Evaluation and Peak Experience

Alan Dix
Lancaster University
http://www.hcibook.com/alan/
http://www.alandix.com/

evaluation
you’ve designed it, but is it right?

purpose

Two types of evaluation

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>formative</td>
<td>improve a design</td>
</tr>
<tr>
<td>summative</td>
<td>say “this is good”</td>
</tr>
<tr>
<td>investigative</td>
<td>gain understanding</td>
</tr>
</tbody>
</table>

points of comparison

• measures:
  - average satisfaction 3.2 on a 5 point scale
  - time to complete task in range 13.2–27.8 seconds
  - good or bad?
• need a point of comparison
  - but what?
  - self, similar system, created or real??
  - think purpose ...
• what constitutes a ‘control’
  - think!!

types of knowledge

• descriptive
  - explaining what happened
• predictive
  - saying what will happen
  - cause → effect
  - where science often ends
• synthetic
  - working out what to do to make what you want happen
  - effect → cause
  - design and engineering

different kinds of evaluation

endless arguments
  - quantitative vs. qualitative
  - in the lab vs. in the wild
  - experts vs. real users (vs UG students!)
really
  - combine methods
    - e.g. quantitative – what is true & qualitative – why
  - what is appropriate and possible
mechanism

- reduction reconstruction
  - formal hypothesis testing
  + may be qualitative too
  - more scientific precision
- wholistic analytic
  - field studies, ethnographies
  + 'end to end' experiments
  - more ecological validity

when does it end?

in a world of perpetual beta ...

real use is the ultimate evaluation

- logging, bug reporting, etc.
- how do people really use the product?
- are some features never used?

from evaluation to validation

validating work

your work

evaluation

- experiments
- user studies
- peer review

generative artefacts

artefact

evaluation

- toolkits
- devices
- interfaces
- guidelines
- methodologies

singularity
different people
different situations
different designers

(pure) evaluation of generative artefacts is methodologically unsound
validating work

justification vs. validation

... breaking the rules
interaction design at the edge

designing for peak experience

kinds of experience
(e.g. mobile apps)

• can’t do it without …
  – phone itself, address book

• can’t live without …
  – compelling apps
    e.g. iPhone snooker, photo sharing
    ... designing for peak experience ...

baked beans
  – staple food
  – good enough for everyone
    e.g. word processor – corporate decisions

Mars bar
  – favourite chocolate bar
  – best for some
    e.g. video game, web email – personal decisions

• different disciplines
  – mathematics: proof = justification
  – medicine: drug trials = evaluation

• combine them:
  – look for weakness in justification
  – focus evaluation there

• expert opinion
• previous research
• new experiments

• experiments
• user studies
• peer review

yea, I know
very cultural
designing for peak experience

who wins?

- good enough product
- peak product
- different users
- quality of experience

Good enough products never win for any user, some peak product always better

technology driven design?

user needs

user-centred design starts here

product

technology

route to market

Product must be:
- wanted by users
- achievable with technology
- have route to ‘market’ (not necessarily selling)
- all need to be there
- ... but can start anywhere

designing for peak experience

how to do it:
- traditional interface design
  - user profiles, central personas, average and typical, process and methods, from need to solution
- design for peak experience
  - individual user, niches, extreme personas, specific and eclectic ideas and inspiration, from concept to use

when to do it:
- individual choice, user experience, the long tail: many applications for smaller groups