

## Keynote Customer Experience / UX Method

The Keynote methodology takes a new approach to customer experience testing by combining the critical elements of traditional usability testing, market survey research, and log analyses into a single evaluation in a fast, cost-effective manner.

Large samples of users evaluating a Web site remotely provide more reliable and representative data than traditional usability testing. The task-based context allows for meaningful interpretation and aggregation of users' behavior. This approach has the unique advantage of structuring the different aspects of customer experience, including brand positioning and expectations, users' behavior, and subjective experience of the site itself, into one analysis.

Keynote's research methodology is thus able to provide accurate, reliable data on the Web user's experience and how that experience can be improved. Thus, site managers discover how they should allocate resources for maximum impact, while designers obtain insight into why particular features and functions are not working as planned and how best to modify them.

### Keynote Research Applications and Usage

Many different methods can be used to conduct Web research, but it is essential to choose the right tools to accomplish specific research objectives. Keynote outlines a Research Roadmap following the product development

cycle to help its clients select from the different tools and methodologies available. Stages of the product development cycle can be divided into four general categories:

- Discovery
- Exploration
- Evaluation
- Tracking

The Keynote applications appropriate for each stage are discussed below.

### Market Discovery: Assessing Market Needs to Determine Strategic Direction

Before any program begins, successful ventures start with a deep understanding of the market in order to develop a winning strategy. Discovery and market strategy formulation include the following processes:

- Setting specific business objectives
- Identifying and assembling core business metrics (sales, leads, traffic or support center data)
- Conducting competitive assessments/benchmarking

(baseline)

- Determining drivers of customer behavior/customer insight
- Determining the most important segments

The following Keynote applications are appropriate for these objectives:

### **Keynote Competitive Research**

Keynote Competitive Research provides an in-depth understanding of your customers, competitors, and markets. (see the KCR Home Page for more details)

KCR studies compare top industry sites directly and show how people are currently using the Web for a particular product category. These studies provide comprehensive coverage of an industry, sophisticated analytics, drivers of customer behavior, and best practices. KCR studies objectively and comprehensively measure a website's performance against its key competitors.

### **Brand and Value Proposition Research**

Through a custom engagement, Keynote explores the customer's perception of your brand in the marketplace, isolating the online experience to uncover how exposure to your site impacts brand image - before and after customers interact with your site.

Keynote will provide you with accurate and reliable data for measuring the effectiveness of your online presence. You will also receive actionable insights that will allow you to make changes that immediately enhance your brand image and separate you from the competition.

### **Competitive Intelligence Research**

Keynote Competitive Intelligence Research provides a critical baseline

that enables you to expose your competition's weakness and learn from their strengths. Our head-to-head evaluations - conducted without the competition ever finding out - let you see how your site is performing versus your competitors using core customer experience metrics such as: satisfaction, organization, visual appeal, ease-of-use, likelihood to use/return/purchase, and brand perceptions.

Competitive Intelligence Research will provide you with detailed analysis on competitive sites revealing what is and is not working and actionable recommendations that will help set you apart from the competition.

### **Web Site UX Assessment**

Keynote Web Site UX (User Experience) Assessment Research provides a clear picture of how customers are actually experiencing your site. It helps you eliminate obstacles that cause drop-offs and provides actionable recommendations on how to acquire more customers at a lower cost and drive sales by creating a truly user-friendly experience.

Keynote measures the effectiveness of your Web site against a core set of success factors. Your target customers actually test-drive your site and return invaluable feedback on how your site can better meet their needs.

### **Web Site Design Research**

Keynote Web Site Design Research provides actionable insights that will help you prototype and concept test your ideas before they are put into development. Whether you are launching a new Web site, releasing new products, or pilot-testing new services, our expert team of Web UX researchers will help you every step of the way.

Keynote allows you to accurately gauge reactions with in-depth evaluations from your target audience. As a result, you gain a better understanding of how new functionality and modifications to your site impact customer experience and your bottom line.

### **Continuous Benchmarking Research**

Keynote Continuous Benchmarking Research establishes key metrics to assess the effectiveness of your Web site over an extended period of time. You will be able to see exactly where improvements need to be made and after implementation, how they pay off.

To accurately measure improvements on a site redesign or product enhancement, Keynote will work with you to run before-and-after evaluations or choose to run multiple evaluations on a quarterly basis.

### **Customer Acquisition Research**

Keynote Customer Acquisition Research enables you to follow panelists as they seek information and products on the Web without directing them to a specific Web site. By learning exactly how users navigate the Web and perform transactions, you can make sure you are reaching out to your customers in the most effective way.

### **WebEffective**

WebEffective™ is the technology platform that enables all of the research methods and study types in Keynote's arsenal. It is the only customer experience research tool today that allows you to capture your customer's complete online experience, including client-side interactions with dynamic Web 2.0 sites.

WebEffective is driven by Keynote's extensive infrastructure which includes one of the most

representative panels of over 160,000 real online users from a cross-section of demographics, languages, and broadband and dial-up connections. With WebEffective you also have the flexibility to use other user panels of your choosing or to intercept visitors to your site.

WebEffective combines the best of market research, usability labs and web analytics providing usable data from virtually any website. It supports large user samples to provide reliable information on their attitudes and behavior.

Web 2.0 is changing user's expectations for web design and navigability, and WebEffective provides you with information that allows you to confidently improve the user experience on your site.

### **Keynote Research Panel**

Keynote has its own research panel, the Keynote Research Panel (KRP), which currently consists of over 160,000 Web users with thousands of individuals joining each month. The panel is managed to provide the full spectrum of Internet users and to meet high-demand categories, such as high-income shoppers or business decision-makers.

To combat sampling biases associated with self-selection, Keynote recruits panelists through a broad range of sources including word-of-mouth, e-marketing campaigns, affiliate programs, and invitations on portal sites. Keynote also proactively recruits under-represented groups. The Keynote research team runs internal evaluations periodically to ensure that data from the KRP are similar to data that would be obtained from a true probability sample.

The possibility of unknown biases associated with nonrandom samples is typically of greater concern to the

scientific purist than to the business decision-maker, but there may be some practical implications. For example, panelists who join the KRP are likely to be more Web-savvy than the general population and be more open-minded to exploring new Web businesses. As is true of all market research sampling, sample results must be interpreted with the caveat that statistical estimation assumes a true random sample from the population of interest.

### **Samples for Specific Evaluations**

Keynote can construct samples from the KRP to model the online population or particular target markets by using quotas on particular attributes. For example, if the Web population were 56 percent men and 44 percent women, a 200-person sample would contain quota targets of 112 men and 88 women. Keynote would then randomly sample men and women from the KRP until these quotas are filled.

Although most clients utilize the KRP for their sampling needs, clients may also recruit participants directly from their own Web sites, customer lists, or RDD (random digit dialing) samples. In all cases, Keynote works with each client to generate a sample that approximates the client's target customer profile.

### **Sample Size and Statistical Significance**

Keynote employs larger samples (typically from 200 to 800 individuals) than traditional usability tests, which tend to use five to eight people. Larger sample sizes provide better insights into brand positioning and a more comprehensive picture of customer experience that businesses require to make strategic decisions. These large samples comprise a variety of users, ensuring the representation of many perspectives and the ability

to estimate the magnitude of problems by the percentage of users who encounter them.

Large samples can be analyzed with statistical tests to more accurately interpret comparisons. Larger samples make statistical tests more sensitive for detecting possible differences among groups, or between observed and expected results. Typically, with a sample size of 200, statistics such as chi-squared, t-tests and regression will detect statistically significant differences with a confidence level of 95 percent. In addition, a sample size of 200 is large enough for a statistically meaningful analysis of different types of Web users.

What does "statistically different" mean? Sometimes differences arise simply because of sampling errors rather than true differences between populations. A confidence level of 95 percent means that it is highly unlikely that a particular difference between distributions arose from sampling error alone, rather than true underlying differences.

### **Response Rate**

All market research surveys, usability tests, and customer ratings are affected by response rates—the percentage of invited members of the tester pool who respond within the specified data collection period. Response rate is important, because those who respond may be different from those who did not, rendering the sample less representative. For example, those who choose to participate in an evaluation may be people who are more enthusiastic than usual and so may not fully represent the entire population. Response rates for Keynote evaluations are similar to industry averages for email-related surveys. Response rates are monitored and regulated to ensure that impact on the representativeness of the

sample is minimized.

### **Panelist Variables**

Keynote collects demographic data from each member of the KRP at the time of registration. Keynote then surveys panelists several times per year to update their profiles and obtain more specific data on their purchasing behaviors and interests. Panelists must be at least 13 years of age to participate (18 if the panelist resides in the European Union). In obtaining, storing, and sharing panelist data, Keynote complies with guidelines established by the World Association of Opinion and Marketing Research Professionals (ESOMAR) and TRUSTe.<sup>1</sup>

### **Panelist Incentives**

Keynote offers panelists a token of appreciation for completing an evaluation, such as a gift certificate or opportunity to donate to charity. Keynote panelists sometimes receive incentives to participate in periodic surveys, such as an entry into a cash sweepstakes. However, many Keynote panelists report that curiosity and a desire to improve customer experience motivate them more than the evaluation incentives do. To discourage “professional” panelists, Keynote prevents individuals from participating in an evaluation more than once a month and more than eight times in a 12-month period.

### **Fraud Checks**

Keynote screens all panelist registrations for possible fraud before admission to the KRP, eliminating panelists who sign up more than once under different names or provide obviously false information. The completion rate is the percentage of invited respondents who complete the evaluation and make a good faith effort to complete the objectives. Keynote routinely excludes panelists that do not meet this

“reasonable effort” criterion from the final sample. Of those who respond, the percentage of panelists completing the evaluation with usable data is very high, minimally impacting representativeness of the sample.

### **Scenario-Based Testing**

Like traditional usability tests, Keynote employs scenario-based techniques, where panelists are asked to pursue a structured set of objectives. This approach has the advantage of making both aggregate and individual behavior interpretable. Because users who are provided goals may behave differently from users who have their own goals, Keynote stresses using scenarios and objectives that are similar to what a real customer would encounter. Findings from Keynote evaluations employing scenario-based procedures are highly consistent with findings from log analyses of actual customers, providing confidence that scenario-based testing biases are minimal.

### **Methodological Issues and Concerns**

The success of the Keynote approach depends on the quality of the questions, the sampling procedures, and the validity of the evaluation methods.

### **User Objectives and Questions**

Keynote maintains a library of evaluation scripts with particular objectives and supporting questions that are common to most Web sites (e.g., registration, searching, understanding the core value proposition of the business, or estimating user satisfaction). The Keynote research team conducts internal analyses to determine which questions are most likely to provide insights into the customer experience. In this way, clients gain the benefit from Keynote’s

collective experience with previous evaluations.

## Comparing Multiple Web Sites

Comparing multiple Web sites presents special methodological concerns.

### Within-Subjects Designs

In Keynote's Comparative Evaluation solution, panelists evaluate more than one Web site and attempt the same set of tasks on each site. Each participant thus serves as his or her own control, completing the same objectives on both the client site and competitors' sites (this is called a "within-subjects" design, as opposed to a "between-subjects" design, where each site is evaluated with a different group of panelists). This within-subjects design allows for a direct comparison of users' experiences on the two sites and provides more power for statistical comparisons. The different sites provide an immediate comparison for all statistics. For example, a Web site may have a registration failure rate of 20 percent, which might seem adequate by industry standards. However, a Comparative Evaluation may reveal that the same group of people showed a failure rate of only 2 percent on a competing site, revealing the possibility of significant improvement.

### Learning Effects and Counterbalancing

Keynote's Comparative Evaluation solution eliminates learning effects by counterbalancing the order in which the panelists evaluate the Web sites. Learning effects, also known as order bias, occur when panelists learn to do tasks better on the second Web site and thus do not accurately portray the usability of the second Web site. In order to eliminate potential learning effects,

Keynote reverses the order of the sites for half of the panelists. Thus, Keynote can attribute any differences in customer experience to differences between the sites and not to the order in which panelists evaluated the sites.

### Sample Size for Within-Subjects Designs

Comparative Evaluations employ a within-subjects design such that each panelist evaluates both sites, which eliminates error variance due to individual differences. Although reduced error variance means that even a small sample size is likely to detect statistically significant differences between the sites, Keynote recommends a sample size of 200. The sample size is necessary for meaningful descriptive statistics of market research data to inform strategic business concerns, and to do between-subjects analyses on the first-site data when there are order effects.

### Previous Exposure Effects

Previous exposure effects are similar to learning effects. Panelists who have had previous experience with a Web site in their personal history may evaluate a site differently from someone who is less familiar with it. For example, it would not be surprising that success rates show a leading site as more usable than a less established site, even if the lesser-known site were more user-friendly. Web users may be more experienced with the leading site's procedures and find them easier to accomplish. Panelists can be asked to report their familiarity with the site to determine whether previous exposure may be a factor in their site preferences. Controlling for previous exposure in constructing samples (e.g., selecting panelists who have equal levels of experience with all the sites evaluated) must be weighed against the need to

realistically measure brand awareness among competing sites.

### **Between-Subjects Designs**

Clients can use Keynote evaluations to compare different versions of the same Web site or compare different Web sites, creating a true experiment with random assignment to alternative design solutions. For example, a client may want to compare two prototype registration processes, or compare two different sites' registration processes. Keynote can send a sample of panelists to different versions of the same site with the different registration processes. This method leads to definitive answers as to which design is most effective.

### **Cross-Sectional and Longitudinal Designs**

Web sites can be compared before and after design changes to ensure that design changes produced the intended improvements in customer experience. In a cross-sectional approach (different individuals but similar sample parameters at two different points in time), or alternatively.

## **The Keynote Evaluation Development Team**

Keynote evaluation methodology is developed and refined by consultants, who work directly with clients, as well as with a team of research experts who conduct internal evaluations.

### **Keynote Consultants**

Keynote supplements its core offering through the work of its Professional Services division, which includes a team of in-house consultants. Keynote consultants add value to client projects by guiding evaluation strategy and design, assisting in the interpretation of results, and making recommendations based

upon those findings. Consultants have training in both market research and usability methods, as well as in advanced Keynote evaluation practices. Each consultant has extensive experience developing evaluations for a variety of clients. Their collective experience informs the research team, to allow a continual fine-tuning of the methods and cataloguing of best practices and benchmarks.

### **Keynote Research Team**

The Keynote research team defines and refines the methodology, conducts primary research on best practices and compiles industry benchmarks. This team conducts internal research and stays abreast of scientific developments in measuring customer experience to continually refine the Keynote methodology. In addition, the research team analyzes aggregate results across evaluations to identify best practices and establish benchmarks for comparison with particular evaluation results.<sup>11</sup>

## About Keynote

Keynote Systems (NASDAQ "KEYN") is the global test and measurement company for mobile communications and online business performance.

As an independent and trusted third-party, Keynote provides IT and marketing executives with an unbiased view into their Internet services from around the world. For over a decade, we have been providing measurement data and testing capabilities that allow companies to understand and improve their customer's online and mobile experience.

Keynote Systems, Inc.  
777 Mariners Island Blvd.  
San Mateo, CA 94404

[www.keynote.com](http://www.keynote.com)

---

---

---

<sup>i</sup> For the full text of these documents, see [http://www.esomar.nl/codes\\_and\\_guidelines.html](http://www.esomar.nl/codes_and_guidelines.html) and [http://www.truste.org/webpublishers/pub\\_resourceguide.html](http://www.truste.org/webpublishers/pub_resourceguide.html).

<sup>ii</sup> For review of recommended survey development practices, see G. A. Churchill, *Marketing Research: Methodological Foundations* (New York: Dryden Press/Harcourt Brace College Publishers, 1999). For a review of recommended usability test development process, see J. S. Dumas and J. C. Redish, *A Practical Guide to Usability Testing* (Portland, Ore.: Intellect, 2000).



The trademarks of Keynote Systems, Inc. include Keynote®, DataPulse®, CustomerScope®, Keynote CE Rankings®, Keynote Customer Experience Rankings®, Perspective®, Keynote Red Alert®, Keynote Traffic Perspective®, Keynote WebEffective®, The Internet Performance Authority®, MyKeynote®, SIGOS®, SITE®, keynote™, The Mobile & Internet Performance Authority™ and all related trademarks, trade names, logos, characters, design and trade dress are trademarks or registered trademarks of Keynote Systems, Inc. in the United States and other countries and may not be used without written permission. PR-04-03-SB1