

EPIC2020

Ethnographic Praxis in Industry Conference

CASE STUDY

Scaling Experience Measurement

Capturing and Quantifying User Experiences across the Real Estate Journey

REBECCA J HAZEN, PhD, *Zillow Group*

GENNY MANGUM, *Zillow Group*

TOM SOUHLAS, *Zillow Group*

Zillow is undergoing a major evolution, transitioning from serving as the world's largest digital marketplace for real estate advertising into an end-to-end platform to support customers across the phases of buying, selling, and renting homes. As Zillow expands into more transactional spaces, the company has recognized the need to develop a clear and actionable understanding of users and their experiences as they interact with our products and services. To address this need, we set out to establish an Experience Measurement program to provide organization-wide visibility into how well our Zillow experiences meet users' needs and expectations as they progress through their real estate journey. This program will enable teams to gain insights at the intersection of attitudinal and behavioral experience data and lead us to our end goal of empowering informed decision-making across all levels of the experience and organization.

In this paper, we provide an overview of our approach to establishing an Experience Measurement program at Zillow. We focus on a small subset of Experience Outcomes (XOs) in an initial feasibility study to develop a program that would scale and drive impact in assessment and decision making across lines of business. Finally, we share lessons learned throughout the process of developing and validating the framework and discuss the impact of this work on current and future organizational outcomes.

Keywords: user experience measurement, experience outcomes, scaling experience measurement.

INTRODUCTION

Zillow is a real estate platform that provides customers with access to a broad range of real estate information and services for buying, selling, and renting homes. In recent years, Zillow Group has launched increased services to support customers through the more transactional elements of the real estate journey including rental applications and payments (Zillow Rentals), home loans and refinance (Zillow Home Loans), title and escrow services (Zillow Closing Services), and even buying and selling homes directly to and from Zillow Offers (figure 1).



Figure 1. Zillow Group Services (select)

Motivation

Our interest in creating an Experience Measurement program grew out of a deep commitment to our customers. As a company, we strive to provide high-quality user experiences, and empower our customers as they navigate the transactional journey to unlock life’s next chapter. This dedication is modeled through several of our core values (“Customers are our North Star”) as well as in one of our company-wide Objectives and Key Results (OKRs) (“Deliver an integrated, joyful customer experience”) (Workfront 2020). As we take on new and greater challenges and provide more complex features and services, we must hold ourselves accountable to our customers and deliver high-quality experiences. To do so, we must first define what makes for a good experience, and then measure the quality of those experiences in a comprehensive and actionable manner. As a result, we set out to create an Experience Measurement program that would shape how we think about customer problems and enable us to use measurement constructs to drive high-quality design and decision-making activities.

Related Work

The value of understanding and measuring user experiences, and the importance of user-centered metrics, has been an active area of investigation in recent years. As a result, research and product teams around the world have developed and disseminated new frameworks and approaches to quantifying these elements. One such example is the HEART (Happiness, Engagement, Adoption, Retention, and Task success) framework developed at Google (Rodden 2010), which was designed as a means to leverage user-centered attitudinal and behavioral metrics across five categories to measure progress toward business objectives. This framework took the approach of developing an actionable understanding of how products and services are performing for the user through a synthesis of user-centered attitudinal and behavioral metrics and doing so at scale. Another example is the development of a measurement framework centered around Customer Experience Outcomes that was recently demonstrated at Amazon Prime Video (Morris and Gati 2018). In this paper, Customer Experience Outcomes (CXOs) were formulated based on an understanding of the jobs that customers were doing, developed through a series of in-depth qualitative and ethnographic studies. The authors then took a survey-based approach, coupled with behavioral analysis, to measure each of the prioritized CXOs at scale.

After examining our own user needs and business objectives, we decided to extend and apply similar concepts at Zillow in creating an experience measurement program centered around experience outcomes. For us, the program's success hinged on our ability to capture and objectively assess the quality of experiences as they occur across digital, physical, and human-based touchpoints. This meant identifying what was important and made for a good experience, as well as defining metrics and developing measurement plans to reflect the nature of these experiences. However, the challenge was how to make this a reality in a manner that would scale and keep up with the fast-paced and complex nature of our business.

Program Objectives

Throughout this work, our objective was to develop a program to measure the quality of our user experiences and to empower stakeholders to use this information to drive high-quality design and decision-making activities. We aimed to provide stakeholders with a method of understanding how we are doing within our user experiences and a scaffold for digging deeper into the data surrounding these experiences to contextualize the results. We strived to support these stakeholders in evaluating the impact of design decisions, incorporating these framework elements into a range of evaluative studies from simple usability studies to large A/B tests. And finally, we wanted this framework to support prioritization efforts within product spaces and across lines of business.

This paper describes our process for developing the Experience Measurement program, and then walk through our experiences in an initial feasibility study. We share how the innate complexities of the activities and services offered often required us to rely on different qualitative and ethnographic techniques to investigate experiences and how we worked to achieve scale across diverse domain areas. Next, we discuss lessons learned as we refined our overall process, focusing on depth, efficiency, and ability to scale across the different experience areas at Zillow. We also talk about the value and role of ethnography to inform design and business decisions at Zillow. Finally, we share scenarios surrounding how this framework has been used to date at Zillow in order to (1) empower others to conduct consistent and high-quality measurement research, (2) inform the contextual underpinnings of dozens of attitudinal and behavioral metrics across the business, and (3) contribute to an expanding framework for how organization-wide prioritization will be conducted in the future at Zillow.

APPROACH

Experience Outcomes

The goal of this program was to develop a scalable framework for objectively measuring the quality of experiences across the end-to-end customer journey. The foundations of this Experience Measurement program are grounded in a partner construct known as Experience Outcomes (XO). Drawing inspiration from the jobs-to-be-done framework (Morris and Gati 2018, Christensen 2003, Christensen 2016), the basic unit of XOs are essentially the “jobs” that customers are attempting to accomplish and are employing our products and services to do. XOs are constructed through the extensive synthesis of research insights drawn from

interviews, observations, and other qualitative and quantitative user research conducted over time.

At Zillow, XOs are intentionally solution-agnostic and are designed to reflect the different personas or customer types who interact with our products and services. Each XO is accompanied by a detailed summary that describes ‘what does success look like for this XO,’ and ‘what makes for a good experience when interacting with elements that support this XO.’ We also include a description of how we support these elements within our current user experiences. This information not only helps to further contextualize the XO but also establishes the foundation for what will be measured in the Experience Measurement program.

The Zillow XO framework was developed by user experience research and design teams and was deployed across the company approximately nine months in advance of diving into developing and validating the Experience Measurement program. By that point, XOs had become a common currency of sorts, helping to provide context and create a shared understanding of the problems we were aiming to address through our solutions. With XOs now widely accepted and thus the foundation built, we decided that it was time to take the next step and fully operationalize the XOs into more actionable concepts that could ultimately be measured.

1. Developing a Framework

With these objectives in mind, we set out to create a measurement program that would transform our existing XOs into actionable concepts and operationalized measurement strategies. The result was an Experience Measurement framework (see figure 2).

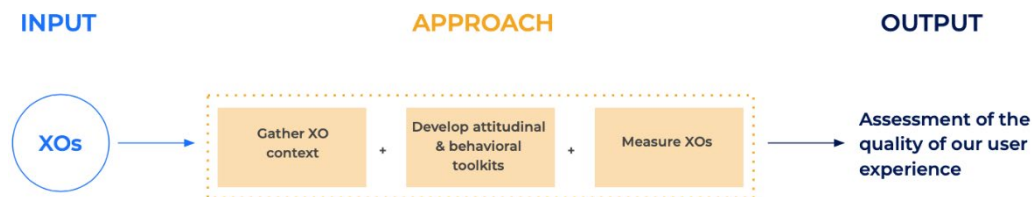


Figure 2. Experience Measurement Program Framework.

1.1 Defining Key XOs: The first element of the Experience Measurement framework consists of defining the prioritized or key XOs. At Zillow, we have over 150 different XOs representing the jobs being conducted by the various personas across the phases of the end-to-end real estate journey. We decided to focus on a prioritized set as operationalizing measurement plans for all XOs at once would be logistically impossible. Prioritization efforts involved a cross-team synthesis of research findings and insights. We then selected a prioritized set of XOs based on their importance to the customer and the opportunity space for the business.

1.2 Defining Success Criteria and Mapping to Zillow Experiences: In our framework, XOs are accompanied by success criteria: a set of high-level concepts that further define what makes for a good experience for a given XO. These criteria not only provide important context to those working in the space, but they also define the critical elements of success that we will measure at a more granular level later in the program. Once we define our success criteria, we then formally map out how and where these elements occur within our current user experiences, both on and off the Zillow website. This process helps to validate the criteria and lays the groundwork for future toolkit development, as decisions around how, where, and when to capture user data are considered.

1.3 Developing Attitudinal Toolkits: Early in developing this program, we determined that a comprehensive understanding of the user experience requires a synthesis of both attitudinal and behavioral inputs. As a result, we set out to create a set of attitudinal and behavioral toolkits to create structure and empower our teams with the tools required to develop well-constructed measurement plans. Attitudinal toolkits are designed to define approaches to capturing attitudinal data in support of our overall understanding and assessment of the effectiveness of our experiences and solutions. These toolkits consist of survey questions, measuring the high-level success criteria and XO satisfaction, and targeting strategies outlining where and when in the user experience these questions should be deployed to customers.

1.4 Developing Behavioral Toolkits: We also developed behavioral toolkits to guide the types of behavioral data to be collected and analyzed in order to support our overall understanding. The behavioral toolkits were developed in partnership with data science and analytics partners and were aimed at understanding and analyzing behavioral metrics and components closely associated with the success elements of the experience.

1.5 Measuring on Experiences: The final element of the program involved capturing data to drive analysis. In addition to the goals outlined in the program objectives, the outcomes of these analyses are intended to inform the prioritization of key XOs, reinitiating the framework process.

2. Validating the Framework

Our next step in this work was to conduct a small feasibility study to validate the framework, and generate lightweight, scalable protocols to facilitate framework adoption across the product spaces and experience areas at Zillow (figure 3).

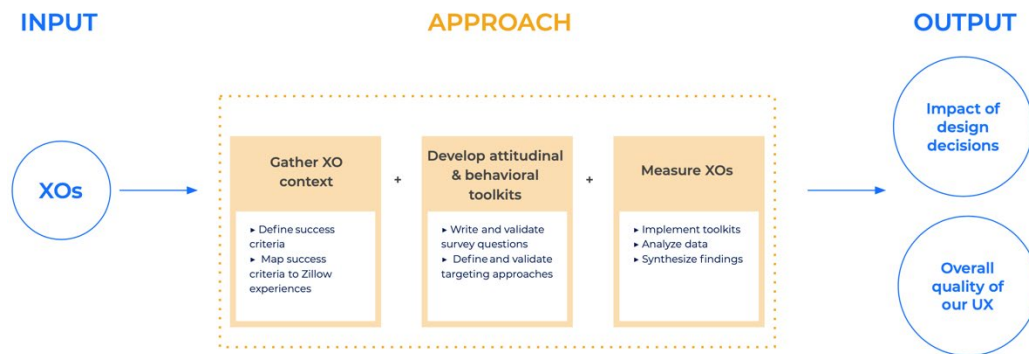


Figure 3. Experience Measurement Framework, expanded.

2.1 Defining Key XOs: We selected a small sample of XOs to focus on for the initial feasibility study. Unlike our typical prioritization process, we selected these XOs based on the anticipated value that they would provide toward our efforts to validate the framework and scale the supporting protocols. We chose certain XOs because we felt that they would generally be representative of other prioritized XOs based on the type and frequency of tasks involved. Others were selected based on potential challenges and opportunities that we anticipated, such as those involving smaller populations or doing more complex activities. We acknowledged that many of the activities associated with each of the initial XOs were taking place as on-site activities; we are currently in the process of planning a separate feasibility study examining the application of this framework on how we measure experiences that are taking place beyond the Zillow website and in scenarios that involve both on- and off-site experiences.

For the purpose of this paper, we will focus on a single XO: ‘Help renters find homes based on stated criteria.’

2.2 Defining Success Criteria and Mapping to Zillow Experiences: Although we had previously developed success criteria for the majority of the XOs on our map, we determined that a new approach to defining these criteria was necessary for us to fully operationalize each XO for the purpose of this program. We went through a process of re-establishing contextualized success criteria for each XO, building a shared understanding with design and product partners of what success looks like, and what makes for a good experience. This approach relied on a variety of qualitative techniques, including interviews and contextual inquiry, to understand and explore these questions and capture the necessary context surrounding how users approach the task.

For the XO ‘help renters find homes based on stated criteria,’ we conducted semi-structured interviews with relevant individuals, coupled with observations to learn about the activities they were doing. We spoke with individuals who used Zillow as their primary resource for these jobs, and those who either did not prefer or did not use Zillow to capture a range of input and experiences. Through these activities, we probed on what this XO meant to users, and how they go about the task(s), and what makes for a good experience in doing so. We also took the approach of examining a parallel example, asking participants to

talk through similar questions for a similar activity in an entirely different domain. We found that this type of activity helped explore tacit elements of the experience and surface elements that people may value, but not fully realize or readily verbalize. We then examined what, if anything from these explorations, also applied to the rental scenario, then reconciled findings.

After exploring these questions and synthesizing the resulting insights into contextualized success criteria, we next examined the manifestations of these elements within our own user experiences. This activity aimed to create a shared understanding of where, when, and how these elements take place to inform survey targeting decisions for our toolkits and provide context to design and product stakeholders when interpreting findings from the XO measurement activities.

This manifestation activity was driven by insights arising from observational data and was conducted in collaboration with stakeholders from design, product, marketing, and analytics teams. In the case of the XO ‘help renters find homes based on stated criteria,’ this was relatively straightforward as the behaviors comprising this XO were easy to observe through standard user sessions. We had other XOs, however, that proved to be much more complicated. For these XOs, we had to rely on non-traditional techniques to indirectly “observe” users as scenario-based sessions failed to capture the range and authenticity of natural user behaviors. In one case, we opted to leverage FullStory (FullStory, Atlanta, GA), a digital analytics experience platform previously implemented within several product areas at Zillow, as it allowed us to remotely observe reconstructed user sessions associated with the XO. Although not initially implemented with this use case in mind, we found the tool to be useful to facilitate these observations as we felt that it would provide unique access to behavioral data and reduce biases associated with observation and scenario-based task observation. FullStory provided access to a great amount of unique and unobtrusive observational data, however, we acknowledged that the data was limited to digital interactions and did not provide any insight into verbal or emotional displays that were occurring as the users interacted with the user experience. To address the limitation, we then triangulated insights gathered from these observations against findings from the interview studies to analyze these emotions and behaviors and form a complete picture of the user experience. This activity helped to solidify our understanding and mapping of these manifestations before further operationalizing the metrics.

2.3 Developing Attitudinal Toolkits: We next started developing our attitudinal toolkits to assess the quality of our experiences and the effectiveness of our solutions. This involved creating survey questions and targeting strategies that we could use to capture attitudinal data about these user experiences. In our framework, the type and number of questions included may depend on the measurement scenario and objective. In each case, we took the approach of measuring at the level of success criteria element and XO satisfaction.

One important component of these toolkits is the ability to directly capture attitudinal data in the context of the user experience that is being assessed. Recognizing that a “one-size-fits-all” approach to data collection would be unlikely to meet the needs of the broader organization, we invested a significant amount of time and effort into determining both how and when to best capture data. This approach was technically challenging, but we believed it would provide us with higher quality, and more relevant, actionable data. The core element of this understanding was at what point in the experience has the user had sufficient

exposure to be able to confidently answer these questions. Determining when, where, and how to deploy in-context attitudinal data collection relied heavily on insights coming out of the observational studies and interviews, and behavioral data extracted in partnership with our data science partners. We also had to keep in mind the importance of data quality and minimize the impact on the overall user experience when developing these strategies.

In the case of our XO ‘help renters find homes based on stated criteria,’ we decided to use a site intercept survey that would prompt users for responses to a set of attitudinal measurement questions directly in the context of performing the activities of interest. The corresponding targeting strategy outlined details including the page to display the survey on, the amount of time delay before launch, and variables such as the number of pages visited prior to survey eligibility. We also worked with our data science partners to determine sample size and study duration, based on existing behavioral data. We then ran a qualitative study to validate both the survey questions (validity, reliability, comprehension) and targeting strategies before launching data collection.

2.4 Developing Behavioral Toolkits: The process of developing behavioral toolkits involved examining existing behavioral metrics, taking into consideration the user experience lens. Along with our product and analytics partners, we examined existing metrics and worked to determine how well each captured the elements of success we had previously outlined. In the case that metrics were missing or ill-defined, we proposed new metrics. Because behavioral metrics are often used to evaluate and communicate success across the business, we felt it essential to examine and validate these metrics.

2.5 Measuring on Experiences: The final activity was measuring on the user experiences, according to the strategies we had developed. For this study, this meant launching site intercept surveys and capturing corresponding attitudinal and behavioral data. For the feasibility study, we captured over 5,000 complete responses across our desktop and mobile web platforms. We analyzed these attitudinal data in conjunction with basic behavioral metrics to develop a high-level assessment for the XO and then explored relationships to identify additional research opportunities to contribute to our overall understanding (research synthesis).

LESSONS AND IMPACT

Lessons Learned

Although we have had many successes throughout this work, we have also encountered several learning opportunities along the way. One lesson that surfaced early was the value of gaining buy-in and alignment with cross-disciplinary partners early in the process and continuing to foster that sense over time. We found that establishing initial buy-in through reflecting value and alignment with the things important to each of the stakeholder groups, and then building on that sense by including stakeholders in conversations and research activities throughout building toolkits provided benefit beyond expectation. Having participation from these groups along the way also made it easier for stakeholders to see potential applications, and to contribute to shaping the overall program and further informing outcomes. We also found that working from an existing framework (XOs) that

many people across the company were already familiar with helped gather early support/buy-in.

One aspect that we had to spend a relatively significant amount of time and energy on was finding the balance between quality and flexibility in creating repeatable research protocols to truly achieve scale. This program is a massive undertaking that requires contributions from dozens of researchers and product teams across domains of the business. Because we did not have the luxury of having everyone drop everything and focus on building out toolkits for a month or two, making these activities easy to adopt as a part of current research flows was important. This was initially a challenge, however, we found success through identifying opportunities to more naturally incorporate these elements within common research activities and were able to create and demonstrate variations on the research protocols, depending on factors like time and availability of resources. This made adoption much easier to envision, and also helped to make rapid progress within a more lightweight and streamlined process.

Similarly, because moving fast is important, it was also incredibly beneficial to leverage existing knowledge and tools whenever possible. This helped us to more quickly operationalize the XOs and gain access to insights that might otherwise be difficult to capture.

Organizational and Business Impact

Though still relatively early in this work, we have already seen a significant amount of buy-in across the organization, including within Senior Leadership, and an impressive rate of adoption among research and product teams. Since the initial feasibility study, we have empowered researchers across our team to engage in the process of building out attitudinal toolkits for a set of nearly 40 prioritized XOs across several key lines of business, using consistent and scalable research protocols. We are also scaling behavioral toolkits for these same XO areas. This has allowed us to move quickly in introducing this program to a broad range of teams, promoting the value of user experience measurement and XOs across the organization. We have also seen teams sharing findings related to early measurement work; through this, are beginning to see evidence of how the XOs and measurement can be used on a larger scale to help inform planning and prioritization efforts among design, product, and business teams.

Another major outcome that we are seeing is the value of leveraging insights gathered throughout qualitative and ethnographic studies to encourage conversation and evaluation of existing metrics in place across the organization. This served to inform the contextual underpinnings of dozens of attitudinal and behavioral metrics relied upon by team across the entire business. Historically, teams have relied on these types of metrics business they are easy to communicate, and quickly and concisely convey meaning. By encouraging conversations and reframing these metrics around the concept of user experience and measurement on XOs elements, we are further reinforcing the value of user experience as a higher-level factor in our metrics and success as an organization.

CONCLUSION

We started out in this work with a goal to develop an actionable understanding of how well we are meeting the needs and expectations of customers as they interact with our products and services across the real estate journey. As a business, and within individual product spaces, there is also a need to know what to prioritize or focus on next, often in the face of competing priorities. We felt strongly that grounding these decisions in the user experience aligned with our values and would enable more informed decision making. Through this work, we developed a scalable Experience Measurement program for defining survey strategies and driving high-quality contextual data collection across experiences in the real estate customer journey. In leading with a qualitative, ethnographically-informed approach, and leveraging data and technology solutions already in place, we were able to move fast to meet organizational needs without compromising on research depth or quality. This approach also served to build empathy and strengthen the cross-disciplinary partnerships that were essential to the success of our ability to measure these experiences. We believe that this framework will empower teams to quickly and effectively define and capture data to evaluate the success of product, design, and business decisions, a capability critical to our evolution as a company.

AUTHORS AND CORRESPONDENCE

Correspondence or requests for additional information can be sent to the primary author (RH) via LinkedIn ([LinkedIn.com/in/RebeccaHazen5349](https://www.linkedin.com/in/RebeccaHazen5349)) or email rebeccahaz@zillowgroup.com.

REFERENCES CITED

- Christensen, Clayton M, Taddy Hall, Karen Dillon, and David S. Duncan. 2016. "Spotlight on consumer insight: Know Your Customers' "Jobs to Be done." Harvard Business Review.
- Christensen, Clayton M and Michael E. Raynor. 2003. "The Innovator's Solution: Creating and Sustaining Successful Growth." Boston, MA: Harvard Business Review Publishing Corporation.
- Morris, Lauren and Rebecca Gati. 2018. "Humanizing Quant and Scaling Qual to Drive Decision-Making." *Ethnographic Praxis in Industry Conference Proceedings*: 614–630.
- Rodden, Kerry, Hilary Hutchinson, and Xin Fu. 2010. "Measuring the User Experience on a Large Scale: User-Centered Metrics for Web Applications." *Proceedings of CHI 2010*.
- Workfront. 2020. "What are OKRs?" <https://www.workfront.com/strategic-planning/goals/okr>