

교사용 평가문항 서비스 사이트의 개선을 위한 사용편의성 평가 사례 연구

- 한국교육과정평가원의 평가문항 서비스 시스템을 중심으로 -

박 형 국(KT 미래기술연구소 선임연구원)

홍 선 주(한국교육과정평가원 부연구위원)

《 요 약 》

본 연구에서는 사용편의성에 영향을 미치는 요인으로서 효과성, 효율성, 학습용이성, 시스템 완성성, 만족도 등에 비추어 초·중·고 교사에게 평가문항을 제공하는 평가문항 서비스 사이트의 사용편의성을 평가하였다. 검사 방법으로는 반성적 검사 방법(retrospective testing)을 사용하였고, 검사에는 6명의 초·중·고 교사들이 참여하였다. 우선 참여 교사들에게 평가문항 서비스 사이트를 사용하여 필요한 문항을 검색 및 활용하는 과제들을 소리 내며 생각하기(think aloud)를 하면서 수행하게 하였고, 이를 녹화 분석하여 1차 분석 자료를 생성하였다. 녹화된 비디오를 보면서 인터뷰를 수행하는 자극회상기법(stimulated recall)을 사용하여 연구자들은 참여 교사들과 함께 비디오 자료를 보면서 1차 분석된 내용을 교사들에게 확인하고 추가적인 개선 사항들을 얻어내었다. 이 인터뷰 역시 녹화하였으며, 인터뷰 녹화 자료와 1차 분석 자료를 기반으로 2차 분석을 수행하였다. 1~2차 분석을 통하여 사이트의 주요 메뉴와 기능별 사용상의 문제점들을 밝히고, 이에 대한 개선안을 제안하였다.

주제어 : 사용편의성, 반성적 검사, 소리 내며 생각하기, 자극회상기법, 문제은행

I. 서론

1. 연구의 배경

최근 정보기술의 눈부신 발전으로 인간 생활의 많은 부분이 정보통신기술(ICT: Information and Communication Technologies)에 의존하면서 이루어지고 있다. 특히 정보를 찾고 지식을

생성해 내는 분야에서는 인간과 정보통신기술 간의 상호작용의 질이 인간의 수행 절차 또는 그 결과에 큰 영향을 미친다. 이렇게 사람과 기술 간 상호작용의 효율성이 중요시되기 때문에 사용자 인터페이스 설계(user interface design)에서 사용편의성(usability)을 향상시키려는 꾸준한 노력이 이루어지고 있다(Shneiderman & Plaisant, 2004).

이러한 노력들은 민간 기업이 고객 만족 최우선이라는 마케팅 전략을 사용하는 상황에서는 더욱 중요한 분야로 자리 잡고 있다. 가격과 기능, 성능이 똑같은 경우라면 고객을 감동시킬 수 있는 다른 요소는 디자인이라고 많이 강조되고 있으며, 감성마케팅의 중요한 한 분야로 자리 잡고 있다. 제품의 디자인은 심미적인 디자인과 사용자 인터페이스 디자인으로 다시 나눌 수 있으며, 사용자 인터페이스 디자인은 심미적인 부분과 다르게 필수적인 요소일 수 있다. 사용자 인터페이스는 사용자에게 어려움을 줄 수도 있고 감동을 줄 수 있는 기능적인 면에서 다루어져야 할 분야이다.

그렇다면 사용자의 편의성을 극대화하기 위한 인터페이스 설계는 그 중요성이 상업적인 제품에만 국한되는가? 학교현장에서의 컴퓨터 활용과 관련하여 교육용 소프트웨어나 웹사이트는 교사가 직관적으로 그 사용법을 인지하고 쉽게 사용할 수 있도록 설계되어 있는가? 김영애와 김혜숙(2005)의 연구에 의하면 2004년 12월 기준 초·중·고등학교의 교사 1인당 컴퓨터 수는 1.3대 정도로 나타났으며, 초·중·고등학교의 교사들은 수업시간 중 약 50% 이상을 ICT를 활용하여 수업을 진행한다고 응답하였다. 이렇게 ICT는 이제 더 이상 일부 교사만이 사용하는 것이 아닌 모든 교사가 수업과정 또는 수업을 준비하는 과정에서 사용해야 하는 중요한 도구가 되고 있다. 그러나 기술의 발전과 새로운 기술의 도입이 항상 인간의 문제를 저절로 해결해 주지는 못한다(Norman, 1993; Shneiderman, 1998). 경우에 따라서는 새로운 문제를 야기해 혼란을 가중시킬 수도 있다. ICT의 활용에도 그러한 측면이 있다. 학교현장에서 ICT의 활용은 교사나 학생의 학습이라는 측면에서 어느 정도 성과를 가져왔음에도 불구하고, 한편으로는 많은 교사들에게 여전히 어려움과 혼란을 야기하고 있는 것도 사실이다. 그러나 이러한 문제를 해결하기 위한 노력은 주로 교사들에 대한 정보소양 교육에 편중되어 왔고, 교사들이 활용하는 교육용 소프트웨어나 웹사이트 개발에 반영하려는 노력은 많지 않았다고 할 수 있다. 새로운 기술의 구현이나 적용이 목적이 될 뿐 사용자인 교사와 학생들의 입장이 간과되거나 시간, 예산, 개발의 편의성을 이유로 사용자 편의성에 대한 고려가 간과된다면 ICT의 교육적 활용은 성공하기 어려울 것이다.

지난 2002년 한국교육과정평가원의 교수학습개발센터 발족과 함께 2006년 현재까지 교실수업을 지원하기 위한 노력의 일환으로 교사를 대상으로 하는 많은 교수·학습 관련 자료, 평가 문항, 그리고 최근 들어 이러닝 콘텐츠까지 다양한 형태의 교수·학습 지원 자원들(resources)이 개발되어 한국교육과정평가원 교수학습개발센터 웹사이트(KICE-TLC: classroom.kice.re.kr)를 통해 학교현장에 서비스되고 있다. 이러한 서비스 중 본 연구에서 대상으로 하는 서비스는

KICE TLC의 하나의 메뉴이기도 하지만 동시에 독자적인 사이트로 운영되고 있는 ‘평가문항 서비스 시스템’이다.

평가문항 서비스 시스템은 1만6천8백여 개의 평가문항을 데이터베이스로 구축하고 검색엔진을 갖춘 웹사이트를 통해 문항을 서비스하는 일종의 문제은행 시스템이다. 따라서 평가문항 서비스 시스템 웹사이트(이하 평가문항 서비스 사이트)는 교사가 현장에서 유용하게 활용할 수 있는 방식으로 양질의 평가문항을 제공할 수 있어야 하며, 문항의 효율적인 검색과 추출 및 저장 등 문제은행으로서 갖추어야 할 기본적인 기능들(Ward & Marray-Ward, 1994)을 갖추어야 한다.

그러나 평가문항 서비스 사이트 구축 이래 2005년까지 약 4년 동안 평가문항의 지속적인 개발과 질 제고에 관심과 투자가 집중되어 왔던 것에 비해, 이 문항들을 실제로 교사가 쉽게 접근하고 활용할 수 있도록 사이트를 보완하고 개선하려는 노력은 상대적으로 적었다. 그렇다고 사이트가 전혀 개선되지 않은 것은 아니다. 2005년도의 경우 전국 12개 연구협력 학교를 지정하고 해당 학교의 교사를 중심으로 사이트에 대한 모니터링을 실시한 바 있다. 모니터링의 주된 초점이 평가문항의 질 개선에 주어지긴 하였지만 이와 아울러 문항사용에 어떤 어려움이 있는지 사이트 사용에 관한 모니터링도 함께 이루어졌다. 그 결과 사이트 접근성의 문제라든가 일부 기술적인 문제와 같은 거시적인 수준에서 사이트 사용에 불편을 초래하는 요인들은 개선이 되었다. 그러나 여전히 교사들이 이 평가문항 서비스 사이트를 이용해 문항을 검색한 후 이를 저장하거나 시험지를 만들어 활용하는 것이 현재의 시스템으로서는 쉽지 않다는 문제가 제기되어 왔다(윤현진 외, 2005). 이렇게 볼 때, 평가문항 서비스 사이트가 교사의 입장에서 사용하기 편리하도록 인터페이스가 설계되어 있지 않다면, 아무리 제공되는 문항의 질이 좋고 그 양이 방대하더라도 고객인 교사들에게 외면당할 수 있는 것이다. 이러한 맥락에서 사이트의 사용편의성(usability) 향상은 반드시 이루어져야 할 당면 과제라고 할 수 있다.

2. 연구의 목적

교육용 소프트웨어 및 웹사이트의 양적인 증가는 어느 정도 이루어졌고 최근 질적인 향상을 위하여 사용편의성에 대한 관심과 중요성이 증가하고 있다(나일주·허균·유병민, 2003). 그러나 아직까지는 교육용 소프트웨어와 관련된 사용편의성 검사와 관련된 실제적인 연구들이 매우 적은 실정이며, 교사를 위한 웹사이트에 대한 사용편의성 검사는 더욱이 찾아보기 힘든 상황이다. 사용편의성 평가는 평가자가 누구인가에 따라 전문가에 의한 평가와 이용자에 의한 평가로 구분한다. 전문가 평가는 빠른 시간 내에 문제점을 발견할 수 있다는 장점이 있는 반면, 이용자 평가는 전문가의 편견을 배제하고 이용자 입장에서 실제 사용상의 문

제점을 지적해 낼 수 있다는 장점이 있다. 본 평가문항 서비스 사이트의 경우 그동안 실제 이용자인 교사들의 사용이 어렵다는 지적을 꾸준히 해왔으므로, 이용자에 의한 사용편의성 평가를 통해 그 어려움을 구체적으로 밝혀 이에 대한 개선안을 제안하고자 하였다.

이에 본 연구는 이용자에 의한 사용편의성 검사를 실시하여 평가문항 서비스 사이트의 사용편의성에 관련된 문제점들을 확인하고 해결을 위한 정보를 추출하여 사이트를 개선하는데 그 목적이 있다. 이를 위해 본 연구는 첫째, 사용편의성 검사의 일반적인 평가 준거 및 검사 기법들을 개관하고, 둘째, 본 연구를 위해 설정한 준거가 검사에 어떻게 활용되었으며 또한 본 연구에 적용된 기법이 어떻게 검사의 과정을 안내하였는지 그 시행 절차를 상세히 기술하고, 셋째, 준거에 비추어 분석한 검사 결과와 이로부터 도출한 사이트 개선안을 제시하고자 한다.

교사를 위한 웹사이트의 양적 증가와 이들 사이트의 사용편의성에 대한 높아진 관심에 비해 아직까지는 사용편의성 검사와 관련된 실제적인 연구들이 부족한 실정임을 감안할 때, 실제 사이트 이용자인 교사들에 의한 본 사용편의성 검사의 결과는 향후 교사를 대상으로 하는 소프트웨어나 웹사이트의 설계에 시사점을 제시할 수 있을 것이다.

Ⅱ . 이론적 배경

1. 사용편의성(Usability)

국제표준기구(ISO: International Standards Organization)는 사용편의성(Usability)에 대하여 조작적 정의를 내리고 있는데, 사용편의성을 ‘목표된 특정 사용자들이 목표된 특정한 상황 속에서 목표된 특정 과제를 성취하는데 있어서의 효과성, 효율성, 만족감’이라고 설명하고 있다.

사용편의성은 인간컴퓨터 상호작용(HCI: Human Computer Interaction) 분야의 발전으로 그 역할이 더욱 중요하게 되었다. HCI는 여러 분야에서 그 주요 근원을 가지고 왔다. 20세기 초 산업 생산성을 높이려고 최초 시도했던 산업 공학(industrial engineering), 2차 세계 대전 당시 인간이 조정하는 기계 설계에서의 문제점들을 해결하려는 노력에서 파생된 인간공학(human factors), 19세기 말 사람의 의식과정을 실험적으로 연구한 인지과학(cognitive psychology) 등이 특히 주요한 근원을 제공한 분야들이다(Sarmento, 2004). 여기에 언급된 HCI의 주요 근원이 되는 세 가지 분야의 목적을 종합하자면 인간의 인지과정을 연구하여 기계 설계에 반영함으로써 인간과 기계 간의 부조화를 줄여, 인간과 기계 간의 부적절한 상호작용으로 인한 실수를 줄이고, 인간의 특정 과제 수행에 있어서 효과성과 효율성을 높이는 것이다.

HCI가 추구하는 주요 목적은 사용편의성에서도 추구하는 바이다. 사용자들이 어느 특정 웹사이트를 그 본래의 목적을 수행하기 위해서 사용할 때 자신이 해오던 과제 수행과정과 다른 방식의 사용법을 제공한다면, 더욱이 사용자 입장에서 그 사용법을 익히기가 어렵다면, 사용자는 그 웹사이트를 사용하면서 많은 스트레스를 받을 것이다. 또한 새로운 웹사이트의 목적은 일반적으로 현재까지 과제수행을 해오던 방식보다 더 손쉽고 빠르게 원하는 목표를 성취하는 것이기 때문에 이러한 요구를 만족시키지 못한다면 사용편의성이 뛰어난 웹사이트라고 말하기 어려울 것이다.

이와 같은 사용편의성을 제품이 잘 제공하고 있는지 평가하는 것을 사용편의성 검사라고 한다. 일반적 의미의 사용편의성 검사는 제품을 사용하는 사람이 그 제품을 사용하여 수행하고자 하는 과제를 빠르고 쉽게 수행할 수 있는가를 검사하는 방법으로 이해할 수 있다(Dumas & Redish, 1993). 그렇기 때문에 교육, 훈련, 수행 지원, 정보 접근에 필요한 상호작용적인 컴퓨터 소프트웨어를 설계·개발하는 데 있어서는 사용편의성 검사가 필수적인 것으로 여겨지고 있다. 특히, 이러한 컴퓨터 소프트웨어를 개발하는 과정에서 또는 개발 후 수정 및 보완하는 과정에서 사용편의성 검사는 유용한 정보를 제공하는 형성평가 기술로 주목을 받고 있다(Hughes, 1999; Reeves & Hedberg, 2003).

2. 사용편의성 검사의 준거

Rubin(1994)은 웹사이트를 유용하다고 말할 때는 네 가지의 요인을 포함한다고 설명하였다. 유용한 웹사이트는 ① 내용과 그 기능이 유용한 것(usefulness)이어야 할 뿐만 아니라, ② 그 사이트를 사용하여 사이트가 원래 의도한 목적을 달성할 수 있어야 하며 그 목적을 달성하는 것이 최소한의 노력으로 이루어질 수 있는 효과성(effectiveness)이 있어야 하고, ③ 그 사이트의 사용법을 배우기 쉬워야 하고(learnability), ④ 사용자가 대단히 만족(high user satisfaction)해야 한다(Rubin, 1994). 이 네 가지는 웹사이트에 대한 사용편의성 검사를 수행할 때 평가의 하위 항목이 되는 요인들이라고 할 수 있다.

Nielson(1993)은 일반적인 컴퓨터 시스템의 사용편의성 검사를 위하여 다섯 가지의 요인을 제안하였다. 그 다섯 가지 요인들은 ① 프로그램의 사용법을 배우기 쉬운가와 관련된 학습용이성(learnerability), ② 일단 프로그램을 배우면 사용자가 얼마나 생산적일 수 있는가를 다루는 효율성(efficiency of use), ③ 프로그램을 일정 기간 동안 사용하지 않다가 다시 사용하였을 때 사용법을 얼마나 잘 기억해 내는가를 말하는 기억용이성(memorability), ④ 시스템을 사용하면서 사용자가 얼마나 많은 실수를 하게 되는가와 일단 실수를 하였을 때 다시 정상 회복을 할 수 있는가와 관련된 실수와 실수로부터의 정상 복귀 용이성(Few and Noncatastrophic Errors), ⑤ 사용자가 시스템을 사용하면서 느끼게 되는 전반적인 수준을 말하는 만족 수준

(level of satisfaction) 등이다.

이 밖에 이성흠, 권숙진, 이영민(2000)은 Guillemette(1995), Lindgarrd(1994), Nielsen(1990), Reed(1992), Shackel(1991) 등의 연구를 종합하여 학습용 소프트웨어의 사용편의성 검사를 위한 범주를 다섯 가지로 제안하였다. 범주는 ① 프로그램을 처음 사용하는 학습자도 주어진 학습 과제를 쉽고 효과적으로 수행할 수 있도록 프로그램이 설계 및 개발되어야 한다는 학습가능성(learnability), ② 학습자가 특정한 학습목표를 쉽게 달성할 수 있도록 프로그램이 설계 및 개발되어야 한다는 성과 지향성(performance effectiveness), ③ 프로그램 속에 학습자의 학습 과정을 도울 수 있는 전략들을 미리 만들어 놓아야 한다는 융통성, ④ 학습자가 작은 실수를 범하더라도 쉽게 기능을 회복하여 학습을 지속할 수 있어야 한다는 실수허용 및 시스템의 완전성(error tolerance and system integrity), ⑤ 학습자가 즐겁게 프로그램을 사용할 수 있어야 하고 미적으로도 만족해야 한다는 사용자 만족도가 그것이다.

이러한 평가 요소들은 편의성 검사의 목적과 유형에 따라 선택적으로 사용될 수 있다(이성흠 외, 2000). 따라서 상기 요소들과 그 개념을 본 연구에서 사용편의성 검사의 대상으로 하는 평가문항 서비스 사이트와 관련지어 효과성, 효율성, 학습용이성, 실수허용과 시스템의 완전성, 그리고 사용자 만족도라는 다섯 가지의 준거로 정리해 보았다. 이 다섯 가지 요소들은 본 연구의 사용편의성 검사 자료를 분석하는 하나의 틀로서 사용될 수 있다.

- 1) 효과성은 Rubin(1994)이 언급한 유용성과 효과성, 그리고 이성흠 등(2000)이 언급한 성과 지향성을 종합하는 준거로서, 사이트가 해당 사이트를 통해 수행하고자 하는 과제들을 잘 수행할 수 있도록 만들어져 있는가, 또 사용자가 사이트를 통해 그 과제들을 수행하는 것이 쉬운가를 의미한다.
- 2) 효율성은 Nielson(1993)이 언급한 요인으로서, 사용법을 완전히 익혔다면 과제를 수행할 때 최소한의 작업과 노력으로 원하는 목표를 달성할 수 있도록 설계되었는가를 의미한다.
- 3) 학습용이성은 여러 문헌들에서 언급된 요인으로서 해당 사이트를 사용하는 방법을 배우는 것이 얼마나 쉬운가? 즉, 사용 초심자에서 사용 전문가가 되는 것이 얼마나 쉬운가를 의미한다.
- 4) 실수허용과 시스템의 완전성은 Nielson(1993)과 이성흠 등(2000)이 언급한 요인으로서, 사용자가 사이트 사용 중 작은 실수를 범하더라도 쉽게 기능을 회복하여 사용을 지속할 수 있는가를 의미한다.
- 5) 사용자 만족도는 여러 문헌들에서 언급된 사용자의 만족을 의미하는 것으로서 사용자가 사이트에 대해서 가지게 되는 태도를 포괄적으로 의미한다. 사이트를 사용하는 것이 사용자에게 심적 부담을 주지 않는 즐거운 일인가, 그리고 사이트의 외관에 대해 미적으로도 만족스러운가를 의미한다.

3. 사용편의성 검사의 기법들

Carstens와 Patterson(2005)은 사용편의성 검사의 방법을 Nielson(1993)과 Lazar(2001)에 바탕을 두고 수행 측정(performance measurement), 소리 내며 생각하기 프로토콜(thinking aloud protocol), 코칭 방법(coaching method), 반성적 검사(retrospective testing), 구성적 상호작용(constructive interaction), 질문지 방법(questionnaires) 등 6개의 범주로 설명하였다. 수행 측정은 해당 웹사이트나 프로그램이 목적으로 하는 과제의 수행 성과와 관련된 것들을 측정한다. 사용자에게 의해서 정해진 시간 내에 성공적으로 수행한 과제의 수, 과제를 수행하는 데 걸린 시간, 수행 중에 발생한 실수나 오류의 수, 오류에서 다시 회복하는 데 걸린 시간 등이 측정의 대상이 될 수 있다.

소리 내며 생각하기 프로토콜(think aloud protocol)은 사용자가 해당 과제를 수행하면서 단기기억 내에서 일어나고 있는 인지과정에 대하여 말하게 하여 그 사용자가 사용하는 것을 관찰하면서 동시에 사용자의 말을 통하여 해당 웹사이트나 프로그램의 사용편의성을 검사하는 방법이다(Carstens & Patterson, 2005; Reeves & Hedberg, 2003). 웹사이트를 사용하면서 인지과정을 사용자가 말로 표현하기 때문에 평가자는 평가하려는 웹사이트나 프로그램의 사용시 의도한대로 사용자가 사용하는지 또는 어떤 문제점을 겪는지 바로 관찰을 하거나 유추해 낼 수 있다.

코칭 방법은 사용자가 제품을 사용하면서 생기는 질문을 하게 하고 답변을 받을 수 있게 하여 프로그램의 문제점이 무엇인지, 어떻게 문제점을 해결할 지를 알아내는 방법이다(Carstens & Patterson, 2005). 사용자와 질문과 답변을 주고받는 과정을 통해 평가자는 제품의 디자인을 변경하거나 제공할 도움말을 구성하는 데 필요한 자료를 수집할 수 있다.

질문지법은 사용자가 테스트에 참가한 후 질문지를 통해서 사용자들로부터 직접적인 그들의 의견을 수집하는 방법이다(Carstens & Patterson, 2005). 소리 내며 생각하기 세션을 거친 후 바로 질문지를 사용함으로써 관찰로 얻어진 자료 외에 더 많은 자료를 얻을 수 있는 가장 손쉬운 방법 중의 하나이다.

반성적 검사는 소리 내며 생각하기 방법으로 사용자가 검사에 참가하고, 이 때 사용자의 발화된 언어 자료, 사용자가 보고 있는 화면, 사용자의 몸짓, 손짓 등을 녹화해 두었다가 평가자가 녹화된 영상을 사용자와 함께 보면서 사용자에게 검사 당시 왜 특정 말을 하거나 행동을 보였나 알아내는 방법이다(Carstens & Patterson, 2005). 이렇게 피험자가 컴퓨터 프로그램을 사용하여 과제를 수행하는 과정을 영상으로 기록해 둔 것을 피험자가 다시 보면서 자신의 인지과정을 회상하여 당시 무엇을 생각하고 있었는지, 무엇을 하고 있었는지, 또 그런 것들을 왜 했었는지에 대한 자료를 수집하는 기법을 자극회상기법(stimulated recall)이라고 한다. 녹화 영상을 통해 제시되는 시각적 또는 청각적 단서(cues)들이 피험자가 실제 과제수행

중 인지과정을 회상할 수 있는 자극자(stimulant)로 사용된다는 의미에서 자극회상기법이라 불린다. 따라서 반성적 검사 방법은 주로 소리 내며 생각하기 방법으로 수집한 자료를 평가자가 일차적으로 분석한 후, 자극회상기법을 활용한 인터뷰를 통해 그 분석이 올바른지 실제 사용자에게 확인하는 방법으로 많이 사용된다. 실제로 이 인터뷰를 하면서 평가자는 분석 내용을 확인할 수 있을 뿐 아니라, 자신이 미처 알아내지 못한 프로그램에 대한 문제점이나 특징을 사용자로부터 들을 수 있어 유용한 정보를 수집할 수 있다.

구성적 상호작용은 두 명의 사용자가 함께 웹사이트를 사용하여 과제를 수행하게 하고 자료를 수집하는 방법이다(Carstens & Patterson, 2005). 구성적 상호작용 방법은 과제 수행 중 사용자 간 상호작용을 통해 서로 발견하지 못한 정보를 교환하는 과정에서 나타나는 새로운 자료를 얻을 수 있다는 장점이 있다.

위 방법들 중 본 연구에서는 반성적 검사가 사용되었다. 반성적 검사를 실시하여 소리 내며 생각하기와 자극회상기법을 모두 적용함으로써 풍부한 자료를 확보하고자 하였다.

Ⅲ. 연구 절차 및 방법

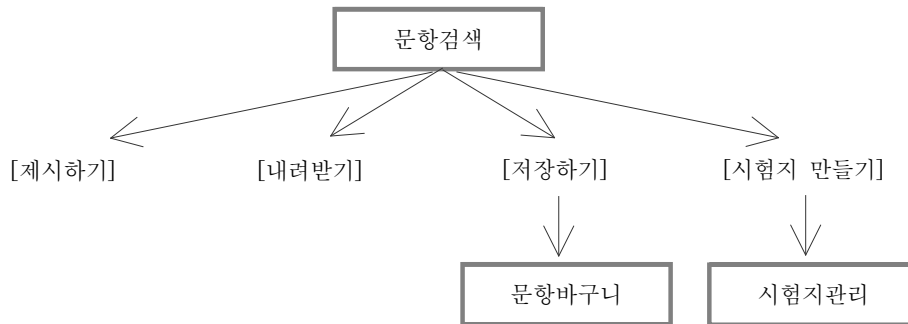
본 사용편의성 검사는 한국교육과정평가원의 평가문항 서비스 사이트를 대상으로 수행되었다. 이에 사이트의 목표 사용자인 현장 교사들을 참여자로 하였으며, 사이트의 주요 기능을 참여자가 직접 사용해 볼 수 있도록 몇 개의 과제를 개발하여 검사 도구로 활용하였다. 검사는 실험실 환경에서 진행되었으며, 그 절차에 있어서는 반성적 검사의 방법을 따랐다. 검사 자료의 분석은 사이트의 사용과 관련한 사용상의 문제점을 찾아내는 것을 목적으로 하였다. 이에 발견한 문제점을 상술한 5개 준거에 비추어 어떠한 측면에서 사용편의성을 저해하는지 알아보고, 저해 요인을 제거하는 방식으로 개선안을 제안하였다.

1. 검사 대상 : 한국교육과정평가원의 평가문항 서비스 사이트

한국교육과정평가원에서는 2002년부터 시작하여 2006년 현재까지 5년 동안 평가문항 개발 및 적용, 평가문항 서비스 사이트 구축 및 보완 등을 통해 학교 현장의 교수·학습과 평가 상황에 맞춰 평가문항을 제작하고 보급해 왔다(조난심 외, 2002; 양길석 외 2003; 윤현진 외, 2004; 윤현진 외, 2005). 그 결과 2002년 중등 수학 평가문항을 시작으로 사이트를 개설한 이래, 현재는 초·중·고 학교급 및 10개 교과에 걸쳐 총 16,811문항이 개발과 검토의 과정을 거쳐 학교 현장에 제공되고 있다.

평가문항 서비스 사이트는 문항검색, 문항바구니, 시험지관리의 3개 주 메뉴와 제시하기,

내려받기, 저장하기, 시험지 만들기의 4대 기능 및 각 기능에 딸린 하위 메뉴들로 구성되어 있다. 사이트의 구조를 간략히 도식화하면 다음과 같다.



〔그림 1〕 평가문항 서비스 사이트의 구성

사용자는 문항검색 메뉴를 활용하여 학년, 교과, 단원별로 문항을 검색할 수 있을 뿐만 아니라 문항 유형, 난이도, 성취 기준별로 상세한 검색을 수행할 수 있다. 검색을 통해 용도에 맞는 문항을 추출한 후 선택한 문항들을 ① 교실에서 프로젝터를 활용하여 스크린에 띄워 활용하거나[제시하기], ② 한글 파일로 자신의 로컬 컴퓨터에 저장한 후 변형하여 활용하거나[내려받기], ③ 사이트에서 제공되는 자신의 문항바구니에 저장해 두거나[저장하기], ④ 별도의 편집 작업 없이 자동으로 시험지로 제작할 수 있다[시험지 만들기]. 따라서 평가문항 서비스 사이트는 기존에 일반화되어 있는 인쇄 매체나 한글 파일(CD 자료) 중심의 평가문항 보급과 차별화되는 것으로서, 웹 기반으로 문항을 서비스하여 교사라면 누구나 언제 어디서나 평가문항을 제공받을 수 있게 하였다.

2. 참여자

일반적으로 사용편의성 검사는 5~6인의 피험자만 있으면 문제점의 85%를 발견할 수 있는 효과적인 방법론으로 알려져 있다(Nielson, 1993). 따라서 평가문항 서비스 사이트의 목표 사용자인 현장 교사 중에서 사이트 사용 경험의 유무 및 교직 경력의 구분에 따라 전체 사용자 집단을 대표하는 6개의 하위 집단을 구성하고, 각 집단에서 한 명의 교사를 선정하는 방식으로 총 6인의 참여자를 선정하였다.

〈표 1〉 참여한 교사의 구성

구 분		교직 경력		
		10년 미만	10년 이상~20년 미만	20년 이상
사용 경험	유	고등학교 사회 교사	초등학교 교사	초등학교 교사
	무	중학교 국어 교사	고등학교 영어 교사	고등학교 수학 교사

평가문항 서비스 사이트의 사용 경험을 집단 구성의 하나의 준거로 삼은 이유는 사용 경험에 따라 사이트 사용의 숙련도가 다르므로 인해 참여자가 발견하는 문제점이 다를 수 있기 때문이었다. 또한 교직 경력을 주요하게 본 이유는 경력의 차에 따라 교사가 본 사이트를 통해 지원받고자 하는 부분이 다를 수 있고, 그로 인해 특정 기능과 관련하여 요구 사항 및 해당 기능의 구현 수준에 대한 기대치가 다를 수 있기 때문이었다.

3. 검사 도구

사용편의성 검사에서는 참여자가 사이트를 실제로 사용하는 동안 부딪히는 문제점들을 수집해야 하므로(Nielson, 1993; Reeves & Hedberg, 2003), 참여자가 사이트를 충실하게 사용해 보는 것이 중요하다. 따라서 평가문항 서비스 사이트의 주요 4대 기능인 ‘제시하기’, ‘내려받기’, ‘저장하기’, ‘시험지 만들기’를 중심으로 수행과제를 개발하여 참여자에게 제공하였다. 참여자가 4개의 과제를 완수하기 위해서는 사이트의 주 메뉴인 ‘문항검색’, ‘문항바구니’, ‘시험지관리’ 등을 모두 사용해야 함은 물론 부수적인 그 밖의 기능들도 모두 사용하도록 과제를 구성하였다. 또한 수행과제가 참여자의 실제 사용 상황을 충실히 반영하도록 참여자의 학교급과 교과에 따라 문항검색 조건을 다르게 지정하였다. 다음 <표 2>는 참여자가 중학교 국어 교사였던 경우에 제시된 과제의 예이다.

〈표 2〉 중학교 국어 교사에게 제시된 과제의 예

구분	과제
과제 1 : 제시하기	중학교 2학년 국어과 ‘문학의 표현’ 단원에서 난이도가 보통인 3개의 평가문항을 선택하여 제시하기
과제 2 : 저장하기 (문항바구니)	중학교 3학년 국어과 ‘음운의 변동’ 단원에서 4개의 문항을 선택하여 바구니에 저장하기
과제 3 : 내려받기	중학교 1학년 국어에서 키워드 ‘문학’을 포함한 문항을 모두 찾아 한글 파일로 저장하기
과제 4 : 시험지 만들기	중학교 3학년 국어 ‘쓰기’ 영역에서 난이도가 보통인 평가문항을 찾아 이중 몇 개의 문항을 선별하여 시험지로 만들기

4. 검사 절차

본 사용편의성 검사를 시행하는 데 있어서 사용편의성 검사 시 자극회상기법을 사용한 조 일현과 김영은(2006)의 연구에서 보고한 실험실 환경과 동일한 환경을 갖추어 검사를 실시하였다. 실험실은 일반적인 동영상 촬영 용도의 스튜디오에 참여자가 평가문항 서비스 사이트를 사용하기 위한 개인용 컴퓨터를 설치하고, 참여자의 얼굴 표정을 촬영하기 위한 비디오 카메라와 참여자의 마우스 조작을 촬영하기 위한 비디오카메라 각 1대씩, 참여자의 컴퓨터 모니터를 캡처(capture)하기 위한 비디오카메라 1대, 그리고 참여자의 발화된 진술을 기록하기 위한 마이크 1대가 설치되었다.

검사의 절차는 반성적 검사(Carstens & Patterson, 2005)의 방법을 따라서 참여자가 실험실에서 소리 내며 생각하기 방법으로 검사에 참여하게 하고, 이 때 3대의 비디오카메라 1대의 마이크를 사용하여 참여자의 표정, 참여자의 손동작, 참여자가 보고 있는 화면, 참여자의 발화된 진술을 녹화 및 녹취하였다. 이후 3개의 동영상 자료와 1개의 음성 자료를 단일 영상으로 편집하였고, 이 영상을 연구자들이 보면서 1차 분석을 한 후, 다시 참여자와 함께 이 영상을 보면서 1차 분석을 바탕으로 참여자를 인터뷰하였다. 구체적인 실험 절차는 다음과 같다.

가. 사전 안내

사용편의성 실험에 앞서 참여자에게 전화와 이메일을 통해 검사의 목적과 절차에 관한 사전 안내를 제공하였다. 실험 당일에는 참여자가 사전 안내 사항을 이해하고 있는지 점검하기 위한 질의응답 시간을 3~5분 정도 가졌다. 이를 통해 참여자가 실험의 목적과 절차를 이해하고 실험 중 소리 내며 생각하기를 편안하게 할 수 있도록 유도하였다.

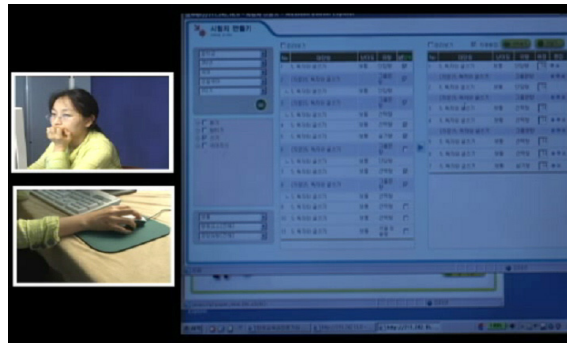
나. 본 실험

참여자가 총 6명이므로 실험도 6차례에 걸쳐 개별적으로 진행되었다. 실험의 주요 활동은 참여자가 실험실에서 사이트를 사용하여 주어진 과제를 해결하면서 소리 내며 생각하기 방법으로 검사에 임하는 것과, 실험 동안 참여자가 발생시키는 모든 행동 및 언어정보를 실험실 장비를 통해 녹화 및 녹취하는 것이다. 실험은 참여자의 사이트 친숙도와 컴퓨터 숙련도에 따라 30분에서 1시간 이상까지 소요되었다.

다. 자극자 개발 및 인터뷰 준비

실험을 통해 생성된 3개의 동영상 자료 즉, 참여자의 표정, 참여자의 손동작, 참여자가 보

고 있는 화면 자료들과 1개의 음성 자료 즉, 참여자의 발화된 진술 자료를 단일 영상으로 편집하였다. 이 편집 영상은 연구자들의 1차 분석을 위한 자료로 사용되며, 이후 자극회상기법에 의한 인터뷰 시 참여자에게 시청각적 단서를 제공하여, 참여자가 실제 과제수행 동안의 인지과정을 회상할 수 있도록 돕는 자극자로 사용되는 데 목적이 있다. [그림 2]는 자극자의 예시이다.



[그림 2] 3개의 소스를 녹화하여 구성한 자극자

이 자극자를 토대로 실험 결과에 대한 1차적인 분석을 실시함과 동시에 참여자별 인터뷰 질문들을 추출하였다.

라. 인터뷰

자극회상기법을 이용한 인터뷰를 수행하였다. 이 기법은 자극자를 이용함으로써 참여자가 과제수행 동안의 자신의 생각 등을 회상해 내도록 도와, 당시 참여자의 인지과정을 추적해 봄으로써 사이트 사용 중 사용자가 부딪히는 문제점들을 찾아내고자 하는 데 그 사용 목적이 있다. 따라서 참여자가 인터뷰 질문에 답하는 동안 자극자, 그 중에서도 해당 질문이 추출된 화면을 보면서 기억을 재생하여 답하도록 하였다. 인터뷰는 실험 시간의 1.5배가량 소요되었다. 이 인터뷰 자료는 향후 사용편의성 검사 결과 분석 시 가장 중요한 정보원이 되기 때문에 인터뷰의 전 과정은 녹화 및 녹취되었다.

5. 분석 절차

사용편의성 검사의 결과 분석은 2차에 걸쳐 진행되었다. 1차 분석은 참여자의 실험 참여 동영상에 대한 분석이었고, 2차 분석은 참여자의 인터뷰 응답에 대한 분석이었다. 본 검사를 통해 수집한 분석 대상 자료가 이처럼 질적 자료임을 감안할 때 자료 분석에 분석자의 주관

적인 판단을 완전히 배제할 수는 없으므로, Patton(2002)이 언급한 다수 분석자 검증법(analyst triangulation)을 사용함으로써 연구의 신뢰도(credibility)를 높이하고자 하였다. 따라서 분석에는 본 연구자들을 포함하여 교육공학 전공자 3인이 참여하였다. 분석은 3인의 분석자가 모여 몇 개의 참여자 검사 자료를 샘플로 함께 분석한 후, 개별적인 분석에 들어갔으며 이후 이 결과를 종합하여 하나의 분석 결과를 제시하는 방식으로 이루어졌다.

가. 1차 분석 : 실험결과 분석

연구자들이 자극자를 보면서 실험결과를 분석하였다. <표 3>은 참여자 1인에 대한 분석 중 일부를 발췌한 것이다. 칼럼은 왼쪽부터 순서대로 첫째, 참여자가 수행 중인 과제의 번호, 둘째, 동영상 즉 자극자의 재생 시간, 셋째, 실험 중 참여자가 보인 행동에 대한 기술 및 참여자가 발화한 진술에 대한 전사, 넷째, 이를 토대로 분석자가 파악한 사이트 사용상의 문제점, 마지막으로, 이 문제점에 대해 참여자의 확인을 받기 위한 인터뷰 질문이다.

〈표 3〉 실험결과 분석의 예

과제	인덱스	관찰 행동 및 발화된 진술	문제점	확인 질문
2번	0254	문항바구니로 잘못 들어갔네. → 창을 닫아버림.	네비게이션이 어려워서, 이전으로 돌아가기 위해 창을 닫고 새로 시작함.	왜 창을 닫으셨나요? 단지 않고 다른 메뉴로 옮기면 안 되었나요?
	0424	[미리보기]에서 화면에 출력되는 게 내가 선택한 문항인가? 아니네.	[미리보기]는 선택 여부에 상관없이 모든 문항을 제시함.	선택문항만 미리 보려는 의도로 [미리보기]를 사용하셨나요?
3번	0746	어떤 목표로 만들어진 문제인지... [미리보기] 안에서도 리스트 내용을 보여주면 좋은데...	[미리보기]에서는 문항리스트에서 제공되는 문항정보가 제공되지 않음.	문항정보를 [미리보기]에서도 볼 필요가 있으신가요?

나. 2차 분석 : 인터뷰 결과 분석

자극회상기법을 이용한 인터뷰를 시행하는 동안, 참여자의 응답 음성과 응답 시 표정, 응답하는 동안 모니터에 제시되었던 자극자 화면을 마이크 1대와 비디오카메라 2대를 사용하여 녹화 및 녹취한 후 자극자를 개발할 때와 같은 방식으로 단일 영상으로 편집하였다. 이 동영상을 보면서 인터뷰 결과를 분석하였다. 다음 <표 4>는 참여자 1인에 대한 분석 중 일부를 발췌한 것이다(원 자료는 엑셀 파일로 아래의 각 칼럼에 주어진 일련번호 순으로 일렬로 배열되어 있음). 1차 분석 양식에 이어서 인터뷰 분석을 추가한 것으로, 7번째 칼럼부터

시작하여 순서대로 영상 재생 시간, 질문에 대한 참여자의 응답, 참여자가 추가적으로 제시한 의견, 이를 토대로 분석자가 제안하는 잠정적 개선안을 담고 있다.

〈표 4〉 인터뷰 분석의 예

1) 자극자	2) 인덱스	3) 관찰 행동 및 발화된 진술	4) 문제점	5) 확인 질문
과제 2	0254	문항바구니로 잘못 들어갔네. → 창을 닫아버림.	네비게이션이 어려워서, 이전으로 돌아가기 위해 창을 닫고 새로 시작함.	왜 창을 닫으셨나요? 단지 않고 다른 메뉴로 옮기면 안 되었나요?
6) 인터뷰영상	7) 인덱스	8) 참여자 응답	9) 참여자 추가 의견	10) 개선안
과제 2	0148	일단 잘못했다고 생각해서 닫았고, 돌아가기 버튼도 없고 해서...그냥 창을 닫았죠.	돌아가기 버튼이 없다는 게... 사이트가 전체적으로 불편하네요. 창을 닫았다 다시 시작하면 어쨌건 검색 조건 설정부터 다시 다 해야 하니...	전체적으로 네비게이션이 편한 방향으로 재설계함.
1) 자극자	2) 인덱스	3) 관찰 행동 및 발화된 진술	4) 문제점	5) 확인 질문
과제 2	0424	[미리보기]에서 화면에 출력되는 게 내가 선택한 문항인가? 아니네.	[미리보기]는 선택 여부에 상관없이 모든 문항을 제시함.	선택문항만 미리 보려는 의도로 [미리보기]를 사용하셨나요?
6) 인터뷰영상	7) 인덱스	8) 참여자 응답	9) 참여자 추가 의견	10) 개선안
과제 2	0522	그렇죠.	[미리보기]가 정확히 무엇인지 모르겠더라고요. 물론 여러 번 해본 다음에 알았지만...	[미리보기] 기능을 보다 잘 표현하는 명칭 고려함. 선택 문항만 보는 옵션 등 기능 재설계함.

※ 위첨자로 표시된 숫자는 원래 표의 칼럼을 의미한다. 지면의 한계로 원래 엑셀 파일의 10개의 칼럼을 5개의 칼럼으로 나타내고 있다.

다. 분석 종합

참여자별 실험 및 인터뷰 결과 분석을 완료한 후 보고된 문제점들을 중심으로 유사하거나 관련이 있는 문제들을 분류해서 묶어 주고, 6인의 행동과 진술, 그리고 인터뷰 응답 등을 종합해서 문제점을 다시 기술하였다. 다음으로는 이 문제들이 상술한 사용편의성 감사의 5가지 준거 즉, 효과성, 효율성, 학습용이성, 실수허용과 시스템의 완전성, 사용자 만족도에 비추어 어떤 측면에서 본 사이트의 사용편의성에 영향을 미치는지 알아보고자 하였다. 이러한

과정을 통해 문제점에 대한 개선안을 준거를 충족시키는 방향으로 제안하고자 하였다. 다만 5가지 준거 중 사용자 만족도의 경우에는 관찰이나 인터뷰를 통해서 밝혀진 문제점들을 가지고 분석자가 판단하는 데 한계가 있었다. 여기에서는 인터뷰 시 참여자가 특별히 불편함을 밝혀 지적한 경우에 한하여 사용자 만족도가 낮은 것으로 판단하였다. 다음은 분석 종합 중 일부를 발췌한 것이다. 칼럼은 왼쪽부터 순서대로 첫째, 문제점 분류, 둘째, 문제점의 구체적 기술, 셋째, 준거 관련성, 넷째, 개선안이다.

〈표 5〉 분석 종합의 예

분류	문제점	준거 관련성	개선안
>> 문항검색 단원 설정 시 트리 (tree) 구조 사용법 파악의 어려움	(관찰) 컴퓨터에 익숙하지 않은 몇몇 사용자의 경우, 검색창에 있는 트리의 기능을 이해하지 못하여 ‘단원’ 단위의 검색을 수행하지 못하고 ‘학기’ 단위 검색을 한 후 그 결과 내에서 단원을 직접 보고 선택함. (인터뷰) 트리 구조 사용법(단원을 선택하는 방법)을 쉽게 알 수 있었다면 좀 더 검색이 용이했을 수 있지만... 어쨌거나 검색을 통해 최종적으로 원하는 것을 찾아내는 데 큰 문제는 없음.	▶효율성 저해 요인 ▶낮은 학습용이성	⇒트리 사용 방법 직관화 및 단순화 : 트리를 모두 풀어서 제공하는 것은 학습 시에는 인지적 과부하를 발생시킬 우려가 있으나, 본 사이트와 같이 검색이 중요한 경우에는 화면이 허용하는 한 풀어 제공함. 또한 트리의 상위 항목이 선택되면 자동으로 하위 항목이 선택되도록 함.
>> 메인화면 [미리보기] 연결 문제가 독립적으로 제시됨.	(인터뷰) 지문 하나에 연결된 문제들이 여러 개 있는 경우, [미리보기]로 볼 경우 연결 문제가 나오지 않고 문제 하나하나가 독립되어 나타남. 반면, 제시하기로 볼 경우에는 연결되어 나타남.	▶해당 기준 없음 : 시스템을 보다 정교화시키는 노력이 요구되는 경우임.	⇒ 지문 하나에 연결된 2개 이상의 문항은 연결되어 움직이도록 함. 한 문항만 선택해도 연결된 문항이 따라오게 하고, 따라온 상태에서 사용자가 문제를 삭제할 수 있도록 함.
>> 문항바구니 내에 문항바구니 내에 폴더 관리 기능 부재	(인터뷰) 문항바구니 안에 문항들을 분류해서 따로 보관할 수 있는 폴더를 만들 수 있으면 좋겠음. 매번 다른 이유에서 선택하고 저장한 문항들이 한데 섞여 있어서 불편함.	▶효율성 저해 요인 ▶사용자 만족도 저하	⇒문항 바구니 기능 강화 : 문항 바구니 안에서 문항 관리가 가능하도록 폴더의 생성, 이동, 삭제 등을 할 수 있는 기능을 추가함.
>>시험지 만들기 [만들기창] 배점 설정 자동 해제	(관찰) 배점 입력 후 다른 기능을 잠시 수행하고 나면 설정해 놓은 배점이 사라지고 default 상태(배점이 주어지지 않은 상태)로 돌아감. 백(back) 버튼 같은 기능이 없어서 입력해 놓은 페이지가 회복이 안 되고 처음부터 다시 입력해야 함.	▶낮은 시스템의 완전성 ▶효율성 저해 요인	⇒ 배점 입력은 한번 입력한 후 사용자가 새로운 입력을 하기 전까지는 그대로 유지되도록 수정함.

IV. 결과

상술한바 지금까지의 사용편의성 검사를 통해 다수 분석자에 의한 다단계 분석으로부터 도출한 검사 결과를 정리하면 다음과 같다. 첫째, 평가문항 서비스 사이트의 메인화면 인터페이스와 관련하여 8건의 문제점이 발견되었는데, 대부분은 문항검색 결과를 제시하거나 검색한 문항을 미리 보는 방식과 관련하여 사용자 개개인의 요구를 반영할 수 있는 융통성을 두지 않은 데서 오는 문제들이었다. 이러한 문제는 사이트의 효율성을 저하시키고 사용자 만족도를 떨어뜨리므로 사용자 편의 도모를 위해 기존 기능의 개선 및 새로운 기능의 추가가 요구되었다. 또한 인터페이스의 학습용이성이 높지 않은 문제가 있어 재설계가 요구되었다.

둘째, 본 서비스를 일종의 문제은행 서비스라고 볼 때 가장 기본이 되는 기능인 문항검색 기능과 관련하여 5건의 문제점이 발견되었는데, 학습용이성이 낮은 검색창 인터페이스로 인하여 사용자가 검색 기능을 사용하여 효과적인 검색을 수행하지 못하는 문제들이었다. 아울러 검색 기능 수행 시 시스템의 실수허용 수준이나 시스템의 완전성이 미흡한 문제도 있었으며, 검색 조건과 관련하여 오히려 사용자의 불편을 초래할만한 세부적인 항목까지를 선택하도록 요구하는 문제도 있었다. 이러한 문제들은 복합적으로 효과성을 저해하고, 효율성은 물론 사용자 만족도를 떨어뜨려 검색창 인터페이스에 대한 전면적인 재설계가 요구되었다.

셋째, 평가문항 서비스 사이트에서 문항을 검색한 후 검색 결과로 제시되는 문항 중 일부 문항을 선택하여 ‘저장하기’를 수행하면 개인의 문항바구니에 선택 문항이 보관되는데, 이 문항바구니와 관련하여 3건의 문제점이 발견되었다. 현재는 문항바구니가 문항을 단순히 저장해 두는 저장소 이외의 다른 부가적인 역할을 수행하지 못함으로 인해, 바구니에 저장되어 있는 문항들을 관리할 수 없어 문항 관리의 효율성은 물론 사용자 만족도를 저하시키는 문제들이었다. 따라서 사용자 요구를 반영하여 부가적인 기능을 추가하고, 나아가 사이트의 다른 기능, 예를 들면 ‘시험지 만들기’와 연동할 수 있도록 개선이 요구되었다.

넷째, 본 사이트에서 가장 필요한 기능인 반면 가장 사용이 어렵다고 지적된 ‘시험지 만들기’와 관련하여 11건의 문제점이 발견되었다. 문제점은 시험지를 제작하기 위한 ‘만들기창’에서 제공하는 하위 기능들이 본연의 기능을 완벽하게 소화해 내지 못하는 데서 오는 문제들과 시험지 제작 후 만든 시험지를 보기 위한 ‘시험지창’에서 편집 기능을 지원하지 않는 데서 오는 문제들이었다. 더욱이 두 창 모두 인터페이스의 설계가 직관적이지 않아 학습용이성이 낮음으로 인해 사용자가 그 기능이나 용도를 쉽게 파악할 수 없었다. 이러한 문제들은 복합적으로 효과성 및 효율성을 저하시키고 사용자 만족도를 떨어뜨렸다. 이에 전자의 경우 하위 기능들이 본래 의도한 기능을 충실히 수행할 수 있도록 프로그램의 완성도 향상이 요구되었고, 후자의 경우 사용자 요구에 부응하여 편집 기능을 추가할 필요가 있었다. 또

한 두 창 모두 직관적인 인터페이스로의 재설계가 요구되었다.

다섯째, 평가문항 서비스 사이트에서 제공하는 문항을 교사가 프로젝터를 활용하여 교실에서 형성평가 용도로 활용할 수 있도록 ‘제시하기’ 기능을 두고 있는데, 이 제시하기와 관련하여 1건의 문제점이 발견되었다. 이는 문항을 제시하는 창에서 제공하는 버튼들이 직관적이지 못하게 설계되어 있어 버튼의 기능에 대한 학습용이성이 다소 떨어지는 문제로 버튼의 재설계가 요구되었다. <표 6>은 검사 결과를 요약한 것이다.

〈표 6〉 검사 결과의 요약

문제점 요약		준거 관련성	개선안 요약
메인화면 인터페이스	<ul style="list-style-type: none"> - 전반적인 메뉴 구성의 문제 - [검색결과제시] 관련 <ul style="list-style-type: none"> 선택 누적 불가, 소팅 규칙, 전체 선택 버튼 미활용, 문항의 정체성(identity) 파악의 어려움 - [미리보기] 관련 <ul style="list-style-type: none"> 기능의 미사용, 문항 선택 불가, 문항정보 제공 요구, 전체 문항에 대한 미리보기 수행 	<ul style="list-style-type: none"> ▶ 효율성 저해 요인 ▶ 학습용이성 저해 요인 ▶ 사용자 만족도 저해 요인 	<ul style="list-style-type: none"> - 메뉴/기능 등 인터페이스 재설계 - 검색 결과가 사용자 개개인의 요구에 따라 다양하게 제시될 수 있도록 [검색결과제시] 리스트에 각종 기능 추가 - 미리보기가 사이트의 다른 4대 주요 기능들과 동급의 기능이 되도록 기능 확장·개선
문항 검색	<ul style="list-style-type: none"> - 검색 범위가 달라진 경우에도 동일한 검색창의 외관 적용으로 기능이 다름을 파악하기 어려움 - 검색 버튼 및 상세 검색의 위치 - 단원 설정 시 트리 구조 사용법 파악의 어려움 - 검색 조건 설정 자동 해제 - 제시된 검색 조건 중 일부 조건 미활용 	<ul style="list-style-type: none"> ▶ 효과성 저해 요인 ▶ 효율성 저해 요인 ▶ 학습용이성 저해 요인 ▶ 실수허용과 시스템의 완전성 저해 요인 ▶ 사용자 만족도 저해 요인 	<ul style="list-style-type: none"> - 문항 검색을 위한 검색창 인터페이스 (검색 범위에 따른 외관 변화, 각종 검색 방법들의 사용 방식, 선택 항목의 제시 위치, 트리 구조 등 포함) 재설계 - 검색 조건 선택 시 선택항목 최소화
저장하기 및 문항바구니	<ul style="list-style-type: none"> - 문항바구니 내에 폴더 관리 기능 부재 - 문항바구니에 저장되어 있는 문항에 대한 정보 표기 요구 - ‘시험지 만들기’와 연동이 되지 않음 	<ul style="list-style-type: none"> ▶ 효율성 저해 요인 ▶ 사용자 만족도 저해 요인 	<ul style="list-style-type: none"> - 문항바구니에 폴더 관리 기능 추가 등 바구니 기능 강화 - 바구니 속 문항들을 시험지로 만들 수 있도록 문항바구니에 시험지 만들기 연동
시험지 만들기	<ul style="list-style-type: none"> - [만들기창] 관련 <ul style="list-style-type: none"> 미리보기 중복 및 발견의 어려움, 배점 입력이 필수 항목임. 배점 설정 자동 해제, 수동 배점 입력, 직관적이지 못한 아이콘 디자인, 연결 문제가 독립적으로 제시됨. - [시험지 보기창] 관련 <ul style="list-style-type: none"> 시험지 제목 부여의 어려움, 시험지 편집 불가, 시험지 쪽 번호 삽입 요구, 직관적이지 못한 아이콘/버튼 설계 	<ul style="list-style-type: none"> ▶ 효과성 저해 요인 ▶ 효율성 저해 요인 ▶ 학습용이성 저해 요인 ▶ 실수허용과 시스템의 완전성 저해 요인 ▶ 사용자 만족도 저해 요인 	<ul style="list-style-type: none"> - 만들기창에 현재 포함되어 있는 하위 기능들(문항 선택, 편집, 배점 입력 등)이 제 기능을 충분히 수행하도록 프로그램의 완성도 향상 - 만들기창 인터페이스 재설계 - 시험지 편집이 용이하도록 시험지창 인터페이스(제목/쪽번호 입력, 줄 간/문항 간 간격 조정, 용지 설정 등) 재설계
제시하기	<ul style="list-style-type: none"> - [제시하기창] 관련 <ul style="list-style-type: none"> 직관적이지 못한 버튼 설계 	<ul style="list-style-type: none"> ▶ 학습용이성 저해 요인 	<ul style="list-style-type: none"> - 직관적인 버튼 설계

지금까지 검사를 통해 발견한 문제들은 크게 세 범주로 구분될 수 있다. 첫 번째는 특정 기능과 관련하여 현 상태에서는 사용자가 사용법을 파악하는 데 어려움을 느끼는 경우가 있었다. 두 번째는 특정 기능과 관련하여 해당 기능이 작동하는 수준이 미흡하여 기능을 보장할 필요성이 제기되었다. 마지막으로 현재 사이트에서는 제공하고 있지 않은 기능 또는 특정 기능과 관련한 하위 메뉴를 신설할 것이 요구되었다. 이 문제점들은 앞서 2005년도에 개선되었다고 언급한 사이트 접근성의 문제¹⁾라든가 한글 연동과 관련한 기술적인 문제²⁾(윤현진 외, 2005)에 비해서 사이트의 사용편의성에 직결되는 문제들이며 구체적인 수준으로 어려움을 기술하고 있다.

지난 4년 동안 평가문항 서비스 사이트는 조금씩 진화하여 왔다. 단순히 문항을 검색·추출·저장하는 서비스를 넘어서 현재는 웹상에서 시험지를 바로 만들어 출력하는 서비스까지 제공하고 있다. ‘시험지 만들기’ 기능은 사용자가 자신의 컴퓨터에 문항을 내려받은 후 편집 작업을 거쳐 시험지를 만드는 과정을 자동화한 것으로서, 웹상에서 별도의 편집 작업 없이 바로 시험지 형태로 만들어 줌으로써 평가 문항을 지면으로 활용하는 과정의 효율성을 높였다(윤현진 외, 2005). 이처럼 사이트의 개선은 현장 교사들의 편의를 도모하는 방향으로 진행되어 왔다. 따라서 본 검사 결과에 대한 개선안을 제안하는데 있어서도 교사들의 요구를 최대한 수용하여, 학습용이성이 떨어지는 각종 인터페이스를 재설계하고, 필요한 기능을 확장 및 개설하여 사이트의 사용편의성을 높이고자 하였다.

V. 논의 및 제언

본 사용편의성 검사는 사용자 집단을 대표하는 표집 사용자와 함께 평가문항 서비스 사이트의 사용상의 문제점을 밝히고 이에 대한 사이트 개선 방안을 제안하였다. 앞서 제시한 사용편의성 검사 결과와 관련하여 지면의 한계상 <표 6>에 담지 못한 상세한 문제점과 그에 대한 구체적인 개선안은 앞서 제시한 <표 5>의 양식으로 작성되어 이에 따른 사이트 개선작업이 진행 중에 있다. 이렇게 함으로써 평가문항 서비스 사이트의 사용편의성 향상을 통해 사이트의 질적 향상을 꾀하고자 하였다. 지금까지 반성적 검사 방법에 따라 문제점을 수집하고 증거에 비추어 개선안을 제안하는 과정을 토대로 다음과 같은 논의를 덧붙이는 바이다.

첫째, 반성적 검사 방법을 사용하면 두 가지 다른 정보원 즉, 실험실 관찰과 사후 인터뷰

- 1) KICE-TLC를 통해 평가문항 서비스 사이트를 찾아가기까지 다단계의 메뉴를 거쳐게 되어 있어 사이트 접근이 어려움.
- 2) 평가문항 서비스 사이트에서 문항 다운로드 시 한글 파일 형태로 문항을 받게 되어있으나 사이트에서 지원하는 한글 버전이 현장 학교의 실정에 맞지 않아 다운로드가 불가능한 경우가 자주 발생함.

를 통해 참여자로부터 각기 다른 종류의 진술³⁾을 얻음으로써 풍부한 자료 수집이 가능하다고 알려져 있다(Bowers & Snyder, 1990). 실제로 실험실 관찰을 통해서도 참여자의 행동 관찰과 소리 내며 생각하기에 의한 진술을 얻었는데, 이를 통해 주로 사이트의 문제점들을 발견할 수 있었다. 소리 내며 생각하기에 의한 진술은 절차 진술⁴⁾이 대부분으로, 행동 관찰이 보여주는 참여자의 오류 행동에 대해서 어떻게 그런 오류 행동을 하게 되었는지 보충적인 정보를 제공하였다. 또 다른 정보원인 인터뷰는 크게 두 가지 역할을 하였는데, 우선 실험실 관찰을 통해 연구자가 발견한 문제점들에 대해 실제 참여자로부터 확인함으로써 내적 타당도(internal validity)를 높이는 역할을 하였다(Hughes, 1999; Patton, 2002). 또한 인터뷰를 통해 얻은 참여자의 설명적 진술⁵⁾은 향후 사이트가 사용자 편의를 도모하기 위해서 어떤 기능들을 더 추가하거나 간소화해야 하는지 등과 같은 설계적인 측면의 정보를 제공하였다. 그런데 여기서 한 가지 주목할 만한 점은 참여자의 일반적인 컴퓨터 활용능력의 수준 차에 따라 두 가지 정보원 중 특정 하나가 상대적으로 정보수집에 더 유용한 경우가 있었다는 점이다. 컴퓨터 활용능력이 높은 참여자의 경우, 실험실 관찰을 통해서도 많은 정보를 얻지 못했고 반면 사후 인터뷰를 통해서 특정 기능 하나하나가 어떻게 개선되면 좋겠다는 구체적인 개선안에 관련한 많은 정보를 얻을 수 있었다. 이와 달리, 컴퓨터 활용능력이 낮은 참여자에게서는 실험실 관찰을 통해 사이트 사용의 어려움, 특히 어떤 오류로 인해 참여자가 인터페이스를 정확히 이해하고 사용하는 데 어려움을 겪게 되는지와 관련한 많은 정보를 얻을 수 있었고, 반면 사후 인터뷰를 통해서도 관찰에 대한 연구자의 해석을 확인하고 부연하는 수준의 정보를 얻을 수 있었다. 이렇게 볼 때 Hughes(1999)의 연구와 조일현과 김영은(2006)의 연구 등에서 언급한 바와 같이, 사용편의성 검사에서 보다 타당하고 신뢰성 있는 정보를 확보하기 위해서는 다양한 정보원을 활용하는 것이 바람직하다는 것을 확인할 수 있었다. 즉, 소리 내며 생각하기 방법과 자극회상기법은 피험자의 진술을 얻는 시점이 다르고 얻을 수 있는 진술의 종류도 다르기 때문에 두 방법을 연계하여 순차적으로 사용하는 반성적 검사를 차용함으로써 다양하고 풍부한 피험자 진술을 확보하는 것이 바람직하다.

둘째, 교육정보화라는 국가정책의 일환으로 학교 현장에 보급된 ICT는 2000년까지 그 인프라 구축을 완료하고, 현재까지 ICT 활용을 위한 교원 연수 실시 및 콘텐츠 보급 등 지속적인 지원하에 교수·학습 방법의 혁신에 기여하고 있다(김영애 외, 2004). 그러나 아직까지 ICT를 활용하는 교사의 컴퓨터 활용능력은 연령별, 학교급별, 과목별 특성으로 인해 그 수준

3) Bowers와 Snyder(1990)는 소리 내며 생각하기와 자극회상기법을 비교하는 연구에서 각 방법을 통해 얻은 피험자들의 진술(statement)을 분석하여 진술문의 성격을 6개로 범주화하였다(절차 진술, 절차와 설명 진술, 설명 진술, 읽기 진술, 설계 진술, 기타 진술).

4) 피험자가 단순히 자신이 무엇을 하고 있는지를 진술하는 것

5) 피험자가 절차 진술을 넘어 과제 전반에 걸쳐 자신이 사용하고 있는 전략에 대해 설명적으로 진술하는 것

에 있어서 많은 차이를 보인다. 따라서 평가문항 서비스 사이트의 개선안을 제안하는 과정에서 개선의 수준과 관련하여 분석자 간 많은 논의를 요했다. 본 사이트의 경우 전국 38만여 명의 교사를 서비스 대상으로 하는데, 어느 수준의 컴퓨터 활용능력을 기준으로 사용편의성을 고려해야 하는지 의사결정에 어려움이 있었다. 실제로 본 검사에 참여한 참여자 간에도 컴퓨터 활용능력의 수준 차에 따라서 검사의 수행 시간이 많게는 2배가량 차이가 났다. 더욱이 검사 결과 한 참여자를 통해 심각하게 지적된 문제점이 다른 참여자에게서는 큰 문제점이 아닌 것으로 파악되는 상황도 발생하였다. 그러나 본 검사의 실시배경에는 평가문항 서비스 사이트를 누구나 사용하기 쉽게 만들어 보다 많은 현장 교사들이 양질의 평가문항을 접할 수 있게 한다는 목적이 있으므로, 최소 수준의 컴퓨터 활용능력을 갖춘 사용자라도 큰 어려움 없이 사이트를 사용할 수 있어야 한다는 원칙을 세우고 이 수준의 참여자들을 기준으로 문제점을 정리하고 개선안을 제안하였다. 따라서 검사를 마친 참여자들의 사이트에 대한 전반적인 의견은 ‘사용을 기피할 정도로 사용편의성이 낮은 것은 아니다’이었음에도 불구하고 분석 결과는 많은 문제점과 많은 개선안을 담고 있다.

끝으로, 본 연구에서 다른 바와 같이 평가문항 서비스 사이트는 그동안 사용편의성과 관련한 문제점으로 인해 현장의 교사가 사용하는데 어려움이 있었던 것은 사실이다. 그러나 이 사이트가 장기간에 걸쳐 다수의 교과 및 평가전문가들이 참여하여 구축한 방대한 평가문항 데이터베이스를 갖추고 있으며, 이 문항들에 대한 질 관리 또한 현장 교사들의 참여로 지난 몇 년간 꾸준히 이루어져 왔다는 점에서는 높이 평가할 만하다. 본 연구는 평가문항 서비스 사이트의 사용편의성 향상에 직결되는 많은 문제점들과 개선안을 심도 있게 다룬 반면, 해당 사이트의 장기적인 발전 방향을 제시하려는 노력이 상대적으로 부족했던 면이 있다. 이에 연구를 마치면서 향후 평가문항 서비스 사이트의 발전 방향과 관련하여 다음과 같은 점을 제언하는 바이다.

첫째, 평가문항 서비스 사이트가 단순히 현장의 교사들에게 엄선된 문항을 제공하는 서비스에 그치지 말고, 문항의 제작 방법이나 제작 시 유의점 등에 관한 자료를 함께 제시한다면 교사의 평가전문성 향상에도 기여할 수 있을 것이다. 상술한 바와 같이 평가문항 서비스 사이트는 그 자체가 독자적인 사이트로 운영되고 있지만 더 크게는 KICE-TLC의 한 구성요소라고 볼 수 있다. 이렇게 본다면 KICE-TLC에서는 교사의 전문성 향상을 위한 다양한 자원들을 이미 서비스해 오고 있으므로, 이와 연계하여 평가문항 서비스 사이트에서도 평가문항뿐만 아니라 교사의 평가전문성 향상을 위한 다양한 자료들을 함께 제공하는 방안을 고려할 수 있을 것이다.

둘째, 현재의 평가문항 서비스 사이트는 교사가 사이트에서 검색한 문항을 활용하여 평가를 실시하는 경우, 평가 결과에 따른 교수·학습 측면에서의 지원이 미흡한 것이 현실이다. 현장의 교수·학습 활동을 충실히 지원하기 위해서는 문항 분석 자료, 정·오답에 대한 상

세한 해설과 설명, 학생들의 오류 유형과 원인 등을 제공함은 물론, 평가 결과에 따른 후속 조치로서 KICE-TLC에서 서비스하는 특정 자원과의 유기적인 연결이 이루어져야 할 것이다. 이렇게 함으로써 평가문항 서비스 사이트를 평가문항 서비스뿐 아니라 교실 수업과 관련한 모든 교수·학습 활동을 총체적으로 지원하는 종합적인 서비스 사이트로 발전시켜 나갈 수 있을 것이다.

참 고 문 헌

- 김영애 · 양재명 · 김혜숙(2004). **교육정보화 현황분석 연구**. 한국교육학술정보원.
- 김영애 · 김혜숙(2005). **2005년도 교육정보화 현황분석 연구**. 한국교육학술정보원.
- 나일주 · 허균 · 유병민(2003). WBI 완성도 향상을 위한 Usability Test 활용 방안. **컴퓨터교육학회 논문지**, 6(3), 109-117.
- 양길석 · 양명희 · 조영미 · 민병곤 · 윤준채(2003). **국민공통기본 교육과정 교과별 평가 도구 개발 연구(Ⅱ)**. 한국교육과정평가원.
- 윤현진 · 김명화 · 송미영 · 홍선주(2005). **국민공통기본교육과정 교과별 평가 도구 개발 및 적용 연구(Ⅳ)**. 한국교육과정평가원.
- 윤현진 · 양길석 · 김명화 · 강신천(2004). **국민공통기본교육과정 교과별 평가 도구 개발 및 적용 연구(Ⅲ)**. 한국교육과정평가원.
- 이성흠 · 권숙진 · 이영민(2000). 효과적인 컴퓨터 기반 학습 프로그램 개발을 위한 편의성 검사의 이론적 기초. **교육공학연구**, 16(1), 163-178.
- 조난심 · 양명희 · 조영미 · 이성흠 · 김수동(2002). **국민공통기본 교과별 평가 도구 개발 연구(Ⅰ)**. 한국교육과정평가원.
- 조일현 · 김영은(2006). 사용성 검사 기법으로서 회상 자극법(Stimulated Recall)의 타당성과 실용성에 관한 연구. **교육정보미디어연구**, 12(2), 71-75.
- Bowers, V. A. & Snyder, H. I. (1990). Concurrent versus retrospective verbal protocol for comparing window usability. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 1270-1274.
- Carstens, D. S. & Patterson, P. (2005). Usability study of travel websites. *Journal of Usability Studies*, 1(1), 47-61.
- Dumas, S. J. & Redish, J. C. (1993). *A practical guide to usability testing*. Norwood, NJ: Ablex.
- Guillemette, R. A. (1995). The evaluation of usability in interactive information systems. In J. M. Carey (ed.), *Human factors in information systems: Emerging theoretical bases* (pp. 207-221). Norwood, NJ: Ablex.
- Hughes, M. (1999). Rigor in usability testing. *Technical Communication, Fourth Quarter*, 488-494.
- Lazar, J. (2001). *User-centered web development*. Boston: Jones and Bartlett.
- Lindgaard, G. (1994). *Usability testing and system evaluation*. London: Chapman & Hall.
- Miller, G. A. (1956). The magical number seven, plus-or-minus two: Some limits on our capacity for processing information. *Psychology Review*, 63, 81-97.

- Nielsen, J. (1990). Evaluating hypertext usability. In D. H. Jonassen & H. Mandle (eds.), *Designing hypermedia for learning* (pp. 147-168). Berlin: Springer-Verlag.
- Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.
- Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Reading, MA: Addison-Wesley.
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Reed, S. (1992). Who defined usability? *You do! PC/Computing*, 5(12), 220-232.
- Reeves, T. C. & Hedberg, J. G. (2003). *Interactive learning systems evaluation*. New Jersey: Educational Technology Publications.
- Rubin, J. (1994). *Handbook of usability testing*. New York: John Wiley & Sons.
- Sarmiento, A. (2004). Issues of human computer interaction. *Information Management* 17(3/4), 22-23. Retrieved July 25, 2006 from ProQuest.
- Shackel, B. (1991). Usability-Context, framework, definition, design, and evaluation. In B. Shackel & S. J. Richardson (eds.), *Human factors for informatics usability* (pp. 21-37). Cambridge: Cambridge University.
- Shneiderman, B. (1998). *Designing the user interface. Strategies for effective human-computer interaction* (3rd ed). Reading, MA: Addison-Wesley.
- Shneiderman, B. & Plaisant, C. (2004). *Designing the user interface: Strategies for effective human-computer interaction* (4th ed.). Boston: Pearson.
- Ward, A.W. & Murray-Ward, M. (1994). Guidelines for the development of item banks. An NCME instructional module. *Educational Measurement: Issues and Practice*, 13(1), 34-39.

• 논문 접수 : 2006년 10월 10일 / 수정본 접수 : 2006년 11월 15일 / 게재 승인 : 2006년 11월 27일

ABSTRACT

A Case Study on the Usability Test for the Item Service Site for Teachers

Hyung-Kook Park(Senior Researcher, KT Advanced Technology Laboratory)

Sun-Joo Hong(Associate Research Fellow, Korea Institute of Curriculum & Evaluation)

This case study implemented a usability test to improve the Item Service Site that provides about 16,800 items for K-12 teachers. It reviewed the concept of usability and usability test, and the general criteria and methods that are generally used for usability testing. This study recruited 6 participants and employed the retrospective testing method. It discussed the usability of the site in terms of effectiveness, efficiency, learnerability, error tolerance and system integrity, and user satisfaction.

The participants performed the tasks of searching items, collecting items, classifying items, downloading items, and making/revising tests using the items while doing 'think aloud' and their behaviors and utterances were video-recorded. Through the video analysis, the researchers produced the 1st round findings and developed the interview questions. Then, the researchers used the stimulated recall to interview the participants. The interviews were conducted to get the confirmation on the 1st round findings from the participants, and to seek the additional comments about problems and improvements. These interviews were also video-recorded and analyzed. The 2nd round findings were produced based on the 1st round findings and the stimulated-recall interviews. Through the analysis, the usability problems were identified and the suggestions were devised based on the criteria mentioned above.

Key Words : usability, retrospective test, think aloud, stimulated recall, item bank