

A Study on the Application and Functional Improvement of an Interactive Automatic Program Assessment System for Enhancing Computational Thinking in Elementary School Students

HakNeung Go*

*Teacher, Songjeong Jungang Elementary School, Gwangju, Korea

[Abstract]

This study investigates the effects of applying an Interactive Automatic Assessment System (IAAS) in text-based programming education for elementary school students and aims to improve the system. The results showed that the IAAS group exhibited a statistically improvement in computational thinking compared to the control group ($t=2.146$, $p=.038$). Furthermore, an analysis of computational thinking improvement across different learning levels (high, middle, and low) revealed that the experimental group demonstrated improvement in all three subgroups ($p<.05$). In contrast, the control group showed significant improvement only in the middle-level subgroup ($p=.003$), thereby confirming the potential of IAAS for individualized learning based on learning levels. Issues identified during the study's application, such as multi-cell configurations and limited recommendation features, were addressed by implementing single-cell integration and level-specific problem recommendation functions. The validity of these improvements was secured through expert validation ($N=8$, $I-CVI\geq 0.78$). This study is significant for confirming the effectiveness and individualization potential of IAAS in elementary text-based programming education, as well as for improving the system from a user-centered perspective.

▶ **Key words:** Online Judge, Interactive Automatic Assessment System, Differentiated Instruction, Programming, Computational Thinking

[요 약]

본 연구는 초등학생의 텍스트 기반 프로그래밍 학습에서 상호작용형 프로그램 자동 평가 시스템(Interactive Automatic Assessment System, IAAS)의 적용 효과를 분석하고 시스템을 개선을 목적으로 한다. 연구 결과, IAAS 활용 집단은 통제집단 대비 컴퓨팅 사고력이 유의미하게 향상되었다 ($t=2.146$, $p=.038$). 또한 학습 수준(상/중/하)에 따른 컴퓨팅 사고력 향상 정도를 분석한 결과, 실험 집단은 상위,중위,하위 집단에서 컴퓨팅 사고력이 유의미하게 향상된 반면($p<.05$), 통제집단은 중위 집단에서만 유의미한 향상을 보여($p=.003$), IAAS가 학습 수준에 따른 개별화 학습의 가능성을 확인하였다. 연구 적용 과정에서 발견된 다중 셀 구성, 제한된 추천 기능 등의 문제를 개선하여 단일 셀 통합 및 수준별 문제 추천 기능을 구현했으며, 전문가 검증($N=8$, $I-CVI\geq 0.78$)으로 타당성을 확보하였다. 본 연구는 초등 텍스트 프로그래밍 교육에서 IAAS의 효과와 개별화 가능성을 확인하고 사용자 중심의 시스템을 개선하는데 의의가 있다.

▶ **주제어:** 온라인 저지, 상호작용형 프로그램 자동 평가 시스템, 개별화 수업, 프로그래밍, 컴퓨팅 사고력

- First Author: HakNeung Go, Corresponding Author: HakNeung Go
- *HakNeung Go (snddl3@outlook.kr), Songjeong Jungang Elementary School
- Received: 2025. 04. 24, Revised: 2025. 06. 07, Accepted: 2025. 06. 07.
- This paper excerpts and summarizes parts of the first author's Ph.D. dissertation from Korea National University of Education.

I. Introduction

현대 사회에서 컴퓨팅은 단순한 기술을 넘어 문제 해결의 핵심 역량으로 자리매김하고 있다. 의료, 과학, 금융 등 다양한 분야에서 컴퓨팅 기술을 활용하여 효율성을 증대하고 혁신적인 발견을 가능하게 하였다[1]. 컴퓨팅 사고력은 문제를 이해하고 해결책을 만들어 컴퓨팅 시스템을 통해 구현되도록 하는 인지 과정으로, 컴퓨팅을 최대한 활용하기 위해 필수적인 역량이다[2]. 이에 컴퓨팅 사고력의 중요성이 강조되었 있으며 영국, 에스토니아, 인도, 핀란드 등 세계 주요 국가에서는 초등학교 저학년부터 컴퓨팅 사고력을 함양하기 위한 정보 교육을 정규 교육과정에 반영하고 있으며, 대한민국 역시 소프트웨어 교육 지침을 시작으로 컴퓨팅 사고력을 교육과정에 반영하였다[3].

컴퓨팅 사고력은 문제의 핵심을 파악하고 해결하기 위한 알고리즘을 설계하는 추상화와 알고리즘을 프로그래밍 언어를 통해 구현하고 오류를 수정하는 자동화로 구성되어 있다[4]. 프로그래밍은 이러한 추상화와 자동화를 직접적으로 적용하여 컴퓨팅 사고력을 향상시키는 것으로 연구되었다[5-6]. 프로그래밍은 논리적이고 형식적이며 오류 없이 작성해야 한다. 블록 기반 프로그래밍 언어(이하 블록 언어)는 시각적인 단서가 제공되어 사용하기 쉽고 오류 발생이 적어 배우기 쉬워 프로그래밍 입문 과정에서 많이 활용된다[7]. 반면 텍스트 기반 프로그래밍 언어(이하 텍스트 언어)는 추상적 표현으로 엄격한 구문 규칙과 복잡한 구조를 따라야 하지만 고차원적 사고를 가능하게 하고 효율적이고 복잡한 문제를 해결할 수 있다[8].

한편 텍스트 언어 입문 과정에서 나타나는 특징 중 하나는 학생들의 성취도가 정규분포가 아닌 쌍봉형 분포를 이루는 것이다. 즉, 성취도가 높은 집단과 낮은 집단이 정규분포보다 더 많이 나타난다[9-10]. 따라서 텍스트 언어 입문과정에서는 이러한 특징을 고려할 필요가 있다. 하지만 국내에서 연구된 텍스트 언어의 입문 과정 중 학습자의 특징을 고려하지 않고 사용법과 기초 문법을 익힌 후 다른 교과와의 융합 교육, 알고리즘 교육, 피지컬 교육 등에 집중하는 경향을 보였다. 이러한 학생 특성을 고려하지 않은 학습은 어려움을 겪는 학습자에게는 쉽게 포기하게 만들 수 있으며 잘하는 학생들은 흥미를 잃게 만들 수 있다. 따라서 텍스트 언어 입문 교육에서 학습자의 수준을 고려한 교육이 필요하다.

개별화 학습은 동일한 학습 목표를 달성하기 위해 학습자의 수준, 흥미, 스타일을 고려하여 학습 경로를 달리하는 것이다[11]. 프로그래밍 교육에서 개별화 학습을 지원

하는 대표적인 도구는 프로그램 자동 평가 시스템(이하 온라인 저지)이 있다. 온라인 저지는 문제에 대한 답을 코드로 작성하고 제출하면 실시간으로 평가받을 수 있는 시스템으로 학습자가 자신의 수준에 맞춰 학습을 지원한다[12]. 대부분의 온라인 저지는 웹에서 활용할 수 있으며 문제 제시, 코드 작성, 평가 결과 확인이 분리되어 있어 코드를 작성하는데 문제를 암기하고 있어야 하거나, 결과를 확인하는데 작성한 코드를 기억해야 하는 외재적 인지부하가 증가하여 본유적 인지부하가 줄어 학습이 잘 이루어지지 않을 수 있다[13]. 반면 주피터 노트북 기반 프로그램 자동 평가 시스템(이하 상호작용형 자동 평가 시스템)은 한 페이지에서 문제 제시, 코드 작성, 평가 결과를 할 수 있어 외재적 인지부하를 줄일 수 있어 본유적 인지부하가 증가해 유의미한 학습이 일어날 수 있다[14].

이에 본 연구에서는 텍스트 언어를 처음 배우는 학생들을 대상으로 상호작용형 자동 평가 시스템 활용이 컴퓨팅 사고력에 미치는 영향을 분석하고자 한다. 또한 상호작용형 자동 평가 시스템은 수준에 따라 학습 경로를 달리할 수 있으므로 개별화 학습의 가능성을 분석하고자 한다. 마지막으로 상호작용형 자동 평가 시스템을 사용한 후 후기를 바탕으로 수정 보완하고자 한다.

II. Preliminaries

1. Cognitive Development and Programming Language

Bruner는 인간의 정보 처리 및 세계 이해 방식을 작동적 표현, 영상적 표현, 상징적 표현의 세 가지 양식으로 구분하였다. 이는 아동의 발달 단계에 따라 지식을 표현하는 방식일 뿐만 아니라, 정보를 처리하고 세상을 이해하는 방식이기도 한다. 작동적 표현, 영상적 표현, 상징적 표현으로 갈수록 고차원적이며 추상적이고 생산성이 높으며, 다른 표현 양식을 통합하는 능력을 갖추게 된다. 따라서 교육과정 설계 시 인간의 정보 처리 방식 발달을 고려해야 하며, 행동적, 영상적 표현 단계를 거쳐 상징적 표현을 학습하도록 함으로써 학생들이 종합적인 맥락에서 복잡한 문제를 해결할 수 있도록 지원해야 한다[15-16].

프로그래밍 교육과 Bruner가 제시한 표현 양식을 비교하면 다음과 같다. 작동적 표현은 언플러그드 활동에 해당하며, 구체적인 활동과 경험을 통해 프로그래밍의 추상적인 개념을 이해하게 돕는다. 영상적 표현은 블록 언어에 해당하며 시각적으로 정보를 제공하는 코드를 통해 직관

적이고 접근하기 쉽게 만들어준다. 상징적 표현은 텍스트 언어로 텍스트 기반 명령어를 사용하여 더 높은 수준의 추상적 사고를 요구한다. 따라서 효과적인 프로그래밍 교육은 학습자의 인지 발달 단계를 고려하며 점진적으로 고차원적인 사고 능력을 함양할 수 있도록 프로그래밍 언어를 고려해야 한다[17-19].

2. Differentiated Instruction

학습자들은 학습 스타일, 과제 양식, 수준 등 각기 다른 특성을 지니고 있음에도 불구하고, 전통적인 수업에서는 학생들에게 동일한 학습 과제를 제공하고 동일한 교수학습 방법, 동일한 매체, 동일한 학습 속도로 학습을 제공하여 개별 학습자를 고려한 학습을 할 수 없었다. 한편, 개별화 수업은 학습자의 다양한 흥미, 선호, 학습 준비도를 고려하여 동일한 학습 목표에 도달하기 위해 내용, 과제, 학습 결과물, 정서/환경을 조정하는 학습이다[11]. 수학 교과에서 교사는 학생에게 다양한 난이도의 문제를 제공하고 학생은 수준에 맞는 문제를 선택하여 해결한 뒤 개별적인 피드백을 받는 방법의 개별화 수업을 통해 학생들의 성취도 및 수학에 대한 긍정적인 태도, 자신감, 주의 집중이 향상됨을 보고하였다[20-21].

개별화 학습을 위한 적응적 테크놀로지 모형은 데이터, 내용, 시점, 방법 4가지 차원으로 정의한다. 구체적으로 데이터는 어떤 데이터로 개별화 할지, 대상은 학습 내용, 제시 형태, 교수학습 지원 중 어떤 내용으로 개별화를 할지, 시점은 개별화 내용을 언제 제시할지, 방법은 개별화를 시스템(개발자 또는 교사), 학습자, 혼합 중 누가 통제할지에 대한 내용이다[22]. 방법 중 학습자에 의한 통제는 학습 내용에 대한 학습자의 자율성을 높여 자기조절능력을 기르고 학습 동기를 부여할 수 있지만, 학생들이 충분한 배경 지식이 부족한 경우 학습의 질이 저하될 수 있다. 교사에 의한 통제는 교사가 학습 자료, 문제, 학습 속도를 결정하는 방식으로, 학생들이 배경지식이 부족하거나 구조를 체계적으로 이해하지 못하는 경우에 적합하다. 하지만 개별화된 자료 제공에 한계가 있으며, 학생들의 다양성을 충분히 반영하는 데 어려움이 있다. 시스템에 의한 통제는 학생들의 수준과 속도를 반영하여 개별화된 학습을 제공할 수 있는 장점이 있으나, 기술 중심으로 접근하는 방식이어서 교육학적 효과의 검증이 부족하다[23].

프로그래밍 입문 과정에서 학생들의 성취도는 잘하는 집단과 어려워하는 집단이 정규분포보다 많은 쌍봉형 구조를 보이는 특징이 있어 학습자의 성취수준을 고려한 학습 방법이 필요하다. 이와 관련하여 온라인 저지는 문제

해결 후 즉각적인 피드백이 제시되어 개별화 학습을 효과적으로 지원할 수 있는 잠재력을 가지고 있다[24].

3. Interactive Automatic Assessment system

웹 기반 온라인 저지는 오랜 기간 운영되어 안정성 및 다양한 문제를 제공하는 등의 장점이 있다. 그러나 프로그래밍 입문 과정을 고려했을 때 다음과 같은 단점이 존재한다. 첫째, 문제 제시 페이지와 코드 작성 페이지가 분리되어 있어 학습자가 코드 작성시 문제를 기억하고 있어야 하며, 이는 문제 해결에 필요한 인지적 자원을 제한하는 외재적 인지부하를 야기할 수 있다. 또한 코드 작성 페이지와 평가 결과 확인 페이지가 분리되어 오답 발생 시 어디가 틀렸는지 파악하고 디버깅하기 위해 코드를 처음부터 다시 작성해야 하는 외재적 인지부하가 발생할 수 있다. 둘째, 오류 발생 시 대부분 오류 유형만을 제시한다. 이는 입문 과정 학습자에게 스스로 수정하는데 어려움을 유발할 수 있다. 셋째, 정답 또는 오답으로 평가 결과만 제시하거나 채점 데이터 별 평가 결과만을 제시한다. 이는 교사가 순회 지도시 피드백을 주는데 피드백을 주는데 어려움이 있다. 또한 코드가 어떻게 작동했는지 파악하는데 어려움이 있을 수 있다. 넷째, 프로그램 개발 도구와 온라인 저지가 분리되어 학습과 평가에 화면 전환을 요구한다[25-26].

반면 상호작용형 자동 평가 시스템은 주피터 노트북의 상호작용형 특징을 활용하여 페이지 이동 없이 한 문서 내에서 프로그램 학습과 평가를 지원한다. 또한 문제 해결시 문제 제시, 코드 작성, 평가를 페이지 전환 없이 실시하며 문자-숫자, 그래프, 표, Turtle 기반의 다양한 평가를 지원한다. 상호작용형 자동 평가 시스템은 웹 기반 온라인 저지와 비교했을 때 다음과 같은 특징을 가지고 있다. 첫째, 페이지 이동 없이 문제 제시, 코드 작성, 평가 결과를 제시하여 코드 작성 및 평가 결과를 확인할 때 외재적 인지부하를 최소화할 수 있다. 둘째, 오류 발생시 traceback 모듈에서 제공하는 피드백을 한글로 제시하며, 오류 발생 위치를 제시하여 오류를 수정할 수 있도록 한다. 셋째, 평가 결과시 채점데이터의 평가 결과와 함께 제시하여 학생이 프로그램이 어떻게 작동되는지 유추할 수 있도록 하고, 교사가 순회지도시 쉽게 확인할 수 있도록 하였다. 하지만 문제가 제한되어 있고 평가시 채점 데이터를 공개하여 학생이 정답을 유추할 수 있는 단점이 있다[14,27-28]. 상호작용형 프로그램 자동 평가 시스템은 fig1과 같다.



Fig. 1. Interactive automatic assessment system

III. The Proposed Scheme

본 연구에서는 블록 언어를 학습한 후 텍스트 언어를 처음 배우는 학생들에게 상호작용형 자동 평가 시스템이 컴퓨팅 사고력에 미치는 영향을 분석하고 학습자 수준에 따라 컴퓨팅 사고력 향상 여부를 분석하고자 한다. 또한 상호작용형 자동 평가 시스템을 적용 후 장단점을 도출하여 이를 발전시키고자 한다.

1. Research Methods

1.1 Participants and Design

연구 대상은 G초등학교 6학년에 재학중인 A학급과 B학급의 학생 각각 23명, 총 46명으로 선정하였다. A학급은 실험 집단으로 상호작용형 자동 평가 시스템을 활용하는 집단이며, B학급은 대조 집단으로 전통적인 방법으로 프로그래밍을 학습한 집단으로 설정하였다. 두 집단은 사전 학습으로

정규 교육과정을 통해 블록 언어인 Entry를 학습하였다.

상호작용형 자동 평가 시스템의 효과성과 수준에 따른 효과를 분석하기 위해 컴퓨팅 사고력 결과를 이용한 양적 검사로 설계하였다. 사전 검사로 컴퓨팅 사고력 검사와 수학 학습 전략 검사 도구를 실시하며 컴퓨팅 사고력 검사를 두 집단을 비교·분석하고 각 집단의 수준을 파악하는데 활용하였다. 수학 학습 전략 검사는 각 집단의 수준을 파악하는데 활용하였다. 이후 교육 프로그램 적용하며 사후 검사로 컴퓨팅 사고력 검사를 실시하여 두 집단을 비교 분석하였다 및 수준에 따른 발달 정도를 분석하였다. 연구 과정은 Table 1과 같다.

Table 1. Environmental Design

G1	O ₁	O ₂	X ₁	O ₃
G2	O ₄	O ₅	X ₂	O ₆

G1 : Experimental Group

G2 : Control Group

O1, O4 : Pre-test(Computational thinking)

O2, O5 : Pre-test(Math learning strategy)

O3, O6 : Post-test(Computational thinking)

X1 : Learning with Interactive Automatic Assessment system

X2 : Learning in traditional ways

1.2 Educational Program

텍스트 언어는 범용 프로그래밍 언어이면서 문법오류를 최소화할 수 있는 Python을 선정하였다. 학습 내용은 고학능과 이영준(2023)이 제시한 기초 프로그래밍 내용 요소를 바탕으로 학습자 수준을 고려하여 구성하였다. 각 차시는 블록 언어를 선행조직자로 한 수업 모형을 적용하였다. 수업 모형은 선행조직자 제시, 학습 과제 제시, 인지구조 강화 단계로 구성되어 있으며, 인지구조 강화단계에서는 학습 과제에 맞춰 프로그래밍-수학 융합 문제를 해결하도록 하였다. Table 2는 차시별 학습 내용 및 인지구조 강화 단계에서 해결할 문제 번호이다[29-30].

Table 2. Educational program

Session	Content	Advanced Organizer	Question ¹⁾
1	Output		1101,1103, 1201,1203
2	Variable, Input		2401,2402, 2403,2404
3	Type Conversion, Arithmetic Operation1		3101,3103, 3107,3108
4	Arithmetic Operation2		3106,3301, 3304,3306
5	Arithmetic Operation3	-	3313,3309, 3311,3312

1) https://github.com/GoHakNeung/jupyter_judge/tree/main/문항

6	Repeat1		5201,5202, 5203,5204
7	Repaat2		5205,5206, 5103,5104
8	Repeat3	-	5101,5102
9	Condition1		3307,3305
10	Condition2		6101,6102, 6105,6203
11	Condition3		6201,6301, 6302,6303
12	Complex Question	-	7102,7103

인지 구조 강화 단계에서 실험집단과 통제집단 모두 동일 유형의 여러 문제를 제시하여 학습 수준에 따른 개별화된 학습 경로를 가능하게 하였다. 이때 실험집단은 상호작용형 자동 평가 시스템을 활용하여 문제를 해결하였고, 통제집단은 주피터 노트북 Code cell을 이용하여 문제를 해결하였다.

1.3 Assessment Instruments

컴퓨팅 사고력 검사도구는 한국교육학술정보원(KERIS)에서 개발한 SW교육 효과성 측정을 위한 초등학생용 컴퓨팅 사고력 검사지(2017)를 활용하였다. 본 검사지는 문제풀이형이며 5지 선다형으로 총 17개 문항으로 구성되어 있다. 사후 검사에서는 문항 배치 및 보기 순서를 조정하여 응답의 무작위성을 확보하였다[31].

수학 학습 전략 검사도구는 한국과학창의재단에서 개발한 수학 클리닉 평가 중 일부 문항을 활용하였다. 본 검사지는 수학 문제 풀이시 문제 이해, 오답 검토, 문제 풀이 전략 등 다양한 학습 전략을 활용하는지 여부를 측정한다[32]. 이는 컴퓨팅 사고력의 문제 이해, 알고리즘 작성, 코드 작성, 디버깅과 관련이 있다. 또한 교수학습모형 중 인지 구조 강화 단계에서 수학 교과와 융합한 문항을 해결하므로 관련성이 있다고 판단하였다. 검사지는 자기보고식이며 5지 선다형으로 총 6문항으로 구성되어 있다. 컴퓨팅 사고력 검사와 수학 학습 전략 검사는 실험집단과 통제집단의 인지적 수준을 파악하고 집단 구분에 활용하였다.

2. Results

2.1 Reliability Verification

컴퓨팅 사고력 검사에서 학생들의 응답 일관성을 평가하기 위해 2모수 로지스틱 모델을 적용한 문항반응이론을 사용하였다. Zombo-Huber Method (Zh값)을 통해 응답

자의 응답 패턴을 분석했으며 사전 검사에서 통제집단 1명(-3.12546), 사후 검사에서 통제집단 1명(-2.80813)이 -2 이하 값으로 측정되어 경우 비정상적인 응답으로 간주하여 분석에서 제외하였다[33]. 수학 학습 전략 검사에서 실험집단과 통제집단 모두 0.8이상의 높은 신뢰도로 확인하였다. 신뢰도 분석 결과는 Table 3과 같다.

Table 3. Reliability test result of Math learning strategy

Area	Question	Cronbach's a		Number of items
Strategy	3,4,5	EG*	0.915	3
		CG**	0.880	
Process	1,2,6	EG	0.846	3
		CG	0.858	
Total		EG	0.941	6
		CG	0.929	

*EG : Experimental Group

**CG : Control Group

2.2 Analysis of Differences Between Groups

사전 검사 결과, 실험집단과 통제집단 간의 컴퓨팅 사고력 점수에 유의미한 차이가 없음을 확인하여 동질집단임을 확인하였다($t = 1.754, p = .087$). 사후 검사 결과, 실험집단과 통제집단 간의 컴퓨팅 사고력 점수에 유의미한 차이가 나타났다($t = 2.146, p = .038$). 이는 텍스트 언어 교육에서 상호작용형 자동 평가 시스템을 활용한 집단이 그렇지 않은 집단보다 컴퓨팅 사고력 향상에 더 효과적임을 확인하였다. 독립표본 t검정 결과는 Table 4와 같다.

Table 4. Result of Pre-test

	Group	M	SD	t	p
Pre	EG	9.13	3.81	1.754	.087
	CG	7.10	3.88		
Post	EG	12.70	2.85	2.146	.038*
	CG	10.71	3.27		

* $p < 0.05$

2.3 Results of Pre-test and Post-test

실험집단과 통제집단의 텍스트 언어 교육 전후 컴퓨팅 사고력 변화를 살펴보기 위해 사전-사후 검사에서 실시했던 컴퓨팅 사고력 검사 결과를 대응표본 t-검정을 실시하였다. 실험집단의 경우 평균이 9.13점에서 12.7점으로 상승하였으며, 통계적으로 유의미한 상승을 확인하였다($t=4.96, p<.001$). 통제집단의 경우 평균이 7.1점에서 10.71점으로 상승하였으며, 이 또한 통계적으로 유의미한 상승을 확인하였다($t=4.65, p <.001$). 이러한 결과는 텍스트 언어 교육이 실험집단과 통제집단 모두에서 컴퓨팅 사고력을 통계적으로 유의미하게 향상시키는 효과가 있음을 시사하였다. 대응 표본 t검정 결과는 Table 5와 같다.

Table 5. Result of Post-test

	Group	M	SD	t	p
EG	Pre	9.13	3.81	4.65	.000***
	Post	12.70	2.85		
CG	Pre	7.10	3.88	4.96	.000***
	Post	10.71	3.27		

***p<0.001

2.4 Effect Analysis by Learning Level

본 연구에서는 상호작용형 프로그램 자동 평가 시스템이 학생들의 개별화 학습에 미치는 영향을 분석하기 위해 학생들을 상위 집단, 중위 집단, 하위 집단으로 구분하여 컴퓨팅 사고력의 향상정도를 분석하였다. 이를 위해 학생들이 사전 검사로 실시한 컴퓨팅 사고력 점수(총 17점)와 수학 학습 전략 검사 결과(6점 ~ 30점)를 기준으로 구분하였으며, 추가적인 판단이 필요한 경우 사전학습으로 실시한 블록 언어를 학습했을 때 성취 정도를 고려하여 구분하였다. 집단별 인원은 Guzdial이 언급한 20% rule을 반영하여 상위 집단과 하위 집단은 각 20%인 5명이며 중위 집단은 나머지 인원으로서 설정하였으며 그 결과는 Table 6와 같다[34].

Table 6. Result of CT, Math strategy test by group

	EG		CG	
	CT	Math	CT	Math
Upper	12	24	12	23
	11	28	11	23
	15	30	12	19
	11*	26	14	25
	13	29	13	29
Middle	11	19	6	20
	7	23	9	24
	8	21	8	20
	10	18	10	17
	11	26	6	22
	8	19	6	28
	11	25	3	28
	12	23	5	27
	11	22	6	27
	13	21	9	26
	7	24	2	24
9	16	-	-	
13	21	-	-	
Lower	7	18	2	30
	4	18	2	24
	1	15	4	26
	1	21	2	18
	4	25	7**	16

*Upgrade by achievement in Entry

**Downgrade by achievement in Entry

집단별 컴퓨팅 사고력 사전-사후 검사는 표본 수가 적어 비모수 통계 분석 방법인 Wilcoxon 부호 검정을 적용하였다. 하위 집단의 경우 실험집단에서 통계적으로 유의미한 향

상이 확인되었으나(Z = 2.03, p = .042), 통제집단에서는 유의미한 향상이 없었다(Z=1.63, p = .102). 중위 집단의 경우 실험집단과 통제집단 모두 통계적으로 유의미하게 향상되었다(실험집단: Z=2.15, p=.032, 통제집단: Z=2.95, p = .003). 상위 집단의 경우 실험집단에서 유의미하게 향상되었으나(Z=2.03, p=.042), 통제집단에서는 유의미한 변화가 없었다(Z=1.36, p =.174). Wilcoxon 분석 결과는 Table 7과 같다.

Table 7. Wilcoxon sign test result by group

	Group	n	Median	SD	Z	p
Lower	EG	Pre	4	2.10	2.03	0.42*
		Post	10	2.18		
	CG	Pre	2	2.40	1.63	.102
		Post	9	2.22		
Middle	EG	Pre	11	1.67	2.15	0.32*
		Post	12	0.84		
	CG	Pre	6	2.38	2.95	.003**
		Post	11	2.15		
Upper	EG	Pre	12	2.97	2.03	0.42*
		Post	16	0.84		
	CG	Pre	12	3.56	1.36	.174
		Post	15	1.52		

*p<0.05, **p<0.01

상호작용형 자동 평가 시스템을 활용하여 학습한 집단은 모든 학습 수준에서 컴퓨팅 사고력에 통계적으로 유의미한 향상이 있음을 확인하였다. 반면 통제집단은 중위 집단에서 유의미한 향상이 있었으나 상위, 하위 집단에서는 유의미한 향상이 없었다. 이는 인지구조 강화단계에서 활용한 상호작용형 자동 평가 시스템이 상위, 중위, 하위 집단의 학습 경로를 개별화하여 학습 결과에 영향을 미친 것으로 판단된다. 즉, 상위 집단은 문제를 신속하게 해결한 후 자동 평가를 통해 결과를 확인하고 피드백을 통해 오답 및 오류를 수정하거나 다른 문제를 해결할 수 있었다. 하위 집단은 오류 발생이 많았으나, 오류 원인이 한글로 제공되고 오류 위치 또한 제공되어 이를 바탕으로 오류를 수정할 수 있었다[14]. 또한 실험집단에 속한 학생들은 자동 평가를 활용하여 결과 확인, 오답 및 오류를 피드백을 바탕으로 스스로 학습할 수 있는 환경이 조성되어 교사가 어려움을 겪는 학생들을 집중적으로 지도할 수 있었던 것으로 판단된다. 반면 통제집단의 경우 상위 집단은 문제 해결 후 평가결과를 기다리거나 확인하지 못하고 다른 문제를 해결하였다. 또한 하위 집단은 실험집단과 마찬가지로 오류 발생시 어려움을 겪었으며, 오류 발생시 traceback 모듈의 메시지만 제공되어 학생들이 이해하기 어려웠던 것으로 판단된다. 이로 인해 학생들은 교사에게 도움을 요청하였고, 교사는 제한된 시간내에 많은 학생들을 지도하기에 어려움이 있었던 것으로 판단된다.

3. Revision and Enhancement of the Interactive Automatic Assessment System

상호작용형 자동 평가 시스템을 활용하여 텍스트 언어 학습할 때 외재적 인지부하 및 오류에 대한 피드백에서 장점이 있었지만 다음과 같은 문제점을 확인하였다. 첫째, 한 문항을 구성하기 위해 Code cell을 3개를 사용한 것은 학생들이 문제의 범위를 인지하는데 어려움을 주었다. 둘째, 학생이 실수로 평가 코드 작성에 사용된 매직 커멘드(%%writefile filename.py)를 수정한 경우 파일이 저장되지 않아 평가가 실시되지 않고 교사가 코드를 수정해주어야 하는 불편함이 있었다. 셋째, 상호작용형 자동 평가 시스템은 교사가 주제와 관련된 다양한 문제를 제공하여 개별화 수업을 지원하지만 제공되는 문제의 제한성과 다양한 학습 경로 제시에 어려움이 있었다.

이러한 문제점을 개선하기 위해 하나의 Code cell 내에서 매직 커멘드 없이 문제 제시, 코드 작성, 평가 및 평가 후 문제 추천 기능이 통합하였다. 주피터 노트북에서 수정된 문제 출력 함수의 실행 과정과 문제 추천 기능은 Fig2와 같다. 문제 제시, 코드 작성, 평가는 Question() 함수로 통합했으며, 문제 번호와 유형을 매개변수로 받도록 하였다. Question() 함수는 HTML 형식으로 문제, 코드 작성 영역(코드 작성을 지원하는 확장 프로그램인 Code Mirror 활용), 제출 버튼을 출력한다. 제출 버튼 클릭 시 Code Mirror에 작성된 코드는 텍스트 파일로 저장되고, 평가 유형 매개변수에 따라 해당 평가 함수가 실행되어 그 결과가 제출 버튼 하단에 출력된다.



Fig. 2. Providing individualized problems in scratch cell

학생들에게 개별화된 문제를 제공하기 위해 주피터 노트북의 스크래치 셀을 활용하였다. 주피터 노트북의 스크래치 셀은 본문과 독립적으로 코드를 작성할 수 있으며, 본문에서 import 한 라이브러리를 활용할 수 있다. 이러한 특징을 살려 Colab 라이브러리를 통해 스크래치 셀을

생성하면서 추천하는 문제 번호를 문제 출력 함수와 함께 출력하도록 하였다. 통합 문제 출력 함수의 평가 결과 출력 시에는 '문제 추천' 버튼을 출력하며, 추천 알고리즘에 따라 스크래치 셀에 새로운 문제를 출력한다. 이 추천 알고리즘은 학생들의 프로그래밍 숙달을 목표로 동일 유형의 문제를 제시하며, 2회 연속 정답 시 한 단계 높은 난이도의 문제, 2회 연속 오답 시 한 단계 낮은 난이도의 문제를 추천한다.

수정한 통합 자동 평가 모듈의 기능성과, 학교 현장에서 사용 가능성, 학생 흥미 요구와 관련하여 타당도를 확인하기 위해 전문가 검토를 실시하였다. 내용 타당도 검토 문항은 선행연구를 토대로 컴퓨터 교육 박사수료를 한 2인과 협의하여 구성했으며 문항의 상세 내용은 Table 8와 같다.

Table 8 Validity questions

Area	Question
Functionality	The Integrated module* will function smoothly in problem presentation, code writing, and code assessment
	The Integrated module will recommend problems using Scratch cells, allowing learners to engage in further learning through them
Usability	The Integrated module will possess the features and structure to be effectively utilized in actual classroom instruction or online learning environments
	The user interface (UI) of the Integrated module will be intuitively and clearly designed so that learners can easily understand and operate it
	The Integrated module will be designed so that students can easily use it and focus on learning without separate training
	It will be easy for teachers to create learning materials using The Integrated module
Student's need	The provision of recommended problems using Scratch cells in the Integrated module will support individualized learning that considers learners' levels during class time
	The Integrated module will provide features that stimulate students' interest in programming and data analysis learning and encourage active participation
	The Integrated module will contribute to enhancing learning effectiveness by providing problems and learning materials of appropriate difficulty for learners' levels
	The Integrated module will contribute to increasing students' satisfaction with programming and data analysis learning and motivating continuous learning?

*The revised and supplemented integrated automatic assessment module

전문가 검토에 참여한 전문가는 컴퓨터교육 교수 2인, 컴퓨터 교육 박사학위 소지자 2인, 컴퓨터 교육 박사 수료 3인, 컴퓨터 교육 박사과정 1인이다. 전문가의 전공, 학위, 교육 경력은 Table 9와 같다.

Table 9. Experts involved in the review

	Job	Education	Career
A	Professor	Ph.D	6
B	Professor	Ph.D	10
C	Teacher	Ph.D	13
D	Teacher	Ph.D	6
E	Teacher	Master	5
F	Teacher	Master	15
G	Teacher	Master	13
H	Teacher	Master	8

각 문항은 ‘매우 부적절하다.’, ‘부적절하다’, ‘적절하다’, ‘매우 적절하다’의 4점 리커트 척도로 구성했으며 영역별로 전문가 의견을 제시할 수 있도록 하였다. 각 전문가들의 검토 결과를 수집하여 I-CVI(Item-Content Validity Index) 산출식을 활용하여 내용 타당도를 분석하였다. I-CVI 계산식은 다음과 같다[35].

$$I-CVI = \frac{n_r}{N}$$

N = 전체 전문가인원, n_r = ‘적절하다’, ‘매우 적절하다’를 준 응답자

전문가 검토에 참여한 인원이 8명이어서 절단값은 0.78로 설정하였다[35]. 문항별 통계 결과는 Table과 같다.

Table 10. Result of I-CVI statics analysis

Area	Item	Mean	SD	I-CVI
Functionality	1	3.63	0.48	1
	2	3.13	0.60	0.875
Usability	3	3.38	0.70	0.875
	4	3.38	0.70	0.875
	5	3.25	0.43	1
	6	3.38	0.70	0.875
	7	3.38	0.48	1
Student's need	8	3.63	0.48	1
	9	3.50	0.50	1
	10	3.50	0.50	1

수정된 통합 자동 평가 모듈에 대한 전문가 검토 결과 모든 영역에서 0.78이상이므로 통합 자동 평가 모듈의 기능성, 수업에서 사용 가능성, 학생의 요구 영역에서 내용 타당도가 갖췄다고 할 수 있다. 전문가 의견 중 문제 추천과 관련하여 수업 중 학생들이 공통문제를 해결하다가 수준에 따라 달리 문제를 해결하는 것은 개별화 및 맞춤형

학습에 기여할 것으로 의견을 주었다. 하지만 난이도에 따른 다양한 문제가 필요하나, 아직 문제가 개발되지 않은 점을 지적하였다.

IV. Conclusions

본 연구는 블록 언어를 학습한 후 텍스트 언어인 Python을 학습하는 초보 학습자를 대상으로 상호작용형 자동 평가 시스템의 효과를 검증하고 이를 개선하고자 하였다. 먼저, Python을 학습한 후 실험집단과 통제집단의 컴퓨팅 사고력 변화를 비교한 결과, 두 집단 모두 사전-사후 검사에서 컴퓨팅 사고력이 향상되었으며, 이는 블록 언어 학습 이후 Python 학습이 컴퓨팅 사고력에 긍정적인 영향을 미침을 시사하였다. 특히, 교육 프로그램 적용 후 동질집단이던 두 집단이 이질집단이 되었는데, 이는 상호작용형 프로그램 자동 평가 시스템이 학습자에게 과제 해결 후 즉각적인 평가 결과와 피드백을 제공함으로써 자기 수준에 맞는 학습 경로를 조정할 수 있도록 하였기 때문으로 판단되었다. 따라서 상호작용형 자동 평가 시스템을 활용한 Python 교육은 일반적인 Python 교육보다 컴퓨팅 사고력 향상에 효과적인 것을 시사한다.

다음으로, 사전 검사를 바탕으로 학습 수준에 따라 상위, 중위, 하위 집단으로 구분하여 컴퓨팅 사고력을 비교한 결과 실험집단은 모든 집단에서 유의미한 향상을 보인 반면, 통제집단은 중위 집단에서만 유의미한 향상을 보였다. 상호작용형 자동 평가 시스템을 활용한 상위집단은 평가 결과를 바탕으로 다음 문제를 해결하거나 오답을 수정하면서 학습하는 등 학습 경로를 스스로 조정하여 컴퓨팅 사고력 향상에 도움이 된 것으로 판단된다. 또한 하위 집단의 경우 오류 발생시 한글로 오류 원인과 그 위치를 제시하여 오류 해결에 도움이 되었고, 상호작용형 자동 평가 시스템이 스스로 학습할 수 있는 학습 환경을 조성해 교사는 도움이 필요한 학생들에게 집중하여 지도할 수 있었기 때문으로 판단된다. 이를 통해 상호작용형 자동 평가 시스템은 학습 수준에 따라 개별화 학습 가능성을 확인하였다.

본 연구에서 활용한 상호작용형 프로그램 자동 평가 시스템을 적용하는데 있어 일부 한계점을 확인하였다. 이에 교육 프로그램 적용 후, 3개의 Code cell을 단일 Code cell로 통합하는 방식으로 평가 모듈을 개선하였다. 또한 평가 결과 출력 후 ‘문제 추천’ 버튼 출력하고 이를 클릭하면 스크래치 셀에 문제를 출력하는 방식으로 문제 추천 기능을 추가하였다. 개선된 통합 자동 평가 모듈은 전문가

검토를 통해 타당성을 확보하였다.

본 연구는 몇 가지 한계점을 지니고 있었다. 첫째, 연구 참여 학생 수가 적어 연구 결과를 일반화하는 데 어려움이 있었다. 특히, 학생을 수준별로 집단을 구성할 때 표본 수가 비모수 검정의 최소 인원이라서 보다 많은 참여자를 확보하여 통계적 검정력을 높일 필요가 있었다. 둘째, 개별화 학습 효과를 평가하기 위해 학습 수준으로 집단을 구분하여 비교하였으나, 진정한 개별화 학습 평가를 위해서는 수준에 따른 집단 비교가 아닌 개별 학습 수준의 변화를 측정할 필요가 있다. 더 나아가, 개별화 학습에 영향을 미치는 요인은 학습 수준뿐만 아니라 흥미, 학습 스타일 등 다양하므로, 이러한 요인들을 자동 평가 시스템에 반영하는 방안에 대한 고려도 필요하다.

REFERENCES

- [1] Bscho, "Key Issues and Strategies in the Computing Industry," *ETRI insight*, 2019, 7, DOI: 10.22648/ETRI.2019.B.000017
- [2] Soyang, "the need for computational thinking," *Korea Information Processing Society Review*, 24(2), pp.4-12. 2017
- [3] Sckang, Sjah, Yhsung, Ysjung, Yakim, Jhseo, and Sypark, "An Empirical Data Analysis Report on the Current Status of Overseas Software Education Operation," *KERIS Issue Report*, 2019.
- [4] Yjlee, S. Hback, J. Hshin, H. Cyoo, I. Kjung, S. Jahn, J. Wchoi and S. Kjeon, "A Basic Study for the Introduction of Computational Thinking in Elementary and Secondary Education," *Korea Foundation for the Advancement of Science & Creativity*, 2014.
- [5] Hector, B. "Review on the teaching of programming and computational thinking in the world," *Frontiers in Computer Science*, Vol. 4, 2022, 4, DOI : <https://doi.org/10.3389/fcomp.2022.997222>
- [6] Skjeon, Yjlee, "The effect of CPS programming model on promoting elementary school students' divergent thinking," *The Journal of Korean Association of Computer Education*, 15(2), pp.1-8. 2012, DOI :10.32431/kace.2012.15.2.001
- [7] Žanko, Ž., Mladenovic, M. & Boljat. "Misconceptions about variables at the K-12 level," *Education and Information Technologies*, Vol. 24, 1251-1268. 2018.
- [8] Jwchoi, and Yjlee, "The analysis of Learners' difficulties in programming Learning," *The Journal of Korean Association of Computer Education*, Vol, 17(5), pp. 89-98, 2014. DOI :10.32431/kace.2014.17.5.009
- [9] Dehnadi, S., Bornat, R, "The camel has two humps ," *Middlesex University, UK*, pp.1-21, 2006
- [10] Guzdial, M., "What makes programming so hard?," <http://home.cc.gatech.edu/csl/uploads/6/Guzdial-blog-pieces-on%20what-is-CSEd.pdf>,
- [11] Yghong, "Differentiated Instruction Within Us," *Seoul Education*, Vol. 239, 2020.
- [12] Wychang and Sskim, "A review on trends of programming (algorithm) automated assessment system and its application," *The Journal of Korean Association of Computer Education*, Vol. 20(1), pp. 13-26, 2017. DOI :10.32431/kace.2017.20.1.002
- [13] Dskim, and Sjkwon, "Designing Handheld Learning Tool Based On Cognitive Load Theory," *The Journal of Korean Association of Computer Education*, Vol. 10(1), pp. 67-75, 2007. DOI :10.32431/kace.2007.10.1.007
- [14] Hkgo, Swkim, Yjlee, "Design and Implementation of a Programming Automatic Assessment System in Jupyter Notebook," *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 13(3), pp. 1080-1086, 2023. DOI: <http://dx.doi.org/10.18517/ijaseit.13.3.1>
- [15] Jelee, "Misunderstandings and Right Understandings Regarding Bruner's Modes of Representation," *Narrative and Educational Research*, Vol. 10(3), pp. 243-262, 2022. DOI :10.25051/jner.2022.10.3.010
- [16] Jelee, Hskang, "In Search of Bruner's Curriculum Theory and Its Implications for Korean Curriculum Reform," *Narrative and Educational Research*, Vol. 8(1), pp.301-324, 2020. DOI :10.25051/jner.2020.8.1.012
- [17] Lin, Y., Liao, H., Weng, S. and Dong, W., "Comparing the effects of plugged-in and unplugged activities on computational thinking development in young children," *Education and Information Technologies*, Vol.29(8), pp. 9541-9574, 2024. <https://doi.org/10.1007/s10639-023-12181-x>
- [18] Griebel, E., Fein, B., Obermüller, F., Fraser, G. and Just, R., "On the applicability of language models to block-based programs," pp. 2374-2386, Melbourne, Australia, March 2023. DOI: <https://doi.org/10.1109/ICSE48619.2023.00213>
- [19] Harper, K. A., Freuler, R. J., Demel, J. T. and Brand, S. H., "Continuing the Comparison Between Graphical-and Text-based Programming Instruction," 2013.
- [20] Shjeong, Gyson, "A Study on the Effectiveness of In-Class Step-By-Step Individualization Programs and feedback for Elementary School Mathematics Classes by Level," *Journal of Elementary Education*, Vol. 38(1), 2022. DOI :10.23103/dnueje.2022.38.1.147
- [21] Stjung, and Jspark, "A Study of the Effect on Mathematical Learning Achievement and Disposition in a Class Using Level Orientated Teaching Method," *The Journal of Elementary Mathematics Education of Korea*, Vol, 13(2). pp. 193-210, 2009.
- [22] Vandewaetere, M. and Clarebout, G., "Advanced technologies for personalized learning, instruction, and performance,"

Handbook of research on educational communications and technology, pp, 425-437, 2014. DOI https://doi.org/10.1007/978-1-4614-3185-5_34

- [23] UNESCO, "Artificial Intelligence and Education: Guidance for Policy-makers" 2023.
- [24] Sjkim, Jhpark, Sbjeon, and Dhseo, "Web-based online judge system for online programming education," pp, 1-3, Las Vegas, NV, USA, January 2022. DOI: <https://doi.org/10.1109/ICCE53279.2022.9730318>
- [25] Bjchoi, Baekjoon Online Judge, <https://help.acmicpc.net/judge/rte>
- [26] GSHS Informatics Teachers. KOISTUDY. <http://koistudy.net/>
- [27] Hngo, Hwmoon, Yjlee, Swkim, "Design and Implementation of a Data Preprocessing Automatic Assessment Module in Jupyter Notebook," International Journal on Advanced Science, Engineering & Information Technology, Vol. 14(6), 2024. DOI : 10.18517/ijaseit.14.6.20233
- [28] F. S. Zamfir and E. Pricop, "On the design of an interactive automatic Python programming skills assessment system," 2022 14th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Ploiesti, Romania, pp. 1-5, 2022, DOI: 10.1109/ECAI54874.2022.9847414.
- [29] Hngo, Yjlee, "Development and application of Python teaching · learning strategies with block programming language as a advanced organizer," Journal of Learner-Centered Curriculum and Instruction Vol. 23(6) 2023, <https://doi.org/10.22251/jlcci.2023.23.6.317>
- [30] Hngo, Jhjeon, Yjlee, "A Study on the Development of Problem Bank for Programming-Math Convergence Education in Programming Automatic Assessment System," Journal of The Korean Association of Information Education, Vol. 27(2), pp. 141-152, 2023. DOI : <http://dx.doi.org/10.14352/jkaie.2023.27.2.141>
- [31] Wglee, Jmkim, Jhseo, "A Study on the Development of Measurement Tools for the Effectiveness of Software (SW) Education in 2017," KERIS, 2017
- [32] Askmath, Kofac, <https://askmath.kosac.re.kr/main.do>
- [33] Paek, I. and Cole, K., "Using R for item response theory model-134 applications," Routledge, 2019.
- [34] Guzdial, M., "Why is it so hard to learn to program," Making Software: What Really Works, and Why We Believe It, O'Reilly Media, 111-124, 2010.
- [35] Tjseong, "Understanding and Application of Modern Basic Statistics," Hakgisa, 2019.

Authors



HakNeung Go received the B.S. degree in Science education in Gwangju National University of Education, Korea in 2015. He completed Ph.D. degree in Computer education from Korea National University of

Education, Korea, in 2023. He is currently a Teacher in the Gadeuk Elementary School, Korea. His research interests include programming, informatics education, online judge, AI education, Data Visualization.