

쓰기장애 위험군 학생의 조기선별 정확성을 위한 그림-단어제시검사의 신뢰도와 예측 타당도 연구: 2년간의 종단자료*

여 승 수**

부산교육대학교 초등특수 전공

손 지 영***

대전대학교 중등특수교육과

《 요 약 》

본 연구는 쓰기 위험군 초등학교 저학년(1,2학년) 학생들을 조기에 선별할 수 있는 목적으로 사용이 가능한 저학년용 쓰기검사 도구를 개발하고 측정학적 적합성을 평가하는 주된 목적을 갖고 있다. 특히 본 연구에서는 그림-단어제시검사를 활용하여 쓰기 위험군 학생의 조기선별 여부를 검증하고 있다. 본 연구에 참여한 연구대상 학생은 A시에 소재하고 있는 B 초등학교 42명이었다. 검사에서 사용된 분석방법은 신뢰도와 예측타당도를 분석하기 위해서 피어슨 상관계수를 사용하였으며 진단의 정확성을 측정하기 위해서 ROC 곡선 분석방법이 사용되었다. 본 연구결과에 따르면, 그림-단어제시검사는 양호한 신뢰도와 타당도를 나타내는 것으로 확인되었다. 또한 ROC 곡선 분석에서 제공된 절단점 점수를 활용할 경우 진단의 정확성이 양호한 것으로 확인되었다. 마지막으로 본 연구의 의의와 제한점에 대하여 논의하였다.

주제어 : 쓰기, 그림-단어제시검사, 교육과정중심측정, 신뢰도, 예측 타당도, 진단의 정확성

* 이 논문은 2012년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 연구되었음 (NRF-2012S1A5A8023659).

** 제1저자

*** 교신저자(sonjy@dju.kr)

1. 서론

1. 연구의 필요성

쓰기는 많은 학령기 학생들에게 있어서 실제로 어려운 과제중 하나이며, 일부 학생들은 쓰기에서의 어려움을 성인이 된 이후에도 지속적으로 경험하고 있다. 국내의 경우 관련된 쓰기 분야의 연구가 부족하여 쓰기장애를 갖고 있는 학생을 추정하기가 어려운 상황이지만, 2002년도에 실시한 미국의 국가학업성취도 결과에 따르면 초등학교 4학년의 경우 72%가 국가가 정한 일정 기준의 쓰기 능력을 성취하지 못하는 것으로 나타났다(Persky, Daane, & Jin, 2003). 그리고 2007년도 미국 국가학업성취도 결과를 살펴보면, 8학년의 65%와 12학년의 75%가 여전히 낮은 수준의 쓰기 능력을 가지고 있다고 보고하고 있다(Salahu-Din, Persky, & Miller, 2008). 이러한 결과를 볼 때, 학령 초기에 쓰기 능력의 부족은 학년이 올라가도 계속 누적적으로 지속되어 쓰기의 어려움을 지속적으로 경험하고 있는 것을 알 수 있다.

쓰기 능력은 학생들이 아는 것을 의사소통 하는 능력과 전반적인 문해력(literacy) 발달에 중요한 영향을 미치는 요소이다(Graham & Perin, 2007). 또한 쓰기는 정보를 분석하고 평가하는 데 사용되면서 학생들이 지식을 통합하고 사고하는 데 영향을 미친다(Shanahan, 2004). 덧붙여 학령기 이후에 쓰기 유창성은 직업 능력의 개발에 영향을 미치는 중요한 기술 중 하나이다(Lienemann. et al., 2006). 미국의 국립쓰기위원회(National Commission on Writing)에서는 쓰기가 학령기 학생들의 교육적 성공에 결정적 영향을 미치는 중요한 요소이므로, 쓰기 교육이 학교 개혁의 중심이 되어야 한다고 주장하기도 하였다(NCW, 2003) 이와 같이 쓰기 능력은 문해력 발달에 중요한 요소이며, 다른 교과영역의 학습을 위해 사용되는 도구적 성격을 가지고 있고, 사회에서 필요한 핵심적인 기술이므로 학령초기부터 쓰기 능력 향상을 위한 노력이 강조되어야 할 것이다(Puranik & Lonigan, 2012).

이처럼 학업 발달 측면에 있어서 쓰기의 중요성을 고려한다면 일선 학교현장에서는 학령 초기에서부터 효과적인 쓰기 교육과 중재 방법이 강조되어야 할 것이다. 특히, 장애 위험군(at-risk) 또는 학습장애를 가진 학생들을 대상으로 하는 쓰기 교육이 중요시 되어야 한다. 왜냐하면 학습부진이나 학습장애를 가진 많은 학생들이 쓰기에서 어려움을 경험하고 있으나, 이러한 쓰기의 문제점들이 초등 고학년이나 중학생이 될 때까지 발견되지 못하는 경우가 있기 때문이다(McMaster, Parker, & Jung, 2012). 따라서 조기선별과 중재를 통해서 쓰기에서 어려움을 경험하는 학생들의 문제가 장기적으로 지속되지 않도록 미리 예방하고 효과적인 중재를 제공해야 할 것이다(Berninger et al., 2008).

조기선별과 중재를 효과적으로 실시하기 위해서는 쓰기에 어려움을 갖는 장애위험군 학생들을 적절히 선별하여 중재를 실시하고 지속적으로 진전도를 점검하여야 한다. 이를 통해 학생들의 쓰기 성취도에서 변화가 있는지를 살펴보면서 중재의 유형과 강도 등을 변화시켜야 할 것이다. 이러한 선별, 중재, 진전도 점검 체제는 중재반응모형(Response to Intervention(RTI) Model)과 연결시켜 생각해볼 수 있다(Haager, Klingner, & Vaughn, 2007). 중재반응모형은 학습에 어려움을 보이는 학생들(위험군 학생들)을 선별하여 이들에게 부가적 교수를 제공하고 이 후에 이러한 학생들의 진전도를 점검했을 때 적절한 반응을 보이지 않을 경우, 특수교육대상자로서 학습장애를 판별하는 것이다(Fuchs et al., 2003). 이러한 모델은 2004년에 통과된 미국의 전장애아교육법(the Individuals with Disabilities in Education Act: IDEA)에서 새로운 학습장애 정의로서 사용되었으며, 이로 인해 미국에서는 학습장애 판별을 위해 중재반응모형을 대안적으로 사용하게 되었다(Fuchs & Fuchs, 2006).

중재반응모형은 장애위험 학생들을 판별하는 보편적인 선별(universal screening), 증거기반의(evidence-based) 중재 실시, 지속적인 진전도 점검, 진전도를 보이지 않는 학생들을 대상으로 하는 집중적이고 개별화된 교수의 네 가지 주요 단계를 가지고 있다(Berninger et al., 2006). 그런데 이러한 중재반응모형의 각 단계들이 효과적으로 이루어지기 위해서는 학생의 진전도를 타당하게 측정할 수 있는 평가도구가 필요하다. 실제 교실 상황에서 교사가 중재반응모형을 적용하기 위해서는 모든 학생들을 대상으로 간편하게 진전도를 점검할 수 있어야 한다. 이를 위해서는 기존의 전통적인 평가 방식의 학업성취도 검사가 아니라, 실제 수업에서 이루어지는 교육과정을 기반으로 하여 짧은 시간에 간단하게 사용할 수 있는 평가도구가 필요하다.

이를 위해 교육과정중심측정(Curriculum-Based Measurement: 이하 CBM)은 교육현장에서 중재반응모형을 적용할 때 학생들의 진전도를 지속적으로 점검할 수 있는 한 가지 방법으로 활용될 수 있을 것이다(Haager, Klingner, & Vaughn, 2007; Fuchs & Fuchs, 2006). 원래 CBM은 미네소타대학 학습장애연구소(Institute for Research on Learning Disabilities(IRLD))에서 1980년대에 개발한 것으로, 주요 교과목에서 학생들의 학업성취 변화 정도를 지속적으로 측정할 수 있는 표준화된 측정방법으로 정의될 수 있다(Deno, 1985). 이러한 CBM은 학생 수행에 대한 신뢰롭고 타당한 지표를 제공할 수 있으며, 짧은 기간에 걸친 학업 향상 정도를 민감하게 측정할 수 있고 그래프의 기울기를 통해 학생 수행의 향상 정도를 명확하게 보여주는 장점을 가진다(Deno, 1985; 1992; 2003). 이러한 CBM의 장점을 활용하면 교육 현장에서 학생들을 조기에 선별하고 중재를 실시하여 진전도를 점검하는 것을 교사들이 좀 더 편리하게 수행할 수 있을 것이다(여승수, 2012a, 2012b; 여승수, 손지영, 2013). 이처럼 CBM이 갖고 있는 검사의 독특한 특성으로

인하여 중재반응모형에서 학습장애 학생을 선별하기 위한 검사로 활용될 수 있다.

쓰기영역의 CBM은 읽기와 수학교과 영역과 마찬가지로 초기에 미네소타대학 학습장애연구소(IRLD)에서 개발되었으며, 구체적으로 이야기나 그림자극을 제시하면 학생들이 3-5분 동안 쓰기를 실시해서 이들의 쓰기 수행을 측정하는 방법으로 이루어졌다(Deno, Mirkin, & Marston, 1982). 구체적으로, 이 CBM 쓰기 검사에서는 학생이 쓴 총 어절 수, 정확한 맞춤법으로 쓴 어절 수, 맞춤법과 문법에 맞추어 정확하게 쓴 어절이 연속된 횟수로 측정을 하였다. 이러한 쓰기 CBM은 초등 고학년 학생들에게 실행하기 쉽고 효율적으로 쓰기 유창성을 평가할 수 있는 신뢰롭고 타당한 지표를 제공하고 있다(McMaster, Parker, & Jung, 2012). 이후에 일부 선행연구들에서도 쓰기에 어려움을 가지는 학생들의 성취도를 타당하게 측정할 수 있는 CBM을 개발하여 평가를 실시하기 시작했다.

최근에 Coker와 Ritchey(2010)는 학령전기와 초등 1학년 233명의 학생들을 대상으로 쓰기 CBM 검사를 개발하여 적용하였다. 학생들을 대상으로 ‘점심식사 때 먹고 싶은 좋아하는 음식에 대해 쓰시오’ 같은 이야기 자극을 제공한 후, 학생들이 문장을 쓰는 형식으로 이루어졌다. 이 검사에서는 학생이 쓴 어절 수, 정확한 맞춤법으로 쓴 어절 수, 정확하게 쓴 어절이 연속된 횟수를 측정하는 방식으로 이루어졌다. 또한 학생들이 쓴 결과물에 대해 질적으로 평가할 수 있도록 하는 루브릭을 제공하여 점수화하였다. 국내에서는 김동일(2008)이 CBM 기반에 근거한 쓰기 검사를 개발하였는데, 완성된 한 개의 지문을 학생에게 제시하고 1분 동안 이어질 내용을 생각하게 하고 3분 동안 쓰기 검사를 실시하도록 구성되어 있는 이야기 자극 검사의 형태였다.

McMaster와 Espin(2007)은 쓰기 CBM과 관련된 선행연구들을 분석하면서, 쓰기 CBM에서 이러한 이야기 자극 검사의 형태는 초등 저학년 보다는 고학년 학생들을 대상으로 실시되었을 때 더 높은 신뢰도와 타당도가 산출된다고 보고하였다. 이러한 분석 결과는 검사의 난이도 측면과 연결시켜 고려해 볼 수 있을 것이다. 왜냐하면 제시된 한 문장을 바탕으로 한 단락의 작문을 완성하는 과제는 저학년 학생들에게 어려운 과제로 여겨질 수 있기 때문이다(McMaster, Du, & Petursdottir, 2009).

이에 Lembke, Deno, & Hall (2003)의 연구, McMaster & Campbell (2008)의 연구, McMaster와 동료들(2009)의 연구들에서는 초등 저학년 학생들을 대상으로 사용하기에 적합한 여러 가지 쓰기 CBM 검사를 개발하였고 그중 한 가지 검사 유형이 바로 그림-단어 제시검사(Picture-Word Prompt measures)였다. 이 검사는 기존의 CBM의 특성을 모두 갖고 있으며 저학년 학생들에게 적합한 검사임을 기존의 연구결과에서 증명해주고 있다. 즉 신뢰도와 타당도에서 뛰어난 검사임을 증명하였다. 이 검사의 실시방법은 검사지에 학생들이 익숙한 그림과 그림에 해당하는 단어를 함께 제시한 후 학생들이 해당 단어가 포함되는 문장을 완성하는 형태이

다. 따라서 기존의 쓰기 검사의 유형과는 달리 장문의 작문을 요구하지 않으며 짧은 문장을 제한된 시간(2-5분)동안 작성하게 되어 초등학교 저학년 학생들에게 적합한 난이도의 검사를 제공할 수 있는 장점을 갖고 있다.

구체적으로 관련된 연구결과를 살펴보면, McMaster와 동료들(2011)은 12주 동안 85명의 1학년 학생들을 대상으로 그림-단어제시검사, 문장따라쓰기(sentence-copying) 검사, 이야기자극(story prompts) 검사를 매주 실시하여 측정하였는데, 세 가지 검사 유형 중에서 그림-단어제시검사가 짧은 기간 동안 학생들의 진전도에 가장 민감한 것으로 나타났다. 비슷한 맥락으로 Parker와 동료들(2011)도 이러한 그림-단어제시검사의 측정학적 특징으로 민감성을 제시하였는데, 구체적으로 쓰기를 배우기 시작하는 학생들을 대상으로 쓰기 수행의 개별적 차이와 진전도를 민감하게 측정할 수 있는 것이 큰 특징이라고 설명하였다.

이와 같이 미국의 경우에는 쓰기에 효과적으로 중재반응모델을 적용할 수 있도록 하는 CBM 검사의 개발과 적용이 지속적으로 이루어지고 있다. 또한 지금까지 쓰기 CBM 검사의 측정학적 적합성에 대한 연구를 반복적으로 실시하여, 이를 바탕으로 신뢰도가 높고 타당한 검사도구 개발과 현장 적용에 집중하고 있는 상황이다. 그러나 국내에서는 여전히 CBM 검사에 관한 연구는 활발히 이루어지지 못하고 있으며, 특히 쓰기 영역의 CBM 검사 개발은 김동일(2008)의 연구를 제외하고는 매우 미비한 상황이다. 더구나 국내에서는 현재 초등학교 저학년 학생들의 쓰기능력을 평가할 수 있는 CBM 유형의 검사가 충분히 개발되지 못하고 있어, 학교 현장에서 쓰기에 어려움이 있는 학생들을 조기에 선별하고 적합한 중재를 제공하는데 있어서 어려움을 겪을 수밖에 없는 안타까운 현실이다.

이에 본 연구에서는 초등 저학년 학생들을 대상으로 사용할 수 있는 쓰기 CBM 검사로서 그림-단어제시검사를 개발, 적용하여 검사도구의 측정학적 적합성(신뢰도와 타당도)을 평가하고자 한다. 특히 본 연구에서는 2년간의 종단 자료(longitudinal data)에 근거한 측정학적 적합성을 평가하려고 한다. 이처럼 2년간의 종단자료를 설계한 중요한 이유는 실제로 그림-단어제시검사를 이용하여 쓰기장애 위험군 학생을 선별할 수 있는 목적으로 사용될 수 있는지의 여부를 좀 더 명확히 검증하기 위함이었다. 가장 이상적인 조기선별 검사는 초등학교 1,2학년 때 실시된 검사결과를 바탕으로 향후 실시된 국가학업성취도 검사 결과를 정확하게 예측할 수 있어야 한다. 우리나라를 비롯하여 전 세계적으로 국가에서 실시되는 첫 번째 학업성취도 결과는 3학년 초반에 실시된다. 이러한 국가학업성취도 측정시기를 고려한다면, 초등학교 1,2학년에 실시된 검사결과와 3학년에서 실시되는 검사간의 타당도를 고려할 필요가 있을 것이다. 그러나 현재까지 진행된 대부분 예측 타당도의 경우, 검사 실시 기간이 한 달 혹은 한 학기로 매우 짧은 제한점을 갖고 있어서 실제로 장애위험군 학생을 선별할 수 있는지의 여부를 검증하기 어려운 제한점을 갖고 있다. 따라서 본

연구에서는 좀 더 정확한 검사의 예측타당도를 평가하고자 2년간의 중단자료에 근거한 검사실시를 계획하였다.

마지막으로 본 연구에서는 의학분야에서 활용되고 있는 ROC(Receiver Operator Characteristic) 곡선 분석방법을 사용하여 진단의 정확성을 검증하고자 한다. ROC 곡선 분석방법은 검사점수를 이용하여 집단을 분류(예, 장애집단 혹은 비장애 집단, 질병 여부) 할 때의 정확성을 평가할 수 있는 분석방법으로 특히 검사의 예측타당도를 평가할 때 사용된다. 따라서 본 연구에서는 그림-단어제시검사가 실제로 쓰기장애 위험군 집단과 비위험군 집단을 정확히 분류할 수 있는지의 여부를 검증하기 위해서 추가적으로 ROC 곡선 분석방법을 사용하였다.

2. 연구문제

본 연구는 쓰기장애 위험군 학생을 조기에 선별할 수 있는 그림-단어제시검사의 측정학적 적합성을 평가하기 위한 목적을 갖고 있으며, 구체적인 연구문제는 다음과 같다.

- 첫째, 2년간의 중단자료에서 나타난 그림-단어제시검사의 신뢰도는 어떠한가?
- 둘째, 2년간의 중단자료에서 나타난 그림-단어제시검사의 예측타당도는 어떠한가?
- 셋째, 2년간의 중단자료에서 나타난 그림-단어제시검사의 진단적 정확성은 어떠한가?

II. 연구 방법

1. 연구대상

본 연구는 A광역시에 속한 B초등학교 1,2학년 학생을 대상으로 2년간의 중단 연구를 실시하였다. B초등학교의 학생 수는 지속적으로 감소하고 있으며 2012년 1학년의 학급 수는 모두 3학급에 불과하였다. 학급당 평균 학생 수는 25명 내외로 구성되어있다. B초등학교에 소속된 1학년의 3학급 중 담임교사의 자발적인 참여를 나타낸 2학급의 학생들만이 본 연구에 최종적으로 참여하게 되었다. 2012년 3월에 처음 검사를 시작할 당시에는 2학급 학생의 전체 학생 수는 48명이었지만, 전학 등의 이유로 최종적으로 6명이 연구에서 제외되었다. 따라서 본 중단자료의 최종 학생

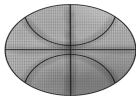
수는 42명이었다. 남녀별 성비를 살펴보면, 남학생이 다소 많은 22명(52.4%)이었으며, 여학생은 이보다 적은 20명(47.6%)이었다.

2. 검사도구

1) 그림-단어제시검사

본 연구에서 사용된 그림-단어제시검사는 McMaster와 Campbell(2008)의 연구에서 개발한 동일한 검사를 한국어에 기반하여 새롭게 개발하였다. 그림-단어제시검사의 구성은 아래의 예시처럼 그림과 단어를 검사지에 동시에 제공하고 있다. 그림과 단어가 제시된 다음 단락에는 학생들이 그림-단어가 포함된 문장을 작성할 수 있는 빈칸이 아래 밑줄로 제공되었다.

* 다음에 주어진 그림의 단어를 포함하는 문장을 완성하십시오.

1) 농구공 

문장: _____

[그림 1] 그림-단어제시검사의 예시

그림-단어제시검사를 개발한 구체적인 과정은 다음과 같다. 먼저 검사에서 사용한 그림과 단어는 저학년 학생의 난이도에 적합한 단어를 선택하기 위해서 초등학교 1,2학년 쓰기 교과서에 포함된 단어를 모두 검토하여 수집한 후 해당 단어를 무작위로 나열한 후 단어를 선정하여 1개의 검사지를 구성하였다. 그림 1과 같이 한 개의 문항에는 한 개의 단어와 그 단어를 나타내는 그림이 함께 제시되어 있다. 한 검사지에는 10개의 그림과 그에 해당되는 단어를 제공하고 있다. 1회차 검사의 예시는 본 연구의 부록에 제공하고 있다.

검사는 매우 간단한 방법으로 실시되었는데, 먼저 2분간 검사가 학급별로 실시된다. 검사실시 전에 검사 앞면에 제공되는 2개의 예문을 먼저 풀어본 후 본 검사에 들어가게 되었다. 학생들은 정해진 검사시간 동안 검사지에 제시된 그림-단어가 포함된 간단한 문장을 작성하였다. 예를 들어, 사과의 그림과 사과가 제시된 문항의 경우 ‘어제 어머니가 사다주신 사과를 먹었다.’와 같은 짧은 문장을 작성할 수 있을 것이다. 본 검사에서 사용한 검사의 채점방법은 기존 연구결과들(Espin et

al., 2000; McMaster & Espin, 2007; Videen, Deno, & Marston, 1982)을 고려하여 ‘맞게 쓰인 어절이 연속된 횡수’ 방법을 사용하였다.

구체적인 채점 방법은 다음과 같았다.

채점방법(맞게 쓰인 어절이 연속된 횡수): 철자, 문법, 내용에 문제가 없는 어절이 연속된 횡수. 철자, 문법, 내용에 문제가 없는 어절이 연속된 횡수에서 그림-단어에서 제시된 단어가 포함된 어절을 작성할 경우 문항 당 0.5점 부여.

2) 예측 타당도 검사도구

본 연구에서는 그림-단어제시검사와 국가학업성취도 쓰기 검사간의 예측 타당도 정도를 검증하기 위해서 2008년에 실시된 ‘초등학교 3학년 국가수준 기초학력 진단평가’ 기출문제를 사용하였다. 본 검사는 한국교육과정평가원 웹사이트(<http://www.kice.re.kr>)의 기출문제 영역에서 문항과 답지를 다운받아서 사용하였다. 진단평가는 총 30문항으로 구성되어 있다. 1-5번까지는 받아쓰기 문항으로 교사가 반복적으로 3회 단어를 불러주고 학생들은 해당 단어를 받아쓰게 되어있다. 6번부터 29번까지의 문항은 주관식과 객관식의 문항들로 구성되었다. 마지막 30번 문항은 주어진 주제에 맞는 짧은 작문을 작성하는 검사로 구성되었다. 검사의 총 만점은 100점이며 문항에 따라 점수의 배점이 차이가 나도록 구성되어 있다. 본 검사에서는 67점을 도달기준 점수로 제공하고 있다. 따라서 본 연구에서도 67점 이하를 받은 학생을 최종적으로 쓰기 위험군 학생으로 분류하였다. 본 연구에서 사용된 “쓰기 위험군”이라는 용어는 쓰기교과영역에서만 위험성을 갖고 있는 학생들이며 이들이 반드시 쓰기장애를 갖고 있는 학생은 아니다. 쓰기 위험군으로 선별된 학생들 중 적합한 증재를 제공함에도 불구하고 반응을 보이지 않을 경우 쓰기장애로 명명할 수 있을 것이다.

3. 검사절차

본 연구에서 실시된 그림-단어제시검사는 2년간 매월(3월-12월까지) 측정되었다. 원칙적으로 매월 3번째 주 수요일에 검사실시를 계획했지만 학교 행사 등의 기타 일정으로 수요일이 어려운 경우도 있었으며, 이러한 경우 3번째 화요일이나 목요일등으로 검사 일정을 변경하여 실시하였다. 검사는 3월부터 학기가 마무리되는 12월까지 실시되었으며 중간에 여름 방학(8월)을 제외한 총 8회의 검사가 매월 실시되었다. 이러한 검사 측정주기는 2년간 실시되었다. 본 연구에서 실시된 검사는 본 논문의 연구자와 특수교육과 박사과정 학생 1명이 함께 검사를 실시하였다. 검사 채점은 특수아동 평가 수업을 이수한 특수교육과 박사과정 학생 1명, 학부 3학년 1

명, 학부 2학년 3명의 학생이 참여하였다. 모든 검사는 2명의 채점자가 각각 1회씩 측정하였으며, 채점자간 검사의 차이점수가 발생할 경우 채점자 및 본 연구의 연구자가 논의를 통해서 합의된 점수를 도출하였다.

예측타당도를 위해서 실시된 진단평가는 검사에 참여한 학생들이 3학년에 올라가기 직전인 2013년 2월 중순에 실시되었다. 검사는 본 연구자와 특수교육과 박사과정학생 1명이 실시하였으며, 본 검사의 채점은 그림-단어제시검사와 마찬가지로 특수아동 평가 수업을 이수한 특수교육과 박사과정 학생 1명, 학부 3학년 1명, 학부 2학년 3명의 학생이 참여하였다. 검사는 2명의 검사자가 각각 실시하였으며 검사점수가 불일치 할 경우 논의를 통해서 합의된 점수를 도출하였다.

4. 분석방법

본 연구에서 사용한 자료분석 방법을 연구문제별로 설명하면 다음과 같다. 첫 번째 연구문제는 그림-단어 제시검사간 동형검사 신뢰도를 분석하는 것이며, 이러한 분석을 위해서 피어슨 상관계수(Pearson product-moment correlation coefficient)를 산출하였다. 두 번째 연구문제는 예측타당도를 밝히는 것으로 첫 번째 연구문제와 마찬가지로 피어슨 상관계수를 사용하였다. 마지막 연구문제는 검사의 진단적 정확성을 분석하는 것이며, 이러한 분석을 위해서 ROC(Receiver Operator Characteristic) 곡선(Curve)을 사용하였다. ROC 곡선 분석방법은 특정한 검사의 절단점 점수(cut scores)를 확인하기 위한 목적으로 사용된다(Swets, 1988). 본 연구에서는 초등학교 3학년 진단평가 결과의 도달 여부를 기준으로 도달된 학생과 실패한 학생을 분류하기 위한 절단점 점수를 확인하기 위해서 사용되었다. ROC 곡선 분석방법에서 절단점 점수의 결정은 민감도(sensitivity)와 특이도(specificity)의 값을 모두 고려하여 결정해야 한다(Swets, 1988). 민감도란 검사도구를 사용하여 실제로 쓰기장애 위험군에 속한 학생을 정확하게 선별할 수 있는 확률을 의미하며, 특이도란 검사도구를 사용하여 위험군에 속하지 않은 학생을 실제로 정확하게 선별할 수 있는 확률로 정의된다. 민감도와 특이도 관계는 반비례 관계이기 때문에 한쪽 값을 증가시키면 다른 쪽 값이 감소하게 된다. 예를 들어, 민감도를 증가시키면 특이도가 감소하게 된다. 따라서 본 연구에서는 민감도와 특이도 값이 모두 높게 산출되는 값을 기준으로 절단점 점수를 산출하였다. 특히 민감도와 특이도중 민감도가 좀 더 높은 값을 기준으로 절단점 점수를 산출하였는데, 그 이유는 쓰기장애 위험군 집단에 포함될 실제 위험군 학생들을 좀 더 정확히 선별하는 것이 조기선별의 목적에 유용하기 때문이다(Riedel, 2007).

ROC 곡선 분석을 통해서 추가적으로 얻을 수 있는 지표는 AUC(Area Under

the Curve)이다. AUC 지표는 절단점 점수에서 산출된 민감도와 특이도 값의 효과 크기(effect sizes)로 해석될 수 있는 지표로 0.5에서 1사이의 값을 산출하게 된다 (Hintze & Silbergliitt, 2005). AUC가 0.5에 가까울수록 선별검사의 정확성이 낮은 것으로 해석되며, 반대로 1에 가까울수록 선별검사의 정확성이 높은 것으로 해석이 가능하다(Swets, 1988).

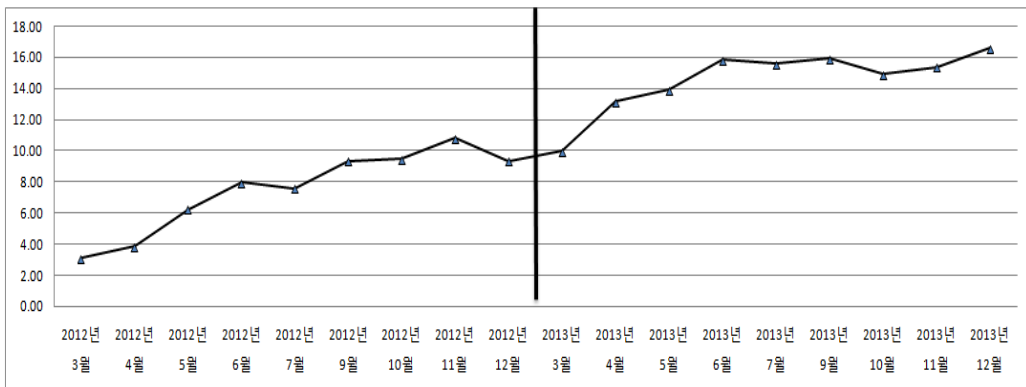
III. 연구 결과

1. 기술 통계치

<표 1>에서는 그림-단어제시검사의 2년간의 월별 기술 통계치를 제공하고 있다.

<표 1> 그림-단어제시검사의 기술 통계치

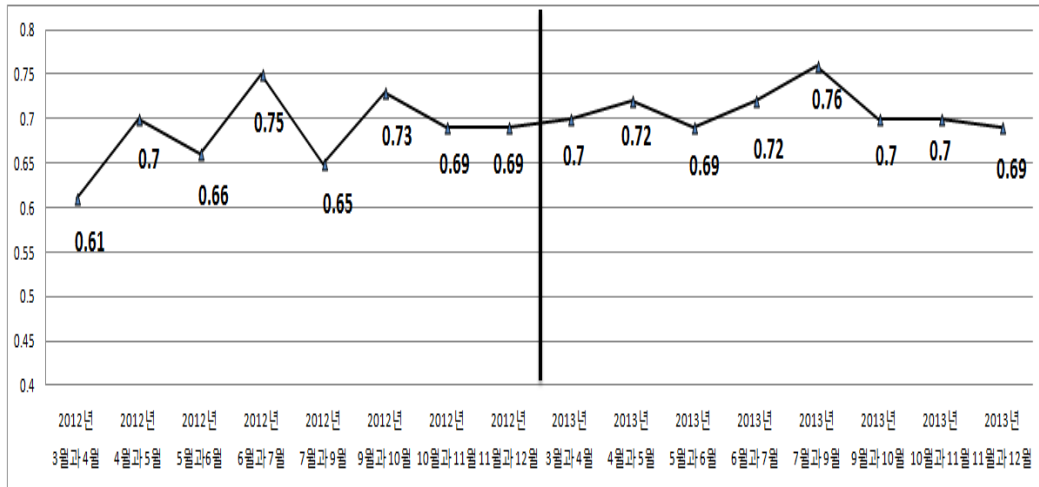
검사	월별 그림-단어제시검사 평균								
	3월	4월	5월	6월	7월	9월	10월	11월	12월
	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차	평균/ 표준편차
1학년	3.14	3.90	6.26	7.98	7.60	9.36	9.45	10.86	9.36
	2.57	3.02	3.44	4.48	3.96	5.84	4.88	5.21	4.43
2학년	9.95	13.19	13.95	15.86	15.62	15.93	14.90	15.38	16.60
	4.21	5.27	5.09	5.05	6.59	5.89	5.15	5.26	5.21



[그림 2] 그림-단어제시검사의 월별 평균 변화

그림-단어 제시검사의 월별 평균의 변화를 살펴보면 처음 초등학교에 갓 입학한 3월에는 대략 3점 정도의 낮은 평균을 얻었지만 2학년 학기말이 되었을 무렵에는 약 17점의 값을 획득하였다. 즉 2년의 기간 동안 학생들은 대략 14점 정도의 쓰기 능력에서 향상을 나타낸 것으로 해석될 수 있을 것이다. 매년 나타나는 평균값의 특징 중 한 가지는 학기 초(3월, 4월, 5월)에는 가파른 평균의 증가가 나타나고 있지만 후반부로 갈수록 평균의 증가량은 크게 나타나지 않는 것으로 나타났다.

2. 신뢰도

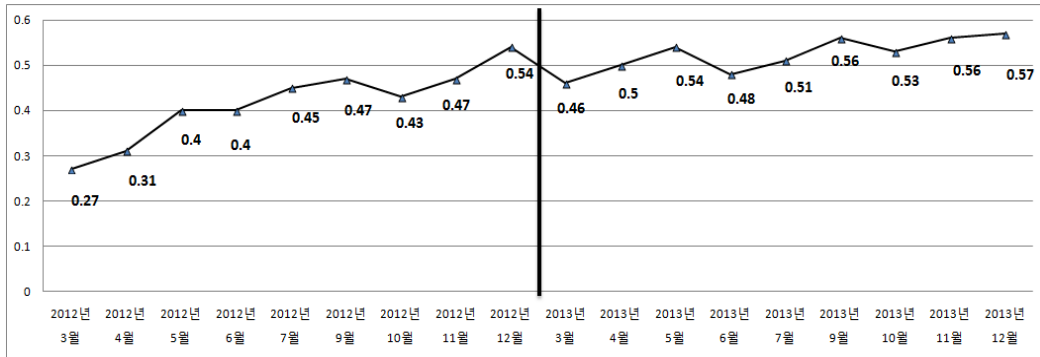


[그림 3] 그림-단어제시검사의 동형검사 신뢰도

인접한 월에서 측정된 두 검사간의 동형검사 신뢰도(alternate-form reliability) 값은 [그림 3]에서 제공하고 있다. 1년차의 1학년 검사결과의 경우 처음 실시된 3월과 4월의 동형검사 신뢰도를 제외하고는 .65-.75의 양호한 신뢰도가 산출되었다. 2년차의 2학년 검사의 신뢰도는 .69-.76의 양호한 신뢰도를 산출하였다. 1학년보다는 2학년의 신뢰도가 조금 더 안정적으로 산출된 것으로 해석된다.

3. 타당도

1) 예측 타당도(준거 타당도)



[그림 4] 그림-단어제시검사의 월별 타당도 지수

국가수준의 학업평가점수인 쓰기 진단평가과 그림-단어제시검사의 예측 및 준거 타당도를 살펴보면, 먼저 1학년의 경우 3월과 4월을 제외하고는 나머지 달에서 측정된 검사에서는 모두 .40이상의 타당도 지수를 산출할 수 있었다. 2학년의 경우 1학년보다는 다소 높은 타당도 지수가 산출되었으며 그 값의 범위는 .46-.56이었다.

2) 진단적 평가의 정확성 분석

ROC(Receiver Operator Characteristic) 곡선) 분석 방법을 사용하여 그림-단어제시검사의 진단적 정확성을 분석하였다. 먼저 민감도와 특이도 값을 모두 고려한 그림-단어제시검사의 월별 절단점 점수는 다음의 <표 2>에서 제공하고 있다. 연구방법에서 설명된 것처럼 민감도와 특이도가 모두 높게 산출되는 절단점 점수를 선별했으며 특별히 민감도가 다소 높은 점수를 최종 절단점 점수로 선정하였다. 검사가 처음 시작된 1학년 3월달의 절단점 점수는 1.5였으며 학년 마지막 달(12월)에 실시된 검사의 절단점 점수는 7.5점이었다. 검사가 진행될수록 점차적으로 절단점 점수도 증가하고 있는 것으로 나타났다.

<표 2> 진단적 평가의 정확성 결과

자료	절단점 점수	민감도	특이도	AUC (Area Under Curve)	분류 정확도 (%)	
2012년 (1학년)	3월 검사	1.5	.72	.50	.68	52
	4월 검사	2.5	.72	.60	.75	69
	5월 검사	4.5	.75	.60	.75	71
	6월 검사	5.5	.75	.60	.78	71
	7월 검사	4.5	.84	.60	.80	78

〈표 2〉

계속

자료	절단점 점수	민감도	특이도	AUC (Area Under Curve)	분류 정확도 (%)	
2012년 (1학년)	9월 검사	5.5	.75	.60	.80	71
	10월 검사	5.5	.84	.60	.81	83
	11월 검사	8.5	.78	.70	.81	76
	12월 검사	7.5	.81	.60	.80	76
2013년 (2학년)	3월 검사	7	.91	.60	.80	80
	4월 검사	10.5	.75	.60	.76	71
	5월 검사	12.5	.72	.70	.77	83
	6월 검사	13.5	.72	.60	.77	69
	7월 검사	13.5	.81	.80	.84	80
	9월 검사	14.5	.75	.70	.82	74
	10월 검사	14.5	.75	.70	.79	74
	11월 검사	13.5	.87	.80	.85	85
12월 검사	14.5	.81	.80	.90	80	

절단점 점수에서 산출된 민감도와 특이도 값을 평가하기 위하여 AUC값을 함께 산출하였다. 1학년의 경우 6월까지 실시된 검사의 점수에서는 .80 미만의 AUC가 산출되었으며 6월 이후에 실시된 검사의 점수에서는 .80이상의 AUC가 산출되었다. 2학년의 경우 4,5,6,10월의 검사를 제외하고는 모두 .80이상의 높은 정확성을 나타냈다.

본 연구에서 산출된 절단점 점수를 사용하여 실제로 쓰기장애 위험군에 포함된 학생을 얼마나 정확히 분류할 수 있는지를 확인하기 위한 확률 결과를 살펴보았다. 먼저, 첫 번째 검사(2012년 3월)에서는 52%로 분류의 정확성이 낮게 산출되었지만 지속적으로 검사가 월별로 실시되면서 분류의 정확성은 70%이상을 나타냈다. 2학년의 분류의 정확성은 1학년보다는 조금 더 높게 나타났으며, 71%–85%사이의 분류 정확성을 나타냈다.

IV. 논 의

그 동안 학습장애 영역에서 쓰기교과는 다른 주요교과 영역과 달리 해당 교과의

중요성에도 불구하고 학교 현장이나 연구 분야에서 낮은 관심을 받고 있었다. 따라서 본 연구에서는 쓰기연구 및 현장에서 활용이 가능한 쓰기검사를 개발하는 주된 목적을 지니고 있다. 특별히 본 연구에서는 초등학교 저학년 학생 중 쓰기장애 위험군에 속할 수 있는 학생을 조기에 선별이 가능한 그림-단어제시검사의 측정학적 적합성을 평가하였다. 그림-단어제시검사가 쓰기 학습장애 학생들에게 유용한 주된 이유는 저학년 학생에게 적합한 난이도의 문항으로 구성되어 있으며, 이러한 특성으로 인하여 기존의 쓰기 검사와는 달리 쓰기장애 위험군 학생을 조기에 선별할 수 있는 목적으로 사용을 할 수 있다. 따라서 본 연구는 실제로 그림-단어제시검사를 사용할 경우 신뢰도와 함께 쓰기장애 위험군 학생을 조기에 선별이 가능한지를 경험적인 자료에 근거하여 검증하였다. 또한 쓰기장애 위험군 학생과 비위험군 학생의 진단적 분류 정확성(diagnostic accuracy)을 검증하기 위해서 ROC 곡선 분석방법을 사용했으며, 학교현장에서 쓰기 영역에서 장애위험군 학생을 조기에 선별하는데 도움이 될 수 있는 절단점 점수를 제공하였다. 구체적인 연구결과에 대한 논의는 다음과 같다.

첫째, 본 연구의 첫 번째 연구문제는 검사의 신뢰도에 대해 2년간의 자료를 이용하여 평가하는 것이었다. 신뢰도 지수는 전반적으로 2학년 보다는 1학년의 지수들이 다소 낮게 나타났다. 이와 같이 1학년의 신뢰도 지수가 낮게 산출된 이유는 검사에 대한 경험이 부족한 이유로 추정할 수 있을 것이다. 신뢰도의 결과는 검사에 대한 일관된 반응으로 검사에 대한 경험이 증가할수록 신뢰도 지수가 증가할 수 있기 때문에 학년이 증가할수록 일반적으로 신뢰도는 증가할 수 있다(Crocker & Algina, 1986). 따라서 검사에 대한 경험이 부족한 초등학교 1학년 학생들의 경우 낮은 신뢰도가 산출되는 것은 충분히 예측이 가능한 결과일 것이다.

본 연구에서는 학년 간 신뢰도 값에 대한 비교와 함께 전반적인 신뢰도 지수가 매우 높지 않게 산출되었다. 즉 .70 초반의 신뢰도 지수가 2년간 산출되었다. 이처럼 그림-단어제시검사의 높지 않은 신뢰도 지수는 이전에 수행된 쓰기 CBM 검사의 신뢰도 결과와는 큰 차이를 나타내지 않는다. 다른 교과영역(예, 수학 CBM, 읽기 CBM)의 검사결과와 비교했을 때 일반적으로 쓰기 CBM 신뢰도는 높지 않은 것으로 나타났다(McMaster & Espin, 2007). 이러한 이유로 Campbell(2010)의 연구에서는 쓰기 CBM 검사의 경우 신뢰도의 값이 .65이상이면 양호한 신뢰도로 평가할 수 있다고 제안하고 있다. 따라서 본 연구에서도 Campbell(2010)의 연구에서 제안한 준거에 근거했을 때 높지는 않지만 양호한 신뢰도 결과로 평가할 수 있을 것이다.

둘째, 본 연구에서는 2년간의 종단자료를 활용하여 예측타당도를 평가하였다. 예측 타당도의 값도 신뢰도와 동일하게 학년이 증가할수록 타당도 지표가 증가하는 것으로 나타났다. 이러한 원인은 타당도 검사와 쓰기 검사간 시간의 차이로 해석이

가능할 것이다. 즉, 두 개의 검사를 실시한 측정기간이 증가할수록 타당도 지수는 감소하게 된다(Crocker & Algina, 1986). 이러한 이유로 인하여 저학년의 경우 높지 않은 타당도가 산출되었으며, 2학년의 경우는 1학년 보다 다소 높은 타당도가 산출되었다. 본 검사에서 산출된 예측 타당도 지수는 1학년의 일부 타당도를 제외하고는 .45이상이며 .57이하로 나타났다. 이러한 타당도 지수의 범위는 동일한 연구를 수행한 McMaster, Du, 그리고 Petursdottir(2009)의 연구에서 산출된 타당도 지수의 값(범위: .52-.60)과도 매우 유사하다. 특히 McMaster, Du, 그리고 Petursdottir(2009)의 연구는 매주 측정된 검사결과를 사용하여 준거타당도를 산출한 점을 고려한다면 2년간의 장기간동안 실시된 검사결과를 이용하여 산출된 본 연구의 타당도 결과는 양호한 것으로 평가할 수 있다.

셋째, 본 연구에서는 ROC 곡선 분석방법을 사용하여 그림-단어제시검사의 분류 정확성을 평가하였다. 먼저 ROC 곡선분석의 민감도와 특이도 값을 고려하여 절단점 점수를 사용하였다. 절단점 점수를 사용할 경우 실제 분류의 정확성이 초기에는 낮은 정확성을 나타내지만 측정횟수가 증가할수록 70-80% 정도로 양호하게 나타났다. 앞서 언급된 것처럼 타당도 지수는 .60정도로 높지 않은 수치이지만 ROC 곡선 분석방법에서 산출된 절단점 점수를 사용할 경우 진단의 정확성이 좀 더 높게 나타나는 것으로 확인되었다. 따라서 그림-단어제시검사를 실제 학교현장에 적용하여 쓰기장애 위험군 학생을 정확하게 선별하기 위해서는 ROC 곡선분석에서 산출된 절단점 점수를 활용하는 것이 진단의 정확성을 향상시키는데 있어서 유용하게 사용될 수 있음을 추정할 수 있다.

본 연구를 통해서 제공할 수 있는 쓰기 학습장애영역의 연구 및 교육적 함의는 다음과 같다.

첫째, 학습장애 영역 중 특히 쓰기교과는 다른 주요 교과(예, 수학, 읽기)와 비교했을 때 그 동안 학교 현장 및 관련된 연구에서도 관심을 받지 못했던 교과 영역이었다. 이러한 이유로 인하여 쓰기 학습장애의 출현율, 중재전략, 진단에 관한 연구들도 활발히 진행되지 않았던 제한점을 갖고 있었다. 이와 같이 소외받은 이유 중 한 가지는 쓰기학습장애 학생들을 연구할 수 있는 적합한 검사 도구가 부재했기 때문일 것이다. 따라서 본 연구에서 개발한 그림-단어제시검사는 특히 초등학교 저학년 학생들의 쓰기능력과 관련된 연구에 활용될 수 있을 것으로 기대된다.

둘째, 본 연구의 주된 연구목적과 마찬가지로 본 연구에서 사용한 그림-단어제시검사를 사용할 경우 쓰기장애 위험군 학생을 조기에 선별이 가능할 것으로 기대된다. 본 연구에서 산출된 신뢰도와 타당도 지수는 양호한 것으로 평가할 수 있으며, 이와 함께 실제 집단을 분류하는 정확성도 양호한 것으로 나타났다. 현재 초등학교 저학년 학생들에게 적합한 쓰기검사도구가 충분치 개발되지 않고 있는 점을 고려한다면, 본 연구에서 개발한 그림-단어제시검사는 학교현장에서 조기선별의 목적으로

유용하게 사용될 수 있을 것으로 기대된다.

셋째, 특히 본 연구에서는 ROC 곡선 분석을 사용하여 조기선별에 도움을 줄 수 있는 절단점 점수를 제공하고 있다. 따라서 일선 초등학교 현장에서 검사를 실시한 후 본 연구에서 제공하는 절단점 점수를 바탕으로 쓰기장애 위험군 학생을 조기에 선별할 수 있을 것이며, 이러한 조기선별은 조기중재를 제공할 수 있는 교육적 근거로 활용될 수 있을 것이다.

넷째, 본 연구에서는 2년이라는 중단자료를 이용하여 예측타당도와 신뢰도 결과를 검증하였다. 지금까지 진행된 대부분의 CBM 쓰기 연구의 경우 한 학기 혹은 한 달이라는 짧은 기간 동안 검사의 측정학적 적합성을 평가하고 있기 때문에 실제로 쓰기 검사를 사용하여 높은 예측 타당도를 기대할 수 있는지의 여부를 확신하기는 어려웠다. 따라서 2년이라는 중단연구 설계를 통해서 검증된 본 연구의 결과는 향후 1-2년 후에 쓰기장애 위험군 집단으로 선별될 수 있는 학생을 선별하기 위한 목적으로 사용이 가능함을 보여주고 있다.

다섯째, 본 연구에서 사용한 그림-단어제시검사는 CBM에 기반한 쓰기 검사 도구이다. 즉, CBM이 갖고 있는 중요한 특성을 모두 공유하고 있다. 예를 들어, 검사 실시시간이 매우 짧은 것과 함께 검사실시방법이 매우 간편하며, 특히 본 검사는 집단검사로 사용이 가능하다. 따라서 본 검사를 실제 초등학교 현장에서 적용할 경우 검사를 위해서 특별히 교육과정을 수정하거나 수업진도의 결손을 초래하지 않으며 아침 자습시간을 통해서 간단히 검사를 실시할 수 있는 장점을 갖고 있을 것이다. 즉 교사가 감당해야 할 추가적 평가의 업무를 최소화할 수 있는 장점을 지니고 있다. 따라서 학습부진을 사전에 예방하고 감소시키고자하는 학교 혹은 지역 교육청의 경우 그림-단어제시검사를 학교 및 학급단위로 실시하여 전체 저학년 학생들의 쓰기 능력을 오랜 기간 동안 지속적으로 모니터링 할 수 있는 장점과 함께 장애위험군 학생을 조기에 선별할 수 있는 경험적 근거를 제공할 수 있다.

본 연구에서 사용한 그림-단어제시검사의 연구를 지속적으로 발전시키기 위해서는 다음과 같은 후속연구들이 지속될 필요가 있다.

첫째, 본 연구는 2개의 학급에 소속된 초등학교 저학년 학생들이 참여하여 검사의 측정학적 적합성을 평가했지만, 향후 진행될 후속연구에서는 다양한 지역에서 더 많은 표본의 학생이 참여할 필요가 있을 것이다. 또한 학생 개인의 환경변인은 학업 성취에 중요한 변인이기 때문에 향후 연구에서는 이러한 점을 고려할 필요가 있을 것이다.

둘째, 본 연구에서는 국가에서 실시한 초등학교 3학년 진단평가를 사용하여 예측타당도와 절단점 점수를 사용했지만 검사의 절단점은 임의로 결정한 점수이기 때문에 신중히 고려할 필요가 있으며, 이러한 이유로 인하여 향후 진행될 연구에서는 다양한 쓰기 준거 타당도 검사가 사용되어 절단점을 점검할 필요가 있다. 이러한 연

구를 통해서 좀 더 정확한 절단점 점수를 확인하고 검증할 수 있을 것이다.

셋째, 본 연구는 검사의 시작을 아동이 제도권 교육에 들어오는 초등학교 1학년으로 한정되었지만 최근에 진행된 외국의 관련 연구의 경우 유치원 학생들을 대상으로 CBM 쓰기 검사의 예측타당도를 검증하고 있다. 따라서 조기선별을 위하여 유치원 학생들을 시작점으로 실시되는 추가적인 예측타당도도 실시될 필요가 있을 것이다.

참고문헌

- 김동일(2008). **기초학습기능 수행평가체계: 쓰기 검사**. 서울: 학지사.
- 여승수 (2012a). 중재반응모형 활용을 위한 수학 CBM 검사의 신뢰도 및 예측 타당도 검사. **학습장애연구**, 9(1), 89-105.
- 여승수 (2012b). 측정시간 차이에 따른 수학 CBM 검사 결과의 비교: 학습장애 선별 모형인 중재반응모형에서의 활용을 위하여. **학습장애연구**, 9(2), 53-69.
- 여승수, 손지영 (2013). 수학 학습장애 진단을 위한 교육과정중심측정(CBM) 기울기의 효율적인 측정주기 분석. **학습장애연구**, 10(1), 69-86
- Berninger, V. W., Rutberg, J. E., Abbott, R. D., Garcia, N., Anderson-Youngstrom, M., Brooks, A., et al (2006). Tier 1 and tier 2 early intervention for handwriting and composing. *Journal of School Psychology*, 44, 3-30.
- Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, 46, 1-21.
- Campbell, H.(2010). The technical adequacy of Curriculum-based measurement passage copying with secondary school English language learners. *Reading & Writing Quarterly*, 26, 289-307.
- Coker, D. R., & Ritchey, K. D.(2010). Curriculum-based measurement of writing in kindergarten and first grade: An investigation of production and qualitative scores. *Exceptional Children*, 76, 175-193.
- Crocker, L., & Algina, J.(1986). *Introduction to classical and modern test theory*. New York, NY: Holt Rinehart & Winston.
- Deno, S. L.(1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219-232.
- Deno, S.(1992). The nature and development of curriculum-based measurement. *Preventing School Failure*, 36(2),5-10.
- Deno, S. L.(2003). Developments in curriculum-based measurement. *Journal of Special Education*, 37(3), 184-192.
- Deno, S. L., Mirkin, P., & Marston, D.(1982). Valid measurement procedures for

- continuous evaluation of written expression. *Exceptional Children Special Education and Pediatrics: A New Relationship*, 48, 368-371.
- Espin, C., Shin, J., Deno, S. L., Skare, S., Robinson, S., & Benner, B.(2000). Identifying indicators of written expression proficiency for middle school students. *Journal of Special Education*, 34, 140-153.
- Fuchs, D., & Fuchs, L. S.(2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41(1), 92.
- Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L.(2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities Construct. *Learning Disabilities Research & Practice*, 18(3), 157-171.
- Graham, S., & Perin, D.(2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99, 445-476.
- Haager, D., Klingner, J., & Vaughn, S.(2007). *Evidence-based reading practices for response to intervention*. Baltimore: Paul Brooks Publishing Co.
- Hintze, J. M., & Silberglitt, B.(2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review*, 34(3), 372-386.
- Lembke, E., Deno, S. L., & Hall, K. (2003). Identifying an indicator of growth in early writing proficiency for elementary school students. *Assessment for Effective Intervention*, 28(3-4), 23-35
- Lienemann, T. O., Graham, S., Leader-Janssen, B., & Reid, R.(2006). Improving the writing performance of struggling writers in second grade. *Journal of Special Education*, 40(2), 66-78.
- McMaster, K. L. & Campbell, H.(2008). New and existing curriculum-based writing measures: Technical features within and across grades. *School Psychology Review*, 37, 550-556.
- McMaster, K. L., Du, X., & Petursdottir, A.(2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities*, 42, 41-60.
- McMaster, K. L., Du, X., Yeo, S., Deno, S. L., Parker, D., & Ellis, T.(2011). Curriculum-based measures of beginning writing: Technical features of the slope. *Exceptional Children*, 77, 185-206.
- McMaster, K. L., & Espin, C.(2007). Curriculum-based measurement in writing: A literature review. *Journal of Special Education*, 41, 68-84.
- McMaster, K. L., Parker, D., & Jung, P.(2012). Using curriculum-based measurement for beginning writers within a response to intervention framework. *Reading Psychology*, 33, 190-216.
- National Commission on Writing(2003). The neglected "R" : The need for a writing

- revolution. Retrieved from <http://www.writingcommission.org/>
- Parker, D. C., McMaster, K. L., Medhanie, A., & Silbergitt, B.(2011). Modeling early writing growth with curriculum-based measures. *School Psychology Quarterly*, 26(4), 290-304.
- Persky, H. R., Daane, M. C., & Jin, Y.(2003). The nation's report card: Writing 2002. Retrieved from <http://nces.ed.gov/nationsreportcard/writing/>
- Puranik, C. S., & Lonigan, C. J.(2012). Early Writing Deficits in Preschoolers With Oral Language Difficulties. *Journal of Learning Disabilities*, 45(2),179-190.
- Riedel, B. W.(2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly*, 42(4), 546-567.
- Salahu-Din, D., Persky, H., & Miller, J.(2008). The nation's report card: Writing 2007. Washington, DC: U.S. department of education national center for education statistics. Retrieved from <http://nces.ed.gov/nationsreportcard/writing/>
- Shanahan, T.(2004). Overcoming the dominance of communication: Writing to think and to learn. In T. L. Jetton & D. A. Dole (Eds.), *Adolescent research and practice* (pp. 59-73). New York, NY: Guilford.
- Swets, J. A.(1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285-1293
- Videen, J., Deno, S. L., & Marston, D.(1982). *Correct word sequences: A valid indicator of proficiency in written expression* (Report No. IRLD-RR-84). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Examining Predictive Validity and Reliability of
Picture-Word Prompt Measure
for Identifying Struggling Beginning Writers
: 2 Year-Longitudinal Data

Yeo, Seungsoo

Busan National University of Education

Son, Ji Young

Daejeon University

<Abstract>

The main purpose of this study was to develop and evaluate writing measures for early screening for first-second graders at risk. Especially, in this study, picture-word prompt measure was used as a measure for identifying struggling beginning writers. In this study, 42 students from B elementary school in the A city were participated. In order to evaluate alternate-form reliability and predictive validity, pearson correlation-coefficients were used, while ROC curve analyses were used for evaluating accuracy of classification. This study showed that the picture-word prompt measure could be considered as an adequate indicator of predicting struggling beginning writers. In addition, the use of ROC curve analyses resulted in increasing the probability of accuracy of classification. Implication for research and practices are presented.

Key Words : writing, picture-word prompt measure, CBM, reliability, predictive validity, accuracy of classification

논문 접수: 2014. 11. 05 심사 시작: 2014. 11. 10 게재 확정: 2014. 12. 15