

모바일 인터랙션 평가에서 실제감 재현에 대한 경험적 연구

An Empirical Study on the Recreation of Reality for Evaluating Mobile Interaction

박정순

천안대학교 디자인영상학부

Park, Jeong-Soon

Division of Design & Media, Cheonan University

신수길

세종대학교 산업디자인학과

Shin, Soo-Kil

Dept. of Industrial Design, Sejong University

• Key words: Mobile Interaction, Mobile Usability Evaluation, Mobility

1. 서론

모바일 시스템은 전형적으로 매우 역동적인 상황에서 사용되며, 이런 상황에서 사용자들은 다양한 물리적 환경에 놓여지게 된다. 따라서 현장 기반의 평가가 모바일 시스템을 위한 바람직한 접근 방법으로 보인다. 그러나 현장에서 사용성을 평가하는 것이 쉽지 않다. 여러 문헌을 종합해 볼 때 현장평가의 어려움을 세 가지로 정리할 수 있다. 첫 번째는 위에서 기술한 사용맥락 중에서 중요한 상황을 포착하기 위해 실제적인 환경을 만들기 복잡하고 까다롭다는 것이며, 두 번째는 현장에서 평가를 수행할 때 관찰기법과 소리내어 말하기 기법과 같이 이미 확립된 기법을 적용하기가 그렇게 쉽지않은 것이다. 마지막 세 번째는 현장 평가의 경우 사용자들이 실험설정에 잠재적으로 영향을 주는 다수의 알려지지 않은 변수들로 이루어진 환경에서 신체적으로 움직이기 때문에 통제에 한계가 있으며 데이터 수집을 까다롭게 한다는 것이다. 실험실 환경 안에서 이러한 애로점들은 눈에 띄게 줄어든다. 사용성 평가가 실험실 환경 안에서 수행될 때 실험 통제와 고품질의 데이터 수집은 문제가 되지 않는다. 그러나 이러한 설정의 약점 중 하나는 실제성의 결여이다.

이와 같은 배경 하에서 본 논문의 목적은 실험실 환경에서 모바일 시스템의 사용성을 평가하기 위해 실험실 평가의 장점은 그대로 유지하면서 위에서 논의한 문제점들을 해결할 수 있는 새로운 기법들을 탐색하는 것이다.

2. 평가방법을 위한 프레임워크

표 1. 신체의 움직임과 네비게이션 여부에 따른 다섯가지 구성

		네비게이션하기 위해 요구되는 주의력	
		무의식적 네비게이션	의식적 네비게이션
신체 움직임 여부	정지	1. 테이블에 앉아 있거나 서있는 경우	
	일정한 속도	2. 러닝머신 위를 일정한 속도로 걸거나 보행기구를 이용해 정해진 속도로 발걸음하는 경우	4. 장애물 때문에 일정하지 않은 경로를 일정한 속도로 걷는 경우
	불규칙한 속도	3. 러닝머신 위를 불규칙한 속도로 걷는 경우	5. 장애물 때문에 일정하지 않은 경로를 불규칙한 속도로 걷는 경우

본 연구에서 제안하는 평가방법은 사용자가 모바일 시스템을 사용하는 동안 사용자의 신체적인 움직임에 초점을 맞추고 있다. 이러한 개념을 기술하기 위하여 다음과 같은 두 가지 차원이 사

용된다.

- 신체 동작의 유형 : 정지된 상태, 일정한 속도로 움직임, 불규칙한 속도로 움직임
- 네비게이션하기 위해 요구되는 주의력 : 무의식적 네비게이션, 의식적 네비게이션

이런 두 가지 차원을 병렬로 배치함으로써 실험실에서의 사용성 평가 설정을 구체화하기 위한 2x3 매트릭스를 만들 수 있다.[표1]

3. 실험설계

실험은 [표 1]에서 기술한 다섯 개의 실험설정과 실제 보행자 거리에서의 평가를 포함한 6가지 방법 각각을 채용하여 일련의 사용성 평가를 실시하였다. 각각의 평가에서 사용자는 모바일 시스템을 사용하면서 특별히 만들어진 다수의 과업들을 해결하였다. 실험에 사용된 모바일 시스템은 컴팩 iPAQ PDA를 위해 실험적으로 만들어진 SMS(Short Messaging Service) 프로그램이다. 이 프로그램은 사용자가 짧은 문자 메시지를 주고 받을 수 있도록 해주며, 추가로 EMS(Enhanced Message Service) 규격을 준수하고 있다. 사용하는 프로그램은 전용 감시 프로그램을 백그라운드로 실행시켜 사용자 행동이나 수행도가 정확하게 측정될 수 있도록 실험을 위해 특수하게 만들어진 것이다. SMS를 실험용 프로그램으로 선택한 이유는 SMS가 모바일 사용자 사이에서 가장 널리 사용되고 있을 뿐 아니라 스크린을 이용한 읽기와 키보드를 이용한 글자의 타이핑을 모두 포함하는 매우 인터랙티브한 기능이기 때문이다. 36명의 피실험자를 대상으로 짧은 메시지를 보내거나 받는 것을 포함한 5개의 과업을 수행하도록 하였다. 주어진 과업을 해결하는 동안 피실험자에게 본인의 상태를 소리내어 말하도록 하였다. 정해진 시간일정을 맞추기 위해 각각의 피실험자에게 10분의 시간이 배정되었고, 각각의 평가에 대해 다음과 같은 세가지 유형의 데이터가 수집되었다.

- 사용성과 관련된 문제점
- 수행도
- 작업부하

4. 실험결과

4.1 도출된 사용성 문제

표 2. 여섯가지 실험설정에 의해 규명된 사용성 문제

	실험 설정					
	1	2	3	4	5	6
평균	10.8	7.5	6.7	6.7	5.2	6.3
표준편차	1.6	1.5	2.0	2.4	2.4	2.1

표 3. 실험에 의해 규명된 사용성 문제의 유형별 개수

	실험 설정						중요도별 개수
	1	2	3	4	5	6	
Critical Level	4	4	3	4	3	3	4
Serious Level	11	11	9	9	9	8	17
Cosmetic Level	19	8	8	8	6	12	32
실험설정 별 총합	34	23	20	21	18	23	53

여섯 개의 실험설정 중 어떤 것도 모든 사용성 문제의 도출을 지원하지 못했다. 탁자에 앉아 모바일 시스템을 사용하는 첫 번째 실험설정에서 다른 방법에서보다 많은 34개의 사용성 문제를 발견했다. 나머지 실험설정 방법의 경우에는 전체의 반도 안되는 사용성 문제만이 규명되었는데, 결정적으로 중요한 수준과 비교적 중요한 수준을 놓고 보면 여섯 개의 실험설정 모두 비슷한 숫자의 사용성 문제 개수를 보이고 있다. 첫 번째 실험설정과 나머지 실험설정 사이의 주요 차이점은 표면적인 사소한 수준의 사용성 문제 개수와 관련이 있다. 즉 다른 실험설정 방법을 사용했을 때보다 탁자에 앉아 모바일 시스템을 사용하는 실험설정에서 두 배 이상이나 되는 표면적 수준의 사용성 문제를 발견해냈다.

4.2 수행도

우리는 각각의 평가방법에서 수집한 수행도 사이에 명확한 차이가 있을 것으로 예상했다. 주요 측정자료는 다섯 개 과업에 대해 각각 측정된 소요시간이다. 수행도 측면에서 우리가 생각한 가설은 최적의 평가방법에서 사용자가 발휘하는 수행도가 현장평가(여섯번째 실험설정)에서 측정한 수행도와 유사할 것이라는 것이다. 여섯 개의 실험설정을 바탕으로 수행한 각각의 과업에 대한 소요시간을 분산분석해본 결과 각 실험설정 사이에 어떤 구조적인 차이점도 발견할 수 없었다. 가장 빠른 과업 완료시간을 기록한 실험설정은 과업에 따라 서로 달랐다. 이와 함께 완료된 과업의 개수, 잘못 버튼을 누른 회수, 과업에 대한 설명을 다시 해달라고 피실험자가 요청한 회수 등을 분석하였는데 각 평가방법 사이를 명확하게 구분지을만한 어떤 차이점도 발견하지 못했다.

4.3 작업부하

표 4. 각 실험설정에서 작업부하와 관련된 주관적 경험

	실험 설정					
	1	2	3	4	5	6
정신적 요구	29	75	204	126	185	148
신체적 요구	92	117	112	118	127	194
노력	52	163	106	228	178	186
전체적인 작업부하	27	35	48	55	48	54

1번부터 6번까지의 실험설정을 쌍대비교한 결과는 다소 유의미한 차이를 보이고 있는데, 러닝머신위를 일정한 속도로 걷는 2번 실험설정을 제외한 모든 실험설정과 비교해 탁자앞에 앉아서 모바일 시스템을 사용하는 1번 실험설정이 정신적 행위를 덜 요구하는 것으로 나타났다. 이것은 새로운 실험설정들과 현장에서의 실험이 똑같이 피실험자에게 더 많은 주의력을 요구한다는 사실을 증명하는 것처럼 보인다. 그러나 신체적 요구와 노력, 전체적인 작업부하와 관련하여 각 실험설정에 대한 심층 비교 결과는 3

번 실험설정에 대한 2번 실험설정의 비교결과를 제외하면 어떤 유의미한 결과도 보여주지 않는다.

5. 결론

본 논문은 실험실 환경에서 모바일 시스템의 사용성을 평가하기 위한 여섯가지 실험설정 방법을 제시한다. 이런 실험설정의 목적은 통제된 환경에서 모바일 시스템을 평가할 수 있는 환경구축이며, 가능한 한 자연스런 사용 상황을 재현하고자 함이다. 다섯가지 실험설정은 이동하는 동안 요구되는 주의력과 신체적 움직임에 대한 이동성을 기초로 한 프레임워크를 바탕으로 만들어졌다. 만들어진 실험설정은 두가지 실험을 통해 평가되었다. 두 실험 모두 보행자 거리에서의 현장평가를 비교기준으로 사용하였다. 실험결과 사용자 수행도 측면에서 기법들 사이에 유의미한 차이는 없었으며, 작업부하의 경우 주어진 과업을 해결하기 위해 정신적, 신체적으로 얼마나 열심히 작업했는가를 의미하는 노력측면과 전체적인 작업부하에서 의미있는 차이를 보여주었다. 그러나 보행자 거리를 걷는 것과 정확히 같은 작업부하를 나타내는 실험설정은 없었다. 사용성 문제의 규명을 얼마나 지원하는가에 관한 측면에서는 단 하나의 유의미한 차이가 있었다. 사용성 문제의 규명에 초점을 맞추었을 때 여섯가지 실험설정 중 가장 단순한 방법이라 할 수 있는 탁자에 앉아 시스템을 평가하는 실험설정이 다른 실험설정 방법들보다 확실히 우수한 결과를 보여주었다. 그러나 차이는 주로 표면적이고 사소한 사용성 문제 수준에서 나타났다. 두가지 실험 모두 명확한 한계점을 가지고 있다. 각각의 실험설정 모두 여섯명의 피실험자를 대상으로 평가되었으며, 실험 B의 일부 실험에서는 단 4-5명의 피실험자만이 참여하였다. 그러나 제한된 피실험자들을 대상으로 여러개의 실험설정을 비교하는 것이 우리의 목적이기 때문에 더 많은 피실험자를 대상으로 할 필요가 있으며, 의미있는 실험설정에 대한 후속실험에서는 피실험자 수를 늘려야 할 것이다.

본 논문에서 제안된 프레임워크와 관련 실험들은 모바일 시스템과 상호작용하는 동안의 신체적인 움직임을 어떻게 재현할 것인가하는 문제를 탐색하는데 초점을 맞추고 있다. 그러나 신체적 움직임과 이동은 모바일 시스템의 사용상황을 구성하는 많은 새로운 요소들 중 하나일 뿐이며, 다른 관련 요소들로는 모바일 시스템 사용에서 사회적, 물리적, 시간적 상황 요소들이 있을 수 있다. 본 논문과 관련된 몇몇 연구를 살펴보면 이러한 요소들이 사용성 평가에 어떻게 영향을 끼치고, 어떻게 실험실에서의 평가에 구체화할 것인가하는 문제를 탐색하고 있다. 이런 다양한 요소들에 대한 고려와 함께 본 논문에서 제시하는 아이디어와 평가방법을 좀 더 구체적으로 다듬기 위해 후속연구와 실험이 필요할 것이다.

참고문헌

- Hart & Staveland, Development of NASA-TLX : Results of Empirical and Theoretical Research, Human Mental Workload, Elsevier, 1988, pp139-183
- NASA, Task Load Index, <http://iac.dtic.mil/hsiac/products.htm#TLX>
- Molich, Comparative Evaluation of Usability Tests In Proceedings of the UPA Conference, 1998, pp189-200