Week 14

Evaluation (Part 3)

Goals

- Describe the key concepts associated with inspection methods
- Explain how to do heuristic evaluation and walkthroughs
- Explain the role of analytics in evaluation
- Describe how A/B testing is used in evaluation
- Describe how to use Fitts' Law – a predictive model

Inspections

- Several kinds
- Experts use their knowledge of users and technology to review software usability
- Expert critiques can be formal or informal
- Heuristic evaluation is a review guided by a set of heuristics
- Walkthroughs involve stepping through a pre-planned scenario noting potential problems

Heuristic evaluation

- Developed by Jacob Nielsen in the early 1990s
- Based on heuristics distilled from an empirical analysis of 249 usability problems
- These heuristics have been revised for current technology by Nielsen and others for:
  - Mobile devices
  - Wearables
  - Virtual worlds
  - Social media
  - …
- Design guidelines form a basis for developing heuristics

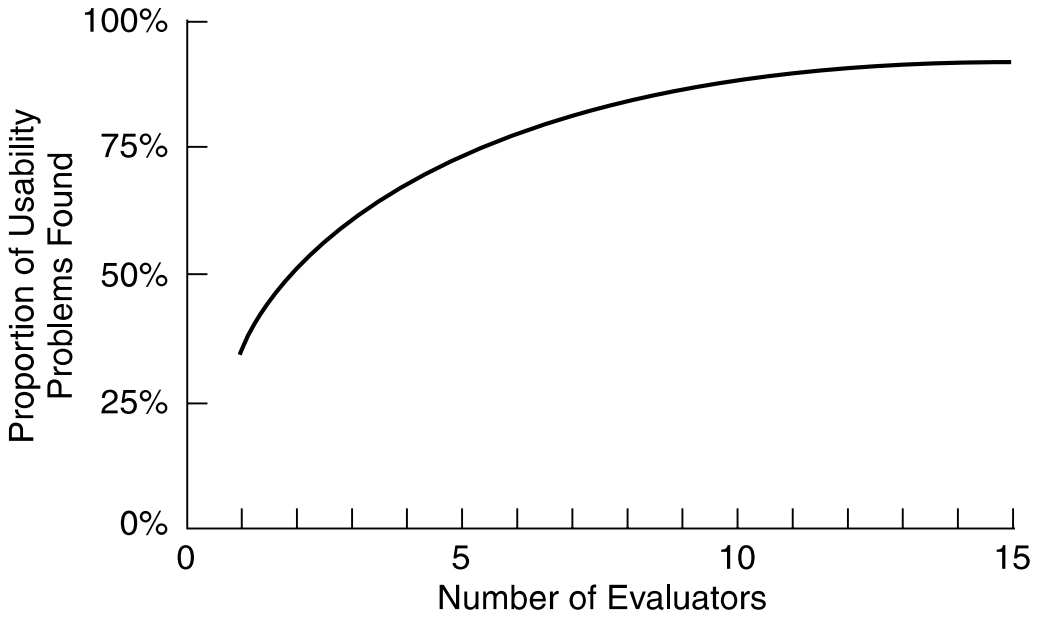
Revised version (2014) of Nielsen's original heuristics *(continued)*

- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, recover from errors
- Help and documentation

Revised version (2014) of Nielsen's original heuristics

- Visibility of system status
- Match between system and real world
- User control and freedom
- Consistency and standards
- Error prevention

Number of evaluators and problems



Curve showing the proportion of usability problems in an interface found by heuristic evaluation using different numbers of evaluators

Source: Nielsen and Mack, 1994. Courtesy of Wiley.

Number of evaluators

- Nielsen suggests that on average five evaluators identify 75-80 percent of usability problems
- Cockton and Woolrych (2001) point out that the number of users needed to find 75-80 percent of usability problems depends on the context and nature of the task problems

Heuristics for websites focus on key criteria

- Clarity
- Minimize unnecessary complexity and cognitive load
- Provide users with context
- Promote positive and pleasurable user experience

Source: Budd, 2007

Doing heuristic evaluation

- Briefing session to tell experts what to do
- Evaluation period of 1-2 hours in which:
  - Each expert works separately
  - Take one pass to get a feel for the product
  - Take a second pass to focus on specific features
- Debriefing session in which experts work together to prioritize problems

Advantages and problems

- Few ethical and practical issues to consider because users not involved
- Can be difficult and expensive to find experts
- Best experts have knowledge of application domain and users
- Biggest problems:
  - Important problems may get missed
  - Many trivial problems are often identified, such as false alarms
  - Experts have biases

Turning design guidelines and golden rules into heuristics

Ask questions like the following:

"Does the application include a visible title page, section or site? Does the user always know where they are located? Does the user always know what the system or application is doing? Are the links clearly defined? Can all actions be visualized directly (i.e., no other actions are required)?"

Granollers, 2018, p. 62

Evaluating for accessibility using Guidelines

- Web Content Accessibility Guidelines (WCAG) (see Lazar et al., 2015)
- Guidelines can be used as heuristics for evaluating websites
- Governments and large corporations have to make their websites accessible by law
- Four key concepts:
  - Perceivable
  - Operable
  - Understandable
  - Robust

Source: WCAG 2.1 at a Glance.

Cognitive walkthroughs

- Focus on ease of learning
- Designer presents an aspect of the design and usage scenarios
- Expert is told the assumptions about user population, context of use, task details
- One or more experts walk through the design prototype with the scenario
- Experts are guided by three questions

The three questions

- Will the correct action be sufficiently evident to the user?
- Will the user notice that the correct action is available?
- Will the user associate and interpret the response from the action correctly?

As the experts work through the scenario, they note problems

Pluralistic walkthrough

- Variation on the cognitive walkthrough theme
- Performed by a carefully managed team
- The panel of experts begins by working separately
- This is followed by a managed discussion that leads to agreed decisions
- The approach lends itself well to participatory design
- Also other adaptations of basic cognitive walkthroughs

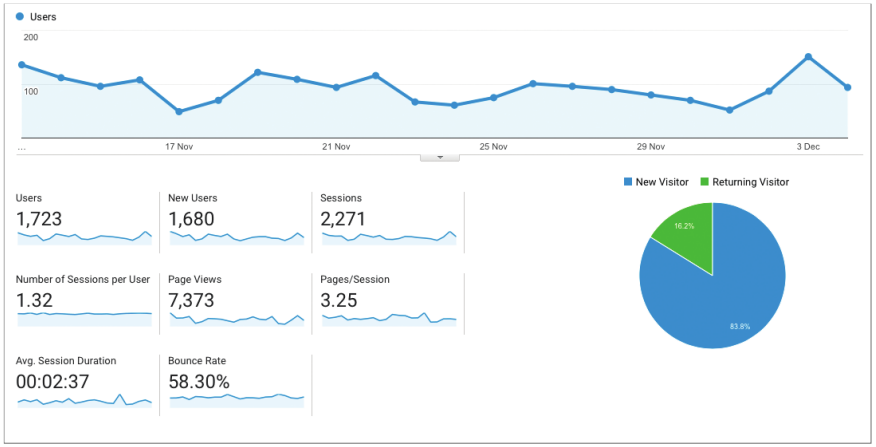
Web Analytics

- A form of interaction logging that analyzes users' activities on website
- Designers use the analysis to improve their designs
- When designs don't meet users' needs, they will not return to the site
- They become one-time users

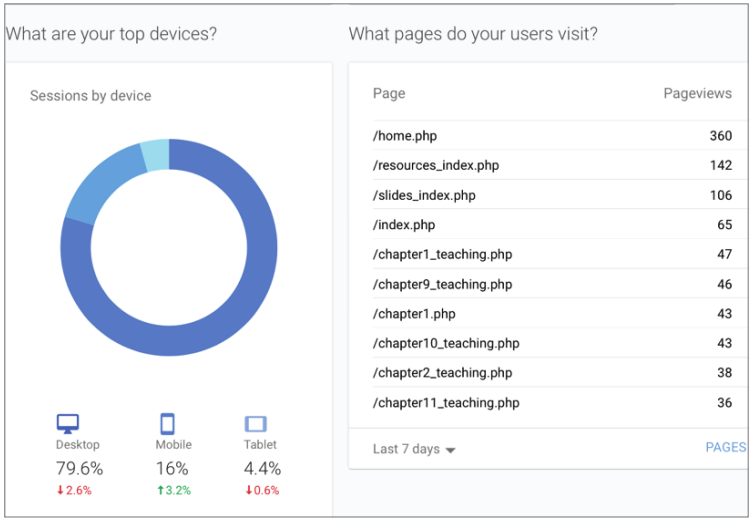
Web Analytics *(continued)*

- Web analytics enable designers to track the activities of users on their site
- They can see how many people come to the site, how long they stay, and where they go
- Web analytics offer designers the “big picture” about how their site performs based on user activity
- One of the most well-known analytics

Segment of Google Analytics for Interaction Design 5e website, December 2018



(a)

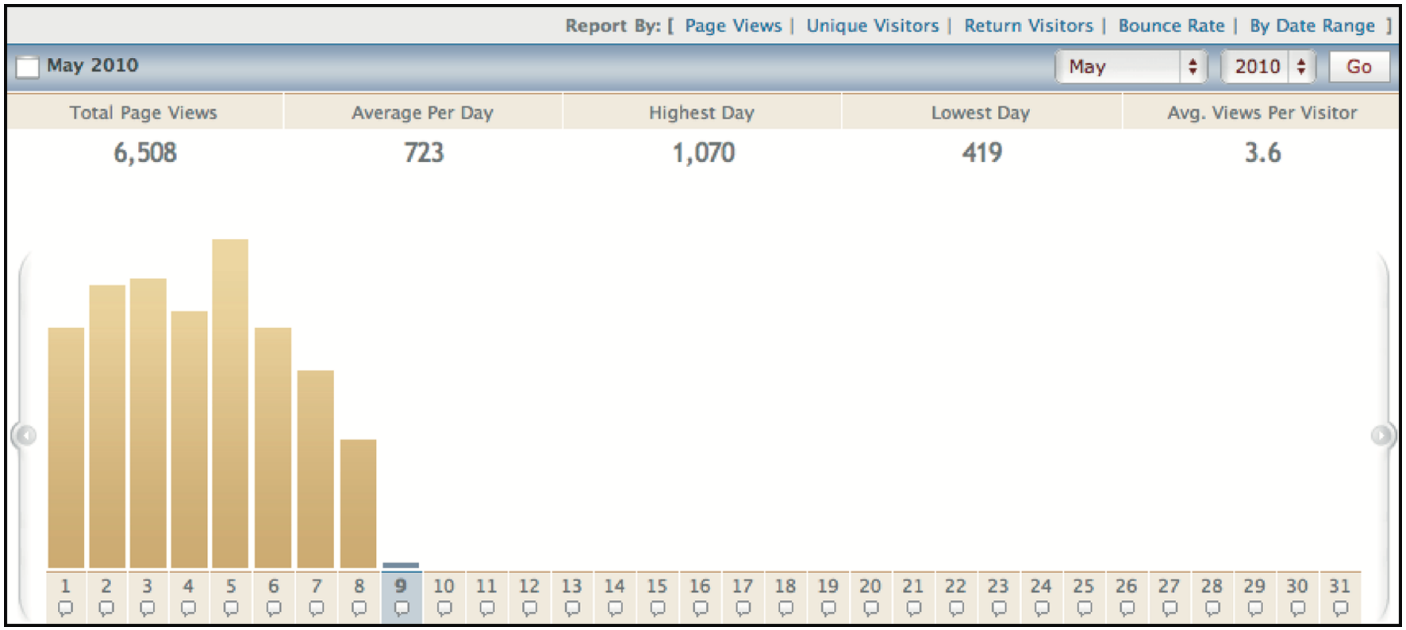


(b)

Segment of Google Analytics for Interaction Design 5e website, December 2018 *(continued)*

Language	Acquisition			Behavior		
	Users ↓	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Duration
	529 % of Total: 100.00% (529)	462 % of Total: 100.22% (461)	642 % of Total: 100.00% (642)	60.28% Avg for View: 60.28% (0.00%)	3.26 Avg for View: 3.26 (0.00%)	00:02:31 Avg for View: 00:02:31 (0.00%)
1. en-us	317 (59.81%)	279 (60.39%)	391 (60.90%)	55.50%	3.80	00:03:02
2. en-gb	44 (8.30%)	34 (7.36%)	52 (8.10%)	63.46%	2.44	00:01:21
3. zh-cn	27 (5.09%)	21 (4.55%)	35 (5.45%)	82.86%	2.40	00:01:31
4. es-es	12 (2.26%)	11 (2.38%)	13 (2.02%)	61.54%	2.08	00:00:32
5. sv-se	11 (2.08%)	9 (1.95%)	13 (2.02%)	69.23%	1.46	00:01:36
6. ko-kr	9 (1.70%)	9 (1.95%)	14 (2.18%)	35.71%	6.29	00:04:10
7. de-de	6 (1.13%)	6 (1.30%)	6 (0.93%)	66.67%	3.33	00:00:25
8. en	6 (1.13%)	6 (1.30%)	6 (0.93%)	83.33%	1.17	00:00:06
9. ar	5 (0.94%)	3 (0.65%)	6 (0.93%)	66.67%	4.17	00:01:00
10. nl-nl	5 (0.94%)	5 (1.08%)	5 (0.78%)	40.00%	2.80	00:01:02

Segment of early VisiStat Analytics from 2010



Source: VisiStat Analytics, 2010

Segment of early VisiStat Analytics from 2010 *(continued)*

	Unique Visitor	Views	Detail
1.	Los Angeles, California	6	
2.	Sharpsburg, Maryland	1	
3.	Phoenix, Arizona	3	
4.	Lemesos, Limassol	2	
5.	Targu-mures, Mures	1	

Where visitors to Mountain Wineries in California come from

Source: VisiStat Analytics, 2010

A/B Testing

- A large-scale experiment
- Offers another way to evaluate a website, application of app running on a mobile device
- Often used for evaluating changes in design on social media applications
- Compares how two groups of users perform on two versions of a design

A/B Testing *(continued)*

- Can involve thousands of users
- May create ethical dilemmas if users don't know they are part of the test
- Care is needed to ensure that other issues are not affecting users' behavior

Predictive models

- Provide a way of evaluating products or designs without directly involving users
- Less expensive than user testing
- Usefulness limited to systems with predictable tasks, for example, voicemail systems, smartphones, and dedicated mobile devices
- Based on expert error-free behavior

Fitts' Law (1954)

- Fitts' Law predicts that the time to point at an object using a device is a function of the distance from the target object and the object's size
- The further away and the smaller the object, the longer the time to locate it and point to it
- It is particularly useful for determining where on a screen to position an object
- Fitts' Law is useful for evaluating systems for which the time to locate an object is important, for example, smartphones, handhelds, and mobile devices

Summary

- Inspections can be used to evaluate requirements, mockups, functional prototypes, or systems
- User testing and heuristic evaluation may reveal different usability problems
- Design guidelines can be used to develop heuristics
- Walkthroughs are a fine-grained focused method for evaluating small parts of a product

Summary *(continued)*

- Analytics involves collecting data about users activity on a website or product to see which parts are used
- A/B testing is a form of large-scale experiment
- Fitts' Law can be used to predict expert, error-free performance for clearly defined tasks with limited key presses, for example, to evaluate keypress sequences for handheld devices and the position of objects on a screen