

수행평가의 이해와 적용

임병빈 · 연준흠
(공주대 · 청주외국어고)

Im, Byung-Bin, & Yeon, Jun-Hum. (1999). Understanding and adaptation of performance assessment. *English Language and Literature Teaching*, 5, 149~189.

Assessment can be defined as any method used to understand the current knowledge that a student possesses. Assessment may influence decisions about grades, placement, instructional needs, and curriculum. Therefore, the purpose of assessment is to identify how the students think and know, to diagnose the difficulties they face, and to reflect the result of assessment on teaching. But the traditional multiple-choice test failed to evaluate and teach higher-level thinking and problem-solving skills. In this context, performance assessment is being required as an alternative assessment to get better understanding about what the students can do as well as what they know. This article points out some weaknesses of traditional assessments, and comments on the theoretical background and necessity of performance assessment. And it presents more specific information about performance assessment and some examples. It is certain that performance assessment is student-centered and future-oriented. But performance assessment can't be the surest and best way of evaluating the students' abilities. We are just on the way of another experimental stage for improving teaching methodology. More supplementary analyses and further improvements are needed.

1. 서론

최근에 새롭게 부각되고 있는 교육현장의 관심사로서 각종 세미나·논문 발표를 포함하여 언론 매체에 등장하고 있는 '수행평가'라는 혁신적인 평가 방법은 아직도 많은 교육종사자에게 다소 낯설은 개념으로 남아 있는 것 같다. 뿐만 아니라 수행평가에 관하여 그 개념을 어느 정도 이해하고 있는 교사라 할지라도 교수학습 과정에서 구체적으로 어떻게 적용해야 하는지 매우 당황해 하는 일종의 새로운 딜레마가 막중한 부담으로 작용하고 있는 현실이다.

한국의 영어 교육 발자취를 돌아보면 이제 한 세기를 넘어 새로운 도약을 위한 나침의 발판을 구축하고 새로운 시대의 국제적 흐름에 발맞추어 보다 효과적으로 영어를 배울 수 있도록 각종의 처방과 치료를 진행하여 온 것으로 볼 수 있다. 몇 가지 예를 들면 '90년대에 즈음하여 '대입 수학 능력 시험'의 영어시험출제에서 전반적으로 의미중심 평가(meaning-oriented test) 방식으로의 전환과 아울러 듣기평가 실시, 초등학교 교육과정의 조기 영어교육 실시, 대규모의 원어민 보조교사 활용, 현직 영어교사 해외 연수, TOEIC/TOEFL 성적의 각종 혜택 반영, 대학 교양영어 교육의 회화 위주 수업, 멀티미디어를 활용한 영어 수업 등을 손꼽을 수 있다(임병빈, 1997). 영어교육에서 수행평가를 적용해 보고자 함도 이와 같은 맥락에서 요구되는 또 하나의 과제로 간주될 필요성이 있는 것 같다. 그러나 수행평가의 교육현장 적용은 여러 가지 문제점을 안고 있다. 첫째, 영어교사 대다수가 수행평가를 제대로 이해하지 못하여 적용에 어려움이 있다. 둘째, 수행평가의 효과적인 실시를 위한 환경이 조성되어 있지 않다(과다한 수업 부담 및 업무 처리, 교내신제도에 의한 성적처리 규정과의 상충성 등). 셋째, 수행평가에 대한 충분한 연수가 필요하다. 넷째, 교사와 학생의 필요에 의한 것이 아닌 타율에 의한, 즉 교육부 지침에 따른 수행평가 실시에 대한 거부감이 팽배해 있다.

따라서 본 연구에서는 현재 우리의 영어교육 현장에서 당면하고 있는 수행평가의 심각한 문제점 해결에 도움을 줄 수 있도록 수행평가의 발전 배경을 알아보고 수행평가의 특징, 유형, 내용 및 수행평가의 문항 예시 등에 관하여 논의해 봄으로써 영어교사가 수업과정에서 수행평가 방식을 효과적으

로 활용하여 종래의 전통적 평가 방식에 의존했던 교육의 효과보다 더 나은 학습 효과를 거둘 수 있도록, 수행평가에 관한 근본적 이해와 적용과정의 주요 사항들을 제시해 보고자 한다.

II. 수행평가의 발전 배경

1. 전통적 평가

1) 전통적 평가에 대한 비판

많은 학생들의 성취도를 평가하기 위하여 객관적인 측정에 의존하는 평가 제도는 아주 적절한 것처럼 보인다. 그 동안 성취 검사는 학생이 알고 있는 것과 할 수 있는 것에 대한 타당하고 신뢰성 있는 평가 도구로 인정받아 왔다. 이 평가는 교과 내용을 학생들에게 가르치는 것이 필수적이며 논리적인 것으로 보였던 시기에 유래하였다. 더욱이 많은 학습 이론가들은 개념이 작은 구성 요소로 분할될 때 효과적인 학습이 이루어진다고 믿었다. 전통적 평가는 이러한 가정과 실천을 반영하였다. 왜냐하면 그러한 평가의 대부분은 전형적으로 특정 교과와 내용 영역을 평가하기 위하여 일련의 선택형 항목을 사용하였기 때문이다. 이러한 평가를 지지하는 사람들은 단편적 지식(예, 특정 지식이나 사실)을 많이 알고 있는 학생이 내용 영역에 대해 더 많이 이해하고 있다고 가정하였다.

성취 검사의 결과는 다양한 목적으로 사용되었고 또 앞으로도 그럴 것이다. 평가 점수는 학생들을 비교하고, 집단이나 특정 프로그램에 배치하고, 학생을 지도하거나 상담하는데 사용되고 있다. 또한 그 결과는 교사나 행정가를 평가하고 심지어는 학교의 교육 과정이나 교육 프로그램의 질을 평가하는데도 사용된다. 이러한 편이성 때문에 매년 수많은 학생들에게 전통적인 평가 방법을 사용하여 평가되고 있다.

전통적인 표준화 성취 검사에 대한 비판은 대단히 많다. 다음 몇 가지 대표적인 예를 살펴보자.

Lipman(1987)은 선택형 문제를 사용한 성취 검사는 복잡한 문제의 해결 기능, 다양한 사고 그리고 학생들 사이의 협동적인 노력을 평가하는데 효과

적이지 못하다고 주장한다. 성취 검사는 또한 소통 기능을 평가하는데도 비 효과적이라고 언급하고 있다. 이와 마찬가지로 Resnick과 Resnick(1989)은 성취 검사는 간단하고, 단편적이며, 피상적인 읽기, 즉 단편적인 글에서 정보를 찾고, 수동적으로 오류를 찾아내고, 선택된 응답으로 다른 사람의 질문에 답하는 것과 같은 특징을 갖고 있다고 주장한다.

Archibald와 Newmann(1988)은 성취 검사에서 높은 점수를 받은 학생이 다른 학생들보다 잘 알고 있으며, 고도의 사고 기능을 갖고 있거나 더 많은 교육을 받았다는 가정에 도전한다. 그들은 성취 검사의 결과가 학교 성적과 어느 정도는 상관 관계가 있을지라도 학문적인 질문, 지식의 통합 또는 새로운 문제 처리 능력을 요하는 수준에서의 수행과는 상관 관계가 없다고 지적하고 있다.

Barth와 Mitchell(1992)은 선택형 위주 평가가 본질적으로 답을 구성하는 것이 아니라 선택하도록 하여 학생을 수동적으로 만들기 때문에 교육을 망치고 있다고 주장한다(p. 14). 더욱이 그들은 규범 위주의 성취 검사가 답이 항상 옳거나 틀린 것이라는 인상을 준다는 사실을 비난한다. 성취 검사의 사용은 이해보다는 암기를 권장하여 궁극적으로는 쓸데없는 것을 위한 공부를 조장한다고 주장한다(p. 15).

Mislevy(1989)는 다음과 같이 지적하고 있다. 교육 평가는 오늘날 그 기초를 위협하는 것으로 보이는 위기에 직면해있다. 본질적인 문제는 표준화 평가 이론에서 의미하는 인간 능력에 대한 견해가 인지 심리학과 교육 심리학에서 주장하는 견해와 양립될 수 없다는 점이다. 학습자는 단순히 새로운 사실과 기술의 축적을 통해서가 아니라 그들의 지식 구조를 재형성하고, 절차를 자동화하고 정보를 처리할 수 있는 능력을 기르고, 사실과 기능이 언제 어떻게 적절한지를 말할 수 있는 모형을 개발함으로써 자신의 능력을 향상시킬 수 있다는 것이다.

Wiggins(1989)는 대규모 평가가 교육이나 학생들의 사고 처리 유형을 유사한 것으로 취급하고 있다고 주장한다. 그러나, 이해를 측정하기 위해서는 학습자의 답을 조사해야 하며, 학생이 완전히 검토되었다는 것을 확인하기 위해서는 평가자와 피평가자 사이에 어느 정도 대화의 가능성이 있어야 한다.

이러한 비판들이 주로 표준화 성취 검사에 초점을 두고 있지만 대부분 교사가 제작한 평가나 출판물에 의해 공급되는 평가에도 적용된다.

2) 내용 중심 교육 과정과 교육에 대한 비판

학교에서 표준화 성취 검사의 지나친 사용에 대한 비판과 함께 너무나 많은 교육 과정과 교육이 내용을 가르칠 목적으로 조직되어 있다고 주장하는 사람들도 있다. 평가와 교육 비판가들의 초점이 다를지라도 그들의 결론은 같다. 모두 고도의 가치가 있는 기능과 지식을 가르치고 평가하는데 실패했다고 주장한다.

오늘날 교육 실천에 대한 비판가들은 대부분 지도가 교사 중심이며, 학생들은 수동적인 학습자로 예상된다고 말한다. 예를 들어, Glickman(1991)은 지난 반세기 동안 교실에서는 거의 바뀐 것이 없다고 주장한다. “수업 시간의 대부분을 교사는 강의하고, 학생은 듣고, 교재를 읽거나 답안지를 채우는 데 시간을 보내고 있다. 지금의 교실을 관찰하는 것은 마치 50년 전의 학생들을 보는 것과 같다.”(p. 5).

교과 교육을 중심으로 조직된 교육 과정과 지도에 특히 비판적인 사람은 Wiggins(1989)이다. 그는 중요한 것을 모두 가르치려는 시도가 교육을 사소하고 잊혀지는 것들의 목록으로 전략시켰다고 말하며 그러한 교육을 한 사람들을 비판한다. 대안으로 학생들에게 적은 양을 잘 알고 행하도록 가르쳐야 한다고 주장한다. 그는 “사고 습관과 고도의 기능”을 길러야 한다고 말한다. Wiggins는 학생들이 생각이 없고, 피상적이며, 값싼 공부에 대해 싫증을 느끼도록 해야 한다고 말한다. 따라서 그는 이러한 것들이 목표가 될 때 교육 과정 설계는 결국 전형적인 틀에서 해방될 것이라고 주장한다(p. 45). Wiggins는 대안으로 교육 과정이 다음과 같은 네 가지 목적을 수행하도록 조직되어야 한다고 주장한다. (1) 학생들이 신중한 질문을 통해서 그들의 피상적인 지식을 심화시킬 수 있는 능력을 갖추게 하며, (2) 그러한 질문을 확실하고 체계적인 지식으로 바꿀 수 있게 하며, (3) 학생들이 고도의 기능을 계발하도록 하고, (4) 중요한 질문에 완전히 몰두하게 하여 중요한 지식을 추구하는데 기쁨을 얻는 방법을 터득하도록 해주어야 한다.

2. 행동주의와 인지주의 관점의 평가

1) 행동주의 관점에 대한 비판

“비행기를 탈 때 선택형의 지필 고사에서 고득점을 얻어 자격을 취득한 비행사의 비행기를 타겠는가? 아니면 수천 마일의 실제 비행 기록을 갖고 있는 비행사의 비행기를 타겠는가?” 이 말은 전통식 평가의 문제점을 제기하는 동시에 새로운 평가에 대한 필요성을 잘 반영하고 있다(Tannenbaum, 1996).

대부분 전통적인 평가의 이론적 가정은 행동주의적 관점에 근거를 두고 있다. 그러나 특히 1990년대에 학습과 학습 평가에 대한 새로운 대안적인 방법의 필요성을 깨닫게 되었다. Gardner(1993)는 지능의 복잡성에 대한 관심이 되살아나고 있다고 주장한다. 다른 연구자들(Perkins, 1981; Gruber, 1985)은 창조성의 개념을 연구하였다. 그들의 연구에 의하면 창의적인 개인은 하나의 정신적 조합을 갖고 있는 것이 아니라 좀더 효율적이고 융통성 있는 방법으로 그들이 갖고 있는 것을 사용한다는 사실을 밝혔다. 그러한 개인은 자신의 행동, 시간 사용 그리고 산출물의 특질에 대해 지극히 반성적(reflective)이라는 점이다(Gardner, 1993). 따라서 외국어와 제 2 언어 학습에서 새로운 평가는 상황이 수행에 미치는 영향을 인정해야 하고 자기 평가에 개인을 관련시키는 것을 포함한 능력을 평가하도록 가장 적절한 상황을 제공해야 한다는 것이다. Brecht와 Walton(1993, p. 2)은 언어 능력을 “국내 또는 외국에서 그 사회의 언어를 사용하여 다른 문화나 언어 사회의 구성원들과 어느 정도 직업적으로 또는 전문적으로 적절한 소통 과업을 수행하는 능력”으로 정의한다. 그들은 또한 소통의 정의를 내리도록 고안된 언어 학습들을 요구하며 이러한 능력이 최적으로 습득될 수 있는 방법, 교수 자료의 개발, 교수 방법의 적용, 교사 교육과 평가, 언어 습득과 관련된 연구 등을 요구하고 있다(pp. 8-9).

많은 교육자들은 대안 평가가 학습자의 학문적 발달에 대한 역동적인 측면을 파악하는 중요한 방법이라는 것을 인정하게 되었다. 대안 평가란 지도 상황에서 사용될 수 있고 학교나 교실에서의 일상적인 활동으로 쉽게 통합될 수 있는 절차와 기법을 말한다(Hamayan, 1995, p. 213). 대안 평가는 학

생들에게 자신이 할 수 있는 것을 보여주도록 요구하는 전략을 사용하기 때문에 제 2 언어로써 영어를 배우는 학습자에게 특히 유용하다. 전통적인 평가와는 달리 학생들은 그들이 회상하고 재생할 수 있는 것보다는 그들이 통합하고 생산해내는 것에 대하여 평가된다. 대안 평가의 주요 목적은 학습자가 어떤 특정 영역에서 실생활의 과업을 어떻게 접근하고, 처리하고 완성하는지에 대한 증거를 수집하는 것이다(Huerta-Marcias, 1995, p. 9). 대안 평가는 일반적으로 다음과 같은 기준을 충족시켜야 한다.

- 학습자를 서로 비교한다기보다는 일정 기간에 걸쳐 개별 학생의 성장을 기록하는데 초점을 둔다.
- 학습자의 단점(그들이 모르는 것)보다는 장점(그들이 알고 있는 것)을 강조한다.
- 학습 유형, 언어 숙달도, 문화적 교육적 배경 및 학습자의 등급 수준을 고려한다.

평가 방법은 학습에 대한 이론에 의해 결정된다. 초기 학습 이론에 따르면 복잡한 고도의 기능은 학습을 일련의 필수적인 기능으로 분류하여 조금씩 습득되어야만 했다. 이는 마치 지식이라는 벽들을 하나씩 쌓아올리는 접근과 같다. 기본 기능이 기계적으로 학습된 다음에 복잡한 이해와 통찰력으로 조합될 수 있다고 가정하고 있다. 그러나 오늘날 인지 심리학에서는 모든 학습이 학습자가 생각하고 적극적으로 정신적 모형을 발전시켜가면서 구성하는 것을 요구한다고 밝히고 있다.

2) 인지주의 관점의 평가

인지적 관점에서 보면 유의미한 학습은 반성적이고, 구성적이며, 자기 통제적이다. 인간은 사실적 정보의 단순한 기록자가 아니라 지식 구조의 창조자로 여겨진다. 어떤 것을 안다는 것은 정보를 단순히 수용하는 것이 아니라 이를 해석하고 이미 갖고 있는 다른 지식과 연관짓는 것이다. 또한 단지 수행하는 방법에 대해서 아는 것뿐만 아니라 언제 수행하고, 새로운 상황에 어

떻게 적용하는 방법을 아는 것도 중요하다. 그러므로 선택형 평가의 전형적인 초점이었던 별개의 정보량은 유의미한 학습 평가에서 그렇게 중요하지 않다. 오히려 중요한 것은 학습자가 복잡한 문제를 해결하기 위하여 상황 속에서 정보를 조직하고 구조화하고 사용하는 능력이다.

종래의 학습 관점과는 대조적으로 인지심리학에서는 학습이 선형적인 것이 아니라 동시에 그리고 각기 다른 속도로 여러 방향으로 진행해 나간다고 제안한다. 모든 연령층과 각기 다른 능력 수준을 지닌 사람들은 끊임없이 개념을 사용한다. 따라서 개념 학습이 특정 연령 이전까지 또는 모든 기본적인 사실이 숙달되기 전까지 미루어져서는 안된다. 사람들은 지식을 습득하는 유형이나 속도, 지식 습득과 수행에 적용할 수 있는 주의력과 기억 능력, 그리고 그들이 창안해낸 개인적인 의미를 밝힐 수 있는 면에서 매우 다양하다.

학습의 본질에 대한 최근 연구에서는 구조 연습과 사실적 지식에 대한 연습을 지나치게 강조하는 교육이 학생들에게 도움이 되지 않음을 분명히 보여주고 있다. 또한 학습과 동기의 통합에 대해서도 정의적이고 초인지적인 기능의 중요성을 강조하고 있다. 예를 들어, 형편없는 사고자와 문제 해결자는 그들이 갖고 있는 특정 기능뿐만 아니라 어떤 과업에서 그러한 기능을 사용하지 못한다는 점에서 우수한 사람들과 구분이 된다고 제안하고 있다. 지식 기능의 습득은 사람을 유능한 사고자나 문제 해결자로 만드는데 충분하지 못하다. 사람들은 기능과 전략에 대한 지식뿐만 아니라 이를 적용하는 능력을 습득할 필요가 있다. 이러한 것들이 평가의 적절한 목표가 되어야 한다.

고도의 인지 능력을 기르는 학습의 사회적 역할이 지난 몇 년 동안 관심의 대상이 되어 왔고 실제 사회의 문제 해결에서는 집단으로 함께 일하는 것을 필요로 하였다. 그러나 대부분의 전통적인 교육이나 평가는 소집단 활동보다는 개별 활동을 강조하였다. 오늘날에는 집단이 여러 면에서 학습을 촉진하는 것으로 상정되고 있다. 효과적인 사고 전략을 만들어내고, 복잡한 과업을 수행해내고, 상호 건설적인 피드백을 제공하고, 비판적인 사고 요소를 평가하는 것이 필요하므로 집단 평가가 중시된다. 다음 표는 행동주의와 인지주의 관점의 차이점을 잘 나타내고 있다.

표 1
 행동주의와 인지주의 관점의 비교(Pietel, et al. 1991)

평가의 중점	행동주의 견해	인지주의 견해
학습자에 대한 견해	수동적으로 반응	적극적으로 반응 지식 환경을 스스로 구성
평가 영역	별개의 독립된 기능	학제간의 통합적 기능
지식에 대한 관점	별개 지식의 축적	사실과 기능 지식의 적용
교육의 중점	최대로 전달	초인지에 관심, 효과적인 자료 제공, 동기 평가 및 자기 진단
평가의 특징	지필고사, 객관식 평가 단답형 속도 측정	참평가, 수행 평가 개방형(한 개 이상의 정답) 고도의 사고 능력 강조 속도를 요구하지 않음
평가의 빈도	한 번의 기회	수시 평가, 장기간의 샘플
평가 대상	개인 평가 결과 중심	개인 및 집단 평가 과정 중심
기법의 사용	기계에 의한 채점	전문가에 의한 채점 채점 기준에 의한 채점
평가 내용	하나의 속성	다차원적인 평가 학습자 능력의 유연성 강조

3. 대안 평가

1) 대안 평가의 필요성

개인의 기능이나 성취에 대한 척도로써 선택형 평가의 가치와 사용에 대해 의문을 제기하면서 능력을 측정하기 위한 보다 공정하고 실질적인 방법에 대한 필요성을 느끼고 직접적이며 참된 평가 형태의 사용을 주장해왔다. 따라서 교육자들 사이에 대안 평가 즉, 지필 고사나 규범적 평가 이외의 다른 방법을 통한 평가에 관심이 모아지게 되었다. 다시 말하면, 실생활에서 접할 가능성이 높은 상황에 자신의 기능을 적용할 수 있는 개인의 능력과 업무와 관련된 특정한 능력에 초점을 두는 평가에 관심이 높아졌다.

창의적인 개인은 한 가지 방법이 아닌 보다 효율적이고 융통성 있는 방법으로 자신이 갖고 있는 지식을 활용한다. 그러한 개인은 자신의 행동, 시간의 사용, 산출물의 질적인 측면 등을 반영한다.

널리 연구되고 있는 대안 중의 하나가 수행에 근거를 둔 평가이다. 수행 평가는 직접적인 관찰과 기능의 적용에 대한 검토이다. 포트폴리오, 발표, 집단 프로젝트 등 다양한 활동을 통한 수행 평가는 단지 선택형이나 지필 고사를 통해 간접적으로 측정한다기보다는 실제 상황에서 학습자의 지식과 기능을 어떻게 사용하는지를 직접 관찰한다.

수행 평가에 대한 생각은 새로운 것이 아니다. 예술, 체육 심지어는 고용이나 훈련 분야에서 오랫동안 사용되어 왔다. 새로운 것은 교육자들과 수행 평가를 사용하는 사람들 사이에 고도의 사고 기능과 문제 해결 능력 그리고 기능 중심의 직업과 관련된 기본적인 기술을 포함한 광범위한 기본 기능을 보다 효과적으로 측정하려는 관심이 커진 것이다.

2) 대안 평가의 종류

평가 제도를 개선해야 한다고 제안하는 사람들은 서로 다르지만 유사한 의미를 지닌 용어를 사용한다. 이러한 용어에는 참평가, 참교육, 수행 평가 등이 포함되어 있다. 먼저 참평가와 참교육에 대하여 살펴보고 수행 평가에 대해서는 항목을 달리하여 좀더 구체적으로 다루기로 한다.

(1) 참평가

참평가(authentic assessment)는 학교에서 이루어지는 과업이 실제 세계에서의 과업과 유사한 경우를 지칭한다(Wiggins, 1989). 참평가의 목적은 실제 상황과 아주 유사한 상황 속에서 다양한 유형의 능력을 평가하는 것이다. 예를 들어, 참평가는 학생들에게 실제 교재를 읽게 하거나, 유의미한 주제에 대하여 실질적인 목적을 갖고 글을 쓰거나, 책에 대하여 토론하고, 일지를 쓰고, 편지를 쓰거나 글을 완성할 때까지 교정하는 것 등과 같은 실제의 과업에 참여하도록 한다. 따라서 재료나 평가 과업 모두 가능한 자연스럽게 보여야 한다. 더욱이 참평가는 완성된 산출물뿐만 아니라 과업 이면의 생각이나 과정에도 가치를 둔다(Wiggins, 1989; Wolf, 1987).

실제 과업을 수행하는 것은 그 자체가 유용하며 적극적인 활동이다. 교사의 관점에서 그러한 과업을 지도하는 것은 가치 있는 기능과 전략에 중점을 두고 있다는 것을 보여준다. 학생들은 중요한 지식과 기능을 실제의 목적을 갖고 어떻게 적용할 것인가에 대하여 배우고 연습하게 된다. 따라서 단순히 정보를 회상하는 것이 아니라 자신이 알고 있는 것을 새로운 과업에 적용시켜야 한다. 예를 들어, 이야기에 나타난 모든 은유를 찾아내라고 하는 것과 작가가 특정한 은유를 사용하는 이유나 그러한 것들이 이야기에 미치는 영향에 대하여 토론하라고 하는 것 사이의 차이점을 고려해 보라. 후자의 경우에 학생들은 교실 밖에서 자연스럽게 이야기하는 것과 마찬가지로 그들의 지식과 기능을 학습에 적용하게 된다.

참평가에 대한 좀 더 자세한 정의는 Wiggins(1989)에 의해 이루어졌다. 그는 세 가지 요인 즉, 과업, 상황 그리고 평가 기준이 평가의 참을 결정한다고 제안한다. 참과업은 학생에게 어떤 수행을 마무리하거나 산출물을 만들어내기 위하여 지식이나 기술을 사용하도록 하는 것이다. 이러한 정의를 토대로 볼 때 공식을 암기하는 것은 참과업이 될 수 없지만 공식을 사용하여 실질적인 문제를 푸는 것은 참과업이 될 것이다.

그는 상황이 가능한 실제적이어야 한다고 제안한다. 전문가, 시민 그리고 소비자 등이 직면하게 되는 상황을 모방하거나 복제해야 한다는 것이다. 학생들에게 사전 지식이 거의 없거나 활동을 마치는 데 시간이 부족하거나 적절한 자원을 반영하지 못하거나 상담할 기회가 없는 시험은 참된 것이 되지

못할 것이다.

마지막으로 참평가는 수행하거나 생산해내는 학습 자료를 판단하는데 사용되는 것과 같은 기준을 사용해서 이루어져야 한다. 한 예로 작문을 평가하는데 사용되는 참기준은 글의 구조나 내용을 우선적으로 고려해야 하고 부차적인 오류(철자, 구두점, 문법)가 일차적인 초점이 되어서는 안된다.

요약하자면, 참평가는 정보를 단순히 회상하거나 인지하는 것이 아니라 알고 있는 것을 적용하도록 한다.

(2) 참교육과 학습

'참교육'이 뜻하는 바는 용어상으로 수행 평가나 참평가와 유사하다. 이 용어가 교육과 학습을 지칭하고 있지만 평가나 지도에서 변화를 요구하는 사람들은 평가와 교육이 통합되어야 한다고 주장하기 때문에 평가의 범위 내에서 논의하는 것이 적절하다. Newmann과 Wehlage(1993)는 유의미한 학업 성취를 가져온 교육을 기술하기 위해 사소하고 쓸모 없는 교육과 대비하여 '참교육'이라는 개념을 사용하였다.

특히 Newmann과 Wehlage는 학생들에게 다음과 같은 세 가지 폭넓은 목표를 성취하도록 도와주는 것이 참교육이라고 주장한다.

- ① 의미를 구성하고 지식을 만들어낸다.
- ② 의미를 구성하기 위하여 학문적인 질문을 사용한다.
- ③ 실제 사회에서 유용한 담화 생산, 산출물, 수행을 향한 목표를 갖고 있다.

그들은 또한 참교육이라는 개념의 이해를 돕기 위해 다섯 가지 기준을 제정하고 있다. 이러한 기준은 어떤 수업이 참이 될 수 있는 정도를 평가하는데 사용될 수 있다.

- ① 학습자가 고도의 사고 기능을 사용한다(학습자는 문제를 해결하고, 종합하고, 설명하기 위하여 지식과 기능을 적용한다).
- ② 개념, 주제 또는 기능에 대한 이해의 깊이를 강조한다.
- ③ 문제나 주제가 실제 사회나 세계와 연계성을 갖는다.

- ④ 교사와 학생 사이에 상호적이며 실질적인 대화가 이루어진다.
- ⑤ 학습 성취를 위한 사회적 지원이 따른다.

Ⅲ. 수행 평가의 정의, 특징 및 유형

1. 수행 평가의 정의

수행 평가(performance assessment)는 흔히 참평가 대신에 또는 참평가와 함께 사용되는 용어이다. 수행 평가는 응답이나 산출물을 창조함으로써 학생들이 자신의 지식과 기능 그리고 전략을 보이도록 요구한다(Wiggins, 1989). 학생들은 다지 선택형에서 고른다고보다는 연구 수행, 보고서 작성, 등장 인물 분석, 등장 인물의 동기에 대한 논의, 학습한 중요한 정보의 창안, 좋아하는 이야기의 극화, 개인적으로 유의미한 이야기 일부의 낭독 등을 통해 자신의 학업 능력을 보일 것이다.

수행 평가 주창자들은 다음과 같은 유형의 평가를 요구한다. 실험 설계 및 실시, 학생들에게 다시 생각하도록 요구하는 논술 쓰기, 정보 통합 및 적용, 과업을 수행하기 위하여 다른 학생과 협력하기, 장치나 기법을 사용하여 숙달도 보이기, 모형 만들기, 지도를 만들고 해석하고 이를 사용하기, 수집하기, 보고서 비판, 시 또는 간략한 이야기 쓰기, 연설하기, 악기 연주, 구술 시험 참가, 포트폴리오 만들기, 체육 기능이나 일정 개발 등이 있다.

수행 평가는 규범 평가에 대한 대안이다. 교사는 학습자가 알고 있는 것과 할 수 있는 것에 대한 보다 풍부하고 완전한 이해를 얻기 위해 수행 평가를 사용할 수 있다. 수행 평가는 학습자에게 그들의 지식과 기능을 보여주는 답안이나 산출물을 만들도록 요구하는 평가 방법이다.

2. 수행 평가의 특징

수행 평가의 주요 특징은 다음과 같다. 1) 응답을 선택한다기보다는 구성한다. 2) 실제 세계를 반영하는 과업에 관한 학습자의 행동을 관찰한다. 3) 학습자의 학습과 사고 유형을 채점에서 밝힌다. 즉, 학습자가 어떻게 학습을 하였는지 또 어떤 전략을 사용했는지를 보여준다. 4) 학습자가 일정 기간에

걸쳐 학습한 것을 보여준다.

수행 평가의 특징과 관련된 사항을 살펴보면 다음과 같다.

1) 중요한 학습 결과를 측정한다. 수행 평가가 중요한 학습 결과를 측정하는 정도는 특정한 평가 문제나 과업에 달려있다. 수행 평가 과업은 직장이나 일상 생활과 관련이 있는 중요한 실제 세계의 수행을 반영한다. 수행 과업은 학교 상황에서 실제적으로 다루어질 수 있는 특정한 교수 방법과 유의미하게 연결되어야 한다.

2) 교육적 활동과 연결될 수 있는 수행에 대해 분명히 기술한다. 수행 평가 과업이 적절한 연령, 실생활에서의 다양한 개념에 접근할 때 교사는 학습자의 수행을 파악할 수 있으며 학습자의 문제 해결 전략을 밝힐 수 있다. 그러나 학생의 능력을 밝히고, 학습자의 전략 행동에 대한 정보를 찾아내고, 이러한 관찰을 특정한 교수 기법과 연관짓는 것은 교사의 기술에 달려 있다.

3) 다양한 교수 모형과 조화를 이룬다. 이론적으로 수행 평가는 다양한 교수 방법으로 실시될 수 있다. 수행 평가를 실시할 때 특히 문제나 어려움을 갖고 있는 학생들에게 다양한 교수 방법을 실험해야 한다.

4) 교사에 의해 시행되고, 채점되고, 해석된다. 평가를 설계하고 시행하고 학습자의 수행을 분석하는데 많은 시간이 걸릴 것이다. 학생수가 많은 교실 상황에서 실시하기에는 많은 어려움이 따를 것이다. 수행 평가 개발자는 교실 상황의 제약 내에서 수행 평가를 토대로 한 계획을 실행하는 방법을 모색할 필요가 있다.

5) 학습 목표를 학생에게 전달한다. 평가가 교수 목표와 보조를 맞출 때 교사는 교육의 방향을 결정하는 평가를 사용할 수 있으며, 학생들은 개인적인 학습 목표를 달성하기 위해 평가를 사용할 수 있다. 그러나 이것은 채점 기준이 분명하고, 구체적이며, 시각적으로 나타나는 정도에 달려 있다.

6) 정확하고 유의미한 정보(즉, 신뢰롭고 타당한)를 제공한다. 수행 평가는 교실 중심 평가의 미래 지향적인 방향을 제시하고 있지만 그런 약속을 성취할 때까지 여러 가지 부수적인 분석과 개발을 요구한다.

학습자가 어떻게 생각하는지를 알아내고, 학습에서 어려움을 겪는 부분을 진단하기 위한 최선의 방법이 평가이다. 학생들에게 자신을 충분히 표현할

수 있는 가능한 많은 영역을 주고 자연스러운 상황에서 학습을 평가하는 것이 바람직하다.

전통적인 성취 평가에서는 어떤 문장에 대한 문법적 해결책을 묻는 선택형 문제가 학습자의 쓰기 기능을 나타낼 수 있다고 가정하는 반면에 수행 평가는 학습자의 쓰기 자체를 대상으로 한다. 전통적인 평가는 정오답이라는 객관적인 개념을 토대로 채점되는 반면에 수행 평가는 분명히 인간적인 판단을 수반하며 평가받는 학생들에게 명백히 밝힐 것을 요구하거나 답을 설명할 수 있도록 하는 대화까지 포함한다(Wiggins, 1989).

대부분 성취 평가는 별개의 언어적 기능과 수적 기능을 측정하는 반면에 수행 평가는 훨씬 더 광범위한 능력을 평가할 수 있다. 살아가는 동안 우리가 직면하는 대부분의 문제가 개방형이고 복잡한 것처럼 수행 평가는 각각의 학생들에게 개인적인 그리고 좀 더 통합된 방법으로 자신의 숙달도를 밝히도록 한다(Archbald & Newmann, 1988).

수행 평가는 특히 성취 검사, 규범 평가 또는 선택형 평가와 비교해볼 때 학습자에게 주는 장점이 많이 있다.

1) 수행 평가는 단순히 학습자가 알고 있는 것이 아니라 학습자가 할 수 있는 것을 평가한다. 기능을 적용할 수 있는 능력에 대한 관심이 높아지면서 수행 평가는 주어진 상황에서 기능을 적용할 수 있는 능력을 평가할 수 있는 기회를 제공해준다. 잘 고안된다면 실생활에서의 교육적 수행과 업무 수행 능력을 정확히 반영할 것이다.

2) 수행 평가는 고도의 사고 기능을 개발하는데 초점을 둔다. 고도의 사고 기능은 오늘날 직장에서 필수적인 것으로 밝혀지고 있다.

3) 수행 평가는 학습자를 평가 과정의 중심에 놓는다. 대부분의 전통적인 평가는 학습자와 함께 이루어진다고 보다는 학습자에게 행해지는 것이었다. 학습자는 평가 근거에 대한 정보를 얻지 못하며, 결과를 함께 공유하지 못하며, 평가 결과를 프로그램 계획에 사용하기 위하여 학습자와 함께 이루어지는 노력이 없다. 그러나 수행 평가에서 학습자는 평가 과정의 적극적인 참여자이며 교사와 함께 특정한 평가 과업을 실질적으로 창안한다.

4) 수행 평가는 성취 검사보다 정확성을 더욱 높일 수 있는 잠재력을 갖고 있다. 수행 평가는 학습자가 할 수 있는 것에 초점을 두며 다양한 학습

유형에 대한 융통성을 허용한다. 적절히 개발된다면 성취 검사에서 성행하는 운이 좋은 결과나 임의적인 결과 요인을 배제할 수 있다. 예를 들어, 성취 검사는 일반적으로 한 번의 기회를 주고 획일화된 시간 내에 실시되며 회상 속도를 측정함으로써 느린 답변에는 감점을 준다. 그 결과 속도가 느린 학생은 당연히 낮은 점수를 받게 된다. 이 때 사람의 능력에 대한 추론이 흔히 부정확하게 이루어진다. 낮은 평가 점수가 반드시 그 영역에 관한 지식이나 기능을 갖추지 못했다는 것을 의미하지는 않는다. 다만 그는 느리게 과업을 수행할 뿐이다. 그러나 수행 평가에서는 이러한 단점을 배제시킨다.

5) 수행 평가는 대개 동료나 감독관과의 어느 정도 협력이 필요하다. 수행 평가는 내용뿐만 아니라 과정에도 관심을 둔다. 대부분 개인이나 직장에서 개별 성취와 집단 성취가 균형을 이루는 능력을 요하기 때문에 집단 상황에서 일을 잘 해낼 수 있는 능력을 개발하는 것이 중요하다.

6) 수행 평가는 개인적인 판단을 포함한다. 직장에서 수행 결과를 검토하는 것처럼 수행 평가는 순수하게 객관적인 것이 아니다. 광범위한 연습, 다수의 평가자 그리고 다양한 기법들이 결과의 일관성과 신뢰도를 높이기 위하여 창안되는 동안 수행 평가는 전통적인 선택형 평가보다 실질적으로 전문적인 판단을 요하며 개인적인 판단이 포함된다.

7) 수행 평가는 목표에 대한 명확한 상세화와 훈련이나 교육 프로그램의 목표를 필요로 한다. 명확한 목표는 학습자와 프로그램 참여자에게 기대하는 숙달도 수준을 요구하며, 교육 목표, 내용 및 평가를 일관성 있게 유지해준다.

3. 수행 평가의 유형

수행 평가의 유형은 비교적 간단한 답에서 과제를 제시하거나 보일 수 있는 장기적인 프로젝트까지 다양하다. 이러한 수행은 흔히 학생들에게 고도의 사고 능력을 요하며 여러 가지 언어 기능의 통합을 요구한다. 따라서 일부 수행 평가는 전통적인 평가보다 길고 복잡하다.

수행 평가는 학습자의 수행에 대한 직접적이고 체계적인 관찰이며 사전에 만들어진 수행 기준에 따라 수행을 평가하는 것이다. 이런 유형의 평가에서 학생들은 복잡한 과업을 수행하거나 제작물을 만들도록 요구받으며 과정과 과제 결과에 대하여 평가를 받는다.

수행 평가의 목적은 기능의 실질적인 발표를 통하여 보다 정확하고 실제적인 평가를 제공하는 것이다. 학생이 알고 있는 것뿐만 아니라 자신의 지식과 기능을 적절히 사용할 수 있는지를 평가하는 것이다.

수행 평가의 주된 특징은 학생들의 적극적인 참여를 유도하는 것이다. 제시된 선택지에서 고르는 것이 아니라 학습자 자신이 답을 창안하고 구성하는데 책임을 진다. 교사는 전통적인 지필고사, 다지선택형 또는 진위형 평가에 의존하지 않고 학습자의 지식을 평가할 수 있는 다양한 방법을 고안할 수 있다. 그러나 단순히 수행 평가를 사용해서 학습자를 평가하고 진술하는 것은 그러한 기법의 적용에 도움이 되지 않는다. 전문적인 교사는 다양한 형태의 수행 평가에 대하여 알고 있어야 하며 교실 상황에서 어떻게 사용될 수 있는지 알아야 한다. 다음 <표 2>는 수행 평가를 크게 네 가지 기본 유형으로 구분한 것이다.

표 2
수행평가의 기본 유형과 사례

전통식 평가	수행 평가			
선택형	응답구성형	산출물	수행	과정
<ul style="list-style-type: none"> ◦ 진위형 ◦ 다지선택형 ◦ 짝짓기 	<ul style="list-style-type: none"> ◦ 빈 칸 채우기 (단어, 구) ◦ 단답형 (문장, 단락) ◦ 도형 명칭 부여 ◦ 시각 자료 제시 (개념망, 순서도, 그래프, 표완성, 삽화) 	<ul style="list-style-type: none"> ◦ 논술형 ◦ 연구보고서 ◦ 실험보고서 ◦ 일지 ◦ 이야기 ◦ 연극 대본 ◦ 시 ◦ 포트폴리오 ◦ 전시회 ◦ 모형 ◦ 오디오/비디오 테이프 	<ul style="list-style-type: none"> ◦ 구두 발표 ◦ 음악/운동 ◦ 실험 ◦ 체육 대회 ◦ 연극 공연 ◦ 연기 ◦ 토론 ◦ 음악 발표 	<ul style="list-style-type: none"> ◦ 구두 질문 ◦ 관찰 ◦ 인터뷰 ◦ 회의 ◦ 과정 서술 ◦ 생각 표현 ◦ 학습 기록

교사는 다양한 유형의 수행 평가 방법을 고안할 수 있다. 수행 평가 유형의 몇 가지 예를 좀 더 구체적으로 살펴보면 다음과 같다.

1) 응답 구성형

대부분 단답형이나 빈 칸 채우기 또는 질문에 간략한 응답을 쓰도록 하는 평가를 말한다. 또한 평가에서 개방형 질문을 사용할 수도 있다. 개방형 질문은 학생들에게 선택지에서 고른다고보다는 특정한 문제에 자신의 답을 만들도록 요구하며 그 답변에 대한 생각을 설명하도록 한다. 대부분의 경우에 질문은 하나의 정답만을 갖지 않는다. 예를 들어, 참여자에게 일련의 진술을 비교하거나 대조하도록 하고, 정보를 분류하고, 자신의 선택을 설명하도록 한다. 그러나 각각의 경우에 개방형 질문은 학생들이 자신의 응답을 구성하고, 다른 방법을 통하여 해결을 찾고, 여러 가지 답이 있는 상황에 접근하고, 어떤 문제에 대한 이해의 정도를 밝히는 기회를 제공해준다. 이 모든 것들은 선택형 평가를 통해서서는 거의 불가능한 것들이다.

2) 시각 자료 제시형

학생들에게 말하거나 쓰기를 요구하지 않고 내용 지식을 끌어내기 위하여 교사는 그림, 도형, 모형, 실물, 그래프 또는 차트 등을 그리거나 조작하도록 할 수 있다. 예를 들어, 식민지 시대의 미국에 대하여 공부할 때 교사는 학생들에게 식민지의 지도와 식민지 명칭의 표를 줄 수 있다. 그 다음에 학생들은 적절한 위치에 명칭을 놓을 수 있다. 이러한 명칭 붙이기 활동은 도형, 망 그리고 삽화 등으로 교육 과정 전반에 걸쳐 사용될 수 있다. 나비에 관한 단원을 마치기 위하여 교사는 초보 ESL 학습자에게 나비의 일생을 설명보다는 그림을 그리도록 할 수 있다. 학습자는 어휘 기억에 대한 평가로 자신이 그린 그림이나 도형에 나비의 각 부분을 나타낼 수 있다. 교실에서 나비의 관찰을 기록하거나 나비의 종류, 서식 및 다른 특징에 대한 교실 자료에 대한 이해를 설명하기 위하여 그림으로 제시할 수 있다.

대표적인 시각 자료 제시형의 예로 K-W-L 차트를 이용한 방법이 있다. 많은 교사들이 한 단원을 시작하거나 끝내기 위해 K-W-L 차트(what I know/what I want to know/what I've learned)를 사용하여 성공한 사례가 많이 있다. 단원 학습 전에 이러한 전략은 교사에게 학습자의 배경 지식과 흥미에 대한 이해를 얻게 해준다. 그 다음에는 교사가 학습 내용을 평가하는데 도움이 된다. K-W-L 차트는 학급 활동이나 개인 활동으로 개발될 수도

있다. 제한된 영어 숙달도를 가진 학생들에게는 차트가 모국어나 삽화로 완성될 수도 있다.

K-W-I. 차트의 예

K

Lincoln was important.
His face is on a penny.
He's dead now.
I think Lincoln was a president.
He was a tall person.

W

Why is Lincoln famous?
Was he a good president?
Why is he on a penny?
Did he have a family?
How did he die?

L

Lincoln was president of the U.S.
He was the 16th president.
There was a war in America when Lincoln was president.
He let the slaves go free.
Two of his sons died while he was still alive.

단원 학습 이전에 교사는 학생들에게 주제에 대하여 이미 알고 있는 것 (사전 지식 또는 배경 지식)과 단원 학습 중에 알고 싶은 것에 대하여 질문함으로써 K와 W란을 채우도록 할 수 있다. 이것은 학습자가 단원을 학습하는 동안 집중하고 관심을 갖도록 도와주며 단원 학습 후에 L란을 채우고 무엇인가를 학습했다는 성취감을 줄 수 있다.

3) 논술형

논술형(essay type)은 작문, 묘사, 분석, 설명 또는 요약 등을 통해서 어떤 주제에 관한 학습자의 이해를 평가하는데 오랫동안 사용되어왔다. 논술형은 학생이 주어진 상황에서 사실을 어떻게 잘 사용하며 일관적인 논의를 어떻게 잘 구조화시키는지 보여줄 수 있다. 논술형 문제에 효과적으로 답을 하기 위해서는 비판적인 사고, 분석 그리고 종합 능력이 필요하다.

논술이나 다른 쓰기 평가 방법도 철자, 문법, 통사 및 문장이나 단락 구조를 포함한 학습자의 작문 능력을 평가하는데 사용될 수 있다. 쓰기 평가의 표준화되고 객관적인 채점에 관해서는 상당한 연구가 이루어졌다.

4) 연구 보고서, 전시회, 발표

이런 유형의 공통점은 일정 기간에 걸쳐 일어나는 복잡한 과업에 대한 성취 결과를 평가하는 것이다. 즉, 학습한 내용을 발표하거나 전시하도록 한다. 가장 친숙한 형태의 수행이 모의 직업 인터뷰나 현장 평가이다. 주어진 시간 동안 직장에서의 인터뷰 수행은 분명히 정의된 기준에 따라 평가된다. 발표는 다양한 기능 중심의 기본 기능을 평가하는데 사용될 수 있다. 예를 들어, 학생들에게 지시 책자를 읽도록 한 다음에 그에 일치하는 활동을 수행하도록 할 수 있다. 이런 유형의 연습은 학생들에게 정보를 읽고 이해하고 적용하는 능력을 보이도록 한다. 기본적인 기능의 사용 능력에 대한 강조가 높아지면서 일부 형태의 발표가 평가에서 필수적인 요소로 포함되고 있다. 연구 보고서(project)는 기능이나 지식을 종합적으로 보여준다. 연구 보고서는 광범위한 능력을 요구하며 흔히 초점이 학제간이며 학생들의 주도권과 창의성을 요한다. 교사나 교육받은 채점관이 미리 모든 참여자에게 알려준 기준에 따라 연구 보고서를 채점한다.

연구 보고서의 일부로 학생들에게 교실이나 다른 사람들 앞에서 발표를 하거나 실제 수행을 하게 할 수 있다. 연구 보고서는 학생간 또는 집단간에 경쟁 형태를 띠거나 학생들이 일정 기간에 걸쳐 이루어낸 협동 활동이 될 수도 있다. 집단 연구 보고서는 다수의 학생들에게 계획, 연구, 토론이나 집단 발표를 요하는 복잡한 문제에 대하여 함께 작업을 할 수 있게 한다. 이런 기법은 특히 협동심을 촉진시키고 가치 있는 결과를 강화시키기 때문에 매력적이다.

5) 실험실 활동

학습자가 과학적인 개념을 얼마나 이해하고 과학적인 과정을 잘 수행할 수 있는지를 평가하기 위하여 실험이 사용될 수 있다. 그러한 평가 활동은 학생들에게 가설을 세우고, 계획하고 실험을 실행하고, 결과를 기록하고, 측정이나 평가 기술을 사용하고, 과학적 사실과 개념을 적용함으로써 과학을 하도록 권장한다.

학생들은 과학 실험실에서처럼 일련의 과업을 주어진 시간 내에 개인적으로 또는 집단으로 진행해간다. 질문은 학생들의 사고 전략을 끌어내기 위하여 개방형으로 구성되는 것이 일반적이다.

6) 포트폴리오

포트폴리오(portfolio)는 일정 기간에 걸쳐 이루어진 대표적인 학업 수행의 수집물을 말한다. 포트폴리오는 흔히 학생이 가장 잘한 일을 기록하고, 다양한 유형의 과정 정보를 포함할 수도 있다(예를 들면, 학습자가 작성한 초안, 학습자 자신의 작품에 대한 평가, 부모의 평가). 포트폴리오는 학습자의 능력이나 향상도를 평가하기 위해 사용될 수도 있다. 그러므로 포트폴리오는 어떤 주제에 대한 학습자의 향상도를 밝히기 위해 순서에 따라 완성되는 몇 개의 과제로 구성된다. 포트폴리오에는 초기 계획, 초안, 자기 평가, 동료나 교사의 피드백, 차후 계획 등을 포함할 수 있다.

포트폴리오 평가는 학생의 과제를 수집하는 것이며 주된 목적이 학생의 향상을 밝히는 것이다. 이러한 절차의 사용은 특히 쓰기와 관련된 언어 영역에서 확산되고 있다. 과제에 대한 평가에 학생을 참여시키며 피드백을 제공하기 위한 것이다. 교사와 동료 학생의 검토가 중요하다. 포트폴리오 평가의 가장 큰 장점은 학생들을 독립적인 사고자가 될 수 있도록 하는 것이며, 학습자로서 자율성의 신장이 촉진된다는 점이다.

포트폴리오는 단순히 학생 과업의 수집 그 이상의 의미가 있다는 것을 기억하는 것이 중요하다. 다양한 유형의 포트폴리오가 있다. 예를 들어, 과제 포트폴리오, 수행 포트폴리오, 평가 포트폴리오, 집단 포트폴리오, 지원 포트폴리오 등이 있다. 목적에 따라 다음 항목 중에서 선택할 수도 있다. 창의적

과제, 평가, 퀴즈, 숙제, 연구 보고서와 과제, 구두 발표 테이프, 학생 일기, 자기 평가, 동료나 교사의 논평도 있다.

7) 구두 수행이나 발표

수행 평가에는 인터뷰, 구두 보고, 역할 놀이, 묘사하기, 설명하기, 요약하기, 다시 말하기, 이야기나 교재를 풀어서 말하기 등이 포함된다. 구두 평가는 이해와 사고 기능을 확인하기 위하여 진행 위주로 실시되어야 한다.

영어 숙달도와 내용 지식을 판단하기 위하여 초기 단계의 학생들에게 인터뷰를 실시할 때 교사는 가능한 시각 단서를 사용하고 응답에서 최소한의 영어를 허용하도록 하는 것이 바람직하다. Pierce와 O'Malley(1992)는 학생들이 말하고 싶은 한 두 개의 그림을 선택하도록 하여 특히 학문적인 언어(비교하기, 설명하기, 묘사하기, 분석하기, 가설 세우기 등)나 주제에 적절한 어휘의 사용을 끌어내는 질문으로 학생들을 유도해 나가는 방법을 제안하고 있다.

8) 구두와 쓰기 산출물

ESL 학습자의 향상을 평가하기에 유용한 구두 또는 쓰기 산출물은 내용 영역에 관한 사고 및 학습 측정, 읽기 반응 측정, 쓰기 숙제(구조적이며 창의적인), 대화 일지 및 오디오 비디오 카세트 등이 있다.

“내용 영역 측정”은 학습자가 설명문의 글을 읽을 때 초인지적 전략의 사용을 권장하기 위하여 고안된 것이다. 항목은 “내가 이해한 것과 내가 이해하지 못한 것(생각이나 어휘)”이라는 두 개의 제목이 있는 형태로 만들어질 수 있다.

“읽기 반응 측정”은 문학 작품에 대한 학습자의 반응을 측정하기 위해 사용될 수 있다. 비판적 사고를 권장하는 문학 작품에 대한 어떤 일반적인 또는 특이한 질문에 반응하도록 하거나 지면의 한 쪽에는 간략한 글이 복사되어 있고 다른 쪽에는 글에 대한 학습자의 생각을 쓰도록 할 수도 있다.

초보 단계의 ESL 학습자들은 흔히 설명적인 쓰기 과제가 통제되거나 구조화될 때 성공을 경험하게 된다. 교사는 학습자에게 토론, 브레인스토밍, 망 그리기, 개요 작성 등을 포함한 사전 쓰기 단계를 통하여 학생들을 유도할 수

있다. 독립적인 쓰기 산출물뿐만 아니라 사전 쓰기의 결과도 평가될 수 있다.

“대화 일지”는 교사와 학생 사이의 상호적이며 진행중인 소통 수단을 제공한다. 학습자는 주제 선택을 결정하고 그들 수준의 영어 숙달도로 참여한다. 초보자는 교사에 의해 언급되는 그림을 그릴 수 있다.

“오디오 비디오 카세트”는 학습자의 구두 읽기, 발표, 극, 인터뷰 또는 회의(교사 또는 동료) 등으로 만들어질 수 있다.

9) 집단 수행 평가와 개별 수행 평가

집단 수행 평가는 평가 과업을 집단으로 수행하도록 요구하는 평가를 말한다. 예를 들어, 개별 수행 평가로 알려진 평가가 집단으로 실행되면 집단 수행 평가가 되며, 개별 평가는 해당 집단의 부분으로서 각 개인의 수행을 토대로 이루어진다.

공동 협력 기술과 다른 사람과 함께 일하는 능력이 필수적인 업무 기능으로 점차 인식되고 있다. 그러한 기능을 가르치고 평가하는 한 가지 방법은 학생들이 협동적으로 일하도록 요구하는 프로젝트를 통해서이다. 교내 프로그램 가운데서 가장 보편적인 형태의 집단 프로젝트는 학생 소식이이다. 이는 학생들에게 하나의 산출물을 개발하고 만들어내기 위하여 팀이 되어 일하도록 하며 그들의 기본 기능을 여러 가지 다른 방법으로 적용할 수 있게 해준다(쓰고, 편집하고, 비용을 계산해보는 등). 젊은이들이 함께 사회 봉사 활동이나 다른 형태의 현장 체험에 참여하는 집단 프로젝트는 상황에서 기능을 배우고 평가하는 기회를 제공해준다. 집단 프로젝트는 기본 기능을 직장에서의 실제 과업에 적용함으로써 기본 기술, 공동 협력 그리고 많은 젊은이들을 위해 일하는 것과 연결되도록 도와준다. 이러한 상황에서의 평가는 업무 산출물이나 각 개인이 집단 내에서 어떤 일을 했는지를 나타낼 것이다.

IV. 영역별 수행 평가 항목의 예시

교과 내용이나 학습자의 수준에 따라 교사는 다양한 수행 평가 항목을 선정할 수 있다. 다음은 각 영역별 수행 평가의 예시이다. 수업 시간에 이루어

지는 모든 활동은 수행 과제로 제시할 수 있으며 각 영역이 독립적으로 평가되는 경우도 있지만 대부분 두 개 이상의 과업이 통합하여 이루어지는 경우가 더 자연스러운 활동이 될 수 있다.

듣기

- 들려주는 말을 듣고 어휘를 식별할 수 있다.
- 들은 내용에 알맞은 그림을 선택할 수 있다.
- 들은 내용에 알맞은 그림을 그릴 수 있다.
- 지시한 내용에 따라 행동으로 옮길 수 있다.
- 이야기의 내용을 요약할 수 있다.
- 이야기의 일부를 듣고 결론을 도출할 수 있다.
- 들려주는 말을 듣고 표나 그림을 완성할 수 있다.
- 역할극이나 게임을 듣고 이해할 수 있다.
- 간단한 대화를 듣고 중심 내용을 이해할 수 있다.

말하기

- 자신을 소개할 수 있다.
- 간단한 인사를 주고받을 수 있다.
- 그림이나 실물을 보고 이를 묘사할 수 있다.
- 주어진 상황에 적절한 의사 소통 표현을 말할 수 있다.
- 사물의 위치, 크기, 형태 등을 말할 수 있다.
- 대화를 듣고 자신의 느낌을 말할 수 있다.
- 사물의 유사점과 차이점을 말할 수 있다.
- 글을 읽고 요약하여 말할 수 있다.
- 일정표를 보고 간단히 말할 수 있다.
- 간단한 인터뷰에 답할 수 있다.
- 주어진 상황에서 자신의 의견을 말할 수 있다.
- 역할극을 할 수 있다.
- 구두로 보고하거나 발표할 수 있다.

어휘

- 모르는 단어의 의미를 문맥 단서를 활용하여 추론할 수 있다.
 - 접두사와 접미사의 의미를 알 수 있다.
 - 사전을 활용하여 주어진 어휘에 대한 정의를 제한된 시간 내에 찾을 수 있다.
 - 주어진 단어를 의미망을 활용하여 확대하여 나갈 수 있다.
 - 동의어나 반의어를 알 수 있다.
 - 다의어의 의미를 주어진 문맥에서 찾아낼 수 있다.
 - 어휘 퍼즐을 완성할 수 있다.
 - 제시된 단어를 범주별로 분류할 수 있다.
 - 영영사전식 어휘 정의를 읽고 해당 단어를 알 수 있다.
-

읽기

- 글을 읽고 요지를 찾을 수 있다.
 - 자신의 답에 대한 근거를 말할 수 있다.
 - 글에서 중요한 문장을 찾아낼 수 있다.
 - 세부 사항을 찾아낼 수 있다.
 - 지시문을 읽고 지시대로 할 수 있다.
 - 광고문이나 실용문을 읽고 이해할 수 있다.
 - 글의 일부를 읽고 전후 내용을 추론할 수 있다.
 - 글을 읽고 결론을 도출할 수 있다.
 - 글을 순서대로 놓을 수 있다.
 - 글의 성격과 목적을 이해할 수 있다.
 - 글에 나타난 원인과 결과를 찾을 수 있다.
 - 문학 작품을 읽고 감상할 수 있다.
 - 글에 나타난 사실과 의견을 구분할 수 있다.
 - 이야기를 요약할 수 있다.
 - 글쓴이의 의도를 파악할 수 있다.
 - 영자 신문이나 잡지를 읽고 이해할 수 있다.
-

쓰기

- 간단한 대화나 글을 받아쓸 수 있다.

- 주어진 문장을 환문할 수 있다.
- 자신을 소개하는 글을 쓸 수 있다.
- 그림이나 사물을 보고 글로 쓸 수 있다.
- 통제 작문을 완성할 수 있다.
- 빈 칸 메우기 방법을 통해 글을 만들 수 있다.
- 어휘 퍼즐을 만들 수 있다.
- 주어진 글을 요약하여 쓸 수 있다.
- 영문 서식을 작성할 수 있다.
- 구직 신청서나 이력서를 작성할 수 있다.
- 영문편지나 E-mail을 쓸 수 있다.
- 초대장을 쓸 수 있다.
- 일기를 쓸 수 있다.
- 자신의 생각이나 느낌 또는 의견을 글로 표현할 수 있다.
- 연구 보고서를 쓸 수 있다.

다음은 각 영역별 수행 평가의 예시이다. 교사는 학습자의 수준에 따라 과제의 유형이나 문제를 변형시킴으로써 난이도를 조정할 수 있으며, 평가 척도도 다양하게 설계할 수 있을 것이다.

〈예시 1 - 듣기 수행 과제〉

(가) 활동 목표 : 연산에 관한 내용 듣고 이해하기

(나) 활동 유형 : 개인별 과제 수행

(다) 활동 시간 : 5분

(라) 절차

- ① 학생들에게 자료를 나누어준다.
- ② 교사나 녹음된 음성으로 2회 반복하여 들려준다.
- ③ 듣고 해당 번호에 계산식을 쓴다.

(마) 문제

□ 다음은 수 계산에 관한 내용이다. 잘 듣고 계산식을 쓰시오.

• Listening Script

1. Three plus two equals five.
2. Eight minus one equals seven.
3. Four times twenty five equals one hundred.
4. Twelve divided by two makes six.
5. Three point two plus three point two five makes six point four five.
6. One-half plus a quarter makes three quarters.

①	
②	
③	
④	
⑤	
⑥	

◆ **재검 기준**

평가 관점	평가 척도	평가 기준
답안 내용의 정확성	상	다섯 가지 이상의 내용이 정확하다.
	중	3-4개의 내용이 정확하다.
	하	2개 이하의 내용만 정확하다.

◀정답▶ ① $3 + 2 = 5$ ② $8 - 1 = 7$ ③ $4 \times 25 = 100$ ④ $12 \div 2 = 6$
 ⑤ $3.2 + 3.25 = 6.45$ ⑥ $\frac{1}{2} + \frac{1}{4} = \frac{3}{4}$

〈예시 2 - 듣기 수행 과제〉

- (가) 활동 목표 : 비교 표현 이해하기
 (나) 활동 유형 : 개인별 과제 수행
 (다) 활동 시간 : 3분
 (라) 절차

- ① 학생들에게 자료를 나누어 준다.
- ② 교사나 녹음된 음성으로 2회 반복하여 들려준다.
- ③ 듣고 해당 번호에 숫자를 적는다.

(마) 문제

다음 들려주는 말을 잘 듣고 빈 칸에 해당하는 숫자를 쓰시오.

• Listening Script

Let me tell you about my family. There are five children in our family. My name is Bob and I'm 26 years old. My sister Jane is two years older than I and my brother John is a year older than Jane. Mary is two years younger than I and Bill is only 21.

Name	Age
John	①
Jane	②
Bob	③
Mary	④
Bill	⑤

◆ 재검 기준

평가 관점	평가 척도	평가 기준
답안 내용의 정확성	상	모두 정확하게 맞았다.
	중	3-4개의 내용이 정확하다.
	하	2개 이하의 내용만 정확하다.

◀정답▶ ① 29 ② 28 ③ 26 ④ 24 ⑤ 21

〈예시 3 - 듣기 수행 과제〉

- (가) 활동 목표 : 듣고 행동으로 나타내기
- (나) 활동 유형 : 개인별 또는 조별 과제 수행
- (다) 활동 시간 : 5분

(라) 절차

- ① 개인 또는 조별 학생을 선정한다.
- ② 교사의 지시에 따라 행동으로 나타낸다.
- ③ 교사는 문장을 변형시켜 지시할 수 있다.

(마) 문제

다음 지시에 따라 행동으로 나타내시오.

Stand up.

Stretch.

Touch your right ear(leg, foot, arm, etc).

Touch your both shoulder(ears, knees, feet, etc).

Point to the blackboard(ceiling, floor, eraser).

Point to the right-hand corner of the blackboard.

Walk forward. Stop!

Smile. Laugh!

Shake your right hand.

Make your fist.

Open your hand.

Hit your head with your left hand.

Wave your right hand.

Sit down.

◆ 재검 기준

평가관점	평가 척도			반영 비율
	상	중	하	
정확성	4/5이상이 정확하였다.	3/5정도 정확하였다.	2/5이하만 정확하였다.	70%
반응속도	반응 속도가 매우 빠르다.	반응 속도가 빠른 편이다.	반응 속도가 느린 편이다.	30%

〈예시 3 - 말하기 수행 과제〉

(가) 활동 목표 : 설문 자료 만들기(Making a questionnaire)

(나) 활동 유형 : 개인별 또는 조별 과제 수행

(다) 활동 시간 : 10-15분

(라) 절차

- ① 개인 또는 조별 학생을 선정한다.
- ② 교사는 자료를 나누어주고 해당 항목과 관련된 다양한 질문을 만들도록 한다.
- ③ 학생들의 과업을 돕기 위하여 모형을 제시한다.
- ④ 각 항목에는 다양한 질문이 가능하다는 것을 알려준다.

(마) 문제

□ 다음 각 항목과 관련된 질문을 만들어 보세요.

항 목	질 문
Name	① May I have your name, please?
Birthday	②
City or Town	③
School	④
Family	⑤
Reasons for Studying English	⑥
Plan for the future	⑦
Free-time activities	⑧

◀정답 예시▶ 각 항목과 관련된 질문은 다양하게 만들 수 있다. 예를 들어, ⑤의 경우에는 다음과 같은 표현을 사용하여 가족 사항을 물을 수 있다. Will you tell me about your family? How many members are there in your family? Do you have a sister? How about your family?

◆ 재검 기준

평가관점	평가 척도			반영 비율
	상	중	하	
표현력	7개 항목 이상에서 질문을 만들었다.	4-6개 항목에서 질문을 만들었다.	3개 이하 항목에서 질문을 만들었다.	70%
유창성	휴지나 망설임이 거의 없다	휴지나 망설임이 다소 있었다.	휴지나 망설임의 빈도가 많았다.	30%

(예시 4 - 어휘 수행 과제)

(가) 활동 목표 : Silent Quiz

(나) 활동 유형 : 조별 또는 개인별 과제 수행

(다) 활동 시간 : 10분

(라) 절차

- ① 2인 1조 또는 개인별 학생을 선정한다.
- ② 교사는 자료를 나누어주고 임의의 순서에 따라 행동으로 표현한다.
(조별로 할 경우에는 한 학생은 행동으로 표현하고 다른 학생은 답을 적는다. 퀴즈 I 을 마친 다음에는 역할을 바꾸어 다음 퀴즈를 해결한다.)
- ③ 행동으로 옮기는 교사(또는 학생)는 먼저 임의의 순서를 정하고 번호를 불러가면서 행동에 옮기면 답하는 학생은 해당 단어를 찾아 번호를 적는다.

(마) 문제

□ 행동에 맞는 단어 앞에 번호를 기입하시오.

Silent Quiz I

- | | | |
|-----------|------------|------------------|
| ___ bite | ___ clap | ___ sneeze |
| ___ laugh | ___ tiptoe | ___ shake |
| ___ catch | ___ point | ___ try to reach |
| ___ chew | ___ smile | ___ wave |

Silent Quiz II

___ bend ___ cry ___ throw
 ___ blow ___ frown ___ wink
 ___ climb ___ kick ___ yawn
 ___ cough ___ shrug ___ twist

◆ 재점 기준

평가관점	평가 척도			반영 비율
	상	중	하	
정확성	18개 이상 맞았다.	12-17개가 맞았다.	11 이하만 맞았다	70%
반응속도	반응 속도가 매우 빠르다.	반응 속도가 빠른 편이다.	반응 속도가 느린 편이다.	30%

〈예시 5 - 어휘 수행 과제〉

(가) 활동 목표 : 어휘 퍼즐 풀기(Working a puzzle)

(나) 활동 유형 : 개인별 또는 조별 과제 수행

(다) 활동 시간 : 5분

(라) 절차

- ① 개인 또는 조별 학생을 선정한다.
- ② 교사는 어휘 퍼즐 문제를 준다.
- ③ 개인별 또는 조별로 문제를 해결한다.
- ④ 학생들은 사진을 사용할 수 있다.

(마) 문제

□ 다음 어휘 퍼즐을 완성하십시오.

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								

Across

- 1A What language are you studying now?
 3C There is ___ apple on a table.
 4A What is the same word of "unhappy"?
 4E What is another way to write "OK"?
 6A What is after first?
 7F How much is five minus four?
 8A What is the short form of Texas?
 8D What does a chicken lay?

Down

- A1 What is the opposite of "begins"?
 A6 The past form of "sit" is _____.
 B4 Tommy and Jane _____ my friends.
 C1 How do you feel when a friend sends you an interesting letter?
 D5 I always _____ my homework before I watch TV.
 F1 Susan didn't come to school because she was _____.

- F6 What animal likes to run after cats?
 G4 What is the short form for morning?
 H2 What is the fifth month?
 H6 How many fingers do you have?

◆ 재검 기준

평가 관점	평가 척도	평가 기준
답안 내용의 정확성	상	15개 이상 맞았다.
	중	11-14개 맞았다.
	하	10개 이하 맞았다.

◀정답▶

	A	B	C	D	E	F	G	H
1	E	N	G	L	I	S	H	
2	N		L			I		M
3	D		A	N		C		A
4	S	A	D		O	K	A	Y
5		R		D			M	
6	S	E	C	O	N	D		T
7	A					O	N	E
8	T	X		E	G	G		N

(에서 6 - 읽기 수행 과제)

(가) 활동 목표 : 읽고 그리기(Read and Draw)

(나) 활동 유형 : 개인별, 조별

(다) 활동 시간 : 5-10분

(라) 절차

- ① 읽기 자료를 나누어준다.

② 학습자는 읽고 글의 내용에 맞게 그림을 그린다.

(마) 문제

□ 다음 글을 읽고 내용과 일치하는 그림을 그려보세요.

Below is a description of a spaceship by somebody who said that he had seen it. Read and draw a spaceship according to the description.

The spaceship was about two meters high. It was round at the bottom and pointed at the top. The spaceship was supported by three legs which were taller than the main body of the ship. There was a ladder against the door. I saw a small man with two horns. I remember a mysterious sign on the side of the spaceship. It showed a triangle surrounded by a circle.

◆ 재점 기준

평가관점	평가 척도			반영 비율
	상	중	하	
과업수행	6개 이상 항목이 맞게 그려져 있다.	4-6개 항목이 맞게 그려져 있다.	3개 이하의 항목이 맞게 그려져 있다.	70%
정확성	그림이 정확하게 맞았다.	1-2곳 잘못된 곳이 있다.	3개 이상 잘못된 부분이 있다.	30%

◀정답 예시▶ 다음 사항이 그림에 포함되어야 한다. 1) 우주선 아랫부분이 둥글다. 2) 윗부분은 뾰족하다. 3) 3개의 다리가 있다. 4) 다리가 몸체보다 길다. 5) 사다리가 문에 기대어 있다. 6) 두 개의 뿔이 달린 남자가 있다. 7) 원안에 삼각형이 그려진 마크가 있다.

〈예시 7 - 쓰기 수행 과제〉

(가) 활동 목표 : 빈 칸 메우기(Cloze Test)

(나) 활동 유형 : 개인별

(다) 활동 시간 : 5분

(라) 절차

- ① 여섯 번째 단어마다 생략된 글을 나누어준다.
- ② 학습자는 학습한 내용이나 문맥을 통해서 빈 칸을 완성한다.
- ③ 때에 따라서는 정답이 두 개 이상 허용된다는 것을 알려준다.
(수준이 낮은 학생들에게는 보기를 줄 수 있다.)

(마) 문제

다음 글의 빈 칸을 완성하십시오.

Reasons for Exercising

Why does a person exercise? ① _____ people exercise for their good
② _____. Other people exercise to burn ③ _____ and lose weight.
They want ④ _____ look better. Still other people ⑤ _____ for fun
and relaxation. In ⑥ _____, most people probably exercise for ⑦ _____
three reasons.

◆ **재검 기준**

평가관점	평가 척도			반영 비율
	상	중	하	
과업완성	6개 이상 맞았다.	4-5개 맞았다.	3개 이하 맞았다.	70%
정확성	철자 오류가 거의 없다.	철자 오류가 다소 있다.	철자 오류가 많이 나타난다.	30%

◀정답▶ ① Some(Most) ② health ③ calories ④ to ⑤ exercise
⑥ fact ⑦ all(these)

(예시 8 - 쓰기 수행 과제)

- (가) 활동 목표 : 어휘 퍼즐 만들기(Making a puzzle)
(나) 활동 유형 : 조별 과제 수행
(다) 활동 시간 : 10분

(라) 절차

- ① 조별 인원 및 학생을 선정한다.
- ② 교사는 어휘 퍼즐 문제를 준다.
- ③ 학생들은 조별로 퍼즐 문제를 만든다.
- ④ 학생들은 사전을 사용할 수 있다.

(마) 문제

다음 어휘 퍼즐을 만들어 보시오.

	A	B	C	D	E	F
1	B	O	A	T		C
2	U		L			O
3	Y	E	L	L	O	W
4		A				
5		S			U	N
6	S	T	O	P		O

Across

- 1A What do people use to travel on water?
 3A _____
 6A _____
 5E _____

Down

- A1 _____
 B3 _____
 C1 _____
 F1 _____
 F5 _____

◆ 재검 기준

평가관점	평가 척도			반영 비율
	상	중	하	
과업 완성	7개 이상 맞았다.	5-6개가 맞았다.	4개 이하만 맞았다	60%
표현력	오류가 거의 없고 다양한 표현 사용	오류가 다소 있으며 표현이 다양하지 못함	오류가 많이 있고 표현 능력이 부족함	40%

V. 결론 및 제언

영어 교수학습 과정에서 평가가 차지하고 있는 기능의 비중은 매우 막중하다(임병빈, 1997). 평가와 교수학습의 관련성을 적절하게 설명하고 있는 Oller(1987)의 견해에 의하면, 교수학습과 평가란 불가분의 관계 속에서 상호 작용을 통하여 궁극적으로 학습자에게 보다 나은 학습효과를 도출시킬 수 있는 유기적인 순환체계로 간주된다.

외국어 교육의 교수방법 변천사에서 볼 때, 언어이론과 학습이론의 변화에 따라 시대적으로 다양한 교수방법이 대두되었던 것처럼, 평가에 있어서도 커다란 흐름의 변화 양상을 요약해 보면 결국 '전과학적(prescientific) 시기의 전통식 평가방법', '구조주의 시기의 심리측정-구조주의적 분리식 평가방법', '심리언어학적-사회언어학적 경향의 통합식 평가방법'으로 전개되어 왔다(임병빈, 1995). 한편, 국내의 영어교육 평가 동향을 살펴보면 적용 시기상의 차이는 있을지라도 대체적으로 위에서 언급한 평가 경향을 반영하여 왔다. 즉, 오랫동안 전통식 평가에 안주해오다가 그 후 분리식 평가가 유행하였으며 1990년대에 들어 소위 '의미중심 평가(meaning-oriented test)' 형태의 통합식 평가가 기초를 이루어 왔다. 특히 국내 영어 교육의 실제적 방향을 좌우하는 대학 입학 영어시험(대입 예비고사-학력고사-수능시험)에서도 이와 같은 평가 형식의 반영이 여실히 나타나, 수 년 전부터는 듣기 평가를 포함시킨 통합적 형식의 의미 중심 평가 형태로 자리잡고 있다. 이와 더불어 최근에는 초·중등 학교의 학교 교육과정에서 교수학습 방법 및 평가 방법 개선을 통

한 학습자의 학습 효과를 증진시키기 위하여 교육부는 '수행평가' 제도를 정책적으로 추진하고 있는 상황이다.

물론, 수행평가 방식이 모든 영어교육 상황에 최선의 기대효과를 거둘 수 있는 것이 아님은 자명한 일이며, 현재와 앞으로도 마치 의학 분야에서 기적의 신약처럼 모종의 특효약이 될 수도 없다. 다만 우리는 현재보다 좀 더 나은 학습 효과를 도출시키고자 학습과 관련한 무수한 변인들을 대상으로 부단한 실험을 계속하고 있는 셈이다. 수행평가도 바로 이와 같은 관점에서 인식의 기초를 두고 필요한 변화의 패러다임 속에서 소위 '플러스 알파(+ α)' 효과를 지향해야 할 것이다. 따라서 본 연구에서는 기존의 평가 방식과 교수 방법의 변화 과정에서 수행평가의 필요성 및 발달 배경을 논의하였으며, 교수학습 과정에서 보다 바람직한 수업 및 과제 활동을 통한 학습의 기대 효과를 창출할 수 있는 수행평가의 특징 및 유형을 제시하였다. 동시에 영어의 영역별 수행평가 항목 및 구체적인 의사소통 기능 시험의 예시 문항도 제시되어 현직 교사들에게 수행평가 출제시에 참고가 될 수 있도록 하였다.

새로운 천년을 눈앞에 둔 우리 사회는 모든 분야에서 '무한경쟁', '정보화', '다변화' 등의 술한 표어를 내걸고 전방위적인 노력을 경주하고 있는 시대이다. 따라서, 우리의 영어교육도 새로운 시각인 '수행평가'를 내실있고 체계적인 과정으로 발전시켜 타의에 의한 것이 아닌 자율 의지에 의하여 플러스 알파 효과를 보낼 수 있도록 하나씩 하나씩 문제점을 해결하여 궁극적으로 모처럼 구축되어 가는 영어 교육의 활성화에 긍정적이고 능동적인 촉매 역할을 기대한다.

참 고 문 헌

- 임병빈. (1995) 고등학교 영어시험 출제상의 문제점. *영어교육*, 50(1), 1-2.
- 임병빈. (1997). *영어교수학습*. 보성출판사.
- Archibald, Doug A., & Newmann, Fred M. (1989). *Beyond standardized testing*. Reston, Virginia: National Association of Secondary School Principals.
- Barth, P., & Mitchel, R. (1992). *Elementary education for the 21st century*, Golden Colorado : North American Press.
- Brecht, R., & Walton, R. (1993). *National strategic planning in the less commonly taught languages*. Washington, DC: National Foreign Language Center.
- Dietel, R. L., Herman, J. L., & Knuth, R. A. (1991). What does research say about assessment? NCREL, Oak Brook.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Glickman, Carl. (1991). Pretending not to know what we know. *Educational Leadership*, May, 4-10.
- Gruber, H. (1985). Giftedness and moral responsibility: Creative thinking and human survival. In Horowitz, F., & O'Brien, M., (eds.), *The gifted and the talented: Developmental perspectives*. Washington, DC: American Psychological Association.
- Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics*, 15, 212-226.
- Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked question. *TESOL Journal*, 5, 8-10.
- Lipman, M. (1987). Some thoughts on the formation of reflective education. In J.B. Baron., & R.J. Sternberg. *Teaching-thinking skills: Theory and practice*. pp. 151-161. New York: W. H. Freeman, 1987.

- Mislevy, Robert J. (1989). *Foundations of a new test theory*. Princeton, New Jersey: Educational Testing Service.
- Newmann, Fred M., & Wehlage, Gray G. (1993). Five standards of authentic instruction. *Educational Leadership*, April, 8-12.
- Oller, J. W. Jr. (1987). Practical ideas for language teachers from a quarter century of language testing. *English Teaching Forum*, 25(4), 42-46.
- Perkins, D. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.
- Pierce, L. V., & O'Malley, J. M. (1992). *Performance and portfolio assessment for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Resnick, L. B., & Resnick, D. P. (1989). *Assessing the thinking curriculum: New tools for educational reform*. Pittsburgh, Pennsylvania: Learning Research and Development Center: University of Pittsburgh and Carnegie Mellon University.
- Tannenbaum, Jo-Ellen. 1996. Practical ideas on alternative assessment for ESL students. Eric Digest. No ED395500.
- Wiggins, Grant. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 703-713.
- Wolf, D. P. (1987). Opening up assessment. *Educational Leadership*, 45(4), 24-29.