

# ‘생물교과논리 및 논술’ 에서의 논증글쓰기에 대한 동료평가-자기평가-교수자 평가 간 상관관계 분석

강경희\*

제주대학교

## Correlation Analysis among Peer Evaluation, Self-evaluation and Instructor Evaluation on Argumentation Writing in the 'Biology Logic and Essay'

Kyunghee Kang\*

Jeju National University

**Abstract** : This study analyzed correlation among peer evaluation, self-evaluation, and instructor evaluation of argumentation writing conducted in the 'Biology logic and essay' class. To this end, 14 pre-service biology teachers wrote argumentation four times, and peer evaluation, self-evaluation, and instructor evaluation were conducted on these results. The peer evaluation for argumentation writing averaged 2.70 points, the self-evaluation average was 2.68 points, and the instructor evaluation average was 2.71. Looking at the average of each sub-evaluation area, scientific knowledge was the highest in peer evaluation, and the composition area was the highest in self-evaluation. In addition, the area of scientific knowledge was the highest in the instructor evaluation. As a result of examining the standard deviation, the standard deviation of self-evaluation was the largest. As a result of analyzing the correlation between evaluation types, there was a high correlation between peer evaluation and instructor evaluation, and there was a moderate correlation between self-evaluation and instructor evaluation. On the other hand, there was no significant correlation between peer evaluation and self-evaluation. As a result of analyzing the correlation by evaluation area, there was no significant correlation between peer evaluation and self-evaluation in all four areas of scientific knowledge, scientific argumentation, composition, and expression. On the other hand, in the analysis between peer evaluation and instructor evaluation, there was a significant correlation in the areas of scientific knowledge, scientific argumentation, and composition. In addition, there was a significant correlation between self-evaluation and instructor evaluation only in the area of expression. The results of this study show that peer evaluation is highly likely to be applied in the 'Subject logic and essay' class. It also suggests that it is necessary to subdivide evaluation criteria according to the type of evaluation in order to use peer evaluation and self-evaluation.

**keywords** : Biology logic and essay, argumentation writing, peer evaluation, self-evaluation, instructor evaluation

### I. 서론

교사양성을 위한 사범대학 교육과정에서 ‘교과논리 및 논술’ 과목은 많은 대학교에서 교양 과목으로 운영하고 있는 글쓰기 수업과는 달리 ‘각 교과별 특성에 부합되는 논리적 사고의 근본 법칙 및 논술에 관한

교육(MOEST, 2009)’로 규정되어 있다. 이에 따라 사범대학 각 학과 및 전공에서는 교과교육학 필수과목으로 ‘교과논리 및 논술’ 강좌를 운영하고 있다. 중등학교에서의 과학글쓰기는 2007 개정 과학교육과정에서 제시되었고, 이후 여러 차례에 걸쳐 교육과정이 개정될 때마다 지속적으로 강조되고 있는 교수-학습 활

\* 교신저자: 강경희 (kkh6554@jejunu.ac.kr)

\*\* 이 논문은 2023학년도 제주대학교 교육·연구 및 학생지도비 지원에 의해서 연구되었음.

\*\*\* 2023년 3월 6일 접수, 2023년 4월 7일 수정원고 접수, 2023년 4월 17일 채택

<http://dx.doi.org/10.21796/jse.2023.47.1.75>

동이다. 그러므로 과학교사양성 교육과정에서 예비 과학교사로서 교과 논리성을 익히고 논술지도력을 함양하는 것은 매우 중요하다. 특히 사범대학 교육과정에서 ‘교과논리 및 논술’ 과목을 운영하는 것은 글쓰기를 통해 학교 과학교육과 과학교사 양성교육과정을 연계한다는 점에서 의미가 있다. 특히 일반적인 대학교 글쓰기 수업과는 달리 ‘교과논리 및 논술’ 수업은 표현력에 대한 강조보다는 문제 해결력과 비판적 의사소통 함양에 주안점을 두고 있다(Park, 2019).

특히 생물교육 분야에서는 예비 생물교사의 글쓰기를 대상으로 한 연구들(Kim, 2015; Kim & Kang, 2021; Sim, 2016)이 지속적으로 이루어졌다. 그러나 ‘교과논리 및 논술’ 과목이 교사양성 교육과정에서 매우 중요한 의미를 가지고 있음에도 불구하고 실제 이 교과를 대상으로 한 연구는 제한적으로 실시되었다. 실제 국어교과(Chang, 2019), 수학교과(Kim, 2017), 영어교과(Shin, 2016), 과학교과(Park, 2019), 가정교과(Choi, Jun, & Chae, 2015; Lee & Kim, 2018), 한문교과(Kim, 2014), 유아교과(Jang, 2021; Lee & Chang, 2015) 등 일부 교과교육 분야에서만 관련 연구들이 진행되었다. 또한 이 선행 연구들은 ‘교과논리 및 논술’ 수업 운영 방안과 예비 교사를 대상으로 한 인식 조사를 중심으로 실시되었고, ‘생물교과논리 및 논술’을 대상으로 한 연구는 거의 이루어지지 않았다.

글쓰기 수업에서 동료평가 활용의 가능성은 많은 연구들(Kim, 2009; Lee, 2014; Lee, 2018; Oh, 2009)에서 제시된 바 있다. 동료평가는 동료의 글을 보면서 자신의 부족한 부분을 알 수 있을 뿐만 아니라 학습 과정에 대한 반성적 사고를 가능하게 한다(Lee, 2014). 또한 동료평가의 타당성과 관련해 Kulkarni *et al.* (2013)은 교수자 평가와 동료평가 결과 간 높은 상관관계가 나타남을 주장하였다. Lee & Chong (2014)은 대학생들을 대상으로 한 연구에서 쓰기 수준에 따라 동료평가 피드백의 질이 달라짐을 제시한 바 있다. 이와 같은 연구들과는 달리 동료평가는 동료 간 친밀도 등에 영향을 받기 때문에 공정성이 낮다는 주장(Dochy *et al.*, 1999)도 있다. 이러한 측면에서 동료평가를 평가 전문성의 측면이 아닌 학습 전략의 일환으로 접근할 필요가 있다는 지적이 제기되기도 한다(Lee, 2014). 특히 동료들의 글을 평가하는 것은 평가자 자신의 쓰기 능력 향상에 도움을 줄 수 있다(Lee & Chong, 2014). ‘생물교과논리 및 논술’ 과목은 예비 생물교사들의 논리적인 글쓰기 능력 향상과 함께 글쓰기 지도 역량 함양을 지향하고 있다. 그러므로 이 교과에서 이루어지는 논증 글쓰기에 대한 동료평가를 실시하는 것은 예비 생물교사들의 글쓰기 평가 역량을 기르는 데에도 영향을 미칠 가능성이 있다. 또한 글쓰기에 대한 동료평가 결과가

타당하다면 교수자 평가를 보완하는 평가 방법으로 활용할 수 있을 뿐만 아니라 학습자들의 글쓰기 평가 능력을 높이는 데 기여할 수 있다(Yoo, 2015). 따라서 ‘생물교과논리 및 논술’ 수업에서 이루어지는 글쓰기 활동에서 동료평가가 실제 어떤 양상을 보이는지 살펴보고, 그를 토대로 ‘교과논리 및 논술’에서의 글쓰기 활동에 대한 다양한 평가 방안을 모색하는 것은 중요하다.

자기평가는 학습자 스스로 학습 과정과 성과를 되돌아봄으로써 학습을 점검하는 것이므로 메타인지적 활동이라고 볼 수 있다(Jo, Jung & Nam, 2021). 글쓰기 수업에서 자기평가는 동료평가에 비해 상대적으로 덜 주목받아왔지만(Han, 2019), 고쳐 쓰기라는 측면에서 볼 때 학습자의 글쓰기 향상에 유용하게 활용될 가능성이 있다(Hwang & Hahn, 2020). 또한 학습 평가에서 학습자가 자신의 학습 과정과 성취 정도를 평가할 수 있는 기회를 통해 자기 피드백을 경험하는 것은 매우 중요하다(Butler & Lee, 2010). 특히 자기평가에서는 학습자가 자신의 학습 성과를 어느 정도 솔직하게 평가하느냐가 매우 중요한 의미를 지닌다. 이와 관련해 자기평가에서 과대평가 또는 과소평가의 문제가 있다는 지적도 있다(Campbell & Fehr, 1990). 그러므로 자기평가와 동료평가, 교수자 평가를 상호 비교하여 글쓰기 수업에서 활용 가능성을 탐색하는 것은 중요하다. 특히 ‘교과논리 및 논술’ 과목에서 동료평가, 자기평가를 실시하는 것은 예비 교사들이 평가 주체로서 참여할 수 있다는 점에서 글쓰기 평가 능력을 기르는 데 도움을 제공할 수 있다. 그러므로 본 연구에서는 ‘생물교과논리 및 논술’에서의 논증글쓰기를 대상으로 동료평가, 자기평가, 교수자 평가를 실시하고, 평가 유형별 상관관계를 분석함으로써 향후 ‘교과논리 및 논술’ 교과에서 어떤 평가 방법들을 활용할 수 있는지 살펴보고자 한다.

대학교 수준의 글쓰기 수업에서 평가는 매우 중요한 의미를 지닌다(Lee, 2018). 평가 목표 설정과 평가 방법 등에 따라 글쓰기 수업의 성과를 여러 측면에서 살펴볼 수 있다. 글쓰기 평가가 중요한 의미를 지니는데 비해 실제 대학교 글쓰기 수업에서의 평가에 대한 연구는 제한적으로 이루어졌다(Chong, 2009). 글쓰기 평가와 관련해 글쓰기 결과물이라는 내용적 측면으로만 접근하기 보다는 글쓰기 과정에 학습자 간 토의 등 상호 피드백에 초점을 두어야 한다는 지적도 있다(Lee, 2007). 특히 글쓰기 수업에서 학습자 간 상호작용은 의사소통능력을 향상시키는 데에 효과적이라는 주장도 제기되었다(Kim, 2008). 대학 글쓰기 수업에서 학습자 간 첨삭 활동에 대해 분석한 연구(Um, 2019)에서는 첨삭을 통한 동료평가가 학습자의 글쓰기 능력 향상과 소통의 활성화에 긍정적인 영향을 미친 것

으로 나타났다. 그리고 ‘교과논리 및 논술’ 수업은 예비 교사들 간 상호작용을 강화시키고, 과학 글쓰기에 대한 논의를 활성화하는 데 긍정적인 영향을 미치는 것으로 알려져 있다(Park, 2019). 이와 같이 교사양성 교육과정에서 필수적으로 개설되고 있는 ‘교과논리 및 논술’에서의 글쓰기 활동에 대한 연구들은 효과적인 수업 운영 방안을 모색하는 데 중점을 두고 있다. 그러나 ‘교과논리 및 논술’에서 예비 교사들의 상호작용을 토대로 한 글쓰기 평가에 대한 분석은 거의 이루어지지 않았다. 그러므로 ‘교과논리 및 논술’ 교과가 교사양성 교육과정에서 교과교육학 과목으로 운영된 지 10여 년이 경과한 시점에서 글쓰기 평가 방안에 대한 논의는 필수적으로 요구되고 있다.

특히 글쓰기 평가의 신뢰도와 타당도를 높이기 위한 방안들 중 하나로 평가 활동에 학습자가 직접 참여하는 방식이 제안되었다(Chong, 2009). 특히 ‘영어 교과논리 및 논술’을 수강하는 예비 영어교사를 대상으로 ‘교과논리 및 논술’ 과목에 대한 인식을 조사한 연구(Shin, 2016)에서는 학습자들이 글쓰기 활동에 대해 동료평가가 필요하다고 응답하였다. 그러므로 ‘생물교과논리 및 논술’ 수업에서 예비 생물교사들의 글쓰기 산출물을 대상으로 동료평가, 자기평가, 교수자 평가 등 평가 유형에 따른 양상을 비교하는 것은 교사양성 교육과정에서 ‘교과논리 및 논술’ 교과 운영에 의미있는 시사점을 제공할 수 있을 것이다. 특히 동료평가, 자기평가, 교수자 평가 간 상관관계를 분석함으로써 평가 유형 간 일치도를 알아볼 수 있다(Hwang & Hahn, 2020). 즉 높은 상관관계는 하나의 평가 유형에서 높은 점수를 받은 논증글쓰기가 다른 평가에서도 높은 점수를 받았음을 의미하므로 동료평가나 자기평가가 교수자 평가와 함께 다면적 평가로 활용될 수 있는지 그 가능성을 살펴볼 수 있다. 특히 교수자 1인에 의한 글쓰기 평가는 신뢰도의 측면에서 많은 한계점들을 내포하고 있다(Yoo, 2015). 그러므로 본 연구에서는 평가 유형 간 상관관계 분석을 통해 각 평가 유형이 다른 평가 방법을 보완하는 방안으로서의 가능성을 알아보고자 하였다.

‘생물교과논리 및 논술’에서 동료평가와 자기평가를 글쓰기 피드백을 위한 방안으로 활용할 수 있는지 살펴보는 것은 의미가 있다. 글쓰기 평가에서 동료평가, 자기평가, 교수자 평가 간 상관관계를 분석하는 것은 또한 글쓰기 피드백의 경우 평가 점수도 중요하지만 그에 못지않게 어떤 판단 근거에 따라 평가했는지가 중요하므로 피드백 과정에서 서술형 평가 내용을 분석하는 것은 의미가 크다(Lee & Chong, 2014). 이에 본 연구에서는 다음과 같은 연구 문제를 설정하였다.

첫째, 교과논리 및 논술 수업에서 예비 생물교사들의 글쓰기에 대한 동료평가-자기평가-교수자 평가 간

상관관계를 분석한다.

둘째, 예비 생물교사의 글쓰기에 대한 동료평가와 자기평가 중 서술형 평가 내용의 특징을 분석한다.

## II. 연구 내용 및 방법

### 1. 연구 대상

이 연구는 2022학년도 1학기 J대학교 사범대학 생물교육 전공에 개설된 ‘생물교과논리 및 논술’ 수강생인 예비 생물교사 14명을 대상으로 이루어졌다. 이 예비교사들은 이전 학기 동안 생물교육학 관련 2개 과목을 수강한 경험이 있으므로 교과교육학에 대한 기본적인 내용은 이해하고 있는 것으로 볼 수 있다. 또한 이전 학기 수강 교과목에서 예비 생물교사들의 수업실연을 대상으로 동료평가와 자기평가를 실시한 바 있으므로 동료평가와 자기평가 방법에 대해서는 기본적인 이해가 이루어졌다. 특히 학기 초에 본 연구의 내용과 과정을 수강생들에게 설명하였고, 이에 예비 생물교사들은 모두 자발적인 참여 의사를 밝히는 연구 참여 동의서를 작성하였다. 또한 자기평가와 동료평가에서의 응답 내용을 연구 자료로 활용하는 데에 동의하였다.

### 2. ‘생물교과논리 및 논술’ 수업 내용 및 평가 방법

이 수업은 과학적 추론의 기초와 논리 논술의 이해 등 기본적인 내용과 관찰글쓰기, 설명글쓰기, 연구계획서 작성 등 다양한 유형의 글쓰기 실습으로 구성되었다. 또한 중등학교 생명과학 글쓰기 지도에 관한 내용을 다루었고, 특히 SSI (Socioscientific issues)를 주제로 한 논증글쓰기가 이루어졌다. 각 주별 구체적인 수업 내용은 Table 1에 제시하였다.

이 수업에서는 관찰글쓰기, 설명글쓰기, SSI 논증글쓰기 등 예비 생물교사들의 글쓰기가 총 7회 이루어졌다. 이 중 4회의 논증글쓰기를 대상으로 동료평가와 자기평가, 교수자 평가가 실시되었다. 관찰글쓰기는 학기 초에 실시된 간단한 글쓰기로, 수업 중에 예비 생물교사 모두가 발표하고, 그에 대한 피드백을 제시하는 형태로 진행되었기 때문에 동료평가는 실시하지 않았다. 설명글쓰기는 수업 중 발표와 그에 따른 첨삭이 이루어졌고, 연구계획서 쓰기는 모둠별로 진행되어서 동료평가와 자기평가를 실시하지 않았다.

동료평가와 자기평가, 교수자 평가가 이루어진 논증글쓰기는 수업 중에 실시되었다. 우선 예비 생물교사

Table 1. The composition of 'Biology logic and essay' class

주차	수업 내용	수업 방법
1	강좌 내용 소개 글쓰기 과제 및 루브릭을 포함한 동료평가와 자기평가 방식 제시	강의
2	과학적 사고와 과학적 방법에 대한 이해 관찰 등 탐구기능에 대한 이해와 이를 이용한 글쓰기	관찰글쓰기, 발표, 강의
3	귀납추론과 연역추론, 귀추적 사고와 가설연역적 사고의 이해 독자에 대한 이해와 설명의 유형	설명글쓰기, 발표, 강의
4	정언, 선언, 가언삼단논법과 딜레마 형식적 비형식적 오류	논법 유형별 예시 만들기, 발표, 강의
5	연구계획서 작성법 자료 조사와 인용표기	연구 주제 발표, 강의
6	SSI의 개념과 Toulmin의 논증요소 SSI 논증글쓰기(1)-주제: GMO 작물 재배	글쓰기, 동료평가, 자기평가, 교수자 평가, 강의
7	SSI 교육과 과학적 소양 SSI 논증글쓰기(2)-주제: 인공지능의 활용	글쓰기, 동료평가, 자기평가, 교수자 평가, 강의
8	과학과 핵심역량과 SSI 교육 SSI 논증글쓰기(3)-주제: 원자력발전의 활용	글쓰기, 동료평가, 자기평가, 교수자 평가, 강의
9	논증글쓰기와 토론 SSI 논증글쓰기(4)-주제: 동물실험 허용	글쓰기, 동료평가, 자기평가, 교수자 평가, 강의
10	연구계획서 쓰기 연구 노트 및 연구보고서 작성법	모둠별 글쓰기, 발표, 강의
11	생명과학 교과서에 제시된 글쓰기 문제 검토 중등학교 과학 글쓰기 지도 방안	강의, 토의, 발표
12	글쓰기 유형별 평가 글쓰기 평가 루브릭 작성	강의, 토의, 발표
13	실용적 글쓰기의 이해 대중을 위한 글쓰기	글쓰기, 교수자 첨삭, 강의
14	논리와 논술의 관계 토론과 논술 연계 활동	강의, 토의, 발표
15	기 말 고 사	

들은 자신의 글쓰기를 완성한 후 자기평가를 실시하였고, 해당 글쓰기 결과물을 대학 내 원격수업센터의 학생 토론 게시판에 업로드하였다. 예비 생물교사들의 글쓰기 결과물이 모두 탑재된 후 동료평가가 이루어졌다. 예비 생물교사들은 게시판에 탑재된 동료들의 글쓰기 결과물을 읽고 평가하였고, 완성된 자기평가지와 동료평가지는 오프라인으로 제출하였다. 동료평가가 이루어지는 동안 교수자 평가도 진행되었다. 동료평가표와 자기평가표는 Table 2에 제시된 루브릭에 따라 각 평가 준거별로 3점 만점의 점수 체계로 구성되었고, 서술형으로 전체적인 총평을 기술하도록 구성

되었다.

논증글쓰기의 주제는 SSI의 개념과 예시 등에 대한 학습 후 4개 모둠에서 토의를 통해 추천한 내용으로 선정되었다. 첫 번째 글쓰기 주제는 GMO작물 재배, 두 번째 주제는 인공지능의 활용, 세 번째 주제는 원자력 발전의 활용, 네 번째 주제는 동물실험 허용으로 정하였다. 이 밖에도 인간배아연구, 임상시험, 전기차, 배양육 등 다양한 주제가 모둠 토의 과정에서 제기되었고, 이 중 각 모둠에서 최종 선정한 주제들을 논증 글쓰기 주제로 정하였다.

Table 2. Rubrics on argumentation writing

범주	평가 준거	내용
과학 지식	정확성	제시된 과학 내용지식(과학 단위, 용어, 개념, 원리, 법칙, 이론 등)이 과학적으로 정확하게 해석되고 적용된다.
	관련성	제시된 과학 내용지식(용어, 단위, 개념, 원리, 법칙, 이론)이 과제의 내용과 관련이 높다.
	근거	설명이나 주장을 뒷받침할 수 있는 다양한 근거들이(일상적 경험, 사실, 연구 결과, 이론 등) 제시되고, 제시된 근거가 검증된 자료로서 믿을 수 있다.
과학 논증	추론	제시된 근거나 데이터가 어떻게 주장을 뒷받침하는지 충분히 설명함으로써 제시된 근거로부터 주장이나 결론이 논리적으로 타당하게 도출된다.
	반증	자신의 주장이나 관점과 다른 주장, 대안을 제시하고 그에 대한 논리적인 대응이나 반박이 체계적으로 이루어진다.
구성	조직	글 전체의 문단이 적절하게 나뉘지고, 적절한 논리 접속사를 이용하여 문장이나 문단들이 글의 주제와 논리적으로 일관성 있게 연결된다.
	형식	글의 형식이 과제에서 요구하는 유형에 적합하다.
표현	자료 표현	자료의 의미를 효과적으로 전달하기 위해 표, 그래프 등 적절한 자료의 유형을 선택하여 오류 없이 정확하게 제시한다.
	인용 표기	참고 문헌과 인용표기를 표기 형식에 따라 바르게 표기한다.

### 3. 글쓰기 평가 기준 구성

본 연구에서는 ‘생물교과논리 및 논술’ 수업에서 예비 생물교사들의 글쓰기에 대한 동료평가와 자기평가, 교수자 평가를 실시하기 위해 평가 기준을 구성하였다. 본 연구에서는 선행 연구(Park, 2019)에서 제시된 기준을 일부 수정하여 활용하였다. 논증글쓰기의 주제는 수업 중 토의를 통해 결정되었기 때문에 주제 영역의 항목을 제외하여 루브릭을 구성하였다. 수정한 루브릭은 과학교육학 박사 1인과 현직 생물교사 2인이 검토하였다. 과학교육학 박사 1인은 과학 글쓰기 루브릭 개발 연구를 수행한 경험이 있고, 생물교사 2인은 수업에서 과학 글쓰기를 실시하고 평가를 진행한 바 있다. 3인의 검토 결과를 토대로 논증글쓰기 평가 루브릭을 최종 구성하였고, 이 루브릭은 Table 2에 제시하였다. 특히 본 연구에서는 예비 생물교사들의 글쓰기에 대한 동료평가, 자기평가, 교수자 평가를 비교할 것이므로 모두 동일한 평가 기준에 따라 평가를 실시하였다. 이와 더불어 동료평가와 자기평가에서는 전체적인 총평을 서술형으로 작성하도록 하여 이를 토대로 예비 생물교사들이 동료들의 글쓰기를 평가할 때 어떤 측면에 주목하였는지를 살펴보았다.

### 4. 자료 수집 및 분석

본 연구에서는 동료평가, 자기평가, 교수자 평가를 비교하기 위해 평균 비교 분석과 상관관계 분석을 실시하였다. 각 평가항목 당 3점 만점 점수 체계로 평가하였고, 3점은 우수, 2점은 보통, 1점은 미흡으로 설정하였다. 논증글쓰기를 실시할 때 루브릭의 평가 준거들로 구성된 동료평가표와 자기평가표를 배부하였고, 각 항목별로 3점 점수 척도로 평가하였다. 또한 동일한 방식으로 교수자 평가가 이루어졌다. 평균분석은 글쓰기 결과물에 대한 전체 평점 평균, 평가 항목별 평균, 표준편차를 살펴보았다. 평균 분석은 각 평가 유형에서 나타나는 기준의 높낮이 차이를 살펴보기 위한 것이다. 또한 각 평가 유형 간 Spearman 상관관계를 분석하였다. 상관관계 분석은 동료평가와 자기평가, 교수자 평가 간 결과 유사성이 있는지를 알아보기 위한 것이다. 특히 본 연구는 14명의 예비 생물교사들이 작성한 논증글쓰기 결과물을 대상으로 한 것이기 때문에 그에 대한 동료평가, 자기평가, 교수자 평가 결과를 정규분포로 전제하기 어렵다. 그러므로 등분산을 가정하지 않는 비모수 상관관계를 나타내는 Spearman 상관계수를 활용하였다. 선행 연구(Hwang & Hahn, 2020)에서 나타난 바와 같이 예비 생물교사들의 글쓰기 산출물에 대한 동료평가, 자기평가

가, 교수자 평가의 상관관계가 높을수록 평가 양상이 서로 일치하는 것으로 볼 수 있다. 즉 상관관계 분석을 통해 평가 유형 간 평가 패턴을 비교하고자 했다. 모든 통계처리는 SPSS 24.0을 활용하였다.

또한 동료평가와 자기평가에서는 점수척도 외에 서술형으로 총평을 기술토록 하였다. 서술형 평가에 대한 분석은 귀납적 접근법(Miles & Huberman, 1984)에 기초해 실시하였다. 예비 생물교사들의 서술형 평가 내용은 질적 자료 분석법 중 하나인 귀납적 접근법을 토대로 하여 범주화하였다. 우선 평가 내용으로부터 원자료를 추출한 후 이를 반복적으로 세밀하게 읽어 핵심 요소들을 추출하였다. 추출된 핵심 요소를 토대로 동료평가와 자기평가에서 나타난 특징들을 범주화하였다. 서술형 평가 내용에 대한 질적 분석은 과학교육학 박사 2인에 의해 이루어졌다. 분석자 중 1인은 교과논리 및 논술 강의를 10학기 이상 진행해왔고, 다른 분석자 또한 과학 영재 대상 글쓰기 수업을 실시한 경험이 있다. 또한 분석자들은 글쓰기 프로그램 개발, 글쓰기 유형별 루브릭 개발 등 과학 글쓰기 관련 연구를 수년 간 수행하였다. 서술형 평가 내용에 대해 분석자들은 각각 일차 분석을 한 후 분석 결과를 비교하였고, 상호 검토와 논의를 거쳐 최종 결과를 도출하였다.

### Ⅲ. 연구 결과 및 논의

#### 1. 동료평가-자기평가-교수자 평가 비교

논증글쓰기 평가 영역별 동료평가와 자기평가, 교수자 평가의 평균과 표준편차를 알아보았다. 동료평가는 평균 2.70점인데 비해 자기평가 평균은 2.68점, 교수자 평가 평균은 2.71점으로 나타났다. 자기평가 평균이 다른 평가 유형에 비해 상대적으로 낮게 나타났다. 이와 같은 결과는 선행 연구(Hwang & Hahn, 2020)의 결과와는 다른 경향을 나타내는 것이다. 표준편차를 살펴보면 자기평가의 경우가 가장 크게 나타나 평

균값을 중심으로 개인별 자기평가 점수가 비교적 넓게 분포되었음을 알 수 있다. 이에 비해 동료평가와 교수자 평가는 표준편차가 작게 나타났다. 예비 과학 교사들의 자기평가에서 자기관대화 경향이 나타나지 않는다는 연구(Lee & Kang, 2018)와 비교해 볼 때 수업실연과 글쓰기라는 학습 활동의 차이가 있지만 본 연구에서의 자기평가에서도 자기관대화 반응은 거의 나타나지 않았음을 알 수 있다. 예비 생물교사들이 자기관대화 반응을 나타내지 않았기 때문에 자기평가 결과가 비교적 폭넓게 분포한 것으로 판단된다. 특히 본 연구에서는 자기평가를 작성한 후 동료평가가 이루어졌기 때문에 평가 기준에 대한 보정이 이루어졌을 가능성이 있다. 이에 따라 자기평가 점수의 편차가 상대적으로 높게 나타났다고 볼 수도 있다. 그러므로 평가의 순서를 변경하였을 때 자기평가와 동료평가 각각의 편차가 어떻게 나타나는지 살펴볼 필요가 있다.

하위 평가 영역별 평균을 살펴보면 동료평가에서는 과학지식 측면이 평균 2.84점으로 가장 높았고, 다음으로 구성 영역이 평균 2.75점이었다. 이에 비해 자기평가에서는 구성 영역이 평균 2.84점으로 가장 높았고, 과학지식 영역이 평균 2.73점이었다. 교수자 평가에서는 과학지식 영역의 평균 점수가 2.85점으로 가장 높았고, 다음으로는 구성 영역이 평균 2.78점으로 나타났다. 이 결과에서 나타난 바와 같이 동료평가, 자기평가, 교수자 평가 모두 과학지식 영역과 구성 영역에서 높은 점수를 보였다. 이에 비해 과학 논증과 표현 영역이 상대적으로 가장 낮은 수준을 보였다. 하위 평가 영역에 대한 표준편차 분석에서도 자기평가 결과는 동료평가와 교수자 평가에 비해 넓게 분산된 양상을 나타냈다.

평가 유형 간 관련성을 알아보기 위해 비모수 집단 간 Spearman 상관관계를 분석하였다(Table 4 참조). 분석 결과 동료평가와 교수자 평가는 Spearman's  $\rho$  .672(유의수준 < .01)로 높은 상관관계를 보였다. 또한 자기평가와 교수자 평가 간에는 Spearman's  $\rho$  .515(유의수준 < .01)로 보통의 상관관계를 나타냈다. 이보

Table 3. Mean and standard deviation by evaluation type

평가영역	동료평가		자기평가		교수자 평가	
	평균	표준편차	평균	표준편차	평균	표준편차
과학지식	2.84	.11	2.73	.42	2.85	.12
과학논증	2.62	.22	2.54	.34	2.59	.28
구성	2.75	.13	2.84	.30	2.78	.15
표현	2.59	.15	2.61	.32	2.58	.28
전체	2.70	.10	2.68	.32	2.71	.12

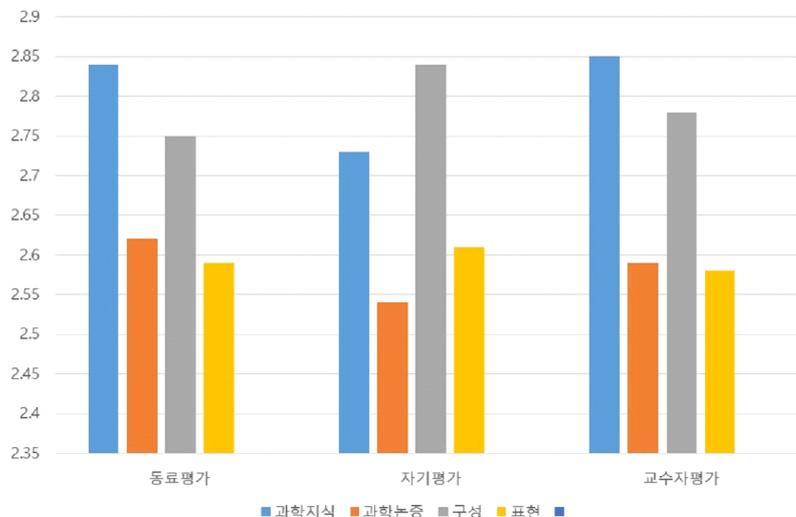


Figure 1. Comparison by evaluation area

다는 약하지만 동료평가와 자기평가 간에는 Spearman’s  $\rho$ 가 .210으로 유의미한 상관관계가 나타나지 않았다. 이 결과는 동일한 글에 대한 자기평가와 동료평가의 양상이 다르게 이루어지고 있음을 보여주는 것이다. 또한 선행 연구(Hwang & Hahn, 2020)의 결과와도 차이가 있다. 이 선행 연구에서는 한 단락 글쓰기에 대한 평가 유형 간 상관관계를 분석한 데 비해 본 연구에서는 논증글쓰기에 대한 평가를 대상으로 했다. 이는 평가 대상인 글쓰기의 유형에 따라 동료평가와 자기평가가 다르게 나타날 수 있음을 시사하고 있다. 본 연구에서 동료평가와 자기평가 간 유의미한 상관관계가 나타나지 않았다는 것은 예비 생물교사들이 논증글쓰기를 평가할 때 일관된 평가 기준을 적용하지 못할 가능성을 보여주고 있다. 이 부분을 보완하기 위해서는 글쓰기 동료평가와 자기평가를 반복적으로 경험하게 하여 명확한 평가 기준을 활용토록 할 필요가 있다. 특히 평가 루브릭에 따라 동료평가와 자기평가 간 상관관계가 어떻게 나타나는지를 반복적으로 살펴보는 시도가 요구된다.

상관관계 분석에서 동료평가와 자기평가가 교수자 평가 결과와 유의미한 상관관계를 나타낸다는 것은 동료평가나 자기평가에서 높은 점수를 받은 글쓰기 결과물은 교수자 평가에서도 높은 점수를 받았고, 한 평가 유형에서 낮은 점수를 받은 경우는 다른 평가 유형에서도 낮은 점수를 받았음을 의미하는 것이다.

선행 연구들(Ahn, 2008; Sadler & Good, 2006)에서는 동료평가와 자기평가 등의 학생평가가 교수자 평가와 높은 일치도를 보였으므로 실제 학습 평가에 활용될 수 있음을 주장하였다. 본 연구의 결과에서는 동료평가와 교수자 평가 간, 자기평가와 교수자 평가 간 상관관계를 보이고 있다. 그러므로, ‘교과논리 및 논술’ 수업의 글쓰기 평가에서 교수자 평가와 함께 동료평가와 자기평가를 활용할 수 있음을 시사하고 있다. 각 평가 유형 간 상관관계 분석 결과는 Table 4에 제시하였다.

Table 5에 제시한 바와 같이 평가 영역별 평가 유형 간 상관관계를 분석한 결과 동료평가와 자기평가는 과학 지식, 과학 논증, 구성, 표현의 네 영역에서 모두 유의미한 상관관계가 나타나지 않았다. 동료평가와 교수자 평가 간 비교에서는 과학 지식, 과학 논증, 구성 영역에서 유의미한 상관관계가 나타났다. 이에 비해 자기평가와 교수자 평가 간에는 표현 영역에서만 유의미한 상관관계를 보였다. 평가 영역별 상관관계 분석에서 동료평가와 자기평가 간 유의미한 상관관계를 나타내지 않았다는 점은 평가 영역별 루브릭에 대한 검토가 필요함을 시사하고 있다. 선행 연구(Hwang & Hahn, 2020)에서 지적한 바와 같이 평가 영역 내 항목들의 동질성이 부족하다면 평가자 개인의 주관적 인식이 영향을 미칠 가능성이 커진다. 그러므로 항목들을 변경하면서 반복적인 상관관계를 분석

Table 4. Comparison of Spearman Correlation Between Evaluation Types

	동료평가	자기평가	교수자 평가
동료평가		.210	.672**
자기평가			.515**

$\rho$ : \*\* < .01

Table 5. Comparison of correlation between evaluation types by evaluation area

	동료 - 자기	동료 - 교수자	자기 - 교수자
과학 지식	.043	.586**	.323
과학 논증	.189	.894**	.213
구성	.068.	.646**	.095
표현	.215	.165	.833**

$p$ : \*\* < .01

하여 자기평가와 동료평가 간 관계를 명확하게 확인할 필요가 있다. 본 연구에서는 평가자 개인 특성에 의해 나타나는 차이와 평가 루브릭에 기인한 차이를 구분하기 위한 설계가 이루어지지 않았으므로 후속 연구를 통한 보완이 요구된다.

하위 평가영역별로 평가유형 간 상관관계를 분석하였다. Table 3에 제시된 바와 같이 과학 지식 영역에서는 교수자 평가 평균이 가장 높았고, 자기평가 평균이 가장 낮았다. Spearman 상관관계 분석 결과 동료평가와 교수자 평가 간 상관관계수 .586 ( $p < .01$ )으로 높은 상관관계를 나타냈다. 자기평가와 동료평가는 .043으로 상관관계가 나타나지 않았다. 또한 자기평가와 교수자 평가는 .323 ( $p < .05$ )로 낮은 상관관계를 나타냈다. 과학 지식 영역의 구체적인 평가 기준은 과학 지식의 정확성과 관련성이다(Table 2 참조). 동료평가의 경우 교수자 평가와 높은 상관관계를 나타낸 것으로 보아 ‘교과논리 및 논술’의 글쓰기에 대한 동료평가에서 과학 지식의 정확성과 관련성에 대한 평가 기준을 활용하는 것이 타당함을 시사하고 있다. 이에 비해 자기평가는 교수자 평가와의 분석 결과 낮은 상관관계를 보였으므로, 자기평가에서의 평가 기준을 동료평가 및 교수자 평가와 동일하게 활용하는 것이 적합한가에 대한 논의가 필요함을 보여주고 있다. 특히 자기평가가 다른 평가 유형과의 상관관계가 낮게 나타난 것은 예비 생물교사들이 평가 기준을 동일하게 활용하지 못하고 있을 가능성을 나타내는 것이므로 글쓰기 중 과학 지식 영역에 대한 자기평가 경험을 확대할 필요가 있다. 선행 연구(Lee, Lee, & Jang, 2015)에서 제시한 바와 같이 자기평가 등의 학생평가를 활용하기 위해서는 평가 기준에 대한 충분한 학습과 함께 많은 평가 경험을 제공하는 것이 효과적이다. 그러므로 자기평가 평가 기준 설정에 대한 검토와 함께 ‘교과논리 및 논술’ 수업에서 학생 평가 활동을 강화할 필요가 있는 것으로 판단된다.

과학논증 영역에서는 동료평가 평균이 가장 높았고, 자기평가 평균이 가장 낮았다. 상관관계 분석 결과 동료평가와 교수자 평가는 상관관계수 .894 ( $p < .01$ )로 매우 높은 상관관계를 보인 반면 자기평가와 교수자 평

가 간에는 .213으로 상관관계가 성립되지 않았다(Table 5 참조). 이 결과에 기초해 볼 때 논증 글쓰기 평가에서 동료평가의 활용 가능성은 높은 것으로 판단된다. 또한 논증 글쓰기 평가 중 논증 관련 평가 기준을 동료평가에서 활용하는 것은 타당하다고 볼 수 있다. 반면에 과학논증 영역에 대한 자기평가 분석 결과는 자기평가 평가 기준으로 논증에 관한 항목을 활용하기 위해서는 평가 기준 설정에 대한 검토가 필요함을 나타내고 있다. 자기평가 결과가 세 가지 유형의 평가 중 가장 낮은 평균을 나타냈고, 동료평가 및 교수자 평가와 상관관계가 성립되지 않은 것은 예비 생물교사들이 자신의 글쓰기에 대해 상대적으로 낮은 인식을 가지고 있음을 보여주는 것이다. 그러므로 다양한 논증 글쓰기를 읽고, 그에 대해 동료들과 토의하고 평가하는 활동을 확대할 필요가 있을 것이다.

Table 5에 나타난 바와 같이 구성 영역에서는 자기평가 평균이 가장 높았고, 다음으로는 교수자 평가, 동료평가 순으로 나타났다. 동료평가와 교수자 평가 간 상관관계수 .646 ( $p < .01$ )로 높은 상관관계를 보였다. 이에 비해 자기평가와 교수자 평가 간에는 .095로 유의미한 상관관계가 나타나지 않았다. 동료평가 결과가 교수자 평가와 높은 상관관계를 나타낸 것은 동료평가에서 구성 영역에 대한 항목들을 평가 기준으로 활용할 수 있음을 나타내고 있다. 또한 동료평가 평균이 가장 낮은 것으로 볼 때 동료평가가 다른 평가 유형보다 기대 수준이 높음을 시사하고 있다. 이에 비해 자기평가는 구성 영역에서 가장 높은 평균을 보였다. 자기평가는 동료평가 및 교수자 평가와 상관관계를 나타내지 않았는데, 이는 동일한 글에 대한 자기평가가 동료평가 및 교수자 평가와는 다른 양상이라는 것을 나타낸다. 특히 이 결과는 과학 논증 글쓰기 평가에서 자기평가를 활용하는 것이 타당한가에 대한 논의의 필요성을 보여주고 있고, 각 평가 유형에 따른 평가 기준의 설계가 고려되어야 함을 시사하고 있다.

표현 영역에서는 자기평가 평균이 가장 높았고, 다음으로 동료평가, 교수자 평가 순이었다. 상관관계 분석 결과에서는 자기평가와 교수자 평가 간 상관관계수가 .833 ( $p < .01$ )로 매우 높은 상관관계가 나타났다.

반면에 동료평가와 교수자 평가 간에는 .165로 나타나 유의미한 상관관계가 나타나지 않았다(Table 5 참조). 표현 영역에 대한 동료평가 결과는 앞서의 결과와 달리 교수자 평가와 유의미한 상관관계가 나타나지 않았다. 글쓰기에서 표현 영역은 주관적인 인식이 많은 영향을 미치므로(Hwang & Hahn, 2020), 동료평가의 결과에 개인적인 인상이 반영되었을 가능성이 있고, 표현 영역에 대한 동료평가 기대 수준이 다른 평가 유형에 비해 상대적으로 높음을 나타내고 있다. 이에 비해 표현 영역에서의 자기평가는 가장 높은 평균을 보였고, 교수자 평가와의 상관관계가 매우 높게 나타났다. 이 결과는 과학 논증 글쓰기 표현 영역 평가에서 자기평가 활용 가능성이 높음을 보여주는 것이다. 한편 자기평가에서도 표현에 대한 주관적 인상이 반영되었을 가능성이 있는 만큼 다양한 글쓰기에 나타난 표현에 대한 동료들 간 의견 교환 활동을 통해 표현 영역 평가 활동을 확대할 필요가 있다.

평가 기준의 하위 영역에 대한 평가 유형 간 상관관계 분석 결과를 볼 때 ‘교과논리 및 논술’ 수업의 글쓰기 결과물에 대한 교수자 평가와 함께 동료평가와 자기평가를 활용할 때 평가 기준에 대한 세부적인 설계가 필요함을 시사하고 있다. 즉 동료평가와 자기평가에서 적용하는 평가 항목을 동일하게 할 것인가 아니면 서로 다르게 할 것인가에 대한 검토가 필요하다고 본다. 본 연구 결과를 토대로 살펴보면 일례로 동료평가의 평가 기준으로는 과학 지식, 과학 논증, 구성 등을 설계하고, 자기평가에는 과학 지식, 표현 영역을 활용하는 등의 재구성도 검토할 필요가 있다고 판단된다. 교수자 평가에서는 지식, 논증, 구성, 표현 영역을 모두 활용하고 동료평가와 자기평가에서는 관련성이 높은 평가 항목을 활용하는 것도 고려할 수 있을 것이다.

## 2. 서술형 평가 내용 분석

동료평가와 자기평가에서 실시한 서술형 평가 내용을 분석하였다. 네 차례에 걸쳐 이루어진 동료평가의 서술형 총평을 평가 내용별로 분석한 결과 총 519개 항목이 나타났고, 피평가자인 예비 생물교사 1인당 평균 37개의 평가 요소가 제시되었다. 동료평가 내용이 가장 많은 경우는 51개였고, 가장 적은 경우는 25개로 나타났다. 과학 지식, 과학 논증, 구성, 표현 네 영역에 대한 평가 내용을 분석한 결과 과학 논증에 대한 내용이 170개(32.8%)로 가장 많은 비중을 차지하였다. 다음으로는 구성 영역 122개(23.5%), 표현 영역 121개(23.3%)였고, 과학 지식에 대한 평가 내용이 106개(20.4%)로 가장 적었다. 과학 논증에 대한 동료평가 내용이 가장 많은 비중을 차지한 것은 평가 대

상 글쓰기가 논증글쓰기였으므로 동료평가를 할 때 예비 생물교사들이 논증 관련 내용을 중점적으로 살펴본 것으로 추정된다. 과학 논증에 대한 서술형 동료평가 내용을 살펴보면 근거의 타당성을 지적하거나 근거에 대한 보강의 필요성을 지적하는 경우가 많았다. 또한 반증의 적절성과 반증 추가 필요성을 제시하는 경우가 많았다. 구성 영역에서는 문단 나누기에 대한 내용이 가장 많이 제시되었다. 또한 문단 간 연결관계와 글의 통일된 체계 등에 대한 내용이 다수를 차지하였다. 표현 영역에서는 효과적인 자료 활용에 대한 내용이 가장 많았고, 다음으로는 어색한 문장과 접속사의 활용 등에 대한 지적이 많았다. 과학 지식 영역에서는 주제에 맞는 과학 개념을 제시했다는 점이 가장 많았고, 다음으로는 과학 지식에 대한 설명이 적절했다는 내용이 많았다. 이에 비해 글에서 제시하고 있는 과학 지식이 너무 어렵다는 지적도 제시되었다. 4회의 논증글쓰기에 제시된 서술형 동료평가의 예시는 아래에 제시하였다.

GMO에 대한 정확한 내용을 제시하였고, 식량 부족과 영양 결핍 등의 근거를 잘 나타내어 쉽게 읽히는 글이었다. 반증도 적절한 것 같은데, 다만 문단 나누기는 좀 고쳐야 할 것 같다. 너무 긴 문단은 적절하게 나누는 게 좋겠다.

(첫 번째 글쓰기에 대한 예비 생물교사 G의 서술형 동료평가)

인공지능의 법적 지위에 대한 내용이 이 주제에 대한 정확한 근거인지 잘 모르겠다. 인공지능과 관련한 윤리적 문제에 대한 검증된 자료들을 제시하는 게 더 좋을 것 같다.

(두 번째 글쓰기에 대한 예비 교사 N의 서술형 동료평가)

자신의 주장에 대한 근거가 구체적으로 잘 나타난 글이다. 구체적인 통계 자료를 활용한 것도 좋아보인다. 반증이 명확하지 않은 것은 개선해야 할 것 같다.

(세 번째 글쓰기에 대한 예비 교사 F의 서술형 동료평가)

흔히 아는 실험인 ‘파블로프의 개’를 다른 측면으로 본 것이 좋았다. 동물실험에 대한 새로운 지식을 알게 되어서 좋았다. 구체적인 반증 사례를 제시한 것이 효과적이었고, 문단 구성도 잘 이루어졌다.

(네 번째 글쓰기에 대한 예비 교사 J의 서술형 동료평가)

앞의 예시에서 나타난 바와 같이 예비 생물교사들의 서술형 동료평가에서는 논증글쓰기 내용에 대한 구체적인 평가 내용이 나타났다. 이는 선행 연구(Oh, 2009)에서 지적한 형식적인 동료평가와는 다른 양상을 보이는 것이다. 특히 Seo & Jeon(2012)에서는 동료평가에서 나타난 피드백 유형을 단순소감형, 오류지적형, 논평형 등으로 분류하였다. 본 연구에서 제시된 서술형 동료평가의 내용은 오류지적형이 가장 많았다. 이에 비해 단순 소감형은 매우 적었다. 예비 생물교사들의 피드백 내용이 대안을 제시하는 수준까지는 나타나지 않았지만 오류 지적과 그에 대한 논평의 형식을 갖추고 있는 것은 바람직한 양상이라고 볼 수 있다. 본 연구에서는 논증글쓰기의 주제를 예비 생물교사들이 모두 토의를 거쳐 선정하였다. 이 과정에서 글쓰기 주제에 대한 이해를 높일 수 있었고, 동료평가에서도 구체적인 지적 사항이 제시된 것으로 판단된다. 이 결과는 ‘교과논리 및 논술’ 교과 운영 시 글쓰기 주제 선정 과정에 예비 교사들을 참여시키는 것이 동료평가 활용 측면에서 효과적일 수 있음을 시사하고 있다.

자기평가에서의 서술형 평가 내용에 대한 분석 결과 평가 항목은 총 131개였다. 매회 논증글쓰기에 대한 자기평가에서 예비 생물교사들은 평균 2.3개의 평가 내용을 제시하였다. 이 중 가장 많은 비중을 차지하고 있는 것은 논증 영역에 대한 내용으로 전체의 48.1%인 63개였다. 이 결과는 자기평가에서 과학 논증 영역 평균이 2.54점으로 가장 낮게 나타난 것과 연관된 것으로 판단된다. 즉 자신이 쓴 글에 대해 논증에 대한 평가가 가장 낮았으므로 서술형 평가에서도 이에 대한 내용을 언급한 것으로 볼 수 있다. 이외에 표현 영역에 대한 내용이 40개로 30.5%였고, 구성 영역에 대한 내용이 28개로 21.4%로 나타났다. 반면 과학 지식 영역에 대한 서술형 자기평가 내용은 나타나지 않았다. 내용 측면에서는 오류지적형이 대부분이었고, 단순 소감형도 일부 나타났다. 논증글쓰기에 대한 서술형 평가에서 동료평가 보다는 자기평가의 평가 내용이 현저히 적은 것은 자기평가가 내포하고 있는 자기주도적 학습력과 반성적 사고(Dochy *et al.*, 1999)가 실제 수업에서 발현되기 위해서는 반복적인 평가 활동이 이루어져야 함을 보여주는 것으로 판단된다. 예비 생물교사들의 서술형 자기평가 내용의 예시는 아래에 제시하였다.

유전자변형 작물을 재배해야 한다는 내 주장을 뒷받침하는 근거를 다 추가했으면 좋았을 텐데 뭔가 아쉬운 글이다. 글의 형식은 논증글쓰기에 잘 맞는 것 같다.

(첫 번째 글쓰기에 대한 예비 생물교사 B의 서술형 자기평가)

자료 조사를 더 해서 근거를 보강해야 할 것 같다. 문단 구성은 읽기 쉽게 잘 된 것 같다.

(두 번째 글쓰기에 대한 예비 생물교사 E의 서술형 자기평가)

문장에서 자료에 대한 인용 표기가 이루어지지 않았다는 점이 아쉽게 생각된다. 근거에 대한 설명이 너무 길었는데 통계 자료 등을 이용하지 못한 게 부족하다.

(세 번째 글쓰기에 대한 예비 생물교사 C의 서술형 자기평가)

내 주장을 설득할 수 있는 근거를 더 많이 찾아야 할 것 같다. 문장 표현과 자료에 대한 인용은 잘 이루어진 것 같다.

(네 번째 글쓰기에 대한 예비 생물교사 I의 서술형 자기평가)

#### IV. 결론 및 제언

본 연구는 ‘생물교과논리 및 논술’ 수업에서 실시된 논증글쓰기에 대한 동료평가와 자기평가, 교수자 평가 간 상관관계를 분석하고, 이를 토대로 ‘교과논리 및 논술’ 글쓰기 활동 평가에 대한 시사점을 도출하는 것이다. 이를 위해 14명의 예비 생물교사들이 네 차례에 걸쳐 논증글쓰기를 하였고, 이에 대한 동료평가와 자기평가, 교수자 평가가 이루어졌다.

논증글쓰기에 대한 동료평가는 평균 2.70점인데 비해 자기평가 평균은 2.68점, 교수자 평가 평균은 2.71점이었고, 자기평가 평균이 다른 평가 유형에 비해 상대적으로 낮게 나타났다. 하위 평가 영역별 평균은 동료평가에서는 과학지식 측면이 평균 2.84점으로 가장 높았고, 자기평가에서는 구성 영역이 평균 2.84점으로 가장 높았다. 또한 교수자 평가에서는 과학지식 영역의 평균 점수가 2.85점으로 가장 높았다. 표준편차 분석에서는 자기평가에서의 표준편차가 가장 크거나 다른 평가 유형에서 보다 더 넓게 분산된 특징을 보였다. 동료평가, 자기평가, 교수자 평가 모두에서 평가 영역 중 과학지식 영역과 구성 영역이 높은 점수를 보였다.

평가 유형 간 상관관계를 분석한 결과 동료평가와 교수자 평가 간에는 높은 상관관계를 보였고, 자기평가와 교수자 평가 간에는 보통의 상관관계를 나타냈다. 평가 영역별 평가 유형 간 상관관계를 분석한 결과 동료평가와 교수자 평가 간에는 과학 지식, 과학 논증, 구성 영역에서 유의미한 상관관계가 나타났다. 또한 자기평가와 교수자 평가 간에는 표현 영역에서

만 유의미한 상관관계를 보였다. 서술형 동료평가에서는 과학 논증과 관련한 피드백이 가장 많았고, 다음으로는 구성 영역, 표현 영역 순으로 나타났다. 본 연구에서 나타난 평가 유형 간 상관관계는 논증 글쓰기 평가 상황과 글쓰기 주제, 예비 생물교사들 간 상호작용 등에 영향받을 가능성이 높다. 즉 이와 같은 요인들의 영향을 완전히 배제시키지 못하였다는 점은 본 연구의 한계점이라고 판단된다.

서술형으로 이루어진 동료평가와 자기평가에서는 과학 논증 영역에 대한 내용이 가장 많았다. 이에 비해 과학 지식에 대한 내용은 가장 적게 나타나 논증 글쓰기에 대한 예비 생물교사들의 평가 관점이 특정 영역에 치중되어 있음을 보여주었다. 본 연구의 결과를 토대로 ‘교과논리 및 논술’ 수업의 논증글쓰기 평가에 대해 살펴볼 만한 점들을 제시하고자 한다.

첫째, 평가 유형에 따른 평가 기준을 세분화할 필요가 있다. 본 연구의 결과에서 나타난 바와 같이 동료평가와 교수자 평가, 자기평가와 교수자 평가 등 평가 유형 간 높은 상관관계를 보이는 평가 영역은 서로 달랐다. 이는 동료평가를 통해 평가하기에 적합한 영역이 있고, 자기평가 활용이 효과적인 영역이 있음을 의미하는 것이다. 그러므로 동료평가, 자기평가, 교수자 평가의 평가 기준을 일률적으로 적용하는 것 보다는 평가 유형별로 평가 기준을 세분화하는 것이 평가의 정확성과 타당성을 높이는 데 효과적일 수 있다. 특히 앞서 제시한 본 연구의 한계점을 보완하여 평가 상황, 모둠 구성, 사전 평가 연습 등에 대한 설계가 체계적으로 이루어질 필요가 있다.

둘째, 논증 글쓰기에 대한 동료평가와 자기평가 평가 기준 설정 과정에 이 평가의 주체인 학습자가 직접 참여하는 방안을 모색하는 것이 효과적이라고 판단된다. 서술형 동료평가와 자기평가를 분석한 본 연구의 결과에 따르면 과학 논증, 구성 영역에서는 동료평가와 자기평가가 활발하게 이루어졌지만 과학 지식 영역에 대한 평가는 매우 제한적이었다. 이는 평가의 주체로 참여한 예비 생물교사들이 논증글쓰기 평가에서 어떤 측면을 중점적으로 고려하였는가를 나타내는 것이다. 그러므로 평가 기준 설정 단계에서부터 학습자와 교수자 간 논의가 이루어진다면 평가 과정에서 학습자들의 적극적인 참여를 유도할 수 있을 것이다. 특히 ‘교과논리 및 논술’ 수업은 교사양성 교육과정 중 필수 교과목이므로 평가 기준 설정에 예비 교사들이 참여하는 것은 평가 전문성 함양에도 도움이 될 것이다.

셋째, 본 연구에서는 논증글쓰기를 대상으로 한 평가를 실시하였으나 예비 생물교사들의 글쓰기 평가 전문성을 높이기 위해서는 다양한 글쓰기 유형에 대한 동료평가, 자기평가를 실시하는 것이 바람직할 것

이다. ‘교과논리 및 논술’ 과목은 예비 교사들의 글쓰기 능력 함양과 함께 향후 학교 현장에 진출했을 때 글쓰기 지도 역량 강화를 목표로 하고 있다. 그러므로 실제 중등학교 수업에서 활용할 수 있는 다양한 유형의 글쓰기를 어떻게 평가할 것인가를 미리 경험하는 것은 중요하다. 그러므로 ‘교과논리 및 논술’ 수업을 통해 글쓰기 유형별로 동료평가와 자기평가, 교수자 평가를 비교 분석하는 후속 연구가 이루어질 필요가 있다.

넷째, 본 연구에서는 평정척도와 서술형 문항을 이용해 동료평가와 자기평가를 실시하였다. 동료평가인 경우 서술형 평가에 다양한 의견들이 제시되었으나, 자기평가에서는 매우 제한적인 의견만 나타났다. 이와 같은 결과는 동료평가와 자기평가의 형식을 다양화해 구체적이고 다각적인 평가가 이루어질 필요가 있음을 시사하고 있다. 본 연구에서는 총평 형식으로 서술하도록 했기 때문에 구체적인 평가 영역별 의견이 제시되지 못한 한계가 있다. 그러므로 후속 연구에서는 총평 형식이 아닌 평가 영역별 서술형으로 제시하는 등 여러 평가 방법을 적용하고 그에 따른 차이를 분석하는 것이 바람직하다.

## 국 문 요 약

이 연구는 ‘생물교과논리 및 논술’ 수업에서 실시된 논증글쓰기에 대한 동료평가와 자기평가, 교수자 평가 간 상관관계를 분석하는 것이다. 이를 위해 14명의 예비 생물교사들이 네 차례에 걸쳐 논증글쓰기를 하였고, 이 결과물들을 대상으로 동료평가와 자기평가, 교수자 평가가 이루어졌다. 논증글쓰기에 대한 동료평가는 평균 2.70점이었고, 자기평가 평균은 2.68점, 교수자 평가 평균은 2.71점이었 다. 하위 평가 영역별 평균을 살펴보면 동료평가에서는 과학지식이 가장 높았고, 자기평가에서는 구성 영역이 가장 높았다. 또한 교수자 평가에서는 과학지식 영역이 가장 높았다. 표준편차를 살펴본 결과 자기평가의 표준편차가 가장 크게 나타났다. 평가 유형 간 상관관계를 분석한 결과 동료평가와 교수자 평가 간에는 높은 상관관계를 보였고, 자기평가와 교수자 평가 간에는 보통의 상관관계를 나타냈다. 반면에 동료평가와 자기평가 간에는 유의미한 상관관계가 나타나지 않았다. 평가 영역별 상관관계를 분석한 결과에서는 동료평가와 자기평가 간에는 과학 지식, 과학 논증, 구성, 표현의 네 영역에서 모두 유의미한 상관관계가 나타나지 않았다. 또한 동료평가와 교수자 평가 간 분석에서는 과학 지식, 과학 논증, 구성 영역에서 유의미한 상관관계가

나타났다. 또한 자기평가와 교수자 평가 간에는 표현 영역에서만 유의미한 상관관계를 보였다. 이 연구의 결과는 '교과논리 및 논술' 수업에서 동료평가의 적용 가능성이 높음을 보여주는 것이다. 또한 동료평가와 자기평가를 활용하기 위해서는 평가 유형에 따른 평가 기준 세분화가 필요함을 시사하고 있다.

**주제어:** 교과논리 및 논술, 논증글쓰기, 동료평가, 자기평가, 교수자 평가

## References

- Ahn, J. Y. (2008). A case study on the application of peer assessment technique in the e-learning environment. *Journal of Korea Entertainment Industry Association*, 2(1), 27-32.
- Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, 27, 5-31.
- Campbell, J. D., & Fehr, B. (1990). Self-esteem and perceptions of conveyed impression: Is negative affectivity associated with greater realism?. *Journal of Personality and Social Psychology*, 58, 122-123.
- Chang, S. M. (2019). A preliminary study for systematizing the contents of "logic and essay writing" in Korean language arts - text mining on 2009~2018 syllabus. *Korean Writing*, 40, 7-43.
- Choi, M. J., Jun, M. K., & Chae, J. (2015). The status of 「Logic and Writing in Home Economics Education」 course and the students' perception about the course. *Journal of Home Economics Education Research*, 27(1), 51-65.
- Chong, H. (2009). The validity and reliability of college writing assessment. *Korean Writing*, 9, 277-305.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 20(3), 331-350.
- Han, L. H. (2019). A study on the method of writing education using self-feedback. *Ratio et Oratio*, 12(2), 147-182.
- Hwang, H., & Hahn, H. (2020). A comparative study of self, peer, and instructor assessment in a single paragraph writing. *Korean Journal of General Education*, 14(1), 107-130.
- Jang, M. Y. (2021). A study on the design of flipped learning instructional design for the subject of 'Logic and Essay in Early Childhood of Education' using YouTube. *The Journal of Humanities and Social Science*, 12(6), 2861-2876.
- Jo, E., Jung, D., & Nam, J. H. (2021). The impact of negotiation-Based peer and self-assessment activities on science-gifted students' modeling. *Journal of the Korean Chemical Society*, 65(6), 455-467.
- Kim, H. M., & Kang, K. H. (2021). An analysis of science core competencies presented in pre-service biology teachers' SSI (Socioscientific Issues) writing. *Biology Education*, 49(1), 122-132.
- Kim, J. K. (2014). A study on the development of teaching-learning methods for <'Logic and Essay Writing' in the field of Chinese classics education> -Through a teaching case of teacher training course in university. *Journal of Korean Classical Chinese Literature*, 29(1), 381-405.
- Kim, N. M. (2009). A peer review analysis for university writing classes. *The Review of Korean Cultural Studies*, 38, 151-180.
- Kim, S. H. (2008). The operation of writing courses for science and engineering students. *Philosophy and Reality*, 79, 286-297.
- Kim, S. Y. (2015). The effects of preservice biology teachers' argumentation on bioethical issues. *Biology Education*, 43(1), 27-36.
- Kim, S. (2017). A case study of flipped learning in a classroom of logic & essay writing on mathematics. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, 7(12), 345-359.
- Kulkarni, C., Wei K. P., Le Huy, Chia, D., Papadopoulos, K., Cheng, J., Koller, D., & Klemmer, S. R. (2013). Peer and self assessment in massive online classes,

- ACM Transactions on Computer-Human Interaction*, 20(6), 1-31.
- Lee, H. (2018). Considerations on a college writing evaluation. *Hansung Language and Literature*, 39, 103-125.
- Lee, H., & Kim, K. (2018). Logic and writing in home economics education development and evaluation of the course “Creative Thinking & Debate”. *Journal of Research in Curriculum Instruction*, 22(2), 126-138.
- Lee, S. H. (2007). *A search for effective science writing curriculum composition: Focusing on the case of Catholic university*. The Korean Society for Thinking and Communication Conference Proceeding, 37-48.
- Lee, S. J. (2014). Usability of peer- and self-assessment in Primary English writing. *English Language Assessment*, 9, 101-121.
- Lee, S. Y., & Chang, Y. H. (2015). A study on the ‘Logic and Essay’ course in early childhood preservice teacher education programs. *Early Childhood Education & Care*, 10(1), 189-208.
- Lee, Y. B., & Chong, H. M. (2014). A study on aspects and validity of peer feedback in college students’ writing. *Korean Writing*, 20, 299-334.
- Lee, Y. K., & Kang, K. H. (2018). The relationship between pre-service science teachers’ practical teaching competence and the agreement of self-evaluation and peer evaluation. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, 8(8), 477-488.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis*. London: SAGE Publication.
- Ministry of Education and Science Technology (2009). *Manual of Practice for Teacher Qualification Examination*. Seoul: Author.
- Oh, T. (2009). The aspects and types of peer assessment on writing by utilizing internet bulletin board. *Journal of CheongRam Korean Language Education*, 40, 235-262.
- Park, J. Y. (2019). A case study of the lesson for the logic and writing education in science education based on the integration teaching-learning methods of the writing education. *Culture and Convergence*, 41(3), 757-800.
- Sadler, F., & Good, E. (2006). The impact of self and peer grading on student learning. *Educational Assessment*, 11(1), 1-31.
- Seo, Y. J., & Jeon, E. J. (2012). A study of peer feedback opinion validity by types in writing. *Korean Language Education Research*, 44, 369-395.
- Shin, S. (2016). Pre-service teachers’ perception of ‘Logic and Writing in English Education’ curriculum. *The Mirae Journal of English Language and Literature*, 21(1), 455-483.
- Sim, J. (2016). A study of preliminary biology teachers’ scientific inquiry skills and logical thinking ability through the activity of science writing. *Biology Education*, 44(1), 114-126.
- Um, S. W. (2019). A study on the model of Korean writing course using flipped learning. *The Korean Journal of Literacy Research*, 10(1), 61-80.
- Yoo, H. J. (2015). Writing lessons evaluation plans for university students utilizing peer-evaluation. *The Journal of Human Sciences Research*, 46, 127-146.

## 저 자 정 보

강 경 희 (제주대학교 부교수)