

# Human Acceptance of Unpredictable Behaviors in Fully Driverless Ride-Hailing Services

Disha Das  
Georgia Institute of Technology  
Atlanta, USA  
ddas71@gatech.edu

Tarushree Gandhi  
Georgia Institute of Technology  
Atlanta, USA  
tgandhi9@gatech.edu

**Abstract**—There has been a growing research on the trust of humans in Autonomous Vehicles, where various surveys have been conducted to evaluate the experience, likeability and trust of humans on self-driving cars. Most of these work focus on examining different factors that determine human trust and reliability on autonomous cars e.g frequency of use, knowledge, ease of learning etc. There also exist studies evaluating ways to help improve human trust on AVs. However, there has been no research so far on comparing acceptance of errors/unexpected behavior of human-driven cars versus autonomous cars. In this work, we aim to determine the gap between leniency towards human drivers and autonomous cars by recording human responses towards mistakes made by both the agents. From this study, we infer that humans are more forgiving and accepting of mistakes made by autonomous cars as compared to that by human drivers.

**Index Terms**—Autonomous Cars, Trust, Taxi, Ride-hailing service, Human Acceptance

## I. INTRODUCTION

Modern life demands efficient use of time as a resource. It is a common instance that a person hailing a taxi is looking to get to their destination reasonably faster than, say, buses or other means of public transport. Also, safety is a major concern. Under such critical constraints, will humans be forgiving and accepting of the numerous possible mistakes that an autonomous car can make?

In this study, we seek to understand surveyor's attitudes towards Robotaxis when pitted against a similar performance by human-driven taxis. Given the flawed performance of a Robotaxis and an equally flawed performance by human-driven taxi, we seek to understand which service a rider is inclined to take and why.

## II. RELATED WORK

In 2018, [3] Waymo launched a self-driving taxi service called "Waymo One" in Metro Phoenix area of Arizona. While the cars were capable enough on the roads, each of them had a human attendant. Two years later, the company has kicked off a fully-driverless taxi service in the same area, this time, open to public [11]. Tesla announced its own fleet of Robotaxis to hit the market in the near future [2]. Likewise, a slew of other AV companies have made the same claim.

Based on Waymo One's feedback reviews [14], only 70% of customers were perfectly satisfied with their rides. Shakiness, excessive braking, weird drop-offs, etc. were among the many causes of dissatisfaction.

[10] described the factors leading to the adoption of driverless cars in closed environments. A survey based approach was used to evaluate the Trust and Adoption of driverless technology by the participants. The key factors influencing driverless car adoption were identified to be performance expectancy, reliability, security, privacy and trust.

Our study is different from the prior work in the sense that we want to uncover the underlying human psychology relating to acceptance of driverless technology and identify any biases for or against the same.

## III. APPROACH

### A. Method

A total of 10 scenarios were created. The circumstance described in each video has been observed or reported by passengers of famous Level 4 and Level 5 company autonomous vehicles. These circumstances were replicated in Unity 3D by taking cues from such reports and Youtube videos. Participants were made into two groups.

- **Group I:** This group was told that the taxi they were sitting in was self-driving. They were shown all the scenarios one by one and asked to rate the driving on several items.
- **Group II:** This group was told that the taxi is driven by a human. They were shown the same scenarios as those of the first group and asked to rate the driving on the same items.

### B. Simulation Platform

The design and planning of the videos in the scenarios was made in Unity 3D [9], a cross-platform game engine that allows users to create 3D and 2D games. In order to simulate traffic dynamics and city ambience, the following asset packages were used:

- *Traffic Dynamics:* Urban Traffic System PRO 2018.1
- *Vehicles and Humans:* Urban Traffic System PRO 2018.1
- *Suburban houses:* UK Terraced Houses
- *City buildings:* Town Constructor Pack
- *Roads:* Low Poly Street Pack

### C. Scenarios

Screenshots of the scenarios can be seen in Fig [1]. The scenario videos were shot from the point of view of the shotgun rider of the taxi. Of these, 3 depicted *high-risk* situations, 3 *low-risk*, and 4 *no-risk*. The no-risk scenarios served as a measure of critical judgement from participants when no error in driving is shown. They also helped in forming a favorable opinion of the taxi driver so that they are not judged too critically.

**Scenario 1 - Sudden Braking (High-risk)** : According to an article [12] published on 30th October, 2020, Waymo cars were involved in 11 rear-end collisions. This occurred due to sudden braking of the Waymo cars when they faced an unexpected circumstance. In this scenario, the taxi is shown riding on a highway in the middle lane. A biker tries to cut into the middle lane quite suddenly, due to which the taxi comes to a complete halt. As a result of this, a vehicle coming at a fast pace behind it crashes into the taxi's rear end. On a highway, this is a reasonable expectation.

**Scenario 2 - Slow Driving (Low-risk)** : Users of Waymo's self-driving taxi service, called Waymo-One, have complained about the constant braking and timid driving of the taxis [14]. Autonomous Vehicles being tested in the Silicon Valley are found driving at speeds way below speed limit on empty roads, causing bottlenecks and much annoyance [15]. To simulate this behavior, we designed this scenario where a taxi is driving way below the speed limit. At a later part of the video, the rear view of the taxi is shown where a line of vehicles are seen honking at the taxi.

**Scenario 3 - Left Turn (No-risk)** : Autonomous Vehicles seem to be doing great on the unguarded left turn front [13]. In this scenario, the taxi is shown making a brilliant left turn.

**Scenario 4 - Overcorrection (Low-risk)** : From 22nd October 2020, Tesla began pushing its Fully-Self Driving Beta features to a select few of its customer base. It was purported to make the auto-pilot mode "zero-intervention" [5]. Days after its release, videos began surfacing of drivers reporting major and minor errors in the judgement of the Tesla AV cars [4]. One such error is presented in this scenario. In this scenario, the taxi is seen making a wide left turn on an empty road and over-correcting its path after the turn. Over-correction can be dangerous for vehicles coming from behind.

**Scenario 5 - Red Light Running (High-risk)** : According to an analysis performed by the AAA Foundation for Traffic Safety, two people are killed every day on U.S. roads by red light runners [6]. According to an article [7], in a simulation, a Waymo car barely missed a car speeding towards it in an intersection. The car's safety driver immediately took over to avoid the crash. In such cases, the car merely slows down

upon detecting a vehicle, just enough to avoid collision. A human driver would have come to a complete stop to avoid unforeseen accidents. This is one of the greatest drawbacks in self driving cars. They lack human intuition. This scenario presents a situation where the taxi goes through a green light. On the right, another car is seen speeding towards the intersection with no sign of stopping. Instead of coming to a complete halt, the taxi continues on with a calculated speed that will just avoid a collision.

**Scenario 6 - Pedestrian Crossing (No-risk)** : In this scenario, the taxi has a green light at an intersection. But when it detects a pedestrian crossing the road, the taxi stops and waits.

**Scenario 7 - Obeying Traffic Lights (No-risk)** : The taxi comes to a halt at a red light, and starts moving once the green light comes on.

**Scenario 8 - Rash Merging (High-risk)** : On 14th February, 2016 in California, a Waymo car was angling towards a lane from a parking spot at 2mph. It's front hit the side of a bus that was coming from behind. The operators stated that the car had detected the bus, but had predicted that it would yield for the Waymo car [1]. This scenario replicates this incident exactly. Human errors and their unpredictable behaviors play a major role in wayward driving in autonomous vehicles.

**Scenario 9 - Smooth Merging (No-risk)** : In this scenario, a smooth merging into the fast lane of a highway is shown. The Taxi overtakes a firetruck and merges into its lane smoothly right behind another vehicle.

**Scenario 10 - Driving Area Misclassification (Low-risk)** : This scenario takes into account a video released by a Tesla car owner [8]. The new Fully Self Driving Beta feature had been pushed into his car and he had taken it for a ride. The car arrives at an intersection where it intends to turn left. The left road has a parking lot attached to it. The car wrongly perceives the parking spot for a lane, drives into a car parked in the parking lot and almost hits it. This situation has been replicated exactly in this scenario.

### D. Participants

40 participants volunteered to fill in a survey. All participants had at least 1 year of driving experience. Group I consisted of 11 females and 9 males. Group II consisted of 11 females and 9 males.

### E. Measures

For each Scenario, participants were asked to rate the driving on the following factors:

- *Skills*: How good was the taxi's driving?
- *Severity*: How severe was the problem in this scenario, if there was any?

- *Blame on the taxi*: If there was a problem, how much is the Taxi to blame?

The results of these factors are presented in Tables 1 through 12. Following are the notations used in these tables.

- $M_R$ : Median of Robotaxi sample
- $M_H$ : Median of Human-driven taxi sample
- $U$ : U-value in Mann-Whitney U-test
- $R$ : Robotaxi
- $H$ : Human driven Taxi
- $r$ : Effect size

Certain other factors were rated using a 5 point Likert scale: *Judgement, Comfort, Safety, Carefulness, Speed, Patience and Alertness*. In order to understand the reasoning behind these ratings and ensure valid responses, an open ended question was asked: *In a few words, describe what you observed*.

After the 10 scenarios, participants were asked to rate the taxi service on a scale of 1 to 5. They were asked if they'd return to this service in future and an open ended question to justify their answer.

#### F. Hypothesis

While there are many metrics to base our hypothesis on, we zero in on favorability rating: user's choice on using the taxi service in the future and their rating of the taxi. Using this metric, we formulate the following hypothesis:

**Hypothesis.** Given identical situations, the Robotaxi has a higher favorability rating than that of the human-driven taxi.

### IV. DISCUSSION

**Scenario 1 results:** From the results of Mann-Whitney U-test given in Table [1], it can be seen that participants blamed the human-driven Taxi more than the Robotaxi for the accident. Answers in the open ended question reveal some of their attitudes. For the Robotaxi, most comments stated that the sudden merging of the biker caused the accident. On the other hand, the human-driven Taxi was blamed for stopping suddenly, or for not detecting the biker early on. From Fig [2], it can be seen that most ratings are comparable in both the cases.

Measures	$M_R$	$M_H$	U	$p(R < H)$	$p(R > H)$	r
Skills	4	4	187.5	0.359	0.641	0.057
Severity	3	3	206	0.57	0.43	0.03
Blame	2.5	1	264.5	0.968	<b>0.033</b>	0.29

TABLE I

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 1

**Scenario 2 results:** As can be seen from Table [2], the Robotaxi was rated highly for its Skill even though its slow driving was a cause for annoyance. No such effect was observed for the human-driven Taxi. The Robotaxi was also praised for its carefulness albeit slow driving. From Fig [3], the human-driven Taxi was rated below the Robotaxi in all aspects of the driving except Patience.

Measures	$M_R$	$M_H$	U	$p(R < H)$	$p(R > H)$	r
Skills	3	2	248	0.913	<b>0.092</b>	0.21
Severity	2	2	190	0.393	0.618	0.043
Blame	4	4	198	0.483	0.529	0.0069

TABLE II

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 2

**Scenario 3 results:** Even though this was a no-error scenario, the human-driven Taxi was rated significantly less than the Robotaxi in all aspects of driving as can be seen in Fig [4]. The following of traffic rules (stopping at red light and making a successful left turn) was more positively perceived for the Robotaxi. For the few that did manage to find cause for complaints even in this scenario, they were likely to blame the human-driven Taxi for a perceived "speeding up near the left turn" as compared to the Robotaxi. This is given by Table [3].

Measures	$M_R$	$M_H$	U	$p(R < H)$	$p(R > H)$	r
Skills	4.5	3.5	243	0.893	<b>0.112</b>	0.19
Severity	1	2	158	0.208	0.902	0.2
Blame	3	3	164.5	<b>0.144</b>	0.863	0.17

TABLE III

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 3

**Scenario 4 results:** Every participant observed that the over-correction can lead to accidents in both the cases. However, from Table [4], the Robotaxi was perceived to have better driving skills than the human-driven Taxi. Even the severity of the over-correction was perceived to be more in case of the human-driven Taxi. This contributes to the same trend that was observed in the previous scenarios, that the Robotaxi is perceived in a better light even though both the vehicles manage to make the same mistakes. In Fig [5], it is seen that both the vehicles were rated almost equally on all aspects.

Measures	$M_R$	$M_H$	U	$p(R < H)$	$p(R > H)$	r
Skills	3	3	242	0.884	<b>0.121</b>	0.18
Severity	2	2	173.5	0.228	0.781	0.12
Blame	4	4	188	0.371	0.64	0.052

TABLE IV

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 4

**Scenario 5 results:** Of all the Scenarios, this is the most dangerous one. Skills of the Robotaxi were rated highly, as shown in Table [5]. In the open-ended comments, participants had blamed multiple characters. The Robotaxi was perceived as not having detected the speeding car from its right. But, the blame was put mostly on the other driver running the red light. On the other hand, most of the blame was put on the human-driven Taxi. Many assumed that running a red light is common and that the car doing so was automatically at fault. However, the human-driven Taxi was blamed to be at fault too because it could clearly see the speeding car on its right but made no attempts to stop. Fig [6] shows the dominance of the Robotaxi on most of the driving aspects over the human-driven Taxi.

**Scenario 6 results:** In this no-error scenario, most of the ratings on driving aspects are comparable, as seen in Fig [6].

Fig. 1. **Screenshots of the scenarios:** 1. Sudden braking, 2. Slow driving, 3. Smooth left turn, 4. Over-correction, 5. Red light running, 6. Pedestrian crossing, 7. Obeying traffic lights, 8. Rash merging, 9. Smooth merging, 10. Driving area misclassification

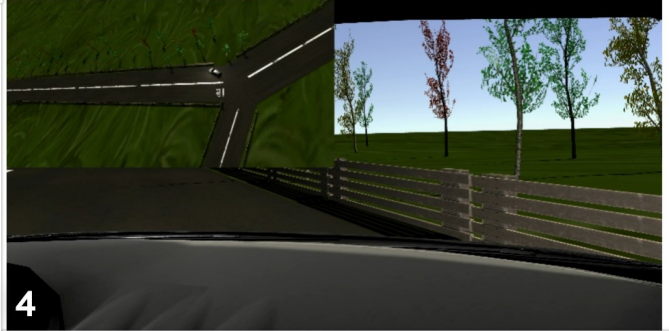
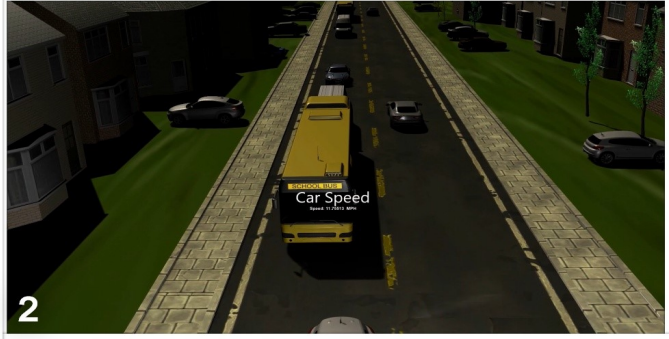




Fig. 2. Scenario 1 ratings

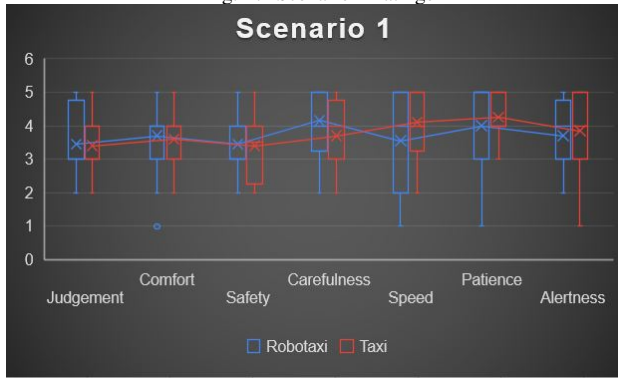


Fig. 3. Scenario 2 ratings

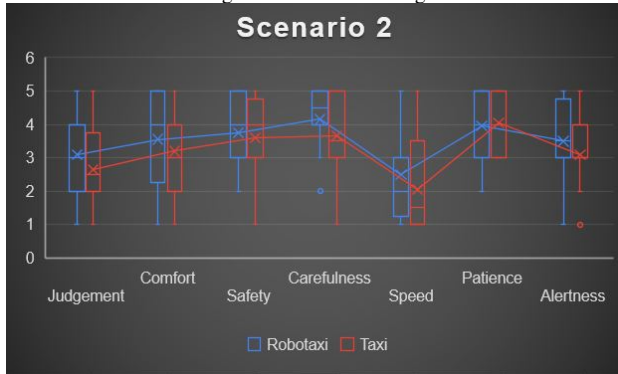
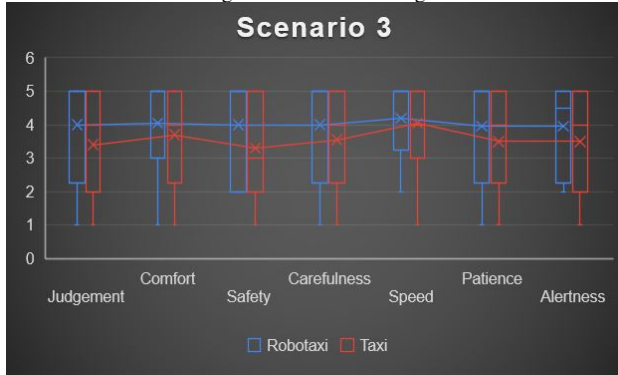


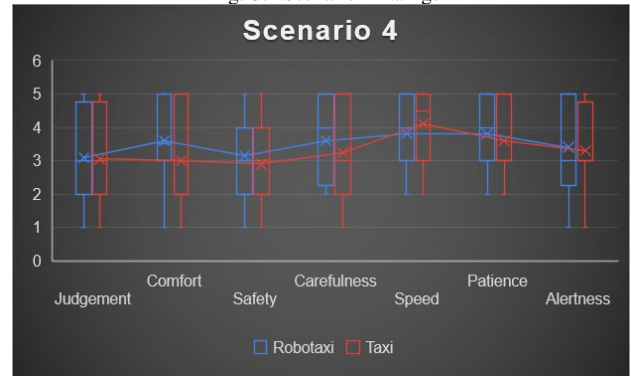
Fig. 4. Scenario 3 ratings



However, we notice something interesting in Table [6]. In this scenario, the pedestrian started crossing the road even though the Taxi had right of way. In case of Robotaxi, the pedestrian was blamed for putting the Robotaxi in a tight spot, and naturally, the Robotaxi was praised for having the sense to stop for him. This effect was not observed in case of the human-driven Taxi.

**Scenario 7 results:** In this no-error scenario, the Robotaxi was praised for its ability to detect the traffic lights and following the rules. On the other hand, many complaints were dug up (even though they weren't apparent from the videos) regarding the speed of the human-driven Taxi and its lack of alertness when it comes to pedestrians wanting to cross the

Fig. 5. Scenario 4 ratings



Measures	M <sub>R</sub>	M <sub>H</sub>	U	p (R<H)	p (R>H)	r
Skills	3	3	173	0.228	<b>0.121</b>	0.12
Severity	3	3	173.5	0.228	0.78	0.12
Blame	3	2.5	186.5	0.354	0.343	0.059

TABLE V

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 5

Fig. 6. Scenario 5 ratings

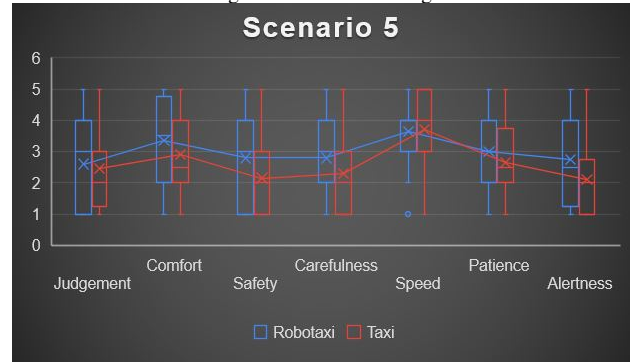


Fig. 7. Scenario 6 ratings



road. From Table [7], this effect is seen as a high tendency to blame the human-driven Taxi over the Robotaxi. Both the services were rated comparably on the driving aspects as seen from Fig [8].

**Scenario 8 results:** This is another dangerous scenario depicting a crash with a slow moving bus. Both the services were reprimanded for their bad judgement in driving. However, even in the absence of any excuse to forgive the Robotaxi, it scored

Measures	M <sub>R</sub>	M <sub>H</sub>	U	p (R<H)	p (R>H)	r
Skills	5	5	204	0.56	0.453	0.019
Severity	1	1	203	0.546	0.467	0.013
Blame	3	3	163	<b>0.12</b>	0.886	0.19

TABLE VI

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 6

Measures	M <sub>R</sub>	M <sub>H</sub>	U	p (R<H)	p (R>H)	r
Skills	5	5	182	0.284	0.727	0.09
Severity	1	1	200	0.509	0.509	0.0035
Blame	3	3	181.5	<b>0.175</b>	0.838	0.16

TABLE VII

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 7

Fig. 8. Scenario 7 ratings



significantly better in the ratings on driving aspect as can be seen from Fig [9]. The Robotaxi was perceived to have better driving skills as compared to the human-driven Taxi and the severity of the accident was exaggerated in case of the human-driven Taxi, as shown in Table [8].

Measures	M <sub>R</sub>	M <sub>H</sub>	U	p (R<H)	p (R>H)	r
Skills	2	1	271	0.988	<b>0.013</b>	0.35
Severity	3	3.5	167.5	<b>0.167</b>	0.841	0.15
Blame	5	5	182	0.267	0.745	0.099

TABLE VIII

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 8

Fig. 9. Scenario 8 ratings



**Scenario 9 results:** Similar to Scenario 7, this was a no-error scenario where the human-driven Taxi got significantly worse ratings than the Robotaxi in perception of driving skills

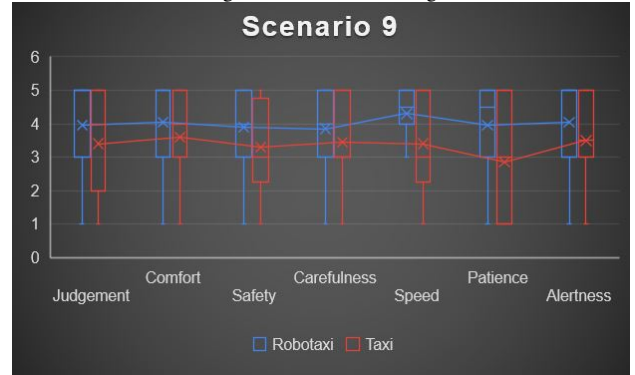
as shown in Fig [10]. Just like in Scenario 7, participants judged the human-driven Taxi's merging skills rather harshly as shown in Table [9], stating that it could have made minor adjustments or that it was impatient while merging into the fast lane. No such observations were made in case of the Robotaxi, pointing to a clear bias in perception of self-driving cars.

Measures	M <sub>R</sub>	M <sub>H</sub>	U	p (R<H)	p (R>H)	r
Skills	4.5	3	267	0.971	<b>0.031</b>	0.3
Severity	1	1	159	<b>0.1</b>	0.9	0.2
Blame	3	3	181.5	0.283	0.728	0.091

TABLE IX

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 9

Fig. 10. Scenario 9 ratings



**Scenario 10 results:** This scenario, where the Taxi nearly hits the rear-end of another parked car right next to a left turn, brings in results that are in stark contrast to each other. Clearly, the human-driven Taxi receives bad ratings in all aspects. The taxi driver is thought to be drunk or having fallen asleep while turning.

On the other hand, the Robotaxi is not blamed for the wide left turn to that extent. Instead, it is noted for its obstacle detection. The ratings received by the Robotaxi are better than the Taxi by a wide margin as can be seen from Fig [11]. In Table [10], we see that the human-driven Taxi is perceived to be far less skilled and is blamed more. The situation is judged to be far more severe in this case.

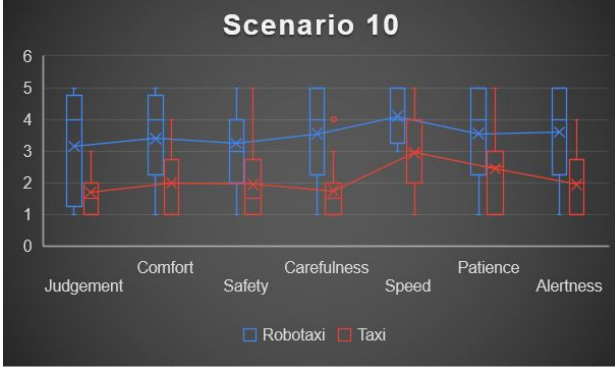
Measures	M <sub>R</sub>	M <sub>H</sub>	U	p (R<H)	p (R>H)	r
Skills	3	2	320.5	0.999	<b>0.0008</b>	0.53
Severity	2	3	121.5	<b>0.014</b>	0.987	0.35
Blame	4	5	155	<b>0.096</b>	0.909	0.21

TABLE X

MANN-WHITNEY TEST COMPARISON FOR SCENARIO 10

**Reaction to High-Risk vs Low-Risk Scenarios:** High risk scenarios are likely to bring out far intense reactions from the passengers involved. But we wanted to find out if there's a difference in the intensity of reactions between the Robotaxi and human-driven Taxi. This was tested out across all the High-risk and Low-risk scenarios. The results are presented in Table [11]. When it comes to low-risk scenarios, participants were generally in favor of the Robotaxi. However, no such

Fig. 11. Scenario 10 ratings



trend was observed in case of high-risk scenarios. It is inferred that when it comes to life-endangering situations, intolerance towards mistakes is almost equally likely in both the taxis.

Measures	High-Risk Scenarios		Low-Risk Scenarios	
	p (R > H)	p (R < H)	p (R > H)	p (R < H)
Skills	0.449	0.552	<b>0.0003</b>	0.999
Severity	0.801	0.200	0.972	<b>0.028</b>
Blame	0.318	0.684	0.868	<b>0.133</b>

TABLE XI

MANN-WHITNEY TEST COMPARISON BETWEEN HIGH-RISK AND LOW-RISK SCENARIOS

**Final results:** Over the scenarios, the Robotaxis enjoyed a clear bias which is reflected in the overall service ratings and the answers to the question: "Will you be willing to use this service in the future?" (seen from Fig [13]). From Table [12], a clear preference of the Robotaxi is seen. Participants are more willing to try the Robotaxi in future inspite of its flaws (as seen from Fig [12]), which fares well for self-driving companies that are looking forward to starting self-driving cab services. From the above results, our hypothesis was supported.

Measures	U	p (R<H)	p (R>H)	r
Service Rating	276	0.986	<b>0.015</b>	0.34
Use in future?	300	0.999	<b>0.0003</b>	0.54

TABLE XII

MANN-WHITNEY TEST COMPARISON FOR FINAL RATINGS

Fig. 12. Consensus on choosing the service in future

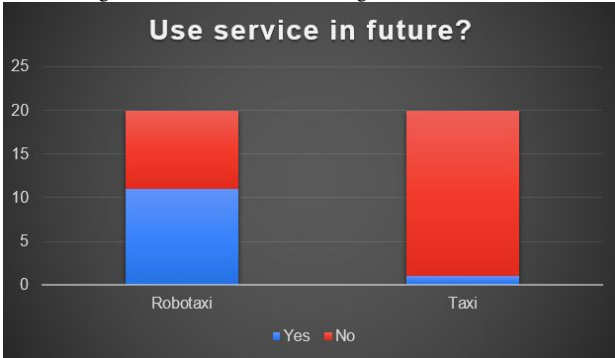


Fig. 13. Overall rating



## V. CONCLUSION

There's a clear bias in judgement between self-driving cars and human driven cars. Given the results from this study, it seems likely that customers may turn to self-driving cab services, given their positive perception of self-driving cars and faith in the autonomous technologies. It can be inferred from the results that even though humans are critical of autonomous car's failures, they are willing to judge the mistakes less harshly. The results show that humans are willing to accept autonomous technology and give it a chance for growth and improvement. AV companies can use this inference to their benefit as it will ensure the companies ample time to develop, test, optimize and improve their driverless technology.

## REFERENCES

- [1] Nellie Bowles. Google self-driving car collides with bus in california, accident report says. *The Guardian*.
- [2] Mike Brown. Tesla robo-taxi: Elon musk outlines 3-step plan to fully-autonomous rides. *Inverse*.
- [3] Matthew DeBord. Waymo has launched its commercial self-driving service in phoenix — and it's called 'waymo one'. *Business Insider*.
- [4] Paul A. Eisenstein. Tesla's beta test of "full self-driving" system worries drivers, pedestrians — and even owners. *NBC News*.
- [5] Eva Fox. Elon musk on tesla fsd: Will release zero-intervention limited beta in a few weeks. *Tasmanian*.
- [6] Andrew Gross. Red light running deaths hit 10 year high. *Newsroom*.
- [7] Andrew Hawkins. Waymo pulls back the curtain on 6.1 million miles of self-driving car data in phoenix. *The Verge*.
- [8] Peter Holderith. Tesla's 'full self driving' beta tech nearly wrecked this model 3 into a parked car. *The Drive*.
- [9] Arthur Juliani, Vincent-Pierre Berges, Esh Vckay, Yuan Gao, Hunter Henry, Marwan Mattar, and Danny Lange. Unity: A general platform for intelligent agents. *CoRR*, abs/1809.02627, 2018.
- [10] Kanwaldeep Kaur and Giselle Rampersad. Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *Journal of Engineering and Technology Management*, 48, 05 2018.
- [11] John Krafcik. Waymo is opening its fully driverless service to the general public in phoenix. *Waymo Blog*.
- [12] Jane Lanhee Lee. Waymo releases phoenix area self-driving car incident score card. *Reuters*.
- [13] Sasha Lekach. This is why self-driving cars suck at making unprotected left turns. *Mashable*.
- [14] Paul Morris. Waymo data shows system's weaknesses. *Seeflection*.
- [15] Faiz Siddiqui. Silicon valley pioneered self-driving cars. some tech-savvy residents don't want them tested in their neighborhoods. *The Washington Post*.