



과학 관련 정서적 영역 검사 도구 활용 및 개선 내용 분석

정수임¹, 신동희^{2*}

¹은계중학교, ²이화여자대학교

Analysing the Use of Test Tools in Science-Related Affective Domain and its Improvements

Sue-Im Chung¹, Donghee Shin^{2*}

¹Eungye Middle School, ²Ewha Womans University

ARTICLE INFO

Article history:

Received 05 February 2019

Received in revised form

16 February 2019

12 March 2019

Accepted 25 March 2019

Keywords:

Science-related affective characteristics, test tool revision, item improvement, research trend, science education assessment

ABSTRACT

This study analyzed the current use of science-related test tools in affective area in Korean academic journals for the last 30 years since 1989 and the cases of modification of the original tools related to the tools. This provides practical information on the development and improvement of test tools in the future and offers the basis for developing valid test tools. The study was conducted in two stages. We analyzed the status of 88 domestic science education researches dealing with testing tools related affective characteristics and 57 domestic and foreign resource tools referenced by them. After that, we compared the characteristics of the five cases in which the original tool was revised and the concrete item correction examples. The results were analyzed as follows: topics that have emerged continuously or that have recently begun to appear in domestic research, the research situations according to their purpose and measurement methods used by the tool, and use state of original tools. The background and key points of revision were analyzed for the five tools, TOSRA, SAI, PISA items, STEBI-A/B, SMQ, whose revision process were disclosed. In addition, the changes of the items were analyzed by comparing the four test items with the questions before and after the revision. As a result of item analysis, item improvement was proceeding in the direction of enhancing readability, clarifying meaning, and changing items according to change of sub constructs. The implications of this study are as follows: notes on interpreting the results obtained from the tool with multiple mixed constructs, the possibility of the testing tool that broadens the horizon of the research, error recognition in the subjective self-report, and notes when applying test tools from other disciplines to science.

1. 서론

지난 50여 년간 과학 교육에서는 표준화된 측정 도구를 개발하는 연구가 활발히 수행되어 왔다(Doran, Lawrenz, & Helgeson, 1994; Tamir, 1998; Britton & Schneider, 2007). 표준화된 검사에 대한 관심이 꾸준히 증가하는 이유를 Liu(2012)는 세 가지로 정리했다. 첫째는 표준화 자료에 기반한 과학 책무성 평가에 대한 관심 증가, 둘째로 질적 연구에 대한 한계 인식과 이로 인한 표준화된 측정값을 포함한 무작위 실험 연구 필요성 대두(NRC, 2002), 마지막으로 학생들의 대안 개념 확인을 위한 대규모 조사의 지속적 필요성 등을 이유로 들었다. 개인적 성취, 프로그램 혹은 교육과정 측면에서 교육의 성과를 평가하는 과정은 단순히 성취의 공과(功過)를 넘어 과학 교육 과정 전반에 환류하면서 지속적으로 발전 가능한 교육과정의 기반을 마련하게 된다. 이때 교육 결과나 학생의 상태를 측정하는 검사 도구는 과학 교육 연구자들이 수립한 가설을 검증하는 수단일 뿐 아니라 또 다른 연구에 영감을 주는 원동력이 될 수 있다.

현재 국내 연구에서 과학 관련 태도를 측정하는 도구로 활발하게 사용되는 TOSRA(Test of Science Related Attitude)(Fraser, 1981)의 초판본은 5개의 척도에 따라 각각 다른 원래 도구(이하 원도구)로

이루어졌다. Fraser(1977)가 밝힌 5개 척도에 대한 원도구는 각각 Ormerod(1971), Meyer(1969), White & Mackay(1974), SCPESTM(1973), 그리고 Loughton & Wilkinson(1968)의 연구 결과를 채택하고 수정했다고 한다. 이는 여러 검사 도구들로부터 영감을 받아 개발된 검사 도구가 타당화되고 다른 사회·문화권으로 확장되며, 나아가 오랜 기간 동안 다수의 연구자들에게 채택되어 수많은 연구를 산출한 사례 중 하나다. 검사 도구의 개발 과정에서 연구자들은 해당 도구의 모체가 되는 원도구를 언급하게 되는데, 실제 과학 교육 현장에서 활발히 사용되는 많은 검사 도구들이 기존 원도구의 문항이나 구인들을 참고해서 수정했다고 밝혔다. 예를 들어, 과학에 특화된 학습에 대한 동기를 측정하는 SMTSL(Students' Motivation Towards Science Learning)(Tuan, Chin, & Shieh, 2005)은 일반적 학습에 대한 동기 검사 도구인 MSLQ(Motivated Strategies of Learning Questionnaire)(Pintrich & DeGroot, 1990), PALS(Patterns of Adaptive Learning Survey)(Midgley, Maehr, & Urdan, 1993), MMI(Multidimensional Motivational Instrument)(Uguroglu, Schiller, & Walberg, 1981)의 문항을 과학에 맞추어 질적으로 수행된 선행 연구 결과를 추가해 조정했다. 이렇게 기존 검사 도구를 연구자의 목적에 맞도록 수정해서 새로운 검사 도구를 개발하는 일련의 과정을 분석해 보면 검사 도구를

* 교신저자 : 신동희 (donghee@ewha.ac.kr)
http://dx.doi.org/10.14697/jkase.2019.39.2.263

활용하는 과학 교육 연구에 대한 몇 가지 시사점을 얻을 수 있다. 첫째, 타당하게 개발된 검사 도구는 새로운 검사 도구를 개발하기 위한 자양분이 될 수 있으며, 이를 통해 새로운 잠재 변수를 다루는 연구를 창출해낼 수 있다. 둘째로, 수많은 연구 목적에 딱 맞는 검사 도구를 발견하기 어려운 현실에서 기존 검사 도구를 수정하거나 개선할 때 지침이 되는 다양한 사례와 정보를 얻을 수 있다. 한편, 25년여 만에 개정을 시도한 SAI(Scientific Attitude Inventory)는 동 기간 동안 일어난 과학 교육적 변화 뿐 아니라 사회 문화적 변화에 부응해야 함을 개정 배경으로 밝혔다(Moore & Foy, 1997). 이들이 25년 전 개발한 검사 도구를 수정하게 되는 중요한 계기는 동료 연구자들이 SAI를 활용한 연구 결과를 지속적으로 축적해 왔고, 이를 근거로 제시한 비평들(Munby, 1983; Nagy, 1978)을 연구자들이 적극 수용해서 수정했다. 이는 검사 도구를 개발한 이후 지속적으로 나타난 문제점을 개선하고 보완하는 과정을 드러냄으로써, 후속 연구자들이 검사 도구 개선을 중심으로 한 연구에 도움이 되는 정보를 제공했으며, 지속 가능한 검사 도구의 조건과 면모를 구체화했다.

Liu(2012)는 1990년 이후 개발된 과학 관련 표준화 검사 도구를 조사한 결과를 개념적 이해, 인지적 논리, 태도, 과학의 본성, 학습 환경, 교사의 신념과 수행 등으로 분류해서 나타냈다. ERIC에서 검색한 총 229편의 관련 연구에서 검사 도구 개발에 대한 연구 49편을 추출했다. 이중 18편은 개념적 이해와 인지적 논리 등 인지적 영역에 해당하고 이외에 태도 11편, 학습 환경 9편, 교사의 신념과 수행 6편, 과학의 본성 5편으로 나타났다. 이들이 추출한 검사 도구 관련 연구가 1990년 이후 20여 년 간 개발된 모든 과학 교육 관련 검사 도구를 망라하지 못한다 해도 검사 도구 개발 동향에 대한 중요한 시사점을 얻을 수 있다. 과학 교육에서 인지적 영역 뿐 아니라 정의적 영역 특성을 측정하고자 하는 관심이 지속되고 있다는 점이다. 과학 교육에서 정의적 영역의 대표적 특성으로 ‘태도’를 중심으로 한 연구가 활발히 진행되어왔지만, 연구자에 따라 다양한 특성-태도, 흥미, 가치, 선호, 자아 개념, 인성, 도덕성, 포부 수준 및 동기, 성취 동기, 불안, 사회성, 정서 등을 정의적 영역으로 분류하기도 한다(Anderson & Bourke, 2000; So *et al.*, 2000; Kim, Jung, & Shin, 2015; Shin *et al.*, 2017a). 이렇게 태도 이외에 다양한 정의적 특성을 주제로 활발한 연구가 진행되기 위해서는 각 정의적 특성을 구체적으로 정의하고 측정할 수 있는 검사 도구가 필요하다.

전통적으로 과학 교육에서는 정의적 측면보다 인지적 측면의 성취를 더 강조하는 경향이 있었다(Kim *et al.*, 2014). 그러나 과학을 구축하는 과정 또는 과학을 학습하는 과정에서는 과학자나 학습자의 정의적 측면이 공명(共鳴)할 때, 과학 지식을 현실화하고, 그 지식 체계를 학습을 통해 효과적으로 전수할 수 있게 된다. 이는 과학 지식의 구축 과정에서 과학자의 호기심, 의지, 자부심, 개방성, 협동성, 비판적 태도 등과 같이 정의적 측면이 관여하며(Kim *et al.*, 2014), 학생의 인지적 성취와 정의적 특성이 높은 상관성이 있다는 연구 결과(Shen & Pedulla, 2000; Jürges, Schneider, & Büchel, 2005; Park, 2007; Seo, Choi, & Kim, 2007; Kim *et al.*, 2009; Kim & Seo, 2011; Kim, Kim, & Park, 2014; Ku *et al.*, 2016)로써 뒷받침된다. 최근 과학 교육은 정의적 성취를 복돋으려는 다양한 노력을 기울이고 있지만, 평가의 측면에서 정의적 특성을 측정하기 어려운 방법상 문제점을 안고 있다(Sung, 2010). 더불어 정의적 성취 결과를 개인과 과학 교육 과정

및 프로그램의 향상을 위해 쓸모 있게 활용하는 방안을 모색해야 할 과제도 있다. 따라서 정의적 특성을 측정하는 검사 도구의 개발과 개선은 거시적으로는 인지적·정의적으로 균형 잡힌 과학 교육을 구현하는 책무성과 관련한 문제가 된다. 동시에, 미시적으로 개인의 상태와 성취를 진단해서 학습에 환류하면서 학습의 향상 뿐 아니라 향후 개인의 성장과 사회적 성취에 영향을 미칠 수 있다.

과학 관련 정의적 영역을 다루는 검사 도구에 대한 국내 동향 연구는 과학 교육 평가 동향 연구(Chung & Shin, 2016), 정의적 영역 검사 도구에 대한 조사 연구(Shin *et al.*, 2017a) 등이 있다. Chung & Shin(2016)은 30년간 국내 학술지에서 과학 교육 평가 연구를 분석한 결과, 정의적 영역 평가를 정량화, 객관화하려는 시도가 상대적으로 적었음을 지적했고, 최근 들어 정의적 영역을 구체화하려는 연구가 많이 나타나고 있음을 보고했다. Shin *et al.*(2017a)은 최근 10년 간 2개 학술지를 분석해 정의적 영역을 다룬 검사 도구 연구를 추출했고, 이들을 과학 관련 태도, 과학 학습 동기, 자아 개념, 학습 정서, 진로 포부 등으로 나누어 검사 도구의 하위 구인을 자세히 소개했다. 이들의 연구는 과학과 정의적 영역 평가 연구의 전반적 경향을 분석하거나(Chung & Shin, 2016), 정의적 검사 도구의 개요 및 현황(Shin *et al.*, 2017a)을 횡단적으로 다루고 있어서 검사 도구의 수정이나 개선과 같이 종단적 변화 과정을 파악하기에는 한계가 있다. 검사 도구가 수정·보완되는 과정을 쫓아가며 분석한 사례를 통해 새로운 검사 도구의 제작과 개발에서 마주치는 시행착오를 줄일 수 있고, 문제의 원인이 될 내용을 조기에 차단하거나 이해할 수 있게 된다. 또한 검사 도구가 참고한 원래의 검사 도구를 연구자들의 다양한 연구 목적에 따라 변경, 조정하고 타당화하는 사례는 측정할 수 있는 정의적 특성의 범위를 확장할 수 있도록 후속 연구에 영감을 제시할 수 있다.

따라서 본 연구에서는 과거 30년간(1989~2018년) 과학과 정의적 영역 평가에서 검사 도구와 관련한 국내 연구가 진행된 동향을 분석했다. 더불어 정의적 영역 검사에 활용된 원도구를 거슬러 조사해서 현황을 파악하고, 특히 원도구가 개정되는 과정에서의 변화 경향과 수정되는 문항 사례를 구체적으로 제시함으로써 타당한 정의적 영역 평가 도구 개발을 위한 근거와 시사점을 도출했다.

II. 연구 방법

1. 분석 대상

본 연구는 1989년부터 2018년까지 30년 동안 국내 학술지에 발표된 연구 중 검사 도구를 중심으로 한 과학과 정의적 영역 평가 관련 논문을 분석했다. 우선 한국교육학술정보원의 데이터베이스(RISS) 검색을 통해 나타난 연구들 중 정의적 영역 평가를 내용으로 하는 연구들을 정리했다. 검색어는 ‘과학 교육’, ‘과학’, ‘물리’, ‘화학’, ‘생물’, ‘지구과학’의 영역 내용을 ‘평가’, ‘측정’, ‘검사’, ‘정의적’과 차례로 조합했으며, 검색 결과 온전히 인지적 영역 평가만 다룬 경우는 분석에서 제외했다. 연구의 제목에 검사 도구가 명시적으로 나타나지 않는 경우를 분석에 포함할지 여부는 검사 도구 활용이 연구에 중요한 방법이었거나 분석 결과와 밀접하게 연계되는 경우로 제한했다. 예를 들어 ‘우리나라 중학교 2학년 학생들의 과학에 대한 정의적 태도 특성 탐색’ (Kwak, 2017)은 분석 대상에 포함했다. 총 분석 대상

논문은 88편이다.

다음은 선정된 88편의 논문에서 참고한 것으로 언급한 원래의 검사 도구를 정리한 후, 특별히 전체 문항의 개정 과정이 논문이나 연구 보고서로 발표된 경우 구체적 수정 사례를 조사했다. 국내 연구 88편에서 언급한 주요 원도구로 국내 연구 28편, 국외 연구 29편을 분석했고, 개정 과정이나 문항이 공개된 5개의 검사 도구, 즉 TOSRA, SAI, PISA2006, PISA2015, STEBI-A/B(Science Teaching Efficacy Belief Instrument-A/B), SMQ(Science Motivation Questionnaire) 등에서 개정 주요 내용을 분석했으며, 문항으로 TOSRA를 제외한 4개 도구를 분석했다(Table 1).

Table 1. Number of papers and items analyzed

국내 연구	주요 원도구		연구 수	개정 과정 분석	
	국내	국외		문항 수	
				개정 전	개정 후
88	28	29	5	145	97

2. 분석 영역 선정 및 분석 방법

본 연구는 크게 두 단계로 진행되었다. 첫 번째는 정의적 영역 검사 도구를 중심으로 진행된 국내 과학 교육 연구 현황을 짚어보면서 검사 도구 활용의 제한점이나 개선점을 점검한다. 두 번째는 이들이 활용한 원도구를 조사하고 도구가 수정되고 개선되는 과정에 대한 구체적 사례를 제시한다. 도구 개선 과정의 실제 사례는 과학 교육 현장에서 원도구가 연구 목적에 맞도록 수정되어 활용될 때 타당하게 기능할 수 있는 실질적 근거를 제시해 줄 수 있다. 따라서 실제 과학 교육 현장에서 정의적 영역 검사가 어떤 목적과 방법으로 무엇을 측정하며, 어떻게 변화해 가는지 파악하는 것이 급선무다. Table 2에서 국내 연구 현황은 88편의 연구를 분석해서 연구 목적, 도구의 평가 방법, 주제, 도구 출처를 귀납적으로 유형화했다. 이들이 기존의 도구를 활용한 경우 원도구가 다른 주제, 공개 형태, 대상 과목(영역)을

조사했다. 널리 사용되는 원도구 중 개정 과정이 공개된 5개의 사례에 대해서는 개정된 주요 내용과 함께 문항 수준에서 수정된 사례를 구체적으로 분석했다(Table 2). 단, 1977년 판 TOSRA(Fraser, 1977)는 전체 문항이 공개되지 않아 문항 분석 사례에서 제외했다.

과학 관련 정의적 영역 검사 도구를 활용한 국내 과학 교육 연구의 선정에는 연구자 중 1인이 연구 제목을 보면서 인지적 영역을 다루는 경우, 검색 결과에서 중복된 경우, 학술 발표 초록인 경우, 검사 도구와 관련 없는 경우 등을 제외했다. 인지적 영역과 정의적 영역의 평가가 혼재된 경우 정의적 영역에 제한해서 분석 대상에 포함했다. 이렇게 선정된 88편의 논문을 읽은 후 연구자 2인이 협의를 통해 Table 3과 같이 연구 목적, 주제 등을 범주화했다. 특히 국내 연구에서 활용한 원도구는 연구 방법에서 명시적으로 언급되어 있는 경우로 한정했고, 언급된 원도구에서 참고한 원래 도구는 분석 대상 연구에서 명시적일 경우에 함께 분석했다. 예를 들면 Kim & Lee(2018)에서 영재 학생들의 정서 반응을 측정할 ‘간편형 과학 학습 정서 검사지’는 Lee(2017)의 도구를 활용했지만, 이는 Pekrun *et al.*(2002)의 AEQ(Achievement Emotions Questionnaire)의 문항을 번역한 후 선별해서 적용했음을 명시적으로 밝혔기 때문에 원도구는 AEQ로 판단했다. 국외에서 개발한 원도구의 문항을 단순 번역, 번안한 경우는 원도구를 그대로 사용한 경우로, 원도구 문항에 자체 개발한 문항을 추가한 경우는 원도구를 수정한 경우로 분류했다.

한편, 국내 연구에서 빈번히 사용하는 원도구로 언급한 검사 도구 중 개발 과정이 공개된 과학과 정의적 영역 검사 도구는 1977년과 1978년 판 TOSRA(Fraser, 1977; Fraser, 1978), STEBI-A와 B(Enochs & Riggs, 1990; Riggs & Enochs, 1990), SAI와 SAI II(Moore & Sutman, 1970; Moore & Foy, 1997), SMQ와 SMQ II(Glynn & Koballa, 2006; Glynn, Taasobshirazi, & Brickman, 2007, 2009; Glynn *et al.*, 2011), 그리고 PISA2006과 PISA2015 문항(OECD, 2007, 2016)으로 조사되어 이들의 구인 변화 및 개정 주요 사항을 분석했다. 특히 개정 전 후 문항의 변화를 비교했고, 수정된 사유를 Table 2와 같이 유형화했다.

Table 2. Analysis criteria

범주		세부 범주			
국내 연구 현황	연구 목적*	1. 도구 개발·검증·수정	2. 정보 제공	3. 현상 조사	4. 효과 검증
	도구의 평가 방법	1. 자기보고식 평정(리커트)	2. 질문·서술	3. 포트폴리오	4. 타인 평정
	주제*	1. 태도(일반) 2. 태도(특정 주제, 대상) 3. 다수 요인 4. 효능감 5. 동기 6. 자기/동료 평가 7. 신념 8. 확산성 9. 인식 10. 감정과 감성 11. 과제 집착력 12. 과학 경험 13. 성취·학습 정서 14. 자기 주도성 15. 행동 평정 16. 흥미 17. 과학 교실 문화 18. 과학 교육 능력 19. 과학 윤리성 20. 몰입 21. 불안 22. 심리적 학습 환경 23. 인식			
	검사 도구 출처	1. 원도구	2. 원도구 수정	3. 자체 개발	
원도구	주제	국내 연구 현황의 주제와 같음			
	공개 형태	1. 학술지 논문	2. 학위 논문	3. 기타(연구 보고서, 서적)	
	대상 교과 영역	1. 과학	2. 타과목(영역)	3. 공통	
개정 과정 사례	기본 사항	1. 하위 영역	2. 연구자	3. 측정 개념	
	개정 주요 내용	구인의 변화, 평가 방법의 변화, 문항 수, 타당화 과정 등			
	문항 수	1. 삭제	2. 신설		
	표현 명료화	1. 단순 교정 2. 복문·중문의 단문 교체 3. 의미 중복 구, 절, 수식어 제거 4. 간단한 문장 교체			
	의미 구체·명확화	1. 반응 제한 표현 제거 2. 지시문 연계 수정 3. 구체적 상황 제시 4. 의미 명확한 표현 5. 강조 기호 추가			
주요 사항 교체	1. 쉬운 단어·용어 교체 2. 주요 용어 교체 3. 성차별적 언어 교체 4. 시제 교체				

* 분석 대상 연구 결과 귀납적으로 분류한 범주임.

III. 연구 결과 및 논의

1. 과학과 정의적 영역 검사 도구 활용 연구 현황

가. 주제별 연구 현황

1989년부터 2018년까지 30년 동안 학술지에 발표된 국내 연구 현황을 조사했다. 분석 기간 동안 시기에 관련 없이 전체적으로 고르고 빈번하게 나타난 측정 개념은 ‘태도’다. 우리나라 과학과 교육과정에서 정의적 특성을 향상하고자 하는 목표는 인지적 성취와 함께 항상 교육 목표의 두 축을 이루어 왔지만, 이를 평가하고자 하는 시도는 비교적 뒤늦게 시작되었다. Table 3은 국가 수준 교육과정 문서에서 과학과 정의적 영역 평가를 어떻게 다루고 있는지 정리한 것이다. 제4차 교육과정서부터 ‘과학에 대한 태도’가 명시되었고, 제5차에는 ‘태도’(제5차)로 변경되었다가 제6차 교육과정 이후 ‘과학적(인) 태도’가 현재까지 지속되고 있다. 2007년 개정 교육과정서부터는 ‘과학에 대한 흥미와 가치 인식’ 등의 평가 예시를 통해 사실상 ‘과학에 대한 태도’ 또한 함께 제시되고 있다(MOCE, 1981; MOCE, 1987; MOE, 1992; MOE, 1997; MOEHRD, 2007; MOEST, 2011; MOE, 2015).

Kim, Jung, & Shin(2015)은 30여 년간 국내 과학 교육 학회지를 분석한 결과, 정의적 영역을 크게 태도와 기타 영역으로 분류한 바 있다. 그들은 태도 영역을 과학적 태도 외에 과학, 과학 학습, 과학 교수 관련으로 재분류했다. Shin *et al.*(2017a)도 10년간의 정의적 영역 검사 도구를 조사하면서 태도 영역을 과학에 대한 태도, 과학 학습에 대한 태도, 과학적 태도로 분류한 바 있다. Table 4는 88편의 분석 대상을 연구 목적과 연구에 나타난 평가 방법에 따라 주제별로 정리

한 것이다. 하나의 연구에서 다수의 연구 목적이나 평가 방법을 활용할 수 있으므로 중복된 수치로 나타나기도 하지만, 유독 태도는 일반 영역(24편)과 특정 주제 및 대상(12편)을 통틀어 가장 많이 나타났다. 이는 ‘태도’라는 개념이 연구자에 따라 다양하게 정의되기 때문에 오히려 분명히 정의하기 어려운 측면(Koballa, 1988; Krynowsky, 1988)이 있음을 의미하기도 한다. 특히 과학에 대한 태도를 정의하는 한 예로 Hasan & Billeh(1975)는 ‘느낌, 의견, 신념, 정서적 반응’으로 Munby(1980)는 ‘느낌과 신념’과 관련된 것으로 언급한 바 있다. 따라서 다양하고 복합적인 개념을 ‘태도’로 단순화함에 따라 태도는 다양한 하위 구인을 통합적으로 대표하는 용어로 기능했다고 볼 수 있다. PISA 2006 문항에서 과학에 대한 자아 개념, 과학의 즐거움, 과학에 대한 가치, 도구적 동기 등을 태도로 나타낸 사례가 그 예다(OECD, 2007). 일반적인 과학에 대한 혹은 과학적 태도는 특정 주제와 대상에 대한 태도(12편)로 응용되기도 했다. 과학의 세부 영역이나 주제 예컨대 환경 오염, 실험 실습, 양면 가치, 융합, 물리 이미지 등에 대한 태도나 특정 학습-형성 평가, 수행 평가 등에 대한 선호나 인식 등 주제 특이적 연구와 함께 예비 교사, 영재 학생 등의 대상 특이적 연구가 이에 해당한다. 연구자의 관심이 반영된 특정 주제에 대한 태도를 측정하기 위해 검사 도구를 자체 개발하는 경우(Jeong & Kim, 1993), 국내외에서 개발된 도구를 활용하는 경우(Kim, Hong, & Chang, 2001), 일반적 태도 도구를 특정 주제에 맞도록 수정하는 경우(Lee, Choi, & Nam, 2000) 등이 나타났다.

2000년대 이후로는 효능감(8편), 동기(5편), 자기/동료 평가(4편), 성취 및 학습 정서(4편), 확신성(3편) 관련 연구가 지속적으로 등장하고 있다. Table 3에서 제6차 교육과정 이후(1995년~)로 국가 교육과정에서도 과학적 태도 이외에 정의적 영역에 대한 구체적인 예시로서 흥미, 가치 인식, 적극성, 협동성, 인성과 감성 등이 명문화됐다(MOE,

Table 3. Science-related affective domain assessments mentioned in National Science Curriculum

교육과정	시행 연도 (중 1학년 기준)	국가 수준 교육과정 문서의 특징	평가 관련 특이 사항
제4차 교육과정	1984	<ul style="list-style-type: none"> 국가 교육과정에 ‘평가’ 항목 신설 태도 평가 언급 ‘과학에 대한 태도’를 평가 	
제5차 교육과정	1989	<ul style="list-style-type: none"> 평가 영역에 대한 방법 예시 ‘태도’를 평가 	<ul style="list-style-type: none"> 1994년 대학수학능력시험
제6차 교육과정	1995	<ul style="list-style-type: none"> ‘지도 및 평가 상의 유의점’에서 평가 항목으로 독립 용어 변화: 태도→ 과학적인 태도 ‘과학적인 태도’의 구체적 특성 예시: 학습 과정에서 계속 탐구하려는 의욕, 상호 협동, 증거를 존중하는 태도 등 	<ul style="list-style-type: none"> 수행평가 실시 TIMSS1995 TIMSS1999
제7차 교육과정	2001	<ul style="list-style-type: none"> 용어 변화: 지식→기본 개념 제6차 교육과정과 유사 	<ul style="list-style-type: none"> 국가수준학업성취도 평가 TIMSS 2003/2007 PISA 2003/2006 /2009
2007년 개정	2010	<ul style="list-style-type: none"> 정의적 영역의 예시 변화: 과학에 대한 흥미와 가치 인식, 과학 학습 참여의 적극성, 협동성, 과학적으로 문제를 해결하는 태도, 창의성 등 과학적 태도 외에 과학에 대한 태도 언급 자유 탐구 신설 	
2009년 개정	2011	<ul style="list-style-type: none"> 평가의 성취 기준 신설 	<ul style="list-style-type: none"> 성취 평가제 TIMSS 2011/2015 PISA 2012/2015
2015년 개정	2018	<ul style="list-style-type: none"> 독립 항목에서 ‘교수 학습 및 평가의 방향’의 하위 항목으로 체제 변경되었으나 보다 구체적, 세분화된 내용 인성과 감성 함양에 도움이 되는 소재와 상황 발굴 협동심 함양을 위한 모듈 평가 	

1992; MOE, 1997; MOEHRD, 2007; MOEST, 2011; MOE, 2015). 태도가 주를 이루었던 정의적 영역에서 측정 개념이 보다 다양해졌다. 이는 국내외 연구에서 사회 인지 학습 이론(Bandura, 1977, 1986), 학습 동기 이론(Keller, 1984; Brophy, 1988), 사회 정서 학습(Elias *et al.*, 2008; Merrell & Gueldner, 2010) 및 성취 정서 이론(Pekrun *et al.*, 2002) 등이 소개되면서 학습에 영향을 미치는 정의적 특성들에 대한 관심이 구체화되었기 때문이다.

그밖에 세부적 주제에 대한 연구는 1~2편정도 나타나고 있다. 흥미, 감정·감성 등은 기존 태도를 측정하는 도구 속에 포함되어 있었던 흥미나 감정 구인을 추출해 효과를 검증하거나 도구를 수정하는 연구(Lee & Woo, 1996; Kwak & Shin, 2014)와 함께 도구를 자체적으로 개발하는 연구(Jeong, 2007)로 나타났다. 이는 기존 태도 검사 도구에 성격이 다른 구인들이 혼재되어 있음을 단적으로 보여주며, 측정 개념과 구인을 명료화할 필요성을 제시한다. 한편 과학 경험(Kim & Ryu, 2010; Shin *et al.*, 2017b), 과학 교실 문화(Chun *et al.*, 2015), 심리적 학습 환경(Lee & Kim, 1998), 과학 윤리성(Kim & Kim, 2016) 등 학습이나 학업 성취와 관련한 다양한 맥락을 밝히려는 시도에도 관심이 지속되고 있었다.

나. 연구 목적 및 도구의 평가 방법에 따른 연구 현황

연구 목적은 4가지로 분류했다. 도구를 개발해서 검증하거나 수정하는 도구 관련 연구와 함께 정보 제공, 현장에 대한 조사 및 분석, 그리고 수업이나 프로그램의 효과를 검증하는 연구가 이에 해당한다. 앞서 분석한 23개의 주제별로 각 연구 목적에 따른 연구의 빈도수를 정리했다(Table 4). 가장 많이 나타난 것은 도구의 개발·검증·수정(35편) 관련 연구였다. 다음으로 집계된 효과 검증(30편)과 현상 분석 및 조사(20편) 연구는 검사 도구가 활용된 현황을 나타낸다. ‘태도(일반)’의 경우 개발·검증·수정 연구(7편)도 많지만 그보다 현상을 조사하거나 효과를 검증하는 등의 활용 관련 연구(14편)가 더 많이 나타나고 있다. 특정 주제나 대상에 대한 태도 검사 도구를 개발한 연구는 8편으로, 같은 기간 더 많이 개발되었음에도 그를 활용한 연구(4편)는 일반 태도보다 적게 나타난 결과와 대비된다. 주제나 대상이 특정된 검사 도구는 활용 가능 범위가 다소 축소될 수 있다는 점에서 설명 가능한 결과다. 이를 다른 측면에서 말하면, 특정 주제나 대상에 대한 연구자의 관심이 지속되지 않으면, 후속 연구로 진행되지 않고 개발에만 그치는 1회용 검사 도구로 사장될 가능성이 있다. 반면, ‘효능감’의 경우 도구 개발 연구는 2편이지만, 도구를 활용한 연구는 6편으로 나

Table 4. Research organized by topics according to research purpose and measurement method from 1989 to 2018

주제	주제별 연구 목적 및 평가 방법에 따른 정의적 영역 평가 관련 연구 수(총 88편)											
	연구 목적					소계	평가 방법					소계
	도구 개발 검증 수정	정보 제공	현상 조사	효과 검증	자기보고식 평정		질문 서술	포트폴리오	타인 평정	기타		
태도(일반) ^{1, 2}	7	3	3	11	24	19	1	3	1	3	27	
태도(특정 주제, 대상)	8	0	1	3	12	10	0	0	0	2	12	
다수 요인	1	6	2	1	10	2	0	0	1	7	10	
효능감 ¹	2	0	4	2	8	8	0	0	0	0	8	
동기	2	0	0	3	5	4	1	0	0	0	5	
자기/동료 평가 ²	1	0	3	0	4	2	1	0	2	0	5	
성취 정서, 학습 정서 ²	2	0	1	1	4	4	1	0	0	0	5	
확신성 ¹	1	0	2	0	3	3	0	0	0	0	3	
인식	0	0	2	1	3	2	1	0	0	0	3	
신념 ¹	0	0	1	1	2	1	1	0	0	0	2	
감정과 감성	2	0	0	0	2	2	0	0	0	0	2	
과제 집착력 ²	2	0	0	0	2	1	0	0	1	1	3	
과학 경험	2	0	0	0	2	2	0	0	0	0	2	
자기 주도성 ^{1, 2}	0	0	0	2	2	2	1	0	0	0	3	
행동 평정	1	0	0	1	2	0	0	0	2	0	2	
흥미 ^{1, 2}	0	0	0	2	2	2	1	0	0	0	3	
과학 교실 문화	1	0	0	0	1	1	0	0	0	0	1	
과학 교육 능력	0	0	0	1	1	0	0	1	0	0	1	
과학 윤리성	1	0	0	0	1	1	0	0	0	0	1	
몰입	0	0	0	1	1	1	0	0	0	0	1	
불안	1	0	0	0	1	1	0	0	0	0	1	
심리적 학습 환경	1	0	0	0	1	1	0	0	0	0	1	
인성	0	0	1	0	1	0	0	0	1	0	1	
계	35	9	20	30	94	69	8	4	8	13	102	

1. 연구 목적에서 중복 집계된 경우
2. 평가 방법에서 중복 집계된 경우

타났다. 이는 국외에서 개발된 원도구를 그대로 채택해서 연구에 활용했기에 나타난 결과이며, 효능감에 대한 연구자들의 관심이 활발히 진행되고 있음을 의미한다. 대조적인 두 사례에서 얻을 수 있는 시사점은 타당한 검사 도구의 존재가 연구자들의 영감을 북돋으며 후속 연구를 촉진할 수 있는 원동력이 된다는 점, 그래서 타당하게 개발된 검사 도구의 등록과 검색, 인출이 체계적으로 관리될 필요가 있다는 점이다. 도구를 활용하려는 연구자가 연구 목적에 적합한 기존 검사 도구를 찾지 못할 경우 목적에 맞도록 스스로 도구를 제작해서 사용하게 된다. 이렇게 도구 개발이 주 목적이 아닐 경우 타당화 과정은 내용 타당도 정도로 약화되는 경우(Lee & Hur, 1998; Hwang *et al.*, 2001; Park & Paik, 2001; Lee & Hann, 2001; Kim, Jeong, & Hwang, 2005; Hong, 2014)가 발견되기도 한다. 최근 들어 과학 교육에서도 검사 도구의 측정학적 중요성이 강조되는 추세이며, 검사 도구를 타당화하려는 연구가 활발히 진행되고 있다(Ha & Lee, 2013; Woo, Rachmartullah, & Ha, 2017; Park *et al.*, 2017; Shin *et al.*, 2017; Shin, Ha, & Lee, 2016). 연구자가 많은 노력을 들여 개발한 도구를 등록하고 관리하는 체계를 구축한다면 다른 연구자들이 자신의 연구 목적에 맞는 타당한 도구를 검색한 후, 최소한의 수정을 거쳐 이를 활용할 수 있을 것이다. 이로써 후속 연구로 발전시키고 촉진할 가능성도 높아질 것이다.

Table 5는 연구 목적에 따라 검사 도구를 어떻게 활용했는지 정리한 것이다. 기존에 개발된 원도구를 활용한 경우, 기존 도구에 새롭게 추가·변형되는 문항이나 구인들을 포함해 수정한 경우, 자체 개발한 경우 등 셋으로 분류했다. 도구를 검증하는 연구에서는 국외에서 개발한 검사 도구를 국내 현실에 맞도록 타당화 혹은 검증(Ha & Lee, 2013; Jo & Ku, 2013; Jo, 2011; Woo, Rachmartullah, & Ha, 2017)하는 경우(4편)가 나타났으며 이러한 경우에는 원도구를 번역한 수준에서 그대로 이용했다. 원도구를 수정하는 경우는 새로운 구인이나 문항을 추가 혹은 기존 도구에서 일부 구인만 선별하는 경우(Woo & Chung, 1994; Lee & Woo, 1996; Choi, Kim, & Kim, 2007; Kim & Ryu, 2010), 검사 대상의 변경에 맞춰 용어나 맥락을 수정하는 경우(Mun *et al.*, 2014; Jeon, 2014), 우리나라 실정에 맞추어 번역본을 타당화하는 경우(Park, 1997; Park *et al.*, 2017), 여러 원도구의 공통 요인을 추출하는 경우(Chun *et al.*, 2015) 등의 사례가 나타났다. 검사 도구를 자체 개발하는 경우는 원도구를 참고한 경우(Kim, Chung, & Jeong, 1998; Kim, Hong, & Chang, 2001; Kim & Kim, 2016)와 원도구의 도움 없이 선행 연구 결과나 학생으로부터 귀납적으로 도출된 내용을 근거로 제작한 경우(Jeong & Kim, 1993; Chung, Yeun, & Hur, 1994; Lee & Kim, 1998; Jeong, 2007; Song & Choi, 2017)가 나타났다. 예를 들어 Song & Choi(2017)는 학생들이 작성한 물리에 대한 이미지를 바탕으로 요인 분석을 통해 구인을 설정하고

집단 차이를 분석하는 방법으로 도구를 개발했다. 연구의 목적이 과학 교육 현장에서 나타나는 현상을 조사하거나 효과를 검증하는 경우에는 주로 원도구를 그대로 사용하거나(23편) 수정해서 사용하는(10편) 경우가 대부분이었지만, 소수 대상에 대한 간단한 인식을 조사하는 경우(Kim & Kim, 1999; Park & Paik, 2001; Park, Kim, & Jeon, 2007), 특수한 연구 설계를 적용한 경우(Hwang *et al.*, 2001)에는 자체 개발한 도구를 활용하기도 했다.

분석 대상 연구에서 활용한 검사 도구의 평가 방법은 크게 자기보고식 평정(리커트), 질문·서술, 포트폴리오, 타인 평정, 기타로 분류했다(Table 4). 자기보고식 평정은 리커트식 척도, 순위 매기기 등이 해당되며, 분석 대상 연구 중 리커트식 척도는 3점부터 7점까지 나타났고 5점이 가장 많이 나타났다. 포트폴리오는 일반적인 포트폴리오 외에 회고록이나 반성적 저널 등을 포함했고, 타인 평정은 루브릭에 근거해 학생의 수행을 관찰하는 경우로 분류했다. 기타 영역에는 신경 생리학적 방법인 시신 추적기(Kim *et al.*, 2012)를 포함했으나, 정보 제공을 목적으로 검사 도구가 드러나지 않는 대부분의 연구를 포함시켰다. 하나의 연구에 다양한 평가 방법을 활용한 경우는 중복해서 처리했다. 지난 30년 동안 과학과 정의적 영역 평가에서 가장 많이 활용된 평가 방법은 자기보고식 평가(69편)이며 특히 수치로 스스로를 평정하는 리커트 척도를 주로 활용했다. 질문과 서술을 활용하는 연구(8편), 포트폴리오(4편)는 주로 영재 학생, 교사, 특정 학생 등 주로 소수 인원을 대상으로 하는 현상 조사 및 효과 검증 연구(Byeon *et al.*, 1998; Lee & Hann, 2001; Park, Kim, & Jeon, 2007; Yoo, 2000; Ko *et al.*, 2013)에 한정되었다. 도구를 검증하거나 개발하는 연구에서는 거의 예외 없이 리커트 척도를 사용하는 자기보고식 형태를 사용했다. Liu(2012)는 척도의 간격과 시작점이 다른 점수를 더해서 총점으로 나타내는 리커트 방식에 대한 대안적 방법으로 원점수를 잠재적 점수로 변환하는 라쉬 모형 사용을 예로 들었다. 한편, 자기보고식 평정이나 질문 및 서술을 활용한 평가는 모두 학생의 주관적 자기 보고에 의존한다는 측면에서 측정학적 타당성에 관심을 가질 필요가 있다. Chung & Shin(2017)은 과학과 정의적 영역 검사 도구와 학생들이 상호작용하는 과정에서 측정 개념에 대한 상이한 이해와 함께 사회적 바람직성과 같은 심리적 기제가 작용해서 검사 결과에 영향을 줄 수 있음을 보고한 바 있다. 최근 국외의 검사 도구를 국내내용으로 타당화하는 과정에서 사회·문화적 배경을 고려하거나, 다양한 집단에 대해 차별적 문항 기능(differential item function: DIF)을 검사하는 연구들이 나타났다(Park *et al.*, 2017; Shin, Ha, & Lee, 2016). 그 배경에는 심리 측정학적 연구 결과를 과학 교육에 도입해서 검사 도구에 포함된 다양한 오차를 감소시키려는 시도가 이루어지고 있음을 알 수 있다.

Table 5. Use of test tool according to research purpose

연구 목적	연구 목적에 따른 검사 도구 활용 연구 수(총 76편)			계
	원도구	원도구 수정	자체 개발	
도구 개발, 검증, 수정	4	13	18	35
정보 제공	1	0	0	1
현상 조사	12	3	3	18
효과 검증	11	7	4	22
계	31	20	25	76

주: 연구 대상 88편 중 검사 도구를 활용하는 연구는 76편을 대상으로 분석

다. 국내 연구에서 활용한 원도구 현황

Table 6은 분석 대상 연구에서 언급된 원도구를 이들이 공개된 형태와 다루는 영역과 함께 정리한 것이다. 23개의 주제에 따라서 어떤 원도구를 사용했고, 어떤 경로로 얻을 수 있었으며, 과학에 특화된 도구인지 여부를 알 수 있다. 여러 주제에 걸쳐 나타나는 검사 도구는 TOSRA(1978, 1981)로 태도(일반)과 태도(특정 주제 및 대상)과 함께 감정과 감성에서도 이 도구를 활용하고 있다. TOSRA는 과학자, 사회, 탐구, 과학 학습, 취미, 직업 등 광범위한 과학 관련 맥락을 아우르는 7개의 하위 영역을 다루고 있어서 어떤 주제라도 변형하거나 발췌해서 응용할 수 있는 여지가 있다. 그밖에 Pintrich & De Groot(1990)의 MSLQ(Motivated Strategies for Learning Questionnaire)는 신념과 자기 주도성에서 각각 원도구로 언급된 경우가 있었다.

연구자들이 자신의 연구에 참고하거나 활용한 원도구는 학술지 논문을 통해서 공개된 경우가 32편으로 가장 많았고, 학위 논문을 참고한 경우가 14편, 연구 보고서나 서적을 출처로 한 경우가 13편으로 나타났다(Table 6). 다른 연구자의 학위 논문으로부터 원도구를 가져왔다고 밝혔지만, 실제로는 학술지에 게재가 되어 있었던 7편의 경우를 제외하면 14편 중 나머지 7편은 순수하게 학위 논문에서 개발한 원도구를 사용했다. 검사 도구를 개발한 학위 논문은 학술지에 발표되고 후속 연구로 연계되며 발전될 수 있다는 점에서 검사 도구의 다양한 출처로서 기능할 수 있지만 한편으로는 타당한 도구인지 검증하는 과정이 함께 강화되어야 한다. 또한 학술지에 공개되지 않은 학위 논문이라도 검사 도구의 측정 개념, 수검 대상, 하위 구인, 타당화 수준 등을 등록·관리하는 체제로 구축할 필요가 있다. 연구자는 타당화 수준을 보완한다는 전제하에 검사 도구와 관련한 후속

연구로 연계할 수 있으며, 이로써 양질의 검사 도구가 개발 이후에 채택되지 못하고 사장되는 현상을 방지할 수 있다.

주제별로 언급된 원도구는 TOSRA, STEBI-B, R-CBC, STAQ, CARS, SMQ II, SAI와 같이 과학 과목에 특화되어 개발된 경우가 34편으로 가장 많았으며, 어느 과목이나 적용 가능한 범교과적 검사 도구는 20편으로 Pajares(1996), Bandura(1997), Schunk(1991)의 효능감 문항, MSLQ, AEQ 등이 있다(Table 6). 타 과목이나 영역에서 개발된 검사 도구를 과학과에 도입한 경우도 5편으로 나타났다. 가정(Yu & Yoon, 2001), 수학(Jeon, 2011), 신경학 및 생리학(Byeon *et al.*, 2013; Reach *et al.*, 2005), 윤리학(Lee & Kim, 2013) 영역에서 사용된 검사 도구가 과학 과목에 맞도록 수정되었다. 특히 Byeon *et al.*(2013)의 시선 추적기는 정의적 특성을 객관적, 정량적으로 측정하기 위한 방법으로 신경 생리학적 방법을 제안했다는 점에서 자기보고식 방법과 차별화되었다. 비교적 오래전부터 논의되어 연구가 진행되어 온 태도와 같은 정의적 특성은 과학에 특화된 검사 도구가 대부분이다. 2000년대 이후부터 국내 연구에서 소개되기 시작한 효능감, 신념, 동기, 확산성 등은 교과에 관련 없이 범교과적 영역을 다루는 검사 도구와 함께 과학에 특화된 도구가 동시에 혼재한다. 이는 타 학문 영역에서 시작된 외부의 연구 문제가 과학 안으로 들어와 적용되어 확산되는 과도기적 과정에 있음을 의미한다. 반대로 감정과 감성이나 과학 경험과 같은 주제는 기존에 있던 과학 관련 검사 도구로부터 정의적 특성을 추출하는 형태로 연구가 진행되었다. 전자는 타 영역 연구 결과와 과학 관련 내용을 잘 연계해서 과학과에 특화된 검사 도구를 개발하는 방향으로, 후자는 다양한 정의적 특성을 구체화해서 측정 구인이 보다 명확해지는 방향으로 연구가 진행되고 있다.

Table 6. Original tool used in domestic studies

대상 교과	주제	공개 형태에 따른 원도구		
		학술지 논문	학위 논문	기타(연구 보고서 및 서적)
과학	태도(일반)	<ul style="list-style-type: none"> · Simpson & Troost의 STAQ(1982) · 이운환 외 과학에 관련된 태도 검사(1990) · 이경훈과 우종욱의 과학에 대한 태도(1996) · 정완호 외, 정진우 외의 과학 관련 태도(1994, 1998) · 김효남 외의 정의적 특성 문항(1998) · Krough & Thomsen의 R-CBC(2005) 	<ul style="list-style-type: none"> · 이영미의 정의적 특성 평가 도구(1997) · 최형숙의 평가 관점(1998)* 	<ul style="list-style-type: none"> · Fraser의 TOSRA(1981) · 이범홍, 김영민의 과학에 대한 태도(1983) · 김주훈과 이양락의 과학에 대한 태도(1984) · McDonough의 과학에 대한 부모의 지원(1995) · PISA 문항 · TIMSS 문항
	태도(특정 주제 및 대상)	<ul style="list-style-type: none"> · Dreyfus & Roth의 인식 검사(1991) · 정은영과 김영수의 환경 오염에 대한 태도 평가(1993) · Siegel & Ranney의 CARS(2003) 	<ul style="list-style-type: none"> · 김은량의 TOSRA 번역본(1991) · 강인숙의 검사 도구(1998) · 김경미의 과학 학습 태도 검사(1999)* 	<ul style="list-style-type: none"> · Fraser의 TOSRA(1981)
	효능감	<ul style="list-style-type: none"> · Enochs & Riggs의 STEBI-B(1990) · Dalgety, Coll, & Jones의 CAEQ(2003) · Uzuntiryaki & Aydin의 화학자기효능감 측정 도구(2009) 		
	신념	<ul style="list-style-type: none"> · Schwartz <i>et al.</i>의 VOSI(2008) 	<ul style="list-style-type: none"> · 임성민의 물리 학습에서 인지적 신념과 동기의 평가(2001)* · 팽애진의 과학 교수에 대한 설문지(2004)* 	
	다수 요인			<ul style="list-style-type: none"> · 위스콘신 대학의 SPA, SPS
	동기	<ul style="list-style-type: none"> · Glynn <i>et al.</i>의 SMQ II(2011) 		
	감정과 감성	<ul style="list-style-type: none"> · Fraser의 TOSRA(1978) · Moore & Sutman의 SAI(1970) 		
	과학 경험		<ul style="list-style-type: none"> · 신영준의 과학 경험도(2000)* 	<ul style="list-style-type: none"> · Klopfer의 탐구 기능(1971)
	자기 주도성		<ul style="list-style-type: none"> · 신영준의 과학 경험도(2000)* 	<ul style="list-style-type: none"> · Klopfer의 탐구 기능(1971)

	효능감	· Schunk의 효능감(1991) · Pajares의 효능감 문항(1996) · Bandura의 효능감(1997)
	신념	· Pintrich & De Groot의 MSLQ(1990)
	자기/동료 평가	· 배호순의 교사의 수업 전개 중심도로 한 평 가의 준거 체계(1992) · 박용현의 교사 자질 평정 척도(1994) · 원효현의 교사의 수업 평가(1997)
	행동 평정	· Renzulli <i>et al.</i> 의 SRBCSS(2009) · 박혜정의 영재 선발 루브릭(2011) · 신윤주의 영재 행동 체크리스트(2011)
범교과	다수 요인	· Spencer & Spencer의 역량 사전(2003)
	동기	· Keller의 수업 교재 동기 검사(1993)
	확신성	· Bruttomesso <i>et al.</i> (2003)
	자기 주도성	· Pintrich & De Groot의 MSLQ(1990)
	인식	· 윤진경의 설문지(2010)
	과학 교실 문화	· 이정은과 김진모의 실험 공동체 요인 척도 (2008)
	성취 학습 정서	· Pekrun <i>et al.</i> 의 AEQ(2011)
	몰입	· 이은주의 학습 몰입 검사(2001)
	불안	· Zuckerman & Odom의 SAS(1965)
	인성	· McDonald, Moore, & Freehil의 영재아 식 별(1982)
타교과 (영역)	태도 (특정 주제 및 대상)	· 유인영의 가정과 교사의 수행평가에 대한 태도(2000)*
	확신성	· Reach <i>et al.</i> (2005)
	인식	· 전소영의 설문지(2011)
	과제 집착력	· 변정호의 시선 추적 연구(2013)*
	과학 윤리성	· 김영한과 이승하의 청소년 도덕성 진단 검 사 도구(2012)

* 분석 대상 연구가 원도구의 출처를 학위 논문으로 밝혔으나, 해당 원도구가 이후에 학술지에 공개된 경우

2. 정의적 영역 검사 도구 개정 사례

가. 검사 도구 개정 사례

검사 도구는 학생의 성취, 수행, 상태, 특성을 측정함으로써 과학 학습 평가를 구현하는 중요한 수단이다. 검사 도구가 개발되면 예전에 감지하기 어려웠던 개인의 특성과 성취를 가시화할 수 있어서 개인적 성취와 더불어 학습을 돕고 프로그램의 효과를 평가할 수 있게 된다. 앞서 과학과 정의적 영역 도구를 활용한 연구 현황에서 살펴본 바에 의하면 하나의 원도구를 중심으로 연구 목적에 따라, 연구자의 관점과 관심에 따라, 사회·문화적 변화에 따라 다양하고 지속적으로 변주되고 있음을 알 수 있다. 동일한 검사 도구가 개정되는 다양한 이유와 그로 인해 변화된 전후 결과를 분석하는 연구는 앞으로 계속될 검사 도구 개발과 수정을 다루는 연구들에 실질적 정보와 조언을 제공해 줄 수 있다. Table 7은 국내 연구 현황 88편을 분석한 결과, 이들이 원도구로 참고한 검사 도구들 중 명시적으로 개정판이 공개된 검사 도구 5편에 대해 하위 영역, 연구자, 측정 개념, 개정 주요 사항 등을 정리한 것이다. 수검 대상자의 변경을 반영한 수정본과의 시간 차이가 비교적 짧았던 STEBI-A/B에서부터 25년여의 시대와 사회상 변화를 반영해서 개정된 SAI와 SAI II까지 다양한 기간에 걸친 개정 사례를 제시했다.

우선 하위 영역이나 구인의 변화가 두드러진 경우는 TOSRA 1977

판과 TOSRA 1978판, PISA2006과 PISA2015, SMQ와 SMQ II이다. TOSRA는 5개의 척도에 2개의 척도를 신설해서 기존 척도의 성격이 유지되는 수준에서 보완되었다. 과학이 주 영역이었던 2006년과 2015년 PISA에서는 동일한 내용을 물어보는 하위 항목은 5항목만 유지되었고 ‘과학 직업 기대’와 ‘과학에 대한 인식론적 신념’ 등 2항목이 신설되었다. 이때 10여년의 시대 변화를 반영한 문항 변화 사례도 발견된다. 예를 들면 ‘과학 활동’ 중 2006년의 ‘라디오 듣기’가 2015년에는 삭제되고 컴퓨터 프로그램을 통한 모의실험, 블로그 활용, 생태 기관 웹사이트 방문 등이 추가되었고, ‘과학에서 일반적 흥미’는 인체 생물학, 천문학, 물리, 화학 등으로 표현되던 내용들이 2015년에는 생물계, 우주와 우주의 역사, 운동과 힘, 에너지와 에너지 전환 등으로 표현되었다(OECD, 2007; OECD, 2016). PISA와 같은 국제 학업 성취도 분야에서 다루는 정의적 영역은 연구자의 관심, 주도적 연구 동향, 시대 상황 등에 대한 연구자들의 의사결정에 따라 달라지는 경향이 있었다. Glynn *et al.*(2011)은 SMQ를 SMQ II로 개정하면서 구인 타당도 향상을 위해 6개 구인을 5개로 줄였다고 밝혔다. 특히 SMQ의 개발자들은 2006년 이후 2011년 까지 지속적으로 구인을 수정하는 연구를 진행함으로써 검사 도구의 변화 과정을 살펴볼 수 있는 사례를 보았다(Glynn, Taasobshirazi, & Brickman, 2007, 2009). 이들에 의하면 SMQ II에서 삭제된 ‘평가 불안’ 구인은 ‘자기 효능감과 평가 불안’(Glynn, Taasobshirazi, & Brickman, 2009)이라는 과도기적 단계를 거쳐 최종 ‘자기 효능감’으로 확정되었다.

검사 도구의 점수 산출 방식을 변경한 경우는 TOSRA와 SAI다. 1977년판 TOSRA는 0점~4점 점수 부여, 2개 중 강제 선택, 5점 척도 리커트 방식 등 3가지 점수 방식을 활용했으나 1978년판에서는 리커

1) SAI II의 개정을 발표한 Moore & Foy(1997)의 연구에서는 연구가 처음 투고된 1995년을 기준으로 25년이라고 명시했다.

트 5점 척도로 일원화했다. SAI는 4점 척도 리커트 방식이었으나 Nagy(1978)의 비판을 수용하면서 SAI II에서는 중간 반응을 허용하는 5점 척도로 변경했다.

각 도구의 타당화 과정은 내용을 검토하는 과정과 통계적 방법으로 타당화하는 과정을 거친다. 예를 들어 TOSRA(Fraser, 1978)는 전문가와 교사 집단이 내용 타당도와 안면 타당도를 검토한 후 신뢰도와 변별 타당도를 검사했다. SAI(Moore & Sutman, 1970)는 전문가 집단의 검토와 더불어 특별히 학생 집단이 개발 과정에 참여했고, 구인 타당도를 증명하기 위해 서로 다른 수업 처치를 받은 3집단의 사전 사후 SAI 점수를 비교했다. 그러나 SAI II(Moore & Foy, 1997)는 높고 낮은 점수를 받은 학생들의 집단 차이로 검사 도구의 변별 타당

도를 증명하려 했다. STEBI-A/B는 내용에 대해 전문가의 검토를 받은 후, 통계적으로 확인적 요인 분석과 다른 검사 도구-SPI(the Subject Preference Inventory)(Markle, 1978)와의 상관을 비교해서 공인 타당도를 살펴보았다. 현장 평가 결과 요인 분석에 의해 2문항을 제거하기도 했다. SMQ(Glynn & Koballa, 2006)가 SMQ II(Glynn *et al.*, 2011)로 개편되는 과정에는 학생 집단과 교사 집단이 참여했고, 2009년의 연구(Glynn, Taasobshirazi, & Brickman, 2009)에서 구인 타당도 개선의 필요성이 제기되면서 탐색적 요인 분석과 확인적 요인 분석을 거쳐 타당화했다.

Table 7. Contents modified when original tools are revised

		연구 주제	측정 개념	개정 주요 사항
T O S R A	TOSRA(1977)	TOSRA(1978)	Fraser	과학 관련 태도
	<ul style="list-style-type: none"> 과학의 사회적 영향 탐구에 대한 태도 과학적 태도의 수용 과학 수업의 즐거움 과학 야의 수업에 대한 흥미 	<ul style="list-style-type: none"> 과학의 사회적 영향 과학자의 평범함 탐구에 대한 태도 과학적 태도의 수용 과학 수업의 즐거움 과학에서 취미적 관심 과학에서 직업적 관심 		
S A I	SAI(1970)	SAI II(1997)	Moore & Sutman	과학적 태도
	<ul style="list-style-type: none"> 긍정적 인지적 부정적 인지적 긍정적 정서적 부정적 정서적 (위치 진술 내용: 과학 지식의 불변성, 과학의 제한성, 객관성과 지적 정직성, 과학의 가치, 과학의 본성에 대한 대중의 이해, 과학자의 일)	<ul style="list-style-type: none"> 원도구 비판 수용: 어려운 단어, 중간 반응 신설 필요성, 문항 곤란도와 형태 위치 진술 유지하면서 태도 진술의 가독성 향상, 성차별적 언어 수정 어려운 진술 수정: 아이디어 생성 활동, 객관적인, 현상 등 4점 척도를 5점 척도로 수정, 중간 반응 신설 태도 진술 수 감소 		
P I S A	PISA(2006)	PISA(2015)	OECD	태도
	<ul style="list-style-type: none"> 과학에서 자아 효능감 과학에서 일반적 흥미 과학 관련 활동 과학의 즐거움 과학의 도구적 동기 과학의 일반적 가치 과학 탐구에 대한 지지 과학의 개인적 가치 과학에서의 자아 개념 과학 주제 학습에서 흥미 과학을 잘하는 것의 중요함 인식 과학 학습의 미래 지향 동기 	<ul style="list-style-type: none"> 과학에서 자아 효능감 과학에서 일반적 흥미 과학 활동 과학의 즐거움 과학 학습에 대한 도구적 동기 과학 직업 기대 과학에 대한 인식론적 신념 		
S T E B I	STEBI-A(1990)	STEBI-B(1990)	Riggs & Enochs	과학 교수 효능감 신념
	<ul style="list-style-type: none"> 개인 과학 교수 효능감 신념(자기 효능감 차원) 과학 교수 결과 기대(결과 기대 차원) 			
S M Q	SMQ(2006)	SMQ II(2011)	Glynn & Koballa	과학 동기
	<ul style="list-style-type: none"> 내적 동기 외적 동기 성취 지향 자기 결정 확신성(자기 효능감) 평가 불안 	<ul style="list-style-type: none"> 내적 동기 자기 결정(의지) 자기 효능감 직업 동기 성취(성적) 동기 	<ul style="list-style-type: none"> 구인 타당도 향상 필요(6개 구인 5개로 수정) 사회적 인지 이론과 탐색 요인 분석 결과에 근거 수정됨. 과학 전공과 비과학 전공 등 과학을 배우는 모든 학생 대상 확대 과학 전공, 비전공의 포커스 집단, 과학 교사에 의한 평가, 파일럿 테스트, 문헌 조사 등 활용 교사들이 포커스 집단의 응답을 읽고 원 문항에 대해 토의하고 선정함. 선정된 문항은 두 집단에게 파일럿 테스트 후 문항 선정 과학 학습 중으로만 한정해서 “대학 과학을 수강하는 중에”라는 지시문을 제시함. 짧은 시간에 끝나도록 문항 수를 줄임. 	

Table 8. Change of Items when original tools are revised

변경 내용	도구 개정될 때 변경된 문항 수							
	SAI →	SAI II	STEBI-A→	STEBI-B	SMQ →	SMQ II	PISA 2006→	PISA 2015
	(60문항)	(40문항)	(25문항)	(23문항)	(30문항)	(25문항)	(30문항)**	(9문항)
문항 수	문항 삭제	20		2		15		21
	문항 신설	·		·		10		·
	단순 교정	1		·		6		·
언어 표현 명료화	복문·중문의 단문 교체	6(3)*		·		2		·
	의미 중복된 구, 절, 수식어 제거	7		·		·		·
	간단한 문장 교체	14(3)*		·		2		·
의미 구체화 명확화	반응 제한 표현 제거	3(3)*		·		1		·
	지시문 연계 수정	·		·		2		·
	구체적 상황 제시	1		·		·		·
	의미 명확한 표현	·		·		·		1(1)*
주요 사항 교체	강조 기호 추가	·		·		·		9(6)*
	쉬운 단어·용어 교체	3(3)*		·		1		·
	주요 용어 교체	·		·		·		5(5)*
	성 차별적 언어 교체	3(2)*		·		·		·
	시제 교체	·		10		·		·
	수정 없는 문항	9***		13		1		0

* ()안의 숫자는 중복 적용된 문항 수를 의미함.

** PISA 2006의 6개 구인(자아 개념, 즐거움, 일반적 가치, 개인적 가치, 도구적 동기, 미래 지향 동기)의 PISA 2015에서의 변동만 조사함.

*** Moor & Foy(1997)는 수정 없는 문항을 10개라고 했으나, “It is useless to listen to a new idea unless everybody agrees with the idea.”를 “It is useless to listen to a new idea unless everybody agrees with it.”으로 수정했고, 이를 수정으로 취급하지 않았음. 본 연구에서는 이를 ‘단순 교정’으로 분류했으므로 수정 없는 문항을 9개로 나타냄.

나. 검사 도구의 문항 개정 사례

분석 대상인 5편의 검사 도구는 모두 리커트 4점이나 5점 척도로 구성된 문항의 형태로 구현되어 있다. 검사 도구가 개정될 경우 가시적 변화는 문항에서부터 시작된다. Table 8은 4편의 원도구가 개정될 때 나타난 문항 변화와 그에 따른 문항 수를 정리한 것이다. 1977년 판 TOSRA의 전체 문항이 공개되지 않아 문항 분석 사례에서는 제외했다. 분석 대상 원도구 모두 문항 수가 감소되었으므로 삭제된 문항은 공통적으로 나타나고, 구인이 재구조화된 SMQ II에서는 25문항 중 10문항이 신설되었다. 문항이 삭제되거나 신설되었던 이유는 하위 영역(구인)의 변화(Fraser, 1978; Glynn *et al.*, 2011; OECD, 2016), 타당화 과정에서의 통계 결과 반영(Enochs & Riggs, 1990), 간편화(Moor & Foy, 1997) 등이 있었다. TOSRA의 문항 수는 원래 5개 척도에서 한 척도 당 6~11개로 일정하지 않았지만, 7개 척도로 수정하면서 각 척도 당 10문항으로 균등 조정했다. Fraser(1978)는 개발 과정에서 문항은 더 많은 14개였지만, 전문가와 교사의 안면 타당도, 가독성, 명료성 등을 고려해서 선별했다고 밝혔다. PISA 문항은 과학의 특정 주제와 관련된 내용을 묻는 경우를 제외하고 일반적 진술로 표현된 6개 구인 30문항만을 분석 대상으로 했다.2) 분석 대상 구인

6개 영역 중 과학의 즐거움과 도구적 동기의 2개 구인만 남고 4개 구인이 삭제됨에 따라 21문항이 삭제되었다. SMQ는 ‘평가 불안’에 대한 구인이 삭제되는 등 구인의 재조정에 의해 문항 수가 달라졌다(Glynn *et al.*, 2011). 대상 변경으로 인해 문항을 수정한 STEBI-B는 타당화 과정의 요인 분석 결과 2문항을 삭제했고, SAI와 SAI II는 보다 간편하게 사용할 수 있도록 20문항을 삭제해서 문항 수를 대폭 감소시킨 경우다(Table 8).

문항을 이루는 언어 표현 측면에서는 읽고 이해하기 쉽도록 가독성 측면을 고려한 사례가 나타났다. 주로 복문이나 중문으로 이루어진 복잡한 문장을 단문으로 수정하거나, 의미가 중복된 표현을 제거하고 최대한 간단한 문항으로 교체했다. 이렇게 가독성에 있어 큰 변화는 원본과 개정본의 시간 차이가 비교적 크게 나타난 SAI와 SAI II, SMQ와 SMQ II에서 나타났다. 의미를 구체화하거나 명확하게 하는 측면에서는 반응을 제한하는 불필요한 표현 제거, 지시문과 연계한 명확한 반응을 유도, 구체적 상황 제시, 의미 명확한 표현 뿐 아니라 혼동하지 말아야 하는 용어에 강조 기호(<, >)를 추가한 경우도 나타났다(Table 8).

문항을 이루는 문장이나 의미 외에 주요하게 교체한 사항은 어려운 단어를 쉬운 단어로 교체한 경우, 주요 용어를 교체한 경우, 성 차별적 언어를 수정한 경우, 시제를 교체한 경우 등이 나타났다. 25년 만에 개정된 SAI II에서 과학자를 남성 대명사로 표현한 문항이 성을 특정하지 않는 문항으로 교체되었고, STEBI-B가 예비 교사를 대상으로 한다는 점에서 현재 시제를 미래 시제로 수정한 경우가 그 사례다. PISA 2015 문항에서는 일반 과학(broad science)과 학습으로서 과학

2) 예를 들어 과학에 대한 효능감 문항은 “식품의 겉면에 적힌 과학 정보 해석하기”와 같이 특정 주제 8개의 진술에 대한 동의 정도를 표현하는 형태로 제시되었다. 이러한 주제들은 시대에 따라 변할 수 있는 예시적 측면이 있으므로 분석 대상 문항에서 제외했다. 과학 탐구에 대한 지지, 과학에서 일반적 흥미, 과학 주제 학습에서 흥미, 과학을 잘 하는 것의 중요성, 과학 관련 활동 등도 같은 이유로 분석 대상 문항에서 제외했다.

(school science)을 혼동하지 않도록 예전에는 단순히 과학(science)으로 표현하던 용어를 수정했다.

Table 9는 Table 8에서 원도구가 개정될 때 변경되는 내용의 구체적인 문항 사례를 정리한 것이다. 정관사를 삭제하거나 문장 안에서 반복되는 단어를 대명사로 바꾸는 단순 수정에서부터 문장의 구조나 구성 요소를 부분적, 전체적으로 바꾸는 통사적 수정까지 다양한 사례가 나타났다. 문항에서 쓰이는 중문이나 복문은 의미적으로 두 개 이상의 문장이 들어있기 때문에 각각의 의미 문장에 대한 동의 수준이 다를 수 있어서 학생의 반응을 어렵게 할 수 있다. De Vellis(2017)는 강한 어조, 짧은 형태이며 모호하지 않은 선언적 진술, 전문 용어가 없는 간단한 문장 등을 좋은 문항의 특징으로 권장했다. “Science may be described as being primarily an idea-generating activity”와 같이 모호한 진술은 “Ideas are the important result of science”와 같이 전문 용어가 없는 짧고 선언적 진술로 수정되면서 학생들의 분명한 반응을 유도할 수 있다.

PISA 2015 문항에서는 2006년도 문항과 달리 ‘science’와 관련한 단어에 강조 기호(< >)를 씌워서 <broad science>, <school science> 등으로 나타냈다. 이는 ‘과학’이라는 용어를 학습으로서 배우는 과학

과 일상에서 사용하는 과학으로 서로 구별해보도록 주목하게 한다. 또한 기존에 ‘science’라고 쓰던 단어를 ‘<broad science>’로 수정해서 광의로서 과학(broad science)과 특정 과목으로서 과학(general science, integrated science, comprehensive science)을 구별하고자 했다(Song et al., 2013). 이는 연구자가 ‘과학’을 과학 그 자체, 학교 과목이나 과학 교수, 과학 쟁점, 과학자의 일, 과학 직업 등 여러 가지 의미로 쓰고 있으며, 학생은 ‘과학’의 의미를 검사 도구 개발자의 의도와 다르게 혼동하는 경우가 발견된다는 연구 결과(Krynowsky, 1988; Chung & Shin, 2017)와 관련 있다.

25년여 만에 SAI II로 개정을 시도한 Moore & Foy(1997)는 he, him, his와 같이 여성이 부재한 채 남성만 존재하는 문항들은 현 시대의 사고와 언어를 반영하지 못하므로 개정할 필요가 있다고 여겼다. 특히 모든 문항에서 과학자 혹은 과학을 하려는 사람을 남성 대명사로 나타낸 문항을 삭제하거나, 성을 특정하지 않는 문장으로 수정했다(Table 9). STEBI-A와 B의 관계는 엄밀히 개정본 보다는 수검 대상 변경을 반영해 문항을 수정한 경우에 해당한다. 수검 대상이 현직 교사에서 예비 교사로 바뀔 때, 과학 교수를 발휘할 시점에 따라 효능감의 시제를 미래로 교체한 경우다(Enochs & Riggs, 1990). 예비 교사

Table 9. Cases of item change when original tools are revised

변경 내용	수정 사례		특징	
	원도구 문항	개정된 문항		
언어 표현 명료화	단순 교정	I enjoy learning the science.	I enjoy learning science.	<ul style="list-style-type: none"> 반복 단어 대명사 교체 정관사 삭제
	복문·중문의 단문 교체	Science is so difficult that only highly trained scientists can understand it.	Only highly trained scientists can understand science.	<ul style="list-style-type: none"> 주절과 종속절에 대한 동의 수준이 다를 수 있음.
	의미 중복된 구, 절, 수식어 제거	The day after day search for scientific knowledge would become boring for me.	The search for scientific knowledge would be boring.	<ul style="list-style-type: none"> 수식어 제거 및 구체화할 의도 없는 개념에 중점 두지 않도록 개념, 용어를 간략화
	간단한 문장 교체	Science may be described as being primarily an idea-generating activity.	Ideas are the important result of science.	<ul style="list-style-type: none"> 단어를 최소화한 간단한 문장으로 교체
의미 구체화 명확화	반응 제한하는 불필요한 표현 삭제	Every citizen should understand science because we are living in an age of science.	Every citizen should understand science.	<ul style="list-style-type: none"> 반응을 제한하는 불필요한 제한어를 삭제해서 분명한 반응을 유도
	지시문 연계 수정	I believe I can master the knowledge and skills in the science course.	I believe I can master science knowledge and skills.	<ul style="list-style-type: none"> 지시문에서 상황을 제한한 문구 “When I am in a college science course,”를 제시했으므로 개별 문항에서 생략
	구체적 상황 제시	A important purpose of science is to help man to live longer.	A major purpose of science is to produce new drugs and save lives.	<ul style="list-style-type: none"> 인간 수명 연장이라는 막연한 상황에서 신약 개발, 생명 구조라는 구체적 상황 제시
	의미 명확한 표현	I will learn many things in my school science subject(s) that will help me get a job.	Many things I learn in my <school science> subject(s) will help me to get a job	<ul style="list-style-type: none"> 문장 구조가 바뀌어 의미가 명확해짐. 의미가 미묘하게 달라지면서 객관적 사고에 도움을 줌.
	강조 기호 추가	Studying my school science subject(s) is worthwhile for me because what I learn will improve my career prospects.	Studying my <school science> subject(s) is worthwhile for me because what I learn will improve my career prospects	<ul style="list-style-type: none"> school science를 <괄호>속에 넣어 broad science와 구분하도록 강조
주요 사항 교체	쉬운 단어·용어 교체	When one asks questions in science, he gets information by observing natural phenomena.	Scientific questions are answered by observing things.	<ul style="list-style-type: none"> 어려운 단어 수정, 압축적으로 제시
	주요 용어 교체	I like reading about science.	I like reading about <broad science>	<ul style="list-style-type: none"> science를 broad science 교체
	성 차별 언어 교체	If one scientist says a theory is true, all other scientists will believe him.	If one scientist says an idea is true, all other scientists will believe it.	<ul style="list-style-type: none"> 과학자를 남성으로 가정하는 성차별 언어 수정
	시제 교체	I generally teach science ineffectively.	I will generally teach science ineffectively.	<ul style="list-style-type: none"> 연구 대상이 현직 교사에서 예비 교사로 확장됨에 따라 시제를 미래형으로 교체

가 “I generally teach science ineffectively.”와 “I will generally teach science ineffectively.”에 동의하는 정도는 다를 수 있기 때문이다.

IV. 결론 및 시사점

본 연구는 1989년부터 2018년까지 30년 동안 국내 학술지에 발표된 과학 관련 정의적 영역 검사 도구 활용 현황과 이와 관련한 검사 도구의 수정 내용을 분석했다. 이로써 향후 검사 도구 개발과 개선 과정에 대한 실질적 정보를 제공하고 정의적 영역 평가와 관련한 과학 교육 연구에 시사점을 주고자 했다. 연구는 크게 두 단계로 진행되었는데, 정의적 영역 검사 도구를 중심으로 한 국내 과학 교육 연구 88편의 현황을 분석한 후, 이들이 참고한 국내외 원도구 57건을 찾아 그중 원도구가 개정되었던 사례 5건의 특징과 구체적 문항 수정 사례를 비교했다. 본 연구의 결론은 다음과 같다.

첫째, 과학과 정의적 영역 검사 도구 관련 국내 연구에서 다루는 주제는 분석 기간 초반부터 태도를 중심으로 한 연구가 현재까지 지속됨과 동시에 최근에는 효능감과 확신성, 동기, 자기/동료 평가, 정서 등으로 보다 구체화되었다. 태도라는 개념이 연구자마다 다양하게 정의되면서 다양하고 복잡한 개념을 통합해서 태도로 나타낸 측면이 있다. 사회 인지 학습 이론, 학습 동기 이론, 사회 정서 학습 이론, 성취 정서 이론 등의 연구가 소개됨에 따라 이를 과학과 정의적 영역에 반영하려는 시도가 활발히 진행되고 있고, 이를 반영한 주제로 구체화되고 있다. 특히 기존 태도 척도에 포함되어있던 하위 구인을 추출해 활용하는 연구도 나타나서 복합적 태도 척도가 보다 세부적으로 분화하고 있음을 알 수 있다.

둘째, 국내 연구에서 정의적 영역 검사 도구를 활용한 연구 목적을 조사한 결과, 연구 주제의 지속성에 따라 연구 목적의 분포 양상이, 연구 목적에 따라 원도구의 수정 범위가 다르게 나타났다. 분석 기간 초기부터 정의적 영역의 대표 주제로 자리 잡아 지속적으로 연구된 태도 연구의 경우, 개발·검증·수정 연구보다 그를 활용하는 효과 검증 및 현상 조사 연구가 더 많이 나타났다. 반면 비교적 최근에 나타난 정의적 영역 주제의 경우 아직은 개발 연구가 그를 활용한 연구보다 많은 편이다. 전자는 도구를 개발한 이후 해당 주제에 대한 효과 검증과 현상을 조사하면서 안정적으로 후속 연구로 이어지는 경우로 보인다. 후자는 연구 주제를 도입하는 초기 연구로서 검사 도구를 활용하는 후속 연구로 지속될지 여부에 따라 연구 목적의 분포 양상은 전자와 달라질 수 있다. 개발이후 그를 활용한 후속 연구가 나타나지 않는 경우, 효과 검증과 현상 조사 목적의 연구는 더 이상 증가하지 않았던 경우가 그 예에 해당한다. 한편, 검사 도구 개발을 목적으로 하는 연구는 그들이 참고한 원도구를 수정하거나 원도구의 도움 없이 자체 개발하는 경우가 많았다. 현장 조사 및 효과 검증 연구는 원도구를 그대로 이용하거나 수정하는 경우가 많아서 연구 목적에 따라 검사 도구에 들이는 노력과 활용하는 정도가 다르게 나타났다.

셋째, 검사 도구의 평가 방법 측면에서 주로 자기보고식 리커트 척도를 사용함에 따라 과학 교육 현장에서도 심리 측정학적 연구 결과에서 도출된 문제들을 타당도 저해 요인으로 인식하고 이를 해결하려는 시도를 진행하고 있었다. 주관적 자기 보고에 기반한 평가 방법에 대한 한계, 리커트 점수를 원점수로 사용할 때 측정학적 오차, 차별적 문항 기능 등을 해결하려는 노력 등이 함께 시도되면서 과학 교육

계에서도 측정학적 타당성을 확보하려는 노력을 활발히 개선하고 있다.

넷째, 국내 정의적 영역 검사 도구가 참고한 원도구를 수록한 국내외 연구들을 조사한 결과, 많은 연구자들이 선호하거나 집중해서 선택하는 검사 도구가 몇 가지 있으며, 원도구가 다루는 정의적 영역은 과학 관련 특성 뿐 아니라 타학문으로부터 유래한 범교과적 특성까지 다양하게 나타났다. TOSRA와 같이 광범위한 과학 관련 맥락을 하위 영역으로 포함하는 도구, 어느 과목이나 적용 가능한 범교과적 내용을 다루는 검사 도구는 많은 연구에서 채택되어 활용되었는데, 어떤 주제라도 내용에 맞춰 수정 가능하거나 발췌가 용이하기 때문으로 추측된다. 특히 최근 활발히 연구되는 주제는 범교과적 정의적 영역을 다루는 경우가 많은데, 타학문의 연구 주제가 과학으로 확산되는 과도기적 단계와 과학에 특화하는 과정의 한 단면을 볼 수 있다. 한편, 원도구의 출처는 학술지 논문이나 서적 및 연구 보고서를 통해 공개된 경우도 있으나, 학위 논문을 출처로 한 경우도 상당수 있었다. 연구자의 노력으로 개발된 양질의 검사 도구를 사장시키지 않으려면 출처에 관련 없이 쉽게 검색하고 활용할 수 있는 검사 도구 등록 관리 체계를 마련할 필요성이 제기된다.

다섯째, 도구가 개정되는 원인에는 연구 목적 변화, 연구자의 관점과 관심의 이동, 사회·문화적 요구 등이 있으며 개정 과정에서 하위 영역 및 구인, 점수 산출 방식, 타당화 과정, 구체적 문항 등이 수정되었다. 하위 영역이나 구인을 수정하는 경우는 시대의 흐름에 따라 연구자의 관심 영역이 달라져 하위 구인을 현저하게 바꾸는 경우, 기저 이론을 보완함에 따라 구인을 신설하거나 재조정하는 경우가 나타났고, 점수 산출 방식은 리커트 척도의 점수 조정서부터 여러 가지 점수 산출 방식을 일원화하는 방법으로 변경되었다. 개정 과정에서 타당화 과정은 더욱 강화되었는데, 통계적 방법으로는 공인 타당도나 변별 타당도, 구인 타당도와 함께 요인 분석을 활용했다. 내용 타당도는 전문가 집단 뿐 아니라 학생을 참여시키는 방법을 병행하기도 했다.

여섯째, 검사 도구의 문항을 개정하는 과정에서는 문항 수, 언어 표현 명료화, 의미 구체화 및 명확화, 주요 사항 교체 등의 사례가 나타났다. 문항 수는 하위 구인의 변화에 따라 종속되지만, 주로 간편화되거나 통계 결과를 반영해서 삭제되면서 줄어드는 경향이 있었다. 문항을 이루는 언어 측면에서 표현을 간결하게 수정해서 가독성을 증진시켰고, 의미 측면에서는 의미를 구체화하거나 명확하고 혼동하지 않도록 장치를 마련하기도 했다. 또한 쉬운 단어 사용, 성차별적 언어 수정, 시제 교체, 명확한 용어 교체 등이 수정 사례로 나타났다. 개정본과의 시간 차이가 큰 경우, 시대 변화상을 반영한 맥락이 나타나기도 했다.

이상 과학 관련 정의적 영역 검사 도구 활용 현황과 개선 사례를 분석한 결과로부터 도출한 시사점은 다음과 같다. 첫째, 정의적 특성을 측정할 때 대표적으로 사용하는 태도 검사 도구가 다루는 구인이 다소 상이한 성격을 가진 경우가 있음을 인식하고 활용 목적에 따라 이를 잘 구별해서 사용할 필요가 있다. 검사 도구 안에서 태도는 흥미, 가치, 정서, 동기 등이 혼재된 개념으로 나타나며, 이들 구인이 향하는 대상 또한 과학 그 자체, 과학 학습, 일상 과학, 과학자의 일 등으로 다양하다. 이들을 통합해서 ‘태도’로 나타내고 해석해도 활용 목적에 부합할 수 있는지 면밀하게 검토한 후 사용해야 한다. 정의적 영역을 다루는 주제가 최근 들어 구체적 특성으로 분화하는 경향이 있음을

고려할 때 향후 정의적 영역 검사 도구는 측정하려는 개념을 명확히 밝히고 사용자에게 하위 구인의 활용 범위와 제한점에 대한 정보를 제시할 필요가 있다.

둘째, 잠재 변인을 측정하는 검사 도구의 존재는 관심 있는 연구자가 향후 연구의 구체적 결과를 예상할 수 있도록 사고의 지평을 확장하는 역할을 할 수 있다. 본 연구 결과, 개발된 검사 도구는 개발에 그쳐서 사장되지 않고 다양한 연구에 활용될 수 있도록 후속 연구와 연계하는 체계를 갖출 필요가 있음을 강조했다. 개발된 검사 도구를 등록해서 검색 및 활용 가능한 자료로 구축하면, 개별 연구자가 자신의 연구 목적에 맞는 검사 도구를 채택해서 최소한의 수정으로 이를 활용할 수 있게 된다. 이는 후속 연구를 촉진하고 발전시키는 기폭제가 될 수 있다. 학술지와 연구 보고서, 출판물 뿐 아니라 학위 논문 등 다방면으로 개발되는 검사 도구를 출처에 따라 정리해서 자료를 축적하면, 개발에만 그치고 사장되는 검사 도구의 활용 가능성을 높일 수 있을 것이다.

셋째, 타학문 영역에서 개발된 검사 도구를 과학에 특화할 경우 보다 세심한 타당화 과정이 요구될 수 있다. 범교과적으로 적용할 수 있는 정의적 특성이 과학에 특화되는 과도기적 과정이 본 연구에서 보고되었다. 이 때 타학문 영역에서 측정했던 개념을 과학 관련 개념으로 대체할 때 구인이 변화되지 않는지, 내용이 잘못 인식되지 않는지 전문가의 검토와 함께 수검자의 반응을 알아보려는 노력도 필요할 것이다. 통계적 타당화 과정에서도 통계 결과에 따라 삭제되는 문항의 특성을 정리하고 다른 문항과 차별적으로 작용하는 이유를 분석한다면, 향후 문항 제작의 유용한 사례로서 활용될 수 있을 것이다.

넷째, 대부분 자기보고식 리커트 척도를 사용하는 평가 형태를 고려할 때, 주관적 자기 보고에 의존하는 과학과 정의적 영역 평가의 측정학적 타당성 확보 방안을 보다 적극적으로 모색해야 한다. 이를 위해서는 선행 연구된 측정 심리학의 제 문제들을 과학 교육의 관점에서 재조명하려는 시도를 해볼 수 있다. 주로 방법적 측면에서 통계적 접근을 통해 오차를 제거하려는 단계에서 더 나아가 과학 교육의 특수성과 사회 문화적, 심리적 원인 때문에 나타나는 오차 등, 과학 관련 정의적 특성의 내용적 측면에서 검사 도구의 타당성을 점검해볼 필요가 있다. 과학 교육 평가 현장에서 간과했던 심리·사회·문화적 이유로 인한 오차, 검사 도구 자체의 결함, 학생의 과학 특성으로 발생할 수 있는 오차의 양상과 원인에 통합적으로 접근함으로써 주의 깊게 검사 도구를 타당화하는 노력이 병행되어야 할 것이다.

마지막으로 검사 도구와 도구를 이루는 문항의 이력은 앞으로 개발될 도구와 문항 제작에서 시행착오를 줄이고 노력을 단축할 수 있는 본보기가 될 것이다. 과학과 정의적 영역 평가에 대한 연구자의 관심이 다양해지면서 앞으로 연구 목적에 맞도록 원도구를 수정해서 활용하는 경우는 더욱 증가할 것으로 예상된다. 검사 도구와 그를 이루는 문항의 변천과 역사는 이들 연구자들에게 타당한 검사 도구를 채택하고 수정할 수 있도록 방향성을 제시할 것이며, 실증 자료에 근거한 연구 풍토를 진작시키는 데 도움이 될 것이다.

국문요약

이 연구는 1989년 이후 30년간 국내 학술지에 발표된 과학 관련 정의적 영역 검사 도구 활용 현황과 이와 관련한 검사 도구의 수정

사례와 내용을 분석했다. 이로써 향후 검사 도구 개발과 개선 과정에 대한 실질적 정보를 제공하고, 타당한 검사 도구 개발을 위한 근거를 마련했다. 연구는 두 단계로 진행되었다. 정의적 영역 검사 도구를 중심으로 한 국내 과학 교육 연구 88편의 현황과 이들이 참고한 국내외 원도구 57건의 현황을 분석했다. 그 후 원도구가 개정되었던 사례 5건의 특징과 구체적 문항 수정 사례를 비교했다. 연구 결과 국내 연구 현황에서는 지속적으로 나타나는 주제와 최근에 출현하기 시작한 주제, 연구 목적과 평가 방법에 따른 연구 상황, 원도구 활용 현황 등이 분석되었다. 분석 대상 연구의 원도구 중 개정 과정이 공개된 5건의 검사 도구인 TOSRA, SAI, PISA 문항, STEBI-A/B, SMQ 등에 대해 개정 배경과 주요 사항을 분석했다. 개정 전후의 문항이 공개된 4개의 검사 문항을 비교해서 문항의 변천 과정을 분석했다. 분석 결과 문항 개선은 하위 구인의 변화에 따른 신설과 삭제, 가독성 향상, 의미를 명료화하는 방향으로 진행되고 있었다. 본 연구 결과의 시사점으로 여러 구인이 혼재된 검사 도구에서 얻은 결과 해석상 유의점, 연구의 지평을 넓혀주는 검사 도구의 가능성, 주관적 자기 보고에서 나타나는 오차 인식, 타학문에서 유래한 검사 도구를 과학에 적용시킬 때 유의할 점 등을 제안했다.

주제어 : 과학 관련 정의적 영역 특성, 검사 도구, 문항 개선, 검사 도구 개정, 연구 동향, 과학 평가

References

- Anderson, L. W., & Bourke, S. F. (2000). *Assessing affective characteristics in the schools*. New York: Routledge.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1986). *Social foundation of thoughts and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). Self-efficacy. *Harvard Mental Health Letter*, 13(9), 4.
- Britton, E. D., & Schneider, S. A. (2007). Large-scale assessments in science education. In S. K. Abell & N. G. Lederman(Eds.), *Handbook of research on science education*. Mahwa, NJ: Lawrence Erlbaum.
- Brophy, K. A. (1988). On motivating students. In D. Berliner & B. Rosenshine(Eds.), *Talks to teachers*(pp. 201-245). New York: Random House.
- Byeon, J., Kim, S., Lee, I. Kwon, Y. (2013). A study on attention type and shift of task commitment in secondary students' biological classification-eye-tracking approach. *Biology Education*, 41(2), 254-266.
- Choi, S., Kim, S. Y., & Kim, S. W. (2007). The development of instruments to assess attitudes toward science of students and their parents. *Journal of the Korean Association for Science Education*, 27(3), 272-284.
- Chun, E., Na, J., Jung, 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.
- Chung, W., Yeun, B., & Hur, M. (1994). The development of an instrument for measuring scientific attitudes of elementary school children. *Journal of the Korean Association for Science Education*, 14(3), 265-271.
- De Vellis, R. F. (2017). *Scale development: Theory and applications*(4th Edition). Los Angeles: Sage.
- Doran, R. L., Lawrenz, F., & Helgeson, S. (1994). Research on assessment in science. In D. L. Gabel(Ed.), *Handbook of research on science teaching and learning*. New York: Macmillan Publishing Company.
- Elias, M., Parker, S., Kash, M., Weissberg, R., & O'Brien, M. U. (2008). *Social and emotional learning, Moral education and character education: A comparative analysis and a view toward convergence*.

- In L. P. Nucci, T. Krettenauer, & D. Narváez(Eds.), *Handbook of moral and character education*. New York: Routledge.
- Enochs, L., & Riggs, I. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics*, 90(8), 694-706.
- Fraser, B. J. (1977). Selection and validation of attitude scales for curriculum evaluation. *Science Education*, 61(3), 317-329.
- Fraser, B. J. (1978). Development of a test of science-related attitudes. *Science Education*, 62(4), 509-515.
- Fraser, B. J. (1981). *TOSRA: Test of science related attitudes handbook*. Melbourne: The Australian Council for Educational Research.
- Glynn, S., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors, *Journal of Research in Science Teaching*, 48(10), 1159-1176.
- Glynn, S., & Koballa, T. (2006). Motivation to learn in college science. In J. J. Mintzes & W. H. Leonard(Eds.), *Handbook of college science teaching*(pp. 25-32). Arlington, VA: National Science Teachers Association Press.
- Glynn, S., Taasoobshirazi, G., & Brickman, P. (2007). Nonscience majors learning science: A theoretical model of motivation. *Journal of Research in Science Teaching*, 44, 1088-1107.
- Glynn, S., Taasoobshirazi, G., & Brickman, P., (2009). Science motivation questionnaire: Construct validation with nonscience majors. *Journal of Research in Science Teaching*, 46(2), 127-146.
- Ha, M., & Lee, J. (2013). The item response, generalizability, and structural validity for the translation of Science Motivation Questionnaire II (SMQ II). *Journal of Learner-Centered Curriculum and Instruction*, 13(5), 1-18.
- Hasan, O. E., & Billeh, V. Y. (1975). Relationship between teachers' change in attitudes toward science and some professional variables. *Journal of Research in Science Teaching*, 12(3), 255-261.
- Hong, J. (2014). The case study of self and peer evaluation for pre-service biology teachers for teaching practice. *Biology Education*, 42(3), 290-303.
- Hwang, S., Kim, H., Yoo, J., & Park, S. (2001). Ninth graders' self -assessment of scientific process skills in open investigation. *Journal of the Korean Association for Science Education*, 21(3), 506-515.
- Jeon, J. (2014). Development and construct validation of the achievement emotions questionnaire-Korean middle school Science(AEQ-KMS). *Journal of the Korean Association for Science Education*, 34(8), 745-754.
- Jeon, S. (2011). A Study on the cognition of gifted students and parents about observation and recommendation system for selection of mathematics and science gifted students(Master's thesis). Korea University, Seoul, Korea.
- Jeong, J. (2007). Development of an instrument to measure elementary and secondary school students' scientific emotions and a survey of scientific emotions in the generation of biological hypotheses. *Biology Education*, 35(1), 11-26.
- Jeong, E., & Kim, Y. (1993). Development of an instrument to assess middle school students' attitude toward environmental pollution. *Journal of the Korean Association for Science Education*, 13(2), 271-281.
- Jo, S. (2011). A study on the use of gifted behavior rating scale in teacher's recommendation for the gifted education. *The Journal of the Korean Society for the Gifted and Talented*, 10(2), 5-30.
- Jo, S., & Ku, J. (2013). An exploration in the measurement and application of attitude towards science based on social constructivism. *Journal of the Korean Association for Science Education*, 33(2), 466-477.
- Jürges, H., Schneider, K., & Büchel, F. (2005). The effect of central exit examinations on student achievement: quasi-experimental evidence from TIMSS Germany. *Journal of the European Economic Association*, 3(5), 1134-1155.
- Keller, J. (1984). The use of the ARCS model of motivation in teacher training. In K. E. Shaw(Ed.), *Aspects of educational technology XVII. Staff development and career updating*. New York: Nichols.
- Kim, S., Byeon, J., Lee, I., & Kwon, Y. (2012). An eye tracking study on test-item solving of science scholastic achievement focused in elementary school students. *Journal of Learner-Centered Curriculum and Instruction*, 12(1), 65-78.
- Kim, H., Chung, W., & Jeong, J. (1998). National assessment system development of science-related affective domain. *Journal of the Korean Association for Science Education*, 18(3), 357-369.
- Kim, M., Hong, J., & Chang, N. (2001). Development of the instrument for assessing the degree of ambivalence of the attitudes of high school students. *Biology Education*, 29(1), 27-33.
- Kim, S., Jeong, S., & Hwang, Y. (2005). Development and application of self-evaluation test items for secondary science teacher professional growth. *Journal of the Korean Association for Science Education*, 25(7), 736-745.
- Kim, S., Jung, C., & Shin, D. (2015). Research trends of science related attitude in the Korean major journals of science education. *Journal of Learner-Centered Curriculum and Instruction*, 15(12), 179-200.
- Kim, H., & Kim, C. (1999). The effects of portfolio assessment on elementary school students' science knowledge, inquiry ability and science attitudes. *Journal of the Korean Association for Science Education*, 19(1), 19-28.
- 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, K., Kim, S., Kim, M., Kim, S., Kang, M., Park, H., & Jung, S. (2009). Comparative analysis of curriculum and achievement characteristics between Korea and high performing countries in PISA & TIMSS. Seoul: Korea Institute for Curriculum and Evaluation.
- Kim, S., Kim, K., & Park, J. (2014). The effect of mathematics achievement on changes in mathematics interest and values for middle school students. *Journal of Research in Curriculum Instruction*, 18(3), 683-701.
- Kim, J., & Lee, H. (2018). Emotional response of gifted students in the class based on the evaluation of core competencies. *Journal of Learner-Centered Curriculum and Instruction*, 18(4), 225-246.
- Kim, Y., Park, Y., Park, H., Shin, D., Jung, J., & Song, S. (2014). *World of science education*. Seoul: Book's hill.
- Kim, N., & Ryu, C. (2010). Development and application of tool for measuring high school students' scientific experience. *Journal of the Korean Earth Science Society*, 31(3), 276-287.
- Kim, S., & Seo, H. (2011). Self-regulated learning ability related to science inquiry skill and affective domain of science in middle school students. *Journal of Science Education*, 35(2), 307-323.
- Ko, M., Kim, E., Heo, J., & Yang, I. (2013). Elementary school teachers' beliefs of inquiry and practice of science performance assessment. *The Korean Society of Earth Science Education*, 6(2), 124-135.
- Koballa, T. R. (1988). Attitude and related concepts in science education. *Science Education*, 72(2), 115-126.
- Krynowsky, B. (1988). Problems in assessing student attitude in science education: A partial solution. *Science Education*, 72(4), 575-584.
- Ku, J., Kim, S., Lee, H., Cho, S., & Park, H. (2016). OECD programme for international students assessment: An analysis of PISA 2015 Results. Seoul: Korea Institute for Curriculum and Evaluation.
- Kwak, H., & Shin, Y. (2014). The effects of formative assessment using mobile applications on interest and self-directedness in science instruction. *Journal of the Korean Association for Science Education*, 34(3), 285-294.
- Kwak, Y. (2017). Exploration of features of Korean eighth grade students' attitudes toward science. *Journal of the Korean Association for Science Education*, 37(1), 135-142.
- Laughton, W. H., & Wilkinson, W. J. (1968). Pupils' attitudes to science teaching. *Education in Science*, 26, 31-33.
- Lee, C. (2017). The development of a concise science achievement emotions questionnaire for secondary school students(Master's thesis). Kongju University, Kongju, Korea.
- Lee, H., Choi, K., & Nam, J. (2000). The effects of formative assessment with detailed feedback on students science achievement, attitude, and interaction between teacher and Students. *Journal of the Korean Association for Science Education*, 20(3), 479-490.
- Lee, M., & Hann, A. (2001). Effect of portfolio assessment in elementary science teaching. *Journal of Korean Elementary Science Education*, 20(1), 107-122.
- Lee, S., & Hur, M. (1998). The development of middle school students' attitude evaluation instrument in laboratory work. *Journal of Research in Curriculum & Instruction*. 2(2), 194-213.
- Lee, J., & Kim, B. (1998). Development of the psychological learning environment instrument generated by science teachers in the science instruction. *Journal of the Korean Association for Science Education*, 18(3), 313-325.
- Lee, S., & Kim, Y. (2013). A study of the development of the Korean youth moral test: standardization of the test and categorization of moral type. *Journal of ethics education studies*, 31, 237-259.
- Lee, K., & Woo, J. (1996). A study of valid measurement in science related attitudes(II): To develop an affective component of attitudes toward science scale. *Journal of the Korean Association for Science Education*, 16(2), 190-199.
- Liu, X. (2012). Developing measurement instruments for science education research. In B. J. Fraser, G. T. Kenneth, & C. J. McRobbi(Eds), *Second International Handbook of Science Education*, Dordrecht; New York: Springer.
- Markle, G. C. (1978). Assessing the validity and reliability of the subject

- preference inventory with preservice elementary teacher. *Journal of Research in Science Teaching*, 15, 519-522.
- Merrell, K., & Gueldner, B. (2010). *Social and emotional learning in the classroom: promoting mental health and academic success*. New York: Guilford publishing.
- Meyer, G. R. (1969). *A test of interest*. Milton, Queensland: Jacaranda Press.
- Midgley, C., Maehr, M., & Urdan, T. (1993). *Patterns of adaptive learning survey*. Ann Arbor, MI: University of Michigan Press.
- Ministry of Culture and Education. (1981). *Middle school education curriculum*. Notification No. 1981-442 of MOCE. Seoul: MOCE.
- Ministry of Culture and Education. (1987). *Middle school education curriculum*. Notification No. 87-9 of MOCE. Seoul: MOCE.
- Ministry of Education. (1992). *Middle school curriculum*. Notification No. 1992-11 of MOE. Seoul: MOE.
- Ministry of Education. (1997). *Middle school curriculum reference book*. Notification No. 1997-15 of MOE. Seoul: MOE.
- Ministry of Education and Human Resources Development. (2007). *Middle school curriculum*. Notification No. 2007-79 of MOEHRD. Seoul: MOEHRD.
- Ministry of Education Science, and Technology. (2011). *Science curriculum*. Notification No. 2011-361 of MOEST. Seoul: MOEST.
- Ministry of Education. (2015). *Science curriculum*. Notification No. 2015-74 of MOE. Seoul: MOE.
- Moore, R., & Sutman, F. (1970). The development, field test and validation of an inventory of scientific attitudes. *Journal of Research in Science Teaching*, 7, 85-94.
- Moore, R., & Foy, R. (1997). The scientific attitude inventory: A revision(SAI II). *Journal of Research in Science Teaching*, 34(4), 327-336.
- Mun, K., Mun, J., Shin, S., & Kim, S. (2014). Development and application of high school students' 'physics self-efficacy'. *Journal of the Korean Association for Science Education*, 34(7), 693-701.
- Munby, H. (1980). An evaluation of instruments which measure attitudes to science. In C. McFadden(Ed.), *World Trends in Science Education*. Halifax, Nova Scotia: Atlantic Institute of Education.
- Munby, H. (1983). Thirty studies involving the "Scientific Attitude Inventory": What confidence can we have in this instrument? *Journal of Research in Science Teaching*, 20, 141-162.
- Nagy, P. (1978). Subtest formation by cluster analysis of the Scientific Attitude Inventory. *Journal of Research in Science Teaching*, 15, 355-360.
- National Research Council(2002). *Scientific research in education*. Washington, DC: National Academic Press.
- OECD (2007). *PISA 2006: Science competencies for tomorrow's world*, vol 1. Paris: OECD Publishing.
- OECD (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris: OECD Publishing.
- Ormerod, M. B. (1971). The social implications' factor in attitudes to science. *British Journal of Educational Psychology*, 41, 335-338.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578.
- Park, C. (2007). The trend in the Korean middle school student's affective variables toward mathematics and its effect on their mathematics achievement. *The Mathematical Education*, 46(1). 19-31.
- Park, M., Kim, Y., & Jeon, D. (2007). The effect of factors in assessment on the science learning motivation of high achieving students. *Journal of the Korean Association for Science Education*, 27(7), 623-630.
- Park, E., Lee, S., Rachmatullah, A., & Ha, M. (2017). Validating the translated version of CARS(Changes in Attitude About the Relevance of Science), exploring variables related to CARS Scores, and constructing two equivalent test sets of CARS. *Journal of Science Education*, 41(2), 179-194.
- Park, H., & Paik, S. (2001). The effects of lessons adopting portfolio assessment regarding feedback on elementary school student's scientific knowledge, inquiry ability, and their perception. *Journal of the Korean Association for Science Education*, 21(1), 22-29.
- Park, S. (1997). Development and validation of the Korean science teaching efficacy beliefs instrument (K-STEBI) for prospective elementary school teachers. *Journal of the Korean Association for Science Education*, 16(2), 205-224.
- Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36-48.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91-106.
- Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.
- Reach, G., Zerrouki, A., Leclercq, D., & d'Ivernois, J. F. (2005). Adjusting insulin doses: From knowledge to decision. *Patient Education and Counseling*, 56(1), 98-103.
- Riggs, I., & Enochs, L. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, 74(6), 625-637.
- School Council Project for Evaluation of Science Teaching Methods(1973). *A modified version of the science pupil opinion poll*. University of Leicester.
- Schunk, D. (1991). Self-efficacy and academic motivation. *Journal Educational Psychologist*, 26(3-4), 207-231.
- Seo, J., Choi, J., & Kim, Y. (2007). Comparison of life learning skills of gifted science students and normal students in high school. *Biology Education*, 35(1), 61-72.
- Shen, C., & Pedulla, J. (2000). The relationship between students' achievement and their self-perception of competence and rigour of mathematics and science: A cross-national analysis. *Assessment in Education*, 7(2), 237-253.
- 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., Kang, H., Kwak, Y., Kim, H., Lee, S., & Lee, S. H. (2017a). A comparative analysis of the test tools in science-related affective domains. *Biology Education*, 45(1), 41-54.
- Shin, Y., Kwak, Y., Kim, H., Lee, S. Y., Lee, S. H., & Kang, H. (2017b). 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.
- So, K., Shim, K., Lee, H., & Chang, N. (2000). Study on attitude of science gifted and talented middle school students toward science. *Journal of the Korean Association for Science Education*, 20(1), 166-173.
- Song, M., Choi, H., Rim, H., Park, H., & Son, S. (2013). *OECD programme for international student assessment: Establishing PISA 2015 field trial(Research report RRE 2013-6-2)*. Seoul: Korea Institute for Curriculum and Evaluation.
- Song, Y., & Choi, H. (2017). Development and application of measurement tools for physics image using the semantic differential method. *Journal of the Korean Association for Science Education*, 37(6), 1051-1061.
- Sung, T. (2010). *Modern educational evaluation(3rd ed.)* Seoul: Hakjisa.
- Tamir, P. (1998). Assessment and evaluation in science education: Opportunities to learn and outcomes. In B. J. Fraser & K. G. Tobin(Eds.), *International handbook of science education*. Dordrecht, The Netherlands: Kluwer Academic.
- Tuan, H., Chin, C., & Shieh, S. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education*, 27(6), 639-654.
- Uguroglu, M., Schiller, D., & Walberg, H. (1981). A multidimensional motivational instrument. *Psychology in the Schools*, 18, 279-285.
- White, R. T., & Mackay, L. D. (1974). Tests of perception of scientists and self. Paper presented to the forty seventh annual meeting of the National Association for Research in Science Teaching, Chicago, April.
- Woo, H., & Chung, Y. (1994). A study on elementary school students' and middle school students' attitudes toward environmental problems. *Journal of the Korean Association for Science Education*, 14(2), 225-235.
- Woo, M., Rachmatullah, A., & Ha, M. (2017). Validating overconfidence biases measurement using genetics and evolution concept test and exploring gender and academic track differences. *Journal of Learner-Centered Curriculum and Instruction*, 17(12), 27-51.
- Yoo, S. (2000). Using portfolios as tools to implement preservice teachers' professional development and reflective thinking in early science education. *Journal of Early Childhood Education*, 20(1), 163-184.
- Yu, I., & Yoon, I. (2001). Teacher's attitude and practical application of performance assessment in the home economics education, *Korean Home Economics Education Association*, 13(1), 93-106.

저자 정보

정수임(은계중학교 교사)

신동희(이화여자대학교 교수)