

과학과 자기보고식 정의적 영역 평가의 정확성에 영향을 주는 요소 탐색

정수임¹, 신동희^{2*}¹은계중학교, ²이화여자대학교

Exploring the Factors Influencing on the Accuracy of Self-Reported Responses in Affective Assessment of Science

Sue-Im Chung¹, Donghee Shin^{2*}¹Eungye Middle School, ²Ewha Womans University

ARTICLE INFO

Article history:

Received 13 February 2019

Received in revised form

12 March 2019

2 May 2019

Accepted 5 May 2019

Keywords:

science education assessment, science-related affective characteristic, self-reporting test tool, response bias, error in test, science-specific error, improvement of test tool

ABSTRACT

This study reveals the aspects of subjectivity in the test results in a science-specific aspect when assessing science-related affective characteristic through self-report items. The science-specific response was defined as the response that appear due to student's recognition of nature or characteristics of science when his or her concepts or perceptions about science were attempted to measure. We have searched for cases where science-specific responses especially interfere with the measurement objective or accurate self-reports. The results of the error due to the science-specific factors were derived from the quantitative data of 649 students in the 1st and 2nd grade of high school and the qualitative data of 44 students interviewed. The perspective of science and the characteristics of science that students internalize from everyday life and science learning experiences interact with the items that form the test tool. As a result, it was found that there were obstacles to accurate self-report in three aspects: characteristics of science, personal science experience, and science in tool. In terms of the characteristic of science in relation to the essential aspect of science, students respond to items regardless of the measuring constructs, because of their views and perceived characteristics of science based on subjective recognition. The personal science experience factor representing the learner side consists of student's science motivation, interaction with science experience, and perception of science and life. Finally, from the instrumental point of view, science in tool leads to terminological confusion due to the uncertainty of science concepts and results in a distance from accurate self-report eventually. Implications from the results of the study are as follows: review of inclusion of science-specific factors, precaution to clarify the concept of measurement, check of science specificity factors at the development stage, and efforts to cross the boundaries between everyday science and school science.

1. 서론

최근 들어 취업 시장과 대학 입학 선발 전형에서 지원자의 잠재 가능성을 종합적으로 평가할 수 있는 전형 방법들이 예전의 단순 지식 평가를 대체하고 있다(Lee & Moon, 2015; Kim *et al.*, 2015; Yoo, Lee, & Kim, 2012). 인성과 적성, 직무 적성, 역량 평가와 같은 필기시험 뿐 아니라 토론, 면접, 발표, 프로그램 참여 등으로 전형 방법을 다양화하면서 지원자의 인지적 능력 뿐 아니라 정의적 특성도 종합적으로 평가한다. 이들 평가에서 널리 사용되는 자기보고식 반응을 기반으로 한 표준화 검사는 개인의 정의적 특성을 판단하는 다른 대안적 자료보다 유용(Kim & Lee, 2006; Hough *et al.*, 1990)하지만 참여자의 진실한 반응을 전제한다는 점에서 어느 정도 한계가 있다. 특히 경향성이 있거나 편향된 반응, 정확성을 담보하지 못한 응답 등은 도구의 타당도와 신뢰도에 영향을 미치는 문제로 인식된다. 따라서 자기보고식 검사 결과를 근거로 중요한 의사 결정이 이루어지는 교육학과 심리학, 사회 과학과 정신 의학 등의 분야에서는 이러한 문제를

감소시키기 위한 방안을 다방면으로 탐색하고 있다(Bae, Lee, & Ham, 2015; Choi, *et al.*, 2015; Ferrando & Chico, 2001; Joo, *et al.*, 2001; Reise & Flannery, 1996; Paulhus, 1984). 사실, 자기보고식 검사로 측정된 점수가 눈에 보이지 않는 학생의 능력이나 특성을 정확히 나타내는지의 의문은 평가의 본질에 대한 문제로 심리 측정과 교육 측정 분야의 오랜 관심 주제였다. 응답자의 진점수에는 측정 오차가 묻어 있어 결국 관찰 점수만 알 뿐 아무것도 모르는 것과 마찬가지로(Rust & Golombok, 2014)이 진점수 이론의 의미이자 출발점이다. 학생들이 자기보고식 검사에서 얻은 점수는 관찰 점수이므로 진점수는 오차를 이해해야 추정이 가능하다. 결국 정확한 자기 보고를 방해하는 요소는 오차를 밝히는 선행 연구들에서 그 실마리를 찾을 수 있다.

Reise & Flannery(1996)는 자기보고식 심리 검사에서 통계적으로 수행된 반응자 적합도 검사에서 부적합 반응이 나타나는 원인을 반응 왜곡, 무선 반응, 특성 분산(traitedness variation)으로 분류했다. Hui & Triandis(1985)는 이러한 비전형적 반응의 경향성을 묵중 경향성, 반응 범위, 극단 반응으로 정리했다. 이들의 관심은 자기보고식 검사

* 교신저자 : 신동희 (donghee@ewha.ac.kr)

** 이 논문은 정수임의 2018년도 박사 학위 논문의 데이터를 활용하여 재구성하였음.

<http://dx.doi.org/10.14697/jkase.2019.39.3.363>

에서 나타나는 부적합 반응을 탐지하기 위한 적합도를 고안하는데 있었다. 이러한 시도는 문제점을 탐지하고 유형화·가시화함으로써 주관적 자기보고식 검사 도구의 한계를 극복할 수 있는 방향을 제시했고 해결책을 모색하는 후속 연구로 활발히 연계되고 있다(Ferrando & Chico, 2001; Zickar & Drasgow, 1996). 정확한 자기 보고에 영향을 미치는 경향성을 찾으려는 노력은 오래 전으로 거슬러 올라간다. Cronbach(1946)는 같은 내용이라도 다른 상황이나 조건에서 일관성 있게 다른 응답을 하는 경향을 반응 세트(response set)로 정의하며 측정 구인(構因)에 상관없이 검사 상황에 따라 정확하지 않은 결과를 산출하는 결과에 주목했다. 반응 왜곡(response bias)은 측정하려는 목적과 관련 없이 다른 기준에 의해 응답하려는 경향성이며 정확한 자기 보고를 방해하는 체계적 경향성이다(Cronbach, 1946; Hui & Triandis, 1985; Schulz *et al.*, 2008). Messick(1991)은 반응 왜곡을 검사에 상관없이 개인적 특성 때문에 나타나는 반응 양식(response style)과 검사 상황에서 일시적 방해 요인이나 동기적 요인 때문에 발생하는 반응 세트(response set)로 분류했다. 이 두 가지 분류 방식에 의하면, 앞에 언급한 비전형적 반응을 반응 왜곡 현상의 범주에 넣어 재조직할 수 있다. 반응 세트는 검사를 실시하는 상황에 따라 발생하는 현상이므로 검사를 해야 할 필요성을 느끼지 못한 채 아무렇게나 반응하는 무신 반응, 어떤 동기 때문에 현실을 왜곡하거나 위장해야 하는 반응 등은 반응 세트에 해당한다. 반응 양식은 개인의 검사 상황에서 비교적 안정적인 개인의 특성이며 측정 경향성과 함께 반응 범위나 극단 반응을 포함하는 특성 분산이 이에 해당한다. 가장 대표적 반응 왜곡의 사례는 사회적 바람직성이다.

사회적 바람직성은 검사에서 측정하는 내용이 사회적으로 인정받거나 바람직하다고 판단되는 항목에 동의함으로써 자신을 긍정적으로 묘사하려는 경향성이다(Crowne & Marlow, 1960; Stricker, 1963; Zerbe & Paulhus, 1987; Paulhus, 2002). 피험자는 진실하게 자신을 보고하기보다 사회적, 규범적 압력에 부응해서 문항에 응답하려는 경향성을 보이므로 좁게는 개인의 정확한 특성 파악이 어렵거나 나아가서는 개념을 측정하려는 시도와 개념 간 상관성 연구에 장애가 되기도 한다. 사회적 바람직성이 반영된 반응 결과는 교육적, 사회적 현상의 이해를 어렵게 하고, 현상 간 상관이나 인과적 유추에 오류를 유발해서 잘못된 결과를 도출할 수 있다(Crowne & Marlow, 1960; Cronbach, 1946). 최근 들어 성격 검사, 심리 검사나 상담 분야, 선발과 채용을 위한 인·적성 평가 등 검사 결과의 파급력이 큰 분야에서 실용적 목적으로 반응 왜곡을 감지하려는 시도가 활발히 수행되고 있는 이유다(Son, Cha, Kim, 2007; Au, 2007). 정의적 특성을 주로 자기보고식 검사 문항을 사용해서 측정하는 과학과 정의적 영역 평가에서 학생들의 반응을 왜곡시킬 수 있는 여러 요인들 사회적 바람직성을 포함한 반응 양식과 반응 세트에 대한 비판적 고찰이 필요한 시점이다.

한편, Scheuneman(1984)은 검사에서 편파의 원인을 탐색하는 이론적 틀을 소개하면서 진점수와 측정 오차에 차별 효과 항을 추가한 관찰 점수 모형을 제안했다. 차별 효과 항은 문항이나 검사 자체의 특성이나 결함 요인과 피험자 특성 요인으로 이루어진다. 이는 Chung & Shin(2017)이 과학 관련 정의적 특성 검사에서 발견한 학생들의 두 가지 반응 유형과 관련 된다. 검사 개발자가 의도한 개념을 다르게 인식하는 피험자 반응은 Scheuneman(1984)이 차별 요인으로 제시한

문항 및 검사 자체의 요인과 관련 있고, 진짜 속마음과 달리 표현하는 반응 유형은 피험자 집단의 특성 요인에 해당한다. Scheuneman(1984)은 도구의 결함의 예로 문항이나 지시 사항에 있어 불충분한 설명, 모호한 지시, 혼동 용어 등의 불확실성과 함께 집단에 따라 차별적 매력을 지닌 오답 등을 지적했다. 도구를 이루는 개념과 언어가 불확실하고 모호할 때, 다양한 경험과 성격을 지닌 학생들은 각자 경험한 역사 안에서 마땅하다고 생각되는 상황을 떠올리고 그에 맞는 반응을 결정한다. 비슷한 문화와 상황을 공유했던 학생 집단이라면 특정한 행동을 더욱 많이 선택할 가능성이 있는데, 이는 집단에 따라 차별적 기능을 하는 문항으로 작용할 수 있다. 문화와 상황에 따라 행동에 가하는 제약이 다르면 특정한 행동으로 정적 혹은 부적 강화를 받으면서 집단 차이가 발생하기 때문이다(Anastasi, 1988). 이처럼 도구의 결함으로 인해 집단에 따라 다른 반응들이 유도되는 경우는 정확한 자기 보고에 영향을 미치는 오차 요소에 해당한다.

지금까지 심리 측정학계에서 밝혀내고 제기한 일반적 문제들은 과학 관련 정의적 평가에서도 마땅히 적용될 것이나, 과학 교과에서 특이하게 나타나는 문제들은 과학 교육적 관점에서 밝혀내야 할 몫이다. 최근 과학 교육계에서도 보다 타당하고 신뢰로운 정의적 영역 검사 도구를 개발하려는 시도가 활발히 진행되고 있다. Chung & Shin(2016)의 과학 교육 평가 동향 연구에서는 검사 도구를 창안하면서 통계적 검증 뿐 아니라 이를 현장에 투입해 적용하는 종합적 연구가 증가하는 추세를 보고했다. 특히 과학 관련 정의적 검사 도구를 개발하는 과정에서 측정학적 중요성이 부각됨에 따라 다양한 통계 방법을 사용해서 타당화하는 연구가 발표되었다. 내적 일관성과 함께 타당성을 검증하는 방법으로 요인 분석과 함께 구조 방정식 분석 및 모형의 적합도를 산출했다(Shin, Ha, Lee, 2016; Shin, Ha, Lee, 2014; Shin *et al.*, 2017; Chun *et al.*, 2015). 문항 반응 이론으로 자료를 처리하는 경우는 일반화 타당도로서 차별적 문항 기능을 수행하고(Shin, Ha, Lee, 2016; Shin, Ha, Lee, 2014), 기술통계와 집단 간 차이 분석 및 상관 분석으로 타당화한 경우(Kim & Kim, 2016)도 있었다. Shin, Ha, & Lee(2014)는 과학 관련 정의적 영역 검사로서 ‘융합에 대한 태도 검사 도구’를 개발하면서, 집단별 비교와 함께 Rasch 모형을 적용한 차별적 문항 기능을 확인했다. 이는 검사 도구의 일반화 타당도를 확보(Messick, 1995)하기 위한 방안이며, 구조 방정식 분석을 통해 측정 모형의 동등성 검증을 실시하기도 했다. 이상 과학과의 정의적 영역 검사 도구는 주로 자기보고식의 리커트 형태가 사용되며 과거에 비해 측정학적 양호도가 중시되면서 각종 통계 기법을 활용하는 경우가 증가했으나, 자기보고식 검사 문항이 지니는 오류에 대한 명시적 관심은 거의 드러나지 않았다.

과학과 정의적 평가는 태도, 가치, 동기, 신념, 흥미 등이 특정 자극 대상인 ‘과학’과 관련해서 학생들이 나타내는 반응을 다룬다. 학생들은 일상과 학습의 경험을 통해 과학에 대한 자신의 관점과 정서, 선호를 형성하고 과학의 특성과 본성에 대한 나름의 이해를 내면화하고 있다. 이들이 지닌 과학 관련 인지적·정의적 인식을 측정하기 위해, 검사 도구는 과학을 중심으로 한 다양한 맥락을 문항으로 제시하면서 이에 반응하는 학생들의 인식을 포착한다. 본 연구에서는 과학 관련 개념이나 인식을 측정하려할 때 학생이 지닌 과학 특성, 본성에 대한 인식 등이 원인이 되어 나타나는 반응을 ‘과학 특이적 반응’으로 보았다. 학생이 나타낸 과학 특이적 반응은 측정 구인과 부합하는 경우도

있으나, 측정 구인을 방해할 경우 오차로 작용하면서 정확성에 영향을 미칠 수 있다. 과학 및 과학 학습에 대한 효능감, 선호, 흥미, 중요성 인식 등을 묻는 질문에 반응할 때 학생들이 지닌 과학 인식과 경험은 반응의 방향과 크기를 결정한다. 이때 학생이 지닌 과학 특성을 이끌어내어 측정하는 것이 과학 관련 검사 도구의 중요한 역할이다. 이는 비단 과학과 뿐 아니라 특정 대상, 예컨대 수학이나 사회과 같은 교과를 중심으로 측정하는 모든 검사 도구에 내재한 속성이므로 대상에 대한 특이적 반응이 특별히 검사 도구의 타당도에 영향을 미칠 것으로 문제시 하지 않았다. 그러나 과학 관련 자기보고식 문항에 반응하려는 순간, 학생이 지닌 과학 특성 때문에 오히려 문항이 측정하려는 목적이나 구인에 맞지 않는 반응, 자기 보고의 정확성을 저해하는 반응이 나타난다면, 측정하려는 대상의 특성 때문에 나타나는 오차인지 여부를 타당도와 관련해 점검해 볼 필요가 있다. 이는 검사 도구의 목적이 과학 관련 정의적 특성을 측정하려 하기 때문에 상대적으로 간과해왔던 오차 요소일 수 있다. 학생이 내면화한 과학 특성이 반응에 어떤 방향으로 작용하는지 그 과정을 살펴보아야 한다. 이로써 특정 영역에서 정의적 특성과 인식을 측정하려는 평가 도구는 보다 정확하고 타당하게 개선될 수 있도록 수정할 대상과 방향을 특정할 수 있게 된다.

본 연구에서는 학생들이 과학과 정의적 영역 검사에 응할 때 과학 특이적 반응으로 오히려 정확한 자기 보고에 영향을 받게 되는 사례를 탐색하고자 한다. 이에 앞서 ‘정확한 자기 보고’의 의미는 측정하려는 구인에 대한 정확하거나 솔직한 반응으로 정의했다. 이는 오차 요소가 개입되지 않은 가상적 반응 혹은 그 결과로 나타난 점수를 의미한다. 정답이 있는 인지적 검사 도구에 비해 정의적 영역은 정확함의 영역을 특정하기 어렵다는 측면에서, 측정하려는 구인과 다른 반응은 같은 점수로 나타나더라도 정확한 자기 보고에서 벗어난 반응으로 간주될 수 있다. 이러한 사례들로부터 측정하려는 구인에서 벗어나 정확한 자기 보고에 오차로 작용하는 요소들을 유형화함으로써 향후 평가 도구 개발과 개선에 실용적 정보를 제공하고, 특히 과학 특이적 반응에 의한 오차가 학생들이 경험한 일상과 학습을 통해 형성되었다는 점에서 과학 교과 교육 측면의 시사점을 제안한다. 이를 위한 본 연구의 문제는 다음과 같다.

첫째, 과학 관련 정의적 영역 평가에서 학생이 지닌 과학에 대한 관점, 경험, 학습 등의 과학 특성이 검사 도구와 상호작용하며 나타나는 과학 특이적 반응 사례는 어떠한 양상으로 나타나는가?

둘째, 과학 특이적 반응이 오히려 정확한 자기보고를 방해하며 영향을 미치는 사례는 어떠한 경우가 있으며, 이들로부터 도출된 오차 요소는 무엇인가?

II. 연구 방법

1. 분석 대상

연구 참여자는 경기도 중소 도시의 인문계 남녀 공학 고등학교 1, 2학년 학생 649명이다(Table 1). 이들은 과학 관련 정의적 특성과 심리 특성에 대한 검사를 실시했고, 검사 전후 각 10명, 34명 등 총 44명이 검사와 관련한 내용으로 면담에 참여했다. 연구에 참여한 학생들은 공통 교육 과정 대상자인 1학년 328명과 선택 교육 과정에 들어서 과학 교육 과정의 변화가 뚜렷한 2학년 321명이다. 특히 2학년 학생은 인문·사회 과정과 자연·공학 과정으로 진로 과정이 달라지면서 과학에 대한 도구적 동기가 구분될 뿐 아니라 이수해야 하는 과학 과목의 수, 시간 등의 차이가 분명히 드러나는 집단이다. 따라서 서로 다른 진로 과정을 선택한 학생들의 과학에 대한 인식 차이와 그로 인한 반응을 분석할 수 있다.

검사 전 실시한 사전 면담에 참여한 학생들은 성별, 학업 성취도를 고려했고, 오랜 시간 소요되는 발성 사고 면담에 성실하게 참여하기 희망하는 10명의 학생을 선정했다. 검사 후 면담 참여 학생 34명은 검사 결과 <과학>과 <과학 과목>의 용어 반응, 사회적 바람직성 반응, 특이적 반응 경향(극단, 중간 반응) 등을 종합적으로 고려해서 선정했다. 검사에 참여한 649명의 학생들은 검사 참여 및 학업 성취도 제출에 대한 보호자의 동의서를 제출했고, 면담 참여자는 면담 내용의 녹취와 전사에 대한 별도의 보호자 동의서를 제출했다. 면담에 참여한 학생들의 이름은 모두 가명으로 처리했다.

2. 연구 절차

이 연구는 2017년 11월부터 2018년 2월까지 수행되었으며, 분석 자료 중 인지적 성취 자료는 연구 시점보다 약 3~8개월 앞선 2017년 1학기말 학업 성취도를 사용했다. 1, 2학년 학생 733명에게 연구자가 직접 연구에 대한 설명을 실시한 후 검사 참여와 학업 성취도 제공에 학부모와 학생의 동의를 득한 649명의 자료를 분석했다. 검사를 실시하기 전 10명의 학생들을 대상으로 실시한 사전 면담에서는 검사 문항에 대한 피험자의 인식을 사전에 알아보기 위해 문항을 소리 내어 읽고, 점수를 주는 과정을 이야기하도록 했다(Abd-El-Khalick *et al.*, 2015; Fives *et al.*, 2014; Osborne, Simon, & Tytler, 2009). 학생들이 이야기한 내용은 검사에 반영되어 수정·보완되었는데, 리커트 4점에서 5점 척도로 바꾸어 중간 반응을 알아보도록 척도를 조정하고, 설문에서 연상하는 과목을 조사하는 2문항을 추가했다. 총 649명에게 정의적·심리적 특성 등을 알아보는 검사를 실시한 후, 검사 결과를 근거로 34명의 면담 참여자를 선정했다. 면담 참여 학생의 선정에는

Table 1. Research participants

	연구 참여자 수(명)											계		
	1학년			2학년(인문·사회 과정)			2학년(자연·공학 과정)			남	여			
	남	여	계	남	여	계	남	여	계					
정의적 특성 등 검사	113	115	328	57	121	178	158	85	243	328	321	649		
면담	검사 전	·	·	·	·	·	5	5	10	26	18	34		
	검사 후	4	3	7	·	4	4	17	6	23				

같은 내용의 문항임에도 <과학>을 <과학 과목>으로 교체했을 때 점수 차이가 크거나 차이가 없는 경우, 사회적 바람직성이 높거나 낮은 경우, 극단 반응이나 중간 반응이 우세한 경우, 검사 중 연상 과목이 문항마다 다르다고 응답한 경우 등을 종합적으로 고려했다. 면담 내용은 모두 녹음되고 전사되었으며, 면담 시간은 사전 면담의 경우 총 5회, 1인당 평균 면담 시간은 29분, 사후 면담은 총 21회, 1인당 평균 면담 시간은 21분 소요되었다. 649명의 검사 결과 자료와 학업 성취도의 양적 자료는 PASW Statistics 18.0 프로그램에 의해 기술 통계와 집단 차이 및 상관 분석을, 44명의 면담 자료는 근거 이론의 개방 코딩 방법을 부분 활용해서 개념과 범주를 추출했다. 양적 자료와 질적 자료 분석을 통해 정확한 자기 보고에 영향을 미치는 과학 특이적 요소를 도출했다.

3. 분석 자료

Table 2는 연구에서 분석한 양적 자료와 질적 자료의 내용을 정리한 것이고 Table 3은 검사에서 사용한 문항을 구인 별로 정리한 것이다. 양적 자료로는 PISA 2006의 태도 문항 26문항을 이용해서 정의적 특성을 측정했다. 특히 이들 문항 중 <과학>이라는 용어를 <과학 과목>으로 대체해서 용어 혼동 여부를 탐지해 보는 6문항을 추가했다. 예를 들어 ‘과학은 사회에 유익한 것이다’라는 문항을 ‘과학 과목은 사회에 유익한 것이다’로 교체한 문항을 추가하고 이들 짝 문항에 응답한 차이를 비교해 보는 것이다. 상당수의 학생들은 두 문항을 같은 것으로 인식해서 같은 점수를 주거나, 점수를 다르게 주는 경우 두 용어에 대한 다른 인식을 드러내기도 했다. 학생들의 심리적 특성을 탐지하는 사회적 바람직성 문항은 미국과 독일에서 성인을 대상으로

로 타당화한 SDS-17(Stöber, 2001)을 한국인에 타당화한 SDS-9(Bae, Lee, Ham, 2015)을 활용했고, 본 검사에서 신뢰도는 .692다(Table 2, Table 3). 사회적 바람직성 문항의 예로 ‘나는 때때로 나의 불쾌한 기분을 남들에게 푼다’와 같은 문항에 ‘전혀 그렇지 않다’고 응답하는 학생은 사회에서 좋은 인상으로 보이고 싶은 심리가 강하다고 판단하고, 사회적 바람직성 척도는 높게 나타난다. 이밖에도 사전 면담 참여 학생들의 의견을 반영해서 설문에 응할 때 특정 과목을 연상하는 지 여부를 선택형으로 질문에 포함했다.

학업 성취도는 2017년 1학기의 학기말 성취도 점수를 분석했다. 2번의 지필 평가와 수행 평가를 합산한 점수를 T점수로 환산했다. 1학년은 과학, 2학년 인문·사회 과정 학생들은 지구과학 I, 자연·공학 과정 학생들은 한 학기 동안 물리 I, 화학 I, 생명과학 I, 지구과학 I 을 모두 이수했으므로 이들의 T점수를 모두 활용했다(Table 2). 이로써 인문·사회 과정 학생들은 성적이 산출된 최근 과목을 지구과학 I으로 인지했고 이를 근거로 검사에 임했다는 반응이 많았다.

Table 4는 학생들이 과학과 관련한 정의적 특성 검사 과정에서 느끼는 어려운 점, 과학과 학교 과학에 대한 생각, 반응에 대한 개인적 경향성 등을 면담한 질문들을 정리한 것이다. 사전 면담에서는 주로 검사에서 사용한 과학과 과학 과목에 대한 생각을 집중적으로 질문했고, 사후 면담에서는 검사 도구와 개인적 특성에 초점을 맞추어 질문했다. 녹음되어 전사된 내용은 연구자 중 1인이 1차 부호화한 내용을 공동 연구자가 검토 후 협의해서 확정하고, 현상을 명명한 1차 부호를 연구자 2인이 2~3차 부호로 묶어서 좁혀 나갔다. 분석 결과는 질적 연구 수행 경험이 있는 과학 교육 전문가 5인의 검토를 2회에 걸쳐 받았다. 1차 협의회에서는 1차 부호화가 타당하게 명명되었는지 비교했고, 검토자간, 검토자와 연구자간 의견이 다른 경우 맥락 이해를

Table 2. Analysis materials

분석 대상	내용
정의적 영역 검사 및 사회적 바람직성*, 기타 인식 등	- 정의적 영역 검사 26문항(PISA 2006 태도 문항 발췌)(Lee, Sohn, & No, 2007) - 사회적 바람직성 9문항(SDS-9)(Bae, Lee, & Ham, 2015) - 용어 교체 6문항(PISA 2006 태도 문항 중 <과학>을 <과학 과목>으로 교체) - 설문 시 특정 과목 인식 2문항
양적 자료	- 기술 통계, 집단 차이, 상관관계 분석 및 문항 적합도 분석 - PASW Statistics 18.0, WINSTEPS 3.58.1(Linacre, 2005)
학업 성취도	- 649명의 과학 과목(과학, 물리 I, 화학 I, 생명과학 I, 지구과학 I)의 1학기말 성취도 - 1학년은 과학, 2학년 인문과정은 지구과학 I, 2학년 자연과정은 4과목(물리 I, 화학 I, 생명과학 I, 지구과학 I) 성취도 활용 - 표준화 점수 활용(T점수 환산)
질적 자료	학생 면담 - 검사 전 면담 10명, 검사 후 면담 34명, 총 44명 면담 실시 - 검사 전 면담 결과로 검사 도구 수정 - 검사 결과로 검사 후 면담 대상자 선정 - 근거 이론의 개방 부호화 방법 활용해서 부호화, 범주화

* 사전 면담에서 Stöber(2001)의 SDS-17을 발성 사고로 적용했으나, 문항 수가 많다는 의견에 따라 SDS-9으로 대체함.

Table 3. Contents of questionnaire

항목	내용	형태
과학 관련 정의적 영역 특성*	과학 과목에 대한 자아 개념, 즐거움(내적 동기), 과학의 가치(일반적, 개인적), 도구적 동기	리커트 5점
용어 혼동 탐지 문항	<과학>을 <과학 과목>으로 교체, 과학의 가치 문항 중 6문항 변형	리커트 5점
사회적 바람직성(SDS-9)**	긍정형 바람직성, 부정형 바람직성	리커트 5점
기타 인식 등	설문 상황에서 떠올리는 과학 및 과학 과목	선택형

* 각 구인의 신뢰도(Cronbach α)는 .856~.942 범위에 있음.

** Cronbach α= .692

Table 4. Contents of questions for semi-structured interview

구성		질문 내용
사전 면담	문항 발생 사고	문항을 소리 내어 읽은 후 답을 고르고, 그렇게 반응한 과정을 소리 내어 말하기
	용어 인식	설문 중 떠올린 과학 영역 및 과목, 세부 과목 연상 여부 및 연상 과목, 세부 과목 중 스스로 생각한 과학에 가까운 과목의 종류와 그렇게 생각한 이유, 내가 생각하는 과학, 학교 과학과 내가 생각한 과학의 차이와 공통점, 학교 과학과 내가 생각한 과학에 대한 느낌과 생각, 과학이 타과목과 구별되는 특이한 측면, 설문에서 어려운 점 등
	반응 왜곡	스스로 생각하는 과학 성적과 느낌의 정도, 면담이나 설문에서 솔직한 반응 여부, 솔직하게 반응하기 어려운 문항이나 내용, 솔직한 마음보다 점수를 높거나 낮게 준 문항과 그 이유, 점수를 높거나 낮게 주는 전반적 경향성 등
사후 면담	용어 구분 인지	검사 도구에서 <과학>과 <과학 과목> 용어 혼재 인식 여부, 설문 시 지시문 읽었는지 여부
	검사 도구	<과학>과 <과학 과목>에 대한 차이 인식, 내가 생각하는 과학, 학교 과학과 내가 생각한 과학의 차이와 공통점, 과학 과목에서 떠올린 세부 과목, 특정 과목을 연상하게 되는 문항 존재 여부
	개인적 특성	자신의 과학 성적과 느낌 인식, 진로와 과학 연계 여부, 설문에 응할 때 솔직성에 대한 생각, 솔직하게 반응하기 어려운 문항이나 내용, 솔직한 마음보다 점수를 높거나 낮게 준 문항과 그 이유, 점수를 높거나 낮게 주는 전반적 경향성, 진로 선발 과정에서 실시되는 설문에 응답하는 경향성 등

위한 전사본을 검토한 후 수정했다. 2차 회의에서는 개념, 하위 범주, 범주가 타당하게 구성되었는지 검토하고 과학 특이적 요소를 확정했다.

III. 연구 결과 및 논의

학생들은 일상과 학습을 통해 경험한 과학을 내면화하면서 점차 과학에 대한 저마다 독특한 관점을 형성해 간다. 이러한 학생의 과학 경험과 관점은 과학 관련 정의적 특성을 측정할 때 검사 도구와 상호 작용하며 과학 특이적 반응을 나타낼 수 있다. 학생들의 과학 특이적 반응을 알아보기 위해 과학 관련 정의적 특성과 심리적 특성 등을 검사한 양적 자료와 검사 전후 학생을 면담한 질적 자료를 활용했다. Table 5는 검사 전 실시한 발생 사고 면담 참여자 10명과 검사 후 실시한 면담 참여자 34명 등 총 44명과의 면담에서 선정한 172개의 대화 장면에서 도출한 개념과 그들을 범주화한 수를 정리한 것이다. 도출한 개념은 29개이며 이를 7개의 하위 범주로, 다시 3개의 범주로 묶었다.

Table 5. Number of scenes, concepts, categories derived from students' interview

분석 내용	선정 장면 수			도출 개념 수	하위 범주 수	범주 수
	검사 전	검사 후	계			
과학 특이적 요소	58	114	172	29	7	3

Table 6. Categories and concepts derived from student interviews

범주	하위 범주	개념
과학의 특성	과학을 보는 관점	과정(방법)으로서의 과학, 산물(지식 체계)로서의 과학, 실천으로서의 과학
	과학의 특성과 본성	자연 현상을 다루는 과학, 과학 학문의 위계성, 탐구 방법 중시, 과학 지식 체계의 내적 일관성, 과학과 기술의 연계, 과학의 잠정적 특성, 과학의 사회성 및 연관성, 과학과 창의성, 과학의 논리성, 과학의 객관성, 과학의 도구성, 과학의 비판적·양면적 측면, 과학에 대한 주관적 인식
개인의 과학 경험	과학 동기	진로 선택 및 선발 상황, 과학 관련 직업에 대한 관심
	과학 경험과 검사의 상호작용	과학 관련 개인적 경험 연상, 과학 관련 개인적 경험에 의한 심리적 반응, 개인적 경험과 일반 인식의 차이, 이론 vs 실생활, 특정 과목에 대한 자기 효능감
검사 도구 속 과학	과학과 삶에 대한 인식	일상생활 관련성
	과학 관련 개념의 불명확성	측정 대상 개념 모호(분리된 과학 과목), 측정 대상 개념과 상황 모호
	과학 용어 혼동	측정 대상 개념 인식, 측정 대상 개념 용어 혼동, 측정 대상 인식과 특정 맥락에서 혼동 가능성

Table 6은 학생들의 면담에서 나타난 범주, 하위 범주, 개념을 정리한 것으로 검사에 반영된 학생들의 과학에 대한 생각과 경험을 정리한 것이다. 특히 이 개념들이 검사가 측정하려는 구인이나 목적과 관련 없이 검사에 영향을 미칠 수 있는 상황에 주목했다. 상황과 내용이 유사한 개념을 묶어 최종 3개의 범주-과학의 특성, 개인의 과학 경험, 검사 도구 속 과학 등이 나타났다. 이들 세 범주는 각각 과학 본질적 측면, 학습자 측면, 검사 도구 측면에 상응한다. 개인의 과학 경험은 검사 상황에서 일시적, 지속적 오차를 나타내는 반응 왜곡(response bias)과, 검사 도구 속 과학은 Scheuneman(1984)의 차별적 기능 문항의 원인 중 검사 도구의 결함과 관련 있어 보인다.

1. 학생들의 정확한 자기 보고에 영향을 미치는 과학의 특성

과학 관련 정의적 평가에서 학생들이 지닌 과학 특성은 도구가 측정하려는 구인과 부합하는 경우가 대부분이지만, 그로 인해 정확한 자기 보고가 불확실해지는 사례도 발견되었다. 이는 과학의 본질적 측면과 관련 있다. 학생들이 일상과 학습의 경험을 통해 체화한 과학

Table 7. Students' perception toward science

하위 범주	개념	면담 사례
과학을 보는 관점	과정(방법)으로서의 과학	· 김우진 : (과학은) 탐구하는 거죠 공부하는 그냥 어, 이해도 포함되지만 외우는 게 거의 많다고 생각해요. 탐구는 저희가 직접 궁금하게 느끼고 생각하는 거를, 그것을 직접 답을 찾아가는 과정? 실험 같은 거? 그걸 탐구라고 생각해요. 과학은 (공부보다는) 탐구 쪽에 더 가깝다고.
	산물(지식 체계)로서의 과학	· 김원진 : 어떻게 보면 눈앞에 일어나는 거를 글로 기록하거나 아니면 수식 같은 걸로 일반화하는 게 과학이 아닐까요?
	실천으로서의 과학	· 정진민 : 과학은 과학자들이 모여서 뭔가 생성 원리를 찾거나 아니면 그걸 직접 응용해서 어떤 실용적인 뭔가를 만들거나 그렇게 직접 하는 거로 생각했어요.
	자연 현상을 다루는 과학	· 강도민 : 과학은 그냥 이렇게 주변에 있는 모든 자연 현상이 되었든, 아니면 뭐 존재하는 자체가 다 과학이랑 연관되어 있잖아요. 그냥 그런 것들 자체를 다루는 걸 전 과학이라고 생각을 해요.
	과학 학문의 위계성	· 장영표 : (과학은) 수준이 다르잖아요. 낮은 거에서부터, 높은 건 지금의 II과목일 수도 있는데, 어떤 사람이 높게 생각하는 건 대학교 과정일 수도 있고.
	탐구 방법 중시	· 구진욱 : 과학은 연구를 하는 건데, 연구는 실험 같은 게 중요한 거잖아요.
	과학 지식 체계의 내적 일관성	· 강도민 : 저희가 학교에서 배우는 게 4가지가 있으면 그것들이 서로 다 상호 호환적이고 다 연결이 되잖아요. 이것들이 결국에 모여서 과학이라는 큰 틀을 만들어준 거니까.
과학의 특성과 본성	과학과 기술의 연계	· 김영현 : 과학은 기술적으로 컴퓨터 만들고 이런 것처럼, 기술하는 데에 사용하는 것 같고.
	과학의 잠정적 특성	· 이세진 : 저희가 발견한 것도 있고, 아직 모르는 것도 있고 공존하고 있다 생각해요. 발견된 것도 이제 달라질 수도 있는 거고, 그래서 과학은 계속 개발되어야 해요.
	과학의 사회성 및 연관성	· 김영나 : 과학은 계속 발전하고 사회랑 연관된 거잖아요. · 이세진 : 과학은 점점 쌓아져 온 어려운 그런 거. 그게 또 한 사람이 한 게 아니고 여러 사람이 모여서 그렇다고 확인해줘야 하는 거잖아요.
	과학과 창의성	· 서노을 : 과학은 정해지지 않고 창의성이랑 합쳐져서 자기 마음대로.
	과학의 논리성	· 최현정 : 일상에서 쓰는 과학은 논리적인 거.
	과학의 객관성	· 윤승호 : 과학은 말로 증명할 수 있는 거? 그러니까 국어는 사람의 주관적인 생각이 좀 많이 들어가잖아요. 근데 과학은 누가 봐도 매칭이 되어서 사람들이 아, 이렇게 이해할 수 있는 거라고 생각해요.
	과학의 도구성	· 장산희 : 과학은 제가 좋아하는 분야를 탐구해서 그걸 베이스로 해서 사회생활에 나갔을 때 뭐, 지진이나 그런 거를 연구해서 돈을 벌 수 있는 그런 과학.
	과학의 비판적·양면적 측면	· 최원식 : 과학은 긍정적인 것도 부정적인 것도 다 생각하는 거죠. 과학이 제일 발전할 때가 보면 전쟁 중일 때거든요. 긍정적인 지금 뭐 에어컨이나 휴대폰 이런 것처럼 삶을 좀 더 편안하게 편하게 해주는 거는 긍정적인 거구요.
	과학에 대한 주관적 인식	· 장영표 : 과학은 그냥 다 우연의 일치인 것 같아요. 그러니까 자연의 이치를 알아가는 우연의 결과. · 김명진 : 대체로 생명과학은 뭔가 다른 과목하고 달라서 이해가 아니라 암기라고 생각해가지고.

은 다소 주관적이어서 객관적 관점과 일치하지 않는 경우가 있지만 학생들은 ‘과학은 무엇이며 어떤 특성을 지닌다’는 나름의 인식과 느낌을 지니며, 이에 근거해서 반응한다. 과학의 본질적 측면에서 학생들이 이야기한 하위 범주를 ‘과학을 보는 관점’과 ‘과학의 특성과 본성’으로 묶었다(Table 7).

가. 과학을 보는 관점

Kim et al.(2014)은 과학을 구성하는 주된 요소로 지식과 방법, 그리고 이를 구현하는 실천으로 보았고, 이를 지식 체계로서의 과학, 방법으로서의 과학, 실천으로서의 과학으로 정리했다. 이들이 제시한 과학의 측면은 학생들이 과학을 보는 관점에서도 드러났다. 김원진 학생에게 과학은 자연 현상을 “글로 기록하거나 수식으로 일반화”해서 나타나는 산물이며 지식 체계다. 김우진 학생은 과학을 탐구로 보았고, “직접 답을 찾아가는 과정”으로 인식하며 그 방법으로 실험을 언급했다. 정진민 학생은 과학을 “과학자들이 모여서 직접 하는 거”라고 이야기했다. 이 학생은 과학의 주체가 과학자이며, 그 일을 하기 위해 “과학자가 모여서 실용적인 뭔가를 만든다”고 했다. 이는 이론이나 지식을 생성하는 과학이라기보다 ‘기술’의 영역에 더 가까워 1980 년대에 새롭게 등장한 관점인 ‘실천으로서의 과학’을 떠올리게 한다.

20세기 강력한 과학 이미지를 형성했던 이론 의존성에서 벗어나 실험 실의 고유성을 인정하고 실험 기기를 통해 새로운 효과나 현상을 만들어 냈는데, 이때 과학자 혼자서 아닌 다양한 사람들의 협력을 기초로 한 협동이 이루어진다(Hong, 1994; Pickering, 1992).

학생들 대부분은 과학을 지식 체계가 종합된 학문으로 보는 경향이 강했고, 탐구나 실험, 기술을 떠올리며 과학만의 독특한 특성으로 인식했다. 과학을 학문으로 생각할 때 나타나는 문제는 검사 도구 관련 오차와 연결되어 용어를 혼동해서 반응하게 된다. 학생들은 과학 과목을 ‘지식’으로 이루어진 학문으로 생각하고 과학은 그보다 범위가 넓은 학문으로 생각하므로 둘을 같은 것으로 혼동했다. 서노을 학생은 과학은 지식이라는 산물이므로 학교에서 배우는 지식의 집합체인 과학 과목과 동일하다는 생각을 나타내며 두 개념을 구별하지 않고 있음을 알 수 있다.

서노을: 전 (과학과 과학 과목에 점수를) 똑같이 줬어요. 지금 배우는 과학 과목이 나중에 쓰는 과학이랑 똑같다고 생각해서. 지금 배우는 과학 과목이 이제 곧 과학이라는 말이잖아요. 그럼 이제 학교에서도 과학 과목 배우듯이 이제 그게 지식으로 넘어가는데, 과학에 대한 지식이 되니까 똑같다고 생각했어요.

(1학년 서노을 학생 면담 중)

나. 과학의 특성과 본성

Table 7에서 과학은 어떠한 특성을 지녔으며 그 본성은 어떠한지에 대한 학생 인식이 나타난다. Kim *et al.*(2014)은 과학의 특성을 탐구 방법 중시, 개방적 시험 가능성(testability), 지식 체계의 내적 일관성, 자연 현상을 대상으로 한다는 점을 들었다. 이와 관련되어 학생들이 이야기한 개념은 ‘탐구 방법 중시’, ‘과학 학문의 위계성’, ‘자연 현상을 다루는 과학’, ‘과학 지식 체계의 내적 일관성’이다. 과학이 다루는 대상은 “자연 현상과 존재하는 것”(강도민)이며 과학을 연구하기 위해서는 “실험이 중요한”(구진욱) 역할을 한다고 했다. 학생들은 또 “상호 호환되어 연결되는 것들이 모여 큰 틀”(강도민)이 되는 지식의 일관성과 “낮은 수준에서 높은 수준”(장영표)으로 전개되는 위계성을 인식했다.

이러한 개념들은 학계의 일반적 의견과 대체로 일치하지만, 학생들의 경험 안에서 다양하게 변주되며 다소 주관적 관점이 나타나기도 했다. 과학이 “자연의 이치를 알아가는 우연의 결과”(장영표)라거나, “대체로 다른 과목과 다른” 특성을 지닌 어떤 과목(김명진)이 있다는 생각을 통해 과학의 전형(典型)을 학생의 마음속에서 특정하기 어려운, 다양성의 모습을 추측할 수 있다. 실제로 학생들은 학교에서 배우는 과학 과목들의 지식 체계가 서로 연관됨을 인식하지만, 동시에 차이점도 함께 인식하면서 세부 과목에 대한 각자의 선호가 나타나기도 했다. 상당수의 학생들은 과학 과목을 문항의 맥락과 자신의 경험에 따라 세부 과목으로 대체했고, 그 과목의 선호에 따라 반응했다. Chung & Shin(2017)은 선호하는 과목에 따라 정의적 성취가 달리 나타나는 집단 차이를 보고한 바 있다. 과학 과목을 세부 과목으로 대체해서 반응한 결과가 통계적으로 의미 있는 차이를 만들어 내지 않거나, 학생 스스로 인식하지 못한 채 반응한다 할지라도 검사의 타당성에 대한 원론적 문제가 남는다. 학생들이 과학을 세부 과목으로 생각하고 반응하는 것을 검사 목적과 타당도에 비추어 허용할 수 있는지의 여부다. 과학 지식 체계의 내적 일관성과 탐구 방법, 자연 현상을 대상으로 삼는다는 측면에서 세부 과학 과목들은 과학의 특성을 공유한다. 그러나 이를 인식하는 학생들이 지식의 내적 일관성보다는 세부 과목의 차이점에 더 주목한다면 과목에 따라 다른 반응을 유도할 수 있다는 점도 고려해야 한다.

한편, Schwartz *et al.*(2004)은 과학의 본성에 대해 합의된 7가지 관점을 과학 지식의 잠정성, 주관성, 과학적 방법에 대한 믿음, 이론과 법칙 사이의 관계, 관찰과 추리의 관계, 과학자의 창의력, 사회 문화적 요인 등으로 정리했다(Lederman *et al.*, 2002; NSTA, 2000). 이들이 정리한 결과와 부합해서 학생들에게 나타난 과학의 본성 관련 개념은 과학과 기술의 연계, 과학의 잠정적 특성, 과학과 사회의 연관성, 과학의 사회성 및 합의 등이다(Table 7). 이 중 과학과 기술의 연계와 과학의 사회성 및 합의는 Ratcliffe & Grace(2003)가 제시한 과학의 본성에 대한 견해를 참고했다. 이들이 제시한 14가지 견해 중에는 과학과 기술의 관계, 과학의 진화 대 혁명적 변화, 과학 지식의 개방적 수용 등을 포함하고 있어 학생들의 인식과 대응된다. 과학이 기술에 사용된다고 한 김영현은 과학에서 기술로 향한 일방적 관계 인식에 그치긴 했지만 과학과 기술을 연계했고, 이세진은 과학 지식은 “발견된 것도 달라질 수 있어서 계속 개발되어야” 한다는 잠정성을 표현했다. 특히 학생들은 사회 안에서 영향을 주고받는 과학(김영나)과 과학자

들의 사회 안에서 합의되거나 점진적으로 진화하는 측면의 과학 지식(이세진)을 인식하기도 했다.

학생의 마음 속에 형성된 과학의 본성이 자의적으로 해석되거나, 반대로 과학의 본성에 대치되는 생각이 나타날 때 학생들은 과학에 대한 주관적 관점으로 반응하기 쉽다. 학생들은 과학이 상상력과 창의성을 필요로 하는 영역임을 인지했으나 그들에게 창의성은 고정적 틀에 얽매이지 않는 자유로운 것이다. 따라서 과학은 “정해지지 않고 창의성이랑 합쳐져서 자기 마음대로”(서노을) 해도 되는 것이다. 서노을의 이러한 생각은 과학 과목과 과학의 인식 차이를 알아보는 과정에서 드러났는데, 과학에 대한 왜곡된 관점과 같은 맥락의 반응으로 연결될 수 있다. 즉 자유롭게 자기 마음대로 해도 되는 과학에 긍정적으로 반응하지만, “딱딱 정해져 있는”(박은영) 과학 과목에 부정적 인식을 갖게 되는 불일치 현상의 원인이 된다.

다른 사례로 과학과 사회, 과학과 기술이 서로 영향을 주고받는다 생각하는 사회나 기술에 대한 학생의 주관적 인식과 결합해 복합적으로 작용하기도 한다. 단순히 과학의 일반적 가치를 묻는 질문이지만, ‘사회’와 ‘기술’이라는 용어와 함께 제시되면 최원식과 같이 기준에 지닌 생각에 비추어 반응 점수를 조정하기도 했다.

최원식: 사람들의 생활 여건과 사회, 이 차이라서. 사람한테는 그게 좋다고 생각을 하고 그런데 사람들이 모여 있는 사회에는 부정적인 것이 있을 수 있다고.

(2학년 자연·공학 과정 최원식 학생 면담 중)

Kim & Lee(1996)는 과학이 사회 문제와 관련되면서 이를 잘 인식하는 학생들에게 동시에 나타나는 긍정성과 부정성을 양면 가치 태도로 정의한 바 있다. 학생들이 과학의 본성과 관련하여 과학과 과학 지식의 특성으로 꼽은 것은 객관성, 논리성, 도구성, 비판적 사고 외에도 양면적 측면을 이야기 했는데 이는 과학과 사회 문제를 연계할 때 특히 잘 나타난다. 비판적 사고와 양면적 측면은 하나의 주제에 대해 긍정성과 부정성을 동시에 떠올리면서 반응을 망설이는 경향으로 나타나는 데, 최원식은 사회와 과학을 함께 생각하는 문항에 대해서는 같은 구인을 측정하는 다른 문항과 달리 점수를 낮추는 반응을 했다.

Table 8은 사회와 과학을 포함한 진술이 포함된 ‘과학의 일반적 가치’ 구인에 대한 학생들의 반응 결과로, 검사에 참여한 649명 중 정의적 성취가 높은 320명의 이상 반응 경향을 포착한 것이다. 정의적 성취가 높은 학생들이 ‘과학의 일반적 가치’ 문항에 부정 혹은 중간 응답한 학생 현황을 나타낸 것이다. 과학의 일반적 가치는 다섯 개의 구인 중 가장 높은 평균 점수(4.0)를 나타내어 정의적 성취가 낮은 학생이라도 비교적 높게 점수를 주는 구인이다. 그러므로 정의적 성취가 높은 학생이라면 해당 문항들에 비교적 높은 긍정 응답(4점, 5점)이 예상된다. 그러나 같은 구인인 12번~14번과 달리 15번과 16번은 15.9~20.9%의 학생들이 부정·중간 응답(3점, 2점, 1점)을 했다. 이는 구인이 같은 다른 문항들이 9.1%~11.9%로 나타난 결과를 상회한다. 앞서 최원식이 ‘사회’라는 단어에 부정적으로 반응한다는 진술과 함께 과학의 긍정적, 부정적 측면을 동시에 헤아려 그저 좋게 반응할 수 없게 되더라는 맥락과 일치한다. 균형감 있는 시선으로 과학의 양면을 동시에 평가하다 보면 15번과 16번은 과학의 일반적 가치를 묻는 문항으로 단순하게 생각되지 않았을 것이다. 두 문항의 이상 반응은 Table 9에서 Rasch 모형으로 분석한 결과 모형과 실제

Table 8. Students' negative or intermediate response to the 'general value of science' items with high affective achievement

정의적 성취가 높은 학생들(N=320) 중 부정·중간 응답 학생 수(%)				
12번	13번	14번	15번	16번
일반적으로 과학과 기술의 발전은 사람들의 생활 여건을 개선한다.	과학은 우리가 자연 세계를 이해하는데 중요하다.	과학과 기술의 발전은 경제 발전에 도움이 된다.	과학은 사회에 유익한 것이다.	일반적으로 과학과 기술의 발전은 사회에 이익이 된다.
35(10.9)	29(9.1)	38(11.9)	67(20.9)	51(15.9)

Table 9. Item fit analysis on 'general value of science' by Rasch model

문항	원점수	문항 측정치(곤란도)	오차	내적합도*	외적합도*	점이연 측정 상관계수**
12	4.0	-0.92	0.06	1.23	1.26	0.55
13	4.0	-0.74	0.06	1.17	1.16	0.58
14	4.1	-1.20	0.06	1.26	1.24	0.53
15	3.9	-0.82	0.06	1.46	1.76	0.48
16	4.0	-1.25	0.06	1.47	1.75	0.46

* 내적합도와 외적합도는 1.0을 기준으로 지수의 값이 0.5~1.5이면 적합한 문항으로 간주함(Linacre, 2005)

**피험자의 문항 점수와 피험자 총점의 상관계수로 문항 변별도 지수와 유사하게 해석됨(Seol, Kim, & Lee, 2006)

반응이 잘 맞지 않는 부적합 지수가 나타난 결과와 일치한다. Table 9는 문항에 응답한 학생들의 반응이 문항 반응 이론 모형에 얼마나 들어맞는지를 판정하는 문항 적합도를 정리한 것이다. 15번과 16번은 외적합도가 각각 1.76, 1.75로 학생들의 반응이 모형의 기대에서 벗어나 있음을 알 수 있다. 즉 피험자의 능력이나 특성에 비추어 곤란도가 낮았던 문항임에도 점수를 부정적으로 주었음을 지시하고 있어 해당 문항이 측정하고자 하는 일반적 가치를 제대로 측정하고 있는지 의문을 제기할 수 있다.

한 경향성을 나타내며 정확한 자기 보고에 영향을 미치는 체계적 오차로 작용하기도 한다. 이러한 학생의 특성으로 인해 검사 상황과 동기 측면에서, 개인적 특성 측면에서 반응 왜곡이 나타나는 사례를 개념으로 도출해서 Table 10에 정리했다. 과학 동기는 반응 세트에, 과학 경험과 검사의 상호작용 및 과학과 삶에 대한 인식은 반응 양식에 해당한다.

가. 과학 동기

과학을 장차 직업으로 하려는 동기를 지닌 학생은 그렇지 않은 학생보다 과학 관련 정의적 평가에 좀 더 적극적 반응을 보였다. 이들은 과학 관련 직업에 대한 관심이 상대적으로 높고 진로를 선택할

2. 학생들의 정확한 자기 보고에 영향을 미치는 개인의 과학 경험

학생들이 과학 경험과 학습을 통해 내면화한 특성은 개인마다 일정

Table 10. Some cases of students' responses with 'personal science experience'

하위 범주	개념	면담 사례
과학 동기	진로 선택 및 선발 상황	· 김태민 : 전 무조건 진로 쪽으로 그냥 바로 긍정적으로 다 할 것 같아요. 연구자 : 거짓말이라도 할 수 있다는 얘기니? 김태민 : 거짓말은 아닙니다. 아까 과학 과목하면 생물을 생각했다 했잖아요. 여기서는 그냥 아예 물리, 화학을 좋아하니까 그걸로 생각하고 과학 과목을 체크할 것 같아요. 지금은 급하니까 싫은 거는 빼고 좋은 거 위주로 하면 아예 거짓말은 (아니죠).
	과학 관련 직업에 대한 관심	· 용지호 : 그랬으면 좋겠어서, 나에게 쓸모가 있었으면 좋겠고 내가 좋아했으면 좋겠어서, 더 맞아, 나는 좋아하지. 나 과학 과목 미래에도 당연히 필요하지 하고 조금 더 높은 점수를 주는 경우가 있는 것 같아요.
과학 경험과 검사의 상호작용	과학 관련 개인적 경험 연상	· 김빛나 : 제 진로랑 관계되고 의료 기술 이런 거 떠올리면. 이런 뭐지? 사회 이익이 된다 이런 거 읽으면 생명(과학)이 생각났어요.
	과학 관련 개인적 경험에 의한 심리적 반응	· 황정혜 : 질문에 따라 다른 것 같아요. 사회 발전이 되고 유익한 건 맞는데. 제가 답을 잘 쓰고 개념을 잘 이해하는 거에는 (낮게).
	개인적 경험과 일반 인식의 차이	· 박은영 : 나, '나'라고 딱 명시가 되어 있으니까. '나는~'이라고 하면 약간 더 생각하게 되고, 오래 걸리는 (중략) 나가 아니고 일반적으로 이런 거는 쉽게 했는데. 그냥 바로 제 생각.
	이론 vs 실생활	· 안정훈 : 물리 시간에 배우는 뉴턴, 아인슈타인 이런 걸 배우는데, 그게 실생활에서 그대로 적용이 되는 게 아니어서 조금 이론이라는 그런 이미지가 강하게 나타났다고 할까.
과학과 삶에 대한 인식	특정 과목에 대한 자기 효능감	· 권주영 : 좋아하는 거 위주는 아니고, 4개 중에 제일 그나마 물리를 제일 잘한다고 생각해서 물리로 생각하고 그렇게 (풀었죠).
	일상생활 관련성	· 오진우 : 앞으로 과학과 함께 살아가야 하잖아요. 지금도 그렇고. 사회에 유익한 이런 것들, 전부 다 제가 써야 되니까 계속. 그래서 그냥 앞으로 쉬워질 것 같고 또 그래야 하니까 전부 매우 그렇다 로 한 거죠.

때나 검사와 선발 상황에서 뚜렷한 목표 의식을 나타냈다. 김태민 학생은 진로 선택과 선발 상황에서 평소 자신의 반응보다 더 긍정적인 반응을 한다면 “싫은 거는 빼고 좋은 거 위주로 하면 거짓말은 아니라고” 정색하며 말했다(Table 10). 과학을 분리된 세부 과목으로 생각하면 어떤 경우든 긍정적으로 맞아 떨어지는 과목이 있어 자신의 반응을 합리화하는 근거로 삼았다. 반면 인문·사회 계열인 임호수 학생은 “문과라서 관련이 없어서 솔직하게” 반응한다며, 진로와 선발에 반영된다면 달리 반응할 것이라 말했다. 용지호 학생은 자연·공학 계열로 진로를 선택한 후 과학에 대해 “그랬으면 좋겠다”는 희망을 마음속에 되뇌며, 그 희망에 따라 조금 더 높은 점수로 반응했다. London(1997)과 Noe & Ford(1992)는 진로 동기가 높은 사람은 지속적인 노력을 통해 미래에 대한 긍정적 태도를 유지하려는 경향이 있다고 보고했다. 과학 관련 진로를 희망하는 학생은 특히 진로 선택과 선발 상황에서 제공되는 과학 관련 정의적 영역 평가에서 의식적 혹은 무의식적으로 긍정 왜곡하려는 경향을 나타낼 수 있으며, 미래 자신의 희망에 긍정적 태도를 유지하려는 경향으로 자성적 예언과 함께 긍정 반응을 유도할 수 있다.

Table 11은 과학의 도구적 동기 수준이 다른 학생들의 진로 과정 분포를 나타낸 것이다. 과학 동기 집단의 분류는 도구적 동기 점수를 가장 낮은 점수부터 오름차순으로 42.2%, 15.6%, 42.2%로 끊어 각각 하, 중, 상 집단으로 설정했다. 하-과학 동기 집단과 상-과학 동기 집단에서 인문·사회 과정 학생과 자연·공학 과정 학생들의 분포는 반대 양상이다. 상-과학 동기 집단에서는 자연·공학 학생이 50%로 우세하지만, 하-과학 동기 집단에서는 인문·사회 학생이 43.1%로 가장 많았다. 이공 계열로 진로를 결정한 학생들 중 과학 동기가 높은 학생

들이 인문·사회 계열 학생들보다 많았다.

Table 12는 과학의 도구적 동기 집단의 인지적·정의적 성취 및 사회적 바람직성에 대한 집단 차이를 나타낸 것이다. 과학 동기가 높은 학생들은 다양한 심리적 특성이 개입하면서 긍정 상향 반응한다는 면담 내용과 같이 정의적 특성의 다른 4개 구인에서 유의미하게 높은 점수가 나타났다. 인지적 측면에서도 과학 동기가 높은 집단의 과학 학업 성취도가 유의미하게 높은 결과가 나타나 과학 동기, 과학 관련 정의적 성취, 과학 인지적 성취의 상관을 추측할 수 있다. 그러나 과학 동기 수준 집단의 차이가 반응 왜곡이 반영된 결과인지, 집단의 특성 차이인지, 혹은 양쪽 모두인지 이 자료들만으로 판단할 수는 없다. 한편, 과학 동기 집단의 사회적 바람직성 척도를 나타낸 것으로 상-과학 동기 집단의 사회적 바람직성이 다른 집단에 비해 높게 나타났다. 이는 과학 동기 집단의 반응에 측정 구인과 무관한 심리적 특성이 개입되었을 가능성을 시사하는 사례지만 그 자체로 반응 왜곡의 분명한 증거라 할 수 없다. 심리학 분야에서는 구인 타당도에 영향을 미치는 반응 왜곡을 연구하고자 인위적 실험 조작, 동기 상황이 다른 집단의 비교, 자연스러운 검사 상황에 대해 통계적으로 접근 하고 있지만 연구 방법의 결함이 지적되면서, 추론할 뿐 확실한 증거를 제시하지 못하고 있다고 했다(Son, Cha, & Kim, 2007).

나. 과학 경험과 검사의 상호작용 및 과학과 삶에 대한 인식

학생들의 과학 경험과 정의적 영역 검사가 상호작용하면서 나타내는 특이적 반응을 ‘과학 경험과 검사의 상호작용’으로 묶었다. 학생들

Table 11. Distribution of students in career course by level of instrumental motivation score in science

	과학 동기 집단의 학생 수(%)		
	하-과학 동기	중-과학 동기	상-과학 동기
1학년(n=228)	90(32.8)	35(34.7)	103(37.6)
인문·사회(n=178)	118(43.1)	26(25.7)	34(12.4)
자연·공학(n=243)	66(24.1)	40(39.6)	137(50.0)
전체(N=649)	274(100.0)	101(100.0)	274(100.0)

Table 12. Group differences of affective·cognitive achievement and psychological scale according to level of instrumental motivation score in science

		과학 동기 집단의 정의적, 인지적 성취와 심리적 지수			F
		하-과학 동기	중-과학 동기	상-과학 동기	
정의적 성취	자아 개념	2.7	3.1	3.7	115.403***
	즐거움	2.8	3.2	3.8	148.343***
	일반적 가치	3.7	3.8	4.3	53.723***
	개인적 가치	2.8	3.3	3.9	233.622***
인지적 성취 (학업 성취도)	과학	47.9	47.8	54.1	13.310***
	물리	46.0	48.1	53.1	14.315***
	화학	47.0	48.3	52.6	9.101***
	생명과학	47.1	48.8	52.7	9.051***
	지구과학	48.8	49.2	52.5	6.969**
	평균(T점수)	48.5	48.6	53.2	20.361***
심리적 지수	사회적 바람직성	3.5	3.6	3.7	9.525***

** $p < .01$, *** $p < .001$

이 검사에 응할 때 개인의 과학 경험으로부터 연상되는 것과 이로 인한 심리적 반응이 각기 다르며, 학생 개인과 일반 인식의 차이를 혼동하기도 하고 특정 과목에 대한 효능감이 검사에 영향을 주기도 한다. 또한 과학과 삶의 밀접성을 인식하는 학생들의 이야기를 개념으로 정리했다(Table 10).

학생들은 과학을 대상으로 하는 문항과 만날 때 자신이 예전에 겪은 과학 경험을 회상했다. 특히 문항의 형식이나 맥락에서 자기 관련성이 높은 경우(황정혜, 박은영) 문항에 대한 태도나 집중도가 달라졌다. 박은영 학생은 이를 “약간 더 생각해서 오래 걸리거나”, “바로 제 생각”을 나타낸다고 말했다. 박은영 학생처럼 자신을 드러내야 하는 문항에서 집중도가 높아지거나 시간이 오래 걸리는 현상을 성실한 반응으로 간주해서 정확한 자기 보고를 지향한다고 단정하기는 어렵다. Markus(1977)는 과거 경험에서 형성된 자신에 대한 인지적 일반화를 자기 도식(self-schema)으로 정의했다. 자기 도식이 있는 사람은 자신에 대한 정보를 빠르게 처리하고, 자기 도식과 부합하는 새로운 정보는 빠르게 처리하는 반면, 모순될 경우 저항을 느낀다(Sweeney & Moreland, 1980). 따라서 반응을 왜곡하려는 의도로 문항에 응답하는 사람은 자기 도식과 일치하지 않기 때문에 인지적 노력들 뿐 아니라 감정적 갈등이 발생하고, 이는 반응 시간 지연으로 나타난다(Holden & Hibbs, 1995; Holden *et al.*, 1992; McDaniel & Timm, 1990; Markus, 1977; Munsterberg, 1908). 문항에 거짓으로 응답할 때 시간이 더 오래 걸린다는 연구들과 달리 시간이 더 단축되거나 차이가 없다는 상반된 결과(Kluger, Reilly, & Russell, 1991; George & Skinner, 1990; Hsu, Santelli, & Hsu, 1989)도 보고되지만 대체로 왜곡과 반응 시간이 상호작용함에 동의하고 있으며, 두 변인을 조절하는 변수에 대한 연구로 진행되는 추세다(Bruntti *et al.*, 1998; Holden & Kroner, 1992). Vasilopoulos, Reilly, & Leaman(2000)은 인상 관리를 하는 경우 직무 친숙도가 높을 때 시간이 더 많이 걸리는 반응을 확인했다. 따라서 개인의 과학 경험과 밀착되어 자기 관련성이 높다고 판단되는 문항은 그렇지 않은 문항보다 집중하는 경향이 자주 나타남으로써 사회적 바람직성과 같은 심리 요인이 개입되기 쉽다.

한편 안정훈은 “이론이라는 이미지가 강하게 나타나는”것은 과학 학습으로, 실생활 관련은 과학으로 생각해서 반응한다고 했다. 이 학생은 이전의 과학 경험을 중심으로 형성된 뚜렷한 자기 도식을 지녔으며, 이러한 흑백의 도식들이 다소 과장된 반응으로 나타났다. 분명한 자기 도식을 가진 학생은 자신을 판단의 준거로 삼아 문항에 대해 신속하게 정보를 처리하고 반응한다. 이때 과학에 대해 형성한 경험적 도식이 어떤 정서와 결합해 있느냐에 따라 반응의 방향이 달라질 수 있다. 이밖에도 문항에서 떠올리는 개인 경험에 따라 특정 과목에 한정되거나(김빛나) 자의적 해석을 하는 경우, 효능감이 높은 과목으로 대체해서 반응하는 경우(권주영) 등이 나타났다.

한편 학생들이 지닌 과학과 삶에 대한 인식은 과학에 대한 전체 느낌을 좌우하지만, 그 전반적 느낌 때문에 정확한 자기 보고를 방해하는 경우도 나타났다. 학생들은 자신의 경험과 삶 속에서 과학이 주는 유익함과 편리함을 충분히 인식하고 과학을 일상과 분리할 수 없는 삶의 일부로 여겼다. 오진우는 현재 과학의 위상이 미래에도 계속될 것으로 전망하고 긍정했다. 이 학생은 검사에서 모든 문항에 대해 ‘매우 그렇다’로 반응했고, 연구자는 불성실 반응 여부를 질문했다. 그는 세부 내용을 자세히 판단하기보다 “앞으로 쉬워질 것 같고

또 그래야 하나까”라며 스스로 설정한 당위성에 맞추어 모든 문항에 동의했다. 이때 작용한 것은 과학과 삶을 뗄 수 없는 것으로 인식함으로써 긍정적으로 느끼고 생각하려는 의지였다.

3. 학생들의 정확한 자기 보고에 영향을 미치는 검사 도구 속 과학

검사 도구 측면에서 과학 특이적 반응은 도구 속에 존재하는 과학 관련 개념의 불명확성과 그로 인한 용어 혼동을 다룬다. 과학 관련 정의적 영역 평가 도구는 개념과 개념을 담는 언어로 이루어지며, 문장의 형태로 제시되었다. 평가 도구는 검사자의 의도를 피험자에게 정확하고 분명하게 제시해서 검사 목적에 맞는 반응을 이끌어내야 한다. 도구를 구성하는 개념과 언어, 문항의 맥락이 명확하지 못하거나, 차별적 반응을 유도하거나, 오해해서 반응할 여지가 있다면 정확한 자기 보고를 담보할 수 없으며, 그 결과를 의미 있게 해석할 수 없다. 특히 측정하려는 구인과 관련 있는 중심 개념이 모호하고 문항 맥락이 다의적으로 해석된다면 피험자 반응의 의미를 특정할 수 없게 된다. 본 연구에서 사용한 PISA 2006의 과학 관련 정의적 특성 문항에서 측정하는 중심 개념은 과학 혹은 과학 과목에 대한 자아 개념, 즐거움, 일반적 가치, 개인적 가치, 도구적 동기다. 학생들이 이야기한 과학은 일상의 과학, 학교 과학(과학 과목)이 혼동되기도 하고, 여러 과목으로 분리된 과학 과목 중 하나로 대체되기도 한다. 또 개인이 겪은 과학과 과학 학습 경험은 문항의 맥락에서 다의적으로 해석되기도 한다. 따라서 학생이 도구와 상호작용하는 과정에는 학생이 지니고 체화한 과학 특이성이 반영되기도 하며, 이러한 특성이 자기보고식 과학 관련 정의적 평가의 정확성에 영향을 주기도 한다.

Table 13은 학생들이 지닌 과학 특이적 생각 때문에 도구와 반응하면서 나타나는 대표적 사례와 이를 범주화한 개념을 정리한 것으로 이들 5개의 개념은 ‘과학 관련 개념의 불명확성’과 ‘과학 용어 혼동’의 하위 범주로 묶였다. 과학 관련 개념의 불명확성은 측정 대상 개념인 과학, 과학 과목이 학생에게 다양하게 인식되거나 문항의 상황이 다르게 인식되는 경우다. 그 결과로 과학 용어의 혼동이 나타나는데, 개념을 구별해서 잘 인식하는 경우, 혼동하는 경우, 특정 맥락에서 혼동하는 경우들이 각 개념에 해당한다.

학생들이 잘 구별하지 못하거나 분명치 않게 인식하고 반응하는 주요한 측정 개념은 과학과 과학 과목이다. 임호수는 4개의 과목으로 분리된 과목 중 문항의 상황과 자신의 선호에 따라 다른 과목을 대입했고, 이전하는 문항의 맥락에서 연상되는 과목을 선별적으로 대체해서 반응했다. 그 결과는 문항마다 다른 대상 개념에 대한 반응으로 나타날 수 있는데, 이는 과학 관련 정의적 특성을 측정하려는 단순한 목적을 지닌 문항 개발자의 예상과 다른 결과로 나타날 수 있다. Table 14는 검사에서 ‘과학’이나 ‘과학 과목’의 용어가 나올 때 무엇을 떠올리며 응답하는지 질문한 결과다. 학생들은 과학의 지식 체계가 서로 연관됨을 인식했지만, 동시에 차이점도 인식했다. 상당수 학생들(55.5%)이 과학을 세부 과목으로 대체하거나 문항마다 다른 과목을 연상한다고 했다. 문항 개발자는 학생들이 과학의 내적 지식 체계의 일관성을 잘 인식해서 종합적으로 응답할 것이라 예상하지만, 학생들은 자신의 선호, 정서, 효능감, 최근 학습 과목 등을 떠올려 반응하기도 했다.

Table 13. Cases of students' responses with 'science in test tool'

하위 범주	개념	면담 사례
과학 관련 개념의 불명확성	측정 대상 개념 모호 (분리된 과학 과목)	· 임호수 : 저는 대부분은 지구과학이었는데, 풀다가 막 즐긴다거나 약간 그런 느낌이 나오면 생물이 같이 떠올라서, 일단 지구과학이 한 60~70% 정도 되지 않았을까요?
	측정 대상 개념과 상황 모호	· 이진하 : 저는 우주까지는 생각 안 했고요. 보통 자연이라고 하면 생각나는 그런, 약간 숲 같은 그런 이미지들 있잖아요. 과목으로 생각하자면 생물.
과학 용어 혼동	측정 대상 개념 인식	· 민준호 : 제가 생각하는 게 다르니까. 과학 이런 거는 연구자들이 모여서 하는 거라고 생각하고 과학 과목은 학교에서 배우는 걸 생각하니까 따로따로 나오니까 아, 이렇구나 하면서 답을 다르게 해요.
	측정 대상 개념 용어 혼동	· 윤승호 : 생명이나 물리 이런 게 과학 과목이고 그게 합쳐지는 게 과학인 줄 알았어요. 과학 과목은 과학에서 체계적으로 이렇게 파트 별로 나뉜진 게 과학 과목이라고 생각하고 그런 게 다 합쳐져서 통틀어서 말하는 게 과학인 것 같아요.
		· 최현정 : 과학 문제가 수능 칠 때 하는 입시 문제예요? 아니면, 하늘은 왜 파란지 나뭇잎은 왜 초록색인지 같이 일상생활에서 찾아보는 과학 의문점을 직접 해결하는 거예요?
측정 대상 인식과 특정 맥락에서 혼동 가능성	· 이진하 : 과학이라고 하면 엄청 다른 세계 같은 데서 연구원들이 힘들게 연구하고 그런 어려운 거고 과학 과목은 그걸 갖다가 학생들이 이해하기 쉽게, 좀 쉽게 풀이해 놓은 듯한. 그런.	

Table 14. Distribution of students related to specific science area or subjects in science-related affective test

구분	검사에서 연상한 과학 영역 응답 학생 수(%) (N=649)						계
	물리	화학	생명과학	지구과학	과학	문항마다 다름	
과학	37(5.7)	29(4.5)	70(10.8)	82(12.6)	289(44.5)	142(21.9)	649(100.0)
과학 과목	37(5.7)	24(3.7)	84(12.9)	87(13.4)	273(42.1)	144(22.2)	649(100.0)

세부 과목으로 대체했을 때 반응의 결과는 Table 15와 같이 연상 과목에 따른 집단 차이로 나타났다. 선호 여부에 따라 세부 과목의 정의적 성취가 달라진다는 연구 결과(Chung & Shin, 2017)도 보고되었고, 학생이 과목을 연상하게 되는 원리에 따라 검사 결과는 영향을 받을 수 있다. Table 15에서 집단의 차이가 나타나는 이유는 주로 물리를 연상하는 집단의 정의적 성취가 다른 집단보다 높게 나타나고, 지구과학을 연상하는 집단의 정의적 성취가 낮기 때문이다. 이는 지구과학 연상 집단을 이루는 구성원 중 과학에 대한 동기 수준이 낮은 인문·사회 과정 학생이 많으며, 설문을 실시한 해당 학기에 이수를

완료한 과목으로 최근까지 학습한 기억이 많이 남아있던 경우를 반영한 결과로 추측된다(Table 2).

Table 14에서 주목해야 할 집단은 '문항마다 다름'에 응답한 집단이다. 21.9%와 22.2%로 대략 20% 내외의 학생들이 문항마다 다르게 세부 과목을 대체해서 반응한다. 문항마다 다른 과목의 연상은 측정 개념이 문항마다 달라진다는 점에서 일관성 없는 반응으로 나타날 수 있다. Table 16은 과학 영역 연상 집단을 세부 영역, 과학, 문항마다 다름의 세 집단으로 나누어 5개 구인의 내적 일관성 신뢰도(Cronbach α)를 나타냈다. Cronbach α 는 문항들 간 상관의 평균을 기초로 산출

Table 15. Group differences by specific science area related to groups in affective achievement

구분	검사에서 연상한 과학 영역 집단의 정의적 영역 성취						F
	물리 (n=37)	화학 (n=29)	생명과학 (n=70)	지구과학 (n=82)	과학 (n=289)	문항마다 다름 (n=142)	
자아 개념	3.6	3.3	3.1	3.0	3.2	3.1	2.938*
즐거움	3.7	3.3	3.2	3.1	3.3	3.3	2.542*
일반적 가치	4.0	4.0	3.9	3.8	4.0	4.0	0.627
개인적 가치	3.7	3.4	3.3	3.2	3.4	3.3	2.592*
도구적 동기	3.8	3.5	3.3	3.0	3.4	3.2	5.078***

* $p < .05$, *** $p < .001$

Table 16. Inter-item consistency by subconstructs in specific science area related to groups

연상 집단	검사 문항의 구인 별 내적 일관성 신뢰도(Cronbach α)					
	자아 개념 (6문항)	즐거움 (5문항)	일반적 가치 (5문항)	개인적 가치 (5문항)	도구적 동기 (5문항)	전체 (26문항)
세부 영역(n=218)	.939	.933	.891	.873	.946	.966
과학(n=289)	.939	.920	.839	.859	.949	.953
문항마다 다름(n=142)	.881	.889	.838	.816	.916	.929
전체(N=649)	.930	.919	.861	.856	.942	.955

하므로 ‘문항마다 다름’ 집단의 α 계수가 다른 집단보다 작은 것은 이 집단의 문항 간 상관성이 다른 집단에 비해 작음을 의미한다. 문항마다 기준으로 생각하는 영역이나 과목이 다르다면 반응의 양상도 달라질 수 있다. 이 집단의 α 계수가 상대적으로 작아서 문항마다 다르게 반응하는 경향성을 추측할 수 있지만, 상대적으로 작다는 것 외에 판별 기준이 없어서 의미 있는 차이인지 여부는 분명하지 않다. 그러나 문항마다 다른 과목을 대체하는 학생들의 같은 구인에 대한 상관성이 다른 집단에 비해 일관적으로 낮다는 점은 주목할 필요가 있다.

이상과 같이 과학을 하위 과목으로 대체한 결과를 과학에 대한 종합적 인식이나 태도로 간주해도 좋은지에 대한 합의된 의견을 찾아보기 어렵다. 또는 이러한 현상을 인지하지 못했거나 이에 수반되는 영향을 무시할 수 있다고 간과하고 있을 것이다. 특히 과학 관련 정의적 특성을 측정함에 있어 과학의 특성을 잘 이해하고 내면화한 학생들이 오히려 검사 목적에 반하는 반응을 하게 된다면 검사 도구의 타당성을 저해할 수 있다.

한편, 측정하려는 구인의 중심 개념인 과학과 과학 과목에 대한 인식은 민준호나 이진하처럼 전문가와 학생의 영역에서 연구하거나 학습하는 것으로 구별해서 인식할 수도 있고 윤승호처럼 동음이의어 관계인 과목으로서 과학과 혼동할 수도 있다(Table 13). 과학과 과학 과목을 구별하더라도 두 용어의 정체를 ‘학문’으로 인식하는 한, 특정 맥락에서 두 용어를 같은 것으로 생각하고 반응할 수 있다. 최현정이 ‘과학 문제’를 ‘과학 입시 문제’와 ‘과학에 대한 의문점’인지 궁금해 하듯이 일상에서 흔히 사용되는 언어가 중의적으로 해석될 여지가 있을 때 자의적으로 해석하고 반응할 여지가 있다. 학생들의 경험 속에 형성된 과학 개념은 도구 속 과학과 상호작용하며 검사자가 의도하지 않은 반응으로 나타날 수 있다.

IV. 결론 및 시사점

본 연구는 자기보고식 검사를 통해 과학 관련 정의적 영역을 평가할 때 검사 결과에서 나타나는 주관성의 양상을 과학 특이적 측면에서 드러내었다. 과학 관련 개념이나 인식을 측정할 때 학생이 지닌 과학 특성, 본성에 대한 인식이 원인이 되어 나타나는 반응을 과학 특이적 반응으로 정의했다. 그 중에서 과학 특이적 반응이 특별히 측정 구인을 방해하거나 정확한 자기 보고를 벗어나게 하는 경우에 대해 탐색했다. 학생을 면담한 질적 자료와 정의적 특성 및 심리적 특성을 검사한 양적 자료로부터 과학 특이적 요소로 인한 오차 결과를 도출해, Figure 1의 도식으로 나타냈다. Figure 1은 자기보고식 과학 관련 정의적 영역 평가에서 과학 관련 개념을 측정하면서 정확성에 영향을 주는 경우를 과학 특이적 측면에서 정리한 것이다. 정확한 자기 보고 영역을 중심으로 설정한 과녁 모양의 도식으로 나타냈다. 44명의 학생 면담을 통해 학생들의 과학 인식과 그로 인한 반응이 드러나는 172개의 장면을 선정했고, 그 중 29개의 개념, 7개의 하위 범주, 3개의 범주(Table 6)로 학생이 지닌 과학 특이성에 의한 반응 사례를 도출했다.

과학 특이적 반응 사례는 측정하려는 구인과 부합하는 경우도 있었지만, 측정하려는 바와 상관없는 반응을 유도해서 정확한 자기 보고를 방해하는 경우도 나타났다. Figure 1에서 나타난 세 범주는 과학의 특성, 개인의 과학 경험, 검사 도구 속 과학으로 각각 과학의 본질적

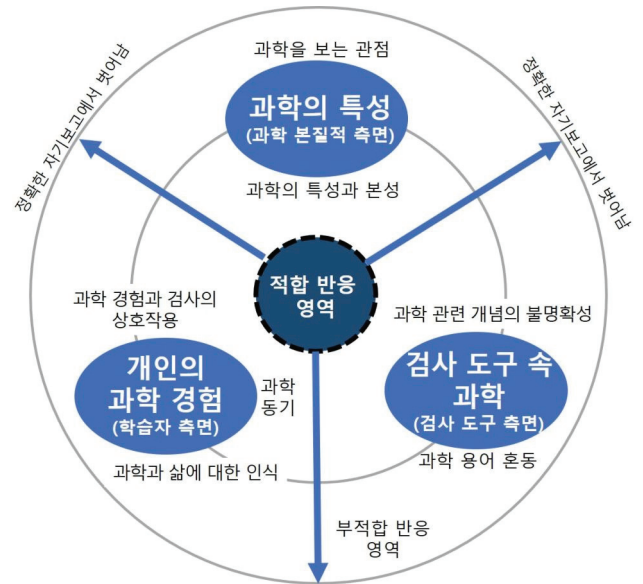


Figure 1. Science-specific factors that interfere with accurate self-report

측면, 학습자 측면, 검사 도구 측면에서 포함되어 있는 과학 특이성을 정리한 것이다. 학생들이 일상과 학습을 통해 내면화한 과학을 보는 관점과 과학의 본성 및 특징 인식은 ‘과학의 특성’으로, 개인의 과학 경험과 검사가 상호작용하며 나타나는 현상, 경험을 통해 지니는 과학과 삶에 대한 인식, 과학과 관련한 동기 등은 ‘개인의 과학 경험’으로 범주화했다. 마지막 범주는 도구 속에 존재하는 과학 관련 개념의 불명확성과 그로 인한 용어 혼동을 다루며 ‘검사 도구 속 과학’으로 명명했다.

중심의 적합 영역은 학생의 과학 관련 특성에 맞게 정확히 반응한 영역이지만, 이는 다소 가설적 영역이며 경계는 모호하다. 학생들이 반응한 점수는 문항에 대한 자기 인식이 나타난 결과이며, 엄밀하게 오차가 반영된 관찰 점수라는 점에서 ‘정확한 자기보고’는 다소 제한적 관념일 수 있다. 학생의 반응이 자신이 지닌 과학 특성을 잘 반영해서 문항의 목적에 부합할 경우 부적합 영역이 축소되면서 학생의 반응 영역은 적합 영역으로만 나타날 것이다. 그러나 과학의 특성을 포함한 문항이 검사가 측정하려는 목적을 방해하는 경우가 나타났고, 학생이 지닌 과학 특성 때문에 구인과 관련 없는 반응으로 연결되는 경우가 있었다. 학생은 정확하게 자기를 보고했다고 인식하지만, 검사자의 입장에서는 정확한 자기 보고에서 벗어났다고 판단하게 된다. 원인은 검사 도구 속 문항의 개념이나 상황이 과학의 특성을 나타내면서 측정 구인에 집중하지 못하게 방해하거나, 학생이 내면화한 과학 특성과 반응하면서 검사자가 예상하지 못한 결과를 산출했기 때문이다. 이 경우 ‘검사 도구 속 과학’과 ‘개인의 과학 경험’은 정확한 자기보고에 영향을 미치는 요소로 작용한다.

이상의 연구 결과와 결론을 통해 과학과 정의적 영역 평가에서 정확한 자기 보고의 주관성을 인식함으로써 과학 교육 현장에서 간과되었던 측정 오차들을 구체적으로 드러냈다. 이렇게 가시화된 오차들은 정확한 검사 도구를 개선하는 실용적 측면의 정보를 제공할 뿐 아니라 학생 개인에 대한 풍부한 이해를 가능케 한다. 또한 자기 보고를 방해하는 오차 요소들은 학생들이 경험한 일상과 학습을 통해 형성되었다는 점에서 과학 교과 교육 측면의 시사점을 얻을 수 있다.

이상 연구로부터 얻은 시사점은 다음과 같다.

첫째, 과학이 중심 개념인 정의적 특성을 측정하는 검사에서는 학생 개인이 지닌 과학 특이성 때문에 측정하려는 구인을 타당하게 측정하지 못하는 역설적 상황이 발생할 수 있다. 과학의 속성과 특성을 연상하게 하는 문맥에 집중하다가 측정 대상과 관련 없는 응답을 하는 사례, 과학과 삶을 밀접하게 생각해야 한다는 생각 때문에 세부 내용을 따지지 않는 사례, 개인이 지닌 과학 경험에 따라 자의적으로 해석하는 사례, 과학 동기가 뚜렷한 학생이 좋은 인상을 주려는 심리 특성으로 인한 왜곡 사례 등이 발견되었다. 따라서 검사 문항이 학생들이 지닌 과학 특이성을 발현시킬 수 있는 상황을 제공하는지 면밀히 따져볼 필요가 있다.

둘째, 검사에서 측정하려는 개념을 잘 정의하고 명확히 하려는 노력과 주의가 필요하다. ‘측정 개념을 명확히 한다’는 의미는 개념을 정의하는 첫 단계를 비롯해서 학생에게 인식되는 수준, 검사 실시 상황까지 포함한다. 검사자는 과학을 학습 상황의 과학과 구분할 수 있다면 과학과 과학 과목이라는 개념을 비교적 잘 구분할 것이라 예상했지만, 이들이 학생에게 인식되는 수준에서는 경우의 수가 여럿 발생했다. 처음부터 둘을 구분하지 않는 경우, 분리된 세부 과학이나 과목으로 대체하는 경우 등이 그 예다. 개념 구별과 주의 집중을 유의시키기 위해서는 검사 실시 전 안내문과 지시문을 강조해서 활용하는 방법, 안내자의 명시적 지시에 따르도록 하는 방법 등이 있다. 그러나 학생 수준에서 발생하는 측정 개념 관련 오차, 예컨대 자의적 인식으로 반응하는 사례를 적합한 영역 반응으로 수용할 것인지 여부는 검사자나 개발자의 합의가 필요한 부분이다. 검사의 목적과 타당성에 비추어 수용할 수 없을 경우에는 용어의 수정이나 명시적 용어로의 대체가 필요하다.

셋째, 과학을 대상으로 한 정의적 특성 평가에는 과학 특이적 반응을 유도하는 요소가 있다. 측정하려는 구인과 과학 특이적 반응이 상충하는 사례가 발견된다면 이를 수정하거나 제거해야 한다. 사전에 이러한 요소를 판단하려면 일차로 과학 교육 전문가의 검토를 요하며, 문항 개발 이후라도 같은 구인 내에서 상관이 낮거나 다른 문항과 차별적인 이상 반응 등이 발견된다면 ‘과학 특이성’ 요소로 인한 오차 반응인지 점검해 본다.

마지막으로 과학 교과 교육적 측면에서, 과학과 과학 과목에 대한 학생들의 인식 차이와 이로 인한 개념과 용어 혼동 사례의 원인을 고찰할 필요성이 있다. 이러한 현상의 원인에는 학생들이 일상과 학교에서 경험하는 과학과 과학 학습의 격차, 과학과 과학 학습에 대한 왜곡된 관점, 그로 인한 학습 진입의 어려움과 부정적 감정, 부정적 감정으로 인한 학습 진입 좌절 등의 많은 현상이 관련되어 있다. 특히 과거 일상과 학교 문화에 괴리를 느꼈던 학생들의 경계를 넘는 어려움은 현재에는 다소 다른 양상으로 재현된다. 과학 교육계의 학생 친화 노력과 과학 대중화 운동으로 일상에 깊이 파고든 과학은 어렵고 머리 아픈 학교 과학(과학 과목)과 대비된다. 학생들은 일상에서의 과학은 쉽고 재미있는 것, 학교 과학은 어려운 것으로 인식하면서 이분화된 과학과 과학 과목에 대한 인식을 지니게 되었다. 일부 학생들은 신기한 현상은 과학이고, 이에 대한 원리는 과학 과목이라는 왜곡된 관점으로 반응하기도 했다. 일상과 과학을 친숙하게 연계하려는 교육적 접근이 오히려 학교 과학을 낯설게 느끼는 원인으로 작용한 경우다. 이제는 일상 과학과 학교 과학 사이의 경계를 넘으려는

노력이 요구된다.

학생 반응이 의도한 특성만을 드러낼 것이라 가정하는 도구 개발자의 의도와 달리 학생 수준에서 도구 속 문항은 스스로 형성한 자기 인식, 심상, 도식에 비추어 평가된다. 학생의 반응은 측정 대상 구인에 대한 인식 뿐 아니라 그들만의 고유한 사고, 정서, 행동을 반영한다. 평가의 목적이 궁극적으로 학생 개인의 발전과 정치(定置), 과학 교육과 교육 정책의 개선에 있다면, 학생 개인에 대한 온전하고 풍부하고 정확한 이해가 선행된 후 반응 결과를 해석해야 할 것이다.

국문요약

이 연구는 자기보고식 검사를 통해 과학 관련 정의적 영역을 평가 하려할 때 검사 결과에서 나타나는 주관성의 양상을 과학 특이적 측면에서 밝혔다. 과학 관련 개념이나 인식을 측정하려할 때 학생이 지닌 과학 특성, 본성에 대한 인식이 원인이 되어 나타나는 반응을 과학 특이적 반응으로 정의했다. 그 중에서 과학 특이적 반응이 특별히 측정 구인을 방해하거나 정확한 자기 보고를 벗어나게 하는 경우에 대해 탐색했다. 고등학교 1, 2학년 649명의 정의적 특성 및 심리적 특성을 검사한 양적 자료와 학생 44명을 면담한 질적 자료로부터 과학 특이적 요소로 인한 오차 결과를 도출했다. 학생이 일상과 과학 학습 경험으로부터 내면화한 과학에 대한 관점과 과학 특성은 검사 도구를 이루는 문항들과 상호작용한다. 그 결과 과학의 특성, 개인의 과학 경험, 검사 도구 속 과학이라는 세 측면에서 정확한 자기 보고를 방해하는 요소가 발견되었다. 과학 본질적 측면과 관련 있는 과학의 특성은 학생들이 과학을 보는 관점과 주관적으로 인식한 과학의 특성이 측정하려는 구인에 관계없이 문항에 반응하도록 한다. 학습자 측면에서 개인의 과학 경험은 학생이 지닌 과학 동기, 과학 경험과의 상호작용, 과학과 삶에 대한 인식으로 구성된다. 마지막으로 도구적 측면에서 검사 도구 속 과학은 과학 개념의 불명확성으로 인한 용어 혼동으로 연결되며 정확한 자기보고를 방해할 수 있다. 본 연구 결과에 의한 시사점으로 검사 문항에서 과학 특이적 요소의 포함 여부 검토, 측정 개념을 명확히 하기 위한 주의점, 개발 단계에서의 과학 특이성 요소 검토, 일상 과학과 학교 과학의 괴리를 줄이려는 노력 필요 등을 제안했다.

주제어 : 과학 관련 정의적 영역 평가, 자기보고식 검사 도구, 검사의 오차 반응, 반응 왜곡, 과학 특이적 요소, 검사 도구 개선

References

- Abd-El-Khalick, F., Summers, R., Said, Z., Wang, S., & Culbertson, M. (2015). Development and large-scale validation of an instrument to assess Arabic-speaking students' attitudes toward science. *International Journal of Science Education*, 37(16), 2637-2663.
- Anastasi, A. (1988). *Psychological testing*(6th ed.). New York: Macmillan & Co.
- Au, Y. (2007). A search on social desirability according to administered mode and demonstrable condition of a psychology testing. *Journal of Educational Evaluation*, 20(4), 235-258.
- Bae, B., Lee, D., & Ham, K. (2015). Validation of the Korean short-version of social desirability scale(SDS-9) using the Rasch model. *Korean Journal of Counseling*, 16(6), 177-197.
- Brunetti, D., Schlotmann, R., Scott, A., & Hollrah, J. (1998). Instructed faking and MMPI-2 response latencies: The potential for assessing response validity. *Journal of Clinical Psychology*, 54, 143-153.

- Choi, J., Hwang, S., Pai, D., Hwang, S. T., & Kim, Y. (2015). Diagnostic efficiency of personality disorder screening tool; The Korean version of self-report standardized assessment of personality-abbreviated scale: preliminary validation study. *Journal of the Korean Neuropsychiatric Association*, 54(4), 534-541.
- Chun, E., Na, J., Joung, Y., & Song, J. (2015). Development and application of the measuring instrument for the analysis of science classroom culture from the perspective of 'community of practice'. *Journal of the Korean Association for Science Education*, 35(1), 131-142.
- Chung, S., & Shin, D. (2016). Trends of assessment research in science education. *Journal of the Korean Association for Science Education*, 36(4), 563-579.
- Chung, S., & Shin, D. (2017). Cases of discrepancy in high school students' achievement in science education assessment: Focusing on testing tool in affective area. *Journal of the Korean Association for Science Education*, 37(5), 891-909.
- Cronbach, L. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6(3), 475-494.
- Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Ferrando, P., & Chico, E. (2001). Detecting dissimulation in personality test scores: A comparison between person-fit indices and detection scales. *Educational and Psychological Measurement*, 61, 997-1012.
- Fives, H., Huebner, W., Bimbaum, A., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education*, 98(4), 549-580.
- George, M., & Skinner, H. (1990). Using response latency to detecting accurate response in a computerized lifestyle assessment. *Computers in Human Behavior*, 6, 167-175.
- Holden, R., & Hibbs, N. (1995). International validity of response latencies for detecting fakers on a personality test. *Journal of Research in Personality*, 29, 362-372.
- Holden, R., & Kroner, D. (1992). Relative efficacy of differential response latencies for detecting faking on a self-report measure of psychopathology. *Psychological Assessment*, 4, 170-173.
- Holden, R., Kroner, D., Fekken, G., & Popham, S. (1992). A model of personality test item response dissimulation. *Journal of Personality and Social Psychology*, 63, 272-279.
- Hong, S. (1994). Interaction between science and technology: Technology as knowledge and science as practice. *The Quarterly Changbi*, 22(4), 329-350.
- Hough, L., Eaton, N., Dunnette, M., Kamp, J., & McCloy, R. (1990). Criterion-related validity of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581-595.
- Hsu, L., Sanetelli, J., & Hsu, J. (1989). Faking detection validity and incremental validity of response latencies to MMPI subtle and obvious items. *Journal of Personality Assessment*, 53, 278-295.
- Hui, C. H., & Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly*, 49, 253-260.
- Joo, Y., Kim, Y., Jeong, S., Shin, M., & Lee, C. (2001). Relationships between subjective symptoms and objective psychopathology in patients with schizophrenia. *Journal of the Korean Neuropsychiatric Association*, 40(4), 667-678.
- Kim, S., & Kim, H. (2016). Development of a science ethicality test for elementary school students. *Journal of the Korean Association for Science Education*, 36(1), 1-13.
- Kim, H., & Lee, S. (1996). Secondary students' attitudes toward science-technology related issues in Korea. *Journal of the Korean Association for Science Education*, 16(4), 461-469.
- Kim, M., & Lee, H. (2006). A study of faking on normative and ipsative measures of personality for personnel selection. *The Korean Journal of Industrial and Organizational Psychology*, 19(3), 371-393.
- Kim, J., Jeong, H., Kim, Y., & Cho, Y. (2015). A study on application of NCS recruiting systems in public organization - based on NCS recruiting performance during the first half year in 2015. *Journal of Skills and Qualifications*, 4(1), 65-84.
- Kim, Y., Park, Y., Park, H., Shin, D., Jung, J., & Song, S. (2014). World of science education. Seoul: Book's hill.
- Kluger, A. Reilly, R., & Russell, C. (1991). Faking biodata tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology*, 76, 889-896.
- Lederman, N., Abd-El-Khalick, F., Bell, R., & Schwartz, R. (2002). Views of NOS questionnaire toward valid and meaningful assessment of learners' conceptions of NOS. *Journal of Research in Science Teaching*, 39, 497-521.
- Lee, J., & Moon, S. (2015). An analysis on the university entrance examination system's change process and main contents of occasional-regular admissions. *Contemporary Educational Research*, 27, 97-130.
- Lee, M., Sohn, W., & No, U. (2007). The Results from PISA 2006. Seoul: Korea Institute for Curriculum and Evaluation.
- Linacre, J. (2005). A user's guide to Winsteps Rasch-model computer programs. Retrieved from www.winsteps.com.
- London, M. (1997). London's career motivation theory: An update on measurement and research. *Journal of Career Assessment*, 5(1), 61-80.
- Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, 35, 63-78.
- McDaniel, M., & Timm, H. (1990). Lying takes time: Predicting deception in biodata using response latency. Paper presented at the American Psychological Association Annual Convention, Boston.
- Messick, S. (1991). Psychology and methodology of response styles. In R. E. Snow, & D. E. Willey (Eds.), *Improving inquires in social science* (pp. 161-200). Hillsdale, NJ: Erlbaum.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5-8.
- Munsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York: Doubleday, Page & Co.
- National Science Teachers Association(NSTA) (2000). NSTA position statement of the nature of science. Retrieved July 12 2003, from <http://www.nsta.org/159&psid=22>
- Noe, R., & Ford, J. (1992). Emerging issues and new directions for training research. In G. Ferris, & K. Rowland (Eds.), *Research in personnel and human resources management* (pp. 345-384). Greenwich, CT: JAI Press.
- Osborne, J., Simon, S., & Tytler, R. (2009). Attitudes towards science: An update. Paper presented at the annual meeting of the American Educational Research Association, San Diego, California.
- Paulhus, D. (1984). Two component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 589-609.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braum, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Pickering, A. (1992). *Science as practice and culture*. Chicago: The University of Chicago.
- Ratcliffe, M., & Grace, M. (2003). *Science education for citizenship: teaching socio-scientific issues*. Maidenhead, UK: McGraw-Hill International.
- Reise, S. P., & Flannery, W. P. (1996). Assessing person-fit on measures of typical performance. *Applied Measurement in Education*, 9(1), 9-26.
- Rust, J., & Golombok, S. (2014). *Modern psychometrics: The science of psychological assessment*(3rd). London: Routledge.
- Scheuneman, J. D. (1984). A theoretical framework for the exploration of causes and effects of bias in testing. *Educational Psychologist*, 19(4), 219-225.
- Schluf, Boaz, Hattie, J., & Dixon, R. (2008). Factors affecting responses to Likert type questionnaires: Introduction of the ImpExp, a new comprehensive model. *Social Psychology of Education*, 11(1), 59-78.
- Schwartz, R., Lederman, N., & Crawford, B. (2004). Developing views of nature of science in authentic context: An explicit approach of bridging the gap between nature of science and scientific inquiry. *Science Education*, 88, 610-645.
- Seol, H., Kim, D., & Lee, S. (2006). Validation of the emotional empathy scale using Rasch rating scale model. *Journal of Education Evaluation*, 19(2), 179-201.
- Shin, S. Ha, M., & Lee, J. (2014). Difference analysis between groups and the generalizability of the instrument for measuring high school students attitude toward convergence. *Journal of Learner-Centered Curriculum and Instruction*, 14(5), 107-124.
- Shin, S., Ha, M., & Lee, J. (2016). The development and validation of instrument for measuring high school students' STEM career motivation. *Journal of the Korean Association for Science Education*, 36(1), 75-86.
- Shin, Y., Kwak, Y., Kim, H., Lee, S., Lee, S. H., & Kang, H. (2017). Study on the development of test for indicators of positive experiences about science. *Journal of the Korean Association for Science Education*, 37(2), 335-346.
- Son, E., Cha, J., & Kim, A. (2007). Test of construct equivalence of personality inventory in low and high socially desirable responding groups. *The Korean Journal of Social and Personality Psychology*, 21(2), 71-87.
- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, 17(3), 222-232.

Stricker, L. J. (1963). Acquiescence and social desirability response styles: Item characteristics, and conformity. *Psychological Reports*, 12, 319-341.

Sweeney, P., & Moreland, R. (1980). Self-schemas and the perseverance of beliefs about the self. Paper presented at the meeting of American Psychological Association, Montreal.

Vasilopoulos, N., Reilly, R., & Leaman, J. (2000). The influence of job familiarity and impression management on self-report measure scale scores and response latencies. *Journal of Applied Psychology*, 85, 50-64.

Yoo, D., Lee, J., & Kim, H. (2012). A study on the comparative analysis of the historical transformation process on employment pattern and characteristics by the period of Korea major enterprise. *The Korean Academy of Business History*, 27(4), 33-58.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in

organizational behavior: A reconception. *Academy of Management Review*, 12, 250-264.

Zickar, M., & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement*, 20, 71-87.

저자 정보

정수임(은계중학교 교사)

신동희(이화여자대학교 교수)