



수행평가, 필요한가? 그리고 가능한가?

임 재 훈 (한국교육과정평가원)

수행평가의 필요성이나 취지 자체를 부정하는 교사는 별로 없다. 어렵고 복잡한 논의를 끌어들이지 않고, 수학 교사들에게 친숙하고 간단한 논리를 써서 생각해 보아도 수행평가의 필요성을 어느 정도는 인정할 수 있다. 약간의 논리를 사용하여 생각해 보자.

학생의 머리 속에 들어 있는 눈에 보이지 않는 지식을 눈에 보이는 문제에 정답을 하거나 못하는 행동으로 미루어 짐작하려고 하는 것이 평가이다. 평가를 할 때, 우리는 '학생이 알고 있다면, 문제를 풀 것이다' 라는 가정하에 평가를 한다.

$$p : \text{안다} \rightarrow q : \text{푼다(True)}$$

그런데, ' $p : \text{안다} \rightarrow q : \text{푼다}$ ' 는 참인 명제로부터 안심하고 말할 수 있는 것은 무엇인가? 학생이 문제에 정답을 했다고 해서 학생이 알고 있다는 결론을 내려도 좋은가? $p \rightarrow q$ 가 성립한다고 해서 그 역인 $q \rightarrow p$ 가 성립하는 것은 아니다. 그러므로, 학생이 문제에 정답을 했다는 사실로부터 그 학생이 알고 있다는 결론을 논리적으로 끌어낼 수 없다.

$$q : \text{풀었다} \xrightarrow{???} p : \text{안다}$$

' $p : \text{안다} \rightarrow q : \text{푼다}$ ' 는 명제가 참이라는 사실로부터 논리적으로 알 수 있는 것은 그 대우 명제인 ' $\sim q : \text{못풀었다} \rightarrow \sim p : \text{모른다}$ ' 는 명제가 참이라는 것이다. 평가 문제에 대한 학생의 반응으로부터 확실히 알 수 있는 것은 '학생이 무엇을 알고 있는가' 가 아니라 '학생이 무엇을 모르고 있는가' 이다.

$$\sim q : \text{못 풀었다} \rightarrow \sim p : \text{모른다(True)}$$

학생이 무엇을 얼마나 알고 있는지를 점수화하는 것이 아니라 '학생이 무엇을 얼마나 모르고 있는지를 알아 내어 그것을 보완해 주려는 것' 이 논리적으로 볼 때 평가의 본연의 기능이라고 할 수 있다. 지금까지의 객관식 선택형 문항으로도 학생이 무엇을 모르고 있는지를 알아 내어 그것을 보완해 주는 평가의 논리적인 기능은 어느 정도 수행될 수 있다.¹⁾ 물론, 객관식 선택

형 문항은 모르면서도 우연히 정답을 고를 가능성이 있기 때문에, 학생이 무엇을 모르고 있는지를 모두 알려 주지 못하는 단점은 있다. 객관식 선택형 평가가 완전한 것은 아니다. 그러나 그것이 나쁜 평가도 아니다. 채점에 들어가는 시간이나 경비·채점의 객관성에 대한 요구 등 현실적 여건을 고려하면, 객관식 평가는 경제적이고 현실적인 평가 방안이라고도 할 수 있다.

이상과 같이, 평가의 논리적인 기능은 학생의 아는 정도를 알려 주는 ‘포지티브’한 데 있기보다 학생의 모르는 정도를 알려 주는 ‘네가티브’한 데에 있지만, 인간의 마음 속에는 여전히 평가를 통해 ‘학생이 무엇을 얼마나 알고 있는지’를 알고자 하는 욕구가 있다. 이와 관련하여 다음과 같은 개연적 추론을 생각해 보자.

$$\frac{p \rightarrow q}{q} \\ \therefore p \text{는 보다 신뢰할 만하다.}$$

폴리아라는 수학자가 든 다음 예는 개연적 추론이 어떠한 것인지 잘 보여 준다.

만약 우리가 육지에 접근하고 있다면 우리는 자주 새들을 볼 것이다.
지금 우리는 새들을 보고 있다.

\therefore 우리가 육지에 접근하고 있다고 믿을 만하다.

개연적 추론에서 문제가 되는 것은 확실성의 정도이다. 다음을 보자.

학생이 아래의 문제 1에 정답을 했을 때보다 문제 2에 정답을 했을 때, 교사는 학생이 알고 있을 것이라는 추측을 좀 더 자신있게 할 수 있을 것이다. 모르고도 우연히 맞는 번호를 골랐을 가능성이 문제 1에서는 50%이지만 문제 2의 경우는 20%로 낮기 때문이다.

문제 1. $2(x+4)-3=7$ 인 x 는 다음 중 어느 것인가?

- ① 1 ② 2

문제 2. $2(x+4)-3=7$ 인 x 는 다음 중 어느 것인가?

- ① -1 ② 0 ③ 1 ④ 2 ⑤ 3

그러나, 아무리 선택지를 많게 해도 모르면서 우연히 정답을 고를 가능성을 완전히 없앨 수는 없다. 이런 점에서 보면, 객관식 선택형 문제보다는 다음과 같은 서술형 주관식 문제

1) 실제로 우리가 이러한 기능으로 평가를 사용하였는지는 아니면 등수 매기기에 사용했는지는 별개의 문제이다.

에 학생이 정답을 했을 때 교사는 학생이 알고 있다는 추측을 좀 더 자신있게 할 수 있다.

문제 3. $2(x+4)-3=7$ 인 x 를 구하여라. 풀이 과정을 써라.

이로써 우리는 서술형 주관식 문제의 필요성에 도달했는데, 서술형 주관식 평가는 수행평가의 한 가지 방법으로 거론되고 있는 평가 방법이다.

한편, 수행평가에서는 다양한 평가 방법을 사용할 것을 권장하고 있다. 개연적 추론 논리를 따라 조금만 생각하면 왜 그런 주장이 나오는지 짐작할 수 있다. 개연적 추론에서 추론의 확실성의 정도는 한 가지 징조만 나타났을 때보다 여러 가지 징조가 나타났을 때 증가한다. 다음 예를 보자.

만약 우리가 육지에 접근하고 있다면, 우리는 자주 새들을 볼 것이고 쓰레기가
떠내려 오는 것을 볼 것이다
지금 우리는 새들을 보고 있다.

∴ 우리가 육지에 접근하고 있다고 믿을 만하다.

만약 우리가 육지에 접근하고 있다면, 우리는 자주 새들을 볼 것이고 또한 쓰레기가
떠내려 오는 것을 볼 것이다.
지금 우리는 새들을 보고 있다. 그리고 떠내려 오는 쓰레기를 보고 있다.

∴ 우리가 육지에 접근하고 있다고 앞의 경우보다 더 믿을 만하다.

이러한 생각을 평가에 적용하면, 한 가지 평가 자료보다는 다양한 평가 방법에 의해 얻은 다양한 평가 자료를 가지고 추론할 때 학생이 무엇을 얼마나 알고 있는지 더 확실하게 추론할 수 있을 것이라는 생각이 나올 수 있을 것이다. 우연의 일치인지 모르나, 수행평가에서는 한 가지 평가방법이 아니라 서술형 주관식, 포트폴리오, 관찰 면접법, 프로젝트 등 여러 가지 다양한 평가 기법을 사용할 것을 적극 권장하고 있다.

이상 수행평가의 필요성에 대하여 논리적으로 간단히 생각해 보았다. 이제 현 시점에서 수행평가를 전면적으로 실시하는 것이 가능한지에 대해서 생각해 보겠다.

어떤 교육개선 운동이든 그것이 성공하기 위해서 갖추어야 할 필요 조건이 있다. 그것은 그 교육개선 운동이 교사들의 마음을 슬프게 해서는 안된다는 것이다. 아무리 좋은 취지의 교육개선 운동이라도 그것이 학교 현장에서 직접 그것을 실행해야 할 교사들의 심기를 불편하게 만든다면, 그 운동은 실제 교육 개선을 위해 절대적으로 요청되는 '교사들 개개인의 교육 체계의 실질적인 변화'를 가져오지 못할 가능성이 높다. 몇몇 정책 입안자들이나

운동가들이 소리 높여 개선을 부르짖을지라도, 현장의 교육 시스템은 거의 바뀌지 않게 된다. 겉으로는 무엇인가 큰 변화가 일어나는 것처럼 요란하지만 실질적으로 변하는 것은 별로 없다는 것을 이제까지 정부 주도로 추진된 여러 교육개선 운동을 경험하면서 많은 교사들이 체득하게 되었다고 말하면 지나친 말일까? 어쩌면 수행평가의 앞길에도 비슷한 운명이 기다리고 있는지 모른다.

수행평가는 몇 년 전부터 계속 이야기되어 온 것이고, 수행평가 방법으로 거론되는 평가 방법들도 그리 새로운 것은 아니다. 이미 학교 현장에서는 서술형 주관식 평가를 실시해 오고 있다. 숙제 내주고 노트 검사하는 것을 좀 더 세련된 형태로 다듬어 하면 수행평가에서 말하는 포트폴리오가 될 수 있다. ‘평소점수’를 주어 오던 것을 체크 리스트를 가지고 체계적으로 하면 관찰법이 될 수 있다.

그러나, 수행평가는 이제까지 없던 아주 새로운 평가 방법인 듯이, 또 기존의 평가 방법과 대비되면서 기존의 평가 방법은 나쁜 것이고 수행평가만 바람직한 평가 방법인 듯이 소개되고 받아들여지는 경향이 있는 것 같다. 새로운 주장에 힘을 부여하기 위해서 기존의 것을 ‘나쁜’ 것으로 몰아붙일 필요가 있다는 것은 이해할 수 있다. 정책 입안자나 연구자의 입장에서는 자신이 입안한 계획이나 주장이 기존의 보다 아주 조금 더 나은 것이라고 소개하는 것보다 기존의 것과는 다른 혁신적인 것이며 기존의 것이 지닌 문제점을 일소할 수 있는 아주 좋은 것이라고 선전하는 것이 재정적 지원을 포함한 여러 가지 지원을 받는 데 유리하다.

그러나 새로운 방법의 ‘좋은’ 점을 기존의 ‘나쁜’ 점에 대비적으로 부각시켜 교사들을 압박하는 것은, 교사들에게 “당신들은 이제까지 교육을 잘못 해온 것이요. 그러니 이제부터는 이리이러하게 고치시오” 라는 메시지를 암암리에 전달한다. 교사들은 교육 개선과 관련된 급진적인 주장들로부터 자신들이 이제까지 해온 교육은 나쁜 것이었으며 그러한 교육을 해온 자신들은 죄인이라는 메시지를 전달받는 것이다. ‘현재의 교육은 잘못된 것이며, 현재의 나쁜 교육을 새로운 방법을 도입해 좋은 교육으로 바꾸겠다’는 생각은 교사들을 슬프게 만든다.

교육개선 운동이 ‘교사 개개인의 교육체계’라는 교육의 실질적인 하부구조를 변화시키는 성공을 이루기 위해서는 기존의 교육을 ‘나쁜’ 것으로 보는 사고방식을 버릴 필요가 있다. 기존의 교육을 나쁜 것으로 보는 사고방식은 필연적으로 마음을 조급하게 만든다. 나쁜 것을 하루라도 더 나쁜 상태 그대로 있게 둘 수는 없는 노릇 아닌가! 이러한 생각은 새로운 방법이 도입될 수 있는 여건이 어느 정도 이루어져 있는지 자세히 따져보지 않은 채 모든 것을 하루 아침에 뜯어 고치려는 급진적인 방식의 개혁을 선호하게 만든다.

현재의 교육이 아무리 잘못되고 나쁜 것으로 보일지라도, 좀 더 따뜻한 시각으로 교육을 볼 필요가 있다. 기본적으로 교육은 그것이 교육이라면 일단 좋은 것이다. 이렇게 볼 때, 마음에 약간의 여유를 가질 수 있다. 물론 세상에 완전한 것은 없는 법이니, 기존의 교육 또한 개선될 점을 언제나 갖고 있다고 보아야 한다. 그러나 이 때의 사고방식은 '나쁜 것을 좋게(bad → good)'가 아니라 '좋은 것을 더 좋게(good → better)' 바꾸는 것이다.

수행평가가 마치 전에 없다가 새로 생겨난 것인 양, 지금까지의 나쁜 평가를 대치할 좋은 평가인 양 주장될 필요는 없다. 또한, 교사들도 수행평가에 대해서 지나친 부담을 가지지 않는 것이 좋을 것 같다. 자신이 할 수 있는 선에서 한 가지 정도를 더 해 보는 것으로 충분할 것이다. 예를 들어, 학생들을 조를 짜서 프로젝트 과제를 한 학기에 하나 정도 하게 해볼 수 있을 것이다. 수학 교사의 입장에서는 한 학기에 프로젝트 하나를 내주는 것이지만, 모든 교과가 수행평가를 위해 프로젝트를 내주면, 학생 입장에서는 한 학기에 열 개도 넘는 프로젝트를 해야 한다. 또한, 점수의 실질 반영 폭도 크게 하지 않는 것이 낫다. 예를 들어, 10 점을 반영한다면 프로젝트를 수행하여 제출하면 7~8 점을 기본 점수로 주고 아주 잘 한 경우 10 점을 주는 것이 적절할 것이다. 점수 차이를 크게 하면 우리 나라처럼 교육열이 높은 곳에서는 부모, 형, 누나 등 전가족이 프로젝트에 매달리게 될 수도 있고, 채점하는 교사의 입장에서도 채점상의 부담이 가중된다. 초등학교에서는 별 문제가 안될 지 모르지만, 내신 성적이 상급학교 진학에 결정적인 요인이 되는 중등학교에서 교사가 떠안게 되는 채점상의 부담은 적지 않을 것이다.

수행평가의 실시는, 수학교사의 입장에서는 “수행평가 방법을 한두 가지 해보니 나의 교육 체계가 이전에 비해 조금 더 좋아진 것 같아 좋더라”는 느낌을 가질 수 있고 학생 입장에서는 “프로젝트를 하면서 재미있었다”는 정도의 경험을 하게 하는 것을 목적으로 해서 시작하면 충분할 것이다. 해보아서 좋고 흐뭇하고 가능하다면, 하지 말라고 해도 교사들이 알아서 더 하려고 할 것이다.

정책 당국자들은 현 시점에서 수행평가를 (중등학교까지) 전면적으로 실시할 여건이 과연 갖추어져 있는지 솔직하게 생각해 볼 필요가 있다. 수행평가 문항이 하루 이틀에 하나씩 만들 수 있는 것인가? 수업과 행정 업무 부담을 현재처럼 다 지고 있으면서 수행평가 문항 개발에 충분한 시간을 투여할 수 있는 교사가 몇 명이나 될까? 교사들에게 알아서 수행평가 문항을 만들어서 쓰라고 하는 것은, 고대 이집트의 독재자가 히브리 백성에게 벽돌을 만들 재료를 주지 않으면서 알아서 벽돌을 만들어내라고 하던 것과 별반 다르지 않다. 당장 수행평가를 (중등학교까지) 전면적으로 실시하고자 한다면, 채점 기준표 및 모범 답안이 포함된 수행평가 문항을 충분히 제공해 주어야 한다. 제공된 채점 기준표에 따라 많은 경우

수백 명이나 되는 학생들의 답안을 채점하는 일만도 교사들에게는 충분히 벅찬 일이다.

또한, 현재 수행평가 연수는 교사들을 대상으로만 이루어지고 있는 것 같은데, 이와 병행하여 감사 업무를 수행하는 일반직을 대상으로 한 연수가 이루어져야 할 것 같다. 수행평가는 기본적으로 평가에 있어 '교사의 주관성'을 인정하는 것이다. 아무리 채점 기준표를 잘 만들어도 예를 들어, 1점보다는 더 잘했는데 2점보다는 못한 답안들이 있기 마련이다. 교사도 사람인 이상 이런 답안을 채점할 때 어떤 것은 1점을 주고 어떤 것은 2점을 주는 경우가 생기기 마련이다. 이런 것을 감사에서 문제 삼는다면 무사할 교사가 없을 것이며, 겉으로만 마지못해 수행평가를 할 뿐 아무도 진심으로 수행평가를 하려고 하지 않게 될 것이다.

끝으로, 수행평가와 같이 평가를 통해 교육을 개선하려는 시도에 대해 생각해 보겠다.

평가에는 고유의 이론이 있을 수 있을까? 평가에 관한 기법적인 이론은 있을 수 있어도 그것을 넘어서는 고유의 교육적 이론이라고 할 만한 것은 있기 어렵다. 평가는 기본적으로 교수 학습에 종속되어야 하기 때문이다. 가르치고 배운 것을 가르치고 배운 대로 평가하는 것이므로, 평가는 원칙상 교수 학습에 종속되고 그에 일치되어야 한다. 비유컨대, 평가 이론은 본질상 교수 학습 이론에 종속되는 '따름정리'라고 할 수 있다.

이 점에서 보면, 평가방법을 바꿈으로써 교수 학습을 변화시키겠다는 생각은 일종의 본말전도로 볼 수 있다. 평가와 관련하여 우선 중요한 것은 평가방법을 바꾸는 것이 아니라 평가가 교육에서 수행해야 할 본연의 기능과 위치를 제대로 자리매김해주는 것일 지도 모른다. 평가가 이끌어가는 교육은 '입시 위주의 교육'이든 '수행평가 위주의 교육'이든 별로 바람직한 것이 아니다.

또한, 수행평가를 도입하면 학교 수학교육에 무엇인가 본질적인 진보나 개선이 이루어질 것이라고 기대하는 것도, 일종의 본말전도이며 평가의 힘을 지나치게 맹신하는 데서 비롯되는 착각이다. 수행평가의 도입은 수학교육의 개선에 도움이 될 수 있을 지 모르나, 학교 수학교육의 본질적인 진보를 위한 보조적인 역할이라면 몰라도 중심적인 역할까지 기대하기는 어렵다. 수행평가가 피타고라스 정리를 더 의미있게 가르치고 배우는 것을 보장해 주는 것도 아니며, 함수 개념의 이해가 더 잘 이루어지도록 보장해 주는 것도 아니다.

수행평가에 대해 지나친 기대도 부담도 거부감도 가질 필요가 없다. 어쨌든 수행평가가 강조되고 지원되는 것은 현재의 시류이다. 시류를 거부할 특별한 이유가 있지 않다면 애써 시류를 거부할 필요는 없다. 또한 몇 년 후면 사그러들 지도 모르는 시류에 지나치게 민감할 필요도 없다. 다만, 주어진 이 시류를 각각의 수학 교사들이 자신의 교육 체계를 현재보다 '조금 더 낫게' 만드는 기회로 지혜롭게 활용하면 그것으로 충분한 것이다.