

Testing Your Alexa Skill With Real World Users



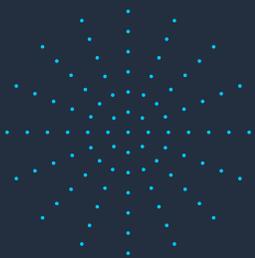
Introduction



When designing and developing an Alexa skill, the goal is to build an experience that is engaging, valuable, and easy for the customer to use.

While following best practices and testing the skill yourself is important, the best way to understand how your skill will be received is to test, early and often, with real users. This provides a window into how people unfamiliar with your skill will perceive, explore, and understand it, and almost always uncovers overlooked issues and opportunities for improvement.

Testing your skill with real world users requires preparation, coordination, and the right tools. It's important to make sure you've got testers who are representative of your target customers, that you provide testers with the right instructions, and that you're able to capture as much valuable data from the testing sessions as possible.



In this User Acceptability Testing Spotlight, we'll cover the **important aspects of user acceptability testing and how you can incorporate them into your skill development process**, to ensure you'll launch a skill your customers will both understand and enjoy.

What Is **User Acceptability Testing** and What Are the Different Types?

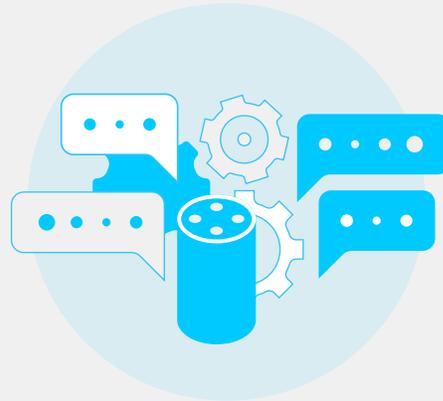
User acceptability testing means putting your skill in front of real people, observing how they interact with it, interviewing them afterwards, and then iteratively improving your skill based upon the data you collect. It can, and should, be integrated into every phase of skill development, but how you implement it depends upon where you are in the skill development process.

There are two main types of user acceptability testing:



“WIZARD OF OZ” TESTING

The name of this type of user testing references the “man behind the curtain” from the movie “The Wizard of Oz.” This is user testing you can do when your skill is in the idea phase, and before you’ve built or developed anything. You first write out some sample dialogues and responses for your skill, and then ask a user to interact with your intended skill based upon these samples. However, instead of responses coming from Alexa, the responses come from a person filling Alexa’s role, and providing responses based upon the samples. This type of user testing can be useful in determining what initial user expectations for your skill will be, and how those expectations will be expressed.



BETA TESTING

Once you’ve started developing your skill, you’ll want to test what you’re building with real users. You can do this before the skill is publicly available by using the beta testing features from Alexa. Beta testing allows you to invite people to try your skill before the skill is certified and released. Those people you invite must first accept your invitation and can then use your skill on any Alexa-enabled device registered to their account. This lets you get feedback on features you’ve built before the entire skill is ready for release.

Best Practices for User Acceptability Testing

The feedback and ideas you get from a well-designed beta test can be extremely valuable.

There are many factors that go into setting up, conducting, and analyzing a beta test, and it's important to follow best practices when doing so, to make sure the feedback you get is as valuable as possible, and you're able to effectively apply it to build the best skill you can.



GETTING THE RIGHT TESTERS

- The most critical factor in user acceptability testing is the testers themselves. The testers should be representative of the skill's target customers, have no greater familiarity with the skill than the target customers will have, and be willing to provide open and honest feedback.
- A good rule of thumb when it comes to finding user acceptability testers is the easier it is to find them, the less valuable they will be. It might be easy to get your colleagues and coworkers to test the skill, but they're unlikely to be representative of your target customers. Friends and family may be willing to test your skill as well, but unwilling to give you critical feedback.
- The best testers are those who represent a cross section of your target customer population. You can locate testers matching these criteria by doing your own recruitment, or by utilizing a testing company that provides their own pool of testers.



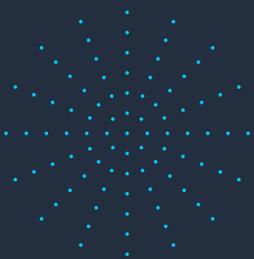
WATCH AND LEARN FROM WHAT YOUR TESTERS DO

- When conducting a user acceptability test, the most valuable data you receive comes from directly observing testers as they interact with your skill.
- Understanding exactly what they do, and when, provides insight into user expectation, confusion, and desire that lets you pinpoint which interactions can be improved, and how. Real user observation enables you to prepare your skill for the range of potential utterances your customers might use, and understand not just what your customer will want, but how they will express it.
- One way to get this data is to directly observe your testers while they use your skill. When doing this, it's important to be as unobtrusive as possible – people behave very differently if they know they're being watched in real time, and you want the tester experience to be as close as possible to the customer experience.
- Another way to observe a testing session is to record it. This has the advantages of being less intrusive and providing you with a record to which you can refer and review. To get these recordings, you'll either need for your testers to come to you and use your equipment, or for them to use their own equipment and send you the recording.



ASK THE RIGHT QUESTIONS

- After a tester has used your skill, you have a great opportunity to ask them questions about the experience and learn from their responses. While this data cannot replace direct observation, it can valuably supplement it.
- When asking questions about the experience, it's important to phrase them so you can understand what the customer really thinks. For example, people tend to misunderstand how they would behave, and will frequently say they will do things they don't end up doing. This issue can come into play if you ask, for example, whether the tester thinks they would use your skill. One way to get around this is, if your skill is attempting to improve a task your customer is currently doing, to ask the tester how they currently perform the task, and whether using the skill is a better experience.
- It's important to be aware of question fatigue – testers get tired of asking questions, and answers towards the end of a long list of questions tend to be answered in haste and without much thought.
- Beware of multiple questions that ask essentially the same thing. For example, asking the customer what they liked most about the skill, and then asking for their favorite part of the skill, are very similar questions, and many testers won't know the difference, or share the same understanding of what the difference is.
- Keep your questions few and clear, and make sure they count.



“To design the best UX, pay attention to what users do, not what they say. Self-reported claims are unreliable, as are user speculations about future behavior.”

– Jacob Nielsen, Nielsen Norman Group

Tools and Resources



Pulse Labs is a company providing a platform and a panel for testing Alexa skills. The Pulse Labs testing platform lets skill developers set up beta test invitations, instructions, and post-test questions. Testers can then use the platform to test any skill, and complete transcripts from the testing sessions, including audio inputs from the testers, are provided back to the developer so they can understand everything that happened during a testing session. Pulse Labs also maintains a large, diverse, and vetted panel of beta testers, which you can use to test your skill with your target customer demographic, and get valuable feedback.

Learn more at: <https://www.pulselabs.ai/testing/>