

교육과정평가연구
The Journal of Curriculum and Evaluation
2023, Vol. 26, No. 4, pp. 47~74
DOI: <https://doi.org/10.29221/jce.2023.26.4.47>

초·중학생 대상 빈칸 메우기 검사를 통한 교사의 교육용 텍스트 이독성 수준 평정의 타당성 분석¹⁾

류수경 (한국교육과정평가원 부연구위원)*
박태준 (한국교육과정평가원 연구위원)
박은아 (한국교육과정평가원 선임연구위원)
변태진 (광주교육대학교 과학교육과 조교수)
한정아 (한국교육과정평가원 전문연구위원)
하명진 (한국교육과정평가원 전문연구위원)
김지혜 (경기도교육연구원 연구위원)
최소영 (부산대학교 조교수)**

요약

본 연구는 교사가 교육용 텍스트의 이독성 수준을 평정한 결과가 학생 입장에서의 이독성 수준을 적절하게 반영하고 있는지 검증하는 것을 목적으로 한다. 이를 위해 다음과 같은 일련의 절차를 거쳤다. 먼저 교사가 이독성 수준을 평정한 텍스트 중에서 검사용 텍스트를 선정한 후 서답형의 빈칸메우기 검사지를 개발하였다. 이어서 초등학교 6학년과 중학교 3학년 학생을 대상으로 예비검사 총 286명 및 본검사 총 421명을 대상으로 빈칸메우기 검사를 시행하고, 교사 채점단이 서답형 답안을 채점하도록 하였다. 평정 및 채점 결과에 대한 통계 분석으로는 Rasch 모형을 통해 교사의 텍스트 수준 평정 서열을 도출한 다음, Spearman's Rho를 활용한 서열 상관관계 분석을 통해 교사의 텍스트 수준 평정 서열과 학생 대상 이독성 검사의 문항 난이도 서열 간의 상관관계를 분석하였다. 그 결과 초등학교 대상 본검사 결과는 0.8154, 중학생 대상 본검사 결과 0.8374로 상관관계가 높게 나타나 교사의 텍스트 이독성 수준 평정의 타당성이 확보되었음을 확인할 수 있었다. 이를 토대로 향후 더 발전적인 이독성 지수 개발의 타당성을 확보하기 위해서 학생 대상 이독성 검사를 다각화하여 활용하는 방안에 대해 제안하였다.

주제어: 이독성 지수, 빈칸 메우기 검사, 교사의 이독성 수준 평정, 상관관계 분석, Rasch 모형

1) 이 논문은 박태준 외(2022) 'KICE 이독성 지수 (KICE Readability Index) 자동측정 프로그램 설계 및 개발 연구 (I)(한국교육과정평가원 연구보고 RRC 2022-9)'의 내용 중 학생 대상 텍스트 이독성 검사를 통한 교사의 텍스트 수준 평정 타당성 분석에 관련된 부분을 수정·보완한 것임.

* 제1저자, bluemoon14@kice.re.kr

** 교신저자, soyoungchoi@pusan.ac.kr

I. 서 론

맞춤형 교수·학습이 강조되면서 학습자에게 적절한 수준의 텍스트를 제공하는 것은 국어과를 비롯한 모든 교과 학습에서 중요한 화두이다. 학습자의 흥미, 그리고 능력 수준에 맞는 텍스트를 제공하면 학습 동기를 고취시킬 수 있고 학습 목적 읽기 능력을 길러 교과 학습의 효과를 높일 수 있기 때문이다. 이러한 목적에서 교과용 텍스트의 이독성을 양적으로 평가하는 이독성 지수 개발과 질적으로 평가하는 교사의 이독성 지수 평정에 관한 연구는 지금까지 꾸준히 진행되어 왔다. 이독성 지수와 교사의 이독성 평정은 교육용 텍스트의 수준을 판단하는 데 상호보완적인 역할을 하기 때문에 한국어 언어 구조에 알맞은, 그리고 교육용 목적에 알맞은 이독성 평가 도구가 개발되고 널리 보급될 필요가 있다.

공신력 있는 이독성 지수가 부재한 국내 여건에서, 교사의 이독성 수준 평정은 표층적인 텍스트 요인만으로 평가할 수 없는 다양한 요인을 종합하여 질적으로 평정할 수 있기에 학교 현장의 활용도가 높다(최소영 외, 2022). 또한 교사의 이독성 수준 평정은 양적 평가 도구인 이독성 지수를 개발하는 데에도 매우 중요한 역할을 할 수 있다(최소영, 박태준, 2022). 이독성 지수는 준거 텍스트의 수준을 기준으로 새로운 텍스트의 수준을 예측하므로 준거 텍스트의 수준을 평정하는 것이 타당하게 이루어져야 하는데, 이때 교사의 전문적인 감식안이 준거 텍스트의 수준 평정하는 데 주요한 역할을 하기 때문이다.

그러나 교사와 학생 간에 교육용 텍스트의 난도 평정에 차이가 있다는 연구(김지은, 2012), 학생들을 대상으로 텍스트 난도 평정을 시행하는 연구(김해인, 2019) 등을 통해 실제 학생들의 눈높이를 고려한 이독성 지수에 대한 요구가 높아지고 있다. 교사는 학생들의 독해력과 어휘력에 대한 이해가 깊고 교과 학습 내용의 위계를 충분히 이해하고 있으므로 교육용 텍스트의 이독성 수준을 평정할 수 있는 전문성이 있지만, 교육용 텍스트의 주된 독자는 학습자라는 점에서 이독성 지수를 개발할 때에는 실제 학생들이 교육용 텍스트를 읽고 이해하는 정도와 교사의 텍스트 수준 평정이 얼마나 상관이 있는지 확인하여 교사의 학년 수준 평정이 타당한지 확인할 필요가 있다. 이는 텍스트의 수준을 교사의 평정에만 의존하지 않고 학습자의 이해 정도와 연계하여 이독성 지수 개발의 타당성을 높이고자 하는 노력이기도 하다.

본 연구에서는 KICE 이독성 지수(KICE Readability Index)를 개발하는 과정에서 교사의 수준 평정이 이루어진 일부 텍스트를 활용하여 학생 대상 텍스트 이독성 검사 결과를 실시하고, 교사의 수준 평정 서열과 문항의 난이도 서열 간의 상관 정도를 확인하여 준거 텍스트에 대한 교사의 학년 수준 평정의 타당성을 분석하고자 한다. 구체적인 방법으로는 학생 대상 텍스트 이독성 검사 결과에 대해 Rasch 모형을 적용하여 문항의 난이도 특성을 파악한 다음 서열 상관관계분석을 통해 교사의 수준 평정 서열과 학생 대상 텍스트 이독성 검사 결과에서 나타난 텍스트의 난이도 서열 간의 관계를 분석하는 것이다. 교사가 평정한 텍스트 수준과 학생 검사 결과 나타난 텍스트 난이도 사이에 상관관계가 높다는 것은 교사 평정의 타당성을 입증하는 방법이 될 수 있다.

II. 학생 대상 이독성 검사에 관한 선행연구 검토

1923년 미국에서 최초의 이독성 공식(Lively & Pressey, 1923)이 발표된 이후, 영어 텍스트에 적용되는 이독성 공식이 다양하게 개발되었다. 이를 두고 Mesmer(2008)는 이독성 측정 도구가 발표된 시기 및 주요 특징에 따라 크게 1세대와 2세대로 구분하였다. 대부분의 1세대 도구는 텍스트 내적 요인에 주목하여 텍스트의 수준을 위계화하는 데에 치중하였다면 2세대 도구는 텍스트의 이독성 지수를 산출할 뿐만 아니라 해당 텍스트를 읽기에 적합한 학생들의 독해 수준이 어느 정도인지 알려주거나 학생의 독해 능력에 따른 맞춤형 텍스트를 제시해 준다.

학생의 독해 능력에 따라 맞춤형 텍스트를 제시할 수 있었던 것은 학생들이 준거 텍스트를 얼마나 읽고 이해하는지를 확인할 수 있는 학생 대상 이독성 검사가 있었기 때문이다. 일찍이 이독성 지수 개발에서 학생 대상 이독성 검사를 적용한 것은 The New Dale-Chall Readability Formula(Chall & Dale, 1995)이다. 이 연구는 초기 공식인 Dale-Chall Readability Formula(Dale & Chall, 1948)를 개정하면서 '기초 어휘 3,000개 목록'을 업데이트하였고, 기존에 시행된 빈칸 메우기 검사 결과와 연계함으로써 텍스트의 문장 길이와 어휘 수준을 분석하면 학생이 읽었을 때 예상되는 빈칸 메우기 점수와 학년 수준을 각각 산출할 수 있도록 설계되었다. 이는 빈칸 메우기 점수를 예측하도록 회귀방정식을 개발하고 이를 학년 수준으로 환산하여 제시함으로써 가능하였다. 이와 같이 텍스트의 양적 요인으로부터 빈칸 메우기 점수를 예측하는 공식은 텍스트 이독성을 단지 텍스트 간의 수준 위계화에만 두지 않고 실제 학생이 읽고 이해하는 정도와 연계하려는 시도였다.

국내에서는 조용구(2016)가 The New Dale-Chall Readability Formula를 참고하여 이독성 지수 개발 연구 과정에서 학생 대상 이독성 검사를 실시하였다. 이 연구에서는 문장의 길이와 명사 비율을 통해 학생의 이독성 검사에서의 빈칸 메우기 점수를 예측하는 이독성 공식을 개발하고, 빈칸 메우기 점수를 학년 수준으로 변환해서 해석할 수 있도록 하였다.

이후 영어권에서는 컴퓨터 기술의 발전으로 자연어 처리 기술이 정교화 되고 텍스트를 수월하게 분석할 수 있게 되면서 다양한 2세대 이독성 측정 도구가 나타났다. 개발의 주체가 개인 연구자에서 대형 출판사 중심으로 옮겨 가면서 대규모 텍스트 데이터베이스를 기반으로 이독성 자동측정 도구가 개발되었다.

현재 영어권 국가에서 가장 널리 사용되고 있는 이독성 지수인 Lexile은 학생의 읽기 능력과 텍스트의 난도를 동일한 척도(200~1700L) 상에서 측정함으로써 실제 읽기 능력 평가 결과에 근거한 학생의 읽기 능력 수준에 맞는 텍스트를 선택할 수 있다(Mesmer, 2008). ATOS(Advantage TASA Open Standard for Readability)는 Lexile과 함께 전 세계 영어권 국가에서 가장 널리 사용되고 있는 이독성 지수 중 하나이며, 학생의 독해 능력에 따라 Accelerated Reader(AR)라는 독서 프로그램을 연계하여 활용하도록 안내하고 있다. 이처럼 Lexile과 ATOS가 학생들의 독해 능력에 따라 맞춤형 텍스트를 제시해 줄 수 있는 것은 기존의 표준화된 대규모 읽기 평가 데이터와 연계되어 개발되었기 때문이다. 즉, 이독성 지수를 개발하기 위해 텍스트만 수집하는 것이 아니라 해당 텍스트를 읽은 학생의 반응 자료까지 고려한 텍스트의 이독성 특성을 통해 학생들의 독해 결과를 예측할 수 있는 모델을

구축하는 것이다.

DRP(Degrees of Reading Power)는 백분율로 표시되는 척도(0~100)로 텍스트의 이독성을 측정하며, 숫자가 높을수록 어려운 텍스트를 의미한다. DRP는 빈칸 메우기 검사를 통해 텍스트 난도를 측정하는 Bormuth cloze 평균 공식(Bormuth, 1971)에 그 기반을 두고 있다. 기존의 표준화된 읽기 평가와 연계하여 학생의 읽기 능력 수준을 측정하는 Lexile과는 달리, DRP는 텍스트 내에 빈칸이 포함된 형태(modified cloze passages)의 자체적인 평가를 통해 학생 수준과 텍스트 난도를 평가한다(Mesmer, 2008; Nelson et al., 2012).

지금까지 선행 연구를 참고할 때 이독성 지수를 개발하는 과정에서 학생 대상 텍스트 이독성 검사가 필요할 때 활용할 수 있는 방식은 빈칸 메우기 검사와 표준화 독해 검사가 있다. 그중에서도 빈칸 메우기 검사가 이독성 지수 개발 과정에서 학생 대상 텍스트 이독성 검사로 많이 활용되는 것은 무엇보다 텍스트 특성에 기반한 독해 결과를 확보할 수 있기 때문이다. 일반적인 독해 표준화 검사는 사실적 이해와 추론적 이해뿐만 아니라 비판적 이해, 감상적 이해, 창의적 이해 등 텍스트를 이해하는 데에 요구되는 사고 능력을 다방면에서 평가한다면, 빈칸 메우기 검사는 앞뒤 문맥을 통해 빈칸을 메우게 함으로써 사실적 이해와 추론적 이해 능력에 초점을 두고 독해 능력을 평가할 수 있다. “읽고 이해하기 쉬운 정도(이성영, 2011: 170)”를 뜻하는 이독성은 그 수준을 평정할 때에 과제 요인과 독자 요인이 크게 작용하는 ‘비판적 이해나 감상적 이해, 창의적 이해’보다는 텍스트 특성과 긴밀하게 연결되는 ‘사실적 이해’와 ‘추론적 이해’ 수준에서 초점을 두고 판단하는 것이 적절하다. 또한 독해 표준화 검사는 독해 결과를 평가하는 반면, 빈칸 메우기 검사에서는 독해 결과와 과정을 모두 평가할 수 있고, 문맥을 활용하도록 한다는 점에서 학습자의 읽기 전략 사용을 유도하고 촉진할 수 있다는 이점도 있다(박수자, 2001). 이러한 점에 기초하여 조현재·김영수(2000)에서는 빈칸 메우기 검사를 통해 생물 교과서의 이독성 측정하는 방안을 제안하고 검증하였고, 조용구·이경남(2013)의 연구에서도 사실적 읽기와 추론적 읽기에 초점을 두고 읽기 능력 평가를 위하여 빈칸 메우기 검사 도구를 개발하였다.

이에 본 연구에서는 서로 다른 학습자가 동일한 텍스트를 읽고 얼마나 쉽고 어렵게 느끼는지를 파악할 수 있는 빈칸 메우기의 형태로 초·중학생 대상 텍스트 이독성 검사를 개발하였다. 검사 문항 개발 시 선다형 문항의 적용도 검토하였으나 문항 풀이 과정에서 오답지 매력도의 영향을 고려하여, 조용구(2016)의 연구에서와 같이 응시자가 직접 답안을 기입하는 서답형 문항으로 문항 유형을 최종 결정하였다.

III. 연구 절차 및 방법

본 연구는 KICE 이독성 지수(KICE Readability Index)를 개발하는 과정에서 교사의 수준 평정이 이루어진 일부 텍스트를 활용하여 학생 대상 텍스트 이독성 검사 결과를 실시하고 교사의 학년 수준 평정의 타당성을 분석한다. 이때, 교사의 텍스트 수준 평정은 KICE 이독성 지수 개발 과정에서 수집한 교육용 텍스트 데이터베이스 전체에 대해 이루어졌으며, 학생 대상 텍스트 이독성 검사는 교사의 수준

평정이 이루어진 일부 텍스트를 활용하여 이루어졌다.

먼저, KICE 이독성 지수 개발을 위한 텍스트 데이터베이스의 수준 평정은 기존의 교과서 수록 학년을 그대로 활용하지 않고 범교과적 ‘이독성 평정 점검표’를 개발하여 교사 2인이 연이어 텍스트의 수준을 평정하도록 하였다. 이때 이독성 평정 점검표는 최소영 외(2021)에서 개발한 국어과 중심의 ‘이독성 질적 평가 준거표’를 간소화하고 국어과 외 교과 텍스트의 특성을 반영하여 개발한 것으로, 평정의 일관성과 편이성 확보를 위해 마련한 것이다. 평정 기준의 제공이 교사의 이독성 평정에 유효한 도움이 된다는 점은 류수경 외(2022)에서 검증된 바 있다.

평정은 초등학교 교사와 중학교 교사가 이독성 평정 점검표에 기반하여 각자의 학교급에 맞게 초등학교 교과서 수록 텍스트, 중학교 교과서 수록 텍스트를 평정하도록 설계했다. 이때 텍스트당 교과 교사 1명(1차 평정)과 국어과 교사 1명(2차 평정)의 2명에게 평정되도록 배치하였다. 즉 1개의 텍스트는 이독성 평정 점검표를 기반으로 1차적으로 해당 교과 교사에게 평정되고 2차적으로 국어과 교사에게 평정되게 함으로써, 텍스트의 교과별 특성 및 글에 대한 분석 전문성의 양 측면이 모두 반영되도록 한 것이다.

초등학교 수록 텍스트 1차 평정에는 국어과 교사 9명, 사회과 교사 3명, 과학과 교사 3명, 2차 평정에는 국어과 교사 20명으로, 총 35명이 참여하였다. 중학교 수록 텍스트 1차 평정에는 국어과 교사 9명, 사회과 교사 6명, 과학과 교사 4명이, 2차 평정에는 국어과 교사 24명이 참여하여 총 43명이 평정에 참여하였다. 상기한 텍스트 이독성 수준 평정 구도를 정리하여 제시하면 <표 1>과 같다.

<표 1> 텍스트 이독성 수준 평정 구도

구분	초등학교 수록 텍스트 평정	중학교 수록 텍스트 평정
공통 기반	이독성 평정 점검표	
1차 평정	국어과 교사: 9명 사회과 교사: 3명 과학과 교사: 3명	국어과 교사: 9명 사회과 교사: 6명 과학과 교사: 4명
2차 평정	국어과 교사: 20명	국어과 교사: 24명
계	35명	43명

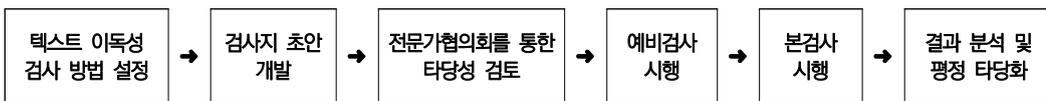
교과 교사와 국어과 교사 2인 체제의 평가 구조를 유지하여 교사 평정을 실시했을 때 평정자 간 신뢰도를 확인하기 위해 초등학교급과 중학교급의 전체 텍스트에 대해 1차 평정 및 2차 평정 간 상관분석을 실시하였다.

<표 2> 교사의 텍스트 이독성 수준 1차 및 2차 평정 간 상관분석 결과

Pearson 상관계수	0.959
유의확률(P-value)	0.000

분석 결과, <표 2>와 같이 Pearson 상관계수가 0.959로 나타났고, 이는 유의수준 0.05에서 통계적으로 유의하였다. 이와 같이 텍스트에 대한 교과 교사의 1차 평정과 국어과 교사의 2차 평정 결과 간 상관계수가 높게 산출됨에 따라 교과 텍스트 평정의 평정자 간 신뢰도가 높음을 알 수 있다.

학생 대상 텍스트 이독성 검사를 통한 교사의 텍스트 이독성 수준 평정 타당성 분석은 다음 [그림 1]과 같은 절차로 진행되었다. 전문가협의회 및 예비검사를 통해 검사 도구를 개발하고, 학생 및 학교 표집, 감독교사 대상의 온라인 워크숍 등을 거쳐 학생 대상 텍스트 이독성 검사를 시행하였다. 세부 채점 기준을 마련하여 학생의 응답 결과에 대한 복수 채점을 진행하였으며, 이후 결과를 분석하는 순으로 진행하였다.



[그림 1] 연구 절차

텍스트 이독성 검사 학년으로는 초등학교 6학년과 중학교 3학년을 선정하였다. 이들은 해당 학교급의 교육과정을 이수하는 최종 학년에 해당하므로, 학교급별 첫 학년에서부터 마지막 학년에 이르기까지의 모든 텍스트를 읽고 이해할 수 있을 것으로 기대되는 학년이기 때문이다. 그리고 층화 군집 표집 방법(stratified cluster sample design)을 활용하였으며, 지역 규모를 층화변인으로 사용하였고, 텍스트 이독성 검사 대상은 학교 단위로 선정하였다.

초등학생과 중학생을 대상으로 실시된 예비검사 및 본검사 시행 개요는 <표 3>과 같다.

<표 3> 예비검사 및 본검사 시행 개요

검사	학교급	참여 학생 수	검사 일정
예비검사	초등학교	136명	2022.05.02. ~ 2022.05.13.
	중학교	150명	2022.05.23. ~ 2022.06.03.
본검사	초등학교	201명	2022.06.20. ~ 2022.07.08.
	중학교	220명	2022.07.04. ~ 2022.07.15.

예비검사는 대도시, 중소도시, 읍면지역에서 각각 2개교씩, 즉 초등학교와 중학교에서 6개교씩 총 12개교를 표집하여 전국의 각 권역별 학교를 고르게 포함하도록 했으며, 초등학교 6학년 136명, 중학교 3학년 150명이 참여하였다. 본검사는 이보다 규모를 확대하여 대도시, 중소도시, 읍면지역에서 각각 3개교씩, 즉 초등학교와 중학교에서 9개교씩 총 18개교를 표집하였고, 표집 학교 내에서는 1개 학급씩 무선 표집하여 초등학교 6학년 201명, 중학교 3학년 220명이 참여하였다.

학교급별 예비검사는 2022년 5~6월, 본검사는 6~7월에 실시되었다. 이때 학교급별 중간고사 및 기말고사 일정과 겹치지 않도록 사전에 제시된 검사 기간 내에서 각 표집 학교에 따라 학사 일정을 고려하여 검사 시행일을 선택하도록 하였으며, 선정된 학급 내의 학생 수 정보를 수합하여 인원수에 맞

취 검사지를 인쇄 및 우편 발송함으로써 지필검사의 형태로 검사를 진행하도록 하였다. 또한 검사 시행 전 감독교사를 대상으로 한 온라인 워크숍을 개최하는 등 검사 사전 준비를 철저히 하였다.

초등학생과 중학생 대상의 검사 시행이 마무리된 후에는 검사지를 모두 수합하고 검사지에 수기된 학생 답안을 엑셀 파일에 입력하여 채점을 효율적으로 할 수 있도록 준비하였다. 채점은 교사 채점단을 구성하여 전문가협의회로 진행하였으며, 유사 답안도 정답으로 인정하는 허용 단어 채점법을 사용하고 각 빈칸 문항에 대해 정답은 1점, 오답은 0점으로 채점하였다. 예비검사 및 본검사 채점 시행의 개요를 제시하면 <표 4>와 같다.

<표 4> 예비검사 및 본검사 채점 개요

검사	학교급	참여 교사 수	채점 일정
예비검사	초등학교	10명(국어 4명, 사회3명, 과학 3명)	2022.05.13.
	중학교	10명(국어 4명, 사회3명, 과학 3명)	2022.06.10.
본검사	초등학교	10명(국어 4명, 사회3명, 과학 3명)	2022.08.11.
	중학교	5명(국어 5명)	2022.08.08.
		6명(사회 3명, 과학 3명)	2022.08.11.

예비검사 결과 채점을 위한 전문가협의회는 2022년 5~6월에 실시되었으며, 초등학교 교사 10명, 중학교 교사 10명이 채점위원으로 참여하였다. 본검사 결과 채점을 위한 전문가협의회는 8월에 실시되었으며, 학생들의 응답 결과 채점을 위해 초등학교 교사 10명, 중학교 교사 11명이 참여하였다.

전문가협의회에서는 채점 기준을 마련하고 복수의 채점위원이 채점을 진행하였다. 초반에는 동일 문항에 대해 2명 이상의 채점위원이 복수 채점을 실시함으로써 마련된 채점 기준이 잘 적용되고 있는지 점검하였으며, 이후 각 채점위원마다 분량을 구분하여 개별적으로 채점을 진행하였다. 마지막 단계에서는 전체 문항에 대해 채점위원 간 채점 결과를 상호 검토하는 과정을 거쳐 최종 채점 결과를 확정하였다.

IV. 초·중학생 대상 텍스트 이독성 검사 도구 개발

1. 검사지 개발을 위한 논의

본 연구에서는 초·중학생 대상 텍스트 이독성 검사지를 개발하기 위해 교과 교육 전공 박사 및 해당 학교급 교사로 구성된 전문가협의회를 통해 대상 텍스트 선정, 빈칸 위치 설정 및 문항 배열 등에 대해 다각적인 논의를 진행하였다. 이때 대상 텍스트는 교과 학습을 위한 교육용 텍스트임을 고려하여 국어, 사회, 과학 교과서 텍스트를 대상으로 하였다. 검사지 개발 과정에서 주요하게 고려한 사항을 정리하면 다음과 같다.

첫째, 검사용 텍스트 선정 시 학년별 위계, 주제의 다양성, 주요 학습 내용, 텍스트 길이(글자 수) 등을 고려하였다. 이중 학년별 위계를 우선적으로 고려하여, 각 교과와 최저 학년에서부터 최고 학년 교과서에 수록된 텍스트를 선정하되, 교사의 평정 결과에 따라 선정하였다. 이에 초등학교 검사용 텍스트는 국어 1~6학년, 사회와 과학 3~6학년 교과서에 수록된 텍스트 중 교사가 2~7학년 수준으로 평정한 텍스트들이 선정되었으며, 중학교 검사용 텍스트는 국어 7~9학년, 사회 7학년 및 9학년, 과학 7~9학년에 수록된 텍스트 중 교사가 7~10학년으로 평정한 텍스트 및 초등학교 검사지와 연계 위해 초등학교 6학년에 수록되고 7학년으로 평정된 텍스트들이 선정되었다. 즉, 교사의 텍스트 수준 평정의 위계가 고르게 드러나도록 텍스트를 선정하였다. 또한 국어, 사회, 과학 교과서에서 다양한 주제들이 선정되도록 노력하였는데, 국어과에서는 유사 주제 중복을 피하기 위해 사회와 과학 소재를 제외한 언어, 음식, 전통문화, 음악 등 인문·문화·예술 관련 텍스트를 선정하였고, 사회과는 지리, 일반사회(정치, 경제, 법, 사회문화), 역사, 과학과는 물리, 화학, 생물, 지구과학 등 하위 영역 분야의 주제들이 고르게 포함되도록 안배하여 특정 영역에 대한 배경지식 변인이 문제해결에 유리하게 작동하지 않도록 하였다. 이뿐만 아니라 성취기준 또는 핵심 개념을 이해하는 데 필수적인 텍스트를 선택하였으며, 시각자료 의존도가 낮은 텍스트를 선정함으로써 텍스트에 집중하여 문제를 해결할 수 있도록 문항을 구성하였다. 또한 텍스트 길이가 이독성 검사 결과에 미치는 영향을 최소화하기 위해 텍스트의 길이 역시 고려하였는데, 초등학교의 경우 300~400자 내외, 중학교는 500~600자 내외로 글자 수를 한정하였고, 가능한 문단 수도 3개의 문단이 넘지 않도록 하였다.

둘째, 각각의 텍스트에서 빈칸의 위치는 교육용 텍스트의 학습 목적 읽기 능력이라는 본 연구의 목적을 반영하여 읽기를 통한 정보 확인과 읽기를 통한 교과 내용 요소 이해를 중심으로 빈칸을 설정하였다. 즉, 글을 읽고 이해하였는지를 확인하기 위해 앞에서 제시된 정보를 되받는 부분, 그리고 텍스트의 주제나 핵심적 학습 내용과 관련이 깊은 중요한 부분을 중심으로 빈칸을 위치시켰다. 특히, 사회과와 과학과 교과서의 글은 개념과 개념에 대한 설명 구조의 글이 주를 이루고 있는데, 본 검사지에서는 배경지식으로 문제를 해결하기보다는 텍스트 전체의 맥락 속에서 빈칸을 작성할 수 있도록 하였다. 이뿐만 아니라 텍스트 내용의 난이도에 맞추어 빈칸의 난이도 위계까지 고려하였는데, 학생들이 텍스트의 전체적인 내용을 파악한 후 텍스트에 포함된 정보를 사실적으로 이해하고 텍스트 내에서 힌트를 얻어 빈칸을 작성하는 것에서부터 맥락에 맞는 어휘를 추론하여 작성할 수 있는 것까지 빈칸의 난이도가 다양할 수 있도록 검사지를 개발하였다.

셋째, 학교급과 학년을 고려하여 텍스트 수와 빈칸 문항 수를 결정하고 검사지 내 제시 순서를 논의하여 배열하였다. 초등학교 검사지의 경우 총 20개 텍스트에 40개 빈칸 문항, 중학교의 경우 총 25개 텍스트에 50개 빈칸 문항으로, 각 텍스트에 2개씩 빈칸 문항을 포함하도록 구성하였다. 이는 각 텍스트별 빈칸 문항이 1개로 제한될 경우에는 빈칸의 어휘 난도가 곧 텍스트의 난도가 될 우려가 있고, 각 텍스트에 빈칸 문항이 많이 포함될 경우 제한된 검사 시간 내에 이독성 수준을 파악할 수 있는 텍스트 수가 감소하게 된다는 점을 고려한 것이다. 검사지에 텍스트를 배열하는 순서는 빈칸 메우기 검사 형식이 익숙하지 않을 학생의 인지적 부담을 고려하여 인지적 부담이 적은 순부터 제시되도록 하였다. 텍스트의 교과서 수록 학년과 교사의 평정 학년을 1차 기준으로 하여 저학년부터 고학년 텍스트 순으로 배치하였으며, 동일한 학년 수준의 텍스트인 경우에는 교과(국어과, 사회과, 과학과 텍스트 순)와

〈표 8〉 중학생 대상 이독성 검사용 텍스트 목록

번호	텍스트 제목/제재	수목 학년	저자	출처
1	〈서울의 굴절〉*	6	집필진	교육부(2020). 초등학교 국어 6-1 나. 미래엔. 222-226.
2	경제 성장	6	집필진	교육부(2019). 초등학교 사회 6-1. 지학사. 104-106.
3	빛의 굴절	6	집필진	교육부(2020). 초등학교 과학 6-1. 천재교과서. 103.
4	〈고추, 김치의 색깔을 바꾸다〉	7	전국지리 교사모임	남미영 외(2020). 중학교 국어 1-2. 교학사. 48-50.
5	〈만화와 포장지도 예술이 되지〉	7	전성수	노미숙 외(2020). 중학교 국어 1-1. 천재교육. 66-68.
6	문화의 세계화	7	집필진	최성길 외(2018). 중학교 사회 1. 비상교육. 84.
7	생물 다양성	7	집필진	김호련 외(2020). 중학교 과학 1. 동아출판. 110-111.
8	기후	7	집필진	박형준 외(2018). 중학교 사회 1. 천재교과서. 20-21.
9	민주 정치의 발전	7	집필진	모경환 외(2018). 중학교 사회 1. 금성출판사. 167.
10	부력	7	집필진	노태희 외(2020). 중학교 과학 1. 천재교과서. 86-89.
11	〈좋은 소음도 있다? 백색 소음 효과〉	8	배명진	신유식 외(2020). 중학교 국어 2-2. 미래엔. 124-125.
12	〈로봇도 권리가 있을까〉	7	한기호	류수열 외(2020). 중학교 국어 1-1. 금성출판사. 53-58.
13	〈게임이 우리에게 주는 혜택〉	8	윤형섭	류수열 외(2020). 중학교 국어 2-2. 금성출판사. 23-24.
14	태양	8	집필진	임태훈 외(2020). 중학교 과학 2. 비상교육. 110-111.
15	물질의 용해	8	집필진	김성진 외(2020). 중학교 과학 2. 미래엔. 208-211.
16	〈관계는 첫인상부터 시작된다〉	7	이철우	김진수 외(2020). 중학교 국어 1-1. 비상교육. 58-60.
17	〈한글의 창제〉	8	이익섭	이은영 외(2020). 중학교 국어 2-2. 동아출판. 102-105.
18	도시 경관	9	집필진	김진수 외(2018). 중학교 사회 2. 미래엔. 146.
19	물가 안정	9	집필진	박형준 외(2018). 중학교 사회 2. 천재교과서. 91.
20	노동권 침해	9	집필진	이민부 외(2018). 중학교 사회 2. 박영사. 21.
21	자석과 전기	9	집필진	김호련 외(2020). 중학교 과학 3. 동아출판. 225-227.
22	성운	9	집필진	김성진 외(2020). 중학교 과학 3. 미래엔. 256-257.
23	세포	9	집필진	노태희 외(2020). 중학교 과학 3. 천재교과서. 188-189.
24	〈인간의 서식지를 예감하다〉	9	김찬호	남미영 외(2020). 중학교 국어 3-1. 교학사. 68-75.
25	사람 세력과 봉당 형성	9	집필진	김형중 외(2020). 중학교 역사 2. 금성출판사. 108-109.

* ‘〈 〉’로 표기된 부분은 텍스트 제목이며, ‘〈 〉’로 표기되지 않은 부분은 텍스트 제재에 해당함.

V. 학생 대상 텍스트 이독성 검사 결과 분석 및 교사의 교육용 텍스트 수준 평정 타당성 분석

Rasch 모형은 문항 난이도를 모수로 하여 피험자 능력을 추정하는 문항반응 모형으로 (Hambleton, Swaminathan & Rogers, 1991: 13-14), 이를 통해 학생의 능력뿐만 아니라 문항간의 난이도 서열을 알 수 있다. 현재 영어권에서 사용되고 있는 대표적인 이독성 지수인 Lexile과 ATOS 개발 과정에서도 Rasch 모형이 사용된바 있으며(Milone, 2014; Smith et al. 1989), 학생 대상 검사에 참여한 응시자 규모를 고려하여 본 연구에서도 Rasch 모형을 사용하여 학생 검사 결과를 분석하고 문항의 난이도 분포를 확인하였다.

학생 대상 텍스트 이독성 검사 결과 분석은 정답과 오답으로 구분된 자료를 바탕으로 하기에 이분(dichotomous) Rasch 모형을 적용하였다. 이때 Rasch 모형은 문항반응이론(Item Response Theory, 이하 IRT)의 한 종류이므로, IRT의 기본 가정인 지역 독립성(local independence) 가정과 일차원성(uni-dimensionality) 가정이 충족되는지 확인하였다. 분석에는 통계 프로그램인 jamovi version 2.2.5와 jMetrik version 4.1.1이 사용되었다.

1. 초등학생 대상 텍스트 이독성 검사 결과

초등학생 대상 텍스트 이독성 검사 결과를 Rasch 모형 분석으로 분석하여 문항 분리 지수(Separation index)와 신뢰도 지수를 산출하였다. 이는 Rasch 모형으로 문항의 난이도를 추정하였을 때 난이도 분포를 확인하고, 다른 집단의 학생이 동일한 문항으로 검사를 했을 때에도 문항 난이도가 일관되게 나타날 수 있는가를 보여주는 것이다.

〈표 9〉 초등학생 대상 텍스트 이독성 검사 문항 분리 지수 및 신뢰도

분리 지수	6.0801
신뢰도	0.9737

〈표 9〉와 같이 초등학생 대상 텍스트 이독성 검사의 분리 지수는 6.0801로, 이는 문항들이 난이도 측면에서 약 6개의 수준으로 구분될 수 있다는 의미이다. 이를 통해 초등학생 대상 텍스트 이독성 검사는 쉬운 문항부터 어려운 문항이 고르게 개발된 것을 확인할 수 있다. 함께 제시된 신뢰도는 Rasch 분석 결과 추정된 문항 난이도의 정확성을 의미하며 0.9737로 높게 나타났다.

초등학생 대상 텍스트 이독성 검사에 대한 Rasch 모형 분석의 결과를 문항별로 살펴보면 〈표 10〉과 같다. 이는 초등학교 6학년을 대상으로 텍스트 이독성 검사를 시행한 결과 나타난 정답률, 난이도, 측정 표준오차, 모형 적합도 등을 제시한 것이다. 모형 적합도를 판단하는 지표로는 Infit Mean Square와 Outfit Mean Square가 있는데, Outfit Mean Square의 경우 ‘이상치(outlying scores)’

의 영향에 상대적으로 더 민감하다는 단점이 있으므로(Fan & Bond, 2019, p.94), 본 연구에서는 Infit Mean Square를 기준으로 모형 적합도를 판단하였다.

〈표 10〉 초등학교 대상 텍스트 이득성 검사 Rasch 분석 결과

텍스트 번호	문항 번호	정답률	난이도(logit)	측정 표준오차	모형 적합도 (infit mean square)
1	1	0.945	-3.2822	0.324	1.051
	2	0.950	-3.3918	0.338	0.923
2	3	0.915	-2.7668	0.267	1.048
	4	0.876	-2.284	0.227	0.886
3	5	0.652	-0.7432	0.162	1.161
	6	0.886	-2.391	0.235	0.873
4	7	0.881	-2.3366	0.231	1.017
	8	0.776	-1.4700	0.183	1.022
5	9	0.587	-0.4151	0.157	1.168
	10	0.886	-2.3910	0.235	1.015
6	11	0.925	-2.9177	0.282	1.039
	12	0.935	-3.0874	0.301	0.887
7	13	0.647	-0.7172	0.161	0.931
	14	0.801	-1.6435	0.19	1.060
8	15	0.746	-1.2777	0.176	0.956
	16	0.761	-1.3719	0.179	1.126
9	17	0.373	0.6275	0.159	0.957
	18	0.945	-3.2822	0.324	0.874
10	19	0.726	-1.1572	0.172	1.243
	20	0.836	-1.9145	0.204	1.044
11	21	0.841	-1.9566	0.206	0.932
	22	0.756	-1.3401	0.178	1.007
12	23	0.512	-0.0529	0.155	0.982
	24	0.687	-0.9302	0.166	0.858
13	25	0.886	-2.391	0.235	0.926
	26	0.388	0.552	0.158	1.131
14	27	0.876	-2.284	0.227	0.944
	28	0.587	-0.4151	0.157	1.032
15	29	0.597	-0.4644	0.157	0.967
	30	0.572	-0.3417	0.156	0.942
16	31	0.716	-1.0988	0.170	0.920
	32	0.552	-0.2448	0.155	0.922

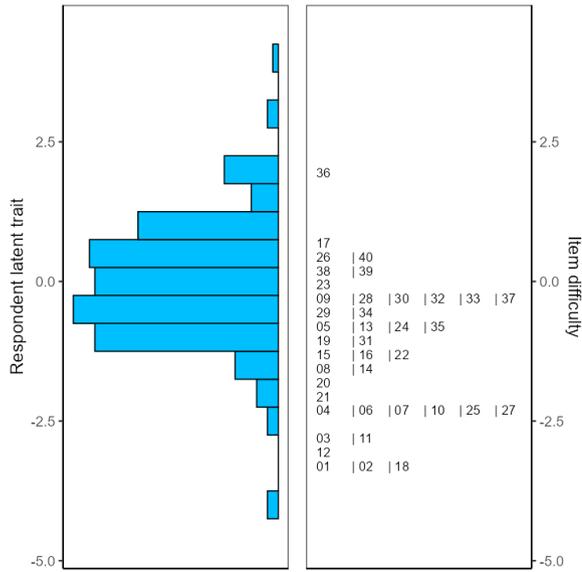
텍스트 번호	문항 번호	정답률	난이도(logit)	측정 표준오차	모형 적합도 (infit mean square)
17	33	0.542	-0.1966	0.155	0.951
	34	0.617	-0.5641	0.159	1.068
18	35	0.662	-0.7957	0.163	0.853
	36	0.159	1.9587	0.205	1.036
19	37	0.572	-0.3417	0.156	1.027
	38	0.438	0.3066	0.156	1.012
20	39	0.473	0.1382	0.155	1.046
	40	0.418	0.4038	0.156	1.069

먼저 정답률을 살펴보면 초등학생 대상 텍스트 이독성 검사에서 가장 정답률이 높은 빈칸은 2번 문항이었으며(국어 '비사치기', 정답률 0.950), 가장 정답률이 낮은 빈칸은 36번 문항이었다(과학 '기후 변화', 정답률 0.159). 전반적으로 정답률은 낮은 것부터 높은 것까지 다양하게 분포하였다. Rasch 모형 분석을 통해 추정된 빈칸 문항의 난이도는 등간 척도(equal interval scale)인 logit으로 표시되는데, 이 숫자가 클수록 어려운 빈칸임을 의미한다. 학생의 문항별 정답률이 높을수록 빈칸 문항의 난이도 logit값이 작고, 정답률이 낮을수록 난이도 logit값이 크게 나타났다.

마지막 열의 infit mean square는 데이터에 대한 Rasch 모델의 적합도(fit)를 나타낸다. infit mean square 값이 0.8~1.2 범위 내에 있으면 모형 적합도가 적절한 것으로 간주되며(Wright & Linacre, 1994), 과대적합(infit mean square < 0.8)인 문항들이 있을 경우 '지역 문항 독립성(Local item independence)' 가정이 위배될 가능성이 있고, 과소적합(infit mean square > 1.2)인 문항들이 있을 경우 '일차원성(unidimensionality)' 가정이 위배될 가능성이 있다(Fan & Bond, 2019).

위의 <표 10>에 제시된 바와 같이, 초등학생 대상 검사에서는 Fan & Bond(2019)의 기준에 따르면 과대적합(infit mean square < 0.8)에 해당하는 문항이 없어 문항반응이론의 적용을 위해 필요한 '지역 문항 독립성' 가정이 위배되지 않았음을 확인할 수 있다. 지역 문항 독립성을 추가적으로 확인하기 위해 잔차의 상관관계를 의미하는 Q_3 (Yen, 1984) 지표를 확인한 결과 문항의 잔차 간 상관계수가 0.3 이상인 값이 존재하지 않아 '지역 문항 독립성' 가정이 위배되지 않았음을 다시 확인할 수 있었다. 과소적합(infit mean square > 1.2)의 경우 총 40문항 중 한 문항(19번)만이 1.243으로 기준을 약간 벗어난 것으로 나타나 '일차원성' 가정에 큰 문제가 없는 것으로 판단되었다.

Rasch 분석을 통하여 초등학생 대상 이독성 검사 문항이 학생 능력의 수준에 맞게 적절하게 평가하고 있는지 판단하였다. 학생들의 읽기 능력과 빈칸 문항의 난이도를 같은 logit이라는 척도를 사용하여 그 분포를 도식화하면 상대적 비교가 용이하고, 학생이 특정 문항의 정답을 맞힐 가능성을 예측할 수 있다. 예를 들어, 학생 능력 로짓값이 문항 난이도 로짓값과 같으면, 학생이 그 문항의 정답을 맞힐 가능성이 50% 정도임을 뜻하며, 학생 능력 logit값이 문항 난이도 logit값보다 높으면 학생은 그 문항의 정답을 맞힐 능력이 충분히 있고, 반대로 학생 능력 logit값이 문항 난이도 logit값보다 낮으면, 학생은 그 문항의 정답을 맞힐 능력이 부족하다고 볼 수 있다(McNamara, 1996).



[그림 2] 초등학교 학생 읽기 능력 및 문항 난이도 분포도

[그림 2]는 초등학교 학생 대상 텍스트 이독성 검사에 참여한 6학년 학생들의 읽기 능력(그림 왼쪽)과 빈칸들의 난이도(그림 오른쪽)를 동일한 척도(logit)에 분포시킨 것이다. 그림 왼쪽의 학생 능력 분포와 오른쪽의 빈칸 문항들의 난이도를 각각 놓고 살펴보면 학생 능력 분포와 문항 난이도 분포가 모두 특정한 수준에 치우침 없이 비교적 넓은 범위로 고르게 분포되어 있음을 알 수 있다. 그런데, 학생 능력과 문항 난이도를 비교해서 살펴보면, 학생 능력치 하단에 대응하는 위치에 분포된 빈칸 문항은 많고, 학생 능력치 상단에 대응하는 위치에 분포된 빈칸 문항은 적다. 즉, 전반적인 학생의 능력에 비해 쉬운 문항이 많고, 어려운 문항이 적었던 것이다. 이와 같은 결과는 초등학교 학생 대상 텍스트 이독성 검사가 2~7학년 수준으로 평정 범위가 넓은 텍스트들로 구성되어서 6학년 학생의 능력 수준보다 쉬운 빈칸, 쉬운 텍스트들이 많이 포함되었기 때문인 것으로 짐작된다.

2. 중학생 대상 텍스트 이독성 검사 결과

중학생 대상 텍스트 이독성 검사 결과를 Rasch 모형 분석으로 분석하여 문항 분리 지수(Separation index)와 신뢰도 지수를 산출하였다.

<표 11> 중학생 대상 텍스트 이독성 검사 문항 분리 지수 및 신뢰도

분리 지수	7.6162
신뢰도	0.9831

〈표 11〉은 중학생 대상 텍스트 이독성 검사의 문항 분리 지수(Separation index)와 신뢰도 지수이다. 분리 지수는 7.6162로, 이는 문항들이 난이도 측면에서 약 7.6개의 수준으로 구분될 수 있다는 것을 의미하며, 중학생 대상 이독성 검사지 역시 쉬운 문항부터 어려운 문항이 고르게 개발된 것을 알 수 있다. 문항 난이도에 대한 신뢰도는 0.9831로 높게 나타났다.

중학생 대상 텍스트 이독성 검사에 대한 Rasch 모형 분석의 결과를 문항별로 살펴보면 〈표 12〉와 같다. 이는 중학교 3학년을 대상으로 텍스트 이독성 검사를 시행한 결과 나타난 정답률, 난이도, 측정 표준오차, 모형 적합도 등을 제시한 것이다.

〈표 12〉 중학생 대상 텍스트 이독성 검사 Rasch 분석 결과

텍스트 번호	문항 번호	정답률	난이도(logit)	측정 표준오차	모형 적합도 (infit mean square)
1	1	0.841	-2.1039	0.207	1.052
	2	0.636	-0.6931	0.159	0.986
2	3	0.818	-1.9012	0.196	0.986
	4	0.586	-0.4216	0.155	1.182
3	5	0.632	-0.6679	0.159	0.956
	6	0.373	0.6895	0.157	1.132
4	7	0.909	-2.8954	0.260	1.030
	8	0.905	-2.8294	0.254	0.944
5	9	0.773	-1.5454	0.182	0.949
	10	0.441	0.3306	0.153	0.919
6	11	0.655	-0.7955	0.161	1.033
	12	0.827	-1.9799	0.200	0.920
7	13	0.527	-0.1131	0.153	1.423
	14	0.500	0.0271	0.153	1.024
8	15	0.841	-2.1039	0.207	1.128
	16	0.714	-1.1480	0.169	1.157
9	17	0.609	-0.5434	0.157	1.082
	18	0.918	-3.0365	0.272	0.796
10	19	0.923	-3.1122	0.279	0.723
	20	0.850	-2.1913	0.211	0.834
11	21	0.509	-0.0196	0.153	0.861
	22	0.527	-0.1131	0.153	1.020
12	23	0.595	-0.4701	0.156	0.917
	24	0.409	0.4962	0.154	1.082
13	25	0.477	0.1436	0.153	1.225
	26	0.682	-0.9539	0.164	1.046

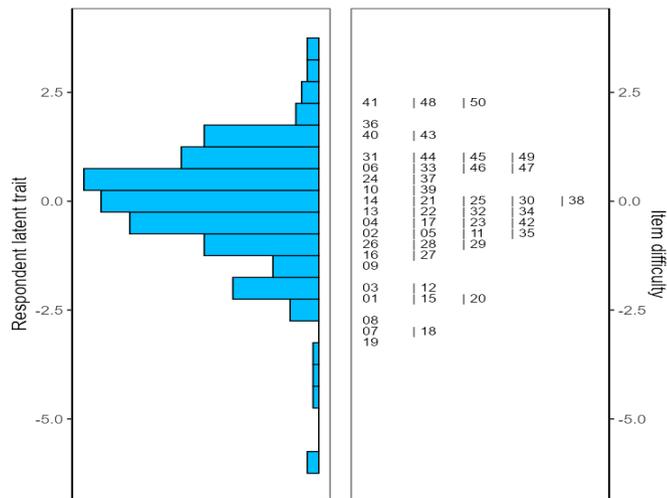
텍스트 번호	문항 번호	정답률	난이도(logit)	측정 표준오차	모형 적합도 (infit mean square)
14	27	0.732	-1.2644	0.172	0.830
	28	0.691	-1.0082	0.165	0.799
15	29	0.668	-0.8739	0.162	1.077
	30	0.482	0.1203	0.153	0.916
16	31	0.327	0.9406	0.161	0.840
	32	0.573	-0.3495	0.155	0.916
17	33	0.359	0.7635	0.158	0.896
	34	0.532	-0.1365	0.153	0.936
18	35	0.655	-0.7955	0.161	1.130
	36	0.195	1.7922	0.186	0.965
19	37	0.405	0.5201	0.155	1.251
	38	0.477	0.1436	0.153	0.984
20	39	0.441	0.3306	0.153	0.975
	40	0.236	1.4989	0.175	0.989
21	41	0.145	2.2149	0.207	1.170
	42	0.618	-0.5929	0.158	0.855
22	43	0.241	1.4684	0.174	0.996
	44	0.327	0.9406	0.161	0.965
23	45	0.327	0.9406	0.161	0.993
	46	0.373	0.6895	0.157	0.840
24	47	0.377	0.665	0.156	0.892
	48	0.145	2.2149	0.207	1.084
25	49	0.323	0.9664	0.161	0.903
	50	0.136	2.3029	0.212	0.988

먼저 정답률을 살펴보면 중학생 대상 텍스트 이독성 검사에서 정답률이 가장 높은 빈칸은 19번(과학 '부력', 정답률 0.923), 가장 낮은 빈칸은 50번 문항으로(사회 '사립 세력과 봉당 형성', 정답률 0.136), 전반적으로 높은 것과 낮은 것이 다양하게 분포되었다. 학생의 문항별 정답률이 높을수록 Rasch 모형 분석을 통해 추정된 빈칸 문항의 난이도 logit의 숫자가 작고, 정답률이 낮을수록 난이도 logit 숫자가 크게 나타났다.

Rasch 모델의 적합도(fit) 측면에서 infit mean square 값이 0.8~1.2 범위에 있을 때 적절한 모형 적합도로 판단하였으며(Wright & Linacre, 1994), 과대적합(infit mean square < 0.8)인 문항이 있는 경우 '지역 문항 독립성(Local item independence)' 가정이 위배될 가능성이 있고, 과소적합(infit mean square > 1.2)인 문항이 있는 경우 '일차원성(unidimensionality)' 가정이 위배될 가능성이 있다(Fan & Bond, 2019). 분석 결과, 과대적합에 해당하는 문항이 50개의 빈칸 문항 중 3개, 과소적합에 해당하는 문항이 3개 있었으나, 모두 기준을 근소하게 벗어나는 수치로서 '지역 독립성' 및

‘일차원성’과 같은 Rasch 모형 적용을 위한 기본 가정이 충족되는 것으로 판단하였다. 과대적합 문항인 18, 19, 28번 문항의 Infit Mean Square 값은 각각 0.796, 0.723, 0.799로 기준인 0.8과 근소한 차이를 보이며, 과소적합 문항인 13, 25, 37번 문항의 Infit Mean Square 값은 각각 1.423, 1.225, 1.251로 기준 1.2에서 약간 벗어나는 수치를 보였다. 이와 같은 결과에 따라 지역 문항 독립성을 추가적으로 확인하기 위해 잔차의 상관관계를 의미하는 Q_3 (Yen, 1984) 지표를 확인한 결과, 모든 문항의 잔차 간 상관계수(Q_3)가 0.3 미만으로 나타나 ‘지역 문항 독립성’ 가정이 위배되지 않았음을 알 수 있었다.

Rasch 분석을 통하여 중학생 대상 텍스트 이독성 검사 문항이 학생 능력의 수준에 맞게 적절하게 평가하고 있는지 판단하였다. 중학생 대상 텍스트 이독성 검사에 참여한 중학교 3학년 학생들의 읽기 능력과 빈칸 문항의 난이도를 logit 척도를 사용하여 분포를 도식화한 것이다.



[그림 3] 중학생 읽기 능력 및 문항 난이도 분포도

[그림 3]은 중학생 대상 텍스트 이독성 검사에 참여한 중학교 3학년 학생들의 읽기 능력(그림 왼쪽)과 빈칸들의 난이도(그림 오른쪽)를 동일한 척도(logit)에 분포시킨 것이다. 그림 왼쪽의 학생 능력 분포와 오른쪽의 빈칸 문항들의 난이도를 각각 놓고 살펴보면 학생 능력 분포가 문항 난이도 분포보다 더 넓게 퍼져 있음을 알 수 있다. 특히 학생 능력치 하단에 대응하는 위치에 분포된 빈칸 문항이 나타나지 않았는데, 이를 통해 읽기 능력이 하위에 해당하는 학생들에게 이 검사가 다소 어려웠다는 점을 알 수 있다. 이와 같은 결과는 중학생 대상 텍스트 이독성 검사가 7~10학년 수준으로 평정된 비교적 평점 범위가 좁은 텍스트들로 구성되어서 중학교 3학년에서 읽기 능력이 하위에 해당하는 학생이 읽고 정답을 맞히기에 쉬운 텍스트가 없었던 것으로 짐작된다.

3. 교사의 교육용 텍스트 수준 평정 타당성 분석

Rasch 모형 분석을 통해 문항의 난이도 분포를 살펴본 결과 초등학교와 중학교 모두 해당 학년 학생이 정답을 맞히기 쉬운 문항에서 어려운 문항이 고르게 분포하고 있음을 알 수 있었다. 이를 바탕으로 Spearman's Rho를 활용한 서열 상관관계 분석을 통해 교사의 수준 평정 서열과 학생 대상 텍스트 이독성 검사 결과에서 나타난 텍스트의 난이도 서열 간의 관계를 분석하였다. 이때, 교사의 수준 평정 서열은 최소영 외(2021)에서 개발한 이독성 평가 준거표를 간소화하여 교사 2인이 평정한 값의 평균으로 산출하였고, 텍스트의 난이도 서열은 텍스트 1개에 빈칸 문항이 2개이므로 2개 빈칸 문항 난이도 추정치(logit) 평균으로 산출하였다.

서열 상관관계 분석을 통해 산출되는 상관계수가 높을 경우, 교사의 평정을 통해 수준을 서열화한 결과와 실제 학생 대상 텍스트 이독성 검사를 통해 산출된 텍스트의 난이도 수준 서열이 유사하다는 결론을 내릴 수 있다. 이와 같이 서열 상관관계 분석을 통한 이독성 지수 타당성 분석은 대표적인 영어권 이독성 지수 6종의 텍스트 난이도 예측력을 검증한 Nelson et al.(2012) 연구에도 적용된 바 있다.

초등학생용 검사지에 포함된 총 20개의 텍스트에 대해 교사가 평정한 수준과 학생 검사 결과 추정된 텍스트 난이도를 산출하고, Spearman's Rho를 통해 교사 평정 서열과 학생 텍스트 난이도 서열 간의 상관관계를 분석하였다.

〈표 13〉 교사의 텍스트 수준 평정과 초등학생 대상 이독성 검사의 텍스트 난이도 비교

텍스트 번호	교과	텍스트 제목/제재	교사 평정 수준(학년)	학생 검사 난이도(logit)
1	국어	〈비사치기〉	2	-3.337
2	국어	〈좋은 습관을 기르자〉	3	-2.525
3	사회	교통수단 발달	3.5	-1.567
4	과학	자석	3	-1.903
5	과학	지구	3.5	-1.403
6	국어	〈갯벌을 보존해야 하는 까닭〉	3.5	-3.003
7	국어	〈전통 음식의 우수성〉	6.5	-1.180
8	사회	지도 응용 프로그램	4	-1.325
9	과학	〈탄산수 거품의 최고 높이 예상하기〉	4	-1.327
10	과학	폐수	4.5	-1.536
11	국어	〈한지들이〉	5.5	-1.648
12	국어	〈한글이 위대한 이유〉	5	-0.492
13	국어	〈자연을 닮은 우리 악기〉	5	-0.920
14	사회	기온	5.5	-1.350
15	사회	고려 건국	5	-0.403
16	국어	〈서울의 궁궐〉	7	-0.672
17	사회	국회	6	-0.380

텍스트 번호	교과	텍스트 제목/제재	교사 평정 수준(학년)	학생 검사 난이도(logit)
18	과학	기후 변화	6.5	0.582
19	사회	경제 성장	7	-0.018
20	과학	빛 굴절	7	0.271

〈표 14〉 교사의 텍스트 수준 평정과 초등학생 대상 이독성 검사의 텍스트 난이도 간의 서열 상관관계

Spearman's Rho	0.8154
유의확률 (P-value)	0.000

그 결과 〈표 14〉에 제시된 바와 같이, 교사들이 평정한 텍스트 수준과 학생 검사 결과 산출된 텍스트 난이도 간의 서열 상관관계(Spearman's Rho)는 0.8154로 높게 나타났으며, 이는 $P < 0.001$ 수준에서 통계적으로 유의하였다. 따라서, 본 연구에서 적용한 교사의 초등학교급 텍스트 수준 평정의 타당성을 확인할 수 있었다.

다음으로 중학생용 검사지에 포함된 총 25개의 텍스트에 대한 교사 평정 수준과 학생 검사 결과 추정된 텍스트 난이도를 산출하고, Spearman's Rho를 통해 교사 평정 서열과 학생 텍스트 난이도 서열 간의 상관관계를 분석하였다.

〈표 15〉 교사의 텍스트 수준 평정과 중학생 대상 이독성 검사의 텍스트 난이도 비교

텍스트 번호	교과	제재	교사 평정 난이도(학년)	학생 검사 난이도(logit)
1	국어	〈서울의 궁궐〉	7	-1.399
2	사회	경제 성장	7	-1.161
3	과학	빛의 굴절	7	-0.011
4	국어	〈고추, 김치의 색깔을 바꾸다〉	7	-2.862
5	국어	〈만화와 포장지도 예술이 되지〉	7	-0.607
6	사회	문화의 세계화	7	-1.388
7	과학	생물 다양성	7	-0.043
8	사회	기후	7.5	-1.626
9	사회	민주 정치의 발전	7.5	-1.790
10	과학	부력	7.5	-2.652
11	국어	〈좋은 소음도 있다? 백색 소음 효과〉	8	-0.066
12	국어	〈로봇도 권리가 있을까〉	8	0.013
13	국어	〈게임이 우리에게 주는 혜택〉	8	-0.405
14	과학	태양	8	-1.136
15	과학	물질의 용해	8	-0.377
16	국어	〈관계는 첫인상부터 시작된다〉	8.5	0.296

텍스트 번호	교과	제재	교사 평정 난이도(학년)	학생 검사 난이도(logit)
17	국어	〈한글의 창제〉	9	0.314
18	사회	도시 경관	9	0.498
19	사회	물가 안정	9	0.332
20	사회	노동권 침해	9	0.915
21	과학	자석과 전기	9	0.811
22	과학	성운	9	1.205
23	과학	세포	9.5	0.815
24	국어	〈인간의 서식지를 예감하다〉	10	1.440
25	사회	사람 세력과 봉당 형성	10	1.635

〈표 16〉 교사의 텍스트 수준 평정과 중학생 대상 이독성 검사의 텍스트 난이도 간의 서열 상관관계

Spearman's Rho	0.8374
유의확률 (P-value)	0.000

그 결과 〈표 16〉에 제시된 바와 같이, 교사들이 평정한 텍스트 수준과 학생 검사 결과 산출된 텍스트 난이도 간의 서열 상관관계는 0.8374로 높게 나타났으며, 이는 $P < 0.001$ 수준에서 통계적으로 유의하였다. 따라서 중학생을 대상으로 시행한 텍스트 이독성 검사를 통하여 본 연구에서 적용한 교사의 중학교급 텍스트 수준 평정의 타당성을 확인할 수 있었다.

VI. 결론 및 제언

본 연구는 교사가 교육용 텍스트의 이독성 정도를 학년 수준으로 평정한 텍스트를 활용하여 학생 대상 텍스트 이독성 검사를 실시하고, 교사의 학년 수준 평정 서열과 학생 대상 이독성 검사 문항 난이도 서열 간의 상관관계 분석을 통해 교사의 텍스트 이독성 평정의 타당성을 분석하였다.

연구 내용을 요약하면 다음과 같다. 학생 대상 텍스트 이독성 검사는 초등학교 6학년과 중학교 3학년을 대상으로 빈칸 메우기 검사로 시행되었다. 텍스트는 교사의 평정 결과에 따른 수준별 텍스트를 다양한 주제 분야를 고려하여 선정하였으며, 각 텍스트에서 빈칸의 위치는 배경지식으로 문제를 해결하기보다는 문맥 단서를 통해 추론할 수 있는 곳에 배치하였다. Rasch 모형 분석으로 문항의 난이도 분포를 살펴본 결과 초등학생 대상 검사와 중학생 대상 검사 모두 쉬운 문항과 어려운 문항이 고르게 분포하고 있었다. 교사의 수준 평정 서열과 학생 대상 이독성 검사의 문항 난이도 서열 간의 상관관계를 분석한 결과 초등학생 대상 검사 결과는 0.8154, 중학생 대상 검사 결과는 0.8374로 높게 나타나 교사의 텍스트 이독성 수준 평정의 타당성을 확인하였다.

이상의 연구 결과를 바탕으로 향후 후속 연구에서는 학생 평가 실시 방법을 다각적으로 모색해 볼 필요가 있다. 본 연구에서는 서술형 빈칸 메우기 검사를 통해 학생 대상 텍스트 이독성 검사를 실시하였다. 이는 빈칸 메우기 검사의 전통적인 방법으로 이독성 평가 도구 개발 연구에서 학생 평가를 연계할 때 자주 사용된 방법이지만, 제한된 평가 시간 안에 다룰 수 있는 텍스트 수가 한정되고 채점에 드는 시간과 비용이 크기 때문에 대규모 인원을 평가하기 어렵다. 학생들이 느끼는 쉽고 어려운 정도를 반영한 이독성 지수를 개발하기 위해서는 선다형 빈칸 메우기 검사, 일반 독해 평가 문항, 리커트 척도를 활용한 난도 평정 등 다양한 학생 평가 방법을 개발할 필요가 있다.

또한 학생들의 독해 평가를 통해 학생 독해 점수를 예측하는 이독성 모형 개발을 고려할 수 있다. Lexile이나 ATOS와 같은 2세대 영어 이독성 평가 시스템은 학생들의 독해 점수에 따라 적절한 수준의 텍스트를 제시해 준다. 이것이 가능한 이유는 방대한 읽기 평가 데이터베이스를 바탕으로 개발되었기 때문이다. 대규모 읽기 평가 데이터베이스가 있다면, 교사가 평정한 학년 수준을 예측하는 모형과 별개로 학생의 독해 점수를 예측하는 모형을 추가로 개발할 수 있다. 학생의 독해 점수를 예측하는 모형이 개발되면 새로운 텍스트에 대한 독해 평가 점수를 바탕으로 학생에게 좀더 쉬운 텍스트, 혹은 좀더 어려운 텍스트를 맞춤형으로 제시해 줄 수 있을 것이다. 학생의 독해 능력과 텍스트 수준의 직접적 연계를 통한 맞춤형 교육의 다각화를 위해, 본 연구를 토대로 후속 연구가 지속될 필요가 있다.

참고문헌

- 김지은(2012). 읽기 교과서의 설명 텍스트 난도에 대한 교사와 학생의 인식 차이에 관한 연구. **한국초등국어교육**, 49, 33-64.
- 김해인(2019). 중학생들의 설명텍스트 난도 평가. **문화와융합**, 41(4), 1255-1278.
- 류수경, 길호현, 박태준, 최소영(2022). 교사의 이독성 평가 전문성 신장을 위한 평가자 훈련 워크숍 사례연구. **독서연구**, 62, 141-185.
- 박수자(2001). 읽기 지도의 이해. 서울: 서울대학교출판부
- 박태준, 박은아, 류수경, 한정아, 최소영, 변태진(2022). **KICE 이독성 지수 (KICE Readability Index) 자동측정 프로그램 설계 및 개발 연구(I)**. 한국교육과정평가원.
- 이성영(2011). 읽기(독서)에서의 교육 내용 위계화: 초등 교과서의 이독성 비교 연구 -국어, 사회, 과학 교과서를 중심으로. **국어교육학연구**, 41, 169-193.
- 조용구, 이경남(2013). 읽기 능력 측정을 위한 빈칸 메우기 검사 도구의 개발. **독서연구**, 30, 465-491.
- 조용구(2016). 독자에게 적합한 글을 선정하기 위한 이독성 공식의 개발. 박사학위논문, 한국교원대학교.
- 조현재, 김영수(2000). 생물 교과서의 이독성(易讀性) 측정 도구로서 Cloze 검사의 적용 방법에 대한 연구. **생물교육**, 28(4), 348-355.
- 최소영, 길호현, 박태준, 류수경(2022). 교육용 텍스트의 이독성 질적 평가 준거 개발 및 타당화 연구. **국어교육**, 176, 1-57.
- 최소영, 박태준(2022). 텍스트 이독성 질적 평가 요인의 설명력 분석 - 교육용 정보 텍스트를 중심으로 -. **국어교육**, 179, 251-294.
- 최소영, 박태준, 류수경, 장지혜, 길호현, 이인화(2021). **교육용 텍스트의 이독성 평가 준거 개발 연구**. 한국교육과정평가원.
- Bormuth, J. R. (1971). *Development of Standards of Readability : Toward a Rational Criterion of Passage Performance*. Final Report.
- Chall, J. S. & Dale, E. (1995). *Readability revisited: the new Dale & Chall readability formula*. Cambridge, Massachusetts: Brookline Books.
- Dale, E., & Chall, J. S. (1948). A formula for predicting readability: Instructions. *Educational research bulletin*, 27(2), 37-54.
- Fan, J., & Bond, T. (2019). *Applying Rasch measurement in language assessment: Unidimensionality and local independence*. *Quantitative Data Analysis for Language As*

- essment Volume I*. London, England : Routledge.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc.
- Lively, B. A., & Pressey, S. L. (1923). A method for measuring the vocabulary burden of textbooks. *Educational administration and supervision*, 9(7), 389-398.
- McNamara, T. F. (1996). *Measuring second language performance*. New York: Longman Publishing Group.
- Mesmer, H. A. E. (2008). *Tools for matching readers to texts: Research-based practices*. New York: Guilford Press.
- Milone, M. (2014). *Development of the ATOS readability formula*. Renaissance Learning.
- Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*. Council of Chief State School Officers, Washington, DC.
- Smith, D., Stenner, A. J., Horabin, I., & Smith, M. (1989). *The Lexile scale in theory and practice: Final report for NIH Grant HD-19448*. Washington, Washington DC: MetaMetrics, Inc.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370. Retrieved from <http://www.rasch.org/rmt/rmt83b.htm>. (검색일: 2023. 9. 30.)
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125-145.

[교과서 목록]

- 교육부. (2020). **초등학교 과학 3-1**. 서울: (주)천재교과서.
- 교육부. (2020). **초등학교 과학 4-1**. 서울: (주)천재교과서.
- 교육부. (2020). **초등학교 과학 5-2**. 서울: (주)천재교과서.
- 교육부. (2020). **초등학교 과학 6-1**. 서울: (주)천재교과서.
- 교육부. (2020). **초등학교 국어 1-2나**. 서울: (주)미래엔.
- 교육부. (2020). **초등학교 국어 3-1나**. 서울: (주)미래엔.
- 교육부. (2020). **초등학교 국어 3-2가**. 서울: (주)미래엔.
- 교육부. (2020). **초등학교 국어 4-1나**. 서울: (주)미래엔.

- 교육부. (2020). **초등학교 국어 5-1나**. 서울: ㈜미래엔.
- 교육부. (2020). **초등학교 국어 5-2나**. 서울: ㈜미래엔.
- 교육부. (2020). **초등학교 국어 6-1가**. 서울: ㈜미래엔.
- 교육부. (2020). **초등학교 국어 6-1나**. 서울: ㈜미래엔.
- 교육부. (2019). **초등학교 사회 3-1**. 서울: 지학사.
- 교육부. (2019). **초등학교 사회 4-1**. 서울: 지학사.
- 교육부. (2019). **초등학교 사회 5-1**. 서울: 지학사.
- 교육부. (2019). **초등학교 사회 5-2**. 서울: 지학사.
- 교육부. (2019). **초등학교 사회 6-1**. 서울: 지학사.
- 김성진, 조용근, 최미화, 김홍석, 장철한, 김혜경, 권효식, 오현선, 구향모, 강희정, 김대준, 이진우, 류형근, 문무현, 이유진, 유명익. (2020). **중학교 과학 2**. 서울: ㈜미래엔.
- 김성진, 조용근, 최미화, 김홍석, 장철한, 김혜경, 권효식, 오현선, 구향모, 강희정, 김대준, 이진우, 류형근, 문무현, 이유진, 유명익. (2020). **중학교 과학 3**. 서울: ㈜미래엔.
- 김호련, 김상협, 김민성, 이연숙, 황성용, 이성하, 노동규, 임혁, 배미정, 이태원, 권오성, 박권태, 송성재. (2020). **중학교 과학 3**. 서울: 동아출판.
- 김호련, 김상협, 김민성, 이연숙, 황성용, 이성하, 노동규, 임혁, 배미정, 이태원, 권오성, 박권태, 송성재. (2020). **중학교 과학 1**. 서울: 동아출판.
- 노태희, 이봉우, 김선경, 장종목, 강석진, 임희연, 양찬호, 박재근, 민진선, 배영혜, 오필석, 김연귀, 박창용. (2020). **중학교 과학 1**. 서울: 천재교과서.
- 노태희, 이봉우, 박재근, 김선경, 민진선, 장종목, 배영혜, 강석진, 오필석, 임희연, 김연귀, 양찬호, 박창용. (2020). **중학교 과학 3**. 서울: 천재교과서.
- 임태훈, 백종민, 남경운, 강태욱, 강대훈, 이복영, 장효순, 황인신, 김미경, 이용철, 고현덕, 신미영. (2020). **중학교 과학 2**. 서울: 비상교육.
- 김진수, 이경옥, 박채형, 박현숙, 박수진, 강양희, 박상희, 정대승, 김봉규, 오경준, 강송연, 이경숙. (2020). **중학교 국어 1-1**. 서울: 비상교육.
- 남미영, 허철구, 공규택, 김아진, 류상하, 박경하, 백덕현, 양연규, 여의기, 이푸르니, 임성규, 임영규, 전혜옥, 정태기, 정형근, 황재진. (2020). **중학교 국어 1-2**. 서울: 교학사.
- 남미영, 허철구, 공규택, 박미영, 백덕현, 양연규, 임성규, 정태기, 정형근, 최주옥, 황재진. (2020). **중학교 국어 3-1**. 서울: 교학사.
- 노미숙, 주진택, 안수진, 김호태, 신해연, 황희중, 소정섭, 장선영, 강영미, 류신형, 노수경. (2020). **중학교 국어 1-1**. 서울: 천재교육.
- 류수열, 전명재, 이동민, 최동진, 정지영, 강호정, 나미나, 박인규, 박수현, 김민선. (2020). **중학교**

- 국어 2-2.** 서울: 금성출판사.
- 류수열, 전명재, 이동민, 최동진, 정지영, 강호정, 박용진, 나미나, 박인규, 박수현. (2020). **중학교 국어 1-1.** 서울: 금성출판사.
- 신유식, 정미선, 이필규, 김영찬, 전경원, 윤인희, 박선주, 김정희, 이은화, 나단비. (2020). **중학교 국어 2-2.** 서울: ㈜미래엔.
- 이은영, 박순영, 박정호, 권가용, 박성희, 이재찬, 허단비. (2020). **중학교 국어 2-2.** 서울: 동아출판.
- 김진수, 문대영, 조성호, 김숙, 문승규, 이강준, 최영아, 이희원, 손영찬, 오두환, 김신정, 이은주, 박진민, 양설, 이상급. (2018). **중학교 사회 2.** 서울: ㈜미래엔.
- 김형종, 장문석, 박범희, 고재연, 고진아, 김현성, 우정애, 이대희, 우현진, 맹수용. (2020). **중학교 역사 2.** 서울: 금성출판사.
- 모경환, 이윤호, 강대현, 김현경, 이수화, 황미영, 조철기, 승현아, 김영일, 서정현, 윤민주, 나유진. (2018). **중학교 사회 1.** 서울: 금성출판사.
- 박형준, 신정엽, 이봉민, 서현진, 김현철, 박서연, 이정식, 김봉수, 조영매, 이해란, 고인석, 신정아, 김찬미. (2018). **중학교 사회 1.** 서울: 천재교과서.
- 박형준, 신정엽, 이봉민, 서현진, 김현철, 박서연, 이정식, 김봉수, 조영매, 이해란, 고인석, 신정아, 김찬미. (2018). **중학교 사회 2.** 서울: 천재교과서.
- 이민부, 조영달, 김왕근, 김기남, 김도영, 김태환, 박세구, 박찬선, 박철용, 이병인, 정명섭, 최종현. (2018). **중학교 사회 2.** 서울: ㈜박영사.
- 최성길, 최원희, 강창숙, 박상준, 최병천, 조일현, 권태덕, 이수영, 조철민, 조성백, 김상희, 강봉균, 정민정, 김연주. (2018). **중학교 사회 1.** 서울: 비상교육.

· 논문접수 : 2023.10.05. / 수정본접수 : 2023.10.27. / 게재승인 : 2023.11.13.

ABSTRACT

**A Study on the Validation of Teacher's Educational Text
readability Evaluation through a Cloze Test for Elementary
and Middle School Students**

Ryu, Sukyeong

Associate Research Fellow, Korea Institute for Curriculum and Evaluation

Park, Taejoon

Research Fellow, Korea Institute for Curriculum and Evaluation

Park, Eunah

Senior Research Fellow, Korea Institute for Curriculum and Evaluation

Byun, Taejin

Assistant Professor, Gwangju National University of Education

Han, Junga

Researcher, Korea Institute for Curriculum and Evaluation

Ha, Myeungjin

Researcher, Korea Institute for Curriculum and Evaluation

Kim, Jihye

Researcher, Gyeonggi Institute of Education

Choi, Soyoung

Assistant Professor, Pusan National University

The goal was to determine whether the teacher's evaluation of text readability aligns with the students' perspective. Two sets of cloze tests were created and administered to measure the agreement between teacher-evaluated text readability and the difficulty derived from students' test results. A total of 286 students for field trial and 421 students for main survey those who are 6th grade elementary school and 3rd grade middle school participated in the tests.

Rasch model analysis was used to examine the distribution of text difficulty, which was found to be evenly distributed. Spearman's Rho was employed to analyze the correlation between the teacher's evaluation sequence and the text difficulty sequence in the students' tests. The results indicated a high correlation of 0.8154 for elementary school students and 0.8374 for middle school students, confirming the validity of the text readability test and evaluation conducted in this study.

Key Words: readability index, cloze test, teacher-evaluated text readability, correlation analysis, Rasch model analysis