

다자 중심적 사용성 평가 방법의 구조화 과정*

김정룡*** · 이하연** · 최영철***

The Structuring Process of Multi-Centered Usability Evaluation Method

Jung-Yong Kim***, Ha-Yeun Lee**, Young-Cheol Choi***

ABSTRACT

In order to minimize a biased view of evaluators in usability test, a multiple evaluation method so called Multi-Centered Usability Evaluation Method(MCUEM) was suggested and structured in this study. To use MCUEM, the existing evaluation methods were re-categorized into 'developer and evaluator-centered', 'evaluator-centered', 'user-centered', 'evaluator and user-centered', and the 'developer, evaluator, and user-centered methods. Furthermore, in order to verify the relevance of MCUEM, the usability of mobile phone was tested in this study. In results, it was found that different evaluation methods generated greatly different results, that showed the risk of using a particular usability method due to different perspectives of evaluators. At the same time, it was found that MCUEM could greatly minimize such biases. Therefore, it was concluded that using MCUEM could help the usability engineer design a test protocol with a minimal risk of bias due to different perspective of evaluators.

Keyword: Usability, Multi-Centered Usability Evaluation Method(MCUEM), Mobile phone

1. 서 론

사용성의 개념적인 발전과 더불어 수십 여종의 평가 방법이 제안되었다. 그리고 이러한 평가 방법들은 각각 다른 특징과 관점을 가지고 있어 그 선택에 따라 평가 결과가 중요한 영향을 받게 되었다. 이에 따라 평가 방법에 대한 분류와 장단점을 분석하는 연구가 이루어졌다.

Nielsen(1995)은 인터페이스를 평가하는 기본적인 네 가지 방법을 자동화된(automatic) 방법, 실증적(empirical) 방법, 구조화된(formal) 방법 그리고 비구조화된(informal) 방법으로 장점과 단점을 기술하고 사용주기 단계에 따라 이를 구분하였다. 그리고 Hom(1998)은 평가 방법을 사용자

인터페이스에 대하여 평가하는 검사 방법(inspection method), 사용자가 시스템이나 프로토타입 등을 사용하면서 작업을 수행하면서 알아보는 실험적 방법(testing method), 평가자가 대화를 나누거나, 질문을 하여 사용자의 선호도, 요구사항 및 시스템의 이해도 등에 대한 정보를 얻는 질문 방법(inquiry method)으로 분류하고, 그에 대한 세부평가 항목을 제시하였다. Perlman(1996)은 검사 방법(inspection), 관찰과 비디오 녹화 방법(observational skills and video), 프로그램 방법(program instrumentation), 그리고 설문조사 방법(questionnaires)으로 분류하였다.

일반적으로 사용성 평가 방법의 선택은 예산, 기간, 평가를 실시할 때의 시스템이나 제품의 설계 단계, 이용 가능한 사용자의 수 및 사용성 평가자의 경험에 영향을 받는다(임

*본 연구는 한양대학교 과학기술분야 특별교비 장학금 수혜자의 연구 결과를 토대로 작성된 것임.

LG전자 MC사업본부, *한양대학교 산업공학과

교신저자: 김정룡

주 소: 426-791 경기도 안산시 상록구 사1동 1271, 전화: 031-400-4276, E-mail: jykim@eml.hanyang.ac.kr

치환, 1996)고 이해되고 있어, 위에 언급한 측정 방법의 차이를 정확하게 이해하고 이를 적절히 사용하여야만 충분한 사용성 정보를 추출할 수 있다.

그러나, 다양한 평가 방법이나 복수의 평가자가 참여하여 사용성 평가를 하는 경우에도 평가의 객관성이 자동적으로 보장된다고 볼 수 없다. 쉬운 예로, 제품을 제조한 사람이 직접 다양한 평가 방법을 동원하여 사용성을 평가하는 경우나, 평가 전문가가 실제 제품 사용경험이 충분치 않은 상태에서 다양한 평가 방법을 동원하여 평가할 경우에는 설령 복수의 평가 방법을 사용했다 할지라도, 최종 사용자의 관점의 차이에 의해 최종 의견이 달라질 수 있다. 한편, 아무리 다양한 복수의 사용자를 참여 시켜 다양한 평가를 한다 하더라도 전문성의 결여로 중요한 사용성의 결함을 발견하지 못하는 경우도 발생한다.

그러므로 본 연구에서는 사용성 평가에서 흔히 발견할 수 있는 평가 결과의 편향성을 어떻게 하면 최소화 할 수 있을까 하는 데 초점을 두고 이를 해소하기 위한 방법을 구조화하기 위한 방법을 제시하였다. 다시 말해, '평가의 역할을 맡은 주체의 주관적 관점에 따른 결과의 편향성'을 극복하기 위하여 본 연구에서는 '다자 중심적 사용성 평가 방법(Multi-Centered Usability Evaluation Method: MCUEM)'을 구조화하여 제안하였다. 이러한 방법은 일반 사용성 평가자나 선행 연구에서 개념적으로 사용되었거나 일부 방법에 적용되었을 수 있으나, 이를 구체적으로 명명하고 용어를 정의하고, 방법을 구조화한 연구 결과를 찾을 수 없어, 본 연구에서 이를 시도하게 되었다. 또한 다자 중심적이지 않은 사용성 평가를 하는 경우 어떠한 편향성을 보일 수 있는지, 그와 동시에 어떻게 이를 극복할 수 있을지를 실험을 통하여 밝혀 보았다.

2. 연구 방법

2.1 평가 주체의 정의

'다자 중심적 사용성 평가 방법(MCUEM)'에서 제품을 평가하는 '주체'에 대한 개념적 정의를 다음과 같이 정리하였다.

첫째, 제품을 생산하면서 제작자의 입장에서 그 제품의 사용성이 좋은지를 평가하는 경우를 '개발자(developer)'가 주체가 된다고 정의한다. 예로서, 제조 현장에서 제조업자가 사용성 평가를 담당하는 경우나, 독립적으로 최종 제품까지를 책임지는 소프트웨어 프로그래머들이 이에 해당한다고 볼 수 있다. 둘째, 시제품 또는 완제품을 중심으로 사용성 평가 전문가의 의견을 묻는 경우를 '전문 평가자(evaluator)'가 주체가 된다고 정의한다(이후 '평가자'로 약칭). 예로서,

제품에 대한 평가를 인간공학자를 포함한 사용성 평가 전문가가 담당하는 경우가 이에 해당된다고 볼 수 있다. 셋째, 중간 또는 완제품에 대해 소비자를 직접 섭외하여 평가하는 경우를 '일반 사용자(user)'가 주체가 된다고 정의한다(이후 '사용자'로 약칭). 예로서, 대부분의 제조회사에서 소비자를 대상으로 실시하는 제품 평가가 이에 해당한다고 할 수 있다. 다시 말해, 평가의 주체는 평가 결과에 심리적인 영향을 주도적 또는 실질적으로 미치게 되는 사람을 의미하고, 단순히 시간적인 참여나 평가 작업의 노동 분담정도가 크다는 이유만으로 평가의 주체로 정하지 않는다.

2.2 분류된 평가 주체의 새로운 조합

평가의 주체를 3부류로 정의하였으므로 이에 따라 평가 주체별 평가 방법이 3종류로 나뉘어 질 것으로 예상하였으나, 실제 사용되는 평가 방법에는 2개 이상의 복수 주체가 참여하는 경우가 있어, 이러한 경우를 모두 고려하였다. 그 결과, 다양하게 조합된 평가 주체를 찾아내었고 이에 대한 설명은 다음과 같다.

1. [개발자+평가자]가 주체: 우선 개발자 단독으로 전문적인 사용성 평가를 실시하는 경우는 기존 연구의 예에서 발견하지 못했다. 그러므로 개발자 단독의 평가를 분류에서 제외하였다. 대신, 본 연구에서는 전문 평가자가 설계 제작한 도구를 가지고 개발자가 직접 평가를 실시하는 경우를 발견하였고 이 때, 두 개의 관점이 동시에 평가 결과에 중요한 영향을 미칠 수 있다고 분류하여[개발자+평가자]를 주체로 분류하였다.
2. [평가자]가 주체: 많은 평가 방법이 이에 해당하며, 제품의 구조와 기능에 대해 이해가 깊은 전문가가 주관적 또는 객관적인 도구를 사용하여 주도적으로 평가를 하게 되는 경우이다.
3. [사용자]가 주체: 이 경우에도 전문 평가자의 역할이 있기는 하나, 주로 평가를 진행하는 역할에 한정되며 제품에 대한 의견은 사용자가 주도적으로 표현하게 되는 경우이다.
4. [평가자+사용자]가 주체: 사용자가 중심이 되어서 평가 하되 그 조건과 상황을 전문 평가자가 구체적으로 관여하고 통제함으로써, 전문 평가자의 관점과 태도가 결과에 주요한 영향을 미치는 경우이다.
5. [개발자+평가자+사용자]가 주체: 복합적 검사(Pluralistic Walkthrough)와 같은 경우에는 개발과 평가에 참여한 모든 관련자가 동시에 참여하는 과정이 담겨 있어 매우 예외적이기 때문에 따로 분류하였다.

2.3 평가 주체에 따른 기존 사용성 평가 방법의 재분류

위에서 분류된 평가 주체 구분에 따라 기존의 사용성 평가 방법이 본 연구에서 연구자에 의해 재분류 되었고 그 내용은 다음과 같다.

Guideline Checklist-1(특정 설계 지침에 따른 검사): 사용성 전문 평가자들이 특정 제품군을 대상으로 제작한 가이드라인을 개발자가 독자적으로 사용할 수 있다는 점 때문에 주체를 [개발자+평가자]로 분류.

Heuristic Evaluation(경험적 검사): 평가자들의 경험에 의한 직관적 사고를 통하여 제품의 기능과 형태를 검사하는 방법으로 다수의 일반 소비자를 평가자로 하지 않고 선택된 전문가를 중심으로 하므로 주체를 [평가자]로 분류.

Cognitive Walkthrough(인지적 검사): 정보인지처리과정에 대한 이해와 경험을 가지고 있는 전문 평가자가 제품 사용 시 사용자가 겪을 인지적 과정을 예측하는 방법이므로 주체를 [평가자]로 분류.

Guideline Checklist-2(일반적 설계 지침에 따른 검사): 일반적 설계 지침을 특정 제품에 적용하기 위해서는 체크리스트를 특정 제품에 응용할 수 있는 정도의 전문성을 갖춘 평가자이어야 하므로 주체를 [평가자]로 분류.

Standard Inspection(표준 검사): 특정 제품 또는 인터페이스에 대해 제안된 설계 표준을 제대로 따르고 있는지를 전문가가 평가하므로 [평가자]로 분류.

Consistency Inspection(일관성 검사): 같은 개발 체계를 가진 복합적인 제품의 일관성을 전문 평가자가 집중적으로 검토해야 하므로 [평가자]로 분류.

Feature Inspection(기능검사): 제품의 기능을 중심으로 평가하며 사용자가 과제를 수행하는 동안에 이를 관찰하여 기능을 검사한다. 이때 사용자가 참여하기는 하나 평가의 심리적 주체가 평가자이므로 [평가자]로 분류.

Expert Review(전문가 평가): 문자 그대로 충분한 경험을 가진 전문가 수준의 사람이 제품에 대해 형식에 얽매이지 않고 평가하는 방법이므로 [평가자]로 분류.

Thinking Aloud(사고 발생법): 일반 사용자들이 제품을 사용하면서 발생하는 생각과 느낌을 말로 표현한 것을 정리하는 방법이므로 전문 평가자는 정확한 기록자의 역할을 하므로 주체를 [사용자]로 분류.

Co-discovery Method(동시 발생법): 두 명의 일반 사용자가 제품을 사용하면서 서로 도움을 받는 상황을 연출하고 이 때의 생각을 표현하게 되므로 평가자의 역할은 기록자의 범위를 벗어나지 않으므로 주체를 [사용자]로 분류.

Performance Measurement(수행도 측정): 지정된 과제에 대해 사용자의 수행시간, 오류의 수 등의 정량적 데이터를 측정하게 되므로 주체를 [사용자]로 분류.

Logging Actual Use(사용 기록 평가): 컴퓨터를 이용하여 시스템을 사용할 때 사용자의 작업 내용이 자동적으로 기록되므로 주체를 [사용자]로 분류.

Questionnaire(설문조사): 미리 작성된 질문에 의해 사용자가 서면으로 답변하는 방법으로 설문지 설계에 따라 평가자의 주관이 개입될 여지가 다소 있으나 사용자의 주관적인 의견과 선호도를 주로 조사하므로 주체를 [사용자]로 분류.

Focus Group(집중 토의): 사용자 그룹을 형성하여 한명의 중재자에 의해 제품에 대한 토론을 실시하고 평가자는 그 결과를 요약하는 역할을 수행하므로 의견의 심리적 주체는 [사용자]로 분류.

Card Sorting(카드 분류법): 사용자의 생각의 잠재적인 구조, 정신 모형 등을 발견하기 위해, 다수의 사용자가 카드를 자신이 보기 적합하게 분류하도록 하므로 판단의 주체는 [사용자]. 카드 제작은 전문 평가자가 관여하나 판단의 주도권은 사용자에게 있음.

Affinity Diagram(공감 도표법): 다양한 개념이 적힌 메모지를 사용자들이 팀을 만들어 팀별로 분류하도록 하므로 [사용자]가 주체.

Observation(관찰법): 제품 사용 현장에 방법으로 사용자가 자유롭게 정보를 창출할 수 있는 방법이므로 주체를 [사용자]로 분류.

Interview(면접): 이 방법은 그 구체적인 질문내용을 평가자가 만들고 직접 실험에 참여하여 평가자가 구두로 질문하므로 평가자의 관점과 태도가 반영될 여지가 크고, 이와 동시에 사용자의 의견이 직접적인 결과로 반영되므로 [평가자+사용자]로 분류.

Question-asking Protocol(발문법): 제품 사용 중에 사용자의 생각을 표현하는 방법으로, 그 전에 평가자가 제품의 특정 항목을 지정하고 직접 질문을 하게 되므로 [평가자+사용자]로 분류.

Contextual Inquiry(전후 관계의 질문법): 면접법과 유사

하나 사용자의 행동의 동기나 전략에 대해 평가자가 관찰하고 질문하는 방법이므로 [평가자+사용자]로 분류.

Pluralistic Walkthrough(복합적 검사): 개발자, 일반 사용자, 전문 평가자들이 함께 만나서 토의하여 다양한 범주의 문제점을 발견하는 방법이므로 [개발자+평가자+사용자]로 분류.

위의 분류를 표1에 정리하였다.

표 1. 본 연구에서 제안한 심리적 평가의 주체에 따른 평가 방법의 재분류

Center of Usability Evaluation (사용성 평가의 주체)	Method(방법)
Developer+Evaluator (개발자+평가자)	Guideline Checklist-1 (특정 설계 지침에 따른 검사)
	Heuristic Evaluation(경험적 검사)
Evaluator (평가자)	Cognitive Walkthrough (인지적 검사)
	Guideline Checklist-2 (일반적 설계 지침에 따른 검사)
	Standard Inspection(표준 검사)
	Consistency inspection (일관성 검사)
	Feature Inspection(기능검사)
	Expert Review(전문가 평가)
	Thinking aloud(사고 발성법)
	Co-discovery method (동시 발성법)
User (사용자)	Performance measurement (수행도 측정)
	Logging actual use (사용 기록 평가)
	Questionnaire(설문 조사)
	Focus group(집중 토의)
	Card sorting(카드 분류법)
	Affinity Diagram(공감 도표법)
	Observation(관찰법)
Evaluator+User (평가자+사용자)	Interview(면접)
	Question-asking protocol (발문법)
	Contextual inquiry (전후 관계의 질문법)
Developer+Evaluator+User (개발자+평가자+사용자)	Pluralistic walkthrough (복합적 검사)

2.4 MCUEM의 평가 방법 선정 원칙

MCUEM을 사용하기 위해서는 첫째, 기본적으로 본 연구

에서 재분류한 평가 주체 중 복수의 평가 주체를 선정한다. 둘째, 각각의 주체가 복수의 평가 방법을 사용하여야 한다. 이 때 선정 기준은 평가 방법의 특성과 목적에 따라 상호 보완적인 것을 선택한다. 이는 사용성 평가의 정밀도를 향상시킬 수 있는 장점이 있다. 단, [개발자+평가자]의 경우는 현재 1개의 방법만이 선정되어 있으므로 예외로 한다. 이런 두 가지 원칙을 지키되 평가자의 판단에 의한 추가적 평가 방법의 선정을 제한하지 않는다.

2.5. MCUEM의 검증을 위한 실험

본 연구에서는 MCUEM을 사용하지 않을 경우 발생할 수 있는 결과의 편향성을 조사하기 위해 휴대전화의 예를 들어 실제 사용성 평가를 실시하였다.

2.5.1 실험에 사용된 평가 방법의 선정

평가 방법의 선정은 MCUEM 설계자의 주도로 이루어지나, 위에서 언급한 선정 원칙을 벗어나지 않도록 하였다. 선정된 방법과 선정 이유를 명시하였다.

1. [개발자+평가자] 중심방법

Guideline Checklist(특정 설계 지침에 따른 검사): 개발자와 평가자가 동시에 주체적으로 관여하는 방법 중 유일한 방법이므로 본 연구에서 선정되었다. 특히 본 실험에서는 휴대전화의 인터페이스 설계를 위한 특정한 사용성 평가 checklist를 전문 평가자가 설계하였고 이를 현장의 개발자가 사용하도록 하였다. 본 실험을 위해 설계된 checklist를 표 2에 정리하였다.

2. [평가자] 중심방법

Heuristic Evaluation(경험적 검사): 이 방법은 제품의 기능과 구조가 가지는 심각한 문제점을 포함한 대다수의 사용성 문제점을 발견할 수 있다(Jeffries et al., 1991)는 장점이 있어 선택하였다.

Cognitive Walkthrough(인지적 검사): 수행도 검사 시나리오에 따라 실험 참가자가 겪었을 인지적 과정을 탐색하기 위하여 선택하였다. 경험적 검사에서 제품의 기능별 또는 구조별 평가가 이루어지는 반면 수행순서와 과정에 대해 문제점을 발견할 수 있어 상호보완적인 방법으로 사용되었다.

3. [사용자] 중심방법

Performance Measurement(수행도 측정): [사용자] 중심의 평가 방법 중에는 정량적이고 객관적인 자료를 얻을 수 있다는 장점이 있어 선택하였다.

Questionnaire(설문조사): 수행도 측정이 사용자의 객관적인 정보를 추출하는데 반해 [사용자] 중심의 주관적인 의견을 추출할 수 있어 수행도 측정결과를 보완할 수 있다고 판단되어 선택하였다.

4. [평가자+사용자] 중심방법

Question-asking Protocol(발문법): 사용자들의 주관적인 의견과 선호도를 평가하기 위해 발문법을 선택하였다. 이 방법은 실험 중에 질문을 던지게 되므로 실험참가자들이 기억용량에 구애 받지 않고 실시간 정보를 획득할 수 있다는 장점이 있어 선택하였다.

Interview(면접): 수행도 평가 후에 문제점으로 예상되는 제품의 기능이나 이미지에 대해 심도 있는 정보를 획득할 수 있어 발문법에서 얻는 즉흥적인 답변을 보완할 수 있다는 장점이 있어 선택하였다.

5. [개발자+평가자+사용자] 중심방법

Pluralistic Walkthrough(복합적 검사): 복합적 검사 방법을 채택하여 각 평가의 주체를 모두 포함하여 종합적인 정보를 얻기 위해 선택하였다.

2.5.2 실험 과정

1. [개발자+평가자] 중심의 평가

휴대전화 사용성 평가를 위해 기존의 분류 항목을 재구성하여 만든 checklist(이하연 외, 2002)를 사용하였다(표 2). Checklist의 분류 항목은 학습성, 논리성, 접근성, 사용자 조작의 유연성, 일관성, 항행 용이성, 정보표현의 적절성, 유사성과 은유성, 오조작 예방성, 시인성, 미적 완성도, 시스템, 하드웨어의 조작 및 인지 용이성이다. Checklist는 위의 분류에 따라 53개의 문항으로 구성되어 개발자에게 제공되었다. 즉, 전문 평가자에 의해 독립적으로 만들어진 항목을 사용하여, 개발자가 독립적으로 평가하였다.

2. [평가자] 중심의 평가

사용성 평가의 경험이 있고 단말기의 구조와 기능에 대해 이해하고 있는 교수 1인, 박사급 연구원 2인, 본 저자 1인이 전문가 평가에 참여하였다. 메뉴, 화면구성 및 표현 그리고, 하드웨어의 형태와 기능의 5항목을 중심으로 평가하되 평가자에게는 참고로 checklist와 수행도 평가 및 설문조사에서 지적된 항목들이 참고로 제공되었다. 이 과정에서 유사한 평가항목에 대해 어떤 다른 평가가 나오는지 조사할 수 있었다. 또한 수행도 평가에서 지적된 항목에 대해 인지적 검사를 실시하도록 하여 사용자 본인이 표현하지 못했던 실수의 인지 심리적 과정에 대해 추정해 보도록 하였다.

3. [사용자] 중심의 평가

수행도 측정과 설문조사를 실시하였다. 평가에 참여한 사용자는 20~30대의 남녀로 총 18명으로 하였으며, '초보자', '중급자', '고급자' 각 6명으로 구성하여 실시하였다. 수행도 측정은 휴대폰의 일반적 기능과 더불어 무선 인터넷 기능에 대한 수행도를 주로 평가하였으므로 실험 참가자의 숙련도를 무선 인터넷 사용횟수에 따라 분류하였다. 즉, 초보자는 무선 인터넷을 한 달에 4회 이하로 접속하고, 평균 1회 접속

표 2. [개발자+평가자] 중심의 방법에서 사용된 Checklist

논리성	1. 사용자가 쉽게 이해할 수 있는 메뉴구조인가?
	2. 사용자가 메뉴의 이름을 보고 세부 항목을 예측할 수 있는가?
	3. 사용자가 메뉴의 아이콘을 보고 세부 항목을 예측할 수 있는가?
	4. 과제 수행의 단계가 논리적인가?
	5. 메뉴와 세부 내용이 정확히 일치하는가?
	6. 같은 수준의 메뉴가 논리적으로 뚜렷이 구분되는가?
접근성	1. 재미있는 메뉴로의 접근이 용이한가?
	2. 과제 수행의 단계수가 많은가?
학습성	1. 사용법이 사용자들에게 친숙한가?
	2. 전문용어의 사용을 지양하고 사용자들에게 친근한 용어를 사용하였는가?
	3. 사용된 단어가 보편적이고 객관적인가?
사용자 조작의 유연성	1. 실수를 교정하기가 용이한가?
	2. 과제 수행 과정이 가지고 있는 기대와 일치하는가?
항행용이성	1. 메뉴간의 이동이 용이한가?
	2. 원하는 항목으로 갈 수 있도록 화면상에 해당 버튼에 대한 정보를 명확히 알려주고 있는가?
	3. 한 화면에 입력해야할 항목이 많은 경우, 항목간의 이동을 쉽게 할 수 있는가?
일관성	1. 중요한 메시지가 화면의 일정부분에 일관성 있게 표현되고 있는가?
	2. 용어의 사용이 일관성이 있는가?
	3. 조작버튼의 기능에 대한 이미지가 일관성 있게 표현되고 있는가?
	4. 같은 의미의 아이콘을 일관성 있게 사용하고 있는가?
유사성과 은유성	1. 단어가 의도한 의미를 쉽게 알 수 있는가?
	2. 아이콘이 주는 이미지가 적절하였는가?
정보표현의 적절성	1. 단순한 아이콘을 사용하였는가?
	2. 아이콘이 무엇을 의미하고 명령하는지 구체적으로 이해할 수 있는가?
	3. 아이콘으로 이해한 내용과 실제로 수행된 기능이 정확히 일치하는가?
	4. 간단한 단어로 의사를 전달하고 있는가?
	5. 메뉴를 설명하고 있는 단어가 무엇을 의미하고 명령하는지 구체적으로 이해 할 수 있는가?
	6. 단어를 보고 이해한 내용과 실제 수행된 기능이 정확히 일치하는가?
	7. 맞춤법과 띄어쓰기가 맞는가?
미적완성도	1. 폰트의 이미지가 기능과 일치하는가?
	2. 사용된 색상이 적절한가?
	3. 단말기의 외형이 세련되었는가?
	4. 하드웨어 구성 간의 디자인이 조화로운가?
시인성	1. 글자를 쉽게 알아볼 수 있는가?
	2. 선택하려는 메뉴가 눈에 쉽게 띄는가?
	3. 아이콘과 배경화면이 뚜렷이 구별되는가?

표 2. [개발자+평가자] 중심의 방법에서 사용된 Checklist

시인성	4. 시스템의 현재 상태에 대한 정보를 제공하는가?
	5. 중요한 정보가 한 눈에 들어오는가?
	6. 중요한 정보를 텍스트로 표현하였는가?
오조작 예방성	1. 화면마다 실행 취소 기능(이전, 상위, 취소 등)이 존재하는가?
	2. 피드백 메시지의 내용이 간결하고 정확한가?
	3. 피드백 메시지의 내용이 초급자나 고급자 모두에게 도움이 될 수 있는가?
	4. 사용자가 심각한 오류를 범하려 할 때 시스템이 이것을 알려주는가?
	5. 오류를 범했을 때 제공되는 오류메시지가 오류를 수정할 수 있도록 도움을 주는가?
시스템	1. 시스템의 수행속도가 느리다는 인상을 주지 않는가?
	2. 인터넷 사용 시 수행속도가 느리다는 인상을 주지 않는가?
	3. 시스템 내부의 오류는 없는가?
	4. 시스템 내부 오류 발생 시 제공되는 메시지는 적절한가?
하드웨어 조작 및 인지용이성	1. 방향키의 사용이 편리한가?
	2. 방향키의 배치는 적절한가?
	3. 단축키의 사용은 편리한가?
	4. 단축키의 배치는 적절한가?
	5. 버튼의 아이콘은 알아보기 쉬운가?

시간이 10분 이하인 사용자, '고급자'는 무선 인터넷을 한달에 15회 이상 접속하고, 총 접속시간이 1시간 이상인 사용자, 이 두가지 조건을 만족시키지 못한 사용자를 중급자로 정의하였다.

실험기구로써는 'TO'(개발상품명) 단말기 두 대와 수행시간 측정을 위하여 스톱워치 그리고 실험 전 과정을 녹화하기 위하여 캠코더가 사용되었다.

실험 과정은 실험 참가자의 아이콘의 인지도에 대한 설문조사(새로운 아이콘 디자인이 추가되어 이를 평가에 포함), 실험 전 조작 방법 학습 및 사용자 개인 신상에 대한 설문지 작성, 본 실험 중의 수행도 측정 및 실험 후의 설문지 작성으로 구성하였다. 수행도 측정의 측정 변수로써 과제 수행하는 데 걸린 시간, 오류의 수, 오류의 종류와 반복 오류의 종류를 측정하였다.

4. [평가자+사용자] 중심의 평가

수행도 평가 과제 수행 중에 발문법을 사용하여 하나의 과제를 수행하고 다음 과제에 영향을 미치지 않는 범위 내에서 실험참가자의 의견을 수집하였다. 실험 후에 설문조사를 마치고 설문에서 좋지 않은 평가를 받은 문항을 중심으로 평가자가 사용자를 대상으로 면접을 실시하였다.

5. [개발자+평가자+사용자] 중심의 평가 방법

종합적인 검사를 시행함에 있어서 개발자와 평가자, 사용

자가 모두 한 자리에 모여 토의하기 힘든 현실을 고려하여 개발자가 제시한 문제점을 가지고 전문 평가자들과 면접 시 사용성의 문제를 가장 많이 지적한 2인의 사용자가 토의를 진행하였고, 추가적으로 문제점이 지적되었다.

3. 연구 결과

3.1 [개발자+평가자] 중심의 평가 결과

Checklist의 13개의 분류 항목에서 총 53문항 중 개발자는 메뉴의 구조가 부적절하다는 것과 용어의 표현 및 의미가 부적절하다는 것, 화면이 전환되는 속도가 느리다는 문제점을 포함하여 총 11개의 문제점을 지적하였고 그 내용은 표 5에 요약되었다.

3.2 [평가자] 중심의 평가 결과

경험적 검사 방법과 인지적 검사 방법으로 다섯 항목에 대한 총 146개의 문제점이 도출되었으며 각 항목 별 문제점의 수와 주 문제점의 내용은 표 3과 같다. 주 메뉴 항목에서는 주 메뉴의 전체적인 순서에 대한 논리성 부족과 언어 표현 문제에서 많은 문제들이 발견되었으며 화면 구성 및 표현에 대한 평가에서는 단말기 소프트웨어의 전반적인 표현 및 구성에 대한 27개의 문제점이 발견되었다. 그리고 인터넷 접속 환경에서는 특히 파일을 다운로드 받는 과정에서의 13개의 문제점이 나타났다. 하위 메뉴 수행 시 관련된 화면 및 기능에 대한 문제점은 74개가 발견되었다(표 3). 이를 항목 별로 문제점을 재분류하여 32개 항목을 표 5에 요약하였다.

표 3. 평가자 중심 결과 발견된 문제점 수와 문제점 개요

항목	발견된 문제점 수	주문제점
주메뉴	17개	주 메뉴의 구조와 순서의 논리성 부족 메뉴 이름 표현의 부적절성
화면구성 및 표현	27개	소프트웨어 키의 표현 및 사용의 모호성 전반적인 화면의 표현이 오류를 유도
하드웨어	15개	시스템의 불안정성 하드웨어 디자인 및 버튼과 아이콘의 디자인 불명확성
인터넷 접속환경	13개	인터넷 접속시 메뉴의 부정확성 파일 다운로드 시 화면 표현의 부적절성
하위 메뉴 별 평가	74개	각 하위 메뉴에서의 다양한 세부적인 문제점 도출

3.3 [사용자] 중심의 평가 결과

3.3.1 수행도 측정

18명의 실험 참가자가 총 21개의 단계별 과제를 완료하는 데 걸린 수행시간을 평균하였다. 그리고 과제 수행을 완료하지 못한 실험 참가자의 수가 많은 과제의 순서대로 과제별 평균 수행시간과 실패수를 나열하였다(표 4). 캐릭터 다운로드의 경우 시스템의 불안정으로 인한 인터넷 접속 불량으로 완수하지 못한 참가자가 많았고 시작메뉴의 등록 및 편집에서 과반수의 참가자가 과제를 완수하지 못하여 시작

표 4. 수행도 평가의 결과

Task 번호	작업(Task) 내용	수행시간 평균 (분:초)	참가자의 인지오류에 의한 실패수	시인성, 조작성, 시스템 불안정 등 기기오류에 의한 실패수
1	대기화면 변경	0:23	4	2
2	방송채널 On/Off	0:58	15	8
3	벨소리를 무음램프로 전환	0:32	5	3
4	테마 변경	0:22	7	2
5	알람 설정	1:15	5	14
6	착신거절 설정	1:00	11	1
7	시작메뉴 등록	1:19	14	1
8	시작메뉴 확인	0:12	4	·
9	시작메뉴 편집	0:52	7	·
10	전화번호부 등록	1:08	2	5
11	그룹 변경	0:54	1	9
12	자동지역번호 설정	0:36	12	3
13	마이폴더 내 파일삭제	0:58	14	3
14	일정 추가	1:21	5	14
15	화면 밝기 조정	0:44	9	2
16	ez-I접속	0:11	2	·
17	게임 다운받기	3:38	11	1
18	메시지 보내기	2:17	7	4
19	자기번호찾기	0:26	6	1
20	총통화시간 조회	0:24	8	·
21	벨소리 유무확인 및 다운로드	3:33	·	4
22	캐릭터 유무확인 및 다운로드	3:09	·	2
23	다운받은 멜로디 삭제	1:03	16	·
24	통화목록 확인 및 전화번호부 등록	0:53	2	5
25	대기화면의 초기문구 변경	1:22	·	·
26	전화번호(7712)로 찾기	0:24	2	3
Total			169	87

메뉴의 전반적인 인터페이스에 문제점이 있음이 지적되었다. 표 5에 총 26개의 문제점을 항목별로 요약하였다.

표 5. 다자중심적 사용편의성 평가 방법의 항목별 문제점 도출 결과

문제점 항목 (52항목)	◆ 문제점이 발견되었음			
	[개발자+평가자] 중심	[평가자] 중심	[사용자] 중심 수행도 측정 설문 조사	[평가자+사용자] 중심 [개발자+평가자+사용자] 중심
대기화면변경		◆	◆	◆
방송채널 On/Off		◆	◆	◆
테마변경		◆	◆	
시작메뉴변경		◆	◆	◆
게임 다운		◆	◆	◆
시작메뉴 등록		◆		◆
초기문구변경		◆		
전화번호부 검색		◆		
시작메뉴 편집		◆		◆
인터넷 접속		◆		◆
벨소리 다운로드		◆		◆
캐릭터 다운로드		◆		◆
다운 받은 벨소리 지우기		◆		◆
통화목록확인, 등록		◆	◆	
전화번호부 등록		◆	◆	
자기번호 찾기		◆	◆	
일정 추가		◆	◆	
알람 설정		◆	◆	
그룹 변경		◆	◆	
화면밝기 조정		◆	◆	
착신거절 설정			◆	◆
메시지 보내기			◆	◆
자동지역번호 설정			◆	◆
마이폴더 내 파일 삭제			◆	◆
총통화시간 조회			◆	
무음램프로 전환			◆	
메뉴구조 및 이름의 부적절성	◆	◆		◆
용어사용 및 의미	◆	◆		◆
화면 전환되는 속도	◆	◆		◆
하드웨어 버튼 디자인	◆	◆		◆
피드백 메시지 내용	◆	◆		◆
아이콘의 식별성	◆	◆		◆
조작 키의 수	◆	◆		
중요한 정보의 시인성	◆	◆		
재미있는 메뉴 접근성	◆	◆		
아이콘의 일관성	◆	◆		
한 화면의 메뉴 수	◆			◆

표 5. 다자중심적 사용편의성 평가 방법의 항목별 문제점 도출 결과

◆ 문제점이 발견되었음

문제점 항목 (52항목)	[개발자+평가자] 중심		[사용자] 중심		[평가자+사용자] 중심	
	중심	중심	수행도 측정	설문 조사	중심	중심
하드웨어 아이콘 식별		◆		◆		
버튼 사용의 용이		◆		◆		
단말기 버튼의 디자인 만족도				◆		
아이콘의 만족도				◆		
사용법				◆		
향후 구입 의향				◆		
단말기 디자인의 만족도				◆		
사용 용이성				◆		
전반적인 사용 만족도				◆		
실수 교정의 용이성				◆		
마이폴더 존속 여부						◆
삭제기능 존속 여부 (마이폴더)						◆
단축키 기능의 메뉴 수준 차이						◆
시작메뉴의 존속 여부						◆
하이퍼 컷 화면의 커서 위치						◆
평가방법별로 발견된 문항수	11개	32개	26개	24개	14개	5개

3.3.2 설문조사

실험 전에 아이콘의 인지도에 대한 설문조사에서 총 64개의 아이콘 중 14개의 아이콘의 의미를 참가자들의 과반 수 이상을 파악하지 못하였다. 실험 후, 단말기에 대한 전반적인 인상 및 사용에 대한 총 43문항의 설문조사에서 24문항이 좋지 않은 점수를 획득하였다(표 5 참조).

3.4 [평가자+사용자] 중심의 평가 결과

발문법과 면접을 통하여 사용자의 심리적 상황을 보다 면밀하게 이해할 수 있으며, 이를 통하여 수행도가 낮았던 항목에 관한 원인 분석의 힌트로 사용하였다. 이 방법으로 14 문항에서 문제점이 제시되었다(표 5 참조).

3.5 [개발자+평가자+사용자] 중심의 평가 결과

복합적 검사를 실시한 결과 다섯 가지의 문제 항목이 지적되었다.

4. 토 의

표 5에서 보여주는 바와 같이, 지적된 문제점의 항목의 내용과 개수는 어떤 평가 주체를 사용했느냐에 따라 차이를 보였다. 즉, 전체 문제점 중 [개발자+평가자] 중심의 방법은 21.2%를, [평가자] 중심의 방법은 61.5%를, [사용자] 중심의 방법은 82.7%를, [평가자+사용자] 방법은 26.9%를, 그리고 [개발자+평가자+사용자] 중심의 방법은 9.6%를 발견할 수 있었다. 이를 통하여 평가 주체 변화에 따른 편향성이 평가 주체에 따라 매우 다양하게 발생한다는 것을 확인할 수 있었다. 한편, [사용자] 중심의 방법만으로도 전체 문제점의 82%를 발견한 것을 볼 때, 정밀한 사용성 평가를 필요로 않은 경우에는 [사용자] 중심의 평가로만으로도 만족할 수 있다고 볼 수도 있을 것이다. 그러나 문제는 전문성과 장기간의 학습이 요구되는 제품의 경우 사용자들이 내재되어 있는 문제점을 발견하지 못할 소지가 많아 [사용자] 중심의 방법론만으로 양질의 평가를 할 수 있다고 결론내리기는 어렵다고 볼 수 있다.

한편, 본 연구에서는 휴대전화만을 대상으로 실험하여 그 평가 결과의 해석은 다소 제한적이 될 수밖에 없다. 그리고 각 평가 주체에 따른 평가 방법 선정 과정과 Checklist 제작 과정에서 연구자의 주관적인 경험과 판단이 사용되어, 이를 보다 객관화 할 수 있는 방법에 대한 논의가 필요하다. 그럼에도 불구하고, 이러한 방법을 통해 다양한 사용성 문제점을 발견할 수 있게 되어, 본 연구에서 제한한 구조화 과정이 일단 다자 중심적 평가 방법의 목적에 부합했다고 평가 되었다.

다자 중심적 사용성 평가 방법(MCUEM)은 평가의 주체 변화에 따른 편향성을 최소화 하려는 현장 평가자들의 요구에 부응하기 위해 실용적 목적으로 구조화되었다. 아직도 한정된 제품을 전제로 만들어졌다는 한계와, 일부 사용성 평가 방법이 포함되지 않아, 구조화 과정을 수정 없이 모든 제품에 그대로 사용하는 데는 다소 무리가 있을 수 있으나, 유사 제품에는 적용이 가능하다고 판단된다. 결국 최종적인 사용성 평가 방법의 선택은 경제적 비용과 소요기간에 따라 좌우 될 수 있으나 각 방법에 대한 경제적 비용에 대한 예측은 본 논문에서 다루지 못하였다. 그러나 본 연구에서 표현된 정의와 분류 및 선정 원칙들은 여타 다양한 제품의 사용성 평가 방법을 설계하는데도 일관되게 사용될 수 있을 것으로 기대한다.

5. 결 론

사용성 평가 방법에 따라 발생하는 결과의 편향성을 최소화

화하기 위해 일반적으로 인식되어 왔던 복수 주체에 의한 평가를 다자 중심적 사용성 평가 방법(MCUEM)으로 명명하고 이를 구조화 하였고 그 효용성을 실험 예를 들어 검증하였다. 이를 위해 평가의 심리적 주체-개발자, 평가자, 사용자-를 정의하고, 이들 평가 주체간의 조합을 새롭게 제안하였고, 이 조합에 따라 사용성 평가 방법들을 새롭게 분류하였다. 이러한 구조에 따라 평가 대상을 휴대전화로 선정하고 평가를 실시하였다.

평가 결과, 평가의 주체가 변화할 때 결과의 편향성이 발생하는 것을 확인할 수 있었고, 이러한 편향성이 구조화된 다자 중심적 사용성 평가 방법으로 효율적으로 발견되고 수정될 수 있다는 것을 확인하였다.

따라서 추후 제품이나 시스템의 사용성을 평가할 경우 평가 방법의 목적이나 도구적 차이에 의한 평가 방법의 선택과 함께, 다양한 주체가 관여된 평가 방법의 선정이 필요하고, 이 과정에서 보다 구조화된 형태의 다자 중심의 평가 방법을 사용하는 것이 중요하다고 생각한다. 본 연구에서 제안된 다자 중심적 평가 방법의 구조와 실행 방법은 추후 지속적인 연구를 통해 보다 구체적이고 실용적으로 발전될 수 있을 것이다.

참고 문헌

이하연, 김정룡., "휴대전화 인터페이스에 대한 다중심적 사용편의성 평가", 대한인간공학회 추계학술대회, 용평리조트, 11월 29일, p259-262, 2002.
 임치환, "사용성 평가 방법들의 비교 및 선택", 한국공업경영학회지, Vol. 19, No. 37, 87-94, 1996.

Nielsen, J., "Usability Inspection Methods", Proceedings ACM CHI Conference, http://acm.org/sigchi/chi9_5/proceedings/tutors/jn_bdy.htm, 1995.
 Hom, J., "Usability Toolbox," <http://jthom.best.vwh.net/usability/>, 1996.
 Perlman, G., "Practical Usability Evaluation", *Proceedings ACM CHI Conference*, 1996.
 Jeffries, R., Miler, J. R., Wharton, C. and Uyeda, K. M., "User Interface Evaluation in the Real World: A Comparison of Four Techniques", *Proceedings ACM CHI Conference*, 119-123, 1991.

◎ 저자 소개 ◎

❖ 김 정 룡 ❖
 한양대학교 기계공학과 졸업
 Ohio State Univ. 산업공학 석사
 Ohio State Univ. 산업공학 박사
 한양대학교 정보경영공학과 교수
 관심분야: 생체역학, 인지공학, 인체공학적 제품설계

❖ 이 하 연 ❖
 한양대학교 산업공학과 석사
 현재 LG전자 MC사업본부 연구원
 관심분야: 인간공학, 인지공학, HCI

❖ 최 영 철 ❖
 한양대학교 산업공학과 석사과정
 관심분야: 인간공학, 인지공학, HCI

논문 접수 일 (Date Received) : 2005년 03월 23일
 논문게재승인일 (Date Accepted) : 2005년 05월 21일