

# Evaluation and Inspection Methods

---

USER INTERFACE

COURSE 2020/2021

# Usability Testing

---



# Usability Testing

---

- Involves recording performance of **typical users** doing typical tasks.
- Controlled settings.
- Users are observed and timed.
- Data is recorded on video & key presses are logged.
- The data is used to calculate performance times, and to identify & explain errors.
- User satisfaction is evaluated using questionnaires & interviews.
- Field observations may be used to provide contextual understanding.

# Experiments & Usability Testing

---

- Experiments test hypotheses to discover new knowledge by investigating the relationship between two or more variables.
- Usability testing is **applied experimentation**.
- Developers check that the system is usable by the intended user population for their tasks.
- Experiments may also be done in usability testing.

# Usability Testing

---

- Goals and questions focus on how well users perform tasks with the product.
- Comparison of products or prototypes is common.
- Focus is on time to complete task, number and type of errors
- Data collected by video and interaction logging.
- Testing is central.
- User satisfaction questionnaires and interviews provide data about users' opinions.

# Testing conditions

---

- Usability lab or other controlled space.
- Emphasis on:
  - selecting representative users;
  - developing representative tasks.
- 5-10 users typically selected.
- Tasks usually around 30 minutes.
- Test conditions are the same for every participant.
- Informed consent form explains procedures and deals with ethical issues.

# Types of data

---

- Time to complete a task.
- Time to complete a task after a specified time away from the product.
- Number and type of errors per task.
- Number of errors per unit of time.
- Number of times online help and manuals accessed.
- Number of users making an error.
- Number of users successfully completing a task.

# How many participants is enough for user testing?

---

The number is a practical issue.

Depends on:

- schedule for testing.
- availability of participants.
- cost of running tests.

Typically 5-10 participants.

Some experts argue that testing should continue until no new insights are gained.



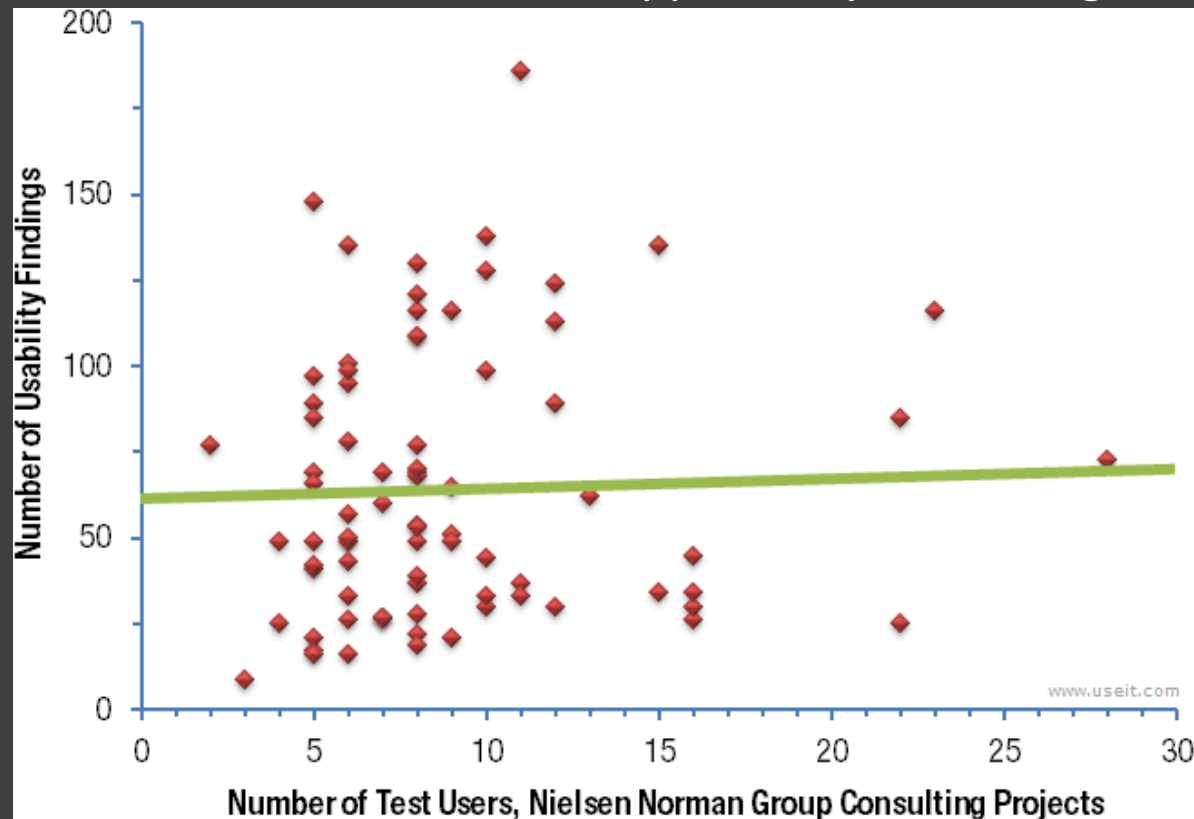
# Number of evaluators

---

- Nielsen suggests that **on average 5 evaluators** identify 75-80% of usability problems.
- Cockton and Woolrych (2001) point out that the number of users needed to find 75-80% of usability problems **depends on the context and nature of the problems**.

# How many participants is enough for user testing?

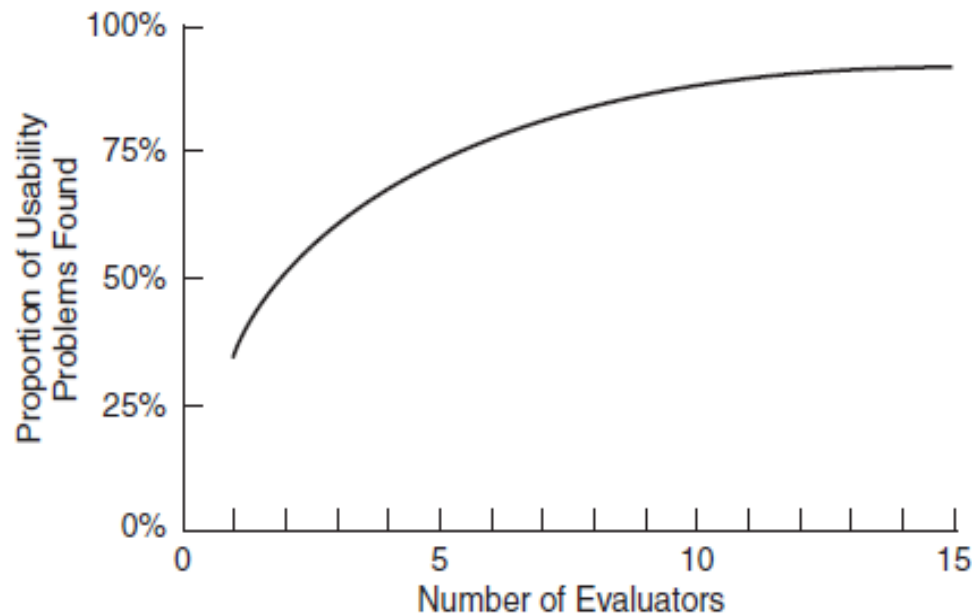
Testing more users doesn't result in appreciably more insights.



www.useit.com

<https://www.nngroup.com/articles/how-many-test-users/?lm=how-to-conduct-usability-studies&pt=report>

# How many participants is enough for user testing?



**Figure 15.1** Curve showing the proportion of usability problems in an interface found by heuristic evaluation using various numbers of evaluators. The curve represents the average of six case studies of heuristic evaluation

*Source:* Usability Inspection Methods, J. Nielsen & R.L. Mack ©1994. Reproduced with permission of John Wiley & Sons Inc.

# UQ Usability Laboratory

---



<http://www.uqul.uq.edu.au/>

# Portable Lab Equipment

---



<https://www.nngroup.com/articles/traveling-usability-lab/>

# Portable equipment for use in the field

---



**Figure 14.4** The Tracksys system being used with a mobile device camera that attaches to a flexible arm, which mounts on a mobile device, and is tethered to the lab

*Source:* Courtesy of Harry Brignull.



# Mobile head-mounted eye tracker

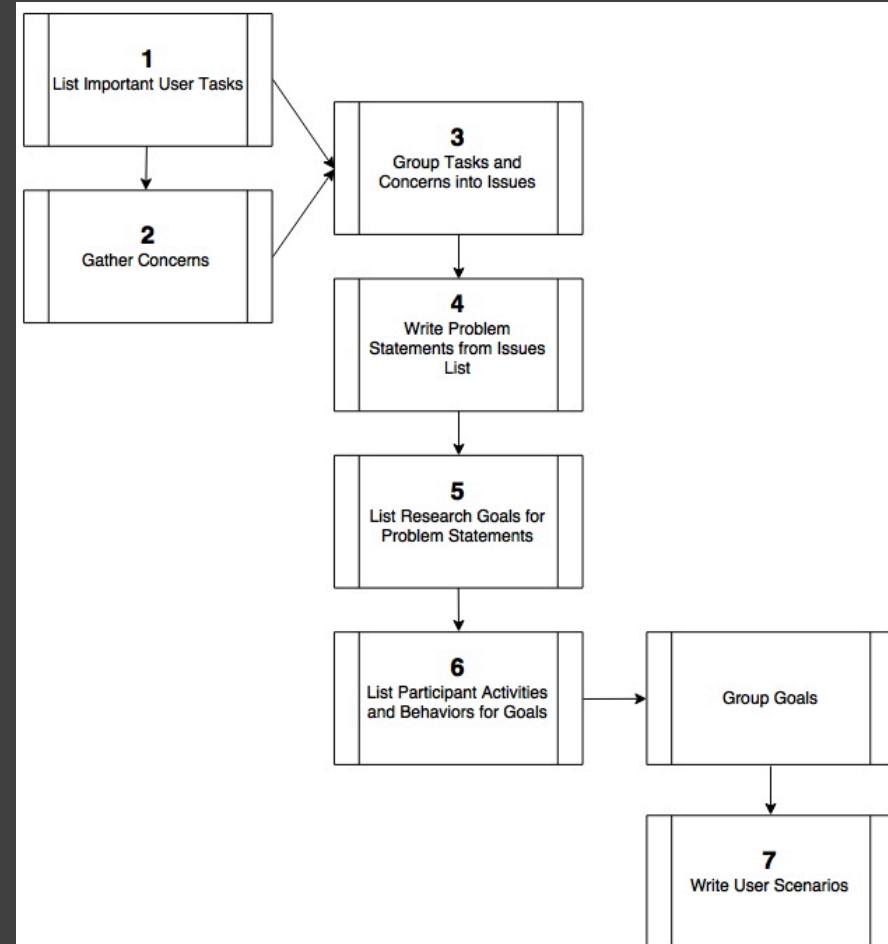


**Figure 14.5** The mobile head-mounted eye-tracker

*Source:* Picture courtesy of SensoMotoric Instruments (SMI), copyright 2010.

# A 7-Step Method

1. Determine the most important user tasks.
2. Discover which system aspects are of most concern.
3. Group items from 1 & 2, then sort issues by importance to users and organization.
4. For each issue, formulate a problem statement.
5. For each problem statement, list research goals.
6. For each research goal, list participant activities and behaviors.
7. For each group of goals, write user scenarios.



<https://www.nngroup.com/articles/ux-research-goals-to-scenarios/?lm=how-to-conduct-usability-studies&pt=report>



# Usability testing the iPad

---

- 7 participants with 3+ months experience with iPhones
- Signed an informed consent form explaining:
  - what the participant would be asked to do;
  - the length of time needed for the study;
  - the compensation that would be offered for participating;
  - participants' right to withdraw from the study at any time;
  - a promise that the person's identity would not be disclosed; and
  - an agreement that the data collected would be confidential and would be available to only the evaluators
- Then they were asked to explore the iPad
- Next they were asked to perform randomly assigned specified tasks

# Examples of the tasks

---

App or website	Task
iBook	Download a free copy of <i>Alice's Adventures in Wonderland</i> and read through the first few pages.
Craigslist	Find some free mulch for your garden.
eBay	You want to buy a new iPad on eBay. Find one that you could buy from a reputable seller.
<i>Time</i> Magazine	Browse through the magazine and find the best pictures of the week.
Epicurious	You want to make an apple pie for tonight. Find a recipe and see what you need to buy in order to prepare it.
Kayak	You are planning a trip to Death Valley in May this year. Find a hotel located in the park or close to the park.

**Table 14.1** Examples of some of the tests used in the iPad evaluation (adapted from Budiu and Nielsen, 2010).

Source: Copyright Nielsen Norman Group, from report available at <http://www.nngroup.com/reports/>.

# Example of the equipment

---



**Figure 14.6** The setup used in the Chicago usability testing sessions

Source: Copyright Nielsen Norman Group, from report available at <http://www.nngroup.com/reports/>.

# Problems and actions

---

- Problems detected:
  - Accessing the Web was difficult.
  - Lack of affordance and feedback.
  - Getting lost.
  - Knowing where to tap.
- Actions by evaluators:
  - Reported to developers.
  - Made available to public on [nngroup.com](http://nngroup.com).
- Accessibility for all users important

# Experimental Design

---

- Test hypothesis.
- Predict the relationship between two or more variables.
- Independent variable is manipulated by the researcher.
  - about users (gender, age, ...), kind of information to show to the users, devices to use, ...
- Dependent variable influenced by the independent variable.
  - % of people ending the tasks, time performance, # of clicks over an element, ...
- Typical experimental designs have one or two independent variables.
- Validated statistically & replicable.

# Experimental Design

---

- **Different participants** (Between-subjects design) - single group of participants is allocated randomly to the experimental conditions.
- **Same participants** (Within-subjects design) - all participants appear in all conditions.
- **Matched participants** - participants are matched in pairs, e.g., based on expertise, gender, etc.

# Experimental Design

---

Design	Advantages	Disadvantages
Different	No order effects	Many subjects & individual differences a problem
Same	Few individuals, no individual differences	Counter-balancing needed because of ordering effects
Matched	Same as different participants but individual differences reduced	Cannot be sure of perfect matching on all differences

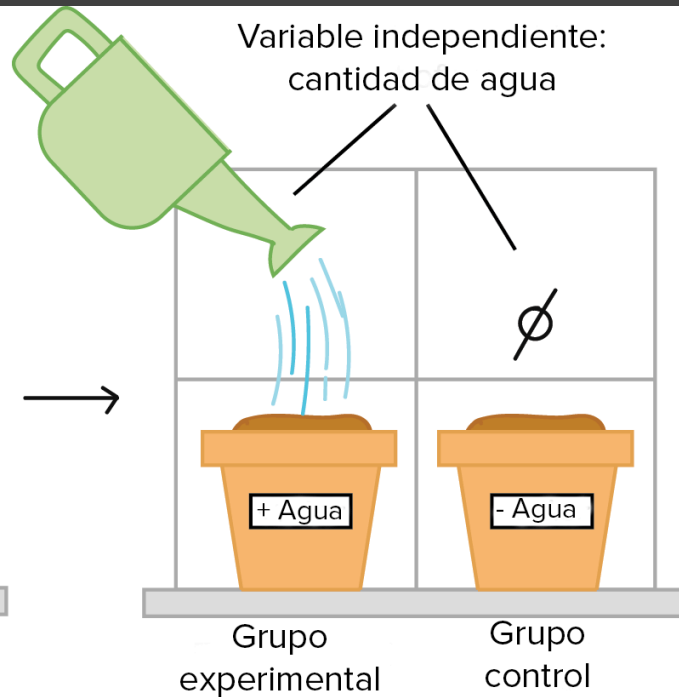
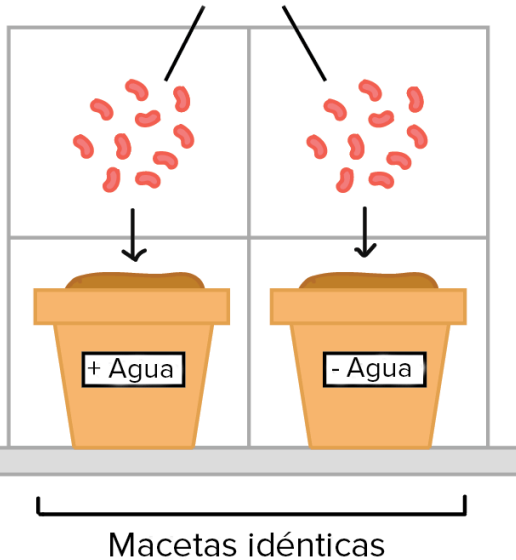
# Experimental Design

10 seeds in each pot

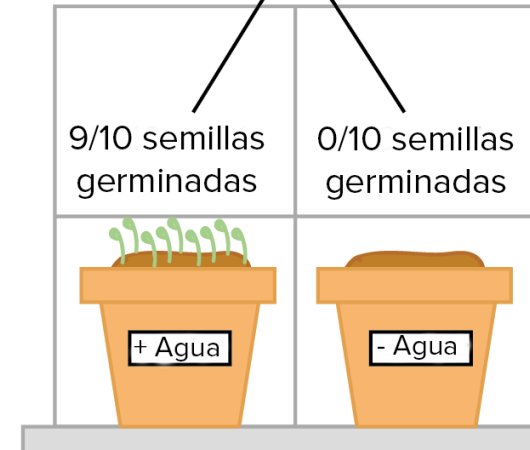
Independent variable:  
amount of water

Dependent variable:  
Proportion of germinated seeds

10 semillas de chícharo  
en cada maceta



Variable dependiente:  
fracción de semillas germinadas



Identical pots



# Field studies

---

- Field studies are done in **natural settings**.
- **“In the wild”** is a term for prototypes being used freely in natural settings.
- Aim to understand **what users do naturally** and how technology impacts them.
- Field studies are used in product design to:
  - identify opportunities for new technology;
  - determine design requirements;
  - decide how to best introduce new technology;
  - evaluate technology in use.

# Technology for context-aware field data collection

---



**Figure 14.7** An example of a context-aware experience sampling tool running on a mobile device

Source: From Cogdill, K. (1999) "MedlinePlus Interface Evaluation: Final Report". Reproduced by permission of Prof. Keith Cogdill.

# An in the wild study: UbiFit Garden



**Figure 14.8** UbiFit Garden's glanceable display: (a) at the beginning of the week (small butterflies indicate recent goal attainments; the absence of flowers means no activity this week); (b) a garden with workout variety; (c) the display on a mobile phone (the large butterfly indicates this week's goal was met)

Source: From Consolvo, S., McDonald, D.W., Toscos, T. et al (2008) "Activity sensing in the wild: a field trial of UbiFit garden". In: *Proceedings of CHI 2008*, ACM Press, New York, p. 1799.

# Data collection & analysis

---

- Observation & interviews
  - Notes, pictures, recordings
  - Video
  - Logging
- Analyses
  - Categorized
  - Categories can be provided by theory
    - Grounded theory
    - Activity theory

# Data presentation

---

- The aim is to show how the products are being appropriated and integrated into their surroundings.
- Typical presentation forms include:
  - Vignettes,
  - Excerpts,
  - Critical incidents,
  - Patterns, and narratives.

# Inspections

---

- Several kinds.
- **Experts** use their knowledge of users & 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, etc.
- Design guidelines form a basis for developing heuristics.

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



# Heuristic evaluation

---

- Heuristics for websites focus on key criteria (Budd, 2007)
  - Clarity
  - Minimize unnecessary complexity & cognitive load
  - Provide users with context
  - Promote positive & pleasurable user experience

# Heuristic evaluation

---

1. Establish an appropriate list of heuristics.
2. Select your evaluators (different from the end users, typically usability experts and preferably with domain expertise).
3. Brief your evaluators about what to do and cover.
4. First evaluation phase (around 2 hours to use the product freely).
5. Second evaluation phase (applying the heuristics individually).
6. Record problems (you or the evaluators).
7. Debriefing session among the evaluators to list problems and suggest potential solutions.

[https://www.interaction-design.org/literature/article/heuristic-evaluation-how-to-conduct-a-heuristic-evaluation?utm\\_source=facebook&utm\\_medium=sm](https://www.interaction-design.org/literature/article/heuristic-evaluation-how-to-conduct-a-heuristic-evaluation?utm_source=facebook&utm_medium=sm)

# Advantages and problems

---

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

# Cognitive walkthroughs

---

- Focus on **ease of learning**.
- Designer presents an aspect of the design & 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 4 questions.

# The 4 questions

---

- Will the user try and achieve the **right outcome**?
- Will the user notice that the **correct action** is available to them?
- Will the user **associate** the correct action with the outcome they expect to achieve?
  - Steve Krug, says in his book “Don’t Make me Think” – “It doesn’t matter how many times I have to click, as long as each click is a mindless, unambiguous choice.”
- If the correct action is performed; will the user see that **progress** is being made towards their intended outcome?

# Cognitive walkthroughs

---

- Example: the log in process on a website
- Tasks
  - Open browser
  - Navigate to site
  - Click login button
  - Enter user name in user name field
  - Enter password in password field
  - Click the login button

# Pluralistic walkthrough

---

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

# Evaluation using analytics

- A method for evaluating user traffic through a system or part of a system.
- Many examples: Google Analytics (chapter 7), Visistat (shown below), Learning Analytics.
- Times of day & visitor IP addresses

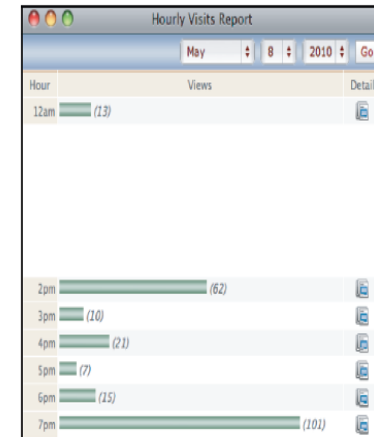


Figure 15.5 Clicking on May 8 provides an hourly report from midnight until 10.00 p.m. (only midnight and 2.00 p.m.–7.00 p.m. shown)

Source: <http://www.visistat.com/tracking/monthly-page-views.php>

Display By: Geographic Location			
	Unique Visitor	Views	Detail
1.	Los Angeles, California	6	[icon]
2.	Sharpsburg, Maryland	1	[icon]
3.	Phoenix, Arizona	3	[icon]
4.	Lemesos, Limassol	2	[icon]
5.	Targu-mures, Mures	1	[icon]

Figure 15.6 Clicking on the icon for the first hour in Figure 15.5 shows where the IP addresses of the 13 visitors to the website are located

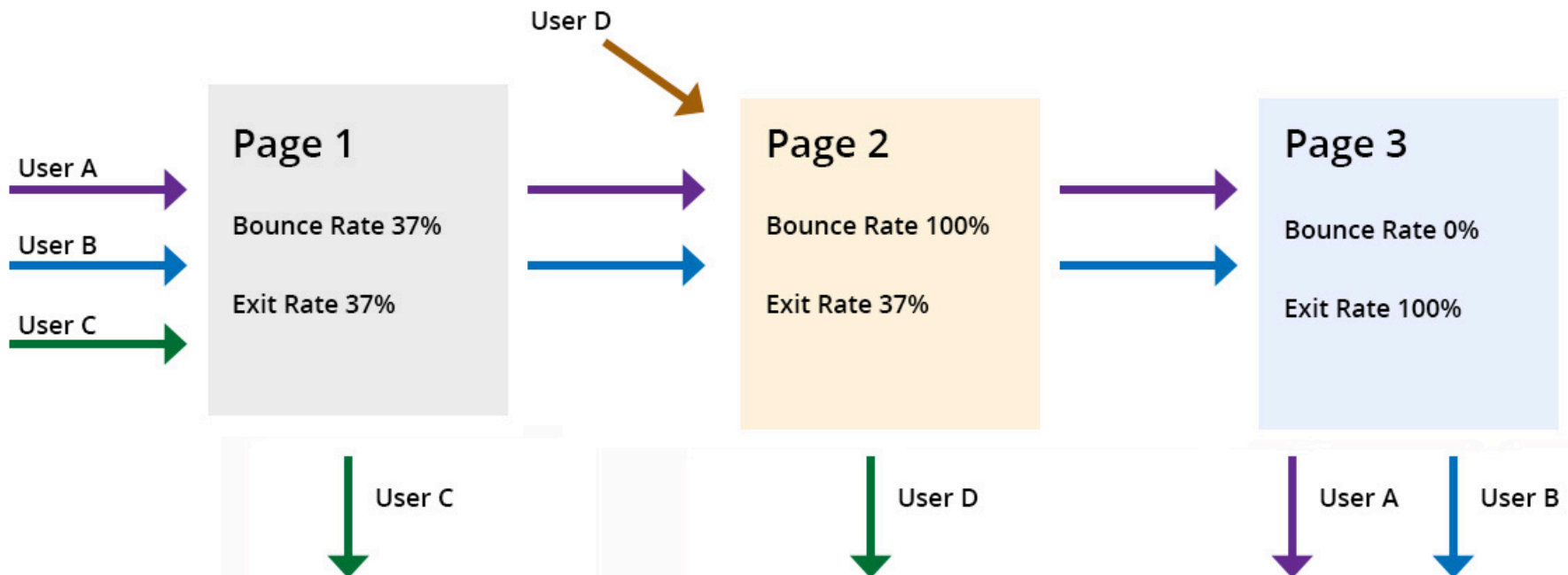
Source: <http://www.visistat.com/tracking/monthly-page-views.php>



# Evaluation using analytics

## Example: Google Analytics

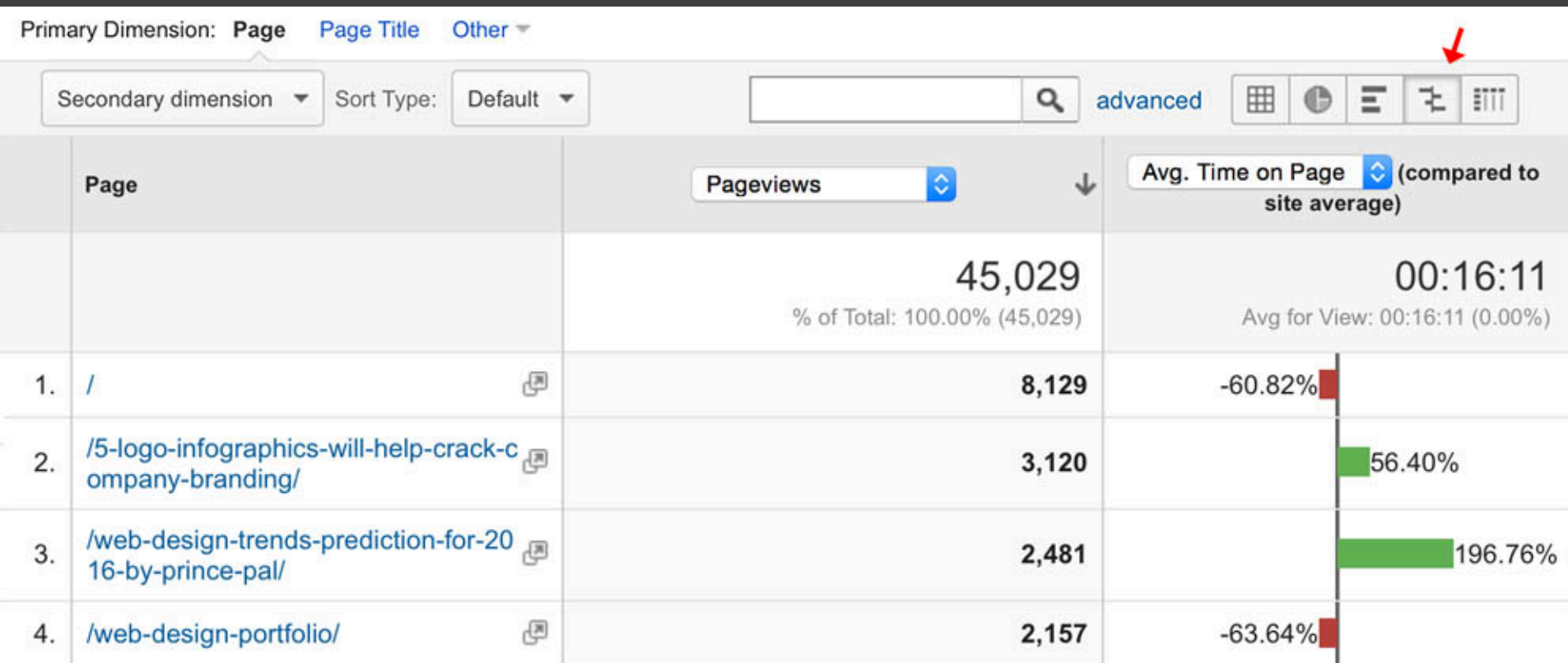
- Bounce Rate = how many users leave the website after visiting it partially (bounces / entrances)
- Exit Rate = how many users leave the website (exits / entrances)



# Evaluation using analytics

## Example: Google Analytics

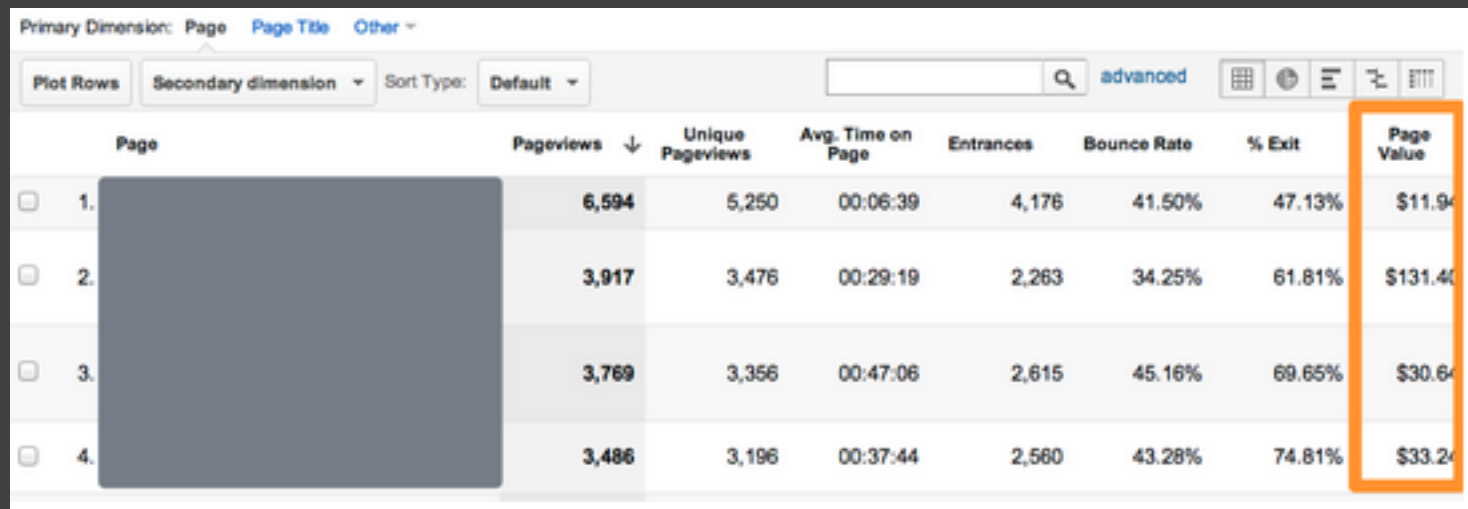
- Average Time spent by users visiting the website



# Evaluation using analytics

## Example: Google Analytics

- Page Value = a price given to a particular web page used to calculate the usability. In case of ecommerce websites, it deals with transaction returns whereas for other websites, target returns is what the primary focus is.



The screenshot shows the Google Analytics interface with a table of page metrics. The 'Page Value' column is highlighted with an orange box. The table has the following columns: Page, Pageviews, Unique Pageviews, Avg. Time on Page, Entrances, Bounce Rate, % Exit, and Page Value. The data is sorted by Page Value in descending order.

	Page	Pageviews	Unique Pageviews	Avg. Time on Page	Entrances	Bounce Rate	% Exit	Page Value
<input type="checkbox"/>	1.	6,594	5,250	00:06:39	4,176	41.50%	47.13%	\$11.94
<input type="checkbox"/>	2.	3,917	3,476	00:29:19	2,263	34.25%	61.81%	\$131.40
<input type="checkbox"/>	3.	3,769	3,356	00:47:06	2,615	45.16%	69.65%	\$30.64
<input type="checkbox"/>	4.	3,486	3,196	00:37:44	2,560	43.28%	74.81%	\$33.24

# Bibliografía

---

Interaction Design: Beyond Human - Computer Interaction. Jenny Preece, Yvonne Rogers, Helen Sharp. J. Wiley & Sons, 2002.

Material available at [www.id-book.com](http://www.id-book.com)