



전공능력진단도구 개발 및 표준화에 관한 연구* - A 대학 교직과 사례를 기반으로 한 간접평가에서 수행형 직접평가 문항으로의 전환 연구를 중심으로

허지원¹

《 요 약 》

본 연구는 대학의 전공능력진단도구가 대부분 자기점검형의 간접평가 방식으로 이루어지고 있는 점을 문제로 지적하고, 그 대안으로 수행형 문항을 적용한 직접평가 방식의 전공능력 진단도구를 개발하고자 한 연구이다. 이를 위해 사범계열 및 교직과를 연구대상으로 하여, 선행연구 및 사례분석 결과와 전문가인터뷰, 자문위원회, 전공 배경에 대한 이해를 기반으로 예비문항을 개발하고, 예비검사를 실시하였다. 예비문항은 안면타당도를 검토하기 위해 사범계열 및 교직이수 학생에게 검사를 실시한 후, 그 결과를 반영하여 문항을 수정·보완한 후, 내용 전문가에게 내용타당도를 평가하도록 하였다. 그 결과 내용타당도 지수(CVI)는 .96으로 높게 나타났으며, 6개 수행형 문항 중 CVI 1.0이 5개 문항(83.3%), CVI .87이 1개 문항(16.7%)이었다. 이렇게 개발된 직접평가를 위한 수행형 문항을 자기점검형 문항으로 구성되었던 기존 검사와 관련성 비교를 하였고 이를 통해 의미 있는 시사점을 도출하고자 하였다. 이상의 연구 결과를 통해 학생들의 전공능력을 보다 정확히 측정하기 위한 방법으로 수행형 직접평가방식의 평가를 제안하며, 일련의 연구 과정을 표준 매뉴얼로 제시하였다. 이 연구의 결과가 다양한 전공에서 직접평가 방식의 문항 개발 시 활용 가능한 자료가 되기를 기대한다.

주제어 : 전공능력진단도구, 직접평가, 간접평가, 수행형 문항, 교직역량진단도구

1. 남부대학교 교수, hjw123@nambu.ac.kr

I. 서론

지식기반사회로 대변되는 4차 산업혁명 시대에 탄력적으로 대응하고, 미래사회를 선도할 수 있는 인재를 배출하기 위해 많은 대학들은 교육과정을 핵심역량을 기반으로 한 역량기반 교육과정으로 대대적인 개편을 시도하고 있다. 이는 직업 현장에서 필요로 하는 역량을 지닌 인재 배출의 요구에 비해 그동안의 대학 전공교육과정이 현장의 수요와 요구를 제대로 반영하지 못하고 있다는 자성과 평가에서 기인한다 할 수 있다(교육부, 2019). 이러한 사회의 변화와 실질적 수요를 반영하여 대학들은 교육과정 개편을 시도하고 있지만 대부분의 대학들은 형식적인 수준에서의 교육과정 개편에 그치고 있는 실정이다. 즉, 대학은 대학 단위에서의 기본적인 역량과 정의를 제시하고, 학과에서는 학과의 특성을 반영하여 전공능력을 규정하는 유연성이 요구되에도 불구하고, 많은 대학들이 전공능력을 설정하는데 미흡한 것으로 나타났다(최현철, 김수동, 김경표, 2020). 방법적인 면에서도 대학마다 다른 절차에 따라 전공능력을 설정하고 있으며, 반드시 필요한 절차를 생략하거나 간소화함으로써 역량기반 교육과정을 시행하는 대학 간에도 교육의 효과와 질적인 격차가 발생하고 있다(이화영, 2021).

‘무엇을 아는가’에서 사회 적응을 위해 실제로 ‘무엇을 할 수 있는가’가 중요한 사회로 변화하였고, 이러한 사회변화의 요구를 반영해 2015 개정 교육과정에서는 그동안 문서상의 논의로만 있어 왔던 역량중심의 교육과정을 실제 국가수준 교육과정에 반영하였다. 이 파장은 고등교육기관인 대학에까지 급속도로 영향을 미쳤으며, 대학의 교육목표와 인재상에 맞는 핵심역량을 설정하고 이를 구현하기 위해 앞서 방법상의 미비점을 지적하기는 하였으나, 많은 대학들은 발 빠르게 전공능력과 관련한 하위역량을 설정하고 그에 맞는 교육과정을 개편하였다(허지원, 2021). 따라서 대학들은 자체적으로 시행한 교육과정 개편의 효과를 검증하기 위해 핵심역량진단도구 및 전공능력진단도구 등의 학습능력을 진단하기 위한 평가도구들을 개발하게 되었다.

하지만 이와 관련된 연구들(김동일 외, 2009; 김문경 외, 2021; 김연희 외, 2010; 박수정, 송영수, 2016; 이경화 외, 2011; 성민경, 2015; 손원빈, 최화숙, 2019; 손진아 외, 2021)을 분석하였을 때, 이들 연구에서 개발된 진단도구는 대부분 자기점검형의 간접적 평가방식을 띄고 있었다. 최근 들어 다양한 전문가 그룹과 교육평가 관련 학술대회 등을 통해 선다형 지필평가의 문제점을 극복하고 학생들의 다양한 능력을 평가할 수 있는 대안적 평가 방법에 대한 논의가 활발히 진행되고는 있다(신민희, 2012). 하지만 그럼에도 불구하고 여전히 현실에서는 직접평가 실행이 그리 일반적이지 않은 것으로 보인다. 이는 간접평가 방식에 비하여 개발과 평가 시 많은 시간과 노력을 요하기 때문일 것이다(이수진, 2011).

전공능력진단평가는 학습자들이 특정 전공 분야에서 보유해야 할 지식, 기술, 능력, 태도 등을 얼마나 함양하였는가를 정량·정성적으로 측정하고 평가하는 것을 말한다. 이러한 평가는 학습자들의 학습 수준을 파악하고 필요한 개선 사항을 식별하기 위한 기초자료가 되기도 하고, 산업 현장에 적합한 전문 인재를 양성하였는지를 검증하기 위한 목적으로 사용되기도 한다. 그렇다면 전공능력진단평가는 학습자들의 전공 지식과 능력을 신뢰할 수 있는 방법으로 평가되어야 할 것이다. 그런데 과연 지금의 자기점검형 간접평가 방식으로 진행되는 전공능력평가가 과연 신뢰로운 평가라고 할 수 있는가에 대해 의문을 제기한다.

이에 본 연구는 기존의 전공능력진단도구가 학생들의 전공능력을 직접적으로 평가하지 못하고 있는 현실을 문제로 지적하면서, 진단검사의 현황을 조사하고, 이 과정에서 도출된 의미있는 결과를 바탕으로 A 대학 교직과의 간접평가 방식으로 만들어진 기존 전공능력진단평가 문항을 수행형 직접평가 문항으로 전환하는 연구 과정을 사례로 제시하여 교직과를 포함한 다양한 전공에서 간접평가방식의 전공능력진단도구를 수행형 직접평가 방식으로 개선하고자 할 경우 활용할 수 있는 표준 모델을 제시하는 것을 연구의 목적으로 삼았다.

연구의 범위는 전공능력진단도구가 학생 중심 교육 강화와 학과 인증제도의 도입 및 대학 평가체계 구축 등의 이유로 이미 개발과 시행이 오래전부터 진행되어온 현 상황을 고려하여 전공능력진단도구를 새롭게 개발하거나 문항 내용에 집중한 것이 아닌, 기존 도구를 개선하는, 즉, 간접평가 문항에서 직접평가 수행형 문항으로의 개발 절차에 초점을 두었음을 명확히 밝힌다. 이러한 집중된 연구 범위의 설정은 보다 명확하고 구체적인 결과를 도출하여 실제 간접평가에서 수행형직접평가로의 도구 개발을 계획하는 연구자나 교육 현장에 직접적인 도움을 줄 수 있을 것이라고 기대한다.

이러한 연구 목적을 달성하기 위한 구체적인 연구 절차는 다음과 같다.

첫째, 기존 자기 점검형(5점 리커트척도) 전공능력진단도구의 문제점을 개선할 수 있는 직접평가방식의 전공능력진단도구를 개발한다. 둘째, 개발된 전공능력진단도구에 기초하여 해당 전공의 재학생들을 대상으로 전공능력을 진단한다. 셋째, 결과분석을 통해 직접평가의 효과를 증명하고, 본 연구에서 수행된 절차와 방법을 기반으로 다양한 전공영역에서 표준 모델로 활용할 수 있는 직접평가 방식의 전공능력진단도구 개발 표준 매뉴얼을 개발한다.

연구방법으로는 현재 대학생을 대상으로 진행되고 있는 대표적인 국내·외 진단검사와 관련된 문헌을 먼저 분석한다. 해외 사례로는 미국의 CLA, OECD의 AHELO, 호주의 GSA를 살펴보고, 국내 사례로는 한국직업능력연구원의 K-CESA, 한국산업인력공단의 NCS, 경기대학교와 건양대학교의 핵심역량진단평가 개발 사례를 중심으로 분석하여 시사점을 도출한다. 이를 토대

로 직접평가를 위한 수행형 전공능력진단도구 문항 및 평가기준 개발, 예비검사 실시, 최종 문항 및 평가기준 완성, 학과별 직접평가 전공능력진단 도구 개발을 위한 표준 매뉴얼 및 제반 요소 제안 순으로 연구를 수행하고자 한다. 이때 연구대상은 A 대학교의 교직원역량진단도구를 활용하고자 한다.

II. 이론적 배경

1. 용어의 개념 및 정의

가. 전공능력진단도구

전공능력은 ‘전공역량’과 동일한 개념으로 사용되고 있으나, 대학 기본역량 진단에서 전공역량 대신 ‘전공능력’이라는 용어를 사용함에 따라 현재 많은 대학에서 전공역량을 전공능력으로 변경하여 사용하고 있다(이화영, 2021). 전공능력을 그대로 풀이하면, 전공과 관련한 능력이라 할 수 있는데, 이에 대한 자세한 개념 정의를 위해 학습역량의 개념을 살펴보면, 학습역량은 학습활동을 통하여 개인의 학업성취를 이룰 수 있게 하는 개인의 능력으로 학습역량이 높다는 것은 높은 학업성취를 이룰 수 있다는 것을 의미한다고 할 수 있다(양홍권, 2016). 그렇다면 전공능력이란, 전공과 관련된 학습활동을 통하여 개인의 전공 능력을 함양하고 개인의 능력으로써 전공 능력이 높다는 것은 높은 전공 능력을 발휘할 수 있다는 것을 의미한다고 볼 수 있다. 손진아 외(2021)는 전공능력이란, 각 단과대학 또는 계열, 학과의 관련 분야에서 요구하는 직무, 과업, 역할을 수행하는 데 필요한 지식, 기술, 태도를 포함하는 복합적, 종합적 능력, 이재호(2022)는 어떤 분야나 학문을 전문적으로 연구하는 능력으로 전공능력을 정의하였다.

따라서 본 연구에서는 ‘전공능력’을 어떤 분야나 학문 혹은 계열 및 학과의 관련 분야에서 요구하는 직무, 과업, 역할을 수행하는 데 필요한 지식, 기술, 태도를 포함하는 복합적, 종합적인 능력으로 정의하고, ‘전공능력진단도구’를 이러한 전공능력을 학습자가 얼마나 잘 이해하고, 전공 직종에서 직무를 성공적으로 수행하는 데 필요한 능력을 소유하고 있는지 측정하는 표준화 검사로 정의하고자 한다.

나. 직접평가

직접평가(direct evaluation)에 대한 개념 정의를 살펴보면 다음과 같다.

먼저, 문헌에서 소개하는 개념 정의를 살펴보면, 첫째, 교육심리학 용어사전(2000)에서는 직접평가를 피검자 스스로가 자신의 지식이나 기능을 나타낼 수 있도록 산출물을 만들거나, 행동으로 나타내거나, 답을 작성(또는 구성)하도록 함으로써 지식이나 기능을 직접적으로 측정·평가하는 행위라고 말한다. 이러한 특성 때문에 참평가, 대안적 평가, 수행평가라고도 부른다. 둘째, 교육평가 용어사전(2004)에서는 직접평가를 학생의 특정한 지식이나 능력을 평가하기 위해서 학생 스스로 자신의 지식이나 능력을 산출물 혹은 행동으로 나타내도록 그 산출물과 행동을 직접적으로 평가하는 것으로 말한다. 직접 수행을 평가하는 특성 때문에 대안적 평가 방법에 속한다. 예를 들어, 영어회화 능력을 평가하기 위해서 지필 검사를 하는 것이 아닌 영어로 회화하는 것을 직접 관찰하여 평가한다. 셋째, 한국어 평가론(2006)에서는 직접평가를 구두시험이나 작문 시험과 같이 학습자의 언어 수행능력을 직접적으로 측정하는 것을 의미하며, 간접평가는 학습자의 언어수행 능력을 간접적으로 포착하거나, 간접적인 언어 지식을 측정하는 평가를 의미하며 그 예로써 지필평가의 대표적인 다지선다형 시험이나 빈칸 채우기 등을 들고 있다.

권정화(2020)는 초기 영유아 언어발달 평가에서 부모보고식 평가와 직접평가 간의 비교 연구에서 직접평가를 평가 대상자에게 적절한 평가도구를 사용하여 평가 기준에 맞게 직접적으로 평가를 실시하고, 그 결과를 기록하여 채점하는 방법으로 직접평가를 정의하기도 하였다.

이상의 직접평가에 대한 개념 정의에서 발견한 공통점은 직접평가를 설명하기 위해 ‘산출물 혹은 행동으로 나타내는 것’, ‘직접적으로 측정·평가하는 것’, ‘수행능력’이라는 개념을 사용하였다는 점이다. 이 중 ‘수행능력’은 학교 현장에서 사용하는 직접평가와 밀접한 의미 관계를 가지는 것으로 보인다.¹⁾

이상에서 수행평가는 직접평가의 성격을 가장 잘 나타내는 평가방식으로 소개 되어짐을 알 수 있다.

다음으로 직접평가는 간접평가와 대비되는 개념으로, 그리고 간접평가의 대안으로 사용되기도 한다. 교육부에서 발표한 교육과정 해설에서 이를 비교하여 개념 설명을 한 사례를 제시하면 다음과 같다.

1) 이를 증명하는 자료로 교육부(1998)에서 발표한 직접평가에 대한 정의를 소개하면 다음과 같다.
 - 간접적인 방법보다는 직접적인 방법을 중시하는 평가
 - 도덕성을 지필시험이나 구두시험이 아니라 관찰을 통해 평가하는 것이 직접평가이다.
 - 수행평가는 가능한 한 직접평가의 성격을 풍부하게 포함하려고 함.

‘직접평가’란 언어활동에 관한 지식이 아니라 실제 언어활동을 대상으로 언어 능력의 특성 및 수준을 평가하는 것을 말한다. 언어활동을 직접 수행하게 하여 평가하는 것으로, 실제로 말하고, 듣고, 쓰고, 읽는 활동을 하게 하여 그 수행 과정이나 결과를 바탕으로 언어 능력의 특성과 수준을 판단한다. 언어활동이라는 직접적인 자료로 그 능력을 평가한다는 점에서 직접평가라 한다. 이에 비해 ‘간접평가’는 언어활동에 관한 지식을 평가하거나 기타 간접적인 방법으로 언어 능력을 알아보는 방법이다. 말하기나 쓰기 등의 평가에서 선다형이나 진위형, 완성형 등의 방식으로 언어 능력을 평가하는 것을 말한다.

[출처: 교육과학기술부(2008). 국어과 교육과정 해설. p.180-183]

모 지방자치단체의 교육청 연수원에서는 간접평가를 수업 전에 이루어진 학습자의 과거 학습 경험이 얼마나 잘 입력되어 있는지를 평가하는 의미로, 직접평가는 수업 중에 이루어지는 평가로써 수행, 관찰, 과정을 평가하는 개념으로 직접평가와 간접평가를 비교·설명하기도 하였다.²⁾ 마지막으로 국어사전에서 영역별로 정의하는 직접평가에 대한 정의를 살펴보면 다음과 같다.

- 직접 평가(直接評價)[정보·통신] 표본화에 근거하거나 자동 검사에 의하여 데이터 집합의 품질을 평가하는 방법
- 직접 평가(直接評價)[교육] 실제적이고 자연적인 의사소통 상황에서 언어 사용 능력을 직접적으로 측정하는 방식의 평가. 언어를 사용할 수 있는 상황과 조건을 만들어 주고 실제로 언어를 사용하는 모습을 평가한다.
- 직접 평가 방법(直接評價方法)[환경] 환경 가치를 평가하는 방법 가운데 경제 주체들의 행동을 관찰하여 평가하는 간접 평가 방법과 달리 경제 주체들의 주관적 가치를 이해 당사자에게 직접 문의하여 조사·분석하는 평가 방법.
- 직접 피해 평가 (直接被害評價)[군사] 항공 관측, 항공 사진 또는 직접 관측에 의하여 실제 타격 지점의 피해 상황을 직접 확인하여 평가하는 일.

국어사전에서 정의하는 직접평가 개념 역시 분야에 따라 다소의 차이는 있으나, ‘직접’, ‘실제’, ‘행동’, ‘평가’와 같은 용어를 공통적으로 사용하고 있음을 확인할 수 있다.

이상의 문헌조사를 통해 살펴본 직접평가의 개념과 정의를 바탕으로, 본 연구에서는 직접평가

2) 간접평가를 수업 전에 이루어지는 학습자의 과거 학습경험의 평가와 이를 통한 수업계획을 위한 근거자료로 활용하는 의미에서 수업 전의 간접평가로, 직접평가를 수업 중에 이루어지는 평가로써 평가 결과를 토대로 교사는 학생들의 학습 상황을 이끌어 가는 근거자료로 활용하는 의미에서 간접평가와 직접평가를 사용하고 있었다.

를 학생의 지식이나 능력을 평가하기 위해서 그것을 산출물 혹은 행동과 같이 직접적인 수행능력을 통해 측정·평가하는 것을 뜻하며, 간접평가와 대비 되는 즉, 간접평가에서의 문제점을 해결할 수 있는 대안적인 평가 방법으로써, 직접적인 수행을 평가하는 것으로 정의한다.³⁾

2. 선행연구 분석

먼저, 대학생 전공능력진단도구 개발과 관련된 선행연구의 경우 초창기에는 주로 하위역량을 도출하는 데 집중하였다. 관련 선행연구로는 김동일 외(2009), 김연희 외(2010), 이경화 외(2011), 성민경(2015), 김지숙(2016), 손원빈, 최화숙(2019), 이화영(2021) 등이 있다.

이후 전공능력진단도구의 평가측면에 집중한 연구들이 진행되었는데, 먼저 김필성(2018)은 역량기반 학과 교육과정을 통해 평가하되 교수자의 업무 과중을 방지하기 위하여 현재 실시하고 있는 졸업요건(논문, 졸업작품 등)이나 교과목에서의 평가를 최대한 활용하는 방안을 제시하였다. 서금택과 윤용관(2019)은 5개 대학의 역량기반 교육과정 사례를 종합하여 보고하면서 학생 전공역량 진단과 관련하여 학생 설문조사가 실시된다고 밝혔다. 이민정과 김수동(2018)은 전공별 역량기반 교육목표 학습성과 평가서를 개발하고 시행한 결과를 보고하였다. 최현철 등(2020)은 D 대학에서 전공교육 CMS(Course Management System) 위원회를 설립하고 교육과정의 구성부터 운영, 평가 및 평가결과 분석까지 실시하는 사례를 보고하였다. 이 가운데 그들은 학과별 교육목표 성취도 평가를 제안하였다. 선행연구들의 전공교육과정 측정 방안을 대조하면 다음의 <표 1>과 같다.

<표 1> 자기조절학습 검사 도구의 내용 선행연구에서의 전공능력 측정 방안

선행연구에서의 측정 방안	김필성 (2018)	서금택, 윤용관 (2019)	이민정, 김수동 (2018)	최현철, 김수동, 김경표(2020)
학과별 교육목표 성취도 평가 방식				○
설문조사 방식		○		
학습성과 평가서 작성 방식			○	
논문, 졸업작품 등에 의한 평가 방식	○			
선정된 교과목의 시험 활동 평가	○			

3) 문헌연구를 통해 직접적인 수행을 평가하기 위해 쓰이는 문항의 유형으로는 수행형, 관찰형, 인터뷰(면접)형, 에세이형, 시뮬레이션형 등 이 있는 것으로 조사되었다.

다음으로 수행형 진단도구를 활용하여 운영한 사례 중 구체적인 사례를 소개하고, 구현하며 평가하는 과정을 다루고 있는 논문들을 제시하면 다음과 같다.

먼저, 학교 사례로, 이재현, 홍성실(2004)은 문제해결 능력을 평가하기 위한 수행형 평가 도구를 제안하였다. 학습자들에게 주어진 실제 문제를 제시하고 이를 해결하도록 요구하였는데, 결과뿐만 아니라 그 과정을 평가하였고, 이를 통해 학습자들의 문제해결 능력과 창의성 평가가 가능하다고 주장하였다. 다음, 박주현, 강미경(2016)은 프로젝트 기반 학습에서 사용되는 수행형 평가도구를 제시하였다. 학습자들이 주어진 주제나 문제를 해결하기 위해 프로젝트를 수행하고, 이를 통해 학습자의 문제해결 능력과 협업 능력 등을 평가하였다. 마지막으로 김현우, 정지원(2012)은 직무관련 프로젝트를 수행하고 평가하는 수행형 평가 도구를 제안하며, 실제 직무와 관련된 프로젝트를 수행하여 학습자의 직무 능력과 역량을 평가할 수 있다고 주장하였다.

다음, 대학 사례로, Chen, H. L., Liao, C. C., & Chang, Y. H.(2011)의 연구가 있다. 이 논문은 대만의 대학 수업에서 포트폴리오를 활용하여 학습자의 학습성과를 평가하는 방법을 제시하며, 포트폴리오를 사용하여 학생들의 학습경험을 추적하고 평가하는 방법을 연구하였다. 이 포트폴리오는 대표적인 수행형 평가 방법 중 하나로 특히 대학에서 많이 활용 되었는데, 박상희 외(2016)는 한국 대학에서 포트폴리오를 활용하여 학생들의 학습 성과와 능력을 평가하는데 대한 사례를 소개하고 있다. 한 대학의 학생들이 자기주도학습 능력을 개발하고 평가하기 위해 포트폴리오를 활용하는 방법을 탐구하였는데, 포트폴리오를 통해 학생들의 학습 활동, 학습 목표 설정, 학습 과정 기록 등이 평가되며, 이를 통해 학생들의 자기주도학습 능력을 평가하고 개선하는 데 포트폴리오를 활용할 수 있다고 주장하였다.

3. 진단도구 운영 사례

대학생 학습능력을 평가하기 위해 개발된 세계의 주요 진단 도구로는 OECD의 AHELO(Assessment of Higher Education Learning Outcomes), 미국의 CAAP(Collegiate Assessment of Academic Proficiency and Progress), MAPP(Measure of Academic Proficiency and Progress), CLA(Collegiate Learning Assessment), 캐나다의 NGS(National Graduate Survey), 호주의 GSA(Graduate Skills Assessment)와 CEQ(Course Experience Questionnaire), 브라질의 ENADE(Exame Nacional de Desempenho dos Estudantes), 멕시코의 EXANI-III(Examen Nacional de Ingreso al Posgrado) 등이 있다.

국내의 대표적인 진단도구로는 한국직업능력연구원의 K-CESA(Korea Collegiate Essential

Skills Assessment), 고용노동부와 한국산업인력공단의 NCS(National Competency Standards)가 있고, 대학 자체적으로 진단검사도구 개발에 앞장섰던 사례로는 성균관대학교의 SCCA, 숙명여자대학교의 SM-CESA, 건양대학교의 핵심역량진단평가, 동국대학교의 D-CESA 등이 있다.

이 중 대표적인 검사로 알려져 있으며, 수행형 검사를 실시하여 본 연구의 목적인 직접평가를 위한 전공능력진단도구를 개발하는 데 있어 참고할 수 있는 해외 사례로 OECD의 AHELO, 미국의 CLA, 호주의 GSA와 국내 사례로 한국직업능력연구원의 K-CESA, 고용노동부와 한국산업인력공단의 NCS, 국내대학(경기대학교, 건양대학교) 등의 역량진단평가에 대해 분석해보았다.

이상과 같이 본 연구의 목적인 수행형 문항을 바탕으로 하는 전공능력진단도구 개발을 위해 조사한 선행연구들의 국내·외 사례를 종합하여 볼 때, 다음의 공통점을 발견할 수 있었다.

첫째, 간접평가 방식인 점검형 문항과 직접평가 방식으로 대변할 수 있는 수행형 문항이 진단도구에 함께 쓰였다.⁴⁾ 즉, 기존 간접평가 방식은 낮은 신뢰도 및 타당도를 가지고 있는 단점이 있지만 반대로 수행형의 경우 답안 작성의 어려움으로 인한 불성실성, 그리고 채점 및 관리의 어려움이 있기 때문에, 간접평가와 직접평가 방식을 혼합하여 도구를 개발하였다. 둘째, 문항은 구체적인 문제상황과 현장 관련성이 높은 자료를 제시하였고, 단순 창의력을 요하는 문제가 아닌 이론을 바탕으로 하여 피험자의 역량을 종합적으로 발휘하여 문제를 해결할 수 있도록 개발하였다. 셋째, 직접적인 능력을 측정하기 위해 다면평가 방법을 적용하였다는 점이다. 대표적인 문항 유형으로는 수행형, 관찰형, 인터뷰(면접)형, 에세이형, 시뮬레이션형 등이 있었다. 측정하고자 하는 능력을 평가할 때 효과적인 방식의 수행형 문항을 다양하게 선택하여 적용하였다. 넷째, 개발 절차가 매우 복잡하기 때문에 많은 연구진들이 대거 투입된 대규모 프로젝트로 진행되었고, 척도개발에 큰 비용과 시간이 소요되었다는 점이다. 다섯째, 교육을 받기 이전의 상태와 교육을 받은 이후의 상태를 비교하였다. 즉, 해당 교육기관에 입학하기 전 혹은 막 입학한 시기에 학생의 수준과 특정 기간 동안 교육을 받은 이후의 학생의 능력을 측정하여 교육의 성과를 비교하였다는 점이다.

이상의 선행연구 분석을 통해 얻은 시사점은 직접평가를 위한 예비문항 개발에 많은 도움이 될 것으로 기대된다.

4) 본 연구에서 직접평가는 기존 자기 점검형 방식의 간접평가와 대비되는 개념으로 사용하며, 그 방법으로 수행형 문항을 제시한다.

4. 수행형 평가의 한계

수행형 직접평가는 학습자가 실제로 문제를 해결하거나 작업을 수행함으로써 그들의 실제 능력을 직접 평가할 수 있다는 점에서 본 연구의 과제인 자기점검형 간접평가방식의 문제점을 보완할 수 있는 대안으로 제시하였다.

수행형 직접평가는 실제 역량평가와 개별화된 평가가 가능하다. 학습자의 실제 역량과 능력을 평가할 수 있는 방법으로, 단순한 지식 습득에 그치지 않고 실제 응용능력을 평가할 수 있다 (Garrison, D. R., & Ehringhaus, M., 2007). 또한 개별 학습자의 특성과 수준에 맞춰서 평가할 수 있어서, 학습자마다 다양한 평가 기준을 적용할 수 있다(Sadler, D. R., 1989)는 장점을 가지고 있다. 하지만 이점과 함께 연구자는 수행형 문항이 갖는 한계점도 있음을 알고 개발을 수행하여야 할 것이다. 이는 다음과 같다.

첫째, 평가자의 주관적인 판단이 개입될 수 있으며, 이로 인해 평가의 공정성과 객관성이 상실될 수 있다(Birenbaum, M. 2006). 즉, 수행형 평가에서는 학습자의 작업이나 프로젝트를 개별적으로 평가하기 때문에 평가자의 주관이나 편향이 불가피하게 작용할 수 있으며, 이는 평가의 공정성을 해치고 학습자들에게 불공정한 결과를 초래할 수 있다.

둘째, 수행형 평가는 설계와 실시에 많은 시간과 비용이 소요되며, 특히 대규모로 평가를 실시할 때는 이러한 비용이 더욱 증가할 수 있다(Airasian, P. W., 1994). 수행형 평가는 학습자들이 실제 작업을 수행하고 결과물을 제출하는 과정이 필요하기 때문에 시간과 비용이 많이 소요될 수 있는데 특히 대규모 수업이나 다수의 학습자를 대상으로 할 경우 더 많은 리소스가 필요하다.

셋째, 수행형 평가는 일반화가 어렵다. 수행형 평가에서는 학습자의 작업이나 프로젝트를 특정한 상황에 대한 응용으로 평가하기 때문에 이를 일반화하여 다양한 상황에 적용하기 어려울 수 있다.

넷째, 평가의 일관성을 유지하기 어렵다. 수행형 평가에서는 학습자들이 다양한 작업을 수행하기 때문에 평가자들 간에 일관된 평가 기준을 유지하기 어려울 수 있다. 이는 평가의 신뢰성을 저하시킬 수 있고, 이로 인해 학습자들에게 혼란을 야기할 수도 있다.

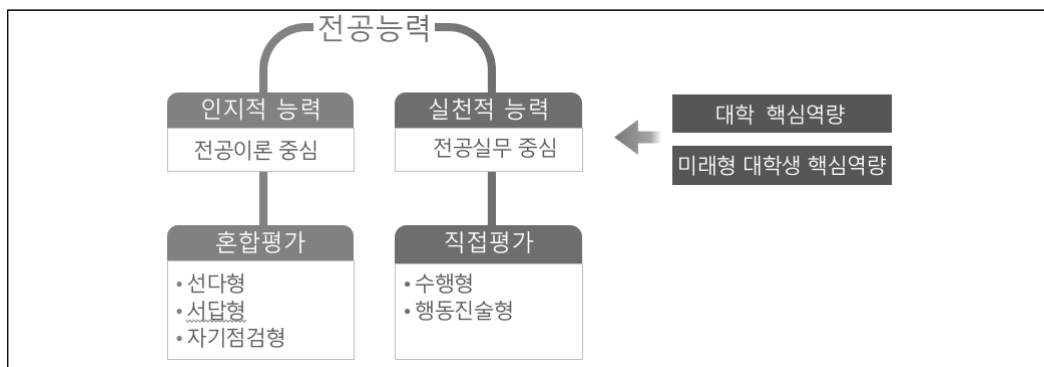
따라서 본 연구에서는 이를 보완하기 위해 채점자 간 평가방법을 선택하고, 전면 수행형 문항 도입보다는 실천적 능력을 측정하는 문항에 대해서 먼저 수행형으로 교체하는 점진적 개선 방안을 제안하였다. 본 연구물을 참고로 수행형 진단도구를 개발하고자 하는 연구진은 반드시 이상과 같은 수행형 평가의 한계점을 알고 이를 보완할 수 있는 방안 모색을 함께 하여야 할 것이다.

III. 수행형 전공능력진단도구 문항 개발

1. 직접평가를 위한 수행형 전공능력진단도구 프레임 개발

수행형 전공능력 진단도구 문항 개발을 위한 틀(frame)을 제시하기 위해 A 대학교에서 개발된 학과별 전공능력진단도구의 내용을 분석하였다. 그리고 그와 함께 단일한 대학생 전공기초능력 진단도구로서의 통합성과 체계성 확보를 위해 다양한 계열의 학과 공동으로 2023.7.11.~2023.7.20.에 걸쳐 회의를 개최하며 연구 프로세스와 결과 보고 방식에 대한 통일성을 기하고자 노력하였다. 그 결과 진단하고자 하는 전공능력은 크게 인지적 능력과 실천적 능력으로 나눌 수 있으며, 또다시 그 행위진술을 중심으로 하위영역에 대한 평가요소를 인지적 요소, 정의적 요소, 행동적 요소로 구성되어 있음을 확인할 수 있었다.

더불어 전공능력은 학생이 대학 졸업 후 변화하는 미래 산업현장에 잘 적응하여 필수 직무능력을 발휘할 수 있는 역량을 말하며, 전공능력진단도구는 그 결과를 진단하고자 함이므로, 진단도구 개발 시 미래형 대학생 핵심역량 혹은 대학 자체 핵심역량과의 연관성도 고려하는 것도 필요하다. 이를 그림으로 나타내면 [그림 1]과 같다.



[그림 1] 전공능력진단도구 개발 프레임

본 연구는 직접평가를 위한 수행형 문항의 전공능력진단도구를 개발하는 것을 목적으로 하나 다음의 이유를 통해 도입 초기부터 수행형 문항을 전면 도입하기보다는 인지적 능력과 실천적 능력으로 전공능력진단 영역을 분류하여 실천적 능력 즉, 전공실무 중심의 평가가 필요한 영역에 대해서 우선 직접평가 방식을 도입하고 점진적으로 그 범위를 확장하는 프레임을 제안한다.

그 이유는 다음과 같다.

먼저 교육목표 분류 측면에서 ‘인지적 능력’은 사람이 정보를 수집하고 처리하며, 그 결과를 기억하고 이해하는 능력을 의미하며, 지각, 기억, 사고, 문제해결 능력 등이 이에 해당하고 주로 지각력 테스트, 지능검사 등을 통해 측정된다. ‘실천적 능력’은 개념이나 지식을 실제 상황에서 적용하고 실행하는 능력을 의미한다. 이는 어떤 작업을 수행하고 문제를 해결하는 데 필요한 능력을 말하는 것으로 기술적 능력, 손재주, 응용력, 문제해결 능력, 의사소통 능력 등이 해당하고, 실습 과제, 실제 업무수행, 시뮬레이션 등을 통해 평가될 수 있다.

따라서 인지적 능력을 확인하기 위한 문항보다는 실천적 능력에 해당하는 전공실무 중심의 문항 즉, 행동적 요소를 수행형 직접평가 문항으로 먼저 전환하여 평가하는 것을 제안한다.

다음 실천적 측면에서 점진적 도입을 주장하는 이유는, 첫째, 도입 초기부터 전면 수행형 평가를 도입할 경우 평가자와 피평가자 모두에게 상당한 부담이 되기 때문이다. 둘째, 간접평가 방식으로 효과적으로 진단이 가능한 인지적 능력에 해당하는 하위 영역이 있기 때문이다. 셋째, 수행형 검사를 위한 기반 요소가 갖추어지지 않았기 때문이다. 넷째, 수행형 검사를 위한 기반 요소가 갖추어질 시간이 필요하기 때문이다.

따라서 기존의 간접평가 방식에서 직접평가 방식으로의 도입을 처음 시도하는 교육 현장에서는 수행형 평가방식을 전면 도입 하기보다는 실천적 능력에 해당하는 영역에 한 해 수행형 문항을 설계하는 것을 제안한다. 하지만 인지적 능력을 진단하는 경우에도 채점의 수월성으로 많이 활용되는 자기점검형 체크리스트만을 고집하기보다는 실제로 수행하는 구체적 행동에 근거하여 직무수행을 측정할 수 있는 행위기준평가형 문항을 도입하는 등 평가목적에 맞는 다양한 문항을 구성하여 혼합 평가하는 것을 제안한다.

2. 예비 검사 문항 개발

가. 직접평가를 위한 수행형 문항 개발 절차

직접평가를 위한 수행형 문항 개발을 위해 일차적으로, 기존에 개발된 전공능력진단도구 문항을 인지적 능력과 실천적 능력으로 분류하였다. 앞서 전공능력진단도구 개발 프레임에서 제시한 점진적 도입과 평가목적에 맞는 혼합평가를 위해 기존에 개발된 간접평가 방식의 진단도구를 재점검하여 이 중 실천적 능력에 해당하는 문항을 1차 전문가 협의회를 통해 추출하였다. 이렇게 추출된 문항을 수행형 문항으로 개발하고, 개발된 예비문항에 대한 내용타당도 검사를 위해 2차

전문가 협의회를 개최하였다. 이 과정에서 도출된 수정내용을 보완하여 예비문항을 완성하고, 예비 검사 대상자를 선발하여 실제 예비 검사를 시행하고, 그 결과를 통해 검사소요 및 채점과 같은 운영방식의 효율성을 검토하고 안면 타당도 검사, 준거 타당도 검사, 신뢰도 검사 등을 시행하여 최종적으로 전공능력진단도구를 완성하였다.

나. 직접평가를 위한 수행형 예비문항 개발

1) 문항 개발을 위한 전공 배경 이해

전공능력진단도구를 개발하기 위해서는 해당 전공(학과)에 대한 이해가 수반 되어야 한다. 교원양성을 목적으로 하는 교직과는 교원자격증 취득을 위해서 교직과목을 22학점 이상 취득해야 하며, 그중에서도 교육실습 영역에서 2과목 4학점 이상을 취득해야 한다. 현행 교육실습은 학교현장실습(수업실습, 참관실습, 실무실습 등)과 교육봉사활동을 포함하고 있다. A 대학교 교원양성기관인 교직과의 교육목표는 ‘바른 인성을 갖춘 현장 리더형 교사 양성’이며, 이를 위해 갖추어야 할 교직 역량과 그 하위역량은 <표 2>와 같다.

<표 2> A 대학교 교직 역량

교직역량	하위역량
교육 과정 · 수업 역량	교육과정
	수업
학생 지도 역량	생활지도
	진로지도
학교 공동체 운영 역량	학교 및 학습경영
	교직실무
교직관리 역량	교직인성
	자기관리

또한 A 대학교는 2019학년도 교직역량기반 성과분석을 통해 교육실습을 위해 필요한 주제와 영역으로 ‘생활지도’, ‘행정처리 업무’, ‘지도안 작성’ 및 ‘수업 시연’을 도출하였고, 이 네 가지 영역이 전공과목에서 골고루 다루어지고 있음을 확인하였다. 따라서 교원자격증 취득을 위한 교직과목 중 교육실습 과목을 실천적 능력을 신장시키기 위한 과목으로 분류하고, 이에 필요한 ‘생활지도’, ‘행정처리 업무’, ‘지도안 작성’, ‘수업시연’ 영역을 평가하는 방법으로 직접평가를

도입하여 그에 필요한 수행형 문항 설계가 필요하다는 사실을 도출하였다.

2) 문항 개발을 위한 전문가 협의회

예비문항 개발을 위해 전문가 협의회를 개최하였다. 전문가 협의회는 사범계열 교수, 교육효과 측정 및 평가 분야 전문가, 현장 교사 등 총 11명으로 구성하였다. 자세한 내용은 <표 3>과 같다.

〈표 3〉 전문가 패널 구성표

번호	구분	소속	직위	비고	번호	구분	소속	직위	비고
1	A	A 대 사범 계열	교수		7	G	A 대 사범 계열	교수	
2	B		교수		8	H	교육효과 측정 및 평가 전문가	교육학박사	
3	C		교수		9	I		교육학박사	
4	D		교수		10	J	성남00 고등학교	교사	경력 10년 이상
5	E		교수		11	K	산본00 고등학교	교사	
6	F		교수						

먼저, 교직과의 전공능력진단도구에서 실천적 능력을 측정하기 위한 문항을 추출하기 위해 1차 전문가 협의회를 개최하였다. 그 결과 기존 31개의 자기점검형 5점 척도 문항에서 수행형 문항으로 재구성하였을 경우, 보다 정확한 능력 측정이 가능하다고 판단되는 문항으로 6개의 문항이 추출되었다. 이때 기존 문항과의 연계성을 확보하고 질 높은 문항을 개발하기 위해 A 대학교 교직역량진단도구 문항 개발 선행연구자들을 포함시켜 총 7명의 전문가 의견 수렴 절차를 거치도록 하였다. 또한 직접평가를 위한 수행형 문항 및 평가기준의 내용타당도를 검증하기 위한 전문가 협의회도 추가로 개최하였다.

3) 직접평가를 위한 수행형 문항 개발

A 대학교의 자기점검형 교직역량진단도구의 문항은 총 31개 문항이다. 이 중 직접평가가 필요한 실천적 능력 - ‘생활지도’, ‘행정처리 업무’, ‘지도안 작성’, ‘수업시연’- 에 해당되는 문항을 델파이 조사를 통해 도출하여 직접평가를 위한 수행형 문항으로 재설계하였다. 조사에 참여한 전문가 패널은 교직담당 및 사범계열 교수 총 7인으로 구성되었으며, ‘생활지도’, ‘행정처리 업무’, ‘지도안 작성’, ‘수업시연’ 영역에 해당하는 문항을 도출하고, 이들 중 특히 수행형 문항을

적용하여 직접평가를 실행해야 한다고 생각되는 문항을 선택하게 하고, 이 중 공통된 의견이 모아진 문항을 바탕으로 <표 4>와 같이 직접평가 방식으로 바뀌어야 할 문항을 도출하였다.

이렇게 도출된 문항을 앞서 도출한 선행연구의 시사점과 전공 배경에 대한 이해 및 개발 프레임을 반영하여 4~6번 및 28번 문항은 수행형으로, 15번 문항은 시뮬레이션형으로, 30번 문항은 관찰 및 에세이형으로 문항개발을 하기로 결정하였다.

<표 4> 직접평가가 필요한 문항 도출 결과

교직역량	순번	문항	문항 유형	
교육과정 · 수업역량	교육과정	1	4. 학습자 수준에 맞는 교육내용을 선정하고 수업을 설계하여 전개해 나갈 수 있다.	수행형
	수업	2	5. 수업설계 원리를 알고 수업에 적용할 수 있다.	수행형
		3	6. 다양한 교수매체의 유형을 알고 학습자에게 효과적인 교수매체를 활용할 수 있다.	수행형
학생지도역량	생활지도	4	15. 부적응 학생, 정서행동고위험군 학생, 다문화 학생의 특성을 이해하고, 상담할 수 있다.	시뮬레이션 평가
교직관리역량	자기관리	5	28. 과학기술의 발전 및 교육정책 변화를 이해하고 적절하게 대응할 수 있는 교육매체 활용 능력을 습득할 수 있다.	수행형
		6	30. 교육봉사활동을 통해 봉사를 실천할 수 있다.	관찰, 에세이

먼저, 4~6번 문항은 수업설계 원리(이론)를 기반으로 수업을 설계하고, 적절한 교수매체를 활용하여 수업을 전개해 나갈 수 있는 ‘교육과정·수업역량’, ‘교직관리역량’을 진단할 수 있는 능력을 측정하기 위한 내용이므로 수업지도안 설계 및 모의수업 진행을 통한 종합형 수행형 문항으로 재구성하였다. 실제 교사 임용후보자 선정 경쟁 2차 시험에서 수업지도안 설계 및 모의수업을 진행하고 있어, ‘교육과정·수업역량’을 평가하기에 좋은 방법이라고 판단하였다.

다음으로 15번 문항의 경우 ‘학생지도역량’을 평가하는 문항으로써 부적응 학생, 정서행동고위험군 학생, 다문화 학생의 특성을 이해하고 상담할 수 있는 능력을 측정하기 위한 내용으로 이는 ‘시뮬레이션형’ 문항으로 재설계하였다.

마지막으로 교육 환경 변화를 이해하고 적절하게 대응할 수 있는 교육매체 활용 능력을 묻는 28번 문항과 교육봉사활동을 통해 봉사를 실천할 수 있는지를 묻는 30번 문항은 ‘교직관리 역량’ 중 ‘자기관리’를 평가하기 위한 문항으로 18번 문항은 수행형 문항으로, 30번 문항은 평가요소 중 정의적 요소를 평가하기 위한 문항으로 관찰형과 에세이형으로 문항을 재설계하였다.

이렇게 개발된 예비검사 문항의 안면타당도를 검토하기 위해 사범계열 및 교직이수 학생 7명에게 검사를 실시한 후, 그 결과를 반영하여 문항을 수정·보완한 후 내용 전문가 총 4명(현장교사 2명, 교육과정 및 측정 분야 박사 2명)에게 내용타당성을 평가하도록 하였다. 각 문항에 대해 ‘매우 적절하다(4점)’, ‘적절하다(3점)’, ‘적절하지 않다(2점)’, ‘전혀 적절하지 않다(1점)’의 4점 Likert 척도로 평가하도록 하여 내용타당도 지수(Content Validity Index, CVI)를 산출하였다. 그 결과 평균 CVI는 .96으로 높게 나타났으며, 6개 수행형 문항 중 CVI 1.0이 5개 문항(83.3%), CVI .87이 1개 문항(16.7%)이었다.

내용타당도 검증 시 검사 대상자들의 수준에 맞는 용어의 선택과 간결하고 명확한 문제 제시, 난이도 조절 등에 대해서는 의견을 따로 기입하도록 하였다. 그 결과 28번 문항의 경우 비록 ‘교직원역량’을 측정하기 위한 문항으로 분류되어 있으나, 그 내용이 ‘교육과정·수업역량’과 더 밀접한 관련이 있으며, 다양한 교수매체의 유형을 알고 학습자에게 교수매체를 활용할 수 있는지를 묻는 6번 문항과 내용이 중복되어 ‘교육과정·수업역량’을 측정하는 문항에 포함시키는 것이 좋겠다는 의견이 도출되어 최종적으로 4~6번과 28번 문항을 대과제 1문항으로, 15번과 30번 문항을 각각 소과제 1문항씩 도출하였다. 이때 30번은 관찰형과 에세이형 두 가지 방식으로 제안하였다.⁵⁾ 수행형 문항을 평가하기 위한 평가기준(루브릭) 개발도 동일한 내용전문가를 대상으로 내용타당도 평가를 실시하였다. 그 결과 검사 대상 수준에 적합한 성취기준 제시, 평가준거의 구체화 등을 통해 평가기준안을 재수정하였다. 더불어 직접평가의 특성상 학생들에 따라 다양한 응답이 제시될 경우를 고려하여 예비 검사를 통해 수행된 자료를 검토하여 최종적으로 평가기준(루브릭)을 도출하였다. 그리고 이때 평가기준에 따른 점수 배점은 기존 자기점검형 문항과의 연관성을 고려하여 동일하게 5점으로 구성하였다.⁶⁾

선행연구 분석을 통해 채점과 관련하여서는 2인(이상)의 공동 채점을 권장한다. 그 이유는 첫째, 평가의 객관성과 신뢰성을 높이기 위함이고, 둘째, 평가자 사이에 발생할 수 있는 의견 불일치를 좁히기 위함이다. 따라서 문항 4번과 같이 평가기준에 주관이 개입될 수 있는 문항의 경우나 전공심화 능력을 평가해야 하는 경우에 동일문항에 다수의 평가자를 배정하는 것이 좋고 할 수 있다. 이렇게 하여 도출된 수행형 진단도구의 최종 문항과 평가기준(루브릭)의 일부를 제시하면 [그림 2]와 같다.

5) 본 연구의 최종 목적은 직접평가를 위한 수행형 문항 개발 표준 매뉴얼을 제시하는 것이며, 이를 위한 과정을 직접 수행하여 시행착오를 줄이고, 해당 전공에 적용해 볼 수 있는 프로세스를 제안하는 것이다. 따라서 본 연구에서 제안하는 문항 유형 이외의 문항을 내용에 따라 적절히 적용하여 개발 가능하다.

6) 본 연구에서는 직접평가의 도입을 적용하는 단계에서는 전면 수행형 문항으로 개발하기보다 직접평가와 간접평가를 혼합한 혼합형 문항을 제안하고 있기 때문에 5점으로 점수를 편성하였을 경우 기존 문항의 평가 점수와 합산하여 결과 분석하기에 용이할 것으로 판단하여 5점 만점의 평가기준(루브릭)을 제시하였다.

<수행형 문항>

과제1 : 교육과정안(수업지도안) 작성하기 60분
(준비물: 교과서, 컴퓨터)

- 조건
- 수업지도안 작성을 위해 자신의 전공 교과(서)의 일부 단원을 선정
- 학습자 수준에 맞게 교육 내용을 선정하고 조직
- 수업설계 시 교육과정 설계의 일반원리를 적용
- 적절한 교수매체를 활용하여 수업지도안을 작성

과제2 : 수업 시연하기(모의수업) 15분(장소: 강의실)
[교사 임용 후보자 선정 경쟁 2차 시험의 내용을 참고하여 개발하였음]

- 수업지도안을 바탕으로 수업의 전개 과정부터 수업 시연할 것
- 교사와 학생의 활동이 구체적으로 드러나도록 시연할 것
- 실생활과 연관지어 활동을 마무리하는 과정을 포함할 것
- 강의식 수업으로 시연할 것
- 교수매체를 활용할 것
- 적정량의 판서를 확보할 것

평가 기준		
평가 항목	평가 요소	배점
수업설계원리(이론)를 적용하여 교육과정(학습지도안)을 설계하고, 이를 바탕으로 모의수업을 진행하는 능력(교육과정·수업역량, 교직관리역량)	① 교육과정의 기본요소가 수업지도안에 모두 반영되었는가?	5개 만족 5
	② 학습목표·내용선정 및 조직·평가가 일관되게 구성되었는가?	4개 만족 4
	③ 평가는 학습목표 달성 여부를 확인할 수 있는 내용과 방법으로 설계되었는가?	3개 만족 3
	④ 내용을 효과적으로 전달할 수 있는 교수매체를 활용하였는가?	1~2개 만족 2
	⑤ 이상의 수업설계를 통해 (모의)수업을 전개하였는가?	자의적미참여 1

<시뮬레이션형 문항>

- 시나리오 1 -

나는 00중학교 신입교사이며, 1학년 1반의 담임을 맡고 있습니다. 최근 학급 학생 중 한 명인 영철이가 결석이 잦고, 수업 활동에 잘 참여하지 않으며, 앞드려 자는 등 무기력한 모습을 자주 보입니다. 위(Wee) 클래스에 요청하여 진단을 해본 결과 스트레스 지수가 매우 높게 나타났으며, 자살충동 지수 또한 매우 높게 나타나 정서행동 고위험군 학생으로 분류되었습니다. 담임교사로서 부적응 및 고위험군 학생인 영철이를 더이상 방치해서는 안 된다고 판단하여 학과 시간 이후 영철이와 상담을 하기로 하였습니다.

- 시나리오 2 -

나는 00중학교 신입교사이며, 1학년 1반의 담임을 맡고 있습니다. 지금은 신학기로 아직까지 학급 학생들이 서먹한 상황입니다. 특히 학급 학생 중 한 명인 수미는 아버지가 캐나다 사람입니다. 따라서 수미는 서구적인 외모를 가졌는데, 이 때문인지 반 아이들은 수미를 멀리하고, 수미역시 선뜻 아이들에게 잘 다가가지 못하는 것 같습니다. 따라서 담임교사인 나는 학급 학생들에게 다문화 이해 교육을 실시하고, 수미와도 상담을 하기로 하였습니다.

[방법 : 영철, 수미 역을 맡을 별도의 인력을 섭외하거나 평가자가 역할을 대신할 수 있다.]

평가 기준		
평가 항목	평가 요소	배점
부적응 학생, 정서행동, 고위험군 학생, 다문화 학생의 특성을 이해하고 상담할 수 있는 능력(학생지도역량)	① 내담자와 신뢰를 형성할 수 있는 태도와 자세를 갖추었는가?	4개 만족 5
	② 부적응 학생과 정서행동 고위험군 학생을 위한 프로그램 및 지역센터와의 연계 제도에 대한 정보를 알고 적절히 안내하였는가?	3개 만족 4
	③ 다문화 학생의 특성을 이해하기 위한 기본적 소양과 지식을 준비하고 있는가?	2개 만족 3
	④ 비밀상담을 원칙으로 하나 사태의 심각성을 인지할 경우 학부모와 학교에 적절한 안내 조치를 취하였는가?	1개 만족 2
		자의적미참여 1

[그림 2] 수행형 진단도구의 최종 문항과 평가 기준

IV. 수행형 전공능력진단도구 개발 결과

1. 예비 검사 실시

가. 연구 대상

직접평가를 위한 수행형 전공능력진단도구의 경우, 채점 시 선다형 검사에 비하여 채점에 요구되는 시간이 다소 길기 때문에 채점 기간을 고려하였으며, 검사의 효율성을 위해 정규고사 기간을 이용해 2023년 6월 13일부터 6월 28일까지 16일간 예비검사를 실시하였다. 이를 위해 A 대학교 사범계열 및 교직과정 이수 학생을 검사 대상으로 표집하였으며, 예비검사 표집은 교직을 배우는 전공 계열을 고르게 분포하도록 하였고, 원활한 검사를 위해 4학기 이상의 고학년 학생을 표집하는 것을 원칙으로 하였다. 그 결과 교직과정을 이수 중인 사회복지학과(4명), 향장미용학과(3명), 식품영양학과(5명), 호텔조리학과(2명), 간호학과(3명) 학생 17명과 사범계열 재학생인 유아교육과(5명), 초등특수교육과(14명), 유아특수교육과(7명) 학생 26명으로 총 43명을 표집하였다. 전체 43명의 학생은 예비검사에 모두 참여하였고, 성실하게 응답하였다. 연구대상에 대한 일반적인 특성은 <표 5>와 같다.

〈표 5〉 예비검사 표본의 특성

구분		인원(명)	비율(%)
성별	남	14	32.6
	여	29	67.4
	합계	43	100.0
학년	3학년	37	86.0
	4학년	6	14.0
사범계열	유아교육과	5	11.6
	초등특수교육과	14	32.5
	유아특수교육과	7	16.3
학과	간호학과	3	7.0
	사회복지학과	4	9.3
	향장미용학과	3	7.0
	식품영양학과	5	11.6
	호텔조리학과	2	4.6

나. 예비검사 실시 절차

예비검사는 기말고사 기간을 이용해 강의실에서 검사를 시행하였다. 검사 시간은 문항의 내용과 문제해결 과업의 양을 고려하여 15분~60분으로 설정하였다. 학생들은 우선 기존의 자기점검형 방식인 간접평가 검사에 응답한 후, 수행형 검사에 응답하였다.

수행형 검사의 예비검사 결과는 교육학 박사인 교직과 교수에 의해 채점되어 산출되었다. 채점은 예비검사 기간과 동일하며, 채점자 간 신뢰도를 확보하기 위해 한 명의 학생에 대해 두 명의 채점자가 채점하도록 하였다.

두 명의 채점자로 제한한 이유는 채점자가 많아질수록 각 채점자 간의 의사소통이 어려워지고, 채점의 일관성과 신뢰성이 감소할 수 있기 때문이다(박상희 외, 2016). 물론 채점자 수를 2명 이상으로 확대하면 다수의 채점자들 간의 의견을 종합하여 보다 신뢰성 있는 평가가 가능할 수도 있다. 그러나 채점자가 증가할수록 채점의 일관성을 유지하기 위한 채점 교육과 관리에 더 많은 노력과 자원이 필요하며, 채점자 간의 일치도를 유지하는 것이 어려워질 수 있고, 소요되는 시간과 비용도 증가할 수 있기 때문이다. 또한 채점자 선정 시에는 해당 분야의 전문 지식과 채점 경험이 있는 사람을 선정하고, 채점 교육을 제공하여 채점의 일관성을 유지하는 것도 필요하다.

채점 이후에는 수행형 검사 결과에 대한 기초통계 및 기존 검사와의 관계를 분석하였다.

2. 예비 검사 결과 연구 대상 및 자료 수집

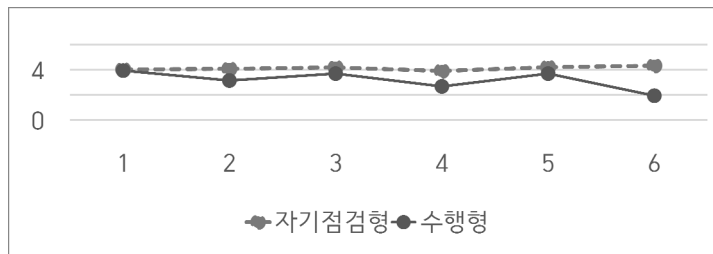
가. 기존 검사와의 관련성

수행형 문항으로 재구성된 전공능력진단도구가 기존의 검사와 비교하여 어떤 특성을 가지는지 비교하기 위해 첫째, 동일 학생들에게 기존 자기점검형 평가와 개발된 수행형 평가에 모두 응시하게 하여 문항정답률을 비교하였다. 이때 자기점검형 문항과 수행형 문항의 수가 다르므로 수행형 평가지의 내용을 관련성 있는 자기점검형 문항에 맞게 세분화하여 평가하고, 그 평균값을 비교하였다. 자기점검형과 수행형 평가 간 정답률 차이가 어떤 분포를 나타내는지 <표 6>과 같이 확인하고, 차이 값을 산출하였다.

〈표 6〉 자기점검검형과 수행형 문항 간 평균 정답률 비교와 차이

교직역량		검사	자기점검검형	수행형	차이
교육과정	교육과정	1	4.02	3.93	0.09
		2	4.07	3.14	0.93
수업역량	수업	3	4.18	3.69	0.49
		4	3.90	2.67	1.23
학생지도 역량	생활지도	5	4.20	3.69	0.51
		6	4.32	1.93	2.39
평 균			4.12	3.18	0.94

먼저, 교육과정·수업역량을 묻는 1~3번 평가에서의 평균 문항정답률은 각각 5점 만점에 자기점검검형의 경우 4.02, 4.07, 4.18, 수행형에서는 3.93, 3.14, 3.69로 두 문항 간 평가 점수 차이가 있는 것으로 나타났다. 즉, 학생들은 자신이 교육과정을 설계하고, 설계한 수업을 실제 전개해 나갈 수 있는 능력이 높다고 평가하였으나, 실제 수행평가에서는 그보다 낮은 점수가 나타났다. 이를 [그림 3]과 같이 비교하였다.



[그림 3] 자기점검검형과 수행형 문항의 평균 정답률 비교와 차이

특히, 학생지도역량(상담능력)을 묻는 4번 검사와 교직관리역량(봉사실천)을 묻는 6번 검사의 경우 기존 자기점검검형의 검사로 평가하였을 경우 각각 3.90, 4.32로 답하였으나, 이를 수행형(시뮬레이션형)과 실제 봉사 경험 횟수로 평가하였을 경우, 2.67과 1.93으로 문항정답률이 크게 차이 나는 것으로 나타났다.

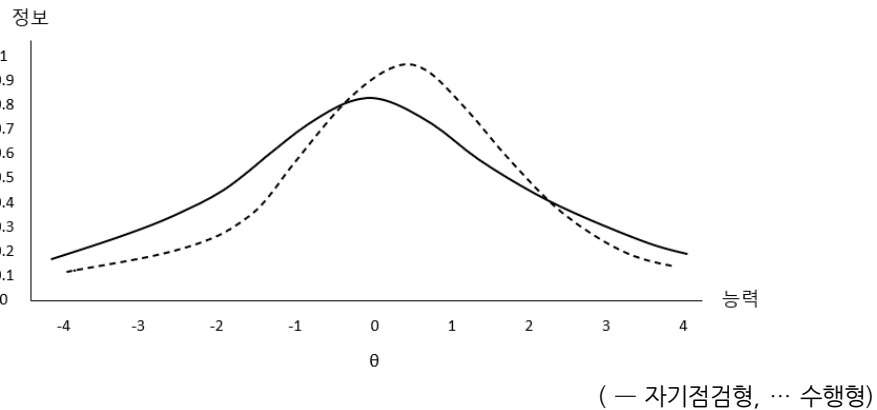
다음으로, 자기점검검형 평가와 수행형 평가를 비교·분석하기 위하여 문항반응이론 중 검사정보 함수 수식을 활용해 분석하였다.

$$I(\theta) = \sum_{j=1}^n I_j(\theta)$$

여기서, $I(\theta)$: 능력수준 θ 에서 검사정보의 양
 $I_j(\theta)$: 능력수준 θ 에서 문항 i 에 대한 정보의 양
 N : 검사에서 문항들의 수

수식에서 제시하는 검사정보 함수는 검사유형별로 제공하는 정보의 차이를 비교할 수 있게 한다. 자기점검형의 선다형 문항과 수행형 문항의 수가 다르기 때문에 검사정보 함수의 평균값을 사용하여 분석하였다.

[그림 4]는 자기점검형 문항보다 수행형 문항의 곡선이 더 높게 나타났으며, 이는 더 많은 정보를 제공하고 있는 것을 보여주는 것이다. 즉, 수행형 문항들이 자기점검형의 간접평가보다 학생들의 전공 능력을 더 정확하게 측정해 줄 수 있는 문항들이라는 의미이다.



[그림 4] 자기점검형의 선다형 문항과 수행형 문항 간 정보함수 비교

그리고 [그림 4]에서 수행형의 경우에 오른쪽으로 약간 치우치고 있는 것을 확인할 수 있는데, 이는 능력이 높은 학생들이 수행형 문항으로 평가할 경우 자기점검형 문항보다 정확하게 능력을 측정할 수 있다는 것을 의미한다. 따라서 수행형 문항으로 구성된 전공능력진단도구가 학생들의 능력 평가를 위해 더 적합한 문항들임을 보여주고 있다고 해석할 수 있다.

이와 같은 결과를 바탕으로 자기점검형 평가와 수행형 평가 간의 비교를 통해 도출한 특성을 정리하면 다음과 같다.

첫째, 학생들은 자신의 능력에 대해 관대한 평가를 하는 것으로 나타났으며, 특히 실천적 능력을 측정하는 검사에서 이러한 차이는 더욱 크게 나타났다. 둘째, 능력이 우수한 집단일수록 수행

형 문항을 사용하였을 때 자기점검형문항 보다 정확하게 능력을 추정할 수 있다. 셋째, 수행형 문항이 자기점검형 문항보다 더 많은 정보를 제공할 수 있다.

이상과 같이 수행형 문항이 전공능력을 진단하는 데 보다 정확성을 도출할 수 있다는 것을 자기점검형과 수행형 문항 간의 평균 정답률 비교와 차이, 문항 간 정보함수를 비교·분석한 결과를 통해 확인할 수 있었다. 이는 본 연구가 지적하는 자기점검형의 간접평가 방식의 문제점을 해결하는 대안으로써의 수행형 직접평가방식이 효과가 있다는 사실 검증으로 매우 의미 있는 결과라 할 수 있다.

나. 채점자 간 일치도

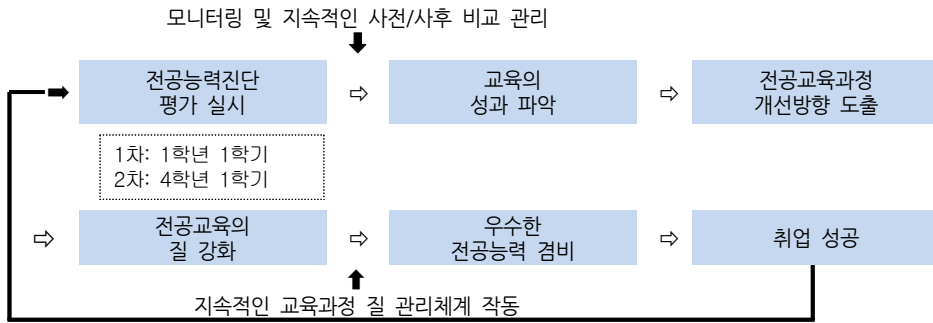
하나의 답안을 2명의 채점자가 채점하였고, 이들 두 점수 간의 단순 상관계수를 구하여 채점자 간 일치도를 확인하였다. 그 결과 수행형 문항에 대한 채점자 간 상관관계는 <표 6>과 같으며, .82로 비교적 양호한 신뢰도를 보여 채점자들 사이에서 평가점수가 크게 차이 나지 않는다는 것을 알 수 있다.

〈표 6〉 채점자 간 일치도

	채점자1	채점자2
채점자1		.82**
채점자2	.82**	

*P < .05, **P < .01

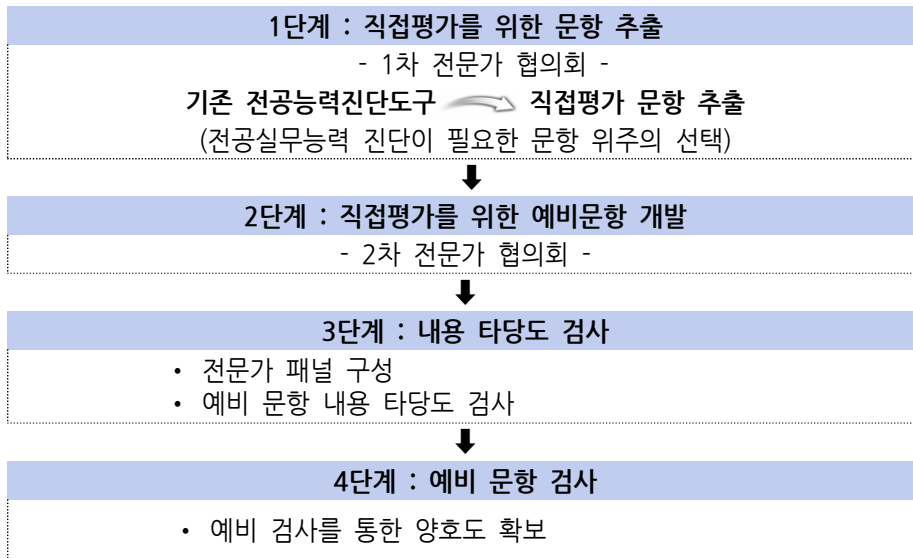
마지막으로 이상의 전공능력진단도구 개발 및 결과분석을 바탕으로 지속적인 교육과정 질 관리체계를 구축하여 전공교육의 질을 강화하고 학생들에게 우수한 전공능력을 함양시켜 취업 경쟁력을 강화하고 대학의 교육혁신을 이끌 수 있도록 환류 체계 구축을 제안한다. 이를 그림으로 나타내면 [그림 5]와 같다.



[그림 5] 직접평가 방식의 전공능력진단도구 활용 및 환류 체계

V. 요약 및 제언

본 연구는 대학생의 전공별 필수 능력의 하위 영역에서 직접평가를 위한 문항 및 채점 기준 등 진단도구를 개발하고자 한 연구이다. 연구진이 각 전공(학과)별 진단도구를 모두 개발하여 제시하기에 각 전공이 가지고 있는 고유한 학문체계와 특수성을 이해하고 적용하는 데 한계가 있어 A 대학교 교직과(정)를 연구 대상(전공)으로 삼고, 일련의 연구과정을 통해 각 전공별 진단도구 개발을 위한 표준화 모델을 [그림 6]과 같이 제시한다.



- 예비검사 대상자 선발
- 안면 타당도 검사
- 준거 타당도 검사
- 신뢰도 검사

↓ 수정·보완

직접평가를 위한 수행형 전공능력진단도구 완성

[그림 6] 직접평가를 위한 수행형 전공능력진단도구 개발 표준화 모델

이는 다음과 같은 특성을 지닌다.

첫째, 대학생이 해당 전공 분야의 직업인으로서 갖추어야 할 핵심적인 전공능력을 측정하고자 하므로 업무수행에 필수적으로 요구되는 인지적 능력과 실천적 능력을 측정한다. 둘째, 전공분야와 관련된 다양한 분야에서 요구되는 능력 중 공통적으로 요구되는 필수 전공능력을 측정한다. 셋째, 학생 개인의 특성에 대한 최대 정보를 얻기 위해 단편적 지식보다는 종합적 능력을 측정할 수 있는 선다형, 서답형, 자기 점검표(Likert 척도) 등의 혼합형 검사로 구성하고, 특히, 실천적 능력은 수행평가 방식으로 설계한다. 넷째, 신입생을 대상으로 1학년에 검사를 실시하고, 동일학년이 4학년이 될 때 검사를 실시함으로써 대학의 전공교육을 통하여 개인의 전공능력이 변화한 정도와 대학교육의 성과를 파악한다. 다섯째, 개발된 직접평가를 위한 수행형 전공능력진단도구를 효과적으로 활용할 수 있도록 향후 직접평가를 위한 기반 요소를 추출하여 제안한다. 제안된 제반 요소는 크게 시스템 도입·운영, 문항 모듈 구성 방안, 채점 인력 구축을 제시한다.

본 연구의 기대효과는 다음과 같다.

첫째, 전공능력진단도구 직접평가 검사문항을 개발함으로써, 학생들의 전공능력을 보다 정확히 평가할 수 있는 도구를 마련할 수 있다. 둘째, 진단 결과를 토대로 학과 및 대학의 교육의 질을 강화하고 교육과정을 정비하며 그 결과를 추적·평가할 수 있는 시스템 개발의 기반을 마련할 수 있다. 셋째, 매년 전공능력진단 검사를 시행함으로써 전공 교육과정 개선을 위한 기초자료를 확보하고, 교육혁신의 성과지표로 활용할 수 있다.

마지막으로 본 연구의 제한점으로는 연구의 범위를 새롭게 진단도구를 개발하는 경우가 아닌, 기존에 개발된 진단도구를 고도화(수정·보완)하는 경우를 연구 범위로 한정 지었다는 점이다. 그 이유는 전공능력진단도구가 대부분의 대학에서 개발되어 평가 결과를 통해 학과(전공) 학생들의 전공능력을 진단하고, 그 결과를 교육과정 개발에 환류하는 등 이미 개발과 활용이 고착된 상황이므로 기존 진단도구를 수정·보완하고자 하는 상황에서 도움을 줄 수 있는 연구 결과를

제시하고자 한 것이다. 하지만 교육혁신의 미명 하에 융복합전공, 다전공 등 신규 학과(전공)가 많이 생겨나는 상황에서 전공능력진단도구 역시 새로이 개발될 수 있다는 점을 고려하지 못한 점이 본 연구의 아쉬운 점이라 할 수 있다. 따라서 이를 후속 연구 주제로 삼아 한계점을 보완하고자 한다.

이상의 표준화 모델 제안과 특성을 통해 학생들의 전공능력을 보다 정확히 측정하기 위한 방법으로 수행형의 직접평가 방식을 제안하고 이를 표준 매뉴얼로 제시하여 다양한 전공에서 직접평가 방식의 문항 개발 시 활용 가능한 자료가 되기를 기대한다.

※ 논문 투고일: 2024. 1. 1. ※ 논문 수정일: 2024. 2. 27. ※ 게재 확정일 : 2024. 2. 28.

〈참고문헌〉

- 강승혜 외(2006). **한국어 평가론 : 한국어 교육 총서**. 국립국어원 한국어세계화재단. 서울: 태학사.
- 교육부(1998). **초·중등학교 교과별 수행평가의 실제(1) 총론**. ORM 99-3-1 세종: 교육부
- 교육과학기술부(2008). **국어과 교육과정 해설**. 서울: 교육과학기술부
- 권정화(2020). **초기 영유아 언어발달 평가에서 부모보고식 평가와 직접평가 간의 비교 연구**. 대구대학교대학원. 석사학위청구논문.
- 김수동(2018). 고등교육에서의 역량기반 교육과정을 위한 평가연구. **예술인문사회 융합 멀티미디어 논문지**, 8(5), 759-767. 인문사회과학기술융합학회.
- 김필성(2018). 한국 D대학교의 역량기반 학과 교육과정의 성과평가 방법 개발에 대한 사례 연구. **예술인문사회융합멀티미디어 논문지**, 8(3), 221-229. 인문사회과학기술융합학회.
- 김현우, 정지원 (2012). **중등교육 직무능력을 평가하는 교사수행평가도구의 개발 및 검증**. **교육평가연구**, 21(2), 163-173. 한국교육평가학회
- 박상희, 이희정, 김미경(2016). 학생 자기주도학습 능력진단을 위한 포트폴리오 평가방안 연구. **한국평가평가학회지**, 23(1), 한국평가평가학회
- 박주현, 강미경(2016). 과학영재학교 학생들의 문제해결과제 수행시 평가방법 연구. **교육과정 평가연구**, 19(4), 142-66. 한국교육과정평가학회
- 양홍권(2016). **대학생용 학습역량 검사도구 개발 연구**. **평생교육·HRD연구**, 12(1), 29-64. 숭실대학교 한국평생교육·HRD연구소.
- 이재현, 홍성일(2004). 고등학교 문제해결능력 평가를 위한 개념구조 및 실증연구. **교육평가연구**, 17(3), 29-64. 한국교육평가학회
- 이재호(2022). **초등특수교육 전공능력진단도구 개발 연구: 예비 특수교사를 중심으로**, **학습자 중심교과교육연구**. 22(7), 253-271, 학습자중심교육학회.
- 서금택, 윤용관(2019). 대학 핵심역량과 학과 전공역량 중심의 교육과정 개발 사례. **예술인문사회 융합 멀티미디어 논문지**, 9(7), 163-173. 사단법인 인문사회과학기술융합학회.
- 손진아, 이정희, 김선형(2021). 미용 전공 학생의 전공능력 진단 도구 개발 및 타당성 연구. **한국미용학회지**, 27(3), 577-592, 한국미용학회.
- 신민희(2012). 학습성과 수행평가를 위한 루브릭 개발과 적용에 관한 연구. **교육공학연구**, 15(5), 108-118. 한국교육공학회.
- 이수진(2011). 쓰기 직접 평가를 위한 텍스트 분석 평가의 실행 방안. **작문연구**, 13(13), 142-166. 한국작문학회.
- 이화영(2021). **대학에서의 전공능력 도출 모형 개발 및 타당성 검증**. 건국대학교 대학원 박사학위 청구논문.
- 최현철, 김수동, 김경표(2020). 전공역량 기반 교육과정 체계 고도화 연구. **문화와 융합**, 42(2), 423-443. 한국문화융합학회.
- 한국교육심리학회(2000). **교육심리학 용어사전**. 서울: 학지사
- 허지원(2021). ‘역량기반 교육’ 개념체제의 이해와 조건 탐색을 통한 과제 제시. **핵심역량교육 연구**, 6(1), 1-22. 한국핵심역량교육학회.
- Garrison, D. R., & Ehringhaus, M. (2007). *Blended learning and online tutoring: Planning learner support and activity design*. *Blended Learning in Higher Education: Framework, Principles,*

and Guidelines.

Sadler, D. R. (1989). *Formative assessment and the design of instructional systems*. Instructional Science.

Birenbaum, M. (2006). *Assessment preferences and their relationship to learning strategies and orientations*. Assessment & Evaluation in Higher Education.

Airasian, P. W. (1994). *Assessment in the classroom*. The Phi Delta Kappan.

(사이트)

교육부. 대학혁신 지원 방안 <https://www.moe.go.kr/boardCnts/view.do?boardID=294/>

국어사전. <https://wordrow.kr>

광주광역시 교육청 연수원 <https://eduup.gen.go.kr>

〈Abstract〉

Development and Standardization of Major Competency Assessment Tools: Transitioning from Indirect Evaluation to Performance-Type Direct Evaluation questions Focusing on the Case of Education Department at University A

Heo, Jiwon¹

This study criticizes the fact that most university major competency assessment tests are conducted using the self-check indirect evaluation methods. In response to that, this study aims to develop a direct evaluation method for major competency assessment tests by incorporating performance-based questions as an alternative. To achieve this, the study focuses on the education and teaching departments conducting preliminary test development based on previous research, case analysis, expert interviews, advisory committees' recommendations, and an understanding of the major background. Preliminary tests were developed and a pilot test was conducted, which involved the education and teaching departments students to assess its validity. The results were used to modify and improve the test items, and the content experts evaluated the content validity.

The content validity index (CVI) was high at 0.96, with 5 out of 6 performance-based categories with a CVI of 1.0 (83.3%) and 1 category with a CVI of 0.87 (16.7%). The study then compared the developed performance-based test categories for direct evaluation with the self-check categories of the existing assessment to derive meaningful implications.

Based on the results of this study, a proposal is made to use performance-based direct evaluation methods to more accurately measure students' the major competencies. This study also presents a standardized manual for the entire research process. We hope that the results of this research will serve as valuable resources for developing direct evaluation categories for various majors

Keywords : Major competency diagnostic tool, direct assessment, indirect assessment, performance-based questions, teaching competency diagnostic tool

1. Associate Professor, Nambu University, hjw123@nambu.ac.kr