

교육과정평가연구
The Journal of Curriculum and Evaluation
2025, Vol. 28, No. 1, pp.185~214
DOI: <https://doi.org/10.29221/jce.2025.28.1.185>

텍스트 마이닝을 통한 초·중등 교사의 과정중심평가에 대한 인식 연구

권순보 (한국교원대학교 시간강사)*
박수원 (풍양초등학교 교사)**

요약

본 연구는 초·중등교사들의 과정중심평가에 대한 인식을 탐구하기 위해 수행되었다. 이를 위해 교사 255명을 대상으로 서술형 설문조사를 실시하고, 텍스트 마이닝 기법을 활용하여 분석하였다. 연구 결과, 교사들은 과정중심평가를 학생의 학습 과정에서 이루어지는 평가로 인식하고 있으며, 이를 통해 개별 학생의 성장을 지원할 수 있다고 보았다. 그러나 실행 과정에서 현실적인 제약이 존재하는 것으로 나타났다. 초등교사들은 다양한 교과를 평가해야 하는 부담을, 중등교사들은 평가의 객관성과 피드백 제공의 어려움을 주요한 난점으로 지적하였다. 또한, 과정중심평가의 개념에 대한 교사들의 이해 수준이 다양하여, 평가의 의미와 실행 방식에 대한 혼란이 존재하는 것으로 분석되었다. 연구 결과를 바탕으로, 초등과 중등의 교육 환경 차이를 반영한 맞춤형 정책 지원과 과정중심평가의 명확한 개념과 안내가 필요함을 제언한다.

주제어: 과정중심평가, 초등교사, 중등교사, 텍스트 마이닝

* 제1저자, sbkwon@knue.ac.kr

** 교신저자, golilong@naver.com

I. 서론

과정중심평가는 ‘학습 결과에 대한 평가’의 대안으로 등장한 ‘학습을 위한 평가’의 한 부분으로 학생의 메타인지를 계발하고 지원하는 ‘학습 과정으로서의 평가’를 의미한다 (Earl, 2013). 교육부와 한국교육과정평가원(2017)에서는 과정중심평가를 ‘교육과정의 성취기준에 기반한 평가 계획에 따라 교수·학습 과정에서 학생의 변화와 성장에 대한 자료를 다각도로 수집하여 적절한 피드백을 제공하는 평가’라고 정의하고 있다. 이를 반영한 2022 개정 교육과정에서는 학교 교육에서 평가 시 ‘학습 결과만이 아니라 결과에 이르기까지의 학습 과정을 확인하고 환류’하여야 하고, ‘학생이 자신의 학습 과정과 결과를 스스로 평가할 수 있는 기회를 제공’해야 한다고 강조한다(교육부, 2022).

과정중심평가는 결과중심평가와 대비되는 개념이다. 이는 평가 패러다임의 확장이자, 분절되어 있던 교육과정과 교수·학습, 평가의 일체화를 의도한 것이다. 과정중심평가란 용어가 사용된 초기 출처인 김순남 외(2013)의 「창의인재 육성을 위한 학생평가 정책 연구」에서는 창의적 인재 양성을 위해 결과 중심의 평가 방식을 해소해야 한다는 문제 의식에서 과정중심평가의 도입 필요성을 제기한다. 국가 교육과정에서는 2015 개정 교육과정부터 과정중심평가의 취지를 반영하여, 평가 운영 방법으로 적극 권장하고 있다(교육부, 2015). 2015 개정 교육과정 총론에서 ‘학습의 과정을 중시하는 평가’가 제시된 것이 그 예이다.

2015 개정 교육과정 총론에서 과정중심평가가 제시된 이후 개념에 대한 관심이 높아졌다. 과정중심평가에 관한 선행연구로 교사와 학생의 실행 경험을 심층적으로 다룬 연구(김영실, 2022; 김유정 외, 2019; 신보미, 이경연, 2018; 신혜진 외, 2017; 정혜진, 지은림, 2020; 전현욱, 2019; 전현욱, 이형연, 2019; 홍수향 외, 2017), 정책의 적용 여부와 평가 방식의 활용 현황을 정량적으로 분석한 연구(고현, 2019; 김석우, 2021; 김석우 외, 2021; 반재천 외, 2018), 과정중심평가의 의미와 정책을 분석한 연구(김정민, 2018; 박정, 2019), 과정중심평가의 연구동향을 분석한 연구(최훈원, 최윤정, 2024) 등이 이어진다. 선행연구들에서 과정중심평가는 학생의 변화와 성취를 지속적으로 파악하여 맞춤형 피드백을 제공하는 평가(고현, 2019; 김석우, 2021; 박지현 외, 2020), 수업과 평가가 통합되어 학습의 전 과정에서 이루어지는 평가(김석우 외, 2021; 김유정 외, 2019; 김정민, 2018; 김진희, 2020), 교육과정의 성취기준을 기반으로 평가를 계획하고 학습자의 목표 달성 정도를 다각도로 확인하는 평가(전현욱, 이형연, 2019; 홍수향 외, 2017) 등으로 설명된다.

이는 현장에서 과정중심평가가 학습 결과를 평가하는 방식에서 벗어나, 학습 과정을

포함하여 학생의 전반적인 성장을 지원하는 평가 방식으로 인식되고 있다는 것을 보여준다. 그러면서 한편으로는 과정중심평가가 수행평가의 또 다른 이름으로 인식되거나, 학생에게 교사의 피드백이 제공된 평가 방식과 같이 단순한 기법으로 인식되고 있다는 것을 보여준다. 또한 선행연구들에서는 교사들이 과정중심평가의 도입 취지는 공감하나, 현실적인 어려움과 제약으로 효과적으로 적용하고 있지 못하다는 지적이 이어진다(김영실, 2022; 반재천 외, 2018; 신혜진 외, 2017).

이에 본 연구에서는 현장 교사들이 과정중심평가를 어떻게 이해하고 있고, 실제 어떠한 형태로 실천하고 있으며, 실행에 어려움을 겪는 이유가 무엇인지를 종합적으로 탐구해보고자 한다. 최훈원과 최윤정(2024)의 과정중심평가 연구동향 분석에 따르면, 학교 급별, 교과별, 특정 시기 전후 등으로 연구 주제에 차이를 보인다. 이는 현 시점에서 과정중심평가에 대한 현장 교사들의 인식을 면밀히 파악하는 것이 현장의 평가 여건과 구조를 살펴보기 위해 필요하다는 것을 보여준다. 또한 개념에 대한 교사들의 인식을 분석하여, 현 시점에서 과정중심평가를 현장에 일반화할 수 있는 조건을 탐색해볼 수 있을 것이다. 더 나아가 초중과 중등교사들의 인식을 별도로 분석하여 비교해본다면, 각각의 교육 환경에서 과정중심평가를 수행하는 데 있어서의 어려움과 이를 극복하기 위한 정책 방안이 드러날 것이다.

이를 위해 본 연구에서는 텍스트 마이닝 기법을 활용해 과정중심평가에 관한 초·중등 교사들의 인식을 분석하고자 한다. 텍스트 마이닝은 비정형화된 자연어 데이터를 연구 목적에 적합하도록 텍스트 정보로 추출하여 가공하는 분석 방법이다(Kumar & Paul, 2016). 서술형 설문조사를 텍스트 마이닝으로 분석하는 방법은 응답자의 서술형 데이터를 정량화하면서도 진술의 맥락을 보존할 수 있고, 선택형 설문조사가 간과할 수 있는 심층적인 인식과 숨겨진 패턴을 드러내는 데 유리하다. 또한 이 기법은 방대한 양의 데이터를 분석하여 사회 집단의 인식을 이해할 수 있다는 점에서 초·중등교사들의 과정중심평가에 관한 인식을 분석하는 유용한 방법이 될 것이다. 본 연구의 연구 문제는 다음과 같다.

첫째, 교사들은 과정중심평가를 어떻게 이해하고 있으며, 이는 국가 교육과정에 부합하는가?

둘째, 학교 현장에서 과정중심평가를 실천하는 과정에서 당면하는 문제점은 무엇인가?

셋째, 과정중심평가를 현장에 안착하기 위한 조건은 무엇인가?

II. 이론적 배경

1. 평가의 세 가지 접근 방법

평가는 교수·학습을 바꿀 수 있는 잠재력을 지닌다. 평가는 그 자체로 학생의 학습과 성취도를 높일 수 있는 도구가 된다. 이러한 평가에 관한 인식은 학생 평가에 대한 패러다임의 변화로 가능하다(김순남 외, 2013). 기존에 평가를 시험과 등치시키고, 시험으로 사회를 통제하거나 사회적 이동을 위한 수단으로 삼았던 인식으로부터 벗어나야 하기 때문이다. 이는 학습 측정을 위한 평가에서 학습을 위한 평가로, 결과 중심적 평가에서 과정 중심적 평가로의 전환을 의미한다(김순남 외, 2013).

평가에 대한 세 가지 접근 방법이 존재한다. 학습 결과에 대한 평가(assessment of learning), 학습을 위한 평가(assessment for learning), 학습 과정으로서의 평가(assessment as learning)가 그것이다(Earl, 2013). 하나씩 살펴보면, 먼저 학습 결과에 대한 평가는 학교에서 주를 이루는 전통적인 평가 방식이다. 다른 학생과의 비교를 통해 상대적 위치를 나타내거나, 학습 결과를 시험을 통해 증명하여 학생과 학부모에게 학생의 발전 정도를 드러내 보여주는 데 목적을 둔다. 이를 위한 평가는 신뢰할 수 있고 정당한 것이어야 한다. 교사들은 학생과 학부모가 합리적이고 정당한 선택을 할 수 있도록 학생의 역량 또는 성취도를 타당하게 진술하기 위한 평가를 실시하여야 한다.

다음으로 학습을 위한 평가는 전통적 평가 방식에 대한 대안적 관점을 제공한다. 전통적 평가 방식이 주로 총괄평가로 이루어진다면, 학습을 위한 평가는 형성평가로 이루어지는 경우가 많다. 학습을 위한 평가는 교수·학습 활동을 수정하거나, 학생들의 이해 정도를 교사에게 제공하기 위해 설계된다. 학생의 개별적인 학습 요구를 파악하여 다음 단계의 활동에 반영할 수 있도록 한다. 우리나라에서 결과중심평가의 대안으로 사용되는 과정중심평가 개념은 상당 부분 학습을 위한 평가의 의미를 반영한 것이라 할 수 있다.

마지막 학습 과정으로서의 평가는 학습을 위한 평가의 한 부분이다. 학습 과정으로서의 평가는 평가를 학생의 메타인지를 계발하고 성찰을 지원하는 과정으로 사용한다. 주요 평가자가 교사에서 학생으로 전환되어, 학습 과정과 평가를 잇는 연결자로서 학생 역할을 강조한다. 학습 과정으로서의 평가에서 학생들은 자신의 학습을 성찰하고 비판적으로 분석한다. 자기 평가가 핵심이다. 성공적인 학습 과정으로서의 평가는 학생들이 학습 과정에서 성찰하는 질문을 하게 하고, 학습과 수행에 관한 다양한 전략을 스스로 고려하도록 하여야 한다. 우리나라 국가 교육과정 총론에서는 평가에 관해 세 가지 접근 방법을 총체적으로 제시하면서, 학습을 위한 평가와 학습 과정으로서의 평가의 설계와 운영을 강조한다(교육부, 2015; 2022).

2. 관련 개념과의 구분

2015 개정 교육과정 총론에서는 ‘학습의 과정을 중시하는 평가를 강화하여 학생이 자신의 학습을 성찰하도록 하고, 평가 결과를 활용하여 교수·학습의 질을 개선’시키고자 하며, ‘학습의 결과뿐만 아니라 학습의 과정을 평가하여 모든 학생의 교육 목표에 성공적으로 도달’할 수 있어야 한다고 명시한다(교육부, 2015). 이는 국가 교육과정에 앞서 제시한 학습을 위한 평가와 학습 과정으로서의 평가 개념이 동시에 담겨 있다는 것을 보여준다.

그러나 과정중심평가에 대한 개념적 혼란이 존재한다. 그것이 ‘학습 과정을 평가’한다는 측면서 형성평가와, ‘학습 과정을 통해 평가’한다는 측면에서 수행평가와 의미가 중복된다는 것이다(김정민, 2018; 신혜진 외, 2017; 정혜진, 지은림, 2020). 이로 인해 학교 현장은 개념들에 대한 혼란을 갖거나, 구분 없이 혼재해 사용하는 경향을 보이게 된다. 개념들 간에 명확한 이해를 갖지 못한 채 혼용되다 보니, 실제 과정중심평가를 수행하고 있다 하여도 그것에 대한 확신을 갖지 못하는 문제가 발생하는 것이다. 따라서 과정중심평가와 유사 용어들 간의 개념적 차이를 짚어갈 필요가 있다.

먼저 형성평가는 ‘학생의 발전이나 향상을 점검하기 위해 그 과정에서 이루어지는 평가’로 정의된다(김정민, 2018). 과정중심평가는 교사에게 진단 정보를 제공하고, 학생에게 피드백 결과와 성찰 근거를 제공한다는 측면에서 형성적(formative)이라 할 수 있다. 과정중심평가에서는 학생의 학습 결과뿐만 아니라 그 과정을 평가하기 위해 형성평가를 활용하기 적합하다. 다만, 과정중심평가에서는 학습 과정뿐만 아니라 학습 결과에 대한 평가도 이루어질 수 있기 때문에 이를 형성평가와 등치시키기에는 무리가 있다. 과정중심의 형성평가 뿐만 아니라 과정중심의 총괄평가 역시 존재 가능한 것이다.

다음으로 수행평가는 ‘학생 개개인이 처한 상황을 고려하여, 평가 과제 수행 과정과 결과에 대한 포괄적인 정보를 수집하는 평가 방식’이라고 정의된다(박도순, 1995). 과정중심평가는 학습의 결과뿐만 아니라, 학습 과정을 평가의 주요 대상으로 삼는다는 측면에서 수행평가의 특징을 잘 설명하는 개념이다. 그래서 과정중심평가의 의도를 수업장에서 가장 잘 구현할 수 있는 평가 방식 중 하나가 수행평가라 설명된다(교육부, 한국교육과정평가원, 2017). 다만, 학습 과정을 평가하기 어려운 결과 중심의 수행평가도 존재할 수 있고, 학습 과정과 별개로 과제형으로 제시되는 수행평가도 존재할 수 있다. 따라서 수행평가의 특징 중 하나가 과정중심평가이고, 과정중심평가의 실천 방법 중 하나가 수행평가라 말할 수 있지만, 과정중심평가가 곧 수행평가라 말할 수는 없다.

3. 과정중심평가에 관한 교사들의 인식

선행연구에서 교사들의 과정중심평가에 대한 인식은 크게 두 가지 긍정과 세 가지 부정으로 구조화된다. 첫 번째 긍정적 인식으로 교사들은 과정중심평가의 필요성에 공감한다(고현, 2019; 김영실, 2022; 김진희, 2020; 신혜진 외, 2017). 교사들은 과정중심평가가 학생들의 학습 과정에 대한 이해를 돕고 성장을 지원할 수 있다는 점에 긍정적으로 평가한다. 기존의 결과 중심 평가에서 벗어나 학습 과정에서 학생에게 의미 있는 피드백을 제공하고, 수업과 평가의 통합을 가능하게 한다는 점 역시 과정중심평가의 교육적 가치를 높여준다. 두 번째 긍정적 인식으로 교사들은 과정중심평가를 통해 교육과정과 수업, 평가의 통합 가능성을 기대한다(고현, 2019; 김유정 외, 2019; 김정민, 2018). 과정중심평가는 교육과정과 수업, 평가가 하나로 통합되어 운영되는 방식으로 학생 중심의 수업을 강화하고, 수업 중 평가를 통해 학생 개개인의 특성에 맞는 학습 지원이 가능하게 할 수 있다고 본다. 이는 평가가 그 자체로 수업의 일부가 되기 때문에 갖게 되는 이점이다.

한편, 첫 번째 부정적 인식으로 교사들은 과정중심평가의 실행에 현실적 어려움이 존재한다고 여긴다(고현, 2019; 반재천 외, 2018). 교사들은 과정중심평가의 취지에는 공감하나, 평가를 실천하는 과정에서 여러 현실적 어려움을 겪고 있다. 특히 평가의 객관성과 공정성에 대한 문제로 인해 성적 평가에 대한 학생, 학부모의 반발이 우려된다. 두 번째 부정적 인식으로 교사들은 과정중심평가를 효과적으로 운영하기 위해서는 충분한 평가 전문성이 필요하고 여긴다(고현, 2019; 김유정 외, 2019). 그들은 스스로 과정중심평가의 구체적 방법에 대한 지식과 기술이 부족하다고 여기며, 평가 도구 활용과 피드백 제공에 대한 충분한 연수와 지원이 필요하다고 여긴다. 세 번째 부정적 인식으로 교사들은 과정중심평가를 운영하기 위한 충분한 정책 지원이 필요하다고 여긴다(김영실, 2022; 김진희, 2020; 반재천 외, 2018; 신혜진 외, 2017). 교사들은 과정중심평가의 성공적 정착을 위해 행정 업무 경감, 평가 도구와 자료의 제공, 현장 밀착형 연수 등의 정책 지원을 요구한다. 그들은 구체적인 지원이 부족한 상태에서 부여된 새로운 평가 방식에 부담을 느끼고, 일부 교사들은 형식적으로 과정중심평가를 운영하기도 한다고 지적한다.

선행연구에서 드러난 과정중심평가에 대한 교사들의 인식은 그 도입 취지와 의미는 이해하나 현실적인 어려움과 지원 부족으로 효과적으로 적용하고 있지 못하는 것으로 정리할 수 있다. 또한 교사들은 과정중심평가를 평가의 장면에 국한하지 않고, 학습 과정에 대한 포괄적이고 다차원적인 접근으로 이해하고 있으며, 학생 개개인의 학습 성장을 지원하고 수업의 질을 향상시키는 중요한 도구로 인식하고 있다는 점을 알 수 있다. 다만, 선행연구들은 과정중심평가 도입 초기의 혼란이 있던 시기에 이루어진 것들이 많

다. 따라서 과정중심평가가 보편적으로 활용되고 있는 최근에 교사들의 인식을 본 논문을 통해 살펴보고자 한다.

III. 연구 방법

1. 연구 대상자

2024년 1월부터 2월까지 전국 초·중등 교사를 대상으로 설문을 진행하여 수집된 자료를 분석하였다. 본 연구에서는 구글 설문을 활용하여 온라인 방식으로 설문을 진행하였으며, 응답자의 신뢰성을 확보하기 위해 인증을 거친 교사 커뮤니티, 대학원 수업을 수강하는 현직 교사 등을 대상으로 자료를 수집하였다. 설문조사에는 총 255명의 교사가 응답하였으며, 이 중 초등학교 교사는 143명, 중등학교 교사는 112명으로 구성되었다.

본 연구에서 확보한 143명의 초등학교 교사, 112명의 중등학교 교사의 응답은 기존 LDA를 활용한 선행연구들과 비교했을 때 상대적으로 충분한 표본 규모를 갖추고 있다. 예를 들어, 곽민호와 서영진(2020)은 57명의 대학생 동료평가 데이터를, 신은혜(2022)는 15명의 과학교사 면담 데이터를, 최진수와 정혜원(2024)은 국내외 프로세스 데이터 관련 논문 66편을, 장윤선(2023)은 초·중등학생 자기평가 연구 논문 113편의 초록을 활용하여 LDA 분석을 수행하였다. 본 연구는 전국 단위의 초·중등 교사를 대상으로 표본을 수집하였으며, 자료 수집 과정에서 신뢰성을 확보하기 위해 노력했다는 점에서 일정 수준의 타당성을 확보했다고 볼 수 있다.

설문 문항은 '과정중심평가에 대한 이해', '현장에서의 과정중심평가 실행 방법', '과정중심평가 실행 시 당면하는 문제점' 등을 포함하였다. 설문 응답 결과, 255명의 교사가 모든 문항에 성실히 답변한 것으로 확인되었으며, 응답자의 배경 정보는 <표 1>에 제시되어 있다. 본 연구는 전국 단위로 표본을 수집하여 다양한 교사 집단을 반영하였으며, 이를 통해 과정중심평가 실행에 대한 교사들의 인식을 폭넓게 탐구할 수 있었다.

<표 1> 설문에 참여한 교사들의 배경 정보

단위 : 명(%)

변인	구분	교사		
		빈도(명)	비율(%)	
성별	남	78	30.6	
	여	177	69.4	
소속 교육청	강원도교육청	5	2.0	
	경기도교육청	126	49.4	
	경상남도교육청	4	1.6	
	경상북도교육청	6	2.4	
	광주광역시교육청	7	2.7	
	대구광역시교육청	7	2.7	
	대전광역시교육청	10	3.9	
	부산광역시교육청	12	4.7	
	서울특별시교육청	39	15.3	
	세종특별자치시교육청	3	1.2	
	울산광역시교육청	6	2.4	
	인천광역시교육청	10	3.9	
	전라남도교육청	2	0.8	
	전라북도교육청	2	0.8	
	제주특별자치도교육청	1	0.4	
	충청남도교육청	8	3.1	
	충청북도교육청	7	2.7	
급별	초등	초등	139	54.5
		특수	4	1.6
	중등	가정	6	2.4
		과학	13	5.1
		국어	19	7.5
		기술	2	0.8
		미술	3	1.2
		사회	14	5.5
		수학	16	6.3
		영어	24	9.4
		음악	4	1.6
		제2외국어	6	2.4
		체육	2	0.8
특수	3	1.2		
교직경력	5년 이하	39	15.3	
	6~10년	64	25.1	
	11~15년	68	26.7	
	15~20년	40	15.7	
	20년 이상	44	17.3	

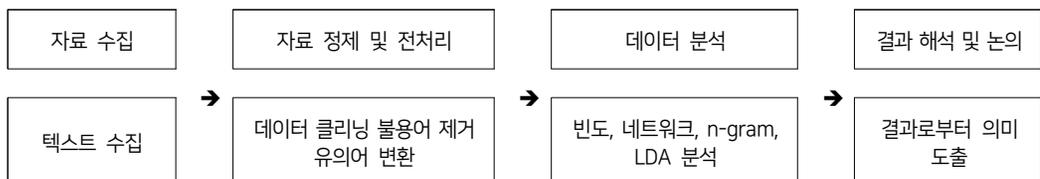
2. 연구 절차 및 데이터 전처리

본 연구에서는 분석을 수행하기 위해 R을 활용하였다. R은 전통적인 통계 분석뿐만 아니라 텍스트 데이터 분석을 위한 다양한 패키지를 제공한다. 본 연구에서는 텍스트 데이터의 전처리부터 네트워크 분석, 토픽 모델링, 그래프 시각화에 이르기까지의 전 과정을 R을 통해 수행하였으며, 사용된 주요 패키지는 <표 2>에 제시하였다.

<표 2> 본 연구에서 사용한 R 패키지 기능

패키지	기능
dplyr	데이터 프레임 변환 및 행과 열의 추가, 삭제
ggplot2	데이터 시각화를 위한 그래프 생성
KoNLP	형태소 분석기를 활용한 한글 텍스트의 형태소 추출
tidyr	데이터 구조 변환 및 결측치 처리
tidytext	텍스트를 토큰화하여 분석, 토큰 단위 데이터 처리
tidyverse	데이터 분석을 위한 여러 패키지의 모음, 데이터 처리 및 시각화
topicmodels	문서-단어 행렬(DTM)을 활용한 LDA 모델링
stringr	정규 표현식을 이용한 문자열 처리, 불용어 제거 및 유의어 처리
widyr	동시 출현 분석을 위한 단어 쌍 빈도 계산

[그림 1]은 연구 절차를 요약하여 보여준다. 연구의 첫 단계는 데이터 수집으로, 교사들이 입력한 설문 데이터를 엑셀 파일로 저장한 뒤 CSV 형식으로 변환하여 분석에 활용하였다. 두 번째 단계는 데이터 전처리로, 연구자들은 데이터를 정제하고 불용어를 제거한 후 유의어를 변환하여 분석 가능한 형태로 구성하였다. 불용어 제거에 대해 예를 들면 ‘이거’, ‘이름’, ‘습니다’와 같이 빈번하게 등장하지만, 분석 목적에 적합하지 않은 단어들을 불용어 목록에 포함하여 제외하였다. 유의어 처리와 관련하여서는 ‘아이들’과 같은 단어를 ‘학생’으로 변환하였다. 세 번째 단계는 데이터 마이닝 단계로, 이 단계에서 빈도 분석을 비롯하여 키워드 네트워크 분석과 n-gram 분석, 그리고 LDA 분석을 시행하였다. 마지막 단계에서는 이전 분석에서 도출된 결과를 질적 데이터와 비교하며, 연구의 의미를 해석하고 도출하였다.



[그림 1] 연구 절차

Cooper 등(2004)은 복잡한 상황에서 시행되는 교육정책은 다양한 관점에서 검토되어야 한다고 제안한다. 본 연구에서는 Cooper 등(2004)의 다차원 교육정책분석 모형을 차용하여 서술형 설문 문항을 설계하였다. 문항은 과정중심평가에 대한 이해(규범적 차원), 과정중심평가의 실행(기술적 차원), 실행에서 당면하는 문제점(구조적 차원), 실행에 필요한 여건 및 환경(구성적 차원)으로 구성하였다. 구체적인 서술형 문항 질문은 다음과 같다.

과정중심평가란 무엇이라 생각하는가? 과정중심평가를 실행하고 있는가? 그렇다면 어떠한 방식으로 실행하는가? 과정중심평가를 실행하고 있지 않다면 그 이유는 무엇인가? 과정중심평가를 실행하는 데 있어 어려운 부분은 무엇이라 보는가? 과정중심평가를 실행하기 위해 필요한 여건과 환경은 무엇이라 보는가?

응답 자료에 대한 기술통계치는 <표 3>에 제시되어 있으며, 해당 값은 텍스트 데이터 전처리 후의 결과를 반영한 것이다. 초등학생의 평균 응답 길이는 49.29로 나타났으며, 표준편차는 37.56이었다. 어휘 수는 646개, 전체 단어 수는 2,207개로 집계되었다. 중학생의 경우, 평균 응답 길이는 62.81이었으며, 표준편차는 36.31이었다. 어휘 수는 676개, 전체 단어 수는 2,146개로 확인되었다.

<표 3> 기술통계

	초등	중등
응답 수	143	112
응답 길이 평균	49.29	62.81
응답 길이 표준편차	37.56	36.31
어휘 수	646	676
총 단어 수	2,207	2,146

3. 키워드 네트워크 분석

텍스트 데이터를 활용하여 특정 주제와 관련된 핵심 키워드를 추출하고, 단어 간 동시 출현 빈도를 기반으로 관계를 분석하는 기법을 키워드 네트워크 분석이라고 한다. 동시 출현 관계는 특정 범위 내(예: 문장, 문서)에서 키워드가 함께 나타나는 것을 의미한다. 예를 들어, $T_1 \sim T_n$ 내의 텍스트에서 특정 키워드 집합 $k_1 \sim k_m$ 이 출현한다고 가정할 때, k_i 와 k_j 가 같은 텍스트에서 등장할 경우 두 키워드는 동시 출현 관계에 있다고 하며, 이때 텍스트 내에서 해당 키워드의 동시 출현 빈도는 C_{ij} 로 표시된다(이수상, 2014).

네트워크 구조에서 키워드들은 노드(node)로 표현되며, 이들 간의 관계는 링크(link)나 엣지(edge)로 나타낸다(백영민, 2023). 네트워크 분석에서는 이러한 링크가 동시 출현

관계를 나타내며, 이를 통해 키워드 간의 관계망을 시각화할 수 있다. 키워드 네트워크 내에서는 특정 노드들의 밀집된 집합을 커뮤니티(community)라고 한다. 커뮤니티는 연결성이 강한 노드의 그룹을 의미하며, 이는 단어 간 반복적인 상호작용을 나타낸다. 커뮤니티 분석은 각 그룹이 가지는 속성과 관계를 분석하고, 특정 주제나 맥락에 대한 이해를 돕는다(최진호 외, 2011; Girvan & Newman, 2002). 이 방법의 장점은 단어들 간의 관계를 구조적으로 파악할 수 있다는 점이다. 특히, 관련 개념 간의 상호작용을 통해 개념의 역할을 구체적으로 탐구할 수 있다(이수상, 2014).

본 연구에서는 키워드 간 동시 출현 데이터를 활용하여 네트워크를 구성하였다. 먼저 키워드 간 동시 출현 빈도를 토대로 연결 관계를 설정하였다. 이때 의미 있는 관계만 남기고자 연결망 내에서 단절된 키워드를 제거하고, 중복된 관계는 간소화하였다. 네트워크 내 주요 키워드를 탐색하기 위해 연결 중심성 지수(degree centrality)를 활용하였다. 연결 중심성은 특정 키워드가 다른 키워드와 맺고 있는 관계의 총합을 의미하며, 변수 간 관계에서 중심성 정도를 평가하는 중요한 척도로 활용된다(정승환 외, 2014; 홍세희 외, 2019). 네트워크 상에서 연결 중심성이 높은 키워드는 다른 키워드와의 관계망에서 핵심적인 역할을 수행하며, 의미 구조를 형성하는 주요 개념으로 작용한다. 동시출현빈도는 widyr 패키지의 pairwise_count() 함수를 사용하였으며, 분석 단위로는 각 사람이 3개의 질문에 응답한 답변을 합친 문서를 기준으로 삼았다. 따라서 동일 문서 내에서 함께 나타난 단어를 한 쌍으로 보고 동시출현빈도를 계산하였다.

본 연구에서는 Louvain 알고리즘을 분석 알고리즘으로 적용하였다. Louvain 알고리즘은 네트워크 내 자연스럽게 형성되는 관계 구조를 바탕으로, 밀접하게 연결된 키워드들을 하나의 커뮤니티로 묶고, 그렇지 않은 키워드는 느슨하게 결합하는 방식으로 배치한다. 이 알고리즘은 모듈성(Modularity) 값을 최대화하는 방향으로 커뮤니티를 탐지하며, 최적의 커뮤니티 구조를 도출한다. 모듈성은 네트워크 내 연결 밀도가 높은 집단과 낮은 집단을 구분하는 지표로, 값이 클수록 명확한 커뮤니티 구조를 지닌다는 것을 의미한다. 모듈성은 아래 식으로 정의되며, Louvain 알고리즘은 Clauset et al.(2004)의 그리디 알고리즘(greedy algorithm)을 기반으로 한다. 초기에는 각 노드를 개별 커뮤니티로 설정한 뒤, 인접 노드로 옮겨가며 모듈성 증가 폭을 평가한다. 이후 새롭게 형성된 커뮤니티들을 하나의 노드로 묶어 새로운 네트워크를 구성하고, 이 과정을 반복하다가 모듈성 값이 더 이상 증가하지 않으면 최종 커뮤니티 구조로 확정한다(Blondel et al., 2008).

$$M = \frac{1}{2m} \sum_{i,j} (A_{ij} - \frac{k_i k_j}{2m}) \delta(c_i, c_j)$$

- A_{ij} : 노드 i, j 를 연결하는 엣지의 가중치
- k_i : 노드 i 에 연결된 엣지의 수
- c_i : 노드 i 가 속한 커뮤니티
- m : $\frac{1}{2} \sum_{i,j} A_{ij}$, 네트워크의 총 엣지 가중치의 절반
- $\delta(u, v)$ 함수 : $u = v$ 일 때 1, 그렇지 않으면 0

4. n-gram 언어 모델

n-gram은 n개의 연속된 단어를 확률적으로 나타내는 기법이다. 이 방법은 단순성과 구현의 용이성을 장점으로 하지만, n의 값에 제한이 있어 긴 문장의 정보를 충분히 포착하는 데 한계가 있다. 문장은 단어들의 집합으로 구성되며, 하나의 문장에 대한 확률은 해당 문장에서 n개의 단어가 함께 등장할 확률을 의미한다. 구체적으로, 문장의 확률은 첫 번째 단어의 발생 확률, 첫 번째 단어가 발생한 조건에서 두 번째 단어의 발생 확률, 그리고 첫 번째와 두 번째 단어가 발생한 조건에서 세 번째 단어의 발생 확률을 순차적으로 곱하여 계산된다(김기현, 2019).

전통적인 통계 기반 언어 모델은 카운트를 기반으로 확률을 계산하는데, 문장의 길이가 길어질수록 해당 문장이 말뭉치(corpus)에서 나타나지 않을 가능성이 커진다. 또한, 반복된 곱셈으로 인해 확률값이 매우 낮아지는 문제가 발생할 수 있다. 이를 개선하기 위해, 다음 단어의 출현 확률을 추정할 때 전체 단어를 모두 고려하지 않고 특정 개수의 단어만 보고 확률을 계산할 수 있다. 이때, 기준이 되는 단어의 수는 n-gram에서 n-1에 해당한다. 예를 들어, n이 1일 경우에는 유니그램(unigram), 2일 경우에는 바이그램(bigram), 3일 경우에는 트라이그램(trigram)으로 불리며, n이 4 이상일 경우 해당 숫자를 앞에 붙여서 명명한다. 본 연구에서는 바이그램을 활용하여 연속된 두 단어의 관계를 분석하였다.

본 연구에서는 바이그램 분석을 활용하여 초·중등 교사들이 과정중심평가에 대해 인식하는 주요 개념들의 관계성을 파악하고자 하였다. 바이그램 분석은 연속된 두 단어의 동시 출현 빈도를 기반으로 단어 간의 연관성을 탐색하는 방법으로, 교사들의 응답에서 나타나는 개념 간의 의미적 연결망을 시각적으로 분석하는 데 유용하다. 이는 단어의 개별적 빈도 분석만으로는 포착하기 어려운 개념 간 상호작용과 잠재적 의미 구조를 드러낼 수 있는 강점을 가진다.

바이그램 분석의 타당성은 연구 문제와의 적합성에서 찾을 수 있다. 본 연구는 교사들의 인식 속에서 과정중심평가가 어떻게 개념적으로 구성되어 있는지를 탐색하는 것을 목적으로 한다. 바이그램 분석은 교사들이 실제로 사용하는 언어를 기반으로 개념 간 관계를 도출함으로써, 설문조사나 인터뷰와 같은 전통적 방법으로는 발견하기 어려운 잠재적 인식 구조를 밝히는 데 기여한다. 또한, 시각화를 통해 단어 간의 연결을 직관적으로 확인할 수 있어, 과정중심평가에 대한 교사들의 인식을 보다 심층적으로 분석할 수 있다.

5. LDA

토픽 모델링 기법은 문서 집합에서 숨겨진 주제를 찾아내는 과정으로, 대표적인 알고리즘으로는 LDA(Latent Dirichlet Allocation)가 있다. 경험적 베이즈 모수 모형(parametric empirical Bayes model)인 LDA는 디리클레(Dirichlet) 분포를 적용해 알고리즘의 복잡성을 줄이는 방식으로 설계된다. LDA에서는 다항 분포(multinomial distribution)가 우도 함수(likelihood function)로 사용되며, 토픽 비율에 대해 설정된 사전 분포(prior distribution)가 디리클레 분포의 켈레 분포(conjugate distribution)이기 때문에 사후 분포(posterior distribution) 또한 디리클레 분포를 따른다(유진은, 2021; Blei et al., 2003). 이러한 특성 덕분에 LDA는 복잡한 계산 과정을 효율적으로 처리할 수 있으며, 문서 내에서 특정 단어들이 주제를 나타내는 방식을 파악할 수 있게 한다.

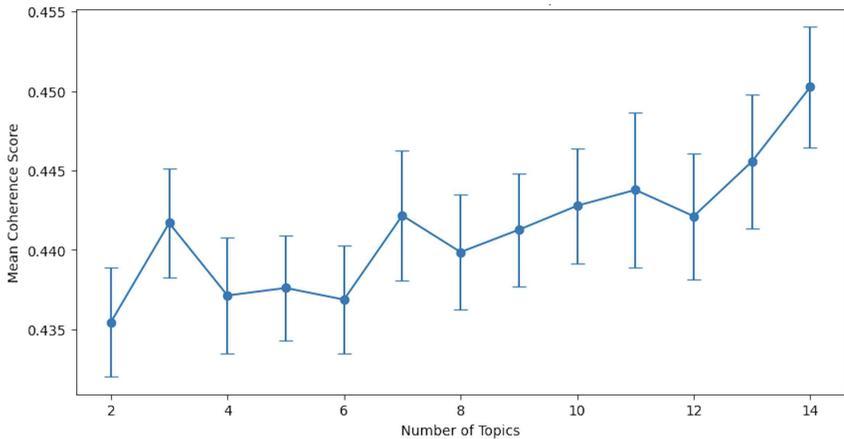
본 연구에서는 LDA를 수행하기에 앞서, 전처리된 초등 및 중등 텍스트 자료를 대상으로 파이썬의 Gensim 라이브러리에 내장된 coherence 함수를 활용하여 응집도(coherence) 값을 산출하였다. 응집도 값이 클수록 토픽이 의미론적으로 더 일관성이 있음을 의미한다(강현아, 임희석, 2021; Newman et al., 2010). 응집도는 100회 반복 산출하여 그 평균을 구하였으며, 토픽 수 2개에서 14개까지의 변화 양상을 95% 신뢰구간과 함께 제시하였다. 관련 결과는 [그림 2]와 [그림 3]에 제시하였다.

[그림 2]의 초등 텍스트 분석 결과, 토픽 수가 14개일 때 응집도 값이 가장 높게 나타났다. [그림 3]의 중등 텍스트 분석에서는 토픽 수가 13개일 때 가장 높은 응집도 값을 보였다. 그러나 이와 같이 토픽 수가 많아질수록 개별 토픽이 지나치게 세분화되어, 오히려 해석 가능성이 저하될 수 있다(유예림, 김지은, 2023; Cao et al., 2009). 본 연구에서는 응집도 결과를 참고하되, 텍스트 자료의 특성을 고려하여 최적의 토픽 수를 결정하였다.

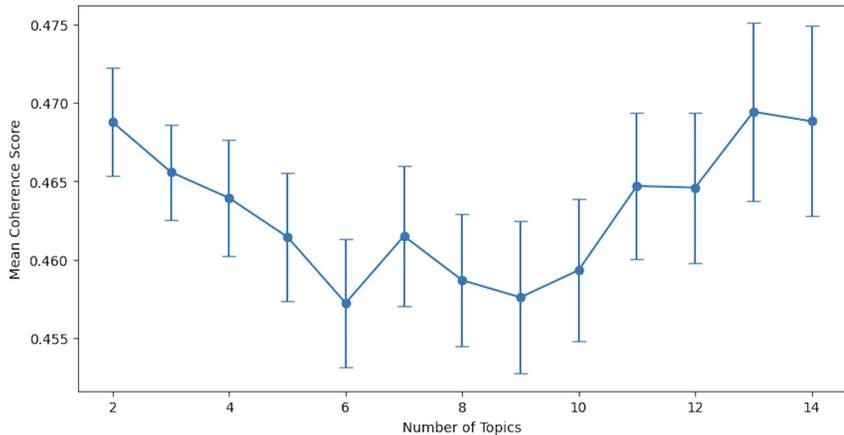
[그림 2]에서는 토픽 수가 3개와 7개일 때 상대적으로 높은 응집도 값을 나타냈으며, [그림 3]에서는 토픽 수가 2개에서 6개까지 점진적으로 감소하는 경향을 보였다. 이는

본 연구에서 활용된 질문이 총 3개로 구성되어 있고, 이 질문들을 기반으로 작성된 응답이 제한적이었음을 감안할 때 토픽의 다양성을 확보하기 어려운 특성과 관련이 있다. 이러한 점을 종합적으로 고려하여, 본 연구에서는 초등 및 중등 텍스트 모두에서 3개의 토픽을 최종적으로 선정하였다.

전처리가 완료된 데이터를 초등과 중등으로 구분하여, 각 데이터에 대한 문서-단어 행렬(Document-Term Matrix, DTM)을 생성하였다. 이후 DTM을 통해 LDA 모델을 적용하여 초등과 중등 데이터 각각에서 3개의 토픽을 도출하였다.



[그림 2] Topic Coherence Score(초등)



[그림 3] Topic Coherence Score(중등)

IV. 연구 결과

1. 빈도 분석

본 연구에서는 텍스트 데이터 내에서 명사, 동사, 형용사, 부사를 포함한 단어의 출현 빈도를 살펴보았다. 연구자들은 데이터를 정제하여 상위 20개의 단어를 초등과 중등으로 구분하여 <표 4>, <표 5>와 같이 제시하였다. 최상위 빈도 어휘는 초등의 경우 평가(239회)와 학생(238회)이며, 과정(106회), 수업(79회), 교사(70회), 결과(66회), 어려움(46회), 시간(44회), 과정중심평가(40회), 성장(40회)가 그 뒤를 이었다. 중등의 경우 평가(197회)와 학생(171회)가 최상위 빈도 어휘를 보였고, 과정(88회), 수업(66회), 교사(57회), 결과(54회), 학습(50회), 시간(48회), 과정중심평가(46회), 피드백(46회) 순으로 나타났다.

<표 4> 빈도 분석 결과(초등)

순위	단어	빈도	빈도 비율 ¹⁾	순위	단어	빈도	빈도 비율
1	평가	239	8.48	11	학습	38	1.35
2	학생	238	8.44	12	성취기준	33	1.17
3	과정	106	3.76	13	피드백	32	1.13
4	수업	79	2.80	14	도달	28	0.99
5	교사	70	2.48	15	학급	28	0.99
6	결과	66	2.34	16	관찰	26	0.92
7	어려움	46	1.63	17	교육	25	0.89
8	시간	44	1.56	18	방법	24	0.85
9	과정중심평가	40	1.42	19	학습과정	24	0.85
10	성장	40	1.42	20	학급당	23	0.82

<표 5> 빈도 분석 결과(중등)

순위	단어	빈도	빈도 비율	순위	단어	빈도	빈도 비율
1	평가	197	7.31	11	수행	40	1.48
2	학생	171	6.34	12	성장	35	1.30
3	과정	88	3.26	13	학습과정	33	1.22
4	수업	66	2.45	14	필요	30	1.11
5	교사	57	2.11	15	학습자	30	1.11

1) 특정 단어가 전체 응답에서 차지하는 비율을 백분율로 나타낸 값

순위	단어	빈도	빈도 비율	순위	단어	빈도	빈도 비율
6	결과	54	2.00	16	중심	26	0.96
7	학습	50	1.85	17	제공	24	0.89
8	시간	48	1.78	18	기준	22	0.82
9	과정중심평가	46	1.71	19	관찰	21	0.78
10	피드백	46	1.71	20	다양	20	0.74

2. 네트워크 분석

본 연구에서는 네트워크 분석으로 <표 6>, <표 7>과 같이 어휘 간 연결성을 파악하였다. 연결 중심성이 가장 높은 단어로 초등의 경우 평가와 학생이 나타났고, 이어서 과정, 수업, 교사, 결과, 어려움, 시간, 과정중심평가가 나타났다. 중등의 경우 연결 중심성이 높은 단어로는 평가, 학생, 과정, 결과, 시간, 피드백, 교사, 수업, 학습, 과정중심평가 순으로 나타났다. 네트워크 분석을 통해 도출한 결과를 그룹으로 묶어 시각화하면 [그림 4]와 [그림 5]와 같이 드러난다.

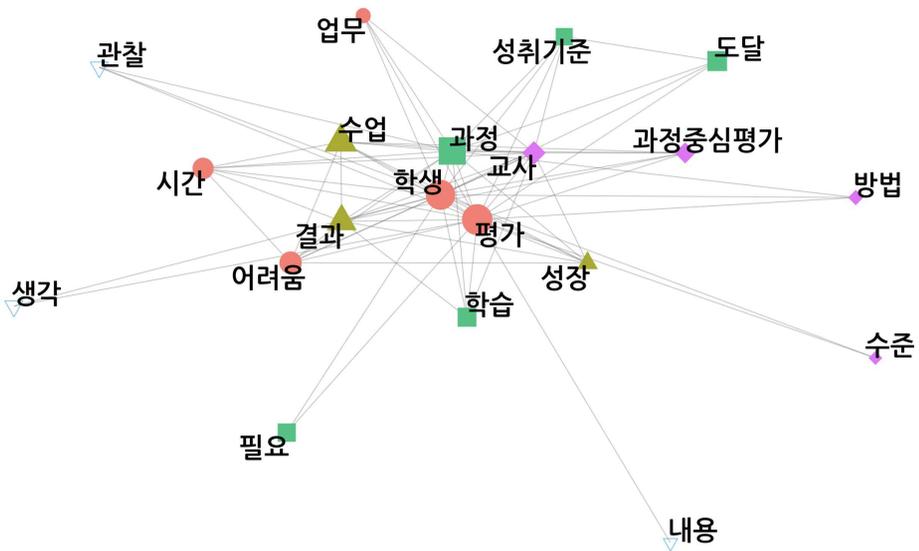
<표 6> 연결 중심성 분석 결과(초등)

순위	단어	중심성	그룹	순위	단어	중심성	그룹
1	평가	0.898	1	11	학습	0.445	3
2	학생	0.847	1	12	필요	0.430	3
3	과정	0.726	3	13	성취기준	0.410	3
4	수업	0.616	2	14	성장	0.382	2
5	교사	0.576	5	15	방법	0.381	5
6	결과	0.545	2	16	업무	0.379	1
7	어려움	0.544	1	17	수준	0.365	5
8	시간	0.507	1	18	생각	0.364	4
9	과정중심평가	0.498	5	19	관찰	0.361	4
10	도달	0.467	3	20	내용	0.351	4

<표 7> 연결 중심성 분석 결과(중등)

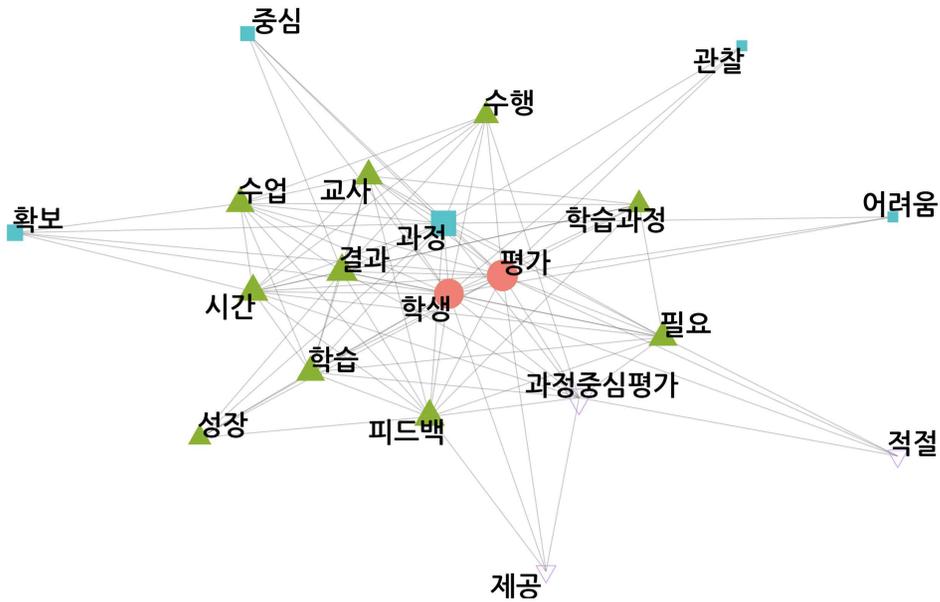
순위	단어	중심성	그룹	순위	단어	중심성	그룹
1	평가	0.939	1	11	필요	0.333	2
2	학생	0.818	1	12	학습과정	0.242	2
3	과정	0.515	3	13	수행	0.242	2
4	결과	0.424	2	14	성장	0.182	2
5	시간	0.394	2	15	확보	0.152	3

순위	단어	중심성	그룹	순위	단어	중심성	그룹
6	피드백	0.394	2	16	제공	0.121	4
7	교사	0.333	2	17	중심	0.121	3
8	수업	0.333	2	18	적절	0.121	4
9	학습	0.333	2	19	어려움	0.091	3
10	과정중심평가	0.333	4	20	관찰	0.091	3



[그림 4] 네트워크 연결망(초등)

초등 네트워크 분석 결과인 [그림 4]를 구체적으로 살펴보면, 가운데 ‘평가’와 ‘학생’이 위치하여 중요 키워드로 작용하고 있다. 평가와 학생은 과정, 도달, 성취기준, 피드백, 업무, 방법, 내용, 생각 등과 연결되어 있다. 중등의 네트워크 분석 결과인 [그림 5]에서도 ‘평가’와 ‘학생’이 가운데 위치하여 중요 키워드로 작용하고 있다. 이들은 결과, 과정, 시간, 교사, 피드백, 학습과정, 확보, 수행, 제공 등과 연결되어 과정중심평가를 시행하는데 필요한 요소들을 보여준다.



[그림 5] 네트워크 연결망(중등)

3. 바이그램 분석

본 연구는 n-gram 기법 중 바이그램을 활용하여 연속된 두 단어의 빈도를 분석하였다. 초등은 <표 8>, 중등은 <표 9>에서 각각 상위 20개 연속 단어 쌍을 보여준다. 여기서는 명사 뿐만 아니라 동사, 형용사, 부사도 함께 나타나 단어가 도출된 의미를 파악하기 쉽다. 초등의 경우, 과정-평가(28회), 학급-학생(12회), 학생-성장(12회), 수업-평가(11회), 평가-평가(11회), 평가-학생(11회), 학생-개개인(9회), 결과-과정(8회), 성취기준-도달(8회), 수업-과정(8회), 학급당-학생수(8회), 많다-학생(7회), 시간-부족(7회) 등이 상위 빈도 단어 쌍으로 나타났다. 중등의 경우, 피드백-제공(17회), 과정-평가(13회), 평가-평가(11회), 학생-변화(9회), 변화-성장(8회), 중심-평가(8회), 학생-학습(8회), 너무-많다(7회), 정확-평가(7회), 평가-학생(7회) 등이 상위 빈도 단어 쌍으로 나타났다.

바이그램은 연속하는 단어 쌍을 나타내므로 [그림 6], [그림 7]과 같이 노드로 연결된 단어를 함께 읽으면 교사들이 지닌 과정중심평가에 대한 인식을 알 수 있다. 초등 데이터인 [그림 6]에서는 ‘학생’과 ‘평가’이 상위 연속 빈도 단어로 드러난다. 먼저 ‘학생’은 학생-학급당, 학생-성장, 학생-개개인, 학생-많다, 학생-학급 등과 연결된다. 다음으로 ‘평가’는 평가-중심, 평가-생각, 평가-위하다, 평가-수업, 평가-과정, 평가-결과, 평가-학습과정과 연결된다. 이외에 시간-부족, 피드백-제공, 관찰-기록, 성취기준-도달 등이 연속 빈도 단어로 등장한다.

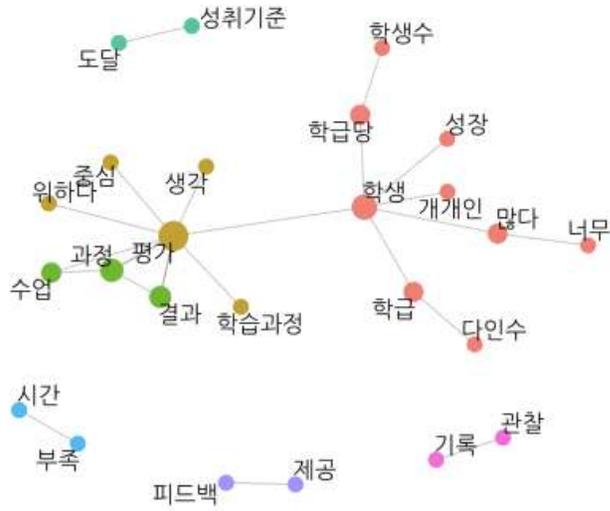
중등 데이터인 [그림 7]에서도 ‘학생’과 ‘평가’가 상위 연속 빈도 단어로 드러난다. 먼저 ‘학생’은 학생-변화, 학생-성장, 학생-학습과 연결된다. 다음으로 ‘평가’는 수업-평가, 평가-정확, 평가-중심, 평가-과정과 연결된다. 이외에 적절한-피드백, 피드백-제공, 너무-많다, 학생수-학급당이 연속 빈도 단어로 나타난다.

<표 8> 바이그램 연속 단어 빈도 결과(초등)

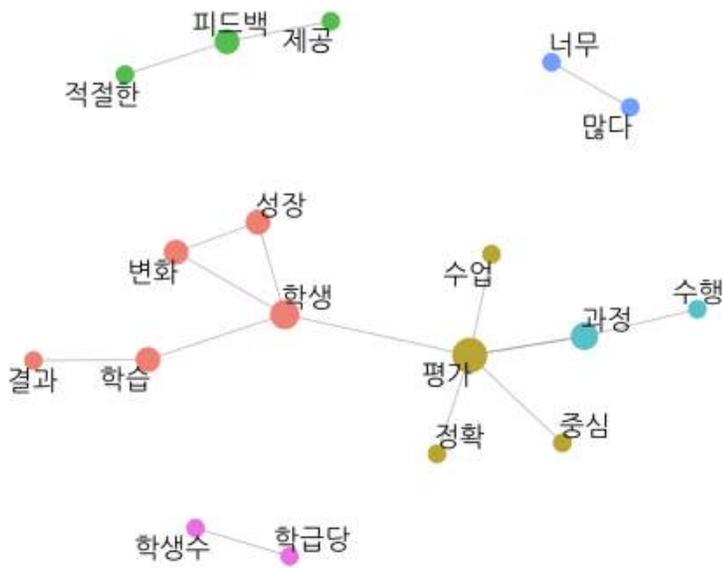
순위	단어1	단어2	빈도	순위	단어1	단어2	빈도
1	과정	평가	28	11	학급당	학생수	8
2	학급	학생	12	12	많다	학생	7
3	학생	성장	12	13	시간	부족	7
4	수업	평가	11	14	평가	결과	7
5	평가	평가	11	15	평가	과정	7
6	평가	학생	11	16	평가	생각	7
7	학생	개개인	9	17	피드백	제공	7
8	결과	과정	8	18	학급당	학생	7
9	성취기준	도달	8	19	결과	평가	6
10	수업	과정	8	20	관찰	기록	6

<표 9> 바이그램 연속 단어 빈도 결과(중등)

순위	단어1	단어2	빈도	순위	단어1	단어2	빈도
1	피드백	제공	17	11	수업	평가	6
2	과정	평가	13	12	수행	과정	6
3	평가	평가	11	13	적절한	피드백	6
4	학생	변화	9	14	평가	과정	6
5	변화	성장	8	15	학급당	학생수	6
6	중심	평가	8	16	학생	성장	6
7	학생	학습	8	17	학습	결과	6
8	너무	많다	7	18	결과	과정	5
9	정확	평가	7	19	결과	평가	5
10	평가	학생	7	20	과정	일어나다	5



[그림 6] 바이그램 연결망(초등)



[그림 7] 바이그램 연결망(중등)

4. 토픽 모델링

본 연구에서는 LDA 분석을 통해 초등과 중등 각각 세 가지 토픽을 추출하였다. [그림 8]과 [그림 9]는 각각 초등과 중등의 텍스트 데이터에서 추출한 토픽을 시각화한 것이다. 데이터 내에서 해당 단어나 나타날 확률이 높을수록 단어 막대의 길이가 길다. 본 연구에서는 토픽명 선정에 근거를 확보하기 위해 두 단계를 거쳤다. 먼저 토픽 Probabilities에 제시되는 Gamma 값이 높은 설문 응답 문장을 검토하였다. 다음으로 전문가들과의 상호 검증을 진행하였다. 연구자들이 대표 설문 응답 문장을 검토하여 초안을 작성한 뒤, 전문가 그룹의 검토와 자문을 거쳐 최종 토픽명을 확정하였다. 이들에게 각 토픽의 주요 단어와 이를 대표하는 설문 응답 문장을 제공하여 초안의 적절성과 일관성을 평가받았다. 전문가 그룹에 대한 자세한 정보는 <표 10>과 같다.

<표 10> 전문가 그룹 정보

대상	직급	학교급	교직경력	전공	학위
박○○	교사	초등	14년	교육평가	박사
박○○	교감	초등	22년	교육과정	박사
육○○	교사	초등	10년	교육과정	박사
이○○	교사	중등	18년	교육과정	박사

초등에서는 토픽 1에서 평가, 과정, 결과, 업무 등이 나타났다. 토픽 내 단어와 설문 본문을 고려하여 해당 토픽명을 ‘결과를 넘어 과정’으로 명명하였다. 토픽 2에서는 수업, 교사, 과정중심평가, 어려움, 도달 등이 등장하였고, 토픽명은 ‘교사의 어려움’으로 부여하였다. 토픽 3에서는 학생, 성장, 학습, 성취기준, 피드백 등이 나타났고, 토픽명은 ‘학생의 성장’으로 설정하였다. 아래는 초등교사의 대표 설문 응답 중 일부이다.

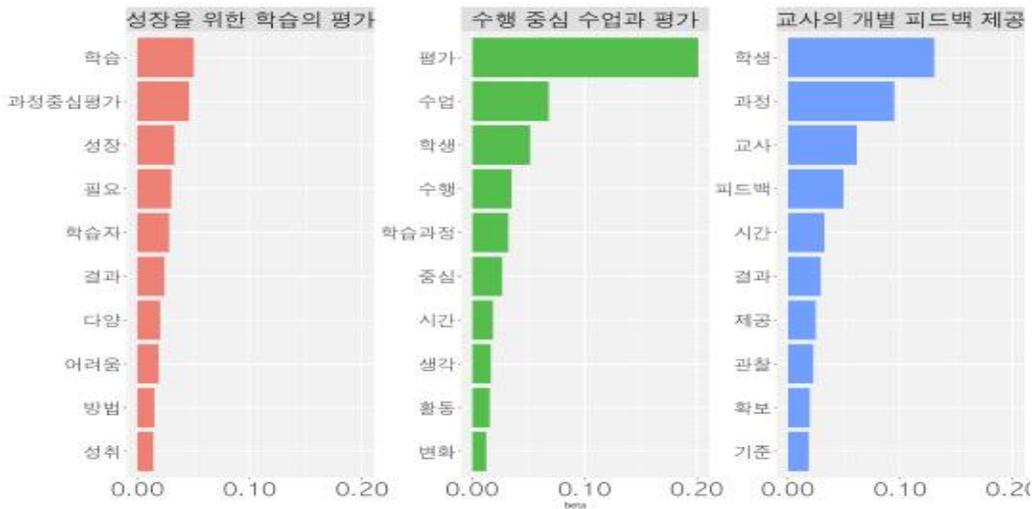
학급당 학생수를 조정하여 교사가 학생들을 관찰하고 실시간 피드백을 줄 수 있는 환경을 조성해야 한다. 현재처럼 한 반에 20명이 넘는 학생을 대상으로는 실시간 관찰 및 평가, 피드백은 사실상 불가능하다. (초등교사 #ET85)

중등에서는 토픽 1에서 학습, 과정중심평가, 성장, 필요, 학습자 등이 나타났고, 토픽명은 ‘성장을 위한 학습과 평가’로 정하였다. 토픽 2에서는 평가, 수업, 학생, 수행, 학습과정 등이 등장하였고, 토픽명은 ‘수행 중심 수업과 평가’를 부여하였다. 토픽 3에서는 학생, 과정, 교사, 피드백, 시간, 결과 등이 나타났고, 토픽명은 ‘교사의 개별 피드백 제공’으로 명명하였다. 아래는 중등교사의 대표 설문 응답 중 일부이다.

개별 학생의 수행 과정을 교사가 관찰하고 그에 맞는 피드백을 바로 제공하는 데에 양적 질적 한계가 있다. 특히 학습 과정에서 개별 학생에 대한 즉각적인 피드백을 제공하기가 어렵다. 문제는 학생들이 이러한 평가 방식을 선호하지 않으며, 객관식이 대부분인 집필평가에 비해 타당도와 신뢰도가 떨어진다고 생각하는 경우가 많다. (중등교사 #ST53)



[그림 8] LDA 분석 결과(초등)



[그림 9] LDA 분석 결과(중등)

V. 결론 및 제언

본 연구는 현장 교사들이 과정중심평가를 어떻게 이해하고 있고, 실제 어떠한 형태로 실천하고 있으며, 실행에 어려움을 겪는 이유가 무엇인지를 종합적으로 탐구해보고자 하였다. 이를 위해 텍스트 마이닝 기법을 활용해 과정중심평가에 대한 초·중등교사들의 인식을 분석하였다. 이를 통해 과정중심평가가 이루어지는 실제 현장의 평가 여건과 구조를 살펴보고, 과정중심평가가 일반화되기 위한 정책 방안을 드러내 보이려고 하였다.

연구 결과는 세 가지 논의점으로 정리할 수 있다. 첫째, 교사들이 과정중심평가를 수행하기에 직면하는 현실적 어려움이 드러난다. 빈도 분석에서 드러난 상위 빈도 어휘 목록에서 교사들이 과정중심평가를 이해하는 방식과 현실적 제약으로 인한 어려움을 짐작할 수 있다. 바이그램 분석에서도 너무-많다, 많다-학생, 시간-부족과 같은 부정적 의미의 연속 단어 쌍이 상위 빈도로 드러난다. 이는 선행연구에서 교사들이 과정중심 평가의 필요성에 공감하나(신혜진 외, 2017; 김진희; 2020; 김영실, 2022), 실행에 어려움을 겪고 있다는 결과(반재천 외, 2018; 고현, 2019)와 맥을 같이 한다.

주목할 점은 초등과 중등의 서로 다른 교육 환경 차이로 인해 실행에서의 어려움에 차이점을 보인다는 것이다. 초등교사들은 다양한 교과를 개별 학생에 맞게 평가해야 한다는 부담을 갖는다. 중등교사들은 평가의 객관성과 공정성, 적절한 피드백 방식에 고민을 갖고 있다. 이는 선행연구에서 상세히 드러내지 못한 초·중등 교원들의 실행에 따른 어려움의 차이를 보여준다. 바이그램 분석 결과 초등의 경우 학급-학생, 시간-부족과 같은 단어 쌍이 등장하는 것에 비해, 중등의 경우 피드백-제공, 정확-평가, 수행-과정과 같은 단어 쌍이 등장한다. 이는 학생 발달 단계에 따른 평가 목적의 차이와 입시와의 연계 정도에 의한 차이라고 보인다. 입시의 부담이 덜하고, 학생 개개인의 성장에 집중할 수 있는 초등에서는 중등에 비해 결과중심평가의 비중을 줄일 수 있다. 다만, 교사 1인이 다양한 교과를 동시에 가르쳐야 하는 초등교사가 모든 교과에서 학생 개개인의 특성을 반영한 평가를 실시하는 것이 쉽지 않다고 여겨진다. 상대적으로 평가의 객관성과 공정성이 강조되고, 학생 상호 간의 경쟁 의식이 강한 중등에서는 교사의 피드백으로 인한 평가 결과의 변화까지 민감한 사안이 될 수 있다.

학생의 학습과정을 매시간 관찰하여 기록을 남기고 평가를 하는 것이 실제로 다양한 과목을 수업하면서 병행하기에 교사 입장에서 굉장히 번거롭고 부담스러운 작업이다. (초등교사 #ET85)

개인 학생의 기준이 다르기 때문에 내신성적과 연계했을 때 평가의 주관성과 불공정성이라는 문제점이 있음. (중등교사 #ST76)

둘째, 교사들은 과정중심평가의 개념을 다양한 방식으로 이해하고 있다. 네트워크 분석 결과, 초·중등 모두 ‘평가’와 ‘학생’이 가운데 위치하였고, 두 키워드가 과정, 도달, 성취기준, 피드백, 업무, 변화, 개개인, 중심 등과 연결되었다. 설문 응답 내용에서도 일부 교사들은 과정중심평가를 학생의 학습 과정 전반을 포괄해 이해하는 반면, 일부 교사들은 피드백을 제공하거나 변화 과정을 관찰하는 방식으로 한정해 이해하는 것을 알 수 있다. 선행연구에서 과정중심평가의 개념 혼란을 지적했듯이(신혜진 외, 2017; 김정민, 2018; 정혜진, 지은림, 2020), 교사들 간의 합의된 인식을 위해 명확한 개념 정의와 안내가 필요하다.

수업의 과정 중 평가가 이루어지고 학생들에게 피드백을 제공하는 것 (초등교사, #ET59)

과정중심평가는 학생의 학습과정과 성장을 중심으로 평가하는 교육평가 방법입니다. 이 방법은 교육과정의 성취기준에 기반하여 교수와 학습 활동 중에서 학생의 변화와 성장에 대한 자료를 다양한 방법으로 수집하고, 이를 토대로 적절한 피드백을 제공하는 평가를 의미합니다. (중등교사 #ST101)

주목할 점은 유사 개념들 간의 혼란 뿐만 아니라, 과정중심평가란 개념의 범위로 인한 혼란에서도 원인을 찾을 수 있다는 것이다. 협의의 시각에서 과정중심평가는 결과중심평가와 대비되는 학습 과정 중에 이루어지는 평가로 이해된다. 그러나 광의의 시각에서 과정중심평가는 시험으로 사회를 통제하거나 사회적 이동을 위한 수단으로 삼았던 평가관으로부터 벗어나 학생의 성장을 도모하는 학습과 평가 과정 전반으로 이해될 수 있다. 전자에서 과정중심평가와 대비되는 결과중심평가가 후자에서는 과정중심평가가 이루어지기 위한 하나의 요소가 될 수 있는 것이다. 다만, 이러한 과정중심평가의 다층적 분석은 개념을 깊이 있게 이해하는 데 도움을 줄 수 있으나, 학교 현장에 일관된 인식을 심어주기에 한계를 지닌다.

셋째, 교사들이 ‘학습 결과로서의 평가’에 대비되는 ‘학습을 위한 평가’에 대한 이해도는 높으나, ‘학습 과정으로서의 평가’에 대한 인식이 부족할 수 있다는 점을 보여준다. 2022 개정 교육과정 총론에서는 ‘학습 결과만이 아니라 결과에 이르기까지의 학습 과정을 확인하고 환류해야 한다’는 ‘학습을 위한 평가’ 개념과 ‘학생이 자신의 학습 과정과 결과를 스스로 평가할 수 있는 기회를 제공해야 한다’는 학습 과정으로서의 평가 개념이 동시에 들어있다(교육부, 2022: 12). 그러나 본 연구의 분석 결과, 교사들이 ‘학습을 위한 평가’에 대한 이해가 높은 것에 반해, ‘학습 과정으로서의 평가’에 관한 인식이 부족한 것으로 드러났다. 학습으로서의 평가는 평가의 주체를 교사에서 학생으로 확장하고, 학생이 성찰을 통해 자신의 학습 상황을 점검하고 조정해나갈 수 있도록 평가를 설계하는 방식이므로, ‘메타인지’, ‘성찰’, ‘자기평가’, ‘스스로’와 같은 키워드가 관련되나

주요하게 등장하지 않았다. 설문 작성 내용에서도 학습 과정으로서의 평가에 관한 인식을 찾아보기 어려웠다.

이상의 연구 결과를 바탕으로 두 가지 정책을 제언한다. 하나는 초등과 중등의 서로 다른 교육 환경 차이를 이해한 맞춤형 정책 지원이 필요하다는 것이다. 초등의 경우 다양한 교과를 모두 평가해야 한다는 어려움을 드러내고, 중등은 개별 학생에게 맞춤형 피드백을 제공하는 방식에 대한 고민과 객관성을 담보한 평가 방식에 대한 어려움이 드러난다. 초등의 경우 과정중심평가의 실행에 관한 충분한 현장 연수가 필요하다. 범 교과 개념을 중심으로 교과를 통합해 평가하는 등 변화된 평가 방식을 직접 경험할 수 있는 기회가 제공되어야 한다. 선도교사를 중심으로 지역별 실행 연수를 활성화하고, 우수 사례를 발굴해 온라인으로 공유하며, 구체적인 평가 과정을 워크북 형태로 제공하는 등 개별 교사의 평가 전문성을 증진할 정책 지원이 필요하다. 중등의 경우, 디지털 기술을 활용해 객관성과 공정성을 담보하면서 적절한 개별 피드백을 제공할 방안을 강구해야 한다. 최근 현장에는 AI를 활용해 학생의 서술형 평가 결과를 분석하고 맞춤형 피드백을 제공하는 연구가 활발히 이루어지고 있다. 관련 연구회 등을 통해 우수 사례를 모델링하고, 현장에 널리 적용될 수 있도록 지원할 필요가 있다.

다른 하나는, ‘학습 과정으로서의 평가’의 의미를 담은 명확한 개념 정의와 안내가 필요하다는 것이다. 개념에 관한 혼란을 줄이기 위해, 국가(교육부)나 지역(교육청) 차원에서 과정중심평가에 대한 명확한 정의를 내리고, 일관적인 정책을 시행해나갈 필요가 있다. 특히 국가 교육과정이 지향하는 온전한 과정중심평가가 학교 현장에 일반화되기 위해서는 ‘학습 과정으로서의 평가’에 대한 현장 인식을 증진해 나가야 한다. 이는 2022 개정 교육과정이 강조하는 학습자 주도성과도 관련된다. 학생이 자신의 학습 과정과 결과를 스스로 평가할 수 있도록 평가가 설계될 경우, 교사의 교수와 학생의 학습 역시 학습자가 주도하는 방식으로 이어질 것이다. 이를 위해 ‘학습을 위한 평가’를 넘어 ‘학습 과정으로서의 평가’ 관점이 현장에 시급히 안내되어야 한다.

끝으로 본 연구는 과정중심평가에 관한 교사들의 전반적인 인식을 탐구한 것으로, 세부 응답 내용에 관한 실제 사례 분석이나 제안한 정책의 구체적인 구현 방안에 대한 논의가 부족하다. 과정중심평가를 우수하게 실천하는 교사들의 사례 분석이나 초·중등교사의 실제 운영 간 차이 분석, 제안된 정책의 구체적 실현 방안 제안 등이 연구된다면 과정중심평가가 우리나라 학교 현장에 정착하는데 도움이 될 것이라 생각한다. 이러한 연구들이 후속 연구로 이어지길 기대한다.

참고문헌

- 강현아, 임희석(2021). 토픽모델링 및 주성분 분석 기반 검색 질의 유형 분류 연구. 정보처리학회논문지. 소프트웨어 및 데이터 공학, 10(6), 223-234. <https://doi.org/10.3745/KTSDE.2021.10.6.223>
- 고현(2019). 교사별 과정중심평가에 대한 인식 및 실태와 지원방안 탐색. 학습자중심교과교육연구, 19(9), 1137-1164. <https://doi.org/10.22251/jlcci.2019.19.9.1137>
- 곽민호, 서영진(2020). 잠재디리클레할당 (LDA) 을 활용한 대학생의 동료평가 기준에 관한 탐색적 연구. 교양교육연구, 14(6), 323-337. <https://doi.org/10.46392/kjge.2020.14.6.323>
- 교육부(2015). 2015 개정 교육과정 총론. 세종: 교육부.
- 교육부, 한국교육과정평가원(2017). 과정을 중시하는 수행평가 어떻게 할까요? 서울: 교육과정평가원(ORM 2017-19-2).
- 교육부(2022). 2022 개정 교육과정 총론. 세종: 교육부.
- 김기현(2019). 김기현의 자연어 처리 딥러닝 캠프 (파이토치 편): 딥러닝 기반의 자연어 처리 기초부터 심화까지. 서울: 한빛미디어.
- 김석우(2021). CBAM 모형을 활용한 중학교 교사의 과정중심평가 관심도 및 실행도 분석. 학습자중심교과교육연구, 21(4), 527-549. <https://doi.org/10.22251/jlcci.2021.21.4.527>
- 김석우, 장카이, 서원석(2021). 중학교 교사가 지각하는 과정중심평가에 대한 IPA 분석. 학습자중심교과교육연구, 21(17), 429-441. <https://doi.org/10.22251/jlcci.2021.21.17.429>
- 김순남, 강이희, 김병찬, 박삼철, 유진은, 이은송, 전명남, 조훈희(2013). 창의인재 육성을 위한 학생평가 정책 연구: 국제 사례를 중심으로. 서울: 한국교육개발원(OR 2013-09).
- 김영실(2022). 초등학교 교사들의 과정중심평가 실천 방식과 수행 유형 분석. 교육과정평가연구, 25(2), 139-166. <https://doi.org/10.29221/jce.2022.25.2.139>
- 김유정, 장원형, 홍훈기(2019). 텍스트 네트워크 분석법을 활용한 2015 과학과 교육과정 평가 분석 및 논의: 과정 중심 평가와의 연계를 중심으로. 교육과정평가연구, 22(3), 225-250. <https://doi.org/10.22799/jce.2019.22.3.009>
- 김정민(2018). 과정중심평가의 개념과 교육적 의의 탐색. 학습자중심교과교육연구,

- 18(20), 839-859. <https://doi.org/10.22251/jlcci.2018.18.20.839>
- 김진희(2020). 대학교 교수의 과정중심평가에 대한 인식 및 실태. *학습자중심교과교육연구*, 20(10), 1283-1305. <https://doi.org/10.22251/jlcci.2020.20.10.1283>
- 박도순(1995). 교육평가 분야 : 연구의 최근 동향과 수행평가의 과제. *중앙교육진흥연구소*, 7(7), 120-128.
- 박정(2019). 과정중심평가를 위한 학생 자기평가 의미 탐색. *교육평가연구*, 32(3), 421-440. <https://doi.org/10.31158/JEEV.2019.32.3.421>
- 박지현, 진경애, 김수진, 이상아(2018). 과정 중심 평가 내실화 방안 연구. 한국교육과정평가원 연구보고서(ORM 2018-49).
- 박지현, 진경애, 김수진, 이상아(2020). 단위학교의 과정 중심 평가 운영 지표 개발 연구. *교육과정평가연구*, 23(2), 157-181. <https://doi.org/10.29221/jce.2020.23.2.157>
- 반재천, 김선, 박정, 김희경(2018). 교사별 과정 중심 평가에 대한 교사의 인식. *교육과정평가연구*, 21(3), 105-130. <https://doi.org/10.22799/jce.2018.21.3.005>
- 백영민(2023). R 기반 네트워크 분석. 서울: 한나레아카데미.
- 신보미, 이경언(2018). 교사 학습 공동체의 과정중심 평가도구 검토 사례 분석. *교육과정평가연구*, 21(4), 73-98. <https://doi.org/10.29221/jce.2018.21.4.73>
- 신은혜(2022). 과학 교사의 과학·AI융합교육 실천 동기와 인식. *현장과학교육*, 16(3), 396-410. <https://doi.org/10.15737/ssj.16.3.202208.396>
- 신혜진, 안소연, 김유원(2017). 과정 중심 평가 활용의 정책적 분석 - 서울특별시 소재 중학교 교사의 수행평가 활용 사례를 중심으로-. *교육과정평가연구*, 20(2), 135-162. <https://doi.org/10.22799/jce.2017.20.2.006>
- 유예림, 김지은(2023). 토픽 모델링을 활용한 「교육평가연구」 학술지의 연구 동향 분석. *교육평가연구*, 36(3), 553-579. <http://dx.doi.org/10.31158/JEEV.2023.36.3.553>
- 유진은(2021). (AI시대) 빅데이터 분석과 기계학습. 학지사.
- 이수상(2014). 언어 네트워크 분석 방법을 활용한 학술논문의 내용분석. *정보관리학회지*, 31(4), 49-68. <http://dx.doi.org/10.3743/KOSIM.2014.31.4.049>
- 장윤선(2023). LDA 토픽모델링을 활용한 국내 초·중등학생의 자기평가 연구동향 분석. *교육과정평가연구*, 26(3), 295-318. <https://doi.org/10.29221/jce.2023.26.3.295>
- 정혜진, 지은림(2020). 교실평가에서 이루어지는 과정중심평가의 특성과 의미: 서울시 소재 중학교 사례를 중심으로 한 질적연구. *교육평가연구*, 33(4), 915-944. <https://doi.org/10.31158/JEEV.2020.33.4.915>

- 정승환, 황예담, 송영수(2014). 핵심어 네트워크 분석(network analysis)을 통한 국내 HRD 연구동향 탐색. *HRD연구*, 16(3), 1-33. <https://doi.org/10.18211/kjhrdq.2014.16.3.001>
- 전현욱(2019). 과정중심평가 만들어 가기에 관한 참여적 실행연구: 세종시 사과초등학교를 중심으로. *학습자중심교과교육연구*, 19(24), 695-725. <https://doi.org/10.22251/jlcci.2019.19.24.695>
- 전현욱, 이형연(2019). 과정중심평가 수행 방식에 관한 문화기술적 사례연구: 경기도 초록초등학교를 중심으로. *학습자중심교과교육연구*, 19(8), 123-154. <https://doi.org/10.22251/jlcci.2019.19.8.123>
- 최진수, 정혜원(2024). 텍스트 마이닝을 활용한 국내외 프로세스 데이터 연구 동향 분석. *교육과정평가연구*, 27(3), 197-221. <https://doi.org/10.29221/jce.2024.27.3.197>
- 최진호, 김희수, 임남규(2011). 기술예측을 위한 특허 키워드 네트워크 분석. *한국지능시스템학회논문지*, 17(4), 227-240.
- 최훈원, 최윤정(2024). 키워드 네트워크 분석을 활용한 과정중심평가의 연구동향 분석. *교육과정평가연구*, 27(2), 251-277. <https://doi.org/10.29221/jce.2024.27.2.251>
- 홍세희, 조기현, 이현정, 손수경, 김효진, 윤미리, 강윤경(2019). 키워드 네트워크 분석을 활용한 청소년의 회복탄력성 연구동향. *한국청소년연구*, 30(2), 273-301. <https://doi.org/10.14816/sky.2019.30.2.273>
- 홍수향, 장인실, 김태선(2017). 합의적 질적 연구를 통한 초등교사의 과정중심평가에 대한 인식. *교육과정연구*, 35(4), 47-69.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, 72(7), 1775-1781. <https://doi.org/10.1016/j.neucom.2008.06.011>
- Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70(6), 066111. <https://doi.org/10.1103/PhysRevE.70.066111>

- Cooper, B. S., Frusarelli, L. D., & Randall, E. V. (2004). *Better policies, better schools: Theories and applications*. Boston: Allyn and Bacon.
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning, second edition*. 온정덕, 윤지영 역(2022). *학습 과정으로서의 평가: 교실 평가로 학습 극대화*. 서울: 학지사.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821 – 7826. <https://doi.org/10.1073/pnas.122653799>
- Kumar, A., & Paul, A. (2016). *Mastering text mining with R: Extract and recognize your text data*. Packt Publishing.
- Newman, D., Lau, J. H., Grieser, K., & Baldwin, T. (2010). Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics* (pp.100–108).

논문접수 : 2025. 1. 3. / 수정본접수 : 2025. 2. 4. / 게재승인 : 2025. 2. 11.

ABSTRACT

A study of elementary and secondary teachers' perceptions of process-focused assessment through text mining

Kwon, Soonbo

Part-time lecturer, Korea National University of Education

Park, Sue won

Teacher, Poongyang Elementary School

This study examines the perceptions of elementary and secondary school teachers regarding process-focused assessment. A total of 255 teachers participated in an open-ended survey, and the collected data were analyzed using text mining techniques. The findings indicate that teachers perceive process-focused assessment as an evaluation that takes place during the learning process and view it as a tool to support individual student growth. However, practical constraints were identified in the implementation process. Elementary school teachers reported difficulties in assessing multiple subjects, while secondary school teachers highlighted challenges related to ensuring objectivity and providing feedback. Additionally, there was considerable variation in teachers' understanding of the concept of process-focused assessment, leading to confusion about its definition and execution. Based on the findings, this study suggests the need for policy support tailored to the distinct educational environments of elementary and secondary schools, as well as the establishment of a clear definition and guidelines for process-focused assessment.

Key Words: Process-focused Assessment, Text Mining, Elementary Teachers, Secondary Teachers