

지리과 서답형 문항의 평가 방안과 채점 기준에 관한 연구

-NAEP 4학년 서답형 문항 답안의 채점 사례를 중심으로-

장의선

한국교육과정평가원

I. 서론

학생들의 단순 지식에서부터 생각을 조직하고 통합하며, 상황에 맞게 표현하는 고차적 사고력에 이르기까지 폭넓은 스펙트럼의 학습 결과에 대해 추정할 수 있게 하는 서답형 문항은 대부분의 교육자들에게 유용한 평가 도구로 인식되고 있다. 그러나 평가는 신중하게 제작되어야 하며, 공정성을 증진하고 편견을 감소하는 방식으로 채점되어야 한다. 평가의 목적이 학생들의 허를 찌르거나 부주의한 실수를 처벌하는 것이 아니기 때문이다. 오히려 평가의 목표는 학생의 인지적인 학업 성취 수준을 포함한 종합적인 사고활동에 대해 적절한 추정을 할 수 있게 하는 통찰력을 얻는 것이다.

이러한 중요성에도 불구하고 지리교육에서 서답형 문항의 개발이 활발하게 이루어지지 못했던 가장 큰 이유는 명확한 채점 기준 설정이 곤란하기 때문으로 추측된다. 오늘날 한국 교육의 현실에서 오답시비 없는 서답형 문항을 만들고, 누가 보아도 공정하고 객관적인 결과 처리를 위한 채점 기준을 적용해야 하는 부담감이 컸던 것이다.

한편, 학교 현장에서는 1990년대 이후 학생이 교수의 대상이 아닌 교수·학습의 주체로 인식되면서 교실 수업의 개선과 변화로 인한 평가 방식의 전환이 요구되고 있다. 평가 방식의 전환에는 문항 유형과 출제 방식, 채점 기준, 점수 산출 방식 등의 제반 여건에 대해 많은 전문가들에 의한 충분한 논의가 필요하다. 하지만 각 교육 기관에서는 충분한 논의 없이 수행 평가나 서답형 평가 등과 관련된 정책을 학교 현장에 무리하게 진행시켰다. 실제로 일부 교육청에서 2005년부터 정기고사의 서답형 평가 비율을 30%로 정한 것을 시작으로 해마다 단계적으로 확대하여 50%까지 실시할 것을 권고하였으며, 2010년에는 경기도 교육청이 시행을 권고하면서 거의 모든 시도에서 제도화된 평가 형태로 자리 잡게 되었다. 그러나 현재 중·고등학교 현장에서의 서답형 평가는 아직까지 교육적으로 의미 있는 평가 방식으로 제 자리를 찾지 못하고 있다. 교사 연수 매뉴얼이나 자료집의 형태로 문항 사례가 보급되고 있지만, 채점에 대한 신뢰도 확보 방안을 체계적으로 마련하지 못해 서답형 평가의 활성화는 여전히 미흡한 편이다.

따라서 이 연구에서는 먼저 서답형 문항의 채점 시스템과 채점 시 고려 사항을 살펴보고, 우리보다 앞서 서답형 문항의 채점 시스템을 체계적으로 적용해 온 미국 지리교육계의 서답형 문항 답안과 채점의 구체적 사례를 검토하면서 우리나라에 대한 그 적용 가능성과 과제를 모색하고자 한다. 다만 서답형 답안과 채점 사례를 제시함에 있어서 NAEP(The National Assessment of Educational Progress)의 지리영역 4학년 서답

형 문항을 분석하기로 하였다. NAEP는 미국의 국가수준 학업성취도 평가에 해당하는 시험으로서 서답형 답안 채점과 관련한 오랜 경험을 축적하고 있기 때문에 우리나라의 서답형 답안 채점 시스템의 체계적 발전에 유용한 시사점을 줄 수 있을 것으로 판단되었기 때문이다.

II. 서답형 문항 채점 시스템과 채점 시 고려 사항

1. NAEP 서답형 문항의 채점 시스템

NAEP 문항은 크게 선다형(Multiple Choice Items)과 서답형(Constructed-Response Items)으로 구성되는데, 선다형은 4지 선다형이며, 서답형은 다시 단답형(Short Constructed-Response)¹⁾과 확장형(Extended Constructed-Response)으로 구분된다. 본 연구에서는 서답형 문항에 대한 답안과 채점 사례에 한정하여 살펴해보도록 하겠다.

서답형 문항의 채점에 대해서는 일반적으로 두 가지 접근 방법이 논의되는데, 하나는 세부 요소로 구분하여 점수를 부여하는 분석적 채점 방식(analytical method, 또는 point method)이고, 다른 하나는 답안 내용을 전반적으로 읽은 후 전체적인 느낌과 판단에 의하여 점수를 부여하는 총체적 채점 방법(holistic scoring method, 또는 rating method)이다. 전자는 정답이나 모범 답안이 구체적으로 제시되어 채점의 객관성과 신뢰성이 민감할 때, 후자는 학생의 반응이 다양한 수준으로 나타날 수 있는 보다 복합적인 문항을 채점할 때 활용되는 경향이 있다(양길석, 2006: 19). 보통 서답형 문항에서는 이 두 가지 채점 방식이 적절하게 사용되는데, NAEP의 채점 가이드에 설명된 것과 같이 답안 작성에 제시된 조건을 모두 충족하면 '목표 도달', 일부를 충족하면 '보통'이라는 채점 상세 기준은 짐작하건대, 분석적 채점 방식을 염두에 둔 것으로 보인다.

NAEP는 학생들의 답안에 대해 전자 이미지 처리 및 채점 시스템을 이용하여 점수를 부여하고 있는데 이러한 채점 과정은 일반적으로 채점 가이드의 개발과 검토, 시험 채점 및 운영, 사전 보정, 평정자 훈련, 실제 채점 및 사후 관리로 진행된다(그림 1 참조). 특히 2·3 단계의 점수가 부여되는 단답형과 4 단계 이상의 점수가 부여되는 확장형을 포함하는 서답형 문항의 경우에는 학생 답안을 캡처한 이미지를 사용하여 훈련된 평정자에 의해 채점된다. 제한된 시간 내에 높은 수준의 정확성과 신뢰성을 가진 많은 수의 서답형 문항 채점은 NAEP의 성공에 필수적이므로 서답형 문항에 대한 신뢰성 높은 효율적 채점을 보장하기 위해 NAEP에서는 다음과 같은 조취를 취하고 있다.

먼저 평가 틀²⁾에 제시된 기준과 일치하는 명시적 채점 가이드를 개발하고, 자격을 갖춘 경험이 풍부한 평정자³⁾의 모집과 자격시험을 통한 능력 검증 및 훈련에 많은 노력을 기울인다. 또한 평정자들이 문항지

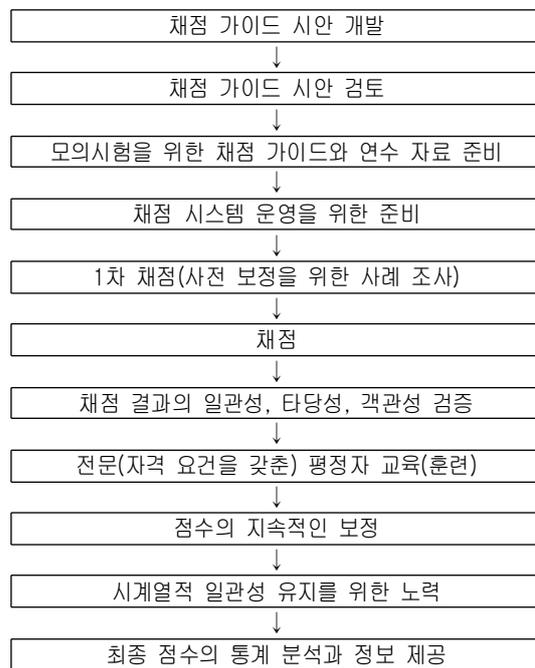
1) 단답형의 경우 대체로 'short-answer item'이라고 표기하지만, NAEP에서는 'short-constructed response'라고 표기하므로 이를 따르기로 한다.

2) 일반적으로 평가 틀은 평가 내용과 방법을 계획하고 결정해주는 지침서의 역할을 함으로써 평가의 일관성 유지를 위한 기준이 되는 것이다.

3) NAEP 평정자의 자격 요건은 4년제 대학 또는 석사 학위 이상의 자격을 소유해야 하며 채점 및 교육 유경험자가 우선적이다. 또한 일부 과목에서는 전문적 지식에 대한 자격시험을 거쳐야 하며, 기존에 치러졌던 대부분의 서답형 문항에 대한 채점 훈련의 결과들은 개별 평정자들에게 피드백 된다. NAEP 채점 팀은 일반적으로 평정 관리자(책임자)의 주도 아래 10~12명의 평정자들로 구성되어 있다.

배송과 관련하여 번거롭지 않도록 학생들의 답안을 스캔한 이미지로 전송하는 이미지 프로세싱과 전산화된 점수 부여 방식의 채점 시스템을 사용한다.

이 때, 평정자들은 각 학생들의 답안을 분리된 개별 컴퓨터 화면에서 보고 점수를 부여하는데, 채점 중인 학생의 다른 문항에 대한 답안은 볼 수 없도록 프로그램화 되어 있다. 즉, NAEP 서답형 문항의 채점은 하나의 문항에 대한 모든 학생들의 답안들이 채점될 때까지 다른 문항에 대한 어떠한 학생들의 답안도 볼 수 없도록 각각의 문항 별로 이루어지므로 평정자들 역시 한 번에 하나의 문항과 하나의 채점 가이드에 대해 훈련을 받는다. 이 방법은 전자 이미지 학생 답안에 대한 채점에 매우 효율적인 것으로 알려져 있다. 채점 후, 평정자들은 재 채점을 포함하여, 지속적인 신뢰도 검사를 통해 이전 몇 년 동안의 채점과의 일관성을 모니터링하는 한편, NAEP 평가 전문가는 이러한 결과에 대해 반복적인 모니터링을 토대로 평정자의 의사 결정 자질을 평가하고 있다. 이 모든 훈련, 채점 및 질적 관리 절차는 기술 보고서에 상세히 기록하고 있다 (NCES, 2011).



< 그림 1 > NAEP 서답형 문항의 채점 절차

특히 NAEP는 채점 과정에서 평정자들의 점수에 대한 신뢰도를 확보하기 위해서 다음과 같은 세 가지 방법을 사용하고 있다. 첫째, 평정 책임자는 각 평정자들의 평점을 본인이 직접 채점한 결과와 비교한다. 각 평정자의 채점 결과 중 5% 정도를 재 채점하여 검토하고 일관성이 있는지 점검한다. 둘째, 각 평정자들은 채점 과정을 상세하게 살펴보고, 점수화 과정에 대해 보정하는 과정을 거침으로써 점수의 안정성을 유지한다. 예를 들어 각 평정자가 점수를 부여한 후에 어느 정도의 휴식을 가진 다음, 채점 가이드를 숙지한 후 다시 채점하게 한다. 셋째, 전국 단위의 경우에는 25%, 주 단위에서는 5%의 반응을 표집하여 평정자 간

신뢰도를 계산한다(NCES, 2011).

한편, 서답형 문항에 대해서는 문항에 따라 점수 부여 단계가 달라지는데 보통 작문은 6개, 그 외에는 과목과 문항에 따라 채점 가이드에 제시된 대로 2~5단계의 점수 수준이 부여된다. 예를 들어 5단계의 점수 부여 수준을 살펴보면, 오답(incorrect), 미흡(minimal), 보통(partial), 우수(satisfactory), 그리고 탁월(extended)의 등급으로 부여되며, 3단계의 점수 부여는 수준 미달(unacceptable), 보통(partial), 목표 도달(complete)의 등급으로 부여되기도 한다. 점수 부여 단계와 기준에 대해서는 사례 제시 부분에서 보다 상세하게 다루기로 하겠다.

2. 서답형 문항 채점 시 고려 사항

선택형 평가와 달리 서답형 평가의 성공 여부는 문항 자체의 질뿐만 아니라 채점의 질에 의해서도 결정된다. 문항의 질적 수준이 양호하더라도 채점의 공정성과 신뢰성이 미흡하다면 그 결과는 교육적으로 의미가 없기 때문이다. 따라서 서답형 문항 답안에 대한 보다 신뢰롭고 합리적인 채점을 위해 고려할 사항으로서 다음과 같은 점들이 일반적으로 논의되고 있다.

첫째, 서답형 문항을 채점할 때 측정하고자 하는 평가 요소(학습 목표)를 기준으로 답안 작성 내용을 평가해야 한다. 측정되는 평가 요소와 상관없는 답안 내용들이 점수에 영향을 주도록 해서는 안 된다는 의미이다. 예를 들어 특정 교과와 수업 시간에 주어진 평가 점수는 교육자들이 특정 주제에서 학생 지식에 대해 추정할 수 있도록 하므로 문법 오류나 철자 오류 등으로 인해 점수를 낮추면 이러한 목적을 저해할 것이며, 평가의 타당성을 훼손할 수 있기 때문이다. 다시 말해, 만약 올바른 철자의 답안에만 점수가 주어진다면 답안 내용 중 틀린 철자가 포함된 학생은 학습 결과와 철자법에 대한 평가가 뒤섞이게 되어 정확한 성취 수준과 시험 점수를 받는데 방해 받을 수 있다. 또한 채점하는 동안 철자 오류에 대해 무시하고자 노력한다 하더라도 틀리게 표현된 단어가 의도된 답을 표현했는지 아닌지에 대해 결정해야 하는 문제가 여전히 남는다.

둘째, 정답으로 가능하다고 생각할 수 있는 유사어 또는 정답으로 허용될 수 있는 다른 표현을 예측하는 것도 중요하다(Williams, 1991: 22). 또한 학생들은 몇 가지 답안 중에 적절하다고 생각할 수 있는 일부만을 작성할 수 있는데, '두 가지 중 하나만 옳은 답안', '틀린 답안', '적절하지 못한 의견', '오답', 또는 이와 유사한 답안이 포함되기도 한다. 따라서 이러한 답안에 대해서는 채점이 곤란하므로 채점에 대한 명확한 준거를 문두에 밝히거나 채점 가이드에 제시하는 것이 바람직하다.

셋째, 서답형 문항을 채점할 때, 인종, 서툰 필체 등에 의한 편견의 가능성이 큰 관심사이기도 하다. 채점에서 편견을 줄이고자 할 때 몇 가지 전략을 구상할 수 있다. Kan(2007)과 Odendahl(2007)은 모두 채점 규정(rubric)의 사용을 제안하고 있는데 이는 채점 규정들이 모든 학생들이 지속적으로 동일한 기준에 따라 평가될 수 있도록 도와준다고 확신하기 때문이다. 또한 Tuckman(1994)은 채점은 학생의 이름을 가리고 완료해야 하며, 하나의 학생에 대해서 가능한 두 명 이상의 서로 다른 동료들의 채점 성적을 가지고 논의하는 것이 장려되어야 한다고 진술하였다. 학생들의 이름이 드러난 시험지가 채점 될 때는 인종 및 성별이 영향을 미치는 것으로 판명되었기 때문이다(Schafer, Gange, & Lissitz, 2005).

넷째, 하나의 문항이 또 다른 문항의 평가에 영향을 주는 가능성을 줄일 수 있는 방법으로

Popham(2006)은 다음 문항으로 넘어가기 전에 하나의 문항에 대한 모든 학생들의 답안을 채점하는 것을 제안하였다. 즉, 첫 번째 질문에 대한 점수는 두 번째 질문의 채점으로 넘어가기 전에 테스트에 응시한 학생들 전체에 걸쳐 끝나 있어야 함을 의미한다(Tuckman, 1994: 71). 이러한 방법은 앞서 밝혔듯이 현재 NAEP의 채점 시스템에서도 사용되고 있다.

마지막으로 학생들이 서답형 문항에 답할 때 그들의 언어를 사용할 수 있으므로 어떤 답안들은 읽기가 어렵거나 판독하기 어려울 수 있기 때문에 서답형 문항은 채점하기에 시간이 많이 소비되고 어려운 경우가 있다. 또한 발문과 정답이 명료하게 구성되어 있지 않다면 정답으로 고려될 수 있는 하나 이상의 답지를 가질 수 있기 때문에 이러한 점 역시 채점이 어려운 이유가 될 수 있다(Ory, 1983: 22).

이 외에 채점에 영향을 주는 요소로는 다음과 같은 내용들이 논의되기도 하였다. 먼저 답안 작성 시 직접적인 경험으로부터 도출된 증거를 사용하는 것이 교과 지식을 제시하는 것 보다 더 높이 평가되는 경우, 언어적 장벽 때문에 그들의 성취 정도를 제대로 증명할 수 없는 경우, 특정 주제에 대해서는 이야기식(서술식) 응답들이 설명식의 답안들 보다 더 높은 점수를 받아 온 경우, 응답 장르에 대한 개별 평가자들의 편견이 드러날 경우 등이다. 이러한 사례들은 서답형 문항 채점의 정당성에 대해 생길 수 있는 우려이며 또한 채점에 대한 신뢰도가 손상될 수 있는 경우이기도 하다(Schafer, Gange, & Lissitz, 2005: 22).

뿐만 아니라 일부 영악한 학생들은 옳은 답안이 아니면서도 마치 그럴 듯하게 보이는 의사(擬似) 답안(bluffing)을 작성함으로써 높은 점수를 받아내는 경우가 있다. 문두를 다른 표현으로 바꿔 쓰거나 의미를 잘 모르면서 기본 개념들을 적절히 섞어서 서술하는 유형이 대표적인 예이다. 물론 이러한 의사 답안을 작성하기 위해서는 일정 수준의 테크닉(기술)이나 그 주제에 대한 단편적인 지식이 요구되지만, 문항에 대한 옳은 답안이 아닌 한 점수를 허용해서는 안 된다(Waugh & Gronlund, 2013: 139).

이 같은 맥락에서 많은 사람들은 하나의 답안 내용을 두 번 채점할 때 같은 점수를 산출하기 어렵다는 채점의 주관성 때문에 신뢰성에 문제가 있다고 주장하기도 하고, 또 여러 평가자들 사이에서 서로 다른 채점 결과가 나올 수도 있다며 서답형 문항의 채점에 대한 공정성에 회의를 갖기도 한다. 교육자들 역시 채점에 소요되는 노력과 시간, 채점 결과에 대한 신뢰성뿐만 아니라 문항 및 학생들의 답안 내용에 대한 광범위한 통계자료와 샘플을 만들기에 적합하지 않다는 이유 등으로 선택형에 비해 서답형 문항들을 선호하지 않을 수 있다.

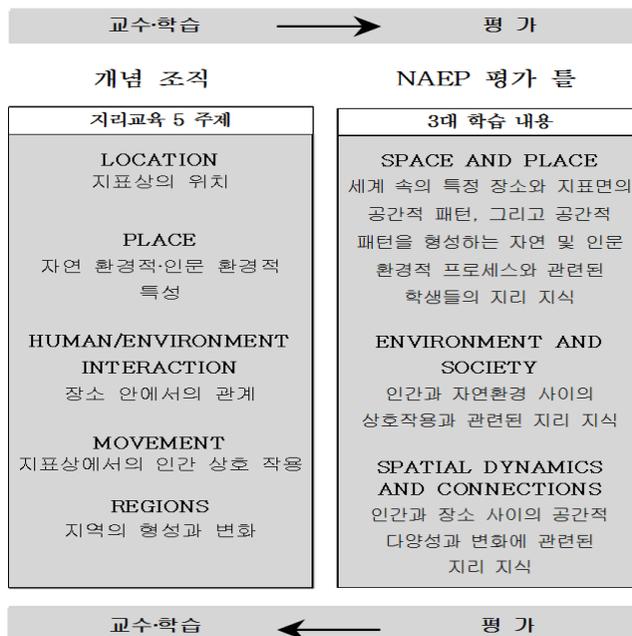
그러나 최근에 발표된 '중등학교 학사관리 선진화 방안'(2011. 12. 13 교육과학기술부 보도자료)은 이러한 단점에도 불구하고 서답형 문항에 주목하도록 요구한다. 특히 '학생이 스스로 자신의 생각을 구성하고 이를 표현함으로써 창의성, 문제해결력 등을 신장시키기 위해 서술형 평가와 수행평가의 질을 제고하도록 하고 있다는 점, 그리고 성취평가제의 도입으로 교사들이 엄격한 상대평가의 압박으로부터 다소 자유로워진 상태에서 다양한 평가방법을 구상할 수 있을 것'이란 점은 바람직한 서답형 평가의 실현 가능성을 그만큼 높여준다고 판단된다. 따라서 서답형 문항 답안에 대한 상세하고도 명확하면서 주의 깊은 채점 방안 역시 절실하다 하겠다.

Ⅲ. 지리과 서답형 문항 답안 및 채점 사례

1. NAEP의 지리과 평가 영역

1) NAEP의 지리과 평가 영역

NAEP 지리과 서답형 문항은 기본적으로 NAEP 평가 틀에 의해 제작되고 채점된다. 이러한 NAEP 평가 틀에서 지리과 평가 영역은 내용 영역과 행동 영역이라는 두 차원으로 이루어져 있다. 이것은 곧 내용에 대한 이해 정도와 인지 능력을 함께 평가하겠다는 의미이다. NAEP 지리 평가 틀의 내용 영역은 '공간과 장소(Space and Place)', '인간과 환경(Environment and Society)', 공간 관계와 공간 변화(spatial Dynamics and Connections)의 세 영역으로 구성된다. 그러나 알려져 있듯이 미국의 지리 교수·학습에서는 1984년 초·중등학교에 대한 지리 교육 지침이 제시된 이후, 지리 교육 5 주제가 지리를 가르치는 대부분의 교사와 학생들에게 지리 내용 조직과 학습을 위한 강력하고도 유용한 도구로 활용되고 있다. 반면, NAEP 지리 평가 틀의 세 가지 내용 영역은 NAEP의 지리 평가로 측정되어야 할 주제와 세부 사항을 명확히 하기 위해 제공되고 있다. 세 가지 내용 영역은 학생들의 학습 결과를 측정하는 평가 시행과 결과에 대한 보고를 구조화하기 위해 설계되었다고 볼 수 있다(그림 1 참조, NAGB, 2010: 8-10).



< 그림 2 > 지리교육 5 주제와 NAEP 지리과 3대 학습 내용 영역

NAEP 지리 평가 틀의 세 내용 영역은 각각 주요 평가 내용과 구체적인 평가 요소의 사례, 그리고 학습 후 도출될 수 있는 결과로 구조화 될 수 있다(표 1 참조).

<표 1> NAEP 지리 평가 틀의 내용 영역별 구조

구분	주요 평가 요소	사례	학습 결과
공간과 장소 Space and Place	<ul style="list-style-type: none"> 특정 장소의 위치와 정보, 그 장소에 영향을 미치는 이슈 및 원인의 파악과 관련된 지리 지식 	<ul style="list-style-type: none"> 지표면에서 지진이 일어나는 지역의 위치나 다국적 기업의 공장과 작물 재배 지역, 기후 지역의 분포 등 	<ul style="list-style-type: none"> 지표면의 공간적 패턴과 그것을 형성하는 자연 및 인문 환경적 프로세스에 대한 공간적 관점 함양 자신과 시민적 삶에서 직면 할 문제 이해 및 해결에 유용
인간과 환경 Environment and Society	<ul style="list-style-type: none"> 자연환경과 인간의 거미줄 같이 복잡한 상호작용 속에서 어떠한 방법으로 인간이 자연환경에 의존하거나 적응, 또는 자연환경을 변화시키는지에 대한 이해 	<ul style="list-style-type: none"> 바람에 의한 침식 감소를 위해 나무를 심는 것과 같이 인간에 의한 자연의 변형이 가져오는 긍정적 결과, 또는 지하수원지 위에 쓰레기 매립지를 설치함에 따른 부정적인 결과 등 	<ul style="list-style-type: none"> 인간과 환경의 상호작용과 관련된 지식을 통해 이로 인한 긍정적 또는 부정적인 잠재적 결과에 대해 학생들이 이해할 수 있음
공간 관계와 공간 변화 Spatial Dynamics & Connections	<ul style="list-style-type: none"> 인간과 장소간의 상호 작용의 결과로서 공간 관계 및 공간 변화에 대한 지리적 관점과 사고력 	<ul style="list-style-type: none"> 공간 관계에 영향을 미치는 요소: 문화와 경제적 다양성, 정치적 변화와 긴장 상황, 이주, 여행, 관광, 사상(이데올로기)과 기술적 혁신, 전염병의 확산 등 	<ul style="list-style-type: none"> 공간 관계를 탐구하고 이해하는 즉, 거리상으로는 먼 지역일지라도 공간적으로는 연결된 위치일 수 있다는 공간적 관계의 영향력을 예측하는 능력 함양

(NAGB, 2010: 13-15)

각 내용 영역별 문항 분포 비율을 보면, 약 40%가 ‘공간과 장소’에 대한 내용 범위를 평가하고, ‘인간과 환경’, ‘공간 관계와 공간 변화’ 내용 범위를 각각 30%의 비율로 평가한다. 여기에서 ‘공간과 장소’ 영역이 40%로 높은 것은 이와 관련된 기본 어휘, 지식, 기능이 지리의 기초를 형성하기 때문이라고 밝히고 있다(NAGB, 2010: 9). 이러한 NAEP 지리 평가 내용 영역별 문항 배분 비율은 미국의 K-12 교육 과정에서 지리 교육의 현실을 반영한다고 볼 수 있다. 즉, ‘공간과 장소’ 내용 영역의 문항 비율이 높은 것은 대부분의 교사가 다른 두 내용 영역에 비해 ‘공간과 장소’ 영역에 더 강한 기반을 갖추고 있기 때문이다. 더구나 공간과 장소에 대한 개념들은 지리교육 5 주제에서 가장 먼저 제시되는 두 가지, 즉 위치와 장소에 의해 명확하게 표현되는데, 특히 이 두 가지 주제는 지리 학습의 기초가 될 뿐만 아니라 대부분의 지리 교사들에게 익숙한 내용이기 때문이기도 하다. ‘인간과 환경’, ‘공간 관계와 공간 변화’의 평가 내용에 대해서는 현대사회 현상과 관련하여 더 많은 무게를 둘 필요성을 반영하여 두 영역 모두 30%의 문항 출제 비율을 유지하고 있다. 이는 학생들이 현대 생활의 복잡한 문제를 해결하기 위해서는 이 두 영역 내에서 다루어지는 내용을 반드시 알아야만 하기 때문이다.

한편, NAEP의 행동 영역은 학습자의 지적 행동 목표를 지식 그 자체와 지식에 대한 기능으로 구분한 Bloom의 인지적 영역을 거의 유사하게 적용하고 있다. 따라서 NAEP 지리 평가 틀의 행동 영역은 지식(knowing), 이해(understanding), 적용(applying)의 세 가지 범주로 구성되며, 이들 역시 각각 주요 평가 내용과 구체적인 평가 요소의 사례, 그리고 학습 후 도출될 수 있는 결과로 구조화 될 수 있다(표 2 참조).

< 표 2 > NAEP 지리 평가 틀의 행동 영역별 구조

구분	주요 평가 요소	사례	학습 결과
지식 (Knowing)	무엇인가? 어디인가? • 지리교과의 '개념'과 '지식' • 요구되는 기능: '관찰 기능'과 '기록 기능'	• 장소의 이름이나 특정 국가의 고유 자원을 기억하거나, 여러 국가 간의 무역 패턴에 대한 정보를 찾을 수 있는 능력	• 경관에 대한 차이점을 관찰 할 수 있거나(알거나) 그것에 대한 질문에 사실적 정보 수준의 지식으로써 답할 수 있음
	왜 여기에 있는가? 어떻게 갈 것인가? 그 의미는 무엇인가? • 정보의 의미와 맥락 이해, 정보 간의 관계 파악, 해당 정보를 사용하여 지표면의 패턴과 프로세스 설명 능력	• 육지와 해양의 기온 차가 발생하는 것에 대한 원리를 설명 할 수 있거나, 아열대 해안 지역에서 왜 낮에는 육지 쪽으로, 밤에는 바다 쪽으로 바람이 부는지 이해할 수 있는 능력 등	• 지리적 현상에 대한 개념과 원리 및 지리적 기능을 사용하여 과제 해결
적용 (Applying)	지식과 이해는 지리 문제를 해결하는데 어떻게 사용할 수 있나? • 분류, 가설 설정, 귀납적·연역적 추리, 문제 해결 모형 사용 • 고차적 사고력 요구	• 산성비 문제의 해결책 마련 과정 • 특정 도시에 가장 적합한 폐기물 처리 프로그램 검토 • 관련 분야 또는 경쟁 요인과의 이해 관계에 대한 평가검토를 통해 공항과 발전소 건립에 유리한 입지 요건 분석 • 학교나 이웃 상점으로 가는 경로 구상과 같은 일상생활 문제의 해결을 위한 지리적 분석 기능 사용	• 문제에 대한 종합적 이해와 실행 가능한 해결책을 제안하는 과정에서 다양한 지리적 도구와 기능을 사용할 수 있음

(NAGB,2010: 37-38)

지리과 평가 틀의 행동 영역은 서답형 문항에서 각 영역의 평가 요소에 따라 서로 다른 사고 과정과 기능을 토대로 답안 작성을 요구한다(장의선, 2012: 948). 따라서 행동 영역별 평가 요소는 서답형 문항 제작뿐만 아니라 채점에 대해서도 중요한 영향을 미친다. 어떠한 행동 영역을 평가하느냐에 따라 문항 구성 방법과 채점 가이드의 내용이 달라지기 때문이다. '지식' 영역의 경우는 매우 명확하고 간결할 수 있지만 '이해'와 '적용' 영역으로 갈수록 위 사례에서도 알 수 있듯이 문항 구성에는 더 많은 고민과 절차가 필요해지는 한편, 채점 가이드의 내용은 더 복잡하고 다양한 고려 사항들이 제시될 수 있다. 왜냐하면 학생들은 '지식' 영역에 대한 명확한 답안보다는 자신의 사고를 표현하는 '이해'와 '적용' 영역에서 답안 작성의 자율성을 상대적으로 더 넓게 보장 받기 때문이다. 따라서 평가 틀의 내용 영역과 행동 영역은 서답형 문항의 제작뿐만 아니라 채점에서도 매우 중요한 고려 사항이 된다.

2. 서답형 문항 답안과 채점 사례

본 논문에서는 NAEP 2001, 2010년에 공개된 초등학교 4학년 대상 서답형 18개 문항⁴⁾을 분석 대상으로 삼았으며, 점수 부여 단계에 따른 대표적 사례를 제시하고자 한다. 서답형 문항에 대한 점수 부여 단계는 위에서 언급했듯이 여러 단계가 있지만 지리과에서는 주로 3단계를 가장 보편적으로 사용하는 것으로 나타났다. 즉, 분석 대상 18개 문항 중 12문항(66.7%)이 모두 3단계의 점수로 채점 되었으며, 4단계는 5문항

4) 현재 NAEP에서 공개되고 있는 문항은 1994, 2001, 2010년에 시행되었던 문항의 일부이다. 본 연구에서는 NAEP의 평가 틀과 채점 방안에 대한 부분적 수정·보완이 이루어진 후 시행된 2001년, 2010년 공개 문항을 대상으로 분석하였다.

(27.8%), 나머지 1문항(5.5%)이 2단계 점수로 채점되어 이러한 특징을 잘 보여준다.

3단계의 점수 부여는 주로 '목표 도달'(Complete), '보통'(Partial), '수준 미달'(Unacceptable)로 제시되고, 4단계는 '목표 도달'(Complete), '근접 도달'(Acceptable), '보통'(Partial), '수준 미달'(Unacceptable), 2단계는 '목표 도달'(Complete), '수준 미달'(Unacceptable)로 제시되고 있다. 흥미로운 점은 점수 부여 단계가 많을 수록 고난도 문항이라고 예측하기 쉬운데 실제 난도를 살펴보면 비례관계는 성립하지 않았고, 오히려 2~3 단계의 점수를 부여하는 문항에서 고난도 문항이 많았다.

1) 2단계 점수 부여 채점 사례

G7 12번이 2단계 점수로 채점된 문항 사례이다. 문항 유형으로는 단답형으로 분류되며 '공간과 장소' 영역에서 '지식' 기능을 평가하고 있다. 즉, 지리부도의 아시아 지도를 참고하여 내륙 국가의 이름을 쓰도록 구성하였다. 다만 정답률이 23.1%였던 만큼 학생들이 상당히 어려워했던 문항으로 판단되는데 이는 도해력과 함께 발문에 제시된 'landlocked'라는 용어가 초등학교 4학년 학생들에게는 다소 생소하였기 때문인 것으로 추정된다(표 5, 표 6 참조).

< 표 5 >. 2단계 점수 부여 문항과 채점 상세 사례

문항 예시	
<p>Questions 11–12 are based on the map of Asia on pages 78 and 79 of the atlas.</p> <p>12. Look again at the map of Asia. On the line below, write the name of an Asian country that is landlocked.</p> <p>_____</p>	
채점 가이드의 상세 내용	
<p>폴이의 단서 학생들은 지도에서 아시아의 내륙 국가를 파악할 수 있다.</p>	
점수 단계	내용
목표 도달	답안이 내륙 국가를 올바르게 파악함.
수준 미달	답안이 내륙 국가를 파악하지 못함.
모범 답안 예시	아프가니스탄, 아르메니아, 아제르바이잔, 부탄, 카자흐스탄, 키르기스스탄, 라오스, 몽골, 네 팔, 타지키스탄, 투르크메니스탄, 우즈베키스탄
유의 사항	하나 이상의 국가를 답안에 제시한 경우, 각 국가는 내륙 국가이어야 함.

5) NAEP에서는 '근접 도달' 단계의 점수 부여 용어를 'Essential', '수준 미달' 단계의 점수 부여 용어를 'Inappropriate'로 사용하다가 2010년부터 각각 'Acceptable', 'Unacceptable'로 변경하여 사용하고 있다. 따라서 본 연구에서도 'Acceptable, 'Unacceptable'로 통일하여 사용하고자 한다.

< 표 6 > 각 점수 단계별 학생 답안과 채점 결과 해설

목표 도달 단계의 학생 답안

12. Look again at the map of Asia. On the line below, write the name of an Asian country that is landlocked.

Mongolia

채점 결과 해설: 아시아에서 내륙 국가의 이름을 올바르게 작성하였다.

수준 미달 단계의 학생 답안

Questions 11-12 are based on the map of Asia on pages 78 and 79 of the atlas.

12. Look again at the map of Asia. On the line below, write the name of an Asian country that is landlocked.

Pakistan

Questions 11-12 are based on the map of Asia on pages 78 and 79 of the atlas.

12. Look again at the map of Asia. On the line below, write the name of an Asian country that is landlocked.

Russia China KAZAKHSTAN

채점 결과 해설: 첫 번째 답안의 파키스탄은 해안이 아라비아 해(Arabian Sea)에 접해 있다. 두 번째 답안에서 내륙 국가의 이름인 카자흐스탄이 제시되어 있지만 다른 두 국가는 아니기 때문에 학생은 내륙국가라는 개념을 정확하게 이해하고 있지 못하다.

2) 3단계 점수 부여 채점 사례

가장 많은 비중을 차지하는 3단계의 점수 수준을 가진 문항 사례로는 C5 8번 문항의 채점 상세 내용과 학생 답안을 살펴보고자 한다. C5 8번 문항은 ‘공간 관계와 공간 변화’에 해당하는 내용영역에서 ‘적용’ 능력을 측정하고자 한 단답형 문항이다. 즉 지도 자료를 토대로 해당 지역의 새로운 공항 입지에 대한 의사 결정 능력을 평가하도록 구성되었는데 이러한 구성은 지리도해기능을 평가하는 방안에도 의미 있는 시사점을 줄 수 있다. 난도는 다소 평이한 중난도(정답률 약 47%) 문항이었으며, 문두의 술어에 ‘그리시오(Draw)’와 ‘설명하시오(Explain)’가 함께 사용되었다는 점이 특징적이다(표 7 참조).

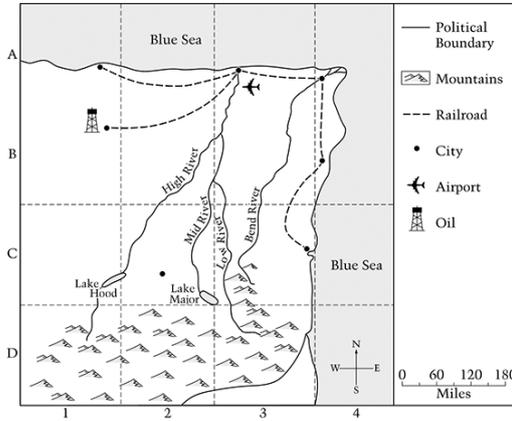
채점 상세 내용에 따라 ‘목표 도달’ 단계의 점수를 부여받은 학생 답안을 살펴보면 <표 8>과 같다. 답안에서는 채점 상세에 제시된 바와 같이 지리적으로 적절한 장소에 공항을 입지시키고 그 입지 선택에 대해서도 논리적인 방법으로 설명하고 있다. 즉 사람들이 휴양지로 여행하기 위해 그들이 사는 곳에 공항을 위치시키기 때문에 공항이 도시 근처에 입지해야 함을 논리적으로 잘 설명하고 있다(표 8 참조).

‘보통’ 단계의 학생 답안 사례 두 가지는 모두 공항의 입지는 점수가 부여 될 수 있지만 설명이 논리적이 지 못해 점수가 인정되지 않은 경우이다(표 9 참조).

<표 7>. 3단계 점수 부여 문항과 채점 상세 사례

문항 예시

Questions 6 - 8 are based on the map below.



8. Suppose a second airport were going to be built in this area. Draw an airport symbol on the map in the place where you think the new airport should go.

Then, in the space below, explain why you put the second airport where you did.

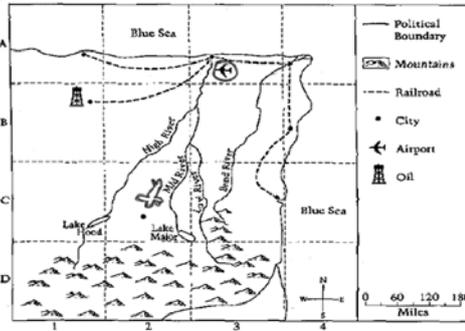
채점 가이드의 상세 내용

점수 단계	내용
목표 도달	지리적으로 적절한 장소에 공항을 입지시키고, 입지 선택을 논리적인 방법으로 설명함.
보통	a) 지리적으로 적절한 장소에 공항을 입지시켰지만, 입지 선택에 대해 비논리적으로 설명하거나 설명을 하지 않음. 또는, b) 공항을 적절한 위치에 입지시키지 못했지만 입지 선택에 대해 논리적으로 설명함
수준 미달	지리적으로 적절한 장소에 공항을 입지시키지 못함. 또는 입지 선택에 대해 비논리적으로 설명함.
모범 답안 예시	<p>지리적으로 적절한 위치(Geographical Appropriate Place):</p> <ol style="list-style-type: none"> 바다(Blue Sea), 산맥, 비행기의 동체와 날개가 철도, 유전(oil rig), 강, 호수에 걸쳐있는 경우를 제외한 모든 장소 공항의 위치는 도시와 겹쳐질 수 있음. <p>설명(Explanations):</p> <ol style="list-style-type: none"> 도시에 접근하기 위해서 해변, 산악지역, 호수 또는 하천과 같은 휴양지에 접근하기 위해서

< 표 8 > 3단계 점수 중 목표 도달 단계의 학생 답안과 채점 결과 해설

목표 도달 단계의 학생 답안

8. Suppose a second airport were going to be built in this area. Draw an airport symbol on the map in the place where you think the new airport should go. Then, in the space below, explain why you put the second airport where you did.



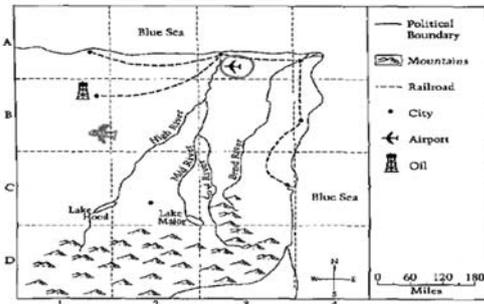
I put it by a city
so the people who live
there have an airport
if there going to travel.

채점 결과 해설: 지리적으로 적절한 장소에 공항을 입지시키고, 입지 선택을 논리적인 방법으로 설명함

< 표 9 > 3단계 점수 중 보통 단계의 학생 답안과 채점 결과 해설

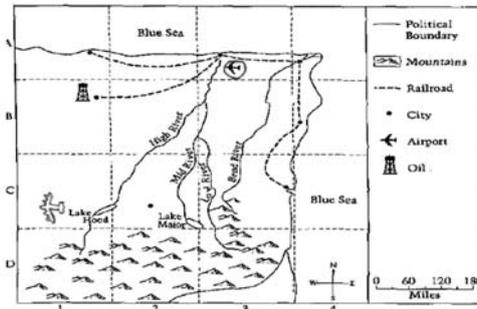
보통 단계의 학생 답안

8. Suppose a second airport were going to be built in this area. Draw an airport symbol on the map in the place where you think the new airport should go. Then, in the space below, explain why you put the second airport where you did.



I put the second airport there because
there are no mountains or rivers in
the way to fly off.

8. Suppose a second airport were going to be built in this area. Draw an airport symbol on the map in the place where you think the new airport should go. Then, in the space below, explain why you put the second airport where you did.



I put the second airport there because
the planes from the other airport might
need to refuel and people by the second
airport might need a lift.

채점 결과 해설: 두 답안의 설명은 정답으로 인정되지 않는다. 첫 번째 답안에서는 강이나 산이 항공기가 비행하는 것과 어떻게 관련되어 있는지에 대한 설명이 없기 때문에 비행장을 그곳에 입지시킨 이유로서 점수를 부여할 수 없다. 두 번째 답안에서는 다른 공항을 출발해서 오는 비행기는 이 공항까지의 비행 거리가 180마일이 되지 않기 때문에 연료를 다시 주입할 필요가 없고 '이 공항 근처의 사람들이 이용할 필요가 있다'는 서술은 비행장이 왜 그곳에 입지해야 하는가에 대한 설명으로 충분하지 않다.

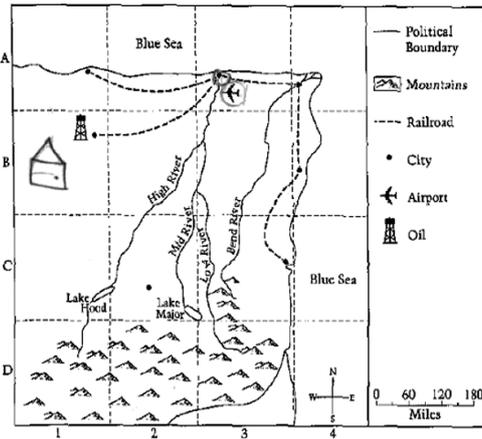
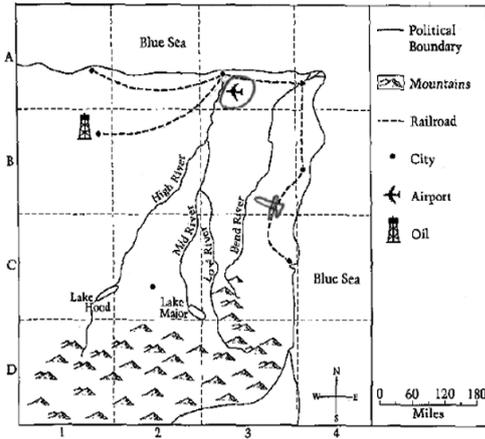
‘수준 미달’ 단계의 학생 답안 사례는 공항의 기호를 사용하지 않거나 공항에 대한 입지를 지리적으로 찾지 못하고 공항 입지에 대한 설명도 논리적이지 못한 경우이다(표 10 참조).

< 표 10 > 3단계 점수 중 수준 미달 단계의 학생 답안과 채점 결과 해설

수준 미달 단계의 학생 답안

8. Suppose a second airport were going to be built in this area. Draw an airport symbol on the map in the place where you think the new airport should go.
Then, in the space below, explain why you put the second airport where you did.

8. Suppose a second airport were going to be built in this area. Draw an airport symbol on the map in the place where you think the new airport should go.
Then, in the space below, explain why you put the second airport where you did.



I put the airport there because it is by the other side of the blue sea

because it is a lot of place there.

채점 결과 해설: 첫 번째 답지에서는 비행기가 철로 위에 위치하므로 점수를 부여할 수 없다. 또한 ‘Blue Sea’의 반대편으로 날아가야 하기 때문이다’라는 서술은 공항의 입지 이유로서 볼 수 없다. 두 번째 답지에서는 ‘짐’ 기호가 사용되고 있어 공항으로 인정할 수 없다. ‘많은 장소가 있다’라는 기술은 공항이 그 곳에 입지해야 하는 이유를 설명하는 내용이 아니므로 점수를 부여할 수 없다.

3) 4단계 점수 부여 채점 사례

4단계의 점수 수준을 가진 문항들은 대부분 단순 선지식보다는 개념에 대한 이해와 적용 능력을 평가 요소로 할 뿐만 아니라 지도에 대한 해석과 활용을 함께 해야 문항 풀이가 가능하도록 구성된 점이 특징적이었다. 이들 중 가장 어려운 문항 중 하나인 G8 15번(정답률 25.9%)의 경우를 사례로 살펴보면 먼저 평가 내용 영역은 ‘공간과 장소’, 행동 영역은 ‘적용’ 능력에 해당한다. 그러나 답안 작성과정에서는 축척과 범례 개념에 대한 이해뿐만 아니라 주어진 조건에 맞는 마을 지도를 직접 작성하도록 하는 도해기능 평가까지 포함하고 있어 학생들이 많이 어려워했던 문항으로 추정된다(그림 3 참조). 따라서 G8 15번은 평가 요소가 다양한 만큼 채점 상세 내용 역시 상당히 많은 내용을 제시하고 있었다. 뿐만 아니라 학생들이 작성한 다양한 답안에서 점수 부여에 대한 판단이 필요할 것으로 예상되는 상황에 대해 매우 치밀하게 준비한 점은 우리나라 서답형 문항 채점 기준과 비교할 때 매우 의미 있는 시사점을 주고 있다(표 11 참조).

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runt River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

< 그림 3 > 4단계 점수 부여 유형 사례: G8 15번

< 표 11 > G8 15번 문항에 대한 점수 단계와 채점 상세 내용

점수 단계	내용
목표 도달	범례에 있는 4가지 요소를 모두 정확한 위치에 표시하고, 정확한 방향으로 축척에 맞게 길이와 넓이를 표현함.
근접 도달	범례에 있는 4가지 요소를 모두 정확한 위치에 표시하였지만 축척을 틀리게 표현함. 범례에 있는 4가지 요소 중 3가지를 정확한 위치에 표시하면서 축척을 정확하게 표현함.
보통	범례에 있는 4가지 중 1~2개 요소를 정확한 위치에 표시하고 축척을 정확하게 표현함. 또는, 범례에 있는 4가지 중 3개의 요소를 정확한 위치에 표시하면서 축척을 틀리게 표현함.
수준 미달	범례에 있는 4가지 요소를 모두 틀린 위치에 표시하고, 축척과 방향도 모두 틀리게 표현함. 또는, 축척만 바르게 표현함. 또는, 범례에 있는 4가지 중 1~2개의 요소를 정확한 위치에 표시하였지만 축척을 틀리게 표현함. 물론 이 단계의 학생들도 문장으로 된 정보를 지도로 표현할 능력이 있음을 보여주려 하겠지만 정확한 표현이 아니면 점수를 받을 수 없음.
채점 시 유의 사항	경계 <ul style="list-style-type: none"> • 일부 지그재그 형태의 경계가 있어도 허용함(한 블록의 절반 이내에서) • 만약 경계를 그리지 않았다면, 문항에 제시된 전체 격자망을 4구역(북서 구역, 남서 구역, 북동 구역, 남동 구역)으로 나누고 있는 것이라고 판단해야 함. • 만약 학생이 표현한 범례 요소가 경계 밖에 그려져 있다면 전체 격자망을 경계로 간주하고 있다고 판

<p>단해야 함(이 경우 축척 표현은 틀린 것임). 이 경우 최대 점수 3.</p> <ul style="list-style-type: none"> • 만약 경계가 그려져 있지는 않지만 학생이 분명히 경계를 인식하고 있다고 간주된다면, 다시 말해서 축척이 정확하고 범례에 있는 4가지 요소가 정확하게 표시되어 있다면 점수는 3. • 경계는 반드시 경직된 선일 필요는 없음: 경계는 실선일 수도 있고, 도로나 강 끝에 그려진 선일 수도 있으며, 네 모퉁이에 있는 점일 수도 있음. • 만약에 경계선이 정확한 거리보다 더 길게 그려져 있다 하더라도 마을의 사방의 경계가 제 위치에 그려져 있다면 감점하지 않음. <p>학교와 공원</p> <ul style="list-style-type: none"> • 학교와 공원의 기호는 한 칸의 격자망의 1/4의 반보다 크게 그려야 한다. 왜냐하면 이 기호들은 동, 서, 남, 북 방향을 고려할 필요가 없기 때문임. • 기호의 반 이상이 경계 안에 그려져 있어야 함. • 동일한 기호가 1회 이상 사용되고 있다면 점수에 반영하지 않음. • 답안에 기호 대신 '학교' 또는 '공원'이라고 글로 쓰여 있다면 점수를 부여함. <p>도로와 하천</p> <ul style="list-style-type: none"> • 도로와 하천은 마을의 한쪽 끝에서 다른 쪽 끝까지 그려져 있어야 함(만약 끝 부분에서 격자의 반만 그리고 있어도 인정함). 만약 도로와 하천의 선들이 블록 끝의 경계를 지나 더 길게 그려져 있다면 점수는 3. • 도로와 하천을 같은 방향으로 겹쳐 그리고 있더라도 인정함. • 도로와 하천을 추가로 더 그리고 있더라도 Runt River와 Main Street로 생각되는 것들이 적절하게 그려져 있다고 간주된다면 점수를 부여할 수 있음. • 도로가 전체적으로 동서 방향으로 그려져 있다면 약간의 남북 방향이 표현되어 있더라도 인정함. • 도로는 첫 번째와 마지막 블록에만 그리는 것이 아니라 반드시 해당되는 모든 블록 안에 그려져야 함. 그렇지 않으면 점수를 부여할 수 없음. • 도로는 한 가닥의 실선으로 그려져 있더라도 인정함 • 하천은 동서 방향보다는 상대적으로 남북 방향으로 그려져야 함. • 하천은 첫 번째와 마지막 블록에만 그리는 것이 아니라 반드시 해당되는 모든 블록 안에 그려져야 함. 그렇지 않으면 점수를 부여할 수 없음. • 만약 하천이 물결 모양의 실선으로 표현되어 있더라도 하천 기호로 간주함. <p>다음은 인정하지 않음</p> <ul style="list-style-type: none"> • Main Street이나 Runt River가 기호로 표현되지 않았거나, • 학교 및 공원의 위치를 연결한 가상의 선과 평행한 경우(학교나 공원은 반드시 북동쪽 혹은 남서쪽에 그려져야 함).

위의 채점 상세 내용에 따라 '목표 도달' 수준의 점수를 부여받은 학생 답안 사례와 채점 결과 해설은 <표 12>와 같다. 채점 상세 내용에 제시된 것과 같이 문항에 제시된 모든 조건을 충족하는 답안 사례들이 다(표 12 참조).

G8 15번 문항에서 점수 수준이 '근접 도달' 단계인 학생 답안과 채점 결과 해설은 <표 13>과 같다. 특히 각 답안에 대한 채점 결과 해설에서 해당 점수 수준을 부여한 이유에 대하여 상세하게 설명하고 있어 점수를 받은 학생뿐만 아니라 시험에 대비해야 하는 교사와 학생의 이해를 돕고 있는 점이 인상적이다(표 13 참조).

'보통' 단계의 학생 답안은 해당 점수 수준에서 볼 수 있는 가장 전형적인 사례를 제시하고 있는데, 이는 매우 상세하게 설명되어 있는 채점 결과 해설을 통해서도 알 수 있다(표 14 참조).

< 표 12 > 4단계 점수 중 목표 도달 단계의 학생 답안과 채점 결과 해설

목표 도달 단계의 학생 답안

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runt River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Key

- S School
- Street
- P Park
- River

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runt River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Key

- S School
- Street
- P Park
- River

채점 결과 해설: 두 답안은 모두 정확한 범례의 기호와 축척을 사용하여 글로 된 정보를 지도로 잘 옮기고 있다. 두 번째 답안에서 Runt River는 남북 방향으로 곧장 그려져 있지는 않지만 이런 유형의 답안은 점수를 부여할 수 있다. 왜냐하면 방위가 북서-남동 방향까지는 전환되고 있지 않기 때문이다.

< 표 13 > 4단계 점수 중 근접 도달 단계의 학생 답안과 채점 결과 해설

근접 도달 단계의 학생 답안

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runt River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Key

- S School
- Street
- P Park
- River

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runt River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Key

- S School
- Street
- P Park
- River

채점 결과 해설: 첫 번째 답안에서 학생은 마을에 범례에 제시된 4가지 모든 요소를 정확히 제 위치에 그리고 있다. 그러나 마을은 축척에 맞게 그려져 있지 않고 사각형도 아니다. 두 번째 답안에서는 마을이 축척에 맞게 그려져 있고, Main Street과 학교, 공원이 적절하게 그려져 있다. 하지만 Runt River는 남북 방향이 아니라 동서 방향에 가깝게 그려져 있으므로 점수를 부여할 수 없다.

< 표 14 > 4단계 점수 중 보통 단계의 학생 답안과 채점 결과 해설

보통 단계의 학생 답안

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runk River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Scale = 1 mile

Key: School (S), Street (—), Park (P), River (≡)

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runk River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Scale = 1 mile

Key: School (S), Street (—), Park (P), River (≡)

채점 결과 해설: 두 답안은 '보통' 단계의 점수 수준을 보여주는 대표적인 사례이다. 첫 번째 답안은 축척을 표현하고 있지 않고 전체 격자망을 사용하고 있다. 이 경우 정답을 완성하는 요소로서 축척에 대한 표현이 누락되어 있으므로 완전한 답안 내용 중 3개의 요소만 포함되어 있는 것이다. 학교와 공원은 정확한 위치의 격자에 표현되어 있고 Main Street은 동서방향으로 그려져 있다. 하천은 정확하지 않고 건물들은 무시되어 있다. 두 번째 답지는 축척을 표현하고 있으며 Main Street과 하천도 정확하게 그려져 있다. 그러나 공원과 학교는 위치가 잘못 표시되어 있다.

'수준 미달' 단계의 답안에서는 해당 학생이 축척과 범례에 대한 개념 이해뿐만 아니라 지리도해기능이 상당히 미흡한 수준임을 알 수 있다. 채점 결과 해설에서는 이러한 상황에 대해서도 매우 상세하게 이유를 설명하고 있는 것으로 나타났다(표 15 참조).

< 표 15 > 4단계 점수 중 수준 미달 단계의 학생 답안과 채점 결과 해설

수준 미달 단계의 학생 답안

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runk River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Scale = 1 mile

Key: School (S), Street (—), Park (P), River (≡)

15. LITTLE TOWN

- Width: 4.0 miles east to west
- Length: 3.0 miles north to south
- Main Street runs east to west through the town.
- The school is on the northeast side of town.
- Phelps Park is on the southwest side of town.
- Runk River runs north to south through the town.

On the grid below, each square is one mile wide and one mile long. Draw a map of Little Town on the grid. Draw the town's borders. Then, use the symbols in the key below to draw the features listed above.

Scale = 1 mile

Key: School (S), Street (—), Park (P), River (≡)

채점 결과 해설: 이 두 답안을 작성한 학생은 상당한 혼란에 빠져 있음을 볼 수 있다. 첫 번째 답안에서는 축척이 없고 여러 개의 학교와 공원, 하천을 그리고 있다. 두 번째 답지에서는 마을이 집에서 출발하여 하천과 공원을 거쳐 학교에서 끝나는 식으로 마치 도로처럼 표현되어 있다.

IV. 요약 및 시사점

NAEP의 채점 시스템과 답안 채점 사례를 통해 다음과 같은 시사점을 얻을 수 있었다.

첫째, 채점의 신뢰성 확보 방안에 대한 시사점이다. 평정자들의 채점 전문성 제고를 위한 훈련 시스템의 상설화와 지속적인 채점 보정 작업, 그리고 시계열적 일관성을 위한 채점 점검 등 NAEP의 채점 시스템에서는 채점의 신뢰성을 확보하기 위하여 다양한 방법을 지속적으로 실행하고 있었다.

둘째, 치밀한 채점 상세 내용을 제시한 채점 가이드의 개발에 대한 시사점이다. 키워드 중심의 기계적 채점 방식이 아니라 학생들이 답안에서 나타날 수 있는 다양한 변수들을 고려하여 채점 상세 내용을 치밀하고 구체적으로 제시하고 있었다. 채점 상세의 내용이 간략하고 보기 쉽도록 구성된 것이 아니라 학생들이 작성할 수 있는 답안 내용에 대한 다양한 경우를 여러 가지 제시하다보니 자칫 복잡하고 채점을 곤혹스럽게 할 수도 있지만 지속적인 평정자 훈련을 통해 이러한 점 역시 극복 가능 했을 것으로 사료된다.

셋째, 문항 유형에 따른 점수 단계 구분과 각 점수 단계별 채점 사례 및 채점 결과 해설 제공에 대한 시사점이다. 2~4단계로 구분되는 다양한 점수 단계가 있어 문항 유형에 따라 적합한 채점 가이드의 상세 내용을 개발 할 수 있을 뿐만 아니라 해당 점수 부여에 대한 근거를 채점 결과 해설에 제공하고 있어 시험을 준비하는 학생과 교사에게 유용한 자료로 활용되고 있다.

우리나라의 경우에도 일부 교육청 수준이나 국가수준 학업 성취도평가에서의 서답형 평가 도입으로 채점에 대한 많은 논의와 발전이 이루어져 왔다. 그러나 상설화 된 평정자 조직과 전문적 훈련이 미비한 상태에서 채점과 부여 점수에 대한 신뢰도는 확신하기 어렵다. 물론 전국의 교사를 대상으로 한 인력풀 구축 등으로 평정자의 자격 조건을 어느 정도 담보하고 있지만 채점 전문성 확보를 위한 NAEP 평정자의 자격 요건과 훈련 시스템에 비추어 본다면 평정자 전문성 확보와 관리를 위한 다양한 방안과 노력이 더욱 필요하다고 판단된다.

한편, 채점 시스템의 체계화라는 명목 하에 대규모 컴퓨터 채점 시스템을 단기간 안에 무리하게 추진하고 있음을 알 수 있다. 일례로 지난 국가수준 학업 성취도평가의 채점 시 발생했던 컴퓨터 전산 시스템이 안정적이지 못해 많은 채점자들이 혼란스러워 하기도 하였다. 또한 키워드 중심의 기계적 채점 경향에도 문제가 있다. 현재 우리나라의 교육 현실에서는 대부분의 서답형 문항 채점에서 학생들의 답안에서 나타날 수 있는 다양한 경우의 수를 고려하지 않고 단지 몇 개의 키워드 중심으로 점수를 산정하는 경향이 있다. 이 같은 맥락에서 대규모 서답형 평가에 대해서도 채점자가 사람이 아닌 컴퓨터 시스템으로 가능할 것이라는 연구도 이목을 끌고 있다.

그러나 지금까지 여러 교육자들에 의해 논의되어 왔듯이 서답형 문항의 본래 취지는 학생들의 사고 기능 평가에 있음을 견지해야 할 것이다. 몇 개의 키워드를 중심으로 이루어지는 채점은 학생의 답안이 정답인지, 의사답안인지, 오답인지, 일부만 적절한지 등의 측면에서 서답형 문항의 채점 방식으로는 근본적인 문제가 있음을 고려해야 한다. 따라서 대규모의 서답형 답안에 대한 채점이 정확하고 신속하게 이루어지기 위해서는 많은 시간과 노력이 요구되겠지만 앞서 언급한 전문적인 평정자 조직과 훈련, 문항 유형에 따른 치밀한 채점 가이드의 개발 등 채점 시스템의 체계적이고도 지속적인 발전이 필요하다 하겠다.

< 참고 문헌 >

- 교육과학기술부(2011). 중등학교 학사관리 선진화 방안. 교육과학기술부 보도자료(2011. 12. 13).
- 양길석(2006). 서술형·논술형 평가 문항의 제작. 서술형·논술형 평가 예시문항 자료집(심화선택 일반사회 경제/사회·문). 3-26. 서울특별시교육청.
- 장의선(2012). 지리과 서답형 문항의 주요 유형에 관한 연구-NAEP의 지리과 4학년 문항을 사례로-. 대한지리학회지. 47(6). 934-954.
- C. Keith Waugh·Norman E. Gronlund(2013). *Assessment of Student Achievement(10ed)*, Pearson, New Jersey.
- Kan, A.(2007). An Alternative Method in the New Educational Program from the Point of Performance-based Assessment Rubric Scoring Scales. *Educational Sciences: Theory & Practice*, 7(1), 144-152. Retrieved from Education Research Complete database.
- NAGE(2010). *Geography Framework for the 2010 National Assessment of Educational Progress*.
- NCES(2011). *The Nation's Report Card: Geography 2010*, Institute of Education Science, U.S. Department of Education, Washington, D.C.(<http://nces.ed.gov/nationsreportcard/>).
- Odendahl, N.(2007). Put Your Tests to the Test. *T H E Journal*, 34(1), 46-47. Retrieved from Education Research Complete database.
- Ory, J.(1983). *Improving Your Test Questions*. Retrieved from ERIC database.
- Popham, W.J.(2006). *Assessment for educational leaders*. Boston, MA: Pearson.
- Schafer, W., Gagné, P., & Lissitz, R.(2005). Resistance to Confounding Style and Content in Scoring Constructed-Response Items. *Educational Measurement Issues & Practice*, 24(2), 22-28. Retrieved from Education Research Complete database.
- Tuckman, B.(1994). Tips for using essay tests. *Education Digest*, 59(5), 71. Retrieved from Education Research Complete database.
- Williams, J.(1991). *Writing Quality Teacher-Made Tests: A Handbook for Teachers*. Retrieved from ERIC database.