User Interaction COMPSCI2031

Dr Florian Weidner Florian.Weidner@glasgow.ac.uk
Dr Ilyena Hirskyj-Douglas ilyena.hirskyj-douglas@glasgow.ac.uk

Recap: What we did last yesterday

- Models of Interaction in Human-Computer Interaction
- Fitts Law Task
- A/B testing
- Reading: Models of Interaction

User Interaction Topics

- ✅ HCI History and Introduction
- ✅ Usability and Heuristics
- ✅ Heuristic Evaluation and Human Cognition
- ✅ Human Perception and Capabilities
- ✅ Experimental Design & Variables Research
- ✅ Personas and Scenarios
- ✅ Surveys in HCI
- ✅ Ethnography
- ✅ Statistical Methods
- ✅ Theories in HCI & User-Centered Design
- ✅ Models of Interaction
12. Large Scale and Mobile HCI
13. Various Users and Ethics
14. Revision & Example Exams & Quiz

Large-Scale & Mobile HCI Studies

Lecture 12

Large-Scale Mobile HCI Studies

- Mobile HCI studies take many forms:
  - e.g. text entry, gestures, AR, usage studies, privacy…
- How do you perform studies?
- Quantitative analysis
  - e.g. time taken to complete task / error rates…
- Qualitative analysis
  - e.g. interviews, ethnography, opinions of experience…

Into the wild

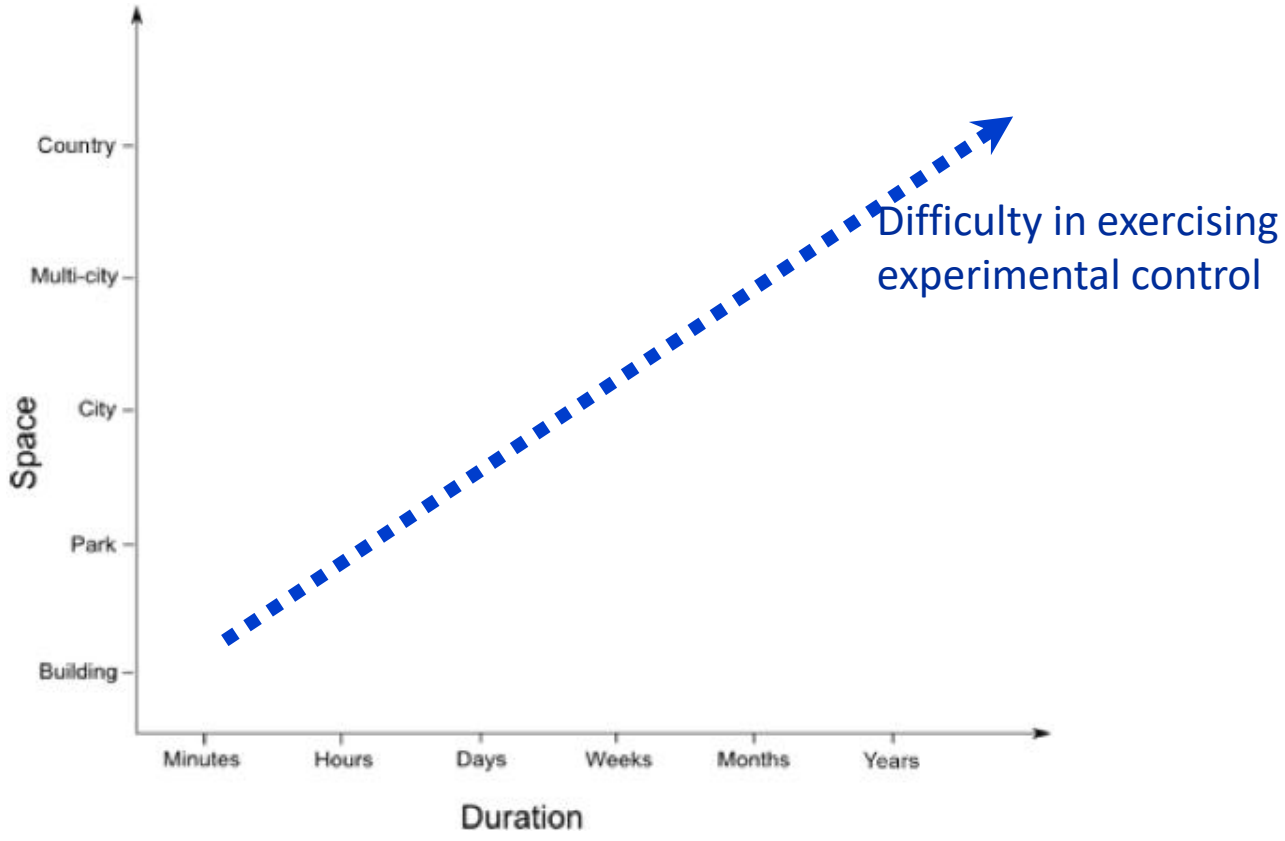
- Early / ‘traditional’ experiments all done in the lab
- Easy to observe, control, eliminate confounding variables
- But?
  - Unrepresentative?
  - Technology’s eventual intended context of use?

→ Recent studies often performed in realistic settings: ‘in the wild’

Into the wild

- What forms of evaluation are used in the wild?
- Even if outside a lab, often still direct observation
  → Videos, interviews, …
- Often still using evaluator-supplied hardware

Challenges:
Space & Duration of Trials



Problems?

- Control vs. Scale?
- Resources?

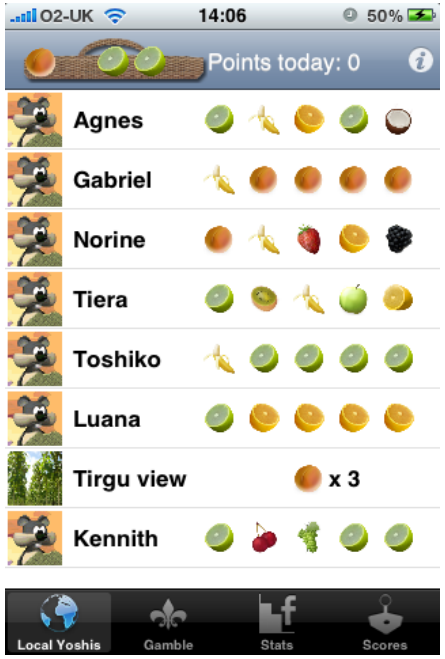
How? Do research via app stores!

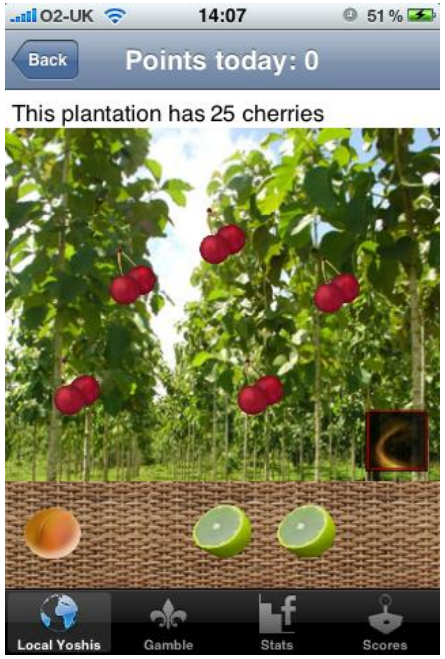
- Put software you want to study on app stores.
- Participants using their own devices.
  - Already experts with hardware; no training
  - No extra device to carry; already have with them always
  - No fixed end date
- Potentially very large numbers of users, from all over the world.
  - Chosen to use the app; more representative?

Research via app stores

- Problematic? → Control vs. Scale
- Not all good news?
  - Don’t meet users
  - Can’t directly observe users
  - Can we still gather qualitative data?
  - Internal vs external validity
  - Additional ethical challenges?

Example - Hungry Yoshi



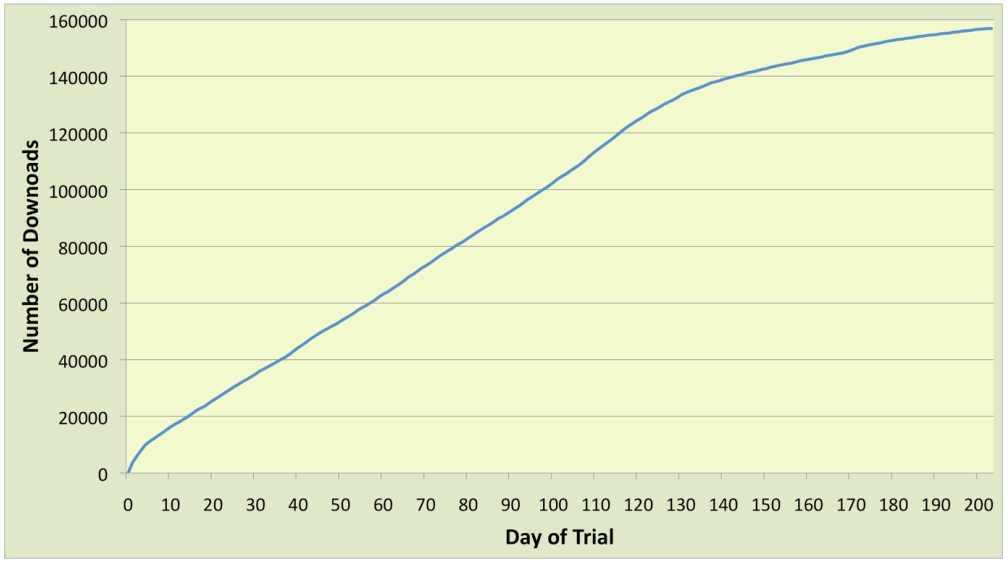


Game using Wi-Fi access points as game resource

For reference: https://doi.org/10.1007/978-3-642-12654-3_13

Example - Hungry Yoshi

- Investigating use of app stores in running mobile HCI trials
- > 300,000 downloads; new users quite steadily acquired
- Only worked for around 2 years, as iOS updates blocked functionality



Example - Hungry Yoshi

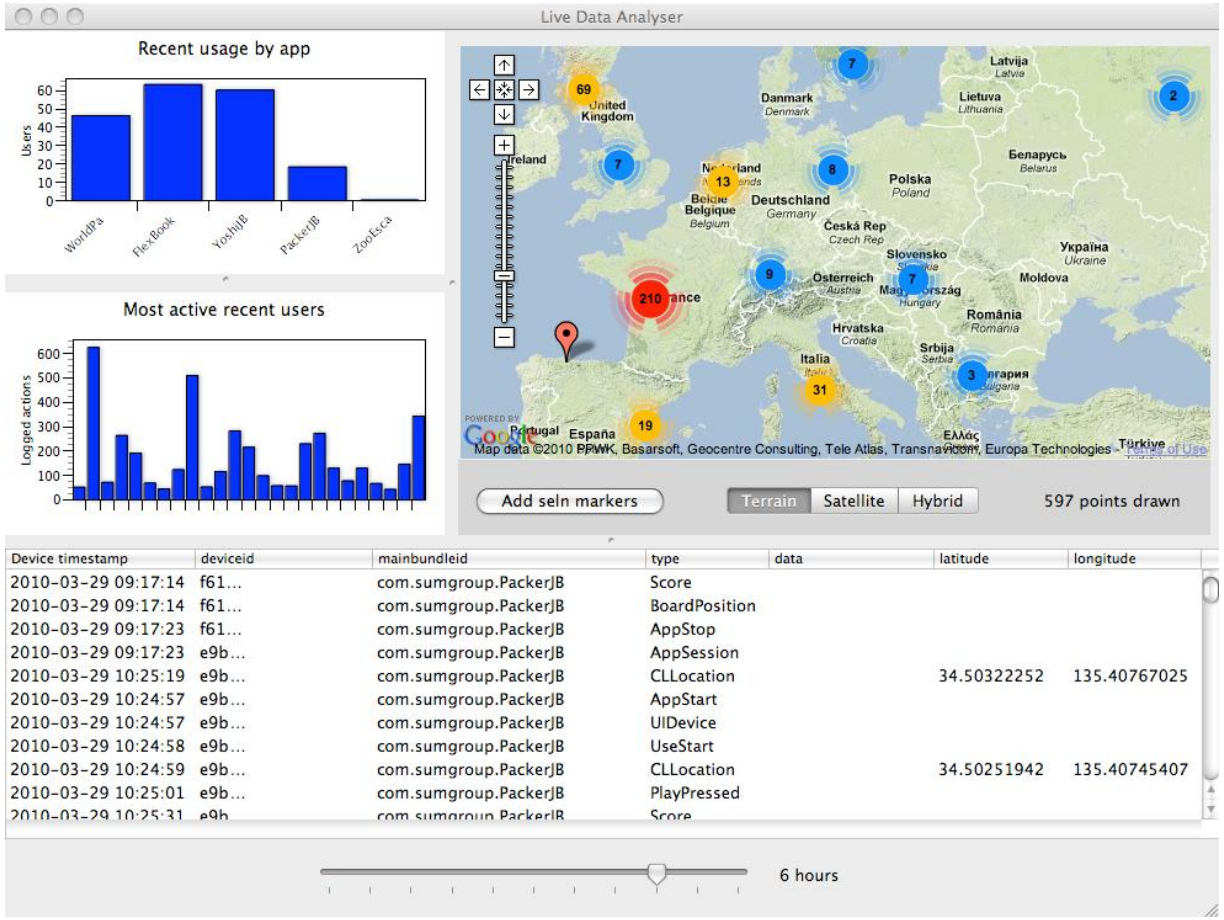
- Global user base
- Only have locations from those users who agree to supply it

University of Glasgow

Data logging

- Can tell how/if people are using our software around the world?
- Need to record (“log”) info on use while apps are running
- Hungry Yoshi used logging framework with phone and server parts
- Capture user interactions (button taps, screen changes), device/sensor data (accelerometer, Wi-Fi connections)
- Write to locally-stored files on device
- Opportunistic uploading over Internet

Data visualization



- Have tools!
- SGVis - live analysis of usage

Qualitative Evaluation

- Would this type of trial be possible at a distance?
- Questionnaires?
  - Answers selected from lists or type into textbox
  - Other types of ‘task’, e.g. become a Facebook friend
  - 19% of players responded
  - Server-side, so instant updates

+ Paid telephone interviews

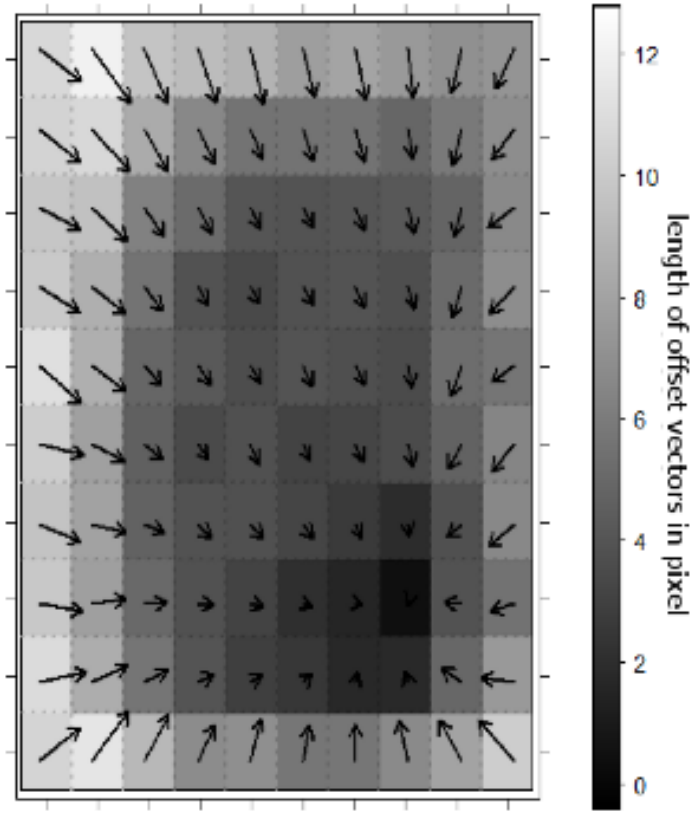
Example: Hit It!

- Android game: touch objects on the screen
- 91,731 installations
- 120,626,225 touch events
- Looked at error rates for different sizes and locations (Fitt's Law?)

For reference: https://dl.acm.org/doi/10.1145/2037373.2037395

Example: Hit It!

- Found that touch positions are skewed
- Could create function that shifts touch input to compensate
- Updated game to use compensation
  - Error reduced by 7.8%



Large Scale Trials: Difficulties

- Verification of user info…
  - Are people telling the truth or just ticking random buttons?
  - About age, gender, opinions...?
  - A way to know?
- Compatibility: Trial software installed on large variety of devices
  - Especially on Android
  - Differences in OS, CPU, screen sizes…
  - All could be confounding variables

Large Scale Trials: Difficulties

- Collecting qualitative data is difficult
  - Very short questionnaire answers
  - Phone / Skype calls
    - Most users probably don’t speak your language
    - Difficult to arrange - time zones
- Mass of quantitative data; harder to get qualitative
  - Could see what was happening, but not why

Solution? Hybrid Methodology

- Hybrid approach: combining ‘mass participation’ & local deployments
- Concurrent large scale and small-scale trials
- App released to general public and local users
- Some aspects of trial best suited to each group
- More solid ethical practice?

Example: Predictor

- World Cup Predictor app
- Released 1 week before football World Cup
  - 11 locally-recruited users
  - 10,806 through app store

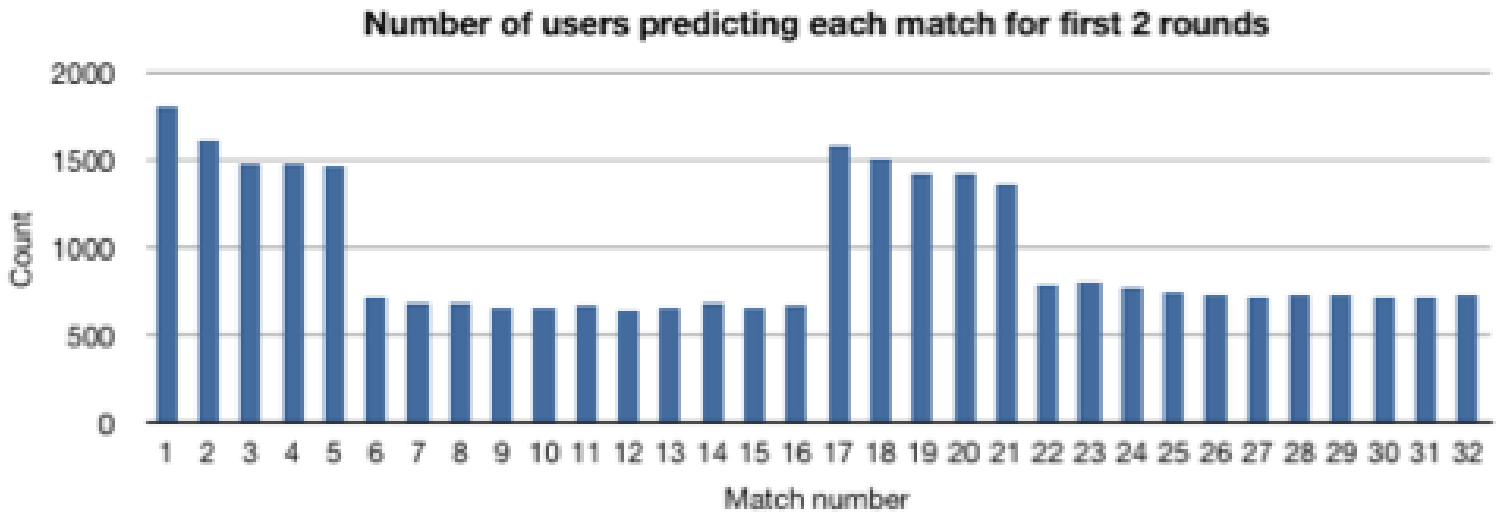
Screenshot of

For reference: https://dl.acm.org/doi/10.1145/2207676.2208588

Hybrid Trials: Benefits

1. Use the small to explain the large
2. Use the large to verify the small
3. More solid ethical practice

Use the Small to Explain the Large



Use the Large to Verify the Small

- If a system is trialled among small group of locally-recruited participants...
  - Do results generalise to population at large?
  - ‘Outlier’ users
    - Do we have user(s) showing unusual behaviour?
    - Could skew the results of the study
- Experimenter effects?
  - Subtle conscious or unconscious cues an experimenter gives participants
  - Could affect users’ performance in the trial
  - Less likely in large-scale trial?
    - User interaction with evaluators generally far lower

Use the Large to Verify the Small

- Looking at one feature of app: head-to-head challenge
- Local users:
  - Head-to-head uptake: 73%
- Global users:
  - Head-to-head uptake: 0.8%
- Local trial alone would have led to misleading results

University of Glasgow

Questions?
Comments?
Concerns?

User Interaction Topics

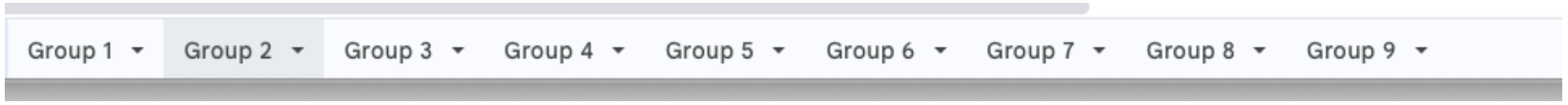
- ✅ HCI History and Introduction
- ✅ Usability and Heuristics
- ✅ Heuristic Evaluation and Human Cognition
- ✅ Human Perception and Capabilities
- ✅ Experimental Design & Variables Research
- ✅ Personas and Scenarios
- ✅ Surveys in HCI
- ✅ Ethnography
- ✅ Statistical Methods
- ✅ Theories in HCI & User-Centered Design
- ✅ Models of Interaction
- ✅ Large Scale and Mobile HCI
13. Various Users and Ethics
14. Revision & Example Exams & Quiz

Large vs Small Studies

- Open the following google sheet link (link on teams/Moodle): https://docs.google.com/spreadsheets/d/1A8Vi6E0AWeB2hXj0wIYDPJbs2clbAH7C4LozMO1mIJ4/edit?usp=sharing

- Step 1: Put the data you collected yesterday from Fitt's study into the link into the 'Everyone data' tab [5 mins]

- Step 2: In your group look at the Fitts study data, copy this into your individual group and have a go at some analysis. For example:
  - If you order by device, and work out the mean, does this match your own data from yesterday?
  - Is the Standard Deviation (SD) and mean different in error timings than your own group or individual score?
  - Is your data skewed? (Hint: Test for normality, this can be done in Excel, here is a guide).
  - What statistics could you do to analyse this data? (Hint: T-test/Welch Test, which can also be done in Excel)
  - Write down your findings in your team space



- [30 mins]

Discussion: Large vs Small Studies

University of Glasgow

No Reading ☺
Please catch up if you are behind