

ORIGINAL ARTICLE

지구과학 수업에서 진단 및 형성평가 활용을 위한 스마트 맞춤형 평가(SPA) 시스템의 개발 및 효과

손준호¹ · 김종희^{2*}

(¹문산초등학교 · ²전남대학교)

Development and Effects of Smart Personalized Assessment(SPA) System for Using of Diagnostic and Formative Assessment in Earth Science Classes

Jun Ho Son¹ · Jonghee Kim^{2*}

(¹Munsan Elementary School · ²Chonnam National University)

ABSTRACT

The purpose of this study is to develop SPA system using diagnostic and formative assessment in earth science classes in order to discuss its effect on learning achievement and self-directed learning attitude. For this purpose, we developed total management system, app for teachers, and app for students. This research was practiced to 76 students in 5th grade. The results are as follows. Firstly, the group taking a class used by app for diagnostic and formative assessment had an effect of improving learning achievement. However, as for learning achievement long term endurance test, the group taking a class using app for diagnostic and formative assessment had no effect. Secondly, the group taking a class using apps for diagnostic and formative assessment had an effect of improving learners' self-directed learning attitude. As for self-directed long-term endurance test, the group taking a class using app for diagnostic and formative assessment had an effect. In conclusion, I hope that this SPA system might apply to the science classes as it is a system that will satisfy the needs of both teachers and students, giving much needed feedback to students.

Key words : diagnostic assessment, formative assessment, application, learning achievement, self-directed learning attitude

1. 서론

지식 정보화 사회의 가속화는 학교 현장에 새로운 교육적 요구를 증가시켰으며, 이로 인해 주요 선

진국에서는 국가나 주 단위 차원에서 체계적으로 학교 교육의 질을 관리하는 등 학교 교육의 역할을 한층 강조하고 있다(최원호와 정은영, 2008; 반재천과 신선희, 2011). 학교에서는 교육의 질적 개선과

Received 19 March, 2016; Revised 4 April, 2016; Accepted 14 April, 2016

*Corresponding author : Jonghee Kim, Chonnam National University, 77

Yongbongro Buk-gu Gwangju, 61186, Korea

Phone: +82-10-6520-5452

E-mail: earthedu@jnu.ac.kr

본 논문은 손준호의 2014년도 박사 학위논문의 내용을 발췌 정리하였음

© The Korean Society of Earth Sciences Education . All rights reserved.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License

(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

학습자의 학습 기회 확대가 중요함을 인지하여 정보통신기술을 활용함으로써 긍정적인 교육적 효과를 가져 올 것으로 기대하고 있다(김용과 손진곤, 2011). 최근에는 모바일교육 콘텐츠가 등장하여 디지털기기를 활용해 학습하는 노매딕 학습자(nomadic learner)에게 맞춤형 교육 정보를 제공할 수 있게 되었고, e-러닝의 한계였던 ‘anytime, anywhere’를 극복할 수 있는 가능성에 주목하게 되었다(김민정과 황지현, 2007). 특히 스마트 기기에서 활용할 수 있는 앱(application; app)은 학습자의 능동적인 참여를 가능하게 하여 학습에 대한 이해력과 자기주도적 학습 능력을 향상시킬 수 있게 하였다(이상홍, 2010). 이는 학습자의 능동적인 학습을 원하는 정보화 사회(Joseph *et al.*, 1999)에서 스마트 기기를 활용한 교육의 활성화는 수요자의 요구에 대한 만족과 함께 학교 교육의 질을 향상시켜 국가 경쟁력을 강화시킬 수 있는 방법일 수 있다. 교육과학기술부(2011) 역시 지식 정보화 사회에 맞는 ‘스마트교육 추진 전략’에서 온라인 평가 체제 구축과 같은 교실 혁명을 스마트교육으로 추진하겠다고 이미 밝힌 바 있다. 평가의 경우, 우리나라에서는 국가 수준의 교육과정에 그 구체적인 방향과 방법을 제시하여 국가 차원에서 교육의 질을 관리하고자 노력하고 있는데(황윤환, 2008), 이를 구현하기 위해 스마트 평가(smart assessment)의 역할을 강조하고 있다(조재춘과 임희석, 2012). 이처럼 학습에서 평가를 강조하고 있는 이유는 평가가 과학의 기본 개념에 대한 학습자의 이해를 증진시킬 수 있는 도구이자 성공적인 수업을 위해서 교사와 학습자 모두에게 중요한 교수·학습 과정의 일부이기 때문이다(황윤환, 2008; Marzano, 2003; Scherer, 2007; William, 2007). 그러므로 과학교사는 평가를 통해 학습자의 교육목표 달성 여부를 확인할 수 있기 때문에 진단 및 형성평가를 수업 시간에 활용하여 학생들이 학습목표에 정확하게 도달했는지 반드시 확인해야 한다.

교사는 진단평가(diagnostic assessment)를 통해 학습 시작 전에 학습과제와 관련해 학습자가 반드시 알아야 하는 구체적인 기능과 선수 개념의 학습 정도에 대한 출발점 행동을 확인해야 한다(김종희, 2003). 그리고 이렇게 해서 얻어진 학습자에 대한

다양한 정보를 활용해 학습자의 강·약점을 고려한 수업을 계획함으로써 교사와 학생에게 많은 도움을 줄 수 있다(이보영, 2011; Alderson, 2005; Cohen, 1993; Harrison, 1983). 형성평가(formative assessment)도 진단평가와 함께 교수·학습 과정에서 그 중요성이 날로 부각되고 있다. OECD(2005)에서는 교실에서 학습을 개선하기 위해 형성평가의 활용 방안을 제시한 바 있는데, 이처럼 형성평가를 강조하는 이유는 1년에 한 두 차례 실시하는 총괄평가 형태의 표준화 검사로는 학습자 개개인의 정확한 학습 상태를 신속·정확하게 파악할 수 없기 때문이다(Stiggins, 2005). 하지만 이러한 중요성에도 불구하고 학교 현장에서는 학생 수의 과다, 시간의 부족, 준비의 어려움, 채점과 결과 분석 및 피드백 제공에 따른 교사의 큰 부담감 때문(남정희 외, 1999; 조성수, 2003)에 진단 및 형성평가를 수업 시간에 제대로 활용하지 못하는 실정이다. 왜냐하면 과학교사가 학생들에게 진단 및 형성평가를 실시하여 실시간 피드백을 제공할 수 있도록 도움을 줄 수 있는 평가 시스템이 보급되지 않았기 때문이다. 선행 연구 결과를 보면 동형검사 문항을 활용한 진단 및 형성평가의 실시간 피드백이 반영된 평가시스템에 관한 연구는 부족한 편이다(김계형, 2013; 김은수, 2013; 김홍식과 권기태, 1999; 문일수 외, 2004; 박동우, 2011; 박성범, 2012; 유선경, 2004). 그러므로 교사에게는 평가 실시에 따른 부담을 덜어주고, 학습자에게는 맞춤형 피드백을 통해 학업성취도 및 자기주도적 학습 태도를 신장시킬 수 있는 평가시스템이 개발되어 학교 현장에서 이를 효율적으로 활용할 수 있도록 도움을 주어야 할 것이다.

따라서 이 연구에서는 지구과학 수업에서 진단 및 형성평가를 보다 효과적으로 활용하기 위해 지구과학의 교육과정 분석을 통해 평가 결과에 대한 피드백을 강조한 ‘스마트 맞춤형 평가’(smart personalized assessment; SPA) 시스템을 개발해 그 효과를 검증하고자 다음과 같이 연구문제를 설정하였다.

첫째, 지구과학 수업 시간에 진단 및 형성평가를 SPA 시스템으로 활용하면 학습자의 학업성취도에 어떠한 영향을 미치는가?

둘째, 지구과학 수업 시간에 진단 및 형성평가를 SPA 시스템으로 활용하면 학습자의 자기주도적 학습 태도에 어떠한 영향을 미치는가?

II. 연구 방법

1. 연구 절차

이 연구를 위한 전체 연구의 절차는 다음과 같다 (Fig. 1).

2. 연구 대상

이 연구는 학교장과 학년부장 및 해당 학급 담임 교사와 학부모의 동의, 무선 환경 구축 및 스마트기가 준비되어야 진행할 수 있다. 이에 연구자는 여러 학교의 협조를 구했고, 그 중 2개 학교에서 수업이 가능하다는 연락이 왔다. 이 중 무선 환경 및 스마트기에 대한 준비가 상대적으로 쉬운 학교를 선택하였고, 5학년 6개 반 중 4개 반의 담임교사가 연구에 참여하기로 하였다. 하지만 1개 반의 경우, 일부 학부모가 동의하지 않았으며 특수반 학생이 포함되어 있어 연구에서는 제외시켰다. 이렇게 하여 광주광역시 소재 S 초등학교 5학년 3개 반 76명을 최종 연구 대상으로 선정하였다.

3. 연구 설계

무작위로 실험집단과 통제집단을 선정하였고, 3월 초 학교에서 실시했던 교육청 주관 과학과 진단평가 결과를 활용해 집단 간 동질성을 검증하였다. 그리고 다음과 같이 연구 설계를 하였다(Fig. 2).

실험집단	O ₁	X ₁	O ₂	O ₃
	O ₄	X ₂	O ₅	O ₆
통제집단	O ₇		O ₈	O ₉

O₁, O₄, O₇ : 사전검사(교육청 진단평가, 자기주도적 학습 태도)

X₁ : 수업 처치(진단 및 형성평가를 앱으로 실시)

X₂ : 수업 처치(진단 및 형성평가를 학습지로 실시)

O₂, O₅, O₈ : 사후검사(학업성취도, 자기주도적 학습 태도)

O₃, O₆, O₉ : 추후검사(학업성취도 장기과제, 자기주도적 학습 태도 지속성)

Fig. 2. Experimental design

실험집단과 통제집단의 교사 변인을 최소화하기 위해 연구자가 11차시씩 33시간의 수업과 검사지를 투입한 3시간을 포함하여 36시간씩 3개 반 총 108시간을 진행하였다. 또한 1개의 실험집단에게는 SPA 시스템 사용 안내를 위해 1시간가량 별도의 교육시간을 가졌다. 1~9차시의 수업은 교육과정 분석 및 학습자의 특성을 고려하여 과학교육 전문가 1인, 지구과학교육 박사과정 2인과 석사과정 2인과 함께 재구성한 내용으로 진행하였다. 1차시의 분량은 40분이며 8~9차시는 시간표를 조정하여 연속차시로 80분간 진행하였다. 진단평가 문항의 해결 시간은 평균 5~10분, 형성평가 문항의 해결 시간은 평균 5분, 학업성취도 검사 문항의 해결 시간은 평균 40분

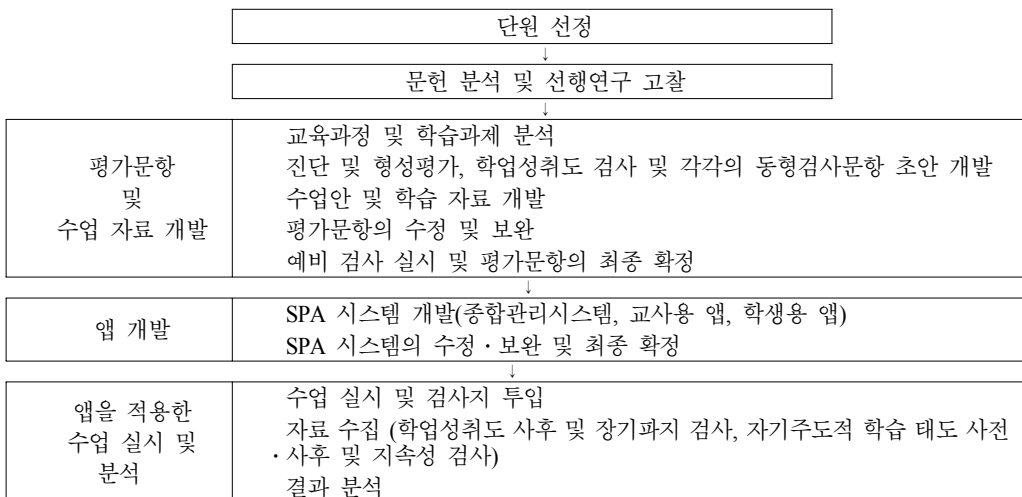


Fig. 1. Procedure of the study

내외였다. 또한 자기주도적 학습 태도 검사지의 해결 시간은 평균 20분 내외였다.

4. 평가 문항의 탑재

과학개념이 가장 많이 포함되어 있는 2007 개정 교육과정 5학년 1학기 ‘지구와 달’ 단원을 선정하여 차시별 내용에 맞는 평가 문항을 개발하였는데, 진단 및 형성평가 문항은 각각의 동형검사 문항을 각 12개씩 개발하였다. 진단평가 문항과 진단평가 동형검사 문항의 동질성을 검증하였고(Mann-Whitney $U = 237.00, p = .715$), 형성평가 문항은 학습자가 차시 수업에서 배운 내용을 수업 중에 확인할 수 있도록 단일 차시는 1~5개 문항을, 연속 차시는 6개의 문항을 개발하였다. 형성평가 문항 및 형성평가 동형검사 문항은 각 차시별 학습목표에 따른 내용을 중심으로 각 28개 문항을 개발해 그 동질성을 검증하였다($t = 3.022, p = .938$). 이렇게 개발한 평가 문항은 SPA 시스템에 탑재하였다(Fig. 3).

5. SPA 시스템의 개발

진단 및 형성평가를 실시간 활용할 수 있도록 한 SPA 시스템은 ① 종합관리시스템, ② 교사용 앱, ③

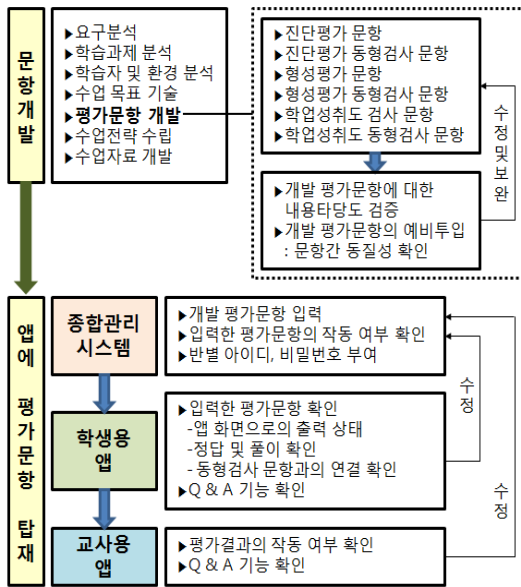


Fig. 3. Application of diagnostic and formative assessment using SPA to classes

학생용 앱으로 개발하였다.

가. 종합관리시스템

종합관리시스템은 학습자가 사용할 모든 평가문항을 교사가 직접 제작하고 관리할 수 있도록 하기 위해 웹페이지 형태로 개발한 것이다. 교사가 종합관리시스템을 활용하면 평가문항의 업로드와 문항의 수정 및 삭제가 가능하며, 학습자의 평가 결과를 한 눈에 볼 수 있다. 종합관리시스템에서는 선다형, 다답형, 주관식, 정오답형 등 다양한 형태의 문항을 제작할 수 있으며, 각 반별로 교사가 그룹을 설정하여 학습자를 관리할 수 있다. 종합관리시스템은 공지사항 관리, 시스템 로그기록, 사용자관리, 반(그



Fig. 4. Total management system



Fig. 5. App for teachers

롭)관리, 평가관리, 문제은행, 통계관리의 메뉴로 구성되어 있다(Fig. 4).

나. 교사용 앱

교사용 앱에서는 학습자들의 전반적인 경향을 확인할 수 있도록 하는 것이 주된 핵심이다(Fig. 5). 교



Fig. 6. App for students

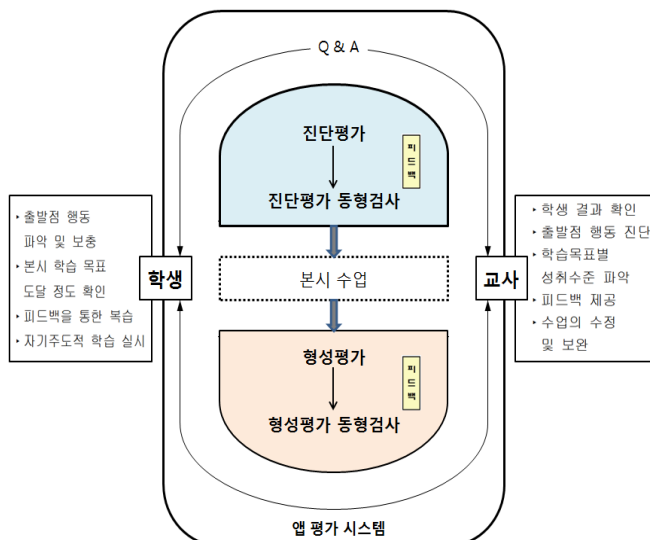


Fig. 7. Assessment system using apps in class

사용 앱을 활용하면 학습자의 개별 성적을 확인하거나 특정 문항에 대한 오답 경향을 확인할 수 있는데, ‘보기’의 메뉴를 클릭하면 학생용 앱에서 학습자가 풀고 채점된 답안지를 그대로 볼 수 있다. 또한 평가 일시, 점수, 문제풀이 상태 등 필터링을 통해 원하는 결과를 볼 수 있어 신속한 경향 파악이 가능하다.

다. 학생용 앱

학생용 앱은 진단 및 형성평가로 구성되어 있는데, 둘 다 학습자들이 문항을 풀면 자동 채점이 되고, 틀린 문항에 대해서는 피드백이 제공되어 이를 학습한다. 그런 후 제시된 동형검사 문항을 해결하면 앞의 방법과 동일하게 반복되는데 이처럼 학생용 앱의 핵심은 피드백에 있다. 여기에 ‘카카오톡’이나 ‘메시지 보내기’ 기능과 연동시켜 묻고 답하기(Q&A) 기능을 활용할 수 있도록 함으로써 학습자의 궁금증 해소에도 도움이 되도록 하였다. 진단 및 형성평가 문항을 차시별로 탑재하였는데, 학습자가 문항을 해결하면 녹색과 적색 신호등으로 정오 여부를 알려준다. 또한 틀린 문항의 경우 ‘생각해 볼까요?’라는 해설을 제공하여 자세히 읽고 스스로 학습할 수 있도록 하였다. 그리고 동형검사 문항을 해결한 후 그 결과를 서버로 전송하도록 하였고, 모든 문항은 반드시 풀어야만 다음 단계의 학습을 할 수 있도록 하였다. 학생용 앱의 전체적인 내용은 다음과 같다(Fig. 6).

6. 수업 처치

개발한 SPA 시스템에 평가문항을 탑재하여 교사용 앱과 학생용 앱을 수업 시간에 적용하였는데, 그 절차는 다음과 같다(Fig. 7).

위의 앱 평가 시스템은 진단 및 형성평가를 앱으로 활용함으로써 동형검사 문항을 통한 피드백과 교사와 학습자 간의 실시간 상호작용을 강조한 본시 수업 절차이다. 연구자는 학습자에게 최적화된 평가문항과 함께, 결과에 따른 피드백 자료를 실시간으로 제공해 주어 스스로 학습할 수 있게 하였다.

7. 검사 도구

학업성취도 검사 문항은 손준호와 김종희(2015)

가 개발한 20개 문항을 사용하였다. 학업성취도 검사 문항의 신뢰도는 Cronbach's α 가 .805, 학업성취도 동형검사 문항의 신뢰도는 Cronbach's α 가 .807이었다. 한편 자기주도적 학습 태도 검사지는 김중윤(2007)이 총 25개의 문항으로 수정하여 사용한 사전·사후 검사지를 사용하였다. 이 검사지에서 측정하고자 하는 하위변인은 학습 기회의 개방성, 학습에 대한 주도성, 학습에 대한 애착, 문제 해결 기술로써 전체 신뢰도는 Cronbach's α 는 .896이었으며, 하위변인인 학습기회의 개방성에 대한 Cronbach's α 는 .872, 학습에 대한 주도성의 Cronbach's α 는 .834, 학습에 대한 애착의 Cronbach's α 는 .849, 문제 해결 기술의 Cronbach's α 는 .899이었다.

8. 자료 처리 및 분석

이 연구에서는 SPSS 21.0 통계프로그램을 활용하여 비모수 검증을 실시하였다. Kruskal-Wallis 검증은 세 집단 간의 동질성을 확인할 때, Mann-Whitney 검증은 학업성취도와 자기주도적 학습 태도의 변화를 확인할 때, Wilcoxon 검증은 학업성취도 사후 검사와 학업성취도 장기파지 검사간의 효과를 확인할 때 사용하였다. 또한 자기주도적 학습 태도의 사전 검사, 사후검사, 지속성 검사 간의 효과를 확인하고자 반복측정 변량분석을 사용하였다.

III. 연구 결과 및 논의

1. 학업성취도에 미치는 효과 분석

연구문제 1의 검증을 위해 실험집단과 통제집단의 동질성을 확인하였는데(Table 1), Kruskal-Wallis 검증 결과 두 집단은 동질집단임을 확인하였다($p>.05$).

첫째, 진단 및 형성평가를 학습지로 실시한 수업이 학습자의 학업성취도에 미치는 효과를 확인하였는데, 검증 결과 Mann-Whitney $U = 196.50$, $p = .015$ 로 유의수준 .05에서 통계적으로 유의미한 차이가 있었다(Table 2). 그리고 진단 및 형성평가를 앱으로 실시한 수업이 학습자의 학업성취도에 미치는 효과를 확인하였는데, 검증 결과 Mann-Whitney $U = 124.50$, $p = .000$ 로 유의수준 .01에서 통계적으로 유의미한 차이가 있었다(Table 3). 또한 진단 및 형성

Table 1. Results of homogeneous group comparison for verification of research issues 1

집단	<i>N</i>	<i>M</i>	<i>SD</i>	χ^2	<i>df</i>	<i>p</i>
실험(앱)	25	71.04	13.86	.512	2	.774
실험(학습지)	25	73.60	14.41			
통제	26	71.85	15.50			

Table 2. Results of diagnostic and formative assessment using papers

집단	<i>N</i>	<i>M</i>	<i>SD</i>	Mann-Whitney <i>U</i>	<i>p</i>
실험(학습지)	25	12.24	2.49	196.50	.015*
통제	26	10.31	3.86		

* $p < .05$

Table 3. The result of diagnostic and formative assessment using an app

집단	<i>N</i>	<i>M</i>	<i>SD</i>	Mann-Whitney <i>U</i>	<i>p</i>
실험(앱)	25	14.40	1.80	124.50	.000*
통제	26	10.31	3.86		

* $p < .01$

Table 4. Results of diagnostic and formative assessment using an app and papers

집단	<i>N</i>	<i>M</i>	<i>SD</i>	Mann-Whitney <i>U</i>	<i>p</i>
실험(앱)	25	14.40	1.80	142.50	.001*
실험(학습지)	25	12.24	2.49		

* $p < .01$

평가를 앱이나 학습지로 실시한 수업이 학습자의 학업성취도에 미치는 효과를 확인하였는데, 검증 결과 Mann-Whitney $U=142.50$, $p=.001$ 로 유의수준 .01에서 통계적으로 유의미한 차이가 있었다(Table 4).

위의 결과, 진단 및 형성평가를 학습지로 실시한 집단과 앱으로 실시한 집단 모두 그렇지 않은 집단에 비해 학업성취도가 향상되었다. 또한 진단 및 형성평가를 앱으로 실시하여 수업을 한 집단은 학습지로 실시하여 수업을 한 집단에 비해 학업성취도가 더 많이 향상되었다. 전반적으로 학습자에게는 진단 및 형성평가의 실시가 학업성취도 향상에 도움을 줄 수 있는데, 이는 학습자가 진단평가를 통해 빠르고 정확한 피드백을 제공받으며, 교사는 진단평가에 근거한 체계적인 교수활동을 실시하여 효율적인 학습을 전개할 수 있기 때문이다(이보영, 2011). 이처럼 진단평가는 개념체제의 연결고리 역할을 해준다. 하지만 실제 학교 현장에서는 진단평가 문항

을 제작하는데 꽤 많은 시간이 걸리므로 교사들이 적극적으로 시간 투자를 한다는 것은 현실적으로 매우 어려운 문제이다(Alderson, 2005). 그러므로 과학교사가 진단평가를 실시하는데 있어 부담이 되지 않는 범위 내에서 효율적으로 활용할 수 있는 방법으로 SPA 시스템처럼 앱을 활용한 진단평가의 실시가 그 대안이 될 수 있다. 무엇보다도 앱을 활용하면 시·공간적인 제약 없이 학습자가 스스로 진단평가 문항을 풀 수 있고, 교사는 학습자의 시험 결과를 직접 채점할 필요 없이 자동 채점을 통해 그 결과를 원하는 목적에 맞게 선택하여 분석할 수 있다는 장점을 갖고 있기 때문이다.

형성평가 역시 학습자의 학업성취도 향상에 많은 도움을 준다. 형성평가는 학습자가 더 많은 역할에 참여하도록 하는 것이 특징인데, 이렇게 함으로써 자신의 장·단점을 어떻게 적용해야 하는지를 알게 되어 학업성취도가 향상될 수 있기 때문이다(조성수, 2003). 이처럼 형성평가는 수업 목표 달성을 위

해 학습자를 위한 피드백에 초점을 둔 평가인 것이다. 하지만 교실에서 실시되는 형성평가는 학습자가 평가문항의 내용에 대해 비판적인 토의를 해 보는 기회가 부족한데(Black, 1993), 이는 기존의 형성평가가 자신의 학습능력을 점검해 보고, 피드백을 통해 정확한 학습 개념을 형성하여 자신감을 갖도록 하는 기능보다는 학업성취도 평가의 연장선상에서 주로 실시되어 왔기 때문이다. 그래서 그동안 학습지 형태로 지속되어 왔던 형성평가는 피드백이라는 순기능보다는 순위나 성적에 반영되는 평가로서의 역기능을 많이 수행해 왔으며, 이로 인해 학습자는 형성평가에 대한 부정적인 시각을 갖게 되어 학업성취도의 향상이라는 원래의 목적을 달성하지 못한 경우가 많았다. 그러나 형성평가를 SPA 시스템처럼 앱으로 실시한 경우 학습지에 비해 학습자의 더 많은 참여와 역할을 수행할 수 있기 때문에 학업성취도 향상에 도움을 줄 수 있다. 연구자가 개발한 SPA

시스템에서 형성평가는 각 차시마다 가장 중요시되는 학습목표에 맞게 문항을 개발해 적용하였고, 각 문항마다 해설을 제공하여 학습자가 틀린 문항에 대해 읽어보고 이해할 수 있도록 구성하였다. 그리고 틀린 문항에 대해서는 동형검사 문항을 제공하여 이를 해결하도록 유도하였으며, SNS를 활용하여 실시간 질의·응답을 통해 피드백을 받는 시스템으로 구성함으로써 평가의 과정에서 학습자의 적극적인 참여를 이끌어 낼 수 있었다.

둘째, 실험집단(앱)과 실험집단(학습지) 및 통제집단 간 학업성취도 장기파지 검사를 실시하였는데, Wilcoxon 검증 결과는 다음과 같다(Table 5, Table 6).

위의 결과, 통제집단의 경우는 $Z = -.270, p = .787$ 로 사후검사와 장기파지 검사 간에 유의미한 차이가 없었다. 진단 및 형성평가를 학습지로 실시하여 수업을 한 집단의 경우도 역시 $Z = -.300, p = .764$ 로 사후검사와 장기파지 검사 간에 유의미한 차이가

Table 5. Scores of achievement test about post-test and long term test

		실험(앱)	실험(학습지)	통제
사후검사	<i>M</i>	14.40	12.24	10.31
	<i>SD</i>	1.80	2.49	3.86
장기파지 검사	<i>M</i>	11.76	12.20	10.11
	<i>SD</i>	3.32	3.68	3.85

Table 6. Results of post and long term achievement test

집단	결과	장기파지 -사후	<i>N</i>	평균 순위	순위 합	<i>Z</i>	<i>p</i>
실험(앱)		음의 순위	18	12.86	231.50	-2.853	.004*
		양의 순위	5	8.90	44.50		
		동률	2				
		합계	25				
실험(학습지)		음의 순위	12	9.42	113.00	-.300	.764
		양의 순위	8	12.13	97.00		
		동률	4				
		합계	25				
통제		음의 순위	12	12.71	152.50	-.270	.787
		양의 순위	13	13.27	172.50		
		동률	1				
		합계	26				

* $p < .05$

없었다. 하지만 진단 및 형성평가를 앱으로 실시하여 수업을 한 집단의 경우는 $Z = -2.853$, $p = .004$ 로 사후검사와 장기과제 검사 간에 유의미한 차이가 있었다. 따라서 진단 및 형성평가를 앱으로 실시하여 수업을 한 집단은 통제집단 보다는 학업성취도가 지속되지만, 진단 및 형성평가를 학습지로 실시하여 수업을 한 집단보다는 학업성취도가 지속되지 않았다. 다시 말해 진단 및 형성평가를 앱으로 실시한 경우는 단기 학업성취도 검사에서는 확실히 향상된 결과를 나타냈지만, 12주 후 장기과제 검사에는 통제집단보다는 효과가 있음에도 진단 및 형성평가를 학습지로 실시한 집단에 비해서는 학업성취도가 오랫동안 지속되지 못하였다. 이는 앱으로 학습을 할지라도 지속적인 반복학습의 노출효과가 없었거나, 접속의 번거로움으로 인해 재접속의 비율이 낮아져 복습의 기회가 적었기 때문으로 생각한다.

2. 자기주도적 학습 태도에 미치는 효과 분석

연구문제 2의 검증을 위해 실험집단과 통제집단의 동질성을 확인하였는데(Table 7), Kruskal-Wallis 검증 결과 두 집단은 동질집단임을 확인하였다 ($p > .05$).

첫째, 진단 및 형성평가를 학습지로 실시한 수업이 학습자의 자기주도적 학습 태도에 미치는 효과를 확인하였는데, Mann-Whitney 검증 결과

Mann-Whitney $U = 218.50$, $p = .045$ 로 유의수준 .05에서 통계적으로 유의미한 차이가 있었다. 특히 자기주도적 학습 태도의 하위변인인 ‘학습에 대한 주도성’에서는 유의수준 .05에서 통계적으로 유의미한 차이가 있었으나, ‘학습기회의 개방성’, ‘학습에 대한 애착’, ‘문제 해결 기술’에서는 유의수준 .05에서 통계적으로 유의미한 차이가 없었다(Table 8). 또한 진단 및 형성평가를 앱으로 실시한 수업이 학습자의 자기주도적 학습 태도에 미치는 효과를 확인하였는데, Mann-Whitney 검증 결과 Mann-Whitney $U = 89.00$, $p = .000$ 으로 유의수준 .01에서 통계적으로 유의미한 차이가 있었다. 특히 자기주도적 학습 태도의 하위영역인 ‘학습기회의 개방성’, ‘학습에 대한 주도성’, ‘학습에 대한 애착’, ‘문제 해결 기술’ 모두 유의수준 .01에서 통계적으로 유의미한 차이가 있었다(Table 9). 그리고 진단 및 형성평가를 앱이나 학습지로 실시한 수업이 학습자의 자기주도적 학습 태도에 미치는 효과를 확인하였는데, Mann-Whitney 검증 결과 Mann-Whitney $U = 206.50$, $p = .039$ 로 유의수준 .05에서 통계적으로 유의미한 차이가 있었다. 특히 자기주도적 학습 태도의 하위변인인 ‘학습기회의 개방성’, ‘학습에 대한 애착’에서는 유의수준 .05에서 통계적으로 유의미한 차이가 있었으나, ‘학습에 대한 주도성’, ‘문제 해결 기술’에서는 유의수준 .05에서 통계적으로 유의미한 차이가 없었다(Table 10).

Table 7. Results of homogeneous group comparison for verification of research issues 2

영역	집단	N	M	SD	χ^2	df	p
학습기회의 개방성	실험(앱)	25	20.56	3.39	1.66	2	.437
	실험(학습지)	25	20.20	5.31			
	통제	26	21.88	4.40			
학습에 대한 주도성	실험(앱)	25	20.68	3.44	3.88	2	.824
	실험(학습지)	25	20.96	5.18			
	통제	26	21.15	4.10			
학습에 대한 애착	실험(앱)	25	23.00	5.54	2.99	2	.224
	실험(학습지)	25	23.92	5.12			
	통제	26	25.35	5.10			
문제 해결 기술	실험(앱)	25	20.84	3.04	1.88	2	.391
	실험(학습지)	25	21.76	4.45			
	통제	26	21.85	2.88			
전체	실험(앱)	25	85.08	13.67	1.65	2	.439
	실험(학습지)	25	86.84	18.64			
	통제	26	90.23	15.14			

Table 8. Results of diagnostic and formative assessment using papers

영역	집단	<i>N</i>	<i>M</i>	<i>SD</i>	Mann-Whitney <i>U</i>	<i>p</i>
학습기회의 개방성	실험(학습지)	25	21.68	3.97	222.50	.052
	통제	26	19.31	3.94		
학습에 대한 주도성	실험(학습지)	25	21.92	4.13	219.00	.045*
	통제	26	19.19	3.84		
학습에 대한 애착	실험(학습지)	25	24.32	4.32	282.00	.415
	통제	26	23.08	3.27		
문제 해결 기술	실험(학습지)	25	22.40	3.64	225.00	.058
	통제	26	20.46	2.97		
전체	실험(학습지)	25	90.32	14.76	218.50	.045*
	통제	26	81.69	10.05		

**p* < .05

Table 9. Results of diagnostic and formative assessment using an app

영역	집단	<i>N</i>	<i>M</i>	<i>SD</i>	Mann-Whitney <i>U</i>	<i>p</i>
학습기회의 개방성	실험(앱)	25	23.68	3.00	115.50	.000*
	통제	26	19.31	3.94		
학습에 대한 주도성	실험(앱)	25	23.32	3.13	122.00	.000*
	통제	26	19.19	3.84		
학습에 대한 애착	실험(앱)	25	27.12	3.95	134.50	.000*
	통제	26	23.08	3.24		
문제 해결 기술	실험(앱)	25	23.44	3.25	151.00	.001*
	통제	26	20.46	2.97		
전체	실험(앱)	25	97.56	11.55	89.00	.000*
	통제	26	81.69	10.05		

**p* < .01

Table 10. Results of diagnostic and formative assessment using an app and papers

영역	집단	<i>N</i>	<i>M</i>	<i>SD</i>	Mann-Whitney <i>U</i>	<i>p</i>
학습기회의 개방성	실험(앱)	25	23.68	3.00	206.50	.039*
	실험(학습지)	25	21.68	3.97		
학습에 대한 주도성	실험(앱)	25	23.32	3.13	247.00	.201
	실험(학습지)	25	21.92	4.13		
학습에 대한 애착	실험(앱)	25	27.12	3.95	194.50	.022*
	실험(학습지)	25	24.32	4.32		
문제 해결 기술	실험(앱)	25	23.44	3.25	216.00	.316
	실험(학습지)	25	22.40	3.64		
전체	실험(앱)	25	97.56	11.55	206.50	.039*
	실험(학습지)	25	90.32	14.76		

**p* < .05

위의 결과, 학습지의 경우 학습자에게 흥미와 호기심을 불러일으키는 데는 한계가 있을 수 있지만, 학습지 형태로 진단평가를 실시하는 것은 학습자들이 스스로 부족한 부분을 알 수 있어 어느 정도 학습준비를 한 후 수업에 참여하도록 하는 습관 형성에 도움을 주었다고 생각한다. 그래서 형성평가를 통해 어려운 문제를 포기하지 않고 해결하려는 노력이 어느 정도 습관화되었기 때문에 자기주도적 학습 태도 향상에 도움이 되었을 것이다. 특히 학업성취도가 향상되기 위해서는 많은 조건들이 갖추어져야 하겠지만, 학습을 계획하고 시작하려는 ‘학습에 대한 주도성’의 향상이 학업성취도의 향상과 연관됨을 확인할 수 있었다.

진단 및 형성평가를 앱으로 실시한 경우 학습자의 자기주도적 학습 태도 향상에 많은 효과가 있었는데, 이는 연구문제 1의 연구 결과를 뒷받침한다고 볼 수 있다. 즉 진단 및 형성평가를 앱으로 실시하였을 경우, 학습자는 학습에 대한 높은 관심과 자기주도적인 학습 욕구를 보여 학습문제를 해결하는데 필요한 배경지식을 사용하는 능력이 향상되었는데 이러한 변화는 학업성취도의 향상이라는 결과로 이어졌다.

진단 및 형성평가를 앱으로 실시한 집단은 진단 및 형성평가를 학습지로 실시한 집단보다 자기주도적 학습 태도의 향상에 유의미한 차이가 있었는데, 자기주도적 학습 태도의 하위변인 중 ‘학습기회의

Table 11. Scores of self-directed learning attitude test about pre-test, post-test and long term test

집단	영역	학습기회의 개방성	학습에 대한 주도성	학습에 대한 애착	문제 해결 기술	전체	
사전	실험(앱)	M	20.56	20.68	23.00	20.84	85.08
		SD	3.39	3.44	5.54	3.04	13.67
	실험(학습지)	M	20.20	20.96	23.92	21.76	86.84
		SD	5.31	5.18	5.12	4.45	18.64
	통제	M	21.88	21.15	25.35	21.85	90.23
		SD	4.40	4.10	5.10	2.88	15.14
사후	실험(앱)	M	23.86	23.32	27.12	23.44	97.56
		SD	3.00	3.13	3.95	3.25	11.55
	실험(학습지)	M	21.68	21.92	24.32	22.40	90.32
		SD	3.97	4.13	4.32	3.64	14.76
	통제	M	19.31	19.19	23.08	20.46	81.69
		SD	3.94	3.84	3.24	2.97	10.05
지속성	실험(앱)	M	21.84	22.04	24.56	22.48	94.44
		SD	3.86	4.08	4.15	3.65	7.36
	실험(학습지)	M	19.20	18.64	22.80	19.64	80.28
		SD	2.78	2.32	3.69	3.18	9.77
	통제	M	19.11	19.50	22.65	20.65	81.92
		SD	3.14	3.11	3.99	3.39	12.14

Table 12. Results of self-directed learning attitude test score by repeated analysis variance between group and test period

Source	SS	df	MS	F	p
피험자간					
집단(A)	2469.609	2	1234.805	5.071	.009*
오차	17776.321	73	243.511		
피험자 내					
검사 시기(B)	845.464	2	422.732	3.504	.033**
A × B	3668.101	4	917.025	7.602	.000*
오차	17612.179	146	120.631		

* $p < .01$, ** $p < .05$

개방성'과 '학습에 대한 애착'에서 유의미하게 향상된 연구 결과는 스마트러닝이나 ICT 활용 수업이 대체적으로 자기주도적 학습 태도를 향상시킨다고 밝힌 선행연구 결과와도 일치한다(김운기, 2012; 김중윤, 2007; 이용섭, 2004). 왜냐하면 스마트러닝이나 ICT 활용 수업에서는 학습자가 수업에 직접 참여하여 교사와 학습자 간에 쌍방향 의사소통을 활용한 학습과 평가가 가능하기 때문이다. 특히 스마트러닝은 스마트 기기를 활용하여 인터넷 기반에 익숙한 학습자의 특성을 반영한 다중 방향의 학습자 중심 학습을 지향하고 있는데, 이 연구에서 개발한 SPA 시스템 역시 이러한 특성을 잘 반영하고 있다. 앱 기반의 형성평가 시스템은 개별학습을 위한 기회를 제공해 줄 수 있지만, 학습지 형태로 제공되는 형성평가 시스템은 학습자의 입장에서는 수동적이며 지루하고 개방적인 소통이 미흡하다고 볼 수 있다. 그러므로 학습지 형태의 형성평가를 학습자에게 제공한다 할지라도 학습에 흥미를 느끼지 못하거나 오히려 부담감을 주어 학습에 대한 애착을 느끼지 못하여 자기주도적 학습 태도 향상이 어려울 수 있다. 이는 진단 및 형성평가를 앱으로 실시한 집단이 학습지로 실시한 집단에 비해 학업성취도가 향상되었다고 한 연구문제 1의 결과를 뒷받침한다. 이처럼 학업성취도가 향상된 까닭을 자기주도적 학습 태도의 영역과 연관 지어 해석해 볼 수 있는데, 자기주도적 학습 태도의 하위변인 요소가 많이 향상될수록 학습자의 학업성취도 역시 더욱 향상되는 것을 확인할 수 있다.

둘째, 실험집단(앱)과 실험집단(학습지) 및 통제집단 간 자기주도적 학습 태도의 지속성 검사를 실시하여 반복측정 변량분석으로 검증한 결과는 다음과 같다(Table 11, Table 12).

위의 결과, 집단과 검사 시기의 상호작용 효과는 $F=7.602(df=4, p=.000)$ 로 유의수준 .05에서 효과가 있었다. 따라서 진단 및 형성평가를 앱으로 실시하여 수업을 한 집단은 진단 및 형성평가를 학습지로 실시하여 수업을 한 집단과 아무런 처치를 하지 않고 수업을 한 집단에 비해 자기주도적 학습 태도를 지속적으로 향상시키는 데 보다 효과적이었다. 이는 앱을 학습에 지속적으로 사용할 경우, 학습자가 활용하는 습관이 형성되어 자기주도적 학습 태도에 어느 정도 영향을 미쳤기 때문이며, 자기주도적 학습 태도의 모든 하위변인에서 다른 집단에 비

해 높은 평균값을 유지한 것 또한 이러한 해석을 뒷받침한다고 할 수 있다. 따라서 SPA 시스템을 활용하게 되면 자신의 학습을 스스로 조절함으로써 자기주도적 학습 태도의 지속성에 도움을 줄뿐만 아니라 동시에 학업성취도의 향상으로 나타날 수 있다.

IV. 결론 및 제언

이 연구는 지구과학 수업에서 진단 및 형성평가의 효율적인 활용을 위해 SPA 시스템을 개발하고, 이를 지구과학 수업 시간에 활용했을 때 학습자의 학업성취도와 자기주도적 학습 태도에 미치는 효과를 살펴보고자 하였다.

1. 결론

이 연구의 결론은 다음과 같다.

첫째, 지구과학 수업 시간에 진단 및 형성평가를 효율적으로 활용하기 위해 SPA 시스템을 개발하였다. 학습자의 출발점 행동을 파악할 수 있는 진단평가 문항과 학습목표에 맞는 구체적인 형성평가 문항 및 동형검사 문항을 개발하여 피드백을 강화하였다. SPA 시스템은 사용자의 편리성과 효율성을 강조하고자 ① 종합관리시스템, ② 교사용 앱, ③ 학생용 앱으로 구분하여 개발하였다. 종합관리시스템은 학생 등록, 공지사항, 문제은행 및 평가관리, 통계관리 등의 실행이 가능하며, 다양한 형태의 문항을 제작·수정할 수 있도록 하였다. 또한 종합관리시스템을 통해 교사는 출제 및 채점에 대한 부담에서 벗어날 수 있게 됨으로써 진단 및 형성평가의 적극적인 활용을 가능하게 하였다. 교사용 앱은 교사가 학습자의 평가 결과를 전체 혹은 개인별로 확인하여 학습자의 특성에 따른 맞춤식 지도를 할 수 있도록 하였다. 학생용 앱은 신속하고 정확한 채점 결과를 통해 학습자의 적극적인 참여를 유도하였으며, 진단 및 형성평가 문항 중 틀린 문항에 대해서는 해설과 함께 동형검사 문항을 제공하였다. 이와 더불어 SNS를 활용한 질의·응답을 통해 실시간 피드백을 제공하는 등 학습자의 적극적인 참여를 강조하였다.

둘째, 진단 및 형성평가를 SPA 시스템으로 실시한 수업이 학습자의 학업성취도와 자기주도적 학습

태도에 미치는 효과를 검증하였다. 그 결과, 진단 및 형성평가를 실시하면 학습자의 학업성취도 향상에 효과가 있었는데, 학습지보다는 앱으로 실시했을 때 학습자의 학업성취도 향상에 더 효과적이었다. 하지만 장기과제 검사에서는 앱을 실시한 집단의 평균 점수가 학습지로 실시한 집단의 평균점수 보다 낮게 나왔다.

셋째, 진단 및 형성평가를 학습지로 실시했을 때는 학습자의 ‘학습에 대한 주도성’의 하위변인에서 효과가 있었으나, 앱으로 실시했을 때는 ‘학습기회의 개방성’, ‘학습에 대한 주도성’, ‘학습에 대한 애착’, ‘문제 해결 기술’의 모든 하위변인에서 효과가 있었다. 특히 앱으로 평가를 실시한 집단은 학습지로 실시한 집단에 비해 ‘학습기회의 개방성’, ‘학습에 대한 애착’의 하위변인에서 보다 더 효과적이었다. 자기주도적 학습 태도의 지속성 검사에서도 진단 및 형성평가를 앱으로 실시했을 때 모든 하위변인에서 지속적인 효과가 나타났다. 이를 통해 진단 및 형성평가를 앱으로 실시한 수업은 전반적으로 학습자의 학업성취도와 자기주도적 학습 태도 향상에 효과가 있었으며, 자기주도적 학습 태도의 경우는 장기적으로 그 효과가 지속되었다. 따라서 이 연구에서 개발한 SPA 시스템을 수업 시간에 활용하면 학습자의 학업성취도 뿐만 아니라 자기주도적 학습 태도까지도 향상시킬 수 있으므로 과학수업에 보다 적극적으로 활용할 필요가 있다.

2. 제언

이상의 결론을 바탕으로 제언을 하면 다음과 같다.

첫째, 이 연구에서는 연구자 개인이 평가 시스템의 개발과 운영을 단독으로 진행하다 보니 많은 어려움이 있었다. 따라서 교사에게 다양한 학습자의 특성에 맞춘 평가문항 제작과 실시간 피드백이 가능한 가칭 ‘교사를 위한 DIY 평가용 시스템(앱)’ 제작에 대한 지원을 교육 지자체 및 중앙 교육부처에서 해야 할 것이다. 왜냐하면 SPA 시스템 이외에도 지역적 혹은 학교급별 특성에 따라 다양한 평가시스템을 제안할 수 있기 때문이다.

둘째, SPA 시스템을 활용한 평가시스템을 진단 및 형성평가의 활용에만 그치지 말고 수행평가, 자율학습 및 탐구 보고서 등에 활용하도록 함으로써

지속적인 학업성취도 향상 및 자기주도적 학습 태도의 하위변인 향상에 더 많은 노력을 기울여야 할 것이다.

셋째, 디지털교과서에서 맞춤형 피드백을 강화한 평가시스템이 강조되고 있는 시점에서 학계와 교육기관에서는 이에 대한 연구와 개발이 지속되어야 할 것이다. 덧붙여 이 연구에서 개발한 SPA 시스템이 디지털교과서 콘텐츠의 한 부분으로 자리매김될 수 있도록 관심을 가져야 할 것이다.

참고문헌

- 교육과학기술부(2011). 인재대국으로 가는 길 스마트교육 추진 전략 대통령 보고서.
- 김계형(2013). 초등학교 곤충의 한살이 학습을 위한 안드로이드 기반 어플리케이션의 설계 및 구현. 한국교원대학교 교육대학원 석사 학위 논문.
- 김민정, 황지현(2007). 국내외 사례를 통한 모바일 교육 콘텐츠연구-국내와 국외 사례를 중심으로. 한국디자인포럼, 15(1), 79-87.
- 김용, 손진곤(2011). 스마트폰 활용을 위한 초·중등 교육용 e-러닝 시스템 설계에 관한 연구. 인터넷정보학회논문지, 12(4), 135-143.
- 김운기(2012). 스마트교육이 자기주도학습능력에 미치는 효과. 국민대학교 교육대학원 석사 학위 논문.
- 김은수(2013). 앱 저작도구를 이용한 학교학습용 앱 개발 및 학교 시스템 연구. 인하대학교 교육대학원 석사 학위 논문.
- 김종윤(2007). e-Learning 활용 문제중심 학습이 초등학생의 자기주도적 학습태도와 문제 해결력 신장에 미치는 효과. 한국교원대학교 교육대학원 석사 학위 논문.
- 김종희(2003). 달과 행성의 위상작도 모듈을 활용한 수업이 개념변화에 미치는 효과. 부산대학교 대학원 박사 학위 논문.
- 김홍식, 권기태(1999). 멀티미디어 저작도구와 데이터베이스를 이용한 웹 기반 형성평가 방안에 관한 연구. 컴퓨터교육학회 논문지, 2(4), 157-167.
- 남정희, 성을순, 엄재호, 김경희, 최병순(1999). 형성평가에 대한 과학교사들의 인식 및 실태. 과학

- 교육논문집, 43(6), 720-727.
- 노원경(2009). 학생이 지각한 교수전략과 학습전략 및 학업성취도 간의 관계. 한국교원교육연구, 26(4), 169-190.
- 문일수, 김한일, 김철민, 박찬정(2004). 형성평가시 개별피드백을 제공하는 시스템. 컴퓨터교육학회 논문지, 7(1), 89-96.
- 박동우(2011). 스마트 앱을 활용한 보충학습 시스템 개발. 동아대학교 교육대학원 석사 학위 논문.
- 박성범(2012). 성취도 분석 기반의 반복연습형 온라인 형성평가 시스템 설계 및 구현. 한국교원대학교 교육대학원 석사 학위 논문.
- 반재천, 신선희(2011). 다수준 맥락변인이 학업성취도에 영향에 미치는 효과 추이 분석: 2003-2009 국가수준 학업성취도 평가를 중심으로. 교육평가연구, 24(3), 685-712.
- 손준호, 김종희(2015). 동형검사를 활용한 진단 및 형성평가가 초등과학 수업에 미치는 효과 : ‘지구와 달’ 단원을 중심으로. 한국과학교육학회지, 35(4), 619-628.
- 유선경(2004). 교수방법의 효율화를 위한 웹 기반 진단평가 시스템 설계 및 구현. 이화여자대학교 교육대학원 석사 학위 논문.
- 이보영(2011). 진단평가 실시 유형이 고등학교 영어 교수 활동과 학업성취도 및 학습자 인식에 미치는 영향. 이화여자대학교 교육대학원 석사 학위 논문.
- 이상홍(2010). 스마트폰과 공학교육: 스마트폰과 대학 교육 2.0. 공학교육, 17(2), 10-13.
- 이용섭(2004). 초등학교 과학과 ‘지구’ 분야의 ICT 활용 수업모듈 개발과 그 효과. 부산대학교 대학원 박사 학위 논문.
- 이정숙(2013). 교육용 앱의 평가도구 개발 연구. 아주대학교 교육대학원 석사 학위 논문.
- 조성수(2003). 형성평가에서 교정적 피드백 유형이 학업성취도와 학습태도에 미치는 효과. 서강대학교 교육대학원 석사 학위 논문.
- 조재준, 임희석(2012). 교수-학습 활동과 학습자의 특성을 고려한 스마트교육 개념모델. 컴퓨터교육학회 논문지, 15(4), 41-49.
- 최원호, 정은영(2008). 2006년 국가수준 학업성취도 평가에서 과학 성취도와 학생 관련 배경변인의 관계. 한국과학교육학회지, 28(8), 848-859.
- 황윤한(2008). 교실 수업에서의 평가와 학습의 연계. 교육과정연구, 26(4), 47-72.
- Alderson, J. C. (2005). Diagnosing Foreign Language Proficiency: The Interface Between Learning and Assessment. New York: Continuum.
- Black, P. J. (1993). Formative and Summative Assessment by Teachers, Studies in Science Education, 21(1), 49-97.
- Cohen, A. D. (1993). Assessing Language Ability in the Classroom (2nd ed.). Boston, MA: Heinle & Heinle.
- Harrison, A. (1983). A Language Testing Handbook. London: Macmillan Press.
- Joseph, G. R., Martinez, P. A., & Nancy C. (1999). Teacher Effectiveness and Learning for Mastery, Master Teachers, Mastery Learning, Group Work in Education. Journal of Educational Research, 92(5), 279.
- McMillian, J. H. (2007). Classroom Assessment; Principles and Practice for effective Instruction(4rd ed.). Boston; Pearson Education, Inc.
- OECD (2005). Formative assessment: Improving learning in secondary classroom. Paris : The Author.
- Scherer, M. (2007). An Answer for the Long Term. Educational Leadership, 65(4), 7.
- Stiggins, R. (2005). From Formative Assessment to Assessment for Learning: A Path to Success in Standards-based Schools. Phi Delta Kappan, 87(4), 34-38.
- William, D. (2007). Changing Classroom Practice. Educational Leadership, 65(4), 36-42.