

쓰기성취검사 개발의 예비 연구

윤 희 령 한 종 철

연세대학교

본 연구는 위계적이고 다면적인 쓰기능력에 대한 성취검사의 개발을 목적으로 한 예비연구로서, 쓰기성취검사를 제작하고, 신뢰도와 타당도를 검증하여 그 적용가능성을 알아보기 위해서 수행되었다. 쓰기성취검사는 쓰기의 발달이론적 진단모형에 근거하여 글씨 검사, 맞춤법 검사, 및 작문 검사의 세 소검사로 구성되었다. 각 소검사의 모든 검사문항은 국민학교 교사의 자문을 받아서 선발되었으며, 글씨 검사 및 맞춤법 검사는 국민학교 2, 4, 6학년 아동 264명을 대상으로 예비검사 단계를 거쳐 본검사를 제작하였다. 제작된 쓰기성취검사의 신뢰도와 타당도를 알아보기 위해 2, 4, 6학년 국민학생 256명에게 검사를 실시하였다. 신뢰도 검증에서, 쓰기성취검사의 세 가지 소검사의 내적합치도(Cronbach's α)는 양호하였으며, 글씨 검사의 평정자간 일치도와 검사-재검사 신뢰도의 상관계수는 모두 중위상관 이상을 나타내었다. 타당도 검증을 위해서 검사수행과 교사평정 및 관련준거검사, 학년, 국어성적과의 상관을 보고, 성별인 및 가정환경변인에 따른 수행의 차이를 분석하였고, 그 결과 모든 준거 및 구인과의 상관이 유의미하였고, 성별인에 따른 집단간 수행의 차이가 유의미하였고, 가정환경변인에 따른 집단간 수행차이가 나타나지 않았다. 이들 결과들은 제작된 쓰기성취검사가 쓰기능력에 대한 하나의 신뢰롭고 타당한 검사임을 시사해 주고 있다. 추후연구를 통해 세련화와 표준화 작업이 이루어지면 본 검사는 쓰기장애를 포함하여 쓰기문제를 가진 아동들을 위한 봉사전달에 유용하게 활용될 수 있을 것이다.

언어는 사람 이외의 다른 동물들에게는 없는 고유하며 창조적인 의미 상징체계이다. 언어에는 말하기, 듣기, 읽기, 쓰기의 네 가지 의사소통 기능이 있으며, 이 기능들은 계층적인 위계를 가지고 있다. 이 중에서 쓰기기능은 언어 위계의 가장 높은 단계에 있으며, 연구자들은 쓰기를 필자의 지식 및 특정 관심 문제에

관한 그들의 생각이나 감정을 반영하는, 가시적이며, 쉽게 읽을 수 있고, 이해가능한 언어를 산출하는 문제해결과정이라고 정의한다(Berninger & Hooper, 1993). 글로써 의사전달을 효과적으로 할 수 없는 개인은 심각한 표현 문제의 고통을 받고 있으며(Luftig, 1987), 쓰기표현의 효율성 여부는 교육의 일차적 담

당기관인 학교에서의 아동의 적응과 학업성취에 상당한 영향을 미친다. 한편, 쓰기곤란의 범위는 글쓰기를 싫어하는 것에서부터 글씨쓰기, 문제 명료성 및 유창성, 맞춤법 및 작문 등의 문제에 이르기까지 다면적이고 광범위하다. 그러므로, 쓰기능력을 평가할 때는 이러한 다양하며 위계적인 쓰기 요소들을 고려하여야 한다. 쓰기평가와 관련된 대표적인 쓰기요소들은 다음과 같다.

대표적인 쓰기요소들과 평가도구들

글씨(handwriting) 의사소통을 글로 하는 많은 경우에 연필과 종이 등의 필기구가 사용된다. 연구자들은 대부분의 글씨쓰기결함은 소수의 글씨오류 때문이라는 것을 지적했다(Greene & Petty, 1967). Mann, Suiter, 및 McClung(1987)은 일반적인 글씨 쓰기 오류의 유형을 글자 형태, 띄어쓰기, 기울기(slant), 선의 질과 배치, 그리고 속도 등의 6가지 범주로 분류하였다. 우리나라에서는 홍은경(1987)이 유아를 대상으로 한글 쓰기에 있어서의 오류의 경향과 쓰기간에 따른 쓰기시간과 글자의 안정성을 연구하였다.

맞춤법(spelling) 맞춤법은 구어적 또는 문어적으로 단어 내의 낱자들의 순서를 정확하게 인지, 회상, 재생산, 또는 획득하는 능력으로 정의된다(Graham & Miller, 1983). 즉, 구어적 또는 문어적으로 한 단어의 철자를 맞출 때는, 음운과 문자를 적절히 일치시킬 수 있어야 한다(Beers & Henderson, 1977; Henderson, 1985). 연구자들(Graham & Miller, 1983; Henderson, 1985)이 확인한 맞춤법에 필요한 기술들에는 ① 음운 간의 변별, ② 각 음운의 문자 선택치의 확인, ③ 구어의 음절을 확인하고, 쓸 때 그것들을 재생하는 것, ④ 단어에서 강세를 인지하는 것, ⑤ 어근, 복합어, 접두사, 접미사 등 형태소를 인지하는 것, ⑥ 어떻게 단어가 결합되고 다른 단어를 만들기 위해서 재결합되는가에 대한 형태소 규칙을 아는 것 등이 있다. 학생이 이들 기술들에 더욱 숙달될 수록, 단어를 정확하게 쓰는 데 더 효율적이다(Luftig, 1989). 맞춤법 문제의 심각성, 성질 및 오류 양태를 진단하

기 위해 사용되는 표준화된 가용한 한국어 맞춤법 검사에는 기초학습기능검사(박경숙, 윤점룡, 및 박효정, 1989) 내의 철자재인검사가 있으며, 이는 재인방식의 비언어적 응답검사이다.

작문(composition) 문어를 통한 의사소통인 작문은 의사소통의 가장 복잡한 형태이다. 능숙한 작문을 하기 위해서는 내적인 발상을 문어적 형태로 변환시킬 수 있어야 할 뿐만 아니라 어휘를 적절하게 구사하여 읽고 쓰며, 다른 사람들이 이해할 수 있는 글을 쓸 수 있어야 한다. 이런 기술들에서의 결합은 글을 통해 효과적으로 의사소통하는 데 있어서의 문제를 야기할 수 있다. 또한 효과적인 문어표현을 위해서는 작문(composition)과 사필(transcription) 기술의 2가지 범주화된 능력이 요구된다(Polloway, Patton, & Cohen, 1983; Smith 1982). 작문은 아이디어를 글의 형태로 배치하는 것이며(Hammill, 1986; Smith, 1982), 이것은 필자가 표현하고자 하는 사고와 개념이 있고, 이를 엮어내는 조직기술 및 어휘기술이 있어야 하고, 내적 사고를 글의 형태로 변환시키려는 욕구(동기)가 있음을 시사한다. 그리고 글을 옮겨 적는 사필에는 맞춤법, 문법과 문장부호에 대한 지식과 같은 기계적인 요소들이 필요하며, 완성된 산물이 좀더 좋아질 수 있도록 하기 위한 재고와 비평이 포함된다. 한편 현재 가용한 표준화된 한국어 작문검사는 없다.

쓰기장애

쓰기문제를 아동기 발달장애의 한 유형으로 보는 경우가 있다. 이는 문어적 표현의 장애(disorder of written expression : 이하 쓰기장애)라 하는 것으로 학습장애(Learning Disability : LD)의 한 하위 유형으로 분류된다. 학습장애는 아동기 발달 장애의 하나로서, 아동의 실제 수행 수준과 아동의 잠재력에 근거한 기대 수준과의 불일치를 그 특징으로 하고, 그 하위유형에는 쓰기장애, 읽기장애, 수학장애, 기타 학습장애 등의 4가지가 있다. DSM-IV에서 명기된 쓰기장애의 핵심적 특징은 개인의 신체 연령, 측정된 지능, 및 연

령에 적절한 교육이 주어졌을 때의 개인 표준화 검사나, 쓰기기술 기능평가도구로 측정된 쓰기기술이 기대치보다 저조하다는 것이다. 학습장애를 포함하여 학습문제를 가진 학생들은 글을 통하여 의사소통하고, 감정과 경험을 말하고, 타인을 설득하는 데 있어 다른 아동들보다 더 많은 곤란을 경험한다. 정상적 성취 또래들과 비교할 때, 이들은 발상을 전개하고 조직화하는 전략 및 쓰기 과정을 조절하고, 산출된 원문의 질의 감찰에 관한 지식에 대해 미성숙하고, 덜 개념화시키고 있다(Berninger & Whitaker, 1993).

쓰기문제의 발달이론적 진단

쓰기 문제를 발달이론적으로 보는 연구자들은 이를 발달적 맥락에서 이해하고, 각 단계별로 평가하는 변인이 달라져야 한다고 보았다(Berninger, Mizokawa & Bragg, 1991). 그들은 아동이 발달해감에 따라 신경학적인 변인, 언어적 변인, 인지적 변인의 영향을 차례로 받게 되는 쓰기성취의 제약이론을 주장하였으며, 이들의 연구를 통해 후에 Berninger 와 Swanson (1994)은 이 이론을 변형시켰다. 즉, 구속물들은 불연속적인 단계로 작용한다기보다, 이들 다양한 구속물들이 모든 쓰기 단계에 통해 작용하지만, 그 상대적인 가중치가 각 발달단계를 통해 변화한다는 것이다. 예로서, 국민학교 저학년은 신경발달적 제약의 영향력이 가장 크고, 국민학교 고학년은 언어적 구속물의 영향력이 증가하고, 중학교에서는 인지적 구속물(계획과 번역/교정)의 영향력이 증가하게 된다.

이러한 발달이론적 맥락에서 연구자들은 특수글씨장애(specific handwriting disability), 특수맞춤법장애(specific spelling disability), 특수작문장애(specific compositional disability), 그리고 다요인 쓰기기술장애(disabilities in multiple component writing skills) 등의 4가지 쓰기장애의 하위 진단을 제시하였다(Berninger et al., 1991).

학업 및 쓰기성취에 영향을 미치는 요인들

학업 및 쓰기에 영향을 미치는 인구통계학적 요인에 대한 연구는 활발히 이루어져왔다. 많은 연구자들이 인구통계학적 요인과 성취 불일치와의 관계에 초점을 맞추어 연구하였으며, 저성취자들은 일반적으로 어머니의 학력 수준이 낮으며, 가족 크기가 크고, 가난한 가정의 출생순위가 늦은 학생들 중에 많다고 간주되었다(Alberman, 1973; Bell, Abrahamson, & McRae, 1977; Belmont & Belmont, 1978; Broman, Bien, & Shaughnessy, 1985; Davie, Butler, & Goldstein, 1972; Davis, 1984; Werner, 1980). 그러나 이러한 일반화는 의심스러운 점이 많은데, 예를 들어, 사회경제적 지위(SES; socioeconomic status)와 성취불일치와 상관관계는 어떤 연구에서는 정적인 상관관계가 있는 것으로 나타났으나(Frankel, 1960; Roberts, 1962; Shaw & Brown, 1957), 다른 연구에서는 상관이 별로 없는 것을 볼 수 있다(Morrow & Wilson, 1961; Passow & Goldberg, 1958). 더욱이 역상관이 있는 것으로 나타난 연구도 있다(Curry, 1961; McQuary, 1954). 출생순위와 성취불일치의 관계에 대한 연구도 유의미한 상관에서 상관이 없음까지 다양한 결과를 보였다(Frankel, 1960; Passow & Goldberg, 1958; Oakland & Stern, 1989; Roberts, 1962; Shaw & Brown, 1957).

한편 Berninger와 Fuller(1992)는 쓰기를 처음 시작하는 국민학교 저학년을 대상으로 구어적 유창성(verbal fluency), 글씨를 잘 쓰는 정자적 유창성(orthographic fluency), 그리고 작문(written composition)에 있어서의 성차를 연구하였는데, 이들의 연구에서 남학생은 구어적 유창성에서 더 나은 수행을 보였고, 여학생은 정자적 유창성과 작문에서 산출된 단어와 질의 수가 더 많았다. 이 연구는 쓰기를 처음 시작하는 단계에서, 정자적 유창성이 언어적 유창성보다 더 중요할 수 있으며, 소년들이 쓰기장애의 유병률이 더 높을 수 있음을 시사하는 것이었다.

연구 목적

학교장면에서의 쓰기표현은 전학교 연령에 대해 모든 아동의 쓰기표현이 좀 더 향상될 수 있도록 도움을 제공이 요구되는 중요한 의사전달 기재이다. 학교장면에서 빈약한 쓰기능력 때문에 학습곤란을 경험하고, 심리교육적 평가가 의뢰되는 아동 수가 증가하고(National Assessment of Educational Progress, 1980 ; Scardamelia & Bereiter, 1986) 있음에도 불구하고 쓰기과정 및 쓰기장애에 대한 연구는 학교장면에서 그 중요성에 비하여 아직 초기 단계에 있다(Berninger et. al., 1991). 또한 학습장애에 대한 관심과 더불어 쓰기표현의 장애를 포함한 학습장애에 연구가 증가되고 있는 실정이라는 하나, 현재 우리나라에서 이루어진 학습장애에 대한 연구들로는 쓰기장애아에 대한 구체적인 도움을 주기에 미흡한 점이 많다. 따라서 본 연구에서는 기존의 연구 발견점들을 바탕으로 국민학교 연령을 대상으로 한 쓰기성취검사를 개발하여 쓰기성취를 다면적이고 위계적으로 평가할 수 있는 도구를 제작하고, 이를 국민학생들에게 실시하여 그 신뢰도와 타당도에 대한 기초적 자료를 수집하여, 이들 결과를 통해 그 적용가능성을 알아보고자 하였다.

방 법

연구 1 : 쓰기성취검사의 제작

대 상

서울 시내 2개국교의 교사와 학생을 대상으로 쓰기성취검사 제작을 위한 연구를 수행하였다. 표집의 대표성을 위하여 서울시내 9개 교육구 중 학교의 등급을 고려하여 A와 C급 학교를 각각 하나씩 임의로 선정하고, 각 학교에서 2, 4, 6학년 각 1개반 씩을 표집하였다. 교사 2명의 자문을 통해 쓰기성취검사의 세 가지 소검사의 문항선정과 수정을 하였으며, 학생 264명(남: 146명, 여: 118명)의 대상으로 글씨 예비검사

와 맞춤법 예비검사의 문항분석을 위한 연구를 수행하였다.

검사의 구성

글씨 검사 글씨가 정확하고 알아보기 쉽게 쓰였는지를 측정하는 검사이다. 예비검사에서는 한 음절로 된 무의미 글자 자극 60 문항을 학생들이 네모칸에 쓰도록 하였다. 그 결과, 무의미 글자가 친숙한 자극이 아니라는 점과, 쓰기칸에 대한 응답 양식이 전체 글씨 쓰기를 대표할 수 없다는 문제 등이 제기되어 본 검사에서는 유의미한 문장으로 글자자극을 선정하고, 쓰기칸을 대신하여 3칸의 줄에 응답하도록 하였다.

본검사의 문장 자극은 17개의 표본 동요 자극에서 일차적으로 선정하고, 이들 중에서 자모 낱자 및 자모 조합이 골고루 들어있는 지를 고려해서 최종적으로 한 소절을 선정하였다. 검사를 실시할 때는 칠판에 정해진 문장을 검사자가 정서한 후, 학생들이 5분 동안 이 문장을 검사지의 글씨검사란에 베끼도록 하였다. 글씨 산물의 평가는 글자의 형태와 안정성(홍은경, 1987), 띄어쓰기, 기울기, 선의 질, 글자 크기 및 배치의 다섯가지 오류 범주(Luftig, 1989)에서 이루어졌다.

맞춤법 검사 첫번째, 본검사를 위한 예비검사 60 문항의 제작을 위하여 먼저 “국민학교 교육용 어휘”(이웅백, 1987)와 “한국어 교육용어휘 연구”(이충우, 1994)에서 낱말의 난이도와 친숙도를 고려하여 총 130 개를 선정하고, 김민수, 고영근, 이승재 및 임홍빈(1993), 미승우(1993), 원영섭(1993), 그리고 최기호(1994) 등을 참조하여 선정된 어휘가 문장 내에서 쓰이도록 문항자극화한 다음, 맞춤법 항목별로 어휘를 분류하고, 각 항목당 1~3개의 문항을 취하여 총 60개의 문항자극을 선정하고, 마지막으로 국민학교 교사 2명의 자문을 통해 국민학교 실정에 맞도록 수정, 교정하여 예비문항화하였다.

두번째, 예비검사 60문항으로부터 본검사를 위한 20문항의 선정을 위해서 내적합치도, 전체 문항 상관, 문항난이도, 문항변별도, 그리고 논리적 내용 선정의

5단계를 거쳤다. 맞춤법 예비문항의 내적합치도는 .92였고, 모든 문항들의 전체 문항 상관은 모두 .90 이상으로 다른 문항들과 수행이 불일치하는 문항은 없었다. 난이도가 80% 이상인 문항과 20% 이하인 문항 25개를 제외시켰으며, 맞춤법 수행이 우수한 상위 27%와 이를 잘 못하는 하위 27%로 상위집단과 하위 집단으로 나누고, 실시된 모든 문항에 대하여 이 두 집단간 χ^2 검증을 실시하여 문항변별도를 산출하여 χ^2 의 확률이 .05 이하인 1개 문항을 제외시켰다. 마지막으로 단어의 학년 수준과 맞춤법 규정항목이 고루 표집되었는지를 고려하여 본문항 20개를 선정하였다.

작문 검사 문항 자극은 미국의 학습장애 진단 검사인 성취진단검사(Diagnostic Achievement Battery-2: DAB-2, Newcomer, 1990) 내의 작문 검사의 삽화 자극을 한국화하여 사용하였다. 이 삽화는 '토끼와 거북이'를 현대적으로 재구성한 연속적인 3개의 그림으로 구성되어 있으며, 본 연구에서는 이 삽화를 검사지에 제시하고 이를 토대로 그림이 제시된 순서에 따라 아동에게 15분 동안 글짓기할 시간을 주었다. 작문의 평가는 질적 기준과 양적 기준의 두 기준에서 채점되었다. 질적 기준은 DAB-2의 작문평가의 질적 기준 9개 중 한국어에 적합한 8개를 선정하였으며, 그 구체적 내용은 (a) 대상이 기술되었는가?; (b) 이야기의 순서(처음, 중간, 끝)가 있는가?; (c) 인물이나 차에 적절한 명명을 하였는가?; (d) 유머를 사용하였는가?; (e) 회사, 경주, 또는 후원자에 적절한 명명을 하였는가?; (f) 도덕적/철학적 주제가 있는가?; (g) 문단이 사용되었는가?; (h) 대화가 사용되었는가? 등이다. 그리고, 양적인 평가의 기준은 산출된 총 단어수, 총 문장수의 두 가지 기준을 선정하였다.

연구 2 : 검사에 대한 신뢰도와 타당도 연구

대 상

연구 2는 서울시의 2개 국교 2학년(90명), 4학년(87명), 6학년(79명) 학생 256명(남: 128명, 여: 128명)을

대상으로 수행하였다. 가정환경자료는 2개 국교 중 8 국민학교 132명을 대상으로 부모의 학력과 형/누나의 유무에 대해 수집되었다.

검사도구 및 절차

제작된 쓰기성취검사의 신뢰도와 타당도를 검증하기 위하여 1995년 5월 13일부터 5월 27일에 걸쳐 검사를 시행하였다. 1회 조사에서는 검사자가 학급에 들어가서 BGT 검사(이상로와 변창진 및 이희도 1986; Hutt 1985; Koppitz 1975), 제작된 쓰기성취검사, 그리고 기초학습기능검사의 철자재인 소검사(박경숙과 윤점룡 및 박효정, 1989)의 순서로 지필검사형태로서 집단실시하였다. 쓰기성취검사의 소요시간은 약 30분(글씨검사: 5분, 맞춤법 검사: 10분, 작문검사: 15분)이었다. 검사시간은 예비연구를 통해 경험적으로 결정되었다.

예비검사 쓰기성취검사가 끝난 후 각 담임교사와 시각디자인 전문가 1인에게 글씨검사 수행에 대한 A(4), B(3), C(2) 및 D(1)의 4점 척도의 평정을 의뢰하였다. 또한 담임교사에게는 아동에 대한 가정환경조사서와 국어교과의 성적의 기록을 의뢰하였다. 작문 검사의 수행은 학년별로 중학교 국어교사 3명을 통해 A(4), B(3), C(2), 및 D(1)의 4점 평정을 의뢰하여 자료를 수집하였다. 2회 조사를 통해 쓰기성취검사의 검사-재검사 신뢰도 검증을 하였는데 1회 연구대상자 중 8 국민학교의 132명에 대해 1회 조사 2주 후에 제작된 쓰기성취검사를 재실시하였다.

자료분석

신뢰도 검증을 위해서 각 소검사의 내적합치도, 검사-재검사 신뢰도 계수, 글씨검사 채점의 평정자간 일치도를 산출하였다. 타당도 검증을 위해서 준거관련 검사 및 구인과의 상관을 검증하고, 성변인과 가정환경변인에 따른 집단간 수행의 차이를 검증하였다.

글씨검사의 평정자간 신뢰도는 1차 조사의 검사응답 중 각 학년에서 15개의 응답을 무선표집하여 검

증하였고, 검사-재검사 신뢰도 산출, 그리고 가정환경 변인 및 국어성적과의 상관 분석은 사 국민학교학생 132명을 대상으로 이루어졌으며, 이들 이외의 통계분석은 전체 연구대상자에 대하여 이루어졌다. 모든 통계분석을 위해 SPSS-PC Windows 프로그램이 사용되었다.

결 과

신뢰도

쓰기성취검사의 각 하위검사에 대한 내적 합치도 계수 Cronbach's α 는 글씨 검사는 .83, 맞춤법 검사는 .87, 그리고 작문 검사는 .90이었다. 글씨 검사 평정의 평정자 간 일치도 Pearson 상관계수(r)은 통계적으로 모두 유의미하였으며($p < .001$), 구체적으로 형

태와 안정성, .65~.67; 띄어쓰기, .93~.98; 선의 질, .60~.65; 행의 기울기, .60~.81; 크기 일관성, .45~.74이었다. 본 연구의 피검자들 중 132명에 대해서 2 주일의 간격을 두고 재검사를 실시한 결과, 검사-재검사 신뢰도는 글씨검사에서는 중위상관을 가지고 있는 것으로 나타났고(형태와 안정성, .54; 띄어쓰기, .58; 선의 질, .59; 행의 기울기, .65; 크기 일관성, .47), 맞춤법 검사에서는 .88, 작문 검사는 질적 기준에서는 .57, 단어수와 문장수의 두 양적 기준에서는 모두 .80이었다.

타당도

교사평정 및 준거관련 검사와의 상관

쓰기성취의 준거평정치로 사용된 글씨 및 작문수행에 대한 교사평정이 객관적인가를 알아보기 위해 122

표 1. 각 하위 검사와 기준 측정도구 및 평정의 상관(N=253)
(교사평정과 국어성적은 학년 변인의 영향을 통제함)

	글씨교사 평정	작문교사 평정	BGT	기초철자 재인	국어성적 (N=130)
글씨 검사					
형태와 안정성	-.40***		.26***	-.34***	-.29**
띄어쓰기	-.35***		.28***	-.37***	-.14
선의 질	-.56***		.36***	-.47***	-.28**
행의 기울기	-.28***		.23***	-.41***	-.18*
크기 일관성	-.40***		.35***	-.44***	-.17*
맞춤법 검사			-.38***	.77***	.32***
작문 검사					
질적 기준		.34***	-.19**	.49***	.32***
양적 기준					
단어 수		.58***	-.24***	.57***	.24**
문장 수		.44***	-.23***	.49***	.23**

글씨검사와 BGT는 오류 점수

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

명의 자료를 대상으로 한 글씨 검사의 평정자간 일치도는 .53~.86($p < .001$)이었고, 작문 검사의 교사평정 일치도는 .75~.76 ($p < .001$)이었다. 규준평정과의 상관분석에서 학년별로 평정된 글씨 및 작문 검사의 교사평정은 학년변인을 통제하였다. 결과는 표 1에 제시되어 있으며, 구체적으로 살펴보면 다음과 같다.

글씨 검사는 오류점수로 나타나므로 같은 오류 검사인 BGT와의 정적 상관을 제외하고는 모두 부적 상관을 보였으며, 교사평정과의 상관이 있어서 중위상관을 보인 선의 질과의 상관(-.56, $p < .001$)을 제외하고는 다른 하위기준들과의 상관은 중위상관보다 낮게 나타났다(형태와 안정성, -.40, $p < .001$; 띄어쓰기, -.35, $p < .001$; 행의 기울기, -.28, $p < .001$; 크기 일관성, -.40, $p < .001$). 그리고 BGT 점수와 글씨 검사의 상관은 모두 통계적으로 유의하였으나($p < .001$), 중위상관보다 낮게 나타났다(형태와 안정성, .26; 띄어쓰기, .28; 선의 질, .36; 행의 기울기, .23; 크기 일관성, .35).

맞춤법 검사는 기초학습기능검사의 철자재인검사 점수와는 상관이 높게 나타났으나(.77, $p < .001$), BGT

점수와는 중위 상관보다 낮게 나타났다(-.38, $p < .001$). 작문검사는 작문교사평정과의 상관이 있어서 단어 수에서는 중위상관을 나타냈으며(.58, $p < .001$), 질적기준(.34)과 문장 수(.44)는 중위 상관보다 낮게 나타났다($p < .001$). 작문 검사와 BGT는 중위상관보다 낮게 나타났다(질적기준, -.19, $p < .01$; 단어수, -.24, $p < .001$; 문장수, -.23, $p < .001$).

학년과의 상관

구인타당도는 검사가 측정하는 특성이 그 검사가 근거한 이론적 모형을 반영하는 정도와 관련되어 있다. 연구에서 쓰기성취검사의 구인타당도를 알아보기 위하여 학년과의 상관, 교육과정 중 국어교과성적과의 상관, 소검사 간의 상관 등을 알아보았다.

표 2에 쓰기성취검사의 모든 소검사 하위기준들에서 학년이 올라갈 수록 수행을 잘하게 됨을 나타내고 있다. 학년 변인이 성취점수에 어떤 효과를 주는 지 알아보기 위해 일원 변량분석을 실시하였고, 사후 검증으로 Scheffé 검사를 .05수준에서 실시하였다. 분석 결과 글씨검사에서는 형태와 안정성[$F(2.251)=$

표 2. 학년별 쓰기성취 점수의 평균과 표준편차

학 년	2(N=88)	4(N=86)	6(N=80)	F
글씨 검사				
형태와 안정성	6.37 (3.40) ^a	5.03 (3.12) ^b	3.91 (2.35) ^b	14.03*
띄어쓰기	1.28 (1.81) ^a	.73 (1.51) ^b	.15 (.45) ^c	13.67*
선의 질	2.25 (.72) ^a	1.80 (.73) ^b	1.40 (.59) ^c	32.50*
행의 기울기	2.31 (.70) ^a	1.80 (.68) ^b	1.50 (.60) ^c	31.96*
크기 일관성	2.19 (.66) ^a	1.90 (.63) ^b	1.64 (.70) ^c	14.80*
맞춤법 검사	6.58 (3.24) ^a	11.67 (3.37) ^b	14.88 (3.48) ^c	130.86*
작문 검사				
질적인 기준	2.37 (1.49) ^a	4.01 (1.34) ^b	4.71 (1.65) ^c	55.41*
양적인 기준				
단어 수	32.73 (16.99) ^a	61.66 (24.48) ^b	99.35 (37.51) ^c	124.98*
문장 수	5.13 (3.46) ^a	7.88 (3.29) ^b	12.59 (5.62) ^c	66.74*

() 안은 표준편차

* $p < .0001$

같은 영문자는 $p < .05$ 수준에서 집단 간 차이가 없다는 것을 의미한다.

14.03], 띄어쓰기[F(2,251)=13.67], 선의 질[F(2,251)=32.50], 행의 기울기[F(2,251)=31.96], 크기 일관성[F(2,251)=14.80]의 모든 기준에서 2, 4, 6학년 집단의 점수 차이가 유의미하였다($p < .001$). 맞춤법 검사[F(2,251)=130.86, $p < .001$]에서도 집단 간 차이가 유의미하였다.

또한 작문 검사에서도 질적기준[F(2,251)=55.41], 양적기준의 단어 수[F(2,251)=124.98], 문장 수[F(2,251)=66.74]에서 모두 학년집단 간 점수 차이가 유의미하였다($p < .001$). Scheffé 검증 결과 글씨 검사의 형태와 안정성 기준을 제외하고는 2학년과 4학년, 2학년과 6학년, 그리고 4학년과 6학년 간에 모두 차이가 있었다. 형태와 안정성 기준의 경우는 2학년과 4학년, 2학년과 6학년 간의 차이가 있었지만, 4학년과 6학년 간에는 차이를 보이지 않았다.

또한 학년과의 상관을 살펴 볼 때, 표 3에서처럼 학년과 유의한 상관이 있었다($p < .001$). 각 소검사별로 분석해 보면, 글씨 검사는 학년과의 상관이 중위 상관보다 낮게 나타났다(형태와 안정성, -.32; 띄어쓰기, -.31; 선의 질, -.45; 행의 기울기, -.45; 크기 일관성, -.32). 글씨 검사는 오류 점수로 나타나므로 부적인 상관이 있다. 맞춤법 검사와 학년과는 높은 상

관을 보였다(.71). 작문 검사는 양적 기준 중의 단어 수와 학년과 높은 상관이 있었으며(.71), 학년과 질적 기준(.54) 및 양적 기준 중의 문장 수(.58)의 상관에서는 중위 상관이 발견되었다. 결과적으로 각 소검사 하위기준들에서 학년과의 상관이 유의하였으며, 이는 쓰기성취검사의 구인타당도를 지지하는 결과이다.

국어성적과의 상관

국어성적은 쓰기성취와 관련이 있을 것으로 기대되는 학업지표이다. 앞에서 제시된 표 1에서 국어성적과 쓰기성취검사의 각 소검사들 간의 상관이 나타나 있다. 띄어쓰기를 제외한 모든 소검사에서 국어성적과의 상관이 통계적으로 유의미하였다. 글씨 검사는 오류점수로 나타나기 때문에 국어성적과는 부적인 상관을 보였다(형태와 안정성, -.29, $p < .01$; 선의 질, -.28, $p < .01$; 행의 기울기, -.18, $p < .05$; 크기 일관성 -.17, $p < .05$). 맞춤법 검사와(.32, $p < .001$) 작문 검사의 모든 하위기준들(질적기준, .32, $p < .001$; 양적기준의 단어수, .24, $p < .01$; 양적기준의 문장 수, .23, $p < .01$)의 상관은 유의미하였다.

표 3. 소검사 및 하위기준 점수들과 학년과의 상관

소검사 및 하위기준	r
글씨 검사 형태와 안정성	-.32*
띄어쓰기	-.31*
선의 질	-.45*
행의 기울기	-.45*
크기 일관성	-.32*
맞춤법 검사	.71*
작문 검사 질적 기준	.54*
양적 기준	
단어 수	.71*
문장 수	.58*

글씨 검사는 오류 점수 * $p < .001$

표 4. 성별 쓰기성취검사 점수 평균(N=251)

	남(N=125)	여(N=126)	t
글씨 검사			
형태와 안정성	5.91(3.18)	4.37(2.96)	3.97***
띄어쓰기	.90(1.74)	.59(1.15)	1.66
선의 질	2.06(.76)	1.60(.71)	4.91***
행의 기울기	2.08(.73)	1.67(.70)	4.50***
크기 일관성	2.12(.69)	1.71(.65)	4.90***
맞춤법 검사	10.05(4.61)	11.74(4.70)	-2.88**
작문 검사			
질적요인	3.56(1.74)	3.75(1.81)	-.83
양적요인			
단어 수	58.00(34.02)	68.91(41.46)	-2.28*
문장 수	7.77(4.76)	8.96(5.50)	-1.83

() 안은 표준편차, * $p < .05$ ** $p < .01$ *** $p < .001$

성별 쓰기성취검사 점수

제작된 쓰기성취검사 점수와 아동의 성변인에 따른 수행의 차이를 분석하였으며, 이는 표 4에 제시되어 있다. 여학생의 수행이 남학생의 수행보다 많은 기준들에서 유의하게 높은 것이 확인되었다. 글씨 검사에서는 띄어쓰기기준을 제외하고, 형태와 안정성($t_{249}=3.97, p < .001$), 선의 질($t_{249}=4.91, p < .001$), 행의 기울기($t_{249}=4.50, p < .001$), 크기 일관성($t_{249}=4.90, p < .001$) 등의 하위기준들에서 여학생의 수행이 더 우수하였다. 맞춤법 검사에서도 여학생의 점수가 더 높았으며($t_{249}=-2.88, p < .01$), 작문 검사에서는 단어 수에서($t_{249}=-2.28, p < .05$) 여학생의 수행이 우수하였고, 다른 기준들에서는 차이가 유의미하지 않았다.

가정환경 변인별 쓰기성취검사 점수

제작된 쓰기성취 검사로 가정환경 변인과 쓰기성취와의 관계를 살펴보았다. 분석에 사용된 가정환경 변인은 부모의 학력 수준과 형이나 누나의 유무 여부였다. 아버지의 학력은 대졸이상, 고졸, 중졸이하의 세 수준으로 나누고, 어머니의 학력은 고졸이상과 중졸이하의 두 수준으로 나누었다. 결과, 부모의 학력 수준에 따라서는 쓰기성취에 차이가 없는 것으로 나타났다. 형이나 누나의 유무에 따른 집단간 소검사별 수행에서도 집단간 점수차이가 발견되지 않았다.

논 의

제작된 쓰기성취검사의 적용가능성을 알아보기 위해 서울 시내 국민학교 2, 4, 6학년 학생 256명에 대해 검사를 실시하였다. 검사의 신뢰도는 문항의 내적합치도, 채점자간 평정의 일치도, 그리고 검사-재검사 신뢰도를 산출하여 검증하였다. 제작된 각 소검사 내 문항들 간의 동질성 정도를 알아보기 위해 산출한 Cronbach α 계수는 글씨 검사에서 .83, 맞춤법 검사에서 .87, 작문 검사에서 .90으로 내적합치도가 양호하였다. 글씨 검사의 평정자간 일치도를 소검사별로 살펴보면 글씨 검사의 평정자간 일치도는 높은 일치

도를 보인 띄어쓰기(.93~.98)를 제외하고는 대체적으로 중상위의 상관이 나타났다(.45~.81, $p < .001$). 검사-재검사 간의 Pearson 상관계수(r)를 산출한 결과, 글씨 검사는 중위상관 이상이 나타났고(.47~.65), 맞춤법 검사는 높은 상관이 나타났으며(.88), 작문 검사에서는 중상의 상관이 나타난 질적 기준(.57)에서보다 두 양적 기준의 상관이 높았다(단어수, .80; 문장수, .80). 이러한 결과로부터 맞춤법 검사 및 작문 검사의 양적 기준은 시간 간격에 상당히 안정적임을 확인할 수 있었다. 신뢰도 검증에서 글씨 검사는 성취검사가 요구하는 만족할 만한 신뢰도를 보이지 못했다. 이것은 채점 절차와 검사절차에서 그 원인을 찾아 볼 수 있을 것이다. 본 연구에서 글씨 검사는 검사자가 칠판에 글씨검사 자극을 판서하고 5분간 이를 베끼도록 지시한다. 다른 두 소검사들이 비교적 표준화된 수행시간을 가지고, 능력검사적 성격인데 반해 글씨 검사는 속도검사적이어서 시간변인을 측정할 수 없는 집단검사로는 평가상의 한계가 있었다. 또한 칸응답보다는 줄응답에서 채점시 객관적인 평정이 어려워지는 문제가 있다. 한편, 작문 검사 하위기준들의 1, 2차 점수들 간의 상관에서는 질적 기준의 상관이 다른 두 기준들에 비해 시간 변화에 덜 안정적이었다. 이것은 단어 수나 문장 수와 같은 양적 기준은 주어진 검사 시간 때문에 상한계선을 가지는 데 반해, 주제라든가, 인물이나 대상에 대한 명명과 같은 질적 기준은 퇴고(revision)의 과정에서 향상될 수 있는 점을 생각해 볼 수 있다. 즉, 2주의 시간 간격이 잠재적인 퇴고의 효과로 작용하여 작문수행이 질적으로 향상되었을 가능성이 시사된다.

쓰기성취검사의 타당도의 검증은 쓰기검사수행과 교사평정 및 관련 준거검사, 학년, 교과과정 중 국어 교과성과 상관, 그리고, 성변인 및 가정환경변인에 따른 집단간 수행의 차이의 분석을 통해 이루어졌다. 각 준거와 쓰기성취검사의 소검사들 간의 상관은 모두 유의미하였다. 특히, 맞춤법 검사와 기초학습기능 검사의 철자재인검사의 상관은 다른 상관에 비해 매우 높은 수준이었다. 이는 이 두 검사가 정답이 분명한 객관적인 평가도구라는 데서 그 원인을 찾을 수

있으며, 제작된 맞춤법 검사의 타당도를 지지하는 증거가 될 수 있다. 한편, 관련준거 검사 중에서 BGT는 쓰기성취검사와의 상관관이 가장 낮은 범위에 있었다. 그러나, BGT가 학업성취에 대해서 보이는 상관의 범위가 $-.34$ (이상로 등, 1986)인 점을 감안할 때 이러한 상관 범위(.23~.38, $p < .001$)가 결코 관련이 미약함을 시사한다고는 볼 수 없다. 또 다른 설명으로는 본 연구에서 선택한 BGT 채점에는 Koppitz의 발달적 채점법이 적용되었으며, 이것은 5~10세의 아동에 적합한 것으로 본 연구에서 표집된 2, 4, 6학년 중 2, 4학년 만이 신뢰롭게 평정될 수 있는 채점 체계였다. 따라서 4학년 이상에서는 채점에 변별도가 없을 것으로 가정할 수 있으며, 이런 영향 때문에 쓰기성취검사와의 상관관이 낮아졌을 것이라고 가정할 수 있다. 앞으로의 연구에서 준거 평정과의 상관 분석을 위하여 국민학교 연령 집단을 모두 포괄할 수 있는 검사도구를 사용하는 것이 필요할 것이다. 글씨와 작문 검사의 교사평정치와 쓰기성취점수와의 상관은 중위상관 이하로 나타났다. 담임교사나 중학교 국어교사의 평가는 비교적 국민학교 아동 평가에 공정할 것으로 볼 수 있지만 이를 객관적 평가도구로 사용하기에는 부족한 점을 간과할 수 없다. 따라서, 앞으로의 연구에서는 교사평정치를 수집할 때에도 보다 객관적인 평가기준을 통한 평정을 의뢰하는 것이 요구된다.

쓰기성취점수의 2, 4, 6학년에서의 집단 간 수행의 차이를 분석한 결과 모든 소검사와 하위기준에서 학년간 의미있는 차이가 발견되었으며, 학년이 올라감에 따라 수행이 향상되고 있었다. 이는 쓰기의 발달 이론적 모형을 지지하는 결과로 볼 수 있으며, 제작된 검사의 타당도를 입증하는 증거였다. 국어성적과 쓰기성취점수와의 상관분석 결과, 국어성적과 글씨 검사 내 띄어쓰기를 제외한 모든 기준들에서 중위보다 낮은 상관관이 나타났다($p < .05$). 이 결과로부터 국어교과인 쓰기능력에 대한 평가가 충분히 교과성적에 반영되지 못하고 있음을 가정해 볼 수 있다.

아동의 성변인에 따른 쓰기성취에 대한 영향은 여학생의 수행이 남학생의 수행보다 많은 소검사 기준

들에서 유의하게 높은 것이 확인되었다. 이는 여학생이 남학생 보다 문어적 표현력이 더 우수하다는 선행연구를 지지하는 또 결과이며, 제작된 쓰기성취검사의 타당성을 지지하는 자료라고 할 수 있을 것이다(Berninger et al., 1992). 가정환경 변인과 쓰기성취와의 관계를 살펴보았는데, 최근의 연구 결과(Oakland et al., 1989)와 일치한 결과를 얻었다. 이런 결과를 통해 쓰기성취검사의 또 다른 지지 증거를 발견할 수 있었다.

각 소검사별로 연구의 의의와 요구되는 추후 연구과제는 다음과 같다.

글씨 검사는 신뢰도와 타당도가 그다지 높지는 않았으나 집단 검사로서 실시되었음을 감안한다면 글씨 쓰기에 대한 비교적 객관적인 평가도구라고 볼 수 있다. 아동의 글씨를 평가하는 것은 아동의 글씨 결함을 확인하는 것 뿐만 아니라, 최근 인지과학 등의 분야에서 이루어지는 음성인식과 마찬가지로 글자체를 인식하는 연구와 성인의 글자체의 변형과 그 발달 과정 등을 연구하는 데에도 기초자료로 쓰일 수 있을 것이다. 추후 연구에서는 첫째, 보다 객관적인 채점 기준과 방법을 마련해야 하겠다. 이를 위해서는 하위 기준들 중 평정에 가장 영향력 있는 변인을 밝혀내거나, 각 하위기준이 글씨 평정에 미치는 영향에 관한 경로를 파악하는 연구들이 요구된다. 둘째, 본 연구가 집단으로 실시되었기 때문에 고려되지 못한 속도 등의 변인들을 반영하기 위한 개인검사 연구가 요구된다.

맞춤법 검사는 비교적 높은 신뢰도와 타당도를 가지고 있었으며, 이는 국민학생의 맞춤법 성취를 평가하고, 치료 및 교정하는 연구 등에 직접적으로 사용될 수 있음을 시사한다. 앞으로의 연구에서는 맞춤법을 측정하는 방법들을 다양화하고, 이들 각각의 검사양식에 따른 장단점 등을 파악하는 일이 요구된다. 이들을 아동의 맞춤법 능력을 측정하는 데 사용하므로써 좀 더 다각적인 맞춤법 능력의 측정이 이루어질 수 있을 것이다.

작문 검사는 가장 높은 단계에 있는 쓰기표현인 작문 능력을 쓰기성취의 종합적인 검사 도구 내에서 측

정하였다는 데 그 의의가 있으며, 국민학생의 작문능력에 대한 연구자료로 사용할 수 있다. 앞으로의 연구에서는 인지 능력을 측정하는 채점 기준이 강화되어야 하고, 다양한 작문 양식으로 평가하는 연구가 필요할 것이다.

요약하면, 본 연구에서는 글씨, 맞춤법, 작문의 세 측면을 쓰기표현의 대표적 요소로 보고 이들 능력을 각각 측정할 수 있는 쓰기성취검사를 개발하였다. 이는 쓰기표현의 다면성과 위계성을 반영하는 것이다. 본 연구의 의의는 한글에 있어서 최초로 쓰기의 여러 측면을 체계적이고 통합적으로 평가하기 위한 도구를 개발하였으며, 이를 국민학생에게 실시하여 그 적용가능성에 대한 기초적인 자료를 수집한 점에서 찾을 수 있다. 추후 연구를 통하여 검사가 세련화되고 표준화 작업이 이루어진다면 쓰기장애를 포함한 쓰기문제를 가진 아동들을 위한 평가 및 지도 등의 봉사전달을 좀더 효과적이고 효율적으로 제공할 수 있을 것이다.

참고문헌

- 김민수 · 고영근 · 이승재 · 임홍빈. (1993). 금성판 국어대사전. 서울: 금성출판사.
- 미승우. (1993). 새 맞춤법과 교정의 실제(증보판). 서울: 어문각
- 박경숙 · 윤점룡 · 박효정. (1989). 기초학습기능검사. 서울: 한국교육개발원.
- 원영섭 편. (1993). 초·중·고 국어 교과서에 나타난 띄어쓰기·맞춤법 용례. 서울: 세창출판사.
- 이상로 · 변창진 · 이회도. (1986). BGT 성격진단법. 서울: 중앙적성연구소.
- 이용백. (1987). 국민 학교 교육용 어휘. 서울: 국어연구소
- 이충우. (1994). 한국어 교육용어휘 연구. 서울: 국학자료원.
- 최기호. (1994). 한글맞춤법 새 길라잡이. 서울: 토담.
- 홍은경. (1987). 유아의 한글 쓰기에 있어서의 오류 및 글자의 안정성에 관한 연구. 중앙대학교 석사학위 청구논문.
- Alberman, E., (1973). The early prediction of learning disorders. *Developmental Medicine and Child Neurology*, 15, 202-204
- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders*(4th ed.) Washington, DC, American Psychiatric Association.
- Beers, J., & Henderson, E. (1977). A study of development orthographic concepts among first grade children. *Research in the Teaching of English*, 11, 133-148.
- Bell, A.E., Abrahamson, D.S., & McRae, K.L. (1977). Reading retardation: A 12 year prospective study. *Journal of Pediatrics*, 91, 363-370.
- Belmont, I., & Belmont, L. (1978). Stability or change in reading achievement over time: Developmental and educational implications. *Journal of Learning Disabilities*, 11, 31-39.
- Berninger, V.W., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology*, 30, 363-382.
- Berninger, V.W., & Hart, T. (1993). From research to clinical assessment of reading and writing skills in primary and intermediate grade students. *Journal of Educational Psychology*, 85(3), 478-508.
- Berninger, V.W., & Hooper, S.R. (1993) Prevention and remediating writing disabilities: Interdisciplinary frameworks for assessment, consultation, and intervention. *Journal of School Psychology*, 29, 590-594.
- Berninger, V.W., & Swanson, H.L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In E. Butterfield (Ed.), *Children's writing: Toward a*

- process theory of development of skilled writing*.
Greenwich, CT : JAI Press.
- Berninger, V.W., & Whitaker, D. (1993). Theory-based branching diagnosis of writing disabilities. *School Psychology Review*, 22(4), 623-642.
- Berninger, V.W., Mizokawa, D.T. & Bragg, R. (1991). Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychology*, 29, 57-79.
- Broman, S., Bien, E., & Shaughnessy, P. (1985). *Low achieving children : The first seven years*. Hillsdale, NJ : Erlbaum.
- Curry, R.L. (1961). Certain characteristics of under- and overachievers. *Peabody Journal of Education*, 39, 41-45.
- Davie, R., Butler, N., & Goldstein, H. (1972). *From birth to seven : A report of the National Child Development Study*. London : Longman.
- Davis, F. (1984). Understanding underachievers. *American Education*, 20, 12.
- Frankel, E. (1960). A comparative study of achieving and underachieving high school boys of high intellectual ability. *Journal of Educational Research*, 53, 172-180.
- Graham, S., & Miller, L. (1983). Spelling research and practice : A unified approach. In E.L. Meyern, G.A. Vergason, & R.J. Whelan (Eds.), *Promising practices for exceptional children*. Denver : Love.
- Greene, H.A., & Petty, W.T. (1967). *Developing language skills in the elementary schools*. Boston, MA : Allyn and Bacon.
- Hammill, D.D. (1986). Problem in the written composition. In D.D. Hammill & N.R. Bartel (Eds.), *Teaching students with learning and behavior problems* (3rd ed.) Boston : Allyn and Bacon.
- Hammill, D.D., & Larsen, S. (1988). *Test of Written Language*(2nd ed.). Austin, TX : Pro-Ed.
- Henderson, E. (1985). *Teaching spelling*. Boston, MA : Houghton Mifflin.
- Hutt, M.L. (1985). *The Hutt Adaptation of the Bender-Gestalt Test. Rapid Screening and Intensive Diagnosis*. 4th ed. Orlando : Grune & Stratton.
- Koppitz, E.M. (1975). *The Bender Gestalt Test for Young Children. Vol. II. Research and Application, 1963-1973*. NY : Grune & Stratton.
- Luftig, R.L. (1989). *Assessment of learners with special needs*. Boston, MA: Allyn and Bacon.
- Mann, P.H., Suiter, P.A., & McClung, R.M. (1987). *Handbook in Diagnosing-Prescriptive Teaching* (3rd ed.). Boston, MA : Allyn and Bacon.
- McQuary, J.P. (1954). Some differences between over- and underachievers in college. *Educational Administration and Supervision*, 40, 117-120.
- Morrow, W.R., & Wilson, R.C. (1961). Family relations of bright high-achieving and underachieving high-school boys, *Child Development*, 32, 501-510.
- National Assessment of Educational Progress. (1980). *Writing Achievement : Result from the third national writing assessment* (Vol. 1 : 17-year-olds ; Vol. 2 : 3-year-olds ; Vol. 3 : 9-year-olds). Denver : National Assessment of Educational Progress. (Eric Document Reproduction Service No. ED 196 042, ED 196 043, and ED 196 144)
- Newcomer, P.L. (1995). *Diagnostic Achievement Battery*(2nd ed.). Austin, TX : Pro-Ed.
- Oakland, T. & Stern, W. (1989). Variables associated with reading and math achievement among a heterogeneous group of students. *Journal of School Psychology*, 27, 127-140.
- Passow, A.H., & Goldberg, M.L. (1958). Study of underachieving gifted. *Educational Leadership*, 16, 121-125.

- Polloway, E.A., & Patton, J.R., & Cohen, S.B. (1983). Written language for mildly handicapped students. In E.L. Regen, G.A. Vergason, & R.J. Whelan (Eds.) *Promising practices for exceptional children*. Denver : Love.
- Roberts, H.E. (1962). Factors affecting the underachievement of bright high school students. *Journal of Educational Research*, 56, 176-183.
- Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In M.Wittrock (Ed.), *Handbook on research on teaching* (3rd de. ; pp 778-803). New York ; Macmillan.
- Shaw, M.C., & Brown, B.J. (1957). Scholastic underachievement of bright college students. *Personnel and Guidance Journal*, 36, 195-199.
- Smith, F. (1982). *Writing and the writer*. New York : Holt, Rinehart & Winston.
- Weiner, E.S. (1980). Diagnostic evaluation of writing skills. *Journal of Learning Disabilities*, 13, 48-53.
- Werner, E.E. (1980). Environmental interaction in minimal brain dysfunctions. In H.E. Rie & E.D. Rie(Eds.), *Handbook of minimal brain dysfunction : A critical view*. New York : Wiley.

Development of Writing Achievement Test : A Preliminary Study

Hee-Ryoung Yun and Jong Chul Han

Department of Psychology
The Graduate School of Yonsei University

The purpose of this study was to develop a writing achievement battery and to evaluate its applicability. The study was divided in two phases. The primary focus of Study 1 was the development of the Writing Achievement Battery consisting of three subtests: Handwriting, Spelling, and Composition. The items of the three subtests were selected through consultation with elementary school teachers. The preliminary Handwriting and Spelling subtests were administered to 264 elementary school children in 2nd, 4th, and 6th grade. Through this procedure the response method of the Handwriting subtest was determined and the items of the spelling subtest were selected through item-analysis.

Study 2 was conducted to examine the reliability and validity of the Writing Achievement Battery constructed in Study 1. The battery was administered to 256 elementary school children. To examine the reliability of the WAB, Cronbach's α coefficient, the consistency between evaluators in the Handwriting subtest, and the test-retest correlation coefficient were calculated. The consistency of the H: .83, S: .87, C: .90, suggested an internal consistency of the subtest. The .45 to .98 consistency between evaluators indicated that consistency was above average. The test-retest coefficients .47 to .65, .89, .57 to .80, for H, S, C, respectively indicated that the Spelling subtest was generally stable and the other two subtests were less stable.

For the criterion-related validity, the scores from the WAB were compared with those of the teachers evaluations of the Handwriting Subtest and the Composition Subtest, the BGT, the Spelling Recognition subtest of the Basic Learning Functioning Battery, and the students' grade in the subject of Korean. There were significant correlations between the three subtests of the WAB and the criteria and constructs. The group differences according to gender and SES(Socio-Economic Status) were also analyzed. There were significant group differences across gender and there were no group differences in SES. This evidence supported the validity of the WAB.