

온라인 평가 시스템을 위한 평가문항 분류 기준에 관한 연구

이재원*, 최은만**

*한국교육학술정보원

**동국대학교 IT학부 컴퓨터공학전공

e-mail: jaewon@keris.ac.kr

Jae-Won Lee*, Eun Man Choi**

*Korea Education and Research Information Service

**Dept of Computer Science and Engineering, Dongguk University

요 약

온라인 평가 시스템의 발전으로 오프라인을 통한 전통적인 시험에서 수준에 맞는 다양한 평가가 가능한 온라인 평가가 폭넓게 수행되고 있다. 하지만 평가문항에 대한 체계적인 유형 분류가 미흡하여 평가문항을 온라인 평가 시스템을 통해 표현하고 이기종 시스템간의 평가문항을 원활하게 공유하지 못하고 있다. 이 논문은 국내외 평가문항 유형 분류를 분석하고 온라인 상에서 평가문항을 효과적으로 표현하기 위한 명확한 문항 분류 기준을 제시한다.

1. 서론

「교육은 백년대계」라 하는 말은 먼 앞을 내다보고 교육에 투자하는 것이 현명한 일이라는 것을 의미함과 동시에 교육의 효과성 측정이 어렵고 또한 중요함을 내포하고 있다. 교육의 효과를 측정하는 대표적이고 전통적인 방법이 시험이다. 시험에 사용하는 시험지는 교육의 효과성 측정뿐만 아니라 교육을 수행하기 전에 피평가자의 수준을 측정하는 도구로 오랫동안 이용하고 있다. 시험지는 세부적으로 피평가자를 평가하기 위한 문항으로 구성되어 있기 때문에, 효과적인 평가를 위해서는 올바른 문항을 만드는 것이 중요하다.

온라인 평가 시스템은 전통적인 평가 절차를 온라인 상에서 구현한 시스템이다. 온라인 평가 시스템에서도 효과적인 평가를 실시하기 위해서는 피평가자에게 올바른 평가문항을 적절하게 제시하는 것이 중요하다.

체계화된 평가문항의 분류와 정교한 메타데이터 모델의 설계는 온라인 평가 시스템에서 문항을 효율적으로 표현할 수 있게 한다. 하지만, 평가문항을 분류하는 기준이 명확하지 않고 문항에 대한 정교한 메타데이터 모델이 없기 때문에, 온라인 평가 시스템에서 문항을 표현하는 방식이 비효율적인 실정이다.

따라서, 이 논문에서는 온라인 평가 시스템을 위한 평가문항의 분류 기준을 명확하게 제시하고 문항

에 대한 정교한 메타모델을 제시한다. 이를 위해서 먼저 2장에서 국내외의 관련 연구 활동을 분석하여 문제점을 분석한 후에 3장에는 개선한 분류 기준을 제시하고 4장에서 제시한 분류 기준을 이용하여 기존 메타데이터 모델의 개선안을 제시한다.

2. 관련 연구 활동

온라인 평가를 위한 국내 문항과 국외 문항의 유형은 서로 구별할 수 있는 명백한 차이점을 가지고 있다. 국내에서는 문항을 워드프로세스의 다양하고 강력한 기능을 이용하여 문항을 개발하고 있는 반면에, 국외에서는 워드프로세스의 기능에 의존하지 않고 독립된 문항 저작 도구에서 다양한 기능을 제공하고 있다. 그래서, 국내의 문항 분류 기준은 저작도구는 워드프로세스의 다양한 기능을 제공하고 있는 반면에 국외의 문항 저작도구는 웹문서를 작성하는 정도의 기능만을 제공하고 있다.

2.1. 해외 평가문항

Kathleen은 선택 항목의 제약도와 복잡도를 기준으로 이러닝 평가 서비스를 위한 평가 문항을 [그림 1]과 같이 28가지 유형으로 분류하였다.[1] 28가지는 선택 항목의 유형에 따라 분류한 것이며, 문항을 평가자나 개발자 관점에서 피평가자의 반응을 분석한 결과이다.

IMS GLC¹⁾는 문항을 보다 개발자 관점에서 피

평가자의 반응을 분석하여 3가지 유형으로 분류하였다.[2] 3가지 유형은 단순문항, 복합문항, 적응문항으로서, 피평가자와의 상호작용이 하나인 문항을 단순문항이라 하고 피평가자와의 상호작용이 두 개 이상인 문항을 복합문항이라 하고, 첫 번째 질문에 대한 피평가자의 반응에 따라 새로운 질문이 나타나는 문항을 적응문항이라고 정의한다.

서에 기술된 것처럼 24가지 유형간의 중복을 인정하고 있는데, 원인을 분석해 본 결과 두 가지 평가문항 분류의 기준이 적용되었기 때문임을 확인할 수 있었다. 즉, 24가지 문항 유형은 문항 개발자 관점의 선택항목 유형이란 기준과 피평가자 관점의 답안 작성 유형이란 기준을 적용하였다.

<표 1> 24가지 국내 평가문항 유형 분류

		Intermediate Constraint Item Types						
		Most Constrained			Least Constrained			
Less Complex	Fully Selected	1. Multiple Choice	2. Selection/Identification	3. Reordering/Rearrangement	4. Substitution/Correction	5. Completion	6. Construction	7. Presentation/Portfolio
		More Complex	Fully Constructed	1A. True/False (Ikadyna, 1994, p.54)	2A. Multiple True/False (Ikadyna, 1994, p.55)	3A. Matching (Istaitieh, 1998, p.234; Ikadyna, 1994, p.55)	4A. Interlinear (Ikadyna, 1994, p.65)	5A. Single Numerical Constructed (Parshall et al., 2002, p.87)
		1B. Alternate Choice (Ikadyna, 1994, p.53)	2B. Yes/No with Explanation (McDonat, 2002, p.110)	3B. Categorizing (Bennett, 1993, p.44)	4B. Sore-Finger (Ikadyna, 1994, p.67)	5B. Short-Answer & Sentence Completion (Istaitieh, 1998, p.237)	6B. Figural Constructed Response (Parshall et al., 2002, p.87)	7B. Demonstration, Experiment, Performance (Bennett, 1993, p.45)
		1C. Conventional or Standard Multiple Choice (Ikadyna, 1994, p.40)	2C. Multiple Answer (Parshall et al., 2002, p.2; Ikadyna, 1994, p.63)	3C. Ranking & Sequencing (Parshall et al., 2002, p.2)	4C. Limited Figural Drawing (Bennett, 1993, p.44)	5C. Close-Procedure (Istaitieh, 1998, p.242)	6C. Concept Map (Shavelson, R.L., 2001; Chung & Baker, 1997)	7C. Discussion, Interview (Bennett, 1993, p.45)
		1D. Multiple Choice with New Media (Distractors) (Parshall et al., 2002, p.87)	2D. Complex Multiple Choice (Ikadyna, 1994, p.57)	3D. Assembling Proof (Bennett, 1993, p.44)	4D. Bug/Fault Correction (Bennett, 1993, p.44)	5D. Matrix Completion (Istaitieh, S., 2002, p.225)	6D. Essay (Page et al., 1995, 561-565) & Automated Editing (Istaitieh et al., 2001, pp.1-66)	7D. Diagnosis, Teaching (Bennett, 1993, p.4)

[그림 1] Kathleen에 의한 28가지 유형 분류

Mark J. Girel은 평가문항을 온라인 상에서 효과적으로 표현하기 위한 속성계측모델과 문항모델을 제시하였다.[3][4] 그러나, 수학과목의 특정 분야에 대해서만 적용한 모델이라는 한계를 가지고 있다.

2.2. 국내 평가문항

한국교육과정평가원은 연구를 통해 평가문항 유형은 총 4가지로 분류하고 있다.[5] 이 연구에서는 4가지 유형은 선택형, 단답형, 서술형, 수행형이며, 선택형으로 세부적으로 정답형과 다답형으로 구분하고 있다. 하지만, 이와 같은 4가지 평가 문항 분류는 피평가자 관점에서 답안을 작성하는 기준을 적용하였기 때문에, 개발자 관점에서 온라인 평가 서비스를 개발하여 활용하는데 적용하기 어렵다.

개발자 관점에서 국내 평가문항을 분류하고자 한 시도로서 한국교육과정평가원의 평가문항을 추출하여 결과 문항 샘플을 추출·분석한 연구가 있었다. 연구에 따르면 평가문항의 유형을 <표 2> 같이 24가지로 분류하고 있다.[6] 그리고, 24가지 평가문항의 유형은 선택형, 단답형, 서술형, 수행형의 4가지 분류 유형을 포함하면서 보고서형, 실기형, 서술·논술형 등의 20가지 유형이 추가되어 있다. 그러나, 보고

평가 유형	설명 (답지의 요구에 따른 분석)
선택형	·단일, 이중, 다중 선택형
단답형	·단어나 구 중심의 정답 입력
논술형	·주장과 근거가 포함되도록 함
서술형	·주장과 근거가 포함되기 보다는 지식적인 것을 기술하는 수준임
수행형	·실제 과제를 수행하듯이 정답 작성
보고서형	·실험이나 실습을 마치고 작성하는 보고서 형태의 정답지
실기형	·실기의 결과에 대한 정답지
서술·논술형	·일부는 주장과 근거가 포함되며, 일부는 지식적인 것이 기술됨
연결형	·상호 관련이 있는 것을 연결 함
선택형·관찰	·제시되는 문항을 관찰하고 선택 함 (듣기, 보기 등의 예시 문항)
체크리스트형·서술형	·체크도 할 수 있고, 지식적인 것을 서술하기도 함
체크리스트형	·단순히 체크하도록 함
작품분석형	·만들어진 모형이나 작품을 보고 분석하도록 함
실습·보고서형	·실습한 결과 보고서를 작성하는 답지
서술형·조사보고서형	·지식적인 것에 대한 서술과 함께 조사한 것을 보고서로 작성할 수 있는 답지
실습형	·실습을 요구하는 답지
연구보고서형	·개인이 조사하고 연구한 결과를 기록하도록 하는 답지
보고서·실습·실습형	·보고서 작성, 실험·실습이 동시에 있는 답지
실기형 (듣고 대답하기)	·듣고 대답하기
실기형 (정보의 공백 활동)	·정보의 공백을 이용한 정답 반응
실기형(조사하기)	·조사하기
감상일지 및 관찰형	·감상일지를 작성하거나 관찰하는 답지
실기·관찰·학습지형	·실기·관찰·학습지가 결합된 문항
에세이형	·에세이를 작성하도록 하는 답지(창작)
NULL(비분류 문항)	·아무 유형에 포함되지 않은 독특한 문항

위와 같이 국내에서 평가문항에 대한 유형 분류는 피평가자 관점을 기준으로 하며 명료한 개발자 관점의 기준에서 정의한 문항 유형의 제시가 부족한 실정이다.

3. 2단계 평가문항 분류 기준

국내 평가문항 유형 분류의 독특한 점은 첫째로 답안의 형태가 평가문항의 분류에 중요한 기준인 점이고, 둘째로 평가문항 분류 기준의 세분화가 명확

1) IMS GLC: IMS Global Learning Consortium

하게 완료되어 않았다는 점이다. 전자와 같은 답안의 형태에 따른 평가문항의 분류는 국내에서 오랫동안 견지된 기준으로서 전통적인 시험지를 구성하는 문항을 구분할 때 유용하였지만, 항목의 선택 유형을 보다 더 다양하게 할 수 있고 피평가자의 상호작용이 가능한 온라인 환경에서 문항을 표현하기 위해서는 미흡한 점이 있다. 따라서, 전자의 미흡점을 개선하고자 국내에서도 선택 항목의 유형과 피평가자와의 상호작용에 따른 문항 분류를 시도한 바 있으나 전자의 전통적인 분류 기준과 혼합되어 문항을 세분화할 수 있는 분류 기준을 제시하지 못하였다.

전자와 후자의 문제점을 극복하기 위해서 <표 2> 과 같이 답안 형태에 의한 기준과 선택 항목 형태에 의한 기준을 2단계로 구분하는 방안을 제시한다.

<표 2> 2단계 평가문항 분류 기준

답안 유형 (1단계)	선택 항목 유형 (2단계)
선택형	CheckBoxList
	RadioButtonList
	ImageMap
	DropDownList
	CheckBoxList
	RadioButtonList
	ImageMap
	DropDownList
	Combo Box
	CheckBoxList
단답형	RadioButtonList
	ImageMap
	TextBox
	Drag & Drop
서술형	Ranking
	MultiTextBox
수행형	Fileupload
	Drag & Drop

그리고, 2단계 평가문항 분류 기준을 Kathleen에 의한 28가지 분류 유형에 적용하면 [그림 2]와 같다.

1. Multiple Choice	2. Select One / Multiple	3. Responding / Response	4. Select Multiple / Composite	5. Composite	6. Constructive	7. Process / Problem
14. True / False	25. Multiple True / False	34. Matching	41. Interface	51. Step / Multiple	61. Open-Ended / Multiple Choice	71. Project
18. Answer Choice	28. Test / Multiple / Embedded	38. Categorizing	48. Soap-Flager	58. Short-answer & Sentence Completion	68. Filled Constructed Response	78. Demonstration / Expertise
10. Constructive / Standard	20. Multiple Answer	30. Ranking & Sequencing	40. Limited Free / Drag	50. Close-Ended	60. Conceptual	70. Discussion / Interview
11. Multiple Choice / Multiple / Multiple	21. Composite / Multiple Choice	31. Reasoning / Proof	41. Bug Fix / Correction	51. Match / Composite	61. Essay / Constructed	71. Diagnosis / Teaching



[그림 2] 28가지 유형에 2단계 분류기준 적용

2단계 분류 기준은 답안 유형 기준과 선택 유형 기준의 연관성을 두면서도 독립성을 제공하기 때문에 [그림 2]와 같이 문항에 대한 그룹핑과 세분화를 명확하게 할 수 있다.

4. 평가문항 모델

4.1. 평가문항 메타데이터 모델

온라인 평가 시스템에서 효과적으로 평가문항을 표현하기 위해서는 평가문항의 분류 기준과 함께 평가문항에 대한 메타데이터 모델도 필요하다.

한국교육과정평가원의 연구에서는 평가문항의 메타데이터 표준안을 <표 3>과 같이 정의하고 있다.^[5]

<표 3> KICE의 평가문항 메타데이터 모델

메타명	설명
1. 파일명	문항 관리 차원
2. 학교급	5 단계 교육급 (초, 중, 고1, 고2(선택), 고3(선택))
3. 학년	해당 학년
4. 학기-단계	1학기, 2 학기 (단계형 교과목의 경우에만 일부 적용)
5. 교과(과목)	예) 국어, 영어, 수학, 과학, 사회
6. 교과서	국어만 해당 예) 국어, 국어생활
7. 영역	과목별로 지정한 영역(교과내용 영역) 예) 듣기, 쓰기, 읽기, 말하기, 지리, 역사, 생물, 지구과학 등
8. 대단원	과목의 대단원
9. 소단원	과목의 소단원
10. 내용요소	평가 문항의 내용
11. 행동요소	각 교과마다 필요한 행동 요소
12. 형식문항유형 (답안유형)	선택형, 단답형, 서술형, 수행형 등의 문항의 형식에 따른 분류
13. 내용문항유형 (선택문항)	예) 수행형 (실기형, 보고서형, 창작형 등 내용 수행에 따른 분류) 문항의 내용에 따른 분류
14. 난이도	어려움, 보통, 쉬움 (상 중 하)
15. 검토자	출제자 책임 의식을 위한 문항 검토자
16. 검토일	update
17. 문항	문항내용 (문항으로 통일)
18. 정답 및 해설	정답 또는 예시답안 및 해설
19. 채점기준	서술형, 수행형 등을 위한 채점 기준
20. 유의사항	문항 사용과 관련된 기타 내용
21. 지문	문항의 지문
22. 주제어 검색	주제어 검색을 지원
23. 문항 출처	출제자 책임 의식을 위한 문항의 출처

<표 3> 23가지 메타데이터 요소 중에서 문항을 분류하는 유형에 대한 요소는 「12. 형식문항 유형」 과 「13. 내용문항 유형」 이다. 그리고, 형식문항 유형은 2단계 분류 기준의 답안 유형과 동일하며, 내용문항 유형은 2단계 분류 기준의 선택 문항 유형과 유사하다. 그리고, 메타데이터 표준안에는 내용문항 유형의 구분 기준이 명확하지 않다. 하지만, 온라인 평가 시스템에서 문항을 효과적으로 표현하기 위해서는 선택 문항 유형이 세분화되어야 하므로, 「13. 내용문항 유형」 을 2단계 분류 기준의

「선택 항목 유형」으로 <표 4>와 같이 대체할 수 있다.

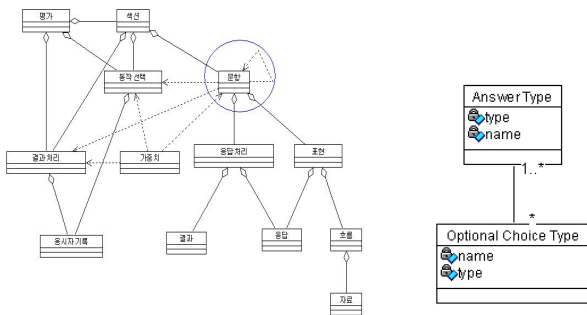
<표 4> 대체된 평가문항 메타데이터 모델

메타명	설명
12. 답안 유형	선택형, 단답형, 서술형, 수행형
13. 선택 항목 유형	ex) MultipleText, FileUpload (서술형)

4.2. 평가문항 메타모델

메타모델은 모델링 언어의 모델이다.[7] 그리고, 모델은 복잡한 대상을 단순화하여 대상의 이해와 대상의 조작을 돕는다. 평가문항 메타모델은 다양한 평가문항을 단순화하여 평가문항의 표현과 조작을 돕는다.

온라인 평가문항 시스템을 메타모델로 표현하면 [그림 3]과 같다. 2단계 평가문항 분류 기준은 [그림 3] 메타모델의 문항 영역에 해당하며 [그림 4] 메타모델과 같이 세분화된다.



[그림 3] 온라인 평가문항 시스템 메타모델 [그림 4] 평가 메타모델

평가문항 메타모델은 개발자가 온라인 평가 시스템 상에서 기존의 평가문항을 적용하거나 새로운 평가문항을 개발할 때 체계적이고 효과적인 적용과 개발이 가능하다.

5. 결론

시험으로 대표되는 전통적인 평가 절차가 IT 기술의 발전에 따라 온라인 상에서 구현되고 있다. 그리고, 온라인 상에서는 전통적인 평가에서 제공하지 못한 다양한 선택 문항을 제시하고 다양한 상호작용을 구현할 수 있다. 따라서, 전통적인 시험지 내의 문항을 분별하는 기존의 분류 기준으로 온라인 상에서 제공하는 문항을 분별하는데 한계가 있다.

이런 한계를 극복하기 위해서 이 논문에서는 관련 연구를 바탕으로 답안 유형과 선택 항목 유형의 연관성과 독립성을 유지하는 2단계 분류 기준을 제시

하였다. 그리고, 평가문항 메타데이터 모델에 2단계 분류 기준을 적용하는 개선 모델을 제시하였다. 온라인 평가 시스템에서 2단계 분류 기준의 개념으로 문항을 표현한다면 보다 효율적으로 표현할 수 있다.

앞으로 온라인 평가 시스템에 2단계 분류 기준과 개선한 메타데이터 모델의 적용단계를 세분화하고 절차를 구체화하여 객관적인 효과성을 측정해야 한다.

참고문헌

[1] Kathleen Scalise, Bernard Gifford(2006), "Computer-Based Assessment in E-Learning: A framework for Constructing 'Intermediate Constraint' Questions and Tasks for Technology Platforms, Volume 4, Number 6·June 2008. JTTLA.

[2] Steve (2006), "IMS GLC QTI Implementation ver 2.1 public draft Specialization", <http://www.imsglobal.org>, IMS GLC.

[3] Mark J. Gierl, Changjiang Wang, Jiawen Zhou(2008), "Using the Attribute Hierarchy Method to Make Diagnostic Inferences about Examinees' Cognitive Skills in Algebra on the SAT", Volume 6, Number 6·February 2008, JTTLA.

[4] Mark J. Gierl, Jiawen Zhou, & Cecilia Alves(2008), "Developing a Taxonomy of Item Model Types to Promote Assessment Engineering", Volume 7, Number 2·December 2008, JTTLA.

[5] 조난심 외(2005), "자기 학력 진단 서비스 시스템 구축 운영-교과별 평가문항 재가공 및 문항 DB 구축", 연구보고 CRI 2005-1-6, 한국교육과정평가원.

[6] 강신천 외(2006), "평가문항 상호운용성 표준화 방안 연구", 연구보고 KR 2006-11, 한국교육학술정보원.

[7] Stephen J. Mellor, Kendall Scott, et al(March 2004), "MDA Distilled - Principles of Model Driven Architecture", ISBN 0-201-78891-8, Addison-Wesley.