

딱따구리 과제에서 초등예비 교사들의 가설 평가 지식에 대한 분석을 통한 가설 평가 능력 지수 산출식의 개발

이준기 · 이일선 · 권용주

(한국교육대학교)

Development of the Quotient Equation of the Hypothesis Evaluating Ability by Analysis of the Pre-service Elementary School Teachers' Knowledges for Evaluating Hypothesis on a Woodpecker Task

Lee, Jun-Ki · Lee, Il-Sun · Kwon, Yong-Ju

(Korea National University of Education)

ABSTRACT

The purpose of this study was to invent a quotient equation which could quantitatively evaluate individual's hypothesis evaluating ability. The equation was induced by the analysis of the classification types about hypothesis evaluation knowledges generated by 15 pre-service elementary school teachers and the construction of the quotient equation on hypothesis evaluating ability. The hypothesis evaluation task administered to subjects was dealt with the woodpecker behavior. The task was initiated by generating hypothesis on the following question: 'Why don't woodpecker have brain damage after pecking wood?' Subjects then were asked to design and perform experiments for testing hypothesis. Finally they were asked to evaluate their own hypothesis based on the collected, analyzed and interpreted data. The knowledges generated from their evaluating hypothesis were analyzed by 4 major categories (richness, type, level and accuracy). Then, a general equation which could quantitatively and systematically evaluate individual's hypothesis evaluating ability was invented by an inductive process. After combining all the categories the following quotient equation was proposed; ' $VQ = \sum(TEn \times AE_n) \times LE$ '. According to this results, woodpecker task and hypothesis evaluating ability quotient equation (VQ) which invented in this study could be applied to a practical use of measuring students' ability of scientific hypothesis evaluation.

Key words : hypothesis evaluating ability, pre-service elementary school teacher, woodpecker task, hypothesis evaluating ability quotient equation (VQ)

I. 서 론

하루가 다르게 지식과 정보의 양이 늘어나는 21세기 지식정보화 사회에서는 새로운 지식의 양산 못지않게, 생성된 지식들을 올바른 평가를 통해 정당화하는 것이 더욱 중요해지고 있다. 이러한 지식

의 정당화 과정에서 반드시 필요한 과정이 '가설 평가(hypothesis evaluation)' 과정이다(Klahr & Dunbar, 1988). 가설 평가 과정은 과학적 지식의 생성과정 중 연역적 추론 방법을 통하여 결과 지식과 결론 지식을 생성하는 단계로, 앞서 귀납과 귀추의 추론 과정을 거치면서 생성된 지식들이 절차적 지식과의

상호 작용을 통해 진술되는 마지막 과정이라고 할 수 있다(권용주 등, 2003). 그러므로 가설 평가는 가설을 포함하는 잠정적인 설명을 검증해 내는 과정으로써, 단지 잠정적인 설명에 지나지 않던 명제에 대해서 정당성을 부여하는 과정이므로 많은 창의력이 요구된다. 또한, 가설 평가는 새로운 과학지식의 생성에 있어 가장 핵심적인 과정 중의 하나로 여겨지고 있다(권용주 등, 2003; McPherson, 2001).

가설 평가에 대한 선행 연구는 많지 않은데, 권용주 등(2003)은 과학적 지식 생성 과정에 대한 연구에서 가설 평가는 연역적인 사고 과정의 일부이며, 결과 수집 → 가설 평가 → 결론 도출 과정을 거쳐 이루어진다고 보았으며, 과학자들에 대한 심층 면담을 통한 양일호 등(2006)의 연구에서는 결과 수집 → 가설 평가 및 결론 진술의 단계가 있음을 논의한 바 있다. 또한, 동물행동학자들을 연구한 이안나 등(2007)은 설명의 평가 과정을 결과 수집 → 기준-결과 비교 → 설명 판단 → 결론 도출로 보고한 바 있다. 한편, 다른 연구자들은 증거 평가라는 용어를 통해 설명 체계를 채택 혹은 기각하기 위한 증거의 취사 선택 과정에서의 개념 변화 과정이나 선개념의 영향 등을 연구하였다(박종원 등, 1993; Kuhn *et al.*, 1988; Dawson & Rowell, 1987). 하지만 이러한 연구들은 대개 갈등과 불일치 상황에 기반을 둔 개념 변화에 그 목적이 있고, 소재가 물리학이나 지구과학 내용만을 다루고 있는 경우가 대부분이었다. 또한, 상당부분 가설 평가의 사고 과정이나 이를 통한 개념 변화 연구에만 집중되어 있고, 능력의 평가 부분에 대해서는 연구 사례를 찾기가 힘들다. 때문에 학생들이 수행 평가 등을 통해서 자신만의 가설을 평가해 보고, 이를 기록·발표 하여도 정량적인 방법으로 평가하기가 어려운 실정이다.

더 나아가 지금까지 가설 평가 능력은 통합 탐구 능력의 하위 범주로서 과학적 탐구 능력 검사 도구 내의 일부 선다형 문항의 형태로 간접 검사되어왔다. 이러한 검사들은 연구자에 의해 제공되는 자료를 바탕으로 주어진 가설의 타당성을 평가하는 방식으로, 응답은 객관식 선다형답지를 통해 이루어져 왔다(예를 들어, 권재술과 김범기, 1994; 송경혜 등, 2004; 우종욱과 이항로, 1995; 정완호 등, 1993). 하지만 이와 같은 선다형 방식의 평가로는 자신이 직접 수행한 가설 평가가 아닌 다른 사람이 수행한 가설 평가가 타당하게 진행되었는지에 대해서는 부분적으로

평가가 가능하지만, 자신이 직접 창의적으로 가설 평가를 수행하는 능력에 대해서는 평가가 거의 불가능하다는 제한점이 있다. 즉, 이들 도구의 대부분이 이미 연구자에 의해 외부로부터 가설이나 결과를 제공하고 이의 타당성을 묻거나, 미비한 부분을 보완하는 형태의 문항으로 구성되어 있어서 실제로 과학자들이 하는 바와 같이 자신이 직접 생성한 가설을 자신이 얻은 데이터를 통해 평가하는 진정한 의미의 가설 평가 능력(이안나 등, 2007)을 측정하기는 어려웠으므로 이들 선행의 평가 방식은 가설 평가 능력에 대한 진정한 평가로 보기는 어렵다는 것이다. 또한 선행의 평가 도구는 과학 탐구 능력 또는 가설 평가 능력의 전반적인 측면에 알 수 있을지 모르지만, 그중 가설 평가 능력의 하위 요소들이 어떤 측면으로 얼마나 신장되었는지 같은 세부적인 부분을 파악하기는 어렵다.

따라서 가설 평가에 대한 학생들의 정성적인 지식을 정량적인 수치인 지수(quotient)로 변환시켜 산출할 수 있다면 수행 평가를 포함한 과학 탐구 능력의 평가에서 그 효용 가치는 높을 것이다. 실제로 최근 권용주 등(2007)은 학생들의 가설 생성 능력을 측정할 수 있는 ‘가설 설명 지수(HQ)’ 산출식을 고안하여 제시한 바 있다. 따라서 가설 평가 능력 역시 이러한 지수를 산출할 수 있는 적절한 식이 개발된다면 효율적이고 정량적인 측정에 도움을 줄 수 있으리라 기대된다.

이에 이 연구에서는 가설 평가 능력을 측정하는데 적합한 과제를 소개하고, 초등 예비 교사들이 가설 평가 활동을 통해 생성한 다양한 가설 평가 지식들을 유목화 하여 다차원적으로 분석한 후 이들의 가설 평가 능력을 평가할 수 있는 ‘지수산출식’을 개발하였다. 또한 개발된 ‘가설 평가 능력지수 산출식’을 적용하여 실제로 실험에 참여한 예비 교사들의 가설 평가 능력 점수를 구해 보았다.

II. 연구 방법

1. 연구의 대상

이 연구의 목적은 학생들의 가설 평가 능력을 정량적이고 객관적으로 가늠할 수 있는 지수 산출식을 개발하는 것이다. 보다 효과적인 연구 목적 달성을 위해, 중부권 소재 교원 양성 대학의 3학년에 재

학 중인 초등과학교육 전공 대학생 15명을 대상으로 하였다.

2. 과제 개발 절차

이 연구에 사용된 ‘딱따구리 과제’는 딱따구리가 단단한 나무 등치를 매우 빠른 속도로 쪼는 데도 두뇌에 손상을 입지 않는 이유에 대한 가설을 세우고, 실험을 수행한 뒤 얻은 데이터를 바탕으로 자신의 가설을 평가해 보는 과제이다. 이 과제들을 개발하기 위해서 연구자들은 먼저 10가지의 생물체를 소재로 한 가설 평가 과제들을 1차로 고안하였고, 생물교육 전문가 3인과 생물교사 4인과의 정기적인 세미나와 워크숍을 통해서 4개의 잠정 과제를 선정하였다. 그런 뒤 잠정적 과제들을 연구 대상이 아닌 학생 10명을 대상으로 사전 조사를 실시하여 적합성 여부를 분석하였다. 그런 다음, 분석 결과에 따라 최종적으로 가장 적합한 1개의 과제를 선정된 뒤, 이를 수정하여 투입하였다.

3. 가설 평가 과제

1) 딱따구리 과제의 원리

딱따구리는 딱따구리과(科)의 새에 대한 통칭이다. 삼림에 살며 날카롭고 단단한 부리로 나무에 구멍을 내어 그 속의 벌레를 잡아먹거나 집을 짓고 산다. 이들은 1초에 18~20번씩 하루에 12,000번 나무를 쪼고, 이때의 충격은 1,000 g이나 된다(Gibson, 2006). 만약 인간이 이런 상황에 처해진다면, 아기의 경우 ‘흔들린 아기 증후군(shaken baby syndrome)’에 걸리거나, 성인의 경우 무하마드 알리처럼 뇌손상을 입을 것이다. 하지만 엄청난 속도와 힘으로 나무를 쪼지만 딱따구리는 뇌손상을 입지 않는다(Kalat, 2004).

그렇다면 딱따구리가 나무를 쪼면서 두뇌를 다치지 않는 까닭은 무엇일까? 딱따구리가 뇌를 다치지 않는 까닭은 ‘머리와 부리’가 완벽한 직선을 이루는 것과, 이 각도를 유지하면서 나무에 매달려 있도록 도와주는 딱따구리의 긴 끈지깃과 독특한 발가락 구조(그림 1) 때문이기도 하다.

만약 이 각도가 약간이라도 틀어지게 된다면 뇌는 바로 손상될 것이다(Oda et al., 2006). 또한, 딱따구리는 위쪽 부리 부분에 스펀지 같은 연골 조직이 있어 충격을 흡수해 주고(May et al., 1979), 설골

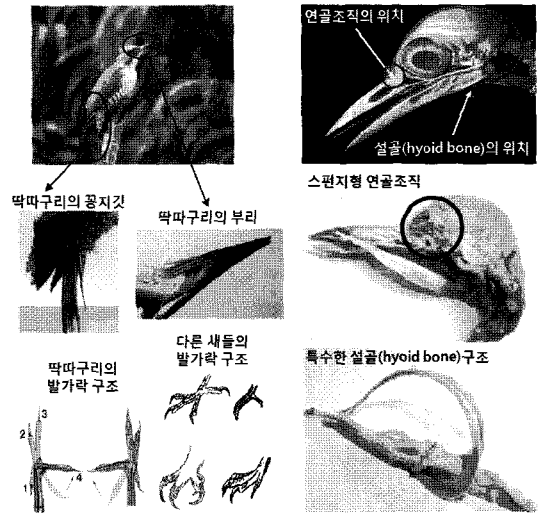


그림 1. 딱따구리의 독특한 신체 구조 (Oda et al., 2006)

(hyoid apparatus)의 독특한 구조로 인해 나무를 쪼는 순간 뇌를 반대 방향으로 당겨 충격을 덜 받는 것이다(Oda et al., 2006). 사람이 이러한 반복적 충격을 머리에 받으면 뇌가 파손되겠지만, 딱따구리는 그렇지 않은 이유를 찾아 검증하고, 이를 통해 얻은 데이터를 바탕으로 자신의 가설을 평가해 보도록 하는 것이 이 과제의 목적이라 할 수 있다.

2) 준비물 및 현상 제시 방법

딱따구리 과제의 투입을 위한 준비물 및 현상 제시 방법을 요약하면 표 1과 같다.

3) 과제 수행 과정

준비된 영상 자료를 시청한 후, 교사는 표 1에 제시된 도구들을 사용하여, 학생들과 딱따구리 되어 보기 놀이를 통해 현상을 제시한다. 이때 학생들이 자연스럽게 ‘우리(인간)는 머리가 아픈데 왜 딱따구리는 매번 큰 충격을 받으면서도 아무 이상 없이 살 수 있을까’라는 인과적인 의문이 생성되도록 안내한다. 그런 다음, 교사는 이에 대해 유사한 경험을 발표하도록 하고, 이를 통해 의문에 대한 설명이 가능하도록 잠정적인 설명인 가설을 생성하도록 한다. 최종적으로 스스로 생성한 가설을 어떻게 하면 검증할 수 있을지 설계토록 함으로써 학생들이 자신만의 실험을 능동적으로 고안하게 지도한다. 마지막으로 학생이 직접 자신이 설계한 방법으로 실험을 수행하여 얻은 결과를 활동지에 기록하게 하고,

표 1. 딱따구리 과제에 준비물 및 유의 사항

제시 단계	준비물	유의 사항
동영상 파일		① 딱따구리가 나무를 쪼는 모습을 담은 생체 동영상 제시 ② 딱따구리가 물체를 쪼는 동작과 소리가 강조된 영상을 제시
현상 제시	생일모자(고깔), 종이에 그린 과녁	① 체험 게임을 통해 ‘딱따구리 되어보기’ 활동 (고깔모자를 앞쪽으로 하여 새의 부리처럼 쓰고, 고개를 세차게 흔들며 과녁을 쪼면서, 딱따구리가 나무를 쪼는 행동을 따라 하게 한다) ② 이때 머리가 어지럽다고 할 만큼 빠른 속도와 큰 각도로 쪼게 한다.
가설 검증*	스펀지, 투명박스, 물, 날계란 (혹은 메추리알) 유리병, 날계란, 주방용 세제 또는 꿀 (점성 강한 것), 물 또는 점성 없는 액체	이는 학생의 가설이 연골 스펀지 조직에 관한 것일 경우에 한함 이는 학생의 가설이 뇌척수 액의 점성에 관한 것일 경우에 한함
가설 평가	결과 기록용 활동지	설계한 실험을 통해 얻은 결과를 수집한 후 자신의 가설을 평가

*가설 검증의 경우는 생성한 가설의 내용에 따라 다양해 질 수 있으므로 가장 빈번히 나타나는 2가지의 예시만을 제시하였음. 실제로 수행하지 않고 설계만을 알아볼 때는 필요 없는 준비물임.

이를 통해 앞서 생성한 자신의 가설이 맞는지를 평가하도록 지도한다.

4. 생성된 가설 평가 지식의 분석 및 분류들 고안절차

연구자는 15명의 예비 교사들이 의문 생성 과제를 수행하면서 생성한 15개의 가설 평가 지식들을 귀납적인 방법으로 분석하였다. 즉, 예비 교사들이 생성한 가설 평가 지식을 검토하여 공통성과 차이점에 따라 선행 연구들을 근거로 분류 기준들을 개념화하여 기술하였다. 그리고 각각의 분류 기준들을 종합하여 과학적 가설 평가를 분류할 수 있는 분류들을 고안하였고, 이 분류들에 따라 연구 대상 예비 교사들이 생성한 가설 평가 지식들을 분류하였다. 이 과정에서 연구자의 분석 신뢰도를 확보하기 위해서 생물교육전문가 2명과 현장 고등학교 생물교사 4명을 분석에 참여시켰다. 확인 분석 과정에서 6명의 분석 참여자들은 3개의 개별 분류 기준에 따라 연구자가 제시한 6개의 가설 평가 지식을 독립적으로 분류하였다. 분류 결과, 모두 6개의 분석 사례 중 1개의 사례를 제외한 모든 분류에서 동일한 분석 결과를 제시하였다. 그리고 1개의 서로 다른 분석도 연구자와의 협의 후에는 모두 동일한 유형으로 분류하였다.

5. 가설 평가 능력 지수 산출식의 개발 절차 연구에 참여한 예비 교사들이 생성한 가설 평가

지식에 대한 분류들을 기초로 유도된 가설 평가 능력 지수 산출식의 개발을 위한 R&D 절차는 그림 2와 같다.

그림 2와 같이 연구자들은 먼저 1차 일반식을 개발하여 과학교육 전문가와 현장 교사들을 대상으로 세미나를 통해 1차 일반식의 단점을 개선하였다. 이렇게 수정된 2차 일반식을 예비 교사들이 생성한 가설 평가 지식들에 적용하여 일반식의 타당성과 신뢰성을 재차 점검한 후 수정·보완하여 최종적인 일반식을 개발하였다.

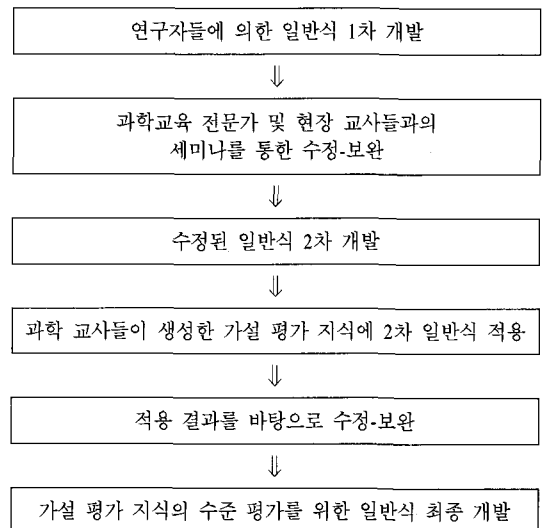


그림 2. 가설 평가 능력의 정량적 평가를 위한 일반식 개발의 절차

III. 연구 결과 및 논의

1. 생성된 가설 평가 지식 분석

딱따구리 과제에서 연구자들은 연구 대상 학생들에게 자신의 가설에 따라 이를 검증할 수 있는 실험을 설계하여 실제로 그 방법에 의해 실험을 수행하도록 하였다. 그리고 실험을 통해 얻은 결과들을 활동지에 기록한 뒤, 이것들을 토대로 자신이 생성한 가설에 대해 평가를 수행하도록 하였다. 이 과제에서 연구에 참여한 초등 예비 교사들은 모두 자신의 가설에 맞는 적절한 실험(검증 방법)을 설계한 뒤 얻은 결과물을 통해 가설을 평가하였다. 이들이 생성한 15개의 가설 평가 지식 중 특징적인 것을 예시하면 아래와 같다.

예시 1. 가설은 옳다. 부리와 머리 사이의 물렁뼈는 충격을 줄이는데 관여하는 것으로 볼 수 있다.

예시 2. 쇼바가 잘 되어있는 고급 승용차를 타면 승차감이 매우 좋다. 또한, 물리 시간에 배운 충격량에 관한 공식 $I = F \times t$ 를 생각해 보면, 힘은 시간에 반비례하므로, 부리와 두개골 사이의 연골은 시간을 증가시켜 충격을 줄일 것이다. 따라서 가설은 옳다.

예시 3. 우리가 세운 가설은 틀린 것으로 평가되었다. 이 결과에서는 물을 채운 병의 경우 3회와 5회 실험에서 깨지는 메추리알이 발생하였을 뿐만 아니라, 뇌척수 액은 사람에게도 있다. 그러므로 뇌척수액과 같은 유체는 딱따구리의 충격 흡수 방식과 무관하다고 할 수 있다.

예시 4. 이 실험에서 얻은 결과로 볼 때에, 가설은 옳다. 가설이 맞으려면, 스펀지 조직의 두께가 메추리알이 깨지지 않을 정도의 임계 두께 이상이 되어야 한다. 이 결과에서는 부착된 스펀지의 두께에 따라 메추리알의 파손 정도가 감소되다가 1.5 cm 이상부터는 손상이 전혀 없었다. 이를 통해 딱따구리의 두개골과 윗부리 사이에 있는 스펀지 조직은 충격을 흡수하여 두뇌 손상을 방지하는 역할을 하고 있음을 알 수 있다.

1) 가설 평가를 위한 근거의 유형

위의 예시들은 연구자가 예비 교사들에게 제시한 과제를 해결하면서, 그들이 자신들의 가설을 평가한 내용의 일부이다. 가설의 평가는 탐구의 단계상 정당화(justification)의 단계에 속하므로(권용주

등, 2003) 무엇을 통해 정당화하는가에 대한 문제는 매우 중요하다. 예시 1의 경우, 자신이 생성한 가설이 옳다고 평가는 하고 있으나, 이에 대한 판단 근거(evidence)를 제시하지 못하고 있다. 평가는 반드시 평가 근거와의 상호 작용 속에 이루어지므로(Kuhn et al., 1988) 단순히 진위만을 언급한 예시 1의 경우는 올바른 평가가 이루어지지 못한 것으로 볼 수 있다.

한편, 예시 2, 3, 4의 경우는 서로 다른 근거를 제시하면서 가설의 진위 여부를 평가하고 있다. 예시 2는 가설이 옳다는 것에 대한 근거로 고급 승용차에 승차했던 자신의 경험과 물리학의 충격량 공식을 제시하였다. 여기서 가설을 평가하는데 사용된 근거들은 자신의 실험 결과가 아닌 예전의 경험이나 암기된 이론과 같은 외부로부터 도입된 것임을 알 수 있다. 이와 같은 방법으로 평가를 할 경우(이론 기반 평가; theory-based evaluation), 실험 결과는 의미가 없어지며 선입견(bias)에 의한 잘못된 평가로 이어져 그릇된 결론을 내릴 가능성이 높다(Koslowski, 1996; Kuhn et al., 1988).

예시 4의 경우, 자신의 가설이 옳다는 판단에 대한 근거를 자신의 실험 결과로부터 얻어진 규칙성 지식만을 통해 제시하고 있다. 이처럼 결과 혹은 결과들 간의 관계만으로 가설을 평가할 경우(증거 기반 평가; evidence-based evaluation)는 예시 2에서보다 정량적이고 과학적인 평가가 이루어질 수 있다(Koslowski, 1996; Kuhn et al., 1988). 위의 예시에서 가설을 평가하기 위한 근거의 유형은 크게 1) 근거 제시를 못하는 경우, 2) 근거를 제시하고 출처가 결과 외적인 경우 그리고 3) 근거를 제시하고 출처가 결과 내적인 경우로 정리해 볼 수 있다. 예시 3의 경우는 2)와 3)의 유형이 혼합된 형태로 볼 수 있다.

Amsel과 Brock(1996)은 증거 평가 기능의 발달적 차이를 알아보기 위한 실험을 수행했는데, 아동과 성인을 대상으로 식물 그림 자료들을 제시하고 가설을 평가하는데 사용하는 증거의 출처를 연구를 수행했다. 이 연구에서 아동들은 성인보다 사전 신념과 결론 자료에 강한 영향을 받고 있었다. 또한, 성인들은 결과로부터 생성된 규칙성 지식(식물의 건강과 변수들 사이의 관계)을 통해 판단을 정당화하는 반면, 아동들은 사전 신념이나 경험을 근거로 하거나 혹은 사전 신념에 일치하는 일부 변수만을 통해 판단하려는 경향을 보였다. 이러한 내용을 종합하여 가설 평가의 유형에 따른 점수화가 용이하도록

록 유목화해 보면 표 2와 같다.

2) 가설 평가를 위한 근거의 유창성

많은 선행 연구자들은 가설 평가가 연역적 사고 과정의 일부임을 밝히고 있다(권용주 등, 2003; Lawson, 1995). 연역적 추론 과정은 주어진 전제들로부터 타당한 결론을 이끌어 내는 사고 과정을 의미한다. 이는 외부의 개입 없이 전제(premise)들 내의 논리적 사고를 통해, 이의 통합으로만 유도되는 닫힌 추론이라는 점(close-ended)에서 외부로 열려진(open-ended) 귀납적 추론과는 차별성을 갖는다(Knauff et al., 2006). 즉, 가설 평가 과정에서 평가 기준과 수집된 결과를 비교/판단한 뒤 이들을 통합하여 결론 지식에 이르게 되는데, 이때 가설 평가 기준은 연역 추론에서의 대전제 역할을, 수집된 실험 결과들은 전제의 역할을 하게 된다(권용주 등, 2003).

다른 일반적인 연역 추론에서처럼 가설 평가에서도, 전제의 수가 많을수록 이의 통합을 통해 생성되는 결론 지식의 질은 보다 우수해질 것이다. 예를 들어, 아무런 근거 없이 평가를 하는 예시 1보다는 2가지 근거를 드는 예시 2와 3에서 얻은 결론 지식의 신빙성이 더 높을 것이다. 이러한 내용을 종합하여 가설 평가의 유창성을 점수화가 용이하도록 유목화해보면 표 3과 같다.

3) 가설 평가의 정확성

가설을 평가하는 데에 있어서 유형과는 다른 또

표 2. 가설 평가의 근거 유형 배점 기준

구분	배점
평가의 근거가 없는 것	0
평가의 근거가 있고, 근거의 출처가 결과 외적인 것	1
평가의 근거가 있고, 출처가 결과 내적인 것	2

표 3. 가설 평가의 근거 유창성 배점 기준

구분	배점
1개의 근거만을 사용하여 가설을 평가한 경우	1
2개의 근거를 사용하여 가설을 평가한 경우	2
...	...
n개의 근거를 사용하여 가설을 평가한 경우	n

하나의 특성이 있다. 이것은 바로 평가를 위해 제시된 근거가 결과의 정보들을 얼마나 담아내고 있는가이다. 선행 연구들에서도 가설 평가에서의 근거제시에 대한 논의가 있어 왔는데, 박종원 등(1993)의 연구에 의하면 학생들은 선개념의 영향으로 인해 무관련 증거를 제시하거나, 결과 중의 일부만을 증거로 제시하는 경우가 있다고 하였다.

전문적인 과학자들조차도 이런 행위를 할 때가 있는데, Holton(1978)은 밀리컨이 1913년 시행한 실험의 실제 데이터를 그의 랩노트를 분석하여 조사해 본 결과, 140개의 데이터 중 80개의 데이터를 폐기하고 60개의 데이터만으로 가설을 평가한 것으로 나타났다. 폐기의 이유는 단지 뭐가 잘못되었을 것이라는 막연한 의심 때문이었다고 한다. 위의 예에서 볼 수 있듯이 '결과의 내용을 얼마나 충분히 반영하여 평가하는가'라는 내용은 해석 오류를 막고 올바른 가설 평가를 이끌어주는데 중요한 요소이다.

한편, 결과로부터 규칙을 끌어내었다고 하더라도 이에 대한 판단을 내릴 때, 실험 설계 단계에서 생성한 '가설 평가 기준'과의 비교가 이루어지지 않는다면 올바른 가설 평가가 이루어지지 않을 것이다. 이는 실제의 과학자들에 대한 연구의 예에서 많이 찾아 볼 수 있는데, 이안나 등(2007)의 연구에서도 과학자들은 가설 평가시 반드시 기준-결과 비교 과정을 거치는 것으로 나타났다.

예시 2를 살펴보면, 가설의 정당화를 위해 제시한 근거들에 실험을 통해 얻은 결과가 전혀 반영되어 있지 않고, 평가 기준과의 비교 역시 나타나 있지 않다. 예시 3의 경우에도 이 실험을 수행한 학생 C가 수집한 결과로는 시행 3회차와 5회차에서 메추리알이 파손된 것은 알 수 있으나, 사람의 머리에 뇌척수 액이 있다는 것은 결과 내용과는 상관이 없다. 뿐만 아니라 예시 3의 경우 결과에 나타난 모든 변인들의 관계를 고려한 지식이 아닌 단편적인 부분(3회차와 5회차)만을 제시함으로써, 불완전한 정보를 바탕으로 평가가 이루어졌다. 평가 기준과의 비교는 예시 2와 3 모두에서 이루어지지 않았다. 마지막으로 예시 4의 경우는 근거로 제시한 '스펀지의 두께에 따른 메추리알의 파손 정도'라는 규칙성 지식이 실험 결과를 충분히 반영하였으며, 평가 기준과의 비교도 나타내었다. 위의 예시들에서 가설 평가의 정확성은 결과 내용의 반영 정도에 따라, 1) 근거가 결과를 반영하지 못하는 경우, 2) 근거가 결

과를 부분적으로 반영하는 경우, 그리고 3) 근거가 결과를 전체적으로 반영하는 경우로, 평가 기준과의 비교 여부에 따라 1) 평가 기준과의 비교가 없는 경우, 2) 평가 기준과의 비교가 나타난 경우로 정리해 볼 수 있다.

이해를 돕기 위하여, 모든 요소가 잘 나타난 예시 4를 통해 평가 기준, 가설 평가의 근거(결과), 결론이 어떻게 나타났는지 명시해 보면 아래와 같다.

- 가설이 맞으려면, 스펀지 조직의 두께가 메추리알이 깨지지 않을 정도의 임계두께 이상이 되어야 한다. [가설 평가 기준-실험을 설계하면서 예측을 통해 설정해 둔 기준]
- 이 결과에서는 부착된 스펀지의 두께에 따라 메추리알의 파손 정도가 감소되다가 1.5cm 이상부터는 손상이 전혀 없었다. [가설 평가 근거-실험 실행 수행을 통해 얻은 결과 혹은 외부 근거]
- 이를 통해 딱따구리의 두개골과 윗부리사이에 있는 스펀지 조직은 충격을 흡수하여 두뇌 손상을 방지하는 역할을 하고 있음을 알 수 있다. [결론-가설의 평가를 통한 일반화]

이러한 내용을 종합하여 정확성에 따른 점수가 용이하도록 유목화해 보면 표 4와 같다.

4) 가설 평가의 수준

많은 과학적 사고 과정에 관한 선행 연구들은 ‘가설 검증(hypothesis testing)’과는 달리 ‘가설 평가(hypothesis evaluation)’ 과정은 가설을 단순히 채택할 것인가 기각할 것인가를 판단함과 아울러 이에 대한 일반 진술인 결론을 도출하는 과정으로 완결

표 4. 가설 평가의 정확성 배점 기준

구분	배점
근거가 결과를 반영하지 못함	1
평가 기준 비교 없음	2
평가 기준 비교	3
근거가 결과를 부분적으로 반영	4
평가 기준 비교 없음	5
평가 기준 비교	6
근거가 결과를 전체적으로 반영	

됨을 보고하고 있다(권용주 등, 2003; 양일호 등, 2006; 이안나 등, 2007). 특히, 이러한 결론 도출은 설명의 범위를 확대하는 일반화의 과정으로 차후의 연구에 순환적으로 영향을 주기 때문에 선행 연구들에서 공통적으로 보고되었으며, 중요한 위치를 차지하고 있다.

위에 제시된 가설 평가 지식들 중 예시 1, 2의 경우는 선행 연구에 의해 제시된 과학적 사고의 단계 중 단순히 설명(가설)판단까지만 수행한 반면, 예시 3, 4는 가설의 진위를 판단함과 아울러 이에 대한 결론을 도출해 내고 있다. 선행 연구 결과들에 비추어 보았을 때, 예시 1, 2의 경우보다는 3, 4의 경우가 보다 가설을 평가하는 데에 있어, 보다 깊은 수준의 사고에 까지 도달한 것으로 볼 수 있다. 따라서 수준에 따라 가설 평가는 1) 결론을 도출하지 못한 경우와 2) 결론을 도출한 경우로 나누어 볼 수 있다. 이러한 내용을 종합하여 가설 평가의 수준에 따른 점수화가 용이하도록 유목화해 보면 표 5와 같다.

이상의 논의에서 나타난 바와 같이 과학적 가설 평가는 유형, 수준, 유창성, 정확성이라는 항목에 의해서 분류될 수 있다. 이를 정리하여 제시하면 그림 3과 같다.

그림 3에 제시된 분류 기준에서 유형을 가로 방향으로, 유창성을 세로 방향으로 배열함으로써 2차

표 5. 가설 평가의 수준 배점기준

구분	배점
결론을 도출하지 못함	1
결론을 도출 함	2

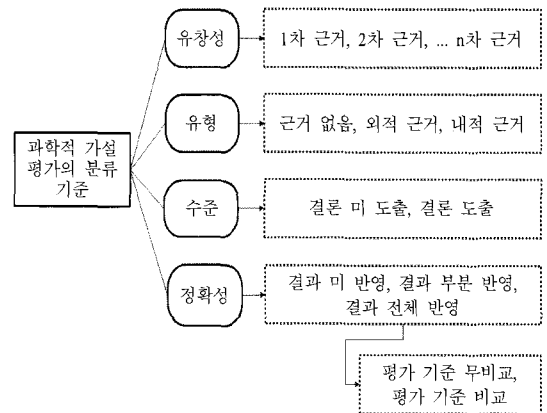


그림 3. 과학적 가설 평가의 분류 기준

원 평면 위에 가설 평가의 분류 기준들을 나타낼 수 있다. 또한, 가로축의 가설 평가의 유형을 수준에 따라 재분류하고, 세로축의 유창성에 따라 나타나는 평가의 근거를 그 정확성에 따라 재분류하면 표 6과 같은 가설 평가의 분류틀로 나타낼 수 있다.

예시 3의 가설 평가 사례를 분류틀에 적용해 보면 표 7과 같다. 예시 3은 2개의 근거를 사용하여 가설을 평가했으며, 근거 1은 내적 근거이며, 근거 2는 외적 근거이다. 또한, 평가 기준과의 비교는 이루어지지 않았고, 결론을 도출하였다.

2. 가설 평가 능력 지수(VQ) 산출식의 고안

위에서 분석한 바와 같이 연구에 참여한 초등 예비 교사들이 생성한 가설 평가 지식은 크게 평가의 유형, 평가의 정확성, 평가의 수준, 평가 근거의 유

창성이라는 4가지 측면에서 평가될 수 있을 것이다. 이상의 내용들을 근거로 개인의 가설 평가 능력을 평가할 수 있는 지수를 산출하는 방법을 식으로 나타내면 그림 4와 같다. 이때, 각 단위항의 연산 방식은 개발된 분류틀을 토대로 하여 이에 대한 확률적 동시성 여부에 의해 곱의 법칙과 합의 법칙에 준하여 유도된 것이다.

예비 교사들의 다양한 가설 평가 지식들을 귀납적으로 분석하여 얻어진 가설 평가의 분석틀(표 6)과 가설 평가 능력 지수 산출식(그림 4)을 예비 교사 B의 가설 평가 지식인 예시 2와 예비 교사 D의 가설 평가 지식인 예시 4의 사례에 실제로 적용하여 가설 평가 능력지수를 구해 보면 다음과 같다.

예비 교사 B의 가설 평가 지식인 예시 2를 분석해 보면 평가의 근거는 제시하였으나, 경험이나 암

$$VQ = \sum(TEn \times AE_n) \times LE$$

가설 평가 능력 지수 = {첫 번째 평가 근거 (평가 근거의 유형 × 평가 근거의 정확성) + ... + n 번째 평가 근거} × 가설 평가의 수준

(VQ: hypothesis evaluation ability quotient, TE: type of hypothesis evaluation, LE: level of hypothesis evaluation, AE: accuracy of hypothesis evaluation, n: nth evidence for evaluation)

그림 4. 가설 평가 능력지수 산출식

표 6. 가설 평가 지식 분류틀

유창성	유형		근거 없음		외적 근거		내적 근거	
			결론 미도출	결론 도출	결론 미도출	결론 도출	결론 미도출	결론 도출
1차 근거	결과 미반영	평가 기준 무비교						
		평가 기준 비교						
	결과 부분 반영	평가 기준 무비교						
		평가 기준 비교						
	결과 전체 반영	평가 기준 무비교						
		평가 기준 비교						
...								
n차 근거	결과 미반영	평가 기준 무비교						
		평가 기준 비교						
	결과 부분 반영	평가 기준 무비교						
		평가 기준 비교						
	결과 전체 반영	평가 기준 무비교						
		평가 기준 비교						
계								

기된 지식에 의한 외부 근거이므로 두 개의 근거 모두 유형값은 1, 각 근거의 정확성은 실험 결과를 반영하고 있지 않고 평가 기준과의 비교가 없으므로 모두 1이다. 수준은 결론 도출을 하지 못하였으므로 1이다.

$$VQ_{\text{예비 교사B}} = \{(1 \times 1) + (1 \times 1)\} \times 1 = 2$$

예비 교사 D의 가설 평가 지식인 예시 4를 분석해 보면, 평가의 근거를 제시하였고 자신의 실험으로부터 유래된 내적근거이므로 유형값은 2, 정확성은 실험 결과를 전체를 반영하고 있고, 평가 기준과의 비교가 있으므로 6이다. 수준은 결론까지 도출하였으므로 2가 된다.

$$VQ_{\text{예비 교사D}} = \{(2 \times 6) + (2 \times 6)\} \times 2 = 48$$

두 학생 모두 2가지 근거를 제시하면서, 가설을 평가하였지만 예비 교사 B와 D의 가설 평가 능력지수 차이는 크게 나타났다. 이와 같이 유도된 ‘가설 평가 능력 지수’ 산출식을 적용하여 연구에 참여한 초등 예비 교사들이 설계한 가설 평가 지식을 분석

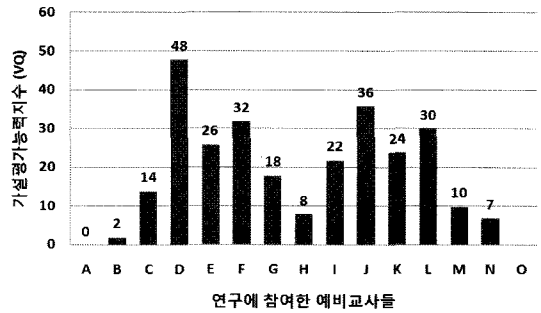


그림 5. 연구에 참여한 초등 예비 교사들에 대한 가설 평가 능력지수 적용의 예

하는 데에 적용해 보면 그림 5와 같은 결과를 얻을 수 있다.

IV. 결론 및 교육적 적용

이 연구에서는 초등 예비 교사들에게 생명 현상과 관련된 과제를 제시하고 그들이 스스로 생성한 가설에 대한 평가 결과를 기록하게 하고, 이 가설 평가 지식들을 귀납적으로 분석하여 특성을 살펴보았다. 또한, 이들 특성을 종합하여 개인의 가설 평가

표 7. 예시 3의 가설 평가에 대한 가설 평가 지식 분류틀

유창성	유형		근거 없음		외적 근거		내적 근거	
			결론 미도출	결론 도출	결론 미도출	결론 도출	결론 미도출	결론 도출
1차 근거	결과 미반영	평가 기준 무비교						
		평가 기준 비교						
	결과 부분반영	평가 기준 무비교						○
		평가 기준 비교						
	결과 전체반영	평가 기준 무비교						
		평가 기준 비교						
2차 근거	결과 미반영	평가 기준 무비교						○
		평가 기준 비교						
	결과 부분반영	평가 기준 무비교						
		평가 기준 비교						
	결과 전체반영	평가 기준 무비교						
		평가 기준 비교						
계			0	0	0	1	0	1

능력을 정량적으로 평가할 수 있는 산출식을 고안하였다. 이 연구 결과를 통해 다음과 같은 결론을 얻을 수 있다.

첫째, 이 연구에서는 15명의 예비 교사들이 생성한 15개의 가설 검증 결과 진술문을 유형, 수준, 정확성, 유창성이라는 측면에서 분류해 볼 수 있었다. 먼저 가설 평가 지식의 유형 측면에서는 평가 근거가 없는 유형, 평가 근거가 외적인 유형, 평가 근거가 내적인 유형의 3가지로 분류되었다. 가설 평가의 수준면에서는 결론 미도출형과 결론 도출형으로 나누어 볼 수 있었다. 정확성 면에서는 근거가 결과를 반영하지 못하는 경우, 근거가 결과를 부분적으로 반영하는 경우 그리고 근거가 결과를 전체적으로 반영하는 경우로 분류할 수 있었으며, 이들을 다시 평가 기준과의 비교가 없는 경우와 평가 기준과의 비교가 나타난 경우로 나누어 볼 수 있었다. 이러한 결과의 종합을 통해 얻은 가설 평가의 분류들은 다른 예비 교사들이나 학생들의 가설 평가 지식을 파악하고 평가하는 데에 유용하게 적용될 수 있을 것으로 기대된다.

둘째, 이 연구를 통해 예비 교사들이 생성한 가설 평가 지식들에서 나타나는 중요한 요소(유형, 수준, 근거의 유창성)들을 근거로 개발된 가설 평가 능력 지수(VQ) 산출식은 손쉽게 학생들의 가설 평가 능력을 측정할 수 있는 새로운 방법으로 적용될 수 있을 것이다. 지수 산출식의 적용을 통해 한 개인의 능력 신장도 비교 및 집단 간 비교 혹은 집단 내 개인 차의 비교 등 학생들의 질적 자료에 대한 다양한 방식의 정량적 평가가 가능해질 것으로 전망된다.

이 연구의 한계와 의의를 살펴보면 다음과 같다. 본 연구에서 제시하고 있는 산출식은 적은 인원의 예비 교사들을 대상으로 단일 과제만을 통해 구성된 수식인 만큼 다양한 변인들에 대해 고려되지 못한 한계점을 지닌다. 산출식이 보다 타당하고 신뢰로 우려면 공인된 검사 도구와의 비교나 대단위 집단에 대한 적용을 통한 추후 연구가 이루어져야 할 것이다. 또한, 논문에서 사용되고 있는 0~4 혹은 1~3과 같은 등간 척도는 절대값을 지니는 계수(coefficient)가 아니며, 식의 활용을 보여주는 하나의 예시일 뿐이다. 각각에 대한 하위 사고 단위와 그에 따른 인지적 부담을 계측하여 절대값을 지니는 계수를 추출 및 확인하기 위한 후속 연구가 필요하다.

다만, 이 연구는 표준화 검사를 통한 도구 개발

연구라기보다는 가설 평가라는 질적 자료를 정량적인 자료로 변화시킬 수 있는 산출식을 유도하기 위한 시도라는 데에 그 의의가 있다. 이와 같은 방법을 보다 정교화 하여 활용한다면, 학생들의 수행평가나 실험보고서와 같은 정성적이고 질적인 자료들에 대하여 정량적으로 가능해 볼 수 있는 기회가 보다 많아질 것이다.

V. 참고문헌

- 권용주, 이준기, 정진수(2007). 과학고등학교 학생들이 전분 실험에서 생성한 생물학 가설의 설명경향 분석. *중등교육연구*, 55(1), 275-298.
- 권용주, 정진수, 박윤복, 강민정(2003). 선언적 과학 지식의 생성 과정에 대한 과학철학적 연구 - 귀납적, 귀추적, 연역적 과정을 중심으로 -. *한국과학교육학회지*, 23(3), 215-228.
- 권재술, 김범기(1994). 초·중학생들의 과학 탐구능력 측정도구의 개발. *한국과학교육학회지*, 14(3), 251-264.
- 박종원, 장병기, 윤혜경, 박승재(1993). 중학생들의 빛과 그림자에 대한 증거평가. *한국과학교육학회지*, 13(2), 135-145.
- 송경례, 이항로, 임철헌(2004). 초등학교 고학년 학생의 과학 탐구능력 측정을 위한 평가도구 개발. *한국과학교육학회지*, 24(6), 1245-1255.
- 양일호, 정진수, 권용주, 정진우, 허명, 오창호(2006). 과학자의 과학지식 생성 과정에 대한 심층 면담 연구. *한국과학교육학회지*, 26(1), 88-98.
- 우중옥, 이항로(1995). 고등학생의 지구과학 탐구능력 측정을 위한 평가도구 개발. *한국과학교육학회지*, 15(1), 92-103.
- 이안나, 권용주, 정진수, 양일호(2007). 동물행동학자의 연구 활동에서 나타나는 연구 단계, 사고 과정, 행동 양식 및 생성 지식에 관한 연구. *한국생물교육학회지*, 35(3), 361-373.
- 정완호, 허명, 은경용(1993). 국민학생의 과학 탐구능력 측정을 위한 평가도구 개발. *한국과학교육학회지*, 13(1), 80-91.
- Amsel, E. & Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, 11, 523-550.
- Dawson, C. & Rowell, J. (1987). The use of data in problem solving: The ways, when and wherefores. *Research in Science Education*, 17, 1-10.
- Gibson, L. J. (2006). Woodpecker pecking: how woodpeckers avoid brain injury. *Journal of Zoology* 270(3), 462 - 465.
- Holton, G. (1978). Subelectrons, presuppositions, and the Mil-

- likan-Ehrenhaft dispute, In Holton, G. *The scientific imagination: Case studies*. Cambridge, Cambridge University Press.
- Kalat, J. W. (2004). *Biological Psychology, 8th Ed.* Belmont, CA: Wadsworth Publishing.
- Klahr, D. & Dunbar, K. N. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12(1)*, 1-48.
- Knauff, M., Famgmeier, T., Ruff, C. C. & Slotsky, V. (2006). fMRI-evidence for a three-stage-model of deductive reasoning. *Journal of Cognitive Neuroscience, 18*, 320-334.
- Koslowski, B. (1996). *Theory and evidence: the development of scientific reasoning*. Cambridge, MA: The MIT Press.
- Kuhn, D., Amsel, E. & O'Loughlin, M. (1988). *The development of scientific thinking skills*. New York, NY: Academic Press, Inc.
- Lawson, A. E. (1995). *Science teaching and the development of thinking*. Belmont, CA: Wadsworth Publishing Company.
- May, P. R., Fuster, J. M., Haber, J. & Hirschman, A. (1979). Woodpecker drilling behavior. An endorsement of the rotational theory of impact brain injury. *Archives of Neurology, 36*, 370-373.
- McPherson, G. R. (2001). Teaching & learning the scientific method. *The American Biology Teacher, 63(4)*, 242-245.
- Oda, J., Sakamoto, J. & Sakano, K. (2006). Mechanical evaluation of the skeletal structure and tissue of the Woodpecker and Its Shock Absorbing System. *JSME International Journal Series A, 49*, 390-396.