

2018

Bosch Annotate

design guidelines

Contents

<i>4</i>	First Time User Experience
<i>6</i>	Gatcha!
<i>10</i>	Diamonds
<i>14</i>	XP, Levels and Badges
<i>18</i>	Task Design
<i>20</i>	A Communication Platform

First Time User Experience

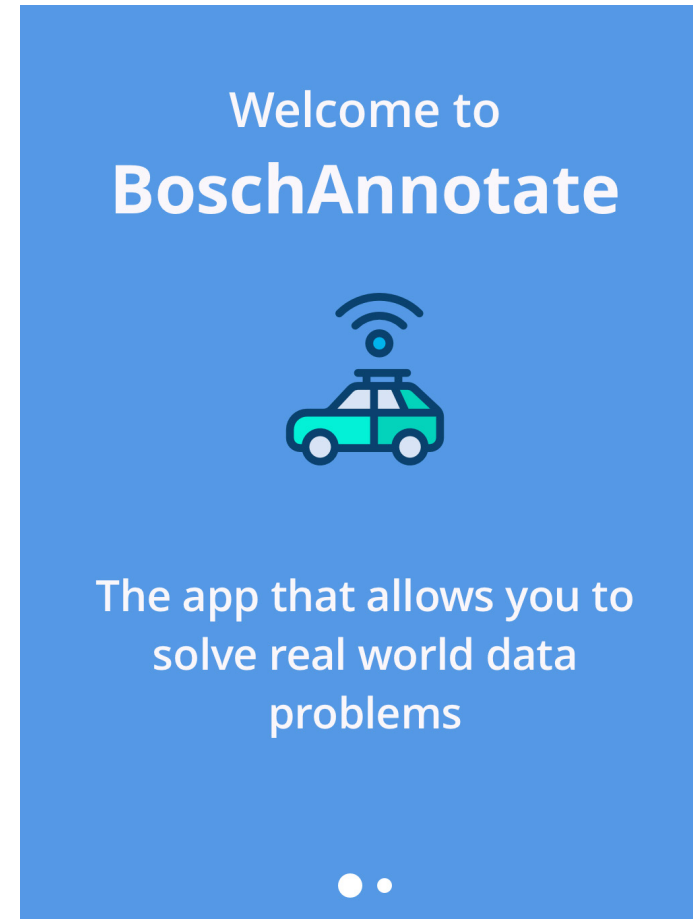
What is the purpose of this app?

Solving Real-World Data Problems

From our annotator survey and from the paper, "Experiments on Motivational Feedback for Crowdsourced Workers" by researchers at the University of Maryland and IBM we learned that providing background and information on why workers are performing the tasks they are about to do, can be a great way to motivate users.

We tackled this problem by designing two onboarding screens. The purpose of these screens was to give users a high level understanding of what the project was about and to indirectly coerce users into wanting to use the app, by making them feel proud about it.

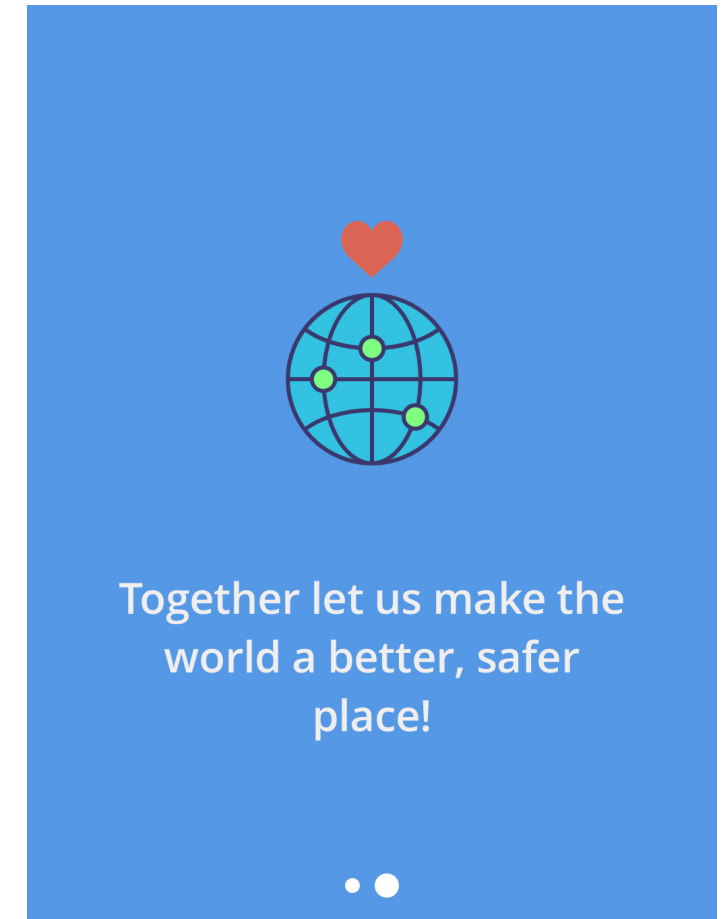
The first screen, was designed to make users feel like they are a critical part of this future Bosch envisions, and that the problems they would be solving are difficult, important and real.



Making the World a Better Place

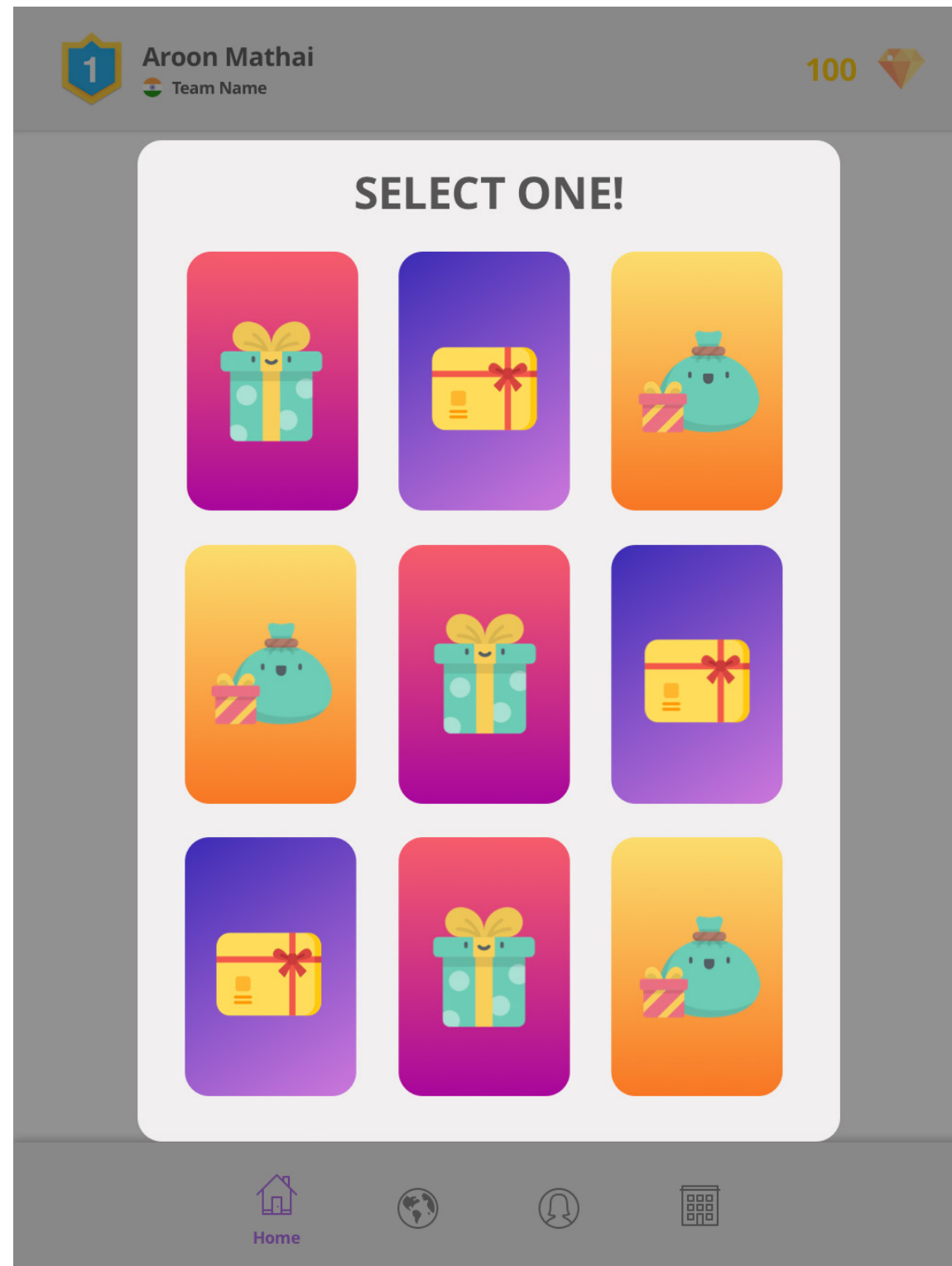
The second onboarding screen tries to tell users that by performing these tasks and by using the app, they would be driving humanity to a preferred state and in general, making the world a better place.

To summarize, the first screen tells users why Bosch needs them and why they should use the app and the second screen tells users what would be the result of them using the app.



1: Gatcha!

Slot Machine Mini-Feature to Drive Retention



Description

Gatcha Mini-Feature

Questions:

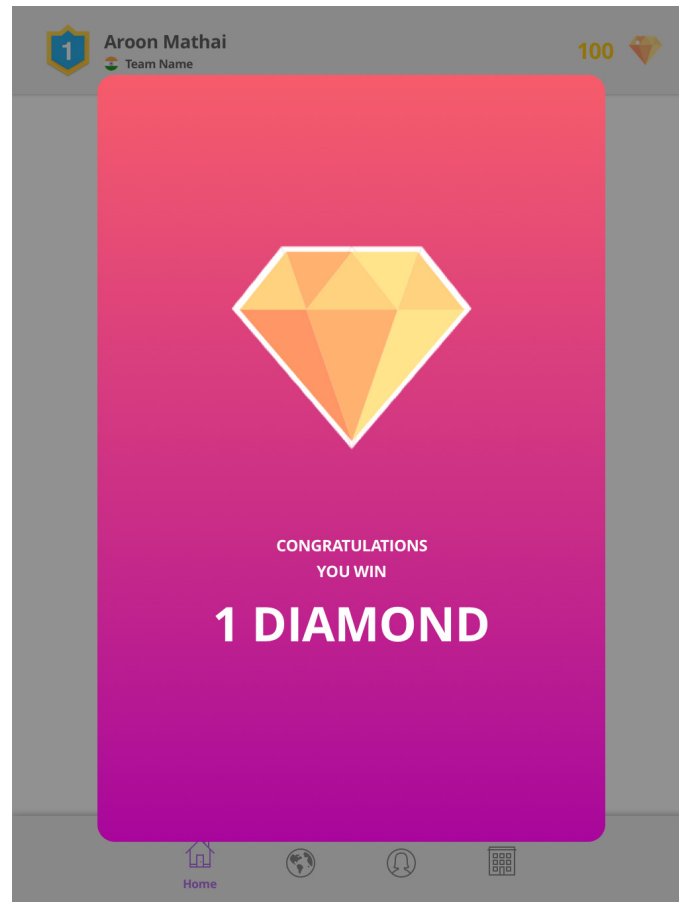
How do we drive retention?
Can we use variable reward schedules to create new customer behaviors?

Design summary:

From Nir Eyal's book 'Hooked' one factor that contributes towards user retention is offering them variable rewards using a variable reward schedule.

In this feature, users are given 9 cards and are allowed to select one, resulting in a surprise gift. Over time users will learn that these gifts range from less exciting and useful to more exciting and useful. Users see this feature once every 24 hours, motivating them to return to the app every day.

The key though is give users different rewards every day. In our prototype, we designed two categories of cards to solve this problem - A bad card and a good card. By strategically giving users different categories of cards each time they open the app, Bosch could maximise retention.



In our prototype we associated rewards to diamonds which can be converted into actual US dollars. More on this later, but every card can be thought of as a payment to users by Bosch, which is why it becomes essential to rotate between good and bad cards, instead of only supplying good cards.

This is directly in line with the variable reward “schedules” that B.F. Skinner researched and called a variable ratio schedule. It’s called “variable” because you don’t reward the behavior every time. You vary how often the person gets a reward when they do the target behavior. And it’s called “ratio” because you give a reward based on the number of times a person has done the behavior (rather than, for example, rewarding someone based on time - for example, giving a reward the first time the person does the behavior after 5 minutes has elapsed).

In a variable ratio schedule you may decide that you are going to reward the behavior, on average, every five times the person does the behavior, but you vary it, so sometimes you give the reward the third time they do the behavior, sometimes the seventh time, sometimes the second time, etc. It averages out to every five times.

In the beginning, Bosch would reward users every time they enter the app (In our prototype this would through a good card). This is called *continuous reinforcement*. Once the behavior is established, however, Bosch would then switch to only rewarding them every three or five or seven times on average.



These variable rewards have two prominent benefits:

They result in the most instances of the behavior than any of the other schedules, and

They result in behaviors that Skinner said were “hard to extinguish”, which means that the behavior persists over time, even when rewards aren’t being given any more (In our prototype this would mean users getting bad cards).

2: Diamonds

A virtual currency that can be traded for real world currency



Description

A virtual in-app currency that can be converted into real-world currency.

Questions:

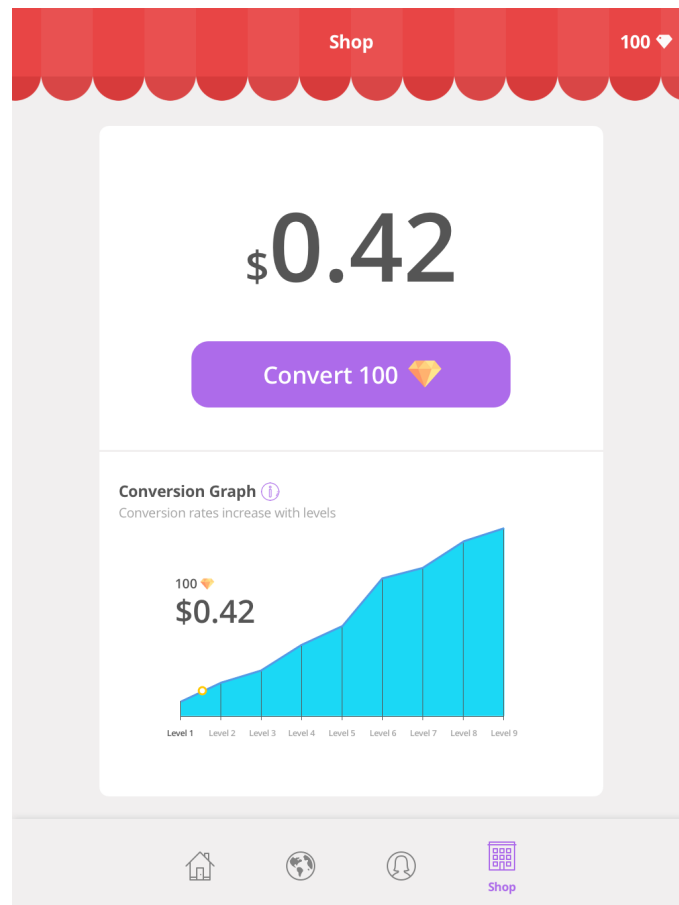
How can we get users to invest in the app so as to maximize retention?

Design summary:

Several studies have shown that expending effort on a task seems to commit us to it. For example, when buying a lottery ticket, players are able to either choose their own numbers or play a set of digits generated randomly. Certainly, choosing either option has no effect on the odds of winning. Traditional thinking would predict that the less effortful path would be the one users prefer.

However, the opposite is true. Despite the considerable effort required to pick the lottery numbers, players who choose their own numbers play more. This phenomenon isn't just about a skewed perception of luck. According to a classic study by Ellen Langer, even when players are explicitly told their chances of winning, they choose to trade worse odds for the ability to play the numbers they spent the time and effort picking.

After a user has been triggered into action and duly rewarded, the investment phase is where the user is asked to do work and starts building commitment. It is here that the user is prompted to put something of value back



into the system, typically in the form of time, money, physical effort, social capital, or personal data.

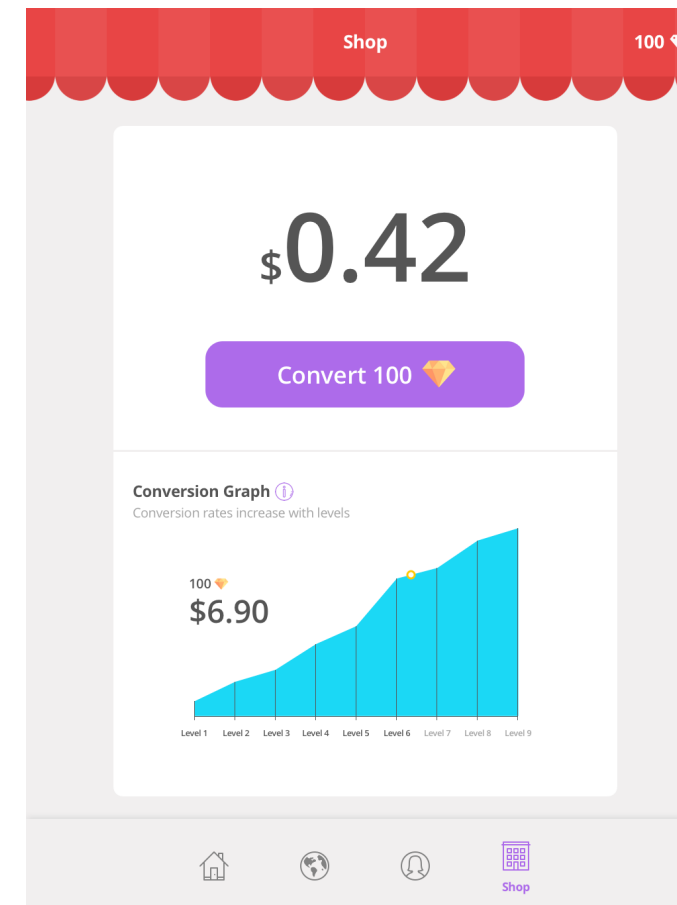
We try to do this in our prototype using diamonds. Users earn experience points (XP) on completing tasks. Once users earn a certain amount of XP they level up and each time they level up they earn diamonds.

To recapitulate, users can earn diamonds in two ways. One, through daily mini-games and two, by levelling up.

What makes diamonds so valuable is that they at any point can be converted into real-world currency. In the adjoining screenshot, a user can click the purple button to convert the 100 diamonds they have accumulated over time into \$0.42.

However, the goal is to keep users invested in the app for as long as possible. The more time users go without cashing in their diamonds, the more successful the app would be. For this reason, we designed the conversion graph.

The graph indicates that a user a little above level 1 can convert their 100 diamonds into \$0.42.

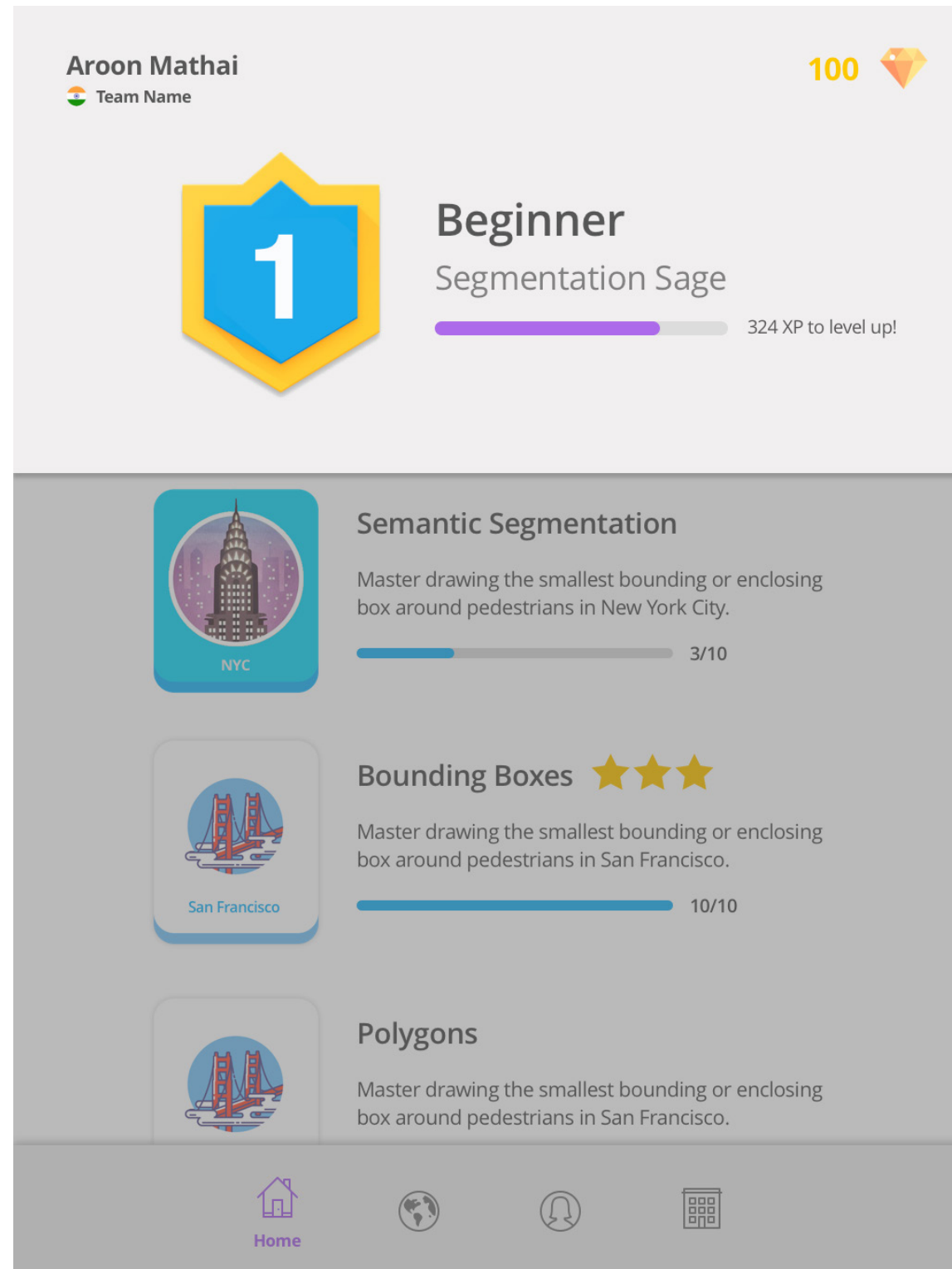


By interacting with the graph, a user can also learn that they can convert those same 100 diamonds for a lot more actual money at higher levels. In the prototype, the graph shows that at a little above level 6, a user can convert those same 100 diamonds for \$6.90 instead of a measly \$0.42.

This motivates users to invest these diamonds over time and subsequently to stay active in the application for as longer.

3: XP, Levels and Badges

An in-depth user flow



Description

A mechanic borrowed from popular social games to give feedback to users about the quality of their work.

Questions:

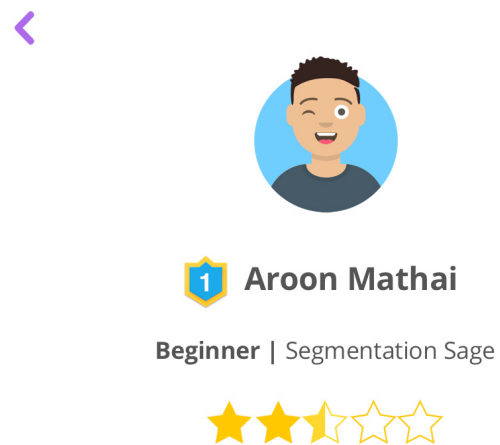
What is the most delightful way to provide feedback to users?

Design summary:

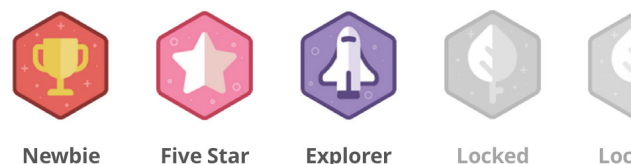
Progression is an important design element, which leads users to a feeling of mastery and accomplishment. It can be represented in different ways. Progress bars and levels are two of the most common ways of representing progression. But other mechanics can also be used like badges or achievements, unlocking new challenges, etc.

Progress tracking can be an important tool to help Bosch understand their annotators. It can help recognize what tasks the user has already completed and adjust the difficulty level of the next tasks accordingly to the current user skill. In an ideal situation the app would begin with small and easy tasks for new users increase them in size and difficulty at exactly the same rate as the user becomes more skilled.

As users complete more tasks, they earn more points and level up; that in turn allows them to work on more difficult tasks. These points that can be looked at as a measure of user progress are called experience points (XP) in our prototype.



Badges



XP is an estimate of the users skill within the application. It is important to note that XP and skill are not the same thing and don't always increase together.

XP is a more private form of progress tracking. In our app we use it as a form of feedback to users. Users earn between 0 and 5 stars for each task based on the quality of their work. The higher the quality per task, the higher the XP they earn. We also designed a more public form of progress tracking with levels and badges.

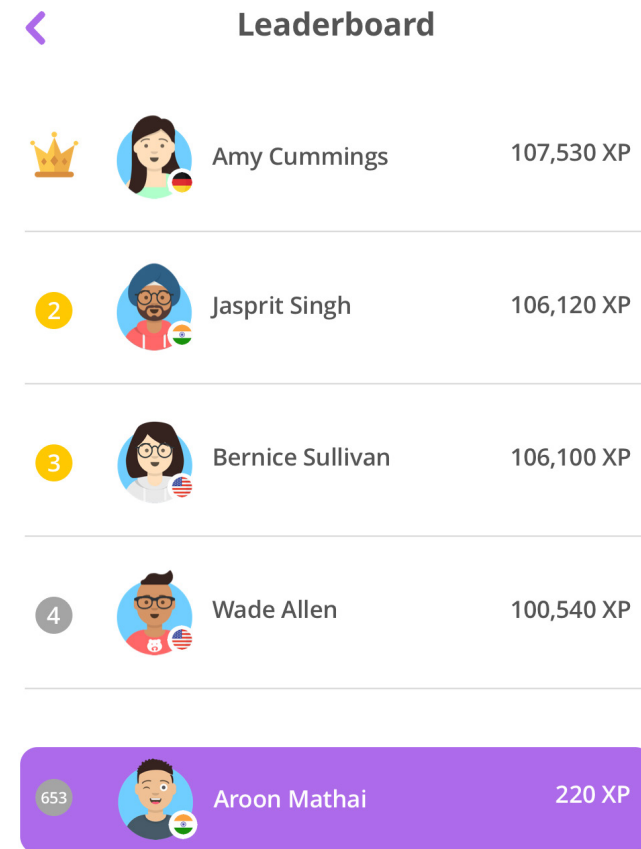
As users accumulate more XP they level up. Levelling up becomes more difficult the higher the level. For example, levelling up from level 1 to level 2 would require less XP than while levelling up from level 2 to level 3. Users' numeric levels are displayed on their public profiles so as to keep them accountable across the community.

Users also stand the chance to win visual badges which like their levels is also displayed on their public profiles. Here are a few badges we designed for the prototype:

Newbie: A badge users earn on completing their first task in the app

Five Star: A badge users earn on receiving 5 quality stars on any task for the first time.

Explorer: A badge users earn on trying more than one types of tasks.



Building upon the idea of keeping people accountable by keeping some information public to the entire community of annotators, we came up with a community leaderboard.

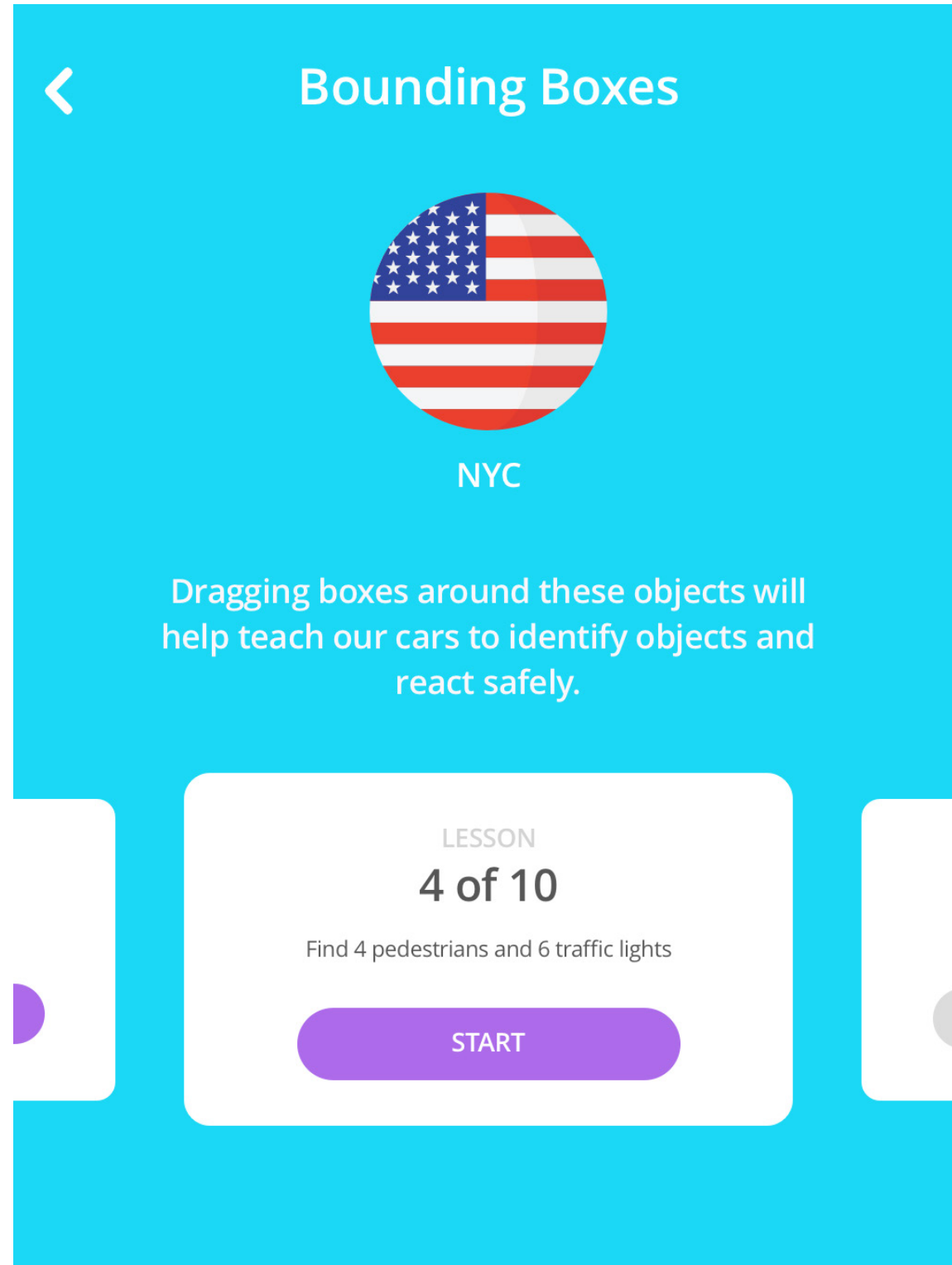
Apart from keeping users accountable, we hope to create a sense of competition amongst users. In the leaderboard we designed, users compete based on two factors:

One, based on XP. In the adjoining screenshot Amy has the highest XP and is on top of the leaderboard. Our user, Aroon who has managed to accumulate only 220 XP is ranked 653.

Two, based on nationality. Adding their national flag along with their profile picture is a cheap yet highly effective way to create a sense of competition amongst users.

4: Task Design

Designing tasks to reduce cognitive load



Description

Breaking tasks into bite sized sub-tasks to reduce cognitive load on annotators

Questions:

How can we reduce cognitive load on users?
How can we make tasks seem less boring?

Design summary:

Annotation tasks are generally perceived to be tedious and boring. One way for these tasks to be perceived as less boring is to break them down into different sub tasks.

For example, a more intensive bounding box or segmentation tasks can be followed by a less intensive tasks like a short quiz or maybe a task that asks users to check the accuracy of a task performed by one of their peers.

An additional advantage of such 'peer reviews' is an overall increase in quality in tasks across the community and quicker feedback to annotators. From our research we found that it is extremely critical to provide accurate and quick feedback to users, and designing tasks like this will help solve this problem.

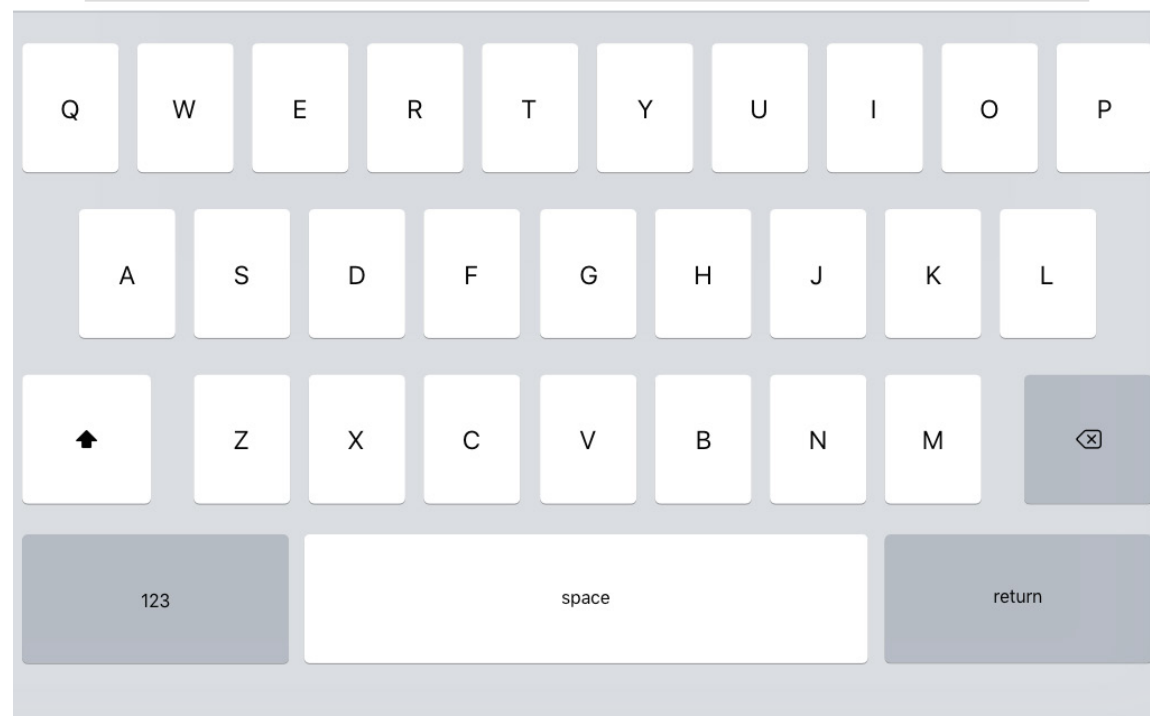
5: A Communication Platform

Promoting conversation between Annotators and Developers

Task Complete



Note to Developer



Description

Providing a platform to encourage conversation and collaboration between annotators and developers.

Questions:

How can we promote one-to-one feedback and collaboration between annotators and developers?

Design summary:

One important learning from our Annotator surveys was that a lot of times annotators complete tasks but are not sure whether they performed the task correctly or not. Often times, they misunderstand the prompt and perform tasks incorrectly.

These users are categorized as scammers unfairly. We designed a communication platform that annotators could use if they wanted to, to talk to developers on completing a task.

