

2018년 국가수준 성인의 쓰기 능력 평가 결과 분석

장지혜 서울대학교 국어교육과 박사 수료(제1저자)

서보영 서울대학교 국어교육과 박사 후 연구원(제2저자)

김종철 서울대학교 국어교육과 교수(교신저자)

- I. 머리말
- II. 이론적 배경
- III. 연구 방법
- IV. 연구 결과
- V. 논의 및 제언

I. 머리말

평균적으로 인간이 하루 동안 듣고, 말하고, 읽고, 쓰며 사용하는 단어가 10만 개 정도라고 한다. 하루 동안 호흡하는 횟수는 2만 3천여 번, 맥박이 뛰는 횟수는 10만 번 정도라고 할 때, 언어는 인간의 삶에 있어 호흡이나 심장의 고동만큼이나 큰 비중을 차지하고 있다고 할 수 있다. 굳이 언어의 사용 빈도를 맥박의 빈도와 비교하지 않더라도, 대화의 기술, 설득의 전략, 비판적 읽기 등 언어 능력이 개인이 성공적인 삶을 영위하는 데 있어 핵심적인 역량임은 재론의 여지가 없다. 사회적인 차원에서도 탈권위주의 시대로 향해갈 수록 국어 능력에 기초한 구성원들의 소통의 중요성은 더욱 커지고 있으며, 이러한 측면에서 국민의 국어 능력을 진단하고 이를 신장하기 위한 정책을 마련하는 것은 긴요한 문제라고 할 수 있다.

‘국민의 국어 능력 평가(NAKLA, National Assessment of Korean Language Ability)’는 국어기본법과 그 시행령에 근거하여 국민의 국어능력 수준을 진단하고 국어정책 수립을 위한 기초 자료 확보를 위해 국립국어원에서 5년 주기로 실시하는 성인 대상의 대단위 조사 연구이다. 2013년 듣기,

말하기, 읽기, 쓰기, 문법 다섯 영역에 대해 원년 조사가 실시된 바 있으며(김종철·민병곤·박현정·정희창·김봉순·박재현 외, 2013), 해당 연구는 일반 국민을 대상으로 대규모 표본 집단을 선정하여 조사하였다는 점과, 국민의 국어능력 수준을 절대적 등급 지표로 환산하여 구체적인 해석 결과를 제시하였다는 점에서 의의를 지닌다. 5년을 주기로 하여 2018년 두 번째 조사가 실시되었다(민병곤·김종철·구본관·손원숙·김혜정·박재현 외, 2018). 2013년에는 20세에서 59세까지의 성인 3,000명을 대상으로 다섯 영역을 조사하였다면, 2018년에는 2013년 연구와의 연속성을 고려하되, 20세에서 69세의 성인 5,500명을 대상으로 하고, 기존 다섯 영역에 디지털 영역의 조사를 추가하여 조사 대상, 규모, 영역을 확대하였다.

본 연구는 민병곤 외(2018)의 연구 중 쓰기 영역의 평가 도구 및 평가 결과를 제시하고, 우리나라 성인의 쓰기 능력 실태와 특성에 대한 해석을 통해 성인의 쓰기 능력 신장을 위한 방안과 후속 연구의 방향에 대해 논의하고자 한다.

쓰기는 고차원적인 문제 해결 능력을 요하는 언어 활동으로, 고려시대의 과거시험에서부터 현대의 대입 논술 고사, 학위 논문을 비롯한 학술적 글 쓰기에 이르기까지 사회적으로 그 가치와 권위가 인정되어 온 국어 능력의 정수(精髓)이다. 뿐만 아니라 인스턴트 메신저를 활용한 메시지의 송수신이 하루 평균 인당 10~20 건이라는 조사 결과(과학기술정보통신부, 2018: 12), 매체의 소통에 있어 소비자가 아닌 생비자(生費者)로의 역할 전환은 쓰기 능력이 국민의 일상에서 차지하는 중요도가 증가하고 있음을 보여주며, 고차적 문식력에서 일상적 문해력에 이르기까지 쓰기 능력에 대한 면밀한 진단과 이에 대한 처방이 시행되어야 할 필요성을 보여준다.

쓰기 능력 평가의 측면에서 2013년 연구는 그간 국민을 대상으로 한 대규모 평가에서는 시행되기 어려웠던 직접 평가를 시행하였으며, 타당한 쓰기 평가 도구와 척도를 마련하였다는 점에서 의의를 지닌다. 본 연구는 2013년 원년 조사와의 연계성 속에서 성인 3,000명을 대상으로 대규모 직접 평

가를 실시하였다는 점과, 이를 통해 국민의 쓰기 능력 변화 추이를 연속성 있게 살펴보았다는 점에서 유의미하다.

II. 이론적 배경

쓰기 능력 평가에 있어 선결되어야 할 질문은 첫째 ‘왜 평가하는가’, 둘째 ‘무엇을 평가할 것인가’, 셋째 ‘어떻게 평가할 것인가’일 것이다. 첫 번째 질문에 대한 답이 ‘국민의 쓰기 능력 진단과 정책 마련’이라면 두 번째 질문을 해결하기 위해서는 본 연구에서 상정하고 있는 쓰기 능력의 의미를 명확히 한정할 필요가 있다.

쓰기 능력은 복잡한 인지 과정으로 구성되어 있어 개념 규정이 쉽지 않다. 인지적 측면뿐만 아니라, 수사학적 측면, 언어적 측면, 문화적 측면이 모두 포함된 개념이기 때문이다. 논자에 따라 쓰기 능력에 대한 다양한 정의가 존재하고(Bereiter, 1980; 박영목, 1999; Wiegle, 2002; 최미숙·원진숙·정혜승·김봉순·이경화·전은주 외, 2008; 가은아, 2011), 형식주의, 인지주의, 사회구성주의 등 쓰기를 바라보는 관점에 따라서도 쓰기 능력의 정의가 달라질 수 있겠으나, 본고에서는 특정한 이론에 기대지 않고 ‘다양한 작문 맥락에 적합한 글의 내용을 생성하고, 효과적으로 조직, 표현하는 능력’으로 쓰기 능력을 규정하고자 한다. 이는 해당 개념 규정이 ‘어떻게 평가할 것인가’의 문제를 해결하기 위한 가교가 된다는 점에서, 개념에 대한 철학적이고 포괄적인 접근보다 하위 요인을 중심으로 구체적으로 기술하는 방식을 취한 것이다. 쓰기의 정의적 영역을 쓰기 능력으로 보는 관점(조재윤, 2009)도 존재 하나 정의적 영역은 능력이라기보다는 ‘태도’ 범주에 가까운 것으로 보아 쓰기 능력의 요인을 인지적 영역으로 한정하였다. 아울러 자신의 글의 내용과 구조를 점검하고 조정하는 상위인지적 영역은 과정적인 특성을 지닌 까닭에

쓰기 수행의 결과물인 텍스트만으로는 판단하기 어렵다는 점에서 본고의 개념 규정에는 반영하지 않는다.

대규모 쓰기 평가에서 가장 중요한 것은 타당도와 신뢰도를 고려하여 평가 도구와 방법을 마련하는 일이며, 이는 ‘어떻게 평가할 것인가’ 하는 세 번째 문제와 연관된다. 즉, 높은 타당도의 평가 과제 개발, 높은 신뢰도의 평가 척도 및 채점 방식 마련이 쓰기 평가의 유의미성을 결정한다고 할 수 있다. 관련하여 비글(Weigle, 2002)에서도 쓰기 평가에서 가장 중요한 요인으로 ‘평가하고자 하는 쓰기 형태를 명확하게 드러내는 과제, 구인이 포함된 평가 척도(rating scale), 평가 척도를 준수하는 평가자’를 제시한 바 있으며, 본 연구의 평가 도구 설계에 있어서도 이러한 점을 우선적으로 고려하였다.

대학수학능력시험, 국어능력인증시험(ToKL) 등과 같은 선별을 위한 국내 대규모 쓰기 평가의 경우 경제적이고 신뢰성이 높아 간접 평가 방식이 주로 사용된다. 국가 수준 학업 성취도 평가(NAEA)의 경우 간접 평가 방식과 직접 평가 방식을 혼용하고 있으나 답안 길이가 짧고 응답이 제한된 단답형이나 문장 수준의 서술을 요구하고 있어 본질적 의미의 직접 평가로 보기 어렵다. 쓰기 지식이 아닌 쓰기 능력을 측정할 경우 간접 평가의 방식은 그 타당도가 높지 않다(박영목, 1999; Yancey, 1999; Wiegle, 2002; 박영민, 2011; 이수진, 2011; 김종철 외, 2013; 장성민·민병곤, 2016)는 한계를 고려할 때 쓰기 능력의 평가는 직접 평가 방식이 합당하다. 이때 특정한 하나의 수사적 목적을 달성하는 데 요구되는 작문 기능이 다른 수사적 목적을 달성하는 데는 효과적으로 사용될 수 없다(박영목, 1999: 23)는 점에서 총체적인 쓰기 능력의 진단을 위해서는 여러 장르의 쓰기 과제를 개발해야 한다는 점도 인식해야 한다.

다음으로, 쓰기 평가 척도는 평가 목표 및 평가하고자 하는 능력과 밀접한 관련이 있다. 쓰기 평가 척도로 ‘내용과 사고, 조직, 표현과 문체’를 제시한 Purves(1984) 이후 전통적으로 ‘내용, 조직, 표현’은 글쓰기의 질을 설명하는 주된 준거가 되어왔다(Wolcott & Legg, 1998; 박영목, 1999; 정희모·이

재성, 2009, 김성숙, 2011). 국어능력과 관련한 또다른 국내 대규모 평가인 한국어능력시험(TOPIK) 쓰기 영역 문항의 평가 범주 또한 '내용(내용 및 과제 수행), 조직(글의 전개 구조), 언어 사용(어휘, 문법, 맞춤법)'을 평가 준거로 삼고 있으며, 미국 국가 수준 교육 성취도 평가(NAEP, National Assessment of Educational Progress)의 쓰기 영역에서도 '내용의 깊이와 다양성, 내용 조직의 체계성, 표현의 정확성과 다양성' 등을 평가 준거로 삼는다(남민우, 2011). 이때 중요한 것은 쓰기 평가 준거가 쓰기 목적이나 장르에 따라 상이하게 나타날 수 있다는 점이다(노창수, 1987; 김정자, 1992, 서수현, 2008). 쓰기 맥락이 글의 내용, 조직, 표현에 영향을 미친다는 점에서 친교적인 글, 설명적인 글, 논증적인 글 등 여러 장르의 쓰기 과제 개발 시 각 장르별 특성을 고려하여 평가 척도를 세부적으로 마련해야 한다.

끝으로, 쓰기 채점에서 우선적으로 고려할 것은 평가 방식의 설정이다. 채점 방식은 크게 총체적 평가 방식(holistic scoring)과 분석적 평가 방식(analytic scoring)으로 나눌 수 있는데, 총체적 평가는 하나의 글을 평가할 때 개별적인 평가 요소보다는 글 전체에 초점을 두어 평가하는 방식 (Wolcott & Legg, 1998: 71)이며, 분석적 평가는 단일한 점수 부여 대신 평가 목적에 따른 기준들과 다양한 쓰기의 측면에 따라 평가하는 방식(Weigle, 2002: 114)이다. 효율성의 측면에서 유용하기 때문에 일반적인 평가 상황에서는 총체적 평가가 널리 쓰이고 있으나, 총체적 평가 방식은 집단 내에서 상대적인 비교를 하게 되는 경향이 있기 때문에 객관성이나 신뢰도가 떨어질 수 있다는 것과 응답이나 수행에 대한 이해와 진단이 어렵다는 것이 단점으로 지적된다(Huot, 1990; 지은림, 1999; 한국교육평가학회, 2004). 따라서 신뢰도 확보와 더불어 실제의 쓰기 수행에 대한 진단이 요구되는 대규모 쓰기 평가의 경우에는 채점자의 주관을 최대한 배제할 수 있는 분석적 평가 방식을 택하는 것이 요구된다.

한편, 총체적 평가와 분석적 평가에서 평가자들의 평가 결과가 어떻게 다르게 드러나는지를 제시한 정희모·이재성(2009)의 논의는 염밀한 척

도 마련의 필요성을 제시해 준다는 점에서 주목할 만하다. 해당 연구에 따르면 총체적 평가에서는 평가자들이 글의 주제, 초점, 내용, 조직에 중점을 두는 반면, 분석적 평가에서는 어휘, 문장·문법에 높은 점수를 주는 경향이 있다. 이는 분석적 평가에서는 다른 평가 항목과의 평가 비중을 맞추기 위해 평가 항목 간 가중치가 필요함을 함의하는 것이다. 뿐만 아니라 일반적으로 제2언어 학습자 대상의 글쓰기가 아닌 모어 필자의 글쓰기에서는 표현적 오류가 없는 글 보다는 맵락에 적합한 내용으로 구성된 글을 좋은 글로 판단한다. 비글(Weigle, 2002: 189-192) 또한 제2언어 학습자 대상의 글쓰기에서는 ‘내용·조직(content/organization)’과 ‘언어 사용(language use)’ 두 항목의 점수를 각각 10점씩 설정한 반면, 학문 목적 외국어 학습자, 모어 학습자 대상의 글쓰기에서는 ‘내용-조직-언어 사용’의 점수를 ‘10-10-5’의 비율로 평가하였다. 관련하여 국내 쓰기 평가 관련 연구에서 ‘내용-조직-표현’ 항목에 ‘5-3-2’, ‘7-5-5’ 등 각각의 가중치를 두어 평가한 논의들을 참조할 수 있다(김종철 외, 2013: 74; 박영민·가은아, 2009: 185, 장성민·민병곤, 2016: 174).

그러나 분석적 채점 방식이 언제나 신뢰도를 보장하는 것은 아니다. 김라연(2007)에서는 분석적 평가를 수행했음에도 평가 집단 간(교사 집단, 대학생 집단, 학생 집단) 유의한 상관이 없는 결과가 도출되었는데, 이는 분석적 채점에도 채점자의 사전 협의와 훈련이 필수적임을 말해주는 결과이다. 또한 권태현(2012)에서도 별도의 채점자 훈련이나 협의가 없는 경우 분석적 채점이 총체적 채점보다 채점자 간 신뢰도가 매우 낮을 수 있음을 시사한 바 있다. 이는 채점 방식과 무관하게, 숙련되지 않은 채점 방식이 채점자 신뢰도를 떨어뜨릴 수 있는 것으로 해석할 수 있다. 따라서 대규모 쓰기 평가의 경우 분석적 채점 방식을 채택하되 평가자 간 신뢰도를 높이기 위해 평가자 훈련 및 협의가 필수적임을 인식하고, 이러한 과정을 통해 쓰기 평가 기준과 이를 구체화한 평가 예시문 등을 세밀하게 마련해야 한다는 점을 고려해야 한다.

III. 연구 방법

1. 연구 대상

본 연구에서는 전체 국민이라는 모집단을 대표할 수 있는 표본 집단의 선정이 무엇보다 중요하다. 표본 집단의 여러 가지 인구학적 특성이 모집단의 그것과 일치해야만 ‘국민의 쓰기 능력 진단’이라는 연구 목적이 객관적으로 달성될 수 있기 때문이다. 다만, 표본 설계 과정에서는 표본의 대표성이라는 측면과 함께 조사의 난이도와 조사 비용, 기간 등의 현실적인 여건을 함께 고려해야만 했다. 그 결과 총화 집락 추출과 응답자 할당 추출의 2가지 단계로 표본추출을 진행하였다.

총화추출은 모집단을 서로 겹치지 않은 그룹들, 즉 층(strata)들로 나누고 각 층에서 표본을 뽑는 방법이다. 본 연구에서는 전체 모집단을 시도 단위로 나눈 후, 다시 그 안에서 동지역과 읍면 지역을 구분했다. 각 층 내에서는 크기비례확률추출(Sampling with probabilities proportional to size)을 통해 읍·면·동을 선정하였다. 크기비례확률추출은 규모가 큰 집락일수록 표본으로 선정될 확률이 비례해서 높아지게 된다. 특히 집락들의 규모가 심하게 차이 나거나 대규모 집락이 전체 모집단에 미치는 영향이 클 경우에는 각 집락의 크기에 비례하여 뽑는 방법인 크기비례확률추출법이 널리 사용되고 있다.

추출된 집락(읍·면·동) 중 최종적으로 조사대상자(개인)를 선정하는 과정에서는 할당추출법을 사용했다. 할당추출은 모집단을 구성하는 하부 집단(예를 들어 성별이나 연령대)의 구성비에 일치하도록 표본을 추출하는 방식이다. 본 연구에서는 개인의 국어능력과 밀접한 연관성을 지녔다고 볼 수 있는 ‘성별(남/여), 연령대(20대/30대/40대/50대/60대), 직업(정신노동/육체노동/기타), 학력(고졸 미만/고졸/대재 이상)’을 기준으로 한 할당추출을 진행하

였다. 이미 모집단을 충화하는 단계에서 지역 변인을 고려했기 때문에 표본 추출 시 고려한 변인은 ‘지역, 성별, 연령, 학력, 직업’의 5개로 구성된다. 이렇게 해서 본 조사의 모집단은 전국 17개 시도에 거주하는 만 20세~69세의 남녀 3,000명이었으며, 표본별 분포 현황은 다음과 같다.

〈표 1〉 표본별 분포 현황

단위: 명(%)

성별		연령대					
남	여	20대	30대	40대	50대	60대	
1,521 (50.7)	1,479 (49.3)	553 (18.4)	591 (19.7)	696 (23.2)	692 (23.1)	468 (15.6)	
지역 규모			직업군			학력군	
대도시	중소 도시	읍면 지역	정신 노동	육체 노동	기타	고졸 미만	고졸
1,320 (44.0)	1,158 (38.6)	522 (17.4)	835 (27.8)	1,283 (42.8)	822 (29.4)	435 (14.5)	979 (32.6)
1,586 (52.9)							

2. 평가 도구

평가의 목적이 선별이 아닌 진단이라는 점에서 평가 문항 개발시 괴험자의 ‘실제의 쓰기 능력’을 평가하는 데 중점을 두었다. ‘실제의 쓰기 능력’이란 수사적인 상황에 대한 인식, 글의 목적을 고려한 적합한 내용 및 형식의 선정, 사회적 의사소통을 고려한 표현 등을 함의하고 있는 것으로 사회적 행위로서의 실제성을 의미한다(Bawarshi & Reiff, 2010/2015: 103 - 106). 이에 따라 친교적, 설명적, 논증적 글쓰기를 각각 선정하여 쓰기 능력을 조사하고자 하였다.¹⁾ 각 담화 유형의 구체적인 하위 장르 선정에 있어 공문서, 제안서

1) 쓰기의 담화 유형은 크게 설명적 글쓰기, 설득적 글쓰기, 친교적 글쓰기, 자기 표현의 글쓰기로 나눌 수 있다. 이 중 자기 표현의 글쓰기에는 일기, 자서전 등이 해당하는 것으로 공

등과 같이 전문적 장르 인식을 요하는 담화 유형보다는 보편적으로 전 연령 층이 실제 생활에서 친숙하게 사용하는 담화 유형인가를 고려하였고, 결과적으로 감사의 편지, 설명문, 논설문을 통해 자연스러운 쓰기 능력을 조사하고자 하였다. 특히 감사의 편지의 경우, 현재의 문식 상황에서 자필 편지보다는 이메일, 문자 메시지 등을 쓰는 빈도가 높은 것을 감안하여 감사의 이메일을 보내는 것으로 설정하였다.

이때 문제 상황을 드러내는 언어적 표현에 대한 물이해나, 혹은 평가의 조건에 대한 주의집중 부재가 문제되지 않도록 하였으며, 각 문항에 대한 내용, 조직, 표현의 채점 준거 또한 의사소통의 목적에 기여하는가를 기준으로 마련하였다.

이상의 내용을 반영하여 최종적으로 제시된 문항 및 평가 기준은 다음과 같다. 2번 문항(300자 설명적 글쓰기, 80점), 3번 문항(400자 논증적 글쓰기, 130점)의 경우 다음 주기(5년 후)에 가교 문항으로 활용될 수 있어 여기에서 는 공개하지 않는다.

W-1. 탁구 동호회의 차기 회장으로 선출되어 회원들에게 감사의 인사를 담아 전자우편(이메일)을 보내고자 합니다. 동호회 회장이라고 생각하고 아래 내용의 본문을 완성해 보십시오. (3문장으로 작성. 본문에 기술된 문장은 답안에 중복 기술하지 않음)

탁구사랑 동호회 회원 여러분께
<u>안녕하세요, 차기 회장으로 선출된 ○○○입니다.</u>

동호회 회장 ○○○ 올림

적인 조사 상황에 적합하지 않아 제외하였다.

평가 항목	채점 요소	채점 척도				
		매우 미흡	미흡	보통	우수	매우 우수
내용: 20점 (50%)	-목적(감사 인사)에 알맞고, 상황에 적합한 내용을 선정하였다. -통일성을 갖추어 불필요하거나 벗어난 부분이 없게 내용을 구성하였다. -양적인 조건을 충족했다.	4	8	12	16	20
조직: 12점 (30%)	-감사의 인사 - 취임에 대한 포부 - 끝인사의 구조가 잘 갖추어졌다. -담화 유형의 구조적 특성에 맞게 구성하였다.	2	5	7	10	12
표현: 8점 (20%)	-문장이 적법하며 어휘 사용이 적절하다. -맞춤법(띄어쓰기 포함)을 지킨다. -독자를 고려하여 예의바르고 자연스러운 표현을 활용하였다.	2	3	5	6	8

문항별 난이도와 변별도는 고전 검사 이론에 터하여 분석하였다. 문항 반응 이론을 통해 난이도와 변별도를 분석할 경우, 5점 척도의 등급 반응 자료로 결과가 산출되는 쓰기 영역의 각 문항에 대한 난이도가 각 등급 척도별로 분석되어 개별 문항에 대한 문항 정보를 파악하는 데 어려움이 있다고 보았기 때문이다. 서답식 문항 3개로 이루어진 쓰기 영역의 고전검사이론에 따른 난이도와 변별도는 다음과 같이 분석되었으며, 전체 신뢰도는 0.522로 나타났다.²⁾

- 2) 고전 검사 이론에서 난이도는 평균 점수를 총점으로 나눈 값으로 계산되며, 변별도는 문항 점수와 해당 영역 점수 간의 상관계수로 산출된다. 이때, 각각의 난이도 지표가 의미하는 바를 언어적으로 편별하는 기준으로는 Cangelosi(1990)에서 제시한 기준을, 변별도의 편별 기준으로는 Ebel(1965)에서 제시한 기준을 사용하였다(Ebel: 1965, 성태제, 1998 재인용). 해당 편별 기준은 다음과 같다.

〈난이도 분포에 대한 해석〉

문항 난이도	문항 평가
.25 이하	어려운 문항
.25 ~ .75	적절한 문항

〈표 2〉 쓰기 영역의 문항 분석 결과

문항번호	응답자수 (명)	정답률 (%)	난이도	변별도	해당 문항 제거 시 신뢰도	신뢰도
W-1	3,000	66.67	0.667	0.338	0.528	0.522
W-2	3,000	45.84	0.458	0.426	0.297	
W-3	3,000	44.82	0.448	0.448	0.353	

먼저, 난이도는 0.667에서 0.448까지로 형성되어 모든 문항이 적절한 것으로 나타났으며, 변별도 또한 1번 문항의 경우 변별도가 있고, 2, 3번 문항의 경우 모두 변별도가 높은 문항인 것으로 확인되었다.

다음으로, 각 문항의 하위 요소별 난이도와 변별도 분석 결과는 다음과 같다.

〈표 3〉 쓰기 영역의 하위 요소별 분석 결과

문항번호	하위 요소	사례수	난이도	변별도	해당 문항 제거 시 신뢰도
W-1	내용	3,000	0.707	0.827	0.761
	조직	3,000	0.448	0.798	0.608
	표현	3,000	0.427	0.736	0.776

.75 이상	쉬운 문항
--------	-------

〈변별도 분포에 대한 해석〉

문항 변별도 지수	문항 평가
.40 이상	변별력이 높은 문항
.30 ~ .39	변별력이 있는 문항
.20 ~ .29	변별력이 낮은 문항
.10 ~ .19	변별력이 매우 낮은 문항
.10 미만	변별력이 없는 문항

W-2	내용	3,000	0.619	0.709	0.849
	조직	3,000	0.485	0.790	0.571
	표현	3,000	0.458	0.687	0.689
W-3	내용	3,000	0.638	0.603	0.742
	조직	3,000	0.444	0.611	0.496
	표현	3,000	0.485	0.656	0.587

하위 요소별 난이도는 0.444에서 0.707까지 형성되어 모두 적절한 수준이었으며, 변별도는 0.603에서 0.827까지 형성되어 모든 하위 요소에서 높은 변별력을 가지고 있는 것으로 분석되었다. 이와 같은 결과에 기초해 볼 때, 각 문항별로 설정한 하위 채점 요소들이 조사 대상자의 쓰기 능력을 측정하는데 유의미한 것으로 판단할 수 있다.

3. 연구 절차

1) 자료 수집

본 연구는 피험자의 의견 개진이 아닌 능력 평가라는 특성을 띠고 있고, 피험자 1인당 1시간 이상의 응답 시간이 소요되는 등 조사의 난도가 매우 높은 편이다. 피험자로부터 성실한 응답을 이끌어 내고, 원활한 조사 진행을 위해 피험자를 직접 만나 평가를 진행하는 조사원은 방문 면접조사 경험이 풍부한 전문 조사원 중에서 선발했다. 특히 수도권의 경우에는 2017년과 2018년의 예비 조사에 참여한 조사원을 우선적으로 선발해 본 조사에 다시 투입했다. 조사 방식은 1:1 방문 면접조사로 진행됐으며, 2018년 7월에서 11월에 걸쳐 CAPI(Computer-Aided Personal Interview)³⁾를 통해 자료를

3) CAPI는 면접원이 조사 대상자의 응답을 종이 설문지가 아닌 PC를 통해 수집하는 것으로, 현재 많은 조사에서 CAPI를 통해 자료를 수집하고 있다

수집했다.

본 연구는 변화하는 문식력을 반영하여 컴퓨터를 활용한 작문 상황을 제시하였다. 국제성인역량조사(PIAAC, Programme for the International Assessment of Adult Competencies) 경우에도 컴퓨터 활용 능력에 따라 컴퓨터 평가와 서면 평가 중 한 가지를 응답자가 수행하고 있는데, 사회인구적 요인(연령, 교육수준, 이민배경, 성별)이 고려되었을 때, 서면 평가를 실시한 응답자의 점수와 컴퓨터 평가를 실시한 응답자의 점수에 조직적인 차이(systematic difference)가 없었다(한국직업능력개발원, 2013). 또한 컴퓨터를 활용한 작문은 실제 작문 상황의 실제성을 담보하고 손글씨 쓰기의 번거로움으로 인한 거부감을 해소할 수 있는 방안이 될 수 있다. 따라서 컴퓨터로 쓰기를 기본 평가 방식으로 채택하되, 활용 능력에 따라 손글씨 쓰기를 허용하였다. 손글씨 쓰기의 경우 2013년 본 조사와 동일한 원고지를 활용하였다.

2) 문항 채점

조사 결과 쓰기 영역에서는 3,000개의 응답 데이터를 확보하였고, 3개의 문항이 포함되어 있으므로 총 9,000개의 답안을 대상으로 하여 대규모 채점을 수행하였다.

이러한 방대한 양의 채점에 있어 가장 크게 고려한 것은 채점의 신뢰도를 확보하는 일이었으며, 이에 따라 본 연구에서는 전문성을 갖춘 대규모 채점단을 섭외하였다. 채점단은 쓰기를 전공하고 있는 서울대학교 대학원 박사 과정생 및 석사 학위를 소지한 경력 3년 이상의 국어 교사, 학사 학위를 소지한 경력 5년 이상의 국어 교사 10인으로 구성되었으며 2017년 조사 도구 개발 연구, 2018년 예비 조사에 참여하였던 채점자 중 신뢰도가 높았던 채점자를 우선 선발하여 채점에 대한 안정성과 연속성을 확보하고자 하였다.

채점자 간 신뢰도 제고를 위해 채점 전 채점단 및 평가 전문가를 포함한 워크숍을 실시하였다. 워크숍은 채점자내 훈련과 채점자간 훈련의 2단계로 이루어졌다. 채점자내 훈련 단계에서는 교육 평가 및 채점 관련 학제 전문가

와 채점단이 모여 평가의 목적과 방향 등의 맥락 등에 대한 공동 이해의 과정을 거친 후, 평가 및 채점과 관련된 강의 및 토론을 실시하는 것으로 진행되었다. 채점자내 훈련이 끝난 후 채점자간 훈련에서는 우선 각 문항별로 구체적인 채점 기준 및 채점 척도, 채점 예시문 등을 통해 채점 준거를 명료화하는 것으로 시작하였으며, 각 문항별로 전체 문항 반응의 1%인 30개 이상의 예시 답안을 통해 예비 채점을 실시하였다. 이때의 예시 답안은 2018년 예비 조사의 채점시 변별이 어려웠던 답안 및 본조사에서 기준에 대해 논의 할 만한 문항을 중심으로 선제하여 제공되었다. 채점자간 훈련은 미리 설정한 채점 기준이 다루지 못한 예외적인 반응이나 기준을 발견하여 이를 수정함으로써 적절하고 명료한 채점 준거를 마련할 수 있도록 하였으며, 채점자간 상호 일치성이 확보될 때까지 장시간 반복 수행되었다.

채점의 효율성과 신뢰도를 확보하기 위해, 본 연구에서는 하나의 문항에 대해 여러 명의 채점자들이 교차 채점하게 하여 채점의 신뢰도를 확보하고자 하였으며, 공통채점과 무선채점을 혼합하여 채점을 설계하였다. 공통채점은 각 문항별로 5.5%(165개X3문항)의 답안에 대해 10인의 모든 채점자가 교차 채점하도록 설계한 것이며, 공통채점 대상을 제외한 나머지 답안은 가능한 채점자의 무선조합(random combination)을 만들어 반복 배분하여 해당 영역의 전문가 2인이 교차 채점하는 방식이었다. 문항당 2인의 교차 채점 방식의 타당성과 적합성을 검증하기 위해 교차채점 및 공통채점 방식으로 진행하였다. 뿐만 아니라 채점 기간 중 1회의 중간 점검 및 통계적 분석을 통해 채점자 간 신뢰도 및 채점 경향을 확인하였으며, 온라인 협의를 통해 지속적으로 부족한 지점을 보완하는 방식으로 채점 방식의 타당성을 보완하였다.

쓰기 영역의 3개 문항, 3,000명 조사 대상자에 대한 채점 결과 모든 문항에서 0.8 이상의 높은 상관계수가 확인되어 채점자 간 신뢰도가 매우 높은 것으로 분석되었다.⁴⁾

4) 구체적 신뢰도 자료는 <부록>에 제시하였다.

다음으로, 〈표 4〉는 쓰기 채점에 대한 채점자 간 평균 점수 차이를 비교한 것을 보여준다. 채점자에 따라서 각 문항의 평균 점수가 통계적으로 차이가 있는지를 분석한 결과 문항 1에서는 10명의 채점자 사이에 다소 차이가 있는 것으로 나타났지만 문항 2와 문항 3에서는 채점자 간 점수 차이가 통계적으로 유의미하지 않았다. 문항 1의 경우 이러한 차이는 피조사자의 수가 3,000 명의 대규모 집단이기 때문에 발생한 것으로 보이며, 실제 평균 점수의 차이는 1.4점의 작은 수준이어서 이러한 결과를 신뢰할 수 있는 것으로 분석되었다.

〈표 4〉 쓰기 채점에 대한 채점자 간 평균점수 차이 비교(반복측정ANOVA)

	채점자 평균										F	유의도	사후 검정
	1	2	3	4	5	6	7	8	9	10			
W-1	26.26	26.76	26.25	26.25	25.36	25.67	25.64	25.38	25.98	26.65	6.442	0.012	9<3,4,6 <8,5,7 <2,1,10
W-2	38.84	38.23	38.27	37.75	38.50	38.46	38.90	39.40	36.21	39.12	0.136	0.713	-
W-3	63.73	64.07	60.61	61.33	61.50	60.44	61.35	61.37	60.79	67.73	2.104	0.149	-

3) 등급화

김종철 외(2013)에서는 조사된 점수가 무엇을 의미하는지에 대한 해석을 위해 ‘수정된 앵고프(Modified Angoff) 방식’을 활용하여 분할 점수(cut-score)를 산출하고 국민의 절대적인 국어 능력 수준을 제시한 바 있다. 본 연구에서는 2013 연구와의 연계성을 위해 2013년 결과와의 동백분위 동등화(equipercentile equating)를 시행하여 분할 점수를 산출하였다.⁵⁾ 동등화를 위해 LEGS(Linking with Equivalent Groups or Single Group Design) 프

5) 본 연구의 목적은 우리나라 국민의 쓰기 능력을 평가하고자 함이며 이에 따른 모집단은 2013년과 2018년 모두 우리나라 국민으로 동일하다고 볼 수 있고 각 영역 별 조사 대상자 수 또한 충분하다. 이러한 조사 특성을 고려한다면 점수의 백분위를 기준으로 조사 대상자들의 능력을 비교하는 동백분위 동등화를 적용할 수 있다.

로그램을 이용하였으며, 동백분위 결과는 <부록>에 제시하였다.

이상에 제시된 본 연구의 연구 절차를 정리하면 다음과 같다.

〈표 5〉 본 연구의 연구 절차

주요 과제	세부 내용	기간
기개발된 조사 도구 보완 및 전문가 워크숍	<ul style="list-style-type: none">• 기초 조사(2017)에서 제안한 조사 도구의 문항 내용 및 실행 차원 보완	2018.04. ~ 2018.05.
조사 대상 표본 추출 및 조사 시행 방식 결정	<ul style="list-style-type: none">• '성별, 연령대, 지역 규모, 학력군, 직업군'의 5개 변인 설정• 인구 분포 통계를 활용한 2단계 층화 표본 추출(stratified sampling) 방식을 적용하여 조사 대상의 대표성 확보• 피달된 조사원이 태블릿 피시를 활용하여 1:1 개별 조사 (태블릿 피시 사용이 불가한 경우를 대비하여 종이 답안지 예비)	2018.04. ~ 2018.05
예비 조사 시행	<ul style="list-style-type: none">• 300명을 대상으로 예비 조사 실시• 수도권을 대상으로, 층화 표본 추출 방식으로 조사 대상 선정	2018.05. ~ 2018.06.
예비 조사 결과 분석 및 평가도구 최종 설계	<ul style="list-style-type: none">• 주관식 문항 채점 진행을 통한 문항 보완 및 채점자의 숙련도 제고• 예비 조사 결과를 바탕으로 조사 도구, 조사 시행 방식, 조사 대상의 타당도 확인 및 본 조사 실행 계획 설계• 국어교육 관련 전문가 자문 및 연구진 워크숍을 통한 내용적·실행적 보완	2018.06. ~ 2018.07.
본 조사 시행	<ul style="list-style-type: none">• 3,000명(전체 영역 5,500명)을 대상으로 본 조사 실시• 제주도를 포함한 전국적 분포를 고려하여, 2단계 층화 표본 추출 방식으로 조사 대상 선정	2018.07. ~ 2018.11.
본 조사 결과 채점	<ul style="list-style-type: none">• 주관식 문항 채점을 위한 '평가 항목, 채점 요소, 채점 척도' 보완 및 확정• 채점자 신뢰도 확보를 위한 채점단 워크숍• 채점자 간 신뢰도 확보를 위한 중간 검토 및 평가 항목 간 상관도 분석	2018.11. ~ 2018.12.
본 조사 결과 분석	<ul style="list-style-type: none">• 전체 결과 도출	2018.12.

IV. 연구 결과

1. 전체 결과

쓰기 영역의 분석 대상 3,000명의 평균 점수는 121.61점(250점 만점)이었으며, 최솟값 54점, 최댓값 237.50점을 보였고 표준 편차는 30.89이었다. 쓰기 영역의 평균 점수 121.61점은 수준별 등급으로는 2수준에 해당되어, 조사 대상 국민의 평균적인 쓰기 능력은 '2수준'에 해당하는 것으로 해석될 수 있다

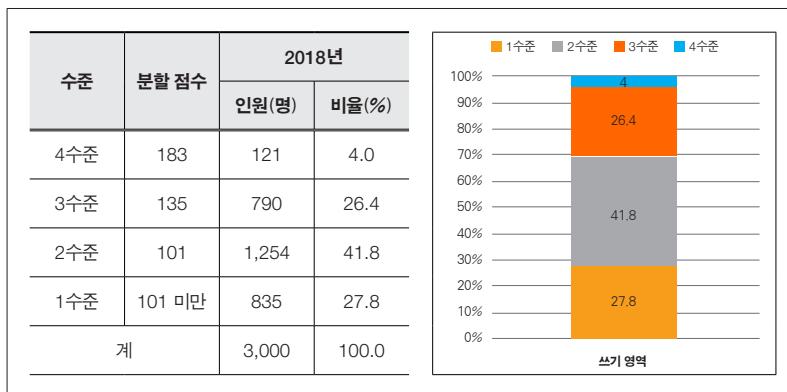
〈표 6〉 쓰기 영역의 기술 통계

사례수	만점	평균	표준편차	최솟값	최댓값
3,000	250	121.61	30.89	54.00	237.50

쓰기 영역의 점수 분포를 분할 점수를 활용한 절대적 수준 등급으로 나누어 보면, 4수준이 121명(4.0%), 3수준이 790명(26.4%), 2수준이 1,254명(41.8%), 1수준이 835명(27.8%)으로 산출되어 전반적으로 국민들의 쓰기 능력이 다른 영역에 비해 저조한 것으로 분석할 수 있다. 특히, 2수준 이하의 비율이 69.6%(2수준: 41.8%, 1수준: 27.8%)나 된다는 점에 주목할 필요가 있다.⁶⁾

6) 국내외 국어(언어) 능력 조사에서 성취 수준을 등급화하고 있다. PIAAC의 경우, 도달 정도에 따라 '1수준 이하-1수준-2수준-3수준-4수준-5수준'의 6수준으로 분할하고 있으며, PISA의 경우 '1b 수준-1a 수준-2수준-3수준-4수준-5수준-6수준'의 7수준으로 분할하고 있다. 국내의 국가수준 학업성취도 평가에서는 '기초 미달-기초-보통-우수'의 4수준으로, 미국의 National Assessment of Educational Progress에서는 '기본(Basic)-숙달(Proficient)-우수(Advanced)'의 3수준으로 분할하기도 한다. 본 연구의

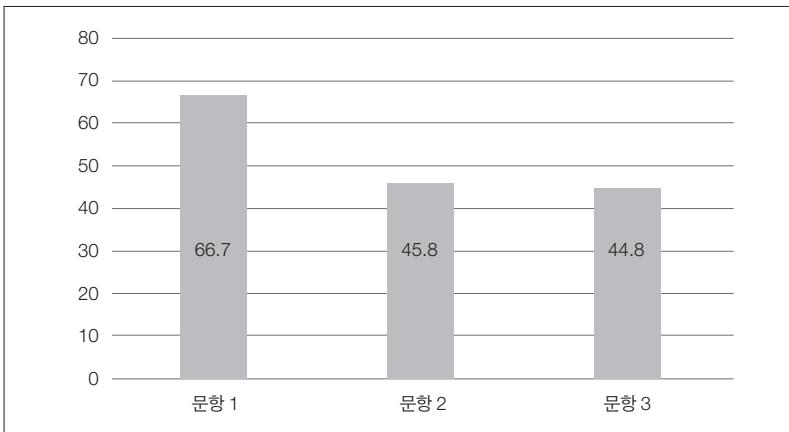
〈표 7〉 쓰기 영역의 수준 등급별 분포



또한, 쓰기 영역에서 조사하였던 3개 문항에 대해 정답률을 비교해 보면 〈그림 1〉과 같은데, 문항 2와 문항 3의 정답률(45.8%, 44.8%)이 문항 1의 정답률(66.7%)에 비해 낮음을 확인할 수 있다. 문항 1의 전자 우편을 통한 친교적 글쓰기는 짧은 문량의 일상적인 글쓰기 장르여서 어려움을 느끼지 않

원년 연구에 해당하는 2013년 조사에서는 많은 문항으로 구성된 조사의 경우 등급 구간이 세분화될수록 타당도 유지가 어렵다는 점에서 ‘우수-보통-기초-기초 미달’의 네 수준으로 등급을 설정하였다. 이는 국가 수준 학업 성취도 평가 등의 대규모 국어능력 평가에서 일반적으로 활용하는 방식이기도 하며, 본 연구는 원년 연구의 설정 방식을 준용하여 네 수준으로 등급을 설정하되, 수준의 명칭을 다소 중립적으로 제시할 필요가 있다고 판단해 ‘4수준-3수준-2수준-1수준’으로 제시하였다. 본 연구에서 설정한 등급별 국어능력의 일반적 수준은 다음과 같다.

등급	국어능력의 수준
4수준	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 대부분(80% 이상) 획득한 수준
3수준	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 상당 부분(50% 이상 80% 미만) 획득한 수준
2수준	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 부분적으로(20% 이상 50% 미만) 획득한 수준
1수준	국민 공통 기본 교육과정 국어 과목을 이수한 국민에게 기대되는 국어능력을 미흡하게(20% 미만) 획득한 수준



〈그림 1〉 쓰기 영역의 문항별 정답률

는 데 비해, 문항 2의 설명적 글쓰기나 문항 3의 논증적 글쓰기의 경우 상대적으로 친숙성이 덜하고, 분량 면에서도 문항 2는 300자 이상의 긴 글, 문항 3은 400자 이상의 긴 글을 요구하고 있어 긴 분량의 글쓰기에 대한 어려움을 느끼고 있음을 확인할 수 있다.

2. 변인별 분석 결과

1) 성별 차이

쓰기 영역의 성별에 따른 평균 점수 차이는 여성이 남성에 비해 평균 점수가 근소하게 높은 것으로 나타났지만, 이는 통계적으로 유의미한 차이는 아니었다.

〈표 8〉 쓰기 영역의 성별 차이

배경변인	평균점수		t	유의도
	남성	여성		
성별	121.07	122.16	-0.970	0.332

〈표 9〉 쓰기 영역 수준 등급의 성별 분포

수준	남성		여성		합계	
	인원(명)	비율(%)	인원(명)	비율(%)	인원(명)	비율(%)
4수준	59	3.9	62	4.2	121	4.0
3수준	394	25.9	396	26.8	790	26.4
2수준	644	42.3	610	41.2	1,254	41.8
1수준	424	27.9	411	27.8	835	27.8
계	1,521	50.7	1,479	49.3	3,000	100.0

2) 연령별 차이

쓰기 영역의 연령대에 따른 평균 점수 차이는 통계적으로 유의한 차이가 있는 것으로 나타났다. 가장 평균 점수가 높은 집단은 20대였으며, 60대가 가장 낮았다.

〈표 10〉 쓰기 영역의 연령대별 차이

배경변인	평균점수					F	유의도	사후비교
	20대	30대	40대	50대	60대			
연령대	127.58	126.47	121.96	118.44	112.56	21.251	0.000	60대<50대<40대<30대,20대

〈표 11〉 쓰기 영역 수준 등급의 연령대별 분포

수준	20대		30대		40대		50대		60대		합계	
	명	%	명	%	명	%	명	%	명	%	명	%
4수준	41	7.4	33	5.6	22	3.2	21	3.0	4	0.8	121	4.0
3수준	170	30.7	189	32.0	184	26.4	169	24.4	78	16.7	790	26.4
2수준	211	38.2	236	39.9	313	45.0	283	40.9	211	45.1	1,254	41.8
1수준	131	23.7	133	22.5	177	25.4	219	31.7	175	37.4	835	27.8
계	553	18.4	591	19.7	696	23.2	692	23.1	468	15.6	3,000	100.0

3) 지역 규모별 차이

지역 규모에 따른 쓰기 영역의 평균 점수는 중소도시가 가장 높았으며, 읍면 지역이 가장 낮은 것으로 나타났다. 이는 통계적으로도 유의미한 차이를 보이는 것으로 확인되었다.

〈표 12〉 쓰기 영역의 지역 규모별 차이

배경변인	평균점수			F	유의도	사후비교
	대도시	중소 도시	읍면 지역			
지역 규모	120.84	123.86	118.56	6.055	0.002	읍면, 대도시 < 중소도시

〈표 13〉 쓰기 영역 수준 등급의 지역 규모별 분포

수준	대도시		중소 도시		읍면 지역		합계	
	명	%	명	%	명	%	명	%
4수준	46	3.5	54	4.7	21	4.0	121	4.0
3수준	352	26.7	327	28.2	111	21.3	790	26.4
2수준	538	40.7	481	41.5	235	45.0	1,254	41.8
1수준	384	29.1	296	25.6	155	29.7	835	27.8

계	1,320	44.0	1,158	38.6	522	17.4	3,000	100.0
---	-------	------	-------	------	-----	------	-------	-------

4) 직업군별 차이

직업군에 따른 쓰기 영역의 평균 점수는 정신노동, 기타, 육체노동 순으로 나타났고,⁷⁾ 통계적으로 유의미한 차이가 있었으며, 육체노동·기타 집단과 정신노동 집단의 차이가 사후비교에서 유의미한 것으로 확인되었다.

〈표 14〉 쓰기 영역의 직업군별 차이

배경변인	평균점수			F	유의도	사후비교
	정신노동	육체노동	기타			
직업군	127.07	118.87	120.42	18.973	0.000	육체노동, 기타 <정신노동

〈표 15〉 쓰기 영역 수준 등급의 직업군별 분포

수준	정신노동		육체노동		기타		합계	
	명	%	명	%	명	%	명	%
4수준	49	5.9	35	2.7	37	4.2	121	4.0
3수준	262	31.4	315	24.6	213	24.1	790	26.4
2수준	338	40.5	538	41.9	378	42.9	1,254	41.8
1수준	186	22.3	395	30.8	254	28.8	835	27.8
계	835	27.83	1,283	42.8	882	29.4	3,000	100.0

5) 학력군별 차이

고졸 미만, 고졸, 대재 이상의 학력군별로 쓰기 영역의 조사 결과의 차이를 분석한 결과는 다음과 같다. 학력군에 따른 쓰기 영역의 평균 점수는

7) 기타 직업군에 속하는 세부 직업으로는 주부, 학생, 무직 등이 있다.

대재 이상 집단이 가장 높게 나타났으며 고졸 미만 집단이 가장 낮은 것으로 나타났다. 집단간 평균점수의 차이에서도 통계적으로 유의미한 차이가 있었으며, 이 차이는 고졸 미만, 고졸, 대재 이상 집단의 차이가 모두 유의미한 것으로 분석되었다.

〈표 16〉 쓰기 영역의 학력군별 차이

배경변인	평균점수			F	유의도	사후비교
	고졸 미만	고졸	대재 이상			
학력	110.70	118.32	126.63	55.572	0.000	고졸미만<고졸<대재이상

〈표 17〉 쓰기 영역 수준 등급의 학력군별 분포

수준	고졸 미만		고졸		대재 이상		합계	
	명	%	명	%	명	%	명	%
4수준	2	0.5	28	2.9	91	5.7	121	4.0
3수준	67	15.4	230	23.5	493	31.1	790	26.4
2수준	198	45.5	417	42.6	639	40.3	1,254	41.8
1수준	168	38.6	304	31.0	363	22.9	835	27.8
계	435	14.5	979	32.6	1,586	52.9	3,000	100.0

3. 2013년과 2018년 쓰기 능력의 추이 비교

본 연구에서는 2013년의 영역별 원 점수 평균을 2018년의 척도에 따라 환산하였으며, 2018년을 기준 년도로 삼아 2013년과 2018년도의 쓰기 능력 수준의 추이를 살펴보고자 한다.

쓰기 영역의 평균 비교를 위해 2013년의 평균값을 2018년으로 환산한

값은 아래의 〈표 18〉와 같다.⁸⁾

〈표 18〉 2013년 원 점수 평균의 동백분위 환산값과 2018년 원 점수 평균

영역	2018년의 척도로 환산한 2013년 원 점수 평균	2018년 원 점수 평균
쓰기	120.2 (122.3)*	121.29

*괄호 안은 2013년 원 점수 평균

2013년도의 평균 점수를 2018년의 평균 점수와 비교하기 위해 동백분위 동등화한 결과 2013년의 쓰기 영역 평균 점수는 120.2로 나타났다. 즉 쓰기 영역의 2013년의 원 점수 평균인 122.3점은 2018년에 120.2점에 해당한다. 2018년도의 쓰기 평균 점수는 121.29로 2013년과 비교하였을 때에 국민의 전체적인 쓰기 능력이 약 1점 정도 상승하였다고 이해할 수 있다. 이는 2018년 조사에 평균 점수가 가장 낮은 60대의 조사 결과가 포함된 것을 감안한다면, 그 상승값의 유의미성이 더 커진다고 볼 수 있다.

쓰기 영역에서 2013년부터 2018년까지의 쓰기 능력의 변화를 살펴보면 아래의 〈표 19〉와 같다.

〈표 19〉 2013년과 2018년의 쓰기 능력의 변화

수준	분할 점수	2013년		2018년	
		인원(명)	비율(%)	인원(명)	비율(%)
4수준	183	55	3.6	121	4.0
3수준	135	341	22.7	790	26.4
2수준	101	829	55.3	1,254	41.8

8) 동백분위 동등화는 2013년 조사와의 일관성을 위해 통합형 검사지로 검사한 표본 500개를 제외한 2,500개의 응답 데이터를 바탕으로 설계되었다.

1수준	101 미만	275	18.4	835	27.8
계		1,500	100.0	3,000	100.0

쓰기 영역은 2013년에는 4수준의 비율이 3.6%, 3수준이 22.7%, 2수준 55.3%, 1수준이 18.4%로 산출되어 전반적으로 국민들의 쓰기 능력은 보통과 기초 수준에 포진되어 있는 것으로 나타났다. 2018년에는 4수준의 비율 (+0.4%p), 3수준의 비율 (+3.7%p)이 소폭 증가하는 양상을 보였고, 2수준은 2013년에 비하여 감소 (-13.5%p)하였으며, 1수준은 증가 (+9.4%p)하였다. 이에 따라 2013년에 2수준, 3수준, 1수준, 4수준의 순서로 비율이 높았던 것과 비교하였을 때, 2018년에는 3수준보다 1수준의 비율이 근소한 차이로 높아졌다. 전체적인 평균은 2013년과 유사한 수준이나 3, 4수준 및 1수준의 비율이 동시에 증가한 것으로 보아, 국민의 국어능력 중 쓰기 능력은 그 편차가 다소 발생하고 있음을 알 수 있다.

1) 성별 쓰기 능력 변화 추이

2013년과 2018년 쓰기 영역에서 각 수준별 조사 대상자 비율의 변화를 성별로 살펴보면 〈표 20〉과 같다.

〈표 20〉 쓰기 영역 수준 등급의 성별 분포

영역 (2013)	수준	남성		여성	
		인원(명)	비율(%)	인원(명)	비율(%)
쓰기 (2013)	우수	29	3.8	26	3.5
	보통	164	21.4	177	24.1
	기초	428	55.8	402	54.7
	기초 미달	146	19.0	130	17.7
	계	767	51.1	735	48.9

쓰기 (2018)	4수준	59	3.9	62	4.2
	3수준	394	25.9	396	26.8
	2수준	644	42.3	610	41.2
	1수준	424	27.9	411	27.8
	계	1,521	50.7	1,479	49.3

4수준의 비율은 2013년의 경우 남성이 3.8%, 여성이 3.5%로 남성의 비율이 조금 더 높았으나, 성별 간 유의할 만한 차이는 보이지 않았다. 2018년에는 남성이 3.9%, 여성이 4.2%로 여성의 비율이 조금 더 높지만 이 또한 성별 간 유의할 만한 차이는 아니었다. 3수준의 경우 2013년에는 남성이 21.4%, 여성이 24.1%로 여성이 다소 높았으며, 2018년에도 남성이 25.9%, 여성이 26.8%로 여성이 높지만 성별 간 차이는 소폭 줄었다. 또한 남성과 여성 모두 3수준의 비율이 2013년에 비해 소폭 상승하였다. 2수준 비율을 살펴보면, 2013년 남성이 55.8%, 여성이 54.7%로 남성의 비율이 조금 더 높았고, 2018년에도 남성이 42.3%, 여성이 41.2%로 남성이 근소하게 더 높으며, 두 성별 모두 2013년에 비해 2수준 비율이 큰 폭으로 감소하였다. 2013년 1수준의 경우 남성이 19%, 여성이 17.7%로 남성의 비율이 조금 더 높았으나, 2018년에는 남성이 27.9%, 여성이 27.8%로 거의 유사한 수준을 보이며, 두 성별 모두 2013년에 비해 1수준 비율이 큰 폭으로 증가하였다.

2) 연령대별 쓰기 능력 변화 추이

2013년과 2018년 쓰기 영역에서 각 수준별 조사 대상자 비율의 변화를 연령대별로 살펴보면 <표 21>과 같다.

〈표 21〉 쓰기 영역 수준 등급의 연령대별 분포

영역	수준	20대		30대		40대		50대	
		명	%	명	%	명	%	명	%
쓰기 (2013)	4수준	19	6.1	16	4.1	16	3.8	3	0.8
	3수준	94	29.9	101	25.6	80	19.0	65	17.5
	2수준	171	54.5	220	55.8	227	54.0	211	56.9
	1수준	30	9.6	57	14.5	97	23.1	92	24.8
	계	314	20.9	394	19.6	420	28.0	371	24.7
쓰기 (2018)	4수준	41	7.4	33	5.6	22	3.2	21	3.0
	3수준	170	30.7	189	32.0	184	26.4	169	24.4
	2수준	211	38.2	236	39.9	313	45.0	283	40.9
	1수준	131	23.7	133	22.5	177	25.4	219	31.7
	계	553	21.8	591	23.4	696	27.5	692	27.3

20대의 경우 4수준(+1.3%p)과 3수준의 비율(+0.8%p)은 2013년보다 약간 높거나 유사한 비율을 보이나, 2수준의 감소폭(-16.8%p)과 1수준의 증가폭(+14.1%p)이 다른 연령대에 비해 크게 나타났다. 또한, 4수준의 비율은 대부분의 연령대에서 소폭 증가하였으나, 40대는 2013년 3.8%에서, 2018년 3.2%로 오히려 소폭 감소하는 경향을 보였다. 2013년에 20대에서 50대로 갈수록 4수준이 감소하는 경향은 2018년에도 유지되나, 40대에 비해 50대의 4수준 비율이 유의하게 낮았던 2013년과 달리, 2018년에는 40대와 50대가 유사한 비율을 보였다. 다른 연령대에서 1수준의 비율이 큰 폭으로 증가한 것과 달리 40대는 1수준의 비율 변화(+2.3%p)가 소폭이었다.

한편, 3수준의 경우 유사한 비율을 보이는 20대를 제외하고, 30, 40, 50대에서 증가하였으며, 특히 2013년에는 20대에서 50대로 갈수록 3수준이 감소하는 경향(20대: 29.9%, 30대: 25.6%, 40대: 19%, 50대: 17.5%)을 보인 반면, 2018년에는 20대보다 30대에서 3수준 비율이 소폭 높고, 이후 연령대 증가에 따라 감소하는 것을 확인할 수 있었다(20대: 30.7%, 30대: 32.0%, 40

대: 26.4%, 50대: 24.4%). 같은 견지에서 2013년에는 1수준 비율 또한 20대에서 50대로 갈수록 증가하였으나(20대: 9.6%, 30대: 14.5%, 40대: 23.1%, 50대: 24.8%), 2018년에는 30대가 가장 낮고, 이후 연령대 증가에 따라 증가하였다(20대: 23.7%, 30대: 22.5%, 40대: 25.4%, 50대: 31.7%).

전반적으로 20대에서 50대로 갈수록 쓰기 능력이 감소하는 경향은 2013년과 같지만, 2013년과 달리 2018년 조사에서는 20대와 30대 간 쓰기 능력에 유의한 차이를 보이지는 않았다. 또한 40대는 다른 연령대와 달리 4 수준의 비율이 소폭 줄었으나, 1수준 증가폭도 다른 연령대보다 적어 전체적인 편차가 줄었다는 점도 특징적이다.

3) 지역 규모별 쓰기 능력 변화 추이

2013년과 2018년 쓰기 영역에서 각 수준별 조사 대상자 비율의 변화를 지역 규모별로 살펴보면 <표 22>와 같다.

<표 22> 쓰기 영역 수준 등급의 지역 규모별 분포

영역	수준	대도시		중소 도시		읍면 지역	
		명	%	명	%	명	%
쓰기 (2013)	우수	31	4.5	16	2.8	8	3.3
	보통	158	22.7	130	23.1	53	21.8
	기초	403	57.9	284	50.4	143	58.8
	기초 미달	104	14.9	133	23.6	39	16.0
	계	696	46.3	563	37.5	243	16.2
쓰기 (2018)	4수준	46	3.5	54	4.7	21	4.0
	3수준	352	26.7	327	28.2	111	21.3
	2수준	538	40.7	481	41.5	235	45.0
	1수준	384	29.1	296	25.6	155	29.7
	계	1,320	44.0	1,158	38.6	522	17.4

2013년부터 2018년까지 쓰기 능력의 변화를 지역 규모별로 세분하여 살펴보면 다음과 같다. 4수준의 경우 중소 도시(+1.9%p)와 읍면 지역(+0.7%p)에서는 소폭 증가하였으며, 대도시의 경우 소폭 감소하였다 (-1.0%p). 3수준의 경우 대도시(+4.0%p)와 중소 도시(+5.1%p)에서는 소폭 증가한 반면, 읍면 지역에는 근소하게 감소(-0.5%p)하였다. 2수준은 모든 지역 규모에서 큰 폭으로 감소하였으며, 그중 대도시의 감소폭이 두드러졌다 (-17.1%p), 1수준은 대도시(+14.2%p)와 읍면 지역(+13.7%p)에서는 큰 폭으로 증가한 반면, 중소 도시에서는 미미한 증가폭(+2.0%p)를 보였다.

전반적으로 2013년에 비해 중소도시의 쓰기 능력은 상향하였고, 대도시는 그 편차가 다소 커졌으며, 읍면 지역의 쓰기 능력은 1수준의 비율이 큰 폭으로 증가함에 따라 다소 하향하였음을 조심스럽게 추정할 수 있다.

4) 직업군별 쓰기 능력 변화 추이

2013년과 2018년 쓰기 영역에서 각 수준별 조사 대상자 비율의 변화를 직업군별로 살펴보면 <표 23>과 같다.

<표 23> 쓰기 영역 수준 등급의 직업군별 분포

영역	수준	정신노동		육체노동		기타	
		명	%	명	%	명	%
쓰기 (2013)	우수	22	5.3	16	2.6	17	3.5
	보통	97	23.3	114	18.8	130	27.0
	기초	221	53.1	349	57.7	260	54.1
	기초 미달	76	18.3	126	20.8	74	15.4
	계	416	27.7	605	40.3	481	32.0

쓰기 (2018)	4수준	49	5.9	35	2.7	37	4.2
	3수준	262	31.4	315	24.6	213	24.1
	2수준	338	40.5	538	41.9	378	42.9
	1수준	186	22.3	395	30.8	254	28.8
	계	835	27.8	1,283	42.8	882	29.4

2013년부터 2018년까지 쓰기 능력의 변화를 지역 직업군별로 세분하여 살펴보면, 4수준의 경우 모든 직업군에서 2013년과 2018년에 큰 변화를 보이지 않았다. 그러나 3수준에서는 정신노동(+8.1%p)과 육체노동(+5.8%p)에서 증가한 반면, 기타에서는 다소 감소(-2.9%p)하여, 2013년에는 3수준의 분포가 ‘기타>정신노동>육체노동’으로 나타났으나, 2018년에는 ‘정신노동>육체노동>기타’로 나타났다. 2수준은 모든 직업군에서 큰 폭으로 감소하였으며, 1수준은 모든 직업군에서 증가하였으나, 정신노동의 변화폭 (+4.0%p)에 비해 육체노동(+10.0%p)과 기타(+13.4%p)에서 큰 폭의 증가율을 보여 2013년에는 1수준의 분포가 ‘육체노동>정신노동>기타’ 순으로 나타난 반면, 2018년에는 ‘육체노동>기타>정신노동’ 순으로 나타났다.

2013년에는 쓰기 능력의 차이에 있어 ‘기타>정신노동>육체노동’의 경향을 확인할 수 있었으나, 2018년에는 정신노동의 쓰기 능력이 소폭 상향하고, 기타의 쓰기 능력이 다소 하향함에 따라 ‘정신노동>기타>육체노동’의 경향을 보였다.

5) 학력군별 쓰기 능력 변화 추이

2013년과 2018년 쓰기 영역에서 각 수준별 조사 대상자 비율의 변화를 학력군별로 살펴보면 <표 24>와 같다.

〈표 24〉 쓰기 영역 수준 등급의 학력군별 분포

영역	수준	고졸 미만		고졸		대재 이상	
		명	%	명	%	명	%
쓰기 (2013)	우수	1	0.6	10	1.9	44	5.4
	보통	34	20.0	101	19.4	206	25.4
	기초	95	55.9	299	57.4	435	53.7
	기초 미달	40	23.5	111	21.3	125	15.4
	계	170	11.3	521	34.7	810	54.0
쓰기 (2018)	4수준	2	0.5	28	2.9	91	5.7
	3수준	67	15.4	230	23.5	493	31.1
	2수준	198	45.5	417	42.6	639	40.3
	1수준	168	38.6	304	31.0	363	22.9
	계	435	14.5	979	32.6	1,586	52.9

2013년부터 2018년까지 쓰기 능력의 변화를 학력군별로 세분하여 살펴보면, 4수준은 모든 학력군에서 변화가 미미하여 비슷한 비율을 유지하고 있다. 3수준은 고졸 미만에서는 감소(-4.6%p)한 반면, 고졸(+4.1%p) 및 대재 이상(+5.7%p)의 경우 증가하였다. 2수준은 모든 학력군에서 큰 폭으로 감소하였으며, 1수준은 모든 학력군에서 큰 폭으로 증가하였다. 2013년에는 4수준은 ‘대재 이상>고졸>고졸 미만’, 3수준은 ‘대재 이상>고졸 미만>고졸’, 2수준은 ‘고졸>고졸 미만>대재 이상’, 1수준은 ‘고졸 미만>고졸>대재 이상’의 분포를 보이는 반면, 2018년에는 4수준과 3수준은 ‘대재 이상>고졸>고졸 미만’, 2수준과 1수준은 ‘고졸 미만>고졸>대재 이상’의 분포를 보인다.

2013년에는 쓰기 능력의 차이에 있어 ‘대재 이상>고졸 미만’, ‘대재 이상>고졸’의 경향을 확인할 수 있었으나, 2018년에는 고졸 미만의 쓰기 능력이 다소 하향함에 따라 ‘대재 이상>고졸>고졸 미만’과 같이 모든 집단 간 유의한 차이가 있음을 확인할 수 있다.

4. 수준별 응답 예시

여기에서는 공개된 1번 문항을 대상으로 채점 척도에 따라 평가 점수를 합산한 결과 상, 중, 하로 분류된 피조사자의 답안을 소개하고자 한다. 문항 1에서는 탁구 동호회 회장 취임 감사 인사를 담은 이메일을 3문장으로 작성 할 것을 요구하였는데, 다음의 세 자료는 각각 모든 평가 요소에서 매우 우수, 보통, 매우 미흡의 점수를 받은 피조사자의 예시 답안들이다.

〈표 25〉 쓰기 문항 1 上 예시 답안

파일번호	성별	연령대	지역규모	직업	학력	쓰기수준
33038	남	50대	대도시	육체노동	대재이상	3수준

회원여러분들 미흡한 저를 뺏아주셔서 감사합니다. 회원들의 화합과 단합을 위해 노력하고 동호회가 즐겁고 좋은 모임이 될수 있도록 소중한 인연을 잘 이어가겠습니다. 모두들 감기 조심하세요

파일번호	성별	연령대	지역규모	직업	학력	쓰기수준
43057	여	20대	읍면지역	정신노동	대재이상	4수준

저를 믿고 뺏아주신 회원여러분들께 다시 한 번 감사 인사드립니다. 앞으로 우리 동호회의 발전을 위해 노력할테니 회원 여러분들의 관심 부탁드립니다.갑자기 추워진 날씨 감기 조심하시고 다음 모임 때 뵙겠습니다.

위의 답안들은 내용, 조직, 표현 모든 영역에서 ‘매우 우수’ 수준의 평가를 받은 답안이다. 우선 내용 측면에서는 글의 목적에 맞게 감사의 인사가 포함되었으며, ‘동호회 회장 취임’이라는 상황 맥락을 고려하여 구체적인 계획과 함께 앞으로의 포부나 당부를 포함하였다. 또한 전자 우편이라는 양식적 특성에 따라 독자를 고려하여 끝인사를 포함한 점과 세 문장이라는 양적인 조건이 충족된 점도 내용 생성 능력의 우수성을 보여준다. 담화 유형의 구조적 특성에 맞게 세 문장이 유기적으로 연결되어 있으며, 맞춤법과 문장 구조 또한 우수한 것으로 확인되었다.

다음 답안은 모든 영역에서 보통의 평가를 받은 답안이다. 첫 번째 답안

의 경우 감사의 인사는 포함되어 있지만 양적인 조건을 충족하지 못하였고, 이에 따라 글의 완결성이 부족하여 조직 측면에서도 좋은 평가를 받지 못하였다. 두 번째 답안의 경우 상황 맥락을 고려하여 앞으로의 포부와 회원들에 대한 당부를 포함하였지만 글의 목적인 '감사 인사'가 누락되어 내용 영역에서 보통의 평가를 받았으며, 세 문장으로 서술되어 있으나 문장 간 응집성이 부족하고 문장의 호응에 오류가 있어 조직과 표현 측면에서도 보통의 평가를 받았다.

〈표 26〉 쓰기 문항 1 中 예시 답안

파일번호	성별	연령대	지역규모	직업	학력	쓰기수준
4057	남	20대	대도시	정신노동	고졸	2수준

저를 선출해 주셔서 정말 감사합니다. 앞으로 회장으로 있는 동안에 동호회의 발전을 위해 최선을 다하겠습니다.

파일번호	성별	연령대	지역규모	직업	학력	쓰기수준
43034	여	60대	중소도시	기타	고졸미만	3수준

앞으로 회원님들과 즐거운 탁구 동호회가 되도록 노력하겠습니다. 탁구를 하며 건강한 생활을 합시다. 부족한 점은 여러분이 채워주시고 도와주세요.

다음은 모든 영역에서 '매우 미흡'에 해당하는 평가를 받은 답안이다.

〈표 27〉 쓰기 문항 1 下 예시 답안

파일번호	성별	연령대	지역규모	직업	학력	쓰기수준
9057	여	20대	대도시	육체노동	고졸	2수준

앞으로 잘 부탁드립니다

파일번호	성별	연령대	지역규모	직업	학력	쓰기수준
25041	남	50대	중소도시	정신노동	고졸	1수준

우리모임은 친목도모가 아니고 건강을 위한 모임이기 때문에 빠짐없이 참여하시는게 최선이라 생각합니다

두 답안 모두 한 문장으로 기술되어 제시된 양적인 조건에 크게 못 미친다. 이에 따라 또한 글의 목적인 감사 인사가 내용에 포함되어 있지 않으며, 글의 시작이나 마무리에 대한 인식이 보이지 않아 조직 능력도 매우 미흡한 것으로 판정되었다. 다만 제시된 예시 답안과 같이 표현의 측면에서 미흡과 매우 미흡의 평가를 받은 답안이라 할지라도 높은 비율의 답안에서 독자를 고려하여 적절하게 높임 표현을 사용한 것이 특징적이다. 이는 친교적 글쓰기 장르가 이메일, 문자 메시지, SNS 등을 통해 일상적으로 수행하는 친숙한 쓰기 양식이라는 점과 무관하지 않은 것으로 보인다.

V. 논의 및 제언

1. 논의

본 연구에 따르면 국민들의 평균적인 쓰기 능력은 '2수준'이며, 평균 점수는 250점 만점에 121.61점(만점 대비 48.64%)으로 듣기, 말하기, 읽기, 디지털, 문법 영역의 평가 결과와 비교할 때 가장 낮았다. 또한 쓰기 능력의 신장이 요구되는 2수준과 1수준의 비율이 전체의 69.6%라는 결과에서, 국민의 3분의 2 이상의 쓰기 능력이 미흡한 수준이라는 해석이 가능하다. 이는 전체적인 결과로는 2013년 조사와 유사한 양상이지만, 2수준 이하의 비율이 2013년 73.7%에 비해 근소하게 감소하였으나, 보충 교육이 시급한 것으로 보이는 1수준의 비율(2013년: 18.4% / 2018년: 27.8%)만 놓고 보았을 때는 큰 폭의 증가세가 두드러진다는 점에서 차이를 보인다. 이러한 현상은 국민들의 쓰기 능력의 저하가 장기간 심화되어 갈 가능성을 암시하는 것이라는 점에서 문제적이다.

이에 따라 쓰기 능력이 저조한 원인을 분석하고 이에 대한 대처 방안을

마련하기 위해 변인별 쓰기 능력 변화 추이를 중심으로 특징적인 내용을 살펴보자 한다. 성별 쓰기 능력의 경우 여성이 남성에 비해 근소하게 높지만 성별 간 유의할 만한 차이는 아니며 2013년과 유사한 결과라고 할 수 있다. 이는 우리나라 성인의 언어 능력의 경우 성별 차이가 크며, 남성의 언어 능력이 더 높다는 국제성인역량조사(PIAAC) 결과(한국직업능력개발원, 2013: 14)와는 배치되는 것이기도 하다. PIAAC의 경우 쓰기 문항이 포함되어 있지 않다는 점, 해당 조사가 실시된 2011년 이후에도 지속적으로 고등교육기관에 재학하는 여성의 비율이 높아지고 있다는 점, 학습자 대상 연구 결과에서는 대부분 여학생이 남학생보다 높은 쓰기 능력을 보고하고 있다는 점 등을 함께 고려해볼 수 있다.

연령대별 쓰기 능력 변화의 경우 모든 연령대에서 1수준의 비율이 증가하였으나, 그중 20대의 증가폭은 2.5배 수준의 오름세를 보였다. 20대에서 60대로 갈수록 쓰기 능력이 저하되는 경향은 2013년과 동일하지만 다른 연령대의 쓰기 능력이 소폭 상향된 것에 비해 20대의 쓰기 능력이 다소 하락한 것은 면밀한 고찰이 필요하다. 구어적 속성을 지녀 주로 한 두 문장으로 작성하는 인스턴트 메신저나, 페이스북·트위터 등 130~140자 등으로 쓰기 분량이 제한되는 단문 쓰기에 익숙해진 쓰기 상황으로 인해 300~400자 이상의 긴 글쓰기에 서툰 것이 하나의 이유로 지적될 수 있다. 피험자들의 쓰기 답안을 질적으로 살펴본 결과, 2번 문항과 3번 문항의 경우 작성된 답안의 길이가 2013년에 비해 전반적으로 다소 짧아진 것을 확인할 수 있었는데 이는 손글씨 조사에서 컴퓨터 기반 조사로 쓰기 방식이 변화된 것을 고려해야 하겠으나 문단을 나누어 길게 작성하는 글쓰기에 대한 미숙성에 기인하였을 가능성을 배제할 수 없다.

20대의 쓰기 능력 저하에 대한 또 한 가지 원인으로 대학 진학률의 감소를 고려할 필요가 있다. 2018년 대학 진학률은 69.7%로 2008년 83.8% 이후 지속적인 감소 추세를 보인다(교육부, 2018). 즉 2018년 조사시의 20대 피험자의 평균 학력이 2013년 20대 피험자의 평균 학력보다 낮아졌을 가능성

을 의미하는 것이다. 학력과 쓰기 능력 간 유의한 상관 관계가 있음(김종철 외, 2013)을 고려할 때, 대학 진학률의 감소가 쓰기 능력의 저하로 이어지지 않도록 고등교육기관이 아니어도 지속적으로 한 편의 완결된 글을 쓰고, 학습할 수 있는 기회가 제공될 필요가 있다. 또한 일반적인 역량이 연령이 증가함에 따라 저하된다는 것을 감안하면, 20대의 쓰기 능력이 저하되는 경향은 남은 생애 주기의 총체적인 쓰기 능력의 저하를 가져올 수 있다는 점에서 중등 교육이 완료된 이후 쓰기 능력이 급격히 하락하지 않도록 지속적이고 주의 깊은 관리가 요구된다.

아울러, 30, 40, 50대의 쓰기 능력은 소폭 상향화 된 것으로 보이는데 2013년도 조사 당시의 연령대 사람들의 연령이 높아짐에 따른 것으로 추정 할 수 있으나, 2013년도 조사 당시의 연령대 사람들이 5년 뒤에도 같은 수준을 보이는지를 좀 더 면밀히 추적할 필요가 있고, 이 연령대가 학교 교육을 받았을 때 어떤 교육적 상황이었는지도 앞으로 찾아볼 필요가 있을 것이다.

지역 규모별 쓰기 능력 변화의 경우 2013년에 비해 대도시와 읍면 지역에서 1수준의 비율이 2배가량 증가하였음을 확인할 수 있는데, 특히 읍면 지역은 다른 지역과 달리 3수준의 비율 감소도 동반되어 쓰기 능력이 다소 하향하였음을 추정할 수 있다. 이는 직업군별 쓰기 능력 변화와 함께 살펴볼 필요가 있다. 직업군의 경우에도 정신노동과 달리 육체노동과 기타 직업군에서 1수준의 비율이 큰 폭으로 증가하였기 때문이다. 이는 읍면 지역 거주자나 육체노동·기타 직업군의 경우 생활이나 직업상에서 전자기기를 활용 한 문자 활동, 한 편의 완성된 글을 작성하는 경험이 도시 지역 거주자나 정신노동 직업군에 비해 상대적으로 적기 때문인 것으로 볼 수 있다. 앞서 살펴본 바와 같이 일상적으로 자주 수행하는 ‘친교적 글쓰기’의 경우 두드러지게 정답률이 높았던 점도 평소의 쓰기 경험 빈도나 기회가 쓰기 능력에 긍정적 영향을 미칠 수 있음을 보여준다.

학력군별 쓰기 능력 변화의 경우에도 2013년에는 고졸 집단과 고졸 미만 집단의 차이가 유의하지 않았으나, 본 연구에서는 고졸 미만의 쓰기 능력

이 다소 하향함에 따라 모든 집단 간 유의한 차이를 확인할 수 있었다. 고졸 미만 집단의 경우 다른 집단과 달리 4수준, 3수준, 2수준의 비율 모두가 하락 하였고, 1수준의 비율만이 큰 폭으로 상향되었다. 이는 국민의 쓰기 능력의 양극화가 심해질 수 있다는 점을 의미하는 것이라는 점에서 문제적이다. 국민의 쓰기 능력이 소통의 가능성과 직결되는 것이라고 할 때, 집단 간 소통의 부재가 심화되는 문제 상황을 함의하는 것으로 볼 수 있기 때문이다. 이에 따라 쓰기 능력과 관련된 정책적·교육적 대처를 강화할 필요가 있으며, 특히, 2수준 이하를 대상으로 기존의 교육을 재점검하고 특화된 정책과 국어 교육 과정을 마련할 필요가 있다.

2. 제언

본 연구의 목적인 ‘국민의 쓰기 능력 진단과 정책 마련’은 궁극적으로 국민의 쓰기 능력 신장을 지향하는 바, 끝으로 쓰기 능력의 신장을 위한 대안과 후속 연구의 방향에 대해 고찰하며 논의를 마무리하고자 한다.

쓰기 능력 평가 결과를 통해 알 수 있는 가장 중요한 문제는 전반적인 국민의 쓰기 능력 수준이 매우 낮다는 점이다. 이를 해결하기 위해 우선적으로 정책적·교육적 대처가 요구되는 2수준 이하 집단을 벤인별로 살펴보면 연령대별로는 50대 이상, 지역 규모별로는 읍면 지역, 직업군별로는 육체노동, 학력군별로는 고졸 미만일 것이다. 그러나 쓰기 영역의 경우 평균이 가장 높았던 20대, 중소 도시, 정신노동, 대체 이상 집단마저도 2수준 이하의 비율이 60% 이상이기 때문에 특정 집단이 아닌 대국민 쓰기 교육을 시행하고, 기존에 시행했던 정책을 면밀히 검토하여 구체적인 대안을 마련할 필요가 있다.

이와 관련하여 민병곤 외(2018)에 제시된 국어 학습 유무에 따른 쓰기 영역의 평균 점수가 국어 학습 참여 집단의 경우 비참여 집단에 비해 13.87 점 높은 것으로 확인되었고, 이것이 통계적으로 유의미한 차이였다는 논의

는 주목할 만하다. 다양한 쓰기 학습 기회의 제공, 쓰기 능력 수준에 맞춘 교육 프로그램의 운영, 성인을 대상으로 한 프로그램 참여 독려를 통한 쓰기 학습 문화의 정착 등이 핵심적인 대안이 될 수 있음을 보여주는 결과이기 때문이다.

또한 실제 작성된 답안을 질적으로 확인한 결과 쓰기 문항 중 설명적 글 쓰기의 경우 양적으로 내용을 생성하는 데 급급하여 문장의 질이 떨어지고 묘사의 목적이 온전히 달성되지 못하는 경향을 확인할 수 있었으며, 논증적 글쓰기의 경우에도 근거를 3가지 이상 쓴 답안은 거의 없고, 반론을 고려한 답안도 매우 드물었던 것을 확인할 수 있었다. 반면 친교적 글쓰기에서는 비교적 내용적, 수사적으로 우수한 답안이 다수 확인되었다. 이는 일상적으로 접할 기회가 많아 친숙한 쓰기 활동일수록 뛰어난 수행 능력을 보일 가능성 이 높아짐을 의미한다. 이를 위해서는 다양한 장르의 쓰기 활동에 성인들이 스스로 참여하도록 장려하는 것이 요구되며, 이는 온전한 한 편의 글을 쓰는 능력이 개인의 성공적인 삶의 기회를 확대할 수 있음에 대한 인식이 사회적으로 공유됨으로써 가능해질 수 있다. 예컨대 2017년 도입된 국민청원제도의 경우 30일 동안 20만 명 이상의 국민이 추천한 청원에 대해 정부와 청와대 관계자가 직접 답변하는 형식을 취하는데, 이로 인해 건의하는 글쓰기 장르에 대한 친숙성, 적절한 근거를 들어 독자를 고려한 한 편의 글을 쓰는 것의 중요성에 대한 인식이 다소 확대된 것이 한 예이다. 이처럼 쓰기를 통한 사회적 참여를 통해 쓰기 능력을 끊임없이 활용할 수 있도록 국가적 차원에서의 권장이 필요하다. 뿐만 아니라 한 편의 글쓰기가 교실에서 시작되고 종결되며 교사만을 유일한 독자로 삼는 현재의 쓰기 교육이 쓰기를 통한 사회적 소통 교육, 개인적·사회적으로 의미 있는 쓰기 교육으로 변화된다면 중등 교육이 종결된 이후에도 주체적으로 쓰기 활동을 지속해나갈 가능성을 높일 수 있을 것이다.

평가 방법의 차원에서 5년 주기의 해당 연구가 실효성을 갖기 위해서는 국민의 쓰기 능력 실태를 있는 그대로 볼 수 있도록 실제성이 확보된 조

사 도구가 중요하다. 실제의 쓰기 맥락에서 읽기와 쓰기가 분절적으로 수행되지 않으며, 링크와 리트윗의 방식으로 타인의 글이나 자료를 활용하여 글을 쓰는 빈도가 높아지고 있다는 점에서 주어진 지문을 활용하여 쓰는 문항을 개발하는 것도 고려해 볼 수 있다. 읽기 능력과 별도로 쓰기 능력만을 측정할 수 있는가의 문제는 더 고민해 보아야 하겠으나, 주어진 자료를 바탕으로 자신의 생각을 뒷받침하도록 함으로써 쓰기에서의 자료 활용 능력과 적절한 근거 제시 능력 등 총체적이고 실제적인 쓰기 능력을 평가할 수 있다는 점에서 고려해 볼 필요가 있다.

또한 쓰기는 인지적 부담이 큰 활동이기 때문에 간접평가 방식과 달리 직접평가의 경우 피험자의 태도가 작성된 글의 질에 큰 변인으로 작용할 수 있다. 답안들을 질적으로 살펴본 결과 답안의 분량이나 완성도의 편차가 다소 확인되었는데, 이러한 점에서 피험자의 성실도를 제고할 수 있는 방안이 필요할 것이다. 예컨대 현재는 쓰기 평가를 다른 영역 평가와 통합하여 수행하고 있으나, 다른 영역과 별도로 일정이나 시간을 배분하여 충분한 시간을 두고 수행하거나, 태블릿으로 입력을 하는 단계에서 채점자가 주안점으로 볼 부분들이 무엇인지(목표 측정 능력)를 간략하게 명시해주거나, 자신이 작성한 답안이 요구된 분량의 어느 정도를 충족하였는지 화면에 표시되게 하는 등 최선의 수행을 이끌어낼 수 있도록 하는 조사 방법이 요구된다.

아울러 후속 연구에서는 피험자의 변인에 대한 섬세한 분석 연구가 필요하다. 민병곤 외(2018)에서 제시한 바와 같이 최종 전공, 독서량, 인터넷 이용량 등의 변인에 따라 쓰기 능력의 유의미한 차이가 발견되지 않았는데 그렇다면 쓰기 능력에 영향을 미치는 핵심적 요인은 무엇이며, 쓰기 능력과 관련한 정책 마련시 더욱 주의 깊게 살펴야 하는 부분은 무엇인지 세밀하게 해석할 필요가 있다.

* 본 논문은 2019.10.31. 투고되었으며, 2019.11.19 심사가 시작되어 2019.12.9. 심사가 종료되었음.

참고문헌

- 가은아(2011), 「쓰기 발달의 양상과 특성 연구」, 한국교원대학교 박사학위논문.
- 주용완·이계남·최영준·윤은정(2018), 『2017 인터넷 이용 실태조사 요약 보고서』, 과천: 과학기술정보통신부; 서울: 한국인터넷진흥원.
- 교육부(2018), 『간추린 교육통계』, 세종: 교육부.
- 권태현(2012), 「쓰기 채점 방식에 따른 국어교사의 채점 신뢰도 비교 연구: 고등학생의 비평문 채점을 중심으로」, 『교원교육』 28(3), 113-134.
- 김라연(2007), 「총체적 쓰기 평가와 분석적 쓰기 평가의 상관 연구: 고등학생의 논설문 평가를 중심으로」, 『이중언어학』 35, 103-126.
- 김성숙(2011), 「학문 목적 기초 한국어 쓰기 능력 평가 척도 개발과 타당성 검증」, 연세대학교 박사학위논문.
- 김정자(1992), 「쓰기 평가 방법 연구: 평가기준의 설정과 그 적용을 중심으로」, 서울대학교 석사학위논문.
- 김종철·민병곤·박현정·정희창·김봉순·박재현·주재우·이관희·이지수·박은진·정재미(2013), 『2013년 국민의 국어 능력 평가』, 서울: 국립국어원.
- 남민우(2011), 「국어과 학업성취도 평가 도구의 국제 비교 연구」, 『국어교육연구』 40, 1-36.
- 노창수(1987), 「글짓기 영역의 기능평가론」, 『국어교육』 61, 37-64.
- 민병곤·김종철·구본관·손원숙·김혜정·박재현·김승현·서보영·이문수·장지혜·이효정·차경미·강지영·김동섭·허모아(2018), 『2018년 국민의 국어 능력 실태 조사』, 서울: 국립국어원.
- 박영목(1999), 「작문 능력 평가 방법과 절차」, 『국어교육』 99, 1-29.
- 박영민(2011), 「국가 수준 학업성취도 평가 개선을 위한 '쓰기' 평가 이론의 정립」, 『국어교육』 134, 87-112.
- 박영민·가은아(2009), 「인문계 고등학생의 쓰기 지식과 쓰기 수행의 상관 및 성별, 학년별 차이 연구」, 『국어교육』 128, 175-202.
- 서수현(2008), 「요인 분석을 통한 쓰기 평가의 준거 설정에 대한 연구」, 고려대학교 박사학위논문.
- 성태제(1998), 「문항 제작 및 분석의 이론과 실제」, 서울: 학지사.
- 이수진(2011), 「쓰기 직접 평가를 위한 텍스트 분석 평가의 실행 방안」, 『작문연구』 13, 142-166.
- 장성민·민병곤(2016), 「대학수학능력시험 "국어" 작문 영역 평가의 타당성 검토: 선다형 점수와 수행형 점수의 상관관계를 중심으로」, 『작문연구』 29, 163-209.
- 정희모·이재성(2009), 「대학생 글에 대한 총체적 평가와 분석적 평가의 결과 비교 연구」, 『청람어문교육』 39, 251-273.
- 지은립(1999), 「사회과 보고서 수행평가를 위한 총체적 채점과 분석적 채점의 비교」, 『교육평가연구』 12(2), 11-24.

- 최미숙·원진숙·정혜승·김봉순·이경화·전은주·정현선·주세형(2017), 『국어 교육의 이해: 국어 교육의 미래를 모색하는 열여섯 가지 이야기』(3판), 서울: 사회평론.
- 한국교육평가학회(2004), 『교육평가용어사전』, 서울: 학지사.
- 한국직업능력개발원(2013), 『한국인의 역량, 학습과 일: 국제성인역량조사(PIAAC) 보고서』, 서울: 한국직업능력개발원.
- Bawarshi, A. S. & Reiff, M. J. (2015), 『장르 : 역사·이론·연구·교육』, 정희모·김성숙·김미란·박정하·유혜령·김기란·한래희·이윤빈·오현지·전성우·강지은·이세라 (역), 광명: 경진출판(원서출판 2010).
- Bereiter, C. (1980), Development in writing, In L. W. Gregg & E. R. Steinberg(Eds.), *Cognitive Process in Writing*, Hillsdale, NJ: LEA.
- Hout, B. (1990), "The Literature of direct writing assessment: Major concerns and prevailing trends". *Review of Educational Research* 60(2), 237-263.
- Purves, A. C. (1984), "In search of an internationally valid scheme for scoring compositions", *College composition and communication* 35(4), 426-438.
- Weigle, S. C. (2002), *Assessing Writing*, Cambridge: Cambridge University Press.
- Wolcott, W. & Legg, S. M. (1998), *An Overview of Writing Assessment*, Urbana, IL: NCTE.
- Yancey, K. B. (1999), "Looking back as we look forward: Historicizing writing assessment", *College Composition and Communication* 50(3), 483-503.

2018년 국가수준 성인의 쓰기 능력 평가 결과 분석

장지혜 · 서보영 · 김종철

본 연구는 국립국어원에서 시행한 2018년 국민의 국어 능력 실태 조사 중 쓰기 영역의 평가 결과를 분석하고 논의하는 것을 목적으로 한다. 조사는 만 20세~69세의 성인 남녀 3,000명을 대상으로 시행되었으며 인구 분포 통계를 활용한 2단계 충화 표본 추출(stratified sampling) 방식을 적용하여 조사 대상의 대표성을 확보하였다. 또한 '성별, 연령대, 지역 규모, 학력군, 직업군'의 5개 변인을 설정하여 각 변인별로 실제 인구 분포 비율에 비례하게 조사 대상이 표집될 수 있도록 표본을 추출하였다.

영역별 평균 원 점수는 쓰기 영역이 121.61(250점 만점)을 보였다. 조사 대상 국민의 평균적인 쓰기 능력은 2수준에 해당하며, 특히 2수준 이하 비율이 전체의 69.6%여서 전반적인 쓰기 능력이 저조한 것으로 나타났다.

본 조사에서는 2013년과 2018년의 조사도구 간 동등화를 바탕으로 연도별 국어능력의 결과를 상호 비교하는 것이 가능하도록 설계하였으며 동백 분위 동등화(equipercentile equating)의 방식을 통해 2013년과 2018년 국어능력을 비교하였다. 그 결과 2018년도의 쓰기 영역 평균 점수는 2013년도 쓰기 영역 평균 점수 120.2와 비교하였을 때 약 1점 정도 상승하였다. 전체적인 평균은 2013년과 유사한 수준이나 3, 4수준의 비율은 소폭 증가한 반면 1수준의 비율이 큰 폭으로 증가한 것으로 보아, 국민의 국어능력 중 쓰기 능력은 상하위 수준 집단 간 편차가 다소 발생하고 있음을 알 수 있다. 연령 대별로는 50대 이상, 지역 규모별로는 읍면 지역, 직업군별로는 육체노동, 학력군별로는 고졸 미만에 우선적으로 정책적 · 교육적 대처가 요구되나, 그밖

의 모든 집단에서도 2수준 이하의 비율이 60% 이상이므로 특정 집단이 아닌 대국민 쓰기 교육과 구체적 대안 마련이 필요하다.

이러한 논의 결과를 바탕으로, 다음의 네 가지 과제를 제안하였다. 첫째, 다양한 쓰기 학습 기회 및 쓰기 능력 수준에 맞춘 교육 프로그램의 제공. 둘째, 쓰기를 통한 사회적 참여를 통해 생애에 걸쳐 쓰기 능력을 활용할 수 있도록 하는 국가적 차원에서의 장려. 셋째, 실제성이 확보된 조사 도구 및 최선의 수행을 이끌어낼 수 있는 조사 방법의 마련. 넷째, 피험자 성인에 대한 분석을 통해 정책 마련의 토대 확보 등이다.

핵심어 쓰기 능력, 쓰기 평가, 쓰기 수준, 쓰기 교육, 국어 능력, 국어 능력 평가, 국가수준 국어 능력 평가

ABSTRACT

An Analysis of the 2018 National Assessment of the Korean Writing Ability of Adults

Jang Jihye · Seo Boyoung · Kim Jongcheol

The purpose of this study was to analyze and discuss the results of the 2018 assessment of Korean language ability of adults, focusing on writing ability. This survey was conducted with 3,000 adult men and women aged 20 to 69.

The results showed that the average writing score was 121.61 in 2018, about one point higher than in 2013. This performance is generally very poor.

In the 2018 survey, the cut-off score of the 2013 study was applied, but the score of the division was newly formulated through the equalization of the survey tools, and each grade was classified into levels 4, 3, 2, and 1. In addition, except for the variables newly introduced in 2018, the change in the participants' language ability was confirmed based on the variables set in 2013, namely sex, age, regional scale, education, and occupation. In 2018, the ratio of level 1 increased(+9.4%p), so the order of ratio of levels 2, 1, 3, and 4, compared to the ratio of levels 2, 3, 1, and 4 in writing in 2013. The overall average was similar to that of 2013 but the ratios of levels 3, 4 increased slightly, and 1 increased greatly, indicating that there was a slight variation in the writing ability of the Korean adults.

When the participants were grouped by variable and ranked according to their writing skills, the scores of participants in their 20s, 30s, were higher than the scores of participants in their 40s and 50s. In terms of region size, the scores of participants in the urban areas were higher than the scores of participants in rural areas. The ranking by occupation showed that the scores of mental workers were higher than the scores of manual workers. Then, the ranking by education showed that univer-

sity graduates had higher scores than high school graduates and middle school graduates. However, there was no difference between the scores of men and those of women.

Based on the findings, we propose the following. First, to improve writing ability among the general public, practical educational programs should be developed. Second, social participation through writing should be encouraged at the national level and people should be encouraged to use their writing skills throughout their lifetime. Third, a large-scale writing assessment environment should be created to improve the validity, reliability, and authenticity of the evaluation. Finally, the subject variables should be closely analyzed to inform policy-making.

KEYWORDS Writing Ability, Writing Assessment, Writing Level, Writing Education, Korean Language Ability, Assessment of Korean Writing Ability, National Assessment of Korean Writing Ability

〈부록 1〉

쓰기 채점자 간 신뢰도

	채점자 1	채점자 2	채점자 3	채점자 4	채점자 5	채점자 6	채점자 7	채점자 8	채점자 9	채점자 10	문항 번호
채점자 1	-	0.865	0.852	0.885	0.857	0.855	0.880	0.862	0.879	0.892	w-1
	-	0.870	0.865	0.847	0.869	0.883	0.853	0.826	0.860	0.864	w-2
	-	0.863	0.851	0.855	0.859	0.837	0.839	0.842	0.867	0.847	w-3
채점자 2	-	-	0.858	0.865	0.869	0.882	0.834	0.842	0.840	0.867	w-1
	-	-	0.871	0.881	0.883	0.865	0.850	0.870	0.859	0.859	w-2
	-	-	0.904	0.903	0.902	0.909	0.883	0.861	0.924	0.888	w-3
채점자 3	-	-	-	0.864	0.855	0.866	0.836	0.838	0.869	0.885	w-1
	-	-	-	0.869	0.886	0.883	0.820	0.836	0.867	0.856	w-2
	-	-	-	0.901	0.875	0.908	0.894	0.869	0.882	0.870	w-3
채점자 4	-	-	-	-	0.842	0.905	0.863	0.896	0.893	0.894	w-1
	-	-	-	-	0.885	0.862	0.835	0.880	0.874	0.849	w-2
	-	-	-	-	0.895	0.918	0.887	0.881	0.907	0.853	w-3
채점자 5	-	-	-	-	-	0.895	0.848	0.913	0.848	0.839	w-1
	-	-	-	-	-	0.865	0.883	0.897	0.872	0.875	w-2
	-	-	-	-	-	0.901	0.902	0.895	0.899	0.874	w-3
채점자 6	-	-	-	-	-	-	0.847	0.856	0.871	0.916	w-1
	-	-	-	-	-	-	0.845	0.852	0.889	0.872	w-2
	-	-	-	-	-	-	0.881	0.896	0.921	0.891	w-3
채점자 7	-	-	-	-	-	-	-	0.821	0.840	0.850	w-1
	-	-	-	-	-	-	-	0.837	0.828	0.873	w-2
	-	-	-	-	-	-	-	0.865	0.877	0.861	w-3
채점자 8	-	-	-	-	-	-	-	-	0.875	0.870	w-1
	-	-	-	-	-	-	-	-	0.847	0.867	w-2
	-	-	-	-	-	-	-	-	0.891	0.844	w-3
채점자 9	-	-	-	-	-	-	-	-	-	0.899	w-1
	-	-	-	-	-	-	-	-	-	0.846	w-2
	-	-	-	-	-	-	-	-	-	0.904	w-3
채점자 10	-	-	-	-	-	-	-	-	-	-	w-1
	-	-	-	-	-	-	-	-	-	-	w-2
	-	-	-	-	-	-	-	-	-	-	w-3

〈부록 2〉

쓰기 영역 동백분위 결과

2018년 점수	2013년 점수
	동백분위방법
54	50
55	50
56	50
57	50
58	50
59	50
60	50
61	50
62	50
63	50
64	50
65	51
66	53
67	54
68	54
69	54
70	57
71	58
72	59
73	60
74	61
75	61
76	62
77	63
78	66
79	68

80	69	113	109
81	70	114	111
82	71	115	113
83	73	116	114
84	76	117	116
85	77	118	118
86	78	119	121
87	79	120	122
88	79	121	124
89	81	122	126
90	82	123	128
91	83	124	129
92	84	125	131
93	86	126	132
94	87	127	134
95	87	128	135
96	89	129	137
97	90	130	138
98	90	131	139
99	92	132	140
100	93	133	141
101	94	134	143
102	96	135	143
103	97	136	145
104	98	137	145
105	99	138	147
106	100	139	148
107	101	140	149
108	103	141	151
109	104	142	153
110	105	143	153
111	105	144	154
112	107	145	155

146	156	179	189	212	214
147	157	180	190	213	216
148	158	181	191	214	217
149	159	182	192	215	217
150	160	183	193	216	217
151	161	184	193	217	220
152	162	185	193	218	220
153	163	186	194	219	222
154	163	187	194	220	222
155	164	188	195	221	222
156	165	189	197	222	224
157	166	190	197	223	224
158	167	191	198	224	224
159	168	192	198	225	224
160	169	193	199	226	224
161	169	194	201	227	224
162	170	195	202	228	224
163	172	196	202	229	224
164	173	197	204	230	224
165	174	198	205	231	231
166	174	199	206	232	231
167	176	200	208	233	231
168	177	201	208	234	231
169	177	202	209	235	231
170	178	203	211	236	231
171	179	204	211	237	240
172	180	205	211	238	240
173	181	206	213		
174	182	207	213		
175	183	208	213		
176	184	209	213		
177	186	210	214		
178	187	211	214		