

Week 12

Conducting A User Study

HCI 이론 및 실습 2018 Spring

Human-Computer Interaction+Design Lab _ Joonhwan Lee

Experimental User Study

Usability Testing vs Experiments

Usability testing

- Improve products
- Few participants
- Results inform design
- Usually not completely replicable
- Conditions controlled as much as possible
- Procedure planned
- Results reported to developers

Experiments for research

- Discover knowledge
- Many participants
- Results validated statistically
- Must be replicable
- Strongly controlled conditions
- Experimental design
- Scientific report to scientific community

Experimental User Study

- ✦ 실험은 왜 하는가?
 - ✦ 과학적 진실이 무엇인지 확인하기 위해서
 - ✦ Research Statement 가 ‘참’인지 평가하기 위해서
- ✦ 실험의 예: 뚱뚱한 사람은 혈압이 높다 (명제)
 - ✦ 다양한 연구 방법이 있을 수 있다.
 - ✦ 의사의 임상 데이터를 분석 (기존에 존재하는 데이터)
 - ✦ 사람들을 모집한 후, 혈압을 측정
 - ✦ 가장 이상적인 방법은 전세계의 모든 사람들의 혈압을 수집하는 것 (가능?)
 - ✦ 샘플 모집하여 실험

Experimental User Study

- ♦ 인터랙션에서 특정 요소를 살펴보기 위한 컨트롤된 실험
- ♦ 실험자는 실험을 위한 hypothesis 를 만든다
- ♦ 여러개의 실험 조건이 고려될 수 있는데, 그들 실험 조건은 조작된 변인에 의해 조금씩 차이가 만들어 짐 (eg., 20 menu items vs. 7 menu items)
- ♦ 서로 다른 조건에 반응하는 행동 방식 (eg., reaction time) 을 측정한다

Experimental Factors

- ◆ Subjects

- ◆ 어떤 사람들을 실험 대상으로 선택할 것인지, 어느 정도의 사람을 대상으로 실험할 것인지
(eg., dashboard re-design for elders)

- ◆ Variables

- ◆ 조작하고 측정해야 하는 것들 (eg., grey on yellow vs. green on yellow or RT)

- ◆ Hypothesis

- ◆ 실험을 통해 알고 싶은 것

- ◆ Experimental design

- ◆ 실험을 어떻게 설계하고 진행할 것인지

Subjects

- ♦ How many subject do we need?
 - ♦ for most usability test, 5 is enough to cover more than 85% of usability problems
 - ♦ Qualitative study: 5 - 15
 - ♦ Quantitate study: 20 is enough
 - ♦ <http://www.useit.com/alertbox/20000319.html>
 - ♦ <http://www.useit.com/alertbox/20040719.html>
 - ♦ http://www.useit.com/alertbox/quantitative_testing.html
 - ♦ <http://www.measuringusability.com/blog/five-history.php>

Testing Condition

- ✦ Usability lab or other controlled space.
- ✦ Emphasis on:
 - ✦ selecting representative users;
 - ✦ developing representative tasks.
- ✦ Tasks usually last no more than 30 minutes.
- ✦ The test conditions should be the same for every participant.
- ✦ Informed consent form explains procedures and deals with ethical issues.

Some Type of Data

- ✦ Time to complete a task.
- ✦ Time to complete a task after a specified time away from the product.
- ✦ Number and type of errors per task.
- ✦ Number of errors per unit of time.
- ✦ Number of navigations to online help or manuals.
- ✦ Number of users making a particular error.
- ✦ Number of users completing task successfully.

Variables & Hypothesis

- ✦ independent variable (IV) : 독립변수
 - ✦ 서로 다른 조건을 만들기 위한 것들
 - ✦ e.g. interface style, number of menu items
- ✦ dependent variable (DV) : 종속변수
 - ✦ 실험을 통해 측정가능한 것들
 - ✦ e.g. time taken, number of errors.
- ✦ 결과의 예측 (prediction of outcome)
 - ✦ IV 와 DV 의 관계를 예측
 - ✦ e.g.: error rate 은 폰트 사이즈가 줄어들면 증가할 것이다.

Experimental Design

- ♦ 실험 디자인 혹은 실험 설계
 - ♦ within groups design
 - ♦ 각각의 참여자들은 모든 condition 을 수행하여 실험한다
 - ♦ condition에 계속 노출됨에 따라 학습효과가 일어날 가능성 발생
 - ♦ 참여자 집단의 variation 에 따라 실험 결과가 영향을 미칠 가능성이 적다
 - ♦ between groups design
 - ♦ 각각의 참여자는 하나의 condition 만 수행
 - ♦ 학습효과 없음
 - ♦ 더 많은 참여자가 필요해서 비용이 많이 듦
 - ♦ 참여자 집단의 variation 이 결과를 왜곡되게 만들 수 있음.

Analysis of Data

- ♦ 다양한 통계 분석 기법이 필요
 - ♦ 데이터의 형태에 따라 통계 분석 기법이 달라짐
 - ♦ 변량의 비교: eg., t-test vs. anova
 - ♦ 예측모델: linear model analysis
- ♦ 분석의 결과로 찾는 것
 - ♦ condition 1 vs. condition 2
 - ♦ 차이가 있는지?
 - ♦ 차이는 통계적으로 유의미한 것인지?
 - ♦ 차이는 얼마나 큰지?
 - ♦ 가정은 얼마나 정확한지?
 - ♦ 차이를 만드는 이유는? → 추론, 상상력을 동원하여 데이터의 의미를 분석

Study 2:

Evaluating MOVE Design

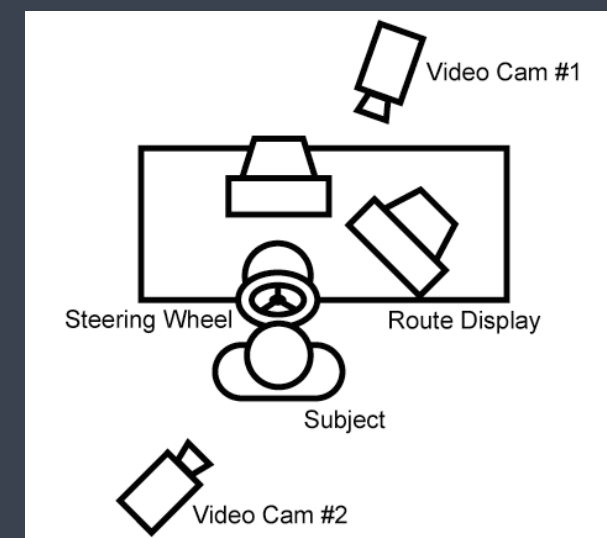
- ▶ Purpose: to evaluate feasibility and effectiveness of prototype design
- ▶ Map reading performance study
 - compare MOVE with the most optimized current static map (LineDrive)
- ▶ Hypothesis:
MOVE presentation methods can reduce the number of glances and fixation times to comprehend information
→ *reduce perceptual load*



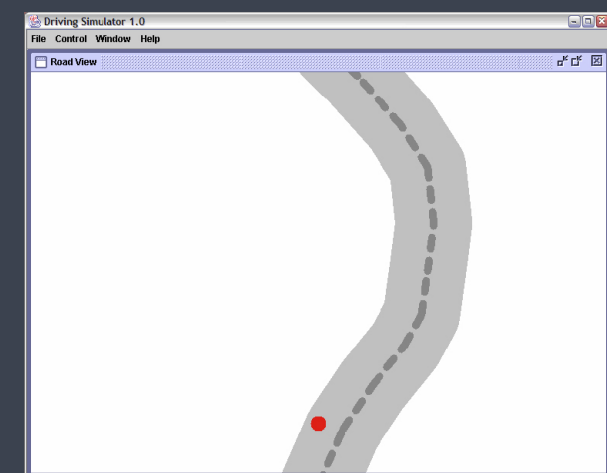
Study 2: Evaluating MOVE Design

Study Overview

- ▶ Dual task study
- ▶ Simple simulated driving task to saturate attention plus navigation display
- ▶ Subjects were told to maintain a central position on the road and prompted to glance at the navigation system and verbally report what was seen



Simulated driving task



Study 2: Evaluating MOVE Design

Procedure

- ▶ 20 participants (12M, 8F; aged 19-56)
- ▶ Conditions (counter balanced)
 - Baseline - check primary task performance without map display
 - Static Route Map: LineDrive
 - 4 MOVE presentation styles (ZC, ZC+R, R, ZC+O)
 - ZC w/o car location cursor - to compare with static map
- ▶ Measures
 - Total number of glances per task
 - Total fixation time
 - Average distance off from the road

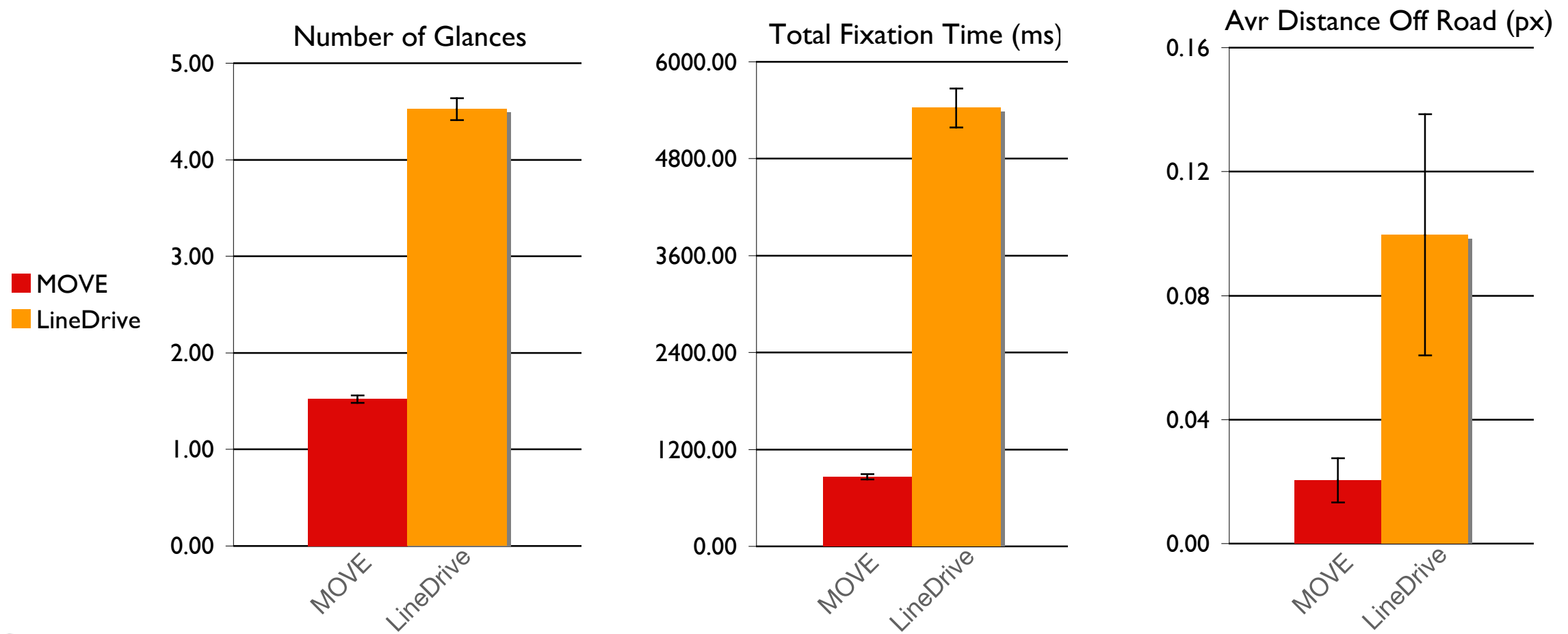


Study 2: Evaluating MOVE Design

Results & Discussion

MOVE vs. LineDrive (lower is better)

All significant at 5% significance level

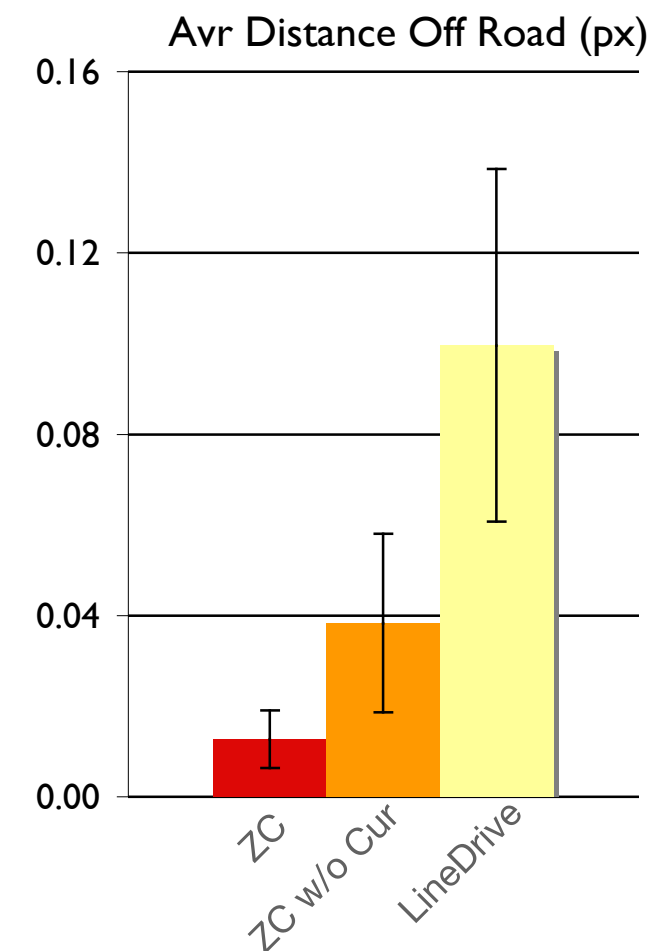
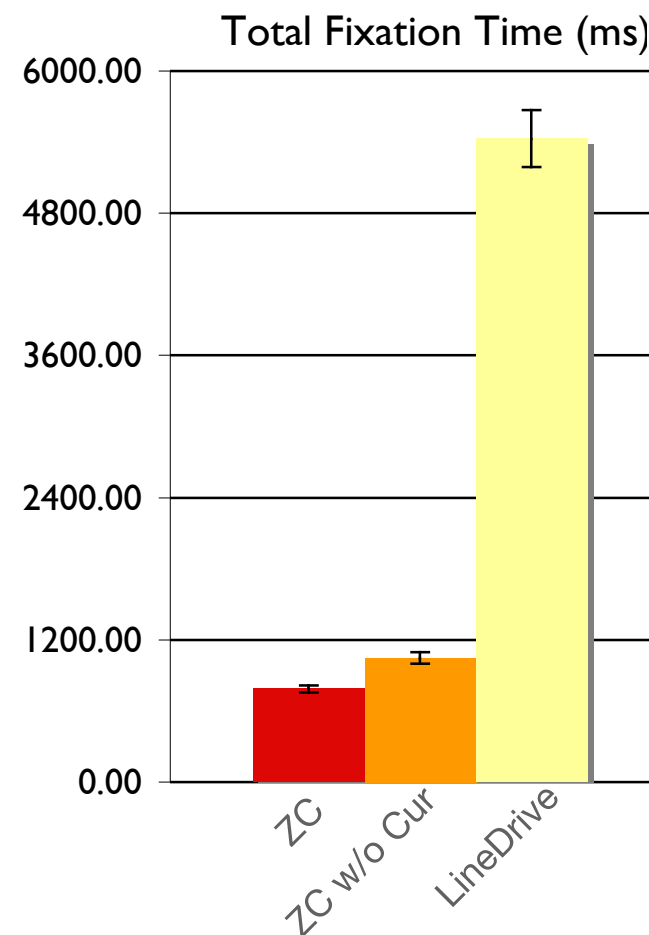
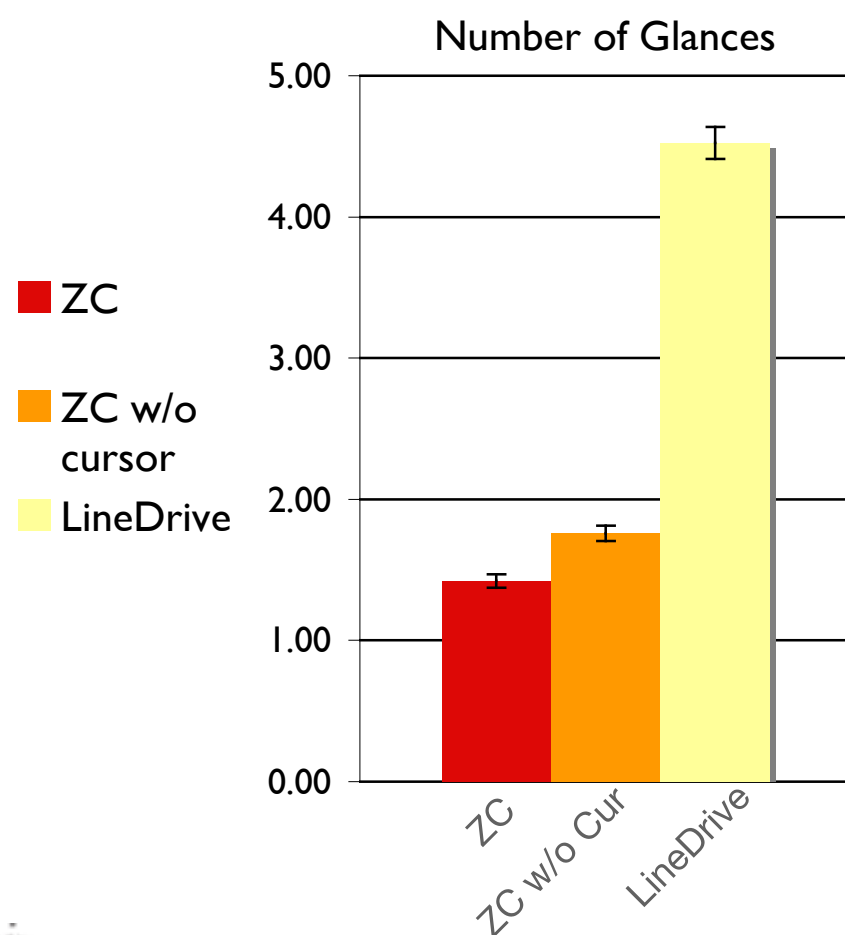


Study 2: Evaluating MOVE Design

Results & Discussion

Merit of Cursor (lower is better)

All significant at 5% significance level

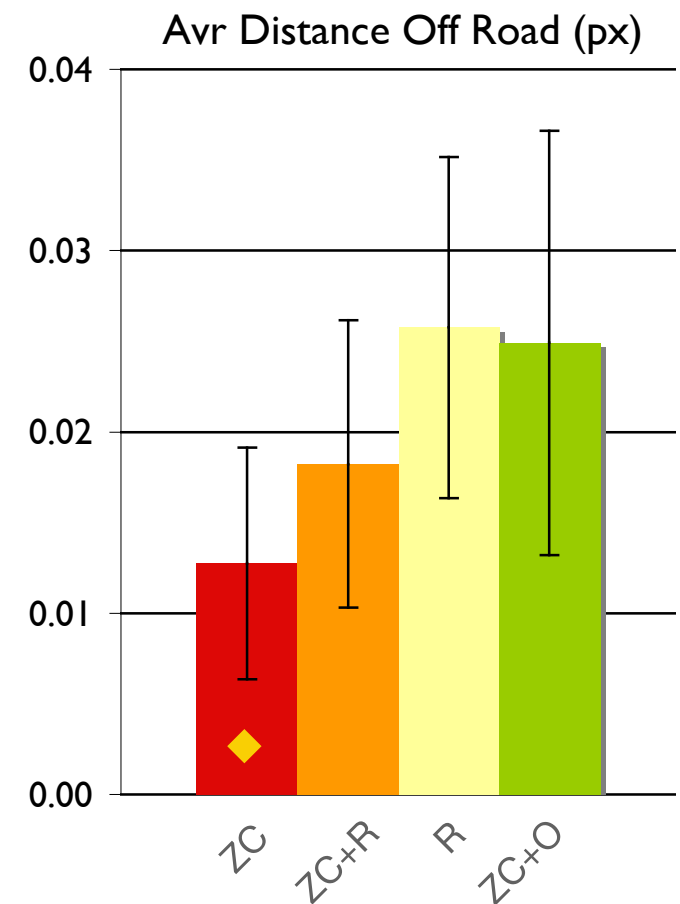
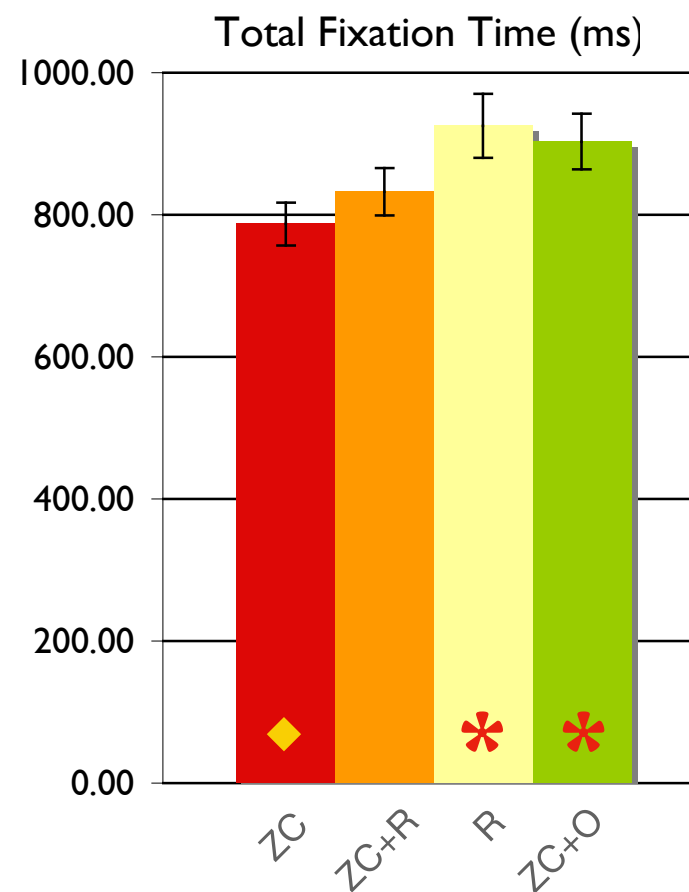
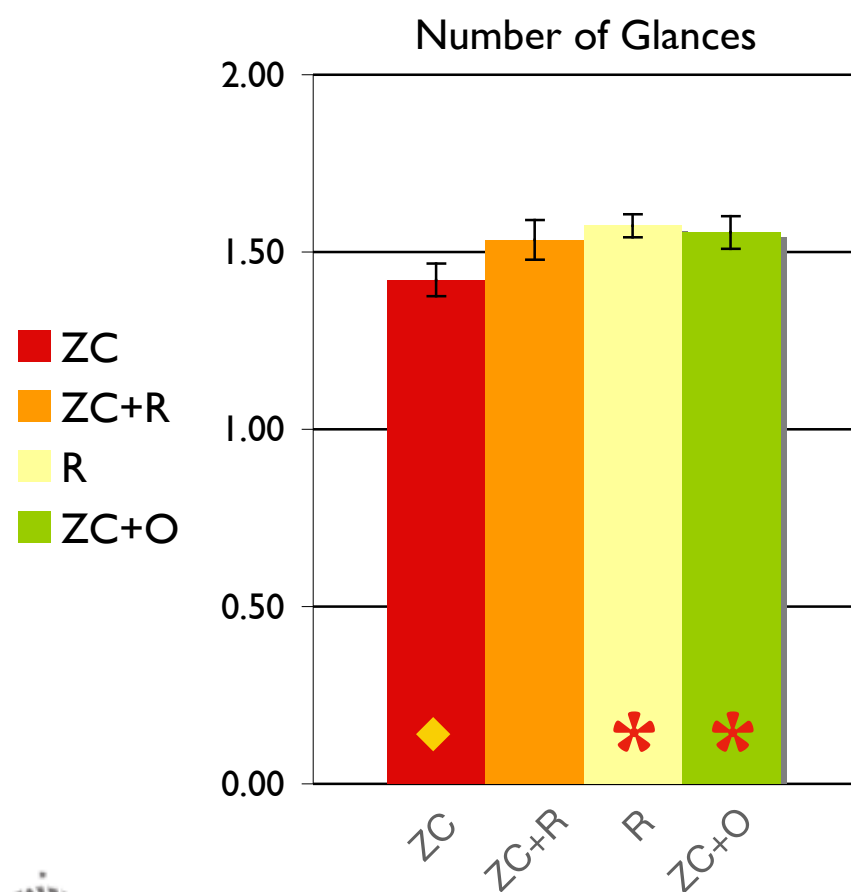


Study 2: Evaluating MOVE Design

Results & Discussion

MOVE 4 Presentation Styles (lower is better)

◆ used for baseline comparison
* All significant at 5% significance level



Brainstorming: User Study

- ✦ People who use a mouse and keyboard together will be faster to fill out a form than keyboard alone.
- ✦ 스터디 디자인을 해보자
 - ✦ Hypothesis?
 - ✦ Population?
 - ✦ Procedure?
 - ✦ Two types?
 - ✦ Between vs Within
 - ✦ Data Analysis?

Experimental Studies on Groups

- ✦ 그룹을 대상으로 하는 실험은 단일 유저를 대상으로 하는 것 보다 훨씬 어려움.
- ✦ 여러 문제점들이 발생
 - ✦ 피험자 그룹의 선정
 - ✦ 실험을 할 태스크의 선택
 - ✦ 데이터 수집
 - ✦ 분석

Subject Groups

- ♦ 많은 수의 피험자는 실험의 비용이 올라가고, variation이 커지며, 안정화된 데이터를 얻기까지의 시간이 많이 걸림.
- ♦ 많은 사람들의 시간을 맞추는 것도 어려움.
- ♦ 보통, 3-4개의 그룹을 대상으로 실험 → 그 이상은 실험을 운용하기가 어려워짐.

Data Gathering

- ✦ 그룹 실험의 데이터는 다양한 방법으로 수집 가능
- ✦ 몇대의 비디오 카메라를 설치해서 피험자들의 협업과정을 녹화, 분석
- ✦ 실험용 어플리케이션을 만들어 피험자들의 행위의 log를 기록
- ✦ 그룹으로 모아진 데이터에는 개개인의 개별적인 데이터가 포함되어 있으므로 synchronization 의 이슈가 발생할 수 있다. → time code 등의 사용으로 문제를 해결해야
- ✦ 그룹 프로젝트라 할지라도 개개인의 perspective 가 중요.

Physiological Methods

- Eye tracking
- Physiological measurement

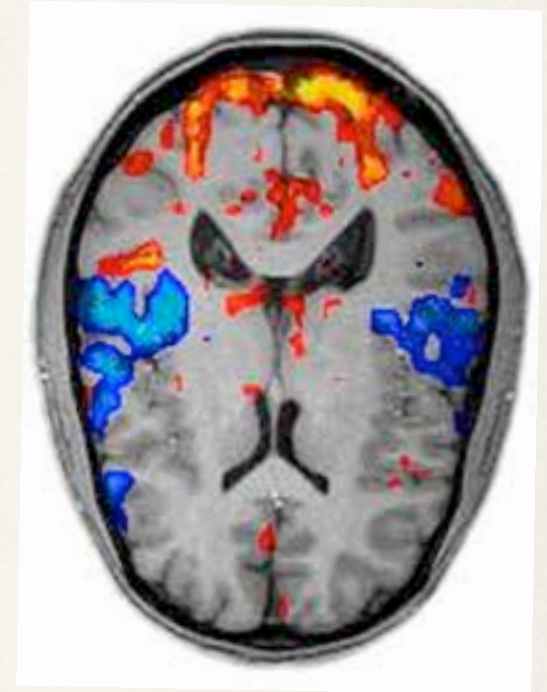
Eye Tracking

- ♦ 머리에 부착하거나 책상위에 올려놓고 화면에서의 눈의 위치를 추적하는 장치
- ♦ 눈의 움직임은 디스플레이가 요구하는 cognitive process 의 양을 반영한다
 - 눈이 계속 머무르는 곳은 무엇인가 하기 위해 cognitive process 가 이루어지는 곳
- ♦ 측정하는 것
 - ♦ fixation: 눈이 머무르는 위치. 시간과 횟수는 디스플레이를 바라볼 때의 어려움의 수준을 반영
 - ♦ scan path: 눈이 정보를 찾기 위해 움직이는 경로



Physiological Measurements

- ✦ 신체의 변화와 관련된 정서적 반응을 측정
- ✦ 사용자가 인터페이스를 쓸 때의 반응 (reaction)을 알고, 그 때의 감정 상태 (예: 불편하다, 어렵다 등) 을 파악하는데 도움이 될지도 모른다
- ✦ measurements
 - ✦ heart activity, including blood pressure, volume and pulse.
 - ✦ electrical activity in muscle: electromyogram (EMG)
 - ✦ electrical activity in brain: electroencephalogram (EEG)
 - ✦ flow of oxygenated blood in brain: functional magnetic resonance imaging (fMRI)



Questions...?
