

최종보고서

2025년 국민의 국어능력 실태 조사

- [제3차] 쓰기 영역 -

2025. 11.

서울대학교 산학협력단

산학협력단장

제 출 문

국립국어원장 귀하

국립국어원과 체결한 연구용역 계약에 따라 ‘2025년 국민의 국어능력 실태 조사’에 관한 연구 보고서를 작성하여 제출합니다.

■ 사업 기간: 2025년 3월 ~ 2025년 11월

2025년 11월 28일

연구책임자: 구본관(서울대학교 국어교육과 교수,
서울대학교 국어교육연구소 겸무연구원)

연구 기관: 서울대학교 국어교육연구소, (주)한국갤럽조사연구소

연구책임자: 구본관

공동연구원: 민병곤, 김혜정, 김승현, 조진수, 박성석, 최소영, 장지혜, 이성준,
장호원, 백형근

보조연구원: 이상재, 허모아, 안이진, 이정민, 이진현, 지서영, 이혜린, 김고은

2025년 국민의 국어능력 실태 조사

본 연구의 목적은 2024년에 개발한 쓰기 영역의 조사 도구를 보완하여 쓰기 영역의 국어능력 본조사를 실시하고, 그 결과를 분석하여 국민의 쓰기 능력 향상을 위한 정책 방안 수립의 기초 자료를 마련하는 데 있다. 또한, 2023년부터 2025년에 걸쳐 실시된 제3차 실태 조사 6개 영역의 조사 결과를 종합적으로 분석하고, 2013년의 제1차 조사, 2018년의 제2차 조사, 2023년부터 2025년까지 이루어지는 제3차 조사의 결과를 연계하여 분석하였다.

제3차 국민의 국어능력 실태 조사가 영역별 연차 조사의 방식으로 시행됨에 따라 본 연구의 본조사는 쓰기 영역을 대상으로 시행되었다. 본조사의 조사 도구는 2024년에 이루어진 기초 연구 및 예비 조사에 기반을 두어 설계하되, 전문가 자문을 비롯하여 예비 조사 결과에서 확인된 문항의 변별도와 곤란도 등의 정보를 바탕으로 보완 후 선정하였다. 이에 더하여, 본 연구에서는 쓰기 영역에 특화된 심층 배경 변인을 설정하여 쓰기 능력을 보다 심층적으로 살펴보고자 하였다. 이에 따라 인구통계학적 정보 관련 변인 외에 글쓰기에 대한 인식, 글쓰기 빈도, 긍정적인 글쓰기 경험 빈도, 쓰기 과정에서 생성형 인공지능 활용 빈도, 글쓰기 조력자의 존재 등을 쓰기 영역 심층 변인으로 설정하였다.

본조사는 2025년 6월부터 8월까지의 두 달가량에 걸쳐, 전국 17개 시도의 만 20세~69세의 성인 3,000명을 대상으로 하여 1:1 방문 면접 조사로 실시하였다. 조사 참여자 선정 시 인구 분포 통계를 활용한 2단계 층화 표본 추출 및 크기 비례 확률 추출법을 적용하여 조사 대상의 대표성을 확보하였다. 이에 더해, 조사의 안정성 및 신뢰도 제고 방안 탐색 차원에서 50명을 대상으로 집합 조사를 추가적으로 실시하였다. 따라서 본조사의 총인원은 3,050명이며, 단 조사 결과에 대한 주요 분석 대상은 모집단의 대표성을 확보한 3,000명의 응답으로 하고 집합 조사 50명의 응답은 방문 면접 조사 방식과의 비교를 위해 활용되었다.

쓰기 영역의 원점수 평균은 169.99점(300점 만점, 표준편차 42.11)이었으며, 능력 점수(T점수) 평균은 50점(100점 만점, 표준편차 8.60)이었다. 수준별 인원 분포의 경우 4수준이 335명(11.2%), 3수준이 791명(26.4%), 2수준이 1,216명(40.5%), 1수준이 658명(21.9%)으로 나타났다. 변인별로 성별에 따른 능력 점수 차이는 통계적으로 유의하지 않았다. 연령대의 경우 능력 점수가 30대, 20대, 40대, 50대, 60대 순으로 높게 나타났으며, 이러한 차이가 통계적으로 유의하였다. 거주 지역 규모의 경우 대도시, 중소 도시, 읍면 지역의 순으로 능력 점수가 높게 나타났고, 이러한 차이가 통계적으로 유의하였다. 직업군에 따라서는 정신노동, 기타, 육체노동의 순으로 능력 점수가 높게 나타났으며 이러한 차이가 통계적으로 유의한 것으로 나타났다. 학력의 경우 대재 이상, 고졸, 고졸 미만의 순으로 능력 점수가 높게 나타났으며 이러한 차이가 통계적으로 유의한 것으로 나타났다.

또한 이 연구에서는 3차 실태 조사 결과를 바탕으로 수정된 앵고프 방법을 활용하여 각 조

사 영역의 성취 수준별 분할 점수를 새로 산출하였다. 수준 설정에는 국어교사, 국어교육 전공 대학원생, 국어교육 전공 교수 각 4인이 전문가 패널로 참여하였다. 수준 설정 협의회는 문항 유형에 따라 선택형 문항과 수행형 문항으로 나뉘어 2회 실시되었다. 수준 평정은 3회의 라운드로 이루어졌으며, 패널에게 각 라운드 전 수행 수준 기술문, 본조사 기술 통계, 라운드 판정 결과에 대한 정보를 제공하였다. 성취 수준 도출 결과 듣기 영역은 4수준 40.6%, 3수준 27.7%, 2수준 25.7%, 1수준 6.0%, 듣기·말하기 영역은 4수준 23.3%, 3수준 34.1%, 2수준 25.6%, 1수준 17.0%, 말하기 영역은 4수준 18.1%, 3수준 29.8%, 2수준 32.3%, 1수준 19.9%, 읽기 영역은 4수준 33.0%, 3수준 28.5%, 2수준 24.9%, 1수준 13.7%, 문법·규범 영역은 4수준 29.6%, 3수준 37.3%, 2수준 23.5%, 1수준 9.6%, 쓰기 영역은 4수준 11.2%, 3수준 26.4%, 2수준 40.5%, 1수준 21.9%로 나타났다.

위의 성취 수준 도출 결과에 따라 제3차 국민의 국어능력 실태 조사에서 실시된 6개 영역의 결과에 대한 종합 결과 분석을 실시하였다. 우선 인구통계학적 변인인 성별, 연령대, 지역 규모, 직업군, 학력과 관련하여, 듣기 영역에서는 성별에 따른 성취 수준의 등급별 인원 분포 차이는 통계적으로 유의하지 않았으나, 연령대, 거주 지역의 규모, 직업군, 학력 변인에서는 유의한 차이가 나타났다. 이어서 듣기·말하기 영역에서 또한 성별에 따른 성취 수준의 등급별 인원 분포 차이는 통계적으로 유의하지 않았으나, 연령대, 거주 지역의 규모, 직업군, 학력 변인에서는 유의한 차이가 나타났다. 그 반면 말하기 영역에서는 성별을 포함하여 연령대, 거주 지역의 규모, 직업군, 학력 변인의 모든 변인에서 성취 수준의 등급별 인원 분포에 통계적으로 유의한 차이가 나타났다. 한편 읽기 영역에서는 성별에 따른 성취 수준의 등급별 인원 분포 차이는 통계적으로 유의하지 않았으나, 연령대, 거주 지역의 규모, 직업군, 학력 변인에서는 유의한 차이가 나타났다. 문법·규범 영역에서는 성별에 따른 성취 수준의 등급별 인원 분포 차이는 통계적으로 유의하지 않았으나 연령대, 거주 지역의 규모, 직업군, 학력 변인에서는 유의한 차이가 나타났다. 쓰기 영역에서는 성별에 따른 성취 수준의 등급별 인원 분포 차이는 통계적으로 유의하지 않았으나 연령대, 거주 지역의 규모, 직업군, 학력 변인에서는 유의한 차이가 나타났다.

영역별 심층 변인에 관련해서는, 듣기 영역, 듣기·말하기 영역, 말하기 영역 모두에서 ‘직접 만나서 하는 대화’, ‘누리소통망(SNS)이나 메신저로 하는 대화’ 등과 같은 심층 변인에 따라 유의한 점수 차이가 나타났다. 대면 대화의 빈도가 높을수록, 누리소통망(SNS)이나 메신저로 하는 대화의 빈도가 높을수록 듣기, 듣기·말하기, 말하기와 같은 구어 의사소통 능력 점수가 높은 것으로 나타났다. 읽기 영역에서는 ‘읽기 빈도’와 ‘1회 읽기 분량’, ‘읽기의 가치에 대한 인식’ 등과 같은 심층 변인에 따라 유의한 점수 차이가 나타났다. 읽기 빈도가 잦을수록, 한번에 읽는 분량을 가리키는 1회 읽기 분량이 길수록 읽기 능력 점수가 높은 것으로 나타났다. 문법·규범 영역에서는 ‘어문 규범과 관련된 정보 탐색 빈도’ 및 ‘<온라인가나다> 이용 빈도’에 따라 유의한 점수 차이가 나타났으나, 매우 자주 찾아보거나 이용했다는 응답의 빈도가 매우 낮았다. 쓰기 영역에서는 우선 학력이 ‘대재 이상’인 경우 그렇지 않은 경우에 비해 쓰기 능력 수준이 높게 나타났으며, 비자발적(공적) 글쓰기 빈도가 높을수록 쓰기 능력 수준이 높은 것으로 나타났다. 또한 학력이 ‘고졸 미만’인 경우 그렇지 않은 경우에 비해 쓰기 능력 수준이 낮게 나타났으며, 지역이 대도시인 경우 그렇지 않은 경우에 비해 쓰기 능력 수준이 높게 나타났

고, 자기 이해와 생각·감정 표현을 위해 글쓰기가 필요하다고 인식할수록 쓰기 능력 수준이 높은 것으로 나타났다.

한편, 본 연구에서는 제1차와 제2차 조사에서의 가교 문항과 제2차와 제3차 조사에서의 가교 문항의 정답률을 비교하고, 정답률의 증감 정도를 효과 크기로 계산하여 국어능력의 추이를 개략적으로 파악하였다. 제1차(2013), 제2차(2018), 제3차(2023-2025) 국민의 국어능력 실태 조사를 비교한 결과, 전체적인 국어능력은 약 10여 년간 소폭 상승한 것처럼 보이지만 효과 크기 기준에서는 뚜렷한 변화로 보기 어렵다. 제1차와 제2차 조사 비교에서는 듣기, 읽기, 쓰기, 문법 영역이 모두 변화가 거의 없는 수준으로 나타났고, 말하기만 작은 규모의 상승이 확인되었다. 제2차와 제3차 조사 비교에서도 대부분의 영역에서 효과 크기가 작은 효과 크기의 기준에 미치지 못해 의미 있는 변화로 해석되기 어렵다. 특히 제3차 조사는 영역 구성과 조사 방식이 일부 조정되었음에도 국민의 국어능력 자체는 크게 변하지 않은 것으로 나타났으며, 다만 쓰기와 말하기 영역에서 상대적으로 완만한 향상 가능성이 제기된다. 종합적으로 볼 때 세 차례 조사 기간 동안 국민의 국어능력 전반은 안정적으로 유지되어 왔으며 영역별 미세한 차이는 존재하나 전체적 변화는 제한적이라고 보인다.

위의 분석을 통해 실태 조사 결과의 시사점을 도출하고 국어 정책으로 활용할 수 있는 방안을 영역별로 제시하였다. 듣기, 듣기·말하기, 말하기 영역에서는 연령대가 높아질수록 4수준의 비율이 줄고 1수준의 비율이 늘어나는 경향이 있다는 점, 지역 규모 및 매체 사용 양상에 따라 듣기, 듣기·말하기, 말하기 능력 점수에 차이가 나타난다는 점을 바탕으로 세대별 언어 경험의 차이, 지역 간 언어문화 자원의 불균형, 디지털 기반 소통 환경의 급격한 변화, 의사소통을 통한 상호작용 경험의 차이 등을 고려하여 구어 의사소통 능력을 지원할 수 있는 정책을 마련해야 한다는 점을 밝혔다. 특히 이러한 요인 간의 격차는 단기에 해소될 문제가 아니므로 데이터 기반 정책 피드백 체제를 구축하여 지속적인 지원이 필요함을 제안하였다. 읽기 영역에서는 연령대가 높아질수록 4수준의 비율이 줄고 1수준 비율이 늘어나는 경향이 나타난다는 점, 읽기 능력 점수가 매체 사용 양상에 따라 상이하게 나타난다는 점을 바탕으로 연령층과 디지털 격차를 고려한 문해력 교육 정책이 필요하다는 점을 밝혔다. 특히 디지털 기술 발전과 정보 환경의 변화 속에서 기존의 문자 중심 문해력의 의미가 비판적, 디지털, 시민 문해로 확장되고 있음을 지적하며, 근거 기반의 비판적 읽기 교육, 전인적 읽기 교육, 평생 문해 정책과 같은 교육적 및 정책적 접근이 필요함을 제안하였다. 문법·규범 영역에서는 연령대가 높아질수록 4수준 비율이 줄고 높은 연령대에서 1수준 비율이 상대적으로 높게 나타난다는 점과 학력에 따라 1수준 비율에 차이가 크다는 점을 바탕으로 노년층 교육에 대한 정책 지원과 학력 격차를 고려한 지원책 마련이 필요하다는 점을 지적하였다. 또한 국립국어원에서 제공하고 있는 국어사전, 어문 규정, <온라인가나다> 등의 문법·규범 관련 웹페이지를 사용자 친화적으로 개편하여 접근성과 활용도를 높여야 함을 제안하였다. 쓰기 영역에서는 쓰기 능력의 양극화가 뚜렷해지고 있으며, 특히 세대 간, 거주 지역 간, 학력 간 격차가 두드러진다는 점을 바탕으로 노년층의 ‘성인기 이후의 쓰기 실천 유지’를 위한 학습 환경 조성이 필요하며 고등교육 접근 기회의 확대 정책과 기초 문해력 강화 정책이 병행되어야 한다는 점, 직업군별·지역별로 접근 가능한 실용형 교육 모델을 구축해야 한다는 점, 쓰기 기능뿐만 아니라 쓰기 인식 및 쓰기 태

도를 제고할 수 있는 쓰기 교육이나 쓰기 경험을 제공하고 관련한 쓰기 문화를 조성해야 한다는 점을 제언하였으며, 향후 디지털 매체 글쓰기 역량에 대한 정교한 논의가 이루어질 필요가 있음을 언급하였다.

주요어: 국어능력, 국어능력 평가, 국민의 국어능력, 듣기 능력, 말하기 능력, 듣기·말하기 능력, 읽기 능력, 문법·규범 능력, 쓰기 능력

차 례

제 1 장 서 론

1.1. 연구의 목적 및 필요성	2
1.2. 연구의 범위	5
1.3. 연구의 추진 과정	7

제 2 장 쓰기 영역 본조사

2.1. 조사 도구의 확정	10
1) 최종 문항 선정	10
2) 문항별 수정 사항	11
3) 영역 배경 변인 구성	16
(1) 인구통계학적 변인	16
(2) 쓰기 영역 심층 변인	17
2.2. 표본 집단 추출	20
1) 방문 면접 조사	20
(1) 모집단 검토	20
(2) 모집단 층화	21
(3) 표본 크기 결정	21
(4) 집락 크기 결정	22
(5) 크기비례확률추출법을 통한 집락 추출	26
(6) 집락 내 개인 할당 추출	27
(7) 할당 추출을 위한 변인 설정	27
(8) 표본 배분	28
(9) 할당 변인별 표본 수	30

차 례

2) 집합 조사	32
2.3. 본조사 방식	33
1) 조사 방식의 타당성 확보 방안	33
(1) 방문 면접 조사와 집합 조사	33
(2) 태블릿 피시 활용과 종이 설문지 활용	33
(3) 조사 도구 개발 및 전자설문지 구현	33
(4) 쓰기에 특화된 TAPI 시스템 개발	34
2) 방문 면접 조사	35
(1) 조사 방법	35
(2) 조사원 선발 및 교육	36
(3) 방문 면접 조사 준비 및 실시	38
(4) 실사 품질관리	38
(5) 사후 검증	39
(6) 응답자 신뢰도 제고 및 불만 대응	40
3) 집합 조사	41
2.4. 문항 채점	42
1) 채점 설계	43
2) 채점단 협의회	43
3) 세부 채점 기준	46
4) 채점 방법	52
5) 채점자 신뢰도 분석	53

차 례

2.5. 본조사 결과 분석	55
1) 방문 면접 조사 결과	55
(1) 쓰기 능력 조사 결과 분석	55
(2) 쓰기 영역 문항 카드	57
(3) 배경 변인에 따른 쓰기 능력의 차이 분석	65
2) 조사 방식에 따른 결과 비교	67
2.6. 외국의 자국어 능력 조사와의 비교 분석	68

제 3 장 실태 조사 성취 수준 도출

3.1. 진행 절차	74
1) 참여 전문가 패널의 선발	74
2) 수준 설정 협의회 진행 과정	75
3) 수준별 최소 능력 보유자 협의	76
4) 분할 점수 산출 방법	81
3.2. 성취 수준 도출 결과	84
1) 영역별 분할 점수	84
2) 평정 신뢰도 분석	85

차 례

제 4 장 제3차 실태 조사 6개 영역 종합 결과 분석

4.1. 성취 수준 재도출에 따른 분석 결과	90
1) 듣기, 듣기·말하기, 말하기 영역 결과	90
(1) 듣기 영역 결과	90
(2) 듣기·말하기 영역 결과	95
(3) 말하기 영역 결과	100
2) 읽기 영역 결과	105
3) 문법·규범 영역 결과	110
4) 쓰기 영역 결과	114
4.2. 영역별 심층 변인에 대한 분석	119
1) 듣기, 듣기·말하기, 말하기 영역 심층 변인 분석	119
2) 읽기 영역 심층 변인 분석	134
3) 문법·규범 영역 심층 변인 분석	142
4) 쓰기 영역 심층 변인 분석	144

제 5 장 제1차~제3차 실태 조사 비교

5.1. 제1차~제3차 조사 시기 국민의 국어능력 추이 요약	150
5.2. 제1차 조사 시기 대비 제2차 조사 시기 국민의 국어능력 추이	154
5.3. 제2차 조사 시기 대비 제3차 조사 시기 국민의 국어능력 추이	157

차 례

제 6 장 제언

6.1. 실태 조사 결과의 시사점	160
1) 듣기, 말하기, 듣기·말하기 영역	160
2) 읽기 영역	162
3) 문법·규범 영역	164
4) 쓰기 영역	165
6.2. 국어 정책 활용 방안	169
 참고 문헌	176
Abstract	178
[붙임 1] 본조사 문항 최종본: 쓰기 영역	183
[붙임 2] 본조사 조사원 지침서	191

표 차례

<표 I -1> 연구의 좌표	2
<표 I -2> 연구 내용의 범위(쓰기 영역)	5
<표 I -3> 연구 내용의 범위(종합 및 연계 분석)	6
<표 I -4> 단계별 추진 일정 및 세부 사업 내용	7
<표 II -1> 조사 도구의 구성	10
<표 II -2> 문항 검토 및 선정 절차	10
<표 II -3> 자문 전문가 및 자문 기간	10
<표 II -4> 쓰기 영역 최종 선정 문항	11
<표 II -5> 방문 면접 조사 및 집합 조사 개요	20
<표 II -6> 표본 크기 결정	22
<표 II -7> 17개 시도별 읍면동 현황(2025년 3월 주민등록인구통계 기준) ..	22
<표 II -8> 본조사 추출 집락 및 집락별 조사 완료 표본 수	22
<표 II -9> 크기비례확률추출을 이용한 동(집락) 추출 과정(예시)	27
<표 II -10> 표본 추출을 위한 변인 설정	28
<표 II -11> 2025년 3월 기준 표본 할당표	29
<표 II -12> 시도별, 성별, 연령별 표본 수	30
<표 II -13> 시도별/학력별 표본 수	31
<표 II -14> 시도별/직업별 표본 수	32
<표 II -15> 집합 조사 표본 할당표	32
<표 II -16> 본조사(방문 면접 조사) 진행 절차	35
<표 II -17> 본조사(방문 면접 조사) 조사원 교육 결과	37
<표 II -18> 지역 사무소별 검증 현황	40
<표 II -19> 집합 조사 진행 과정	41

표 차례

<표 II-20> 집합 조사 진행 일시 및 장소	42
<표 II-21> 집합 조사 진행	42
<표 II-22> 채점 자료 할당	43
<표 II-23> 본조사 채점단	43
<표 II-24> 쓰기 1번 문항 세부 채점 기준	46
<표 II-25> 쓰기 2번 문항 세부 채점 기준	48
<표 II-26> 쓰기 3번 문항 세부 채점 기준	49
<표 II-27> 쓰기 4번 문항 세부 채점 기준	51
<표 II-28> 쓰기 영역 문항별 채점자 간 신뢰도(총점 기준)	54
<표 II-29> 쓰기 영역 문항별*채점 기준별 채점자 간 신뢰도	54
<표 II-30> 영역별 원점수 요약(쓰기)	55
<표 II-31> 쓰기 영역 문항 특성 분석 결과	55
<표 II-32> 쓰기 영역 능력 점수의 성별 차이	66
<표 II-33> 쓰기 영역 능력 점수의 연령대별 차이	66
<표 II-34> 쓰기 영역 능력 점수의 지역 규모별 차이	66
<표 II-35> 쓰기 영역 능력 점수의 직업군별 차이	66
<표 II-36> 쓰기 영역 능력 점수의 학력별 차이	67
<표 II-37> 조사 방식에 따른 쓰기 영역 능력 점수의 차이	67
<표 II-38> 평가 영역 및 대상의 비교	69
<표 II-39> NAPLAN 평가 준거 (서사)	69
<표 II-40> NAPLAN 평가 기준 (설득)	70
<표 II-41> NAEP 평가 준거	70
<표 II-42> NSBWT 평가 준거	71

표 차례

<표 II-43> 디지털 평가의 비교	71
<표 III-1> 전문가 협의회 일정	74
<표 III-2> 참여 전문가 구성	74
<표 III-3> 국어능력의 일반적 수준(국립국어원·서울대학교 국어교육연구소, 2013) ...	75
<표 III-4> 듣기 영역 수준별 최소 능력 보유자의 수행 수준 기술문	76
<표 III-5> 읽기 영역 수준별 최소 능력 보유자의 수행 수준 기술문	77
<표 III-6> 문법·규범 영역 수준별 최소 능력 보유자의 수행 수준 기술문 ·	77
<표 III-7> 듣기·말하기 영역 수준별 최소 능력 보유자의 수행 수준 기술문 ..	78
<표 III-8> 말하기 영역 수준별 최소 능력 보유자의 수행 수준 기술문	79
<표 III-9> 쓰기 영역 수준별 최소 능력 보유자의 수행 수준 기술문	80
<표 III-10> 라운드별 제공 정보	82
<표 III-11> 분할 점수 산출	84
<표 III-12> 평정 라운드별 기대 정답률 평정 범위(객관식 문항)	85
<표 III-13> 평정 라운드별 기대 정답률 평정 범위(수행형 문항)	86
<표 IV-1> 듣기 영역의 능력 점수(T점수) 요약	90
<표 IV-2> 듣기 영역의 수준 등급별 인원 분포	90
<표 IV-3> 듣기 영역 능력 점수의 성별 차이	91
<표 IV-4> 듣기 영역 수준 등급별 인원 분포의 성별 차이	91
<표 IV-5> 듣기 영역 능력 점수의 연령대별 차이	91
<표 IV-6> 듣기 영역 수준 등급별 인원 분포의 연령대별 차이	92
<표 IV-7> 듣기 영역 능력 점수의 지역 규모별 차이	92
<표 IV-8> 듣기 영역 수준 등급별 인원 분포의 지역 규모별 차이	92
<표 IV-9> 듣기 영역 능력 점수의 직업군별 차이	93
<표 IV-10> 듣기 영역 수준 등급별 인원 분포의 직업군별 차이	93

표 차례

<표 IV-11> 듣기 영역 능력 점수의 학력별 차이	93
<표 IV-12> 듣기 영역 수준 등급별 인원 분포의 학력별 차이	94
<표 IV-13> 인구통계학적 변인에 따른 듣기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과	94
<표 IV-14> 듣기·말하기 영역의 능력 점수(T점수) 요약	95
<표 IV-15> 듣기·말하기 영역의 수준 등급별 인원 분포	95
<표 IV-16> 듣기·말하기 영역 능력 점수의 성별 차이	96
<표 IV-17> 듣기·말하기 영역 수준 등급별 인원 분포의 성별 차이	96
<표 IV-18> 듣기·말하기 영역 능력 점수의 연령대별 차이	96
<표 IV-19> 듣기·말하기 영역 수준 등급별 인원 분포의 연령대별 차이	97
<표 IV-20> 듣기·말하기 영역 능력 점수의 지역 규모별 차이	97
<표 IV-21> 듣기·말하기 영역 수준 등급별 인원 분포의 지역 규모별 차이 ·	97
<표 IV-22> 듣기·말하기 영역 능력 점수의 직업군별 차이	98
<표 IV-23> 듣기·말하기 영역 수준 등급별 인원 분포의 직업군별 차이	98
<표 IV-24> 듣기·말하기 영역 능력 점수의 학력별 차이	98
<표 IV-25> 듣기·말하기 영역 수준 등급별 인원 분포의 학력별 차이	99
<표 IV-26> 인구통계학적 변인에 따른 듣기·말하기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과	99
<표 IV-27> 말하기 영역의 능력 점수(T점수) 요약	100
<표 IV-28> 말하기 영역의 수준 등급별 인원 분포	100
<표 IV-29> 말하기 영역 능력 점수의 성별 차이	101
<표 IV-30> 말하기 영역 수준 등급별 인원 분포의 성별 차이	101
<표 IV-31> 말하기 영역 능력 점수의 연령대별 차이	101

표 차례

<표 IV-32> 말하기 영역 수준 등급별 인원 분포의 연령대별 차이	102
<표 IV-33> 말하기 영역 능력 점수의 지역 규모별 차이	102
<표 IV-34> 말하기 영역 수준 등급별 인원 분포의 지역 규모별 차이	102
<표 IV-35> 말하기 영역 능력 점수의 직업군별 차이	103
<표 IV-36> 말하기 영역 수준 등급별 인원 분포의 직업군별 차이	103
<표 IV-37> 말하기 영역 능력 점수의 학력별 차이	103
<표 IV-38> 말하기 영역 수준 등급별 인원 분포의 학력별 차이	104
<표 IV-39> 인구통계학적 변인에 따른 말하기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과	104
<표 IV-40> 읽기 영역의 능력 점수(T점수) 요약	105
<표 IV-41> 읽기 영역의 수준 등급별 인원 분포	105
<표 IV-42> 읽기 영역 능력 점수의 성별 차이	106
<표 IV-43> 읽기 영역 수준 등급별 인원 분포의 성별 차이	106
<표 IV-44> 읽기 영역 능력 점수의 연령대별 차이	106
<표 IV-45> 읽기 영역 수준 등급별 인원 분포의 연령대별 차이	107
<표 IV-46> 읽기 영역 능력 점수의 지역 규모별 차이	107
<표 IV-47> 읽기 영역 수준 등급별 인원 분포의 지역 규모별 차이	107
<표 IV-48> 읽기 영역 능력 점수의 직업군별 차이	108
<표 IV-49> 읽기 영역 수준 등급별 인원 분포의 직업군별 차이	108
<표 IV-50> 읽기 영역 능력 점수의 학력별 차이	108
<표 IV-51> 읽기 영역 수준 등급별 인원 분포의 학력별 차이	109
<표 IV-52> 인구통계학적 변인에 따른 읽기 능력 점수(T점수) 설명 모형의 중다회귀분석 결과	109

표 차례

<표 IV-53> 문법·규범 영역의 능력 점수(T점수) 요약	110
<표 IV-54> 문법·규범 영역의 성취 수준 등급별 인원 분포	110
<표 IV-55> 문법·규범 영역 능력 점수의 성별 차이	110
<표 IV-56> 문법·규범 영역 성취 수준 등급별 인원 분포의 성별 차이	111
<표 IV-57> 문법·규범 영역 능력 점수의 연령대별 차이	111
<표 IV-58> 문법·규범 영역 성취 수준 등급별 인원 분포의 연령대별 차이	111
<표 IV-59> 문법·규범 영역 능력 점수의 지역 규모별 차이	112
<표 IV-60> 문법·규범 영역 성취 수준 등급별 인원 분포의 지역 규모별 차이	112
<표 IV-61> 문법·규범 영역 능력 점수의 직업군별 차이	112
<표 IV-62> 문법·규범 영역 성취 수준 등급별 인원 분포의 직업군별 차이	113
<표 IV-63> 문법·규범 영역 능력 점수의 학력별 차이	113
<표 IV-64> 문법·규범 영역 성취 수준 등급별 인원 분포의 학력별 차이	113
<표 IV-65> 인구통계학적 변인에 따른 문법·규범 능력 점수(T점수) 예측 모형의 중다회귀분석 결과	114
<표 IV-66> 쓰기 영역의 능력 점수(T점수) 요약	115
<표 IV-67> 쓰기 영역의 수준 등급별 인원 분포	115
<표 IV-68> 쓰기 영역 능력 점수의 성별 차이	115
<표 IV-69> 쓰기 영역 수준 등급별 인원 분포의 성별 차이	116
<표 IV-70> 쓰기 영역 능력 점수의 연령대별 차이	116
<표 IV-71> 쓰기 영역 수준 등급별 인원 분포의 연령대별 차이	116
<표 IV-72> 쓰기 영역 능력 점수의 지역 규모별 차이	117
<표 IV-73> 쓰기 영역 수준 등급별 인원 분포의 지역 규모별 차이	117
<표 IV-74> 쓰기 영역 능력 점수의 직업군별 차이	117

표 차례

<표 IV-75> 쓰기 영역 수준 등급별 인원 분포의 직업군별 차이	118
<표 IV-76> 쓰기 영역 능력 점수의 학력별 차이	118
<표 IV-77> 쓰기 영역 수준 등급별 인원 분포의 학력별 차이	118
<표 IV-78> 인구통계학적 변인에 따른 쓰기 능력 점수(T점수) 설명 모형의 중다회귀분석 결과	119
<표 IV-79> 듣기 영역 능력 점수의 직접 만나서 하는 대화 빈도별 차이 ..	120
<표 IV-80> 듣기·말하기 영역 능력 점수의 직접 만나서 하는 대화 빈도별 차이	121
<표 IV-81> 말하기 영역 능력 점수의 직접 만나서 하는 대화 빈도별 차이 ..	121
<표 IV-82> 듣기 영역 능력 점수의 누리소통망(SNS)이나 메신저 대화 빈도 별 차이	122
<표 IV-83> 듣기·말하기 영역 능력 점수의 누리소통망(SNS)이나 메신저 대 화 빈도별 차이	122
<표 IV-84> 말하기 영역 능력 점수의 누리소통망(SNS)이나 메신저 대화 빈 도별 차이	123
<표 IV-85> 듣기 영역 능력 점수의 음성 통화 대화 빈도별 차이	124
<표 IV-86> 듣기·말하기 영역 능력 점수의 음성 통화 대화 빈도별 차이 ·	124
<표 IV-87> 말하기 영역 능력 점수의 음성 통화 대화 빈도별 차이	125
<표 IV-88> 듣기 영역 능력 점수의 영상 통화 대화 빈도별 차이	125
<표 IV-89> 듣기·말하기 영역 능력 점수의 영상 통화 대화 빈도별 차이 ·	126
<표 IV-90> 말하기 영역 능력 점수의 영상 통화 대화 빈도별 차이	126
<표 IV-91> 듣기 영역 능력 점수의 발표·강연 듣기 빈도별 차이	127
<표 IV-92> 듣기·말하기 영역 능력 점수의 발표·강연 듣기 빈도별 차이 ·	127
<표 IV-93> 말하기 영역 능력 점수의 발표·강연 듣기 빈도별 차이	128

표 차례

<표 IV-94> 듣기 영역 능력 점수의 발표·강연하기 빈도별 차이	128
<표 IV-95> 듣기·말하기 영역 능력 점수의 발표·강연하기 빈도별 차이 ...	129
<표 IV-96> 말하기 영역 능력 점수의 발표·강연하기 빈도별 차이	129
<표 IV-97> 듣기 영역 능력 점수의 토의(회의)에서 듣고 말하기 빈도별 차이 ·	130
<표 IV-98> 듣기·말하기 영역 능력 점수의 토의(회의)에서 듣고 말하기 빈도별 차이 ·	130
<표 IV-99> 말하기 영역 능력 점수의 토의(회의)에서 듣고 말하기 빈도별 차이 ·	131
<표 IV-100> 듣기 영역 능력 점수의 토론 듣기 빈도별 차이	131
<표 IV-101> 듣기·말하기 영역 능력 점수의 토론 듣기 빈도별 차이	132
<표 IV-102> 말하기 영역 능력 점수의 토론 듣기 빈도별 차이	132
<표 IV-103> 듣기 영역 능력 점수의 토론에서 말하기 빈도별 차이	133
<표 IV-104> 듣기·말하기 영역 능력 점수의 토론에서 말하기 빈도별 차이 ...	133
<표 IV-105> 말하기 영역 능력 점수의 토론에서 말하기 빈도별 차이	134
<표 IV-106> 매체 사용 빈도(종이 신문, 잡지)별 읽기 영역 점수의 차이 ·	135
<표 IV-107> 매체 사용 빈도(텔레비전, 라디오)별 읽기 영역 점수의 차이 ·	135
<표 IV-108> 매체 사용 빈도(인터넷 포털(검색용))별 읽기 영역 점수의 차이 ·	136
<표 IV-109> 매체 사용 빈도(누리소통망(SNS), 메신저)별 읽기 영역 점수의 차이	136
<표 IV-110> 매체 사용 빈도(온라인 커뮤니티(인터넷 카페, 블로그))별 읽기 영역 점수의 차이	137
<표 IV-111> 매체 사용 빈도(온라인 동영상 플랫폼, OTT)별 읽기 영역 점수 의 차이	137
<표 IV-112> 매체 사용 빈도(게임, 웹툰, 웹소설)별 읽기 영역 점수의 차이 ...	137
<표 IV-113> 읽기 빈도별 읽기 영역 점수의 차이	138

표 차례

<표 IV-114> 독서를 하지 못하는 이유별 읽기 영역 점수의 차이	139
<표 IV-115> 읽기에 대한 가치에 따른 읽기 영역 점수의 차이	139
<표 IV-116> 읽기 효능감에 따른 읽기 영역 점수의 차이	140
<표 IV-117> 읽기 분량별 읽기 영역 점수의 차이	140
<표 IV-118> 디지털 기기 사용 시간별 읽기 영역 점수의 차이	141
<표 IV-119> 주로 이용하는 매체별 읽기 영역 점수의 차이	142
<표 IV-120> 독서 문화 참여 경험별 읽기 영역 점수의 차이	142
<표 IV-121> 문법·규범 영역 능력 점수의 어문 규범 정보 탐색 빈도별 차이 ..	143
<표 IV-122> 문법·규범 영역 능력 점수의 온라인 사전 이용 빈도별 차이 ..	144
<표 IV-123> 문법·규범 영역 능력 점수의 온라인가나다 이용 빈도별 차이 ...	144
<표 IV-124> 쓰기 영역 심층 변인 목록	145
<표 IV-125> 쓰기 영역 심층 변인의 기술 통계	145
<표 IV-126> 쓰기 영역 심층 변인(쓰기를 하지 못하는 이유)의 기술 통계 ...	145
<표 IV-127> 쓰기 능력 수준 및 전체 심층 변인 간 상관	146
<표 IV-128> 심층 변인에 따른 쓰기 능력 점수(T점수) 예측 모형의 중다회귀 분석 결과	146
<표 IV-129> 인구통계학적 변인 및 심층 변인에 따른 쓰기 능력 점수(T점 수) 예측 모형의 중다회귀분석 결과	147
<표 V-1> 제1차 조사와 제2차 조사의 영역별 국어능력 비율 추이	150
<표 V-2> 듣기 영역 제1차, 제2차, 제3차 수준별 비율	151
<표 V-3> 말하기 영역 제1차, 제2차, 제3차 수준별 비율	151
<표 V-4> 읽기 영역 제1차, 제2차, 제3차 수준별 비율	152
<표 V-5> 쓰기 영역 제1차, 제2차, 제3차 수준별 비율	152

표 차례

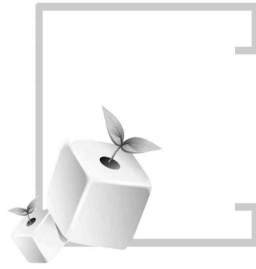
<표 V-6> 문법 및 문법·규범 영역 제1차, 제2차, 제3차 수준별 비율	153
<표 V-7> 제1차~제2차 시기 국민의 국어능력 추이 분석 결과	155
<표 V-8> 제2차~제3차 시기 국민의 국어능력 추이 분석 결과	157
<표 VI-1> 쓰기 영역의 수준 등급별 분포의 조사 시기별 비교	166
<표 VI-2> 심층 변인 중 ‘글을 충분히 쓰지 못하는 이유’에 대한 학력별 기술 통계(단위: %)	167
<표 VI-3> 심층 변인 중 ‘비자발적(공적) 글쓰기 빈도’에 대한 학력별 기술 통계(단위: %)	167

그림 차례

[그림 I -1] 연구 목적 및 세부 목표	4
[그림 II -1] 모집단 자료 현황	20
[그림 II -2] 모집단 층화 기준	21
[그림 II -3] 크기비례확률추출 예시	26
[그림 II -4] 할당 변인	28
[그림 II -5] 조사 방식	34
[그림 II -6] 쓰기 영역 TAPI 시스템 예시	35
[그림 II -7] 조사원 교육 체계	36
[그림 II -8] 조사원 교육 예시	37
[그림 II -9] 실사 통제 시스템 예시	39
[그림 II -10] 사후 검증 절차 및 질문지	40
[그림 II -11] 응답자 배부용 집합 조사 안내문	42
[그림 II -12] 사전 채점 연습	45
[그림 II -13] 채점 협의회	45
[그림 II -14] 채점 결과 입력 화면	52
[그림 II -15] 채점단 모바일 메신저 화면	53
[그림 II -16] 영역별 원점수 요약(쓰기)	55
[그림 II -17] 쓰기 영역 1번 문항의 ICC	56
[그림 II -18] 쓰기 영역 2번 문항의 ICC	56
[그림 II -19] 쓰기 영역 3번 문항의 ICC	57
[그림 II -20] 쓰기 영역 4번 문항의 ICC	57
[그림 III -1] 수준 설정 협의회 진행 과정	76
[그림 III -2] 평정 양식	82
[그림 III -3] 기술 통계	83
[그림 III -4] 분포표	84

그림 차례

[그림 V-1] 제1차~제3차 조사 시기 국민의 국어능력 추이 요약	154
[그림 V-2] 제1차~제2차 시기 국민의 국어능력 변화의 효과 크기에 대한 숲 그림 ..	156
[그림 V-3] 제2차~제3차 시기 국민의 국어능력 변화의 효과 크기에 대한 숲 그림 ..	158



제 1 장

서 론



1.1. 연구의 목적 및 필요성

‘2025년 국민의 국어능력 실태 조사’의 연구 목적은 다음과 같다. 첫째, 2024년에 개발된 조사 도구를 활용하여 쓰기 영역의 국어능력 본조사를 실시한다. 또한 그 결과 분석을 바탕으로 국민의 쓰기 능력 향상을 위한 정책 방안을 수립한다. 둘째, 2023년부터 2025년에 걸쳐 실시된 제3차 실태 조사의 6개 영역의 조사 결과를 종합적으로 분석한다. 셋째, 2013년의 제1차 조사, 2018년의 제2차 조사, 2025년까지의 제3차 조사의 결과를 연계하여 분석한다.

표 < I -1>에서 확인할 수 있듯, 본 조사는 연차별 실태 조사의 연속선상에 있다. 본 연구는 ‘국어능력 평가를 위한 기초 연구(2012)’, ‘2013년 국민의 국어능력 평가(2013)’, ‘국민의 국어능력 실태 조사 도구 개발(2017)’, ‘국민의 국어능력 실태 조사(2018)’, ‘국민의 국어능력 실태 조사 개선 기초 연구(2022)’, ‘2023년 국민의 국어능력 실태 조사(2023)’, ‘2024년 국민의 국어능력 실태 조사(2024)’를 바탕으로, ‘2025년 국민의 국어능력 실태 조사’로 이어지는 영역별, 연차별 실태 조사에 해당한다. 따라서 국어능력에 대한 기존 연구, 특히, ‘2024년 국민의 국어능력 실태 조사(2024)’와의 연속성을 중시하여, 해당 연구를 통해 기개발된 쓰기 영역 실태 조사 도구를 보완 및 활용함으로써 국어능력 추이 분석의 기틀을 마련한다.

이러한 조사 시행의 근거는 국어기본법에 규정되어 있는바, 국민의 국어능력 향상과 국어의 발전 및 보전, 국어 환경의 개선을 위한 국어능력 조사의 필요성이 명시되어 있다. 국어기본법 제9조 실태 조사에 관한 법령에서는 “문화체육관광부장관은 국어 정책의 수립에 필요한 국민의 국어능력, 국어 의식, 국어 사용 환경 등에 관한 자료를 수집하거나 실태를 조사할 수 있다.”라고 규정하고 있으며, 제23조 국어능력의 검정에 관한 법령에서는 “문화체육관광부장관은 국민의 국어능력 향상과 창조적인 언어생활의 정착을 위하여 국어능력을 검정할 수 있다.”라고 언급하고 있다. 또한 국어기본법 제6조 국어 발전 기본계획의 수립에 관한 법령에서는 “문화체육관광부장관은 국어의 발전과 보전을 위하여 5년마다 국어 발전 기본계획(이하 “기본계획”이라 한다)을 수립·시행하여야 한다.”라고 규정하며 국어의 발전을 위한 국어능력 조사가 5년을 주기로 시행되어야 할 필요가 있음을 밝히고 있다.

<표 I -1> 연구의 좌표

제1차 조사 정례적 조사 체계 마련	2013년 국민의 국어능력 평가(2013)	
	■평가 영역(5영역) 설정: 듣기, 말하기, 읽기, 쓰기, 문법 ■영역별 가교 문항의 설정 ■영역별 국어능력의 절대적 정도 측정, 국어능력 수준의 등급화	
↓		
제2차 조사 조사 도구 개발 및 안정화	국민의 국어능력 실태 조사 도구 개발(2017)	2018년 국민의 국어능력 실태 조사(2018)
	조사 도구의 내용 문제 조사 도구 개발상의 문제	■타당도, 변별도, 신뢰도 낮은 문항 재설계 ■(문법/읽기) 디지털 문식력 측정 항목 보완 ■전국 단위의 표본 선정, 연령대 확대 ■국어능력 측정을 위한 변인 다양화

		■채점자 간 신뢰도 강화, 주관식 채점 정밀화
	조사 도구 구현 및 실행상의 문제	■컴퓨터 기반의 조사 도구 구현 ■충분한 조사 시간 확보 및 사례 상향

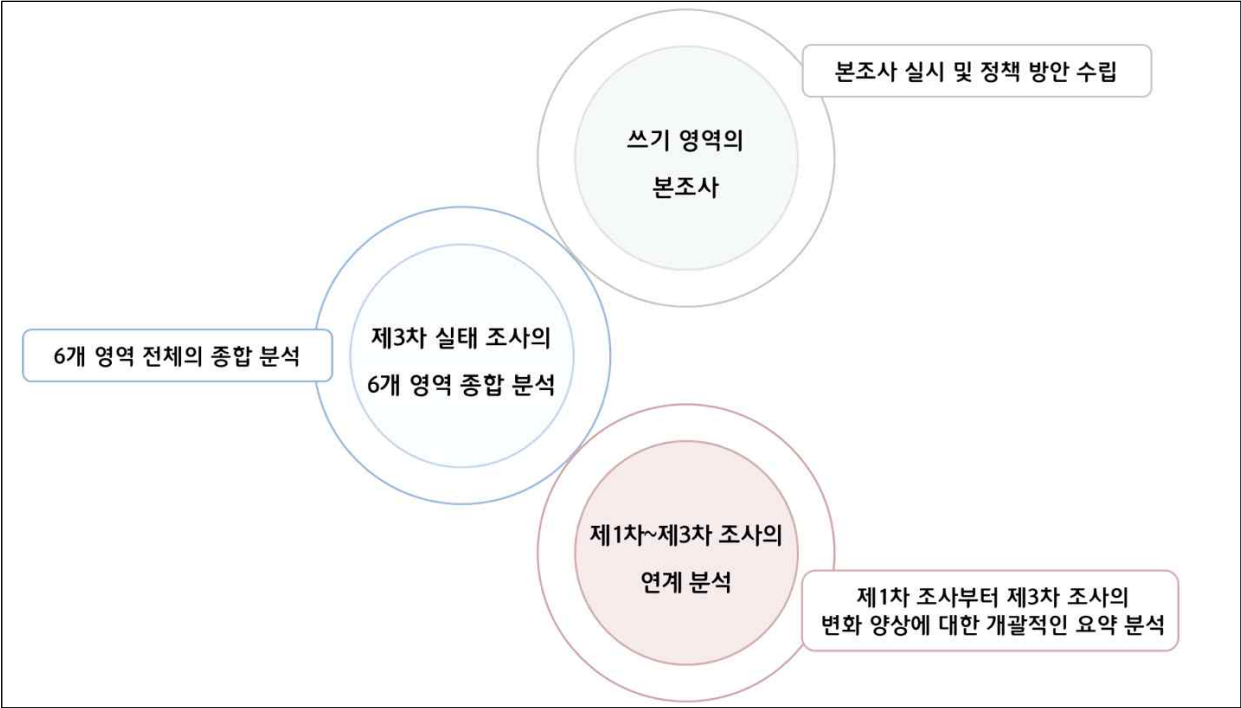
↓

제3차 조사	국민의 국어능력 실태 조사 개선 기초 연구(2022)	2023년 국민의 국어능력 실태 조사(2023)	2024년 국민의 국어능력 실태 조사(2024)
실태 조사 고도화	■영역별 ‘분할 조사’ 실시: “듣기·말하기” 영역 초점화 ■구어 의사소통의 실제성 강화 - 다양한 담화 유형 고려 - 듣기, 말하기 영역의 연계 고려 - 동영상 활용한 문항 구성 - 서답형 문항 개발 ■변화된 국어 사용 환경 반영 및 디지털(미디어) 리터러시 강화	■영역별 ‘분할 조사’ 실시: “읽기”, “문법·규범” 영역 초점화 ■변화된 국어사용 환경(디지털 리터러시 등 디지털 언어 환경)을 고려 ■문항 내 텍스트 및 담화 맥락의 실제성 강화 ■‘규범’을 하위 영역으로 추가하여 ‘문법’ 영역을 ‘문법·규범’ 영역으로 개편	■영역별 ‘분할 조사’ 실시: “쓰기” 영역 초점화 ■변화된 국어사용 환경(디지털 리터러시 등 디지털 언어 환경)을 고려 ■문항 내 텍스트 및 담화 맥락의 실제성 강화
대규모 실태 조사의 본조사 시행	2023년 국민의 국어능력 실태 조사(2023) ■기개발된 조사 도구 보완 - 담화 맥락의 실제성 반영 - 구어 담화의 역동성과 쌍방향적 소통의 특성 반영 - 피험자의 귀속적 요인 영향 배제 ■3,000명 대상의 대규모 듣기·말하기 영역의 실태 조사(본조사) 시행 ■국어능력 평가를 위한 정책 방향 입론	2024년 국민의 국어능력 실태 조사(2024) ■기개발된 조사 도구 보완 - 담화 맥락의 실제성 반영 - 의사소통 환경의 변화 반영 - 피험자의 귀속적 요인 영향 배제 ■5,050명 대상의 대규모 읽기, 문법·규범 영역의 실태 조사(본조사) 시행 ■국어능력 평가를 위한 정책 방향 입론	2025년 국민의 국어능력 실태 조사(2025) ■기개발된 조사 도구 보완 - 담화 맥락의 실제성 반영 - 의사소통 환경의 변화 반영 - 피험자의 귀속적 요인 영향 배제 ■3,050명 대상의 대규모 쓰기 영역의 실태 조사(본조사) 시행 ■국어능력 평가를 위한 정책 방향 입론

2025년 국민의 국어능력 실태 조사는 2023년, 2024년에 이어서 전 국민을 대상으로 실시되는 연도별, 연차별 대규모 조사이다. 본 연구에서는 쓰기 영역의 본조사를 시행하고, 본조사 시행 결과를 심층 분석함으로써 국민의 쓰기 능력의 변화 추이를 탐색하고 국어능력 평가를 위한 정책 방향을 입론한다. 아울러 2025년의 본조사가 제3차 조사의 마지막 차시라는 점을 고려하여 제3차 조사의 6개 영역 조사

결과를 종합하고, 제1차 조사에서부터 제3차 조사에 이르는 국어능력 변화 추이를 요약적으로 제시한다. 또한 조사 결과의 시사점을 활용할 수 있는 방안을 제시한다.

이러한 연구 목적을 달성하기 위한 연구 목표는 [그림 1-1]과 같다.



[그림 1-1] 연구 목적 및 세부 목표

[그림 1-1]에서 제시한 영역별 세부 연구 목표를 체계화하면 다음과 같다.

(1) 쓰기 영역의 본조사
상위 목표: 본조사 실시 및 정책 방안 수립 하위 목표 1. 기개발된 조사 도구의 보완 및 세부 평가 계획 수립 하위 목표 2. 본조사 결과 도출 및 분석 하위 목표 3. 국민의 쓰기 능력 실태 조사 결과의 활용 방안 마련

첫 번째로, 쓰기 영역에서는 2024년에 개발된 예비 조사 결과를 분석하여 쓰기 조사 도구를 보완한다. 이를 바탕으로 쓰기 영역의 구체적인 평가 계획을 수립한다. 또한 (1)-2에서는 쓰기 영역의 본조사 실시 후 그 결과를 다양한 변인에 따라 도출하여 쓰기 능력의 실태를 파악할 수 있는 객관적인 지표를 제시한다. 마지막으로 (1)-3에서는 본조사의 분석 결과를 활용할 수 있는 방안을 마련한다.

(2) 제3차 실태 조사의 6개 영역 종합 분석

목표: 6개 영역 전체의 평가 하위 영역과 영역별 변인에 대한 심층 분석

두 번째로, 각 영역별 평가 결과를 변인과 관련지어 분석함으로써 국어능력 실태에 관한 보다 세부적인 정보를 확인한다.

(3) 제1차~제3차 조사의 연계 분석

목표: 제1차 조사부터 제3차 조사의 변화 양상에 대한 개괄적인 요약 분석

세 번째로, 제1차부터 제3차까지 이루어진 조사 양상의 변화를 개괄적으로 파악할 수 있도록 요약 분석한다. 단, 보다 심층적인 분석은 국립국어원 제안요청서에 따라 2026년 진행으로 예정되어 있으므로 본 사업에서는 제1~3차 조사의 변화 양상을 볼 수 있는 개괄적인 요약 분석만 제시한다.

이러한 세부 목표의 달성에 따라 기대되는 본 연구의 의의는 다음과 같다.

우선, 변화된 국어 사용 환경에 따른 쓰기 영역의 역량 설정을 통한 국어능력을 구체화하는 데에 기여할 수 있다. 또한 조사 체계 및 지표, 조사 방법, 환류 방안 등을 고려한 조사 설계 방안을 체계화하여 쓰기 영역 조사 도구를 안정화 및 정교화할 것으로 기대할 수 있다. 그리고 본 연구는 대규모의 국민의 국어능력 실태 조사를 시행한 결과의 심층 분석 및 해석을 바탕으로 쓰기 영역에 특화된 정책을 제안하고 평가 시스템을 구축할 것을 기대할 수 있다.

이와 더불어, 연도별, 영역별로 시행하는 국민의 국어능력 실태 조사에 대한 체계적 관리의 기틀을 마련하고, 국내의 다양한 국어능력 평가 및 국제적인 대규모 언어 능력 조사 사업과 관련된 위상을 확보할 것이라 기대할 수 있다.

1.2. 연구의 범위

본 연구에서 수행하는 연구의 범위를 제시하자면 쓰기 영역 본조사는 <표 I-2>, 종합 및 연계 분석은 <표 I-3>과 같다.

<표 I-2> 연구 내용의 범위(쓰기 영역)

기개발된 조사 도구의 보완 및 전문가 협의회	■ 실태 조사 연구에서 개발한 조사 도구를 내용적·실행적 차원에서 보완
	■ 제1차, 제2차 국민의 국어능력 실태 조사 및 2024년 연구 성과를 평가하고, 2025년 실태 조사 계획 수립을 위한 시사점을 도출
조사 도구 설계 및 제작	■ 국민의 국어능력 신장을 목표로 하는 정책적 시사점 도출에 관여하는 변인을 설정하기 위한 전문가 협의회
	■ 2024년 예비 조사의 결과를 바탕으로 본조사에 사용할 1배수 문항 선별 및 보완
	■ 단일형 검사지로 4문항을 40분간 응시하도록 설계

	문항 형식	문항 유형	문항 수	시간
	서답형	설명적 글쓰기	1	10분
		논증적 글쓰기	1	15분
		친교적 글쓰기	1	5분
		디지털 매체 글쓰기	1	10분
	합계		4	40분
평가 실행 방식 수립 및 본조사 실시	- 평가 사전 시연 단계를 거쳐 문제점 분석 및 보완 - 태블릿 피시 사용이 불가한 경우를 대비하여 종이 답안지 제작			
	‘세대, 지역, 직업, 성별, 학력’ 변인을 고려한 표집 집단의 설정(층화 표본 추출) - 태블릿 피시를 통한 조사 방식 안정화 - 조사 요원 훈련 및 능숙도 제고			
	‘세대, 지역, 직업, 성별, 학력’ 변인을 고려하여 3,050명 표본 추출 - 전문 조사 요원을 활용한 1:1 방문 면접 조사(3,000명) - 평가의 안정성 및 신뢰도 제고를 위한 집합 조사(50명)			
본조사 결과 채점 및 결과 분석	- 본조사 자료에 대한 채점을 위한 채점단 협의회 진행 - 채점 척도에 따라 세부 구인별로 점수를 부여하는 분석적 채점을 통해 채점 신뢰도 제고			
	- 전체 결과 도출 2025년 국민의 국어능력 실태를 제시할 수 있는 국민의 객관적인 지표 도출 - 세부 결과 도출 - 고전 검사 이론 및 문항 반응 이론에 바탕한 문항별 난이도 및 변별도 분석(3,000명) - 성별, 연령대, 지역 규모, 직업군, 학력 등 배경 변인에 따른 분석(3,000명) - 조사 방식에 따른 분석(방문 면접 조사 3,000명, 집합 조사 50명 결과 비교)			
연구진 협의회 및 정책 방향 입론	분석 결과에 대한 이론적 점검 및 분석 결과의 활용 방안 마련을 위한 연구진 협의회			
	- 결과 분석을 토대로 국민의 국어능력 향상을 위한 정책 수립 방향 입론 - 주기별 국어능력 평가의 추이 분석을 위한 기반 마련			

<표 I -3> 연구 내용의 범위(종합 및 연계 분석)

제3차 실태 조사 성취 수준 도출	- 제1차 조사에서의 분할 점수 및 성취 수준 도출 방식 검토 - 참여 전문가 패널의 선발 및 사전 교육 - 앵고프 협의회를 통한 영역별 성취 수준 설정 및 분할 점수 산출
제3차 실태 조사 6개 영역 종합 결과 분석	- 듣기, 듣기·말하기, 말하기, 읽기, 문법·규범, 쓰기의 총 6개 영역의 평가 결과를 종합적으로 분석 - 성취 수준 재도출에 따른 영역별 분석 결과 도출 - 영역별 심층 변인에 대한 분석
제1차~제3차 실태 조사 비교	2013년에 실시된 제1차 조사부터 2025년에 마무리되는 제3차 조사까지의 조사에 대해 조사 도구와 조사 방식 등을 비교하고 조사 결과를 요약적으로 분석 - 세부 결과 도출 - 제1차~제3차 조사 시기 국민의 국어능력 추이 요약 - 제1차 조사 시기 대비 제2차 조사 시기 국민의 국어능력 추이 파악 - 제2차 조사 시기 대비 제3차 조사 시기 국민의 국어능력 추이 파악

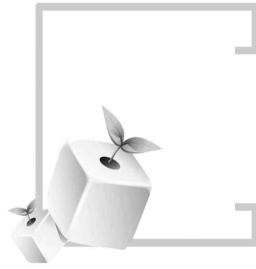
1.3. 연구의 추진 과정

본 연구의 추진 경과를 제시하자면 <표 I -4>와 같다.

<표 I -4> 단계별 추진 일정 및 세부 사업 내용

단계	추진 일정	세부 사업 내용																					
1단계 기개발된 조사 도구의 보완	착수 ~ 4월 30일	- 계약 및 연구 착수 ‘2013년 국민의 국어능력 평가’, ‘2018년 국민의 국어능력 실태 조사’, 2024년 연구에서의 예비 조사 성과 및 한계 분석 2025년 실태 조사 계획 수립을 위한 반영 사항 도출 [국립국어원 보고] - 착수 보고																					
2단계 설문 배경 변인 설정	5월 1일 ~ 5월 10일	국민의 국어능력 신장을 목표로 하는 정책적 시사점 도출에 관여하는 변인을 설정하기 위한 연구진 협의 및 전문가 자문																					
3단계 조사 도구 최종 설계	5월 중순	2024년 예비 조사의 결과를 바탕으로 조사 도구 설계 - 전문가 자문 및 자문 의견 반영 - 조사 도구 최종 설계(단일형 검사지) <table><tr><th>문항 형식</th><th>문항 유형</th><th>문항 수</th><th>시간</th></tr><tr><td rowspan="4">서답형</td><td>설명적 글쓰기</td><td>1</td><td>10분</td></tr><tr><td>논증적 글쓰기</td><td>1</td><td>15분</td></tr><tr><td>친교적 글쓰기</td><td>1</td><td>5분</td></tr><tr><td>디지털 매체 글쓰기</td><td>1</td><td>10분</td></tr><tr><td colspan="2">합계</td><td>4</td><td>40분</td></tr></table> [국립국어원 보고] - 중간 보고(본조사 계획 보고)	문항 형식	문항 유형	문항 수	시간	서답형	설명적 글쓰기	1	10분	논증적 글쓰기	1	15분	친교적 글쓰기	1	5분	디지털 매체 글쓰기	1	10분	합계		4	40분
문항 형식	문항 유형	문항 수	시간																				
서답형	설명적 글쓰기	1	10분																				
	논증적 글쓰기	1	15분																				
	친교적 글쓰기	1	5분																				
	디지털 매체 글쓰기	1	10분																				
합계		4	40분																				
4단계 평가 실행 방식 수립 및 조사 도구 제작	5월 말 ~ 6월 초	‘세대, 지역, 직업, 성별, 학력’ 변인을 고려한 표집 집단의 설정(층화 표본 추출) - 태블릿 피시를 통한 조사 방식 안정화 - 조사 요원 훈련 및 능숙도 제고 - 사전 시연 단계를 거쳐 문제점 분석 및 보완 - 태블릿 피시 사용이 불가한 경우를 대비하여 종이 답안지 제작																					

5단계 본조사 실시	6월 18일 ~ 8월 12일	▪ ‘성별, 연령, 지역, 학력, 직업’ 변인을 고려하여 3,050명 표본 추출(20~69세) ▪ 전문 조사 요원을 활용한 1:1 방문 면접 조사(3,000명) ▪ 평가의 안정성 및 신뢰도 제고를 위한 집합 조사(50명)
↓ ↓ ↓		
6단계 채점단 협의회	8월 18일	▪ 채점 신뢰도 제고를 위해 쓰기 조사 도구 및 문항, 채점 척도, 채점 수행 방법 교육 및 채점 훈련을 통한 세부 사항 협의(20명)
↓ ↓ ↓		
7단계 본조사 결과 채점	8월 말 ~ 9월 초	▪ 채점 설계와 방법에 따라 채점 진행
↓ ↓ ↓		
8단계 본조사 결과 도출 및 분석	8월 초 ~ 9월 말	▪ 전체 결과 도출 - 2025년 국민의 국어능력 실태를 제시할 수 있는 국민 3,050명의 객관적인 지표 도출 ▪ 세부 결과 도출 - 인구통계학적 변인 등에 대한 심층 분석 - 쓰기 영역 결과 도출 및 방문 면접 조사와 집합 조사 결과 비교
↓ ↓ ↓		
9단계 연구진 협의회 및 성취 수준 분석	10월 초 ~ 11월 초	▪ 분석 결과에 대한 이론적 점검 및 분석 결과의 활용 방안 마련을 위한 연구진 협의회 ▪ 분할 점수 및 성취 수준 재도출
↓ ↓ ↓		
10단계 국어능력 평가를 위한 정책 방향 입론 및 심층 분석	11월 초 ~ 11월 30일	▪ 결과 분석을 토대로 국민의 국어능력 향상을 위한 정책 수립 방향 입론 ▪ 제3차 실태 조사 종합 결과 분석 및 제1차~제3차 실태 조사 연계 분석 [국립국어원 보고] ▪ 용역 사업 최종 보고



제 2 장

쓰기 영역 본조사



2.1. 조사 도구의 확정

1) 최종 문항 선정

본조사 도구의 경우 2024년 예비 조사 결과의 양호도를 바탕으로 최종 문항을 확정하였다. 이때 2024년 예비 조사 당시 개발된 문항 중에서 본조사에 사용할 1배수 문항을 선별 및 보완(4문항)하였다. 조사 방식은 일원화된 조사 도구를 사용하여 응시자가 하나의 검사지에 응시하도록 하였다. 또한, ‘1:1 방문 면접 조사’와 ‘집합 조사’ 방식에 따른 결과치 비교를 위해 ‘개별 가구 방문 평가’와 ‘소규모 내방형 평가(집합 조사)’를 진행하였다. 방문 면접 조사와 집합 조사의 조사 도구 구성은 각각 <표 II-1>과 같다.

<표 II-1> 조사 도구의 구성

방문 면접 조사 - 3,000명			집합 조사 - 50명		
영역	문항 수	시간	영역	문항 수	시간
쓰기	4문항	40분	쓰기	4문항	40분

※ 쓰기 4문항에는 가교 문항 1개가 포함됨

2025년 본조사에서는 2024년 예비 조사에서 설계한 쓰기 영역 본조사의 조사지 구성 방식에 따라 쓰기 영역에서 사용할 최종 문항을 선정하였다. 먼저 공동 연구원들이 예비 조사의 결과를 분석하고, 결과 및 과정에서 드러난 문제를 검토하였으며(2025년 4월), 이후 전문가의 자문 및 검토를 의뢰하였다(2025년 5월).

<표 II-2> 문항 검토 및 선정 절차

단 계	1차 검토	⇒	2차 검토	⇒	3차 검토	⇒	4차 검토
	내부 검토		외부 검토		내부 검토		내부 검토
내 용	■발주 기관 요청 검토 ■문항 수정 및 1.2배수에서 최종 문항 선정		■1배수 문항에 대한 전문가 자문 검토		■연구 책임자 검토 ■발주 기관 검토 의견 반영 및 최종 문항 확정		■디자인 수정 및 TAPI 판본 검토

서면 자문을 통해 전문가 견해를 반영하였으며, 연구진 내에서 수정 방향을 협의하였다. 회의에 참여한 전문가는 다음과 같다.

<표 II-3> 자문 전문가 및 자문 기간

전문가	소속	자문 기간
권태현	충북대학교 국어교육과	2025.04.28. ~ 2025.05.10.

유상희	한국교육과정평가원
-----	-----------

전문가 협의회 후 공동 연구원은 전문가 의견 및 예비 조사를 통해 도출된 문항별 난이도, 변별도, 답지 반응을 등의 자료를 바탕으로 본조사에서 활용할 조사 도구 문항을 선별하였다. 선별된 최종 문항은 <표 II-4>와 같다.

<표 II-4> 쓰기 영역 최종 선정 문항

대영역	중영역	소영역(평가 요소)	선정 문항 번호	조사 방식
쓰기	가. 설명적 글쓰기	설명문 쓰기	W-1	서답형 (주관식)
	나. 논증적 글쓰기	논설문 쓰기	W-2	
	다. 친교적 글쓰기	사과문 쓰기	W-3	
	라. 디지털 매체 쓰기	온라인 게시판 요청문(복합 매체) 쓰기	W-4	

2) 문항별 수정 사항

W-1의 경우 가교 문항으로서의 특성을 고려하여 예비 조사 문항에 대한 수정은 하지 않았으나, 다른 문항에서 제시하고 있는 채점 항목들과의 일치도를 제고하는 방향으로 채점 항목의 세부적인 정교화 작업을 수행하였다. 이는 가교 문항이 지닌 측정의 연속성과 안정성을 보장하면서도 전체 평가 체계 내에서의 일관성을 확보하기 위한 것이다.

영역	쓰기	예비 조사	1번 (공통)	본조사	W-1
수정 내용	수정 사항				
	전		후		
채점 항목 및 예시 답안 수정	[채점 항목 - 조직] - 대상에 대한 독자의 이해를 도울 수 있는 내용 전개 방식을 사용하였다. [예시 답안] 얼굴은 전체적으로 계란형이며 초승달 같은 눈썹에 작은 입을 가진 미인의 모습이다. 얼굴은 고르게 흰색을 띠고 이마와 양 볼에 연지, 곤지를 찍었다. 두발은 얼굴 상반부를 테두리같이 둘러 귀를 덮고 양 볼 끝까지 내려 드리웠으며 뽕같이 양쪽 머리 위에 쪽을 졌다. 웃는 듯한 실눈과 검은 눈썹, 반쯤 벌린 입을 하고 있고 코 양 옆으로 팔자 주름이 표현되어 있다. 결발, 눈썹, 눈, 양 볼 할 것 없이 좌우 대칭을 이		[채점 항목 - 조직] - 대상에 대한 독자의 이해를 도울 수 있는 내용 전개 방식(묘사, 비교, 대조, 분석 등)을 사용하였다. [예시 답안] 얼굴은 전체적으로 계란형이며 초승달 같은 눈썹에 작은 입을 가진 여성의 모습이다. 얼굴은 고르게 흰색을 띠고 이마와 양 볼에 연지, 곤지를 찍었다. 두발은 얼굴 상반부를 테두리같이 둘러 귀를 덮고 양 볼 끝까지 내려 드리웠으며 양쪽 머리 위에 뽕같이 쪽을 졌다. 웃는 듯한 실눈과 검은 눈썹, 반쯤 벌린 입을 하고 있고 코 양 옆으로 주름이 잡혀 있어 광대가 도드라져		

	루고 있으며 중앙의 우뚝 솟은 날씬한 코가 조화를 이루고 있다.	보인다. 눈썹, 눈, 양 볼 할 것 없이 좌우 대칭을 이루고 있으며 중앙의 우뚝 솟은 날씬한 코가 다른 부분들과 조화를 이루고 있다.
--	-------------------------------------	--

W-1의 경우 채점 기준(조직 항목)과 예시 답안의 수정이 있었다. 채점 기준의 경우 ‘독자의 이해를 도울 수 있는 내용 전개 방식’에 대한 예로 ‘묘사, 비교, 대조, 분석’을 상세 기술하였다. 예시 답안은 일부 문법 오류(띄고 → 띄고, 양 옆 → 양옆)를 수정하였으며, 탈의 모양을 상세히 설명하는 내용을 추가하여 수정하였다.

W-2의 수정 사항은 다음과 같다.

영역	쓰기	예비 조사	2번 (가형)	본조사	W-2
수정 내용	수정 사항				
	전		후		
문두 수정					
채점 항목 및 예시 답안 수정	[채점 항목 - 표현] - 사례 깊고 효과적인 표현을 사용하였다 [예시 답안] <찬성> 요즘 아동의 출입을 제한하는 식당과 카페가 늘어나고 있습니다. 이에 대한 반대 의견도 많지만 저는 이러한 운영 방침에 대해 찬성합니다. 어린이 제한 구역 방침은 업주들 입장에서 아이들의 부주의한 행동으로 발생할 수 있는 피해와 보상 책임을 예방하기 위한 것입니다. 실제로 아이가 음식점에서 뛰어다니다 종업원과 부딪혀 다친 사건에서 종업원 책임이 있다는 판결도 있었습니다. 업주 입장에서 이렇게 자신이 입을 경제적 손실을 사전에 방지하는 것은 당연한 일이며, 자신의 가게를 어떻게 운영할지는 업주 본인의 자유에 해당합니다. 또한 아이의 잘못된 행동을 제지하지 않는 부모들이 있다는 점을 고려할 때, 이러한 운영 방침은 다른 고객들이 조용하고 편안한 환경에서 서비스를 받을 권리를 보장해 줍니다.		[채점 항목 - 표현] - 독자를 고려한 효과적인 표현을 사용하였다. [예시 답안] <찬성> 요즘 아동의 출입을 제한하는 식당과 카페가 늘어나고 있습니다. 이에 대한 반대 의견도 많지만 저는 이러한 운영 방침에 대해 찬성합니다. 어린이 제한 구역 방침은 업주들 입장에서 아이들의 부주의한 행동으로 발생할 수 있는 피해와 보상 책임을 예방하기 위한 것입니다. 실제로 아이가 음식점에서 뛰어다니다 종업원과 부딪혀 다친 사건에서 종업원에게 책임이 있다는 판결도 있었습니다. 업주 입장에서 이렇게 자신이 입을 경제적 손실을 사전에 방지하는 것은 당연한 일이며, 어린이의 권리 보장도 물론 중요하지만, 자신의 가게를 어떻게 운영할지를 결정할 수 있는 업주의 권리 또한 중요합니다. 또한 아이의 잘못된 행동을 제지하지 않는 부모들이 있다는 점을 고려할 때,		

	어린이 제한 구역 운영 방침은 단순히 어린이를 배제하는 것이 아니라 업주의 운영의 자유와 개인의 휴식권을 보장하는 수단입니다. 따라서 저는 이에 찬성합니다.	이러한 운영 방침은 다른 고객들이 조용하고 편안한 환경에서 서비스를 받을 권리를 보장해 줍니다. 어린이 제한 구역 운영 방침은 단순히 어린이를 배제하는 것이 아니라 업주의 운영의 자유와 개인의 휴식권을 보장하는 수단입니다. 따라서 저는 어린이 제한 구역 운영 방침에 찬성합니다.
--	--	---

W-2의 경우 예비 조사 문항에서는 ‘자신의 생각을 글로 써 보세요’라는 포괄적인 발문을 사용하였으나, 이를 개선하여 ‘주장하는 글’이라는 텍스트 유형을 명시적으로 제시함으로써 학습자가 작성해야 할 글의 성격과 방향을 분명히 하였다. 이러한 수정은 전체 조사 도구 내에서 문항 간 발문의 일관성을 확보하려는 목적에서 이루어진 것이다.

채점 기준(조직 항목)은 ‘설득력을 높일 수 있는 내용 전개 방식’에 대한 예로 ‘문제-해결, 인과, 유추, 분석’을 상세 기술하였다. 채점 기준(표현 항목)의 경우 예비 조사에서는 ‘사려 깊고 효과적인 표현’이었던 표현 항목의 채점 준거를, ‘논증적 글쓰기’ 평가의 일반적인 구인을 고려하여 ‘독자를 고려한 효과적인 표현’으로 명확하게 제시하였다. 예시 답안의 경우 일부 문장을 윤문하였으며(종업원 책임 → 종업원에게 책임이), 반론을 고려하는 문장을 포함하여(‘어린이의 권리 보장도 물론 중요하지만’ 등) 예시 답안을 수정하였다. 또한 결론에서 자신의 주장을 다시 한번 강조하는 내용인 ‘저는 이에 찬성합니다’를 ‘저는 어린이 제한 구역 운영 방침에 찬성합니다’와 같이 명시적으로 서술하는 방향으로 수정하였다.

W-3의 수정 사항은 다음과 같다.

영역	쓰기	예비 조사	3번 (가형)	본조사	W-3
수정 내용	수정 사항				
	전		후		
문두 및 본문 수정					

채점 항목 및 예시 답안 수정	[예시 답안] (안녕하세요. 테니스 사랑 동호회 총무 000입니다. 다음 달 28일(토)에 열릴 친선 경기 날짜를 21일(토)로 잘못 공지하였습니다.) 이번 테니스 동호회 간 친선 경기 날짜를 잘못 공지하여 불편을 끼쳐드린 점에 대해 머리 숙여 사과의 말씀을 드립니다. 총무로서 각종 행사 일정을 꼼꼼하게 챙겼어야 하는데, 저의 착오로 이런 일이 발생한 점에 대해 죄송스럽고 또 부끄럽습니다. 다시금 이런 일이 반복되지 않도록 각종 행사 일정을 꼼꼼하게 챙기며 테니스 사랑 동호회 총무로서 책임을 다하겠습니다. (다시 한번 진심으로 사과드립니다.)	[예시 답안] (안녕하세요. ‘테니스 사랑’ 동호회 총무 000입니다. 다음 달 28일(토)에 열릴 친선 경기 날짜를 제가 21일(토)로 잘못 공지하였습니다.) 친선 경기 날짜를 잘못 공지하여 불편을 끼쳐드린 점에 대해 머리 숙여 사과의 말씀을 드립니다. 총무로서 각종 행사 일정을 꼼꼼하게 챙기지 못한 저의 책임입니다. 다시금 이런 일이 반복되지 않도록 각종 행사 일정을 꼼꼼하게 챙기며 테니스 사랑 동호회 총무로서 책임을 다하겠습니다.(다시 한번 진심으로 사과드립니다.)

발문의 경우 채점 기준과의 정합성을 확보하고 다른 문항과의 일관성을 유지하기 위해 분량 조건을 ‘3문장 이상’으로 구체적으로 명시하였다. 이러한 분량 기준의 명시화는 학습자에게 명확한 작성 지침을 제공하면서 채점자가 객관적이고 일관된 기준을 적용하기 위함이다. 아울러 본문의 ‘친선 경기 날짜를 21일(토)로 잘못 공지’를 ‘친선 경기 날짜를 제가 21일(토)로 잘못 공지’로 수정하여 행동 주체인 ‘제가’를 명시하였다. W-3의 예시 답안은 사과하기라는 글의 목적과 잘못의 주체를 명확히 드러낼 수 있도록 ‘제가’, ‘저의 책임’ 등의 표현을 명시하는 방향으로 수정되었다.

마지막으로 W-4 문항의 수정 사항은 다음과 같다.

영역	쓰기	예비 조사	4번 (가형)	본조사	W-4
수정 내용	수정 사항				
	전		후		
조건 및 사진 수정					

채점 항목 및 예시 답안 수정	[예시 답안] ① 상품에 문제가 있다는 것을 직접 보여 주기 위해서입니다. / 곰팡이 핀 굴이 배송되었다는 것을 입증하기 위해서입니다. ② (2번 사진 선택 시 2점)	[예시 답안] (4-1), (4-2) ① ‘안 신전하면 전액 환불’이라고 쓰여 있는 사진이기 때문에 전액 환불 요청의 타당성을 강조하기 위해 선택하였다. / 상품에 문제가 있으니 환불해 주어야 한다는 것을 전달하고 싶다. ② 굴에 핀 곰팡이가 잘 보여서 선택하였다. / 상품에 문제가 있으니 환불해 주어야 한다는 것을 전달하고 싶다.

		(4-3) (안녕하세요.) 오늘 배송받은 굴에 곰팡이가 피어 있습니다. 본 제품은 신선하지 않은 경우 전액 환불을 보장하고 있으며, 곰팡이로 인해 굴을 먹을 수 없는 상황입니다. 따라서 빠른 시일 내에 반품 처리를 해 주시기를 요청드립니다.
--	--	--

먼저, W-4-1, W-4-2의 발문에서 다른 문항과의 일관성을 고려하여 ‘1문장으로 작성’을 ‘1문장으로 작성할 것’으로 수정하였다. 다음으로 사진의 경우, 예비 조사 단계에서는 적절한 사진과 적절하지 않은 사진을 함께 제시하고 학습자가 적절한 사진을 선택하도록 문항을 구성하였다. 그런데 사진 선택 능력과 ‘쓰기’ 능력 간 간섭으로 인해 변별도가 떨어지는 문제가 발생하여, 이를 해결하기 위해 문항을 전면적으로 재설계하였다. 개선된 문항은 주어진 맥락에 적절한 두 사진을 제시하되, 해당 사진을 선택한 이유와 해당 사진을 통해 전달하고자 하는 내용을 구체적으로 진술하는 방식으로 설계하였다.

또한 디지털 문항의 예시 답안을 실제 응답을 근거로 상세히 수정하였다. 예비 조사 단계에서는 정답이 될 수 있는 ①번 사진과 오답인 ② 사진을 제시하고, ② 사진을 고른 경우 오답으로 간주하여 최하점인 2점으로 평정하였다. 그러나 본조사에서는 둘 중 어느 것을 선택하더라도 정답이 될 수 있는 두 장의 사진을 제시하되, 각 사진을 골라 활용하는 이유를 논리적으로 타당하게 설명할 수 있는지를 평가할 수 있도록 설계하였다. 따라서 ‘상품에 문제가 있음을 강조하기 위해’ 등 타당한 이유를 들어 환불 요청을 뒷받침하고 있는지를 평가할 수 있도록 예시 답안을 수정하였다.

3) 영역 배경 변인 구성

2025년 국민의 국어능력 본조사에서는 배경 변인과 국민의 국어능력 사이의 관계를 심층적으로 살펴보기 위해 배경 변인을 구성하였다. 특히 쓰기 영역의 경우 지난 주기 조사와 비교할 때 생성형 인공지능 활용 글쓰기가 도입되어 점차 대중화되고 있는 쓰기 환경의 변화 양상을 설문에 반영하고자 하였다. 이에 따라 쓰기에 대한 인식, 쓰기 장르별 쓰기 빈도와 쓰기 환경 등을 설문에 포함하는 한편, 글쓰기 시 챗지피티(ChatGPT) 등 생성형 인공지능 이용에 관한 설문을 포함하여 문항을 구성하였다.

(1) 인구통계학적 변인

● 해당되는 항목에 V나 ○로 표시하여 주십시오.

변인	발문 및 내용			
1. 시·도	① 서울	② 부산	③ 대구	④ 광주
	⑤ 인천	⑥ 대전	⑦ 울산	⑧ 세종
	⑨ 경기	⑩ 강원	⑪ 충북	⑫ 충남
	⑬ 전북	⑭ 전남	⑮ 경북	⑯ 경남
	⑰ 제주			
2. 읍면동 코드			3. 지역 구분	① 대도시 ② 중소 도시 ③ 읍면 지역

4. 성별	① 남자 ② 여자	5. 나이	● 귀하의 현재 나이는 어떻게 되십니까? 만 _____ 세
(연령대) =>응답자에게 노출되지 않음	① 20-29세 ② 30-39세 ③ 40-49세 ④ 50-59세 ⑤ 60~69세		
6. 직업	① 관리자 ② 전문가 및 관련 종사자 ③ 사무 종사자 ④ 서비스 종사자 ⑤ 판매 종사자 ⑥ 농림어업 종사자 ⑦ 기능원 및 관련 기능 종사자 ⑧ 장치, 기계조작 및 조립 종사자 ⑨ 단순 노무 종사자 ⑩ 직업 군인 ⑪ 주부 ⑫ 무직/퇴직/기타 ⑬ 학생		
7. 최종 학력	① 무학 ② 초등학교(국민학교) 졸업 ③ 중학교 중퇴 ④ 중학교 졸업 ⑤ 고등학교 중퇴 ⑥ 고등학교 재학 ⑦ 고등학교 졸업 ⑧ 대학교 재학 ⑨ 대학교 졸업 ⑩ 대학원 재학 ⑪ 대학원 졸업		

(2) 쓰기 영역 심층 변인

(1) 귀하는 글쓰기를 어떻게 생각하십니까?					
	매우 그렇지 않다	그렇지 않다	보통이다	그렇다	매우 그렇다
나 자신의 생각이나 감정을 이해하고 표현하기 위해 필요하다	①	②	③	④	⑤
일상이나 학업, 업무를 원활히 수행하기 위해 필요하다	①	②	③	④	⑤
다른 사람과 좋은 관계를 맺기 위해 필요하다	①	②	③	④	⑤
(2) 귀하는 얼마나 자주 비자발적 글쓰기를 하십니까?					
* 비자발적 글쓰기: 공문서 쓰기, 업무를 위한 서류나 보고서 쓰기, 과제 글쓰기 등과 같은 공식적인 글쓰기					
전혀 쓰지 않음	1년에 2~3회	1개월에 2~3회	1주일에 2~3회	거의 매일	
①	②	③	④	⑤	

(3) 귀하는 얼마나 자주 자발적 글쓰기를 하십니까?

* 자발적 글쓰기: 일상 기록, 일기 쓰기, 편지 쓰기, 블로그 쓰기 등과 같은 일상적인 글쓰기

전혀 쓰지 않음	1년에 2~3회	1개월에 2~3회	1주일에 2~3회	거의 매일
①	②	③	④	⑤

(4) 귀하는 글쓰기를 통해 긍정적 감정(기쁨, 성취감, 보람, 유용함 등)을 느끼는 경험이 얼마나 자주 있으십니까?

전혀 없음	1년에 2~3회	1개월에 2~3회	1주일에 2~3회	거의 매일
①	②	③	④	⑤

(4-1) 글쓰기를 통해 긍정적인 감정을 느끼는 가장 주된 이유는 무엇입니까?

표현 자체의 재미와 즐거움	다른 사람들과의 소통과 관계 형성	실제 삶의 문제 해결(타인 설득이나 상황 개선, 공적·사적 업무 처리 등)	성찰이나 새로운 깨달음 획득	타인의 인정이나 칭찬
①	②	③	④	⑤

(5) 귀하가 자신의 생각만큼 충분히 글을 쓰지 못하고 있다면, 그 이유는 무엇입니까?

글쓰기에 흥미가 없어서	글을 어떻게 써야 할지 몰라서	인터넷 검색, 동영상 시청 등을 더 좋아해서	시간이 없어서	다른 사람들이 내 글을 읽는 게 부담스러워서
①	②	③	④	⑤

(6) 귀하가 쓰기 과정에서 생성형 인공지능을 얼마나 자주 활용하십니까?

* 생성형 인공지능: 챗지피티(ChatGPT), 뮌 등

전혀 활용하지 않음	한두 번 활용해 봄	가끔 활용함	자주 활용함	거의 매번 활용함
①	②	③	④	⑤

(7) 귀하는 귀하의 가정에서 귀하의 글에 대해 함께 이야기를 나눌 사람이 있거나, 글쓰기 도움을 받을 수 있습니까?

* 1인 가구 등 동거인이 없는 경우 응답하지 않음

매우 그렇지 않다	그렇지 않다	보통이다	그렇다	매우 그렇다
①	②	③	④	⑤

(8) 귀하는 귀하가 속한 사회(지역, 직장, 학교)에서 귀하의 글에 대해 함께 이야기를 나눌 사람이 있거나, 글을 쓸 때 도움을 받을 수 있는 사람 또는 프로그램이 있습니까?

매우 그렇지 않다	그렇지 않다	보통이다	그렇다	매우 그렇다
①	②	③	④	⑤

2.2. 표본 집단 추출

본 연구의 조사 시행 방식은 ‘1:1 방문 면접 조사’와 ‘집합 조사’ 방식에 따른 결과치 비교를 위해 ‘ 개별 가구 방문 평가’와 ‘소규모 내방형 평가(집합 조사)’를 진행하였다.

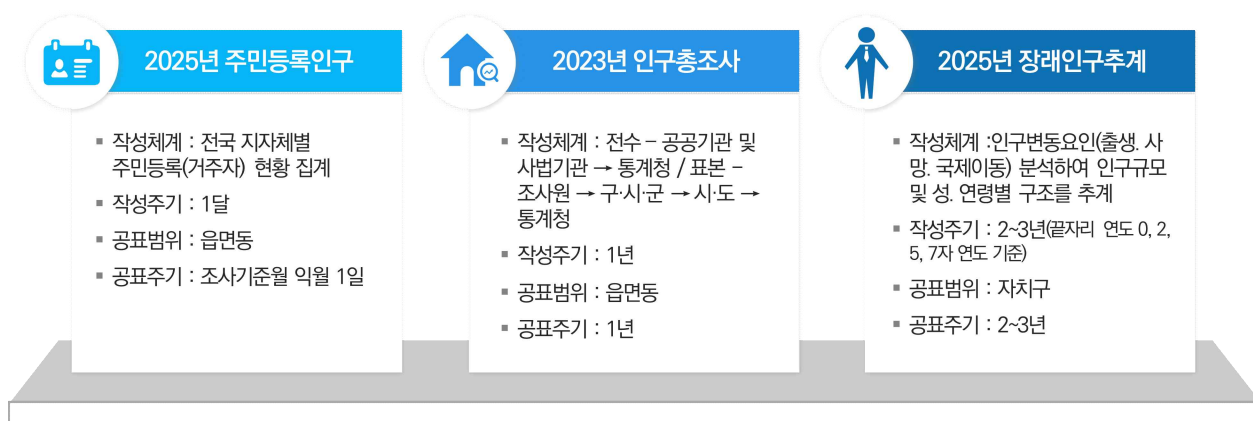
<표 II-5> 방문 면접 조사 및 집합 조사 개요

구분	방문 면접 조사	집합 조사
모집단	만 나이 20세 이상 69세 이하 일반 국민	
조사 대상	전국 만 나이 20세 이상 69세 이하 일반 국민	수도권 거주 만 나이 20세 이상 69세 이하 일반 국민
조사 기간	2025년 6월 19일 ~ 8월 10일	2025년 8월 7일
조사 지역	전국 17개 시도	수도권(서울, 인천, 경기)
표본 크기	3,000명	50명
표본 배분	모집단의 인구특성(성, 연령, 지역, 직업, 학력)에 비례하도록 표본 할당	성, 연령대별 동일 표본 할당
조사 방법	태블릿 피서를 이용한 대인면접조사(TAPI: Tablet PC Aided Personal Interview)	

1) 방문 면접 조사

(1) 모집단 검토

정확한 표본 설계와 모수추정을 위해서는 모집단인 20대에서 60대까지 국민의 실제 거주지별 정확한 인구통계가 필요하다.



[그림 II-1] 모집단 자료 현황

현재 우리나라에서는 ① 주민등록인구통계, ② 인구총조사, ③ 장래인구추계 총 세 가지의 인구통계

정보가 대표적인 모집단 자료로 활용된다. 2023년 인구총조사는 인구 규모, 분포 및 구조, 가구, 주택에 관한 특성을 반영한 통계자료지만, 기준시점이 2023년으로 조사 시점과 현재 시점의 차이가 있어 실제 정보가 다를 가능성이 존재한다. 2025년 장래인구추계는 2020년 인구총조사 결과를 기반으로 인구 변동 요인(출산력, 사망력, 인구이동 등)을 고려하여 작성된 통계자료로, 전국, 시도별 단위로 공표하고 있어 읍면동 단위 층화가 필요한 본조사의 모집단으로는 적절하지 않다. 반면, 주민등록인구통계는 “주민등록법”에 의한 주민등록 인구 및 세대현황에 대하여 전국 단위 시도, 시군구, 행정동 및 법정동(리)별 통계를 매달 말일 기준으로 공표하여 시의성과 활용성 면에서 우수하다. 따라서, 본조사에서는 모집단으로 행정안전부의 주민등록인구통계를 활용하며, 거주자를 기준으로 한 통계자료를 활용하였다. 이에 따라 본 조사의 목표 모집단은 전국 만 20~69세 성인남녀이며, 조사 모집단은 조사 시점 최신 행정안전부의 주민등록인구통계 자료 기준 대한민국에 거주하는 만 20~69세 성인남녀로 정의한다.

(2) 모집단 층화

본조사의 표본은 집락으로 층화 후, 집락을 먼저 추출하고, 집락에서 조사단위인 개인은 할당 추출하는 방식으로 설계한다. 층화추출은 모집단을 서로 겹치지 않는 그룹들, 즉 층들(strata)로 나누고 각 층에서 표본을 뽑는 방법이다.

모집단 층화를 위해서는 먼저 전국을 17개 시도로 1차 층화한 후, 17개 시도별로 다시 그 지역의 특성을 고려하여 2차 층화를 실시한다. 2차 층화는 읍면지역이 없는 서울, 광주, 대전을 제외한 나머지 14개 시도는 동부와 읍면부로 구분하여 총 31개 층으로 층화하였다.

1차 층화

17개 시도단위 층화

2차 층화

동부/읍면부로 층화
(읍면 지역이 없는 서울, 광주, 대전 제외)

1차 층화	2차 층화		총의 개수
광역시도	서울, 광주, 대전	-	3
	부산, 대구, 인천, 울산, 세종, 경기, 강원, 충북, 충남, 전북, 전남, 경북, 경남, 제주	동부	28
		읍면부	
총			31

[그림 II-2] 모집단 층화 기준

(3) 표본 크기 결정

표본 크기를 결정하는데 고려해야 할 요소는 생산할 주요 통계의 허용 목표 오차의 크기, 사용할 수 있는 예산의 규모와 통계 생산 단위의 수준 등이다. 본조사에서의 주요 통계 생산 단위와 허용 목표 오차, 예산 규모 등을 종합적으로 고려하여 전체 표본규모는 3,000명으로 결정하였다. 전국적으로 3,000명을 조사할 경우 95% 신뢰수준에서 표본 오차는 ±1.8%이다.

<표 II -6> 표본 크기 결정

구분	내용
표본 크기	3,000명
표본 오차	비율 추정 시 95% 신뢰수준 $\pm 1.8\%$

(4) 집락 크기 결정

조사의 모집단 조사단위(unit)는 개인이지만 각 개인의 목록을 모두 나열하고 있는 추출틀을 확보하기 어려운 상황이므로 조사단위들이 모여 있는 집락(cluster)의 추출틀을 확보하고, 이 집락에서 개인을 추출하는 방식을 채택한다. 이 경우 집락추출(cluster sample)이 최소의 비용으로 더 많은 정보를 얻기 위한 효과적인 방법이 될 수 있다(Richard et al., 2012: 266).

따라서 본조사에서는 일정한 구역으로 나눌 수 있으며 관련 정보를 비교적 획득하기 쉬운 집락인 읍면동을 먼저 추출하고 여기에서 조사단위 표본을 얻는 집락추출을 진행하였다.

<표 II -7> 17개 시도별 읍면동 현황(2025년 3월 주민등록인구통계 기준)

지역	읍면동 수	지역	읍면동 수
서울	426	강원	194
부산	206	충북	153
대구	152	충남	210
인천	160	전북	243
광주	96	전남	323
대전	82	경북	335
울산	55	경남	310
세종	24	제주	43
경기	603	전국	3,615

*전국 출장소 포함(단, 거주자가 0명인 출장소 제외)

집락의 크기와 집락에서 추출되는 표본의 크기는 상호 관련성이 있는데, 일반적으로 사회조사에서는 목표 표본 수를 먼저 결정하고 조사를 진행하게 되는데, 만일 집락의 수를 늘린다면 집락에서 추출되는 표본 수는 줄어들게 되고, 반대로 집락의 수를 줄인다면 집락별 표본 수는 늘어나게 된다.

본조사에서는 하나의 집락에서 10명의 표본을 조사하는 것을 원칙으로 하여, 다음과 같이 300개의 집락(읍면동)을 추출한 후 3,000명을 조사하였다.

<표 II -8> 본조사 추출 집락 및 집락별 조사 완료 표본 수

총 구분	집락(읍면동)	
서울	서울특별시 종로구 송인제2동	서울특별시 양천구 신정7동
	서울특별시 용산구 청파동	서울특별시 강서구 화곡제1동
	서울특별시 성동구 왕십리제2동	서울특별시 강서구 가양제1동
	서울특별시 성동구 금호4가동	서울특별시 강서구 방화제1동

층 구분	집락(읍면동)	
	서울특별시 광진구 중곡제1동 서울특별시 광진구 자양제3동 서울특별시 동대문구 제기동 서울특별시 동대문구 장안제2동 서울특별시 중랑구 면목제4동 서울특별시 중랑구 중화제2동 서울특별시 중랑구 신내2동 서울특별시 성북구 정릉제2동 서울특별시 성북구 월곡제2동 서울특별시 강북구 송천동 서울특별시 강북구 우이동 서울특별시 도봉구 도봉제1동 서울특별시 노원구 월계1동 서울특별시 노원구 하계1동 서울특별시 노원구 상계2동 서울특별시 은평구 녹번동 서울특별시 은평구 대조동 서울특별시 은평구 신사제2동 서울특별시 서대문구 신촌동 서울특별시 서대문구 남가좌제2동 서울특별시 마포구 용강동 서울특별시 마포구 망원제2동 서울특별시 양천구 목3동 서울특별시 양천구 신월6동	서울특별시 구로구 구로제3동 서울특별시 구로구 개봉제2동 서울특별시 금천구 독산제1동 서울특별시 금천구 시흥제4동 서울특별시 영등포구 당산제2동 서울특별시 영등포구 신길제6동 서울특별시 동작구 상도제1동 서울특별시 동작구 사당제3동 서울특별시 관악구 성현동 서울특별시 관악구 중앙동 서울특별시 관악구 신림동 서울특별시 서초구 서초2동 서울특별시 서초구 반포3동 서울특별시 서초구 양재2동 서울특별시 강남구 삼성2동 서울특별시 강남구 도곡1동 서울특별시 강남구 세곡동 서울특별시 송파구 방이1동 서울특별시 송파구 석촌동 서울특별시 송파구 문정2동 서울특별시 송파구 잠실4동 서울특별시 강동구 고덕제1동 서울특별시 강동구 천호제3동 서울특별시 강동구 둔촌제1동
부산_동부	부산광역시 중구 보수동 부산광역시 동구 좌천동 부산광역시 부산진구 연제동 부산광역시 부산진구 당감제1동 부산광역시 동래구 수민동 부산광역시 동래구 사직제2동 부산광역시 남구 대연제6동 부산광역시 북구 구포제1동 부산광역시 북구 덕천제1동	부산광역시 해운대구 중제1동 부산광역시 해운대구 반여제2동 부산광역시 사하구 괴정제4동 부산광역시 사하구 다대제1동 부산광역시 금정구 장전제2동 부산광역시 강서구 명지1동 부산광역시 연제구 연산제2동 부산광역시 수영구 망미제1동 부산광역시 사상구 덕포제2동
부산_읍면부	부산광역시 기장군 정관읍	
대구_동부	대구광역시 동구 신암2동 대구광역시 동구 해안동 대구광역시 서구 내당4동 대구광역시 남구 봉덕3동 대구광역시 북구 산격2동 대구광역시 북구 태전1동	대구광역시 수성구 범어1동 대구광역시 수성구 황금2동 대구광역시 수성구 고산2동 대구광역시 달서구 용산1동 대구광역시 달서구 월성2동 대구광역시 달서구 도원동
대구_읍면부	대구광역시 달성군 화원읍	대구광역시 달성군 다사읍서재출장소
인천_동부	인천광역시 중구 영종1동 인천광역시 미추홀구 송의1.3동 인천광역시 미추홀구 도화1동 인천광역시 미추홀구 주안8동 인천광역시 연수구 동춘1동 인천광역시 연수구 송도4동 인천광역시 남동구 구월3동 인천광역시 남동구 만수4동 인천광역시 남동구 논현2동	인천광역시 부평구 부평5동 인천광역시 부평구 갈산1동 인천광역시 부평구 십정1동 인천광역시 계양구 작전1동 인천광역시 서구 검암경서동 인천광역시 서구 청라3동 인천광역시 서구 가좌2동 인천광역시 서구 오류왕길동
인천_읍면부	인천광역시 옹진군 영흥면	

층 구분	집락(읍면동)	
광주	광주광역시 서구 농성1동 광주광역시 북구 용봉동 광주광역시 북구 두암1동 광주광역시 서구 양3동	광주광역시 서구 화정4동 광주광역시 남구 월산동 광주광역시 북구 신안동 광주광역시 북구 양산동
대전_동부	대전광역시 동구 가양2동 대전광역시 중구 오류동 대전광역시 서구 정림동 대전광역시 서구 월평2동	대전광역시 서구 둔산2동 대전광역시 유성구 온천2동 대전광역시 유성구 관평동 대전광역시 대덕구 덕암동
울산_동부	울산광역시 중구 우정동 울산광역시 남구 신정2동 울산광역시 남구 삼호동	울산광역시 동구 방어동 울산광역시 동구 남목3동 울산광역시 북구 효문동
울산_읍면부	울산광역시 울주군 범서읍	
세종_동부	세종특별자치시 아름동	
세종_읍면부	세종특별자치시 금남면	
경기_동부	경기도 수원시 장안구 정자2동 경기도 수원시 권선구 세류2동 경기도 수원시 권선구 금곡동 경기도 수원시 권선구 곡선동 경기도 수원시 팔달구 우만2동 경기도 수원시 영통구 원천동 경기도 수원시 영통구 망포2동 경기도 성남시 수정구 태평1동 경기도 성남시 수정구 고등동 경기도 성남시 중원구 상대원2동 경기도 성남시 분당구 정자2동 경기도 성남시 분당구 야탑3동 경기도 성남시 분당구 운중동 경기도 의정부시 장암동 경기도 의정부시 송산3동 경기도 안양시 만안구 안양1동 경기도 안양시 만안구 석수2동 경기도 안양시 동안구 관양동 경기도 안양시 동안구 호계3동 경기도 부천시 원미구 역곡2동 경기도 부천시 원미구 중3동 경기도 부천시 소사구 소사본동 경기도 부천시 소사구 송내2동 경기도 부천시 오정구 신흥동 경기도 광명시 하안1동 경기도 평택시 중앙동 경기도 평택시 비전1동 경기도 평택시 동삭동 경기도 안산시 상록구 일동 경기도 안산시 상록구 본오1동 경기도 안산시 상록구 안산동 경기도 안산시 단원구 신길동 경기도 고양시 덕양구 원신동 경기도 고양시 덕양구 창릉동 경기도 고양시 덕양구 행주동	경기도 고양시 일산동구 식사동 경기도 고양시 일산동구 백석2동 경기도 고양시 일산서구 일산3동 경기도 고양시 일산서구 송포동 경기도 구리시 갈매동 경기도 남양주시 호평동 경기도 남양주시 다산1동 경기도 오산시 중앙동 경기도 오산시 신장2동 경기도 시흥시 은행동 경기도 시흥시 정왕1동 경기도 시흥시 연성동 경기도 군포시 군포2동 경기도 군포시 광정동 경기도 의왕시 청계동 경기도 하남시 갑일동 경기도 하남시 미사3동 경기도 용인시 기흥구 신갈동 경기도 용인시 기흥구 기흥동 경기도 용인시 기흥구 동백3동 경기도 용인시 수지구 신봉동 경기도 용인시 수지구 상현2동 경기도 파주시 금촌3동 경기도 파주시 운정3동 경기도 이천시 증포동 경기도 김포시 김포본동 경기도 김포시 장기동 경기도 화성시 진안동 경기도 화성시 화산동 경기도 화성시 동탄4동 경기도 화성시 동탄7동 경기도 광주시 탄벌동 경기도 광주시 신현동 경기도 양주시 회천3동 경기도 여주시 중앙동
경기_읍면부	경기도 남양주시 와부읍 경기도 남양주시 화도읍	경기도 화성시 봉담읍 경기도 화성시 향남읍

층 구분	집락(읍면동)	
	경기도 남양주시 별내면 경기도 파주시 문산읍 경기도 이천시 부발읍 경기도 안성시 대덕면	경기도 화성시 정남면 경기도 포천시 소흘읍 경기도 연천군 왕징면 경기도 양평군 용문면
강원_동부	강원특별자치도 춘천시 후평3동 강원특별자치도 춘천시 신사우동 강원특별자치도 원주시 태장1동	강원특별자치도 강릉시 중앙동 강원특별자치도 동해시 천곡동 강원특별자치도 속초시 노학동
강원_읍면부	강원특별자치도 원주시 흥업면 강원특별자치도 영월군 영월읍	강원특별자치도 인제군 북면
충북_동부	충청북도 청주시 상당구 용암2동 충청북도 청주시 서원구 성화.개신.죽림동 충청북도 청주시 흥덕구 봉명2.송정동	충청북도 충주시 교현.안림동 충청북도 제천시 용두동
충북_읍면부	충청북도 청주시 흥덕구 강내면 충청북도 충주시 엄정면	충청북도 증평군 증평읍 충청북도 음성군 맹동면
충남_동부	충청남도 천안시 동남구 청룡동 충청남도 천안시 서북구 쌍용3동 충청남도 천안시 서북구 부성1동	충청남도 보령시 대천3동 충청남도 서산시 동문1동 충청남도 당진시 당진2동
충남_읍면부	충청남도 천안시 서북구 성거읍 충청남도 아산시 배방읍 충청남도 서산시 인지면	충청남도 당진시 송악읍 충청남도 서천군 마서면 충청남도 예산군 삽교읍
전북_동부	전북특별자치도 전주시 완산구 평화2동 전북특별자치도 전주시 완산구 효자5동 전북특별자치도 전주시 덕진구 송천1동 전북특별자치도 군산시 초촌동	전북특별자치도 군산시 소룡동 전북특별자치도 익산시 어양동 전북특별자치도 전주시 덕진구 호성동
전북_읍면부	전북특별자치도 김제시 용지면 전북특별자치도 부안군 변산면	전북특별자치도 김제시 백산면
전남_동부	전라남도 목포시 상동 전라남도 여수시 삼일동 전라남도 광양시 중마동	전라남도 목포시 이로동 전라남도 광양시 골약동
전남_읍면부	전라남도 광양시 광양읍 전라남도 고흥군 두원면 전라남도 영광군 영광읍	전라남도 여수시 돌산읍 전라남도 고흥군 대서면
경북_동부	경상북도 포항시 남구 상대동 경상북도 포항시 북구 용흥동 경상북도 경주시 선도동 경상북도 김천시 지좌동	경상북도 구미시 원평동 경상북도 구미시 임오동 경상북도 영주시 휴천2동 경상북도 문경시 점촌5동
경북_읍면부	경상북도 포항시 북구 흥해읍 경상북도 안동시 남후면 경상북도 상주시 모동면	경상북도 칠곡군 왜관읍 경상북도 울진군 죽변면 경상북도 포항시 남구 호미곶면
경남_동부	경상남도 창원시 의창구 팔룡동 경상남도 창원시 성산구 중앙동 경상남도 창원시 성산구 웅남동 경상남도 창원시 마산회원구 석전동 경상남도 창원시 진해구 자은동 경상남도 진주시 상대동	경상남도 진주시 가호동 경상남도 김해시 내외동 경상남도 김해시 활천동 경상남도 김해시 장유3동 경상남도 거제시 옥포2동 경상남도 양산시 삼성동
경남_읍면부	경상남도 진주시 내동면 경상남도 김해시 진영읍 경상남도 양산시 물금읍	경상남도 양산시 하북면 경상남도 남해군 남해읍 경상남도 합천군 합천읍
제주_동부	제주특별자치도 제주시 화북동 제주특별자치도 제주시 노형동	제주특별자치도 서귀포시 대천동
제주_읍면부	제주특별자치도 제주시 한림읍	

(5) 크기비례확률추출법을 통한 집락 추출

집락 추출 방식은 크게 단순확률추출과 크기비례확률추출(Sampling with probabilities proportional to size, 이하 PPS)로 구분할 수 있다. 단순확률추출에서는 각 집락이 추출될 확률이 모두 같다. 그러나 집락들의 규모가 심하게 차이가 나거나 대규모 집락이 전체 모집단에 미치는 영향이 클 경우에는 각 집락의 크기에 비례하여 뽑는 방법인 크기비례확률추출법을 권장한다.

예를 들어, 10개의 집락으로 이루어진 모집단이 있고, 총인구수는 1,000명으로 가정할 경우, 단순확률추출 시에는 10개 집락 중 하나를 선택하므로 모든 집락의 추출 확률이 동일하다. 반면, 크기비례확률추출(PPS)을 실시할 경우에는 상대적으로 크기가 큰 집락이 추출될 확률이 그만큼 높아진다. A동부터 J동까지로 구성된 층이 있고, 이 층에서 3개의 집락(동)을 추출한다고 가정하면, 크기비례확률추출은 아래와 같은 방법으로 진행한다. 따라서 과거 조사와 동일하게 집락추출 방식은 크기비례확률추출법을 이용하였다.



[그림 II-3] 크기비례확률추출 예시

- 가장 먼저 각 집락의 인구수를 기준으로 누적 인구수를 계산함
- 추출 간격 계산: 전체 누적 인구수 / 추출 집락 수
- 본 예시에서 추출 간격은 500 ($1,500 / 3 = 500$)
- 첫 번째 추출 간격 사이의 난수를 추출한 뒤, 이 난수가 속한 집락을 1차 추출 지점으로 선정. 이후부터는 일정한 간격(추출 간격, 등간)에 따라 표본을 추출
- 첫 번째 추출 간격인 0부터 500 사이의 난수를 추출해 시작점 확정: 무작위로 120이 선정 되면, 누적 인구수에서 120이 포함된 'A동'이 최초로 추출됨
- 최초 기준 수치에 추출 등간인 매 500을 더하여 얻은 값을 포함하고 있는 집락을 조사 대상으로 결정

<표 II -9> 크기비례확률추출을 이용한 동(집락) 추출 과정(예시)

집락 동	인구수	누적 인구수	추출 등간 500 가정 시
A동	200	200	추출(시작점: 120)
B동	200	400	
C동	200	600	
D동	100	700	추출
E동	100	800	
F동	100	900	
G동	100	1,000	
H동	100	1,100	
I동	200	1,300	추출
J동	200	1,500	
계	1,500		

(6) 집락 내 개인 할당 추출

본조사에서는 조사 여건과 이전 조사 방식과의 일관성을 고려하여 할당 추출로 설계하였다. 할당추출법(Quota Sampling)이란, 표본의 주요 변수별 분포를 미리 할당한 후, 해당 조건의 표본을 찾아서 채워나가는 방식으로 비확률표본추출 방법의 하나이다. 표본의 원하는 분포를 얻어낼 수 있고, 조사 비용 및 시간을 절약할 수 있다는 장점 때문에 사회조사에서 여전히 활용되고 있다.

(7) 할당 추출을 위한 변인 설정

일반적으로 조사 대상 집단의 표본을 추출하기 위한 변인이 다양하고, 변인별 분할이 세분화될수록 조사 결과에 대해 보다 정밀한 해석이 가능하다.

그렇지만 연구의 실행 가능성을 고려하여 본 연구에서는 다음과 같이 5개 변인(성별, 연령대, 지역, 직업, 학력)을 할당 변수로 선정하고 변인별 분할도 지나치게 세분화하지 않도록 한다. 변인의 구분이 지나치게 세분화하면, 각각의 변인에 속하는 표본 집단의 수가 너무 적어져 연구실행의 타당성을 잃을 위험이 있기 때문이다.

기본 할당 변수 : 3way (성 x 연령 x 지역)		직업, 학력 등 추가 고려 변수	
모집단 분포가 알려진 변수 & 층간 이질적+층내 동질적인 변수 & 통제 가능한 변수		- 사무직, 고객 서비스 등 특정 직업군은 언어적 능력이 중요한 역할을 하며, 반면 육체적인 노동을 주로 하는 직업군에서는 언어적 능력에 상대적으로 덜 중요할 수 있음 - 또한, 일반적으로 학력이 높을수록 국어 능력이 높은 경향을 보임 - 따라서 성, 연령, 지역 외에 국어 능력에 영향을 미칠 수 있는 직업과 학력을 추가 할당 변수로 고려하고자 함 - 다만, 변인의 구분이 지나치게 세분화하면, 각각의 변인에 속하는 표본 집단의 수가 너무 적어져 조사의 타당성이 떨어질 수 있고, 해당 응답자를 찾는 데에 상당한 시간과 비용이 소모되므로, 전체 용역 기관과 예산을 고려해야 함	
할당 변수	층화 기준		
성(gender)	여자, 남자		
연령(age)	20대, 30대, 40대, 50대, 60대		
거주지역(area)	31개층		

[그림 II -4] 할당 변인

본 연구에서는 다음의 5개 변인 및 변인별 요소를 적용하였다.

<표 II -10> 표본 추출을 위한 변인 설정

변인		구분
기본 할당 변수	성별	남자, 여자
	연령	20대, 30대, 40대, 50대, 60대
	지역	17개 시도(서울, 대전, 광주를 제외한 지역은 동부, 읍면부로 구분)
추가 고려 할당 변수	직업	-정신노동(①관리자, ②전문가 및 관련 종사자, ③사무 종사자) -육체노동(④서비스 종사자, ⑤판매 종사자, ⑥농림어업 종사자, ⑦기능원 및 관련 기능 종사자, ⑧장치/기계조작 및 조립 종사자, ⑨단순 노무 종사자, ⑩직업 군인) -기타(⑪주부, ⑫무직/퇴직/기타, ⑬학생)
	학력	고졸 미만, 고졸, 대졸 이상

(8) 표본 배분

층별로 성별, 연령 구간별, 지역을 기준으로 조사 할당표를 선정하며, 직업과 학력을 추가 할당하는 방식¹⁾으로 표본 할당표를 작성하였다.

변인별 표본 수는 연구 목적을 고려해 비례 배분이나 임의 배분 중 한 가지 방법을 선택할 수 있으며, 본조사에서는 전체 국민을 대표하는 표본을 추출해야 하므로 모집단의 인구특성에 비례하도록 표본을 할당하였다.

예컨대, 전국을 17개 시도로 나누고, 서울과 광주, 대전을 제외한 지역은 동부와 읍면부로 다시 구분하여 31개 지역, 성별과 연령은 20대 남/여, 30대 남/여, 40대 남/여, 50대 남/여, 60대 남/여로 총 310개 구간으로 설정하였다.

1) 본 연구의 고려 변인은 다섯 개로, 이들을 모두 엄격하게 맞추어 진행하려면 변인별 분할이 지나치게 세분되어 해당 응답자를 찾는 데에 상당한 시간과 비용이 소요된다. 따라서 ‘성별, 지역, 연령’ 변인은 각 읍면동 표본 할당표에 정확히 맞추어 진행하되, ‘직업과 학력’ 변인은 할당 인원 대비 약간의 차이를 허용하는 방식으로 조사를 진행하였다.

<표 II-11> 2025년 3월 기준 표본 할당표

구 분 (단위: 명)	합계	20대		30대		40대		50대		60대	
		남	여	남	여	남	여	남	여	남	여
합계	3,000	254	236	284	263	318	305	355	350	312	323
서울	560	49	56	59	59	55	56	59	61	50	56
부산_동부	180	14	14	16	15	18	18	21	21	20	23
부산_읍면부	10	1	1	1	1	1	1	1	1	1	1
대구_동부	120	10	9	11	9	12	12	15	15	13	14
대구_읍면부	20	2	1	2	2	3	2	2	2	2	2
인천_동부	170	14	13	17	16	19	18	19	20	17	17
인천_읍면부	10	1	1	1	1	1	1	1	1	1	1
광주	80	7	7	7	7	9	9	9	10	7	8
대전	80	8	7	8	7	8	8	9	9	8	8
울산_동부	60	5	4	6	5	7	6	7	7	6	7
울산_읍면부	10	1	1	1	1	1	1	1	1	1	1
세종_동부	10	1	1	1	1	1	1	1	1	1	1
세종_읍면부	10	1	1	1	1	1	1	1	1	1	1
경기_동부	700	60	56	71	66	79	77	81	82	63	65
경기_읍면부	120	10	8	11	8	13	11	16	14	15	14
강원_동부	60	5	4	5	5	6	6	8	7	7	7
강원_읍면부	30	2	2	2	2	3	2	4	3	5	5
충북_동부	50	5	4	5	4	5	5	6	6	5	5
충북_읍면부	40	3	2	4	3	4	3	5	5	6	5
충남_동부	60	6	5	6	5	7	7	7	7	5	5
충남_읍면부	60	4	3	5	4	7	5	8	7	9	8
전북_동부	70	6	6	6	5	7	7	9	9	7	8
전북_읍면부	30	2	1	2	2	3	2	5	3	5	5
전남_동부	50	5	4	4	4	5	5	7	6	5	5
전남_읍면부	50	4	3	3	3	5	4	7	6	8	7
경북_동부	80	7	5	7	6	9	9	10	10	8	9
경북_읍면부	60	4	3	4	4	6	5	8	7	10	9
경남_동부	120	10	8	10	9	13	13	15	16	13	13
경남_읍면부	60	4	3	4	4	6	6	8	7	9	9
제주_동부	30	2	2	3	3	3	3	4	4	3	3
제주_읍면부	10	1	1	1	1	1	1	1	1	1	1

(9) 할당 변인별 표본 수

각 시도의 성별, 연령대 표본 수는 2025년 3월 기준 주민등록인구통계를 기준으로 비례 할당하였으며, 실제 조사된 표본 수와 모집단 통계는 아래와 같다. 일반적으로 표본조사에서는 전체 모집단 비율 추정 과정에서 실제 모집단 구성비를 정확히 반영하고자 할 경우 가중치 보정 등의 사후층화 과정을 거친다. 그러나 본조사에서는 표본 설계 단계에서부터 모집단 통계를 기준으로 한 비례할당추출을 사용했고, 조사 과정에서도 시도별, 성별, 연령별 할당표를 최대한 지켜가며 조사를 진행하였다. 그 결과 실제 조사된 표본의 지역별, 성별, 연령대별 구성비는 모집단 구성비와 비슷하게 나타난다. 따라서 추가적인 가중치 보정을 고려할 필요는 없어 보인다.

<표 II-12> 시도별, 성별, 연령별 표본 수

대분류	구분	표본		모집단 통계(20~69세) (2025년 3월 주민등록인구통계)	
		표본 수(명)	구성비(%)	인구수(명)	구성비(%)
합계		3,000	100.0	36,503,421	100.0
시도별	서울	560	18.7	6,869,386	18.8
	부산	190	6.3	2,288,868	6.3
	대구	140	4.7	1,673,492	4.6
	인천	180	6.0	2,213,653	6.1
	광주	80	2.7	999,158	2.7
	대전	80	2.7	1,040,614	2.9
	울산	70	2.3	797,824	2.2
	세종	20	0.7	268,809	0.7
	경기	820	27.3	9,945,806	27.2
	강원	90	3.0	1,042,997	2.9
	충북	90	3.0	1,118,052	3.1
	충남	120	4.0	1,469,144	4.0
	전북	100	3.3	1,168,244	3.2
	전남	100	3.3	1,182,457	3.2
	경북	140	4.7	1,716,012	4.7
	경남	180	6.0	2,245,569	6.2
	제주	40	1.3	463,336	1.3
성별	남자	1,523	50.8	18,551,964	50.8
	여자	1,477	49.2	17,951,457	49.2
연령별	20대	490	16.3	5,861,737	16.1
	30대	547	18.2	6,595,951	18.1
	40대	623	20.8	7,631,158	20.9
	50대	705	23.5	8,633,311	23.7
	60대	635	21.2	7,781,264	21.3

각 시도의 학력별 표본 수는 가장 최신 자료인 2020년 인구주택총조사 자료를 기준으로 비례 할당하였다. 조사 시점과 인구주택총조사의 시점 간에 약간의 차이가 있으나, 20년 인구주택총조사가 활용 가능한 모집단 정보 중 가장 최신의 자료였으며, 만 20~69세 인구의 학력 분포가 5년 사이에 크게 달라지지 않았을 것으로 가정했다. 실제 조사된 표본 수와 모집단 통계는 아래 표와 같으며, 표본 집단의 학력별 분포가 모집단 분포 대비 최소 80%를 맞추도록 하였다.

<표 II -13> 시도별/학력별 표본 수

구분	표본 수(명)			모집단 통계(20~69세) (2020년 인구주택총조사)		
	고졸 미만	고졸	대재 이상	고졸 미만	고졸	대재 이상
합계(%)	9.3	33.8	56.8	9.5	34.0	56.6
합계(명)	280	1,015	1,705	3,409,724	12,235,255	20,370,831
서울	39	153	368	483,633	1,881,817	4,529,585
부산	21	59	110	268,115	791,698	1,322,134
대구	15	44	81	181,520	536,867	995,577
인천	16	72	92	193,440	845,432	1,053,172
광주	6	26	48	79,039	323,346	629,627
대전	8	23	49	91,975	316,623	654,189
울산	7	28	35	77,968	314,773	424,631
세종	1	4	15	12,721	50,918	173,312
경기	62	283	475	715,141	3,242,896	5,411,683
강원	10	34	46	138,717	377,419	528,163
충북	11	36	43	137,227	418,728	554,856
충남	15	45	60	174,031	564,703	684,122
전북	12	37	51	143,681	433,680	612,969
전남	15	39	46	171,412	448,893	530,516
경북	19	54	67	243,574	676,433	850,817
경남	20	67	93	260,107	868,972	1,149,549
제주	3	11	26	37,423	142,057	265,929

각 시도의 직업별 표본 수는 조사 준비 시점에서 가장 최신 자료인 2025년 1분기 경제활동인구조사 자료를 기준으로 비례 할당하였다. 다만 경제활동인구조사에서 직업별 취업자 통계는 만 15세 이상 국민 중 경제활동인구를 기준으로 통계치를 집계한다는 점에서 본조사의 모집단과 차이를 보인다. 따라서 경제활동인구조사 자료 중 연령대별 경제활동인구수와 연령대별 경제활동참여율 등의 자료를 가공하고, 이를 직업별 취업자 통계에 반영하는 방식으로 만 20~69세 인구 중 취업자(정신노동, 육체노동)와 비취업자(비경제활동인구, 실업자)의 규모를 추정했다. 그런 다음 각 직업별 분포에 비례하도록 시도별 표본을 할당하였다.

실제 조사된 표본 수와 모집단 통계는 아래 표와 같으며, 표본 집단의 직업별 분포가 모집단 분포 대비 최소 80%를 맞추도록 하였다.

<표 II-14> 시도별/직업별 표본 수

구분	표본 수(명)			모집단 통계(20~69세) (2024년 1분기 경제활동인구조사, 단위: 천 명)		
	정신노동	육체노동	기타	정신노동	육체노동	비경제활동 인구 (추정치)
합계(%)	30.5	40.7	28.8	30.5	40.6	28.9
합계(명)	914	1,221	865	11,218	14,924	10,637
서울	217	175	168	2,748	2,099	2,070
부산	52	77	61	630	917	733
대구	37	57	46	437	686	544
인천	58	72	50	689	907	632
광주	21	34	25	285	422	330
대전	25	32	23	315	427	325
울산	15	32	23	175	366	255
세종	9	6	5	126	74	72
경기	287	310	223	3,471	3,848	2,744
강원	19	42	29	230	470	317
충북	22	44	24	278	581	295
충남	28	62	30	383	745	423
전북	23	47	30	270	557	342
전남	17	56	27	195	625	305
경북	29	71	40	356	882	503
경남	41	87	52	506	1,084	644
제주	14	17	9	124	234	104

2) 집합 조사

집합 조사에서는 수도권 지역의 20대 남/여, 30대 남/여, 40대 남/여, 50대 남/여, 60대 남/여의 총 10개 구간으로 설정하고, 구간별로 동일 크기의 표본을 할당하였다. 표본 규모가 50명으로 크지 않으므로, 지역과 직업, 학력 변수는 고르게 배분하였다.

<표 II-15> 집합 조사 표본 할당표

구분 (단위: 명)	합계	20대		30대		40대		50대		60대	
		남	여	남	여	남	여	남	여	남	여
합계	50	5	5	5	5	5	5	5	5	5	5

당일 참석 취소를 고려하여 연령별로 1명의 예비 인원을 추가로 모집하였으며, 해당 예비 인원 5명을

포함한 총 55명이 집합 조사에 참여하였다.

2.3. 본조사 방식

1) 조사 방식의 타당성 확보 방안

(1) 방문 면접 조사와 집합 조사

본 조사는 방문 면접 조사 3,000명과 집합 조사 50명을 조사하는 복합적 조사이다. 방문 면접 조사의 타당성을 확보하기 위해 질문을 명확하게 작성하고, 응답자가 이해하기 쉽게 질문을 구성하였다. 또한 체계적인 조사원 교육을 진행하여 조사원들이 조사 방법과 지침에 대해 정확히 설명할 수 있도록 하였다.

집합 조사의 타당성을 확보하기 위해 참여자 모집 시 특정 연령대나 성별에 쏠리지 않도록 균형 잡힌 샘플링을 진행하였다(성, 연령별 동일 표본 배분). 또한, 2024년에 실시한 집합 조사 절차상의 문제점을 분석하여 집합 조사에 필요한 세부 지침을 마련하고 태블릿 피시를 이용한 대인면접조사(TAPI: Tablet PC Aided Personal Interview) 시스템을 보완하였다. 세부적으로, 응답자용 지침을 1쪽짜리의 안내문 형식으로 제작하여 응답자가 조사 전, 충분히 숙지할 수 있도록 하며, 연구원이 안내문을 기반으로 직접 교육을 시행하였다. 이를 통해 동일한 기준에서 조사를 진행할 수 있도록 하여 응답자의 편향을 최소화할 수 있도록 하였다.

또한, 태블릿 피시의 전자설문지에는 문항별 시간 제한 및 글자 수 제한이 명시되어 있고, 더불어 남은 시간과 현재 작성한 응답의 총글자 수가 실시간으로 표시되게끔 구현하였으나, 종이로 진행하면 이를 응답자가 스스로 확인할 수가 없어, 집합 조사 대상자 중 종이 설문지를 활용하는 응답자는 연구원이 옆에서 직접 문항별 시간을 확인하여 응답자에게 알리도록 하였다.

(2) 태블릿 피시 활용과 종이 설문지 활용

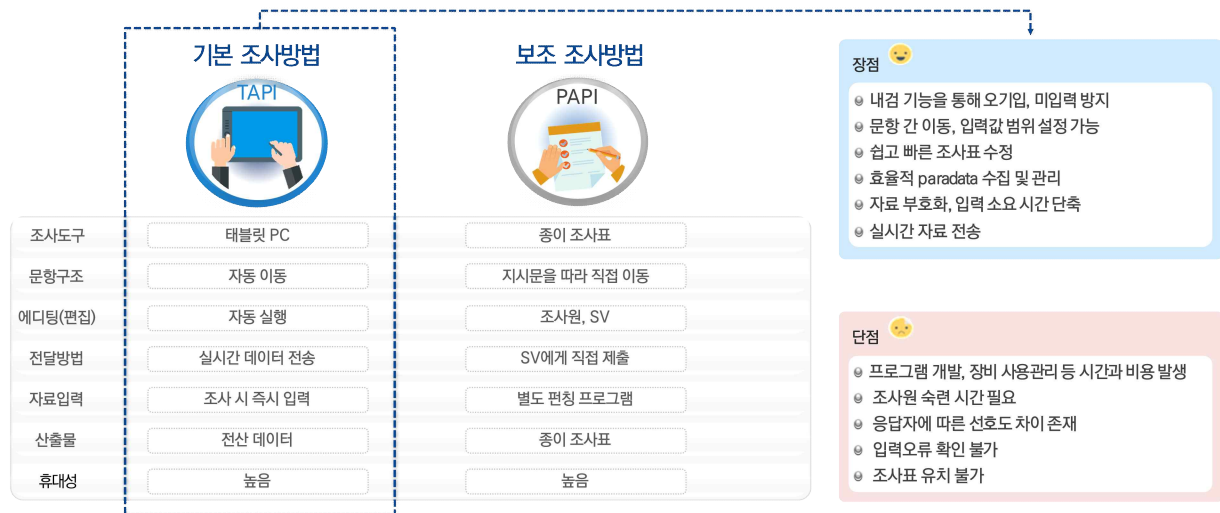
본 조사는 원칙적으로 태블릿 피시를 활용한 전자설문지를 기반으로 진행되지만, 태블릿 피시에 익숙하지 않은 고령층 등 일부 응답자가 원하는 경우 종이 설문지를 활용할 수 있도록 하였다. 태블릿 피시를 활용한 조사에서는 출제자의 의도대로 전자설문지를 정확하게 구현하며, 실사 시작 전에 사전 테스트를 진행하여 오류나 응답자의 불편이 없게 하였다.

종이 설문지를 활용한 조사에서는 태블릿 피시를 활용한 전자설문지와 동일한 조건에서 조사가 진행될 수 있도록 문항별 시간 제한과 글자 수 제한을 설정하고, 조사원이 응답 시간을 확인하여 응답자가 일관되게 조사를 진행할 수 있도록 하였다. 종이 설문지에 기재된 내용은 조사원이 직접 태블릿 피시에 입력하며, 설문지 회수 이후에는 연구원이 이를 대조하여 결과의 정확성을 확인할 수 있도록 하였다. 이를 통해 중복 입력이나 누락을 방지하며, 이를 통해 조사 방식의 일관성과 자료의 신뢰성을 확보하였다.

(3) 조사 도구 개발 및 전자설문지 구현

한국갤럽조사연구소는 서울대 국어교육연구소가 개발한 조사 도구를 온라인 설문지로 구현하는 역할

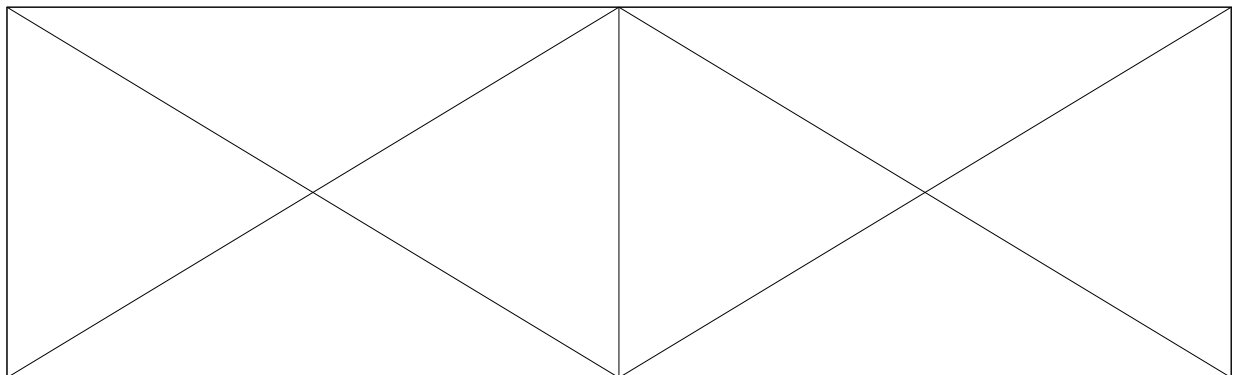
을 수행하였다. 원활한 조사 진행과 평가의 신뢰도 제고를 위해 조사 시작 전, 모든 도구 제작을 완료하고, 국립국어원, 서울대 국어교육연구소와 함께 시연 및 검토 과정을 거쳤다. 검토 과정에서 나타난 문제점을 해결하여, 조사 도구 제작을 완료하였다.

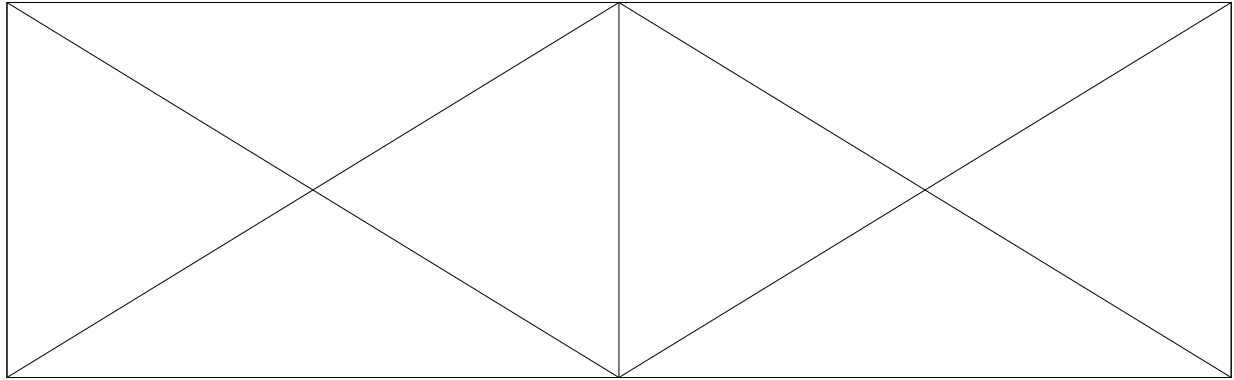


[그림 II-5] 조사 방식

(4) 쓰기에 특화된 태블릿 피시 대인면접조사(TAPI) 시스템 개발

2025년 국민의 국어능력 실태 조사를 위해 한국궤립조사연구소는 본조사에 특화된 TAPI 시스템을 개발하였다. 2024년 예비 조사 TAPI 시스템을 기반으로 글자 크기, 디지털 문항 전환 등 응답자 편의성 차원에서 시스템을 보완하여 본조사를 실시하며, 2024년 집합 조사에서 발생한 문제점(문항별 시간 제한, 글자 수 제한 등)을 보완하여 응답자 편의성과 실사 품질관리 기능을 강화하였다.





[그림 II-6] 쓰기 영역 TAPI 시스템 예시

2) 방문 면접 조사

(1) 조사 방법

본조사는 1:1 방문 면접 조사로 진행하며, 태블릿 피시의 디지털 검사지를 통해 자료를 수집하고, 본 조사에 특화된 TAPI(Tablet Aided Personal Interview) 시스템을 활용하였다. 화면 속 링크나 아이콘, 메뉴를 클릭하여 글을 읽거나 스크롤을 내리는 등의 디지털 문항을 구현하였으며, 이는 2024년 예비 조사에서도 활용한 기능으로 본조사를 위한 시스템을 개발하여 응답자 편의성과 실사 품질관리 기능을 강화하였다.

또한 조사 과정에서 태블릿 피시의 배터리 방전으로 인해 조사가 중단될 위험이 있거나 응답자가 태블릿 피시를 이용한 조사 방식을 부담스러워할 때는 종이 설문지로 조사를 진행한 다음, 조사원이 응답 결과를 TAPI 시스템에 입력하도록 하였다.

다만, 외국인이나 의사소통이 어려운 장애인은 조사 대상에서 제외하며, 조사에 끝까지 응답한 자료만 유효한 표본으로 활용하였다. 선정된 응답자가 모든 문항에 대한 평가를 완료하면, 조사원이 응답자에게 사례품을 전달하였다.

<표 II-16> 본조사(방문 면접 조사) 진행 절차

구분	주요 내용
조사원 선발	한국갤럽 소속 경력 5년 이상의 전문 조사원 중 국민의 국어능력 실태 조사 수행 경험 및 TAPI 조사 경험이 있는 조사원을 선발
조사원 교육	선발된 조사원에게 조사의 목적 및 내용, 실사 설계, 조사 시 유의 사항, TAPI 조작법 등 교육 시행
조사원 관리	- 한국갤럽조사연구소 실사 관리시스템을 통해 조사원이 전송한 조사 진행 상황 및 내용 모니터링 - 검증 결과에 대한 통보 및 주의를 통해 조사원 통제 및 감독, 조사 진행상의 문제점 및 해결 방안 안내
검증	전화 검증팀 검증(전체 조사 대상 중 20%를 임의로 선정하여 응답자에게 전화 검증)

(2) 조사원 선발 및 교육

본조사는 전국 단위의 대규모 조사이며, 국어능력에 대한 평가로 쓰기 영역으로 구성되어 난도가 높고, 응답자의 적극적인 협조가 필요한 조사이므로, 반드시 숙련된 조사원이 투입되어야 한다.

약 2개월의 실사 기간을 고려하여 조사원 1인당 4개 정도의 읍면동, 즉 약 40명을 조사하는 것을 기준으로 필요 조사원 수를 산출하였다.

조사원 1인당 평균 4개 읍면동(40명)을 전담할 수 있도록 계획하여 총투입되어야 하는 조사원 수는 약 75명이지만 조사원 대체나 추가 투입을 고려하여 초기에 90명의 조사원을 선발하였다.

조사를 성공적으로 수행하기 위해 2024년 국민의 국어능력 실태 조사 수행 조사원과 태블릿 피시를 활용한 가구 방문 면접 조사 경험이 풍부한 조사원을 선발하여 투입하여 조사 품질을 제고하고자 하였다.

조사원 교육은 참여 연구진이 직접 지역 실사사무소가 위치한 지역을 방문하여, 지역별로 조사원을 한 번에 모아 집체교육을 실시하였다. 조사원 교육 시에는 지역별 실사 슈퍼바이저 및 에디터도 참여하도록 하며, 예비 조사원을 포함한 조사 참여 조사원 전원이 참석하였다.

조사원 교육 내용은 이론 중심 교육과 실습 중심 교육으로 구분하여 진행하며, 특히, 조사원이 피험자가 되어 본 연구의 조사 도구를 직접 경험하게 하여 조사원의 이해 부족으로 인한 오류 발생 가능성을 낮추었다.

■ 5단계 교육

1단계	2단계	3단계	4단계	5단계
일반교육	설문교육	TAPI 교육	Role-Playing 교육	조사원별 교육 평가
- 조사 취지 및 개요 - 조사 수행 배경 및 목적 - 조사 중요성 강조 - 응답 거부 시 대응 방안 - 우수 현장조사사례 공유	- 설문지침서를 활용한 교육 진행 - 각 문항별 주요 개념 - 주요 문항 및 현장 에디팅 방법	- TAPI에 대한 기본적인 사항 및 사용법 설명 - 할당 관리 방법 - 응답자 정보 작성 - 쓰기 조사를 위한 키보드 연결 방법/오류 해결 방법	- 조사원 2명씩 짝을 이뤄 Role-Playing 실시 - 조사 대상자 컨택 단계에서부터 조사 실시, 거절 시 응대요령까지 모든 과정 포함	- 교육이 완료된 후 교육 내용 전반에 대한 조사원 평가표 작성 - 조사원 평가표 채점 후 80점 미만인 조사원은 재교육 실시

■ 권역별로 표준화된 교육 실시

- 교육은 본사 연구원이 진행
- 표준화된 교육을 위해 각 지역별로 집체교육 실시
- 지역별로 동일한 연구원이 진행하여 표준화된 교육안으로 지역간 편차 발생 예방
- 조사원 교육 시에는 실사 슈퍼바이저 및 에디터 뿐 아니라 예비조사원을 포함한 조사 투입 인력 전원이 참석하도록 함

▼ 조사도구

조사용 태블릿 PC	설문지	응답자용 IRB 동의서	키보드
답례품	협조 공문	조사원 지침서	조사원증

[그림 II-7] 조사원 교육 체계

05 조사 설계

01 총 표본수 300명

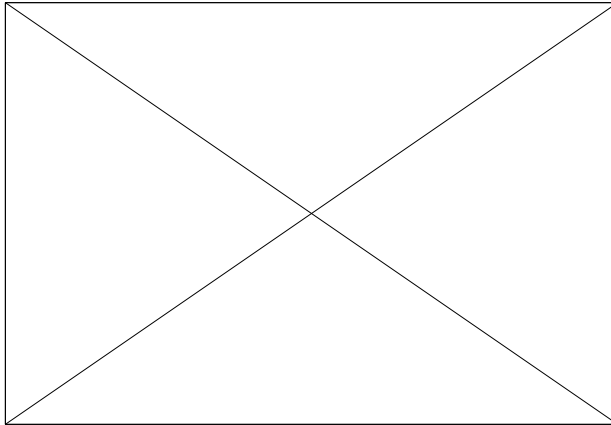
02 수도권에서 28개의 읍/면/동 추출

03 각 읍면동에서 10명 또는 11명의 응답자 추출 (만20-69세 남녀)

04 지역, 성별, 연령대, 직업, 학력을 고르게 조사

※ 성별×연령 10세별 + 직업 + 교육수준 할당이 있으므로, 조사 진행 과정에서 수퍼바이저와 긴밀한 소통 필요

지역	서울, 인천, 경기(동부, 읍면부)
성별	남자, 여자
연령대	20대~60대, 10세 간격
직업	1) 정신노동 : 관리자, 전문가 및 관련 종사자, 사무 종사자 2) 육체노동 : 서비스 종사자, 판매 종사자, 농업·어업·축산업 종사자, 기능원 및 관련 기능 종사자, 장차·기계 및 조립 종사자, 단순 노동 종사자, 군인 등 3) 기타 : 주부, 학생, 무직
교육수준	중졸 이하, 고졸, 대졸 이상



[그림 II -8] 조사원 교육 예시

03 설문 진행을 위해 기억할 사항

이것은 반드시 기억해야 합니다!!

조사원의 역할과 준수해야 할

- 모든 문항은 응답자가 스스로 읽고 풀어야 함
- 응답자가 어려워하는 문항을 더 쉽게 설명해 주거나, 힌트를 주는 행위 금지
- 응답자 본인의 국어 능력대로 참실하게 응답할 수 있게 관리

모든 응답자가 동일한 환경에서 평가에 임해야 함

- 설문지와 동일한 순서로 평가 진행
- 시험과 비슷한 국어 능력을 평가하는 조사이므로, 주변 소음이 없는 조용한 환경에서 진행

평가 시간을 관리해야 함

- 각 영역과 문항별 제한 시간이 정해져 있음
- 최대한 평가 시간에 맞춰 진행해야 함

응답자의 문제 풀이 과정에 소중해야 함

- 복수의 응답을 체크 하거나 문제를 건너뛰는 것, 문제를 제대로 이해하지 않고 답을 선택하는 일이 없도록 해야 함
- 응답자가 참여를 꺼려도 최대한 능력 안에서 참여가 가능하도록 유도, 설득

정답을 잘 맞추는 것보다 **객관적인 실력을 파악**하는 것이 중요

04 설문 진행 방법 : 쓰기

쓰기 영역(4문항)

- 모든 문항은 하나의 답만 있음. 동시에 진행되는 경우 2개의 보기에 표시하거나, 문항을 건너뛰는 일이 없도록 함
- 태블릿 자체 키보드나 블루투스 무선 키보드를 통한 입력(타입)을 우선으로 함(무엇이든 경우 응답자가 종이에 자필로 작성 후 조사원이 태블릿PC에 입력)
- 종이에 쓴 내용이나 타이핑 과정에서 오답이나 어법에 맞지 않은 문장이 있을 경우에도 교정을 하지 않고 그대로 입력
- 각 유형별 4번 문항(디지털 문항)은 응답자가 직접 태블릿PC를 터치하거나 입력하며, 대리 입력은 불가(단, 조사원이 클릭 버튼 등 간단한 화면 구성에 대한 설명은 가능)
- 사전에 응답시간 및 문항에 대한 고지 필수. 문항별 분량(300자, 400자, 1문장, 3문장 등)이 정해져 있음. 분량을 최대한 지켜 답안을 작성하도록 안내
- 주어진 시간을 초과하여 응답한 경우 조사원 메모에 반드시 기입

STEP 01 설문지 전달 및 시간 고지

STEP 02 제한된 시간 내 문제를 풀 수 있도록 통제

STEP 03 응답이 충실히 되었는지 확인

조사원 교육은 효과를 보다 높이기 위해 조사원을 모아 각 지사에서 다음과 같이 집체교육을 실시하였다. 본조사에 투입된 조사원은 총 90명이다.

<표 II -17> 본조사(방문 면접 조사) 조사원 교육 결과

실사 지점	해당 지역	교육 일시	교육 장소	참여 인원
서울	서울, 인천, 경기, 강원, 제주	6/19(목) 10시	갤럽 본사	45명
대전	대전, 세종, 충남, 충북	6/19(목) 15시	갤럽 대전지사	11명
대구	대구, 경북	6/20(금) 10시	갤럽 대구지사	10명
부산	부산, 울산, 경남	6/20(금) 15시	갤럽 부산지사	14명
광주	광주, 전남, 전북	6/23(일) 13시	갤럽 광주지사	10명

(3) 방문 면접 조사 준비 및 실시

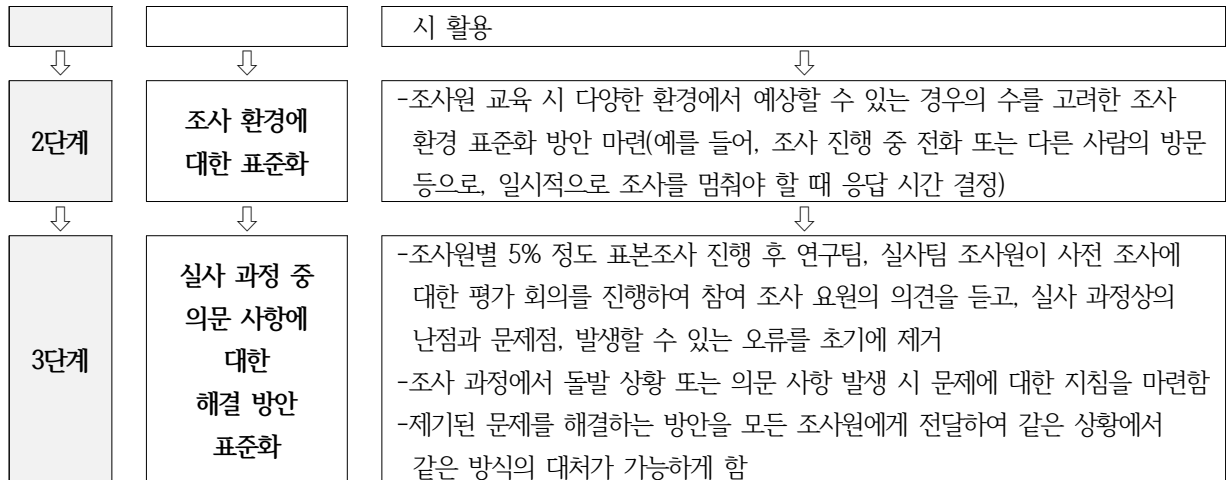


(4) 실사 품질관리

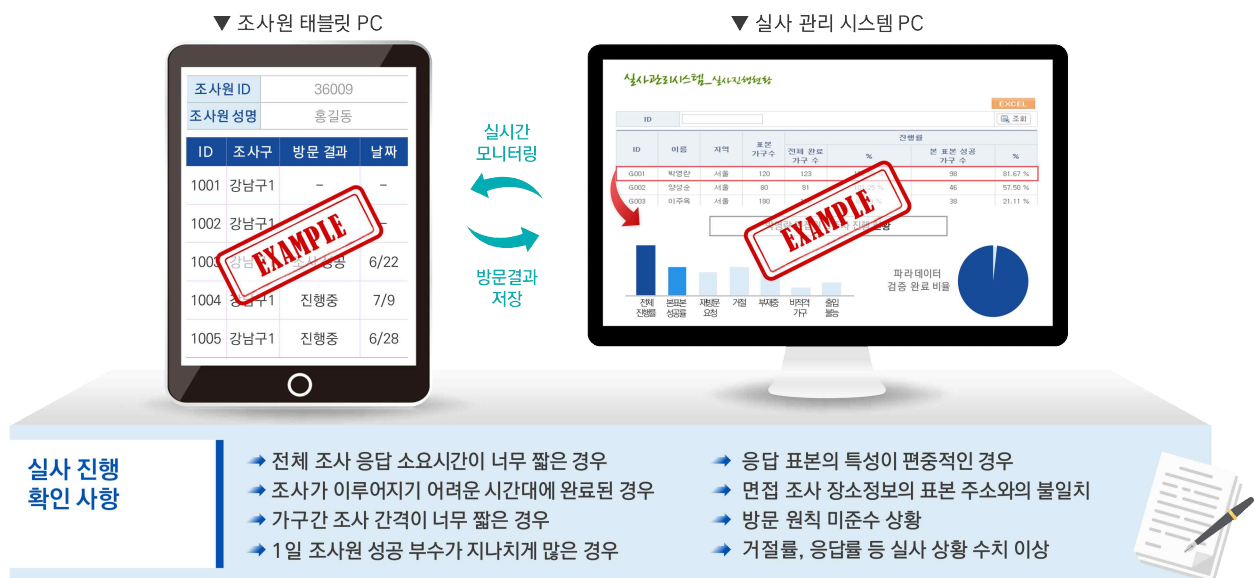
조사원에게는 표준화된 지침과 절차를 제공한다. 다만, 가구 방문 면접 조사의 특성상, 조사 과정에서 제삼자의 방문, 전화 통화, 가사 일 등으로 인해 조사 진행이 일시 중단될 수 있으며, 낯선 사람을 집에 들이지 않으려는 경계 심리로 인해 공동주택의 로비나 휴게실, 공원 벤치 등에서 조사가 진행되는 등 여러 가지 상황이 발생할 수 있다.

이처럼 개별 방문 면접 조사에서는 피험자들이 모두 동일한 환경에서 평가를 받기 어려우므로, 비표본 오차를 최소화하기 위해 조사 방식과 절차를 표준화하기 위한 노력이 필요하다. 이를 위해 다음과 같은 표준화된 절차를 수행하였다.

1단계	조사원 교육 표준화	-표준화된 교육안/비상 상황 지침을 제작하여 조사원들에게 배포 -동일한 교육자에 의한 교육 -교육 시 교육안을 영상 제작하여 조사 요원 교육 및 기존 조사 요원의 재교육
-----	------------	--



조사 진행 상황에 대한 일일 현황 집계, 보고를 위해 일일 실사 통제 시스템을 구축하여 활용하였다. 실사 통제 시스템을 통해 조사원의 개별적 대처나 해결이 어려운 상황 등 문제 상황 발생을 파악하고, 문제 상황 해결을 위한 신속히 대처하여 실사 품질을 제고하였다.



[그림 II -9] 실사 통제 시스템 예시

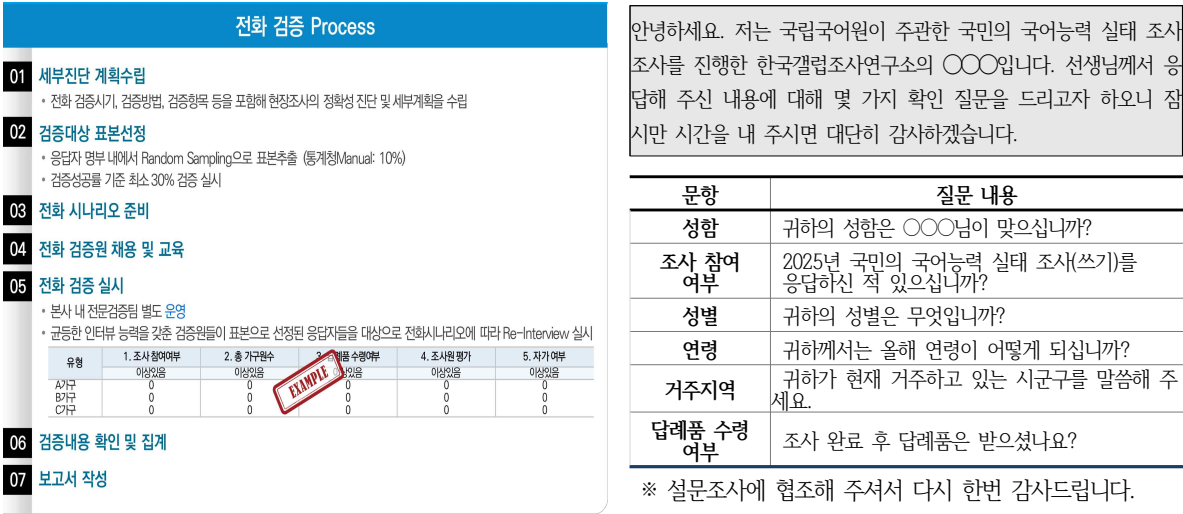
(5) 사후 검증

조사원이 완성해서 전송한 조사 결과에 대해서 실사부서와는 독립된 검증팀이 사후 검증을 진행한다. 일반적인 설문조사의 전화 검증 비율은 10% 내외이고 국가데이터처 역시 10%에 대해 사후 검증을 시행하나, 본조사의 경우 국가데이터처 기준의 3배에 해당하는 총조사 표본의 30%에 대한 전화 검증을 완료하는 것을 목표로 사후 검증을 진행하였다. 검증 내용은 응답자의 주소지 및 조사 참여 여부, 연령대 등 주요 사실에 대한 재확인으로 구성되며, 추가로 답례품 수령 여부 등을 질문하였다. 사후 검증 결

과 문제가 발견된 표본의 경우 실사 관리자의 재확인 절차를 거치며, 재확인 절차에서도 ‘허위 기재’ 등 오류가 발견되면 해당 표본을 폐기하고 다른 응답자를 선정하여 재조사하는 과정을 거쳤다. 의도적인 허위 기재가 확인된 조사원이 진행한 설문지를 전수 재검증하였다.

<표 II -18> 지역 사무소별 검증 현황

실사 사무소	해당 지역	표본 수			검증 내용 이상 있음
		완료	검증	비율	
서울	서울, 인천, 경기, 강원, 제주	1,690	507	30.0%	0
부산	부산, 울산, 경남	440	132	30.0%	0
대구	대구, 경북	280	84	30.0%	0
광주	광주, 전남, 전북	280	84	30.0%	0
대전	대전, 세종, 충남, 충북	310	93	30.0%	0



[그림 II -10] 사후 검증 절차 및 질문지

(6) 응답자 신뢰도 제고 및 불만 대응

기본적으로 조사원의 교육 시 조사 과정에서 발생할 수 있는 상황이나 의문 사항에 대한 지침을 주 나, 본 연구는 일반적인 설문조사와는 그 성격이 달라서 조사를 진행하는 과정에서 다양한 의문이나 문 제를 제기할 가능성이 높다. 이러한 문제를 해결하기 위해 다음과 같은 대응 방안을 마련하였다.

[1단계]	■ 한국갤럽 연구팀과 실사 감독원(supervisor) 간 주기적 소통 매주 실사팀으로부터 주간 진행 상황을 보고받아 조사 진행률에 대해 점검. 또한 정기 회의를 통해 조사 원의 어려움과 피험자들의 불만 사항을 파악하여 문제 상황 발생을 미리 방지
[2단계]	■ 서울대 국어교육연구소와 한국갤럽의 협조 체계 구축 한국갤럽은 실사팀으로부터 보고받은 문제 상황 중 조사 기관 차원에서 해결이 어려운 문제는 서울대 국 어교육연구소와 공유하고, 해결책을 찾도록 함

[3단계]	■ 국립국어원의 협력 체계 구축 피험자의 거부감이 예상되는 본조사에서 예상할 수 있는 문제 상황 중 하나는 조사(평가)에 대한 의구심과 개인 정보 노출에 대한 우려임. 이에 대비해 국립국어원의 협조 요청 공문과 함께 문의 전화 등이 있으면 조사를 진행하고 있는 것이 사실이라는 확인을 해 줄 수 있는 협력 체계 구축
-------	--

실사 진행 시 본조사를 위한 콜센터 운영을 통해 조사 문의 및 불만 사항에 대해 신속히 응대하고자 하였다.

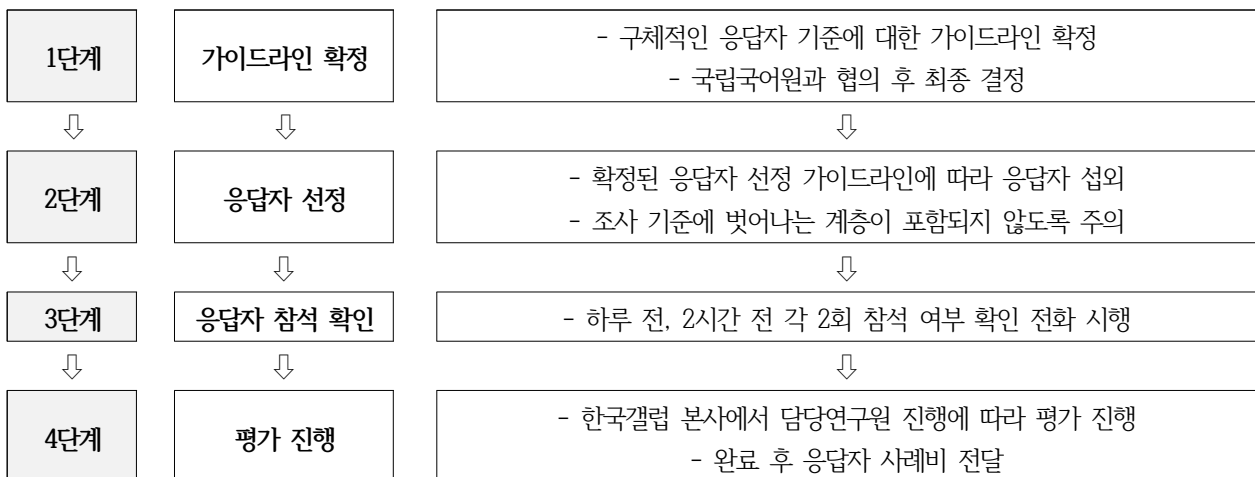
콜센터 상담원은 조사 참여 과정에서 발생하는 문의와 요청에 적극 대응하여 불만을 해결하고, 필요 시 연구원이 직접 응대함으로써 응답률 제고, 조사의 공신력, 정확성 등을 제고하였다.

3) 집합 조사

본조사에서는 디지털 환경의 변화와 함께 ‘디지털 문항’이 포함되어 있어 집합 조사에서도 방문 면접 조사와 동일하게 태블릿 피시를 활용한 자기기입식 조사로 진행하였다. 또한 방문 면접 조사와 동일한 조건에서 진행하기 위해 디지털 문항을 제외한 문항에서 응답자가 원하는 경우 종이 설문지를 보조적으로 활용할 수 있도록 했다.

집합 조사는 응답자에게 조사 안내문을 배부한 후, 서울대 국어교육연구소 연구원이나 한국갤럽조사연구소 담당연구원의 주도로 조사 진행 방법 및 조사 시 유의 사항을 약 10분간 설명했다. 조사에 대한 안내 이후 조사 시작은 동시에 진행하며, 종료 시각도 동일했다. 다만, 집합 조사는 방문 면접 조사와 조사 방식에 따른 차이를 파악하기 위해 실시하는 조사이므로, 평가 시간만 안내 후 풀이 과정에서는 별도의 통제를 하지 않았다. 당일 참석 취소를 고려하여 5명의 예비 인원을 포함한 총 55명이 집합 조사에 참여하였다.

<표 II-19> 집합 조사 진행 과정



<표 II -20> 집합 조사 진행 일시 및 장소

집합 조사 일시	조사 장소	참여 인원
8월 7일 10시 30분 ~ 11시 30분	한국갤럽 본사 1층 세미나실	24명
8월 7일 14시 00분 ~ 15시 00분	한국갤럽 본사 1층 세미나실	10명
8월 7일 19시 00분 ~ 20시 00분	한국갤럽 본사 1층 세미나실	21명

2025년 국민의 국어능력 실태 조사 - 쓰기 영역 -

조사 소개

이 조사는 문화체육관광부 국립국어원이 주관하고, 서울대학교 국어교육연구소와 대한국립조사연구소가 수행합니다. 이 조사의 목적은 우리 국민의 전반적인 국어능력을 파악하는 데 있습니다. 귀하를 평가하려는 목적이 아니므로, 부담 없이 편안한 마음으로 응답해 주시면 됩니다. 귀하께서 답변하신 자료는 국어 정책을 수립하는 데 큰 도움이 됩니다.

조사 대상/내용

본 조사는 방문 면접조사와 집합 조사로 나뉘어 진행됩니다. 방문 면접조사는 전국에 거주하는 만 20~69세의 남녀 3,000명을 대상으로 실시되며, 귀하께서 참여하시는 집합 조사는 수도권에 거주하는 만 20~69세의 남녀 50명을 대상으로 실시됩니다. 조사 내용은 쓰기과 관련한 주관식 문항으로 구성되어 있으며, 설명적 글쓰기, 논증적 글쓰기, 친교적 글쓰기, 디지털 매체 글쓰기의 총 4개 문항과 글쓰기와 관련한 응답자 배경 질문 8문항이 포함됩니다.

비밀 보장

본 조사에서는 주소(시도), 성별, 나이, 직업, 학력 등의 개인정보를 수집합니다. 수집된 개인정보는 통계적 자료 분석의 목적으로만 사용되며, 이름과 연락처(휴대폰) 등 개인식별정보는 데이터 품질 관리를 위한 사후 검증 자료로만 활용됩니다. 또한, 조사 과정에서 수집된 모든 자료는 개인정보보호법 제18조에 의해 비밀이 철저히 보장됩니다.

주관 기관 국립국어원

조사 방법 태블릿 피씨(PC)를 활용한 집합 조사

조사 대상 수도권 거주 만 20~69세 남녀 50명

수행 기관 서울대학교 국어교육연구소/대한국립조사연구소

조사 내용 쓰기 4문항, 글쓰기와 관련한 응답자 배경 질문 8문항

조사 시간 1시간

조사 진행

- 동의를 제출**
 - 연구 참여자는 설명문을 충분히 숙지한 후, 동의서 2장에 각각 서명하여 주시기 바랍니다.
 - 연구 참여자를 동의서는 본인이 보관하며, 연구자를 동의서는 조사 종료 후 제출합니다.
- 쓰기 문항 응답 시 유의사항**
 - 총 4문항, 제한 시간 40분이 주어지며, 각 문항별 제한 시간을 반드시 준수해 주시기 바랍니다.
 - 태블릿 피씨(PC)로 응답할 경우 문항 상단에 타이머가 표시되오니, 시간을 확인하며 응답해 주세요.
 - 각 문항에는 지정된 분량이 있습니다. (1번은 300자, 2번은 400자 등) 분량 미달로 인한 불관련한 답변 작성에 없도록 시간을 충분히 활용하여 답변 구성과 입력에 집중해 주세요.
 - 입력한 글의 분량은 태블릿 화면에 실시간으로 표시됩니다.
- 입력 방식 관련 유의사항**
 - 응답자는 직접 자판을 이용하여 답안을 입력해야 합니다. 단, 자판 이용이 어려운 고령자 등 부득이한 경우에는 자필(손글씨)로 종이에 작성할 수 있습니다.
 - 디지털 매체 글쓰기 문항(4번)의 경우, 피할자가 직접 태블릿을 터치하거나 입력하여 문제를 해결해야 하며, 대리 입력은 허용되지 않습니다. 단, 화면 구성(예 버튼 위치 등)에 대한 간단한 질문은 허용되오니 도움이 필요한 경우 손을 들어 알려 주시기 바랍니다.
- 응답 종료 후 퇴실 안내**
 - 모든 문제 풀이가 끝났더라도 제한 시간(40분) 이전에 퇴실은 불가능합니다.
 - 40분이 지나고 응답을 마친 경우에 한해 퇴실 가능합니다.
- 조사 참여 시간제**
 - 본 조사에 '모두 응답'한 경우에 한하여 5만 원의 사례비(현금)가 지급됩니다. 다만, 중도에 포기하더라도 '쓰기 문항 2번'까지 응답한 경우에 한해서는 2만 원의 사례비가 지급됩니다.
 - 조사에 참여해 주셔서 감사합니다.

[그림 II -11] 응답자 배부용 집합 조사 안내문

<표 II -21> 집합 조사 진행

내용	소요 시간	비고
참석자 확인 및 접수	30분	-
조사 진행 안내 및 유의 사항 전달	10분	한국갤럽조사연구소 연구원 진행
조사 진행	50분	쓰기 영역 40분 / 배경 질문 10분 내외
합계	90분	-

2.4. 문항 채점

1) 채점 설계

채점 설계는 대규모 채점 자료의 수집이 체계적이고 신뢰성을 보장하기 위해 요구된다. 채점 설계는 채점 방법을 고려하여 피험자가 작성한 답안을 채점자에게 안정적으로 할당하는 것과 채점 진행 과정에서 신뢰도 영향 요인을 줄이기 위해 알맞은 채점 모형을 적용하는 것으로 이루어진다. 채점 자료 할당과 채점 모형이 효과적으로 기능하여 채점이 신뢰롭게 이루어지기 위해서는 채점에 참여하는 채점자들은 글을 평가하는 균질한 관점을 형성할 수 있도록 배경 변인을 고려하여 선정할 필요가 있으며 채점 척도 적용을 위한 훈련 프로그램을 제공할 필요가 있다.

올해 쓰기 본조사의 채점은 한 피험자의 답안을 2명의 채점자가 채점하는 짝 채점(pair scoring) 방법이 적용되었다. 적용한 채점 방법을 통해 신뢰할 수 있는 채점 결과를 확보하기 위하여 채점자 선정 과정에서 2024년 예비 조사에서 실시한 라쉬 모형 분석 결과 중 채점 적합도가 부적합으로 나타난 채점자를 제외하였다. 다음으로 채점자 훈련 과정에서 채점자 간 협의와 채점 척도 및 세부 기준에 대한 숙지 교육의 효과를 확인하였을 때 급내상관(ICC)이 .97($p<.000$)로 매우 높아 대규모 조사 채점에 적합한 수준인 점을 확인하였다. 채점자의 편성은 짝을 이룬 두 채점자의 세부 전공을 다르게 하여 전공 특성에 따른 편향을 줄이고자 하였다.

채점 대상 피험자는 체계적 할당 방법으로 편성하였다. 이는 단기간에 진행하는 채점의 효율성을 높이고, 표집 과정에서 무선적으로 인구 통계에 비례하여 표본 집단이 추출되어 응답 수집 순서의 규칙성이 없다는 점을 고려한 것이다. 체계적 할당 방식에 따라 각 채점자는 개인별로 305명의 피험자가 4개 문항에 작성한 답안을 채점하도록 채점 계획을 수립하였다.

<표 II-22> 채점 자료 할당

조사	유형	피험자 (명)	채점자 (명)	채점자별 피험자 (명)
본조사	공통	3,050	20	305

2) 채점단 협의회

본조사 채점에는 총 20명이 참여하였으며, 국어교육 전공(문법, 화법, 문학) 석사 및 박사과정 재학생과 수료생, 졸업생으로 구성되었다. 선정 과정에서는 국어교육 전공자를 기본으로 이전에 실시한 2022년 국립국어원 ‘국민의 국어능력 실태 조사 개선 기초 연구’의 1, 2차 예비 조사, ‘2023년 국민의 국어능력 실태 조사’의 본조사, ‘2024년 국민의 국어능력 실태 조사’의 예비 조사 채점 등 대규모 평가 참여 경험이 있는지를 고려하였으며, 이를 통해 채점 운영의 안정성과 연속성을 확보하고자 하였다.

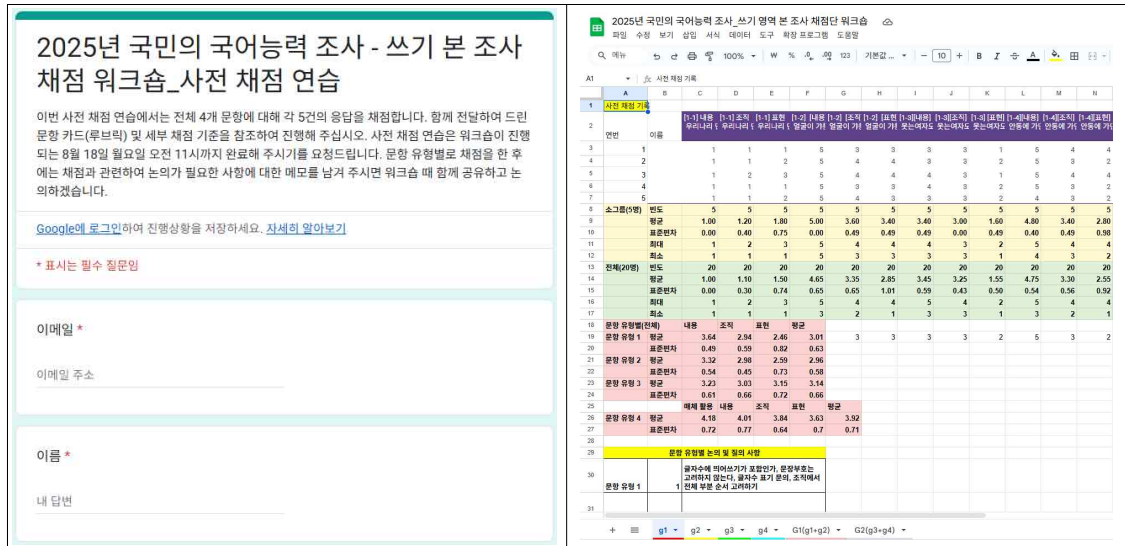
<표 II-23> 본조사 채점단

채점자 ID	전공	학력
24kr01	문법 교육	박사 수료
24kr02	문학 교육	박사 수료

24kr03	문학 교육	석사 졸업
24kr04	문법 교육	석사 수료
24kr05	화법 교육	박사 수료
24kr06	문법 교육	석사 과정
24kr07	화법 교육	박사 수료
24kr08	문법 교육	박사 과정
24kr09	문법 교육	박사 졸업
24kr10	문학 교육	박사 수료
24kr11	문학 교육	박사 수료
24kr12	문법 교육	석사 수료
24kr13	화법 교육	박사 수료
24kr14	문법 교육	박사 졸업
24kr15	화법 교육	석사 수료
24kr16	문법 교육	석사 과정
24kr17	문학 교육	박사 수료
24kr18	화법 교육	박사 수료
24kr19	문학 교육	박사 수료
24kr20	화법 교육	박사 수료

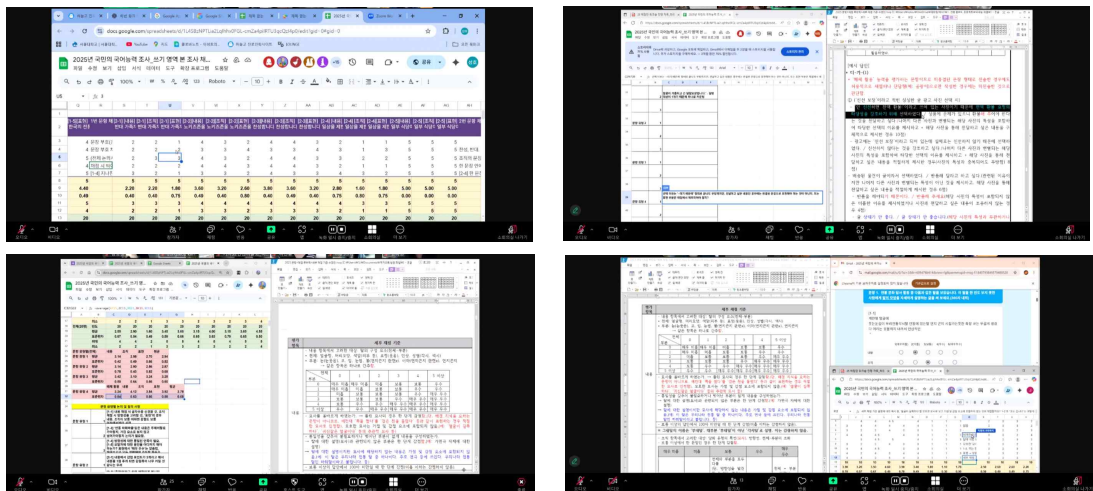
대규모 쓰기 평가에서 채점자 훈련은 채점자 신뢰도 확보와 타당한 채점 결과 산출을 위한 필수 조건이다. 채점자들은 채점 수행 과정에서 각 문항 유형별 채점 척도를 기준으로 피험자가 작성한 글의 준거별 수준을 인식하고 알맞게 판정해야 한다. 채점자들은 채점자 훈련을 통해 자신의 채점 수행의 현재적 특징을 파악하고, 평가 목적에 부합하는 온건한 채점 수행을 할 수 있도록 준비가 되어야 한다. Li et al. (2023)에서는 채점자 훈련 과정으로 척도 기술문 읽기, 척도 기술문에 대한 토의, 평가 영역 및 범위에 대한 설명 듣기, 루브릭에 대한 질의응답 및 논의, 사례 찾기(benchmarking)를 수행한 사례를 제시한 바 있으며, 본 과업에서도 채점 신뢰도 확보를 위한 협의회를 개최하여 채점자들이 채점 수행을 위한 지식 형성, 인식 고양, 실천 능력 습득을 위한 과정으로 활용한다. 이 과정을 통해 채점자들은 동일 수준의 글에 대해 일관된 채점을 수행할 수 있으며, 다양한 글 사례를 검토함으로써 내적 채점 기준을 형성하고, 공동의 글 사례 검토를 통해 문항별 채점 지침을 개발할 수 있다.

채점단 협의회는 본조사 자료의 수집량과 시기, 채점 기간, 결과 분석에 필요한 소요 기간을 고려하여, 사전에 사전 채점 연습을 실시한 후 본 협의회를 진행하는 방식으로 운영하였다. 먼저 사전 채점 연습은 채점자들에게 할당된 쓰기 본조사의 문항과 채점 척도에 대한 기초적인 이해를 형성하고, 채점 연습 경험을 통해 글과 채점 척도를 연계하는 과정에서의 개인적인 문제를 확인하는 것을 목적으로 실시하였다. 사전 채점 과제의 수행 기간은 2025년 8월 14일부터 8월 18일까지였으며, 채점자들에게는 해당하는 검사지에 응답한 피험자 5인이 작성한 답안과 점수 및 채점 관련 논의 사항을 입력할 수 있는 링크, 채점 척도를 전달하였다.



[그림 II-12] 사전 채점 연습

이후 진행된 채점 협의회는 사전 채점 연습 과정에서 발견한 문항 유형별 채점 척도 적용의 어려움이 나 판단하기 어려운 글 자료의 문제를 채점자들이 공유하고, 공동으로 협의하는 과정을 통해 최선의 채점 척도 활용 방안을 도출하는 것을 목적으로 진행되었다. 채점 협의회는 2025년 8월 18일 저녁 6시 30분부터 11시까지 줌(Zoom)을 통해 실시간 비대면 방식으로 진행되었다. 회의에서는 사전 채점 연습 결과에 대한 소집단 협의, 협의 내용을 바탕으로 한 실시간 채점 연습, 실시간 채점 연습 결과에 대한 검사지 유형별 협의, 세부 채점 세부 기준 정리를 진행하였다.



[그림 II-13] 채점 협의회

채점 협의회에서 진행한 소집단 협의에서는 사전 채점 과제 수행 결과를 함께 확인하면서 각 글 사례에 대해 준거별로 가장 엄격하거나 관대한 채점을 한 경우에 해당하는 채점자를 중심으로 채점 척도 적용에서 고려한 사항이나 판단의 이유를 공유하도록 하였다. 다음으로 실시간 채점 연습에서는 소집단 협의를 바탕으로 새로운 피험자의 글 사례에 대하여 채점을 연습하였으며, 직후에는 소집단에서 함께 그 결과를 확인하면서, 아직 채점자 간의 인식 차이가 있는 문항 유형이나 글 사례에 나타난 특징 등에

대해 협의를 진행하였다. 이 과정에서 채점자별로 각 준거와 글 특성에 대한 엄격성 차이를 재점검하고 문항별 세부 채점 기준에 대한 협의를 진행하였다.

3) 세부 채점 기준

채점자 및 연구진 간 협의를 거쳐 최종적으로 설정된 세부 채점 기준은 다음과 같다. 1~3번 문항의 경우 조사의 연속성을 고려하여 2018년 2차 조사의 세부 채점 기준을 준용하고자 하였으며, 일부 모호하거나 타당성이 다소 부족한 세부 채점 기준을 중심으로 수정하였다.

<표 II -24> 쓰기 1번 문항 세부 채점 기준

평가 항목	세부 채점 기준																																																	
내용	- 내용 항목에서 고려한 대상: 탈의 구성 요소(전체-부분) * 전체: 얼굴형, 머리모양, 색깔(피부 등), 표정(웃음), 인상, 성별(각시, 색시) * 부분: 눈(눈웃음), 코, 입, 눈썹, 볼(연지곤지 관련x), 이마(연지곤지 관련x), 연지곤지 → 같은 항목은 하나로 간주함.																																																	
	<table><tr><th> 전체 부분 </th><th>0</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5 이상</th></tr><tr><td>0</td><td>매우 미흡</td><td>매우 미흡</td><td>미흡</td><td>보통</td><td>보통</td><td>우수</td></tr><tr><td>1</td><td>매우 미흡</td><td>미흡</td><td>보통</td><td>보통</td><td>우수</td><td>우수</td></tr><tr><td>2</td><td>미흡</td><td>보통</td><td>보통</td><td>보통</td><td>우수</td><td>매우 우수</td></tr><tr><td>3</td><td>보통</td><td>보통</td><td>보통</td><td>우수</td><td>매우 우수</td><td>매우 우수</td></tr><tr><td>4</td><td>보통</td><td>우수</td><td>우수</td><td>매우 우수</td><td>매우 우수</td><td>매우 우수</td></tr><tr><td>5 이상</td><td>우수</td><td>우수</td><td>매우 우수</td><td>매우 우수</td><td>매우 우수</td><td>매우 우수</td></tr></table>	전체 부분	0	1	2	3	4	5 이상	0	매우 미흡	매우 미흡	미흡	보통	보통	우수	1	매우 미흡	미흡	보통	보통	우수	우수	2	미흡	보통	보통	보통	우수	매우 우수	3	보통	보통	보통	우수	매우 우수	매우 우수	4	보통	우수	우수	매우 우수	매우 우수	매우 우수	5 이상	우수	우수	매우 우수	매우 우수	매우 우수	매우 우수
	전체 부분	0	1	2	3	4	5 이상																																											
	0	매우 미흡	매우 미흡	미흡	보통	보통	우수																																											
	1	매우 미흡	미흡	보통	보통	우수	우수																																											
2	미흡	보통	보통	보통	우수	매우 우수																																												
3	보통	보통	보통	우수	매우 우수	매우 우수																																												
4	보통	우수	우수	매우 우수	매우 우수	매우 우수																																												
5 이상	우수	우수	매우 우수	매우 우수	매우 우수	매우 우수																																												
- 묘사를 올바르게 하였는가. → 틀린 묘사의 경우 한 단계 감점함.(단, 배경 지식을 요하는 문항이 아니므로, 예컨대 ‘쪽을 졌다’를 ‘검은 천을 둘러다’ 등과 같이 표현하는 경우 적절한 묘사로 인정함), 모호한 묘사는 가점 및 감점 요소에 포함되지 않음.(예: ‘얼굴이 길쭉하다’, ‘귀신같은 얼굴이다’ 등의 주관적 묘사 등)																																																		
- 통일성을 갖추어 불필요하거나 벗어난 부분이 없게 내용을 구성하였는가. → 탈에 대한 설명(묘사)과 관련되지 않은 부분은 한 단계 감점함.(예: 가면극 자체에 대한 설명) → 탈에 대한 설명이지만 묘사에 해당하지 않는 내용은 가점 및 감점 요소에 포함되지 않음.(예: 이 탈은 우리나라 전통 탈 중 하나이다. 주로 연극 등에 쓰인다. 우리나라 전통 탈인 하회탈이라고 불립니다. 등)																																																		
- 보통 이상의 답안에서 100자 미만일 때 한 단계 감점(미흡 이하는 감점하지 않음). ※ 그림탈의 이름은 ‘부네탈’. 대부분 ‘부네탈’이 아닌 ‘각시탈’로 설명. 이는 감점하지 않음.																																																		
조직	- 조직 항목에서 고려한 대상: 담화 유형의 특성(묘사), 방향성, 전체-부분의 조화 - 보통 이상에서 한 문장인 경우 한 단계 감점함. <table><tr><td>매우 미흡</td><td>미흡</td><td>보통</td><td>우수</td><td>매우 우수</td></tr></table>	매우 미흡	미흡	보통	우수	매우 우수																																												
매우 미흡	미흡	보통	우수	매우 우수																																														

	<table><tr><td>- 주제 이탈(아예 묘사가 안 들어간 경우) 경위담화 유형의 구조적 특성도 포함)</td><td>전체 + 부분 + 주제 이탈 문장 포함된 경우</td><td>전체와 부분을 모두 다룸 (단, 방향성을 발견할 수 없는 경우) ex) 등근 모양 + 눈과 연지근지 하얗고 연지근지가 있다. 웃는 인상이고 눈썹이 진하다. (전체적인 모습과 세부적인 모양을 방향성 없이 뒤섞어 이야기하는 경우)</td><td>전체 + 부분 + 방향성을 갖춤</td><td>전체 + 부분 + 방향성 + α(문단으로서 구조적 특성 및 완결성[중심+뒷받침 문장, 처음-중간-끝 등, 문장 간 응집성])</td></tr></table>	- 주제 이탈(아예 묘사가 안 들어간 경우) 경위담화 유형의 구조적 특성도 포함)	전체 + 부분 + 주제 이탈 문장 포함된 경우	전체와 부분을 모두 다룸 (단, 방향성을 발견할 수 없는 경우) ex) 등근 모양 + 눈과 연지근지 하얗고 연지근지가 있다. 웃는 인상이고 눈썹이 진하다. (전체적인 모습과 세부적인 모양을 방향성 없이 뒤섞어 이야기하는 경우)	전체 + 부분 + 방향성을 갖춤	전체 + 부분 + 방향성 + α(문단으로서 구조적 특성 및 완결성[중심+뒷받침 문장, 처음-중간-끝 등, 문장 간 응집성])				
- 주제 이탈(아예 묘사가 안 들어간 경우) 경위담화 유형의 구조적 특성도 포함)	전체 + 부분 + 주제 이탈 문장 포함된 경우	전체와 부분을 모두 다룸 (단, 방향성을 발견할 수 없는 경우) ex) 등근 모양 + 눈과 연지근지 하얗고 연지근지가 있다. 웃는 인상이고 눈썹이 진하다. (전체적인 모습과 세부적인 모양을 방향성 없이 뒤섞어 이야기하는 경우)	전체 + 부분 + 방향성을 갖춤	전체 + 부분 + 방향성 + α(문단으로서 구조적 특성 및 완결성[중심+뒷받침 문장, 처음-중간-끝 등, 문장 간 응집성])						
	<유의 사항> * 설명의 방향성: ①, ②가 동시에 나타나야 함. ① 전체적 방향성: 전체→부분 or 부분→전체 ② 부분적 방향성: 위→아래 or 아래→위 * 전체적 방향성과 통일성은 너무 엄격하게 보지 말 것. ex) 전체에서 부분의 방향으로 설명하고, 다시 전체적 인상으로 마무리하는 경우(전체-부분-전체) 좋은 흐름이라고 볼 수 있기 때문에. * 탈에 대한 묘사 후 느낌을 덧붙인 답변의 경우, 이를 정의적 묘사로 보고 감점하지 않음.									
표현	- 표현 항목에서 고려한 대상: 맞춤법, 어휘(문장 부호는 고려하지 않음)									
	<table><tr><th>매우 미흡</th><th>미흡</th><th>보통</th><th>우수</th><th>매우 우수</th></tr><tr><td colspan="3"> •아래 요인이 포함될 때마다 ‘우수’에서 한 단계씩 감점 - 문장의 호응 오류 - 맥락에 부적절한 어휘나 맞춤법 3개 이상 오류(5개 이상 오류 시 두 단계 감점, 동일 유형 오류 시 하나로 산정) - 종결 어미의 오류(문장 미종결, 종결어미 불일치) - 문장 길이의 적정성 오류(문장 하나를 열거형 등으로 지나치게 길게 쓴 경우, 각 문장이 지나치게 짧은 경우) </td><td> (기준점) 우수 주술 호응이 틀린 곳이 없으나, 띄어쓰기 + 맞춤법 1~2개 틀림 </td><td> •아래 요인이 포함될 때 ‘우수’에서 한 단계 가점 - 띄어쓰기 및 맞춤법 정확함 - 객관적인 표현 및 풍부한 어휘 - 수사적 인식(친절한 설명 등) </td></tr></table>	매우 미흡	미흡	보통	우수	매우 우수	•아래 요인이 포함될 때마다 ‘우수’에서 한 단계씩 감점 - 문장의 호응 오류 - 맥락에 부적절한 어휘나 맞춤법 3개 이상 오류(5개 이상 오류 시 두 단계 감점, 동일 유형 오류 시 하나로 산정) - 종결 어미의 오류(문장 미종결, 종결어미 불일치) - 문장 길이의 적정성 오류(문장 하나를 열거형 등으로 지나치게 길게 쓴 경우, 각 문장이 지나치게 짧은 경우)			(기준점) 우수 주술 호응이 틀린 곳이 없으나, 띄어쓰기 + 맞춤법 1~2개 틀림
매우 미흡	미흡	보통	우수	매우 우수						
•아래 요인이 포함될 때마다 ‘우수’에서 한 단계씩 감점 - 문장의 호응 오류 - 맥락에 부적절한 어휘나 맞춤법 3개 이상 오류(5개 이상 오류 시 두 단계 감점, 동일 유형 오류 시 하나로 산정) - 종결 어미의 오류(문장 미종결, 종결어미 불일치) - 문장 길이의 적정성 오류(문장 하나를 열거형 등으로 지나치게 길게 쓴 경우, 각 문장이 지나치게 짧은 경우)			(기준점) 우수 주술 호응이 틀린 곳이 없으나, 띄어쓰기 + 맞춤법 1~2개 틀림	•아래 요인이 포함될 때 ‘우수’에서 한 단계 가점 - 띄어쓰기 및 맞춤법 정확함 - 객관적인 표현 및 풍부한 어휘 - 수사적 인식(친절한 설명 등)						
	- 한 문장이면 매우 미흡 - 띄어쓰기는 감안하지 않으나 오류가 심각한 경우 한 단계 감점									

* ‘내용’과 ‘조직’ 항목에서 모두 보통 이하의 점수를 받으면 ‘표현’ 항목도 보통 이하의 점수를 부여함.(오류 1~2개 보통, 3~4개 미흡, 5개 이상 매우 미흡)

- ‘내용’과 ‘조직’ 항목이 모두 우수나 매우 우수면 표현은 틀린 개수대로 채점함.

* 문제의 정보를 따르지 않은 경우 오답 처리. 오답은 세 항목(내용, 조직, 표현) 모두 매우 미흡으로 처리함.

<표 II -25> 쓰기 2번 문항 세부 채점 기준

평가 항목	세부 채점 기준				
내용	- 내용 항목에서 고려한 대상: 논증의 구성 요소(주장, 이유, 근거, 반론 고려), 타당성				
	매우 미흡	미흡	보통	우수	매우 우수
	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 감점 - 주장이 명확하지 않음. - 이유/근거에서 지칭하는 대상이나 그 의미가 명확하지 않음. - 타당성 없는 이유/근거를 제시함.(다소 엄격하게 채점) - 논점 이탈 내용(논제 및 주장에 부합하지 않는 내용)이 포함됨.	(기준점) 주장 + 이를 뒷받침하는 타당한 이유/근거로 서술	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 가점 - 반론 고려 내용이 포함됨. - 이유/근거가 둘 이상 서술되어 있음.(이유/근거를 한 문장에 짧게 쓰거나 유사한 경우는 이유/근거 1개로 봄) - 이유/근거에 타당한 부연 설명이 있음.(동어 반복은 인정하지 않음)		
- 각각 변별되는 이유/근거가 셋 이상인 경우 매우 우수 - 논점이 완전히 이탈한 경우 매우 미흡 - 150자 미만일 때 한 단계 감점(모든 등급)					
조직	- 조직 항목에서 고려한 대상: 담화 유형의 특성(논증), 문단 구분, 응집성(문장 간 연결)				
	매우 미흡	미흡	보통	우수	매우 우수
	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 감점 문장 간의 연결이 긴밀하지 않음. 주제 이탈 문장이 포함되는 등 글의 완결성이 없는 경우	(기준점) 세 문장으로 서술되었으며, 중심 문장과 뒷받침 문장으로 구성됨. (“찬성합니다.”, “반대합니다.”도 문장 개수에 포함되나, “반대. ~(이유/근거)~” 등과 같은 개조식 서술은 안 됨.)	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 가점 네 문장 이상 서술되었으며 문장 간의 연결이 자연스러움.(네 문장 이상인 것을 판단할 때는 “찬성합니다.”, “반대합니다.”는 제외. 세 문장 이하인 경우는 문장 개수에 포함함.) 서론-본론-결론의 형식을 갖추고, 문단을 적절하게 구분함.		
- 한 문장이면 매우 미흡					
표현	- 표현 항목에서 고려한 대상: 맞춤법, 어휘(문장 부호는 고려하지 않음)				
	매우 미흡	미흡	보통	우수	매우 우수

•아래 요인이 포함될 때마다 ‘우수’에서 한 단계씩 감점 - 문장의 호응 오류 - 맥락에 부적절한 어휘나 맞춤법 3개 이상 오류(5개 이상 오류 시 두 단계 감점, 동일 유형 오류 시 하나로 산정) - 종결 어미의 오류(문장 미종결, 종결어미 불일치) - 문장 길이의 적정성 오류(문장 하나를 열거형 등으로 지나치게 길게 쓴 경우, 각 문장이 지나치게 짧은 경우)	(기준점) 주술 호응이 틀린 곳이 없으나, 띄어쓰기 + 맞춤법 1~2개 틀림	•아래 요인이 포함될 때 ‘우수’에서 한 단계 가점 - 띄어쓰기 및 맞춤법 훌륭함 - 수준 높은 어휘 및 풍부한 표현 - 독자를 고려한 효과적인 표현
- 한 문장이면 매우 미흡 - 띄어쓰기는 감안하지 않으나 오류가 심각한 경우 한 단계 감점		

* ‘내용’과 ‘조직’ 항목에서 모두 보통 이하의 점수를 받으면 ‘표현’ 항목도 보통 이하의 점수를 부여함.(오류 1~2개 보통, 3~4개 미흡, 5개 이상 매우 미흡)

* ‘내용’과 ‘조직’ 항목이 모두 우수나 매우 우수면 표현은 틀린 개수대로 채점함.

* 문제의 정보를 따르지 않은 경우 오답 처리. 오답은 세 항목 모두 매우 미흡으로 처리함.

<표 II -26> 쓰기 3번 문항 세부 채점 기준

평가 항목	세부 채점 기준				
	매우 미흡	미흡	보통	우수	매우 우수
매체 활용	이유가 적절하지 않거나 미진술 시 2점	사진의 특성이 포함되지 않은 미흡한 이유를 제시하였거나 사진과 전달하고 싶은 내용이 조응하지 않는 경우 4점	관련된 이유이지만 다른 사진과 변별되는 특성이 아닌 것을 제시하고, 해당 사진을 통해 전달하고 싶은 내용을 제시한 경우 6점	다른 사진과 변별되는 해당 사진의 특성을 포함하여 타당한 선택의 이유를 제시하고 + 해당 사진을 통해 전달하고 싶은 내용을 적절하게 제시한 경우 (사진의 특성과 중복되어도 무방함) 8점	다른 사진과 변별되는 해당 사진의 특성을 포함하여 타당한 선택의 이유를 제시하고 + 해당 사진을 통해 전달하고 싶은 내용을 구체적으로 제시한 경우 10점
	매우 미흡	미흡	보통	우수	매우 우수
내용	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 감점 - 반품 요청(목적)의 내용이 드러나지 않음. - 주제에서 벗어나는 내용이 포함되거나 통일성이 없음.		- 불량 굴 반품 요청의 표현이 포함됨.	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 가점 - 문제 상황 설명 - 요구 근거, 문제로 인한 손실, 상품 불량으로 인한 유감 표현(항의) - 상황(판매자를 대상으로 한 온라인 게시글 작성)에 알맞은 내용 선정(예: 정	

			중한 인사 또는 판매자의 신뢰도, 단골 고객임 등을 언급(항상 좋은 상품에 감사드립니다), 판매자가 해당 주문을 빨리 확인할 수 있도록 상품 정보나 주문 내역의 요약 등)		
	- 양적인 조건을 충족하였는가: 3문장 미만이면 한 단계 감점 - 판매자를 대상으로 하는 온라인 게시글로서 적절하지 않은 내용이 포함된 경우 한 단계 감점 - 불량 상품 반품 요청과 무관한 내용인 문장이 포함된 경우 한 단계 감점(정중한 인사, 단골 고객임을 언급한 것 등은 관련된 내용으로 판정함) - 타당한 정보를 상상력을 발휘하여 추가한 경우는 인정함. - ‘반품 요청’ 내용이 없지만 가점 요소가 있는 경우 미흡에서부터 가점				
	매우 미흡	미흡	보통	우수	매우 우수
조직	- 1문장(반품 요청 내용 미포함)	- 1문장(‘반품 요청’ 또는 ‘요구 근거(상품 문제 제기)’ 포함의 구성) - 2 문장	- 2문장(‘반품 요청 - 요구 근거(상품 문제 제기)’ 포함의 구성) - 3문장 또는 그 이상	- 3문장 이상일 때, ‘문제 상황 설명 - 요구 근거 또는 문제로 인한 손실 - 문제 해결 요청(해결 방안)’ 등 3가지 이상의 내용으로 구성됨.	3문장 이상일 때, - ‘문제 상황 설명 - 요구 근거 또는 문제로 인한 손실 - 문제 해결 요청(해결 방안)’의 구조가 잘 갖추어졌으면서, 문장의 연결과 흐름이 자연스럽게(제시 순서는 무관하나 나름의 논리적 흐름이 있거나 응집성 장치를 활용함.) 완결성이 있음.
	- 표현 항목에서 고려한 대상: 맞춤법, 어휘(문장 부호는 고려하지 않음)				
	매우 미흡	미흡	보통	우수	매우 우수
표현	•아래 요인이 포함될 때마다 ‘우수’에서 한 단계씩 감점 - 문장의 호응 오류 - 맥락에 부적절한 어휘나 맞춤법 3개 이상 오류(5개 이상 오류 시 두 단계 감점, 동일 유형 오류 시 하나로 산정) - 종결 어미의 오류(문장 미종결, 종결어미 불일치) - 한 문장인 경우			(기준점) 주술 호응이 틀린 곳이 없으나, 띄어쓰기 + 맞춤법 1~2개 틀림	•아래 요인이 포함될 때 ‘우수’에서 한 단계 가점 - 띄어쓰기 및 맞춤법 훌륭함 - 수준 높은 어휘 및 풍부한 표현 - 독자를 고려한 예의 바른 표현이나 효과적인 표현이 사용됨.
	- 띄어쓰기는 감안하지 않으나 오류가 심각한 경우 한 단계 감점				

	- 마침표 미표기는 감점하지 않음. - 한 문장이더라도 틀린 것이 없다면 '보통'으로 판정 - '내용'과 '조직' 항목에서 모두 보통 이하의 점수를 받으면 '표현' 항목도 보통 이하의 점수를 줄 것.(오류 1~2개 보통, 3~4개 미흡, 5개 이상 매우 미흡) - '내용'과 '조직' 항목이 모두 우수나 매우 우수면 표현은 틀린 개수대로 채점함. - 주어나 목적어가 빠지는 등의 심각한 비문이 아니라면 어색한 표현은 어느 정도 허용적으로 채점함.
추가 사항	* 문제의 정보를 따르지 않은 경우 오답 처리. 오답은 세 항목 모두 매우 미흡으로 처리함.

<표 II -27> 쓰기 4번 문항 세부 채점 기준

세부 채점 기준						
평가 항목	채점 요소	매우 미흡	미흡	보통	우수	매우 우수
내용	주제와 목적(사과)에 알맞게 작성되었는가?	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 감점 - 사과의 표현(목적)이 드러나지 않음. - 주제에서 벗어나는 내용이 포함되거나 통일성이 없음.		사과의 표현이 포함됨.	•아래 요인이 포함될 때마다 ‘보통’에서 한 단계씩 가점 - 책임의 인정이나 문제 발생 원인 표현, 또는 잘못에 대한 유감 표현이 포함됨. - 문제 해결이나 태도 개선 의지 표현, 또는 문제 해결 방안이나 대안이 포함됨.	
	양적인 조건을 충족하였는가?	3문장 미만이면 한 단계 감점				
	• 타당한 정보를 상상력을 발휘하여 추가한 경우는 적절한 답으로 간주함. • 사과의 인사가 없어도 가점 요소가 있으면 ‘미흡’에서부터 가점					
조직	담화의 구조적 특성에 적합한가? (‘사과-책임 인정-문제 해결’의 구성이어야 함)	- 1문장(사과, 책임 인정, 문제 해결 등의 내용 미포함)	- 1문장 ‘사과 또는 책임 인정’ 포함의 구성 - ‘사과’를 포함한 2문장의 구성	- 2문장 ‘사과-책임 인정’ 포함의 구성 - ‘사과’를 포함한 3문장 또는 그 이상의 구성	3문장 이상일 때, ‘사과+책임 인정+문제 해결’ 등 3가지 이상의 내용으로 구성됨.	3문장 이상일 때, - ‘사과+책임 인정+문제 해결’ 등 3가지 이상의 내용으로 구성되어 있으면서, 문장의 연결과 흐름이 자연스럽고(나름의 논리적 흐름이 있거나 응집성 장치를 활용함) 완결성이 있음.
표현	정확한 문장과 올바른 맞춤법을	- 표현 항목에서 고려한 대상: 맞춤법, 어휘(문장 부호는 고려하지 않음)				
		매우 미흡	미흡	보통	우수	매우 우수

사용하고 있는가? (띄어쓰기의 경우 관대하게 보고, 어휘, 구어, 맞춤법만 체크함)	•아래 요인이 포함될 때마다 ‘우수’에서 한 단계씩 감점 - 문장의 호응 오류 - 맥락에 부적절한 어휘나 맞춤법 3개 이상 오류(5개 이상 오류 시 두 단계 감점, 동일 유형 오류 시 하나로 산정) - 종결 어미의 오류(문장 미종결, 종결어미 불일치) - 한 문장인 경우	(기준점) 주술 호응이 틀린 곳이 없으나, 띄어쓰기 + 맞춤법 1~2개 틀림	•아래 요인이 포함될 때 ‘우수’에서 한 단계 가점 - 띄어쓰기 및 맞춤법 훌륭함 - 수준 높은 어휘 및 풍부한 표현 - 독자를 고려한 예의 바른 표현이나 효과적인 표현이 사용됨.
- 띄어쓰기는 감안하지 않으나 오류가 심각한 경우 한 단계 감점			

- ‘내용’과 ‘조직’ 항목에서 모두 보통 이하의 점수를 받으면 ‘표현’ 항목도 보통 이하의 점수를 줄 것.(오류 1~2개 보통, 3~4개 미흡, 5개 이상 매우 미흡)

- ‘내용’과 ‘조직’ 항목이 모두 우수나 매우 우수면 표현은 틀린 개수대로 채점함.

- 주어나 목적어가 빠지는 등의 심각한 비문이 아니라면 어색한 표현은 어느 정도 허용적으로 채점함.

* 문제의 정보를 따르지 않은 경우 오답 처리. 오답은 세 항목 모두 매우 미흡으로 처리함.

4) 채점 방법

채점단 협의회 이후, 쓰기 본조사의 채점은 재택 채점 방식으로 진행되었다. 채점자들은 개인별 구글 시트(Google Spreadsheets) 링크에 접속하여 할당된 피험자의 문항별 답안을 확인하고 준거별로 채점을 진행하였다. 채점 화면에는 점수 입력 양식 외에 해당하는 문제와 글, 글자 수 정보를 함께 제공하였다.

2025_KA 쓰기 본 조사_25kr04

파일 수정 보기 삽입 서식 데이터 도구 확장 프로그램 도움말

Q 메뉴 100% W % 123 Arial 10 B I A

A1	A	B	C	D	E	F	G	H	I
1	ID	1	글자수	내용(50%)	조직(30%)	표현(20%)			
2	11	우리나라 전통 탈이머 사람얼굴과비슷하게 생겼고 탈을쓰기도하고 직접만져보면서 체험을해보며 모양에대해서 더 깊게 알수있습니다	68	1	1	1			
3	23	이 탈은 안동화회탈입니다 탈의 종류는 여러가지가 있습니다 여러가지 탈 중에 이 탈은 웃고 있고요 얼굴에 빨간 점이 볼에 2개 이마에 1개 있습니다 화장을 한 느낌으로 보아 살짝 여자의 얼굴을 하고 있는 거 같습니다	119	3	3	3			
4	33	사람얼굴과 비슷하게 생겼고,탈의모습이 편안하고 인자하게 생겼으며 탈에 연지곤지찍은게 독특하며 기억에오래남을것같고 재미있습니다	70	2	2	1			
5	45	얼굴은 가름하고 둥근모양의 게란형이고 눈썹은 길고 알아서 초승달 같습니다. 이것은 각시탈이라고 흔히들 신부의 얼굴을 표현한 탈입니다. 눈은 웃는 모양으로 순하고 부드러운 전통적인 인상을 줍니다.	109	3	3	4			
6	51	이탈은 하회탈이라 부르는 탈로 전국 노래 자랑을 진행하는 개그맨 남희석의 별명이기도 합니다. 눈은 반달 모양으로 웃고 있으며 코는 화살처럼 길쭉하고 이마와 양볼에 연지 곤지가 찍혀 있습니다. 입술도 같은 분홍색이며 입꼬리가 살짝 올라가서 보는 사람도 기분 좋아지는 웃는상 입니다 전체적인 얼굴형은 게란형이고 단정하게 빗어넣긴 헤어스타일을 하고 있습니다	194	5	4	4			
7	57	이 탈은 한국 전통의 탈이라고 합니다. 한국 정서의 맞게 선한 눈매와 얼굴 그리고 선한 모습이 우리나라의 선조들의 얼굴과 비슷함을 느낄수 있습니다. 탈의 모양은 두눈은 동그렇게 웃는 모양이고 눈썹도 눈과 어울리게 순한 모습을 보입니다. 그리고 이마와 두 볼에는 빨간색 동그렇게 점을 찍은 모습입니다. 그리고 머리는 검은색으로 모자를 쓴듯한 모양이고 머리의 길이로 턱선까지 오는 단발 모양의 머리 모양입니다. 그리고 입술은 빨간색으로 웃는 모습을 하고 있어 처음 보는 사람도 무섭지 않고 친근감을 보이는 선한 모습을 하고 있고 턱선도 U자형 턱선으로 아름다운 곡선의 모양을 하고 있습니다..	329	5	3	5			
8	64	하얀 얼굴에 웃는 얼굴의 하회탈입니다. 연지곤지 찍는 모습이 부드럽고 인상적입니다. 게란형 얼굴을 가지고 있다. 옛날 신부의 얼굴을 연상 시킨다. 환하게 웃는 모습이 행복해 보인다.	100	3	3	3			

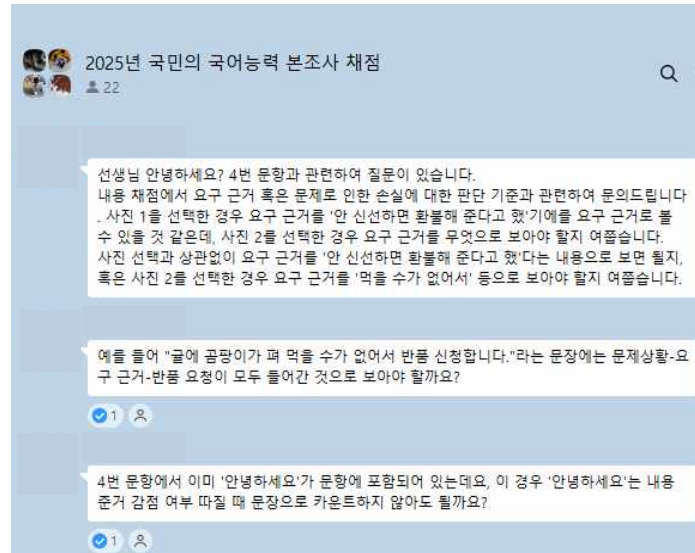
평가 문항

평가 항목

내용: (50%)

[그림 II-14] 채점 결과 입력 화면

채점 진행 중에는 채점자들이 발견한 특이한 글 사례나 채점 척도 적용 과정에서의 어려움과 같은 문제가 있는 경우 연구진과 소통할 수 있도록 하였으며, 그 결과를 채점단과 즉각적으로 공유하기 위하여 모바일 메신저(카카오톡)를 활용하였다.



[그림 II-15] 채점단 모바일 메신저 화면

본조사의 채점은 8월 19일부터 9월 5일까지 17일간 진행되었다. 채점을 종료한 후에는 채점 자료에 결측이나 다른 오류가 없는지를 확인하는 검수 과정을 진행한 후 본조사 채점 결과를 확정하였다.

5) 채점자 신뢰도 분석

쓰기 영역 4개 문항에 대한 채점 결과, 총점 기준으로 채점자 간 신뢰도²⁾가 다음과 같이 산출되었다. 1번 문항에서는 $x=0.89$, $r=0.90$, 2번 문항에서는 $x=0.77$, $r=0.77$, 3번 문항에서는 $x=0.80$, $r=0.81$, 4번 문항에서는 $x=0.62$, $r=0.64$ 로 채점자 간 채점 일치 정도가 모두 높은 수준이었다.³⁾

2) 본 연구에서는 대규모 쓰기 채점 상황과 채점 공정성 확보, 채점 설계의 경제성을 고려하여 채점자 훈련이나 선정 연구에 주로 활용되는 채점자 내 신뢰도가 아니라 채점자 간 신뢰도를 채점자 신뢰도로 보고하였다. 이는 채점단이 모두 채점자 훈련에 참여한 점과 세부 채점 기준을 제공한 점, 채점 문제 상황에 대한 모니터링과 연구진 질의가 가능했던 점을 고려한 것이다.

3) Cohen의 Kappa 계수 해석 기준(Landis & Koch, 1977)에 따르면, $0.41 \leq x \leq 0.60$ 이면 ‘중간 정도의(moderate)’, $0.61 \leq x \leq 0.80$ 이면 ‘상당히 높은(substantial)’, $0.81 \leq x \leq 1.00$ 이면 ‘거의 완벽한(almost perfect)’ 신뢰도에 해당한다. Pearson의 적률상관계수를 해석할 때에는 관행적으로 0.5~0.7이면 ‘뚜렷한(large/strong)’ 상관, 0.7~0.9이면 ‘매우 강한(very strong)’ 상관, 0.9~1.0이면 ‘거의 완벽한(nearly perfect)’ 상관이라고 말한다.

<표 II-28> 쓰기 영역 문항별 채점자 간 신뢰도(총점 기준)

문항	N	Cohen의 Kappa 계수(κ)	Pearson의 적률상관계수(r)
쓰기 1번	3,050	0.89	0.90
쓰기 2번	3,050	0.77	0.77
쓰기 3번	3,050	0.80	0.81
쓰기 4번	3,050	0.62	0.64

쓰기 영역 4개 문항 각각에서 채점 기준별로는 채점자 간 신뢰도가 다음과 같이 산출되었다. 1번 문항에서는 내용에서 $\kappa=0.87$, $r=0.87$, 조직에서 $\kappa=0.78$, $r=0.78$, 표현에서 $\kappa=0.72$, $r=0.72$ 로 채점자 간 일치 정도가 높은 수준이었다. 2번 문항에서는 내용에서 $\kappa=0.68$, $r=0.69$, 조직에서 $\kappa=0.73$, $r=0.73$, 표현에서 $\kappa=0.63$, $r=0.63$ 으로 채점자 간 일치 정도가 높은 수준이었다. 3번 문항에서는 내용에서 $\kappa=0.78$, $r=0.78$, 조직에서 $\kappa=0.70$, $r=0.73$, 표현에서 $\kappa=0.58$, $r=0.59$ 로 채점자 간 일치 정도가 수용 가능한 수준이었다. 4번 문항에서는 매체에서 $\kappa=0.56$, $r=0.57$, 내용에서 $\kappa=0.55$, $r=0.56$, 조직에서 $\kappa=0.57$, $r=0.60$, 표현에서 $\kappa=0.51$, $r=0.51$ 로 채점자 간 일치 정도가 수용 가능한 수준이었다.

<표 II-29> 쓰기 영역 문항별*채점 기준별 채점자 간 신뢰도

문항	채점 기준	N	Cohen의 Kappa 계수(κ)	Pearson의 적률상관계수(r)
쓰기 1번	내용	3,050	0.87	0.87
	조직	3,050	0.78	0.78
	표현	3,050	0.72	0.72
쓰기 2번	내용	3,050	0.68	0.69
	조직	3,050	0.73	0.73
	표현	3,050	0.63	0.63
쓰기 3번	내용	3,050	0.78	0.78
	조직	3,050	0.70	0.73
	표현	3,050	0.58	0.59
쓰기 4번	매체	3,050	0.56	0.57
	내용	3,050	0.55	0.56
	조직	3,050	0.57	0.60
	표현	3,050	0.51	0.51

이상에서와 같이, 쓰기 영역에서는 각 문항에서의 채점자 간 신뢰도가 총점 기준으로는 모두 높은 수준이었고, 세부 채점 기준별로도 모두 수용 가능한 수준이었으므로, 이러한 채점 결과를 토대로 후속 분석을 수행하는 데 문제가 없다고 판단하였다.

2.5. 본조사 결과 분석

1) 방문 면접 조사 결과

(1) 쓰기 능력 조사 결과 분석

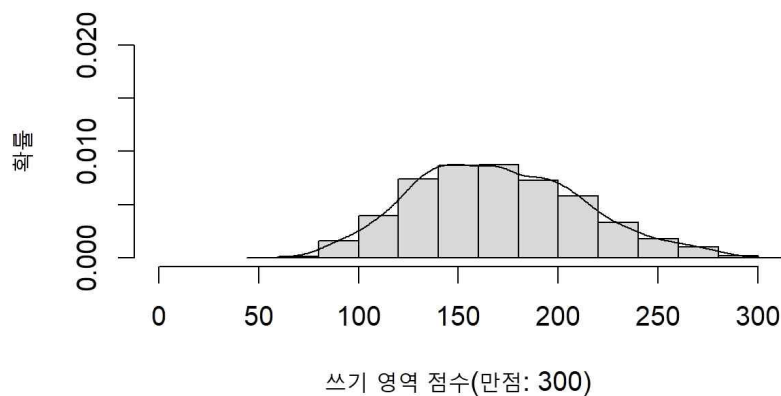
쓰기 영역의 원점수 요약 즉, 평균 및 표준편차, 범위(최솟값 및 최댓값) 등의 정보는 다음과 같다.

<표 II-30> 영역별 원점수 요약(쓰기)

영역	N	만점	평균	표준편차	최솟값	최댓값
쓰기	3,000	300	169.99	42.11	66.80	295.40

3,000명의 쓰기 영역 점수(만점 300점) 중 가장 낮은 점수는 66.80점, 가장 높은 점수는 295.40점이었고, 평균은 169.99점, 표준편차는 42.11로 분석되었다.

쓰기 영역의 문항 특성을 분석한 결과는 다음과 같다.



[그림 II-16] 영역별 원점수 요약(쓰기)

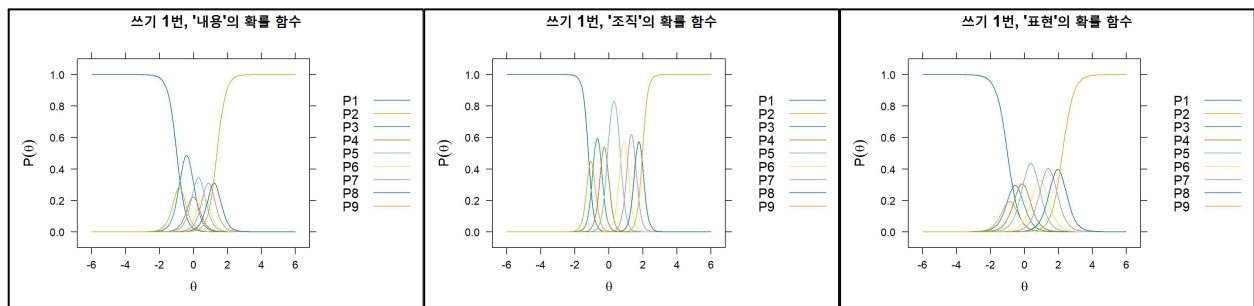
<표 II-31> 쓰기 영역 문항 특성 분석 결과

문항 번호	N	CTT		IRT				신뢰도	
		변별도	난이도	변별도	SE	난이도	SE	제거	포함
쓰기 1번	3,000	.75	.54	1.44	.06	.56	.07	.61	.67
쓰기 2번	3,000	.77	.54	2.30	.12	.38	.05	.53	
쓰기 3번	3,000	.66	.56	.94	.05	.47	.13	.65	
쓰기 4번	3,000	.66	.63	1.02	.05	-1.12	.19	.63	

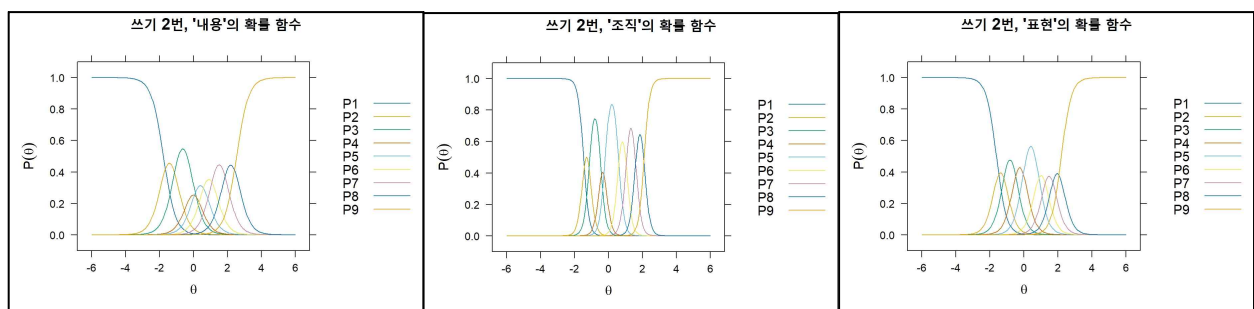
먼저, 고전 검사 이론(CTT)에 따른 문항별 변별도는 전체 문항이 0.66~0.77로 나타나, 모든 문항의 변별도가 '높음'(0.4 이상)인 것으로 확인되었다. 고전 검사 이론(CTT)에 따른 문항별 난이도(용이도)는 '중간'(0.4 이상 0.6 미만)인 문항이 3개(1번, 2번, 3번), '쉬움'(0.6 이상 0.8 미만)인 문항이 1개(4번)로 나타났다.

다음으로, 문항 반응 이론(IRT)에 따른 문항별 변별도는 '거의 없음'(0.35 미만)이나 '낮음'(0.35 이상 0.65 미만)에 해당하는 문항은 없었고, '매우 높음'(1.70 이상)은 1개(2번), '높음'(1.35 이상 1.70 미만)은 1개(1번), '적절함'(0.65 이상 1.35 미만)은 2개(3번, 4번)로 확인되었다. 문항 반응 이론(IRT)에 따른 문항별 난이도(곤란도)는 '쉬움'(-2.0 이상 -0.5 미만)인 문항이 1개(4번), '중간'(-0.5 이상 0.5 미만)인 문항이 2개(2번, 3번), '어려움'(0.5 이상 2.0 미만)인 문항이 1개(1번)로 나타났다. 이는 성인 전체를 대상으로 하는 국민의 국어능력을 평가하기에 적절한 문항 수준으로 볼 수 있다.

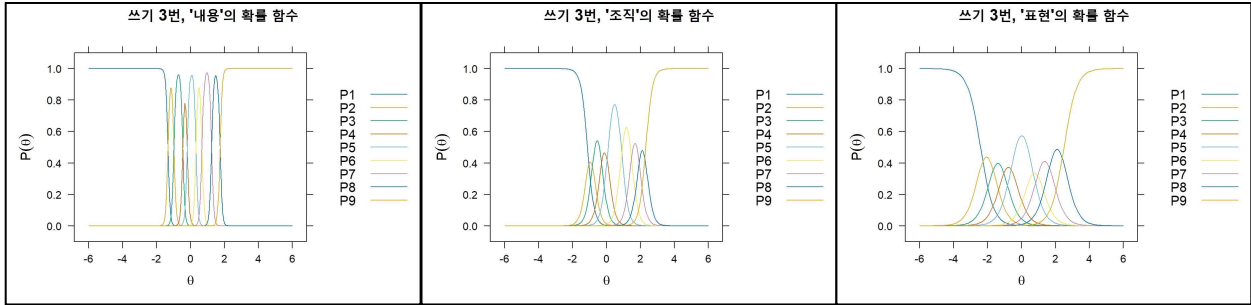
이상에서 설명한, 문항 반응 이론(IRT)에 따른 쓰기 영역 문항 특성의 분석 결과를 아래의 문항 특성 곡선(ICC)을 통해 시각적으로 확인할 수 있다.



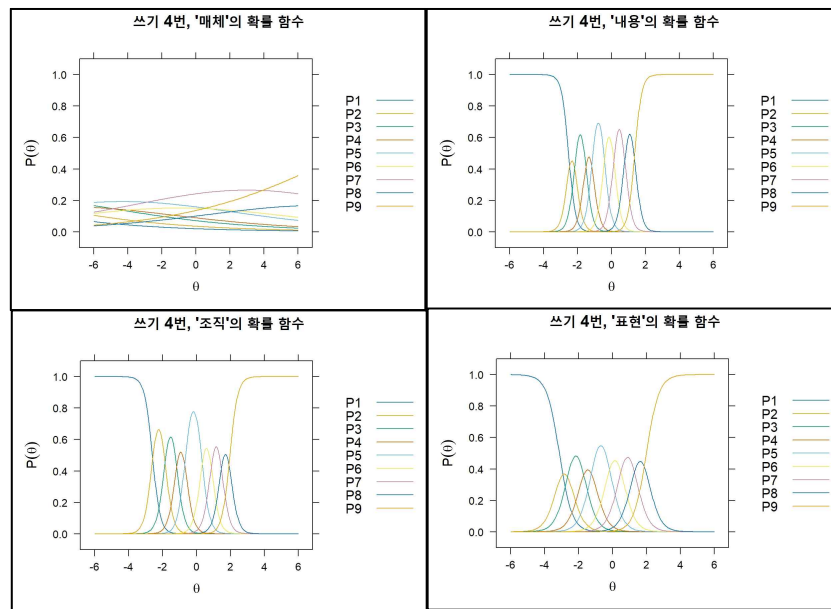
[그림 II-17] 쓰기 영역 1번 문항의 ICC



[그림 II-18] 쓰기 영역 2번 문항의 ICC



[그림 II-19] 쓰기 영역 3번 문항의 ICC



[그림 II-20] 쓰기 영역 4번 문항의 ICC

(2) 쓰기 영역 문항 카드

2025 국민의 국어능력 실태 조사	W-1	문항 정보	설명
평가 문항	 		

자료 출처																																																				
평가 요소 텍스트 정보	설명적 글쓰기 - 묘사글																																																			
예시 답안	평가 항목		채점 요소(2018 채점 요소 일부 정교화)			채점 척도																																														
						매우 미흡	미흡	보통	우수	매우 우수																																										
	내용: 40점 (50%)		- 설명 대상의 전체와 부분을 잘 이해할 수 있는 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.			8	16	24	32	40																																										
	조직: 24점 (30%)		- 대상에 대한 독자의 이해를 도울 수 있는 내용 전개 방식(묘사, 비교, 대조, 분석 등)을 사용하였다. - 내용의 배열이 체계적(설명의 방향성을 가지고 차례대로 설명)이며, 응집성과 완결성을 갖추었다.			5	10	14	19	24																																										
	표현: 16점 (20%)		- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 객관적이고 명확한 표현을 사용하였다.			3	6	10	13	16																																										
본조사 결과	CTT 분석 결과		변별도(a): 0.75					난이도(용이도; b): 0.54																																												
			<table><tr><td>없음</td><td>매우 낮음</td><td>낮음</td><td>있음</td><td>높음</td></tr><tr><td>0.0</td><td>0.1</td><td>0.2</td><td>0.3</td><td>0.4</td></tr><tr><td>$\leq a <$</td><td>$\leq a <$</td><td>$\leq a <$</td><td>$\leq a <$</td><td>$\leq a \leq$</td></tr><tr><td>0.1</td><td>0.2</td><td>0.3</td><td>0.4</td><td>1.0</td></tr></table>					없음	매우 낮음	낮음	있음	높음	0.0	0.1	0.2	0.3	0.4	$\leq a <$	$\leq a <$	$\leq a <$	$\leq a <$	$\leq a \leq$	0.1	0.2	0.3	0.4	1.0	<table><tr><td>매우 어려움</td><td>어려움</td><td>중간</td><td>쉬움</td><td>매우 쉬움</td></tr><tr><td>0.0</td><td>0.2</td><td>0.4</td><td>0.6</td><td>0.8</td></tr><tr><td>$\leq b <$</td><td>$\leq b <$</td><td>$\leq b <$</td><td>$\leq b <$</td><td>$\leq b \leq$</td></tr><tr><td>0.2</td><td>0.4</td><td>0.6</td><td>0.8</td><td>1.0</td></tr></table>					매우 어려움	어려움	중간	쉬움	매우 쉬움	0.0	0.2	0.4	0.6	0.8	$\leq b <$	$\leq b <$	$\leq b <$	$\leq b <$	$\leq b \leq$	0.2	0.4	0.6	0.8	1.0
			없음	매우 낮음	낮음	있음	높음																																													
			0.0	0.1	0.2	0.3	0.4																																													
	$\leq a <$	$\leq a <$	$\leq a <$	$\leq a <$	$\leq a \leq$																																															
0.1	0.2	0.3	0.4	1.0																																																
매우 어려움	어려움	중간	쉬움	매우 쉬움																																																
0.0	0.2	0.4	0.6	0.8																																																
$\leq b <$	$\leq b <$	$\leq b <$	$\leq b <$	$\leq b \leq$																																																
0.2	0.4	0.6	0.8	1.0																																																
IRT 분석 결과		변별도(a): 1.44					난이도(곤란도; β): 0.56																																													
		<table><tr><td>거의 없음</td><td>낮음</td><td>적절함</td><td>높음</td><td>매우 높음</td></tr><tr><td>0.00</td><td>0.35</td><td>0.65</td><td>1.35</td><td>1.70</td></tr><tr><td>$< a <$</td><td>$\leq a <$</td><td>$\leq a <$</td><td>$\leq a <$</td><td>$\leq a$</td></tr><tr><td>0.35</td><td>0.65</td><td>1.35</td><td>1.70</td><td></td></tr></table>					거의 없음	낮음	적절함	높음	매우 높음	0.00	0.35	0.65	1.35	1.70	$< a <$	$\leq a <$	$\leq a <$	$\leq a <$	$\leq a$	0.35	0.65	1.35	1.70		<table><tr><td>매우 쉬움</td><td>쉬움</td><td>중간</td><td>어려움</td><td>매우 어려움</td></tr><tr><td>$\beta <$</td><td>-2.0</td><td>-0.5</td><td>0.5</td><td>2.0</td></tr><tr><td>-2.0</td><td>$\leq \beta <$</td><td>$\leq \beta <$</td><td>$\leq \beta <$</td><td>$\leq \beta$</td></tr><tr><td></td><td>-0.5</td><td>0.5</td><td>2.0</td><td></td></tr></table>					매우 쉬움	쉬움	중간	어려움	매우 어려움	$\beta <$	-2.0	-0.5	0.5	2.0	-2.0	$\leq \beta <$	$\leq \beta <$	$\leq \beta <$	$\leq \beta$		-0.5	0.5	2.0		
		거의 없음	낮음	적절함	높음	매우 높음																																														
0.00	0.35	0.65	1.35	1.70																																																
$< a <$	$\leq a <$	$\leq a <$	$\leq a <$	$\leq a$																																																
0.35	0.65	1.35	1.70																																																	
매우 쉬움	쉬움	중간	어려움	매우 어려움																																																
$\beta <$	-2.0	-0.5	0.5	2.0																																																
-2.0	$\leq \beta <$	$\leq \beta <$	$\leq \beta <$	$\leq \beta$																																																
	-0.5	0.5	2.0																																																	

2025 국민의 국어능력 실태 조사		W-2		문항 정보		논증	
평가 문항							
자료 출처							
평가 요소 텍스트 정보	논증적 글쓰기 - 논설문 쓰기						
채점 기준 예시 답안	평가 항목	채점 요소	채점 척도				
			매우 미흡	미흡	보통	우수	매우 우수
	내용: 65점 (50%)	- 주장을 분명히 제시하고, 타당하고 충분한 이유나 근거를 들었다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	13	26	39	52	65
	조직: 39점 (30%)	- 서론-본론-결론의 형식을 갖추고, 문단을 적절하게 구분하였다. - 설득력을 높일 수 있는 내용 전개 방식(문제-해결, 인과, 유추, 분석 등)을 사용하였다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	8	16	23	31	39
	표현: 26점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려한 효과적인 표현을 사용하였다.	5	10	16	21	26

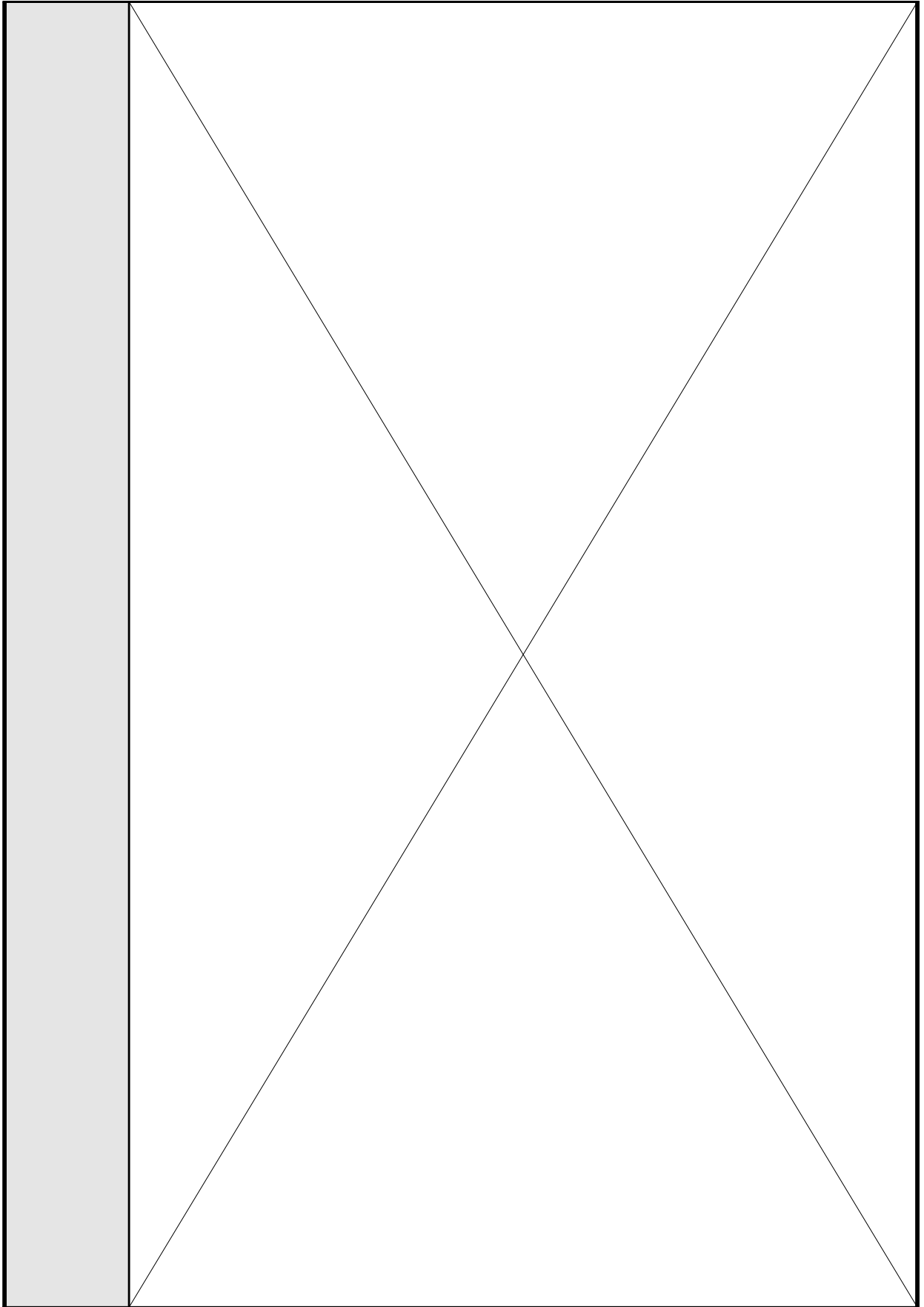
본조사 결과	CTT 분석 결과	변별도(a): 0.77					난이도(용이도; b): 0.54				
		없음	매우 낮음	낮음	있음	높음	매우 어려움	어려움	중간	쉬움	매우 쉬움
		0.0 $\leq a <$ 0.1	0.1 $\leq a <$ 0.2	0.2 $\leq a <$ 0.3	0.3 $\leq a <$ 0.4	0.4 $\leq a \leq$ 1.0	0.0 $\leq b <$ 0.2	0.2 $\leq b <$ 0.4	0.4 $\leq b <$ 0.6	0.6 $\leq b <$ 0.8	0.8 $\leq b \leq$ 1.0
	IRT 분석 결과	변별도(a): 2.30					난이도(곤란도; β): 0.38				
		거의 없음	낮음	적절함	높음	매우 높음	매우 쉬움	쉬움	중간	어려움	매우 어려움
		0.00 $< a <$ 0.35	0.35 $\leq a <$ 0.65	0.65 $\leq a <$ 1.35	1.35 $\leq a <$ 1.70	1.70 $\leq a$	$\beta <$ -2.0	-2.0 $\leq \beta <$ -0.5	-0.5 $\leq \beta <$ 0.5	0.5 $\leq \beta <$ 2.0	2.0 $\leq \beta$

2025 국민의 국어능력 실태 조사		W-3	문항 정보	천교
평가 문항				
자료 출처				

평가 요소 텍스트 정보		친교적 글쓰기 - 사과문					
채점 기준 예시 답안	평가 항목	채점 요소	채점 척도				
			매우 미흡	미흡	보통	우수	매우 우수
	내용: 20점 (50%)	- 주제와 목적(잘못된 공지에 대한 사과)에 부합하고, 상황(동호회 회원을 대상으로 한 공적인 게시글 작성)에 알맞은 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	4	8	12	16	20
	조직: 12점 (30%)	- ‘문제 상황(사과 내용) 설명과 사과 표현 - 책임의 인정이나 문제 발생 원인 제시- 문제 해결이나 태도 개선 의지 표현/문제 해결 방안이나 대안의 제시’의 구조가 잘 갖추어졌다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	2	5	7	10	12

	<table><tr><td>표현: 8점 (20%)</td><td>- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 예의 바르고 효과적인 표현을 활용하였다.</td><td>2</td><td>3</td><td>5</td><td>6</td><td>8</td></tr></table>	표현: 8점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 예의 바르고 효과적인 표현을 활용하였다.	2	3	5	6	8			
표현: 8점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 예의 바르고 효과적인 표현을 활용하였다.	2	3	5	6	8					
본조사 결과	CTT 분석 결과	변별도(a): 0.66					난이도(용이도; b): 0.56				
		없음	매우 낮음	낮음	있음	높음	매우 어려움	어려움	중간	쉬움	매우 쉬움
	0.0 $\leq a <$ 0.1	0.1 $\leq a <$ 0.2	0.2 $\leq a <$ 0.3	0.3 $\leq a <$ 0.4	0.4 $\leq a \leq$ 1.0	0.0 $\leq b <$ 0.2	0.2 $\leq b <$ 0.4	0.4 $\leq b <$ 0.6	0.6 $\leq b <$ 0.8	0.8 $\leq b \leq$ 1.0	
	IRT 분석 결과	변별도(a): 0.94					난이도(곤란도; β): 0.47				
거의 없음		낮음	적절함	높음	매우 높음	매우 쉬움	쉬움	중간	어려움	매우 어려움	
		0.00 $< a <$ 0.35	0.35 $\leq a <$ 0.65	0.65 $\leq a <$ 1.35	1.35 $\leq a <$ 1.70	1.70 $\leq a$	$\beta <$ -2.0	-2.0 $\leq \beta <$ -0.5	-0.5 $\leq \beta <$ 0.5	0.5 $\leq \beta <$ 2.0	2.0 $\leq \beta$

2025 국민의 국어능력 실태 조사	W-4	문항 정보	디지털
평가 문항			



자료 출처																																									
평가 요소 텍스트 정보	온라인 게시판 글 쓰기 - 반품 요청하기																																								
채점 기준 예시 답안	<table><tr><th rowspan="2">평가 항목</th><th rowspan="2">채점 요소</th><th colspan="5">채점 척도</th></tr><tr><th>매우 미흡</th><th>미흡</th><th>보통</th><th>우수</th><th>매우 우수</th></tr><tr><td>매체 활용: 10점 (20%)</td><td>- 표현 의도를 고려하여 적절하게 복합양식적 특성(이미지 및 텍스트 항목)을 활용하였다.</td><td>2</td><td>4</td><td>6</td><td>8</td><td>10</td></tr><tr><td>내용: 20점 (40%)</td><td>- 주제와 목적(불량 상품 반품 요청)에 부합하고, 상황(판매자를 대상으로 한 온라인 게시글 작성)에 알맞은 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.</td><td>4</td><td>8</td><td>12</td><td>16</td><td>20</td></tr><tr><td>조직: 10점 (20%)</td><td>- '문제 상황 설명 - 요구 근거 또는 문제로 인한 손실 - 문제 해결 요청(해결 방안)'의 구조가 잘 갖추어졌다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.</td><td>2</td><td>4</td><td>6</td><td>8</td><td>10</td></tr><tr><td>표현: 10점 (20%)</td><td>- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 정중하고 명확한 표현을 활용하였다.</td><td>2</td><td>4</td><td>6</td><td>8</td><td>10</td></tr></table>	평가 항목	채점 요소	채점 척도					매우 미흡	미흡	보통	우수	매우 우수	매체 활용: 10점 (20%)	- 표현 의도를 고려하여 적절하게 복합양식적 특성(이미지 및 텍스트 항목)을 활용하였다.	2	4	6	8	10	내용: 20점 (40%)	- 주제와 목적(불량 상품 반품 요청)에 부합하고, 상황(판매자를 대상으로 한 온라인 게시글 작성)에 알맞은 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	4	8	12	16	20	조직: 10점 (20%)	- '문제 상황 설명 - 요구 근거 또는 문제로 인한 손실 - 문제 해결 요청(해결 방안)'의 구조가 잘 갖추어졌다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	2	4	6	8	10	표현: 10점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 정중하고 명확한 표현을 활용하였다.	2	4	6	8	10
	평가 항목			채점 요소	채점 척도																																				
		매우 미흡	미흡		보통	우수	매우 우수																																		
	매체 활용: 10점 (20%)	- 표현 의도를 고려하여 적절하게 복합양식적 특성(이미지 및 텍스트 항목)을 활용하였다.	2	4	6	8	10																																		
	내용: 20점 (40%)	- 주제와 목적(불량 상품 반품 요청)에 부합하고, 상황(판매자를 대상으로 한 온라인 게시글 작성)에 알맞은 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	4	8	12	16	20																																		
	조직: 10점 (20%)	- '문제 상황 설명 - 요구 근거 또는 문제로 인한 손실 - 문제 해결 요청(해결 방안)'의 구조가 잘 갖추어졌다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	2	4	6	8	10																																		
표현: 10점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 정중하고 명확한 표현을 활용하였다.	2	4	6	8	10																																			

본조사 결과	CTT 분석 결과	변별도(a): 0.66					난이도(용이도: b): 0.63				
		없음	매우 낮음	낮음	있음	높음	매우 어려움	어려움	중간	쉬움	매우 쉬움
		0.0	0.1	0.2	0.3	0.4	0.0	0.2	0.4	0.6	0.8
		$\leq a <$ 0.1	$\leq a <$ 0.2	$\leq a <$ 0.3	$\leq a <$ 0.4	$\leq a \leq$ 1.0	$\leq b <$ 0.2	$\leq b <$ 0.4	$\leq b <$ 0.6	$\leq b <$ 0.8	$\leq b \leq$ 1.0
	IRT 분석 결과	변별도(a): 1.02					난이도(곤란도: β): -1.12				
거의 없음		낮음	적절함	높음	매우 높음	매우 쉬움	쉬움	중간	어려움	매우 어려움	
0.00		0.35	0.65	1.35	1.70 $\leq a$	$\beta <$ -2.0	-2.0	-0.5	0.5	2.0	
$< a <$ 0.35		$\leq a <$ 0.65	$\leq a <$ 1.35	$\leq a <$ 1.70		$\leq \beta <$ -0.5	$\leq \beta <$ 0.5	$\leq \beta <$ 2.0	$\leq \beta$		

(3) 배경 변인에 따른 쓰기 능력의 차이 분석

본 항에서는 배경 변인에 따른 국민의 쓰기 능력의 차이 분석 결과를 다룬다. 배경 변인에 따른 쓰기 능력 수준의 차이를, 문항 반응 이론 분석에 따라 추정된 ‘능력 점수(T점수)’에 대한 차이 검정을 통해 살펴보고자 한다.

쓰기 영역에서의 성별, 연령대, 지역 규모, 직업군, 학력과 같은 배경 변인에 따른 차이 분석 결과는 다음과 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수는 남성(50.03)이 여성(49.97)보다 근소하게 높았으나 통계적으로 유의한 차이는 아닌 것으로 확인되었다($p=.84$).

<표 II -32> 쓰기 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성		t	df	p
	평균	SD	평균	SD			
성별	50.03	8.71	49.97	8.49	.20	2,998	.84

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 30대(51.20), 20대(50.97), 40대(50.70), 50대(49.73), 60대(47.84) 순으로 높게 나타나 연령대가 높을수록 능력 점수는 낮아지는 경향성이 보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 60대, 30대와 50대, 30대와 60대, 40대와 60대, 50대와 60대 간 대응별 차이가 유의한 것으로 분석되었다.

<표 II -33> 쓰기 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		F	df1	df2	p
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	50.97	8.42	51.20	9.09	50.70	8.40	49.73	8.65	47.84	8.03	15.69	4	2,995	<.001

다음으로, 거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 대도시(50.57), 중소 도시(49.91), 읍면 지역(48.64)의 순으로 높게 나타났고, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 읍면 지역, 중소 도시와 읍면 지역 간 대응별 차이가 유의한 것으로 분석되었다.

<표 II -34> 쓰기 영역 능력 점수의 지역 규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역					
	평균	SD	평균	SD	평균	SD	F	df1	df2	p
지역 규모	50.57	7.67	49.91	9.41	48.64	8.69	8.80	2	2,997	<.001

직업군에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 정신노동(51.36), 기타(49.83), 육체노동(49.14)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타 간 차이가 유의한 것으로 분석되었다.

<표 II -35> 쓰기 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타					
	평균	SD	평균	SD	평균	SD	F	df1	df2	p
직업군	51.36	8.97	49.14	8.24	49.83	8.55	16.95	2	2,997	<.001

끝으로, 학력에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 대재 이상(51.26), 고졸(48.84), 고졸 미만(46.51)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다.

<표 II-36> 쓰기 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	46.51	7.38	48.84	8.22	51.26	8.76	52.18	2	2,997	<.001

2) 조사 방식에 따른 결과 비교

쓰기 영역에서는 조사 방식(방문 면접 조사 혹은 집합 조사)에 따른 피조사자의 능력 점수(원점수) 차이가 $\alpha < .05$ 수준에서 통계적으로 유의하였다($p < .01$). 집합 조사 참여자의 평균이 방문 면접 조사 참여자보다 높은 것을 확인할 수 있었다. 그 효과 크기 cohen's $d=1.03$ 으로 큰 효과 크기(>0.8)에 해당함을 확인할 수 있었다.⁴⁾

<표 II-37> 조사 방식에 따른 쓰기 영역 능력 점수의 차이

능력 점수(원점수)						t-검정 결과			
방문 면접 조사			집합 조사			t	df	p	
N	평균	SD	N	평균	SD				
3,000	169.99	42.11	50	213.33	48.66	-7.20	3,048	<.01	

한편, 이 조사에서는 태블릿 피시 사용이 어려운 피험자에 한하여 종이 인쇄물에 응답할 수 있도록 하였다. 그러나 종으로 응답한 피험자의 수가 방문 면접 조사의 경우 3,000명 중 15명, 집합 조사의 경우 50명 중 0명에 불과하였다. 따라서 조사 매체에 따른 차이를 분석하기에는 무리가 있다.

4) Cohen(1988)에서는 효과 크기(d)가 0.2이면 작은(small) 효과 크기, 0.5이면 중간(medium) 효과 크기, 0.8이면 큰(large) 효과 크기라고 제안한다.

2.6. 외국의 자국어 능력 조사와의 비교 분석

언어 능력은 각 개인이 일상생활을 영위하고 타인과 소통하며 긍정적인 관계를 맺는 것을 넘어, 한 시민으로서 가져야 할 핵심 역량의 집합이며, 나아가 국가 경쟁력의 지표라고도 할 수 있다. 이에 따라 OECD, IEA 등 국제기구에는 PISA, PIAAC, PIRLS 등 다양한 조사 도구를 이용하여 국가 간 리터러시 수준을 측정해 왔다. 그러나 이들 평가는 공통적으로 읽기 중심의 수용적 언어 능력을 측정하는 데 초점을 두고 있으며, 표현적 언어 능력과 관련한 쓰기 능력은 상대적으로 간과되어 왔다. 이 과정에서 국제적 리터러시 조사는 정보의 수용 능력은 정밀하게 진단해 왔지만, 언어를 활용하여 사고를 조직, 표현하는 능력은 충분히 다루어 오지 못했다고 평가할 수 있다.

이와 달리, 일부 국가에서는 자국민의 국어능력을 진단하기 위해 쓰기 능력을 평가하는 진단 도구를 발전시켰다. 대표적인 사례로 미국의 NAEP(National Assessment of Educational Progress), 호주의 NAPLAN(National Assessment Program - Literacy and Numeracy), 그리고 노르웨이의 NSBWT(The Norwegian National Sample-Based Writing Test)를 살펴볼 수 있다. 세 평가 모두 국가적 차원에서 쓰기 능력을 표준화된 방식으로 평가하고 있다. NAEP는 설득, 설명, 서사(경험)의 세 장르를 설정하여 글쓰기 능력을 평가하였으며, NAPLAN은 서사와 설득 중심, NSBWT는 설득, 설명, 서사(상상)의 장르를 중심으로 쓰기 능력을 진단하였다.

국제적 흐름을 볼 때, 이들 평가들은 단순한 기능적 쓰기를 넘어 사회적 소통 행위로서의 쓰기 수행 능력을 평가하려고 시도하였다는 공통점이 있다. 독자를 고려하여 자신이 표현하고자 하는 내용을 선정하고, 공동체의 관습을 고려하여 조직함으로써 적절하게 소통하는 실제적 쓰기 능력을 평가하고자 한 것이다. 본 연구의 <국민의 국어능력 실태 조사>(NAKLA)는 이러한 세계적 흐름을 따르면서도 문식 환경의 실제성 및 변화 양상을 고려하여, 설명적·논증적·친교적·디지털 매체 글쓰기라는 네 개 영역을 중심으로 국민의 실제 언어 사용 능력을 체계적으로 진단하려고 시도하였다.

특히 최근의 국제 문식성 평가는 급격히 변화하는 디지털 환경을 반영하기 위해 변화하고 있다. OECD의 PISA는 2015년부터 기존 종이 기반 읽기 평가에서 벗어나 컴퓨터 기반 역량 평가를 실시하고 있으며(OECD, n.d.), 미국의 NAEP 쓰기 역시 컴퓨터 기반 평가로 전환하였다(NCES, n.d.-a). 이와 동시에 디지털 매체 텍스트의 특징인 복합양식성을 반영해야 한다는 문제의식도 커졌는데, 일례로 Unsworth et al. (2019)에서는 호주의 NAPLAN에서 단일 양식 장르만을 다루고 있다는 점을 지적하며, TIMSS나 PISA 등의 국제적 추세를 따라 복합양식 텍스트의 평가를 보다 적극적으로 포함해야 한다고 제안한 바 있다. 국민의 국어능력 실태 조사는 이러한 점을 반영하여 디지털 매체를 활용한 쓰기, 사진을 포함한 복합양식 자료 활용, 온라인 맥락에서의 책임 있는 소통 능력을 포함하는 통합적 역량을 측정하고자 하였다.

우선 평가 영역 및 대상 측면에서 외국의 자국어 능력 조사와 비교하여 분석하자면 다음과 같다. NAPLAN, NAEP, NSBWT는 모두 학교급 내의 학습자를 대상으로 시행하고 있으며, 평가 영역은 주로 설명, 서사, 설득의 세 장르를 중심으로 구성되어 있다. NAPLAN은 초·중등 학습자(3, 5, 7, 9학년)를 대상으로 서사와 설득 장르의 쓰기를 평가하고 있으며, NAEP는 4, 8, 12학년 학습자를 대상으로 설명, 설득, 서사 장르를 모두 평가하되 서사 중에서는 경험(experience)에 관한 쓰기 능력을 중심으로 평가하고 있다. NSBWT에서는 5, 8학년 학습자를 대상으로 하여 설득, 설명, 서사 장르를 모두 평가하고 있으며 서사는 상상(imaginary)에 관한 서술을 중심으로 평가하고 있다. 이들 평가는 학교 교육과정을 중심으로 학습자의 쓰기 능력을 진단하는 데 초점을 두고 있다.

본 연구에 해당하는 NAKLA과 외국의 자국어 능력 조사의 차이점은 크게 세 가지이다. 먼저 서사 장

를 대신하여 친교적 쓰기 장르를 채택하였다. 이는 일상적인 소통 맥락에서 빈번하게 이루어지는 쓰기 수행 상황을 포괄함으로써 보다 실제적인 언어 능력을 평가한다는 점에서 의미가 있다. 나아가 평가 영역에서 디지털 매체 글쓰기를 별도의 영역으로 평가함으로써 쓰기 상황의 실제성을 높이고, 대규모 쓰기 평가에서 선도적인 구성을 취하였다. 또한 NAKLA의 대상은 만 20세 이상 69세 이하의 일반 국민으로, 특정 학년이 아닌 국민 전체의 언어 사용 능력을 평가한다는 점에서 차별적이다. 요약하면, NAKLA는 학령 인구 중심의 리터러시 역량을 진단하는 것을 넘어 전 세대를 포괄하는 국민의 시민적·사회적 소통 능력을 통합적으로 평가함으로써 기존의 쓰기 평가보다 더 폭넓은 언어 능력을 평가할 수 있다는 점에서 의의가 있다고 할 수 있다.

아래의 <표 II-34>는 본 연구의 조사 도구(NAKLA)와 각국의 쓰기 평가 도구를 영역과 대상 측면에서 비교한 표이다.

<표 II-38> 평가 영역 및 대상의 비교

평가(실시 국가)	영역	대상
NAPLAN (호주)	서사, 설득	3, 5, 7, 9학년 학습자
NAEP (미국)	설득, 서사(경험), 설명	4, 8, 12학년 학습자
NSBWT (노르웨이)	설득, 설명, 서사(상상)	5, 8학년 학습자
NAKLA (한국, 본 연구)	설득, 설명, 친교, 디지털	만 20세 이상 69세 이하 일반 국민

다음으로, 평가 준거를 바탕으로 외국의 자국어 능력 조사와 비교해 보자면 다음과 같다. 평가 준거가 공개된 NAPLAN과 NAEP, NSBWT와의 비교를 통해 본 연구의 평가 준거의 특성을 보다 명료히 제시하고자 한다. 우선 NAPLAN(ACARA, 2023)의 평가 준거를 살펴보면 <표 II-35>과 같다.

<표 II-39> NAPLAN 평가 준거 (서사)

번호	평가 준거	점수 범위	설명
1	독자	0-6	글쓴이가 독자를 인식하고, 독자를 끌어들이며, 독자에게 영향을 미치는 능력
2	텍스트 구조	0-4	이야기의 구조를 조직하는 능력으로, 도입-전개-결말을 적절하고 효과적으로 구성하는 능력
3	아이디어	0-5	이야기의 아이디어를 창의적으로 선택하고 구성하는 능력
4	인물과 배경	0-4	인물의 묘사 및 발전, 시간-장소-분위기의 감각을 형성하는 능력
5	어휘	0-5	언어 선택의 폭과 정확성
6	응집성(cohesion)	0-4	지시어, 대체어, 연결어, 연상어 등을 활용해 글 전체의 관계와 흐름을 통제하는 능력
7	문단 구성	0-2	독자가 내용을 쉽게 따라갈 수 있도록 문단을 적절히 구분하고 조직하는 능력
8	문장 구조	0-6	문법적으로 정확하고 구조적으로 완전하며 의미 있는 문장을 생성하는 능력
9	문장 부호	0-5	문장 부호를 정확하고 적절하게 사용하여 글의 가독성을 높이는 능력
10	철자	0-6	철자의 정확성 및 사용된 단어의 난이도 수준

널리 통용되는 쓰기 평가 지표인 ‘내용-조직-표현’ 영역에 따라 구분해 보면, 1~4번은 ‘내용’, 6~7번은 ‘구조’, 5번과 8~10번은 ‘표현’ 영역에 해당한다고 볼 수 있다. 서사라는 장르적 특수성으로 인해 4번과 같은 기준이 추가로 도입된 것으로 보인다.

설득 문항의 평가 기준은 <표 II-36>와 같다(ACARA, 2025).

<표 II-40> NAPLAN 평가 기준 (설득)

번호	평가 준거	점수 범주	주요 평가 초점 (Skill Focus)
1	독자	0-6	글쓰기가 독자를 인식하고, 독자의 관심을 끌며 설득하는 능력
2	텍스트 구조	0-4	설득적 글의 구조적 요소(서론, 본론, 결론)를 적절하고 효과적으로 조직하는 능력
3	아이디어	0-5	설득적 주장을 위한 아이디어를 선택하고, 관련성을 유지하며 구체적으로 발전시키는 능력
4	설득 전략(persuasive devices)	0-4	독자를 설득하기 위해 사용하는 설득적 장치의 다양성과 효과적인 활용 능력
5	어휘	0-5	문맥에 적절한 어휘를 폭넓고 정교하게 사용하는 능력
6	응집성(cohesion)	0-4	지시어, 생략, 접속사, 대체어, 연상어 등을 활용해 글 전체의 논리적 연결을 유지하는 능력
7	문단 구성	0-3	독자가 논리 전개를 따라갈 수 있도록 문단을 구분하고 조직하는 능력
8	문장 구조	0-6	문법적으로 정확하고 구조적으로 완전하며 의미 있는 문장을 생산하는 능력
9	문장 부호	0-5	문장 부호를 정확하고 적절히 사용하여 글의 가독성을 높이는 능력
10	철자	0-6	철자의 정확성과 사용된 단어의 난이도 수준

설득적 글쓰기의 평가 준거 역시 서사 글쓰기와 유사하게 구성되어 있으며, 서사에서 ‘인물과 배경’을 평가하였던 4번 준거가 ‘설득 전략’으로 제시된 것을 확인할 수 있다.

한편 NAEP의 평가 준거 또한 ‘내용-조직-표현’의 세 영역으로 구성되어 있다(박혜영 외, 2020: 81-82 참조). ‘내용 생성’, ‘내용 조직’, ‘언어 사용과 쓰기 관습’이라는 세 개 대영역으로 구성되어 있으며, 하위 평가 준거로는, 내용 생성에 있어 ‘글 내용의 복잡도와 깊이’, ‘사고력’, ‘내용의 상세함’, ‘적절한 예시’를 제시하고 있다. NAPLAN이나 NSBWT와는 달리 ‘독자’를 별도의 항목으로 구분하지는 않은 것으로 보인다.

<표 II-41> NAEP 평가 준거

영역	평가 준거
내용 생성	글 내용의 복잡도와 깊이
	사고력
	내용의 상세함
	적절한 예시
내용 조직	글의 구조
	응집성과 초점
언어 사용과 쓰기 관습	문장 구조와 표현의 다양성
	단어 선택
	목소리와 어조
	문법 및 언어 규범

한편, NSBWT의 평가 준거는 다음의 다섯 가지이다(Skar et al., 2017).

<표 II-42> NSBWT 평가 준거

작가-독자 상호작용	•작가와 독자는 적절한 방식으로 소통하는가? •작가는 독자의 정보 요구를 어느 정도 수용했는가?
내용(content)	•내용은 과제와 어느 정도 관련되어 있는가? •내용은 어느 정도 정교하고 (해석 불가)인가?
텍스트 구조	•텍스트가 과제에 지정된 의사소통 맥락과 관련이 있는가? •텍스트는 논리적인 구조를 갖추고 있는가?
언어 사용	•문장 변화와 언어 사용은 적절한가?
표현 역량	•문법, 철자법, 구두법을 적절히 사용하였는가?

NAPLAN과 마찬가지로 ‘독자’(‘작가-독자 상호작용’) 항목을 별도로 두어 구분하고 있으며, 이외에 ‘내용’, ‘텍스트 구조’, ‘언어 사용·표현 역량’을 평가하고 있다. 본 연구인 NAKLA에서도 기본적으로 ‘내용-조직-표현’의 대영역을 채택하되, 디지털 매체 쓰기 문항의 경우 ‘매체 활용’에 관한 평가 항목을 별도로 두어 학습자의 복합양식 리터러시를 측정하고자 하였다.

다음으로 디지털 매체 활용 역량 평가 측면을 외국의 자국어 능력 조사와 비교하기 위해 최근 디지털 매체 활용 역량에 관한 평가 동향을 요약하면 <표 II-39>와 같다. NSBWT의 경우 표준화된 CBT를 도입한 적 없으며, 현재 시행되고 있지 않아 본 분석에서 제외하였다. 2015년부터 선도적으로 CBT를 실시하고 있는 PISA 및 2021년부터 전면적으로 CBT를 실시하고 있는 PIAAC을 함께 비교하였다.

<표 II-43> 디지털 평가의 비교

	시행 방식	디지털·복합양식 평가 반영
NAPLAN (호주)	CBT (3학년 제외)	점진적 복합양식 텍스트 평가 강화
NAEP (미국)	CBT	컴퓨터 및 태블릿 PC 이용한 평가
PISA (OECD)	CBT-PBT 병행	디지털 환경에서의 학습(Learning in the Digital World)
PIAAC (OECD)	CBT	디지털 문제해결능력(PS-TRE, Adaptive Problem Solving)
NAKLA (한국, 본 연구)	CBT	디지털 매체 글쓰기

NAPLAN은 지필로부터 CBT(Computer-Based Test)로 전환을 완료하였으며, 학습자가 접하는 텍스트 유형을 복합양식적으로 확장하려고 지속적으로 시도하고 있다(Unsworth et al., 2019). 단, 3학년의 경우 컴퓨터 기반 평가가 아닌, 종이 기반 평가를 실시하고 있다(ACARA, n.d.). NAEP 쓰기 평가에서는 2011년부터 전면 컴퓨터 기반 테스트를 채택하였으며, 2017년에는 태블릿을 이용하여 쓰기 평가를 진행하고 있다(NCES, n.d.-b).

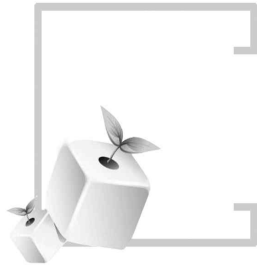
PISA에서는 읽기 리터러시 틀에 다중 출처(multiple-source), 동적 텍스트(dynamic text)를 포함하여 디지털 환경에서의 정보 탐색 및 통합 능력을 측정해 왔다. 2025년에는 ‘디지털 환경에서의 학습(Learning in the Digital World)’을 통해 디지털 도구를 활용한 지식 구축 및 문제 해결 능력을 별도로 평가하였다(OECD, 2019). PIAAC 역시 성인 대상 평가의 1주기에서 PS-TRE(기술 환경에서의 문제 해결)를(OECD, 2012), 2주기에서는 적응적 문제해결(Adaptive Problem Solving)을 도입하여 복잡한 디지털 맥락에서의 정보 탐색 및 통합 능력을 측정하였다. 이전의 PS-TRE가 디지털 기술을 활용하여

의사소통하는 능력에 중점을 두었다면, 2주기에서는 여전히 이 능력을 중요하게 여기면서도, “변화하는 정보 환경에 사고와 추론을 조정(adjust)할 수 있는 적응적 사고력(adaptive reasoning)”(OECD, 2017: 8)을 강조하는 방향으로 나아갔다고 할 수 있다.

한편, 본 연구(NAKLA)에서 개발한 평가 도구에서는 컴퓨터를 도구로 활용하는 것을 넘어 평가의 한 영역으로 디지털 매체 글쓰기를 설정함으로써 매체 리터러시 자체를 평가 대상으로 삼았다는 점이 특징적이다. 이는 여러 쓰기 및 읽기 평가에서 받아들여지고 있는 평가 환경의 디지털화와는 달리, 디지털 매체 공간에서 이루어지는 소통 능력을 평가의 대상으로 삼았다는 점에서 의미가 있다. 본 연구의 ‘디지털 매체 글쓰기’는 매체 그 자체를 의미와 구조, 표현을 변화시키는 요인으로 인식하고 다양한 매체 형식을 결합하여 의미를 생산하는 능력을 평가하는 데 목적을 둠으로써 새로운 시도를 하였다고 볼 수 있다. 특히 복합양식성을 핵심 개념으로 삼아 사진과 텍스트 등 이질적인 양식의 통합이 평가 기준에 포함되었다는 점에서 의의가 있다고 평가할 수 있다.

이상의 분석 내용을 요약하자면 다음과 같다. <국민의 국어능력 실태 조사>에서 개발한 문항(이하 NAKLA)은 기존의 학교급 내 학습자를 대상으로 한 쓰기 평가와 달리, 평가 대상과 내용의 범위를 확장하여 국민의 실제 언어 사용을 진단하는 데 목적을 두었다. 먼저 평가 영역 측면에서 서사적 글쓰기를 대신해 친교적 글쓰기를 도입함으로써 일상적 상황에서 이루어지는 소통을 포함하였고, 평가 대상을 성인 전반으로 확대하여 국민 전체의 국어 사용 역량을 평가하고자 하였다. 즉, 학업 성취도 중심의 리터러시 평가에서 나아가 시민적, 사회적 소통 역량을 포함하는 포괄적 언어 능력 평가로서 설계되었다는 의의가 있다.

또한 NAKLA는 ‘디지털 매체 글쓰기’를 독립된 평가 영역으로 설정하여, 단순히 컴퓨터 또는 태블릿 PC를 도구로 사용하는 것을 넘어 미디어 리터러시 자체를 평가 대상으로 삼았다. 이는 매체 환경을 언어적 의미 구성의 핵심 요소로 인식하고 다양한 매체 형식을 결합하여 의미를 생산하고 의사소통하는 능력을 측정하려는 시도라는 점에서 선도적이라고 평가할 수 있다.



제 3 장

실태 조사 성취 수준 도출



3.1. 진행 절차

본 조사의 분할 점수 산출은 대표적인 기준 설정 방법인 앵고프 방법(angoff method) 중, 자료에 대한 토의 과정을 수반하는 ‘수정된 앵고프 방법’을 적용하였으며, 성취 기준의 이해관계자인 전문가 집단에 의해 성취 수준 설정이 이루어졌다. 이를 위한 전문가 협의회는 평정의 용이성을 고려하여 문항 수와 소요 시간, 선택형(객관식) 문항과 수행형 및 서답형(주관식) 문항 유형의 차이에 따라 두 차례로 나누어 실시하였다.

<표 III-1> 전문가 협의회 일정

비고	일자	해당 영역	참여 전문가
1차 협의회	8월 6일(수)	듣기, 읽기, 문법·규범	12명
2차 협의회	11월 1일(토)	듣기·말하기, 말하기, 쓰기	

1) 참여 전문가 패널의 선발

분할 점수의 산출 과정에 참여할 전문가 패널은 ‘국어교육 분야 전문가’를 일차적인 요건으로 하였다. 이는 본 조사의 핵심 개념인 ‘국어능력’의 형성과 직접적인 관련을 맺고 있는 학문 영역이 국어교육학이기 때문이며, 이에 따라 국어교육 전공자를 중심으로 패널을 구성하였다. 그런데 특정 전공으로만 이루어진 전문가 집단이 성취 수준 결정에 참여할 경우 집단 편향이 발생할 수 있어, 분할 점수의 일반성을 담보하기 어렵다. 따라서 본 연구에서는 전문가 집단의 구성을 국어교육의 실천가이자 연구자인 교사, 대학원생(박사 수료 이상), 전공 교수의 비율이 균형 있게 분포되도록 하였으며, 동시에 다양한 배경을 지닌 전문가들이 참여할 수 있도록 구성의 다양성을 확보하였다.

교사의 경우 경력 및 지역적 다양성을 고려하여, 중고등학교에서 국어교육 경력 5년 이상을 보유한 4명을 선정하였다. 대학원생은 국어교육학 전공으로 박사과정을 수료한 연구자 4명을 선정하였고, 교수는 국어교육학 전공자로서 화법 교육, 쓰기 교육, 문법 교육, 읽기 교육 각 분야의 전문가를 1명씩 포함해 총 4명을 위촉하였다.

한편, 1차 협의회 이후 진행된 2차 협의회에서는 개인 사정으로 인해 4명의 전문가 패널이 교체되었다. 교체된 인원의 선정도 1차 협의회와 동일한 기준(전공, 경력, 전문성)을 충족하는지를 고려하여 패널 구성의 일관성과 전문성 수준을 유지하고자 하였다. 이에 따라 2차 협의회 역시 교사·연구자·교수의 비율이 균형 있게 유지되었으며, 수준 설정 과정의 절차적 타당성을 확보하였다. 이러한 과정을 거쳐 최종적으로 분할 점수 도출 협의회에 참여한 전문가 집단의 구성은 다음과 같다.

<표 III-2> 참여 전문가 구성

구분	인원(명)	전문가 집단의 구성	
		1차 협의회	2차 협의회
교사	4	경기(고, 22년), 서울(중, 15년), 서울(중, 6년), 충청(고, 5년)	서울(중, 20년), 서울(중, 15년), 서울(중, 6년), 충청(고, 5년)

대학원생	4	국어교육학 전공 박사 수료 이상	국어교육학 전공 박사 수료 이상
교수	4	화법 교육, 문법 교육, 읽기 교육, 쓰기 교육	화법 교육(2인), 문법 교육, 쓰기 교육

2) 수준 설정 협의회 진행 과정

협의회 진행 과정은 크게 사전 준비와 평정 전 최소 능력 보유자의 수행 특성 기술문 작성, 실제 분할 점수 판정 과정으로 이루어진다. 먼저, 사전 준비 단계에서 전문가 패널은 이번 연구의 전반적인 성격과 조사 방식, 조사 대상 등에 대한 교육을 받는다. 또한, 실제 분할 점수의 산출 과정이 어떠한 방식으로 진행되며, 이 과정에서 전문가 집단이 어떠한 역할을 해야 하는지에 대해서도 구체적으로 안내받는다. 이에 더해, 해당 영역의 조사 도구 문항들을 직접 풀이하고, 내용과 수준을 검토하는 과정을 통해 해당 조사 도구에 대해 심층적으로 이해할 수 있도록 사전 활동을 제시하였다. 선택형 문항을 대상으로 하는 1차 협의회의 사전 활동에서는 패널에게 조사 영역별 문항에 대한 예상 정답과 기대 정답률(0~100%)을 입력하도록 하였으며, 다분형 문항을 대상으로 하는 2차 협의회의 사전 활동에서는 수집된 본조사 응답 사례를 총점 기준으로 사분위수로 나누고, 평균과 각 사분위수 구간에 위치하는 다섯 사례의 답안을 듣거나 읽고, 그에 대한 예상 정답과 기대 정답률(0~100%)을 입력하도록 하였다. 사전 활동의 참여 결과는 협의회 당일에 이루어지는 수행 수준 기술문(PLD) 작성을 위한 협의의 참고 자료로 제공하였다.

협의회 당일에 이루어지는 수행 수준 기술문 작성 과정에서는 먼저 본 연구에서 사정한 각 수준별 국어능력의 의미가 무엇인지를 제시하고, 본조사 참여자의 인구통계학적 특성(지역, 연령, 학력, 직업 등)에 관한 정보를 제공하였다.

<표 III-3> 국어능력의 일반적 수준(국립국어원·서울대학교 국어교육연구소, 2013)

등급	국어능력 수준
우수	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 대부분(80% 이상) 획득한 수준
보통	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 상당 부분(50% 이상 80% 미만) 획득한 수준
기초	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 부분적으로(20% 이상 50% 미만) 획득한 수준

국어능력 수준 및 인구통계학적 특성 제시 후 3명씩 구성된 네 개의 소집단을 중심으로 실질적인 토의가 이루어지도록 하였다. 소집단 편성은 패널 구성 비율이 반영되도록 하였다. 즉, 하나의 소집단은 교사 1명, 박사 수료 이상 연구자 1명, 교수 1명으로 구성되었다. 그리고 전공 영역의 다양성을 확보하기 위해, 각 소집단에는 반드시 한 명 이상의 다른 영역의 전공자가 포함되도록 하였다.⁵⁾

5) 특히 2차 협의회에서는 전문가 구성 비율을 유지하는 동시에, 1차 협의회에 참여했던 전문가 2명과 새롭게 교체된 전문가 1명을 한 소집단으로 배치하였다. 이를 통해 교체된 패널이 기존 논의의 흐름을 충분히 공유하고 이해할 수 있도록 하였으며, 협의 과정의 연속성과 판단의 일관성을 확보하고자 하였다.

이러한 방식으로 구성된 전문가들은 네 개의 소집단에 편성되었으며, 각 평정자에게 1~12의 번호를 부여하였다. 전문가들은 소집단 구성원들과 공동 과제를 수행하며 상호작용하고, 전체 판정 결과를 참고하면서 평정에 참여하였다.

1차 협의회		2차 협의회	
시간	일정	시간	일정
오전 9:00	워크숍 안내	오전 9:00	워크숍 안내
9:15	PLD 작성사례 검토	9:15	PLD 작성 사례 검토
9:30	듣기 영역 문항 카드 검토 및 수준별 최소능력보유자의 PLD작성	9:30	듣기·말하기 영역 문항, 채점 척도, 답안 사례 검토 및 수준별 최소능력보유자의 PLD작성
10:30	읽기 영역 문항 카드 검토 및 수준별 최소능력보유자의 PLD 작성	10:30	말하기 영역 문항, 채점 척도, 답안 사례 검토 및 수준별 최소능력보유자의 PLD작성
11:30	문법 영역 문항 카드 검토 및 수준별 최소능력보유자의 PLD 작성	11:30	쓰기 영역 문항, 채점 척도, 답안 사례 검토 및 수준별 최소능력보유자의 PLD작성
오후 12:15	점심 시간	오후 12:15	점심 시간
13:15	PLD 작성 결과 확인 및 협의(소그룹)	13:15	PLD 작성 결과 확인 및 협의(소그룹)
13:45	Round 1 평정	13:45	Round 1 평정
15:00	쉬는 시간	15:00	쉬는 시간
15:10	본조사기술 통계 및 1 Round 결과 확인 및 협의(중그룹)	15:10	본조사 기술 통계 및 1 Round 결과 확인 및 협의(중그룹)
15:40	Round 2 평정	15:40	Round 2 평정
16:40	쉬는 시간	16:40	쉬는 시간
16:50	Round 2 결과 확인 및 협의(전체)	16:50	Round 2 결과 확인 및 협의(전체)
17:20	Round 3	17:20	Round 3
18:00	마무리	18:00	마무리

[그림 III-1] 수준 설정 협의회 진행 과정

3) 수준별 최소 능력 보유자 협의

앵고프 방식은 전문가 패널이 상정하는 최소 능력 보유자의 상에 근거하여 분할 점수 판정이 이루어지므로, 각 등급에서 이를 적절히 설정하는 것이 매우 중요하다. 최소 능력 보유자에 대한 협의는 본 조사 도구에서 설정한 1~4 수준(기초미달-기초-보통-우수)의 기본적인 개념을 바탕으로 이루어졌으며, 작성의 목표로는 수준별 분할이 이루어지는 4(우수)-3(보통), 3(보통)-2(기초), 2(기초)-1(기초미달) 수준의 최소 능력 보유자가 조사 문항을 통해 나타낼 수행에서의 특징을 기술하는 방식으로 이루어졌다. 협의를 돕기 위해 패널들이 사전 과제로 제출한 문항별 예상 정답(수행형 문항의 경우 응답 사례별 예상 점수)과 기대 정답률(수행형 문항의 경우 기대 점수 확률)의 제출 결과를 종합하여 공유하였으며, 이를 바탕으로 각 소그룹에서는 문항에 관한 토의를 진행하였다. 소그룹 협의를 통해 각 영역의 수준별 최소 능력자가 수행할 것으로 기대되는 표준으로 수행 수준 기술문(Performance Level Descriptor)을 작성하였다.⁶⁾

<표 III-4> 듣기 영역 수준별 최소 능력 보유자의 수행 수준 기술문

영역	수준	수준별 최소 능력자의 수행 수준
듣기	4-3	강연, 토론, 인터뷰 등 다양한 담화 상황에서 화자의 의도와 주요 내용을 파악하고 비판적으로 수용할 수 있는 최소한의 능력을 갖추고 있다. 명시적·함축적 정보를 구별하며 주장을 논리적으로 분석하고, 매체의 표현 효과나 편향성을 인식하는 최소한의 능력을 갖추고 있다. 긴 구어 담

6) 이때 수행형 문항(듣기·말하기, 말하기, 쓰기) 영역의 경우, 문항 특성에 따라 내용·조직·표현·전달·매체 등 세부 준거별로 수행 수준 기술문을 작성하였다.

		화에서도 세부 내용을 집중해 이해하며, 전문 용어가 사용되거나 낯선 주제를 다룬 담화도 파악할 수 있는 최소한의 능력을 갖추고 있다. 정보를 재구성하고 문제 해결을 위한 대안을 제시할 수 있는 최소한의 능력을 갖추고 있다.
	3-2	일상적이고 익숙한 주제의 구어 담화에서 발화의 중심 내용과 화자의 의도를 어느 정도 파악할 수 있으며, 강연이나 인터뷰 등에서 주장과 근거의 관련성을 일부 이해할 수 있는 최소한의 능력을 갖추고 있다. 매체의 특성을 파악하고 표현 효과와 내용을 부분적으로 이해할 수 있는 최소한의 능력을 갖추고 있다. 길이가 길지 않고 일상적인 용어가 사용되거나 친숙한 주제를 다룬 담화를 파악할 수 있는 최소한의 능력을 갖추고 있다. 제한적인 수준에서 비판적 사고와 종합적 분석을 할 수 있는 최소한의 능력을 갖추고 있다.
	2-1	일상적인 어휘로 구성된 짧은 담화에 대해 제한적으로 이해할 수 있는 최소한의 능력을 갖추고 있다. 강연이나 인터뷰 등의 담화에서는 화자의 의도 추론이나 정보 간의 관계를 파악할 수 있는 최소한의 능력을 갖추고 있다.

<표 III-5> 읽기 영역 수준별 최소 능력 보유자의 수행 수준 기술문

영역	수준	수준별 최소 능력자의 수행 수준
읽기	4-3	기사·설명문 등 정보 글의 핵심을 파악하고 비판적으로 이해할 수 있으며, 복합 매체 자료에서 시각·청각적 요소의 의미와 표현 효과를 파악하고 비판적으로 이해할 수 있는 최소한의 능력을 갖추고 있다. 수필과 소설을 통해 경험, 감정, 상황을 정교하게 이해하고, 전자우편 등 실용문에서는 의도와 표현 효과를 파악할 수 있는 최소한의 능력을 갖추고 있다. 긴 텍스트의 세부 내용을 파악할 수 있으며, 사회문화적 맥락을 고려하여 내용을 이해할 수 있는 최소한의 능력을 갖추고 있다. 글의 목적과 표현 방식의 효과를 분석하고, 추론적으로 글을 이해할 수 있으며, 글쓴이의 관점을 바탕으로 비판적인 해석을 수행할 수 있는 최소한의 능력을 갖추고 있다.
	3-2	정보 글, 문학, 실용문 등 다양한 유형의 글을 이해하고 글쓴이의 의도를 파악할 수 있는 최소한의 능력을 갖추고 있다. 공익 광고나 홍보 영상과 같은 복합 매체 자료에서 시각적, 청각적 요소의 표현 효과를 일부 이해할 수 있으며, 주제를 제한적으로 파악할 수 있는 최소한의 능력을 갖추고 있다. 간단한 텍스트의 세부 내용을 파악할 수 있으며, 담화 맥락과 배경지식을 고려하여 중심 내용을 부분적으로 파악할 수 있는 최소한의 능력을 갖추고 있다. 제한적인 수준에서 표현 효과를 파악하고, 제한적으로 추론적, 비판적으로 이해할 수 있는 최소한의 능력을 갖추고 있다.
	2-1	친숙한 주제나 자주 접하는 유형의 짧은 글에서 내용을 부분적으로 이해하고 글쓴이의 의도를 제한적으로 파악할 수 있는 최소한의 능력을 갖추고 있다. 간단한 형식의 복합 매체 자료에서 부분적으로 주장이나 정보를 파악할 수 있는 최소한의 능력을 갖추고 있다. 간단한 텍스트의 세부 내용을 제한적으로 파악할 수 있는 최소한의 능력을 갖추고 있다.

<표 III-6> 문법·규범 영역 수준별 최소 능력 보유자의 수행 수준 기술문

영역	수준	수준별 최소 능력자의 수행 수준
문법·규범	4-3	어휘, 문장, 담화 차원에서 국어 어휘와 문법 요소의 문법적 쓰임과 의미를 이해하고, 맥락에 따른 미묘한 의미 차이를 구분하여 인식하고 이를 규범에 맞게 사용할 수 있는 최소한의 능력을 갖추고 있다. 관용 표현의 맥락적 의미를 파악하고 유의어를 맥락에 맞게 선택하며 문장 성분 간의 호응도 파악할 수 있는 최소한의 능력을 갖추고 있다. 조사 및 어미의 의미를 민감성을 가지고 인식할 수 있으며, 지시어와 접속 표현 등의 적절성을 판단할 수 있는 최소한의 능력을

		갖추고 있다. 표준 발음, 한글 맞춤법, 외래어 표기법 등 국어 규범 전반에 대해 민감성을 지니고 있으며, 이를 바탕으로 적절한 표현을 선택할 수 있는 최소한의 능력을 갖추고 있다.
	3-2	맥락에 따른 섬세한 의미 차이 파악이나 복잡한 어문 규범 적용은 제한적인 수준에서 수행하지만, 일상적이거나 기본적인 맥락에서 국어의 어휘와 문법 요소가 지닌 의미와 쓰임은 파악할 수 있는 최소한의 능력을 갖추고 있다. 친숙한 어휘는 활용할 수 있으며, 유의어 간의 의미 차이를 제한적으로 판단할 수 있는 최소한의 능력을 갖추고 있다. 조사와 어미, 지시어와 접속 표현 일상적인 수준에서 사용되었을 경우에는 적절성을 판단하지만 복잡한 맥락에서 사용된 경우 적절성을 제한적으로 판단할 수 있는 최소한의 능력을 갖추고 있다. 표준 발음과 한글 맞춤법은 일부 알고 있으나 복잡한 음운 변화가 일어나는 경우 올바른 발음 인식이 제한적이며, 자주 사용하는 외래어 표기는 알고 있으나 외래어 표기 원칙은 부분적으로 적용할 수 있는 최소한의 능력을 갖추고 있다.
	2-1	일상적인 맥락에서 기본적인 어휘와 문법 요소를 제한적으로 인식하고 사용할 수 있는 최소한의 능력을 갖추고 있다. 기본적인 유의어, 어미와 조사, 지시어와 접속어 등의 의미와 쓰임을 대략적으로만 인식하며, 표기와 발음 간의 관계 인식이 제한적이고 기초적인 수준에서 제한적으로 표기할 수 있는 최소한의 능력을 갖추고 있다.

<표 Ⅲ-7> 듣기·말하기 영역 수준별 최소 능력 보유자의 수행 수준 기술문

영역	수준	수준별 최소 능력자의 수행 수준	
듣기·말하기	4-3	내용	다양한 담화 상황에서 상대의 감정과 기대를 고려하여 말할 수 있는 최소한의 능력을 갖추고 있다. 대화의 목적, 맥락, 문제 상황을 정확히 파악하여 의도에 맞게 적절한 해결 방안을 제시할 수 있는 최소한의 능력을 갖추고 있다. 핵심 정보가 충분하고 구체적으로 드러나도록 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다.
		조직	다양한 담화 상황에서 목적에 맞게 내용을 체계적으로 조직해 논리적으로 전개할 수 있는 최소한의 능력을 갖추고 있다. 담화의 맥락과 흐름을 파악하여 상대의 이전 발화와 자연스럽게 연결할 수 있는 최소한의 능력을 갖추고 있다. 발화를 적절하게 시작하고 분명하게 마무리할 수 있는 최소한의 능력을 갖추고 있다.
		표현	청자와 맥락에 적합한 어휘·표현을 선택하고, 예의 있게 말할 수 있는 최소한의 능력을 갖추고 있다. 어법에 맞고 불필요한 반복 없이 자연스럽게 말할 수 있는 최소한의 능력을 갖추고 있다.
		전달	발음이 정확·명료하며, 속도·어조를 조절해 유창하게 말할 수 있는 최소한의 능력을 갖추고 있다.
	3-2	내용	다양한 담화 상황에서 상대의 말을 이해하여 적합하게 말할 수 있는 최소한의 능력을 갖추고 있다. 대화의 목적, 맥락, 문제 상황을 어느 정도 파악하여 일반적으로 기대되는 응답을 할 수 있는 최소한의 능력을 갖추고 있다. 핵심 정보를 일부 포함하며, 부분적으로 구체성을 갖춰 표현할 수 있는 최소한의 능력을 갖추고 있다.
		조직	다양한 담화 상황에서 목적을 고려하여 내용을 조직할 수 있는 최소한의 능력을 갖추고 있다. 담화에서 부분적으로 논리성과 체계성을 갖추어 말할 수 있는 최소한의 능력을 갖추고 있다. 상대의 이전 발화에 맞게 담화 흐름을 부분적으로 연결할 수 있는 최소한의 능력을 갖추고 있다.
		표현	청자와 맥락에 적합한 어휘를 부분적으로 사용할 수 있는 최소한의 능력을 갖추고 있다. 어법을 제한적으로 지키며 말할 수 있는 최소한의 능력을 갖추고 있다.
		전달	발음이 부분적으로 정확하나 이해 가능한 수준이며, 제한적인 수준에서 속도, 어조를 조

			절하여 말할 수 있는 최소한의 능력을 갖추고 있다.
	2-1	내용	대화의 목적, 맥락을 부분적으로 이해하여 문제에 적합한 응답을 제한적으로 발화할 수 있는 최소한의 능력을 갖추고 있다. 상대방의 말을 부분적으로 이해하여, 상대의 감정에 제한적으로 공감하는 최소한의 능력을 갖추고 있다. 대화의 목적과 상황을 극히 부분적으로 고려하여 대화를 구성하는 최소한의 능력을 갖추고 있다.
		조직	다양한 담화 상황에서 대략적 내용을 조직할 수 있는 최소한의 능력을 갖추고 있다. 담화에서 극히 제한적으로 논리성과 체계성을 갖추어 말할 수 있는 최소한의 능력을 갖추고 있다. 담화 흐름에 극히 부분적으로 연결된 말을 할 수 있는 최소한의 능력을 갖추고 있다.
		표현	청자와 맥락에 적합한 어휘를 극히 부분적으로 사용할 수 있는 최소한의 능력을 갖추고 있다. 어법을 극히 제한적으로 지키며 말할 수 있는 최소한의 능력을 갖추고 있다.
		전달	극히 제한적으로 발음, 속도, 어조를 갖추어 말할 수 있는 최소한의 능력을 갖추고 있다.

<표 III-8> 말하기 영역 수준별 최소 능력 보유자의 수행 수준 기술문

영역	수준	수준별 최소 능력자의 수행 수준	
말하기	4-3	내용	화제 및 주제, 맥락, 담화 유형을 고려하여 내용을 선정할 수 있는 최소한의 능력을 갖추고 있다. 청자의 특성을 고려하여 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다. 주장하는 말하기에서 쟁점을 바탕으로 주장을 정확하게 구성하고 타당한 근거를 제시할 수 있는 최소한의 능력을 갖추고 있다. 설명하는 말하기에서 다양한 설명 방법(정의, 구분, 예시 등)을 들어 말하는 최소한의 능력을 갖추고 있다. 자료를 적절히 활용하여 구체적인 내용을 충분히 마련할 수 있는 최소한의 능력을 갖추고 있다.
		조직	담화 유형을 고려하여 선정한 내용을 논리적이고 체계적으로 조직할 수 있는 최소한의 능력을 갖추고 있다. 담화 표지 등을 활용하여 시작과 끝이 드러나고 일관성 있게 조직할 수 있는 최소한의 능력을 갖추고 있다.
		표현	청자의 수준을 고려하여 어휘와 어법을 적절하게 구사하는 최소한의 능력을 갖추고 있다. 말하기 상황을 고려하여 맥락에 적합한 어휘와 어법, 문장 구조를 사용하여 효과적으로 말할 수 있는 최소한의 능력을 갖추고 있다.
		전달	발음이 명료하고 정확하며, 속도·어조를 적절하고 능숙하게 조절하여 유창하게 말할 수 있는 최소한의 능력을 갖추고 있다.
	3-2	내용	화제 및 주제, 맥락, 담화 유형을 부분적으로 고려하여 기본적인 내용을 선정할 수 있는 최소한의 능력을 갖추고 있다. 청자의 특성을 부분적으로 고려하여 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다. 주장하는 말하기에서 주장을 구성하고 기본적인 근거를 제시할 수 있는 최소한의 능력을 갖추고 있다. 설명하는 말하기에서 간단한 설명 방법을 사용하여 말하는 최소한의 능력을 갖추고 있다. 자료를 제한적으로 활용하여 내용을 마련할 수 있는 최소한의 능력을 갖추고 있다.
		조직	담화 유형을 제한적으로 고려하여 내용을 조직할 수 있는 최소한의 능력을 갖추고 있다. 시작과 끝이 부분적으로 드러나게 내용을 조직할 수 있는 최소한의 능력을 갖추고 있다.
		표현	청자의 수준을 부분적으로 고려하며 어휘와 어법을 제한적인 수준으로 구사할 수 있는 최소한의 능력을 갖추고 있다. 말하기 상황을 고려하여 맥락에 적합한 어휘와 어법, 문

			장 구조를 사용하여 말할 수 있는 최소한의 능력을 갖추고 있다.
		전달	발음이 부분적으로 정확하나 이해 가능한 수준이며, 제한적인 수준에서 속도, 어조를 조절하여 말할 수 있는 최소한의 능력을 갖추고 있다.
	2-1	내용	화제 및 주제를 극히 제한적으로 고려하여 간단한 내용을 선정할 수 있는 최소한의 능력을 갖추고 있다. 청자의 특성을 극히 부분적으로 고려하여 간단한 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다. 주장하는 말하기에서 극히 제한적인 주장을 제시하고 부분적인 이유를 제시할 수 있는 최소한의 능력을 갖추고 있다. 설명하는 말하기에서 극히 제한적인 설명 방법을 사용하여 말하는 최소한의 능력을 갖추고 있다. 자료를 극히 제한적으로 활용하여 내용을 마련할 수 있는 최소한의 능력을 갖추고 있다.
		조직	담화 유형을 극히 제한적으로 고려하여 내용을 조직할 수 있는 최소한의 능력을 갖추고 있다.
		표현	어휘와 어법을 극히 제한적인 수준으로 구사할 수 있는 최소한의 능력을 갖추고 있다. 말하기 상황을 부분적으로 고려하며 기초적인 어휘와 어법, 문장 구조를 제한적으로 사용하여 말할 수 있는 최소한의 능력을 갖추고 있다.
		전달	극히 제한적으로 발음, 속도, 어조를 갖추어 말할 수 있는 최소한의 능력을 갖추고 있다.

<표 III-9> 쓰기 영역 수준별 최소 능력 보유자의 수행 수준 기술문

영역	수준	수준별 최소 능력자의 수행 수준	
쓰기	4-3	내용	주제, 글의 목적과 유형(설명, 논증, 사과, 요청), 예상 독자 등 작문의 맥락을 명확히 이해하여, 이에 적합한 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다. 주장하는 글에서 쟁점에 대한 주장을 분명하게 제시하고, 주장에 대한 타당한 이유와 근거를 제시할 수 있는 최소한의 능력을 갖추고 있다. 설명하는 글에서 대상의 특성을 고려한 적절한 설명 방법(묘사, 비교, 대조, 분석 등)을 사용하여 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다.
		조직	글의 내용을 체계적으로 조직·배열하여 응집성 있는 글을 쓸 수 있는 최소한의 능력을 갖추고 있다. 문장 간, 문단 간 연결이 유기적이고 통일성 있게 조직할 수 있는 최소한의 능력을 갖추고 있다.
		표현	독자와 목적 등 맥락을 고려하여 적절한 어휘와 문장을 사용하고, 글쓰기 관습을 지켜 어문 규범(맞춤법·띄어쓰기·문장부호)에 맞게 글을 쓸 수 있는 최소한의 능력을 갖추고 있다. 적절한 문장 길이, 문장 성분의 자연스러운 호응, 일관된 문체로 표현할 수 있는 최소한의 능력을 갖추고 있다.
		매체	표현 의도와 글의 내용을 고려하여 복합양식 자료를 효과적으로 활용할 수 있는 최소한의 능력을 갖추고 있다.
	3-2	내용	주제, 글의 목적과 유형(설명, 논증, 사과, 요청), 예상 독자 등 작문의 맥락을 부분적으로 이해하여, 글의 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다. 주장하는 글에서 쟁점에 대한 주장을 제한적으로 제시하고, 주장에 대한 부분적인 이유를 제시할 수 있는 최소한의 능력을 갖추고 있다. 설명하는 글에서 대상의 특성을 부분적으로 고려한 간단한 설명 방법을 사용하여 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다.
		조직	글의 내용을 제한적으로 조직·배열할 수 있는 최소한의 능력을 갖추고 있다. 제한적

			인 수준에서 문장 간, 문단 간 연결을 할 수 있는 최소한의 능력을 갖추고 있다.
		표현	글의 목적을 고려하여 어휘와 문장을 사용하고, 어문 규범(맞춤법·띄어쓰기·문장부호)을 부분적으로 지키며 글을 쓸 수 있는 최소한의 능력을 갖추고 있다. 문장 길이, 문장 성분의 호응, 일관된 문체 등을 부분적으로 고려하여 표현할 수 있는 최소한의 능력을 갖추고 있다.
		매체	표현 의도와 글의 내용을 고려하여 복합양식 자료를 부분적으로 활용할 수 있는 최소한의 능력을 갖추고 있다.
	2-1	내용	주제, 글의 목적과 유형(설명, 논증, 사과, 요청), 예상 독자 등 작문의 맥락을 극히 부분적으로 이해하여, 간단한 내용을 구성할 수 있는 최소한의 능력을 갖추고 있다. 주장하는 글에서 주장이나 이유를 극히 제한적으로 제시할 수 있는 최소한의 능력을 갖추고 있다. 설명하는 글에서 간단한 설명 방법을 사용하여 내용을 제한적으로 구성할 수 있는 최소한의 능력을 갖추고 있다.
		조직	글의 내용을 극히 제한적으로 조직·배열할 수 있는 최소한의 능력을 갖추고 있다. 극히 제한적인 수준에서 문장 간, 문단 간 연결을 할 수 있는 최소한의 능력을 갖추고 있다.
		표현	어문 규범(맞춤법·띄어쓰기·문장부호)을 일부 지키며 글을 쓸 수 있는 최소한의 능력을 갖추고 있다. 문장 길이, 문장 성분의 호응, 일관된 문체 등을 극히 제한적으로 고려하여 표현할 수 있는 최소한의 능력을 갖추고 있다.
		매체	표현 의도를 제한적으로 고려하여 복합양식 자료를 극히 부분적으로 활용할 수 있는 최소한의 능력을 갖추고 있다.

4) 분할 점수 산출 방법

본 조사의 분할 점수는 전문가 패널 12명의 문항별 기대 정답률 판정이 전체 3번의 라운드(평정 회차)를 거치면서 조정되는 과정을 통해 산출된다. 이때 3번의 라운드는 <기대 정답률 판정 → 정보 피드백 → 토의 → 기대 정답률 판정 → 정보 피드백 → 토의 → 기대 정답률 판정>으로 진행되었다. 기대 정답률 판정은 개별 문항에 대해 이루어졌으며, 반복된 판정 작업과 관련된 자료 제공, 이에 관한 토의 과정에서의 의견 교환을 통해 전문가들 간의 편차를 줄이는 과정을 거쳤다. 협의회에서 진행된 기대 정답률 판정을 통한 분할 점수 산출의 구체적인 방법은 다음과 같다.

첫째, 기대 정답률의 판정은 객관식 문항(듣기, 읽기, 문법·규범)의 경우에는 평정 회차별로 각 영역 문항에 대한 성취 수준 최소 능력 보유자가 정답을 맞힐 확률을 0~100%(5% 단위)로 표기한다. 수행형 문항(듣기·말하기, 말하기, 쓰기)의 경우 각 문항은 내용·조직·표현·전달·매체 등 세부 준거별로 채점되므로, 각 준거에 대해 해당 성취 수준의 최소 능력 보유자가 해당 점수를 받을 척도별 확률을 0~100%(5% 단위)로 표기한다.

객관식 문항 평정 양식(듣기)

문항 번호	Round1		
	우수-보통	보통-기초	기초-기초미달
L-1			
L-2			
L-3			
L-4			
L-5			
L-6			
L-7			
L-8			
L-9			
L-10			
L-11			
L-12			
L-13			
L-14			
L-15			

수행형 문항 평정 양식(쓰기)

문항 번호	평가항목	점수	Round1		
			우수-보통	보통-기초	기초-기초미달
W-1-공통	내용	5점(매우 우수)			
		4점(우수)			
		3점(보통)			
		2점(미흡)			
		1점(매우 미흡)			
	조직	합계	0	0	0
		5점(매우 우수)			
		4점(우수)			
		3점(보통)			
		2점(미흡)			
		1점(매우 미흡)			
	표현	합계	0	0	0
		5점(매우 우수)			
		4점(우수)			
		3점(보통)			
		2점(미흡)			
		1점(매우 미흡)			
		합계	0	0	0

[그림 III-2] 평정 양식

둘째, 평정을 하는 각 라운드 사이에는 판정에 도움을 줄 수 있는 자료(정보)가 제공된다. 1라운드 종료 후 2라운드 시작 전에 각 영역의 원점수 평균과 문항별 곤란도·변별도 등 기술 통계 자료를 제공한다. 2라운드 종료 후에는 전체 전문가 집단의 2라운드 평정 결과 평균과 분포표(기준 정보)를 제공한다.

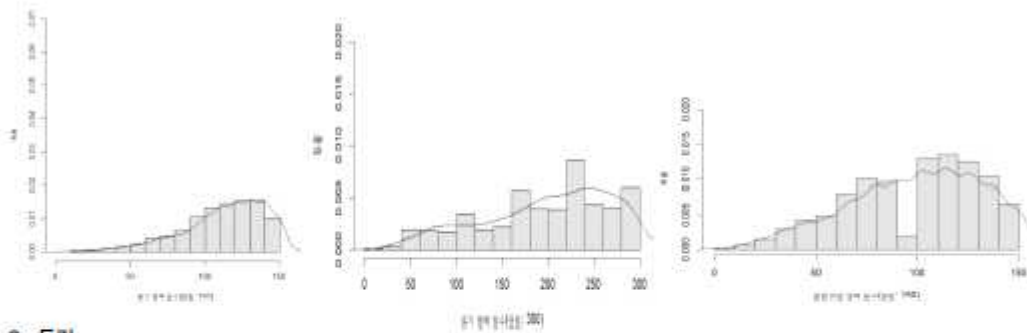
<표 III-10> 라운드별 제공 정보

선택형 문항 (듣기, 읽기, 문법·규범)	1라운드 → 2라운드	▪ 본조사의 문항별 기술 통계
	2라운드 → 3라운드	▪ 전체 패널의 2차 판정 결과의 평균값, 중앙값 및 분포표
수행형 문항 (듣기·말하기, 말하기, 쓰기)	1라운드 → 2라운드	▪ 본조사 문항의 준거별 기술 통계
	2라운드 → 3라운드	▪ 전체 패널의 2차 판정 결과의 평균값, 중앙값 및 분포표

제3차 국민의 국어능력 실태 조사
듣기/읽기/문법 및 규범 영역 기술 통계 자료

1. 총괄(영역별 원점수 요약)

영역	N	만점	평균	표준편차	최솟값	최댓값
듣기	3,033	150	114.11	27.02	10	150
읽기	5,000	300	197.93	69.36	15	300
문법·규범	5,000	150	98.46	32.93	0	150



2. 듣기

영역	문항	N	곤란도	해석	변별도	해석
듣기	L-1	3,033	0.55	중간	0.44	높음
듣기	L-2	3,033	0.83	매우 쉬움	0.44	높음
듣기	L-3	3,033	0.93	매우 쉬움	0.35	있음
듣기	L-4	3,033	0.77	쉬움	0.44	높음
듣기	L-5	3,033	0.76	쉬움	0.47	높음
듣기	L-6	3,033	0.72	쉬움	0.49	높음
듣기	L-7	3,033	0.97	매우 쉬움	0.27	낮음
듣기	L-8	3,033	0.47	중간	0.50	높음
듣기	L-9	3,033	0.84	매우 쉬움	0.52	높음
듣기	L-10	3,033	0.81	매우 쉬움	0.43	높음
듣기	L-11	3,033	0.83	매우 쉬움	0.49	높음
듣기	L-12	3,033	0.64	쉬움	0.55	높음
듣기	L-13	3,033	0.83	매우 쉬움	0.39	있음
듣기	L-14	3,033	0.73	쉬움	0.46	높음
듣기	L-15	3,033	0.72	쉬움	0.37	있음

[그림 Ⅲ-3] 기술 통계

평가항목	수준	점수	평정별 빈도																				분포							
			0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	평균	중앙값	절사평균	표준편차	최소값	최대값	범위
내용 선정	우수·보통	5점(매우 우수)	0	0	0	1	2	0	3	0	3	0	0	1	0	0	0	2	0	0	0	0	0	40.4	35.0	39.0	21.1	15	80	15-80 (65)
		4점(우수)	0	0	0	1	1	0	2	0	5	0	2	0	1	0	0	0	0	0	0	0	37.9	40.0	38.0	12.2	15	60	15-60 (45)	
		3점(보통)	1	1	2	0	5	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	19.2	20.0	19.0	11.7	0	40	0-40 (40)	
		2점(미흡)	9	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.1	0.0	1.0	4.3	0	15	0-15 (15)	
		1점(매우 미흡)	11	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.0	0.0	1.4	0	5	0-5 (5)	
		합계																												
	보통·기초	5점(매우 우수)	2	1	2	1	3	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	20.4	17.5	17.5	19.7	0	70	0-70 (70)	
		4점(우수)	0	0	0	1	3	0	4	2	1	0	0	0	1	0	0	0	0	0	0	0	30.4	30.0	29.0	11.4	15	60	15-60 (45)	
		3점(보통)	1	0	0	1	1	0	1	3	1	1	3	0	0	0	0	0	0	0	0	0	33.8	35.0	35.5	14.9	0	50	0-50 (50)	
		2점(미흡)	2	2	4	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	12.5	10.0	12.0	9.7	0	30	0-30 (30)	
		1점(매우 미흡)	7	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.9	0.0	2.5	3.8	0	10	0-10 (10)	
		합계																												
기초·기초미달	5점(매우 우수)	8	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8.3	0.0	4.5	16.0	0	55	0-55 (55)		
	4점(우수)	3	1	4	0	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	12.9	10.0	12.5	11.3	0	30	0-30 (30)		
	3점(보통)	0	0	1	0	1	1	3	3	3	0	0	0	0	0	0	0	0	0	0	0	30.8	32.5	32.0	8.6	10	40	10-40 (30)		
	2점(미흡)	0	1	1	0	1	0	1	0	5	1	1	1	0	0	0	0	0	0	0	0	34.6	40.0	35.5	14.8	5	55	5-55 (50)		
	1점(매우 미흡)	2	1	4	0	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	13.3	10.0	13.0	9.2	0	30	0-30 (30)		
	합계																													
내용 구성	우수·보통	5점(매우 우수)	0	0	0	1	2	2	1	1	2	0	1	0	0	0	0	1	1	0	0	0	37.9	32.5	36.0	20.2	15	80	15-80 (65)	
		4점(우수)	0	0	0	0	2	0	2	1	3	1	1	0	1	1	0	0	0	0	0	0	39.6	40.0	39.0	13.5	20	65	20-65 (45)	
		3점(보통)	1	1	0	1	5	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	20.0	20.0	21.0	9.4	0	30	0-30 (30)	
		2점(미흡)	9	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.1	0.0	1.5	3.8	0	10	0-10 (10)	
		1점(매우 미흡)	11	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.0	0.0	1.4	0	5	0-5 (5)	
		합계																												
	보통·기초	5점(매우 우수)	2	1	4	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	19.2	10.0	16.0	20.3	0	70	0-70 (70)	
		4점(우수)	0	0	0	1	4	3	1	1	0	0	1	1	0	0	0	0	0	0	0	0	28.3	25.0	27.0	12.0	15	55	15-55 (40)	
		3점(보통)	1	0	0	0	1	1	1	2	1	0	2	1	2	0	0	0	0	0	0	0	38.3	37.5	40.0	17.2	0	60	0-60 (60)	
		2점(미흡)	2	2	4	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	11.7	10.0	11.0	9.0	0	30	0-30 (30)	
		1점(매우 미흡)	9	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.5	0.0	1.5	4.8	0	15	0-15 (15)	
		합계																												
기초·기초미달	5점(매우 우수)	8	1	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8.3	0.0	4.5	15.9	0	55	0-55 (55)		
	4점(우수)	2	3	3	0	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	11.7	10.0	11.0	9.4	0	30	0-30 (30)		
	3점(보통)	0	0	1	1	1	0	3	1	3	2	0	0	0	0	0	0	0	0	0	0	31.7	32.5	32.5	11.1	10	45	10-45 (35)		
	2점(미흡)	0	1	2	0	0	0	1	3	3	1	1	0	0	0	0	0	0	0	0	0	31.3	35.0	32.0	14.2	5	50	5-50 (45)		
	1점(매우 미흡)	2	1	1	1	4	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	17.1	20.0	18.5	11.4	0	40	0-40 (40)		
	합계																													

[그림 III-4] 분포표

셋째, 이러한 정보를 토대로 소집단 토의를 거친 뒤 다시 기대 정답률 판정을 한다. 이러한 방식으로 세 번의 라운드를 거치면서, 마지막 3라운드에서 패널들이 최종 판정한 기대 정답률의 평균값을 토대로 최종 분할 점수를 산출한다.

3.2. 성취 수준 도출 결과

1) 영역별 분할 점수

수준 설정 협의회에 참여한 전문가 패널이 1·2차 협의회를 통하여 조사 영역별 문항에 대해 수준별로 기대 정답률(0~100%)을 평정한 결과 산출된 분할 점수는 다음과 같다.

<표 III-11> 분할 점수 산출

영역	수준	분할 점수	빈도	비율
듣기	4-3	128.0	1,232	40.6
	3-2	101.0	841	27.7
	2-1	66.6	779	25.7
	1-	-	181	6.0
읽기	4-3	243.9	1,649	33.0
	3-2	187.1	1,425	28.5
	2-1	105.6	1,243	24.9
	1-	-	683	13.7
문법·규범	4-3	117.0	1,482	29.6
	3-2	88.7	1,863	37.3
	2-1	53.5	1,175	23.5
	1-	-	480	9.6
듣기·말하기	4-3	55.4	696	23.3
	3-2	49.8	1,016	34.1
	2-1	43.0	762	25.6

	1-	-	508	17.0
말하기	4-3	122.3	538	18.1
	3-2	101.0	885	29.8
	2-1	75.3	959	32.3
	1-	-	590	19.9
쓰기	4-3	222.9	335	11.2
	3-2	182.2	791	26.4
	2-1	135.3	1,216	40.5
	1-	-	658	21.9

수준 설정 협의회에서 진행된 최종 3라운드 평정 결과를 바탕으로 수준별 최소 능력자의 기대 정답률 (기대 점수 확률) 평균값에 문항 및 준거별 배점을 곱해 분할 점수를 산출하였다. 1차 협의회에서 150점 만점인 듣기 영역의 4-3 수준 분할 점수는 128.0점, 3-2 수준 분할 점수는 101.0점, 2-1 수준 분할 점수는 66.6점으로 나타났다. 300점 만점인 읽기 영역에서 4-3 수준 분할 점수는 243.9점, 3-2 수준 분할 점수는 187.1점, 2-1 수준 분할 점수는 105.6점이었다. 150점 만점인 문법·규범 영역에서 4-3 수준 분할 점수는 117.0점, 3-2 수준 분할 점수는 88.7점, 2-1 수준 분할 점수는 53.5점으로 나타났다.

수행형 문항에 대해 실시한 2차 협의회 결과 결과는 다음과 같다. 60점 만점인 듣기·말하기 영역에서 4-3 수준 분할 점수는 55.4점, 3-2 수준 분할 점수는 49.8점, 2-1 수준 분할 점수는 43.0점이었다. 150점 만점인 말하기 영역에서 4-3 수준 분할 점수는 122.3점, 3-2 수준 분할 점수는 101.0점, 2-1 수준 분할 점수는 75.3점이었다. 300점 만점인 쓰기 영역에서 4-3 수준 분할 점수는 222.9점, 3-2 수준 분할 점수는 182.2점, 2-1 수준 분할 점수는 135.3점으로 나타났다.

2) 평정 신뢰도 분석

협의회에 참여한 12명의 전문가 패널의 의견이 라운드를 거치면서 조율되어 일치하는 방향으로 조정 이 되었는지를 확인하고자 기대 정답률 평정값의 범위를 통해 평정 신뢰도를 확인하였다.

<표 III-12> 평정 라운드별 기대 정답률 평정 범위(객관식 문항)

영역	수준	1 라운드			2 라운드			3 라운드		
		최소	최대	범위	최소	최대	범위	최소	최대	범위
듣기 (15문항)	4-3	60.0	93.7	33.7	78.3	92.7	14.3	78.7	91.7	13.0
	3-2	34.3	80.3	46.0	55.3	80.7	25.3	57.7	77.3	19.7
	2-1	6.3	51.7	45.3	16.3	63.0	46.7	33.0	57.7	24.7
읽기 (20문항)	4-3	59.3	96.0	36.8	72.3	92.8	20.5	72.5	88.0	15.5
	3-2	31.0	81.3	50.3	51.5	75.3	23.8	54.5	70.5	16.0
	2-1	6.5	59.3	52.8	15.3	50.8	35.5	25.0	44.0	19.0
문법· 규범 (13문항)	4-3	56.5	92.3	35.8	69.2	88.5	19.2	72.3	85.0	12.7
	3-2	30.4	76.5	46.2	43.8	71.2	27.3	51.2	68.8	17.7
	2-1	7.3	54.2	46.9	12.7	53.5	40.8	26.9	43.8	16.9

<표 III-13> 평정 라운드별 기대 정답률 평정 범위(수행형 문항)

영역	수준	척도	1 라운드			2 라운드			3 라운드		
			최소	최대	범위	최소	최대	범위	최소	최대	범위
듣기·말하기 (2문항, 4준거)	4-3	3	40.6	89.4	48.8	52.5	92.5	40.0	71.9	89.4	17.5
		2	10.0	57.5	47.5	7.5	42.5	35.0	10.6	26.9	16.3
		1	0.0	10.6	10.6	0.0	5.6	5.6	0.0	5.0	5.0
	3-2	3	10.6	57.5	46.9	26.9	82.5	55.6	41.3	68.1	26.9
		2	31.3	76.3	45.0	16.9	65.0	48.1	28.8	49.4	20.6
		1	2.5	26.9	24.4	0.0	16.9	16.9	0.0	10.6	10.6
	2-1	3	0.0	31.9	31.9	11.9	68.8	56.9	26.3	50.0	23.8
		2	16.3	58.8	42.5	20.6	60.0	39.4	32.5	53.8	21.3
		1	25.6	81.3	55.6	7.5	45.6	38.1	10.0	31.9	21.9
말하기 (3문항, 4준거)	4-3	5	10.8	69.6	58.8	16.7	82.5	65.8	22.1	50.4	28.3
		4	20.0	57.5	37.5	15.8	61.3	45.4	24.2	50.8	26.7
		3	3.3	45.8	42.5	0.0	36.7	36.7	12.1	30.8	18.8
		2	0.0	10.8	10.8	0.0	9.6	9.6	0.0	10.0	10.0
		1	0.0	4.6	4.6	0.0	1.7	1.7	0.0	5.0	5.0
	3-2	5	0.4	36.7	36.3	0.0	72.5	72.5	7.5	28.3	20.8
		4	12.5	45.8	33.3	15.0	43.8	28.8	17.9	35.8	17.9
		3	16.3	64.2	47.9	0.8	59.6	58.8	25.8	49.2	23.3
		2	2.9	33.3	30.4	0.0	30.8	30.8	6.7	22.5	15.8
		1	0.0	12.5	12.5	0.0	10.8	10.8	0.0	10.8	10.8
	2-1	5	0.0	8.8	8.8	0.0	60.0	60.0	0.0	13.8	13.8
		4	0.0	17.9	17.9	0.0	31.3	31.3	4.2	20.0	15.8
		3	10.0	36.7	26.7	10.0	43.8	33.8	24.2	39.6	15.4
		2	29.6	69.6	40.0	1.3	62.5	61.3	25.0	46.3	21.3
		1	7.5	46.3	38.8	0.0	38.3	38.3	5.4	23.3	17.9
쓰기 (4문항, 3~4준거)	4-3	5	8.1	67.3	59.2	11.2	45.4	34.2	14.2	36.2	21.9
		4	18.8	59.2	40.4	21.5	58.5	36.9	27.3	52.7	25.4
		3	6.9	53.5	46.5	14.2	49.6	35.4	15.4	43.5	28.1
		2	0.0	16.2	16.2	0.0	16.2	16.2	0.4	15.8	15.4
		1	0.0	6.2	6.2	0.0	3.5	3.5	0.0	5.0	5.0
	3-2	5	0.0	35.4	35.4	0.0	29.6	29.6	1.2	24.6	23.5
		4	8.5	43.8	35.4	10.4	38.5	28.1	14.6	37.7	23.1
		3	18.8	67.3	48.5	24.2	63.8	39.6	27.7	55.8	28.1
		2	5.4	37.7	32.3	5.8	36.2	30.4	7.7	30.8	23.1
		1	0.0	14.6	14.6	0.0	12.7	12.7	0.0	11.2	11.2
	2-1	5	0.0	13.1	13.1	0.0	15.8	15.8	0.0	13.1	13.1
		4	0.0	21.9	21.9	0.0	28.5	28.5	0.8	28.1	27.3
		3	11.5	46.5	35.0	13.8	43.5	29.6	19.6	39.6	20.0
		2	20.8	67.7	46.9	20.8	66.2	45.4	21.5	50.8	29.2
		1	6.9	43.8	36.9	7.3	38.8	31.5	8.8	31.9	23.1

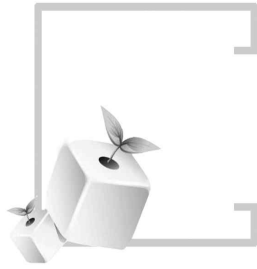
전체 12명의 전문가 패널이 듣기 15문항, 읽기 20문항, 문법·규범 13문항, 듣기·말하기 2문항(4개 준거, 3점 척도), 말하기 3문항(4개 준거, 5점 척도), 쓰기 4문항(3~4개 준거, 5점 척도)에 대해 최소 능력 보유자의 기대 정답률을 0~100%로 평정한 결과 중 항목별로 최솟값과 최댓값의 차이를 확인한 결과, 라운드를 거듭하면서 전반적으로 최소 기대 정답률과 최대 기대 정답률의 차이가 줄어들어 전문가 패널 간 일치도가 향상된 것으로 나타났다.

객관식 문항의 평정 결과, 듣기 영역 15개 문항에서는 1~3라운드에 걸쳐 4-3 수준(4 수준 최소 능력 보유자)의 범위가 33.7, 14.3, 13.0으로 줄어들었으며, 3-2 수준(2 수준 최소 능력 보유자)의 범위가 46.0, 25.3, 19.7로 줄어든 것으로 나타났다. 2-1 수준의 경우에는 범위가 45.3, 46.7, 24.7로 2라운드에서 소폭 범위가 커졌다가 3라운드에서 대폭 차이가 줄어든 것으로 나타났다. 읽기 영역 20개 문항에서는 1~3라운드에 걸쳐 4-3 수준의 범위가 36.8, 20.5, 15.5로 줄어들었으며, 3-2 수준에서도 50.3, 23.8, 16.0으로 범위가 줄어들었으며, 2-1 수준에서도 52.8, 35.5, 19.0으로 일관되게 줄어들었다. 문법·규범 영역 13개 문항에서는 1~3라운드에 걸쳐 4-3 수준의 범위가 35.8, 19.2, 12.7로 줄어들었고, 3-2 수준이 46.2, 27.3, 17.7로 줄어들었으며, 2-1 수준에서는 46.9, 40.8, 16.9로 범위가 일관되게 줄어든 것으로 나타났다.

수행형 문항의 평정 결과, 듣기·말하기 영역 2개 문항, 4개 준거, 3점 척도에서 4-3 수준의 1~3라운드 범위 평균은 35.6, 26.9, 12.9이었으며, 3-2 수준의 범위 평균은 38.8, 40.2, 19.4, 2-1 수준의 범위 평균은 43.3, 44.8, 22.3으로 모두 3라운드에서 범위가 대폭 줄어든 것으로 나타났다. 말하기 영역 3개 문항, 4개 준거, 5점 척도에서 4-3 수준의 1~3라운드 범위 평균은 30.8, 31.8, 17.8이었으며, 3-2 수준은 32.1, 40.3, 17.8이었고, 2-1 수준은 26.4, 44.9, 16.8로 모두 3라운드에서 범위가 대폭 줄어든 것으로 나타났다. 쓰기 영역 4개 문항, 3~4개 준거, 5점 척도에서 4-3 수준의 1~3라운드 범위 평균은 33.7, 25.2, 19.2였으며, 3-2 수준에서는 33.2, 28.1, 21.8, 2-1 수준에서는 30.8, 30.2, 22.5로 모두 3라운드에서 범위가 줄어든 것으로 나타났다.

기대 정답률 평정을 실시한 1차 협의회 영역 중에 1~3라운드 평정 사이의 범위 평균 변화 폭이 가장 크게 나타난 것은 읽기 영역의 3 수준(3-2) 최소 능력 보유자에 대한 평정이었으며, 듣기 영역의 4 수준 최소 능력 보유자에 대한 평정과 2 수준 최소 능력 보유자에 대한 평정은 상대적으로 범위 평균 변화의 폭이 작았던 것으로 나타났다. 척도별 기대 확률 평정을 실시한 2차 협의회 영역 중에 1라운드~3라운드 평정 사이의 범위 평균 변화 폭이 가장 크게 나타난 것은 듣기·말하기의 2 수준(2-1) 최소 능력 보유자에 대한 평정이었으며, 듣기·말하기 영역은 다른 다분 문항이 사용된 영역에 비해 전체 수준에서 1~3라운드 범위 평균 차이가 큰 것으로 나타났다.

이상의 수준 설정 협의회에서의 평정 범위에 대한 분석 결과를 통해 이번 2025년 국민의 국어능력 실태 조사의 조사 영역별 성취 수준 설정 결과, 전문가들의 의견이 수렴하는 경향이 있음을 확인하였으며, 따라서 도출된 각 영역의 수준별 분할 점수를 신뢰할 수 있다고 판단된다.



제 4 장

제3차 실태 조사 6개 영역 종합 결과 분석



4.1. 성취 수준 재도출에 따른 분석 결과

제3차 조사의 듣기, 듣기·말하기, 말하기 영역 조사 결과에 대한 기본적인 사항은 국립국어원·서울대학교 국어교육연구소(2023)에, 읽기 영역, 문법·규범 영역 결과에 대한 기본적인 사항은 국립국어원·서울대학교 국어교육연구소(2024)에 제시되어 있으나, 2025년 연구에서 성취 수준을 재도출하였기 때문에 성취 수준의 등급별 인원 분포와 그에 따른 통계 분석 결과를 새로 제시할 필요가 있다. 종합 분석이라는 점을 고려하여 능력 점수(T점수)를 기준으로 한 분석 결과 중 일부는 기존과 동일하더라도 한 번 더 제시하였고, 그에 더해 성취 수준 재도출에 따른 통계 분석 결과를 함께 제시한다.

1) 듣기, 듣기·말하기, 말하기 영역 결과

(1) 듣기 영역 결과

듣기 영역의 분석 대상은 3,033명이고 원점수의 만점은 150점으로, 조사 결과 원점수 평균은 114.11점, 표준편차는 27.02로 나타났다. 듣기 영역의 능력 점수를 T점수로 변환하면 평균 49.84점(100점 만점) 표준편차는 8.06이었고, 최솟값 21.79점, 최댓값 62.50점의 범위를 보이는 것으로 나타났다.

<표 IV-1> 듣기 영역의 능력 점수(T점수) 요약

영역	N	만점	평균	표준편차	최솟값	최댓값
듣기	3,033	100	49.84	8.06	21.79	62.50

듣기 영역의 수준 등급별 인원 분포는 4수준이 1,232명(40.6%), 3수준이 841명(27.7%), 2수준이 779명(25.7%), 1수준이 181명(6.0%)으로 나타났다.

<표 IV-2> 듣기 영역의 수준 등급별 인원 분포

수준	인원(명)	비율(%)
4수준	1,232	40.6
3수준	841	27.7
2수준	779	25.7
1수준	181	6.0
계	3,033	100.0

이러한 듣기 영역에서의 배경 변인에 따른 차이 분석 결과는 다음과 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수는 남성(50.09)이 여성(49.59)보다 근소하게 높았지만 통계적으로 유의한 차이는 아닌 것으로 분석되었다($p=.09$). 이러한 결과는 듣기 영역의 국민의 국어능력에서 남성과 여성 간에 차이가 없을 가능성이 큼을 시사한다.

<표 IV-3> 듣기 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성		t	df	p
	평균	SD	평균	SD			
성별	50.09	8.00	49.59	8.12	1.07	3,031	.09

성별에 따른 수준 등급별 인원 분포를 비교한 결과, 남성과 여성의 1~4수준 비율은 대체로 비슷하였고, 실제로 비율 차이가 통계적으로 유의하지 않은 것으로 분석되었다(Wilcoxon 검정: $p=.29$). 이러한 결과 또한 듣기 영역의 국민의 국어능력에서 남성과 여성 간에 차이가 없을 가능성이 큼을 시사한다.

<표 IV-4> 듣기 영역 수준 등급별 인원 분포의 성별 차이

변인	수준	남		여		순위합 검정 결과	
		명	%	명	%	W	p
성별	4수준	636	41.6	596	39.6	1,125,704	.29
	3수준	419	27.4	422	28.0		
	2수준	381	24.9	398	26.5		
	1수준	92	6.02	89	5.9		

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 20대(52.21), 30대(51.98), 40대(50.55), 50대(48.93), 60대(45.64) 순으로 높게 나타나 연령대가 높을수록 능력 점수는 낮아지는 경향성이 엿보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 30대 간 차이를 제외한 모든 연령대 간 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 연령대가 높을수록 듣기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-5> 듣기 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		F	df1	df2	p
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	52.21	7.98	51.98	7.73	50.55	7.29	48.93	7.80	45.64	7.70	61.79	4	3,028	<.001

연령대에 따른 수준 등급별 인원 분포를 비교한 결과, 20대에서 60대로 연령대가 높아질수록 특히 4수준 비율이 줄어드는 패턴($53.8\% > 51.3\% > 43.4\% > 36.1\% > 19.2\%$)이 뚜렷하게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 30대, 40대와 50대 간 차이를 제외한 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 연령대가 높을수록 듣기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-6> 듣기 영역 수준 등급별 인원 분포의 연령대별 차이

변인	수준	20대		30대		40대		50대		60대		순위합 검정 결과		
		명	%	명	%	명	%	명	%	명	%	χ^2	df	p
연령대	4수준	321	53.8	302	51.3	272	43.4	220	36.1	117	19.2	241.11	4	<.001
	3수준	141	23.6	155	26.3	181	28.9	188	30.8	176	28.9			
	2수준	112	18.8	114	19.4	154	24.6	160	26.2	239	39.2			
	1수준	23	3.9	18	3.0	20	3.2	42	6.9	78	12.8			

다음으로, 거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 대도시(50.59), 중소 도시(50.36), 읍면 지역(46.96)의 순으로 평균 점수가 높게 나타나 거주 지역의 규모가 클수록 듣기 영역 능력 점수가 높은 경향성이 엿보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 읍면 지역, 중소 도시와 읍면 지역의 차이가 유의한 반면, 대도시와 중소 도시의 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과는 읍면 지역의 거주자들이 다른 지역의 거주자들에 비해 듣기 영역의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-7> 듣기 영역 능력 점수의 지역 규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
지역 규모	50.59	7.82	50.36	7.60	46.96	8.95	44.12	2	3,030	<.001

거주 지역의 규모에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 대도시나 중소 도시에 비해 읍면 지역에서 4수준 비율이 10%p 이상 낮게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 읍면 지역, 중소 도시와 읍면 지역의 차이가 유의한 반면, 대도시와 중소 도시의 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과 또한 읍면 지역의 거주자들이 다른 지역의 거주자들에 비해 듣기 영역의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-8> 듣기 영역 수준 등급별 인원 분포의 지역 규모별 차이

변인	수준	대도시		중소 도시		읍면 지역		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
지역 규모	4수준	554	43.8	513	42.0	165	30.1	64.22	2	<.001
	3수준	350	27.7	353	28.9	138	25.2			
	2수준	303	24.0	311	25.5	165	30.1			
	1수준	58	4.6	43	3.5	80	14.6			

다음으로, 직업군에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 정신노동(51.76), 기타(49.37), 육체노동(48.52)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로

분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타의 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 직업군에 따라 듣기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 듣기 영역의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-9> 듣기 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
직업군	51.76	7.53	48.52	7.90	49.37	8.46	46.97	2	3,030	<.001

직업군에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 정신노동에 비해 육체노동이 4수준 비율이 낮고, 2수준 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 직업군에 따라 듣기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 듣기 영역의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-10> 듣기 영역 수준 등급별 인원 분포의 직업군별 차이

변인	수준	정신노동		육체노동		기타		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
직업군	4수준	514	51.0	379	32.5	339	39.5	93.49	2	<.001
	3수준	271	26.9	347	29.7	223	26.0			
	2수준	190	18.8	353	30.2	236	27.5			
	1수준	33	3.3	88	7.5	60	7.0			

다음으로, 학력에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 대재 이상(51.40), 고졸(47.84), 고졸 미만(43.45)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 학력이 높을수록 듣기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-11> 듣기 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	43.45	7.52	47.84	8.12	51.40	7.56	137.71	2	3,030	<.001

학력에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 학력이 높을수록(즉, 대재 이상, 고졸, 고졸

미만) 4수준 비율이 높게 나타났는데(즉, 순서대로 48.4% > 30.1% > 11.1%), 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 학력이 높을수록 듣기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-12> 듣기 영역 수준 등급별 인원 분포의 학력별 차이

변인	수준	고졸 미만		고졸		대재 이상		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
학력	4수준	22	11.1	267	30.1	943	48.4	227.67	2	<.001
	3수준	52	26.1	261	29.5	528	27.1			
	2수준	85	42.7	280	31.6	414	21.3			
	1수준	40	20.1	78	8.8	63	3.2			

한편 듣기 능력 수준을 예측하는 인구통계학적 변인을 확인하기 위하여 중다회귀분석 및 단계별 회귀분석을 수행하였다. 중다회귀분석을 수행하기 위한 예측 변수로는 성별, 연령, 지역 규모, 직업군, 학력 등 주요 인구통계학적 변인을 투입하였다. 이때, 연령 변인은 연속형 변수로 처리하였으며, 나머지 변인들은 범주형 변수로 처리하여 더미 변수로 변환하였다. 반응 변수로는 문항 반응 이론 분석에 따라 추정된 듣기 영역의 ‘능력 점수(T점수)’를 사용하였다.

회귀 모형에 투입할 예측 변수를 선정하고자 R의 step() 함수를 이용하여 단계별 변수 선택을 수행하였다. 이때, 성별, 연령, 지역 규모, 직업군, 학력 등 투입한 모든 인구통계학적 변인이 듣기 영역 능력 점수(T점수)를 통계적으로 유의미하게 예측하는 것으로 나타났다. 이때, 통계적으로 유의미한 인구통계학적 변인을 포함하고 설명력이 높은 회귀 모형을 최종 모형으로 선택하였다. 최종 선택된 회귀 모형의 중다회귀분석 결과는 다음과 같다.

<표 IV-13> 인구통계학적 변인에 따른 듣기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	54.800	NA	0.602	91.0	< 2e-16 ***
연령	-0.133	-0.222	0.011	-11.9	< 2e-16 ***
읍면 지역	-2.706	-0.129	0.355	-7.6	3e-14 ***
정신노동	1.211	0.0707	0.312	3.9	1e-14 ***
고졸 미만	-2.397	-0.074	0.604	-4.0	7e-05 ***
대재 이상	1.853	0.110	0.336	5.5	4e-08 ***
잔여 표준오차	7.5 (df=3,027)				
R^2	0.15	조정된 R^2		0.14	
F 통계량	100 (5, df=3,027)	p		< 2.2e-16 ***	

*** $p < .001$ ** $p < .01$ * $p < .05$

최종 회귀 모형의 설명력(R^2)은 15.0%로 비교적 낮게 나타났으며 통계적으로 유의미한 수준이었다($p < .001$). 연령이 높아질수록 듣기 능력 점수(T점수)가 약 0.13점씩 낮아질 것으로 예측되었다($\eta^2 = 0.10$; 중간 효과 크기). 다음으로, 지역 규모가 읍면 지역인 경우에는 다른 지역의 경우에 비해 듣기 능력 점

수(T점수)가 약 2.71점 정도 낮아질 것으로 예측되었다($\eta^2=0.02$; 작은 효과 크기). 다음으로, 직업군이 정신노동인 경우에는 다른 직업군의 경우에 비해 듣기 능력 점수(T점수)가 약 1.21점 정도 높을 것으로 예측되었다($\eta^2<0.01$; 효과 크기 거의 없음). 한편, 학력이 고졸 미만인 경우 다른 학력군의 경우에 비해 듣기 능력 점수(T점수)가 2.40점 정도 낮아질 반면($\eta^2<0.01$; 효과 크기 거의 없음), 대재 이상인 경우에는 다른 학력군의 경우에 비해 1.85점 정도 높아질 것으로 예측되었다($\eta^2<0.01$; 효과 크기 거의 없음).

(2) 듣기·말하기 영역 결과

듣기·말하기 영역의 분석 대상은 2,982명(3,033명에서 결측치 51명 제외)이고 원점수의 만점은 60점으로, 조사 결과 원점수 평균은 49.70점, 표준편차는 6.95로 나타났다. 듣기·말하기 영역의 능력 점수를 T점수로 변환하면 평균 49.80점(100점 만점) 표준편차는 8.58이었고, 최솟값 19.38점, 최댓값 61.25점의 범위를 보이는 것으로 나타났다.

<표 IV-14> 듣기·말하기 영역의 능력 점수(T점수) 요약

영역	N	만점	평균	표준편차	최솟값	최댓값
듣기·말하기	2,982	100	49.80	8.58	19.38	61.25

듣기·말하기 영역의 수준 등급별 인원 분포는 4수준이 696명(23.3%), 3수준이 1,016명(34.1%), 2수준이 762명(25.6%), 1수준이 508명(17.0%)으로 나타났다.

<표 IV-15> 듣기·말하기 영역의 수준 등급별 인원 분포

수준	인원(명)	비율(%)
4수준	696	23.3
3수준	1,016	34.1
2수준	762	25.6
1수준	508	17.0
계	2,982	100

이러한 듣기·말하기 영역에서의 배경 변인에 따른 차이 분석 결과는 다음과 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수는 여성(49.84)이 남성(49.77)보다 매우 근소하게 높았지만 통계적으로 유의한 차이는 아닌 것으로 분석되었다($p=.81$). 이러한 결과는 듣기·말하기 영역의 국민의 국어능력에서 남성과 여성 간에 차이가 없을 가능성이 큼을 시사한다.

<표 IV-16> 듣기·말하기 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성		t	df	p
	평균	SD	평균	SD			
성별	49.77	8.46	49.84	8.71	.242	2,980	.81

성별에 따른 수준 등급별 인원 분포를 비교한 결과, 남성과 여성의 1~4수준 비율은 대체로 비슷하였고, 실제로 비율 차이가 통계적으로 유의하지 않은 것으로 분석되었다(Wilcoxon 검정: $p=.28$). 이러한 결과 또한 듣기·말하기 영역의 국민의 국어능력에서 남성과 여성 간에 차이가 없을 가능성이 큼을 시사한다.

<표 IV-17> 듣기·말하기 영역 수준 등급별 인원 분포의 성별 차이

변인	수준	남		여		순위합 검정 결과	
		명	%	명	%	W	p
성별	4수준	336	22.3	360	24.4	1,136,064	.28
	3수준	514	34.1	502	34.0		
	2수준	402	26.7	360	24.4		
	1수준	254	16.9	254	17.2		

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 20대(51.00), 30대(50.84), 40대(50.27), 50대(49.45), 60대(47.51) 순으로 평균 점수가 높게 나타나 연령대가 높을수록 듣기·말하기 영역 능력 점수는 낮아지는 경향성이 엿보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 50대, 20대와 60대, 30대와 60대, 40대와 60대, 50대와 60대의 차이가 유의한 것으로 분석되었다. 이러한 결과는 연령대가 높을수록 듣기·말하기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-18> 듣기·말하기 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		F	df1	df2	p
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	51.00	8.12	50.84	7.81	50.27	8.16	49.45	8.84	47.51	9.42	16.66	4	2,977	<.001

연령대에 따른 수준 등급별 인원 분포를 비교한 결과, 20대에서 60대로 연령대가 높아질수록 특히 4수준 비율이 감소하는 패턴(28.9 ≍⁷⁾ 29.4% > 24.9% > 19.5 > 14.3%)이 뚜렷하게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 30대, 20대와 40대, 30대와 40대 간 차이를 제외

7) 이와 관련하여 30대의 4수준 비율이 20대의 4수준 비율보다 약간 높게 나타났으나 이는 통계적으로 유의하지는 않으므로 대략적으로 같음을 나타내는 기호(≍)를 사용하여 제시하였다.

한 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 연령대가 높을수록 듣기·말하기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-19> 듣기·말하기 영역 수준 등급별 인원 분포의 연령대별 차이

변인	수준	20대		30대		40대		50대		60대		순위합 검정 결과		
		명	%	명	%	명	%	명	%	명	%	χ^2	df	p
연령대	4수준	170	28.9	170	29.4	154	24.9	117	19.5	85	14.3	54.92	4	<.001
	3수준	217	36.8	209	36.2	212	34.3	199	33.1	179	30.0			
	2수준	125	21.2	128	22.2	168	27.2	172	28.6	169	28.4			
	1수준	77	13.1	71	12.3	84	13.6	113	18.8	163	27.4			

다음으로, 거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 대도시(50.20), 중소 도시(49.69), 읍면 지역(49.13)의 순으로 평균 점수가 높게 나타나 거주 지역의 규모가 클수록 능력 점수가 높은 경향성이 엿보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.05$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 읍면 지역의 차이가 유의한 것으로 분석되었다. 이러한 결과는 거주 지역의 규모가 클수록 듣기·말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-20> 듣기·말하기 영역 능력 점수의 지역 규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
지역 규모	50.20	8.34	49.69	8.73	49.13	8.77	3.142	2	2,979	<.05

거주 지역의 규모에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 대도시에 비해 읍면 지역에서 4수준 비율이 10%p 이상 낮게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p=.049$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 읍면 지역, 중소 도시와 읍면 지역의 차이가 유의한 반면, 대도시와 중소 도시의 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과 또한 읍면 지역의 거주자들이 다른 지역의 거주자들에 비해 듣기 영역의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-21> 듣기·말하기 영역 수준 등급별 인원 분포의 지역 규모별 차이

변인	수준	대도시		중소 도시		읍면 지역		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
지역 규모	4수준	330	26.5	278	23.3	88	16.2	2	6.04	.049
	3수준	422	33.9	409	34.3	185	34.1			
	2수준	304	24.4	308	25.8	150	27.6			
	1수준	189	15.2	199	16.7	120	22.1			

다음으로, 직업군에 따른 능력 점수(T점수) 차이 분석 결과, 정신노동(50.77), 육체노동(49.34), 기타(49.30)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타의 차이가 유의한 것으로 분석되었다. 이러한 결과는 직업군에 따라 듣기·말하기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 듣기·말하기 영역의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-22> 듣기·말하기 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
직업군	50.77	7.92	49.34	8.91	49.30	8.79	9.47	2	2,979	<.001

직업군에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 정신노동에 비해 육체노동이 4수준 비율이 낮고, 2수준 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 육체노동과 기타 집단 간 차이를 제외한 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 직업군에 따라 듣기·말하기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 듣기·말하기 영역의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-23> 듣기·말하기 영역 수준 등급별 인원 분포의 직업군별 차이

변인	수준	정신노동		육체노동		기타		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
직업군	4수준	247	24.8	259	22.7	190	22.5	14.87	2	<.001
	3수준	379	38.0	352	30.9	285	33.8			
	2수준	248	24.9	292	25.6	222	26.3			
	1수준	124	12.4	237	20.8	147	17.4			

다음으로, 학력에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 대재 이상(50.40), 고졸(49.15), 고졸 미만(45.99)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 학력이 높을수록 듣기·말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-24> 듣기·말하기 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	45.99	0.71	49.15	0.30	50.40	0.19	24.74	2	2,979	<.001

학력에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 학력이 높을수록(즉, 대재 이상, 고졸, 고졸 미만) 4수준 비율이 높게 나타났는데(즉, 순서대로 25.7% > 20.1% > 15.0%), 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 학력이 높을수록 듣기·말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-25> 듣기·말하기 영역 수준 등급별 인원 분포의 학력별 차이

변인	수준	고졸 미만		고졸		대재 이상		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
학력	4수준	29	15.0	174	20.1	493	25.7	38.44	2	<.001
	3수준	48	24.9	271	31.2	697	36.3			
	2수준	52	26.9	237	27.3	473	24.6			
	1수준	64	33.2	186	21.4	258	13.4			

한편 듣기·말하기 능력 수준을 예측하는 인구통계학적 변인을 확인하기 위하여 중다회귀분석 및 단계별 회귀분석을 수행하였다. 중다회귀분석을 수행하기 위한 예측 변수로는 성별, 연령, 지역 규모, 직업군, 학력 등 주요 인구통계학적 변인을 투입하였다. 이때, 연령 변인은 연속형 변수로 처리하였으며, 나머지 변인들은 범주형 변수로 처리하여 더미 변수로 변환하였다. 반응 변수로는 문항 반응 이론 분석에 따라 추정된 듣기·말하기 영역의 ‘능력 점수(T점수)’를 사용하였다.

회귀 모형에 투입할 예측 변수를 선정하고자 R의 step() 함수를 이용하여 단계별 변수 선택을 수행하였다. 이때, 연령, 지역 규모, 직업군, 학력 등은 듣기·말하기 영역 능력 점수(T점수)를 통계적으로 유의미하게 예측하는 변인으로 나타난 반면, 성별은 통계적으로 유의미한 설명력을 보이지 못했다. 이에 따라 통계적으로 유의미한 인구통계학적 변인만을 포함하고 설명력이 높은 회귀 모형을 최종 모형으로 선택하였다. 최종 선택된 회귀 모형의 중다회귀분석 결과는 다음과 같다.

<표 IV-26> 인구통계학적 변인에 따른 듣기·말하기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	52.783	NA	0.595	88.699	< 2e-16 ***
연령	-0.076	-0.120	0.012	-6.208	6.45e-10 **
대도시	0.591	0.033	0.314	1.879	0.060.
정신노동	0.901	0.050	0.335	2.688	0.007 **
고졸 미만	-2.189	-0.063	0.674	-3.247	0.001 **
잔여 표준오차	8.5 ($df=2,977$)				
R^2	0.03	조정된 R^2		0.029	
F 통계량	23 (4, $df=2,977$)		p		< 2e-16

*** $p < .001$ ** $p < .01$ * $p < .05$. $p < .1$

최종 회귀 모형의 설명력(R^2)은 3.0%로 비교적 낮게 나타났으며 통계적으로 유의미한 수준이었다($p < .001$). 연령이 높아질수록 듣기·말하기 능력 점수(T점수)가 약 0.08점 정도 감소할 것으로 예측되었다

($\eta^2=0.02$; 작은 효과 크기). 다음으로, 지역 규모가 대도시인 경우에는 다른 도시 유형의 경우에 비해 듣기·말하기 능력 점수(T점수)가 0.59점 높아질 것으로 예측되었고($\eta^2<0.01$; 효과 크기 거의 없음), 직업군이 정신노동인 경우에는 다른 직업군의 경우에 비해 듣기·말하기 능력 점수(T점수)가 0.90점 높아질 것으로 예측되었다($\eta^2<0.01$; 효과 크기 거의 없음). 한편, 학력이 고졸 미만인 경우 다른 학력군의 경우에 비해 듣기·말하기 능력 점수(T점수)가 2.19점 정도 낮아질 것으로 예측되었다($\eta^2<0.01$; 효과 크기 거의 없음).

(3) 말하기 영역 결과

말하기 영역의 분석 대상은 2,972명(3,033명에서 결측치 61명 제외)이고 원점수의 만점은 150점으로, 조사 결과 원점수 평균은 97.32점, 표준편차는 25.63으로 나타났다. 말하기 영역의 능력 점수를 T점수로 변환하면 평균 50.15점(100점 만점), 표준편차는 9.11이었고, 최솟값 18.73점, 최댓값 76.98점의 범위를 보이는 것으로 나타났다.

<표 IV-27> 말하기 영역의 능력 점수(T점수) 요약

영역	N	만점	평균	표준편차	최솟값	최댓값
말하기	2,972	100	50.15	9.11	18.73	76.98

말하기 영역의 수준 등급별 인원 분포는 4수준이 538명(18.1%), 3수준이 885명(29.8%), 2수준이 959명(32.3%), 1수준이 590명(19.9%)으로 나타났다.

<표 IV-28> 말하기 영역의 수준 등급별 인원 분포

수준	인원(명)	비율(%)
4수준	538	18.1
3수준	885	29.8
2수준	959	32.3
1수준	590	19.9
계	2,972	100

이러한 말하기 영역에서의 배경 변인에 따른 차이 분석 결과는 다음과 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수는 여성(50.64)이 남성(49.67)보다 높았으며 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.005$). 이러한 결과는 말하기 영역의 국민의 국어능력에서 여성이 남성보다 그 능력이 더 높을 가능성이 큼을 시사한다.

<표 IV-29> 말하기 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성		<i>t</i>	<i>df</i>	<i>p</i>
	평균	SD	평균	SD			
성별	49.67	8.87	50.64	9.32	2.91	2,970	<.005

성별에 따른 수준 등급별 인원 분포를 비교한 결과, 여성이 남성보다 4수준과 1수준의 비율이 높고, 3수준과 2수준의 비율은 낮았는데, 이러한 비율 차이가 통계적으로는 유의하지 않은 것으로 분석되었다(Wilcoxon 검정; $p=.09$).

<표 IV-30> 말하기 영역 수준 등급별 인원 분포의 성별 차이

변인	수준	남		여		순위합 검정 결과	
		명	%	명	%	<i>W</i>	<i>p</i>
성별	4수준	239	15.9	299	20.3	1,141,779	.09
	3수준	461	30.7	424	28.8		
	2수준	505	33.7	454	30.8		
	1수준	295	19.7	295	20.0		

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 30대(52.25)에서 평균 점수가 가장 높았고 20대(52.21), 40대(51.17), 50대(48.65), 60대(46.43) 순으로 평균 점수가 높게 나타나 전반적으로 연령대가 높을수록 말하기 영역 능력 점수는 낮아지는 경향성이 엿보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대, 30대, 40대 간 차이를 제외한 다른 연령대 간 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 연령대가 높을수록 말하기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-31> 말하기 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	52.21	9.06	52.25	9.02	51.17	8.79	48.65	8.48	46.43	8.78	48.6	4	2,967	<.001

연령대에 따른 수준 등급별 인원 분포를 비교한 결과, 20대에서 60대로 연령대가 높아질수록 특히 4수준 비율이 줄어드는 패턴($28.4\% > 24.4\% > 19.3\% > 10.5\% > 8.0\%$)이 뚜렷하게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 30대, 30대와 40대 간 차이를 제외한 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 연령대가 높을수록 말하기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-32> 말하기 영역 수준 등급별 인원 분포의 연령대별 차이

변인	수준	20대		30대		40대		50대		60대		순위합 검정 결과		
		명	%	명	%	명	%	명	%	명	%	χ^2	df	p
연령대	4수준	167	28.4	141	24.4	120	19.3	63	10.5	47	8.0	223.76	4	<.001
	3수준	188	32.0	207	35.8	203	32.6	177	29.6	110	18.8			
	2수준	145	24.7	151	26.1	202	32.5	232	38.8	229	39.1			
	1수준	88	15.0	79	13.7	97	15.6	126	21.1	200	34.1			

다음으로, 거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 대도시(51.53), 중소 도시(50.1), 읍면 지역(47.07)의 순으로 평균 점수가 높게 나타나 거주 지역의 규모가 클수록 말하기 영역 능력 점수가 높은 경향성이 엿보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 거주 지역의 규모가 클수록 말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-33> 말하기 영역 능력 점수의 지역규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
지역 규모	51.53	8.27	50.1	9.26	47.07	9.82	46.41	2	2,969	<.001

거주 지역의 규모에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 대도시나 중소 도시에 비해 읍면 지역에서 4수준과 3수준 비율이 낮게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 거주 지역의 규모가 클수록 말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-34> 말하기 영역 수준 등급별 인원 분포의 지역 규모별 차이

변인	수준	대도시		중소 도시		읍면 지역		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
지역 규모	4수준	265	21.4	206	17.2	67	12.4	97.93	2	<.001
	3수준	410	33.1	364	30.5	111	20.6			
	2수준	404	32.6	373	31.2	182	33.8			
	1수준	159	12.8	252	21.1	179	33.2			

다음으로, 직업군에 따른 능력 점수(T점수) 차이 분석 결과, 정신노동(52.31), 기타(50.08), 육체노동(48.3)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타, 육체노동과 기타 간 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 직업군에 따라 말하기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 말하기

영역의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-35> 말하기 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
직업군	52.31	8.67	48.3	8.93	50.08	9.29	53.53	2	2,969	<.001

직업군에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 정신노동의 경우 육체노동 및 기타와 비교하여 4수준과 3수준의 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 나타났다. 이러한 결과 또한 직업군에 따라 말하기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 말하기 영역 국어 능력이 높을 가능성이 큼을 시사한다.

<표 IV-36> 말하기 영역 수준 등급별 인원 분포의 직업군별 차이

변인	수준	정신노동		육체노동		기타		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
직업군	4수준	233	23.3	151	13.3	154	18.4	131.99	2	<.001
	3수준	368	36.9	277	24.4	240	28.6			
	2수준	288	28.9	390	34.3	281	33.5			
	1수준	109	10.9	318	28.0	163	19.5			

다음으로, 학력에 따른 능력 점수(T점수) 차이 분석 결과, 대재 이상(51.72), 고졸(47.91), 고졸 미만(44.39)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과는 학력이 높을수록 말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-37> 말하기 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	44.39	8.87	47.91	8.76	51.72	8.84	98.63	2	2,969	<.001

학력에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 학력이 높을수록(즉, 대재 이상, 고졸, 고졸 미만) 4수준과 3수준의 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 나타났다. 이러한 결과 또한 학력이 높을수록 말하기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-38> 말하기 영역 수준 등급별 인원 분포의 학력별 차이

변인	수준	고졸 미만		고졸		대재 이상		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
학력	4수준	5	2.7	102	11.8	431	22.5	209.6	2	<.001
	3수준	36	19.1	214	24.7	635	33.1			
	2수준	58	30.9	309	35.6	592	30.9			
	1수준	89	47.3	242	27.9	259	13.5			

한편, 말하기 능력 수준을 예측하는 인구통계학적 변인을 확인하기 위하여 중다회귀분석 및 단계별 회귀분석을 수행하였다. 중다회귀분석을 수행하기 위한 예측 변수로는 성별, 연령, 지역 규모, 직업군, 학력 등 주요 인구통계학적 변인을 투입하였다. 이때, 연령 변인은 연속형 변수로 처리하였으며, 나머지 변인들은 범주형 변수로 처리하여 더미 변수로 변환하였다. 반응 변수로는 문항 반응 이론 분석에 따라 추정된 말하기 영역의 ‘능력 점수(T점수)’를 사용하였다.

회귀 모형에 투입할 예측 변수를 선정하고자 R의 step() 함수를 이용하여 단계별 변수 선택을 수행하였다. 이때, 성별, 연령, 지역 규모, 직업군, 학력 등 투입한 모든 인구통계학적 변인이 말하기 능력 점수(T점수)를 통계적으로 유의미하게 예측하는 것으로 나타났다. 통계적으로 유의미한 인구통계학적 변인을 포함하고 설명력이 높은 회귀 모형을 최종 모형으로 선택하였다. 최종 선택된 회귀 모형의 중다 회귀분석 결과는 다음과 같다.

<표 IV-39> 인구통계학적 변인에 따른 말하기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	53.09	NA	0.806	65.85	< .001 ***
성별	1.598	0.088	0.325	4.91	< .001 ***
연령	-0.118	-0.174	0.013	-9.17	< .001 ***
대도시	1.335	0.072	0.346	3.86	< .001 ***
읍면 지역	-2.168	-0.092	0.445	-4.87	< .001 ***
육체노동	-0.869	-0.046	0.401	-2.17	0.03 *
정신노동	1.488	0.077	0.425	3.50	< .001 ***
대재 이상	2.007	0.106	0.394	5.09	< .001 ***
고졸 미만	-1.918	-0.051	0.711	-2.70	0.007 **
잔여 표준오차	8.52 (df=2,963)				
R^2	0.1259	조정된 R^2		0.1235	
F 통계량	53.32 (8, df=2,963)		p		< .001

*** p <.001 ** p <.01 * p <.05

최종 회귀 모형의 설명력(R^2)은 12.59%로 비교적 낮게 나타났으며 통계적으로 유의미한 수준이었다(p <.001). 성별이 여성일 경우에는 남성인 경우에 비해 말하기 능력 점수(T점수)가 약 1.6점 정도 높아질 것으로 예측되었다(η^2 <0.01; 효과 크기 거의 없음). 다음으로, 연령이 높아질수록 말하기 능력 점수(T점수)가 약 0.12점씩 낮아질 것으로 예측되었다(η^2 =0.07; 중간 효과 크기). 다음으로, 지역 규모가 대

도시인 경우에는 다른 도시 유형의 경우에 비해 말하기 능력 점수(T점수)가 약 1.34점 정도 높아질 것으로 예측되는 반면($\eta^2=0.02$; 작은 효과 크기), 지역 규모가 읍면 지역인 경우에는 다른 도시 유형에 비해 2.2점 정도 낮아질 것으로 예측되었다($\eta^2=0.01$; 작은 효과 크기). 다음으로, 직업군이 육체노동인 경우에는 다른 직업군의 경우에 비해 말하기 능력 점수(T점수)가 약 0.9점 정도 낮아질 반면($\eta^2=0.02$; 작은 효과 크기), 정신노동인 경우에는 다른 직업군의 경우에 비해 약 1.5점 정도 높아질 것으로 예측되었다($\eta^2<0.01$; 효과 크기 거의 없음). 한편, 학력이 대재 이상인 경우에는 다른 학력군의 경우에 비해 2점 정도 높아지는 반면($\eta^2=0.01$; 작은 효과 크기), 고졸 미만인 경우 다른 학력군의 경우에 비해 말하기 능력 점수(T점수)가 1.9점 정도 낮아질 것으로 예측된다($\eta^2<0.01$; 효과 크기 거의 없음).

2) 읽기 영역 결과

읽기 영역의 분석 대상은 5,000명이고 원점수의 만점은 300점으로, 조사 결과 원점수 평균은 197.93점, 표준편차는 69.36으로 나타났다. 읽기 영역의 능력 점수를 T점수로 변환하면 평균 49.53점(100점 만점) 표준편차는 9.10이었고, 최솟값 25.20점, 최댓값 67.33점의 범위를 보이는 것으로 나타났다.

<표 IV-40> 읽기 영역의 능력 점수(T점수) 요약

영역	N	만점	평균	표준편차	최솟값	최댓값
읽기	5,000	100	49.53	9.10	25.20	67.33

읽기 영역의 수준 등급별 인원 분포는 4수준이 1,649명(33.0%), 3수준이 1,425명(28.5%), 2수준이 1,243명(24.9%), 1수준이 683명(13.7%)으로 나타났다.

<표 IV-41> 읽기 영역의 수준 등급별 인원 분포

수준	인원(명)	비율(%)
4수준	1,649	33.0
3수준	1,425	28.5
2수준	1,243	24.9
1수준	683	13.7
계	5,000	100

읽기 영역에서 성별, 연령대, 지역 규모, 직업군, 학력과 같은 인구통계학적 변인에 따른 차이 분석 결과는 다음과 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수는 남성(49.55)이 여성(49.50)보다 근소하게 높았으나 통계적으로 유의한 차이는 아닌 것으로 확인되었다($p=.86$).

<표 IV-42> 읽기 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성		t	df	p
	평균	SD	평균	SD			
성별	49.55	9.16	49.50	9.03	-.18	4,998	.86

성별에 따른 수준 등급별 인원 분포를 비교한 결과, 남성과 여성의 1~4수준 비율은 대체로 비슷하였고, 실제로 비율 차이가 통계적으로 유의하지 않은 것으로 분석되었다(Wilcoxon 검정: $p=.29$). 이러한 결과 또한 읽기 영역의 국민의 국어능력에서 남성과 여성 간에 차이가 없을 가능성이 큼을 시사한다.

<표 IV-43> 읽기 영역 수준 등급별 인원 분포의 성별 차이

변인	수준	남		여		순위합 검정 결과	
		명	%	명	%	W	p
성별	4수준	859	33.8	790	32.1	3,072,448	.29
	3수준	730	28.7	695	28.3		
	2수준	588	23.1	655	26.6		
	1수준	364	14.3	319	13.0		

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 20대(51.19), 30대(50.75), 40대(49.88), 50대(48.97), 60대(47.44) 순으로 평균 점수가 높게 나타나 연령대가 높을수록 읽기 영역 능력 점수는 낮아지는 경향성이 보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 30대와 40대, 40대와 50대 간 차이를 제외한 모든 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-44> 읽기 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		F	df1	df2	p
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	51.19	9.46	50.75	9.41	49.88	9.12	48.97	8.62	47.44	8.56	26.62	4	4,995	<.001

연령대에 따른 수준 등급별 인원 분포를 비교한 결과, 20대에서 60대로 연령대가 높아질수록 특히 4수준 비율이 줄어드는 패턴(42.8% > 39.9% > 34.2% > 28.8% > 22.8%)이 뚜렷하게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 30대, 40대와 50대 간 차이를 제외한 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 연령대가 높을수록 읽기 영역의 국민의 국어능력이 낮을 가능성이 큼을 시사한다.

<표 IV-45> 읽기 영역 수준 등급별 인원 분포의 연령대별 차이

변인	수준	20대		30대		40대		50대		60대		순위합 검정 결과		
		명	%	명	%	명	%	명	%	명	%	χ^2	df	p
연령대	4수준	353	42.8	356	39.9	362	34.2	341	28.8	237	22.8	95.89	4	<.001
	3수준	204	24.8	232	26.0	310	29.2	373	31.5	306	29.4			
	2수준	166	20.1	193	21.6	245	23.1	309	26.1	330	31.7			
	1수준	101	12.3	111	12.4	143	13.5	160	13.5	168	16.1			

다음으로, 거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 중소 도시(49.72), 대도시(49.56), 읍면 지역(48.95)의 순으로 평균 점수가 높게 나타났으나, 이는 통계적으로 유의한 차이는 아닌 것으로 확인되었다($p=.12$).

<표 IV-46> 읽기 영역 능력 점수의 지역 규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
지역 규모	49.56	9.19	49.72	9.29	48.95	8.29	2.10	2	4,997	.12

거주 지역의 규모에 따른 수준 등급별 인원 분포를 비교한 결과, 대도시나 중소 도시에 비해 읍면 지역에서 4수준 비율이 약 4%p 이상 낮게 나타나고, 2수준의 비율이 10%p가량 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.05$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 중소 도시와 읍면 지역의 차이가 유의한 반면, 대도시와 중소 도시, 대도시와 읍면 지역의 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과는 읍면 지역의 거주자들이 다른 지역의 거주자들에 비해 읽기 영역의 국어능력이 다소 낮을 가능성이 큼을 시사한다.

<표 IV-47> 읽기 영역 수준 등급별 인원 분포의 지역 규모별 차이

변인	수준	대도시		중소 도시		읍면 지역		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
지역 규모	4수준	726	33.3	687	34.0	236	29.5	6.40	2	.04
	3수준	630	28.9	588	29.1	207	25.9			
	2수준	508	23.3	466	23.1	269	33.6			
	1수준	316	14.5	279	13.8	88	11.0			

직업군에 따른 능력 점수(T점수) 차이 분석 결과, 정신노동(51.18), 기타(49.08), 육체노동(48.63)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타 간 차이가 유의한 것으로 분석되었다.

<표 IV-48> 읽기 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
직업군	51.18	9.42	48.63	8.62	49.08	9.19	37.16	2	4,997	<.001

직업군에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 정신노동에 비해 육체노동이 4수준 비율이 낮고, 2수준 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 직업군에 따라 읽기 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 읽기 영역 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-49> 읽기 영역 수준 등급별 인원 분포의 직업군별 차이

변인	수준	정신노동		육체노동		기타		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
직업군	4수준	617	41.1	584	28.5	448	30.9	60.75	2	<.001
	3수준	418	27.8	621	30.3	386	26.6			
	2수준	274	18.3	556	27.1	413	28.5			
	1수준	192	12.8	289	14.1	202	13.9			

다음으로, 학력에 따른 능력 점수(T점수) 차이 분석 결과, 대재 이상(51.04), 고졸(48.15), 고졸 미만(45.33)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-50> 읽기 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	45.33	7.91	48.15	8.58	51.04	9.24	112.43	2	4,997	<.001

학력에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 학력이 높을수록(즉, 대재 이상, 고졸, 고졸 미만) 4수준 비율이 높게 나타났는데(즉, 순서대로 40.5% > 25.5% > 14.4%), 실제로 이러한 비율 차이

가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 학력이 높을수록 읽기 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-51> 읽기 영역 수준 등급별 인원 분포의 학력별 차이

변인	수준	고졸 미만		고졸		대재 이상		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
학력	4수준	66	14.4	438	25.5	1,145	40.5	198.38	2	<.001
	3수준	126	27.5	525	30.6	774	27.4			
	2수준	170	37.1	498	29.0	575	20.4			
	1수준	96	21.0	257	15.0	330	11.7			

한편 읽기 능력 수준을 설명하는 인구통계학적 변인을 확인하기 위하여 중다회귀분석 및 단계별 회귀 분석을 수행하였다. 중다회귀분석을 수행하기 위한 설명 변수로는 성별, 연령, 지역 규모, 직업군, 학력 등 주요 인구통계학적 변인을 투입하였다. 이때, 연령 변인은 연속형 변수로 처리하였으며, 나머지 변인들은 범주형 변수로 처리하여 더미 변수로 변환하였다. 반응 변수로는 문항 반응 이론 분석에 따라 추정된 읽기 영역의 ‘능력 점수(T점수)’를 사용하였다.

회귀 모형에 투입할 설명 변수를 선정하고자 R 프로그램에 기본적으로 내장되어 있는 stats 패키지에서 제공하는 step() 함수를 이용하여 단계별 변수 선택을 수행하였다. 연령, 지역 규모, 학력은 읽기 영역 점수(T점수)를 통계적으로 유의미하게 설명하는 변인으로 나타난 반면, 성별, 직업군은 통계적으로 유의미한 설명력을 보이지 못했다. 이에 따라 통계적으로 유의미한 인구통계학적 변인만을 포함하고 설명력이 높은 회귀 모형을 최종 모형으로 선택하였다. 최종 선택된 회귀 모형의 중다회귀분석 결과는 다음과 같다.

<표 IV-52> 인구통계학적 변인에 따른 읽기 능력 점수(T점수) 설명 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	50.344	NA	0.618	81.453	<2e-16***
연령	-0.035	-0.0515	0.011	-3.177	0.001**
육체노동	-0.631	-0.034	0.268	-2.352	0.0187*
대재 이상	2.294	0.125	0.308	7.444	4.16e-12***
고졸 미만	-2.625	-0.083	0.478	-5.493	4.38e-08***
잔여 표준오차	8.9 (df=4,995)				
R^2	0.046	조정된 R^2		0.046	
F 통계량	61 (5, df=4,995)	p		<2e-16	

*** $p < .001$ ** $p < .01$ * $p < .05$

최종 회귀 모형의 설명력(R^2)은 4.6%로 비교적 낮게 나타났지만 통계적으로 유의미한 수준이었다($p < .001$). 연령이 높아질수록 읽기 능력 점수(T점수)가 낮아질 것으로 예측되었다($\eta^2 = 0.02$; 작은 효과 크기). 직업군이 육체노동인 경우에는 다른 직업군에 비해 읽기 능력 점수(T점수)가 0.63점 정도 낮아질 것으로 예측되었다($\eta^2 < 0.01$; 효과 크기 거의 없음). 학력이 대재 이상인 경우에는 다른 학력군의 경우에

비해 읽기 능력 점수(T점수)가 2.29점 정도 높아질 것으로 예측된 반면($\eta^2=0.02$; 작은 효과 크기), 고졸 미만일 경우 다른 학력군의 경우에 비해 읽기 능력 점수(T점수)가 2.62점 낮아질 것으로 예측되었다($\eta^2<0.01$; 효과 크기 거의 없음).

3) 문법·규범 영역 결과

문법·규범 영역의 분석 대상은 5,000명이고 원점수의 만점은 150점으로, 조사 결과 원점수 평균은 98.46점, 표준편차는 32.93으로 나타났다. 문법·규범 영역의 능력 점수를 T점수로 변환하면 평균 50점(100점 만점), 표준편차는 8.59였고, 최솟값 25.13점, 최댓값 64.35점의 범위를 보이는 것으로 나타났다.

<표 IV-53> 문법·규범 영역의 능력 점수(T점수) 요약

영역	N	만점	평균	표준편차	최솟값	최댓값
문법·규범	5,000	100	50	8.59	25.13	64.35

문법·규범 영역의 성취 수준 등급별 인원 분포는 4수준이 1,482명(29.6%), 3수준이 1,863명(37.3%), 2수준이 1,175명(23.5%), 1수준이 480명(9.6%)으로 나타났다.

<표 IV-54> 문법·규범 영역의 성취 수준 등급별 인원 분포

수준	2024년	
	인원(명)	비율(%)
4수준	1,482	29.6
3수준	1,863	37.3
2수준	1,175	23.5
1수준	480	9.6
계	5,000	100.0

문법·규범 영역에서의 성별, 연령대, 지역 규모, 직업군, 학력과 같은 인구통계학적 변인에 따른 차이 분석 결과는 아래와 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수는 남성(50.06)이 여성(49.94)보다 근소하게 높았으나 통계적으로 유의한 차이는 아닌 것으로 확인되었다($p=.63$).

<표 IV-55> 문법·규범 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성		t	df	p
	평균	SD	평균	SD			
성별	50.06	8.54	49.94	8.65	.48	4,998	.63

성별에 따른 성취 수준 등급별 인원 분포를 비교한 결과, 남성과 여성의 1~4수준 비율은 대체로 비슷하였고, 실제로 비율 차이가 통계적으로 유의하지 않은 것으로 분석되었다(Wilcoxon 검정; $p=.56$). 이러한 결과 또한 문법·규범 영역의 국민의 국어능력에서 남성과 여성 간에 차이가 없을 가능성이 큼을 시사한다.

<표 IV-56> 문법·규범 영역 성취 수준 등급별 인원 분포의 성별 차이

변인	수준	남		여		순위합 검정 결과	
		명	%	명	%	W	p
성별	4수준	761	29.9	721	29.3	3,095,937	.56
	3수준	944	37.2	919	37.4		
	2수준	601	23.7	574	23.3		
	1수준	235	9.2	245	10.0		

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 20대(51.34), 30대(50.95), 40대(50.65), 50대(49.47), 60대(48.07) 순으로 평균 점수가 높게 나타나 연령대가 높을수록 문법·규범 영역 능력 점수는 낮아지는 경향성이 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 30대, 20대와 40대, 30대와 40대 간 차이를 제외한 모든 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-57> 문법·규범 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		F	$df1$	$df2$	p
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	51.34	8.66	50.95	8.88	50.65	8.56	49.47	8.24	48.07	8.35	23.94	4	4,995	<.001

연령대에 따른 성취 수준 등급별 인원 분포를 비교한 결과, 20대에서 60대로 연령대가 높아질수록 특히 4수준 비율이 줄어드는 패턴($36.9\% > 35.3\% > 32.5\% > 24.9\% > 21.4\%$)이 뚜렷하게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대·30대·40대 간에는 유의한 차이가 나타나지 않았으나, 이들 집단과 50대·60대 간, 그리고 50대와 60대 간에는 모두 유의한 차이가 확인되었다.

<표 IV-58> 문법·규범 영역 성취 수준 등급별 인원 분포의 연령대별 차이

변인	수준	20대		30대		40대		50대		60대		순위합 검정 결과		
		명	%	명	%	명	%	명	%	명	%	χ^2	df	p
연령대	4수준	304	36.9	315	35.3	345	32.5	295	24.9	223	21.4	89.03	4	<.001
	3수준	297	36.0	317	35.5	383	36.1	489	41.3	377	36.2			
	2수준	148	18.0	173	19.4	246	23.2	291	24.6	317	30.5			
	1수준	75	9.1	87	9.8	86	8.1	108	9.1	124	11.9			

거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 중소 도시(50.55), 대도시(49.78), 읍면 지역(49.21)의 순으로 평균 점수가 높게 나타났다. $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 중소 도시, 중소 도시와 읍면 지역 간 차이가 유의한 것으로 분석되었으며, 대도시와 읍면 지역의 차이는 통계적으로 유의하지 않은 것으로 확인되었다.

<표 IV-59> 문법·규범 영역 능력 점수의 지역 규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
지역 규모	49.78	8.50	50.55	8.94	49.21	7.86	8.27	2	4,997	<.001

거주 지역의 규모에 따른 성취 수준 등급별 인원 분포를 비교한 결과, 대도시나 중소 도시에 비해 읍면 지역에서 4수준 비율이 약 4~5%p 이상 낮게 나타나고, 2수준의 비율이 4~5%p가량 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p = .018$). 다만, $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 거주 지역의 규모에 따른 대응별 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과는 지역 규모에 따라 문법·규범 영역의 전반적인 성취 수준 등급 간의 차이는 존재하지만 그 차이가 충분히 크지 않은 것으로 볼 수 있다.

<표 IV-60> 문법·규범 영역 성취 수준 등급별 인원 분포의 지역 규모별 차이

변인	수준	대도시		중소 도시		읍면 지역		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
지역 규모	4수준	646	29.6	636	31.5	200	25.0	8.05	2	.018
	3수준	801	36.7	747	37.0	315	39.4			
	2수준	512	23.5	444	22.0	219	27.4			
	1수준	221	10.1	193	9.6	66	8.3			

직업군에 따른 능력 점수(T점수) 차이 분석 결과, 정신노동(51.49), 기타(49.57), 육체노동(49.21)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타 간 차이가 유의한 것으로 분석되었다.

<표 IV-61> 문법·규범 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
직업군	51.49	8.53	49.21	8.35	49.57	8.81	33.41	2	4,997	<.001

직업군에 따른 성취 수준 등급별 인원 분포를 비교한 결과, 특히 정신노동에 비해 육체노동이 4수준 비율이 낮고, 2수준 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분

석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타 직업군의 차이가 유의하였으며, 육체노동과 기타 직업군의 차이는 유의하지 않은 것으로 확인되었다. 이러한 결과 또한 직업군에 따라 문법·규범 영역의 국민의 국어능력에 차이가 있으며 특히 정신노동 종사자가 문법·규범 영역의 국어능력이 높을 가능성을 시사한다.

<표 IV-62> 문법·규범 영역 성취 수준 등급별 인원 분포의 직업군별 차이

변인	수준	정신노동		육체노동		기타		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
직업군	4수준	555	37.0	524	25.6	403	27.8	61.79	2	<.001
	3수준	546	36.4	789	38.5	528	36.4			
	2수준	277	18.5	531	25.9	367	25.3			
	1수준	123	8.2	206	10.0	151	10.4			

학력에 따른 능력 점수(T점수) 차이 분석 결과, 대재 이상(51.47), 고졸(48.56), 고졸 미만(46.34)의 순으로 평균 점수가 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-63> 문법·규범 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	46.34	8.43	48.56	8.23	51.47	8.51	111.98	2	4,997	<.001

학력에 따른 성취 수준 등급별 인원 분포를 비교한 결과, 학력이 높을수록 4수준 비율이 높게 나타났는데(고졸 미만 16.4% < 고졸 22.1% < 대재 이상 36.4%), 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 학력이 높을수록 문법·규범 영역의 국민의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-64> 문법·규범 영역 성취 수준 등급별 인원 분포의 학력별 차이

변인	수준	고졸 미만		고졸		대재 이상		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
학력	4수준	75	16.4	379	22.1	1,028	36.4	210.91	2	<.001
	3수준	129	28.2	681	39.6	1,053	37.3			
	2수준	185	40.4	470	27.4	520	18.4			
	1수준	69	15.1	188	10.9	223	7.9			

한편, 문법·규범 능력 수준을 예측하는 인구통계학적 변인을 확인하기 위하여 중다회귀분석 및 단계별

회귀분석을 수행하였다. 중다회귀분석을 수행하기 위한 예측 변수로는 성별, 연령, 지역 규모, 직업군, 학력 등 주요 인구통계학적 변인을 투입하였다. 이때, 연령 변인은 연속형 변수로 처리하였으며, 나머지 변인들은 범주형 변수로 처리하여 더미 변수로 변환하였다. 반응 변수로는 문항 반응 이론 분석에 따라 추정된 문법·규범 영역의 ‘능력 점수(T점수)’를 사용하였다.

회귀 모형에 투입할 예측 변수를 선정하고자 R의 `olsrr` 패키지에서 제공하는 `ols_step_both_p()` 함수를 이용하여 단계별 변수 선택을 수행하였다. 연령, 지역 규모, 학력은 문법·규범 영역 능력 점수(T점수)를 통계적으로 유의미하게 예측하는 변인으로 나타난 반면, 성별과 직업군은 통계적으로 유의미한 설명력을 보이지 못했다. 이에 따라 통계적으로 유의미한 인구통계학적 변인만을 포함하고 설명력이 높은 회귀 모형을 최종 모형으로 선택하였다. 최종 선택된 회귀 모형의 중다회귀분석 결과는 다음과 같다.

<표 IV-65> 인구통계학적 변인에 따른 문법·규범 능력 점수(T점수) 예측 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	50.29	NA	0.59	85.62	<.001 ***
연령	-0.023	-0.036	0.01	-2.22	0.026 *
대도시	-0.880	-0.051	0.26	-3.39	<.001 ***
읍면 지역	-1.039	-0.044	0.35	-2.95	0.003 **
대재 이상	2.663	0.154	0.28	9.39	<.001 ***
고졸 미만	-1.988	-0.067	0.45	-4.42	<.001 ***
잔여 표준오차	8.394 (df=4,994)				
R^2	0.0468	조정된 R^2		0.0459	
F 통계량	49.04 (5, df=4,994)	p		<.001	

*** p <.001 ** p <.01 * p <.05

최종 회귀 모형의 설명력(R^2)은 4.68%로 비교적 낮게 나타났으며 통계적으로 유의미한 수준이었다(p <.001). 연령이 높을수록 문법·규범 능력 점수(T점수)는 낮아지는 경향성을 확인할 수 있었다($\eta^2=0.02$; 작은 효과 크기). 지역 규모가 대도시인 경우 그렇지 않은 경우에 비해 문법·규범 능력 점수(T점수)가 0.88점 정도 낮을 것으로 예측되고($\eta^2=0.0008$; 효과 크기 거의 없음), 읍면 지역인 경우에도 그렇지 않은 경우에 비해 문법·규범 능력 점수(T점수)가 1.039점 정도 낮을 것으로 보인다($\eta^2=0.0019$; 효과 크기 거의 없음). 한편, 학력이 대재 이상인 경우에는 그렇지 않은 경우에 비해 문법·규범 능력 점수(T점수)가 2.663점 정도 높을 것으로 예측된다($\eta^2=0.02$; 작은 효과 크기). 반면 고졸 미만인 경우에는 그렇지 않은 경우에 비해 문법·규범 능력 점수(T점수)가 1.988점 정도 낮을 것으로 보인다($\eta^2=0.0039$; 효과 크기 거의 없음).

4) 쓰기 영역 결과

쓰기 영역의 조사 결과, 다분 문항 반응 모형의 일종인 등급 반응 모형(Graded Response Model)을 통해 추정된 능력 점수(T점수)와 수정된 앵고프 방법을 통해 판정한 수준 등급별 인원 분포를 전체적으로 요약하면 다음과 같다.

쓰기 영역의 분석 대상 3,000명의 능력 점수(T점수) 평균은 50.00점(100점 만점), 표준편차는 8.60이

었고, 최솟값 21.96점, 최댓값 82.60점의 범위를 보이는 것으로 나타났다.

<표 IV-66> 쓰기 영역의 능력 점수(T점수) 요약

영역	N	만점	평균	표준편차	최솟값	최댓값
쓰기	3,000	100	50.00	8.60	21.96	82.60

쓰기 영역의 수준 등급별 인원 분포는 4수준이 335명(11.2%), 3수준이 791명(26.4%), 2수준이 1,216명(40.5%), 1수준이 658명(21.9%)으로 나타났다.

<표 IV-67> 쓰기 영역의 수준 등급별 인원 분포

수준	2025년	
	인원(명)	비율(%)
4수준	335	11.2
3수준	791	26.4
2수준	1,216	40.5
1수준	658	21.9
계	3,000	100.0

쓰기 영역에서 성별, 연령대, 지역 규모, 직업군, 학력과 같은 배경 변인에 따른 차이 분석 결과는 다음과 같다. 먼저, 성별에 따른 능력 점수(T점수) 차이 분석 결과, 쓰기 영역의 평균 점수는 남성(50.03)이 여성(49.97)보다 근소하게 높았으나 통계적으로 유의한 차이는 아닌 것으로 확인되었다($p=.84$).

<표 IV-68> 쓰기 영역 능력 점수의 성별 차이

변인	능력 점수(T점수)				t-검정 결과		
	남성		여성				
	평균	SD	평균	SD	t	df	p
성별	50.03	8.71	49.97	8.49	.20	2,998	.84

성별에 따른 수준 등급별 인원 분포를 비교한 결과, 남성과 여성의 1~4수준 비율은 대체로 비슷하였고, 실제로 비율 차이가 통계적으로 유의하지 않은 것으로 분석되었다(Wilcoxon 검정: $p=.90$). 이러한 결과 또한 쓰기 영역 관련 국민의 국어능력에서 남성과 여성 간에 쓰기 영역 점수의 차이가 없을 가능성이 큼을 시사한다.

<표 IV-69> 쓰기 영역 수준 등급별 인원 분포의 성별 차이

변인	수준	남		여		순위합 검정 결과	
		명	%	명	%	<i>W</i>	<i>p</i>
성별	4수준	180	11.8	155	10.5	1,127,433	.90
	3수준	388	25.5	403	27.3		
	2수준	616	40.4	600	40.6		
	1수준	339	22.3	319	21.6		

다음으로, 연령대에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 30대(51.20), 20대(50.97), 40대(50.70), 50대(49.73), 60대(47.84) 순으로 높게 나타나 연령대가 높을수록 쓰기 영역의 능력 점수는 낮아지는 경향성이 보였으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 20대와 60대, 30대와 50대, 30대와 60대, 40대와 60대, 50대와 60대 간 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-70> 쓰기 영역 능력 점수의 연령대별 차이

변인	능력 점수(T점수)										F-검정 결과			
	20대		30대		40대		50대		60대		<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
연령대	50.97	8.42	51.20	9.09	50.70	8.40	49.73	8.65	47.84	8.03	15.69	4	2,995	<.001

연령대에 따른 수준 등급별 인원 분포를 비교한 결과, 20대에서 60대로 연령대가 높아질수록 특히 1수준 비율이 늘어나는 패턴($18.0\% < 18.3\% < 18.9\% < 23.5\% < 29.3\%$)이 다소간 엮보였는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p < .001$). 그러나 $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 60대와 모든 다른 연령대 간의 대응별 차이만이 유의한 것으로 분석되었다. 이러한 결과는 쓰기 능력이 젊은 연령대보다 노령 연령대에서 특히 더 낮을 가능성이 큼을 시사한다.

<표 IV-71> 쓰기 영역 수준 등급별 인원 분포의 연령대별 차이

변인	수준	20대		30대		40대		50대		60대		순위합 검정 결과		
		명	%	명	%	명	%	명	%	명	%	χ^2	<i>df</i>	<i>p</i>
연령대	4수준	65	13.3	81	14.8	69	11.1	77	10.9	43	6.8	58.63	4	<.001
	3수준	138	28.2	150	27.4	193	31.0	185	26.2	125	19.7			
	2수준	199	40.6	216	39.5	243	39.0	277	39.3	281	44.3			
	1수준	88	18.0	100	18.3	118	18.9	166	23.5	186	29.3			

다음으로, 거주 지역의 규모에 따른 능력 점수(T점수) 차이 분석 결과, 쓰기 영역의 평균 점수가 대도시(50.57), 중소 도시(49.91), 읍면 지역(48.64)의 순으로 높게 나타났고, 실제로 이러한 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 읍면 지역, 중소 도시와 읍면 지역 간 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-72> 쓰기 영역 능력 점수의 지역 규모별 차이

변인	능력 점수(T점수)						F-검정 결과			
	대도시		중소 도시		읍면 지역		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
지역 규모	50.57	7.67	49.91	9.41	48.64	8.69	8.80	2	2,997	<.001

거주 지역의 규모에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 대도시가 중소 도시나 읍면 지역에 비해 1수준 비율이 7%p가량 낮고 3수준 비율이 7%p가량 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 대도시와 중소 도시, 대도시와 읍면 지역의 차이가 유의한 반면, 중소 도시와 읍면 지역의 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과는 대도시 거주자들이 다른 지역 거주자들에 비해 쓰기 영역의 국어능력이 높을 가능성이 큼을 시사한다.

<표 IV-73> 쓰기 영역 수준 등급별 인원 분포의 지역 규모별 차이

변인	수준	대도시		중소 도시		읍면 지역		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
지역 규모	4수준	128	9.8	160	13.0	47	10.0	15.73	2	<.001
	3수준	395	30.4	291	23.7	105	22.3			
	2수준	549	42.2	475	38.6	192	40.9			
	1수준	228	17.5	304	24.7	126	26.8			

직업군에 따른 능력 점수(T점수) 차이 분석 결과, 쓰기 영역의 평균 점수가 정신노동(51.36), 기타(49.83), 육체노동(49.14)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타 간 차이가 유의한 것으로 분석되었다.

<표 IV-74> 쓰기 영역 능력 점수의 직업군별 차이

변인	능력 점수(T점수)						F-검정 결과			
	정신노동		육체노동		기타		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
직업군	51.36	8.97	49.14	8.24	49.83	8.55	16.95	2	2,997	<.001

직업군에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 정신노동에 비해 육체노동과 기타가 4수준과 3수준 비율이 낮고, 2수준과 1수준 비율이 높게 나타났는데, 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정: $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 정신노동과 육체노동, 정신노동과 기타의 차이가 유의한 반면, 육체노동과 기타의 차이는 유의하지 않은 것으로 분석되었다. 이러한 결과 또한 직업군에 따라 쓰기 영역의 국어능력에 차이가 있으며 특히 정신노동 종사자가 육체노동 종사자보다 능력이 높을 가능성이 큼을 시사한다.

<표 IV-75> 쓰기 영역 수준 등급별 인원 분포의 직업군별 차이

변인	수준	정신노동		육체노동		기타		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
직업군	4수준	137	15.0	106	8.7	92	10.6	30.74	2	<.001
	3수준	262	28.7	308	25.2	221	25.5			
	2수준	351	38.4	504	41.3	361	41.7			
	1수준	164	17.9	303	24.8	191	22.1			

끝으로, 학력에 따른 능력 점수(T점수) 차이 분석 결과, 평균 점수가 대재 이상(51.26), 고졸(48.84), 고졸 미만(46.51)의 순으로 높게 나타났으며, 실제로 이러한 차이가 통계적으로 유의한 것으로 분석되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다.

<표 IV-76> 쓰기 영역 능력 점수의 학력별 차이

변인	능력 점수(T점수)						F-검정 결과			
	고졸 미만		고졸		대재 이상		F	df1	df2	p
	평균	SD	평균	SD	평균	SD				
학력	46.51	7.38	48.84	8.22	51.26	8.76	52.18	2	2,997	<.001

학력에 따른 수준 등급별 인원 분포를 비교한 결과, 특히 학력이 높을수록 4수준 비율이 높게 나타났는데(대재 이상 13.9% > 고졸 8.6% > 고졸 미만 3.9%), 실제로 이러한 비율 차이가 통계적으로 유의한 것으로 분석되었다(Kruskal-Wallis 검정; $p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Bonferroni) 결과, 모든 대응별 차이가 유의한 것으로 분석되었다. 이러한 결과 또한 쓰기 영역의 국어능력은 학력이 높을수록 높을 가능성이 큼을 시사한다.

<표 IV-77> 쓰기 영역 수준 등급별 인원 분포의 학력별 차이

변인	수준	고졸 미만		고졸		대재 이상		순위합 검정 결과		
		명	%	명	%	명	%	χ^2	df	p
학력	4수준	11	3.9	87	8.6	237	13.9	88.51	2	<.001
	3수준	44	15.8	245	24.1	502	29.4			
	2수준	129	46.2	430	42.3	657	38.5			
	1수준	95	34.1	254	25.0	309	18.1			

한편, 쓰기 능력 수준을 설명하는 인구통계학적 변인을 확인하기 위하여 중다회귀분석 및 단계별 회귀분석을 수행하였다. 중다회귀분석을 수행하기 위한 설명 변수로는 성별, 연령, 지역 규모, 직업군, 학력 등 주요 인구통계학적 변인을 투입하였다. 이때, 연령 변인은 연속형 변수로 처리하였으며, 나머지 변인들은 범주형 변수로 처리하여 더미 변수로 변환하였다. 반응 변수로는 문항 반응 이론 분석에 따라 추정된 쓰기 영역의 ‘능력 점수(T점수)’를 사용하였다.

회귀 모형에 투입할 예측 변수를 선정하고자 R의 step() 함수를 이용하여 단계별 변수 선택을 수행하였다. 쓰기 영역에서 성별, 직업군의 예측력은 유의하지 않아 제외되었고, 학력, 지역, 연령은 유의미한 예측 변인으로 검증되었다. 이에 따라 통계적으로 유의미한 인구통계학적 변인만을 포함하고 설명력이 높은 회귀 모형을 최종 모형으로 선택하였다. 최종 선택된 회귀 모형의 중다회귀분석 결과는 <표 IV-78>과 같다.

<표 IV-78> 인구통계학적 변인에 따른 쓰기 능력 점수(T점수) 설명 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	50.14	NA	0.77	64.94	<.001***
대재 이상	1.99	0.23	0.37	5.33	<.001***
고졸 미만	-2.00	-0.23	0.59	-3.39	<.001***
대도시	0.83	.010	0.31	2.66	.01**
연령	-0.03	-0.00	0.01	-2.29	0.02**
잔여 표준오차	8.44 (df=2,995)				
R^2	0.038	조정된 R^2		0.036	
F 통계량	29.28 (df=4, 2,995)	p		<0.001	

*** $p < .001$ ** $p < .01$ * $p < .05$

최종 회귀 모형의 설명력(R^2)은 3.8%로 낮게 나타났지만 통계적으로 유의미한 수준이었다($p < .001$). 학력은 대재 이상인 경우 그렇지 않은 경우에 비해 쓰기 능력 수준이 높게 나타났으며($\eta^2 = 0.03$; 작은 효과 크기), 쓰기 능력 점수(T점수)가 약 2점 정도 높아진다. 학력이 고졸 미만인 경우 다른 학력군의 경우에 비해 쓰기 능력 수준이 낮게 나타났으며($\eta^2 = 0.006$; 효과 크기 거의 없음), 쓰기 능력 점수(T점수)가 약 2점 정도 낮아진다. 또한 지역은 대도시인 경우가 다른 지역의 경우에 비해 쓰기 능력 수준이 높게 나타났으며($\eta^2 = 0.002$; 효과 크기 거의 없음), 쓰기 능력 점수(T점수)는 0.8점 정도 높다. 연령대의 경우 연령대가 높아질수록 쓰기 능력 수준이 낮게 나타났으며($\eta^2 = 0.002$; 효과 크기 거의 없음). 따라서 인구통계학적 예측 변인 가운데에서는 학력 가운데서도 특히 ‘고등교육(대학)’의 이수가 쓰기 능력의 차이를 불러온다는 점에 주목할 필요가 있을 것으로 보인다.

4.2. 영역별 심층 변인에 대한 분석

1) 듣기, 듣기·말하기, 말하기 영역 심층 변인 분석

듣기, 듣기·말하기, 말하기 영역의 심층 변인에 따른 능력 점수(T점수)의 평균 차이를 분석하기 위하여 주요 구어 의사소통 참여 빈도별 일원배치 분산분석(one-way ANOVA)을 수행하였으며, 평균 점수 차이가 통계적으로 유의한 경우에는 셰페(Scheffé) 보정 방법을 적용하여 사후 검정을 실시하였다.

(1) (가족, 직장 동료 등 어떠한 사람과 함께든지) 직접 만나서 하는 대화

① 듣기 능력의 차이

먼저, (가족, 직장 동료 등 어떠한 사람과 함께든지) 직접 만나서 하는 대화 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(50.92)에 응답한 집단의 능력 평균 점수가 다른 집단에 비해 높게 나타났다. 그 반면 ‘거의 하지 않는다’(42.94)에 응답한 집단의 능력 평균 점수가 가장 낮게 나타났다. 결과적으로 대면 대화 빈도가 감소할수록 듣기 능력 점수의 평균값도 대체로 낮아지는 경향이 나타났다.

<표 IV-79> 듣기 영역 능력 점수의 직접 만나서 하는 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	c		bc		b		b		a					
	거의 하지 않음 (n=43)		별로 하지 않음 (n=146)		보통 (n=373)		자주 함 (n=801)		매우 자주 함 (n=1,670)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
대면 대화 빈도	42.94	6.15	47.07	8.75	48.16	8.87	49.24	8.28	50.92	7.50	25.65	4	3,028	<.001
사후 검정 해석	매우 자주 함(a) > 자주 함(b) = 보통(b) ≥ 별로 하지 않음(bc) ≥ 거의 하지 않음(c)													

② 듣기·말하기 능력의 차이

먼저, (가족, 직장 동료 등 어떠한 사람과 함께든지) 직접 만나서 하는 대화 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(50.47)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 더 높게 나타났다.

<표 IV-80> 듣기·말하기 영역 능력 점수의 직접 만나서 하는 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		b		b		b		a					
	거의 하지 않음 (n=43)		별로 하지 않음 (n=146)		보통 (n=372)		자주 함 (n=798)		매우 자주 함 (n=1,668)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	대면 대화 빈도	45.46	9.48	47.30	9.09	49.02	9.43	49.25	8.64	50.47	8.22	10.03	4	3,022
사후 검정 해석	매우 자주 함(a) > 자주 함(b) = 보통(b) = 별로 하지 않음(b) = 거의 하지 않음(b)													

③ 말하기 능력의 차이

먼저, (가족, 직장 동료 등 어떠한 사람과 함께든지) 직접 만나서 하는 대화 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다(하루에 약 3번 이상)’(50.86)에 응답한 집단의 능력 평균 점수가 가장 높았으며, ‘거의 하지 않는다(한 달에 약 1~2번 미만)’(44.19)에 응답한 집단의 능력 평균 점수가 가장 낮게 나타났다. 결과적으로, 대화 빈도가 감소할수록 능력 평균 점수도 낮아지는 경향이 확인되었다.

<표 IV-81> 말하기 영역 능력 점수의 직접 만나서 하는 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	c		bc		b		b		a					
	거의 하지 않음 (n=43)		별로 하지 않음 (n=146)		보통 (n=373)		자주 함 (n=800)		매우 자주 함 (n=1,668)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df 1	df 2	p
	대면 대화 빈도	44.19	8.18	47.73	9.11	48.95	8.84	49.59	9.24	50.86	9.13	12.01	4	3,025
사후 검정 해석	매우 자주 함(a) > 자주 함(b) = 보통(b) ≥ 별로 하지 않음(bc) ≥ 거의 하지 않음(c)													

(2) 누리소통망(SNS)이나 메신저로 하는 대화

① 듣기 능력의 차이

다음으로, 누리소통망(SNS)이나 메신저로 하는 대화 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준

에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(51.50)에 응답한 집단의 능력 평균 점수가 다른 집단에 비해 높게 나타났다.

<표 IV-82> 듣기 영역 능력 점수의 누리소통망(SNS)이나 메신저 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		b		b		b		a					
	거의 하지 않음 (n=186)		별로 하지 않음 (n=294)		보통 (n=504)		자주 함 (n=787)		매우 자주 함 (n=1,262)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	누리소통 망(SNS), 메신저 대화 빈도 사후 검정 해석	47.58	8.69	48.27	8.79	48.10	8.42	49.41	7.91	51.50	7.39	27.12	4	3,028
매우 자주 함(a) > 자주 함(b) = 별로 하지 않음(b) = 보통(b) = 거의 하지 않음(b)														

② 듣기·말하기 능력의 차이

다음으로, 누리소통망(SNS)이나 메신저로 하는 대화 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(50.73)에 응답한 집단의 능력 평균 점수는 ‘별로 하지 않는다’(49.56)에 응답한 집단과 비교하여 같거나 더 높은 것으로 나타났고, ‘자주 한다’(49.19), ‘보통이다’(48.85), ‘거의 하지 않는다’(48.13)에 응답한 집단과 비교하여 더 높은 것으로 확인되었다.

<표 IV-83> 듣기·말하기 영역 능력 점수의 누리소통망(SNS)이나 메신저 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		ab		b		b		a					
	거의 하지 않음 (n=186)		별로 하지 않음 (n=293)		보통 (n=503)		자주 함 (n=787)		매우 자주 함 (n=1,258)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
누리소통 망(SNS), 메신저 대화 빈도	48.13	9.48	49.56	8.72	48.85	9.01	49.19	8.79	50.73	8.03	8.09	4	3,022	<.001
사후 검정 해석	매우 자주 함(a) ≥ 별로 하지 않음(ab) ≥ 자주 함(b) = 보통(b) = 거의 하지 않음(b)													

③ 말하기 능력의 차이

다음으로, 누리소통망(SNS)이나 메신저로 하는 대화 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(51.5)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 가장 높게 나타났다.

<표 IV-84> 말하기 영역 능력 점수의 누리소통망(SNS)이나 메신저 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		b		b		b		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=503)		(n=787)		함					
	(n=186)		(n=294)						(n=1,260)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
누리소통 망(SNS), 메신저 대화 빈도	47.22	8.65	48.65	8.85	49.2	9.31	49.42	9.09	51.5	9.09	16.38	4	3,025	<.001
사후 검정 해석	매우 자주 함(a) > 자주 함(b) = 보통(b) = 별로 하지 않음(b) = 거의 하지 않음(b)													

(3) 음성 통화로 하는 대화

① 듣기 능력의 차이

다음으로, 음성 통화로 하는 대화 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘별로 하지 않는다’(51.86)에 응답한 집단의 능력 평균 점수가 가장 높게 나타났고, ‘거의 하지 않는다’(45.68)에 응답한 집단의 능력 평균 점수가 가장 낮게 나타났다. 음성 통화 대화 빈도에 따른 듣기 능력 점수의 평균값에는 뚜렷한 경향성이 드러나지 않았는데, 이는 통제되지 않은 외생변수의 개입을 시사한다.

<표 IV-85> 듣기 영역 능력 점수의 음성 통화 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	c		a		b		ab		b					
	거의 하지 않음 (n=67)		별로 하지 않음 (n=223)		보통 (n=602)		자주 함 (n=919)		매우 자주 함 (n=1,219)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	F	df1	df2	p										
음성 통화 대화 빈도	45.68	8.92	51.86	7.86	49.91	9.06	50.21	8.18	49.38	7.29	9.57	4	3,028	<.001
사후 검정 해석	별로 하지 않음(a) ≥ 자주 함(ab) ≥ 보통(b) = 매우 자주 함(b) > 거의 하지 않음(c)													

② 듣기·말하기 능력의 차이

다음으로, 음성 통화로 하는 대화 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(50.35), ‘보통이다’(49.66)에 응답한 집단의 능력 평균 점수는 ‘자주 한다’(49.37), ‘별로 하지 않는다’(49.36)에 응답한 집단과 비교하여 같거나 더 높은 것으로 나타났다. 한편, ‘거의 하지 않는다’(46.07)에 응답한 집단의 능력 평균 점수는 그 밖의 집단과 비교하여 더 낮은 것으로 확인되었다.

<표 IV-86> 듣기·말하기 영역 능력 점수의 음성 통화 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		ab		a		ab		a					
	거의 하지 않음 (n=67)		별로 하지 않음 (n=223)		보통 (n=600)		자주 함 (n=918)		매우 자주 함 (n=1,219)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	음성 통화 대화 빈도 사후 검정 해석	46.07	9.22	49.36	8.87	49.66	8.74	49.37	8.80	50.35	8.23	5.15	4	3,022
매우 자주 함(a) = 보통(a) ≥ 자주 함(ab) = 별로 하지 않음(ab) ≥ 거의 하지 않음(b)														

③ 말하기 능력의 차이

다음으로, 음성 통화로 하는 대화 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .01$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘별로 하지 않는다’(51.24), ‘자주 한다’(50.25)에 응답한 집단의 능력 평균 점수는 ‘보통이다’(50), ‘매우 자주 한다’(49.88)에 응답한 집단과 비교하여 같거나 더 높은 것으로 나타났다. 한편, ‘거의 하지 않는다’(46.43)에 응답한 집단의 능력 평균 점수는 가장 낮은 것으로 확인되었다. 다만, 집단 간 능력 평균 점수 차이 간에는 일정한 경향성이 두드러지게 나타나지 않았는데, 이는

통제되지 않은 외생변수의 개입 가능성을 시사한다.

<표 IV-87> 말하기 영역 능력 점수의 음성 통화 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		ab		a		ab					
	거의 하지		별로 하지		보통		자주 함		매우 자주 함					
	않음		않음		(n=602)		(n=919)		(n=1,219)					
	(n=67)		(n=223)											
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
음성 통화 대화 빈도 사후 검정 해석	46.43	8.16	51.24	8.32	50.00	9.25	50.25	9.60	49.88	8.96	3.79	4	3,025	<.01
별로 하지 않음(a) = 자주 함(a) ≥ 보통(ab) = 매우 자주 함(ab) ≥ 거의 하지 않음(b)														

(4) 영상 통화로 하는 대화

① 듣기 능력의 차이

다음으로, 영상 통화로 하는 대화 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘거의 하지 않는다’(50.58)에 응답한 집단의 평균 점수가 가장 높게 나타났고, ‘자주 한다’(47.25)에 응답한 집단의 평균 점수가 가장 낮았다. 대체적으로 영상 통화 대화 빈도가 높아질수록 듣기 능력 점수의 평균값이 감소하는 경향이 나타났다.

<표 IV-88> 듣기 영역 능력 점수의 영상 통화 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		ab		bc		c		bc					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=457)		(n=210)		함					
	(n=1,299)		(n=959)						(n=108)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
영상 통화 대화 빈도 사후 검정 해석	50.58	8.14	49.91	7.70	49.11	8.34	47.25	8.33	48.42	7.20	10.12	4	3,028	<.001
거의 하지 않음(a) ≥ 별로 하지 않음(ab) ≥ 보통(bc) = 매우 자주 함(bc) ≥ 자주 함(c)														

② 듣기·말하기 능력의 차이

다음으로, 영상 통화로 하는 대화 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과,

대화 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p=.13$).

<표 IV-89> 듣기·말하기 영역 능력 점수의 영상 통화 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주 함					
	않음		않음		(n=456)		(n=210)		(n=108)					
	(n=1,296)		(n=957)											
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
영상 통화 대화 빈도	49.73	8.59	49.47	8.64	50.66	8.42	49.20	8.59	49.55	8.95	1.76	4	3,022	.13

③ 말하기 능력의 차이

다음으로, 영상 통화로 하는 대화 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 대화 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p=.421$).

<표 IV-90> 말하기 영역 능력 점수의 영상 통화 대화 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=457)		(n=210)		함					
	(n=1,299)		(n=956)						(n=108)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
영상 통화 대화 빈도	50.17	9.15	50.11	9.10	49.93	9.33	50.03	9.64	48.39	8.44	0.97	4	3,025	.421

(5) 발표·강연 듣기

① 듣기 능력의 차이

다음으로, 발표·강연 듣기 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 듣기 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘별로 하지 않는다’(51.82)와 ‘자주 한다’(51.18)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 같거나 더 높은 것으로 나타났다. 그 반면 ‘거의 하지 않는다’(48.43)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 같거나 더 낮은 것으로 나타났다. 발표·강연 듣기 빈도에 따른 듣기 능력 점수의 평균값에는 뚜렷한 경향성이 드러나지 않았는데, 이는 통제되지 않은 외생변수의 개입을 시사한다.

<표 IV-91> 듣기 영역 능력 점수의 발표·강연 듣기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		ab		a		ab					
	거의 하지 않음 (n=1,593)		별로 하지 않음 (n=765)		보통 (n=425)		자주 함 (n=185)		매우 자주 함 (n=65)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
발표·강연 듣기 빈도	48.43	8.07	51.82	7.33	50.83	8.49	51.18	7.75	50.83	7.77	27.77	4	3,028	<.001
사후 검정 해석	별로 하지 않음(a) = 자주 함(a) ≥ 매우 자주 함(ab) = 보통(ab) ≥ 거의 하지 않음(b)													

② 듣기·말하기 능력의 차이

다음으로, 발표·강연 듣기 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 듣기 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p=.06$).

<표 IV-92> 듣기·말하기 영역 능력 점수의 발표·강연 듣기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지 않음 (n=1,591)		별로 하지 않음 (n=762)		보통 (n=424)		자주 함 (n=185)		매우 자주 함 (n=65)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	발표·강연 듣기 빈도	49.34	8.69	50.39	8.38	49.95	8.50	49.80	8.46	50.80	9.66	2.28	4	3,022

③ 말하기 능력의 차이

다음으로, 발표·강연 듣기 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 듣기 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘거의 하지 않는다’(48.37)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 더 낮은 것으로 나타났다.

<표 IV-93> 말하기 영역 능력 점수의 발표·강연 듣기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		a		a		a					
	거의 하지 않음 (n=1,592)		별로 하지 않음 (n=764)		보통 (n=424)		자주 함 (n=185)		매우 자주 함 (n=65)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	F	df1	df2	p										
발표·강연 듣기 빈도	48.37	9.00	52.02	8.89	51.90	8.94	51.29	9.75	52.08	8.82	29.27	4	3,025	<.001
사후 검정 해석	매우 자주 함(a) = 별로 하지 않음(a) = 보통(a) = 자주 함(a) > 거의 하지 않음(b)													

(6) 발표·강연하기

① 듣기 능력의 차이

다음으로, 발표·강연하기 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘별로 하지 않음’(51.29)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 더 높은 것으로 나타났다. 그 반면 ‘거의 하지 않음’(49.63)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 같거나 더 낮은 것으로 나타났다. 발표·강연하기 빈도에 따른 듣기 능력 점수의 평균값에는 뚜렷한 경향성이 드러나지 않았는데, 이는 통제되지 않은 외생변수의 개입을 시사한다.

<표 IV-94> 듣기 영역 능력 점수의 발표·강연하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		b		b		ab					
	거의 하지 않음 (n=2,235)		별로 하지 않음 (n=475)		보통 (n=209)		자주 함 (n=89)		매우 자주 함 (n=25)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df 1	df 2	p
발표·강연 하기 빈도	49.63	7.98	51.29	8.04	49.36	8.45	48.42	8.70	49.85	7.66	5.11	4	3,028	<.001
사후 검정 해석	별로 하지 않음(a) ≥ 매우 자주 함(ab) ≥ 거의 하지 않음(b) = 자주 함(b) = 보통(b)													

② 듣기·말하기 능력의 차이

다음으로, 발표·강연하기 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p = .12$).

<표 IV-95> 듣기·말하기 영역 능력 점수의 발표·강연하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지 않음 (n=2,231)		별로 하지 않음 (n=474)		보통 (n=208)		자주 함 (n=89)		매우 자주 함 (n=25)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	발표·강연 하기 빈도	49.70	8.56	50.34	8.33	48.89	9.23	49.01	9.04	52.39	9.20	1.85	4	3,022

③ 말하기 능력의 차이

다음으로, 발표·강연하기 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘별로 하지 않는다’(52.12)에 응답한 집단의 능력 평균 점수가 가장 높았으며, ‘거의 하지 않는다’(49.51)에 응답한 집단의 능력 평균 점수가 가장 낮은 것으로 나타났다. 다만, 집단 간 능력 평균 점수 차이 간에는 일정한 경향성이 두드러지게 나타나지 않았는데, 이는 통제되지 않은 외생변수의 개입 가능성을 시사한다.

<표 IV-96> 말하기 영역 능력 점수의 발표·강연하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		ab		ab		ab					
	거의 하지 않음 (n=2,233)		별로 하지 않음 (n=475)		보통 (n=208)		자주 함 (n=89)		매우 자주 함 (n=25)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	F	df 1	df 2	p										
발표·강연 하기 빈도	49.51	9.05	52.12	9.33	50.55	9.48	50.82	9.05	50.93	9.70	8.43	4	3,025	<.001
사후 검정 해석	별로 하지 않음(a) ≥ 매우 자주 함(ab) = 자주 함(ab) = 보통(ab) ≥ 거의 하지 않음(b)													

(7) 토의(회의)에서 듣고 말하기

① 듣기 능력의 차이

다음으로, 토의(회의)에서 듣고 말하기 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘거의 하지 않는다’(48.35)에 응답한 집단의 능력 평균 점수가 그 밖의 집단에 비해 더 낮게 나타났다.

<표 IV-97> 듣기 영역 능력 점수의 토의(회의)에서 듣고 말하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		a		a		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=444)		(n=213)		함					
	(n=1,673)		(n=653)						(n=50)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
토의(회의)														
듣고 말하기 빈도	48.35	7.98	51.75	7.34	51.41	8.48	51.61	7.76	53.33	6.86	34.09	4	3,028	<.001
사후 검정 해석	매우 자주 함(a) = 별로 하지 않음(a) = 자주 함(a) = 보통(a) > 거의 하지 않음(b)													

② 듣기·말하기 능력의 차이

다음으로, 토의(회의)에서 듣고 말하기 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 빈도별 평균 점수 차이가 있는 것으로 나타났으나($p < .01$), $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과 집단 간 차이는 확인되지 않았다.

<표 IV-98> 듣기·말하기 영역 능력 점수의 토의(회의)에서 듣고 말하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=443)		(n=213)		함					
	(n=1,669)		(n=652)						(n=50)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
토의(회의)														
듣고 말하기 빈도	49.21	8.68	50.31	8.55	50.31	8.46	50.59	8.25	51.74	8.33	4	4	3,022	<.01
사후 검정 해석	매우 자주 함(a) = 자주 함(a) = 별로 하지 않음(a) = 보통(a) = 거의 하지 않음(a)													

③ 말하기 능력의 차이

다음으로, 토의(회의)에서 듣고 말하기 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘거의 하지 않는다’(48.54)에 응답한 집단의 능력 평균 점수가 그 밖의 집단과 비교하여 더 낮게 나타났다. 다만, 집단 간 능력 평균 점수 차이 간에는 일정한 경향성이 두드러지게 나타나지 않았는데, 이는 통제되지 않은 외생변수의 개입 가능성을 시사한다.

<표 IV-99> 말하기 영역 능력 점수의 토의(회의)에서 듣고 말하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		ab		a		a		ab					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=443)		(n=213)		함					
	(n=1,672)		(n=652)						(n=50)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
토의(회의)														
듣고 말하기 빈도	48.54	9.02	51.43	9.01	52.2	8.88	52.63	9.40	52.03	8.59	26.79	4	3,025	<.001
사후 검정 해석	자주 함(a) = 보통(a) ≥ 매우 자주 함(ab) = 별로 하지 않음(ab) ≥ 거의 하지 않음(b)													

(8) (방송 토론 등의) 토론 듣기

① 듣기 능력의 차이

다음으로, (방송 토론 등의) 토론 듣기 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 듣기 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이가 검정(Scheffé)을 실시한 결과, ‘매우 자주 한다’(51.93), ‘별로 하지 않는다’(51.49), ‘보통이다’(50.45)를 선택한 집단의 듣기 능력 점수가 그 밖의 집단에 비해 같거나 더 높게 나타났다.

<표 IV-100> 듣기 영역 능력 점수의 토론 듣기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	ab		a		a		ab		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=418)		(n=186)		함					
	(n=1,636)		(n=737)						(n=56)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
토론 듣기 빈도	48.82	8.14	51.49	7.38	50.45	8.39	50.29	8.24	51.93	6.85	16.28	4	3,028	<.001
사후 검정 해석	매우 자주 함(a) = 별로 하지 않음(a) = 보통(a) ≥ 자주 함(ab) = 거의 하지 않음(ab)													

② 듣기·말하기 능력의 차이

다음으로, (방송 토론 등의) 토론 듣기 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 듣기 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p = .08$).

<표 IV-101> 듣기·말하기 영역 능력 점수의 토론 듣기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지		별로 하지		보통		자주 함		매우 자주					
	않음		않음		(n=415)		(n=186)		함					
	(n=1,634)		(n=736)						(n=56)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df 1	df 2	p	
토론 듣기 빈도	49.43	8.57	50.49	8.36	49.70	8.90	49.46	8.79	50.50	9.36	2.09	4	3,022	.08

③ 말하기 능력의 차이

다음으로, (방송 토론 등의) 토론 듣기 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 듣기 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘별로 하지 않는다’(51.85), ‘보통이다’(51.05)에 응답한 집단의 능력 평균 점수는 ‘매우 자주 한다’(50.69), ‘자주 한다’(50.52)에 응답한 집단과 비교하여 같거나 더 높게 나타났다. 한편, ‘거의 하지 않는다’(48.9)에 응답한 집단의 능력 평균 점수가 가장 낮은 것으로 확인되었다. 다만, 집단 간 능력 평균 점수 차이 간에는 일정한 경향성이 두드러지게 나타나지 않았는데, 이는 통제되지 않은 외생변수의 개입 가능성을 시사한다.

<표 IV-102> 말하기 영역 능력 점수의 토론 듣기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		a		ab		ab					
	거의 하지 않음 (n=1,635)		별로 하지 않음 (n=736)		보통 (n=418)		자주 함 (n=185)		매우 자주 함 (n=56)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df 1	df 2	p
토론 듣기 빈도	48.90	9.10	51.85	8.92	51.05	9.36	50.52	8.96	50.69	9.27	15.23	4	3,025	<.001
사후 검정 해석	별로 하지 않음(a) = 보통(a) ≥ 매우 자주 함(ab) = 자주 함(ab) ≥ 거의 하지 않음(b)													

(9) 토론에서 말하기

① 듣기 능력의 차이

끝으로, 토론에서 말하기 빈도에 따른 듣기 영역 능력 점수(T점수) 차이 분석 결과, 말하기 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p = .1$).

<표 IV-103> 듣기 영역 능력 점수의 토론에서 말하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	거의 하지 않음 (n=2,194)		별로 하지 않음 (n=493)		보통 (n=247)		자주 함 (n=72)		매우 자주 함 (n=24)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	토론 말하기 빈도	49.76	8.04	50.54	7.68	48.93	8.77	50.05	8.33	51.13	8.67	1.95	4	3,028

② 듣기·말하기 능력의 차이

끝으로, 토론에서 말하기 빈도에 따른 듣기·말하기 영역 능력 점수(T점수) 차이 분석 결과, 말하기 빈도별 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p=.9$).

<표 IV-104> 듣기·말하기 영역 능력 점수의 토론에서 말하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과				
	a		a		a		a		a						
	거의 하지		별로 하지		보통		자주 함		매우 자주						
	않음		않음		(n=247)		(n=72)		함						
	(n=2,191)		(n=493)						(n=24)						
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
토론															
말하기	49.66	8.56	49.95	8.53	49.90	8.95	50.43	8.37	49.96	11.22	0.27	4	3,022	.9	
빈도															

③ 말하기 능력의 차이

끝으로, 토론에서 말하기 빈도에 따른 말하기 영역 능력 점수(T점수) 차이 분석 결과, 말하기 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p<.001$). $\alpha<.05$ 수준에서의 대응별 차이 검정 (Scheffé)을 실시한 결과, ‘별로 하지 않는다’(52.06)에 응답한 집단의 능력 평균 점수가 가장 높았고, ‘거의 하지 않는다’(49.57)에 응답한 집단의 능력 평균 점수가 가장 낮게 나타났다. 다만, 집단 간 능력 평균 점수 차이 간에는 일정한 경향성이 두드러지게 나타나지 않았는데, 이는 통제되지 않은 외생변수의 개입 가능성을 시사한다.

<표 IV-105> 말하기 영역 능력 점수의 토론에서 말하기 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		ab		ab		ab					
	거의 하지 않음 (n=2,194)		별로 하지 않음 (n=493)		보통 (n=247)		자주 함 (n=72)		매우 자주 함 (n=24)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
토론 말하기 빈도	49.57	9.07	52.06	9.20	50.05	9.48	50.25	9.37	50.56	9.14	7.48	4	3,025	<.001
사후 검정 해석	별로 하지 않음(a) ≥ 매우 자주 함(ab) = 자주 함(ab) = 보통(ab) ≥ 거의 하지 않음(b)													

2) 읽기 영역 심층 변인 분석

읽기 영역의 심층 변인에 따른 능력 점수(T점수)의 평균 차이를 분석하기 위하여 일원배치 분산분석(one-way ANOVA)을 수행하였으며, 평균 점수 차이가 통계적으로 유의한 경우에는 셰페(Scheffé) 보정 방법을 적용하여 사후 검정을 실시하였다.

(1) 다양한 유형의 매체 사용 빈도

먼저, 다양한 유형의 매체 중에서 종이 신문, 잡지 등 중심 매체의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 사용한다(하루에 5번 이상)’(50.65)와 ‘전혀 사용하지 않는다(한 달에 한 번도 사용하지 않음)’(50.20)에 응답한 집단의 능력 점수 평균이 ‘거의 사용하지 않는다(한 달에 약 1~2번)’(49.53), ‘어느 정도 사용한다(하루에 1~4번)’(48.21), ‘보통이다(일주일에 약 1~2번)’(47.37)에 응답한 집단과 비교하여 평균 점수가 같거나 높게 나타났다. 다만 ‘매우 자주 사용한다(하루에 5번 이상)’에 응답한 집단의 사례 수가 다른 집단에 비해 매우 적으므로($n=36$), 해석에 유의가 필요하다 하겠다.

<표 IV-106> 매체 사용 빈도(종이 신문, 잡지)별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		ab		ac		abc		a					
	전혀 사용 않음 (n=2,567)		거의 사용 않음 (n=1,466)		보통 (n=669)		어느 정도 사용함 (n=262)		매우 자주 사용함 (n=36)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
신문, 잡지 사용 빈도 사후 검정 해석	50.20	8.51	49.53	9.25	47.37	10.06	48.21	10.14	50.65	9.29	14.58	4	4.995	<.001
매우 자주 사용함(a) = 전혀 사용 않음(a) ≥ 거의 사용 않음(ab) ≥ 어느 정도 사용함(abc) ≥ 보통(ac)														

다음으로, 다양한 유형의 매체 중에서 텔레비전, 라디오 등 시청각 중심 매체의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘보통(일주일에 약 1~2번)’(51.74)에 응답한 집단의 능력 점수 평균이 가장 높게 나타났다. 그 반면 ‘매우 자주 사용함(하루에 5번 이상)’(47.93)에 응답한 집단의 능력 점수 평균이 가장 낮게 나타났다. 전통 매체(올드 미디어) 중에서도 문자 매체와 달리, 시청각 혹은 영상 매체인 텔레비전과 라디오 사용 빈도가 높아질수록 읽기 능력 점수 평균값은 대체적으로 낮아지는 경향이 나타났다.

<표 IV-107> 매체 사용 빈도(텔레비전, 라디오)별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	ab		ab		a		b		c					
	전혀 사용		거의 사용		보통		어느 정도		매우 자주					
	않음 (n=66)		않음 (n=196)		보통 (n=975)		사용함 (n=2,125)		사용함 (n=1,638)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
TV, 라디오 사용 빈도 사후 검정 해석	51.65	8.28	51.33	8.85	51.74	9.43	49.51	9.20	47.93	8.47	30.57	4	4.995	<.001
보통(a) ≥ 전혀 사용 않음(ab) = 거의 사용 않음(ab) ≥ 어느 정도 사용(b) > 매우 자주 사용(c)														

다음으로, 다양한 유형의 매체 중에서 신매체(뉴 미디어)인 인터넷 포털(검색용)의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 사용함(하루에 5번 이상)’(50.30) 및 ‘어느 정도 사용함(하루에 1~4번)’(49.55)에 응답한 집단의 평균 능력 점수가 그 밖의 항목에 응답한 집단의 능력 평균 점수에 비해 높게 나타났다. 대체적으로 인터넷 포털(검색용) 사용 빈도가 감소할수록 읽기 능력 점수의 평균값도 낮아지는 경향이 나타났다.

<표 IV-108> 매체 사용 빈도(인터넷 포털(검색용))별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		b		b		a		a					
	전혀 사용		거의 사용		보통		어느 정도		매우 자주					
	않음		않음		보통		사용함		사용함					
	(n=88)		(n=225)		(n=673)		(n=1,844)		(n=2,170)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
인터넷														
포털	47.16	8.35	47.73	8.97	47.86	8.86	49.55	9.40	50.30	8.84	13.36	4	4,995	<.001
사용 빈도														
사후 검정	매우 자주 사용(a) = 어느 정도 사용(a) > 보통(b) = 거의 사용 않음(b) = 전혀 사용 않음(b)													
해석														

다음으로, 다양한 유형의 매체 중에서 누리소통망(SNS), 메신저 서비스의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .05$). 그러나 $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, 누리소통망(SNS), 메신저 사용 빈도에 따른 읽기 능력 점수 평균값에는 실질적인 차이가 나타나지 않았다. 이러한 결과가 나온 원인은 집단 간의 차이가 크지 않고, 집단 내에서의 분산이 크기 때문으로 추정할 수 있다.

<표 IV-109> 매체 사용 빈도(누리소통망(SNS), 메신저)별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	전혀 사용		거의 사용		보통		어느 정도		매우 자주					
	않음		않음		사용함		사용함		사용함					
	(n=279)		(n=368)		(n=854)		(n=1,720)		(n=1,779)					
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
누리소통망(SNS), 메신저 사용 빈도	48.66	8.52	48.65	8.51	49.12	8.85	49.78	9.14	49.79	9.35	2.61	4	4,995	.034
사후 검정 해석	매우 자주 사용(a) = 어느 정도 사용함(a) = 보통(a) = 전혀 사용 않음(a) = 거의 사용 않음(a)													

다음으로, 다양한 유형의 매체 중에서 온라인 커뮤니티(인터넷 카페, 블로그)의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘보통(일주일에 약 1~2번)’(50.61), ‘어느 정도 사용함(하루에 1~4번)’(50.20), ‘거의 사용 않음(한 달에 약 1~2번)’(50.16)에 응답한 집단의 읽기 능력 점수 평균이 상대적으로 높게 나타났다. 그 반면 ‘매우 자주 사용함(하루에 5번 이상)’(46.83)에 응답한 집단의 능력 점수 평균이 가장 낮게 나타났다. 다만, 집단 간 능력 평균 점수 차이 간에는 일정한 경향성이 두드러지게 나타나지 않았는데, 이는 다른 변수의 개입 가능성을 시사한다.

<표 IV-110> 매체 사용 빈도(온라인 커뮤니티(인터넷 카페, 블로그))별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		a		a		a		c					
	전혀 사용 않음 (n=708)		거의 사용 않음 (n=864)		보통 (n=1,207)		어느 정도 사용함 (n=1,418)		매우 자주 사용함 (n=803)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
온라인 커뮤니티 사용 빈도 사후 검정 해석	48.61	8.03	50.16	8.65	50.61	9.07	50.20	9.48	46.83	9.21	17.18	4	4,995	<.001
보통(a) = 어느 정도 사용함(a) = 거의 사용 않음(a) ≥ 전혀 사용 않음(b) ≥ 매우 자주 사용(c)														

다음으로, 다양한 유형의 매체 중에서 온라인 동영상 플랫폼, 인터넷 동영상 서비스(OTT)의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘전혀 사용 않음’(47.03)에 응답한 집단의 읽기 능력 점수 평균이 그 밖의 집단에 비해 같거나 더 낮게 나타났다.

<표 IV-111> 매체 사용 빈도(온라인 동영상 플랫폼, OTT)별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	b		ab		a		a		a					
	전혀 사용 않음 (n=268)		거의 사용 않음 (n=391)		보통 (n=1,152)		어느 정도 사용함 (n=1,917)		매우 자주 사용함 (n=1,272)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
동영상 플랫폼, OTT 사용 빈도	47.03	7.68	48.90	9.25	49.26	9.02	50.06	9.36	49.68	8.88	7.506	4	4,995	<.001
사후 검정 해석	어느 정도 사용(a) = 매우 자주 사용(a) = 보통(a) ≥ 거의 사용 않음(ab) ≥ 전혀 사용 않음(b)													

끝으로, 다양한 유형의 매체 중에서 게임, 웹툰, 웹소설의 사용 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 매체 사용 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘매우 자주 사용함(하루에 5번 이상)’(47.36)에 응답한 집단의 읽기 능력 점수 평균이 그 밖의 집단에 비해 낮게 나타났다.

<표 IV-112> 매체 사용 빈도(게임, 웹툰, 웹소설)별 읽기 영역 점수의 차이

변인	능력 점수(T점수)					F-검정 결과
	a	a	a	a	b	

	전혀 사용 않음 (n=1,831)		거의 사용 않음 (n=1,099)		보통 (n=1,057)		어느 정도 사용함 (n=700)		매우 자주 사용함 (n=313)		F	df1	df2	p
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
게임, 웹툰														
웹소설	49.82	8.29	49.54	9.51	49.47	9.22	49.78	10.04	47.36	9.24	5.103	4	4,995	<.001
사용 빈도														
사후 검정	전혀 사용 않음(a) = 어느 정도 사용(a) = 거의 사용 않음 = 보통(a) > 매우 자주 사용(b)													
해석														

(2) 읽기 빈도

읽기 빈도에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 읽기 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘거의 매일’(56.24)에 응답한 집단의 읽기 능력 평균 점수가 가장 높았다. 그 반면 ‘전혀 읽지 않음’(46.56)을 선택한 집단의 읽기 능력 평균 점수는 가장 낮게 나타났다. 전체적으로 읽기 빈도가 높아질수록 읽기 능력 점수 또한 함께 높아지는 경향이 나타났다.

<표 IV-113> 읽기 빈도별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	e		d		c		b		a					
	전혀 읽지 않음 (n=1,432)		1년에 2~3회 (n=1,527)		1달에 2~3회 (n=1,005)		1주일에 2~3회 (n=599)		거의 매일 (n=437)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
읽기 빈도	46.56	8.88	47.70	8.79	51.45	8.52	53.14	8.27	56.24	6.81	167.3	4	4,995	<.001
사후 검정 해석	거의 매일(a) > 1주일에 2~3회(b) > 1달에 2~3회(c) > 1년에 2~3회(d) > 전혀 읽지 않음(e)													

(3) 읽기를 하지 못하는 이유

다음으로, ‘글 읽기를 하지 못하는 이유’에 대한 가설적 원인을 다섯 가지 정도로 추려, 중복 선택이 가능하도록 조사 문항을 구성하고 특정 선지를 선택한 경우를 ‘해당’ 집단, 선택하지 않은 경우를 ‘비해당’ 집단으로 구분하여 두 집단 간 평균 차이를 독립표본 t검정 방식으로 검정하였다. 그 결과, ‘습관 부족’과 ‘흥미 부족’을 제외한 나머지 변인들에서는 해당 이유의 선택 유무에 따른 읽기 능력 점수의 평균 값에 통계적으로 유의한 차이가 나타났다. 우선 ‘직장, 학업 등으로 인한 시간 부족’을 선택한 집단의 능력 점수(50.73)는 이를 선택하지 않은 집단의 능력 점수(48.82)에 비해 높았다. ‘내용 이해 어려움’을 선택한 집단의 능력 점수(43.91)는 이를 선택하지 않은 집단의 능력 점수(50.35)에 비해 낮게 나타났다.

‘독서보다 핸드폰, 게임 선호’를 선택한 집단의 능력 점수 평균(51.78)은 그렇지 않은 집단의 능력 점수 평균(48.85)에 비해 높게 나타났는데, 결과 해석 시 분석에 영향을 미치지만 통제되지 않은 외생변수 개입 가능성을 고려해야 할 필요가 있다.

<표 IV-114> 독서를 하지 못하는 이유별 읽기 영역 점수의 차이

변인	능력 점수(T점수)				t-검정 결과		
	비해당		해당		t	df	p
	평균	SD	평균	SD			
직장, 학업 등으로 인한 시간 부족	48.82 (n=3,141)	9.09	50.73 (n=1,859)	8.98	7.24	3939.4	<.001
내용 이해 어려움	50.35 (n=4,360)	8.78	43.91 (n=640)	9.20	-16.63	819.3	<.001
습관 부족	49.46 (n=2,439)	9.27	49.59 (n=2,561)	8.93	0.51	4961.4	0.612
흥미 부족	49.41 (n=3,068)	9.67	49.71 (n=1,932)	8.10	1.17	4611.6	0.243
독서보다 핸드폰, 게임 선호	48.85 (n=3,845)	9.39	51.78 (n=1,155)	7.60	10.87	2310.8	<.001

(4) 읽기의 가치에 대한 인식

읽기에 대한 가치 인식에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 인식에 따른 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘수준 높은 삶을 위해 필수’(52.95)에 응답한 집단의 읽기 능력 점수 평균이 통계적으로 가장 높게 나타났다. 이를 뒤따라 ‘학업과 진로에 도움이 됨’(50.87), ‘글 읽기가 아니더라도 정보를 얻을 수 있음’(47.64), ‘시간 보내기용으로 좋음’(47.25), ‘글 읽기는 다른 일에 방해가 됨’(42.90) 순서로 읽기 능력 점수 평균값이 낮아졌다.

<표 IV-115> 읽기에 대한 가치에 따른 읽기 영역 점수의 차이

변인		능력 점수(T점수)										F-검정 결과			
		d		c		c		b		a					
		글 읽기는 다른 일에 방해가 됨 (n=252)		글 읽기가 아니더라도 정보를 얻을 수 있음 (n=1,395)		시간 보내기 용으로 좋음 (n=806)		학업과 진로에 도움이 됨 (n=1,241)		수준 높은 삶을 위해 필수 (n=1,306)					
		평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
읽기에 대한		42.90	9.60	47.64	8.57	47.25	9.83	50.87	8.59	52.95	7.89	125.2	4	4,995	<.001

가치 인식	
사후 검정	수준 높은 삶을 위해 필수(a) > 학업과 진로에 도움이 됨(b)
해석	> 글 읽기가 아니라도 정보를 얻을 수 있음(c) = 시간 보내기용으로 좋음(c) > 글 읽기는 다른 일에 방해가 됨(d)

(5) 읽기 효능감

자신의 읽기 능력 수준에 대한 인식, 즉 읽기 효능감에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 인식에 따른 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘어떤 글이든지 어려움 없이 이해함’(52.86)에 응답한 집단의 읽기 능력 점수 평균이 가장 높게 나타났다. 이어서 ‘글 내용을 대체로 이해함’(51.36), ‘가끔 이해할 수 없는 글이 있음’(47.03)이 그 뒤를 따랐고, ‘사진이나 그림 위주의 글은 이해할 수 있음’(42.15)과 ‘글자는 읽어도 내용을 이해하지 못함’(41.18)에 응답한 집단의 읽기 능력 점수 평균이 집단 간 차이 없이 다른 집단들에 비해 가장 낮게 나타났다.

<표 IV-116> 읽기 효능감에 따른 읽기 영역 점수의 차이

능력 점수(T점수)																													
변인	d		d		c		b		a		F-검정 결과																		
	글자는		사진이나		가끔 이해할		글 내용을		어떤																				
	읽어도		그림 위주의		수 없는		대체로		글이든지																				
	내용을		글은 이해할		글이 있음		이해함		어려움 없이																				
	이해하지		수 있음						이해함																				
못함																													
(n=203)		(n=432)		(n=1,089)		(n=2,218)		(n=1,058)																					
평균		SD		평균		SD		평균		SD		F		df1		df2		p											
읽기		41.18		9.08		42.15		8.93		47.03		8.21		51.36		8.14		52.86		8.63		227.3		4		4,995		<.001	
효능감																													
사후 검정		어떤 글이든지		어려움 없이		이해함(a)		>		글 내용을		대체로		이해함(b)		>		가끔 이해할		수 없는		글이 있음(c)							
해석																													

(6) 1회 읽기 분량

1회 읽기 분량에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 인식에 따른 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘교과서 5쪽 이상’(53.73)에 응답한 집단의 읽기 능력 점수 평균이 가장 높게 나타났으며, 1회 읽기 분량이 줄어들수록 읽기 능력 점수가 낮아지는 경향이 나타났다.

<표 IV-117> 읽기 분량별 읽기 영역 점수의 차이

변인	능력 점수(T점수)	F-검정 결과
----	------------	---------

	d		d		c		b		a					
	전혀 안 읽음		짧은 요약 글		교과서 1쪽		교과서 2~4쪽 정도		교과서 5쪽 이상					
	(n=1,092)		(n=1,104)		(n=922)		(n=829)		(n=1,053)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p
읽기 분량	46.83	9.60	47.61	8.79	49.13	8.63	50.75	8.24	53.73	8.21	1046	4	4,995	<.001
사후 검정 해석	교과서 5쪽 이상(a) > 교과서 2~4쪽 정도(b) > 교과서 1쪽(c) > 짧은 요약 글(d) = 전혀 안 읽음(d)													

(7) 디지털 기기 사용 시간

디지털 기기 사용 시간에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 인식에 따른 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘하루 5시간 이상~하루 종일’(53.41)에 응답한 집단의 읽기 능력 점수 평균이 가장 높게 나타났다. 그 반면 ‘연락 외에 전혀 사용 없음’(45.94)을 선택한 집단의 점수 평균은 가장 낮았다. 다만 디지털 기기 사용 시간에 따른 읽기 능력 점수에는 주된 경향이 나타나지 않았는데, 이는 심층적인 추가 분석의 필요성을 시사한다.

<표 IV-118> 디지털 기기 사용 시간별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	d		c		b		c		a					
	연락 외에 전혀 사용 않음 (n=423)		하루 1시간 정도 (n=1,488)		하루 3시간 정도 (n=2,355)		하루 5시간 정도 (n=558)		하루 5시간 이상 ~ 하루 종일 (n=176)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
디지털 기기	45.94	9.49	49.13	8.61	50.44	9.02	48.20	9.55	53.41	8.23	35.06	4	4,995	<.001
사용 시간 사후 검정	하루 5시간 이상~하루 종일(a) > 하루 3시간 정도(b) 													

(8) 주 이용 매체

주로 이용하는 매체별 읽기 영역 점수(T점수) 차이 분석 결과, 인식에 따른 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과, ‘짧은 동영상’(52.07)을 선택한 집단의 읽기 능력 점수 평균이 가장 높게 나타났다. 그 뒤를 이어 ‘긴 동영상’(50.66), ‘글로 된 인터넷 자료’(50.56)를 선택한 집단의 읽기 능력 점수 평균이 높게 나타났으며,

나머지 ‘책, 신문 등의 인쇄 매체’(47.66)와 ‘텔레비전’(47.35)을 선택한 집단의 읽기 능력 점수 평균은 가장 낮은 것으로 드러났다.

<표 IV-119> 주로 이용하는 매체별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	c		c		b		a		b					
	책, 신문 등의 인쇄 매체 (n=213)		텔레비전 (n=1,825)		글로 된 인터넷 자료 (n=1,124)		짧은 동영상 (n=803)		긴 동영상 (n=1,035)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
	F	df1	df2	p										
주로 이용하는 매체	47.66	10.71	47.35	8.75	50.56	9.48	52.07	8.09	50.66	8.75	53.95	4	4,995	<.001
사후 검정 해석	짧은 동영상(a) > 긴 동영상(b) = 글로 된 인터넷 자료(b) > 책, 신문 등의 인쇄 매체(c) = 텔레비전(c)													

(9) 독서 문화 참여 경험

마지막으로 독서 문화 참여 경험에 따른 읽기 영역 점수(T점수) 차이 분석 결과, 인식에 따른 평균 점수 차이가 통계적으로 유의하지 않은 것으로 확인되었다($p=.209$).

<표 IV-120> 독서 문화 참여 경험별 읽기 영역 점수의 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		a		a		a		a					
	참여 경험 없음		한두 번 참여 경험 있음		가끔 참여함 (연간 1회)		자주 참여 (연간 1회 이상)		회원 또는 정기적 참여					
	(n=4,126)		(n=551)		(n=207)		(n=100)		(n=16)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
참여 경험	49.46	8.95	49.91	9.59	49.07	10.52	51.31	8.71	48.25	10.13	1.47	4	4,995	.209

이상으로 심층 변인에 따른 읽기 능력 점수(T점수)의 차이 분석 결과를 보고하였다. 대체로 문자 중심의 문식 환경에 노출될수록, 읽기에 대한 중요성을 인식하고 있을수록 읽기 능력도 높다는 결과를 보였다. 다만, 전체적으로 독자의 문식 환경이 질적으로 열악해지고 있는 중에도 인터넷, 스마트폰 등 신매체(뉴 미디어) 중심의 매체 환경에 익숙할수록 읽기 능력이 높은 것으로 보고된 점은 특기할 만하다. 앞으로 새로운 읽기 세대의 교육과 변화된 읽기 환경 조성에 대한 정책 제언이 이뤄지기 위해서는 후속 심층 연구가 요구된다 하겠다.

3) 문법·규범 영역 심층 변인 분석

문법·규범 영역의 심층 변인에 따른 능력 점수(T점수)의 평균 차이를 분석하기 위하여 일원배치 분산 분석(one-way ANOVA)을 수행하였으며, 평균 점수 차이가 통계적으로 유의한 경우에는 Scheffé 보정 방법을 적용하여 사후 검정을 실시하였다.

(1) 어문 규범(맞춤법, 띄어쓰기 등)과 관련된 정보 탐색 빈도

어문 규범과 관련된 정보 탐색 빈도에 따른 문법·규범 영역 능력 점수(T점수) 차이 분석 결과, 탐색 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .001$). ‘매우 자주 찾아본다(하루에 1번 이상)’(53.26), ‘거의 찾아보지 않는다(한 달에 약 1~2번)’(51.57)에 응답한 집단의 능력 평균 점수가 ‘자주 찾아본다(일주일에 약 4~6번)’(51.17), ‘보통이다(일주일에 약 1~3번)’(50.77), ‘전혀 찾아보지 않는다(한 달에 한 번도 검색하지 않음)’(49.2)에 응답한 집단과 비교하여 더 높은 것으로 나타났고, $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과 이러한 차이가 통계적으로 유의한 것으로 확인되었다. 단, 통계 분석 결과 해석 시 ‘매우 자주 찾아본다’로 응답한 사람의 수($n=6$)가 극히 적다는 점에 유의해야 한다.

<표 IV-121> 문법·규범 영역 능력 점수의 어문 규범 정보 탐색 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과				
	ab		a		ab		ab		a						
	전혀		거의		보통		자주 찾아봄		매우 자주						
	찾아보지		찾아보지		(n=397)		(n=108)		찾아봄						
	않음		않음						(n=6)						
(n=3,168)		(n=1,321)													
평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p
어문 규범 탐색 빈도 사후 검정 해석	49.2	8.51	51.57	8.47	50.77	8.89	51.17	5.57	53.26	5.72	19.57	4	4,995	<.001	
매우 자주 찾아봄(a) = 거의 찾아보지 않음(a) ≥ 자주 찾아봄(ab) = 보통(ab) = 전혀 찾아보지 않음(ab)															

(2) 국립국어원 제공 서비스 이용 빈도

먼저, 국립국어원 제공 서비스 중 <표준국어대사전> 또는 <우리말샘>의 이용 빈도에 따른 문법·규범 영역 능력 점수(T점수) 차이 분석 결과, 이용 빈도별 평균 점수 차이는 통계적으로 유의하지 않은 것으로 확인되었다($p = .29$).

<표 IV-122> 문법·규범 영역 능력 점수의 온라인 사전 이용 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과				
	a		a		a		a		a						
	전혀 이용 않음 (n=3,715)		거의 이용 않음 (n=935)		보통 (n=293)		자주 이용함 (n=53)		매우 자주 이용함 (n=4)						
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD	F	df1	df2	p	
온라인															
사전	49.92	8.48	50.47	8.84	49.42	9.08	50.65	8.79	47.62	11.79	1.25	4	4,995	.29	
이용 빈도															

다음으로, <온라인가나다>의 이용 빈도에 따른 문법·규범 영역 능력 점수(T점수) 차이 분석 결과, 탐색 빈도별 평균 점수 차이가 통계적으로 유의한 것으로 확인되었다($p < .01$). ‘매우 자주 이용한다(하루에 1번 이상)’(60.19), ‘전혀 이용하지 않는다(한 달에 한 번도 검색하지 않음)’(50.15)에 응답한 집단의 능력 평균 점수가 ‘거의 이용하지 않는다(한 달에 약 1~2번)’(49.88), ‘자주 이용한다(일주일에 약 4~6번)’(48.79), ‘보통이다(일주일에 약 1~3번)’(48)에 응답한 집단과 비교하여 더 높은 것으로 나타났다. $\alpha < .05$ 수준에서의 대응별 차이 검정(Scheffé)을 실시한 결과 이러한 차이가 통계적으로 유의한 것으로 확인되었다. 단, 앞서 어문 규범 관련 정보의 탐색 빈도에서와 마찬가지로, 이 경우에도 통계 분석 결과 해석 시 ‘매우 자주 이용한다’로 응답한 사람의 수($n=2$)가 극히 적다는 점에 유의해야 한다.

<표 IV-123> 문법·규범 영역 능력 점수의 온라인가나다 이용 빈도별 차이

변인	능력 점수(T점수)										F-검정 결과			
	a		ab		ab		ab		a					
	전혀 이용 않음 (n=3,951)		거의 이용 않음 (n=752)		보통 (n=239)		자주 이용함 (n=56)		매우 자주 이용함 (n=2)					
	평균	SD	평균	SD	평균	SD	평균	SD	평균	SD				
											F	df1	df2	p
온라인 가나다 이용 빈도 사후 검정 해석	50.15	8.49	49.88	8.84	48	9.17	48.79	8.62	60.19	0.96	4.58	4	4,995	<01
매우 자주 이용함(a) = 전혀 이용 않음(a) ≥ 거의 이용 않음(ab) = 자주 이용함(ab) = 보통(ab)														

4) 쓰기 영역 심층 변인 분석

이번 연구에서는 이전 2013년 제1차 조사 및 2018년 제2차 조사에 비해 배경 변인에 대한 설문 항목을 증배하여, 쓰기 능력과 관련한 쓰기 인식, 쓰기 실태, 쓰기 환경 등에 대한 심층적 조사를 실시하였다. 해당 조사 결과를 바탕으로 심층 변인이 수준별 쓰기 능력 점수에 미치는 영향을 확인하기 위해 중다회귀분석 및 단계별 회귀분석을 수행하였다. 중다회귀분석을 수행하기 위한 설명 변수로는 글쓰기에 대한 인식, 글쓰기 빈도(비자발적/자발적), 긍정적인 글쓰기 경험 빈도, 쓰기를 하지 못하는 이유, 쓰

기 과정에서 생성형 인공지능 활용 빈도, 글쓰기 조력자의 존재 여부(가정 내/사회 내)를 투입하였다. 이때, 쓰기를 하지 못하는 이유는 범주형 변수로 처리하였으며, 나머지 변인은 수치형 변수로 처리하였다. 쓰기 영역에서 조사된 심층 변인의 목록은 표 <IV-124>와 같다.

<표 IV-124> 쓰기 영역 심층 변인 목록

심층 변인	설명	변인 유형
DQ1_1	글쓰기에 대한 인식(자기 이해 및 생각과 감정 표현 목적)	수치형
DQ1_2	글쓰기에 대한 인식(일상, 학업, 업무 수행 목적)	수치형
DQ1_3	글쓰기에 대한 인식(관계 형성 목적)	수치형
DQ2_1	글쓰기 빈도(비자발적/공적)	수치형
DQ3_1	글쓰기 빈도(자발적/사적)	수치형
DQ4_1	긍정적인 글쓰기 경험 빈도	수치형
DQ5_1	쓰기를 하지 못하는 이유	범주형
DQ6_1	쓰기 과정에서 생성형 인공지능 활용 빈도	수치형
DQ7_1	글쓰기 조력자의 존재(가정 내)	수치형
DQ8_1	글쓰기 조력자의 존재(사회 내)	수치형

범주형 변인인 ‘쓰기를 하지 못하는 이유’ 변인을 제외한 나머지 변인들의 평균은 최저 1.978에서 3.990, 표준편차는 최소 .715에서 최고 1.243, 왜도는 절댓값 최소 .090에서 최고 .703, 첨도는 절댓값 최소 .336에서 최고 1.233으로 나타났다. 쓰기 영역 심층 변인의 기술 통계는 <표 IV-125>, <표 IV-126>과 같다.

<표 IV-125> 쓰기 영역 심층 변인의 기술 통계

심층 변인	표본 수	평균	표준편차	왜도	첨도
DQ1_1	3,000	3.990	.715	-.703	1.233
DQ1_2	3,000	3.899	.856	-.667	.418
DQ1_3	3,000	3.735	.791	-.518	.336
DQ2_1	3,000	2.748	1.243	.164	-.972
DQ3_1	3,000	2.515	1.096	.283	-.706
DQ4_1	3,000	2.434	1.009	.294	-.616
DQ6_1	3,000	1.978	1.007	.602	-.748
DQ7_1	2,946	2.844	.985	-.143	-.872
DQ8_1	3,000	2.810	.972	-.090	-.836

<표 IV-126> 쓰기 영역 심층 변인(쓰기를 하지 못하는 이유)의 기술 통계

심층 변인	표본 수	빈도(비율)				
		흥미가 없어서	어떻게 써야 할지 몰라서	인터넷, 동영상 등이 더 좋아서	시간이 없어서	다른 사람이 읽는 게 부담돼서
DQ5_1	3,000	541 (18%)	608 (20%)	1,249 (42%)	469 (16%)	133 (4%)

(1) 전체 심층 변인 분석

국민의 쓰기 능력 수준에 영향을 주는 예측 변인을 탐색하기 위해 심층 변인을 활용하여 단계적 회귀 분석(stepwise) 방식으로 중다회귀분석을 실시하였다. 각 예측 변인이 국민의 쓰기 능력을 유의하게 예측하는지 중다회귀분석의 방식으로 분석한 결과는 <표 IV-127>, <표 IV-128>와 같다.

<표 IV-127> 쓰기 능력 수준 및 전체 심층 변인 간 상관

	θ	DQ1_1	DQ1_2	DQ1_3	DQ2_1	DQ3_1	DQ4_1	DQ6_1	DQ7_1	DQ8_1
θ	1	.131	.092	.078	.160	.095	.081	.046	-.046	.026
DQ1_1		1	.482	.509	.173	.173	.109	.102	.080	.042
DQ1_2			1	.483	.266	.156	.117	.137	.117	.106
DQ1_3				1	.146	.141	.134	.052	.076	.022
DQ2_1					1	.495	.436	.358	.203	.263
DQ3_1						1	.644	.268	.213	.203
DQ4_1							1	.251	.250	.287
DQ6_1								1	.215	.276
DQ7_1									1	.593
DQ8_1										1

<표 IV-128> 심층 변인에 따른 쓰기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과

	변인	계수	표준화 계수	표준 오차	t	유의 확률
모형	(Intercept)	45.456	NA	.098	46.518	<.001
	DQ2_1	.103	.148	.014	7.591	<.001
	DQ7_1	-.093	-.107	.020	-4.621	<.001
	DQ1_1	.076	.063	.022	3.401	<.001
	DQ8_1	.049	.055	.021	2.330	.020
결정계수	$R^2=0.034$, 조정된 $R^2=0.032$					
모형 적합도 검정	$F(4, 2941)=25.557$, $p<.001$					

자발적(사적) 글쓰기 빈도(DQ3_1), 긍정적인 글쓰기 경험 빈도(DQ4_1), 글을 못 쓰는 이유(DQ5_1), 쓰기 과정에서 생성형 인공지능 활용 빈도(DQ6_1)의 예측력은 유의하지 않아 제외되었고, 쓰기에 대한 인식, 비자발적(공적) 글쓰기 빈도, 글쓰기 조력자의 존재(가정 내/사회 내)는 유의미한 예측 변인으로 검증되었다. 글쓰기 빈도의 경우 비자발적(공적) 글쓰기 빈도가 높을수록(DQ2_1) 쓰기 능력 수준이 높게 나타났다($\eta^2=0.02$; 작은 효과 크기). 또한 글쓰기에 대한 인식 중 자기 이해와 생각·감정 표현을 위해 글쓰기가 필요하다고 생각할수록 쓰기 능력 수준이 높게 나타났다($\eta^2=0.008$; 효과 크기 거의 없음). 글쓰기 조력자의 존재 여부와 관련하여 가정 내에 글쓰기 조력자가 존재할수록(DQ7_1) 쓰기 능력 수준이 낮게 나타났으며($\eta^2=0.005$; 효과 크기 거의 없음), 사회 내에 글쓰기 조력자가 존재할수록(DQ8_1) 쓰기 능력 수준이 높게 나타났다($\eta^2=0.002$; 효과 크기 거의 없음). 전체 모형은 유의하게 나타나기는 하였으나, 설명력 3.4%로 매우 낮은 수준이었다($R^2=0.034$). 따라서 심층 변인 가운데에서는 글쓰기 빈도 가운데서도 특히 ‘비자발적(공적) 글쓰기 빈도’가 쓰기 능력의 차이를 불러온다는 점에 주목할 필요가 있으며, 효과 크기가 거의 없기는 하지만 가정 내 글쓰기 조력자의 존재 여부와 사회 내 글쓰기 조력자의 존재 여부가 쓰기 능력에 대해 서로 상반되는 예측력을 보이는 이유도 면밀히 살펴볼 필요가 있다.

(2) 인구통계학적 변인과 심층 변인 종합 분석

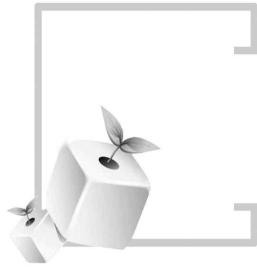
보다 심층적인 분석을 위해 인구통계학적 변인과 심층 변인을 모두 투입하여 중다회귀분석한 결과는 다음 표 <IV-129>와 같다.

<표 IV-129> 인구통계학적 변인 및 심층 변인에 따른 쓰기 능력 점수(T점수) 예측 모형의 중다회귀분석 결과

변인	회귀 계수	표준화 계수	표준 오차	t	유의 확률
(상수)	47.05	NA	0.98	47.90	<.001***
대재 이상	1.95	0.23	0.36	5.40	<.001***
비자발적 글쓰기 빈도	0.68	0.08	0.14	4.84	<.001***
가정 내 글쓰기 조력자	-0.87	-0.10	0.16	-5.31	<.001***
고졸 미만	-2.16	-0.25	0.57	-3.76	<.001***
대도시	0.84	0.10	0.31	2.71	.007**
쓰기 인식-자기 이해·표현	0.58	0.07	0.22	2.59	.010***
잔여 표준오차	8.34 (df=2,939)				
R^2	0.054	조정된 R^2		0.052	
F 통계량	27.8 (df=6, 2,939)	p		<0.001	

*** p <.001 ** p <.01 * p <.05

전체 모형의 설명력은 5.4%로 낮은 수준이었으나($R^2=0.054$), 통계적으로 유의한 수준이었다(p <.001). 우선 학력이 ‘대재 이상’인 경우 그렇지 않은 경우에 비해 쓰기 능력 수준이 높게 나타났으며($\eta^2=0.03$; 작은 효과 크기), 비자발적(공적) 글쓰기 빈도가 높을수록(DQ2_1) 쓰기 능력 수준이 높게 나타났다($\eta^2=0.008$; 효과 크기 거의 없음). 인구통계학적 변인을 함께 고려했을 때 ‘비자발적(공적) 글쓰기 빈도’의 효과 크기가 크게 줄어든 점(0.02→0.008)에 주목할 필요가 있다. 대재 이상 학력과 비자발적(공적) 글쓰기 빈도의 상관이 $r=0.422$ 로 높은 편이라는 점을 고려할 때, 비자발적(공적) 글쓰기를 많이 해서 쓰기 능력이 높아지는 효과는 미미하며, 대재 이상의 학력자일수록 비자발적(공적) 글쓰기를 많이 하다 보니 비자발적(공적) 글쓰기와 쓰기 능력 간 상관이 높게 나타난 것으로 분석할 수 있다. 또한 가정 내에 글쓰기 조력자가 존재할수록(DQ7_1) 쓰기 능력 수준이 낮게 나타났으며($\eta^2=0.008$; 효과 크기 거의 없음), 학력이 ‘고졸 미만’인 경우(EDU2) 그렇지 않은 경우에 비해 쓰기 능력 수준이 낮게 나타났다($\eta^2=0.005$; 효과 크기 거의 없음). 앞서 심층 변인 중 ‘사회 내 글쓰기 조력자’가 유의미한 예측 변인으로 나타났으나, 학력이 높을수록 사회 내 글쓰기 조력자가 증가한다는 점에서 종합 분석에서는 예측력이 유의하지 않은 것으로 나타나 해당 변인은 제외되었다. ‘가정 내 글쓰기 조력자’의 경우, 쓰기 능력과 부적 상관을 보이는데, 이는 조사 대상자들이 가정 내의 글쓰기 문화 조성의 측면보다 ‘글을 쓸 때 도움을 받는다’는 측면에 주목하여 응답하였을 가능성을 고려해 볼 수 있다. 또한 ‘연령대’ 변인이 제외된 것 또한 연령이 높을수록 ‘비자발적 글쓰기 빈도’가 낮아진다는 점을 함께 고려할 필요가 있다. 아울러 지역이 ‘대도시’인 경우(REGION1) 그렇지 않은 경우에 비해 쓰기 능력 수준이 높게 나타났으며($\eta^2=0.003$; 효과 크기 거의 없음), 자기 이해와 생각·감정 표현을 위해 글쓰기가 필요하다고 생각할수록(DQ1_1) 쓰기 능력 수준이 높게 나타났다($\eta^2=0.002$; 효과 크기 거의 없음).



제 5 장

제1차~제3차 실태 조사 비교



5.1. 제1차~제3차 조사 시기 국민의 국어능력 추이 요약

제1차 국민의 국어능력 실태 조사는 2013년에, 제2차 조사는 2018년에 실시되었다. 제1차 조사는 20세~59세의 성인 남녀 3,000명을 대상으로, 조사 도구를 ‘가’ 형과 ‘나’ 형으로 구분하여 듣기, 말하기, 읽기, 쓰기, 문법의 다섯 영역에 걸쳐 실시되었다. 제2차 조사는 20세~69세의 성인 남녀 5,500명을 대상으로, ‘가’ 형, ‘나’ 형 및 ‘통합형’ 검사지를 사용하여 듣기, 말하기, 읽기, 쓰기, 문법, 디지털의 여섯 영역에 걸쳐 실시되었다.

제2차 국민의 국어능력 실태 조사가 이루어진 후에는 동백분위 동등화를 거쳐 제1차 조사와 비교하였고, 2013년 제1차 실태 조사 보고서와 2018년 제2차 실태 조사 보고서에 제시된 제1차 조사와 제2차 조사의 영역별 국어능력 비율 추이는 다음과 같다.

<표 V-1> 제1차 조사와 제2차 조사의 영역별 국어능력 비율 추이

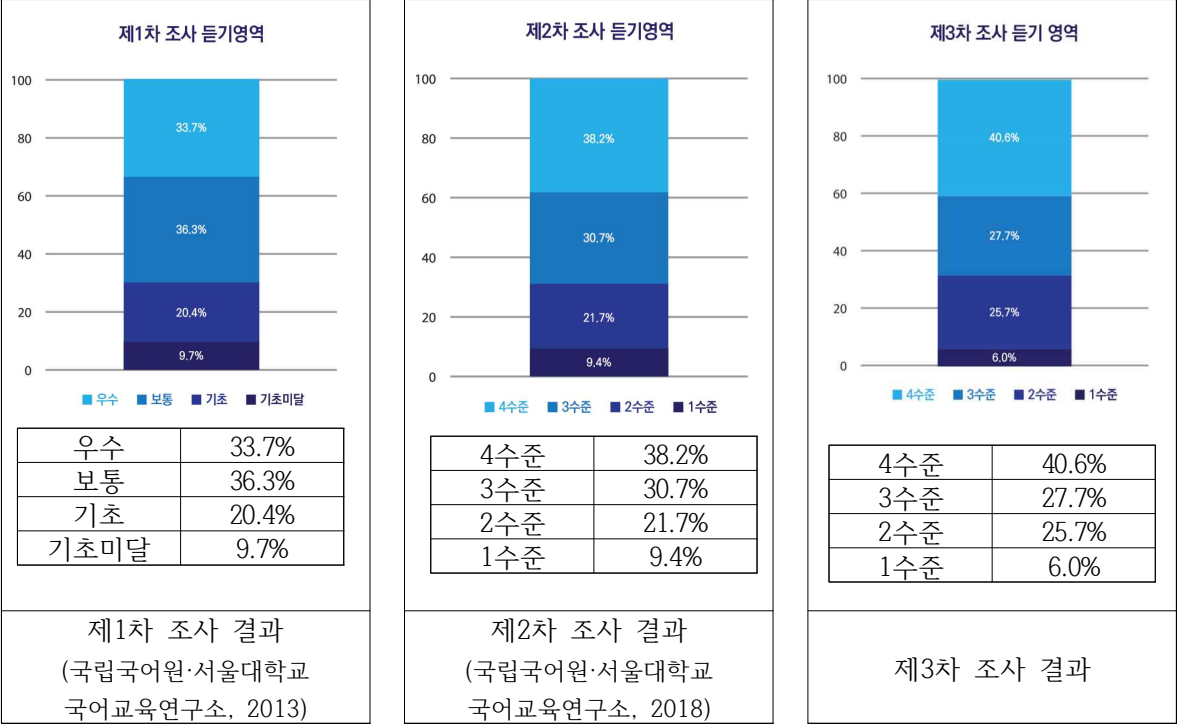
		2013년				2018년			
		(국립국어원·서울대학교 국어교육연구소, 2013)				(국립국어원·서울대학교 국어교육연구소, 2018)			
		우수	보통	기초	기초 미달	4수준	3수준	2수준	1수준
듣기	조사 대상자 수(명)	478	514	288	137	1,147	919	652	282
	비율(%)	33.7	36.3	20.4	9.7	38.2	30.7	21.7	9.4
말하기	조사 대상자 수(명)	40	470	702	205	139	799	1,469	532
	비율(%)	2.8	33.2	49.5	14.5	4.7	27.2	50.0	18.1
읽기	조사 대상자 수(명)	662	1,110	882	262	1,160	2,117	1,821	402
	비율(%)	22.7	38.1	30.2	9.0	21.1	38.5	33.1	7.3
쓰기	조사 대상자 수(명)	55	341	829	275	121	790	1,254	835
	비율(%)	3.6	22.7	55.3	18.4	4.0	26.4	41.8	27.8
문법	조사 대상자 수(명)	522	1,095	949	350	1,326	2,290	1,274	610
	비율(%)	17.9	37.5	32.5	12.0	24.1	41.6	23.2	11.1

제3차 국민의 국어능력 실태 조사는 연차별로 영역을 나누어 실시하였고, 2023년에 듣기, 듣기·말하기, 말하기 영역, 2024년에 읽기 영역 및 문법·규범 영역, 2025년에 쓰기 영역 실태 조사가 이루어졌다. 2018년 조사와 비교하면 듣기·말하기 영역이 신설되었고, 문법 영역은 문법·규범 영역으로 변경되었다. 또한 디지털 영역은 영역 층위에서 사라졌고, 언어 환경과 매체 환경의 변화를 반영하여 전 영역에 디지털 요소를 반영하는 변화가 이루어졌다. 2023년 듣기, 듣기·말하기, 말하기 영역은 만 20세~69세의 성인 남녀 약 3,000명을 대상으로 검사 도구를 일원화하여 실시하였다. 2024년 읽기 영역과 문법·규범 영역은 만 20세~69세의 성인 남녀 약 5,050명을 대상으로 하되, 방문 면접 조사 5,000명, 집합 조사 50명으로 조사 방식을 구분하였으며, 검사 도구를 ‘가’ 형과 ‘나’ 형으로 나누어 실시하였다. 2025년 쓰기 영역은 만 20세~69세의 성인 남녀 약 3,050명을 대상으로 하되, 방문 면접 조사 3,000명, 집합 조사 50명으로 조사 방식을 구분하였으며, 검사 도구를 일원화하여 실시하였다. 2025년에는 수정된 앵고프 방식을 사용하여 성취 수준별 분할 점수를 새로 산출하였으며, 이에 따른 제3차 국민의 국어능력 실태 조사의 영역별 수준 비율은 4장에서 제시한 바와 같다.

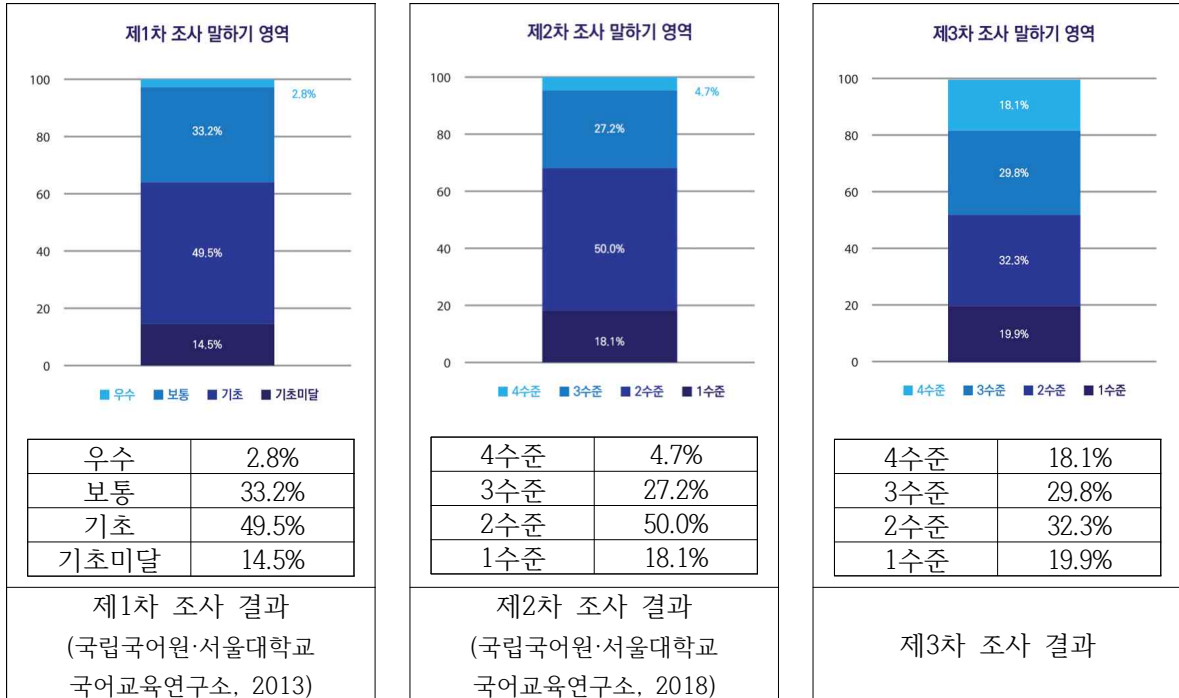
제3차 국민의 국어능력 실태 조사에서는 앞서 언급한 대로 언어 환경, 매체 환경의 변화를 고려하여 디지털 요소를 전 영역에 반영하였고, 듣기·말하기 영역의 신설, 문법에서 문법·규범으로의 영역 변화

등 영역 차원의 변화도 이루어졌다. 또한 제3차 조사에서는 수정된 앵고프 방식으로 성취 수준별 분할 점수를 새로 산출하였기 때문에, 제3차 조사 결과 산출된 수준별 비율을 이전 조사와 동일한 선상에서 비교하기는 어렵다. 단, 제1차부터 제3차까지의 조사 결과를 정리하는 차원에서 영역별로 듣기, 말하기, 읽기, 쓰기, 문법/문법·규범 영역의 수준별 비율을 나란히 제시하면 다음과 같다.

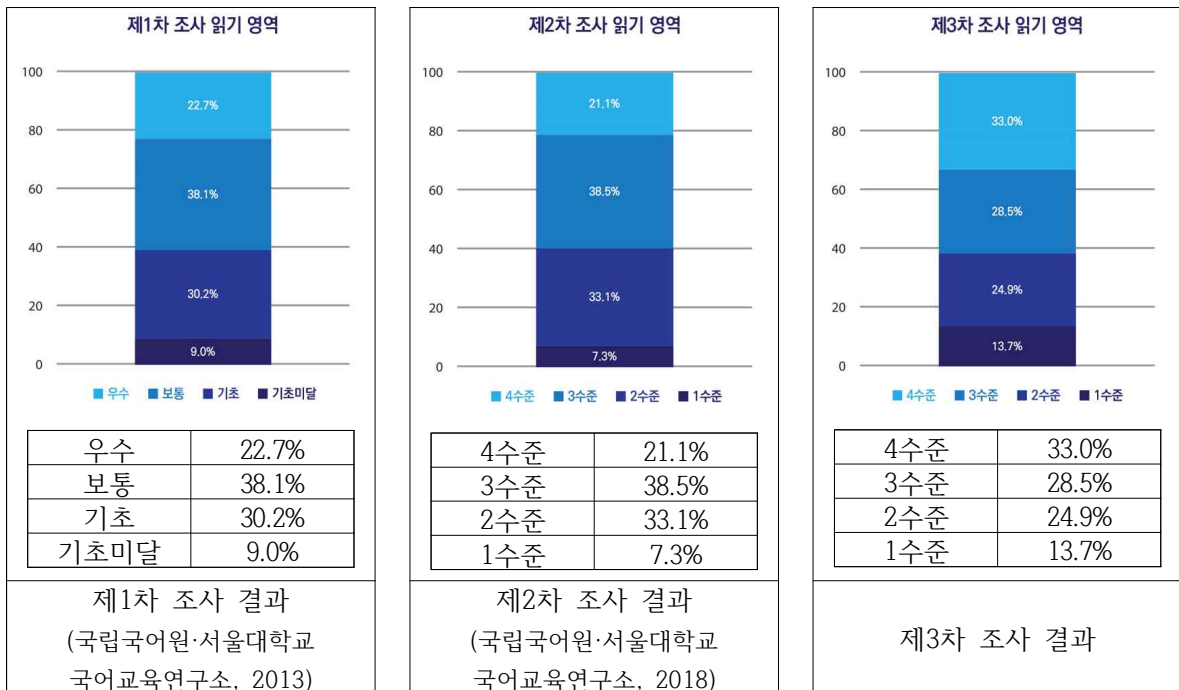
<표 V-2> 듣기 영역 제1차, 제2차, 제3차 수준별 비율



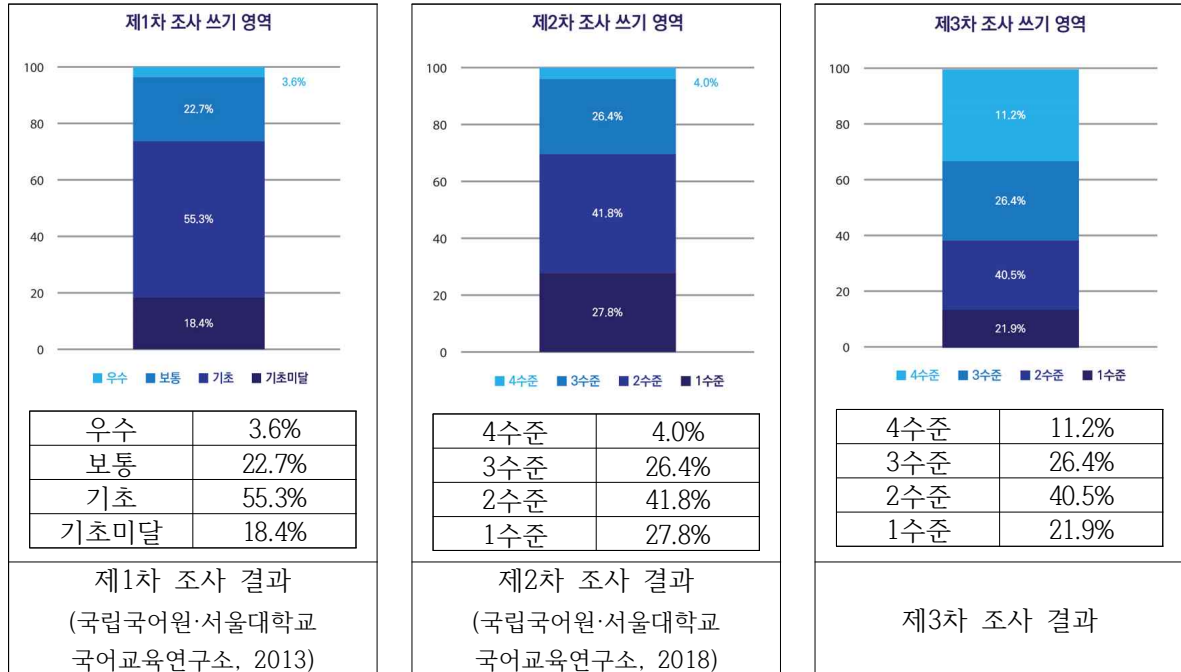
<표 V-3> 말하기 영역 제1차, 제2차, 제3차 수준별 비율



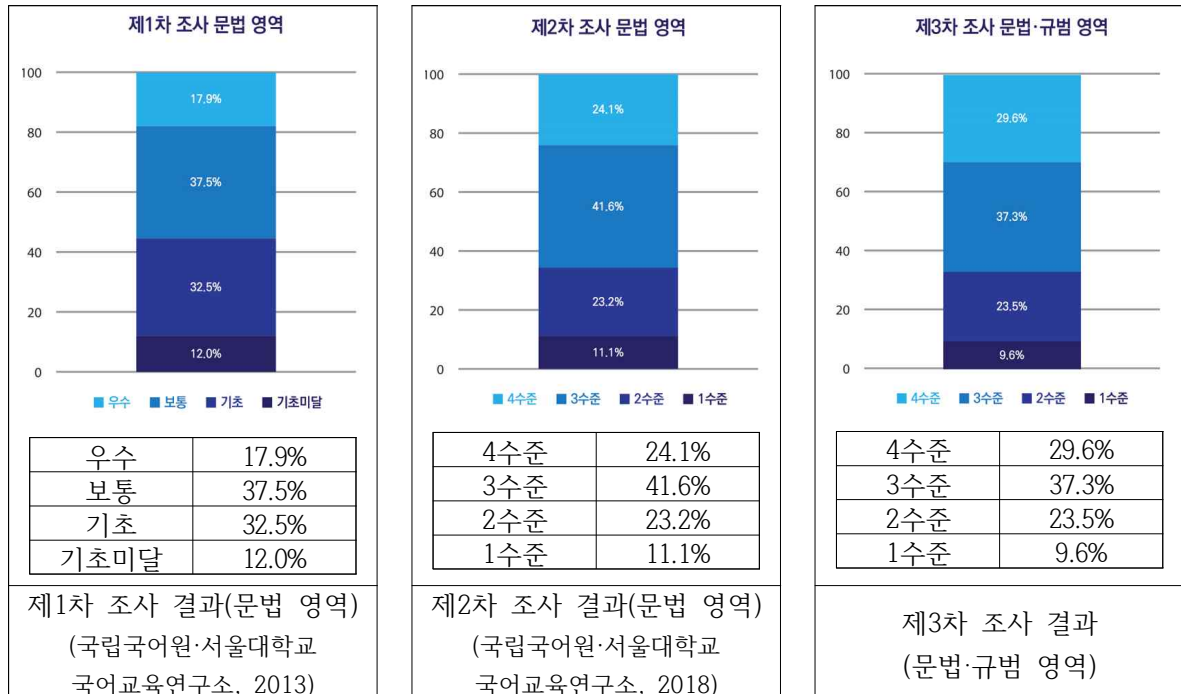
<표 V-4> 읽기 영역 제1차, 제2차, 제3차 수준별 비율



<표 V-5> 쓰기 영역 제1차, 제2차, 제3차 수준별 비율

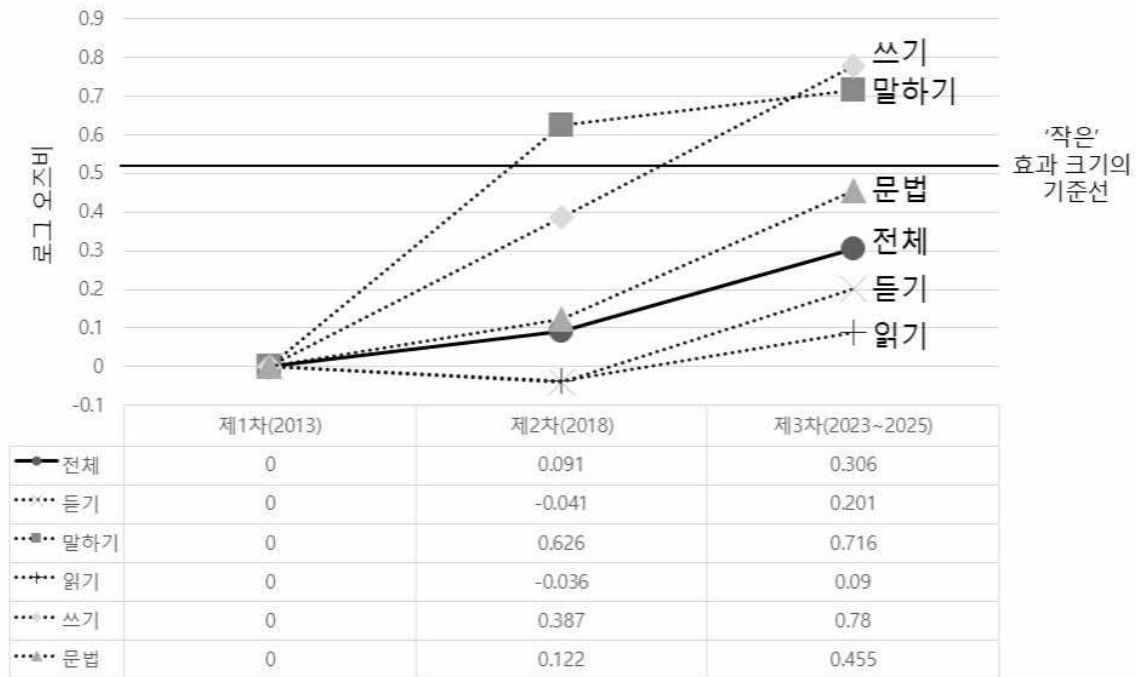


<표 V-6> 문법 및 문법·규범 영역 제1차, 제2차, 제3차 수준별 비율



본 절에서는 제1차~제3차 국민의 국어능력 실태 조사 결과를 비교하여 지난 10여 년간 국민의 국어 능력에 변화가 있었는지, 있었다면 어떻게, 얼마나 변화하였는지를 파악해 보고자 하였다. 이를 위해 제

1차와 제2차 조사에서의 가교 문항 10개와 제2차와 제3차 조사에서의 가교 문항 15개에서의 정답률을 비교하고, 정답률의 증감 정도를 효과 크기로 계산하여 파악해 보았다. 이러한 분석 결과를 아래의 그림에서 종합적으로 파악할 수 있다.



[그림 V-1] 제1차~제3차 조사 시기 국민의 국어능력 추이 요약

위 그림을 살펴보면, 제1차 조사 시기로부터 제3차 조사 시기에 이르는 약 10여 년간, 전체적인 국민의 국어능력은 다소간 상승한 것처럼 보이나 이러한 변화가 작은 효과 크기 이상에 해당하는 의미 있는 크기의 변화라고 보기는 어려운 것을 알 수 있다. 다만 세부 영역별로 살펴보았을 때에는, 쓰기와 말하기 영역에서 지난 10여 년간 의미 있는 상승이 있었으리라고 짐작이 된다. 이상과 같은 종합적인 결론에 이르기까지 그 토대가 된, 세부적인 분석 방법과 결과에 관해서는 5.2절과 5.3절에서 설명할 것이다.

5.2. 제1차 조사 시기 대비 제2차 조사 시기 국민의 국어능력 추이

본 절에서는 2013년에 시행된 제1차 조사 대비 2018년에 시행된 제2차 조사에서 국민의 국어능력이 달라졌는지, 달라졌다면 어떻게 달라졌는지를 분석한 결과를 다룬다. 이를 위해 제1차와 제2차 조사에서 공통적으로 사용한 가교 문항 10개에서 각각 효과 크기를 메타분석(meta analysis)의 방법을 준용하여 계산하였다. 효과 크기의 계산 방법은 구체적으로 다음과 같다.

우선, 각 문항별로 제1차 조사에서의 오즈(odds) 대비 제2차 조사에서의 오즈 비율에 해당하는 오즈비(odds ratio; OR)를 계산하였다. 오즈는 어떤 사건이 일어나지 않을 확률(즉, 1-p) 대비 사건이 일어날 확률(p)을 나타내는 수학적 개념이다. 현재의 분석에서는 문항의 정답을 맞히는 것이 사건이 일어난 경우라고 간주하여 오즈를 정답을 맞히지 못할 확률(=오답자 수/응시자 수=1-정답자 수/응시자 수) 대

비 정답을 맞힐 확률(=정답자 수/응시자 수)로 계산한다. 공통 분모인 응시자 수는 소거되므로 결국, 현재의 분석에서 오즈는 각 문항에서의 정답자 수(1차 조사 정답자 수=a; 2차 조사 정답자 수=c)를 오답자 수(1차 조사 오답자 수=b; 2차 조사 오답자 수=d)로 나누어 계산된다. 각 문항마다 제1차 조사에서의 오즈(=a/b)와 제2차 조사에서의 오즈(=c/d)를 계산한 후 제2차 조사에서의 오즈를 제1차 조사에서의 오즈로 나누면 오즈비(=bc/ad)를 산출할 수 있다. 이러한 오즈비가 0~1이라면 이는 제1차 대비 제2차 조사에서의 정답률이 하락하였음을 의미하고, 이에 따라 해당 문항과 관련한 국민의 국어능력이 하락하였을 수 있음을 시사한다. 오즈비가 1보다 크다면 이는 제1차 대비 제2차 조사에서의 정답률이 상승하였음을 의미하고, 이에 따라 해당 문항과 관련한 국민의 국어능력이 상승하였을 수 있음을 시사한다. 오즈비에 로그를 취한 로그 오즈비(lnOR)를 기준으로 삼을 경우, 로그 오즈비가 음수라면 국민의 국어능력이 하락하였을 수 있음을, 로그 오즈비가 양수라면 국민의 국어능력이 상승하였을 수 있음을 시사한다.

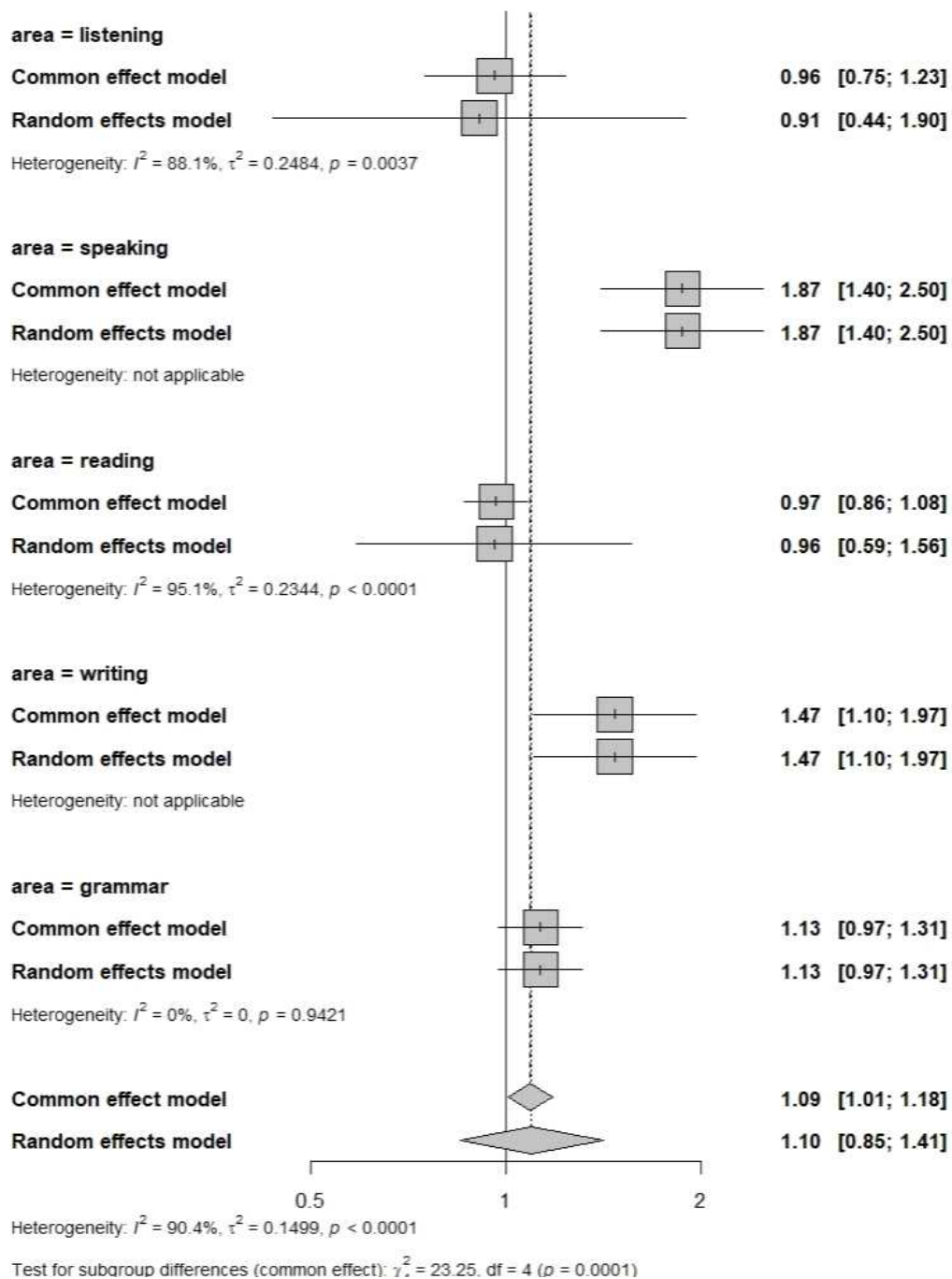
단순히 오즈가 1보다 작거나 크지(혹은, 로그 오즈비가 0보다 작거나 크지)만으로도 국민의 국어능력이 하락 혹은 상승하였다고 선불리 단정하기는 어렵다. 최소한, 오즈의 95% 신뢰구간이 1을 포함하지 않을 만큼(혹은, 로그 오즈비의 95% 신뢰구간이 0을 포함하지 않을 만큼) 오즈의 값이(혹은, 로그 오즈의 절댓값이) 충분히 커야 국어능력의 하락 혹은 상승을 논할 수 있다. 나아가, 효과 크기의 해석 기준을 제안한 Cohen(1988)에 따라 $0.51 \leq |\log OR| < 1.25$ 일 때 작은 효과 크기, $1.25 \leq |\log OR| < 1.90$ 일 때 중간 효과 크기, $1.90 \leq |\log OR|$ 일 때 큰 효과 크기로 해석하는 것이 더욱더 바람직하다. 따라서 본 연구에서는 로그 오즈비의 절댓값이 0.51에 미치지 못한다면 효과 크기가 거의 없는 것으로 해석하여 사실상 변화가 없는 것으로 해석하고자 하였다.

아래의 표에, 듣기, 말하기, 읽기, 쓰기, 문법 영역별로 각 가교 문항의 정답률이 제1차와 제2차 조사에서 어떻게 나타났는지, 이에 따라 오즈비(95% 신뢰구간 포함)와 로그 오즈비가 얼마였는지, 그 효과 크기를 어떻게 해석할 수 있는지, 전체 효과 크기 계산을 위한 가중치는 얼마이며, 전체 효과 크기는 얼마인지를 정리하여 제시하였다. 결과를 살펴보면, 듣기, 읽기, 쓰기, 문법 영역의 효과 크기는 ‘거의 없음’($|\log OR| < 0.51$)으로 나타났는데, 이는 제1차 대비 제2차 조사에서 해당 영역과 관련한 국민의 국어능력이 거의 변화가 없었을 것임을 시사한다. 이와 달리 말하기 영역의 효과 크기는 ‘작은’ 효과 크기($0.51 \leq |\log OR| < 1.25$)로 나타났는데, 이는 제1차 대비 제2차 조사에서 해당 영역과 관련한 국민의 국어능력이 소폭 상승했을 수 있음을 시사한다. 전 영역을 포괄한 전체 효과 크기는 $OR=1.095$ 이고, 그 95% 신뢰구간은 0.849~1.413으로 1을 포함하고 있어, 전반적으로 제1차 조사 대비 제2차 조사에서의 정답률은 변화가 없었던 것으로 보인다. 이는 제1차 조사 시기인 2013년과 비교해 제2차 조사 시기인 2018년에 국민의 국어능력이 거의 변화가 없었으리라는 점을 시사한다.

<표 V-7> 제1차~제2차 시기 국민의 국어능력 추이 분석 결과

영역	문항 번호 (제2차 본조사 기준)	응시자 수		정답률		오즈비 (OR)	95% 신뢰구간	로그 오즈비 (logOR)	효과 크기 해석	가중치 (랜덤 모형)
		제2차	제1차	제2차	제1차					
듣기	L1	3,000	200	76.5%	83.9%	0.620	[0.421, 0.913]	-0.478	거의 없음	8.9%
	L2	3,000	200	78.9%	73.9%	1.314	[0.947, 1.824]	0.273	거의 없음	9.5%
	전체(랜덤 모형)					0.960	[0.748, 1.233]	-0.041	거의 없음	
말하기	S3	2,939	200	59.5%	43.9%	1.871	[1.401, 2.497]	0.626	작은	9.8%

읽기	R1	5,500	400	77.9%	64.0%	1.984	[1.602, 2.457]	0.685	작은	10.4%
	R2	5,500	400	49.3%	58.6%	0.690	[0.562, 0.848]	-0.371	거의 없음	10.5%
	R3	5,500	400	58.3%	66.1%	0.721	[0.582, 0.892]	-0.327	거의 없음	10.4%
	R4	5,500	200	61.3%	65.2%	0.853	[0.635, 1.147]	-0.159	거의 없음	9.8%
	전체(랜덤 모형)					0.965	[0.863, 1.081]	-0.036	거의 없음	
쓰기	W2	3,000	200	52.1%	42.6%	1.472	[1.102, 1.965]	0.387	거의 없음	9.8%
문법	G1	5,500	400	66.4%	63.8%	1.124	[0.910, 1.388]	0.117	거의 없음	10.4%
	G2	5,500	400	65.2%	62.2%	1.136	[0.921, 1.401]	0.128	거의 없음	10.4%
	전체(랜덤 모형)					1.130	[0.974, 1.311]	0.122	거의 없음	
전체 효과 크기(랜덤 모형)						1.095	[0.849, 1.413]	0.091	거의 없음	



[그림 V-2] 제1차~제2차 시기 국민의 국어능력 변화의 효과 크기에 대한 숲 그림

5.3. 제2차 조사 시기 대비 제3차 조사 시기 국민의 국어능력 추이

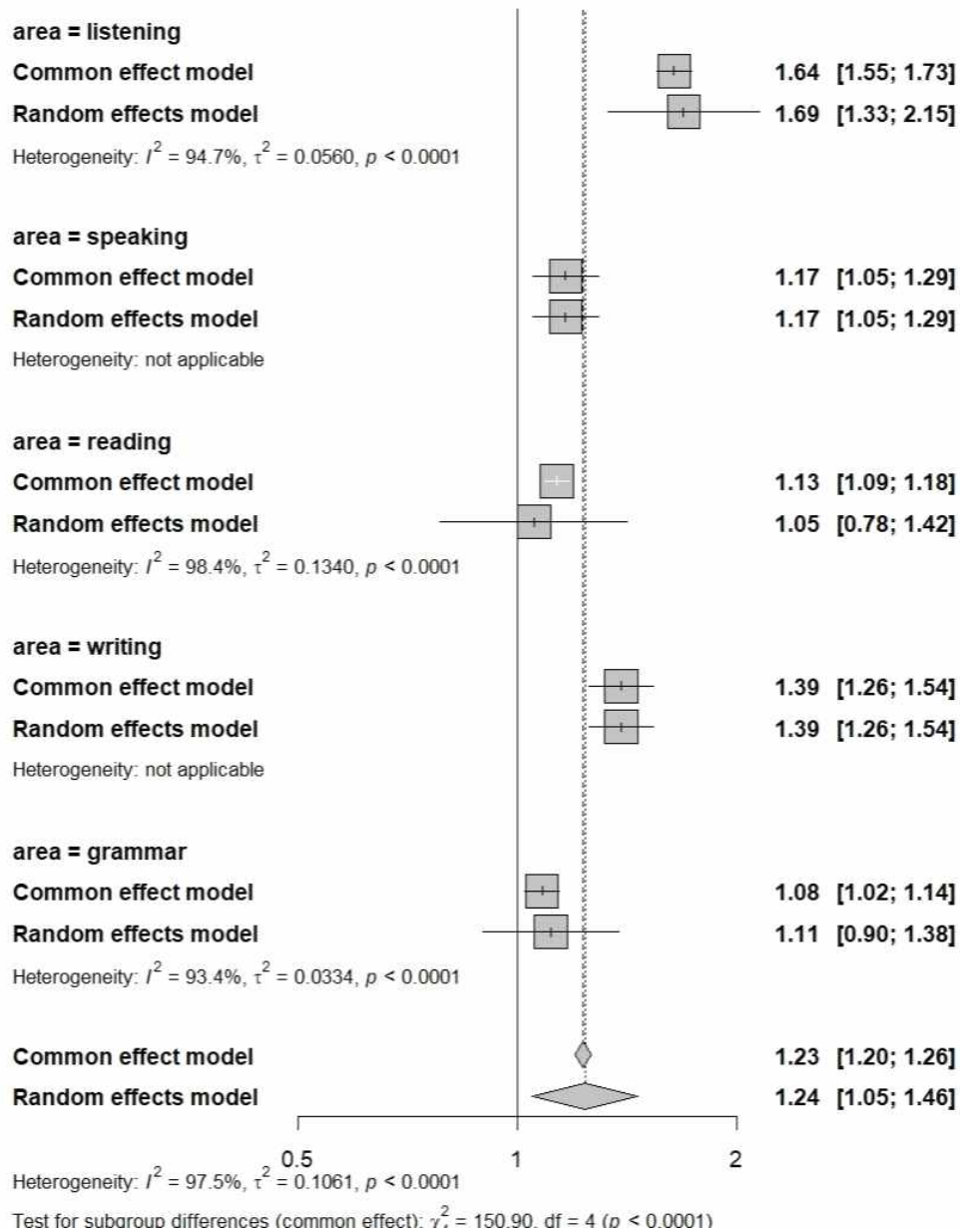
본 절에서는 2018년에 시행된 제2차 조사 대비 2023년~2025년에 시행된 제3차 조사에서 국민의 국어능력이 달라졌는지, 달라졌다면 어떻게 달라졌는지를 분석한 결과를 다룬다. 이를 위해 제2차와 제3

차 조사에서 공통적으로 사용한 가교 문항 15개에서 각각 효과 크기를 메타분석(meta analysis)의 방법을 준용하여 계산하였다. 효과 크기의 계산 방법은 앞의 항에서 밝힌 바와 같다.

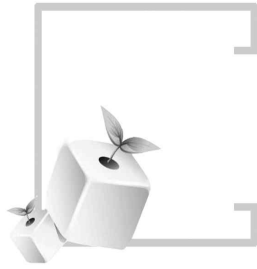
아래의 표에, 듣기, 말하기, 읽기, 쓰기, 문법 영역별로 각 가교 문항의 정답률이 제2차와 제3차 조사에서 어떻게 나타났는지, 이에 따라 오즈비(95% 신뢰구간 포함)와 로그 오즈비가 얼마였는지, 그 효과 크기를 어떻게 해석할 수 있는지, 전체 효과 크기 계산을 위한 가중치는 얼마이며, 전체 효과 크기는 얼마인지를 정리하여 제시하였다. 결과를 살펴보면, 듣기, 말하기, 읽기, 쓰기, 문법 영역의 효과 크기는 모두 ‘거의 없음’($|\log OR| < 0.51$)으로 나타났는데, 이는 제2차 대비 제3차 조사에서 각 영역과 관련한 국민의 국어능력이 거의 변화가 없었을 것임을 시사한다. 전 영역을 포괄한 전체 효과 크기는 $OR=1.240$ 이고, 그 95% 신뢰구간은 1.049~1.465이었다. 비록 OR 이 1보다 크다는 점이 통계적으로 지지될 수 있다고는 하나, 로그 오즈비의 절댓값이 0.215로 작은 효과 크기의 기준인 0.51에 미치지 못하므로, 전반적으로 제2차 조사 대비 제3차 조사에서의 정답률은 변화가 없었던 것으로 보인다. 이는 제2차 조사 시기인 2018년과 비교해 제3차 조사 시기인 2023~2025년에 국민의 국어능력이 거의 변화가 없었으리라는 점을 시사한다.

<표 V-8> 제2차~제3차 시기 국민의 국어능력 추이 분석 결과

영역	문항 번호 (제3차 본조사 기준)	응시자 수		정답률		오즈비 (OR)	95% 신뢰구간	로그 오즈비 (logOR)	효과 크기 해석	가중치 (랜덤 모형)
		제3차	제2차	제3차	제2차					
듣기	L1	3,033	3,000	55.0%	49.7%	1.237	[1.118, 1.368]	0.213	거의 없음	6.7%
	L2	3,033	3,000	82.5%	68.5%	2.167	[1.919, 2.446]	0.773	작은	6.6%
	L10	3,033	3,000	81.5%	69.7%	1.916	[1.699, 2.160]	0.650	작은	6.6%
	L12	3,033	3,000	64.3%	52.8%	1.610	[1.452, 1.785]	0.476	거의 없음	6.7%
	전체(랜덤 모형)					1.642	[1.554, 1.735]	0.496	거의 없음	
말하기	S1	2,983	2,939	64.0%	60.4%	1.166	[1.049, 1.295]	0.154	거의 없음	6.7%
읽기	R1	5,000	5,500	75.0%	77.9%	0.851	[0.777, 0.931]	-0.161	거의 없음	6.7%
	R2	5,000	5,500	63.0%	49.3%	1.750	[1.619, 1.893]	0.560	작은	6.7%
	R4	5,000	5,500	70.0%	61.3%	1.473	[1.358, 1.597]	0.387	거의 없음	6.7%
	R5	5,000	5,500	72.0%	80.0%	0.643	[0.577, 0.717]	-0.442	거의 없음	6.6%
	R6	5,000	5,500	50.0%	48.5%	1.062	[0.970, 1.163]	0.060	거의 없음	6.7%
	R7	5,000	5,500	75.0%	76.7%	0.911	[0.820, 1.013]	-0.093	거의 없음	6.6%
	전체(랜덤 모형)					1.134	[1.093, 1.177]	0.126	거의 없음	
쓰기	W1	3,000	3,000	54.0%	45.8%	1.389	[1.255, 1.538]	0.329	거의 없음	6.7%
문법	G1	5,000	5,500	85.0%	83.8%	1.096	[0.986, 1.218]	0.092	거의 없음	6.7%
	G2	5,000	5,500	86.0%	81.9%	1.357	[1.221, 1.508]	0.305	거의 없음	6.7%
	G3	5,000	5,500	69.0%	70.5%	0.931	[0.857, 1.012]	-0.071	거의 없음	6.7%
	전체(랜덤 모형)					1.082	[1.023, 1.143]	0.079	거의 없음	
전체 효과 크기(랜덤 모형)						1.240	[1.049, 1.465]	0.215	거의 없음	



[그림 V-3] 제2차~제3차 시기 국민의 국어능력 변화의 효과 크기에 대한 숲 그림



제 6 장

제언



6.1. 실태 조사 결과의 시사점

1) 듣기, 말하기, 듣기·말하기 영역

우선 듣기 실태 조사 결과의 시사점을 논의하자면 다음과 같다. 국민의 국어능력 실태 조사에서 듣기 영역의 결과는 다양한 인구통계학적 요인과 매체 활용 행태가 듣기 능력에 복합적으로 영향을 미친다는 점을 보여 준다. 특히 연령대, 지역 규모, 직업군, 학력이 통계적으로 유의미한 예측 변인으로 작용하였으며, 매체 활용 양상 또한 듣기 능력과 관련이 있었다. 이러한 결과는 언어 환경의 변화 속에서 듣기 능력이 단순한 언어 이해 능력을 넘어 디지털 매체와 결합된 복합적 언어 처리 능력으로 작용하고 있음을 보여 주는 것이라 할 수 있다.

듣기 영역의 인구통계학적 변인 분석 결과에 따른 시사점을 살펴보자면 다음과 같다. 첫째, 연령이 높아질수록 듣기 능력 점수가 하락하는 경향은 세대 간 언어 환경의 차이와 청각적 정보 처리 능력의 감소가 복합적으로 작용한 결과로 볼 수 있다. 이는 연령의 증가에 따른 생리적 요인뿐 아니라 정보 접근성과 미디어 환경의 차이에서 비롯된 인지적 요인도 함께 고려되어야 함을 의미한다. 특히 20~30대는 영상과 음성 등이 결합된 복합매체를 통한 정보 환경에 익숙한 반면, 60대 이상은 단일 매체를 통한 언어 소통에 익숙하며 음성보다는 문자언어를 통한 정보의 습득과 이해에 더 익숙한 경향이 있다. 이러한 차이는 세대별 듣기 교육과 언어 경험의 격차로 이어질 수 있으므로, 향후 성인 대상 국어능력 향상 프로그램의 내용을 구성할 때에는 세대별 언어 노출 환경에 따른 차이를 고려할 필요가 있다.

둘째, 읍면 지역이 대도시나 중소 도시에 비해 듣기 능력 수준이 낮은 결과는, 지역 간 언어 문화 접근성의 불균형을 반영하는 결과로 볼 수 있다. 읍면 지역에서는 방송, 강연 등 다양한 담화 유형에 따른 자극의 빈도가 낮고 회의, 토론 등 공적 담화에 참여할 기회가 적은 경우가 많다. 이는 단순히 미디어 접근성의 문제를 넘어서 지역 사회 내에서 언어문화 생태계가 상대적으로 협소할 수 있다는 점을 시사하는 것이다.

셋째, 학력이 높을수록 듣기 능력이 높게 나타난 결과는 듣기 능력이 단순한 언어적 기능이 아니라 인지적 통합 능력과 학습 경험의 축적과 밀접하게 연관되어 있음을 보여 주는 것이다. 고졸 미만 집단의 듣기 능력이 낮게 나타난 것은 언어적 이해 과정에서 추론하기, 비판적 이해하기 등과 같은 높은 수준의 인지 능력이 충분하게 발달되지 못한 결과로 해석해 볼 수 있다. 이러한 점은 성인 문해력 정책에서 듣기 능력을 읽기와 함께 이해 중심 문해력(literacy)의 한 부분으로 중요하게 다룰 필요가 있음을 시사한다.

다음으로, 듣기 영역 심층 변인 분석 결과에 따른 시사점은 다음과 같다. 듣기 영역 심층 변인 분석 결과에 따르면 직접 만나서 하는 대화, 누리소통망(SNS)이나 메신저로 하는 대화 빈도가 높은 집단의 듣기 능력 평균 점수가 다른 집단에 비해 높은 것으로 나타났으며, 토의(회의)에서 듣고 말하기를 거의 하지 않는다고 응답한 집단의 듣기 능력 평균 점수가 다른 집단에 비해 더 낮은 것으로 나타났다. 이는 다른 사람과 소통하는 빈도와 상호작용에 참여하는 기회여부가 듣기 능력에 영향을 주는 것으로 분석해 볼 수 있다. 매체 사용 빈도에 따른 시사점에서도 확인할 수 있는 바와 같이, 듣기 능력은 다른 사람과의 소통에 적극적으로 참여하고 상호작용하는 과정에서 자신에게 필요한 정보를 능동적으로 선택하고 반응하는 경험을 통해 향상되는 것으로 파악할 수 있다. 따라서 듣기 능력 향상을 위해서는 대면 의사소통을 포함하여 매체를 통한 온라인 소통 등 다양한 의사소통에 참여할 수 있는 기회와 경험을 제공하는 것이 필요할 것이다.

이어서 말하기, 듣기·말하기 영역의 결과를 바탕으로 한 시사점을 제시하자면 다음과 같다. 국민의 국

어능력 실태 조사에서 말하기 영역의 결과는 개인의 사회적 배경, 언어 활동 참여 정도, 구어 기반 상호작용 환경이 말하기 능력의 형성과 유지에 복합적으로 작용한다는 점을 보여 준다. 특히 연령대, 지역 규모, 직업군, 학력이 모두 통계적으로 유의미한 예측 변인으로 확인되었으며, 구어적 상호작용의 빈도, 다양한 담화 유형 노출 경험, 직무 관련 언어 요구 수준 등이 말하기 능력의 차이를 설명하는 핵심적인 요인으로 나타났다. 이러한 결과는 말하기 능력이 단순한 발화 기술이나 선천적인 소통 능력이 아니라 사회문화적 맥락 속에서의 소통을 통해 획득되는 사회적 언어 자본과 밀접하게 연관된 능력임을 시사한다고 볼 수 있다.

한편 듣기와 말하기 능력은 상호 보완적이며 동시에 작동하는 상호작용 기반의 의사소통 능력이라는 점을 고려하여 통합적으로 접근할 필요가 있다. 이번 실태 조사 결과에서도 듣기 영역에서 확인된 요인들이 말하기 영역에서도 동일하게 나타났으며, 특히 언어 사용 환경, 사회적 상호작용의 기회, 직업 상황에서의 언어 요구 수준 등이 듣기·말하기 능력에 복합적으로 영향을 미치는 것으로 분석되었다. 이는 듣기와 말하기를 독립적인 기능이라기보다는 상호협력적인 언어 수행 능력으로서 통합적으로 접근해야 함을 시사하는 것이라 할 수 있다. 다만 듣기·말하기 영역에 대한 분석 내용은 말하기 영역의 분석과 유사하였으므로 이 장에서 통합하여 제시하였다.

말하기, 듣기·말하기 영역의 인구통계학적 변인 분석 결과에 따른 시사점을 살펴보자면 다음과 같다. 첫째, 연령이 증가할수록 말하기 능력 수준이 하락하는 경향은 구어 소통 상황에서 반응 속도가 감소하는 것과 더불어 세대별 언어 활동의 참여 양상 차이에서 기인하는 것으로 볼 수 있다. 20~30대의 경우 직장 내에서 보고, 회의 등의 다양한 말하기 상황에 참여하는 기회가 많은 반면, 60대 이상의 집단은 구어적 상호작용의 기회가 일상적이고 제한적인 관계로 축소되는 경향이 있다. 또한 고령층은 문자 기반의 정보 처리 방식에 더욱 익숙하여 즉각적인 발화나 논리적인 내용 구성, 발화 상황에서의 조정과 협력 등을 요구하는 구어 의사소통에 어려움을 느낄 가능성이 상대적으로 크다고 볼 수 있다.

둘째, 읍면 지역의 말하기 능력 점수가 대도시나 중소 도시에 비해 낮은 것은 공적인 구어 소통 상황에 참여할 수 있는 기회가 지역 간에 차이가 있음을 의미하는 것으로 해석할 수 있다. 읍면 지역은 도시 지역에 비해 다양한 직종과 교류할 수 있는 기회나 회의·토론·강연 등의 담화 상황에 참여할 기회가 제한적일 가능성이 크다. 이러한 차이는 단순히 말할 기회의 유무를 넘어, 지역에 따라 다양한 담화 유형을 경험할 수 있는 언어 문화적 자원의 양과 질에 차이가 있다는 것을 보여 주는 것이라 할 수 있다.

셋째, 학력이 높을수록 말하기 능력이 높게 나타난 것은 말하기 능력이 단순히 유창하게 말하기와 같은 기술의 차원이 아니라 고차적인 인지 능력과 상호작용 능력의 통합으로 이루어지는 것임을 의미하는 것이라 할 수 있다. 대재 이상의 학력 집단은 학교나 직무 상황에서 발표, 보고, 토의, 토론 등의 다양한 담화를 반복적으로 수행하고, 이러한 수행 경험이 누적됨으로써 말하기 능력이 향상될 수 있다. 이에 비해 고졸 미만의 집단에서 말하기 능력 점수가 다소 낮게 나타난 것은 구어 수행에 필요한 인지적 전략이나 표현 능력을 갖추기 위한 기회가 충분하지 않았음을 의미하는 결과로 해석해 볼 수 있다.

넷째, 직업군에 따른 차이는 말하기 능력의 차이와 밀접하게 관련되어 있었는데, 특히 정신노동 직군에서 말하기 능력이 상대적으로 높은 것은 직무 특성상 설명, 보고, 회의, 협상 등과 같이 업무 상황에서 말하기를 기반으로 한 소통이 많은 비중을 차지하기 때문이다. 반면 육체노동 직군에서는 업무 상황에서 상대적으로 복합적인 정보 전달을 중심으로 한 구어 소통에 참여할 기회가 적을 가능성이 있으며, 이러한 직무 수행 상황에 따른 언어 요구의 차이가 언어 능력의 차이를 나타나게 하는 요인으로 작용할 가능성이 있음을 알 수 있다.

말하기, 듣기·말하기 영역의 심층 변인 분석 결과에 따른 시사점을 살펴보자면 다음과 같다. 말하기 영역의 심층 변인 분석 결과 대면 대화의 빈도, 음성 통화, 담화 참여 수준 등 상호작용 기반의 변인이 말하기 능력에 영향을 미치는 것으로 나타났다. 이는 말하기 능력이 다양한 담화 상황에 참여하고 상호

작용함으로써 강화될 수 있는 능력임을 시사하는 것으로 볼 수 있다.

첫째, 직접 대면 대화 빈도가 높은 집단일수록 말하기 능력 점수가 높게 나타난 것은 구어 상호작용을 통한 협력적 의미 구성, 비언어적 표현 해석, 즉각적인 반응 제시 등과 같이 역동적인 담화 상황에 적극적으로 참여하는 것이 말하기 능력에 큰 영향을 미친다는 것을 보여 주는 결과라 할 수 있다. 특히 말하기 능력 점수가 가장 낮은 집단과 가장 높은 집단 간의 능력 차이가 유의미하게 나타났다는 점은 말하기 능력이 지속적인 사용을 통해 강화되는 능력임을 시사하는 것이다.

둘째, 음성 통화 빈도 또한 말하기 능력과 관련이 있는 것으로 나타났다. 음성 통화는 비대면 상황에서 상대의 의도를 정확하게 파악하고 즉각적으로 반응하며, 적절하게 자신의 생각을 전달하는 것이 중요하다. 이처럼 직접적인 구어 의사소통에의 참여가 결과적으로 말하기 능력의 강화에 영향을 미치는 것이다.

셋째, 발표·강연을 듣는 경험의 부족이 말하기 능력의 낮은 수준과 연관되는 점이나 토의·회의 참여 경험의 부재가 말하기 능력 발달을 저해하는 요인으로 나타났다는 점도 주목할 만한 결과이다. 결국 다양한 담화 상황에 노출되고, 이러한 담화에 참여하는 언어 수행 경험이 말하기 능력에 직접적인 영향을 준다는 것을 알 수 있다.

이상의 결과는 말하기 능력이 단순한 언어 기능이 아니라 공적 담화 참여 경험, 상호작용 참여 정도, 다양한 담화 상황의 노출 정도 등에 의해 강화되는 사회적 능력이라는 것을 보여 주는 것이라 할 수 있다.

2) 읽기 영역

우선 읽기 영역의 전반적 경향과 읽기 특징과 관련해 시사점을 제언하자면 다음과 같다. 읽기 영역의 수준별 분포를 살펴보면 전체의 약 33%가 4수준에 속하나, 여전히 약 38% 이상이 1수준 및 2수준에 머무르고 있다. 읽기 영역의 주목할 만한 특성을 몇 가지 짚어 보면 다음과 같다.

첫째, 연령대별 차이와 관련해 세대 간 격차가 뚜렷하다는 점이다. 연령별 점수를 보면, 연령이 높을수록 읽기 능력이 감소되는 경향이 뚜렷하게 보인다. 20대의 읽기 능력 점수(51.19점)가 가장 높고, 60대(47.44점)가 가장 낮다. 이는 각 수준별 인원 분포를 비교한 결과와도 상통하는데, 20대에서 60대로 연령대가 높아질수록 4수준 인원 비율은 크게 줄어드는 반면(42.8%→22.8%), 1수준 인원 비율은 조금씩 늘어나는 것(12.3%→16.1%)을 통해 알 수 있다.

이는 어쩌면 당연한 현상으로, 학교 교육기(period of school education)를 벗어난 뒤에는 시간의 경과에 따라 문식 환경에서부터 멀어지고, 인지력도 저하되면서 읽기 점수가 자연스럽게 점진적으로 낮아짐을 보여 준다. 그러나 동시에, 학교에서 사회로 진출한 후에 성인의 문해력이 개선되거나 유지되지 않는다는 것은 사회와 일상 삶의 문해 환경이나 문식 자원이 좋지 않거나 적어도 나아지도록 뒷받침되지 않는다는 것을 방증한다.

둘째, 읽기의 대상이 되는 매체와 그 활용 양상이 빠르게 변화하고 있다는 점이다. 매체 유형별 사용 빈도를 보면, 종이책(인쇄 자료)은 확연히 줄고, 빠른 속도로 인터넷 웹 기반 텍스트 또는 동영상 활용 비율이 높아지고 있다. 예컨대, 온라인 동영상 플랫폼과 인터넷 동영상 서비스(63.7%), 인터넷의 포털사이트 검색(80.3%)과 커뮤니티(44.5%) 등을 통해 매일 1번 이상 읽기 활동을 경험하는 것으로 나타났다.

또한 디지털 기기 사용 시간(매체 이용 시간)에 따른 읽기 능력을 살펴보면, 디지털 기기를 ‘하루 5시간 이상~하루 종일’ 사용하는 집단이 메신저 외에 거의 사용하지 않는 집단보다 읽기 점수가 훨씬 높은 것으로 나타났다. 특히 읽기 빈도가 ‘거의 매일’인 집단의 능력 평균 점수가 가장 높은 점, 아울러 전통

매체(올드 미디어)인 영상·시청각 매체 이용이 많은 집단일수록 읽기 점수가 낮게 나타날 뿐 아니라, 신매체(뉴 미디어)인 디지털 매체 환경에 익숙한 집단의 읽기 점수가 상대적으로 높다는 점 등을 종합할 때, 이러한 현상은 매체 사용의 양상이 연령대와 밀접하게 관련되어 있음을 추정케 한다. 즉, 읽기 능력이 높은 젊은 청장년층들(20~30대)에서는 디지털 기기 사용이 활성화되어 있고, 읽기 능력이 낮은 것으로 나왔던 고령층(60대)에서는 디지털 기기 사용 비율이 상대적으로 낮다는 점을 떠올리면 충분히 추론할 수 있다. 오히려 이는 디지털 기기를 많이 사용한다고 해서 반드시 읽기 능력이 떨어지는 것은 아니라는 점, 즉 디지털 기기의 사용이 읽기 능력과 반드시 부적 관계는 아니라는 점을 보여 준다.

이를 통해 디지털 읽기 생태 환경(digital reading ecology)이 가속화되고 있고, 그 결과 ‘읽기 양식의 다층화’가 진행 중임을 알 수 있다. 이는 또한 디지털 매체 중심의 읽기 환경에 빠르게 적응한 젊은 세대와 전통적 문해 환경에 익숙한 중장년층 간의 ‘디지털 격차형 문해력(digital literacy divide)’ 현상을 시사한다.

셋째, 독서의 수준과 관련해 독서의 질이 하락하고 있다는 점이다. 먼저, 긴 글을 읽지 않는 현상이 두드러진다. 1회 독서 시 읽기 분량을 조사했더니, 전혀 읽지 않는 경우를 포함하여, 읽더라도 대부분 카드 뉴스 정도의 짧은 요약 글에서 블로그 길이의 교과서 1쪽~4쪽을 읽는다는 답변이 전체의 78.9%이다. 즉 교과서 5쪽 이상의 대체로 긴 글을 읽는 경우는 21.1% 정도이다. 인터넷상의 글이라 하더라도, ‘컴퓨터 첫 화면(1 node)에 나타난 글’을 기준으로 할 때, 스크롤바를 한참 내려야 하는 긴 분량의 글은 잘 읽지 않는다는 것이다. 이는 소셜 미디어 플랫폼(예: 엑스, 틱톡, 인스타그램)의 성장으로 인해 사람들이 긴 글보다는 280자 미만의 트윗, 짧은 비디오 클립 등에 익숙해지는 현상이 읽기 능력에도 영향을 미치고 있다(Macwan et al., 2024; Carr, 2010)는 연구 결과에 의해서도 뒷받침된다.

다음으로는 독서의 횟수도 적다. 읽기를 ‘거의 매일’ 하는 경우(8.7%)나 일주일 2~3번 읽는 경우(12%)는 20% 남짓이다. 동시에 매체 선호 유형 조사에서는 종이책이나 텔레비전(뉴스, 드라마)에 비해 인터넷 기반 디지털 자료나 동영상(OTT)을 본다는 응답자의 비중이 높았다. 이는 문자 중심 자료를 읽는 애독자가 갈수록 줄어든다는 최근의 경향을 설명해 준다.

이렇듯 읽기 빈도나 읽기 분량 등은 읽기 능력과도 상관성이 높은데, 자주 읽을수록 또 긴 글을 읽을수록 읽기 능력 점수가 높은 것으로 나타났다. 반면 거의 읽지 않거나, 읽더라도 카드 뉴스처럼 글자 수가 적은 짧은 글만 읽는 경향이 계속되면 이는 읽기 능력을 떨어뜨릴 수 있고, 이렇게 낮아진 읽기 능력으로 인해, 갈수록 어렵고 긴 글은 잘 읽지 못하게 되는 악순환을 낳는다.

이처럼 긴 글을 읽는 독서 경험을 통해 문자를 해독하고 머릿속에서 의미를 개념화하거나 이미지로 표상하는 과정에서 고도의 정보처리 능력이 요구되는데, 매체의 변화로 이러한 유의미한 경험의 양이 줄고 있다면, 이러한 깊은 읽기(deep reading)와 지속적 몰입(sustained engagement) 경험의 감소가 전반적으로 문해력 저하의 요인으로 작용하고 있지 않은지 의심해 볼 수 있다.

다음으로, 사회·교육적 변인과 관련한 읽기 양상을 살펴보자면 다음과 같다. 우선, 학력에 따른 읽기 능력이다. 학력이 높을수록 읽기 점수가 높다. 예컨대, 대학 재학 이상(51.04점)이 고졸(48.15점)과 고졸 미만(45.33점)보다 유의하게 높다. 또한 어떤 글이든지 어려움 없이 이해한다고 인식하고 있는, 즉 읽기 효능감이 높은 집단은 읽기 점수도 높다. 또한 수준 높은 삶을 위해서는 읽기가 중요하다고 생각하는, 즉 읽기 가치관이 높은 집단은 읽기 점수도 가장 높다. 따라서 대체로 읽기 능력이 높다면 읽기 효능감이나 읽기에 대한 가치관이 형성되어 있을 가능성이 높고, 따라서 읽기를 삶의 질과 직결된 핵심 역량으로 인식하고 있다고 볼 수 있다.

다음으로는 직업별 읽기 능력이다. 정신노동직(사무·전문직)이 육체노동직보다 평균 약 2.5~3점 높으며, 4수준 비율도 41.1% 대 28.5%로 큰 차이가 보인다. 이는 직업 환경이 텍스트 처리와 정보 활용의 빈도와 연관되어 있음을 보여 주면서, 동시에 읽기 능력은 단순 언어 능력이 아니라 ‘지식 노동’의 기초

가 되는 역량임을 재확인시켜 준다.

끝으로 지역 규모에 따른 구성원들의 읽기 능력이다. 대도시·중소 도시는 읍면 지역보다 4수준 비율이 약 4~5%p 더 높다. 이는 지역 간 교육 접근성 및 독서문화 기반 차이를 시사하며, ‘지역 문해 격차(local literacy gap)’ 현상을 보여 준다. 따라서 지역 간 문해력 격차를 완화시킬 수 있는 생활 기반 문해, 공공 독서 인프라 확충 등이 필요하다.

또한, 매체·디지털 환경 변화에 따른 읽기 양상과 관련한 시사점은 다음과 같다. 우선, 매체별로 읽기 능력에 미묘한 차이가 있다. 앞서 언급한 것처럼, 종이 매체(신문·잡지)를 ‘자주 이용하는 집단’이 읽기 점수가 유의하게 높았으며(그 역은 성립하지 않음), 반대로 텔레비전, 라디오 등 전통 매체(올드 미디어)를 과다하게 이용하는 응답자는 낮은 점수(47.93점)로 나타났다. 그러나 디지털 기기를 많이 사용한다고 해서 무조건 낮은 읽기 점수를 받은 것은 아니다. 디지털 매체 중에서도 ‘인터넷 포털(검색형)’과 같은 정보탐색형 매체는 읽기 능력에 긍정적 영향(상위집단 50.30점)을 끼치는 반면, 게임·웹툰·웹소설과 같은 유희형(엔터테인먼트용) 매체를 과다 이용한 집단은 읽기 점수가 낮다(47.36점). 이는 ‘정보 중심적 읽기’는 문해력 유지에 긍정적이나, ‘오락 중심 읽기’는 문해력 약화로 연결됨을 시사한다.

다음으로, 주로 이용하는 매체에 따라라도 읽기 능력의 점수 차이가 생겼다. ‘짧은 동영상(52.07점)’ 이용자의 읽기 능력 점수는 비교적 높으나, ‘책·신문’(47.66점), ‘텔레비전’(47.35점) 이용자는 가장 낮았다. 즉 높은 디지털 기기 보급률과 함께 15초 이하의 짧은 동영상이 인지 부담도 적고, 접근 가능성도 높아서 특히 20~40대의 젊은 세대에게 틱톡, 쇼츠 등의 짧은 동영상은 상당히 흔하게 그리고 많이 접하는 매체 장르인 반면, 텔레비전이나 라디오와 같은 전통 매체(올드 미디어)의 경우에는 60대 이상에게 친숙한 매체이다. 따라서 짧은 동영상 이용자의 읽기 점수가 높다는 것이 이해된다. 하지만, 이상의 논의와 함께, 읽기 분량이 줄어들수록 읽기 능력이 저하되고 있는 양상으로 미루어 보아, ‘영상 기반 문해력’이 일정 수준의 정보 해석 기능을 대체할 수는 있겠지만, 비판적 독해와 같은 깊이 있는 읽기에는 한계가 있음을 보여 준다. 무엇보다 비판적 독해와 같은 깊이 있는 사고를 요하는 읽기 능력을 신장하기 위해서는 전통적 의미의 긴 글 읽기가 필수적이라고 하겠다.

3) 문법·규범 영역

제3차 국민의 국어능력 실태 조사의 ‘문법·규범 영역’은 제1차와 제2차 국민의 국어능력 실태 조사의 ‘문법 영역’에 ‘규범’을 영역 수준에서 추가한 평가 영역이다. 기존의 문법 영역도 중영역이 ‘규범, 어휘, 문장, 담화’로 구성되어 규범 관련 문항이 포함되어 있었지만, 제3차 조사에서는 규범을 대영역 수준으로 올려 영역명을 ‘문법·규범’으로 변경하고, ‘어휘, 문장, 담화’를 문법 대영역의 중영역으로 삼고 ‘한글 맞춤법, 표준어 규정, 외래어 표기법/국어의 로마자 표기법’을 규범 대영역의 중영역으로 삼았다는 점에서 기존 조사와 차이가 있다.

문법·규범 영역의 성취 수준 등급별 인원 분포는 4수준이 29.6%, 3수준이 37.3%, 2수준이 23.5%, 1수준이 9.6%로 나타나, 3수준의 비율이 가장 높았고, 4수준과 2수준이 그다음으로 높게 나타났다. 1수준의 비율은 다른 수준에 비해서는 가장 낮은 전체의 10%에 조금 못 미치는 정도로 나타났지만, 1수준에 해당하는 문법과 규범 능력이 매우 제한적임을 고려할 때 1수준의 비율을 줄이기 위한 지속적 노력이 필요하다고 할 수 있다.

인구통계학적 변인별 분석 결과에 따르면 성별에 따른 성취 수준의 등급별 인원 분포 차이는 통계적으로 유의하지 않았으나, 연령대, 거주 지역의 규모, 직업군, 학력에서는 유의한 차이가 나타나 이에 대한 논의도 필요하다.

연령대의 경우 연령대가 높아질수록 4수준의 비율이 줄어드는 패턴이 나타남에 주목할 수 있다. 20대의 경우 4수준 비율이 36.9%인데 50대의 경우 4수준 비율이 24.9%, 60대의 경우 4수준 비율이 21.4%이다. 50대와 60대 간 성취 수준 등급별 인원 분포에 유의한 차이가 있었다는 점도 중요하다. 50대는 41.3%로 3수준의 비율이 가장 높고 4수준과 2수준이 각각 24.9%, 24.6%로 비슷한 정도인데 비해, 60대는 3수준 비율이 36.2%로 50대에 비해 낮고, 2수준 비율이 30.5%로 상대적으로 높게 나타났다. 60대의 경우 1수준 비율도 11.9%로 연령대 중 가장 높게 나타났다. 이러한 점을 고려할 때, 생애 주기에 따라 각 연령대에 필요한 문법과 규범 교육이 필요하며, 특히 60대 이상 중장년층을 대상으로 한 맞춤식 교육이 필요하다.

거주 지역의 규모에 따른 능력 차이를 살펴본 결과 읍면 지역이 다른 지역에 비해 4수준의 비율이 낮고 2수준의 비율이 높은 것으로 나타났다. 특히, 읍면 지역의 2수준 비율이 다른 지역에 비해 높게 나타났다라는 점은 지역 규모에 따라 문법·규범 능력의 차이가 왜 발생하는지에 대한 심층 연구가 필요하다는 점을 시사하며, 이에 근거한 정책적 지원 방안 마련도 필요함을 보여 준다. 그 외에도 특히 1수준의 비율이 높은 경우를 유의 깊게 살펴야 한다. 학력의 경우 1수준의 비율이 대재 이상 7.9%, 고졸 10.9%인데 비해 고졸 미만의 경우 15.1%로 매우 높게 나타났다. 1수준에 해당하는 경우 일상생활의 제한적인 규범적인 국어생활에서도 어려움을 겪기 때문에, 이를 대상으로 한 지원책 마련이 요구된다.

문법·규범 영역의 경우 어문 규범(맞춤법, 표준어 등)과 관련된 정보 탐색 빈도와 국립국어원 제공 서비스 이용 빈도를 심층 변인으로 삼아 조사와 분석을 실시하였다. 어문 규범 정보를 ‘매우 자주 찾아본다(하루에 1번 이상)’고 응답한 집단의 평균이 가장 높았으나, 응답자가 6명으로 극히 적어 유의미한 해석은 어려웠다. 또한, 어문 규범 정보를 ‘거의 찾아보지 않는다(한 달에 약 1~2번)’고 응답한 집단의 평균이 그다음으로 높게 나타났다는 점과 어문 규범 정보를 ‘전혀 찾아보지 않는다(한 달에 한 번도 검색하지 않음)’고 응답한 집단의 평균이 가장 낮게 나타났다는 점은 어문 규범 정보를 찾는 태도와 문법·규범 능력의 관련성을 보여 주는 것일 가능성이 있으나, ‘보통이다(일주일에 약 1~3번)’, ‘자주 찾아본다(일주일에 약 4~6번)’로 응답한 집단의 평균이 ‘거의 찾아보지 않음’으로 응답한 집단의 평균보다 낮아 이들 사이의 관계가 간단하지 않음을 알 수 있다. 국립국어원에서 제공하는 <온라인가나다>의 이용 빈도의 경우 매우 자주 이용한다고 응답한 집단의 평균이 가장 높았으나, 응답자가 2명으로 극히 적어 유의미한 해석은 어려우며, 매우 자주 이용한다고 응답한 집단과 전혀 이용하지 않는다고 응답한 집단의 평균이 다른 집단보다 높게 나타났다. 이처럼 응답자 수가 충분하지 않아서 일반화하기 어려운 점과 일정한 경향성을 확인하기 어려운 점으로 인해, 문법·규범 능력과 국립국어원에서 제공하는 문법·규범 관련 서비스 이용의 관계를 확정적으로 설명하기 위해서는 차기 실태 조사 등을 통한 데이터 보강이 필요하다.

4) 쓰기 영역

우선 쓰기 영역 조사 결과의 전반적 경향을 살펴보자면 다음과 같다. 쓰기 영역의 분석 대상 3,000명의 원점수 평균은 169.99점(300점 만점)이었으며, 최솟값 66.80점, 최댓값 295.40점을 보였고, 표준편차는 42.11이었다. 쓰기 영역의 평균 점수 169.99점은 수준별 등급으로는 2수준에 해당되어, 조사 대상 국민의 평균적인 쓰기 능력은 ‘2수준’에 해당하는 것으로 해석할 수 있다.

쓰기 영역의 수준별 분포를 살펴보면, 4수준이 335명(11.2%), 3수준이 791명(26.4%), 2수준이 1,216명(40.5%), 1수준이 658명(21.9%)으로 산출되었다. 이전 시기 조사와의 수준 등급별 분포와 비교할 때 ‘국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 대부분(80% 이상) 획득한

수준'에 해당하는 4수준에 해당하는 비율이 증가한 것을 확인할 수 있으며(제2차 조사: 4.0%, 제3차 조사: 11.2%), 앞 장에서 제시한 바와 같이 제1차~제3차 실태 조사를 연계 분석한 결과에서도, 국민 전반의 쓰기 능력이 다소간 상승한 것으로 나타났다.

<표 VI-1> 쓰기 영역의 수준 등급별 분포의 조사 시기별 비교

수준	2013년(제1차 조사)		2018년(제2차 조사)		2025년(제3차 조사)	
	인원(명)	비율(%)	인원(명)	비율(%)	인원(명)	비율(%)
4수준	55	3.6	121	4.0	335	11.2
3수준	341	22.7	790	26.4	791	26.4
2수준	829	55.3	1,254	41.8	1,216	40.5
1수준	275	18.4	835	27.8	658	21.9
계	1,500	100.0	3,000	100.0	3,000	100.0

그러나 수준 등급별 분포를 면밀히 살펴보면, '국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 미흡하게(20% 미만) 획득한 수준'에 해당하는 1수준의 비율이 21.9%에 달하는데, 이는 국어능력 조사 대상인 다른 영역의 1수준 비율(듣기: 6.0%, 말하기: 19.9%, 듣기·말하기: 17.0%, 읽기: 13.7%, 문법·규범: 9.6%)에 비해 높은 수치라는 점에서 주목할 필요가 있다. 일반 국민 5명 중 1명 이상이 쓰기 영역에서 어려움을 겪고 있으며, 모든 문식 능력 중 쓰기 능력이 가장 미진하다는 것을 보여 주는 결과로 해석할 수 있기 때문이다. 특히 2013년 제1차 조사와 비교할 때 4수준에 해당하는 비율의 증가와 동시에 1수준의 비율 또한 높아졌다는 것은(제1차 조사: 18.4%, 제2차 조사: 27.8%, 제3차 조사: 21.9%), 전체 평균이 상승한 것과는 별개로 국민 간 쓰기 역량 격차가 심화되고 있음을 시사하는 것으로 볼 수 있다. 모든 국민이 기본적인 시민 역량으로서 자신의 생각을 글로 표현할 수 있는 최소한의 문식 능력을 갖추어야 한다는 점에서 1수준에 해당하는 국민의 쓰기 능력 증진을 위한 정책적인 지원 마련이 필수적일 것으로 보인다.

다음으로는 배경 변인에 따른 쓰기 능력의 차이를 살펴보겠다. 첫째, 연령대에 따른 쓰기 영역의 평균 점수(T점수)는 전반적으로 연령대가 높아질수록 낮아지는 경향을 보이며, 구체적으로는 30대(51.20), 20대(50.97), 40대(50.70), 50대(49.73), 60대(47.84) 순으로 나타났다. 또한 20대에서 60대로 연령대가 높아질수록 1수준 비율이 늘어나며(18.0% < 18.3% < 18.9% < 23.5% < 29.3%), 특히 60대의 경우 모든 다른 연령대와 유의한 차이를 보이는 것을 확인할 수 있다. 연령의 증가에 따른 생리적, 인지적 변화를 고려할 때 자연스러운 결과로도 볼 수 있으나, 30대가 가장 높은 능력 점수를 보인다는 점, 1수준과 2수준의 비율만 보았을 때, 30대(57.8%) < 40대(57.9%) < 20대(58.6%) < 50대(62.8%), 60대(73.6%)의 순으로 높아진다는 점은 주목할 만하다. 이는 비자발적(공적) 글쓰기 빈도가 높을수록 쓰기 능력이 높아진다는 결과와 함께 고려할 때, 쓰기 능력이 단지 교육 수준의 결과가 아니라 사회적 활용 경험에 의해 강화되는 실천적 역량임을 보여 준다고 할 수 있기 때문이다. 특히 유의한 차이는 아니지만 교육 현장에서 멀어지는 30대의 능력 점수가 20대의 능력 점수가 더 높다는 점, 1, 2수준의 비율은 20대가 30, 40대보다 더 높다는 점은 20대가 다른 연령대보다 상대적으로 텍스트보다는 영상이나 단문 위주의 의사소통에 익숙해 있다는 점과 더불어, 직무 경험을 바탕으로 한 쓰기 경험이나 문서 위주의 소통 방식에 대한 숙련도가 낮다는 점을 함께 고려할 필요가 있을 것이다. 즉, 문해력 정책 마련 시 단순한 재교육에 그치지 않고 직업적, 사회적 환경에서 일정 수준의 공적 쓰기 활동을 지속적으로 수행하는 경험을 위한 사회적 학습 환경을 제공하는 것이 필요할 것으로 보인다.

둘째, 거주 지역 규모를 중심으로 살펴볼 때 대도시 거주자의 쓰기 능력이 중소 도시나 읍면 지역보다 유의하게 높다는 것은 대도시의 경우 교육 및 문화적 인프라가 밀집되어 있어 쓰기와 읽기를 비롯한 여러 문식 경험을 다양하게 축적할 수 있는 환경이라는 점과 밀접하게 관련된다. 특히 4수준의 비율은 중소 도시(13.0%) > 읍면 지역(10.0%) > 대도시(9.8%)로, 대도시가 다소 낮은 반면, 1수준의 비율은 읍면 지역(26.8%) > 중소 도시(24.7%) > 대도시(17.5%)로 낮게 나타났다. 이는 대도시 거주자의 전반적인 쓰기 능력 수준이 높게 나타난 것이 개인의 고등교육 수준이나 역량 차 때문이라기보다, 문화·정보 자원이 풍부한 생활 환경을 통해 쓰기 학습 및 실천 기회가 균등하게 제공됨으로써 집단 내 격차가 완화될 수 있음을 함의하는 것으로도 볼 수 있다.

셋째, 학력의 경우 전체 변인 중 쓰기 능력에 대한 예측력이 가장 높게 나타났는데, 이는 쓰기가 단순한 기술이 아니라, 추상적 사고·논리 구성·비판적 읽기 능력과 깊이 연관된다는 점을 반영한다. 대학교육은 보고서, 논문, 에세이 등 구조화된 글쓰기 훈련의 기회를 제공하는데, 이러한 교육 경험이 개인의 사고 체계와 언어 표현력에 장기적인 영향을 미칠 수 있기 때문이다. 더불어 ‘대재 이상’이라는 학력 변인과 ‘비자발적(공적) 글쓰기’ 변인의 상관도가 높다는 점($r=0.422$)을 고려할 때, 대재 이상의 학력자일수록 비자발적(공적) 글쓰기 기회가 많고, 이에 따라 쓰기 능력이 신장되었을 가능성도 함께 고려해야 한다. 심층 변인 중 ‘글을 충분히 쓰지 못하는 이유’에 대해 ‘대재 이상’ 집단은 ‘인터넷 검색, 동영상 시청 등을 더 좋아해서’라고 응답한 비율이 48.7%, ‘글을 어떻게 써야 할지 몰라서’가 14.5%인 데 반해, ‘고졸 미만’ 집단은 ‘인터넷 검색, 동영상 시청 등을 더 좋아해서’가 16.8%, ‘글을 어떻게 써야 할지 몰라서’가 36.8%에 해당한다는 점, ‘비자발적 글쓰기 빈도’에서 ‘전혀 쓰지 않음’에 응답한 비율이 ‘대재 이상’ 집단은 8.3%인 데 반해, ‘고졸 미만’ 집단은 54.6%에 달한다는 점도 이러한 해석을 뒷받침한다.

<표 VI-2> 심층 변인 중 ‘글을 충분히 쓰지 못하는 이유’에 대한 학력별 기술 통계(단위: %)

수준	사례 수(명)	글쓰기에 흥미가 없어서	글을 어떻게 써야 할지 몰라서	인터넷 검색, 동영상 시청 등을 더 좋아해서	시간이 없어서	다른 사람들이 내 글을 읽는 게 부담스러워서
고졸 미만	280	36.4	36.8	16.8	6.8	3.2
고졸	1,015	24.0	25.4	36.6	10.3	3.6
대재 이상	1,705	11.4	14.5	48.7	20.2	5.1

<표 VI-3> 심층 변인 중 ‘비자발적(공적) 글쓰기 빈도’에 대한 학력별 기술 통계(단위: %)

수준	사례 수(명)	전혀 쓰지 않음	1년에 2~3회	1개월에 2~3회	1주일에 2~3회	거의 매일
고졸 미만	280	54.6	25.7	12.1	4.6	2.9
고졸	1,015	29.3	31.5	26.4	10.0	2.8
대재 이상	1,705	8.3	19.7	29.6	27.8	14.5

따라서 학력 격차로 인한 언어 격차를 줄이기 위해, 비(非)대학 교육 경로에서도 공적 글쓰기 훈련 기회를 제공해야 하며, 직업적 환경 외에도 사회·문화적 활동을 통해 글쓰기를 자연스럽게 경험할 수 있는 환경을 조성할 필요가 있다.

넷째, 쓰기에 대해 ‘나 자신의 생각이나 감정을 이해하고 표현하기 위해 필요하다’고 인식하는 사람

일수록 쓰기 능력이 유의하게 높았는데, ‘일상이나 학업, 업무를 원활히 수행하기 위해 필요하다’, 혹은 ‘다른 사람과 좋은 관계를 맺기 위해 필요하다’와 같이 쓰기를 외재적 목적을 위한 것으로 인식하는 경우에는 쓰기 인식이 쓰기 능력과 유의한 관련을 보이지 않는다는 결과와 비교할 때 주목할 만한 점이다. 이는 필자의 쓰기 신념을 ‘전달적(transmissional) 신념’과 ‘상호작용적(transactional) 신념’으로 구분하여, 상호작용적 신념이 강한 필자는 글쓰기를 단순한 정보 전달의 행위로 보지 않고 독자와의 의미 구성 과정으로 인식함으로써 인지적·정서적 관여 수준이 높으며 쓰기 수행과 결과물의 질도 우수하다는 점을 밝힌 화이트와 브루닝(White & Bruning, 2005)의 연구 결과와도 일치하는 결과라고 할 수 있다. 초중등 국어과 교육과정에서 ‘지식·기능·태도’를 고르게 교수·학습하는 것을 목표로 삼고 있듯, 성인 쓰기 교육의 목표 또한 쓰기 기능적 측면에만 둘 것이 아니라, 쓰기를 통한 자기 표현 기회를 제공함으로써 쓰기의 가치를 내면화하고 쓰기에 대한 긍정적 인식을 함양할 수 있도록 해야 할 것이다.

이어서, 조사 영역에 따른 쓰기 능력의 차이를 살펴보자면 다음과 같다. 국민의 국어능력 실태 조사의 경우 국민의 실제 언어 사용을 진단하는 데 목적이 있다. 이를 위해 본 연구에서는 사회적 소통 행위로서의 쓰기 수행 능력을 평가하고자 하였으며, 담화 목적에 따라 ‘설명적, 논증적, 친교적’ 글쓰기 영역을 설정하고, 최근의 실제적 문식 환경을 반영하여 복합양식 자료 활용 및 온라인 맥락에서의 책임 있는 소통 능력을 포함하는 ‘디지털 매체 글쓰기’ 영역까지 포함하여 총 네 가지 영역에 대한 쓰기 능력 조사가 시행되었다.

문항 특성 분석 결과 고전 검사 이론(CTT)에 따른 문항별 난이도(용이도)는 ‘설명하는 글 쓰기(W1)’, ‘논증적 글 쓰기(W2)’, ‘친교적 글 쓰기(W3)’의 경우 각각 0.54, 0.54, 0.56으로 ‘중간’(0.4 이상 0.6 미만)에 해당하며 유사하게 나타났으나, 이전 시기 조사와 달리 이번 시기에 새롭게 추가한 ‘디지털 매체 글쓰기(W4)’ 문항의 경우 0.63으로 ‘쉬움’(0.6 이상 0.8 미만)으로 나타났다. ‘친교적 글쓰기(W3)’와 비교할 때, 각각 ‘요청’ 및 ‘사과’를 위한 일상적인 글쓰기 장르라는 점, 3문장 이상의 분량을 작성하는 짧은 글쓰기라는 점에서 유사함에도 ‘디지털 매체 글쓰기’의 난이도가 더 낮았다는 것은 특기할 만하다. 문항 자체의 특성에 대한 면밀한 고찰이 함께 이루어져야 하겠으나, 기본적으로 국민들의 디지털 매체 글쓰기 능력이 낮지 않음을 방증하는 결과로 볼 수 있기 때문이다. 또한 디지털 매체 글쓰기 문항의 평가 항목별 결과를 세부적으로 비교해 보았을 때에도 ‘매체 활용’ 항목의 점수가 전통적인 글쓰기의 평가 항목인 ‘내용, 조직, 표현’ 점수보다 높은 것을 확인할 수 있었는데, 이 또한 누리소통망(SNS)나 이모티콘을 활용한 메시지 등을 통해 복합매체 활용 글쓰기의 친숙도가 높아짐에 따라 그 역량도 함께 높아진 것으로도 해석할 수 있을 것이다. 추후 복합매체 활용 글쓰기의 구인을 보다 세부적으로 탐색하고, 디지털 매체 글쓰기 역량과 전통적 글쓰기 역량과의 관계를 탐색할 수 있는 문항 개발 및 평가 결과 간 비교 분석이 이루어질 필요가 있다.

6.2. 국어 정책 활용 방안

지금까지 제3차 국민의 국어능력 실태 조사의 결과를 분석하고 각 영역별로 어떠한 시사점을 얻을 수 있는지 확인하였다. 본 절에서는 이상의 논의를 바탕으로 국어 정책 차원에서 고려해야 할 사항을 듣기, 듣기·말하기, 말하기, 읽기, 문법·규범, 쓰기 영역 순으로 제언한다.

우선 듣기, 듣기·말하기, 말하기 영역의 조사 결과를 보면 우리나라 국민들의 구어 의사소통 능력에 세대별 언어 경험의 차이, 지역 간 언어문화 자원의 불균형, 디지털 기반 소통 환경의 급격한 변화, 의사소통을 통한 상호작용 경험의 차이 등의 요인이 복합적으로 작용하면서 영향을 미치고 있는 것으로 나타났다. 특히 고령층은 정보 처리 속도의 저하와 더불어 디지털 기반 음성, 영상 환경에 충분히 적응하지 못하는 경향을 보였으며, 젊은 층은 다양한 매체를 활용한 소통에 익숙하였고 다양한 공적 말하기 상황에 비교적 많이 노출되고 참여하고 있었으나 짧고 단편적인 소통에 치우치는 경향이 있었다. 읍면 지역에서는 다양한 담화에 참여하는 경험이 제한적이어서 구어 의사소통 능력을 함양할 수 있는 기회가 도시 지역에 비해 상대적으로 부족하였다. 이러한 소통의 상황은, 구어 의사소통 능력을 이해하기 위해서는 사회적 참여, 직무 수행의 상황, 매체 환경의 변화에 따른 적응 능력 등과 같은 요인을 고려하여야 할 필요가 있음을 의미하는 것이다. 따라서 국민의 구어 의사소통 능력을 위한 국어 정책을 마련할 때 다음과 같은 점을 고려해야 할 것이다.

첫째, 세대별로 상이한 구어 의사소통 방식과 언어적 경험을 고려한 맞춤형 구어 의사소통 교육이 필요하다. 연령이 높아질수록 듣기, 말하기 능력이 하락하는 경향이 나타났는데, 이를 단순히 고령층의 노화에 따른 능력의 감소로 보기보다는 조사 결과에서도 알 수 있듯이 매체 환경의 변화 및 사회적 상호작용 기회의 감소에 따른 결과로 보는 것이 적절하다. 특히 고령층은 문자 기반의 정보 처리에 익숙하여 영상이나 음성을 통한 소통에 다소 어려움을 겪는 경향을 보이므로, 이들을 위한 교육은 디지털 기반의 실생활 의사소통 능력 중심으로 구조화될 필요가 있다. 반면 젊은 층은 디지털 활용 능력이 뛰어나지만 짧은 영상이나 메신저 중심으로 소통을 하는 경향을 보이므로, 깊이 있는 듣기나 체계적인 말하기 능력이 약화될 가능성이 있다. 따라서 이들을 위한 교육은 단순한 기술적 소통보다는 협력적 의미구성 능력, 다양한 공적 담화 수행 능력, 공적 언어에 대한 감수성을 강화하는 방향으로 이루어질 필요가 있을 것이다. 중장년층의 경우, 직무 환경의 변화와 업무 의사소통 구조의 복잡화로 인하여 구어 소통 능력이 직무 역량과 밀접하게 연계되고 있다. 그러므로 이들에게는 직무 상황에서의 의사소통 능력 강화를 위한 교육 프로그램을 제공함으로써 직무 의사소통에서 자기 효능감을 높일 수 있도록 하는 접근이 필요하다.

둘째, 지역 간 언어문화 자원의 격차를 줄이기 위한 지역 기반 언어 정책이 필요하다. 읍면 지역에서는 공적 담화에 참여하거나 경험할 수 있는 기회가 적고, 다양한 언어 공동체와 상호작용할 기회가 도시 지역에 비해 다소 부족하다. 이는 매체의 접근성 문제를 넘어서 지역의 언어문화 생태계가 협소하다는 것을 의미하며, 이러한 문화자원의 격차가 듣기, 말하기 능력의 약화로 이어지고 장기적으로 지역 주민들의 사회적 상호작용이나 직무 상황에도 영향을 줄 수 있다. 따라서 지역별 국어문화센터를 중심으로 지역 주민들의 실제 구어 의사소통 능력을 함양할 수 있는 다양한 교육 프로그램을 제공할 필요가 있다. 또한 지역 학교, 도서관, 평생학습관 등과 연계한 언어문화 함양 프로젝트 등의 운영을 지원함으로써 지역 사회에서 다양한 담화 상황에 참여하는 기회를 제공하려는 노력도 이루어져야 할 것이다.

셋째, 디지털 사회로의 전환에 대응하여 새로운 형태의 구어 의사소통 평가 체제를 마련할 필요가 있다. 오늘날의 듣기, 말하기 상황은 단순히 음성 언어를 기반으로 하는 소통에서 나아가 복합 양식의 통

합적 처리 능력을 요구한다. 따라서 기존의 단방향적 듣기나 형식적인 말하기 평가로는 실제의 의사소통 능력을 충분히 파악하거나 평가하기에는 한계가 있다. 디지털 환경에서는 청자가 어떻게 정보를 선별하고 해석하는지, 화자가 어떤 방식으로 자료를 구성하여 발화하는지가 중요하므로, 이러한 능력을 타당하게 평가하기 위한 방안을 개발할 필요가 있는 것이다. 이는 단순히 평가 도구의 개발을 넘어서 국민의 구어 의사소통 능력을 시대 변화에 맞추어서 정의하고 그 기준을 새롭게 정립하는 과정이 됨으로써, 장기적으로 국민의 구어 능력을 함양하는 데 기여하는 방안을 마련하는 데 도움을 줄 것이다.

넷째, 구어 의사소통 능력의 세대·지역·학력·직업 간 격차는 단기에 해소될 문제가 아니므로, 데이터 기반 정책 피드백 체제를 구축하여 지속적으로 지원하는 것이 필요하다. 현재의 언어 사용은 온라인과 오프라인 환경에서 끊임없이 변화하고 소통 방식도 빠르게 변화하고 있으므로, 정책이 효과적으로 마련되고 운영되기 위해서는, 정기적인 실태 조사, 지역별 언어문화 지표 개발 등이 요구된다. 특히 구어 의사소통 능력은 단순한 기술이 아니라 사회적 환경, 직업 상황, 디지털 매체의 변화에 민감하게 반응하는 능력이므로, 세대별 상호작용 방식의 변화, 직업군 기반 언어적 요구 분석, 지역의 언어문화 자원의 분석, 디지털 매체의 사용 유형 등을 통합적으로 분석하는 국가 차원의 데이터 체계를 구축하는 것이 중요하다. 이러한 데이터 체계를 바탕으로 한 정책 모니터링 시스템을 개발, 운영하면 국어 정책이 단발성 사업에 그치는 것이 아니라 사회적 불평등, 지역 격차, 디지털 접근성 문제와 연동된 구조적인 문제들을 해결해 나갈 수 있을 것으로 기대된다.

읽기 영역의 경우 오늘날 한국의 읽기 상황이 빠른 디지털 환경으로의 변화 속에서 ‘깊은 읽기’가 감소되고, 단편적 정보 소비가 확산되며, 세대 간 문해 격차가 심화되고 있음을 알 수 있다. 또한 세대별 경향으로, 젊은 세대는 디지털 탐색형 읽기가 강점이라면, 고령층은 전반적으로 독해력이 약화되고 있음도 발견할 수 있다. 이를 바탕으로, 국어 정책 제안을 위한 시사점을 정리하면 다음과 같다.

첫째, 세대별 측면에서 각 연령별 요구와 수준에 맞는 맞춤형 문해교육이 필요하다. 앞서 언급했듯이, 연령이 높을수록 읽기 능력이 감소되는데, 이런 현상을 당연시해서는 안 될 것이다. 읽기 영역 분석 결과를 살펴보면 전체적으로 인터넷 웹 기반 자료의 비중이 늘면서 자연스럽게 글보다는 영상 중심의 정보 습득 비율이 높아졌다. 아울러 글 읽기가 다른 일에 방해가 된다고 인식한 비율이 상대적으로 높게 나타나기도 하였다. 이러한 현상은 읽기의 가치관, 즉 읽기를 중요하게 생각하지 않는 인식을 보여 준다. 특히 20대의 경우는 읽기 점수는 다른 연령층에 비해 높지만, 짧은 영상 중심의 정보 습득 경향이 높고, 읽기의 중요성에 대한 인식이 부족할 수 있다는 점에 유의할 필요가 있다. 한편 고령층(50~60대)의 읽기 점수 저하는 단순한 인지 감소라기보다는 변화되는 매체 환경에서의 적응 문제로 바라봐야 한다. 따라서 노년층 읽기 교육은 단순히 글자(한글)를 익히고 글 내용을 파악하는 독해 훈련이 아니라 정보 탐색과 평가를 중심으로 하는 ‘디지털 기반 정보 문해력(digital literacy)’ 교육으로 전환되어야 함을 시사한다.

둘째, 전인 교육적 측면에서, 인지 발달만큼 중요한, 정의적(情意的) 발달을 돕는 읽기 교육이 요구된다. 읽기 효능감, 읽기 가치관 등의 정의적인 읽기 태도 형성은 평생 독자로서의 성장, 독서 문화 저변화에 필수적이다. 따라서 장기 몰입형 혹은 공감형 독서 프로그램을 마련할 필요가 있다. 학교 교육기를 벗어난 청년층이 기습적인 읽기 능력과 긍정적 읽기 태도를 바탕으로 하여, 평생 독자로서 읽기를 교양과 놀이로서 즐기며 계속해서 성장할 수 있도록 하는 것은 교육적으로 매우 중요한 지점이다. 이를 위해서도 온라인이나 오프라인 등을 통해서 모여서 함께 읽고 생각을 나누는 사회적 독서와 같은 읽기에의 참여와 실천 프로그램이 마련될 필요가 있다.

이뿐만 아니라, 장년층을 위한 독서 정책에서는 어떤 글이나 정보도 잘 습득할 수 있다는 자기 인식, 즉 읽기 효능감(reading efficacy)이나, 읽기가 일상적 삶이나 직업 세계에서도 매우 중요하다는 읽기에 대한 인식을 심어 주는 메타인지적 접근이 요구된다. 예를 들면, 40~50대 연령층은 가정과 직업 세계에

서도 전환기에 있을 가능성이 높다. 이때 ‘어떤 글이든지 어려움 없이 이해할 수 있다’는 자신감을 길러 주고, 이를 바탕으로 ‘수준 높은 삶을 위해’ 읽기가 필요함을 알고, 읽기를 삶의 질 향상과 연계된 문화적 가치로 인식시키는 교육이 중요하다. 이를 위해서는 읽기를 통해 업무 능력을 개선할 수 있는 업무 능력 신장 독서 프로그램을 바탕으로, 읽기 효능감을 함께 신장시킬 수 있는 프로그램 등을 계획할 수 있다.

또한 청년층은 기술적 문해력(technical literacy)보다 정서적·인지적 자신감을 높일 수 있는 정의적 차원의 접근이 요구된다. 예를 들면, 긴 호흡으로 한 권의 책을 지속적으로 끝까지 읽는 경험을 하게 하거나, 그 과정에서 세심한 읽기(close reading)나 지속적 읽기 몰입의 경험을 누릴 수 있도록 하는 것이 중요하다. 특히나 활동력이 높은 청년층의 경우에는 독서 행위로의 참여와 실천으로 연결될 수 있도록 다른 사람과 읽은 글에 대해 감상을 나누며 경험을 공유하는 ‘사회적 독서(social reading)’, ‘공유 독서(shared reading)’와 같은 교육 프로그램을 개발하여 보급할 필요가 있다.

셋째, 사회적 요구 측면에서는 학력, 직업, 지역에 따른 읽기(문해력) 불평등 완화를 위한 평생 문해 정책이 필요하다. 현대 사회에서 문해력은 단순한 읽기·쓰기 능력을 넘어, 정보를 해석하고 비판적으로 활용하는 핵심 역량으로 자리 잡고 있다. 그러나 디지털 전환, 노동시장 양극화, 지역 간 교육자원의 불균등 분포로 인해 문해력의 수준은 사회·경제적 지위와 밀접하게 연결되고 있다. 고학력층과 저학력층, 전문직과 단순직, 수도권과 비수도권 간의 문해력 격차가 확대되고 있으며, 특히 중장년층과 고령층은 디지털 문해력(digital literacy)에서 상대적으로 취약한 것으로 나타난다. OECD의 ‘성인역량조사(PIAAC)’ 결과에서도 한국은 성인 문해력 수준의 세대 간, 학력 간 격차가 크고, 사회적 이동 가능성을 제약하는 요인으로 작용하고 있다(OECD, 2016: 24-26).

따라서 디지털 기술 발전과 정보환경의 변화 속에서, 기존의 ‘문자 해독 중심 문해’는 ‘비판적, 디지털, 시민 문해’로 확장되어야 하고, 문해력은 노동시장 참여, 소득 수준, 사회참여에 직접적 영향을 미치므로, 문해 격차 완화는 사회적 불평등 해소의 기초가 되어야 한다. 또한 고령층·저학력층의 디지털 문해력 부족은 정보 접근의 장벽이 되어 사회적 고립을 심화시키기 때문에, 지역 간 공공도서관, 평생학습관, 독서문화 기반의 불균형을 완화함으로써 지역사회의 사회자본(social capital)을 강화할 수 있어야 한다. 이에 따라 정부 차원의 평생 문해 정책은 ‘사회적 불평등 완화, 포용적 사회 실현’을 위한 목적으로 이뤄져야 한다.

넷째, 디지털 전환 대응 측면에서, 동영상이나 누리소통망(SNS) 중심 매체 환경에서의 사실 확인(팩트 체크)을 위한 정보 판별 읽기 훈련이나, 근거 기반 비판적 읽기 교육이 강화될 필요가 있다. 앞서 살펴 보았듯이, 읽기 및 문식 환경의 급속한 변화로 읽기의 매체가 디지털화되고 노출되는 읽기 환경이 달라지면서, 세대별로 읽기의 방식과 양상에 차이가 발생하였다. 그 결과 읽기 환경은 텍스트 중심에서 멀티모달(multimodal), 즉 영상·이미지·누리소통망(SNS) 중심으로 빠르게 이동했다고 볼 수 있다. 정보 소비의 형태가 짧고, 시각적인 콘텐츠 중심(쇼트폼, 릴스, 틱톡 등)으로 변화하면서, 학습자들은 긴 호흡의 읽기나 논리적 사고를 전개하기보다 단편적 정보 처리에 익숙해지면서, 매체의 신뢰성이나 타당성 혹은 공정성 등에 대한 검증 과정 없이, 쉽게 편견과 고정관념에 물들게 된다.

디지털 매체는 정보 접근을 확장시켰지만, 동시에 가짜 뉴스(fake news), 확증 편향(filter bubble), 정보 조작(manipulation) 등의 위험을 키웠다. 따라서 단순히 정보를 파악하는 것을 넘어, ‘출처의 신뢰성 판단, 정보의 의도 파악, 텍스트 간 비교·통합, 논리적 추론 및 근거 평가’와 같은 비판적 문해력(critical literacy) 교육이 필수적이다. OECD도 디지털 환경에서 학생들이 정보의 신뢰성을 평가하고 맥락을 이해하는 능력이 점점 더 중요해지고 있으며, 이러한 ‘비판적 읽기(critical reading)’가 학업 성취와 시민 참여의 핵심 요소임을 강조하고 있다(OECD, 2021: 10-19).

이를 바탕으로 다음과 같은 정책 예시를 제안할 수 있겠다.

1. 세대별 맞춤형, 전인 교육 문해교육 프로그램 개발

[전체] 세대 공감 문해력(Generation Literacy Link)

[하위]

◎ “디지털 문해력 60부터”

: 노년층 맞춤형 디지털 문해·독서 융합 프로그램

◎ “책과 세상을 잇는 다리”

: 일상과 가치의 균형을 되찾는 치유형 독서 프로그램

◎ “느린 독서, 깊은 사고”

: 짧고 빠른 읽기에서 벗어나, 깊고 느린 독서를 통해 사고력·정서력 강화

2. 격차 해소, 포용 사회를 위한 문해교육 프로그램 개발

[전체] 국민 평생 문해 기본계획 (National Lifelong Literacy Plan)

[하위]

◎ “평생 문해, 모두의 문해 네트워크”

: 지역 간 문해교육 기관 및 시민단체 협력 기반의 연계 프로그램

◎ “삶을 바꾸는 읽기, 모두를 위한 문해”

: 지역·세대 간 연대와 공동체적 학습을 강조하는 실천 중심 읽기 프로그램

3. 비판적 사고력 개발을 위한 문해교육 프로그램 개발

[전체] 현명하게 판단하는 시민 역량 신장 프로그램

[하위]

◎ “디지털 시대, 읽는 힘이 사고력이다.”

: 알고리즘과 상업적 정보 구조에 휩쓸리지 않고 스스로 의미를 구성하는 능력 함양 프로그램

◎ “디지털 문해력 인증제(DLC, Digital Literacy Certificate)”

: 정보 불평등 완화, 허위 정보 확산 방지, 시민 비판의식 능력의 인증을 통한 이력제 도입

요컨대 국민 전체의 평균 점수는 안정적이지만, 연령·학력·직업·매체습관에 따른 구조적 격차가 확인 되었으며, 이는 포용적 문해력(inclusive literacy) 교육으로 나아가기 위한 정책적 대응을 요구한다. 따라서 국어능력이 단순한 언어 기술이 아니라, 시대 변화에 적응하는 핵심 생존 역량(literacy as adaptive capacity)임을 인식하고, 데이터에 기반한 정책을 구체적으로 계획하는 것이 중요하다.

문법·규범 영역의 경우 연령대가 높아질수록 4수준 비율이 줄어드는 경향이 나타나고 높은 연령대에서 1수준 비율이 상대적으로 높게 나타나므로, 노년층 교육에 더 큰 정책 지원이 요구된다. OECD가 10년 주기로 실시하는 국제성인역량조사(PIAAC)의 문해력 영역 결과에서 볼 수 있듯이, 연령의 상향에 따른 문해력 저하는 우리나라뿐이 아닌 세계 공통적인 현상이다. 국민의 국어능력 실태 조사를 통해서도 같은 결과가 확인된 만큼, 국립국어원에서 성인 교육에 더욱 힘쓰되 특히 노년층 친화적인 교육 프로그램을 개발할 필요가 있다. 국립국어원에서 주최하고 있는 국어문화학교 등의 성인 교육 프로그램에서 노년층을 특화한 교육이 이루어질 수 있도록, 옛날과 오늘날이 다른 맞춤법 내용을 다루거나 어법 교육용 텍스트에 노년층 문화를 소재로 하는 등의 노력을 더할 수 있을 것이다. 또한 노년층의 디지털 문해력 수준이 높지 않음을 함께 고려하여, 디지털 문서 작성 시 유용한 문법·규범 관련 기초적 능력 향상 프로그램, 국립국어원 웹서비스에서 국어 지식을 확인하고 활용하는 방법에 대한 안내 프로그램 등의

개발이 필요할 것이다.

한편, 문법·규범 영역의 심층 변인에 대한 분석에서 본조사의 응답자 수가 충분하지 않아서 일반화하기 어려운 점과 일정한 경향성을 확인하기 어려운 점으로 인해, 문법·규범 능력과 국립국어원 서비스 이용의 관계를 확정적으로 설명하는 데에도 어려움이 있었다. 따라서 차기 실태 조사 등을 통한 데이터 보강이 필요하며, 근본적으로는 국립국어원 제공 서비스의 사용자 규모가 전반적으로 늘어나면서 양질의 사용 경험이 충분히 축적되어야 국립국어원의 문법·규범 관련 서비스와 문법·규범 능력의 관계를 내실 있게 설명할 수 있을 것이다.

따라서 데이터 확보를 위한 기반을 다짐과 동시에 대국민 교육적 성격을 강화하기 위해서는, 무엇보다 국립국어원에서 제공하고 있는 국어사전, 어문 규정, <온라인가나다> 등의 문법·규범 관련 웹페이지를 사용자 친화적으로 개편하여 접근성과 활용도를 제고할 필요가 있다. 이를 위해 ‘표준국어대사전’과 ‘우리말샘’의 통합 검색 페이지, 국어사전과 어문 규정 내용 및 <온라인가나다>의 연계 항목 등의 신설을 고려할 수 있다. 현재 국립국어원 웹사이트에는 각 사전을 다른 웹페이지로 제공하고 있으나, 이는 사용자의 접근성과 편의성을 저하하는 원인이 되며 사용자가 두 사전의 장점을 비교하며 활용하기에도 불편함이 있다. 또한 국어사전과 어문 규정이 웹사이트에서 완전히 다른 메뉴로 구별되어 있는 것은 개념적으로 적합하나, 국어사전에 찾은 내용이 어떠한 어문 규정과 관련되는지 찾아보기 어려운 구조가 된다. 따라서 사용자가 여러 국어사전을 한꺼번에 확인하기 편리하면서, 검색한 단어와 관련된 어문 규정 내용 및 그에 대한 <온라인가나다>의 해설도 한 페이지에서 확인할 수 있도록 개선할 필요가 있다. 이때 ‘네이버 국어사전’이 여러 사전을 한 페이지에서 확인할 수 있도록 구분된 탭(tap)을 제공하여 사용자가 여러 웹페이지를 각각 열 필요가 없도록 서비스하고 있으며, 어문 규범 관련 내용도 함께 제공하고 있음을 참고할 수 있다. 이상에 제안한 국립국어원 웹사이트 개편 방안을 통해 접근성과 활용도를 제고하면, 비단 문법·규범 능력과의 연관성 분석의 기반을 확보하는 것뿐만 아니라 대국민 교육적 효과를 기대할 수 있을 것이다.

쓰기 영역의 경우 2025년 국민의 쓰기 능력 조사 결과는 평균 수준의 향상에도 불구하고 국민 간 쓰기 능력의 양극화가 뚜렷해지고 있음을 보여 준다. 특히 세대 간, 거주 지역 간, 학력 간 격차가 두드러진다는 점에서 이를 완화할 수 있는 다음과 같은 국어 정책이 요구된다.

첫째, 고등교육 수준의 글쓰기 교육 접근 기회의 확대와 기초 문해력 강화 정책이 병행되어야 한다. 조사 결과는 고등교육(대학) 이수 여부가 쓰기 능력의 차이를 불러오는 가장 중요한 요인임을 명확히 보여 준다. 이는 학력 수준이 낮은 집단(고졸 미만, 고졸)이 쓰기 능력 취약 계층이 될 가능성이 높음을 의미한다. 따라서 고등교육 이수 기회가 부족한 성인 및 노령층을 대상으로 하는 쓰기 능력 향상 프로그램(문해 교육)을 체계적으로 확대하고 지원하는 정책이 필요하다. 특히 쓰기 능력은 학습된 지식이 아니라 반복적 실천과 경험의 결과로 유지되기 때문에, 직장을 떠나거나 글쓰기 기회가 줄어드는 노년층에게는 공적 글쓰기 경험을 대체할 사회적 학습 환경을 조성해야 한다. 말하자면 ‘성인기 이후의 쓰기 실천 유지’를 지원하는 평생 교육 설계가 필요하다는 것이다. 예컨대, 국가평생교육진흥원에서 주관하는 「성인 문해교육 지원사업」이나 지역자치단체별 「평생학습도시 사업」 등과 연계하여 ‘생활 글쓰기 아카데미’, ‘시민 리터러시 교실’과 같은 실습형 프로그램을 마련할 수 있다. 관련하여 핀란드의 평생 문식력 전략(National Literacy Initiative, 2017~)을 참고할 수 있는데, 핀란드에서는 해당 정책을 통해 성인을 대상으로 “읽고-쓰기-생각하기”의 순환 구조를 강화하고, 지역 도서관을 기반으로 한 쓰기 교육 워크숍을 운영하고 있다. 단순히 기초 문해력 교육에 그치지 않고 디지털 리터러시와 비판적 사고력을 포함한 종합적 문해력 강화를 목표로 한다는 것이 특히 주목할 만한 점이다. 각 지역 커뮤니티센터에서 시사 문제나 사회적 쟁점과 관련한 주제를 중심으로 ‘공적 글쓰기 세미나’를 개최하여, 시민이 글쓰기를 사회적 참여 행위로 인식하도록 설계한 것이다. 앞서 밝혔듯 쓰기가 논리적이고 비판적인 사고와 깊이 연관

된다는 점에서, 실제성(authenticity)을 기반으로 한 이러한 쓰기 프로그램은 쓰기에 대한 접근성을 높임과 동시에, 삶의 제 문제들을 다면적·복합적으로 살피는 고등 수준의 사고력을 요하는 쓰기의 기회를 제공할 수 있다는 점에서 유의미하다.

둘째, 직업군별·지역별로 접근 가능한 실용형 교육 모델을 구축할 필요가 있다. 육체노동, 서비스업 종사자 등 글쓰기 노출이 적은 집단을 대상으로, 직업교육 및 재취업 교육 과정에 실용적인 쓰기 교육을 포함하여, 자연스러운 쓰기 연습 기회를 제공할 필요가 있으며, 읍면 지역이나 중소 도시에서도 접근 가능한 평생교육 시설과 디지털 기반 학습 도구를 확충하여, 지역 간 문해 격차를 실질적으로 완화할 필요가 있다. 영국과 핀란드의 사례처럼, 지역 도서관을 “지역공동체 글쓰기 공간(Community Writers’ Space)”으로 활용하여, 시민 저널, 생활 글쓰기 공모전, 지역 의제 글쓰기 대회 등을 주기적으로 운영하는 것도 가능하다.

셋째, 쓰기 기능뿐만 아니라 쓰기 인식 및 쓰기 태도를 제고할 수 있는 쓰기 교육이나 쓰기 경험의 제공 및 쓰기 문화 조성이 요구된다. 쓰기 교육의 목적이 삶의 제 국면에서 스스로 글을 쓰는 ‘자발적 평생 필자’를 만드는 데 있다고 할 때, 자발적인 필자가 되기 위해서는 쓰기 흥미, 쓰기 동기, 쓰기 효능감 등과 같은 쓰기에 대한 긍정적 태도의 함양이 필수적이다. 쓰기에 대한 긍정적 태도가 자발적인 쓰기 수행을 불러오고, 쓰기에 대한 실제 수행 경험이 쓰기 능력의 신장으로 이어지는 선순환 구조를 만들기 위해서는 우선 국민이 쓰기에 대한 긍정적 태도를 갖출 수 있게 하는 교육 경험이나 쓰기 경험, 또는 쓰기 문화를 마련할 필요가 있다. 예컨대 2017년에 만들어졌던 청와대 ‘국민 청원 게시판’의 경우, 청와대 누리집에 국민이 청원을 등록하고 30일 동안 20만 개 이상의 동의를 받으면 정부나 청와대 관계자들이 청원에 대한 답변을 제공하는 정책이었으며, 2017년 8월부터 2022년 2월까지 총 5억 1,600만 명이 방문하였고 일일 평균 670건의 청원이 게시되었다. 이는 국민이 자신의 경험과 사회적 주제를 글로 표현할 수 있는 공간을 확대함으로써, 쓰기 능력을 공공 담론 참여의 기반으로 발전시킨 사례라고 할 수 있다. 자신의 글을 통해 사회문화적 의제에 참여하거나, 정책을 바꾸는 등 실제 삶의 문제를 해결하는 경험은, 국민들에게 쓰기의 실제성이나 쓰기를 통한 소통의 힘을 체험케 함으로써 쓰기 효능감이거나 쓰기의 가치에 대한 인식을 새롭게 하는 기회를 마련한 것으로 볼 수 있다.

넷째, 향후 디지털 매체 글쓰기 역량에 대한 보다 정교한 논의가 필요하다. 디지털 매체 글쓰기는 국민의 실제 언어 사용 환경을 가장 직접적으로 반영한다는 점에서 중요성이 커지고 있으나, 이번 조사는 디지털 매체 글쓰기 능력 평가를 위한 첫 시도로서 현재의 문항 구성만으로는 실제적 디지털 매체 글쓰기 능력이나 전통적 글쓰기 능력과의 관계를 체계적으로 분석하는 데 한계가 있었다. 따라서 향후 국민의 국어능력 실태 조사는 디지털 매체 글쓰기의 특성을 더욱 정밀하게 반영할 수 있는 문항 개발 연구를 강화하고, 복합매체 요소 활용 능력·온라인 책임성·정보 처리 및 재구성 역량 등 핵심 세부요인을 구분한 평가 체계를 마련해야 한다. 아울러 전통적 글쓰기와 디지털 글쓰기 간의 상관 구조를 탐색하여 두 역량의 관계를 종합적으로 진단할 수 있는 복합 평가 모델을 구축함으로써 국민 문식 역량 변화를 보다 정확하게 파악할 수 있는 분석 기반을 갖출 필요가 있다. 또한 이번 조사에서 국민의 디지털 매체 글쓰기 역량이 낮지 않게 나타난 만큼, 성인 문해 교육에서 디지털 기반 글쓰기 경험을 전통적 글쓰기로 자연스럽게 이행할 수 있는 교육 모형을 도입하는 것도 가능할 것으로 보인다. 단문 중심의 누리소통망(SNS) 글쓰기나 이모티콘 활용 등 디지털 친숙도를 발판으로 삼되, 이를 구조화된 문장 생성·논리 전개·설명·논증적 글쓰기와 연결하는 단계형 교육 프로그램을 구성하여 디지털 친화성과 전통적 문식 역량 간의 균형 있는 발달을 유도해야 한다.

지금까지 제3차 국민의 국어능력 실태 조사 결과를 바탕으로 듣기, 듣기·말하기, 말하기, 읽기, 문법·규범, 쓰기의 각 영역별로 앞으로의 국어 정책 실현 방안에 대해 제언하였다. 국민의 국어능력 실태 조사는 국민의 국어능력을 주기적으로 파악하여 국민의 국어능력 향상을 위해 필요한 정책 마련의 기초

자료가 된다는 점에서 그 중요성을 확인할 수 있다. 본 연구에서는 제3차 실태 조사에 관한 종합 분석 결과를 토대로 정책 제언을 하였으며, 차후 제1차에서 제3차 시기 조사 결과의 연계성을 심층 분석하고 이를 바탕으로 향후 조사 설계를 고도화하는 작업이 이루어져야 할 것이다.

참고 문헌

- 국립국어원·서울대학교 국어교육연구소(연구 책임자: 김종철)(2013), 2013년 국민의 국어능력 평가, 서울: 국립국어원.
- 국립국어원·서울대학교 국어교육연구소(연구 책임자: 민병곤)(2017), 국민의 국어능력 실태 조사 도구 개발, 서울: 국립국어원.
- 국립국어원·서울대학교 국어교육연구소(연구 책임자: 민병곤)(2018), 2018년 국민의 국어능력 실태 조사, 서울: 국립국어원.
- 국립국어원·서울대학교 국어교육연구소(연구 책임자: 구본관)(2022), 국민의 국어능력 실태 조사 개선 기초 연구, 서울: 국립국어원.
- 국립국어원·서울대학교 국어교육연구소(연구 책임자: 구본관)(2023), 2023년 국민의 국어능력 실태 조사, 서울: 국립국어원.
- 국립국어원·서울대학교 산학협력단(연구 책임자: 김성규)(2006), 국어능력 조사 방법 연구, 서울: 국립국어원.
- 국립국어원·서울대학교 산학협력단(연구 책임자: 장소원)(2012), 국어능력 평가를 위한 기초 연구, 서울: 국립국어원.
- 박혜영·천해주·김영서·김주환(2020). 미국의 초·중등 작문 평가의 특징과 시사점 탐색 - 국가 및 주 수준 표준화 작문 평가를 중심으로. *작문연구* 44, 73-107.
- ACARA (Australian Curriculum, Assessment and Reporting Authority) (n.d.). Understanding online assessment . National Assessment Program - Literacy and Numeracy (NAPLAN)
<https://www.nap.edu.au/naplan/understanding-online-assessment>](<https://www.nap.edu.au/naplan/understanding-online-assessment>)
- ACARA (Australian Curriculum, Assessment and Reporting Authority) (2023). *NAPLAN 2023 narrative writing marking guide*. ACARA.
- ACARA (Australian Curriculum, Assessment and Reporting Authority) (2025). *The Australian National Assessment Program - Literacy and Numeracy (NAPLAN) assessment framework*. ACARA.
- Carr, N. G. (2010). *The Shallows: What the Internet Is Doing to Our Brains What the Internet Is Doing to Our Brain*, W. W. Norton & Company.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Science Second Edition*. Lawrence Erlaum Associates.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159-174.
- Li, L., Hu, J., Dai, Y., Zhou, P., & Zhang, W. (2023). Depth-Perception-Based Representation in Holistic Rating on ESL Essay Writing. *Reading & Writing Quarterly* 40(3), 211-235.
- Macwan1, C., Chandni J.V., & Jignesh V. (2024). Exploring the influence of reels and short videos on the reading and listening habits of generation Z: A comprehensive study. *International Journal of Business and Management Practices (IJBMP)* 2(3), 309-333.
- NCES (National Center for Education Statistics) (n.d.-a). The Nation's Report Card - Data & Analysis. <https://nces.ed.gov/nationsreportcard/dba/>

- NCES (National Center for Education Statistics) (n.d.-b). Writing | The Nation's Report Card.
<https://nces.ed.gov/nationsreportcard/writing/>
- OECD (2012). *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing.
- OECD (2016). *Skills Matter: Further Results from the Survey of Adult Skills (PIAAC)*. OECD Publishing.
- OECD (2017). *Adaptive problem solving: Moving towards a new assessment domain in the second cycle of PIAAC*. OECD Education Working Papers No. 156.
- OECD (2019). *PISA 2018 Assessment and Analytical Framework*, OECD Publishing.
- OECD (2021). *21st-Century Readers: Developing Literacy Skills in a Digital World*. OECD Publishing,
- OECD (n.d.). PISA 2025: Learning in the Digital World.
<https://www.oecd.org/en/topics/sub-issues/learning-in-the-digital-world/pisa-2025-learning-in-the-digital-world.html>
- Richard L. S., William M. I., R. L. O., & Kenneth G. G. (2012). 표본조사의 이해와 활용, 김영원 류제복·박진우·홍기학 역, 서울: 교우사.
- Unsworth, L., Cope, J., & Nicholls, L. (2019). Multimodal literacy and large-scale literacy tests: Curriculum relevance and responsibility. *The Australian Journal of Language and Literacy* 42(2), 128-139.
- White, M. J., & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary educational psychology* 30, 166-189.

<Abstract>

The 2025 National Assessment of Korean Language Ability

The purpose of this study is to refine the survey tool for the writing domain, developed in 2024, and to conduct the main survey of Korean language proficiency in this domain. The analysis of the survey results aims to provide foundational data for formulating policy measures to enhance the writing ability of the public. In addition, this study offers a comprehensive analysis of the results from all six domains of the third National Assessment of Korean Language Ability conducted from 2023 to 2025 and links these findings to those from the first survey in 2013 and the second survey in 2018.

The third National Assessment of Korean Language Ability, conducted annually by domain, focused on the writing domain in this particular study. The assessment instrument for this study was designed based on the foundational research and preliminary assessment conducted in 2024, but was selected based on expert advice and the results of the preliminary assessment, including the discrimination and difficulty of the questions, the response rate, and the analysis of differences by variables. In addition, this study expanded the assessment background variables to examine the writing ability in more depth. In addition to the variables related to demographic information, variables such as perceptions of writing, writing frequency, frequency of positive writing experiences, frequency of using generative AI during writing, and the presence of writing assistants were included as in-depth variables.

The main assessment was conducted over a two-month period from June to August 2025, using a one-on-one door-to-door interview survey of 3,000 adult males and females aged 20-69 living in 17 cities across the country. Two-stage stratified sampling and probability proportional to size sampling were used to select participants, utilizing demographics to ensure representativeness of the population. In addition, in order to explore ways to improve the stability and reliability of the survey, 50 people were added to the collective assessment. As such, the total number of respondents in the main assessment was 3,050, but the statistical analysis of the results is based on the responses of 3,000 respondents to ensure the representativeness of the population, and the responses of 50 respondents from the collective assessment are used for comparison with the door-to-door interview method.

The average score for the writing domain was 169.99 (out of 300, standard deviation

42.11), and the average ability score (T-score) was 50 (out of 100, standard deviation 8.60). Distribution by proficiency was as follows: 11.2% (335) at Level 4, 26.4% (791) at Level 3, 40.5% (1,216) at Level 2, 21.9% (658) at Level 1. By variable, in writing domain, the difference by sex was not statistically significant. The difference by age group was statistically significant, with proficiency scores highest among those in their 30s, followed by those in their 20s, 40s, 50s, and 60s. The difference in proficiency scores by size of residential area was statistically significant, with proficiency scores highest in large cities, followed by small- and medium-sized cities and rural areas. By occupation, proficiency scores were highest for mental labor, followed by other, and then manual labor, and these differences were statistically significant. For education, proficiency scores were in the order of college or higher, high school diploma, and less than high school diploma, and these differences were statistically significant.

Using the results of the third National Assessment of Korean Language Ability, this study also recalculated cut scores to grade proficiency levels in each domain using a modified Angoff method. A panel of 12 experts—four Korean language teachers, four graduate students majoring in Korean language education, and four university faculty members in the same field—participated in the standard-setting process. Deliberations were conducted twice by item type, covering selected-response and constructed-response, and level judgments proceeded through three rounds, during which panelists were provided with performance descriptors, descriptive statistics from the main survey, and results of previous rounds. The resulting proficiency-level distributions were as follows: for listening, 40.6% at Level 4, 27.7% at Level 3, 25.7% at Level 2, and 6.0% at Level 1. For listening and speaking, 23.3% at Level 4, 34.1% at Level 3, 25.6% at Level 2, and 17.0% at Level 1. For speaking, 18.1% at Level 4, 29.8% at Level 3, 32.3% at Level 2, and 19.9% at Level 1. For reading, 33.0% at Level 4, 28.5% at Level 3, 24.9% at Level 2, and 13.7% at Level 1. For grammar and norms, 29.6% at Level 4, 37.3% at Level 3, 23.5% at Level 2, and 9.6% at Level 1. For writing, 11.2% at Level 4, 26.4% at Level 3, 40.5% at Level 2, and 21.9% at Level 1.

Based on these proficiency-level distributions, this study conducted a comprehensive analysis of all six domains of the third assessment. For demographic variables—sex, age, region size, occupation, and education—the listening and the listening and speaking domains showed no statistically significant differences by gender but significant differences by the other variables. In the speaking domain, statistically significant differences were found for all variables, including sex. In the reading domain, differences by sex were not statistically significant, but differences by age, region size, occupation,

and education were statistically significant. In the grammar and norms domain, sex differences were not statistically significant, while differences by age, region size, occupation, and education were statistically significant. In the writing domain, sex differences were again not statistically significant, while differences by age, region size, occupation, and education were statistically significant.

In the listening, listening and speaking, and speaking domains, statistically significant differences in proficiency scores were found depending on in-depth variables such as the frequency of face-to-face conversation and the frequency of conversation through social networking services or messaging platforms. The results showed that proficiency scores in the listening, listening and speaking, and speaking domains were higher when the frequency of face-to-face conversation was greater and when the frequency of conversation through social networking services or messaging platforms was higher. In the reading domain, significant differences were observed based on domain-specific variables such as reading frequency, amount of reading per session, and perceptions of the value of reading. Reading proficiency tended to be higher among individuals who read more frequently and who read larger amounts per session. In the grammar and norms domain, significant differences were observed based on the frequency of searching for information related to grammar norms and the frequency of using the National Institute of Korean Language's Online-Ganada service, although the proportion of respondents who reported using these resources "very frequently" was notably low. In the writing domain, writing proficiency tended to be higher among individuals with a college education or above and among those who engaged more frequently in non-voluntary(public) writing. Conversely, proficiency tended to be lower among individuals with less than a high school education. Writing proficiency was also higher among those living in large cities and among individuals who perceived writing as necessary for self-understanding and for expressing thoughts and emotions.

This study also compared the proportion of correct responses for the anchor items shared between the first and second assessments and between the second and third assessments, and estimated effect sizes to identify general trends in Korean language proficiency over time. A comparison across the first(2013), second(2018), and third(2023-2025) National Assessment of Korean Language Ability shows that, although overall proficiency appears to have increased slightly over the past decade, the changes are not substantial when evaluated using effect sizes. Between the first and second assessments, listening, reading, writing, and grammar showed little to no change, with only a small increase observed in speaking. Between the second and third assessments as well, effect sizes for most

domains did not reach the threshold for a small effect, making it difficult to interpret the differences as meaningful. Despite some adjustments to domain composition and assessment procedures in the third cycle, the overall level of national Korean language proficiency has remained largely stable, with only modest potential improvements suggested in writing and speaking. Taken together, these findings indicate that national Korean language proficiency has remained generally steady across the three cycles, with only limited domain-specific variation.

Based on these analyses, this study presented domain-specific implications and policy recommendations for the use of survey findings in Korean language policy. For the listening, listening and speaking, and speaking domains, the results showed that the proportion of adults at Level 4 decreased and the proportion at Level 1 increased with age, and that proficiency scores varied depending on region size and patterns of media use. These findings highlight the need for policies that support oral communication abilities by accounting for generational differences in language experience, regional disparities in linguistic-cultural resources, rapid changes in digital communication environments, and differences in opportunities for interaction. Given that these gaps cannot be resolved in the short term, the study suggests establishing a data-driven policy feedback system to ensure sustained support. In the reading domain, the decline in the proportion of adults at Level 4 and the increase in the proportion at Level 1 with age, along with differences in reading proficiency according to patterns of media use, point to the need for literacy education policies that address age-related and digital divides. The study further emphasized that, amid technological advancement and a shifting information environment, traditional print-based literacy has expanded to encompass critical, digital, and civic literacy. Accordingly, it recommends evidence-based critical reading instruction, holistic literacy education, and lifelong literacy policies. In the grammar and norms domain, the decrease in the proportion of adults at Level 4 and the relatively high proportion at Level 1 among older adults, as well as substantial differences by education level, indicate a need for policies that support older adults and address educational disparities. The study also recommends improving the accessibility and usability of grammar and norms-related online resources such as the National Institute of Korean Language's dictionary, orthographic guidelines, and the Online-Ganada service through user-centered redesign. In the writing domain, growing polarization in writing proficiency—particularly across generations, regions, and education levels—indicates a need for learning environments that help older adults maintain writing practices beyond early adulthood, alongside expanded access to higher education and foundational literacy programs. The study also recommends developing practical instructional models tailored to

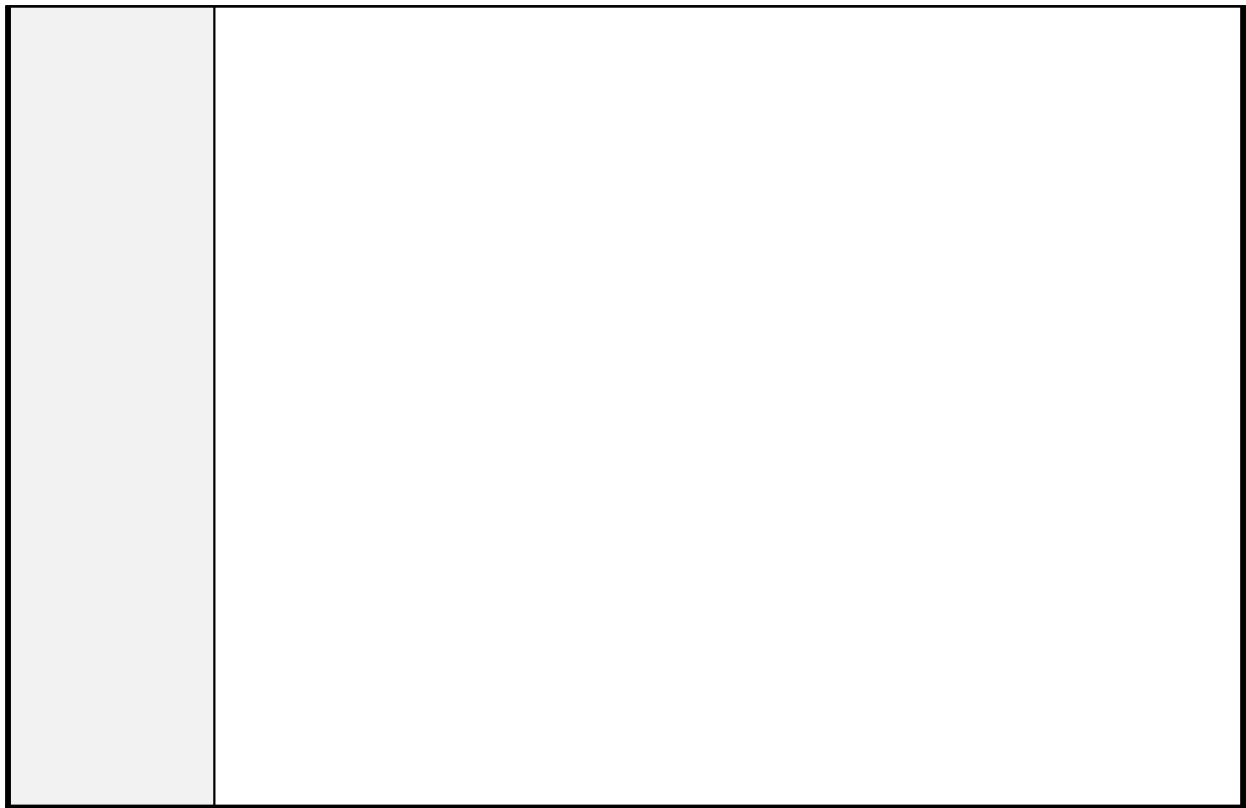
different occupational and regional groups, promoting writing education that enhances not only writing skills but also writing awareness and attitudes, and cultivating a broader writing culture. Finally, the study notes the need for further systematic discussion of writing proficiency in digital media contexts.

Key-words: Korean language proficiency, Korean language proficiency assessment, national Korean language proficiency, listening ability, speaking ability, listening and speaking ability, reading ability, grammar and norms ability, writing ability.

[붙임 1] 본조사 문항 최종본: 쓰기 영역

2025 국민의 국어능력 실태 조사		W-1	문항 정보		설명		
평가 문항							
자료 출처							
평가 요소 텍스트 정보	설명적 글쓰기 - 묘사글						
출제 의도	해당 문항은 설명적 글쓰기 능력을 평가하기 위한 문항으로 난이도는 중으로 설정하였다. 화제에 대한 정보를 독자가 이해할 수 있도록, 구체적이고 객관적으로 상술하면서 300자 내외의 완결된 글을 작성할 수 있는지를 평가하고자 하였다.						
해설	평가 항목	채점 요소(2018 채점 요소 일부 정교화)	채점 척도				
			매우 미흡	미흡	보통	우수	매우 우수
	내용: 40점 (50%)	- 설명 대상의 전체와 부분을 잘 이해할 수 있는 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	8	16	24	32	40
	조직: 24점 (30%)	- 대상에 대한 독자의 이해를 도울 수 있는 내용 전개 방식(묘사, 비교, 대조, 분석 등)을 사용하였다. - 내용의 배열이 체계적(설명의 방향성을 가지고 차례대로 설명)이며, 응집성과 완결성을 갖추었다.	5	10	14	19	24
	표현: 16점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 객관적이고 명확한 표현을 사용하였다.	3	6	10	13	16

2025 국민의 국어능력 실태 조사		W-2	문항 정보		논증		
평가 문항							
자료 출처							
평가 요소 텍스트 정보	논증적 글쓰기 - 논설문 쓰기						
출제 의도	해당 문항은 논증적 글쓰기 능력을 평가하기 위한 문항으로 난이도는 상으로 설정하였다. 자신의 주장을 분명히 제시하고 타당한 이유와 근거를 들 수 있는지, 400자에 해당하는 한 편의 글을 여러 문단으로 구성하여 응집성 및 완결성 있게 작성할 수 있는지, 독자를 고려한 효과적인 표현을 통해 글의 설득력을 높이고 있는지 등을 평가하고자 하였다.						
정답 및 해설	평가 항목	채점 요소	채점 척도				
			매우 미흡	미흡	보통	우수	매우 우수
	내용: 65점 (50%)	- 주장을 분명히 제시하고, 타당하고 충분한 이유나 근거를 들었다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	13	26	39	52	65
	조직: 39점 (30%)	- 서론-본론-결론의 형식을 갖추고, 문단을 적절하게 구분하였다. - 설득력을 높일 수 있는 내용 전개 방식(문제-해결, 인과, 유추, 분석 등)을 사용하였다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	8	16	23	31	39
	표현: 26점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려한 효과적인 표현을 사용하였다.	5	10	16	21	26



2025 국민의 국어능력 실태 조사		W-3		문항 정보		친교	
평가 문항							
자료 출처							
평가 요소 텍스트 정보	친교적 글쓰기 - 사과문						
출제 의도	해당 문항은 생활과 밀접하게 관련된 쓰기 상황에서의 친교적 글쓰기 능력(사과하는 글쓰기)을 평가하기 위한 문항으로, 난이도는 하 수준으로 설정하였다. 주제와 목적, 상황을 고려하여 3문장 이상의 짧은 글을 작성할 수 있는지 평가하고자 하였다.						
정답 및 해설							
	평가 항목	채점 요소	채점 척도				
			매우 미흡	미흡	보통	우수	매우 우수

내용: 20점 (50%)	- 주제와 목적(잘못된 공지에 대한 사과)에 부합하고, 상황(동호회 회원을 대상으로 한 공적인 게시글 작성)에 알맞은 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	4	8	12	16	20
	조직: 12점 (30%) - '문제 상황(사과 내용) 설명과 사과 표현 - 책임의 인정이나 문제 발생 원인 제시- 문제 해결이나 태도 개선 의지 표현/문제 해결 방안이나 대안의 제시'의 구조가 잘 갖추어졌다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	2	5	7	10	12
	표현: 8점 (20%) - 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 예의 바르고 효과적인 표현을 활용하였다.	2	3	5	6	8

2025 국민의 국어능력 실태 조사	W-4	문항 정보	디지털
평가 문항			

자료 출처							
평가 요소 텍스트 정보	온라인 게시판 글 쓰기 - 반품 요청하기						
출제 의도	해당 문항은 디지털 환경에서 필자가 사회적 참여 행위로서 능동적인 표현 활동을 수행할 수 있는지를 평가하기 위한 문항으로, 난이도는 중으로 설정하였다. 표현 의도를 고려하여 적절하게 복합양식적 특성을 활용하는 능력이 있는지, 온라인 게시판이라는 소통 공간의 특성과 ‘반품 요청’이라는 목적에 부합하도록 내용을 생성하고 조직하며, 독자를 고려한 표현을 사용할 수 있는지, 자신의 생각을 다양한 기호 체계를 활용하여 표현하였는지 등을 평가하고자 하였다.						
정답 및 해설	평가 항목	채점 요소	채점 척도				
			매우 미흡	미흡	보통	우수	매우 우수
	매체 활용: 10점	- 표현 의도를 고려하여 적절하게 복합양식적 특성(이미지 및 텍스트 항목)을 활용하였다.	2	4	6	8	10

	(20%)						
	내용: 20점 (40%)	- 주제와 목적(불량 상품 반품 요청)에 부합하고, 상황(판매자를 대상으로 한 온라인 게시물 작성)에 알맞은 내용을 선정하였다. - 불필요하거나 벗어난 부분이 없게 통일성을 갖추어 내용을 구성하였다. - 양적인 조건을 충족하였다.	4	8	12	16	20
	조직: 10점 (20%)	- '문제 상황 설명 - 요구 근거 또는 문제로 인한 손실 - 문제 해결 요청(해결 방안)'의 구조가 잘 갖추어졌다. - 내용의 배열이 체계적이며, 응집성과 완결성을 갖추었다.	2	4	6	8	10
	표현: 10점 (20%)	- 문장이 자연스러우며, 어휘 사용이 적절하다. - 맞춤법(띄어쓰기 포함)과 글쓰기 관습을 준수하였다. - 독자를 고려하여 정중하고 명확한 표현을 활용하였다.	2	4	6	8	10

[붙임 2] 본조사 조사원 지침서



1. 조사 연혁
2. 조사 개요
3. 조사 내용
4. 조사 설계

01 조사 연혁

= 18년까지 5년 주기로 국어 전 영역에 대한 조사 실시 → 22년부터 영역별로 구분하여 특정 영역의 예비조사와 본조사가 2년에 걸쳐 이루어지는 방식으로 변경, 25년도까지 영역별로 국어능력 실태조사 진행

연도	내용
2013	2012년 예비조사 실시 후 2013년 1차 국어능력 실태조사 실시 (듣기, 말하기, 읽기, 쓰기, 문법 등 전 영역)
2018	2017년, 예비조사 실시 후 2018년 2차 국어능력 실태조사 실시 (듣기, 말하기, 읽기, 쓰기, 문법 등 전 영역)
2022	[예비조사] 듣기, 말하기 영역 예비조사 실시(수도권 대상)
2023	[본조사] 듣기, 말하기 영역 본조사 실시(전국 대상) [예비조사] 읽기, 문법 영역 예비조사 실시(수도권 대상)
2024	[본조사] 읽기, 문법 영역 본조사 실시(전국 5,000명 대상) [예비조사] 쓰기 영역 예비조사 실시(수도권 300명 대상)
2025	[본조사] 쓰기 영역 본조사 실시(전국 3,000명 대상)

| 3

02 조사 개요

모집단	• 수도권 거주 만 나이 20세 이상 69세 이하 일반 국민
조사 지역	• 전국 17개 시도
표본 크기	• 3,000명
표집방법	• 크기비례확률추출법을 통한 집락(읍면동) 추출
조사방법	• 태블릿PC를 이용한 대인면접조사(TAPI: Tablet PC Aided Personal Interview)
조사도구	• 태블릿PC 인터넷(보조적으로 종이 설문지 활용)
조사 일정	• 2025년 8월 8일까지
주관기관	• 국립국어원 / 서울대학교 국어교육연구소

| 4

03 조사 내용

구분	내용		
	영역	문항 수	제한 시간
 문항 구성	쓰기	4문항	40분
	응답자 배경 질문	8문항	-
진행 방법	1) 직접 자판을 입력하여 답안을 작성하도록 하되, 부득이한 경우(예컨대 고령자) 자필로 작성하고 이를 조사원이 자판으로 태블릿에 입력 이 경우 반드시 '조사원 메모'에 기입 2) 디지털 매체 글쓰기 문항의 경우(4번) 응답자가 직접 태블릿을 터치하거나 입력하여 해결하도록 하고 대리 입력은 하지 말 것. 다만 화면 구성(예컨대 클릭 버튼 등)에 대한 간단한 설명은 응답자에게 해도 무방함 3) 쓰기는 태블릿 문항에 제시된 응답시간 및 분량에 대한 고지 필수. 분량이 각각 300자와 400자인 문항의 경우 이에 미달한 짧은 글로 답안 작성을 완료하는 경우가 없도록 시간을 충분히 사용하여 구성과 입력을 할 수 있도록 안내 4) 문항은 난이도별로 쉬운 문항에서 어려운 문항 순으로 배열되어 있으므로 가능한 순서대로 해결할 수 있도록 진행		

| 5

04 조사 설계

01	총 표본수 3,000명
02	전국 300개 읍/면/동 추출
03	각 읍면동에서 10명의 응답자 추출 (만20~69세 남녀)
04	지역, 성별, 연령대, 직업, 학력을 고르게 조사

※ 성별×연령 10세별 + 직업 + 교육수준 할당이 있으므로, 조사 진행 과정에서 슈퍼바이저와 긴밀한 소통 필요

지역	서울, 인천, 경기(동부, 읍면부)
성별	남자, 여자
연령대	20대~60대, 10세 간격
직업	1) 정신노동 : 관리자, 전문가 및 관련 종사자, 사무 종사자 2) 육체노동 : 서비스 종사자, 판매 종사자, 농림 어업 숙련 종사자, 기능원 및 관련 기능 종사자, 장차·기계 및 조립 종사자, 단순 노무 종사자, 군인 등 3) 기타 : 주부, 학생, 무직
교육수준	중졸 이하, 고졸, 대졸 이상

| 6

제 2 장

조사 도구 준비 및 조사 진행

1. 조사 준비물
2. IRB 조사참여 동의 절차
3. 설문 진행을 위해 기억할 사항
4. 설문 진행 방법

Copyright © 2025 Gallup Korea. All rights reserved.

01 조사 준비물

= 조사원은 다음의 준비물을 반드시 지참하고 조사를 실시

조사용 태블릿 PC 및 키보드	질문지	응답자용 IRB 동의서
	 태블릿PC에 거부감을 느끼는 경우만 사용 *종이로 진행한 경우 조사표 제출 필수	
조사원 지침서	조사원증/협조 공문	답례품
		

02 IRB 조사참여 동의 절차

- 응답자가 IRB 동의서 설명문을 충분히 숙지 후 서명할 수 있도록 안내
- 조사 시작 전 IRB 동의서를 먼저 받아야 하며, 동의 거부 시 조사에 참여할 수 없음



| 9

03 설문 진행을 위해 기억할 사항

“

이것은 반드시 기억해야 합니다!!

”



정답을 잘 맞추는 것보다 객관적인 실력을 파악하는 것이 중요

| 10

04 설문 진행 방법

※ 쿼터 체크 : 직업/학력

- = 직업은 정신 노동/육체 노동/기타의 세 가지 유형, 학력은 중졸 이하, 고졸, 대학 재학 이상의 세 가지 유형으로 나뉘어 있습니다.
- = 각 쿼터 유형은 설문지 표지의 직업과 학력 문항에 표시된 알은 줄로 구분됩니다.

직업	① 관리자	② 전문가 및 관련 종사자	③ 사무 종사자	정신 노동
	④ 서비스 종사자	⑤ 판매 종사자	⑥ 농림어업 종사자	
	⑦ 기능원 및 관련 기능 종사자	⑧ 장치, 기계조작 및 조립 종사자		육체 노동
	⑨ 단순 노동 종사자	⑩ 직업 군인		기타
	⑪ 주부	⑫ 무직/퇴직/기타	⑬ 학생	

학력	① 무학	② 초등학교(국민학교) 졸업	③ 중학교 중퇴	중졸 이하
	④ 중학교 졸업	⑤ 고등학교 중퇴	⑥ 고등학교 재학	
	⑦ 고등학교 졸업			고졸
	⑧ 대학교 재학	⑨ 대학교 졸업	⑩ 대학원 재학	대학 이상
		⑪ 대학원 졸업		

11

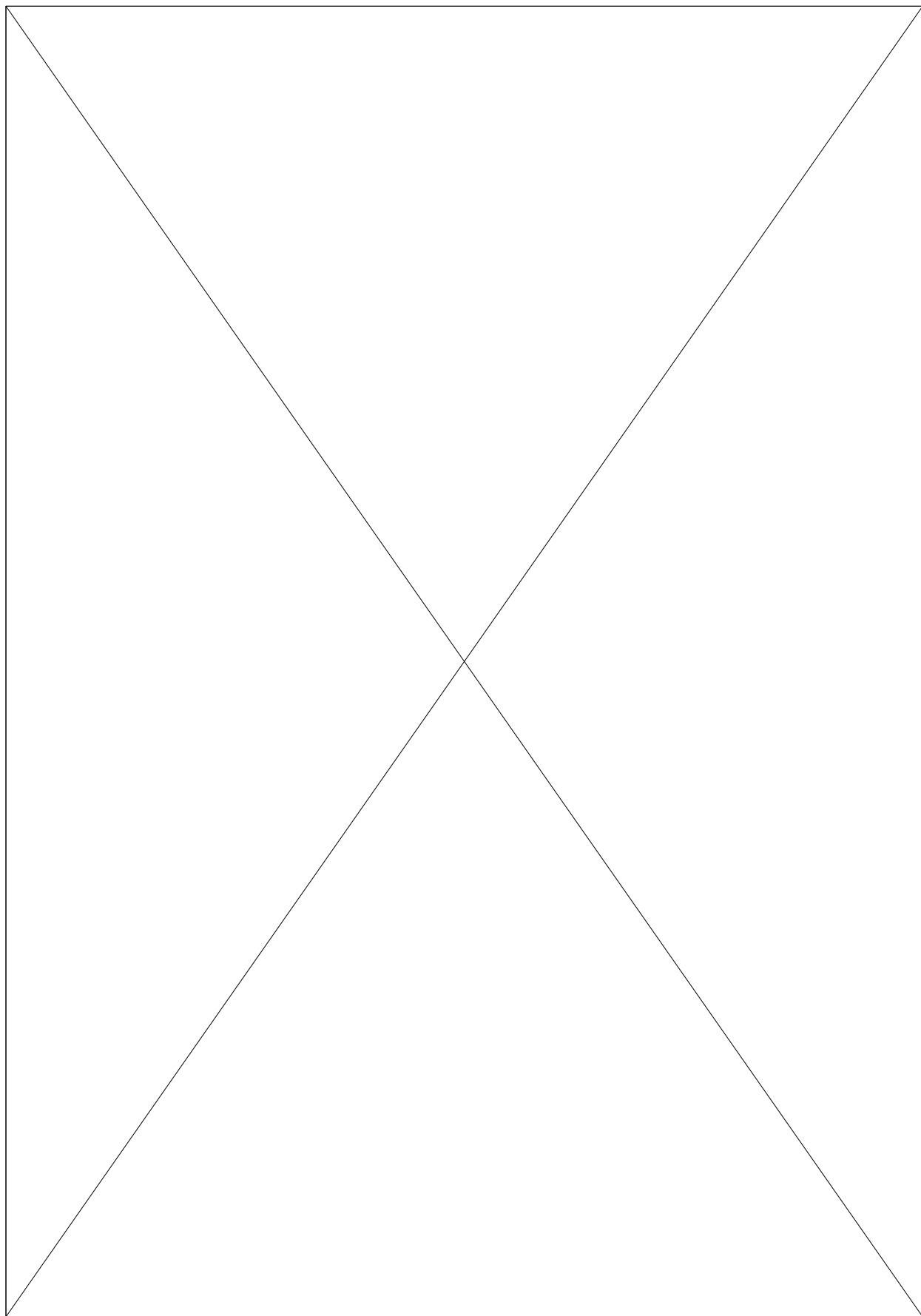
04 설문 진행 방법

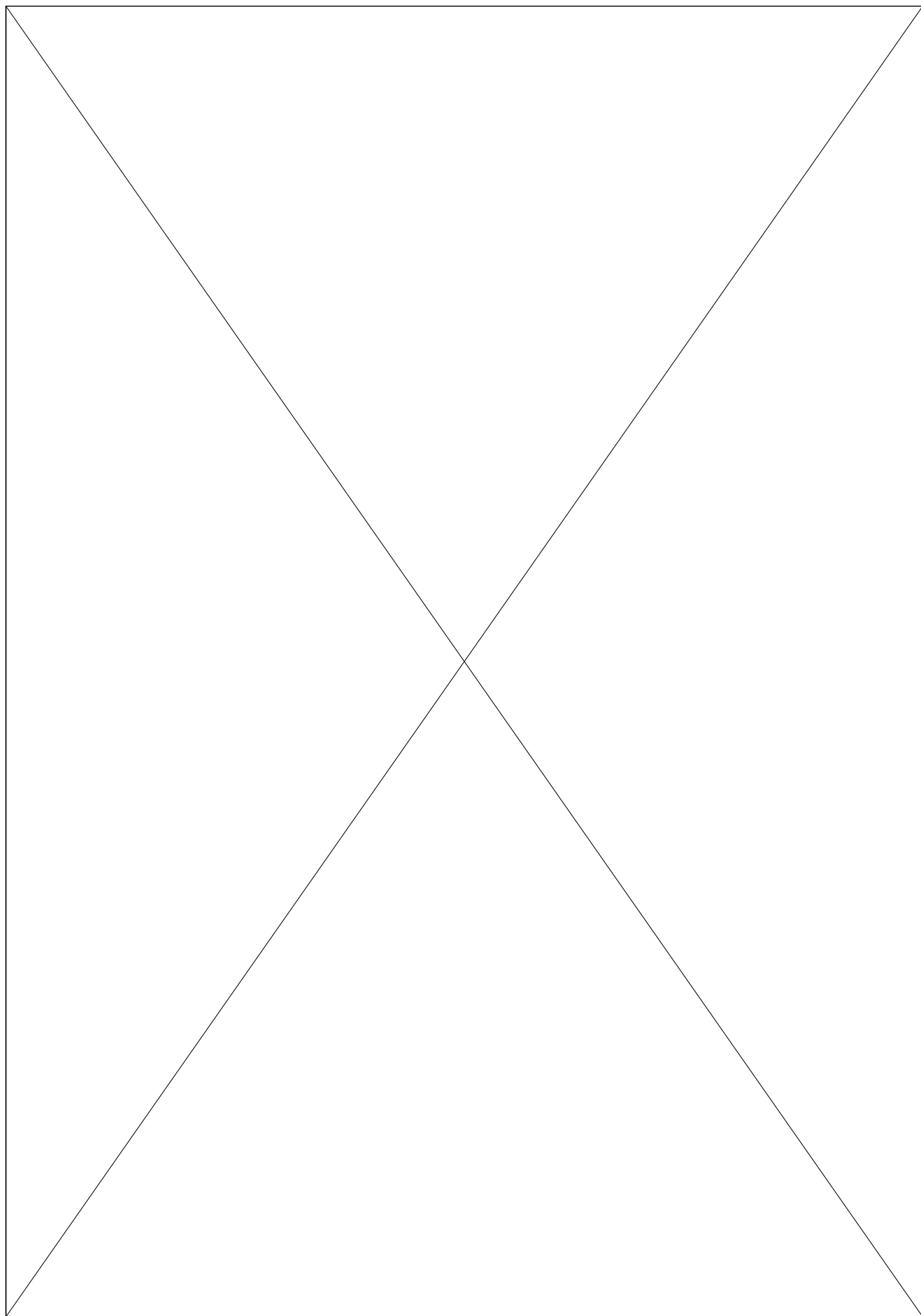
※ 한국표준직업분류 활용

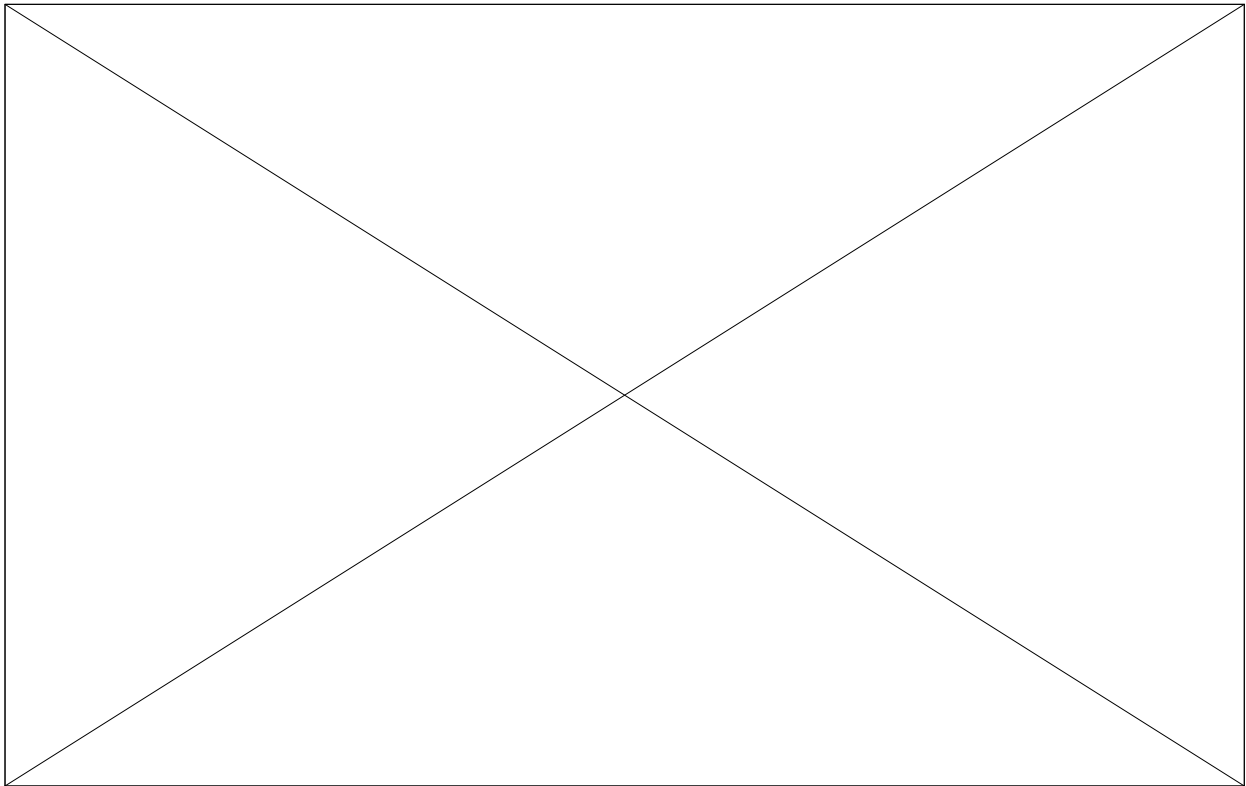
- = 직업 구분은 통계청의 한국표준직업분류 기준에 따릅니다. (자세한 분류 기준은 한국표준직업분류 총표도 참고)

직업분류	설명	대표 직업
관리자	정부, 기업, 단체 또는 그 내부 부서의 정책과 활동을 기획, 지휘 및 조정하는 직무를 수행	고위 공무원, 기업체 임원 및 부서장, 기업체 부서 관리자(총무, 인사, 재무, 마케팅 부서 등), 학교 교장/교감, 유치원 원장, 병원 원장
전문가 및 관련 종사자	특정 분야의 전문지식과 경험을 바탕으로 개념과 이론을 이용하여 해당 분야에 대한 연구·개발, 자문, 지도(교수) 등 전문 서비스를 제공하는 업무 수행	의사, 약사, 간호사, 간호조무사, 치위생사, 영양사, 물리치료사, 안경사, 교수, 교사, 강사, 관사, 변호사, 세무사, 회계사, 자산 운용가, 부동산 중개인, 연구원, 프로그래머, 시스템 개발자, 디자이너, 건축가, 항공조종사, 사회복지사, 작가, 기자, 아나운서, 배우, 가수, 성악자
사무 종사자	관리자, 전문가 및 관련 종사자를 보조하여 경영방침에 의해 사업계획을 입안하고 계획에 따라 업무추진, 문서처리가 주 직무	일반 공무원, 사무원(경영/인사 및 교육 훈련/지재 관리/운송/회계관리 사무원 등), 비서, 고객 상담 및 모니터 요원, 전화상담원
서비스 종사자	조리 및 음식 관련, 보안 관련, 돌봄 및 보건·복지 관련, 이·미용, 훈련·정리 등 서비스를 제공하는 업무 수행	경찰관, 소방관, 주차단속원, 돌봄서비스 종사자, 요양보호사, 간병인, 미용사, 승무원, 조리사, 웨이터, 음식서비스 종사자, 장례 상담원
판매 종사자	영업활동을 통해 상품이나 서비스를 판매하거나 매장에서 계산하는 업무 등을 수행	소규모 상점 경영자, 영업사원, 보험중개인, 매장 계산원, 노점 판매원, 휴대폰 단말기 판매원, 상품/지중자 대여원, 행사(만화)도우미
농림·어업 숙련 종사자	자기 계획과 판단에 따라 작물을 재배·수확하고 동물을 번식·사육(단, 숙련종사자를 도와 농사일/임업 및 어업과 관련한 단순한 일을 수행하는 사람은 단순노동 종사자로 분류)	곡식/채소 등 작물 재배원, 조경원, 낙농업 관련 종사원, 가축 사육 종사원, 동물원 사육사, 벌목원, 조림원 및 산림경영원, 어업 관련 양식원, 어부 및 해녀
기능원 및 관련 기능 종사자	광업, 제조업, 건설업 분야에서 관련된 지식과 기술을 응용하여 금속을 성형하고 각종 기계류 설치 및 정비, 설유, 수공예 제품과 도자, 금속 및 기타 제품을 가공	제철사, 김치(밀반찬) 제조 종사원, 제단사, 제분사, 의복 수선원, 가구 제조 및 수리원, 용접원, 정비원(자동차, 냉장냉동장비, 가전제품 등), 전기공, 건설 관련 기능 종사자(목공, 미장공, 도배공 등), 배관공
장치·기계 조작 및 조립 종사자	기계를 조작하여 제품을 생산, 고도의 자동화된 산업용 기계 및 장비를 조작하고 부분품을 가지고 제품을 조립하는 업무, 운송장비의 운전업무도 포함	기계 조작원, 자동차(무용)/전기부품 조립원, 기관사, 버스 운전사, 택시 운전사, 트럭/특수차 운전사, 대리운전자, 선박 승무원 등
단순노동 종사자	주로 수공구의 사용과 단순하고 일상적이며, 어떤 경우에는 상당한 육체적 노력이 요구되고, 거의 제한된 형의와 판단인을 필요로 하는 업무를 수행	건설 단순 종사원, 하역 및 적재 단순 종사원, 이삿짐 운반원, 택배원, 음식배달원, 상품포장원, 청소원, 아파트 경비원, 기사육아 도우미, 주방보조원, 주유원, 가스 점검원, 주차안내원, 농림어업 단순 종사자
군인	직업군인	영관급(소령, 중령, 대령) 이상 장교, 위관급 장교(소위, 중위, 대위), 부사관, 기타 군인
기타	주부, 무직/퇴직/기타, 학생	

12







04 설문 진행 방법 : 쓰기

쓰기 영역(4문항)

- 모든 문항은 하나의 답만 있음. 종이로 진행하는 경우 2개의 보기에 표시하거나, 문항을 건너뛰는 일이 없도록 함
- 태블릿 자체 키보드나 블루투스 무선 키보드를 통한 입력(타이핑)을 우선으로 함(부득이한 경우 응답자가 종이에 자필로 작성 후 조사원이 태블릿PC에 입력)
- 종이에 쓴 내용이나 타이핑 과정에서 **오탈자나 어법에 맞지 않은 문장이 있을 경우에도 교정을 하지 않고 그대로 입력**
- 4번 문항(디지털 문항)은 응답자가 직접 태블릿PC를 터치하거나 입력하며, 대리 입력은 불가(단, 조사원이 클릭 버튼 등 간단한 화면 구성에 대한 설명은 가능)
- 사전에 응답시간 및 분량에 대한 고지 필수, 문항별 분량(300자, 400자, 1문장, 3문장 등)이 정해져 있음, **분량을 최대한 지켜 답안을 작성하도록 안내**
- 주어진 시간을 초과하여 응답한 경우 조사원 메모에 반드시 기입
- 종이 조사표로 조사한 경우, 응답 내용을 TAPI에 입력하고 종이 조사표는 **실사관리자에게 제출 필요**



18

제 3 장

상황별 대응 요령

1. 상황별 대응 (1)
2. 상황별 대응 (2)
3. 상황별 대응 (3)

Copyright © 2025 Gallup Korea. All rights reserved.

01 상황별 대응 (1)

상황 01 조사의 취지에 대해 물어볼 때

- 이 조사는 단순히 시험을 보는 것이 아니라 현재 국민의 국어능력 실태를 파악하는 조사의 한 방법이라고 설명합니다.
- 조사기관과 목적을 이해시키기 위해 공문을 최대한 활용하고 정부 차원의 공신력 있는 조사라는 점을 강조합니다.
- 개인의 성적이 공개되는 것이 아니고, 연구 목적으로 통계 처리가 되므로 편하게 임할 수 있도록 당부합니다.



내가 학교 졸업한지가 몇 년인데 또 시험을 봐야 하죠?
저 국어 못하는데 제 점수로 망신주려고요?

시험이 아니라 조사의 한 가지 형태입니다.
정부 차원의 실태 조사이며, 개인의 점수는 절대 공개되지 않습니다.



02 상황별 대응 (2)

상황 02 응답자가 평가 도중에 그만두려 하거나 어려운 문항을 건너뛰려고 할 때

- 평가 시간이 길어지고, 응답자가 긴 시간 자신의 생각을 표현해야 할 경우 평가를 꺼리게 될 가능성도 있습니다.
- 문제를 제대로 읽지 않고 답을 하거나 일부 문항을 건너뛰는 일은 없도록 합니다.
- 평가를 중도에 그만두면 이제까지의 응답도 모두 무효가 되므로, 조금만 더 시간을 내어 협조해줄 것을 당부합니다.



이렇게 길게 쓰라고요? 너무 귀찮은데.. 이걸 안 할게요.
이 문항은 너무 어려워서 하기 싫어요.

중간에 평가가 중단되면 이제까지 응답하신 내용이 무의미 합니다.
조금만 힘을 내셔서 평가를 마쳐주세요.



21

03 상황별 대응 (3)

상황 03 응답자가 외부 상황에 의해 조사를 일시적으로 중단해야 할 때

- 평가 도중 다른 사람으로부터 연락이 오거나 손님이 방문하여 일시적으로 조사가 중단되어야 할 수도 있습니다.
- 조사를 일시적으로 멈추고, 중단 시점의 시간을 확인하여 조사를 재개할 때 시간 배분이 잘못되는 일이 없도록 합니다.



잠시 중요한 전화를 받아야 해서요..
밖에 초인종 소리가 나는데요, 잠시만요.

네 알겠습니다. 잠시 후에 진행할게요.
(다시 돌아오면) 조금 전에 2번까지 마쳐셨고요, 남은 시간은 25분 입니다.



22

<기획·연구>

국립국어원 김미현 학예연구사
국립국어원 안진 연구원

<연구 참여자>

연구책임자 구본관(서울대학교 국어교육과 교수,
서울대학교 국어교육연구소 겸무연구원)

공동연구원 민병곤(서울대학교 국어교육과 교수,
서울대학교 국어교육연구소 겸무연구원)

김혜정(경북대학교 국어교육과 교수)

김승현(홍익대학교 국어교육과 교수)

조진수(경인교육대학교 국어교육과 교수)

박성석(춘천교육대학교 국어교육과 교수)

최소영(부산대학교 국어교육과 교수)

장지혜(춘천교육대학교 국어교육과 교수)

이성준(성균관대학교 학부대학 초빙교수)

장호원(㈜한국궤립조사연구소 부장)

백형근(㈜한국궤립조사연구소 부장)

연구보조원 이상재(서울대학교 국어교육과 대학원 박사 과정)

허모아(서울대학교 국어교육과 대학원 박사 과정)

안이진(서울대학교 국어교육과 대학원 박사 과정)

이정민(서울대학교 국어교육과 대학원 박사 과정)

이진현(서울대학교 국어교육과 대학원 박사 과정)

지서영(서울대학교 국어교육과 대학원 박사 과정)

이혜린(서울대학교 국어교육연구소 연구실장)

김고은(㈜한국궤립조사연구소 차장)

조사 수행 (주)한국궤립조사연구소

발행인: 국립국어원장

발행처: 국립국어원

서울시 강서구 금남화로 154

전화 02-2669-9775, 전송 02-2669-9727

인쇄일: 2025년 11월 28일

발행일: 2025년 11월 28일

인 쇄: 가람문화사

※ 이 보고서는 국립국어원의 용역비로 수행한 ‘2025년 국민의 국어능력 실태 조사’
사업의 결과물을 발간한 것입니다.