Dinosaur Educator Toolkit

Mara Varady, Rebecca Henry, Len Shirase, Gauri Govil, Adam Mason

Field Museum of Natural History - Chicago

SI 487
UX Validation Report

April 6, 2018

Study Design & Methodology

Research Questions & Metrics

For our validation study, we wanted to evaluate participants' interest to engage with our design. In addition, we wanted to make sure the titles to our activities were clear in presenting what each activity would be. In general, we wanted to validate that our design was fun, interesting, and clearly presented the learning objectives we are trying to teach. The three research questions that we focused on specifically to evaluate these goals were:

1. Do the names of the islands communicate the idea of learning about dinosaur bone structure and bone function?
2. Are the activities interesting and informative?
3. What things are missing from the toolkit?

The metrics used to help answer these questions were more qualitative in nature as we used small focus group testing as our primary form of validation. Due to the small sample size of this tests, traditional statistical analyses would not be able to yield significant results. Therefore, the metrics we used included quotes from the participants about their thoughts on a given activity, whether certain activities were confusing or clear to the participants, and if they were able to grasp the learning objective presented.

Study Approach

Since our project had no previous interface or user experience to test against, A/B testing and other large scale comparison tests would not work for our project. Another thing we had to consider was our primary audience. Since we are designing primarily for elementary and middle school kids, we wanted the participants of our validation study to also be kids. Surveys and other traditional methods of reaching larger audiences would not work since testing with kids is much harder logistically. Remote testing would also be difficult because it would also require a parent or adult to take time out of their schedule to setup a call. In general, getting a large group of kids to test with is extremely difficult because some parents may be unwilling to have their kids to participate, kids are often in school, and other scheduling and logistical issues that may prevent them from participating.

Based on these factors, our study approach was to do focus group testing on a small number of kids we knew (Gauri's brother and his two friends) so that we could get feedback from our primary audience and persona. We conducted this focus group in

person, with all three kids in the same session so that we could get immediate feedback without any barriers of technology. With this method, we could also get more genuine feedback and also see their non-verbal reactions to our design.

In terms of testing with other people that are readily available to us, we concluded that testing with college-age students would not give us meaningful results. The questions we made for our focus group were aimed towards kids and would not be as meaningful to ask to college students. For example, one of the main goals and research questions we wanted to answer was testing kids' engagement and interest in our design and concept. Asking college students about engagement to a platform aimed at children does not make too much sense. In terms of other potential test groups such as teachers (secondary persona), we are planning on setting up testing with the teachers we met in Chicago next week. We would like to set up calls and have one-on-one sessions with the teachers and ask them about their thoughts on our design and the activities. However, if the scheduling with these teachers does not work, then we will send out a survey to the same teachers.

Procedures & Materials

We conducted testing in a group setting with one laptop computer displaying our prototype. We improved upon our mid-fidelity prototypes and incorporated feedback from our previous focus group session with educators to produce hi-fidelity prototypes for this testing session. In the first step of the session, the participants were individually asked what they knew about dinosaurs and where they had learned this information, whether from prior lessons in school, on TV, in the movies, or another experience. Next, the participants were shown the landing page and asked to read a description about the toolkit, and then asked what their perceptions were of the toolkit based off of the description to gauge participant interest. From here, the participants were asked to navigate to the page with the activity names. The participants were then individually asked what they felt each of the activity names meant, and to guess what they would be doing in each activity.

Then, the participants were asked to navigate to the first activity--"Bone Battle" and read the activity description. After reading the description, the participants were individually asked if they felt the description had sufficient information about the activity and if they believed any information missing. Following this, participants navigated through the activity and observed the design, information, and at the end, were asked what their takeaway from the activity was. After completing the first activity, participants were asked to navigate to the second activity--"Function Fight" and read the activity description. After reading the description, the participants were individually asked if they

felt the description had sufficient information about the activity and if there was any information missing, similar to the first activity. Following this, participants navigated through the activity and observed the design, information, and at the end, were asked what their takeaway from the activity was.

We primarily asked the participants to complete tasks related to navigating through the prototype because we wanted to make the process very intuitive and to make sure it flowed well. Examples of survey questions we asked them such as "What do you initially think of when you hear 'Bone Battle'?" and "What is one thing you learned after doing this activity", were primarily used to understand whether students were engaged in the activity as well as whether they were able to extract the key learning objective from each activity.

Recruitment

We recruited our testing participants through a team member. The participants consisted of Gauri's brother and his two friends all age 12 and in the 7th grade. We realize that a potential bias of data collected from this group is selection bias as we did not choose the subjects randomly and the subjects have a personal connection with the tester. Another possible bias that could have occurred is groupthink bias as all participants were tested together. It is possible that this may have led to more generalized responses to our questions. In an ideal situation the team would have conducted testing with randomly chosen middle school participants. We would have tested participants one-on-one in a more standardized setting like in their classrooms/schools. In the analysis of our data we do acknowledge that more extensive testing with middle-schoolers would need to be done to confirm our conclusions. The Field Museum has confirmed to us that in the future of this project they would be able to conduct this testing with students themselves. We use the conclusions of the data we collected to guide the further development of the digital prototypes before they are handed off to The Field Museum.
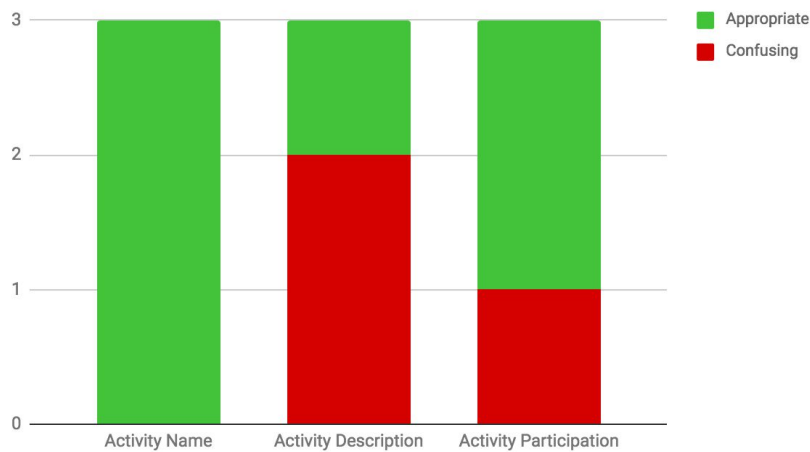
Analysis

Demographics

Our sample consisted of three 12-year-old participants all in the 7th grade. All participants had some previous general knowledge of both dinosaurs and turkeys. They learned about dinosaurs mostly from TV and they learned about turkeys from a combination of school and real life. They did not have such extensive knowledge of either as to not benefit from a product like Dino's Alive. All participants said they

enjoyed hands-on learning in the classroom and would be more open to lesson plans that included it. None were particularly interested in science or history in school but thought dinosaurs would be a "cool" subject to learn about.
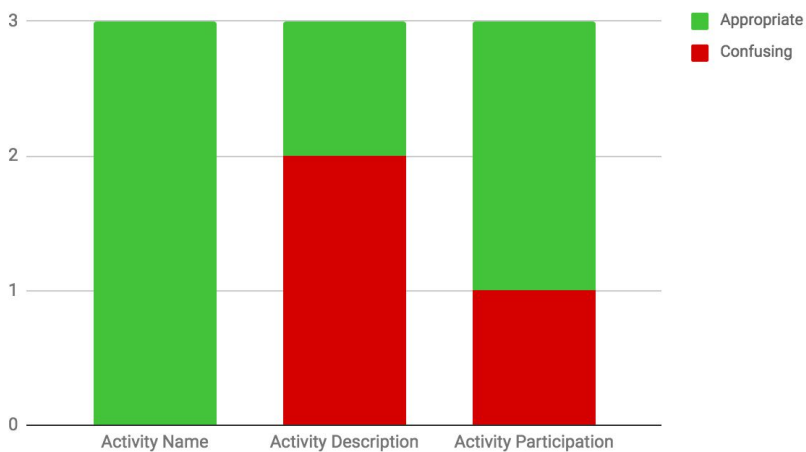
Statistics

The goal of our product is to create the kind of learning experience that kids like our participants would be interested in. Comparative tests like A/B testing were not suitable for our tests. In addition, large scale surveys would not work as well since our primary audience is children in elementary school and grade school. Our user testing was focused on interviewing three kids and having them run through our digital prototype. The following graphs summarize the students' perceptions about the title of each activity, the descriptions of each activity, as well as overall activity engagement.

Mentioned above, our testing consisted of interviewing three participants about our digital prototype. Therefore, our testing had no quantitative data to conduct any statistical analysis.

Insights

As stated before, we found that the names we used for the activities seemed to mislead the participants we tested with. They thought that there would be action and fighting in the activities, whereas they were more comparison activities. That said, they had less of an issue with "Bone Battle" and were more confused with "Function Fight." Specifically, participants thought that Bone Battle meant "specific dinosaur bones battling" and "two animals battling each other", and Function Fight meant "different features of dinosaurs fighting" and "robots or making functions to fight".

Overall, however, the participants seemed generally interested in our activities based on the descriptions and wanted to learn more, and they also found the navigation to be very intuitive, however one student mentioned that "[they] would want more specific instructions on what you to do and look for" in the environment. While participating in the activity itself, students rated both activities at between 8-9 on a scale of 1-10 from extremely boring to very engaging. However, two students mentioned that they wanted the ability to zoom in and look around SUE's skeleton, specifically saying that they wanted to be able to "click on the bones, take some bones out to compare them and [they] really want[ed] to manipulate bones" and "add more objects in the second activity, as well as [be able to] zoom into the environment". We found their comments to be reassuring, as this will be a feature available in the final design through Sketchfab; it was just not a feature available in the prototype. Regarding the second activity, we also asked the students if they had any ideas as to the best ways to summarize the information in the activity and one student mentioned including a "summary quiz", which we will incorporate in the final design.

It was also useful testing with participants that are our target age group, as we were able to ask what they knew about dinosaurs and turkeys, in order for us to gauge what other toolkit users might know. The participants seemed to have general familiarity with these two creatures, but they were limited in their knowledge and were eager to learn more. We can use this to infer that future users will have similar understanding, and we might want to teach about things that are a little more in-depth but nothing too detailed.

Adjustments

From our analysis, we can pinpoint adjustments that should be implemented for the final design. The disconnect between the name and the activity illustrated that the name alone does not adequately describe the activity. The way we will alleviate this is by changing the name of "Function Fight" which is the name that caused the most problems. However, it is important to note that our prototype did not have a hover functionality that would have allowed the participants to see a description of the island. This does not take away from the importance of this change as the description is hidden at first glance and a misleading title would set the wrong expectations going into the activity whether participants saw a description or not. A possible alternative to "Function Fight" would be "Adaptation Attack" which emphasises the idea of adapting to an environment more clearly.

In addition, we saw that some participants wanted clearer instructions going into the activity. This is a difficult issue to fix as we do not want to over instruct and have no room for interpretation but we do see a need and value in a slightly more guided approach. Because SketchFab is primarily a model viewing platform, there are not many restrictions or guidelines we can implement directly into the environments which emphasizes the importance of our overall web interface having clear and understandable instructions. This design change would likely be implemented by The Field Museum, as they are handling the copy of the interface. We will ensure that our insights are clearly explained to The Field Museum team in order for this issue to be best alleviated.

In our prototype the summary page for each activity was fairly barebones which left a lot of room for improvement. While the participants thought that the summary design adequately explained the takeaways for the first activity, they felt as though the same design could not be translated to the second activity which covered more complex material. When we asked the participants for suggestions on how to summarize the material, one student mentioned including a "summary quiz." We had thought of this previously but did not incorporate it in the prototype, so we will implement it in our final design.