

## 아동 인지 정서 모니터링 도구 CEMO-A의 개발 및 타당화\*


<b>박 주 리</b> 서울여자대학교 교육심리학과 박사과정	<b>임 명 지</b> 서울여자대학교 교육심리학과 박사과정	<b>김 자 은</b> 서울여자대학교 교육심리학과 석사과정
<b>손 인 영</b> 서울여자대학교 교육심리학과 석사과정	<b>장 수 민</b> 서울여자대학교 교육심리학과 석사과정	<b>송 현 주<sup>†</sup></b> 서울여자대학교 교육심리학과 교수

본 연구는 아동의 정신 건강 평가와 조기 선별을 위한 디지털 모니터링 도구인 CEMO-A의 개발 및 타당성 검증을 목적으로 한다. CEMO-A는 고유의 생태학적 평가 방식과 기능성 게임화를 적용하여 일상적 환경에서도 자연스럽게 타당한 평가가 가능하다. 또한, ‘지속적 주의, 작업 기억, 인지 통제, 실행 기능, 정서 민감성, 정서 인식’ 등 여러 인지 및 정서 요소를 통합적으로 평가하여 주요 정신 장애를 선별할 수 있도록 설계되었다. 연구는 6-10세 일반 아동 100명을 대상으로 하였으며, CEMO-A와 기존의 표준화된 심리 검사(아동 대상 지능 검사 및 신경심리 검사, 보호자 대상 질문지) 결과를 통해 구성 타당도와 공준 타당도를 검증하였다. 분석 결과, 인지와 정서 기능을 독립적이면서도 통합적으로 평가하는 CEMO-A의 구성적 타당성이 입증되었으며, 기존의 표준화된 검사와의 비교를 통해 공준 타당성도 확인하였다. 본 연구는 CEMO-A가 아동의 정신 건강 상태를 효율적으로 평가할 수 있는 유용한 디지털 도구임을 시사하며, 향후 조기 선별 및 예방적 개입을 돕는 효과적인 방법이 될 수 있음을 제안한다.

주요어 : CEMO-A, 디지털 평가, 디지털 정신 건강 도구, 정신 건강 선별, 인지 및 정서 평가

\* 본 연구는 보건복지부 재원으로 한국보건산업진흥원의 보건의료기술연구개발사업 지원을 받아 이루어졌음(과제명: 디지털 표현형 기반 아동청소년의 심리상태 모니터링 및 중재 플랫폼 구축, 고유번호: HI22C0775).

† 교신저자(Corresponding Author) : 송현주 / 서울여자대학교 교육심리학과(현 심리·인지과학학부) 교수 / (10797) 서울시 노원구 화랑로 621 50주년 기념관 613호 / Tel: 02-970-5888 / E-mail: jayoo1004@swu.ac.kr

 Copyright ©2025, Clinical Psychology in Korea: Research and Practice  
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

최근 몇 년간 디지털 정신 건강(digital mental health) 기술의 발전은 정신 건강 및 심리 치료 분야에 큰 변화를 가져왔다. 이러한 기술은 크게 디지털 평가 도구와 디지털 치료제로 구분될 수 있으며, 각기 정신 건강 문제를 평가하고 치료적 개입을 하는 데 있어 높은 접근성과 효율성을 지닌 혁신적인 방법으로 주목받고 있다(Marcu et al., 2022; Spadaro et al., 2023). 특히, 디지털 정신 건강 평가 도구는 정신 장애를 조기에 탐지하고 전문적인 개입의 필요성을 파악하기 위한 효과적인 방법으로, 제한된 환경에서 시행하는 기존의 표준화된 평가 방식에 비해 활용 가능성과 생태학적 타당성이 높은 것으로 보고되었다(Park et al., 2024; Song et al., 2020). 아동의 경우 정신 건강 문제의 장기적이고 광범위한 영향을 예방하고 적절한 개입을 제공하기 위한 조기 선별이 더욱 중요하므로, 이러한 디지털 평가 도구 개발의 필요성이 크다. 하지만 문헌 검토 연구(Martin-Key et al., 2022)에 의하면, 이러한 디지털 평가 분야는 아직 초기 단계로, 단순히 설문지를 디지털화한 경우가 다수이며, 정교하고 포괄적인 진단 기술을 사용한 체계적인 연구는 매우 부족한 실정이다.

디지털 평가에서는 아동의 집중도와 참여도를 높이기 위해 기능성 게임화(serious gaming)를 적용할 수 있다. 기능성 게임은 사용자가 목표에 도달하기 위해 여러 단계, 활동, 평가를 거치는 방식으로 설계된 가상 플랫폼으로, 교육이나 의료 같은 실제 세계의 중요한 목적을 위해 게임 원칙과 개념을 적용하는 방식을 의미한다(EI Kayed et al., 2024). 아동은 이러한 게임을 통해 자연스럽게 인지 및 정서적 능력을 평가받을 수 있다. 이는 전통적인 검사 방식에 비해 아동의 부담을 줄이고, 더 현실적

인 데이터를 얻을 수 있는 장점이 있다. 선행 연구에 의하면, 개인의 인지 기능을 평가하기 위해 게임화된 평가 방법을 활용할 경우, 개인의 동기 부여 향상에 영향을 줄 뿐만 아니라, 시험 불안을 감소시켜 최선의 성과를 얻을 수 있다(Lumsden et al., 2016). 또한 게임화를 활용할 경우 참여자의 동기 부여 및 참여에 긍정적인 영향을 주면서도 검사의 정확도를 유지할 수 있음이 시사되었다(Vermeir et al., 2020). Jablonská 외(2024)는 4-8세 아동의 인지 기능을 평가하는 상호작용 게임의 사용성 평가에서, 해당 게임들이 목표한 인지 기능을 타당하게 측정하며 아동들에게 긍정적인 반응을 얻었다고 보고하였다.

디지털 정신 건강에 대한 임상적 고찰 논문에서는, 특히 진단적 경계가 명확하지 않거나 발달 과정의 영향을 고려해야 하는 집단에서 초-진단적 접근이 중요함이 강조되었다(Smith et al., 2023). 또한, 정서와 인지를 대립되는 것으로 간주하고 별개로 평가하던 과거와 달리, 이 두 과정은 복합적으로 상호작용하며 관련 뇌 영역들이 강하게 연결되어 있기 때문에 통합적으로 평가할 필요가 있다는 점이 반복적으로 제기되었다(Barbas & García-Cabezas, 2017). 예를 들어, 최근 신경과학 연구에서는 후측 안와전두피질과 전측 대상피질이 전전두피질과 강하게 연결되어 있어, 이를 통해 정서와 인지 과정이 상호 연계되어 작용함을 나타냈다(Barbas & García-Cabezas, 2017). 자율신경 정신생리학 연구도 인지와 정서가 신체적인 각성 상태와 통합적이고 양방향적으로 연결되어 있음을 강조했으며(Quadt et al., 2022), 정동적인 스트룹 과제와 기능적 자기공명영상(fMRD)을 활용한 연구에서는 정서와 인지가 공통된 신경 네트워크를 공유한다는 점이 시사

되었다(Raschle et al., 2017). 부정적 정서는 그 정서의 강도와 상황에 따라 실행 기능을 저하시킬 수도 있고, 반대로 도움이 되기도 하는 등, 정서와 인지 간의 상호작용은 이중적인 효과를 나타낸다(Crivelli et al., 2024; Rodas et al., 2024). 이러한 연구 결과는 특정 상황에서 정서적 요인이 인지적 통제를 방해하거나 촉진하는 요인으로 작용할 수 있음을 시사한다. 아동의 뇌는 아직 발달 과정에 있어 성인보다 인지와 정서 정보를 더욱 밀접하게 연결하여 처리하는 경향이 강하므로, 이 두 가지 기능을 통합적으로 평가하는 것이 더욱 중요하다(Silvers et al., 2014; Spear, 2013).

인지와 정서 기능 평가 시, 아동기에는 연령에 따른 발달적 차이도 고려할 필요가 있다. 전두엽 실행 기능을 비롯한 인지 기능의 발달은 수십 년 동안 광범위하게 연구되어 왔으며, 아동기에 유의미한 발달이 이루어진다는 점이 반복적으로 확인되어 왔다(Klenberg, 2015). 예를 들어, 6-8세 초기 학령기 아동을 대상으로 한 연구에서 실행 기능의 유의미한 연령 차이가 확인되었다(Canet Juric et al., 2013). 7-12세 아동을 대상으로 한 연구에서도, 주의력, 단기 기억, 작업 기억 과제 모두에서 연령 간 유의미한 차이가 나타났다(Egami et al., 2015). 이처럼 인지 기능이 발달하는 과정에서, 정서 기능 또한 연령에 따라 변화하는 양상을 보인다. 그러나, 정서 인식 능력의 발달에 대한 연구는 세부적인 연구 방법에 따라 연령 간의 차이에 있어 상이한 결과가 나타났으며, 최근 연구(Riddell et al., 2024)는 아동의 정서 구별 능력에 대해 체계적으로 정리한 연구가 소수에 불과하다는 점을 지적하였다. 기존 연구를 종합적으로 검토한 Herba와 Phillips(2004)는 유아기에도 우수한 정서 인식 능력이 나타나며,

이 시기에는 역동적인 얼굴 표정과 음성 정보가 중요한 역할을 한다고 보고하였다. 또한, 유아기 이후에도 정서 표현 인식 능력이 향상되지만, 특정 발달 단계에서 급진적인 변화가 나타나기보다는 점진적으로 발달하는 경향을 강조하였다. 유사한 맥락에서, Widen(2013)은 아동이 얼굴 표정을 통해 정서를 명명하는 능력이 비교적 초기에 발달하여, 평균적으로 74개월(6세 2개월)경에는 기본적인 여섯 가지 정서(행복, 슬픔, 분노, 두려움, 놀람, 역겨움)를 구별하여 명명할 수 있다고 보고하였다. 특히 행복, 슬픔, 분노, 두려움은 보다 이른 시기에 발달하는 것으로 나타났다. 한편, Chronaki 외(2015)의 연구에서는 얼굴 및 음성 자극을 활용한 정서 변별 과제에서 연령 간 유의미한 차이가 확인되었다. 그러나 이러한 차이는 6세 미만의 미취학 아동과 20세 이상의 성인을 포함한 결과에서 확인된 것이며, 6세에서 9세의 저연령 아동과 10세에서 11세의 고연령 아동 간에는 정서 인식 정확도의 차이가 나타나지 않았다. 이와 유사하게, West 외(2020)의 연구에서도 음성-얼굴 정서 일치 과제에서 7세에서 9세 사이의 연령 간 유의미한 상관성이 확인되지 않았다.

본 연구는 아동 및 청소년의 정신 장애 위험성을 조기에 탐지할 수 있는 평가용 앱인 CEMO-A(Cognitive Emotion Monitoring-Assessment)를 개발하고, 타당성을 검증하기 위해 진행하였다. CEMO-A는 아동의 자연스러운 참여를 유도하는 게임 요소 및 고유한 생태학적 평가 방식을 도입하여, 인지 및 정서 기능을 일상적 맥락에서 종합적으로 평가할 수 있도록 설계되었다. 타당성 검증을 위해, 6세에서 10세 일반 아동을 대상으로 CEMO-A와 기존의 표준화된 심리 검사를 실시하고 비교·분석하였다.

이를 통해 CEMO-A가 아동의 인지 및 정서 상태를 정확히 평가할 수 있는지 확인하고, 기존 심리 평가 도구와 비교하여 변별력과 민감도를 검토하고자 한다. 본 연구를 통해 CEMO-A의 타당성이 입증됨으로써, 정신 건강 문제의 조기 탐지와 평가에 기여할 수 있게 될 것으로 기대한다.

## 방 법

### 연구대상

본 연구는 아동용 인지·정서 모니터링 앱인 CEMO-A의 타당성을 검증하기 위해, 6세에서 10세 사이의 일반 아동을 대상으로 진행되었다. 연구에 필요한 연구 참여자 수는 G\*Power 프로그램을 통해 계산하였으며, 상관 분석을 위한 표본 크기 산출을 기준으로, 효과 크기  $r = 0.30$ , 유의수준  $\alpha = 0.05$ , 검정력 0.80의 조건에서 최소 84명의 참여자가 필요함을 확인하였다(Faul et al., 2009). 이에 따라 중도 탈락을 고려하여 총 100명을 모집하였으며, 각 연령군(6, 7, 8, 9, 10세)의 남녀 각각 10명씩 균등하게 구성하고자 하였다.

그러나 결과 분석 과정에서 연령별 비교를 위한 ANOVA 분석을 추가하게 되었으며, G\*Power를 통해 ANOVA 분석의 적절한 샘플 크기를 산출한 결과, 효과 크기  $f = 0.25$ (중간 효과 크기), 유의수준  $\alpha = 0.05$ , 검정력 0.80의 조건에서 최소 200명의 참여자가 필요함이 확인되었다. 따라서 본 연구의 표본 크기는 상관 분석을 수행하는 데에는 적절하나, ANOVA 분석의 경우 검정력이 제한될 가능성이 있다.

또한, 7세 연령군 모집 과정에서 담당 연구

원이 성별 분포를 정확하게 확인하지 못하여 2명이 남아 대신 여아로 모집되는 오류가 발생하였다. 해당 오류는 연구 검사가 완료된 후 모집 데이터를 검토하는 과정에서 확인되어, 7세 연령군은 성비 균형이 맞지 않게 되었다.

연구 대상자는 지역 커뮤니티 및 SNS에 모집 공고를 게재하여, 참여 희망자 100명을 선착순으로 모집하였다. 연구 대상이 미성년자임을 고려하여, 법정대리인과 아동 본인의 동의를 모두 받은 경우에 한해 연구에 참여할 수 있도록 하였다. 연구 대상자 제외 기준은 경계선 지적 기능 또는 지적 장애(지능지수 80 미만)에 해당하거나, 시력 또는 청력에 제한이 있는 경우로 설정하였다. 중도 탈락 기준은 연구 참여 중 아동 또는 보호자가 참여 의사를 철회한 경우로 하였으나, 실제로 중도 탈락자는 없었다.

각 연령군에서 20명씩 총 100명이 연구에 참여했으며, 이 중 9세 여아 1명은 경계선 지적 기능에 해당하는 전체지능지수(Full Scale Intelligence Quotient[FSIQ] = 74)가 확인되어 분석 대상에서 제외하였다. 또한 10세 남아와 9세 여아 각각 1명은 CEMO-A 앱의 일시적인 오류로 인해 데이터가 저장되지 않아 분석에서 제외되었다.

한편, 보호자 보고 및 SCQ(Social Communication Questionnaire) 결과를 고려했을 때, 연구 대상자 중 자폐 스펙트럼 장애(Autism Spectrum Disorder [ASD])로 의심되는 아동은 포함되지 않았다. 그러나 CBCL(Child Behavior Checklist)의 DSM(Diagnostic and Statistical Manual of Mental Disorders) 진단 척도(70T 이상)를 기준으로 ADHD(Attention-Deficit/Hyperactivity Disorder)가 의심되

는 아동은 4명, 우울 장애가 의심되는 아동은 2명, 불안 장애가 의심되는 아동은 5명이 포함되었다. 이 중 한 아동은 ADHD와 불안 장애 모두에서 CBCL 기준 임상군에 해당하였다.

## CEMO-A의 개발

### 개발 과정

본 연구팀에서 개발한 CEMO-A는 태블릿 기반의 기능성 게임형 인지 및 정서 평가 앱이다. 이는 디지털 치료제를 수년 간 연구해 온 심리학 교수이자 임상심리전문가인 교신저자를 주축으로, 상담 및 임상심리 전공 석사 및 박사 과정 연구원, 그리고 디지털미디어학과 교수와 외부 개발 업체 등과의 협업을 통해 총 3단계에 걸쳐 개발되었다.

**1단계: 개념적 프레임워크 및 평가 항목 선정.** CEMO-A의 개발은 타당화된 기존의 디지털 평가 도구인 CoCon(Song et al., 2020)의 구조를 기반으로, 아동의 인지 및 정서 기능을 종합적으로 평가할 수 있는 개념적 프레임워크를 구축하는 과정으로 시작되었다. 이러한 평가 항목과 주요 측정치는 교신저자의 다년간 디지털 정신건강 기술에 대한 연구 및 임상 경험을 바탕으로, 문헌 고찰 및 선행 연구 검토를 통해 설계되었다. 또한, 교신저자와 소아정신건강의학과 교수 1인이 아동의 우울, 불안, 스트레스 반응 등에 대한 평가의 임상적 적절성을 검토하며 내용 타당도를 검증하였다.

CEMO-A는 본 시행 전에 아동이 각 게임을 충분히 이해할 수 있도록 연습 시행을 먼저 진행하며, 각 게임과 소게임의 정답 수와 오답 수는 자동으로 기록되도록 개발되었다. 정

답률에서 오답률을 차감하여 최종 점수(final score [FS])를 산출하며, 이 점수는 ‘지속적 주의(sustained attention [SA]), 작업 기억(working memory [WM]), 인지 통제(cognitive control [CC]), 실행 기능(executive function [EF]), 정서 민감성(emotion sensitivity [ES]), 긍정정서 인식(positive emotion recognition [PER]), 슬픔 인식(sad recognition [SdR]), 분노 인식(anger recognition [AgR]), 불안 인식(anxiety recognition [AxR])’ 등의 핵심 측정치를 도출하는 데 활용된다. 또한, 핵심 측정치를 바탕으로 ‘주의력-결핍/과잉행동 장애(ADHD), 자폐 스펙트럼 장애(ASD), 우울 장애(depressive disorder [DEP]), 불안 장애(anxiety disorder [ANX]), 스트레스(stress [ST])’와 같은 선별 측정치가 산출되도록 구성되었다. 선별 측정치의 경우, 점수가 낮을수록 해당 문제가 있을 가능성이 높음을 나타낸다.

**2단계: 프로토타입 개발 및 자극 검토.** 소프트웨어 개발 전문가 및 외부 개발업체와 협력하여 CEMO-A의 초기 프로토타입을 구현하였다. 사용자 경험(user experience [UX])과 직관적인 인터페이스(user interface [UI])를 고려하여 설계했으며, 아동이 쉽게 흥미를 느끼고 적극적으로 참여할 수 있도록 하기 위해, 선택한 동물 캐릭터가 숲속 대원이 되어 다양한 뱃지를 획득해 가는 이야기 형식을 도입하였다.

아동의 정서 및 인지 기능을 평가하기 위한 얼굴 표정 및 목소리 음성 자극은 석사 및 박사 과정 연구원들이 전문 디자이너나 성우와의 협업을 통해 개발하였다. 얼굴 자극은 아동, 청소년, 성인의 세 가지 연령대와 남녀 성별을 반영하여 총 12인의 얼굴 표정 자극을 제작하였으며, ‘행복, 놀람, 부끄러움, 불안, 슬

품, 화남, 무표정(중립)’의 일곱 종류의 정서를 포함하였다. 눈과 눈썹의 각도, 입 꼬리의 각도와 형태 등을 수치로 조정하여 정서 강도를 두 단계로 구분하였다. 음성 자극은 전문 성우들이 전문 녹음실에서 녹음하였으며, 녹음 과정에서 연구원들의 지침과 피드백을 즉석에서 반영하여 진행하였다. 자극 문장은 중립적인 내용의 짧은 문장으로 구성되었으며, 각 문장은 성인 및 아동 남녀의 목소리로 녹음되었다. 이 또한, 다양한 정서 상태를 반영하도록 연기되었으며, 정서 강도는 강한 정서와 약한 정서의 두 가지 조건으로 구분하여 제작되었다.

자극의 정서 표현 정확성과 강도를 평가하기 위해 여대생 10명을 대상으로 설문 시행하였다. 참가자들은 각 자극의 정서 및 강도를 평가하였으며, 정서 정확도는 다중 객관식 문항으로, 강도 평가는 10점 리커트 척도로 측정되었다. 이를 바탕으로 타당성이 높은 자극을 선별하고, 수정 작업을 거쳤다. 이후, 전문가 검토를 통해 자극의 정서 표현 정확성, 강도 적절성, 자극의 일관성을 추가적으로 평가하였다. 검토에는 심리학 교수 1인, 임상심리 또는 디지털 치료제 분야에서 연구 및 실무 경험이 있는 박사 과정생 2인, 그리고 석사 과정생 5인이 참여하였다. 이러한 전문가들의 피드백을 반영하여 최종 자극을 선정하였다.

**3단계: 타당도 검증 및 파일럿 테스트.** 초기 프로토타입 개발 후, 석사 및 박사 과정 연구원들이 UX/UI, 게임 구조, 측정 방식, 오류 등을 검토하였다. 연구원들은 사용자 경험과 측정 방식의 타당성을 평가하고, 개선이 필요한 부분에 대한 피드백을 제공하였다. 또

한, 6세 및 8세 아동 2인을 대상으로 타당성 평가를 진행하였다. 이를 통해 앱의 사용성, UX/UI, 디자인의 적절성 등을 검증하였다. 이러한 검토와 내부 테스트를 거쳐 총 42회의 앱 업데이트가 이루어졌으며, 이 과정에서 기술적 오류를 수정하고 안정성을 확보하였다. 파일럿 테스트는 ADHD나 아스퍼거 진단을 받은 아동 14명을 포함한 6-10세 아동 50명을 대상으로 이루어졌으며, 임상군과 정상군 간의 데이터를 비교함으로써 선별 측정치의 타당성을 평가하였다. 이러한 결과를 바탕으로 CEMO-A의 측정 방식을 최종적으로 보완하였다.

#### 생태학적 평가 방식

CEMO-A는 기존의 구조화된 검사 환경에서 단회적으로 시행되는 평가 도구들과 달리, 생태학적 심리 평가 방식을 반영하여 보다 유연한 환경에서 평가가 이루어질 수 있도록 설계되었다. 이는 아동이 특정한 검사 상황에 제한되지 않고, 반복적인 참여를 통해 자연스러운 평가 데이터를 도출할 수 있도록 지원하는 방식이다.

기존의 평가 방식은 단순히 한 번의 성공 또는 실패로 결과를 결정하는 제한적인 방법을 적용하는 경우가 많았다. 반면 CEMO-A는 정답률과 오답률을 통해 산출하는 최종 점수(FS) 및 ‘반복적 로그인을 허용하는 생태학적 상황의 심리 평가를 위한 변형된 계단식 절차 평가 방법(특허 고객번호: 220070060341)’에 기반한다. 이에 따라 특정 검사 환경에 국한되지 않으며, 유동적인 상황에서 반복적으로 참여하더라도 적절한 평가 결과를 도출할 수 있다. 이러한 접근 방식은 Song 외(2020)가 개발한 모바일 앱 ‘CoCon’의 평가 모델과 유사하

다. CoCon은 적응형 계단 절차(adaptive staircase procedure) 및 고유한 알고리즘을 적용하여 아동 및 청소년의 인지 기능을 평가하였으며, 신뢰성과 타당성이 입증된 평가 도구로 확인되었다. CEMO-A도 이러한 절차와 알고리즘을 사용해, 연속된 정답 수가 일정 수에 도달하면 난이도가 상승하고, 일정 수의 오답이 계속될 경우 난이도가 하락하는 방식으로 진행된다. 또한, CEMO-A는 기능성 게임화를 적용하여 아동의 몰입도를 높이고, 보다 현실적인 데이터를 수집할 수 있도록 설계되었다.

#### 측정도구

##### **CEMO-A 인지 정서 모니터링-평가 앱**

CEMO-A는 총 7개의 게임으로 구성되며, 이중 3개의 게임은 2-4개의 하위 소게임으로 이루어졌다. 이는 연구팀에서 준비한 갤럭시 태블릿을 통해 시행되었으며, 소요 시간은 약 1시간이었다.

**Find the Target! (FT, 잘 찾아보세요!).** 첫 번째 게임은 두 개의 소게임으로 구성된다. 미리 제시된 목표 자극만을 선택하는 연속 수행 검사(Continuous Performance Test [CTP]) 방식의 일종으로, 지속적인 주의력을 평가한다. 첫 번째 소게임(FT1)에서는 ‘공’ 그림, 두 번째 소게임(FT2)에서는 얼굴 표정 그림이 자극으로 제시된다. 두 번째 소게임은 정서 민감성도 평가할 수 있도록 추가 설계되었다.

**Whose Voice? (WV, 누구 목소리인지 찾으세요!).** 두 번째 게임은 정서 민감성 및 정서 인식 능력을 평가하고자 설계되었다. 이 게임은 목소리를 통해 다른 사람의 정서를 민감하

게 알아차리고 변별하는 능력을 측정하는 방식이다. 아동들은 정서가 포함된 목소리를 들은 후, 주어진 네 개의 얼굴 표정 중에서 해당 정서와 일치하는 그림을 선택하는 과제를 수행하게 된다.

**Remember! (R, 순서를 따라하세요!).** 세 번째 게임은 아동의 작업 기억을 평가하기 위해 설계되었으며, 총 4개의 소게임으로 구성되어 있다. 첫 번째(R1)와 세 번째(R3) 소게임에서는 자극을 일시적으로 기억한 후 제시된 보기 중에서 순서대로 선택해야 하며, 두 번째(R2)와 네 번째(R4) 소게임은 기억된 정보를 정신적으로 조작하여 역순으로 선택하는 방식이다.

**Stop Signal! (SS, 점프! 점프! 멈추세요!).** 네 번째 게임은 인지 통제 능력을 평가하기 위해 설계되어, 순간적으로 주의를 집중하고 불필요한 충동적 반응을 억제하며 적절한 행동을 선택하는 능력을 측정한다. 2개의 소게임 중 첫 번째 소게임(SS1)에서는 청각 신호가 주어지면 반복하던 행동을 멈추고 신속하게 새로운 반응을 해야하며, 두 번째 소게임(SS2)에서는 청각 또는 시각적인 신호가 불규칙하게 제시되어 상황에 따라 적절한 반응을 해야 한다.

**Driving a Car! (DC, 화살표를 따라 운전하세요!).** 다섯 번째 게임도 인지 통제 능력을 평가하기 위해 설계되었다. 이 게임은 화면에 제시되는 자동차 경주 장면을 보고 있다가 제시되는 시각적 자극에 따라 반응해야 하며, 청각적인 신호가 제시되는 경우에는 이에 적합한 반응을 해야 한다

**What was It? (WI, 무엇이었는지 기억하세요!).** 여섯 번째 게임은 인지적 유연성과 주의 전환 능력을 비롯한 실행 기능, 그리고 정서 민감성을 평가하기 위해 설계되었다. 색깔, 얼굴 표정, 목소리 자극 중 2가지나 3가지가 제시된 이후, 이 중 하나에 대한 질문을 듣고 답을 선택해야 한다.

**Cross the River (CR, 징검다리를 건너세요!).** 마지막 게임은 실행 기능을 평가하기 위한 것으로, 복잡한 목표를 달성하기 위해 계획, 추론, 주의 전환, 인지적 유연성 등의 능력을 복합적으로 측정한다. 이 게임에서는 자극의 개수, 색깔, 모양이 각기 다른 주사위들로 구성된 징검다리를 보고, 규칙에 맞는 주사위만 선택해 올바른 경로를 찾아야 한다. 이 때 규칙에 부합하더라도 아래칸부터 위칸까지 이어지지 않는 주사위는 선택하지 않아야 한다.

#### 한국 웨슬러 아동 지능검사 5판

전반적인 지적 능력을 평가하기 위해, Wechsler(2014)가 출판하고 광금주와 장승민(2019)이 국내 실정에 맞게 표준화한 한국 웨슬러 아동지능검사 5판(Korean Wechsler Intelligence Scale for Children - Fifth Edition [K-WISC-V])을 사용하였다. K-WISC-V는 6세 0개월부터 16세 11개월까지의 아동을 대상으로 개별적으로 실시하는 최신판 지능 검사로, 지능 이론뿐 아니라 인지 발달, 신경 발달, 인지 신경과학, 학습 과정에 대한 최근 심리학 연구들을 기반으로 제작되었다. 본 연구에서는 10개의 기본 소검사를 실시하여, 전체 지능 지수(FSIQ) 및 5개의 하위 지표(언어이해, 시공간, 유동추론, 작업기억, 처리속도) 점수 등을 산출하였다.

#### Stroop color and word test children's version 스트룹 아동 색상-단어 검사

전두엽에서 담당하는 인지적인 억제 과정의 효율성을 평가하기 위해, 스트룹 아동 색상-단어 검사(STROOP)를 사용하였다. 스트룹 검사는 Golden(1975)에 의해 개발되었으며, 본 연구에서 사용한 한국판은 신민섭과 박민주(2007)가 5세부터 14세까지 아동을 대상으로 표준화하였다. 스트룹 검사는 단어를 자동적으로 글자 그대로 읽는 단어 과제, 의식적인 노력을 기울여 색상을 명명하는 색상 과제, 그리고 단어의 색과 글자가 일치하지 않는 조건에서 자동화된 반응을 억제하고 색상을 명명해야 하는 색상-단어 과제로 구성된다(신민섭, 박민주, 2006). 한 과제 당 20개씩 5열로 100개의 자극이 제시되며, 45초의 제한 시간 내에 정확하게 읽어낸 자극의 수로 점수를 산출한다. 단어 과제는 ‘빨강’, ‘초록’, ‘파랑’ 단어가 검정색 글자로 제시된다. 색상 검사는 ‘XXXX’ 문자가 빨간색, 초록색, 파란색으로 제시되며, 색상-단어 과제는 ‘빨강’, ‘초록’, ‘파랑’ 단어가 불일치하는 색상으로 제시된다. 모든 자극은 무작위 순서로 배열되어 있으며, 어떤 단어나 색상도 열 내에서 연이어 반복되지 않는다. 정상 기준 집단에서 2주 간격의 검사-재검사 신뢰도 계수는 단어점수 .76, 색상점수 .66, 색상-단어 점수 .72로 나타났다(신민섭, 박민주, 2006).

#### CCTT \_Children's Color Trails Test 아동 색 선로 검사

지각 추적 능력, 정신 운동 속도, 순차적 처리 능력, 분할 주의력 등을 평가하기 위해, 신민섭과 구훈정(2007)이 표준화한 아동 색 선로 검사(CCTT)를 사용하였다. CCTT는 5세에서 15

세까지 시행 가능하며, 1부터 15까지의 숫자를 순서대로 연결하는 CCTT-1, 그리고 숫자와 색상 사이를 전환해가면서 각각의 순서를 따라가야 하는 CCTT-2로 구성되어 있다. 소요 시간은 약 10분 정도이며, 표준화 연구에서 2주 간격의 검사-재검사 신뢰도 계수는 CCTT-1은 .50, CCTT-2는 .75로 나타났다(신민섭, 구훈정, 2007).

**CBCL 6-18 \_Child Behavior Checklist for Ages 6-18 아동청소년 행동평가척도 부모용**

행동 및 정서적인 문제를 평가하기 위해, 오경자와 김영아(2010)가 표준화한 아동청소년 행동평가척도(CBCL)를 사용하였다. 해당 설문지는 6세에서 18세인 아동청소년에 대해 주 양육자가 작성하는 보호자용 질문지이다. CBCL은 내재화 문제(불안, 우울, 신체증상 등)와 외현화 문제(규칙위반, 공격행동 등)를 포함하는 증후군 척도, DSM 진단 척도, 문제행동특수 척도, 그리고 적응 척도 등으로 구성되어 있다. 총 120개 문항의 3점 리커트 척도이며, 응답시간은 약 15분-20분 정도 소요된다. 세부 측정 척도의 내적 일관성을 평가하는 Cronbach's  $\alpha$  계수는 .62-.95 사이에 분포되어 있다(오경자, 김영아, 2010).

**SCQ \_Social Communication Questionnaire 사회적 의사소통 설문지: 현재**

자폐 스펙트럼 장애(ASD) 관련 증상을 확인하기 위해, Rutter 외(2003)가 개발하고 유희정(2008)이 국내에서 표준화한 사회적 의사소통 설문지(SCQ)를 사용하였다. SCQ는 현재형과 일생형으로 두 가지 형태가 있는데, 본 연구에서는 지난 3개월의 행동에 관한 것을 평정하는 현재형 활용하였다. 현재형은 별도의 절

단점이 제공되지는 않으나, ASD 증상의 대략적인 정도를 알 수 있는 지표로 사용될 수 있다. 이는 총 40문항으로 구성되며, '사회적 상호작용', '의사소통', '제한적, 반복적, 상동적 행동 패턴' 등의 하위 영역으로 구분되어 있다. 설문지의 응답은 보호자가 간단하게 '예/아니오'로 답하는 형태이며, 응답에는 약 15분 정도 소요된다. 표준화 연구에서 내적 일관성을 분석한 결과, Cronbach's  $\alpha$  계수는 5-6세에서는 .84, 7~10세에서는 .89로 나타났다(유희정, 2008).

**연구 절차**

본 연구는 서울여자대학교 연구윤리위원회의 승인을 받은 후 진행되었으며(승인 번호: SWU IRB - 2022A - 53, 승인 날짜: 2023년 2월 22일), 6세에서 10세의 일반 아동 100명을 대상으로 하였다. 연구 참여자는 온라인 게시판과 SNS를 통해 모집된 자발적 참여자였고, 지능지수가 80 이상이며 시각이나 청각 장애가 없는 아동이 연구 대상으로 선정되었다. 최종 분석에서는 3명의 아동이 제외되어 총 97명의 데이터를 분석하였다.

참여를 희망한 아동과 보호자에게는 연구 내용을 충분히 설명하였고, 아동에게는 쉬운 표현으로 절차를 안내하였다. 보호자와 아동 모두의 서면 동의를 받았으며, 연구 도중 언제든지 동의를 철회할 수 있음을 고지하였다.

검사는 일대일 대면 방식으로, 방해 받지 않는 조용한 상태인 서울여자대학교 교육심리학과 실습실에서 진행되었다. 검사를 실시한 연구원들은 상담 및 임상심리 전공 석사 과정 이상인 연구원들로, 임상심리전문가 및 정신건강임상심리사 1급 자격을 소지한 박사 과

정생에 의한, 충분한 교육과 지도감독 하에 진행하였다. 아동은 먼저 CEMO-A 앱을 시행한 후, 기존의 아동용 심리 검사(K-WISC-V, Stroop, CCTT)를 순서대로 시행하였다. CEMO-A 앱을 통한 평가 과정에서 연구원은 회원가입 및 로그인 절차에 있어 도움을 제공했고, 이후 아동이 앱의 안내에 따라 과제를 수행하였다. 쉬는 시간을 제외한 검사 소요 시간은 총 2시간 30분 정도였으며, 아동의 피로도와 집중력을 고려하여 적절한 휴식 시간을 제공하며 진행하였다. 보호자는 CBCL 6-18, SCQ와 같은 보호자 보고용 질문지를 작성하였다.

연구 참여에 대한 보상으로 3만원이 지급되었으며, 표준화된 심리 검사에 대한 간략한 보고서도 제공되었다. 모든 개인 식별 정보는 안전하게 관리되었으며, 연구 종료 후 3년이 경과한 시점에 폐기될 예정이다. 연구 자료 분석을 위해 연구 대상자에게는 개별 아이디 번호가 부여되어 식별 불가능한 상태로 코딩 처리되었다.

#### 자료 분석

본 연구는 아동의 인지 정서 평가용 앱인 CEMO-A의 타당성을 검증하기 위한 목적으로 진행되었다. 이를 위해 일반 아동 100명을 대상으로 CEMO-A를 시행하고, 기존의 아동용 심리 검사 및 보호자 보고용 질문지와 같은 표준화된 심리 검사를 추가로 실시하였다. 연구 자료의 정리와 통계 분석은 IBM SPSS Statistics 30.0을 사용하여 수행하였으며, 데이터 시각화를 위해 RStudio를 활용하여 violin plot과 box plot을 결합한 그래프를 생성하였다.

먼저 기술 통계 및 빈도 분석 방법을 통해

연구 대상자의 연령과 성별 분포 등 기본적인 특성, 그리고 CEMO-A와 표준화된 심리검사 결과의 분포를 확인하였다. 타당도 검증에서는 구성 타당도와 공존 타당도를 평가하였다. 첫째, CEMO-A가 인지 및 정서 기능을 얼마나 잘 반영하는지 확인함으로써 구성 타당도를 검증하였다. 이를 위해 측정치 간의 상관 분석을 실시하여, 이론적으로 유사하거나 상이한 측정치 간의 관련성을 확인하였다. CEMO-A의 정서 인식 측정치(PER, SdR, AgR, AxR)는 데이터 분포가 특정 값에 집중되어 있어 비모수적인 방법인 Spearman 상관계수를 사용하였다. 나머지 변인들은 VIF 값이 10 미만으로 다중공선성 문제가 없음을 확인했으며, Kolmogorov-Smirnov (K-S) 검정 결과는 정규성을 충족하지 않았으나,  $N = 97$ 로 표본 크기가 충분하므로 중심극한정리에 근거하여 Pearson 상관계수를 적용하였다. 또한, 인지 측정치는 연령이 증가할수록 발달하고 정서 측정치는 연령에 따른 유의미한 차이가 없는 발달적 특성을 확인하기 위해, 연령 집단 간의 차이 분석을 실시하였다. 차이 검증에서는 등분산성이 충족된 경우 일원배치 분산분석(one-way ANOVA)과 Tukey의 HSD 사후 검정을 사용하였으며, 등분산성이 충족되지 않은 경우에는 Welch's ANOVA와 Games-Howell 사후 검정을 적용하였다. 정서 인식 측정치는 비모수적인 방법인 Kruskal-Wallis와 Bonferroni 수정 사후 검정을 적용하였다. 아울러, 연령별 점수 분포와 발달적 경향성을 보다 직관적으로 파악하기 위해, violin plot과 box plot을 결합하여 시각화하였다. 둘째, CEMO-A와 표준화된 심리 검사 결과들 간의 상관 분석을 실시하여 공존 타당도를 검증하였다. 이를 통해 동일한 개념을 측정하는 요소 간의 높은 관련성과 상이

한 개념을 측정하는 요소 간의 변별력이 있는지의 여부를 확인하였다. 이 또한 Pearson 상관계수와 Spearman 상관계수를 사용하여 분석하였다.

마지막으로, CBCL의 DSM 진단 척도에서 ADHD, 우울, 불안 장애가 의심되는 총 10명의 아동이 연구 결과에 미친 영향을 검토하기 위해, 해당 아동들을 제외한 후 동일한 방식으로 주요 분석을 수행하였다. 이를 통해 임상군으로 의심되는 아동의 포함 여부가 연구 결과에 미치는 영향을 평가하고, 주요 결과의 강건성을 확인하였다.

## 결 과

### 인구통계학적 정보

연구 대상자의 연령과 성별 분포는 표 1, 그리고 표준화된 심리 검사 결과를 확인하기 위한 기술 통계 분석 결과는 표 2에 제시하였다. 연구 대상은 총 100명이 모집되었으나, 제외 기준에 해당하는 1명과 CEMO-A 데이터 저장 오류가 발생한 2명을 제외한 97명의 데이터가 분석에 사용되었다.

참가자들의 K-WISC-V의 전체지능지수(FSIQ) 평균은 108.53( $SD = 11.89$ )이며, 하위 다섯 개의 지표 점수들의 평균은 모두 중간(100) 이상이었다. 신경심리검사인 STROOP의 색상-단어(CW) 및 CCTT-2의 T점수 평균은 각각 50.40( $SD = 14.90$ )과 50.09( $SD = 10.58$ )에 해당하였다.

표 1. 연구 대상자 연령 및 성별 분포 ( $N = 97$ )

성별	연령					전체
	6세	7세	8세	9세	10세	
남	10	8	10	10	9	47
여	10	12	10	8	10	50
전체	20	20	20	18	19	97

표 2. 표준화된 심리 검사에 대한 기술 통계 분석 결과 ( $N = 97$ )

	M	SD	최소값	최대값
WISC_FSIQ	108.53	11.89	81	131
STROOP_CW	50.40	14.90	18	103
CCTT-2	50.09	10.58	18	69
CBCL_TP	52.15	9.29	30	76
SCQ	2.49	2.27	0	14

주. WISC\_FSIQ = 웨슬러 아동 지능검사 5판\_전체지능지수(Full Scale IQ); STROOP\_CW = 스트룹 아동 색상-단어 검사\_색상-단어(Color-Word); CCTT-2 = 아동 색 선로 검사 2; CBCL\_TP = 아동청소년 행동평가척도 부모용\_문제행동 총점(Total Problems); SCQ = 사회적 의사소통 설문지

빈도 분석 결과, CBCL 상에서 문제행동 총 점을 기준으로 11명(11.3%)이 임상 범위인 64T 이상에 해당하였다. 또한, CBCL의 DSM 진단 척도를 기준으로 ADHD(4명), 우울 장애(2명), 불안 장애(5명)가 의심되는 아동이 포함되었으며, 이 중 1명은 ADHD와 불안 장애 모두에서 기준을 초과하는 점수를 보였다. SCQ 점수의 범위는 0에서 14점이었으며, 90% 이상의 참가자(90명)가 5점 이하의 낮은 점수를 나타냈다.

구성 타당도

CEMO-A의 핵심 측정치들 간의 상관 분석 결과는 표 3에 제시하였다. 인지 기능 측정인 SA, WM, CC, EF 간에는 유의미한 정적 상관이 확인되었다, SA-WM :  $r = .37, p < .01$ ; SA-CC :  $r = .46, p < .01$ ; SA-EF :  $r = .48, p < .01$ ; WM-CC :  $r = .27, p < .01$  ; WM-EF :  $r = .37, p < .01$  ; CC-EF :  $r = .40, p < .01$ . 정서 민감성(ES)은 인지 기능 측정치인 SA( $r =$

$.47, p < .01$ ), WM( $r = .20, p < .05$ ), CC( $r = .34, p < .01$ ), EF( $r = .40, p < .01$ )뿐만 아니라 정서 인식 측정치인 SdR( $r_s = .36, p < .01$ ) 및 AxR( $r_s = .71, p < .01$ )와도 유의미한 상관을 보였다. 한편, 정서 인식 측정치 중 SdR과 AxR 간( $r_s = .22, p < .05$ )에 유의미한 정적 상관이 있었고, AgR은 SdR( $r_s = -.37, p < .01$ ) 및 AxR( $r_s = -.54, p < .01$ )과 유의미한 부적 상관을 나타냈다. PER은 참가자 중 93.8%(91명)가 최고점(1.00)을 받았으며, 이로 인해 다른 주요 측정치들과의 상관이 유의하지 않았다.

CEMO-A 점수의 발달적 경향성을 확인하기 위해 연령 집단 간의 차이를 분석한 결과는 표 4와 같다. 첫 번째로 각 소개임들의 최종 점수에 대한 결과를 보면, 인지적인 요소를 포함한 게임인 FT1( $F(4, 92) = 3.18, p < .05$ ), FT2( $F(4, 92) = 3.26, p < .05$ ), R1( $F(4, 92) = 3.21, p < .05$ ), R2( $F(4, 92) = 5.23, p < .01$ ), R3( $F(4, 92) = 4.10, p < .01$ ), R4( $F(4, 92) = 9.11, p < .01$ ), DC( $F(4, 92) = 4.82, p < .01$ ),

표 3. CEMO-A 핵심 측정치들 간의 상관 분석 결과 (N = 97)

	SA <sup>1</sup>	WM <sup>1</sup>	CC <sup>1</sup>	EF <sup>1</sup>	ES <sup>1</sup>	PER <sup>2</sup>	SdR <sup>2</sup>	AgR <sup>2</sup>
WM <sup>1</sup>	.37**	-						
CC <sup>1</sup>	.46**	.27**	-					
EF <sup>1</sup>	.48**	.37**	.40**	-				
ES <sup>1</sup>	.47**	.20*	.34**	.40**	-			
PER <sup>2</sup>	.09	.02	.06	-.10	.09	-		
SdR <sup>2</sup>	.10	-.14	.04	-.11	.36**	.13	-	
AgR <sup>2</sup>	-.04	-.05	-.06	.08	-.12	-.12	-.37**	-
AxR <sup>2</sup>	.06	.03	.05	-.10	.71**	.09	.22*	-.54**

주. \* $p < .05$ , \*\* $p < .01$ ; <sup>1</sup> Pearson 상관계수, <sup>2</sup> Spearman 순위 상관계수; SA = Sustained Attention; WM = Working Memory; CC = Cognitive Control; EF = Executive Function; ES = Emotion Sensitivity; PER = Positive Emotion Recognition; SdR = Sad Recognition; AgR = Anger Recognition; AxR = Anxiety Recognition

표 4. CEMO-A의 소개임 최종 점수에 대한 연령 간의 차이 분석 결과 (N = 97)

		Levene's	F	p	Post-hoc	평균 점수 오름차순
소개임	FT1 <sup>2</sup>	3.74**	3.18	.022	-	6, 7, 8, 10, 9
	FT2 <sup>2</sup>	2.74*	3.25	.020	6 < 9*	6, 7, 8, 10, 9
	WV <sup>1</sup>	2.39	1.42	.235	-	9, 8, 6, 10, 7
	R1 <sup>1</sup>	1.54	3.21	.016	6 < 10*	6, 7, 8, 9, 10
	R2 <sup>2</sup>	14.11**	5.23	.002	6 < 8, 9, 10*	6, 7, 8, 9, 10
	R3 <sup>1</sup>	2.41	4.10	.004	6, 7 < 9*	6, 7, 8, 10, 9
	R4 <sup>1</sup>	.38	9.11	.000	6 < 8, 9, 10*, 7 < 9, 10*	6, 7, 8, 9, 10
	SS1 <sup>2</sup>	5.42**	.92	.460	-	6, 7, 8, 10, 9
	SS2 <sup>1</sup>	2.12	.97	.426	-	7, 9, 6, 8, 10
	DC <sup>2</sup>	2.99*	4.82	.003	6 < 9, 10*	6, 7, 8, 9, 10
	WI <sup>2</sup>	26.00**	6.80	.000	6 < 7, 8, 9, 10*	6, 8, 7, 9, 10
	CR <sup>1</sup>	.53	2.62	.040	6 < 9*	6, 8, 7, 10, 9
	핵심 측정치	SA <sup>2</sup>	3.48*	4.28	.005	6, 7 < 9*
WM <sup>2</sup>		6.46**	11.79	.000	6 < 8, 9, 10*, 7 < 9, 10*	6, 7, 8, 9, 10
CC <sup>2</sup>		3.87**	4.51	.004	6, 7 < 10*	6, 7, 8, 9, 10
EF <sup>2</sup>		3.06*	7.15	.000	6 < 7, 8, 9, 10*	6, 8, 7, 10, 9
ES <sup>2</sup>		4.33**	1.55	.240	-	6, 8, 9, 7, 10
PER <sup>3</sup>			5.50	.240	-	7, 8, 9, 6=10
SdR <sup>3</sup>			6.06	.195	-	8, 10, 7, 9, 6
AgR <sup>3</sup>			15.51	.004	6 < 7, 10*	6, 9, 8, 7, 10
AxR <sup>3</sup>		3.67	.453	-	9, 10, 8, 7, 6	
선별 측정치	ADHD <sup>2</sup>	3.022*	11.77	.000	6 < 7, 8, 9, 10*, 7 < 9, 10*, 8 < 9*	6, 7, 8, 10, 9
	ASD <sup>2</sup>	4.74**	4.95	.002	6 < 7, 9, 10*	6, 8, 7, 9, 10
	DEP <sup>1</sup>	.53	3.22	.016	7 < 9*	7, 8, 10, 6, 9
	ANX <sup>1</sup>	.77	.29	.884	-	6, 7, 8, 10, 9
	ST <sup>1</sup>	1.08	1.90	.118	-	7, 6, 8, 10, 9

주. \* $p < .05$ , \*\* $p < .01$ ; <sup>1</sup> 일원배치 분산분석(one-way ANOVA) 및 Tukey의 HSD 사후 검정; <sup>2</sup> Welch's ANOVA 및 Games-Howell 사후 검정; <sup>3</sup> Kruskal-Wallis 및 Bonferroni 수정 사후 검정; FT = Find the Target, WV=Whose Voice; R = Remember; SS = Stop Signal; DC = Driving a Car; WI = What was It; CR = Cross the River; SA=Sustained Attention; WM = Working Memory; CC = Cognitive Control; EF=Executive Function; ES = Emotion Sensitivity; PER = Positive Emotion Recognition; SdR = Sad Recognition; AgR=Anger Recognition; AxR=Anxiety Recognition; ADHD = Attention-Deficit/Hyperactivity Disorder; ASD = Autism Spectrum Disorder; DEP = Depressive Disorder; ANX = Anxiety Disorder; ST = Stress.

WI( $F(4, 92) = 6.80, p < .01$ ), CR( $F(4, 92) = 2.62, p < .05$ )에서는 연령 간 유의미한 차이가 확인되었으며, 연령이 증가할 수록 점수가 높아지는 경향성을 보였다. 그러나 SS1( $F(4, 92) = .92, m$ )과 SS2( $F(4, 92) = .97, m$ )에서는 각각 50.5%(49명)와 45.5%(44명)의 참가자가 최고 점(1.00)에 해당하여 연령 간 유의미한 차이는 없었다. 정서적인 요소만 포함한 게임인 WV는 연령 간 유의미한 차이가 발견되지 않았다,  $F(4, 92) = 1.42, p > .05$ .

그림 1은 CEMO-A의 소게임 중 연령별 점수 차이가 나타나는 방식이 서로 다른 5개 소게임의 점수 분포를 violin plot과 box plot으로 나타낸 것이다. FT1, DC, WI는 연령이 증가할 수록 점수가 최상위 값에 집중되는 경향을 보여, 일부 연령대에서 천장 효과(ceiling effect)가 관찰되었다. FT2와 R4는 연령에 따른 점수 차이가 유의미하게 나타난 대표적 소게임으로, 연령이 증가할수록 점수 향상이 뚜렷했다. 그리고, 정서 측정치인 WV는 연령 간 점수 차이가 유의미하게 나타나지 않은 변인으로, 정서 인식 능력은 연령에 따른 변화를 보이지 않는다는 점을 시각적으로 확인할 수 있다.

두 번째로 핵심 측정치 점수를 보면, 인지 기능 측정치인 SA( $F(4, 92) = 4.28, p < .01$ ), WM( $F(4, 92) = 11.79, p < .01$ ), CC( $F(4, 92) = 4.51, p < .01$ ), EF( $F(4, 92) = 7.15, p < .01$ )에서 연령 간 유의미한 차이가 있었으며, 연령이 증가할 수록 점수가 높은 양상이었다. 정서 기능 측정치인 ES( $F(4, 92) = 1.55, m$ ), PER( $H(4) = 5.50, m$ ), SdR( $H(4) = 6.06, m$ ), AxR( $H(4) = 3.67, m$ )에서는 연령 간 유의미한 차이가 없었다. 단, AgR의 경우 6세에서만 유의미하게 낮은 점수가 나타났으며,  $H(4) = 15.51, p < .01$ , 다른 연령대에서는 발달적 경

향이 관찰되지 않았다.

마지막으로 선별 측정치를 살펴보면, 그림 2는 CEMO-A의 선별 측정치 점수에 대한 연령별 분포를 violin plot과 box plot으로 나타낸 것이다. ADHD( $F(4, 92) = 11.77, p < .01$ )와 ASD( $F(4, 92) = 4.95, p < .01$ ) 점수는 연령 간 유의미한 차이가 있었으며, 연령이 증가할 수록 점수가 높아지는 경향을 보였다. 반면 ANX( $F(4, 92) = 15.51, p > .05$ ) 및 ST( $F(4, 92) = 1.90, p > .05$ ) 점수는 연령에 따른 유의미한 차이가 나타나지 않았으며, 시각화된 데이터에서도 연령에 따른 분명한 변화 양상이 관찰되지 않았다. DEP 점수는 7세와 9세 간의 차이가 통계적으로 유의미하게 나타났으나,  $F(4, 92) = 3.22, p < .05$ , 전체적인 연령 분포를 고려했을 때 일관된 증가 또는 감소 경향이 확인되지는 않았다.

추가적으로, CBCL의 DSM 진단 척도에서 ADHD, 우울, 불안 장애가 의심되는 10명의 아동을 제외한 후 동일한 분석을 수행한 결과, 전체적인 패턴과 통계적 유의성에 의미 있는 변화가 없었다. 이는 연구 결과가 일부 아동의 임상적 특성에 의해 크게 좌우되지 않으며, 본 연구에서 확인된 결과가 비교적 일관되게 유지됨을 시사한다.

### 공존 타당도

CEMO의 공존 타당도, 즉 수렴 및 변별 타당도를 검증하기 위해 CEMO의 영역(소게임 최종 점수, 핵심 측정치, 선별 측정치)별 점수들과 표준화된 검사 결과 간의 상관관계를 분석하였다.

첫째, 소게임 최종 점수와 표준화된 검사 간의 상관분석 결과는 표 5에 제시하였다. 지

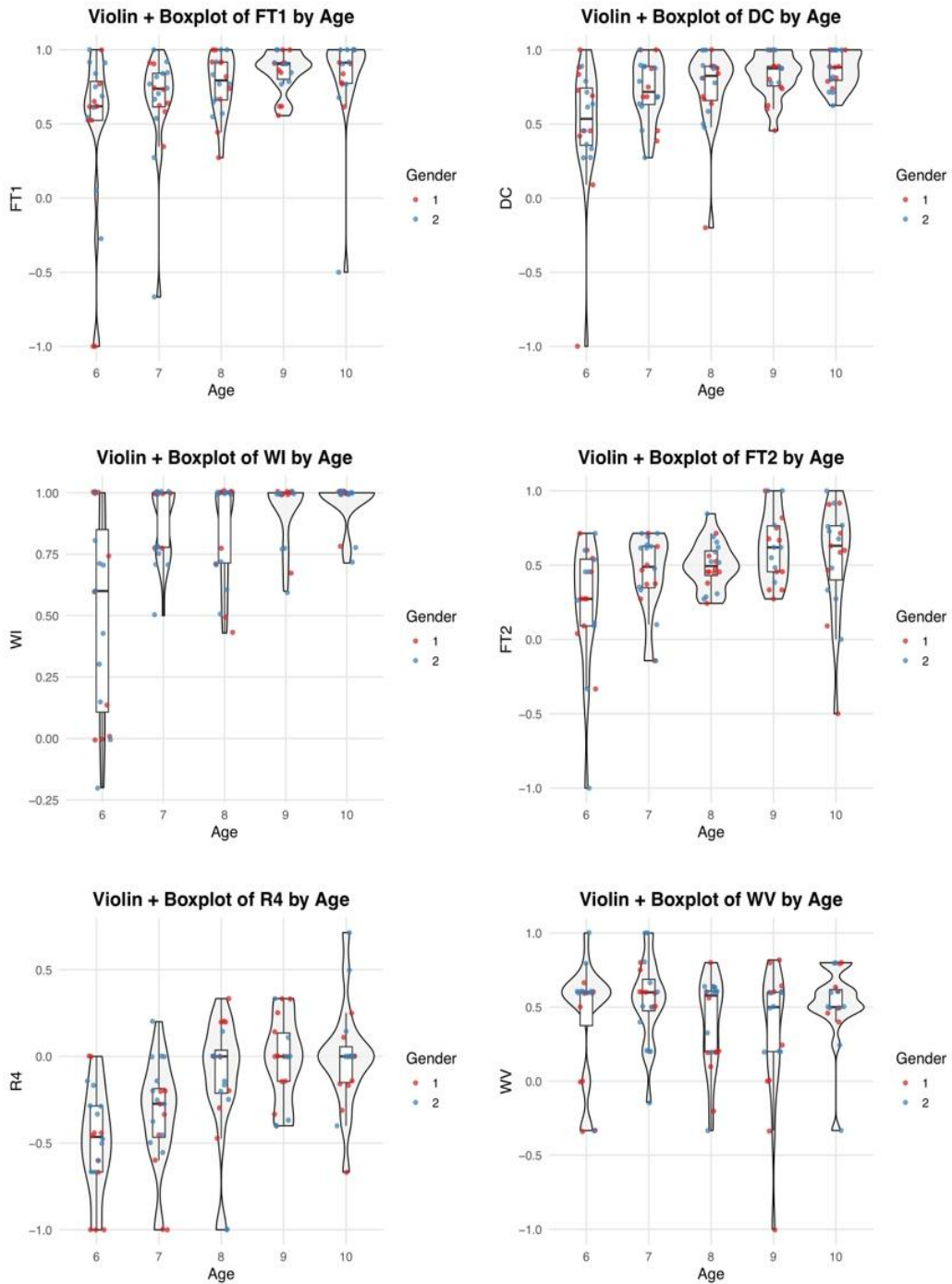


그림 1. CEMO-A 소개임 최종 점수의 연령별 분포

주. FT = Find the Target; DC = Driving a Car; WI = What was It; R = Remember; WV = Whose Voice.

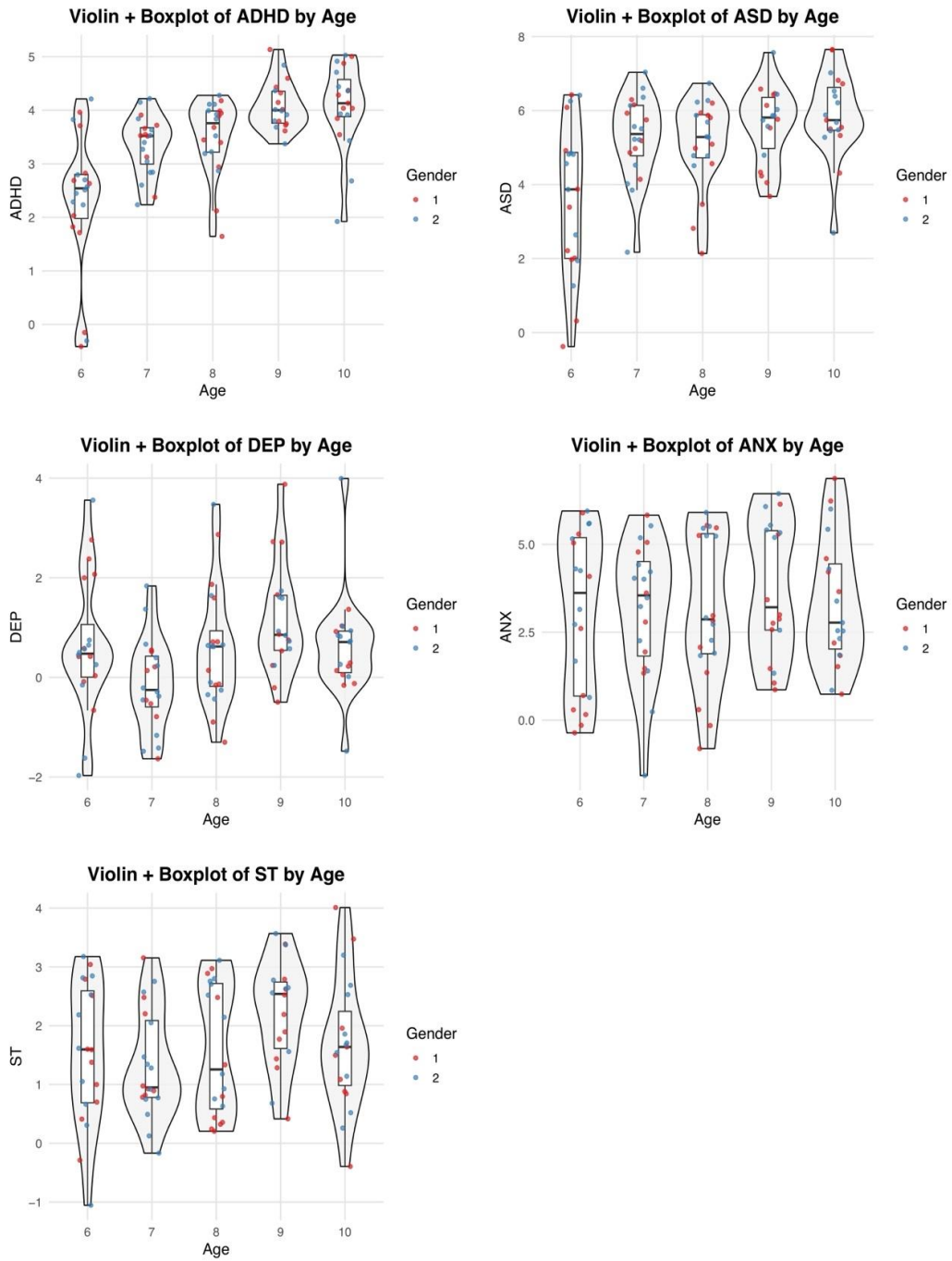


그림 2. CEMO-A 선별 측정치 점수의 연령별 분포

ADHD = Attention-Deficit/Hyperactivity Disorder; ASD = Autism Spectrum Disorder; DEP=Depressive Disorder;  
 ANX = Anxiety Disorder; ST = Stress

표 5. CEMO-A 소개임 최종 점수와 표준화된 검사 간의 상관 분석 결과 (N = 97)

	FT1	FT2	WV	R1	R2	R3	R4	SS1	SS2	DC	WI	CR
WISC_FSIQ	.28**	.29**	-.03	.34**	.53**	.13	.18	.10	.16	.29**	.26**	.42**
WISC_GAI	.23*	.22*	-.06	.34**	.45**	.17	.08	.11	.18	.21*	.22*	.39**
WISC_CPI	.32**	.32**	.07	.27**	.42**	.08	.32**	.11	.07	.37**	.31**	.40**
STROOP_CW	.23*	.18	-.05	.22*	.25*	.08	.32**	-.15	.07	.26**	.13	.20
CCIT-2	.03	.05	-.04	.08	.16	-.02	.31**	-.04	.13	.25*	.18	.10
SCQ	.10	.23*	.03	-.04	-.01	.25*	.06	-.01	-.16	-.06	.13	-.02
CBCL_AP	.00	-.01	-.04	-.01	.00	-.01	.05	.05	.06	.07	-.01	.00
CBCL_AxP	-.08	.03	-.01	-.05	.08	-.03	-.09	.15	.00	-.20*	-.02	-.04
CBCL_SP	-.14	.04	-.07	.04	-.11	.05	.08	.05	.08	-.01	-.02	-.02
CBCL_ADHP	-.06	-.04	-.02	-.06	-.27**	.07	-.09	.07	-.13	-.22*	-.19	-.18
CBCL_ODP	-.09	-.01	.04	-.00	-.03	.04	-.08	.07	-.02	-.04	-.03	.03
CBCL_CP	-.14	-.16	-.04	-.05	-.07	.13	.05	-.08	.04	.04	-.03	-.06
CBCL_TAF	.06	.08	.05	.21	.15	-.04	.17	-.14	.11	-.01	.22*	.27*
CBCL_SC	.05	.06	.13	.10	.03	-.10	.02	.07	.09	-.02	.09	.25*
CBCL_AP	.15	.08	-.01	.30**	.28**	.08	.29**	-.18	.16	.11	.26*	.20

주. \*  $p < .05$ , \*\*  $p < .01$ ; FT = Find the Target; WV = Whose Voice; R = Remember; SS = Stop Signal; DC = Driving a Car; WI = What was It; CR = Cross the River; WISC\_FSIQ = 웨슬러 아동 지능검사 5판\_전체지능지수; GAI = 일반 능력 지표; CPI = 인지 효율 지표; STROOP\_CW = 스트룹 아동 색상-단어 검사\_색상-단어 점수; CCIT-2 = 아동 색 선로 검사 2; SCQ = 사회적 의사소통 설문지; CBCL\_AP = 아동정소년 행동평가척도 부모용\_정서 문제; AxP = 불안 문제; SP = 신체화 문제; ADHP = 주의력-결핍/과잉행동 문제; ODP = 반항행동 문제; CP = 품행 문제; TAF = 적응척도 총점; SC = 사회성; AP = 학업 수행.

능 검사의 FSIQ는 인지적 요소를 포함한 소계임인 FT1( $r = .28, p < .01$ ), FT2( $r = .29, p < .01$ ), R1( $r = .34, p < .01$ ), R2( $r = .53, p < .01$ ), DC( $r = .29, p < .01$ ), WI( $r = .26, p < .01$ ), CR( $r = .42, p < .01$ )과 유의미한 정적 상관을 보인 반면, 정서적인 요소만 포함한 소계임인 WV( $r = -.03, ns$ )와는 유의한 상관이 나타나지 않았다. FSIQ와 유의한 상관을 나타낸 소계임들은 GAI와 CPI 모두에서 유의한 상관을 보인 한편, R4는 FSIQ( $r = .18, ns$ ), GAI( $r = .08, ns$ )와 유의한 상관이 없었지만 CPI( $r = .32, p < .01$ )와는 유의미한 정적 상관을 나타냈다. STROOP의 CW는 FT1( $r = .23, p < .05$ ), R1( $r = .22, p < .05$ ), R2( $r = .25, p < .05$ ), R4( $r = .32, p < .01$ ), DC( $r = .26, p < .01$ )와 유의미한 상관을 보였으며, R4( $r = .31, p < .01$ )와 DC( $r = .25, p < .05$ )는 CCTT-2와도 유의한 상관을 보였다. SCQ는 FT2( $r = .23, p < .05$ ) 및 R3( $r = .25, p < .05$ )와 유의미한 정적 상관을 보였다. CBCL의 DSM 진단척도 중 주의력-결핍/과잉행동 문제는 R2( $r = -.27, p < .01$ ), DC( $r = -.22, p < .05$ )와 유의미한 부적 상관을, 불안 문제는 DC( $r = -.20, p < .05$ )와 부적 상관을 나타냈다. CBCL 적응척도 상에서 총점은 WI( $r = .22, p < .05$ ), CR( $r = .27, p < .05$ )와 유의미한 정적 상관을 나타냈다. 그리고 적응척도의 소척도 중 사회성은 CR( $r = .25, p < .05$ )과, 학업수행은 R1( $r = .30, p < .01$ ), R2( $r = .28, p < .01$ ), R4( $r = .29, p < .01$ ), WI( $r = .26, p < .05$ )와 정적인 상관을 보였다. CBCL의 문제행동 증후군 척도 및 문제행동 특수 척도 프로파일에서는 유의미한 상관이 나타나지 않았다.

둘째, 핵심 측정치와 표준화된 검사 간 상관 분석 결과는 표 6에 제시하였다. 인지 기

능 측정치부터 살펴보면, 지능검사의 FSIQ와 GAI, CPI는 인지 기능 측정치 SA, WM, CC, EF와 유의미한 정적 상관을 보였다. FSIQ-SA :  $r = .41, p < .01$  ; FSIQ-WM :  $r = .32, p < .01$  ; FSIQ-CC :  $r = .33, p < .01$  ; FSIQ-EF :  $r = .43, p < .01$ . STROOP의 CW는 SA( $r = .27, p < .01$ ), WM( $r = .27, p < .01$ ), EF( $r = .21, p < .05$ )와, CCTT-2는 CC( $r = .27, p < .01$ )와 유의한 정적 상관을 나타냈다. CBCL의 문제행동 증후군 척도에서 사고 문제는 CC( $r = -.20, p < .05$ )와 유의미한 부적 상관을 보였다. CBCL의 DSM 진단 척도에서 주의력-결핍/과잉행동 문제는 CC( $r = -.22, p < .05$ ), EF( $r = .23, p < .05$ )와, 품행문제는 SA( $r = -.21, p < .05$ )와 유의미한 부적 상관을 나타냈다. CBCL 적응 척도에서는 총점 및 사회성이 EF( $r = .30, p < .01$ ;  $r = .22, p < .05$ )와, 학업수행은 EF( $r = .28, p < .01$ ) 및 WM( $r = .25, p < .05$ )와 유의한 상관을 보였다. 한편 정서 기능 측정치를 살펴보면, 지능 검사 중에서는 CPI와 ES( $r = .33, p < .01$ ) 간에만 유의미한 정적 상관이 확인되었고, CCTT-2는 AgR( $r = -.21, p < .05$ )과 유의한 부적 상관을 나타냈다. CBCL에서는 불안/우울( $r_s = -.24, p < .05$ ), 불안 문제( $r_s = -.23, p < .05$ ), 강박 증상( $r_s = -.26, p < .05$ )이 PER과 유의한 부적 상관이 있었고, 사고 문제( $r_s = .20, p < .05$ )와 강박 증상( $r_s = .21, p < .05$ )은 AgR과 유의한 정적 상관을 보였다. CBCL의 문제행동증후군 척도 종합점수 세 가지는 유의미한 상관이 나타나지 않았다.

셋째, 선별 측정치와 표준화된 검사 간의 상관 분석 결과는 표 7에 제시하였다. FSIQ는 ADHD( $r = .50, p < .01$ ), ASD( $r = .37, p < .01$ ), ANX( $r = .20, p < .05$ ), ST( $r = .27, p < .01$ )와 유의미한 상관을 나타냈다. 그러나

표 6. CEMO-A 핵심 측정지와 표준화된 검사 간의 상관 분석 결과 (N = 97)

	SA <sup>1</sup>	WM <sup>1</sup>	CC <sup>1</sup>	EF <sup>1</sup>	ES <sup>1</sup>	PER <sup>2</sup>	SDR <sup>2</sup>	AGR <sup>2</sup>	AxR <sup>2</sup>
WISC_FSIQ	.41**	.32**	.33**	.43**	.17	.02	.07	-.08	.02
WISC_GAI	.33**	.31**	.26**	.38**	.05	.00	.03	.00	-.09
WISC_CPI	.46**	.26**	.38**	.45**	.33**	.02	.15	-.18	.18
STROOP_CW	.27**	.27**	.20	.21*	.10	.09	-.14	-.01	.02
CCIT-2	.10	.15	.27**	.17	.07	-.03	.16	-.21*	.04
SCQ	.13	.15	-.11	.06	.19	.00	-.18	.05	.07
CBCL_pAD	-.08	.06	-.07	-.07	-.01	-.24*	-.09	.15	-.03
CBCL_pWD	-.03	.06	.05	.12	.09	-.12	-.06	.01	.09
CBCL_pSC	.02	.09	.10	-.01	.09	-.12	-.17	.11	.03
CBCL_pSP	-.16	-.03	-.13	-.14	-.08	-.10	-.19	.20*	-.12
CBCL_pTP	-.15	-.03	-.20*	-.09	-.02	-.20	-.19	.04	.08
CBCL_pAP	-.07	.01	-.07	-.09	.01	-.11	-.15	.07	.00
CBCL_pRB	-.05	-.05	.02	.03	.03	.07	-.10	.07	.01
CBCL_pAB	-.17	-.03	-.10	-.13	-.02	.13	-.08	-.01	.12
CBCL_pOP	-.03	-.14	-.06	-.14	.00	-.08	-.07	.10	-.02
CBCL_AFP	-.03	.03	.06	-.00	-.02	-.10	-.15	.02	.04
CBCL_AxP	-.06	-.00	-.13	-.04	.01	-.25*	-.04	.13	-.07
CBCL_SP	-.11	.04	-.00	-.02	.00	-.14	-.15	.11	-.04
CBCL_ADHIP	-.12	-.06	-.22*	-.23*	-.04	-.10	-.17	.07	.05
CBCL_ODP	-.09	-.02	-.04	.01	.04	.14	-.14	.03	.09
CBCL_CP	-.21*	.08	-.01	-.06	-.13	.03	-.15	.14	-.09

표 6. CEMO-A 핵심 측정지와 표준화된 검사 간의 상관 분석 결과 (N = 97) (계속)

	SA <sup>1</sup>	WM <sup>1</sup>	CC <sup>1</sup>	EF <sup>1</sup>	ES <sup>1</sup>	PER <sup>2</sup>	SdR <sup>2</sup>	AgR <sup>2</sup>	AxR <sup>2</sup>
CBCL_OCP	-.14	-.02	-.12	-.06	-.02	-.26*	-.12	.21*	-.06
CBCL_PTSP	-.12	.03	-.13	-.06	.00	-.19	-.15	.13	.07
CBCL_SCT	.08	.16	.05	.10	.06	-.09	-.03	.12	-.14
CBCL_TAF	.11	.13	.00	.30**	.07	.05	.02	.06	-.04
CBCL_SC	.08	.01	.05	.22*	.17	.04	.03	.02	.10
CBCL_AP	.17	.26*	.11	.28**	.02	.02	-.06	.15	-.17

주. \*  $p < .05$ , \*\*  $p < .01$ ; <sup>1</sup> Pearson 상관계수; <sup>2</sup> Spearman 순위 상관계수; SA = Sustained Attention; WM = Working Memory; CC = Cognitive Control; EF = Executive Function; ES = Emotion Sensitivity; PER = Positive Emotion Recognition; SdR = Sad Recognition; AgR = Anger Recognition; AxR = Anxiety; WISC\_FSIQ = 웨슬러 아동 지능검사 5판\_진체지능지수; GAI = 일반 능력 지표; CPI = 인지 효율 지표; STROOP\_CW = 스트룹 아동 색상-단어 검사\_색상-단어 점수; CCTT-2 = 아동 색 선로 검사 2; SCQ = 사회적 의사소통 설문지; CBCL\_pAD = 아동정소년 행동평가척도 부모용\_불안/우울; pWD = 위축/우울; pSC = 신체 증상; pSP = 사회적 미성숙; pTP = 사고 문제; pAP = 주의집중 문제; pRB = 규칙 위반; pAB = 공격 행동; pOP = 기타 문제; AP = 정서 문제; AxP = 불안 문제; SP = 신체화 문제; ADHP = 주의력-결핍/과잉행동 문제; ODP = 반항행동 문제; CP = 품행 문제; OCP = 강박증상; PTSP = 외상후스트레스 문제; SCT = 인지속도 부진; TAF = 적응척도 총점; SC = 사회성; AP = 학업 수행.

표 7. CEMO-A 선별 측정치와 표준화된 검사 간의 상관 분석 결과 (N = 97)

	ADHD	ASD	DEP	ANX	ST
WISC_FSIQ	.50**	.37**	.17	.20*	.27**
WISC_GAI	.43**	.27**	.17	.07	.17
WISC_CPI	.53**	.49**	.13	.39**	.42**
STROOP_CW	.30**	.22*	.18	.12	.13
CCTT-2	.25*	.19	.11	.18	.28**
SCQ	.05	.13	.05	.16	.15
CBCL_pAD	-.07	-.04	-.20*	-.10	-.18
CBCL_pWD	.08	.12	-.06	.07	.04
CBCL_pSC	.05	.09	-.06	.01	-.04
CBCL_pSP	-.17	-.13	-.03	-.17	-.18
CBCL_pTP	-.16	-.10	-.09	-.06	-.14
CBCL_pAP	-.09	-.04	-.01	.00	-.02
CBCL_pRB	.00	.03	.00	-.03	-.06
CBCL_pAB	-.15	-.08	.01	-.02	-.07
CBCL_pOP	-.13	-.08	-.10	-.07	-.14
CBCL_AfP	.02	.00	-.03	-.06	-.10
CBCL_AxP	-.09	-.05	-.18	-.06	-.12
CBCL_SP	-.03	-.00	-.05	-.03	-.06
CBCL_ADHP	-.24*	-.17	-.02	-.03	-.09
CBCL_ODP	-.04	.02	.00	.02	-.03
CBCL_CP	-.07	-.09	.05	-.14	-.14
CBCL_OCP	-.12	-.07	-.28**	-.15	-.23*
CBCL_PTSP	-.11	-.05	-.14	-.07	-.15
CBCL_SCT	.12	.11	-.03	-.04	.00
CBCL_TAF	.21	.16	-.01	-.01	.02
CBCL_SC	.15	.19	-.05	.13	.08
CBCL_AP	.28**	.17	.04	-.06	.01

주. \* $p < .05$ , \*\* $p < .01$ ; ADHD = Attention-Deficit/Hyperactivity Disorder; ASD = Autism Spectrum Disorder; DEP = Depressive Disorder; ANX = Anxiety Disorder; ST = Stress; WISC\_FSIQ = 웨슬러 아동 지능검사 5판\_전체지능지수; GAI = 일반 능력 지표; CPI = 인지 효율 지표; STROOP\_CW = 스트룹 아동 색상-단어 검사\_색상-단어 점수; CCTT-2 = 아동 색 선로 검사 2; SCQ = 사회적 의사소통 설문지; CBCL\_pAD = 아동청소년 행동평가척도 부모용\_불안/우울; pWD = 위축/우울; pSC = 신체 증상; pSP = 사회적 미성숙; pTP = 사고 문제; pAP = 주의 집중 문제; pRB = 규칙 위반; pAB=공격 행동; pOP = 기타 문제; AfP=정서 문제; AxP = 불안 문제; SP = 신체화 문제; ADHP = 주의력-결핍/과잉행동 문제; ODP = 반항행동 문제; CP = 품행 문제; OCP = 강박증상; PTSP = 외상후스트레스 문제; SCT = 인지속도 부진; TAF = 적응척도 총점; SC = 사회성; AP = 학업 수행.

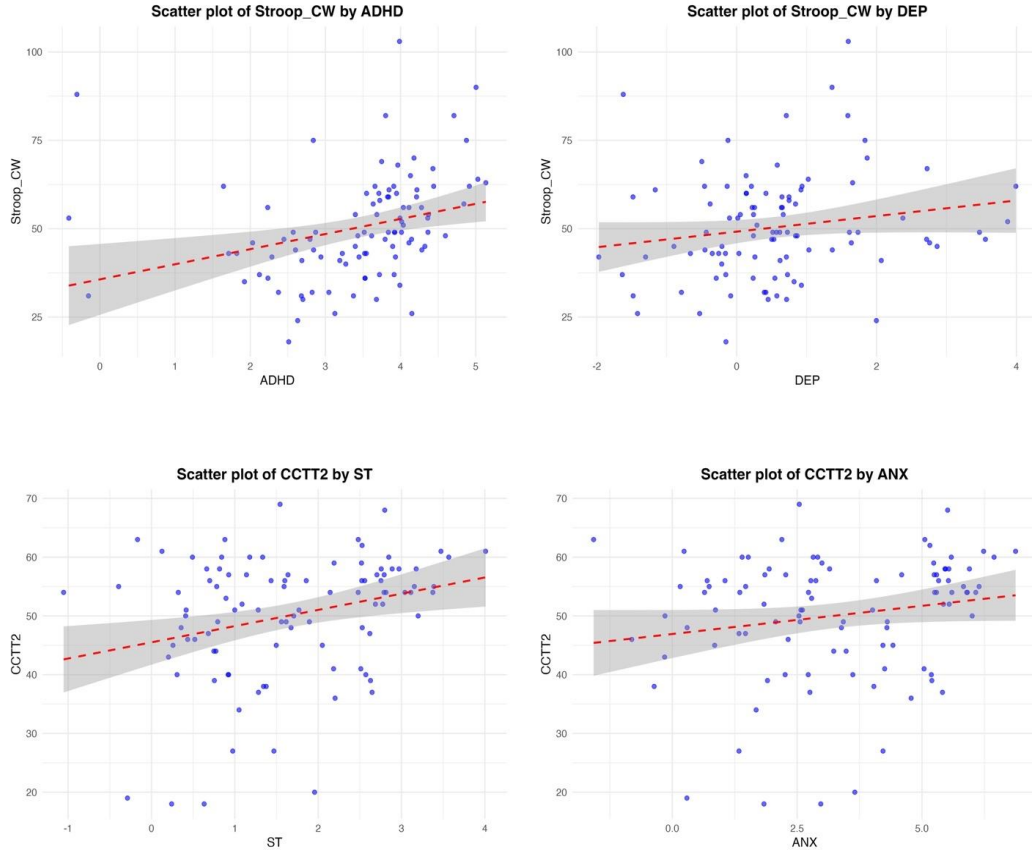


그림 3. CEMO-A 선별 측정치와 신경심리검사 점수의 산포도

주. ADHD = Attention-Deficit/Hyperactivity Disorder; DEP = Depressive Disorder; ANX = Anxiety Disorder; ST=Stress; STROOP\_CW = 스트룹 아동 색상-단어 검사\_색상-단어 점수; CCTT-2 = 아동 색 선로 검사 2.

ADHD와 ASD는 GAI 및 CPI 모두와 유의미한 상관을 나타낸 반면, ANX와 ST는 CPI와만 유의미한 상관을 보였다, GAI-ADHD :  $r = .43$ ,  $p < .01$ ; GAI-ASD :  $r = .27$ ,  $p < .01$ ; CPI-ADHD :  $r = .53$ ,  $p < .01$ ; CPI-ASD :  $r = .49$ ,  $p < .01$ ; GAI-ANX :  $r = .07$ ,  $ns$ ; GAI-ST :  $r = .17$ ,  $ns$ ; CPI-ANX :  $r = .39$ ,  $p < .01$ ; CPI-ST  $r = .42$ ,  $p < .01$ . STROOP의 CW는 ADHD( $r = .30$ ,  $p < .01$ ), ASD( $r = .22$ ,  $p < .05$ )와 유의한 상관을, CCTT-2는 ADHD( $r =$

$.24$ ,  $p < .05$ ), ST( $r = .28$ ,  $p < .01$ )와 유의미한 상관을 보였다. 그런데 그림 3에 제시한 산점도를 확인한 결과, 극단값(outlier)이 전체적인 상관 구조에 미치는 영향이 크다는 점이 확인되었으며, 특정 데이터 하나를 제외한 후 다시 분석한 결과 상관계수가 다소 증가하였다. 이처럼 특정 극단값으로 인해 Pearson 상관계수가 과소평가될 가능성이 있어, 보완적으로 비모수 검정인 Spearman 순위 상관계수를 추가적으로 확인하였다. Spearman 상관 분석

결과, STROOP의 CW는 ADHD( $r_s = .50, p < .01$ ), ASD( $r_s = .34, p < .01$ )와, CCTT-2는 ADHD( $r_s = .27, p < .01$ ), ST( $r_s = .27, p < .01$ )와 유의미한 상관을 보였으며, 상관계수가 증가하는 양상을 보였다. 또한, Spearman 분석에서는 추가적으로 STROOP의 CW와 DEP( $r_s = .22, p < .05$ ), CCTT-2와 ANX( $r_s = .21, p < .05$ ) 간의 유의미한 상관이 확인되었다. CBCL의 DSM 진단 척도에서는 주의력-결핍/과잉행동 문제가 CEMO-A의 ADHD( $r = -.24, p < .01$ )와 유의미한 상관이 있었고, CBCL 적응 척도의 학업 수행은 ADHD( $r = .28, p < .01$ )와 유의미한 상관을 나타냈다. 그리고 CBCL의 문제행동중후군 척도와 문제행동 특수 척도에서는 불안/우울이 DEP( $r = -.20, p < .05$ )와, 강박 증상은 DEP( $r = -.28, p < .01$ ) 및 ST( $r = -.23, p < .05$ )와 유의미한 부적 상관을 보였다. CBCL의 문제행동중후군 척도 종합점수 세 가지는 유의미한 상관이 없었다.

이에 더하여, 동일한 분석을 ADHD, 우울, 불안 장애가 의심되는 아동 10명을 제외한 후 반복하여 수행한 결과, 유의미한 변화가 나타나지 않았다. 이는 본 연구에서 확인된 CEMO-A의 공존 타당도가 일부 연구 대상자의 특성에 의해 크게 영향을 받지 않으며, 비교적 안정적인 측정 도구로 기능할 가능성을 의미한다.

## 논 의

본 연구는 CEMO-A의 구성 타당도와 공존 타당도를 분석하여 인지 및 정서 기능 평가의 타당성을 검토하였다. 우선 CEMO-A의 핵심 측정치 간 상관 분석을 통해 구성 타당도를

평가한 결과, 인지 기능 및 정서 인식이라는 두 영역 내의 각 측정치 간에 유의미한 상관이 확인된 반면, 두 영역 간에는 유의미한 상관이 나타나지 않았다. 이는 CEMO-A가 인지와 정서 기능을 독립적으로 평가할 수 있는 구성적 타당성을 입증하는 결과로 해석될 수 있다. 또한, ES는 두 영역 모두와 유의미한 상관을 나타냈는데, 이는 정서 민감성이 감정의 미세한 변화를 포착하고 해석하는 데 있어 인지적 요소도 필요로 하는 개념임을 반영하여 타당한 결과로 볼 수 있다. SdR과 AxR 간에 정적인 상관이 나타난 반면, AgR과 SdR, AgR과 AxR 간에는 부적인 상관이 나타났다. 이는 우울과 불안이 유사한 기제를 공유하는 반면, 분노와 우울/불안 간에는 주의 및 해석 편향이 발생하여 다른 정서 인식이 약화되는 양상으로 해석될 수 있다(Klein et al., 2018; Maoz et al., 2017). 한편, PER은 천장 효과(ceiling effect)로 인해 다른 정서 기능 측정치와 유의미한 상관을 나타내지 않은 것으로 여겨진다. 이로 보아 자극 제시 시간이나 반응 시간을 줄이는 등의 방법을 통해 PER의 난이도를 조정하면, 변별력이 더 증가할 것으로 기대된다.

다음으로, CEMO-A의 소게임 점수와 측정치들의 연령 차이를 분석하여 구성 타당도를 평가하였다. 그 결과, 인지적인 영역에서는 연령 간 유의미한 차이와 함께 연령이 증가할수록 점수가 높아지는 경향성이 뚜렷하여, 연령에 따른 인지 기능 발달 특성이 확인되었다. 다만, FT1, SS, DC, WI은 연령이 증가할수록 천장 효과가 관찰되어, 이러한 과제에서는 난이도 조정을 통해 연령에 따른 변별력을 확보할 필요성이 시사되었다. 정서적 영역에서는 연령 간 유의미한 차이가 나타나지 않아, 발달

적인 현저한 변화 없이 안정된 양상을 보였다. 이러한 발달적 차이는 CEMO-A의 구성적 타당성을 추가적으로 입증하는 결과로 해석될 수 있겠다.

마지막으로, 표준화된 심리 검사 결과와의 상관 분석을 통해 CEMO-A의 공존 타당도를 평가하였다. CEMO-A의 인지적 영역은 지능 검사, 신경심리 검사, 그리고 보호자가 보고한 학업 수행 능력과 유의미한 상관을 나타내어 신뢰할 수 있는 평가 지표로 활용될 가능성을 입증하였다. 특히 CEMO-A의 소계임 중 기하학적 이미지를 거꾸로 기억해야 하는 R4는 GAI와 유의미한 상관이 없는 반면 CPI 및 신경심리 검사 결과와 유의한 상관을 나타냈다. 이는 R4가 일반적인 지적 능력이 아닌 보다 순수한 전두엽 실행 기능을 평가하는 과제일 가능성을 시사하며, 이러한 실행 기능의 결함을 나타내는 ADHD 및 ASD를 선별하는 데 있어 핵심적으로 활용될 수 있을 거라 기대된다. 또 흥미로운 점은, FT2와 R3가 SCQ와 유의미한 정적 상관을 보였다는 점이다. 이는 얼굴 표정 그림(FT2)과 기하학적 이미지(R3)를 활용한 과제에서 높은 성취를 보이는 아동들이 보호자 보고 상 자폐 스펙트럼 관련 증상이 많은 양상임을 나타낸다. 선행 연구(Mottron et al., 2006)에 따르면, ASD를 지닌 사람들은 시각 피질의 후방 영역을 주로 활성화하여 복잡한 시각 과제에서 세부적인 정보 처리를 수행하는 경향이 있으며, 이는 단순하거나 정적인 시각 자극을 더 잘 구별하는 능력과 관련이 있다. 추후 연구를 통해 FT2와 R3가 이러한 ASD의 특성을 타당하게 평가할 수 있음을 확인하고 이를 선별 측정치에 반영한다면, ASD 선별에 대한 CEMO-A의 정확성과 민감성이 더 증가할 것으로 예상된다.

CEMO-A의 정서 요소가 포함된 소계임 및 정서 기능 핵심 측정치는 인지적 능력과 유의미한 상관을 보이지 않아, 정서와 인지 기능 간 상대적인 변별 타당성이 확보되었음을 확인하였다. 또한, AgR은 분할 주의력 평가인 CCTT-2와 유의한 부적 상관을 나타내어, 주의 전환이 어려운 경우 분노 인식이 강해질 수 있음을 나타냈다. 이는 주의 전환이 어렵거나 인지적 경직성이 높은 경우, 분노 자극에 대한 주의 편향이 강화될 수 있음을 시사하며, 선행 연구(Premkumar et al., 2008)의 결과와 일치하는 결과이다. 선별 측정치 중 DEP는 STROOP의 CW와, ANX 및 ST는 CCTT-2 및 CPI와 정적 상관을 보였고, DEP 및 ST는 CBCL의 강박 증상과 유의한 부적 상관을 나타냈다. 이러한 결과는 우울할수록 억제 기능의 어려움이 증가하고, 불안 및 스트레스가 높을수록 분할 주의력과 인지 효율성이 저하되며, 우울 및 스트레스 수준이 높을수록 강박 관련 증상이 증가하는 경향을 시사한다. 이는 선행 연구(Moniz et al., 2016) 결과와 일치하며, CEMO-A의 정서 관련 측정치의 공존 타당성을 지지하는 근거가 될 수 있다. 이 중 일부 결과는 극단값의 영향을 보완하기 위해 추가적으로 시행한 비모수 검정에서만 유의하게 나타난 결과이지만, 향후 연구에서 더 많은 연구 대상을 확보하고 임상군을 포함한다면, 이러한 관계를 보다 명확하게 입증할 수 있을 거라 기대된다. 한편, 현재 CEMO-A의 정서 기능 측정치에서 점수 분포가 제한적으로 나타나 변별력이 다소 낮은 경향이 확인되었다. 특히 일부 측정치의 경우 천장 효과로 인해 개별 차이를 정교하게 구별하기 어려운 문제가 발생할 수 있다. 이를 보완하기 위해, CEMO-A의 고도화 과정에서 정서 기능을 측정

하는 소게임의 문항 수를 늘리고 반응 시간이 나 자극의 난이도를 조정하여 평가의 정밀도를 높일 필요가 있겠다. 이러한 개선이 이루어진다면, 정서 기능 평가의 신뢰성과 변별력이 더욱 강화되어 보다 정교한 정신 건강 평가 도구로서의 역할을 수행할 수 있을 것으로 기대된다.

앞서 언급한 바와 같이 CEMO-A와 표준화된 심리 검사 점수들 간에 유의미한 상관이나 나타났지만 상관계수의 크기는 전반적으로 낮거나 중간 정도였는데, 다음과 같은 요인이 영향을 미쳤을 것으로 여겨진다. 첫째, 본 연구는 정상 아동을 대상으로 진행되었기 때문에, 보호자 평정 척도인 CBCL을 비롯한 검사 결과가 하위 범위에 집중되는 경향이 있었다. 정신 건강 문제가 두드러지지 않는 집단에서 점수 범위가 협소해지면서 변산성이 줄어들어 상관계수가 낮아졌으며, 이로 인해 일부 극단값도 더 크게 영향을 미치게 된 것으로 보인다. 이를 보완하기 위해 Spearman 상관 분석을 추가로 수행한 결과 상관계수의 폭이 다소 증가했으나, 향후 임상군을 포함한 더 큰 표본 대상 연구를 통해 보다 광범위한 점수 분포에서 재평가할 필요가 있겠다. 두 번째, CEMO-A와 기존 검사들은 평가 방식의 차이가 존재하기에 직접적인 일대일 대응이 어려워 상관계수가 크게 나타나지 않을 수 있다. 특히 보호자 평정 척도는 보호자의 주관적 관찰에 의존하기 때문에, 아동의 미묘한 문제는 충분히 반영되기 어렵다. 기존 연구에서도 보호자 척도와 수행 기반 척도 간 상관계수가 낮게 나타나는 경향이 확인되었다(Magimairaj, 2018; Tan et al., 2018). 또한, 정형화된 평가 과제인 지능 검사와 신경심리 검사는 게임화된 방식의 디지털 평가 과제와 정확히 일치하지 않을 수 있다. 선행

연구에서도 게임화된 정도가 클수록 기존 과제와의 상관계수가 낮아지는 경향이 보고되었으며, 실제 행동 예측력은 오히려 독립적인 설명력을 가질 수 있음을 시사했다(Lee et al., 2023; Song et al., 2020).

본 연구는 CEMO-A의 타당성을 검증하는데 있어 의미 있는 결과를 제시했으나, 몇 가지 제한점이 존재하며 이를 보완하기 위한 추가적인 연구가 필요하다. 첫째, 본 연구의 표본 크기는 상관 분석을 수행하는 데에는 적절하지만, 연령별 비교를 위한 ANOVA 분석에서 검정력이 충분하지 않을 수 있다. G\*Power 분석 결과 200명의 참여자가 요구되었으나, 본 연구의 표본 수는 이에 미치지 못하였다. 또한, 모집 과정에서 7세 연령군의 성비가 균형을 이루지 못해, 연구 결과 해석에 일부 제한이 있을 수 있다. 따라서 향후 연구에서는 보다 충분한 표본을 확보하고 성별 균형을 고려하여 연구의 신뢰성을 높이는 한편, 연령별 차이를 보다 명확히 검증하여 CEMO-A의 발달적 타당성을 체계적으로 평가할 필요가 있다. 둘째, 본 연구는 정상 범주의 인지 기능을 가진 아동들을 대상으로 진행되었으며, 정신적 및 신경발달적 문제를 나타내는 임상 집단이 포함되지 않았다. 이러한 표본의 제한성으로 인해 CEMO-A의 평가 결과가 다양한 아동에게 일반화될 수 있는지 검토할 필요가 있다. ADHD, ASD, 우울 및 불안 장애 등 다양한 임상군을 연구 대상에 포함하여 연구를 진행한다면, CEMO-A가 특정 정서적, 인지적, 행동적 특성을 얼마나 정확하게 평가하고 선별할 수 있는지 보다 명확히 검증할 수 있을 것이다. 특히, CEMO-A는 일상적 맥락에서의 심리 상태 모니터링 및 조기 선별을 목적으로 설계된 도구이며, 본 연구는 일반 아동만을 대상

으로 하였다. 따라서 CEMO-A의 선별 측정치는 임상 장면에서의 선별 및 진단적 활용에 한계가 있으며, 임상 집단을 포함한 후속 연구를 통해 민감도와 특이도를 체계적으로 검증할 필요가 있다. 셋째, CEMO-A는 반복적 로그인을 통한 생태학적 평가가 가능하도록 설계되었으나, 본 연구에서는 연구 설계상 단회성 평가 방식으로 진행되었다. 향후 연구에서는 실제 환경과 유사한 조건에서 반복적 로그인 기능을 활용한 평가를 수행하여, CEMO-A의 실질적인 활용성과 강점을 보다 체계적으로 검증할 필요가 있다. 넷째, 본 연구는 CEMO-A의 평가 도구로서의 기본적인 타당성을 검증하는 데 초점을 맞춘 횡단 연구로, 예측 타당도는 검토하지 못하였다. 따라서 향후 연구에서 장기적인 종단 연구를 통해 CEMO-A 점수가 아동의 정신 건강 상태를 적절히 예측할 수 있는지 확인하는 과정이 필요하다. 이를 통해 CEMO-A가 조기 개입과 예방적 평가 도구로서 활용될 수 있는 가능성을 보다 명확하게 평가할 수 있을 것이다. 다섯째, 대규모 정상 집단을 대상으로 한 연구를 통해 연령별 규준을 마련할 필요가 있다. 아동의 인지 기능은 각 연령대 마다 발달적 차이가 크기 때문에, 연령에 따른 발달적 특성을 반영한 세밀한 규준이 요구된다. 이를 통해 CEMO-A가 연령별로 기대되는 인지 및 정서 기능 수준을 보다 정확히 평가할 수 있으며, 특정 연령대에 적합한 기준을 마련함으로써 평가 결과의 신뢰성과 타당성을 높일 수 있다. 이러한 규준이 마련되면, 각 연령대의 정상 범주와 이상치를 보다 명확히 구분할 수 있어, 조기 선별과 맞춤형 개입에 더욱 효과적인 도구로 활용될 수 있을 것이다.

본 연구는 CEMO-A가 아동의 인지 및 정서

기능을 평가하는 도구로서 기본적인 타당성을 갖추고 있으며, 특히 디지털 정신 건강 평가 도구로서의 가능성을 시사한다는 점에서 의의가 있다. CEMO-A는 고유의 방식을 적용한 기능성 게임형 디지털 평가로, 기존의 표준화된 심리 검사와 차별화된 평가 접근 방식이다. 이는 평가 과정에서 아동의 흥미와 참여도를 높이고, 보다 일상적인 환경에서도 쉽게 활용될 수 있는 생태학적 타당성을 갖춘 디지털 평가 도구로서의 가능성을 지닌다. 또한, CEMO-A는 인지와 정서 기능을 독립적으로 평가하면서도, 두 기능 간의 상호작용을 함께 고려할 수 있도록 설계된 평가 도구이다. 이러한 점은 기존 디지털 평가 도구와 차별화되는 특징으로, 아동의 다양한 인지 및 정서적 특성을 보다 정교하게 분석하고 통합적으로 평가하는 데 기여할 수 있다. 기존 표준화된 검사와의 수렴 및 변별 타당도 분석 결과, CEMO-A는 특정 정신 건강 상태 및 인지·정서적 특성을 탐지하는 평가 도구로 활용될 수 있음을 시사하였다. 다만, 연구 결과를 통해 볼 때 CEMO-A의 타당성과 신뢰도를 강화하기 위해 보완이 필요한 부분들이 확인되었으며, 향후 고도화 과정을 거칠 예정이다. 이를 통해 CEMO-A가 디지털 기반 조기 선별 도구로서 활용되어 기존 평가 도구를 보완하는 중요한 역할을 할 수 있을 것으로 기대된다.

## 참고문헌

- 곽금주, 장승민 (2019). K-WISC-V의 실시와 채점 지침서. 인싸이트.  
[https://inpsy.co.kr/psy/item/view/KWISC5\\_CO\\_TG](https://inpsy.co.kr/psy/item/view/KWISC5_CO_TG)

- 신민섭, 구훈정 (2007). 아동 색 선로검사. 학지사.  
[https://inpsyt.co.kr/psy/item/view/CCTT\\_CO\\_TG](https://inpsyt.co.kr/psy/item/view/CCTT_CO_TG)
- 신민섭, 박민주 (2006). 한국판 아동용 스트룹 색상·단어 검사의 표준화 연구 [포스터 발표]. 한국심리학회 연차 학술발표논문집, 2006(단일호), 628-629.
- 신민섭, 박민주 (2007). 스트룹 아동 색상-단어 검사. 학지사.  
[https://inpsyt.co.kr/psy/item/view/STROOP\\_CO\\_SG](https://inpsyt.co.kr/psy/item/view/STROOP_CO_SG)
- 오경자, 김영아 (2010). ASEBA 아동청소년 행동평가척도 매뉴얼. 휴노.  
[https://www.aseba.co.kr/info/05.aspx?scode=CBC\\_L6](https://www.aseba.co.kr/info/05.aspx?scode=CBC_L6)
- 유희정 (2008). 한국판 사회적 의사소통 설문지 (K-SCQ). 학지사.  
[https://inpsyt.co.kr/psy/item/view/SCQ\\_C\\_CO\\_PG](https://inpsyt.co.kr/psy/item/view/SCQ_C_CO_PG)
- Barbas, H., & García-Cabezas, M. Á. (2017). Prefrontal cortex integration of emotion and cognition. In M. Watanabe (Ed.), *The prefrontal cortex as an executive, emotional, and social brain* (pp. 51-73). Springer.  
[https://doi.org/10.1007/978-4-431-56508-6\\_4](https://doi.org/10.1007/978-4-431-56508-6_4)
- Canet Juric, L., Richards, M. M., Introzzi, I., Andrés, M. L., & Urquijo, S. (2013). Development patterns of executive functions in children. *Spanish Journal of Psychology*, 16, Article e41. <https://doi.org/10.1017/sjp.2013.44>
- Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology*, 33(2), 218-236.  
<https://doi.org/10.1111/bjdp.12075>
- Crivelli, D., Acconito, C., & Balconi, M. (2024). Emotional and cognitive “route” in decision-making process: The relationship between executive functions, psychophysiological correlates, decisional styles, and personality. *Brain Sciences*, 14(7), Article e734.  
<https://doi.org/10.3390/brainsci14070734>
- Egami, C., Yamashita, Y., Tada, Y., Anai, C., Mukasa, A., Yuge, K., Nagamitsu, S., & Matsuishi, T. (2015). Developmental trajectories for attention and working memory in healthy Japanese school-aged children. *Brain & Development*, 37(10), 840-848.  
<https://doi.org/10.1016/j.braindev.2015.02.003>
- El Kayed, J. E., Akl, T., Massoud, C., Margossian, C., Fayad, H., Fares, G., Felefly, T., & Rihana, S. (2024). Serious game for radiotherapy training. *BMC Medical Education*, 24(1), Article e463.  
<https://doi.org/10.1186/s12909-024-05430-1>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160.  
<https://doi.org/10.3758/BRM.41.4.1149>
- Golden, C. J. (1975). A group version of the Stroop Color and Word Test. *Journal of Personality Assessment*, 39(4), 386-388.  
[http://dx.doi.org/10.1207/s15327752jpa3904\\_10](http://dx.doi.org/10.1207/s15327752jpa3904_10)
- Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child*

- Psychology and Psychiatry*, 45(7), 1185-1198.  
<https://doi.org/10.1111/j.1469-7610.2004.00316.x>
- Jablonská, M., Fajnerová, I., & Nekovářová, T. (2024). Cognitive screening for children: Piloting a new battery of interactive games in 4-to 8-year-old children. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 18(4). Article e4. <https://doi.org/10.5817/CP2024-4-4>
- Klein, A. M., de Voogd, L., Wiers, R. W., & Salemink, E. (2018). Biases in attention and interpretation in adolescents with varying levels of anxiety and depression. *Cognition and Emotion*, 32(7), 1478-1486.  
<https://doi.org/10.1080/02699931.2017.1304359>
- Klenberg, L. (2015). *Assessment and development of executive functions in school-age children* [Unpublished doctoral dissertation, University of Helsinki, Institute of Behavioural Sciences]. University of Helsinki.  
<https://www.researchgate.net/publication/335146719>
- Lee, R. S. C., Albertella, L., Christensen, E., Suo, C., Segrave, R. A., Brydevall, M., Kirkham, R., Liu, C., Fontenelle, L. F., Chamberlain, S. R., Rotaru, K., & Yücel, M. (2023). A novel, expert-endorsed, neurocognitive digital assessment tool for addictive disorders: Development and validation study. *Journal of Medical Internet Research*, 25, Article e44414.  
<https://doi.org/10.2196/44414>
- Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games*, 4(2), Article e11.  
<https://doi.org/10.2196/games.5888>
- Magimairaj, B. M. (2018). Parent-rating vs performance-based working memory measures: Association with spoken language measures in school-age children. *Journal of Communication Disorders*, 76, 60-70.  
<https://doi.org/10.1016/j.jcomdis.2018.09.001>
- Maoz, K., Adler, A. B., Bliese, P. D., Sipos, M. L., Quartana, P. J., & Bar-Haim, Y. (2017). Attention and interpretation processes and trait anger experience, expression, and control. *Cognition and Emotion*, 31(7), 1453-1464.  
<https://doi.org/10.1080/02699931.2016.1231663>
- Marcu, G., Ondersma, S. J., Spiller, A. N., Broderick, B. M., Kadri, R., & Buis, L. R. (2022). The perceived benefits of digital interventions for behavioral health: Qualitative interview study. *Journal of Medical Internet Research*, 24(3), Article e34300.  
<https://doi.org/10.2196/34300>
- Martin-Key, N. A., Spadaro, B., Funnell, E., Barker, E. J., Schei, T. S., Tomasik, J., & Bahn, S. (2022). The current state and validity of digital assessment tools for psychiatry: Systematic review. *JMIR Mental Health*, 9(3), Article e32824.  
<https://doi.org/10.2196/32824>
- Moniz, M., de Jesus, S. N., Gonçalves, E., Viseu, J., Baptista, A. S., & Pacheco, A. (2016). Computerized Victoria Stroop Test in adult unipolar depressed patients and healthy subjects: Influence of age and gender. *Psychology*, 7(1), 1-5.  
<https://doi.org/10.4236/psych.2016.71001>

- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27-43.  
<https://doi.org/10.1007/s10803-005-0040-7>
- Park, S. E., Chung, J., Lee, J., Kim, M. J. B., Kim, J., Jeon, H. J., Kim, H., Woo, C., Kim, H., & Lee, S. A. (2024). Digital assessment of cognitive-affective biases related to mental health. *PLOS Digital Health*, 3(8), Article e0000595.  
<https://doi.org/10.1371/journal.pdig.0000595>
- Premkumar, P., Cooke, M. A., Fannon, D., Peters, E., Michel, T. M., Aasen, I., Murray, R. M., Kuipers, E., & Kumari, V. (2008). Misattribution bias of threat-related facial expressions is related to a longer duration of illness and poor executive function in schizophrenia and schizoaffective disorder. *European Psychiatry*, 23(1), 14-19.  
<https://doi.org/10.1016/j.eurpsy.2007.10.004>
- Quadt, L., Critchley, H., & Nagai, Y. (2022). Cognition, emotion, and the central autonomic network. *Autonomic Neuroscience: Basic & Clinical*, 238, Article e102948.  
<https://doi.org/10.1016/j.autneu.2022.102948>
- Raschle, N. M., Fehlbauer, L. V., Menks, W. M., Euler, F., Sterzer, P., & Stadler, C. (2017). Investigating the neural correlates of emotion-cognition interaction using an affective stroop task. *Frontiers in Psychology*, 8, Article e1489.  
<https://doi.org/10.3389/fpsyg.2017.01489>
- Riddell, C., Nikolic, M., Dusseldorp, E., & Kret, M. E. (2024). Age-related changes in emotion recognition across childhood: A meta-analytic review. *Psychological Bulletin*, 150(9), 1094-1117. <https://doi.org/10.1037/bul0000442>
- Rodas, J. A., Leon-Rojas, J., & Rooney, B. (2024). Mind over mood: Exploring the executive function's role in downregulation. *Frontiers in Psychology*, 15, Article e1322055.  
<https://doi.org/10.3389/fpsyg.2024.1322055>
- Rutter, M., Bailey, A., & Lord, C. (2003). The Social Communication Questionnaire (SCQ). Western Psychological Services.  
<https://www.wpspublish.com/scq-social-communication-questionnaire.html>
- Silvers, J. A., Buhle, J. T., & Ochsner, K. N. (2014). The neuroscience of emotion regulation: Basic mechanisms and their role in development, aging, and psychopathology. In K. N. Ochsner & S. M. Kosslyn (Eds.), *The Oxford handbook of cognitive neuroscience, Vol. 2. The cutting edges* (pp. 52-78). Oxford University Press.
- Smith, K. A., Blease, C., Faurholt-Jepsen, M., Firth, J., Van Daele, T., Moreno, C., Carlbring, P., Ebner-Priemer, U. W., Koutsouleris, N., Riper, H., Mouchabac, S., Torous, J., & Cipriani, A. (2023). Digital mental health: Challenges and next steps. *BMJ Mental Health*, 26(1), 238, Article e300670.  
<https://doi.org/10.1136/bmjment-2023-300670>
- Song, H., Yi, D. J., & Park, H. J. (2020). Validation of a mobile game-based assessment of cognitive control among children and

- adolescents. *PLOS one*, 15(3), Article e0230498.  
<https://doi.org/10.1371/journal.pone.0230498>
- Spadaro, B., Martin-Key, N. A., Funnell, E., Benáček, J., & Bahn, S. (2023). Opportunities for the implementation of a digital mental health assessment tool in the United Kingdom: Exploratory survey study. *JMIR Formative Research*, 7, Article e43271.  
<https://doi.org/10.2196/43271>
- Spear, L. P. (2013). Adolescent neurodevelopment. *Journal of Adolescent Health*, 52(2), S7-S13.  
<https://doi.org/10.1016/j.jadohealth.2012.05.006>
- Tan, A., Delgaty, L., Steward, K., & Bunner, M. (2018). Performance-based measures and behavioral ratings of executive function in diagnosing attention-deficit/hyperactivity disorder in children. *ADHD Attention Deficit and Hyperactivity Disorders*, 10(4), 309-316.  
<https://doi.org/10.1007/s12402-018-0256-y>
- Vermeir, J. F., White, M. J., Johnson, D., Crombez, G., & Van Ryckeghem, D. M. (2020). The effects of gamification on computerized cognitive training: Systematic review and meta-analysis. *JMIR Serious Games*, 8(3), Article e18644.  
<https://doi.org/10.2196/18644>
- Wechsler, D. (2014). WISC-V: Technical and interpretive manual. NCS Pearson, Incorporated.  
<https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/wisc-v>
- West, M. J., Angwin, A. J., Copland, D. A., Arnott, W. L., & Nelson, N. L. (2020). Cross-modal emotion recognition and autism-like traits in typically developing children. *Journal of Experimental Child Psychology*, 191, Article e104737.  
<https://doi.org/10.1016/j.jecp.2019.104737>
- Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, 5(1), 72-77.  
<https://doi.org/10.1177/1754073912451492>

원고접수일 : 2024. 12. 02.

수정원고접수일 : 2025. 03. 12.

게재확정일 : 2025. 03. 13.

## Development and Validation of the CEMO-A: A Tool for the Cognitive and Emotional Monitoring - Assessment of Children \*

Juri Park<sup>1)</sup>      Myeongji Im<sup>1)</sup>      Jaeun Kim<sup>2)</sup>  
Inyoung Son<sup>2)</sup>      Sumin Jang<sup>2)</sup>      Hyunjoo Song<sup>3)†</sup>

<sup>1)</sup>Department of Educational Psychology, Seoul Women's University, Ph.D. Student

<sup>2)</sup>Department of Educational Psychology, Seoul Women's University, M.A. Student

<sup>3)</sup>Department of Educational Psychology, Seoul Women's University, Professor

This study aims to validate the initial efficacy of CEMO-A, a digital monitoring tool designed for assessing children's mental health and early screening. CEMO-A integrates an ecologically valid assessment approach and serious gaming elements, enabling naturalistic and valid assessments in everyday environments. It comprehensively evaluates sustained attention, working memory, cognitive control, executive function, emotion sensitivity, and emotion recognition to detect major mental health disorders. The study involved 100 typically developing children aged 6 to 10 years. CEMO-A's performance was compared with standardized psychological assessments, including intelligence tests, neuropsychological tests, and caregiver-report questionnaires, to examine construct and concurrent validity. Results indicated that CEMO-A independently and integratively assessed cognitive and emotional functions, demonstrating construct validity, while comparisons with existing standardized tests confirmed concurrent validity. These findings suggest that CEMO-A is a valuable digital tool for efficiently evaluating children's mental health and holds potential for early screening and preventive intervention.

*Keywords* : CEMO-A, digital assessment, digital mental health tool, mental health screening, cognitive and emotional evaluation, gamification, ecological validity

---

\* This study was funded by the Ministry of Health and Welfare and supported by the Health and Medical Technology Research and Development Project of the Korea Health Industry Development Institute (Task Unique Number: HI22C0775).

† Corresponding Author: Song Hyunjoo / Professor, Department of Educational Psychology, Seoul Women's University, Professor / (10797) Room 613, 50th Anniversary Hall, 621 Hwarang-ro, Nowon-gu, Seoul, Korea / Tel: 02-970-5888 / Email: jayoo1004@swu.ac.kr