

예비교사의 작문 평가 문식성 관련 연구 동향) - 주제범위 문헌고찰의 적용

김혜연²⁾

〈 목 차 〉

1. 서론
2. 이론적 배경
3. 연구 방법
4. 연구 결과
5. 결론

1. 서론

학생의 글을 평가하는 일은 국어과 교사들에게 가장 중요한 과업 중 하나지만, White(1994)가 지적한 바와 같이 작문 평가는 상당히 어렵고 까다로운 과업이다. 국어과 교육과정(교육부, 2015, 2022)은 물론 대다수의 작문학자들도 작문 평가에 관한 한 선다형이나 단답형 문항 등을 통해 간접적으로 학습 성취도를 판단하기보다는 한 편의 완결된 글에 대한 직접 평가(direct assessment)가

1) 이 논문은 2024년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2024S1A5A2A01028017)

2) 동국대학교-서울, 부교수, hyeyoun.kim@dgu.ac.kr

이루어져야 한다는 입장을 취해 왔다. 그러나 학생 글을 직접 평가하는 일은 많은 시간과 노력을 요하는 일이며 전문적 교수내용 지식을 갖추어야 하다 보니, 실제 교육 현장에서 이를 실행하는 일은 쉽지 않다. 그럼에도 불구하고, 많은 수의 교사들이 사범계열 대학의 교육과정을 거치면서도 작문 평가에 대해 제대로 준비하지 못한 채 현장에 투입되고 있다(Crusan, Plakans, & Gebril, 2016; Mertler, 2004; Popham, 2009). 즉, 현직교사들의 작문 평가가 제대로 이루어지지 않은 이유 중에는 작문 평가 자체의 본질적인 어려움도 있겠지만, 작문 평가를 효율적으로 이끌어가는 방법이나 작문 평가 능력 향상을 위한 실천들이 교사 교육을 통해 제대로 이루어지지 않은 탓도 있다.

몇몇 연구들은 이와 관련하여 ‘교사의 평가 문식성’(teacher assessment literacy)이라는 개념을 채택하고 교사 평가의 전반적인 현황 조사 및 (예비)교사를 위한 교육 모형 개발을 시도하였다(Crusan et al., 2016; Wang, Wang, & Huang, 2008; Mertler, 2004; Volante & Fazio, 2007). 평가 문식성이란, 교육 평가에 대한 근본적인 이해를 바탕으로 학생 수행의 과정과 결과, 교수학습 맥락, 학습과 발달에 대한 관점 등을 종합적으로 고려하여 평가에 임할 수 있는 능력을 의미한다(Stiggins, 1991, 1995). 교사가 적절한 평가 능력을 갖추기 위해서는 단지 강의실에서 지식 학습 위주로 평가를 배우는 데 그쳐서는 안 되며, 다양한 맥락을 고려하는 실제적 실천과 평가자로서의 역할에 대한 본질적인 이해가 필요하다고 볼 수 있다. 더불어, 교사 평가 문식성의 함양은 교사되기의 핵심 중 하나라는 점에서 예비교사를 위한 교육과정에서부터 시작되어야 할 것이다(Crusan et al., 2016; Volante & Fazio, 2007). 교사가 평가 문식성을 갖추는 일은 단시일 내에 이루어지기 어려우며, 여러 과정들의 통합적 운영(예. 작문교육과 교육평가)에 따라서 이루어질 필요도 있기 때문에 예비교사의 교육과정에 관련 실천이 포함되어야 효율적이면서도 실질적인 교사 교육이 가능하다.

국내에서도 예비교사의 평가 능력 관련 연구들은 상당히 활성화된 편이다. 그러나 관련 국내 연구들은 대다수가 예비교사의 신뢰성 있는 작문 채점

에 중점을 두었으며(박영민, 2012, 2021; 박영민·최숙기, 2009; 박찬홍, 2019; 백유진·박영민, 2023), 신뢰성 이외의 측면에 대해서는 국외에 비해 연구가 매우 부족하다. 또한 예비교사의 평가 능력 신장을 위한 본격적인 연구 수행도 거의 없었다. 즉, 작문 평가와 관련하여 다양한 국면과 맥락들이 존재하나, 많은 연구들이 작문 채점의 신뢰성과 정확성에 집중되어 있으며 그 밖의 부문들이 고루 탐구되지 못하였다는 점에서, 예비교사의 작문 평가 문식성을 균형 있게 향상시키기 위한 교육적 근거를 마련하기에는 부족하다. 그러나 Weigle(2002: 2)에서도 지적하였듯이, 교사의 작문 평가 실천은 학생의 성취도를 객관적으로 판단하는 일을 넘어서서 작문 평가를 실시하는 근본적인 이유와 목적이 무엇인지를 염두에 두는 방식으로 나아가야 한다.

이상의 논의들을 종합해 볼 때, 예비교사의 작문 평가 문식성과 관련하여 균형 있는 후속 연구를 독려하기 위하여, 그리고 교사교육 실천을 위한 근거를 마련하는 차원에서, 관련 연구들을 체계적으로 고찰하고 어떠한 연구들이 어느 정도의 성과를 냈는지 포괄적으로 정리할 필요가 있다. 본 연구에서는 주제범위 문헌고찰(scoping review)의 방법론을 도입하여 관련 연구들을 엄정한 기준에 의거하여 수집하고 최종 선정된 문헌들을 대상으로 예비교사의 작문 평가 문식성 관련 연구 동향을 분석하여 후속 연구나 교육적 적용의 기반으로 삼고자 한다. 주제범위 문헌고찰은 체계적 문헌고찰과 함께 특정 주제와 관련된 연구 동향을 정리하고 분석할 때 활용하는 연구 방법론으로서, 연구들 간의 관계를 맺고 전반적인 경향성을 파악하고자 할 때 유용하다(Arksey & O'Malley, 2005). 특히 수집 문헌의 범위를 국내로만 제한하지 않고 국외 연구로까지 확장함으로써, 국내외 간 연구 동향 비교도 가능케 하고자 한다. 최근 들어 (예비)교사의 작문 평가 관련 연구들이 국내외 막론하고 부쩍 활성화된 데 비해 연구 성과를 정리하는 시도가 거의 없었으므로, 해당 분야의 연구 동향의 중간 정리의 성격으로서도 이 연구의 의미를 찾을 수 있다.

2. 이론적 배경

1) 예비교사의 평가 문식성

교사들은 종종 자신에게 주어진 시간의 50%까지 평가에 할애하기도 한다는 점에서, 평가는 교사 활동의 중추라 할 수 있다(Plake, 1993: 21). 시간의 문제를 넘어서서, 교실 평가는 교사의 다양한 역량에 영향을 미치는 중대한 일이기도 하다. 학습자의 수준을 잘 파악하는 적절한 평가를 수행할 수 있다면, 이는 다시 교사의 지도와 학습자의 동기, 성취 수준의 향상을 불러온다(Brookhart, 2011).

그러한 의미에서 여러 연구들이 교사의 평가 능력 신장이 중요하다는 점을 크게 강조하여 왔다. 특히 Stiggins(1991) 이래로, 교사의 평가 능력과 관련하여 평가 문식성(assessment literacy)의 개념을 도입한 학자들이 많아졌다. 오늘날 문식성은 전통적 의미의 이해하고 표현하는 능력에만 국한되는 것이 아니라, 이해 및 표현 활동에 대한 태도와 정서, 그리고 이를 둘러싼 맥락과 관점 및 사회적 실천을 모두 포괄하는 확장된 개념으로 지칭된다(정혜승, 2008). 즉, 교사의 평가 능력 신장에 대한 관심이 특정 기능에만 국한될 우려도 있다는 점에서, 평가 문식성으로 개념화하는 학자들은 교사의 기본적인 평가 능력은 물론, 평가에 대한 지식, 관점 및 성찰, 다양한 맥락에 대한 고려 및 사회적 실천의 의미를 모두 포괄하려는 의도를 담고 있다(Crusan et al., 2016; Malone, 2013; Mertler, 2004; Popham, 2009; Siegel & Wissehr, 2011; Stiggins, 1991). 예컨대, Popham(2009)은 상당수의 현직교사들이 교육평가의 근원적 의미를 잘 모르거나 망각하고 있음을 지적하면서, 평가 문식성은 점수를 잘 매기는 것 이상의 의미를 지녀야 하며 학습자의 학습 성과와 교육의 질을 향상시킨다는 평가의 근원적 목표와 밀접하게 관련되어야 함을 역설하였다.

이렇듯 교사의 평가 문식성 관련 연구들이 활성화됨과 동시에, 평가 문식성

신장이 예비교사를 위한 교육 프로그램의 일환으로 본격적으로 이루어질 필요가 있다는 주장 역시 강화되었으나, 실증적 뒷받침이 될 만한 연구는 많지 않았다. Mertler(2004)는 예비교사와 현직교사들을 대상으로 교실 평가 문식성 검사(classroom assessment literacy inventory)를 실시한 결과, 적절한 평가 도구 선택, 채점 및 결과 해석, 결과를 통한 판정, 비윤리적이거나 불합리한 작문 판별 등의 항목에서 두 그룹 간 유의미한 차이를 확인하였다. 아울러 이 연구는 예비교사의 평가 문식성을 신장시키기 위한 별도의 노력이 필요함을 일깨웠다는 점에서도 의의를 지닌다. Volante와 Fazio(2007)는 예비교사의 평가 문식성을 향상시켜야 할 필요성에 대해 좀 더 구체적으로 알아보기 위하여, 평가의 주요 목적, 다양한 평가 도구의 활용, 훈련의 필요성, 평가 문식성 향상 방법 등의 다양한 항목별로 설문 조사를 실시하였다. 이들에 따르면, 상당수의 예비교사들이 평가에 대해 결과 중심으로 생각하는 경향이 있었으며 형성평가의 목적의식은 잘 고려하지 않았다. 또한 이들 중 대다수가 구체적인 교실 내 평가 실천에 대한 별도의 강의 개설이 필요하다고 응답하였다. Wang 외(2008)는 평가 문식성에 대한 관심이 교수내용 지식(Pedagogic Content Knowledge, 이하 PCK)의 전형적이고 실질적인 실천이 될 수 있음을 강조하면서, 예비교사의 평가 문식성 향상을 도모하기 위한 웹 기반 프로그램을 개발하기도 하였다. Shulman은 일찍이 교수내용 지식 개념을 강조하면서, 교사에게는 “내용과 교육이 결합된 담당 과목 고유의 지식”(Shulman, 1987:8)을 전문적으로 이해하는 것이 매우 중요하다고 보았다. 평가 문식성 개념은 “평가에 대한 지식과 신념”을 포함하고 있기 때문에 전형적인 교수내용 지식의 일환으로 볼 수 있다(Wang et al., 2008:449).

그러나 교과 일반에 걸친 교사 평가 문식성만으로 교수내용 지식의 요건을 갖추었다고 보기는 어려울 듯하다. Shulman(1986, 1987)은 ‘내용 없는 기술’(content-free skill)과 ‘순수 내용 지식’(mere content knowledge)의 양극단이 모두 교육적으로 바람직하지 않으며 교과별 내용 지식에 따라 달라질 수 있

는 교과 고유의 교수법이 필요하다고 보았다. 그러므로 일반적 차원에서 ‘교사 평가 문식성’으로 지칭되는 것을 교수내용 지식으로 보기는 어려우며, 각 교과 영역 내의 예비교사 평가 문식성 관련 연구가 더욱 의미를 지닐 것으로 여겨진다.

이러한 차원에서, Siegel과 Wissehr(2011), 그리고 Crusan 외(2016) 등은 교과목 고유의 (예비)교사 평가 문식성의 문제를 다루었다는 점에서 의미를 지닌다. Siegel과 Wissehr(2011)는 과학과 예비교사들을 대상으로 질적 연구를 수행하여 평가 도구, 원칙, 목적 등과 관련된 전반적인 과학과 예비교사의 평가 문식성을 탐구하였다. 이들은 질적 분석을 통해 흥미로운 결과를 제시하였는데, 연구에 참여한 과학과 예비교사들은 다양한 평가 방식에 대해 배웠으며 적절한 평가관을 인지하였음에도 불구하고 실제 평가 수행에서는 그들의 평가관을 반영하는 평가를 수행하지 못하였으며 새로운 평가 방식 대신 전통적인 평가 방식으로 회귀하는 경향이 있었다. 즉, 기존의 예비교사 교육 프로그램이 평가 관련 요소들을 포함하고 있었어도 실제로 그 효과가 나타나는 못하였음을 확인할 수 있다. Crusan 외(2016)는 제2언어 현직교사들을 대상으로 설문 조사를 실시하여, 이들이 교사 교육을 받았을 때 작문 평가를 어떤 방식으로 배웠는지, 작문 평가와 관련된 신념이 무엇인지 등을 알아보았다. 또한 이들은 26%의 교사들이 거의 혹은 전혀 작문 평가를 배우지 않은 채 현장에 임하고 있는 것으로 나타났다는 점에서 문제를 제기하였으며, 이 조사 내용을 기반으로 예비교사 교육을 위한 제언을 제공하였다.

(예비)교사의 평가 문식성 관련 연구들 중에는 일반적인 작문 평가 능력 이외에도 작문 평가에 대한 관점과 철학, 평가 실천에 대한 인식 등을 좀 더 집중적으로 다룬 연구들도 있었다. 대표적으로, Lam(2019)의 연구에서는 경력이 많은 현직교사들도 대안적 평가에 대해 긍정적 반응을 보이면서도 평가의 본질적 목적과 관련하여 심도 깊은 인식을 지니지는 못한 것으로 파악하였다. 교사들은 대체로 학습의 평가(assessment of learning)나 학습을 위한 평

가(assessment for learning)에 대해서는 부분적으로나마 잘 인식하고 있었으나, 학습으로서의 평가(assessment as learning)에 대해서는 제대로 이해하지 못하고 있었다. 또한 Malone(2013) 역시 평가 문식성의 본질적인 함양이 중요함을 강조하면서, 제2언어 교사와 대규모 언어시험의 공인 평가자들을 대상으로 질적 연구를 수행하였다. 이 연구에서는 교육 현장에서 언어 시험에 대한 본질적인 이해가 제대로 되지 않은 채 해당 결과를 활용하고 있다고 문제를 제기하면서, 평가자들이 필수적이라고 여기는 기술적 정보를 포함하는 일과 교사들의 실제적인 요구를 충족시키는 것 사이의 간극을 확인하였다. 김혜연(2015)에서는 예비교사의 글쓰기 요구에 대한 인식을 분석하여 이들이 교사와 학생의 모순된 정체성 속에서 작문 지도와 평가에 대한 막연한 부담과 책임감을 함께 느끼고 있음을 확인하기도 하였다. 즉, 평가에 대한 관점이나 인식과 관련하여 공통적으로 예비교사 대상 교육을 강화할 필요성이 제기되고 있다.

2) 예비교사의 작문 채점과 피드백 훈련

작문 채점과 작문 피드백은 작문 평가의 양대 축이라고 할 수 있다. 그러므로 앞서 평가 문식성 차원의 논의에 집중하였던 연구들 이외에도, 예비교사 대상 작문 채점이나 피드백 관련 연구들은 모두 평가 문식성의 일환으로서 중요하게 다룰 필요가 있다. 국내외를 막론하고 교사의 작문 채점과 피드백에 대해서는 수많은 연구자들이 오랜 기간에 걸쳐 탐구하여 왔다(김혜연, 2016). 최근 들어, 예비교사의 채점과 피드백 관련 연구들도 많이 활성화 되었으며, 특히 예비교사의 채점과 피드백 능력을 향상시킬 수 있는 방법을 제안하거나 교육 방법의 효과를 검증하는 연구들도 나타났다.

여러 연구들에서 예비교사 대상 작문 채점 혹은 피드백 수행을 별도로 훈련할 필요가 있음을 지적하면서, 다양한 방안을 제시하고 그 효과를 검증하

기도 하였다. 먼저 예비교사 대상 채점 교육과 관련하여, Dempsey, Pytlik Zillig, Bruning(2009)은 예비교사들의 쓰기 평가 능력과 효능감을 향상시키기 위한 도구를 개발하여 실제 학생들의 글 채점 실습을 실시하였는데, 이를 통해 채점 능력과 평가에 대한 효능감 모두 유의미한 향상을 보였다. Youngs와 Bird(2010)의 연구에서는 채점 루브릭과 전형적 사례에 대한 분석 과정을 포함하는 예비교사 대상 채점 교육 프로그램을 실시하여 채점 신뢰도 측면의 효과를 입증하였다.

한편, 예비교사 대상 피드백 훈련에 대한 연구들 역시 여럿 수행되었다. Ferris(2007)는 제2언어 예비 작문교사들을 대상으로 한 세미나 강좌에서 작문 피드백 훈련을 실시하였으며, 그 과정 및 절차, 의의를 논하였다. 또한 Junqueira와 Payant(2015)는 제2언어 초보 교사의 글 교정 피드백의 사례 탐구를 통해, 예비교사의 믿음과 실제 피드백 수행이 불일치함을 밝혔으며 이로써 별도의 피드백 훈련이 필요함을 밝혔다. Peck과 Kavanagh(2024)는 체계적 문헌고찰을 적용하여 예비교사의 피드백 수행과 관련된 최근 10년간 연구물들을 수집하고 이로부터 교사교육 실천을 위한 시사점을 도출하였다. 이들에 따르면, 피드백 관련 최신 담론이 이론적으로 활발히 다루어진 데 비해 실제로 예비교사들이 그러한 최신 피드백을 실행하도록 지원하는 방법에 대한 설명은 잘 다루어지지 않았음을 밝혔다.

여러 연구들에서 예비교사의 작문 채점 훈련, 혹은 피드백 훈련만 단독으로 다루기도 하였으나, 몇몇 연구들은 작문 채점과 피드백을 함께 포함하는 교육 프로그램에 대해 탐구하였다. 대표적으로 DelleBovi(2012)는 중등 예비교사들을 대상으로 총체적 채점의 다양한 측면에 대한 이론과 실제 적용을 학습하게 한 후 채점 루브릭 설계 및 피드백 실천으로 이어지는 일련의 프로그램을 구성하고 해당 프로그램에 대한 예비교사와 교사교육 담당자의 반응을 질적으로 분석하였다.

국내에서는 예비교사의 채점이나 피드백 교육과 관련된 연구가 양상에 대

한 분석에 비해 매우 적은 편이었으며, 몇몇 관련 연구들도 모두 채점 훈련에만 관련된 것들이었다. 박영민(2021)의 연구에서는 예비교사가 실제 학생 글을 채점하는 연습을 수행할 수 있는 프로그램을 개발하고 기술적 실천을 위한 방안을 제시한 바 있다. 이윤빈(2015) 역시 예비교사의 작문 채점 능력 신장을 목적으로 작문 평가 워크숍을 개발하고 해당 워크숍의 효과를 검증하기도 하였다.

3. 연구 방법

1) 연구 설계

이 연구는 예비교사의 평가 문식성 관련 국내외 연구들을 주제범위 문헌고찰의 방법론에 의거하여 수집하고 질적 내용분석에 의해 해당 문헌들을 분류함으로써 전반적인 연구 동향을 파악하는 것을 목적으로 한다. 주제범위 문헌고찰(scoping review)은 해당 연구 분야를 뒷받침하는 핵심 개념과 주요 출처, 사용 가능한 근거 유형을 빠르게 연결 짓는 것을 목표로 하며, 해당 분야에 대한 포괄적인 검토에 적합하다(Arksey & O'Malley, 2005). 특정 주제와 관련하여 엄정한 기준으로 기존 연구물을 수집하고 이를 바탕으로 해당 주제의 연구 동향을 파악하는 것을 목적으로 한다는 점에서 체계적 문헌고찰(systematic review)과도 매우 유사하다. 그러나 주제범위 문헌고찰은 좀 더 넓은 범위의 주제와 관련하여 전반적인 동향 파악을 목적으로 질적 접근을 취하는 경향이 높은 데 반해, 체계적 문헌고찰은 더욱 초점화된 주제를 바탕으로 데이터의 추출 및 합성을 통해 양적 분석을 실시하는 데 더 적합할 수 있다(서현주·김수영, 2018).³⁾ 본 연구에서 예비교사의 평가 문식성 연구 동향을 살펴보고자 하

3) 국어교육 분야에서도 여러 연구들에서 이 두 가지 문헌고찰 방법론에 의거하여 특

는 궁극적인 목적은 연구들 사이의 간극이나 비어있는 지점 등을 확인하여 후속 연구를 위한 출발점을 제공하는 데 있다는 점에서 주제범위 문헌고찰의 방법론 적용이 필요하다고 판단되었다.

2) 연구 절차

주제범위 문헌고찰은 연구 질문 도출, 연구물 검색 및 문헌 선정, 자료 기입, 결과의 수집·요약·보고, 자문의 절차를 거칠 것을 권고한다(서현주·김수영, 2018; Arksey & O'Malley, 2005). 필수 절차와 선택 가능한 단계를 모두 고려한 본 연구의 절차는 다음과 같다.

(1) 연구 질문 설정

먼저, 주제범위 문헌고찰을 실시하는 궁극적인 목적 및 연구 범위와 관련하여 연구 질문을 다음과 같이 도출하였다. 첫째, 예비교사의 평가 문식성 관련 연구들의 발행연도와 연구 영역 및 대상에 있어서 어떠한 분포를 보이는가? 둘째, 예비교사의 평가 문식성 관련 연구들의 세부 주제와 연구 방법에 있어서 어떠한 분포를 보이는가? 셋째, 예비교사의 평가 문식성 관련 연구들의 국내외 분포에는 어떠한 공통점과 차이점이 나타나는가?

정 분야 혹은 주제와 관련된 연구 동향을 파악하고자 하였다. 교과서 선정 기준 연구 동향(김은성·김호정·남가영·박재현, 2014)을 필두로 하여, 질문하며 읽기, 텍스트 난도, 읽기 부진 등 독서 영역의 주제뿐만 아니라(권이은·이수진·송정윤, 2021; 김해인, 2016; 이수진, 2020), 빅데이터나 신경과학의 교육적 활용을 위한 문법 분야의 탐구(강지영·오지은, 2021; 김정은·강지영, 2023), 그리고 협동 작문, 작문 유형화, 대학 계열별 글쓰기 등 작문 영역 동향 분석에 이르기까지(김혜연, 2017, 2022; 안상희, 2020), 체계적 문헌고찰과 주제범위 문헌고찰을 활용하여 특정 주제의 연구 동향을 살펴본 연구들이 꽤 있었다.

(2) 문헌 검색 및 선정

다음으로, 연구 질문 해결에 적합한 문헌들을 엄정하고 객관적으로 선정하기 위하여, 일정한 기준을 마련하여 연구물을 검색하고 여러 절차와 단계를 거쳐서 해당되는 연구들을 선택하였다. 주제범위 문헌고찰에서는 체계적 문헌고찰과 마찬가지로 검색 엔진의 선택 및 검색어 기준을 마련하여 일차적인 검색을 통해 연구물들을 수집한 후, 포함 및 제외 기준을 적용하여 제목, 키워드, 초록, 논문 원문 등을 단계에 따라 검토하면서 적합한 연구물로 범위를 좁혀나가야 한다.

〈표 1〉 문헌 선정 절차

단계	국내 연구	국외 연구
검색 대상 데이터베이스 선정	한국학술정보(KISS) 누리미디어(DBpia)	ScienceDirect, Taylor & Francis
검색어와 수식 적용	예비교사, 작문, 평가를 모두 포함하도록 수식 지정	preservice teacher, writing, assessment를 모두 포함하도록 수식 지정
포함 및 제외 기준 검토 및 적용	● 1차 검토는 연구 분야 및 출판물 등의 검색 범위를 좁히거나, 제목, 핵심어, 초록에 대한 검토를 통해 관련 문헌을 선정 ● 2차 검토는 원문 검토를 통해 최종 선정 문헌을 결정	

본 연구에서는 국내 연구의 경우 한국학술정보의 KISS와 누리미디어의 DBpia를, 국외 연구의 경우 Elsevier의 ScienceDirect와 Taylor & Francis를 검색 데이터베이스로 선정한 후 ‘예비교사’와 ‘작문’, ‘평가’를 모두 포함하도록 검색어를 조합하여 검색을 실시하였다.⁴⁾ 이에 따른 최초 검색 결과

- 4) ‘작문 평가 문식성’은 본 연구의 분석 기준이자 추구하는 방향에 해당하므로, 문헌 검색은 작문 평가로 넓게 잡되 문헌 2차 검토에서 해당 방향성이 고려되도록 기준을 잡았다. 또한 향후 수집된 문헌에 대한 질적 내용분석 시 ‘작문 평가 문식성’을 고려하도록 기준을 설정하였다.

는 국내의 경우 총 176건, 국외의 경우 총 1,842건으로 집계되었다(검색 일시, 2024년 7월 12일).

최초 검색된 문헌들은 검색어만을 통해 기계적으로 수집된 결과를 모두 보여주므로 여러 기준들을 적용하여 검색 범위를 좁히거나 문헌들을 검토하여 적합한 수집 대상 연구인지를 직접 확인하는 절차를 거쳐야 한다. 이를 위하여 포함 및 제외 기준을 마련하고 그에 따른 검토를 거쳐 최종 분석 대상 문헌들을 결정하였다. 체계적 문헌고찰에서는 포함 및 제외 기준도 사전에 결정하는 경향이 강하지만, 주제범위 문헌고찰에서는 연구물 검토 과정에서 귀납적으로 정해나가기도 한다(서현주·김수영, 2018: 17). 이와 함께, 주요 연구물의 참고문헌 검토를 통해 적합한 연구물이 추가되기도 하는데, 이 역시 포함 및 제외 기준을 동일하게 적용하여야 한다. 본 연구에서는 연역적 방식과 귀납적 방식을 혼용하여 <표 2>와 같이 포함 및 제외 기준을 적용하였다.

〈표 2〉 포함 및 제외 기준

단계	기준	내용
중복 제거	중복 수집된 문헌 제거	데이터베이스 간 중복, 참고문헌 검토를 통한 추가 수집에서의 중복 등을 모두 제거
1차 검토	포함 기준	제목·핵심어·초록에 검색어가 포함된 경우(국외), 학문 분야(교육학, 언어학), 동료심사제 학술지 논문
	제외 기준	학위 논문, 비동료심사제 학술지 논문, 리뷰, 기사, 학술대회 발표 초록
2차 검토	포함 기준	예비교사의 작문 평가 양상이나 평가에 대한 인식, 수행 평가를 비롯하여 교수학습과 평가를 모두 포함하는 경우, 언어 과목이 아니어도 작문이 포함된 경우, 현직 교사 대상이어도 작문 평가 교육이 주제인 경우, 제2언어 작문 평가를 다루는 경우
	제외 기준	연구 대상만 예비교사이고 교사교육 관점 연구가 아닌 경우, 예비교사 대상이어도 교수학습만 다루는 경우, 예비교사의 동료 평가, 학생의 예비교사 평가, 작문과 관련 없는 예비교사 대상 연구, 현직 교사의 평가 양상, 평가 일반에 대한 논의, 예비교사의 수업 평가

1차 검토는 교육학 및 작문 분야의 동료심사제 학술지 논문 여부를 검색 기준의 재설정 등을 통해 범위를 좁히는 방식으로 진행되었다. 2차 검토에서는 연구 질문에서 다루고자 하는 범위를 엄정하게 포함하는 연구물 선정에 위하여, 제목·핵심어·초록은 물론 원문에 대한 검토까지 거쳐서 최종 연구 대상 문헌을 확정하게 되었다.⁵⁾

이때, 작문 평가의 주체는 예비교사로 한정하였으므로 현직교사의 작문 평가는 수집 대상에서 제외하였다. 그러나 예비교사와 현직교사의 비교, 예비교사의 작문 평가에 대한 현직교사의 견해, 현직교사의 작문 평가 교육 등은 모두 예비교사의 작문 평가를 직·간접적으로 포함하고 있으므로 모두 수집 대상에 포함시켰다. 예비교사의 작문 평가를 부분적으로 포함하는 연구들도 모두 대상에 포함되었으나, 작문 교수학습만 다루는 경우는 제외되었다. 또한 과학이나 사회 등 국어 관련 교과목이 아니어도 ‘작문’에 대한 평가를 다루고 있다면 대상에 포함되었으며, 제1언어와 제2언어 대상도 모두 포함되었다. 그밖에도, 예비교사의 동료 평가, 예비교사에 대한 학생의 평가, 예비교사의 수업 평가 등은 본 연구 질문과 관련이 없으므로 수집 대상에서 제외되었다. 이상의 포함 및 제외 기준에 따라 중복 제거, 1차 검토, 2차 검토의 과정을 거쳐 최종 선정된 수집 문헌은 국내 24건, 국외 31건으로 확정되었다.

(3) 자료 기입 및 결과 보고

자료 기입의 과정에서는 연구 질문의 구체적인 답안을 찾기 위하여 각 연구

5) 국내보다 국외의 연구물 검색 건수가 월등하게 많았기에, 국외 연구물의 필터링 과정에서 제목·핵심어·초록에서만 검색어 검색을 실시하는 것으로 한 단계 더 제약을 두었다. 국내의 경우에는 숫자가 많지 않아서 그러한 제약을 두지 않은 채 검색 최초 결과에 포함된 문헌들을 모두 검토하였다. 그러한 이유로, 국외의 경우 1차 검토를 통해 142건으로 범위를 크게 줄일 수 있었으나, 국내의 경우 121건으로 좁히는 데 그쳤다.

물의 특징을 여러 기준에 따라 정리하였다. 본 연구에서 활용한 자료 기입의 기준은 <표 3>과 같다.

주제범위 문헌고찰에서는 연구물의 속성을 판단할 때 질적 관점에서 재해석할 것을 권장한다는 점을 고려하여(Arksey & O'Malley, 2005), 본 연구의 자료 기입에서는 질적 내용분석(qualitative content analysis)에 입각한 분석 기준을 도입하였다. 질적 내용분석에서는 코딩의 기준을 귀납적으로 도출할 수 있지만, 일정한 코딩 단위에 따라 상호배제와 전체포괄(Mutually Exclusive and Collectively Exhaustive, 이하 MECE)의 원칙을 따를 것을 권장한다(Krippendorff, 2013; Schreier, 2012). 이에 따라, 본 연구에서는 하나의 문헌을 코딩 단위로 삼아 각 문헌마다 <표 3>의 기준별로 하나씩의 코드 내용을 기입하였다. 이때 기준별로 하나씩만 적용되었으며, 어떤 문헌도 하나의 코드를 배정받지 못한 기준이 존재하지 않도록 하였다.

<표 3> 문헌의 자료 기입 기준

분류	기준		세부 항목
연구 일반	발행연도		최근 10년 이내(2015-2024년), 최근 10-20년(2005-2014년), 최근 20-30년(1995-2004년), 최근 30년 이전(1994년 이전)
	연구 분야		제1언어 대상, 제2언어 대상, 구별 없음
	연구 방법		양적 연구, 질적 연구, 혼합 연구, 도구 개발, 이론 연구
	연구 목적		양상 확인 및 문제 발견, 원인 파악, 문제 발견 및 원인 파악, 해결책 제시(근거 기반), 해결책 제시(근거 없음)
	연구 대상	학교급	초등교육, 중등교육, 초·중등교육, 대학교육, 구별 없음
연구 주제		자료 수집·분석 대상	예비교사, 예비교사와 현직교사의 비교, 현직교사의 의견, 선행연구, 없음
		채점	채점 양상, 채점 교육, 채점에 대한 인식, 채점 양상 및 인식
		피드백	피드백 양상, 피드백 교육, 피드백에 대한 인식, 피드백 양상 및 인식, 피드백 교육 및 인식
	채점과 피드백		채점과 피드백 양상, 채점과 피드백 교육, 채점과 피드백 교육 및 인식
	평가 일반		평가 계획 및 설계, 평가 능력이나 자질, 평가에 대한 인식 및 태도, 평가 능력과 그에 대한 인식, 평가자 협의

수집된 문헌의 자료 기입은 우선 ‘연구 일반’ 차원과 ‘연구 주제’ 차원으로 나누어 볼 수 있다. 연구 일반 차원에서는 발행연도, 연구 분야, 연구 방법, 연구 목적, 연구 대상 등 일반적인 주제범위 문헌고찰의 자료 기입 기준들에서 흔히 볼 수 있는 기본적인 정보들을 위주로 각 연구물의 성격을 코딩하였다.

한편, 연구 주제 차원의 기준들은 본 연구의 핵심 주제인 ‘평가 문식성’ 동향과 더욱 밀접하게 관련된다. 최종 수집된 문헌들의 일차적 검토를 통해, 예비교사의 작문 평가 문식성과 관련된 연구들은 기본적으로 채점, 피드백, 평가 일반에 대한 연구들로 나뉠 수 있음을 파악하게 되었다. 앞선 기준들과 마찬가지로 연구 주제 차원에서도 MECE 원칙에 따라, 작문 채점, 작문 피드백, 작문 채점과 피드백, 작문 평가 일반 중 하나의 영역에만 속할 수 있도록 코딩 기준을 설정하였으며, 해당 기준을 다시 상세화하여 양상 확인, 교육적 적용, 인식이나 태도 등의 세부 항목 중에서 다시 선택할 수 있도록 구성하였다.

4. 연구 결과

1) 분석 문헌의 일반적 동향

이 연구는 주제범위 문헌고찰의 방법으로 예비교사의 평가 문식성 관련 연구들의 국내외 동향을 확인하고 이를 통해 향후 후속 연구의 다양한 가능성을 탐색하는 것을 목적으로 한다. 이를 위하여, 앞서 언급한 바와 같이 연구의 일반적인 기준에 따른 동향과 예비교사 평가 문식성의 세부 연구 주제 관련 동향을 구별하여 질적 내용분석을 실시하였다.

〈표 4〉는 연구 일반의 동향 중에서 발행연도, 연구 분야, 연구 대상 관련 국내와 국외 동향을 함께 보여주고 있다. 검색 방법에 따른 제약을 고려하더라도, 그간 예비교사의 작문 평가 문식성 관련 국내 연구들이 국외 못지않게 매

우 활발히 이루어졌음을 짐작할 수 있다.

먼저, 문헌의 발행연도 측면에서는 국내외 모두 최근 20년 사이에 대부분의 연구들이 수행되었음을 확인할 수 있다. 분석 대상 문헌들 중에서 최근 10년 이내에 수행된 연구들은 국내 14건(58%), 국외 14건(45%)이었으며, 최근 10-20년 사이에 수행된 연구들은 국내 10건(42%), 국외 11건(35%)이었다. 그러나 최근 20년 이전에는 국내 연구는 전무하였던 데 반해, 국제적으로는 1990년대에 5건의 연구가 수행되었는데 (예비)교사의 평가 문식성에 관한 논의가 나타나기 시작한 시기와 일치한다(Stiggins, 1991; Plake, 1993). 즉, 예비교사의 작문 평가 관련 논의들은 1990년대에 평가 문식성 관점의 도입 이후 급격히 확장되었다고 볼 수 있다. 국내 연구는 국외에 비해 약 10여년 정도 뒤늦게 시작되었으나, 최근 20년 사이 동향을 놓고 볼 때, 국내 연구가 매우 활성화 되었다는 점을 확인할 수 있다.

〈표 4〉 분석 문헌의 발행연도, 연구 분야 및 대상 관련 동향

기준		국내	국외
발행 연도	2015-2024년(최근 10년 이내)	14건	14건
	2005-2014년(최근 10-20년 사이)	10건	11건
	1995-2004년(최근 20-30년 사이)	-	4건
	1994년 이전	-	2건
연구 분야	제1언어 분야	22건	14건
	제2언어 분야	2건	10건
	구별 없거나 확인 불가	-	7건
연구 대상	학교급	초등교육	2건
		중등교육	8건
		초·중등교육	-
		대학교육	3건
		구별 없거나 확인 불가	10건
	탐구 대상	예비교사 대상	19건
		예비교사와 현직교사 비교	2건
		현직교사의 견해 조사	4건
		선행 연구	3건
		실증적 대상 없음	3건

연구 분야와 관련하여, 국내외 모두 제2언어보다는 제1언어 분야에서 예비교사의 평가 문식성 관련 연구들이 더욱 활발히 수행되었다. 다만, 국내에서는 대다수인 22건(92%)의 연구들이 제1언어, 즉 국어교육 분야에서 수행되었던 데 반해, 국외에서는 14건(45%)의 연구만이 제1언어 분야에서 수행되었고 10건(32%)은 제2언어 분야에서 수행되었다. 그리고 전체 연구 중 7건(23%)은 제1언어와 제2언어의 구별 없이 수행되거나 확인이 불가하였는데, 이론이나 도구 개발을 목적으로 삼은 연구(Ferris, 2007; Hodges et al., 2019)거나, 여러 교과와 예비교사를 대상으로 삼은 연구(Grainger, Heck, & Carey, 2018) 등이 이에 해당하였다. 국내에 비해 유독 국외 연구들에서 제1/제2언어 여부가 구별되지 않은 경우가 많았던 이유는, 국내의 경우 모어 화자 대상 국어교육과 제2언어로서의 한국어교육이 대학 편제상으로도 비교적 명료하게 구별되는 데 반해, 국외 연구 분야에서는 다양한 편제로 인해 그 구별이 명료하지 않거나 통합적으로 교육하는 경우들도 많기 때문인 것으로 추측된다.

다음으로, 연구 대상은 다시 학교급과 자료 수집 및 분석의 대상으로 나누어 분석되었다. 학교급 차원에서도 국내외 간 큰 차이는 없었으나, 사범대학 편제의 특성으로 인한 차이가 나타나기도 하였다. 예비교사 교육 기관을 초·중 대상과 중등 대상으로 명확히 구별하는 국내 경우와 달리, 상당수 국외 연구가 이루어진 미국의 경우 초·중등을 구별하지 않는 교육 기관이 많았기에 별도로 ‘초·중등교육’을 코딩하여야 했다. 그러나 편제와 상관없이 몇몇 국외 연구들에서는 초·중 또는 중·고 교육 예비교사를 대상으로 하였음을 명확히 밝힌 경우들도 있어서 논문에서 최종적으로 표현한 바에 따라 학교급 구별을 실시하였다. 국외 연구에서는 대학의 예비교수자들을 대상으로 한 작문 평가 문식성 관련 연구들도 존재하였지만, 국내에서는 전무하였다. 한편, 학교급 차원에서도 구별이 없거나 확인이 불가한 경우들이 꽤 있었는데, 이 역시 이론적 분석이나 도구 개발에 해당하는 경우도 있었지만(박영민, 2021; Hodges et al., 2019; Stiggins, 1991; Xu & Brown, 2016), 문헌에 예비교사의 대상 학교급 정

보를 충분히 실어주지 않은 경우들도 있었다(Hedgcock & Lee, 2017).

연구 대상과 관련하여, 학교급 이외에도 자료 수집 및 분석의 대상이 누구인지에 따라 연구의 유형을 분류하기도 하였다. 예비교사의 평가 문식성 연구들의 탐구 대상은 국내의 경우 예비교사를 대상으로 하여 자료를 수집하고 분석한 연구들이 15건(63%)으로 가장 많은 비중을 차지하였으며, 그 다음으로 현직교사와 예비교사 모두를 대상으로 자료를 수집한 후 비교 분석한 연구들이 6건(26%)으로 나타났다. 실증적 연구 대상 없이 연구자 직관에 의거한 제언만 담은 연구들도 3건(13%) 있었다. 국외의 경우에도 예비교사 대상 자료 수집 및 분석 연구들이 19건(61%)으로 가장 많았으나, 그 외에도 다양한 경로를 통해 수집된 자료들을 연구에 활용한 것으로 나타났다. 현직교사와 예비교사의 현황 비교는 국내에 비해 상대적으로 많지 않았으나(2건, 6%), 그밖에도 예비교사 대상 작문 평가 교육에 대한 현직교사의 견해 분석(4건, 13%)이나 예비교사 대상 작문 평가 교육 관련 선행 연구 검토(3건, 10%) 등 더욱 다양한 연구 대상을 바탕으로 예비교사 작문 평가 교육을 입체적으로 조망해보고자 하는 시도들이 있었다. 예를 들어, Crusan 외(2016)의 연구에서는 비록 현직교사들을 대상으로 설문조사를 실시하였지만, 해당 문항들 중 상당수가 작문 평가 교육 경험에 대한 것들이었다는 점에서 예비교사 평가 문식성 신장을 위한 해결 방안 마련에 직·간접적으로 참고할 수 있었다. 한편, 선행 연구를 주요 자료로 삼은 연구들은 엄정하고 객관적인 문헌조사 방법론을 적용하여 예비교사 평가 문식성 관련 연구 동향을 살펴보기도 하였다. 예를 들어, Peck과 Kavanagh(2024)는 체계적 문헌고찰을 통해 예비교사의 작문 피드백 능력 향상을 위한 현장 연구들을 분석함으로써, 작문교육 이론 적용에 대한 예비교사들의 인식과 실천의 간극을 확인하였다. 국외의 경우에도 실증적 연구 대상 없이 제언만 담긴 연구들이 있었는데 모두 ‘교사의 평가 문식성’을 개념화하고자 하는 시도들이었다(Stiggins, 1991, 1995; Popham, 2009).

〈표 5〉 분석 문헌의 연구 방법 및 목적 관련 동향

기준		국내	국외
연구 방법	양적 연구 방법론	16건	7건
	질적 연구 방법론	4건	10건
	혼합 연구 방법론	1건	4건
	교육 도구, 측정 도구의 개발	1건	4건
	문헌 탐색, 이론적 탐구	2건	6건
연구 목적	양상의 기술, 문제 발견 및 분석	16건	14건
	문제 분석 및 원인 파악	3건	4건
	해결책 제시(실증적 근거 기반)	2건	9건
	해결책 제시(근거 없음)	1건	4건

다음으로, 〈표 5〉는 연구 일반 동향 중에서도 연구 방법 및 연구 목적과 관련된 수집 문헌들의 성격을 정리한 것이다. 먼저, 연구 방법과 관련하여, 국내외 간의 차이가 다분히 나타났다. 국내의 경우 양적 연구가 16건(67%), 질적 연구가 4건(17%), 혼합 연구가 1건(4%)으로 양적 연구가 크게 우세한 것으로 나타났지만, 국외의 경우 양적 연구가 7건(23%), 질적 연구가 10건(42%)으로 혼합연구가 4건(13%)으로, 상대적으로 더 고르게 나타난 가운데 질적 연구가 좀 더 우세한 경향을 보였다.

또한, 연구 목적과 관련하여, 국내(16건, 67%)와 국외(14건, 45%) 모두 양상이나 문제점의 기술적 분석에 중점을 둔 연구가 가장 많은 수를 차지하였다. 대표적인 사례로, 예비교사의 작문 채점 수행 자료를 분석하여 일관성 및 엄격성 등을 분석한 일련의 국내 문헌들을 들 수 있다(박영민, 2012; 박찬홍, 2019). 문제의 분석뿐만 아니라 그 원인까지 파악하고자 한 연구들은 국내 3건(13%), 국외 4건(13%)으로 파악되었는데, 대체로 질적 연구나 혼합 연구를 통해 사태를 심층적으로 파악하고자 한 경우들이 많았다(서수현, 2012; 장소영·박묘영, 2022; Dempsey et al., 2009; Volante & Fazio, 2007). 앞선 동향과 달리, 실증적 근거를 기반으로 한 해결 방안 제시 연구 동향은 국내 2건(8%), 국외 9건(29%)으로 상

당한 차이가 나타났다. 국외의 해결 방안 제시 연구들에서는 예비교사의 작문 평가 능력 향상을 위한 교육 방법을 개발하거나(Hodges et al., 2019), 작문 평가 능력 향상을 위한 방법을 적용하고 그 효과를 측정하거나(Grainger et al., 2018), 예비교사를 위한 작문 평가 교육 현장에서 활용할 수 있는 교수 방법을 소개하는 등(Hodges et al., 2019) 다양한 해법을 확인할 수 있었다.

2) 분석 문헌의 연구 주제 관련 동향

문헌의 일반적 동향 파악에 이어서, 예비교사 평가 문식성 관련 연구의 현황을 좀 더 심층적으로 파악하기 위하여 각 문헌들이 어떠한 연구 주제를 다루고 있는지를 중심으로도 질적 내용분석을 실시하였다. <표 6>은 분석 대상 문헌의 모든 연구 주제들을 귀납적으로 코딩하는 과정을 통해 도출한 기준을 적용하여, 국내와 국외 동향을 분석·비교한 것이다.

연구 주제 기준에 따른 분석에서 모든 문헌들은 4개의 기준(작문 채점, 작문 피드백, 작문 채점과 피드백, 작문 평가 일반) 하의 총 16개의 세부 항목 중 하나에만 속하며, 어떤 항목에도 속하지 않는 연구물이 없도록 코딩되었다. 먼저, 기준별 분포의 차원에서도 국내와 국외 연구 동향에 뚜렷한 차이가 나타났다. 국내의 경우, 예비교사 평가 문식성 관련 연구 상당수가 작문 채점과 관련된 연구인 것으로 밝혀졌다(17건, 71%). 그 다음으로 작문 평가 전반에 대한 일반적 논의들이 5건(21%)으로 나타났으며, 작문 피드백 관련 연구는 2건(8%)에 그쳤다. 국내에서는 예비교사의 작문 채점과 피드백을 동시에 다룬 연구는 한 건도 없었다. 그러나 국외의 경우, 작문 평가 전반에 걸친 논의 혹은 작문 평가 문식성 관련 논의들이 17건(55%)으로 가장 많은 수를 차지하였으며, 나머지 기준과 관련해서는 작문 채점 연구가 4건(13%), 작문 피드백 연구가 6건(19%), 작문 채점과 피드백을 함께 다룬 연구가 4건(13%)으로 대체로 고르게 분포하였다. 이상의 분포를 통해, 국내 연구에서는 예비교사의 작문 평가 문식성과 관

〈표 6〉 분석 문헌의 연구 주제 관련 동향

기준	세부 항목	국내	국외
작문 채점	작문 채점 양상 분석	8건	1건
	작문 채점에 대한 인식 분석	2건	-
	작문 채점 양상 및 인식 분석	3건	-
	작문 채점 교육 방법 적용	4건	3건
	총계	17건	4건
작문 피드백	작문 피드백 양상 분석	2건	1건
	작문 피드백 양상 및 인식 분석	-	3건
	작문 피드백 교육 방법 적용	-	2건
	총계	2건	6건
작문 채점과 피드백	작문 채점과 피드백 교육 방법 적용	-	2건
	작문 채점과 피드백 교육 방법 적용 및 그에 대한 인식 분석	-	2건
	총계	0건	4건
작문 평가 일반	작문 평가 계획 및 설계	1건	3건
	작문 평가에 대한 인식 및 태도	2건	2건
	작문 평가 능력 및 자질	1건	5건
	작문 평가 능력 및 인식과 태도	-	7건
	작문 평가자 간 협의	1건	-
	총계	5건	17건

련하여 작문 채점 위주로만 연구가 진행되었으며 피드백 혹은 채점과 피드백을 함께 다루는 연구들을 거의 찾아보기 힘들었다는 점을 지적할 수 있다.

각 세부 항목의 분포와 관련하여 대표적인 사례들을 중심으로 살펴보면 다음과 같다. 작문 채점 위주로 연구가 진행된 국내 문헌들을 살펴보면, 예비교사의 작문 채점 엄격성이나 일관성 관련 연구들이 가장 많이 수행되었다(박영민, 2012; 박영민·최숙기, 2009; 박찬홍, 2019; 백유진·박영민, 2023). 작문 채점에 대한 인식 분석은 주로 작문 평가 효능감 조사를 통해 연구가 이루어졌다(가은아, 2018; 박영민, 2010). 작문 채점 양상 및 그에 대한 인식 분석 연구도 있었는데, 해당 연구에서는 예비교사들로 하여금 작문 채점을 수행하게 하는 데 그치지

않고 자기 성찰일지를 작성하게 함으로써, 문제를 파악하는 데 그치지 않고 그 원인도 유추할 수 있게 하였다(장소영·박묘영, 2022). 작문 채점 교육 방법을 적용하고 그 효과를 확인한 연구들은 국내외에 고르게 분포하였다. 대표적으로 이윤빈(2015)은 예비교사를 대상으로 쓰기 평가 워크숍을 실시하여 사전·사후 비교를 통해 그 효과를 검증하였다. 국외 연구인 Youngs와 Bird(2010)에서는 예비교사 대상 채점 교육 프로그램을 실시하여 채점자 신뢰도 측면의 효과를 보여주기도 하였다.

다음으로, 작문 피드백 연구는 국외에서 훨씬 더 활발히 수행되었는데, 대표적으로 Koc과 Ily(2016)는 예비교사 피드백의 양상 분석에 그치지 않고 피드백 수행에 대한 참여자 인식도 함께 조사함으로써 그 실효성을 판단하고자 하였다. 또한 Peck과 Kavanagh(2024)는 예비교사 대상 작문 피드백 교육 방법 관련 연구들에 대한 문헌 탐구를 통해 효과적인 교육 방법을 총괄적으로 정리하기도 하였다. 국외 연구 중에는 작문 피드백뿐만 아니라 채점과 피드백을 동시에 적용한 연구들 역시 종종 찾아볼 수 있었는데, 모두 교육적 적용과 관련된 연구들이었다. 특히 Dempsey 외(2009)의 연구는 채점과 피드백 수행의 비계 역할을 할 수 있는 프로그램을 개발·적용할 뿐만 아니라 그 효과와 참여자 인식을 확인함으로써 예비교사 평가 문식성의 실질적 향상을 꾀하였다는 점에서 의미가 있다.

한편, 작문 평가 일반의 세부 항목에 따른 동향에서도 국내외 간에 큰 차이를 확인할 수 있었다. 각 세부 항목별 대표적인 사례들을 꼽아보자면, 먼저 Stiggins(1995)는 작문 평가 계획 및 설계를 위한 5가지 기준을 확립하고 이를 상세히 소개하였다. 또한 작문 평가에 대한 인식 및 태도와 관련하여, Siegel과 Wissehr(2011)는 예비교사들의 교수철학과 성찰을 담은 일지를 질적으로 분석하여 평가도구와 목적에 대한 이해, 평가 도구 활용 계획 등을 심층적으로 검토하였다. 작문 평가 능력 및 자질과 관련해서는 국내에 비해 다양한 경향성을 확인할 수 있었는데, Mertler(2004)는 작문 평가 문식성을 측정할 수

있는 도구를 개발하여 현직교사와 예비교사 간 차이를 비교 분석하였으며, Hedgcock과 Lee(2017)의 연구에서는 예비교사의 장르 인식 능력이 교육에 따라 어떻게 영향을 받는지, 그리고 장르 인식이 작문 평가에 어떠한 영향을 미치는지를 혼합 연구를 통해 확인하였다. 국외 연구들에서는 예비교사의 작문 평가 능력만 다루기보다는 작문 평가에 대한 전반적 인식과 함께 종합적인 차원으로 다룬 경우가 더 많았는데, 대표적으로 Xu와 Brown(2016)은 체계적 문헌고찰을 통해 선정된 연구물들을 중심으로 작문 평가 문식성 교육을 위한 구성 요소와 영향 요인을 추출하였다. 마지막으로, 작문 평가 일반 관련 세부 기준 중 유일하게 ‘작문 평가자 간 협의’ 관련 연구는 국외 검색에서는 확인되지 못하였고 국내에서 한 건 확인되었다(서수현, 2012). 그러나 작문 교육 현장에서 평가자 간 협의는 필수적이면서도 쉽지 않은 과정이라는 점에서 관련 연구가 더욱 필요해 보인다.

5. 결론

예비교사들이 진정한 작문 평가 문식성을 갖추기 위해서는 학생의 글을 타당하고 정확하게 채점하는 능력을 기르는 데 그치는 것이 아니라, 채점 결과의 의미를 이해하고 이를 바탕으로 학생들에게 필요한 부분이 무엇인지를 파악한 후 다양한 맥락을 고려하여 피드백을 실시하는 등 궁극적으로는 올바른 평가관을 바탕으로 학생의 작문 실력 향상을 꾀하는 일련의 과정을 터득해야 한다. 그러나 예비교사의 작문 평가 문식성과 관련된 논의들이 학술적으로도 충분히 이루어지지 않은 상황에서 무작정 이와 관련된 실천을 주장하기도 어렵다.

이에, 이 연구는 예비교사 대상 작문 평가 문식성 함양을 위한 이론적 기반 마련과 관련 후속 연구의 균형 있는 성장을 목적으로, 국내외 4개 데이터베이

스를 대상으로 주제범위 문헌고찰을 실시하고 최종 선정된 55개의 문헌들에 대해 발행연도, 연구 분야, 대상, 방법, 목적, 주제 등의 기준을 중심으로 연구의 동향을 파악하였다. 특히 국내와 국외를 동일한 기준으로 분석하여 국내외 간 연구 동향의 차이도 비교 분석할 수 있었다. 연구 결과를 종합적으로 정리하고 논의하는 동시에, 향후 예비교사 평가 문식성 관련 국내 후속 연구에 대한 제언을 제시하는 것으로 결론을 마무리하고자 한다.

우선, 국내외를 막론하고 예비교사의 평가 문식성 관련 연구들은 최근 20년 사이 가장 왕성하게 수행되었으며, 국외의 경우 평가 문식성 담론이 등장한 1990년대 이후로 급속도로 성장하였음을 확인할 수 있다. 작문 평가가 교실 현장에서 차지하는 위상이나 중요성을 고려할 때, 이러한 흐름이 지속되고 더욱 확장될 필요가 있어 보인다.

다음으로, 연구 분야 및 대상, 연구 방법과 관련하여 국내외 연구 동향에는 차이가 나타났다. 국내에서는 제1언어 분야의 연구들이 압도적으로 많았던 데 반해, 국외에서는 제1언어와 제2언어 분야 사이에 문헌 수 차이가 크지 않았다. 또한 연구 대상 차원에서, 국내의 경우 중등교육 예비교사 대상 연구가 대다수를 차지하였지만, 국외에서는 압도적인 차이까지는 아니었으며 대학교육 차원의 예비교수자 관련 연구들도 몇몇 있었다. 국내의 예비교사 양성 과정이 초등과 중등 대상으로 명확히 구별되고, 상당수의 연구자들이 중등교육 분야에 종사하고 있다는 점이 그 이유일 수 있다고 추측된다. 아울러 국외 연구에서는 비록 적은 수이지만 현직교사의 의견이나 선행연구의 문헌고찰 등 연구 분석 대상이 좀 더 다양함을 확인할 수 있었다. 국내에서도 더욱 다양한 연구 대상을 통해 평가 문식성의 여러 요소를 입체적으로 탐구할 필요가 있어 보인다.

연구 방법 측면에서도 차이가 나타났는데, 국내는 양적 연구가, 국외는 질적 연구가 더 많이 수행되었다. 이는 국내의 경우 예비교사의 채점 엄격성이나 일관성 관련 연구들이 많이 수행되었던 반면, 국외의 경우 평가 문식성의

다양한 측면들을 조명하고자 하는 시도들, 그리고 평가에 대한 인식과 수행의 불일치 등 질적 차원의 규명이 많았던 데 기인한다. 물론 채점 엄격성과 일관성 등의 주제는 예비교사 채점교육과 관련하여 핵심적인 주제 중 하나이기는 하다. 그러므로 향후 국내 후속 연구와 관련하여, 양적 연구도 지속하는 가운데, 예비교사의 작문 평가 문식성과 관련된 다양한 질적 차원의 관심도 더욱 확장될 필요가 있어 보인다. 또한 국내의 양적 연구들도 대체로 채점의 신뢰성 차원에 집중되어 있지만 그 외의 양적인 측면에 대한 탐구들로 더욱 확장될 필요도 있다.

한편, 연구 목적과 관련한 동향에서는 국내외 차이가 크지 않았다. 다만 국내외 막론하고 해결책 제시와 관련하여 근거가 없는 제언형 연구도 꽤 있었다는 점에서, 실증적 근거를 기반으로 한 해결책 제시, 즉, 예비교사 교육 프로그램을 마련하고 그 효과를 검증한 연구들이 더욱 많이 수행될 필요가 있다.

마지막으로, 연구 주제와 관련하여, 국내외 연구 동향에서 상당한 차이를 확인할 수 있었다. 국내 연구에서는 작문 피드백에 비해 작문 채점 관련 연구가 월등하게 많이 이루어졌으나, 국외의 경우 작문 채점과 피드백이 비슷한 정도로 탐구되었으며 특히 작문 채점과 피드백을 함께 포함하는 예비교사 대상 연구들도 있었다는 점에 주목할 필요가 있다. 물론 일개 연구의 초점은 작을수록 주제가 선명해질 수 있다는 점에서 채점이나 피드백 단독에 대한 연구도 의미가 크지만, 작문 평가 문식성의 관점에서는 이 모든 항목들이 어떻게 조화를 이루는지도 중요하게 다루어야 한다는 점에서 채점과 피드백을 함께 실시하는 연구도 필요하다. 또한 작문 채점과 피드백 이외에도, 작문 평가의 지식이나 능력을 측정하고 평가에 대한 관점 및 인식을 확인하며 예비교사가 자신의 수행을 성찰하는 과정을 주제로 삼은 연구들도 국내에 비해 국외가 월등히 많았다. 이 역시 진정한 작문 평가 문식성의 함양을 위해 꼭 다루어야 할 주제라는 점에서 관련 국내 연구의 확장을 기대해 본다.

문헌고찰의 방법론에서 아무리 엄정하고 객관적인 자료 수집의 절차를 거친다고 해도 검색어 검색만으로는 모든 주요 문헌들을 포함하기는 어렵다. 그러나 연구자 직관에 의존하는 자료 수집과 달리 객관적 기준에 의거하여 여러 단계의 필터링을 거친 수집과 선정의 절차를 포함한다는 점에서 그 의미를 찾을 수 있다. 또한 선정된 문헌들에 대하여 질적 내용분석의 방법론에 의거한 코딩을 실시하여 문헌 분석의 신뢰도와 타당도를 확보하고자 하였다는 점에서도 이 연구의 의의를 찾을 수 있다.

참고문헌

- 가은아(2018), 예비 국어교사의 쓰기 평가 지식과 쓰기 평가 효능감 분석, 리터러시 연구 9(4), 한국리터러시학회, 125-160.
- 강지영, 오지은(2021), 신경과학적 방법론을 적용한 교과교육 연구의 동향 탐색 - 문법교육적 시사점을 중심으로, 국어교육 175, 한국어교육학회, 273-324.
- 교육부(2015), 국어과 교육과정, 교육부 고시 제2015-74호 [별책5].
- 교육부(2022), 국어과 교육과정, 교육부 고시 제2022-33호 [별책5].
- 권이은, 이수진, 송정윤(2021), '질문'에 대한 읽기 교육 연구 동향: 체계적 문헌 분석의 방법을 활용하여, 독서연구 58, 109-142.
- 김은성, 김호정, 남가영, 박재현(2014), 체계적 문헌 고찰을 통한 국어 교과서 선정 기준 연구, 국어교육연구 33, 서울대학교 국어교육연구소, 391-417.
- 김정은, 강지영(2023), 언론 빅데이터 활용 연구 분석에 기반한 문법 교수·학습 방안 탐색, 국어교육연구 51, 서울대학교 국어교육연구소, 143-189.
- 김해인(2016), 텍스트 난도 연구의 동향과 과제, 한국어문교육 19, 고려대학교 한국어문교육연구소, 193-220.
- 김혜연(2015), 예비 초등 교사의 글쓰기 요구 인식 - 정체성 인식과의 관련성을 중심으로, 국어교육연구 36, 서울대학교 국어교육연구소, 425-462.
- 김혜연(2016), 교사의 작문 피드백 관련 국외 연구 동향 및 주요 쟁점 - 네트워크 분석과 문헌 고찰의 적용, 국어교육연구 37, 서울대학교 국어교육연구소, 95-136.
- 김혜연(2017), 협동 글쓰기의 국내외 연구 동향에 대한 체계적 문헌 고찰, 작문연구 34, 한국작문학회, 7-49.
- 김혜연(2022), 작문 유형화 연구의 국제 동향에 대한 체계적 문헌 고찰, 새국어교육 132, 한국국어교육학회, 211-247.
- 박영민(2010), 예비국어교사의 쓰기 평가 효능감 분석, 청람어문교육 42, 청람어문교육학회, 181-207.
- 박영민(2012), 예비국어교사의 중학생 논설문 평가에서 발견되는 엄격성 및 일관성의 특성, 국어교육학연구 43, 국어교육학회, 253-283.
- 박영민(2021), Python GUI 기반 학생 글 채점 연습 프로그램의 구성과 활용, 작문

- 연구 48, 한국작문학회, 61-91.
- 박영민, 최숙기(2009), 현직 국어교사와 예비 국어교사의 쓰기 평가 비교 연구, *교육과정평가연구* 12(1), 한국교육과정평가원, 123-143.
- 박찬홍(2019), 예비 국어교사의 쓰기 평가 효능감과 쓰기 평가 수행에 대한 연구, *국어문학* 72, 국어문학회, 489-516.
- 백유진, 박영민(2023), 다국면 Rasch 모형을 활용한 예비 국어교사와 현직 국어교사의 쓰기 평가 특성 차이 분석, *청람어문교육* 92, 청람어문교육학회, 49-81.
- 서수현(2012), 쓰기 평가 협의 과정에 나타난 쓰기 평가자의 인식 연구, *국어교육학연구* 44, 국어교육학회, 335-367.
- 서현주, 김수영(2018), 주제범위 문헌고찰이란 무엇인가?, *보건의료기술평가* 6(1), 한국보건의료기술평가학회, 6-21.
- 안상희(2020), 대학의 계열별 글쓰기 교육에 대한 체계적 문헌 고찰, *한국어문교육* 30, 고려대학교 한국어문교육연구소, 69-96.
- 이수진(2020), 국어교육분야에서의 읽기 부진 연구 동향: 주제범위 문헌고찰(Scoping Review), *독서연구* 57, 한국독서학회, 145-169.
- 이윤빈(2015), 쓰기 평가 워크숍이 예비 교수자의 쓰기 평가 전문성 신장에 미치는 효과, *작문연구* 27, 한국작문학회, 83-117.
- 장소영, 박묘영(2022), 중등 예비영어교사의 영어 작문 채점에 영향을 미치는 요인: 자기 성찰일지 분석을 중심으로, *Studies in English Education* 27(4), 글로벌 영어교육학회, 569-598.
- 정혜승(2008), 문식성(Literacy)의 변화와 기호학적 관점의 국어과 교육과정 모델, *교육과정연구* 26(4), 한국교육과정학회, 149-172.
- Arksey, H., & O'Malley, L.(2005), Scoping studies: towards a methodological framework, *International Journal of Social Research Methodology* 8(1), 19-32.
- Brookhart, S. M.(2011), Educational Assessment Knowledge and Skills for Teachers, *Educational Measurement: Issues and Practice* 30(1), 3-12.
- Crusan, D., Plakans, L., & Gebriel, A.(2016), Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices, *Assessing Writing* 28, 43-56.
- DelleBovi, B. M.(2012), Literacy instruction: From assignment to assessment,

- Assessing Writing* 17(4), 271-292.
- Dempsey, M. S., PytlikZillig, L. M., & Bruning, R. H.(2009), Helping preservice teachers learn to assess writing: Practice and feedback in a Web-based environment, *Assessing Writing* 14(1), 38-61.
- Ferris, D.(2007), Preparing teachers to respond to student writing, *Journal of Second Language Writing* 16(3), 165-193.
- Grainger, P. R., Heck, D., & Carey, M. D.(2018), Are Assessment Exemplars Perceived to Support Self-Regulated Learning in Teacher Education? *Frontiers in Education* 3(60), 1-9.
- Hedgcock, J. S., & Lee, H.(2017), An exploratory study of academic literacy socialization: Building genre awareness in a teacher education program, *Journal of English for Academic Purposes* 26, 17-28.
- Hodges, T. S., Wright, K. L., Wind, S. A., Matthews, S. D., Zimmer, W. K., & McTigue, E.(2019), Developing and examining validity evidence for the Writing Rubric to Inform Teacher Educators (WRITE), *Assessing Writing* 40, 1-13.
- Junqueira, L., & Payant, C.(2015), “I just want to do it right, but it's so hard”: A novice teacher's written feedback beliefs and practices, *Journal of Second Language Writing* 27, 19-36.
- Koc, B., & Ilya, A.(2016), Exploring pre-service language teachers' perceptions and actual practices of giving feedback in micro-teaching, *Procedia -Social and Behavioral Sciences* 232, 421-429.
- Krippendorff, K.(2013), *Content analysis: An introduction to its methodology* (3rd. ed.), Sage.
- Lam, R.(2019), Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong, *System* 81, 78-89.
- Malone, M. E.(2013), The essentials of assessment literacy: Contrasts between testers and users, *Language Testing* 30(3), 329-344.
- Mertler, C. A.(2004), Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education* 33(1), 49-64.
- Peck, L., & Kavanagh, S. S.(2024), Learning to give feedback to student writers: A

- systematic review of research on field-based experiences for pre-service writing teachers, *Teaching and Teacher Education* 144, article 104602.
- Plake, B. S.(1993), Teacher assessment literacy: Teachers' competencies in the educational assessment of students, *Mid-Western Educational Researcher* 6(1), 21-27.
- Popham, W. J.(2009), Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice* 48(1), 4-11.
- Schreier, M.(2012), *Qualitative content analysis in practice*, Sage.
- Siegel, M. A., & Wissehr, C.(2011), Preparing for the Plunge: Preservice Teachers' Assessment Literacy, *Journal of Science Teacher Education* 22(4), 371-391.
- Shulman, L. S.(1986), Those who understand: Knowledge growth in teaching, *Educational Researcher* 15(2), 4-14.
- Shulman, L. S.(1987), Knowledge and teaching: Foundations of the new reform, *Harvard educational review* 57(1), 1-23.
- Stiggins, R. J.(1991), Assessment literacy, *Phi Delta Kappan* 72, 534-539.
- Stiggins, R. J.(1995), Assessment literacy for the 21st century, *Phi Delta Kappan* 77(3), 238-245.
- Volante, L., & Fazio, X.(2007), Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development, *Canadian Journal of Education* 30(3), 749-770.
- Wang, T.-H., Wang, K.-H., & Huang, S.-C.(2008), Designing a Web-based assessment environment for improving pre-service teacher assessment literacy, *Computers & Education* 51(1), 448-462.
- Weigle, S. C.(2002), *Assessing writing*, Cambridge University Press.
- White, E. M.(1994), Issues and problems in writing assessment, *Assessing Writing* 1(1), 11-27.
- Xu, Y., & Brown, G. T. L.(2016), Teacher assessment literacy in practice: A reconceptualization, *Teaching and Teacher Education* 58, 149-162.
- Youngs, P., & Bird, T.(2010), Using embedded assessments to promote pedagogical reasoning among secondary teaching candidates, *Teaching and Teacher Education* 26(2), 185-198.

예비교사의 작문 평가 문식성 관련 연구 동향 - 주제범위 문헌고찰의 적용

김혜연(동국대학교-서울)

교실 현장에서 작문 평가의 위상과 중요성이 증대됨에 따라, 교사교육의 차원에서 도 예비교사의 작문 평가 문식성을 제대로 함양하는 일이 관건으로 떠올랐다. 이 연구는 예비교사의 작문 평가 문식성 관련 연구 동향을 확인하고 국내외 간 동향을 비교 분석함으로써, 실질적인 교사교육 실천과 더욱 균형 잡힌 후속 연구 수행을 위한 기반을 마련하고자 하였다. 이를 위하여, 주제범위 문헌고찰의 방법론을 도입하여 연구 질문을 도출하고 국내외의 4개 데이터베이스를 대상으로 연구물을 수집하였다. 또한 포함 및 제외 기준을 적용하여 최종 대상 문헌을 국내 24건, 국외 31건으로 확정하였다. 질적 내용분석에 의거하여 선정된 문헌에 대한 자료 기입을 실시하였으며, 이에 따라 발행연도, 연구 분야, 연구 대상, 연구 방법, 연구 목적, 연구 주제 등을 기준으로 분석이 이루어졌다. 분석 결과에 따르면, 발행연도 및 연구 목적에서는 국내외 간 동향이 유사하게 나타난 반면, 연구 분야, 연구 대상, 연구 방법 측면에서는 국내외 간 차이가 꽤 나타났다. 연구 주제 차원에서도 국내 연구의 경우 예비교사의 작문 채점에만 집중되는 경향이 컸으나, 국외의 경우 채점과 피드백이 대체로 균형 있게 탐구되었으며 작문 평가의 관점이나 인식, 평가 수행에 대한 성찰 등 작문 평가 문식성의 다양한 측면을 다룰 수 있는 주제들이 고르게 분포하였다. 작문 평가 문식성의 개념이 추구하는 궁극적인 방향을 고려할 때 국내 연구가 좀 더 다양화 될 필요성이 제기되었다.

주제어 : 작문 평가, 평가 문식성, 채점, 피드백, 예비교사, 교사교육

Research Trends on Preservice Teachers' Writing Assessment Literacy

– A Scoping Review

Hyeyoun Kim (Dongguk University-Seoul)

As the significance and importance of writing assessment in classroom settings have increased, fostering writing assessment literacy among preservice teachers has emerged as a crucial task in teacher education. This study aims to identify research trends related to preservice teachers' writing assessment literacy and compare domestic and international trends to lay the groundwork for practical teacher education and more balanced future research. According to a scoping review, studies were collected from four domestic and international databases. After applying inclusion and exclusion criteria, 24 domestic and 31 international studies were selected as the final literature sample. The results revealed that while the trends in publication year and research objectives were similar between domestic and international studies, notable differences were found in research fields, subjects, and methods. In terms of research topics, domestic studies tended to focus primarily on reliability of preservice teachers' writing assessment scoring, whereas international studies explored both scoring and feedback in a more balanced manner. Additionally, topics addressing various aspects of writing assessment literacy, such as perspectives, perceptions, and reflections on writing assessment, were more evenly distributed in international research.

Keywords : Writing evaluation, Assessment literacy, Scoring, Feedback, Preservice teacher, Teacher education

이 논문은 2024년 9월 20일 투고 완료되어 2024년 9월 25일 편집위원회에서 심사위원을 선정된 뒤 2024년 10월 10일까지 심사를 완료하여 2024년 10월 16일 편집위원회에서 게재가 결정된 논문임.