Waymo Simulated Driving Behavior in Reconstructed Fatal Crashes within an Autonomous Vehicle Operating Domain

John M. Scanlon, Kristofer D. Kusano, Tom Daniel, Christopher Alderson, Alexander Ogle, Trent Victor

Waymo, LLC

Abstract

Preventing and mitigating high severity collisions is one of the main opportunities for Automated Driving Systems (ADS) to improve road safety. This study evaluated the Waymo Driver's performance within real-world fatal collision scenarios that occurred in a specific operational design domain (ODD). To address the rare nature of high-severity collisions, this paper describes the addition of novel techniques to established safety impact assessment methodologies. A census of fatal, human-involved collisions was examined for years 2008 through 2017 for Chandler, AZ, which overlaps the current geographic ODD of the Waymo One fully automated ride-hailing service. Crash reconstructions were performed on all available fatal collisions that involved a passenger vehicle as one of the first collision partners and an available map in this ODD to determine the pre-impact kinematics of the vehicles involved in the original crashes. The final dataset consisted of a total of 72 crashes and 91 vehicle actors (52 initiators and 39 responders) for simulations. Next, a novel counterfactual "what-if" simulation method was developed to synthetically replace human-driven crash participants one at a time with the Waymo Driver. This study focused on the Waymo Driver's performance when replacing one of the first two collision partners. The results of these simulations showed that the Waymo Driver was successful in avoiding all collisions when replacing the crash initiator, that is, the road user who made the initial, unexpected maneuver leading to a collision. Replacing the driver reacting (the responder) to the actions of the crash initiator with the Waymo Driver resulted in an estimated 82% of simulations where a collision was prevented and an additional 10% of simulations where the collision severity was mitigated (reduction in crash-level serious injury risk). The remaining 8% of simulations with the Waymo Driver in the responder role had a similar outcome to the original collision. All of these "unchanged" collisions involved both the original vehicle and the Waymo Driver being struck in the rear in a front-to-rear configuration. These results demonstrate the potential of fully automated driving systems to improve traffic safety compared to the performance of the humans originally involved in the collisions. The findings also highlight the major importance of driving behaviors that prevent entering a conflict situation (e.g. maintaining safe time gaps and not surprising other road users). However, methodological challenges in performing single instance counterfactual simulations based solely on police report data and uncertainty in ADS performance may result in variable performance, requiring additional analysis and supplemental methodologies. This study's methods provide insights on rare, severe events that would otherwise only be experienced after operating in extreme real-world driving distances (many billions of driving miles).

1. Introduction

1.1 Safety of Level 4 Automated Driving Systems

The latest annual roadway traffic fatalities are at nearly 36 thousand in the United States and approximately 1.35 million globally (National Center for Statistics and Analysis 2020; World Health Organization 2018). No operational design domain (ODD) for human-driven passenger vehicles is immune to the occurrence of fatal collisions. Human drivers must be able to plan and safely execute their own navigational plan, while also being prepared to detect and react to sudden, unexpected actions taken by other road users. The role of human driver behavior as a critical reason for crash causation cannot be understated, and will continue to be so for years to come. Human error, often by a driver, is the most widely cited causative factor that leads to crashes (Singh 2015).

One of the advantages of SAE Level 4 Automated Driving System (ADS) technology, hereinafter referred to as ADS (SAE J3016, 2018), is its potential for driving performance that surpasses that of a human in avoiding and mitigating collisions. Research into the potential safety benefits of ADS has highlighted major contributions due to the fact that ADS vehicles are designed to follow traffic laws (e.g., adhere to posted speed limits, obey traffic controls) and cannot engage in the kinds of human activity (e.g., distraction, drowsiness, and intoxication) responsible for a large proportion of the degraded driving performance of human drivers (Dobberstein et al., 2018, Mueller et al., 2020). Even so, ADS will share roads with human drivers and will encounter highly variable human driving performance, including severe human errors, for the foreseeable future. Additionally, ADS technology designers will seek to design a vehicle that behaves predictably for other human road users, and is robust against failure modes that pose risk of collision or injury, in order to avoid offsetting these aforementioned advantages. Therefore, designers consider both operating in a way that does not initiate conflicts and, when possible, mitigate or avoid collisions due to the variable and dangerous behavior of human drivers.

Numerous commercial entities are developing ADS for fully automated operation without any human intervention within a specified ODD. Some initial crash involvement data have been published from public road testing of ADS (Schwall et al., 2020) and advanced driver assistance systems (ADAS) that operate with the human fully engaged in the driving task (Cicchino 2017, Isaksson-Hellman and Lindman 2015). Estimates are that over 10 billion miles of actual driving in the United States would be required to make statistically significant assertions regarding ADS technology's efficacy in preventing fatal collisions (Kalra and Paddock 2016; Lindman et al., 2017). These estimates highlight the fact that high-severity, real-world collisions are rare, and that drawing conclusions about the efficacy of the ADS is challenging under most deployments planned for the near future if one were to only rely on real-world driving mileage. Throughout 6.1 million miles of public road driving in 2019, the Waymo Driver did not experience a single high severity collision (as determined by a 10% or higher risk of serious injury to any party involved) (Schwall et al., 2020). Using historical crash data, a representative set of human drivers operating over similar driving distances would also likely experience few high severity events. Yet, preventing and mitigating these infrequent high severity collisions is one of the main opportunities for ADS to improve road safety. To address the rare nature of high-severity collisions, this paper introduces novel techniques to established safety impact assessment methodologies. This study focused on fatal collisions because they have high societal cost (Blincoe et al., 2010) and have been identified as a priority for elimination by world-wide "Vision Zero" initiatives (Johansson, 2009; Tingvall et al., 1999). Using simulation of reconstructed fatal collisions, from the same ODD in which the ADS is deployed, we assessed the system's capacity to (a) avoid inducing the observed human-driven fatal collisions and (b) respond to the actions of another human actor who induced the collisions.

Waymo is an autonomous driving technology company that operates a fully autonomous ride-hailing service, Waymo One, in the East Valley of Phoenix, AZ. Waymo recently published the various methodologies it uses to determine safety readiness of its ADS (Webb et al., 2020), presenting that no single methodology is sufficient by itself to evaluate ADS safety. As such, Waymo uses multiple complementary methods to assess the safety of the hardware, behavior, and operations used in its commercial ADS service. The current paper presents a novel scenario-based testing methodology within the ADS behavior layer. Collision avoidance testing is a type of simulated scenario-based testing within the set of methodologies Waymo has previously published (Webb et al., 2020). The collision avoidance testing program evaluates the Waymo Driver in thousands of situations of varying severity where urgent evasive braking and/or steering would likely be required to avoid a collision. The purpose of the collision avoidance testing scenario database is to be broad and general, whereas the current study simulated reconstructed collisions within a specific geographic ODD and severity level (fatal collisions). Counterfactual simulations of reconstructed fatal collisions can be used to supplement the existing scenario database, which consists of conflict scenarios from naturalistic driving databases, crash databases, and ADS on-road testing.

1.2 Overview of Prospective Safety Impact Methodology

Prospective safety impact methods aim to estimate the effect of a safety system before it is widely deployed. The prospective safety impact methodology approach aims to predict future effects, as opposed to a retrospective study, which observes the benefits of a safety system after it has been widely deployed in the field. The prospective effect can be measured by estimating the number of collisions avoided and injuries it will prevent. Najm and daSilva (2000) originally proposed a methodology for prospective estimation of safety impact as a function of (a) conflict exposure and (b) collision avoidance.

The ISO 21974-1 definition of a conflict is a "situation where the trajectory(ies) of one or more road users or objects (conflict partner) led to one of three results: a crash or road departure, a situation where an evasive manoeuvre(s) was required to avoid a crash or road departure, or an unsafe proximity between the conflict partners" (ISO TR21974-1). Collision avoidance refers to the evasive maneuvers of road users, that is "any action performed by any conflict partner to change its trajectory or speed in an attempt to avoid or reduce the severity of a potential crash, avoid or reduce the severity of a road departure, or regain vehicular control after a loss of control" (ISO TR21974-1). In this study, conflict avoidance is defined as vehicle control that does not use urgent evasive maneuvers to avoid entering into a conflict. Another way to describe conflict avoidance action in more simple terms is "normal" or "defensive" driving, such as maintaining safe time gaps and not surprising other road users. If a conflict state is not avoided, collision avoidance action, that is, urgent evasive maneuvering, is required to mitigate or avoid a potential collision.

The prospective safety impact technique relies on two primary principles. First, any eventual collision and potential injury can be avoided by simply avoiding the conflict altogether. Computing safety impact due to conflict avoidance can be achieved by examining conflict exposure rates during field testing (Najm and daSilva 2000). Second, if a conflict was not prevented, collision avoidance performance dictates the effectiveness of the system's ability to take evasive action to avoid or mitigate the crash.

This prospective estimation approach was first introduced to estimate the collision avoidance potential of new driver assistance technologies. The technologies studied using this technique include, but are not limited to, anti-lock brakes / electronic stability control (Riexinger et al., 2019; Blower, 2013), forward collision avoidance (Van Auken et al., 2011; Kusano & Gabler, 2012), lane departure prevention (Gordon et al., 2010; Kusano & Gabler, 2015; Scanlon et al., 2016), intersection assistance systems (Bareiss et al., 2019; Sander & Lubbe, 2018; Scanlon et al., 2017), vulnerable road user, such as pedestrians and cyclists, collision prevention systems (Jermakian & Zuby, 2011; Yanagisawa et al., 2017; Gruber et al., 2019; Haus & Gabler 2020; Haus et al., 2019; Rosen et al., 2010), and adaptive cruise control with lane centering (Bärgman & Victor 2020). Most of these prospective estimations were applied to advanced driver assistance systems (ADAS), where the human driver is responsible for control of the vehicle, but the system can intervene if an imminent collision scenario presents itself. Accordingly, these past studies largely seek to address the collision avoidance component of Najm's proposed safety impact approach. Most of these previous studies do not consider faults or failures in the introduced technology (e.g., electrical or mechanical faults, failures in perception), and have limitations in simulation of realistic sensor performance (e.g., variance in sensor performance). Methods to minimize faults are covered by functional safety elements of Waymo's safety readiness determination (Webb et al., 2020).

Counterfactual simulation studies involving ADS, as opposed to ADAS, require a slightly different approach because unlike ADAS, ADS does not have a human driver in the loop. When an ADS is controlling the entire driving task, the conflict avoidance, in addition to the collision avoidance, of a particular scenario will be inherently unique to that ADS.

Past safety impact methods for ADS have used a scenario database populated with scenarios inspired by human-involved collisions. These scenario databases are then used as the basis for simulations of the ADS. See Riedmaier et al. (2020) for an in-depth survey of scenario-based testing for ADS. Several consortia and standards bodies are working on scenario-based testing. The PEGASUS project developed a methodology to evaluate an ADS, focusing on a highway automation system (PEGASUS). The Advanced Vehicle Technology (AVT) Consortium, hosted by the Massachusetts Institute of Technology (MIT), is a collaborative research effort between automakers, insurance companies, tier-1 suppliers, and research organizations developing and testing methods for safety benefits estimation. P.E.A.R.S. is a consortium striving to standardize prospective safety evaluations (Prospective Effectiveness Assessment for Road Safety). The International Standards Organization (ISO), Technical Committee 22 on road vehicles, standards committee on safety impact testing, working group 7 on traffic analysis methodology (ISO/TC 22/SC 36/WG 7) is developing standards on "Prospective safety performance assessment of pre-crash technology by virtual simulation".

The current study contributes to the state of the art in two notable aspects. First, we introduce a novel method for constructing simulations of an ADS replacing humans in human-involved, real-world collisions. Crash reconstructions are used as the basis for constructing simulation scenarios, where each human actor is systematically replaced by the ADS and the vehicle trajectories are aligned such that a similar collision scenario to the reconstructed collision scenario occurs in simulation. This direct substitution allows for a one-to-one comparison with the original human operator to answer the question: how well would the ADS have performed if it was to come across the same fatal collision scenarios encountered by human drivers? Second, this methodology evaluates ADS performance in the role of both the conflict initiator and responder. As noted above, ADS will operate on shared roads with human drivers for the foreseeable future. Therefore, it is instructive to also evaluate the potential for an ADS to not initiate collisions, while also avoiding collisions as a responder to other vehicles' driving behavior.

This paper will investigate the safety impact of the Waymo Driver in crashes that are representative of what human drivers have experienced in the East Valley of Phoenix, AZ geographic ODD but may not represent the unique collection of crashes that an ADS system would experience if widely deployed. Therefore, the results of this study should not be viewed as a comprehensive safety impact estimate. The methodology produces a single representation and outcome for each simulation. Modeling variability in the simulated results introduced by uncertainties in the crash reconstructions as well as uncertainties in ADS performance is out of scope for this paper. Finally, the safety impact in this study is presented relative to the humans originally involved in the crashes. Future comparisons to some targeted baseline, such as an attentive human or ADAS-equipped vehicle, could provide additional context for ADS performance.

1.3 The ODD dilemma

The aforementioned prospective studies have most often relied on counterfactual simulation of nationally representative samples of crashes. For example, widely-used crash databases include the U.S. Fatal Accident Reporting Systems (FARS), National Automotive Sampling System (NASS), more recently the Crash Report Sampling System (CRSS) and Crash Investigation Sampling System (CISS), the German In-Depth Accident Study (GIDAS), and others. The simulations have relied on crash reconstructions of the pertinent components for evaluating system performance, e.g., vehicle kinematics and actor dimensions, and some models of the automated driving features and human operator behavior (if involved in the driving task).

The ODD dilemma is the tradeoff between using large, robust datasets (e.g., aforementioned nationally representative databases) and small, ODD-specific datasets like the one used in this study. The larger datasets may be representative of populations (e.g., the entire country) but not representative of a specific ODD. The smaller ODD-specific samples are representative of that ODD but may be smaller in size, and thus have larger variability potentially missing unobserved crashes.

One challenge with using larger representative crash data in counterfactual simulations for ADS is the need to address the representativeness of the sample for the particular crash population within the ADS ODD. The current study overcomes this challenge by using a census of fatal collisions from the intended deployment area of the specific ADS. This has the advantage of being retrospectively representative of the ODD without having to account for nationally representative sampling schemes. A nationally representative sample is advantageous for producing a robust estimate of a potential systems effect but, for more localized assessments, should be adapted for the ADS ODD. For ADS, like the Waymo Driver, that are deployed in a limited geographic area with distinct road network, environmental, and driving behavior characteristics, using a nationally representative sample has, therefore, an unknown effect on the safety impact's precision. In order to use a nationally representative sample to estimate performance in a specific ODD, subsetting or weighting of cases based on variables (e.g., population density, road type, speed limit, weather conditions, etc.) is required. Using a geographic ODD-specific sample reduces representativeness uncertainty.

To illustrate this point regarding the effect of different geographic ODDs on crash distributions, consider the differences based on geographic location in Figure 1. It shows the distribution of crash type from FARS 2010 to 2018 by location: the city of Chandler, AZ, the cities that make up the Phoenix (PHX) metropolitan area excluding Chandler, and the rest of the U.S. The crashes are normalized within the locations. As such, Figure 1 illustrates the differences in distributions of crash types in different geographic locations. Overall, the Phoenix area has a lower proportion of single vehicle (road departure, loss of control) and higher proportion of pedestrian, cyclist, and animal collisions compared to the rest of the U.S. The city of Chandler has more cross traffic (turning across path, straight crossing path) collisions compared to the Phoenix area. Outside of collision type, one might expect other parameters, such as roadway geometry, driving behaviors, and crash causation factors, to be ODD-specific. The ensuing safety impact will therefore be upon the ODD-specific distribution of crashes.

There are, of course, challenges with using an ODD-specific sample of crashes. In the current study, a small dataset of geographic ODD-specific collisions demonstrates the Waymo Driver's capabilities to avoid human fatal collisions. Given the size of the current geographic ODD-specific sample of crashes, using the results of this study alone to judge system performance would risk overfitting the system design to these crashes. For this reason, any collision that Waymo empirically observes, either through on-road testing or external crash databases, is considered for inclusion within the larger Collision Avoidance Testing scenario database, which consists of thousands of scenarios (Webb et al., 2020).

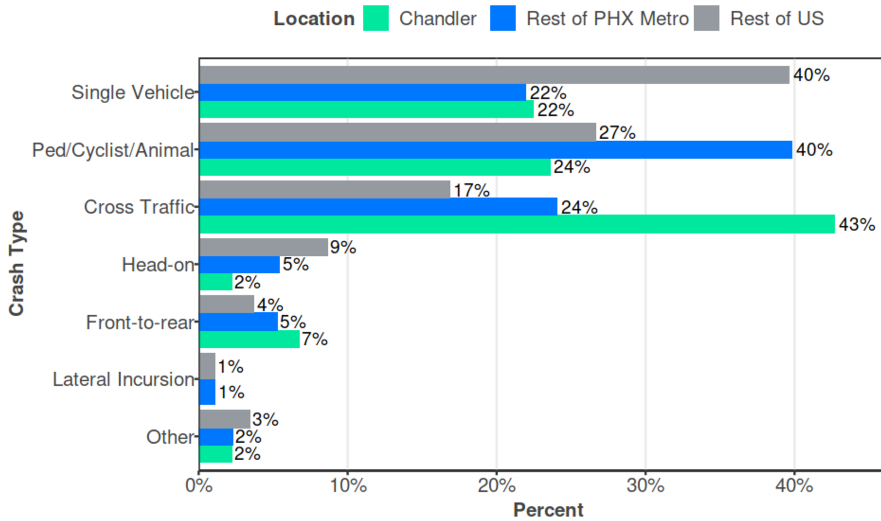

Figure 1. Fatal Crash Types by Location from FARS 2010 - 2018.

1.4 Objective and Research Questions

The current study extends previously implemented safety impact methodology techniques in order to evaluate the Waymo Driver's collision and mitigation performance if placed into real-world fatal collision scenarios. A census of actual fatal collision events from one of the Waymo Driver's ODDs are considered. Two research questions are posed. First, does the ADS avoid initiating the collision scenario? Second, how effective is the ADS at responding to the potential collision scenario initiated by another party?

2. Methods

2.1 Approach Overview

This study's methodology consisted of 5 primary steps, which are depicted in Figure 2. First, all available fatal collisions occurring within Chandler, AZ over a 10-year period (2008-2017) were identified, and materials relevant for a collision reconstruction were requested from the Arizona Department of Transportation (ADOT). Second, collision reconstructions were performed to determine the pre-crash kinematics leading up to the crash among other relevant parameters. Third, the initiator and responder roles were identified in each scenario. Fourth, simulations were performed to evaluate the Waymo vehicle's performance within the reconstructed fatal collision scenarios for each role. Fifth, the results were compiled across all collisions for both the initiator and the responder roles, and crash and injury prevention estimates were generated. Each step will be discussed in the following sections.

2.2 Data Source

ADOT compiles information on every police reported collision that occurs within the state. This publicly available data source (https://azdot.gov/) contains high-level, non-personally identifiable information about the collision,

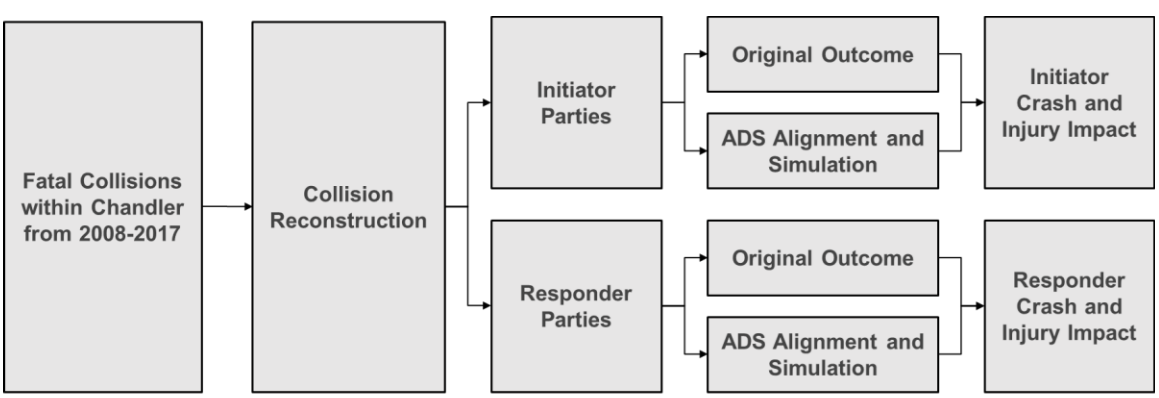

Figure 2. Outline of Fatal Crash Simulations.

environment, parties, and resulting injuries as it is available from police reports and related materials. This study relied on aggregated data tables from ADOT to identify all reported fatal collisions over the course of a 10-year period (2008-2017) that occurred within Chandler, Arizona. Chandler has an area of 65 square miles which is within the current Waymo One commercial service where the public can hail a fully automated vehicle without an autonomous specialist present.[1]

A total of 107 fatal collisions were identified over this time. Public data requests were made to the Arizona Department of Transportation, Arizona Department of Public Safety, and the Chandler Police Department to acquire materials relevant for collision reconstruction, including (if available):

- Police report(s)
- Scene diagrams
- Photographs
- Witness statements
- Event data recorder (EDR) reports
- Other miscellaneous reconstruction-relevant materials (e.g. surveillance footage)

The amount of information and data available for reconstruction varied from case to case. The fidelity of the reconstruction, and subsequent conclusions drawn from any counterfactual simulation, are directly tied to the quality of the underlying data being relied upon. A detailed discussion of data quality and its impact on the current study is addressed later in this paper.

Inclusion Criteria

This study evaluated the Waymo Driver's performance if placed into the role of one of the first two collision partners involved in a given collision sequence. This study protocol had several selection criteria for including an individual collision role within the *simulation case set*. First, only the first two collision partners (up to one initiator role and one responder role) were considered. Accordingly, all single vehicle collisions were simulated up to one time. Second, only class-1 and class-2 vehicles that were not towing any objects were considered in this current study. This was done to only evaluate the Waymo Driver performance within roles that were dynamically similar to the Waymo vehicle. Third, the incident must have occurred in an area with a valid Waymo map. This map availability requirement had three parts. The first requirement was that the map data available at the time of this study must have been similar to the roadway at the time of the original crash. For example, construction may have greatly altered the road structure making the map not match the road at the time of the crash. The second requirement was that, for a given role, the Waymo vehicle must be traveling within a mapped lane and traveling to some targeted mapped lane along a drivable path. This primarily excluded scenarios where the original vehicle may be traveling from a mapped lane onto some unmapped area, such as a private road or driveway. The third requirement was that the collision must have occured on roadway(s) with a speed limit of 45 miles per hour or less. This requirement was used to reflect Waymo's current fully autonomous ODD.

2.3 Role Classification

This study aimed to simulate the first two collision partners involved in the collision sequence (one actor if the collision sequence involved only a single party). Each collision partner has their own unique role in the collision sequence.

Initiator Role

Every collision in the study has an initiator party that took some initial, unexpected action that led to the eventual collision. This is not a designation of fault or right-of-way, but rather, an identification of the party that performed some initiating movement. Accordingly, making this distinction in this study required the consideration of two components. What actions taken were surprising or unexpected? And, which actor initially performed this surprising action? For example, consider a vehicle performing an unprotected left turn across the path of oncoming vehicles. If a collision was to occur during this unprotected turning maneuver, the initiating action in the collision sequence would generally be the movement by the left turning vehicle to make the turn, that is, the driver was expected to yield and the action to go initiated the series of collision events. It should be noted that the initiator party was not necessarily one of the initial collision partners. For example, some non-contacted party may have taken some initial, surprising action that induced

[1] The full boundaries of the Waymo One geographic ODD do not include the entire city of Chandler and also include some other cities in Phoenix's East Valley.

the series of collision events. In this study, the initiator party was always identified to be one of the initial collision partners.

Responder Role

Other parties involved in the collision sequence, if present, were then given the responder role. In this role, the actor must take some action in response to the series of events caused by the initiator party's initial, surprising action. In the left turn across path scenario outlined above, the approaching vehicle is the responder and must then react to the decision by the driver to perform the unprotected left turn. In the unprotected U-turn example, the responding vehicle approaching from behind must react to this stopped vehicle - however urgently - in order to evade a potential collision. For collision sequences involving more than two actors, subsequent actors would also fall into a responder role.

Implications of the Role on the Research Question Being Addressed

It is worth noting that the role being simulated directly influences the research question being addressed. Replacing the initiator party and simulating the event helps answer the research question: does the ADS avoid initiating the collision scenario? This can be accomplished through a number of proactive measures, including maintaining vehicle control or following road rules. The research question addressed by replacing the responder party is: how effective is the ADS at responding to some potential collision scenario initiated by another party?

2.4 Collision Reconstruction

A third-party engineering firm was contracted to perform the collision reconstructions according to reconstruction industry best practices. Reconstruction was performed for each case in order to generate a single representation of the pre-crash collision sequence based on the available evidence. The reconstructions were performed by trained experts without instruction or knowledge of the ADS design. The reconstructionists were also kept blind to the intended simulation purpose of the current study and were unaware of the identity of the simulating party (Waymo).

Every case had a unique collection of materials received to rely upon for performing the respective reconstruction. Additionally, the tooling utilized, such as the reliance on collision simulation software (e.g., HVE Software[2]), was dictated by the available materials. The following elements were determined through the reconstruction process:

(a) the pre-crash kinematics of each actor involved in the collision sequence,

(b) the pre-crash kinematics of any other actors/objects that were relevant to the collision events,

(c) relevant actor and object dimensions and inertial properties, and

(d) any traffic signal phase timings, if present.

The fifth and final reconstruction element required for counterfactual ADS simulation was the collision environment. The road conditions were first noted for every scenario. Every collision with coded roadway conditions considered in this current study was indicated to have dry surface conditions in the police-reported documentation materials. Scaled, orthonormal aerial images were used in the reconstruction process for determining the path of all relevant agents and objects within a geographic coordinate system. These kinematics-based data were then directly ingested into Waymo's simulation environment. The Waymo simulation environment used in this study contained a three-dimensional map. The three-dimensional map, which was generated from sensor data equipped on a Waymo test vehicle driving through the area, contains road characteristics as well as off-road structures and objects, such as trees, poles, and buildings. Reference points taken from within the custom Waymo simulation environment were used to ensure the alignment and orientation of the aerial images used during the crash reconstruction phase. A visual verification process by a human was performed using the reconstruction output to ensure proper replication of the reconstruction within Waymo's simulation environment.

2.5 Counterfactual Simulation

Simulation Platform

All reconstructed actors meeting this study's selection criteria were replaced and simulated within Waymo's simulation platform. The simulation platform is designed to provide a virtual testing environment that serves as a digital twin of the real-world driving environment. This platform is used for examining the Waymo Driver's behavior under various

[2] Engineering Dynamics Company, LLC 2020

conditions, and enables the testing of several key components of the Waymo design. The latest fully autonomous software intended for use in our Waymo One service at the time of writing (February 2021) was used for all simulations.

First, sensor simulation was utilized to represent realistic *perception* performance within the specific collision environment. While the collision reconstructions used to generate the simulations described the global positions of all relevant actors, the ADS is limited by the sensors in what external objects it can detect. This sensor simulation includes accounting for sensor range, field of view, sweeping behavior, and latency. Sensor simulation also replicates any inter-sensor delay as the system cycles through the available sensors and gains confidence in the identification and attributes of any perceived object. Additionally, the system was tested within a scaled environment based on three-dimensional map data. This environment contained salient three-dimensional scene elements, such as off-road obstructions, observed by the Waymo Driver when having previously traveled through the exact location where the collision occurred.

Second, the Waymo Driver's behavior layer is used in the simulation platform to control the simulated vehicle. The same behavior logic used in the on-road ADS deployment is replicated in the simulation environment. The behavior logic uses the simulated sensor data described above, the three-dimensional map data, and the simulated vehicle dynamics to perform the driving task including path planning, making predictions about other road users, and controlling the vehicle.

Pre-crash Alignment

An important piece of the ADS simulation procedure was *alignment* of the scenario, which refers to establishing the initial conditions of the simulated Waymo Driver. The alignment strategy allowed for the synchronization and recreation of events in a way that enabled a comparison of the ADS and the original vehicle involved in the crash. In this study, when the human was replaced, the Waymo Driver was allowed to control the vehicle as it was designed to operate. The difference between how the Waymo Driver would choose to operate and the actions of the original human driver could alter the timing of a potential collision between the two vehicles. For example, the crash reconstruction may have found that the human driver was traveling 20 mph above the posted speed limit prior to the collision. The Waymo Driver is designed to not travel above the speed limit. Replacing the Waymo Driver in this hypothetical crash at some pre-collision location could cause the simulated vehicles to no longer collide, but at the very least, would lead to a fundamentally different collision scenario. Therefore, the alignment process described in this section was used to align the Waymo Driver trajectory with the human-driven trajectory so that, absent any collision avoidance behavior, a similar collision scenario as was experienced by the original human driver would occur. It was essential that this alignment procedure allowed the Waymo Driver to behave as it normally would had it organically encountered this potential collision scenario in the real-world.

The alignment procedure for an exemplar two-party collision scenario is shown in Figure 3. This hypothetical scenario shows the responder traveling straight through an intersection and the initiator traveling straight and running a stop sign on a perpendicular road. In this example scenario, the responder vehicle is replaced with the Waymo Driver ADS. In general, this alignment methodology was used for all responder role simulations and some initiator role simulations. The potential collision partner had some set of intended pre-crash kinematics derived from the crash reconstruction. The replacement vehicle had some set of pre-evasive kinematics before enacting some evasive action. The procedure on this example scenario was as follows:

1. In stage one of the alignment process, the original reconstruction was ingested into the simulator.
2. In stage two, the location of the replacement vehicle during any crash avoidance action was determined.
3. In stage three, the intended kinematics of the replacement vehicle were extrapolated assuming the absence of the collision partner.
4. In stage four, the kinematics information from stage three was used to identify where the replacement vehicle and the collision partner would have overlapping trajectories, assuming that both the replacement vehicle and collision partner continued along their trajectories without influencing each other. In summary, both actors were assumed to continue traveling in the same manner as before any reconstructed avoidance maneuvers, that is, they were unresponsive to each other.
5. In stage five, using the information from stage four, the Waymo Driver was simulated and aligned so as to result in an overlapping trajectory with the collision partner in space and time assuming the absence of the collision partner. If present, all non-collision partner road users and simulated objects were included.
6. In stage six, a starting location for the Waymo Driver along this planned trajectory, substantially prior to the emergence of the collision partner, was selected. This

starting location was chosen to ensure the collision partner was out of sensor range of the Waymo Driver, allowing it to reach a steady state well before the interaction. The scenario was then simulated forward, and the Waymo Driver was enabled to take action as it normally would in the presence of the collision partner. In the Waymo Driver simulated scenario, the other actor was assumed to be unresponsive, i.e., no collision avoidance action modeled. This unresponsive driver modeling was done regardless of any avoidance action taken by that other actor in the reconstructed real-world event. This is generally a conservative modeling technique as drivers often take some sort of collision avoidance action prior to imminent collisions (Kusano & Gabler, 2013; Scanlon, 2017; Scanlon et. al., 2015).

The alignment procedure for the initiator role sometimes varied from the above procedures for several specific initiator role scenarios. For single vehicle collisions, mostly drift-out-of-lane road departure or loss-of-control crashes, the time point of the initial unexpected action (e.g., initial drift from lane or loss-of-control) by the initiator vehicle was replicated. This varied from Figure 3 in that the alignment procedure in stage four and five targeted the time point of the initial surprising movement. This procedure was performed in this manner to capture the location of the event and the relative position of other road users and objects. The second predominant variation was in initiator roles that involved running a red light. In these events, the stage five alignment procedure was performed with the ADS being artificially given a green light. In the actual simulation of the event (stage six), the signal phase would be set to its actual state at the time of the collision (red light).

This alignment procedure is a novel methodology developed for this study that enabled an evaluation of how the Waymo Driver would have behaved when replacing a human driver in the circumstances of the reconstructed scenario. This research topic of scenario alignment for simulation is an active area with no agreed upon best practice. Although the methods presented here covered the scenarios encountered in the examined crash reconstruction, additional methods or assumptions might be needed for other crash modes or configurations.

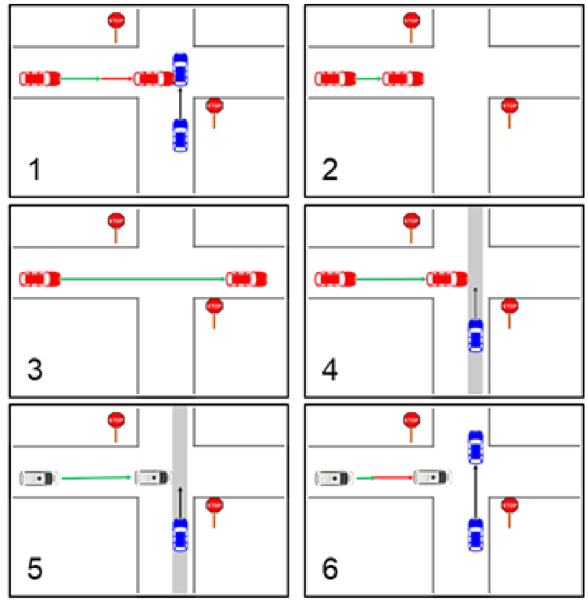

Figure 3. A depiction of the ADS-equipped alignment procedures. The red vehicle (approaching from left) represents the vehicle being replaced by the ADS (white vehicle). The blue vehicle entering the intersection from the bottom is the collision partner. Green arrows designate pre-evasive kinematics, whereas the red arrow highlights evasive kinematics. In stage (1), the original reconstruction is ingested in the simulator and its quality ensured. In stage (2), the location of the replacement vehicle during any crash avoidance maneuver is determined. In stage (3), the intended kinematics of the replacement vehicle are extrapolated assuming the absence of the collision partner. In stage (4), an overlap in spacetime trajectories is determined between the replacement vehicle and collision partner. In stage (5), the Waymo Driver is set to an overlapping trajectory in spacetime with the collision partner. In stage (6), the Waymo Driver is placed far enough prior to the collision partner that it is able to reach steady state before any interaction, and the scenario is then simulated forward to evaluate the Waymo Driver

Modeling Uncertainties in Traffic Signal Phase Timings

Traffic signal phase changes occasionally played a role in a given scenario. Additionally, this signal phase timing was not always readily apparent from the reconstruction materials but would be influential in the performance of the ADS. In the event the traffic signal phase state or change timing could be determined from the available evidence, it was modeled as determined in the collision reconstruction.

For cases involving a signal phase change from red to green as the replacement vehicle is approaching the traffic signal, the scenario was modeled as a green light throughout the approach. These red-to-green phase changes often involve the vehicle slowing for the red and then accelerating following the phase change. To be conservative in our assessment, a green light throughout the approach keeps the Waymo Driver at a maximum speed, thereby reducing the available window for response action to the non-right-of-way actor.

Another variation observed in this dataset was a traffic signal phase change from green to amber state. Varying the timing of when this phase change occurred would influence the ADS's decision to stop for the light or pass through the intersection. This study considered all relevant potential timings in 0.5 second increments. If any event(s) resulted in a collision, the variation with the highest predicted injury level was used in result reporting.

2.6 Severity Assessment

Every simulated role was evaluated according to four potential outcomes: avoided, mitigated, unchanged, and exacerbated:

- Avoided collisions were those in which the simulated vehicle roles were effective at preventing the collision altogether.
- Mitigated simulations were simulations that resulted in a collision with a lower collision severity than the original reconstruction.
- Unchanged simulations had an identical (or substantially similar) collision geometry and speeds as the original reconstruction.
- Exacerbated simulations were collisions with higher collision severity than the original reconstruction.

Collision severity was determined based on the collision dynamics of the first collision event. For single vehicle events, the single party contact event was considered. For multi-party vehicle-to-vehicle events, the first two collision partners were used. Vehicle-to-vehicle and single vehicle collision dynamics were assessed by computing delta-v (change in velocity as a result of collision) and principal direction of force (PDOF) using an adaption of the Kudlich-Slibar impact model (Brach et al., 2011; Kudlich, 1966). Impact speeds and orientations were used to assess collision severity for collisions involving vulnerable road users (e.g., pedestrians, bicyclists, and motorcyclists).

Collision dynamics were then used in conjunction with injury risk models to determine the collision severity. The risk of serious injury, that is, a maximum Abbreviated Injury Severity score of 3 or greater (MAIS3+) (Gennarelli & Wodzin, 2008), at the party-level for each of the collision partners was calculated. Party-level risk for a vehicle was computed by assuming same occupancy as the original subject collision. Regression models based off of NASS-CDS and CISS databases were used to establish the relationship between p(MAIS3+) and delta-v at various PDOF values (Prasad et al., 2015; Funk et al., 2008; Viano et al., 2008). For vulnerable road users, a collection of party type specific injury risk models were used to assess risk of serious injury (Nie et al., 2013; Tominaga et al., 2002; Rosen, 2013; Fredriksson et al., 2010). The maximum party-level serious injury risk was taken as the overall collision severity.

Occasionally, the first event was not the most severe event in the collision sequence. For these collisions, if the collision geometry and impact speeds were unchanged for the simulated first event, the collision would have been considered unchanged. However, there were not any residual collisions after simulating the Waymo Driver (i.e., collisions persisting after simulation), where the first event was not the most severe event.

Collision severity is dependent on a number of additional scenario-specific factors, including seating position, vehicle occupancy, occupant attributes, restraint use, and vehicle inertial properties, among other factors. This study focused on the Waymo Driver, that is, the ability of the ADS sensing and software systems to avoid and mitigate potential collisions rather than passive factors contributing to severity risk, such as vehicle or occupant attributes. Accordingly, the Waymo Driver vehicle was assumed to have identical inertial properties (i.e., mass and moment of inertia) as the original human driven vehicle. Injury severity was evaluated based on driver injury risk and any occupant information needed to compute injury risk (e.g., seatbelt use) was taken from the original scenario.

2.7 Repeatability of Simulation Results

ADS rely on complex, interdependent systems of software modules to monitor and react to potentially hazardous scenarios. Some components of this system, such as behavior prediction, planning algorithms, and perception, can exhibit non-determinism due to the nature of some of the algorithms employed. Additionally, the simulation environment can contribute to this non-deterministic effect due to asynchronous message passing between modules. This non-determinism can have an effect on the kinematics of the vehicle and timing of actions in the simulations. The effects of non-determinism in the simulation environment are continually monitored to ensure they have a minimal effect on the results of the simulations. Non-determinism in algorithms and simulation environments is a recognized challenge in testing automation systems and ADS (Baron et al., 2020; Koopman & Wagner, 2016).

This study accounted for this variable effect by simulating each case a total of five times. The “worst-case” scenario was used in the results of this study, as defined by the collision that occurred with the highest estimated injury risk as presented previously. The results of the five simulations showed only minor variations in outcomes, which suggested that the simulations in this case set did not trigger divergent behaviors in the ADS (e.g., different decisions whether to brake to steer to avoid a collision).

3. Results

3.1 Simulation Case Set

A total of 107 fatal collisions that occurred within Chandler, Arizona were identified from 2008 to 2017. Public records requests for the crash-related documentation were made to ADOT, the Chandler P.D, and the Arizona Dept of Public Safety for all 107 collisions. Relevant materials were available and provided for a total of 92 collisions. One of these collisions took place on a roadway that had changed substantially from the time of the crash and was excluded from further analysis due to a lack of environment for performing the simulation.

A breakdown of the collision partners in the 91 collisions available is shown in Figure 4. The plurality of collisions (28.6%) were vehicle-to-vehicle. Vehicle-to-motorcycle collisions accounted for approximately one-fourth (26.4%) of the events. Other vulnerable road user collisions in the dataset were vehicle-to-pedestrian and vehicle-to-bicyclist, which accounted for 18.7% and 4.4% of the collisions, respectively. The remaining observed collisions were single vehicle (12.1%), single motorcycle (5.5%), and vehicle-to-heavy vehicle (4.4%).

As previously discussed in the methodology section, the first two collision partners in any collision scenario were considered for counterfactual simulation. The number of roles simulated was dependent on the number of vehicles involved in the first collision event. The 91 collisions had a total of 166 collision partners (90 initiators and 76 responders) eligible for simulation. First, a total of 56 collision partners (26 initiators and 30 responders) were excluded due to failing to meet the non-towing and passenger vehicle (class-1 or class-2)

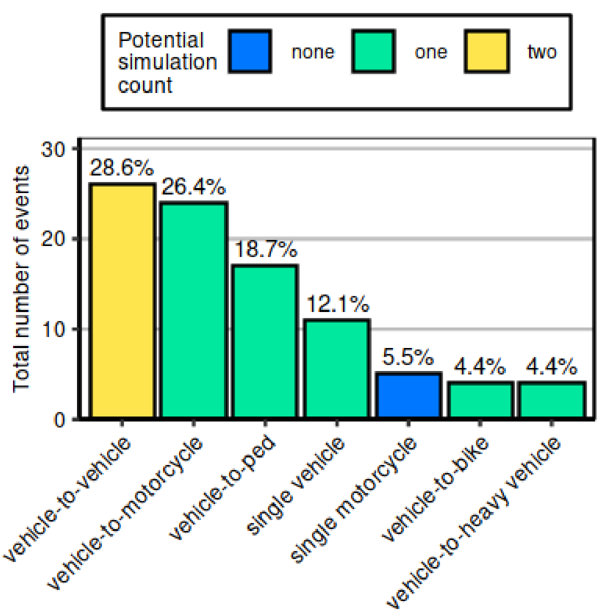

Figure 4. A breakdown of the dataset by the collision partners involved in the first collision event. The potential simulation count is defined by the number of class-1 or class-2 vehicles involved.

Table 1. High-level scenario categories used in the current study with definitions for each.

High-level Scenario	Description
Pedestrian	Involves a pedestrian actor as one of the collision partners
Cyclist	Involves a bicyclist actor as one of the collision partners
Single Vehicle	Involves a single vehicle or motorcycle actor in a collision event with an object, structure, or parked vehicle
Front-to-Rear	A primarily longitudinal event whereby some trailing vehicle or motorcycle actor approaches the rear of some lead vehicle or motorcycle actor (commonly referred to as a rear-end)
Intersection	Involves vehicle or motorcycle actors intersecting or turning into paths as a result of changing or crossing over roadways
Head-on	Involves vehicle or motorcycle actors approaching one another from opposing directions on the same trafficway (requires an actor to be moving counter to the flow of traffic)

requirements. Parties that were not replaced with the Waymo Driver included passenger vehicles with trailers (2), motorcyclists (29), pedestrians (17), bicyclists (4), and heavy vehicles (4). Second, an additional 19 collision partners (12 initiators and 7 responders) were removed due to the aforementioned roadgraph inclusion criteria (i.e., a drivable, unchanged roadgraph since the time of the crash must have been available; 45-mph roadway or under). The final dataset consisted of a total of 72 crashes and 91 vehicle actors (52 initiators and 39 responders) for simulations.

3.2 Safety Impact Assessment

This study's results were analyzed and categorized with respect to several high-level scenario contexts. A general summary of these categories is shown in Table 1. These categories are used to define the first event in the collision sequence, and are not intended to describe subsequent collision events that may have taken place. All of these scenarios are inspired by and based on definitions commonly used in the industry (Najm et al., 2007; Radja et al., 2019).

Figure 5 shows the crash outcomes after simulation with the Waymo Driver by initiator and responder role. All (52 actors; 100%) of the simulations in which the simulated Waymo Driver replaced the initiator role resulted in potential navigation of the scenario without a collision. Conversely, when placed into the responder role (39 actors), the simulated Waymo Driver was able to potentially avoid 82% (32 actors) of collisions and mitigate an additional 10% (4 actors) of collisions. The remaining 8% (3 actors) of responder scenarios resulted in an unchanged collision, all of which were in the front-to-rear struck mode. None of the simulated scenarios resulted in the ADS vehicle exacerbating the collision severity.

Intersection

Intersection collisions were the most common fatal collision scenario in the simulation case set. Three unique intersection scenario types were observed: straight crossing path (SCP), left turn across path / lateral direction (LTAP/LD), and left turn across path / opposite direction (LTAP/OD). These three scenarios are estimated to account for 73% of all intersection crashes and 93% of fatal intersection crashes annually within the United States (Scanlon 2017). A depiction of these three variations can be found in Figure 6.

As stated above, when the Waymo Driver was simulated within the initiator role, all collisions were avoided. All initiators in these intersection collision scenarios were either (a) traveling through the intersection on a red light or (b) making an unprotected crossing maneuver (straight traveling or left turn). A breakdown of scenario frequency by type can be found in Figure 7. The five SCP initiator scenarios were split between those occurring at signalized (four) and stop sign-controlled (one actor) intersections. All of the signalized intersection SCP initiator role simulations were avoided by not running the red light. The single remaining SCP initiator case occurred at a two-way stop-sign controlled intersection, and the Waymo Driver avoided any collision by simply yielding for the stop sign and executing proper gap selection when performing an unprotected straight maneuver. The eight

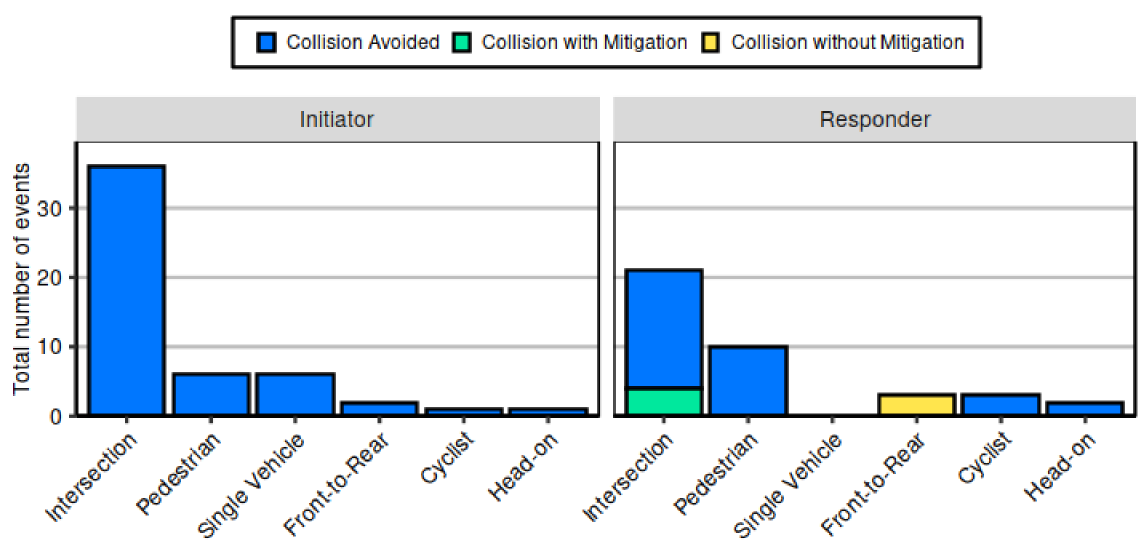

Figure 5. Simulated crash outcomes with Waymo Driver replacing humans in 72 fatal crashes (52 responder and 39 initiator roles).

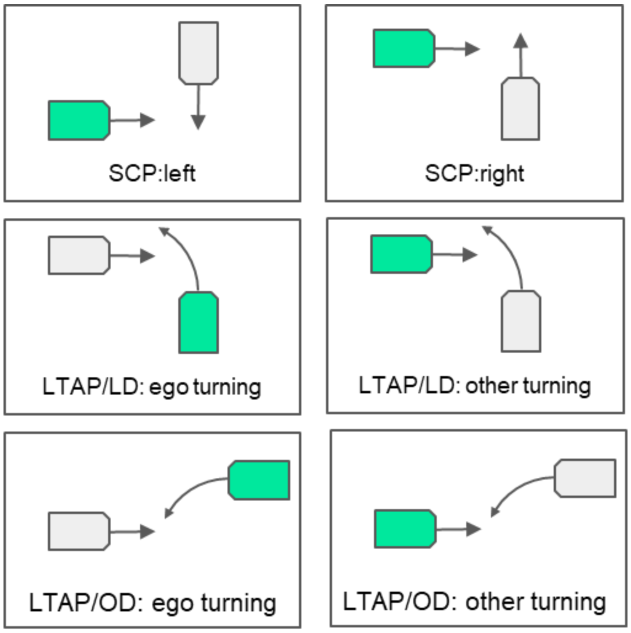

Figure 6. A depiction of all intersection scenario types observed in the current study. The ego vehicle (green) is the role being referenced in each category and refers to the Waymo Driver in simulated outcomes. The other vehicle (grey) is the collision partner within the scenario.

LTAP/LD initiator scenarios were evenly divided between simulations with the initiator turning left and simulations with the initiator going straight. All four LTAP/LD initiator turning simulations involved the vehicle performing an unprotected left turn from a two-way stop-controlled intersection. The four initiators going straight LTAP/LD simulations occured at signalized intersections and involved the initiator running a red light. The LTAP/OD scenario with the initiator role turning was the most frequently simulated scenario and generally involved the Waymo Driver needing to execute proper gap selection during an unprotected left turn. There were four LTAP/OD initiator scenarios with the other vehicle turning, and all of these simulations were the result of the initiator running a red light.

Responder role intersection scenarios were successfully avoided 81% of the time (17 actors) and the remaining 19% of simulations resulted in a mitigated collision (4 actors). Figure 8 provides a breakdown of the results by the scenario type. The majority (5 out of 6 actors) of SCP scenarios involved a red light runner initiator, and the remaining simulations involved the initiator making an unprotected straight crossing maneuver from a stop sign. All six SCP responder simulations were successfully avoided by the Waymo Driver. The LTAP/LD scenario type had six total simulations (four turning left and two traveling straight). The four LTAP/LD responder turning simulations required the Waymo Driver to avoid a red light runner approaching from the left, which the ADS successfully avoided the collision in all simulations. The two LTAP/LD scenario simulations with the responder going straight required the Waymo Driver to avoid an initiator that made an unprotected left turn, which

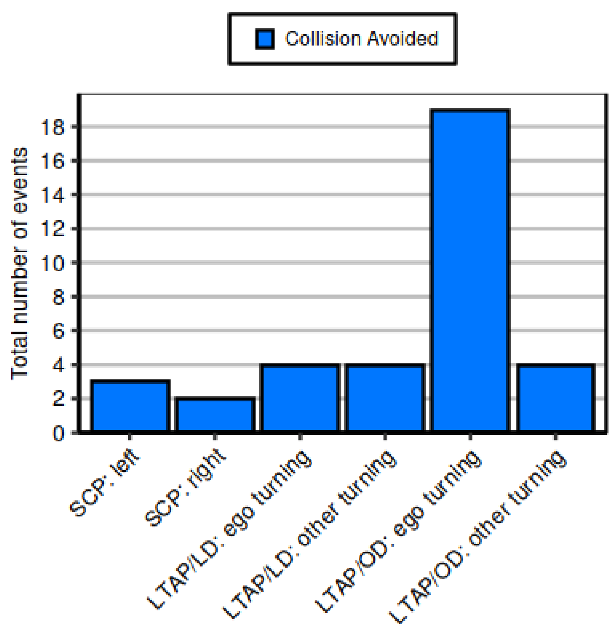

Figure 7. Outcome by scenario type for intersection initiator role simulations.

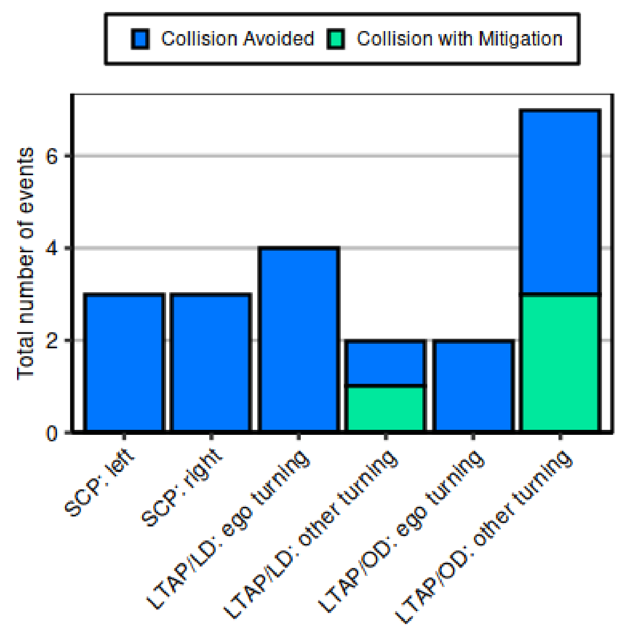

Figure 8. Outcome by scenario type for intersection responder simulations.

the ADS successfully did for one simulations and mitigated the collision in the second simulation. The majority (7 out of 9 actors) of LTAP/OD responder simulations involved the responder going straight. These simulations generally required the Waymo Driver to avoid a turning vehicle that made an unprotected left turn, and the ADS was successful in avoiding the initiator the majority of simulations (57%; 4 out of 7 actors). The remaining two LTAP/OD responder simulations involved the Waymo Driver making a protected left turn when the straight traveling initiator ran a red light. The Waymo Driver successfully avoided a collision in both of these simulations.

Table 2 highlights the reduction in maximum, crash-level risk of a serious (C-MAIS3+) outcome and the responder's impact speeds from the four simulations performed with the Waymo Driver in the responder role which were mitigated. To restate, C-MAIS3+ considers the probability of serious injury to the occupants in both vehicles and is the maximum computed risk. All four cases involved the initiator making an unprotected left turn across the path of the Waymo Driver. The initiator in the single LTAP/LD simulation (case 1) was a motorcyclist, while the initiators in the remaining three cases were passenger vehicles. The probability of serious injury was reduced in these cases by 1.3 times to 15 times in the Waymo Simulated outcome when compared to the original collision. The degree of responder impact speed reduction from the original reconstruction (human driver) to the simulation (Waymo Driver) ranged from 19% to 66%.

Table 2. Comparison of Injury Risk in Original and Waymo Driver Replacement Simulations in the Responder Role. The impact speed of the responder human-driven vehicle in the original case and of the Waymo Driver in the simulated case are presented in parentheses.

Mitigated Case	C-MAIS3+ Original Case (Impact speed)	C-MAIS3+ Waymo Simulated (Impact Speed)
Case 1 (LTAP/LD: other turning)	48% (43 mph)	37% (35 mph)
Case 2 (LTAP/OD: other turning)	29% (51 mph)	12% (29 mph)
Case 3 (LTAP/OD: other turning)	11% (39 mph)	0.7% (13 mph)
Case 4 (LTAP/OD: other turning)	29% (42 mph)	13% (24 mph)

As detailed in the methods section, these injury risk estimates are made using statistical models developed from crash data. Fatality risk is difficult to model because fatalities are rare events that often have unique or extreme causes. Even though some models take into account occupant age, underlying health issues before a crash can greatly influence the mortality of occupants in crashes. The advantage of using these injury risk estimates, however, is they make it easier to compare the potential effect of a system across several crash modes in safety impact study like this one. For example, comparing impact speeds or delta-V in different crash modes is not a fair comparison because injury risk is higher in side impacts than in frontal impacts for an equivalent delta-V.

Single Vehicle

Approximately 7% (6 actors) of the simulations in the case set were single vehicle collisions. In all six of these collisions, the simulated role was the initiator. All six of these simulated roles involved the vehicle drifting off the road. In all of these single vehicle simulated scenarios, successful avoidance of a collision by the simulated Waymo Driver was a result of maintaining vehicle control within the designated lane of travel.

Pedestrian

Pedestrian collisions were the second most (18%; 16 actors) common scenario type in the simulation case set. Every initiator role simulated (6 actors) was avoided by the simulated Waymo Driver. Of those simulations, two involved the initiator drifting out of lane and striking a pedestrian, and the Waymo Driver successfully avoided a collision in each simulation by simply maintaining vehicle control within the lane of travel. Three scenarios occurred at intersections and were prevented by yielding the right-of-way (2 actors) and not running a red light (1 actor). A final simulation involved a skateboarding pedestrian holding onto the back of a vehicle as it sped forward. The Waymo Driver avoided this scenario by not proceeding in the presence of the pedestrian.

Most (10 actors; 63% of total) pedestrian simulations within the case set were in the responder role. All of these responder role pedestrian simulations were successfully avoided. A pictorial of these events by scenario type can be found in Figure 9. The majority (9 actors) of these simulations were straight crossing path pedestrian collisions. During all of these scenarios, the simulated vehicle was traveling straight and a pedestrian traveled perpendicularly across the path of

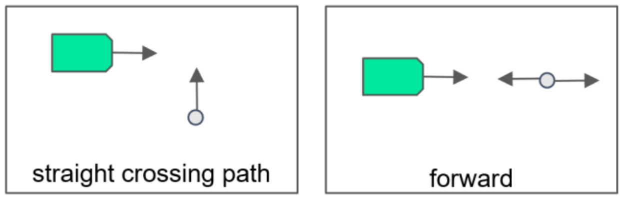

Figure 9. Responder pedestrian scenarios present in the simulation case set.

the vehicle without right of way. All of these scenarios involved the pedestrian crossing the street outside of any designated crosswalk, or crossing at an intersection crosswalk while the oncoming traffic had a green light. The final simulation involved a pedestrian in a travel lane, which is referred to as a forward event. This forward responder simulation involved a pedestrian that was jogging on a 45-mph road in the same direction as traffic.

Front-to-Rear

Front-to-rear collision scenarios accounted for 5% (5 actors) of the total simulations. All (2 actors) of the simulated initiator roles were avoided. The initiator vehicle was always the trailing (striking) vehicle in the front-to-rear configuration. Both of those simulations required the Waymo Driver to maintain spatiotemporal boundaries behind a stopped lead vehicle at a red traffic signal.

All (3 actors) of the simulated rear-end scenarios were unchanged when the Waymo Driver replaced the responder, that is, the collision configuration in the simulated role was identical to the original scenario. All of these simulations had the responder in the lead vehicle (struck) position. One of the scenarios involved the responder being stopped at a red light before being struck from behind. One simulation involved the responder being struck as it began to accelerate following a red-to-green traffic signal phase change. The final scenario involved the responder being struck from behind while traveling on a roadway at a constant speed. In all of these scenarios, the Waymo Driver behaved similarly to the original human driver.

Cyclist

There were four simulations involving a cyclist. All of these simulations were successfully avoided. One simulated perspective was in the initiator role and involved the original vehicle drifting from the lane of travel and striking a cyclist within a bike lane. Successful avoidance in this simulation was achieved through maintaining vehicle control within the lane of travel. The remaining three simulations were in the responder role. Two responder role simulations involved a cyclist making an unprotected straight crossing path maneuver across the path of the ego vehicle. The final responder simulation involved the vehicle needing to yield when making a right turn for a cyclist traveling in the wrong direction.

Head-On

Three head-on actors were included in the simulation case set. The single initiator simulation involved a vehicle that lost control and traveled out of lane into oncoming traffic. Avoiding this collision in simulation required the Waymo Driver to maintain vehicle control within the lane of travel. The two responder role head-on scenarios required the Waymo Driver to respond to a vehicle moving laterally from the oncoming lane into the lane of travel of the simulated role. One of those simulated roles was avoided by braking and allowing the initiator to pass in front, and the second scenario was avoided by nudging the vehicle within the lane of travel to the right and allowing the oncoming vehicle to pass behind.

4. Discussion

4.1 Simulated Safety Impact Performance

This study evaluated the performance of the simulated Waymo Driver on a historical set of fatal collisions experienced by human drivers in Waymo's current commercial fully autonomous ODD. The methodology entailed simulating the sequence of events leading up to the collision with the Waymo Driver individually replacing each of the class 1 and class 2 (light) vehicle actors in the collision event. The goal of this exercise was to use counterfactual "what-if" simulations to evaluate the ability of the Waymo driver to mitigate and prevent these real-world, ODD-specific fatal collisions.

Avoided Collisions

Within any driving scenario, conflict avoidance is the first line of defense for preventing and mitigating a potential collision. As discussed previously, conflict avoidance behavior in this study is defined as vehicle control to avoid a conflict which does not require urgent, evasive maneuvers (i.e., hard braking or steering). Collision avoidance is vehicle control that does require evasive maneuvers to avoid a collision In this study, we defined an urgent, evasive maneuver as braking at a deceleration greater than 0.5 g or steering with a vehicle yaw rate above 8.3 degrees/second.

These thresholds were taken from collision avoidance maneuvers from real-world intersection crashes with vehicles instrumented with event data recorders with available pre-crash information (Scanlon et al., 2015).

Replacing the original initiator with the ADS evaluated the ADS's ability to avoid a potential conflict. In each of the initiator simulations, the ADS was successfully able to avoid the conflict altogether by successfully executing one of a few key conflict avoidance behaviors, such as following road rules, maintaining safe time gaps and not surprising other road users. An urgent, evasive maneuver was neither required nor performed by the Waymo Driver in these simulations. This is in contrast to the original human initiators inducing the fatal collisions in the current study, who either took no avoidance action or had to perform collision avoidance action.

Within the responder role, the Waymo Driver successfully avoided four out of every five (82%) collisions. This study found that in 62.5% of successfully avoided responder role simulations the Waymo Driver avoided a collision by using conflict avoidance behaviors without requiring any urgent, evasive maneuver. The remaining 37.5% of avoided responder collision scenarios required some urgent, evasive maneuver to fully avoid contact. The conditions in the residual mitigated and unaffected collisions are discussed in the next section.

Residual Collisions

Residual collisions persisted in seven simulations when replacing the responding human driver with the Waymo Driver. With the exception of the simulated rear-end struck front-to-rear simulations, all simulated Waymo-responder scenarios that resulted in a collision had some evasive action by the Waymo Driver. In contrast, human drivers in many real-world collisions have been observed to take little or no avoidance maneuvers prior to the collision (Kusano & Gabler, 2013; Scanlon, 2017; Scanlon et. al., 2015), which may be due to distraction, impairment, and/or drowsiness (National Center for Statistics and Analysis, 2020; Singh, 2015). Unavoided simulated collisions occurred in three scenarios: Front-to-Rear, Intersection, and Head-On.

All three of the simulated responder role front-to-rear collisions were unchanged. All of these scenarios involved the initiating vehicle approaching from the rear. The (responding) lead vehicle was stopped at a traffic light in one scenario, just beginning to accelerate at a traffic light (following a red-to-green phase change) in a second scenario, and was traveling at constant speed for the remaining scenario. These front-to-rear responder roles offer little to no opportunity for conflict or collision avoidance, especially for those where the responding vehicle is stopped.

All (four simulations) of the residual intersection collisions were mitigated by the Waymo Driver. All of these scenarios were the result of a vehicle making an unprotected left turn across the path of the straight traveling Waymo Driver. In one of those cases, the Waymo Driver was approaching from the left, and in the other three cases, the Waymo Driver was approaching from the opposite direction. Occupants involved in the original collision configurations were 1.3 to 15 times more likely to sustain a serious injury than in the Waymo Driver simulated outcome. In these mitigated scenarios, the inability to fully avoid the collision was a result of limited available time for evasive action. The time window for collision avoidance in all of these scenarios was a function of the initiating vehicle turning in front of the Waymo Driver at the last moment and the need for the responding Waymo Driver to both perceive the left turning kinematics and execute appropriate action. Off-road sight obstructions did not play a role in the ability of the Waymo Driver to react to these left turn across path mitigated collisions. The state of the surrounding traffic was generally unknown, therefore it is unknown whether or how on-road sight obstructions affected the visibility in these crashes.

In concurrence with historical ADS driving data (Schwall et al., 2020), the current study's results demonstrate that residual collisions due to deviations in human driving behavior are expected in a mixed fleet of ADS vehicles and human drivers. The simulated driving performance of the Waymo Driver in this study suggests that mitigation is possible in many real-world collisions in response to human driving deviations and errors. This study shows that one source of this benefit is attained through a combination of conflict and collision avoidance. In some responder role simulations, conflict avoidance completely avoided a collision. In other responder role simulations, collision avoidance was employed to either avoid or mitigate the severity of a collision.

High-Severity Collision Rarity and Scenario-Based Testing Methodologies

High-severity collisions are rarely encountered during real-world driving. As discussed previously, making conclusive, retrospective statements about ADS efficacy in high-severity collisions requires many billions of miles. However, superior driving performance during high-severity collision scenarios is a key objective in ADS design.

Scenario-based testing, a core component of Waymo's Safety Methodologies for Readiness Determination (Webb et al., 2020), is one way to supplement insights from other safety methodologies, such as real-world driving miles and hazard analysis. Scenario-based testing relies on targeted evaluations of the ADS within challenging driving scenarios that could generate a collision. Waymo has developed its Collision Avoidance Testing program using available human data (naturalistic driving studies, crash databases), as well as with challenging events encountered during ADS testing. By artificially placing the Waymo vehicle into high-severity, human-induced collision scenarios in the current study, these results serve as a demonstration of the broader scenario-based testing program by evaluating performance on a targeted, representative sample of human-involved fatality crashes.

Although promising and positive, the results and methodology presented in this paper are not all-encompassing with respect to the Waymo Driver's performance in high-severity collisions. First, this study focused on a subset of high-severity collisions. There are many factors that can influence whether a serious collision results in a fatality, such as the underlying health of the occupants involved before the collision and vehicle structural and restraint performance. We expect there are similar potential benefits for ADS in other high severity collision populations (e.g., serious injury), but these benefits were not studied here as this study focused on fatal collisions. Other high severity assessments are included in scenario-based verification methodologies employed by Waymo (Webb et al., 2020). Second, this study presents the performance of the Waymo Driver in collision situations initiated by humans, but not those collisions that could potentially be initiated by the ADS. Through several other safety methodologies, Waymo minimizes the risk of an ADS failure causing a collision (Webb et al., 2020). Hazard analysis systematically analyzes risk of system failure and helps to ensure mitigations are in place for potential failures. Simulated deployments utilize a combination of on-road testing and re-simulation of historical logs to find situations that may lead to ADS-specific failures. The Collision Avoidance Testing described above also serves as a continuous evaluation of the Waymo Driver's ability to avoid collisions.

Data Representativeness

As noted in the introduction, one motivation for this study was to use a sample of collisions from the intended geographic ODD of the ADS. This ODD-specific census of fatal crashes alleviates the challenge of weighting or subsetting nationally representative crash data, which may have an unknown effect on the precision of the safety impact estimates. Additionally, by utilizing real-world fatal collisions from the ODD, the dataset captures various region-specific, relevant factors for ADS performance, such as causative mechanisms and roadway geometries. As noted, a widely-accepted methodology for effectively constructing nationally representative sets that correspond to a specific ODD is still an area of ongoing research. Furthermore, large, nationally-representative samples of fatal crash reconstructions do not generally exist. Therefore, the authors believe this study's approach contributes to the current state of the art for ODD-specific, high severity assessment of an ADS.

4.2 Limitations and Future Challenges

Human-Induced Collision Datasets for Counterfactual Simulation

Real-world human collision datasets, and particularly high-severity collisions, are a valuable tool for counterfactual simulation. Real-world collisions provide examples demonstrating where human drivers have failed to prevent and mitigate collisions. High severity, human-induced collisions provide descriptions of the types of collisions that lead to severe injuries and help provide visibility into the nature of these rare events. As human drivers will continue to initiate real-world collisions, the ADS will continue to encounter these collisions due to human performance deviations. Early evidence suggests that the Waymo Driver's collisions may be predominantly within the responder role to these human performance deviations (Schwall et al., 2020).

In line with Webb et al. (2020) and Schwall et al. (2020), we recognize this approach of simulating human-induced collisions is not sufficient *by itself* for demonstrating safety readiness across all possible conflict scenarios that the ADS may experience. The collisions experienced by a future deployed ADS can be thought of as two sets of collisions. The first are those collisions initiated by some other party or entity, where the Waymo Driver is in the responder role. As stated, this current dataset provides some evidence of performance within this category. That is, many of these responder role human collision scenarios will persist in a widely deployed ADS fleet. However, there is additional potential for novel responder role scenarios to appear in a future ADS fleet. These novel scenarios could be a function of a number of modalities, including, but not limited to, human drivers misinterpreting the ADS intentions, or drivers behaving differently around ADS vehicles. Additionally, not all human responder collision scenarios would persist in future ADS collision scenarios. Specifically, the actions by the initiator, such as the decision to make an unprotected turn, may have been a function of the responder parties behavior, such as traveling over the posted speed limit.

The second potential set of collisions are those initiated by the ADS. The potential failure modes of an ADS may be different than that of a human driver. Real-world human collisions provide examples of how humans currently initiate conflicts, which is useful for designing scenario coverage, but this alone does not indicate all modalities where an ADS may induce a collision when deployed. Waymo has multiple safety methodologies for uncovering potential conditions for ADS-initiated collisions (Webb et al., 2020). This includes, but is not limited to, a comprehensive hazard analysis, hardware layer fault detection and response, cybersecurity countermeasures, large-scale simulated (virtual) deployment of the vehicle over billions of miles, and a gradual approach beginning with trained vehicle operators to safely deploy the ADS into the fleet.

In addition to these potentially unique ADS-initiated scenarios, this study does not address the variability in ADS performance based on perception in the simulated human-induced crashes. The sensor simulation used in this study attempts to provide a realistic model of the sensors field of view, latencies, and classification performance. The sensor simulation is validated against its ability to reproduce the sensor performance observed in actual recorded sensor data. This sensor simulation is designed to produce the nominal or mean performance of the sensors. Future work should take into account the distribution of potential sensor performance which could affect perception performance and the behavior of the ADS. Our assumption is that variability of the ADS performance is narrower (less variability) than that of human drivers, given that the ADS cannot engage in the kinds of human activity (e.g., distraction, drowsiness, and intoxication) that results in degraded driving performance.

Implications of Responder Behavior in Simulated Collisions

With 92% of the simulated responder role collisions potentially avoided (82%) or mitigated (10%), an implication of the results is that the human driver in the actual responder role may have missed an opportunity to proactively respond to the actions of the initiator vehicle. This is indicated by the finding that 61% of avoided responder role simulations occurred without ADS collision avoidance action (as defined in previous sections).

One hypothesis that may explain the lack of proactive response in the human responder role (the original cases) is that there may have been an element of inattentive or distracted human driving by the responding parties in these fatal collisions. This would lead to a reduced likelihood of the responder taking proactive conflict avoidance action. While this can be studied for minor/moderate level collision scenarios (using video monitoring of human responder) in available naturalistic driving datasets (e.g. SHRP2), fatal collisions are exceedingly rare, and a tremendous number of naturalistic driving miles would be required to confirm this hypothesis. Another explanation could lie in any inaccuracies in the reconstructed trajectories, a known limitation of reconstructing collisions using only police reported information. Because the presence of other road users (to the extent there were any) are, for the most part, not captured in the simulations, the effect of other traffic or vehicle occlusions on the behavior of the ADS was not taken into account in the simulations. The effect of this deviation on the presented results can be the focus of future sensitivities studies.

Data Quality for Reconstruction

Counterfactual ADAS and ADS simulation for purposes of prospective safety benefits estimation has historically been executed using a wide range of available data sources. Each case within each of these data sources comes with its own unique set of crash-related documentation materials that directly influences the accuracy of the collision reconstruction. Any inaccuracies in the reconstruction may have an effect on the estimated safety impact of the ADS.

This study's approach relied on a single representation of the collision events based on crash reconstruction. The one exception to this was collisions involving unknown traffic light phase timing during which a wide spectrum of timings was incrementally explored in simulation. This single representation of a collision event captured many nuances of the scenario, including roadway geometry, sight obstructions, and traveling speeds. Overall, an ODD-specific baseline comparison between the original vehicle operator and the Waymo Driver was achieved, which was the primary objective of the current study.

There is, of course, some uncertainty in the reconstruction of each collision. In the absence of high-fidelity sensing systems, such as a full ADS sensor suite or birds-eye cameras, variability exists within various elements of the collision reconstruction, including, but not limited to, party kinematics, party dimensions/inertial attributes, and the presence of other uninvolved road users. Although many previous studies have relied on single representations of collision scenarios for counterfactual simulation (Gruber et al., 2019; Hamdane et al., 2015), one methodology for enhancing confidence in ADS performance in the scenario is to vary uncertain conditions in a probabilistic manner (Bareiss et al., 2019; Scanlon et al., 2017; Schachner et al., 2020). This approach allows the uncertain components of the collision reconstruction to be probabilistically resampled in a way that captures the collision scenario as it actually occurred by simply exploring the spectrum of ways that the scenario could have occurred. The challenge with this approach is that the parameter space for varying reconstruction representations is scenario-specific and can have many degrees of freedom. A single representation technique was used in this study in order to lay the groundwork for ODD-specific ADS counterfactual simulation. However, future work will consider uncertainty modeling techniques as a component of this simulation protocol. A second approach is to use the collision reconstruction as a "seed" for building scenario test sets (Menzel et al., 2018). This methodology uses the reconstructed scenario as a base condition and varies components of the reconstruction to explore ways that the scenario may have played out differently. Scenario-inspired approaches are a core component of Waymo's Collision Avoidance Testing Program (Webb et al., 2020).

Target Performance Requirements

This study evaluated the simulated performance of the Waymo Driver in comparison to the original human actor involved in the scenario. The methods accomplished this study's research objective of evaluating the performance of the Waymo Driver in real-world, high-severity collision scenarios. Evaluating an ADS with respect to the actual involved human actor gives insight into the real-world safety impact of that technology, that is, the potential effectiveness of the technology on overall crash and injury prevention if deployed widely into the population of the studied ODD (Najm et. al. 2000). A future challenge, which was outside the scope of the current study, is how ADS evaluators may additionally be interested in the performance with respect to some other target baseline performance. One example of this technique is in Waymo's Safety Methodologies and Safety Readiness Determinations, where the Waymo Driver is evaluated against the performance of a simulated reference agent as a part of the scenario-based testing programs (Webb et al., 2020). This reference model, which is empirically-defined from naturalistic driving data, serves to evaluate our performance criteria to advance the overall safety of the ADS by providing a comparison to human performance.

Another limitation of the current study is that the vehicle fleets in the crash dataset taken from years 2008 to 2017 do not have wide adoption of ADAS, such as Automated Emergency Braking (AEB). As future vehicle fleets evolve, more vehicles will have ADAS. This future vehicle fleet may have different crash characteristics than the populations studied in this retrospective crash data. Future work could consider a baseline that also includes a reference human driver with an ADAS equipped vehicle.

Collision Mitigation Effect

In this study we did not rigorously explore the safety impact of the ADS and collision partner vehicle in the event that a collision does occur. Future studies may analyze factors such as vehicle mass, restraint use, occupant positions, and crashworthiness, which can serve to further influence injury risk.

Sensitivity of Crash Avoidance Result to Variability

The crash avoidance results found during these simulations could be sensitive to variability in the system or environment, especially for intersection collisions in the responder role. In these intersection responder role simulations, such as the straight crossing path and left turn across path scenarios, the initiator enters into a conflict with the responder after violating a traffic signal or making an unprotected crossing maneuver. These scenarios tend to have minimal spatial and temporal opportunity for the responder to react. Variability, such as sensor noise, variation in the relative positions and speeds of the actors, and human perception-reactions could change the outcomes in the simulations presented in this study. For example, some intersection crashes that were avoided had small spatiotemporal margins. Alternatively, some scenarios that were mitigated may have been avoided under slightly different conditions. Given the potential variabilities discussed above, some scenarios are close to the boundary of collision and near-miss. Modeling these sources of variability is an ongoing research effort. As with other sources of uncertainty, such as the reconstruction uncertainty discussed previously, the sensitivity of these safety impact methodologies should be considered in future work.

As mentioned previously, the algorithms and simulation environment can exhibit non-determinism. The simulations in this study, however, were not greatly affected by this non-determinism. In the repeated simulations, no scenarios had varying outcomes. For example, there was no scenario where some simulations resulted in a crash, while other repetitions did not. For those simulations that did result in a collision, the injury risk predictions did not vary greatly. Of the 4 responder role simulations that resulted in mitigated collisions, there was almost no discernible change in injury risk based on the 5 repeated simulations (less than 10^{-5} percent difference). The observed variability between repeated simulations is small and the approach taken in this study to use the worst case simulation outcome is conservative.

Generalizability and Validity of Results

As discussed in the previous sections, there are methodological limitations that will introduce variability of the benefits estimates presented in this study. In this previous discussion we proposed some methodological improvements to better quantify this variability (e.g., parameterization of possible collision scenarios, comparison with additional human baselines). In this section, we discuss the generalizability and validity of the current study's results given this uncertainty.

Conflict avoidance, without evasive maneuvers, was responsible for most of the benefits observed in this study. The results showed that 100% of initiator role and 82% of responder role collisions were avoided by the simulated Waymo Driver. In these avoided collisions, there was no collision avoidance (i.e., evasive maneuvers) required in 100% of initiator role and 63% of responder role simulations. The ability of the ADS to avoid a potential collision with conflict avoidance alone suggests that the ADS was able to perceive and react in a timely manner to the emergence of the conflict partner without ever needing to take crash avoidance action. Given this conflict avoidance result, even if the previously discussed limiting factors of the methodology introduced some variability in the available time window for avoidance action, there may still be an opportunity for conflict and collision avoidance or mitigation if scenario conditions were more or less favorable. One measure of safety with high validity is real-world driving on public roads. As noted before, prospective, counterfactual approaches like the one used in this study are important tools to help assess ADS performance in fatal collisions, which would otherwise take many billions of real-world driving miles to start to evaluate in a statistically significant way. Although future methodological improvements will enhance the precision and accuracy of the safety benefit estimates, the results of the current study demonstrates the potential for ADS systems, like the Waymo Driver, to improve traffic safety outcomes in fatal collisions involving human drivers by practicing conflict avoidance.

5. Conclusion

This study estimated the safety impact of the Waymo Driver within a census dataset over a 10 year period of real-world reconstructed fatal collision scenarios. All scenarios came from Waymo's current commercial fully automated geographic ODD in Phoenix's East Valley. Counterfactual "what-if" simulations were performed to evaluate the simulated performance of the Waymo Driver within these scenarios. Up to two collision partners were simulated independently to evaluate the ADS's capacity to (a) avoid initiating the collision scenario and (b) respond to the actions of an initiating actor.

The simulated Waymo Driver prevented the initiation of every fatal collision in the dataset without performing urgent evasive maneuvers. This result highlights the importance of the Waymo Driver's conflict avoidance capacity to comply with road rules and react properly to potential conflict partners. Although the Waymo Driver is predicted to avoid the observed human collision scenarios in the initiator role, the collection of potential failure modes for an ADS may be different than that of a human driver. In the responder role (the Waymo Driver replacing the driver reacting to the actions of the crash initiator), the system was estimated to prevent 82% of collisions and mitigate an additional 10%. The majority (63%) of the avoided responder scenarios were prevented without the need for collision avoidance action, which highlights conflict avoidance performance of the ADS, such as timely response action (e.g., gradually slowing down) to the incursion of the initiating actor. The accuracy of the predicted mitigation magnitude was generally limited by various factors, including the abruptness of the initiator's actions, sight obstructions, and/or high travel speeds by the initiator parties. The remainder of the responder scenarios (8%) were unchanged from the original scenario. Every unchanged scenario was a front-to-rear scenario with the Waymo Driver being struck from behind.

This study presents a scenario-based testing methodology for assessing ODD-specific safety performance in fatal collisions. This methodology is instructive because high-severity collisions are rare and evaluating effectiveness in these scenarios through public road driving alone is not practical given the gradual nature of ADS deployments. As noted, there are inherent uncertainties introduced by a lack of information in reconstructed crashes derived from police reported crashes that will introduce variability into these safety benefit estimates. Although future methodological improvements and sensitivity studies may serve to enhance the precision and accuracy of the safety benefit estimates, the simulated results of the current study show potential for ADS to improve traffic safety outcomes in otherwise fatal collisions involving human drivers.

6. References

Bareiss, M., Scanlon, J., Sherony, R., & Gabler, H. C. (2019). Crash and injury prevention estimates for intersection driver assistance systems in left turn across path/opposite direction crashes in the United States. *Traffic injury prevention, 20(sup1)*, S133-S138. https://doi.org/10.1080/15389588.2019.1610945

Bärgman, J., & Victor, T. (2019). Holistic assessment of driver assistance systems: how can systems be assessed with respect to how they impact glance behaviour and collision avoidance?. *IET Intelligent Transport Systems, 14(9)*, 1058-1067. DOI: https://doi.org/10.1049/iet-its.2018.5550

Baron, W., Sippl, C., Hielscher, K. S., & German, R. (2020, May). Repeatable Simulation for Highly Automated Driving Development and Testing. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*.

Blincoe, L., Miller, T. R., Zaloshnja, E., & Lawrence, B. A. (2015). *The economic and societal impact of motor vehicle crashes*, 2010 (Revised) (No. DOT HS 812 013). National Highway Traffic Safety Administration.

Blower, D., & Woodrooffe, J. (2013). Real-World Safety Effect of Roll Stability Control (No. 2013-01-2392). *SAE Technical Paper*. https://doi.org/10.4271/2013-01-2392

Brach, R. M., & Brach, M. (2011). *Vehicle Accident Analysis and Reconstruction Methods, Second Edition*. SAE International.

Cicchino, J. B. (2017). Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates. *Accident Analysis & Prevention, 99*, 142-152.

Dobberstein, J., Lich, T., Schmidt, D. (2018). *Accident data analysis-remaining accidents and crash configurations of automated vehicles in mixed traffic.* OSCCAR Project Deliverable D1.1. http://osccarproject.eu/wp-content/uploads/2020/04/OSCCAR_D_1.1.pdf

Fredriksson, R., E. Rosén and A. Kullgren (2010). Priorities of pedestrian protection—A real-life study of severe injuries and car sources. *Accident Analysis and Prevention 42(6)*, 1672-1681.

Funk, J. R., Cormier, J. M., & Gabler, H. C. (2008). Effect of delta-V errors in NASS on frontal crash risk calculations. In *Annals of Advances in Automotive Medicine/Annual Scientific Conference (Vol. 52, p. 155)*. Association for the Advancement of Automotive Medicine.

Gennarelli, T. A., & Wodzin, E. (2008). *Abbreviated injury scale 2005: update 2008*. Russ Reeder, 200.

Gordon, T., Sardar, H., Blower, D., Ljung Aust, M., Bareket, Z., Barnes, M., ... & Theander, H. (2010). *Advanced Crash Avoidance Technologies (ACAT) Program - Final Report of the Volvo-Ford-UMTRI Project: Safety Impact Methodology for Lane Departure Warning - Method Development and Estimation of Benefits.* (No. DOT HS 811 405). United States. National Highway Traffic Safety Administration.

Gruber, M., Kolk, H., Tomasch, E., Feist, F., Klug, C., Schneider, A., ... & Fredriksson, A. (2019, June). The effect of P-AEB system parameters on the effectiveness for real world pedestrian accidents. In *Proceedings of International Technical Conference on the Enhanced Safety of Vehicles (ESV)*.

Hamdane, H., Serre, T., Masson, C., & Anderson, R. (2015). Issues and challenges for pedestrian active safety systems based on real world accidents. *Accident Analysis & Prevention*, *82*, 53-60. https://doi.org/10.1016/j.aap.2015.05.014

Haus, S. H., & Gabler, H. C. (2019). The Potential for Active Safety Mitigation of US Vehicle-Bicycle Crashes. In *Proceedings of the Future Active Safety Technology Towards Zero Traffic Accidents (FAST-Zero)*. Presented at the FAST-Zero, Blacksburg, VA.

Haus, S. H., Sherony, R., & Gabler, H. C. (2019). Estimated benefit of automated emergency braking systems for vehicle–pedestrian crashes in the United States. *Traffic injury prevention, 20(sup1)*, S171-S176. https://doi.org/10.1080/15389588.2019.1602729

Isaksson-Hellman, I., & Lindman, M. (2015). Real-world performance of city safety based on Swedish insurance data. *In Proceedings of the 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV)* in Göteburg, Schweden.

ISO/TR 21974–1:2018 Naturalistic driving studies – Vocabulary – Part 1

Jermakian, J. S., & Zuby, D. S. (2011). *Primary pedestrian crash scenarios: factors relevant to the design of pedestrian detection systems*. Insurance Institute for Highway Safety, Arlington, VA. https://www.iihs.org/topics/bibliography/ref/1888

Johansson, R. (2009). Vision Zero–Implementing a policy for traffic safety. *Safety Science, 47(6)*, 826-831.

Kalra, N., & Paddock, S. M. (2016). *Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?*, RAND Corporation, Santa Monica, Calif. RR-1478-RC, 2016. As of February 02, 2021: https://www.rand.org/pubs/research_reports/RR1478.html

Koopman, P., & Wagner, M. (2016). Challenges in autonomous vehicle testing and validation. *SAE International Journal of Transportation Safety, 4(1)*, 15-24.

Kudlich, H. (1966). *Beitrag zur Mechanik des Kraftfahrzeug-Verkehrsunfalls*. Technische Hochschule Wien.

Kusano, K. D., & Gabler, H. C. (2015). Comparison of expected crash and injury reduction from production forward collision and lane departure warning systems. *Traffic injury prevention, 16(sup2)*, S109-S114. https://doi.org/10.1080/15389588.2015.1063619

Kusano, K. D., & Gabler, H. C. (2013, October). Real-world driver crash avoidance maneuvers in rear-end collisions using event data recorders. In *Proceedings of the Road Safety and Simulation International Conference* in Rome, Italy.

Kusano, K. D., & Gabler, H. C. (2012). Safety benefits of forward collision warning, brake assist, and autonomous braking systems in rear-end collisions. *IEEE Transactions on Intelligent Transportation Systems, 13(4)*, 1546-1555. https://doi.org/10.1109/TITS.2012.2191542

Lindman, M., Isaksson-Hellman, I., & Strandroth, J. (2017, September). Basic numbers needed to understand the traffic safety effect of automated cars. In *Proceedings of the IRCOBI Conference*.

Prospective Effectiveness Assessment for Road Safety. Accessed 20 January, 2020, from https://pearsinitiative.com/.

PEGASUS (n.d.). *PEGASUS Research Project*. Accessed 20 January, 2020, from https://www.pegasusprojekt.de/en/about-PEGASUS

Markkula, G., Madigan, R., Nathanael, D., Portouli, E., Lee, Y. M., Dietrich, A., ... & Merat, N. (2020). Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic. *Theoretical Issues in Ergonomics Science, 21(6)*, 728-752.

Menzel, T., Bagschik, G., & Maurer, M. (2018, June). Scenarios for development, test and validation of automated vehicles. In 2018 IEEE Intelligent Vehicles Symposium (IV) (pp. 1821-1827).

Mueller, A. S., Cicchino, J. B., & Zuby, D. S. (2020). What humanlike errors do autonomous vehicles need to avoid to maximize safety?. *Journal of safety research, 75*, 310-318.

Najm, W. G., & daSilva, M. P. (2000). Benefits estimation methodology for intelligent vehicle safety systems based on encounters with critical driving conflicts. In *ITS America 10th Annual Meeting and Exposition: Revolutionary Thinking, Real ResultsIntelligent Transportation Society of America* (ITS America).

Najm, W. G., daSilva, M. P., & Wiacek, C. J. (2000). Estimation of crash injury severity reduction for intelligent vehicle safety systems. SAE World Congress, 1859-1865. https://doi.org/10.4271/2000-01-1354

Najm, W. G., Smith, J. D., & Yanagisawa, M. (2007). *Pre-crash scenario typology for crash avoidance research.* (No. DOT-VNTSC-NHTSA-06-02). United States. National Highway Traffic Safety Administration.

National Center for Statistics and Analysis. (2020, November). *Summary of Motor Vehicle Crashes: 2018 Data. Traffic Safety Facts*. (Traffic Safety Facts. No. DOT HS 812 961). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812961

Nie, J., Li, G., Yang, J., Zhou, X., Zhang, C., Yu, X., ... & Wang, M. (2013, September). A study of injury risk of bicyclist and pedestrian in traffic accidents in Changsha of China. In *Proceedings of The 5th Expert Symposium on Accident Research (ESAR)*. Hanover, Germany.

Prasad, P., Dalmotas, D., & Chouinard, A. (2015). Side Impact Regulatory Trends, Crash Environment and Injury Risk in the USA. *Stapp Car Crash Journal*, 59.

Radja, G. A., Noh, E. Y., & Zhang, F. (2019). *Crash Investigation Sampling System 2017 Analytical User's Manual*. (No. DOT HS 812 803). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812803

Riedmaier, S., Ponn, T., Ludwig, D., Schick, B., & Diermeyer, F. (2020). Survey on Scenario-Based Safety Assessment of Automated Vehicles. *IEEE Access, 8*, 87456-87477.

Riexinger, L., Sherony, R., & Gabler, H. (2019). Has Electronic Stability Control Reduced Rollover Crashes? (No. 2019-01-1022). SAE Technical Paper. https://doi.org/10.4271/2019-01-1022

Rosén, E., Källhammer, J. E., Eriksson, D., Nentwich, M., Fredriksson, R., & Smith, K. (2010). Pedestrian injury mitigation by autonomous braking. *Accident Analysis & Prevention, 42(6)*, 1949-1957.

Rosén, E., (2013). Autonomous Emergency Braking for Vulnerable Road Users. (2013, September) *In Proceedings of the IRCOBI Conference*.

SAE J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. (2018). https://doi.org/https://doi.org/10.4271/J3016_201806

Sander, U., & Lubbe, N. (2018). Market penetration of intersection AEB: Characterizing avoided and residual straight crossing path accidents. *Accident Analysis & Prevention, 115*, 178-188.

Scanlon, J. M. (2017). Evaluating the Potential of an Intersection Driver Assistance System to Prevent US Intersection Crashes (Doctoral dissertation, Virginia Tech).

Scanlon, J. M., Kusano, K. D., & Gabler, H. C. (2015). Analysis of driver evasive maneuvering prior to intersection crashes using event data recorders. *Traffic injury prevention, 16(sup2)*, S182-S189.

Scanlon, J. M., Kusano, K. D., & Gabler, H. C. (2016). The influence of roadway characteristics on potential safety benefits of lane departure warning and prevention systems in the US vehicle fleet. In *Transportation Research Board 95th Annual Meeting* (No. 16-1893).

Scanlon, J. M., Sherony, R., & Gabler, H. C. (2017). Injury mitigation estimates for an intersection driver assistance system in straight crossing path crashes in the United States. *Traffic injury prevention, 18(sup1)*, S9-S17.

Singh, S. (2015). *Critical reasons for crashes investigated in the national motor vehicle crash causation survey.* (No. DOT HS 812 115). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115

Schwall, M., Daniel, T., Victor, T., Favaro, F., & Hohnhold, H. (2020). *Waymo Public Road Safety Performance Data*. arXiv preprint arXiv:2011.00038. https://arxiv.org/abs/2011.00038

Tingvall, C., & Haworth, N. (1999, September). Vision Zero-An ethical approach to safety and mobility. In *6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000*.

Tominaga, S., & Sakurai, M. (2002). Fundamental Analysis of Motorcyclist Injury Risk Using A Statistical Model Based on Real-world Crashes (No. 2002-32-1830). SAE Technical Paper.

Van Auken, R. M., Zellner, J. W., Chiang, D. P., Kelly, J., Silberling, J. Y., Dai, R., ... & Sugimoto, Y. (2011). *Advanced Crash Avoidance Technologies (ACAT) Program - Final Report of the Honda-DRI Team, Volume I: Executive Summary and Technical Report.* (DOT HS 811 454A). National Highway Traffic Safety Administration.

Viano, D. C., & Parenteau, C. S. (2008). Serious injury in very-low and very-high speed rear impacts (No. 2008-01-1485). SAE Technical Paper.

Webb, N., Smith, D., Ludwick, C., Victor, T., Hommes, Q., Favaro, F., ... & Daniel, T. (2020). *Waymo's Safety Methodologies and Safety Readiness Determinations*. arXiv preprint arXiv:2011.00054. https://arxiv.org/abs/2011.00054

World Health Organization. (2018). Global status report on road safety 2018: Summary (No. WHO/NMH/NVI/18.20). World Health Organization. https://www.who.int/publications/i/item/9789241565684

Yanagisawa, M., Swanson, E., Azeredo, P., & Najm, W. (2017). *Estimation of potential safety benefits for pedestrian crash avoidance/mitigation systems.* (No. DOT-VNTSC-NHTSA-15-XX). United States. National Highway Traffic Safety Administration.

