



## 광합성 문제 해결에 대한 과잉 효능감과 과잉확신: 선다형과 서답형의 비교 연구

하민수<sup>1</sup>, 이준기\*

<sup>1</sup>뉴욕주립대학교 스톤브룩, 전북대학교

### Over-Efficacy in Problem Solving and Overconfidence of Knowledge on Photosynthesis: A Study of Comparison Between Multiple-Choice and Supply-Type Test Formats

<sup>1</sup>Minsu Ha, Jun-Ki Lee\*

<sup>1</sup>State University of New York at Stony Brook, Chonbuk National University

#### ARTICLE INFO

##### Article history:

Received 19 August 2013

Received in revised form

18 November 2013

24 December 2013

Accepted 23 January 2014

##### Key words:

over-efficacy, overconfidence, assessment format, multiple-choice test, supply-type test

#### ABSTRACT

This study aimed to explore the over-efficacy in problem solving and overconfidence of knowledge of students performing assessments in two different test formats: multiple-choice and supply-type. Two hundred and four female middle school students participated in this study. Multiple-choice and supply-type formats of tests on photosynthesis were used, and each item contained scales indicating one's self-efficacy on problem-solving and confidence of knowledge. The results showed that the correlation coefficients of performance between the two different assessment formats were less than 0.5 and the correlation coefficients between efficacy/confidence and actual performance were less than 0.45. Moreover, students tended to exhibit more over-efficacy and overconfidence in multiple-choice formats. The percentage of over-efficacy and overconfidence was higher in the group that completed the multiple-choice test first followed by the supply-type assessment than in the group that started with the supply-type followed by the multiple-choice assessment. From this study, it can be suggested that more use of supply-type assessment is required in science education. If test administrators require the combination of both multiple-choice and supply-type in an assessment, the supply-type assessment format should come first so that students can maintain the appropriate level of efficacy and confidence. In addition, science educators need to develop new learning programs to enhance students' self-monitoring skills of their problem-solving ability and knowledge.

## I. 서론

문제 해결 과정에서 자신이 해당 문제를 해결할 수 있을 것이라는 자아 효능감과 자신의 답이 맞을 것이라는 확신은 성공적인 문제 해결을 위하여 중요한 요소 일 수 있다. 그 이유는 자아 효능감과 확신이 낮을 경우 문제 해결에 능동적이지 못하게 되며, 지속적인 문제 해결을 위한 동기가 부족할 수 있기 때문이다(Johnson & Fowler, 2011). 하지만 높은 수준의 자아 효능감과 확신이 무조건 좋은 것은 아니다. 그 이유는 자아 효능감과 확신이 높을 경우 자신이 틀렸을 수도 있다는 반성적인 사고가 부족할 수 있으며, 성급하게 문제를 해결하여 오히려 문제 해결을 그칠 수 있기 때문이다. 예를 들어서 과학 문제를 해결하는데 잘못된 답을 선택하거나, 전혀 다른 의미로 이해하여 오답을 작성하는 등과 같이 문제 해결을 실패했음에도 불구하고 자신의 문제 해결의 실패를 인지하지 못할 수 있다. 이런 경우를 과잉효능감(over-efficacy), 과잉확신(over-confidence)이라 정의한다(Langendyk, 2006; Nowell & Alston, 2007). 이와 같이 과잉효능감이나 과잉확신이 높은 수준으로 나타나는 경우에는 자신감과 같은 장점에 비하여 신중하지 못한 문제 해결과 같은 단점 역시 중요하게 고려되어야 한다(Johnson & Fowler, 2011).

일례로 Dunlosky & Rawson(2012)는 과잉확신이 낮게 나타나는

학생의 경우 지식을 더 오랫동안 유지하고 있었으며, 과잉확신이 높은 학생들은 성취도가 상대적으로 더 낮았다고 보고하였다. 자신의 지식에 대한 확신이 높거나 문제 해결에 대한 지나친 효능감을 가지고 있을 경우 자신의 능력을 과장하게 되어 올바른 의사결정을 그르치게 된다. 예를 들어서 Ha *et al.*(2012)는 의사 결정 과정에서 두 가지 모델을 제시하였는데, 문제 해결에서 의사결정자가 가지고 있는 합당한 지식에 근거하여 해결하는 것과 의사결정자의 직관적인 느낌에 의존하는 모델로 구분된다. 직관적인 느낌, 확신 등에 근거한 의사 결정 과정은 대개 실수를 유발하게 되며, 그런 실수는 반복적으로 나타날 가능성이 높다. 그러므로 지나친 효능감과 확신은 효율적인 문제 해결과 의사 결정 능력 함양의 장애 요인으로 인식되어 왔다. 이런 중요성에 근거하여 과잉확신에 대한 연구는 교육학뿐만 아니라 의학과 경제학 등에서도 다양하게 연구되어 왔다(Langendyk, 2006; Nowell & Alston, 2007). 지식에 대한 지나친 확신에 대한 선행 연구들을 잠시 요약하면, 남자가 여자에 비하여 이런 과잉확신이 더 많이 나타난다고 보고되었다(Pallier, 2003). 또한 일부 예외 국가가 있던 하지만 전반적으로 아시아권 학생들이 서구권 학생들에 비하여 더 높은 과잉확신 경향이 있다고 알려졌다(Yates *et al.*, 1997).

문제 해결에 대한 자기 효능감과 확신 그리고 과잉효능감과 과잉확신에 관한 연구를 수행하고자 하는 이 연구에서 주목하는 또 다른

\* 교신저자: 이준기(junki@jbnu.ac.kr)

<http://dx.doi.org/10.14697/jkase.2014.34.1.1.0001>

변인은 ‘문항 유형’이다(Zakay & Glicksohn, 1992). 문제 해결에 있어서 어떤 형태의 문제를 접하는가에 따라 효능감과 확신은 달라질 수 있다. 익숙한 문제나 쉬운 문제일 경우 효능과 확신은 높을 것이며 평소 접하지 않거나 어려운 문제를 해결할 경우에는 효능과 확신이 낮을 것이다.

이 연구에서는 선다형(multiple-choice test)과 서답형(supply-type test) 문항에서 학생들의 자기 효능감과 확신에 대한 조사를 하고자 하는 이유는 최근 서답형 문항의 중요성이 크게 부각되고 있기 때문이다(Beggrow *et al.*, in press). 특히 우리나라에서 서답형 문항의 중요성이 부각되는 이유 중 하나는 우리나라 학생들의 서답형 문항의 성취도가 선다형 문항의 성취도에 비교하였을 때 다른 국가의 학생들보다 상대적으로 낮다는 보고가 있었기 때문이다(Choi *et al.*, 2011). 서답형 문항을 해결하기 위해서는 다양한 능력들이 요구된다. 먼저 서답형 문항을 수행하기 위해서는 ‘회상(recall)’ 능력이 필요하다(Opfer *et al.*, 2012). ‘회상’ 능력은 문제에 포함되어 있는 정보를 이용하여 피평가자의 인지에 있는 생각, 단어, 설명 등을 찾아내어 그것을 밖으로 표출시키는 능력이다. 서답형 문항에서는 ‘답지’가 없기 때문에 회상 능력은 필수적으로 요구된다. 반면 선다형 문항의 경우에는 문제에 제시된 정답과 오답지를 통하여 자신이 가지고 있는 정보를 연결시키기만 하면 된다. 아무런 정보가 없는 상황에서 지식을 회상해 내야 하는 것은 주어진 문항과 답지를 비교하면서 어떤 답지가 더 정답에 가까운지 판단하는 것에 비하여 더 어려울 수밖에 없다. 또한 지식의 ‘회상’을 위해서는 장기 기억의 용량도 커야 되며, 복잡하고 어려운 과정을 수행하기 위한 동기 역시 높아야 된다(Opfer *et al.*, 2012). 그 외에도 서답형 문항이 단답형이 아닌 긴 문장을 서술하는 형태일 경우 글쓰기 능력 역시 요구된다(Ha *et al.*, 2011).

문항의 형태와 확신에 관한 선행연구를 보면, 선다형에서 지식에 대한 확신이 지나치게 높을 경우 틀린 답을 선택하는 비율이 많아질 수 있다는 보고가 있었다(Zakay & Glicksohn, 1992). 선다형에서 과잉 확신이 문제가 되는 이유를 간단히 설명하면, 일반적으로 문제 해결자는 선택한 답지에 대한 긍정적인 증거를 수집하려는 경향이 강하며, 그와 반대로 선택한 답지를 부정하는 증거를 수집하려고 하지 않으려는 성향이 있기 때문이다(Koriat *et al.*, 1980). 선다형 검사에서 답이라고 생각하는 보기를 임의적으로 선택한 이후부터 그 답지를 정답일 것이라는 가정 하에 긍정적인 증거를 수집하려고 노력하며, 그 과정에서 선택한 답지를 더 확신하게 되는 것이다. 이와 같은 관점에서 본다면, 서답형의 경우 문항에 제시되는 정보가 상대적으로 적기 때문에 자신이 생각한 답지를 지지하는 증거를 찾을 수 없다. Koriat *et al.*(1980)의 연구는 그 이전에 수행된 Oskamp(1965)의 연구와 일맥상 통하는데, Oskamp(1965)은 정보가 많이 제시된 경우 일반적으로 과잉 확신이 발생한다고 하였다. Oskamp(1965)의 연구에 비추어보았을 때에도 정보가 상대적으로 많은 선다형 보다 정보가 상대적으로 적은 서답형에서 학생들의 과잉확신은 낮게 나타날 수 있을 것이다.

우리나라 학생들에게도 선다형 문항에서 과잉확신이 많이 나타날 수 있을 것으로 이 연구에서 추측하는 이유는 일반적으로 교육 현장에서 서답형 보다 선다형 문항 유형을 더 많이 활용하고 있기 때문이다. 실제로 우리나라 학생들의 서답형 문항의 문제 풀이가 선다형에 비하여 상대적으로 더 어려운 이유에 대하여 과학교육학자들은 그 문제의 심각성을 이해하고 있다. 예를 들어, Hong *et al.*(2001)은 우리나라

학생들의 설명 구성 능력이 알고 있는 내용 지식의 양에 비하여 상대적으로 낮음을 지적하였고, Park(2010)은 서답형 문항으로 측정된 학생들의 능력 정도는 선다형 문항으로 측정된 능력의 70% 정도 밖에 미치지 못하였음을 지적하였다. 서답형 평가가 선다형 평가에 비하여 과학적 능력을 확인하는데 더 효과적임에도 불구하고(Park & Hong, 2002), 우리나라 학생들은 오랫동안 선다형 문항에 노출되어 선다형 문항에 더 높은 성취도를 보이는 것으로 추측할 수 있다(Lee, 2008; Lee, 2011). 최근 이런 문제점을 해결하기 위하여 서답형 평가 도구의 개발과 평가 준거를 개발을 장려하며, 서답형 문항의 인간 채점을 대체할 수 있는 컴퓨터 채점 시스템을 개발하기도 하였다(Nehm *et al.*, 2012).

문제 해결에 대한 효능감과 지식에 대한 확신, 그리고 그것들의 지나침과 평가 방법에 따른 차이 등 다양한 연구 문제들이 제기될 수 있다. 또한 우리나라에서는 자아 효능감에 관한 연구는 상대적으로 많이 이루어진 반면, 과잉효능감, 과잉확신에 관해서는 상대적으로 미진하게 이루어져 왔다. 이 연구에서는 이와 같은 필요성을 근거로 중요한 연구 문제들을 확인하고자 한다.

첫 번째로 이 연구에서 확인하고자 하는 것은 선다형과 서답형 문항의 성취도의 상관관계 수준이다. 상관관계 수치를 확인하면 선다형 문항과 서답형 문항의 성취도에서 공통적으로 설명되는 능력의 부분이 어느 정도인지 확인할 수 있다. 또한 상관관계에서 설명력을 확인하면 서답형 평가에서 더 요구되는 지식의 회상 능력, 장기 기억 능력, 글쓰기 능력 등이 어느 정도로 요구되는지도 간접적으로 확인할 수 있다. 앞서 이루어진 선택형과 서답형 문항 점수의 상관관계 연구를 살펴보면 두 형태의 검사도구의 상관관계는 아주 높지 않다는 것을 확인할 수 있다. Ha *et al.*, (2012)의 연구를 보면, 우리나라 예비생물교사들의 자연선택에 관한 지식을 선다형 검사도구를 사용하여 측정하였을 때와 서답형 검사도구를 사용하여 측정하였을 때 상관관계가 0.52 ( $p < 0.01$ )로 나타났다. Nehm & Schonfeld (2008)의 연구에서는 미국 대학생들의 자연선택 개념을 선다형과 서답형으로 평가한 결과의 상관관계는 0.58 ( $p < 0.01$ )이었다. 이 두 편의 연구를 통하여 일반화 할 수는 없지만 이 정도의 상관관계가 나타난다면 선다형 평가와 서답형 평가 결과에서 공통부분은 약 40%가 채 되지 않는다. 즉, 상당히 많은 부분들이 다른 요인들에 의하여 설명되고 있음을 확인할 수 있다. 예를 들어 회상 능력, 장기기억, 응답에 대한 동기 등이 그런 요인들 중 하나일 것이다(Opfer *et al.*, 2012). 이 연구 문제와 함께 알아볼 것은 선다형과 서답형의 성취도 상관관계에 비하여 두 평가 방법의 자기 효능감, 확신의 상관관계는 어떻게 다른가 하는 점이다. 문제 해결의 성공 여부와 그것에 대한 효능감 또는 확신의 심리적 수준은 다를 것이다. 지식의 수준은 망각과 같은 과정에 의하여 낮아질 수 있지만 효능감과 확신은 개인의 심리적 상태로서 외부 환경에 따라 쉽게 변하지 않는 안정된 특성이 있다(Luszczynska *et al.*, 2005). 그러므로 상관관계 비교를 통하여 지식, 효능감, 확신에 대한 보다 명확한 이해를 가질 수 있을 것이다.

두 번째로 확인하고자 하는 것은 선다형과 서답형 중 어떤 도구에서 더 높은 과잉효능감과 과잉확신을 보이는지에 관한 것이다. 우리나라 학생들은 오랫동안 선다형 문항에 익숙해져 왔으며, 선다형에서 더 높은 성취도 수준을 보인다. 그러므로 선다형 문항에서 ‘잘 할 수 있을 것이다’라는 효능감과 ‘내 답지는 맞을 것이다’라는 확신이 더 높을 것으로 예상된다. 또한 그런 심리적 상태는 비록 자신이 틀린 답지를

선택하였음에도 유지될 가능성이 높을 수 있을 것이다. 더욱이 앞서 Koriat *et al.*(1980)의 연구와 Oskamp(1965)의 연구를 통하여 설명하였듯이, 자신이 선택한 응답에 대한 긍정적인 정당화 과정이 상대적으로 쉽고, 정보의 양이 많은 선다형 평가에서 학생들의 과잉확신은 더 많이 나타날 수 있을 것으로 예상된다. 그러므로 이 가설은 실험적으로 확인될 필요가 있을 것이다.

세 번째로 확인하고자 하는 연구 문제는 문제 해결 전 자기 효능감과 문제 해결 후 확신에 대한 차이가 평가 도구에 따라 어떤 차이점이 나타나는지이다. Koriat *et al.*(1980)의 연구와 Oskamp(1965)의 연구를 통해 논의한 바와 같이, 선다형은 문제 해결 과정에서 자신이 생각하는 답지에 긍정적인 증거들을 수집하여 자신이 생각하는 답지에 ‘적합한 모델’을 가지게 되며, 그로 인하여 문제 해결 후 답지에 대한 확신이 문제를 해결하기 이전에 생각하는 자신의 능력보다 높을 수 있다. 하지만 서답형의 경우 문항에서 제시되는 정보의 양이 상대적으로 적기 때문에 자신이 생각하는 답지를 지지하는 증거를 찾을 가능성은 낮을 수밖에 없으며, 답지 작성 전과 후 모두 동일한 수준의 자기 평가가 이루어질 것이다. 이 연구 가설 역시 이 연구를 통하여 확인할 것이다.

마지막으로 확인하고자 하는 연구 문제는 선다형과 서답형에 노출되는 시기상의 문제이다. 우리나라 학생들은 중간고사나 기말고사를 치루기 이전에 많은 문제들에 노출된다. 앞서 논의한 바와 같이 선다형에 노출될 경우 더 높은 효능과 확신이 생길 수 있으며, 높은 수준의 효능감과 확신은 일부 과잉효능감과 과잉확신으로 나타날 수 있다. 반대로 서답형에 먼저 노출될 경우 낮은 수준의 효능감과 확신을 유지할 수 있고, 차후에 실시하는 평가에서도 과잉효능과 과잉확신이 나타날 가능성이 줄어들 수 있다. 만약 이 가설이 지지 받는다면, 평가 문항을 어떤 순서로 제공해야 되는지에 관한 시사점을 찾을 수 있을 것이다. 이 연구 문제를 해결하기 위하여 선다형을 먼저 수행하고 그 이후에 서답형을 수행한 집단과 서답형을 먼저 수행하고 그 이후에 선다형을 수행한 집단을 비교하여 분석해 보아야 한다.

자신의 지식을 점검하고 문제 해결을 조절하는 능력은 성공적인 자기 주도 학습과 문제 해결에 있어 중요한 요소이다(de Bruin & van Gog, 2012). 하지만 문제 해결에 있어 성공하지 못하는 능력을 가지고 있음에도 불구하고 지식에 대한 과잉확신은 성공적인 학습에 장애요인이 될 수 있다. 이 연구는 문제 해결에서 효능감과 확신, 과잉효능감과 과잉 확신이 선다형과 서답형 평가에서 어떻게 나타나는지 확인하고 과학교육에 시사점을 찾는 것을 목적으로 한다.

## II. 연구 방법

### 1. 검사도구

이 연구를 위하여 사용한 검사도구는 광합성 개념을 확인하는 검사도구로 Marmaroti & Galanopoulou(2006)가 개발하여 그리스 중학생들의 광합성 개념을 확인하기 위하여 사용한 도구를 사용하였다. 이 검사도구의 원본은 8문항의 선다형으로 구성되어 있으며, 이 검사도구의 답지를 제거하고 일부 문항을 서답형에 어울리도록 변환하여 동일 개념을 측정하는 서답형 검사도구를 만들었다. 검사도구의 번역은 생물교육 박사과정생이 하였으며, 수정하여 만든

두 가지 형태의 광합성 지식 검사도구는 생물교육 전문가, 생물교육 박사 과정생, 생물교사의 검토를 통해 확정하였다. 이 검사도구의 채점에서 일부 문항은 부분 점수를 인정하였다. 두 검사도구의 내적 일관성 신뢰도(Cronbach alpha)는 선다형이 0.649, 서답형이 0.731이었다(부록 참조).

문제 해결 전에 수행하는 자아 효능감 측정 도구는 자아 효능감에 관한 문헌을 근거로 하여(Bandura, 1997; Pajares, 1996; Schunk, 1991) 제시된 문제를 성공적으로 해결할 수 있는지 묻는 질문에 5단계 리커트 척도를 사용하였다. 자아 효능감에 관한 척도의 신뢰도(Cronbach alpha)는 선다형 문항이 0.930, 서답형 문항이 0.903이었다. 문제 해결 후 자신이 선택한 답지에 대한 확신을 알아보기 위하여 Bruttomesso *et al.*(2003)과 Reach *et al.*(2005)이 사용한 문항을 번역하였고 척도는 5단계 리커트 척도를 사용하였다. 이 검사도구의 신뢰도는 선다형이 0.915, 서답형이 0.906이었다.

### 2. 연구 설계 및 참여자

앞서 소개한 연구 문제들을 해결하기 위하여 종단과 횡단 방법 모두 사용하였다. 먼저 횡단 분석을 위하여 집단간 동질성이 확보된 참여자를 선별하였다. 집단간 동질성을 확보하기 위하여 반편성 과정에서 우열반이 없으며, 한 명의 교사로부터 생물 수업을 받으며, 성별이 여학생인 학교를 선정해야 했다. 여학생 집단을 선정하는 이유는 성별에 따라 효능감과 확신의 차이가 있으며 남학생에서 과잉확신이 더 나타난다는 연구를 근거하였다(Pallier, 2003). 상대적으로 과잉확신의 분포가 적은 여학생의 경우에도 위에서 제시한 가설들이 성립할 경우 남학생을 대상으로도 그 연구 결과를 적용할 수 있기 때문에 여학생을 선정하였다. 이 조건을 만족하는 중학교를 선정하고 6개 반, 204명의 여학생이 이 연구에 참여하였다. 6개 반은 전학년도 성적을 기준으로 학생을 배치하여 반별 성취도의 차이가 없다. 또한 이 연구에서 수행한 광합성 개념 검사에서도 수행 능력의 차이는 없었다( $p > 0.05$ ).

두 번째로 종단 분석을 위하여 6개 반 학생을 두 집단으로 구분하였다. 첫 번째 집단(A 집단)은 선다형을 먼저 수행하고 그 이후에 서답형을 수행하였다. 두 번째 집단(B 집단)은 서답형을 먼저 수행하고 그 이후에 선다형을 수행하였다. 그러므로 연구에 참여한 학생 모두 선다형과 서답형의 두 종류의 검사지를 모두 수행하였다. 두 번에 걸친 평가에서 사용한 검사도구는 앞서 소개한 선다형 문항과 그에 해당하는 효능감과 확신 척도이다. 참여자는 광합성 개념 학습을 수행 한 후 첫 번째 검사를 수행하였으며, 첫 번째 검사와 두 번째 검사의 시간적 간격은 약 6개월이었다. 또한 6개월 동안 광합성에 관한 추가 학습은 없었다.

### 3. 분석 방법

분석방법에 대한 설명은 각 연구문제 별로 제시한다. 첫 번째 연구 문제를 해결하기 위하여 선다형과 서답형 문항과 그에 따른 효능감과 확신의 척도의 상관관계는 Pearson 상관관계를 이용하여 분석하였다. A 집단과 B 집단의 결과를 구분하여 제시하였다. 두 번째 연구 문제를 해결하기 위하여 빈도 조사를 실시하였다. 먼저 부분 점수조차 획득하지 못한 완벽한 오답을 제시하였음에도 그 문항에 대하여 효능감과 확신을 ‘그렇다’와 ‘매우 그렇다’로 응답한 빈도를 오답의 수로 나누어

비율을 조사하였다. 이 값의 수치가 ‘과잉 효능’, ‘과잉 확신’의 비율이다. 그 비율을 A 집단과 B 집단별로 구분하여 조사하였고, 전체 학생을 기준으로도 제시하였다. 세 번째 연구 문제를 해결하기 위하여 효능감과 확신에 대한 대응표본 t-검정을 실시하였다. 마지막 연구 문제를 해결하기 위하여 A 집단과 B 집단을 독립표본 t-검정으로 비교하였다. A 집단은 선다형을 먼저 수행하고 충분한 시간이 경과 후 (6개월 후) 서답형을 실시하였으며, B 집단은 서답형을 먼저 수행하고 그 이후에 선다형을 수행하였다. 약 6개월 후 광합성 지식의 성취도가 낮아지는 만큼 효능감과 확신도 낮아질 것으로 예상되는데 그 감소의 비율의 차이를 중심으로 논의할 것이다. 모든 통계 분석은 PASW 18.0으로 분석하였다.

### III. 연구 결과 및 논의

#### 1. 선다형과 서답형, 그에 따른 효능감과 확신 변인 사이의 상관관계

서론에서 제시한 첫 번째 연구 문제인 선다형과 서답형, 그에 따른 효능감과 확신 변인 사이의 상관관계를 비교하기 위하여 Table 1에 변인간 상관관계 수치를 제시하였다. 상관관계 수치는 선다형을 먼저 하고 서답형을 후에 한 집단(A 집단), 서답형을 먼저하고 선다형을 후에 한 집단(B 집단)으로 구분하여 제시하였다. 선다형과 서답형 평가 결과의 상관관계를 먼저 보면 A 집단에서 0.417( $p < 0.01$ ), B 집단에서 0.492( $p < 0.01$ )로 채 0.5가 되지 않는다. 그러므로 두 평가 방법에 의하여 공통적으로 설명되는 부분은 약 25% 정도이며 70% 이상은 다른 요소로 설명될 수 있다는 것이다. 서론에서 소개한 Ha et al.(2012)과 Nehm & Schonfeld (2008)의 연구에서도 상관관계가 각각 0.52 ( $p < 0.01$ ), 0.58( $p < 0.01$ ) 정도로 나타났다. 이들 연구에서 사용한 진화 개념 검사도구는 학생들의 오개념을 바탕으로 제작되어 문장 형태로 이루어진 개념 검사지 형태의 선다형과 글로서 설명을 해야 되는 형태의 서술형 평가이다. 그에 반하여, 이 연구에서 사용한 검사 도구는 선행연구에서 사용한 도구에 비하여 그 형태가 상대적으로 단순한 지식 검사 도구이며, 서답형 형태로 단답형으로 제시하는 형태이다. 이와 같은 점을 근거로 생각하였을 때, 이 연구에서 서답형 문항을 수행 할 때 설명으로 풀어서 작성해야 되는 글쓰기 능력은 크게 요구되지 않았을 수 있다. 이와 같은 관점에서 보았을 때, 선다형과 서답형 평가가 공통적으로 설명하지 못하는 약 75%의 능력 중의 상당 부분은 지식의 ‘회상’ 능력과 밀접한 관계를 맺고 있을 것으로 생각해 볼 수 있다. 학생들이 선다형 문항을 해결할 때 문항에 제시된 정답과 오답을 보고 자신이 가지고 있는 개념과 비교하는 과정을 통하여 해결하는 것에 비하여, 서답형에서는 정답과 오답에 대한 아무런 정보 없이 자신의 기억에서 몇 가지 대안을 찾아낸 후 그 중에서 정답으로 생각되는 것을 선택하는 형태로 문제 풀이가 이루어진다. 선다형과 서답형의 주요한 차이점은 답지가 평가자에 의하여 제시되는 것인지, 피평가자 스스로 답지를 자신의 기억 속에서 만들어내는 것인지 일 것이다. Karpicke & Blunt (2011)는 지식을 회상하는 훈련을 통하여 학습을

촉진할 수 있다고 하였다. 하지만 앞서 논의한 바와 같이 우리나라 학생들이 경험하는 대부분의 평가 방법은 선다형이기 때문에 지식을 회상하는 경험이 많이 부족할 수 있다. 학생들은 서답형 평가에서 낮은 성취도를 보이게 되는 이유도 여기에 근거할 수 있다(Choi et al., 2011).

Table 1의 결과에서 두 번째로 논의할 것은 선다형과 서답형에 대한 평가 결과에 대한 상관관계 보다 선다형 평가와 서답형 평가에 대한 효능감과 확신의 상관관계가 더 높다는 것이다. 선다형과 서답형 평가에 대한 효능감과 확신의 상관관계는 A 집단이 각각 0.600( $p < 0.01$ ), 0.685( $p < 0.01$ )이며, B 집단은 각각 0.624( $p < 0.01$ ), 0.611( $p < 0.01$ )이었다. 앞서 살펴본 선다형과 서답형의 성취도 상관관계에 비하여 상당히 높은 수준이다. 이 결과는 서론에서 제시한 가설을 지지하는데, 지식의 수준은 시간적 차이와 같은 외부 조건에 의하여 쉽게 달라질 수 있지만 효능감과 확신은 상대적으로 외부 조건에 의하여 변하지 않는 특성이 있다. 이와 같은 현상은 잠재적으로 문제가 될 수 있는데, 쉬운 형태의 문제를 접하는 과정에서 생긴 높은 수준의 효능감과 확신이 차후에 어려운 문제를 접할 때 그대로 유지될 경우 반성적인 문제 해결을 못하고 과잉효능감과 과잉확신으로 나타날 수 있다는 것이다. 특히 우리나라 학생들은 오랫동안 선다형에 노출되어 왔고 그로 인하여 선다형에 대한 높은 효능감과 확신이 있다. 하지만 상대적으로 어려운 서답형에서 그런 높은 수준의 효능감과 확신은 오히려 문제 해결을 그르칠 수 있다. 이 논의는 Table 2와 Table 4에서 다시 이루어 질 것이다.

세 번째로 Table 1의 결과에서 논의하고자 하는 것은 개념측정도구로 평가한 결과를 자신이 얼마나 정확하게 예측하는가 하는 점이다. 그 결과를 보면 A 집단에서 선다형 평가 결과에 대한 자기 효능감과 확신의 상관관계는 각각 0.297( $p < 0.01$ ), 0.378( $p < 0.01$ ), 서답형 평가 결과에 대한 자기 효능감과 확신의 상관관계는 각각 0.394( $p < 0.01$ ), 0.417( $p < 0.01$ )로 나타났다. 마찬가지로 B 집단에서 선다형 평가 결과에 대한 자기 효능감과 확신의 상관관계는 각각 0.372( $p < 0.01$ ), 0.378( $p < 0.01$ ), 서답형 평가에서는 0.414( $p < 0.01$ ), 0.487( $p < 0.01$ )로 나타났다. 이 정도의 상관관계에서 설명력은 30%에도 채 미치지 못한다. 자신의 능력에 대한 객관적 평가 능력은 상당히 낮다고 할 수 있다. 자신의 능력에 대한 정확한 판단은 성공적인 자기 조절 학습을 위하여 매우 중요한 요소이다\*(de Bruin & van Gog, 2012; Pajares & Miller, 1997). 그러므로 문제 해결 능력에 대한 효능감과 지식에 대한 확신의 정확도를 높이는 프로그램을 개발하여 학생들에게 제공할 필요성이 있을 것이다. Dunlosky & Rawson(2012)는 과잉확신을 줄이기 위한 방법에 관한 연구를 수행하면서 두 집단을 비교하였는데, 실험 집단은 참여자의 초기 답지와 정답을 동시에 보여주며 자신의 응답을 다시 평가하게 하였고 통제집단은 참여자의 초기답지를 다시 보여주고 다시 평가하게 하였다. 이 연구 결과에서 보여주는 것은 자신의 답지를 정답과 비교하는 반성적 모니터링 훈련을 통하여 과잉확신을 줄일 수 있다는 것이다. 이런 방법은 교육현장에서 쉽게 적용 할 수 있는데, 정답을 가리고 문제를 풀고, 자신의 정답에 대한 확신을 점검한 후 바로 가려 놓았던 정답을 보는 훈련도 과잉확신을 줄일 수 있다고 한다(Lipko et al., 2009). 학생들은 정답을 확인하고 바로 수용하기보다 자신의 오답과 비교하는 활동을 통하여 Koriat et al.(1980)이 강조한 부정적 증거에 대한 인식을 높여야 할 것이다.

\* 이 주제는 학습과 교수 방법 전문 학술지인 Learning and Instruction의 2012년 특별호 주제였다. 해당호를 확인하면 다양한 정보를 확인할 수 있다.

Table 1. The Pearson correlation coefficients among variables(\*\*p &lt; 0.01)

	(1)	(2)	(3)	(4)	(5)	(6)
Group A (n = 100)	Multiple-choice test score (1)	1.000				
	MCT efficacy (2)	0.297**	1.000			
	MCT confidence (3)	0.378**	0.928**	1.000		
	Supply-type test score (4)	0.417**	0.369**	0.423**	1.000	
	STT efficacy (5)	0.330**	0.600**	0.604**	0.394**	1.000
	STT confidence (6)	0.297**	0.651**	0.685**	0.417**	0.933**
Group B (n = 101)	Multiple-choice test score (1)	1.000				
	MCT efficacy (2)	0.372**	1.000			
	MCT confidence (3)	0.378**	0.905**	1.000		
	Supply-type test score (4)	0.492**	0.386**	0.350**	1.000	
	STT efficacy (5)	0.366**	0.624**	0.595**	0.414**	1.000
	STT confidence (6)	0.429**	0.622**	0.611**	0.487**	0.964**

## 2. 선다형과 서답형에서 나타나는 과잉효능, 과잉확신의 비율 차이

Table 2는 학생들이 오답을 선택했음에도 불구하고 자신이 정답을 선택할 수 있을 것이라는 효능감과 자신이 생각한 답지가 맞을 것 같다는 확신에 대한 척도에 '④ 자신 있다', '⑤ 많이 자신 있다', 또는 '④ 확신 한다', '⑤ 많이 확신 한다'에 선택한 비율을 보여준다. Table 2를 보면 선다형에서 더 많은 과잉효능감과 과잉확신이 나타난다. 이 결과는 앞서 Koriat *et al.*(1980)과 Oskamp (1965)의 논의를 통해 제시된 가설을 지지한다. 선다형의 경우 서답형에 비하여 정답과 오답에 대한 더 많은 정보가 제공되며 그런 정보를 통하여 학생들은 자신이 생각한 답지에 대한 긍정적인 증거를 수집하게 된다. 이런 과정에서 학생들은 자신의 답지가 맞을 것이라는 믿음을 강화하게 되는 것이다. 반면 서답형의 경우에는 그런 정보가 없기 때문에 높은 효능감과 확신이 생길 수 없는 것이다.

## 3. 선다형과 서답형에서의 효능감과 확신의 차이

Table 3에서는 선다형과 서답형에서 학생들이 응답한 효능감과 확신을 비교하였다. 효능감은 주어진 문제를 간략하게 살펴보고 그것을 해결 할 수 있을지 판단하는 것이며, 확신은 문제 해결 후 자신이

선택한 답지가 정확한지에 대한 스스로의 판단이다. 두 척도는 실제로 아주 높은 상관관계를 보이지만 ( $r > 0.9$ ), 그 수준의 차이는 약간 다르다. Koriat *et al.*(1980)의 설명에 근거로 살펴보면, 일반적으로 피평가자들은 선다형 문항을 풀이하는 과정에서 임시적으로 선택한 답지를 지지하는 증거들을 찾게 되며, 지지하지 않는 증거들은 회피하게 된다. 그런 과정에서 자신이 생각하는 답지에 대한 믿음이 더 강화될 수 있다. 이 가설은 Table 3에서 제시된 결과로 확인할 수 있는데, 선다형 문항에서는 효능감에 비하여 확신의 수준이 더 높은 것을 확인할 수 있는데 반해 서답형의 경우 그렇지 않다. 서답형 평가는 정답과 오답에 대한 정보가 주어지지 않기 때문에 자신이 생각한 답지를 지지하는 증거를 찾을 수 없다. 그러므로 효능감과 확신은 차이가 없을 것이다. 이 가설은 Table 3에 나타난 결과들이 지지하는데 두 집단 모두 서답형에서는 효능감과 확신의 차이가 없었다.

## 4. 선다형과 서술형 평가의 순서에 따른 과잉효능, 과잉확신의 차이

마지막으로 선택형 평가와 서술형 평가의 순서에 따른 학생들의 효능감과 확신의 수준 차이를 확인하였다. 앞서 서론에서 밝힌 바와 같이 이 분석의 연구 가설은 선다형을 먼저 수행한 집단일 경우 높은 수준의 효능감과 확신을 경험하고 차후에 수행하는 서답형에도 상대적으로 높은 수준의 효능감과 확신을 보일 수 있다는 것이다. 이 연구

Table 2. The ratios of over-eficacy (OE) and over-confidence (OC)

		Group A (n = 100 * 8 items)			Group B (n = 104 * 8 items)			Total (n = 204 * 8 items)		
		# of wrong answers	# of OE/OC	Ratio	# of wrong answers	# of OE/OC	Ratio	# of wrong answers	# of OE/OC	Ratio
MCT	OE	193	49	25.4%	263	40	15.2%	456	89	19.5%
	OC		70	36.3%		69	26.2%		139	30.5%
STT	OE	375	24	6.4%	296	13	4.4%	671	37	5.5%
	OC		26	6.9%		9	3.0%		35	5.2%

Table 3. Paired sample t-tests between efficacy and confidence

			M	SD	t	sig.	Cohen's d
Group A	MCT (n = 99)	Efficacy	25.19	7.66	-2.89	0.005	0.11
		Confidence	26.04	7.38			
	STT (n = 100)	Efficacy	16.71	7.50	-0.93	0.354	0.03
		Confidence	16.96	7.82			
Group B	MCT (n = 103)	Efficacy	22.48	8.21	-4.32	0.000	0.17
		Confidence	23.85	8.11			
	STT (n = 101)	Efficacy	18.02	7.09	0.44	0.659	-0.01
		Confidence	17.93	7.12			

가설을 확인하기 위하여 앞서 논의한 Table 2(과잉효능감과 과잉확신의 비율)로 다시 돌아가서 살펴보자. Table 2에서는 A 집단(선다형을 먼저 수행한 집단)과 B 집단(서답형을 먼저 수행한 집단)의 과잉 효능감과 과잉 확신이 제시되어 있다. 서답형을 먼저 한 집단(B 집단)이 선다형을 먼저 한 집단(A 집단)에 비하여 과잉효능감의 비율은 59.8%(15.2%/25.4%) 수준, 과잉확신의 비율은 72.2%(26.2%/36.3%) 수준이다. 하지만 실제로 선다형 평가와 서답형 평가에서 감소한 지식의 수준은 이 정도보다 높다. Table 4를 보면 서술형을 먼저 한 집단이 선다형을 먼저 한 집단에 비하여 선다형 점수의 비율은 89.1%(9.53점/10.70점)로 효능감과 확신의 수준의 감소에 비하면 상당히 많은 지식이 유지되어 있었다(Table 2 참조).

반면에, 선다형을 먼저 수행한 집단에서 서답형 점수는 Table 4에 제시된 바와 같이 73.8%(6.18점/8.37점) 수준이지만 과잉효능감과 과잉확신은 Table 2에 제시된 바와 같이 각각 145.5%(6.4%/4.4%)와 230.0%(6.9%/3.0%)로 지식의 감소에 비하여 더 높은 수준으로 유지되어 있었다. 이 결과는 Table 1(선다형과 서답형, 그리고 그에 따른 효능감과 확신 변인 사이의 상관관계)에서 제시한 논의와 일맥상통하는데, 선다형을 먼저 수행할 경우 과잉효능감과 과잉확신을 경험하게 되고, 그런 효능감과 확신이 6개월 뒤에 동형의 서답형 평가를 수행할 때도 남아있게 될 수 있다. 반대로, 서답형을 먼저 수행할 경우 낮은 수준의 과잉효능감과 과잉확신을 경험하고 6개월 뒤 선다형을 수행하게 되더라도 여전히 상대적으로 낮은 수준의 과잉효능감과 과잉확신을 유지하는 것이다. 이 결과는 Oskamp(1965)의 논의를 통해서도 이해할 수 있는데, 6개월 전 선다형 평가를 받은 집단과 서답형 평가를 받은 집단의 정보의 양의 차이는 선다형 평가를 받은 집단이 훨씬 더 많을 것이다. 그러므로 선다형에 많이 노출되었을 경우 자신이 많은 것을 안다고 느껴 문제 해결에 대한 과잉효능감과 과잉확신이 더 많이 나타나는 것이다.

다시 Table 4로 돌아와서 논의해보자. Table 4는 선다형과 서답형을 사전과 사후에 각각 실시한 두 집단에 대한 비교이다. 앞서 연구 방법에서 제시한 바와 같이 두 집단은 균등한 반환성, 동일한 생물교사, 전부 여학생으로 구성되어 성별에 대해서도 동질의 집단으로 이해할 수 있다. 또한 각 집단은 3개 반씩을 포함하는데 각 집단 내에서 반별 성취도 차이는 유의미하지 않았다( $p > 0.05$ ). 그러므로 이

두 집단을 비교함으로써 사전-사후 효과를 비교할 수 있는 것이다.

앞서 설명한 바와 같이 선다형 평가에서 A 집단과 B 집단은 6개월의 시간적 차이가 존재한다. 반대로 서답형 평가에서도 B 집단과 A 집단 역시 6개월의 시간적 차이가 존재한다. 두 집단의 차이는 어떤 평가를 먼저 했는가 하는 것이다. 평균값을 비교하면 6개월의 시간적 차이 동안 선다형, 서답형 평가 결과 모두 낮아지게 된다. 또한 그에 따른 효능감과 확신 역시 낮아지게 된다. 하지만 통계 결과를 확인하면 서답형의 경우 6개월 동안 효능감과 확신이 유의미하게 낮아지지 않았지만(유의도 각각 0.232과 0.358), 선다형에서는 6개월 동안 효능감과 확신이 유의미하게 낮았다(유의도 각각 0.016과 0.036). 이 연구 결과 역시 앞서 Table 2에서 논의한 바와 일맥상통하는데, 선다형을 먼저 수행할 경우 높은 수준의 자기 효능감과 확신을 경험하였고 그것이 유지되어 서답형을 먼저 한 집단에 비교하여 더 높은 수준의 효능감과 확신을 유지하는 것으로 유추할 수 있다.

#### IV. 교육적 함의 및 제언

학습자 스스로 자신의 지식수준을 반성적으로 점검하고 효율적인 학습방법을 찾을 수 있는 능력은 자기 조절 학습을 위하여 중요하다. 최근 인지 심리학자들의 연구에 의하면 자신의 문제 해결 능력을 과장되게 이해하는 과잉효능감, 과잉확신은 성취도 수준을 낮추고 올바른 판단을 흐리게 하는 것과 같은 문제점을 야기한다고 알려졌다. 이 연구는 문제 해결에 대한 효능감과 확신, 과잉효능감과 과잉확신에 대한 4가지 연구 문제를 설정하고, 실험적인 방법으로 연구 문제를 확인하였다.

연구결과 선다형과 서답형의 상관관계는 0.5를 넘지 않았으며, 학생들이 생각하는 자신의 문제 해결 능력과 실제 능력 간의 상관관계 역시 0.45가 되지 않았다. 선다형 문제 해결 과정에서 높은 효능감과 지식의 확신을 보였다. 또한 선다형에서 문제 해결에 실패하였음에도 불구하고 문제 해결을 성공적으로 했을 것이라는 과잉효능, 자신이 선택한 답지가 맞다고 느끼는 과잉확신의 비율이 훨씬 높았다. 더욱이 선다형 문항을 먼저 수행한 집단에서는 그렇지 않은 집단에 비하여 과잉효능감과 과잉확신의 비율이 더 높게 나타났다.

Table 4. Independent t-tests of MCT and STT scores, efficacy, and confidence between Group A and Group B

	Group	N	M	SD	t	sig.	Cohen's d
MCT score	A	100	10.70	3.32	2.39	0.018	0.33
	B	104	9.53	3.68			
MCT efficacy	A	99	25.19	7.66	2.43	0.016	0.34
	B	103	22.48	8.21			
MCT confidence	A	100	26.12	7.39	2.11	0.036	0.30
	B	104	23.83	8.07			
STT score	B	104	8.37	3.48	4.90	0.000	0.69
	A	100	6.18	2.84			
STT efficacy	B	103	17.94	7.09	1.20	0.232	0.17
	A	100	16.71	7.50			
STT confidence	B	101	17.93	7.12	0.92	0.358	0.13
	A	100	16.96	7.82			

이 연구의 결과들을 통하여 얻을 수 있는 결론은 다음과 같다. 첫 번째로 과학교육평가에서 선다형과 서답형 중 되도록 서답형을 사용해야 한다는 것이다. 실제로 이런 움직임은 과학교육 현장에서 이루어지고 있다. 이 연구에서 사용한 검사도구와 같이 단답형일지라도 학생들의 서답형 평가에서 나타나는 성취도 수준은 선다형에 비하여 낮다(Choi *et al.*, 2011). 또한 선다형 평가로 학생들의 지식 회상 능력, 설명 구성 능력 등 서답형이나 설명형 평가에 비하여 확인할 수 없는 요인들이 많다. 하지만 최근의 컴퓨터 기술의 발달로 서답형 평가에 대한 컴퓨터 채점이 가능해지고 있다(Ha *et al.*, 2011; Kim *et al.*, 2013; Nehm *et al.*, 2012). 서답형 평가를 사용할 경우 학생들의 다양한 능력을 측정할 수 있을 뿐만 아니라, 학생들 역시 평가 과정에서 지식을 회상해 보는 경험을 하여, 장기 기억을 촉진하고, 학업 성취도를 높이는 결과를 만들어 낼 수 있다.

두 번째로 생각해 볼 문제는 여러 가지 현실적인 이유로 선다형 문항을 사용해야만 한다면 언제 사용해야 하는가에 대한 문제이다. 이 연구 결과에서 보여주듯이 선다형 문항을 먼저 수행할 경우 높은 수준의 효능감과 확신을 경험하고, 많은 정보를 획득하게 되어 사후에 서답형 문항을 접하였을 때에도 높은 수준의 자신감과 지니친 확신이 유지될 수 있다. 이렇게 잘못된 지식을 가지고 있음에도 그것이 옳다고 믿는 과잉효능감과 과잉확신이 장기적으로 성취도 발달의 악영향을 미치고, 학습을 소홀히 하게 되는 등의 부작용을 양산할 수 있기 때문에 선다형 문항에 대한 노출은 최소화하는 것이 좋으며 만약 동일 개념에 대하여 선다형과 서답형을 같이 수행해야 된다면 선다형을 사후에 수행하는 것이 좋을 것이다. 특히, 중간고사와 기말고사와 같이 학생들의 능력치를 객관적으로 비교해야 되는 평가와 달리 그럴 필요가 없는 진단 평가, 형성 평가, 자율 학습에서의 문항들은 가급적이면 선다형보다 서답형을 사용할 필요가 있다. 또한 서답형 평가가 더 많은 인지 능력을 요구하여 학생들의 자신감을 감퇴시킬 수 있다는 의견이 있을 수도 있는데, 진단평가나 형성평가와 같은 평가들은 평가 결과가 학업성취도에 반영되지 않으므로 평가에 대한 스트레스가 생성되지 않을 수 있어 학생들의 정신 건강적인 측면에서도 나쁘지 않을 것이다. 더욱이 지속적으로 경험하였을 때 자신감 역시 향상된다는 측면에서도 긍정적으로 판단된다.

세 번째로 효능감과 확신은 지식의 유지에 비하여 상대적으로 안정적으로 유지될 가능성이 있다는 것이다. 이와 같이 외부 환경에 따라 잘 변하지 않는 심리적 요인들은 외부 자극(예를 들어 학습) 주더라도 변화되지 않을 가능성이 높다. 그러므로 자신의 능력을 정확하게 판단하는 능력을 함양시키기 위해서는 많은 노력들이 수반되어야 할 것이다. 앞서 Dunlosky & Rawson(2012)의 연구에서 확인하였듯이, 학생들이 자신의 능력을 정확하게 판단하게 하는 교육방법은 의외로 간단할 수 있다. 예를 들어서, 교사는 학생들이 선택한 답지에 대하여 그 이유를 적게 하거나, 그 답지가 정답이 아닐 가능성에 대하여 학생들에게 질문할 수 있을 것이다. 사후 학생들이 자신이 생각한 답지와 정답을 비교하면서 왜 그런 실수를 하였는지 그 답지가 맞다고 생각했는지 점검하게 함으로써 정확한 자기 판단 능력을 향상시킬 수 있는 것이다. 또한 학생들 스스로도 자신의 답지와 정답을 비교하여 생각하는 훈련을 하거나, 항상 자신의 답지가 틀렸을 가능성에 대하여 인정하는 노력을 한다면 정확한 자기 판단 능력은 향상될 수 있다(Lipko *et al.*, 2009). 또한 이런 형태의 방법을 쉽게 할 수 있도록 웹 기반 평가 도구를 개발할 수도 있을 것이다.

마지막으로 이 연구 결과의 해석에 대한 몇 가지 제한점과 향후 연구에 대한 제언을 하고자 한다. 앞서 서론에서 논의한 바와 같이 과잉효능감과 과잉확신에는 성차가 존재한다. 이 연구는 선행연구에 근거하여 과잉효능감과 과잉확신이 상대적으로 낮다고 알려진 여학생을 대상으로 하여 연구 결과의 보편타당성을 높이고자 하였다. 하지만 우리나라 학생들의 과잉효능감과 과잉확신에 대한 성차 연구는 향후 반드시 필요할 것으로 판단된다. 두 번째, 이 연구에서는 중학생들을 대상으로 실시하였다. 자신의 능력을 판단하는 초인지적인 능력에는 여러 가지 지적, 정의적 능력들이 포함되어 있을 것이고 그런 능력들은 유아기에서 성인에 걸쳐 지속적으로 변화될 것이다. 그러므로 이 연구에서 확인한 결과들을 직접적으로 저학년과 고학년에 적용해서는 안 될 것이며, 그런 점을 확인할 수 있는 횡단 연구가 이루어져야 할 것이다. 또한, 문제의 형태 역시 하나의 변인이 될 수 있다. 이 연구에서 사용한 광합성 검사도구는 단순한 지식을 단답형으로 질문하는 것이다(부록 참조). 하지만 보다 복잡한 형태의 설명형 검사도구일 경우 또 다른 변인이 존재할 수 있다. 문제의 복잡성과 학생들



의 과잉효능감과 과잉확신 역시 향후 연구될 필요성이 있다. 학생들의 자아 효능감과 확신에 영향을 주는 다양한 개인적인 성향을 이 연구에서는 확인하지 않았지만 향후 연구는 이 부분에 대한 연구가 필요할 것으로 판단된다. 마지막으로 면담 등과 같은 다른 방법론을 활용하여 과잉효능감과 과잉확신의 메커니즘을 확인하는 연구 역시 필요할 것으로 판단된다.

## 국문요약

이 연구는 과잉효능감과 과잉확신이 선다형과 서답형 평가에서 어떻게 나타나는지에 조사하였다. 이 연구를 위하여 204명의 여자 중학교 학생들이 참여하였다. 선다형과 서답형으로 된 광합성 개념에 대한 검사도구를 사용하였으며, 각 문항마다 자신의 효능감과 확신을 척도로 나타내게 하였다. 연구결과 선다형과 서답형의 상관관계는 0.5가 채 되지 않았으며, 학생들이 생각하는 자신의 문제 해결 능력(효능감)과 실제 능력 간의 상관관계 역시 0.45가 되지 않았다. 학생들은 선다형 문제 해결 과정에서 지식에 대한 확신을 높이는 경향을 보였다. 선다형 문항을 먼저 수행한 집단에서는 그렇지 않은 집단에 비하여 과잉효능감과 과잉확신의 비율이 더 높게 나타났다. 이 연구 결과는 학생들을 평가하는 과정에서 서답형의 사용을 늘릴 것을 제안한다. 또한 두 형태의 검사도구를 동시에 사용하게 될 경우 선다형 검사도구를 사후에 사용하여 학생들이 적절한 수준의 효능감과 확신을 유지할 수 있도록 해야 될 것이다.

## 감사의 글

마지막까지 논문의 완성도를 높여주신 익명의 심사자님들과 긴 기간 동안 자료 수집에 협조해 주신 주미현 선생님에게 감사의 말씀을 전합니다.

## References

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Beggrow, E. P., Ha, M., Nehm, R. H., Pearl, D., & Boone, W. J. (in press). Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance? *Journal of Science Education and Technology*.
- Bruttomesso, D., Gagnayre, R., Leclercq, D., Crazzolaro, D., Busata, E., d'Ivernois, J. F., Casiglia, E., & Baritussio, A. (2003). The use of degrees of certainty to evaluate knowledge. *Patient Education and Counseling*, 51(1), 29-37.
- Choi, K., Lee, H., Shin, N., Kim, S. W., & Krajcik, J. (2011). Reconceptualization of scientific literacy in South Korea for the 21st century. *Journal of Research in Science Teaching*, 48(6), 670-697.
- de Bruin, A. B., & van Gog, T. (2012). Improving self-monitoring and self-regulation: From cognitive Psychology to the classroom. *Learning and Instruction*, 22(4), 245-252.
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271-280.
- Ha, M., Haurly, L. D., Nehm, R. H. (2012). Feeling of certainty: Uncovering a missing link between knowledge and acceptance of evolution. *Journal of Research in Science Teaching*, 49(1), 95-121.
- Ha, M., Nehm, R. H., Urban-Lurain, M. & Merrill, J. E. (2011). Applying computerized scoring models of written biological explanations across courses and colleges: Prospects and limitations. *CBE-Life Science Education*, 10, 379-393.
- Hong, M. Y., Park, C., & Kim, S. (2001). An analysis of science achievement of the third international mathematics and science study-repeated(TIMSS-R). *Journal of the Korean Association for Research in Science Education*, 21(2), 328-341.(in Korean)
- Johnson, D. D., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, 477(7364), 317-320.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772-775.
- Kim, M, H., Noh, E. H., & Sim, J. H. (2013). The application of automatic scoring program to supply-type items of basic competency test. *The Journal of Curriculum and Evaluation*, 16(1), 137-160.(in Korean)
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 107-118.
- Langendyk, V. (2006). Not knowing that they do not know: self-assessment accuracy of third-year medical students. *Medical education*, 40(2), 173-179.
- Lee, C. Y. (2011). Objective Test Allowing Multiple Choice for the Correct Answer with Penalty for Incorrect Choices. *School Science Journal*, 5(1), 17-26. (in Korean)
- Lee, K. S. (2008). The dominance of objective examinations and the state intervention in Korea. *Philosophy of Education*, 34, 245-275.(in Korean)
- Lipko, A. R., Dunlosky, J., Hartwig, M. K., Rawson, K. A., Swan, K., & Cook, D. (2009). Using standards to improve middle-school students' accuracy at evaluating the quality of their recall. *Journal of Experimental Psychology: Applied*, 15, 307-318.
- Luszczynska, A., Gutiérrez-Doña, B., & Schwarzer, R. (2005). General self-efficacy in various domains of human functioning: Evidence from five countries. *International Journal of Psychology*, 40(2), 80-89.
- Marmaroti, P., & Galanopoulou, D. (2006). Pupils' understanding of photosynthesis: a questionnaire for the simultaneous assessment of all aspects. *International Journal of Science Education*, 28(4), 383-403.
- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, 45(10), 1131-1160.
- Nehm, R. H., Ha, M., & Mayfield, E. (2012). Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations. *Journal of Science Education and Technology*, 21(1), 183-196.
- Nehm, R. H., Ha, M., & Mayfield, E. (2012). Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations. *Journal of Science Education and Technology*, 21(1), 183-196.
- Nowell, C., & Alston, R. M. (2007). I thought I got an A! Overconfidence across the economics curriculum. *The Journal of Economic Education*, 38(2), 131-142.
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744-777.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of consulting psychology*, 29(3), 261-265.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66, 543-578.
- Pajares, F., & Miller, M. D. (1997). Mathematics self-efficacy and mathematical problem solving: Implications of using different forms of assessment. *The Journal of Experimental Education*, 65(3), 213-228.
- Pallier, G. (2003). Gender differences in the self-assessment of accuracy



on cognitive tasks. *Sex Roles*, 48(5-6), 265-276.  
 Park, C., & Hong, M. Y. (2002). A relative effectiveness of item types for estimating science ability in TIMSS-R. *Journal of the Korean Association for Research in Science Education*, 22(1), 122-131.(in Korean)  
 Park, J. (2010). Constructive multiple choice testing system. *British Journal of Educational Technology*, 41(6), 1054-1064.  
 Reach, G., Zerrouki, A., Leclercq, D., & d'Ivernois, J. F. (2005). Adjusting insulin doses: From knowledge to decision. *Patient Education and*

*Counseling*, 56(1), 98-103.  
 Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26, 207-231.  
 Yates, J. F., Lee, J. W., & Bush, J. G. (1997). General knowledge overconfidence: cross-national variations, response style, and "reality". *Organizational behavior and human decision processes*, 70(2), 87-94.  
 Zakay, D., & Glicksohn, J. (1992). Overconfidence in a multiple-choice test and its relationship to achievement. *The Psychological Record*, 42(4), 519-524.

[부 록]

※ 선다형 문항 예시

\* 아래 문제를 읽은 후 응답하세요. “아래 질문을 잘 풀 수 있는지” 자신의 느낌을 선택해주세요  
 ① 자신 없다 ② 아주 약간 자신 있다 ③ 약간 자신 있다  
 ④ 자신 있다 ⑤ 많이 자신 있다

\* 아래 문제를 풀어주세요

1. 식물의 어느 부분에서 광합성이 진행되나요? (            ) ① 잎 ② 녹색인 모든 부분 ③ 식물 전체 ④ 뿌리
---

\* 문제를 푼 후에 느낌을 적어주세요 내가 적은 정답에 어느 정도 확신하나요?  
 ① 확신 없다 ② 아주 약간 확신 한다 ③ 약간 확신 한다  
 ④ 확신 한다 ⑤ 많이 확신 한다

※ 서답형 문항 예시

\* 아래 문제를 읽은 후 응답하세요. “아래 질문을 잘 풀 수 있는지” 자신의 느낌을 선택해주세요  
 ① 자신 없다 ② 아주 약간 자신 있다 ③ 약간 자신 있다  
 ④ 자신 있다 ⑤ 많이 자신 있다

\* 아래 문제를 풀어주세요

1. 식물의 어느 부분에서 광합성이 진행되나요?
----------------------------

\* 문제를 푼 후에 느낌을 적어주세요 내가 적은 정답에 어느 정도 확신하나요?  
 ① 확신 없다 ② 아주 약간 확신 한다 ③ 약간 확신 한다  
 ④ 확신 한다 ⑤ 많이 확신 한다