Methods of UX Testing

TABLE OF CONTENTS

WHY UX TESTING MATTERS

We consider UX testing **integral** to creating world-class products that users love. Testing is a necessary expense. It has the ability to reduce the long term cost of a project, while delivering superior performance to end users. For this reason, it yields an enormous value.

There are three reasons why we recommend testing:

1. *To reduce risk.* Reducing risk of failure normally increases the chances your product will succeed.

2. *To reduce cost.* Investing in testing with users ensures that issues get caught sooner, good ideas are introduced faster, and unnecessary product design and development efforts are reduced. In the end, ROI is delivered through minimizing unnecessary iterations and increasing adoption.

3. *To make the product better.* Creating user-centric, award-winning UX typically involves testing.

At Fresh, we conduct UX testing where it makes economic sense to do so. Our goal is pragmatism. Varying levels of testing accompany varying levels of expected ROI (Return on Investment). For larger projects, often with a bigger budget, more testing might make sense – we'd recommend more testing for a 100M e-commerce business than we would for a 1M service business with a marketing website.

UX testing should be **integrated** and understood as a part of the scientific design process. Integrating testing leads to meaningful insights and feedback. These findings inform and validate design decisions that have serious and long-lasting implications.

8 METHODS OF UX TESTING

At Fresh, we utilize 8 Methods of UX Testing as a part of our design process. Not every project requires every type of test. Depending on the context of the project or the nature of the problem being solved, certain types of testing are recommended. In our experience, the following tests have proven to be valuable:

Psychological User Tests

Trust Test
Basic 5 second trust test to measure confidence and credibility

Impression Tests
Basic 5 second impression test to measure gut reactions, including a keyword summary

Comparison/Preference Test
Comparison/Preference tests of 2 or more alternatives with statistical results
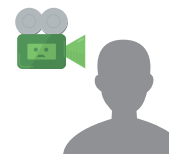
Blur Tests
Blurring images to evaluate what is appropriately or inappropriately calling attention on the page

Traditional User Tests

Time Based User Tasks
Preparing user tasks for key pages and user stories, observing whether users are successful and measuring the time required to complete each task

User Observation with Tasks
Preparing user tasks for key pages and observation of users, often in the target demographic, often with video capture for post-analysis

User Surveys
Various questions to identify pain points, get suggestions for opportunities, rank features, and more

Automated Visitor Screen Recording
See the user's experience via screen recording, analyzing interactions and pain points

TRUST TEST

Basic 5 second trust test to measure confidence and credibility

Trust Tests focus on identifying messaging and credibility issues. They're valuable because they reveal the initial reactions users have to a particular website or product. Would users actually use the site or buy a product? What degree of confidence would they have in using it?

We conduct Trust Tests on *existing websites.* Tests are conducted after a test plan has been created and a hypothesis has been formed. Results are analyzed along with data gathered during user interviews, expert reviews, and web analytics. The test is used to further explore or validate a hypothesis.

E-Commerce Example
A HOME PAGE TRUST TEST WITH 100% FEMALE AUDIENCE SHOWED SIGNIFICANT LACK OF TRUST

8 of 20 said they wouldn't trust the site and 3 said maybe, revealing a 55% overall lack of trust.

Only 37.5% of the people in target demographic (40-60 female age range) showed a lack of trust, revealing that the younger demographic has the most lack of trust.

Example

Fresh conducted a Trust Test on a website homepage to reveal key findings about whether users trusted the website. Short trust tests provide insights into user confidence and site credibility.

IMPRESSION TEST

Basic 5 second impression test to measure gut reactions, including a keyword summary

Impression Tests also focus on identifying messaging and credibility issues. Though similar to a Trust Test, Impression Tests involve more open first impressions with a commentary and a keyword summary that can highlight key reactions users are having. Designers come away with specific words and phrases users use to describe their gut reactions.

Results are analyzed along with data gathered during user interviews, expert reviews, and analytics. This test is used to further validate or invalidate whether design and messaging are working together

We conduct Impression Tests on *existing sites* and *new designs*.

Example

In a 5 second impression test, 20 users are asked to imagine a scenario and look at a specific homepage. In this instance, we collected their first impressions, highlighting concerns in red and strengths in green.

COMPARISON/PREFERENCE TEST

Comparison/Preference tests of 2 or more alternatives with statistical results

Comparison/Preference Tests are conducted to evaluate two or more alternatives, such as two design comps or your design versus a competitor's. These quick A/B tests at the design stage can quickly indicate which options work best, helping designers make decisions quickly and with confidence.

User data is captured via online surveys with users in the target demographic OR in-person, with an observation of the user analyzing the two designs.

We conduct Comparison Tests on *existing sites* or *new designs.* Comparison Tests are conducted after a hypothesis has been formed. Results are analyzed along with data gathered during user interviews, expert reviews, and analytics. The test is used to validate or invalidate the hypothesis.

In the example below, we validated that a new design was chosen significantly more than the competitor, whereas with the old site, the competitor was chosen more often.

CHOSEN
29 times

SUCCESS RATE
90.6%

AVG CHOICE TIME
19.7 seconds
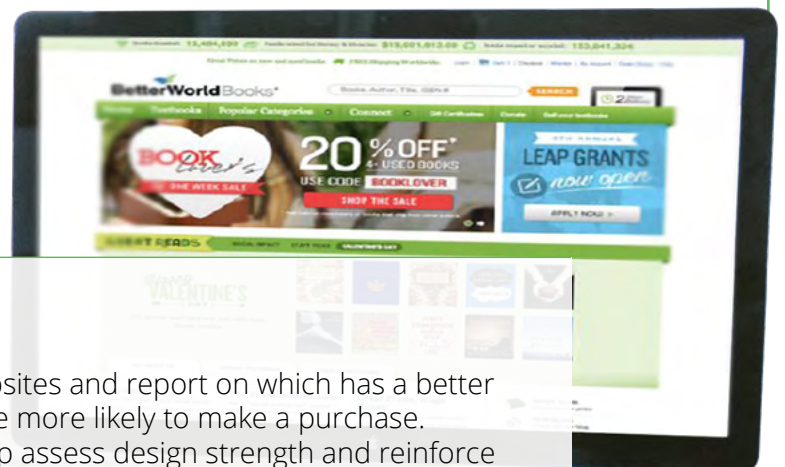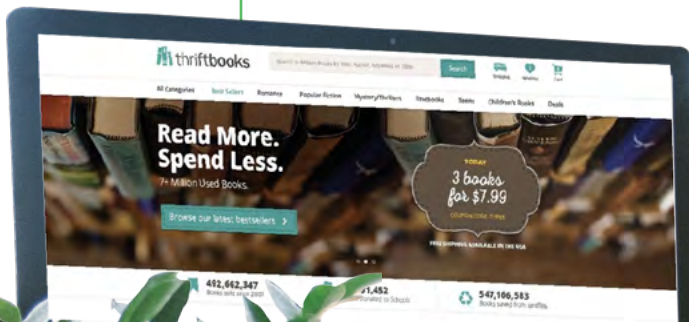
CHOSEN
3 times

SUCCESS RATE
9.4%

AVG CHOICE TIME
29.3 seconds

Example

Users are asked to compare two websites and report on which has a better experience and from which they were more likely to make a purchase. Statistics gathered from the tests help assess design strength and reinforce decisions.

BLUR TEST

Blurring images to evaluate what is appropriately or inappropriately drawing attention on the page

Blur Tests are used to analyze whether or not users are able to identify important images and Calls to Action (CTAs).

The test is subtle and makes it easy to decipher if we are using color and layout appropriately to call attention to what's most important. Is the layout and hierarchy attracting users' attention to the most important areas of the site?

The blurred images are created using Photoshop, or another graphic design tool, to layer over an existing screenshot. Copy, images, and other important site features are visually obscured while colors remain. The test is conducted to see if, when obscured, the CTAs or design elements are striking enough to draw a user's attention.

We conduct Blur Tests on *existing sites* or *new designs.* Results are analyzed along with data gathered during user interviews, expert reviews, and analytics.

Example

Using a Blur Test, we observe whether users are able to identify key Calls to Action. In this instance, it was concluded that Zappos' subtle use of bright contrast draws attention to important areas.

TIME BASED USER TASKS
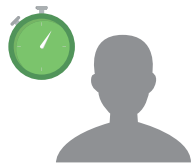
Preparing user tasks for key pages and user stories, observing whether users are successful and measuring the time required to complete each task

Time Based Tasks are used to investigate the flow of the design. We seek to test the design, correcting layout and work flow issues to minimize the amount of time and thought that go into completing a task.

The format is simple: users are given specific tasks to complete when using the site. Researchers

identify whether users are able to complete these tasks successfully and in how much time. The time and efficiency directly relate to usability.

We conduct Time Based Tasks on the current site, new designs, prototypes, and the new site. The data is used to clearly understand what's working efficiently and what's not.

Example

Time Based Tasks below show that users completed the task, but a few people had issues. Because this task is essential to using the site effectively, looking into issues with the tasks that take a significant amount of time is important.

Completed

Not Completed
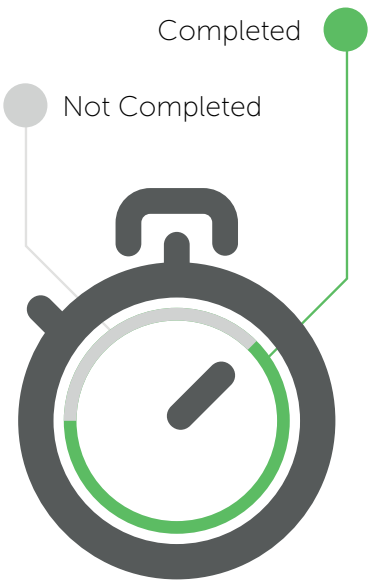
PARTICIPANT	YES	NO	TIME (S)
USER 1	1	0	10
USER 2	1	0	20
USER 3	1	0	15
USER 4	1	0	10
USER 5	1	0	45
USER 6	1	0	10
USER 7	1	0	15
USER 8	0	1	85
USER 9	1	0	65
USER 10	1	0	24
AVERAGE			24

USER OBSERVATION WITH TASKS
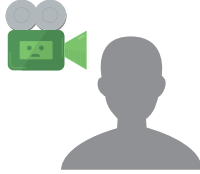
Preparing tasks for key pages and observation of users, often in the target demographic, often with video capture for post-analysis

User Observation allows us to capture the user's experience in completing the different tasks. This is directly related to time based tasks, and adds an extra layer of understanding for designers.

Video recordings, with users talking through the process, allow the team to capture both the screen interactions and the emotional response.

We conduct Time Based Tasks on the current site, new designs, prototypes, or new experience. The data is used to evaluate large and small pain points and opportunities for improvement.

Example

The team can gather both an audio and video recording of users as they complete tasks. This could include a user's desire to see a more prominent "About" section, an emotional reaction to aesthetics, confusion about layout, or a desire for a more streamlined path to checkout.
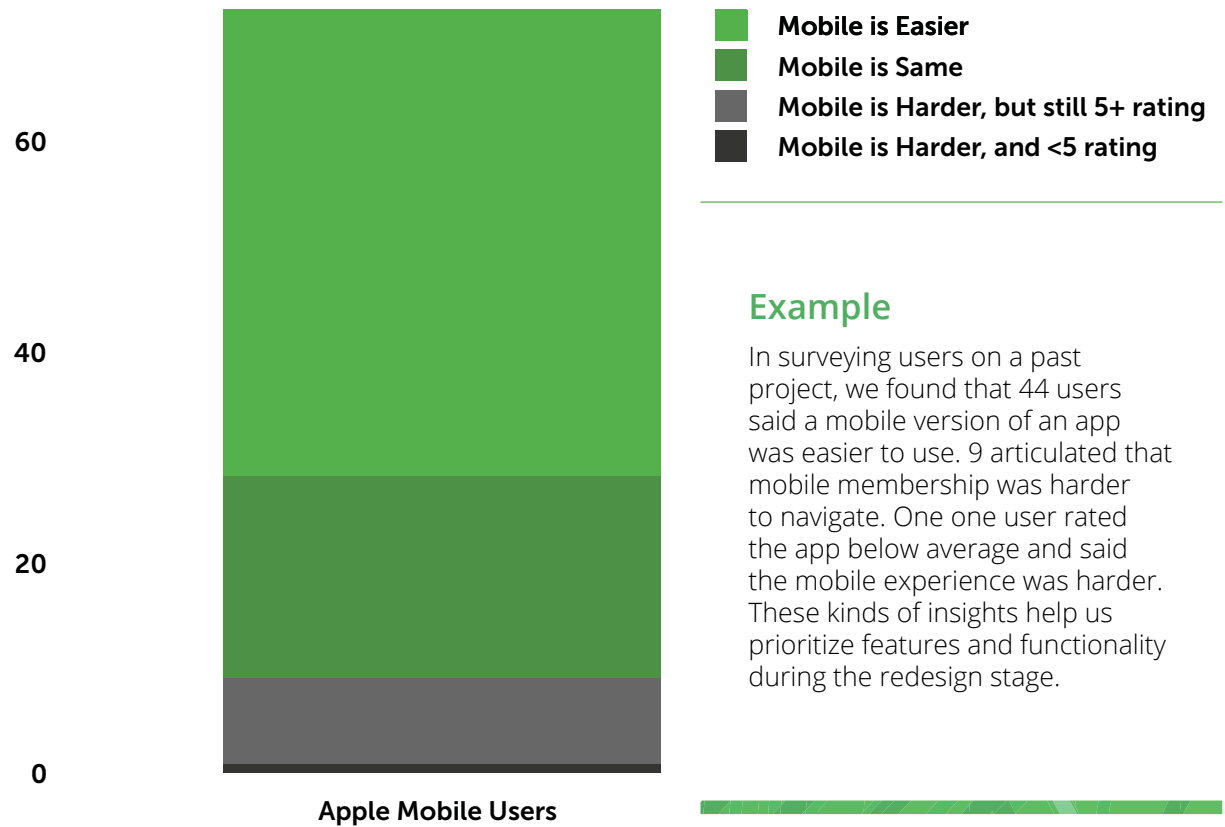
USER SURVEYS

Various questions to identify user pain points, get suggestions for opportunities, rank features, and more

User Surveys are issued to gather quantitative and qualitative data about users and their reactions to a design or concept. Well-written surveys offer the opportunity to quickly and easily gather feedback from a larger sample of users.

Using survey software such as Google and Survey Monkey, we gather user responses to identify

actionable areas for improvement. Due to the way that surveys are constructed and distributed, this process allows us to tap into a variety of issues.

We distribute and analyze surveys during the *Design Testing & Research phases,* yielding valuable data that leads to informing the design process.

■ Mobile is Easier
■ Mobile is Same
■ Mobile is Harder, but still 5+ rating
■ Mobile is Harder, and <5 rating

Example

In surveying users on a past project, we found that 44 users said a mobile version of an app was easier to use. 9 articulated that mobile membership was harder to navigate. One one user rated the app below average and said the mobile experience was harder. These kinds of insights help us prioritize features and functionality during the redesign stage.

Apple Mobile Users
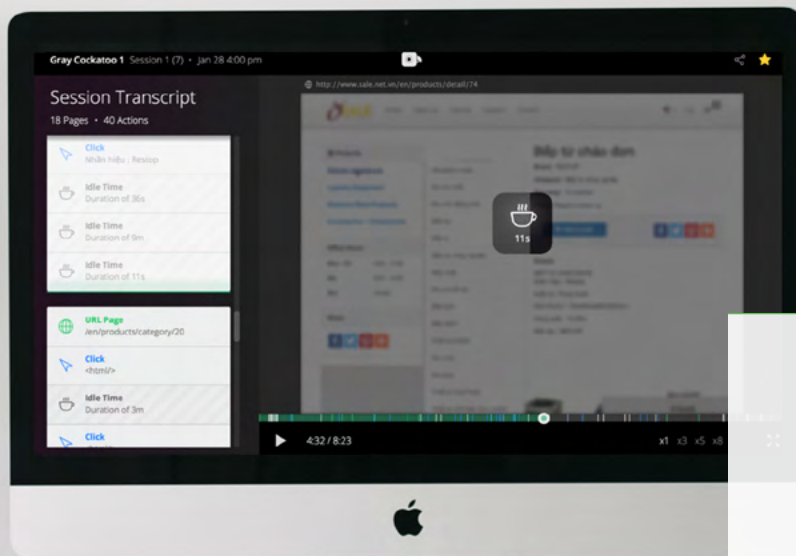
AUTOMATED VISITOR SCREEN RECORDING

Visualize the user's experience via screen recording, viewing complex interactions and product pain points

Automated Visitor Screen Recording is used to visualize the user's experience through a screen recording, viewing complex interactions and product pain points. The method goes beyond analyzing what users are doing, and gives some insight into why and how they are doing it.

Software solutions (like Jaco) do require adding a small embed code to your site to start recording user interactions and UX pain points.

We conduct Automated Visitor Screen Recording during the *Design Testing & Research phases*. The data is used to uncover flow issues, see real click behavior, and inform the design process overall.

Example

To analyze our own website, we used Automated Visitor Screen Recording. Given the ability to visualize UX – the time users spent on each page, where they moved their mouse, how they navigated – we've been able to continue improving our site post-launch.

WHEN TO TEST

There is an optimal point for testing a feature of a product. Testing should be conducted at the right moments throughout the design process.

For example, visual design should ideally come after wireframe layouts have been tested. We recommend testing early iterations of a website or application before new design work has been done. Otherwise, there might be inherent architectural design flaws that no amount of visual design can fix.

We test as soon as it's necessary. Issues identified later are more expensive to resolve.

"When" you should test connects to "why" you should test. Testing should take place because it makes economic sense to do so. To get the greatest return from results, testing early is strategic to re-design work. Later testing can validate design improvements.

Methods and Stages

METHODS	Current State Experience	Redesign: Mockup	Redesign: Prototype	New Experience
Trust Test	YES	YES		YES
Comparison Test	YES	YES		YES
Impression Test	YES	YES		YES
Blur Test	YES	YES		YES
User Testing Time Based Tasks	YES		YES	YES
User Observation with Tasks	YES		YES	YES
User Surveys	YES			YES
Automated Visitor Screen Recording	YES			YES

CONCLUSION

Conducting no user testing is a risky proposition. The reality of UX is that it is a fuzzy process of well-educated trial and error. We test because testing decreases the probability of failure. It also decreases the scale and impact of failure.

Although the testing process is extensive, it can be modified depending on the project. Testing goes a long way towards determining the success of a design. We recommend developing a concrete process that allows you to test consistently with one or more methods. By challenging your assumptions, designing for the end user, and following a formal process, you can reduce the costs associated with testing and get the ultimate return on value.

To reiterate, user testing is **integral** to delivering a world-class product. Yes, it's important to be pragmatic and economical when it comes to testing, but we consider it an essential cost of creating a high-quality user experience.

Testing should also be **integrated**. As shown in the previous table, testing doesn't strictly fall at the beginning, middle, or end of the design cycle. Rather, it should be a key component of design throughout the entirety of the process.
Test to reduce risk, reduce cost, and make your product better. We recommend these **8 Methods of UX Testing** as a strong place to start.