로 경험할 수 있는 교수 설계가 필요하다. 예를 들어 학습자가 디지

털 도구 활용의 성과를 스스로 성찰하는 활동이나 실생활과 연계된 디지털 과제를 꾸준히 제공하는 방안을 고려할 수 있다. 아울러 후속 연구에서는 가치인식 변동에 영향을 미치는 구체적 맥락 요인(과제 유형, 교사 피드백, 또래 상호 작용 등)을 체계적으로 탐색하고 관찰된 변동이 측정 도구 자체의 불안정성에 기인한 것인지 혹은 구인 자체의 역동적 변화를 반영하는 것인지를 종단 측정불변성 검증 등을 통해 구분할 필요가 있다.

4. Limitations and Suggestions for Future Research

본 연구는 다음과 같은 제한점을 갖는다.

첫째, 표본 및 연구 설계와 관련된 제약이 있다. 연구 참여자가 특정 지역 소재의 한 개 초등학교 5학년으로 한정되어 결과의 일반화에 한계가 있으며 다양한 지역·학년·학교급을 포함한 후속 연구가 필요하다. 본 자료는 실험집단에 대한 교육 프로그램 처치가 포함된 맥락에서 수집되었다. 집단 × 시점 상호작용이 유의하지 않았고($p_2 = .001 - .008$) 통제집단 보조분석에서도 전체 표본과 유사한 패턴이 확인되었으나, 보고된 시간적 안정성은 자연적 발달과 교육 경험이 포함된 실제 맥락에서의 안정성으로 해석되어야 한다. 또한 참여자가 9개 학급에 소속되어 있으나 학급단위의 군집 구조를 명시적으로 모델링하지 않았으므로 학급 내 비독립성으로 인해 표준오차가 과소추정되었을 가능성이 있으며 이는 보고된 p값이 실제보다 낮게 산출되어 유의성 판단에 영향을 미쳤을 수 있다. 본 연구에서는 학급수(9개)가 다중모형의 안정적 추정에 불충분하여 적용하지 않았다. 후속연구에서는 충분한 수의 학급을 확보하여 다중모형 또는 군집-강건 표준오차를 적용할 필요가 있다. 아울러 완전사례 분석(listwise deletion)을 적용하여 3시점 모두 응답한 157명만을 분석에 포함하였다.

탈락 분석에서 완전사례와 탈락자 간 사전 점수 차이가 유의하지 않아 무선 결측(MCAR)에 가까운 것으로 판단하였으나, 이는 관찰 변인에 기반한 제한적 점검이며 Little's MCAR 검정 등 체계적 검증이 실시되지 않았다. 결측 메커니즘이 MAR 또는 MNAR일 경우 완전사례 분석에 의한 결과가 편향될 수 있으므로 후속 연구에서는 완전정보최대우도법(FIML)이나 다중대체(multiple imputation) 등을 적용하여 결과의 강건성을 확인할 필요가 있다.

둘째, 외부 준거 및 측정 방법의 제약이 있다. 본 연구에서는 서론에서 논의한 바와 같이 여러 후보 준거변인 중 이론적 관련성과 구인 변별성, 측정 도구의 가용성을 종합적으로 고려하여 메타인지적 자기조절을 외부 준거로 선정하였

다. 그러나 단일 준거에 기반한 타당도 근거는 다원적 근거 축적의 관점에서 제한적이며 특히 ADLS와 메타인지적 자기조절 모두 자기보고식 척도로 측정되어 공통방법편의(common method bias)가 상관을 과대추정하였을 가능성을 배제할 수 없다. 관찰된 상관의 시점 간 일관성은 공통방법편의만으로 설명되기 어려우나 후속연구에서는 디지털 수행 평가, 교사 평정, 학업 성취 등 응답 방법이 상이한 외부 준거를 포함할 필요가 있다. 또한 두 변인의 척도 단위가 상이하므로(ADLS: 문항 평균, 메타인지적 자기조절: 합산 점수) Fig.1에서는 비교 편의를 위해 변환을 적용하였으나 원점수 수준에서의 평균 변화량 직접 비교에는 제약이 있다.

셋째, 분석 방법론의 제약이 있다. 본 연구의 시간적 안정성 분석은 관찰 점수 수준의 Pearson 상관과 반복측정 분산분석에 기반하여 측정 오차를 분리하지 못하였다. 또한 평균 수준 변화 해석에 전제되는 종단 측정불변성(longitudinal measurement invariance)이 사전에 검증되지 않았으며 본 연구의 표본 규모(N = 157)와 문항 수(20문항)를 고려할 때 3시점 종단 확인적 요인분석의 안정적 추정이 제한적이었다. 따라서 본 연구에서 보고된 평균 변화는 관찰 점수 수준의 변화로 해석되어야 하며 후속 연구에서는 급내상관계수(ICC), 잠재성장모형, 교차지연 패널모형 등을 적용하여 잠재 수준의 안정성과 변화 구조를 보다 정밀하게 검토하고 충분한 표본을 확보하여 종단 불변성을 검증한 후 잠재 평균 비교를 실시할 필요가 있다.

넷째, 가치인식 요인에 대한 탐색적 결과의 해석에 유의할 필요가 있다. Steiger(1980) 검정에 의한 시점 간 상관 비교는 탐색적으로 실시되었으며 다중비교 보정이 적용되지 않았으므로 가치인식의 중간-사후 감소($p = .044$)는 대규모 표본에서의 반복 검증이 요구된다. 가치인식의 상대적으로 낮은 안정성의 원인을 본 연구의 자료만으로 규명하기 어려우며 후속 연구에서 가치인식의 변동에 영향을 미치는 맥락 요인(교실 환경, 과제 유형, 교수 방법 등)을 체계적으로 탐색할 필요가 있다.

VI. Conclusions

본 연구는 초등학교 학생의 디지털 소양 정의적 영역 척도(ADLS)의 준거관련 타당도 및 시간적 안정성 근거를 3시점 종단 자료를 통해 검토하였다. ADLS 총점은 메타인지적 자기조절과 모든 시점에서 유의한 정적 상관을 보였으며 Steiger 검정 및 Benjamini-Hochberg 보정 적용 시에도 시점 간 상관 크기의 유의한 차이가 관찰되지 않아

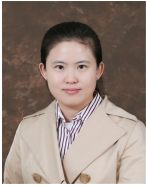
준거관련 타당도에 대한 초기 근거가 확인되었다. 시간적 안정성(순위 안정성)은 ADLS 총점과 자기효능감에서 비교적 적절한 수준으로 나타났고 실천경험은 중간 수준이었다. 평균 수준에서는 사전에서 중간 시점에서의 증가 후 유지되는 양상이 관찰되었다. 다만 가치인식 요인은 준거 관련 타당도와 순위 안정성 모두에서 상대적으로 낮은 수치를 보여 상황 의존적 특성을 포함한 추가 검토가 요구된다. 이상의 결과는 ADLS가 구인 타당도에 더해 준거관련 타당도 및 시간적 안정성에 대한 초기 근거를 추가적으로 확보하였음을 시사한다. 다만 본 연구의 결과는 관찰 점수 수준의 분석에 기반하며 종단 측정불변성이 사전 검증되지 않았으므로 이러한 해석은 후속 연구에서 잠재 변수 수준의 분석을 통해 보완될 필요가 있다. 본 연구의 결과에 기반할 때 ADLS는 학교 현장에서 초등학교 학생의 디지털 소양 정의적 영역 발달 수준을 주기적으로 진단하고 교육 프로그램의 효과를 시점에 걸쳐 객관적으로 평가하는 반복 측정 도구로 활용될 가능성이 있으나 이를 위해서는 종단 측정불변성 검증을 포함한 추가적인 측정학적 근거의 축적이 선행될 필요가 있다. 나아가 ADLS를 통해 측정된 종단 데이터는 교육과정 개선을 위한 실증적 근거를 제공할 수 있으며 후속 연구에서 다양한 학년·학교급·지역으로 적용 범위를 확장하고 종단 측정불변성 및 잠재 수준의 분석을 보완함으로써 디지털 소양 교육의 체계적 발전에 기여할 것으로 기대한다.

REFERENCES

- [1] Ministry of Education. Elementary and Secondary School Curriculum: General Guidelines, MOE Notification No. 2022-33, South Korea, 2022.
- [2] R. M. Bohlin. "The Affective Domain: A Model of Learner-Instruction Interactions," ERIC, 1998.
- [3] E. Kim, Y. Lee. "Development of an Affective-Domain Measurement Instrument for Elementary Students' Digital Literacy and Validation Based on Exploratory Factor Analysis," The Journal of Korean Association of Computer Education, Vol. 28, No. 12, pp. 129-141, 2025. DOI: 10.32431/kace.2025.28.12.012
- [4] M. T. Kane. "Validating the interpretations and uses of test scores," Journal of educational measurement, Vol. 50, No. 1, pp. 1-73, 2013. DOI: 10.1111/jedm.12000
- [5] D. R. Saunders. "ON THE STABILITY OF PSYCHOLOGICAL MEASURES," ETS Research Bulletin Series, Vol. 1954, No. 1, pp. i-13, 1954.
- [6] E. Öztürk Ölmezer. "Contemporary motivation theories in educational psychology and language learning: An overview," The International

- Journal of social sciences, Vol. 3, No. 1, pp. 33-46, 2012.
- [7] A. Ferrari. "DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe," Publications Office of the European Union, pp. 1-45, 2013. DOI: 10.2788/52966
- [8] A. Ferrari. "Digital Competence in Practice: An Analysis of Frameworks," Publications Office of the European Union, 2012. DOI: 10.2791/82116
- [9] J. Park. *The Relationship between the Creative Problem Solving Ability in Science and Students' Metacognition*, Master's thesis, Department of Elementary Education, Ewha Womans University, pp. 1-106, 2005.
- [10] T. Annevirta, M. Vauras. "Developmental changes of metacognitive skill in elementary school children," *The journal of experimental education*, Vol. 74, No. 3, pp. 195-226, 2006. DOI: 10.3200/JEXE.74.3.195-226
- [11] A. M. Kosnin. "Self-regulated learning and academic achievement in Malaysian undergraduates." *International Education Journal*, Vol. 8, No. 1, pp. 221-228, 2007.
- [12] D. Dangin, A. H. Setyawan and E. Hartati. "Metacognitive awareness and self-regulated learning toward students' academic achievement: prediction study," *International Journal of Social Science and Human Research*, Vol. 6, No. 11, 2023. DOI: 10.47191/ijsshr/v6-i11-24
- [13] D. Bawden. "Information and digital literacies: a review of concepts," *Journal of Documentation*, Vol. 57, No. 2, pp. 218-259, 2001. DOI: 10.1108/EUM0000000007083
- [14] N. Law, D. J. Woo, J. De la Torre and G. Wong, "A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2," UNESCO Institute for Statistics, pp. 1-146, 2018.
- [15] R. Vuorikari, S. Kluzer and Y. Punie, "DigComp 2.2: The Digital Competence Framework for Citizens - With New Examples of Knowledge, Skills and Attitudes," Publications Office of the European Union, 2022. DOI: 10.2760/115376
- [16] M. C. Martínez-Bravo, C. Sádaba Chalezquer and J. Serrano-Puche. "Dimensions of digital literacy in the 21st century competency frameworks," *Sustainability*, Vol. 14, No. 3, pp. 1867, 2022.
- [17] R. Schroeder, E. S. Cahoy. "Valuing information literacy: Affective learning and the ACRL standards," *portal: Libraries and the Academy*, Vol. 10, No. 2, pp. 127-146, 2010.
- [18] T. Bouffard-Bouchard. "Influence of self-efficacy on performance in a cognitive task," *J Soc Psychol*, Vol. 130, No. 3, pp. 353-363, 1990. DOI: 10.1080/00224545.1990.9924591
- [19] N. J. Collazo, J. Elen and G. Clarebout. "To use or not to use tools in interactive learning environments: a question of self-efficacy?" *The Literacy Information and Computer Education Journal*, Vol. 1, No. 1, pp. 810-817, 2012.
- [20] A. N. H. Afandi, S. R. Kusumaningrum, R. S. I. Dewi and R. Pristiani. "Digital Literacy Questionnaire Instrument: Based on the Integration of Elementary School Students' Characteristics," *International Journal of Elementary Education*, Vol. 8, No. 2, pp. 344-353, 2024. DOI: 10.23887/ijee.v8i2.76773
- [21] B. Çetin. "Metacognition and Self-Regulated Learning in Predicting University Students' Academic Achievement in Turkey." *Journal of Education and Training Studies*, Vol. 5, No. 4, pp. 132-138, 2017.
- [22] K. C. Mathabathe, M. Potgieter. "Manifestations of metacognitive activity during the collaborative planning of chemistry practical investigations," *International Journal of Science Education*, Vol. 39, No. 11, pp. 1465-1484, 2017. DOI: 10.1080/09500693.2017.1336808
- [23] D. Lucangeli, M. C. Fastame, M. Pedron, A. Porru, V. Duca, P. K. Hitchcott and M. P. Penna. "Metacognition and errors: the impact of self-regulatory trainings in children with specific learning disabilities," *ZDM*, Vol. 51, No. 4, pp. 577-585, 2019.
- [24] J. Shaha, R. Rajendran, "Analyzing the Impact of Metacognition Prompts on Learning in CBLE", *Proceedings of the 16th International Conference on Educational Data Mining*, Bengaluru, India, pp. 561-563, Jul. 2023.
- [25] R. Azevedo. "Computer environments as metacognitive tools for enhancing learning," *Educational Psychologist*, Vol. 40, No. 4, pp. 193-197, 2005. DOI: 10.1207/s15326985sep4004_1
- [26] J. A. Greene, R. Azevedo. "The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments," *Educational psychologist*, Vol. 45, No. 4, pp. 203-209, 2010. DOI: 10.1080/00461520.2010.515935
- [27] T. Raykov, G. A. Marcoulides. "Evaluation of true criterion validity for unidimensional multicomponent measuring instruments in longitudinal studies," *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 24, No. 4, pp. 599-608, 2017. DOI: 10.1080/10705511.2016.1172486
- [28] T. Raykov. "Examining temporal stability of scale validity in longitudinal studies," *Multivariate Behavioral Research*, Vol. 41, No. 3, pp. 401-426, 2006. DOI: 10.1207/s15327906mbr4103_6
- [29] M. S. B. Yusoff. "ABC of content validation and content validity index calculation," *Education in medicine journal*, Vol. 11, No. 2, pp. 49-54, 2019.
- [30] H. Beckerman, M. E. Roebroek, G. J. Lankhorst, J. G. Becher, P. D. Bezemer and A. Verbeek. "Smallest real difference, a link between reproducibility and responsiveness," *Quality of Life Research*, Vol. 10, No. 7, pp. 571-578, 2001.
- [31] M. von Krause, S. T. Radev, A. Voss, M. Quintus, B. Eglhoff and C. Wrzus. "Stability and change in diffusion model parameters over two years," *Journal of Intelligence*, Vol. 9, No. 2, pp. 26, 2021. DOI: 10.3390/jintelligence9020026
- [32] J. A. DeSimone. "New techniques for evaluating temporal consistency," *Organ Res Methods*, Vol. 18, No. 1, pp. 133-152, 2015. DOI: 10.1177/1094428114553061

Authors



Eun-Ji Kim received the B.Ed. degree in Elementary Education from Jinju National University of Education, Korea, in 2011, and the M.Ed. degree in Elementary Computer Education from Korea National University of

Education, Korea, in 2019. She has completed the coursework for the Ph.D. program in Computer Education at Korea National University of Education. Ms. Kim has been an elementary school teacher since 2011. She is interested in computational thinking, digital divide, and digital literacy.



Young-Jun Lee received the B.S. degree in Computer Science from Korea University, Korea, in 1988, and the Ph.D. degree in Computer Science from the University of Minnesota, USA, in 1994.

Dr. Lee has been a Professor in the Department of Computer Education at Korea National University of Education, Cheongju, Korea, since 2003. He is interested in intelligent systems, learning sciences, information education, and artificial intelligence education.