



## "4-second bursts" Quasi-experiment example



**Why: Motivation**  
**What: Principles**  
**How: Practices**



## My objective

- *is to encourage you to do field experiments in your student and professional careers,*
- *to show that its possible, that it makes sense*
- *to help you distinguish poor research from proper one*



## "Post-modern usability"

**Arnold Lund, Microsoft Research**

"If Jakob is right and everyone should and easily can be trained in the craft, how valuable can the profession be?"

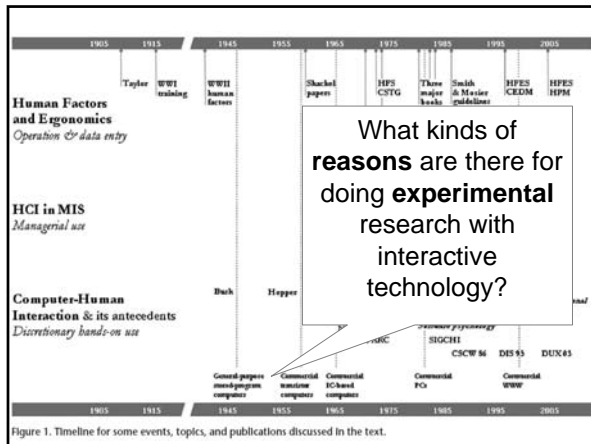
"Usability is a real discipline and a few hours of training, a user aid with heuristics, and a book with interesting stories of usability problems doesn't turn an engineer, marketing person, or manager into a usability professional."

"The technology needs to be treated as the extension and expression of the user"



## Why: Motivations





## The conundrum of interactive technology

Almost anything can be imagined, yet there are hard limitations to what is possible.



## Research interests in evaluating prototypes

To gauge

- Performance against an **ideal**
- **Difference** to a previous/competing version
- **Suitability** for the purpose it was designed for
- **Change** in users' lives
- User **experience / approval**

To help in decision-making



Reason must approach nature in order to be taught by it. It must not, however, do so in the character of a **pupil** who listens to everything that the teacher chooses to say, but of an appointed **judge** who compels the witness to answer questions which he has himself formulated.

(Emmanuel Kant, Critique of Pure Reason, 1790, p. 20)



**What:**

## Principles of quasi-experimenting



## What is an experiment anyway?

To experiment means to cause a change in the object of study and observe the consequences



## Basic terminology

- Phenomenon
  - An event or object or feature in the world, or a relationship thereof, that is independent of the researcher
- Experiment
  - Attempt to find out causal relationships in a phenomenon by causing a controlled change in it
- Independent variable
  - Variations in the thing that is being manipulated by the experimenter
- Dependent variable
  - Measures in the outcome, effect of the manipulation
- Operationalization
  - Concrete way to implement the measurement of phenomenon and causing the change



## Types of experiments

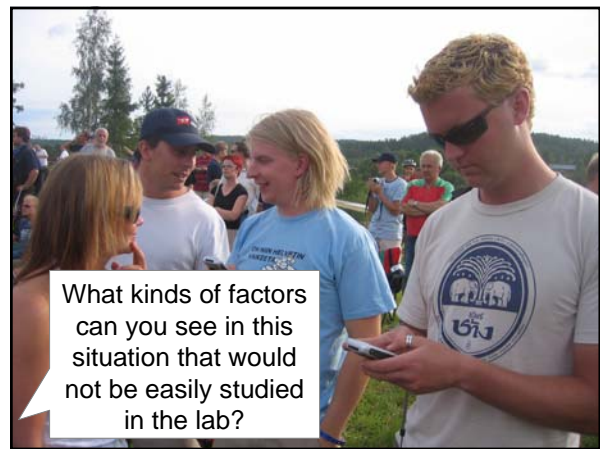
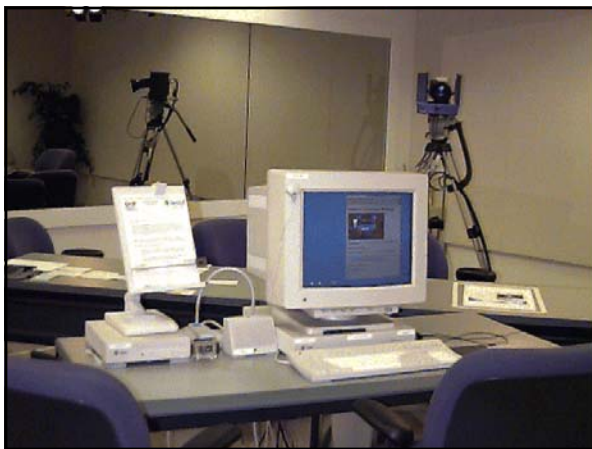
1. **Laboratory experiments** are used in usability testing, for example.
2. **Analogue experiments** deploy simulations and emulation of real-world conditions to improve generalizability. For example, the 1990s trend of decorating usability labs like living rooms was an attempt to stage aspects of real use situations.
3. In **quasi-experiments**, an experimental intervention is carried out in the field even in the face of limited randomization and control.
4. **Natural experiments** are “after-the-fact” quasi-experiments, wherein the variation of a causal agent has taken place naturally. For example, two naturally formed user groups could be compared in terms of some variable that differs between them—say, comparing adolescents to adults in their adoption of a mobile service.



What type of an experiment is this?  
(Amazon changes layout for 10% of users, records shopping tendencies)

## When to go to the field?

1. **The environment affects interaction through the user.** For example, a user can interrupt an interactive task to respond to an abrupt event caused by a colleague.
2. **The environment affects interaction through the computer.** For example, a context-aware service can represent surrounding resources such as printers and people, interpreting their presence from network or Bluetooth data.
3. **The relationship can also be two-directional.** For example, in use of a location-aware mobile map, the user's movement changes the state of the environment as well as the state of the map.



## Ceteris paribus

Experiments rely on the ceteris paribus logic:

*Since all other things were equal, we can infer that X caused Y.*



## Quasi-experiments

The theory of quasi-experiments was founded on the admission of the *imperfection* of field experiments.

The point is that the degree of control is limited *and should be treated as such*.



## Control

- **Control** means organizing the experiment so that the result can be attributed to the independent variable and not to something else
- Types of control
  - Direct intervention
  - Pre-selection
  - Guidance
- All experiments are limited
- Understanding, explicating, and addressing these limitations is the business of quasi-experimenting



## Validity

1. **Statistical inference**
  - Can we infer that the observed change is reliable?
2. **Internal**
  - Was it the IV that caused the observed change in the DV?
3. **Construct**
  - Is the IV really the cause of the DV?
4. **External**
  - Do the results generalize?



## Typical threats to validity

- Random irrelevancies in data due to “noise” or inaccuracies of sampling
- Systematic omissions in data
- Systematic biases in certain contexts of use
- Systematic interactions between user groups and contexts
- Systematic bias due to the intervention
- Systematic carryover effects
- Incorrect statistical tests
- Locality of results



**Designing a quasi-experiment  
about addressing  
nuisance factors**



What kind of nuisance factors could hamper a lab-based usability test of the new TKK website?

### Controlling nuisance variables

1. Keep constant for everybody
2. Vary systematically, include in the experimental design as an IV
3. Randomize
4. Randomize, analyze effects post hoc

Common techniques

- Single-blind
- Double-blind
- Deception
- Disguised experiment
- Multiple researchers
- Postexperimental debriefing
- Experimenter-expectancy control groups
- Unrelated-experiment technique
- Effect "simulator" group

### Some other typical nuisance factors

- Experimenter-expectancy effect
  - Experiment guides/leads behavior
- Demand characteristics
  - Subject guesses the purpose of the experiment
- Subject-predisposition effects
  - "Being a good subject"
  - "Screw your effect"
  - "Evaluation apprehension"
  - "Faithful subjects"
- Placebo

You arrange this experiment at TKK with high school students. How would you control for the evaluation apprehension effect?

### How: Practices

### General stages

- Identifying a problem
- Identifying goals for the study
- Experimental design
- Implementation
- Analysis
- Interpretation

## Research plan

= Plan of assigning subjects to experimental conditions and statistical inference of that experimental design

In quasi-experimenting, elimination of nuisance factors is emphasized



## 2. Method

### 2.1. Participants

Nine experienced web users (one female, eight males) volunteered for the *experimental group*. All participants used the WWW daily (at least for one hour per day) in their work, had at least college education, and five or more years of daily experience with graphical browsers. They were naïve to the purpose of the experiment. All participants had normal or corrected-to-normal vision. Nine other experienced web users, matched in age and years of web-related experience to the experimental group, volunteered for the *control group*.

### 2.2. Materials

The materials consisted of 15 web pages with financial, political, or educational content. In order to make sure that the participants, who were all from Finland, would have no prior experience with the pages, they were selected from the domains of Australia and New Zealand. The content area of all materials always included text, a maximum of two graphical elements, and one to three headers (in bold typeface or in bigger font size than main text). All link panels were arranged verti-

## Pilot

### Why?

- Unforeseen events during a trial
- Calibration of measurements
- Calibration of IV
- Assessing the costs of the experiment
- Assessing N



## Emerging equipment

- Mobile ESM
- Multi-source recording equipment
- Logging software



## Ways to minimize costs

- Field research is 2-8 times more expensive than lab research
- How can you minimize costs?
  - Minimize per-participant costs. Maximize number of data points acquired per subjects. Use multiple repetitions and multiple dependent variables.
  - Minimize equipment costs. Get your own.
  - Minimize training costs. Dedicate employees for this work.
  - ...



**A few more principles to keep in mind...**



## Normal criteria

- Objective
  - The result does not depend on the opinion of the researcher
- Replicable
  - Somebody else could run the same experiment and get the same results
- Public
  - Anybody can take your results and examine them against your procedure
- Self-correcting



“In appraising the merits of theories, it is more important to ask whether they constitute **adequate solutions to significant problems**

than it is to ask whether they are 'true,' 'corroborated,' 'well-confirmed' or otherwise justifiable within the framework of contemporary epistemology.” (Laudan, 1977, p. 14).



## Goals for any research

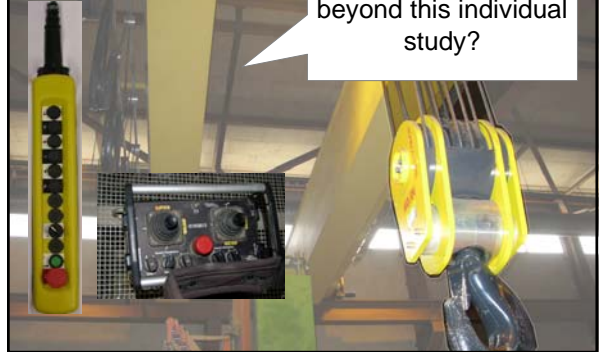
To provide a solution

- to a problem that is **important**
- that can solve **many** other problems
- that solves those problems **effectively**

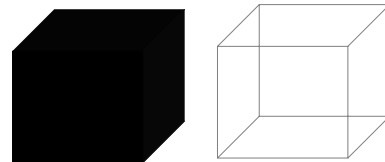


## Crane control interface

How could your study help Konecranes beyond this individual study?



## Locality of results is a typical sin





You're planning a study of an information kiosk at Kämp. You are advised to use logging and questionnaires. What kind of **qualitative** data could you combine to improve your understanding of walk-up-and-use behavior?

