

루브릭 기반 국가기술자격 실기평가 연구

A Study on Practice Test of The National Technical Qualification based on Rubric

윤관식*, 최명란**

Gwan-Sik Yoon*, Myung-Ran Choi**

요약

오늘날 산업 현장은 조직에 능동적으로 적응하기 위해서 종업원들의 다양한 개성을 존중하고, 창의성 및 문제 해결 등을 최대한 발휘할 수 있는 역량들을 요구한다. 이러한 역량들은 실습교육을 통한 수행과정에서 주로 획득될 수 있기 때문에 이러한 수행을 적절하게 판단할 수 있는 평가방법의 개선과 새로운 평가체제를 필요로 한다. 루브릭은 교수-학습과정과 평가를 통합시킬 수 있고, 교사와 학생간의 의사소통의 역할을 담당하고, 교수-학습 전 과정을 안내한다. 본 연구에서는 직업훈련기관의 교수-학습과정에서 빈번하게 이루어지는 실습교육과 국가기술자격시험에서 실시하는 실기시험의 연계와 객관성 확보를 위한 평가도구를 루브릭을 통해 모색하고자 하였다.

이에, 본 논문에서는 국가자격시험의 실기시험 평가도구의 개발 시 고려되어야 하는 사항들을 제시하고, 루브릭의 개발절차를 제시하였다. 또한 현재 국가자격시험에서 적용하고 있는 루브릭의 문제점과 그 개선책을 논의하였다.

Key Words : Rubric, National technical qualification, Practical test

ABSTRACT

The purpose of this study is to identify and develop a rubric for assessing the practice test in national technical qualification with validity and objectivity. There are three specific purposes for this study. First, we suggest some issues to consider when we develop a rubric. Second, we suggest the general procedure and main factors in developing a rubric. Third, we identify some problems and suggest some improvement measures for the rubric in national technical qualification examination. The procedure for developing a rubric is as follows; decisions of task, the relation with goal, classify of performance level, and validity. Also, there are five main factors as follows; contents, criteria, practicality, credibility, and objectivity.

* 한국기술교육대학교(gwansik@kut.ac.kr)

** 한국기술교육대학교(bestnancho@naver.com)

교신저자: 윤관식 한국기술교육대학교(gwansik@kut.ac.kr)

접수일자 : 2012년 10월 26일

수정일자 : 2012년 11월 28일

확정일자 : 2012년 12월 7일

I. 서론

직업능력개발을 위한 직업훈련의 교육내용은 크게 이론과 실습으로 구분되며, 실습에 대한 강조는 직업훈련의 목적과 방향을 지시한다. 그러나 직업훈련에서 실습이 직업생활에 필요한 단순 기능을 익히는 기능교육이라기 보다는, 장기적으로 직업생활의 질 향상을 위한 실습교육을 통하여 끊임없이 관찰, 조사하고, 연구와 탐구를 통하여 현장의 문제를 해결할 수 있는 자기주도적 학습을 지향하는 교육·훈련이어야 한다는 것이다.

오늘날 산업 현장은 조직에 능동적으로 적응하기 위해서 종업원들의 다양한 개성을 존중하고, 창의성 및 문제해결 등을 최대한 발휘할 수 있는 역량들을 요구한다. 이러한 역량은 교수자 주도에 의한 지식 전달보다는 학습자가 주체가 되어 스스로 학습하는 자기주도적 학습이 이루어져야만 획득될 수 있다. 이러한 역량들은 실습교육을 통한 수행과정에서 주로 획득될 수 있기 때문에 이러한 수행을 적절하게 판단할 수 있는 평가방법의 개선과 새로운 평가체제를 필요로 한다.

새로운 평가체제에서 평가할 내용은 과제의 구체화를 통해서 표현된다. 과제에 대한 학습자의 수행이나 결과는 루브릭(rubric)에 의해서 점수화된다(Arter, 2000). 즉 루브릭은 학습자들의 수행 결과로써 획득한 지식 혹은 능력을 어떠한 기준으로 평가할 것인지를 계획한 평가도구이다. 특히 루브릭은 실습과제에 대한 학습자의 수행 능력을 평가할 수 있도록 평가내용과 평가할 수 있는 수행준거, 그리고 수행 수준을 제시한 객관화된 평가도구이다. 한지영(2004)은 루브릭이 교수-학습 전 과정을 안내한다고 하였다.

2012년 현재 국가기술자격법에 따라 발급되는 국가기술자격 종류는 26개 분야 512개 종목이다. 기술·기술 분야가 481종목, 서비스 분야가 31종목이다. 등급별로는 기술·기능 분야 최고 등급인 기능장·기술사가 각각 27개와 84개, 기사 103개, 산업기사 110개, 기능사 157개다. 서비스 분야는 1~3급 종목이 23개, 단일 등급 종목이 8개다. 실기시험 없이 이론과 면접으로 자격을 부여하는 기술사를 제외하고는 모든 국가기술자격은 실기시험을 통하여 수험생을 평가한다. 이는 직업훈련 교육기관의 교수-학습 수행과정에서 실기교육은 매우 중요하다는 것을 반증한다. 또한 국가자격검정에서 실기시험 평가에 대

한 객관성을 확보하기 위한 연구는 거의 전무하다.

본 연구는 직업훈련기관의 교수-학습과정에서 빈번하게 이루어지는 실습교육과 각종 국가자격시험에서 실시하는 실기시험의 연계와 객관성 확보를 위한 평가 도구를 모색하는 것이다. 수행을 강조하는 직업훈련기관의 실습교육에서 학습자 주도적인 방향을 모색하는 것과 동시에 실습교육의 결과로서의 수행 작품에 대한 평가의 객관성 확보를 위한 평가 방법으로서의 루브릭에 대한 개념적 이해 연구는 후속되는 연구를 위하여 필요하다.

구체적인 연구문제는 다음과 같다.

첫째, 국가기술자격시험의 실기시험 평가도구의 개발 시 고려되어야 할 사항을 제시한다.

둘째, 국가기술자격시험의 실기시험의 평가도구로서의 루브릭을 구성하는 요인과 구성하는 방법적인 절차를 확인하고 제시한다.

셋째, 현재 국가기술자격시험에서 활용하고 있는 채점기준표(루브릭)의 문제점과 그 개선책을 논의한다.

II. 루브릭(rubric)의 개념적 이해.

1. 루브릭의 개념

루브릭(rubric)은 등급평가기준의 집합이다. 즉, 학습자가 과제를 수행할 때 나타내는 반응을 평가하는 기준의 집합이다. ‘루브릭’이라는 용어는 라틴어인 ‘Rubrica(Red Earth)’에서 유래하였고, 원래 법을 토대로 어떤 상황을 공정하게 판단하는 과정을 서술한 것을 말한다(Wiggins, 1996:154). 굤리치(Goodrech, 1996)는 교육학에서의 루브릭을 ‘하나의 과제나 활동을 통하여 학습자들이 만든 작품을 준거에 의해 목록화하고 등급을 결정하고, 점수화하기 위한 도구’라고 하였다.

보통 루브릭은 학습자의 수행을 평가하기 위한 범주를 나타내는 과제영역(dimensions)이 있으며, 과제의 각 영역은 수준별 표로 구성된다. 표는 어떤 경우에 그 수준에 해당되는지를 상세히 기록하고 있다.

허먼과 아시바커 그리고 윈터스(Herman, Aschbacher & Wintes, 1992)는 첫째, 수행을 판단하는 기준이 되는 과제영역은 학습자들이 수행이나 산출물을 어떠한 특성 혹은 영역에 따라 평가할 것인가에 관한 것이며 둘째, 각 특성 혹은 영역은 차원을 변별하고 평정하는 수치나 범주(scale or values)를 몇 개로 구분할 것인가를 결정해야 하며 셋째, 각각

의 수행을 판단할 기준을 기술하여야 한다고 하였다.

이와 같은 루브릭은 학습자의 수행과정과 성취정도를 보다 객관적으로 평가하기 위한 도구로서의 역할과 학습을 촉진하는 교수의 도구로서의 역할을 하게 된다. 즉, 평가자들에게는 평가 장면에서 각각의 수행 수준의 특징에 대한 정보를 명세화 하여 제공함으로써 학습자의 수행을 효과적으로 평가할 수 있도록 해주며, 학습자들에게는 본인의 수준이 어느 정도인지에 대한 분명한 피드백을 제공해 줌으로써 앞으로 수행능력을 향상시키기 위하여 무엇이 필요한지를 구체적으로 파악할 수 있도록 해 준다. 더 나아가 학습자들은 평가에 자연스럽게 참여하게 되고 그 과정에서 학습의 초점이 무엇인지 분명히 알고 자기주도적 학습으로 진행할 수 있도록 도와준다.

루브릭이 효과적이라면 수행 역량을 전체적으로 조망하고 각각의 수행 수준에 맞추어서 세분하고 자세하게 묘사해야 한다. 위긴스(Wiggins, 1996)는 루브릭의 전반적인 특징을 다음과 같이 설명하였다. 첫째, 루브릭은 수행자의 수행들을 효과적으로 변별할 수 있도록 타당하고 신뢰성이 있어야 한다. 둘째, 루브릭의 각 영역에서 수준별 기술(descriptor)은 각 수행의 수준을 최적으로 기술하고, 가장 핵심적인 특성을 반영한 용어를 사용한다. 셋째, 루브릭을 통하여 학습자의 장점을 인정하고 단점을 수정하게 되는 질적 정보이어야 한다. 넷째, 척도의 가장 높은 점수는 가장 우수한 본보기가 되기에 루브릭은 준거지향적이다. 따라서 가장 높은 수준에 아무도 도달하지 못할 수도 있고, 대부분의 학습자가 가장 낮은 수준일 수도 있다. 다섯째, 척도 상에서 각 수준에 대한 기술(descriptor)의 변화는 연속선상에서 설명되어야 한다. 여섯째, 루브릭이 구체적일수록 보다 타당한 결과를, 단일하고 간단한 루브릭일수록 신뢰할만한 결과가 나타난다.

루브릭은 정확하게 무엇을 평가할 것인지 확인하고, 평정 척도와 점수 분포를 분명히 한 후 각각의 수행 수준에 대해 그 수행 수준의 독특한 특징을 객관적으로 서술해야 한다. 검토 단계로서 수행 수준의 연속성을 제대로 고려하고 있는지, 동일한 기준을 적용하고 있는지, 수행 수준간의 구별은 명확한지, 각 수행 수준에 대한 점수 부여는 신뢰성이 있는지를 검토하여 수행하여야 한다.

2. 교육현장에서의 루브릭의 활용

본 연구의 분석 대상인 루브릭은 국가기술자격시험을 준비하기 위한 직업훈련과정에서 학습자의 수

업참여를 촉진하며 교수적 지원 도구로서의 역할이 강화되도록 개발할 평가도구이다.

직업훈련기관의 성과가 자격증취득률, 취업률, 중도탈락률로 평가되는 상황이다. 따라서 직업훈련기관 등과 같이 국가기술자격시험에 맞춰진 교수-학습활동(1년 내지 6개월 과정)에 루브릭이라는 평가도구가 갖는 아래와 같은 역할들이 충분히 반영된다면 보다 개선된 질 높은 교육과정 운영이 될 것이다.

아더와 맥타이(Arter & McTighe, 2001)는 학습의 결과물이나 수행을 판단할 수 있도록 해주는 준거에는 체크리스트(checklist)와 수행리스트(performance list), 그리고 루브릭이 있지만, 그 중 루브릭이 학습 결과물과 복잡한 수행을 묘사하고 정의할 수 있는 준거와 수행수준을 제시함으로써 각 항목에 배점을 주는 방식의 체크리스트와 수행리스트의 단점을 보완할 수 있다고 하였다.

(1) 루브릭을 통한 학습전략

굳리치(Goodrich, 1999; 노원경, 2008)는 ‘루브릭을 왜 사용하는가’에 대한 질문에 다음과 같이 제시하였다. 첫째, 루브릭은 평가에 대한 준거와 기준을 비밀로 하던 관례에서 벗어나 교수-학습에 대한 기대를 높인다. 둘째, 학습자 스스로 학습을 조절하고 모니터링 하는 작용에 대하여 자기평가가 가능하게 한다. 셋째, 학습자들이 수행의 준거를 잘 이해할 수 있도록 함으로써, 바람직한 수행을 가능하게 한다고 하였다. 이러한 이유로 루브릭은 학습자의 초인지능력을 개발하는데 도움을 주며, 자기평가를 통해서 학습자들의 자기주도적 학습태도를 향상시킬 수 있으며, 학습자에게 자기지향적인 피드백을 제공해주고, 학습자의 비판적 사고력을 개발하는데 도움을 준다.

브라운(Brawn, 1988)은 초인지의 기능을 집행기능(executive function)으로 표현하면서 이러한 초인지를 통해서 ‘나는 무엇을 알고 있는가?’, ‘어느 정도 잘 할 수 있는가?’, ‘어떤 순서로 이 과제를 수행할 것인가?’, ‘지금까지 잘 해왔는가?’, ‘앞으로 잘 할 수 있을 것인가?’와 같은 역할을 수행할 수 있다고 한다. 굳리치에 의하면 루브릭이 이러한 초인지를 길러 줄 수 있는 학습전략이라고 말하는 것이다.

현재의 교수-학습이 의도한 대로 진행되고 있는지, 앞으로 어떤 과정을 거쳐야 하는지, 자격증을 취득하려면 어느 수준에 올라야 되는지를 루브릭을 통해 제공 받게 될 것이다.

(2) 학업적 자기효능감

학업적 자기효능감(Academic self-efficacy)을 높일 수 있는 평가도구인가? 루브릭은 자기효능감을 높일 수 있는 도구라고 말할 수 있다. 자기효능감(自己效能感)은 특정한 문제를 자신의 능력을 통해 성공적으로 해결할 수 있다는 자기 자신에 대한 신념이나 기대감이다. 높은 자기 효능감을 과제에 대한 집중과 지속성을 유지해 성취수준을 높일 수 있다.

반두라(Bandura, 1986)는 학업적 자기효능감을 학습자가 학업적 과제 수행을 위해 필요한 행위를 조직하고 실행해 나가는 자신의 능력에 대해 내리는 판단으로 학습자가 새로운 상황에서 새로운 지식과 기술을 학습하고 수행할 수 있도록 유도하는 원동력이라 하였다. 이러한 자기효능감은 목표한 과업을 달성하기 위하여 필요한 행동을 계획하고 수행할 수 있는 자신의 능력에 대한 자신감을 갖게 되며, 가장 중요한 것은 과거 수행에서의 성취경험에서 나온다고 하였다. 효능감을 평가하면서 학습자들은 역량, 투입된 노력, 과제의 난이도, 교사의 지원, 그리고 성공과 실패의 수와 패턴을 요소로써 고려하고 있다.

학습자가 자신의 수행을 관찰하고 자신이 목표하는 기준에 비추어 자신의 수행을 판단하여, 긍정적이라면 새로운 목표를 설정하고, 부정적이라면 자신의 목표에 도달하기 위해 부가적인 행동을 이끌 수 있기 때문이다. 즉, 자기효능감이 높은 학습자와 같이 자신이 통제하고 다룰 수 있다고 생각하는 도전적인 과제들을 루브릭의 명시된 수준에 따라 밝아나갈 수 있을 것이다.

(3) 학습자의 수행평가

김소연(2009)은 루브릭의 가장 큰 장점은 교수·학습활동과 평가의 통합이라고 하였다. 교수·학습활동과 평가의 통합은 교사와 학습자 모두에게 지속적인 반성적 사고의 기회를 제공하며, 장기적인 측면에서 보다 나은 수업활동을 통해 교사와 학습자가 협력자 관계가 될 수 있는 구조를 제공한다. 따라서 평가는 교사와 학습자 모두에게 교육목표를 공유할 수 있도록 계획되어야 한다고 하였다. 교육의 목적과 목표, 교수 방법, 그리고 기대효과가 통일이 교수-학습의 보다 질 높은 교육경험을 제공한다는 것이다.

윤관식(2011)은 수행평가란 ‘학습자의 작품이나 활동을 직접 관찰하고, 관찰된 결과를 전문적(주관적)으로 판단함으로써 이루어지는 평가’라고 정의하였다. 평가의 대부분이 결과에 근거해 가치판단이 내려지는 가운데, 루브릭은 점수위주의 단순한 평가 피드백을 벗어나 과정을 포함하여 보다 객관적이고 학습

상황에 대한 준거별 자세한 내용까지 제공받을 수 있다. 수행평가는 전문적(주관적)인 평가라는 점에서 객관성과 공정성을 확보하는 것이 중요하다. 이에 많은 연구자들이 루브릭을 통해 이 부분의 문제를 해결하고자 하였다. 루브릭은 평가자들이 안정적이고 객관적으로 신뢰받을 수 있는 평가를 하는데 도움을 준다고 보기 때문이다.

또한 루브릭은 수행평가 중에서도 수준 높은 학습자의 능력을 이끌어 낼 수 있는 평가도구라 할 수 있을 것이다. 김정덕(2009)은 학습자들 간에 성적이나 능력이 차이가 나더라도 루브릭의 평가준거들이 명확하게 설명해 높은 급간들은 학습자 자신의 상태와 교수자가 요구하는 최상의 수준과 비교하여 판단하게 한다는 점에서 교수자가 학습자들을 개별적으로 지도하는데 보다 도움이 된다고 하였다. 평가와 교수-학습의 연장선에서 루브릭이 교수자에게 지도의 편리성을 가져다줄 수 있다는 것이다.

III. 연구의 방법

1. 문헌연구

현재 루브릭은 전통적인 평가에서 사용하는 점수(score) 혹은 등급(grade)에 상응하는 개념으로 학습자가 수행 과제에서 드러낸 수행 결과물의 수준을 규명하고 판단하기 위한 채점 방법으로서의 평가도구로 활용되고 있다. 또한 루브릭은 학습자에게 교수의 수행 과제에 대한 기대치를 사전 안내한다는 점에서 실습수업에 학습도구로 적용되고 있고, 이러한 필요에 의해 개발되고 있다. 본 연구에서는 이러한 루브릭에 관한 내용들을 문헌연구를 통하여 고찰함으로써, 국가자격시험의 평가도구로 사용되고 있는 루브릭에 초점을 맞추어 평가도구의 개발 시 고려되어야 할 사항을 제시하였다. 또한 ‘어떤 목적 하에 루브릭을 활용하는가?’에 중점을 두어 루브릭을 구성하는 요인과 구성하는 방법적인 절차를 제시하였다.

이 연구를 진행하기 위하여 K공단에 개발절차와 활용에 관한 루브릭의 모든 자료를 요청하였지만, 연구에 참고할 수 있는 자료는 평가도구로 실제 사용되고 있는 5개의 국가기술자격별 실기시험 채점기준표(방수기능사, 용접기능사, 광고도장기능사, 이용사, 자동차정비기능사)이다. 채점기준표는 1. 채점상의 유의사항, 2. 채점기준표, 3. 채점표로 구성되어 있다. 제한된 자료를 통해 국가기술자격 실기시험의 평가도구를 점검한다는 것은 무리가 따르겠지만, 루브릭

을 개발하는데 있어 전제되어야 할 사항과 개발단계를 문헌연구를 통해 제시함으로써 사용되고 있는 평가도구의 개선을 꾀하고자 하였다.

그림 1. 국가기술자격 실기시험 채점기준표 및 채점표(자동차 정비기능사)
Fig. 1. Rubric of national technical qualification test

2. 설문조사

문헌연구를 통하여 제시한 검토지침을 가지고 직업능력개발훈련교사 42명을 통해 국가기술자격 실시시험 채점표에 대한 설문조사를 실시하였다. 현재 직업훈련기관에 종사하지는 않지만 교사경력이 있는 분들을 포함하고 있다. 이 중 25(60%)명의 대상자는 지속적으로 국가 자격증 심사위원으로 참여하고 있는 분들이며, 대상자들 중 95%이상이 교육기관에서

5년 이상의 경력을 가지고 있다.

대상자들에게 K공단에서 사용하고 있는 국가기술 자격 실기시험 채점표의 ‘채점상의 유의사항’, ‘채점기준표’, ‘채점표’를 제시하여 설명한 후, 설문지에 응답하도록 하였다. 하지만 설문대상자의 훈련직종(전공)에 해당하는 모든 국가기술자격 실기시험 채점표를 제시하지 못하였기에 직종별이 아닌 전체 채점기준표에 대한 의견으로 수렴하였다. 설문지는 일반사항과 국가자격시험 루브릭의 ‘내용적 측면’과 ‘준거의 명료성’, ‘실용성’, ‘신뢰성’, 그리고 ‘객관성의 정도’를 묻는 항목으로 구성되었다. 설문지는 스티긴스(Stiggins, 2005)¹⁾의 루브릭 검토지침을 참고하여 ‘객관성의 정도’항목을 추가하여 재구성하였다.

표 1. 설문대상의 일반적 특성
Table 1. General characteristic of those questioned

구분	항목	빈도 (명)	백분율 (%)
성별	남	20	47.6
	여	22	52.4
현재 소속기관	민간직업훈련기관	7	16.7
	공공직업훈련기관	18	42.9
	사설 학원	3	7.1
	대학기관	5	11.9
	기업체	3	7.1
교육기관의 제직기간	기타	6	14.3
	5년 미만	2	4.9
	5-10년 미만	14	34.1
	10-15년 미만	9	22.0
	15년-20년 미만	12	29.3
	20년 이상	4	9.8
훈련 직종	기계	7	16.7
	전기, 전자	3	7.1
	컴퓨터, 사무자동화	4	9.5
	정보통신	1	2.4
	서비스(미용 등)	10	23.8
	조리, 제고제빵	7	16.7
	기타	10	23.8
국가 자격	기능사	16	41.0
	산업기사	6	15.4
	기사	8	20.5
	기능장	7	17.9
국가 자격증 심사위원 참여	기술사	2	5.1
	있음	25	59.5
	없음	17	40.5

IV. 루브릭 개발의 고려사항과 절차의 제시

문헌연구를 통해 국가자격시험 실기시험 평가도구의 개발 시 고려되어야 할 사항과 평가도구로서의

1) 출처: Stiggins(2005). Student-involved assessment for learning, pp. 159-160.

루브릭을 구성하는 요인과 구성하는 방법적인 절차를 제시하였다.

1. 루브릭 개발에 따른 고려사항

본 연구에서 제시하고자하는 루브릭을 구성하는 요인과 구성하는 방법적인 절차를 결정하기 위해서는 다음과 같은 사항들이 고려되어야 할 것이다.

첫째, 목표에 따른 타당한 항목들이 평가준거로 선정되어야 한다. 교육과정에서 나타난 학습결과를 모두 평가의 대상으로 삼는 것은 불가능하다. 그러므로 학습결과의 대표적인 표본이라 할 수 있는 목표를 근거로 한 평가준거가 마련되어야 한다.

둘째, 평가의 준거가 되는 목표는 피평가자들과의 상호작용을 통한 공유된 목표이어야 한다. 교육목표는 학습활동을 통해서 변화 또는 획득되기를 바라는 학습자의 의도된 성과를 반영시킨 것이다. 따라서 학습자가 해야 할 일이 무엇인지, 성취해야할 목표가 무엇인지가 충분한 논의를 통해 인지되었을 때, 학습자가 학습에 참여하는 동기나 관심이 높아지고, 도달해야 할 수준이 명확해 질 수 있다.

셋째, 가장 잘한 수준과 가장 못한 수준의 준거가 타당하여야 한다. 자격증을 취득하기 위한 평가도구라고 한다면 자격증에 근접한 일정한 수준의 학습자들이 사용하는데 있어 그 등급에 준한 수행의 묘사가 적절해야 한다.

넷째, 평가항목 수가 적정하여야 하고, 수준별 등급의 격차가 타당해야 한다. 루브릭의 평가는 여러 평가항목에 있어서 피평가자의 능력이 해당되는 각 수준에 부여된 점수의 합산으로 이루어진다. 따라서 각 항목의 수준별 등급의 격차가 타당해야 종합적인 점수에서 항목별 능력차이가 있는 피평가자들 간의 형평성을 유지할 수 있다.

다섯째, 수행의 수준별 등급은 다양한 질적 특징이 묘사되어야 하며, 그것을 측정할 수 있는 지표가 마련되어야한다. 학습과정과 작품, 결과물이 동일한 수준으로 평가될 수 있도록 다양한 질적 특징으로 묘사된 각 수준별 등급은 가능하면 양화할 수 있도록 명료하고 객관적으로 규정하여야 한다.

여섯째, 각 단계의 지표는 자기평가의 도구로 활용할 수 있도록 구체적이어야 한다. 지표자체가 피드백의 역할을 수행하여야 한다. 이는 루브릭의 각 단계의 지표가 자기주도적 학습의 동기부여로 이어질 수 있도록 한다는 것이다. 학습자는 평가결과를 근거로 하여 그 다음 단계로 무엇을 하여야 하는지 쉽게 파악할 수 있어야 한다.

2. 루브릭 개발을 위한 절차 및 준거의 제시

본 연구에서 제시하는 루브릭의 개발단계는 루브릭이라는 평가도구를 ‘어떤 목적 하에 활용할 것인가?’를 고려하여 그림 2에서와 같이 제시하고자 한다. 첫째, 산업현장에서 필요로 하는 역량들을 실습교육을 통한 수행과정에서 획득하고자 할 경우, 이 교육과정에서의 수행을 적절하게 판단할 수 있는 평가의 도구로서의 루브릭의 개발이다. 둘째, 각종 국가자격시험에서 실시하는 실기시험의 평가의 객관성 확보를 할 수 있도록 수행수준의 지표를 제시할 수 있도록 함으로써, 시험을 준비하는 교수-학습과정에서 활용하도록 하는 루브릭의 개발이다.

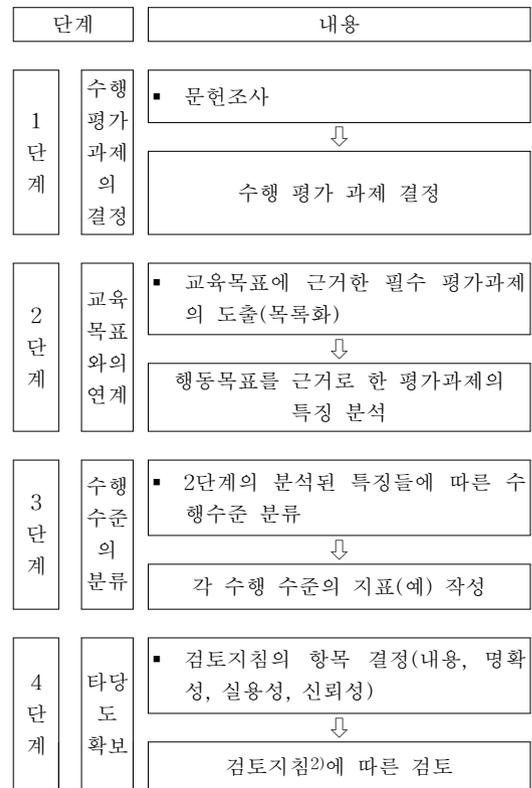


그림 2. 루브릭 개발의 단계
Fig. 2. The process of rubric development

1단계는 수행평가 과제의 결정단계이다. 본 연구에서의 루브릭은 국가자격시험에 대비한 교육과정의 실습교육을 전제하기에 ‘처리하거나 해결해야할 문제’인 과제가 그 중심이 된다. 이 과제는 학생들의

2) 표 2의 루브릭 검토 지침

표 2. 루브릭의 검토 지침
Table 2. Review guidelines of rubric

항목	내용
내용적 측면	1. 루브릭의 항목이 관련 교육 기준(교육목표)에 부합되는가?
	2. 가장 잘한 수준과 가장 못한 수준까지의 수준별 준거가 적절하게 제시되었는가?
	3. 수행의 수준별 등급은 핵심적 특성들이 잘 묘사되었는가?
준거의 명확성	1. 각 평가자간 동일한 수준으로 평가될 수 있도록 양화된 측정지표가 마련되어 있는가?
	2. 양화될 수 없는 항목들의 각 수준은 정확한 용어로 정의되고, 자세히 묘사되어 있는가?
	3. 익숙해진다면 오랜 기간 일관성 있게 사용할 수 있는 평가도구로써 제작되어 있는가?
실용성	1. 교수자와 학습자가 공유하고 적용하기에 적절한 수의 평가항목인가?
	2. 루브릭의 평가결과를 바탕으로 다음 교수-학습의 활동으로 무엇을 하여야 하는지 자연스럽게 이어질 수 있는가? (교수-학습의 활동의 단계 제시)
	3. 수행이나 작품개선을 위한 피드백의 역할을 수행할 수 있는가?(자기평가의 도구)
신뢰성	1. 개별평가자들이 동의할 수 있도록 학습자의 수행 수준에 대한 루브릭의 초점이 타당하고 명확한가?
	2. 수행에 대한 묘사가 적절하여 이를 사용하는 학습자들의 학습을 도와줄 수 있는가?(피평가자의 동의)
객관성의 정도	1. 모든 평가자들은 사전에 회의를 통하여 점수 부여의 객관성을 확보하기 위하여 의견을 나누는가?
	2. 평가 결과에 대한 객관성은 확보되는가?
	3. 평가 결과의 객관성에 대한 피평가자의 생각은 어떠한가?

문제해결력을 증진시키고, 자율적인 학습태도와 습관을 증진시키며, 사회와 산업학교의 협력관계의 향상을 가져오는 과제이다.

2단계는 교육목표와의 연계이다. 루브릭의 평가도구는 발달적 교육관에 근거한 준거지향평가(criterion referenced evaluation)에 기반을 두고 있다고 볼 수 있다. 여기에서 준거는 목표이다. 무엇을 가르치고, 무엇을 배웠는가를 하는 것의 근거가 될 수 있는 목표를 절대적인 기준으로 삼아 학습자들의 성취도를 확인하고자 함이다. 성공적인 교수-활동이 전개되도록 목표로써 제시되는 ‘무엇을 얼마나 성취하였는가?’의 문제를 과제와 연계시킴으로써 자격증 부여의 자격기준을 설정하는 것이다. 합리적이고 타당한 성취기준이나 자격기준을 어떻게 정할 것이냐의 문제는 쉽지 않다. 수행해야할 필수 과제를 관측가능한 행동의 서술로 나타내는 행동목표로 설정하고, 이들 목표들을 목록화 함으로써 구체화하고 체계화시킬 수 있을 것이다.

3단계는 수행수준의 분류단계이다. 전 단계에서는 학습자가 알고 있는 지식이나 기능을 나타내 보일 수 있도록 어떤 산출물로 나타내거나 행동으로 표현하도록 행동목표를 통해 요구된다면, 이 단계에서는 이들 목표들을 기준으로 수준을 나누고 각 수행 수준별로 지표(예)를 작성하는 것이다. 각 수행수준의 등급을 나눠 지표를 제시함으로써 이를 통해 가장 잘한 수준과의 차이를 구체적으로 파악할 수 있을

것이다. 이는 교수-학습활동을 개선하는데 필요한 정보를 제공해 줄 수 있다. 수행의 수준별 등급은 다양한 질적 특성이 묘사되고, 그것을 측정할 수 있는 지표가 마련되도록 작성하기 때문에, 왜 목표에 도달하지 못했는가에 대한 자세한 정보를 수집할 수 있을 것이다. 즉 평가의 내용과 방법 그 자체를 교수-학습활동을 바람직한 방향으로 유도하는 수단으로 이용할 수 있을 것이다.

4단계는 타당성 확보의 단계이다. 작성된 평가도구는 표 2에서 제시하고 있는 내용적 측면, 준거의 명확성, 실용성, 신뢰성, 객관성의 정도 항목을 통해 점검해야 한다. 개발된 평가도구는 그것이 평가하려고 하는 내용을 어느 정도 충실히 측정하고 있는가를 보여주는 타당성(내용적 측면, 준거의 명확성)뿐만 아니라 신뢰성과 객관성, 실용성을 점검함으로써 좋은 평가도구가 갖추어야 할 요건들을 충족하고 있는가를 확인받아야 한다. 첫째, 내용적 측면은 성과의 질을 판단하는데 다루어져야 할 특성이며, 수행과 관련된 모든 특성들을 포괄한다. 둘째, 준거의 명확성은 이해하기 쉬우며, 동일하게 해석할 수 있도록 수행이 적절하고 명확하게 묘사되어야 하며, 각 수행수준이 묘사하는 예를 쉽게 생각해 낼 수 있어야 한다. 셋째, 실용성은 적용하기에 편리한 정도를 의미한다. 루브릭 사용이 학습자의 향상을 돕는 구체적인 교수법을 결정하는데 관련이 있어야 한다. 넷째, 신뢰성은 내용이 타당하며 믿을만하다고 판단할 수 있

는가를 의미한다. 다섯째, 객관성은 모든 대상자들에게 공평하게 실시되어야 한다는 것으로 교사의 채점 방식과 평가도구가 갖는 지시사항 및 문항에 대한 학생들의 해석 방법에 관련이 있다. 학생들이 평가에서 요구하는 것이 무엇인지를 파악하지 못한다거나 문항의 공정성에 관해 의심을 한다면 이 시험은 믿을 수 없는 것으로 인정될 것이다.

V. 기존 루브릭의 분석

루브릭의 개발단계 중 4단계에 이르러 타당도의 확보를 위한 검토지침을 제시하였다. 설문지를 통해 현재 사용되고 있는 국가기술자격 실기시험 채점기준표가 이 검토지침에 얼마만큼 부합하고 있는가를 확인하였다.

표 3에서 살펴보면 전체적으로 점수가 낮게 평가되었다. 평균 3.05~3.55의 분포를 보인다. ‘루브릭의 항목이 관련 교육 기준(교육목표)에 부합되는가?’의 항목은 비교적 긍정적인 인식을 보이고 있다. 그러나 절반가량의 응답자가 ‘그렇지 않다’와 ‘보통이다’의 응답을 한 것으로 보아 이 항목 또한 기대치를 충분히 만족시키지는 못하고 있다.

가장 낮은 평가를 받은 항목은 ‘수행의 수준별 등급은 핵심적 특성들을 잘 묘사하고 있는가?’와 ‘양화될 수 없는 항목들의 각 수준은 정확한 용어로 정의되고, 자세히 묘사되고 있는가?’이다. 제공된 채점기준표를 살펴보면, 자세한 묘사이기보다 ‘미흡한 정도에 따라’, ‘충실하게 나타낸 정도에 따라서’, ‘적합한

정도에 따라’, ‘상태가 올바르게’, ‘색채를 표현한 정도에 따라’, ‘태도나 작업자세가 올바르게 경우’, ‘상태가 완벽하면 5점, 약간 미숙하면 2점, 기타 0점’, ‘양호하면 3점, 보통이면 2점, 기타 0점’ 등으로 표현되어 있어 준거가 모호하게 표현된 관계로 위의 두 항목에서 낮은 점수를 받을 수밖에 없었다고 본다. 양화되어 표현되거나, 양화될 수 없었다면 각 단계에 대한 특성이 자세히 묘사되어있어야 되지만 이 기준이 충분히 반영되지 못한 것으로 보인다.

어느 자격종목은 모든 항목에 평가의 준거를 제시하지 않은 상태에서 항목내용을 기준으로 할 것을 지시하며 ‘만점 5점, 중간점수 3점, 최하점 0점’ 등으로 채점 점수만을 구분하고 있기도 하다. 이는 전문가에 의한 평가가 전제되기에 사용될 수 있는 채점기준표일 것이다. 그러나 이 채점기준표는 자격시험을 준비하는 학습자와 교수자의 입장에서는 많은 부분이 부족하다고 본다. 이 평가도구를 교수-학습의 과정에서 학습자의 수업참여를 촉진하고 교수적 지원 도구로서의 역할이 되게 하기 위해서는 분류한 만점, 중간, 최하의 수준에 대한 명확한 준거제시와 자세한 묘사가 필요하다.

타당성, 신뢰성, 실용성, 객관성의 분류로 본다면 하위점수 1,2,3을 모두 포함하고 있는 타당성의 항목이 상대적으로 부족하다고 볼 수 있다. 평가도구의 조건에서는 기본적으로 전제되어야 되는 항목이기에 앞으로의 개선이 필요하다고 본다. 신뢰성, 실용성, 객관성의 항목들은 비슷한 점수분포를 보이고 있다. ‘객관성의 정도’의 항목 중 평가 결과의 객관성에

표 3. 검토 지침에 따른 국가자격시험 루브릭 평가도구의 평가
Table 3. Evaluation for rubric of national technical qualification test using review guidelines

구분	전혀 그렇지 않다	그렇지 않다	보통이다	그렇다	매우 그렇다	평균 (M)	표준편차 (SD)
루브릭의 항목이 교육목표에의 부합	0(0.0)	3(7.1)	17(40.5)	18(42.8)	4(9.5)	3.55	0.678
수준별 준거의 적절성	0(0.0)	11(26.2)	18(42.8)	13(30.9)	0(0.0)	3.05	0.754
수행의 수준별 핵심적 특성들의 묘사	1(2.4)	8(19.0)	19(45.2)	12(28.6)	2(4.8)	3.14	0.891
양화된 측정지표의 마련	1(2.4)	6(14.3)	17(40.5)	15(35.7)	3(7.1)	3.31	0.886
정확한 용어로 정의와 자세한 묘사	0(0.0)	10(23.8)	20(47.6)	11(26.2)	1(2.4)	3.07	0.769
일관성 있게 사용할 수 있는 도구	0(0.0)	7(16.7)	20(47.6)	12(28.6)	3(7.1)	3.26	0.818
적절한 수의 평가항목	0(0.0)	8(19.0)	16(38.1)	16(38.1)	2(4.8)	3.29	0.825
교수-학습의 활동의 단계 제시	0(0.0)	5(11.9)	16(38.1)	19(45.2)	2(4.8)	3.43	0.760
수행이나 작품개선을 위한 피드백의 역할	1(2.4)	6(14.3)	17(40.5)	16(38.1)	2(4.8)	3.29	0.853
개별평가자들의 동의	0(0.0)	10(23.8)	15(35.7)	15(35.7)	2(4.8)	3.21	0.860
피평가자들의 동의	0(0.0)	7(16.7)	16(38.1)	17(40.5)	2(4.8)	3.36	0.807
사전회의를 통한 객관성 확보	0(0.0)	6(14.3)	13(30.9)	20(47.6)	3(7.1)	3.48	0.823
평가 결과에 대한 평가자의 객관성 확보	0(0.0)	8(19.0)	13(30.9)	18(42.8)	3(7.1)	3.38	0.872
평가 결과의 객관성에 대한 피평가자의 생각	0(0.0)	10(23.8)	14(33.3)	14(33.3)	4(9.5)	3.29	0.933

표 4. 국가자격시험 심사위원 참여여부에 따른 루브릭에 대한 평가의 차이검증
Table 4. T-test of evaluation regarding the rubric which is participation of a judge national technical qualification test

구분		빈도수	평균	표준편차	t	유의확률
루브릭의 항목이 교육목표에의 부합	참여	25	3.84	.746	.572*	.003
	비 참여	16	3.13	.619		
수준별 준거의 적절성	참여	25	3.16	.800	.847	.255
	비 참여	16	2.88	.719		
수행의 수준별 핵심적 특성들의 묘사	참여	25	3.24	.879	.121	.402
	비 참여	16	3.00	.894		
양화된 측정지표의 마련	참여	25	3.32	.945	.139	.928
	비 참여	17	3.29	.849		
정확한 용어로 정의와 자세한 묘사	참여	25	3.08	.862	2.012	.932
	비 참여	17	3.06	.659		
일관성 있게 사용할 수 있는 도구	참여	25	3.36	.907	2.494	.358
	비 참여	17	3.12	.697		
적절한 수의 평가항목	참여	25	3.16	.987	6.927	.158
	비 참여	16	3.50	.516		
교수-학습의 활동의 단계 제시	참여	25	3.40	.816	.443	.693
	비 참여	16	3.50	.730		
수행이나 작품개선을 위한 피드백의 역할	참여	25	3.36	.952	2.735	.544
	비 참여	16	3.19	.750		
개별평가자들의 동의	참여	25	3.20	.913	.439	.862
	비 참여	16	3.25	.856		
피평가자의 동의	참여	25	3.48	.872	1.959	.182
	비 참여	16	3.13	.719		
사전회의를 통한 객관성 확보	참여	25	3.52	.963	3.780	.763
	비 참여	16	3.44	.629		
평가 결과에 대한 평가자의 객관성 확보	참여	25	3.56	.917	.574	.129
	비 참여	16	3.13	.806		
평가결과의 객관성에 대한 피평가자의 생각	참여	23	3.22	1.043	2.561	.499
	비 참여	11	3.45	.688		

대한 피평가자의 생각은 어떠한가?’에 관한 질문을 교사가 보는 관점에서 응답하도록 함으로써 의견의 차이가 커진 것으로 보인다(SD=.933).

표 4는 응답자들이 국가자격시험 실기시험에 심사위원으로 참여한 경험여부(참여25명, 비 참여17명)에 따라 채점기준표 항목에 대하여 어떤 인식의 차이를 보이는가를 알아본 것이다. 이는 심사위원의 경험이 없었던 분들이 루브릭에 대한 접근이 부족함으로 해서 채점기준표에 대한 판단을 다르게 하고 있지는 않은가를 확인하기 위함이다.

첫 번째 항목을 제외한 모든 항목에서 p-value(유의확률)값이 0.05보다 크기 때문에 두 집단 간 인식 정도의 차이가 없다고 볼 수 있다. 다만 ‘루브릭의 항목이 교육목표에 부합되는냐’의 질문에는 t값이 .572이고, p-value는 .03으로 $\alpha = 0.05$ 에서 두 집단 간 인식의 차이가 있는 것으로 나타났다. 심사위원으로 참여했던 응답자가 좀더 ‘루브릭의 항목이 교육목

표에는 더 부합되고 있다’고 보고 있다. 전체적으로도 다른 항목보다는 상대적으로 높은 수치를 보여주고 있는데, 이는 내용전문가로서 본래 측정하고자 의도한 바를 측정하고 있다고 전문적 지식에 근거해 판단하고 있다고 보아진다. 비 참여 응답자는 채점기준표를 가지고 직접 심사에 참여하지는 않았지만, 실기시험 평가의 채점기준표는 가르치는 교사의 입장에서 충분히 검토하고 있어야 하는 상황이기에 나머지 항목에 대해서는 심사위원의 경험유무에 상관없이 두 집단 간 유의미한 차이가 없는 답변을 한 것으로 보인다.

한 가지 덧붙이자면 연구의 자료로 사용한 국가기술자격 실기시험 채점기준표는 채점의 기준을 제시하고 있는 ‘채점기준표’와 점수를 기입하는 ‘채점표’가 따로 분리되어 있다. 평가자가 유의사항과 채점기준에 대한 사항을 완전히 숙지하지 않는다면, 평가를 진행하는데 있어 많은 불편을 야기할 것으로 보인다. 집계를 위해 추후에 한 번 더 정리하더라도 채점기

준과 채점표가 연계된 도구로 제시되었으면 하는 의견이다.

VI. 결론 및 제언

현재 사용되고 있는 국가기술자격 실기시험 채점 기준표는 평가 항목과 각 항목의 배점을 나열하여 각 항목별로 다르게 배점을 주는 방식으로 체크리스트보다는 훨씬 많은 융통성을 부여하고 있지만, 각 항목 배점에 대한 정확한 용어로 정의되고 구체적인 묘사가 제시되어 있는 것이 부족하기 때문에 수행리스트(performance)에 더 가깝다고 할 수 있다.

설문지를 통한 결과에서 가장 부족하다고 지적되었던 항목들이 개선되었을 때 수행리스트가 아닌 진정한 루브릭의 명칭으로 채점기준표의 평가도구를 사용할 수 있을 것이다. 채점기준표(루브릭)는 학습자에게 반응에 대한 구체적인 방법과 수준을 구체적으로 제시해 주는 평가지침이다. 이것을 바탕으로 학습자들이 스스로의 능력에 비추어 어느 수준의 성취가 가능한가를 진단하고, 과제를 수행하는 과정에서 충분히 동기유발이 가능할 것이다.

이 연구는 루브릭을 국가기술자격 실기시험 평가의 도구로써 접근하였다. 그것은 '어떤 목적 하에 활용할 것인가?'를 고려하였다는 것이다.

그래서 이 연구에서 제시하는 루브릭의 특징은 다음과 같다. 첫째, 목표로써 제시되는 '무엇을 얼마나 성취하였는가?'의 문제를 과제와 연계시킴으로써 자격증 부여의 자격기준을 설정하는 것이다. 수행해야 할 필수 과제를 관측가능한 행동의 서술로 나타내는 행동목표로 설정하고, 목록화 함으로써 구체화하고 체계화시키고자 한다.

둘째, 일반적으로 루브릭을 통해 학습자의 활동을 양적 정보보다는 질적 정보로써의 장점을 더 강조하고 있지만, 자격시험의 도구로써의 루브릭이기에 그 수준에 관한 부분을 가능하면 양화시키도록 하고 양화될 수 없다면 자세하게 묘사하도록 권장한다.

셋째, 루브릭은 준거지향평가도구이기에 가장 우수한 본보기로부터 유도되어 가장 높은 점수가 부여된다. 그래서 가장 높게 설정된 점수를 아무도 받지 못할 수도 있고, 대부분의 대상자가 하위점수를 받을 수도 있다. 하지만 자격시험의 루브릭은 일정 수준의 합격자를 배출할 수 있도록 난이도 조절을 적절하게 하는 것이 전제되어야 한다.

좋은 평가도구는 기본적으로 타당성과 신뢰성이 모두 높아야 하지만, 그 사이에서 균형을 맞추는 것

은 어려운 일이다. 루브릭이라는 평가도구를 보다 구체화 시키면 타당성은 높아지고, 단순하고 간단하면 신뢰성 있는 도구가 되기 때문이다. 그러기에 루브릭을 개발하고 이를 적용하는 것은 쉽지 않지만, 루브릭 외에 타당성과 신뢰성에서 뛰어난 평가도구를 찾는 것도 쉽지 않을 것이다.

앞으로 많은 장점을 가진 루브릭은 본 논문에서 전제한 사항 안에서 최선의 개발이 되도록 하나의 틀을 제시하였듯이, 상황에 따라 많은 사항들이 고려되어 다양하게 활용될 수 있을 것이다.

참고 문헌

- [1] 김소연. (2009). *순환적 예술활동을 이용한 유아미술교육프로그램이 창의성, 유추, 그림감상 능력에 미치는 효과 및 루브릭 개발*. 박사학위논문. 성균관대학교.
- [2] 김정덕. (2009). *중학교 쓰기 수업과 평가를 위한 루브릭 사례 연구*. 박사학위논문. 한양대학교.
- [3] 노원경. (2008). *루브릭을 활용한 학습전략 교육 프로그램 개발 및 효과 연구*. 박사학위논문, 연세대학교.
- [4] 윤관식. (2011). *수업설계*. 한국기술교육대학교.
- [5] 한지영(2004). *기술과 교육평가에서 학습자 중심 루브릭이 학습과정 및 학업성취에 미치는 영향*. 서울대학교 박사학위 논문.
- [6] Arter, J. (2000). Rubrics, Scoring guides, and performance criteria: Classroom Tools for assessing and improving student learning. paper presented at the Annual Meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED446100)
- [7] Arter, J. & McTighe, J. (2001). Scoring rubrics in the classroom; Using performance criteria for assessing and improving student performance. Thousand P란, CA: Corwin Press.
- [8] Bandura. A. (1986). Social foundation of thought and action: A Social cognitive theory. Englewood Cliffs, NJ: Prentice Hall.
- [9] Brown, A. L. (1988). Metacognition: The

development of selective attention strategies for learning from text. In Kamil (Ed.), Directions in reading research and instruction. Hillsdale, NJ: Erlbaum.

- [10] Goodrich, H. (1996). Understanding Rubric. Educational Leadership. 54(4), 14-17.
- [11] Goodrich, H. (1999). The role of instructional rubrics and self-assessment in learning to write: A Smorgasbord of Findings. (ERIC Document Reproduction Service No. ED 431029)
- [12] Herman, J. L., Aschbacher, P. R. & Winters, L. (1992). "A Practical Guide to Alternative Assessment", ASCD, Alexandria, VA
- [13] Stiggins, R. J. (1994). Student-centered classroom assessment. New York: Macmillan.
- [14] Wiggins, G. (1996). What is a rubric? A dialogue on design and use. A Handbook for Student Performance Assessment in an Era of Restructuring. Assn for Supervision & Curriculum; Lslf edition.

윤 관 식 (Gwan-Sik Yoon)

정회원

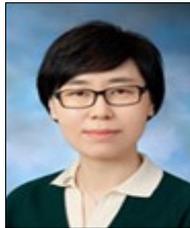


1983년2월: 동아대학교 교육학과교육학사)
 1985년 2월: 경북대학교 교육학과(교육학석사)
 1993년 1월: 플로리다주립대학교(교육공학 박사)
 1994~현재: 한국기술교육대학교 인력개발전문대학원 교수

<관심분야> 교육프로설계 및 개발, 교수법, 이러닝

최 명 란 (Myung-Ran Choi)

정회원



1987년 2월: 중앙대학교 교육학과(교육학사)
 2004년 2월: 한국기술교육대학교 인력개발학과(교육학석사)
 2011년 2월: 한국기술대학교 인력개발학과 (인력개발학박사)
 2004년 2월~현재 한국기술대학교 교양학부 시간강사

<관심분야> competency, evaluation