Notes on Experimental Design

TAIM – M.EIC – 2023

Rui Rodrigues

Experiment Plan Checklist

- Goal of the tests
- Where, When, Duration
- Equipment, Software
- Initial State, System load
- Coordinator / observer / facilitator
- **Participants**
- **Test methods**
- **Tasks**
- **Data to collect**
- Help resources, Facilitator help
- Test completion criteria
- **Success criteria**
- **Experiment Protocol**
- **Pilot tests**

Experiment protocol (guide, script)

- Introduction and objectives
- Informed Consent form
- Pre-test questionnaire
- Tasks
  - Task-specific questionnaire(s)
- Post-test questionnaire
- Interview

Pilot tests

- Anyone
- 2 to 3 people
- Test the procedure
  - Duration
  - Instructions
  - Tasks
  - Questionnaire
  - Etc.

Participants

Participants - Who?

- Representatives of the target users (or approximate)
- Knowing the general tasks, domain and vocabulary
- Balance in gender, age, etc.
  (unless e.g. system is gender-specific)

Participants - How Many?

- Formative:
  - 1 user -> 33% of problems
  - 5 users -> 85%
  - 15 users -> 99%
  - 3 iterations with 5 users -> better than 1 x 15

- **Summative**
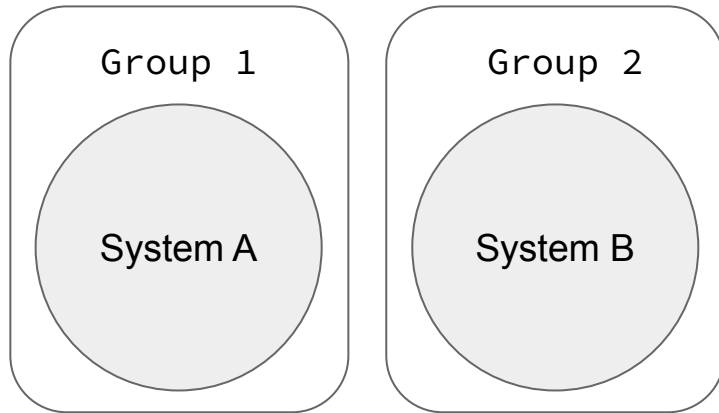 - **At least 10 to 20, to have a sample large enough for statistics**
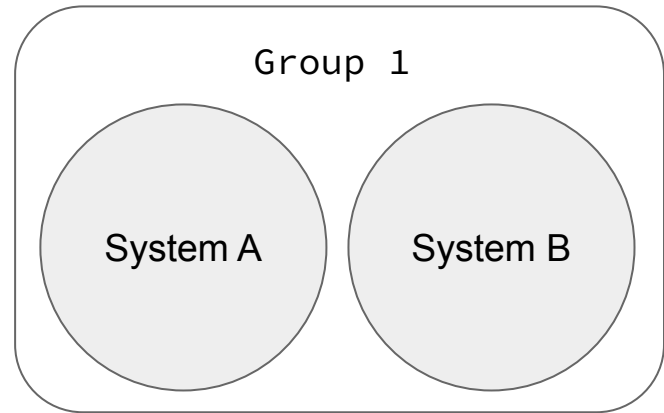
Methods

Methods

- **Formative Assessment (usability problems)**
  - Direct observation in the laboratory
  - Think aloud
  - Wizard-of-Oz
  - Interviews / Questionnaires
- **Summative Assessment (performance measures)**
  - Tests of usability
  - A/B tests

A/B tests

Between-subjects
(inter-groups)

Within-subjects
(intra-group)

Group 1

System A

Group 2

System B

Group 1

System A

System B

Order Effects On Intra-group

- Order may play a role (e.g. learning A, facilitates B)
- So elements should have systems in different orders
  - Randomized
  - Sequenced
  - Latin-Squares

Variables to be tested

- **Dependent variables**
  - Its value depends on the system to be tested
  - Variables measured in the study (e.g. time, errors, SUS)
  - Related to the purpose of the prototype
- **Independent Variables**
  - They do not depend on the variables we are measuring
  - Solution characteristics (e.g. layout, color, etc.)
  - Characteristics of the participants (e.g. age, etc.)

Tasks

Tasks

- Real and representative

- What and not how

- Specific

- Mixes complexities

- Summative (comparative) evaluation

  a. Do not favor one of the solutions

  b. Use the same tasks

Examples

- **Wrong:**
  Choose the option Uber to buy a trip and then write 'Aliados' and select the first option

- **Right:**
  Buy a trip through the Uber service to Aliados

Data

Data types

- **Objective**
  - They do not depend on the predisposition (bias) inherent to the human being
  - E.g., time, errors, heart rate, etc.

- **Subjective**
  - Enhances the user's perception
  - E.g., preference, SUS, etc.

Data types

- **Quantitative (quantity, specific and measurable)**
  - Did you complete the task? Yes/No
  - How long did it take?
  - How many errors?
  - Which one did you prefer? A or B?
- **Qualitative (quality, "open")**
  - "What did you like most about the experience?"
  - "What do you think of the home screen?"

Data Collection

- **Direct observation**
  - Observed users performing tasks
  - In person or video

- **Indirect observation**
  - Diaries – manual
  - Interaction logs – automatic

Instrumentation

- Timestamp

- Player location

- Nr. bullets

- Events (shot, hit, died, ...)

- Where are you aiming?

Measuring Usability

Usability Measures

- Time to complete task
- Nr. of mistakes made
- Nr. of completed tasks
- Nr. of clicks
- Nr. of help queries
- User satisfaction
- Etc.

Standard usability Questionnaires

Post-test

- SUS
  - System Usability Scale
- UEQ
  - User Experience Questionnaire
- NASA TLX
  - Task Load Index
- CSUQ
  - Computer System Usability Questionnaire

Post-task

- SEQ
  - Single Ease Question
- ASQ
  - After-Scenario Questionnaire

System Usability Scale (SUS)

	The System Usability Scale Standard Version	Strongly disagree				Strongly agree
		1	2	3	4	5
1	I think that I would like to use this system.	O	O	O	O	O
2	I found the system unnecessarily complex.	O	O	O	O	O
3	I thought the system was easy to use.	O	O	O	O	O
4	I think that I would need the support of a technical person to be able to use this system.	O	O	O	O	O
5	I found the various functions in the system were well integrated.	O	O	O	O	O
6	I thought there was too much inconsistency in this system.	O	O	O	O	O
7	I would imagine that most people would learn to use this system very quickly.	O	O	O	O	O
8	I found the system very cumbersome to use.	O	O	O	O	O
9	I felt very confident using the system.	O	O	O	O	O
10	I needed to learn a lot of things before I could get going with this system.	O	O	O	O	O

Notes on Experimental Design

TAIM - M.EIC - 2023

Rui Rodrigues