

# 과학적 논증과정 평가를 위한 루브릭 개발

양일호 · 이효정 · 이효녕<sup>1\*</sup> · 조현준<sup>2</sup>

한국교원대학교 · <sup>1</sup>경북대학교 · <sup>2</sup>카이스트 과학영재교육연구원

## The Development of Rubrics to Assess Scientific Argumentation

Il-Ho Yang · Hyo-Jeong Lee · Hyonyong Lee<sup>1\*</sup> · Hyunjun Cho<sup>2</sup>

Korea National University of Education · <sup>1</sup>Kyungpook National University

<sup>2</sup>KAIST Institute for Gifted Students

**Abstract:** The purpose of this study was to develop a rubric for assessing students' scientific argumentation. Through the analysis of relevant literature related to argument in science education for developing rubric, the procedure in development and the category in assessment for rubric were elicited. According to the general procedure in developing rubric, the standard for evaluating the argumentation derived three categories such as a form, contents, and attitude. The form category was further segmented into sub-functions composition, claim, ground, and conclusion in the whole. The category for contents was segmented into sub-functions understanding, credibility, and inference. And the category for attitude was set to sub-functions participatory level and openness. The standard for evaluating sub-functions in each of the categories formed in this way was minutely suggested with five stages. The rubric, which was developed on the basis of literature, was inspected through a regular seminar in one expert in science education and fellow researchers. The rubric, which was developed in the early days, was again modified by being verified on problem and improvement matter after being entrusted to four experts in scientific education. And, the finally-completed rubric indicated to be high with 0.96 in the content validity index by being verified the validity by the four experts in science education. The developed rubric will lead to being able to increase the understanding about demonstration in students, and to being available for being utilized as the criteria for developing the argumentation process program and for evaluating the argumentation activity

**Key words:** scientific argumentation, assessment, rubric, analysis category

### I. 연구의 필요성 및 목적

과학에서 논증과정(argumentation)<sup>1)</sup>은 과학자들의 연구에서 중심적인 역할을 담당하고 있으며 (Newton et al., 1999), 이는 논증과정을 통하여 새로운 이론들이 힘을 얻게 됨으로써 과학적 지식이 발전해 나가기 때문이다(Zohar & Nemet, 2002). 과학자들의 논증과정은 인과적인 의문에 대한 답을 제시할 때 두 가지 이상의 경쟁 이론을 제기함에 따라 발생하는 것으로(Kuhn, 1993), 주로 자신이 지지하는 이론을 뒷받침할 수 있도록 증거를 분석하거나, 과학자 커뮤니티를 통해 다른 의견의 지지자들과 의견을

나누는 활동으로 정의할 수 있다(Hogan & Maglienti, 2001). 따라서 과학교육에서도 “학생들을 과학자처럼”이라는 슬로건 아래 학생들의 과학학습에서 논증과정이 중심적인 활동이 되어야 하며 (Lawson, 2003), 특히 과학적 논증과정을 통한 탐구 활동은 과학적 지식을 구성하는 능력을 향상시키기 위한 핵심적인 도구라 할 수 있다.

학생들은 논증과정을 통해 과학적 명제에 대하여 학습할 뿐 아니라 과학적 지식의 사회적 구성 방법에 대해서도 알 수 있다(Bell & Linn, 2000). 또한 인지 심리학자들에 따르면 논증과정을 통해 개념들 간의 관계와 선지식(prior knowledge)을 알 수 있으므로

\*교신저자: 이효녕(hlee@knu.ac.kr)

\*\*2009.01.15(접수) 2009.02.03(1심통과) 2009.02.12(2심통과) 2009.02.13(최종통과)

\*\*\*이 논문은 2007년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2007-327-B00597).

<sup>1)</sup> argumentation의 번역은 ‘논증과정’ ‘논변활동’ 등으로 연구자마다 다르다. 논증과정이 비교적 널리 사용되고 있고 그 대안이 마땅하지 않다는 점에서 일단 논증과정이라는 용어를 사용하기로 한다.

정보를 인출하는 다양한 경로를 만들어내어 효과적인 기억과 학습을 가능하게 한다(O'Donnell & King, 1998). 논증과정은 현재 가지고 있는 개념의 문제점에 관심을 가지게 하여(Chan, 2001) 대안적인 개념을 적용하고 사고하는 기능을 촉진할 수 있다(Reiser *et al.*, 2001).

그러나 외국의 과학교육 현장에서는 과학교육자들이 과학과 과학학습에서 논증과정의 중요성을 강조했음에도 불구하고 전체 과학 수업의 1% 정도 밖에 논증과정이 활용되지 못하고 있으며(Osborne *et al.*, 2004), 과학 탐구 활동이 일어나는 경우에도 논의(debate)나 논증(argument)은 좀처럼 사용되지 않았다(Newton *et al.*, 1999). 국내의 경우도 외국과 크게 다르지 않다. 예를 들어, 이범홍(1998)은 277명의 중학교 과학교사 중 70% 이상의 교사들이 45분 수업 중 6분 이하로 토론(discussion)을 사용한다고 보고하였다.<sup>2)</sup>

그렇다면 왜 과학수업에서 논증과정이 잘 활용되지 못하는 것일까? 이에 대한 원인은 교사들이 성공적인 논증과정 지도 방법을 제대로 이해하고 있지 못하며 그 결과 논증과정을 시도하기 위한 자신감이 결여되어 있기 때문이고(Driver *et al.*, 2000), 교육과정상의 문제, 교사들의 인식 부족도 그 원인 중 하나라고 하였다(Osborne *et al.*, 2004). 또한 좋은 논증과정에 대한 학생들의 인식 연구 결과, 논증에 관한 지식이 절대적으로 부족하다는 것을 밝혀냈다(Simon *et al.*, 2006). 따라서 학교 교육현장에서 논증과정이 효과적으로 이루어지기 위해서는 첫째, 논증의 질을 개선하기 위한 교육적 장치를 제공하고 둘째, 교사들에게 논증과정상에서 무엇을 찾아야 할 것인지 어떻게 학생들에게 논증을 안내할 것인지를 제공하고 셋째, 학생들이 논증의 본성과 구조에 대해 인식하는 것을 돕고 넷째, 학생이 논증과정에 대하여 어느 정도 진보를 이루었는지 알아내기 위한 평가를 제공해야 한다는 등의 4가지 목표를 달성해야 하며, 이를 위해 논증과정을 평가하고 분석할 수 있는 도구의 개발이 필요하다(Driver *et al.*, 2000).

현재까지 논증과정에 관해 연구해온 학자들은 Toulmin(1958)의 논증구조 패턴(Toulmin's Argumentation Patterns, TAP)을 분석 도구로 많

이 사용하였다. 그러나 이는 독백적 형태의 구조를 띄고 있어 양방향 논증과정을 가르치는데 시사점을 주기 어려우며(Hogan & Maglienti, 2001; Zohar & Nemet, 2002), 실제 학생들의 논증과정에 TAP를 적용하여 요소를 찾는 것이 힘들고(Naylor *et al.*, 2007), 요소를 찾는 것만으로 논증과정의 질과 설득력을 평가하는 것은 어렵기 때문에 교실 상황의 논증과정을 평가하기 위한 새로운 틀이 필요하다(Yore & Treagust, 2006). 이에 따라 Toulmin 이후 많은 학자들은 보다 실용적이고, 논증과정의 산출물 보다는 상호 과정으로서의 논증과정을 평가하기 위한 모델을 만들고자 많은 관점들을 제시하였는데(Sadler & Fowler, 2006), 각 TAP 요소별로 4 척도의 채점기준을 제시하였고(McCann, 1989), TAP의 여러 요소들 중에서 특히 반증 요소의 유무를 중시하는 관점을 제시하여 5수준의 평가 틀을 제시하기도 하였다(Osborne, 2004). 또한 Toulmin의 TAP가 상호 과정적 논증과정 평가에 어려움이 있다는 한계를 극복하기 위해 학생들의 논증과정에서 그룹 활동을 강조하는 Downing model을 개발하였고(Andrews *et al.*, 1993), 내용 일반적인 평가 관점에서 벗어나기 위하여 논증과정 속에 포함된 과학적 지식수준을 평가하고자 했다(Sandoval & Millwood, 2005; Takao *et al.*, 2002). 학자에 따라 초점을 맞추고 있는 부분이 각기 달라서 논증과정 전체를 평가할 때 고려해야 할 많은 관점들을 모두 반영하고 있는 틀을 찾기 힘들다. 따라서 이러한 연구 결과들을 실제 학생들의 과학적 논증과정 활동에 반영하여 보다 나은 교수학습을 실행하기 위해서는 학자들의 논증과정 평가에 관한 관점들을 종합하여 학생들의 논증과정을 평가하는데 사용할 수 있는 틀을 개발하는 연구가 필요하다.

개발된 평가도구의 적용을 통해 시간이 지남에 따라 학생들의 논증과정 능력이 향상되었는지를 알 수 있고 교사들이 학생들의 논증과정을 형식적으로도 이해하게 할 수 있어서 학생들의 논증과정을 보다 활발하게 일어나도록 촉진할 수 있다(Simon *et al.*, 2006). 뿐만 아니라 평가도구는 교사가 학생의 학습을 측정할 수 있도록 수행의 기준을 제시하기 때문에 학생과 교사가 이 기준을 적용함으로써 보다 효과적인 논증과정이 수업에서 나타날 수 있다(Astin,

<sup>2)</sup> 이 연구에서는 'argument(논증)'는 내용에 기여하는 주장, 근거, 보강, 보강 등의 실제적 내용들이며, 'argumentation(논증과정)'은 논증을 통해 해당되는 요소들을 집합적으로 처리하는 과정을 말한다. 학생들에게 토론(discussion)과 논의(debate)를 필요로 하는 과제의 제공을 통해 교사들이 학생들을 논증과정(argumentation)을 통해서 논증(argument)을 구성하는 경험을 갖게 할 수 있다.

1993; Banta *et al.*, 1996).

전통적인 평가에서 점수나 등급이 학생의 수행에 대해 구체적인 정보를 제공하지 못한 반면에 루브릭은 기술적인 평가 항목에 근거해 그 수준과 장단점을 정확하게 특징화시킬 수 있는 장점이 있고, 최상의 단계에서 최하의 단계까지 무엇이 어떻게 평가될 수 있는지를 세분화하여 기술하기 때문에 학생의 수행이 어느 수준에 해당하는가를 결정해 현재 상태, 발달의 가능성과 방향성을 제공할 수 있는 장점이 있다 (Goodrich, 1996). 즉, 루브릭은 논증과정의 목표를 교사와 학생이 모두 명확하게 인식시킬 수 있는 도구이므로 수업이나 활동을 통하여 평가가 바로 이루어질 수 있어 매우 강력한 도구가 될 수 있다.

이 연구는 이와 같은 필요성을 바탕으로 과학적 논증과정 교수와 평가에 도움을 주기 위하여, 기존에 이루어졌던 과학적 논증과정 분석과 평가에 대한 연구들을 검토하고 종합하여 초, 중, 고등학생의 과학적 논증과정 평가를 위한 루브릭을 개발하는데 목적이 있다.

## II. 연구 방법 및 절차

루브릭의 평가 기준은 일반적으로 학생의 반응을 판단하는데 사용될 하나 이상의 기본적인 영역 혹은 준거, 각 영역이나 준거의 의미를 명확히 해주는 정의나 사례, 각 영역이나 준거를 평정하기 위한 수치나 범주, 각 준거의 단계, 단계별 사례 혹은 범례, 단계별 특성을 구분하는 성취 수준 등의 4가지 요소를 포함한다(Herman, Aschbacher & Winters, 1992). 이러한 요소들을 고려하여 루브릭은 Zimmaro(2004)가 제시한 절차에 의해 다음의 5단계로 진행되었다.

### 1. 논증과정 평가 기준 추출

과학교육분야에서 인용지수가 높은 학술지인 *Science Education*, *Journal of Research Science Teaching*, *International Journal of Science Education* 등에 최근 10년간 게재된 논증과정 관련 논문들을 중심으로 이들 논문에서 인용하고 있는 문헌들과 학술연구정보서비스에서 검색을 통해 찾아낸 문헌 등을 참고로 하여 연구자가 논증과정을 분석하는 절차나 평가 관점을 찾고, 실제 논증과정을

평가에서 높은 수준과 낮은 수준이라고 평가된 논증과정들의 공통점을 살펴봄으로써 논증과정에서 중요시 되는 평가 관점을 추출하였다. 논증과정에 관한 연구가 어떻게 이루어져 왔는가를 총체적으로 살펴본 논문들을 토대로 대범주를 생성하고 그 하위 범주로 문헌에서 중시하고 있는 평가 관점들을 범주화하였다.

### 2. 루브릭의 유형 결정

Arter와 McTighe (2001)에 따르면 루브릭은 크게 총체적(holistic), 분석적(analytic)이라는 관점과 과제 특수적, 과제 일반적인 관점의 두 가지 차원으로 나눌 수 있다. 총체적 루브릭은 전반적인 특성에 따라 측정하지만 분석적 루브릭은 수행이나 수행 결과물을 몇 가지 특성이나 차원으로 나누어 복합적으로 분석하는 것이다. 또한 과제 일반적(task general)인 것은 같은 루브릭을 다양한 평가 상황에서 사용하는 것을 말하며 과제 특수적(task specific)인 것은 과제에 따라 각기 다른 루브릭을 사용하는 것을 뜻한다. 따라서 루브릭에 따라 분석하는 방법과 사용하는 상황이 다르기 때문에 평가 목표에 맞는 형태로 결정하였다. (그림 1과 2 참조)

### 3. 핵심 요소들에 대한 명확한 정의

범주화된 평가 관점들에 대한 세부적인 항목들에 대하여 명확하게 정의하는 단계로써 루브릭을 제시함에 있어서 채점자에게 해당 평가 관점에서 무엇을 평가하고자 하는가에 대하여 명확하게 이해할 수 있도록 도와주는 역할을 한다. 제시된 평가 관점들은 모두 문헌 연구를 통해 도출되었다.

### 4. 각 요소에 대한 수행 점수를 판별하는 척도 진술

평가 등급 혹은 평정 척도는 다양한 척도를 포함하고 있는데 평정 척도는 수리적인 것, 질적인 것, 때로는 양자 모두를 포함한다. 수리적 척도는 수행 수준에 따라 점수를 부여하도록 고안된 척도이다. 점수 척도를 몇 개로 나누느냐 혹은 그 간격을 얼마로 하느냐 하는 것은 어떤 목적으로 그것을 사용하느냐에 따라 달라진다. 질적 척도는 수행을 요소의 유무나 비율로

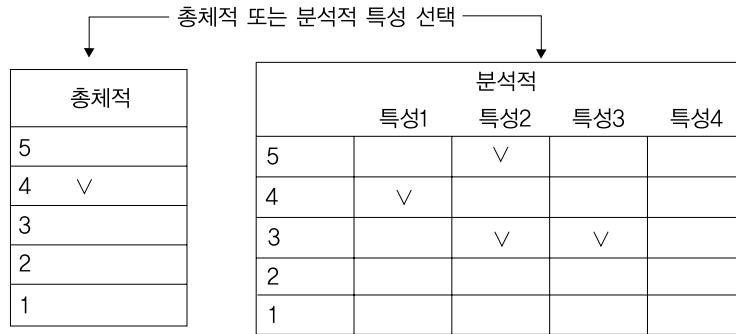


그림 1 총체적 또는 분석적 루브릭(Arter, 2000)

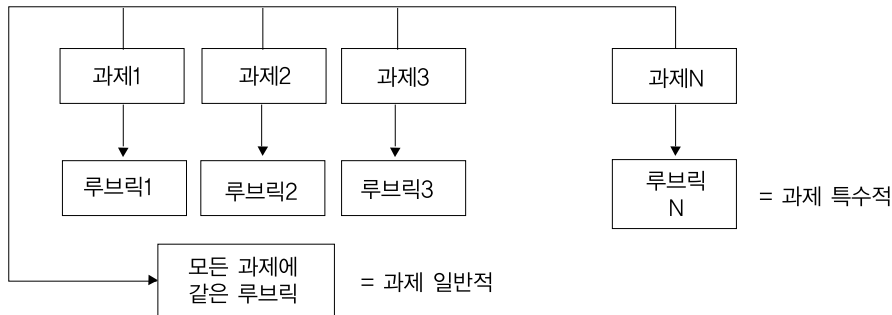


그림 2 일반적 또는 과제 특수적 루브릭(Arter, 2000)

평가하는 양적인 서술 대신 질적 평가기준을 서술하는 것으로 크게 서술적(descriptive)척도와 평가적(evaluative) 척도로 나눌 수 있다. 예를 들어 서술적 척도는 시도하지 않음-부분적인 완성-완성-요구한 것 이상의 완성과 같이 서술적으로 표현하며, 평가적 척도는 A-B-C와 같이 등급(grade)으로 표현하되 이미 아주 우수함, 우수함과 같은 평가적인 의미를 내포한다. 이외에 수리적 척도와 질적 척도를 함께 사용할 수 있으며 좋은 채점 기준은 이 둘을 동시에 포함하는 경우가 많다(Herman, Aschbacher & Winters, 1992). 따라서 이 연구에서도 수행 기준을 제시함에 있어서 질적인 진술을 사용할 것이며 수행 수준 사이를 구분하는데 있어서 상대적인 비교기준이 아닌 절대적인 기준을 사용하며, 상위의 수행 기준은 하위 기준을 만족했음을 전제로 진술하였다.

### 5. 과학 교육 전문가에 의한 내용 타당도 검증

내용 타당도는 과학적 논증과정을 평가하는 범주의

선정이 적절한지, 범주별로 추출된 하위항목과 각 항목에 대한 평가 척도 등이 타당한지에 관한 내용이다. 1차적으로 추출된 요소와 하위 항목에 대하여 동료 연구자들과 여러 차례 수정과정을 거친 후, 최종적으로 개발된 수업 분석 도구는 과학 교육 전문가 4인에 의한 검토를 통해 두 차례에 걸친 수정 작업 뒤 최종적으로 개발된 루브릭에 대한 타당도를 검증받았다. 내용 타당도의 분석은 Likert 척도에 의해 5단계 평정 척도법을 사용하였고, 이들 척도는 매우 적절하다 5, 적절하다 4, 보통이다 3, 부적절하다 2, 매우 부적절하다 1과 같다.

### 6. 루브릭의 예비 적용

루브릭 개발 후 평가도구의 적절성 및 유용성 검증을 위하여 이를 과학 영재의 논증과정을 평가하는 과제에 예비 적용하였다. 적용 대상은 광역시 소재 국립대학교 과학영재교육원 초등과학반 5-6학년 학생 28명이다. 선정된 탐구 실험 주제는 가열된 캔을 수조

속의 물에 거꾸로 집어넣은 후의 현상을 관찰하는 활동이다. 이 주제를 통해 논증과정 수업을 이끌어낸 이유는 과학적 설명을 제안할 때 논리적 근거를 형성할 수 있는 변인들이 다양하고 실험 결과 및 결론에 대해 사전에 학습하지 않은 내용이어야 논증과정이 활발히 이루어질 수 있기 때문이다. 논증과정은 주로 모둠별 결과 발표와 그에 관한 질의응답으로 구성되었는데 이 논증과정을 평가함에 있어서 개발된 루브릭을 예비 적용함으로써 루브릭 개발진과 실제 채점자 간에 지속적인 피드백을 통해 항목에 대한 정의나 평가 척도의 서술이 모호하여 채점자가 잘 이해할 수 없는 부분에 대하여 수정하고 보강하고자 하였다.

### Ⅲ. 논증과정 루브릭 개발

#### 1. 문헌 연구를 통한 논증과정 평가 관점 추출 및 범주화

논증과정에 관한 연구는 크게 철학적인 관점과 심리학적인 관점에 연구로 나누어 볼 수 있는데 (Nussbaum, 2008), 철학적인 관점의 연구들은 크게 형식적(formal) 비형식적(informal) 측면으로 구분할 수 있다. 형식적 측면에서는 삼단 논법으로 대표되는 아리스토텔레스의 논증과정에서 출발하여 논증과정 요소들의 사용 유무를 분석하는 Toulmin 연구가 있고, 비형식적 측면은 일상생활에 보다 가까운 논증

과정으로 결론이 얼마나 반대논증과정에 대해서 잘 반박하고 있는냐로 논증과정을 평가하는 연구가 해당된다. 또한 심리학적인 관점에서의 논증과정은 현재의 개념이 가진 문제점에 주목할 수 있도록 사회 인지적 비평형(sociocognitive conflict)을 초래한다는 관점의 연구와, 개념간의 관계와 선지식을 알 수 있게 해주어 정보를 상기시키는 다양한 경로를 만들어 낸다는 인지적 정교화(cognitive elaboration)의 관점에서 진행되는 현장연구나 실험 연구가 있다. 이를 통해 논증과정에 관한 연구는 논증과정이 갖추고 있는 형식(form)이 무엇인가, 이러한 논증과정이 어떠한 상호과정을 통해 일어나는가, 그리고 개념을 학습하는데 어떤 도움을 주는가에 관한 연구 등 3가지 범주로 수렴될 수 있다는 것을 알 수 있다. 또한 Sampson과 Clark(2008)가 과학교육자들이 과학적 논증과정의 질을 평가하는데 사용하는 분석적인 틀을 살펴본 결과 구조(structure), 정당화(justification) 그리고 내용(content)라는 세 가지 초점을 발견할 수 있었다. 이와 함께 논증과정에 대한 이론적인 개념화를 통해 과학교육에 시사점을 찾고자 했던 Bricker와 Bell(2008)의 연구 과정(그림 3. 참조)에서도 Toulmin의 구조적인 분석이나 형식적 논리학으로 대표되는 논증과정의 형식에 관한 부분, 수사적 분석이나 과학적 추론분석으로 알 수 있는 내용적인 부분, 비계학습을 통해 알 수 있는 태도적인 부분이라는 대범주로 분류할 수 있다.

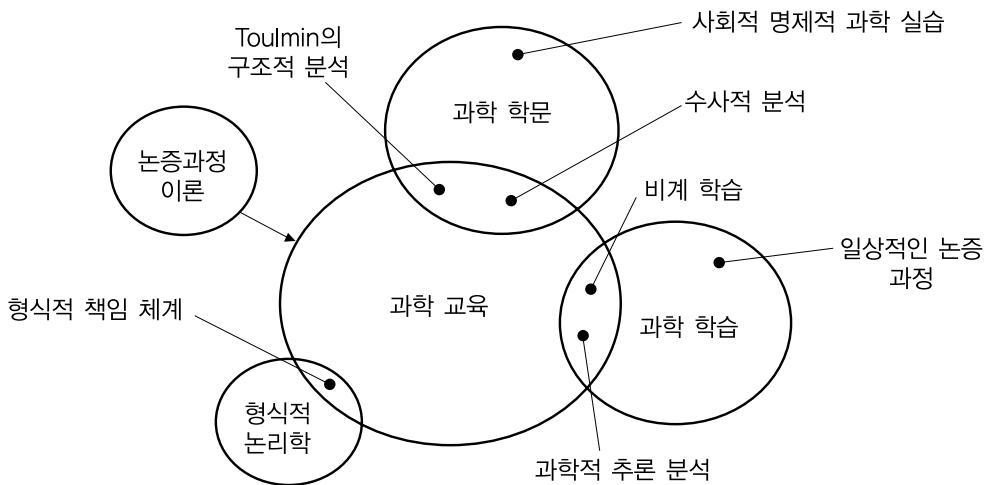


그림 3 과학 교육에서 논증과정에 대한 이론적 개념화 (Bricker & Bell, 2008)

또한 실제 과학적 논증과정 분석에 대한 문헌들을 고찰해본 결과 다양한 분석 관점들을 토대로 아래와 같은 3가지 대범주로 범주화할 수 있었다.

### 1) 형식(form)

Toulmin이 제시한 기준을 바탕으로 하거나 이를 토대로 또 다른 논증과정의 요소를 추출하여 실제 학생들의 논증과정에서 그러한 요소들이 얼마나 나타나는지, 그 요소들을 각각 평가할 때 어떤 관점을 중요시해야 하는지를 분석하는데 초점을 맞추고 있는 문헌들의 평가 관점이 '형식'이라는 대범주이다. Aufscjnaiter 등(2008)은 논증과정의 질을 분석하는데 있어서 초기 논증과정 연구에서는 내용보다는 그것이 포함되고 있는 요소자체의 분석에 초점을 두어 단순한 주장을 넘어 이를 뒷받침할 수 있는 요소들을 가지면 논증과정이 정당화될 수 있다는 것을 밝혔다. 이와 유사하게 Osborne 등(2004)은 Toulmin 틀을 바탕으로 어떤 논증과정이 더 높은 논증과정인가를 결정하는데 있어서 주장, 자료, 보장, 보강 등의 요소와 함께 반증이 논증과정에서 나타나 있는가를 수준을 나누는데 있어 중요하게 여겼고, Erduran 등(2004)도 Toulmin의 틀을 사용하여 전체 학생들과 그룹 학생들의 논증과정을 직접분석하고 평가함으로써, 형식적으로 요소를 갖추고 있는가를 통해 논증과정의 질이 평가될 수 있음을 입증하였다.

Simon 등(2006)에 의하면 학생들도 논증과정을 평가해보도록 한 결과 '증거를 갖고 있는가' 라는 형식적인 요소를 중요시 하였다. 이렇듯 객관적으로 확인할 수 있는 요소들을 찾는 것이 논증과정 분석을 용이하게 해준다는 연구 결과들을 통해 논증과정의 질을 평가할 때 형식적인 측면에 초점을 두는 것이 중요한 관점이 될 수 있다는 것을 알 수 있으며, 논증과정 전문가들조차 언어가 가진 모호함 때문에 논증과정의 질을 평가하는데 직접적인 기준을 제시하여 판단하는 것이 힘들기 때문에 객관적으로 평가 가능한 요소들에 초점을 맞추는 것이 효과적인 평가 방법이 될 수 있음을 알 수 있다.

### 2) 내용(contents)

논증과정에서 가장 중요한 것은 내용 특수적인 지식의 이해와 적용이라고 주장하며, 과학적 지식의 수준이 논증과정상에 가장 큰 영향을 미친다는 관점에

서 논증과정에 나타난 지식의 수준을 분석하는데 초점을 맞추고 있는 것이 '내용'이라는 범주이다. Takao 등(2002)은 특정 과제에 대한 논증과정을 분석함에 있어서 지식의 수준에 따라 그 질이 다를 수 있다고 파악하고, 그 안에 나타난 내용을 단순 관찰에 대한 서술만 하는 수준에서부터 이론적인 모델을 제시하는 수준까지 6단계로 나누어 학생들이 제시한 논증과정을 지식수준에 따라 평가할 수 있음을 제시하였다. 또한 Aufscjnaiter 등(2008)에 의하면 논증과정을 수행하는데 필요한 일반적인 능력을 높이는 것 보다 친밀한 과제를 제시하는 쪽이 논증과정의 빈도를 더 높일 수 있다고 말하였다. Jimenez-Aleixandre 등(2000)이 제시한 관점을 보면 귀납, 연역, 인과, 정의, 분류, 호소, 일관성, 그럴듯함의 관점 등 논증과정에 사용한 설명체계가 어떤 방법을 통해 구성되었는지를 보고 학문적으로 어떤 수준의 증거를 끌어와서 설명하고 있는지를 중요하게 여겼다. 또한 Bazerman(1988)에 의하면 형식적인 분석에 초점을 두는 것은 그 논증과정이 말이 되는지 안 되는지 아는데도 실패할 수 있다고 밝히며 논증과정 구조는 내용 분야와 무관할지 모르지만 논증과정의 질은 내용 분야에 따라 다른 기준에 의해 평가되어야 하므로 학술적인 수준에 따라 분석되어야 한다고 했다. 이러한 관점에서 Sandoval와 Millwood(2005)는 내용 특수적인 용어의 사용을 분석한다거나 이론에 기반한 인과적인 설명을 하는가, 자연과학에서 제시하고 있는 이상적인 생각과 비슷한지 아닌지, 학술적으로 개념을 잘 이해하고 있는지를 평가관점으로 제시하였다. 이처럼 논증과정 분석에 관한 연구들 중 내용 특수적인 과제를 분석하는 논문에서 내용에 나타난 지식의 수준을 중요시하는 관점이 많이 발견되었고, 이러한 분석을 통해 요소의 유무만으로 평가할 수 없는 논증과정의 질을 평가할 수 있다고 주장했다. 따라서 과제 특수적으로 내용의 수준을 평가할 수 있는 관점이 정리된다면 논증과정 교수가 보다 수월해 질 것이라고 생각된다.

### 3) 태도(attitudes)

논증과정은 학생들 간의 상호작용에 의해 정교화되고 발달해간다는 측면에서 논증과정에 참여하는 정도와 마음가짐에 대한 중요성을 강조한 관점이 '태도'라는 범주이다. Mercer 등(1999)은 논증과정에서 상

호작용의 중요성을 강조하면서 상호작용의 차이를 분석할 수 있는 Downing model을 개발하여 7수준으로 학생들의 논증과정을 평가하고자 했다. 이 모델에서는 학생들이 논증과정에 참여하려는 태도에서부터 상대방의 입장을 어떻게 받아들이는가, 그를 바탕으로 자신의 입장을 조정하는지에 대해서도 평가 관점으로 삼았다. 또한 Veerman 등(2002)에 의하면 논증과정상에서 자신이 결정한 입장에 따라서 과학적인 정보를 편파적으로 평가할 수 있으며, 자신의 입장만을 옹호하고 공격하는 성향 때문에 반대 입장의 타당한 증거를 무시하는 특성을 보이는데 이러한 행동이 타협의 여지를 제한하고 지식의 교환을 방해할 수 있다. 또한 논증과정상에서 정당화를 요청하고 그 요청에 대해 응답하는 것은 논증과정에서 중요하며 학생들이 이러한 과정에 적극적으로 참여함으로써 논증과정이 이끌어 나갈 수 있게 되며 이것이 부족하면 논증과정이 중단되는 원인이 될 수 있다(강순민 등, 2006). 따라서 논증과정상에서 열린 마음으로 모든 가능한 증거들을 검토하는 태도적인 측면도 질 높은 논증과정을 이끌어내는데 중요한 요소가 될 수 있다.

## 2. 루브릭의 유형 결정

분석적 루브릭은 총체적 루브릭에 비해 채점 하는

데 더 오래 걸리지만 교사에게 더 많은 정보를 제공하며 학생의 특정한 강점과 약점을 진단하는데 효과적이므로(Waltman *et al.*, 1998), 교수와 평가 두 가지 측면에 도움을 줄 수 있다고 판단되어 분석적 루브릭으로 구성 하였다. 또한 과제 특수적 루브릭은 특정 과제에만 종속되어 있어 구조와의 관련성이 적고 그 때문에 일반화에 제약이 있기 때문에(Messick, 1994), 수행의 질과 핵심적인 특성을 학생들이 이해할 수 있도록 도와줄 수 있는 과제 일반적인 루브릭을 선택하였다(Arter & McTighe, 2001).

## 3. 핵심 항목들에 대한 구체적 정의

앞서 제시한 형식, 내용, 태도의 3가지 대범주에 따라 문헌에 제시된 평가 관점들을 범주화하였고, 이를 동료 연구자들과의 세미나를 통해 1차 수정한 후, 과학 교육 전문가 4인의 검토에 의해 2차 수정한 결과는 그림 4 와 같다.

수정 과정을 자세히 살펴보면 1차 동료 세미나를 통해 수정한 결과 형식적인 측면에 호소라는 항목이 권위에 호소라는 점에서 출처제시의 관점과 겹쳐서 제외되었고, 기술적인 오류 즉, 오타자, 문법 오류와 같은 것을 포괄한다는 점에서 오류로 명칭이 수정되었다. 내용적인 측면에서도 과학적 개념 이해와 용어 사

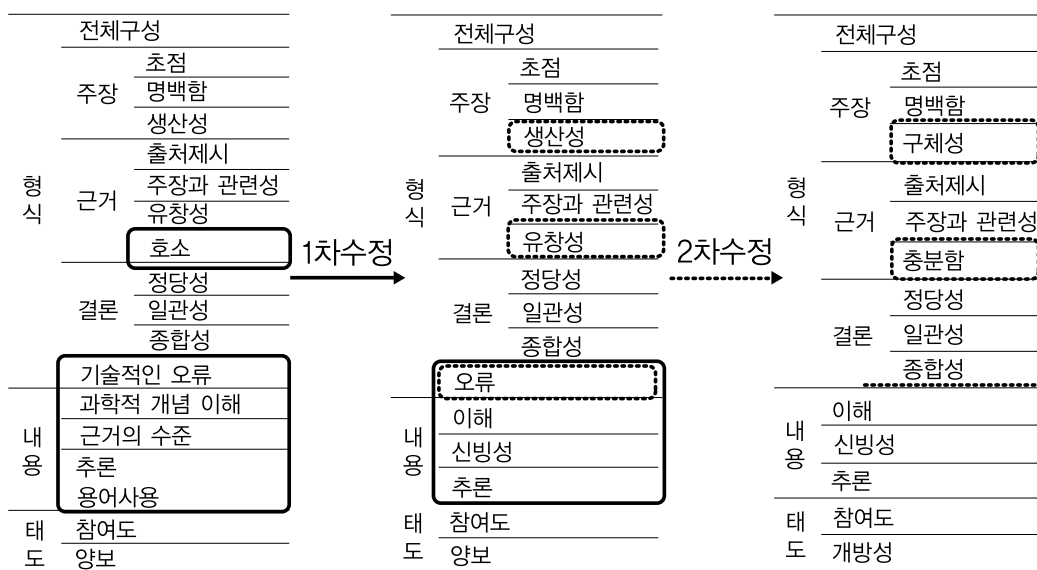


그림 4 논증과정 평가항목 수정과정

용이 증척되는 부분이 있어 이해라는 요소로 통합했고, 근거의 수준은 얼마나 믿을만한가를 나타내는 신빙성이라는 용어로 교체되었다.

과학 교육 전문가 검토에 의한 2차 수정 결과 생산성은 구체성으로 유창성은 충분함으로 보다 직관적으로 의미를 알 수 있는 용어로 교체되었고, 형식적 분야의 오류라는 항목은 과학적 논증과정의 평가에 있어서 중요하지 않은 평가 항목이라는 조언에 따라 삭제하였다. 종합성의 진술이 모호한 점을 지적받아 보다 구체적으로 재 진술 하였고, 그밖에 문맥상 모순이 있는 부분을 조금씩 수정하였다. 완성된 각 항목이 도출된 구체적인 과정과 문헌에서 제시하고 있는 정의는 다음과 같다.

### 1) 형식

우선 Kanselaar 등(2002)은 도입, 전개, 정리로 정의되는 형식적인 구조를 찾거나 Michell(2001)처럼 논증과정 안에서 하위 논증과정이 얼마나 일어나고 있는가, 하위 논증과

정 간의 연결이 잘 이어지고 있는가를 평가하였다. 또한 Chinn(2000)은 논증과정을 개념도로 정리하여 보다 효과적으로 그 논증과정의 구조를 정리하고자 하였다. Schwarz 등(2003)은 논증과정 구조 분석 틀을 제시하면서 논증과정의 질은 내용보다는 구조적 복잡성에 의해 평가될 수 있다고 주장하며, 개인이 생성하는 논증과정의 질은 상황에 의존적이지 않다고 하여 논증과정의 구조를 중시하였다. 이러한 관점들을 묶어 전체 구성이라는 항목을 추출하였다.

또한 논증과정의 요소별로 나누어 분석하는 관점이 있는데 Toulmin의 틀인 TAP에 따라 근거, 주장, 보강, 보강, 한정, 반증으로 나누어 분석한 학자들이 있고(Bell & Linn, 2000; Clark & Sampson, 2006; Jimenez-Aleixandre *et al.*, 2000; Kuhn & Reiser, 2005; McNeill *et al.*, 2006; Simon *et al.*, 2006; Zembal-Saul *et al.*, 2003), 이를 더욱 세분화하여 요소별로 루브릭 평가 기준을 나타낸 연구결과와(McCann, 1989) TAP의 여러 요소 중에서 특히 반증의 유무를 중시하는 관점이 있다(Osborne, 2004).

이에 비해 TAP에 나타난 논증과정 요소를 찾아내는데 어려움을 토로하고 보다 발견하기 쉬운 요소들로 평가 항목을 정하여 이들의 특성을 통해 평가하는

관점이 있다. 이는 크게 대화적 논증과정요소로 구별한 것과 이보다 더 포괄적으로 분석한 관점이 있다. 대화적 논증과정 요소는 강순민 등 (2006)에서 제시한 주장에 대한 질문, 근거에 대한 질문, 단순반박, 근거 반박, 요청 및 요청 응답, 단순 호응, 강화 및 정교화, 메타 질문 등이 있고, Felton과 Herko(2004)이 제시한 입장, 주장, 반대 입장의 주장, 반론, 반증, 한정, 보류와 같은 항목 등이 있다. 또한 요소들을 포괄적으로 분석한 관점으로 Marttunen(1994)은 논증과정에서 나타나는 요소를 크기 주장, 근거, 결론으로 나누었고, Maloney와 Simon(2006)은 토론, 증거 재검토, 증거 명확화 등으로 나누어서 논증과정을 분석하였다.

논증과정 분석을 실행한 연구자들이 논증과정 요소의 복잡성으로 인해 학생들의 논증과정 분석에 있어서 요소를 찾지 못해 평가할 수 없다는 한계를 제시함(Naylor *et al.*, 2007)에 따라 요소별로 분석하는 형식적 측면에 있어서 보다 포괄적인 관점이 적절하다고 판단하여 하위 평가 항목에 들어갈 요소로 Marttunen(1994)가 제시한 논증과정 요소인 주장, 근거, 결론을 사용하였다. 또한 상호과정적 논증과정 요소를 강조한 학자들에게서 중요시되는 평가관점들은 태도 영역에서 다루었다. 평가 관점을 분명히 하기 위해 문헌을 바탕으로 각 항목에서 중요하다고 생각되는 관점들을 세부적으로 추출하여 표 1과 같이 정의하였다. 먼저 전체 구성의 관점은 Osborne 등(2004)이 제시한 논증과정 평가 틀을 기반으로 진술하였고, 주장, 근거, 결론의 기본틀을 제시한 Marttunen(1994)이 제시한 논증과정 평가관점에 있어서 주장의 명백함, 구체성, 초점 등의 항목을 추출하였고 근거는 Maloney와 Simon(2006)의 관점에 따라 근거의 충분성, 출처 제시의 중요성에 대하여 서술하였다. 또한 Sandoval 와 Millwood (2005)도 출처 제시, 근거의 충분성, 주장과의 관련성 등을 제시하였다. 결론은 Aufscjnaiter 등(2008)에 제시된 정당성에 관한 항목과 Hogan 등(2000)이 제시한 종합성 등의 관점에 따라 진술하였다.

### 2) 내용

Kelly와 Bazerman(2003)에 의하면 과학적 지식이 반영된 용어를 사용하는가가 분석의 관점이 될 수 있다고 하였으며 McNeill와 Krajcik(2007)에 의하면



표 1  
형식적 측면의 평가 항목 정의

평가항목	정 의
전체 구성	전체적인 구성이 단순한가 복잡한가, 틀민의 요소인 주장, 근거, 보장, 보강, 한정, 반증 등이 얼마나 드러나 있는가 등의 논증과정 요소 사용 정도를 판단하는 것으로 특히 반증의 표현 여부를 중요시 한다.
명백함	무엇을 말하고자 하는가가 뚜렷하게 진술되어 모든 참여자들이 확인할 수 있고 이해할 수 있을 정도로 명쾌한가를 판단하는 것이다.
주장	<p>초점 주장의 초점이 하나의 논쟁점에 맞춰져 있는가를 보는 것으로 보편적이거나 모호한 진술이 아니라 주제를 정확하게 반영하고 있는가를 판단하는 것이다.</p> <p>구체성 논의 안에서 드러난 부차적인 논제의 수에 따라 판단되는 것으로 얼마나 구체적으로 주제에 접근하여 주장하고 있는가를 살펴보는 것이다.</p>
출처 제시	어떤 방법으로 자료를 언급하는가, 수많은 기록들에 대한 이해와 이를 논증과정에 어떻게 사용하는가에 관한 문제로 인용한 자료를 어떻게 다루는가에 관한 정교화 수준에 대한 문제이다.
근거	<p>주장과 관련성 주장과의 관련 되는 특정 부분에 대하여 자세히 설명하거나 논증과정을 구성하는데 있어서 그 근거가 가진 중요성을 언급하고, 자료를 단지 제시하는데 그치지 않고 그에 관한 자신의 해석을 제시하는가를 평가하는 것이다.</p> <p>충분함 주장을 뒷받침하는 근거가 충분한가를 판단하는 것으로 얼마나 많은 자료를 폭넓게 조사하여 제시하였는가를 통해 평가한다.</p>
정당성	결론이 가진 주요 특성들이 증거를 기반으로 하였거나 그로 인한 추론을 기반으로 하여 이끌어낸 결론인가를 판단하는 것으로 왜 그렇게 생각하는가 또는 어떻게 알았는가 등에 대하여 답을 할 수 있는가의 여부로 판단된다.
결론	<p>일관성 주장과 결론사이의 논리적인 일관성을 가지고 전개되었는가를 판단하는 것으로 정당화 과정이나 설명이 결론 도출에 적합한가를 판단하는 것이다.</p> <p>종합성 반대 입장과 자신의 입장을 조정하여 합의된 결론을 이끌어내는 능력을 판단하는 것으로 이는 고등 사고와 변증법적 사고가 가능한가를 판단하는 것이다.</p>

내용 지식의 이해 정도가 과학적 설명체계를 구성하는데 있어서 매우 중요하다고 하였다. 또한 Jimenez-Aleixandre 등(2000)이 제시한 관점을 보면 논증과정상에서 나타나는 학문적인 활동으로 귀납, 연역, 인과관계, 정의, 분류, 호소, 일관성 등이 있으므로 설명체계를 구성하는데 있어서 어떠한 추론전략을 사용하여 적절한 인과관계를 모색하는가에 따라 평가 될 수 있다고 하였고, Newton 등(1999)에 의하면 논증과정은 개념 이해를 돕기 위한 것이며 높은 수준의 논증과정을 하기 위해서는 지식에 관한 이해가 선행되어야 한다고 하였다. 학생의 논증과정을 평가하는데 있어서 학자들이 공통적으로 그들이 제시한 논증과정자체가 믿을만한 것인가를 평가해야 한다고 제시함으로써 이러한 연구들을 바탕으로 하여 내용적 측면에 하위 평가 항목을 과학적으로 얼마나 잘 이해

하고 있는냐에 따른 평가항목인 주제에 대한 이해와 과학적 용어 사용에 관한 것을 평가하는 이해, 학생의 논증과정이 과학적 사실에 얼마나 부합하는가에 따른 신빙성, 또한 그 연결이 과학적 추론 방법인 연역, 귀납 등의 추론들과 같이 논증과정이 얼마나 논리적으로 전개되었는가를 평가하는 추론 등의 3가지 관점으로 나누게 되었다. 평가 관점을 분명히 하기 위해 문헌을 바탕으로 각 항목에서 중요하다고 생각되는 관점들을 세부적으로 추출하여 표 2와 같이 정의 하였다.

### 3) 태도

먼저 논증과정에서 학생의 참여정도를 중요시하는 관점에서 상대에 대하여 어떤 반응을 보이는가, 상대방의 요구에 대하여 어떻게 대응하는가 등을 평가 관

**표 2**  
내용적 측면의 평가 항목 정의

평가항목	정의
이해	주제에 대한 이해 제시된 주제에 대해 학술적으로 잘 이해하고 있는가의 여부를 알아보고 주제와 관련된 개념들을 과학적으로 정확하게 이해하고 있는가를 평가한다.
	용어 사용 과학적 논증과정에서 해당 주제와 관련된 내용 특수적 용어를 사용하였는가와 그 용어가 적절한 곳에 사용되었는가에 의해 논증과정의 전문성을 판단하는 것이다.
신빙성	논증과정 내용상 가장 중요한 판단 관점으로 믿을만한가, 진실한가, 정확한가와 관련된 것으로 그 근거의 출처가 개인인가, 전문가인가, 연구에 의한 것인가, 합의를 이룬 것인가 등에 의해 평가된다.
추론	추론이란 주장과 증거의 연결을 정당화하는 것으로 주장과 증거를 방어하기에 충분하고 적합한 과학적 방법들을 포함하는 것이다. 따라서 충분한 증거를 사용하여 주장을 뒷받침하고 이를 통해 올바른 선택에 이르게 하는 과정이 논리적인가를 판단하는 것이다.

점으로 삼은 것에는 Mercer 등(1999)의 Downing model이나 강순민 등(2006)의 상호 과정적 논증과정 요소 분석 등을 들 수 있다. 또한 논증과정상에서 타협의 의지를 남겨두고 상대방의 입장이라 할지라도 존중할 수 있는 태도가 질 높은 논증과정을 이끈다는 관점을 제시하였다(Chin & Anderson, 2000; Felton & Kuhn, 2001; Veerman *et al*, 2002; Maloney & Simon, 2006). 이러한 연구들을 바탕으로 하여 태도적 측면에 하위 평가 항목을 참여도와 개방성(concession)로 나누게 되었다. 평가 관점을 분명히 하기 위해 문헌을 바탕으로 각 항목에서 중요하다고 생각되는 관점들을 세부적으로 추출하여 표 3과 같이 정의하였다.

#### 4. 각 요소에 대한 수행 점수를 판별하는 척도 기술

핵심요소에 관한 정의와 문헌에서 제시한 수준을 바탕으로 학생들의 수행을 채점할 수 있는 루브릭을 1점에서 5점까지의 척도로 나타내었다.

**표 3**  
태도적 측면의 평가 항목 정의

평가항목	정의
참여도	기본적으로 토론에 참여하는 태도가 어떠한가를 보는 것으로 자신의 주장과 근거를 제시하는가, 다른 사람의 의견에 어떤 반응을 보이는가, 근거를 평가하고 판단을 내리는가 등의 전반적인 태도를 평가한다.
개방성	주장에 대해 제기한 이의를 받아들이는 것으로 다른 사람의 관점을 신중히 받아들여 열린 마음으로 검토할 수 있는가 하는 것으로 자신과 상반되는 논증과정이라 할지라도 그 타당성을 인정하고 가능한 모든 자료를 검토할 때까지 확정적인 결론을 내리지 않고 판단을 보류할 수 있는 능력을 평가하는 것이다.

#### 1) 형식

##### ① 전체 구성

이 항목은 논증과정 전체의 구성이 잘 짜여져 있는가, 논증과정 구성요소가 나타나 있는가를 평가하는 것이므로 Osborne 등(2004)이 제시한 기준에 따라 척도를 기술하였다(표 4). 이는 Toulmin의 요소를 정확히 수를 세어 판별해내지 못해도 각 요소가 어느 정도 나타나 있는지, 특히 반증이라는 요소, 즉 주장이 옳지 않음을 반대 근거를 들어 증명하는 요소가 얼마나 나타나 있는가의 여부를 통해 파악할 수 있다. 논증과정 요소를 통한 평가에서 반증을 중요시하는 이유는 좋은 논증과정이란 주장에 대한 근거를 포함해야 될 뿐만 아니라 반증이나 한정과 같은 통합적인 장치 또한 필요하다라는 관점에 따른 것이다(Pollack, 1987).

##### ② 주장

이 항목은 논증과정에서 가장 명확하게 판별해 낼 수 있는 주장의 수준에 대해 평가하는 것이므로 이를 판별해낼 수 있는 기준이 되는 명백함, 초점, 구체성

등의 요소들을 토대로 척도를 제시하였다(표 4). 명확하게 진술되어 다른 사람이 이해 가능한가를 가장 기초적인 수준으로 보았고, 진술된 주장이 주제와 관련되어 있는가의 여부를 그 다음 수준으로(Marttunen, 1994), 주제와 관련성을 얼마나 구체적으로 표현하고 있는가를 상위 수준으로 판별하였다(Hogan, 2000). 각각의 관점들은 하위 수준이 선행되어야 상위 수준의 판별이 가능하므로 명백함, 초점, 구체성의 순서로 수준이 나뉠 수 있다.

### ③ 근거

이 항목은 논증과정에서 대부분을 차지하는 근거에서도 내용적인 측면이 아닌 객관적으로 판단 가능한 형식적인 것을 평가하는 것이므로 기준이 되는 근거의 출처를 명시하는 방법, 주장과의 관련성을 표현했는가, 사용한 자료가 충분한가 등을 토대로 척도를 제시하였다(표 4). 먼저 출처를 명시하는 방법과 주장과의 관련성을 표현했는가의 여부는 Sandoval와 Millwood (2005)가 인용한 자료들을 어떻게 다루는가에 따라 나는 5수준을 참조하여 기술하였다. 또한 근거가 충분한가는 Siegel 등(2006)이 제시한 루브릭을 참조하였으며 얼마나 많은 자료를 폭넓게 사용하였는가를 판단하는 것인데, 이는 비단 양적인 측면 뿐 아니라 하나의 자료만 갖고 있다 하더라도 그것이 주장을 뒷받침 하는 것이라면 충분하다고 할 수 있고, 충분한 양의 자료를 가졌다 하더라도 주장에 맞는 자료가 전혀 없을 수도 있으므로 신중하게 판단되어야 한다.

### ④ 결론

이 항목은 논증과정에서 마무리 부분에 해당하는 결론적인 측면을 평가하는 것이므로 정당성, 일관성, 종합성 등의 관점에서 평가될 수 있다. Hogan와 Maglienti(2001)에 의하면 학생들은 결론을 이끌어 내는데 있어서 증거 대신에 개인적인 관점에 의존하는 경향이 있다고 하여, 어떻게 그러한 결론이 나왔는가에 대한 설명을 할 수 있는가의 여부가 결론의 정당성을 입증하는 방법이 될 수 있다. 즉, 근거를 기반으로 주장을 정당화하되 단순한 근거를 기반으로 했는가, 정교한 근거를 기반으로 했는가가 평가 관점이 될 수 있고 반대 입장이 나타났는가 또한 중요한 평가 기준이 될 수 있다(Sadler & Fowler, 2006). 그러나 근거를 기반으로 한 결론이 논리적으로 완전할 것이라

는 기대는 하지 않으므로 주장과 결론 사이의 관계에서 일관성의 측면에도 초점을 맞추어 분석하였다(Marttunen, 1994). 또한 Hogan 등(2000)이 종합성이라는 측면에서 반대 입장이 진술되었는가, 얼마나 진술되었는가를 바탕으로 제시한 수행 기준 또한 참조하여 결론의 수준을 판별할 수 있는 척도를 기술하였다(표 4).

## 2) 내용

### ① 이해

이 항목은 주제에 대하여 과학적으로 이해하고 있는가의 여부와 용어 사용의 수준에 대하여 평가하는 것이다. 따라서 이 같은 점이 잘 반영되어 내용 특수적으로 개발된 Schafer 등(2001)이 제시한 생물학 루브릭을 기반으로 기술하였고, Kelly와 Takao(2002)의 지식의 수준에 따라 평가한 척도를 참조하였고, Kelly와 Bazerman(2003)의 의견에 따라 과학적 용어의 사용 정도 또한 평가 항목으로 포함시켰다(표 5).

### ② 신빙성

이 항목은 논증과정에서 핵심이 될 수 있는 것으로 제시한 근거가 과학적으로 믿을만한 사실인가 하는 여부에 관한 것이다. 근거의 질을 검증할 수 있는 여러 가지 방법 중 어느 수준까지 과학적으로 합의가 이루어진 근거를 사용하는가에 따라 근거의 위계를 나눈 Perella(1987)의 6수준을 바탕으로 척도를 나누었고 Van Gelder(2002)의 의견을 참조하여 진술하였다(표 5).

### ③ 추론

이 항목은 내용에 대한 이해를 바탕으로 각각의 인과적 요소들 안에서 주장에 대한 근거가 명료화된 정도를 평가한다. 이는 Sandoval 와 Millwood(2005)의 개념적 질을 판단하는 관점에 따른 것이며 올바른 인과관계가 성립될 수 있도록 적합한 자료를 사용하였는가를 하는 것을 평가하는 것으로 Siegel 등(2006)이 제시한 수행 척도를 참조하였다(표 5).

## 3) 태도

### ① 참여도

이 항목은 논증과정이 진행되는 동안 얼마나 활발하게 논의가 일어났는가를 평가하는 것이다. Chinn

표 4  
형식적 측면의 루브릭

평가항목	1점	2점	3점	4점	5점
전체 구성	간단한 주장과 반대 주장 또는 주장과 주장으로 이루어진 논증과정이다.	자료, 근거 또는 보장이 있으나 반증은 없는 주장으로 이루어진 논증과정이다.	일련의 주장이나 반대 주장을 통해 논증과정하고, 자료, 보장, 보강 등이 있고 때로 약한 반증이 나타난다.	명백하게 확인 가능한 반증이 있는 주장을 제시하지만 중요한 주장이나 주장들이 있다.	한 가지 이상의 반증으로 보다 확장된 논증과정이 나타난다.
주장	주제에 관련된 주장이 없다.	주제와 관련된 주장이지만 명백하게 진술되어 있지 않다.	주장의 초점이 주제에서 제시하는 하나의 논쟁점에 맞춰져 있고 다른 사람들이 이해할 만하다.	주장이 구체적이고 명백하게 주제에 접근하고 있어 모두가 이해할 수 있다.	부차적인 논제를 인식하여 구체적으로 자신의 주장이 어디에 속해 있는지 알고 명확하게 표현하였다.
근거	근거를 제시하지 않거나 제시한 근거가 주장과 연관성이 없다.	주장과 관련하여 근거가 될 수 있는 자료를 제시하였으나 출처에 대한 언급이 없고 부정확하며 충분하지 못하다.	주장과 관련된 근거를 출처를 언급하여 제시하였지만 명확한 관련성을 제시하는 구체적인 설명이 부족하다.	주장과 관련하여 제시한 근거의 중요성을 밝히고 출처를 언급하였으며 제시한 근거가 주장을 뒷받침하기에 충분하다.	근거를 자세히 해석하여 그 근거의 특정 측면이 주장을 어떻게 뒷받침하는지가 분명히 드러나 있고 그 근거가 충분하고 정확하다.
결론	근거를 사용하지 않고 개인적인 관점에 의존하여 결론을 도출하였다.	근거를 사용하고 있지만 최종결론에도 달하는 과정에 있어서 오류가 있다.	결론의 주요 특성들이 근거에 의해 뒷받침되며 결론을 도출하는 정당화 과정과 그에 관한 설명이 나와 있다.	주장과 결론사이에 논리적 일관성이 있고 정당화 과정이 잘 나타나 있으며 반대 입장이 있을 수 있다는 것을 인식한다.	주장과 결론 사이의 일관성이 있고 정당화되어 있으며 결론에 있어서 예외사항이나 반대 논증과정이 나타나 있다.

과 Anderson(2000)은 논증과정을 협동적으로 구축하였는가를 판단하는데 있어서 최종적인 결론을 제시함에 있어서 얼마나 많은 학생들이 기여하였는가를 기준으로 하며 강순민(2006)에 의하면 상대방에게 근거 제시를 요청하는가 상대의 요청에 대해 응답을 하는가의 여부로 학생들이 적극적이었는가를 판단한다는 관점에 따라 평가 척도를 기술하였다(표 6).

② 개방성

이 항목은 나와 상반되는 의견을 가진 상대방의 의견을 받아들이는 태도가 어떠한가에 관한 것을 평가하는 것으로 열린 마음으로 자신에게 불리한 증거와 유리한 증거 둘 모두를 고려하여 최상의 판단을 하는 것을 최상의 상태로 평가하는 관점이다. Chinn과 Anderson (2000)이 협력적인 논증과정을 목표로 제시한 논증과정 네트워크의 특성 중에서 복합적인 관

점을 취하는가에 대한 5가지 태도를 바탕으로 척도를 진술하였다. 이와 더불어 Maloney와 Simon 등 (2006)은 높은 수준의 논증과정은 증거를 재검토하고 보다 강한 논증과정을 받아들일 준비가 되어있다는 특성을 찾아내었고, Felton과 Kuhn(2001) 한 단계 높은 최상의 논증과정은 두 입장 사이를 협상을 통하여 조정하는 것이라 하였다(표 6).

5. 과학 교육 전문가에 의한 타당도 검증

문헌을 통해 추출한 평가 관점과 그에 해당하는 요소들에 대한 정의와 루브릭에 대하여 과학 교육 전문가 4인에게 검토를 받은 후 수정을 거쳐 내용타당도를 의뢰했다. 그 결과 내용타당도 지수 0.96을 얻었고 내용타당도의 일반적인 수락기준인 0.80(Davis, 1992) 이상이므로 매우 높다고 할 수 있다.

**표 5**  
내용적 측면의 루브릭

평가항목	1점	2점	3점	4점	5점
이해	주제에 대해 이해하지 못하고 과학적 개념을 통해 설명하지 않고 부적절한 용어를 사용한다.	주제에 대한 어떤 이해를 바탕으로 약간의 오류가 있지만 과학적 개념을 통해 설명하고자 하며, 일반적인 용어를 사용하지만 부정확하다.	주제에 대한 기초적인 이해를 바탕으로 제한적이지만 과학적 개념을 사용하여 설명하고, 일반적인 용어를 사용해서 정확하게 표현한다.	주제에 대하여 잘 이해하고 제시된 정보들을 과학적 개념을 사용하여 설명하고, 내용 특수적인 용어를 사용하지만 부정확하다.	주제를 완전하게 이해하고 제시된 정보를 능가하는 과학적 개념을 사용하여 설명하고, 내용 특수적인 용어를 정확하게 사용한다.
신빙성	증명할만한 어떠한 증거도 없이 자신이 제시한 근거가 옳다고 단정한다.	과학적으로 정확하지 않은 사실로 확인되지 않았지만 널리 퍼져 있는 일반 상식을 근거로 사용한다.	학생 개인의 경험적인 실험에 의해 알았던 사실을 근거로 사용한다.	아직 반론의 여지가 남아 있는 과학적 이론을 근거로 사용한다.	자연의 진리로서 확고 불변하다고 학자들에 의해 합의된 과학적 법칙을 증거로 사용한다.
추론	인과관계를 형성할 만큼 충분한 양의 증거를 사용하고 있지 않다.	과학적으로 부정확한 증거를 사용하고 인과 관계에 의해 판단하지 않는다.	과학적으로 부정확한 증거를 사용함으로써 잘못된 인과관계를 형성한다.	과학적 증거를 사용하였으나 인과관계가 불안정한 선택을 한다.	정확한 과학적 증거를 사용하여 선택사항들의 장단점을 판단하여 확실한 인과관계에 의한 선택을 한다.

**표 6**  
태도적 측면의 루브릭

평가항목	1점	2점	3점	4점	5점
참여도	참여가 없거나, 불가능하거나, 내키지 않아한다.	자신의 주장을 표현할 뿐 상대방에 대한 호응이 없다.	자신의 주장과 근거를 제시하고 상대방에 대하여 “응”과 같은 단순 용인만을 표현한다.	상대에게 주장에 대한 근거를 제시하도록 요청하고 그에 대한 평가를 한다..	상대방의 근거 요청에 대하여 응답을 할 수 있고 이 과정을 통해 자신의 논증과정을 점검한다.
개방성	상호작용이 이루어질 수 없을 정도로 자신의 논증과정이 명확하지 않고 주장이 뚜렷하지 않다.	자신의 주장을 표현할 뿐 근거를 제시하지 않아서 상호작용이 일어나지 않는다.	상대방의 논증과정을 인정하지 못하고 자신의 논증과정에 상반되는 근거는 고려하지 않는다.	상대방의 논증과정이라도 그 힘을 인정하지만 자신의 입장을 비판하지 못해 생각을 바꾸지 못한다.	상대방의 논증과정을 바탕으로 가능한 모든 근거를 검토한 뒤 자신의 생각을 바꾸어 대안적인 논증과정을 제시한다.

## 6. 루브릭의 예비 적용

과학영재 학생들의 논증과정 평가에 개발된 루브릭을 예비 적용한 결과 제기된 의문점은 다음과 같았다. 첫째 평가 항목 중 ‘형식’ 부분에 전체 구성/주장/근거/결론이라는 항목을 평가함에 있어서 발표자와 질

문자가 있는데 어디에 중점을 두어 평가해야 하는지에 관한 의문이었다. 이는 채점 대상이 발표자이므로 기본적으로는 발표자에 중점을 두고, 4개의 항목 중 특히 전체 구성에 있어서는 발표자와 질문자간에 상호작용을 통해 보다 높은 수준의 논증과정을 구성할 수 있기 때문에 이 상호작용에 중점을 두어 평가하도록 하였다.

둘째 ‘태도’ 부분에 개방성이라는 항목에서는 ‘상대방의 논증과정을 바탕으로 가능한 모든 근거를 검토한 뒤 자신의 생각을 바꾸어 대안적인 논증과정을 제시한다’는 항목이 최상위 점수를 받도록 되어있지만 발표자의 논리가 정확하고 오류가 없어 다른 모듈의 질의에 모두 응답하고 자신의 주장을 모두에게 관철시킨 경우 이 논증 과정이 하위 수준에 해당되는가에 관한 의문이 있었다. 이에 관한 내용은 앞서 구체적인 서술 부분에 기술되어 있다. 개방성 측면에서 중요시 되는 것은 자신의 주장이 틀릴 수 있음을 인정하고 모든 증거들을 객관적으로 평가할 수 있느냐하는 것이므로 질의 응답과정에서 다른 사람의 의견을 충분히 고려했고, 이미 자신의 논증을 구성하는 과정에 있어서도 반대 의견을 고려하여 발표했기 때문에 자신의 주장을 더욱 견고하게 만들 수 있었다고 볼 수 있으므로 자신의 의견을 바꾸지 않았다고 하더라도 이는 상위점수로 채점을 한다. 이와 같이 예비 적용과정에서 루브릭 개발진과 실제 적용한 교사 사이에 생긴 의문점들을 바탕으로 개발된 루브릭의 평가 척도의 서술을 일부 수정하고 보완하였다.

#### IV. 결론 및 제언

과학은 단순히 자연세계에 대한 사실의 축적이 아니라 자연 세계에 대한 설명을 제공하는 이론 구성활동이 수반되므로 현상에 기초가 되는 인과적 원인을 위한 일시적인 설명들을 제공함에 있어서 이론들은 도전과 반론에 늘 직면해 있게 된다. 이러한 특성을 가진 과학교과를 가장 효과적으로 가르칠 수 있는 방법으로써 논증과정을 통한 교육이 매우 유용하다. 따라서 이 연구에서는 과학과의 초등·중등·고등학교 학생들의 논증과정의 질을 평가하여 과학과 교수학습에 도움을 주기 위한 목적으로 관련 문헌 고찰을 통하여 루브릭을 개발하였다. 이 루브릭은 형식, 내용, 태도의 3가지 범주, 각 범주는 형식에 전제 구성, 주장 근거, 결론, 내용에 이해, 신빙성, 추론, 태도에 참여도, 개방성 등의 세부적으로 평가할 하위항목으로 구성되었고, 각각을 5척도로 수준을 나누어 기술되었다.

이 과학적 논증과정 평가 루브릭은 과학교육에 있어서 다음과 같은 교육적 시사점을 얻을 수 있다.

첫째, 루브릭에 대한 항목을 바탕으로 과학적 논증

과정에 대한 이해를 도울 수 있다.

둘째, 학교에서 실행되고 있는 논증과정의 수준을 평가하여 그 결과에 따라 논증과정을 촉진할 수 있는 프로그램 개발이나 교수전략 개발에 활용될 수 있다.

셋째, 학생들이 자신의 논증과정을 스스로 점검해 보거나 동료 학생의 논증과정을 평가하는데 사용할 수 있으므로 논증과정의 질을 높이는데 기준으로 작용할 수 있다.

따라서 이러한 논증과정 평가 루브릭을 이용하여 학생들의 실제 논증과정을 평가함으로써 더 나은 논증과정 수업을 계획하는 자료를 추출할 수 있다. 이 루브릭은 개인의 논증과정 뿐 아니라 그룹간의 논증과정과 같은 상호작용적 논증과정도 평가할 수 있도록 구성되어 있으므로, 형식이나 내용의 범주를 평가할 때는 각 논증자가 제시하는 개개의 논증과정에 초점을 맞추어 평가하고, 태도의 범주는 상호작용적 논증과정에 보다 초점을 맞추어 평가한다. 이를 통해 논증과정에 참여하는 단위에 따라 그룹간의 논증과정일 때는 주체가 되는 그룹을 평가할 때 이 그룹의 논증과정에 대하여 질의응답을 나누는 과정도 함께 평가할 수 있고, 개인간의 논증과정일 때는 개개인의 점수를 채점하여 평가할 수 있다. 이러한 방법으로 평가도구를 활용하여 학생들의 논증과정 평가에 관한 후속연구가 필요하다 하겠다.

또한 논증 과정의 목표가 설득, 조정, 합의 중 무엇이나에 따라 조금씩 다른 평가 기준이 필요하다는 의견에 따라(Nussbaum, 2008) 이러한 관점을 반영하여 보다 과제 특수적인 루브릭의 개발도 의미가 있다. 이는 단일한 분석관점으로는 학생들의 논증과정 학습에 적절한 시사점을 줄 수 없다는 말이기도 한데, 예를 들어 사회 과학적 상황과 과학적 상황에서 이루어지는 논증과정은 조금 다른데, 여기에서 사회과학처럼 가치문제가 개입되는 것은 찬반이 나누어지고, 과학적 상황의 논증과정은 보다 명확하게 사실을 알기 위한 도전과 요청에 의해서 이루어진다고 생각된다. 따라서 이 두 가지 측면의 논증과정이 분리되어 각각에 맞는 평가 관점에 의해 분석되어야 할 것으로 생각된다. 따라서 보다 특수적인 성격을 많이 반영한 평가 관점에 관한 연구가 추가적으로 필요하다.

## 국문 요약

이 연구의 목적은 학생들의 과학적 논증과정 평가를 위한 루브릭을 개발하는 것이다. 루브릭의 개발을 위해 과학교육에서의 논증과정과 관련된 문헌의 분석을 통해 루브릭의 개발 절차 및 평가의 범주를 도출하였다. 루브릭 개발의 일반적인 절차를 따라서 논증과정 평가의 기준을 형식, 내용, 태도의 세 가지 범주를 문헌을 통해 도출하였다. 형식의 범주에는 전체의 구성, 주장, 근거, 결론의 하위 기능으로 세분화하였고, 내용의 범주에는 이해, 신빙성, 추론으로 하위 기능을 세분화하였으며, 태도의 범주에는 참여도와 개방성을 하위 기능으로 정하였다. 이렇게 구성된 각각의 하위 기능의 평가에 대한 기준을 구체적인 5단계로 제시하였다. 문헌 연구를 토대로 개발된 평가 루브릭은 과학교육 전문가 1인과 동료 연구자들의 정기적인 세미나를 통해 점검 받았다. 개발된 초기의 루브릭은 과학교육 전문가 4인에게 의뢰하여 문제점 및 개선점에 대한 검증을 받아 다시 수정되었으며, 최종 완성된 루브릭은 과학교육 전문가 4인에게 타당도를 검증받아 내용타당도지수 0.96으로 높게 나타났다. 개발된 루브릭을 통해 학생들의 논증과정에 대한 이해를 높일 수 있고, 논증과정 프로그램 개발과 논증과정 평가를 위한 기준으로 활용이 가능할 것이다.

## 참고 문헌

- 강순민, 광경화, 남정희 (2006). 논의 과정을 강조한 교수 학습 전략이 중학생들의 인지 발달, 과학 개념 이해, 과학관련 태도 및 논의 과정에 미치는 영향. *한국과학교육학회*, 26(3), 450-461.
- 이범홍 (1998). 토의 토론 학습과 중등학교 과학 교육. 1997년도 교과교육공동연구 결과보고서 (RR97-II-6), 서울: 한국 학술 진흥 재단.
- Herman, J. I., Aschbacher, P. R., & Winters, I. (1992). A practical guide to alternative assessment. Alexandria, VA: Association for Supervision and Curriculum Development. 김경자(역) (2000). 수행평가 과제 제작의 원리와 실제. 서울: 이화여자대학교 출판부.
- Andrews, R., Costello, P., & Clarke, S. (1993). Improving the quality of argument 5-16: Final Report. Hull, UK: Esmee Fairbairn Charitable Trust/University of Hull.
- Arter, J. (2000). Rubrics, scoring guides, and performance criteria: Classroom tools for assessing and improving student learning. paper presented at the Annual Meeting of American Educational Research Association, ERIC Document Reproduction Service Mo. ED446100.
- Arter, J., & McTighe, J. (2001). Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance. Thousand Oaks, CA: Corwin Press.
- Astin, A. W. (1993). Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education. New York: Macmillan.
- Aufscjnaiter, C. V., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, 45(1), 101-131.
- Banta, T. W., Lund, J. P., Black, K. E., & Oblander, F. W. (1996). Assessment in practice: Putting principles to work on college campuses. San Francisco: Jossey-Bass.
- Bazerman, C. (1988). Shaping written knowledge: The genre and activity of the experimental article in science. Madison: University of Wisconsin Press.
- Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning for the Web with KIE. *International Journal of Science Education*, 22(8), 797-817.
- Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education*, 92(3), 473-498.
- Carr, J.F., & Harris, D.E. (2001).

Succeeding with standards: Linking curriculum, assessment, and action planning. Alexandria, VA: Association for Supervision & Curriculum Development.

Chan, C. K. K. (2001). Peer collaboration and discourse patterns in learning from incompatible information. *Instructional Science*, 29, 443-479.

Chinn, C. A., & Anderson, R. G. (2000). The structure of discussions that promote reasoning. *Teachers College Record*, 100(2), 315-368.

Clark, D., & Sampson, V. (2006). Personally-seeded discussions to scaffold online argumentation. *International Journal of Science Education*, 29(3), 253-277.

Davis, L. (1992). Instrument review: Getting the most from your panel of experts. *Applied Nursing Research*, 5(4), 194-197.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287-312.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88(6), 915-933.

Felton, M. K., & Herko, S. (2004). From dialogue to two-sided argument: Scaffolding adolescents' perspective writing. *Journal of Adolescent & Adult Literacy*, 47(8), 672-683.

Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skill. *Discourse Processes*, 32(2&3), 135-153.

Goodrich, H. (1996). Understanding rubrics. *Educational Leadership*, 54(4), 14-17.

Hogan, K., Nastasi, B. K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussion. *Cognition and Instruction*, 17(4), 379-432.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students and scientists' reasoning about conclusions. *Journal of Research in Science Teaching*, 38(6), 663-687.

Jimenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "Doing the Lesson" or "Doing Science": Argument in high school genetics. London: John Wiley & Sons.

Kelly, G. J., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students use of evidence in writing. *Science Education*, 86, 314-342.

Kelly, G. J., & Bazerman, C. (2003). How students argue scientific claims: a rhetorical-semantic analysis. *Applied Linguistics*, 24(1), 28-55.

Kanselaar, G., Erkens, G., Andriessne, J., Prangma, M., Veerman, A., & Jaspers, J. (2002). Designing argumentation tools for collaborative learning. In P. A. Kirschner, S. J. Buckingham-Shum & C. S. Carr (Eds.), *Visualising argumentation: Software Tools for Collaborative and Educational Sense-making*. (pp. 51-73). London: Springer.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319-337.

Lawson, A. E. (2003). The nature and development of hypothetico-predictive argumentation with implications for science teaching. *International Journal of Science Education*, 25(11), 1387-1408.

Maloney, J., & Simon, S. (2006). Mapping children's discussions of evidence in science to assess collaboration and argumentation. *International Journal of Science Education*, 28(15), 1817-1841.

Marttunen, M. (1994). Assessing argumentation skills among Finnish university students. *Learning and Instruction*, 4, 175-191.



McCann, T. M. (1989). Student argumentative writing: Knowledge and ability at three grade levels. *Research in the Teaching of English*, 23(1), 62-76.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Science*, 15(2), 153-191.

McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In Lovett, M. & Shah, P. (Eds.) *Thinking with data: The proceedings of the 33rd Carnegie symposium on cognition*. Mahwah, NJ: Lawrence Erlbaum Associate, Inc.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal*, 25(1), 95-111.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2): 13-23.

Mitchell, S. (2001). What is this thing called argument? In R. Andrews & S. Mitchell (Eds.), *Essays in argument*. London: Middlesex University Press.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10).

Naylor, S., Keogh, B., & Downing, B. (2007). Argumentation and primary science. *Research in Science Education*, 37, 17-39.

Newton, P., Driver, R. & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education*, 21, 553-576.

Nussbaum, E. M. (2008). Collaborative

discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology*, 33, 345-359.

O'Donnell, A. M., & King, A. (Eds.). (1998). *Cognitive perspectives on peer learning*. Mahwah, NJ: Erlbaum.

Osborne, J. F., Erduran, S., Simon, S., & Monk, M. (2001). Enhancing the quality of argumentation in school science. *School Science Review*, 82(301), 63-70.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Researcher in Science Teaching*, 41(10), 994-1020.

Perella, J. (1987). *The debate method of critical thinking: An introduction to argumentation*. Dubuque, IA: Kendall/Hun.

Pollock, J. L. (1987). Defeasible reasoning. *Cognitive Science*, 11, 481-518.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B., Steinmuller, F., & Leone, T. J. (2001). BGUILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Erlbaum.

Sadler, T. D., & Fowler, S. R. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education*, 90, 986-1004.

Sampson, V., & Clark, D. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, 92(3), 447-472.

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition And Instruction*, 23(1), 23-55.

Schafer, W. D., Swanson, G., Bene, N., &

Newberry, G. (2001). Effects of teacher knowledge of rubrics on student achievement in four content area. *Applied Measurement in Education*, 14, 151-170.

Schwarz, B., & Glassner, A. (2003). The blind and the paralytic: Supporting argumentation in everyday and scientific issues. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 227-260). Dordrecht, the Netherlands: Kluwer.

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to Teach Argumentation: Research and development in the Science Classroom. *International Journal of Science Education*, 28(2-3), 235-260.

Siegel, M. A., Hynds, S., Siciliano, M., & Nagle, B. (2006). Chapter 7. Using rubrics to foster meaningful learning. In McMahon, M., Simmons, P., Sommers, R., Debates, D., & Crawley, F. (Ed.) *Assessment in Science*. (pp. 89-106). Arlington: NSTA press

Takao, A. Y., Prothero, W. A., & Kelly, G. J. (2002). Applying argumentation analysis to assess the quality of university oceanography students' scientific writing. *Journal of Geoscience Education*, 5(1), 40-48.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Van Gelder, T. (2002). Argument mapping with reasonable. *Philosophy and Computers*, 2(1), 85-90.

Veerman, A., Andriessen, J., & Kanselaar, G. (2002). Collaborative argumentation in academic education, *Instructional Science*, 30, 155-186.

Waltman, K., Kahn, A., & Koency, G. (1998). Alternative approaches to scoring: The effect of using different scoring methods on the validity of scores from a performance assessment. CSE Technical Report 488. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Yore, L. D., & Treagust, D. F. (2006). Current realities and future possibilities: language and science literacy-empowering research and informing instruction. *International Journal of Science Education*, 28(2-3), 291-314.

Zemal-Saul, C., Munford, D., Crawford, B., Friedrichsen, P., & Land, S. (2003). Scaffolding preservice teachers' evidence-based arguments during an investigation of natural selection. *Research in Science Education*, 32, 437-463.

Zimmaro, D. M. (2004). Developing grading rubrics. Retrieved from <http://www.utexas.edu/academic/mec/research/pdf/rubricshandout.pdf>

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39, 35-62.