

국민과 함께
고민하는

미래 대입 제도

국민참여위원회
속의

일시

2026.08.29.(토) 14:00 ~ 08.30.(일) 15:30

장소

YBM연수원



대통령 소속
국가교육위원회
National Education Commission

자료를 읽기 전에

이번 국민참여위원회 숙의 토론회의 주제는 다음 두 가지입니다.

- ① 미래사회를 대비한 대입제도 개편 방향성
- ② 내신 서·논술형 평가 확대 및 수능 서·논술형 평가 도입 여부

사전학습자료는 국민참여위원님들이 이 주제를 깊이 있게 논의하실 수 있도록, '숙의의 의미와 절차' 및 '현행 대입제도와 서·논술형 평가에 관한 기본 정보'를 미리 안내하고자 만들어졌습니다. 이 자료는 크게 세 부분으로 이루어져 있습니다.

1장. 숙의의 의미와 절차

숙의가 무엇인지, 이번 대입제도 숙의에 활용하는 '숙의형 공론진단' 방식이 어떤 절차로 진행되는지 살펴봅니다.

2장. 대입제도를 논의하는 이유

국가교육위원회가 추진하고 있는 2028-2037 국가교육발전계획, 미래 사회의 변화 모습 등 대입제도를 둘러싼 논의가 다시 시작된 배경을 알아봅니다.

3장. 서·논술형 평가 알아보기

숙의 주제 중 하나인 서·논술형 평가에 대해 알아봅니다. 3장은 아래와 같은 내용이 포함되어 있습니다.

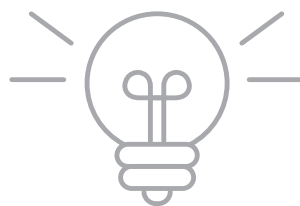
- 서·논술형 평가의 기본 개념, 서·논술형 문항과 선다형 문항의 장·단점
- 고등학교 평가 및 대학수학능력시험에서의 서·논술형 문항 활용 현황
- 고등학교 서·논술형 평가 확대 및 대학수학능력시험 서·논술형 평가 도입 논의의 배경
- 서·논술형 평가의 주요 쟁점
- 서·논술형 평가에서의 AI 활용 가능성

대입제도와 서·논술형 평가에 대해서는 다양한 의견이 제기되고 있습니다. 이에 이 자료집은 여러 입장의 근거가 되는 정보를 균형 있게 전달하고 이해하기 쉬운 표현으로 설명하고자 노력했습니다. 숙의에 참여하시는 국민참여위원께서는 토론회에 참석하시기 전에 이 자료를 충분히 읽고 숙지해 주시기 바랍니다. 사전학습자료집 외에도 관련된 다양한 자료와 정보를 미리 살펴보시면 보다 폭넓은 논의에 도움이 될 것입니다.

“여러분의 소중한 의견이 대한민국 교육의 미래를 만드는 큰 자산이 됩니다.”

1장

숙의·공론화의 의미와 절차



국민과 함께
고민하는

미래
대입
제도

1장. 숙의·공론화의 의미와 절차

1) 숙의(Deliberation)

‘숙의’는 복잡한 사회 문제에 대하여 **정확하고 균형 잡힌 정보**를 바탕으로 **깊이 생각하고**, 다른 사람과 **이성적으로 토론**하여 **공동체를 위한 가장 좋은 방안**을 찾아가는 과정입니다. 이 개념을 연구한 학자들의 설명도 참고할 만합니다.

- 정치철학자 존 롤스(John Rawls)는 개인의 이해관계를 넘어선 ‘공적 이성(public reason)’을 발휘해, 사회 구성원 모두가 받아들일 수 있는 공적 근거를 찾아가는 과정으로 숙의를 설명했습니다.
- 사회철학자 위르겐 하버마스(Jürgen Habermas)는 숙이란 힘이나 권력이 아니라 이성적인 대화 속에서 ‘더 나은 논증의 힘’을 통해 합의에 다다른 과정으로 보았습니다.

2) 공론화(Public consultation)

대의제 민주주의는 국민이 선출한 대표에게 정책 결정 권한을 일정 기간 맡기는 제도입니다. 반면 ‘**공론화**’는 사회적으로 의견이 크게 갈리는 문제나 미래에 중요한 영향을 미칠 정책을 결정할 때, **정부가 국민의 생각과 지혜를 일정 기간 직접 물어 해법을 찾아가는 민주적인 절차**를 말합니다

공론화는 대의제를 대신하는 제도가 아니라 이를 보완하는 절차입니다. 공론화에 참여한 시민들이 내놓은 결론은 그 자체로 법적 구속력을 갖는 최종 결정은 아니지만 정책 결정권자에게 전달되어 정책에 영향을 미치는 ‘권고’의 성격을 가집니다.

이 권고는 대표성을 지닌 시민들이 충분한 정보를 학습하고 토론을 거쳐 만든 ‘숙의된 여론’이라는 점에서 무겁게 받아들여지며, 정부는 이를 존중해 정책에 반영하고자 노력합니다.

3) 왜 하는가?

일반적인 여론조사는 그 순간의 감정이나 단편적인 정보만으로 답하기 쉽다는 한계가 있습니다. 이에 비해 **공론화는 시민들이 객관적인 사실을 먼저 학습**하고, 서로 다른 입장을 듣는 숙의 과정을 거치도록 설계되어 있습니다.

공론화에 참여했던 시민들을 대상으로 한 연구에 따르면, 시민들은 숙의 과정을 거치면서 개인적인 이해관계나 특정 집단의 성향에서 한 걸음 물러나, **사회 전체의 이익과 정책의 타당성을 기준으로 판단하는 경향**을 보인 것으로 나타났습니다.

숙의는 아직 생각이 정해지지 않은 사람에게는 **판단할 수 있는 기준**을 마련해주고, 이미 입장이 뚜렷한 사람에게는 **다른 입장을 이해하고 절차의 정당함을 인정하게 하는 효과**가 있는 것으로 보고됩니다. 이를 통해 사회적 갈등이 완화될 수 있다는 것이 관련 연구들의 공통된 설명입니다.

4) 대표적인 숙의(공론화) 방법

숙의의 방법은 목적과 주제, 참여자의 특성, 예산과 시간 등에 따라 다르게 선택됩니다.

- **공론조사(Deliberative Polling)** : 대표성 있게 뽑은 시민 집단이 학습과 토론을 거치기 전과 후의 의견 변화를 확인하는 방식입니다. 정치학자 제임스 피시킨이 고안했으며, 이 명칭은 스탠퍼드대 숙의민주주의연구소가 등록한 상표입니다.
- **시민의회(Citizens' Assembly)** : 무작위로 뽑힌 시민들이 몇 주에서 몇 달에 걸쳐 학습·숙의한 뒤 정책을 제안하는 방식입니다. 아일랜드 시민의회, 프랑스 기후시민협약 등이 대표적입니다.
- **시나리오 워크숍(Scenario Workshop)** : 다양한 사람들이 모여 앞으로 일어날 수 있는 여러 상황(시나리오)을 그려보고, 각 대안이 타당한지 검토하는 방식입니다.
- **합의회의(Consensus Conference)** : 시민들이 과학기술이나 환경 등 전문적인 주제에 대해 전문가에게 직접 질문하고 확인한 뒤, 합의된 의견을 담은 보고서를 만드는 방식입니다.

이번 대입 토론회에서 활용하는 방법은 제임스 피시킨이 개발한 공론조사 방식을 **우리나라 상황에 맞게 조정**한 '숙의형 공론진단' 방식입니다.

5) 숙의형 공론진단은 어떻게 진행되는가 - 진행절차

- ① **대표성 있는 시민참여단 구성** : 성별, 연령, 의제에 대한 처음 입장 등을 고려해 균형 있게 시민참여단을 구성합니다.
- ② **사전 의견조사** : 숙의를 시작하기 전, 시민들의 기초적인 인식과 선호를 조사합니다.
- ③ **사전 학습 및 전문가 질의응답** : 균형 잡힌 자료집을 학습하고, 서로 다른 입장을 가진 전문가의 발표를 들으며 교차 검증합니다.
- ④ **분임별 소그룹 숙의** : 중립적인 진행자(퍼실리테이터)의 도움을 받아, 모든 참가자가 평등하게 발언하고 서로의 말을 경청합니다.
- ⑤ **사후 의견조사 및 결과 분석** : 숙의를 마친 뒤 최종 판단과 판단이 바뀐 이유를 분석해 정책 결정 과정에 전달합니다.

6) 실제 사례 : 신고리 5-6호기 공론화

2017년 진행된 신고리 5-6호기 공론화는 사회적으로 의견이 크게 갈렸던 국책 사업을 시민 숙의로 다룬 대표적인 사례입니다. 성별, 연령, 의견 분포를 기준으로 무작위로 선정된 시민참여단 500명을 대상으로 2박 3일 종합토론회를 포함한 세 차례의 의견조사가 진행되었습니다.

[표] 신고리 5-6호기 공론화 — 숙의 전후 의견 변화

구분	건설재개	건설중단	판단유보
숙의 전 조사 (471명)	44.7%	30.7%	24.6%
숙의 후 조사 (471명)	57.2%	39.4%	3.3%
변화	+12.5%p	+8.7%p	-21.3%p

집중적인 숙의를 거치는 동안 '판단을 미루겠다'는 응답은 24.6%에서 3.3%로 21.3% 포인트 줄었고, 건설 재개(+12.5%포인트)와 건설 중단(+8.7%포인트)의 응답이 늘었습니다. 어느 한쪽 의견만 늘어난 것이 아니라 양쪽 모두 늘어났다는 점이 특징입니다. **이는 숙의 후에 참여자의 주체성과 책임감이 증가했다는 사실을 보여줍니다.**

숙의 후 조사에서 '판단을 미루지 않고 하나를 선택해 달라'고 묻은 최종 문항에서는 건설 재개 59.5%, 건설 중단 40.5%로 19.0%포인트 차이가 났습니다. 95% 신뢰수준에서 표본오차 ($\pm 3.6\%$ 포인트)를 넘어서는 통계적으로 의미 있는 차이였고, 공론화위원회는 이를 근거로 건설 재개를 정부에 권고했습니다.

7) 이 사례가 보여주는 시사점

첫째, **숙의는 누군가를 설득해 한쪽으로 이끄는 과정이 아니라 판단을 형성해가는 과정**이었다는 점입니다. 종합토론회를 거치며 재개와 중단 응답이 함께 늘고 '판단 유보'만 줄어들었다는 결과는, 숙의가 시민을 어느 한쪽으로 몰아간 것이 아니라 **'아직 판단하지 못한 상태'를 '근거를 가진 판단'으로 바꾸어 놓았다는 것**을 보여줍니다.

둘째, **사람들이 결과를 받아들이게 만든 것은 결론의 내용이 아니라 절차**였다는 점입니다. 시민참여단이 매긴 공론화 과정에 대한 전반적인 만족도는 4점 만점에 3.24점(만족 88.8%)이었습니다. 특히 눈에 띄는 점은, 최종적으로 소수 의견이 된 '건설 중단' 응답자의 만족도(3.28점, 만족 90.4%)가 다수 의견이 된 '건설 재개' 응답자(3.21점, 87.7%)보다 오히려 높았다는 것입니다. 이는 원하는 결론이 나오지 않았더라도 절차가 공정했다고 느끼면 그 결과를 받아들일 수 있다는 것을 보여주는 결과로 해석됩니다.

공론화 이후에도 원자력 정책을 둘러싼 사회적 논의는 계속 이어졌습니다. 이는 공론화가 갈등을 한 번에 없애는 만능 해법은 아니며, **갈등을 좀 더 합리적으로 다룰 수 있는 형태로 바꾸어가는 절차라는 점을 함께 이해할 필요가 있습니다.**

8) 숙의(공론화)에 참여하는 시민의 책임

숙의(공론화)에 참여하는 분들은 특정 집단을 대표해서 온 것이 아니라, **사회 전체를 대신해 판단해 달라는 요청을 받고 참여하는 시민입니다.** 자신의 이익을 앞세우는 개인이 아니라, 공동체를 위해 공적인 판단을 내리는 시민의 자리에 서 있는 것입니다. 성숙한 공론화를 위해 다음과 같은 참여 자세가 필요합니다.

- **열린 마음(Open-mindedness)** : 내가 가진 생각과 경험이 유일한 정답이 아닐 수 있음을 인정하고 새로운 사실이나 설득력 있는 근거를 만나면 언제든지 판단을 바꿀 수 있다는 열린 자세로 참여합니다.
- **공공의 가치 고려(Public Value)** : 개인적·지역적 이해관계나 소속 집단의 입장에서 벗어나 사회 전체와 다음 세대를 함께 고려하는 기준으로 사안을 바라봅니다.
- **사실에 근거한 학습과 확인(Fact-based Learning)** : 확인되지 않은 주장이나 감정적인 구호보다, 제공된 자료와 객관적인 정보를 꼼꼼히 살피고 전문가에게 합리적이고 근거있는 질문을 던집니다.
- **서로 존중하고 경청하기(Mutual Respect)** : 나와 다른 생각을 가진 참여자의 의견도 공동체를 위한 진지한 고민에서 나온 것임을 존중합니다. 다른 사람이 말할 때 끼어들지 않고, 정해진 발언 시간을 지켜 모두에게 평등한 기회를 보장합니다.
- **책임 있게 적극적으로 대화하기(Responsibility)** : 갈등이 불편하다고 해서 침묵하기보다, 궁금한 점은 적극적으로 묻고 함께 받아들일 수 있는 최선의 대안을 찾아갑니다.
- **외부 영향으로부터 독립성 지키기(Independence)** : 특정 단체나 이해관계자의 회유, 압력에 응하지 않습니다. 숙의 기간 중 외부로부터 특정 입장을 요구받았다면 즉시 운영진에게 알려주십시오.
- **비밀 유지와 정보 보호(Confidentiality)** : 분임토의에서 나온 다른 참여자의 발언과 개인정보를 외부에 알리지 않으며, 제공된 자료를 허락 없이 유출하거나 편집·인용해서 배포하지 않습니다. 이는 모든 참여자가 눈치 보지 않고 솔직하게 말할 수 있는 조건이 됩니다.
- **끝까지 성실하게 참여하기(Commitment)** : 숙의의 효과는 전체 과정을 함께 마칠 때 나타납니다. 사전 학습부터 최종 토론회까지 모든 일정에 성실히 참여해 주시길 부탁드립니다.

9) 우리가 함께 만든 대화의 약속을 되새겨 봅니다.

국민참여위원회가 발대식에서 함께 정한 '대화의 약속'을 이번 숙의에서도 함께 지켜나가면 좋겠습니다.

국민참여위원회 - 함께 만들어가는 대화의 약속

존중과 평등	경청과 공감	함께만드는 대화
- 우리는 서로를 동등한 참여자로 존중합니다. - 나이, 성별, 지역, 경험, 신념과 관계없이 누구나 안전하게 말할 수 있습니다. - 상대를 평가하거나 무시하는 표현 대신, 존중의 언어를 사용합니다. - 설득하거나 이기기보다, 나의 경험과 생각을 나누는 데 집중합니다.	- 우리는 말하는 것만큼 듣는 것을 중요하게 생각합니다. - 상대의 이야기를 이해하기 위해 질문합니다. - '내가 옳다'는 확신보다, 다른 가능성을 열어둡니다. - 나와 다른 의견도 공동의 문제해결과정으로 존중합니다.	- 한 사람이 너무 오래 말하지 않도록 배려합니다. - 가능한 사실과 경험에 기반해 이야기 합니다. - 주제에 관해 충분히 학습하고, 준비된 태도로 참여합니다. - 단순한 주장보다 이유와 맥락을 함께 나눕니다.

< 참고 문헌 >

국무조정실. (2024). 「2024 공공갈등관리 매뉴얼」(5차 개정).

국회입법조사처. (2023). 「시민참여 공론화 해외사례와 시사점」(NARS 현안분석 제293호, 김선화·오창룡).

대입제도 개편 공론화위원회. (2018). 「시민의 지혜! 숙의하고 대안을 찾다: 대입제도 개편 공론화 백서」.

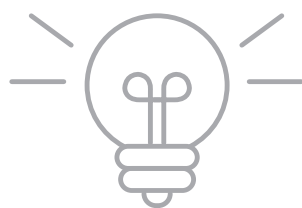
신고리 5·6호기 공론화위원회. (2017). 「숙의와 경청, 그 여정의 기록: 신고리 5·6호기 공론화 백서」.

은재호. (2022). 「공론화의 이론과 실제: 건강한 공론 형성과 진단을 위한 길라잡이」. 박영사.

피시킨, J. S. (2020). 「숙의민주주의」(박정원 역). 한국문화사. (원서출판 2009)

2장

대입제도를 논의하는 이유



국민과 함께
고민하는

미래
대입
제도

2장. 대입제도를 논의하는 이유

1) 2028~2037 국가교육발전계획의 수립

국가교육위원회는 「국가교육위원회 설치 및 운영에 관한 법률」 제11조 등에 근거하여 10년 단위 국가교육발전계획(2028~2037년)을 수립하고 있습니다. 급변하는 미래사회를 선도할 교육비전 및 체제를 마련하여 사회적 합의에 기반한 안정적이고 일관된 교육 정책을 위해 수립되며, 2027년 3월까지 마련할 계획입니다.

이번 숙의를 통해 논의하는 대입제도는 국가교육발전계획(2028~2037년)에 포함될 사항입니다. 최근 이루어진 두 번의 대입제도 개편은 4년 후의 대입제도에 관한 것이었습니다. 2018년에는 2022학년도 대입제도에 대해, 2023년에는 2028학년도 대입제도에 대한 개편을 논의하였습니다. 반면, 이번 숙의는 특정 시점의 대입제도에 대해 논의하는 것이 아닙니다. 2037년까지 현재 대입제도가 유지되어야 할지, 바뀌어야 한다면 어떠한 방향으로 변화해야 할지를 논의합니다.

2) 지금, 우리는 왜 대입제도를 논의하는가?

우리 사회는 인공지능(AI)의 일상화, 저출생·고령화에 따른 인구구조 변화, 다문화사회로의 전환, 기후변화 등 여러 방면에서 큰 변화의 시기를 맞고 있습니다.

초·중·고 학령기 인구(6~17세)는 2000년 810만명에서 2010년 734만명, 2025년 511만명으로 크게 줄었고, 2035년에는 309만명까지(2025년 대비 약 39.5% 감소) 줄어들 것으로 추정됩니다. 과거 한해 100만명이 출생하던 시기의 교육과 현재 한해 25만명이 출생하는 시기의 교육이 같아야 하는지, 아니면 달라져야 하는지 점검이 필요합니다.

인공지능 기술의 급격한 발전은 사회 전반의 변화와 함께 교육의 변화도 요구하고 있습니다. 미국 스탠퍼드 대학이 발표한 'AI 인덱스 보고서'에 따르면, 인공지능은 기본적인 독해와 추론 등 일부 인지평가영역에서 이미 인간의 평균치를 넘어선 것으로 나타났습니다. 기존의 지식 습득과 수행 중심의 일은 점차 인공지능이 대신하고 있으며, 사람의 역할은 인공지능이 만든 결과물을 비판적으로 검토 및 검증하고 다음 단계를 설계하는 방향으로 요구되고 있습니다. 이에 단순히 정답을 찾는 능력보다, 정보를 비판적으로 이해하고 스스로 문제를 해결하는 창의적 사고력을 키우는 교육이 필요하다는 목소리가 높아지고 있습니다.

이제 **10년, 20년 뒤 미래사회를 살아갈 인재를 키우기 위해 교육제도 전반에 대해 다시 생각해 봐야 할 때**입니다. 그리고 그 중심에 평가와 대입제도에 대한 논의가 자리하고 있습니다. 대입제도는 초·중·고 교육 전반에 지대한 영향을 미치며, 대입제도와 연계되지 않은 교육혁신 노력은 현장에 안착하기 어렵기 때문입니다.

3) 대입제도, 국민 의견과 사회적 합의가 중요합니다

대입은 좋은 교육을 받을 기회이자 사회·경제적 기회를 나누는 장치로 기능해왔습니다. 그래서 평가의 공정성과 신뢰성은 교육적 타당성 못지않게 중요합니다. 우리 사회에서 대입의 결과가 대학 진학은 물론, 그 이후의 취업과 사회적 지위에까지 영향을 미치는 상황에서 평가제도는 계층 이동의 사다리가 될 수도 있고, 반대로 기존의 격차를 굳히는 벽이 될 수도 있기 때문입니다.

우리나라의 대입제도는 대학별로 독자적인 시험을 치르던 방식에서 시작해 1969년 예비고사(자격시험) 도입, 1994년 대학수학능력시험 도입, 1997년 학교생활기록부 반영, 2002년 수시모집 도입, 2008년 내신 9등급 상대평가제 도입, 2010년대 입학사정관제(이후 학생부 종합전형으로 발전) 등을 거치며 여러 차례 바뀌어 왔습니다. 대입제도가 계속 바뀌어 온 배경에는 '우수 인재 육성', '공교육 정상화', '공정성 확보' 등 여러 요구에 대한 균형을 맞춰야 했던 사정이 자리하고 있습니다. 평가 방식을 바꾸는 일은 단순히 시험의 형태나 평가결과 제공 방식 등을 바꾸는 데 그치지 않습니다. 평가의 타당성과 신뢰성, 학교 교육의 여건, 공정성에 대한 사회적 신뢰 등 다양한 요소가 함께 갖춰져야 현장에 안착하여 지속될 수 있습니다.

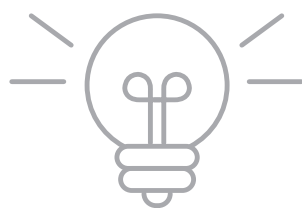
따라서 **대입제도는 학교 현장의 여건을 고려하고, 다양한 이해관계자의 목소리를 경청함으로써 신중하고 단계적으로 사회적 합의를 이루고 신뢰를 쌓아가는 '과정'이 무엇보다 중요합니다.** 이를 위해 국가교육위원회는 학생, 학부모, 교사, 일반국민이 모여 '미래사회를 대비한 대입제도의 방향'을 논의하는 현장토론회를 여러 지역에서 개최하고, 이와 더불어 국민 누구나 참여하여 의견을 개진할 수 있도록 온라인 의견수렴을 진행하고 있습니다. 이번에 진행되는 국민참여위원회 속의 토론회도 이러한 노력의 일환이라고 할 수 있습니다.

< 참고문헌 >

- 안수지, 장채윤, 권승은, 강민영. (2025). 「학령인구 감소에 따른 교육 현장의 변화 및 정책 제언」.
국회미래연구원 인구통계 Brief 25-3호.
- Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). 「AI Index Report 2024」.
Stanford University.
- 대입제도 개편 공론화위원회. (2018). 「시민의 지혜! 숙의하고 대안을 찾다: 대입제도 개편 공론화 백서」.

3장

대입제도와 서·논술형 평가



국민과 함께
고민하는

미래
대입
제도

3장. 대입제도와 서·논술형 평가

평가는 본래 학생이 학습목표에 얼마나 도달했는지를 확인하고, 그 결과를 교사와 학생의 가르치고 배우는 과정에 반영해 학생의 성장을 돕기 위한 것입니다. 하지만 우리나라에서는 평가의 본질적 기능보다는 학생의 순위를 매겨 선발에 활용하는 것이라는 인식이 널리 자리잡아 온 측면이 있습니다. 이러한 인식 속에서 채점결과가 비교적 일정하게 유지되어 공정하다고 여겨지는 지필평가(종이시험)가 중요하게 활용되어 왔습니다.

지필평가는 주로 선택형(선다형) 문항을 중심으로 이루어집니다. 최근 중·고등학교에서 서·논술형 평가를 확대해 나가고 있으나, 여전히 대학수학능력시험은 오지선다형 문항을 유지하고 있습니다. 이하에서는 서·논술형 평가의 개념과 특징, 운영 현황에 대해 알아보고, 대입에서의 서·논술형 평가 논의 배경과 서·논술형 평가 추진 시의 우려와 기대에 대해 살펴보도록 하겠습니다.

1. 서·논술형 문항의 개념과 장·단점

1) 지필평가의 유형

현재 고등학교의 평가는 수업과정에서 이루어지는 다양한 활동을 평가하는 '수행평가'와 정해진 시험기간에 실시되는 '정기고사'로 구성되어 있습니다. 이 가운데 일부 수행평가와 정기고사는 종이시험, 즉 지필평가의 형태로 운영되고 있습니다.

지필평가의 문항 유형을 나누는 기준은 몇가지가 있지만, 대표적으로 학생이 응답을 산출하는 방식에 따라 '선택형'과 '구성형'으로 구분할 수 있습니다.

- 선택형 문항 : 학생이 미리 제시된 답지 중에서 알맞은 답을 선택하도록 하는 문항
(예시) 선다형, 복수선택형, 진위형, 연결형
- 구성형 문항 : 학생이 독자적으로 답안을 작성 또는 구성하도록 하는 문항
(예시) 완성형, 단답형, 서술형, 논술형

정기시험에서의 구성형 문항의 유형은 다음과 같습니다.

(「2022 개정 교육과정에 따른 중학교/고등학교 학생평가 톺아보기」)

- 완성형 문항 : 제시된 빈칸에 적절한 단어, 숫자, 기호 등으로 작성하도록 하는 문항
- 단답형 문항 : 제시된 질문에 적절한 단어나, 구, 숫자, 수식, 그림 등 제한된 형태로 작성하도록 하는 문항
- 서술형 문항 : 교과목에서 학생이 학습한 지식이나 개념, 원리 등을 활용하여 문항에서 요구하는 내용을 정확한 문장 수준으로 작성하도록 하는 문항

- **논술형 문항** : 교과목에서 학생이 학습한 지식이나 개념, 원리 등을 활용하여 문항에서 요구하는 내용을 논리적 구조를 갖춘 완결된 문단이나 글 수준으로 작성하도록 하는 문항

2) 서·논술형 문항의 특성

즉 서·논술형 문항은 주어진 것에서 무엇을 선택하는 것이 아니라 답안을 직접 구성하는 대표적인 문항 유형입니다. 두 문항 유형은 학생들이 답을 직접 구성하며, 이를 통해 분석력, 비판적 사고력, 문제해결력, 창의력 등의 고차원적 사고능력 및 역량을 측정할 수 있습니다.

구분	서술형 문항	논술형 문항
정의 및 특성	· 교과에서 학생이 학습한 지식이나 개념, 원리 등을 활용하여 문항에서 요구하는 내용을 정확한 문장 수준으로 작성하도록 하는 문항 · 학생이 작성한 한 문장 이상의 글을 통해 응답의 정확성과 질을 평가하는 문항	· 교과에서 학생이 학습한 지식이나 개념, 원리 등을 활용하여 문항에서 요구하는 내용을 논리적 구조를 갖춘 완결된 문단이나 글 수준으로 작성하도록 하는 문항 · 근거를 들어 자신의 주장을 논리적이고 설득력 있게 조직하여 작성하도록 하는 문항 · 글의 완결성이나 주장의 완성도를 평가하는 문항
응답 분량	· 응답 분량이 상대적으로 짧음 · 최소 한 문장에서 길게는 서너 문장 수준	· 응답 분량이 상대적으로 길고 복잡함 · 최소 한 문단
평가 하려는 행동 특성	· 정의, 예시, 설명, 요약, 비교·대조, 나열, 분류, 구분, 인과, 분석·해석 등	· 다양한 정보나 지식을 조직하거나 통합하는 능력, 창의적인 사고나 추론 능력 등의 다양한 지식과 사고 능력을 복합적이고 종합적으로 활용할 수 있는 능력 · 제안, 해결방안 제시, 해석·평가, 비평, 증명, 논증 등 주장과 타당한 근거 포함

3) 선다형 문항과 서·논술형 문항의 장·단점

우리나라 지필평가에서는 선다형(선택형) 문항이 오랫동안 사용되어 왔고, 지금도 주되게 사용되고 있습니다. 선다형 문항의 장점과 단점을 아래와 같이 정리할 수 있습니다.

- 정답이 분명하여 **채점의 객관성**이 높습니다.
- 학생들이 짧은 시간에 많은 문항을 풀 수 있습니다. 이에 문항을 많이 출제하여 **교과 전 영역을 촘촘하게 다룰 수** 있습니다.

- 답지를 복합적으로 구성하면 분석, 추론 같은 사고력까지 측정할 수 있습니다.
- 좋은 선다형 문항, 특히 그럴듯한 오답지를 정교하게 구성하는 것이 어려워 **채점은 쉽지만 제작은 어려운 문항 형식**입니다.
- 학생 스스로 답을 구성하는 것이 아니라 선택을 함으로써 **학생 자신의 생각을 구성하고 표현하는 능력을 평가하는데 본질적 제약**이 있습니다.
- 정답을 맞히는 요령이나 반복적인 문제풀이 **훈련에 쉽게 노출**됩니다.

서·논술형(구성형) 문항의 장점과 단점은 아래와 같이 정리할 수 있습니다.

- **학생의 사고 과정과 문제해결력 등 고등역량을 보다 충실하게 평가**할 수 있습니다.
- 학습자가 자신의 생각을 스스로 구성하고 논리적 언어로 표현하도록 요구하기 때문에 **탐구 수업과 역량 중심 교육을 촉진하는 데 유리**합니다. 학생은 자신의 사고과정을 객관적으로 관찰하고 성찰할 기회를 얻을 수 있습니다.
- 정답 여부만으로 학생의 성취를 판단하지 않고 학생이 보여 준 사고와 수행의 정도에 따라 부분점수를 부여함으로써 **학습성과를 보다 세밀하게 평가**할 수 있습니다.
- **평가의 공정성과 신뢰성을 확보하는 것이 상대적으로 어렵습니다.** 채점자의 역량이나 채점 여건이 채점 결과에 영향을 미칠 수 있습니다.
- 답안을 작성하고 채점하는 데 상대적으로 많은 시간이 필요하기 때문에 평가할 수 있는 문항 수가 제한됩니다. 이에 **평가범위가 제한**될 수 있습니다.
- 학생의 고차원적 역량(비판적 사고력, 창의력 등)을 측정하는 타당성 있는 문항과 평가기준을 마련하기 위해서는 **더 높은 평가자의 전문성**이 요구됩니다.

II. 고등학교와 대학수학능력시험에서의 서·논술형 평가 활용 현황

1) 고등학교

(1) 최근의 추세 및 교육청별 지침

최근 우리 교육은 학생의 역량 함양이라는 시대적 흐름에 따라 **서·논술형 평가의 비중을 확대하는 추세**입니다. 이는 2022 개정 교육과정과 고교학점제 도입 등 미래형 교육 체제로의 전환 흐름과 맞닿아 있습니다. 교육부와 교육청에서도 서·논술형 학생평가 도구를 개발·보급하고 서·논술형 평가 연수를 확대하며 학생평가 중앙지원단을 운영하는 등 평가 방법 개선 정책을 꾸준히 시행해 오고 있습니다.

현재 모든 교육청에서는 통상 20~40% 이상의 서·논술형 평가를 포함하도록 권장하거나 의무화하고 있습니다. 다만, 그 비율을 적용하는 방식에는 차이가 있습니다.

(※ 방식과 관련해서는 <부록1> 교육청별 서·논술형 평가 지침 참조)

(2) 서·논술형 평가 운영 현황

한국교육과정평가원에서 고등학교 현장교사 167명을 대상으로 조사한 결과, 정기시험의 서·논술형 평가 반영 비율은 25.1%, 수행평가에서의 서·논술형 평가 반영 비율은 34.7%로 나타났습니다.(한국교육과정평가원, 2025)

아울러 전국 고등학교 중 166개교를 표집하여 분석한 결과, 2024년 국어교과 1학년 1학기 정기시험의 서·논술형 평가 중 응답길이가 단어·어절 수준의 문항은 26.1%, 한 문장 수준인 문항은 45%, 한 문장을 초과하는 수준의 문항은 27.7%, 한 문단 이상 수준의 문항은 1% 정도로 나타났습니다.(한국교육과정평가원, 2024)

2) 대학수학능력시험

대학수학능력시험은 1994학년도 시행된 이후로 선다형을 유지하고 있습니다. 다만 수학 영역에 한해 1997학년도부터 단답형이 도입(30%)되어 현재 운영 중에 있습니다.

III. 서·논술형 평가 도입 및 확대 논의의 배경

서·논술형 평가에 대한 논의가 최근에 갑자기 나온 것은 아닙니다.

2015 개정 교육과정과 2022 개정 교육과정은 자기관리 역량, 지식정보처리 역량, 창의적 사고 역량, 심미적 감성 역량, 협력적 소통 역량, 공동체 역량이라는 6대 핵심역량을 제시하며, 교육의 중심을 '지식의 암기'에서 '지식의 활용'으로 옮겨가고 있습니다. **학교교육은 '무엇을 알고 있는가' 보다 '무엇을 할 수 있는가'를 중시하게 되었으며**, 이에 따라 수업과 평가 방식도 학습자의 삶과 연계된 경험을 강조하는 방향으로 나아가고 있습니다. 서·논술형 평가는 학생의 아이디어와 사고 과정을 직접 포착함으로써 역량 중심 교육을 구현하는 대안으로 계속해서 주목받고 있습니다.

아울러 2023년 12월 교육부가 발표한 '**2028 대학입시제도 개편 확정안**'에도 국가교육위원회 중심의 미래형 대입제도 구상·논의가 필요하며 "**교사의 평가역량을 강화하고 논·서술형 내신 평가의 혁신이 대입과 효과적으로 연계될 수 있도록 선진적 대입기반 구축 지원**"을 언급한 바 있습니다.

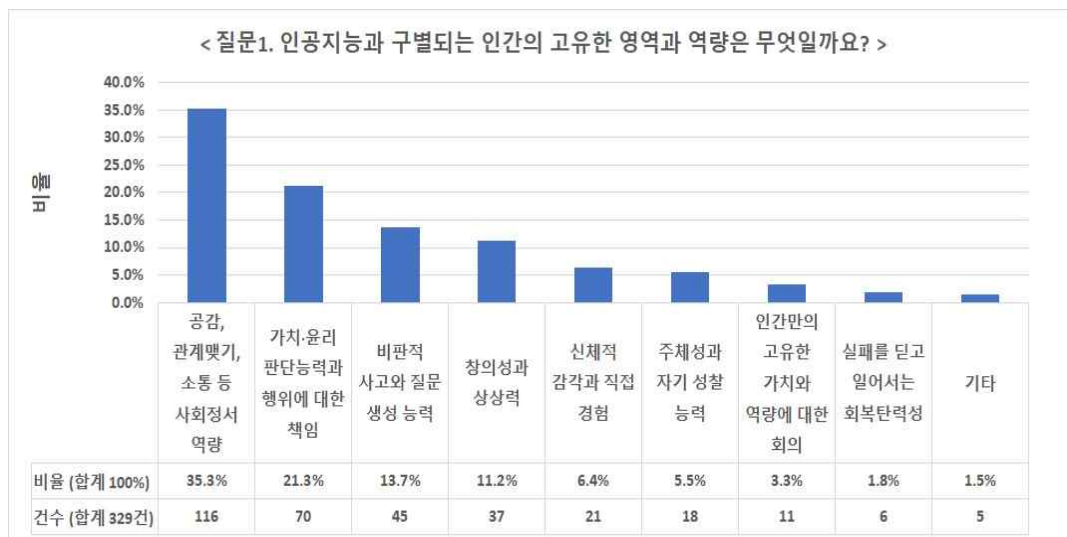
AI시대, 변화하는 인재상과 교육

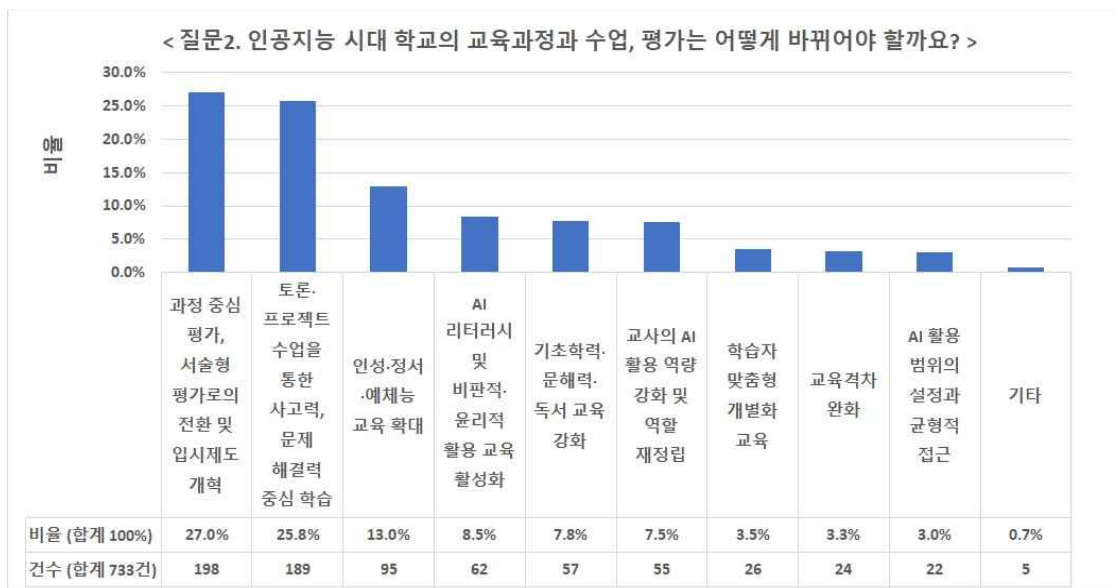
인공지능 기술이 빠르게 발전하면서 우리 삶의 많은 부분에서 변화가 이루어지고 있습니다. 이러한 변화는 인간의 고유한 영역과 역량은 무엇인지, 학교교육은 어떠하여야 하는지 많은 질문을 던지고 있습니다.

국가교육위원회는 이에 대한 답을 구하고자 지난 6월 ‘AI시대, 우리 교육의 방향’을 주제로 대규모 토론회를 개최하였습니다. 전국의 학생, 청년, 학부모, 교육관계자, 일반국민 등 335명이 모여 AI시대의 교육에 대해 생각을 모았습니다.

많은 분들이 공감과 소통 등 사회정서역량, 가치·윤리에 대한 판단능력과 책임감, 비판적사고와 질문하는 능력을 키워야 한다고 말씀하셨습니다. 또한 AI시대에는 과정 중심 및 서술형 평가로의 전환과 입시제도 개혁이 필요하며, 사고력·문제해결력 중심의 학습이 필요하다고 하였습니다.

국민참여위원회 ‘AI 시대 우리교육의 방향’ 토론회 의견 분석





그래서 현재 제도로 충분한가? - 대학수학능력시험의 한계

대학수학능력시험은 지난 33년간 오지선다형 문항을 출제해왔습니다. 그리고 출제의 기반이 되는 교육과정의 내용과 범위가 지속적으로 축소되어 왔습니다. 이러한 상황에서 공정성을 확보하기 위해 기출문제, 사교육 시장에서 다루어진 문제와 유사한 문제를 피해야 한다는 부담이 해마다 더해지고 있습니다. 이에 **교육과정에 충실하면서도 공정성과 평가적 타당성을 갖춘 문항을 개발하는 일은 갈수록 어려워지고 있습니다.**

초·중·고 교육 전반에 큰 영향을 미치는 대학수학능력시험의 특성을 고려할 때, 우리 학생들은 초·중·고 12년 동안 오지선다 문제 풀이의 영향 아래에 있다고 볼 수 있습니다. 오지선다형과 같은 선택형 문항 위주의 평가는 자칫 **문제풀이 기술의 숙달을 고착화하고 토론을 통한 사고의 확장이나 능동적인 탐구학습을 제한하는** 결과를 낳을 수 있습니다. 실제로 우리가 일상에서 마주하는 수많은 문제는 정해진 답안 중 하나를 고르는 선택형 문항만으로 대응하기 어려운 복합적인 성격을 지닙니다.

학습과 수업 변화의 기제로서의 서·논술형 평가

서·논술형 평가를 통해 학교교육이 변화할 수 있다는 기대도 있습니다. 서·논술형 평가가 학교 수업과 긴밀하게 연계될 경우, 학생들은 수업 시간에 정답을 찾는 연습을 넘어 질문하고 토론하며 자신의 생각을 글로 표현하는 경험을 중요하게 여기게 됩니다. 서·논술형 평가가 확대되면 그런 변화가 촉진될 것입니다.

※ (참고) 교육부 '학생평가 중앙지원단'에서 활동하시는 선생님들이 서·논술형 평가 실행 소감(2026)

교사A "서·논술형 평가를 현장에서 실행하면서 느낀점은 선다형 문항에서는 볼 수 없었던 학생들의 사고 과정을 볼 수 있었어요. 선다형에서는 점수가 낮은 학생도 서·논술형에서는 자신의 사고가 충분히 반영된 답안을 쓸 수 있다는 점에 놀랐어요."

교사B "장점이 많은 평가방법이지만 교사의 준비가 많이 필요한 평가방법인 것 같아요. 먼저 가르치고자 하는 내용 요소가 명확해야 한다는 것을 느꼈어요. 서·논술형 평가는 학생들의 사고도 명쾌해지고 깊이 있는 사고가 가능해지기 때문에 평가의 효과와 교육의 효과가 극대화된다고 봐요. 다음으로는 평가 전 수업이 중요하다는 것을 느꼈어요. 평가 전 수업을 통해 이 단원에서 다루는 중요한 문제는 무엇이며 이와 관련된 중요한 내용 요소는 무엇인지가 충분히 가르쳐져야 학생들의 답안의 방향성이 분명해지는 것 같았어요."

IV. 서·논술형 평가의 주요 쟁점

1) 채점 공정성

서·논술형 채점은 왜 어려울까요?

서·논술형 평가와 관련하여 반드시 짚고 넘어가야 할 문제가 있습니다. 바로 채점의 어려움입니다. 선택형 문항은 학생이 고른 답이 정답과 일치하는지만 확인하면 되기 때문에 채점자의 판단이 끼어들 여지가 거의 없습니다. 반면 서·논술형 문항은 학생의 표현 방식과 논리 전개, 배경지식에 따라 답안이 저마다 다르게 나타납니다. 그래서 이 답안이 몇 점 인지는 최종적으로 채점자가 판단하게 됩니다(교육부, 2025). 바로 이 '판단이 개입하는 지점'에서 채점의 어려움이 생겨납니다. 정답이 하나가 아니라는 특성상, 채점의 공정성과 신뢰도 문제가 제기되는 것입니다.

< 관련 언론보도 >

"서·논술형평가는 학생의 사고력과 문제 해결력을 평가할 수 있다는 장점이 있지만, 객관식과 달리 채점자의 판단이 개입될 수 밖에 없어 '누가, 어떤 기준으로 채점하느냐'가 핵심과제로 꼽힌다." (뉴시스, '26.8.8., 「日은 보류한 서·논술형 입시...AI 채점은 공정할까」)

"평가 점수에 따라 대입 당락이 달라지는데 그걸 수용할 학부모가 극히 드물 수 있다. 수능에 대한 행정심판, 행정소송 등 법적 대응이 매우 늘어날 것" (서울신문, '26.8.10., 「서·논술형 시험, 30년 수능 흔드나...①채점 ②수용성 ③사교육 등 쟁점」)

“서논술형 평가는 채점자에 따라 결과가 달라질 수 있는 만큼 이를 보완할 표준화된 평가도구가 필요하다. 현재는 검증된 평가도구가 부재한 상황인데 채점자의 신뢰도 문제가 발생할 수 밖에 없다. 교사의 채점민원 부담이 늘어나게 될 것” (이데일리, '26.7.25., 「서논술형 100%, 학교 혼란 불가피...도입 시기 조정해야」)

서·논술형 채점의 어려움을 어떻게 극복할 수 있을까요?

• 채점기준표(루브릭)를 정교하게 만들고 공유합니다

채점 기준을 미리 구체적인 단계와 예시 답안으로 정리해두면, 채점자의 판단이 개입할 여지 자체를 줄일 수 있습니다. 국가 수준에서 표준화된 채점기준표를 개발해 학교에 보급하고, 예시 답안(anchor paper)을 함께 제공하는 방식이 대표적입니다.

• 본 채점 전에 미리 채점해 봅니다

본격적인 채점에 앞서 일부 답안을 미리 채점해보는 과정(가채점)을 거치면, 예상치 못한 형태의 답안이 있는지 미리 살펴보고 채점 기준을 보완할 수 있습니다. 이 과정은 한 채점자가 스스로 채점 기준을 구체화하는 데도, 여러 채점자가 서로의 눈높이 차이를 미리 맞추는 데도 도움이 됩니다(교육부, 2025).

• 한 답안을 여러 명이 함께 채점합니다

한 답안을 두 명 이상의 채점자가 독립적으로 채점하고(공동채점·교차채점), 점수 차이가 클 경우 제3의 채점자가 다시 검토하는 절차를 두면, 한 채점자의 판단 쓸림이 최종 점수에 그대로 반영되는 것을 막을 수 있습니다(교육부, 2025).

대규모 시험에서도 공정하게 채점할 수 있을까요?

대규모 시험에서 수십만 건의 서·논술형 답안을 공정하게 채점하는 것이 어렵다는 우려가 제기됩니다. 일본은 국가수준의 대입시험에 서술형 문제 도입을 2018년에 추진하였으나, 약 50만명에 이르는 수험생 답안을 공정하고 정확하게 채점하기 위한 현실적인 방안 마련에 실패하였고, 2021년에 서술형 문항 도입이 어렵다고 결론내렸습니다. 반면, 프랑스, 독일, 중국 등 대규모 국가수준 시험에서 서·논술형 평가를 오랜기간 진행해온 나라도 많습니다. 더하여, AI 기술의 발전으로 AI 채점 결과가 훈련된 인간 채점자의 결과와 비교적 높은 수준으로 일치하고 있으며, 일부 교육청에서도 AI 평가지원시스템을 이미 운영 중으로, AI를 채점의 보조도구로 활용하여 채점의 효율성과 일관성을 확보할 수 있다는 주장도 있습니다. AI 활용에 관한 주장은 “V. 서·논술형 평가에서의 AI 활용 가능성”에서 별도로 살펴보겠습니다.

2) 사교육 의존

글쓰기·논술을 위한 새로운 사교육을 추가하게 되면서 사교육비 부담이 늘어난다는 비판이 제기됩니다. 새로운 평가체계가 도입되면 사교육을 통해 대비하게 된다는 것입니다. 어떤 형태의 답안이 좋은 점수를 받는지 모른다는 정보 부족으로 인한 불안감이나, 교과 지식은 같지만 글쓰기 능력으로 인해 불리할 수 있다는 우려도 사교육 수요를 늘리는 요인입니다.

더하여, 개별첨삭을 요구하는 글쓰기, 논술학원은 비용 자체도 높다는 비판도 제기됩니다. 사교육이 더 비싸고 접근하기 어려운 형태로 바뀐다면 그 피해는 이를 감당하는 어려운 가정이 보게 된다는 우려로 이어집니다.

또한, 논술 첨삭이 가능한 전문 인력도 수도권과 대도시에서 집중되어 있어 비수도권·농산어촌 학생이 불리한 지역격차로 이어질 수 있다는 지적도 제기됩니다.

< 관련 언론보도 및 보도자료 >

“원래는 교과목 중에선 영어 학원만 보냈지만, 뉴스를 보고 손 놓고 있으면 안되겠다는 생각이 들었다. 주위 다른 ‘초등맘’들도 본격적으로 논술을 시켜야 하는 게 아니냐고 고민하고 있다.” (연합뉴스, '26.8.16., 「서·논술형으로 사교육 잡는다?...학부모들 “학원 하나 더 다닐 뿐”」)

“개별첨삭과 구술지도는 더 많은 비용이 들고, 전문인력도 수도권과 대도시에서 집중되어 있다. 사교육이 더 비싸고 접근하기 어려운 형태로 바뀐다면, 그 피해는 이를 감당하기 어려운 가정과 비수도권·농산어촌 학생들에게 먼저 돌아갈 것이다(전국중등교사노동조합 보도자료, '26.8.18)”

반면, 사교육 의존이 높아지지 않거나 오히려 낮출 수 있다는 주장도 존재합니다. 중·고교 내신과 수능 유형을 일치시켜 공교육에서 대비할 수 있다는 것입니다.

더하여, 서·논술형 평가를 대비하기 위해 필요한 글쓰기, 토론, 발표는 학교에서 잘할 수 있는 영역이라는 주장도 있습니다. 오히려 선다형 평가가 기출문제를 분석해 패턴을 암기시키고, 반복적인 문제 풀이 훈련을 할 수 있어 사교육 의존이 강했다는 것입니다. 탐구·토론·글쓰기 교육을 하고, 서·논술형 평가를 실시하는 제주 중·고등학교와 대구 초·중학교를 대상으로 하는 연구 결과에 따르면, ‘학원, 과외 등이 학교 수업에 도우며 되느냐’는 설문에 5점 척도(5점 : 매우 그렇다 / 3점 : 보통이다 / 1점 : 전혀 그렇지 않다)에 3점 이하의 결과가 나왔으며, 또한 위와 같은 학교교육에 대한 만족도도 80~90% 이상에 이르는 등 매우 높았다고 합니다(이혜정 외, 2019;2021).

< 관련 언론보도 및 보도자료 >

“답이 정해져 있으니 유형이 나뉘고, 유형이 나뉘니 반복으로 속도를 올리고, 오답의 함정마저 목록으로 외울 수 있습니다. 자기 생각을 세우는 능력은 그렇게 압축되지 않습니다. 사교육이 그 자리를 파고들지 못한다는 뜻이 아닙니다. 다만, 선다형만큼 쉽고 빠르게 상품이 되기는 어렵습니다(교육의봄·사교육걱정없는세상·좋은교사운동 보도자료, '26.8.18)”

3) 학교현장에서의 실현 가능성

서·논술형 평가는 교사에게 정교한 문항 설계와 상당한 정도의 채점 부담을 요구하기 때문에 교육 현장에서 쉽지 않은 도전으로 여겨집니다. 많은 답안을 채점하는 데 드는 시간과 노력도 부담이지만, 채점 결과를 둘러싼 이의제기 역시 교사에게 큰 부담이 될 수 있습니다.

(1) 교육여건

서·논술형 평가는 교사에게 수업에서 토론식 수업과 글쓰기를 일상화하고, 정교한 문항 설계와 함께 채점에 상당한 노력과 시간을 투입할 것을 요구합니다. 새로운 문제를 개발하고, 평가 기준을 마련하며, 교원 연수 등 교육을 받아야 합니다. 하지만, 한 교사가 여러 학급, 많은 학생을 담당하는 현실에서 서·논술형 평가에 대한 전문성을 갖추고, 답안을 세심하게 채점하고 피드백할 현실적인 여건이 부족하다는 것입니다.

이에 대해 교원 1인당 학생수를 줄이고, 행정업무 감축 등을 통해 충분한 출제·채점 시간을 확보하며, 내실있는 연수를 통해 교원 전문성을 향상하고, 교사 내 전문채점관 양성 등이 병행되어야 한다는 목소리가 제시됩니다. 더하여, AI를 평가의 보조도구로 활용하여 채점부담을 경감할 수 있다는 주장도 있습니다.

< 관련 언론보도 및 보도자료 >

“교원단체가 우려하는 핵심은 평가 도입에 따른 ‘업무 과중’과 ‘심리적 압박’일 것입니다. ...(중략)... 전문성 신장 연수에 참여하거나 평가를 진행하는 기간에는, 교사들의 다른 행정 업무를 대폭 축소해 오직 ‘평가와 맞춤형 피드백’에만 집중할 수 있는 여건이 보장되어야 합니다.” (EBS, '26.8.7., 「객관식 넘어 서논술형으로... 성공 조건은?」)

(2) 평가결과에 대한 이의제기 대응체계

교사의 평가 결과에 대한 공정성 시비가 발생하여 민원이나 이의제기로 이어질 수 있다는 우려가 있습니다. 실제로 학교 현장에서는 서·논술형 문항의 채점 결과에 대한 학생·학부모의 민원을 우려하여 채점시 주관적 판단이 개입될 여지가 적은 단답형에 가까운 문항을 출제하게 되는 경우도 많다는 것입니다. 이에 대해, 학교와 교육청 차원의 대응체계를 마련하고, 교사의 평가권을 확립하는 등 교사의 전문적 판단을 존중하고 보호할 제도가 필요하다는 의견이 제시됩니다.

< 관련 언론보도 및 보도자료 >

“지금 쓰는 AI도 채점에서 틀릴 때가 있다. 물론 수년 뒤에는 기술이 발달하겠지만, 정확도 100%라는 게 있겠느냐. 결국 책임은 마지막에 채점하는 교사가 지게 될텐데, 이의 신청을 어떻게 일일이 다 들어줄지 모르겠다.” (연합뉴스, '26.8.16., 「서·논술형으로 사교육 잡는다?...학부모들 “학원 하나 더 다닐 뿐”」)

V. 서·논술형 평가에서의 AI 활용 가능성

AI를 서·논술형 평가에 활용하려는 시도가 국내외에서 이어지고 있습니다. AI는 많은 답안을 빠르게 검토하고 점수와 그 근거를 제시할 수 있지만, 판단이 항상 정확하거나 일관된 것은 아닙니다. 따라서 활용 가능성과 한계를 함께 살펴볼 필요가 있습니다.

1) AI 활용의 가능성과 한계는 무엇일까요?

다음에서는 국내외 주요 연구와 논의를 바탕으로 AI 활용의 가능성과 한계를 네 가지 측면에서 살펴봅니다.

(1) 채점 결과의 정확성과 일관성

AI는 많은 답안에 동일한 채점 기준을 신속하게 적용할 수 있어 대규모 평가에서 교사의 채점을 효율적으로 지원할 수 있습니다¹⁾. 최근 여러 연구에서는 AI 채점 결과가 훈련된 인간 채점자의 결과와 비교적 높은 수준으로 일치하여, 채점 보조 도구로서의 가능성을 보여주었습니다²⁾ 다만 채점 결과는 사용한 AI와 평가 시점에 따라 달라질 수 있습니다³⁾. 일부 연구에서는 AI가 인간 채점자보다 상대적으로 높은 점수를 주거나 특정 점수에 답안을 집중하여, 답안 간 차이를 충분히 구분하지 못한 경우도 확인되었습니다⁴⁾. 따라서

1) Tate et al., 2024; Yan et al., 2024; Jang & Sawaki, 2025

2) 안해연, 2025; Tate et al., 2024; Crossley et al., 2026

3) 이규민 외, 2026; Pack et al., 2024

AI가 답안을 신속하게 검토하는 장점을 활용하되, 중요한 평가에서는 교사가 채점 결과를 다시 확인하는 방식으로 활용할 필요가 있습니다⁵⁾.

(2) 채점 근거와 결과 설명

AI는 답안의 내용과 채점 기준을 바탕으로 점수와 그 근거를 글로 제시할 수 있습니다. 이러한 설명은 교사가 채점 결과를 이해하고 점검하는 데 참고 자료로 활용할 수 있으며⁶⁾, 이를 통해 학생에게 상세한 피드백을 제공할 수 있다는 것입니다. 다만 AI가 답안의 핵심을 놓치거나 일부 채점 기준을 교사와 다르게 적용할 가능성도 보고되었습니다⁷⁾. 따라서 교사가 AI의 설명이 답안과 채점 기준에 맞는지 확인하면서 활용한다면, 점수의 근거를 살펴보고 추가 확인이 필요한 답안을 찾는 데 유용한 보조 자료가 될 수 있습니다.

(3) 교사의 역할과 책임

AI가 채점한 결과와 그 설명은 교사가 답안을 다시 살펴보고 자신의 판단을 점검하는 참고 자료로 활용할 수 있습니다⁸⁾. 그러나 AI가 점수와 이유를 제시하더라도 채점의 정확성과 공정성이 자동으로 보장되는 것은 아닙니다⁹⁾. 최종 점수의 확정과 오류에 대한 재검토는 교사가 담당하고, AI는 교사의 판단을 지원하는 방식으로 활용해야 합니다¹⁰⁾.

(4) 손글씨 답안의 인식

종이에 작성한 답안을 AI로 채점하려면 먼저 손글씨를 디지털 글자로 바꾸는 문자인식 과정이 필요합니다. 한글 손글씨를 인식할 수 있는 기술은 상당히 발전하여, AI Hub 자료에서는 낱글자와 단어의 완전일치율이 95%로 나타났습니다¹¹⁾. 다만 이러한 결과는 통제된 연구 자료, 일반 문서를 대상으로 측정한 것이므로, 실제 학생의 서·논술형 답안에서도 같은 정확성이 나타난다고 단정하기는 어렵습니다. 따라서 실제 학생 답안으로 손글씨 인식 성능을 별도로 점검하고, 인식이 불확실한 부분은 원본 답안과 함께 교사가 확인해야 합니다.

4) 함은혜 외, 2024; 오유경·김경선, 2025; Kim & McCrudden, 2026

5) Jang & Sawaki, 2025

6) Awidi, 2024; Li et al., 2025

7) Johnson & Zhang, 2024; Sawaki et al., 2025

8) Awidi, 2024; Li et al., 2025

9) Johnson & Zhang, 2024; Jang & Sawaki, 2025; Huang et al., 2026

10) 최진영 외, 2025; Jang & Sawaki, 2025

11) 한국지능정보사회진흥원, 2020

2) AI 평가의 전망은 어떠한가요?

생성형 AI 기술이 발전하면서 교육 분야에서 활용 범위도 넓어지고 있습니다. 경제협력 개발기구(OECD)는 교육 목적에 맞게 만들고 효과를 충분히 검증한 생성형 AI가 교육을 개선하는 데 새로운 가능성을 열 수 있다고 전망하였습니다¹²⁾. 미국의 비영리 교육평가 기관인 ETS는 AI가 문항을 만드는 데 그치지 않고, 문항이 평가 목적과 기준에 맞는지 검토하고 평가 자료를 분석하는 데에도 활용되고 있다고 밝혔습니다¹³⁾. 최근 연구에서도 AI에 채점 기준과 점수가 매겨진 예시 답안을 제공하고 채점 방법을 구체적으로 안내했을 때, AI 채점의 정확도와 인간 채점과의 일치 정도가 높아질 수 있는 것으로 나타났습니다¹⁴⁾. 앞으로 AI의 기술이 고도화되고 평가 목적에 맞는 채점 방법이 더욱 정교해진다면, AI는 서·논술형 평가를 보다 안정적으로 지원할 수 있을 것으로 기대됩니다.

3) 어떻게 활용하는 것이 바람직할까요?

앞서 살펴본 것처럼, 최근 여러 연구에서는 AI 채점 결과가 인간 채점과 비교적 높은 수준으로 일치하여 채점 보조 수단으로 활용할 가능성을 보여주었습니다. 그러나 진학이나 선발에 영향을 미치는 중요한 평가에 AI를 활용하려면, 실제 답안으로 충분히 사전 점검하여 인간 채점과의 일치 정도, 반복 채점 결과의 일관성, 특정 학생 집단에 불리한 점수 차이가 나타나지 않는지를 확인해야 합니다. AI와 교사의 판단이 다르거나 점수를 결정하기 어려운 답안은 교사가 다시 살펴봐야 합니다. 또한 학생의 답안과 개인정보를 안전하게 관리하고, AI 사용 여부와 채점 절차, 오류의 재검토 및 이의 제기 절차를 학생에게 안내해야 합니다. 현재로서는 AI가 단독으로 점수를 결정하기보다 교사의 판단을 지원하는 방식으로 활용하는 것이 적절합니다.

VI. 다른나라에서의 서·논술형 평가

다른 나라에서는 우리의 수능과 같은 시험에서도 서·논술형을 활용하고 있습니다. 자주 언급되는 국외의 몇 가지 사례를 살펴보겠습니다.

¹²⁾ OECD, 2026

¹³⁾ Gonthier, 2025

¹⁴⁾ AlGhamdi et al., 2026; Choi et al., 2026; Xue et al., 2025)

프랑스의 바칼로레아

바칼로레아 시험(이하 Bac)은 중등교육 이수 자격과 대학 진학 자격을 동시에 갖는 자격 시험입니다. 일반계 바칼로레아의 최종 성적은 다음과 같이 구성됩니다.

구분	비중	평가
지속평가(학교평가)	40%	고교 1, 2학년의 학교 평가
국가 종결시험(국가시험) - 프랑스어 - 철학 - 전문과목 2개 - Grand oral(구술시험)	60%	국가 수준의 시험 - 필기+구술 - 4시간 논술형 필기 - 과목별 필기, 실기 등 - 구술

우리나라 식 표현으로 하면 학교 내신 40%와 국가 시험 60%의 혼합형 체제라고 볼 수 있습니다(eduscol, 2026-08-18 검색). 채점에는 거의 모든 일선 교사들이 참여합니다. 이들은 채점의 원칙에 대한 설명을 듣고, 모범답안 예와 채점 기준표를 받아 철저히 이에 따라 채점합니다. 프랑스 바칼로레아는 국가 차원의 공식적이고 항구적인 채점 기준이 존재하지 않습니다. 대신 연도별로 교사, 장학관, 교육과정 입안자들 간에 합의가 중요합니다. 일종의 채점자 연수를 통해 실제 답안을 함께 채점해 보면서 서로의 판단을 비교하는 과정을 거칩니다. 채점 후에는 채점자별 평균 비교, 점수 분포 비교, 특정 답안 재검토, 채점자 편차가 큰 경우 원인 분석, 채점 점수 상향 및 하향 조정 등이 이루어집니다(Ministere de l'education nationale, 2026-08-18 검색).

국제 바칼로레아

국제 바칼로레아(International Baccalaureate, 이하 IB)는 국제 교육과정으로 초등(PYP), 중등(MYP) 프로그램이 있고, 대학 입학 자격으로 활용되는 디플로마 프로그램(DP) 및 직업 관련 프로그램(CP)으로 구성됩니다. 이중 우리나라 고등학교 서·논술형 평가 논의와 가장 관련성이 높은 것은 16~19세 대상의 2년 과정인 DP입니다. DP 학생들은 6개 교과군에서 과목을 이수하는데, 평가 방식에 있어 학교 교사에 의한 내부평가와 IB 본부 주관의 외부평가가 결합하는 독특한 구조를 지닙니다. 대부분의 과목은 외부 필기시험의 비중이 가장 높지만, 과제 및 탐구 활동 등 학교 내부 평가도 함께 반영됩니다(IBO, 2026-08-18 검색).

그럼 IB에서는 실제로 서·논술형 평가가 어떻게 이루어질까요? IB의 외부평가의 문항은 짧은 에세이, 구조화된 문제, 단답형, 자료 제시형, 사례 연구, 드물게 선다형을 사용합니다. 내부평가도 역시 구술 발표, 현장조사, 실험, 예술 활동 등 다양한 방식으로 진행됩니다(박혜영 외, 2020). 특히 주목할 점은 평가의 신뢰성을 담보하는 채점 체계입니다. 일선 교사가 전체 학생을 채점하고, 학교에서 일부 표본을 IB 외부채점자에게 제출합니다. 이후 외부채점자가 이를 재채점하면서 교사의 기준과 비교, 조정하는 과정을 거칩니다(IBP 검색 2026-08-18). 이는 시험 결과의 신뢰성을 보장하기 위한 것입니다. 이러한 방식은 프랑스 바칼로레아와 비교했을 때, 상당히 구조화된 채점기준을 바탕으로 기준참조평가, 전문적 판단 및 통계적 품질관리를 복합적으로 적용한다는 점에서 차이가 있습니다.

독일의 아비투어(Abitur)

독일의 아비투어는 일반계 고등학교인 김나지움(Gymnasium)의 졸업시험이자 대학 입학 자격 시험의 역할을 동시에 수행하는 제도입니다. 아비투어는 필기시험과 구술시험으로 구성되며, 학생들이 고등학교 과정에서 선택한 과목과 심화 수준에 따라 시험 과목과 수준이 달라집니다. 시행 시기는 연방주별로 조금씩 다르지만, 대개 4-5월에 필기시험이 치러지고 이후 구술시험이 진행되어 최종 성적은 6-7월 중에 받게 됩니다.

아비투어의 필기시험의 가장 큰 특징은 명확한 요구수준 체계를 갖추고 있다는 점입니다.

- 제1수준 : 재생산 단계로 배운 지식이나 내용을 정확하게 재현하는 수준
- 제2수준 : 재조직과 전이 단계로 배운 지식을 새로운 자료나 문제에 적용하고 분석하고 관계를 파악하는 수준
- 제3수준 : 성찰과 문제 해결 단계

아비투어는 국가 단위의 시험이 아니라 연방주별로 주관하여 시행되므로, 주 간의 학력 격차를 방지하고 시험의 타당성과 난도를 일관되게 유지하는 것이 필요합니다. 이를 위해 공동 아비투어 문항풀을 구축하고 활용함으로써, 각 주의 자율성을 보장하는 동시에 전국적인 평가 기준과 요구수준을 동일하게 유지하고자 노력하고 있습니다(IQB, 2026-08-18 검색). 또, 채점 신뢰성을 위해 기대 수행 및 채점 기준 마련, 복수 채점 및 결과의 조정 등의 장치를 두고 있습니다.

구분	설명
내신평가	고교 3년 과정의 학교 평가
국가수준 대입시험 (아비투어)	학생에 따라 4~5개 과목 시행, 독일어와 수학은 필수 서술형 시험(최소 3개 과목)+구술시험(최소 1개 과목)

이상에서 살펴본 국외 사례를 통해 서·논술형 평가를 안정적으로 운영하기 위해서는 문항 개발, 채점 기준 마련, 평가자 전문성 확보, 채점 조정 및 질 관리 등 체계적인 지원 체제를 갖출 필요가 있음을 알 수 있습니다. 즉, 문항의 타당성과 채점의 신뢰성을 확보하기 위한 다양한 장치를 마련하여 운영할 필요가 있습니다. 다만, 프랑스나 독일의 경우 교육 및 평가를 둘러싼 제도적·사회적 여건이 우리나라와 다르다는 점을 고려하여 이러한 사례의 적용 가능성과 시사점을 살펴볼 필요가 있습니다.

<참고문헌>

- 교육부(2022). 초·중등학교 교육과정 총론. 교육부 고시 제2022-33호 [별책 1]
- 교육부(2025). 고등학교 서·논술형 평가 선도교원 연수 자료.
- 김수진, 김희경, 나우열, 민호기, 성경희, 이미숙, 이민형, 이영미(2025). 중등학교 성취평가 질 제고를 위한 신뢰성 확보 방안(I). 한국교육과정평가원 연구보고 RRE 2025-7.
- 뉴스1(2026.3.25.). 서울교육청, 부산·인천교육청과 'AI 서·논술형 평가지원도구' 공동 운영.
- 박종임, 김동영, 정수진, 신중휘, 이정훈, 김승주(2026). AI 자동채점은 교육개혁의 기회인가, 위협인가: 서·논술형 평가 자동채점이 나아갈 길을 모색하다(ORM 2026-42-7). 한국교육과정평가원.
- 박종임, 김유향, 박지선, 백승주, 정수진, 김진아, 김한승, 김효수, 염동렬, 최봉석, 김민정, 송민섭, 황미애, 황정숙(2024a). 고등학교 서·논술형 평가 도구 개발의 방법과 사례:사회. 한국교육과정평가원 연구자료 ORM 2024-150-1.
- 박종임, 김유향, 박지선, 백승주, 정수진, 김진아, 김한승, 김효수, 염동렬, 최봉석, 권경운, 김형성, 사민선, 최숙기(2024b). 고등학교 서·논술형 평가 도구 개발의 방법과 사례:국어. 한국교육과정평가원 연구자료 ORM 2024-150-2.
- 박종임, 김유향, 박지선, 백승주, 정수진, 김진아, 김한승, 김효수, 염동렬, 최봉석, 윤혜미, 이상조, 이승철, 최승호(2024c). 고등학교 서·논술형 평가 도구 개발의 방법과 사례:수학. 한국교육과정평가원 연구자료 ORM 2024-150-3.
- 박종임, 김유향, 박지선, 백승주, 정수진, 김진아, 김한승, 김효수, 염동렬, 최봉석, 노동규, 왕희정, 임동휘, 최현숙(2024d). 고등학교 서·논술형 평가 도구 개발의 방법과 사례:과학. 한국교육과정평가원 연구자료 ORM 2024-150-4.
- 박종임, 김유향, 박지선, 백승주, 정수진, 김진아, 김한승, 김효수, 염동렬, 최봉석, 김연희, 박지윤, 송민영, 이정우(2024e). 고등학교 서·논술형 평가 도구 개발의 방법과 사례:영어. 한국교육과정평가원 연구자료 ORM 2024-150-5.
- 박종임, 김유향, 박지선, 백승주, 정수진, 김진아, 김한승, 김효수, 염동렬, 최봉석, 맹수용, 박미송, 안민영, 오정현(2024f). 고등학교 서·논술형 평가 도구 개발의 방법과 사례:한국사. 한국교육과정평가원 연구자료 ORM 2024-150-6.
- 박종임, 송민호, 이상하, 박상욱, 최숙기, 이민정(2023). 컴퓨터 기반 서·논술형 평가를 위한 자동채점 방안 설계(II)(연구보고 RRE 2023-7). 한국교육과정평가원.
- 박혜영, 김성숙, 김경희, 이명진, 김광규, 김지영(2019b). 수업-평가 연계 강화를 통한 서·논술형 평가 내실화 방안. 한국교육과정평가원 연구보고 RRE 2019-6.
- 박혜영, 이상하, 이명진, 이미영, 최인봉, 주현우(2020). 평가를 활용한 초·중등학생의 글쓰기 능력 신장 방안 연구(I): 평가틀과 예시문항 개발. 연구보고 RRE 2020-9. 한국교육과정평가원.
- 백순근(2007). 교육측정의 이론과 실제. 서울: 교육과학사.
- 성경희, 김미경, 김유향, 백승주, 유혜영, 이민형, 이재봉, 김한승, 이승택, 최가영, 최봉석(2025). 서·논술형 평가 도구 개발의 방법과 사례. 한국교육과정평가원 연구자료 ORM 2025-91-1~5, ORM2025-92-1~2, ORM 2025-93.

- 성정원, 신병철(2023). ChatGPT를 활용한 서·논술형 평가 자동 채점 가능성 탐색: 세계지리 서·논술형 평가를 중심으로. *한국지리학회지*, 12(3), 415-432.
- 성태제 (2014). *현대교육평가*. 학지사.
- 송민호, 김동영, 박준홍, 신중휘, 이정훈, 정수진, 임윤진(2025). 교과 서·논술형 평가 자동채점을 위한 인공지능 모델 적용 방안 연구(Ⅱ)(연구보고 RRE 2025-5). 한국교육과정평가원.
- 유진은 (2019). *교육평가: 연구하는 교사를 위한 학생평가*. 학지사.
- 이혜정, 이범, 홍영일(2020). 국제 바칼로레아(IB) 프로그램 현장 안착 지원 방안 연구. 대구광역시 교육청 위탁연구보고서
- 이혜정, 조현영, 진성은, 윤혜진, 현승호, 권미애, 홍영일(2022). IB 교육 효과 분석 중단 연구(1년차). 제주특별자치도교육청 위탁연구보고서
- 최진영, 김지수, 김형성(2025). 논술형 평가 전문성 강화를 위한 AI 자동채점 기반 평가자 협의 모델 개발: 교사-AI 협업을 중심으로. *교육정보미디어연구*, 31(4), 1297-1329.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D.(2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901.
- Budzyń, K., Romańczyk, M., Kitala, D., et al.(2025). Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. *Lancet Gastroenterology & Hepatology*, 10(10), 896-903.
- Chen, J., Fife, J. H., Bejar, I. I., & Rupp, A. A.(2016). Building e-rater® scoring models using machine learning methods. *ETS Research Report Series*, 2016(1), 1-12.
- Gronlund, N. E. (1988). *How to construct achievement tests* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Jauhiainen, J. S., & Garagorry Guerra, A.(2025). Generative AI in education: ChatGPT-4 in evaluating students' written responses. *Innovations in Education and Teaching International*, 62.
- Jonsson, A., & Svingby, G.(2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.
- Kim, Y., Hardy, M., Tey, J., Thille, C., & Piech, C. J. (2026). Interpretability from the ground up: Stakeholder-centric design of automated scoring in educational assessments. In *Findings of the Association for Computational Linguistics: ACL 2026* (pp. 37313–37328). Association for Computational Linguistics.
- Klebanov, B. B., & Madnani, N.(2020). Automated evaluation of writing – 50 years and counting. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 7796-7810). Association for Computational Linguistics.
- Li, H., Chen, C. H., Fan, K., Young-Johnson, C., Lim, S., & Feng, Y. (2025). Agreement Between Large Language Models and Human Raters in Essay Scoring: A Research Synthesis. *arXiv*

preprint arXiv:2512.14561.

- Mehrens, W. A. & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology*. (4th ed.). Holt, Rinehart and Winston.
- Pack, A., Barrett, A., & Escalante, J.(2024). Large language models and automated essay scoring of English language learner writing: Insights into validity and reliability. *Computers and Education: Artificial Intelligence*, 6, 100234.
- Stahl, M., Biermann, L., Nehring, A., & Wachsmuth, H. (2024). Exploring LLM prompting strategies for joint essay scoring and feedback generation. In *Proceedings of the 19th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2024)* (pp. 283–298).
- Stemler, S. E.(2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9(4).

<부록 1> 시도교육청별 주요 서·논술형 평가 반영 비율

시·도교육청별 고등학교 학업성적관리 시행지침에 명시된 주요 서·논술형 평가 반영 비율

*지침상 비율과 실제 학교 시행은 다를 수 있습니다

교육청	내용
강원특별 자치도 교육청	모든 과목의 평가는 정기시험과 수행평가를 구분하여 실시하며, 서·논술형 평가를 포함한다. 서·논술형 문항 배점은 과목별 학기말 총 배점(정기시험과 수행평가의 반영비율 환산 점수 합계)의 30%이상 반영한다. 정기시험을 실시하는 경우 매회 서·논술형 문항을 포함하며, 서·논술형만으로 정기시험을 실시할 수 있다.
경기도 교육청	논술형 평가의 반영비율은 학기 단위 성적(정기시험과 수행평가 합산점수, 수행평가100% 실시 과목 점수)의 30% 이상으로 한다.
경상남도 교육청	정기시험에는 서·논술형 평가를 정기시험 배점의 40% 이상 포함하여야 한다. (수행평가의 서·논술형 평가는 '정기시험의 서·논술형 평가 배점 40% 이상'에 포함되지 않음.)
경상북도 교육청	서·논술형 평가는 정기시험 또는 수행평가 영역에 포함하여 실시하며, 학기말 배점 기준의 20% 이상 반영한다. -서·논술형 평가를 정기시험에 포함하여 실시할 경우, 정기시험으로 실시하는 서·논술형 평가는 수행평가에 포함되지 않는다.
광주광역시 교육청	국어, 사회(역사/도덕 포함), 수학, 과학, 영어 교과와 서·논술형 평가의 반영 비율은 학기당 정기시험과 수행평가를 합하여 총 배점의 30% 이상 되어야 하며, 정기시험에서는 서·논술형 평가를 반드시 실시해야 한다.
대구광역시 교육청	국어, 사회(역사/도덕 포함), 수학, 과학, 영어 학기말 환산점 100점 중 20% 이상을 정기시험 내 서·논술형 평가로 출제하여야 한다. 국어, 사회(역사/도덕 포함), 수학, 과학, 영어 교과에서 다음의 요건을 모두 충족하는 경우, 정기시험 내 서·논술형 평가(20%)를 실시한 것으로 갈음할 수 있다. -수행평가 반영 비율을 80% 이상으로 실시 -수행평가 영역 중 논·구술의 과정중심 수행평가를 20% 이상 실시
대전광역시	서·논술형 평가는 정기시험 또는 수행평가 영역에 포함하여 실시할 수 있으며, 반영 비율은 학기 단위 과목 총점(정기시험과 수행평가 환산점수의 합)의 30% 이상

교육청	내용
교육청	으로 한다. 수행평가 비율이 70% 이상인 과목은 학업성적관리위원회의 심의를 거쳐 서·논술형 평가를 실시하지 않을 수 있다.
부산광역시 교육청	선택형 일변도의 평가를 지양하고 학생의 사고력, 문제해결력, 창의력 등 고등 사고 기능을 평가할 수 있는 서·논술형 문항은 국어, 수학, 영어, 한국사, 사회(역사, 도덕 포함), 과학 교과 정기시험에 포함하여 출제하고, 배점은 매 정기시험 만점의 30% 이상 의무 반영하며, 그 외 교과는 반영 여부 및 비율을 교과협의회에서 결정한다.
세종특별 자치시 교육청	서술형·논술형 평가는 정기시험 또는 수행평가 영역에 포함하여 실시할 수 있으며, 한 학기 환산점 100점 중 30% 이상(학기별 1단위로 개설되는 과목과 수행평가 비율이 한 학기 환산점 100점 중 80% 이상인 경우 예외) 반영하여야 한다.
울산광역시 교육청	서·논술형 평가는 정기시험에 포함하여 실시하며, 정기시험 총 배점의 20% 이상 반영한다. 모든 교과는 정기시험을 실시할 경우 서·논술형 평가를 실시하여야 한다. 다만, 다음의 경우에는 서·논술형 평가를 실시하지 않을 수 있다. -학기당 1학점으로 운영되는 경우 -국어, 수학, 영어, 한국사, 사회(역사/도덕 포함), 과학 교과(군) 이외의 교과 중 수행평가 반영 비율이 환산총점의 52% 이상인 경우
전라남도 교육청	(1·2학년) 정기시험을 실시하는 과목은 서·논술형 평가 문항을 매 정기시험 만점의 20% 이상 출제하되, 체육·예술 교과와 전문교과는 10% 이상 출제한다. 이에 대한 구체적인 반영비율 및 방법은 교과협의회에서 자율적으로 정하고, 이를 학업성적관리위원회의 심의를 거쳐 학교장이 최종 결정한다. ※ 국어, 영어, 수학, 사회(역사 포함), 도덕, 과학 교과는 정기시험 또는 수행평가에서 논술형 평가를 학기말 환산점의 5%이상 출제를 권장함. (3학년) 정기시험을 실시하는 과목은 서·논술형 평가 문항을 매 정기시험 만점의 20% 이상 출제하되, 체육·예술 교과와 전문교과Ⅱ는 10% 이상 출제한다.
전북특별 자치도 교육청	서술(논술)형 문항은 국어, 수학, 영어, 사회(역사/도덕 포함), 과학 교과(군)의 경우 정기시험 총 배점의 20% 이상 출제하며, 이외의 교과는 가급적 서술(논술)형 평가 문항을 포함하여 출제한다. ※ 서술(논술)형 배점은 학기별 정기시험 총 배점의 평균 20% 이상이 되어야 함. - 정기시험에서 실시하기 곤란한 논술형 평가를 수행평가에서 실시하고자 하는 경

교육청	내용
	우 배점은 학기당 과목별 총점의 20% 이상으로 한다. 이러한 경우에 정기시험에서는 서답형(단답형, 완성형)의 비율을 시험별 총 배점의 10% 이상으로 한다. - 정기시험을 학기당 2회 실시하는 과목에 한하여, 그 중 1회를 교과별(학년별) 교수학습 및 평가 운영 계획에 따라 서술(논술)형 문항만으로 평가할 수 있다.
제주특별 자치도 교육청	국어, 수학, 영어, 한국사, 사회(역사/도덕 포함), 과학 교과(군)의 과목은 정기시험과 수행평가를 합친 매 학기 환산총점(100점)의 20% 이상을 서·논술형 평가로 반영하여야 한다. 정기시험을 실시하는 교과(군) 중 서답형 평가는 정기시험 배점의 30% 이상으로 하되, 국어, 수학, 영어, 한국사, 사회(역사/도덕 포함), 과학 교과(군)는 서·논술형 평가를 정기시험 배점의 20% 이상으로 출제한다.
충청남도 교육청	국어, 수학, 영어, 한국사, 사회(역사/도덕 포함), 과학 교과(군)는 서·논술형 평가를 정기시험과 수행평가 영역에 포함하여 실시할 수 있으며, 반영 비율은 한 학기 환산점 100점(과목 만점) 중 20% 이상 반영해야 한다.
충청북도 교육청	국어, 수학, 영어, 한국사, 사회(역사/도덕 포함), 과학교과(군)는 서술형·논술형 평가를 한 학기 환산점 100점 중 20%이상 반영하여야 하고, 그 외 교과(군)는 단위학교 학업성적관리위원회에서 결정한다.

※ 서울과 인천의 경우 지침 상에 서·논술형 평가 시행 비율이 별도로 제시되어 있지 않음.

<부록 2> 대입의 주요 전형 요소

우선 대입의 주요 전형 요소를 살펴보겠습니다. 우리나라 대입에서는 다음과 같은 세 가지 전형 요소를 주로 사용합니다. (2028년 대학 신입생 기준)

전형 요소	설명
① 교과 내신 평가 및 학교생활기록부 기재 내용 (고등학교에서 산출)	- 교과 내신 평가 (= 수행평가 + 정기고사) - 성적 표기 : 원점수, 평균, 성취도(A, B, C, D, E), 석차 등급(1등급~5등급), 수강자 수 - 교과 내신 평가 이외의 학교생활기록부 기재 내용 - 교과 활동 및 비교과(예: 동아리) 활동 등 학생의 학교생활 전반을 기록한 자료
② 수능 (국가 수준에서 실시)	- 시험 과목 : 국어, 수학, 영어, 한국사, 통합사회, 통합과학, 성공적인 직업생활, 제2외국어/한문 - 성적 표기 : 표준점수, 백분위, 등급
③ 논술·면접·실기 (대학에서 실시)	대학 자율로 일부 대학에서 실시

이와 같은 세 가지 전형 요소는 그 중 한 가지 또는 몇 가지가 조합되어 대입 전형에 활용됩니다. 여러 가지 조합에 따른 대입 전형의 유형을 정리하면 다음과 같습니다. (2026년 대학 신입생 기준)

모집시기	전형 유형 (전형 요소)		전형 요소 활용 방법	수능 활용
수시 (79.9%)	학교생활기록부 (68.6%)	교과(수행평가+정기고사) (45.0%)	정량 평가	필요시 최저 학력기 준으로 수능 등급 활용
		종합(교과+비교과) (23.6%)	정성 평가	
	(대학별) 논술(3.6%)		교육과정 범위내의 대학별 논술	
	(대학별) 실기(6.3%)		예·체능 실기	
정시 (20.1%)	수능(18.5%)		표준점수, 백분위, 등급	각 대학별 기준에 따라 반영
	(대학별) 실기(1.4%)		예·체능 실기	

This image shows a full page of a document template. It consists of approximately 20 evenly spaced, horizontal dashed lines across the entire width of the page. The background is plain white, and there are no margins, headers, or footers visible. This type of template is commonly used for teaching handwriting to children or as a guide for letter height in professional writing.

[illegible]

This image shows a full page of a document template designed for handwriting practice or general writing. It consists of approximately 20 evenly spaced, horizontal dashed lines running across the entire width of the page. The background is plain white, and there are no margins, headers, footers, or other markings present.

[illegible]