

영어교육평가자의 직업윤리 확보를 위한

기독교적 성찰

신 동 일 (중앙대학교)

I. 들어가는 글

영어시험의 영향력이 너무 커져가고 있다. 속도와 성취로 성장한 한국의 교육현장에 사람들이 모여들면서 검증되지 않은 공부법과 경쟁에서 승리할 수 있다는 처세술이 넘치고 있다. 세계화 정책이 본격적으로 가동되면서 영어공부와 영어시험에 대한 압박이 이제 개인학습자 수준이 아니라 가정과 사회의 재정과 미래를 짓누르고 있는 정도이다. 가르치고 시험을 관리하는 사람이나, 배우고 시험을 준비하는 사람 모두 선의와 정직이 부족하고 물질주의, 성공주의로 인해 부정직은 당연하고 일상적인 교육관행이 되고 있다. 일선 교사가 특정 학생의 답안지를 대리 작성하거나 학원에 시험내용을 유출하는 사건이 일어나고, 수험자는 아예 시험성적을 조작하는 뉴스가 등장해도 크게 놀라지 않는다. 시험은 학성취도를 이끌어내기 위한 건강한 부담감이 되지 못하고, 절제가 없는 경쟁심과 이기심을 부추기고 있다.

무모한 시험의 역사는 성경에도 등장한다. BC 1100년경, 사사기에 등장하는 길르앗 사람들이 에브라임 사람을 색출하기 위해 쉬볼렛(shibboleth: 시냇물) 발음을 요청하고, 시볼렛(sibboleth: 무거운 짐)으로 발음하는 사람들을 모두 죽여버린 기록이 나온다.

그리고 나서 입다는 모든 길르앗 사람들을 불러 모으고 에브라임과 싸웠습니다. 에브라임 사람들이 “너희 길르앗 사람은 에브라임에서 도망친 사람들로써 에브라임과 므낫세 사람 가운데 있는 사람들이다”라고 말했기 때문에 길르앗 사람들이 에브라임을 무찔러버렸습니다. 길르앗 사람들은 에브라임으로 가는 요단강 나루터를 차지했습니다. 에브라임에서 살아남은 사람이 “내가 강을 건너가겠소”라고 말하면 길르앗 사람은 그에게 “네가 에브라임 사람이나?”라고 물어보았습니다. 만약 그가 “아니다”라고 대답하면 그들은 “쉽볼렛”이라고 말해 보라고 했고 그가 “시볼렛”이라고 정확하게 발음하지 못하면 그들은 그를 잡아다가 요단강 나루터에서 죽였습니다. 그때 에브라임 사람은 4만2,000명이나 죽였습니다. (사사기 12:4-6)

사사기 시대처럼 발음 평가로 사람의 생명을 뺏는 수준은 아니더라도 한국의 많은 학생들과 가정이 영어시험으로 진학과 취업에 인생을 걸고 있다. 기독교교육자들은 거짓은 마귀의 것이라고 선포하며 정직 운동(커닝추방운동, 무감독시험)을 시범적으로 시행하기도 했지만 여전히 부정직한 교육평가의 관행을 총체적으로 바꾸진 못하고 있다.

시험의 압박이 커져가면서 시험을 공정하고 타당하게 개발하거나 시행할 수 있는 전문 인력이 주목받고 있다. 그러나 시험 전문가의 기능적 교육과 양적 성장이 진행된 만큼이나 직업윤리교육이 제대로 이뤄지지 않은 터라, 고부담 입시현장과 교실 평가에서 부적절한 관행이 끊임없이 고발되고 있다. 경제의 규모가 성장하는 것과 상관없이 평가자의 전문성, 심지어 대학의 입학사정관 업무마저 그 공신력을 의심 받고 있는 형편이다.

선다형 문항을 이용해서 영어지식을 확일적으로 평가하는 시점에서 의사소통능력, 발화능력, 수학능력, 심지어 잠재력이나 소질을 평가해야 하는 업무가 학교나 기업현장에서 빈번해지고 있지만 대부분 검증되지 않은 원어민감사, 평가자 교육을 전혀 받지 않은 교원, 또한 국가인증 혹은 국제공인이란 모호한 이름으로 포장된 정체불명의 평가기관에 의해 무사안일하게 처리되고 있다. 예를 들어 최근 기업과 학교에서 영어말하기능력, 영어를 통한 문제해결 혹은 수행능력에 큰 관심을 갖고 있지만 문항이나 지침서 제작과정, 채점방법, 채점자교육방법에 관한 규범화된 문서자료는 찾아보기 힘들며 대부분 즉흥적인 구두회의로 시험을 개발하고 관리하는 상황이 여전히 빈번하다.

역사적으로 시험이란 제도는 권력을 만들고 유지시키기 위한 도구로 자주 사용되어 왔는데(Spolsky, 1997), 한국 사회에서도 여러 영어시험들이 그 자체로 하나의 권력으로 성장하고 있다. 시험을 만들고 시행시킬 수 있는 기능적 논의는 넘치지만, 의미, 목적, 윤리성 등에 관한 고민이 교육공동체 안팎에서 충분하지 않다. 시험은 학습(자)을 돕는 손이 되어야 한다. 그런데 학교와 교사를 돕는 시험의 전통이 없다면 교육선진국으로 성장할 수 없다. 누군가를 돕지 못하고 차별하고 구분하고 등수만 나누는 일에 집중하는 곳, 도덕적이지 않은 평가자가 넘치며 누구도 서로를 쉽게 믿지 못하는 악의가 넘치는 교육현장을 안고서는 한국은 결코 교육선진국으로 성장할 수 없다. 한국은 인간미와 도덕성이 결여된 졸부형 국가로 성장하고 있다는 비아냥을 벌써 받고 있음을 기억해야 한다.

이와 같은 문제의식 안에서 본 논문은 영어교육평가자의 직

업윤리성 문제를 제기하고자 한다. 본 논문은 국내외 사례와 윤리규범 기준을 문헌 탐색을 통해 고찰하고 인식론적 배경 안에서 영어교육평가 분야의 윤리 문제를 학술담론으로 유도한다. 그리고 교육현장에서 평가도구를 제작하고 사용할 때 고려해야 할 윤리성의 문제를 기독교적 관점에서 성찰하기로 한다.

II. 영어교육평가자 직업윤리 운동의 배경

1. 실제 국내외 소송 사례와 연구

우선 언어교육평가가 윤리적 문제와 연관이 있는 실제적인 국내외 사례를 몇 가지만 소개한다. Davies(1990)는 평가의 기능이 선별(discrimination)이라면 평가는 항상 도덕성이 결핍되어 있음을 지적했다. 대체로 규모가 크고 민감한 의사결정력이 있는 표준화검사가 학습과정과 일치되지 않는 평가내용을 갖고 있고 서열로 매겨진 평가결과로 수험자에게 부당한 수치심을 느끼게 한다. 교육선진국의 소송 사례를 보면, 그럴 경우 학생의 권리를 침해했다는 윤리적 문제로 연결되곤 한다. 혹은 짧은 시간 동안 진행되는 표준화시험이 학습자의 학업능력, 혹은 영어능력을 적절하게 보여줄 충분한 환경을 제공하지 못한다는 주장도 윤리성 문제와도 연결된다.

1) 호주 빅토리아주 언어 평가 프로젝트

호주 빅토리아주 교육청은 LAP(Learning Assessment Project) 시험을 1995년에 시행했다. 빅토리아주는 초등학교 3

학년, 5학년 학생의 영어, 수학을 포함한 표준 교육과정에 명시된 과목의 수학능력을 평가하는 과정에서, LAP가 학부모와 교사에게 교육정보(예: 보충학습이 필요한 학생들에게 지원을 할 준비)와 학교 단위에서 필요한 평가방법에 관한 정보를 제공한다고 홍보했다. 그리고 평가결과는 비밀로 유지되고 어떤 윤리적 문제도 없을 것이라고 장담했다.

그렇지만 LAP 시험의 윤리성 문제를 학부모, 교사, 교장들이 제기했다(Lynch, 1997). 81% 가정이 LAP 시험의 시행을 반대했고 교사들은 전원이 LAP 시행을 반대했다. 교사에 의한 교실평가는 학생들의 학업수준을 향상시키려는 목적이 있지만, 외부에서 시행되는 평가는 학업성취도 향상보다는 학생들을 서열화시키고 비윤리적으로 차별시킬 가능성이 높다는 이유였다. LAP 시험문항의 유형이 누구에게나 익숙하지 않아 학생들이 알고 있는 것을 적절하게 보여줄 수 없다는 문제점도 지적되었고, 불필요한 학교 간 경쟁, 시험결과를 통한 학교자금 및 학생 유치 사용의 관행도 지적되었다. 학부모나 교사가 시험 자체를 문제가 많다고 보지 않지만 주정부에 의해 기획된 외부시험이 윤리적으로 부적절할 수 있다고 지적한 호주 LAP 사례는 국내 현장에서도 적용할 여지가 많다.

2) 미국 캘리포니아주 선착장 근로자 선발시험

TABLE(Test of Adult Basic Education) 시험은 읽기, 문법, 수학, 독해능력을 평가하며 캘리포니아 남부 항구에서 소수인종 출신자들이 직업을 갖기 위해 준비해야 했다. EEOC(the Equal Employment Opportunities Commission) 단체는 TABLE에서

측정하는 문항들이 선착장 근로자 선발에 어떤 의미가 있는지 주목하였는데, 결국 윤리적 부적절함에 대한 문제 제기가 받아들여지고 무역해운업협회인(The International Longshore and Warehouse Union and the Pacific Maritime Association)는 선착장 근로자를 고용할 때 이 시험을 더 이상 사용하지 않기로 동의했다. 미국 다른 지역에서는 고용의 목적으로 이 시험이 여전히 사용되고 있지만 TABLE 시험 자체의 구성타당도 논의 밖에서 윤리성의 문제로 시험 시행기관이 도전을 받은 것이다.

3) 비윤리적 시험출제와 준비에 관한 국내 사례

국내에서는 특목고나 대학 진학, 취업이나 승진, 유학 등의 목적으로 경쟁적인 시험준비 문화가 형성되었기 때문에 전문성이나 윤리적 책임감 없이 시험을 성급히 개발하여 부적절한 영리를 취득하는 단체가 많다. 시험결과를 부적절하게 활용하거나 시험을 준비하는 과정에서도 비윤리적 관행이 점점 늘어나고 있다. 자연스럽게 시험시행과 준비과정에서 소송마저 빈번하게 발생하고 있다. 예를 들면, 2003년 일본 문부성 초청 국비유학생 선발을 위한 1차 영어시험 문제 가운데 일부 문항이 시중에서 시판되고 있는 문제집에 실린 문항과 동일하거나 비슷한 사례가 있었다. 당시 서울대는 시험을 주관한 국제교육진흥원측에 피해자 구제를 요청하였다. 인터넷에서 시험 문제가 유사하다는 의혹을 제기하는 항의성 글이 쏟아졌는데, 결국 7개 문항이 서울대가 주관하는 영어능력측정시험인 텡스(TEPS)의 기출 문제와 동일하고 3개 문항이 유사한 것으로 밝혀졌다. 텡스 시험의

경우 문제은행 구성단계에서부터 기출문제를 별도 관리, 재출제의 오류를 차단하고 있으나 문제검토 단계에서 이 같은 원칙이 제대로 지켜지지 않았다. 시중에서 문제집이 유통되고 있었고 외부시험의 문항도 그대로 출제한 문항개발자 역시 윤리의식이 부족했다.

또한 2007년 김포외고 입학시험 중 53문항을 유출한 사건이 있었다. 김포외고 입시에는 모두 80문항이 출제됐으며 이 중 20문항은 영어 듣기 문항이었다. 당시 경찰은 문항 유출자 뿐 아니라 문항을 넘겨받은 학원 관계자, 학부모까지 입건했다. 학원 측으로부터 문항을 넘겨 받은 학부모들은 당시 입학시험 당일 새벽에 관련 학원장의 연락을 받고 자녀들을 학원으로 데려가 유출된 문제를 풀게 했다.

2. 윤리와 영어교육평가: 인식론적 배경¹⁾

20세기 후반, 절대적 진리 추구를 통한 객관적, 보편적 원리의 모더니즘은 패러다임의 전환을 맞게 되었다. 더 이상 모더니즘의 틀 안에서는 점점 다양해지고 변수로 가득 찬 현상을 설명하기에는 기존의 접근 관점이 갖는 실증주의적 한계를 보게 된 것이었다. 모더니즘의 규정된 틀에서 벗어나려는 움직임, 즉, 새로운 방법적 절차와 개념, 새로운 인식의 체계가 요구되었다. 이러한 인식의 변화와 더불어 포스트모더니즘의 영향은 철학, 예술, 사회 전 영역에서 일어났다.

¹⁾ 이 부분은 새한영어영문학회에 게재된 신동일(2006) 논문의 일부임을 밝혀둔다.

영어교육평가영역도 예외는 아니었다. 관찰 가능한 고정된 영어능력 등급을 선다형 문항, 과학적 측정방법으로 평가할 수 있다는 입장은 평가의 객관성과 신뢰성을 중시한다. 이와 같은 실증주의적이고 심리측정학적인 관점은 표준화시험이 대세인 영어능력 및 성취도평가 현장을 지배해 왔지만(Davies, 1990) 개인과 다름에 대한 관심을 가지면서 평가에서도 본격적인 해석학적 전통이 생겨나기 시작했다(Hamp-Lyons & Lynch, 1998). 영어능력을 표준화검사와 과학의 언어로 측정할 수 있다는 입장에서부터 윤리의 문제가 영어평가 현장에 부각되는 시점이기도 하다(Hamp-Lyons, 1989). 영어평가영역의 패러다임 전환은 결국 평가방식 자체에 영향을 미치게 된다.

상대주의적 윤리관에서 자기평가(self-assessment)와 자기반성(self-reflection)의 중요성이 강조되면서(Hamp-Lyons, 1997) 20세기 후반 윤리의 문제는 드디어 응용언어학, 영어교육학의 여러 분야, 그 중에서도 영어평가 영역에서 본격적인 논의가 시작된다. 1997년 19번째로 열린 세계언어평가 학술대회 성격의 갖고 있는 언어평가 학술대회 (Language Testing Research Colloquium)의 대회주제가 '언어평가에서의 공정성' (Fairness in Language Testing)이었으며 1997년 Alan Davies가 편집한 *Language Testing* 학술지 14호는 특별호로 윤리와 언어평가의 문제를 다루었다. 상대주의적 존재론을 가지고 자기평가, 성찰일지 등의 평가방법이 적용되면서 처방이 아닌 해석과 이해의 목적을 지닌 언어평가관이 새로운 학문적 전통을 만들기 시작한 것이다.

개인에 대한 해석학적 이해(Denzin & Lincoln, 1994)에 비

중을 두다 보니 수행능력, 학습과정, 언어능력의 다차원기능에 관심을 가진 대안평가가 보다 빈번하게 현장에서 논의되었고 그런 과정에서 평가(자)의 윤리 문제가 큰 관심을 갖게 되었다. 평가의 과정에서 주관성을 피하지 않고 오히려 각기 다른 경험 세계를 가지고 있는 평가대상과 평가자 내부의 주관성의 가치를 새롭게 이해하고 적용했다. 윤리의 문제를 가르치고 확인하지 않으면 상대주의적 관점의 영어평가는 공정함을 확보할 수가 없기 때문이다.

포스트모더니즘 관점의 언어평가는 윤리성을 강조하고 있지만 영어시험의 공정성을 보편적으로 확인할 수 있는 절대적 윤리지침이 없다는 입장을 갖는다(Hamp-Lyons, 1997). 모든 가치가 상대적이란 관점에서 윤리는 다만 지역적이며 일시적이기 때문에 모든 윤리의 가치들이 특정 영어시험의 개발과 시행에 항상 정당화될 수 없는 것이다. 예를 들면 세계어로서 영어, 문화간 의사소통도구로서 영어가 학습되고 사용되면서 미국과 영국은 자신들의 교육적 기준이나 가치를 무차별적으로 제3세계에 주입시키는 것이 비윤리적인 것으로 비판받을 수 있다. 토플(TOEFL)이 구성타당도(construct validity)와 채점 신뢰도를 확보했다고 하지만 윤리적 상대주의 입장에서는 토플의 시행과 관련 학습이 특정 지역 혹은 사회에서 문제가 될 수 있는 것이다. 각 지역과 사회에서는 자신들의 윤리적 토대와 가치를 자유롭게 선택할 수 있고 각기 다른 영어교육평가 상황에서 지역적 주체가 나름대로의 윤리적 규범을 탐색하는 것이 가장 바람직하다.

하지만 이러한 윤리적 상대주의는 현장에서 실증주의 전통으

로 실제 교육정책을 집행하는 집단을 쉽게 설득시키지 못했다. 포스트모더니즘의 영향력으로 ‘무엇이 지식인가’에서 ‘무엇이 지식을 의미하는가’로 철학적 고민이 변화되었고 사용자들이 직접 그들이 알고 싶은 지식을 선택하는 권리를 가진 포스터모더니즘의 지식이 (영어)교육현장에서 홍보되고 있지만 아직도 한국의 (영어)교육은 한국적 지역의 맥락이 차별적으로 주목받지 못하고, 원어민 혹은 지식인 집단에서 제시되는 총체적(grand) 내러티브 지식을 학습하고 모방하는 수준을 벗어나지 못하고 있다. 소수의 지식 독점에서 이제야 표준화모형으로 대중적 영어교육평가 콘텐츠들이 배포되기 시작했으니 현 시점에서 윤리와 윤리적 상대주의를 공론화시키기에는 상당한 어려움이 있다.

그렇다면 언어평가학자인 Davies(1997)가 제시한 윤리관을 국내 영어교육평가 현장에 반영하는 것이 가장 무난할 것으로 판단된다. Davies는 윤리를 전문적인 직업을 갖는 과정 안에서 이해했다. 전문성을 갖기 위해서는 전문 직업을 수행하는 윤리적 규범(code of ethics) 혹은 실행 규범(code of practice)을 이해하고 훈련받아야 한다. 행동강령과 실행규범은 시간이 지나거나 지역에 따라 변화될 수 있고 또 수정되어야 하지만 직업 전문성은 시공간을 막론하고 당시 약속된 규범에 따라 행동하는 윤리의식을 가져야 하는 것이다. 서구의 윤리적 상대주의 담론은 국내 영어교육평가 현장과 간극이 너무 멀다. 영어시험 개발과 시행을 포함한 여러 영어교육평가 현장에서 우선적으로 전달할만한 윤리규범의 문제의식을 현장에서 요구되는 영어평가자의 전문성과 묶어서 담론화시키는 것이 바람직할 것이다.

3. 윤리규범 작업 사례

전병만, 박준언, 안병규, 오준일, 유제명, 이소영, 김신혜 (2005)의 영어교사 평가의 전문성 연구에 윤리적 규범에 관한 지면이 할애되어 있다. 그들은 Davies(1997)의 지적처럼 윤리성 기준을 두 가지로 나누어 설명했다. 첫째는 평가전문가가 보편적으로 지켜야 할 행동강령에 해당되는 윤리적 규범이며 전문가의 도덕성과 이상적 규범을 제시하고 있다. 둘째는 윤리적 규범보다 구체적인 개념으로 평가자가 지켜야 할 최소 행동규범을 제시하여 전문가답지 못한 행동을 명시하는데 초점을 맞추었다. 한국의 영어교사는 공정한 평가에 관한 큰 부담감을 갖고 학생평가를 실행해야 하기 때문에 실행규범을 서둘러 설정하는 것이 바람직할 것이다. 실행규범은 적용의 대상에 대한 고민이 필요한데, 미국의 경우 평가의 모든 관련자에 적용을 주장하는 NCME (National Council on Measurement in Education) Code of Fair Testing Practices in Education 입장과 표준화 시험을 이용하는 평가 관련자에게만 적용을 주장하는 APA Standards for Education and Psychological Testing (American Psychology Association) 입장 차이가 있다. 교실과 학교 단위에서 교사가 개별적으로 시행하는 시험의 경우 윤리적 적용이 어렵다는 주장을 했지만 시험결과의 해석과 적용이 교실에 머물러 있지 않기 때문에 실행규범을 통해 교사의 시험 개발 및 시행에 압력을 행사해야 한다는 입장도 설득력이 있다.

국내의 교육단체나 평가기관에서 평가자를 위한 윤리적 기준을 공식적으로 설정한 선례는 문헌 검색으로 찾을 수 없다. 그리고 일반 교육단체에서는 교원윤리강령이 있는데 평가에 관한 지침은 쉽게 찾아 볼 수 없다. 사도강령, 한국교원윤리강령, 사도헌장, 사도강령 등의 교원윤리강령을 간단하게 정리하면 다음과 같다. (1) 사도강령은 한국문화협회가 1954년에 제정한 것으로 총3장 32항으로 구성되어있고 제1장 면학수행, 제2장 교학시범, 제3장 교직원공으로 되어있다. (2) 한국교원윤리강령은 대한교육연합회가 1958년에 제정한 것으로 교원의 윤리의 책무성에 대하여 제1장 학생, 제2장 가정, 제3장 사회, 제4장 교직원, 제5장 교양에 관하여 윤리를 제정한 것이다. (3) 사도헌장은 대한교육연합회가 1982년에 제정한 것으로 전문(교육의 기능을 개인적 기능, 사회적 기능, 국가적 기능으로 구분)과 5개항의 본문으로 구성되어 있다. (4) 사도강령은 대한교육연합회가 1982년에 제정한 것으로 사도헌장을 구체적으로 풀이한 행동강령으로 전문과 본문 5개의 장으로 되어있다. 제1장 스승과 제자, 제2장 스승의 자질, 제3장 스승의 책임, 제4장 교육자와 단체, 제5장 스승과 사회로 구성되어있다.

교육 뿐 아니라 기업이나 정부에서 시행하는 언어평가현장에서 윤리지침은 쉽사리 찾아볼 수 없다. 예를 들어 건설교통부에서 진행한 항공전문인력(조종사, 관제사)을 대상으로 한 영어말하기평가도구 개발을 위해 50개의 문항으로 구성된 평가자지침서(신동일, 권오량, 2005)가 개발되었다. 그러나 이 역시 개인의 연구용역으로 제안된 평가자의 윤리지침서 수준이며 기관이나 학회 단위에서 제안된 윤리강령은 시도조차 되지 못하

고 있다.

윤리적이지 못한 평가(의 결과)는 수험자, 학부모, 연구자, 학교, 단체 등에 의해 사회적 문제로 제기될 수 있다. 여러 집단에서 교육평가활동에 관한 전문성 뿐 아니라 윤리성 측면의 문제를 제기하고 있는 만큼 영어교육현장은 더 이상 직업윤리의 문제를 간과할 수 없다. 아직은 여러 시험들이 비윤리적으로 사용되고 있다는 고발 수준의 보고서가 나오고 있지만 이제는 윤리의식에 기반을 교사교육, 평가자훈련에 대한 고민이 시작되어야 한다. 윤리규범, 실행규범의 문서화 작업도 좀 더 구체성이 있어야 한다.

다만 국내 교육현장 안에서 규범을 문서로 옮기고 윤리교육을 시행하는 것이 현실적으로 가능할 지 의문이다. 정직은 사람 마음대로 되는 것이 아니기 때문이다. 개인이 의지적으로 결단하고 정치사회적인 캠페인도 필요하다. 그러나 성령이 함께 하지 않고는 경쟁심과 이기심이 가득한 근대 한국의 교육현장에 흠 없는 정직이 사람들의 마음 속에 들어갈 수 있을까? 아직도 국내에서 교육현장을 변화시키는 윤리적 평가의 사례를 찾을 수가 없다. 어디서 이러한 윤리운동을 시작할 수 있을까?

4. 직업윤리교육의 시작

목적과 환경에 따라 조금씩 그 가치가 다르겠지만 시험이 가져야 할 기본적인 속성은 공평함이라고 할 수 있다. 공평해야 교육적인 영향력, 학습의 질에 긍정적인 연계가 생긴다. 공평한

시험은 타당한 시험이라는 측정용어로 자주 대치되는데 보다 알기 쉬운 대중적 용어는 도덕적이고 윤리적인 시험이라고 할 수 있다. 도덕적이고 윤리적인 시험은 비도덕적이고 비윤리적인 시험결과에도 민감한 시험이다. 시험의 구성타당도가 무난한 판정을 받는다 해도 부적절한 역류현상으로 윤리성을 담고 있지 못하다면 평가 자체의 목적(선발, 배치, 진단, 학업성취도 검증)은 퇴색되는 것이다. 영어평가도구 역시 타당하게 사용되려면 시험 개발과정에서부터 그 사용의 결과에 대한 고려가 있어야 하고, 비윤리적이거나 비도덕적인 결과를 수반하게 되면 그 일에 관련된 평가자가 책임을 지는 사회적 약속이 필요하다.

그렇다면 영어시험의 비윤리성은 구체적으로 어떻게 방지될 수 있는가? Davies(1997)가 제안한 윤리적 규범, 실행 규범을 문서화시키고 현장에 제공하면 비윤리적 관행은 방지될 수 있는가? 규범은 얼마든지 만들 수 있다. 그러나 윤리의식은 상부기관의 강요와 정치적 힘으로 쉽게 생기지 않는다. 신뢰공동체가 없고 경쟁적이고 위협적인 교육문화가 지배적이라면 규범을 약속하고 지켜내기는 쉽지 않다.

불신과 불공평의 정서가 남아 있는 영어교육평가 현장에서 누군가는 직업윤리성을 갖기 위한 역할모형을 말아야 하며, 기독교교육자들이 어찌하면 가장 능동적이고 열린 마음으로 그 운동을 시작할 수 있을 것이다. 조선말에 기독교인들이 양반 상놈의 계급사회를 만민은 하나님 앞에 평등하다는 인권을 설파하면서, 또 가문이나 씨족보다는 개인의 책임과 역할을 더 강조하면서, 한국 땅에 새로운 문화의식을 뿌리 내는데 공헌했다. 그 때만큼이나 지금 이 땅의 교육현장에 정직을 기반을 둔 새로운 교육

평가운동에 기독교육자가 참여해야 할 때이다.

기독교육인들이 참여할 직업윤리 운동 과제 중 몇 가지만 살펴보기로 하자. 첫째, 윤리적 교육평가를 감당하려면 개방성과 능동성을 필요로 한다. ‘능동적으로’ 평가 자료를 구성하거나 접근하고 그 내용을 엄밀히 수정보완한 다음 투명하게 교육공동체 안팎에서 공유한다. 미국의 경우는 법률 제정으로 시험 기준의 채택에 직접적인 영향을 주었다. 윤리의 문제에 소홀한 한국도 시험의 사회적 책임을 부각시키기 위해 법률적인 필요조건을 고려할 필요가 있을 것이다. 이 모든 능동적인 활동을 감당하려면 시키는 일만 수행하고, 하던 대로 일을 관행으로는 부족하다. 기독교육인들이 정직과 성실함으로 모범을 보일 수 있을 것이다. 예레미야 선지자는 이스라엘 백성들이 거짓을 말하며 진실되지 못하여, 결국 이웃은 물론 형제도 믿지 못한 사회가 되지 못했다고 책망한다(예레미야 9:3-5). 지금은 거짓을 말하지 않는 능동적인 윤리 운동만이 교육공동체에서 무너진 관계를 회복시킬 수 있다는 선지자적 캠페인이 필요할 때이다.

둘째, 규범을 문서로 만들고 공동체 내부에서 평가의 약속을 지키는 전통을 기독교 학교에서부터 시작할 수 있을 것이다. 정직하지 않고 대충 일하는 곳일수록 약속된 문서가 없다. 문서로 남겨지더라도 허울로만 남아 있을 뿐이다. “악인은 풀같이 생장하고 죄악을 행하는 자는 다 흥왕할지라도 영원히 멸망하리이다”(시편 92:7) 말씀처럼 나쁜 마음, 정직하지 않은 마음, 자신의 명예와 부를 욕심껏 채우려는 마음은 시들어 말라 죽는다. 경쟁의 교실, 정직하지 않은 교육평가 현장에서는 시들러 말라 죽는 풀 뿐이다. 해마다 죽는 풀이 다시 시작해야 하는 것처럼,

사랑과 공의의 규범이 없는 곳에서는 새 선생님이, 새 시험이, 새 교육정책이 반복해서 시작하다가 금방 말라 죽는 일만 반복할 뿐이다. 시편 12절 “의인은 종려나무같이 번성하며 레바논의 백향목같이 발육하리로다” 말씀처럼 의를 지키고 정직한 걸음만을 내딛는 기독교육자들의 교육전통만이 시들지 않을 것이다.

셋째, 시험만 열심히 준비하거나 준비시키고 경쟁에서 이기는 교육만 강조할 것이 아니라 윤리의 문제를 기독교단체에서 교육시켜야 한다. 평가결과들이 학교나 어떤 시스템의 경쟁적인 부문에서 부적절한 판단자료로 사용되지 못하도록 가르쳐야 한다. 예를 들어, 학교현장과 교육적 간극이 큰 외부시험을 학생들이 보게 하고 그 결과를 학교들의 교육수준을 판단하기 위한 자료로 사용한다거나 입시를 위한 도구로 사용한다면 그 시험은 그 내용적 타당성에 상관없이 비윤리적이고 부도덕한 시험이다. 반복적으로 윤리적 측면에서 부적절한 시험(의 사용)으로 말미암아 학생들의 영혼이 다칠 수 있고 교육적 성장이 멈출 수 있다고 자꾸 누군가는 목소리를 높여야 한다. 정직은 하나님의 방식이기 때문에 기독교육자들은 교육현장에서 직업윤리에 대한 책임의식을 가져야 하며 교회와 기독교단체는 그러한 윤리운동에 관심을 가져야 한다.

기독교육자들은 직업윤리성을 확립하기 위한 평가자 각성운동의 동력이 될 수 있다. 나라를 살리는 전문가로서의 소명의식 없이는, 교육자로서의 지성이 거둬나지 않고는 윤리라는 직업의식을 갖기 어렵다. 교육평가현장에서 직업윤리를 확립하기 위해 여러 지혜가 필요하겠지만 우선적으로 그 분 자신이 정직이요,

진실이요 의로우신 하나님 안에서 거하는 연습이 필요하다. 지금까지 시험의 역사가 보여주듯이, 그 분을 통하지 않고 우리 의지와 지성만으로는 윤리운동의 열매를 맺기 어렵다. 정직하고 진실하고 공정할 수 있는 자질의 뿌리는 내 자신이 아니라 빛 되신 하나님임을 기독교교육자들이 명심해야 할 것이다. 그럴 때만 정직의 행동으로 변화가 시작될 것이다.

“너희가 전에는 어두움이더니 이제는 주 안에서 빛이라 빛의 자녀들처럼 행하라 빛의 열매는 모든 착함과 의로움과 진실함에 있느니라”(에베소서 5:8-9) 말씀에 시험을 관리하고 준비시키는 교육평가자들이 도전을 받아야 한다. “이 세대를 본받지 말고 오직 마음을 새롭게 함으로 변화를 받아 하나님의 선하시고 기뻐하시고 온전한 뜻”(로마서 12:2)을 분별하면서 믿음으로 전문가 사역을 행할 때 착하고 선해지며 의롭고 올바른 열매를 맺어질 것이다.

III. 마치며

지금까지 영어교육이란 시장 논리에 가려지고, 학제적 연구주체로도 주목을 제대로 받지 못했던 영어교육평가의 윤리성 문제를 그와 관련된 사례와 윤리규범 등을 통해 고찰해 보았다. 영어교육평가 현장의 윤리 문제를 상대적 인식론으로 풀어가는 것은 현실적으로도 쉽지 않을뿐더러 보편적 진리를 믿고 있는 기독교 세계관의 입장에서도 적절치 않다. 규범으로, 정치적 힘으로 압박하는 것도 쉽지 않다. 본 논문에서는 윤리의 문제를 하나님 공동체에 접속시켜 직업윤리 교육운동의 단초를 심어보

자고 제안했다.

한국에서 영어시험을 시행하는 결과와 역류현상에 대한 전문가의 책임은 영어시험 사용자 뿐 아니라 전체 사회에도 책임이 확대될 수 있다. 선다형 지필시험의 시대에서 의사소통능력, 발화능력, 수행능력, 잠재력과 소질을 평가하는 시대로 넘어가고 있으며 새로운 패러다임의 평가에는 단 하나의 정답이 없다. 시험의 기능적 단면, 측정학적인 접근도 무시할 수 없겠지만 개발과 시행, 준비와 학습의 전 영역에 자발적인 윤리적 의사결정이 그 어느 때보다 중요하다. 이 때 기독교교육자들의 정직에 대한 각성, 교육평가자를 위한 직업윤리 캠페인이 시작되어야 한다.

일반 학교현장에서 윤리운동을 시작되길 기다리기에는 세상이 너무 경쟁적이다. 기독교사 혹은 기독교단체를 중심으로 교육평가자들의 직업윤리 운동이 시작되기를 기대한다. 학생들에게 성경의 진리를 가르치고 동료 교사와 학생들을 진심으로 사랑하고, 불의에 타협하지 않으며 정의로우며, 교육계의 신뢰를 회복하는 일에 힘쓰는 것이 기독교교육자의 정체성이다. 그렇다면 사랑과 공의의 마음을 시험을 만들고 시행하고 학습을 돕는 현장에서도 쓰임 받기를 소망한다.

깨끗한 평가자로서의 마음을 갖고 돕는 평가를 감당하기 위해서는 기독교공동체의 기도가 필요하며 기독교교육자들의 협력네트워크가 요구된다. 미디어 속의 스타나 학원강사의 말을 교사나 교수의 말보다 더 신뢰하는 교육사회 분위기에서 교육 선진성을 갖기 위한 여러 세상적인 노력도 필요하겠지만 깨끗한 윤리 의식을 갖기 위한 기독교지성의 운동이 하나님 보시기에 참 아름다울 것이다.

시험에 관한 비윤리적 관행의 뿌리는 결국 자기의의를 드러내고 싶은 우리의 미숙함 때문이다. 세상의 명예, 재물에 집착할 때 정직하지 않은 상황이 발생한다. 기독교교육자들이 하나님 나라에 속한 자로 이 땅의 교육현장을 다스리는 자일 것이다. 그렇다면 하나님이 우리에게 허락하신 지적자원과 교육현장을 가장 정직하고 성실하게 관리하는 청지기로서 교육평가현장의 부적절한 관행에 의연한 태도를 가져야 할 것이다.

청지기 정체성이 아니라면 내 것을 챙겨야 하고 결국 나와 내 가족만을 위한 부도덕한 범죄를 저지를 수 밖에 없다. 마음의 중심에서 “너희는 먼저 하나님의 나라와 하나님의 의를 구하여라. 그리하면 이 모든 것을 너희에게 더하여 주실 것이다” (마태복음 6:33)라는 말씀처럼 기독교교육자들이 무너져가는 교육현장에 대한 악평을 멈추고 그저 깨끗하고 공평한 땅으로 회복시켜줄 것을 중보할 때 하나님이 새로운 교육운동의 밑알을 허락하실 것으로 믿는다. 공평한 평가자는 하나님 뿐이다. 그 마음을 목상하고 그 마음을 닦고자 하는 깨끗한 기독교인들의 참여가 너무나 절실한 영어교육평가 현장이 우리 눈 앞에 있다. 기독교교육자들의 정직의 전통이 상업주의에 찌든 영어교육 현장으로 넓혀지기를 기대해본다.

인 용 문 헌

신동일. 2006. 윤리적 측면에서 본 영어평가 분야의 전문성.

「새한영어영문학」 48(3), 155-173.

신동일, 권오량. 2005. 항공공사자 영어구술능력평가도구 객관

성 검증을 위한 모형 제안서. 건설교통부 항공안전본부 용역 보고서.

전병만, 박준언, 안병규, 오준일, 유제명, 이소영, 김신혜. 2005. 영어과 교사의 학생평가 전문성 신장을 위한 연구. 「영어교육」 60(2), 3-27.

Davies, A. 1990. *Principles of Language Testing*. Oxford: Blackwell.

Davies, A. 1997. Demands of Being Professional in Language Testing. *Language Testing* 14(3), 328-339.

Denzin, N and Y. Lincoln. 1994. *The Handbook of Qualitative Research*. SAGE: Newbury Park, CA.

Hamp-Lyons, L. 1989. Language Testing and Ethics. *Prospect* 5, 7-15.

Hamp-Lyons, L. 1997. Washback, Impact and Validity: Ethical Concerns. *Language Testing* 14(3), 295-303.

Hamp-Lyons, L and B. K. Lynch. 1998. Perspectives on Validity: A Historical Analysis of Language Testing Conference Abstracts. In A. J. Kunnan (ed.), *Validation in Language Assessment: Selected Papers from the 17th Language Testing Research Colloquium, Long Beach* (253-276). Mahwah, New Jersey.: Lawrence Erlbaum Associations, Inc.

Lynch, B. K. 1997. In Search of the Ethical Test.

Language Testing 14(3), 315-327.

Spolsky, B. 1997. The Ethics of Gatekeeping Tests:
What Have We Learned in a Hundred Years?
Language Testing 14(3), 243-247.