User Experience Research

April, 2011

Alibaba.com®
Global trade starts here.™

User Experience Research

UX research is the key to drive product strategy, optimize design, and assess product health

- Strategic Research drives product strategy by understanding user needs, evaluating new design ideas, discovering product/business opportunities, and inspiring Innovation
- Design Research optimizes design by evaluating, refining, and iterating the design from user's point of view
- Assessment Research evaluates product heath by measuring or comparing against self or competition

User Research

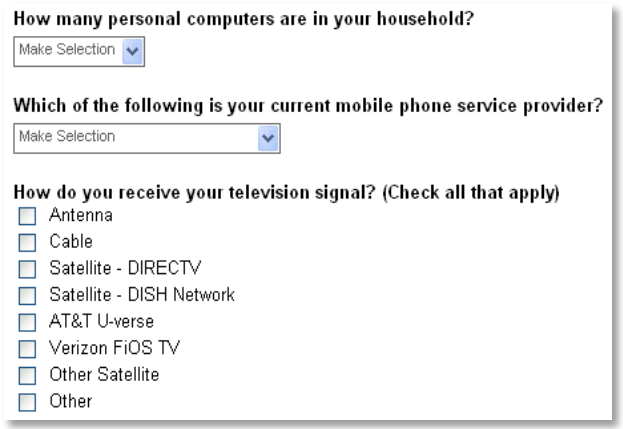
	Strategic Research (Understand & Conceive)	Design Research (Design)	Assessment Research (Develop)
Role	• New ideas and concept discovery • Product and business analysis • Innovation & strategy consulting	• Improve, refine, and iterate the design • Inform design, reduce risk	• Measure or compare against self or competition • Improve quality
Methods	• Ethnography: user needs and real usage • Competitive analysis • Feature and task analysis • Concept development and assessments	• Usability inspections • Participatory design • Rapid, iterative design & testing (RITE method, paper prototyping) • Cart sorting • Usability lab studies	• Usability benchmarking • Data mining analysis • Desirability studies • Product surveys • Customer Satisfaction • A/B test

Research Methodology

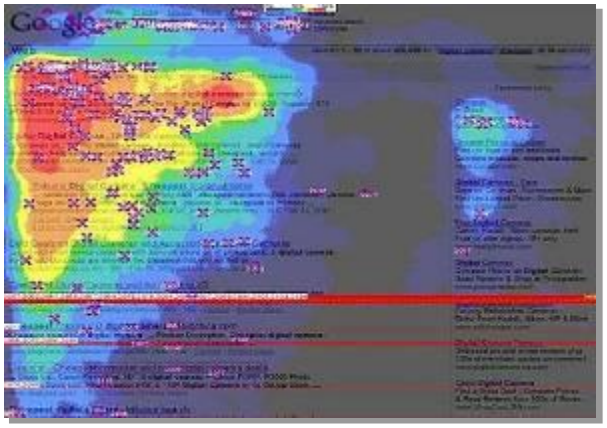
Field Study

Focus Groups

Participatory Design



Surveys



Eye Tracking

Card Sorting

Lab based usability testing

- **Enables direct observation of target users as they interact with product/web site or a design prototype**
  - Observe what users ACTUALLY do, rather than relying on what they SAY they do
  - Understand WHY people do what they do, not just what they do
- **Is a method we use to understand…**
  - Navigation and flow through the user interface
  - Comprehension of labels and text
  - Discoverability of features
- **With 6 – 8 participants we are able to get a good sense for the usability and learnability of the user interface**
- **Can be conducted on the real product/web site, prototypes and even on paper mock-ups**
- **Usability labs are available for rental in silicon valley, Shanghai, and Beijing.**

Test Planning

- **Participant recruiting**
- **Equipment arrangement**
  - Lab/camera rental
  - Software installation
- **Experimental design and test scenarios**
- **Pilot study**
- **Tasks/test sessions**
- **Data analysis**
- **Report**

Participants

- **Define participant profiles**
  - Based on personas, user segmentation, experience level
  - 4-5 participants for each segment
  - More participants are needed for complex study

- **Generate participant screen questionnaire**
  - Questionnaire should be easy to follow, so recruiter with background of the project can identify the right people
  - Work with recruiting firm to recruit and schedule participants
  - Be ready to loose the recruiting criteria when participants are hard to find

Equipment/Lab

The use of Morae enables:

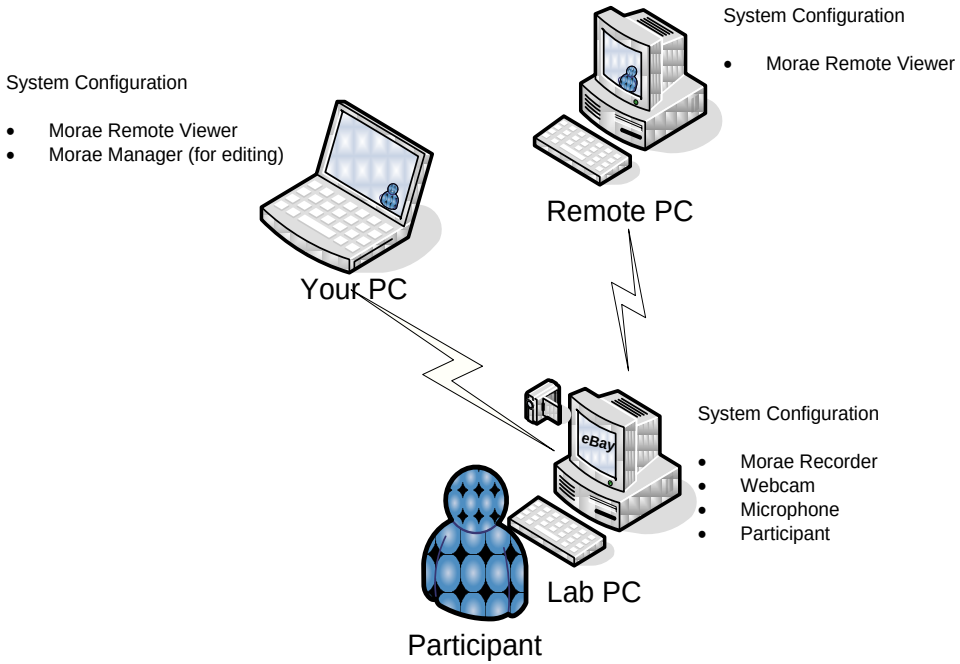
- **Creation of compelling video clips to illustrate usability issues**
- **Recording video transcript of the session for data synthesize**
- **Enables remote viewing of session in real time**

Morae Recorder

- Installed on the lab (participant) PC
- Records and streams video from the participant's PC

Morae Remote Viewer

- Installed on any remote PC
- Allows remote users to view the session
- Allows data logging of the session synced to video



Experimental Design

- **Within-subject design**
  - Each participant test all designs
  - Advantage: requires less participants
  - Disadvantage: learning effect, order confounding, participant fatigue

Design A	Design B
P1, P2, P3, P4, P5, P6	P1, P2, P3, P4 P5, P6

- **Between-subject design**
  - Different group of participants test different designs
  - Advantage: no learning effect, avoid order effect and fatigue
  - Disadvantage: require more participants

Design A	Design B
P1, P2, P3, P4, P5, P6	P7, P8, P9, P10 P11, P12

Tasks

- **Work with customer to understand the goal of the test**
  - Any specific design challenges need to be addressed
- **Generate from the product use cases**
- **Tasks should be written from user's point of view and cover the main functionalities of the design**
- **Define the goal/success but not the process**
- **Start with easy tasks to let participant learn how to participate in the study**

Pilot

- **Pilot is used to test the task design, equipment etc**
  - Prototype should be reviewed before the pilot
- **Real user or team member could be pilot participant**
  - Sometimes, the management team use pilot as the way of quality control

Test Session

- **Think aloud**
  - Ask participants to speak out their thinking, observation, and comments
  - Advantage: rich qualitative data, help understand user's mental model and decision making process
  - Disadvantage: interfere with task times and error rates – longer time and less errors

- **Team**
  - Moderator: the only person interact directly with participant
  - Data logger: record verbal feedback, time, errors. Operate test equipment
  - Observers: encourage customer and the whole design and development team to attend.
  - If time allows, have a debrief session after each test session

Test Session - Process

- **Introduction**
  - If think aloud will be used, ask participants to read aloud the instruction to help them get use to it
  - Emphasize test purpose is the software not the participant
- **Pre-test questionnaires**
- **Perform tasks**
  - Behind glass vs. seat with participants
  - Be aware of the backtalk why you flip the microphone
- **Post-test interviews**
  - Focus on probing why
- **Post-test questionnaire**
  - Such as user satisfaction questionnaire
- **Incentives and thank participant sincerely**
  - Emphasize the feedback will be used to improve products

Data Collected

- **Verbal feedback**
- **Task time**
- **Task completion rate**
  - If a user can not even finish a typical task, there must be something wrong
- **Errors**
  - Error rate
  - Type of user errors
- **Help being used**
  - Times accessing manual/help document
  - Effectiveness of the manual/help document

Data Analysis

Severity rating:

- **Critical**
  - A catastrophic usability problem that causes system failure or unrecoverable loss of data. These problems have no workaround.
  - For example a commonly used input field fails to validate and result in a data corruption.
- **High**
  - A serious condition that impairs the operation, or continued operation, of one or more product functions and cannot be easily circumvented or avoided. The software does not prevent the user from making a serious mistake. The problem is frequent, persistent, and affects many users. Standards are seriously violated.
  - For example, closing the application fails to result in a prompt for the user to save data
- **Medium**
  - A non-critical, limited problem (no data loss or system failure). It does not hinder operation and can be circumvented or avoided. The problem causes moderate confusion or irritation and will impact most users fairly regularly.
  - For example, a tree control that fails to restore to the state previously set by the user.
- **Low**
  - Non-critical problems with the product such as inconsistencies that cause user hesitation or confusion, or small aesthetic issues. Will impact only a small number of users or will only be very infrequently encountered.
  - For example, a button that brings up a dialog requiring further information is missing an ellipsis from its label.
- **Nice to have**
  - An enhancement that would improve the UI, but is outside of the scope of the currently defined release. The enhancement would improve the user experience of the product, but if it was not addressed it would not qualify as a bug.
  - For example, adding a wizard-driven approach to modifying settings that facilitates first time or infrequent use to supplement existing functionality that is already supported.

Data Analysis

- **Focus on observed data**
  - Confirmative bias
  - Favoritism bias
  - Impressing bias
- **Explore why**
- **Simple Statistical Analysis**
  - T-test
  - Chi-square test for categorical data

Report

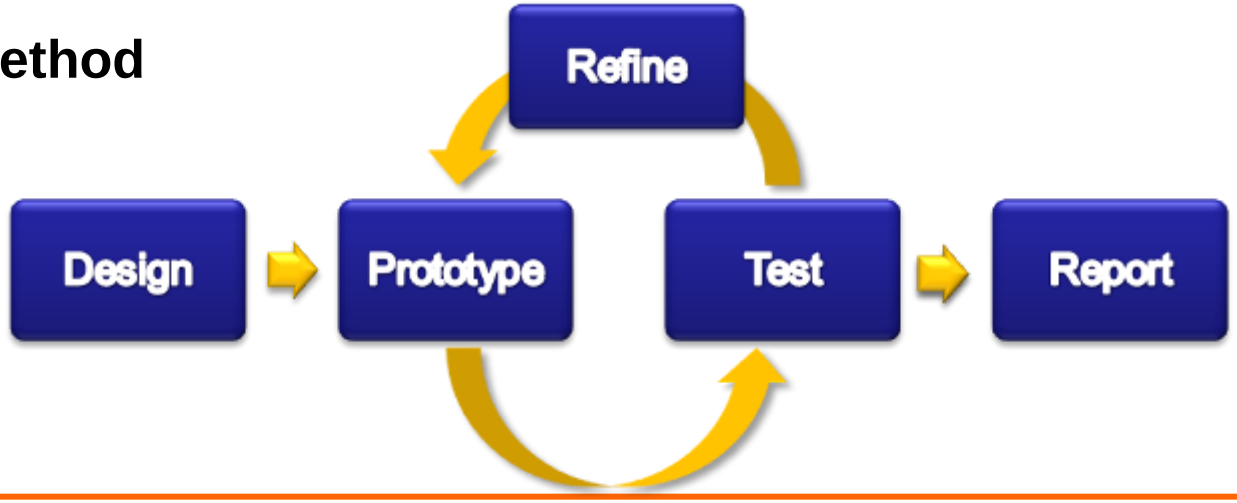
- **Executive summary**
  - Summarize the key findings concisely
  - For people don’t have much time to read
- **Detailed findings**
- **Use of screen shot and call out**
- **Pictures**
- **Video clips**
  - Powerful tool, very convincing
  - Let the team focus on issues rather than validity of small number of participants

RITE Method Usability Studies（快速迭代）

- **Rapid Iterative Test and Evaluation methodology**
  - Focus is on iterative and incremental design improvements

Traditional Waterfall Usability Testing

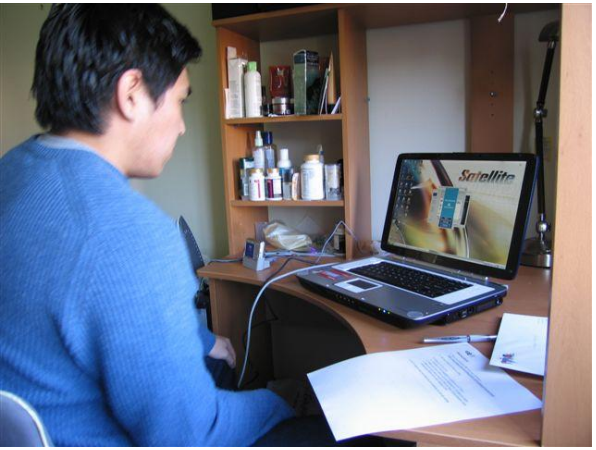
RITE Method



Heuristic Evaluation

- **Visibility of system status**
  - The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
- **Match between system and the real world**
  - The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
- **User control and freedom**
  - Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
- **Consistency and standards**
  - Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
- **Error prevention**
  - Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
- **Recognition rather than recall**
  - Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
- **Flexibility and efficiency of use**
  - Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
- **Aesthetic and minimalist design**
  - Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
- **Help users recognize, diagnose, and recover from errors**
  - Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
- **Help and documentation**
  - Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Field research: observe people in a natural context



We get to see people use product/website…

- On their own computers (PCs, laptops, old, new)
- With various connection speeds (some 56K dial-up)
- To perform their own tasks (which might be different from what you think)
- With their own cameras and workspace (home, internet café, office)
- With life's normal interruptions (telephone, talking parrot, cluttered desk, crying kid)

The focus of field research

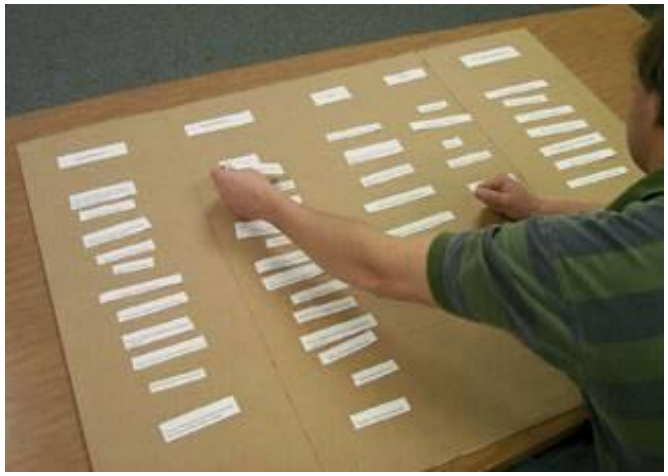
- Questions we focus on in field research are:
- “What is the larger context of use?”
- “What issues exist, and WHY?
- “What can we do to address the issues?”

Card Sorting（卡片分类）

- A great, reliable, inexpensive method for finding patterns in how users would expect to find content or functionality.

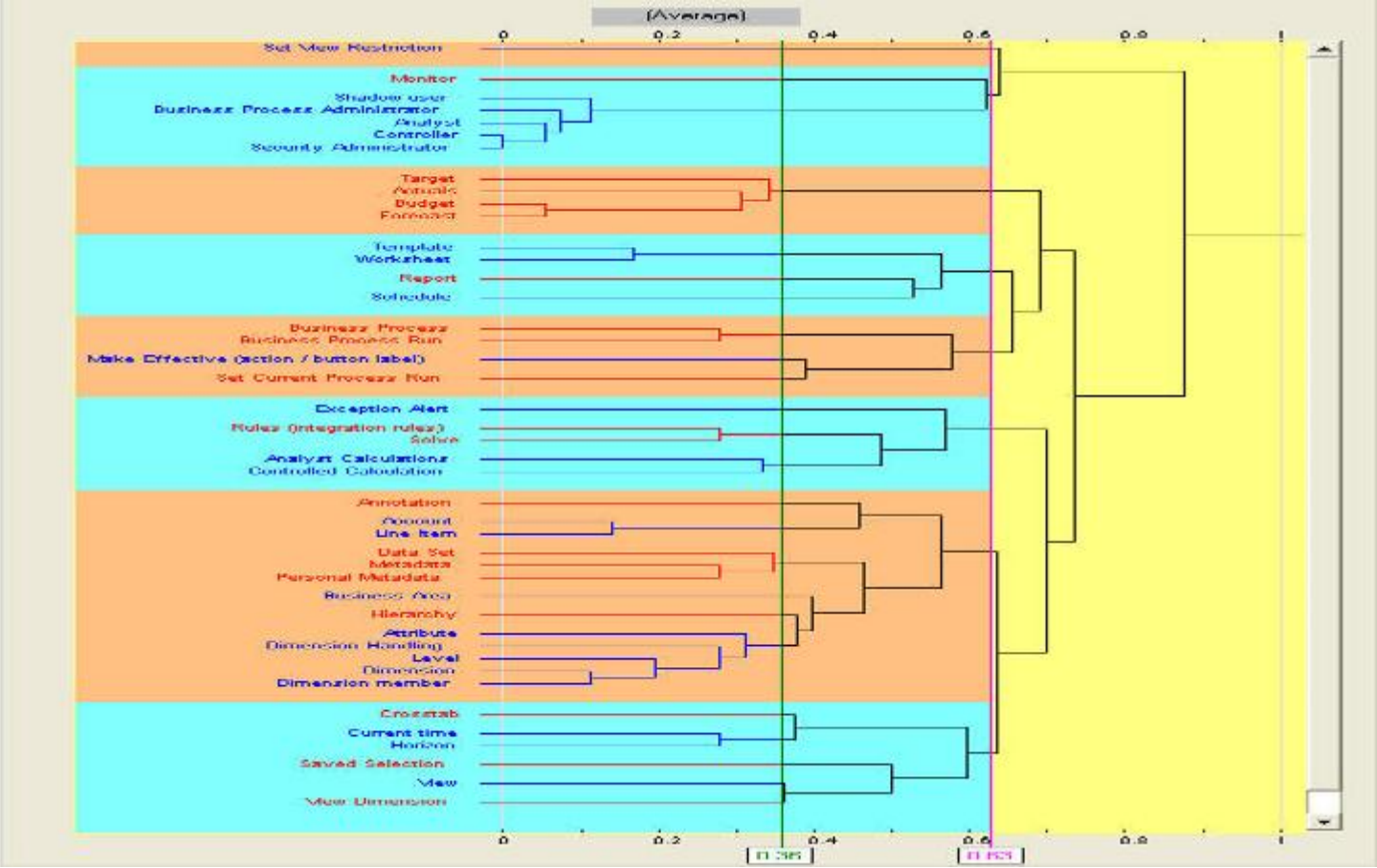
- Those patterns are often referred to as the users' mental model.

- By understanding the users' mental model, we can increase find-ability, which in turn makes the product easier to use.



Sample Card Sorting Results

- **Information is group into major clusters**
- **Helps the design of information architecture, navigation, menu, etc**



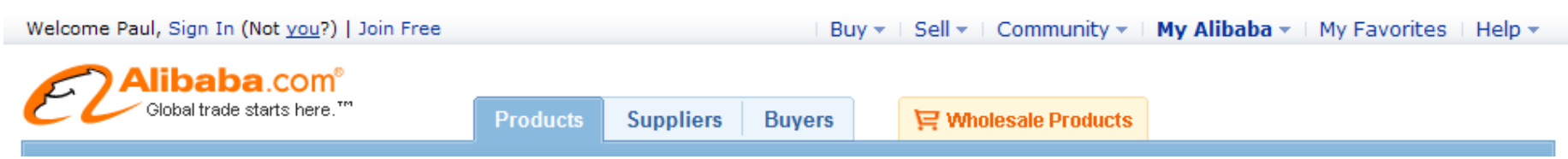
A/B Tests（A/B 测试）

- **How It Works**
  - Randomly display alternative designs for a limited time to users
  - Measure metrics for both designs to capture trend data such as:
    - Conversion rates
    - Revenues
    - Satisfaction
    - Session times

- **Commonly Used To Evaluate:**
  - New home page designs
  - Feature placement on pages
  - Algorithms for recommendation systems
  - Search relevance rankings

- **Only launch alternative design if metrics show statistically significant impact on defined success metrics**

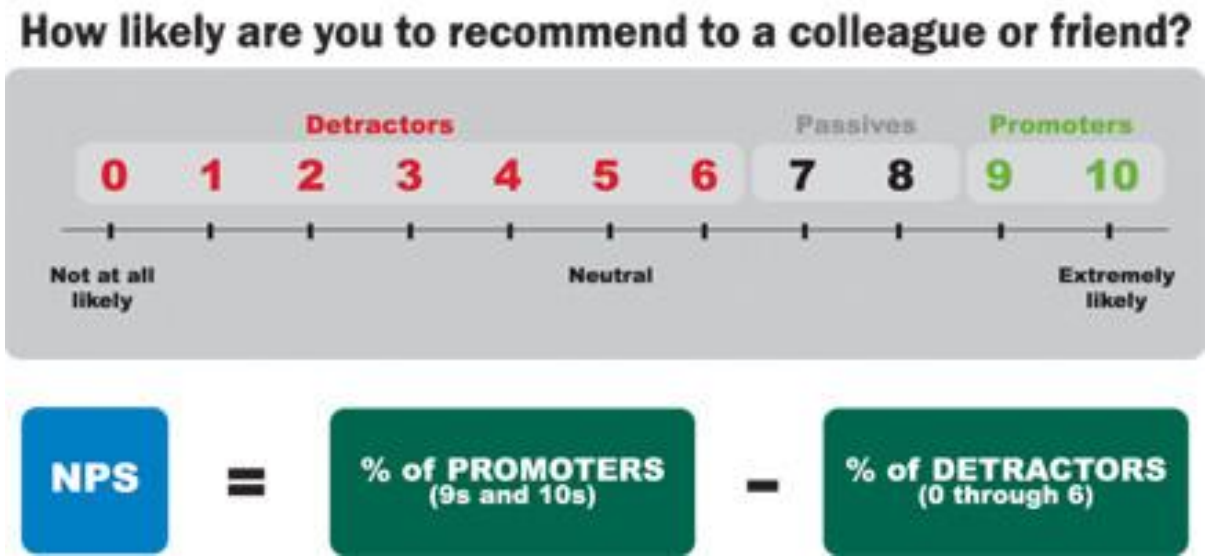
A/B Test Example



- **Alibaba Homepage Copy AB test**

YYYYMMDD	文案	SC首页到访海外session数	从SC首页到访AE首页的海外session数	到访后不再接着往下点的海外session数	到访后当日下单的海外session数
6.15-21日均	Wholesale Products	147285	2224	657	103
				29.53%	**4.62%**
			1.51%	0.45%	0.07%
6.23-29日均	Direct Buy Products	148582	1953	547	96
				28.02%	**4.94%**
			1.31%	0.37%	0.06%
7.1-7日均	Buy on AliExpress.com	150161	2358	630	130
				26.72%	**5.52%**
			1.57%	0.42%	0.09%

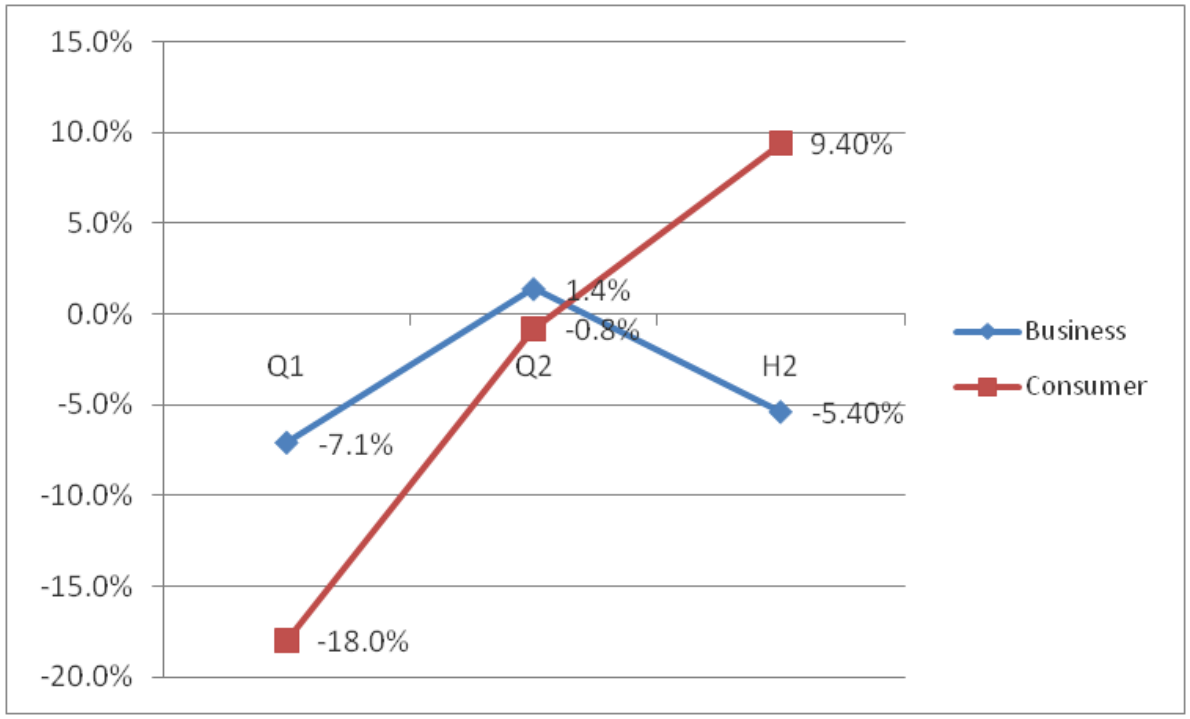
Net Promoter Score Definition（净推荐值）



- **Promoters** (score 9-10 推荐者): loyal enthusiasts who will keep buying and refer others, fueling growth.
- **Passives** (score 7-8 被动者): satisfied but unenthusiastic customers who are vulnerable to competitive offerings.
- **Detractors** (score 0-6 贬损者): unhappy customers who can damage your brand and impede growth through negative word-of-mouth.

NPS Example

- **The NPS score for AliExpress is 3.2%. However, the improvement is mainly from individual users. Business user's NPS is decreased to -5.4%**



NPS Example

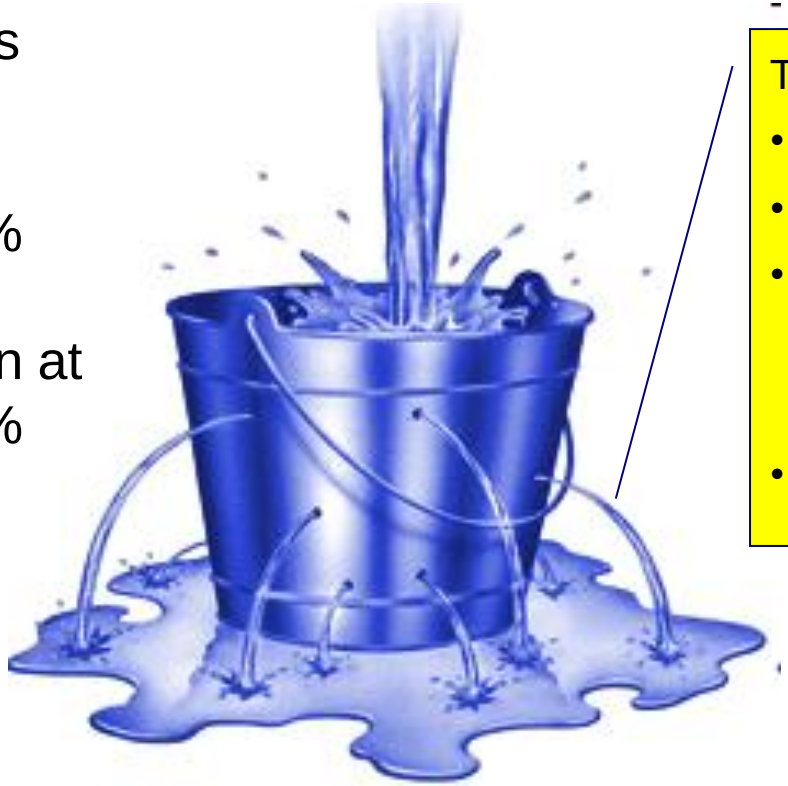
We got a serious leaking bucket problem

Net Promoter Score is only at 3.2% overall

Alibaba.com is at 30%

Comparing to Amazon at 73% and eBay at 71%

SEO, Alibaba Traffic, Marketing



The main issues are:

- Seller would not sell
- Counterfeit products
- Low quality and item different from description
- High shipping cost

Qualitative vs. Quantitative Research Methods

- **Qualitative research such as lab tests and field research give us rich data about**
  - Usability problems
  - Discoverability
  - Navigation
  - Terminology
  - More complex problems

- **Quantitative research such as surveys and usage studies**
  - Tells us about how pervasive a problem might be

- **Ideally, the combination of qualitative and quantitative research give us a more complete picture and is most powerful**
    - Field research or lab test (Why and What)
    - Survey or Live A/B Test (How many)