생물 예비 교사의 선택형과 개방형 문항에서 나타난 자연선택 설명 분석

하민수 · 이준기[1]*
오하이오 주립대학교 · [1]전북대학교

The Analysis of Pre-Service Biology Teachers' Natural Selection Conceptions in Multiple-Choice and Open-Response Instruments

Ha, Minsu · Lee, Jun-Ki[1]*
The Ohio State University · [1]Chonbuk National University

Abstract: Teachers use explanations to communicate important scientific ideas to students. Consequently, all biology teachers should be evaluated to determine how effective they are at constructing and communicating biological explanations. Open response questions are required to detect pre-service biology teachers' abilities to communicate robust and accurate scientific explanations. Nevertheless, multiple-choice questions are typically preferred by educators because of the common drawbacks of using open-response instruments, such as scoring time, inter-rater scoring disagreements, and delayed feedback to test takers. This study aims to measure pre-service biology teachers' competence in building scientific explanations and to investigate how accurately multiple-choice questions predict the results of open-response questions. One hundred twenty four pre-service biology teachers participated in the study and were administered 20 multiple-choice items and three open-response items designed to measure the accuracy and quality of their explanations of evolutionary change. The results demonstrated that pre-service teachers displayed higher competence when tested with multiple choice items than when tested with open response items. Moreover, scores derived from multiple-choice items poorly predicted the scores derived from open-response items. Multiple-choice items were also found to be poor measures of the consistency, purity and abundance of conceptual elements in teachers' evolutionary explanations. Additionally, many teachers held mixed-models composed of both scientific and naïve ideas, which were difficult to detect using multiple-choice formats. Overall, the study indicates that multiple-choice formats are poorly suited to measuring several aspects of biology teachers' knowledge of evolution, including their ability to generate scientific explanations. This study suggests that open-response items should be used in teacher education programs to assess pre-service teachers' explanatory competency prior to being permitted to teach science to children.

Key words: pre-service biology teachers, natural selection, multiple-choice, open-response, scientific explanation

Ⅰ. 서 론

교육의 중요한 목적은 학생들이 배운 지식을 활용하여 핵심 아이디어를 구성하도록 하는데 있다(Dochy, 2001). 그러므로 평가 방법은 지식의 활성화 능력, 문제 해결 능력, 논리적인 설명 구성 능력 등을 평가할 수 있어야 된다. 새로운 지식 평가에 대한 그의 기준은 첫째, 학생들이 상황 의존적 문제에서 다양한 관점으로 문제를 바라볼 수 있어야 하며, 둘째, 지식을 실제 상황에 적용하여 설명할 수 있어야 하며, 셋째, 자신이 현재 가지고 있는 지식을 전이시킬 수 있어야 한다. Vosniadou(1994)는 개념 변화 연구에서 평가의 방법의 선택에 대한 중요성을 역설하면서, 단편적 지식을 묻는 문항과 피평가자의 상황에 의존하여 재구성한 질문 중 후자를 선택할 것을 강조하였다. 예를 들어, '지구의 모양은?' 과 같은 질문에 대해서 학생들은 쉽게 '둥근 모양' 이라고 대답하는 반면, '한쪽 방향으로 계속 걸어갔을 때 끝은 어디일까?' 라는 질문에 학생들은 다른 응답을 보인다. 학생들이 기억을 통하여 단순히 생산해 낸 답안은 그 학생의 개념

*교신저자: 이준기(junki@jbnu.ac.kr)
**2011.03.31(접수) 2011.07.14(1심통과) 2011.08.12(2심통과) 2011.08.23(3심통과) 2011.09.27(최종통과)

구조를 보여주기 힘들다. 답안을 구성하는 것 자체는 그 학생의 개념 구조가 아니라는 것이다. Gitomer, Duschl(2007) 역시 과학 영역에서 평가는 구조화된 핵심 지식들의 조합으로 평가할 것을 강조하였다.

앞서 논의한 학자들의 주장은 단편적 지식의 유무를 확인하는 것보다 학생들이 자신의 지식으로 설명을 구성하는 평가 방식이 학생들의 학습 발달이나 교수 학습 방법을 위한 피드백에 있어서 모두 월등한 장점을 보일 수 있다는 것이다. 하지만 최근 연구는 객관식 또는 선택형 시험에 비하여 논리적이고 구조적인 설명을 생성해 내는 능력, 또는 자신의 지식을 활성화 하여 과학적 설명을 제시하는 개방형 문항에 약한 면모를 보여 왔다. Choi *et al.*(2011)은 TIMSS (Trends in International Mathematics and Science Study)와 PISA(Program for International Student Assessment) 평과 결과를 비교하면서 우리나라 학생들이 객관식에 비하여 문제 해결을 강조한 주관식 평가에서 약한 면모를 보인다고 하였다. 박정 등(2000)의 연구에서도 우리나라 학생들이 내용 지식에 비하여 답안을 구성하는 능력이 부족하다고 논의하였다. Park(2010)의 연구에서도 초등학생들의 서술형 문항의 점수가 선다형 문항에서의 점수에 비하여 약 70%의 선에 그치는 것으로 조사되었다. 박정, 홍미영(2002)은 문항 유형에 따른 과학 능력 추정의 효율성 비교 연구에서 수행형 문항이 선다형 문항에 비하여 보다 효율적으로 측정하고 있음을 논의하였다. 그 이유는 앞서 논의한 바와 같이 상황 의존적인 개방형 문항에서 학생들이 지식을 활성화하고 논리적으로 자신의 단어로 과학적 문장을 구성하는 것과 같이 다양한 요소들이 개방형 문항에 포함되어 있기 때문이다.

이와 같은 우리나라 학생들의 상대적으로 저조한 개방형 문항에 대한 능력 그 동안 학교 현장에서 광범위하게 활용된 선택형 문항에 대비한 기형적인 학습 방법에 기인할 수도 있다. Scouller(1998)는 평가의 형식에 따라 학생들이 평가를 준비하는 전략과 방법이 달라지는 것을 확인하였다. 예를 들어서 소화효소의 이름의 앞 글자의 이름을 따서 기억하는 방법으로 선택형 시험에 대비하는 것과 같은 현상이 이와 같은 문제점을 잘 보여준다. 다시 말하면, 오랫동안 선택형 시험에 노출된 학생들은 선택형 평가에 맞춰진 지식을 획득하고 그런 능력을 강화되어 Dochy(2001), Vosniadou(1994), Gitomer, Duschl(2007)가 강조한 구조화된 설명의 논리적 구성 능력이 약해질 수 있음을 짐작할 수 있다.

설명의 구성 능력에 관한 선행연구를 보면, 전문가와 비전문가의 문제 해결 방식을 비교한 Larkin *et al.*(1980)의 연구에서 전문가들의 지식의 구조화 능력은 비전문가에 비하여 매우 높았다. 이와 같은 양상은 Chi *et al.*(1981)의 연구에서도 동일하게 나타난다. 지식을 구조화하는 능력, 즉 설명을 만들어 내는 능력은 전문성을 보여주며, 설명하는 능력의 훈련은 전문성 신장에 큰 도움이 된다(Chi, 1996). Hammer *et al.*(2005)은 문제해결에서 전문가들은 적절한 지식을 활성화 하여 과학적인 구조의 형태로 조직화 하는 능력을 가지고 있다고 하였다. 이와 같은 설명의 이론적 배경은 diSessa(1988)의 'p-prims (phenomenological primitives)' 이다. 하나의 과학적 설명은 다양한 지식 조각들로 구성되어 있으며 각 지식들이 과학적 설명을 구성하는데 있어 적절하게 활용되었는지 또는 적합하게 사용되었는지에 따라 과학적 설명이 될 수 있으며 또는 비과학적 설명이 될 수 있다.

이와 같은 필요성에 의하여 이 연구에서는 생물예비교사의 과학적 설명 구성 능력을 확인해보고자 한다. 박정 등(2000), Park(2010) 등에서 논의한 초 · 중등 학생에 비하여 생물예비교사는 개방형 평가라는 관점에서 매우 중요한 대상이다. 먼저 그들은 12년간 초 · 중등 교육과정을 통하여 공부해 온 학생들이므로 선택형 검사에 익숙하고 선택형 평가에 맞춰진 지식을 가지고 있을 수 있으며, 또 한편으로 그들은 장차 생물 교사가 되어 과학 지식을 학생들에게 설명해 주어야 되는 대상들이다. 특히 생물 교사는 생물 전문가이자 자신의 인지 구조에 있는 전문 생물 지식을 설명 형태로 구성하는데 있어서도 전문가이어야 한다.

다른 관점으로 보면, 위에서 논의한 과학적 지식의 적용 능력, 오개념이 포함되지 않은 세련된 과학적 설명의 생성 등을 선택형 검사가 예측하고 평가 할 수 있는지 확인할 필요성이 있다. Nehm, Schonfeld (2008, 2010)의 연구에서 선택형 평가는 학생들의 지식수준에 대한 평가 일 뿐 응답의 구조적인 측면이나 오개념과 혼합된 형태의 개념을 확인하는데 한계를 가진다고 하였다. 이와 같은 문제점으로 인하여 다층 형태의 선택형 문제를 개발하고 그 효과를 확인하는 연구가 진행되고 있다(Caleon, Subramaniam,

2010; Tan *et al.*, 2002). 이와 같은 다층형태의 선택형 문제들은 단일 문항에 비하여 학생들의 지식수준과 오개념 수준, 그것의 확신성을 평가할 수 있는 수단으로 더 유용하게 활용될 수 있다. 하지만 앞서 이야기 한 바와 같이 선택형 문항이 가지는 근본적인 문제점은 문항에서 너무 많은 양의 정보를 제시하여 학생들이 자신의 지식을 활성화할 수 있는 기회가 없다는 것이다. 평가도 학습의 연장선이며 지식을 활성화 해 보는 경험은 학습을 더 촉진시킬 수 있다는 관점에서 볼 때 선택형 검사는 이와 같은 근본적인 문제점을 가지고 있다(Pyc, Rawson, 2010). 또한 자신의 언어로 설명을 구성하는 것이 아니라 다른 사람의 언어로 구성된 설명을 비교하여 선택하는 방식이므로 단순히 평가 이외의 과학적 설명을 자신의 글로 구성해 볼 수 있는 기회를 제공하지 않으므로 앞서 논의한 Dochy(2001), Vosniadou(1994), Gitomer, Duschl(2007)가 강조한 평가의 본래 목적을 달성할 수 없다.

이 연구는 생물 예비교사의 선택형 방식의 평가 결과와 서술형 방식의 결과를 비교하여 분석하고자 한다. 구체적으로 이 연구의 목적을 달성하기 위하여 선택한 과학적 개념은 진화 개념이다. 진화 개념은 변이, 변이의 유전성, 경쟁, 과잉 생산, 제한된 환경, 차별적 생존과 번식, 개체군 변이라는 7개의 기본 개념이 논리적인 구조로 결합되어야 하나의 완벽한 설명이 된다(Nehm, Ha, 2011). 또한 목적론, 용불용설, 의인화, 개체 내 적응과 같은 비과학적 설명을 배제시킬 수 있어야 한다. 또한 적절한 과학적 인과관계 용어를 사용하여 과학적 설명을 구성할 때 과학적 진화 설명으로 구분할 수 있다(Nehm, Schonfeld, 2008). 방법론적으로, 진화 개념을 확인하는데 널리 활용되고 있는 선택형 검사 도구(the conceptual inventory of natural selection, Anderson *et al.*, 2002)와 Nehm, Reilly(2007)가 사용한 개방형 검사(Open-response instrument)를 비교할 것이다.

구체적으로 이 연구에서 확인하고자 하는 연구 문제는 아래와 같다.

1. 생물 예비교사들의 선택형과 개방형에서 나타나는 진화 개념의 수준은 어떠한가?
2. 선택형 검사의 점수는 개방형 검사를 어느 정도 예측할 수 있는가?
3. 선택형 검사의 점수는 개념 적용의 일관성, 정교성, 설명의 양을 어느 정도 예측할 수 있는가?

Ⅱ. 연구 방법

1. 참여자

이 연구를 위하여 두 대학교에 재학 중인 생물예비교사 124명이 참여해 주었다. 학년별로는 1학년 28명, 2학년 26명, 3학년 32명, 4학년 38명이 참여하였다. 두 대학교 모두 생물 교사를 양성하는 전문 교육기관이며 학부 단위로 신입생을 모집하지 않으며 학과 단위로 모집한다. 자료의 수집은 3월에 이루어졌으므로 1학년의 경우 대학의 일반 생물학 과정을 수강하기 전이다.

2. 검사 도구

선택형과 개방형 도구를 비교하기 위하여 두 가지 검사 도구가 활용되었다. 선택형 도구는 Anderson *et al.*(2002)이 개발한 CINS(the conceptual inventory of natural selection)이다. CINS는 핀치새, 구피 물고기, 도마뱀의 진화 과정을 제시하면서 학생들의 진화 개념을 묻는 형태로 구성되어 있다. 이 검사 도구는 단어 기반 검사 도구가 아닌 개념 기반 검사도구로 선택형 검사 도구를 개발하기 위하여 수집된 다양한 학생들의 과학적 설명과 비과학적 설명에 근거하여 개발되었다. 그러므로 개방형 검사와 동형의 선택형 검사 도구라 할 수 있다(Nehm, Schonfeld, 2008). 이 검사도구는 변이, 변이의 기원, 종의 기원, 유전성, 과잉생산, 개체군 안정, 제한된 자원, 차별적 생존, 제한된 생존, 개체군 변화의 10개 개념 요소로 구성되었으며 각 개념 요소 별 2문항씩 20문항으로 구성되어 있다. 채점은 맞는 문항을 1점, 틀린 문항을 0점으로 하여 0점에서 20점 분포가 되도록 하였다. 개방형 검사 도구는 Nehm, Reilly (2007)가 개발한 박테리아 진화, 치타 진화, 도롱뇽 진화를 사용하였다. 이 세 문항은 Dochy(2001)가 제시한 바와 같이 실제 상황을 검사 문항에 적용한 것으로 참여자는 자신의 지식을 활용하여 과학적 설명을 구성하는 형태로 응답한다. 개방형 검사 도구의 채점은 Nehm *et al.*(2010)의 채점 기준표를 따라 변이, 변이의 유전가

능성, 경쟁, 과잉 생산, 제한된 환경, 차별적 생존과 번식, 집단의 변화의 7가지 진화 개념을 토대로 분석하였다. 개방형 검사의 채점은 Nehm, Ha(2011)의 방법에 따라 이루어졌다. 채점 방법과 내용에 관한 정보를 공유한 생물교육 박사과정생과 생물교사는 60문항을 토대로 두 채점자간 일치도(Cohen′ kappa)를 확인하였고 각 항목별로 0.8이상의 일치도를 확인 한 후 전체 채점을 하였다. 두 채점자간 일치하지 않는 응답에 대해서는 채점자간 심의를 통하여 최종 채점이 이루어졌다. 두 검사도구의 신뢰도(Chronbach's alpha)는 먼저 선택형 검사가 0.74, 개방형 검사에서 과학적 개념 점수가 0.73, 오개념 점수가 0.48이었다. 동일 채점 기준표를 사용한 Nehm, Ha(2011)의 연구에서도 오개념 점수의 신뢰도는 상대적으로 낮았다.

3. 분석 방법

먼저 첫 번째 연구 문제인 선택형과 개방형 검사에서 나타나는 생물 예비교사의 진화 개념을 확인하기 위하여 두 검사에서 나타나는 난이도와 변별도를 각 항목별로 분석하였다. 난이도는 정답자의 비율로 제시하였으며 변별도는 총점 기준 상 · 하위 학생들의 정답률의 차이로 제시하였다. 두 번째 연구 문제인 선택형 검사 결과의 개방형 검사에 대한 예측도를 확인하기 위하여 사용한 방법은 상관관계 분석(Pearson)과 4분위 분석, 교차분석이다. 선택형 점수를 바탕으로 참여자를 최상위, 상위, 하위, 최하위의 4개 그룹으로 구분하여 각 집단 별로 개방형 검사 결과를 비교분석(ANOVA)하였다. 선택형과 개방형 점수를 토대로 4분위한 집단을 교차 분석하여 두 검사 방법에서 몇 명의 참여자가 동일한 집단에 포함되는지를 확인하였다. 또한 선택형과 개방형에서 다른 결과를 보인 학생들의 응답을 통하여 어떤 문제점을 가지고 있는지 질적으로 분석하였다.

마지막 연구 문제인 선택형 검사의 점수가 어느 수준으로 개념 적용의 일관성, 정교성, 설명의 양을 예측하는지 조사하기 위하여 먼저 개방형 응답의 일관성, 정교성, 설명의 양을 측정하였다. 개념의 일관성을 평가하기 위한 방법으로 7개의 개념 요소가 3개의 개방형 문항 중 몇 개 문항에서 일관성 있게 나타났는지 점수화 하였다. 설명의 정교성을 확인하기 위하여 개방형 문항을 순수한 과학적 설명, 과학적 설명과 오개념이 혼합된 형태의 설명, 순수한 오개념으로 구분하여 점수화 하였다. 마지막으로 설명의 양은 각 개방형 문항에서 참여자가 사용한 단어수를 비교함으로 이루어졌다. 선택형 검사의 예측도는 선택형 점수를 토대로 구분된 4개 집단의 일관성, 정교성, 설명의 양에 관한 점수를 비교하는 방법으로 이루어졌다.

Ⅲ. 연구 결과 및 논의

1. 선택형과 개방형에 나타난 진화 개념

먼저 연구 문제 1을 확인하기 위하여 두 검사도구의 각 문항의 변별도와 난이도를 확인하였다(표 1). 이 분석은 연구에 활용한 검사도구의 각 문항별 타당도 확인보다는 두 검사의 비교를 통해 학생들의 개념을 확인하는데 목적이 있다. 결과를 살펴보면, 문항의 정답률을 나타내는 난이도를 보면 선택형 검사의 경우 평균 50%의 정답률을 보인다. 과잉생산, 차별적 생존, 경쟁, 제한된 자원, 유전성, 변이 개념에서 모두 50% 이상의 난이도를 보이고 있다. 반면 개방형의 경우 차별적 생존, 일부 문항에서의 변이 개념을 제외하고 50% 이하의 난이도를 보인다. Choi *et al.*(in press), 박정 등(2000), Park(2010)의 연구 결과와 같이 생물 예비교사들이 선택형에서 더 높은 문제 해결 능력을 보여주며 상대적으로 자신의 지식을 서술하는 형태에서는 낮은 문제 해결 능력을 보였다.

먼저 서론에서 논의한 바와 같이 생물 예비교사는 생물 전문가의 특성을 보여야 한다. 전문가와 비전문가의 문제 해결 방식을 비교한 Larkin *et al.*(1980)의 연구에서 전문가들의 지식의 구조화 능력은 비전문가에 비하여 매우 높았다. 이와 같은 양상은 Chi *et al.*(1981)의 연구에서도 동일하게 나타난다. 지식을 구조화하는 능력, 즉 설명을 만들어 내는 능력은 전문성을 보여주며, 설명하는 능력의 훈련은 전문성 신장에 큰 도움이 된다(Chi, 1996). Hammer *et al.*(2005)은 문제해결에서 전문가들은 적절한 지식을 활성화 하여 과학적인 구조의 형태로 조직화 하는 능력을 가지고 있다고 하였다. 동물 행동학자들의 연구 활동을 분석한 이안나 등(2007)의 연구를 보면 생물학자들의 연구 과정에서 나타나는 끊임없는 의문 생성과 그에 따른 잠정적 설명의 생성이 생물학 연구에서 얼마나 중요한지 보여준다. 대학의 생물교육이 전문

표 1
선택형과 개방형 검사 도구의 난이도와 변별도

개념	선택형			개방형		
	문항번호	난이도	변별도	문항	난이도	변별도
과잉생산	1	0.61	0.19	세균	0.00	0.00
	11	0.55	0.29	치타	0.00	0.00
				도롱뇽	0.01	0.02
개체군 변이	4	0.38	0.37	세균	0.12	0.18
	13	0.52	0.37	치타	0.06	0.11
				도롱뇽	0.02	0.03
차별적 생존	10	0.61	0.13	세균	0.65	0.40
	18	0.53	0.32	치타	0.68	0.55
				도롱뇽	0.41	0.63
경쟁	5	0.83	0.15	세균	0.01	0.02
	15	0.72	0.40	치타	0.05	0.06
				도롱뇽	0.00	0.00
제한된 자원	2	0.71	0.26	세균	0.00	0.00
	14	0.88	0.18	치타	0.36	0.37
				도롱뇽	0.01	0.02
유전성	7	0.64	0.50	세균	0.17	0.31
	17	0.58	0.35	치타	0.16	0.26
				도롱뇽	0.10	0.19
변이	9	0.50	0.13	세균	0.67	0.44
	16	0.80	0.27	치타	0.31	0.55
				도롱뇽	0.17	0.31
종의 기원	8	0.54	0.50	–	–	–
	20	0.58	0.55			
변이의 기원	6	0.67	0.47	–	–	–
	19	0.57	0.66			
개체군 안정	3	0.90	0.10	–	–	–
	12	0.84	0.00			

화된 생물학자의 양성이라면 생물 지식을 구조화 하여 설명하는 능력의 양성은 매우 중요하다. 평가 방법에 관한 논의로 돌아가서 생물 예비교사의 설명 능력을 점검하기 위해서는 개방형 검사의 활용이 필수적으로 요구된다.

다음으로 제시되는 질문이 왜 이런 현상이 나타나는지 그 원인에 관한 질문이다. 설명을 구성하는 단계에서 지식을 가지고 있지만 활성화 하지 못하거나, 문제와 연관시키지 못하면 지식 요소는 응답에서 나타나지 않는다. 즉 지식을 활성화하기 위해서는 적절한 발판이 필요한 것이다(Chi, 1996). 개체군 변이나 경쟁과 같은 개념 요소를 알고 있지만 그것을 서술하지 않는 이유는 검사지의 질문이 그 지식을 활성화 할 수 있을 만큼의 정보를 가지고 있지 않으며, 피평가자가 그 질문에 대한 답을 구성하는데 있어 그와 같은 개념 요소들이 필요한 것인지 인식하지 못하기 때문이다(Hammer *et al.*, 2005). 사물이나 현상에 대한 개념적 설명이 이루어지기 위해서는 선언적 지식(declarative knowledge)의 재구성에 의한 학생의 능동적인 지식 생성이 반드시 선행되어야 한다(Kwon *et al.*, 2009;

Lee, 2009). 지식 생성 과정에서 다양한 기억 체계가 요구되는데 선택형 평가에 익숙할 경우 지식 생성에서 중요한 역할을 하는 작업 기억 능력을 많이 활용하지 않아 상대적으로 개방형 평가에서 낮은 문제 해결 능력을 보일 수 있다(Lee, Kwon, 2011).

2. 선택형 평가 결과의 개방형 평가 결과에 관한 예측 능력

선택형 평가의 결과가 개방형 평가 결과를 어느 정도 예측할 수 있는지 확인하였다. 이 분석을 위하여 세 가지 통계 분석을 활용하였다. 먼저 Pearson 상관관계분석이다. 선택형 검사도구와 개방형 검사도구의 Pearson 상관관계 분석 결과 상관관계는 개방형 검사 결과의 과학적 개념 점수와 0.518($p < 0.01$), 오개념 점수 결과와 −0.483($p < 0.01$)로 유의미한 상관관계를 보였다. 대략 0.5의 상관관계의 설명력은 25%로 이 결과를 토대로 분석하였을 때 개방형 검사 결과는 25% 정도의 설명력을 가진다고 볼 수 있다. 두 번째 상관관계 분석보다 보다 융통성이 있는 방법으로 두 검사 결과를 토대로 구분된 네 집단의 교차 분석으로 선택형 검사의 예측도를 확인하였다.

먼저 선택형 검사 결과를 바탕으로 최상위, 상위, 하위, 최하위의 4 집단으로 구분하였다. 최상위 집단의 평균 선택형 검사 결과는 20점 만점에 17.29(SD = 1.16) 이며, 상위집단 14.84(SD = 0.64), 하위 집단 11.90(SD = 1.16), 최하위 집단 7.84(SD = 1.77)이다. 최상위집단과 상위 집단의 점수 차이는 2.45이며, 상위집단과 하위 집단의 차이는 2.94, 하위 집단과 최하위 집단의 차이는 4.06점이다. 네 집단 모두 통계적으로 유의미한 차이를 가지고 있다($p < 0.01$). 또한 전체 참여자의 평균 점수는 12.97(SD = 3.75) 미국 2년차 생물 예비교사 12.58(SD=3.98)과 비슷하며(Nehm, Schonfeld, 2007) 미국 생물비전공 대학생 점수 10.42(SD = 3.31)보다 상회한다(Anderson *et al.*, 2002).

연한 회색은 선택형 검사에 비하여 개방형검사에서 상위 집단에 포함된 참여자의 수이며, 진한 회색은 반대로 선택형 검사에 비하여 개방형 검사에서 하위 집단에 포함된 참여자의 수이다. 먼저 이 결과의 Chi 제곱 검정 결과 $X_9 = 60.8$($p < 0.001$)로 집단 간 차이가 있음을 알 수 있다. 또한 두 검사 결과가 한 단계 차이 나는 참여자는 63명으로 전체 50.8%이며 두 단계 이상 차이가 나는 참여자 수는 총 15명으로 12.1%이다.

마지막 분석은 각 집단 별 개방형 검사의 비교이다. 먼저 선택형 검사를 기준으로 구분된 집단 간 개방형 검사의 과학적 개념 검사의 집단 간 비교는 통계적으로 유의미하였다($F_{3,120} = 17.0$, $p < 0.001$). 최상위 집단의 경우 과학적 개념 점수가 6.29(SD:2.07)이며 사후 분석 결과 다른 세 집단에 비하여 통계적으로 유의미하였다. 하지만 상위집단, 하위집단, 최하위 집단의 개방형 검사 결과 점수는 통계적으로 유의미하지 않았다(상위집단: M = 3.84, SD = 2.65; 하위집단: M = 3.23, SD = 2.36; 최하위집단: M = 2.42, SD = 1.86). 개방형 검사에서 나타난 오개념의 점수의 비교에서 집단 간 차이가 나타났다($F_{3,120} = 12.3$, $p < 0.001$). 최상위 집단에서 최하위 집단으로 내려갈수록 단계적으로 오개념의 점수가 높아지는 것을 확인할 수 있다(최상위집단: M = 0.52, SD = 0.85; 상위집단: M = 1.29, SD = 1.22, 하위집단: M = 1.71, SD = 1.16; 최하위집단: M = 2.23, SD = 1.31)

연구 문제 2를 확인하기 위하여 실시한 상관관계 분석, 교차분석, 집단 별 평균 분석 결과 선택형 검사는 일부분 개방형 검사 결과를 예측하고 있으나 매우 높은 수준은 아님을 알 수 있다. 상관관계 분석 결과 약 25%의 설명력을 보이고 있으며, 약 12%의 학생들이 4단계 구분에서 2단계 차이 나는 집단에 포함되는 것을 확인할 수 있다. 또한 선택형 검사에서 최상위 집

표 2
선택형과 개방형 검사 결과의 교차 분석

개방형 \ 선택형	최상위	상위	하위	최하위
최상위	16 (12.9)	9 (7.3)	5 (4.0)	1 (0.8)
상위	10 (8.1)	12 (9.7)	8 (6.5)	1 (0.8)
하위	4 (3.2)	7 (5.6)	12 (9.7)	8 (6.5)
최하위	1 (0.8)	3 (2.4)	6 (4.8)	21 (16.9)

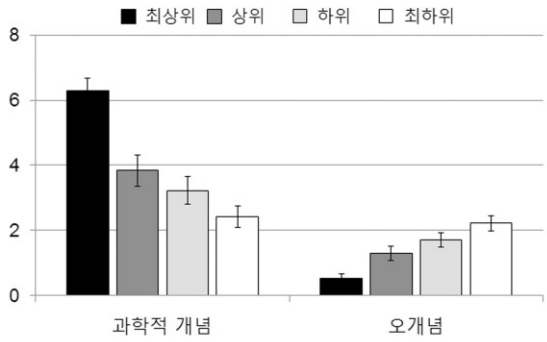

그림 1 *집단 별 개방형 점수 비교*

단을 제외하고 상위, 하위, 최하위 집단의 개방형 점수 결과는 유의미하게 구분되지 않는다.

이상의 결과를 살펴보면 선택형 검사 결과가 일부 개방형 검사 결과를 예측할 수 있으나 상관관계 분석에서 제시된 25%에서 교차분석으로 확인한 50% 정도의 예측 능력을 짐작할 수 있다. 또한 선택형 검사의 경우 최상위 집단을 제외한 나머지 집단에서 개방형 검사 결과의 차이점이 나타나지 않는 것은 선택형 검사가 개방형 검사를 대체할 때 변별도에 문제가 있음을 이해할 수 있다.

선택형과 개방형에서 차이가 많이 나는 응답을 토대로 질적 분석 해 보았다. 먼저 선택형 문항에서 19점으로 최상위 집단인 반면, 개방형 문항에서 과학적 개념 4점, 오개념 2점을 으로 상당히 낮은 점수를 보였다.

세균 문항: 우연적으로 항생제 저항성을 지닌 박테리아가 저항성 없는 박테리아보다 잘 살아남고[차별적 생존] 생식에 참여하였다. 그럼 그 자손들의 비율에서 저항성 있는 박테리아가 증가한다[개체군 변이].

치타 문항: 빠른 치타가 느린 치타보다 먹이를 먼저 먹고 많이 먹고[제한된 자원] 자식을 많이 낳을 확률이 높다[차별적 생존]. 그럼 전체적인 속도가 증가.

도롱뇽 문항: 시력을 갖기 위한 발생과정에서 빛에 의한 자극은 필수적이다[용불용설]. 동굴에서 사는 도롱뇽의 경우에는 빛에 의한 자극이 필요조건이 아니므로 서서히 발생상에서 시력을 지닐 잠재력이 줄어들었을 것이다[적응]. 줄어든 잠재력이 고정된 것이 아닐까?

이 학생의 경우 세 문항 모두 최초 변이가 어디에서 기원했는지 모두 고려하고 있지 않다. 변이와 변이의 유전성은 진화 설명에 있어서 3가지 핵심 개념 중 두 가지이다. 두 번째로 중3과학과 생물Ⅱ 교과서에서 강조하고 있는 과잉생산, 제한된 자원, 경쟁에 대한 언급이 치타 문항에서 한번 언급되고 있다. 마지막으로 생각해 보아야 할 응답은 도롱뇽 문항이다. '시력을 갖기 위한 발생 과정에서 빛의 자극이 필수적이다'라는 말은 과학적으로 바르지 않다. 빛 자극은 시력이 발달하는 과정에서 시력과 관련된 근육의 운동과 관련이 된다. 그럴듯한 과학적 용어로 설명해 놓은 이 응답은 전혀 과학적 설명이 아니다. 또한 '잠재력이 고정되었다'라는 표현은 개체군적 사고가 아닌 개체 내 발달을 진화와 연관 시키는 전형적인 진화 오개념이다(Nehm, Ha, 2011).

위 응답과 대비되는 응답으로 선택형에서 18점으로 최상위 그룹에 속하였으며 개방형 검사에서 11점을 받고 전혀 오개념이 나타나지 않은 응답을 살펴보겠다.

세균 문항: 모든 생물의 특징 중 하나가 돌연변이[변이]이다. 박테리아 역시 여러 세대에 걸쳐 돌연변이가 발생한다. 이 중에서 항생제에 대해 저항성을 가지는 박테리아가 존재하게 되고, 당연히 항생제가 있는 환경에서[제한된 자원] 생존에 유리하다[차별적 생존]. 이러한 형질을 가진 박테리아는 살아남아 많은 후손에 형질을 유전시켜[유전성], 박테리아가 항생제에 저항성을 가지도록 진화된다.

치타 문항: 치타가 여러 세대에 걸쳐 돌연변이[변이]가 일어나고, 이 중에는 빠른 속도로 달리는데 도움을 주는 형질을 가진 치타도 있을 것이다. 이러한 형질을 가진 치타는 상대적으로 더 빠른 속도로 달리게 되고, 먹이를 사냥하거나 위험한 상황에서 도망치는데 유리하여[제한된 자원] 더 많이 생존하게 된다[차별적 생존]. 따라서 더 많은 자손에게 형질을 유전시켜[유전성] 현재의 치타로 진화되어 왔을 것이다.

도롱뇽 문항: 도롱뇽이 세대를 거쳐 돌연변이[변이]가 발생하면서 시력보다는 청각이나 촉각 등이 더 발달한 도롱뇽이 생겼을 것이다. 이러한 시력이 발달하지 않아도 생존에 유리했던 도롱뇽들이 살아남아[차별적 생존] 형질을 유전시켜[유전성] 현재의 도롱뇽으로 진화되었을 것이다.

이 참여자가 작성한 응답을 보면 우리 교과서에서 강조하고 있는 변이, 변이의 유전성, 제한된 자원, 차별적 생존에 대한 응답을 3 문항에 일관성 있게 적용하고 있다. 하지만 여전히 안타까운 것은 중3과학과 생물Ⅱ 교과서에서 강조하고 있는 과잉생산, 경쟁, 개체군 변이와 같은 내용이 나타나고 있지 않다는 것이다.

마지막으로 확인해 볼 응답은 선택형 문항이 16점으로 최상위 집단이지만 개방형 검사에서 낮은 점수와 많은 오개념이 포함된 참여자의 응답이다.

> **세균 문항**: 생명체는 개체 생존과 종족 유지를 통한 자신의 유전자를 다음세대에 전달하기 위한 목적을 달성하기 위해[목적론] 오랜 시간 진화해 왔습니다. 박테리아는 지구상에 존재하면서 곰팡이 등의 다른 생물이 만드는 항생제를 접하게 됩니다. 그 과정에서 자신의 유전자 전달을 달성하기 위해 자신의 생명을 위협하는 항생제를 불활성화 시키거나 분해, 또는 항생제에 민감한 대사경로를 대체하는 대사경로를 사용하는 등의 방법으로 저항성을 발달시켜오게 됩니다[내부의지, 적응].
> **치타 문항**: 더 빨리 달리는 치타들이 사냥에 성공할 확률이 높으므로[제한된 자원] 더 빨리 달리는 치타는 생식 가능한 연령까지 살아남아 자손을 남길 기회가 많아집니다[차별적 생존]. 이러한 자연선택이 지속되어 사냥에 적합하도록 빨리 달릴 수 있는 신체적인 여건을 가진 치타가 많이 살아남아 세대가 거듭됨에 따라 치타는 빨라지게 됩니다.
> **도롱뇽 문항**: 동굴에는 빛이 잘 들지 않으므로 시각은 사용될 기회가 적습니다[용불용설]. 이런 환경에서 시각은 오히려 그 구조를 만들고 유지시키는데 에너지적 비용이 많이 들 뿐만 아니라 쉽게 감염될 수 있는 등 생존에 불리한 결과를 낳을 수도 있습니다. 따라서 시각보다는 다른 감각 기관을 발달시킨 도롱뇽이[내부의지, 적응] 생존에 유리하므로[차별적 생존] 자연선택에 의해 눈은 퇴화될 수 있습니다.

이 참여자의 응답은 다른 참여자들에 비하여 매우 많은 단어를 사용하였다. 하지만 대부분의 단어들은 과학적 설명을 구성하는데 있어서 매우 불필요한 오개념들이다. 먼저 많은 과학교육학자들과 인지 심리학자들이 과학적 설명 구성의 장애요인으로 고려하고 있는 목적론을 사용하고 있으며 '세균이 다양한 방법을 사용하여 저항성을 발달하고 있다'는 설명과 같이 직관적이며 의인화된 설명을 이용하고 있다. 또한 치타 문항에서 '세대가 거듭됨에 따라 치타는 빨라진다.'라는 응답은 개체군 변이의 의미를 내포하고 있는 것처럼 보이지만 세련되지 못한 설명이다. 세대가 거듭됨에 따라 빠른 치타의 비율이 높아진다는 표현이 과학적으로 바른 표현이다. 이 참여자의 응답은 과학적 개념과 비과학적 개념이 혼합된 형태의 전형적인 모습을 보여준다. 과학적으로 그럴듯한 설명이면서도 전혀 과학적이지 않은 설명은 과학적 개념이 형성되면서 이전에 가지고 있던 비과학적 설명이 제거되지 못하고 새로운 과학적 설명과 혼합되는 경우이다. 그러므로 오개념은 지속적으로 잔류할 수 있고 이와 같은 표현들은 학생들에게 전이될 가능성이 있다. 이 혼합 개념은 아래 그림 5에서 다시 논의할 것이다.

3. 선택형 검사결과의 개념 적용의 일관성, 정교성, 설명의 양에 관한 예측 능력

선택형 검사의 점수가 개념 적용의 일관성, 정교성, 설명의 양을 어느 정도 예측할 수 있는지 확인하였다. 먼저 개념의 일관성을 확인하기 위하여 선택형 검사를 바탕으로 구분된 네 집단(최상위집단: 0-25%, 상위집단: 25-50%, 하위집단: 50-75%, 최하위집단: 70-100%)의 각 과학적 개념 요소와 비과학적 개념별 요소의 점수를 비교하였다. 개방형 검사가 3문항으로 구성되었으므로 모든 문항에 개념 요소를 적용시킨다면 최대점은 3점이다. 만약 점수가 2점이라면 약 2 문항에서 그 개념요소가 나타났다고 이해할 수 있다. 그림 2를 보면 선택형 점수 기준 최상위 집단은 변이 개념에서 평균 1.87로 약 2문항에서 변이 개념이 나타났음을 확인할 수 있다. 상위 집단은 평균 1.19로 약 한 문항에서 변이 개념이 나타났으며 하위와 최하위 집단은 각각 0.84, 0.68이었다. 유전성 개념의 경우에는 최상위 집단도 평균 한 문항 이상 나타나지 않았다(최상위: 0.74, 상위: 0.42; 하위: 0.35; 최하위: 0.19). 차별적 생존의 경우 최상위 집단은 평균 점수 2.71로 많은 학생들이 3문항 모두 차별적 생존 개념을 활용하였다. 사후 분석 결과를 보면 변이 개념과 차별적 생존 개념에서 최상위집단을 제외한 상위, 하위, 최하위 집단의 개념의 차이가 나타나지 않는다($p < 0.05$). 유전성 개념 역시 최상위집단과 최하위집단의

경우에만 유의미한 차이가 있다($p < 0.05$).

그림 3에 선택형 검사를 바탕으로 구분된 네 집단의 비과학적 개념 요소 별 점수의 평균 비교가 제시되어 있다. 최상위 집단의 목적론 점수는 0.23으로 가장 낮았으며 상위집단 0.84, 하위집단 0.94로 나타났다. 그 외 용불용설, 내부의지, 적응과 같은 오개념 요소는 평균 점수 0.4점으로 약 7개 응답 중 하나의 응답에 이와 같은 오개념이 나타나고 있음을 알 수 있다. 사후 분석 결과 목적론의 경우 최상위집단을 제외한 3 집단에서 유의미한 차이가 나타나지 않았다($p < 0.05$). 적응 오개념 역시 상위 집단과 최하위집단에서만 유의미한 차이가 나타났을 뿐 나머지 집단에서는 유의미한 차이가 나타나지 않았다($p < 0.05$).

그림 2와 그림 3을 통해 확인할 수 있는 것은 생물 예비교사들이 어느 정도로 일관되게 과학적 개념을 적용할 수 있는지, 또는 어느 정도로 일관되게 오개념 요소를 사용하지 않는지에 관한 것이다. 위의 세균, 치타, 도롱뇽 문항은 문항에서 사용한 생물의 종류만 다를 뿐 동일한 형태의 문제이다. 이와 같은 동형문항에서 어떤 문항에서는 그 개념을 사용하는 반면 다른 문항에서 그것을 사용하지 않는 것은 개념을 일관적으로 사용하지 않는 것으로 개념의 다양한 상황에 적용 능력에 문제가 있는 것으로 이해할 수 있다. Nehm, Ha(2011)는 이와 같은 동형문항에서 표면적인 특징이 지식의 활성화와 관련되어 있으며 이와 같은 문항의 특징을 넘어서 일관성이 있게 응답하는 경우 개념이 체계적으로 조직화 되어 있다고 논의하였다. 이 논의 점을 다시 확인하기 위하여 각 문항별 사

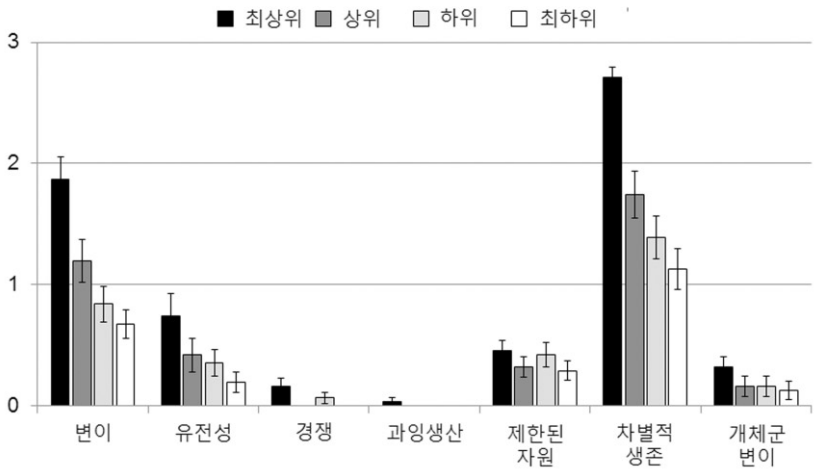

그림 2 *집단 별 개방형 검사에서 나타난 과학적 진화 개념 요소 별 점수*

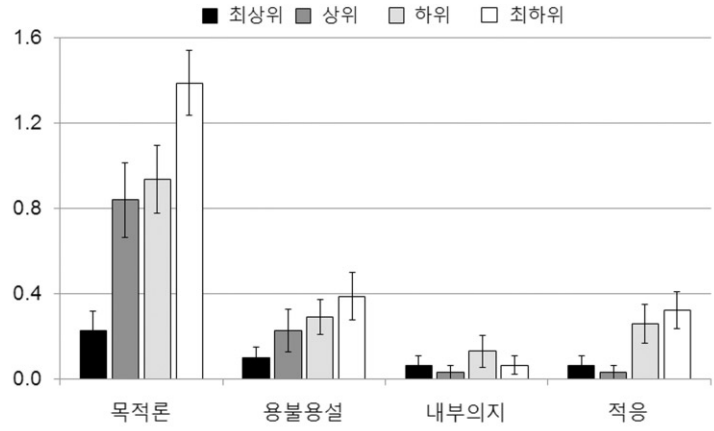

그림 3 *집단 별 개방형 검사에서 나타난 비과학적 개념 요소 별 점수*

용한 과학적 개념의 양을 비교하였다. 그림 4에서 각 집단별로 세 문항에 제시한 과학적 개념 요소의 양을 비교하였다. 네 집단 모두 세균, 치타에서 사용한 개념 요소를 마지막 문항인 도롱뇽 문항에 사용하고 있지 못함을 확인할 수 있다. 최상위 집단은 세균 문항에서 2.23개, 치타 문항에서 2.45개의 과학적 개념 요소를 활용한 반면 도롱뇽 문항에서는 1.61개만 사용하였다. 이와 같은 패턴은 다른 집단에서도 동일하게 나타난다. 위에서 논의한 바와 같이 동형 문항에서 한 문항에서 활용한 과학적 개념 요소를 그대로 다른 개념에 적용시킬 수 없는 것은 설명의 일관성의 측면에서 논의 사항이 될 수 있을 것이다.

그림 5는 앞서 논의한 과학적 개념과 오개념이 혼합된 형태의 개념을 점수화하여 제시하였다. 이 결과는 생물예비교사들의 진화 개념의 세련된 정도를 보여주는 것으로(Nehm, Ha, 2011), 비과학적 개념을 사용하지 않고 순수한 과학적 개념을 활용하여 설명을 구성하였는지 보여준다. 최상위 집단의 경우에 순수한 과학적 모델이 83.9%로 가장 높았지만 여전히 11.8%는 과학적 개념과 오개념을 혼합하고 있었다. 상위 집단 역시 58.1%는 순수한 과학적 개념 요소만, 8.6%는 혼합된 개념, 26.9%는 오개념 요소만 포함된 설명을 보였다. 하위 집단은 45.2%가 과학적 모델, 11.8%가 혼합 모델, 최하위 집단은 34.4%가 과학적 모델, 11.8%가 혼합 모델이다. 전체적으로 평균 10%의 생물 예비교사가 혼합된 형태의 모델을 가지고 있는 것으로 보이며, 그 수치는 선택형 평가에서 최상위 집단의 경우도 마찬가지이다.

앞서 학생들의 응답을 분석한 결과에서 나타난 것과 같이, 혼합개념은 과학 개념 연구에서 꾸준히 연구되고 있다(Ha, Cha, 2009; Nehm, Ha, 2011; Southerland *et al.*, 2001; Wittmann, 1998;

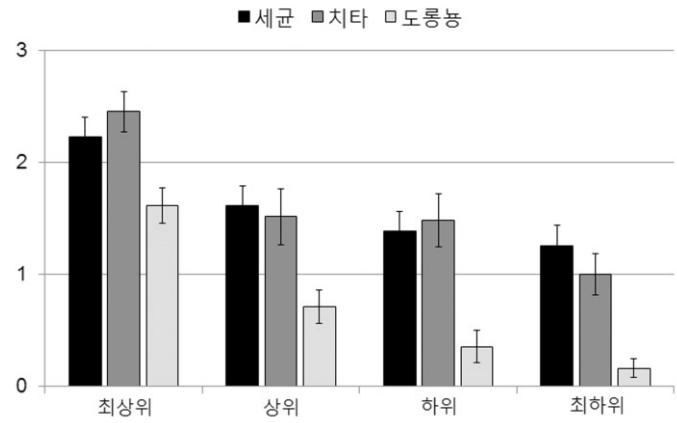

그림 4 *집단별 개방형 문항별 과학적 개념 요소의 양*

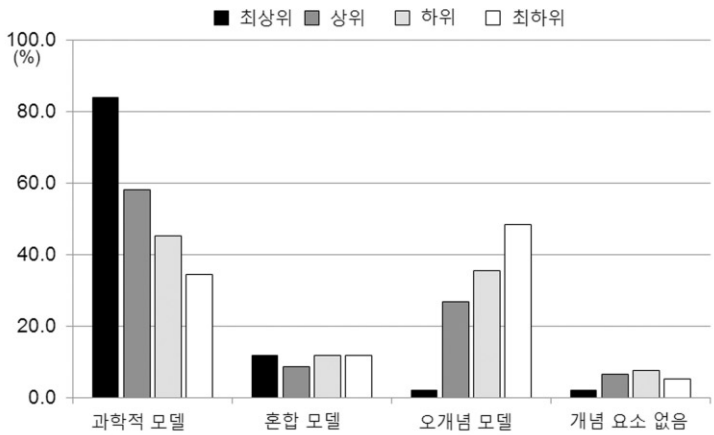

그림 5 *개방형 검사에서 나타난 과학적 모델*

Vosniadou, Brewer, 1992; Vosniadou, 2007). Ha, Cha(2009)는 생물 예비 교사들이 과학적 개념과 비과학적 개념을 동시에 가지고 있으며 때로는 그 개념들을 서로 혼합하여 사용하는 것을 확인하였다. Nehm, Ha(2011)은 미국의 생물 전공 대학생들이 목적론, 용불용설과 같은 오개념이 자연선택의 여러 주제어와 혼합하여 진화에 관한 설명을 구성하는 것을 확인하였다. Southerland *et al.*(2001)의 연구에서도 위의 학생이 가지고 있는 목적론과 과학적 개념의 혼합된 형태를 보고하였다. 이와 같은 혼합 개념은 다른 과학 영역에서도 연구되는데 Wittmann(1998)은 물리 문제 해결에서 과학적 개념과 비과학적 개념을 동시에 활용하여 물리 문제를 해결하는 것을 관찰하였다. Vosniadou, Brewer(1992), Vosniadou (2007)는 지구의 모양에 관하여 학생들이 학교에서 배운 개념과 원래 가지고 있던 오개념을 혼합한 형태를 가지고 있음을 확인하였다. 이와 같은 개념의 혼합은 비과학적 개념이 과학적 개념에 대해서 대체되지 않고 일부 공존하여 인지 구조에 잔존하고 있음을 보여준다. 생물 교사들 역시 비록 과학적 개념을 수용하고 있으나 여전히 비과학적 개념도 수용하고 있을 수 있다.

세 번째로 확인해 볼 것은 선택형 평가의 점수와 설명의 양이다. 설명의 양은 응답자의 설명의 단어의 수로 확인하였다. 단어의 수는 띄어쓰기를 기준으로 변환하였으므로 조사의 사용은 단어 수에 포함되지 않는다. 앞서 확인한 학생들의 응답에서 보여주듯이, 진화적 설명과 관련이 없는 응답들이나 오개념 요소와 같은 불필요한 단어를 많이 사용하는 경우에서 볼 수 있듯이 사용한 단어의 수가 많다고 그 응답이 많은 과학적 개념을 포함하고 있는 것은 아니다. 하지만 일반적으로 생각하였을 때 지식의 양이 많으면 동일 시간에 많은 양의 지식을 활성화 시키고 자신의 응답에 기록할 수 있을 것이다. 먼저 상관관계 결과를 보면 선택형 검사 결과와 개방형 검사 결과의 단어수의 상관관계는 0.285($p < 0.01$)로 유의미한 상관관계를 가진다. 또한 개방형 검사에서 과학적 개념 요소의 개수와 전체 단어수의 상관관계는 0.518($p < 0.01$)이다. 선택형 검사는 단어수를 약 8.1% 예측하고 있으며 개방형 검사에서 과학적 개념 요소를 26.8% 예측하고 있다. 두 번째로 분석한 것은 선택형 검사 결과를 토대로 구분된 네 집단 별 단어수의 평균 비교이다(그림 6). 최상위 집단의 경우 111.4단어를 사용한 반면 나머지 세 집단에서는 유의미한 차이가 나타나지 않았다($p < 0.01$).

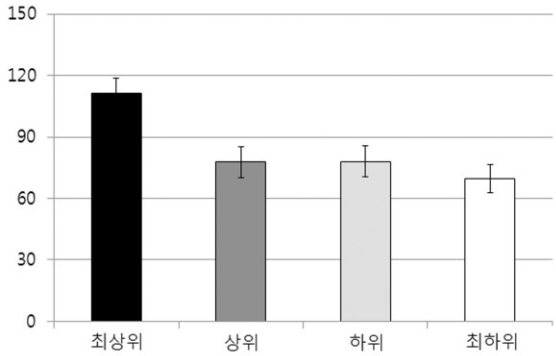

그림 6 *개방형 검사에서의 사용한 단어 수*

Ⅳ. 결론 및 제언

이 연구는 생물 예비교사들의 진화 개념의 구조와 특징을 분석하고 그것을 측정하는 진화 개념 평가 도구의 효용성을 비교하기 위하여 진행되었다. 이 연구의 결과 생각해 볼 수 있는 결론은 다음과 같다.

첫째, 생물 예비교사들은 개방형 검사에 비하여 선택형 검사에 더 높은 점수를 보이고 있다. 생물예비교사는 장차 교실 현장에서 과학적 지식을 설명으로 구성하여 학생들을 이해시켜야 하는 역할을 수행해야 함으로 과학적 설명을 자신의 말로 구성해 내는 능력이 더욱 중요하게 고려되어야 한다. 그러므로 생물 예비교사들의 설명 생성 능력에 대한 다양한 교육 프로그램이 요구되며 그와 같은 능력을 더욱더 적극적으로 평가에 반영해야 할 것이다.

둘째, 선택형 검사 결과는 개방형 검사 결과를 정확하게 예측하지 못하였다. 또한 선택형 검사 도구로 개념 변화 과정을 확인할 수 있는 혼합개념과 같은 중요한 정보를 획득할 수 없었다. 만약 생물 예비교사들의 설명 구성 능력을 확인하고자 할 경우 선택형 검사도구는 제한된 범위에서 사용될 수밖에 없을 것이다. 선택형 검사 도구는 빠른 피드백, 낮은 채점비용과 같은 다양한 장점을 가지고 개방형 검사 도구는 설명 구조를 보다 정교하게 분석할 수 있는 장점을 가진다. 개방형 검사의 채점과 분석에 들어가는 노력이 상당하기 때문에 평가 목적에 맞게 두 검사 도구를 적절히 선택하여 활용해야 할 것이다. 하지만 서론에서 강조

하였듯이 생물 예비교사는 장차 학생들을 교육하는 역할을 담당하기 때문에 교사교육 관점에서 그들의 설명 능력을 확인하기 위해서는 개방형 검사 문항을 사용하는 것이 더 효율적일 것이다.

셋째, 선택형 평가 결과는 개념 요소의 일관성, 순수성, 풍부도를 명확하게 예측하지 못하였다. 이와 같은 결과는 정답을 선택하게 하는 검사도구의 한계점을 보여준다. 향후 개발되는 선택형 검사 도구는 개념 요소의 일관성, 순수성, 풍부도를 보여줄 수 있도록 개발하는 것이 바람직 할 것이다.

국문 요약

생물교사는 자신의 글과 말을 활용하여 지식을 논리적인 과학적 설명의 형태로 변환하여 학생들을 지도한다. 그러므로 생물 예비교사들이 과학적 설명을 구성하는 능력을 확인할 수 있는 개방형 검사가 요구될 것이다. 하지만 채점자간 일치도, 늦은 피드백과 같은 개방형 검사가 가진 문제로 선택형 검사가 대체될 수 있을 것이다. 이 연구는 생물 예비교사들의 과학적 진화 설명 구성 능력을 확인함과 동시에 선택형 진화 개념 검사가 개방형 검사 결과를 어느 정도 예측하는지 두 가지를 조사하였다. 생물 예비교사 124명이 참여하였으며 20문항의 진화 개념 선택형 검사도구와 3문항의 실제 사례를 활용한 개방형 검사 문항을 모든 참여자에게 투입하였다. 연구 결과 생물 예비교사는 선택형 검사 도구에 더 높은 능력을 보였으며 최상위(0 ~ 25%)의 참여자를 제외하고 선택형 검사 도구는 개방형 검사 도구 검사 결과를 예측하지 못하였다. 진화 설명의 일관성, 오개념이 포함되지 않은 순수성, 과학적 설명의 양을 측정한 결과, 선택형 검사 결과는 최상위 집단을 제외하고 다른 집단을 구분하지 못하였다. 또한 이 결과들은 생물 교사들이 과학적 설명의 구성 능력이 지식의 양에 비하여 상대적으로 낮은 것을 확인할 수 있었으며 일부 예비교사들은 오개념과 과학적 개념을 혼합하는 등의 세련되지 못한 과학적 설명을 가지고 있었다. 이 결과는 생물 예비교사들의 개념 평가는 선택형보다 개방형이 더 많은 정보를 제공할 수 있음을 보여주며, 생물 예비교사들을 위한 교사양성 프로그램은 과학적 설명 능력의 신장에 더욱 더 초점을 맞출 것을 제언한다.

참고 문헌

박정, 홍미영 (2002). 문항 유형에 따른 과학 능력 추정의 효율성 비교. 한국과학교육학회지, 22(1), 122-131.

박정, 홍미영, 김성숙, 전현정 (2000). 제3차 수학 · 과학 성취도 국제 비교 연구 (TIMSS-R) 국내 평가 결과 분석 연구Ⅱ. 한국교육과정평가원. 연구보고 RRE 2000-7.

이안나, 권용주, 정진수, 양일호 (2007). 동물 행동학자의 연구 활동에서 나타나는 연구 단계, 사고 과정, 행동 양식 및 생성 지식에 관한 연구. 한국생물교육학회지, 35(3), 361-373.

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. Journal of Research in Science Teaching, 39(10), 952-978.

Caleon, I. S., & Subramaniam, R. (2010). Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students alternative conceptions. Research in Science Education, 40(3), 313-337.

Chi, M. T. H. (1996). Constructing self-explanations and scaffolded explanations in tutoring. Applied Cognitive Psychology, 10(7), 33-49.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5(2), 121-152.

Choi, K, Lee, H, Shin, N., Kim, S. W., & Krajcik, J. (2011). Re-conceptualization of scientific literacy in South Korea for the 21st Century. Journal of Research in Science Teaching. 48(6), 670-697.

diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), Constructivism in the computer age (pp. 49-70). Hillsdale, NJ: Erlbaum.

Dochy, F. (2001). A new assessment era:

different needs, new challenges. Learning and Instruction, 10(1), 11-20.

Gitomer, D. H., & Duschl, R. A. (2007). Establishing multilevel coherence in assessment. In P. A. Moss (Ed.), Evidence and decision making. The 106 th yearbook of the National Society for the Study of Education, Part I (pp. 288-320). Chicago: National Society for the Study of Education.

Ha, M. & Cha, H. Y. (2009). Pre-service teachers' synthetic view on Darwinism and Lamarckism. Paper in the proceedings at 2009 International Conference of National Association for Research in Science Teaching, Garden Grove, CA.

Hammer, D. M., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), Transfer of learning from a modern multidisciplinary perspective (pp. 89-120). Greenwich, CT: Information Age Publishing.

Kwon, Y. J., Lee, J. K., Shin, D. H., & Jeong, J. S. (2009). Changes in brain activation induced by the training of hypothesis generation skills: An fMRI study. Brain & Cognition, 69, 391-397.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208(4450), 1335-1342.

Lee, J. K. (2009). Dissociation of the brain activation network associated with hypothesis-generating and hypothesis-understanding in biology learning: Evidence from an fMRI study. Unpublished Doctoral Dissertation. Cheongwon, Chungbuk: Korea National University of Education.

Lee, J. K., & Kwon, Y. J. (2011). Why traditional expository teaching-learning approaches may founder? An experimental examination of neural networks in biology learning. Journal of Biological Education, 45(2), 83-92.

Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. Journal of Research in Science Teaching, 48(3), 237-256.

Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. BioScience, 57(3), 263-272.

Nehm, R. H., & Schonfeld, I. S. (2007). Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools? Journal of Science Teacher Education, 18(5), 699-723.

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., & Schonfeld, I. S. (2010). The future of natural selection knowledge measurement: A reply to Anderson et al.(2010). Journal of Research in Science Teaching, 47(3), 358-362.

Nehm, R. H., Ha, M., Rector, M., Opfer, J., Perrin, L., Ridgway, J., Mollohan, K. (2010). Scoring Guide for the Open Response Instrument (ORI) and Evolutionary Gain and Loss Test (EGALT). Technical Report of National Science Foundation REESE Project 0909999. Accessed online January 10, 2011 at: http://evolutionassessment.org

Park, J. (2010). Constructive multiple choice testing system. British Journal of Educational Technology, 41(6), 1054-1064.

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. Science, 330(6002), 335.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question

examination versus assignment essay. Higher Education, 35(4), 453-472.

Southerland, S. A., Abrams, E., Cummins, C. L., & Anzelmo, J. (2001). Understanding students' explanations of biological phenomena: Conceptual frameworks or p-prims? Science Education, 85(4), 328-348.

Tan, K. C. D., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. Journal of Research in Science Teaching, 39(4), 283-301.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. Learning and Instruction, 4(1), 45-69.

Vosniadou, S. (2007). Conceptual change and education. Human Development, 50, 47-54.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. Cognitive Psychology, 24, 535-585.

Wittmann, M. (1998). Making sense of how students come to an understanding of physics: An example from mechanical waves. Unpublished doctoral dissertation, University of Maryland.