

☒ 연구논문

교육용 소프트웨어 품질평가를 위한 평가지의 개발
- CD-ROM 타이틀과 인터넷 상의 소프트웨어를 중심으로 -

차운옥

한성대학교 정보전산학부

김명숙

COMKID 교육연구소

Development of the Evaluation Checklists for Quality of
Educational Softwares.

- CD-ROM titles and softwares on the internet -

Woon Ock Cha

Division of Information and Computer Engineering, Hansung University

Myung Sook Kim

COMKID Education Research Institute

Abstract

We have developed three evaluation checklists of educational softwares for users such as, learners, teachers and parents, not for the software developers or designers. This study has especially focused on CD-ROM titles and softwares on the internet.

The followings were selected as the essential attributes that the educational softwares of high quality should possess: ease of use, technical support and interest of learning for learners, educational value and smartness for teachers, packaging integrity, investment value and childproof for parents. The evaluation checklists were developed on the basis of evaluation criteria which were concretion of the essential attributes.

The validity of the evaluation checklists was supported by a theoretical study for the development of checklists and the actual development process. The interrater reliability and the retest reliability were confirmed by statistical methods. Thus, the developed checklists can be a useful tool to evaluate educational

softwares. This study will contribute to improvement of the software quality and help users in selecting educational softwares.

1. 서론

컴퓨터의 발달은 사회 전반에 많은 영향을 미치고 있다. 우리 나라의 교육현장에서 도 컴퓨터를 도입하여 활용하고, 각종 교육용 소프트웨어 개발을 시작한지도 십여 년이 지나가고 있다. 컴퓨터가 보편화되고 교육용 소프트웨어에 대한 수요가 늘어남에 따라, 교육용 소프트웨어 품질평가에 대한 필요성도 증대되었다.

컴퓨터 소프트웨어는 무형의 지적 산물로서, 상품으로 보기 시작한 연륜이 짧아 품질평가가 제대로 이루어지지 않았다. 그러나 상품의 품질이 좋지 못하면 경쟁력을 상실하게 되므로 소프트웨어 품질의 중요성을 인식하게 되었고, 품질 획득을 위한 품질보증활동이 필요하게 되었다. 이러한 필요성은 일반 업무용 소프트웨어 개발분야에서 인식되어 현재 우리 나라에는 컴퓨터 소프트웨어 질 관리를 위한 법적 장치가 마련되어 있다. 소프트웨어 개발 촉진법 제 7조에 의거, 과학기술처는 1991년 4월 1일, 프로그램의 품질 향상과 신뢰성 확보 및 유통 촉진을 위하여 프로그램에 관한 품질보증기준을 공고하고 이를 준수할 것을 권고하였다. 품질보증활동은 개발과정에 대한 평가활동과 개발제품에 대한 평가활동, 즉 인증활동으로 나뉘어 수행된다. 그러나 아직까지는 체계적인 품질보증활동을 수행하고 있는 기업체는 그리 많지 않은 실정이다.

교육용 소프트웨어는 교수-학습자료로서 갖추어야 할 교육적인 기능을 가진 높은 품질 수준을 가지고 있어야 한다. 개발자의 입장에서는 최소의 비용으로 최대 이윤을 얻고자 하기 때문에 비용과 소프트웨어의 품질 수준간에 적당한 타협점을 찾는 방향으로 개발을 하게 된다. 반면 교육용 소프트웨어의 사용자 입장에서는 교육 효과가 높고 교육적 기능이 잘 갖추어진 좋은 품질의 소프트웨어를 원하게 된다. 교육용 소프트웨어는 내용이나 효과, 그리고 이 소프트웨어를 사용하는 사용자 등에서 일반적인 업무용 소프트웨어와 구별되는 특성을 가지고 있으므로 보편적인 소프트웨어 품질평가기준을 적용할 수는 없다.

개발된 교육용 소프트웨어에 대한 올바른 평가를 위해서는 타당성 있고 신뢰성 있는 평가도구가 필요하다. 또한 개발자의 입장에서도 이러한 평가도구는 필요한데 왜냐하면 이것이 개발과정에서의 품질관리기준이 될 수 있기 때문이다. 미국, 일본, 중국 등 여러 나라는 다양한 교육용 소프트웨어 평가도구를 마련하여 품질보증활동을 수행하고 있으나 우리 나라의 경우 교육용 소프트웨어의 품질평가에 관한 연구는 초기상태에 있다고 할 수 있다[1, 4, 5, 9, 10, 11, 12]. 따라서 사용자 측면에서 평가할 수 있는 평가지가 개발되고 평가활동이 보편화되면, 개발자가 품질향상을 위한 노력을 하도록 하는 효과를 가져와서 교육용 소프트웨어의 질을 높이는 데에도 기여하게 되고, 사용자들이 소프트웨어를 구입할 때에도 도움을 줄 수 있게 된다. 결과적으로 시장도 활성화되어 개발자의 개발 욕구를 높여주는 등 많은 효과를 얻을 수 있다.

본 연구의 목적은 사용자, 즉, 학습자, 교사, 학부모의 각 입장에서 교육용 소프트웨어를 평가하기 위한 타당성 있고 신뢰성 있는 평가지를 개발하고자 하는 것이다. 기존의 평가지는 개발자나 교육전문가가 평가를 하도록 마련된 것이어서 평가기준이 너무 어렵고 복잡하여 전문가가 아닌 사람들이 사용하기 어려운 문제점이 있었다. 따라서 소프트웨어를 직접 사용하는 학습자, 교사, 학부모 입장에서 평가할 수 있는 평가도구가 개발되어야만 사용자의 요구를 제대로 반영하는 평가가 이루어질 수 있다.

교육용 소프트웨어 개발 초기에는 소프트웨어를 디스켓에 저장한 형태로 생산하였으나, 컴퓨터 하드웨어 기술의 발전에 따라 멀티미디어 컴퓨터가 보편화되고 저장매체로서는 CD-ROM을 사용하게 되었다. 또, 인터넷을 교육에 활용하는 사례가 점점 많아지고 있으므로 인터넷 상에 올려져 있는 교육용 소프트웨어에 대한 평가도 필요하다. 인터넷 상의 교육용 소프트웨어는 아직은 그 숫자가 많지는 않으나 앞으로 점차 늘어날 것은 당연한 추세이며, 현재로서는 검색속도가 느리다는 문제점이 있지만 앞으로 케이블 TV망, 위성, 초고속 정보통신망 등에 의한 기술적인 발전에 힘입어 이 문제점이 해소되는 것은 시간문제가 될 것이다.

따라서 본 연구의 연구대상 교육용 소프트웨어로는 최신의 CD-ROM 타이틀과 인터넷에 올려진 소프트웨어로 정하였고, 연구대상 교육용 소프트웨어가 대상으로 하는 연령층은 유아에서부터 초등학교 학생까지로 제한하였다. 대상 연령층을 넓게 잡는 경우 중, 고등학교, 성인용 소프트웨어에 대한 평가지는 그 평가항목이 어린이용과는 다소 달라져야 한다.

2. 평가지 개발

2.1 CD-ROM 타이틀용 평가지 개발

소프트웨어의 품질이란 '소프트웨어 전 사용기간 동안에 사용자의 특별한 요구와 기대에 부합하는 정도'로 정의할 수 있고, 품질보증활동이란 '소프트웨어에 요구되는 품질 목표를 달성하기 위한 일련의 계획적이고 체계적인 모든 활동'을 말한다.

소프트웨어 공학 측면에서, 소프트웨어에 대한 품질을 평가하기 위해서는 평가대상 소프트웨어가 요구되는 품질목표와 품질기준을 만족하고 있는지를 점검해야 한다. 품질목표란 '프로그램의 품질을 구현하기 위한 개발 및 평가의 지표가 되는 품질특성'을 말하고, 품질기준이란 '품질목표를 보다 기술적인 측면에서 구체화시킨 품질요소'이다.

교육용 소프트웨어에 대한 사용자용 평가지를 개발하기 위하여, 먼저 교육용 소프트웨어를 '교육목표를 성취하기 위하여 교수, 학습과정에 직접 투입되는 컴퓨터 소프트웨어'로 정의하고 사용자를 학습자, 교사, 학부모 집단으로 나누었다. 각각의 사용자 입장에서 보았을 때 좋은 교육용 소프트웨어가 갖추고 있어야 할 품질특성이 무엇인지를 컴퓨터과학에서의 소프트웨어 공학, 인공지능, 교육학에서의 학습이론, 교수이론 등을 바탕으로 뽑아내었다. 그 다음 각각의 품질특성을 갖추고 있는지를 정확하게 평

가하기 위해 품질기준, 즉, 평가기준을 정하고 각 기준에 알맞은 평가항목을 설정하였다.

2.1.1 품질이 좋은 교육용 소프트웨어가 갖추어야 할 특성

1. 소프트웨어 공학적 특성

일반적인 소프트웨어를 개발할 때 고려해야하는 품질특성을 국제표준기구(International Standard Organization)는 아래와 같이 6가지로 규정하고 있다[송재형, 1990].

기능성(Functionality)	: 사용자의 요구를 만족하는 기능과 특성을 가지고 있는 속성
신뢰성(Reliability)	: 정해진 시간과 조건하에서 정확하고 일관된 결과로 요구된 기능을 수행하는 속성
유용성(Usability)	: 배울 수 있고 사용할 수 있는 속성
효율성(Efficiency)	: 최소의 시간과 기억용량을 소비하여 요구되는 기능을 수행할 수 있는 속성
유지보수성(Maintainability)	: 소프트웨어 변경시 필요한 노력과 관계되는 속성
이식성(Portability)	: 여러 하드웨어 환경에서도 운용가능하도록 쉽게 수정될 수 있는 속성

기능성, 신뢰성, 유용성 특성은 사용자 입장, 효율성, 유지보수성, 이식성은 개발자 입장에서 보았을 때의 품질특성이 된다. 교육용 소프트웨어는 일반 업무용 소프트웨어와는 구별되지만, 사용자 입장에서 봤을 때 품질이 좋은 소프트웨어이기 위해서는 일반적인 소프트웨어의 경우와 마찬가지로 기능성, 신뢰성, 유용성을 갖추고 있어야 한다.

사용자 측면에서 교육용 소프트웨어의 품질을 평가하기 위하여 기능성으로부터는 포장정보의 유용성, 투자가치, 신뢰성으로부터는 기술적 지원도, 아동을 고려한 정도, 유용성으로부터 사용의 용이성 특성을 도출하였다.

2. 교육학적 특성

교육심리학의 교육이론에서, 객관주의 입장에서는 지식이란 고정되어 정체화(identifiable)할 수 있는 존재이고 인지 주체(cognitive subject)로부터는 독립된 것으로 본다. 따라서 학습을 '객관적, 독립적으로 존재하는 실재(reality)를 반영하고 표현해 나가는 과정'으로 정의한다. 이와 같은 관점에서의 학습이론이나 수업이론이 지닌 문제점이 제기되면서 그 대안으로 구성주의 이론이 대두되었다. 구성주의 입장에서는 지식이란 학습자에게 전달되는 정형화된 형태가 아니고, 학습자가 자신의 역사적, 문화적, 사회적 상황을 바탕으로 하여 스스로 능동적으로 구성해나가는 것이라고 본다. 또, 교육을 어떤 사실이나 공식을 가르치는 것으로 보지 않고, 학습자가 스스로 지식을 창조하고 구성해 나가고 수정해 나가도록 격려하는 것으로 본다[강명희 외, 1996].

또, 놀이이론에서는 놀이는 학습자 스스로가 자발적으로 즐기는 행동이며, 학습자가

도전감을 가지고 다양한 경험을 하도록 해 주고 창의적이고 탐구적인 능력이 생기도록 해 주기 때문에 이러한 놀이의 특성을 교육에 접목시키면 교육적 효과와 동기 유발 면에서 뛰어난 효과가 있다고 한다.

교육용 소프트웨어의 품질특성으로서 구성주의 이론으로부터 교육적 가치, 놀이이론으로부터 학습의 흥미도 특성을 도출하였다.

3. 인공지능적 특성

기존의 컴퓨터 보조학습 프로그램(CAI : Computer Assisted Instruction)에서는 교과과정이 고정되어 있고, 학습자의 전체적인 수준에 맞추어 난이도를 몇 가지로 다르게 제시해서 학습자가 자신에게 적합하다고 생각하는 수준을 정해 학습을 하도록 되어 있다. 또 정해진 순서에 따라 사전에 준비된 다수의 문제가 주어지며, 학습자의 응답에 대해 맞고 틀린 것을 비교해 주고 필요한 경우 고정된 한가지의 설명이 주어진다. 즉, 문제의 제시방법이라든지 틀린 답에 대한 설명 등이 고정되어 있어서 학습자 개인 차이에 의한 수업이 될 수 없었으며 학습자가 별로 흥미를 느끼지 못하였다. 따라서 교육효과도 별로 크지 못했다.

이러한 문제점을 보완하고 좀 더 적응력 있는 교육 시스템으로 개발하기 위하여 1970년대에는 CAI에 인공지능기법이 도입되었다. 인공지능의 원리와 기법을 적용해서 만들어진 시스템이 ICAI(Intelligent CAI) 또는 ITS(Intelligent Tutoring System)이다. 교육적 효과가 큰 소프트웨어이기 위해서는 인공지능 측면에서 얼마나 지적인 능력을 지녔는지가 중요한 특성이 된다. 따라서 교육용 소프트웨어의 품질특성으로서 지능성(smartness, intelligent한 정도)을 도출하였다.

한가지 특성을 두 사용자 집단 이상에서 필요한 특성이라고 할 수 있으나, 한 사용자 집단에서 특히 중요하다고 생각하는 특성은 평가를 실시할 평가자가 바로 그 사용자 집단이므로 그 집단에서만 고려하는 것으로 하였다. 각 사용자 입장에서 보았을 때, 좋은 교육용 소프트웨어가 반드시 갖추고 있어야 하는 특성 및 각 특성의 정의는 아래와 같다.

(1) 학습자 입장

- | | |
|---------|--|
| 사용의 용이성 | : 하드웨어나 소프트웨어에 대한 특별한 지식없이 프로그램을 쉽게 사용할 수 있는 정도 |
| 기술적 지원도 | : 프로그램에 사용된 컴퓨터 기술의 질이 학습활동을 효과적으로 뒷받침해 주고 있는 정도 |
| 학습의 흥미도 | : 학습자가 프로그램에 집중해서 재미를 느껴 계속하는 정도 |

(2) 교사입장

- | | |
|--------|--------------------------------------|
| 교육적 가치 | : 구성주의 측면에서 학습 내용 및 교육 방법이 가치가 있는 정도 |
| 지능성 | : 프로그램이 영리하고 지적인 정도 |

(3) 학부모 입장

- 포장정보의 유용성 : 포장에 나타나 있는 정보가 얼마나 유용한가에 대한 정도
 투자가치 : 가격이 적절한지에 대한 정도
 아동을 고려한 정도 : 사용자가 아동이라는 것을 고려한 정도

2.1.2 평가기준 및 평가항목 설정

구체적인 평가기준, 평가항목을 결정하기 위하여 국내[1, 4, 5, 9, 10, 11, 12] 및 국외[14, 18]의 다양한 자료를 검토, 참고하였으며, 전문가들[2, 3]의 의견을 수렴하였다. 개발하고자 하는 평가지는 학습자를 포함한 사용자용이므로 사용자가 평가지를 가지고 소프트웨어를 평가하고자 할 때, 평가항목이 너무 어렵고 많으면 이해하기 어렵고 제대로 평가할 수 없게 된다. 그러므로 평가지는 사용자 집단 별로 한 페이지 내로 작성되어야만 사용자가 별 부담없이 성의 있게 평가항목에 답할 수 있을 것이다. 따라서 평가항목이 너무 복잡하고 많아지지 않도록 품질특성을 평가하는데 있어서 꼭 필요한 평가기준 및 평가항목을 결정하였다. 또한 평가항목에서 사용되는 용어는 될 수 있으면 쉽고 의미가 분명하도록 하였으며, 사용자가 항목의 내용을 제대로 잘 이해하는지와 구체적인 평가척도를 결정하기 위하여 사용자들로 하여금 직접 pilot-test를 하게 한 후, 의견을 조사하여 수정, 보완하는 과정을 거쳤다.

(1) 학습자 입장

<사용의 용이성>

학습자 입장에서 소프트웨어를 구입하여 사용하려고 할 때, 좋은 소프트웨어라면 컴퓨터에 대한 지식이 별로 없어도 어려움 없이 쉽게 사용할 수 있어야 한다. 즉, 사용자 인터페이스가 사용자의 능력과 수준을 고려한 알맞은 형태로 제공되어야 하고, 사용법에 대한 안내가 충분하여야 한다. 사용의 용이성 특성을 평가하기 위한 평가기준과 각 기준에 의한 평가항목은 아래와 같다.

1. 사용자가 최소로 가져야 하는 능력

- 프로그램을 쉽게 설치할 수 있다.
- 글자를 몰라도 사용하는데 지장이 없다(유아용).
- 혼자서 프로그램을 진행시킬 수 있다.

2. 편리성

- 메뉴를 쉽게 사용할 수 있다.
- 학습한 내용을 프린트하기 쉽다.
- 원할 때 프로그램을 중단하거나 다시 시작하기가 쉽다.

3. 사용지침 여부

- 소프트웨어와 함께 들어 있는 설명서는 사용하는데 도움이 된다.
- 도움말이 자세하게 제공된다.
- 학습하다 잘못했을 때 오류 메시지가 나온다.

<기술적 지원도>

정확하고 일관된 결과로 요구된 기능을 수행할 수 있는 능력인 신뢰성은 소프트웨어가 가져야 하는 품질특성 중에서 아주 중요한 것이다. 따라서 교육용 소프트웨어가 학습활동을 보다 효율적으로 지원하기 위해 컴퓨터의 특성을 활용하는 기술적 효과와 효율성을 가지고 있는지를 평가해야 한다. 기술적 지원도 특성을 평가하기 위한 평가기준과 각 기준에 의한 평가항목은 아래와 같다.

1. 안정성

- 그래픽이 깨지지 않는다.
- 시스템 다운이 일어나지 않는다.

2. 신속성

- 화면이 금방 금방 나타난다.

3. 조작성

- 아이콘이 쉽게 눈에 띈다.
- 소리의 크기를 조절할 수도 있고 끌 수도 있다.

<학습의 흥미도>

놀이의 특성을 교육에 접목시켜 교육용 소프트웨어를 개발하고자 하는 전략인 에듀테인먼트(Edutainment)가 가미된 교육용 소프트웨어에서는 학습자의 흥미와 발달정도를 고려한 다양한 놀이활동이 제공되므로 교육적 효과와 동기유발면에서 뛰어난 효과가 있다.

Rieber(1996)도 놀이(play)학습의 학습효과면에서의 중요성을 논하고, 놀이의 개념과 컴퓨터의 마이크로월드(microworld), 시뮬레이션(simulation), 게임(game)의 세 가지 학습환경의 특성이 일치한다고 하였다. 마이크로월드란 관심의 대상이 되는 작지만 완전한 영역을 말하는데, 이것은 시뮬레이션과는 다른 두 가지 특성을 가진다. 첫째로는 학습자에게 대상영역의 간단한 사례를 제공한다는 것이고, 둘째로는 학습자의 인식상태, 감정상태와 조화를 이룬다는 것이다. 이러한 특성이 갖춰지면 학습자는 자율학습(self-regulated learning)이 가능하게 된다. 자율학습이 가능한 마이크로월드의 설계는 시뮬레이션과 게임환경을 구축하면 어느 정도 가능해지는데, 왜냐하면 시뮬레이션은 대상에 대해서 직접적인 연관을 지워주고 게임은 자율학습이 가능하도록 실제적인 방법을 제공해주기 때문이다.

학습의 흥미도 특성을 평가하기 위하여 이상의 에듀테인먼트와 마이크로월드의 개념을 고려한 평가기준과 평가항목은 아래와 같다.

1. 적정성

- 색상이 다양하고 조화롭게 사용되었다.

2. 흥미성

- 프로그램이 매우 재미있다.
- 반복해도 싫증나지 않는다.
- 프로그램을 다시 해 보고 싶다.

3. 프로그램 형식

- 게임 형식이 가미되었다.

4. 연관성

- 프로그램은 실제 상황을 보여준다.
- 어떤 일을 직접 경험해 보는 것과 같은 효과를 준다.

5. 이해성

- 프로그램에 나오는 설명, 그림은 이해하기 쉽다
- 그래픽이 내용을 알기 쉽게 해 주고 매우 재미있다.
- 음향, 음성은 학습에 도움을 준다.
- 새로운 것을 많이 배운 것 같다.

(2) 교사 입장

<교육적 가치>

구성주의 이론에서는 틀에 박힌 너무 잘 짜여진 수업과정에서는 어려운 문제의 해결능력이나 학습자 개개인의 사고력은 향상될 수 없다고 본다. 구성주의 원칙을 따른 교육용 소프트웨어에서는, 교육내용이 자세하게 틀에 박힌 형태로 주어지는 것이 아니라 학습자에게 관련된 자료들을 충분히 제공하여 학습자가 자신의 인지과정에 따라 학습하도록 환경이 만들어져야 한다. 따라서 학습자 스스로가 프로그램의 진행속도를 조절하고 난이도를 조절하도록 하여 스스로 지식을 구성해 나가도록 해야 한다. 또한 피드백이 다양하게 주어져야 하고, 컴퓨터가 즉각적으로 반응하여 학습내용을 이해하는데 도움을 줄 수 있어야 한다. 교육적 가치 특성을 평가하기 위한 평가기준 및 평가항목은 아래와 같다.

1. 다양성

- 내용이 다양하게 제시된다.
- 피드백이 다양하게 주어진다.

2. 효과성

- 컴퓨터가 제시하는 반응은 학습내용의 이해를 도와준다.

3. 자율성

- 프로그램 진행속도를 조절할 수 있다.
- 학습자가 난이도를 조절할 수 있다.

4. 활용성

- 성, 인종, 지역차별에 대한 내용이 없다.
- 교육 지침서가 있다.
- 현 교과과정에 사용할 수 있다.

<지능성>

ICAI 시스템에서는 교사의 전문적인 지식을 이용하므로 인공지능에서의 전문가 시스템의 구조와 특성을 가진다[심임섭, 1994]. ICAI는 가르치는 주제에 관한 지식, 학습

자에 관한 지식, 교수법에 관한 지식 등과 추론능력을 가지고 있다. 학습자 개인의 능력과 수준에 맞는 개별화된 교과내용이 제공되며, 각 단계에서의 학습자의 반응에 따라서 학습자를 가장 효율적인 교수법으로 지도할 수 있는 교수계획이 수립되고 그에 따른 학습내용이 제공된다. 학습자의 반응에 대한 정보를 누적해서 가지고 있으며 사용자에게 그 정보를 제공할 수도 있다.

ICAI에서는 학습자가 필요에 따라서는 재학습을 할 수도 있고 필요 없다고 생각되는 부분은 뛰어 넘을 수도 있으며 언제든지 그만 둘 수도 있다. 또 학습자료와 교수 부분이 분리되어 있어서 학습자 개인에게 다른 종류의 설명이 제공될 수도 있고, 개인수준에 맞는 학습 교정이 가능하며, 다량의 학습문제를 보유하고 있는 데이터베이스로부터 개인의 학습진척 여부에 따라 적당한 문제를 가져올 수도 있다. 따라서 학습자의 능동적인 학습과 개별화 학습이 가능하게 된다.

또한 지능형 교육시스템이 되기 위해서는 학습자와 효율적인 상호작용이 필요하므로 문자, 기호, 음성, 음악, 정지화상 및 동화상 등으로 이루어진 멀티미디어와 하이퍼텍스트 등 첨단 컴퓨터의 성능을 활용할 수 있어야 한다. 멀티미디어 기술은 하이퍼텍스트와 결합하여 하이퍼미디어라는 개념으로 발전하였는데, 다양한 정보를 복합미디어로 표현할 수 있으므로 이를 교육에 이용하는 경우에 보다 다각적이고 이해하기 쉬운 형태로 자료를 구성할 수 있어서 교육효과를 높일 수 있다. 또 정보를 하이퍼텍스트 형태로 구성함으로써, 학습자의 사고과정에 따라 학습자가 정보를 손쉽게 검색할 수 있기 때문에 학습자에게 학습의 통제권을 부여하여 학습과정에 적극적, 능동적으로 참여할 수 있도록 한다. 학습자는 자신의 인지구조에 따라 필요한 정보가 있는 곳으로 언제라도 접근할 수 있으며, 다시 처음의 위치로 돌아오거나 반복해서 보는 것도 가능하다. 또, 앞으로는 음성인식에 대한 연구도 활발히 진행되어 컴퓨터와 학습자 사이에 대화가 가능해져야 한다.

아직까지는 이러한 모든 특징을 가지고 있는 시스템은 많지 않으나, 현재 다양한 인공지능 분야의 배경을 가진 사람들의 연구가 진행 중에 있고 교육용 소프트웨어가 앞으로 지향해야 하는 방향이 지능형 시스템이므로 지능성 특성을 평가하기 위한 평가기준 및 평가항목은 아래와 같다.

1. 개별성

- 프로그램이 학습자의 수준에 따라 난이도를 조절해 준다.
- 학습자의 수준에 따라 피드백이 적절히 조절되어 주어진다.

2. 출력 적정성

- 학습자의 학습과정과 성적을 저장하는 기능이 있다.
- 프로그램에 프린트하는 기능이 있다.

3. 능동성

- 학습자의 생각이나 구상이 프로그램에 반영될 수 있다.
- 제목이나 설명부분을 뛰어넘거나 다시 되돌아 갈 수 있다.

4. 멀티미디어 사용성

- 그래픽, 음향, 동화상, 음악이 학습에 효과적으로 이용된다.

- 프로그램에서 음성이 제공된다.
 - 학습자의 음성을 입력할 수 있다.
5. 하이퍼텍스트 사용성
 - 화면의 어떤 특정부분에 대해 더욱 자세한 내용으로 접근할 수 있다(Hyper 기능).
 6. 상호작용성
 - 학습자가 자신의 답이 맞았는지 곧 알 수 있다.
 - 학습자가 자신의 답이 틀린 것을 곧 알 수 있고 틀린 이유를 프로그램이 설명한다.

(3) 학부모 입장

<포장정보의 유용성>

학부모가 교육용 소프트웨어를 구입할 때에는 소프트웨어 포장에 나타나 있는 정보에 의존하게 되므로, 소프트웨어 품질특성인 기능성을 갖추려면 포장정보가 정확해야 한다. 포장정보의 유용성 특성을 평가하기 위한 평가기준 및 평가항목은 아래와 같다.

1. 대상 연령의 표시여부
 - 겉 포장에 사용 대상연령이 나타나 있다.
2. 학습목표의 표시여부
 - 겉 포장에 학습 목표가 나타나 있다.
3. 학습내용의 표시여부
 - 겉 포장에 학습내용이 설명되어 있다.
 - 겉 포장에 프로그램에 대한 구체적인 정보(예: 컴퓨터 화면)가 있다.

<투자가치>

소프트웨어 품질특성인 기능성을 가지고 있는지를 평가하기 위하여, 소프트웨어의 비용을 담당하는 학부모 입장에서 봤을 때, 소프트웨어의 내용이 가격만큼의 가치가 있는지 평가하는 것이 중요하다. 투자가치 특성을 평가하기 위한 평가기준 및 평가항목은 아래와 같다.

1. 가격의 적절성
 - 가격에 비해 내용이 다양하고 풍부하다.
2. 추천 가능성
 - 다른 사람에게 추천하고 싶다.

<아동을 고려한 정도>

소프트웨어 품질특성인 신뢰성을 가지고 있는지 평가하기 위해서, 사용자가 아동이라는 사실을 고려하여 아동의 입장에서 개발된 소프트웨어인지를 평가한다. 아동을 고려한 정도 특성을 평가하기 위한 평가기준 및 평가항목은 다음과 같다.

1. 사용통제성
 - 자녀가 키보드를 아무렇게나 마구 눌러도 이상이 없다.

- 자녀가 마우스를 여러 번 클릭하거나 하나의 키를 계속 누르고 있더라도 한번의 작동으로 인정한다.
- 특정한 키를 사용하면 부모용 지침내용을 볼 수 있다.

2.1.3 평가기준 및 평가항목 설정에 대한 검토

품질이 좋은 교육용 소프트웨어가 갖추고 있어야 할 특성을 평가하기 위한 평가기준과 평가항목을 컴퓨터 과학과 교육학의 이론을 바탕으로 설정하였다. 물론 완벽하고 절대적인 기준이나 항목은 있을 수 없으며 다른 기준이나 항목이 추가될 수도 있을 것이다. 예를 들어, 학습자 입장에서의 <사용의 용이성> 특성 평가를 위한 기준으로서 입력방법의 일관성, <기술적 지원도> 특성 평가를 위한 기준으로서 출력결과의 정확성, 교사입장의 <교육적 가치> 특성 평가를 위한 기준으로서는 내용의 명료성, 내용 조직의 적절성 등과, 해당 평가기준에 따른 평가항목들을 생각할 수 있다. 본 논문의 대상은 사용자용 평가지이므로 평가 항목이 너무 많고 복잡해지지 않도록 하는 것을 중요한 고려사항으로 삼아 세부적인 기준과 항목은 제외하였다. 또한 <지능성>의 평가 기준인 멀티미디어 사용성을 <학습의 흥미도> 특성에서도 평가기준으로 삼을 수 있으나 중복평가를 하지 않기 위하여 <지능성> 특성에서만 다루었다.

2.2 인터넷 상의 교육용 소프트웨어용 평가지 개발

인터넷에서의 멀티미디어 정보접근 도구 WWW는 학습의 본질을 바꾸어 놓을 수 있는 잠재력을 가지고 있으며, 교수방법의 개발에 이용될 수 있고 효과적인 교수 매체로 쓰일 수 있다[Kinzie, 1996]. 또, 교육학의 구성주의 이론을 실현시킬 수 있는 적절한 환경을 제공해 준다.

따라서 앞으로 급속하게 늘어날 인터넷 상의 교육용 소프트웨어에 대하여서도 그 질의 좋고 나쁨을 사용자 입장에서 평가하는 것이 필요하다. 평가지가 마련되고 학습자 집단, 교사 집단, 학부모 집단으로 구성된 평가단에서 평가한 결과에 따라 우수한 교육용 소프트웨어 사이트를 소개해 줄 수 있으면 인터넷을 통한 교육적 활용효과를 극대화할 수 있을 것이다.

사용자가 인터넷 상의 소프트웨어에 대하여 평가를 하기 위한 평가지도 CD-ROM 타이틀에 관한 평가지와 크게 다르지 않게 작성할 수 있다. 각각의 사용자 입장에서 보았을 때 품질이 좋은 소프트웨어가 갖추고 있어야 하는 특성으로, 학부모 입장에서의 포장정보의 유용성 대신 홈 페이지의 유용성으로 한 것 외에는 CD-ROM 타이틀의 경우와 동일하게 도출하였다. 각각의 특성을 평가하기 위한 평가항목도 몇 가지 항목만 제외하고는 CD-ROM 타이틀에 관한 평가항목과 거의 동일하다. 차이가 나는 항목을 정리하면 아래와 같다.

- 삭제할 평가항목

<사용의 용이성>

- 프로그램을 쉽게 설치할 수 있다.

- 소프트웨어와 함께 들어있는 설명서는 사용하는데 도움이 된다.

<기술적 지원도>

- 화면이 금방 금방 뜬다.

<포장정보의 유용성>

- 겉 포장에 사용 대상연령이 나타나 있다.
- 겉 포장에 학습 목표가 나타나 있다.
- 겉 포장에 학습내용이 설명되어 있다.
- 겉 포장에 프로그램에 대한 구체적인 정보(예:컴퓨터 화면)가 있다.

- 추가할 평가항목

<기술적 지원도>

- 검색속도가 빠르다.

<홈페이지의 유용성>

- 홈페이지에 사용 대상연령이 나타나 있다.
- 홈페이지에 학습목표가 나타나 있다.
- 홈페이지에 학습내용이 설명되어 있다.

3. 평가지 분석

3.1 평가의 실시

본 연구에서 작성된 평가지의 각 항목에 대해서 사용자의 동의 여부와 동의 정도를 표시하도록 하였다. 동의하는 정도에 따라 '매우 그렇다', '그렇다', '약간 그렇다'에 표시하도록 하고, 동의하지 않으면 '그렇지 않다'에, 그 외 평가가 불가능한 경우에는 '대답할 수 없다'에 표시하도록 하였다.

실제로 개발된 평가지에 의한 교육용 CD-ROM 타이틀 평가는 다음과 같은 단계를 거쳐서 이루어질 수 있다.

1. 사용자 집단을 대표할 수 있는 학습자, 교사, 학부모들로 이루어진 평가단을 구성한다.
2. 각각의 사용자 평가단에서 해당 평가를 이용하여 타이틀들을 평가한다.
3. 평가한 타이틀 각각에 대해서 평가자 각자의 평가결과를 정리한다.
4. 평가자들의 점수를 평균해서 타이틀에 대한 평가결과를 알기 쉽게 나타낸다.

이렇게 해서 얻은 평가결과는 사용자들이 소프트웨어를 구입할 때 활용할 수 있게 된다

본 연구에서는 사용자용 평가를 실제로 적용하고 결과를 분석해 보기 위하여 다음과 같이 평가를 실시하였다.

3.1.1 CD-ROM 타이틀 및 평가자 선정

인터넷 상의 교육용 소프트웨어는 아직은 그 숫자가 많지 않기 때문에 구체적인 평가는 최근의 CD-ROM 타이틀에 대한 것만 실시하였다. 따라서 1995년도 이후에 개발된 2개의 미국 타이틀과 1개의 국내 타이틀을 선정하고, 평가단으로는 학습자로서 유치원 또는 초등학교 학생 10명, 학부모로서 선정된 학생의 어머니 10명, 교사로써는 공교육기관 및 컴퓨터 전문 교육센터의 선생님 10명으로 구성하였다. 평가지의 각 항목을 올바르게 이해하도록 하기 위하여 평가자들에게 자세한 설명을 하였으며, 글자를 모르는 미취학 어린이에 대해서는 어머니가 평가지의 내용을 읽어 설명하고 어린이가 답을 하도록 하였다. 타이틀 1, 2에 대해서 5명의 학습자, 5명의 교사, 5명의 학부모가 1주일 간격으로 반복 평가를 하였으며, 타이틀 3에 대해서는 또 다른 5명의 학습자, 5명의 교사, 5명의 학부모가 평가하였다.

3.1.2 평가결과

각각의 사용자별 평가지에 대하여, 각각의 특성별로, Buckleitner(1993)의 방법을 수정한 다음과 같은 공식을 사용하여 평가결과를 점수로 나타내었다.

$$S = [(X + (2Y)/3 + Z/3) / \{n - NA\}] \times 100$$

여기에서,

S = 각 사용자별 소프트웨어 평가 점수(백분율)

n = 각 평가지에 속하는 총 문항의 수

NA = '대답할 수 없다' 에 표시된 수

X = '매우 그렇다' 에 표시된 수

Y = '그렇다' 에 표시된 수

Z = '약간 그렇다' 에 표시된 수 이다.

따라서 모든 항목에 대해 '매우 그렇다'로 표시된 경우 S = 100%가 된다. 타이틀 1에 대한 사용자들의 첫 번째 평가결과는 다음과 같다.

< 표 1 > 타이틀 1에 대한 평가결과

		평가자 1	평가자 2	평가자 3	평가자 4	평가자 5
학 습 자	사용의 용이성	58.3	44.4	38.9	44.4	33.3
	기술적 지원도	66.7	80	66.7	73.3	66.7
	학습의 흥미도	66.7	75.8	57.6	60.6	63.6
교 사	교육적 가치	77.8	96.3	96.3	63.0	88.9
	지능성	66.7	77.8	75.8	54.5	69.7
학 부 모	포장정보의 유용성	33.3	16.7	66.7	41.7	50
	투자가치	66.7	100	100	66.7	100
	아동을 고려한 정도	33.3	66.7	66.7	66.7	66.7

다음으로 타이틀 1의 평가결과를 가지고 각각의 소프트웨어 특성, 사용자별, 전체적인 종합평가 결과를 위해 평가자들이 준 동의 정도를 평균해서 아래와 같이 나타내었다.

타이틀 1

사용의 용이성	(44)◆◆◆◆◆
기술적 지원도	(71)◆◆◆◆◆◆◆
학습의 흥미도	(65)◆◆◆◆◆◆◆
교육적 가치	(84)◆◆◆◆◆◆◆◆
지능성	(69)◆◆◆◆◆◆◆
포장정보의 유용성	(42)◆◆◆◆
투자가치	(87)◆◆◆◆◆◆◆◆◆
아동을 고려한 정도	(60)◆◆◆◆◆◆◆
학습자	(60)◆◆◆◆◆◆◆
교사	(77)◆◆◆◆◆◆◆◆
학부모	(63)◆◆◆◆◆◆◆
종합평가	(65)◆◆◆◆◆◆◆

타이틀 1은 좋은 교육용 소프트웨어가 갖추어야 할 특성 중 사용의 용이성 및 포장정보의 유용성을 제외하고는 전체 100% 중 60% 이상 87%의 동의정도, 종합 평가결과로는 65%의 동의정도를 얻은 비교적 좋은 품질의 제품임을 알 수 있다.

실제로 많은 타이틀에 대한 평가가 실시되어 평가결과에 대한 자료가 많이 축적되면, 특정 점수 범위에 대해서 타이틀의 품질수준을 “매우 우수”, “우수”, “보통”, “좋지 않음” 등으로 판정할 수 있을 것이다.

3.2 타당도 및 신뢰도 검증

개발된 평가지가 교육용 소프트웨어 평가를 위한 좋은 도구이기 위한 필수적인 조건은 타당도(validity)와 신뢰도(reliability)를 갖추고 있는 것이다[8, 16].

3.2.1 타당도

타당도란 평가도구가 목표를 어느 정도로 충실히 측정하고 있는지를 나타내는 척도이다. 본 연구에서 개발된 평가지에 대한 타당도는 평가지 개발을 위한 이론 연구와 실제 개발과정으로부터 뒷받침될 수 있다.

1. 컴퓨터 과학에서의 소프트웨어 공학, 인공지능, 교육학에서의 학습이론, 교수이론으로부터 좋은 교육용 소프트웨어가 갖추고 있어야 할 특성을 도출하였다.

2. 평가기준 및 평가항목을 정하는데 국내 및 외국의 다양한 자료를 참고하였다.
3. 장기간에 걸쳐 교육공학 전공의 교수, 컴퓨터 전문 교육 센터의 연구원 등 전문가의 검증을 거쳐 평가항목의 내용 및 용어에 대한 여러 번의 수정작업을 행하였다.
4. 개발된 평가지를 이용한 pilot-test 과정을 거쳐 평가항목의 표현 및 평가척도를 결정하였다.

3.2.2 신뢰도

신뢰도란 평가도구의 측정상의 정확성 또는 일관성을 나타내는 척도이다. 본 연구의 평가지에 대한 신뢰도를 알아보기 위해 5명씩의 평가자가 평가한 결과의 일치도를 알아보는 평가자간 신뢰도(interrater reliability)와 1차-2차 평가 결과의 일치도를 알아보는 재검사 신뢰도(retest reliability)를 구하였다.

(1) 평가자간 신뢰도

평가자간 평가결과의 일치도를 알아보기 위하여 Fisher(1958)가 제안한 내급간 상관계수(ICC : intraclass correlation coefficient)를 이용하는 Bartko(1966)의 방법을 사용하였다. 이 방법에서는 내급간 상관계수 ρ 에 대한 추정 값을 구하기 위하여 분산분석(Analysis of Variance) 표를 이용한다. 본 연구의 평가결과에 대해서는 Bartko(1966)의 이원 랜덤 모델(Two-way random model)을 적용하였다. 평가하고자 하는 특성내의 분산이 작으면, 즉 평가자간 일치도가 크면, 내급간 상관계수 ρ 가 큰 값을 가지고, 반대로 특성내의 분산이 크면, 즉 평가자간 일치도가 작으면, ρ 는 작은 값을 갖는다.

내급간 상관계수 ρ 에 대한 추정 값으로는

$$R = [MSP - MSE] / [MSP + MSE(k - 1) + k(MSR - MSE) / n] \quad (1)$$

을 구한다. 여기에서

MSP : 특성요인에 대한 평균제곱

MSR : 평가자요인에 대한 평균제곱

MSE : 잔차요인에 대한 평균제곱

k : 평가자 수

n : 특성의 개수이다.

R에 대한 유의성검정은 자유도가 $(n - 1)$, $(n - 1)(k - 1)$ 인 $F = MSP / MSE$ 값을 이용하였다[Bartko, 1966].

타이틀 1에 대한 5명의 학습자가 1차로 사용의 용이성, 기술적 지원도, 학습의 흥미도 특성을 평가한 평가결과로부터 통계 패키지 SAS를 사용하여 다음과 같은 분산분석표를 얻었다[SAS Institute, 1991].

변동요인	자유도	제곱합(SS)	평균제곱(MS)	F값	Pr > F (P값)
평가자	4	365.740000	91.435000(MSR)	2.31	0.1454
특성	2	1990.308000	995.154000(MSP)	25.18	0.0004
잔차	6	316.232000	39.529000(MSE)		

이 표를 이용하여 식 (1)을 계산하면,

$R = [995.154 - 39.529] / [995.154 + 39.529(4) + 5(91.435 - 39.529) / 3] = 0.7708$ 이다.

귀무가설 $H_0 : \rho = 0$ 에 대한 F검정($F = 25.18$)결과는, P값이 0.0004이므로 귀무가설을 기각하게 되고 따라서 $R = 0.7708$ 은 의미가 있다는 것을 나타낸다. 그러므로 평가자간 평가결과는 일치도가 있다.

같은 방법으로 타이틀 1에 대한 교사, 학부모들의 1차 평가결과, 타이틀 1에 대한 학습자, 교사, 학부모들의 2차 평가결과, 타이틀 2에 대한 학습자, 교사, 학부모들의 1차-2차 평가결과, 타이틀 3에 대한 학습자, 교사, 학부모들의 평가결과의 일치도를 같은 방법으로 알 수 있다. 다음 표는 각각의 경우에 대한 F값과 P값을 나타낸다.

< 표 2 > 평가자간 평가결과에 대한 F값과 P값

	타이틀 1				타이틀 2				타이틀 3	
	1차		2차		1차		2차		1차	
	F값	P값	F값	P값	F값	P값	F값	P값	F값	P값
학습자	25.18	0.0004*	4.80	0.0428*	9.51	0.0077*	2.67	0.1290	15.09	0.0019*
교사	41.73	0.0030*	12.14	0.0252*	4.81	0.0934	7.59	0.0511	19.08	0.0120*
학부모	10.05	0.0066*	12.11	0.0038*	6.93	0.0179*	13.14	0.0030*	2.16	0.1858

<표 2>에서 *표시를 한 P값은 0.05보다 작으므로 귀무가설 $H_0 : \rho = 0$ 를 기각하게 된다. 따라서 타이틀 2의 1차 평가에서 교사, 2차 평가에서 학습자, 교사, 타이틀 3에서 학부모 집단만 제외하고는 대체로 평가자간 평가결과의 일치도가 존재하는 것을 알 수 있다.

(2) 재검사 신뢰도

소프트웨어의 특성에 대한 1차-2차 평가결과의 일치도는 평가자간 평가결과의 일치도와 같은 방법으로 알아볼 수 있으므로 여기에서도 Bartko의 방법을 사용하였다. 또 Pearson 상관계수와 Kendall의 순위 상관계수를 SAS 통계 패키지를 사용하여 구하였다. 결과를 정리하면 아래와 같다.

< 표 3 > 1차-2차 평가결과에 대한 F값과 P값

	타이틀 1		타이틀 2	
	F 값	P 값	F값	P 값
사용의 용이성	13.51	0.0136*	7.86	0.0353*
기술적 지원도	6.78	0.0454*	1.27	0.4108
학습의 흥미도	7.50	0.0382*	1.22	0.4247
교육적 가치	14.82	0.0115*	456.62	0.0001*
지능성	52.30	0.0010*	20.03	0.0066*
포장정보의 유용성	24.97	0.0043*	33.70	0.0024*
투자가치	5.00	0.0741	2.33	0.2160
아동을 고려한 정도	24.53	0.0045*	7.50	0.0182*

Bartko 방법에 의한 결과에서는 *표시를 한 P값이 0.05보다 작으므로 타이틀 1에서 투자가치, 타이틀 2에서 기술적 지원도, 학습의 흥미도, 투자가치 특성을 제외하고는 1차-2차 평가결과의 일치도가 존재함을 알 수 있다.

< 표 4 > Pearson과 Kendall의 상관계수

	타이틀 1		타이틀 2	
	Pearson	Kendall	Pearson	Kendall
사용의 용이성	0.863(0.0594)*	0.527(0.2065)	0.802(0.1024)	0.949(0.0230)*
기술적 지원도	0.767(0.1303)	0.598(0.1655)	0.119(0.8483)	0.222(0.6027)
학습의 흥미도	0.806(0.0992)	0.671(0.1172)	0.154(0.8052)	0.222(0.6027)
교육적 가치	0.938(0.0183)*	0.882(0.0459)*	0.996(0.0003)*	0.949(0.0230)*
지능성	0.960(0.0097)*	0.949(0.0230)*	0.977(0.0042)*	0.949(0.0230)*
포장정보의 유용성	0.923(0.2191)	0.949(0.0230)*	0.943(0.0162)*	0.738(0.0770)
투자가치	0.667(0.2191)	0.667(0.1824)	0.408(0.4950)	0.408(0.4142)
아동을 고려한 정도	0.923(0.0254)*	0.714(0.1192)	0.875(0.0523)	0.926(0.0528)

<표 4>에서 괄호 안의 값은 각각의 소프트웨어 특성에 대해서 사용자들의 1차-2차 평가결과의 상관계수에 대한 귀무가설 $H_0 : \rho = 0$ 에 대한 P값을 나타낸다. 따라서 * 표시를 한 P값은 0.05보다 작으므로 H_0 를 기각할 수 있다.

또한, Pearson과 Kendall의 상관계수의 값을 보면 타이틀 1과 타이틀 2에서 대부분의 소프트웨어 특성에 대해 사용자들의 1차-2차 평가결과가 상관관계가 있음을 알 수 있다. 이상에서 1차-2차 평가결과의 일치도가 존재함을 알 수 있다.

평가자간 평가결과의 일치도와 재검사 일치도가 의미하는 바는 개발된 평가지에 의한 결과가 안정성을 보인다는 것이다. 따라서 개발된 평가지의 타당도와 신뢰도 결과로 볼 때 평가지에 대한 특별한 수정은 필요하지 않다고 할 수 있다.

4. 결론

본 논문에서는 사용자 입장에서 교육용 소프트웨어를 평가하기 위한 평가지를 개발하였다. 품질이 좋은 교육용 소프트웨어가 갖추어야 할 특성으로서, 학습자 입장에서 사용하는 용이성, 기술적 지원도, 학습의 흥미도, 교사입장에서는 교육적 가치, 지능성, 학부모입장에서는 포장정보의 유용성, 투자가치, 아동을 고려한 정도를 도출하여 각각의 특성을 평가하기 위한 평가기준 및 각 기준에 알맞은 평가항목을 정하였다.

개발된 평가지에 대한 타당도는 평가지 개발을 위한 이론연구와 실제 개발과정으로부터 뒷받침될 수 있으며, 개발된 평가지에 의한 구체적인 평가결과를 분석한 결과 평가자간 신뢰도 및 재검사 신뢰도를 확인할 수 있었다.

따라서 사용자 집단을 대표할 수 있는 학습자, 교사, 학부모집단을 선정하여 평가단을 구성한 다음, 본 논문에서 개발된 평가를 가지고 교육용 소프트웨어를 평가한 결과를 활용할 수 있게 되면, 소프트웨어의 품질을 높이는 데에도 기여하게 되고 사용자들이 실제로 소프트웨어를 구입하는 데에도 많은 도움을 줄 수 있게 된다. 그렇게 함으로서 교육용 소프트웨어에 의한 교육적 효과를 극대화시킬 수 있다.

본 논문에서는 CD-ROM 타이틀과 인터넷 상의 소프트웨어를 연구대상으로 하였으며, 연구대상 소프트웨어가 대상으로 하는 연령층은 유아에서부터 초등학교 학생으로 제한하였다. 앞으로 디스켓에 저장된 형태의 소프트웨어 및 중, 고등학교 및 성인용 소프트웨어에 대한 평가지에 대한 연구도 이루어져야 할 것이다.

또, 본 연구에서의 평가지에 의한 평가결과와 실제로 교육환경에 적용되었을 때 나타나는 교육효과 등의 관계에 대한 연구도 이루어져야 할 것이다.

참고문헌

- [1] 강명희, 차운옥, 김명숙(1996), 「교육용 소프트웨어 평가 및 인증제도 설립을 위한 기본연구」, 통신개발연구원 정보통신학술 연구과제 보고서
- [2] 권성호, 한양대학교 사범대학 교육공학과 교수
- [3] 김윤정, 신여명, 최준식, COMKID 교육연구소 선임연구원
- [4] 김영수(1990), "한국 CAI 코스웨어에 대한 평가연구," 「교육공학 연구」, 6권, 1호, pp. 115-139
- [5] 박성익(1986), "컴퓨터 소프트웨어 선정 방법," 「새교육」, 1월호, pp. 139-143
- [6] 송재형(1990), "소프트웨어 품질보증 계획과 적용," 「정보과학회지」, Vol. 8, No. 4, pp. 36-40
- [7] 심임섭(1994), "지능형 교육시스템," 「정보과학회지」, Vol. 12, No. 9, pp. 52-69
- [8] 윤순금, 이선희(1994), 「CAI 이론과 실제」, 서울 산업대학교

- [9] 정택희(1992), “교육용 소프트웨어 심의기준 개발연구 - 컴퓨터 보조수업 프로그램 심의기준은 중심으로,” 「교육개발」, 14권, 2호, pp. 52-59
- [10] 정택희 외(1990), 「교육용 소프트웨어 질 관리 방안 연구」, 연구보고 RR 90-18, 한국교육개발원
- [11] 정택희 외(1991), 「교육용 소프트웨어 심의기준 개발연구」, 연구보고 RR 91-14, 한국교육개발원
- [12] 정택희 외(1991), 「교육용 소프트웨어 심의기준 해설서」, 연구보고 RR 91-14-1, 한국교육개발원
- [13] Bartko, J.J.(1966), “The Intraclass Correlation Coefficient as a Measure of Reliability,” *Psychological Report*, Vol. 19, pp. 3-11
- [14] Buckleitner, W.(1993), *High/Scope Buyer’s Guide to Children’s Software*, High Scope Press, Ypsilanti, Michigan
- [15] Fisher, R.A.(1958), *Statistical Methods for research workers*, Hafner, New York
- [16] Guilford, J.P. & Fruchter, B.(1978), *Fundamental Statistics in Psychology and Education*, McGraw-Hill, Kogakusha, LTD
- [17] Kinzie et al(1996), “Frog Dissection via the Worldwide Web : Implications for Widespread Delivery of Instruction,” *Educational Technology Research and Development*, Vol. 44, No.2. pp. 59-69
- [18] Knowledge Adventure(1995), *How to Choose the Best Educational Software*, Knowledge Adventure, Inc., CA
- [19] Lieber, L.P.(1996), “Seriously Considering Play : Designing Interactive Learning Environments Based on the Blending of Microworlds, Simulations, and Games,” *Educational Technology Research and Development*, Vol. 44, No. 2, pp. 43-58
- [20] SAS Institute(1991), *SAS Procedures Guide, Release 6.03*, SAS Institute Inc., Cary, NC