

독서지수 시스템 개발의 과제와 해결 방안에 대한 사례 연구

— 독서능력 검사와 독서지수의 연계에 관한 온독지수
시스템의 사례를 중심으로

차경미 한국교육과정평가원 기획조정본부 부연구위원(제1저자)

최소영 부산대학교 국어교육과 조교수(공동저자)

이상재 서울대학교 국어교육과 박사수료(공동저자)

인정남 충청남도 교육청 교감(공동저자)

신명선 인하대학교 국어교육과 교수(교신저자)

- I. 서론
- II. 독서지수 시스템 개발의 전제와 과제
- III. 독서능력 검사 설계와 MST용 문항 확보
- IV. MST 모델 설계 및 독서지수와의 연계
- V. 결론

I. 서론

최근 읽기 부진만큼이나 ‘비독서(aliteracy)’ 현상이 사회적으로 주목받고 있다. 비독서란 글자를 읽을 수 있음에도 글을 읽지 않는 것으로서(김혜정·허모아, 2021: 10) 비독서 학생은 수업 시간 중 교사에 의해 강요될 때 등 외적 필요에 의해서는 글을 읽지만 그 외에 스스로의 내적 동기에 의해서는 글을 읽지 않는다. 이러한 비독서 현상을 타개하기 위한 하나의 방법으로 수준별 독서 지원 시스템이 제안된다.

수준별 독서 지원 시스템은 교육 현장 기반적인 개념으로서, 실질적, 실제적 연구 주제이다.¹⁾ 수준별 독서 지원 시스템 구축을 위해 통상적으로 제안되는 독서지수(reading index)라는 개념 역시 그러하다. 그 개념이나 위상 등에 대한 이론적 논의만큼, 실제로 이러한 것들을 어떻게 ‘개발’할 것이냐

1) 흔히 실천적 연구(practical research), 현장 중심 연구(field-based research), 또는 실행 연구(action research) 등으로 불리는 교육학의 논의들은 실제성, 맥락성, 참여성 등을 중시한다. 대개 맥락적 접근(context-specificity)을 취하기 때문에 일반화보다는 특정 조건에서의 유의미한 개선을 추구한다.

도 중요하다. 그럼에도 현 국어교육 상황에서 독서지수 개발에 대한 실제적, 수행적 차원의 논의는 부족한 것이 사실이다.²⁾

본고는 독서지수 시스템 개발 과정에서 부딪힐 수 있는 핵심적인 문제들을 어떤 방식으로 해결할 수 있는가에 대한 실제적, 수행적 차원의 연구로서, 사례 연구(case study)에 가깝다. 사례 연구는 연구 대상에 대한 구체적이고 심층적인 정보 제공을 통해 대상에 대한 깊이 있는 이해를 돕는 연구 방법이다.

독서지수 시스템 개발 과정은 독서지수 개발의 원리, 방법, 목적 등에 따라 다양할 수 있다. 이 때문에 영어권의 렉사일(Lexile) 지수와 AR 지수의 체계가 다르고 국내의 LQ 지수나 EBS 및 한국교육과정평가원(KICE)의 논의도 조금씩 다르다. 따라서 독서지수의 개념과 위상 등에 관한 이론적 논의가 아니라 ‘개발’에 관한 실제적 논의라면, 구체적인 사례를 중심으로 논의하는 것이 개발 원리나 방법 등에 대한 이해를 심화하는 데 더 크게 기여하는 경우가 많다.³⁾

-
- 2) 이독성, 텍스트 복잡도, 독서지수 등에 관한 이론적 논의들은 최근 활발하게 이루어지고 있지만 개발에 관한 실질적 논의는 부족한 것이 사실이다. 현재 몇몇 독서지수가 개발된 것으로 알려져 있으나 개발 과정이나 개발 관련 난제와 해결 방안 등을 상세히 다룬 논의는 드물다. 대교 출판사에서 개발한 KReaD 지수의 경우, 조용구·이경남(2020)으로 보고 되었으나 상당히 많은 정보들이 비공개 상태라서 사실상 개발 과정을 알기 어렵다. EBS에서도 독서지수를 개발한 것으로 알려져 있으나 개발 과정 등은 공개되어 있지 않다. 다만 EBS 독서지수 개발의 기반이 된 초창기 연구의 일단을 서혁·이소라·류수경·오은하·윤희성·변경가 외(2013)를 통해 짐작할 수 있을 뿐이다. (주)날말의 LQ 지수의 경우 현재 상업적으로 서비스 중이나 개발 과정을 직접적으로 다룬 논문은 없다. 한국교육과정평가원에서 추진한 연구의 경우, 박태준·김현정·오상철·홍미영·배화순·최소영(2024)으로 마무리된 3개년 연속 과제를 통해 모델 개발 과정이 연구보고서에 상세히 공개되었으나 최종 성과물은 미공개 중이라서 구체적인 성과를 파악하기 어렵다. 충남교육청에서 개발한 온독지수의 경우에는 신명선·이기연·나상수·인정남·차경미(2024)를 통해 그 개발 방법이 비교적 상세히 공개되어 있다.
- 3) 구체적인 실제의 개선을 추구하는 교육학에서 사례 연구가 갖는 의의는 매우 크다. 교육 맥락은 늘 변화하기 때문에 실제의 개선을 원한다면 논의의 시작은 구체적인 현실 맥락과

독서지수 시스템 개발 논의의 경우 실질적인 작업 과정을 통해 여러 과제와 해결 방안이 탐구되는 것이 유리함에도 그간 이와 관련된 연구가 충분히 이루어지지 어려웠던 것은 이러한 연구가 갖는 어려움 때문이다. 개발 논의는 실제로 도서에 지수를 부여하고 문항을 개발하고 학생들의 독서능력을 측정하는 등 지난한 작업 과정을 수행해야 가능하다.

독서지수 시스템 개발 결과는 상업적 이익으로 연결될 수 있기 때문에 대개 그 개발 과정을 있는 그대로 공개하지 않는다는 현실적인 문제도 있다. 널리 알려진 미국의 렉사일 지수, AR 지수, 국내의 LQ 지수나 KreaID 지수 등 많은 독서지수들이 사기업에 의해 주도되어 온 것이 사실이다. 때문에 그 개발 원리와 방법 등에 대한 실제적 과제와 현안 등이 공개되어 있지 않다.

본고는 신명선·이기연·나상수·인정남·차경미(2024)에서 제시한 ‘온독지수(Onthink Reading/O. R. Index)’의 사례를 중심으로 논의를 전개하고자 한다. 온독지수는 현재 국내 도서 214권에 부여되어 있으며 충남교육청의 ‘온시스템’에 구현되어 있다. 온독지수는 지수 산출의 근거를 포함하여 도서의 수준별 위계가 모두 공개되어 있다. 따라서 독서지수 시스템 개발에 관심을 갖고 있는 연구자들이 참조할 수 있는 매우 좋은 사례이다.

본고는 온독지수의 사례를 중심으로 실제 독서지수 시스템 개발 과정에서 부딪힐 수 있는 과제를 해결해 가는 방법을 제안하고자 한다. 독서지수 시스템 개발의 맥락과 과제, 해결 방법 등을 구체적으로 기록하여 추후 독서지수 시스템을 개발하고자 하는 연구자나 교육자들이 이에 대한 깊이 있는 이해에 도달할 수 있도록 논지를 전개하고자 한다.

본고는 크게 다음과 같은 두 가지 독서지수 시스템 개발 관련 과제를 다루고자 한다. 첫째는 독서능력 검사 문항의 설계와 관련 문항 확보 과제이다.

관련 사례로부터 이루어져야 해당 논의의 유의미성이 커지기 때문이다. ‘실제적 개선’을 추구하는 학문 분야들이 이러한 경향을 띠는 경우가 많은데, 경영학이 대표적인 경우이다. 경영학의 핵심 연구 방법 중 하나가 특정 기업에 관한 사례 연구임은 널리 알려져 있다.

예컨대 온독지수처럼 난이도에 따른 도서지수는 부여되어 있으나 정작 독자들은 자신의 독서능력 수준이 어느 정도인지 모를 경우와 관련된다. 이럴 경우 독자의 독서능력을 검사하여 도서지수를 부여해 주어야 한다. 여기서 독서능력 검사 문항 개발 과제가 발생한다. 수준별 도서 목록이 존재한다고 하더라도 독자가 자신의 독서능력 수준을 모른다면 해당 도서 목록의 활용도는 상대적으로 낮아질 수밖에 없다. 마찬가지로 독자가 자신의 독서능력 수준을 안다고 하더라도 자기의 수준에 맞는 적절한 도서 목록을 확보할 수 없다면 독서 활동의 활성화는 쉽지 않다. 따라서 도서의 수준별 위계화가 이루어진 경우에 추가적으로 독자의 독서능력에 따른 수준별 위계화 작업이 이루어져야 한다. 후술되겠지만, 본고에서는 이를 MST와 연계하여 다루고자 한다.

둘째로, 도서지수와 독서지수의 연계 방안 문제를 다루고자 한다.⁴⁾ 도서지수가 이미 설정되어 있는 상황에서, 독자들의 독서능력 검사를 시행하여 결과치를 얻은 뒤에 부딪히는 또 다른 과제가 해당한다. 만일 독자 집단이 소규모라면 전문가들이 독자의 읽기 수준을 가늠하여 적절한 도서지수를 곧바로 부여하는 것도 가능할 것이다. 그러나 온독지수처럼 대규모 독자 집단을 상정해야 하고 질적 검사가 불가능한 상황에서 대규모 양적 검사를 시행해야 하는 경우라면, 독자에게 독서지수를 어떻게 부여하는 것이 타당한가 하는 과제가 부각된다.

본고는 도서지수는 이미 구축되어 있는 상황에서 대규모 학생들을 대상으로 독서지수 시스템을 온라인으로 구축하려는 경우를 구체적인 맥락으로 삼는다. ‘온독지수’가 그러한 경우로서, 이와 같은 상황에서 위와 같은 두 가지 문제를 어떤 방식으로 해결할 수 있는지 온독지수 사례를 기반으로 그 해결 방안의 일단을 탐구해 보고자 한다.

논의는 크게 다음과 같은 방식으로 전개하고자 한다. 먼저 Ⅱ장은 Ⅲ장

4) 도서지수 및 독서지수 용어에 관해서는 Ⅱ장에서 후술된다.

과 IV장의 기반 논의에 가깝다. 몇 가지 주요 용어의 개념, 독서지수 시스템 개발의 과정과 단계, 온독지수에 대한 개관 등을 다루고자 한다. III장에서는 독자의 독서능력 검사를 위한 문항의 개발과 확보 관련 문제를 다루고자 한다. IV장은 MST 설계 방향, 독서지수와 독서지수의 연계 방안 등을 다루고자 한다.

II. 독서지수 시스템 개발의 전제와 과제

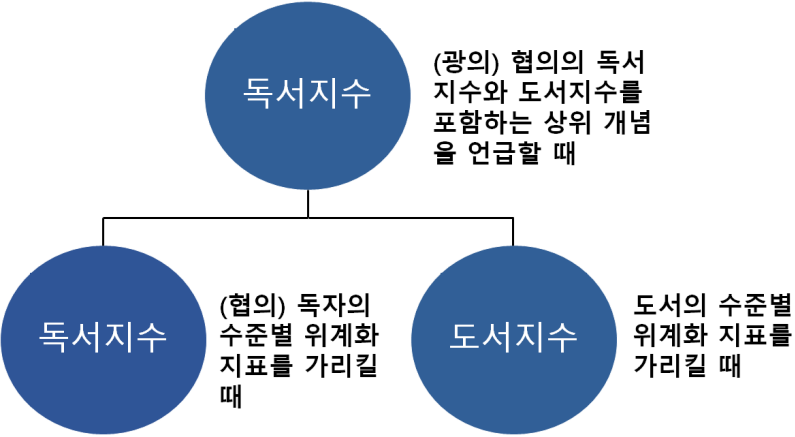
먼저 ‘독서지수(reading index)’와 ‘도서지수(book index)’라는 두 용어를 구분해 두고자 한다. 말 그대로 독서지수가 ‘독자’에 주목한 용어라면 도서지수는 ‘도서’에 주목한 용어인데, 사실상 이 둘을 명확하게 구별할 수 있는가에 대해서는 이견이 존재할 수 있다. 도서의 난이도를 설정하기 위해서는 독자를 가정할 수밖에 없고 독자의 독서능력 수준을 측정하기 위해서는 수준별 도서가 가정되는 것이 통상적이기 때문이다.

본고에서는 독서지수와 도서지수라는 용어가 갖는 위와 같은 개념적 한계를 인정하지만, 논의의 편의상 이 둘을 구분하여 사용하고자 한다. 독서지수는 독자들의 독서능력 수준별 위계화 지표로, 도서지수는 도서의 수준별 위계화 지표를 가리킬 때 사용하고자 한다. 본고의 목적상 이들을 구분하는 것이 논지 전개상 유리한 국면이 많기 때문이다.

그런데 일상적으로 독서지수는 도서지수를 포함하는 상위 개념으로 사용되기도 한다. 이들의 정확한 개념, 그 범위와 한계에 대해서는 차후 깊이 있는 논의가 이루어져야 할 것이다.⁵⁾ 본고에서는 독서지수를 광의의 개념과

5) 독서지수, 이독성(readability), 난이도, 난도, 도서지수 등 다양한 용어가 사용되고 있다. 이들 용어들의 개념적 함의가 다소 다르긴 하지만, 이는 본고의 목적과는 거리가 있으므로

협의를 개념으로 나누어 사용하기로 한다. 독서지수를 광의의 개념으로 사용할 때에는 도서지수를 포함하는 상위 포괄적 개념으로 보고자 한다. 협의의 개념으로 볼 때에는 도서의 수준별 위계화 지표를 나타내는 독서지수에 대응되는 개념으로서 독자의 수준별 위계화 지표를 가리키는 용어로 사용하고자 한다(〈그림1〉 참조). 한편 독서지수와 도서지수를 모두 개발하고 두 지수의 연계까지 이루어진 최종 산출물을 본고에서는 ‘독서지수 시스템’으로 지칭하기로 한다.

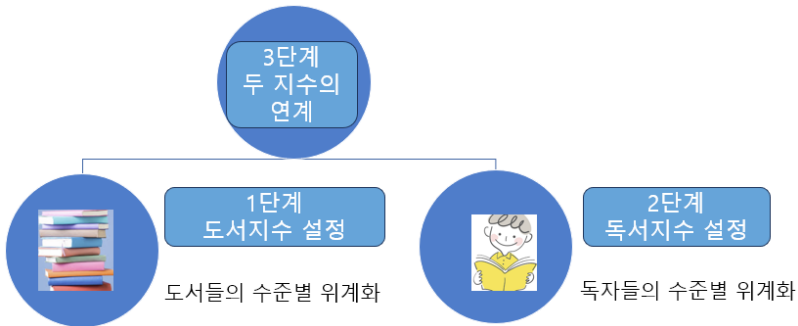


〈그림 1〉 독서지수와 도서지수의 개념 관계

독서지수 시스템 개발에는 다양한 과제가 도출된다. 등급의 개수, 등급별 간격, 등급 설정의 기준, 독서지수와 도서지수의 연계 방안 등등 다수의 연구 과제와 개발 난제들이 부각된다. 이를 고려하여 독서지수 시스템을 개발하기 위해서는 크게 다음과 같은 3단계 과제를 해결해야 한다. 1단계는 도서를 수준별로 분류한 뒤 일정한 체계를 부여하여 도서지수를 완성하는 단계이다. 2단계는 독자들을 수준별로 위계화하는 과정인데, 이 작업은 독서능

로 본고에서는 다루지 않기로 한다.

력 검사를 통해 독자들이 자신의 독서능력 수준을 확인할 수 있도록 하는 단계이다.⁶⁾ 3단계는 1단계와 2단계를 연계하는 작업으로서 독자들이 독서능력 검사를 통해 자신의 독서능력을 확인한 뒤 도서지수와 연계하여 자신에게 맞는 도서가 어느 정도 수준인지를 확인할 수 있게 해 주는 단계이다(〈그림 2〉 참조). 개발 과정에 따라 1단계와 2단계의 순서는 바뀔 수도 있다. 도서지수는 기본적으로 독자의 수준을 고려하여 이루어지며 그 반대도 마찬가지이기 때문이다.⁷⁾



〈그림 2〉 독서지수 시스템 개발의 3단계(가안)

독서지수 시스템 개발의 3단계를 이해하기 위해 영어권의 영향력 있는 독서 지원 프로그램인 르네상스 리딩(Renaissance Reading)사의 Accelerated Reader(이하 AR)을 살펴보자. AR은 크게 Atos Book Level, 즉 BL 지수와 SR(Star Reading) 검사로 구성된다. BL 지수가 도서지수로서 도서의 수

- 6) 예컨대 현행 중학교 2학년 학생이 중2 수준의 독서능력을 지녔다고 단정할 수는 없다. 중2 학생이라 하더라도 고1 수준의 독서능력을 지녔을 수도 있고 초6 수준의 독서능력을 지녔을 수도 있다. 따라서 도서지수가 유용하게 활용되기 위해서는 독자들이 자신의 독서능력을 측정하여 자신의 수준이 어느 정도인지를 명확하게 알 수 있어야 한다.
- 7) 수준별 독서 지원 시스템이 효과적으로 작동하려면 반드시 독서지수와 도서지수가 쌍으로 개발되어야 한다. 그렇지 않으면 교육적 효율성은 매우 저하된다.

준을 나타내는 지표라면, SR 검사는 독자의 독서능력을 확인하기 위한 독서 검사이다. AR은 SR 검사를 통해 독자의 독서능력을 확인한다. 독자의 독서 능력은 BL 지수에 맞게 산정된다.⁸⁾ BL 지수 개발 작업은 <그림 2>의 1단계에 해당하며 SR 검사는 독자들의 수준별 위계화를 위한 도구로서 <그림 2>의 2 단계에 해당한다. SR 검사 실시 후 학생의 독서능력을 BL 지수에 맞추어 지표화하는 작업이 <그림 2>의 3단계에 해당한다고 볼 수 있다.

충남교육청 온시스템에 구축되어 있는 온독지수는 광의의 개념으로 보면 독서지수이다. 그렇지만 아직 독자의 독서능력 위계화 방법을 제시하지 못했기 때문에 엄밀하게 보면 도서지수라고 볼 수 있다. <그림 2>의 개발 과정에서 1단계만 구축되어 있는 상황으로서, 2단계와 3단계 작업이 요청되는 상황이다.

본고는 독서지수 시스템의 2단계와 3단계 개발에 대한 탐구 과정의 일단을 담고 있다. 굳이 온독지수가 아니더라도 2단계와 3단계 과제를 해결해야 하는 상황은 독서지수 시스템 개발 과정에서 늘 맞닥뜨릴 수 있다. 본고는 기구축된 도서지수가 있을 때에 해당 도서지수에 맞게 독서지수를 설정하고 두 지수를 연계하는 실질적 방안에 대한 탐구 기록이다. 본고는 그와 같은 상황에서 취할 수 있는 한 가지 방법론적 해안을 제시하고자 한다.

우리나라에는 다양한 독서 운동 단체나 협회들이 존재하고 이들 중 몇 단체들은 꾸준히 수준별 도서 목록을 제공하고 있다. 이들 목록이 독서지수라고 부를 만큼 체계적인가 여부는 차치하더라도 이러한 목록이 다양하게 존재한다는 점은 고무적이다. 그에 비하면 독서지수 설정 작업은 매우 더딘

8) 'AR 지수'는 미국의 르네상스 리딩에서 개발한 독서지수로서, Atos Book Level, 즉 BL을 가리킨다. 정확하게 말하면 AR은 Accelerated Reader이므로, 르네상스 리딩사에서 제공하는 독서 지원 프로그램을 가리킨다. 이 프로그램에는 읽기 후 퀴즈 프로그램 등이 포함된다. 그러나 흔히들 Atox Book Level을 AR 지수로 통칭하며, BL로도 불린다. AR 지수를 활용하기 위해서는 SR 검사를 봐야 한다. 이 시험은 독서능력을 측정하는 검사로서, 이를 통해 GE 점수를 얻을 수 있다. GE 점수가 나오면 나에게 맞는 영어 도서를 선택하여 고를 수 있다.

것이 사실이다. 상황이 이러할 수밖에 없는 것은 독서지수 설정 작업이 상대적으로 매우 지난한 작업 과정을 요구하기 때문이다.

독서지수를 설정하기 위해서는 학생들의 독서능력 검사 작업이 필수적으로 요구된다. 검사 문항의 개발과 시행, 독서능력의 측정과 해석, 독서 수준의 분류와 체계 등등의 문제가 제기된다. 이때 해당 작업은 기존 도서지수와 연계하여 수행되어야 한다.

독서능력 검사 결과에 기반하여 독서지수를 설정한 뒤 도서지수와 연계하는 작업도 간단치가 않다. 학생이 80점을 맞았다고 했을 때 이 학생의 독서 수준은 어느 정도일까? ‘상’ 수준일까, ‘중’ 수준일까? 기존 도서지수와 어떻게 연계해야 적절할까?

독서지수와 도서지수, 두 지수의 연계 방안은 다양하게 제안될 수 있다. 무엇보다 해당 지수가 특별한 목적을 갖고 있다면 지수별 특징에 맞는 연계 방안이 탐구되어야 한다. 본고는 실제성(reality)과 실행 가능성(practicality)에 초점을 두어 본고가 설정하고 있는 구체적인 맥락에서 현실적으로 적용 가능한 구체적 방안을 탐구하는 데 초점을 두고자 한다.

온독지수는 수준별 도서 목록을 꾸준히 제공하고 있는 국내 5개 독서 협회들이 추천한 총 4,867권을 대상으로 2회 이상 중복 추천된 도서를 말뭉치화한 데에서 출발하였다. 해당 도서들의 수준을 총 110명의 사서 및 사서교사를 동원하여 질적 평정하고 온독지수 산출식⁹⁾에 근거한 양적 검사 결과와의 상관도를 통계적으로 분석하였다. 결과적으로 도서의 수준을 12단계로 나누고 총 214권의 책에 온독지수를 부여하였다(이상 온독지수에 대해서는 신명선 외, 2024 참조). 온독지수가 주목되는 이유는 ‘사고도구어(academic vocabulary)의 다양성과 빈도’ 중심으로만 도서지수를 설정했다는 데에 있다. 통상적으로 이독성에 영향을 미치는 요소(문장의 복잡도, 응집성 등)는 매

9)
$$[0.7(\text{사고도구어 CTTR}) + 0.3(\text{사고도구어 빈도수})] \times 500 + 100$$
(신명선 외, 2024, p. 135)

우 다양한데 이 연구는 단지 사고도구어만으로 온독지수를 설정하였는데도 전문가 질적 평정과 상관도가 0.832에 이르렀다. 이는 사고도구어가 도서의 수준을 위계화하는 매우 중요한 지표임을 보여 준다. 한편 지수 산출의 근거가 공개되어 있는 데다가, 무엇보다 사고도구어라는 하나의 변인만으로 간단하게 도서지수를 설정하여 현장 적용의 편리성과 유용성이 매우 크다.

온독지수 12단계는 다음처럼 구성되어 있다.

〈표 1〉 예상 학년을 고려한 온독지수 체계표

온독지수	예상 학년
1.1~1.9	초등학교 1학년
2.1~2.9	초등학교 2학년
3.1~3.9	초등학교 3학년
4.1~4.9	초등학교 4학년
5.1~5.9	초등학교 5학년
6.1~6.9	초등학교 6학년
7.1~7.9	중학교 1학년
8.1~8.9	중학교 2학년
9.1~9.9	중학교 3학년
10.1~10.9	고등학교 1학년
11.1~11.9	고등학교 2학년
12.1~12.9	고등학교 3학년

본고는 〈표 1〉을 편의상 하나의 도서지수로 간주하고 그에 걸맞은 학생들의 독서능력 검사, 그리고 도서지수와 독서지수의 연계 방안을 다루고자 한다. 다시 말해서, 학생들이 자신의 독서능력을 검사한 뒤 자신의 독서능력이 〈표 1〉의 온독지수 중 어디에 해당하는지를 확인하는 방안을 다루도록 한

다. 이어서 III장에서는 〈그림 2〉 독서지수 개발 과정의 2단계 관련 과제를, IV장에서는 3단계 관련 과제를 논의하기로 한다.

III. 독서능력 검사 설계와 MST용 문항 확보

본고는 앞에서 독서지수 시스템의 효율적인 작동을 위해서는 독서지수와 독서지수의 개발 및 둘 간의 유기적인 연계 작업이 필수적임을 논의하였다. 이러한 연계는 독자의 독서능력을 어떻게 검사하고 해당 결과를 독서지수와 어떻게 연계할 것인가의 실천적 문제로 이어진다. 독서지수 설정을 위한 문항은 독서지수의 성격과 구조를 바탕으로 개발되어야 그 유의미성을 확보할 수 있을 것이다. 이 장에서는 먼저 국내외 독서지수 시스템의 검사 문항이 어떻게 구성되어 있는지를 탐색하고자 한다. 이는 온독지수 기반의 독서능력 검사 문항 개발의 방향성과 특수성을 이해하는 중요한 참고점이 될 것이다. 그리고 본고에서 개발한 독서능력 검사 문항의 구조와 내용, 검사 결과, 문항 확보 과정 등을 제시하고자 한다.

1. 독서능력 검사 설계

1) 국내외 독서지수 시스템의 독서능력 검사 현황

독서지수와 연계하여 독서능력 검사를 오랫동안 시행해 왔으며 정보 확인이 가능한 다음의 세 가지 대표적인 사례를 중심으로 독서능력 검사 방법을 살펴보고자 한다. 첫째 메타메트릭스(Metametrics)의 렉사일 프레임워크(Lexile Framework), 둘째 르네상스 러닝(Renaissance Learning)의 ATOS 지수와 SR(Star Reading) 검사, 셋째 (주)날말의 LQ 검사가 그것이다.

먼저 렉사일 프레임워크는 메타메트릭스에서 개발한 렉사일 지수를 바

탕으로 한다.¹⁰⁾ 렉사일 지수는 미국의 공통 핵심주 교육과정(이하, CCSS)에 포함된 독서지수 중 하나로서 공신력이 있으며 전 세계 영어 교수·학습에서 응용되고 있다. 렉사일 지수는 학생이 검사에 응시함으로써 측정되는 렉사일 읽기 지수 즉 독서지수와, 도서 등의 다양한 텍스트를 자체적으로 분석한 결괏값인 렉사일 도서지수를 연계하고 있다. 독서능력 검사 문항에 포함되는 글도 렉사일 도서지수를 이용한 것이므로 반복적인 시행을 통해 독서능력 검사 결과와 도서 수준을 연계하는 것이 가능하다. 독자 수준과 도서 수준 각각의 측정값은 0~2000L 범위에서 동일한 척도로 연계된다. 또한 최근에는 온라인 기반의 단계적 적응형 검사(Multi-Stage Adaptive Test, 이하 MST)¹¹⁾로 독서능력을 검사하고 있다. 주요 내용은 다음 <표 2>와 같다.

<표 2> 렉사일 프레임워크의 주요 내용

항목	내용	비고
문항 수	• 약 20~30개	
문항 유형	• 크게 5개 유형(그림, 한 문장, 두 문장, 중심 생각, 빈칸 채우기형)으로 구분됨. • 4지 선다형(1개 정답, 3개 오답)	• 독자의 수준 및 특징을 고려하여 서로 다른 유형을 활용해 검사 문항을 구성함.
세트 구성	• 고정된 세트는 없음. • 적응형 검사 방식	• 시험 중 검사자의 응답 결과에 따라 난이도가 조정됨.
검사 시간	• 약 30~45분 • 개인별 검사 속도에 따라 검사 시간이 유동적임.	• 기본적으로 개인의 능력을 파악하기 위한 목적에서, 검사 시간에 제한을 두지 않음.
검사 점수	• 0~2000 Lexile 지수	
기타	• 렉사일 프레임 기반의 검사가 이루어짐. • 문제 은행식 검사 또는 맞춤형 제작 검사도 가능함.	

10) 이하 렉사일(2025) 홈페이지(<https://hub.lexile.com>) 참조
 11) MST는 응시생이 이전 단계에서 보인 능력 수준에 따라 다음 단계의 소검사 또는 문항 수준이 조절되는 검사 방식이다.

다음으로 ATOS 지수는 르네상스 러닝에서 개발한 도서지수로 렉사일 지수와 함께 미국 CCSS에 활용되고 있다.¹²⁾ 텍스트의 수준을 측정하기 위해 렉사일 지수가 단어의 빈도를 활용한 것과 달리, ATOS 지수는 어휘를 등급화하여 활용하였다는 점이 특징적이다. 전술한 것처럼 SR 검사는 ATOS 지수와 연계되어 개발된 독서능력 검사이다. 이 지수는 0.1부터 16까지로 구성된다. 1 이하는 유치원생에, 1은 초등학교 1학년에, 16은 대학교 4학년에 차례로 대응된다. 그리고 렉사일의 독서능력 검사처럼 최근 온라인 기반의 MST를 적용한다. 주요 내용은 다음 <표 3>과 같다.

<표 3> SR 검사의 주요 내용

항목	내용	비고
문항 수	• 두 가지 버전으로 제공됨. ① 확장판(Enterprise): 34개 ② 축소판: 24개	• 축소판의 경우 확장판에 비해 더 자주 시행되어 학습 진도 추적에 사용되기도 함.
검사 영역	• 크게 5개의 영역으로 구분함. ① 어휘(Vocabulary) ② 이해 전략과 의미 구성(Comprehension) ③ 문학 텍스트 분석(Literacy) ④ 작가의 특성(Author) ⑤ 논쟁 및 검사 텍스트(Argument)	
문항 유형	• 3지 선다형(1개 정답, 2개 오답)	
세트 구성	• 고정된 세트는 없음. • 적응형 검사 방식	• 시험 중 검사자의 응답 결과에 따라 난이도가 동적으로 조정됨. • 사실상 무한한 형태의 시험지 제작이 가능함.
검사 시간	• 두 버전 모두 약 20~35분 • 세 가지 시간 제한 유형이 있음. ① 표준(Standard) ② 연장(Extended) ③ 무제한(Unlimited)	• 기본적인 시험 자체는 시간 제한이 없음. • 일부 문항은 시험 진행을 원활하게 하고 보안을 유지하기 위해 시간을 제한하기도 함.

12) 이하 르네상스(2025) 홈페이지(<https://www.renaissance.com>) 참조

검사 점수	• 600~1400으로 조정된 점수를 산출함.
기타	• 문제 은행이 매우 크고, 소프트웨어는 학생이 이미 본 문항을 추적하기 때문에 120일 동안 특정 문항이 동일 학생에게 반복되지 않음.

마지막으로 (주)날말의 LQ 검사는 (주)날말에서 개발한 도서지수인 ‘LQ’와 연계된 검사이다.¹³⁾ 검사 영역은 크게 어휘력과 독해력으로 구성되며, 독해력은 사실적 독해, 추론적 독해, 비판적 독해로 구성된다. LQ 검사의 급수는 다음 <표 4>와 같이 7급에서 1급으로 구성되며, 각 급수별 응시 권장 학년이 제시된다. 상급 수준으로 갈수록 문항 수와 소요 시간이 증가하며, 검사 문항에 포함되는 텍스트 난이도 수준이 급수별로 정해져 있다.

<표 4> LQ 검사 급수별 검사 대상 및 소요 시간

급수	적정 대상	문항 수	소요 시간	검사 문항의 LQ 범위	독자의 LQ 범위
7급	초등 1~2학년 (읽기 준비·학습기)	25	40분	100~480	100~760
6급	초등 3~4학년 (읽기 경험기)	30	40분	350~630	130~930
5급	초등 5~6학년 (읽기 독서기)	30	50분	560~840	210~1040
4급	중등 1~2학년 (분석 독서기)	35	50분	720~1000	370~1300
3급	중등 3~고등 1학년 (전략 독서기)	35	50분	910~1190	470~1470
2급	고등 2~3학년 (독립 독서기)	40	60분	1120~1450	530~1660
1급	대학생·일반인 (전문·교양 독서기)	40	60분	1320~1850	680~1850

13) 이하 (주)날말 홈페이지(<https://www.natmal.com>) 참조

LQ 검사 결과는 도서 추천에 활용되며, LQ 지수 대비 ± 100 구간에 있는 도서를 추천해 준다. 도서 수준에 따른 독서능력 및 동기에 관한 간단한 안내 등도 부가적으로 제공한다. 단, 지수 개발의 원천 자료는 회사의 상업적 목적과 직결되기에 비공개하고 있으며 이에 따라 (주)날말의 학생 검사 문항이나 결과 데이터 등을 본 연구에 참고하기에는 어려움이 있다.

이와 같이 국내외 독서지수 시스템의 독서능력 검사 현황을 검토함으로써 본고에서 얻을 수 있는 시사점을 정리하면 다음과 같다. 첫째, 온라인 기반 검사 선택의 장점이다. 온라인 기반 검사는 검사의 유연성과 확장성을 제고한다. 응시자들은 특정 시간과 공간에 제약받지 않고 개별적으로 응시할 수 있으며, 검사 주체 또한 다양한 난이도의 수준별 검사지를 구성하여 각기 다른 수준의 독자들을 대상으로 정교한 독서능력 위계화 체계를 구축할 수 있다. 또한 확장성의 측면에서, 응시자들의 응시 결과 분석을 토대로 지속적인 진단 및 피드백 체계를 마련할 수 있다. 이러한 점에서 온라인 기반 검사는 독서지수 시스템의 실질적인 운영 가능성을 높이는 데 기여할 수 있을 것이다.

둘째, 렉사일 검사와 SR 검사처럼 MST를 적용함으로써 검사의 기능적 효율성을 확보할 수 있다. MST는 표준화된 선형 검사가 일반적으로 잘 측정하는 중위 수준의 능력뿐 아니라 최하위와 최상위 능력 수준까지 타당하게 측정할 수 있으며, 응시자의 읽기 수준에 맞는 평가를 가능하게 함으로써 응시자의 중도 포기율을 낮추는 데에도 크게 기여하는 것으로 알려져 있다.

셋째, 온라인 환경에서 제한된 시간 내에 안정적인 검사를 실시하기 위해서는 상대적으로 간결한 문항 유형을 활용하는 것이 유리할 것이다. 실제로 앞에서 살펴본 주요 검사들에서도 선다형 중심의 짧고 명료한 문항 유형이 주를 이루고 있다. 선다형 문항은 온라인 기반 MST의 효율성을 극대화하는 데에도 유리하다.

넷째, 독서능력 검사의 문항 설계는 해당 독서지수 시스템이 전제하고 있는 이론적 기반과 교육 방향을 반영해야 한다. 이는 렉사일 프레임워크, SR

검사, LQ 검사 등에서도 일관되게 확인되는 원칙으로서, 독서능력 검사 문항은 해당 독서지수가 추구하는 검사 목적 및 교육적 지향과 조응해야 하는 것이다. 앞서 살핀 세 가지 검사의 문항 수, 문항 유형 등에서 차이가 발생하는 이유는 각각의 독서능력 검사가 해당 독서지수 시스템의 특징을 반영하고 있기 때문이다.

이러한 점에서 본고는 사고도구어 기반 온독지수의 특징을 반영하여 독서능력 검사를 설계할 필요가 있을 것이다. 다시 말해서 기존의 일반적인 독해 능력 검사 문항들과는 달리, 피험자가 얼마나 다양한 사고도구어를 정확히 이해하고 활용할 수 있는가를 포함하는 것이 적절할 것이다.

2) 독서능력 검사의 설계 방향과 MST용 문항 확보

본고는 온라인 기반의 MST라는 검사 방식의 구조적 특성과, 온독지수라는 독서지수의 이론적 기반을 종합적으로 고려하여 독서능력 검사 문항을 설계하고자 한다. 이는 독서능력 검사가 응시자의 독서능력이 어떠한지를 단순히 기술적으로 보고하는 차원을 넘어, 독서지수와 연계된 독서지수 체계의 적합성과 교육적 실효성을 확보하기 위한 실천적 사례로서 전략적 설계 과정으로 자리매김해야 하기 때문이다.

학생들이 온라인에서 자신의 독서능력을 MST로 측정하도록 하기 위해서는 다수의 검사 문항 확보가 필수적이다. 그 문항들은 온독지수의 특성을 반영하고 있어야 하며 수준별(‘상·중·하’ 등)로 구분되어 있어서 학생들의 실력에 맞게 자동적으로 수준별 문항이 제시될 수 있어야 한다. 예컨대 A 학생이 a 문항을 맞췄다면 그다음에는 a 문항보다 더 어려운 문항인 b 문항이 제시될 수 있어야 한다. 만일 B 학생은 a 문항을 틀렸다면 그다음에는 a 문항과 유사한 문항이 제시되거나 그보다 쉬운 문항이 제시되어야 한다.

정리하면, “독서능력 검사 문항 개발과 다수의 문항 확보”라는 과제의 핵심은 ① 독서지수의 특성을 독서능력 검사 설계에 어떻게 반영할 것인가, ② 수준별 문항(‘상·중·하’ 등)을 얼마나 어떻게 확보할 것인가에 있다. ①과

관련하여, 온독지수가 갖는 가장 중요한 특성은 사고도구어의 다양성과 빈도를 기반으로 만들어졌다는 점이므로, 독서능력 검사 설계 시 사고도구어에 대한 이해 여부를 평가 내용에 포함할 수 있다.

②는 사실상 해결하기 매우 어려운 난제에 해당한다. 문항의 수준에 대한 판별은 대규모 인원을 대상으로 수십 년간 검증하더라도 단언하기가 어려울 것이다. 이는 매년 학생들의 수준이 변화하고 있는 데다 동일 시기라 할지라도 주어진 검사 맥락이나 문항 해석 방식에 따라 문항에 대한 정답률과 변별도 등이 변화하기 때문이다.

문항의 수준에 대한 판별은 전문가 집단에 의해 질적으로 이루어질 수도 있다. 그러나 그 판단의 결과를 전적으로 신뢰하기 어려울뿐더러 대규모 인원을 대상으로 할 때에는 실행 가능성이 매우 떨어지기도 한다.

일정한 피검사자 집단을 상정하여 검사를 시행한 뒤 문항별 난이도와 변별도를 확보하는 양적 방법을 사용할 수도 있다. 그러나 이 역시 그 결과를 무조건 신뢰할 수 있는 것은 아니다. 따라서 문항의 수준에 대한 판별은 양적 방법을 사용하거나 질적 방법을 사용하거나 모두 한계를 갖는다는 점을 인정하고 개발에 착수할 수밖에 없다.

이때, 대규모 인원을 대상으로 실시하는 MST에서는 현실적으로 전문가 질적 평정을 도입하기 어렵다. 그런 측면에서 양적 방법, 즉 대규모 검사를 시행하고 통계적 분석 방법을 활용하여 문항의 변별도와 난이도를 확보하는 방법은 매우 번거롭고 작업 과정이 까다롭지만 오히려 현실적인 개발 방법이 될 수 있다.

이처럼 양적 방법은 상기(上記)한 한계를 갖지만, 학생들의 독서능력 검사 데이터가 쌓임에 따라 그 한계점은 상당수 보완될 가능성이 높다는 점에서 고려 대상이 된다. MST 시스템이 갖추어진 뒤 온라인으로 자신의 독서능력을 측정하는 학생의 수가 많아진다면, 다수 데이터 확보에 따라 문항의 변별도와 난이도에 대한 예측력은 높아질 것이기 때문이다.

결과적으로 본고에서는 ① 도서지수의 특성을 독서능력 검사 문항 설계

에 반영하기 위해, 사고도구어 기반 어휘 능력 검사를 독서능력 검사에 포함할 것이다. 그리고 ② ‘상·중·하’ 등의 수준을 알 수 있는 문항을 확보하기 위해 대규모 학생들을 대상으로 실제 독서능력 검사를 시행하여 문항별로 변별도와 난이도를 확보하는 방식을 선택하고자 한다.

한편, MST로 독서능력 검사를 설계할 때에는 다수의 문항 확보가 필수적이다. 나이가 학생의 독서능력은 물론, 과학기술 및 사회 역시 시간에 따라 변화하므로, 검사 문항은 시의적절함과 학문적 타당성을 고려하여 지속적으로 보완되고 수정되어야 한다. 즉 다수의 문항을 확보한 뒤 지속적으로 수정, 보완해 가면서 계속해서 추가 문항을 확보해 나가야 한다. 만일 문항의 수가 적다면 몇 번의 독서능력 검사 응시를 통해 학생들이 문항을 외워버리는 사태가 발생할 수도 있다. 따라서 거대 문제 은행 구축이 필요하다.

문제는 독서지수를 개발하는 첫 단계에서부터 다수의 문항을 확보하는 것이 쉽지 않다는 데에 있다. 일정 시기에 대규모 문항을 개발하는 것 또한 바람직하지 않다. 전술한 것처럼 학문과 사회 등의 변화를 반영하여 문항 수정, 보완 작업이 지속적으로 이루어져야 하기 때문이다. 또 학생들의 변화 양상도 지속적으로 관찰하며 문항을 개발해야 하므로 문항 개발은 시대의 흐름과 발맞춰 가는 것이 바람직하다.

이러한 한계를 인정하고 나면, 남은 과제는 MST를 시행할 수 있는 최소한의 문항 수는 어느 정도인가, 지속적인 추가 문항 개발을 위한 문제 은행 시스템은 어떻게 구축할 것인가 등이 된다. 지속 가능한 문제 은행 시스템의 설계 문제는 또 다른 중요 논의 주제로서, 추후 과제로 넘기도록 한다. 본고에서는 MST 시행을 위한 최소한의 문항 확보 문제만 다루기로 한다.

2. 검사 문항 개발 및 검토

본고에서는 온독지수의 특성을 고려하여 검사 영역을 크게 ‘어휘 능력’ 검사와 ‘독해 능력’ 검사로 나누었다. 어휘 능력 검사는 어휘의 형태와 의미

를 정확하게 알고 문맥에 맞게 이해하고 표현할 수 있는 능력이 있는지를 측정하는 데 초점을 두었으며, 독해 능력 검사는 글의 내용(사실적 이해, 추론적 이해, 비판적 이해), 구조, 표현상의 특징을 깊이 있게 이해할 수 있는 능력이 있는지를 평가하는 데 초점을 두었다. 구체적인 검사 영역 및 하위 요소는 다음 <표 5>와 같다.

<표 5> 어휘 능력 및 독해 능력 검사 영역 및 하위 요소

영역	하위 요소		개념
어휘 능력	형태		단어의 정확한 형태(표기)를 아는가?
	의미	정의	단어의 사전적 정의를 아는가?
		변별	단어들의 의미를 변별하는가?(서로 다른 단어, 다의어 포함)
		의미 관계	유의어, 반의어, 상하의어 등 목표 단어와 의미 관계에 있는 단어를 아는가?
독해 능력	내용 이해	사실적 이해	글에 명시적으로 드러나 있는 내용상의 정보를 이해할 수 있는가?
		추론적 이해	글에 제시된 내용들 간의 관계나 글에 명시적으로 드러나 있지 않은 내용을 맥락에 맞게 설명할 수 있는가?
		비판적 이해	글 내용의 정확성, 적절성, 타당성, 신뢰성, 효용성 등을 판단할 수 있는가?
	구조 이해		글의 논리적 구조와 전개 방식 등 형식적 특징을 이해하고 있는가?
	표현 이해		글에 사용된 다양한 표현의 의미를 이해하고 표현상의 특징을 이해하고 있는가?

어휘 문항은 온독지수의 특성을 고려하여 사고도구어가 상당수 출제될 수 있도록 기획되었다. 어휘들을 사고도구어와 비사고도구어로 나눈다고 할 때에 사고도구어의 비율이 대략 60~70% 정도 되도록 구성하였다. 이때, 사고도구어의 경우 등급이 나누어져 있기 때문에 학교급별, 학년군별 수준을 고려하여 출제하였다. 사고도구어 등급은 신명선·강지영·나상수·이경남·정지호·홍경화 외(2024)를 기반으로 하였다(구체적인 것은 <표 6> 참조).

〈표 6〉 학년군에 따른 사고도구어 검사 범위

학년군	사고도구어 검사 범위(개)					세트 수 (개)	
	1급	2급	3급	4급	계		
[초 3, 4]	30	10			40	2	10
[초 5, 6]	10	20	10		40	2	
[중 1, 2]		10	20	10	40	2	
[중 3, 고 1]			20	20	40	2	
[고 2, 3]			10	30	40	2	

검사용 독해 지문은 렉사일 프레임워크와 마찬가지로, 온독지수가 부여되어 있는 수준별 도서(214권)에서 발췌하였다. 지문의 수준과 문항 수준이 적절히 어울리도록 하기 위함이었다. 컴퓨터의 한 화면에 하나의 문항이 들어갈 수 있도록 하기 위해 지문의 분량은 두 문단 이하로 제한되었으며 ‘1지문 1문항’의 형태를 선택하였다. 그리고 문항 수준에 따라 글자 수의 조정은 이루어지도록 하였다.

책의 분량에 따라 다르지만, 긴 책의 한두 문단을 기반으로 문항을 출제하면 이미 해당 책을 읽은 독자라 하더라도 책에 대한 지식만으로 문항 풀이가 어려울 것으로 예상되었다. 그렇지만 문항 출제 시 책에 대한 지식만으로는 문항 풀이가 불가능하도록 각별히 주의하였다.¹⁴⁾ 그리고 전체 총 문항에서 인문, 사회, 예술, 과학 분야가 골고루 들어가도록 비율을 조정하였다.

검사 문항은 학교급 및 학년군, 문항 난이도, 문항 제시 단계, 검사 범주 및 내용, 지문 내용의 분야 등을 고려한 문항 개발 틀에 따라 개발되었다. 먼

14) 다수의 독서 경험을 가진 독자가 좋은 점수를 받는 것이 합당하다는 관점에서, 해당 책을 읽은 독자가 문항 풀이에 유리하게 출제하는 것이 오히려 더 합리적이라는 주장도 있을 수 있다. 사실상 배경지식은 독서능력의 핵심 기반으로 작용하기도 한다. 실제로 렉사일 지수나 AR 지수의 문항들을 보면 독서 경험이 많은 독자들이 문항 풀이에 더 유리한 문항을 설계하기도 하는 듯하다. 이에 대해서는 추후 더욱 깊이 있는 논의가 필요하다.

저 학교급 및 학년군의 경우 (주)납말의 LQ 지수의 급별 구분, 우리나라의 국가 교육과정 등을 고려하여 [초 3, 4] / [초 5, 6] / [중 1, 2] / [중 3, 고 1] / [고 2, 3]으로 대별하여 제작하였다. 문항의 난이도는 ‘상, 중, 하’로 나누어 개발 하되, ‘상상’과 ‘하하’ 난이도의 보정 문항을 추가로 개발하였다. 또한 학생이 맨 처음 배정받게 되는 핵심 단계(‘중’ 수준)에 해당하는 문항의 수에 가중을 두었다(관련 내용 IV장에서 후술됨).

검사 범주 및 내용은 어휘 능력과 독해 능력의 세부 영역 및 하위 요소가 단계별로 적절히 배치되도록 배분하였다. 모든 문항은 렉사일 프레임워크처럼 4지 선다형으로 개발하였고, 검사 영역의 하위 요소들이 골고루 출제될 수 있도록 영역별 문항 개수를 조정하였다. 총 문항 수는 독자들이 MST에 응시하여 수준에 따른 검사를 진행할 경우 예상되는 경로를 고려하여 총 2세트가 가능하도록 설계하였다(MST 경로는 4장에서 후술됨). 검사 문항 개발 결과를 요약하여 제시하면 다음 <표 7>과 같다.

<표 7> 검사 문항 개발 결과

학년군	난이도	범주	문항 수	
[초 3, 4] [초 5, 6]	상	어휘	8	16
		독해	8	
	중	어휘	12	24
		독해	12	
	하	어휘	8	16
		독해	8	
	상상	어휘	2	4
		독해	2	
	하하	어휘	2	4
		독해	2	
소계			64개	

(중 1, 2) (중 3, 고 1) (고 2, 3)	상	어휘	8	16
		독해	8	
	중	어휘	13	26
		독해	13	
	하	어휘	8	16
		독해	8	
	상상	어휘	2	4
		독해	2	
	하하	어휘	2	4
		독해	2	
소계			66개	
전체 문항 계			326개	

검사 문항 개발에는 교사 20명(초등 10명, 중등 10명)이 참여하였다. 초등 교사의 경우 독서교육에 관심이 많고 전문성이 있는 교원을 충남교육청의 추천을 받아 구성하였으며, 중등 교사의 경우 충남교육청 내 독서교육 관련 전문성 학습공동체 소속 교원 중에 선발하였다. 구성된 교사들을 대상으로 검사 문항 개발 방향 설명회(2024년 10월 10일)를 개최하여 문항 개발의 방향과 원리를 서로 공유하였고, 「문항 개발 지침서」를 책으로 만들어 문항 개발과 수정 작업을 진행하였다. 문항 수정과 보완은 연구진과 문항 출제 교사 간의 협의를 거쳐 실시하였으며, 문항의 난이도 및 변별도, 타당도와 적절성을 비롯하여 문두, 지문, 답지, 선택지 등을 3차례에 걸쳐 수정·검토하였다.

3. MST용 문항 확보를 위한 검사 시행과 결과

1) 검사 시행

독서지수 검사에 사용할 문항의 타당도와 변별도를 확보하기 위해 학생들을 대상으로 실제 검사를 시행하였다. 충남교육청의 협조를 받아 2024년 11월~12월까지 23개 학교(초 17개교, 중 3개교, 고 3개교)의 약 1,600명을 대상으로 실시하였다. 검사 참여 인원은 다음 <표 8>과 같다.

<표 8> 학년군별/문항 난이도별 검사 참여 인원 (단위: 명)

학년군	집단 문항 난이도	A	B	C	계
		'하' 수준 ('하하' 포함) 문항	'중' 수준 문항	'상' 수준 ('상상' 포함) 문항	
[초 3, 4]	초 3	56	45	49	150
	초 4	50	53	48	151
[초 5, 6]	초 5	46	45	67	158
	초 6	54	74	66	194
[중 1, 2]	중 1	50	56	57	163
	중 2	51	56	60	167
[중 3, 고 1]	중 3	56	48	50	154
	고 1	58	42	53	153
[고 2, 3]	고 2	50	63	54	167
	고 3	46	64	59	169
계		517	546	563	1,626

검사 시행 전 교사를 대상으로 사전 교육을 실시하여(2024년 11월 16일) 시행이 원활히 이루어질 수 있도록 하였다. 이때, 검사를 통해 서로 다른

능력을 가진 학생이 동일한 문항을 응시한 결과 얻을 수 있는 ‘평균 정답률 (average accuracy rate)’¹⁵⁾과 변별도 등을 산출하고자 하였다. 따라서 하나의 집단에 하나의 난이도를 가진 문항을 배정하였다. 즉, 동일 학년의 집단 A는 ‘하’ 수준 문항(‘하하’ 포함)에만 응답하였으며, 집단 B는 ‘중’ 수준 문항에만, 집단 C는 ‘상’ 수준 문항(‘상상’ 포함)에만 응답하였다.¹⁶⁾ 난이도에 따른 응시 인원은 유사하도록 조정하였다.

2) 검사 결과 분석

MST에서는 학생 수준에 따라 서로 다른 문항들이 제시되어 학생들이 다 서로 다른 검사지를 받게 된다. 따라서 중요한 것은 개별 문항들 하나 하나의 난이도와 변별도이다. 그러나 지면의 한계상 총 326개 문항 전체의 세부 통계치를 제시하기는 어려우므로 학년군 및 문항 난이도별 검사지의 핵심적인 경향과 문항 예시를 중심으로 분석 내용을 기술하였다.

검사 결과 학년군별, 문항 난이도별 검사지의 정답률 및 변별도를 제시하면 다음 <표 9>와 같다.¹⁷⁾ 모든 학년에서 ‘하하’ 수준으로 갈수록 정답률이 높고, ‘상상’ 수준으로 갈수록 정답률이 낮아 문항 개발 단계에서 난이도 설정은 비교적 잘 이루어진 것을 확인할 수 있다. 다만, 학년군에 따라 동일한 난이도 문항에 대한 정답률에 차이가 있는데, 이는 이번 검사에 응시한 학생 특성 등이 반영된 것일 수 있어 추후 MST 적용을 통해 대규모 학생 데이터

15) 이하에 제시된 ‘정답률’은 모두 평균 정답률을 의미함.

16) 결과적으로 5개 학년군별로 A, B, C 3개 집단에 각각 다른 문항 난이도별 검사지를 제공하여, 총 15개의 세부 집단을 대상으로 한다.

17) 이하 정답률 및 변별도에 대한 해석은 성태제(2019: 282-284)를 따랐다. 문항의 정답률이 0.2 미만이면 ‘매우 어려운 문항’, 0.2~0.4이면 ‘어려운 문항’, 0.4~0.6이면 ‘중간 난이도 문항’, 0.6~0.8이면 ‘쉬운 문항’, 0.8~1.0이면 ‘매우 쉬운 문항’으로 처리하였다. 문항 변별도의 경우 0.1 미만이면 ‘변별력이 없는 문항’, 0.1~0.2이면 ‘변별력이 매우 낮은 문항’, 0.2~0.3이면 ‘변별력이 낮은 문항’, 0.3~0.4이면 ‘변별력이 있는 문항’, 0.4 이상이면 ‘변별력이 높은 문항’으로 처리하였다.

분석 결과를 면밀히 살펴볼 필요가 있다.

변별도의 경우 모든 학년에서 ‘하하’ 수준에서 제일 낮았다가, ‘중’ 수준 또는 ‘상’ 수준에서 제일 높고, ‘상상’ 수준에서 다시 낮아지는 경향성을 보였다. 이는 변별도의 특성상 정답률이 너무 높거나 낮을 경우 변별도가 낮아지는 경향이 있기 때문이다.

〈표 9〉 학년군별/문항 난이도별 정답률 및 변별도

구분	문항 난이도	‘하하’ 수준	‘하’ 수준	‘중’ 수준	‘상’ 수준	‘상상’ 수준	평균
[초 3, 4]	정답률	0.98	0.94	0.83	0.61	0.34	0.74
	변별도	0.09	0.17	0.36	0.53	0.42	0.31
[초 5, 6]	정답률	0.95	0.88	0.72	0.45	0.35	0.67
	변별도	0.19	0.30	0.49	0.46	0.41	0.37
[중 1, 2]	정답률	0.92	0.85	0.71	0.47	0.23	0.64
	변별도	0.20	0.33	0.42	0.54	0.24	0.35
[중 3, 고 1]	정답률	0.92	0.83	0.72	0.49	0.30	0.65
	변별도	0.20	0.36	0.54	0.59	0.38	0.41
[고 2, 3]	정답률	0.92	0.86	0.65	0.44	0.32	0.64
	변별도	0.16	0.33	0.56	0.45	0.13	0.33
평균	정답률	0.94	0.87	0.73	0.49	0.31	0.67
	변별도	0.17	0.30	0.47	0.51	0.32	0.35

검사 결과 분석을 토대로 추가적인 문항 난이도 조정 및 문항 수정 작업을 진행하였다. 먼저, 정답률이 과도하게 높거나 낮은 경우에는 문항의 세부 내용을 수정하였다. 그런가 하면 0.2 이하의 변별도를 가지는 36개 문항을 중심으로 오답지의 매력을 높이거나 정답지를 수정하였다. 이후 충남교육청 소속 교사 전문 공동체의 검토를 받아 검사 문항을 최종 확정하였다.

예시 문항을 수준별로 몇 가지 살펴보면 다음과 같다. 다만 개발 문항들은 실제로 MST에 사용될 예정이므로 문항 전체를 공개하기는 어려워 일부를 기호(○, △, ~ 등)를 사용하여 가렸다. 먼저, [초 3, 4] 학년군의 어휘의 ‘정의’를 파악하고 있는지를 확인하는 ‘하’ 수준 문항은 <그림 3>과 같다. 해당 문항은 비교적 낮은 등급의 사고도구어를 사용하여 출제함으로써 정답률이 0.91로 매우 높고, 변별도는 0.28로 낮게 나타났다.

▶ 다음 밑줄 친 단어의 뜻으로 알맞은 것을 고르세요.

■ 내가 이 책을 읽는 OO은 새로운 지식을 배우기 위해서야.

① 어떤 일을 △△하는 것
 ② 어떤 일이 일어나는 △△
 ③ 어떤 일을 하는 △△나 △△
 ④ 어떤 △△을 통해 얻는 △△ △△

<그림 3> [초 3, 4] ‘하’ 수준 문항 예시

다음으로 [중 1, 2] 학년군의 중 수준 독해 문항 예시는 <그림 4>와 같다. 이 문항의 경우 1개 문단 수준의 지문이 사용되었으며, 지문의 내용을 사실적으로 파악하는 능력을 측정하였다. 문항 정답률은 0.55로 중간 정도이고, 변별도는 0.76으로 높았다.

▶ 이 글의 내용으로 옳지 않은 것을 고르세요.

말하는 사람은 악의 없이 한 칭찬이라고 하더라도 (...중략...) 기득권을 가진 다수자가 소수자와 약자를 집단화해 편견과 고정 관념을 형성하기 때문이에요.

① 칭찬에 악의가 없으면 ~.
 ② 고정 관념이 ~ 좋지 않다.
 ③ 무식과 조롱은 ~ 이루어진다.
 ④ 고정 관념은 ~ 형성된다.

<그림 4> [중 1, 2] ‘중’ 수준 문항 예시

끝으로, [고 2, 3] 학년군 ‘상상’ 수준 문항 중에서 어휘의 ‘의미 관계’ 파악 문항은 <그림 5>와 같다. 해당 문항에는 비교적 높은 수준의 사고도구어가 사용되었으며, 어휘 간 유의 관계와 반의 관계를 동시에 판단해야 했다. 실제로 문항의 정답률은 0.25로 어려웠고, 변별도는 0.3이었다.

▶ 어휘의 의미 관계가 다음과 다른 것을 고르세요.

유동적	능	가변적	↔	고정적
	(유의 관계)		(반의 관계)	

(선택지 생략)

<그림 5> [고 2, 3] ‘상상’ 수준 문항 예시

본고는 실제 학생들을 대상으로 실시한 독서능력 검사를 통해 학년군 별, 문항 난이도별 문항의 정답률 및 변별도를 확보한 뒤, MST 검사용 문항을 2세트 확보하였다. 향후 MST를 적용하여 응시자 수준에 따른 문항 제시 경로와 문항별 응답 데이터의 관련성을 더욱 정교하게 분석한다면, 문항의 측정 타당도와 검사 도구로서의 실용성을 한층 더 확보할 수 있을 것으로 기대된다.

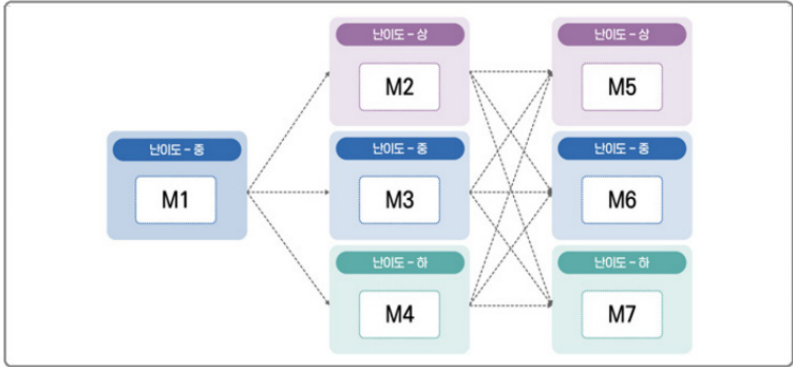
IV. MST 모델 설계 및 도서지수와와의 연계

1. MST 모델 검토

최근 온라인 검사 방식이 확산되면서 MST의 도입도 늘고 있다. MST를 시행한 선행 사례 중에서 읽기 관련 영역을 포함하고 있는 검사의 예로는

OECD의 국제 학업성취도 검사인 PISA와 호주의 자국용 학업성취도 검사인 NAPLAN을 들 수 있다.

PISA는 PISA 2018의 읽기 영역에서부터 MST 방식을 도입하였다. PISA 2018 읽기 영역의 MST는 ‘핵심 단계(Core)-1단계-2단계’의 총 3단계로 설계되었다(서민희·김미림·김완수·이소라·한정아·김성훈 외, 2020: 41). NAPLAN의 MST 역시 3단계로 구성되며, 첫 번째 분기 지점과 두 번째 분기 지점에서의 득점에 따라 배정되는 소검사가 달라진다. 읽기 영역에서는 각 단계의 소검사에 대해 3~4개 동형 검사를 활용한다(서민희 외, 2020: 39). 국내에서는 한국교육과정평가원에서 시행하는 맞춤형 학업성취도 자율검사에서 MST 도입을 계획하고 있다(박인용·김미림·박민호·곽민호·강혜진·박도영, 2023). 이 역시 총 3단계로 구성되며 ‘1-3-3’의 소검사로 구성된다.

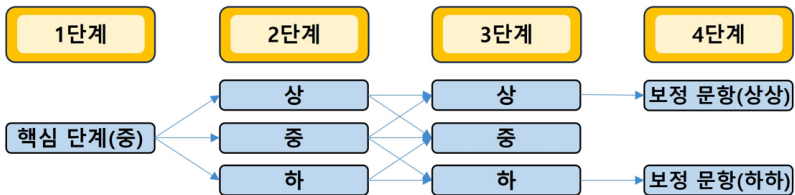


〈그림 6〉 맞춤형 학업성취도 자율검사의 MST 설계(안)(박인용 외, 2023: 12)

〈그림 6〉과 같이 1단계에서는 ‘중’ 수준 난이도의 1개 소검사, 2단계에서는 1단계의 응시 결과에 따라 난이도 ‘상·중·하’로 구성된 3개의 동형 소검사, 3단계에서는 2단계의 응시 결과에 따라 난이도 ‘상·중·하’로 구성된 3개의 동형 소검사로 이어진다.

2. MST 모델 설계

본고는 MST에 관한 선행 연구들에서 공통적으로 선택하여 안정성이 검증된 3단계 구조를 선택하였다. 다만 최하위 학생과 최상위 학생에 대한 세분화된 정보를 얻기 위해 보정 문항이 제공되는 ‘4단계’를 추가하였다. 이로써 ‘3+1’ 단계 체제를 구성하였다. 1단계인 핵심 단계는 MST의 일반적 구성과 동일하게 학생의 기본적인 능력을 확인하기 위한 단계이며 ‘중’ 난이도의 문항으로 검사지를 구성하였다. 2~3단계의 경우 ‘상, 중, 하’ 수준으로 세분화되며, 4단계는 일부 보정 문항으로 일부 학생에게만 제공된다. 이 자료는 영재나 부진아 등을 위해 최종 결과 제공 시 교육 활동 지원 방안으로 활용될 수 있다. 이를 그림으로 나타내면 다음 <그림 7>과 같다.



<그림 7> 본고의 MST 모델에 따른 학생 검사 경로

<그림 7>에 따르면 학생은 자신의 검사 점수에 따라 총 9개의 경로 ①~⑨ 중 하나의 검사지를 배정받게 된다.

- | | |
|----------------------|-----------------|
| ① 중(핵심) - 상 - 상 - 상상 | ② 중(핵심) - 상 - 상 |
| ③ 중(핵심) - 상 - 중 | ④ 중(핵심) - 중 - 상 |
| ⑤ 중(핵심) - 중 - 중 | ⑥ 중(핵심) - 중 - 하 |
| ⑦ 중(핵심) - 하 - 중 | ⑧ 중(핵심) - 하 - 하 |
| ⑨ 중(핵심) - 하 - 하 - 하하 | |

최상위 및 최하위에 대한 추가적 ‘4단계’가 없다면, 경로 ①과 ②, 경로 ⑧과 ⑨에 각각 해당하는 학생들의 수준 및 읽기에 대한 어려움의 양상을 세밀하게 파악하기 어렵다. 기존의 검사 구성으로는 독서능력이 ‘상’ 수준인 학생 내에서도 극도로 최상위에 속하는 학생과 정반대인 학생처럼 양극단에 속한 학생에 대한 검사가 충분하지 않을 수 있다.

한편, 본고는 MST의 난이도 조절 방식에 더하여 각 검사 요소의 중복을 최소화하는 배열을 통해 적은 수의 문항으로도 다양한 독해 능력과 어휘 능력을 효율적으로 검사할 수 있도록 설계하였다. 문항 수가 많아지면 학생의 검사 응시 동기를 저하할 뿐만 아니라, 학생의 응시 피로도가 증가함에 따라 불성실한 응답률이 증가하는 것을 방지하기 어렵다. 이에 Ⅲ장에서 제시한 검사 영역별 문항들이 MST 진행 과정에서 골고루 배열되도록 동일 영역의 문항만 반복해서 제시되지 않도록 문항 제시 순서를 정리하였다. 이러한 문항 배열 방식은 검사의 안정적 시행을 유도하여 검사 결과 데이터의 질을 향상하는 데 기여할 수 있다.

3. 독서지수와 도서지수의 연계

본고에서는 학생의 독서능력에 따라 서로 다른 난이도의 문항 세트를 배정받는 MST 방식을 목표로 하므로, 학생이 배정받은 문항 세트의 난이도에 따라 학생이 받은 점수를 변환하고, 이렇게 변환된 점수를 도서지수와 연계하는 방식을 사용하였다. 본고에서는 이를 위해 독서능력 검사를 통해 도출된 검사지별 문항의 정답률을 산출하여 검사지별 ‘기대 점수(expected score)’와 학생 점수에 따라 독서지수 구간을 설정하였다.

여기서 ‘기대 점수’란 ‘중’ 수준의 검사지를 배정받은 학생이 얻을 것으로 예상되는 점수로서, 이는 검사 결과에서 도출된 학년군별 정답률에 근거하였다. 또한 검사지별 난이도에 따라 학생이 받은 점수에 일정한 가중치를 부여하였는데, 계산식은 다음과 같다.

계산식:

획득 점수 + (총 문항 수 × 문항별 배점 × (기준 정답률 - 난이도별 정답률))

* 기준 정답률: 해당 학년군의 '중' 수준 검사지의 정답률

* 난이도별 정답률: 해당 학년군의 난이도별 검사지('상, 중, 하')의 정답률

예를 들어, 중학교 1, 2학년 '중' 수준 문항의 정답률이 0.7이고, '상' 수준 문항의 정답률이 0.5일 때, 2단계에서 '상' 수준 검사지를 배정받아 총 4 문항의 정답을 맞혔을 경우(20점 획득), $20 + [8 \times 5 \times (0.7 - 0.5)]$ 로 계산되어, 최종 산출 점수는 28점이 된다.

또한 검사지별 기대 점수는 해당 학년군이 나타내는 독서지수의 중앙값과 대응하도록 하였으며, 점수 범위가 $\pm 5\%$ 씩 변할 때마다 ± 10 점의 독서지수가 변동하도록 하였다. 예를 들어, 어떤 독자가 해당 검사지의 기대 점수보다 40% 더 높은 점수를 얻을 경우에는 해당 학년 독서지수의 중앙값에서 80점이 더해진 독서지수를 자신의 적정 독서지수로 진단받게 되는 것이다. 한편, 점수 범위 $\pm 40\%$ 를 벗어나는 값(이상치)에 대해서는 추가적인 독서지수 변환을 제시하지 않았다. 즉, $+45\%$ 의 점수 범위를 가지는 독자 점수는 $+40\%$ 점수 범위를 가지는 것으로 변환된다.

이는 독서능력 검사 결과의 데이터 수가 충분하여 정규 분포를 이루며, 난이도에 따른 점수의 분산이 동일하다는 가정에서 출발한다. 이러한 방식은 독자가 최종적으로 배정받는 검사지의 유형과 상관없이 동일한 계산식으로 독자가 획득한 점수에 대해 가중치를 부여할 수 있다는 장점이 있다.

단계별 검사 결과에 따른 학생의 독서지수 변환 예시를 보이면 다음 <표 10>과 같다.

〈표 10〉 학생의 독서지수 변환 예시
 ((중 1, 2) 학년군 '공통 - 상 - 상' 검사지를 배정받은 학생)

공통		상	
세트 만점	50점	세트 만점	40점
정답률	0.7	정답률	0.5
기대 점수	35점	기대 점수	20점
학생 점수	45점	학생 점수	30점
가중 점수	0점	가중 점수	8점
상		최종 산출 점수	
세트 만점	40점	세트 만점	130점
정답률	0.5	정답률	-
기대 점수	20점	기대 점수	75점
학생 점수	25점	학생 점수	100점
가중 점수	8점	가중 점수	16점
		최종 점수	116점
최종 산출 은독지수 (중학교 1, 2학년) 독서지수의 중앙값 = $340 + 50(\text{가중 점수}) = 390$			

이러한 독서지수 변환 방식은 먼저 난이도별 정답률의 차이를 단순하고 투명하게 보정한다는 점에서 장점이 있다. 이는 추후 학생의 독서능력 검사 결과를 독서지수와 연계하는 방안을 학교 현장에서 사용하거나 실제 프로그램 등에 사용하기에 용이하다는 장점을 가진다. 다만 이러한 방식은 단순 정답률 차이만을 고려하는 선형 보정 방식이므로 문항 간 난이도 분포(편차)를 고려하지 않으며, 정답률 차이가 클 경우에는 점수가 과도보정 되거나 반대

로 과소보정 된다는 한계가 있다. 따라서 추후 MST 적용 시에는 이러한 지점을 충분히 고려하여 독서지수 연계 방안을 지속적으로 수정해 나갈 필요가 있다.

끝으로 본 연구의 결과를 바탕으로 학생의 독서지수에 따른 도서 추천의 결과를 예시하면 다음 <표 11>과 같다.

<표 11> 학생의 독서지수에 따른 추천 도서 예시(일부)

독서지수 (학년 범위)	추천 도서 예시(일부)
독서지수 290 (6학년 4수준~6학년 8수준)	『꽃 아주머니와 비밀의 방』, 『죽이고 싶은 아이』, 『신통방통 흡소핑』, 『고양이가 필요해』, 『당연하게도 너를』, 『이상한 몸 박물관』
독서지수 340 (7학년 7수준~8학년 1수준)	『어, 지금 땅 움직였지?』, 『1인 방송 시작하는 법』, 『우리의 정원』, 『누가 내 모습을 훔쳤을까』, 『어린이가 알아야 할 가짜 뉴스와 미디어 리터러시』
독서지수 390 (8학년 10수준~9학년 4수준)	『101 한국사』, 『해무기의 모든 것』, 『카페, 공장』, 『학교 가기 싫은 날』, 『해 구름 속의 아이』, 『인공지능은 선생님을 대신할까요?』, 『속눈썹, 혹은 잃어버린 잠을 찾는 방법』

V. 결론

본고는 지금까지 독서지수 시스템 개발 과정에서 부딪힐 수 있는 실제적 과제들을 다양하고 구체적인 사례를 기반으로 논의하였으며, 특히 ‘온독지수’라는 독서지수와 연계된 독서능력 검사를 설계하고 실행하여 MST용 독서능력 검사 문항을 확보하고, 독서지수와 독서지수의 합리적 연계 방안을 제안하였다.

본고의 의의를 상술하면 다음과 같다. 첫째, 본고는 실제 독서지수 시스템 개발 과정에서 만나게 되는 실질적 연구 과제들을 독서지수 개발을 위한

독서능력 검사, 독서지수와 도서지수의 연계 등을 중심으로 구체화했다는 의의를 지닌다. 독서지수 개발이라는 실제적 과제를 교육적 의의를 확보하면서도 합리적으로 해결하기 위해 미리 고려되어야 하는 실행상의 요소들과 그 해결 사례에 대해 깊이 있는 이해를 심화할 수 있을 것으로 기대된다. 본고가 독서지수 시스템을 개발하는 첫 단계에서 부딪히는 주요 과제에 대해 연구자들이 취할 수 있는 한 가지 우수 사례로 기록되기를 기대한다.

둘째, 본고는 도서지수와 연계된 독서지수를 개발하기 위해 독서능력 검사를 설계할 때 고려해야 하는 지점들을 구체화하였다. 도서지수의 특성을 고려하여 검사 영역과 하위 요소를 설계하고 문항을 개발할 때 어떤 점들을 고려해야 하는지, 지면의 한계상 충분히 기술하지는 못했지만, 최대한 구체화하고자 노력하였다. 본고를 통해 개발된 독서능력 검사 도구는 독자 수준별 맞춤형 독서 지원 체계의 교육적 실현 가능성을 뒷받침하는 근거 자료로서의 의미를 지닐 수 있을 것이다. 보다 직접적으로는 온독지수 기반의 독서지수 구축의 기반이 될 것이다.

셋째, 온라인 기반 MST용 수준별 문항을 확보하는 구체적인 방안을 제시하였다. MST 적용을 위해서는 수준별 문항을 다수 확보하여 문제 은행을 구축하는 것이 유리하다. 본고에서는 ‘상·중·하’ 수준에 맞는 문항을 개발한 뒤 대규모 검사 시행을 통해 문항의 난이도와 변별도를 확보하여 MST용 문항을 확보하는 방법을 제안하였다.

넷째, 본고는 온라인으로 학생들의 독서능력을 측정하는 한 가지 방법으로서 ‘3+1 단계’ 기반 MST를 설계·제안함으로써, 독서지수 검사의 한 가지 방향을 제시하였다. 본고에서는 MST를 적용함으로써 중위 수준 변별에 초점화된 일반적인 선형 검사보다 학생의 독서능력 수준 변별을 강화했을 뿐만 아니라, ‘3+1 단계’ 설정을 통해 양극단 학생(‘상상’ 수준의 영재나 ‘하하’ 수준의 학생)을 위한 보정 단계를 추가하는 새로운 체제를 도입하여 MST의 장점을 극대화하였다. 이러한 체제는 독서능력의 위계화에 기여할 뿐만 아니라, 학생의 독서능력 수준에 대한 정보를 더욱 세밀하게 제공함으로써

적절한 교육적 지원을 계획하는 데 직접적인 도움을 줄 수 있다. 또한 검사 결과가 지속적으로 누적되면 독서 영재 및 읽기 지연 진단 등 관련 교육 활동을 위한 실증적 데이터 확보가 가능해질 것이다.

다섯째, 본고는 독서지수와 도서지수를 연계하는 방안을 구체적으로 제시하였다. 본고는 대규모 학생을 대상으로 독서능력 검사를 시행한 뒤 그 결과를 분석하여 학생들의 독서능력 수준을 독서지수로 변환함으로써 이를 독서지수와 연계하고자 하였다. 특히 본고에서 제안한 방법은 이론적 타당성을 확보하면서도 비교적 간단하여 다양한 상황에 적용될 수 있을 것이다.

앞서 언급했듯이 독서지수와 도서지수의 연계는 독서지수 개발 과정에서 부딪히는 가장 어려운 난제 중의 하나이다. 국내에서는 아직 공개된 자료도 없고 관련한 깊이 있는 논의도 거의 없는 상황으로 판단된다. 이와 같은 상황에서 본고는 그 해결 방안의 일단을 구체적으로 제시하여 추후 관련 논의의 활성화에 크게 기여할 것으로 판단된다.

한편, 독서지수 시스템은 개발 이후 지속적인 업데이트와 수정 보완 작업이 필수적으로 요구되는 과제라고 판단된다. 대규모 독자들을 대상으로 지속적인 검사를 시행하고 검사 결과를 분석하여 검사 시스템의 수정과 보완이 꾸준히 이루어져야 한다.

본고의 논의 결과는 온독지수 시스템 구축에 실제로 적용될 예정이며, 현재 충남교육청에서 온라인 시스템 구축 작업을 진행하고 있다. 온라인 시스템 구축 작업이 완료된 뒤 다수의 학생들이 이 온독지수 시스템을 활용하게 된다면, 본고는 추후 다수 데이터의 확보를 통해 독서지수 측정 방법 및 도서지수와의 연계 방안을 지속적으로 수정, 보완해 나가고자 한다.

이와 관련하여, 독서지수 시스템 개발 시 미리 독서지수 시스템의 추후 수정 보완을 위한 데이터 확보 방안을 설계하는 것이 중요하다는 점을 강조해 두고자 한다. 독서지수 시스템을 개발할 때는 미리 향후 해당 시스템을 어떻게 수정, 보완해야 하는지 방향을 설정해 두어야 할 것이다. 그리고 수정, 보완을 위한 데이터는 어떻게 확보할지 등을 미리 계획하여 시스템을 구

축해야 한다. 기개발한 독서지수가 독자들에게 실제로 시행된 후 지속적으로 추적 관찰해야 하는 핵심 데이터가 무엇인지를 명확히 하고 이를 미리 정리해 두어야 추후 해당 데이터 확보를 위한 사전 설정 및 사후 관리가 가능할 것이다.

본고에서는 그와 관련하여 다음과 같은 몇 가지를 제안하고자 한다. 이는 특히 본고와 같이 온라인 기반 MST 설계를 선택한 경우에 해당한다. 이는 추가 연구 과제이기도 하다.

첫째, 가장 기초적이고 기본적인 데이터 확보에 해당하는 것으로서, 독자들 개개인에게 부여한 개별 독서지수의 누적적, 시계열적 기록 데이터의 확보가 필요하다. 독서지수 개발이 완료되고 많은 학생들이 이를 활용한다고 가정할 때에 학생들 개개인에게 부여한 독서지수의 개인적 누적 기록은 교육적, 학문적으로 매우 높은 의의를 지닐 것이다. 해당 데이터를 분석한다면 학생들의 독서능력 수준 발달과 수준별 양상에 대한 다양한 논의가 실증적으로 가능해질 것이다. 이때, 온독지수의 경우 온라인 기반 MST 시행으로 관련 데이터 확보가 용이한 한편, 지수 산출의 근거 등이 공개되어 있기 때문에 추후 관련 데이터의 분석과 해석 등 관련 논의의 활성화가 기대된다.

둘째, 학생들이 독서지수를 부여받은 후 어떤 책을 읽었는지에 대한 조사 데이터 확보가 필요하다. 이는 개발한 독서지수의 타당성을 검증하는 한 가지 방식이 될 수 있다. 만일 독자가 독서능력 검사를 통해 독서지수를 부여받은 후에 자신의 독서지수에 맞는 추천 책을 읽고 만족감을 느꼈다면, 해당 추천 시스템을 신뢰하는 경향이 강화될 것이다. 이는 해당 독서지수가 타당하게 설계되었다는 방증이 될 수 있다.

본고에서는 온독지수에 기반을 둔 독서지수를 제공한 후에, 학생의 독서 활동에 대한 조사를 통해 본 독서지수 부여 방식의 타당성과 효과성을 추후 확인 분석하고자 한다. 만일 본고에서 제안한 방식으로 이 검사가 지속적으로 시행되면서 추천 도서에 대한 학생 반응도 누적된다면 검사 결과의 환류에 대한 실천적 제언이 가능해질 것이다. 이를테면, 조사 결과를 바탕으로

학생의 읽기 수준에 가장 적절한 도서 추천 구간(Sweet Spot)에 관한 실증적이며 과학적인 데이터를 확보할 수 있을 것으로 기대된다. 또한 이러한 자료를 바탕으로 추천 도서 목록의 정교화도 꾀할 수 있을 것이다.

셋째, MST 경로 기록을 확보할 수 있도록 처음부터 시스템을 설계할 필요가 있다. 어떤 학생들이 어떤 경로로 검사를 치렀는지, 어떤 경우의 수가 가장 많은지, 가장 적은 경로는 무엇인지 등에 대한 데이터를 확보할 수 있도록 미리 처리해 두어야 한다. 해당 데이터는 온라인 기반 MST 시스템의 수정, 보완 시 유용하게 활용될 수 있다.

온라인 기반 MST로 독서능력 검사를 실시할 경우, 독서지수 시스템 개발 관련 핵심 과제는 단연 최다 문항 확보에 있을 것이다. 만약 동일 학생이 여러 번 재응시를 한다면, 문항에 대한 학습 효과가 발생하여 부정확한 결과를 낼 가능성을 배제할 수 없다. 거대 문제 은행을 구축하고 지속적으로 수정, 보완하는 것은 온라인 기반 MST 시행의 필수 조건으로 보인다.

우리나라의 교육적 현실을 고려할 때 독서지수 개발과 시행은 매우 시급하고 유의미하다. 현재 사기업 등과 달리 공공기관의 노력이나 학계의 개발 관련 논의는 다소 미약한 것이 사실이다. 독서지수 개발이 교육적 타당성을 확보하면서 실제적으로 이루어지기 위해서는 개발 과제들이 충분히 논의될 필요가 있다. 앞으로 실제 ‘개발’ 관련 논의들이 다양하고 깊이 있게 이루어지기를 희망한다.

* 본 논문은 2025.07.30. 투고되었으며, 2025.08.07. 심사가 시작되어 2025.09.08. 심사가 종료되었음.

참고문헌

- 김혜정·허모아(2021), 「청소년기 독서 수행 및 인식에 관한 실태 조사: 비독서의 이유와 스마트폰 사용 양상을 중심으로」 『독서연구』 59, 9-50.
- 박인용·김미림·박민호·곽민호·강혜진·박도영·김완수·임황규(2023), 『국가수준 학업성취도 검사의 단계적 적응형 검사 체제 구축: 검사 설계 및 척도화(연구보고 RRE 2023-8)』, 진천: 한국교육과정평가원.
- 박태준·김현정·오상철·홍미영·배화순·최소영(2024), 『KICE 이독성 지수(KICE Readability Index) 자동측정 프로그램 설계 및 개발(Ⅲ): 텍스트와 어휘 DB(초·중·고등학교급) 정교화 및 자동측정 프로그램 개발(RRT 2024-1)』, 진천: 한국교육과정평가원.
- 서민희·김미림·김원수·이소라·한정아·김성훈·서영숙(2020), 『컴퓨터 기반 국가수준 학업성취도 검사 도입을 위한 검사 설계 방안(연구보고 RRE 2020-3)』, 진천: 한국교육과정평가원.
- 서혁·이소라·류수경·오은하·윤희성·변경아·편지윤(2013), 「읽기(독서) 교육 체계화를 위한 텍스트 복잡도(Degree of Text Complexity) 상세화 연구 (2)」, 『국어교육학연구』 47, 253-290.
- 성태제(2019), 『알기 쉬운 통계분석』, 서울: 학지사.
- 신명선·강지영·나상수·이경남·정지호·홍경화·최소영(2024), 「교과서 기반 등급별 사고도구어 목록 개발 연구」, 『국어교육학연구』 59(1), 97-140.
- 신명선·이기연·나상수·인정남·차경미(2024), 「사고도구어 기반 독서지수 개발 연구: 가능성과 전망」, 『새국어교육』 140, 111-152.
- 조용구·이경남(2020), 「국어 텍스트 분석 프로그램(KReaD 지수)의 개발」, 『독서연구』 56, 222-245.
- (주)날말(2025), LQ 독서력(문해력) 검사, 날말사 홈페이지, 2025. 7. 9., <https://www.natmal.com/views/lq/intro/readtest>.
- Lexile(2025), Understanding Lexile & Quantile Measures: A Guide for Educators, 렉사일사 홈페이지, 검색일자 2025. 7. 9., 사이트 주소 <https://hub.lexile.com/for-educators/>.
- Renaissance(2025), Star Reading: Evidence-based K-12 reading assessments, 르네상스사 홈페이지, 검색일자 2025. 7. 9., 사이트 주소 <https://www.renaissance.com/products/star-reading/>.

독서지수 시스템 개발의 과제와 해결 방안에 대한 사례 연구

— 독서능력 검사와 독서지수의 연계에 관한 온독지수 시스템의 사례를
중심으로

차경미·최소영·이상재·인정남·신명선

본고는 기존에 개발된 사고도구어 기반 독서지수와 학생의 독서능력 지수, 즉 독서지수를 연계하는 방안을 마련한 연구이다. 이를 위해 학생의 독서능력을 독해 능력과 어휘 능력으로 나누고 각 요인별 하위 요소에 대한 검사 문항을 개발한 후 이를 ‘(상상)-상-중-하-(하하)’ 수준의 검사지 형태로 구현하였다. 다음으로 초·중·고등학교 학생 1,626명을 대상으로 검사를 실시하여, 검사지의 난이도에 따른 예상 정답률을 근거로 학생이 획득한 점수에 대한 선형 동등화 방식을 통해 학생의 독서지수를 산출하였다. 이렇게 산출된 독서지수를 독서지수와 연계하여, 학생의 수준에 따른 독서 활동이 이루어질 수 있도록 하였다. 이 연구는 학생의 독서지수를 검사할 수 있는 단계적 적응형 검사(MST) 모델을 개발하고, 이러한 검사 결과를 독서지수와 연계하는 방안을 마련하였다는 점에서 의의를 가진다.

핵심어 독서능력 검사, 어휘 능력 검사, 독서지수, 독서지수, 읽기 수준

A Case Study on the Challenges and Solutions in Developing a Reading Index System

— Linking Reading Ability Tests and Book Index Focusing on O.R.
Index System

Cha Kyungmi · Choi Soyoung ·
Rlee Sangjae · In Jungnam · Shin Myungsun

This study aims to establish a link between a previously-developed Book Index based on academic vocabulary and students' Reading Index. To achieve this, the students' reading abilities were divided into two main components: reading comprehension and vocabulary. Assessment items were developed for each subcomponent and organized into a five-level test form. A preliminary survey was conducted with 1,626 elementary, middle, and high school students. Based on the expected accuracy rates corresponding to item difficulty, a linear equating method was applied to calculate each student's Reading Index. This index was then linked the Book Index to facilitate reading activities appropriate to each student's individual level. The significance of this study lies in its development of a stepwise, adaptive assessment model for evaluating students' Reading Index. It also proposes a method to connect assessment outcomes with the Book Index.

KEYWORDS Reading Ability Test, Vocabulary Ability Test, Reading Index, Book Index, Reading level