

# 시험 지능(Test Intelligence)의 개념화 및 이론적 모델 제안: PRISMA 2020 기반 체계적 문헌 고찰과 주제 합성을 중심으로

김유정 (경기과학고등학교 교사)\*

## 요약

본 연구는 PRISMA 2020 지침에 따른 체계적 문헌 고찰과 주제 합성을 통해 시험 수행 관련 구인들의 공통 내용 영역을 분석하고 ‘시험 지능(Test Intelligence, TI)’의 개념과 가설적 이론 모델을 제안하였다. 연구 결과, 첫째, 시험 수행 관련 구인들은 ‘상황 판단, 정서 인식, 정서 조절, 행동 전략’의 4개 구인과 12개의 핵심 하위 요소(효능감, 중요성, 객관성, 긍정 정서 인식, 부정 활성 정서 인식, 부정 비활성 정서 인식, 과제 초점 유지, 정서 초점 조절, 인지적 재평가, 시간순서 관리, 문항 풀이 전략, 오류 방지)로 수렴되었다. 이를 바탕으로 시험 지능을 ‘학습자가 시험 상황에서 경험하는 다양한 인지적·정서적 반응을 효과적으로 인식 및 조절하고, 행동 전략을 통해 최적의 수행을 이끌어내는 종합적 실행 역량’으로 개념화하였다. 둘째, 본 연구에서는 시험 지능의 4개 구인이 단일 시험 내에서 순차적으로 작동하며 앞선 구인이 다음 구인의 발현에 영향을 미치고, 이전 시험에서 성공적으로 발휘된 시험 지능이 다음 시험 주기의 상황 판단을 강화하는 순환적 구조로 시험 지능의 모델을 제안하였다. 다만, 본 모델은 새로운 구인의 개념과 이론적 구조를 정립하는 연구로서, 향후 내용 타당도 검증과 통계적 타당화를 거쳐 더욱 정교화될 필요가 있다. 본 연구는 그간 학계에서 분절적으로 논의되어 온 시험 수행 관련 구인들을 ‘시험 지능’이라는 통합적 틀로 재구조화하여 학습자가 시험 맥락에서 겪는 역동적인 인지·정서·행동적 기제를 통합적으로 조망하였다는 데 의의가 있다. 아울러 본 연구가 시험 지능을 누적적으로 발달할 수 있다는 가변적 역량으로 개념화한 만큼, 학습자의 상황 판단, 정서 인식, 정서 조절, 행동 전략을 강화할 수 있도록 돕는 실천적 교수학습 방안이 도입될 필요가 있음을 제언한다.

주제어 : 시험 지능, Test Intelligence, PRISMA 2020, 주제 합성, 상황 판단, 정서 인식, 정서 조절, 행동 전략, 인지적 평가, 통제 가치 이론, 시험 수행 자기 효능감, 학업 부력, 메타인지 모니터링, 시험 불안, 성취 정서, 시험 지혜, 시험 전략

\* 제1저자 및 교신저자, lilious0660@gmail.com

## I. 서 론

시험(testing)이란, 학습자의 지식, 기능, 태도 등의 심리적 특성을 표준화된 절차와 기준에 따라 체계적으로 측정하는 도구로(성태제, 2019), 단순히 정답의 수를 집계하기보다는 학습자가 보유한 역량을 얼마나 효과적으로 발휘하는지를 측정하는 역동적이고 복합적인 수행 과정이라 볼 수 있다(Zeidner, 1998). 시험은 학습자의 역량을 진단하는 것을 넘어 대학 입학이나 자격 취득, 취업에 이르기까지 개인의 삶에 결정적인 영향을 미치므로, 교육 심리학과 교육 측정학 분야에서는 이러한 시험의 중요성에 주목하여 시험 수행에 영향을 미치는 요인들을 규명하기 위해 많은 노력을 기울여왔다(Haladyna & Downing, 2004).

전통적인 관점에서 시험 결과는 학습자가 얼마나 학업에 성실히 매진해왔는지를 대변하는 절대적 지표로 간주되어 왔다. 이에 따라 시험 점수가 낮다는 것은 곧 학습자가 그동안 학습에 성실하지 않았음을 증명하는 것이며, 인지적 수준이 비슷한 학습자 간에 성취도의 편차가 발생할 경우, 이는 운이나 시험 당일의 신체적 컨디션과 같은 통제 불가능한 외재적 오염변수 때문인 것으로 여겨졌다. 그러나 다수의 선행 연구들은 평상시 아무리 성실히 학습하고 예비 시험에서 우수한 성적을 유지한 학생이라 할지라도, 시험 상황을 마주했을 때 평소 수준에 미치지 못하는 저조한 성취를 보일 수 있다는 사실을 수십 년에 걸쳐 누적적으로 보고해 왔다.

Hembree(1988)는 시험 불안 관련 선행 연구를 메타 분석한 결과, 학습자의 실제 수행 능력은 높은 시험 불안으로 인해 유의미하게 저해될 수 있음을 증명하고, von der Embse 외(2018)는 대학 입학 등과 직결된 고부담 시험에서 이러한 부정적인 영향력이 더욱 크게 나타남을 보고하였다. 이러한 수행 저하 현상은 정서 불안 뿐만 아니라 학습자의 동기 저하와 전략 부족에 의해서도 발생할 수 있다. Pajares와 Miller(1994)는 자기 효능감이 부족한 학습자의 경우, 인지적 역량이 뛰어남에도 불구하고 시험 상황을 위협으로 인식하여 과제 집착력을 잃고 쉽게 시험을 포기하게 된다고 밝혔으며, Nielsen 외(2019)는 시험을 성공적으로 수행할 수 있다는 사전 신념이 낮을수록 이후의 학업 이탈 가능성이 높아짐을 확인하였다. 또한, Dodeen(2015)은 학습자가 적절한 시험 응시 전략을 갖추지 못할 경우, 자신이 이미 알고 있는 문제도 적절히 대처하지 못해 결과적으로 학업성취가 낮아질 수 있다고 보고하였다.

이러한 논의들은 대학수능능력시험이 개인의 인생에 매우 중요한 영향력을 행사하는 우리나라에서도 중요하게 다루어지고 있다. 선행 연구들에 따르면, 우리나라 학생들은 고부담 시험에서 심각한 수준의 시험 불안을 지속적으로 경험하고 있으며, 이는 학업 성취도를 낮출 뿐 아니라 청소년기의 삶의 질을 심각하게 위협하는 것으로 나타났다(김정현 외, 2014; 백승희, 2025; 정찬호 외, 2004; 최효식 외, 2026). 특히, 완벽주의적 성향을 가진 학생일수록 평가 상황을

심각한 위협으로 지각하는 경우가 많은데, 이는 실전에서 시험 불안을 급증시켜 본래의 실력을 충분히 발휘하지 못하게 한다(김미영, 이창현, 2023). 이러한 연구들은 시험 결과가 학습자의 평소 학업 충실도만을 반영하는 단일 변수가 아니라, 시험이란 특수 상황에서 자신이 지닌 지식과 기능을 발휘하는 데 관여하는 다차원적인 심리적 기제가 작동한 결과임을 시사한다. 만일 학습자의 역량을 효과적으로 발현시키는 심리적 기제를 통제 및 발달 가능한 변수로 정의할 수 있다면, 이는 시험 결과가 우연적인 요인에 의해 결정되는 것이 아니라 교육을 통해 충분히 향상될 수 있음을 의미한다. 더 나아가 학습자가 자신의 지식과 기능을 실전에서 온전히 발휘할 수 있도록 돕는 체계적인 교육의 이론적 토대가 마련될 수 있을 것이다.

따라서 본 연구는 Page 외(2021)의 PRISMA 2020 지침에 따라 시험 수행 관련 문헌을 체계적으로 분석하고, Thomas와 Harden(2008)의 주제 합성(thematic synthesis) 절차를 적용하여 학습자의 잠재적 역량과 실전 수행 역량 사이의 간극에 관여하는 심리적 기제를 통합적으로 설명하는 ‘시험 지능(Test Intelligence, TI)’ 개념과 구성 요인의 작동 방식에 관한 가설적 이론 모델을 제안하고자 하였다. 이는 Duckworth 등(2007)이 분산된 지속성 관련 구인들을 ‘그릿(grit)’이라는 단일 구인으로 통합하여 학업성취에 대한 설명력을 높인 것처럼, 시험 수행에 관여하는 다차원적인 구인들을 시험 지능이라는 통합적 틀로 재구조화하여 시험 수행 역량에 대한 설명력을 높이려는 시도이다. 그릿을 비롯하여 Salovey와 Mayer(1990)의 정서 지능, Earley 와 Ang(2003)의 문화 지능 등 다수의 심리 구인 역시 이론적 개념화에서 출발하여 실증 연구로 정교화되어 온 것처럼, 시험 지능이란 개념적 토대를 정립하는 일은 새로운 구인 연구를 위한 필수적인 첫 단계로서 의의를 지닌다.

본 연구에서 제안하는 시험 지능은 기존의 시험 관련 구인 및 개념들과 다음의 세 가지 측면에서 구별될 수 있다. 첫째, 기존 개념들이 시험 수행과 관련된 특정 심리적 현상을 기술하는 개별적인 구인이라면, 시험 지능은 이처럼 분산된 구인들을 포괄하여 학습자의 시험 수행 역량을 통합적으로 설명하는 상위 역량이라 할 수 있다. 둘째, 기존 구인들이 인지, 정서, 행동 중 특정 차원에 국한되어 현상을 기술해온 반면, 시험 지능은 시험 상황에서 이 세 가지 차원이 어떻게 상호작용하고 조절되는지를 구조적으로 탐색한다는 점에서 차별성을 지닌다. 셋째, 시험 지능은 구인 간에 산재해 있던 분산을 통합하는 상위 역량의 성격을 지니므로, 기존 구인들이 개별적으로는 설명하기 어려웠던 학습자의 시험 수행 역량을 보다 포괄적으로 설명할 수 있을 것으로 기대된다. 이러한 연구 목적을 달성하기 위한 구체적인 연구문제는 다음과 같다.

**연구문제 1.** 학습자의 시험 수행 과정에 영향을 미치는 심리적 요인 및 관련 구인은 무엇인가?

**연구문제 2.** 시험 수행 관련 선행 구인에서 제시하는 하위 구성 요소 간에 공통되는 영역은 무엇이며, 이를 주제 합성에 따라 어떻게 범주화할 수 있는가?

**연구문제 3.** 범주화된 공통 영역들을 바탕으로 도출한 시험 지능(Test Intelligence, TI)의 개념과 통합적 이론 모델은 무엇인가?

## II. 이론적 배경

### 1. 시험 지능(Test Intelligence, TI)

전통적 관점에서 지능(intelligence)은 유전적으로 결정되어 생애에 걸쳐 크게 변하지 않는 고정 불변의 것으로 간주되어 왔다(Spearman, 1904). 따라서 IQ 검사 등의 표준화 검사는 이를 언어, 수학 및 추론 등의 영역으로 수치화하여 협소하게 다루어 왔다. 그러나 Gardner(1983)는 기존의 지능 검사가 인간의 광범위한 잠재력을 담아내지 못한다고 비판하며, 지능을 ‘특정 문화적 맥락에서 가치 있는 문제를 해결하거나 산물을 창출하는 능력’으로 재정의할 것을 제안하였다. 이러한 이론적 전환은 Sternberg(1985)의 실용 지능(practical intelligence), Salovey와 Mayer(1990)의 정서 지능(emotional intelligence), Earley와 Ang(2003)의 문화 지능(cultural intelligence)의 이론적 토대가 되었으며, 이들은 특정 맥락에서 발현되는 복합적인 역량 체계로 교육적 개입을 통해 향상 가능한 가변적 특성을 지닌다는 공통점을 지닌다.

이러한 맥락에서 본 연구는 시험 수행과 관련한 학습자의 역량을 ‘시험 지능(Test Intelligence, TI)’이라 명명하고자 한다. 학습자는 엄격한 시·공간적 제약과 압박 속에서 자신의 역량을 최대한 발휘해야 하므로, 수준이 비슷한 학생들 간에도 심리적 압박을 견뎌내는 정도나 학습 내용을 인출하는 과정에 따라 시험 결과가 달라질 수 있다. 따라서 일반적인 학습 상황에서의 학업 수행 역량과 통제된 평가 상황에서의 시험 수행 역량은 엄연히 구별될 필요가 있다. 또한, 본 연구가 학습자의 시험 수행 역량을 단순한 역량을 넘어 시험 지능이라 명명하는 이유는 다음과 같다.

첫째, 학습자가 시험을 치르는 과정은 인지 처리 능력만으로는 설명할 수 없는 맥락 특수적인 역량 체계로, Gardner(1983)의 맥락 중심적 지능의 정의에 부합한다.

둘째, 시험 상황에서의 대처 역량은 고정된 것이 아니라 경험과 훈련으로 충분히 향상될 수 있음이 시험 불안 완화 및 프로그램, 메타인지적 모니터링 교육, 시험 전략 및 시험

지혜 관련 연구들에 의해 검증되었으며(Dodeen, 2015; Folkman & Lazarus, 1985; Martin & Marsh, 2008; Millman et al., 1965; Schutz & Davis, 2000), 이는 지능을 후천적으로 변화 가능한 역동적 개념으로 바라본 현대 지능 이론의 관점에 부합한다. 결론적으로 시험 수행 역량을 시험 지능이라 명명하는 것은 기존의 확장된 다중 지능 담론과 그 학술적 맥을 함께한다고 볼 수 있다.

## 2. PRISMA 2020

체계적 문헌 고찰은 연구 문제를 중심으로 문헌을 검토하고 사전에 수립된 기준에 따라 선별한 후, 표준화된 절차를 통해 종합함으로써 특정 주제에 관한 최선의 증거를 산출하는 연구 방법이다(Higgins & Green, 2011). 이는 검색 방법과 선별 기준을 사전에 명시함으로써 엄밀성과 투명성을 확보할 수 있다는 이점이 있다. PRISMA(Preferred Reporting Items for Systematic Reviews and Meta-Analyses)는 국제 보건 및 방법론 전문가그룹(PRISMA group)이 2009년 공동 발표한 체계적 문헌 고찰의 표준화된 지침으로, 문헌 검색 기술과 방법론의 발전을 반영하여 2020년에 전면 개정되었다(Page et al., 2021). 본 연구가 PRISMA 2020 지침을 활용한 것은 최신 국제 보고 표준을 엄격히 준수하여 연구자의 주관적 판단을 최소화하고 연구 절차의 재현성과 신뢰성을 확보하기 위함이다. PRISMA 2020은 연구의 흐름도(flow diagram)에 따른 문헌 선별을 권고하며, 검색 결과를 수집하는 ‘식별(identification)’, 제목, 초록 및 주제어를 검토하여 문헌을 골라내는 ‘선별(screening)’, 원문을 상세히 분석하여 기준 부합 여부를 판단하는 ‘적격성 평가(eligibility)’, 최종 분석 문헌을 확정하는 ‘포함(included)’의 4단계로 구성된다.

## 3. 주제 합성(thematic synthesis)

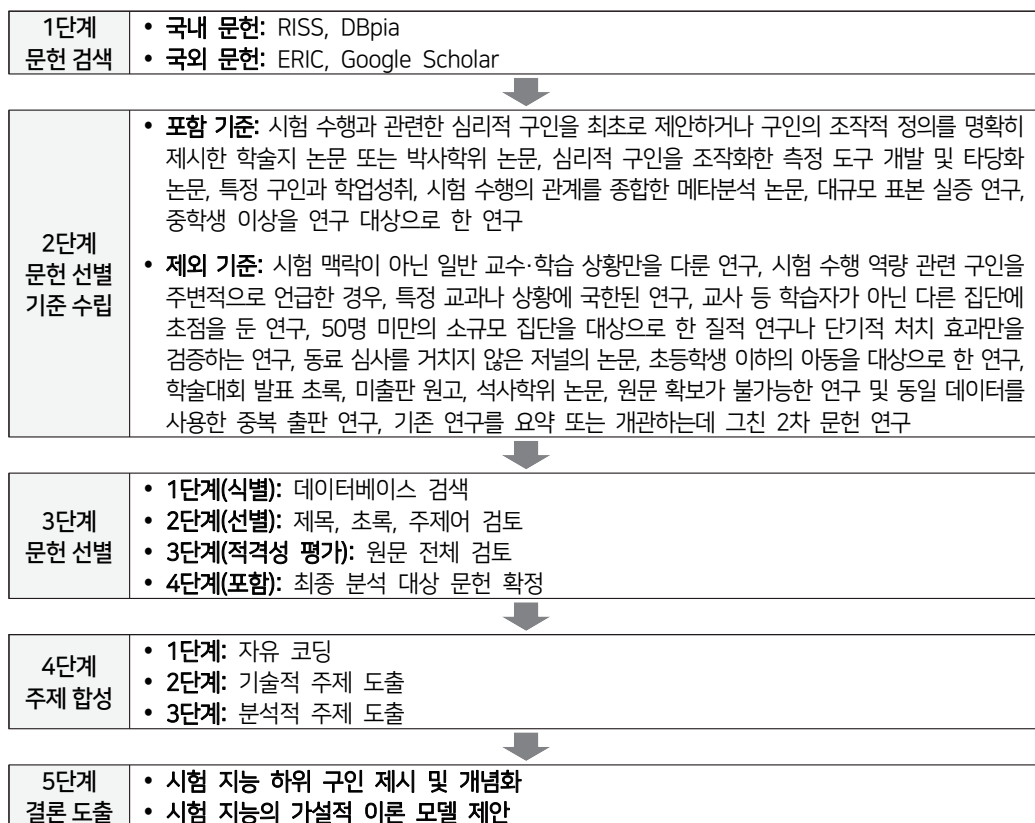
주제 합성은 Thomas와 Harden(2008)이 제안한 문헌 종합 방법론으로, 개별 연구 결과를 집계하는 것을 넘어 문헌에 내재된 개념들을 귀납적으로 코딩하여 상위주제로 통합하는 데 목적이 있다. 이는 체계적 문헌 고찰에 질적 연구의 주제 분석(thematic analysis)을 접목한 것으로, 연구 과정을 투명하고 재현 가능한 절차로 제시한다는 특징을 지닌다. 초기에는 질적 메타 합성을 위해 고안되었으나, 최근에는 특정 구인의 개념 구조와 속성을 탐색하거나 새로운 이론적 틀을 구성하는 문헌 고찰 연구에도 폭넓게 활용되어 왔다(Barnett-Page & Thomas, 2009).

주제 합성은 크게 3단계로 진행된다. ‘자유 코딩(free coding)’ 단계는 문헌을 행(line) 단위로 검토하여 의미 있는 단위마다 코드를 부여한다. ‘기술적 주제(descriptive themes)’

도출 단계는 생성된 코드들을 내용의 유사성에 따라 기술적인 범주로 묶는다. ‘분석적 주제(analytical themes)’ 도출 단계는 기술적 주제에 연구자의 이론적 통찰과 해석을 더해 새로운 개념의 구조를 생성한다. 이처럼 주제 합성은 귀납적 코딩에서 연역적 이론화로 나아가는 체계적인 절차를 거치므로, 분절된 문헌들로부터 상위의 통합적 이론 구조를 도출하는 데 효과적이다.

### III. 연구 방법

본 연구는 PRISMA 2020 지침에 따라 문헌을 체계적으로 분석하고, 주제 합성 방법에 따라 시험 수행과 관련한 기존 구인들의 개념적 관계를 이론적으로 재구조화하여 ‘시험 지능’이라는 통합적 개념을 제시하고 가설적 이론 모델을 제안하였다. 전체 연구 절차를 [그림 1]에 도식화하여 나타내었다.



[그림 1] 연구 절차

## 1. 문헌 검색

본 연구는 국내 문헌 검색에 학술연구정보서비스(RISS)와 DBpia를, 국외 문헌 검색에는 교육 분야 특화 데이터베이스인 ERIC(Education Resources Information Center)을 활용하였다. 본 연구에서 제안하는 ‘시험 지능’은 신생 구인으로, 단일 검색어만으로는 관련 문헌을 수집하는 데 한계가 있다. 따라서 ‘학습자의 시험 수행 과정에 영향을 미치는 심리적 요인 및 관련 구인은 무엇인가?’란 연구 문제를 중심으로, 시험 맥락 속에서 발현되는 인지, 정서, 행동 요인들을 포괄하기 위해 선행 연구에서 개별적으로 다루어져 온 심리적 요인들을 핵심 키워드로 선정하였다.

문헌 검색은 제목(title), 초록(abstract), 주제어(keywords) 범위 내에서 이루어졌으며, 국내 문헌에는 시험 맥락과 심리 기제의 교차 검색을 위해 ((“시험 수행” OR “시험 전략” OR “시험 정서” OR “시험 결과” OR “시험 대처” OR “시험 효능감” OR “시험 불안”) AND (“평가” or “시험”)) 검색식 조합을, 국외 문헌에는 ((“test performance” OR “test-taking strategy” OR “test-taking motivation” OR “test result” OR “test coping” OR “test self-efficacy” OR “test anxiety”) AND (“examination” OR “exam” OR “evaluation”)) 검색식을 사용하였다.

본 연구는 데이터베이스 검색 문헌 외에도 PRISMA 2020이 권고하는 수기 검색(hand search)을 통해 초기 검색에서 벗어난 주요 문헌들을 추가로 확보하였다. 1차 선별 문헌의 참고문헌 목록을 검토하는 역추적 방식(ancestry searching)으로 해당 문헌의 이론적 기반이 되거나 그로부터 파생된 후속 논문들을 Google Scholar를 통해 수집하였고, ERIC이 교육학에 국한되어 있는 한계를 보완하기 위하여 시험 불안이나 시험 정서 등을 심도 있게 다루는 심리학 및 인지과학 분야의 핵심 저널을 직접 탐색(journal browsing)하였다. 이러한 다각도적인 자료 수집을 통해 ‘시험 지능’이 특정 학문 분야에 편중되지 않고 인지, 정서, 행동 영역을 아우르는 다차원적인 구인으로 정립될 수 있도록 하였다.

## 2. 문헌 선별 기준 수립

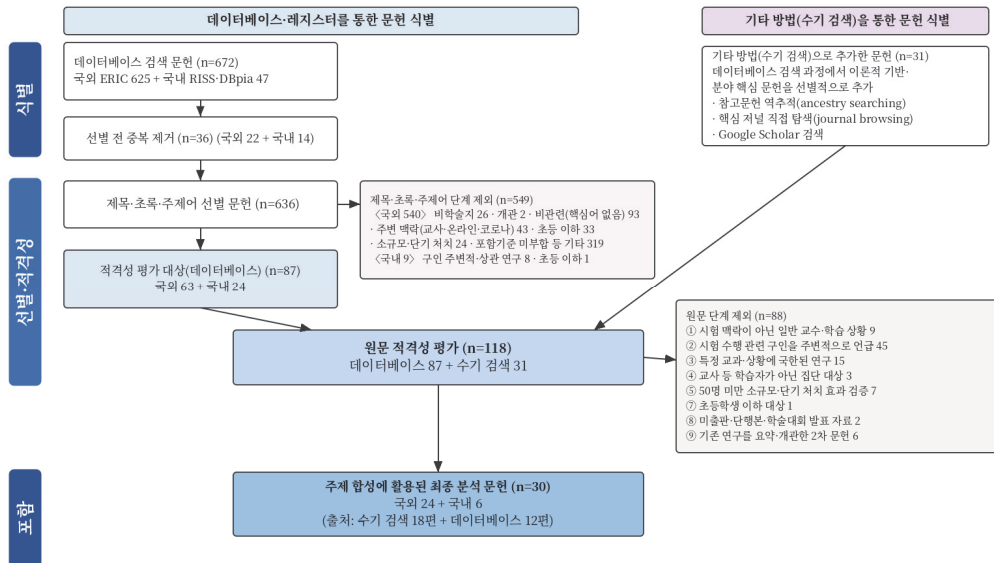
본 연구는 포함 기준과 제외 기준을 엄격히 수립하여 분석 대상 문헌을 선정하였다. 구체적인 포함 기준으로는 첫째, 학술지 게재 논문 또는 박사 학위 논문으로서 시험 수행과 관련한 심리적 구인을 최초로 제안하거나 해당 구인의 조작적 정의를 명확히 제시한 연구, 둘째, 심리적 구인을 조작화하여 측정 도구를 개발하고 타당화한 연구, 셋째, 특정 구인과 학업성취, 시험 수행 간의 관계를 종합한 메타분석 연구, 넷째, 대규모 표본을 활용한 실증 연구, 마지막으로 중학생 이상을 대상으로 한 연구로 한정하였다. 반면, 다음의 기준에 해당하는 문헌은

최종 분석 대상에서 제외하였다. 첫째, 시험 맥락이 아닌 일반 교수·학습 상황만을 다룬 연구, 둘째, 시험 수행 역량 관련 구인을 주변적으로 언급한 연구, 셋째, 특정 교과나 상황에 국한된 연구, 넷째, 교사 등 학습자가 아닌 다른 집단에 초점을 둔 연구, 다섯째, 50명 미만의 소규모 집단을 대상으로 한 질적 연구나 단기적 처치 효과만을 검증하는 연구, 여섯째, 동료 심사를 거치지 않은 저널의 논문, 일곱째, 초등학생 이하의 아동을 대상으로 한 연구, 여덟째, 학술대회 발표 초록, 미출판 원고, 석사학위 논문, 원문 확보가 불가능한 연구 및 동일 데이터를 사용한 중복 출판 연구, 마지막으로 기존 연구를 요약 또는 개관하는 데 그친 2차 문헌 연구이다.

### 3. 문헌 선별

본 연구는 [그림 2]의 PRISMA 2020 흐름도에 따라 식별, 선별, 적격성 평가, 포함의 4단계로 문헌을 선별하였다. 식별 단계에서 데이터베이스 검색을 통해 국외(ERIC) 625편, 국내(RISS, DBpia) 47편으로 총 672편을 확보하고, 중복 36편(국외 22편, 국내 14편)을 제거하여 636편을 식별하였다. 선별 단계에서는 636편의 제목, 초록, 주제어를 검토하여 포함 기준에 부합하지 않는 549편을 제외하고 87편(국외 63편, 국내 24편)을 선별하였다. 한편, 선별된 문헌의 참고 문헌을 역추적하고 심리·인지 과학 분야의 저널을 직접 탐색하여 핵심 문헌 총 31편(국외 27편, 국내 4편)을 추가하였다. 마지막으로 데이터베이스 경로의 87편과 수기 검색 31편을 합한 118편의 원문을 기준에 따라 심층 검토하여 88편을 제외한 최종 30편(국외 24편, 국내 6편; 수기 검색 18편 포함)을 분석 대상으로 확정하였다.

국외 문헌은 초기 식별 대비 최종 포함 비율이 약 38%로 현저히 낮게 나타났는데, 이는 검색어인 ‘test performance’, ‘examination’, ‘evaluation’ 등이 교육뿐 아니라 공학 시험, 프로그램 평가 등 학술 전반에서 쓰이는 고빈도 범용어이기 때문이다. 이로 인해 시험 수행과 무관한 문헌이 대량으로 함께 식별되었고, 제목·초록·주제어 선별 단계에서 이들 대부분이 제외되었다. 반면 국내 문헌은 초기 식별 건수 자체가 적었을 뿐 아니라 (47편), 최종 포함된 문헌 또한 소수(6편)에 그쳤다. 이는 두 가지 요인에 기인한다. 첫째, 국내 검색어(‘시험 불안’, ‘시험 효능감’ 등)는 교육 평가 맥락에 특화되어 초기 관련성은 높았으나, 구인을 최초로 제안하거나 정의하기보다 국외에서 정립된 구인과 척도를 국내 표본에 적용하거나 실증하는 데 집중되어 포함 기준에 부합하지 않은 경우가 많았다. 둘째, 국내 문헌 중 상당수는 수학 불안과 같이 특정 교과에 국한되거나 교사 등 학습자가 아닌 집단 또는 초등학생 이하를 대상으로 하여 제외되었다. 요컨대 국내 문헌 포함이 소수에 그친 것은 연구의 질적 열위가 아니라 국내 연구가 구인의 적용 및 실증에, 국외 연구가 구인의 제안 및 정의에 특화되어 있는 연구 지형의 차이를 반영한 결과이다.



[그림 2] PRISMA 2020 흐름도

#### 4. 주제 합성

본 연구는 Thomas와 Harden(2008)의 주제 합성(thematic synthesis)에 따라 시험 지능의 개념과 하위 구인을 제시하고 가설적 이론 모델을 제안하였다. 1단계는 ‘자유 코딩(free coding)’으로, 최종 선정된 문헌 30편으로부터 연구 유형, 연구 대상, 측정 방식, 핵심 개념의 조작적 정의, 하위 요소 목록 및 정의, 주요 연구 결과, 시험 지능 구인과의 관련성에 관한 내용을 코딩하였다. 2단계는 ‘기술적 주제(descriptive themes)’ 도출로, 공통적으로 등장하는 코딩 내용 단위를 묶어 귀납적으로 범주화하였다. 3단계는 ‘분석적 주제(analytical themes)’ 도출로, 기술적 주제들에 연구자의 이론적 해석을 결합하여 시험 지능의 구인과 개념을 제시하고, 시험 지능에 대한 이론적 모델을 제안하였다.

## IV. 연구 결과

### 1. 문헌 선별 결과

PRISMA 2020 기준에 따라 선별된 최종 분석은 총 30편으로, 해외 문헌 24편은 주로 시험 수행과 관련한 이론을 최초로 정의하거나 이론과 관련한 척도를 개발하는 연구에, 국내 문헌 6편은 개발된 척도를 한국 사정에 맞게 번안하거나 양적으로 실증하는 연구에 집중되어 있었다. <표 1>에 문헌 선별 결과를 요약하여 나타내었으며, 선별된 문헌에 대한 질 평가는 <부록 1>에 수록하였다. 본 연구의 목적은 어떠한 구인의 효과 크기를 추정하는 것이 아니라 신생 구인을 새롭게 정의하고 개념화하는 데 있으므로, 문헌의 질 평가 결과를 기준으로 최종 문헌에의 포함 여부를 결정하지 않았으며, 연구 결과를 해석하는 데 있어 각 문헌의 중요도와 신뢰도의 근거로만 활용하였다.

<표 1> 문헌 선별 결과

| 구인         | 개념           | 학자                      | 주요 내용                                                                                                         |
|------------|--------------|-------------------------|---------------------------------------------------------------------------------------------------------------|
| 인지·<br>동기적 | 인지적 평가       | Folkman & Lazarus(1985) | 학습자가 시험 상황을 위협으로 평가할수록 인지적 간섭과 불안이 증가하여 역량 발현이 저해되고, 도전으로 인식할수록 적극적인 대처가 가능해짐을 실증함.                           |
|            |              | Peacock & Wong(1990)    |                                                                                                               |
|            | 통제-가치 이론     | Pekrun(2006)            | 과제에 대한 통제 가능성과 가치 평가의 상호작용이 정서를 매개하여 수행 결과에 영향을 미침. 특히, 평가 상황은 학습자의 통제 가능성을 극단적으로 위협하여 시험 불안과 같은 부정적 정서를 유발함. |
|            | 시험 수행 자기 효능감 | Nielsen 외(2019)         | 평가 상황을 성공적으로 통제할 수 있다는 주관적 신념으로, 시험 불안으로부터 성취도가 하락하는 것을 차단하는 강력한 보호 기제로 동작함.                                  |
|            | 학업 부력        | Martin & Marsh(2008)    | 학업 부력이 높을수록 시험 상황에서의 좌절과 압박을 극복 가능한 자극으로 인식하고 빠르게 회복할 수 있음.                                                   |
|            |              | Putwain 외(2012)         |                                                                                                               |
|            | 메타인지 모니터링    | Kruger & Dunning(1999)  | 시험 진행 중 자신의 인지 상태와 문항 난이도를 실시간으로 모니터링하는 역량으로, 자신의 상태를 객관적으로 점검하고 문항에 투입되는 자원을 최적화하여 수행 성공력을 높임.               |
|            |              | Coutinho 외(2020)        |                                                                                                               |
|            |              | Hacker 외(2000)          |                                                                                                               |
|            |              | 김남희 외(2024)             |                                                                                                               |

| 구인                  | 개념    | 학자                     | 주요 내용                                                                                                                      |
|---------------------|-------|------------------------|----------------------------------------------------------------------------------------------------------------------------|
| 정서<br>반응<br>및<br>조절 | 시험 불안 | Libert & Morris(1967)  | 시험 불안을 크게 '걱정(worry)'과 '신체적 각성(emotionality)'의 정서 반응으로 구분하여 설명함.                                                           |
|                     |       | Hembree(1988)          |                                                                                                                            |
|                     |       | von der Embse 외(2018)  |                                                                                                                            |
|                     |       | Abbasi & Ghosh(2020)   | 시험 불안이 작업 용량을 고갈시켜 정상적인 인출을 방해하고 수행 능력을 저하시킴.                                                                              |
|                     |       | 임신일, 박병기(2013)         |                                                                                                                            |
|                     |       | 한채연(2026)              |                                                                                                                            |
|                     | 성취 정서 | Pekrun 외(2002)         | 시험 상황에서 발생하는 다양한 정서를 긍정 활성화/비활성화, 부정 활성화/비활성화로 분류함.                                                                        |
|                     |       | Pekrun 외(2011)         |                                                                                                                            |
|                     |       | 도승이 외(2011)            |                                                                                                                            |
|                     | 정서 조절 | Schutz & Davis(2000)   | 시험 상황에서 발생하는 부정적 정서를 관리하고 조절하여 평정심을 유지하는 실행 제어 능력으로, 부정적 정서를 인지적 재평가 전략을 통해 제어함으로써 인지 자원을 온전히 문제 해결에만 활용할 수 있도록 하는 것이 중요함. |
|                     |       | Schutz 외(2004)         |                                                                                                                            |
|                     |       | 오지은 외(2011)            |                                                                                                                            |
|                     |       | 최현정, 정남운(2021)         |                                                                                                                            |
| 전략적<br>행동           | 시험 지혜 | Millman 외(1965)        | 문항 내 힌트 활용 등을 통해 출제자의 의도를 정확히 파악하여 최적의 점수를 이끌어내는 복합적인 능력임.                                                                 |
|                     |       | Rogers & Bateson(1991) |                                                                                                                            |
|                     |       | Rogers & Yang(1996)    |                                                                                                                            |
|                     | 시험 전략 | Dodeen(2015)           | 제한된 시간 내에 시간 배분, 문항 스킵, 소거법 등을 활용하여 시험 실행력을 향상시킬 수 있음. 이는 훈련을 통해 향상 가능한 가변적 특성을 지님.                                        |
|                     |       | Otoun 외(2015)          |                                                                                                                            |
|                     |       | Stenlund 외(2017)       |                                                                                                                            |

## 2. 주제 합성 결과: 시험 지능의 4개 구인 도출

1단계 자유 코딩에서는 총 108개의 의미 있는 단위 코드가 생성되었다. 이후 단위 코드들을 반복적으로 비교·분석하는 과정에서 동일한 현상을 다른 용어로 명명하거나 개념적으로 포함 관계에 있는 코드 25개를 정제하여 83개의 단위 코드를 확정하였다. 2단계 기술적 주제 도출 단계에서는 단위 코드의 내용 유사성에 따라 반복적으로 등장하는 공통적인 내용 단위를 귀납적으로 범주화하여 12개의 하위 요소를 도출하였고, 3단계 분석적 주제 도출 단계에서는 이를 상위 차원으로 묶어 4개 구인으로 수렴하였다. <표 2>는 시험 지능의 4개 구인이 도출되는 과정을, <표 3>은 최종 도출된 시험 지능의 4개 구인과 12개 하위 요소를 나타낸 것이다.

&lt;표 2&gt; 주제 합성 과정(대표 단위 코드만 표기)

| 1차<br>(단위 코드 생성)                      | 2차<br>(기술적 주제 도출) | 3차<br>(분석적 주제 도출) |
|---------------------------------------|-------------------|-------------------|
| 2차 평가, 통제 평가, 시험 자기 효능감, 학업 부력        | 효능감               | 상황 판단             |
| 1차 평가, 가치 평가, 중심성                     | 중요성               |                   |
| 메타인지 정확도, 모니터링 정확도, 자기-타인 인식          | 객관성               |                   |
| 긍정 활성화(즐거움, 희망, 자부심), 긍정 비활성(안도)      | 긍정 정서 인식          | 정서 인식             |
| 시험 불안, 분노, 수치심                        | 부정 활성화 정서 인식      |                   |
| 지루함, 절망                               | 부정 비활성 정서 인식      |                   |
| 과제 초점 유지, 과제 초점 재집중                   | 과제 초점 유지          | 정서 조절             |
| 정서 초점 과정, 자기 구식, 지연 행동, 회파-분산 양식      | 정서 초점 대처          |                   |
| 인지적 재평가 과정, 완벽주의-정서 조절 경로             | 인지적 재평가           |                   |
| 시간 활용 전략, 시간 배분, 검토 시간 확보             | 시간·순서 관리          | 행동 전략             |
| 오류 회피, 추측, 소거법, 단서 활용 전략              | 문항 풀이 전략          |                   |
| 지시문 꼼꼼히 읽기, 공란 남기지 않기, 남은 시간 답 바꾸지 않기 | 오류 방지             |                   |

&lt;표 3&gt; 시험 지능 4개 구인

| 구인       | 개념                                           | 하위 핵심요소                                                  |
|----------|----------------------------------------------|----------------------------------------------------------|
| 상황<br>판단 | 학습자가 시험<br>상황을 인지적으로<br>평가하는 역량              | • <b>효능감</b> : 시험 상황을 잘 통제하며 성공적으로 수행해 낼 것이란 스스로의 믿음     |
|          |                                              | • <b>중요성</b> : 시험의 가치를 중요하게 판단하는 것                       |
|          |                                              | • <b>객관성</b> : 자신의 학습 상태와 시험 준비 수준을 정확하게 평가하는 것          |
| 정서<br>인식 | 시험 상황에서<br>발현되는 자신의<br>정서를 정확히<br>지각하는 역량    | • <b>긍정 정서 인식</b> : 즐거움, 희망, 자부심, 안도                     |
|          |                                              | • <b>부정 활성화 정서 인식</b> : 불안, 분노, 수치심                      |
|          |                                              | • <b>부정 비활성화 정서 인식</b> : 지루함, 절망                         |
| 정서<br>조절 | 학습자가 활성화된<br>정서 반응을<br>의도적으로 관리<br>및 조절하는 역량 | • <b>과제 초점 유지</b> : 시험 중 오직 문제 해결에 인지 자원을 집중하는 것         |
|          |                                              | • <b>정서 초점 조절</b> : 자신의 감정 상태를 모니터링하여 평정심을 되찾는 것         |
|          |                                              | • <b>인지적 재평가</b> : 시험에 대한 해석을 수행에 유리한 상태로 조절하는 것         |
| 행동<br>전략 | 시험에서 높은<br>점수를 획득할 수<br>있도록 돕는<br>구체적인 전략    | • <b>시간·순서 관리</b> : 난이도 별 우선순위를 매겨 자원을 최적화하는 것           |
|          |                                              | • <b>문항풀이 전략</b> : 출제자의 의도를 정확히 파악하여 문항을 해결하는 것          |
|          |                                              | • <b>오류 방지</b> : 지문 오독, 마킹 실수 등에서 비롯된 불필요한 점수 하락을 방지하는 것 |

## 가. 제 1구인: 상황 판단

### 1) 구인 도출의 이론적 근거

Lazarus와 Folkman(1985)의 인지적 평가(cognitive appraisal) 이론에 따르면, 학습자의 수행 성공 여부는 ‘학습자가 과제를 어떻게 평가하는가?’에 좌우된다. 인지적 평가는 상황을 위협, 도전, 손실로 판단하는 1차 평가와 대처 가능한 자원을 추정하는 2차 평가로 구분되는데, 그 결과에 따라 불안, 분노, 죄책감, 슬픔 등의 정서가 유발되어 과제 수행에 영향을 미친다(Folkman & Lazarus, 1985; Peacock & Wong, 1990). Peacock와 Wong(1990)은 학습자가 시험을 위협이 아닌 도전으로 인식하고 그 가치를 중요하게 여기며 자신이 잘 해낼 능력이 있다고 판단할수록 성공 확률이 높아짐을 실증하였다. Pekrun(2006)의 통제가치 이론(Control-Value Theory)에서는 시험 상황에서 학습자의 ‘통제 평가(내가 시험을 통제할 수 있는가?)’와 ‘가치 평가(시험 결과가 나에게 중요한가?)’가 성취 정서 유형을 결정하며, 두 평가가 높을수록 긍정적 정서가 유발된다고 보았다.

Nielsen 외(2019)는 학업 전 시험 자기 효능감(Pre-Academic Exam Self-Efficacy, PAE-SE) 개념을 통해 시험을 성공적으로 수행할 것이란 학습자의 신념이 수행을 강화하는 근거가 될 수 있음을 주장하였다. 한편, Martin과 Marsh(2008)의 학업 부력(academic buoyancy) 개념은 학업에서의 좌절과 시험 압박을 효과적으로 극복하는 학습자의 능력을 의미하는데, 이는 시험 불안의 걱정 요인을 억제함으로써 수행 수준을 높이는 역량이다(Putwain, 2012). 이외에도 메타인지 모니터링(metacognitive accuracy) 연구들은 학습자가 자신의 학습 상태나 준비 수준을 객관적으로 판단할수록 우수한 성취를 나타내고(Coutinho 외, 2020; Hacker 외, 2000; Kruger & Dunning, 1999), 학습자가 스스로를 과대 추정하거나 과소 평가할 경우 정서 인식 및 조절 역량이 왜곡되어 결과적으로 수행에 부정적인 영향을 미친다는 사실을 밝혀내었다(김남희 외, 2024).

### 2) ‘상황 판단’ 구인의 핵심 하위 요소

인지적 평가 관련 연구들은 ‘학습자가 시험 상황을 스스로 어떻게 평가하는가?’에 따라 정서 반응 유형이 결정되고, 이것이 시험 수행에 영향을 미친다는 사실을 일관되게 보고하고 있다. 이러한 내용들을 종합해볼 때, 본 연구는 시험에 대한 학습자의 ‘상황 판단’이 시험 지능의 구인이 될 수 있다고 보았다. 상황 판단의 첫 번째 하위 요소로는 ‘효능감’을 설정하였다. 효능감은 Lazarus & Folkman(1985)의 2차 평가, Pekrun(2006)의 통제 평가, Nielsen 등(2019)의 시험 자기 효능감, Martin와 Marsh(2008)의 학업 부력 개념을 수렴한 것으로, 학습자가 상황을 통제하며 시험을 잘 수행해낼 것이란 스스로의 믿음을 의미한다. 효능감은 긍정 정서를 유발하여 시험을 성공적으로 수행할 수 있게 하므로 시

험 지능의 주요한 하위 요소가 될 수 있다고 보았다. 다음으로 ‘중요성’을 하위 요소로 설정하였다. 이는 Lazarus와 Folkman(1985)의 1차 평가 중심성, Pekrun(2006)의 가치 평가 개념을 수렴한 것으로, 학습자가 시험의 가치를 중요하게 판단하는 지 여부를 나타낸다. 이는 곧 정서 반응의 강도와 동기 수준에 영향을 미치기 때문에 시험 수행 역량 체계에 부합한다고 보았다. 마지막으로 ‘객관성’을 하위 요소로 설정하였다. 객관성은 Kruger와 Dunning(1999)의 메타인지 정확도, Hacker 외(2000)의 모니터링 정확도, 김남희 외(2024)의 자기-타인 인식 개념을 수렴한 것으로, 학습자가 자기 자신의 학습 상태와 시험 준비 수준을 정확하게 평가하는 정도가 시험 수행에 유의미한 차이를 만들어낼 수 있으므로 시험 수행 능력의 근간이 될 수 있다고 보았다.

## 나. 제 2구인: 정서 인식

### 1) 구인 도출의 이론적 근거

학습자는 시험이라는 제한된 시·공간적 제약 속에서 다양한 정서를 경험하며, 이는 인지 자원을 활용하는 방식에 영향을 미쳐 지식과 기능의 발현 정도를 좌우한다. Libert와 Morris(1967)는 평가 상황에서 발현되는 시험 불안(test anxiety)을 인지적 요소인 ‘걱정(worry)’과 생리적·정서적 요소인 ‘각성(emotionality)’으로 구분하여 최초로 개념화하였다. 이후 연구들은 시험 수행을 저하시키는 주된 요소가 걱정 요소임을 일관되게 확인하였으며(Hembree, 1988; von der Embse 외, 2018; 임신일, 박병기, 2013), 시험 불안이 뇌의 작업기억(working memory) 용량을 잠식하여 정보 처리 능력을 저해한다고 설명하였다(Abbasi & Ghosh, 2020). Abbasi와 Ghosh(2020)는 시험 불안을 신체적·인지적·정서적·행동적 반응의 네 가지 차원으로 구조화하였고, Morales-Martinez 외(2025)는 학습자가 시험 유형, 난이도, 목표, 응시 방식 등의 정보를 통합적으로 판단한 결과로 시험 불안이 형성되므로 정서는 인지적 판단의 결과적 산물이라고 주장하였다. 국내에서는 한채연(2026)이 예기불안, 침투적 사고, 자기불신, 정서적 과민성, 사회적 비교, 강박적 확인 경향성, 일상 기능 저해의 7요인으로 구성된 다차원적인 시험 불안 척도를 개발하였다. 한편, 시험 상황에서 학습자가 경험하는 정서는 불안에만 국한되지 않는다. Pekrun 외(2002)는 시험 불안에 편중되어 온 학업 정서 연구를 확장하여 학업 활동에서 발생하는 다양한 정서를 성취 정서(achievement emotion)로 개념화하였고, Pekrun 외(2011)는 이를 성취 정서 질문지(Achievement Emotions Questionnaire, AEQ)로 조작화하였는데, 여기에는 시험 맥락 고유의 정서를 측정하는 하위 척도가 포함되어 있다. 국내에서는 도승이 외(2011)가 한국어판 성취 정서 질문지(K-AEQ)를 개발하고 타당화하여 시험 상황에서 발생하는 9가지 정서(즐거움, 희망, 자부심, 안도, 불안, 분노, 수치심, 지루함, 절망) 구조를 확인하였다. Pekrun(2006)의

통제-가치 이론에 따르면, 긍정 정서는 내적 동기를 강화하고 사고의 폭을 확장하여 유연한 문제 해결을 촉진하는 반면, 부정 정서는 각성 수준에 따라 다른 방식으로 학습자의 수행을 저해한다. 특히, 불안, 분노, 수치심과 같은 부정 활성화 정서는 학습자의 인지 능력을 떨어뜨리고, 지루함, 절망과 같은 부정 비활성 정서는 주의 집중 능력을 방해하거나 노력을 철회하게 만든다. 이처럼 정서 유형이 인지 자원에 관여하는 방식이 질적으로 다른 만큼, 학습자는 자신의 정서를 유형별로 정확히 인식하고 변별할 필요가 있으며 이 역시 시험 지능의 하위 요소가 될 수 있다.

## 2) '정서 인식' 구인의 핵심 하위 요소

시험 상황에서 발생하는 학습자의 정서 그 자체는 무의식적 반응이기 때문에 학습자의 역량으로 간주하기 어렵다. 다만, 앞서 살펴본 바와 같이 발현된 정서는 유형에 따라 인지 자원에 관여하는 방식이 다르므로 이를 학습자가 얼마나 정확하게 지각하고 변별하는지가 시험 수행을 위한 주요 역량이 될 수 있다. 본 연구는 학습자가 시험 상황에서 발현되는 정서를 정확히 지각하고 변별하는 역량을 '정서 인식'으로 명명하고, 이를 시험 지능의 구인으로 설정하였다. 정서 인식의 하위 요소는 학습자가 인식해야 할 정서의 영역을 기준으로 '긍정 정서 인식', '부정 활성화 정서 인식', '부정 비활성화 정서 인식'으로 구분하였다.

첫 번째 하위 요소는 '긍정 정서 인식'으로, Pekrun 외(2011)의 성취 정서 중 긍정 활성화 정서(즐거움, 희망, 자부심)와 긍정 비활성 정서(안도), 도승이 외(2011)의 K-AEQ 긍정 정서 구조 개념을 수렴한 것이다. 학습자가 시험 중 긍정 정서를 자각하면 자신이 안정적이고 통제 가능한 상태에 있음을 확인하여 문제 상황에 유연하게 대처할 수 있으므로 시험 지능의 주요한 요소가 될 수 있다고 보았다. 두 번째는 '부정 활성화 정서 인식'이다. 부정 활성화 정서는 시험 불안(Abbasi & Ghosh, 2020; Liebert & Morris, 1967; Hembree, 1988; Morales-Martinez 외, 2025; von der Embse 외, 2018; 임신일, 박병기, 2013, 한채연, 2026), 분노와 수치심(도승이 외, 2011)처럼 높은 각성 상태에서 인지 자원을 잠식하는 정서로, 학습자가 이를 정확히 자각함과 동시에 일시적인 것으로 받아들이면 인지 자원이 완전히 잠식되는 상황을 막을 수 있어 시험 지능 체계에 포함될 수 있다고 보았다. 세 번째는 '부정 비활성화 정서 인식'이다. Pekrun 외(2011)가 제시한 지루함, 절망과 같은 정서적 무력감은 학습자로 하여금 시험 과제에 온전히 집중하지 못하고 노력할 이유를 찾지 못하게 하지만, 자신의 무력감을 알아차리는 것만으로도 그 원인과 해결 방안을 찾을 동력을 얻을 수 있으므로 이 역시 시험 지능의 하위 요소가 될 수 있다고 보았다.

## 다. 제 3구인: 정서 조절

### 1) 구인 도출의 이론적 근거

시험 상황에서 발생한 부정 정서는 학습자의 의도적 노력과 관리를 통해 수행에 유리한 상태로 전환될 수 있다, Schutz와 Davis(2000)는 시험 응시 중 발생하는 정서 반응을 학습자가 인식하고 이를 과제 수행에 적합한 상태로 관리하는 의도적 과정을 ‘시험 중 정서 조절(Emotional Regulation during Test-taking, ERT)’로 개념화하였으며, Schutz 외(2004)는 이를 인지적 평가 과정, 과제 초점 과정, 정서 초점 과정, 과제 재초점 과정의 네 요인으로 조작화하여 척도로 타당화하였다. 이 연구들은 학습자가 평가 상황을 위협이 아닌 도전으로 전환하여 받아들일 경우, 불안 등의 부정 활성 정서의 강도가 낮아져 본래의 수행 수준을 지켜낼 수 있게 되고, 과제에 온전히 집중하는 과정이 평정심을 회복시켜 수행에 긍정적으로 작용한다는 사실을 실증하였다. 또한, 학습자가 자신의 정서에만 지나치게 몰입하는 과정은 수행에 부정적 영향을 미치지만, 부정적 정서를 억압하거나 회피하지 않고 짧게 표출하거나 수용 가능한 형태로 전환하면 정서가 작업 기억을 점유하는 것을 막아 인지 자원을 다시 과제로 되돌릴 수 있다는 사실을 밝혀내었다. 국내 연구 또한 정서 조절 방식이 수행과 밀접하게 관련됨을 보고하였는데, 오지은(2011)은 남들에게 무능하게 보이고 싶지 않은 동기(수행 회피 목표 지향)가 강한 학습자일수록 시험 상황에서 불안을 더 높게 경험하고 지연 행동을 보이거나 핑계 조건을 만들어 두는 자기방어적·부적응적 전략을 더 많이 사용한다는 점을 밝혔고, 최현정과 정남운(2021)은 평가 염려 완벽주의가 회피-분산적 정서 조절 양식을 매개로 시험 불안을 증폭시킬 수 있음을 규명하여 정서 조절의 방식에 따라 시험 불안의 발현이 달라짐을 보여주었다.

### 2) ‘정서 조절’ 구인의 핵심 하위 요소

정서 조절 관련 연구들은 학습자가 시험 중 발현된 정서를 의도적으로 관리하고 조절하는 역량이 학업 성취도에 긍정적인 영향을 미친다는 사실을 일관되게 보고하고 있다. 이에 본 연구는 ‘정서 조절’을 시험 지능의 구인으로 설정하고 ‘과제 초점 유지’, ‘정서 초점 조절’, ‘인지적 재평가’를 핵심 하위 요소로 제안하였다. 첫 번째 하위 요소는 ‘과제 초점 유지’로, Schutz와 Davis(2000)의 시험 중 과제 초점 과정과 Schutz 외(2004)의 과제 초점 유지 개념을 수렴한 것이다. 이는 시험 중 오직 문제를 해결하는 본래의 목표에 인지 자원을 집중하는 행동을 의미하는데, 이는 인지적 자원이 다른 곳으로 분산되는 것을 막고 수행을 성공적으로 이어갈 수 있게 하므로 시험 지능의 주요한 요소가 될 수 있다고 보았다. 두 번째는 ‘정서 초점 조절’로, Schutz와 Davis(2000)의 정서 초점 과정과 Schutz 외(2004)의 정서 초점 처리, 그리고 오지은(2011)의 자기 구실, 지연 행동과 최현정과 정

남운(2021)의 회피-분산적 정서 조절 양식 개념을 수렴한 것이다. 학습자가 자신의 정서에 과도하게 몰입하지는 않되 부정적 정서를 억압하거나 회피하지 않고 짧게 표출하거나 수용 가능한 형태로 전환하려는 노력은 인지적 자원이 과제에 집중될 수 있게 하여 수행에 긍정적 영향을 미치므로 시험 지능 체계에 포함될 수 있다고 보았다. 세 번째는 ‘인지적 재평가’로, Schutz와 Davis(2000)와 Schutz 외(2004)의 인지적 재평가 과정 개념을 수렴한 것이다. 이는 시험 상황에 대한 해석을 수행에 유리한 방향으로 조절하는 능력으로, ‘이번 시험을 망치면 큰일난다’ 등의 위협적 사고를 ‘나의 실력을 증명하기 위한 하나의 도구일 뿐이다’로 재해석하는 것이 그 예이다. 이처럼 과제를 도전적 평가로 전환하게 되면 부정 활성 정서의 강도가 낮아져 작업 기억이 덜 소모되고, 학습자가 과제에 대한 동기와 집중을 유지함으로써 본래의 수행 수준을 지켜낼 수 있게 된다. 따라서 이 역시 시험 지능의 하위 요소가 될 수 있다고 보았다.

## 라. 제 4구인: 행동 전략

### 1) 구인 도출의 이론적 근거

행동적 차원의 연구들은 시험 상황에서 학습자가 수행을 향상시키기 위해 사용하는 전략과 관련된다. Millman 외(1965)는 시험의 형식적 특성과 시험 상황을 활용하여 교과 지식과는 독립적으로 높은 점수를 획득하는 능력으로서 ‘시험 지혜(test wiseness)’ 개념을 정의하고, 이를 시험의 형식이나 출제자의 의도와 무관한 전략(시간 활용, 오류 회피, 추측, 오답을 제거하는 소거법)과 출제자의 의도나 문항 형식에 의존하는 단서 활용 전략(‘항상·결코·모두’와 같은 한정어, 어간과의 문법적 정합성, 어간과 선택지의 어휘적 유사성, 선택지의 길이·정교성, 유사하거나 정반대인 선택지 쌍, 부조리한 선택지, 정답의 위치 등)의 두 축으로 분류하였다. Rogers와 Bateson(1991)은 고부담 시험 문항에 시험 지혜가 개입할 여지가 있음을 확인하고 시험 지혜 검사를 개발하였으며 시험 지혜 훈련을 정규 중등 교육과정에 포함할 것을 권고하였다. Rogers와 Yang(1996)은 여기서 나아가, 시험 지혜가 학습자의 관련 부분 지식(relevant partial knowledge)과 결합될 때 충분히 발휘될 수 있다는 수행 모형을 제시하며 시험 지혜가 교과 지식과 독립적이며 부분 지식과 상호작용하는 메타 수준의 역량임을 제시하였다.

Dodeen(2015)은 시험 지혜와 유사한 ‘시험 전략(test-taking strategies)’ 개념을 제시하고 시험 문항 예측, 지시문 정독, 전체 시험 훑어보기, 시간 관리, 선다형 문항 응답, 추측, 공란 남기지 않기, 어려운 문항 처리, 시험 후 복기의 아홉 가지로 세분화하였으며, 이를 교수하는 것이 학습자의 시험 점수를 향상시키고 시험 불안과 태도를 개선할 수 있음을 검증하였다. Otoum 외(2015)는 시험 지혜 전략을 응시 시점(답변 전·중·후)에 따라 구조

화하여 측정하였으며, Stenlund 외(2017)는 표준화 시험에서 학습자가 지식을 최대한 활용하는 확장적 전략과 무작위 추측처럼 지식을 우회하는 제한적 전략을 구분하고, 이러한 전략의 사용이 성취 수준과 집단에 따라 체계적으로 다르게 나타남을 확인하였다. 이처럼 행동 전략은 교과 지식과는 별개로 시험 수행 결과를 좌우하며, 시험 상황에서 관찰 및 훈련 가능한 구체적 행동으로 조작화된다는 점에서 시험 지능의 구인이 될 수 있다.

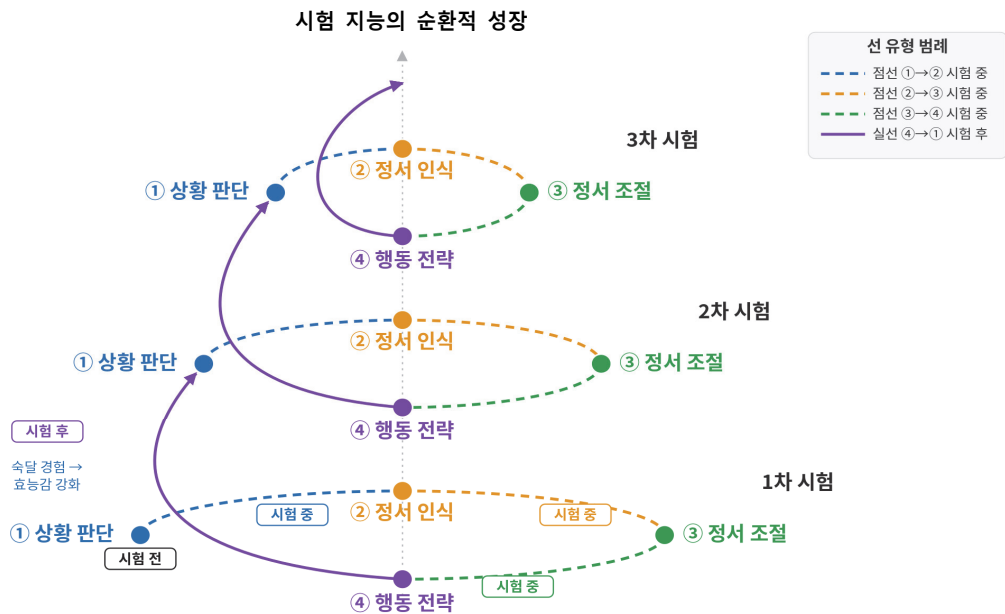
## 2) '행동 전략' 구인의 하위 요소

시험 전략 관련 연구들은 공통적으로 시험에서 높은 점수를 얻을 수 있도록 돕는 구체적인 전략을 제시하고 있다. 이에 본 연구는 '행동 전략'을 시험 지능의 구인으로 제시하고, '시간·순서 관리', '문항풀이 전략', '오류 방지'를 핵심 하위 요소로 설정하였다. 첫 번째 하위 요소는 '시간·순서 관리'로, Millman 외(1965)의 시간 활용 전략, Dodeen(2015)의 시간 관리, Otoum 외(2015)의 난이도에 따른 시간 배분, 검토 시간 확보, 그리고 Stenlund 외(2017)의 시간 관리 행동을 수렴한 것이다. 이는 학습자가 제한된 시간 내에 우선순위를 정하여 자원을 효율적으로 배분함으로써 시간 부족으로 아는 문제도 놓치는 상황을 예방하는 능력에 해당하므로 시험 지능의 주요한 요소가 될 수 있다고 보았다. 두 번째는 '문항풀이 전략'으로, Millman 외(1965)의 오류 회피, 추측, 소거법, 단서 활용 전략, Dodeen(2015)의 시험 문항 예측 및 어려운 문항 처리 방법, Otoum 외(2015)의 핵심어 밀줄, 출제자의 단서를 활용한 추측, 그리고 Stenlund 외(2017)의 추측, 문항 건너뛰기를 수렴한 것이다. 이는 출제자의 의도를 정확히 파악하여 최선의 답을 선택하는 전략으로, 교과 지식이 완벽하지 않더라도 정답률을 높일 수 있는 학습자의 실전 역량이기에 시험 지능 체계에 포함될 수 있다고 보았다. 세 번째 '오류 방지'는 지문 오독, 마킹 실수, 성급한 판단 등을 막기 위한 검토 행위로, Dodeen(2015)의 지시문 꼼꼼히 읽기, 공란 남기지 않기, 시험 후 복기, 그리고 Otoum 외(2015)의 답하기 전 문항을 꼼꼼히 읽기, 모든 문항의 응답 여부 점검, 답안 제출 전 인적 사항 재확인, 남은 시간에 답을 함부로 바꾸지 않기 등을 수렴한 것이다. 이는 학습자의 부주의를 예방하고 점검하여 불필요한 점수 하락을 방지하는 역량이므로 이 역시 시험 지능의 하위 요소가 될 수 있다고 보았다.

## 3. 시험 지능의 이론적 모델

본 연구는 체계적 문헌 고찰과 주제 합성 결과로 도출된 4개 구인을 중심으로 시험 지능을 '학습자가 시험 상황에서 경험하는 다양한 인지적·정서적 반응을 효과적으로 인식 및 조절하고, 행동 전략을 통해 최적의 수행을 이끌어내는 종합적 실행 역량'으로 정의하

고자 한다. 일반적인 지능(IQ)이 ‘학습자가 무엇을 알고 있는가?’를 측정한다면, 시험 지능(TI)은 ‘학습자가 지닌 지식과 기술을 시험 상황에서 얼마나 효과적으로 인출할 수 있는가?’를 측정하는 실천적 지표라 할 수 있다. 다시 말해, 시험 지능은 학습자가 보유한 잠재 역량과 실제 시험에서 발현되는 수행 역량 사이의 간극에 관여하는 심리적 기제를 통합적으로 조망하는 개념인 것이다. [그림 3]은 본 연구에서 제안하는 시험 지능에 관한 이론적 모델을 시각화한 것이다.



[그림 3] 시험 지능의 이론적 모델

본 연구가 제안하는 시험 지능의 이론적 모델에서 학습자의 상황 판단, 정서 인식, 정서 조절, 행동 전략 역량은 시간적 흐름(시험 전·중·후)에 따라 순차적으로 활성화되며, 그 결과가 다음 시험 주기로 환류되는 순환적(cyclic) 구조를 이룬다. 4개 구인을 순차적으로 표현한 것은 앞선 구인이 다음 구인의 발현 방향과 질적 수준에 영향을 미친다는 문헌 고찰 결과를 반영한 것이다. 즉, 학습자가 시험 상황을 어떻게 판단하는지에 따라 발현되는 정서가 달라지고, 그 정서를 어떻게 인식하고 조절하는지에 따라 행동 전략의 효과가 달라질 수 있다. 한편, 순환적 구조는 이전 시험에서 성공적으로 발휘한 시험 지능이 학습자의 효능감을 강화시켜 다음 시험의 상황 판단을 촉진하는 상황을 반영한 것으로, 시험 지능은 평가 경험이 누적됨에 따라 지속적으로 발달할 수 있는 가변적 역량으로 이해될 수 있다.

시험 지능 모델의 가장 근원적인 역량은 ‘상황 판단’으로, 학습자가 시험 상황을 어떻게 평가하는지에 관한 인지적 과정이다. 학습자가 높은 ‘효능감’을 가질수록 시험을 도전적 과제로 인식하여 긍정 정서와 적절한 전략을 이끄는 선순환이 형성된다. 또한, 학습자가 시험의 ‘중요성’을 높게 인식할수록 자부심과 같은 긍정 정서와 동기가 강화되며, 자신의 인지 및 준비 상태를 ‘객관성’ 있게 파악할수록 과도한 불안과 과잉 확신을 억제하며 성취를 높일 수 있다. 본 연구에서 상황 판단 구인을 가장 근원적이라고 본 것은 이 역량이 이후 발현되는 정서에 직접적인 영향을 미치고, 그 정서를 어떻게 인식하고 조절하는지에 따라 수행 결과가 달라지기 때문이다. 따라서 상황 판단 역량의 강화가 시험 수행 향상을 위한 첫 걸음이 될 수 있다고 보았다.

다음으로 연결되는 시험 지능 역량은 ‘정서 인식’으로, 발현된 정서를 학습자가 정확히 지각하고 변별하는 역량이다. 불안, 분노, 수치심과 같은 부정 활성화 정서는 학습자의 작업 기억 용량을 점유하여 수행 능력을 저하시키고, 지루함, 절망 등의 부정 비활성 정서는 무력감을 증대시켜 집중력과 노력을 떨어뜨리는 반면, 희망, 자부심, 즐거움, 안도감 등의 긍정 정서는 인지적 유연성을 높여 수행을 향상시키는 데 기여한다. 이렇듯 발현된 정서에 따라 시험 수행이 전혀 다른 방향으로 진행되는 만큼, 학습자는 자신의 정서를 정확하게 인식하고 변별하는 역량을 갖추는 필요가 있다.

부정적 정서가 일단 발현되었을지라도 학습자는 ‘정서 조절’을 통해 이를 관리할 수 있다. 시험 지능의 세 번째 역량인 ‘정서 조절’은 인식된 정서를 의도적으로 관리하고 수정하는 과정을 의미한다. 학습자는 ‘과제 초점 유지’를 통해 지금 해결해야 할 문제에 모든 인지적 자원을 집중시켜 평정심을 회복할 수 있고, ‘정서 초점 조절’을 통해 자신의 감정을 적절히 표현하거나 보다 완화된 형태로 전환함으로써 감정에 지나치게 매몰되는 상황을 피할 수 있다. 나아가 학습자가 시험 상황에서 위협적인 과제를 극복 가능한 도전적인 것으로 재해석하는 ‘인지적 재평가’를 통해서도 시험 수행을 성공적으로 이어나갈 수 있다. 이렇듯 학습자는 적절한 정서 조절 방법을 스스로 터득하며 시험에 몰입할 수 있는 상태를 스스로 조성해나갈 수 있어야 한다.

시험 지능의 마지막 역량은 ‘행동 전략’이다. 이는 교과 지식과는 별개로 시험에서 높은 점수를 획득할 수 있도록 돕는 구체적인 전략을 의미한다. 이 구인이 시험 지능의 가장 마지막 단계에 위치한다고 해서 그 중요성이 낮은 것은 아니며, 오히려 행동 전략이 온전히 실행되기 위해서는 상황 판단, 정서 인식, 정서 조절에 이르는 일련의 과정이 선행되어야 함을 의미한다. 행동 전략 역량에는 우선순위를 매겨 자원 배분을 최적화하고 시간을 관리하는 ‘시간-순서 관리’, 출제자의 의도를 정확히 파악하고 단서들을 이용하여 정답률을 높이는 ‘문항풀이 전략’, 지문 오독, 마킹 실수, 성급한 판단 등을 방지하기 위한 검토 행위를 통해 불필요한 점수 하락을 방지하는 ‘오류 방지’가 포함된다.

요컨대 본 연구가 제안하는 시험 지능 모델은 상황 판단, 정서 인식, 정서 조절, 행동 전략의 4개 구인이 단일 시험 맥락 안에서 순차적으로 작동하며, 앞선 구인이 다음 구인의 발현을 조건화하는 과정적 구조를 보여준다. 여기서 상황 판단 구인은 정서의 인식과 조절, 나아가 행동 전략의 실행에 이르는 전 과정을 이끄는 근원적 동력으로 작동한다. 또한 이 구조는 단일 시험에서 마무리되는 것이 아니라, 성공적으로 발휘된 시험 지능이 학습자의 효능감을 강화하여 다음 시험의 상황 판단을 촉진하는 순환적 환류를 형성함으로써 시험 지능이 평가 경험의 누적을 통해 지속적으로 발달할 수 있는 가변적 역량임을 시사한다. 이는 학습자의 고정된 능력이나 우연적 요인에 의해 시험 결과가 결정되는 것이 아니라 교육적 개입을 통해 함양 가능한 역량의 산물일 수 있음을 함의한다.

## V. 결론 및 제언

본 연구는 PRISMA 2020 지침에 따라 선행 연구의 시험 수행 관련 내용을 체계적으로 분석한 후, 주제 합성을 통해 ‘시험 지능(Test Intelligence, TI)’이란 신생 구인의 개념과 가설적 이론 모델을 제안하였다. 주요 내용을 정리하면 다음과 같다.

첫째, 시험 수행과 관련한 ‘상황 판단’, ‘정서 인식’, ‘정서 조절’, ‘행동 전략’의 4개 구인과 12개의 핵심 하위 요소를 도출하였다. ‘상황 판단’은 학습자가 시험 상황을 인지적으로 평가하는 능력으로, 시험 상황을 잘 통제하는 가운데 성공적으로 수행할 것이란 믿음인 ‘효능감’과 시험에 부여하는 가치인 ‘중요성’, 그리고 자신의 학습 상태와 시험 준비도에 대한 메타인지적 판단인 ‘객관성’을 포함한다. ‘정서 인식’은 학습자가 시험 상황에서 발현되는 자신의 정서를 정확히 지각하고 변별하는 역량으로, ‘긍정 정서(즐거움, 희망, 자부심, 안도) 인식’, ‘부정 활성화 정서(불안, 걱정, 분노) 인식’, ‘부정 비활성 정서(지루함, 절망) 인식’으로 분류할 수 있다. 학습자는 ‘정서 조절’ 능력을 통해 발현된 정서를 의도적으로 관리할 수 있는데, ‘과제 집중 유지’ 능력을 발휘하여 시험 문제에 모든 인지 자원을 집중하거나 ‘정서 초점 조절’을 통해 자신의 감정을 적절하게 표현하거나 완화된 형태로 전환하고, ‘인지적 재평가’를 통해 위협적인 과제를 극복 가능한 도전으로 재해석함으로써 수행을 성공적으로 이어갈 수 있다. 마지막으로 ‘행동 전략’은 시험에서 높은 점수를 획득할 수 있도록 돕는 구체적인 전략으로, 학습자는 ‘시간·순서 관리’ 역량을 통해 문제 풀이의 우선순위를 정해 시간을 관리하고, ‘문항풀이 전략’을 통해 출제자의 의도를 정확히 파악하고 단서들을 적극적으로 이용하여 정답률을 높일 수 있으며, ‘오류 방지’ 능력을 발휘하여 마킹 실수나 지문 오독 등으로부터 발생하는 불필요한 점수 하락을 방지할 수 있다.

둘째, 본 연구는 도출된 4개 구인을 바탕으로 시험 지능을 ‘학습자가 시험 상황에서 경험하는 다양한 인지적·정서적 반응을 효과적으로 인식 및 조절하고, 행동 전략을 통해 최적의 수행을 이끌어내는 종합적 실행 역량’으로 개념화하였다. 시험 지능은 학습자가 지닌 잠재 역량과 시험 수행 역량의 간극을 설명하는 매개체이므로, 학습자가 시험 상황에서 자신의 잠재력을 얼마나 효과적으로 인출하는지를 설명하는 개념적 지표가 될 수 있을 것으로 기대된다.

셋째, 본 연구는 시험 지능의 순차적·순환적 구조에 관한 가설적 모델을 제안하였다. 이 모델은 시험 지능의 4개 구인이 단일 시험 내에서 순차적으로 활성화되며, 앞선 구인이 다음 구인의 발현 방향과 질적 수준에 직접적인 영향을 미치는 것으로 설명한다. 또한, 이전 시험에서 시험 지능을 성공적으로 발휘한 경험이 다음 시험 주기로 환류되어 상황 판단 요소를 한층 강화하는 순환적 구조를 형성할 수 있음을 보여준다. 이에 따르면, 시험 지능은 고정 불변의 것이 아니라 평가 경험이 누적됨에 따라 지속적으로 발달할 수 있는 가변적 역량으로 이해될 수 있다.

본 연구는 다음의 두 가지 측면에서 의의를 지닌다. 첫째, 기존에 분절적으로 논의되어 온 시험 수행 관련 구인들의 공통 내용 영역을 체계적으로 정리하여 이를 ‘시험 지능’이라는 통합적 틀로 재구조화하였다. 이는 시험 수행 과정에서 학습자가 겪게 되는 역동적인 인지·정서·행동 기제를 통합적으로 조망하는 개념적 토대를 제공한다는 점에서 의의를 지닌다. 둘째, 시험 지능을 누적적으로 발달시킬 수 있다는 가변적 역량으로 개념화한 만큼, 교육 현장에서는 학습자의 상황 판단, 정서 인식, 정서 조절, 행동 전략을 강화시킬 수 있도록 돕는 실천적 교수·학습 방안을 모색할 필요가 있다. 특히 본 연구에서는 시험 지능의 ‘상황 판단’ 구인이 정서 인식과 조절, 행동 전략에 이르는 일련의 과정을 이끄는 근원적인 구인이라고 보았는데, 이에 따르면 학습자의 시험 지능을 함양하기 위한 교육적 개입은 상황 판단 구인을 강화하는 데 우선 순위를 둘 필요가 있다. 예컨대 학습자가 시험 후 오답을 분석할 때, 단순히 교과 내용만을 확인하기보다 풀이 당시 자신의 인지적·정서적·행동적 요인들을 객관적으로 복기하도록 돕는 성찰 도구를 제공하는 방안을 고려할 수 있다. 이러한 성찰 도구를 활용한 다차원적 복기는 학습자로 하여금 수행 실패나 어려움의 원인을 지능 등의 고정적인 요인이나 운과 같은 통제 불가능한 변수가 아니라 스스로 조절하고 보완할 수 있는 유동적인 요인으로 귀인할 수 있게 한다. 이렇게 되면 학습자는 다음 평가 상황을 스스로 통제하고 개선할 수 있다는 신념을 가지게 되며, 이는 상황 판단 구인의 효능감 향상과 밀접하게 연결될 수 있다. 또한, 교사는 시험 불안을 겪는 학습자에게 단순히 긴장하지 말라는 일차원적인 조언을 건네는 것을 넘어 시험 상황에서 학습자가 자신의 감정을 어떻게 모니터링하고 조절할 수 있는지에 관한 구체적인 기법과 시험을 전략적으로 수행하기 위한 구체적인 행동 전략을 제시함으로써 학습자의 시험 지능 발달을 조력할 수 있을 것이다.

본 연구에서 도출된 시험 지능의 4개 구인과 하위 요소, 그리고 시험 지능의 이론적 모델은 문헌 분석을 통해 구성된 가설적 체계로, 경험적 검증을 거치지 않았다는 한계를 지닌다. 따라서 내용 타당도 검증을 거쳐 통계적 타당화로 이어지는 후속 연구가 수행된다면, 시험 지능 모델은 보다 객관적이고 정교한 역량 체계로 자리 매김할 수 있을 것이다.

## 참고문헌

- 김남희, 석보영, 이유락, 강이철(2024). 시험점수 예측과 학습전략 판단에서 자아 평가 유형의 차별 효과. **사고개발**, 20(3), 45-69.
- 김미영, 이창현(2023). 부적응적 완벽주의와 시험불안의 관계: 학업적 자기효능감의 매개 효과. **인문사회**21, 15(1), 43-57.
- 김정현, 김성벽, 정인경(2014). 청소년의 학업스트레스가 학교생활적응에 미치는 영향. **한국청소년연구**, 25(4), 241-269.
- 도승이, 손수경, 변준희, 임지윤(2011). 한국어판 성취정서 질문지(K-AEQ) 개발 및 타당화. **교육심리연구**, 25(4), 945-970.
- 백승희(2025). 초기 청소년들의 학업스트레스, 진로성숙도, 전반적 행복감 간의 종단적 관계. **아동과 권리**, 29(4), 122-143.
- 성태제(2019). **현대교육평가(5판)**. 학지사.
- 오지은, 추상엽, 임성문(2011). 성취목표지향성과 학업평가장면의 회피전략 간 관계: 중학생과 고등학생의 비교. **청소년학연구**, 18(11), 179-213.
- 임신일, 박병기(2013). 국내·외 시험불안 연구의 메타분석. **교육심리연구**, 27(3), 529-553.
- 정찬호, 한중규, 김원, 채정호(2004). 대학 입시 수험생의 정신건강요인과 시험 수행 능력의 관련성. **사회정신의학**, 9(2), 80-86.
- 최현정, 정남운(2021). 평가염려 완벽주의와 시험 불안의 관계에서 수행 회피목표 지향성과 회피-분산적 정서조절양식의 매개효과. **학습자중심교과교육연구**, 21(1), 1093-1117.
- 최효식, 연은모, 백근찬(2026). 청소년이 지각한 성취압력, 학업스트레스, 우울 간 관계에서 가족의사소통의 조절된 매개효과: 잠재조절 구조방정식의 적용. **한국산학기술학회지**, 27(4), 757-768.
- 한채연, 김민용, 김하영, 이민설, 이승현(2026). 대학생 시험불안 척도 개발 및 타당화 연구. **교육과학습: Global Creative Leader**, 16(1), 53-85.
- Abbasi, N., & Ghosh, S. (2020). Construction and standardization of examination anxiety scale for adolescent students. *International Journal of Assessment Tools in Education*, 7(4), 522-534.
- Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: A critical review. *BMC Medical Research Methodology*, 9(1), 1-11.

- Coutinho, M. V., Papanastasiou, E., Agni, S., Vasko, J. M., & Couchman, J. J. (2020). Metacognitive monitoring in test-taking situations: A cross-cultural comparison of college students. *International Journal of Instruction*, 13(1), 407–424.
- Dodeen, H. (2015). Teaching test-taking strategies: Importance and effects. *International Journal of Research in Education and Science*, 1(1), 6–13.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101.
- Earley, P. C., & Ang, S. (2003). Cultural intelligence: Individual interactions across cultures. *Stanford University Press*.
- Folkman, S., & Lazarus, R. S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, 48(1), 150–170.
- Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. Basic Books.
- Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92(1), 160–170.
- Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17–27.
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58(1), 47–77.
- Higgins, J. P. T., & Green, S. (Eds.). (2011). Cochrane handbook for systematic reviews of interventions (Vol. 4). *John Wiley & Sons*.
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., ... & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for information*, 34(4), 285–291.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.
- Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20(3), 975–978.
- Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an

- understanding of students' everyday academic resilience. *Journal of School Psychology*, 46(1), 53-83.
- Millman, J., Bishop, C. H., & Ebel, R. (1965). An analysis of test-wiseness. *Educational and Psychological Measurement*, 25(3), 707-726.
- Morales-Martinez, G. E., Villarreal-Lozano, R. J., & Hedlefs-Aguilar, M. I. (2025). Thinking Styles Underlying the Cognitive Judgements of Exam Anxiety amongst University Students of Engineering. *International Journal of Emotional Education*, 17(1), 21-39.
- Nielsen, T., Friderichsen, I. S., & Hartkopf, B. T. (2019). Measuring academic learning and exam self-efficacy at admission to university and its relation to first-year attrition: An IRT-based multi-program validity study. *Frontline Learning Research*, 7(3), 91-118.
- Otoun, A., Bani Khalaf, H., Bajbeer, A., & Bani Hamad, H. (2015). The level of test-wiseness for the students of arts and science faculty at Sharourah and its relationship with some variables. *Journal of Education and Practice*, 6(29), 102-113.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
- Pajares, F., & Miller, M. D. (1994). Role of self-efficacy beliefs and a short-form of test anxiety in mathematical problemsolving. *Journal of Educational Psychology*, 86(2), 193-203.
- Peacock, E. J., & Wong, P. T. P. (1990). The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine*, 6(3), 227-236.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91-105.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341.
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36-48.

- Putwain, D. W., Connors, L., Symes, W., & Douglas-Osborn, E. (2012). Is academic buoyancy anything more than adaptive coping? *Anxiety, Stress & Coping*, 25(3), 349–358.
- Rogers, W. T., & Bateson, D. J. (1991). The influence of test-wiseness on performance of high school seniors on school leaving examinations. *Applied Measurement in Education*, 4(2), 159–183.
- Rogers, W. T., & Yang, P. (1996). Test-wiseness: Its nature and application. *European Journal of Psychological Assessment*, 12(3), 247–259.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9(3), 185–211.
- Schutz, P. A., & Davis, H. A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist*, 35(4), 243–256.
- Schutz, P. A., DiStefano, C., Benson, J., & Davis, H. A. (2004). The emotional regulation during test-taking scale. *Anxiety, Stress & Coping*, 17(3), 253–269.
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., ... & Henry, D. A. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *bmj*, 358.
- Spearman, C. (1904). “General intelligence,” objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292.
- Stenlund, T., Eklöf, H., & Lyrén, P. E. (2017). Group differences in test-taking behaviour: An example from a high-stakes testing program. *Assessment in Education: Principles, Policy & Practice*, 24(1), 4–20.
- Sternberg, R. J. (1985). Beyond IQ: A triarchic theory of human intelligence. *Cambridge University Press*.
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 45.
- von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, 227, 483–493.
- Zeidner, M. (1998). Test anxiety: The state of the art. *Plenum Press*.

ABSTRACT

## Conceptualization and Proposal of a Theoretical Model for Test Intelligence: A Systematic Literature Review Based on the PRISMA 2020 Guidelines and thematic synthesis

Yu Jung Kim

Teacher, Gyeonggi Science High School

This study employed a systematic literature review and thematic synthesis following the PRISMA 2020 guidelines to analyze the common content domains of test-performance-related constructs and propose the concept and a hypothetical model of Test Intelligence (TI). First, these constructs converged into four constructs—Situational Judgment, Emotion Awareness, Emotion Regulation, and Behavioral Strategy—and twelve core sub-elements (self-efficacy, perceived value, and objectivity; awareness of positive, negative-activating, and negative-deactivating emotions; task-focusing, emotion-focused regulation, and cognitive reappraisal; and time-and-sequence management, item-solving strategy, and error prevention). On this basis, Test Intelligence was conceptualized as a comprehensive executive competence through which learners recognize and regulate the cognitive and emotional responses in test situations and elicit optimal performance via behavioral strategies. Second, the study proposed a hypothetical model in which the four constructs operate sequentially within a single test—each preceding construct influencing the next—while test intelligence exercised in one test strengthens situational judgment in the next, forming a cyclic structure. As a conceptualization study, the model requires further refinement through content and statistical validation. Nonetheless, it reconstructs hitherto fragmented test-performance constructs into an integrative framework that clarifies the dynamic cognitive, psychological, and behavioral mechanisms learners undergo in test contexts and suggests practical instructional strategies to foster the four competencies.

*Keywords:* Test Intelligence, PRISMA 2020, thematic synthesis, situational judgment, emotion awareness, emotion regulation, behavioral strategy, cognitive appraisal, control-value theory, exam self-efficacy, academic buoyancy, metacognitive monitoring, test anxiety, achievement emotions, test-wiseness, test-taking strategy

## 〈부록 1〉 문헌의 질 평가

### 1. 양적 실증 문헌의 MMAT<sup>1)</sup> 질 평가(N=21)

| 저자(연도)                   | MMAT 범주        | 표본(N)              | ① 표집 적절성 | ② 표본 대표성 | ③ 측정 적절성 | ④ 자료 완전성·무응답 | ⑤ 통계 분석 적절성 | 총족 (점수/만점) |
|--------------------------|----------------|--------------------|----------|----------|----------|--------------|-------------|------------|
| Liebert & Morris(1967)   | 양적 기술          | N=151              | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Folkman & Lazarus(1985)  | 양적 비무작위 (중단)   | N=108              | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Peacock & Wong(1990)     | 양적 기술 (척도개발)   | N=100·154·144      | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Rogers & Bateson(1991)   | 양적 비무작위 (문항분석) | 6개 시험·주별 대표 표본     | ●        | ●        | ●        | ○            | ●           | 4/5        |
| Kruger & Dunning(1999)   | 양적 비무작위 (실험)   | N=65·84·140 등      | ●        | ○        | ●        | ●            | ●           | 4/5        |
| Hacker 외(2000)           | 양적 비무작위        | N≈96               | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Martin & Marsh(2008)     | 양적 기술          | N=598 (5개교)        | ●        | ○        | ●        | ○            | ●           | 3/5        |
| 도승이(2011)                | 양적 기술 (척도개발)   | N=560→508          | ●        | ○        | ●        | ○            | ●           | 3/5        |
| 오지은(2011)                | 양적 기술 (조사)     | N=550→500          | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Pekrun 외(2011)           | 양적 기술 (척도개발)   | N=389              | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Putwain 외(2012)          | 양적 기술          | N=298              | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Schutz 외(2004)           | 양적 기술 (척도개발)   | EFA·CFA 복수표본       | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Otoun 외(2015)            | 양적 기술 (조사)     | N=299/1,982        | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Stenlund 외(2017)         | 양적 기술          | N=1,129            | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Coutinho 외(2020)         | 양적 비무작위 (실험)   | N=120/69/172       | ●        | ○        | ●        | ●            | ●           | 4/5        |
| Abbasi & Ghosh(2020)     | 양적 기술 (척도개발)   | N=2,030+402 (19개교) | ●        | ○        | ●        | ○            | ●           | 3/5        |
| 최현정·정남운(2021)            | 양적 기술 (경로분석)   | N=456 (16 제외)      | ●        | ○        | ●        | ○            | ●           | 3/5        |
| 김남희(2024)                | 양적 비무작위 (실험)   | N=153 (2학년)        | ●        | ○        | ●        | ○            | ●           | 3/5        |
| Morales-Martinez 외(2025) | 양적 비무작위 (실험)   | N=671              | ●        | ○        | ●        | ●            | ●           | 4/5        |
| 한채연(2026)                | 양적 기술 (척도개발)   | N=324+413          | ●        | ○        | ●        | ○            | ●           | 3/5        |

\* 범례: ● 충족, ○ 부분 충족, ○ 판단 보류, ○ 미충족.

양적 실증 문헌 21편에 대한 MMAT 평가 결과, 모든 문헌이 표집의 적절성(①), 측정 도구의 적절성(③), 통계 분석의 적절성(⑤)을 충족하였다. 이는 분석에 활용된 양적 실증 연구들이 적절한 표집과 타당화된 측정 도구 사용, 분석 절차를 갖추고 있음을 의미한다. 반면 표본의 대표성(②)과 자료의 완전성·무응답(④)에서는 문헌 간 편차가 나타났다. 표본 대표성의 경우, 대규모 다기관 표본을 활용한 연구가 아닌 특정 학교 학생을 대상으로 한 편의 표

1) Mixed Methods Appraisal Tool(Hong et al., 2018)

집 또는 목적 표집을 수행한 경우 ‘부분 충족’으로 평가하였다. 자료의 완전성·무응답의 경우, 종단 연구나 반복적으로 측정한 연구, 불성실한 응답을 제외한 연구는 ‘부분 충족’으로 평가하고, 응답률이나 결측 처리를 명시적으로 보고하지 않은 연구는 ‘판단 보류’로 평가하였다. 다만 어떤 문헌도 기준을 명백히 ‘미충족’한 경우는 없었으며, 전반적으로 5개 기준 중 3~4개를 충족하는 수준이었다. 이는 편의 표집과 무응답 보고의 부재가 척도·조사 연구에서 흔히 나타나는 특성인 만큼 수용 가능한 수준으로 판단할 수 있다.

## 2. 메타 분석 문헌의 AMSTER-2<sup>2)</sup> 질 평가(N=3)

| 저자(연도)                   | D2<br>프로토콜 | D4<br>검색<br>포괄성 | D7<br>제외목록<br>제시 | D9<br>개별연구<br>RoB | D11<br>통합 방법<br>적절성 | D13<br>RoB<br>해석반영 | D15<br>출판편향<br>검토 | 종합 신뢰도                    |
|--------------------------|------------|-----------------|------------------|-------------------|---------------------|--------------------|-------------------|---------------------------|
| Hembree(1988)            | ○          | ●               | ○                | ○                 | ●                   | ○                  | ○                 | 매우 낮음<br>(Critically low) |
| 임신일·박병기<br>(2013)        | ○          | ●               | ○                | ○                 | ●                   | ○                  | ○                 | 매우 낮음<br>(Critically low) |
| von der Embse<br>외(2018) | ●          | ●               | ●                | ●                 | ●                   | ●                  | ●                 | 보통<br>(Moderate)          |

\* 범례: ● 충족, ● 부분 충족, ● 판단 보류, ○ 미충족. 종합 신뢰도는 7개 핵심 영역(D2·D4·D7·D9·D11·D13·D15)의 충족 정도로 산정함(High/Moderate/Low/Critically low).

메타분석 3편에 대한 AMSTAR-2 평가에서는 문헌 간 신뢰도 편차가 뚜렷하게 나타났다. 상대적으로 최근에 수행된 von der Embse 외(2018)의 연구는 ‘보통’ 수준으로 평가된 반면, Hembree(1988) 및 임신일과 박병기(2013)의 연구는 최신 보고 기준을 충족하지 못하여 ‘매우 낮음’으로 평가되었다. 이는 두 연구가 관련 보고 지침이 정착되기 이전에 수행된 데 기인한 것으로, 연구의 개념적 기여 자체가 무효함을 의미하지는 않는다. 실제로 이들 메타분석을 통해 드러난 시험 불안과 학업성취 간의 부적 관계는 von der Embse 외(2018) 및 다수의 실증 연구에 의해서도 확인된 바 있으므로, AMSTER-2 질 평가에서 낮은 신뢰도를 나타낸 것이 해당 구인에 대한 결론을 좌우하지는 않는다고 판단할 수 있다.

## 3. 이론 문헌의 질 평가(N=6)

Dodeen(2015), Millman 외(1965), Pekrun(2006), Rogers & Yang(1996), Schutz & Davis(2000), Pekrun 외(2002), Pekrun (2006)의 연구는 경험적 자료를 산출하지 않는 이론 문헌이기에 형식적 질 평가 대상에서 제외하였다.

2) A Measurement Tool to Assess systematic Reviews-2(Shea et al., 2017)