

영어어문교육 12권 3호 2006년 가을

ESL 학습자의 쓰기배치고사상의 오배치에 따른 이해도 차이 연구

김 정 태

(University of Illinois at Urbana-Champaign)

Kim, Jung Tae. (2006). Filling the understanding gap of the misplacement of ESL learner's writing placement test. *English Language & Literature Teaching*, 12(3), 147-166.

This study investigates the effect of misplacement in a written Computer-delivered ESL Placement Test (CEPT) context. The study aims to address the following two research questions: a) which scoring rubric features cause the misplacement of ESL learner's writing scores? and b) which scoring rubric features improve ESL learner's writing scores? Thirty-four international examinees took the test and participated in surveys at the University of Illinois. Twelve examinees of them attended the CEPT workshop test. In the workshop test, they carried out self-evaluation on their first essays using a scoring rubric and compared with expert raters' results. After the workshop, the examinees responded to a survey and interview. For the first research question, the results of the survey and interview addressed that the majority disagreed with the raters' rating results. The self-evaluation results also indicated that their misunderstanding of organization feature caused the misplacement. For the second question, the CEPT workshop scores were improved due to the score improvement in the organization feature while the contribution of other features to the total scores was little. Most of the examinees pointed out that a lesson on the scoring rubric enhanced their understanding of the writing features of the rubric so that their placement scores were generally improved.

[misplacement/writing assessment/self-assessment, 오배치/쓰기
평가/자기평가]

I. 서론

영어평가 영역에서 쓰기와 말하기에 대한 중요성은 지속적으로 강조되었으나 실제 평가에 포함된 것은 최근이다. 특히 미국에 소재한 대학교에 입학하기 위해서 수험자들은 TOEFL(Test of English as a Foreign Language)이라는 시험을 보아야 하며, 2000년부터 영작문 평가인 TWE(Test of Written English)는 필수 요소가 되었다. 또한 한국에서 많은 산업체에서 사원들의 영어 실력 평가에 사용되고 있는 TOEIC(Test of English for International Communication)에도 쓰기 평가가 2006년 연말부터 포함된다고 한다. 이러한 쓰기 평가의 중요성은 미국의 각 대학교에서도 마찬가지로 강조되며, 입학생들의 TOEFL 성적 이외에 작문 배치 평가를 시행하여 그 결과에 따라 ESL(English as a Second Language) 교실에 배치하고 있다. 예를 들어 미국 일리노이대학교(University of Illinois at Urbana-Champaign)에서는 ESL Placement Test(EPT)로 모든 외국 입학생들을 대상으로 대학교 수준의 학술 분야의 영작문 실력 평가를 하고 있다. 이처럼 쓰기 평가의 중요성이 강조되는 이유는 수험자들의 실제 학업 또는 업적 성취도와 직결되기 때문이다(Kim, 2003; Steward, 1999).

이러한 추세 속에서 영어 평가 전문가들의 목표는 수험자의 작문 실력을 정확하게 평가하는 쓰기 평가 시스템을 구축하는 것이다. 그 목표는 평가 개발자들이 시험문제를 개발할 때 실제 수험자들의 학습 환경을 최대한 반영하여 시험문제들의 시험타당도(test validity)를 높이는 것이다(Kim, 2003; Lee, 2005). Bachman과 Palmer(1996)는 시험 문제들의 내용이 목표언어단위 영역(Target Language Unit Domain)과 일치시킬 것을 권장했다. Douglas(2000)는 그의 책 “Assessing languages for specific purposes”에서 시험 평가 내용이 측정하고자 하는 언어 영역과 일치할 때 시험의 내용타당도(content validity)가 높아진다고 주장했다.

하지만 영어 평가 개발자들의 노력에도 불구하고 많은 수험자들은 자신의 영작문 평가 점수에 대한 불만을 표출해오고 있다. 즉, 잘못된 시험 결과로 인해 자신이 적절하지 않은 ESL교실에 배치되었다고 주장했으며, 그런 불만을 품은 수험자들의 학습 수행 성취도는 그리 높지 않은 것으로 나타났다(Steward, 1999). 여러 학자들은 시험결과에 대해 하나의 숫자로 표기된 점수를 주기보다 풍부한 채점자의 피드백을 주어야 한다고 주장하고 있으며 그 주장은 더욱 설득력을 얻고 있다(e.g., Bachman 1990; Kim, 2003, Lee, 2005). 특히 Kim(2003)은 그의 논문에서 수험자의 채점기준표(scoring rubric) 요소(feature)에 대한 오해가 그들의 배치에 대한 불만의 중요한 원인이라고 제시하였다.

본 논문에서는 일리노이대학 EPT를 예로 들어 수험자의 시험결과에 대한 오해

가 어떤 영작문 채점기준 요소들에 의해 발생하는지에 대한 규명을 시도하였다. 또한 수험자들에게 영작문 평가 채점기준표 요소들에 대한 교육을 시켰을 경우 어떠한 채점기준 요소들의 점수에 영향을 주며 그 것이 성적에 연결되는 가를 연구하였다.

II. 선행 연구 검토

1. 배치고사의 오배치(misplacement)에 대한 연구

ESL 영작문 평가에서 수험자들의 ‘오배치’에 대해 지금까지 거의 관심을 갖지 않았다. 따라서 선행연구가 미비한 가운데 Steward(1999)는 일리노이대학교의 EPT 수험자들의 오배치에 대해 연구하였다. 먼저 평가 전문가들은 배치고사 결과에 따른 배치의 어려움을 인식하고 ‘오배치’라는 용어를 정의할 필요가 있으며, 시험 타당도의 측면에서 배치가 잘되었는지 또는 잘못 되었는지를 구분하여야 한다(Kim, 2003). Steward(1999)는 오배치라는 용어를 “한 수험자가 배치고사의 결과로 배치된 ESL 교실에서 자신의 영작문 실력에 비해 수업의 수준이 너무 높거나 낮은 경우”라고 정의하였다. 하지만 용어의 모호함으로 인해 Steward는 “정확하게 배치된” 또는 “아마도 오배치된” 이라고 부연하였다. 그의 연구에 따르면 일리노이대학교에서 1998년도 전체 수험자의 약 20%가 그들의 ESL 배치에 대해 동의하지 않았으며, 약 10%의 수험자는 ESL 교실에서 평가고사를 본 결과 EPT 결과와 비교하여 실제로 오배치된 것으로 판명되었다. 수험자들은 대체로 자신들의 영작문 실력을 실제 점수보다 더 높게 평가하였고 이들의 심리적인 오배치는 ESL교실에 계속 출석함에 있어 동기 하락에 큰 영향을 준 것으로 밝혀졌다. Cho(2001)는 EPT 시험결과에 따른 ESL 수업 배치는 수험자가 ESL 수업 수만큼 전공수업을 수강할 수 없으므로 수험자들과 그들의 지도교수들에게 큰 부담을 지워주기 때문에 수험자들의 심리적 또는 실제적 오배치는 EPT 시험결과에 매우 중요한 요인이 되었다고 지적하였다.

배치고사의 오배치는 시험의 타당도의 일종으로써 설명되어야 한다(Bachman, 1990; Wall, Clapham, & Alderson, 1994). 좋은 배치고사를 개발하기 위해 Wall, Clapham과 Alderson(1994)는 다음의 네 가지 시험 타당도 연구 측면을 지적하였다: a) 배치고사의 정확도, b) 시험의 정확성에 대한 수험자의 만족 정도, c) 시험의 내용 적절도, 그리고 d) 시험 신뢰도. Haswell(1998)은 작문 배치고사의 어려움과 배치의 타당도 구현을 위해 오랜 시간이 필요함을 지적했다. 그는 “작문 배치는 언제

나 판단에 직면하며 시간이 지나도 절대로 학생의 적절한 배치에 대해 완전히 판단할 수 없다”(pp. 162-163) 라고 작문 배치고사 후 배치의 어려움을 토로하였다.

2. 분석적 작문 채점 요소들(Analytic Writing Features)

Hughes(1989)는 수험자들에게 하나의 숫자로 된 총점을 전달해 주는 총체적 채점 방식(holistic rating method)과는 달리 분석적 채점 방식(analytic rating method)이 많은 장점이 있다고 주장했다. 그 장점들은 첫째, 분석적 채점 방식이 수험자 개인의 부분적인 작문실력의 비균형적 발전 문제를 해결할 수 있으며, 둘째, 채점자들이 채점과정에서 수험자의 부분적인 작문 능력 중 불필요한 부분을 무시할 수 있고, 셋째, 채점자의 채점결과가 수험자에게 신뢰성을 줄 수 있다는 점이다.

Prochaska(2005)는 영작문평가에 있어서 수험자들의 동료 피드백이 공정할 수 있다고 보고하였다. 특히, Kim(2003)과 Lee(2005)의 연구에서 수험자들은 동료들 간의 피드백을 통해서 문법에 대해 많은 관심이 있음을 밝혔다. 또한 여러 학자들(e.g., Leki, 1990; Nelson & Murphy, 1998; Lee, 2005)은 초보 작문 학습자들은 문법이나 철자의 오류에 많은 관심이 있음을 지적했다. 하지만 Caulk(1994)와 Berg(1999)는 중급 또는 고급 작문 학습자들의 경우 구조 또는 내용에 대해서도 수험자들 상호간에 피드백을 줄 수 있다고 주장하였다.

Sakyi(2000)은 채점자가 수험자의 에세이를 채점할 때 영향을 끼치는 네가지 요인으로 a) 언어표현 사용의 오류, b) 논제 적합성과 저자의 사고 표현, c) 에세이에 대한 채점자 개인적 반응, d) 채점기준표를 들었으며 채점자에게 가장 큰 영향을 미치는 요소로써 내용의 적절성과 언어표현 사용의 오류를 지적하였다. Song과 Caruso(1996)은 내용과 구조 요소가 채점자의 채점에 큰 영향을 미친다고 주장하였으며, Milanovic, Saville과 Shuhong(1996)의 연구에 의하면 채점자들은 어휘와 내용만큼 효율적 의사소통의 여부(Communicative effectiveness)와 과업 구현(Task realization)을 중요시 여긴다고 보고하였다. Cho(1998)는 여러 가지 채점기준 요소들 중 표절에 대해 비영어권 수험자들의 이해가 부족함을 지적하였다. 또한 Lee(2005)는 채점자들 간의 상호 채점자 신뢰도를 일반화 가능성도 이론(G-Theory)과 Facets 분석을 통해서 표절의 신뢰도가 떨어짐을 발견하였다. 즉, 표절 요소가 영작문 채점의 한 요소로서 부적합함을 밝혀냈다. 그 이유는 표절이 채점자들 간의 서로 다른 인식에 기인하며 표절 요소가 포함된 측정평가 설계에 많은 변수가 발생하여 시험 신뢰도를 하락시켰기 때문이다.

3. 연구문제 제시

본 문헌연구를 통해 수험자들의 영작문 채점결과에 대한 오해는 과연 어떠한 채점기준 요소에 의해 발생하며, 그렇다면 수험자들에게 전문 채점자들이 사용하는 채점기준표에 대한 교육을 시켰을 경우 그들의 작문 채점결과에 어떠한 채점기준 요소가 영향을 미치는지를 살펴보는 것은 ESL 작문 교실에 적용할 수 있는 매우 흥미로운 연구가 될 것이다. 따라서 본 연구에서는 다음 두 질문을 제시한다.

- 1) ESL 학습자들의 영작문 채점 결과에 대한 오배치는 어떤 작문 채점기준 요소들을 유발하는가?
- 2) ESL 학습자들의 영작문 실력을 향상 시키는 영작문 채점기준 요소는 어떠한 것들인가?

III. 연구 방법

1. 실험 대상 및 집단

실험대상은 2001년도 여름 일리노이 대학교의 Computer-delivered ESL Placement Test(CEPT)를 치른 학부와 대학원에 입학한 외국인 학생 34명이었다. CEPT 시스템은 저자에 의해 2001년도에 개발되었으며 50분 안에 에세이를 써야하는 CEPT 시험과 약 5시간에 걸쳐 시험을 치루는 CEPT Workshop 시험, 그리고 컴퓨터를 이용한 채점시스템을 포함한다. 이 CEPT 시험은 일반 종이기반 EPT 시험과 함께 수험자가 하나의 시험 형태를 선택하도록 시행되었다. 학교 규정상 CBT TOEFL점수 253점(Paper-based TOEFL 610점)에 미달한 외국인 학생들은 의무적으로 말하기와 쓰기 두 가지 EPT 시험을 일리노이 대학교에 도착하여 보아야한다. 수험자들은 그 결과를 토대로 전공수업과 더불어 적절한 ESL 수업을 수강해야 한다. 대학교 입학 규정상 CBT TOEFL 213점(PBT TOEFL 550) 이상의 외국학생들이 입학하므로 실험집단의 TOEFL 성적은 213점에서 253점 사이의 학생들이다. 수험자들은 50분간 에세이를 쓰는 EPT를 보았으며 시험 후 설문조사에 응했다.

동일한 시기에 위와 같은 실험군에서 12명의 수험자가 두번째 시험인 CEPT Workshop에 참가하였다. 이 시험을 본 수험자들에게는 CEPT 결과와 비교하여 더 나은 점수를 공식 EPT 점수로 인정하는 특혜를 주었으며 이 점수는 수험자들의 ESL 수업 배치에 영향을 줄 수 있으므로 매우 중요하였다. 본 실험에서는 위의 12

명의 수험자들에게 초점을 맞추었으며 그들의 첫 번째 작문시험 결과, 자기평가 결과, 그리고 두 번째 작문시험 결과가 데이터 분석 대상이 되었고 그들에 대한 설문조사 결과와 인터뷰 내용이 분석되었다. 참여자중 10명은 대학원 입학생이었고 2명은 학부 입학생이었다. 그들의 모국어는 6명이 중국어 사용자, 4명이 한국어 사용자, 1명의 스페인어, 그리고 마지막으로 일본어 사용자 1명이었다.

채점은 4명의 일리노이대학교 공인 EPT 채점자들에 의해 컴퓨터로 이루어졌으며 그들은 다년 간 대학교 채점자 교육과 채점에 대한 경험을 쌓아왔다. 채점자들은 저자에 의해 약간의 수정을 거친 학교 공식 채점기준표를 바탕으로 채점을 하였다. 채점기준표에 대한 저자의 수정은 매우 미미한 것으로 채점자들이 보기에 편리하게 채점기준들을 재배치한 것이었다.

EPT 영작문 시험은 총점이 4점으로써 4점인 경우 ESL 수업이 면제되고 3점 이하부터 각 수준에 맞는 ESL 수업을 수강하여야한다. 이때 대학원 수험자의 전공과목을 ESL과목 숫자만큼 수강할 수 없기 때문에 이 시험은 매우 중요하게 인식되어진다.

2. 자료 수집 과정

여기에서는 CEPT, CEPT Workshop 시행과정, 채점 과정, 설문조사, 인터뷰 과정을 설명하였다.

1) CEPT와 채점 절차

이 시험은 하나의 컴퓨터 랩에서 시행되었으며 수험자들은 수험장에 도착하여 시험 감독자의 시험절차와 주의사항에 대한 짧은 안내를 받았다. 그리고 그들은 컴퓨터에 접속해서 10분 길이의 미국인 교수의 동영상 강의를 시청하고 작문 주제에 대한 읽기 자료 하나를 50분 동안 읽으면서 에세이를 작성하였다. 수험자의 에세이는 저자가 개발한 Active Server Page를 이용한 시험 시스템 상에 작성되어서 컴퓨터 랩 내부 네트워크를 이용하여 서버의 마이크로소프트 Access 데이터 베이스 프로그램에 저장되어졌다.

채점 결과는 학교 규정상 24시간 이내에 결과가 통보되어야 하므로 시험이 끝난 직후 4명의 채점자들은 컴퓨터에 접속하여 수험자들의 에세이를 불러내어 컴퓨터 상에서 채점을 시행하였다. 2명의 채점자가 수험자 한명의 에세이를 읽고 채점하였다. 채점 결과는 저자가 수정한 채점기준표에 의거 부분별 점수들(analytic scores) 과 총점, 그리고 수험자의 영작문의 대표적인 약점에 대한 피드백으로 구성되어졌

다. 수험자들에게 영작문 시험 중 철저히 따르도록 명시된 채점기준 요소들은 a) 문법(grammar), b) 구성(organization), c) 소스 사용(use of sources), d) 내용(content), 그리고 e) 표절(plagiarism)로 구성되어졌으며 동일한 배점으로 채점이 되었다. 모든 채점 결과는 컴퓨터에 입력하여 서버에 저장되었다.

2) CEPT Workshop과 자기평가

CEPT를 치룬 수험자 중에서 자신의 시험 점수를 향상시키기를 원하는 수험자 12명이 Workshop 시험에 참가하였다. 이 시험은 5시간에 걸쳐서 다음과 같은 과정으로 시행되었다.

(1) 자기평가

수험자는 먼저 컴퓨터에서 자신의 CEPT 에세이를 불러내어 주어진 채점기준표에 의거 자기평가를 실시하였다. 수험자는 전문채점자가 사용한 동일한 채점기준표를 사용하여 자신의 에세이에 대해 부분점수와 총점, 그리고 대표적인 약점에 대한 자기 피드백을 컴퓨터 상의 채점지에 입력하였다.

(2) 전문 채점자의 채점 결과 열람

수험자들은 컴퓨터 상에서 전문 채점자들의 채점결과를 열람하고 자신의 채점결과와 비교 분석하였다. 두명의 전문 채점가의 채점 결과들과 수험자의 자기평가 채점결과는 컴퓨터 상에 세로로 배치되어 수험자의 비교 분석을 도왔다.

(3) 채점기준표 교육

위의 수험자의 자기평가와 분석 시간이 끝난 후 전문 채점자의 채점기준표 교육을 1시간 동안 시행하였다. 강사는 미국식 영작문의 특성과 각 채점기준 요소들에 대해 설명하고 수험자들과 질의응답 시간을 가졌다.

(4) 에세이 작성과 제출

위의 과정이 끝난 후 수험자는 주어진 작문 주제에 대해 에세이를 작성하였다. 작문 주제는 CEPT와 달랐으나 동등하다고 판명된 주제였다(Liu, 1997). 주어진 시간은 50분이었으며 동영상 보기와 읽기자료 제공 등이 CEPT와 동일하였다. 수험자의 에세이는 컴퓨터 서버에 저장되었다.

(5) 에세이 채점

채점과정은 위의 CEPT와 동일하게 진행이 되었으며 컴퓨터 상에서 진행되었다. 수험자는 24시간 이후 컴퓨터 서버에 접속하여 주어진 아이디와 암호를 사용하여 채점결과를 열람하였다.

3) 설문조사와 인터뷰

수험자들이 CEPT와 CEPT Workshop을 치룬 후 수험자들에게 설문조사를 시행하여 데이터를 수집하였다. CEPT Workshop에 참가한 수험자들에게는 설문조사 직후에 저자가 일대일 인터뷰를 실시하였으며 모든 인터뷰는 오디오 테이프에 녹음되었다. 한국인은 한국어로 인터뷰를 하였고 저자가 영어로 번역하였다. 인터뷰는 수험자의 의견을 끌어내기 위해 자연스러운 대화 형식으로 이루어졌다. 주요 인터뷰 내용은 CEPT 채점 결과에 대한 수험자의 반응과 자기평가와 전문 채점자의 피드백과 이후 수험자의 채점결과에 대한 의견 변화에 대한 것이었다. 또한 채점기준표 교육 이후 달라진 수험자의 의견 변화도 포함되었다.

3. 자료 분석

1) 통계 분석

모든 시험 결과와 설문 결과는 마이크로소프트 엑셀 프로그램에 내장된 통계 기능을 이용하여 분석하였다. 각 결과에 따른 평균과 표준편차를 제시하였으나 샘플 크기가 작아서 가설검증에 따른 통계 분석은 시행하지 않았다.

2) 질적 분석

인터뷰 데이터는 오디오 테이프로 녹화 되었으며 연구자에 의해 분석되었다. 몇몇 중요한 응답은 통계 결과를 뒷받침하기 위해 문자화되어졌다.

IV. 결과 및 논의

1. ESL 학습자들의 영작문 채점 결과에 대한 오배치는 어떤 작문 평가 요소들을 유발하는가?

1) 수험자의 CEPT 채점결과에 대한 동의 정도

표 1은 CEPT 시험결과에 대한 수험자들의 반응을 설문조사 결과로 나타내고 있다. 채점결과와 수험자들의 심리적 자기평가의 일치 정도는 수험자들의 ESL 배치에 대한 불만을 나타내는 정도를 말해준다. 설문조사 결과 수험자의 58.4%(전체 12 응답 중 7)은 채점결과에 대해 동의하지 않고 있음을 보여주며, 41.6%(전체 12 응답 중 5)의 수험자가 채점결과에 수긍하고 있음을 보여준다. 전체 참가자 수와 응답자 수가 다른 것은 많은 참가자들이 응답을 하지 않은 경우와 자료 분석 과정에서 적절하지 않은 응답들을 제외하였기 때문이다.

표 1
수험자의 CEPT 채점 결과에 대한 동의 정도

매우 불일치	불일치	잘모름	일치	매우 일치	계
5	2	0	4	1	12
41.7%	16.7%	0%	33.3%	8.3%	100%

표 2는 CEPT Workshop 중에 실시한 자기평가 후 수험자들의 배치 불일치 정도가 33.3%(전체 12 응답 중 4)로 감소하였으며 50%(전체 12 응답 중 6)은 채점결과에 동의함을 보여준다.

표 2
자기평가 이후 수험자의 CEPT 채점 결과에 대한 동의 정도

매우 불일치	불일치	잘모름	일치	매우 일치	계
3	1	2	5	1	12
25.0%	8.3%	16.7%	41.7%	8.3%	100%

표 3은 CEPT Workshop 과정에서 수험자에 대한 채점기준표 교육 이후 수험자들의 채점결과에 대한 동의 정도를 보여준다. 수험자의 75%(전체 12 응답 중 9)은 채점결과에 수긍하였으며 단지 16.7%(전체 12 응답 중 2)만이 불일치로 남아 있음

을 보여준다.

표 3
채점기준표 교육 이후 수험자들의 채점결과에 대한 동의 정도

매우 불일치	불일치	잘모름	일치	매우 일치	계
2	0	1	5	4	12
16.7%	0%	8.3%	41.7%	33.3%	100%

결론적으로 수험자들의 CEPT 채점결과에 대한 불일치 정도는 최초에 54.8%(전체 12 응답 중 7)에서 자기평가 이후 33.3%(전체 12 응답 중 4)로 줄었으며 채점기준표 교육 이후 다시 16.7%(전체 12 응답 중 2)으로 감소되었음을 보여주었다. 즉, 수험자의 채점기준표에 대한 이해가 높아질 수록 전문 채점가의 채점결과에 대한 동의 정도가 높아지는 것을 보여준다.

2) 수험자의 채점기준표 요소에 대한 자기평가 결과

두번째 설문조사를 통해 약 91.6%(전체 12 응답 중 11)의 수험자들은 자기평가와 채점기준표 교육의 유용성에 대해서 긍정적인 대답을 하였으며, 전문 채점자들의 피드백이 자신들의 영작문의 약점을 확인하는데 큰 도움이 되었다고 대답하였다. 그러면 좀더 자세하게 각 채점기준표 요소 별로 수험자들의 설문조사 결과를 다루어 보자.

먼저 표 4는 각각 23.5%(전체 34 응답 중 8)의 CEPT 수험자들이 채점기준표 중에서 구조와 내용에 대해서 쓰기 교육이 필요함을 피력했다. 또한 20.6%(전체 34 응답 중 7)의 수험자들은 소스 사용에 대하여 17.6%(전체 34 응답 중 6)의 수험자들은 표현 선택에 대한 쓰기 교육의 필요성을 주장하였다. 하지만 표절과 구두법에 대해서 각각 2.9%(전체 34 응답 중 1)의 수험자들이 교육 필요성을 표시하였다. 구두법은 CEPT에 첨가되었다가 Workshop에서 삭제되었다.

표 4
CEPT 수험자들의 채점기준표 요소들에 대한 훈련 필요성

구조	내용	문법	표현 선택	소스 사용	표절	구두법	계
8	8	3	6	7	1	1	34
23.5%	23.5%	8.8%	17.6%	20.6%	2.9%	2.9%	100%

표 5는 CEPT 수험자들의 채점기준표 요소에 대한 전문가의 점수에 대한 동의 정도를 보여준다. 구조에 대해서 41.2%(전체 17 응답 중 7)의 수험자가, 내용에 대해서 35.3%(전체 17 응답 중 6)의 수험자가 동의를 표시하였다. 소스사용과 표절에 대해서는 각각 11.8%(전체 17 응답 중 2)의 수험자가 동의하였으나 문법, 표현 선택, 구두법에는 한명도 동의를 표시하지 않았다.

표 5
CEPT 수험자들의 채점결과에 대한 동의 정도

구조	내용	문법	표현 선택	소스 사용	표절	구두법	계
7	6	0	0	2	2	0	17
41.2%	35.3%	0%	0%	11.8%	11.8%	0%	100%

표 6은 자기평가 후 수험자들이 채점기준표 요소들에 대한 채점 어려움의 정도를 설문조사를 통해 밝힌 것이다. 가장 두드러진 것은 수험자들의 50%(전체 16 응답 중 8)가 구조에 대한 채점이 어렵다고 표시한 것이다. 반대로 문법, 표절, 구두법에 대해서는 한명도 어려움을 토로하지 않았다. 즉, 표 5에서 41.2%가 구조 요소에 대한 채점결과에 동의한 것과 표 6에서 50%가 채점의 어려움을 표시한 것은 수험자들의 구조 요소에 대한 이해 부족으로 볼 수 있다.

표 6
수험자들의 채점기준표 요소에 의한 채점의 어려움 정도

구조	내용	문법	표현 선택	소스 사용	표절	구두법	계
8	3	0	2	3	0	0	16
50.0%	18.8%	0%	12.5%	18.8%	0%	0%	100%

표 7은 수험자들이 자기평가와 채점기준표 교육 후 CEPT Workshop에서 작성한 에세이의 점수에 대해 어떠한 요소의 점수가 총점 향상에 큰 영향을 미칠 것인가라는 질문에 대답한 것이다. 수험자들 중 35%(전체 20 응답 중 7)가 구조 요소를 총점 향상에 긍정적인 요인으로 지목했고, 25.5%(전체 20 응답 중 5)는 내용 요소를, 20.0%(전체 20 응답 중 4)는 표현선택을 지목하였다. 하지만 표절과 구두법 요소에 대해서는 한명도 응답하지 않았다.

표 7
CEPT Workshop 수험자들의 채점기준표 요소 중 점수 향상 예상

구조	내용	문법	표현 선택	소스 사용	표절	구두법	계
7	5	1	4	3	0	0	20
35.0%	25.5%	5.0%	20.0%	15.5%	0%	0%	100%

3) 통계 분석

CEPT Workshop 후 83.4%(전체 12 응답자 중 10명)은 자신들의 점수가 CEPT 결과보다 더 향상되었을 것으로 예측하였으며 그 원인은 채점기준표에 대한 높아진 이해에 있다고 응답하였다. 더 자세히 설명하면 수험자 중 35.0%(전체 20 응답 중 7)는 구조 요소에 대한 점수 향상을 기대하였으며, 25.0%(전체 20 응답 중 5)는 내용 요소에 대해, 20.0%(전체 20 응답 중 4)는 표현선택에 대해서 향상된 점수를 기대하였다.

표 8은 수험자의 CEPT, 자기평가, CEPT Workshop 결과를 평균, 표준편차를 이용하여 비교하였다. 단, 여기서 참가자 12명 중 3명이 자기평가(Self)를 하지 않았기 때문에 결과적으로 9명의 결과만 수록하였다. 수험자들의 점수는 자기평가의 평균이 CEPT보다 0.11점 높아졌으며 표준편차도 비슷한 결과를 보였다. 하지만 이 점수는 가설검증을 하지 않았으므로 유의미성을 평가하기 어렵다.

표 8
수험자의 CEPT, 자기평가, CEPT Workshop 채점 결과 비교

	Mean	Mean Increment	Std. Deviation	N
EPT	2.67	0.00	0.50	9
Self	2.78	+ 0.11	0.67	9
CEPT Workshop	3.00	+ 0.22	0.00	9

표 9는 수험자의 CEPT 점수와 자기평가 점수 간의 분할표(contingency table)이다. 수험자의 88.8% (전체 9명 중 8명)은 동일한 점수를 보였고 단지 11.1%(전체 9명 중 1명)은 다른 점수를 보였다. 이 한명의 수험자의 경우가 실제 점수와 자신의 심리적 점수 결과가 맞지 않는 오배치 현상을 보여준다.

표 9
수험자의 CEPT 채점결과와 자기평가 결과 분할표(Contingency Table)

Self \ CEPT	1	2	3	4
1	0	0	0	0
2	0	3(33.3%)	0	0
3	0	0	5(55.5%)	0
4	0	0	1(11.1%)	0

표 10은 수험자들의 CEPT 점수와 자기평가 점수를 채점기준표의 각 요소별로 평균과 그 차이로 비교하였다. 먼저 구조 요소에 대해서 수험자들의 자기평가 점수가 CEPT 구조 요소 점수보다 0.39점 높게 나온 것을 볼 수 있다. 수험자들은 문법과 내용 요소에 대해서 자기평가를 0.11점과 0.06점 높게 하였으나 소스사용과 표절에 대해서는 자기평가에서 0.11점과 0.34점 낮게 채점하였다.

표 10
Score Improvement for Each Feature Between CEPT and CEPT Workshop

SELF - CEPT	Mean	Difference
SO - EO	2.89 - 2.50	0.39
SG - EG	2.67 - 2.56	0.11
SC - EC	2.78 - 2.72	0.06
SU - EU	2.78 - 2.89	- 0.11
SP - EP	3.22 - 3.56	- 0.34

위의 측정평가적 설명에 덧붙여서 수험자들과의 인터뷰 내용은 영작문 채점기준표 요소에 대한 수험자들의 오해에 대한 이해를 돕는 증거이다. 수험자의 한 사람인 John(pseudonym)은 다음과 같이 설명하였다.

In my opinion, organization was most difficult when I was rating my essay. I was a little angry, because my score was not good. That was 3, but I thought I got 4. I want exempt. The raters gave 3 on organization so my score was not 4. But I could learn about organization during the scoring rubric lesson, and so I just follow the structure of essay so that

the raters want.

John과의 인터뷰에서 알 수 있듯이 수험자들의 채점기준표 요소에 대한 오해는 구조 요소에 대한 오해에서 올 수 있음을 보여준다. 또 다른 수험자 Susan(pseudonym임)과의 인터뷰는 작문에 있어서 구조요소 중 하나인 문장 배치에 대한 문화적인 차이를 보여준다. 비록 문법의 오류가 있더라도 그대로 인용하였다.

I am not familiar with North American writing style because I came from an Asian country. In my country, we just write topic sentence at the end of a paragraph and put examples at first. But raters gave me feedback like I should move the topic sentence to the first. So I became agreed with my score.

결론적으로 수험자들이 자기평가에서 채점기준표를 바탕으로 자신의 에세이를 채점하였을 때 수험자들의 자기평가 점수가 CEPT 채점결과보다 0.11점 높게 나왔다. 이결과는 CEPT 설문조사 결과와 마찬가지로 수험자들이 자신의 작문 실력을 과대평가한 결과였으며 채점기준표의 요소 중 구조 요소에 대한 오해에서 비롯된 것이었음을 보여준다. 수험자들의 자기평가 점수 중 구조에 대한 점수가 전문가들이 채점결과 보다 0.39점 높았던 것이 그 증거이다.

하지만 수험자들은 자기평가와 채점기준표 교육을 거친 후 전문 채점자들의 채점결과에 더욱 동의하게 되었다. 최초에 58.4%(전체 12명 중 7명)의 수험자가 채점결과에 동의하지 않았으며 이결과는 수험자의 심리적인 오배치로 인식되어졌다. 그리고 이러한 심리적 오배치는 실제 ESL 교실 안에서의 수업 참가에 대한 동기부여가 결여되는 원인이 되어왔다(Steward, 1999). 자기평가 후 수험자들의 채점결과에 대한 불만족은 33.3%(전체 12명 중 4명)로 감소하였으며 초기에 채점결과에 동의하지 않던 수험자 7명 중 3명이 결국 전문 채점자들의 채점결과에 동의하게 되었다. 또한 두번째 시험 후 설문조사의 결과는 수험자들의 91.6%(전체 12명 중 11명)가 자기평가와 채점기준표 교육, 전문가의 피드백으로 인해서 자신의 영작문의 약점을 알게 되었다고 보고하였다.

특히 수험자들의 채점기준표 요소 중 구조와 내용에 대한 설익은 지식이 영작문에 약점으로 작용되었음이 발견되었다. 이 두 요소들은 수험자들의 심리적 오배치를 야기하였으며 자기평가 점수가 CEPT 채점 결과보다 0.11점 높게 나온 원인이 되었다.

2. ESL 학습자들의 영작문 실력을 향상 시키는 영작문 채점기준 요소는 어떠한 것들인가?

표 11은 수험자들의 CEPT 점수와 Workshop 점수의 채점기준표 요소들의 평균을 비교한 결과이다. 먼저 표 8 에서 볼 수 있듯이 Workshop 총점은 CEPT 결과보다 0.33점 향상되었다. 이 점수는 ESL 배치에 영향을 줄 수 있는 점수는 아니다. 왜냐하면 채점과정에서 반올림 정책이 사용되기 때문이다. 0.5점에 미달되는 점수는 버림으로 처리되어 점수 향상에 영향을 주지 않는다. 하지만 채점기준표 각 요소별 채점결과는 구조에 대한 평균 상승 폭이 0.83점으로써 수험자의 ESL 배치에 영향을 줄 수 있음을 보여준다. ESL수업 면제 점수인 4점에 근접하는 실력의 수험자들에게 이 구조 요소의 점수는 매우 중요하다. 반면에 다른 채점기준표 요소들인 문법, 내용, 소스사용, 표절은 ESL 배치에 영향을 주는 수준의 점수 향상을 가져오지 못함을 보여준다.

표 11
수험자의 CEPT와 CEPT Workshop 채점 결과 비교

WEPT - CEPT	Mean	Difference
WO - EO	3.33 - 2.50	0.83
WG - EG	2.78 - 2.56	0.22
WC - EC	3.00 - 2.72	0.18
WU - EU	3.00 - 2.89	0.11
WP - EP	3.67 - 3.56	0.11

그리고 표 11에 대한 분할계수(contingency coefficient)들은 표절 요소에 대해 0.46의 낮은 계수를 얻었으며, 나머지 소스사용, 문법, 내용, 구조 요소들에 대해서는 0.70 이상의 계수를 얻었다. 이 계수는 수험자들의 CEPT와 Workshop 결과에 대한 채점자들 간의 상호 일치 정도를 보여주는 것으로써 Cho(1999)와 Lee(2005)의 제안대로 표절 요소는 채점기준표에서 채점자 신뢰도가 극히 낮음을 보여준다.

결론적으로 수험자에 대한 채점기준표 사용과 교육은 전문채점가의 피드백과 더불어 CEPT Workshop 점수의 향상을 가져왔으며 구조 요소가 가장 큰 변인이었음을 알 수 있다. 수험자들의 점수는 CEPT 점수보다 0.33점이 향상되었으며 채점기준표 요소들 중 구조 요소의 점수가 0.83점 향상되어 총점의 향상에 큰 영향을 주었음을 보여준다. 다른 요소들의 점수도 향상되었으나 ESL 배치에 영향을 줄 정도

는 아니었다. 결론적으로, 수험자들에게 짧은 시간의 채점기준표 교육을 시켰을 경우 구조 요소에서 점수가 쉽게 향상되어 총점의 향상을 이끈다는 것을 보여주었다.

V. 결론 및 제언

본 연구에서는 영작문 배치 평가 수험자들이 전문 채점자들의 채점결과에 대해 심리적으로 동의하지 않는 것을 ESL 오배치로 정의하였다. 그리고 수험자들이 채점결과에 동의하지 않는 정도를 자기평가를 통해 수치화 하여 전문 채점자들의 점수와 비교하여 오배치의 정도를 측정할 수 있었다. 수험자들은 자기평가에서 자신의 에세이에 대해 0.11점 높은 과장된 채점을 하였으며 그 원인이 구조 요소에 대해 0.39점 높게 과장 평가한 것 때문임이 밝혀졌다. 즉, 수험자들의 구조 요소에 대한 미숙한 이해가 원인이었다. 즉, CEPT 수험자들의 23.5%(전체 34 응답 중 8)가 구조 요소에 대한 교육 필요성을 토로하였고, 50%(전체 16 응답 중 8)가 구조 요소에 대한 자기평가 채점의 어려움을 토로하였다. 또한 채점기준표 교육 후 수험자들의 35%(전체 20 응답 중 7)가 CEPT Workshop의 자신의 에세이에서 구조 요소에 대한 점수의 향상을 기대하였다. 그 결과 수험자들의 총점은 CEPT 점수보다 CEPT Workshop 점수가 0.33점 향상되었는데 그중 구조 요소에 대한 점수는 0.83점 향상 되어 총점 향상에 지대한 영향을 끼친 것으로 드러났다. 다른 내용, 문법, 소스사용, 표절요소들의 점수 향상 폭은 개별적으로 총점에 영향을 끼치지 못하는 것으로 나타났다.

본 연구의 결과를 바탕으로 한국 영어교육 작문교육 현장에 다음과 같은 몇가지 시사점을 제시하고자 한다. 첫째, 영작문 학습자들에게 대상 영작문 시험의 채점기준표에 대한 구체적이고 조직화된 사전교육이 필요하다. 이러한 채점기준표들은 대체로 온라인 상에 공개되어 있다. 특히, 영미권에 유학을 목표로 하는 학습자들에게 인터넷에 제공되는 TOEFL의 영작문 시험인 TWE 채점기준표에 대한 철저한 사전교육이 필요하다. 또한 영미권 각 대학교의 배치교사의 채점기준표들도 온라인으로 제공되므로 자신이 진학하는 대학교의 학교의 채점기준표에 대한 사전교육을 할 것을 권한다.

둘째, 영어 쓰기 교실 안에서 전문가가 독점하던 채점기준표를 공개하여 학습자들과 공유하여 자기평가의 기회를 더욱 제공할 필요가 있다. 과거 채점기준표는 전문가용과 수험자용이 구분되어 있었다. Kim(2003)과 Lee(2006)은 전문 채점자들의 채점기준표를 수험자들에게 제공하여 자기평가를 시행하였다. 그리하여 수험자들의 채점기준에 대한 오해를 풀면서 전문 채점자들의 채점결과를 더욱 잘 이해하고 자

신의 작문 실력을 향상시키는데 도움이 됨을 보여주었다. 국내 영어 쓰기 교실에서 도 선생님의 채점기준표를 공개하고 학습자들에게 자기평가의 기회를 많이 줄 것을 권장한다.

셋째, 영작문 교실 안에서 학습자들에게 자기평가와 더불어 전문 채점자들의 세부적인 피드백을 활성화할 필요가 있다. 본 연구에서 수험자들은 총점만을 제공 받을 때보다 전문 채점자들의 세부 채점기준표 요소별 점수와 자신들의 약점에 대한 피드백을 제공 받을 때 수험자 자신들의 영작문 약점에 대한 이해도가 높아졌다. 국내 대다수 영작문 교실에서는 교사 1인당 많은 학습자들의 수로 인해 한계가 있을지라도 전문 채점자의 채점기준표 요소별 점수와 상세한 피드백은 학습자들이 자신의 작문 실력의 약점을 파악하는데 큰 도움이 될 것이다.

넷째, 짧은 시간에 영작문 실력을 향상시키려는 의도를 가진 학습자들에게는 영작문 채점기준표에 의거 구조 요소에 대한 집중적인 교육을 권장한다. Kim(2003)과 Weigle(2002)은 작문요소 중 적절한 표현사용 또는 충실한 내용 등의 요인들은 학습자들에게 장시간의 꾸준한 학습이 필요하다고 주장하였다. 하지만 본 연구에서 보여 주었듯이 구조 요소에 대한 짧은 시간 내의 교육이 수험자들의 총점 향상에 큰 영향을 미쳤다. 특히, 유학을 위해 TWE를 준비하거나, 작문 시험이 추가되는 TOEIC을 준비하는 수험자들에게 영작문 채점기준 중 구조 요소에 대한 교육은 많은 도움이 될 것이다. 구체적으로 이 구조 요소에 대한 교육은 수험자의 모국어 수사법(rhetoric)과 북미 대학교 학술 영작문 시스템의 수사법의 차이를 극복하는 데 초점을 맞추어야 할 것이다.

본 연구는 몇 가지 측면에서 한계를 지니고 있다. 첫째, 현실적인 실험 여건의 제약으로 샘플 크기가 매우 작았다. 하지만 보통 북미 대학교의 ESL 교실의 학생 수가 15명 이내인 점을 고려한다면 12명의 수험자 수도 받아들여질 만하나 학습자의 수가 20명을 넘어서는 국내 여건에 비추어 본 연구는 좀 더 많은 샘플 수로 재연될 필요가 있다. 둘째, 본 연구의 수험자들의 모국어와 국적, 문화 등이 동일하지 않기 때문에 동일한 모국어와 문화를 가진 국내의 ESL 교실에 활용하기에 제약이 있을 수 있다.

앞으로의 연구에서는 본 연구의 제약들을 보완하여 국내의 ESL 교실 상황에 맞는 영작문 학습자들의 작문요소에 대한 오해를 연구할 필요성이 대두된다. 또한 학습자들의 영작문 요소에 대한 오해를 어떻게 효과적인 영작문 교육에 연결시킬 것인가도 심도 있는 연구가 요구된다. 마지막으로 많은 영어 작문 학습자들을 대상으로 영작문 요소에 대한 정신측정학적(psychometric) 측면에서의 양적 또는 질적 평가적 접근이 필요하다. 예를 들어, 양적 측정 평가적 측면에서 대량의 샘플 수를 기반으로 다양한 양적통계 방법을 이용하여 아직 밝혀지지 않은 채점요소들을 밝혀내

는 것도 좋은 방법일 것이다.

참 고 문 헌

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Caulk, N. (1994). Comparing teacher and student responses to written work. *TESOL Quarterly*, 28(1), 181-188.
- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215-241.
- Cho, Y. (1998). *Examining the validity of a multiple-choice plagiarism test*. Unpublished master thesis. University of Illinois at Urbana-Champaign.
- Cho, Y. (2001). *Examining a process-oriented writing assessment in a large-scale ESL testing context*. Unpublished doctoral dissertation. University of Illinois-at Urbana-Champaign.
- Douglas, D. (2000). *Language testing for specific purpose*. Cambridge: Cambridge University Press.
- Haswell, R. H. (1998). Searching for Kiyoko: Better mandatory ESL writing placement. *Journal of Second Language Writing*, 7(2), 133-174.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Kim, J. T. (2003). *Computer-delivered ESL placement test at the University of Illinois at Urbana-Champaign*. MA thesis. University of Illinois at Urbana-Champaign.
- Lee, H. (2005). Measuring plagiarism in the second language essay-writing context. *English Language Literature & Teaching*, 12(1), 221-228.
- Leki, I. (1990). Potential problems with peer responding in ESL classes. *The CATESOL Journal*, 3(1), 5-19.
- Liu, H. (1997). *Constructing and validating parallel forms of performance-based writing prompts in an academic setting*. Unpublished doctoral

dissertation, University of Illinois at Urbana-Champaign.

- Milanovic, M. S., Saville, N., & Shuhong, S. (1996). A study of decision-making behavior of composition markers. In M. Milanovic & N. Saville (Eds.), *Studies in language testing 3 Performance testing, cognition and assessment* (pp. 92-114). Cambridge: Cambridge University Press.
- Nelson, G. L., & Murphy, J. M. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27(1), 135-142.
- Prochaska, E. (2005). Korean EFL university students' evaluation of peer review interactions: A social model for evaluating the writing process. *English Language & Literature Teaching*, 11(2), 51-66.
- Sakya, A. A. (2000). Validation of holistic scoring for ESL writing assessment: How raters evaluate compositions. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment* (pp. 129-152). UK: Cambridge University Press.
- Song, B., & Caruso, I. (1996). Do English and ESL faculty differ in evaluating the essays of native English-speaking and ESL students? *Journal of Second Language Writing*, 5(2), 163-182.
- Steward, S. (1999). *An investigation of writing features which have predictive evidence of validity for EPT Placement Levels*. MA Thesis. University of Illinois at Urbana-Champaign.
- Wall, D., Clapham, C., & Alderson, J. C. (1994). Evaluating a placement test. *Language Testing*, 11(3), 321-344.
- Weigle, S. C. (1994). Effects of training on raters of ESL compositions. *Language Testing*, 11(2), 197-223.

예시 언어(Examples in): English

적용가능 언어(Applicable Languages): English

적용가능 수준(Applicable Levels): Elementary/Secondary/College/Higher

김정태

University of Illinois at Urbana-Champaign

411-370 경기도 고양시 일산서구 주엽동 135번지

문촌마을 부영아파트 1506동 205호

Tel: (031) 925-6104

Cell Phone: 010-7239-6104

Email: kim.jungtae@gmail.com

Received in July. 2006

Reviewed by Aug. 2006

Revised version received in Sept. 2006