

초등학교 6학년 디지털 소양 진단 도구의 성취수준 설정 연구¹⁾

김희경 (한국교육과정평가원 선임연구위원)*

박민호 (한국교육과정평가원 부연구위원)**

요약

2022 개정 교육과정에서 디지털 소양이 기초 소양으로 포함됨에 따라 학교교육 전반에서 디지털 소양 교육이 강조되고 있다. 하지만 학교 현장의 교사들이 학생의 수준에 맞는 디지털 소양 교육을 실시하기 위한 정보는 부족한 상황이라고 볼 수 있다. 본 연구에서는 초등학교 6학년을 대상으로 한 KICE 디지털 소양 진단 평가(KICE DiLAP)의 문항 개발 이후에 검사의 측정학적 특성을 점검하고, 향후 교사들이 수월하게 학생의 디지털 소양 수준을 진단할 수 있도록 성취수준 설정을 수행하였다. 이를 위해 진단 도구의 예비시행 및 채점을 진행한 후, 857명의 자료를 확보하여 문항 특성 분석, 총점 및 하위영역 점수 분석을 실시하였고, 최종적으로 디지털 소양에 대한 4단계 성취수준을 설정하였다. 성취수준 설정 결과 우리나라 초등학교 6학년 학생들의 4수준(우수) 비율은 약 20%인 반면, 1수준(노력 요망) 비율도 약 11%인 것으로 나타났다. 또한 성취수준별 대표문항 분석을 통해 각 성취수준별 학생들의 디지털 소양 수준을 구체적으로 파악하여 제시하였다. 이러한 연구 결과는 향후 교사들이 필요에 따라 KICE 디지털 소양 진단 평가(KICE DiLAP)를 사용하여 학생들의 디지털 소양 수준을 진단하고, 결과를 수월하게 해석하여 학습자 수준에 맞는 효율적인 교수·학습 수립에 활용하는 데 유용한 자료가 될 것이다.

주제어 : 2022 개정 교육과정, 디지털 소양, 디지털 리터러시, 디지털 소양 진단, 성취수준 설정

1) 이 논문은 ‘2022 개정 교육과정에 따른 디지털 소양 진단 도구 개발’(전성균 외, 2023) 보고서의 일부 내용을 수정·보완한 것임.

* 제1저자: heekyoung@kice.re.kr

** 교신저자: pmينو85@kice.re.kr

I. 서 론

2022 개정 교육과정에서 학생들이 갖춰야 할 기초 소양의 하나로 디지털 소양을 강조함에 따라 앞으로 학교교육에서 디지털 소양 교육이 취지에 맞게 실행되어야 할 필요성이 부각되고 있다. 그런데 아직 학교 현장에서 디지털 소양 교육을 실시함에 있어 학생의 수준에 알맞은 교육의 내용 및 방법 등을 모색할 수 있도록 구체적인 안내가 마련되지 못하고 있기에 이를 지원하기 위한 연구가 요구되는 상황이다. 2022 개정 교육과정을 기반으로 초등학교 교과교육에서 디지털 소양 교육 적용 시에 예상되는 어려움에 대한 교사 인식 조사에서는 모든 교과에서 초등학교 교사 과반이 디지털 소양 교육 내용의 범위와 목표, 학습자들의 능력치 가늠이 어렵다는 의견을 가진 것으로 나타났다(배주경 외, 2023). 이에 따라 본 연구에서는 단위학교 교사가 학생의 디지털 소양 교육을 위해 필요에 따라 유연하게 사용할 수 있는 디지털 소양 진단 도구의 개발 및 지원이 중요함을 인식하였다. 디지털 소양 진단 도구는 학생들의 디지털 소양 수준을 정확히 파악하고, 이를 바탕으로 교사들이 학생의 수준에 맞는 적절한 디지털 소양 교육 내용 설정 및 교수·학습 자료를 제공할 수 있도록 지원하는 기능을 수행할 수 있을 것이다.

이와 같은 필요성에 의해 한국교육과정평가원에서는 2022 개정 교육과정에서 제시한 디지털 소양 교육목표에 근거하여 학생의 디지털 소양 수준을 진단하기 위한 평가(KICE Digital Literacy Assessment Program, 이하 KICE DiLAP)를 개발하였다(전성균 외, 2023). 2022 개정 교육과정에서는 모든 교과교육에서의 디지털 소양 함양을 강조하였고 디지털 기술에 관련된 지식을 습득하는 것 외에도 실생활 맥락에서 관련된 지식, 기능, 태도를 종합적으로 적용하는 것을 목표로 한다. 2022 개정 교육과정에서 제시한 디지털 소양의 개념은 ‘디지털 지식과 기술에 대한 이해와 윤리 의식을 바탕으로, 정보를 수집·분석하고 비판적으로 이해·평가하여 새로운 정보와 지식을 생산·활용하는 능력’이다(교육부, 2021, p. 13). 전성균 외(2023)에서는 이 정의를 토대로 교육과정 상에 제시한 디지털 소양 교육의 4가지 내용 영역(디지털 기기와 소프트웨어 활용, 디지털 정보의 활용과 생성, 디지털 의사소통과 문제해결, 디지털 윤리와 정보 보호)으로 평가틀을 구성하여 학교현장에서 효율성 있는 디지털 소양 교수·학습을 실시하는 데 실질적이고 타당한 도움을 줄 수 있는 진단 도구를 마련하였다. KICE 디지털 소양 진단 평가를 활용한다면 단위학교 교사들이 교육과정 상의 교육목표에 부합하는 학생들의 디지털 소양 성취수준(예: 우수 / 보통 / 기초 / 노력 요망의 4단계 성취수준)을 수월하게 진단하는 것이 가능할 것으로 기대된다. 이를 위해 본 연구에서는 KICE 디지털 소양 진단 평가에 대한 예비시행을 통해 학생의 응답 자료를 수집하고, 교사들이 학생의 진단 결과를 해석할 수 있도록 4단계 성취수준을 설정하였다.

디지털 소양은 미래 사회를 살아가야 할 학생들에게 필수 역량일 뿐 아니라 2025년부터 디지털 소양이 필요한 AI 디지털교과서가 학교 현장에 도입된다는 점에서 학생의 디지털 소양 수준이 교과 교수·학습의 수월성에도 영향을 끼칠 수 있을 것으로 예상된다. 이러한 상황에서 디지털 소양 진단 평가의 결과 해석을 위해 성취수준을 설정하는 연구는 공교육에서 학생들의 디지털 소양 함양이 체계적으

로 이루어질 수 있도록 모니터링하고 낮은 디지털 소양 수준인 학생의 비율을 감소시키는 데 도움이 될 것이다. 나아가 성취수준별 비율에서 성별, 지역 등에 따른 격차 유무 등을 확인하여 학생들의 디지털 소양에서 교육격차를 예방하거나 감소시킬 방안을 강구하기 위한 기초자료를 제공할 수도 있다.

따라서 본 연구에서는 초등학교 6학년용 KICE 디지털 소양 진단 평가의 성취수준을 설정한 후, 4단계 성취수준별 학생 비율 및 문항 분석 결과를 제시함으로써 각 성취수준에 해당하는 학생의 지식과 수행 수준에 대한 구체적인 진단과 해석을 강화할 수 있는 방안을 탐색하였다.

II. 이론적 배경

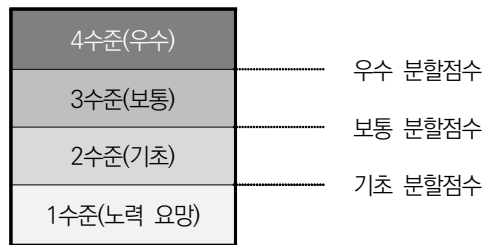
1. 성취수준 설정의 개념 및 필요성

KICE 디지털 소양 진단 평가는 우리나라 학생들의 디지털 소양 수준을 정확히 진단하고, 진단 결과를 활용하여 교사들이 학생의 수준에 맞는 디지털 소양 교육 내용 설정 및 교수·학습 자료를 제공할 수 있도록 지원하는 기능을 갖추어야 한다. 따라서 디지털 소양 진단 평가는 우리나라 학생들의 디지털 소양 수준이 교육목표에 도달했는지 여부를 교육과정의 교육목표 및 내용에 근거하여 판단할 수 있도록 준거참조평가(criterion-referenced assessment)로 사용하는 것이 적절하다. 준거참조평가에서는 학생이 ‘무엇을’ ‘얼마나’ 성취했는가에 초점을 맞추어 ‘무엇을’에 해당하는 교육목표와 ‘얼마나’에 해당하는 성취수준(achievement level, performance level)이 중요한 잣대가 된다(김희경, 조성민, 2013, p. 481).

IEA에서 주관하는 ICILS 2018의 경우에도 컴퓨터·정보 소양(Computer and Information Literacy; CIL)과 컴퓨팅 사고력(Computational Thinking; CT)에 대한 성취수준을 마련하여 학생들의 능력 수준을 판별하고 있다. 컴퓨터·정보 소양은 5단계의 성취수준(1수준 미만, 1수준, 2수준, 3수준, 4수준)으로 설정하고, 컴퓨팅 사고력은 3단계의 성취수준(하, 중, 상)으로 구분하고 있다(IEA, 2020, pp. 4-6). 또한 KERIS에서 주관하는 국가수준 초·중학생 디지털 리터러시 평가에서는 ‘우수’, ‘보통’, ‘기초’, ‘미흡’의 4단계의 성취수준으로 학생들의 디지털 소양을 판단하고 있다(이현숙 외, 2022, p. 31).

이상의 검사들과 같이 KICE 디지털 소양 진단 평가의 결과를 해석하고 활용하기 위해서는 우리나라 초등학교 6학년 학생들의 디지털 소양 수준을 타당하고 신뢰성 있게 판단할 수 있는 성취수준 설정(standard-setting)이 필요하다. KICE 디지털 소양 진단 평가의 결과를 활용할 때, 총점을 비교하는 방식이 아니라, 교육과정 상에 제시된 교육목표에 대한 학생의 도달 정도에 따라 학생의 성취수준을 판단할 필요가 있기 때문이다.

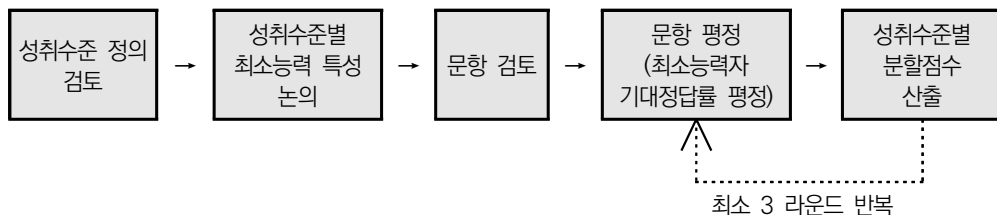
성취수준을 설정하는 것은 학생이 도달한 정도를 몇 개의 범주(성취수준)로 나타내고 이를 구분하는 분할점수(cut-off scores)를 결정하는 것을 의미한다. 학생이 도달한 정도는 ‘도달’과 ‘미도달’의 2개 범주로 구분하거나, 또는 2개 이상의 범주로 구분할 수 있다. [그림 1]과 같이 학생들의 성취수준을 4단계로 구분하는 경우에는 3개의 분할점수가 필요하다.



[그림 1] 4단계 성취수준 분할점수

대표적인 성취수준 설정 방법은 Angoff(1971)에서 제안된 성취수준 설정 방법으로 참여한 전문가 패널(교사, 내용 전문가 등)이 개별 문항을 검토·분석한 후, 해당 성취수준에 포함된 사람들 중 하위 경제선에 겨우 걸쳐 있는 최소능력자의 문항정답률을 추정하는 절차를 통해 해당 성취수준과 그 아래의 성취수준을 구분하는 분할점수를 결정하는 방법이다. 이후 Cizek & Bunch(2007)가 제안한 ‘변형된 Angoff 방법(modified Angoff method)’에서는 1회의 문항정답률 추정 절차로 성취수준을 설정하는 것이 아니라, 반복된 라운드를 통해 전문가들의 문항정답률 추정에 대한 질을 제고한다. 두 번째 라운드부터는 다른 전문가들의 결정이 어떠했는지에 대한 요약적인 정보를 제공받고 전문가 간의 문항정답률 편차가 큰 문항을 중점으로 토의를 진행한다. 이러한 절차를 따르는 변형된 Angoff 방법은 국내외의 대규모 학업성취도 평가, 자격증 시험과 같은 고부담 시험, 공인된 평가 등에서 안정적인 성취수준 설정 방법으로 활용되고 있다. 김창환(2011, p. 149)에서도 변형된 Angoff 방법은 전문가 집단이 문항 내용에 대한 검토 및 분석 절차를 거쳐 성취수준을 설정하는 방법 중에서 가장 자주 쓰이고, 안정적인 방법으로 알려져 있다고 하였다. 국가수준 학업성취도 평가에서도 2002년 변형된 Angoff 방법을 적용하여 성취수준을 설정하였고 이후 2010년, 2015년, 2019년 시점에 재차 변형된 Angoff 방법을 적용하여 성취수준을 재설정한 바 있다(김희경 외, 2019, pp. 68-72).

요약하자면 변형된 Angoff 방법은 기존의 Angoff 방법과 동일하게 각 성취수준의 최소능력자에게 기대하는 문항정답률에 근거하여 분할점수를 산출하는 방법으로, 전문가 패널들은 해당 성취수준 집단에서 분할점수를 가까스로 넘기는 정도의 최소능력 수준에 대한 논의를 거치게 된다(Cizek & Bunch, 2007, pp. 42-56). 다만, 변형된 Angoff 방법에서는 1회의 문항정답률 추정으로 성취수준을 설정하는 기존 Angoff 방법과 달리 [그림 2]와 같이 문항 평정과 성취수준 분할점수를 여러 차례 반복하는 성취수준 설정 절차를 거치게 된다.



• 출처: 김희경 외(2019, p. 76)에서 발췌함.

[그림 2] 변형된 Angoff 방법을 적용한 성취수준 설정 절차

2. 성취수준 설정 선행연구

일반적으로 학습자가 무엇을 얼마나 이해하고 있는지를 중시하는 준거참조평가에서는 평가 결과를 해석할 때, 총점이나 석차로 학생의 수행을 상대적으로 비교하는 방식이 아니라, 학생의 교육목표 도달 정도에 따라 학생의 수행 수준을 판단하므로, 성취수준 설정 절차가 필요하다. Livingston & Zieky(1982)에 의하면 성취수준 설정 절차는 ‘정답’을 맞추는 것이 아니라, 전문가들이 합의 과정을 거쳐 어느 정도의 수준이 기준에 도달했다고 판단하기에 적합한지를 ‘의사결정’하여 분할점수를 만들어내는 절차라고 규정하였다. 따라서 준거참조평가에서 평가 결과를 해석하고 활용하는 데 있어서 성취수준 설정은 반드시 필요한 절차이지만, 전문가들의 합의를 이끌어내는 과정이 어렵고 까다로워 구체적인 절차에 대한 사례 연구가 다수 필요함에도 불구하고, 구체적인 절차가 공개되어 있는 선행연구는 많지 않다.

김희경·조성민(2013)은 초등학교 3학년 ‘기초수학’의 기초학력 진단평가에서 기초학력 도달 여부를 결정하는 성취수준 설정 연구를 수행하였다. 이 연구에서는 성취수준 설정을 위해 총 29문항으로 구성된 초등학교 3학년 ‘기초수학’ 기초학력 진단평가를 시행하여 1,483명의 학생 응답 자료를 수집하고, 6명의 초등학교 교사가 워크숍을 통해 Ebel 방법을 적용하여 기초학력 도달 여부를 결정하는 분할점수를 설정하였다. Ebel 방법에서는 문항을 범주화하여 최소능력자의 문항 정답률을 추정하는데, 문항의 난이도와 내용영역에 따라 문항을 15개의 범주로 묶고, 유사한 문항군에 대해 정답률을 추정하는 과정을 총 3라운드 진행한 후에 최종적인 분할점수를 산출하였다.

김현경·정진수(2018)는 2015년 중학교 3학년 국가수준 학업성취도 평가에서 과학과의 4단계 성취수준(우수/보통/기초/기초미달)을 결정하는 성취수준 설정 연구를 수행하였다. 이 연구에서는 성취수준 설정을 위해 총 40문항으로 구성된 중학교 3학년 학업성취도 평가 검사지를 28명의 전문가 패널이 워크숍을 통해 변형된 Angoff 방법으로 성취수준 설정을 진행하였다. 이 연구의 특징은 성취수준 설정 절차에서 전문가 패널이 개별 문항을 검토하며 정답률을 추정하는 1~3라운드를 진행하기 전에 이전 학업성취도 평가의 문항들을 선별하여, 이에 대한 성취수준별 최소능력자의 정답률을 추정하는 연습을 실시함으로써 전문가들이 문항을 평정하는 데 있어서 익숙해지는 시간을 갖도록 했다는 점이다.

III. 연구 방법

1. KICE 디지털 소양 진단 평가(KICE DiLAP) 개요

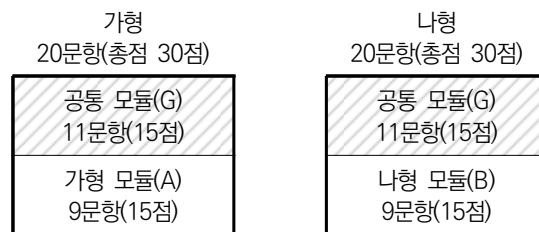
본 연구에서는 우리나라 교육과정을 토대로 디지털 소양 진단 평가를 개발하는 취지를 고려하여 2022 개정 교육과정에서 의도하는 디지털 소양 구성, 즉 4개 영역 및 10개 세부 요소(김자미 외, 2022)를 근거로 아래 <표 1>과 같은 디지털 소양 평가틀을 설정하였다.

〈표 1〉 디지털 소양 평가를

영역	세부 요소	세부 요소 설명
1. 디지털 기기와 소프트웨어의 활용	1-1 디지털 기기의 활용	디지털 기기를 조작하는 데 필요한 기본 원리와 기능을 이해 및 활용한다.
	1-2 소프트웨어의 활용	소프트웨어의 기본 원리와 기능을 이해하고 다양한 작업에서 소프트웨어를 활용한다.
	1-3 인공지능의 활용	다양한 문제해결 과정에 인공지능 기술이 탑재된 도구를 활용한다.
2. 디지털 정보의 활용과 생성	2-1 자료의 수집과 저장	사용 목적을 고려해 자료를 수집하고, 비판적 시각으로 정확성을 평가하여 효율적으로 저장 / 관리한다.
	2-2 정보의 분석과 표현	정보를 효과적으로 전달하기 위해 데이터를 분석, 종합, 시각화한다.
	2-3 디지털 콘텐츠 생성	디지털 미디어를 통해 제공될 수 있는 다양한 유형의 콘텐츠를 생성한다.
3. 디지털 의사소통과 문제해결	3-1 디지털 의사소통	디지털 환경에서 정보를 비판적으로 분석하고, 정보 공유, 의사결정 참여, 협업을 수행한다.
	3-2 디지털 문제해결	문제해결 방안을 구안하고, 디지털 도구를 활용하여 실행한다.
4. 디지털 윤리와 정보 보호	4-1 디지털 윤리	디지털 사회의 성숙한 시민으로서 타인을 배려하고, 예절과 윤리를 실천한다.
	4-2 디지털 정보 보호	자신과 타인의 정보를 보호하기 위한 방법을 실천한다.

* 출처: 김자미 외(2022, pp. 16-17)에서 일부 발췌.

디지털 소양 평가를 하에서 KICE 디지털 소양 진단 평가는 2개의 검사 유형(가형, 나형)으로 개발되었고, 각 검사 유형은 공통문항을 50% 이상 공유하도록 하였다. 문항 형식은 주제 단위로 문항이 세트로 구성되는 모듈형 문항 형식으로 개발하였고, 총 3종의 소검사 모듈(공통 모듈, 가형 모듈, 나형 모듈)을 개발하여 [그림 3]과 같이 가형과 나형에 배치하였다. 그리고 각 모듈 내에는 선다형 외에도 서답형 및 수행형 문항이 포함되어 있다.



[그림 3] KICE 디지털 소양 진단 평가(KICE DiLAP)의 검사도구 설계

KICE 디지털 소양 진단 평가는 실제 디지털 환경에서 학생들이 디지털 기기를 활용하고, 정보를 검색·저장하거나 디지털 콘텐츠를 생성하는 능력을 타당하게 평가할 수 있도록 오프라인의 지필식 평가가 아닌 온라인 평가²⁾ 문항으로 제작되었으며, 예비시행도 온라인 평가로 진행하였다. 예비시행은

2) 디지털 소양 진단 평가 예비시행은 국가수준 학업성취도 평가에서 운영하는 '맞춤형 학업성취도 자율평가(성취)'의 컴퓨터 기반 문항 출제 및 시행 시스템을 활용하였음.

1차, 2차로 구분하여 2023년 6~7월에 걸쳐 시행하였다. 예비시행 대상 학생들은 2가지 검사 유형(가형, 나형) 중 1가지를 배분받았으며, 학교 컴퓨터실에서 검사를 시행하였다.

〈표 2〉에 제시된 바와 같이 1차 예비시행은 소수의 학생(약 50명)을 대상으로 진단 도구의 문항 양도 및 최종 문항 확정을 위한 수정·보완을 목적으로 하였다. 2차 예비시행은 최종 문항 확정이 종료된 후, 학생의 문항 응답 자료를 수집하여 우리나라 초등학교 6학년 학생의 디지털 소양 성취수준 설정을 위한 기초자료로 활용하려는 목적으로 시행되었다. 따라서 2차 예비시행에서는 자료 분석에 충분한 학생 수를 얻고자 962명을 대상으로 검사를 시행하였다.

〈표 2〉 예비시행 대상

구분	학교명	학급 수	학생 수	지역	검사 유형
1차 (2023년 6월)	A초등학교	2학급	50명	세종	가형/나형
2차 (2023년 7월)	B초등학교	8학급	234명	충북	가형
	C초등학교	13학급	407명	대구	나형
	D초등학교	12학급	321명	인천	가형

디지털 소양 진단 평가의 예비시행이 온라인 평가로 진행되었으므로 학생들의 답안은 웹 서버에 바로 저장되는 시스템이다. 이러한 학생들의 답안은 온라인 채점 시스템을 통해 채점자에게 분배되었다. 채점자들이 연습채점을 통해 채점기준이 어느 정도 숙지된 상태에서 본채점을 실시하였으며, 본채점에서는 채점의 안정성 및 신뢰성을 최대한 확보하고자 교차채점 방식으로 진행하였다. 교차채점은 동일한 답안에 대해 채점자 2명이 독립적으로 채점을 한 후, 채점자 간 부여한 점수에 차이가 있을 경우에 재채점을 실시하는 방식이다. 채점 과정을 거친 후, 결시자 등을 제외하는 데이터클리닝 작업 후 857명 사례 수에 대한 자료를 확보하였다.

2. 성취수준 설정 절차

성취수준 설정에 참여하는 전문가 패널은 정보교과 석사 이상의 교사 또는 초등학교에 현재 재직 중이면서 초등학교 6학년 학생을 교육한 경험이 풍부하여 학생들의 디지털 소양 수준을 가늠하기에 능력이 충분하다고 판단되는 교원으로 구성하는 것을 원칙으로 하였다. 성취수준 설정을 위한 전문가 패널 인원은 최소 20명으로 구성하는 것이 적절하나 예산을 고려하여 10명으로 섭외하였다.

성취수준 설정 워크숍의 1일차에는 디지털 소양 진단 도구의 4단계 성취수준(우수/보통/기초/노력요망)의 개념에 대해 토의하는 것으로 시작하였다. 그리고 성취수준별 특성에 대한 이해를 돕기 위해 ICILS 2018 공개 문항을 활용하여 성취수준별 학생들의 정답률을 추정하는 연습을 실시하였다. 디지털 소양 진단 도구와 유사한 성격을 지닌 ICILS 문항을 통한 정답률 추정 연습을 통해 이후 실시할 디지털 소양 진단 도구의 문항 정답률 예측의 정확성을 높이고자 하였다. 동시에 ICILS 2018 공개 문항의 각 성취수준 집단별 정답률을 참조하여 유사한 검사에서의 성취수준 특성에 대한 내면화를 기대하

였다. 또한 디지털 소양 진단 도구를 패널들이 직접 온라인으로 학생과 동일한 환경 하에서 학생의 입장이 되어 문항을 풀어보면서 가형과 나형을 구성하는 모든 문항의 내용, 채점 기준, 난이도에 대해 면밀히 검토할 수 있도록 하였다.

2일차에는 1일차에 어느 정도 형성된 성취수준의 개념을 토대로 각 성취수준에서 경계선에 위치한 최소능력자의 수행 특성을 논의하였다. 논의 과정을 통해 패널들은 성취수준별 최소능력자에 대한 개념을 공유하였다. 성취수준별 최소능력자에 대한 개념을 공유한 이후에는 변형된 Angoff 방법을 활용하여 수준설정을 시작하여 1~2라운드를 진행하였다. 그리고 본 연구에서는 수준설정 과정에서 변형된 Angoff 방법을 좀 더 변형한 새로운 방법을 추가적으로 진행하였다. 변형된 Angoff 방법에서는 다분문항의 경우, 각 부분점수마다 성취수준별 최소능력자의 정답률을 추정해야 한다는 점에서 패널들에게 인지적 부담이 크고 평정 결과 입력에도 번거로움이 있다. 이에 본 연구에서는 이분문항에 대해서는 변형된 Angoff 방법과 동일하게 진행하지만, 다분문항에서는 성취수준별 최소능력자가 얻을 것이라 기대되는 부분점수를 소수점 첫째자리까지 추정하도록 하는 방법을 제안하였다. 그리고 이 방법을 ‘다분문항 기대점수 추정 방법’이라 명명하였다.

1라운드에서는 5명씩 2개 조를 형성하여 디지털 소양 진단 도구 가형과 나형에 포함된 문항 전체에 대해 우수/보통/기초의 성취수준별로 경계선에 위치한 최소능력자의 문항정답률을 추정하였다. <표 3>에는 1라운드에서 10명의 패널이 문항별로 4수준(우수) 최소능력자의 문항 정답률을 추정한 결과의 일부를 제시하였다. 0/1점으로 채점하는 이분점수 문항의 경우에는 문항을 맞힐 확률을 추정하고, 단계적인 부분점수를 부여하는 다분점수 문항은 각 부분점수를 받을 확률을 판정하도록 하였다. 예를 들어 2점 배점의 문항인 2번 문항의 경우에 우수학력의 최소능력자 중에서 2, 1, 0점을 받을 확률을 판정하고, 각 점수를 받을 확률의 총합은 100%가 되어야 한다. 또한 각 점수를 받을 확률은 0, 5, 10, ..., 90, 95, 100%와 같이 5% 단위로 부여하도록 하였다.

〈표 3〉 1라운드 4수준(우수) 최소능력자에 대한 문항정답률 추정 결과 일부

라운드		1	4수준(우수)												
		정답률 예측													
문항 번호	배점	부분 점수	패널 1	패널 2	패널 3	패널 4	패널 5	패널 6	패널 7	패널 8	패널 9	패널 10	중앙 값	평균	
2	2	2점	85	80	90	80	70	85	80	70	90	80	80	81	
		1점	15	20	10	20	20	15	20	30	10	20	20	18	
		0점	0	0	0	0	10	0	0	0	0	0	0	1	
		정답률 합계	100	100	100	100	100	100	100	100	100	100	100	100	
3	1	1점	85	70	85	90	80	85	70	80	60	90	83	80	
4-1	1	1점	70	65	80	90	90	70	70	80	70	90	75	77	
4-2	1	1점	95	65	85	75	80	85	60	85	60	80	80	77	

: (후략)

〈표 3〉의 우측에 10명의 패널이 추정한 확률의 중앙값과 평균값이 제시되어 있고, 이를 각각의 배점 또는 부분점수로 가중하고 검사유형별로 합산하면 두 검사유형(가형, 나형)에 대한 수준별 분할점수가 산출된다. 1라운드에서 각 패널이 최초로 문항에 대해 평정한 예상 정답률은 이후에 진행되는 라운드에 비해 차이가 크게 나기 마련이다. 1라운드 결과 〈표 3〉에서 나타나듯이 각 패널이 부여한 예상 정답률에는 편차가 있는데, 특히 패널 간 예상 정답률의 차이가 큰 문항(50%p 이상)을 중심으로 패널들 간에 토의를 진행한 후 1라운드와 동일한 방식으로 2라운드 평정을 진행하였다.

3일차에는 3라운드를 실시하며 1~2라운드를 통해 어느 정도 성취수준별 최소능력자의 특성에 대해 합의 및 내면화가 이루어진 문항들을 제외하고 패널 간 예상 정답률 차이가 20%p 이상인 문항에 대한 최종 논의 과정을 진행하였다.

IV. 연구 결과

1. 예비시행 결과

예비시행 대상 962명 중 결시자를 제외한 857명 사례 수에 대하여 분석을 실시하였다. 예비시행 결과 분석에서는 진단 도구를 구성하는 문항의 특성 분석과 총점 및 영역별 점수에 대한 분석을 수행하였다. 예비 시행 결과 진단 도구 가형과 나형의 총점 및 평균은 〈표 4〉에 제시하였는데, 가형이 나형에 비해 약간 어려운 것으로 나타났으나, t검정 결과, 통계적으로 유의한 차이는 없는 것으로 나타났다.

〈표 4〉 KICE 디지털 소양 진단 평가(KICE DiLAP)의 총점 및 평균

검사 유형	영역	문항	총점	평균	표준 편차	검사 유형 간 평균 차이 검증
가형	1. 디지털 기기와 소프트웨어의 활용	G2, G7-1, G7-2, A15, A16	30	18.29	6.28	t=-0.789 (p=0.430)
	2. 디지털 정보의 활용과 생성	G4-1, G5-1, G5-2, G6, A11, A12				
	3. 디지털 의사소통과 문제해결	G3, G8-1, A10, A17, A18				
	4. 디지털 윤리와 정보보호	G4-2, G8-2, A13, A14				
나형	1. 디지털 기기와 소프트웨어의 활용	G2, G7-1, G7-2, B17*	30	18.61	5.23	
	2. 디지털 정보의 활용과 생성	G4-1, G5-1, G5-2, G6, B16, B17*				
	3. 디지털 의사소통과 문제해결	G3, G8-1, B13, B14, B15, B18				
	4. 디지털 윤리와 정보보호	G4-2, G8-2, B10, B11, B12				

* 문항 B17은 2가지 영역(1. 디지털 기기와 소프트웨어 활용, 2. 디지털 정보의 활용과 생성)에 포함되는 복합형 문항임.

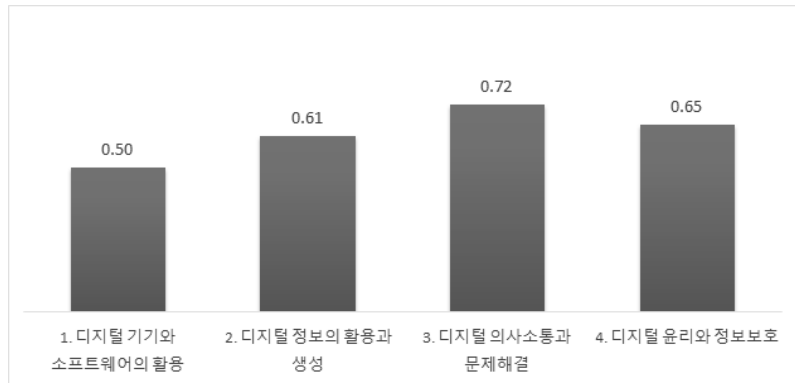
KICE 디지털 소양 진단 평가의 가형과 나형에 배치되었던 문항의 특성을 분석하기 위해 우선 디지털 소양의 4개 영역을 구성하는 문항의 정답률과 부분점수별 비율을 <표 5>에 제시하였다. 배점이 2점 이상이고 부분점수가 있는 다분 문항의 경우, 정답률은 (평균점수)/(배점) 값을 제시하였다. 한편, 문항 1번과 가9번, 나9번은 학생이 온라인 상에서 수행형 문항 풀기에 쉽게 적응할 수 있도록 간단한 지시 사항을 따라 해보는 방식으로 평가 시작과 중간에 배치되었던 연습문제이다. 따라서 <표 5>에 제시된 문항 특성 분석에서 이러한 연습문제들은 제외하였다.

<표 5> 문항 특성 분석 결과

영역	문항번호	배점	부분점수별 비율(%)								정답률
			0점	1점	2점	3점	4점	5점	6점	7점	
1. 디지털 기기와 소프트웨어의 활용 (배점 16)	2	2	8.8	12.7	78.5	—	—	—	—	—	0.85
	7-1	1	43.9	56.1	—	—	—	—	—	—	0.56
	7-2	1	58.3	41.7	—	—	—	—	—	—	0.42
	가15	3	44.2	14.2	18.5	23.1	—	—	—	—	0.40
	가16	2	52.3	39.0	8.7	—	—	—	—	—	0.28
	나17*	7	60.7	2.5	2.8	5.3	13.0	3.9	7.7	4.2	0.24
2. 디지털 정보의 활용과 생성 (배점 18)	4-1	1	33.7	66.3	—	—	—	—	—	—	0.66
	5-1	1	4.8	95.2	—	—	—	—	—	—	0.95
	5-2	2	5.7	24.3	70.0	—	—	—	—	—	0.82
	6	3	17.6	21.8	14.6	46.0	—	—	—	—	0.63
	가11	1	55.1	44.9	—	—	—	—	—	—	0.45
	가12	2	52.6	16.1	31.3	—	—	—	—	—	0.39
3. 디지털 의사소통과 문제해결 (배점 11)	나16	1	25.3	74.7	—	—	—	—	—	—	0.75
	나17*	7	60.7	2.5	2.8	5.3	13.0	3.9	7.7	4.2	0.24
	3	1	27.3	72.7	—	—	—	—	—	—	0.73
	8-1	1	12.8	87.2	—	—	—	—	—	—	0.87
	가10	2	8.4	26.4	65.2	—	—	—	—	—	0.78
	가17	2	31.3	—	68.7	—	—	—	—	—	0.69
	가18	1	26.0	74.0	—	—	—	—	—	—	0.74
	나13	1	53.0	47.0	—	—	—	—	—	—	0.47
	나14	1	15.8	84.2	—	—	—	—	—	—	0.84
4. 디지털 윤리와 정보보호 (배점 7점)	나15	1	16.1	83.9	—	—	—	—	—	—	0.84
	나18	1	47.7	52.3	—	—	—	—	—	—	0.52
	4-2	1	22.4	77.6	—	—	—	—	—	—	0.78
	8-2	1	42.8	57.2	—	—	—	—	—	—	0.57
	가13	1	58.4	41.6	—	—	—	—	—	—	0.42
	가14	1	26.0	74.0	—	—	—	—	—	—	0.74
	나10	1	41.1	58.9	—	—	—	—	—	—	0.59
	나11	1	47.4	52.6	—	—	—	—	—	—	0.53
	나12	1	6.0	94.0	—	—	—	—	—	—	0.94

* 문항 나17은 2가지 영역(1. 디지털 기기와 소프트웨어 활용, 2. 디지털 정보의 활용과 생성)에 포함되는 복합형 문항임.

[그림 4]에는 디지털 소양 진단 평가의 가형과 나형을 구성하는 문항의 영역별 평균 정답률을 제시하였다. 4개 영역의 평균 정답률은 0.50 이상이었고, 그 가운데 가장 평균 정답률이 높은 영역은 '3. 디지털 의사소통과 문제해결'로 0.72의 평균 정답률을 보였다. '1. 디지털 기기와 소프트웨어 활용' 영역의 정답률이 비교적 낮은 이유는 전체 문항 중 정답률이 가장 낮은 가16(정답률 = 0.28) 문항과 나17(정답률 = 0.24) 문항을 포함하고 있기 때문으로 해석된다.



[그림 4] 영역별 평균 정답률

KICE 디지털 소양 진단 평가의 신뢰도와 검사 소요시간을 <표 6>에 제시하였다. 문항 간의 일관성을 평가하는 신뢰도 계수인 Cronbach's α 를 산출한 결과, 가형은 0.8, 나형은 0.7이었다. 학생들의 검사 소요 시간 평균은 가형은 약 29분, 나형은 약 26분인 것으로 조사되었다. 학생들은 40분 동안 디지털 소양 진단 도구를 시행하였으나, 약 5분 정도는 연습문제를 통해 온라인 평가 문항에 적응하는 시간을 가졌으므로 이를 제외하면 디지털 소양 진단 도구를 푸는 데 걸린 실제 평균 소요시간은 21~24분인 것으로 파악된다.

<표 6> 신뢰도 및 검사 소요 시간

검사 유형	학생 수	문항 수	신뢰도 (Cronbach's α)	검사 소요 시간 평균 (분)
가형	572	20	0.8	28.9
나형	285	20	0.7	26.1

2. 문항 평정(최소능력자 기대 정답률 추정) 라운드별 결과

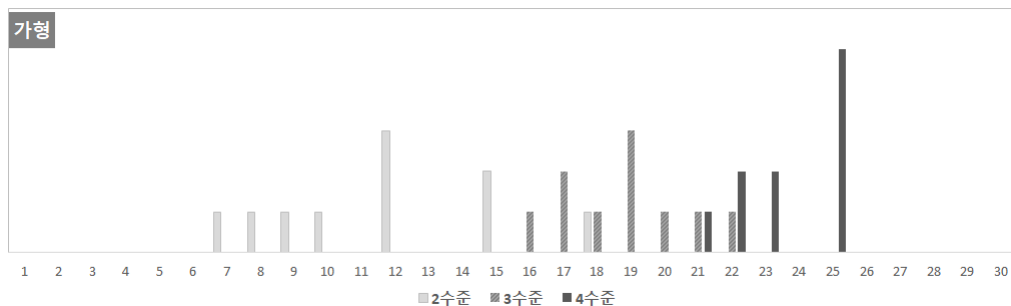
성취수준 설정에 참여한 10명의 전문가 패널은 5명씩 2개조를 편성하여 조별 협의를 진행하였다. 먼저 1라운드에서는 조별 협의 후에 각자 독립적으로 개별 문항을 검토하며 성취수준별 최소능력자의 정답률을 추정하는 절차를 거쳤다. 1라운드 이후부터는 각자 정답률을 추정할 결과에 대한 자료(평균, 중앙값, 최댓값, 최솟값 등)를 제공하고, 조별로 논의를 진행하는 것을 반복하였다.

〈표 7〉에 1라운드에서 변형된 Angoff 방법과 새롭게 제안한 다분문항 기대점수 추정 방법을 통해 산출된 분할점수 평균을 제시하였다. 1라운드에서는 가형의 분할점수들이 나형의 분할점수들보다 다소 높게 추정되는 경향이 나타났다. 변형된 Angoff 방법과 다분문항 기대점수 추정 방법 간에는 큰 차이는 없었으며, 특히 전체 성취수준별 비율에서 4수준(우수)의 비율은 두 방법 간에 차이가 없었다. 구체적으로는 다분문항 기대점수 추정 방법으로 추정된 분할점수가 대체로 변형된 Angoff 방법에 비해 좀 더 낮은 결과를 보였다. 성취수준별 비율에 있어서는 4수준(우수)의 비율은 두 방법 간에 차이가 없었다. 다만, 2수준과 3수준의 경우 학생의 원점수 분포가 밀집되어 있는 점수 구간이라 분할점수의 미세한 위치에 따라 비율에 차이가 나타날 수밖에 없는데, 실제 두 가지 방식(변형된 Angoff 방법/다분문항 기대점수 추정 방법) 간의 분할점수 차이는 1점 이하로 큰 차이가 없었다.

〈표 7〉 1라운드 결과(평균)

방법	성취수준	가형		나형		전체
		분할점수	학생 비율(%)	분할점수	학생 비율(%)	학생 비율(%)
변형된 Angoff 방법	4수준(우수)	24.02	16.78	23.65	17.89	17.15
	3수준(보통)	19.22	29.90	18.50	32.63	30.81
	2수준(기초)	12.26	36.36	12.17	39.65	37.46
	1수준(노력 요망)	-	16.96	-	9.82	14.59
다분문항 기대점수 추정 방법	4수준(우수)	24.40	16.78	23.03	17.89	17.15
	3수준(보통)	18.80	37.24	17.56	40.00	38.16
	2수준(기초)	11.97	31.12	11.36	38.04	32.09
	1수준(노력 요망)	-	14.86	-	8.07	12.60
합계		-	100.00	-	100.00	100.00

〈그림 5〉에 1라운드에서 개별 패널이 예측한 응답률에 기반하여 산출된 분할점수의 분포를 가형 예시로 제시하였다. 1라운드에서는 2수준(기초) 분할점수와 3수준(보통) 분할점수가 18점 부근에서 겹치고 있고, 3수준(보통) 분할점수와 4수준(우수) 분할점수가 21점과 22점 부근에서 겹치고 있음을 확인할 수 있다.



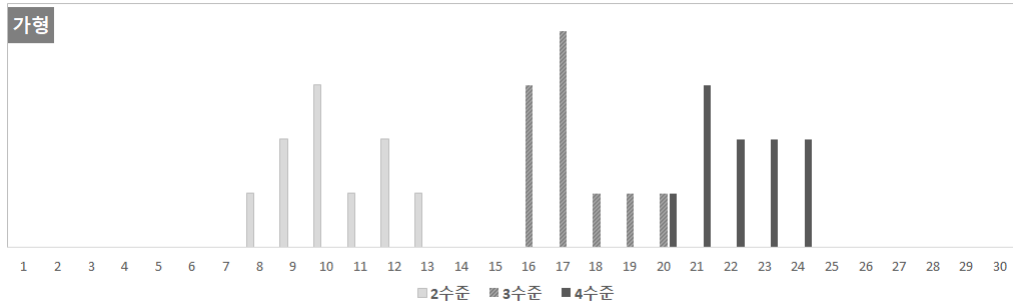
〈그림 5〉 1라운드 분할점수 오버랩(가형)

이어서 <표 8>에 2라운드 결과를 제시하였다. 1라운드 이후에 패널 간에 정답률 예측이 50%p 이상 차이를 보이는 문항을 중심으로 성취수준별 최소능력자의 특성에 대해 논의를 거듭한 후에 2라운드를 진행하였다. 2라운드에서는 문항 외에는 아무런 정보가 제공되지 않았던 1라운드와는 달리 실제 예비시행 학생 데이터를 분석한 결과가 추가로 제공되었다. 추가 정보의 영향으로 2라운드부터는 가형의 분할점수가 나형의 분할점수보다 다소 낮게 나타나기 시작하였다. <표 4>에서 예비시행을 통해 조사된 실제 학생들의 평균 점수는 가형이 나형에 비해 더 낮았기 때문에 가형이 나형에 비해 난이도가 어려웠다고 볼 수 있고, 이를 감안한다면 동일한 수준에 대해 가형의 분할점수가 나형의 분할점수에 비해 다소 낮은 것이 합리적이다. 1라운드에서는 실제 학생들의 예비시행 결과에 대한 정보를 제공하지 않았기 때문에 가형의 분할점수가 더 높았던 것으로 추측되며, 1라운드 이후에 전문가들이 이러한 추가 정보를 접하고 협의를 거치면서 가형의 분할점수가 나형의 분할점수보다 다소 낮게 나타난 것으로 보인다. 또한 가형의 분할점수가 낮아짐으로써 1라운드에 비하여 4수준(우수)의 전체 비율이 17.15%에서 24.39%로 증가하였다. 반면에 1수준(노력 요망) 비율은 14.59%에서 11.44%로 감소하였다.

〈표 8〉 2라운드 결과(평균)

방법	성취수준	가형		나형		전체
		분할점수	학생 비율(%)	분할점수	학생 비율(%)	학생 비율(%)
변형된 Angoff 방법	4수준(우수)	22.58	27.62	23.25	17.89	24.39
	3수준(보통)	17.76	32.69	17.88	40.00	35.12
	2수준(기초)	10.81	26.57	11.28	34.04	29.05
	1수준(노력 요망)	-	13.11	-	8.07	11.44
다분문항 기대점수 추정 방법	4수준(우수)	22.38	27.62	22.66	23.16	26.14
	3수준(보통)	16.69	37.76	17.06	34.74	36.76
	2수준(기초)	10.31	21.50	10.74	36.14	26.37
	1수준(노력 요망)	-	13.11	-	5.96	10.74
합계		-	100.00	-	100.00	100.00

[그림 6]에 2라운드에서 개별 패널이 예측한 정답률에 기반하여 산출된 분할점수의 분포를 가형을 예시로 제시하였다. 2라운드에서는 성취수준별 최소능력자의 문항정답률에 대한 패널들의 의견이 1라운드에 비해 수렴되어 2수준(기초) 분할점수와 3수준(보통) 분할점수가 겹치는 현상은 해소되었으나, 아직 3수준(보통) 분할점수와 4수준(우수) 분할점수가 20점 부근에서 겹치고 있다.



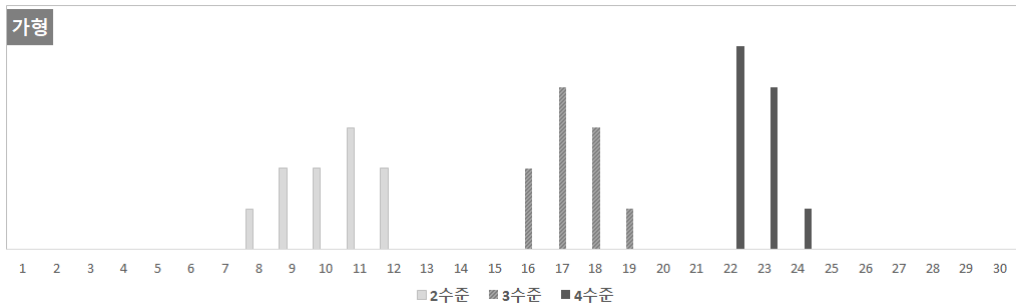
[그림 6] 2라운드 분할점수 오버랩(가형)

이어서 <표 9>에 3라운드 결과를 제시하였다. 2라운드 이후에 패널 간 정답률 예측이 20%p 이상 이견을 보이는 문항들에 대해 재차 논의하며, 각 패널들은 자신의 성취수준별 최소능력자의 특성에 대해 숙고하고 조정하는 시간을 거쳤다. 결과적으로 1라운드 이후 2라운드에서 증가했던 4수준(우수) 비율이 3라운드에서는 1~2라운드의 중간 정도인 20.42%로 정착하였고, 1수준(노력 요망)비율은 2라운드와 동일한 비율로 더 이상 변화를 보이지 않았다.

<표 9> 3라운드 결과(평균)

방법	성취수준	가형		나형		전체
		분할점수	학생 비율(%)	분할점수	학생 비율(%)	학생 비율(%)
변형된 Angoff 방법	4수준(우수)	23.04	21.68	23.31	17.89	20.42
	3수준(보통)	17.81	38.64	17.63	40.00	39.09
	2수준(기초)	10.59	26.57	11.06	34.04	29.05
	1수준(노력 요망)	-	13.11	-	8.07	11.44
다분문항 기대점수 추정 방법	4수준(우수)	22.72	27.62	22.78	23.16	26.14
	3수준(보통)	16.87	37.76	16.94	43.16	39.56
	2수준(기초)	10.22	21.50	10.63	27.72	23.57
	1수준(노력 요망)	-	13.11	-	5.96	10.74
합계		-	100.00	-	100.00	100.00

1~3라운드를 진행하는 동안 패널들은 각 문항에 대한 성취수준별 정답률에 대해 어느 정도 합의를 형성하였는데, 이러한 합의 과정은 [그림 7]에 나타난 바와 같이 최종적으로 패널들의 분할점수 분포가 수준별로 명료하게 분리된 것으로 확인할 수 있다.



[그림 7] 3라운드 분할점수 오버랩(가형)

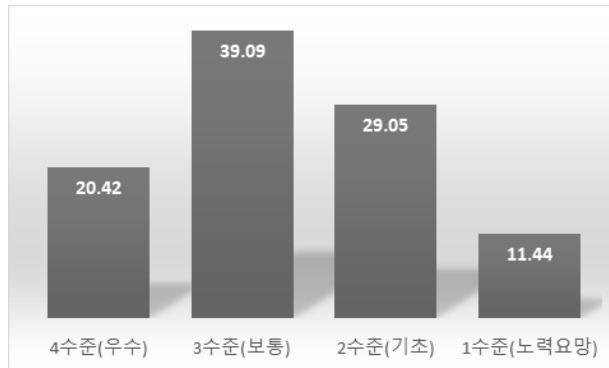
3. 성취수준 설정 최종 결과

본 연구에서는 디지털 소양 진단 평가를 개발하고, 학생들의 디지털 소양 수준을 해석하기 위한 4단계 성취수준을 설정하였다. 성취수준 설정 절차를 거치기 전에 성취수준별 진술문 초안을 작성하고, 성취수준 설정 워크숍에 참여한 전문가 패널들의 최종적인 합의 사항을 참고하여 <표 10>과 같은 성취수준별 진술문 최종안으로 가다듬었다.

<표 10> KICE 디지털 소양 진단 평가(KICE DiLAP)의 4단계 성취수준 정의(최종안)

성취수준	정 의
4수준(우수)	디지털 기기와 소프트웨어(인공지능 포함)의 기본 원리와 기능을 이해하고 능숙하게 활용할 수 있다. 목적에 알맞은 다양한 자료를 검색·수집·저장하고, 디지털 콘텐츠를 실제로 제작하여 정보를 효과적으로 전달할 수 있다. 디지털 환경에서 능숙하게 의사소통하고 정보를 비판적으로 평가할 수 있으며, 상황에 알맞은 방법으로 문제를 해결할 수 있다. 또한 디지털 윤리에 대한 이해를 바탕으로 온라인에서 안전하고 책임 있는 행동을 적극적으로 실천할 수 있다.
3수준(보통)	디지털 기기와 소프트웨어(인공지능 포함)의 기본 원리와 기능을 이해하고 활용할 수 있다. 목적에 알맞은 기본적인 자료를 검색·수집·저장하고, 디지털 콘텐츠를 실제로 제작하여 정보를 전달할 수 있다. 디지털 환경에서 의사소통하고 정보를 일정 수준 비판적으로 평가할 수 있으며, 문제를 일정 수준 해결할 수 있다. 또한 디지털 윤리에 대한 이해를 바탕으로 온라인에서 안전하고 책임 있는 행동을 실천할 수 있다.
2수준(기초)	디지털 기기와 소프트웨어(인공지능 포함)의 기본 원리와 기능을 이해하고 간단히 활용할 수 있다. 목적에 알맞은 자료를 검색·수집·저장하는 데 일부 도움이 필요하며, 디지털 콘텐츠를 제작하기 위한 절차를 이해하고 필요한 자료 선택을 할 수 있다. 디지털 환경에서 의사소통할 수 있으나 정보를 비판적으로 평가하는 데 일부 도움이 필요하며, 간단한 문제를 해결할 수 있다. 또한 디지털 윤리에 대한 이해를 바탕으로 온라인에서 안전하고 책임 있는 행동을 일정 수준 실천할 수 있다.
1수준(노력 요망)	디지털 기기와 소프트웨어(인공지능 포함)의 기본 원리와 기능을 이해하고 활용하는 데 어려움이 있다. 목적에 알맞은 자료를 검색·수집·저장하고, 디지털 콘텐츠를 제작하는 데 어려움이 있다. 디지털 환경에서 의사소통을 할 수 있으나 정보를 비판적으로 평가하거나 간단한 문제를 해결하는 데 어려움이 있다. 또한 디지털 윤리에 대한 이해를 바탕으로 온라인에서 안전하고 책임 있는 행동을 기본적인 수준으로 실천할 수 있다.

즉, <표 10>에 제시된 성취수준 정의를 토대로 본 연구에서는 4단계 성취수준별 분할점수를 설정하였다고 볼 수 있다. [그림 8]은 성취수준 설정 절차의 3라운드에서 얻은 평균에 기반한 분할점수를 적용하여 4단계 성취수준별 학생 비율을 산출한 결과이다. 초등학교 6학년 학생들의 디지털 소양은 4수준(우수) 비율이 20.42%로 나타났고, 1수준(노력 요망)의 비율은 11.44%인 것으로 나타났다.



[그림 8] 성취수준별 학생 비율

V. 결론 및 논의

본 연구에서는 디지털 소양 진단 평가의 문항 개발 이후에 검사의 측정학적 특성을 점검하고, 향후 교사들이 수월하게 평가 결과를 활용할 수 있도록 성취수준 설정을 진행하였다. 본 연구에서 개발한 2023년 디지털 소양 진단 평가의 두 가지 검사 유형(가형/나형)은 각각 20문항(총점 30점)으로 구성되었으며 그 가운데 11문항(15점)은 가교문항으로 배치하였다. 두 가지 검사 유형에 공통으로 배치된 가교문항을 통해 검사 유형 간 결과 비교가 가능하고, 향후 추가적인 검사 유형 개발 시에도 가교문항을 배치함으로써 비교 가능성과 연속성을 유지할 수 있도록 설계하였다.

본 연구에서 개발한 디지털 소양 진단 평가는 실제 디지털 환경에서 학생들이 디지털 기기를 활용하고, 정보를 검색·저장하거나 디지털 콘텐츠를 생성하는 능력을 타당하게 평가할 수 있도록 오프라인 지필식 평가 문항이 아닌 온라인 평가 문항으로 제작되었으며, 예비시행도 온라인 평가로 진행하였다. 962명을 대상으로 예비시행을 실시하였고, 결시자를 제외한 857명의 예비시행 결과 자료를 확보하였다. 예비시행 후에 디지털 소양 진단 도구에 대해 학생 대상으로 간단한 만족도 조사를 실시한 결과, 80% 이상의 학생이 문항 지시의 명확성, 검사 길이의 적정성, 난이도의 적정성에 대해 부정적이지 않았으며, 약 74%의 학생이 전반적으로 디지털 소양 진단 도구에 대해 만족하는 것으로 나타났다. 2022 개정 교육과정에서 디지털 소양 교육을 강조하고 있으나, 실제로 학교현장에서 디지털 소양 교육을 추진하고자 한다면 어느 정도의 난이도와 길이로 구성된 진단 도구 또는 교수·학습 자료를 계획·활용해

야 할지에 대한 기초 자료가 부족한 상황이다. 본 연구에서 제시한 디지털 소양 진단 도구에 대한 만족도 조사는 이러한 상황에서 디지털 소양 진단 도구나 교수·학습 자료의 적정 난이도 및 길이 등에 대해 유용한 기초적인 참고 자료를 제공할 수 있을 것으로 예상된다.

다음으로는 예비시행 결과를 분석함으로써 본 연구에서 개발한 디지털 소양 진단 도구 2가지 검사 유형의 양호도를 검증하였다. 디지털 소양 진단 도구의 2가지 검사 유형에 포함된 29개 문항(공통 11 문항 + 가형 9문항 + 나형 9문항)에 대한 문항 특성을 분석한 결과, 정답률 평균은 0.64였다. 가장 어려웠던 문항은 ‘1. 디지털 기기와 소프트웨어 활용’, ‘2. 디지털 정보의 활용과 생성’의 2개 영역을 모두 포함하도록 복합형 문항으로 출제한 나17번 문항(정답률 = 0.24)이었고, 가장 쉬웠던 문항은 ‘2. 디지털 정보의 활용과 생성’ 영역에 속한 5-1번 문항(정답률 = 0.95)이었다. 나17번 문항이 다른 디지털 콘텐츠를 제작하는 문항과 다른 점은 채점 요소에 학생이 실제로 디지털 콘텐츠의 내용을 구성하고 제작할 수 있는지 여부 외에도 프레젠테이션 템플릿을 활용하여 서식을 목적에 맞게 적절히 조정할 수 있는지 여부를 포함하고 있다는 점이었다. 학생들이 다른 디지털 콘텐츠의 내용 구성을 요구하는 문항에서 크게 어려워하지 않은 점과 비교하면, 디지털 콘텐츠의 내용 구성보다 소프트웨어를 능숙하게 활용하여 서식을 변경하고 조정하는 등의 기능적인 측면을 어렵하다는 사실을 확인할 수 있다. 반면에 가장 쉬운 문항이었던 5-1번 문항은 기사를 읽고 컴퓨터 화면에 알맞은 수치를 채워 넣어 표를 완성하는 문항이다. 디지털 콘텐츠를 생성하는 문항이지만, 소프트웨어(엑셀 등)를 활용하는 문항이 아니므로 스프레드시트 소프트웨어 사용법을 모르더라도 화면에 나타난 표에 정확한 수치를 채워 넣으면 되기 때문에 학생들이 수월하게 정답을 맞힐 수 있었던 것으로 추측된다. 또한 디지털 소양의 4개 하위 영역별 점수를 비교한 결과, ‘3. 디지털 의사소통과 문제해결’ 영역의 평균이 가장 높게 나타났고, 다음으로 ‘4. 디지털 윤리와 정보보호’ 영역의 평균이 높았으며, ‘1. 디지털 기기와 소프트웨어 활용’ 영역의 평균이 가장 낮았다. 이를 통해 학생들은 의사소통을 위한 이메일 등의 기반 지식 이해 및 활용을 대체적으로 쉬워하는 반면, 디지털 콘텐츠 생성 시 엑셀이나 파워포인트 등의 소프트웨어의 활용을 어려워한다는 점을 재차 확인할 수 있었다.

그리고 본 연구에서 개발한 디지털 소양 진단 도구의 2가지 검사 유형에 대한 4단계 성취수준을 설정하였다. 본 연구에서 성취수준 설정을 수행한 목적은 디지털 소양 진단 도구 평가 결과를 통해 학교 현장에서 교사들이 학생의 성취수준 정도를 정확히 진단하고, 디지털 소양 교육 내용 및 방법을 학생의 수준에 알맞게 계획할 수 있도록 지원하고자 함이다. 본 연구에서 개발한 디지털 소양 진단 도구의 평가 결과를 활용할 때, 총점이나 석차를 상대적으로 비교하는 방식이 아니라, 교육과정 상에 제시된 교육목표에 대한 학생의 도달 정도에 따라 학생의 절대적인 성취수준을 판단할 필요가 있기 때문이다. 성취수준 설정을 통해 산출된 성취수준 분할점수를 예비시행 자료에 적용하여 4단계 성취수준별 학생 비율을 확인한 결과, 초등학교 6학년 학생들의 디지털 소양은 4수준(우수) 비율이 20.42%로 나타났지만, 1수준(노력 요망)의 비율도 11.44%인 것으로 나타났다. 서론에서 언급하였듯이 디지털 소양은 미래 사회의 필수 역량이면서 AI 디지털 교과서 도입 시 교과 교수·학습의 수월성에 영향을 끼칠 수 있다는 점을 고려할 때, 공교육에서 학생들의 디지털 소양을 체계적으로 함양하여 1수준(노력 요망) 학생의 비율을 감소시킬 수 있도록 학교 현장에서의 노력과 이를 지원할 정책적 노력이 필요할 것이다.

한편, 방법론적인 측면에서 본 연구의 차별점은 성취수준 설정 절차를 위해 변형된 Angoff 방법을

적용하면서 동시에 새로운 수준 설정 방법을 추가적으로 적용하는 것을 시도했다는 점이다. 변형된 Angoff 방법에서는 다분문항의 경우, 각 부분점수마다 성취수준별 최소능력자의 정답률을 추정해야 한다는 점에서 패널들에게 큰 부담이다. 이에 본 연구에서는 이분문항에 대해서는 변형된 Angoff 방법을 적용하지만, 추정 부담이 큰 다분문항에서는 성취수준별 최소능력자가 얻을 것이라 기대되는 부분점수를 소수점 첫째자리까지 추정하도록 하는 ‘다분문항 기대점수 추정 방법’을 제안하였다. 연구 결과에 의하면, 다분문항 기대점수 추정 방법을 적용한 경우, 변형된 Angoff 방법에 비하여 4수준(우수) 비율은 약간 증가하고, 1수준(노력 요망) 비율은 약간 감소하는 것을 확인할 수 있었다. 그리고 3수준(보통) 비율은 변형된 Angoff 방법과 다분문항 기대점수 추정 방법 사이에 1%p 이내의 차이를 보여 거의 유사한 결과를 산출했다고 볼 수 있다. 다분문항 기대점수 추정 방법은 학교현장의 교사들이 활용하기에 쉽고 간편한 장점이 있으므로, 향후에 좀 더 엄밀한 비교 연구를 거쳐 Angoff 방법과 크게 차이하지 않는 결과를 얻는다면, 대체적인 수준설정 방법으로 활용될 수 있을 것으로 기대된다.

본 연구 결과를 종합하면 KICE 디지털 소양 진단 평가를 통해 초등학교 6학년 학생들의 디지털 소양에 있어 강약점을 파악할 수 있었고, 4단계의 성취수준별로 학생들이 디지털 도구를 활용하여 무엇을 할 수 있고 어떠한 부분이 부족한지에 대한 정교한 해석도 가능하였다. 향후 평가 대상을 현재의 초등학교 6학년에서 학교급과 학년을 확대해 나간다면 초중등 교육 전반에서 학생들의 디지털 소양을 정확하게 진단하는 데에 유용하게 활용될 수 있을 것이다. 그리고 한발 더 나아가 KICE 디지털 소양 진단 평가 결과와 연계한 학생들의 디지털 소양 취약점을 해소하기 위한 교수·학습 방안, 그리고 각 성취수준에 해당하는 학생의 수준을 고려한 효율적인 교수·학습 방안을 강구할 필요성이 제기된다.

참고문헌

- 교육부(2021). **2022 개정 교육과정 총론 주요사항(시안)**. 세종: 교육부.
- 김경희, 박은아, 송미영, 상경아, 김수진, 김희경, 신진아, 서지영, 이채희, 김지영, 김기연, 최숙기 (2012). **중등학교의 성취평가제 운영 방안 연구**. 한국교육과정평가원 연구보고 CRE 2012-8.
- 김자미, 강원석, 김수환, 김재현, 김태영, 김한일, 박선주, 안성훈, 전성균, 전용주, 강신옥, 노경보, 박다솜, 서성원, 송석리, 안득하, 안상진, 이정서, 정웅열, 정종광, 최정원(2022). **2022 개정 정보과 교육과정 시안(최종안) 개발 연구**. 한국과학창의재단.
- 김창환. (2011). 수정된 Angoff 방법을 활용한 교과서 검정 심사에서의 분할점수 설정. **한국교육**, 38(4), 141-162.
- 김현경, 정진수(2018). 과학과 교육과정 개정에 따른 국가수준 학업성취도 평가의 성취수준 설정에 관한 연구. **학습자중심교과교육연구**, 18(5), 305-330.
- 김희경, 김완수, 김수진, 정혜경, 김미림, 김성훈(2019). **국가수준 학업성취도 평가 점수 체제 개선 및 결과 활용도 제고 방안**. 한국교육과정평가원 연구보고 RRE 2019-3.
- 김희경, 조성민(2013). 기초학력 진단평가의 성취수준 설정에 관한 연구: 초등학교 3학년 '기초수학' 사례를 중심으로. **한국학교수학회논문집**, 16(3), 479-498.
- 배주경, 박소영, 이미경, 권점례, 이소연, 이지수(2023). **2022 개정 교육과정에 따른 교과 교육과정 적용 방안 연구(I): 초등학교를 중심으로**. 한국교육과정평가원 이슈페이퍼 ORM 2023-30-5.
- 전성균, 김희경, 이영미, 김영실, 한정아(2023). **2022 개정 교육과정에 따른 디지털 소양 진단 도구 개발**. 한국교육과정평가원 연구보고 RRE 2023-5.
- 이현숙, 이운지, 차현진, 김수환, 나우열, 계보경, 한나라(2022). **2022년 국가수준 초·중학생 디지털 리터러시 수준 측정 연구**. 한국교육학술정보원 연구보고 KR 2022-2.
- 성태제(2010). **현대교육평가**. 서울: 학지사.
- Angoff, W. H. (1971). Scales, norms and equivalent scores. In R. L. Thorndike (ed.), *Educational measurement*, 2nd ed., Washington, DC: American Council on Education.
- Cizek, G., & Bunch, M. (2007). *Standard-setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.

IEA. (2020). *IEA international computer and information literacy study 2018 technical report*. International Association for the Evaluation of Educational Achievement.
Retrieved 19, June, 2024, from
<https://www.iea.nl/publications/technical-reports/icils-2018-technical-report>.
Livingston, S. A., & Zieky, M. J. (1982). *Passing score*. Educational Testing Service.

• 논문접수 : 2024.07.05. / 수정본접수 : 2024.07.31. / 게재승인 : 2024.08.12.

ABSTRACT

A Study on Setting Achievement Levels for Digital Literacy Assessment Tools for Sixth-Grade Elementary Student

HeeKyoung Kim

Senior Research Fellow, Korea Institute for Curriculum and Evaluation

Minho Park

Associate Research Fellow, Korea Institute for Curriculum and Evaluation

With the introduction of digital literacy as a foundational competency in the 2022 revised curriculum, there is an increasing emphasis on digital literacy education through school education. However, information to help teachers provide digital literacy education that meets the students' levels is relatively insufficient. This study examines the psychometric properties of the KICE Digital Literacy Assessment Program(KICE DiLAP) for sixth-grade students, following the development of its items. It also establishes achievement levels to enable teachers to easily diagnose students' digital literacy levels in the future. To this end, a preliminary implementation and scoring of the assessment tool were conducted, and data from 857 students were collected for item characteristic analysis and total and subdomain score analysis. Ultimately, four achievement levels for digital literacy were established. The results indicated that about 20% of sixth-grade students in South Korea reached Level 4 (excellent), while approximately 11% were at Level 1 (needs improvement). Additionally, representative item analysis by achievement level provided a detailed understanding of students' digital literacy at each level. These findings offer valuable data for teachers to use the KICE Digital Literacy Assessment Program(KICE DiLAP) as needed to diagnose students' digital literacy levels and to interpret the results easily, thereby facilitating the development of efficient teaching and learning strategies tailored to learners' levels.

Key Words: 2022 Revised Curriculum, Digital Literacy, Digital Literacy Assessment Program, Setting Achievement Level

