

쓰기 평가 협의 과정에 나타난 쓰기 평가자의 인식 연구*

서수현**

<차 례>

- I. 서론
- II. 이론적 배경
- III. 연구 방법
- IV. 연구 결과
- V. 결론

I. 서론

‘쓰기 평가는 객관적일 수 있을까?’라는 질문은 곧 쓰기 평가를 둘러싼 객관주의와 주관주의의 대립을 떠올리게 한다. 전통적으로 평가에서 중요하게 생각해 왔던 신뢰도의 개념을 우선시하느냐, 새로운 평가의 패러다임에서 주목하고 있는 타당도의 개념을 우선시하느냐에 따라 평가의 방향이나 핵심은 달라진다(정희모, 2010). 그간 쓰기 평가의 주요한 문제는 ‘신뢰도’의 문제였다. 일반적으로 사람들이 평가가 대단히 논리적이고 공정하게 이루어질 것이라고 생각하는 것도 신뢰도에 대한 믿음에 기인한다고 볼 수 있다. 평가자에게 끊임없이 일관되고 엄격한 채점에 대해서 요

* 이 연구는 2012년 광주교육대학교 학술 연구비 지원에 의한 것임.

** 광주교대

구하는 것 또한 마찬가지이다.

쓰기 평가 연구 영역의 초기의 관심사는 신뢰도를 향상하는 데에 있었다. 국내에서의 연구 또한 평가의 신뢰도, 특히 평가자의 신뢰도에 주목해서 이루어져 왔다(박영민·최숙기, 2010a, b, c; 최숙기·박영민, 2011). 이들 선행 연구들은 평가 연구에서 다루어야 할 핵심 개념들을 중심으로 이를 이론적으로 정교화하고 실증적으로 탐색하였다는 데에 의의가 있다. 그러나 훈련 절차를 통해서 신뢰도는 수용 가능한 수준까지 끌어올릴 수 있었으며(Medley, 1982 ; Curran, Beck, Corriveau, & Monti, 1980), 일반화가능도 이론 등의 평가 이론을 활용하여 평가의 최적화 요소들을 탐색(조재윤, 2009)할 수 있었기 때문에, 초기에 비해 신뢰도에 대한 관심은 줄어들었다고 볼 수 있다.

최근 쓰기 평가 연구 영역의 관심사는 타당도의 문제로 전환되고 있는데, 이때 타당도는 다시 구인 타당도, 상황적 타당화, 실제성, 타당도 국면으로서의 결과로 구분할 수 있다(Murphy & Yancey, 2008). 이 연구에서는 쓰기 평가 연구의 변모 양상을 수용하여, 신뢰도의 문제가 아닌 타당도의 측면에서 ‘평가자’에게 주목한다. 심리측정학에 기반한 전통적인 타당도의 개념은 특정 평가가 측정하고자 하는 능력을 얼마나 잘 설명할 수 있는지에 대한 정도성으로 정의되어 왔으나, 최근 구인 타당도에 대한 강조로 인해 타당도의 개념은 평가 도구 그 자체의 문제가 아니라 평가를 계획하고 실행하며 해석하고 사용하는 사회적 행위 과정 전반을 의미하게 되었다(Messick, 1989 ; Mislevy, 2007 ; Moss, 2007). 이렇게 볼 때에 평가자는 평가의 신뢰성을 높이기 위해 훈련을 받아야 하는 대상이 아니라 평가라는 사회적 행위 과정 속에서 끊임없이 고민하고 의사 결정을 하는 주체로 자리매김하게 된다. 이 연구에서 평가자의 판단에 대해 타당도의 개념으로 접근하는 이유는 여기에 있다.

일반적으로 교사는 몇 가지 단계를 거치며 평가를 경험하게 된다. 교사는 자신이 가르칠 내용이 중요하다는 신념을 갖고 이를 수업 내용으로 실행하며, 학생의 수행을 관찰하고 기록하는 등 학생의 수행에 대한 정보를 수집한다. 교사는 이렇게 얻은 결과를 판단하여 해석하고 이를 토대로

학생 및 학부모와 의사소통하게 되는데, 이러한 일련의 과정들은 다시 차 후에 교사의 내적 신념을 강화하는 것으로 작용하게 된다(Hancock, Turbill, & Cambourne, 1994 : 54). 이러한 과정을 고려해 볼 때에 교사가 어떠한 신념을 갖고 있느냐, 즉 무엇이 중요하다고 생각하고 어떠한 것을 평가하고자 하는지는 평가의 출발점이자 귀결점이 될 수 있다는 측면에서 주목할 만하다. 이렇듯 평가자는 자신의 신념에 준하여 수업을 진행하고 그 과정에서 또는 그 이후에 인지적으로 평가 준거를 인식하고 평가를 시작한다. 그런데 평가자가 자신이 인식한 평가 준거를 평가 과정 내내 변함없이 동일하게 적용하지는 않는다. 평가자는 자신의 판단에 대해 비판적으로 점검하고, 새로운 평가 준거와 그에 대한 가설을 세우고 검증해 나가면서 평가 과정을 의미 있는 장으로 만든다. 이는 평가의 객관성과 주관성이 서로 충돌하는 장면이기도 하고, 어떤 특정한 목소리가 우위를 점하는 장면이기도 하면서, 서로 다른 쓰기 평가 담론이 엮이며 기존의 생각을 수정하고 재구성하면서 형성해 나가는 장면이기도 하다. 자신이 가진 인지적 틀과 다른 이의 인지적 틀을 공유하고 비교하면서 나아간다는 면에서 볼 때에, 이는 공동의 이해를 반영하여 개인적 이해를 추구하는 과정이라고도 볼 수 있다.

그러나 실제로 쓰기 평가 과정에서 평가자들이 무엇을 경험하고 있는지, 그러한 경험이 평가자의 인식에 어떠한 영향을 미치는지에 대해서는 아직 연구가 부족한 실정이다. 이에 이 연구에서는 ‘쓰기 평가 과정에서 평가자는 무엇을 경험하고 있는가? 그리고 이를 경험하고 있는 평가자의 인식은 어떻게 구성되는가?’라는 연구 문제에 답하고자 한다. 따라서 이 연구의 직접적인 목적은 평가자의 인식에 대한 탐색에 있다고 할 수 있다. 이 연구의 결과는 쓰기 평가 분야에서 그동안 소홀히 해 왔던 타당도의 문제를 환기하는 데에 일정 부분 기여할 수 있을 것이다. 더 나아가 국어 교사가 갖추어야 할 평가 전문성을 구축하는 데에 단초를 제공할 수 있을 것이다.

II. 이론적 배경

이 연구에서 대상으로 하는 쓰기 평가는 직접 평가로 한정한다. 직접 평가는 학생이 한 편 이상의 텍스트를 쓰고 한 명 이상의 평가자가 미리 기술된 채점 기준표를 활용하여 이를 채점하는 방식을 의미한다. 그런데 평가를 할 때에 평가 문항이나 이에 따른 평가 준거¹⁾가 항상 원활하게 해석되는 것은 아니다. 직접 평가를 수행하는 과정에서는 글에 대한 평가자의 판단이 계속해서 관여하기 때문이다. 쓰기 평가 과정을 ‘평가자 중재 평가’(McNamara, 2000)라고 표현하기도 한다는 것을 떠올려 본다면, 쓰기 평가에서 평가자를 고려하는 것은 자연스러운 일이라 할 수 있다. 이러한 평가자의 판단과 해석에는 여러 요인들이 영향을 미친다(Ruth & Murphy, 1988). 예컨대, 평가자가 출제 의도를 어떻게 규정하느냐에 따라, 그리고 학생 필자들의 의도를 어떻게 규정하느냐에 따라 평가자의 해석은 달라질 수 있다. 주어진 채점 기준표를 토대로 평가를 시작하지만, 평가자는 때로는 선호하는 답화 모델, 단정하지 못한 글씨에 대한 판단, 창의성이나 독창성의 정도에 대한 선호도 등을 평가에 포함시키기도 한다.

쓰기 평가 과정에서 평가자의 판단에 영향을 미치는 요인으로 전통적으로 꼽혀 왔던 것은 평가자의 학문적 배경²⁾이다. 이는 총체적 평가에 영향을 미치는 글의 특성이 무엇인지를 밝혔던 초창기 연구에서부터 주목된 변인이다. 쓰기 평가의 초창기 연구에서는 교사의 담당 교과에 따라 중요하게 생각하는 평가 준거가 상이하게 나타남을 발견하였다. 예컨대, 영어과 교수는 조직, 아이디어의 전개, 문단 조직을 가장 중요한 준거로 생각

1) 평가 준거(evaluation criteria)는 일반적으로 학생이 성취해야 할 행동이 어느 정도여야 성취했다는 증거로 볼 수 있는가를 나타내는 표준을 의미한다(서수현, 2008a : 42).

2) 또한, 평가자가 갖고 있는 언어적·문화적 배경 또한 평가자의 판단에 영향을 미친다. 대학 교수들은 모국어 화자보다 비모국어 화자의 문장 수준의 실수에 대해 관대한 경향이 나타난다. 그러나 이 연구에 참여한 평가자들은 한국어를 모국어로 하고 있으며, 역시 한국어를 모국어로 하는 피평가자를 대상으로 채점에 참여하였기 때문에, 선행 연구에서 언급한 이러한 언어적·문화적 배경 요인은 배제하였다.

하는 데에 비해, 기계과와 과학과 교수는 내용의 수준, 과제의 요구에 대한 충족, 지시문에서 제기한 응답의 정도가 중요한 준거라고 생각한다(Diederich, 1974). 이러한 연구 결과는 국내 연구 결과에 의해서도 뒷받침된다. 국어 교사가 정서 표현의 글을 선호하는 데에 비해, 사회 교사는 설명적인 글과 설득적인 글을, 과학 교사는 설명적인 글을 선호하는 것으로 나타났다. 쓰기 평가에 있어서도 이러한 차이는 나타난다. 국어 교사가 주제의 명료성과 타당성, 사고의 참신성과 창의성을 중요한 준거로 생각하는 데에 비해, 사회 교사와 과학 교사는 내용의 정확성을 가장 중요한 준거로 생각하고 있었다(조성만·이영진·권태현, 2010). 대학생 글쓰기 수업에서 이는 보다 첨예하게 대립하여 전공 영역에 따라 서로 다른 성격의 쓰기를 요구하는 현상으로 나타나는데, 이 역시 담당 교수들의 학문적 배경에 따른 차이에 기인한 것으로 보인다.

평가에 대한 경험 역시 평가자의 판단에 영향을 미치는 중요한 요인이다. 대체로 수업과 평가 경험이 많은 평가자들은 그렇지 않은 평가자들보다 관대한 경향을 보이지만(최숙기·박영민, 2011), 반대되는 연구 결과도 있다(Sweedler-Brown, 1985). 초보 평가자는 전문 평가자에 비해 종종 전통적인 유형에 가치를 부여하는 양상을 보이기도 한다(Weigle, 1999). 이와 같은 문제는 평가자들 사이에 학생들이 쓴 글을 통해 측정하고자 하는 쓰기 능력에 대한 관점이 불일치하기 때문에 발생하는 것으로 해석할 수 있다.

이렇듯, 전문 평가자와 초보 평가자의 평가 과정은 서로 다른 양상을 보이며 차이가 나타나기도 하지만³⁾, ‘평가를 하는 사람’에게 나타나는 일

-
- 3) 쓰기 평가 과정에서 평가자의 판단에 영향을 미치는 요인을 실증적으로 탐색하기 위한 연구들은 대체로 전문 평가자와 초보 평가자의 차이를 비교·대조하는 방식으로 수행되었다. Hout(1993)에서는 전문 평가자와 초보 평가자가 학생들의 글을 읽고 평가할 때에 서로 다른 부분에 초점을 둔다는 결론을 내렸다. 전문 평가자가 글에서 중요하게 생각하고 있는 부분들이 서로 비슷한 데에 비해, 초보 평가자들은 그렇지 않은 모습을 종종 보였다. 뿐만 아니라, 초보 평가자들은 글을 읽는 도중에 자신이 부여하게 될 점수를 미리 언급하기도 하였는데, 이 역시 평가 과정에 다소 부정적인 영향을 미쳤다. 전문 평가자들은 글을 읽는 도중에 판단을 하기보다는 글을 읽은 후에 전체적인 판단을 하는 모습이 두드러지게 나타났다. Pula & Hout(1993) 또한 전문 평가자와 초보 평가자를 대상으로, 평가자가 평가 기준에 대해 어떻게 이해하고 있는지에 따라 평가 결과에 어

반적인 양상을 공유하기도 한다. 평가자는 학생의 글을 읽고 그 글에 대한 이미지를 떠올리고, 개인적인 반응을 하게 되며, 학생의 글이 평가 기준에 얼마나 부합하는지를 점검한다는 것이다. 이러한 과정을 거치면서 평가자는 학생의 글에서 가장 두드러지는 특징을 보이는 부분을 다시 한번 살펴보고 점수를 결정하게 된다. 물론 점수를 결정하는 과정에서는 글에 대해 진단적 판단을 내리거나 다른 글과의 비교 등의 평가 행동을 수행하기도 한다(Wolfe, 2005). 전문성의 여부에 따라 다소 차이는 있겠지만, 이러한 인지적 특성에 대한 설명은 평가자의 평가 과정을 탐색할 수 있는 틀을 제공한다는 데에 의의가 있다.

최근 쓰기 평가에 대한 대안적 관점에서는 여기에 더하여, 평가자의 ‘인식’과 그에 따른 판단에 주목한다. 이러한 입장의 연구자들은 쓰기 평가가 전통적인 심리측정학적 증거에 대한 배타적인 의존에서 벗어나 구성주의자, 전문가 독자 또는 해석학적인 사회적 협상 절차로 나아가는 경향을 지원하는 모형과 설명을 개발해야 한다고 주장한다(Gitomer, 1993 ; Moss, 1992, 1994 ; Broad, 2003). 이와 같은 논의는 쓰기와 쓰기 교육을 둘러싼 담론, 그리고 평가자가 자신을 어떠한 담론 속에 자리매김하고 있는가의 문제와 깊은 관련이 있다. 쓰기를 둘러싼 담론의 변화에 따라 글을 쓰는 학생과 교사의 위치(position), 이들이 생각하는 쓰기 수업의 모습 또한 이전과는 다른 양상을 보이게 되었다. 그리고 이러한 변화는 쓰기 평가의 초점을 변화시키는 데에도 기여하였다. 결국 교사의 쓰기 교육에 대한 인식은 쓰기 평가의 과정 및 결과에 영향을 미치게 된다고 할 수 있다.

Wohlwend(2009)에서는 쓰기와 쓰기 교육에 대한 접근 방식을 ‘성숙’, ‘기능 숙달’, ‘의도성/지향성’, ‘복합양식 장르’, ‘사회적 실천’, ‘사회 정책’으로 나누어 설명하고 있다.⁴⁾ 쓰기를 ‘성숙’의 담론에서 설명을 하면, 쓰

면 영향을 미치는지를 탐색하였다. 그 결과, 평가자는 자신의 글쓰기 경험, 전문적인 체험 훈련의 경험, 쓰기 지도 경험이나 평가 경험 등에 따라 평가 기준을 달리 적용하는 모습을 보였음을 밝힌 바 있다.

4) Ivanič(2004)에서는 쓰기와 쓰기 학습에 대한 접근 방식을 ‘기능’, ‘창의성’, ‘과정’, ‘장르’, ‘사회적 실천’, ‘사회 정책’로 나누고 있으나, Wohlwend(2009)에서 Ivanič(2004)의 논의를 수용한 후 이를 실제적 자료를 통해 재정립하였기 때문에 그 내용 자체는 대동

기는 내적, 정신적 성숙의 연대기적 과정으로 그려지는 아동 발달 이론에 근거하여 설명할 수 있다. 성숙의 담론 속에서 쓰기 교육을 수행하는 교사는 학생들이 준비가 되기 전까지 문식성 실행에 미리 노출시키는 것을 피한다. 이때에 평가는 운동 민첩성과 낱자 이름과 재인을 포함한 준비도에 대한 표시로 기능한다. 이와 같은 모습은 초기 문자 언어 교육에서 쉽게 발견할 수 있다.

‘기능 숙달’이라는 담론은 그간 우리나라의 쓰기 교육을 전반적으로 주도해 왔던 과정 중심 쓰기 교육 그리고 교육과정의 책무성을 강조하는 현재 학교 교육의 흐름과 맞물린다. 이때에 쓰기는 내용 지식과 기능의 실행 및 시연으로 정의된다. 그리고 이는 정부가 제시한 성취 기준에 따른 수행으로 실현된다. 즉, 교육과정에는 일련의 기능을 수행할 수 있도록 성취기준이 나열되어 있고, 학생들은 이러한 지시에 따라 주어진 과정들을 수행하는 것이다. 교사는 학생들의 수행을 진단하고 교정해 주는 역할을 한다. 이러한 관점에서 평가는 교육과정에서 제시한 성취기준에 따른 채점 기준에 따라 이루어진다. 학생 필자가 주어진 교육과정 성취기준에 얼마나 근접하여 쓰기를 수행하는가가 평가의 주된 관심사가 된다.

구성주의 이론에 입각하여 대두된 ‘의도성’의 담론에서, 쓰기는 사회적 메시지를 만들어 내는 학생 필자의 ‘의도’에 가치를 부여하면서 이루어진다. 이러한 관점에서는 학생들이 산출해 낸 결과물보다 그러한 결과물을 통해 말하고자 하는 의도에 관심을 둔다. 교사는 학생들이 자신의 목적에 적합한 메시지를 만들어 낼 수 있도록 격려하고, 다양한 문식 환경에 접근할 수 있도록 관찰하는 역할을 수행한다. 평가 역시 학생들이 만들어낸 메시지에 대한 해석의 입장을 취한다.

이러한 쓰기 담론의 무게 중심은 점차 ‘글’ 자체에만 주목하는 데에서 벗어나 점차 다양한 양식 그리고 그러한 양식들을 통해 실현하고자 하는 가치로까지 옮겨져 간다. ‘복합 양식 장르’ 담론에서 쓰기는 사회적 흥미에 의해 동기화된 복합 양식적 의미 구성에 대한 탐구를 의미한다. 학생

들은 의미를 표현하기 위해 최선의 양식을 탐색하고 좀 더 수준 높은 텍스트를 만드는 설계자의 역할을 수행하며, 교사는 이러한 설계 안에서 의미를 찾고 학교 밖 실행을 위한 교육과정적 공간을 만들어내는 중재자의 역할을 수행한다. 단일 장르를 넘어서는 복합 양식 장르에 관심을 두기 때문에, 평가는 그림, 구두 서사, 연극, 몸짓, 눈짓 등을 포함하며, 메시지를 전달하기 위한 색, 형상, 소리, 행동에 대한 동시 사용에 초점이 맞추어진다.

‘사회적 실행’ 담론에서는 의미의 조정과 공동 구성이라는 쌍방향적 가치를 쓰기에 부여한다. 학생 필자는 공동체의 문화에 참여하기 위해 글을 쓰는 초보자, 동료의 참여자, 사회적 결속의 창조자이며, 교사는 학생 필자에게 문화적으로 반응하는 멘토의 역할을 수행하게 된다. 평가는 학생이 교실, 가족, 공동체 등의 문식 사건 속에서 사회적 관계를 구축하고 참여할 수 있게 만드는 쓰기의 사용에 초점을 두게 된다.

‘사회 정책’ 담론에서 쓰기는 중립적인 언어를 다루는 것이 아니다. 이때의 쓰기는 언어가 갖고 있는 이데올로기적 영향을 다루는 것과 깊은 관련이 있다. 학생들은 교실에 속하는 방법으로 친구에게 편지를 쓰거나 교실의 규칙을 따르는 등의 문식성 정체성을 수행한다. 그리고 교사는 쓰기 수업에서 다루지 않는 것에 대해 비판적으로 바라보면서 학생과 수업을 변화시킬 수 있는 주제로 기능한다. 쓰기와 쓰기 수업의 이데올로기적 측면에 주목하기 때문에, 평가는 학생의 교실 내에서의 사회적 행동 그리고 관행적 수업에 도전하는 교사의 자기 비판적 인식에 초점이 맞추어져 있다.

이렇듯, 쓰기와 쓰기 교육을 둘러싼 담론들은 다양하다. 교사는 이와 같은 담론들 중에 어느 하나의 입장을 취할 수도 있고, 여러 측면을 동시에 고려하면서 취사선택하는 유보적인 입장을 취할 수도 있다. 교사가 쓰기에 대하여 어떠한 인식을 갖고 있느냐, 그리고 필자로서 자신을 어떻게 인식하면서 학생을 대하느냐의 문제는 쓰기 교육의 핵심적 기제로 작용한다(McKinney & Giorgis, 2009). 이러한 인식이 결국은 교사가 수행하는 쓰기 학습의 현재를 설명해 줄 수 있을 뿐만 아니라 그들이 지향하고자 하는 쓰기 학습의 미래도 그려 줄 수 있기 때문이다.

III. 연구 방법

이 연구는 쓰기 평가에 영향을 미치는 전통적 요인인 평가자의 배경이나 경험을 고려하여 설계하되, 평가자의 인식이라는 새로운 요인이 평가 과정에서 드러나는지 살펴보고자 하였다. 이 연구의 참여자는 경력 9년의 중학교 교사 6명과 사범대 4학년 학생 9명이다. 이 연구에서는 전문 평가자와 초보 평가자에게 공통적으로 찾을 수 있는 평가자의 인식 양상을 탐색하고자 하였기 때문에 교사와 예비 교사 모두를 표집하였다⁵⁾. 교사 6명은 모두 여교사로 대학 졸업 후 졸업 중학교에서 근무하였다는 공통점이 있다. 예비 교사인 사범대 학생들은 동일한 학교에 재학 중이나 국어교육이 주전공인 경우도 있고 복수 전공인 경우도 있다. 이들은 남학생 3명과 여학생 6명으로 구성되어 있다.

참여자들이 평가한 글은 중학교 1학년 학생들이 쓴 독후감⁶⁾이었다. 학생들은 ‘글을 읽은 감상을 한편의 글로 써 보자’라는 과제가 제시된 A4 용지에 수업 시간 45분 동안 독후감을 작성하였다. 학생들에게 독후감 쓰기에 대해 별도의 안내는 제공하지 않았다. 독후감을 평가 대상으로 선택한 까닭은 학생들의 배경 지식이 크게 관여하지 않으면서도, 학교에서 빈번하게 사용하는 글의 형태이기 때문이다. 학생 필체가 쓰기 평가에 영향을 미칠 가능성이 있으나 오히려 그것이 실제 평가의 자연스러운 양상에

5) 물론 교사라고 해서 모두 전문 평가자라고 할 수 없다. 그러나 통상적으로 경력 9년 1급 정교사 자격증 소지자로 임용 이후 졸업 중학교에서 근무하였던 교사 평가자에 대해 중학생의 글쓰기를 평가하는 데에 있어 초보 평가자라고 말하기는 어려울 것이다. 이후 자료에서 전문 평가자인 교사 평가자는 ‘가’, ‘나’ 등으로, 초보 평가자인 사범대생 평가자는 ‘A’, ‘B’ 등으로 표기하였다.

6) 익명의 심사위원의 지적대로, 독후감 평가가 쓰기 평가인지에 대해서는 논의의 여지가 있다. 이 연구에서는 학교 현장에서 학생들이 흔히 접할 수 있는 글이면서도 배경 지식에 크게 구애받지 않는다는 점을 고려하여 ‘양의 탈을 쓴 늑대, 아이스크림’이라는 논설문과 ‘항아리’라는 소설을 읽고 쓴 독후감을 평가 대상으로 선정하였다. 그러나 쓰기 평가의 본질적 속성을 보다 정치하게 규명하기 위해서는 학교에서 이루어지는 전형적인 글쓰기에 대한 평가 과정을 탐색하는 연구가 추가적으로 필요하다고 본다.

가깝다고 판단하여, 학생이 직접 손으로 쓴 글을 그대로 보면서 평가하도록 하였다.

참여자들에게는 <표 1>의 평가 준거를 활용하여⁷⁾ 학생들의 글 28편을 개별적으로 평가하게 한 후, 자신의 평가 결과에 대해 동료 참여자들⁸⁾과 협의를 하는 과정을 수행하게 하였다. 이 연구에서 제시한 평가 준거는 다음과 같다.

<표 1> 독후감 쓰기의 평가 준거

평가 준거	채점을 할 때에 고려해야 할 사항
독자에게 주는 즐거움	독자에게 감동을 제공하는가? 독자에게 생각할 만한 내용을 제공하는가?
필자의 가치관 표현	폭넓고 다양한 생각이 드러나는가? 자신만의 독창적 사고가 드러나는가?
구조의 유기성	내용이 유기적인 조직을 갖추고 있는가? 주제를 나타내기 위해 적합하게 구성되어 있는가?
개성적인 표현	필자의 개성이나 특성이 드러나는가? 정서 표현이 참신하게 표현되는가?
내용의 창의성	사물에 대한 필자의 인식과 관점을 파악할 수 있는가? 창의적이고 남다른 시각이 잘 드러나는가?
내용의 진실성	필자의 생각이 솔직히 드러나는가? 필자의 체험이나 정서와 관련된 내용이 잘 나타나는가?

모두 여섯 개의 평가 준거에 대해서는 각각 1점에서 5점까지 점수를

- 7) 평가자들의 자연스러운 평가 과정을 탐색하기 위해서는 별도의 평가 준거를 제공하지 않는 것이 보다 적절할 수도 있다. 그러나 이 연구에서는 두 가지 측면을 고려하여, 평가 준거를 제공하였다. 첫째, 이 연구에서 의도한 평가 방식이 직접 평가임을 고려하였다. 직접 평가의 본질적 성격과 교육적 효용성, 그리고 학교에서 직접 평가가 사용되는 방식에 비추어 생각해 볼 때, 평가 준거를 제공하는 것이 보다 자연스러운 것으로 판단하였다. 둘째, 아무런 평가 준거도 제시하지 않을 경우, 논의의 초점이 모아지지 않아 평가 협의의 과정을 통해 드러나는 평가자의 인식을 탐색하기 어렵다는 연구 설계의 측면을 고려하였다.
- 8) 교사 참여자와 사범대생 참여자는 3명으로 이루어진 자신의 집단에서 각각 쓰기 평가를 수행하였는데, 이는 예비 교사 참여자가 교사 참여자에게 영향을 받는 것을 방지하기 위함이다.

부여할 수 있게 구성되어 있다. 따라서 학생들의 독후감은 5점에서 30점까지의 점수를 획득할 수 있다.

‘독자에게 주는 즐거움’은 자신의 정서를 표현하는 과정을 통해 독자에게 얼마만큼의 즐거움을 제공할 수 있는지와 관련된다. 여기에는 독자의 공감과 반응을 유도한다거나 감동을 주는 글, 독자의 흥미와 관심, 감상을 즐거움을 제공할 수 있는지 등에 대한 판단이 포함된다. ‘필자의 가치관 표현’은 글을 통해 자신의 가치와 사고를 드러낼 수 있는지와 관련된다. 여기에는 폭넓은 사고력과 통찰력, 작품성, 사고의 깊이 등을 적절한 문학적 장치를 통해 표현할 수 있는지 등에 대한 판단이 포함된다. 글 구조와 관련된 ‘구조의 유기성’에는 주제를 나타내기에 적합한 구성이나 구체적인 서술 등을 포함된다. ‘개성적인 표현’은 자신의 감정을 충분히 드러내고 이를 적절하게 표현하는지 등을, ‘내용의 창의성’은 대상에 대한 새로운 인식을 드러내는 창의적이고 남다른 시각이 드러나 있는 글을 쓸 수 있는 등을 평가하는 준거이다. ‘내용의 진실성’은 독후감의 자기 표현적 성격을 고려하여 읽은 책에 대한 자신의 생각을 가감 없이 표현할 수 있는지를 평가하는 준거이다(서수현, 2008a : 105-114).

평가자의 평가 과정을 탐색하는 연구들은 대부분 프로토콜 분석을 연구 방법으로 취한다. 그러나 이 연구에서는 개별 평가자의 인지적·정의적 판단에 주목하기보다는, 개별 평가자가 쓰기 평가에 대해 갖고 있는 생각들은 무엇이고 그것이 실제 평가 과정 속에서 어떠한 양상으로 나타나는지, 그리고 평가자의 인식이 다른 평가자와의 담론 속에서 어떻게 재구성되는지 등에 주목하였기 때문에, 쓰기 평가 과정을 분석하였다.

평가 참여자들의 평가 내용은 mp3 파일 형태로 녹음하였으며, 이후에 전사의 과정을 거쳤다. 전사를 거쳐서 나온 자료는 A4 기준으로 208장이다. 이 연구의 1차적 자료는 평가자들이 학생들의 글을 평가하면서 이루어진 협의의 과정을 전사한 것이다. 연구자는 협의의 과정을 서로 반복적으로 비교하면서 주제별로 범주화하였다. 수집된 자료를 반복적으로 읽고 해석하면서 주요 의미들을 범주화하였으며, 범주화의 과정 속에서 잠재된 의미를 발견하기도 하였고 새로운 내용을 첨가하거나 기존의 내용을 삭제

하기도 하였다. 이 과정에서 연구자의 범주화에 대해 참여자에게 점검을 받아 의미를 확정하기도 하였다. 평가 협의에서 이루어진 대화만으로는 연구자가 평가자의 인식 양상을 단정할 수 없었기 때문에, 필요한 경우 추가적으로 쓰기와 쓰기 평가에 대한 인식을 직접적으로 묻는 질문을 하기도 하였고 필자로서의 자신을 탐색하는 자서전적 글쓰기를 요청하기도 하였다. 따라서 이 논문의 진술 과정 속에는 연구자의 경험과 이해의 폭에 의한 주관성이 개입되었음을 밝힌다.

IV. 연구 결과

1. 쓰기 평가 과정에서 재구성되는 평가자의 인식⁹⁾

이 연구에 참여한 평가자들은 다양한 담론에 위치하고 있었다. 자신이 위치한 담론이 무엇인지를 명시적으로 드러내지 않은 평가자들도 있었으나(A, B, C, F, H 평가자), 대부분의 평가자들은 기능 숙달 담론에 위치하고 있었다(D, E, G, 가, 나, 다, 라 평가자). 소수이긴 하지만 의도성 담론(마 평가자)나 사회 정책 담론(α 평가자)에 위치하는 평가자도 있었다. 대부분이 기능 숙달 담론에 위치하는 데에 대해 평가자들 스스로는 초·중등학교 시절 자신이 받았던 쓰기 교육 경험과 대학에서 이루어지는 교사 교육 과정에서 과정 중심 쓰기 교육을 중시한 영향으로 해석하였다.

쓰기와 쓰기 교육 그리고 쓰기 평가에 대한 이들의 인식은 쓰기 과제

9) 평가는 ① 평가 근거 및 목적 확인 ② 평가 대상 및 내용 확인과 평가 관점, 전략, 방법 결정 ③ 평가 준거 설정과 평가 지표(자료)수집 ④ 평가 기준 및 규칙 결정 ⑤ 평가적 판단 ⑥ 평가 결과의 보고 ⑦ 평가 결과의 활용이라는 여러 국면에 걸쳐 이루어진다(배호순, 1994 : 44). 평가자의 인식을 전반적으로 살펴보기 위해서는 이들 각 국면에 따라 탐색하는 것이 보다 바람직할 것이다. 그러나 이 연구에서는 평가적 판단이라는 특정 국면에 주목하여 자료를 수집하였으며 이를 귀납적으로 범주화한 결과 평가자의 인식이 두드러지게 나타나는 장면은 쓰기 과제 해석, 평가 기준 인식과 점수 부여에 있었기 때문에 이 세 가지로 범주화하였다.

를 어떻게 해석해야 할 지 고민하면서 드러나기 시작하였다. 그리고 평가를 수행하면서 어떻게 점수를 부여해야 하는지 의견을 교환하면서 첨예하게 대립하기도 하고 서로의 입장을 조정하기도 하였다. 평가자들은 자신이 상정한 바람직한 쓰기 교육의 모습에 대해 강하게 주장하기보다 다른 이의 의견을 들으면서 수용하는 양상을 보였다. 이는 특히 예비 교사들 사이에서 두드러지게 드러났는데, 예비 교사들이 자신의 쓰기 교육 전문성에 대한 확신이 없기 때문에 타인의 의견을 쉽게 수용하였기 때문인 것으로 해석된다.

1) 쓰기 과제의 해석 국면

쓰기 평가를 하기 위해서는 우선 쓰기 과제가 필요하다. 쓰기 평가는 쓰기 과제를 설계한 후 학생 필자가 쓰기 과제를 읽고 쓰며 평가자가 이를 평가하는 일련의 단계로 이루어진다. 이때 이 모든 단계에서는 텍스트와 독자(학생과 교사 평가자)의 상호작용이 발생하며 그 과정 속에서 크고 작은 오해도 생기게 된다. 1차적으로 이러한 오해는 쓰기 과제에 대한 해석이 서로 다르기 때문에 빚어지기도 한다. 쓰기 과제를 해석하는 과정에는 평가자뿐만 아니라 여러 평가 주체들의 해석의 과정이 개입된다. 동일한 개인적 경험에 대한 쓰기 과제를 학생과 교사가 어떻게 다르게 해석하는지에 대한 연구(Murphy, Carroll, Kinzer, & Robyns, 1982), 평가 준거에 대해서 학생과 교사가 서로 다른 인식을 하고 있음을 보여주는 연구(Keech & McNelly, 1982) 등은 쓰기 과제를 둘러싼 다양한 해석이 존재함을 잘 보여준다.

평가자들은 쓰기 과제를 평가하기 위해 우선 주어진 쓰기 과제를 해석한다. 이를 위해 평가자들은 학생들에게 부과된 쓰기 과제를 읽고 그 과제의 요구를 파악한다.

A : 나는 이 과제가 뭘까, 생각을 해 봤어. 선생님이 줄거리를 쓰고 감상을 쓰라고 하지 않았을 것 같아. 왜냐면 책을 읽은 감상을 한편의 글

- 로 써 보자고 했으니까 줄거리를 선생님이 포함시키라고 하지 않았을 거 같아. 그냥 일상적으로 독후감을 쓸 때는 줄거리를 쓰고 느낀 점을 쓰라고 배웠기 때문에 애들이 그냥 그 형식으로 쓴 게 아닌가 싶어.
- B : 그런데, 그렇게 쓰는 게 좋은 건 아니잖아. 선생님이 한 편의 글로 써 보자고 과제를 줬는데, 줄거리, 느낀 점 나열이 한 편의 글이라고 할 수는 없지 않아?
- C : 애들이 어려우니까 그대로 따다가 쓰는 경우도 있는 것 같더라.
- A : 그 글이 주장하는 글이다 보니까, 중1이 읽기에 까다롭잖아.
- B : 그래도 나는 애들이 독후감을 너무 판에 박히게 쓴 거에 좋은 점수는 안 주는 게 좋은 것 같아. 교육적으로도, 나쁜 점수를 받아 봐야 반성을 하고 다음에는 그렇게 안 쓰지.
- A : 너무 박하지 않아? 가르쳐 줘야지.
- C : 일단 주어진 과제는 제대로 해석한 거라고 봐야 하지 않을까? 나도 중학교 때 독후감 이런 식으로 썼던 것 같아.

위 사례는 평가자 세 사람 사이에서 이루어진 협의 과정으로, 학생의 글을 채점하기 이전에 ‘독후감’이라는 글의 성격을 규정하고 이를 어떻게 평가할 것인가를 의논하는 과정의 일부이다. A 평가자와 C 평가자가 기존의 독후감 쓰기의 관행과 자신의 과거 수업 경험 등을 들어 관습적으로 쓰는 독후감에 대해 과제를 수행한 것으로 인정하자는 관점을 취하는 데에 비해, B 평가자는 그러한 관행이 잘못된 것이라면 수정을 하는 것이 보다 바람직하다는 입장을 갖고 있다. A 평가자와 C 평가자는 필자로서 자신이 갖고 있는 경험을 평가 상황에도 그대로 적용하는 모습을 보이는 데, 이는 자신이 학창 시절 글을 썼던 경험과 방식이 쓰기를 가르치는 교사들에게 지속적으로 영향을 미친다는 기존의 연구 결과에서 제시한 사례(McKinney & Giorgis, 2009)와도 유사하다. McKinney & Giorgis(2009)에서 학교를 다닐 때에서 자신의 글에 빨간 표시가 가득한 평가를 받았던 경험이 너무나 괴로웠다고 토로하며 교사가 된 후 수업에서도 직접 쓰기를 가르치고 평가하는 것을 부담스러워하는 Mercedes의 사례와 A 평가자와 C 평가자의 반응은 교사의 학창 시절 경험이 이후에도 영향을 미친다는 점에서 크게 다르지 않다. 평가자가 어떠한 쓰기 교육 담론에 위치하는지도

중요하지만, 평가에 앞서 교사에게 이전에 어떠한 국어 수업을 경험했는지에 대해 되돌아보게 함으로써 현재 자신의 수업과 평가에 대해 메타적으로 점검하도록 하는 절차 또한 필요함을 역설적으로 보여주는 사례라 하겠다.

쓰기 평가를 시행하기에 앞서 빚어진 충돌의 양상은 이후 학생들의 글을 확인하면서 다소 조정이 된다. B 평가자가 학생들의 글이 자신의 ‘기대만큼 나오지 않는다.’며 A 평가자와 C 평가자의 논의를 수용하는 양상을 보였기 때문이다. B 평가자 또한 A 평가자와 C 평가자와 마찬가지로 학교 현장에서의 독후감 쓰기는 관행적으로 이루어지고 있다고 생각하고 있었다. 그럼에도 앞으로의 교육에서는 이를 개선해야 한다고 생각하여 ‘독후감을 너무 판에 박히게 쓴 거에 좋은 점수는 안 주는 게 좋은 것’이라는 입장을 취한 것이다. 그러나 학생들의 글을 평가하는 과정에서 자신이 기존에 경험했던 독후감과 유사한 글이 속출하자 ‘현실에 타협’하며 자신의 생각을 수정하였다. 이들 평가자들은 독후감이라는 글의 성격에 대해서는 본질적으로 유사하다고 규정하고 있었으나, 그에 대한 교육적 판단에 대해서는 이견이 있었다. 평가자들의 이와 같은 이견은 한 쪽이 다른 쪽의 생각을 받아들이는 방식으로 마무리된다. 이는 일시적인 판단 유보라고 할 수도 있고, 갈등 및 충돌의 회피로도 해석할 수 있다. 위 사례에서 B 평가자는 ‘교사가 무조건 자기 생각을 고집하는 것은 좋지 않다고 깨달았어요. 학생에게 충분히 가르쳐 주고 이끌어 준 다음에 평가를 하는 게 필요하다고 생각해서 일단 다른 사람들의 의견에 따른 거예요. 그리고 지금 이 글을 쓴 학생들은 일단 제가 가르친 학생들은 아니니까.’라는 말로 타협의 이유를 설명하였다. 예비 교사들의 평가 사례에서는 판단 유보라는 양상으로 나타나지만, 실제 교사들의 협의 과정에서는 쓰기 과제에 대한 인식의 차이로 인해 보다 격렬한 논의가 일어날 수도 있음을 예상할 수 있는 지점이다.

A, B, C 평가자들은 다른 평가자에 비해 자신이 생각하는 쓰기 교육의 이상적 모습을 명시적으로 밝히기를 주저하였으며, 자신의 판단에 대해서도 강하게 주장하기를 꺼렸다. C 평가자는 ‘머리로는 일정 수준에 도달해

야 하고, 기능이나 전략을 가르쳐야 한다고 생각하는데, 애들 글 보니까 또 그것 말고도 쓰기를 가르치고 평가하는 데는 다른 게 있는 것 같기도 해서.’라고 언급하였고, 이에 대해 A와 B 평가자도 부분적으로 동의하는 모습을 보였다. 쓰기 교사로서 특정 답론에 자신을 자리매김하기 위해서는 쓰기 행위에 자체에 대한 이해와 더불어 쓰기 교육에 대한 폭넓은 이해가 필요하며(Ivanič, 2004) 여러 교육적 경험이 수반되어야 이러한 판단이 공고해짐을 보여주는 사례라 할 수 있다.

2) 쓰기 평가 준거에 대한 절대 기준 설정 국면

평가자들은 학생들의 글을 읽고, 점수를 부여한다. 그 과정에서 자신이 생각하는 이상적인 글에 학생들의 글이 얼마나 부합하는지를 판단하기도 하고, 절대적인 글의 기준이 있는지의 여부에 대해 고민하기도 한다.

D: 저는 항상 선생님의 절대 기준이 있어야 한다고 생각하는데, 막상 평가를 다 해 보고 나니까 그 기준은 언제나 환경에 따라서는 조금 조정이 필요하다고 생각했어요. 평가 기준에 있어서는 학생들의 관점도 고려해 봐야겠다는 생각을 많이 가지게 되어서.

E: 저도 이 점수를 쓸 때 처음에 펜으로 먼저 번에 읽은 학생들부터 쪽 썼는데 내 안에서도 기준이 잘 박혀있지 않으면 안 될 것 같아서 이 글 사이에서는 위계가 정해지게 점수를 배열을 하려고 했어요.

특히, 초보 평가자들은 자신이 ‘절대 기준’을 갖고 있어야 한다고 생각한다. 이는 쓰기에 대해 어떻게 생각하든지의 여부와 관계없이 나타나는데, 그것이 ‘평가의 일관성과 관련이 있다’고 생각하기 때문이다. D 평가자와 E 평가자는 ‘기능 숙달’의 답론에 위치하고 있었으며, 학생들이 도달해야 하는 특정한 수준은 마치 형식주의 쓰기 이론에서 이야기하는 ‘이상적인 글’과 같이 공고하다고 생각하고 있었다. 보편적으로 잘 쓴 글이라는 것이 존재하며 이에 도달하기 위하여 학교에서 교육을 해야 한다는 이러한 생각은 곧 평가에 있어서도 절대 평가 체제를 이용하여 가장 잘 쓴

글부터 가장 못 쓴 글까지 언제나 일관되게 채점할 수 있다는 생각과 이어진다. Ivanič(2004)에서는 여전히 많은 교사들이 쓰기에 있어 이러한 전통적인 관점을 고수하고 있으며, 이 담론에 위치한 교사들은 평가에 있어 정확성을 강조하는 경향을 보인다고 언급하였다.

그러나 ‘절대 기준’을 갖고 ‘일관되게’ 평가해야 한다는 초보 평가자들의 생각은 평가 과정과 평가 협의 과정을 거치는 가운데 변화한다. 평가자들은 실제 쓰기 평가 과정을 경험하면서 글 자체의 속성이나 평가 기준 못지않게 글을 쓴 학생 필자의 특성에 대해서도 심사숙고해야 함을 인식하기 시작한다. 평가자들은 학습자의 수행이라는 측면을 고려하면서 학생에게 처음 부여하고자 생각했던 점수를 조정하는 양상을 보이게 된다.

A : 1점이 낮나? 1점이라고 하긴 너무 야박하지 않아? 다른 애들에 비해 서는 자기 생각을 어느 정도 쓴 것 같은데?

C : 왜 1점을 줬는지 모르겠어. 불쌍해서 준 것 같아.

D : 뭔가 다른 애들 것 읽다가 이걸 읽으니까 너무 잘 써서 5점을 줬던 것 같아.

E : 응. 잘 한 것 같애. 이 나이 대 애들에 비해. 그래서 점수를 잘 줬구나.

이렇듯, 평가자들은 학생의 점수를 결정하기 위해서 ‘이 나이 대 애들에 비해’, ‘다른 애들에 비해서는’ 등의 표현을 사용하며 해당 학생 필자를 다른 학생 필자와 계속해서 비교하는 과정을 밟는다. 이는 평가가 학생 필자 간의 수준과 평가자의 해석을 연결하는 방식으로 이루어지면서 절대 평가 체제에서 상대 평가 체제로 전환되고 있음을 나타내는 것이기도 하다. 초보 평가자들은 자신의 평가 과정을 되돌아보면서, 처음에 글 자체만 놓고 평가를 할 때보다 학생들의 수준을 고려할 때에 평가 준거를 다소 관대하게 적용하였음을 인식하게 되었다.

한 가지 흥미로운 점은 교사 평가자들 사이에서는 이와 같은 ‘절대 기준’에 대한 논의가 나타나지 않았다는 점이다. 교사 평가자는 이에 대해 ‘아이들 수준에 대해서는 어느 정도 감을 잡고 있으니까, 절대적인 수준의

글을 요구하는 게 아니라, 이 정도 수준이면 되겠다, 뭐, 그런 생각을 아무래도 좀 하게 되죠.’라는 진술을 통해 학생의 수준을 고려하여 평가하고 있음을 언급하였다. 이러한 양상은 평가 경험이 많을수록 학생의 글에 대해 관대한 양상을 보인다는 선행 연구 결과(최숙기·박영민, 2011)에 의해서도 뒷받침된다. 의사 결정을 위해 동원되는 인간의 판단에는 임의성이 존재할 수밖에 없으며, 그 임의성은 ‘제멋대로의 변덕스러운 것’이라는 의미보다는 ‘심사숙고한 재량의 뜻’이라는 의미를 지닌다(서수현, 2008a : 5)는 점에서, 교사 평가자의 학생 수준에 대한 고려는 교사의 평가 경험에 바탕을 둔 심사숙고의 결과를 반영한 것이기도 하다.

평가자는 해당 학생의 글을 진단하거나 자신이 지금까지 읽은 글과 비교하여 판단하기도 하지만, 평가 준거와 글의 특성을 관련시켜 해석하기도 한다(Wolfe, 2005).

가 : 좀 격양된 말투가 있긴 한데. 이게 말이 됩니까? 사람인 게 부끄럽다.

이런게 좀 그렇긴 한데, 열심히 썼더라구.

나 : 필자의 가치관 표현은 다 5점 준 거 같애.

다 : 너도 5점 줬어? 나도 5점 줬어. 가치관 표현은 잘하는데, 애박에 5점 안 줬어요.

가 : 진짜? 그러면 셋 다 5점인데 5점 주면 되겠다.

나 : 유기성은 떨어져요.

다 : 작은 따옴표를 내가지고. 이걸 하나로 구성한 게 아니라 책의 구조를 하나하나 따가지고 감상만 열거한 거 같아가지고.

위 사례의 평가자들은 ‘필자의 가치관 표현’, ‘유기성’ 등 제시된 구체적인 평가 준거에 근거하여 학생의 글을 해석하고 있다. 가, 나, 다 평가자 모두 쓰기 교사로서의 자신의 모습에 대해 기능 숙달 담론에 위치하고 있었다고 자평하고 있었다. 이들 평가자들은 학생들이 쓰기 과제에서 주어진 지시 또는 교사의 지시를 충실하게 이행하여 성과물을 내는 것이 중요하다고 생각하고 있었다. 그리고 교사는 학생 글에 대한 진단을 바탕으로 그에 따른 적절한 교정적 처방을 해 주어야 한다는 인식을 공유하고

있었다. 따라서 가, 나, 다 평가자들은 평가 과정을 경험하면서 학생 글을 개선하는 데에 어떠한 도움을 주는 것이 좋을지에 대해 고민하는 모습을 보인다.

나 평가자는 ‘애들한테 글을 쓰라고 하면 평가 기준을 잘 주지 않아요. 꼼꼼하게 평가하기보다는 점수만 주는 편이니까. 그런데, 이렇게 해 보니까 나중에 애들한테도 이게 좋을 것 같아요. 자기 글을 자기가 좀 읽어보고 분석해 보면 더 나아질 수도 있고’라는 말로 평가 경험이 자신의 이후 쓰기 수업에 영향을 미칠 수 있음을 언급하였다. 다 평가자 또한 ‘내가 평가를 받는 것이 별로 즐겁지 않아서, 아이들한테 두루뭉술하게 점수를 주었는데, 좀 다시 생각해 봐야겠어요’라는 말로 자신의 쓰기 평가를 반성하는 모습을 보였다. 쓰기 평가에서 평가 준거를 제시하는 것은 학생들의 반응을 일정한 방향으로 설정할 수 있을 뿐만 아니라 자신의 글쓰기에 대한 자기 평가도 수행할 수 있으며, 이는 평가를 통해서 학습을 경험하는 ‘학습을 위한 평가’의 측면에서도 의미가 있다(서수현, 2008b). 나 평가자와 다 평가자는 쓰기 평가 협의를 통해, 구체적인 평가 준거가 갖는 교육적 효용성을 인식하고 자신의 쓰기 수업과 평가에 대한 개선 지점을 모색하고 있었다. 예비 교사들이 학생 필자의 수준을 고려하면서 절대 기준으로 글을 평가하는 것이 곤란하다는 것을 인식하기 시작한 데에 비해, 교사들은 평가 경험을 통해 자신의 수업과 평가를 개선하고자 하였다. 이는 동일한 학생의 글을 평가하더라도, 평가자의 배경이나 경험 등에 따라 다르게 평가될 수 있고, 그 결과도 다르게 활용될 수 있음을 의미한다(Pula & Hout, 1993).

3) 점수 부여의 불일치 국면

평가자들이 쓰기 평가 준거에 대해 일관되게 해석하고, 해당 학생의 글을 다른 학생들과 비교하고 점검하면서 채점한다고 해서, 그 판단이 평가자 내에서 항상 일치되어서 나타나는 것은 아니다. 평가자는 평가 협의 과정에서 총체적 평가와 분석적 평가의 괴리가 나타나는 경우, 또는 글로

쓴 교사의 반응과 분석적 평가가 일치하지 않는 경우 등을 발견하고 이를 반성하기도 한다.

F : 난 5점까지 썼네. 난 끝에 코멘트를 이렇게 달았거든. 문단을 하나도 안 나눠져 있다고 너무 안 나눠져 있다. 그래 놓고는 5점을 썼네.

D : 나는 끝으로 갈수록 점수가 점점 낮아졌어.

F : 왜 그랬지? 점수도 나쁘게 주고, 말도 나쁘게 쓰는 게 좀 부담스러웠나? 뭘 써야 할 지 모르겠어.

D : 나도. 좋은 말을 써야 할 것 같긴 하고 근데 점수는 낮아.

F 평가자는 자신의 평가를 메타적으로 점검하고 동료 평가자들과 협의하는 과정에서 자신의 반응과 점수가 일치하지 않음을 발견하였다. D 평가자 역시 자신의 평가가 일관성이 없음을 발견하고, 학생 글에 어떤 반응을 해야 하는지에 대해 혼란스러워했다. 이처럼 대부분의 교사의 반응은 모호하고 형식적이며 때로는 상반된 내용을 담기도 한다(Straub, 1996). 이와 같은 현상을 해결하기 위해서는 글을 한 번만 읽으면서 바로 평가하는 것이 아니라 읽고 난 후 다시 한 번 읽으면서 평가의 과정을 거치는 것이 필요하다. 그리고 교사에게는 학생에게 어떠한 조언을 해 주는 것이 도움이 될 수 있는지 고민하는 시간도 필요하다.

라 : 좀 글을 다듬어주는 쪽으로 코멘트를 달아야 할 것 같아.

마 : 그렇게 하면 너무 지시하는 것 같지 않아? 애들한테 생각할 시간을 좀 주는 게 더 좋을 것 같아. 같이 얘기해 봐도 좋고.

라 평가자와 마 평가자는 학생 글에 대한 반응에 있어 상반된 입장을 취하고 있다. 라 평가자가 자신이 이상적으로 생각하는 글의 방향으로 학생의 글을 수정해 주어야 한다고 생각하는 데에 비해, 마 평가자는 학생이 스스로 생각할 수 있는 시간을 주는 것이 더 필요하며, 학생의 글에 대해서 교사와 학생이 소통할 필요가 있다고 생각한다. 마 평가자의 인식에 바탕을 둔 이러한 소통이 이루어질 때에, 학생 필자는 자신이 받은 반응을 주체적으로 선택하며, 자신의 학습을 조절할 수 있다(Underwood, 2008).

이는 교사가 학생의 글을 통제하기 보다는 학생들이 자신의 글에 보다 많은 책임을 질 수 있도록 허용해야 한다는 최근의 쓰기 평가 연구의 방향과도 유사하다(Straub, 1996).

점수 부여의 불일치는 평가 과정 전반에 걸쳐 나타나기도 한다. 평가자들은 때때로 처음 채점한 학생의 글에 대해서 상대적으로 높거나 낮은 점수를 부여하기도 하고, 앞뒤 글의 영향을 받아 해당 학생의 글을 평가하기도 하였다.

A : 규열이, 규열이는 의외로 점수가 높네?

C : 나두야. 3,4, 이렇게 썼어. 동수 뒤에 있어서.

A : 그럴 수도 있어.

G : 나는 사실 재민이한테 엄청나게 많은 점수를 줬어. 왜였나면 내가 재민이를 맨 처음에 매겼냐? 맨 처음에 매겼고 글씨도 정연하고.

H : 응, 글씨 잘 썼어.

이와 같은 양상은 평가자들이 학생의 글을 평가할 때에 가장 어려움을 느끼는 점이 평가 기준이 흔들리는 데에 있다는 선행 연구의 결과(박영민, 2009)와도 유사하다. 평가를 하는 과정에서 자신의 평가 기준이 모호해지면서, 다른 학생들과의 상대적인 비교 속에서 점수를 일관되지 못하게 부여하는 것이다. 글씨 또한 평가자들이 점수 부여에 관련된 어려움으로 작용한다. 평가자들은 글씨가 읽기 어려울 때에 학생 글을 읽기 싫어진다고 보고하는데(박영민, 2009), 이는 반대의 경우에서도 나타난다. 즉, 상대적으로 깨끗한 글씨에 높은 점수를 주는 경향을 보이는 것이다.

점수 부여에 대한 불일치는 평가자 사이에서도 나타난다. 평가자마다 엄격성을 적용하는 정도에서 차이를 보이는 것이다. 평가자에 따라 특정한 점수를 부여하는 것에 부담을 느끼기도 하고, 최고점과 최하점을 결정하는 데에 있어 이견이 드러나기도 한다.

H : 내가 최고점과 최하점을 주는 데에 거부감이 있는 것 같애. 5점을 주

는 것도 거부감이 있고 1점을 주는 것도 거부감이 있어서 5점을 잘 안 주는 것 같애.

B : 그러면 최하점을 1점으로 두신 거네요? 전 최하점이 2점입니다. 기준이 다르군요.

지속적인 칭찬과 긍정적인 반응이 학생 글의 질적 수준 향상에 도움이 된다는 연구 결과(Straub, 1996 ; Bardine, Bardine, Deegan. 2000)를 고려할 때에, 최하점인 1점을 부여하지 않고자 하는 이러한 양상은 긍정적으로 해석할 수 있는 여지가 있다. 다만, 교사가 학생의 글에 대해 안내자 및 조언자의 역할을 보다 원만하게 수행하기 위해서는 최하 점수 부여를 피하는 것에 그쳐서는 안 된다. 긍정적인 반응과 함께 학생을 필자로 인정하고 글에 대하여 함께 협의하는 과정이 수반될 때에, 쓰기 교육에서 평가는 보다 의미 있는 교육의 장으로 자리 잡을 수 있을 것이다. 또한, 이러한 차이는 반드시 조정해야 할 오류라기보다 개별 평가자의 특성으로 간주하는 것이 바람직하다(Engelhard, 1994). 평가자가 평가 준거를 활용하여 자신의 평가를 조정할 수 있다면, 평가자 간의 조정도 이루어질 가능성도 있기 때문이다.

2. 쓰기 평가 과정에서 발견하게 되는 평가자 간의 인식 차이

평가자들은 쓰기 과제를 인식하고 평가 준거를 해석하여 평가를 수행하면서 그리고 점수를 부여하면서 서로 간의 차이를 경험하게 된다. 사람마다 글에 대한 인식이 다르고, 평가자마다 학습자와 과제, 평가에 대한 기대가 다르기 때문에 이러한 차이는 자연스레 발생하는 것이라고 볼 수 있다. 이 차이의 근원에는 ‘쓰기’와 ‘쓰기 평가’ 그리고 ‘학습자’에 대한 서로 다른 인식이 존재한다.

라 : 창의성이 없다고는 생각을 안 했는데 여기가 너무 웃겼어.

마 : 창의적이지 않아? 끝을 표어처럼 마무리하는 게.

라 : 사실 이렇게 쓰면 안 되는 거 아니야?

마 : 아니 뭐 독후감이니까 형식이 없는 거지.

라 : 뭐지 이게 하면서 2점을 줬어.

마 : 한 3점 정도 주면 될 거 같애.

위 사례에서 라 평가자가 관습에서 벗어난 학생들의 글에 대해 용인하기 어려워하는 데에 비해, 마 평가자는 학생의 의도나 지향을 읽어내고자 시도하고 있다. 마 평가자에게 자신이 생각하는 쓰기와 학습자의 모습에 대해 설명해 줄 것을 요청한 결과, ‘표현적 쓰기 활동을 통해 자신을 표현하는 것이 중요하다고 생각한다. 무언가 말하고 싶은 것을 드러낼 수 있게 도와주고 그것을 제대로 읽어 주는 것이 교사가 해야 할 일이라고 생각한다.’고 자신이 생각하는 이상적인 쓰기 교육의 모습을 표현하였다. 마 평가자는 ‘의도성’ 담론에 위치한다고 할 수 있다. 이러한 생각을 하는 교사는 학생들이 글을 통해 드러내고자 하는 의도나 내면의 과정이 무엇인지에 관심을 둔다(Wohlwend, 2009). McKinney & Giorgis(2009)의 연구 참여자인 Erika에게서도 이와 유사한 모습이 나타난다. Erika는 쓰기를 통해 학생들에게 흥미와 창의성을 불어넣을 수 있다고 생각하며, 쓰기를 못하는 아이들을 도울 수 있다고 교사로서의 자신을 규정하였다. 표면적으로 볼 때에 글의 수준이 낮다 해도 쓰기를 통해 학습자가 말하고자 하는 것, 그리고 얻고자 하는 것을 보는 것이 더 중요하다고 생각하고 있다는 데에 마 평가자와 Erika의 공통점이 있다.

평가자들은 서로의 의견이 서로 충돌하는 경우, 중간 점수를 채택하는 양상을 종종 보였다. ‘다른 사람의 생각도 일리가 있으니까요,’ ‘좋은 글이라는 게 세부적인 양상은 서로 달라질 수 있잖아.’ 등의 언급을 통해 1차적인 충돌을 회피하고자 하는 의도를 읽을 수 있다. 그렇지만, 그 이면에는 ‘서로 생각이 달라서, 얘기를 해도 해결이 안 날 것 같다.’는 솔직한 고백이 자리한다. 서로의 생각이 달라서 갈등이 빚어질 것으로 예측되는 경우, 의사소통 참여자들은 애매한 언술(equivocation)을 사용하기도 하고, 자신의 의견을 명확히 말하지 않는 방법을 선택하기도 한다(Bavelas, Black,

Chovil, & Mullett, 1990). 평가자들 역시 이러한 경우 자신의 주장을 강하게 피력하기 보다는 중간 점수를 주면서 평가자들 간의 유사성을 찾으려고 노력하는 양상을 보였다.

평가자들은 자신이 생각하는 쓰기 교육의 모습에 대해 비유적 표현을 들어 설명하였는데, 이 과정에서 서로 생각하는 쓰기 교육의 모습이 다를 수 있음을 확인하기도 하였다.

G : 작문 교육은 물수제비 던지는 것 같지 않아요. 연못이나 강에 납작한 돌을 골라 각도를 가늠하며 돌이 수면에서 얼마나 튀기며 멀리까지 나아갈지 고려해 돌을 던지는 행위와 유사한 거죠. 결국, 작문 교육은 학생들이 필자로서 익히고 터득해야 할 것을 제공하고 그 과정을 돕는 거예요. 가르칠 텍스트와 교수 방법을 고민하며 신중히 선택하고 판단하는 것이 돌을 골라 손에 쥐는 것과 같고, 학생들의 반응과 받아들임의 여부를 살피며 학습 모형을 설계하고 예시 답안을 가늠해 보는 것이 수면과 돌의 각도를 재 보는 행위와 같은 거죠. 실제 학생들에게 교육하고 그 성과를 얻는 것이 돌을 던지고 수면 위에 튀기는 돌을 보는 것과 같고요. 물론 이후에 문제점을 파악하고 수정 과정을 거치는 것 또한 다음 번에 더 멀리 돌을 보내기 위해 생각하는 것과 유사해요.

I : 그걸 교사가 다 할 수 있나요? 우리가 어디까지 제공할 수 있나요? 내가 생각하는 작문 교육은 어린 왕자 속 보아뱀 그림과 같아요. 보아뱀 그림은 겉으로는 모자 같지만 실은 보아뱀이 꼬끼리를 삼킨 모습이지요. 이걸 보고 쉽게 추측하는 무리인 아이들이 있고, 이걸 모르는 어른들이 있는 거죠. 즉, 중립적 글쓰기 문제라고 생각한 것도 실은 중립적이지 않을 수 있다는 거예요. 교사가 학생들에게 글쓰기를 위해 경험을 구체화해 줄 때에 어디까지 해 주어야 하는지도 모호해요. 우리가 중립적인 입장에서 애들을 가르치는 것이 가능하긴 할까요?

G 평가자는 자신이 생각하는 좋은 쓰기 수업에 대해서 설명하면서, ‘성취를 할 수 있어야’ 한다는 용어를 종종 사용하기도 하였다. 이러한 평가자는 ‘기능 숙달’을 쓰기 수업의 주요한 목표로 인식하고 있다. 도달해야 하는 목표가 있고 이에 따른 교육 내용과 교수·학습 방법이 결정되며,

목표에 도달하였는지의 여부를 평가하는 것이 가장 중요한 수업의 흐름이라고 생각한다. 이는 문식성에 대한 자율성 모델(autonomous model of literacy, Street, 1999)과도 맥을 같이 한다. 이 모델의 주된 관심은 개인이 문식성을 얼마나 획득하였는지 평가하고, 개인으로 하여금 문식성을 획득하도록 어떻게 교육할 것인가 하는 방법론을 탐색하는 데 놓여있다(정혜승, 2008 : 170). 결국 성취 가능한 가치 중립적인 기능의 연속체로 문식성을 바라보는 이와 같은 입장의 G 평가자는 학생들이 앞으로의 삶을 살아가는 데에 일정 수준 이상의 쓰기 능력을 반드시 갖추어야 한다고 보는 것이다.

반면에, I 평가자는 다른 평가자들과 구분되는 이질적인 담론에서 있다. I 평가자가 위치한 ‘사회 정책’ 담론에서, 쓰기는 문식성 정체성을 수행하는 것이며 교사는 학습자의 상호작용에 대해 비판적으로 바라보는 주체로 기능한다(Wohlwend, 2009). 이는 문식성에 대한 이데올로기 모델(ideological model of literacy, Street, 1999)로도 설명할 수 있다. 학교 공간 안에는 ‘교사가 원하는 방식의 글쓰기’가 존재하기 때문에, 학생 필자들은 종종 자신이 원하는 방식으로 글을 쓰지 못한다(서수현, 2010). I 평가자는 ‘중립적인 입장에서 애들을 가르치는 것이 가능하긴 할까요?’라는 질문을 통해, 학교 공간 안에 존재하는 교사가 원하는 방식의 글쓰기에 대한 우려를 표명한다. 가치 중립적인 교육이 불가능하다는 인식 하에서, I 교사는 쓰기 활동을 수행하는 주체들에 대해 비판적인 인식이 필요하다고 보는 것이다.

쓰기와 쓰기 평가에 대한 서로 다른 인식으로 인해 동일한 글과 동일한 평가 준거로 충분한 연습을 거친 후에 평가를 수행하더라도 상이한 결과를 낼 수 있다. 그렇기 때문에, 평가에 대한 연습과 훈련으로 신뢰도 계수를 끌어올리는 것보다 서로의 목소리에 귀 기울이면서 평가 협의를 수행하는 것이 보다 바람직할 수 있다. ‘쓰기’를 둘러싼 다양한 담론은 교실 수업과 학생의 수행, 그리고 평가를 보다 풍성하게 설명해 줄 수 있는 가능성을 향해 있기 때문이다. 쓰기 평가 과정에서 드러나는 교사의 인식을 탐색하는 것이 필요한 이유도 바로 여기에 있다.

V. 결론

이 연구는 쓰기 평가 협의 과정에서 나타나는 평가자의 인식 양상을 분석하였다. 그 결과, 평가자들은 자신의 신념과 학문적 배경과 쓰기 평가 경험 등에 따라 서로 다른 쓰기 평가 과정을 수행하고 있음을 확인하였다. 평가자들의 쓰기 평가 과정은 여러 국면으로 나누어 살펴 볼 수 있는데, 쓰기 과제의 해석, 쓰기 평가 준거의 해석, 쓰기 평가 이후 점수 부여에 있어서 평가자의 판단은 다양하게 나타났다. 이 과정 속에서 평가자들은 서로 다른 인식과 판단을 하고 있음을 확인하기도 하였고, 이로 인해 갈등을 겪기도 하였으나 대부분 충돌을 회피하거나 자신의 생각을 조정하는 모습을 보였다.

이들이 서로 다른 인식을 하는 근원적 원인으로는 쓰기와 쓰기 평가, 그리고 학습자에 대한 인식이 서로 상이하기 때문이다. 대부분의 평가자들은 ‘기능 숙달’의 관점에서 쓰기와 쓰기 평가를 바라보고 있었다. 특정한 목표에 도달하였는지의 여부를 평가에서 가장 중요한 것으로 인식하고 있었던 것이다. 그럼에도 불구하고, 일부 평가자들은 변화의 주체로 학생을 받아들이고 학생의 글에서 즐거움을 얻으며 그러한 즐거움의 근원을 찾고자 노력하기도 하였다. 이러한 노력은 교실 쓰기 평가에서는 새롭게 확장된 평가 준거를 개발할 필요가 있음을 나타내는 것이기도 하다.

그러나 이 연구는 몇 가지 제한점을 갖는다. 우선, 이 연구에 참여한 예비 교사 및 교사들이 자신의 수업에서 실제로 만나는 학생의 글이 아닌 연구 자료로 주어진 글을 평가했다는 것이다. 교사가 학생에 대한 여러 맥락적 정보를 갖고 있다면, 그에 따라서 평가 과정 및 결과는 달라질 수 있다. 추후 교실 수업과 평가의 실재를 반영한 설계를 통해 쓰기 평가자의 인식을 보다 실증적으로 탐색하는 연구가 요청된다. 또한, 이 연구에 참여한 예비 교사 및 교사들의 경력은 동일하게 통제하였으나 다양한 배경과 경력을 가진 교사들의 평가 협의 상황에서는 다른 인식의 결과가 나올 수 있다. 교사의 배경 변인을 다양화 하여 특정한 배경을 가진 교사들

에게 나타나는 인식의 양상은 무엇인지, 그것과 상관없이 공통적으로 나타나는 인식의 양상은 무엇인지 등에 대해 다층적으로 살펴보는 연구 또한 요청된다.

이 연구의 결과는 교사의 평가 전문성 측면의 연구에 부분적으로 기여할 수 있을 것이다. 교사의 자기 인식, 필자로서의 자기 인식, 독자로서의 자기 인식 등은 교사의 국어 수업과 평가에 대한 실천을 제고하는 데에 도움을 줄 수 있다. 이 연구는 실제 예비 교사와 교사들이 쓰기 평가 과정에서 경험하는 것이 무엇인지에 주목하였으나, 그들의 대화를 통해 드러나는 인식 양상을 탐색하는 데에 그쳤다는 한계가 있다. 교사들이 쓰기 평가 과정에서 겪는 어려움을 규명하고 그 문제를 해결하기 위해서는 실제로 ‘좋은’ 선생님들이 수업과 평가에서 무엇을 하고 있는지에 대해 천착하여 답을 찾고자 하는 연구가 필요하다. 쓰기 평가 과정에서 나타나는 평가자의 인식을 살펴 본 이 연구가 쓰기 교육을 보다 풍성하게 설명할 수 있는 작은 발걸음이 될 수 있기를 기대한다.*

* 본 논문은 2012. 6. 30. 투고되었으며, 2012. 7. 9. 심사가 시작되어 2012. 7. 31. 심사가 종료되었음.

▣ 참고문헌

- 박영민(2009). 「논술 평가자의 주관적 경험 분석」. 『우리어문연구』 33집. 493~523. 우리어문학회.
- 박영민 · 최숙기(2010a). 「국어교사의 설명문 평가에 대한 모평균 추정과 평가 예시문 선정」. 『우리어문연구』 36집. 293~326. 우리어문학회.
- 박영민 · 최숙기(2010b). 「중학생 논설문 평가의 모평균 추정과 평가 예시문 선정」. 『국어교육』 131집. 437~461. 한국어교육학회.
- 박영민 · 최숙기(2010c). 「Rasch 모형을 활용한 국어교사의 쓰기 평가 특성 분석」. 『국어교육학연구』 37집. 367~391. 국어교육학회.
- 배호순(1994). 『프로그램 평가론』. 원미사.
- 서수현(2008a). 「요인분석을 통한 쓰기 평가의 준거 설정에 관한 연구」. 고려대학교 대학원 박사학위 논문.
- 서수현(2008b). 「쓰기 과제 구성 요소의 설정에 대한 연구」. 『국어교육학연구』 33집. 449~472. 국어교육학회.
- 서수현(2010). 「쓰기 수업 경험과 학생 필자의 정체성 구성」. 『한국초등국어교육』 44집. 141~172. 한국초등국어교육학회.
- 정해승(2008). 「문식성(literacy) 교육의 쟁점 탐구」. 『교육과정평가연구』 11(1). 161~185. 한국교육과정평가원.
- 정희모(2010). 「글쓰기 평가에서 객관-주관주의 대립과 그 함의」. 『우리어문연구』 37집. 217~249. 우리어문학회.
- 조성만 · 이영진 · 권태현(2010). 「교과별 좋은 글의 기준 차이에 대한 분석-국어, 사회, 과학 교과를 중심으로」. 『새국어교육』 86호. 353~378. 한국국어교육학회.
- 조재윤(2009). 「일반화가능도 이론을 이용한 쓰기 평가의 오차원 분석 및 신뢰도 추정 연구」. 『국어교육』 128집. 325~357. 한국어교육학회.
- 최숙기 · 박영민(2011). 「논설문 평가에 나타난 국어교사의 평가 특성 및 편향 분석」. 『교육과정평가연구』 14, 1. 201~228.
- Bardine, B. A., Bardine, M. S. & Deegan, E. F. (2000). Beyond the Red Pen : Clarifying Our Role in the Response Process. *English Journal*, 90(1), 94-101.
- Bavelas, J. B., Black, A., Chovil, N., & Mullett, J. (1990). Truths, lies, and equivocation : The effects of conflicting on discourse. *Journal of Language and Social Psychology*, 9, 135-161.

- Broad, R. (2003). *What we really value : Beyond rubrics in teaching and assessing writing*. Logan : Utha State University Press.
- Curran, J. P., Beck, J. G., Corriveau, D. P., & Monti, P. M. (1980). Recalibration of raters to criterion : A methodological note for social skills research. *Behavioral Assessment*, 2, 261-268.
- Diederich, P. B. (1974). *Measuring Growth in English*. Urbana, IL : National Council of Teachers of English.
- Engelhard, G. J. (1994). Examining rater errors in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement*, 31(2), 93-112.
- Gitomer, D. (1993). Performance assessment and educational measurement. In R. Bennett & W. Ward (eds.), *Construction versus choice in cognitive measurement : Issues in constructed response, performance testing, and portfolio assessment* (pp.241-293). Hillsdale, NJ : Lawrence Erlbaum Associates.
- Hancock, J., Turbill, J., & Cambourne, B. (1994). Assessment and evaluation of literacy learning. In S. W. Valencia, E. H. Hiebert, & P. P. Afflerbach (Eds.), *Authentic reading assessment : Practices and possibilities* (46-62). Newark, D E : International Reading Association.
- Huot, B. A. (1993). The influence of holistic scoring procedures on reading and rating student essays. Validating holistic scoring for writing assessment : *Theoretical and Empirical Foundations*, ed. M. M. Williamson and B. Hout. Cresskill, NJ : Hampton Press, Inc.
- Ivanič, R. (2004). Discourses of writing and learning to write. *Language and Education*, 18, 220-245.
- Keech, C. L., & McNelly, M. E. (1982). Comparison and analysis of rater responses to the anchor paper in the writing prompt variation study. In J. R. Gray & L. P. Ruth (eds.), *Properties of writing task : A study of alternative procedures for holistic writing assessment* (260-315). Berkeley : University of California, Graduate School of Education, Bay Area Writing Project.
- McKinney, M., & Giorgis, C. (2009). Narrating and Performing Identity : Literacy Specialists' Writing Identities. *Journal of Literacy Research*, 41(1), 104-149.
- McNamara, T. (2000). *Language Testing*. Oxford University Press.
- Medley, D. M. (1982). Systematic observation. In H. E. Mitzel (Ed.), *Encyclopedia of educational research* (5th ed., 1841-1851). New York : Macmillan.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., 13-103).

London : Collier Macmillan.

Mislevy, R. J. (2007). Validity as design. *Educational Researcher*, 36, 463-469.

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement : Implications for performance assessment. *Review of Educational Research*, 62(3), 229-258.

Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5-12.

Moss, P. A. (2007). Reconstructing validity. *Educational Researcher*, 36, 470-476.

Murphy, S. & Yancey, K. B. (2008). Construct and Consequence : Validity in Writing Assessment. In C. Bazerman(ed), *Handbook of Research of writing : History, Society, School, Individual, Text*. IEA.

Murphy, S., Carroll, K., Kinzer, C., & Robyns, A. (1982). A study of the construction of the meaning of a writing prompt by its authors, the student writers, and the raters. In J. R. Gray & L. P. Ruth (eds.), *Properties of writing task : A study of alternative procedures for holistic writing assessment* (386-468). Berkeley : University of California, Graduate School of Education, Bay Area Writing Project.

Pula, J. J., & Hout, B. A. (1993). A model of background influences on holistic raters. In M. M. Williamson & B. A. Hout(eds.), *Validating holistic scoring for writing assessment : Theoretical and empirical foundations*. 237-265. Cresskill, NJ : Hampton Press, Inc.

Ruth, L., & Murphy, S. (1988). *Designing writing tasks for the assessment of writing*. Ablex Publishing Corporation.

Straub, R. (1996). The Concept of Control in Teacher Response : Defining the Varieties of "Directive" and "Facilitative" Commentary. *College Composition and Communication*, 47(2), 223-251.

Street, B. V. (1999). Literacy and social change : The significance of social context in the development of literacy programmes. In D. A. Wagner (Ed.) (1999), *The future of literacy in a changing world* (55-72). Cresskill, NJ : Hampton Press, Inc.

Sweedler-Brown, C. O. (1985). The influences of training and experience on holistic essay evaluation. *English Journal*, 74(5), 49-55.

Underwood, J. S. (2008). Effective Feedback : Guidelines for Improving Performance. ICLS'08 Proceedings of the 8th international conference on International conference for the learning sciences Volume 2.

- Weigle, S. C. (1999). Investigating rater/prompt interactions in writing assessment : Quantitative and qualitative approaches. *Assessing Writing*, 6(2), 145-178.
- Wohlwend, K. E. (2009). Dilemmas and Discourses of Learning to Write : Assessment as a Contested Site. *Language Arts*, 86, 5, 341-351.
- Wolfe, E. W. (2005). Uncovering rater's cognitive processing and focus using think-aloud protocols. *Journal of Writing Assessment*, 2(1), 37-56.

<초록>

쓰기 평가 협의 과정에 나타난 쓰기 평가자의 인식 연구

서수현

이 연구는 쓰기 평가 과정에서 나타나는 평가자의 인식에 대한 탐색을 목적으로 한다. 예비 교사와 현직 교사의 쓰기 평가 협의 과정을 통해 살펴 본 결과, 평가자들은 쓰기 교육과 쓰기 평가에 대한 각기 다른 담론에 위치하면서 평가를 수행하고 있었다. 이러한 평가자의 위치에 따른 상이한 인식은 쓰기 과제의 해석, 쓰기 평가 준거의 해석, 쓰기 평가 이후 점수 부여 등 여러 국면에 걸쳐 다양하게 나타났다. 평가자들은 평가 과정에서 자신의 인식이 서로 다른 담론에 위치하고 있음을 확인하기도 하였고, 이들 사이에서 충돌을 겪으면서 인지적 갈등을 경험하기도 하였다. 대부분의 경우, 평가자들은 갈등 과정에서 자신의 입장을 관철하기보다 충돌을 회피하거나 자신의 생각을 조정하는 양보의 양상을 보였다. 평가자의 쓰기 교육에 대한 인식은 이와 같이 쓰기 평가의 여러 국면에 다양하게 영향을 미치고 있었다. 이 연구 결과를 단초로 앞으로 쓰기 교육을 실제 수행하는 교사에 대한 폭넓은 연구가 수행되기를 기대한다.

【핵심어】 쓰기 평가, 평가자의 인식, 교사의 인식, 쓰기에 대한 관점, 쓰기 교육에 대한 관점

<Abstract>

Aspects on Awareness of Raters in Writing Assessment

Seo, Soo-hyun

The purpose of this research is to examine aspects on rater's awareness in the direct assessment of writing. This research is on the hypothesis that writing assessment is reflected in the view of writing and the view of learning to write. The analysis explored complicated ways in which rater's belief and academic background and experience associated with the writing assessment aligned and conflicted. Rater's awareness has appeared in a variety of ways to interpreting of writing task and writing evaluation criterion, and grading. There are tensions and contradictions among these views. In order to maximize what they offer to learners, this research suggest that teachers of writing can benefit from being aware of the existence of all views of writing and learning to write and pedagogic practices associated with them.

【Key words】 writing assessment, rater awareness, teacher awareness, view of writing, view of learning to write