

성장중심평가의 취지에 따른 중등학교 평가혁신 실행연구

이형빈 (가톨릭관동대학교 교수)*

요약

이 연구에서는 “입시의 영향력이 강한 중등학교에서는 평가혁신이 어렵다.”라는 암묵적인 비관주의를 극복하고, 현재의 조건 속에서 실현 가능한 중등학교 평가혁신의 범위는 어느 정도이며 이를 일반화하기 위해서는 무엇이 필요한지를 탐색하였다. 이를 위해 연구자는 A중학교 교사들과 일 년 동안 실행연구 계획을 수립하고, 실행 중 성찰을 거듭하였다.

실행연구에 참여한 교사들이 실천한 구체적인 과제는 첫째, 선다형 지필평가 축소하기, 둘째, 수행평가를 진단평가-형성평가-총괄평가의 흐름으로 재구조화하기, 셋째, 피드백과 재도전이 있는 평가 시행하기, 넷째, 핵심 아이디어와 핵심 질문을 중심으로 수업과 평가를 설계하기 등이다. 이를 바탕으로 평가계획 양식과 평가 루브릭을 새롭게 개발하였다. 또한 실행연구 결과를 바탕으로 <A중학교 평가의 원칙과 방향>을 정리하여 모든 교사와 공유하였다.

이러한 실행연구를 다음과 같은 시사점을 얻을 수 있다. 첫째, 평가혁신을 위해서는 학교 공동체적 차원의 실천이 필요하다. 둘째, 평가혁신을 위해서는 이론적 학습과 학교의 실정에 따른 실행연구가 필요하다. 셋째, 평가혁신을 위해서는 학교 관리자의 리더십이 매우 중요하다. 넷째, 평가혁신의 성과가 일반화되기 위해서는 매개적 도구의 역할이 중요하다. 다섯째, 평가혁신을 위해서는 ‘방어적 교육활동’의 풍토에서 벗어나 ‘적극적 교육활동’의 문화를 구축해야 한다. 교육당국은 이를 지원하기 위해 관련 지침을 개정하고, 평가혁신의 의미를 공유하기 위한 여건을 조성해야 한다.

주제어: 평가혁신, 성장중심평가, 실행연구

* 제1저자, party21@hanmail.net

I. 서 론

중등학교에서 평가혁신은 불가능한가? “평가는 학생 개개인의 교육 목표 도달 정도를 확인하고, 학습의 부족한 부분을 보충하며, 교수·학습의 질을 개선하는 데 주안점을 둔다.”(교육부, 2022)라는 국가 교육과정의 규정은 선언적인 문구일 따름인가? 대입의 영향력에 따라 상대평가가 적용되는 고등학교는 논외로 하더라도, 절대평가가 적용되고 있는 중학교에서도 여전히 구태의연한 평가가 이루어지는 이유는 무엇인가?

그동안 교육과정-수업-평가 일체화, 과정중심평가, 성장중심평가 등 평가혁신에 대한 수많은 담론이 쏟아져 나왔다. 하지만 여전히 중등학교 현장에서는 선다형 지필평가가 중심을 이루고 있다. 수행평가 역시 대체로 교수학습과 긴밀하게 연결되지 못한 채 분절적으로 이루어지고 있다. 서·논술형 평가도 고등사고 능력을 평가하기보다는 이른바 ‘무늬만 서·논술형 평가’에 불과하다는 비판에서 벗어나지 못하고 있다(김경희 외, 2024). 교사들은 변별력 확보, 민원 예방, 진도에 대한 부담, 학생수에 따른 부담 등으로 인해 평가혁신의 취지를 살리기 어렵다는 반응을 보이고 있다(신혜진, 안소연, 김유원, 2017; 임태원, 이형빈, 2024).

하지만 이러한 상황에서도 평가혁신을 위해 노력하는 교사도 많다. 문제는 평가혁신을 위한 노력이 개별 교사 차원에 머무는 경우가 대부분이라는 점이다. 평가혁신은 학교 공동체 차원에서 모든 교원이 그 철학과 원리를 공유하는 가운데 이루어져야 한다. 그래야 입시 위주 교육의 오랜 관행을 넘어설 수 있으며, 학생과 학부모도 평가혁신의 취지를 이해하게 된다.

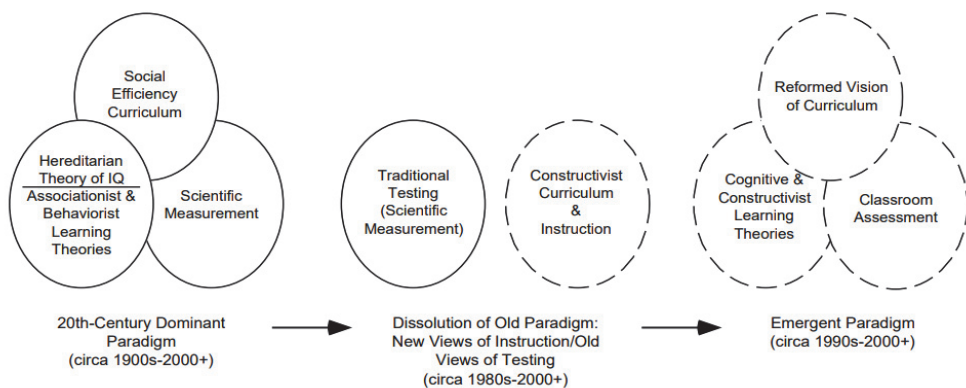
연구자는 이러한 문제의식을 바탕으로 A중학교 교사들과 함께 협력적 실행연구를 진행하였다. 성장중심평가의 취지에 따른 세부 과제를 선정하였고, 이를 실행에 옮기며 그 과정과 결과를 성찰하였다. 이를 통해 “입시의 영향력이 강한 중등학교에서는 평가혁신이 어렵다.”라는 암묵적인 비판주의를 극복하고, 현재의 조건 속에서 실현 가능한 중등학교 평가혁신의 범위는 어느 정도이며 이를 일반화하기 위해서는 무엇이 필요한지를 탐색하였다.

II. 이론적 배경 및 선행연구

1. 평가혁신의 난점

그동안 우리나라 학교에서 교육과정-수업-평가혁신을 위한 노력이 활발히 이루어졌다. 하지만 교육과정과 수업의 변화에 비해 평가의 변화는 상대적으로 더디게 이루어졌다. 수업의 방식은 학생 참여형 수업으로 바뀌었으나, 평가는 여전히 일제식 지필평가 위주이거나, 수행평가를 하더라도 교수학습의 과정과 자연스럽게 연계되지 못하는 경우가 많다. 즉, 수업은 달라졌으나 평가는 변하지 않는, ‘수업과 평가의 분리’ 현상이 지속되고 있다.

이는 비단 우리나라의 문제만은 아니라 여러 나라의 보편적 현상이다. 이를 Shepard(2000)는 [그림 1]로 설명한 바 있다.



[그림 1] 새로운 수업과 전통적 평가 사이의 불일치

Shepard(2000)에 의하면 교육과정-수업-평가 패러다임의 첫 번째 단계는 ‘사회 효율적 교육과정’, ‘행동주의 학습 이론’, ‘과학적 측정’이 결합된 전통적인 방식이다. 두 번째 단계에서는 교육과정과 수업은 구성주의적 관점으로 변하지만, 평가는 여전히 전통적인 방식으로 남아 있다. 이를 ‘새로운 수업과 전통적 평가 사이의 불일치’라고 할 수 있다. 세 번째 단계는 ‘혁신적 교육과정’, ‘구성주의적 학습 이론’, ‘교수학습과 연계된 평가’가 연결되어 있다. 첫째 단계와는 달리 셋째 단계에서는 각각의 영역이 점선으로 표시되어 있는데, 이는 교육과정-수업-평가의 경계가 명확하지 않고 서로 연계되어 있음을 나타낸 것이다.

우리나라의 중등학교는 대체로 두 번째 단계에 있는 것으로 보인다. 혁신학교 등 일부에서는 세 번째 단계를 지향하고 있으나, 대부분의 학교에서는 ‘수업은 달라졌으나 평가는 바뀌지 않은 상황’이라고 할 수 있다. 이러한 현상을 보고한 선행연구는 상당히 축적되어 있다. 혁신학교에서의 교육과정-수업-평가 실태를 종합적으로 연구한 이주연 외(2020)에 의하면 대부분의 혁신학교에서는 수업의 새로운 변화가 이루어지고 있으나 이것이 평가로 연결되는 정도는 낮은 것으로 나타났다.

중요한 것은 이러한 현상의 원인을 분석하고 해결방안을 제시하는 것이다. 이 원인에 대한 분석은 여러 연구물을 통해 종합적으로 확인할 수 있다. 박혜영, 이명애, 이명진(2019)에 의하면 전문가들은 평가혁신을 방해하는 요인으로 ‘교사의 평가 전문성 부족’, ‘교사의 과중한 업무 부담’, ‘대학 입시 정책’, ‘평가의 공정성에 대한 맹신’을 들었다. 신혜진, 안소연, 김유원(2017), 임태원, 이형빈(2024), 권순보, 박수원(2025)에 의하면 교사들은 ‘평가 및 채점 시간 확보의 어려움’, ‘평가의 타당성과 객관성 확보의 어려움’, ‘평가 결과에 대한 민원 발생의 두려움’, ‘평가 결과에 대한 피드백의 어려움’ 등을 평가혁신의 저해 요인으로 여겼다. 이러한 문제를 해결하는 방안으로는 전문가나 현장 교사 모두 ‘평가 전문성 신장’, ‘교사 업무 경감’, ‘교육당국의 연수 및 자료 지원’, ‘입시 제도 개선’ 등을 제시하였다.

하지만 이러한 원인 진단 및 대안이 반복되다 보니 이제 다소 진부한 느낌마저 들게 한다. “입시 때문에, 학생 수가 많아, 민원이 제기될까 봐 어쩔 수 없다.”라는 현장 교사들의 하소연은 상당 부분 사실임이 틀림없겠지만, 다른 한편으로는 자기 합리화를 위한 상투어 같은 인상마저 준다. 입시에서 일정 정도 자유로운 중학교라든가 학생 수가 적은 소규모 학교에서는 여타의 학교에 비해 평가혁신이 어느 정도 이루어지고 있다는 근거도 확인하기 어렵다.

평가혁신은 시간의 흐름에 따라 자연스럽게 이루어지는 것이 아니다. 모든 변화는 ‘객관적 조건의 형성’과 함께 ‘주관적 인식의 변화’ 그리고 ‘구체적인 실천’이 필요하다. 하지만 아쉽게도, 중등학교 평가혁신 실천 사례에 대한 선행연구는 다른 분야에 비해 상대적으로 빈약할 뿐만 아니라 그마저도 일부 교과 혹은 개인적 차원에서의 실천 사례가 대부분이다(양이레, 2018; 육기엽, 2016; 정현철, 2018). 중등학교에서 학교 공동체 차원으로 진행된 평가혁신 실천 사례에 대한 연구는 거의 찾아보기 어렵다.

이에 이 연구에서는 ‘평가혁신을 위한 학교 단위 공동체적 실천’에 주목하고자 한다. 초등학교와 달리 ‘교사별 평가’가 불가능한 중등학교에서는 학교 전체가 평가의 철학과 방향을 공유하고 이를 함께 실천해야만 평가혁신이 가능하다. 설사 교사별 평가가 적용되더라도, 평가는 학교의 공동체적 문화(예를 들어, 경쟁적 문화와 협력적 문화)와 직결되는 문제이기 때문에 단위 학교 차원의 총체적 접근이 필요하다. 또한 평가혁신은 ‘제

도의 문제'이자 '실천의 문제'이다. 제도의 변화는 상대적으로 더딜 수밖에 없다. 그렇다고 하여 제도가 바뀔 때까지 실천을 유예할 수는 없다. 현 제도의 조건 속에서 가능한 평가혁신의 최대치가 어디까지인지를 실천적으로 검증하는 태도가 필요하다. 이런 점에 주목하여, 연구자는 A중학교 교사들과 함께 학교 단위 평가혁신을 위한 실행연구를 진행하였다.

2. 평가혁신의 지향점으로서의 성장중심평가

평가의 목적은 학생이 교육 목표에 도달할 수 있도록 돕는 것이다. 이는 “평가는 학생 개개인의 교육 목표 도달 정도를 확인하고, 학습의 부족한 부분을 보충하며, 교수·학습의 질을 개선하는 데 주안점을 둔다.”라고 규정한 2022 개정 교육과정을 통해서도 확인할 수 있다. 학생이 교육 목표에 도달했다는 것은 곧 전인적인 인간으로 성장하고 있다는 것을 의미한다.

성장중심평가는 ‘학생의 성장과 발달을 촉진하여 참된 학력과 핵심역량을 키워나갈 수 있도록 돕는 평가’(경기도교육청, 2018), ‘학생의 잠재력과 가능성을 확인하고 이를 현실화하기 위해 다양한 기회와 도움을 제공하여 모든 학생이 성장할 수 있도록 돕는 평가’(이형빈, 김성수, 2022) 등으로 개념 정의할 수 있다. 이러한 개념 정의를 통해 볼 때 성장중심평가가 기존의 평가 개념과 차별성을 갖는 지점은 ‘잠재력과 가능성’, ‘기회와 도움 제공’, ‘성장과 발달 촉진’ 등이다. 그렇다면, ‘도움 제공’, ‘잠재력 구현’, ‘발달 촉진’ 등의 구체적인 방법이 무엇이어야 하느냐가 주된 연구와 실천의 대상이 된다.

이 점에 대해서는 주로 특수교육에서 적용되고 있는 ‘역동적 평가’의 개념과 방법을 참고할 수 있다. 일반적인 학생 평가 방식은 제한된 시간 안에, 한 번의 기회를 통해, 걸어서 드러난 능력을 평가하는 ‘정적 평가(static assessment)’라고 할 수 있다. 이와 달리 특수교육 대상 학생에게 적절한 평가는 시간의 제약을 두지 않는 평가, 여러 번 기회를 주는 평가, 교사나 동료의 도움을 바탕으로 이루어지는 평가, 학생의 개별적 특성을 고려한 평가 등이다. 이를 ‘역동적 평가(dynamic assessment)’라 한다. ‘역동적 평가’는 Vygotsky(1978)의 근접발달영역 개념에 기반하고 있다. 즉 ‘이미 형성되어 있는 실제적 발달 수준’보다는 ‘교사와 학생의 역동적 상호작용 속에서 새롭게 발현될 잠재적 발달 수준’에 주목하는 평가이다. 역동적 평가는 인간의 변화 가능성에 대한 신뢰, 잠재적 발달 수준의 지속적인 확장, 교사와 학생의 상호작용에 기반한 다양한 평가 방식을 중시한다(Feuerstein, 1979).

이는 비단 특수교육에만 적용되어야 할 원리가 아니라 모든 교육에 확장되어야 할 바람직한 교육의 원리이다. 이에 따라 평가를 통해 학생의 장단점을 분석하고 이에 적합한

피드백을 제공해야 한다는 인식이 널리 확대되고 있으며(Frey & Fisher, 2011; Hattie & Timperley, 2007), 평가 루브릭을 단순히 채점 기준이 아니라 수행 기준 및 자기 점검의 도구로 활용하도록 해야 한다는 논의도 활성화되고 있다(Brookhart, 2013; Earl, 2013).

우리나라의 중등학교에서도 피드백의 중요성에 대한 인식은 널리 공유된 상황이라 할 수 있다. 일부 학교에서는 이에 더하여 피드백 이후 재도전의 기회를 준다든지, 루브릭을 성장중심평가의 취지에 맞게 개발한다든지, 형성평가와 총괄평가를 백워드 방식으로 구성한다든지 하는 실천 등을 모색하고 있다. 이 연구에서는 이러한 흐름에 주목하여, A중학교를 대상으로 성장중심평가의 취지에 따른 평가혁신 실천 방안을 탐색하고자 한다.

III. 연구대상 및 연구방법

1. 연구대상 학교의 특징

연구자는 A중학교 교원들과 함께 학교 차원의 공동체적 평가혁신을 위한 실행연구를 진행하였다. A중학교는 ○○광역시에 소재해 있으며, 2024년 기준 25학급에 800여 명의 학생과 50여 명의 교원이 근무하는 대규모 학교이다. A중학교 교원에 의하면 이 학교 학생들은 사교육 의존도가 높고 평가 결과에 대한 학부모 민원이 여느 학교에 비해 많다고 한다. 이처럼 A중학교가 처한 환경은 평가혁신을 시도하기에 우호적이지 않다. 50여 명의 교원의 근무하는 대규모 학교이다 보니 교원들의 의지와 자발성을 하나로 모으는 것이 수월하지 않다. 또한 학부모들은 전통적인 오지선다형 지필평가를 선호하는 분위기이다.

이와 반면에 교원들의 평가혁신에 대한 의지와 역량은 높은 편이다. 교장, 교감은 교사들의 전문성 신장을 지원하는 리더십을 발휘하고 있다. 교무부장, 연구부장 등 부장 교사들은 혁신학교 근무 경험을 바탕으로 교사 리더십을 발휘할 역량이 있다. 또한 교사 대부분이 전문적 학습공동체 등을 통해 교사 전문성을 기르고자 하는 학교문화가 형성되어 있다.

이러한 양면성으로 볼 때, A중학교는 이 연구의 취지에 적합한 연구대상이라 할 수 있다. A중학교는 비우호적인 환경에도 불구하고 교사들의 주체적 역량과 전문성에 따라 어느 정도까지 평가혁신을 시도할 수 있는지를 탐구하기에 적합하다. 연구자는 이 점에 착안하여 A중학교 교사들과 함께 실행연구를 진행하였다.

2. 실행연구 방법과 절차

실행연구(action research)는 실행자가 곧 연구자가 되어 문제를 제기하고 이에 대한 해결방안을 스스로 탐색하는 연구방법론이다. 실행연구의 주창자라 할 수 있는 Lewin(1946)은 이를 ‘과학적 법칙에 따라 실천적 행위를 하도록 하는 실험’이라고 정의하고, 그 절차를 ‘계획, 실천, 관찰 및 평가’로 제시하였다. 현장 내부 실행자들이 스스로 실행연구를 수행하기도 하지만, 연구 과정 및 결과를 성찰하기 위해 외부 연구자와 협력적 실행연구를 진행하기도 한다.

연구자는 A중학교의 요청으로 전문적 학습공동체 모임에 함께 참여하며 1년 동안 평가혁신 실행연구를 진행하였다. 실행연구를 위한 전문적 학습공동체에는 전체 50여 명의 교원 중 15명의 교원이 자발적인 의사에 따라 참가하였다. 교장·교감도 이 모임에 빠짐없이 참여함으로써 교사들을 지원하는 방안을 함께 모색하였다. 이들 교원이 먼저 실행연구를 진행하고 그 결과를 전체 교원과 공유하여, 모든 교원이 평가혁신을 적극적으로 실천하도록 유도하기로 하였다.

실행연구 모임은 2024년 3월부터 12월까지 매월 1회 진행되었다. A중학교 교사들은 우선 연구자의 안내에 따라 평가혁신에 대한 이론적 학습을 진행하였다. 그리고 교사들이 자기 교과목의 특성에 맞는 실천 과제를 선정하고, 이를 실제로 실행한 후 그 결과를 다음 모임 시간에 공유하였다. 그리고 이에 대해 동료 피드백을 거친 후 개선 방안을 마련하여 다시 실행하는 과정을 반복하였다. 2학기에도 이와 유사한 방식으로 실행연구를 진행하였고, 2학기 말에는 실행연구의 과정과 결과를 정리한 후 이를 바탕으로 A중학교 교사 모두가 실천해야 할 <A중학교 평가의 원칙과 방향>을 정리하였다. 이러한 실행연구 과정을 <표 1>과 같이 정리할 수 있다.

<표 1> 실행연구 절차

단계	절차	세부 내용	일시
1차 실행	사전 협의	사전 학습	성장중심평가의 이론적 배경
		사전 계획	실행연구계획 수립
			교과별·개인별 실천 과제 수립
	실행 중 성찰	실행 중 성찰	교과별·개인별 실천, 공유, 동료 피드백
		실행 중 성찰	교과별·개인별 실천, 공유, 동료 피드백
		실행 중 성찰	교과별·개인별 실천, 공유, 동료 피드백
	성찰 및 공유	1학기 평가회	실행연구 과정 및 결과 성찰
		공유	전체 교사 공유

단계	절차	세부 내용	일시
2차 실행	실행 중 성찰	실행 중 성찰	교과별 · 개인별 실천, 공유, 동료 피드백
		실행 중 성찰	교과별 · 개인별 실천, 공유, 동료 피드백
		실행 중 성찰	교과별 · 개인별 실천, 공유, 동료 피드백
	성찰 및 공유	학년말 평가회	실행연구 과정 및 결과 성찰
		공유	<A중학교 평가의 원칙과 방향> 정리 전체 교사 공유, 학부모 안내 자료 준비

실행연구 결과에 대한 검증으로 삼각검증(triangulation)을 사용하였다. 삼각검증은 자료에 의한 삼각검증, 연구자 삼각검증, 이론적 삼각검증, 방법론적 삼각검증 등이 있는데(Denzin, 1978), 여기서는 문헌 분석, 참여관찰, 심층 면접 등을 통해 연구대상을 여러 각도에서 바라보는 방법론적 삼각검증을 사용하였다. 연구의 타당도를 확보하기 위해서는 연구자와 연구대상과의 관계도 중요하다. 연구자는 A중학교 교감의 초청으로 실행연구에 참여했으며, 교원 대상 사전 연수와 월1회 컨설팅 등을 통해 교원들과 인간적 라포르를 형성했다. 이를 바탕으로 협력적 실행연구를 진행했기 때문에 연구를 안정적으로 진행하며 그 결과를 검증할 수 있었다.

IV. 평가혁신 실행연구 과정 및 결과

1. 실행연구 과제 선정하기

평가혁신 실행연구의 첫 번째 단계는 교사들이 실천해야 할 구체적인 과제를 선정하는 일이다. 이를 위해 교사들은 우선 연구자와 함께 성장중심평가의 이론적 개념과 원리를 함께 학습하였다. 이를 바탕으로 평가의 어려움을 공유하고 이를 극복하기 위한 과제를 선정하였다.

교사 1: 저는 루브릭을 만드는 게 제일 어려워요. 수행평가 끝나면 학생이나 학부모들이 점수 민원을 넣는 경우가 많은데, 루브릭을 제대로 만들어야 이 점수를 왜 부여했는지 설득이 되겠더라고요. 공부를 해 보니, 루브릭이 단순히 채점 기준만은 아니고, 신뢰도와 타당도를 갖춰야 한다고 하네요. 이 부분이 참 어려워요.

교사 2: 저는 학생들이 배운 내용을 제대로 이해했는지를 논술형으로 평가하고 있어요. 그런데 문제가 너무 어려운지 제대로 쓰지 못하는 학생이 많아요. 그래서 먼저 수업 시간에 간단한 글쓰기를 해 보고,

이를 제대로 된 논술 평가로 연결해야겠다 생각했어요. 이게 형성평가와 총괄평가를 연계하는 거겠죠.

교사 3: 영어 교과에서는 특히 피드백이 중요한데, 이 많은 학생에게 피드백을 일일이 주다 보면 '영혼을 갈아 넣는' 심정이에요. 효율적인 피드백 방법이 뭔지 궁금해요. 우리가 책에서 읽은 것처럼 집단 피드백, 개별 피드백, 동료 피드백 등을 효과적으로 구분하는 방법을 실험해 보려고요.

교사 4: 과학 교과에서는 탐구가 중요한데, 그냥 실험 보고서만 쓰고 나면 아이들에게 뭐가 남아 있을까 싶기도 해요. 요즘 핵심 아이디어, 핵심 질문을 강조하잖아요. 무엇보다 좋은 질문을 잘 만드는 게 중요할 것 같아요. 그래서 핵심 질문을 중심으로 탐구 보고서를 쓰도록 하는 것을 해 보고 싶어요.

교감: 제 마음 같아서는 모든 과목에서 지필평가를 한 번만 보면 좋겠는데, 그걸 학부모에게 이해시키려면 그만큼 수행평가의 타당도와 신뢰도를 확보해야 하겠죠. 어떡하면 학부모들이 지필평가 점수에 목매지 않을 수 있을지, 모든 선생님이 제대로 된 평가를 할 수 있을지 그 시스템을 만들어 봤으면 좋겠어요.

이러한 문제의식을 바탕으로 A중학교 교사들은 연구자와의 논의를 통해 2024년 1년 동안 실행을 통해 연구해야 할 과제를 크게 다섯 가지로 설정하였다.

첫째, 선다형 지필평가를 축소하고 수행평가를 확대하되 이를 형성평가-총괄평가의 방식으로 재구조화하는 것이다. 둘째, 수행평가의 과정에서 학생을 돕는 피드백을 효과적으로 구사하고 재도전의 기회를 부여하는 것이다. 셋째, 핵심 아이디어와 핵심 질문을 중심으로 한 탐구 기반 수업을 진행하고 이를 바탕으로 평가를 진행하는 것이다. 넷째, 이상에서 언급한 평가의 원칙을 평가 루브릭에 반영하고 이를 통해 평가의 타당도와 신뢰도를 확보하는 것이다. 다섯째, 모든 교사의 평가혁신을 유도하는 평가계획 양식을 새롭게 작성하는 것이다.

이상의 과제는 학생의 학습 과정을 돕고 모든 학생을 성장시키는 것을 목적으로 하는 성장중심평가의 취지를 구현하는 방법이다. 하지만 이는 기존의 평가 관행에 비추어 볼 때 대다수 교사에게 낯선 과제이다. 그래서 실행연구에 참여한 교사들이 먼저 이를 실천해 본 후 그 결과를 월 1회 공유하며 문제점을 보완하기로 했다. 그리고 이를 바탕으로 <A중학교 평가의 원칙과 방향>을 정리하고 전체 교사와 공유하기로 하였다.

2. 과제별 실행 및 성찰

가. 선다형 지필평가 축소

2015 개정 교육과정과 2022 개정 교육과정 등 국가 교육과정 문서에서는 “수행평가를 내실화하고 서술형과 논술형 평가의 비중을 확대한다.”라고 명시하고 있다(교육부,

2022). 또한 교육부 훈령에서는 “교과학습의 평가는 지필평가와 수행평가로 구분하여 실시한다. 다만, 학교별 학업성적관리규정에 따라 교과목의 특성상 수업활동과 연계한 수행평가만으로 평가할 수 있다.”라고 하고 있다(교육부, 2024). 따라서 교과별 특성과 학교의 자율적 판단에 따라 지필평가를 실시하지 않거나 1회만 실시하는 것이 가능하다.

하지만 대부분의 중등학교 교과에서는 중간고사, 기말고사라는 이름으로 학기마다 지필평가를 2회씩 실시하고 있다. 교사들은 지필평가를 2회 실시하는 관행과 수행평가 및 논술형 평가를 강화하려는 지향점 사이에서 모순과 갈등을 겪을 수밖에 없다. 이른바 ‘무늬만 논술형 평가’라든가 ‘사실상 과제형 수행평가’가 지속되는 이유가 여기에 있다.

A중학교 실험연구 참여 교사들은 이러한 관행에서 벗어나 수행평가와 논술형 평가를 내실화할 여건을 마련하기 위해 지필평가는 가급적 1회만 실시하거나, 수행평가 100%를 실시하는 것이 바람직하다고 잠정적으로 합의했다. 선다형 평가를 줄이는 대신 수행평가와 논술형 평가를 내실화하는 방안을 구체적으로 마련하기로 했다.

교사 5: 1학기에는 지필평가를 2회 실시했는데, 지필평가를 보고 나면 늘 민원에 시달리게 돼요. 학원에서는 우리가 만든 학습활동지로 예상 문제를 만들고, 시험이 끝나고 나면 학생들에게 민원 넣는 걸 부추기기도 해요. 이 와중에 수행평가는 과정중심으로 해야 하고, 진도는 나가야 하고, 너무 바쁜 거예요. 그래서 2학기에는 지필평가 1번으로 줄였어요. 그랬더니 제가 아이들에게 피드백을 주는 것이 훨씬 수월해졌어요. 아이들도 예전보다 수업에 훨씬 몰입하고 있어요.

지필평가 1회 실시를 경험한 교사들은 대체로 긍정적인 반응을 보였다. 무엇보다도 수행평가를 내실화할 수 있는 시간적·정신적 여유를 확보할 수 있고, 학생들이 수업에 더욱 몰입하는 효과를 얻을 수 있다는 것을 장점으로 여겼다. 하지만 일부 학생은 여전히 지필평가 점수에 연연하는 모습을 보이기도 했다. 이를 극복하는 방안에 대해서도 교사들의 숙의가 이루어졌다.

교사 6: 소위 공부 좀 한다는 애들은 여전히 수행평가보다 지필평가 점수에 연연해요. 수행평가는 대충 해도 점수는 웬만큼 나온다고 생각하고, 지필평가는 몇 문제만 틀려도 점수 차이가 크니까 거기에 올 인하는 것 같아요. 지필평가는 한 번만 보고 30%만 반영되는데, 여전히 아이들은 지필평가에 대한 부담을 크게 느끼는 것 같아요.

교장: 예전에 교수님도 말씀하셨듯이, 지필평가 만점을 꼭 100점으로 하지 않아도 되잖아요? 우리 학교 학업성적관리규정을 바꾸면 되고, 뭐, 교육청에서 뭐라고 하면 제가 책임지면 되니까. 지필평가 만점을 30점, 40점 이렇게 하면 아이들이 점수에 대한 민감도가 훨씬 낮아지지 않겠어요?

교사 7 : 그리고 예전에 우리가 학습했듯이, 수행평가로 평가할 것과 지필평가로 평가할 것을 구분해야 할

것 같아요. 우리 국어과에서는 쓰기, 문학 영역은 수행평가로, 문법, 읽기 영역은 지필평가로 평가하니까 문항수도 줄일 수 있고 아이들이 지필평가에 갖는 부담도 훨씬 줄어들었어요.

이처럼 실행연구 참여 교사들은 모든 영역 모든 단원을 오지선다형 지필평가 반영하는 관행에서 벗어나, 평가의 목표에 따른 평가 방법을 적용하였다. 지식의 암기와 이해는 선다형 지필평가로, 지식의 활용이나 논리적 사고능력은 수행평가나 논술형 평가로 확인함으로써 선다형 지필평가에 대한 과도한 의존에서 벗어날 수 있었다.

또한 지필평가의 만점을 100점이 아닌 실제 반영 비율로 하기로 하였다. 이는 기존의 관행에 비추어 볼 때 매우 획기적인 방안으로, 예컨대 지필평가의 반영 비율이 30%라면 지필평가 만점 역시 30점으로 하는 것이다. 이는 학생의 점수 민감도를 낮춤으로써 평가의 본질을 회복하는 데에 도움이 된다. 현행 교육부 훈령에서는 “지필평가는 가급적 만점으로 출제한다.”라고 규정하고 있다(교육부, 2024). 하지만 ‘가급적’이라는 문구에 대한 명확한 논리적 근거는 찾기 어렵다. 이는 오랜 관행이 굳은 결과로 보인다. 서울특별시교육청은 이를 적극적으로 해석하여 “지필평가 및 수행평가 영역별 만점은 학기말 반영 비율 또는 100점 만점으로 한다.”라고 하고 있다(서울특별시교육청, 2025). A중학교는 서울특별시교육청 소속 학교가 아님에도 교사의 숙의를 거쳐 학업성적관리규정에 지필평가 만점을 실제 반영 비율로 함으로써 교사들이 평가혁신 실천을 뒷받침하는 제도적 장치를 마련하였다.

선다형 지필평가 축소가 곧 평가혁신의 성과로 이어지는 것은 아니다. A중학교 교사들은 선다형 지필평가를 축소함으로써 얻을 수 있는 물리적, 심리적 여유를 바탕으로 수행평가 방식을 새롭게 하고, 피드백과 재도전의 기회를 충분히 제공하고, 이를 바탕으로 학생들의 깊이 있는 배움을 이끌기 위해 노력하기로 했다.

나. ‘진단평가-형성평가-총괄평가 방식’으로 수행평가 재구조화

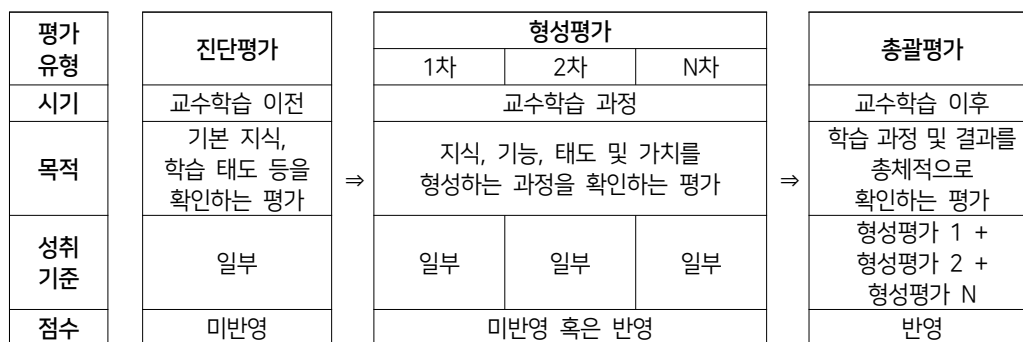
A중학교 실행연구 참여 교사들은 지필평가의 오랜 관행을 극복하는 것 못지않게, 수행평가의 관행을 극복하는 것도 중요하다고 보았다. 어느 학교와 마찬가지로 A중학교에서도 수행평가에 대한 학생의 불만, 교사의 부담이 적지 않았다. 학생들은 수행평가가 또 하나의 시험이나 과제처럼 느껴지고, 특정 시기에 여러 교과와 수행평가가 몰리는 부담을 느끼고 있었다. 실행연구 참여 교사들은 수행평가를 많이 하는 것이 중요한 것이 아니라 수행평가의 취지를 살려 학생의 학습 과정을 돕는 것이 중요하다고 여기고 이를 위한 방법을 모색해 갔다.

교사 7: 수행평가 비율을 70%까지 하는 것은 이번 학기가 처음이요. 학생들에게 꾸준하고 지속적인 학습

습관을 심어주는 첫 학기죠. 지필고사 중심으로 훈련된 학생들에게 이런 체질 개선이 언제 어떻게 어떤 변화를 불러올지 기대되네요. 수행평가를 늘리면서, 사교육을 받지 않은 학생을 최대한 보호하고, 사교육으로 지친 학생이 자기주도적인 학습 의지를 심어주려고 해요.

실행연구 참여 교사들은 교사나 학생들이 수행평가를 부담스럽거나 무의미하게 느끼는 이유가 명확한 목표 의식 없이 여러 가지 수행 과제를 단편적, 분절적으로 시행했기 때문이라고 분석하였다. 이러한 기존 관행을 극복하기 위해서는 수행평가를 진단평가-형성평가-총괄평가 방식으로 재구조화하는 방법을 모색하기로 했다.

교사 8: 우리 교과에서는 3개의 형성평가와 하나의 총괄평가, 총 4단계의 과정중심 수행평가 계획을 만들어 봤어요. 이걸 요리 과정으로 비유할 수 있어요. 형성평가 1단계에서는 A라는 재료를, 2단계에서는 B라는 재료를 다듬는 시간이고, 마지막 총괄평가에서는 그 재료들을 이용해 하나의 요리를 만드는 거죠. 교사는 형성평가 단계별로 피드백으로 주고, 아이들은 그 피드백을 받아 최종적으로 결과물을 제출하는 거죠. 이러한 과정을 거쳐 마지막에 아이들이 조금은 더 성장하지 않을까 하는 바람에서 계획을 세웠어요.



[그림 2] 진단평가-형성평가-총괄평가 흐름의 수행평가 예시

실행연구 참여 교사들은 이러한 수행평가의 흐름을 연구자와 함께 [그림 2]와 같이 도식화해 보았다. 진단평가는 필요한 경우에 실시하고, 형성평가는 교과 및 단원의 특성에 따라 2~3회 진행하고, 총괄평가는 형성평가의 과정을 모두 아우르는 방식으로 진행하는 것으로 했다. 형성평가를 통해 확인한 성취기준을 총괄평가로 다시 한번 확인하는 구조이다. 형성평가의 경우 일부 영역, 예를 들어 모듈별 토론 활동의 경우 협력이라는 취지를 저해하지 않기 위해 점수에 반영하지 않도록 하였다. A중학교 교사들은 이를 교과별, 교사별로 실행을 한 후 이에 대한 성찰을 함께 나누었다. [그림 3]은 국어과 교사들이 실천한 사례이다.

평가 유형	형성평가				총괄평가
	1차	2차	3차		
방법	독서 포트폴리오	독서 토론	토론에 대한 성찰문	⇒	독서 토론 경험을 바탕으로 한 연설
성취 기준	[9국05-01] [9국05-06] [9국02-01]	[9국01-05] [9국01-06] [9국01-09]	[9국02-05] [9국03-04]		[9국05-01] [9국05-06] [9국02-01] [9국01-05] [9국01-06] [9국01-09] [9국02-05] [9국03-04]
점수	반영	미반영	반영		반영

[그림 3] 진단평가-형성평가-총괄평가 흐름의 수행평가 국어과 사례

교사 9: 첫 번째 수행평가는 독서 감상문 쓰기였어요. 이 수행평가는 형성평가 개념이었기 때문에 학생 입장에서 저부담 평가이어야 하나, 제 욕심 때문에 글을 수정하는 과정에서 시간을 너무 많이 들이게 되었어요. 두 번째 수행평가는 각자 쓴 독후감을 바탕으로 독서 토론을 진행했어요. 그리고 마지막 총괄평가에서는 독서, 독후감, 토론 경험을 바탕으로 15분 정도의 연설문을 쓰고 자신의 주장을 발표하도록 했어요. 형성평가에 공을 들여서 그런지 학생들이 총괄평가는 오히려 덜 부담스러워하더라고요. 자신이 선정한 한 권의 책부터 시작해서 마지막 연설에 이르기까지 긴 호흡을 통해 교사가 생각한 것 이상의 것을 얻어가지 않았을까 해요.

이러한 방식은 학생이 학기 말에 도달해야 할 핵심 목표를 중심으로 총괄평가와 형성평가를 유기적으로 연결하는 백워드 설계 방식의 평가계획이라 할 수 있다. 기존의 수행평가가 일회적·단절적 방식의 평가라는 점에서 한계가 있다면, 이러한 방식의 수행평가는 학생이 핵심 목표에 도달할 수 있도록 지속적·연계적으로 이끄는 평가라는 점에서 의미가 있다. 즉, 교수학습 과정에서 형성평가를 통해 학생의 지식, 지능, 태도 및 가치를 형성하도록 돕고, 총괄평가를 통해 이를 최종적으로 확인하는 방식이라는 점에서 성장중심평가의 취지를 잘 살릴 수 있다. 이는 대다수 교사에게는 다소 낯선 방식이어서 시행착오도 없지 않았으나, 실행연구에 참여한 교사들은 이를 단편적인 일회성 과제형 수행평가를 넘어 학생의 깊이 있는 배움과 성장을 돕는 방법으로 여기고 이를 동료 교사와 함께 확대하기로 했다.

다. 피드백과 재도전이 있는 평가

A중학교 실행연구 참여 교사들은 형성평가와 총괄평가를 연계하는 과정에서 자연스럽게 피드백을 단계별로 제공하게 되었다. 그 과정에서 교사들은 기존에 자신이 해왔던 피드백 방식에 어떤 문제점이 있었는지 발견하게 되었다. 그리고 이를 개선하는 방법을 동료와 함께 찾아갔다.

교사 3: 생각해 보니 제가 피드백을 하면서 ‘영혼을 갈아 넣는’ 느낌이 들었던 이유는 아이들의 과제를 걷고 거기에 일일이 답글을 써서 돌려주고 나서 정작 교사의 피드백이 아이들에게 어떤 영향을 주었는지 확인할 방법이 없었기 때문이더라고요. 다시 말해 피드백에 대한 피드백이 없었다고나 할까요? 그냥 제 에너지를 소모적으로 낭비한 셈이었죠. 그래서 이제는 아이들이 학습지를 쓸 때 바로바로 피드백을 말로 해주고, 그 자리에서 고쳐 쓰도록 하고 있어요.

앞에서 언급했듯이 성장중심평가란 ‘다양한 기회와 도움을 제공하여 모든 학생이 성장할 수 있도록 돕는 평가’(이형빈, 김성수, 2022)이며, 이를 위해서는 ‘피드백’이 핵심적인 요소이다. 피드백이 이루어지려면 마땅히 ‘교수학습의 과정’에서 평가가 이루어져야 하며, 이는 ‘학생의 변화와 성장’을 위한 것이어야 한다. 이러한 요소가 유기적으로 연결되지 못한 피드백은 교사 입장에는 ‘또 하나의 업무’, 학생 입장에는 ‘부담스러운 잔소리’가 될 수 있다.

교사 8: 형성평가와 총괄평가를 연결하다 보니 자연스럽게 피드백 이후에 재도전의 기회를 주게 되더라고요. 아이들이 피드백을 받고 그 피드백을 반영해 결과물을 만들어내니까 훨씬 더 좋은 점수를 받게 되더라고요. 처음에는 재도전의 기회를 주는 거를 공부 좀 한다는 애들은 싫어하지 않을까 했는데, 그 이유를 충분히 설명하고 모두에게 기회를 주니까 다들 만족해하더라고요. 아이들이 서로 “야, 너도 할 수 있어. 다시 해봐.”하고 응원하는 모습을 보니까 교사로서도 보람을 느끼고요.

실행연구 참여 교사들은 피드백이 형성평가의 과정에서 자연스럽게 이루어져야 할 뿐만 아니라, 이를 바탕으로 학생에게 ‘재도전’의 기회를 주어야 한다고 보았다. 그래야 피드백이 학생 성장에 실질적으로 도움을 줄 수 있기 때문이다. 이처럼 A중학교 실행연구 참여 교사들은 형성평가와 총괄평가를 유기적으로 연계하고, 이 과정에서 피드백에 따른 재도전의 기회를 줌으로써 성장중심평가의 취지를 구현해 갔다.

라. 핵심 아이디어 및 핵심 질문에 따른 평가

A중학교 실행연구 참여 교사들은 실행연구를 거듭해 갈수록 평가의 본질에 대한 성찰이 깊어졌다. 평가를 통해 학생의 성장을 돕는다는 것은 단지 학생의 성적이 오르도록 하는 것만이 아니라 배움의 깊이가 깊어지게 하는 것을 의미한다고 보았다. 이때 연구자와 교사들이 주목한 개념이 ‘핵심 아이디어’와 ‘핵심 질문’이다.

교감: 제가 우리 선생님들 수업을 참관해 보니 뭔가 허전하고 아쉬운 생각이 많이 들었어요. 선생님들이 수업은 열심히 하시는데, 그 수업을 통해 아이들의 배움이 깊어진다는 느낌은 솔직히 들지 않았어요. 수행평가도 이걸 왜 해야 하나 싶은 게 많고요. 선생님들이 수업과 평가를 계획하실 때 내 수업이 어디로 가야 하는지, 아이들에 대해 무엇을 평가하고 싶은지를 명확히 했으면 좋겠다는 생각이 들어요.

2022 개정 교육과정에서는 “단편적 지식의 암기를 지양하고 각 교과목의 핵심 아이디어를 중심으로 지식·이해, 과정·기능, 가치·태도의 내용 요소를 유기적으로 연계하며 학생의 발달 단계에 따라 학습 경험의 폭과 깊이를 확장할 수 있도록 수업을 설계한다.”라고 하고 있다. 이 취지를 이해하기 위해 실행연구 참여 교사들은 연구자의 도움을 받아 ‘이해중심 교육과정’에서 말하는 ‘핵심 아이디어’와 ‘핵심 질문’에 대해 학습하였다.

‘핵심 아이디어’는 ‘학생들에게 오랫동안 남아 있어야 할 본질적이고 포괄적인 수준의 지식’, ‘심층적 학습을 통해 획득되고, 다른 상황에도 전이 가능한 지식’을 말한다. 또한 ‘핵심 질문’은 ‘학생들이 끊임없이 탐구하며 핵심 아이디어에 도달하도록 이끄는 질문’, ‘학생들의 지적 호기심과 진정성 있는 관심을 불러일으키는 개방적 질문’을 의미한다. 이를 통해 학생들이 ‘진정한 이해’에 도달하도록 하는 것이 ‘이해중심 교육과정’의 취지이다(Wiggins & McTighe, 2000).

Wiggins와 McTighe(2000)는 ‘교과 진도 나가기식 수업’을 극복하기 위해 도입된 ‘활동 중심 수업’ 역시 ‘활동을 위한 활동’에 머무는 경우가 많다고 지적하였다. 즉 단순한 흥미 위주의 학습활동이 끝난 이후 학생에게 실질적으로 깊이 있는 배움이 이루어지지 않는 경우가 많다는 것이다. A중학교 교사들 역시 이러한 지적에 공감하고, 단순히 겉보기 좋은 수행평가나 논술형 평가를 넘어 학생의 깊이 있는 배움과 성장을 확인하는 평가가 이루어져야 한다고 보았다. 그래서 ‘핵심 질문’을 중심으로 한 탐구 수업을 진행하고, 그 결과 학생들이 ‘핵심 아이디어’에 도달했는지를 확인하는 수행평가와 논술형 평가를 진행하기로 했다.

교사 10: 처음에는 핵심 아이디어, 핵심 질문 이런 걸 굳이 왜 작성해야 하는지 이해하지 못했어요. 하지만 실제로 수업해 보니, 아이들이 뭔가 한 번이라도 더 생각해 보고 글을 쓴다는 느낌을 받았어요.

교사 11: 핵심 아이디어 만드는 게 참 어려워요. 그럴 때마다 교수님이 알려주신 대로, “학생들이 수업 시간에 배운 것 중 나머지는 다 잊어도 반드시 머릿속에 남아 있어야 할 깨달음 하나가 뭘까?”를 생각하며 수업을 만들고 이것을 평가로 연결시켜 보려고 해요.

교사가 교육과정을 분석하여 수업과 평가를 계획할 때 핵심 아이디어와 핵심 아이디어를 직접 작성해 보는 것은 매우 의미가 있다. 그래야 학생의 깊은 배움을 이끌 수 있고, 이를 평가를 통해 확인할 수 있다. 하지만 이는 교사에게 매우 어려운 작업이다. 그래서 연구자와 실행연구 참여 교사는 ‘핵심 아이디어란 학습을 마친 후 다른 것은 잊더라도 반드시 학생에게 남아 있어야 할 깨달음’, ‘핵심 질문이란 핵심 아이디어에 도달하도록 돕는 질문’과 같이 쉬운 언어로 생각해 보기로 했다. 그리고 본인들이 직접 핵심

아이디어와 핵심 질문을 작성해 보고 이를 교수학습과 평가에 반영하였다. <표 2>는 A 중학교 과학과 실행연구 참여 교사가 작성한 자료이다.

<표 2> 핵심 아이디어 및 핵심 질문에 따른 수행평가 예시

학습 목표	생물다양성 보전의 중요성을 인식하고 이를 위한 실천 역량 기르기
핵심 아이디어	나(인간)도 다양한 생명체 가운데 하나이다. 생물다양성을 보전해야 미래의 행복도 지속될 수 있다.
핵심 질문	<ul style="list-style-type: none"> ○ 개념 이해를 돕는 질문 <ul style="list-style-type: none"> - 생물이 많이 산다는 것과 생물이 다양하다는 것의 차이는 무엇인가? - 우리 마을에 살고 있는 생물 중 내가 아는 생물 종은 몇 가지일까? ○ 생각을 넓히는 질문 <ul style="list-style-type: none"> - 만약 우리 마을에 생물 종이 줄어든다면 어떤 일이 생길까? ○ 인생을 성찰하게 하는 질문 <ul style="list-style-type: none"> - 생태계의 모든 존재는 본질적으로 동등한 가치를 지니고 있을까? - 나와 길고양이도 동등한 가치를 지닌 존재일까?
수행평가 과제	생물의 분류체계에 따라 우리 마을 생물 종을 분류해 보고, 핵심 질문에 따른 <1인 1생물 탐구 보고서>를 작성하여 발표한다.

교사 4: 이렇게 진행해 보니, 아이들의 실험 보고서의 질이 작년에 비해 훨씬 좋아진 게 사실이에요. 아이들도 생태 문제에 더 진지하게 접근하고 실천하려는 모습을 보이고요. 교사인 저에게도 많은 변화가 생겼는데, 이 활동을 왜 해야 할까, 아이들에게 무엇을 남겨야 하는지 본질적인 질문을 저 스스로에게도 하게 되었어요.

이처럼 ‘핵심 아이디어’와 ‘핵심 질문’을 중심으로 한 탐구 기반 수업과 평가 결과 학생들의 성장을 한층 분명하게 확인할 수 있었다. 하지만 실행연구 세부 과제 중 교사들이 가장 어려웠던 과제가 핵심 아이디어와 핵심 질문을 개발하는 것이었다. 그래서 A중학교 교사들은 이에 대한 추가적인 학습이 필요하다고 보고, 2025년도 2차 실행연구를 통해 이 과제를 좀 더 심화시켜 탐구하기로 하였다.

3. 실행연구 마무리

이상에서 언급한 것처럼, A중학교 실행연구 참여 교사들은 일 년 동안 실행과 성찰을 반복하며 성장중심평가의 취지를 공동체적으로 실현할 방안을 탐구해 왔다. 구체적으로는 선다형 지필평가 축소하기, 진단평가-형성평가-총괄평가의 흐름에 따라 수행평가

재구조화하기, 피드백을 제공하고 재도전의 기회 부여하기, 핵심 아이디어와 핵심 질문을 중심으로 한 수행평가 과제 개발하기 등의 세부 실천 과제를 탐색했다. 이는 실행연구에 참여한 교사만의 실천 과제가 아니라 A중학교 모든 교사의 실천 과제로 일반화되어야 한다. 그래야 교사 개인 차원을 넘어 학교 공동체적 차원의 평가혁신이 이루어질 수 있다.

공동체적 평가혁신이 이루어지려면 이를 자연스럽게 유도하는 매개적 도구가 필요하다. 활동이론에 따르면 주체와 대상, 공동체 사이의 상호작용이 원활하게 이루어지기 위해서는 매개적 도구의 역할이 중요하다(Engeström, 1987). 새로운 도구는 새로운 성과를 낳는 매개 역할을 할 수 있다. 실행연구 참여 교사들은 학기별 평가계획 양식과 평가루브릭을 새롭게 개발하여 자신들이 연구한 내용을 반영하기로 했다. 이를 통해 학교 교사 모두의 공동체적 평가혁신을 유도하기로 했다.

가. 새로운 평가계획 양식 개발

앞에서 언급했듯이, 현재 학교에서 진행되는 평가는 대부분 중간고사와 기말고사 등 지필평가 2회, 분절적 수행평가 3~4회를 진행하는 방식이 고착되어 있다. <표 3>은 대부분의 학교에서 사용하는 평가계획 양식으로, 여기에는 관행적으로 고착된 지필평가와 수행평가 형태가 반영되어 있다.

<표 3> 학교 현장에서 흔히 사용되는 평가계획 양식

유형	지필평가		수행평가		
	중간고사	기말고사	1차	2차	3차
방법					
만점					
반영 비율					
성취기준					
시기					

A중학교 실행연구 참여 교사들은 새로운 평가 방식을 확산하기 위해서는 학교에서 사용하는 평가계획 양식부터 바꾸는 것이 필요하다고 보고, <표 4>와 같은 새로운 평가계획 양식을 개발하였다.

<표 4> A중학교 실행연구에 따라 새롭게 개발한 평가계획 양식

유형	지필평가	수행평가			
횟수					
방법		형성평가			총괄평가
		1차	2차	N차	
배점 (반영 비율)					
성취기준					
시기					

새로운 평가계획 양식은 지필평가는 반드시 2회를 실시해야 한다는 고정관념에서 벗어나 교과와 특성이나 교사의 자율성에 따라 횟수를 달리할 수 있도록 하였다. 또한 영역별 배점을 곧 반영 비율과 동일시함으로써 지필평가 만점을 100점으로 해야 한다는 관념에서도 벗어나도록 하였다. 그리고 수행평가를 형성평가와 총괄평가가 유기적인 흐름으로 연계되는 방식으로 실시하도록 하였다.

나. 새로운 평가 루브릭 활용

평가계획 양식과 마찬가지로 평가 루브릭 역시 평가에 대한 관념을 새롭게 할 수 있는 매개적 도구가 될 수 있다. A중학교를 포함해 대부분의 학교에서 사용하고 있는 수행평가 루브릭은 <표 5>와 같은 방식이다. 하지만 여기에는 학습의 과정을 중시하는 과정중심평가, 학생의 성장을 돕는 성장중심평가의 취지가 반영되어 있다고 보기 어렵다.

<표 5> 학교 현장에서 흔히 사용되는 평가 루브릭

평가 요소	성취수준 (채점 기준)	성취도	배점
		A	4
		B	3
		C	2
		A	3
		B	2
		C	1
		A	3
		B	2
		C	1

일반적으로 평가 루브릭을 수행평가나 논술형 평가를 위한 ‘채점 기준표’로 인식하는

경향이 강하다. 하지만 평가 루브릭은 단순한 채점 기준이 아니라 ‘학생이 배워야 할 것을 제대로 배웠는지에 대한 피드백을 제공해 학습을 지원하는 역할’(Brookhart, 2013)을 해야 하며, ‘학생이 앞으로 어떻게, 어디로 나아가야 하는지에 대한 설명에 비추어 학생의 작업을 분석하도록 하는 도구’(Earl, 2013)가 되어야 한다.

실행연구 참여 교사들은 성장중심평가에 대해 학습하고 구체적인 실현 방안을 탐구해 온 만큼, 이를 평가 루브릭에 구체적으로 반영할 필요가 있다고 보았다. 기존의 평가 루브릭이 사실상 결과물만을 대상으로 평가하는 방식이라면, 새로운 평가 루브릭에는 학습 과정과 결과를 두루 반영해야 하며, 여기에 피드백과 재도전의 절차가 명시되는 것이 바람직하다고 보았다. 또한 인지적 영역뿐만 아니라 정의적 영역에 대한 평가 항목도 평가 루브릭에 반영해야 한다고 보았다. 이러한 논의를 거쳐 <표 6>과 같은 새로운 평가 루브릭을 사용하기로 했다.

<표 6> A중학교 실행연구에서 새롭게 활용한 평가 루브릭

단계	평가 요소		1차 평가			2차 평가	
			도달	미도달		도달	미도달
계획					피드백		
과정							
산출물							
실행 후 성찰	교사의 관찰		정의적 영역 평가 (‘세부능력 및 특기사항’ 기록)				
	자기평가						
	동료평가						

이 루브릭은 우선 학생의 학습 과정을 ‘계획, 과정, 산출, 실행 후 성찰’ 과정으로 구조화하고 각각에 단계에 적합한 평가 요소를 선정하도록 하였다. 이를 통해 Earl(2013)이 말했듯이 ‘학생이 앞으로 어떻게, 어디로 나아가야 하는지’를 친절하게 안내하는 역할을 하도록 하였다. 그리고 각각의 평가 요소에 대해 3단 척도나 5단 척도로 촘촘히 변별하는 것은 현실적으로 가능하지도 않고 절대평가의 취지에도 맞지 않다고 보아 도달/미도달 여부만을 판정하는 것으로 하였다. 또한 ‘피드백과 재도전의 기회 부여’ 과정을 루브릭에 명시함으로써 모든 교사와 학생이 성장중심평가의 취지를 명확하게 인식하도록 하였다.

실행연구에 참여하지 않은 교사들은 새롭게 개발한 평가계획 양식과 평가 루브릭이 다소 낯설거나 부담스러울 수 있다. 따라서 모든 교사에게 이를 강요할 수는 없다. 하지만 새롭게 만든 평가계획 양식과 평가 루브릭이 새로운 평가혁신을 유도하는 매개적 도

구 역할을 할 수 있을 것으로 보고, 이를 개발한 취지를 충분히 공유하는 과정을 거치기로 했다.

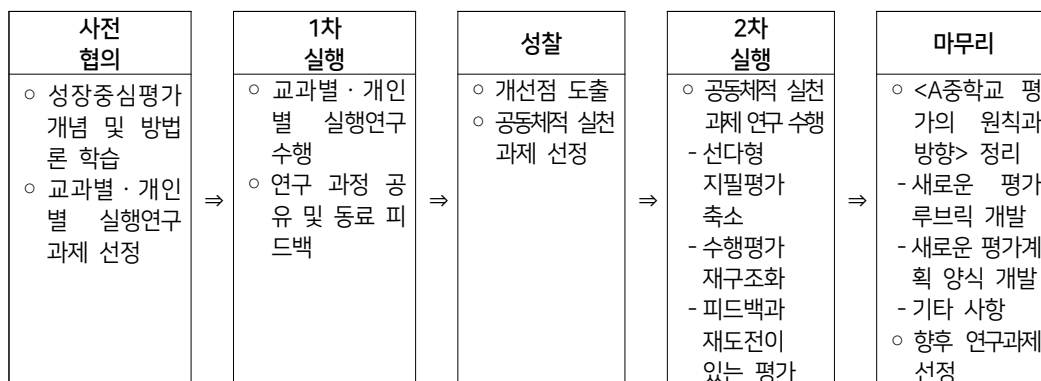
평가 루브릭 개발은 평가의 타당도와 신뢰도 확보를 위해 매우 중요하다. 특히 A중학교처럼 선다형 지필평가를 최소화하고 수행평가와 논술형 평가를 확대하는 경우 더더욱 그러하다. 그런데 적지 않은 학교에서 평가 루브릭을 마치 학생과 학부모의 평가 결과에 대한 민원을 예방하는 차원에서 마련하는 경향이 강하다. 그러한 소극적 차원의 자기 정당화를 넘어 평가의 취지를 제대로 반영하는 루브릭을 개발하는 것은 곧 평가 전문가로서의 교사의 교권을 적극적으로 확보하는 길이기도 하다.

다. 실행연구 결과 공유 및 향후 계획 수립

이상에서 언급한 내용을 중심으로, 실행연구 참여 교사들은 모든 교사가 함께 실천해야 할 세부 과제를 선정하고, 이를 <A중학교 평가의 원칙과 방향>이라는 자료집으로 정리하였다. 여기에는 앞에서 제시한 <표 4>와 <표 6>의 새로운 평가계획 양식과 평가 루브릭을 포함하여 다양한 설명 및 예시 자료가 수록되어 있다.

실행연구 참여 교사들은 2024년 말 학교 교육과정 평가회를 통해 실행연구 결과를 전체 교사에게 발표하고 그 취지를 공유하는 시간을 가졌다. 그리고 이를 2025년 2월 교육과정 워크숍 기간에 전입·신입 교사와 함께 자세히 학습하는 시간을 갖기로 하였다. 2025년에는 이를 모든 교사가 실천해 보고 이를 다시 보완하는 시간을 갖기로 하였다. 이처럼 A중학교에서는 2024년 실행연구팀의 연구 결과를 2025년에 전체 교사를 대상으로 실행연구를 확장하고, 2024년 실행연구팀은 아직 완결 짓지 못한 연구과제를 2025년에 다시 심층적으로 탐구하기로 하였다.

이상에서 언급한 A중학교 실행연구 과정 및 결과를 [그림 4]로 정리할 수 있다.



[그림 4] 실행연구 과정 및 결과

V. 결 론

이 연구는 “입시의 영향력이 강한 중등학교에서는 평가혁신이 어렵다.”라는 암묵적인 비관주의를 극복하고, 현재의 조건 속에서 실현 가능한 중등학교 평가혁신의 범위는 어느 정도이며 이를 현실화하기 위해서는 무엇이 필요한지를 확인하기 위해 진행되었다. 이를 구체적으로 확인하기 위해 연구자는 A중학교 교사들과 일 년 동안 실행연구 계획을 수립하고, 실행 중 성찰을 거듭하며, 이를 일반화할 가능성을 탐색하였다. 이를 통해 확인한 A중학교에서의 성과를 정리하고 이를 통해 확인할 수 있는 시사점을 제안하고자 한다.

1. 실행연구를 통해 확인한 평가혁신의 가능성

대학 입시와 상대평가의 영향력을 직접 체감하고 있는 고등학교는 물론이고 절대평가가 적용되고 있는 중학교에서도 평가혁신에 대한 비관주의가 자리 잡고 있다. 하지만 A중학교 교사들은 이러한 정서를 극복하고 적극적으로 평가혁신을 시도해 보기 위해 실행연구를 진행했다. 연구자는 A중학교 교사들의 실행연구 전반을 관찰하고, 그 과정과 성과에 대해 교사들과 함께 평가를 진행하였다. 이 과정에서 확인한 A중학교 실행연구의 특징과 그 성과는 다음과 같다.

첫째, A중학교 교사들은 공동체적 차원의 학습과 실천을 통해 평가혁신의 성과를 거두었다. 이는 달리 말해 교사 개인 차원의 실천으로는 평가를 둘러싼 여러 가지 어려움을 극복하기 어렵다는 의미이기도 하다. A중학교에서 여러 가지 의미 있는 성과를 거둔 이유는 무엇보다도 공동체적 실천을 통해 든든한 심리적 울타리를 확보함으로써 혼자서는 하기 어려웠던 새로운 시도를 감행하는 내적 동기를 발휘했기 때문이다.

둘째, A중학교 교사들은 학교 실정에 적합한 실행연구 방법론을 적용함으로써 연구와 실천, 성찰을 이어갔다. 평가혁신이 어려운 이유 중 하나는 기존의 관행을 답습하려는 ‘경로 의존성’을 강하게 작동하기 때문이다. 예컨대 지필평가를 학기마다 두 차례씩 실시하는 관행, 지필평가에서 선다형 문항을 다수 출제해 변별성을 확보하고자 하는 암묵적 정서가 대표적이다. A중학교 교사들은 성장중심평가의 개념과 방법에 대한 이론적 학습을 통해 평가에 대한 인식의 지평을 넓혔고, 이를 바탕으로 그동안 상상하지 못했던 새로운 실천을 시도하였다. 또한 모든 교사가 동일한 실천 과제를 일률적으로 시도하지 않고, 교과별·개인별 특성에 따른 실천 과제를 스스로 선정함으로써 교사 주도성을 발휘하였다. 그리고 월 1회 정기적으로 모여 각자의 실천 성과를 공유하고 동료 피드백을

통해 더 좋은 대안을 찾아 나갔으며, 이러한 실천을 토대로 모든 교사가 공동으로 실천해야 할 과제를 찾아 나갔다.

셋째, A중학교 평가혁신에서는 학교 관리자의 리더십도 매우 중요한 역할을 하였다. A중학교 교장·교감은 연구와 함께 실행연구 일정을 계획하고 교사들의 참여를 독려했다. 또한 실행연구 모임에 빠짐없이 참석하여 교사들의 의견을 경청하고, 이 의견을 학교 운영에 반영할 권한을 행사하였다. 이처럼 교육과정 운영의 방향을 제시하고, 구성원을 독려하며, 구성원들의 합의 결과를 학교 운영에 반영하는 학교 관리자의 역할을 ‘변혁적 교육과정 리더십’이라 할 수 있다(Hawthorne et al., 2000). 그동안 학교 관리자의 리더십은 주로 ‘학교 운영 리더십’ 차원에서 논의되어 왔다. 이제는 교육과정, 수업, 평가 등 교사 고유의 영역이라고 분류되어 온 부분에서도 학교 관리자가 어떤 역할을 해야 하는지에 대해 논의가 확장되어야 할 것이다.

넷째, 평가혁신의 성과가 일부 교사에게 국한되지 않고 일반화되기 위해서는 매개적 도구의 역할이 중요함을 확인할 수 있다. 실행연구 참여 교사들이 새로운 평가계획 양식과 평가 루브릭을 활용한 이유가 여기에 있다. 실행연구 참여 교사들은 선다형 지필평가를 지양하고 수행평가의 비율을 높였으며, 수행평가 또한 기존의 일회적·분절적 방식에서 탈피하여 진단평가-형성평가-총괄평가가 유기적으로 연계되는 방식으로 재구조화하였다. 그리고 이러한 흐름을 새로운 평가계획 양식에 반영함으로써 다른 교사들도 이를 시도할 수 있도록 유도하였다. 또한 학습의 과정을 중시하는 과정중심평가, 피드백과 재도전을 통해 모든 학생이 성장할 수 있도록 성장중심평가를 시행해 보고, 이러한 흐름을 한눈에 파악할 수 있는 평가 루브릭을 활용하였다.

평가의 타당도와 신뢰도 확보를 위해서는 평가 루브릭을 제대로 작성하는 것이 매우 중요하다. 특히 A중학교처럼 선다형 지필평가를 지양하고 수행평가와 논술형 평가를 확대하는 경우 더더욱 그러하다. 그런데 적지 않은 학교에서 평가 루브릭을 마치 학생과 학부모의 평가 결과에 대한 민원을 예방하는 차원에서 마련하는 경향이 강하다. 그러한 소극적 차원의 자기 정당화를 넘어 평가의 취지를 제대로 반영하는 루브릭을 개발하는 것은 곧 평가 전문가로서의 교사의 교권을 적극적으로 확보하는 길이기도 하다.

다섯째, A중학교 교사들이 거둔 가장 중요한 성과는 ‘방어적 교육활동’의 풍토에서 벗어나 ‘적극적 교육활동’의 문화를 구축했다는 점이다. 김성천(2024)은 서이초 사건 이후에 교직 문화에 나타난 새로운 흐름을 ‘방어적 교육활동’으로 명명하고 이를 ‘법적·행정적으로 문제가 되지 않는 않지만, 불가피한 이유와 상황에 의해 소극적인 방식으로 직무를 수행하려는 관점·태도·행동으로 인해 학교 교육력의 질 향상으로 이어지지 못하게 하는 상태’로 정의하였다. 방어적 교육활동은 다양한 영역에서 나타나는 현상이지만 특히 평가 영역에서 더욱 두드러지게 나타나는 현상으로 볼 수 있다.

평가와 관련한 ‘방어적 교육활동’은 주로 ‘입시경쟁으로 인한 제도적 한계’와 ‘평가 결과에 대한 민원’, ‘상급 기관의 규제적 관리 감독’ 등에서 발생한다. 특히 코로나 사태와 서이초 사건을 계기로 교사들의 심리가 더더욱 위축됨에 따라 적극적 교육활동을 회피하고 스스로 방어적 교육활동 속에 자신을 보호하거나 자기 합리화 기제를 찾는 경향이 두드러지게 나타나고 있다. 교육적으로 바람직하나 평가 결과에 대한 민원이 제기되기 쉬운 수행평가나 논술형 평가에 애쓰기보다는, 교육적으로 그다지 바람직하다고 보기 어려우나 민원의 여지가 상대적으로 적은 선다형 평가에 쉽게 의존함으로써 스스로를 방어하려는 경향이 대표적인 예이다. ‘입시 때문에 못 하는 것’이 아니라 ‘입시를 핑계로 안 하는 것’이라는 자조적인 비판이 제기되는 이유도 이와 관련이 깊다. 하지만 A중학교 교사들은 이러한 분위기에도 불구하고 자발적으로 평가혁신을 위해 노력했다는 점, 개인의 실천을 넘어 공동체적 해법을 찾아 나갔다는 점, 이 과정에서 평가혁신에 대한 자신감과 자기 효능감을 회복했다는 점이 가장 큰 성과라 할 수 있다.

물론 A중학교의 실행연구에서도 몇 가지 극복해야 할 과제가 있다. 2024년에 진행된 A중학교 실행연구는 외부 연구자와의 협력적 실행연구였다. 향후에는 외부 연구자의 참여 없이도 교원 스스로 실행연구를 진행할 만한 자생력을 확보해야 하는데, 이는 결코 쉬운 과제가 아니다. 또한 실행연구에 참여한 15명의 교원의 성과를 전체 교사와 공유하고 이를 일반화하는 경로를 찾아야 하는 문제가 남아 있다. 이를 위해 A중학교에서는 새로운 평가계획 양식, 평가 루브릭 양식을 통해 연구 성과를 일반화하는 방법을 찾고 있다. 하지만 교원의 변동이 일상적으로 있을 수밖에 없는 공립학교에서 교육활동의 안정성을 확보해야 하는 것은 비단 A중학교만의 문제가 아니다. 따라서 평가혁신 일반화를 위한 교육당국 차원의 지원이 여기에 결합되어야 한다.

2. 평가혁신을 위한 제언

교육의 여러 영역 가운데 평가혁신은 상대적으로 어려운 과제에 해당한다. 앞에서 언급했듯이 ‘수업과 평가의 분리 현상’, 즉 수업은 달라졌으나 평가는 여전히 기존의 관행을 유지하는 현상은 전 세계적으로 보편적인 현상이다(Shepard, 2000). 더욱이 우리나라의 경우에는 입시 위주의 교육풍토, 상대평가의 영향이 오랫동안 자리 잡아 왔다. 중학교의 경우 2012년 성취평가제 도입 이후 절대평가가 적용되었음에도 여전히 상대평가 문화, 선다형 지필평가 위주의 관행이 지속되고 있다.

그러나 우리나라 중등학교에서 평가혁신이 불가능한 것은 결코 아니다. 이를 위해서는 개인적 차원의 실천을 넘어 공동체적 차원의 실천이 이루어져야 하고, 이론적 학습과 실행 중 성찰이 지속되어야 하며, 학교 관리자의 리더십이 뒷받침되어야 하고, 모든 교

사가 평가혁신의 방법론을 공유하도록 하는 매개적 도구가 필요하다. 무엇보다도 평가 혁신에 대한 비관주의, 방어적 교육활동에서 벗어나 적극적인 교육활동 속에 교육의 질을 높이고 이를 통해 교권, 즉 가르침의 진정한 권위를 확보하려는 노력이 필요하다.

그렇다고 하여 교사의 노력만으로 모든 것을 해결할 수 있다는 것은 아니다. 평가는 입시, 대학서열화, 학벌사회와 연결되는 문제이기 때문에 이와 관련된 사회적인 노력이 필요하다. 이는 반드시 해결해야 할 문제이지만 상당한 시간이 소요된다. 이러한 노력과 동시에 현재의 조건 속에서 교사의 평가혁신 실천을 최대한 끌어올리기 위한 교육당국의 지원이 필요하다.

첫째, 평가와 관련된 여러 지침을 개정할 필요가 있다. 이의 큰 방향은 교사에게 더 많은 평가 자율권을 부여하고, 그 자율권을 방임이 아닌 혁신의 노력을 위해 행사할 수 있도록 유도해야 하는 것이다. 그래야 A중학교 실행연구와 같은 과감한 시도가 더 많은 학교에서 이루어질 수 있다. 예컨대 A중학교에서 그러했듯이, 오지선다형 지필평가를 지양하고 수행평가와 논술형 평가를 활성화하는 방향을 유도하는 지침을 교육당국 차원에서 제시해야 한다.

이를 위해 교육당국은 학교 현장에서 진행되고 있는 다양한 평가혁신 사례를 수집하고, 이를 일반화할 수 있는 원리를 도출해서 적극적으로 안내해야 한다. 예를 들어, A중학교는 특히 성장중심평가의 취지에 따라 ‘피드백과 재도전의 기회’를 부여하는 수행평가를 적극적으로 시행했다. 그런데 일각에서는 이렇게 재도전의 기회를 부여하는 방식의 수행평가가 평가의 공정성을 위배한다고 보는 시각도 존재한다. 그러나 “학습 과정을 확인하고 환류하여, 학습자의 성공적인 학습을 지원한다”(교육부, 2022)는 국가 교육과정의 취지를 생각해 볼 때, 피드백과 재도전의 기회를 부여하는 방식의 평가는 오히려 장려되어야 한다. 실제로 서울특별시교육청에서는 “수행평가 결과에 대한 피드백을 제공한 후 학생에게 재도전의 기회를 줄 수 있음”을 명시하여 안내하고 있다(서울특별시교육청, 2025).

둘째, 고등학교 평가혁신을 위한 여건을 마련해야 한다. 고등학교의 평가혁신은 중학교의 평가혁신보다 더 어려운 과제이다. 국가 교육과정에서는 수행평가 내실화와 논술형 평가의 확대를 강조하고 있지만, 고등학교의 상대평가 시스템으로 인해 이러한 취지를 살리는 데에 어려움을 겪고 있다. 2025년부터 기존의 석차 9등급제보다 다소 완화된 형태인 석차 5등급제가 적용되고 있다. 석차 5등급이라는 불가피한 한계를 반영한 고등학교 평가혁신 전략이 마련되어야 하고, 향후 고등학교 절대평가 도입을 위한 전략도 동시에 마련되어야 한다.

셋째, 교육당국은 교사들이 ‘방어적 교육활동’에서 벗어나 ‘적극적 교육활동’으로 나설 수 있도록 법적, 제도적, 심리적 환경을 마련해야 한다. 앞에서 언급했듯이, 평가혁신에

대한 방어적 태도는 ‘입시경쟁으로 인한 제도적 한계’와 ‘평가 결과에 대한 민원’, ‘상급 기관의 규제적 관리 감독’ 등에서 비롯된다. 교사들은 여전히 교육당국이 평가혁신을 지원하기보다는 불필요한 지침, 감사 등을 통해 평가 자율권을 제약한다고 인식하고 있다. 또한 평가 결과에 대한 학부모의 민원으로부터 교사를 보호하는 제도적 장치가 없다고 인식하고 있다. 반대로 교육청에 근무하는 장학사들의 증언에 의하면 교육청에서 다양한 연수 등을 통해 평가혁신의 방향을 제시하고 이를 안내하고 있으나 교사들의 의지가 매우 소극적인 태도를 보인다고 한다.

이러한 인식의 격차는 단지 자신의 처한 위치에 따른 입장의 차이라기보다 여전히 평가와 관련해 ‘공유된 의미’가 교육계에 충분히 자리 잡지 못한 것으로 보인다. Fullan(2015)에 의하면, 교육을 변화시키는 것은 몇 가지 정책이나 지침이 아니라 그와 관련된 의미 공유이다. 예컨대 수행평가나 논술형 평가의 경우 교사에게 심리적 부담이 크기 때문에 이에 대한 의미 공유가 필수적이다. 비록 심리적으로 부담스러운 교육활동이라 할지라도 이의 교육적 의미가 공유되고 이를 내면화한다면 교사 대부분은 방어적 교육활동에서 벗어나 적극적 교육활동에 나서는 전문성과 윤리성을 발휘하게 된다. A중학교의 평가혁신 실행연구가 나름의 성과를 거둔 가장 중요한 이유는 실행 중 성찰을 통해 평가혁신에 대한 의미를 지속적으로 공유했기 때문이다.

물론 평가혁신을 위해 노력하는 교사 중에 입시와 관련된 제도적 한계에 부딪혀 무력감을 느끼는 경우가 많다. 입시 문제는 교육계의 노력만으로 해결할 수 없는 사회적인 문제이다. 하지만 “입시 때문에 평가혁신은 어렵다.”는 인식은 “입시 제도를 개선해야 한다.”라는 사회적 인식으로 확장되어야 하지, “그렇기 때문에 어쩔 수 없이 주입식 교육, 일제식 평가를 할 수밖에 없다.”라는 자기합리화 논리로 왜곡되는 것은 바람직하지 않다. 현재의 조건 속에서 시도할 수 있는 평가혁신의 과제가 무엇인지를 정확히 확인하고, 이를 학교 공동체적 차원에서 실천하고, 이러한 역량이 교육계에 축적되어야 평가제도의 변화(예컨대 고등학교 절대평가 전면 도입)를 이끌어 낼 동력을 형성할 수 있다.

이 연구는 A중학교에서의 실행연구를 통해 이러한 가능성을 확인해 보았다. 향후에는 고등학교에서의 평가혁신에 관한 현장 연구가 이어지기를 기대해 본다.

참고문헌

- 교육부(2022). **초·중등학교 교육과정 총론**. 교육부 고시 2022-33.
- 교육부(2024). **학교생활기록 작성 및 관리지침**. 교육부 훈령 504호.
- 권순보, 박수원(2025). 텍스트 마이닝을 통한 초·중등 교사의 과정중심평가에 대한 인식 연구. **교육과정평가연구**, 25(1), 185-214.
- 김경희, 조지민, 김영은, 김혜숙, 나우열, 박지선, 송민호, 이동욱, 이민형, 임은영, 김선희, 박찬호 (2025). **학생 역량 및 성장 중심의 학생평가 체제 개선 방안**. 연구보고 RRE 2024-8. 한국교육과정평가원.
- 김성천(2024). 교사들의 방어진 교육활동의 양상과 비판적 성찰. **교육비평**, 55, 119-177.
- 박혜영, 이명애, 이명진(2019). 우리나라 미래 초·중등학교 교육평가 방향 탐색. **교육과정평가 연구**, 22(3), 147-171.
- 서울특별시교육청(2025). **2025 중등 학생평가 내실화 계획**. 서울특별시교육청.
- 신혜진, 안소연, 김유원(2017). 과정 중심 평가 활용의 정책적 분석. **교육과정평가연구**, 20(2), 135-162.
- 양이레(2018). 영어과 과정중심수행평가의 인지적, 정의적, 인성적 환류효과. **English Language Assessment**, 13, 121-151.
- 육기엽(2016). 국어과 수업혁신에 따른 평가혁신 사례 고찰. **우리말교육현장연구**, 10(2), 37-70.
- 이주연, 임유나, 정영근, 이경언, 이민형(2020). 혁신학교 교육과정 실행 양상 분석: 교육과정, 수업, 평가를 중심으로. **교육과정평가연구**, 23(2), 1-26.
- 임태원, 이형빈(2024). 미래형 평가에 관한 중등 교사의 인식 및 지원 요구 탐색 연구. **한국교원교육연구**, 41(2), 297-322.
- 이형빈, 김성수(2022). **백워드로 설계하고 피드백으로 완성하는 성장중심평가**. 서울: 살림터.
- 정현철(2018). 체육 교과 수행평가 100% 반영 정책에 따른 학생 평가 쟁점 고찰. **교육평가연구**, 31(2), 463-482.
- Brookhart, S. M. (2013). *How to create and use rubric for formative assessment and grading*. ASCD.
- 장은경 외 역(2022). **루브릭 어떻게 만들고 사용할까?**. 서울: 우리학교.
- Denzin, N. K. (1978). *The research act: A theoretical instruction to sociological methods*. McGraw-Hill.
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize*

- student learning*. Corwin Press. 온정덕, 윤지영 역(2022). **학습과정으로서의 평가: 교실평가로 학습 극대화하기**. 서울: 학지사.
- Engeström, Y.(1987). *Learning by expanding: an activity-theoretical approach to developmental research*. Orienta-Konsultit.
- Feuerstein, R. (1979). *The dynamic assessment of retarded performer: The learning potential assessment device, theory, instruments, and technique*. University Park Press.
- Frey, N., & Fisher, D. (2011). *The formative assessment action plan: Practical steps to more successful teaching and learning*. ASCD. 강정임 역(2021). **피드백, 이렇게 한다**. 서울: 교육을 바꾸는 사람들.
- Fullan, M.(2015). *The new meaning of educational change*. Teachers College Press.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1). 81-112.
- Hawthorne R., Henderson J., & Stollenwerk, D.(2000). *Transformative curriculum leadership* (2nd Ed). Prentice Hall.
- Lewin, K.(1946). Action research and minority problems. *Journal of Social Issues*, 2(4), 34-46.
- Shepard, L.(2000). *The role of classroom assessment in teaching and learning*. University of California.
- Vygotsky, L. (1978). *Mind in society*. Harvard University Press. 정희욱 역(2009). **마인드 인 소사이어티**. 서울: 학이시습.
- Wiggins, G., & McTighe, J.(2000). *Understanding by design* (2nd Ed.). Association for Supervision & Curriculum Development. 강현석 역(2008). **거꾸로 생각하는 교육과정 개발**. 서울: 학지사.

논문접수 : 2025.4.3. / 수정본접수 : 2025.4.24. / 게재승인 : 2025.5.8.

ABSTRACT

Action Research on the Evaluation Innovation in Secondary Schools according to the Purpose of Growth-Based Evaluation

Hyoungbin Lee

Professor, Catholic Kwandong University

This study explored the extent of secondary school evaluation innovation that can be realized under the current conditions and what is necessary to generalize it. To this end, the researcher established an action research plan for a year with teachers of middle school A and repeatedly reflected during the implementation.

The specific tasks implemented by the teachers who participated in the action research were: first, reducing the multiple-choice paper and pencil test, second, restructuring the performative evaluation into diagnostic evaluation-formative evaluation-summative evaluation, third, conducting evaluation with feedback and re-challenge, and fourth, designing instruction and evaluations centered on big ideas and key questions. Based on the results of the action research, the evaluation plan form and evaluation rubric were newly developed. In addition, based on the results of the action research, <the principles and directions of evaluation> were organized and shared with all teachers.

Through this study, the following implications can be obtained. First, it is necessary to practice at the school community level. Second, theoretical learning and action research according to the school situation are necessary. Third, the leadership of school managers is very important. Fourth, the role of an mediated tool is important in order to generalize the performance of evaluation innovation. Fifth, it is necessary to break away from the climate of 'defensive educational activities' and build a culture of 'active educational activities'. To support this, educational authorities should revise guidelines and create conditions to share the meaning of evaluation innovation.

Key Words: evaluation innovation, growth-based evaluation, action research