

사회문화주의 패러다임의 읽기 성취도 평가 설계를 위한 제언

— 2026 NAEP 읽기 평가 체계의 분석 및 함의

김가은 한양대학교 러닝사이언스학과 박사과정(제1저자)

이채운 한양대학교 러닝사이언스학과 박사수료(공동저자)

조병영 한양대학교 국어교육과 교수(교신저자)

* 이 논문은 한양대학교 교내연구지원사업(HY-202100000003540)에 의해 연구되었다.

- I. 연구의 목적
- II. 이론적 배경
- III. NAEP 읽기 평가 체계 분석
- IV. 논의 및 제언

I. 연구의 목적

이 연구는 읽기의 사회문화주의적 관점에 타당한 평가의 개념과 원리를 재편하고, 미국 연방 정부의 ‘2026년 시행 국가교육발전평가 읽기 평가 체계(Reading Framework for the 2026 National Assessment of Educational Progress, NAEP)’의 개발 과정과 내용을 검토함으로써, 미래 세대를 위한 한국의 읽기 평가 개발 방향에 대한 시사점을 도출하는 것에 목적이 있다. 이를 위해 읽기와 학습, 성취와 평가를 바라보는 사회문화주의적 패러다임을 통해서 어떻게 2026 NAEP 읽기 평가가 한층 발전된 교육평가 및 타당도 이론을 반영하는지 분석하고, 향후 한국에서 대단위 읽기 성취도 평가를 설계하는 데 요구되는 기초 논의를 제공하고자 한다.

교육평가는 실질적으로 교육 현장에 가장 심대한 영향을 미치는 교육 행위이다. 또한 교육평가는 학생의 학업 성취 정도를 정밀하게 파악하여 학교 수업 및 교육 실재를 개선하기 위한 교육 정책 마련의 근거를 제공한다. 이 때문에 타당하고 공정하며 학생들의 다양한 언어적·문화적 배경을 포용적으로 고려하여 학습자의 수행 능력을 최선으로 측정할 수 있는 평가에 대

한 관심이 고조되어 왔다(박혜영, 2019; 오규설, 2025; 최미숙, 2004; Afflerbach, 2025; Bennet, 2013) 특히 한 국가의 학생들이 어느 정도의 학업 성취를 보이는지에 관해 정확하게 측정하고 해석하기 위해서는 과학적으로 설계된 국가 수준의 표준화 평가 체계가 반드시 필요하다. 이러한 요구에 부응하여, 지금까지 과학적 요건을 갖춘 대단위 교육평가의 질적 제고와 설계를 위한 논의들은 이미 교육측정 및 학습 과학 분야에서 상당한 정도로 연구가 진척되어 왔다(Mislevy, 2018; Glaser, Chudowski, & Pellegrino, 2001).

이러한 학문적 성취와 더불어, 현대 사회가 점점 더 다양하고 복잡한 읽기 능력을 요구하는 만큼 학생들의 읽기 역량을 정확하게 측정하는 평가 역시, 복잡한 사회문화적 문해 환경과 맥락을 고려하여 과학적으로 설계될 필요가 있다. 특히 최근의 읽기 환경이 급속하게 디지털적으로 전환됨에 따라 실제적인 읽기 자료와 과제를 평가 과정에서 적극적으로 고려해야 한다는 요구가 크다(조병영, 2020; Leu, Kinzer, Coiro, Castek, & Henry, 2017; Sabatini, O'Reilly, Weeks, & Wang, 2020) 또한 문화적 다원화와 역동적인 사회 변화 속에서 독자의 다양성이 현저하게 증대됨에 따라 학습자의 다양한 지적 자원과 언어적 배경, 문화적 경험 등의 환경적 요인이 평가 수행에 영향을 미칠 수 있음을 이해하고, 이를 평가 설계에 적극 반영해야 한다는 점 또한 더욱 강하게 조명되고 있다(Basterra, Trumbull, & Solano-Flores, 2011). 삶의 다양한 국면에서 실질적 학습에 영향을 미치는 다양한 문해 주체, 텍스트, 활동, 맥락 요인을 평가 설계에서 다면적이면서도 유기적으로 고려할 수 있을 때(RAND, 2002), 더욱 타당하고 공정하며 포용적인 평가를 개발하고 실행할 수 있을 것이다.

교육평가에 관한 심층적이고 다원적인 논의는 대단위 평가 문화에 내재된 후기실증주의적 관점(post-positive perspectives)을 사회문화주의적 관점(sociocultural perspectives)으로 전환할 것을 요청하고 있다(Gipps, 2002; Mislevy, 2018). 사회문화주의적 관점에서 학습자의 언어 능력은 사회문화적 맥락에 따라 형성되고 규정되며, 이를 측정하는 상황적 맥락과 도

구 및 자원의 관점에 따라 달라질 수 있다는 점을 강조한다. 마찬가지로 이러한 학문적 패러다임의 전환에 따라 국어과의 인식론 역시 교육과정 및 교과서 구현 과정에서 학습자 개개인의 언어 경험이 각자가 속한 사회문화적 배경에 따라 다르게 형성될 수 있음을 명시하거나, 교수학습 설계 및 실행 과정에서 교실을 구성하는 상황적 맥락이 학습자 개개인의 특성과 상호 작용하여 다층적인 국면을 만들 수 있음을 고려하는 등 언어, 학습, 성취에 관한 전반적인 관점 전환을 적극적으로 고려하고 있다(서혁·박지윤, 2009; 신명선, 2002; 조병영, 2020). 그럼에도 불구하고 학생들의 실제적 삶과 정책 등에 많은 영향을 미치는 대단위 평가는 여전히 최대한 맥락적 요인이 배제된 채, 학생들이 수행할 것이라 기대되는 관찰 가능한 능력의 검증에 치중하여 설계되는 실정이다.

대단위 평가는 학습자의 역량을 특정 기준에 따라 측정하고 이를 학습자의 보편적인 역량으로 간주함으로써 학업 성취를 비롯한 학습의 결과를 표준화하여 비교하고 판단하는 것을 목적으로 한다. 그러나 학습자마다 다르게 가지고 있을 사회문화적 지식과 문화 자산 및 그들의 일상 세계 경험과 맥락을 모두 통제하여 완전무결한 성취도 결과를 보여줄 수는 없다(Afflerbach, 2005). 오히려 이른바 객관적 신뢰도를 최소한도로 희생하면서도, 평가의 과정과 결과에 작용하는 학습자의 사회문화적 맥락이 평가의 설계 원리와 구현 방식, 결과의 해석에 고려될 때 학습자마다 가지고 있는 지적 자산의 다양성을 고려한, 보다 타당하고 유용한 정보를 제공할 수 있는 평가를 개발할 수 있을 것이다.

이렇듯 교육평가에 대한 전환적 관점이 요구되는 시점에서 사회문화적 패러다임을 기반으로 한 대단위 평가에 관해 정치한 논의가 요구되지만, 이에 관한 과학적·분석적 연구는 아직 드물다. 현대 사회의 맥락에 따라 충분한 타당성과 신뢰성을 확보하지 못한 평가는 현시점에서 요구되는 학습자 역량을 제대로 측정하기도 어렵고, 또한 평가 정보를 요청하는 다양한 교육 주체들과도 효과적으로 소통하기 어렵다. 더욱이 타당하지 못한 평가를 통

해 산출된 데이터를 가지고서 교육의 현실태를 진단하고 학습자 성장을 위한 건실한 교육 정책을 마련하기는 더욱 어려울 것이다. 아울러, 과학적 근간이 약하고 시대적 요구에 부응하지 못하는 학업성취도 평가는 전반적으로 교육평가에 대해 이미 진일보한 관점과 요구를 표출하고 있는 교육 주체의 불신을 키우는 데 기여하고 있기도 하다(Emler, 2019; Hargreaves, 2020; Lai, 2023, Martens, Niemann, & Teltemann, 2016). 따라서 현 시점에서 가장 발전된 형태의 사회문화주의적 관점을 기반으로 설계된 국가수준 평가 체계인 2026 NAEP Reading Framework(NAGB, 2022)를 면밀하게 분석함으로써, 향후 한국에서 진행될 논의를 증진할 수 있는 기초 지식을 제공하고자 한다.

이에 이 연구는 다음과 같은 목표와 내용을 포함한다.

- 읽기 평가의 개발과 설계에 관한 교육측정 및 학습과학 분야의 선행 연구를 검토하고, 사회문화주의를 기반으로 하는 타당도 개념을 재구성한다.
- 2026 NAEP Reading Framework가 어떻게 사회문화주의적 관점의 타당도 요구를 반영하고 있는지 개발과 설계의 측면에서 분석한다.
- 2026 NAEP Reading Framework의 설계 원리를 바탕으로 우리나라의 맥락에서 대단위 읽기 평가를 개발할 때 고려해야 할 시사점을 구체적으로 도출한다.

II. 이론적 배경

1. 읽기의 사회문화주의적 관점

읽기는 사회적 관계, 문화적 모형, 권력과 정치, 경험에 관한 관점, 가치

와 태도 등 읽기 주체가 살아가는 세계에 결속된 행위로(Gee, 1989), 결코 탈맥락적(decontextualized)일 수 없다. 읽기라는 인간 행위는 특정한 사회 문화적 맥락 안에서 관습화된 방식으로 일어나는데, 이는 독자가 어떤 사회 문화적 집단에 속해 있는가에 따라 각기 다르게 쌓아 온 배경지식 및 경험이 기호 해독 및 표상의 형성에 복합적으로 영향을 미친다는 것을 의미한다. 이러한 관점에서 볼 때, 읽기는 정보 처리나 의미 구성에서 나아간, 하나의 사회적 실천(social practice)이다.

사회문화주의적 관점을 리터러시(literacy) 현상에 적용한 대표적인 연구자인 Street는 생활세계의 리터러시가 다양한 목적을 기반으로 저마다 달리 벌어지는 방식에 주목하여, 텍스트를 기반으로 읽고 쓰고 대화하는 과정은 반드시 특정 맥락을 기반으로 하는 일련의 ‘실천 행위’라는 점을 조명하였다(Street, 1984). 읽기는 그 행위가 일어나는 사회의 문화적이고 권력적인 구조에서 별도로 추출할 수 없는, 고도로 맥락적인 실천인 것이다. 읽기를 사회문화적인 실천으로 이해하는 관점은 시공간에 따른 다양한 리터러시 사건(literacy event)에 관한 연구 및 ‘누구의 리터러시’가 지배적이거나 주변화되는가에 관한 리터러시 실천(literacy practice)에 관한 연구 등으로 전개되었다(Barton & Hamilton, 1998; Barton, Hamilton, & Ivanic, 2000; Duke, & Martineau, 2007; Gee, 2018; Luke, 2003, Luke, 2004; Purcell-Gates, 1997; Purcell-Gates, Purcell-Gates, Perry, & Briseño, 2011; Street, 1984). 즉, 우리 사회에 존재하는 리터러시는 단일한 구인이 아니라 여러 모습으로 존재하는데, 일부만이 사회문화에 의해 제도로서 인정받으며 하나의 리터러시 양식은 상황적 맥락에 따라 좌우된다는 것이다.

2. 읽기 평가의 사회문화적 요구

읽기를 사회문화적 실천으로 정의할 때, 우리가 그러한 읽기를 가르치고 평가하는 방식 역시 사회문화적이어야 할 것이다(Benavot, Cha,

Kamens, Meyer, & Wong, 1991; Street & Street, 1995). 이와 관련하여, 사회문화주의적 관점에서의 읽기 이해 모형을 개발하고 평가 설계에 적용하고자 한 RAND(2002)는 독해가 텍스트(text), 독자(reader), 활동(activity)이라는 세 요소의 상호적 관계로 이루어지며, 이는 사회문화적 맥락(sociocultural context) 안에 존재하면서 동시에 이 맥락 역시 반복적으로 다른 요소와 상호작용한다는 점을 설명한다. 이 모형에서 사회문화적 맥락은 텍스트를 통해 학습이 이루어지는 구체적인 과제 환경을 시사하면서도, 특정한 문화 집단이나 담화 공동체가 세계를 드러내고 정보를 전송하는 추상적인 환경도 포함한다(RAND, 2002). 평가 설계의 과정에서 이러한 맥락을 고려한다는 것은 실생활에서의 읽기 능력을 측정하고자 하는 읽기 평가의 본질적 목표에 이론적인 정합성을 부여하는 초석이 된다.

사회문화주의적 패러다임을 고려할 때, 읽기 평가의 ‘결과’에 대한 접근 역시 평가가 위치한 사회문화적 맥락을 고려해야 한다. 리터러시 실천은 사회 제도 및 역학 관계에 의해 유형화되며 이러한 주체에 의해 실천되는 특정한 리터러시는 다른 리터러시에 비하여 더 지배적이고 가시적이며 강력하게 영향을 미친다는 점을 고려할 때, 교육 행위의 실천으로서 중요한 지위를 점한 대단위 평가에서 기반으로 삼아야 할 읽기에 관한 관점이 사회문화적이어야 함은 자명하다. 특히 대단위 표준화 읽기 평가는 사회적이고 정치적인 행위라는 점(Moss, Pullin, Gee, & Haertel, 2005)에서 사회문화적으로 타당한 평가의 필요성은 강조된다.

지금까지 대단위 평가에서 읽기는 주로 ‘인지적인 것’ 혹은 ‘구성적인 것’으로 정의되어 왔다. 이러한 평가에서 측정된 읽기 역량은 실생활의 읽기와 유리될 수 있는 것이었으며, 이러한 평가에서 높은 점수를 받은 학생이 실생활에서도 주체적인 목적을 설정하고 적절한 텍스트를 선택하여 추론적이고 비판적인 방식으로 의미를 구성할 수 있는가는 미궁 속에 놓이게 된다. 평가를 통해 파악하고자 하는 것이 지정된 텍스트에 대한 숙달도가 아니라 실제 삶에서 학습자가 얼마나 해당 역량을 잘 수행할 수 있는가에 관한 것이

라면, 맥락을 고려하지 않은 읽기 평가가 충분한 타당도를 확보할 수 없다는 점을 마땅히 추론할 수 있다.

3. 사회문화적으로 타당한 평가

그렇다면 사회문화적인 맥락을 고려한 타당도는 어떠한가? 적절하게 설계하고 개발한 표준화 읽기 평가는 학습자의 실제적인 읽기 역량을 과학적이고 구체적으로 진단하여 학습자 개인에게 읽기 역량 향상에 관한 방향을 제안할 수 있으며, 정책 입안자 및 교육 관계자에게도 교수학습 실천 지원 방안에 관한 근거를 제공할 수 있다. 이를 위해서는 학생들의 실제 읽기 역량을 적절하고 정확하며 공정하게 측정할 수 있도록, 실제 읽기 역량에 영향을 미치는 요인들을 평가 체계에 포함하여 사회문화적으로 타당한 평가를 설계·개발해야 한다.

읽기 평가에 참여하는 학생은 제시된 텍스트를 읽고 과제를 수행하는 과정에서 텍스트에 관해 이해한 내용을 바탕으로 정신적 표상을 구성한다. 하지만 피험자인 학생이 이러한 표상을 얼마나 응집적으로 구성했는지 파악하기 위해서는 제시된 선택지 중 자신의 이해와 가장 유사한 것을 고르도록 하거나, 글로 써보게 하는 등의 과정이 반드시 수반된다. 이 과정에서 개인의 읽기 능력과 무관한 것이 평가 결과에 영향을 주지 않도록 하면서도, 개인의 읽기 역량을 활용할 수 있는 내용 요소를 문항으로 구안하는 것이 그간의 읽기 평가의 타당도 확보에 있어 가장 핵심적인 과제로 논의되었다.

평가의 타당도는 평가가 측정하려는 능력을 정합적으로 측정하고 있는가, 즉 경험적 증거와 이론적 근거가 시험 점수에 기반한 추론 및 행동의 적절성과 정확성을 뒷받침하는 정도에 대한 통합적 판단(Messick, 1990)을 의미한다. 전통적인 평가 이론에서는 평가 타당도를 내용 타당도, 준거 기반 타당도, 구인 타당도로 나누었으나(Reynolds et al., 2010), 타당도는 일원적 개념(unitary concept)이라는 논의(Messick, 1990)에 따라, 평가 타당도 논의

는 구인 타당도(construct validity)를 중심으로 발전되었다. 평가가 측정하고자 하는 추상적인 역량이 구체적인 구인을 통해 충분하고 적절하게 표상되는지, 그것이 적절한 과제를 통해 구현되고 있는지가 타당도의 핵심 관점이 된 것이다. 이러한 관점은 읽기라는 관측 불가능한 사건을 정치하게 추론할 수 있게 하는 이론적 틀에 대한 필요성을 요구한다. 학습자의 실제적 읽기 역량을 측정하고자 한다면, 읽기라는 복잡한 인간의 인지적·정의적 참여 과정에 개입되는 다양한 사회문화적 요인을 고려하는 방식으로 그 구인의 설정이 이루어져야 한다.

더불어 답지 반응 방식을 안내하는 지시적 과정이나 과도한 긴장 및 과제 몰입 정도 등 피험자의 개별적 특성과 같은 평가 외부적 맥락은 읽기 과정 및 결과에 영향을 미쳐 평가 결과의 해석에 난점을 초래한다. 읽기 역량은 학생 개인이 가지고 있는 배경지식의 종류와 양, 제시된 텍스트 및 과제에 몰입하는 정도 등에 의해 크게 달라질 수 있다는 점에서, 읽기 평가의 외부적 맥락의 통제에 관한 엄중한 논의가 요구된다.

한편, 사회문화적으로 타당한 평가에서는 구인의 설계와 구현이 타당하더라도 그 결과가 특정 집단에 유리한 방식으로 편향될 수 있다는 점을 고려해야 한다. 차별적 타당도(differential validity) 가설은 한 평가가 특정 그룹의 개인에게 더 정확한 구인을 예측하도록 설계될 수 있음(Reynolds, 2000)을 강조하며, 타당도 개념에 관한 비판적 확장을 제안한다. 이러한 관점은 평가가 실제 상황 속에서 어떻게 사용되고 어떤 의미를 가지는지에 주목하여 타당도를 판단해야 한다는 논의(Moss, Girard, & Haniford, 2006; Moss, 2003, 2007)와 마찬가지로 사회문화적 역학을 고려하는 관점에 해당한다. 특히 평가는 맥락과 무관하게 독립적으로 존재할 수 없으며 현존하는 모든 시험은 어떤 방식으로든 특정한 문화를 반영한다는 점에서 어느 정도의 문화적 부하를 가지게 되나(Reynolds et al., 2010), 그 부하의 정도가 편향되지 않도록 유의해야 한다. 특정 집단에게만 생소한 내용의 지문을 구성하거나 지적 자산의 다양성을 고려하지 않는 등 평가가 특정 집단의 성취를 저평가

하거나 고평가할 때 그 시험의 타당도는 위협받는다.

평가가 교육의 실제적 맥락에 어떤 영향을 미치는가에 관한 결과 타당도(consequential validity) 역시 대단위 평가에서 고려해야 할 주요한 지점이다(조병영, 2011). 대단위 평가의 시행은 그 결과가 교육 정책 및 지원 등 실제적 삶에 중핵적으로 관련된다는 점에서 매우 권력 역학적이다. 가령 평가 결과 농촌 지역 학생의 성취가 낮다면, 그러한 교육적 현상에 영향을 미쳤을 수 있는 교육 기회와 자원의 차이 등을 결과의 해석에 투영해야 발전적이고 타당한 정책과 교수학습 지원으로 나아갈 수 있다.

타당도에 관한 사회문화주의적 관점은 대단위 표준화 읽기 평가의 개발 및 설계에 있어 세밀한 조정을 요구한다. 2026 NAEP는 변화하는 다양성 시대에 학습자의 읽기 수행을 타당하고 공정하게 측정하기 위해 개정되었다. NAEP는 미국의 국가교육평가위원회(National Assessment Governing Board, NAGB)가 주도하여 4, 8, 12학년 학생들의 학업성취도를 표집하여 측정하는 국가수준 성취도 평가로, 국가의 주요 교육정책을 분석하고 결정하는 데에 중추 역할을 하는 기초 데이터 및 학교급과 학습자 배경 등에 따른 다양한 성취 지표를 제공한다. 2026 NAEP의 읽기 평가는 현시점의 국제적 관점에서 현대 교육측정 및 학습과학의 이론, 사회문화적 요구를 적극적으로 반영하고자 한 학업성취도 평가이다. 따라서 2026 NAEP Reading Framework(NAGB, 2022)를 사회문화적 관점에서의 타당도 논의를 바탕으로 분석하고 검토하여 구체태를 확인하는 것은, 향후 한국의 맥락에 어울리는 표준화 학업성취도 평가 개발의 방향에 관한 실제적 논의에 기틀을 마련할 것이다.

III. 2026 NAEP 읽기 평가 체계 분석

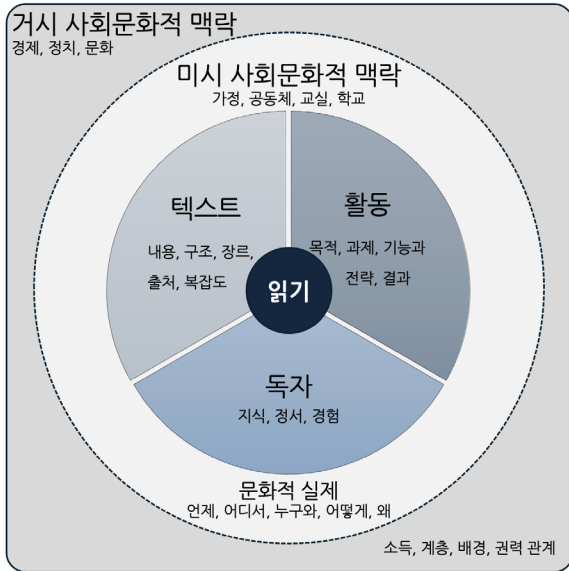
1. 읽기 평가의 사회문화적 구인 모형 설정

사회문화적으로 타당한 읽기 평가를 위해서는 측정의 대상으로 삼는 역량의 구인을 사회문화적 관점에서 설정해야 한다. 읽기의 패러다임은 사회문화적 구성주의로 이동하였으나, 읽기 평가에서는 후기 실증주의적 관점을 바탕으로 하는 패러다임에 충분한 변화가 일어나지 못했다. 하지만 인간의 읽기 과정은 본질적으로 다원적이며 상황 맥락적인바, 2026 NAEP에서는 사회문화적 읽기 모형을 평가들의 바탕으로 삼음으로써 읽기를 “사람의 의미 구성 활동으로서, 맥락 속에서 이루어지는 목적 주도의 활동”이라고 정의한다(NAGB, 2022: 17).

읽기 이해는 텍스트로부터 의미를 만들어내는 과정으로, 다양한 요인의 영향을 받는 복잡한 활동이다. 그 요인에는 다음과 같은 독자의 능력이 포함된다.

- 인쇄물과 복합양식적 텍스트에 참여할 수 있는 능력
- 기초적인 읽기 기술, 언어 지식, 동기와 같은 개인적 자원을 활용하는 능력
- 사회적·문화적 맥락의 범위에 걸친 활동에서 의미를 뽑아내고, 구성하고, 통합하고, 비평하고, 적용하는 능력(NAGB, 2022: 10)

2026 NAEP Reading Framework에서는 ‘읽기 이해’의 하위 구인을 텍스트, 독자, 활동, 맥락으로 정의한다. 이러한 구인은 RAND(2002)의 상호작용적 읽기 학습 논의를 확장한 것으로 볼 수 있는데, 이는 학습이라는 인간의 행위가 사회적 상호작용에 의해 형성되며 문화적 실천의 다양한 측면에 의해 매개된다는 사회문화적 관점(Nasir & Hand, 2006)을 명시한다는 점에서 주목할 만하다. 2026 NAEP에서 확장적으로 정의한 구인을 종합하면 <그림 1>과 같다.



〈그림 1〉 사회문화주의적 읽기 이해 모형

‘맥락’은 읽기 행위가 이루어지는 시공간적 배경뿐 아니라, 독자가 살아 가며 체득하는 다층적인 사회문화적 배경을 포괄한다. 교실, 학교, 가정, 지역공동체 등 읽기 실천에 즉각적 영향을 미치는 문화적 실제들은 ‘미시 사회문화적 맥락’으로서 읽기 실천이라는 현상을 밀접하게 둘러싸고 있다. 한편 독자들은 사회적 상호작용을 통해 리터러시를 경험하는데, 이때 독자가 어떤 문화적 집단이나 담화 공동체에 속해 있는가에 따라 정보를 수용하고 세상을 해석하는 방법이 달라진다. 소득, 종교, 언어, 신념, 이념 등 독자가 어떤 사회문화적 집단과 밀접한가에 따라 독자는 다른 리터러시를 경험하고 다른 역량이 발달한다. 정치·경제·사회적 요인과 같이 독자의 읽기를 비가시적으로 둘러싸고 있지만 강력한 영향을 미치는 조건들은 ‘거시 사회문화적 맥락’으로서 모든 읽기 이해 과정을 감싸고 있다.

맥락에 내재한 ‘독자-텍스트-활동’은 적극적으로 영향을 주고받으며 복합적으로 상호작용한다. ‘독자’는 그들이 마주하는 모든 텍스트와 상호작

용하는 인지적·메타인지적·문화적·동기적·언어적 자원을 활용하는 능동적 존재로서, 자신이 속한 공동체에서 가용한 지식 자원을 통해 읽기 경험을 재구성한다. 이때 ‘텍스트’는 독자와 소통하고자 하는 목적을 가진 저자의 구성물로 정의되며, 텍스트의 복잡도, 양식, 형식 등과 같은 요소는 독자가 텍스트에 관한 정신적 표상을 구성하는 데 영향을 미친다. 이렇게 ‘독자-텍스트’의 교섭적 작용으로 인해 독자마다 경험하는 읽기 ‘활동’은 그 동기와 목적, 사전 경험 등에 따라 개별화되고 다양화된다. 읽기 ‘활동’의 과정에는 텍스트를 주의 깊게 읽으면서 정보를 찾고 회상하며(locate and recall), 내용을 통합하고 해석하고(integrate and interpret), 텍스트를 분석하고 평가하거나(analyze and evaluate), 이를 통해 구성한 의미를 활용하고 적용하는(use and apply) 인지적 행위들이 포함된다. 이 과정에서 요구되는 기호 해독 및 언어적·의미적·메타인지적 텍스트 처리는 독자가 특정 주제에 관한 지식을 증진하고, 실세계 문제에 대한 해결책을 구상하고, 텍스트에 깊이 몰입하는 등 다양한 유무형의 읽기 결과에 영향을 미친다.

이러한 읽기의 네 가지 인지적 목표(탐색 및 회상, 통합 및 해석, 분석 및 평가, 활용 및 적용)는 사회문화적 읽기 모형의 구성 요인에 반영되어 평가 문항으로 구체화된다. ‘맥락 요인은 읽기를 평가의 안팎에 존재하는 요인에 의해 영향을 받는 결과로 이해하는 관점을 문항 설계에 반영할 것을 제안하며, ‘독자’ 요인은 텍스트의 아이디어를 확인하고 적용하도록 요구하는 문항을 통해 축자적·추론적·분석적·비판적 반응과 관련된 일련의 기능과 전략을 다룰 것을 요청한다. ‘텍스트’ 요인은 글과 자료의 부분적 혹은 전체적 특징과 의미를 평가 설계에서 고려할 것을 제안하며, ‘활동’ 요인은 네 가지 이해 목표를 어떻게 평가에 반영할지 결정하기 위해 학문적 맥락, 목적, 텍스트, 과제의 형식, 구조, 본질에 주의를 기울여야 함을 구현한다.

요컨대 2026 NAEP에서는 읽기 이해에 관한 사회문화적 관점을 채택하여 구인 모형을 구성함으로써, 기존 평가에서 읽기를 인지적인 것으로 추정하여 피할 수 없었던 구인 과소표상(construct underrepresentation)의 문

제를 극복하고자 하였다. 또한 보다 과학적 근거에 기반하고 사회적으로 정
의로운 교육 실천을 함의하는 사회문화적 관점의 도입함으로써 대단위 교육
평가의 내용적 충실도와 이론적 타당도를 개선하고자 했다는 점에서 혁신적
평가 체계로 충분히 주목할 만하다.

2. 학문 및 교과 영역에 따른 맥락적 읽기 목적과 텍스트 구성

2026 NAEP는 학생 독자의 읽기 능력을 평가하기 위해 ‘읽기 이해’를
구인으로 두고, 이 정의에 입각한 문항 구성의 기준점을 읽기의 ‘맥락’과 ‘목
적’이라는 두 개의 큰 축으로 제안한다. 여기서 학문 맥락은 문학, 과학, 사회
라는 교과 혹은 지식 영역으로 대별되며, 목적은 이해(reading to develop
understanding, RDU)와 문제해결(reading to solve the problem, RSP)로
구분된다. 이때 이해는 학생들이 텍스트를 주의 깊게 분석하고, 텍스트 안
팸의 정보를 깊이 이해하는 읽기를, 문제해결은 어떤 실제적 문제를 해결하
기 위해 학생들이 여러 개의 텍스트를 읽고 다양한 관점으로 정보를 활용하
도록 하는 읽기를 의미한다. 읽기 맥락은 독자가 읽기 행위에 가지고 들어가
는 목적과 방법을 선택하게 한다는 점(Rouet, Britt, & Durik, 2017)에서, 읽
기의 목적이 학문 맥락 안에서 구체화된다는 점은, 2026 NAEP가 읽기 평가
구인의 사회문화적 타당성을 고려하고 있음을 의미한다.

평가에서 학문 맥락을 고려할 때 텍스트는 매우 중요한 위치에 놓인다.
텍스트는 기본적으로 읽기의 대상이자 문항의 제재이며, 각 학문의 공동체가
어떻게 읽고 쓰는가를 가장 뚜렷하게 드러내는 지표와도 같기 때문이다. 2026
NAEP는 학문 맥락에 따른 대표적인 텍스트 유형을 제시한다. 문학 맥락에는
소설·드라마·영화·시·문학 관련 논픽션 등이, 과학 맥락에는 설명문·논설
문·보고서 등이, 사회 맥락에는 해설서·논설문·역사적 의미가 있는 문서 등
이 해당한다. 텍스트의 구체적 자질은 담화, 언어 구조, 장르 요소와 함께 설명
되는데(NAGB, 2022: 73-74), 이러한 텍스트는 이미지를 동반한 글이나 비디

오를 동반한 하이퍼텍스트와 같이 복합양식의 형태를 포함하며, 학교 안팎의 텍스트를 모두 포괄한다. 이는 같은 학문 영역의 맥락 안에서도 학교의 공식적 수업에서 활용되는 텍스트와 학교 밖 생활 속에서 쓰이는 텍스트가 다를 수 있음을 전제하고, 평가 상황에서 학생이 이를 다양하게 경험하도록 한다는 점에서 눈여겨볼 만하다. 이러한 텍스트는 읽기의 사회문화적 맥락을 반영하며, 구체적 문항의 활동을 통해 독자와 상호작용함으로써 실제적 읽기 활동과 최대한 유사한 경험을 평가 속에서 구현하도록 기능한다.

〈표 1〉은 2026 NAEP Reading Framework에서 제시한 학문 맥락에 따라 어떻게 평가 문항의 읽기 목적이 구체화되고, 평가 제재가 다양하게 구성될 수 있는지를 분석하여 종합한 것이다.¹⁾

먼저, 문학 맥락에서 독자는 시, 서사, 비평과 같은 문학 텍스트를 통해 인간 삶에 관한 주제를 성찰하고 분석하며, 텍스트가 자신의 삶에 어떤 통찰을 주는지 탐색하게 된다. 특히 문학 텍스트는 언어, 수사적 장치, 구조의 보고로서 독자는 자기가 읽은 작품을 다른 작품과 연결하여 문학사의 흐름 속에서 이해하며 자신이 다루는 문제를 현대적 또는 역사적 맥락에 위치시킨다. 문학 영역의 문항은 학교에서 시나 소설을 읽고 토론하는 상황이나, 학교 밖에서 팬픽션과 서평을 창작하거나 SNS에서 작가를 팔로우하고 토론하는 상황을 통해 제시된다.

다음으로, 과학 맥락에서 독자는 자연 세계에 대한 설명, 보고, 주장을 담은 텍스트를 읽고 과학의 개념 및 현상의 원리를 파악하거나, 직접 관찰한 내용을 보고서, 관찰 기록지, 논문 등으로 작성하게 된다. 이 과정은 텍스트를 읽는 능력뿐만 아니라 연구 결과를 도출하고 근거를 들어 결과를 서술하는 방법도 요구한다. 과학 영역의 문항은 학교에서 설명문, 보고서 등의 과학

1) 〈표 1〉에서 ‘읽기 목적’은 보고서 81-82쪽의 내용을, ‘텍스트와 관련된 학생 맥락’ 17-18쪽의 내용을, ‘장르와 텍스트 유형’ 및 ‘담화, 언어 구조, 텍스트 요소’는 73-74쪽의 표를 그대로 가져온 것이다.

〈표 1〉 2026 NAEP Reading Framework의 학문 맥락에 따른 읽기 목적과 텍스트 유형 및 지질 분석 결과

학문 맥락	읽기 목적	텍스트와 관련된 학생 맥락	장르와 텍스트 유형	언어, 언어 구조, 텍스트 요소
문학	-인간의 경험을 이해하기 위해 -다른 사람들을 즐겁게 하기 위해 -개인적이고 사회적인 딜레마를 성찰하고 해결하기 위해 -작가의 작품을 이해하고 사용하여 해석을 발전시키기 위해	〈학교〉 시, 단편 소설, 웹툰, 장편 소설, 영화와 같은 문학 텍스트를 읽고, 창작하며, 이에 대해 토론함. 〈학교 밖〉 독서 동아리에 참여하고, 팬픽션과 서평을 창작하며, 작가를 팔로우하고 토론하고, 애니메이션과 음악으로 문학 작품을 각색하거나 그 외의 것들을 만들.	<div>내러티브</div> -신화, 전설, 우화 -장편 소설 -풍자 문학 -공상 과학 소설 -마술적 사실주의 소설 -판타지 -만화책 (comic book) -만화의 형태로 된 소설 -만화 (manga) -팬픽	-플롯 유형 -인물 유형 -서사적 요소들 -비유적 언어 -관점 -주제 -주변, 대화 -벌음, 용어/단어 선택 -반복, 과장 -회상 -암시 -어조, 아이러니, 역설, 비꼬기 -그림 및 사진, 도표 등 -서사적 시문드스케이프와 같은 복합양식적 요소 -설명 -서술적, 설명적 텍스트 구조
			시	-하이쿠, 소네트, 발라드 (이야기를 담은 시나 노래), (장례식 때 부르는) 만화, 서사시 등
			관련 문법 섹션	-회고록 -전기 (자서전) -문헌 분석 -리뷰, 추천글 -작가 프로파일

<p>과학</p> <p>-자연과 물질적 현상을 이해하기 위해</p> <p>-문제에 대한 해결책을 제시하기 위해</p> <p>-이슈와 아이디어를 탐구하고 토론하기 위해</p> <p>-자신과 사회에 미치는 영향을 고려하기 위해</p>	<p>〈학교〉 설명문, 보고서, 저널의 기사 등과 같은 과학 텍스트를 읽고 창작하고 토론함.</p> <p>공학적 과제에 대한 해결책을 설계하고, 다이어그램과 흐름도를 사용하며, 과학적 현상을 조사하기 위해 단계적 절차를 따름.</p> <p>〈학교 밖〉 게임, 오리, 공예에 참여할 때 과학 텍스트에 참여하며, 과학이나 건강과 관련된 뉴스를 읽고 봄.</p>	<p>-과학 보고서</p> <p>-보도 자료</p> <p>-과학 뉴스 및 심층기사</p> <p>-과학 잡지 기사들</p> <p>-참고 자료 및 휴대용 도감</p> <p>-발견과 관련된 이야기</p> <p>-자기 및 1인칭 시점의 진술</p> <p>-과학에 대한 블로그 및 공동 참여 형태의 것들</p> <p>-대학, 연방과 주 기관, 공식적인 연구팀, 병원 등 과학 웹사이트</p> <p>-원시 데이터</p> <p>-실험실 작성 노트, 과학 저널</p> <p>-절차</p> <p>-출판된 연구 논문</p> <p>-개인적인 의사소통</p>	<p>논거, 비교, 순서 그리고/또는 인과관계를 구성하는 언어적 틀이나 암시</p> <p>추상화 및 명시화</p> <p>-텍스트에 포함된 과학 특수적 단어들의 정의</p> <p>-다의성에 대한 과학 특수적 정의</p> <p>-주장의 단서(~일지도 모른다, 아마도, 나타나다, 제안하다 등)</p> <p>-공간적 및 시간적 지시어</p> <p>-규모와 척도에 대한 언어적, 숫자적 지시어</p> <p>-차트, 표, 그래프, 방정식, 다이어그램, 도식, 모형, 사진, 디지털 인쇄물, 이미지와 같은 시각적, 도표적 요소</p> <p>-모의실험, 타임랩스 사진, 애니메이션과 같은 복합양식적 요소</p>
<p>사회</p> <p>-과거의 사건이 현재에 어떻게 영향을 미치는지에 대해 이해하기 위해</p> <p>-이슈와 아이디어를 탐구하고 토론하기 위해</p> <p>-인간의 동기, 인식, 윤리에 대해 이해하기 위해</p> <p>-사회를 위한 변화를 옹호하기 위해</p>	<p>〈학교〉 1차 및 2차 자료 문서, 교과서의 역사적 이야기, 사례 연구, 시사, 지도, 데이터 법정 사례 등과 같은 사회적 텍스트를 읽은 회고록, 연표, 전기를 읽고 창작하며, 이에 대해 토론함.</p> <p>〈학교 밖〉 사소한 게임, 작품, 시민 활동, 공동체 토론, 자립(self-help), 공동체적 봉사(community service)에 참여할 때 역사 및 사회 텍스트에 참여함.</p>	<p>-신문 기사, 사설, 정치적 만화, 통속적인 글, 블로그, 인구 조사 데이터, 일기, 편지, 연설문, 판매 목록과 기록들, 광고, 공문서/고문서와 같은 역사적, 현대적 문서들</p> <p>-자기 및 저서전</p> <p>-역사적, 현대적 사진과 영상</p> <p>-인구통계학적, 고용 및 교육 수준, 유권자 수 및 투표율 통계, 국내총생산 및 기타 경제 측정 등과 같은 정보를 전달하는 데이터(표, 차트, 그래프, 인포그래픽)</p> <p>-역사적 사회적, 문화적 현상과 추세에 대한 해석적 설명 또는 주장</p> <p>-절차적 텍스트, 공공 사업 공지</p>	<p>논거, 비교, 순서 그리고/또는 인과관계를 구성하는 언어적 틀이나 암시</p> <p>-연대순이나 논거를 표시하는 어휘적 표현</p> <p>추상화 및 명시화</p> <p>-선험의 수사학적 표지</p> <p>-역사적 표현 및 용어</p> <p>-언어와 수사적 장치의 이념적 표지</p> <p>-지도 타임라인, 정치 만화, 사진과 같은 시각적, 도표적 요소</p> <p>-디지털 이야기와 같은 복합양식적 요소들</p> <p>-사건의 모형(어떻게 역사적 사건이 묘사되는지)</p> <p>-공간적(정소, 위치), 시간적(시대, 시간, 순서, 시제) 지시어</p>

텍스트를 읽고 토론하거나 과학 텍스트를 읽고 공학적 과제에 대한 해결책을 설계하는 상황, 또는 학교 밖에서 게임, 요리, 공예를 하면서 과학과 관련된 텍스트를 읽거나 과학 및 건강 관련 뉴스를 읽고 정보를 얻는 상황을 통해 제시된다.

마지막으로 사회 맥락에서 독자는 역사, 지리, 문화 등에 관한 텍스트를 읽고 여러 정보와 생각을 조직적으로 정리하고, 사회적 체계와 사건이 과거부터 현재까지 어떻게 발전하고 변화하는지를 탐구하게 된다. 또한 사회와 관련된 개념과 도구를 활용해 역사적, 현대적 쟁점에 관한 문제를 해결해 나가게 된다. 사회 영역의 문항은 학교에서 교과서의 역사적 이야기, 사례 연구, 법정 사례, 회고록, 연표, 전기 등 다양한 텍스트를 읽고 토론하는 상황, 또는 학교 밖에서 시민 활동이나 공동체 토론에 참여하기 위해 사회 텍스트를 읽는 상황을 통해 제시된다.

학문 맥락을 바탕으로 텍스트를 구성하고 이에 따른 읽기 활동을 설계하는 것은 사회문화적 읽기 구인을 구현하는 것에 있어 핵심적이다. 텍스트는 거시 사회문화적 맥락에 포함된 문화·경제·정치 등의 요소를 반영하는 실체라는 점에서, 학문 맥락에서의 읽기를 평가하는 것은 학생들의 실제 읽기 목적과 활동과 밀접하며 학문의 숙달을 목적으로 하는 좁은 영역에서의 읽기보다 평가 대상으로서 높은 가치를 지닌다.

3. 시나리오 기반의 상황적 평가 과제 설계

앞선 논의의 연장선에서, 읽기 평가는 본질적으로 측정하고자 하는 것이 필연적으로 그것이 발휘되는 상황 맥락에 좌우된다. 평가 과정에서 텍스트를 읽어야 하는 활동이 독자에게 어떤 문제 상황으로 이해되는지, 그 상황이 독자에게 얼마나 중핵적인지는 평가의 과정과 결과에 상당한 영향을 미친다. 이에 읽기 평가에 놓인 리더러시 사건에 맥락화된 구체적 읽기 상황은 구인 타당도와 직접적으로 관련한다.

시나리오 평가는 학생에게 구체적인 상황과 실제적인 목적의식을 부여함으로써 자신의 읽기 행위가 놓이는 맥락적 위치를 안내하고, 이는 평가의 탈맥락성(decontextuality)을 개선하는 데 이바지한다. 이로써 피험자인 학생은 활용되는 리터러시 사건의 의도와 목적에 맞게 평가 과제에 접근할 수 있으며, 평가자는 본래 측정하고자 한 학습자의 읽기 역량을 정밀하게 관찰하고 판단할 수 있게 된다(Sabatini et al., 2020). 이때 평가에서 활용되는 시나리오는 학습자의 실제적인 삶 속에서 리터러시가 요구되는 상황을 중심에 두고 설계되어야 하며, 그러한 상황의 중핵성, 실제성, 역동성을 고려(유상희·옥현진·서수현, 2018; Barton, Hamilton, & Ivanic, 2000; Bloome, Carter, Christian, Otto, & Shuart-Faris, 2005; Purcell-Gates, Anderson, Gagne, Jang, Lenters, & McTavish, 2012)해야 한다. 읽기 평가에서 사용한 시나리오가 학습자의 실제적 읽기 활동과 유리되어 있다면 그 시나리오는 학습자가 리터러시 사건에 참여하는 데에 촉진제로서 기능할 수 없으며, 결국 그것은 탈맥락적 읽기를 수행하는 종전의 평가 경험과 다를 바 없기 때문이다.

2026 NAEP의 읽기 시나리오는 피험자가 평가 과제에 구현된 리터러시 사건에 충분히 몰입하고 참여할 수 있도록 설계되었다. 시나리오는 읽기의 목적과 독자의 역할을 구체적으로 특정하여 독자가 어떤 정체성을 가지고 어떠한 의도로 글을 읽어야 하는지에 대해 명시적으로 고려한다. 이는 구체적인 목적 설정과 독자의 역할 부여가 상황구체적 실천 역량을 측정하는 데 기여한다는 점을 고려하였기 때문이며, 이를 통해 학생들은 평가에 제시된 시나리오에 더 깊게 몰입하여 최적의 문제해결 방법을 모색할 수 있게 된다.

〈그림 2〉는 각 시나리오에서 측정하고자 하는 읽기 역량을 구현한 방식을 명시하는 블록의 구성으로, 학문적 맥락과 넓은 읽기 목적을 바탕으로 구체적인 읽기 목적 및 독자 역할이 결정되며 문항에서 의도하는 구체적이고 특징적인 사건과 행위를 규명하고 있다. 가령 ‘이해’를 목적으로 하는 4학년 문학 영역(맥락) 블록 ‘Hana Hashimoto’에서는 학급 독서 토론에 참여하기 위해 Hana가 어떤 사람인지 탐구하는 과제를 설정하고 있으며(〈그림 3〉 참

블록	학문 맥락	문학/사회/과학	넓은 목적	RDU/RSP
	구체적 목적		독자 역할	
과업	과업별 목적	과업1	...	과업 n
	텍스트			
	이해 문항			
디지털 특성	반응 형식			
	UDEs			

〈그림 2〉 2026 NAEP 읽기 평가들의 블록 구성

조), ‘문제해결’을 목적으로 하는 12학년 사회 영역(맥락) 블록 ‘Hill District’에서는 지역의 역사적 사건과 관련하여 공동체 구성원들이 가진 상이한 관점을 해석하고 조정하기 위해 프로젝트 리더와 협력하여 자신이 이해한 내용을 적용해야 하는 과제를 설정한다(〈그림 4〉 참조).

블록	학문 맥락	문학, 4학년	넓은 목적	RDU	블록 이름	하나
	구체적 과제 목적	학급 책 토론에 참여하기 위해 한나가 어떤 사람인지 배우기				
	독자 역할	책 토론을 준비하기 위해 3명의 학급 친구들과 작업하기				
과업		과업1			과업2	
	과업별 목적	주요 사건과 인물의 생각, 감정, 행동에 대해 배우기			한나가 어떤 사람인지에 대해 쓰기	
	텍스트	Hana Hashimoto, Sixth Violin				
	문항 수	6-7문항			6-7문항	
디지털 특성	반응 형식	• 단일선택형, 복수선택형 • 빈칸 채우기 • 짧은 주관식			• 짧은 주관식 • 관련짓기(matching) • 영역 할당하기(zones)	
	UDEs	과제 기반	교사 캐릭터, 학생 캐릭터, 과제 리마인더, 단어 은행			
		동기부여	교사 캐릭터, 학생 캐릭터			

〈그림 3〉 블록 ‘하나’의 구성(NAGB, 2022: 96)

블록	학문 맥락	사회, 12학년	넓은 목적	RSP	블록 이름	힐 지구
	구체적 과제 목적	현시적이고 역사적인 사건과 관련하여 공동체 구성원들의 관심을 해석하고, 프로젝트 리더와 함께 각자의 작업을 공유하고 통합하기 위해 자료들을 읽고 이해한 바를 적용하기				
과업	독자 역할	각각 다른 관점에서 인지된 공원 계획의 이점과 우려를 합치고 정리하여 프로젝트 리더를 돕기				
	과연별 목적	과업1	과업2	과업3	과업4	두 개의 관점을 선택하고, 다양한 공동체 구성원의 관점을 요약하고, 자신의 관점이 최종 보고서에 포함되도록 공유하기
	텍스트	공원 및 다양한 공동체 구성원이 공원 계획에 대해 어떻게 생각하는지에 대해 배우기 위해 다문서 읽기	피츠버그 힐 지구의 역사와, 그 역사가 공원 설계에 어떻게 관련하는지에 대해 배우기	관점에 따라 이점과 우려를 합쳐 정리하고, 직전에 읽은 텍스트로부터 증거를 제공하기 위해 디지털 노트패드 사용하기		
		• 2개의 웹 페이지 • 4개의 소셜미디어 논평	• 검색 엔진 결과 • 2개의 웹 페이지			
		3-5문항	3-4문항	3-4문항	3문항	
		반응 형식	• 단일선택형, 복수선택형 • 짧은 주관식	• 단일선택형, 복수선택형 • 짧은 주관식 • zone	• 단일선택형, 복수선택형 • 그리드형 • 짧은 주관식	
디지털 특성	빈용 형식을 기반으로 하는 UDEs					과제 캐러티, 목적 설정 진술, 그래픽 조직자, 디지털 노트패드

〈그림 4〉 블록 ‘힐 지구’의 구성(NAGB, 2022: 105)

시나리오를 기반으로 읽기 과제를 설계하는 것은 상황인지(situated cognition) 관점을 바탕으로 한다. 리터러시 사건 속에서 표상되는 의미들은 결코 일반적이거나 범용적이지 않으며 구체적이고 특정한 상황에 따라, 그리고 그 상황에 놓인 독자의 목적과 역할에 따라 특별한 효용을 가지게 된다는 것이다. 이는 인지 과정에 관한 사회문화적 패러다임을 전제하는 것으로, 요컨대 시나리오 기반의 과제 설계는 읽기 행위의 본질인 상황구체성을 타당하게 측정하고자 하는 혁신적 시도인 것이다.

4. 사회적으로 정의로운 보편적 디자인 설계의 실현

읽기 평가에서는 평가에 참여하는 과정과 결과에 영향을 미치지만 읽기 능력과 무관한 요소들이 평가의 결과에 미치는 영향을 최소화하여 타당도를 확보해야 한다. 읽기 평가라는 리터러시 사건에 참여하는 독자는 모두 개별화된 맥락을 가지므로 ‘무관한 요소’는 매우 다양하다. 특히 이러한 요소가 피험자의 읽기 능력을 저평가할 때 그 결과의 타당도가 크게 위협받는다. 예를 들어 피험자가 참조할 텍스트를 발견하지 못했거나 시나리오의 상황을 해결하고자 하는 동기가 저하되었거나 이전 문항의 오답이 다음 문항에 영향을 준다면, 평가 설계에서 측정하고자 했던 총체적인 읽기 능력을 제대로 파악할 수 없게 된다.

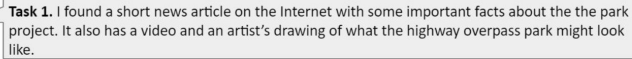
이러한 문제를 최소화하고 피험자의 다양성과 차이를 최대한 고려하기 위해, 2026 NAEP에서는 보편적 디자인 요소(Universal Design Elements, 이하 UDEs)를 배치하여 모든 학생이 과제를 명료히 인식하고, 과제를 끝낼 수 있도록 동기를 부여하고, 과제와 관련한 맥락적 정보를 제공함으로써, 유니버설 디자인 원리(Thompson, Johnstone, & Thurlow, 2002)를 반영한 목적적 평가가 잘 실현될 수 있도록 하고 있다. UDEs는 피험자가 평가에 익숙하지 않음으로써 발생할 수 있는 잡음을 줄이고 평가 내용에 충실히 몰입하도록 지원할 수 있다.

NAEP 2026에서는 세 종류의 UDEs를 활용한다. 먼저 과제 기반 UDEs는 제한된 시간 안에 독자가 평가의 요구를 명료히 이해하고 이용가능한 자원을 충분히 사용하여 읽기 과정에 충분히 인지적으로 참여할 수 있도록 안내하는 역할을 한다. 과제 기반 UDEs에는 독자가 텍스트에 접근하는 과정과 목적 등을 안내하는 일련의 지시문, 독자가 아이디어를 기록하고 다시 검토할 수 있도록 하는 그래픽 조직자, 작업을 완료하기 전 독자가 자신의 접근 방식을 시뮬레이션해 볼 수 있는 예제(Sparks & Deane, 2015) 등이 포함된다. 이러한 요소는 독자가 평가 문항의 요구를 적절히 이해하지 못함으로써 읽기 역량이 저평가되는 문제를 줄인다.

다음으로 동기부여 UDEs는 어려운 과제에 직면한 독자에게 흥미와 몰입, 끈기를 지원하는 역할을 한다. 독자의 동기를 지원하는 것은 읽기 이해에 긍정적인 영향을 미친다(Afflerbach, 2022; Fredricks, Blumenfeld, & Paris, 2004; Guthrie, 1996; Guthrie & Wigfield, 2000; Newmann, 1992; Skinner & Belmont, 1993). 동기부여 UDEs인 ‘과제 캐릭터’는 때로는 동료로서 독자를 응원하고 시나리오 내에서 과제의 중요성을 설명하여 독자의 흥미를 유지하고, 때로는 전문가로서 과제에 관한 정보를 제공한다. 이는 독자가 자신의 읽기 역량을 평가 상황에서 충분히 발휘하는 기능을 한다는 점에서 읽기 평가의 타당도에 기여한다.

마지막으로 정보형 UDEs는 평가와 관련한 구체적인 맥락을 제시함으로써, 피험자가 문항에서 묻고자 하는 바에 접근하도록 돕는 역할을 한다. 평가 항목의 난도를 낮추지 않으면서 피험자의 구인 접근을 안내한다는 점은 평가 목표의 실현 측면에서 실용성을 제공한다. 정보형 UDEs에는 저자나 텍스트에 대한 짧은 설명과 같은 간단한 도입 글, 모호한 단어를 정의하는 간단한 어휘 팝업 등이 있다. 이는 평가의 구인타당도가 지시적 과정에 의해 저해될 수 있음을 고려한 장치이다.

〈그림 5〉에서 ‘힐 지구’의 과제 캐릭터 ‘Kai’가 과제에 대한 상세 사항을 안내하는 구체적 장면을 확인할 수 있다. Kai는 평가 내에서 설정된 프로



Directions: Read the webpage and watch a short video to learn about the park project.

Then, select next to answer several questions about the project.

NEXT

〈그림 5〉 블록 ‘힐 지구’에서의 과업1 전체 안내

젝트의 리더이며, 피험자는 Kai를 도와 몇 가지 과제를 수행하는 역할을 수행함으로써 시나리오에 실제성을 부여받는다. Kai는 뉴스 기사가 공원 프로젝트에 대해 중요한 정보를 담고 있다는 점을 명시적으로 밝히는데, 이는 피험자에게 텍스트의 상황적 중핵성을 제시함과 동시에 해당 텍스트를 주의 집중하여 읽게 하는 동기부여적 장치로 기능할 수 있다. 이때 기사에 포함된 영상과 그림도 중요한 정보원임을 언급함으로써 피험자가 문항에 응답하기 위해 읽어야 하는 정보를 놓치지 않도록 안내한다는 점도 주목할 만하다.

‘힐 지구’의 과업³은 디지털 노트패드를 통해 공원에 대한 이점과 우려를 몇 가지 관점에 따라 정리하는 것을 구체적 목적으로 한다. <그림 6>에서 Kai는 시민들의 공원 계획에 대한 SNS 반응이 어떤 관점에 부합하는지 연결할 것을 요청하며, 시장에게 보낼 보고서에 피험자가 정리한 내용을 포함할 것임을 언급하고 있다. 이는 시나리오 속에서 독자의 역할을 언급하여 피험자의 목적성을 구체화하고 동기를 강화하는 장면으로 볼 수 있다. 한편 그리드의 상단에 위치한 ‘노트패드’는 독자가 필요할 때 클릭하여 다양한 텍스트 출처에서 읽은 세부 사항을 저장하고 조직하고 회상할 수 있도록 돕는다. 컵



I found a lot of different opinions and perspectives on the park in this website. If you can help organize these by topic, I'll add them to our summary report for the Mayor.

Directions. Below are comments from two community members. Determine which perspective best applies to each comment and if that person's comment would be considered a benefit or concern. Then select and drag each comment to the appropriate box in the table in your notepad.

Longtime Hill District Resident Brenda Tate: For Brenda Tate, who has lived on the same block of Webster Avenue in the Hill for all of her 70 years, the park once again will give her the chance to traverse Wylie Avenue to the park then into Downtown and back. "There won't be separation. There will be a clear avenue to come back and forth. It's symbolic," she said.

Ms. Tate, who with her 98-year-old aunt will be attending Friday's groundbreaking, sees positives in the park's construction. "It will be a nice green space, a welcoming space, for people who want to come into the community," she said. *(supportive member of the Hill District)*

City Councilman R. Daniel Lavelle: "What we're going to begin doing [Friday] is finally righting those wrongs of 50 or 60 years ago," added Mr. Lavelle, who represents the Hill. While the park is important, Mr. Lavelle said the greater value lies in providing business and job opportunities within the arena redevelopment for Hill residents and minorities. *(city councilman who represents the Hill district)*

	Benefits	Concerns
Recreational Perspective		
Environmental Perspective		
Economic Perspective		
Historical Perspective		

NEXT

〈그림 6〉블록 '힐 지구' 과업3의 문항 중

퓨터 기반의 평가 상황이 아니라면 독자는 화면에서 읽은 내용 중 주요한 정보를 종이에 메모하는 방식으로 정보를 조직한다는 것을 고려할 때, 노트북드는 과제 기반 UDEs로서 피험자의 과제 접근을 도움과 동시에 실제적인 리더러시 경험을 재현하고 있다는 점을 알 수 있다.

2026 NAEP에서는 응답 형식에도 유니버설 디자인 원리를 적용하여 피험자의 다양성을 고려하고자 한다. 평가가 측정 구인과 무관한 학생의 능력을 요구할 때, 학생이 가지고 있는 개인적인 요인(선천적·후천적 장애 등)에 의해 학생의 읽기 역량이 저평가된다면 이는 타당하지 않기 때문이다(Reynolds, Livingstone, Wilson, & Wilson, 2010). 이를 위해 2026 NAEP에서는 주관식 문항에 응답하기 어려울 경우 음성 녹음을 활용하는 방식을 제안하여, 조정(accommodation)을 기하였다.

2026 NAEP의 UDEs는 읽기 현상을 둘러싼 개인적 차이와 그것의 출처가 되는 사회문화적 다양성을 고려한다. 이는 평가가 사회적으로 정의로운 교육 실천과 평가 기회의 형평성(equity) 실현에 기여할 수 있도록 돕는 초석이라는 점에서 가치를 지닌다. 사회문화주의적 패러다임에서 대단위 평가는 읽기라는 구인을 적절하고 충분히 포괄하는 정확성을 지녀야 하면서

도, 개별 독자가 위치한 사회문화적 지형도를 고려하여 독자 간 차이에 유연하게 대응할 수 있어야 한다.

5. 평가 결과의 보고

시도 및 국가 단위로 이루어지는 평가의 결과는 다양한 방식으로 활용된다. 평가의 결과는 학생 개인의 특성 및 변화를 기술하여 개별 학습자에게 후속 학습의 방향을 제안하지만, 대단위 평가의 결과는 학습자 집단의 특성과 변화 추이에 대한 정보로 가공됨으로써 교육적 자원과 기회에 관한 정책 결정의 기반으로 활용된다는 점에서 사회적 영향력을 미친다.

학습자 집단화의 준거로는 일반적으로 지역 단위, SES·ESCS (Economic·Social·Cultural Status), 성별 등이 활용된다, 이러한 집단의 성취에 관한 정보는 그 자체로 실천적 함의를 제공하지 않는다. 그러한 성취를 보조적으로 설명할 수 있는 인지적·정서적 맥락 변인과 함께 설명될 때 정보의 가치를 얻을 수 있다. 일반적으로 대단위 평가의 결과는 SES와 정적 상관을 보이는 것으로 분석된다(OECD, 2023; OECD, 2024). 특히 PISA, PIACC 등 국제 성취도 평가에서는 읽기 성취도가 낮은 피험자 집단 사이에서 SES에 따른 고착도가 더 크게 나타나는 것으로 보고되었다. 이러한 결과 분석은 중요하지만, SES는 교육적으로 가변적(malleable) 변인이 아니기에, 실질적 변화를 이끌기 위해서는 보다 세밀한 관점에서의 접근을 유도할 수 있는 추가적 정보가 요구된다. 이에 2026 NAEP Reading Framework에서는 학생의 읽기 발달에 영향을 주는 다양한 정보를 세분화하여 수집함으로써, 대중 및 이해관계자에게 포괄적이고 통합적인 교육에 관한 이해를 촉진하도록 하였다. 겉보기에는 고정된 범주로 여겨지는 것이 실제로는 자원 접근성의 상당한 차이를 포함하고 있을 수 있다(NAGB, 2022)는 점을 고려할 때, 세밀한 관점에서의 정보 수집과 상호작용향을 고려한 결과의 해석은 결과 타당도의 핵심 요인이다.

2026 NAEP에서는 성별, SES, 인종/민족, 장애 상태, 영어 학습자 상태 등의 학생의 특성 및 학교의 특성을 수집하고, 성취에 영향을 미쳤을 수 있는 맥락 변인을 고려하여 그 결과를 보고한다. 여타 국제 평가와 비교했을 때 특기할 만한 점은, SES를 학생의 인종/민족 범주 내에서 재구성함으로써 인종/민족 간 성취도 격차를 과대추정하지 않도록 주의했다는 점이다. 가령 백인 학생 집단의 성취도 점수가 흑인이나 히스패닉 학생 집단의 점수보다 높게 나온다면 이는 경제적 자원이나 교육 환경 등 SES에 따른 영향을 크게 받았을 가능성이 높다. 사회경제적 불평등의 효과를 고려하지 않고 단순 평균만을 비교할 경우, 그 평가가 타당하게 구인을 드러낸다고 볼 수 없다.

〈표 2〉 2026 NAEP 설문 인구사회적 카테고리

학생	학년	SES	인종/민족	영어 학습자 상태
	<ul style="list-style-type: none"> • 4학년 • 8학년 • 12학년 		<ul style="list-style-type: none"> • 백인 • 흑인/아프리카계 • 히스패닉/라틴계 • 아시아계 	<ul style="list-style-type: none"> • 현 영어학습자 • 전 영어학습자 • 비 영어학습자
	성별	지역	학습/접근 기회	장애 여부
			<ul style="list-style-type: none"> • 도서관 이용 등 	<ul style="list-style-type: none"> • 특수교육 대상 • 특수교육 대상x
학교	유형		지역사회 특성	지역
	<ul style="list-style-type: none"> • 공립 • 사립 		<ul style="list-style-type: none"> • 도시 • 교외 	

또한 2026 NAEP에서는 평가 결과의 생산적 해석을 위해 성취도를 영어 학습자 범주에 따라 세분화하여 보고한다. 기존에는 평가 시점에서 학습자가 영어 학습자(English Learners)인지 여부를 수집하여 ‘전(前) 영어 학습자’와 ‘비(非) 영어 학습자’의 읽기 성취도가 한 집단에서 보고되었는데, 이 때문에 전 영어학습자의 성취 향상이 포착될 수 없었다. 이에 2026 NAEP에서는 영어 학습자 범주를 ‘현(現) 영어 학습자’, ‘전 영어 학습자’, ‘비 영어 학

습자'로 세분화하여 성취 추이에 관한 생산적인 정보를 제공할 수 있도록 하였다.

2026 NAEP는 학생의 개인적 특성 및 환경적 특성에 대한 설문을 통해 학습자의 읽기에 영향을 줄 수 있는 맥락 변인을 조사한다. 학생 개인적 특성으로는 참여, 동기, 지식, 주도성, 노력, 읽기에 대한 흥미 등이, 학생의 환경적 특성으로는 가정, 공동체, 학교 환경에 대한 학생들의 인식, 소속감, 지지감이 형태로 수집된다. 이러한 요인들은 인구사회학적 변인처럼 고정적이지 않고 가변적이라는 점에서, 읽기 성취에 관한 실질적 발전을 이끌어내고자 하는 정책 입안자 및 교육 관계자에게 교수설계의 심층적 근거를 제공한다. 이러한 정보는 학생, 교사, 관리자의 설문뿐 아니라 학생의 평가 응시 과정에서 얻어지는 과정 데이터를 통해 수집된다.

대단위 평가의 결과를 다양한 맥락적 변인을 바탕으로 해석하고 활용하는 것은, 평가의 사회문화적 위치를 고려하여 실천적 변화를 이끈다는 점에서 사회문화적으로 타당한 평가에서 반드시 요구되는 지점이다. 맥락 변인에 관한 데이터는 특정 집단의 학생들의 성취를 해석하는 것에 관한 실증적 증거를 제공한다는 점에서, 정책 입안자, 교육과정 설계자 및 교사들에게 읽기 이해력 향상을 위한 실행가능한 변화안을 모색할 수 있게 한다.

IV. 논의 및 제언

이 연구는 교육측정 및 학습과학 분야에서 논의된 타당도 이론을 사회문화주의적 관점에서 재배치하였다. 사회문화주의적 관점에서의 읽기 평가 타당도는, 읽기 역량을 구성하는 사회문화적 요인을 반영하는 구인의 설정과 사회문화적 역학 관계를 고려한 결과의 해석 및 활용을 요구한다. 이에 사회문화주의적 관점에서의 타당도에 관한 요구를 가장 잘 반영하고 있다고

판단되는 국가 단위 성취도평가인 2026 NAEP Reading Framework를 통해 그 구체태를 확인하였다. 이 절에서는 논의한 내용을 바탕으로 한국적 맥락에서의 평가 개발에 시사점을 제공하고자 한다.

첫째, 학문적 맥락을 바탕으로 한 시나리오를 제시하여 읽기 평가의 활동을 구성해야 한다. 지문을 통해 표상되는 읽기 경험은 본질적으로 텍스트에 투영되는 학문 공동체의 관습을 반영하는바, 읽기 평가에서 측정하고자 하는 실제적 읽기 역량을 추정하기 위해서는 학문적 맥락의 특성을 바탕으로 한 시나리오를 만들고 이를 바탕으로 구체적 목적과 활동을 구현해야 한다. 이는 독자의 읽기 경험을 입체적으로 구성하여 인지적·정의적 참여에 대한 증거로 작용한다(Guthrie, 1996)는 점에서, 탈맥락적인 평가에서 제기될 수 있는 구인 표상에 관한 타당도 문제를 보완한다.

둘째, 독자의 다양성을 고려하여 평가 참여 과정을 보조할 수 있는 피험자 친화적 요소를 구안해야 한다. 평가의 결과에 뚜렷한 영향 관계를 형성하지만 읽기 역량과 무관한 독자의 특성이나 평가의 지시적 과정은 구인타당도를 위협한다. UDEs와 같은 요소는 피험자가 평가 설계에서 의도한 과제 맥락을 제대로 이해하고 읽기 활동에 참여할 수 있도록 이끌며, 읽기 과정에 덜 몰입하거나 동기를 부여받지 못함으로 인해 저평가될 수 있는 피험자의 읽기 능력을 보존시킨다.

셋째, 평가라는 사건이 놓인 사회문화적 역학 관계를 고려하여 다양한 사회적 변인을 수집하고 보고해야 한다. 2026 NAEP에서는 인구학적·사회적·인지적·정서적 변인을 수집함과 동시에, 학교 및 공동체의 자원과 같은 환경적 변인 및 교사, 교수, 교실의 지지와 같은 사회정서적 변인을 수집하고 있다. 이는 평가의 결과에 영향을 미쳤을 수 있는 다양한 변인을 결과 리포트에 반영하기 위함으로써, 평가 결과가 의미하는 바에 관한 맥락적 이해를 제공하기 위함이다.

한편 국내에서 수집되는 학습자의 사회적 맥락 변인은 성별과 도시 규모로(한국교육과정평가원, 2025), 여타 국제 평가 및 NAEP에서 학습자의 인

중/민족이나 SES를 주요한 맥락 변인으로 조사한다는 점과 대별된다. 성취도 평가의 결과는 학생 개인의 학업적 노력에 대한 결론이기도 하지만, 학생이 속한 공동체의 기회와 자원에 관한 맥락적 설명이기도 하다. 따라서 국가 단위의 성취도 평가라는 중대한 실천 행위가 미칠 사회적 영향력을 고려하여, 학습자의 언어 상태(모어가 한국어인 경우, 모어가 한국어가 아니지만 가정 내에서 한국어를 사용하는 경우, 가정에서 한국어를 사용하지 않는 경우 등)나 사회경제적 지위에 관한 정보를 적극적으로 수집하고 이를 결과 해석에 활용할 필요가 있다. 현대 한국 사회의 통계를 기반으로 사회문화적 맥락 변인을 확장적으로 수집하고 보고하여 활용하는 환류가 발생할 때, 교수학습 지원 및 경제적 자원 분배의 근거가 마련될 수 있을 것이다.

넷째, 맥락 변인의 상호작용항을 고려하여 결과를 해석해야 한다. 학습자는 단일 집단에만 속하지 않는다. 한 명의 학습자는 여러 개의 사회문화적 집단에 속하며 집단들은 서로 얽혀있다. 2026 NAEP에서는 SES를 인종/민족 집단 안에서 재구성함으로써 ‘사회경제적 지위’와 ‘인종/민족’이라는 두 개의 항이 상호작용하고 있다는 실제 사회의 맥락을 고려하고 있다. 이러한 접근은 교육 기회와 학습 자원의 분배에 있어 세밀한 이해를 제공한다. 이주 배경과 SES, 가정 내 자원 등에 관한 정보를 수집하고 각각의 상호작용항을 고려하여 학습자 집단의 성취도를 판단할 때, 실효성 있는 정책의 가능성이 모색될 수 있으리라 기대된다.

마지막으로, 사회정서학습이 표준화 검사와 정적 상관이 있다는 것으로 밝혀졌다는 최근 연구(Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Ha, McCarthy, Strambler, & Cipriano, 2025)를 고려했을 때 사회정서적 변인이 유용한 가변 자원이 될 수 있음에 주목하여 수집 카테고리를 확장해야 한다. 이때 학생들에 대한 사회정서적 지원은 특정 과목 개설의 형태가 아니라 학교에서 보편적으로 경험할 수 있는 문화로서 학생들이 경험할 수 있게끔 하는 것이 중요하다. 학교의 문화가 대화, 협력, 경청 등의 지지감과 관련한 사회정서를 중심으로 구성될 때 읽기 성취에 대한 긍정적 영

향을 기대할 수 있다. 가변적인 자원을 중심으로 학생의 사회문화적 집단 성취를 파악할 때 성취도 평가의 활용 가능성이 더욱 확장될 수 있을 것이다.

국가 단위의 읽기 성취도 평가는 교육의 실행에 관한 실증적인 점검이자 앞으로의 교육 방향에 관한 통계적 지표이다. 읽기라는 인간의 정신 활동이 놓인 사회문화적 관계를 평가의 구인으로 설정하고 문항 및 반응 형식 등을 구안하며 다층적 요인을 고려하여 그 결과를 활용하는 것은, 읽기와 학습 과학, 교육평가에 관한 사회문화적 요구에 대한 응답이자 건강한 읽기 교육에 관한 촉매가 될 것이다. 사회문화주의 패러다임이 국가 단위의 평가에 안착하여 그 실효성을 가질 수 있도록 다양한 실천적 연구가 필요한 시점이다.

* 본 논문은 2025.10.31. 투고되었으며, 2025.11.10. 심사가 시작되어 2025.12.03. 심사가 종료되었음.

참고문헌

- 박혜영(2019), 「차세대 읽기 평가의 방향 탐색: GISA 평가들과 읽기 수행 중재요인 분석을 중심으로」, 『독서연구』 52, 167-198.
- 서혁·박지윤(2009), 「다문화가정 학생의 사회문화적 배경과 읽기 능력에 대한 사례 연구」, 『국어교육학연구』 36, 393-423.
- 신명선(2002), 「사회적 실천 행위로서의 읽기 방법의 설계에 대한 시고」, 『국어교육학연구』 235-264.
- 오규설(2025), 「인공지능 시대의 확장된 비판적 리터러시 평가 개발 연구」, 『독서연구』 77, 9-64.
- 유상희·옥현진·서수현(2018), 「상황 기반 리터러시 평가를 위한 시론: 리터러시 평가에서 '상황'을 고려한다는 것은 어떤 의미인가?」, 『교과교육학연구』 22(4), 264-273.
- 조병영(2011), 「구인 타당도와 결과 타당도 이론이 독서 평가의 해석과 사용에 주는 함의」, 『독서연구』 25, 385-414.
- 조병영(2020), 「뉴리터러시와 전환적 의미 디자인-융복합 문식 환경 속 국어교육의 실천 방향 탐색」, 『국어교육』 171, 31-70.
- 최미숙(2004), 「국어과 성취도 평가의 현황과 개선 방안」, 『국어교육학연구』 20, 237-266.
- 한국교육과정평가원(2025), 『2024년 국가수준 학업성취도 평가 결과』, 충청북도: 한국교육과정평가원.
- Afflerbach, P. (2005), "National Reading Conference policy brief: High stakes testing and reading assessment", *Journal of Literacy Research* 37(2), 151-162.
- Afflerbach, P. (2022), *Teaching readers (not reading): Moving beyond skills and strategies to reader-focused instruction*, NY: Guilford Publications.
- Afflerbach, P. (2025), *Understanding and using reading assessment, K-12*, NY: Guilford Publications.
- Barton, D. & Hamilton, M., (1998), *Local literacies: Reading and writing in one community*, London: Routledge.
- Barton, D., Hamilton, M., & Ivanič, R. (Eds.) (2000), *Situated literacies: Reading and writing in context*, NY: Routledge.
- Basterra, M. d. R., Trumbull, E., & Solano-Flores, G. (2011), *Cultural validity in assessment*, NY: Routledge.
- Bennet, R. E. (2013), "Preparing for the future: What educational assessment must do", *Teachers College Record* 116(11), 1-18.
- Benavot, A., Cha, Y. K., Kamens, D., Meyer, J. W., & Wong, S. Y. (1991), "Knowledge for the masses: World models and national curricula", *American sociological review* 85(100), 1920-1986.

- Blome, D., Carter, S. P., Christian, B. M., Otto, S., & Shuart-Faris, N. (2005), *Discourse analysis and the study of classroom language and literacy events: A microethnographic perspective*, Mahwah, NJ: Erlbaum.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011), "The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions", *Child development* 82(1), 405-432.
- Emler, T. E., Zhao, Y., Deng, J., Yin, D., & Wang, Y. (2019), "Side effects of large-scale assessments in education", *ECNU Review of Education* 2(3), 279-296.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004), "School engagement: Potential of the concept, state of the evidence", *Review of educational research* 74(1), 59-109.
- Gee, J. P. (2018), "Reading as situated language: A sociocognitive perspective", *Theoretical models and processes of literacy*, 105-117.
- Gee, J. P. (1989), "Literacy, Discourse, and Linguistics: Introduction", *Journal of Education* 171(1), 5-17.
- Gee, J. & Gee, J. P. (2007), *Social linguistics and literacies: Ideology in discourses*, London: Routledge.
- Gipps, C. (2002), Sociocultural perspectives on assessment, In G. Wells, & G. Claxton(Eds.) *Learning for life in the 21st century: Sociocultural perspectives on the future of education*, Oxford: John Wiley & Sons.
- Glaser, R., Chudowsky, N., & Pellegrino, J. W. (2001), *Knowing what students know: The science and design of educational assessment*, NW: National Academies Press.
- Guthrie, J. T. (1996), "Educational contexts for engagement in literacy", *Reading Teacher* 49(6), 432-446.
- Guthrie, J. T. & Wigfield, A. (2000), Engagement and motivation in reading, In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, Vol. 3*, NY: Lawrence Erlbaum Associates Publishers.
- Ha, C., McCarthy, M. F., Strambler, M. J., & Cipriano, C. (2025), "Disentangling the Effects of Social and Emotional Learning Programs on Student Academic Achievement Across Grades 1-12: A Systematic Review and Meta-analysis", *Review of Educational Research* 0(0), 1-42.
- Hargreaves, A. (2020), "Large-scale assessments and their effects: The case of mid-stakes tests in Ontario", *Journal of educational change* 21(3), 393-420.
- Lai, Y. (2023). "The double effects of standardized testing on students and environment", *Journal of Education, Humanities and Social Sciences* 8, 1615-1620.
- Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2017), "New literacies: A

- dual-level theory of the changing nature of literacy, instruction, and assessment", *Journal of education* 197(2), 1-18.
- Luke, A. (2004), "On the material consequences of literacy", *Language and Education* 18(4), 331-335.
- Luke, C. (2003), "Pedagogy, connectivity, multimodality, and interdisciplinarity". *Reading research quarterly* 38(3), 397-403.
- Martens, K., Niemann, D., & Teltemann, J. (2016), "Effects of international assessments in education- a multidisciplinary review", *European Educational Research Journal* 15(5), 516-522.
- Messick, S. (1990), *Validity of test interpretation and use*, Princeton: Educational Testing Service.
- Mislevy, R. J. (2018), *Sociocognitive foundations of educational measurement*, NY: Routledge.
- Moss, P. A. (2003), "Reconceptualizing validity for classroom assessment", *Educational Measurement: Issues and Practice* 22(4), 13-25.
- Moss, P. A. (2007), "Reconstructing validity", *Educational researcher* 36(8), 470-476.
- Moss, P. A., Girard, B. J., & Haniford, L. C. (2006), "Validity in educational assessment", *Review of research in education* 30(1), 109-162.
- Moss, P. A., Pullin, D., Gee, J. P., & Haertel, E. H. (2005), "The idea of testing: Psychometric and sociocultural perspectives", *Measurement: Interdisciplinary research and perspectives* 3(2), 63-83.
- NAGB (2022), *Reading Assessment Framework for the 2026 National Assessment of Educational Progress*, Washington, DC: National Assessment Governing Board.
- Nasir, N. I. S. & Hand, V. M. (2006), "Exploring sociocultural perspectives on race, culture, and learning". *Review of educational research* 76(4), 449-475.
- Newmann, F. M. (1992), *Student engagement and achievement in American secondary schools* NY: Teachers College Press.
- OECD (2023), *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*, Paris: OECD Publishing.
- OECD (2024), *Do Adults Have the Skills They Need to Thrive in a Changing World?: Survey of Adult Skills 2023, OECD Skills Studies*, Paris: OECD Publishing.
- Purcell-Gates, V. (1997), *Other people's words: The cycle of low literacy*, US: Harvard University Press.
- Purcell-Gates, V., Anderson, J., Gagne, M., Jang, K., Lenters, K. A., & McTavish, M. (2012), "Measuring situated literacy activity: Challenges and promises", *Journal of Literacy Research* 44(4), 396-425.
- Purcell-Gates, V., Duke, N. K., & Martineau, J. (2007), "Learning to read and write

- genre specific text: Roles of authentic experience and explicit teaching", *Reading Research Quarterly* 42, 8-45.
- Purcell-Gates, V., Perry, K.H., & Briseño, A. (2011), "Analyzing Literacy Practice: Grounded Theory to Model", *Research in the Teaching of English* 45(4), 439-458.
- RAND Reading Study Group (2002), *Reading for understanding: Toward an R&D program in reading comprehension*, Santa Monica, CA: RAND.
- Reynolds, C. R. (2000), "Why is psychometric research on bias in mental testing so often ignored?", *Psychology, Public Policy, and Law* 6(1), 144.
- Reynolds, C. R., Livingston, R. B., Willson, V. L., & Willson, V. (2010). *Measurement and assessment in education*. Upper Saddle River: Pearson E.
- Rouet, J. -F., Britt, M. A., & Durik, A. M. (2017), "RESOLV: Readers' representation of reading contexts and tasks", *Educational Psychologist* 52(3), 200-215.
- Sabatini, J., O'Reilly, T., Weeks, J., & Wang, Z. (2020), "Engineering a 21st Century Reading Comprehension Assessment System Utilizing Scenario-Based Assessment Techniques", *International Journal of Testing* 20(1), 1-23.
- Skinner, E. A. & Belmont, M. J. (1993), "Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year", *Journal of educational psychology* 85(4), 571.
- Sparks, J. R. & Deane, P. (2015), "Cognitively -based assessment of research and inquiry skills: Defining a key practice in the English language arts", *ETS Research Report Series 2015*, 1-55.
- Street, B. V. (1984), *Literacy in theory and practice*, Cambridge: Cambridge University Press.
- Street, J. C. & Street, B. V. (1995), The schooling of literacy, In J. Bourne, M. Briggs, & M. Selinger(Eds.), *Subject learning in the primary curriculum*, London: Routledge
- Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002), *Universal design applied to large scale assessments*. MN: National Center on Educational Outcomes.

사회문화주의 패러다임의 읽기 성취도 평가 설계를 위한 제언

— 2026 NAEP 읽기 평가 체계의 분석 및 함의

김가은·이채윤·조병영

이 연구는 사회문화주의 패러다임의 관점에서 읽기 성취도 평가의 설계 원리를 탐색하고, 이를 바탕으로 한국의 대단위 읽기 평가 체계 개발에 주는 시사점을 제시하는 것을 목적으로 한다. 이를 위해 먼저 교육측정 및 학습과 학 분야의 선행연구를 검토하여 사회문화적 구성주의에 기반한 타당도 개념을 정리하였다. 다음으로, 2026 NAEP의 읽기 체계를 분석하여 해당 체계가 최신의 읽기 이론과 사회적 요구를 어떻게 반영하고 있는지를 개발 및 설계의 측면에서 고찰하였다. 분석 결과, (1) 사회문화주의적 관점을 반영한 읽기 구인 모형, (2) 학문 맥락별 읽기 목적과 텍스트 선정, (3) 시나리오 기반의 과제 설계, (4) 학습자 다양성을 고려한 보편적 디자인 요소, (5) 사회문화적 맥락 변인을 반영한 결과 보고 체계라는 특징을 보였다. 이상의 내용을 바탕으로 현대 한국의 맥락에서 사회문화주의적 읽기 평가를 설계하기 위한 방향을 위해 몇 가지 논점을 제안했다. 본 연구는 읽기 평가를 사회문화적 실천의 한 형태로 재개념화함으로써, 향후 읽기 리터러시 교육 및 평가 정책 설계에 이론적·실천적 근거를 제공한다.

핵심어 사회문화주의, 읽기 평가, 타당도, 2026 NAEP 읽기 체계, 리터러시 교육

Rethinking the Design of Reading Achievement Assessment from a Sociocultural Perspective

— Insights from the 2026 NAEP Reading Framework

Kim Gaeun · Lee Chaeyun · Cho Byeongyoung

This study examines the construct and principles of reading achievement assessment from a sociocultural perspective and offers implications for developing a large-scale reading assessment framework in Korea. To begin, prior research in educational measurement and the learning sciences is reviewed to clarify the concept of validity grounded in sociocultural constructivism. Subsequently, the 2026 NAEP Reading Framework is analyzed to identify how the framework incorporates contemporary theories of reading and addresses current societal demands in its development and design. The analysis highlights five key features: (1) a socioculturally informed construct model of reading; (2) specification of reading purposes and text selection aligned with disciplinary contexts; (3) scenario-based task design; (4) integration of universal design principles to account for learner diversity; and (5) a reporting system that incorporates sociocultural contextual variables. Building on these findings, this study proposes considerations for designing socioculturally oriented reading assessments in the Korean context. By reconceptualizing reading assessment as a sociocultural practice, the study contributes theoretical and practical foundations for future literacy education and assessment policy development.

KEYWORDS Sociocultural theory, reading assessment, validity, 2026 NAEP Reading Framework, literacy education