

산업현장 근로자의 인지 기능 및 정신건강 모니터링을 위한 디지털 플랫폼 연구

양철영*, 조은상**¹, 김상규**², 정성택***

A Study on a Digital Platform for Monitoring Cognitive Function and Mental Health of Industrial Workers

ChulYeong Yang*, Eunsang Jo**¹, Sang-Gyu Kim**², and Sungtaek Chung***

요 약

건설, 제조, 광업 등과 같은 산업현장 근로자들의 인지 및 심리적 상태를 평가하여 안전한 업무 수행을 지원하는 것이 매우 중요하다. 그러나 현재 상용화된 디지털 TBM(Tool Box Meeting) 시스템은 근로자의 인지 및 정신건강 상태를 직접 평가하는 데 한계가 있다. 이에 본 연구에서는 근로자의 인지 기능을 평가하기 위한 6종의 기능성 게임형 콘텐츠와 정신건강 상태를 측정하는 설문을 통합한 디지털 정신건강 모니터링 플랫폼을 제안하였다. 또한, 인지 기능 콘텐츠의 타당성을 검증하기 위해 12명의 자원자를 대상으로 음주 후 인지 기능 저하를 평가하였다. 이에 대한 대응 표본 t-검정 결과, 모든 인지 기능 과제에서 음주 후 수행 저하가 통계적으로 유의하게 나타났다. 결과적으로 제안된 인지 기능 콘텐츠는 실시간으로 근로자의 인지 상태를 평가하여, 산업현장에서의 안전관리 효율성을 향상시킬 수 있는 플랫폼 연구의 기초로 활용할 수 있을 것이다.

Abstract

In industrial environments such as construction, manufacturing, and mining, assessing cognitive and psychological states is imperative. However, existing Digital Toolbox Meeting (TBM) systems primarily focus on risk assessment and preliminary safety training. This study presents a platform integrating six game-based cognitive function assessment tasks and psychological surveys. To evaluate the efficacy of these cognitive function tasks, an experiment was conducted with twelve participants. The t-test results demonstrated statistically significant performance degradation in all tasks following alcohol consumption, with effect sizes measured by Cohen's d exceeding 0.6. As a result the proposed cognitive function content can serve as a foundation for platform development. This research enables real-time evaluation of workers' cognitive states, thereby enhancing safety management efficiency in industrial settings.

Keywords

cognitive function, serious game, industrial site, tool box meeting, TBM, DASS-21, large language model, LLM

* 한국공학대학교 컴퓨터공학과 박사과정

- ORCID: <https://orcid.org/0009-0009-2325-3194>

** 한국공학대학교 바이오헬스융합공학과 석사과정

- ORCID¹: <https://orcid.org/0009-0008-3445-7785>

- ORCID²: <https://orcid.org/0009-0004-6574-6601>

*** 한국공학대학교 컴퓨터공학과(교신저자)

- ORCID: <http://orcid.org/0000-0002-8692-0179>

• Received: Oct. 22, 2025, Revised: Nov. 07, 2025, Accepted: Nov. 10, 2025

• Corresponding Author: Sungtaek Chung

Dept. of Computer Engineering, Tech University of Korea, 237,

Sangidachak-ro, Siheung-si, Gyeonggi-do, Republic of Korea

Tel.: +82-32-8041-0527, Email: unitaek@tukorea.ac.kr

1. 서 론

고용노동부의 2024년 산업재해 통계에 따르면, 전체 산업재해자 수는 142,771명으로 전년 대비 4.4% 증가하였으며 사망자 수도 2,098명으로 4.1% 상승하였다. 이 중 사고 사망자가 39.4%를 차지하였고, 업종별로는 건설업, 제조업, 기타 서비스업, 운수·창고업, 광업 순으로 높은 비율을 보였다[1]. 사고 유형을 살펴보면 낙상 및 충돌 사고는 감소했지만, 물체에 맞음, 끼임, 깔림, 화재·폭발 사고 등은 증가하는 추세를 보인다. 이러한 통계는 정부와 기업이 다양한 안전정책과 관리·감독 강화를 지속해서 시행하고 있음에도 불구하고, 산업현장에서의 재해 발생이 심각한 사회적 문제로 남아 있음을 보여준다.

특히, 건설업, 제조업, 광업 등과 같은 현장 근로자들은 업무 특성상 상시로 다양한 위험에 노출되어 있다. 이들 산업은 중장비 운용, 고소작업, 화기 및 폭발물 취급, 장시간 교대근무, 협소한 공간 작업 등과 같이 재해 발생 가능성이 큰 작업 요소를 포함하고 있으며 이에 따라 단순한 개인의 실수나 장비 결함조차도 치명적인 사고로 이어질 수 있다[2]. 나아가, 산업재해의 주요 원인 중 상당 부분은 근로자의 부주의, 상황 인지 부족, 안전 절차 미준수 등과 같은 인적 요인에서 비롯되는 것으로 알려져 있다.

최근 국내에서는 산업안전보건공단(KOSHA)을 중심으로 작업 전 안전점검회의(TBM, Tool Box Meeting) 제도의 보급이 확대되고 있다[3]. 디지털 TBM은 근로자가 당일 수행할 작업에 내재된 잠재적 위험을 사전에 인지하고, 이에 따른 예방조치를 숙지하도록 지원함으로써 산업재해 예방에 중요한 역할을 한다. 특히, 전통적인 구두 중심의 안전점검회의와 비교할 때, 자료 표준화, 접근성 향상, 기록 관리의 효율성 측면에서 뚜렷한 장점이 있다.

그러나 현재 상용화된 디지털 TBM의 기능은 주로 출결 관리, 전자서명, 안전교육 이수 여부 확인, 위험성 평가 자료 공유 등 행정·기록 중심의 영역에 국한되어 있다. 이러한 한계로 인해 개별 근로자의 인지 기능, 주의 집중 상태, 심리적 스트레스 수준과 같이 실제 현장 안전 수행 능력에 직접 영향을 미치는 요인을 충분히 반영하지 못한다. 다시 말

해, 디지털 TBM은 안전교육과 관리 효율성을 높이는 데 기여하지만, 근로자의 인지·정신적 상태를 기반으로 위험 대응 역량을 평가·관리하는 측면에서는 미흡한 실정이다.

일반적으로 산업현장은 작업환경 변화가 빈번하고, 긴급 상황 발생 시 신속하고 정확한 대응이 더욱 요구된다. 이러한 상황에서 근로자의 안전 수행 능력은 단순한 기술 숙련도에 그치지 않고, 주의 전환, 위험 인지, 의사결정 등과 같은 고차원적 인지 기능에 크게 의존한다. 따라서 산업재해의 주요 원인 중 하나로 근로자의 인지·심리적 요인을 체계적으로 평가할 필요성이 제기되고 있다. 건설 근로자의 인지적 실패(Cognitive failures)가 주의 전환 오류, 지각 오류, 판단 오류 등을 초래하여 위험 행동을 유발하고 사고 발생 위험을 높이는 주요 메커니즘으로 작용한다고 보고하였다[4].

또한, 작업자가 위험 상황에서 시선 분산과 상황 인지 저하를 경험할 때 사고 취약성이 증가한다는 실험적 결과를 제시하였다. 더불어 국제노동기구(ILO, International Labour Organization)와 미국 국립 직업안전보건연구원(NIOSH)은 건설 산업재해의 상당수가 물리적 안전장치의 부재보다 작업자의 인적 요인(Human factor)에 기인한다고 강조한 바 있다[5][6]. 이러한 인적 요인에는 주의력 저하, 위험 인지 부족, 과도한 자신감, 피로 누적, 스트레스, 감정 상태 등 인지·심리적 특성이 포함된다. 특히, 피로 누적, 불안·우울과 같은 심리적 불안정성과 인지 기능 저하 간의 상관성은 최근 여러 연구에서 실증적으로 확인되고 있다. 따라서 고위험 산업현장에서 근로자의 상황 인지 저하 여부를 조기에 파악하기 위해서는 인지 기능과 심리 상태에 대한 정량적 평가 체계 구축이 필요하다.

그리고 교대근무, 직무 스트레스, 장시간 노동으로 인한 근로자의 인지 기능이 업무를 수행하는 데 있어서 유의한 부정적 영향을 미치며 이는 작업 효율성 저하와 안전사고 위험 증가로 직결될 수 있다고 제시하였다[7]. 또한, 직무 스트레스 수준이 높은 근로자일수록 오류 발생 빈도와 안전 수칙 위반율이 높고, 산업재해에 연루될 가능성이 유의하게 증가한다는 실증적 결과를 보고하였다[8]. 이와 같은 결과들은 작업 시작 전에 근로자의 심리적 안정성

과 인지적 준비 상태를 종합적으로 점검하고, 그 결과를 현장 투입 여부 결정 과정에 반영할 필요성을 강하게 시사하는 것이다. 이는 물리적 안전장치와 관리·감독 강화만으로는 줄이기 어려운 인적 요인(Human factor) 기반의 사고를 예방하기 위한 핵심 전략이 될 수 있다.

최근 인지 기능 평가 도구와 심리 상태 측정 도구의 디지털화, 뇌 신경 신호 분석(fNIRS, EEG)을 통한 인지 부하 측정, 정신건강 모니터링, 디지털 콘텐츠 기반 훈련과 같은 접근법들이 주목받고 있다. 예를 들어, 안드로이드 기반 애플리케이션 UX-TMT(User eXperience-Trail Making Test)를 개발하여 기존 종이 기반 TMT를 대체할 수 있는 디지털 도구로서의 신뢰도와 타당성을 검증하였다. UX-TMT를 사용한 경도인지장애 및 치매 집단과 정상 집단을 높은 민감도(0.83)와 특이도(0.92)로 구분할 수 있었으며, MMSE 점수와도 유의한 상관($r = 0.77$)을 보여 디지털 인지평가 도구의 유용성을 입증하였다[9].

또한, EEG, ECG, PPG 등 웨어러블 기반 생체신호를 활용한 스트레스 및 주의·집중 수준 측정 연구를 체계적으로 검토하며 직장인의 정신건강과 집중력 모니터링의 중요성을 강조하였다[10]. 더 나아가 최근에는 디지털 콘텐츠 기반 인지 재활 훈련이 활발히 진행되며 임상 및 일상적 환경에서 인지 기능 향상과 재활 참여도를 높이는 효과적인 접근법으로 주목받고 있다.

이러한 연구 동향을 바탕으로, 본 연구에서는 작업 투입 전 근로자의 인지·심리 상태를 정량적으로 평가하고, 그 결과를 현장 투입 여부 결정에 반영할 수 있는 디지털 TBM 플랫폼을 제안한다. 구체적으로 근로자의 심리 상태를 모바일에서 간편하게 응답할 수 있도록 DASS-21 설문지를 전산화하였으며, 산업현장에서 요구되는 핵심 인지 기능(주의·집중력, 작업 기억력, 시각 반응속도, 집행기능)을 간이 진단할 수 있는 6종의 기능성 게임 콘텐츠를 구현하였다. 나아가 심리 상태와 인지 기능 진단 결과로부터 도출된 다차원 데이터를 통합하여, 대규모 언어모델(LLM) 기반 현장 투입 여부 판단 모듈에 입력함으로써 근로자의 작업 수행 적합성을 종합적으로 분석하고, 관리자에게 객관적이고 표준화된 의사

결정 근거를 제공하고자 하였다.

이와 같은 접근 방식은 기존 TBM의 한계를 보완함과 동시에, 작업 전 근로자 상태를 실시간으로 반영하는 맞춤형 안전관리 체계를 구현할 수 있는 기반을 마련한다는 점에서 의의가 있다.

II. 산업현장 근로자 관리 플랫폼

2.1 시스템 개요 및 아키텍처

본 연구에서 제안하는 산업현장 근로자 관리 플랫폼은 그림 1과 같이 관리자 페이지, 근로자 페이지, 인지 기능 검사 콘텐츠, 서버 및 데이터베이스, 그리고 LLM 기반 종합 소견 생성 모듈로 구성된다.

관리자 페이지는 근로자의 인지 및 정신건강 상태를 효과적으로 모니터링하고, 안전관리 업무를 체계적으로 수행할 수 있도록 설계하였다. 관리자는 플랫폼을 통해 개별 근로자의 인지 기능 검사 결과와 DASS-21 심리검사 결과를 확인할 수 있으며, 이를 기반으로 위험군 근로자에 대한 실시간 모니터링과 조기 개입이 가능하다. 또한, 출결 현황, 작업 일정, 교육 이수 여부를 종합적으로 관리할 수 있으며, 필요하면 공지사항을 게시하여 현장 전체에 신속히 전달할 수 있다.

근로자 페이지는 인지 기능 검사 콘텐츠와 DASS-21 심리검사를 수행할 수 있는 기능을 제공하며, 안전교육 콘텐츠 이수 및 LLM 기반 리포트 조치가 가능하다. 더불어 작업 시작 전 TBM 서명, 위험 작업 허가서 작성, 안전교육 이수 확인 등을 전자서명 방식으로 기록할 수 있어, 기존 문서 기반의 형식적 보고에서 벗어나 효율적인 관리가 가능하다. 이 외에도 공지사항 확인, 회사 일정 조회, 출근 관리 등 현장에서 요구되는 다양한 기능을 지원한다.

인지 기능 검사 콘텐츠와 근로자 페이지의 안전교육 콘텐츠는 Unity 기반으로 제작되었으며, WebGL로 빌드되어 웹 애플리케이션 환경에서 실행 가능하도록 구현되었다. 모든 콘텐츠 수행 결과와 설문 데이터는 중앙 데이터베이스에 저장되며, 안전한 데이터 관리 및 접근 제어를 통해 개인정보가 보호된다. 저장된 데이터는 LLM 종합 소견 생성 모듈의 입력값으로 활용된다.

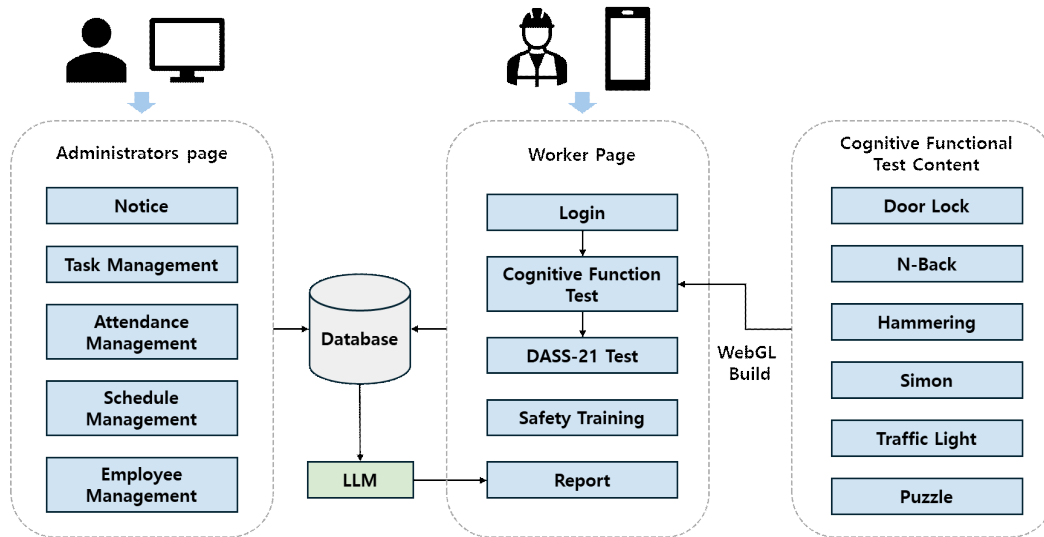


그림 1. 제안한 산업현장 근로자 관리 플랫폼에 대한 시스템 구성도

Fig. 1. System architecture of the proposed industrial site worker management platform

LLM 종합 소견 생성 모듈은 Python(3.11.8) 기반으로 개발되었으며, 근로자의 인지적·심리적 상태를 종합적으로 분석하여 관리자에게 맞춤형 소견 리포트를 제공한다. 이를 통해 관리자는 단순한 보고 관리 수준을 넘어, 데이터 기반의 의사결정을 수행함으로써, 산업현장의 안전성과 효율성을 동시에 향상시킬 수 있다.

2.2 인지 기능 검사 콘텐츠

산업현장에서는 근로자가 다양한 작업 상황과 위험 요소에 적절히 대응하기 위해 주의력(Attention), 집중력(Concentration), 작업 기억력(Working memory), 반응 속도(Reaction time) 등이 필수적으로 요구된다[11].

주의력은 주어진 환경에서 필요한 자극을 선택하고 불필요한 자극을 배제하여 효율적으로 정보처리를 수행하는 핵심적인 능력이다. 특히 고위험 작업 환경의 산업현장에서는 경고음, 시각표지, 동료의 신호 등 안전과 직결되는 자극을 정확히 식별하고 즉각적으로 반응하는 것이 필수적이다. 만약 주의력이 저하될 경우 기계 작동 중 경고음을 놓치거나 안전 절차의 핵심 단계를 누락하여 추락·끼임·충돌과 같은 중대 재해로 이어질 수 있다[12]. 이는 단순한 작업 실수 차원을 넘어 개인의 생명과 직결되고, 작업 전체의 안전 시스템을 위협하는 결과를 초래한다. 따라서 산업현장에서는 작업자 개인의 주의

력 평가 및 향상 훈련, 경고 신호의 명확성과 중복성 확보(청각·시각 병행), 주의 분산을 최소화하는 작업환경 설계가 필수적이다. 이와 같은 다층적 접근을 통해 작업자의 주의력을 유지·보강해야만 산업재해를 예방하고, 효율적인 작업 수행과 안전한 근로 환경을 동시에 확보할 수 있다.

집중력은 주의력을 장시간 유지하며 과업 수행에 인지 자원을 지속해서 투입하는 능력을 의미한다. 산업현장에서는 단조로운 반복 작업이나 장시간 교대근무와 같은 환경에서도 일정한 작업 품질을 유지하고, 긴급 상황에서도 흔들리지 않는 안정적인 수행 능력이 요구된다. 집중력이 저하되면 작업 효율이 떨어질 뿐만 아니라 돌발 위험에 대한 반응속도가 늦어져 사고 위험이 크게 높아진다. 특히 피로와 스트레스는 이러한 집중력 저하를 가속하는 주요 요인으로 작용한다[13].

작업 기억력은 단기적으로 정보를 저장하고 이를 과업 수행에 활용하는 능력이다. 산업현장에서는 안전 절차, 장비 조작법, 위험구역 위치, 공정 순서 등 다양한 정보를 동시에 기억하고 적시에 적용해야 하므로 작업 기억력이 특히 중요하다. 그러나 작업 기억 용량이 부족하거나 인지 부하가 과도할 경우, 절차 누락, 장비 오작동, 위험구역 진입 등과 같은 치명적인 오류로 이어질 수 있으며, 복합 작업(예: 다중 장비 운용, 다단계 절차 수행)에서는 안전성과 직결된다[14].

반응속도는 자극이 제시된 순간부터 이에 대한 반응이 나타나기까지 걸린 시간으로, 인지 시스템의 정보처리 속도와 효율을 평가하는 핵심 지표로 사용된다. 반응속도는 하나의 자극에 대해 즉각적으로 반응하는 단순 반응 속도와 여러 자극 중에서 적절한 반응을 선택해야 하는 복합 반응속도가 있다. 산업현장과 같은 고위험 환경에서는 경고음, 시각 신호, 동료의 지시 등 안전과 직결되는 자극에 신속하게 반응하는 능력이 필수적이다. 만약 반응속도가 지연될 경우, 위험 신호를 인지했다라도 실제 회피 행동이 늦어져 추락·끼임·충돌과 같은 중대 재해로 이어질 수 있다[15].

이에 본 연구에서는 산업현장 근로자의 인지 기능 평가의 대표적인 작업 기억력과 주의·집중력, 반응속도를 측정하는 대표적인 평가 콘텐츠 6종은 TMT-A, N-back Task, Simon Span Task, Visual Reaction Test, Digit Span Test, Go/No-Go Task 프로토콜을 기반으로 구현하였다.

TMT는 신경심리학적 평가에서 가장 표준화된 검사로서, 다양한 신경학적 장애를 판별하는 데 널리 사용되고 있다. 이것은 두 가지 하위 검사(TMT-A, TMT-B)로 구성되었으며, TMT-A는 피검자가 원 안에 표시된 숫자(1, 2, 3 ... 25)를 가능한 한 빠르게 순서대로 연결해야 하는 과제로, 주로 시각적 탐색 능력, 처리 속도, 주의 집중력, 시각-운동 협응 능력을 평가한다. TMT-B는 숫자와 알파벳이 혼합된 원을 교대로 연결하는 과제로, 집행기능 및 인지적 전환 능력을 추가로 측정한다[16].

본 연구에서는 TMT-A 프로토콜을 산업현장의 못 타정 작업과 유사한 형태로 구현하였다. 그림 2(a)는 기존 TMT-A의 구조를 개념적으로 나타낸 것

으로, 피검자가 1부터 25까지 숫자를 순서대로 연결하는 기본 절차를 보여준다. 그림 2(b)는 구현한 콘텐츠 화면으로, 1~25까지 숫자가 무작위로 배치된 못 형태의 오브젝트가 제시된다. 제시된 숫자를 올바른 순서로 클릭하면 숫자가 사라지며 해당 못이 박히고, 잘못된 순서에서는 반응하지 않도록 설계되었다. 모든 못을 순서대로 박는 데 소요된 시간을 수행 지표로 기록하며, 39초 이내는 정상 범위, 40초 이상은 인지 기능 손상의 가능성이 있는 것으로 해석한다[17].

Simon span task는 작업 기억력을 평가하기 위해 널리 사용되는 것으로, 시각적으로 제시되는 자극의 순서를 기억하고 동일하게 재현하는 방식으로 진행된다. 이것은 순차적 정보의 저장 및 갱신 능력, 주의 집중, 단기 기억 처리 속도 등을 측정하는 데 효과적이다. Simon span task의 일반적인 절차는 그림 3(a)와 같이 네 가지 색상 버튼(녹색, 빨간색, 노란색, 파란색)을 기반으로 시작되며, 최초에는 한 가지 색상이 1,000ms 동안 제시되며, 사용자가 이를 정확히 재현할 때마다 제시되는 색상 시퀀스의 길이가 점차 증가한다. 주어진 색상의 순서를 정확히 기억하고 동일한 순서대로 색상을 재현해야 하는 것이다[18].

본 연구에서는 산업현장의 차량 신호대기 상황을 모사한 UI를 적용하여 simon span task 프로토콜을 구현하였다. 그림 3(b)의 화면 상단의 신호등은 무작위 순서로 다양한 색상(녹색, 빨간색, 노란색, 파란색)을 제시하며 각 색상은 1,000ms 동안 점등된다. 피검자는 이를 기억한 후 동일한 순서대로 하단의 색상 버튼을 동일한 순서로 터치해야 하며, 입력된 버튼은 300ms 동안 점등되어 시각적 피드백을 제공한다.



그림 2. 선로 잇기 검사 및 Hammering 콘텐츠 UI
Fig. 2. Trail making test and hammering content user interface (UI)

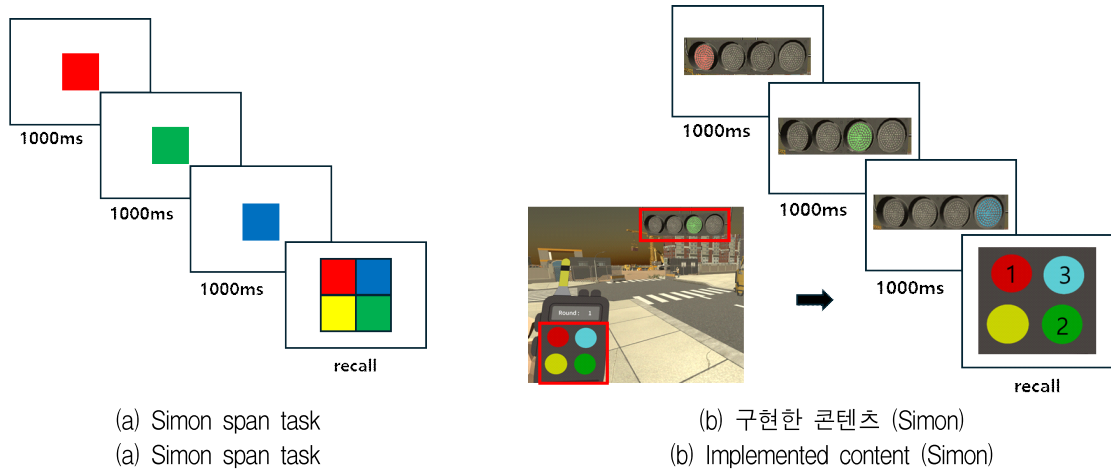


그림 3. Simon span task 및 simon 콘텐츠 UI
Fig. 3. Simon span task and simon content user interface(UI)

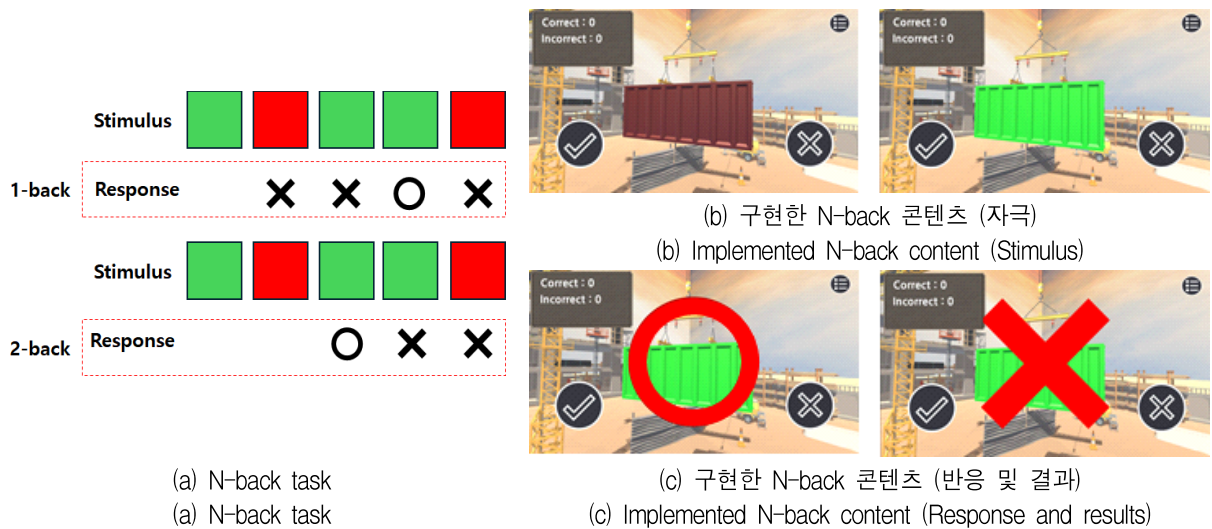


그림 4 N-back task 및 N-back 콘텐츠 UI
Fig. 4. N-back Task and N-back content user interface (UI)

과제는 점차 길어지는 시퀀스를 올바르게 재현하는 방식으로 진행되며, 반응시간과 정반응/오반응 횟수가 수행 지표로 기록된다. 정상 성인의 경우 평균적으로 5~9개의 자극 시퀀스를 정확히 재현할 수 있으며, 4개 이하일 경우 작업 기억 저하 가능성이 있는 것으로 해석한다[19].

N-back task는 작업 기억력 측정 및 훈련, 그리고 인지 부하 특성 분석에 널리 활용되는 대표적 인지 과제이다. 특히 공간 추론 과제인 RAPM(Raven's Advanced Progressive Matrices)와 높은 상관관계를 보이는 것으로 보고되어, 인지 기능 연구 방법으로 주목받고 있다[20]. 본 연구에서는 그림 4와 같이 산업현장의 컨테이너를 시각 자극으로 활용하여

N-back task를 구현하였다. 그림 4(a)는 1-back과 2-back 조건의 기본 구조를 도식화한 것으로, 자극이 차례로 제시되고 피검자가 이전 자극과의 일치 여부를 판단하는 N-back task의 원리를 설명한다. 그림 4(b)는 실제 구현된 콘텐츠에서 제시되는 시각 자극을 보여주며, 건설 현장을 배경으로 크레인에 매달린 컨테이너의 색상이 변화하는 형태로 구성된다. (c)는 피검자가 자극에 대해 반응하는 과정을 나타낸 것으로, 이전 자극과 색상이 일치하면 화면 좌측의 '✓' 버튼을, 불일치하면 우측의 '✗' 버튼을 눌러 응답하도록 설계되었다. 각 반응에 대해 정·오 반응 수를 실시간으로 기록되며, 색상 자극이 제시되는 시점부터 반응 입력까지의 시간이 수행 지표

로 수집된다. N-back task에 관한 선행 연구 결과, 1-back 조건에서는 평균 정확도 0.99, 반응시간 507 ± 176 ms($r = 0.93$)로 보고 되었으며, 2-back 조건에서는 평균 정확도 0.91, 반응시간 547 ± 183 ms($r = 0.85$)로 나타났다. 본 연구에서는 피검자의 수행 결과가 이러한 평균치 이하일 경우 작업 인지 능력 저하 가능성이 있는 것으로 해석하였다[21].

시각 반응 검사(Visual reaction test)는 시각 자극을 인지하여 이를 운동 반응으로 전환하기까지의 시간을 측정하는 검사로, 이 과정에는 높은 수준의 집중력이 요구된다[22]. 반응시간은 개인의 주의력 수준을 반영하며 특히 경계력이 필수적인 직종에서는 짧을수록 바람직하다. 대표적인 시각 반응 검사로는 신호등 검사 프로토콜이 있다. 해당 검사는 한국교통안전공단에서 65세 이상 사업용 차량 운전자를 대상으로 시행하고 있으며 운전자가 신호등 변화 및 돌발 자극에 얼마나 신속하게 반응하는지를 평가한다. 그림 5(a)는 신호등 검사 프로토콜을 도식화한 것으로, 피검자는 가속 페달과 제동 페달을 이용하여 빨간 신호가 제시될 경우 제동 페달을 밟고 초록 신호가 제시되면 가속 페달을 밟는 방식으

로 수행된다.

본 연구에서는 이러한 신호등 검사 프로토콜을 산업현장 상황에 적합한 UI로 변형하여 그림 5(b)와 같이 구현하였다. 화면 상단의 '신호등' 형태의 표적 자극이 무작위로 점멸되며 사용자는 제시된 색상과 같은 색상의 버튼을 즉시 터치해야 한다. 시스템은 반응시간, 정반응 수, 오반응 수를 실시간으로 기록하여 주의·집중력 수준을 정량적으로 평가할 수 있도록 구성하였다. 선행 연구에 따르면, 50세 성인의 정상 반응시간은 남성 344 ± 44 ms, 여성 377 ± 37 ms로 보고되었으며, 본 연구에서는 이 기준보다 느린 반응일 때 주의·집중력 저하로 해석하였다[23].

숫자 외우기 검사(Digit span test)는 주의력, 집중력, 작업 기억을 평가하기 위해 임상과 연구에서 가장 널리 활용되는 고전적 신경 심리검사 중 하나이다. 이 검사는 주어진 숫자열을 단기적으로 기억하고 이를 동일하거나 역순으로 재생하는 과제로 작업 기억의 용량과 유지 능력을 평가할 수 있다. Digit Span Test는 정방향(FDS, Forward Digit Span)과 역방향(BDS, Backward Digit Span)으로 구분된다.



그림 5 신호등 검사 프로토콜 및 Traffic Light 콘텐츠 UI
Fig. 5. Traffic light test protocol and traffic light content user interface (UI)

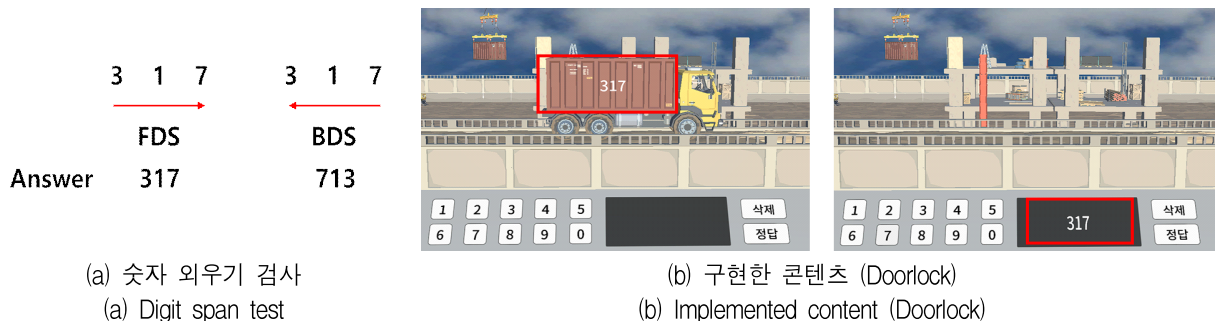


그림 6 숫자 외우기 검사 프로토콜 및 Doorlock 콘텐츠 UI
Fig. 6. Digit span test protocol and doorlock user interface (UI)

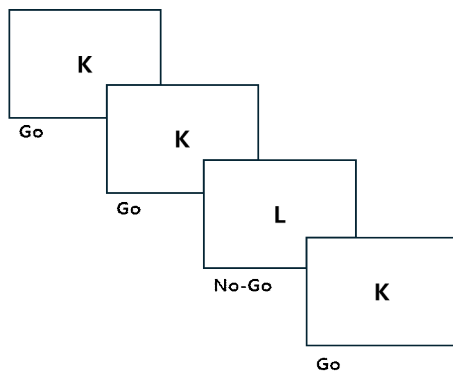
그림 6(a)는 digit span test를 도식화한 것으로, FDS에서는 제시된 숫자열을 그대로 답해야 하며 BDS에서는 숫자열을 역순으로 답해야 한다. 일반적으로 숫자는 1,000ms 간격으로 차례로 제시되며 참가자는 최종 제시 후 주어진 입력창을 통해 해당 숫자열을 재현한다. Span 길이는 2자리에서 시작하여 정답 시 1자리씩 증가하며 동일 길이에서 연속 두 차례 오답을 보이면 과제가 종료된다[24].

본 연구에서는 digit span test 프로토콜을 활용하여 산업현장을 배경으로 한 콘텐츠를 그림 6(b)와 같이 구현하였다. 과제 화물 운송 차량에 제시된 숫자를 기억한 후, 피검자가 화면 하단의 숫자 버튼을 통해 같은 순서로 입력하는 방식이다. 검사는 3자리 숫자열에서 시작하여 정답 입력 시 span 길이가 1자리씩 증가하며, 최대 15자리까지 확장된다. 즉, 피검자는 점차 길어지는 숫자열을 정확히 기억하고 재현해야 하며, 이를 통해 작업 기억 용량과 정보 유지 능력을 평가할 수 있다. Digit span test에 관한 선행 연구 결과는 50세 기준 정상 수행 개수를 남성 6.8 ± 1.8 개, 여성 5.9 ± 1.4 개로 보고되었으며, 본 연구에서는 해당 기준보다 적은 개수를 주의·집중력 저하로 해석하였다[23].

Go/No-Go 과제는 반응 억제 능력과 주의 지속력을 평가하는 대표적인 인지 검사 도구이다. 그림 7(a)는 Go/No-Go Task의 수행 절차를 나타낸 것이다. 과제 수행 전 참가자에게 ‘Go’와 ‘No-Go’ 자극의

대응 규칙(예: 문자 K일 때는 반응, 문자 L일 때는 억제)이 안내되며 이 규칙은 과제 전반에 걸쳐 유지된다. 피검자는 연속적으로 제시되는 자극 중 Go 자극에는 반응하고, 드물게 나타나는 No-Go 자극에는 반응을 억제해야 한다. 수행 결과는 정반응, 오반응(No-Go 자극에서 반응한 경우), 누락(Go 자극에서 반응하지 않은 경우)으로 구분하여 기록된다. 억제 통제의 핵심 지표는 No-Go 시행에서 반응을 잘못 실행할 확률이며 이를 통해 억제 통제와 선택적 주의 수준을 정량적으로 평가할 수 있다[25]. 본 연구에서는 Go/No-Go 과제를 기반으로 두더지 잡기 게임을 모사한 콘텐츠를 그림 7 (b)와 같이 구현하였다. 과제 시작 전 특정 색상의 드럼통을 표적 자극으로 제시하고, 이후에는 표적 자극을 포함한 3가지 색상의 드럼통이 무작위로 등장하도록 구성하였다. 플레이 시간은 60초로 설정하였으며, 피검자가 표적 자극에 반응하면 20점을 부여한다. 또한, 반응 특성에 따라 다음 표적의 제시 간격을 조정하였는데, 정반응 시에는 즉시 다음 표적을 제시하고, 놓치거나(미반응) 오반응 시에는 2초 후에 새로운 자극이 나타나도록 설계하였다[26].

Go/No-Go 과제에 관한 선행 연구 결과는 표적 자극에 대해 1.13초 이내에 반응한 경우를 정상 억제 반응으로 보고되었으며, 본 연구에서는 해당 기준을 표적 자극이 제시되는 시간으로 선정하였다[27].



(a) Go/No-Go task
(a) Go/No-Go task



(b) 구현한 콘텐츠 (Mole)
(b) Implemented content (Mole)

그림 7. Go/No-Go task 및 mole 콘텐츠 UI

Fig. 7. Go/No-Go task and mole content user interface UI

2.3 심리 상태 평가 모듈

심리적 상태를 조기에 선별하거나 평가하기 위해 다양한 자기 보고식 설문 도구가 활용되고 있다. 대표적으로 우울증의 유형과 심각도를 측정하는 BDI(Beck Depression Inventory), 불안 수준을 평가하는 HAM-A(Hamilton Anxiety Scale), 지난 한 달간 개인이 경험한 전반적인 스트레스 지각 수준을 측정하는 PSS(Perceived Stress Scale), 그리고 우울·불안·스트레스의 세 가지 심리 상태를 구분하여 평가하는 DASS(Depression Anxiety Stress Scales)가 널리 사용된다.

본 연구에서는 우울, 불안, 스트레스 각각의 중증도를 정량적으로 산출할 수 있는 DASS를 평가 도구로 선택하였다. 특히, 원형인 DASS-42를 절반으로 축약한 DASS-21은 설문 시간이 짧고 무료라는 장점이 있어, 시간 제약이 있는 실험 환경이나 현장 기반 연구에서 널리 활용된다. DASS-21은 우울, 불안, 스트레스 영역별로 각각 7문항씩 총 21문항으로 구성되며, 각 문항은 지난 1주일간 경험한 증상의 빈도와 강도를 4점 척도로 평가한다[28].

우울 척도는 낮은 긍정적 정서, 흥미 상실, 무가치감 등을 포함하며, 불안 척도는 심박수 증가, 긴장감, 공포와 같은 생리적·정서적 반응을 측정한다. 또한, 스트레스 척도는 지속적인 긴장, 과민성, 부정적 감정 상태를 평가한다. 이러한 점수 체계를 통해 응답자의 심리 상태를 정량적으로 표현할 수 있으며, 연구자는 척도별 점수를 활용하여 우울·불안·스트레스 수준을 명확히 구분하고 통계 분석에 적용할 수 있다.

DASS-21은 한국 근로자 집단에 관한 선행 연구 결과에서 내적 일관성(우울/불안/스트레스 $\alpha = 0.87/0.83/0.83$)과 1주 간격 재검사 신뢰도($ICC[2,1]=0.84/0.94/0.89$)가 모두 우수하게 나타났으며, 삶의 만족도(SWLS)와도 유의한 상관관계를 보였다. 또한, DASS-21을 활용하여 우울, 불안, 스트레스의 중증도에 따른 맞춤형 음악 재생목록을 제공하는 등 여러 산업현장과 플랫폼 연구에서 활용 중이다[29][30]. 이러한 결과를 바탕으로 DASS-21에 대한 검증없이 본 연구의 플랫폼에 적용하였다.

2.4 LLM 기반 종합 판단 모듈

본 연구에서는 DASS-21 심리 평가 결과와 인지 기능 검사 콘텐츠의 수행 결과를 통합적으로 해석하여 근로자의 상태를 평가하기 위해 LLM 기반 소견 생성 모듈을 개발하였다. 해당 모듈은 OpenAI의 GPT-4.1 모델을 활용하였으며, Zero-shot 방식의 프롬프트 엔지니어링 기법을 적용하였다.

DASS-21은 총 21문항으로 구성되며, 우울(Depression), 불안(Anxiety), 스트레스(Stress)의 세 하위 척도로 나뉜다. 각 척도는 7문항으로 이루어져 있으며, 문항은 0~3점으로 채점된다. 하위 척도 점수는 최대 21점, 전체 합산 점수는 최대 63점이다. 산출된 점수는 정상(Normal), 경도(Mild), 중등도(Moderate), 심한(Severe), 매우 심한(Extremely severe)의 다섯 수준으로 분류된다.

우울 척도는 무가치감, 무쾌감, 무기력 등의 정서적 용인을 평가하고, 불안 척도는 자율신경 항진과 공황·공포 반응을 측정하며, 스트레스 척도는 긴장, 과민, 만성적 각성 상태를 평가한다[31].

인지 기능 평가는 여섯 가지 과제(TMT-A, Simon Span Task, N-back, Digit Span, Visual Reaction Test, Go/No-Go)를 기반으로 진행되며, 과제별로 사전에 설정된 Cut-off 기준을 적용하여 정상 범위와 비정상 범위를 구분하였다. 예를 들어, 반응시간·정확도·정오반응 비율 등의 지표가 Cut-off를 초과하거나 미달할 경우, 해당 인지 영역에서 저하로 판정한다.

그림 8은 본 연구에서 개발한 LLM 기반 소견 생성 모듈의 전체 흐름을 도식화한 것이다. 본 모듈은 심리 평가 결과(DASS-21)와 인지 기능 검사 결과를 입력으로 받아, 근로자의 종합적인 정신·인지 상태를 해석하고 현장 투입 가능성에 대한 소견을 자동으로 생성하도록 설계되었다. 모듈은 OpenAI의 GPT-4.1을 기반으로 하며, Python API를 이용한 호출 방식으로 구현되었다. 전체 구조는 두 단계의 프롬프트 체인으로 구성되어 있으며, 각 단계의 처리 과정은 그림 8(a), (b), (c)에 구분되어 나타난다.

그림 8(a)는 모듈의 입력(Input)과 컨텍스트(Context) 구성을 나타낸다. 입력에는 DASS-21 설문 결과와 인지 기능 콘텐츠 수행 점수가 포함되며, 컨

텍스트에는 DASS-21 가이드라인과 콘텐츠별 점수의 정상 기준 및 컷오프(Cut-off) 정보가 포함된다. 이러한 구조를 통해 모델이 해석 과정에서 객관적 기준을 참조할 수 있도록 하였다.

그림 8(b)는 1단계 체인(Chain #1)을 나타내며, 이 단계에서는 첫 번째 프롬프트(Prompt #1)가 입력 데이터와 컨텍스트를 기반으로 작동한다. LLM은 주어진 기준을 참조하여 항목별 상태를 해석하고, 그 결과를 1차 출력(Output #1)으로 생성한다. 이 단계의 주요 목적은 개별 점수 수준에 대한 정량적·정성적 해석을 수행하는 것이다.

그림 8(c)는 2단계 체인(Chain #2)으로, 1단계에서

생성된 Output #1을 추가 컨텍스트로 활용하여 근로자의 전체적인 상태를 통합적으로 분석한다. 두 번째 프롬프트(Prompt #2)는 LLM을 통해 피검자의 심리·인지적 특성과 안전 수행 가능성을 종합적으로 판단하여 최종 소견(Output #2)을 도출한다.

이처럼 설계된 2단계 프롬프트 구조는 DASS-21 결과와 인지 기능 평가를 통합적으로 해석할 수 있게 하며, 컨텍스트 내 가이드라인과 컷오프 기준을 통해 결과의 일관성과 신뢰성을 확보하였다. 각 단계에서 사용된 프롬프트 및 컨텍스트의 구체적인 구성은 표 1과 같이 나타냈다.

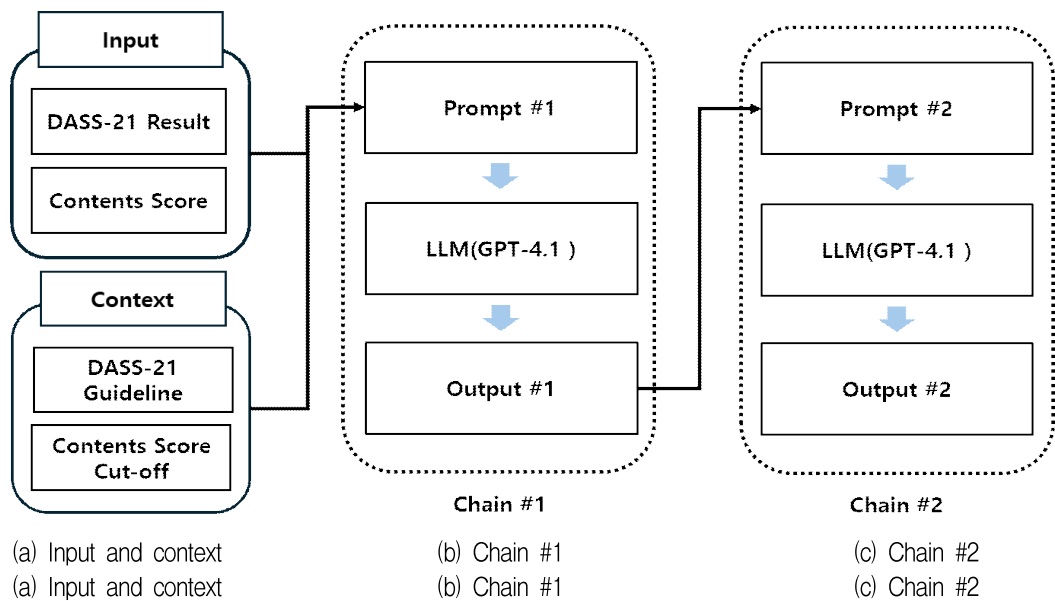


그림 8. 인지 기능 콘텐츠 및 DASS-21 점수를 기반으로 한 소견 생성 모듈의 처리 흐름도.

Fig. 8. Flowchart of the opinion generation module based on cognitive function contents and DASS-21 scores

표 1. 종합 소견 생성을 위한 프롬프트 및 컨텍스트 구성

Table 1. Prompt and context configuration for comprehensive opinion generation

Type	Section	Description
Prompt (zero-shot)	Prompt #1	Analyze each score based on DASS-21 and cognitive cut-offs to determine severity and normality
	Prompt #2	Summarize the overall assessment based on the analysis from Prompt #1, including field-readiness and professional interpretation
Context	DASS-21 guideline	The DASS-21 consists of 21 items, divided equally among three subscales: Depression, Anxiety, and Stress (7 items each). Scores are interpreted across five severity levels: Normal, Mild, Moderate, Severe, and Extremely Severe
	Contents score cut-off	The cognitive assessment includes tasks targeting working memory, attention, and concentration. Each task has defined normal ranges, such as response time accuracy or recall performance

2.5 위험 요소 식별 교육 콘텐츠

산업현장에서 발생하는 다양한 재해를 예방하기 위해서는 작업자가 잠재적 위험 요소를 신속하고 정확하게 식별하는 능력을 갖추는 것이 필수적이다. 본 연구에서는 ‘산업안전보건기준에 관한 규칙’을 기반으로 주요 위험 유형을 시나리오화하여 교육 콘텐츠를 설계하였다. 해당 규칙은 고용노동부가 제정한 대표적 산업 안전 규범으로 건설·제조 등 산업 현장에서 준수해야 할 세부 안전 기준을 포함한다.

이렇게 설계된 교육 콘텐츠는 한국남동발전에서 제공하는 안전보건 자료인 숨은 위험 찾기[32]를 참고하여 Unity 엔진을 활용해 그림 9와 같이 구현하였다. Unity 기반의 Spot the Hazards 콘텐츠는 실제 산업현장과 유사한 상황을 제공함으로써 학습자의 몰입도를 높이고, 사용자가 직접 위험 요소를 탐색하고 선택하는 과정을 통해 학습 효과를 극대화하도록 구현되었다.



그림 9 위험 요소 식별 교육 콘텐츠
Fig. 9. Spot the hazard content

III. 실험 및 구현 결과

3.1 대상자 구성 및 시나리오 설정

본 실험은 산업현장 근로자의 인지 기능 저하 상태를 모의하기 위해 음주 상태에 따른 인지 기능 평가를 수행하였다. 참가자는 총 12명(남 7명, 여 5명)으로 나이는 22세에서 28세(평균 24.6세)였다. 모든 참가자는 연구 참여에 앞서 연구 참여 동의서와 비밀 유지 계약서에 서명하였다. 또한, 알코올 사용 습관을 평가하기 위해 AUDIT(Alcohol Use Disorders Identification Test) 검사를 실시하였으며 참가자 중 알코올 의존으로 평가된 사례는 없었다.

3.2 데이터 수집 절차

음주 실험 전, 모든 참가자의 혈중알코올농도(BAC, Blood Alcohol Concentration)를 알코올 측정기를 통해 확인하여 사전에 음주 이력이 없음을 검증하였다. 이후 피험자는 남녀별 체중에 따라 정량화된 알코올을 20분 이내에 섭취하도록 지시받았다(남성: 체중 1kg당 0.9g, 여성: 체중 1kg당 0.8g). 실험에 사용된 음주는 한국 전통 증류주인 소주였으며 음주 직후 참가자들은 물로 입을 행구었다. 알코올 투여량은 도로교통공단에서 제시한 Widmark 공식을 한국식으로 수정·적용하여 산출하였다. 혈중알코올농도의 이론적 최고치는 다음과 같이 계산된다.

$$C = \frac{A \times 0.7}{P \times R} \quad (1)$$

식 (1)에서 C는 혈중알코올농도의 최고치(%)를 의미하며, A는 피험자가 섭취한 알코올의 양으로 이는 음주량(ml)에 술의 알코올 농도(%)와 알코올의 비중(0.7894)을 곱하여 산출된다. 또한, P는 피험자의 체중(kg)을 나타내며, R은 성별에 따른 계수로서 남성은 0.86, 여성은 0.64가 적용된다. 이 공식을 통해 피험자별 음주량을 체중과 성별에 맞추어 산정하였다. 이후 BAC는 음주 후 약 20분 동안 5분 간격으로 반복 측정하여 최고치 도달 시점을 확인하였다. 인지 기능 평가 콘텐츠는 음주 전·후 각각 다른 일자에 3회씩 수행하도록 하였으며 같은 과제를 반복함으로써 음주에 따른 수행 능력의 차이를 분석하였다[33]. 콘텐츠 수행은 Window 11, Chrome 웹 브라우저에서 진행하였으며, 웹은 nginx 1.24, 서

버는 spring boot 3.2.5, Open JDK 17, 데이터베이스는 MariaDB 8.0.36에서 구동되었다.

3.3 실험 결과

본 연구에서는 음주 전·후 인지 기능 수행 결과의 차이를 검증하기 위하여, 참가자별 3회 수행 평균값을 이용하여 대응 표본 t-검정(Paired samples t-test)을 하였다. 대응 표본 t-검정은 같은 집단의 두 조건(비음주와 음주)을 비교하여 평균 차이가 통계적으로 유의한지를 확인하는 검정 방법으로, 음주의 영향이 실제로 수행 차이를 유발했는지를 평가할 수 있다. 또한, 단순히 유의성(p-value) 여부에 그치지 않고 효과의 크기와 실질적 의미를 함께 제시하기 위해 Cohen's d를 산출하였다. Cohen's d는 두 조건 간 평균 차이를 표준편차로 나눈 값으로, 0.2

는 작은 효과(Small), 0.5는 중간 효과(Medium), 0.8 이상은 큰 효과(Large)로 해석된다[34]. 이를 통해 단순한 통계적 유의성뿐 아니라, 음주가 인지 기능 수행에 미치는 영향의 강도와 실제적 중요성을 종합적으로 검증하고자 하였다.

표 2는 음주 전(M1)과 음주 후(M2) 각각의 인지 기능 수행 결과 및 두 조건 간 차이(M2 - M1)를 나타낸 것이다. 여기서 M2 - M1 값은 음주 후 점수에서 음주 전 3회 평균 점수를 뺀 값으로 산출하였다. 과제별로 음주의 부정적 영향을 판단하는 기준은 다음과 같다. Doorlock과 Simon 과제는 음주 후 수행 개수가 감소한 경우, Hammering과 Traffic Light 과제는 음주 후 반응속도가 느려진 경우, N-back 과제는 정답률이 감소한 경우, 그리고 Mole 과제는 점수가 감소하면 음주의 부정적 영향으로 간주하였다.

표 2. 음주 전(M1)과 음주 후(M2) 인지 기능 과제 수행 결과 및 차이(M2 - M1)

Table 2. Cognitive task results before (M1) and after (M2) alcohol consumption and differences (M2 - M1)

Contents	Doorlock			Hammering			N-back		
Subject	M1	M2	M2-M1	M1	M2	M2-M1	M1	M2	M2-M1
Subject 1	7.3	5.7	-1.6	20.6	24.2	3.6	100.0	98.3	-1.7
Subject 2	12.3	9.3	-3.0	20.4	22.5	2.1	100.0	91.7	-8.3
Subject 3	9.3	9.0	-0.3	19.0	20.6	1.6	100.0	85.0	-15.0
Subject 4	8.3	8.0	-0.3	20.1	26.4	6.3	100.0	88.3	-11.7
Subject 5	6.7	10.3	3.6	27.3	25.0	-2.3	98.3	90.0	-8.3
Subject 6	10.3	9.0	-1.3	27.7	25.6	-2.1	100.0	51.7	-48.3
Subject 7	8.0	6.7	-1.3	27.7	29.1	1.4	100.0	98.3	-1.7
Subject 8	8.0	6.0	-2.0	20.2	21.7	1.5	86.7	90.0	3.3
Subject 9	9.3	11.3	2.0	17.7	18.8	1.1	100.0	100.0	0.0
Subject 10	10.7	8.3	-2.4	21.9	23.9	2.0	83.3	73.3	-10.0
Subject 11	8.7	7.3	-1.4	21.5	19.9	-1.6	100.0	60.0	-40.0
Subject 12	10.7	9.7	-1.0	18.1	21.5	3.4	100.0	100.0	0.0

Contents	Simon			Traffic light			Mole		
Subject	M1	M2	M2-M1	M1	M2	M2-M1	M1	M2	M2-M1
Subject 1	14.3	13.0	-1.3	0.78	0.80	0.02	166.7	143.3	-23.4
Subject 2	13.0	11.3	-1.7	0.71	0.80	0.09	176.7	153.3	-23.4
Subject 3	8.3	5.0	-3.3	0.76	0.91	0.15	156.7	120.0	-36.7
Subject 4	7.7	7.0	-0.7	0.76	0.85	0.09	153.3	130.0	-23.3
Subject 5	9.3	9.7	0.3	0.76	0.83	0.07	183.3	196.7	13.4
Subject 6	9.0	8.3	-0.7	1.00	0.82	-0.18	176.7	136.7	-40.0
Subject 7	9.0	6.7	-2.3	1.00	1.16	0.16	176.7	146.7	-30.0
Subject 8	6.0	5.3	-0.7	0.62	0.85	0.24	136.7	156.7	20.0
Subject 9	14.3	7.7	-6.6	0.74	1.01	0.27	186.7	156.7	-30.0
Subject 10	12.7	12.0	-0.7	0.75	0.76	0.01	243.3	220.0	-23.3
Subject 11	13.7	8.0	-5.7	0.81	0.82	0.01	173.3	163.3	-10.0
Subject 12	18.7	11.3	-7.4	0.76	0.95	0.19	230.0	183.3	-46.7

표 3 음주 전(M1)·후(M2) 인지 기능 과제 수행 비교 (대응 표본 t-검정 결과)

Table 3. Comparison of cognitive task result before (M1) and after (M2) alcohol consumption (Paired-Sample t-Test results)

Measure Contents	M1	M2	t	p	Cohen's d
Doorlock	9.5	8.1	3.79	.009	0.94
Simon	12.0	9.5	2.80	.031	0.75
Hammering (s)	21.1	24.0	-4.50	.004	1.04
Traffic light (s)	0.79	0.89	-3.87	.008	0.89
N-back (%)	97.6	90.7	3.15	.020	0.93
Mole	186.2	156.7	8.57	< .001	0.91

전체 참가자 12명 가운데, 체중에 맞는 알코올을 섭취하였음에도 불구하고 취기 민감도, 인지능력 저하 발생 시점, 개인별 주량 차이 등의 개인적 요인으로 인해 음주 전·후 수행 변화가 뚜렷하지 않은 사례가 일부 확인되었다. 이에 따라 모든 6개 과제에서 부정적 변화가 공통으로 나타난 7명을 최종 분석 대상으로 선정하였다. 표 3은 해당 7명을 대상으로 한 통계 분석으로, 음주가 모든 인지 기능 과제에 부정적인 영향을 미쳤음이 확인되었다. Doorlock 과제에서는 음주 조건에서 수행 개수가 유의하게 감소하였다($t = 3.79$, $p = .009$, $d = 0.94$). Simon 과제 또한 음주 시 수행 개수가 감소하는 경향이 나타났으며($t = 2.80$, $p = .031$, $d = 0.75$), Hammering 과제에서는 음주 조건에서 수행 시간이 유의하게 증가하였다($t = -4.50$, $p = .004$, $d = 1.04$). Traffic Light 과제에서도 음주 시 반응시간이 증가하는 경향을 보였으며($t = -3.87$, $p = .008$, $d = 0.89$), N-back 과제에서는 음주 조건에서 정확도가 유의하게 감소하였다($t = 3.15$, $p = .020$, $d = 0.93$). 마지막으로 Mole 과제에서는 음주 시 점수가 유의하게 감소하는 것으로 확인되었다($t = 8.57$, $p < .001$, $d = 0.91$). 이러한 결과는 음주가 전반적인 인지 기능 전 영역에 걸쳐 부정적인 영향을 미친다는 점을 보여준다. 특히 본 연구에서 사용된 6개 인지 기능 콘텐츠(Doorlock, Simon, Hammering, Traffic Light, N-back, Mole)가 모두에서 유의한 수행 저하가 관찰되었으며, 개인차를 고려하여 유의미한 변화를 보인 집단을 분석한 결과, 효과 크기(Cohen's d) 또한 대부분 중간~대 효과(Medium to large effect size)로 나타나, 음주의 인지 기능 저하 영향이 큰 수준임을 확인하였다.

3.4 근로자 및 관리자 웹앱

본 연구에서 개발한 플랫폼은 근로자용 웹앱과 관리자용 웹앱으로 구분되며 각 사용자 그룹의 역할에 따라 UI/UX와 기능을 차별화하였다.

근로자 웹앱은 메인 페이지에서 6종의 인지 기능 검사 콘텐츠와 DASS-21 설문을 제공하는 것으로 그림 10에서 보여주고 있다. 각 콘텐츠는 처음 수행하는 사용자를 위해 튜토리얼을 제공하며 수행이 종료되면 결과 데이터가 자동으로 중앙 데이터베이스에 저장된다. 콘텐츠 종료 후 사용자는 메인 페이지로 복귀하여 다음 콘텐츠를 이어서 수행할 수 있다. DASS-21 설문은 체크박스 형식으로 순서대로 응답하게 되며, 모든 문항에 응답 완료 시 마지막 항목에서 제출 버튼이 활성화되어 제출할 수 있다.

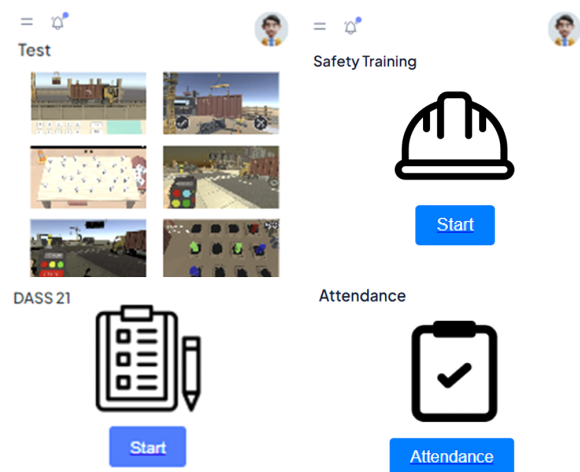


그림 10. 근로자 웹앱 화면

Fig. 10. Worker web app screen

제출된 설문 결과 역시 데이터베이스에 저장된다. 그 외에도 근로자는 출석 체크, 일정 관리, 공지

사항 확인, 비상 연락망 조회, 전자서명 등 산업현장에서 필요한 기본 기능들을 제공받는다.

관리자 웹앱은 근로자의 개인정보 및 출석 관리, 각 콘텐츠 수행 결과와 DASS-21 설문 결과의 조회 기능을 포함한다. 또한, 출석률, 콘텐츠 이행률 등의 데이터를 그래프 형태로 시각화하여 현장 관리자가 직관적으로 파악할 수 있도록 그림 11과 같이 구현하였다. 이와 함께 공지사항 게시, 일정 관리, 전자서명 관리 등 현장 운영과 관리에 필요한 기능들을 제공한다.

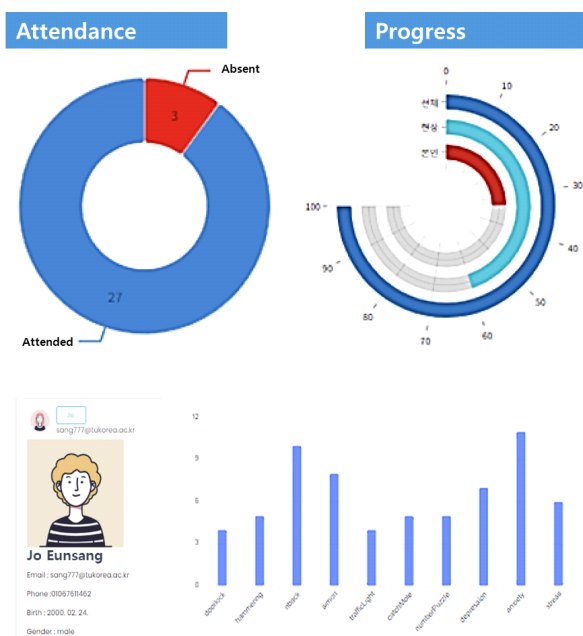


그림 11. 관리자 웹앱 화면(대시보드, 근로자 정보 조회)
Fig. 11. Manager web app screen (dashboard, worker information inquiry)

3.5 LLM 종합 소견 리포트

본 연구에서 개발한 LLM 기반 종합 소견 모듈은 데이터베이스에 저장된 인지 기능 콘텐츠 수행 점수와 DASS-21 설문 결과를 입력값으로 활용하였다. 해당 모듈은 각 점수를 기반으로 컷오프 기준 및 DASS-21 표준과 비교하여 정상/비정상 여부를 판정하고, 이를 표와 그래프 형태로 시각화된 리포트를 제공한다.

그림 12는 DASS-21 결과로, 우울·불안·스트레스 하위 척도별 점수와 판정 결과를 막대그래프로 시각

화한 화면이다. 우울 척도는 0 - 9점을 정상(normal), 10 - 13점을 경도(mild), 14 - 20점을 중등도(moderate), 21 - 27점을 중증(severe), 28 - 42점을 극심(extremely severe)으로 구분하였다. 불안 척도는 0 - 7점을 정상, 8 - 9점을 경도, 10 - 14점을 중등도, 15 - 19점을 중증, 20 - 42점을 극심으로 구분하였으며, 스트레스 척도는 0 - 14점을 정상, 15 - 18점을 경도, 19 - 25점을 중등도, 26 - 33점을 중증, 34 - 42점을 극심으로 정의하였다. 이러한 구간별 점수는 색상 스케일을 통해 시각적으로 구분되도록 구성하였다.

DASS-21 Scores

Subscale	Score	Category
Depression	24	Severe
Anxiety	10	Moderate
Stress	28	Severe

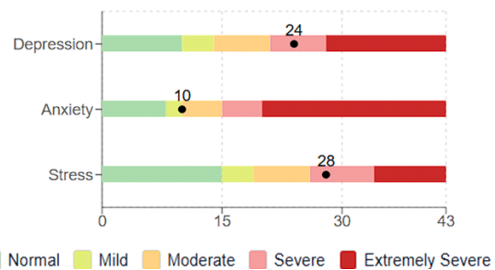


그림 12. 우울, 불안, 스트레스 하위 척도의 점수와 판정 결과를 나타낸 그래프

Fig. 12. Graphs of subscales of depression, anxiety, and stress and test results

그림 13은 인지 기능 콘텐츠 결과를 표와 그래프로 시각화한 화면이다. 인지 기능 콘텐츠 결과표에서는 Door Lock, Hammering, N-back, Simon, Traffic Light, Mole 등 6종의 과제 수행 점수가 기준값과 함께 제시되며, 정상 여부가 표와 그래프를 통해 직관적으로 확인할 수 있도록 나타났다.

본 연구에서 각 과제의 수행 점수는 Doorlock과 Simon은 수행 개수, Traffic Light와 Hammering은 수행 시간, Mole은 점수, N-back은 정확도를 기준으로 산출된다. 과제별 산출 지표가 서로 달라 직접적인 비교가 어려우므로, 본 연구에서는 기준값 대비 백분율 환산(Percentage normalization)을 통해 점수를 표준화하였다.

Cognitive Content Results

Task	Score	Cut-off	Status
Door Lock	4	6	Abnormal
Hammering	19.76	39	Normal
N-back	80	91	Abnormal
Simon	6	5	Normal
Traffic Light	0.83	0.85	Normal
Mole	190	150	Normal

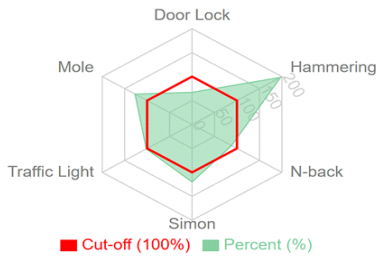


그림 13. 인지 기능 콘텐츠 수행 결과와 기준값 비교: 과제별 점수 및 정상/비정상 판정

Fig. 13. Comparison of cognitive function content performance results and cut-off values: Scores by task and normal/abnormal determination

이를 통해 서로 다른 단위를 가진 인지 기능 과제를 공통 척도에서 비교할 수 있었으며, 수행 안정성과 정상/비정상 여부를 직관적으로 시각화할 수 있었다. 그래프는 이러한 표준화된 점수를 기반으로 한 레이더 차트(Radar chart)로, 각 꼭짓점은 개별 인지 과제를 나타내며, 녹색 영역은 참가자의 수행 수준을, 붉은 선은 각 과제의 기준값(cut-off 선)을 의미한다. 또한, LLM을 활용하여 정량적 점수 해석을 바탕으로 한 종합 소견을 자동으로 생성한다. 예를 들어, 인지 기능 점수가 기준치 이상으로 유지되어 정상 범위임을 보고하는 동시에, DASS-21 결과에서 우울·불안·스트레스가 중등도 수준임을 반영하

여, 피로 누적 및 긴장 상황에서의 안전 위험을 고려해야 한다는 결론을 그림 14와 같이 제시하였다. 이와 같이 모듈은 단순 점수 산출을 넘어 인지 기능 상태와 심리적 위험 요인을 통합적으로 해석한 리포트를 제공한다.

IV. 논 의

본 연구는 고위험 산업현장의 안전관리 체계에 근로자의 인지 및 심리 상태를 통합적으로 반영하는 디지털 TBM(Tool Box Meeting) 플랫폼을 제안하였다. 기존 TBM이 주로 행정적 절차와 형식적 점검에 머물렀던 것과 달리, 본 플랫폼은 인지 기능 검사 콘텐츠와 DASS-21 기반의 심리 평가를 통합하여 근로자 개인의 상태를 현장 투입 전 파악하고, 이를 토대로 안전관리를 제공할 수 있도록 설계되었다.

또한, 인지 기능 평가를 위한 6종의 기능성 게임 콘텐츠(Doorlock, Simon, Hammering, Traffic Light, N-back, Mole)를 개발하고, 알코올 섭취 전후 실험을 통해 콘텐츠의 타당성을 검증하였다. 이 과정에서 각 과제의 수행 정확도, 반응시간, 오류율 등을 지표로 사용하여 음주 조건이 인지 수행에 미치는 영향을 통계적으로 분석하였다. 또한, DASS-21 설문 결과를 병합함으로써, 인지적 수행 저하가 단순한 피로나 심리적 불안 요인과 어떻게 상호작용을 하는지 탐색하였다.

그 결과, 본 연구의 플랫폼은 전통적인 행정 중심 TBM의 한계를 보완하며, 개인 맞춤형 안전관리 체계로의 전환 가능성을 제시하였다. 인지 기능 콘텐츠를 기반으로 한 정량적 평가와 LLM(GPT-4.1) 기반의 자동 소견 생성을 결합함으로써, 관리자는 근로자의 상태를 직관적으로 파악하고 현장 투입 여부를 합리적으로 판단할 수 있게 되었다.

Summary

This report integrates DASS-21 psychological screening with cognitive task results. Overall cognitive function appears within normal limits; attention and executive control relevant for task performance are preserved. On DASS-21, depression and anxiety are in the severe range, indicating a need for emotional management. Stress is moderate-to-severe, suggesting elevated risk under fatigue or high-tension situations. Field deployment is cognitively acceptable; however, counseling and a stabilization program are recommended prior to reassessment.

그림 14. LLM 기반 자동 종합 소견 리포트 예시: 인지 기능 점수와 DASS-21 결과를 통합한 해석

Fig. 14. Example of an LLM-based automatic comprehensive findings report

또한, 산업재해의 원인이 단순한 작업 실수뿐 아니라 피로, 스트레스, 불안 등의 심리적 요인으로도 밀접하게 관련되어 있음을 고려할 때, 본 연구는 근로자의 정신·인지 상태를 수치화하고 관리 체계에 통합하려는 새로운 접근 방향을 제시하였다. 이러한 접근은 정신적 피로 누적이나 인지 기능 저하가 예상되는 근로자를 선제적으로 식별하고 보호함으로써, 산업현장의 예방 중심 안전관리 패러다임으로의 전환을 촉진할 수 있을 것으로 판단된다.

V. 결 론

본 연구는 인지 기능 검사 결과와 심리 평가 데이터를 통합적으로 활용하여, 근로자 상태에 기반한 디지털 TBM 플랫폼의 개념을 제안하고 그 가능성을 검증하였다. 본 플랫폼은 근로자의 인지적·심리적 특성을 반영함으로써 기존 행정 중심의 안전관리 방식에서 벗어나, 개인 맞춤형 안전관리 체계로의 전환 가능성을 보여주었다.

다만, 제시된 컷오프 점수는 기존 검사 프로토콜의 기준값을 그대로 적용한 것으로, 디지털 콘텐츠의 특성과 피험자의 연령대와 개인차를 충분히 반영하지 못한 한계가 있다. 향후 연구에서는 현장 데이터를 지속해서 축적하여 개인화된 기준점을 설정하고, 이를 기반으로 더욱 정밀한 판단 알고리즘을 구축할 필요가 있다. 또한, LLM 기반 소견 생성 모듈은 GPT-4.1 모델을 활용하여 실질적인 분석 및 보고서 생성을 수행할 수 있음을 확인하였으나, 설명 가능성(Explainability) 측면에서는 보완이 요구된다. 모델의 추론 과정을 시각화하거나 근거 중심의 해석 구조를 도입함으로써, 해석력과 신뢰성을 강화하는 후속 연구가 필요하다.

아울러 본 연구는 고위험 산업군 전반을 대상으로 개념적 검증을 수행하였으나, 향후 연구에서는 산업별 특성과 작업환경을 반영한 맞춤형 콘텐츠 개발이 필요하다. 예를 들어 건설현장은 공간 인식 및 주의 집중력, 제조현장은 반응속도가 상대적으로 중요하므로, 산업별 핵심 인지 기능을 중심으로 특화된 콘텐츠를 제공한다면 더욱 실질적인 안전관리 효과를 기대할 수 있을 것이다. 현재 구현된 인지

기능 콘텐츠는 복수의 인지 기능을 제안하고 있지만, 향후 산업군에 따라 요구되는 인지 기능 콘텐츠를 반영한 플랫폼 수행 절차 간소화와 현장에 맞는 수행 주기 및 평가 방식에 관한 추가 연구 수행이 필요하다.

결론적으로, 본 연구는 인지 및 심리 상태 기반의 디지털 TBM 시스템이라는 새로운 접근을 제시함으로써, 산업현장의 안전관리 패러다임을 개인 중심의 예방형 체계로 확장할 가능성을 보여주었다. 향후 축적된 데이터 기반의 지속적 고도화와 현장 적용 연구가 병행된다면, 근로자의 상태를 실시간으로 감지하고 위험을 사전에 예측하는 산업안전 관리 플랫폼으로 발전할 것으로 기대된다.

References

- [1] Ministry of Employment and Labor, "2024 Industrial Accident Statistics", Apr. 2025. <https://oshri.kosha.or.kr/kosha/data/industrialAccidentStatus.do?mode=view&articleNo=455492&article.offset=0&articleLimit=10>. [accessed: Oct. 20, 2025]
- [2] A. S. Howe, et al., "Physical and Psychosocial Correlates of Occupational Physical Injury in the Global Construction Industry: A Scoping Review", *Environmental Health Insights*, Vol. 18, Aug. 2024. <https://doi.org/10.1177/11786302241270371>.
- [3] C. S. Kang and J. W. Lee, "Effect of TBM activities in manufacturing industry on safety factors: Focus on the safety communication, safety importance culture, regulatory procedure culture", *Korean Journal of Safety Culture*, Vol. 28, pp. 139-153, Mar. 2024. <http://doi.org/10.52902/kjsc.2024.28.139>.
- [4] Q. Xiang, G. Ye, Y. Liu, Y. M. Goh, D. Wang, and T. He, "Cognitive mechanism of construction workers' unsafe behavior: A systematic review", *Safety Science*, Vol. 159, pp. 106037, Mar. 2023. <https://doi.org/10.1016/j.ssci.2022.106037>.
- [5] R. Hussain, S. F. A. Zaidi, A. Pedro, H. Lee, and C. Park, "Exploring construction workers' attention

- and awareness in diverse virtual hazard scenarios to prevent struck-by accidents", *Safety Science*, Vol. 175, pp. 106526, Jul. 2024. <https://doi.org/10.1016/j.ssci.2024.106526>.
- [6] National Institute for Occupational Safety and Health, "Construction Program", Centers for Disease Control and Prevention, May 2024. <https://www.cdc.gov/niosh/construction/about/index.html>. [accessed: Oct. 20, 2025]
- [7] P. Bufano, et al., "The effects of work on cognitive functions: a systematic review", *Frontiers in Psychology*, Vol. 15, pp. 1351625, May 2024. <https://doi.org/10.3389/fpsyg.2024.1351625>.
- [8] M. Jung, S. Lim, and S. Chi, "Impact of work environment and occupational stress on safety behavior of individual construction workers", *International Journal of Environmental Research and Public Health*, Vol. 17, No. 22, pp. 8304, Nov. 2020. <https://doi.org/10.3390/ijerph17228304>.
- [9] N. Kokubo, et al., "A new device-aided cognitive function test, User eXperience-Trail Making Test (UX-TMT), sensitively detects neuropsychological performance in patients with dementia and Parkinson's disease", *BMC Psychiatry*, Vol. 18, No. 220, Jul. 2018. <https://doi.org/10.1186/s12888-018-1795-7>.
- [10] P. Traunmüller, et al., "Wearable healthcare devices for monitoring stress and attention level in workplace environments", *arXiv preprint arXiv:2406.05813*, Jun. 2024. <https://doi.org/10.48550/arXiv.2406.05813>.
- [11] K. W. Kim, "An Exploratory Study on Psychological Burden Factors Affecting Workers' Occupational Safety and Health, Occupational Safety and Health Research Institute", Korea Occupational Safety and Health Agency (KOSHA), 2016. <https://share.google/TWZYbybd3wsnsnQqW>. [accessed: Oct. 20, 2025]
- [12] A. M. Zavitsanou and A. Drigas, "Attention and working memory", *International Journal of Recent Contributions in Engineering, Science & IT*, Vol. 9, No. 1, pp. 81-91, Mar. 2021. <https://doi.org/10.3991/ijes.v9i1.19933>.
- [13] E. J. Slattery, et al., "Contributions of working memory and sustained attention to children's reading achievement: A commonality analysis approach", *Cognitive Development*, Vol. 58, pp. 101028, Apr.-Jun. 2021. <https://doi.org/10.1016/j.cogdev.2021.101028>.
- [14] K. Oberauer, "Working memory and attention—A conceptual analysis and review", *Journal of Cognition*, Vol. 2, No. 1, pp. 36, Aug. 2019. <https://doi.org/10.5334/joc.58>.
- [15] R. E. Reigal, et al., "Relationships between reaction time, selective attention, physical activity, and physical fitness in children", *Frontiers in Psychology*, Vol. 10, pp. 2278, Oct. 2019. <https://doi.org/10.3389/fpsyg.2019.02278>.
- [16] J. Llinàs-Reglà, et al., "The Trail Making Test: Association with other neuropsychological measures and normative values for adults aged 55 years and older from a Spanish-speaking population-based sample", *Assessment*, Vol. 24, No. 2, pp. 183-196, 2017. <https://doi.org/10.1177/1073191115602552>.
- [17] J. D. Matarazzo, et al., "Psychometric and clinical test-retest reliability of the Halstead Impairment Index in a sample of healthy, young, normal men", *The Journal of Nervous and Mental Disease*, Vol. 158, No. 1, pp. 37-49, Jan. 1974. <https://doi.org/10.1097/00005053-197401000-00006>.
- [18] F. Mathy, et al., "Developmental abilities to form chunks in immediate memory and its non-relationship to span development", *Frontiers in Psychology*, Vol. 7, pp. 201, Feb. 2016. <https://doi.org/10.3389/fpsyg.2016.00201>.
- [19] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information", *Psychological Review*, Vol. 63, No. 2, pp. 81-97, 1956. <https://psycnet.apa.org/doi/10.1037/h0043158>.

- [20] J. R. Gray, C. F. Chabris, and T. S. Braver, "Neural mechanisms of general fluid intelligence", *Nature Neuroscience*, Vol. 6, No. 3, pp. 316-322, Feb. 2003. <https://doi.org/10.1038/nn1014>.
- [21] S. M. Jaeggi, et al., "The concurrent validity of the N-back task as a working memory measure", *Memory*, Vol. 18, No. 4, pp. 394-412, Apr. 2010. <https://doi.org/10.1080/09658211003702171>.
- [22] A. Jain, et al., "A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students", *International Journal of Applied and Basic Medical Research*, Vol. 5, No. 2, pp. 124-127, May-Aug. 2015. <https://doi.org/10.4103/2229-516X.157168>.
- [23] W. S. Ryu, et al., "An empirical study on quantitative evaluation of cognitive function", *Progress in Medical Physics*, Vol. 21, No. 1, pp. 42-51, Mar. 2010.
- [24] D. L. Woods, et al., "Improving digit span assessment of short-term verbal memory", *Journal of Clinical and Experimental Neuropsychology*, Vol. 33, No. 1, pp. 101-111, Aug. 2011. <https://doi.org/10.1080/13803395.2010.493149>.
- [25] F. Verbruggen and G. D. Logan, "Automatic and controlled response inhibition: associative learning in the go/no-go and stop-signal paradigms", *Journal of Experimental Psychology: General*, Vol. 137, No. 4, pp. 649-672, 2008. <https://doi.org/10.1037/a0013170>.
- [26] T. Tong, et al., "Rapid deployment and evaluation of mobile serious games: A cognitive assessment case study", *Procedia Computer Science*, Vol. 69, pp. 96-103, 2015. <https://doi.org/10.1016/j.procs.2015.10.010>.
- [27] T. Tong, et al., "A serious game for clinical assessment of cognitive status: validation study", *JMIR Serious Games*, Vol. 4, No. 1, pp. 5006, Jan.-Jun. 2016. <https://doi.org/10.2196/games.5006>.
- [28] P. F. Lovibond and S. H. Lovibond, "The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories", *Behaviour Research and Therapy*, Vol. 33, No. 3, pp. 335-343, Mar. 1995. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U).
- [29] D. Jun, et al., "Cross-cultural adaptation and validation of the Depression, Anxiety and Stress Scale-21 (DASS-21) in the Korean working population", *Work*, Vol. 59, No. 1, pp. 93-102, Feb. 2018. <https://doi.org/10.3233/WOR-172661>.
- [30] E. Wang, et al., "Recommendation of music based on DASS-21 (Depression, Anxiety, Stress Scales) using fuzzy clustering", *Electronics*, Vol. 12, No. 1, pp. 168, Dec. 2022. <https://doi.org/10.3390/electronics12010168>.
- [31] J. D. Henry and J. R. Crawford, "The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample", *British Journal of Clinical Psychology*, Vol. 44, No. 2, pp. 227-239, Jun. 2005. <https://doi.org/10.1348/014466505X29657>.
- [32] Korea Energy Agency, "Energy Efficiency Management System", <https://www.koenergy.kr/kosep/hw/fr/cs/cshw21/main.do?menuCd=FN052703>. [accessed: Oct. 20, 2025]
- [33] C. Vinader-Caerols, et al., "Blood alcohol concentration-related lower performance in immediate visual memory and working memory in adolescent binge drinkers", *Frontiers in Psychology*, Vol. 8, pp. 1720, Oct. 2017. <https://doi.org/10.3389/fpsyg.2017.01720>.
- [34] G. M. Sullivan and R. Feinn, "Using effect size—or why the P value is not enough", *Journal of Graduate Medical Education*, Vol. 4, No. 3, pp. 279-282, Sep. 2012. <https://doi.org/10.4300/JGME-D-12-00156.1>.

저자소개

양 철 영 (ChulYeong Yang)



1993년 8월 : KAIST
산업공학과(공학사)
1995년 8월 : KAIST
경영정보시스템(공학석사)
2000년 2월 : KAIST 경영공학
(공학박사)
2025년 9월 ~ 현재 :

한국공학대학교 컴퓨터공학과 박사과정
관심분야 : 디지털 헬스케어

조 은 상 (Eunsang Jo)



2025년 2월 : 한국공학대학교
컴퓨터공학과(공학사)
2025년 3월 ~ 현재 :
한국공학대학교
바이오헬스융합공학과 석사과정
관심분야 : 인공지능(DL), 디지털
헬스케어, 생체신호

김 상 규 (Sang-Gyu Kim)



2024년 0월 : 한국공학대학교
컴퓨터공학과(공학사)
2024년 3월 ~ 현재 :
한국공학대학교
바이오헬스융합공학과 석사과정
관심분야 : 인공지능(DL), 디지털
헬스케어, EEG, 생체신호

정 성 택 (Sungtaek Chung)



1992년 2월 : KAIST 전기 및
전자공학과(공학사)
1995년 2월 : KAIST 정보 및
통신공학과(공학석사)
2000년 2월 : KAIST 전기 및
전자공학과(공학박사)
2004년 3월 ~ 현재 :

한국공학대학교 컴퓨터공학과 교수
관심분야 : AI 의료 신호 및 영상처리, AI 헬스케어