



PROJECT 2: Rate your Lyft

Team: Drake Long, Jai Tudor, & Rhea Manocha

Due Date.....Feb 16, 2017

CGT 17208.....HCDD Learning Studio 1

Professor.....Dr. Gray

TABLE OF CONTENTS

Page	Page Number
Project Overview	3
Secondary Research	4-7
Market Analysis	6
User Demographics	7
User Research	8-9
Personas	10-12
Harper Jenson	11
Dr. Isaac McCray	12
Early Stages	13
Prototype / Testing	14-21
Initial Prototype	15-16
Test Protocol	17
Initial Test	18
Final Prototype	19-20
Final Test	21
Conclusion	22-24
Work Photos	25



PROJECT OVERVIEW

The Challenge

We were faced with the task of redesigning Lyft's current five-star rating system with a better alternative. There are many limitations to a five-star rating system, especially due to the ambiguity of the meaning of each star.

The Solution

Our final solution is a binary rating system that would allow the user to rate their driver during their ride. Then once the ride is over, the user will have the ability to tip leave an additional comment if desired. The final deliverable is a low-fidelity, paper prototype.

The Team

- Drake Long: Researcher / Documenter
- Jai Tudor: Sketcher / Prototyper
- Rhea Manocha: Usability Tester / Prototyper

SECONDARY RESEARCH

We conducted research to develop a more in-depth understanding of the problem we were working on.

How does the current system work?

We started out with looking at how Lyft's current rating system works. Upon the ending of the ride, the user is brought to a screen (Figure 1). The screen gives the user the option to leave a 1-5 star rating, an additional comment, and a tip. Additionally, the user sees the total amount of the ride and is prompted to submit everything.

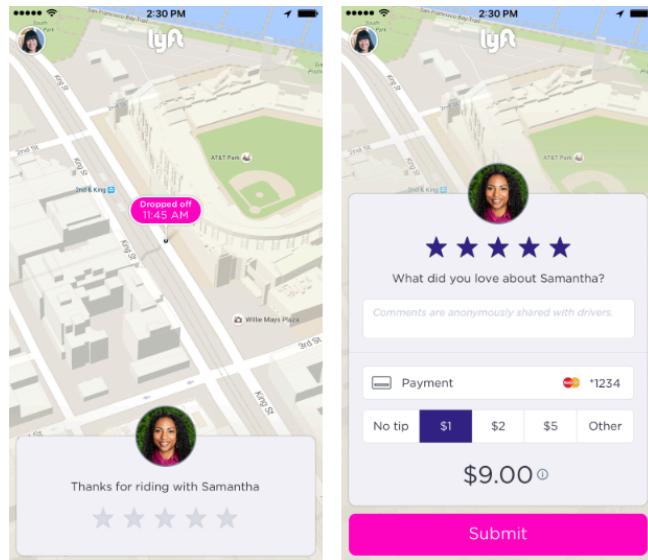


Figure 1

Source: <https://help.lyft.com>

Next, we researched the biggest pain points for the current rating system. We found that there is a disconnect from how users view the 5 star system and how Lyft corporate views it. Lyft drivers must maintain a 4.8 average score rating to keep to prevent being fired. This contradicts what most users think. In theory, Lyft expects drivers to receive mostly 5 star ratings and very few below five-star ratings.

According to David Celis, a software engineer at GitHub, a number of users think 1 star means they hated it, 2 stars means they didn't like it, 3 stars means they liked it, 4 stars mean they really liked it, and 5 stars means they loved it. However, to Lyft, 4 and 5 should represent okay or great. Essentially, if a driver receives only 4-star ratings, they will be fired. However, the average user would think a 4-star rating is good. Celis states, "It is easy to end up with two people who give an item the same three-star rating but actually feel differently about it." Essentially, this means that five-star rating scales are too ambiguous and are fundamentally flawed.

This led us to start thinking about the research behind five-star rating systems and how it compares to binary rating systems.



Five-Star Rating System

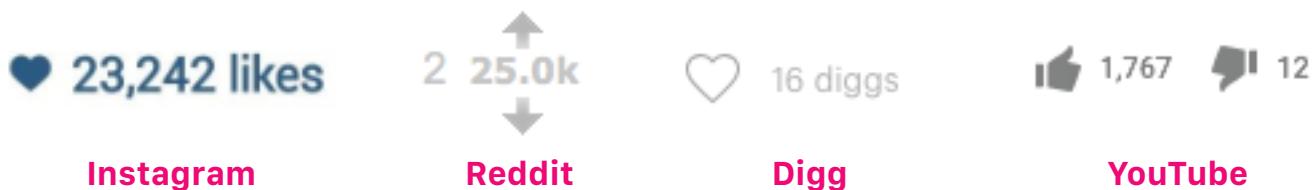


Binary Rating System

Research from Anna E. Bargagliotti of University of Memphis and Lingfang (Ivy) Li of Shanghai University of Finance and Economics suggest that scale systems are subject to human error due to each individuals interpretation of what each notch on the scale means. On the other hand, binary rating systems are less subject to human error because it is easier to understand and is very uniform.

Market Analysis

We know a number of websites use the five-star rating scale, so we looked for websites that use a binary scale. We looked at social media, Instagram currently uses a binary scale with only having the option to like or not like. Reddit and Digg also have a similar system with up votes or down votes and digging or burying. Additionally, we noted that YouTube moved from a five-star system to a binary (thumbs up or thumbs down) system in 2009. According to Margaret Gould Stewart, former YouTube employee, most users were either rating videos 1 star or 5 stars and rarely using the in-between stars. Therefore, the transition to a binary system felt natural. From all of the websites we found, binary data was easier to understand and was very straight-forward.

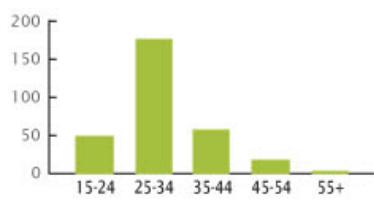


User Demographics

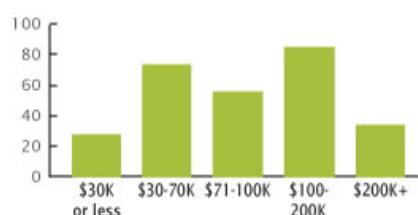
From an academic paper from Lisa Rayle, doctoral students at the University of California Transportation Center, and Vugo rideshare advertising, we were able to see Lyft and Uber's average demographic information.



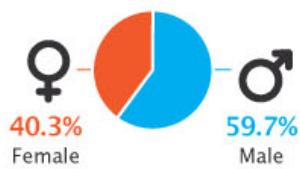
Age⁽¹⁾



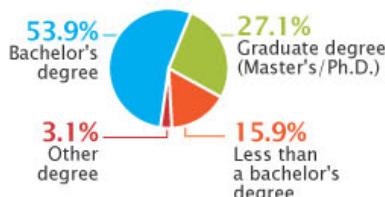
Household Income⁽¹⁾



Gender⁽¹⁾



Education⁽¹⁾



Source: <http://govugo.com/rideshare-passenger-demographics/>

Top 10 Destinations

1. House
2. Work
3. Tourists Destinations
4. Restaurants/Bars
5. Airport
6. Retail Stores
7. Gym
8. Doctor/Hospital
9. Transportation Station
10. Hotels

At this point, we felt we had sufficient secondary research and wanted to conduct our own primary research.

USER RESEARCH

We had conversations with six ride-share app users. The individuals we interviewed, 2 female and 4 male, were all within the 18-25 age range. We wanted to learn about their rideshare experiences, what they do during the ride, and their rating habits.

“

I look at my GPS during the ride all the time. It makes me feel safer, knowing the driver is going the right direction. Also, it lets me know how much longer until my destination.

~ Female | 23 | Professional

”

“

I always forget to rate. I just close the app and go on about my business.

~ Male | 22 | Professional

”

Interview Protocol

Each interview was less of a questionnaire and more of a conversation. We wanted to empathize with the users and really get a feel for each individual's personal experience with the app. Therefore, every user was asked questions tailored to their use habits.

Key Insights (based upon our interviews)

1. During the ride, users mostly stay on their phone; either on social media or watching the Lyft GPS.
2. Users usually leave 5 stars and no comment because it's easier than actually thinking about the ride and fast.
3. Users sometimes forget to rate due to the nature of it being an afterthought.

Connections between user research and secondary research

1. Users usually leave only 1 star or 5 stars, and rarely use the in between stars. This is very similar to the issue Youtube had.
Users are providing binary data to a non-binary system.
2. Our interviewee's typical destinations were restaurants or home.

Now, we felt that we had collected good user data and stories.

From here, we began to create personas that felt realistic.

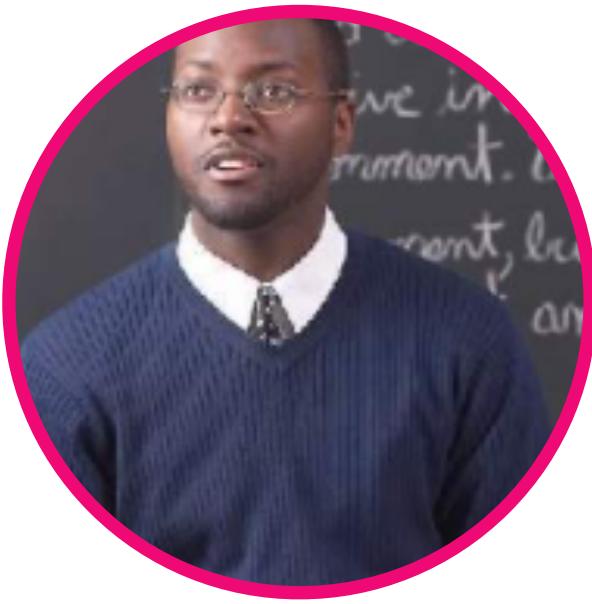
PERSONAS

We developed personas based upon our research, both primary and secondary. We believe the two personas embody our users and we want to design our solution based upon these fictional characters.



Harper Jenson
Female | 27 | Professional | Seattle, WA

Image 1 - Source: <http://www.forbes.com/sites/learnvest/2017/01/31/signing-bonuses-flex-time-and-more-5-work-perks-you-should-negotiate-in-your-next-job-offer/&refURL>



Dr. Isaac McCray
Male | 34 | Professor | Bloomington, IN

Image 2 - Source: <http://www.thenewblackmagazine.com/view.aspx?index=2905>



Harper Jenson

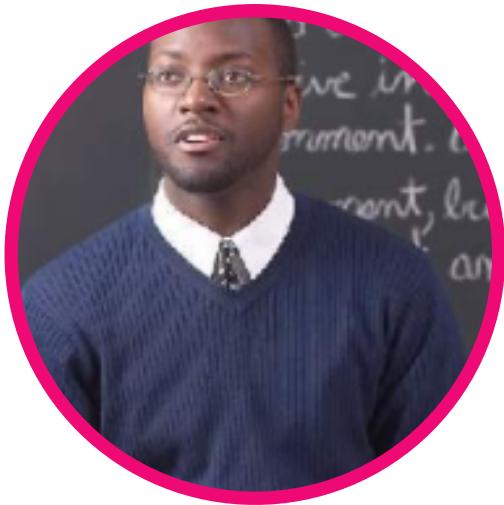
Female | 27 | Seattle, WA

Background

Harper lives in the busy city of Seattle, Washington, where she doesn't have a car because parking is too expensive downtown. She usually walks to her job and other places that are close to her apartment. She works for Nordstrom corporate in their buying office and works with vendors to fill the men's shoe department. Harper is single but is in the dating and nightlife realm. She has been on a few dates with a girl named Hope, but it isn't serious yet. Additionally, she likes to go out to happy hours with her friends on Thursdays at a swanky place in Capitol Hill, another neighborhood in Seattle.

Pain Points

- Sometimes forgets to rate when in a hurry.
- When she does rate, she just gives five stars because it's quicker than actually thinking about the quality of the ride, and she can close the app.
- Never leaves additional comments because it takes too long.



Dr. Isaac McCray

34 | Professor | Bloomington, IN

Background

Dr. McCray is a professor at IU Bloomington, where he teaches linguistics and foreign languages, specializing in Mandarin. He lives about 15 minutes driving from campus. Though he has a car, he enjoys riding his bike to campus when the weather is pleasant. He often uses Lyft when he rides his bike to campus and doesn't want to ride back at nighttime. During the weekdays, he enjoys cooking with his wife Shelda, playing with his four-year old son, Dylan, and walking his golden retriever, Buddy. His hobbies include collecting international currency and repairing the 1974 Dodge Charger he inherited from his father. He enjoys going out to try new local cuisines.

Motivations

- Uses Lyft when going out for drinks or dinner with wife to prevent drunk driving
- When he bikes to work, sometimes will take Lyft back instead of biking at night
- Is currently trying Lyft, which is new to his college town, while Uber has been there for years
- Prefers Lyft because "app feels friendlier"

Pain Points

- Wants to tip, but is saving money to move to bigger house, so feels guilty when he can't tip well.
- When he has had drinks and orders a Lyft, wants quick transaction with as little friction as possible.
- Doesn't feel like 5 star rating accurately depicts ride because he always gives 5 stars, but varies in quality of ride (simple and quick vs. many amenities), but thinks comments take too long.

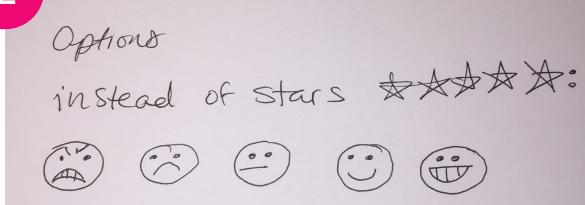
EARLY STAGES

These sketches were created before extensive research. We wanted to get ideas on paper and go from there.

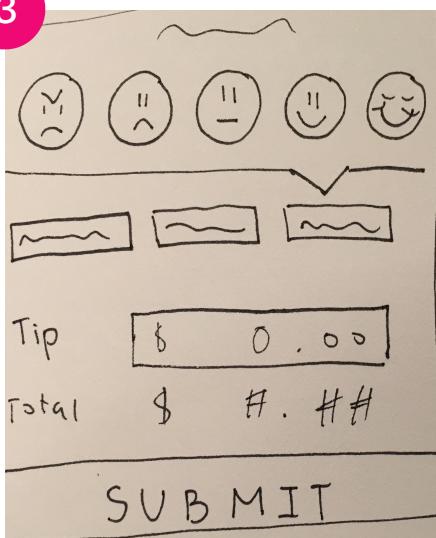
1



2



3



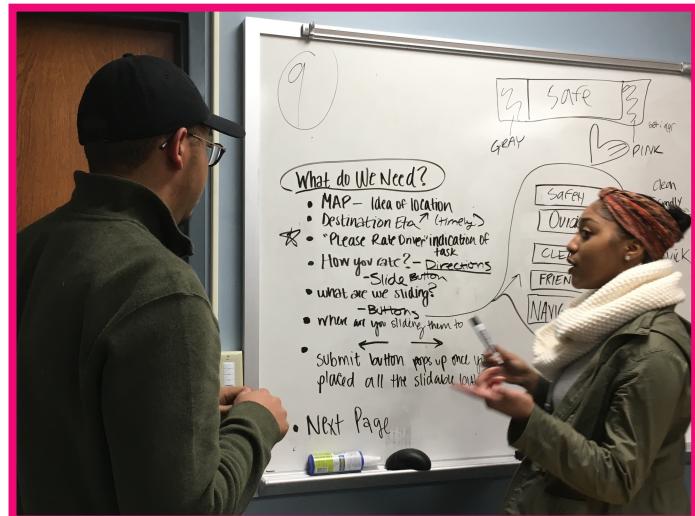
We originally started by thinking about adding categories to the existing 5 star rating system. However, as we continued to research we found that we wanted to move away from a 5 star scale. So, we looked into emoticons instead of stars. We felt that this could convey how the rider felt better than stars could. However, through research, we found that faces are highly subjective and can be interpreted differently by various people. We then sketched a prototype implementing the faces with the categories. Though we decided against the faces, we wanted to explore the option and see what this option would look like on paper.

PROTOTYPING/TESTING

After our research, we revisited our early sketches. We knew we wanted to keep the category aspect, because our research found that users rarely leave additional comments. However, we wanted to eliminate the scale rating entirely, due to the ambiguous nature it provides. Additionally, we wanted to implement the rating system into the ride, instead of leaving it at the end. This will, in theory, prevent users from forgetting to rate once the ride is over.

Whiteboarding

We took to a whiteboard to figure out what are the essential items needed in the rating page. We decided that we needed the user to be able to view the GPS during the ride. Additionally, the user needs to be able to see their ETA. With that in mind, we started making our initial prototype.

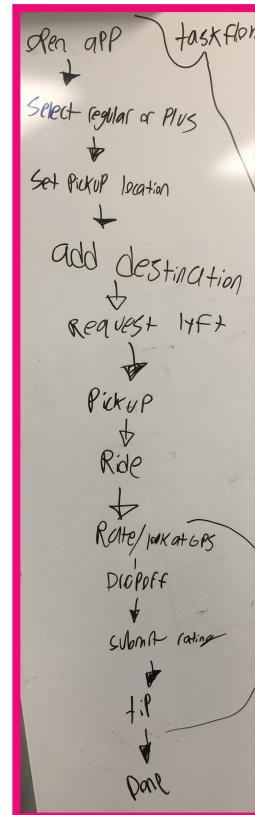


Drake Long

Jai Tudor

User Task Flow

Additionally, we created a user task flow to look at the Lyft experience holistically and determine where our solution would fit into the process. Because we decided to have users rate during the ride, we put our solution after the ride starts and before the ride ends.



Initial Prototype

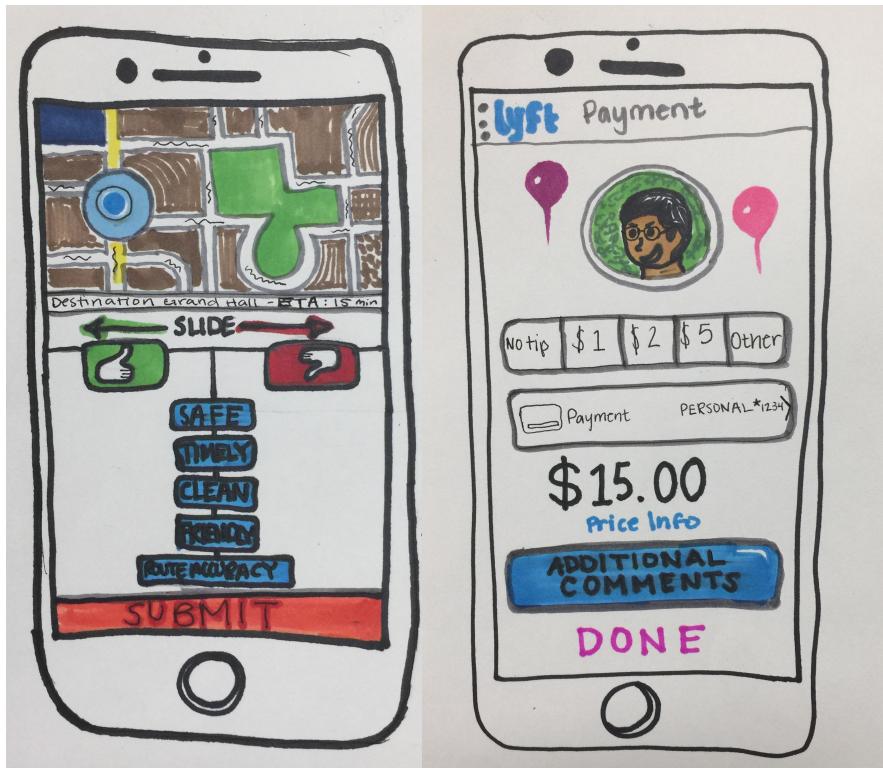


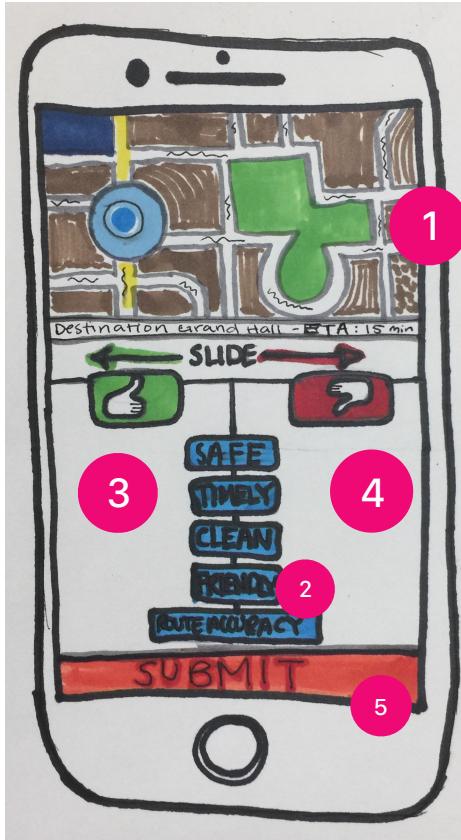
Figure 1

In-ride/rate screen

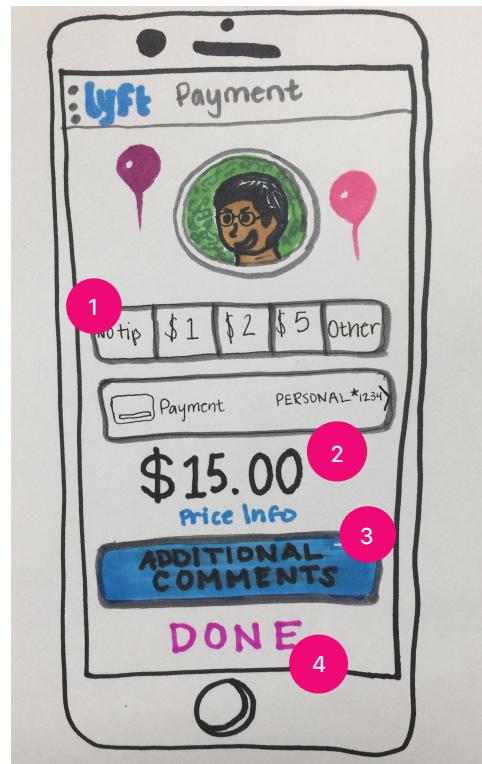
Figure 2

Post-ride/tip screen

We sketched a screen where the user could rate their driver during the ride (figure 1), and still see the essential things they want to see during the ride. The user will not be able to submit the rating until the ride is over. Then once submit is clicked, it will bring them to the ride summary page (figure 2). Here they will be able to tip and leave a comment if they desire to do so. Categories chosen were based on information provided by lyft.com. Lyft calls them the 4 flags, which are navigation, safety, cleanliness, and friendliness. Additionally, we included the category "quick" because our research demonstrated that timeliness was important to rideshare users.



In-ride/rate screen



Post-ride/tip screen

Figure 1

1:

This is the GPS info the user wants to see during the ride. This includes the navigation route, ETA, and destination information.

2:

These are the categories that users use to rate their driver. These can be slid into good or bad on the binary rating system.

3 & 4:

This is the actual binary rating system. If a user slides a category to the left, that is good. If a user chooses to slide a category to the right, that is bad. Additionally, if the user chooses not slide the category, this doesn't effect the driver one way or the other.

5:

This is the submit button. This allows the user to process to the next page. This button will only be active once the ride is over.

Figure 2

1:

This is the tip selector. Users can choose to tip their driver whatever desired amount, or not at all.

2:

This is final price the user will be charged after the ride.

3:

This is an additional comments box that allows the user to leave a comment if they feel they have more to say that couldn't be expressed through the other rating categories.

4:

This is the done button. This is the last step of the rating/tip process. The will be pressed when the ride is completely over and the user leaves the car.

From here, we decided to test this prototype and receive feedback on its usability.

Usability Testing Protocol

Ethics:

Our test began with a testing introduction statement: "Hi, my name is ____ and we are testing our prototype for a redesigned rating system for Lyft. I want to make clear right away is that we're testing the application, not you. There are no wrong answers. Feel free to ask questions as they come about. We will give you a task, and we want to see how you complete it."

Scenario:

You have ordered your Lyft, and have been picked up by your driver. You are in the car, riding to Kroger. Your driver is in a nice Toyota Camry, he is friendly but doesn't talk very much. He drove slowly, followed the GPS, and was very safe.

Tasks:

There were a few tasks given. **1.** Can you please rate your driver during your ride? **2.** Can you proceed to the next page? **3.** Can you tip your driver? **4.** Can you leave an additional comment? **5.** Can you submit your ride?

Questions:

All testers will be asked to complete the same tasks. However, additional questions will be asked depending on the specific person and their interaction with our prototype.

Initial Usability Test

Results:

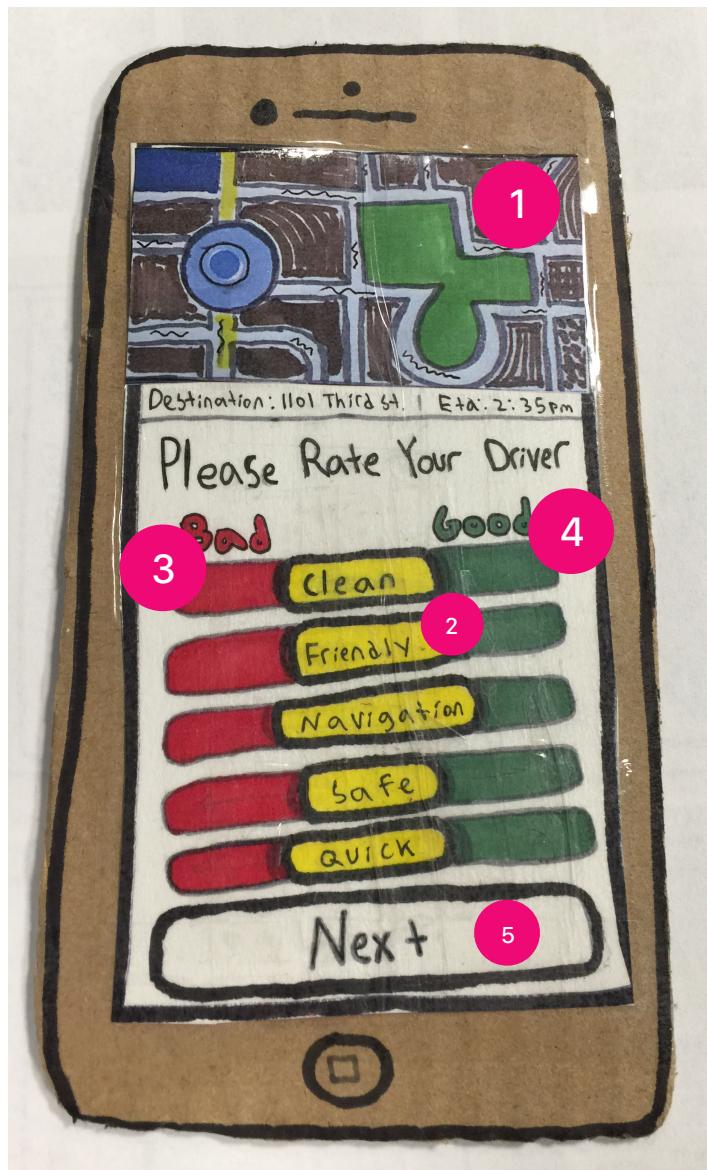
We tested 4 individuals. The individuals we interviewed, 2 female and 2 male, were all within the 18-20 age range. We followed the usability test protocol and asked them to complete the tasks.

1. Can you please rate your driver during your ride?
 - All of the users realized that the categories were supposed to move into good or bad. However, 3 out of the 4 users tried to use them as a drag and drop function. This process took all of them a bit of time to realize what to do with the categories.
2. Can you proceed to the next page?
 - All users were able to click the submit button in a timely manner.
3. Can you tip your driver?
 - All users were able to complete this task without any issues. 2 out of the 4 did not choose to leave a tip. However, that does not affect their ability to complete the task.
4. Can you leave an additional comment?
 - All users were able to click the additional comments box. Although 1 male tester left a comment.
5. Can you submit your ride?
 - All users were able to click the done button in a timely manner.

Results Analysis:

We found that most of the issues came from the first screen. 3 out of the 4 testers ignored the slide functionality and attempted to drag and drop. We took this to mean we need to improve the design to communicate the slide functionality function better.

Final Prototype



In-ride/rate screen

1:

This is the GPS info the user wants to see during the ride. This includes the navigation route, ETA, and destination information. This remain unchanged from the initial prototype.

2:

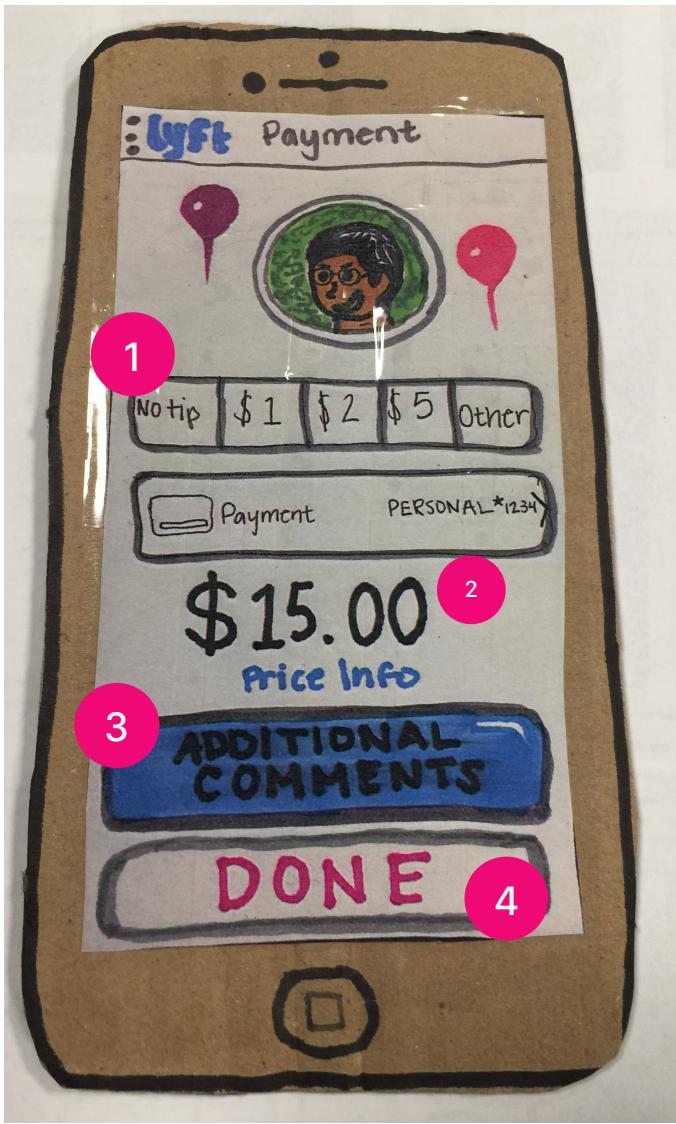
These are the categories that users use to rate their driver. These can be slid into good or bad on the binary rating system. The categories were reorganized in the order that users would encounter them. First riders see if the car is clean, then if the driver is friendly, then if the driver follow the directions. Next, the rider will see if the driver is driving safe, and when you arrive at your destination, did you get there on time.

3 & 4:

This is the actual binary rating system. If a user slides a category to the left, that is good. If a user chooses to slide a category to the right, that is bad. Additionally, if the user chooses not slide the category, this doesn't effect the driver one way or the other. We changed the order and put bad on the left, and good on the right. We found that this is a more natural function, as this is how other binary apps work such as Tinder.

5:

This is the next button. This allows the user to process to the next page. This button will only be active once the ride is over. We changed the word from submit to next to indicate that there is another page after this. This way users won't feel as if they are done after this page.



Post-ride/tip screen

This page of the prototype remained unchanged from the initial prototype. The initial usability test showed that users are able to complete all tasks related to this page efficiently.

1:

This is the tip selector. Users can choose to tip their driver whatever desired amount, or not at all.

2:

This is final price the user will be charged after the ride.

3:

This is an additional comments box that allows the user to leave a comment if they feel they have more to say that couldn't be expressed through the other rating categories.

4:

This is the done button. This is the last step of the rating/tip process. The will be pressed when the ride is completely over and the user leaves the car.

Final Usability Test

Results:

After conducting five usability tests, 2 female and 3 male, we found that users were able to easily navigate through the app without assistance. However, some users were confused or did not comprehend that the rating was during the ride and not only after the ride was over. Due to the limitation of usability testing, we could not simulate the full in-ride experience. If we were to conduct further testing, we would create scenarios that gave examples that emphasized the during aspect of the rating system, because we would like to test the temporal aspect that we manipulated.

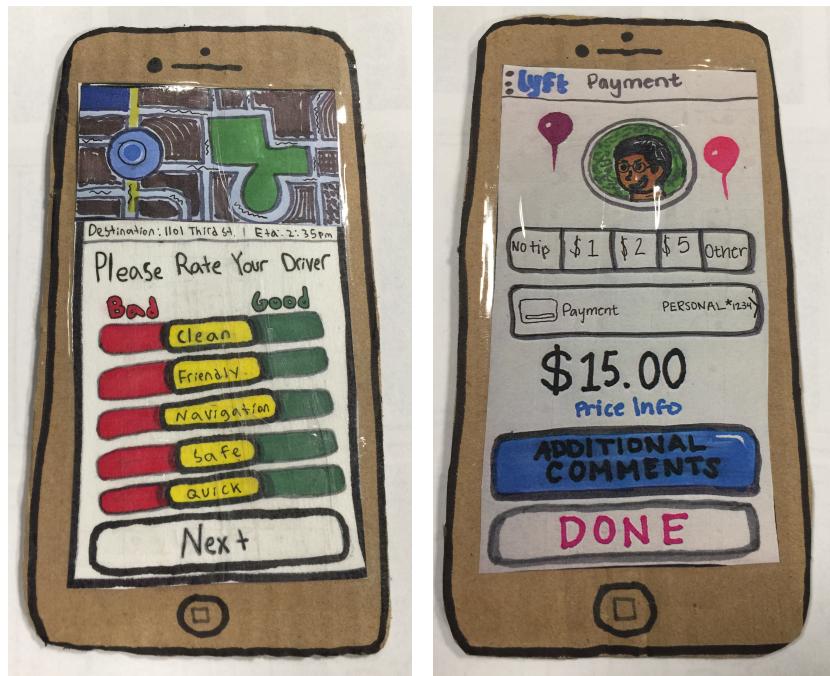
1. Can you please rate your driver during your ride?
 - All of the users realized that the categories were supposed to move into good or bad. Additionally, all users realized the buttons were sliders, and slid them left or right. They were able to complete the task efficiently.
2. Can you proceed to the next page?
 - All users were able to click the submit button in a timely manner.
3. Can you tip your driver?
 - All users were able to complete this task without any issues. Only 1 user chose to leave a tip. However, that does not affect their ability to complete the task.
4. Can you leave an additional comment?
 - All users were able to click the additional comments box. Although 1 female user commented, the rest of the tester said that they never leave a comment.
5. Can you submit your ride?
 - All users were able to click the done button in a timely manner.

Results Analysis:

The usability test results were exactly what we were looking for. All of our testers were completely able to complete all of the tasks efficiently. We were now ready to move on from the testing stages.

CONCLUSION

We were very satisfied with our final usability test results, we decided that would be our final solution. Due to the low-fidelity nature of this project, it will remain a paper prototype.



Rationale

The goal of the case study was to redesign Lyft's current 5-star rating system to allow users to be more specific in their ratings but in a timely manner. Throughout our process, we kept in mind that speed in interaction was the most essential thing. Our design set out to solve three issues we found in the current rating system: **accuracy**, **ambiguity**, and

forgetfulness. These are the 3 problems we found during our research. So, when we created personas, we decided to make these problems pain points for them.

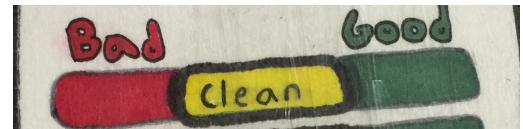
Accuracy: The categories that factor into the drivers rating.

In Lyft's current system, there are no categories to specifically rate the driver on. It is dependent upon the user what matter most and what rating they give overall. We attempted to solve this by introducing categories into the rating page. We have the following categories: clean, friendly, navigation, safe, and quick. This way, when drivers get their feedback at the end of every month, they can see their performance in each category. For corporate measures, 5 categories works out better. Therefore, each good rating adds a star, each bad doesn't add a star, and each neutral adds half a star. This way Lyft corporate can still give drivers an average overall rating.



Ambiguity: The reason for the binary rating system.

We chose to go with a binary rating system to eliminate the ambiguity of the five-star rating system. Our research found that different people view the same amount of stars differently. For example, two people can see 4 stars completely different than the other. Therefore, with a binary system, it is very uniform. The concept of good and bad is very standard and understandable. We hope to eliminate any ambiguity in driver ratings and create a seamless understanding among all users.



Forgetfulness: The act of users forgetting to rate their driver.

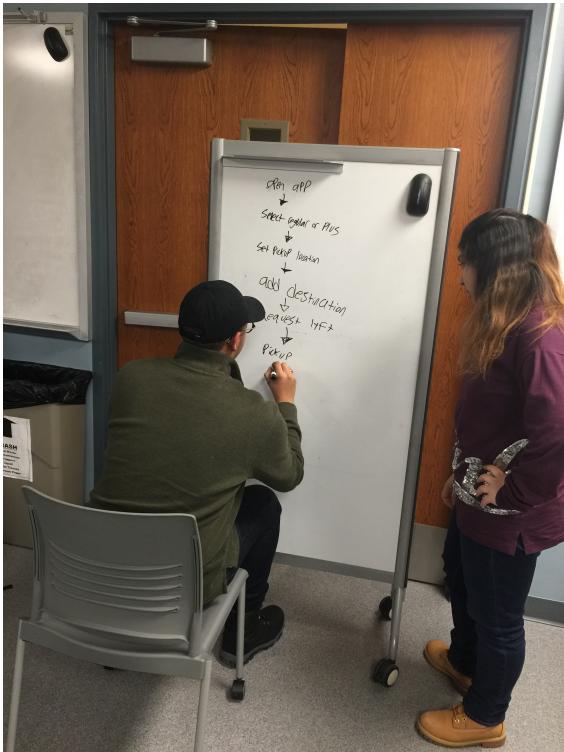
Through our research, we found that many users often forget to rate drivers, especially when they are in a hurry to go to their destination or compromise leaving an accurate rating in favor of quickly submitting their standard rating. We created a journey map to analyze how users interact with the app and the driver. This led us to change the rating system so users could rate the driver during the ride in real time, and change their rating according to how the ride went. Additionally, users don't have to spend time thinking and remembering the ride at the end. This prevents stress and forgetting to rate at the end of the ride, as the users we tested tend to already look at the GPS in the app during the ride.



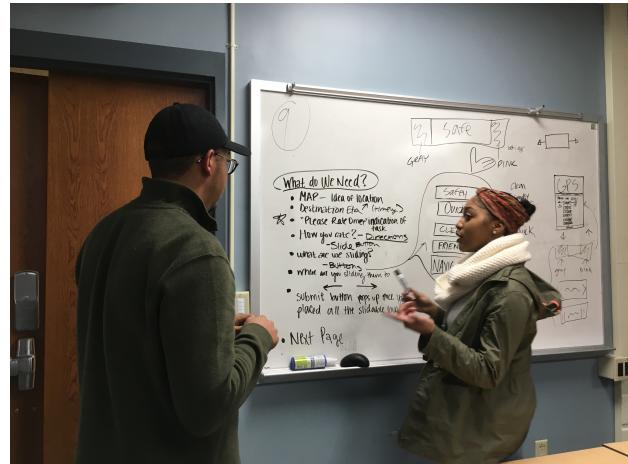
THE END



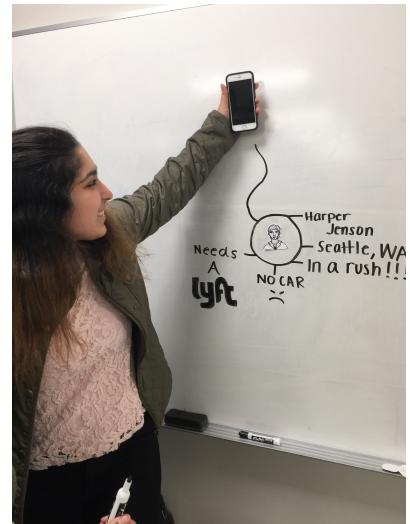
WORK PHOTOS



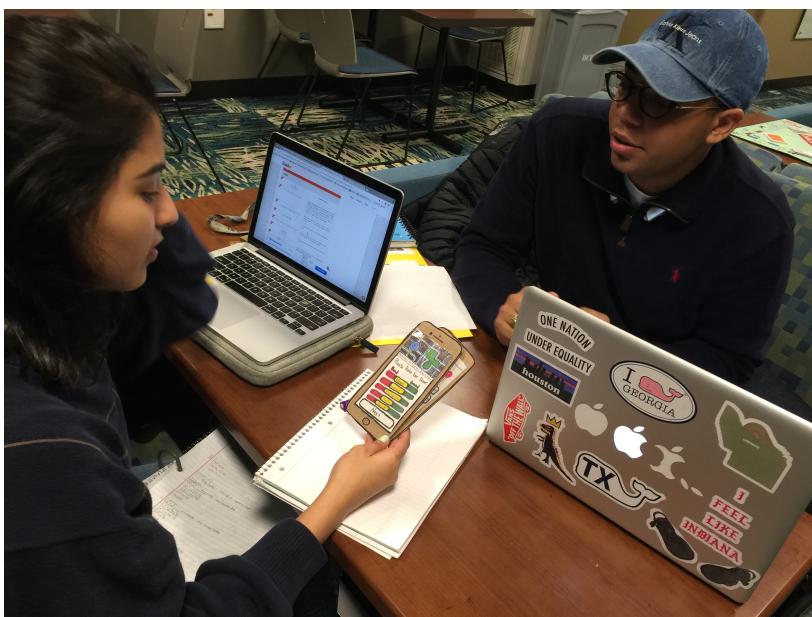
User Task Flow



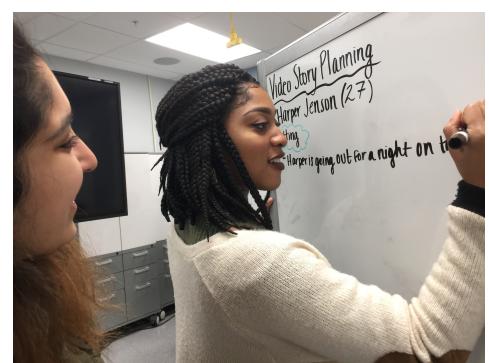
Ideation Process



BTS Pitch Video Shoot



Usability Testing



Pitch Video Planning