


SLAVEN, JANEY FOLK, GREGORY D. ARNONE, RUSSELL BEALE
HUMAN-COMPUTER INTERACTION

Evaluating through user Participation

SLAVEN, JANEY FOLK, GREGORY D. ARNONE, RUSSELL BEALE
HUMAN-COMPUTER INTERACTION

Laboratory studies

- Advantages:
 - specialist equipment available
 - uninterrupted environment
- Disadvantages:
 - lack of context
 - difficult to observe several users cooperating
- Appropriate
 - if system location is dangerous or impractical for constrained single user systems to allow controlled manipulation of use



SLAVEN, JANEY FOLK, GREGORY D. ARNONE, RUSSELL BEALE
HUMAN-COMPUTER INTERACTION

Field Studies

- Advantages:
 - natural environment
 - context retained (though observation may alter it)
 - longitudinal studies possible
- Disadvantages:
 - distractions
 - noise
- Appropriate
 - where context is crucial for longitudinal studies

SLAVEN, JANEY FOLK, GREGORY D. ARNONE, RUSSELL BEALE
HUMAN-COMPUTER INTERACTION

Evaluating Implementations

Requires an artefact:
simulation, prototype,
full implementation

Experimental evaluation

- controlled evaluation of specific aspects of interactive behaviour
- evaluator chooses hypothesis to be tested
- a number of experimental conditions are considered which differ only in the value of some controlled variable.
- changes in behavioural measure are attributed to different conditions

Experimental factors

- Subjects
 - who – representative, sufficient sample
 - not the programmer friend, boss, etc.
 - huge variability in effectiveness (e.g., programmers)
- Variables
 - things to modify and measure
- Hypothesis
 - what you'd like to show
- Experimental design
 - how you are going to do it

Variables

- independent variable (IV)
 - characteristic changed to produce different conditions
 - e.g. interface style, number of menu items
- dependent variable (DV)
 - characteristics measured in the experiment
 - e.g. time taken, number of errors.

Hypothesis

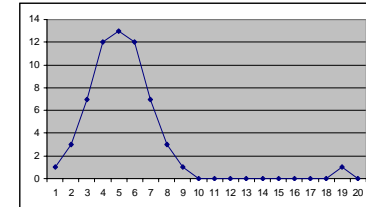
- prediction of outcome
 - framed in terms of IV and DV
 - e.g. "error rate will increase as font size decreases"
- null hypothesis:
 - states no difference between conditions
 - aim is to disprove this
 - e.g. null hyp. = "no change with font size"

Experimental design

- within groups design
 - each subject performs experiment under each condition.
 - transfer of learning possible
 - less costly and less likely to suffer from user variation.
- between groups design
 - each subject performs under only one condition
 - no transfer of learning
 - more users required
 - variation can bias results.

Analysis of data

- Before you start to do any statistics:
 - look at data (e.g. average=5.25 – but 4.9 without outlier)
 - save original data
- Choice of statistical technique depends on
 - type of data
 - information required
- Type of data
 - discrete
 - finite number of values
 - continuous
 - any value



Analysis - types of test

- parametric
 - assume normal distribution
 - robust
 - powerful
- non-parametric
 - do not assume normal distribution
 - less powerful
 - more reliable
- contingency table
 - classify data by discrete attributes
 - count number of data items in each group

Analysis of data (cont.)

- What information is required?
 - is there a difference?
 - how big is the difference?
 - how accurate is the estimate?
- Parametric and non-parametric tests mainly address first of these

Experimental studies on groups

More difficult than single-user experiments

Problems with:

- subject groups
- choice of task
- data gathering
- analysis

Subject groups

larger number of subjects
⇒ more expensive

longer time to 'settle down'
... even more variation!

difficult to timetable

so ... often only three or four groups

The task

must encourage cooperation

perhaps involve multiple channels

options:

- creative task e.g. 'write a short report on ...'
- decision games e.g. desert survival task
- control task e.g. ARKola bottling plant

Data gathering

several video cameras
+ direct logging of application

problems:

- synchronisation
- sheer volume!

one solution:

- record from each perspective

Analysis

N.B. vast variation between groups

solutions:

- within groups experiments
- micro-analysis (e.g., gaps in speech)
- anecdotal and qualitative analysis

look at interactions between group and media

controlled experiments may 'waste' resources!

Field studies

Experiments dominated by group formation

Field studies more realistic:

distributed cognition ⇒ work studied in context
real action is *situated action*
physical and social environment both crucial

Contrast:

psychology – controlled experiment
sociology and anthropology – open study and rich data