Evaluation

Dr Charles Martin

Announcements

- assignment 2 due next Monday
- remember to use your tutorial time for meeting research clusters and collecting data.
- remember to follow the step-by-step guide.
- go fork the repo!

Markdown Formatting Check: There is a CI/CD job that checks your markdown formatting using the `markdownlint-cli` tool. Syntax rules are listed here in our script, rules `MD013` and `MD041` are disabled. All other rules are active.

Plan for the class

- research questions
- about evaluation
- types of evaluation
- planning evaluations
- evaluation by inspection

Research Questions

Research Questions

For the final project you will have to choose a research question to explore.

This is a clear question (one sentence) that guides the design of your research project.

RQs have been called survival beacons because they should guide all aspects of our research plans.

How do we choose a research question and write it clearly?

Important skill for any research activity.

Four steps to write a research question

This framework inspired by Lennart Nacke, everybody’s favourite HCI writing coach on LinkedIn.

1. Outline a broad area of interest
2. Identify a problem that needs solving
3. Justify solving this problem
4. Write the question

To be clear, a research question starts with a question word (what, how, why, can, do, should) and ends with a question mark. It can just be one sentence.

Seems too easy… let’s try it together.

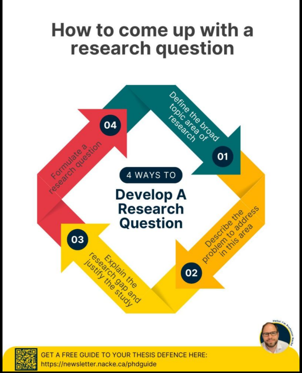

How to come up with a research question
4 WAYS TO
Develop A Research Question
01
Define the broad topic area of research
02
Describe the problem to address in this area
03
Explain the research gap and justify the study
04
Formulate a research question
Hello! I'm Professor Nacke
GET A FREE GUIDE TO YOUR THESIS DEFENCE HERE:
https://newsletter.nacke.ca/phdguide

Figure 1: Lennart Nacke’s 4-step research question framework

A worked research question example

1. haptic wearable interfaces.
2. keeping focussed in complicated meetings
3. lack of awareness in meeting can lead to poor work performance and embarrassment
4. here we go:
   What effects can a haptic wearable interface have on lack of awareness during meetings and later work performance?

Encodes the broad area, the problem, the justification, the context, etc.

Research Question Bingo

Interfaces:

1. haptic feedback gloves
2. AR/VR headset
3. e-textile clothing
4. voice assistant
5. gesture recognition
6. smart headphones
7. ambient light display
8. wearable plants
9. eye-tracking interface
10. multi-touch table

Problems to Solve:

1. family meal planning
2. language learning while commuting
3. caring for houseplants
4. focus during remote work
5. medication schedules
6. teaching kids about conservation
7. organising hobby collections
8. non-verbal communication
9. practicing music in small spaces
10. tracking community events

Activity: Write a Research Question

Let’s write a research question!

Together, let’s Spin the wheels to decide on a broad area and a problem.

Then, decide on a “justification” and write a research question.

Remember that the RQ should include the broad area, the problem, and the justification.

Use the poll everywhere link to suggest research questions and vote on the best ones.

Write for 2-3 minutes, **vote** for 1 minute, then let’s discuss.

Figure 2: PollEverywhere link: https://pollev.com/charlesmarti205

About Evaluation

About Evaluation

Figure 3: Testing things to find out if they work.

What is evaluation?

- **Evaluation:** collecting and analysing data from user experiences with an artefact.
- **Goal:** to improve the artefact's design.
- **Addresses:** functionality, usability, user experience
- Appropriate for all different kinds of artefacts and prototypes.
- Methods vary according to goals.

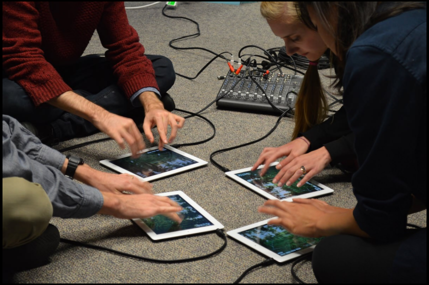


Figure 4: Evaluating iPad apps in 2013 (Martin et al., 2014)

Why is evaluation important?

- **Understanding people**
  - Users may not have the same experiences or perspectives as you do
  - Different users use software differently
- **Understanding designs**
  - Proof that ideas work
  - Understand limitations, affordances, applications
- **Business**
  - Invest in the right ideas
  - Find problems to solve (before production, before next iteration, etc.)
- **Research**
  - Evidence for new interactive systems
  - Empirical proof of hypotheses
  - New knowledge to answer research questions

What should you evaluate/measure?

Does the design do what the users need and want?

Examples:

- **Game App Developers:** Whether young adults find their game fun and engaging compared to other games
- **Government authority:** Whether their online service is accessible to users with a disability
- **Children's talking toy designers:** Whether six-year-olds enjoy the voice, feel of the soft toy, and can use safely

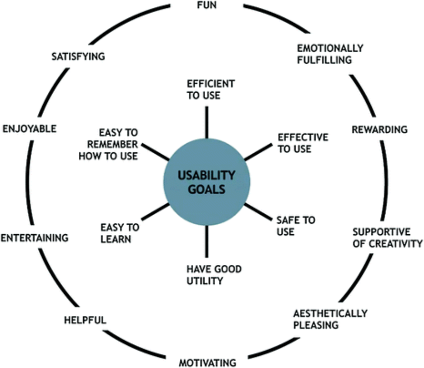


Figure 5: Preece in Raffaele et al. (2016)

Usability and Usability Goals

Six usability goals:

- Effective to use (effectiveness)
- Efficient to use (efficiency)
- Safe to use (safety)
- Having good utility (utility)
- Easy to learn (learnability)
- Easy to remember how to use (memorability)

(Rogers et al., 2023)

Figure 6: Image: dtravisphd on Unsplash

Where should you evaluate your design?

Depends on your evaluation goal!

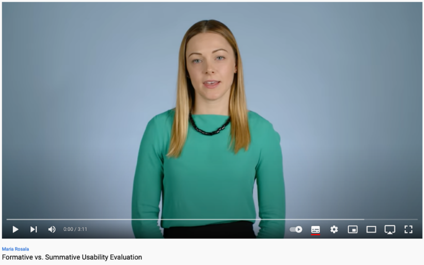
- Lab studies (controlled settings)
- In-the-wild studies (natural settings)
- Remote studies (online behaviour)

Figure 7: Image: Unsplash, UX Indonesia

Formative vs Summative Evaluation

Evaluation serves different purposes at different stages of the design process

- **Formative evaluation:**
  - Assessing whether a product continues to meet users' needs during a design process
  - Early or late stages
- **Summative evaluation:**
  - Assessing whether a finished product is successful
  - Feeds into an iterative design process



Maria Rosala

Formative vs. Summative Usability Evaluation

Figure 8: Formative vs Summative Evaluation NNGroup on YouTube

Types of Evaluation

Controlled settings (e.g., Usability testing)

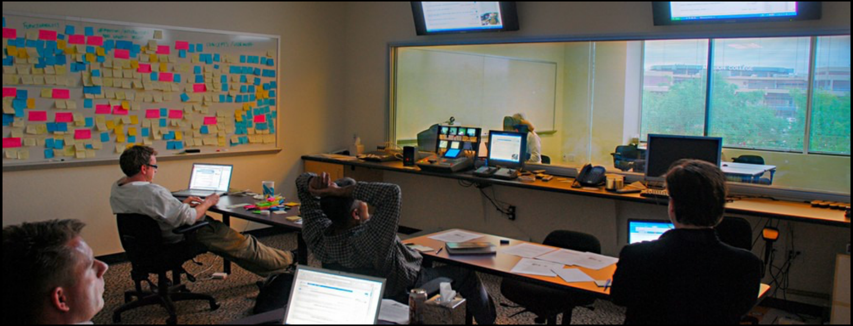


Figure 9: Image Source: Usability Testing (interactiondesign.org)

A controlled evaluation setting is not the normal place for using a technology or for the user to be.

Usability Testing

- **Measures:** numbers or time (e.g., tasks completed, errors made, time taken)
- **Methods:** mixture of methods (e.g., think aloud, observation, interviews, questionnaires, data logging and analytics)
- **Data:** variety of data depending on the methods (e.g., video, audio, facial expressions, key presses, verbal feedback)
- **Settings:** lab + observation room, mobile usability kit, university classroom
- **Number of participants:** 5-12 baseline but more is better

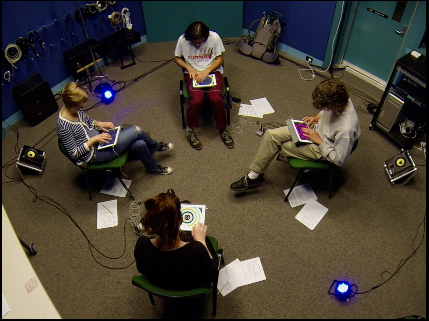
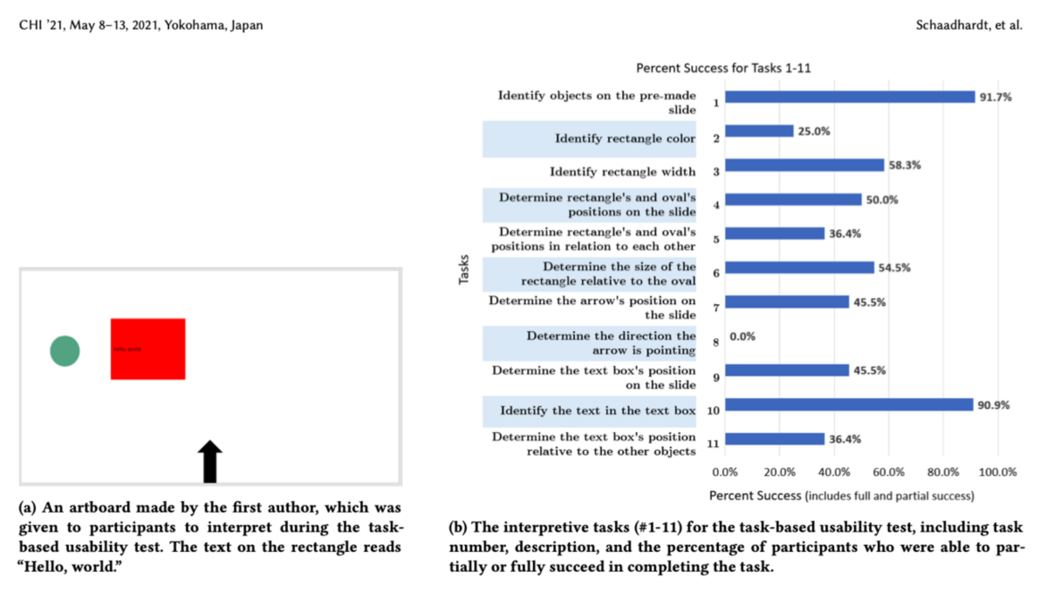


Figure 10: A controlled setting at ANU for testing music apps. Interaction data, timing, audio, video, surveys and interviews were recorded (Martin et al., 2016)

Usability Testing Example



CHI '21, May 8–13, 2021, Yokohama, Japan

Schaadhardt, et al.

Percent Success for Tasks 1-11
Tasks
Identify objects on the pre-made slide
1
91.7%
Identify rectangle color
2
25.0%
Identify rectangle width
3
58.3%
Determine rectangle's and oval's positions on the slide
4
50.0%
Determine rectangle's and oval's positions in relation to each other
5
36.4%
Determine the size of the rectangle relative to the oval
6
54.5%
Determine the arrow's position on the slide
7
45.5%
Determine the direction the arrow is pointing
8
0.0%
Determine the text box's position on the slide
9
45.5%
Identify the text in the text box
10
90.9%
Determine the text box's position relative to the other objects
11
36.4%
0.0%
20.0%
40.0%
60.0%
80.0%
100.0%
Percent Success (includes full and partial success)

(a) An artboard made by the first author, which was given to participants to interpret during the task-based usability test. The text on the rectangle reads "Hello, world."

(b) The interpretive tasks (#1-11) for the task-based usability test, including task number, description, and the percentage of participants who were able to partially or fully succeed in completing the task.

Natural settings (e.g., Field studies)

Evaluating a technology or context of use in the normal setting for the user.

Field studies can:

- Help identify opportunities for new technology
- Establish the requirements for a new design
- Facilitate the introduction of technology or inform deployment of existing technology in new contexts

Helps to establish ecological validity.

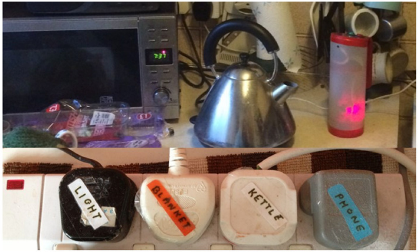


Figure 4. Top: One of the participants got a new kettle, but she still uses the Kettle Mate with the new one. Bottom: Plugs in the older person's home that she has modified with labels and arranged to support her ease of use.

Figure 12: Source: Ambe et al. (2017)

Field Studies

- Goals:
  - Understanding how people interact with technologies in “messy worlds”, how technologies will be integrated into contexts
  - Studying use of existing technologies and impacts of introducing new ones
- Methods: Emphasis on qualitative methods rather than statistical measures e.g., Observations, interviews, diaries, interaction logging
- Duration: No fixed length- can be seconds, months, years
- Paying attention to: Use situations, problems/errors, distractions, patterns of behaviours
- How does your presence and involvement shape engagement? Observation vs participant observation
- Findings: Used for creating thematic analysis, vignettes, narratives, critical incident analysis etc.

Field Studies Example

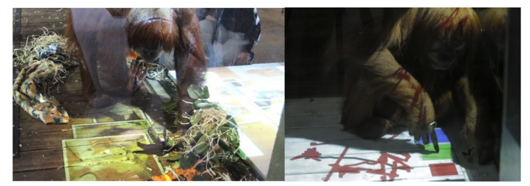


Figure 1 & 2: Orangutan Gabby interacting with two of the interactive projections developed for this project

Figure 13: Co-Designing with Orangutans: Enhancing the Design of Enrichment for Animals (Sarah Webber, Marcus Carter, Wally Smith, and Frank Vetere) Proc. DIS ’20 (Webber et al., 2020)

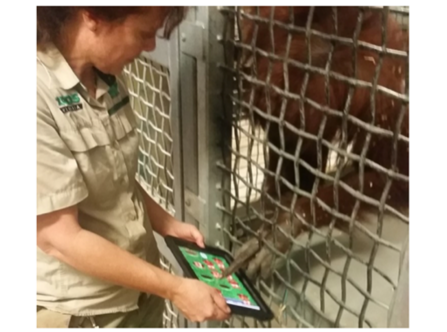


Figure 3: Orangutan Malu interacting with an iPad game.

Figure 14: Design objective 1: Develop a digital installation to provide enhanced, varied enrichment for orangutans at Melbourne Zoo

Opportunistic Evaluations

- quick feedback about a design idea in the early design process
- confirm whether it’s worth developing an idea into a prototype
- informal and doesn’t require lots of time or resources
- not a replacement for formal evaluation
- **care required** with ethics in research (Hons, Master, PhD). Asking supervisors and colleagues for advice vs collecting data to establish findings.

E.g., designers ask colleagues for design feedback: Yichen Wang’s *arMIDI* system early design process with supervisor and colleagues (Wang et al., 2025)

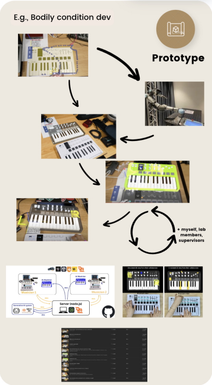


Figure 15: Development process for arMIDI

Which methods to choose from?

The evaluation setting guides certain dimensions of developed artefacts.

- **Combinations of methods** are used for a richer understanding. E.g., usability testing is combined with observations to identify usability problems and how users use the system.

Pros and Cons

- **Controlled settings** allow hypotheses testing on specific features for generalised results.
- **Uncontrolled settings** offer unexpected insights into perception and experience of new technologies in daily life and work

Figure 16: Lab research on AR co-creative system (Wang et al., 2025)

Activity: Evaluating an interactive toy

You're all HCI researchers and we need to evaluate this interactive toy.

We need to choose:

- how we will evaluate the toy?
- in what environment?
- what information do we need and why?
- what *research questions* are being asked?

Talk for 2-3 minutes and then we will hear some answers 🗣️🎤⭐

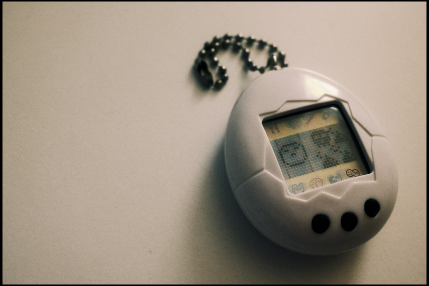


Figure 17: Where and why will we evaluate this toy? (Photo by COSMOH LOVE on Unsplash)

Planning Evaluations

Planning Evaluations

What do we need to keep in mind to plan evaluations?

Figure 18: Planning a study in 2015.

Design and Conduct Issues

- **Reliability:** how well it produces the same results on separate occasions under the same circumstances
- **Validity:** whether the evaluation method measures what it intended to measure
- **Ecological validity:** how the environment in which an evaluation is conducted influences or distorts results
- **Bias:** occurs when the results are distorted
- **Scope:** how much of the findings can be generalised

Ethical Issues

- **Risks:** what are the risks to participants? (e.g., physical harm, reputational risk, distressing conversations, being identified etc)
  - …and how are risks mitigated…
- **Benefits:** what are the benefits to participants? (e.g., none, helping research, fun experience, getting paid, course credit, etc)
- **Consent:** how is informed consent established? (e.g., a participant information sheet and a written form)
- **Data:** how is data stored and who has access to it? what will happen to it over time?

Universities have processes to *approve* the ethical aspects of research that collects data from humans following established rules (National Health and Medical Research Council (NHMRC) et al., 2025).

We don’t go deeply into research ethics in this course but the four issues above

Developing an evaluation plan

- Evaluation Goal/Aims
- Participants
- Setting
- Data to collect
- Methods
- Ethical Considerations and Consent
- Data capture, recording, storage
- Analysis method
- Output(s) of evaluation process

Figure 19: How to evaluate this app?

Labs and Equipment

- tables, chairs
- places for participants and researchers
- Instructions to participants
- Details, equipment for completing tasks
- Data collection equipment: video, audio recording
- In-person / Remote
- Zoom (e.g., COVID), online studies

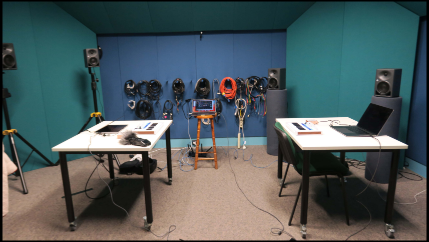


Figure 20: Yichen Wang's human-AI musical collaboration research study setup at School of Music.

Experimental Variables

- **Independent variable:** the condition the researcher controls.
- **Dependent variable:** the outcome we are measuring.
- **Independent vars in HCI:** different interfaces, input devices, software, colours, computer type
- **Dependent vars in HCI:** efficiency, accuracy, subjective satisfaction, ease of learning, physical/cognitive demands

Variables shape your study

- **Tasks:** completing specific tasks, or freely using a technology?
- **Interfaces:** just using one interface, or

Figure 21: Feels like we doing chemistry or biology… (Photo by CHUTTERSNAP on Unsplash)

Hypothesis Testing

E.g.:

A blue backround in the user interface leads to faster task completion.

- Examine the relationship between variables (independent vs. dependent)
- Null and alternative hypotheses guide testing
- Careful experimental design is essential

Hypotheses must be falsifiable and can only be dismissed! (A bit different from the more general “research questions”).

To dismiss or support a hypothesis we generally need significance testing and quantitative methods.

Experiment Design

Which participants test which conditions?

1. Different-participant design: each participant sees one condition.
2. Same-participant design: everybody sees each condition.
3. Matched-participant design: matched groups of participants with a shared trait put into each group

- **Balanced Ordering** is important to counter learning effects.
- **Design choices** affects validity and reliability
- **Data collection:** think back to week 4 lecture, but often includes task

Figure 22: Focus is on the experiment not the design!! (Photo by Girl with red hat on Unsplash)

Table of Experimental Designs

Design	Advantages	Disadvantages
Different participants (between-participants design)	- No order effects	- Requires many participants- Individual differences can affect results- Random assignment helps minimize differences
Same participants (within-participants design)	- Eliminates individual differences between conditions	- Requires counterbalancing- Risk of order effects (e.g., learning or fatigue)
Matched participants (pair-wise design)	- No order effects- Reduces impact of individual differences	- Time-consuming to find matched pairs- May miss other influential

In-the-Wild Studies

- Natural setting, minimal control over participants
- reflecting real-world use unpredictable and complex
- Ethical and practical challenges are greater, e.g., participant consent, privacy, equipment issues, and environmental factors.

Reveal insights about actual use and long-term integration that lab studies often miss.

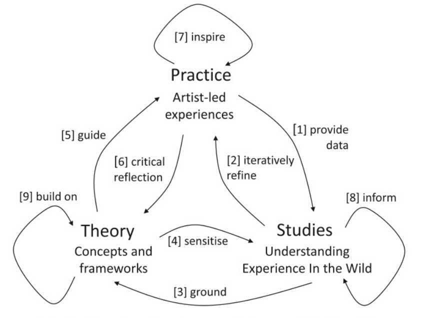


Fig. 1. Overview of performance-led research in the wild.

Figure 23: Overview of performance-led research in the wild. (Benford et al., 2013)

People are complicated

HCI is hard. To do a study, you usually need to:

- Plan the study
- Find participants
- Manage communication with them
- Figure out what to do if they don’t show up
- Managing a study requires some social skills! It’s hard work!

Is there any way to do evaluation *without* users?

Evaluation by Inspection

Evaluation by Inspection

Figure 24: Skip the "users"! Just evaluate against established principles (heuristics) and standards.

Expert Evaluation

- Conducted by designers and design “experts” rather than with end users
- Inspection methods – expert role plays user
- **Heuristic evaluation:** Researchers evaluate whether aspects design adhere to established usability principles (see over)
- **Cognitive walkthroughs:** Simulating user reasoning and problem solving at each step in an interaction sequence (evidence, availability, accessibility of correct action)
- **Analytics:** Understanding user demographics and tracing activities (e.g., number of clicks, duration of sessions etc.)
- **A/B Testing:** Large number of users assigned Design A or B and compare use to test “variable of interest” (e.g., number of clicks on advertising during test period)

Heuristic Evaluations of User Interfaces (video)

Nielsen’s 10 Usability Heuristics

1. **Visibility of system status:** keep the user informed
2. **Match between system and real world:** system uses language and communication familiar to the user, information is natural and logical
3. **User control and freedom:** users make mistakes, there should be “emergency exits” to cancel and return quickly
4. **Consistency and standards:** users should not wonder whether words, situations or actions mean the same thing, follow conventions
5. **Error prevention:** eliminate
6. **Recognition rather than recall:** make elements, actions, and options visible
7. **Flexibility and efficienty of use:** shortcuts to speed up for experts, allow tailored experiences
8. **Aesthetic and minimal design:** less is more, no unnecessary information
9. **Help users recognise, diagnose and recover from errors:** error messages need plain language, and suggest solutions
10. **Help and documentation:** best if explanation is not needed, if it is

Web Design Heuristics

Budd (2007) introduces further heuristics focussed on web, here’s some from the list:

- **Clarity:** Make the system as clear, concise and meaningful as possible for the intended audience.
- **Minimise unneccessary complexity and cognitive load:** Make the system as simple as possible for people to accomplish their tasks.
- **Provide context:** Interfaces should provide people with a sense of context in time and space
- **Promote a pleasurable and positive experience:** people should be treated with

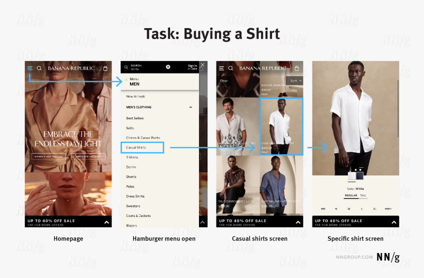
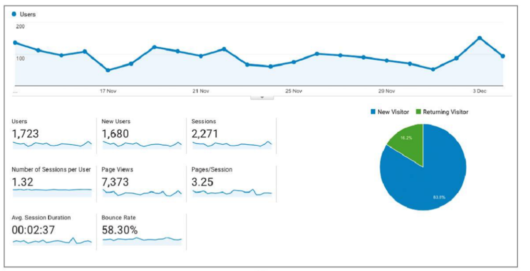


Figure 26: Evaluating a website. Image: nngroup (link)

Shneiderman’s Eight Golden Rules of Design

1. Strive for consistency
2. Seek universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Keep users in control
8. Reduce short-term memory load

Analytics: Evaluation at Scale



Users
1,723

New Users
1,680

Sessions
2,271

Number of Sessions per User
1.32

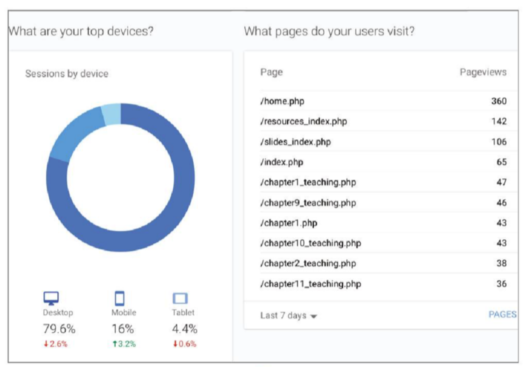
Page Views
7,373

Pages/Session
3.25

Avg. Session Duration
00:02:37

Bounce Rate
58.30%

(a)



What are your top devices?

Sessions by device

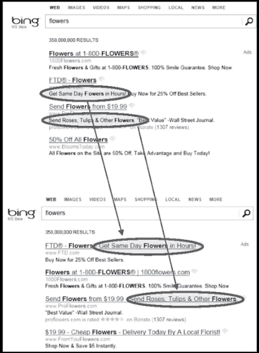
What pages do your users visit?

Page	Pageviews
/home.php	360
/resources_index.php	142
/slides_index.php	106
/index.php	65
/chapter1_teaching.php	47
/chapter9_teaching.php	46
/chapter1.php	43
/chapter10_teaching.php	43
/chapter2_teaching.php	38
/chapter11_teaching.php	36

(b)

A/B Testing

- Large-scale, online controlled experiment used to compare two designs (A = control, B = new design) by measuring user behavior (e.g., click rates), often without users knowing they are part of a study.
- Between-participants design, randomly assigning users to different versions and analyzing outcomes statistically to determine if observed differences are due to the design and not chance.
- Proper setup is critical — running an A/A test first ensures the testing infrastructure is sound, and careful design is needed to avoid misleading results, as shown in real-world examples like Microsoft Office



WEB IMAGES VIDEOS MAPS SHOPPING LOCAL NEWS MORE

bing MS Beta

flowers

358,000,000 RESULTS

Ads

Flowers at 1-800-**FLOWERS**®
1800Flowers.com
Fresh **Flowers** & Gifts at 1-800-**FLOWERS**. 100% Smile Guarantee. Shop Now

FTD® - **Flowers**
www.FTD.com
Get Same Day **Flowers** in Hours! Buy Now for 25% Off Best Sellers.

Send **Flowers** from $19.99
www.ProFlowers.com
Send Roses, Tulips & Other **Flowers**. "Best Value" -Wall Street Journal.
proflowers.com is rated ★★★★★ on Bizrate (1307 reviews)

50% Off All **Flowers**
www.BloomsToday.com
All **Flowers** on the Site are 50% Off. Take Advantage and Buy Today!

WEB IMAGES VIDEOS MAPS SHOPPING LOCAL NEWS MORE

bing MS Beta

flowers

358,000,000 RESULTS

Ads

FTD® - **Flowers** - Get Same Day **Flowers** in Hours!
www.FTD.com
Buy Now for 25% Off Best Sellers.

Flowers at 1-800-**FLOWERS**® | 1800flowers.com
1800Flowers.com
Fresh **Flowers** & Gifts at 1-800-**FLOWERS**. 100% Smile Guarantee. Shop Now

Send **Flowers** from $19.99 - Send Roses, Tulips & Other **Flowers**.
www.ProFlowers.com
"Best Value" -Wall Street Journal.
proflowers.com is rated ★★★★★ on Bizrate (1307 reviews)

$19.99 - Cheap **Flowers** - Delivery Today By A Local Florist!
www.FromYouFlowers.com
Shop Now & Save $5 Instantly.

Figure 27: Original ad title for buying flowers (top) and suggested new title design (below). Source: Kohavi et al. (2022), Cambridge University Press

Predictive Models

Estimate user performance without needing real users, using formulas to assess task efficiency — useful in early design stages or when testing with users is difficult.

Fitts’ Law (Fitts, 1954):

- predicts how long it takes to point at a target based on its size and distance
- helps designers optimize button placement, size, and spacing on screens and devices.
- applications: input methods (e.g., touch, gaze, tilt), mobile and VR, simulating interactions for users with motor impairments

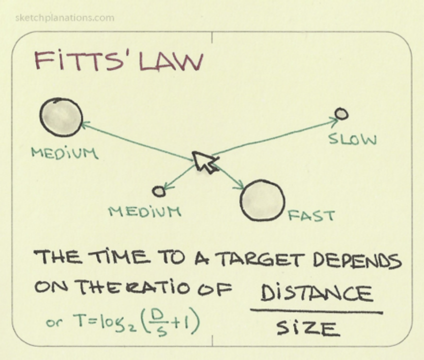

sketchplanations.com
FITTS' LAW
MEDIUM
SLOW
MEDIUM
FAST
THE TIME TO A TARGET DEPENDS ON THE RATIO OF DISTANCE / SIZE
or $T=\log_2(\frac{D}{S}+1)$

Figure 28: Fitt’s Law from Sketchplanations (CC-BY 4.0)

Evaluation after deployment: adoption, use, and non-use

- Adoption/Appropriation/Design-in-use (Ehn, 2008)
- Technology acceptance (Davis, 1989)
- Non-use (Satchell & Dourish, 2009)
- Technology habitation (Soro et al., 2016)
- Technology individuation (Ambe et al., 2017)

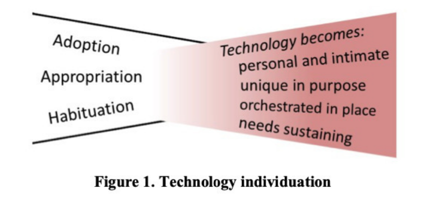


Figure 29: (Ambe et al. 2017)

Questions: Who has a question?

Questions: Who has a question?

Who has a question?

- I can take *cathchbox* question up until 2:55
- For after class questions: meet me outside the classroom at the bar (for 30 minutes)
- Feel free to ask about **any aspect of the course**
- Also feel free to ask about **any aspect of computing at ANU**! I may not be able to help, but I can listen.

Figure 30: Meet you *at the bar* for questions. 🍸🥤🫖☕ Unfortunately no drinks served! 🙃

References

References i

Ambe, A. H., Brereton, M., Soro, A., & Roe, P. (2017). Technology individuation: The foibles of augmented everyday objects. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6632–6644. https://doi.org/10.1145/3025453.3025770

Benford, S., Greenhalgh, C., Crabtree, A., Flintham, M., Walker, B., Marshall, J., Koleva, B., Rennick Egglestone, S., Giannachi, G., Adams, M., Tandavanitj, N., & Row Farr, J. (2013). Performance-led research in the wild. *ACM Trans. Comput.-Hum. Interact.*, *20*(3). https://doi.org/10.1145/2491500.2491502

Budd, A. (2007). *Heuristics for modern WEb application development*. https://andybudd.com/archives/2007/01/heuristics_for_modern_web_application_de

References ii

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Ehn, P. (2008). Participation in design things. *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*, 92–101.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381.

References iii

Martin, C., Gardner, H., & Swift, B. (2014). Exploring percussive gesture on iPads with ensemble metatone. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1025–1028. https://doi.org/10.1145/2556288.2557226

Martin, C., Gardner, H., Swift, B., & Martin, M. (2016). Intelligent agents and networked buttons improve free-improvised ensemble music-making on touch-screens. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2295–2306. https://doi.org/10.1145/2858036.2858269

References iv

National Health and Medical Research Council (NHMRC), Australian Research Council (ARC), & Universities Australia. (2025). *The national statement on ethical conduct in human research (2025)*. https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2025

Raffaele, R., Carvalho, B., Lins, A., Marques, L., & Soares, M. M. (2016). Digital game for teaching and learning: An analysis of usability and experience of educational games. In A. Marcus (Ed.), *Design, user experience, and usability: Novel user experiences* (pp. 303–310). Springer International Publishing.

Rogers, Y., Sharp, H., & Preece, J. (2023). *Interaction design: Beyond human-computer interaction, 6th edition*. John Wiley & Sons, Inc. https://quicklink.anu.edu.au/kv9b

References v

Satchell, C., & Dourish, P. (2009). Beyond the user: Use and non-use in HCI. *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, 9–16. https://doi.org/10.1145/1738826.1738829

Schaadhardt, A., Hiniker, A., & Wobbrock, J. O. (2021). Understanding blind screen-reader users’ experiences of digital artboards. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445242

References vi

Soro, A., Brereton, M., & Roe, P. (2016). Towards an analysis framework of technology habituation by older users. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 1021–1033. https://doi.org/10.1145/2901790.2901806

Wang, Y., Xi, M., Adcock, M., & Patrick Martin, C. (2025). Seeing the sound: Supporting musical collaboration with augmented reality. *Proceedings of the 2025 Conference on Creativity and Cognition*, 99–112. https://doi.org/10.1145/3698061.3726905

References vii

Webber, S., Carter, M., Smith, W., & Vetere, F. (2020). Co-designing with orangutans: Enhancing the design of enrichment for animals. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 1713–1725. https://doi.org/10.1145/3357236.3395559