Journal of Transportation Technologies, 2024, 14, 179-194
https://www.scirp.org/journal/jtts
ISSN Online: 2160-0481
ISSN Print: 2160-0473

Evaluation of the Impact of Driver Behavior on Back of Queues Events in Work Zones Using the SHRP2 Naturalistic Driving Study Data

Shauna Hallmark[1*], Guillermo Basulto-Elias[1], Nicole Oneyear[2], Omar Smadi[1]

[1]Institute for Transportation, 2711 S. Loop Drive, Suite 4700, Ames, Iowa
[2]Iowa Division, Federal Highway Administration, Ames, Iowa
Email: *shallmar@iastate.edu, nicole.oneyear@dot.gov, smadi@iastate.edu

How to cite this paper: Hallmark, S., Basulto-Elias, G., Oneyear, N. and Smadi, O. (2024) Evaluation of the Impact of Driver Behavior on Back of Queues Events in Work Zones Using the SHRP2 Naturalistic Driving Study Data. *Journal of Transportation Technologies*, **14**, 179-194.
https://doi.org/10.4236/jtts.2024.142011

Received: December 23, 2023
Accepted: March 31, 2024
Published: April 3, 2024

Abstract

The SHRP2 Naturalistic Driving Study was used to evaluate the impact of various work zone and driver characteristics on back of queue safety critical events (crash, near-crash, or conflicts) The model included 43 SCE and 209 "normal" events which were used as controls. The traces included representing 209 unique drivers. A Mixed-Effects Logistic Regression model was developed with probability of a SCE as the response variable and driver and work zone characteristics as predictor variables. The final model indicated glances over 1 second away from the driving task and following closely increased risk of an SCE by 3.8 times and 2.9 times, respectively. Average speed was negatively correlated to crash risk. This is counterintuitive since in most cases, it is expected that higher speeds are related to back of queue crashes. However, most queues form under congested conditions. As a result, vehicles encountering a back of queue would be more likely to be traveling at lower speeds.

Keywords

Safety Critical Events, Speed, Following Closely, Glance Behavior

1. Introduction

Rear-end crashes have been noted as one of the predominant types of crashes in work zones, accounting for up to 51% of work zone crashes. FHWA [1] notes crashes in work zones are increasing with a 12% increase in fatal work zone crashes from 2020 to 2021. A number of factors contributing to rear-end crashes have been noted such as location within the work zone. Weng and Meng [2] developed rear-end crash risk models to examine the relationship between rear-end crash risk in the activity area and its contributing factors. Model results indicated that rear-end crash risk at work zone activity areas increases with heavy vehicle percentage and lane traffic flow rate. They also found the lane closest to the work area was prone to higher rear-end crash risk. Additionally, they noted the expressway work zone activity area had much larger crash risk than arterial work zone activity area. Yu *et al.* [3] found a higher severity risk for rear end work zone crash with full access control and for those that occur before the actual work zone area. They also found crashes in areas with undivided medians were more likely to be severe.

Congestion is another major contributing factor to rear end crashes. Ullman *et al.* [4] conducted an in-depth evaluation of work zone crash narratives from the Virginia DOT crash database. Almost 65% of rear-end crashes in work zones were due to slowing/stopping due to work zone presence; 12% were due to slowing/stopping for flagger, police office, or work zone traffic control; and almost 9% were due to changing lanes in work zone. The researchers also estimated that around 24% of all work zone crash types was due to stopping/slowing due to congestion.

Mekker *et al.* [5] evaluated three years of crash and crowd-sourced probe vehicle data to assess the impact of queuing versus free flow conditions. They focused on back of queue (BOQ) crashes rather than just rear-end. Results indicate commercial vehicles were involved in more than 87% of BOQ fatal crashes compared to 39% of all fatal crashes during free flow. They also found the congested crash rate was 24 times higher than the uncongested crash rate. Additionally, 90% of congestion-related crashes were for situations where queues were present for 5 minutes or longer.

Aggressive behavior has been linked to rear-end crash risk in work zones including tailgating (<2 second gap), forced merges, and speeding. A study by Rakotonirainy *et al.* [6] investigated the relationship between rear-end crashes and unsafe following behavior in Queensland, Australia. They evaluated rear-end crashes in general rather than just work zone related crashes. The researchers identified 10 rear-end crash hotspots using safety performance functions and the observed behaviors in those locations. They found tailgating (<2 second gap) occurred in 55.4% of observations.

Ullman *et al.* [7] conducted an observational study of erratic maneuvers in six work zone locations in Texas where queueing was expected to be present. They reported around 2% of observed vehicles engaged in a forced merge and around 1% had a hard braking at one site. Hard braking and forced merge events occurred at other sites, but volumes were not reported so information could not be compared across sites.

Raub *et al.* [8] analyzed patterns for 110 work zone crashes in Illinois. They reported rear-end collisions accounted for 56% of crashes in work zones, and within the work zone area they accounted for 64% of crashes. Officers were asked to comment on factors leading to the crash. Stopping or suddenly slowing was noted for 37% of work zone crashes. Following too closely was the second most cited factor (24%). Distractions in the work zone were noted for 17% of crashes. Dissanayake and Akepati [9] evaluated characteristics of work zone crashes in SWZDI states using a cross-classification method. They reported 9.7% were following too closely (all crashes not just rear-end).

Speeding was also noted as a factor in 52% of rear-end crashes by Raub *et al.* [8]. Dissanayake and Akepati [9] reported 8% and Johnson [10] reported 9% of all work zone crashes were due to speeding. A review of 2014 Fatal Accident Reporting System (FARS) data indicated that about 30 percent of all fatal crashes were speeding related, while 71.4 percent of fatal work zone crashes were speeding related. MnDOT noted the main contributing factors to severe work zone crashes were inattention/distraction (13%), failure to yield (13%), and illegal/unsafe speed (9%) [10]. Li and Bai [11] evaluated crashes in Kansas highway work zones and found 25% of fatal and 18% of injury crashes were coded as too fast for conditions or speeding main contributing factor.

Muttart *et al.* [12] reported most fatal rear-end crashes involve a following vehicle traveling 40 to 70 mph which closes on a lead vehicle with a speed differential greater than 30 mph. They also note that a following distance of less than two seconds also influence rear end crash risk.

Distraction and inattention have also been reported as contributing factors. Raub *et al.* [8] reported distractions accounted for 17% of rear-end work zone crashes and Johnson [10] reported inattention/distraction was the main contributing factor for 13% of all severe work zone crashes.

Distraction, in general, has been shown to increase crash risk. For instance, Klauer *et al.* [13] compared crash/near-crashes in the SHRP 2 data to baseline driving data and found glances of more than 2 seconds away from the roadway double the risk of a crash/near-crash. Fitch *et al.* [14] also used naturalistic driving data to assess the impact of cell phone use on driving performance and found locating/answering a cell phone increased the risk of being involved in a safety critical events (SCE) by 3.65 times and use of a handheld cell phone was associated with 1.39 times increase in SCE.

The objective of the research summarized in this paper was to evaluate back of queue safety critical events (SCE) in work zones using the SHRP2NDS to assess contributing driver and roadway factors. Back of queue scenarios were identified through a review of safety critical events in work zones coded by the Virginia Tech Transportation Institute (VTTI) (crashes, near crashes, or conflicts) as well as a review of time series traces in work zones collected for a related project. This resulted in 46 safety critical events and 283 "normal" events. A mixed Mixed-Effects Logistic Regression model was developed with odds of an SCE as the response variable.

2. Data

2.1. Source

The second Strategic Highway Research Program (SHRP2) Naturalistic Driving

Study (NDS) instrumented the vehicles of naïve drivers with an array of sensors which collected kinematic vehicle data (*i.e.*, speed, acceleration, location) as well as forward, rear, driver face and over the shoulder video views. Over 30 million data miles were recorded for 3,400 participants over the three years of the study. The study took place in Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington.

The SHRP2 Roadway Information Database (RID) was collected simultaneously with the SHRP2 NDS study. Mobile data collection was conducted along with integration of existing roadway and supplemental data acquired from public and private sources such as 511 data and traffic volume.

The Virginia Tech Transportation Institute (VTTI) and processed the SHRP 2 data. A number of crash, near-crash, and conflict events termed as safety critical events (SCE) were coded. Back of queue scenarios were identified through a review of these safety critical events. Additional SCE and baseline scenarios where a driver encountered a back of queue in a work zone were also identified through a review of time series traces in work zones collected for a related project which also utilized the SHRP 2 NDS.

In this related project, work zone times events were identified in the SHPR 2 data primarily through use of 511 data which was collected and archived in the RID. The 511 system allows drivers to receive real-time traffic information on road closures, accidents, route detours, weather alerts, etc. The 511 data for the time period coincident with the SHRP 2 NDS data collection (2011 to 2013) were queried for construction related terms such as "construction", "lane closure", "road work", "maintenance". Potential work zones were flagged and then those which were in place for more than three days retained. Three days was used as a threshold because it was unlikely that a sufficient number of NDS time series traces would be available for short-duration work zones. Only events for active work zones were included. An active work was defined as a work zone having a lane or shoulder closure which excluded work zones that may have consisted or a few construction barrels or signs. GPS position was also available which allowed the traces to be linked to corresponding roadway segments. The forward roadway video clip provided a perspective windshield view for the driver. Static driver characteristics, such as age, gender, years driving, were also available.

In the process of reviewing work zone traces, a number of events were identified where the subject driver encountered a back of queue. Each BOQ event was flagged, and additional data were reduced as described in the following section. Examples of back of queue events are shown in **Figure 1**. Back of queue events included safety critical events as well as events where the driver reacted appropriately in response to the upcoming queue (*i.e.*, slowed without a hard braking). These were termed as baseline or control events.

All potential BOQ SCE were reviewed by the team to ensure they were in a work zone, involved a scenario where the subject driver encountered an actual queue, and involved a hard deceleration. Only those that met these three criteria were included.

Figure 1. Examples of back of queue events (image source: VTTI).

2.2. Data Reduction

Various variables were reduced for each available event from both sources.

Roadway Variables

Non-work zone roadway characteristics of interest were extracted for each time series trace. When roadway characteristics could not be obtained from the roadway information database, they were extracted from Google Earth, the forward view video, or aerial images. Roadway characteristics included the following:

- Number of lanes
- Type of median
- Surface type (asphalt versus concrete)
- Shoulder type
- Speed limit
- Presence of lighting
- Number of uncontrolled intersecting roadways
- Presence and type of traffic control

Work zone configuration and characteristics were coded using the forward view video and included the following:

- Type and location of barriers
- Number of closed lanes
- Presence and type of DMS or other intelligent transportation system (ITS) countermeasures

- Presence of workers
- Presence of equipment
- Lane shifts
- Temporary pavement markings

Environmental conditions such as time of day (*i.e.*, day/night) and weather (raining/not raining) were also coded using the forward roadway view.

Driver Variables

Driver characteristics (*e.g.*, age, gender, years driving, number of violations) were included as variables. Viable events were provided to VTTI, and their analysts coded glance location and distraction. For each event, the driver's glance locations and visual distractions were and included the following:

- Forward
- Left
- Right
- Up
- Down
- Over the shoulder (not shown in the figure, but involved a glance beyond the B pillar)
- Center console
- Steering wheel
- Rear view mirror
- Other (used when blinks, squints, or closed eyes lasted more than 10 frames)
- Missing (used when the eyes were obscured or obstructed for more than 10 frames or when video was missing)

Distractions were only coded when they were associated with a glance away from the forward view. For instance, if a driver was looking forward but talking to a passenger, that was not coded as a distraction. However, if the driver looked to the right at the passenger while talking to him/her, that was coded as a distraction. Distractions were coded as follows:

- Passenger
- Route planning (locating, viewing, or operating a device)
- Moving or dropped object in vehicle
- Animal/insect in vehicle
- Cell phone (locating, viewing, or operating the device)
- IPod/MP3 player (locating, viewing, or operating the device)
- In-vehicle controls
- Drinking/eating
- Smoking
- Personal hygiene
- Other task

Due to sample size, glance location was aggregated as locations where a driver was attending to the roadway task (*i.e.*, forward, rear view mirror) or not attending to the roadway task (*i.e.*, down, back, passenger). Distractions, such as eating, drinking, texting, were coded when they were associated with a glance away from the roadway task. This method of defining distraction as a secondary task associated with a glance away from the driving task was based on the method used by VTTI as well as other researchers [15] [16]. The distraction category included any cell phone related activity where the driver was looking away from the roadway tasks. Cell phone use was identified when possible and noted. This was coded as a treated as a separate variable and did not need to be associated with a glance away from the driving task. Cell phone use included dialing, talking, texting, or handling a cell phone. Hands free cell phone use could not be detected unless it also involved some physical interaction with the phone. As a result, this definition did not include hands free cell phone activity. Distraction, glance data, and cell phone were coded for the six seconds before and six seconds after reaction time. The 6 second window was based on perception reaction time and an assessment of time needed for a driver to execute an evasive maneuver. Distraction, glance, and cell phone were joined to the corresponding time series trace using time stamps. The following variables were reduced:

- **Cellphone use:** subject driver used cellphone at any point 6 seconds before reaction time to 6 seconds after reaction time regardless of glance location
- **Glance**: Subject driver was engaged in a glance of 1 or more seconds away from the forward roadway within the period 6 seconds before reaction time to 6 seconds after reaction time.
- **Cell phone distraction**: Subject driver was engaged in a cell phone task (reaching, texting, talking) which involved a glance away from the forward roadway within the period 6 seconds before reaction time to 6 seconds after reaction time

Several additional driver/vehicle variables were recorded for each BOQ event included the following:

- **Reaction time**: time stamp where the lead vehicle began braking or slowing which suggests a need for the following (subject SHRP2 NDS driver) to also react. The lead vehicle may also have been stopped when the subject vehicle encountered the back of queue and in this case the point at which the subject vehicle would have been able to notice the queue was recorded as reaction time.
- **Incident time:** time stamp when the following subject vehicle took action in response to the lead vehicle (*i.e.*, braking, slowing)
- **Average speed:** average speed for subject vehicle 10 second prior to reaction time
- **Maximum speed:** maximum speed for subject vehicle 10 second prior to reaction time
- **STD:** standard deviation of speed for subject vehicle 10 second prior to reaction time
- **Max acceleration:** the maximum acceleration (recorded in g's) for subject vehicle 10 second prior to reaction time
- **Following:** a subjective measure of following behavior for subject vehicle

o Following closely (<2 seconds)

- Following (2 to 3 seconds)
- Not following (>3 seconds)

3. Methods

3.1. Ethics

This project included use of data from human participants. The team utilized data from the SHPR 2 data NDS. The SHRP 2 NDS data were collected by and are managed by the Virginia Tech Transportation Institute (VTTI). Informed consent was obtained under their institutional review. The team's access to data was covered by the Iowa State University Institutional Board Review Board (Study 19-176-00). Iowa State human subjects research and the activities of the IRB are guided by the ethical principles outlined in the Belmont Report, and by applicable regulations governing human subjects research. The team submitted an application to and received approval from the ISU IRB. The data received from VTTI did not contain any personally identifiable information.

3.2. Model Development

Back of queue scenarios were identified through a review of safety critical events in work zones coded by the VTTI (crashes, near crashes, or conflicts) as well as a review of time series traces in work zones collected for a related project as described in the previous section.

Safety critical events (*i.e.*, crash, near-crash, and conflicts) in the SHRP 2 data were typically classified as a deceleration of 0.5 g or higher and/or an evasive maneuver. Other deceleration thresholds were also considered for the analysis described in this paper (*i.e.*, 0.4 g). Klauer *et al.* [13] evaluated different braking thresholds in the 100 Car Naturalistic Driving Data. They categorized driving behavior as safe (−0.30 to −0.39), moderately safe (−0.40 to −0.49), and unsafe (−0.5 to −0.59). Kusano and Gabler [17] evaluated rear-end crashes from the National Automotive Sampling System / Crashworthiness Data System. They found an average deceleration of 0.52 g. Another study by Aoki *et al* [18] conducted a simulator study where volunteers were subjected to a crash situation. The result from this study also showed an average braking deceleration of 0.39 g. Several other studies have defined hard braking events as ≥ 0.45 g [19] [20] [21]. Wood and Zhang [22] defined crash and near-crash rate of 0.41 g and 0.45 g.

Since some variability existed in what researchers have categorized as the threshold between a near-crash and regular driving event, three different models were developed. The response variable was first defined as a safety critical event using the initial definition of a crash/near-crash of 0.5 g. Models were also developed using a threshold of 0.3 g or 0.4 g. This increased the sample size of cases meeting the criteria for safety critical events as well as increasing the number of predictor variables associated with the additional cases. However, use of the different thresholds resulted in similar results as the first model. Since the definition used by VTTI has been consistently utilized in SHRP 2 analyses, that definition was used and SCE were defined as a crash, an event with a deceleration of 0.5 g or higher, or an event with an evasive maneuver. This resulted in 46 safety critical events (SCE) and 283 "normal" events which were used as controls.

A review of the data indicated that several drivers were represented multiple times in both SCE and control events. This could have been accounted for using a repeated measures variable for drivers in the model. However, most drivers only had one observation and an assessment of initial model results suggested the small sample of drivers with multiple observations was skewing results. Consequently, drivers with more than 2 events were randomly sampled and only 2 events per driver were ultimately included in the model. This reduced the sample to 219 "normal" traces with 43 SCE representing 209 unique drivers.

A Mixed-Effects Logistic Regression model was developed with probability of a SCE as the response variable. Mixed-effects logistic regression models are commonly used for binary responses, often arising in transportation-related problems [23] [24]. Models with mixed effects have two components: fixed and random. The former explains the relationship between the independent variables, while the latter accounts for statistical dependence induced by elements of the same group, *e.g.*, observations from the same individuals or work zones.

Various models were tested using predictor variables which included driver age, driver gender, driver distraction ("Distraction"), cell phone use ("Cellphone), distraction involving a cell phone ("Cell Distraction"), maximum speed before reaction, average speed, roadway type, following behavior, type of work zone (*i.e.*, no closures, shoulder closure, lane closure), type of barrier (*i.e.*, concrete, barriers), and time of day.

A mixed effects logistic regression model was developed to assess the relationship between probability of an SCE and roadway, driver, and work zone characteristics. The variable Y_i was the event type for the i-th trace. For the event type model, the possible values are $Y_{ij} = 0$ if the drive had a "normal reaction" and Y_{ij} if it was a "SCE."

That is,

$$Y_i \sim Bernoulli\left(p_i\right)$$

where the probability of and SCE, p_i, is associated to the independent variables through the logit function:

$$logit\left(p_i\right) = X_i^T \beta,$$

where X_i are the covariate values, and β the fixed parameters. The logit function is defined as

$$logit\left(p\right) = \log\left(\frac{p}{1-p}\right).$$

The logit function facilitates the interpretation of the parameters β, since it represents the log-ratios. The vector β has a size of $k+1$ representing the parameter estimates for the k covariates plus the intercept estimate. If the j

-th entry represents a binary variable (*e.g.*, sex: 1 = male, 0 = female) and $\exp\left(\hat{\beta}_j\right) = 1.02$, then it means that observations with the presence of such variable are 2% more likely to have a near crash reaction.

For both models, stepwise forward selection was used. The selection criterion was the Akaike information criterion (AIC).

4. Results

A Mixed-Effects Logistic Regression model was developed with probability of a SCE as the response variable and driver and work zone characteristics as predictor variables. The final model indicated glances over 1 second away from the driving task and following closely increased risk of an SCE by 3.8 times and 2.9 times, respectively.

The final best model included whether a driver glanced away from the roadway task at least once for 1 or more seconds (Glance +1), following status (Following), and average speed (Avg_Spd) in the 6 seconds before the reaction time. The latter variable was included through a spline to allow it some flexibility. Model fit statistics are provided in **Table 1**. Model results are shown in **Table 2**.

Table 1. Anova results.

Term	F-statistic	df	p-value
Glance + 1	5.7402	1	0.0166
Following	11.0798	2	0.0039
Avg_Spd	5.0076	2	0.0818

Table 2. Model estimates.

Variable	Estimate	Std. Error	z value	Odds Ratio	Pr (>|z|)
(Intercept)	−0.7453	1.1999	−0.6212		0.5345
Glance + 1	1.3339	0.5467	2.4397	3.80	0.0147
Following	−0.3172	0.6054	−0.5239	0.73	0.6003
Following Closely	1.0698	0.5615	1.9052	2.91	0.0568
bs(before_react_avg_speed, degree = 2)^1	−1.3995	2.2424	−0.6241	0.25	0.5326
bs(before_react_avg_speed, degree = 2)^2	−2.4256	1.1141	−2.1772	0.09	0.0295

As noted in **Table 2**, involvement in an SCE was 3.8 more likely if the driver was engaged in a glance away from the roadway task of 1 or more seconds (p = 0.0147). When a driver is following closely (<2 seconds) they are 2.91 times more likely to be involved in an SCE (p = 0.0568) than when not following. Drivers following another vehicle (within 2 to 3 seconds) were less likely to be involved in an SCE, but this difference was not statistically significant (p = 0.6003) and was only included in the model since other conditions for following were included.

The average speed of the subject driver was also significant. Since the relationship is non-linear, the odd ratios cannot be interpreted directly. The relationship is shown graphically in Figure 2. As noted, drivers are more likely to be involved in a SCE at lower speeds than higher speeds. This is counterintuitive since in most cases, it is expected that higher speeds are related to back of queue crashes. However, most back of queue work zone events occur in urban areas where lower speed limits are present than for rural areas and congestion are more likely to result from work zone presence resulting in lower speeds.

It should be noted that the metric only reflected actual speed of subject vehicle. In most cases, work zone speed limit could not be determined, nor could the speed of prevailing vehicles. As a result, while speed was included in the model, speeding could not be determined.

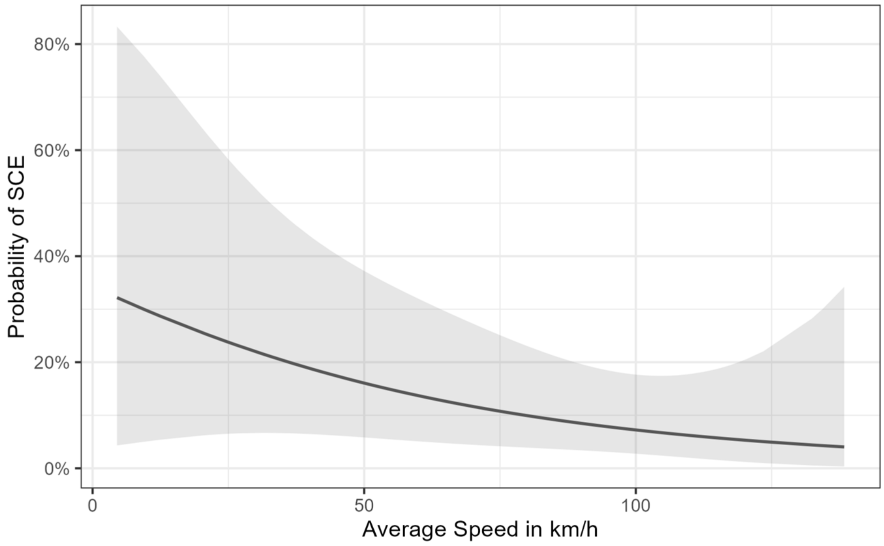

Figure 2. Relationship between average speed and probability of a BOQ safety critical event.

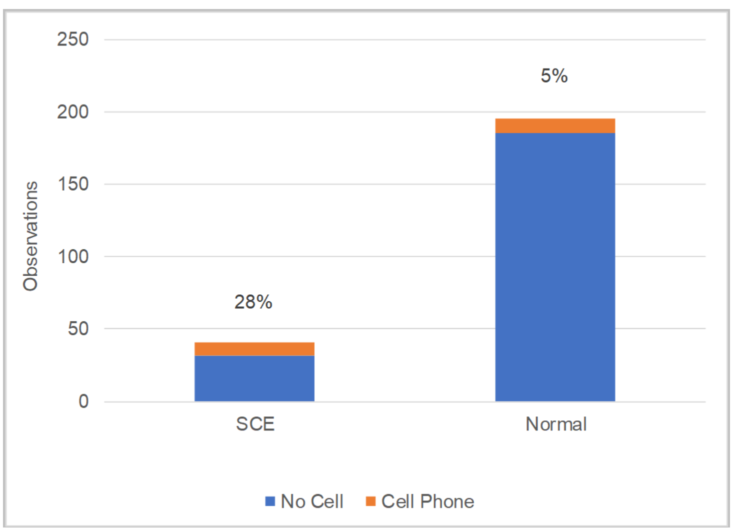

Figure 3. Relationship between cell phone use and safety critical events.

Cell phone use and cell phone distraction were not statistically significant. This may be due to small sample size. Cell phone use was only present in 42 of the 252 events. A simplistic comparison of the data indicates 28% of drivers involved in a back of queue SCE were using a cell phone compared to 5% of driver in the baseline as shown in **Figure 3**. Hence, drivers involved in an SCE were more than 5 times more likely to be engaged in a cell phone task as those involved in a "normal" work zone back of queue event.

5. Discussion

5.1. Limitations

Several limitations were present in the study. The main limitation was sample size. Several thousand traces through work zones were reviewed for a related project and when present a back of queue event was flagged. Even with this quantity of data, the number of back of queue events was small. This resulted in only slightly more than 220 back of queue events. Limited sample size may have impacted the ability to identify relationships. As noted, cell phone use was five times more likely to occur in SCE than for normal back of queue events, but the statistical significance could not be determined. Distraction may also be correlated to back of queue SCE but there were not sufficient distractions to pick up a relationship. Glances away from the forward roadway included glances away with an associated distraction as well as just glances away but the impact of distraction alone could not be confirmed.

Another limitation is that glance location in the SHRP2 data is coded from driver head position rather than use of eye tracking devices. It is only possible to identify glances to general locations rather than to specific objects. Consequently, it was not possible to determine what drivers were looking at. It would have been insightful to determine whether drivers were distracted by work zone elements, such as workers.

5.2. Practical Applications

The study findings can assist transportation agencies in addressing driver behaviors which impact back of queue conflicts. First, the study found activities which engage the driver's attention away from the roadway task for 1 or more seconds increased the likelihood of a back of queue safety critical event in work zones. Additionally, although not statistically significant, there was some evidence that cell use in use in general increased risk. The analyses also reinforce the concept that drivers engaged in glances away from the roadway tasks drove differently in work zones. Coupled with the body of work that has indicated distraction and cell phone use to negatively impact driver behavior, the efficacy of hands free or cell phone laws in work zones is reinforced.

Aggressive driver behavior, in particular speeding and following closely, have also been shown to contribute to rear-end crash risk in work zones. The role of following closely was confirmed by this study. Speeding and following closely may be addressed by queue warning systems (QWS). Other countermeasures such as dynamic speed feedback signs or enforcement may also be effective for these types of behaviors.

6. Conclusion

The main purpose of this study was to assess driver behavior as they approached back of queues in work zones. Back of queue events related to work zones were identified in the SHRP2 NDS. The advantage to the SHRP2 NDS was the ability to review driver behavior before the event. Speed, cell phone use, distraction, and glance location were extracted and included in the analysis. Work zone characteristics such as type (*i.e.*, lane or shoulder closure) and type of barrier present were also coded and used as covariates.

SCE were defined as crash, near -crash, or conflict using a threshold deceleration of 0.5 g or higher and/or an evasive maneuver occurred. Using this definition, the model included 43 SCE and 219 "normal" events which were used as controls. The traces included representing 209 unique drivers.

A Mixed-Effects Logistic Regression model was developed with probability of a SCE as the response variable and driver and work zone characteristics as predictor variables. The final model indicated glances over 1 second away from the driving task and following closely increased risk of an SCE by 3.8 times and 2.9 times, respectively. Average speed was negatively correlated to crash risk. This is counterintuitive since in most cases, it is expected that higher speeds are related to back of queue crashes. In most cases, work zone speed limit could not be determined, nor could the prevailing speed of traffic be determined. As a result, the metric for speed only indicated the speed for the subject vehicle. Whether the vehicle was over the posted work zone speed limit or was traveling too fast for prevailing conditions could not be determined. As a result, there is likely a relationship between speeding and increased work zone safety risk which could not be determined from the model.

Cell phone use was not statistically significant (likely due to sample size). However, a simplistic analysis suggested drivers engage in an SCE were more than five times more likely to be engaged in a cell phone task than drivers involved in a normal back of queue event.

Results are consistent with other studies which have found following closely [4] [6] [8] [9] as a contributor to rear-end crashes. Additionally, studies have indicated that glances away from the roadway task and cell phone use increase crash risk in general [13] [14] [25]. Human factors research in simulated work zones has shown that drivers talking even hands-free were slower to respond, narrowed their eye scanning behavior, and were less likely to check their mirrors in a lane change [26] suggesting a greater likelihood of crashes in work zones when talking on the phone, even hands-free.

No work zone characteristics (*i.e.*, type of barrier, number of lanes closed) were statistically significant in the model. While they are also likely to have an impact on likelihood of a back of queue SCE, there were likely not sufficient instances of a particular work zone characteristic to show statistical relevance.

The main contribution of this study compared to the existing literature is that information about driver behaviors could be included to assess what drivers were doing as they encountered a back of queue. For instance, the study showed that drivers glancing away from the roadway tasks or engagement in cell phone activities increased the likelihood of a rear end SCE in work zones. Most other studies that have assessed rear end crash risk in work zones have utilized crash data which do not include these driver factors. Additionally, this study quantified whether drivers were following closely and able to confirm that following too closely was a major contributing factor.

Disclosure Statement

The authors have no conflict of interest or financial interest to disclose.

Funding Statement

This research was funded by the Minnesota Department of Transportation (MnDOT) and the Federal Highway Administration Implementation Assistance Program (IAP). Results and conclusions do not necessarily represent the official views of the funding organizations.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Federal Highway Administration (FHWA) (2023) FHWA Work Zone Facts and Statistics. Washington, DC. https://ops.fhwa.dot.gov/wz/resources/facts_stats.htm

[2] Weng, J. and Meng, Q. (2011) Analysis of Driver Casualty Risk for Different Work Zone Types. *Accident Analysis and Prevention*, **43**, 1811-1817. https://doi.org/10.1016/j.aap.2011.04.016

[3] Yu, M., Zheng, C. and Ma, C. (2020) Analysis of Injury Severity of Rear-End Crashes in Work Zones: A Random Parameters Approach with Heterogeneity in Means and Variances. *Analytic Methods in Accident Research*, **27**, Article 100126. https://doi.org/10.1016/j.amar.2020.100126

[4] Ullman, G.L., Pratt, M., Fontaine, M.D., Porter, R.J. and Medina, J. (2018) NCHRP Web-Only Document 240: Analysis of Work Zone Crash Characteristics and Countermeasures. National Cooperative Highway Research Program, Washington, DC. https://doi.org/10.17226/25006

[5] Mekker, M.M., Remias, S.M., McNamara, M.L. and Bullock, D.M. (2020) Characterizing Interstate Crash Rates Based on Traffic Congestion Using Probe Vehicle Data. *Joint Transportation Research Program Affiliated Reports*, Paper 31. https://doi.org/10.5703/1288284317119

[6] Rakotonirainy, A., Demmel, S., Watson, A., Haque, M., Fleiter, J., Watson, B. and Washington, S. (2017) Prevalence and Perception of Following Too Close in Queensland. *Proceedings of the* 2017 *Australian Road Safety Conference*, Perth,

10-12 October 2017, 2p.

[7] Ullman, G.L., Fontaine, M.D., Schrock, S.D. and Wiles, P.B. (2001) A Review of Traffic Management and Enforcement Problems and Improvement Options at High-Volume, High-Speed Work Zones in Texas. Ph.D. Thesis, Texas Transportation Institute, Texas A&M University System, College Station, Texas.

[8] Raub, R.A., Sawaya, O.B., Schofer, J.L. and Ziliaskopoulos, A. (2001) Enhanced Crash Reporting to Explore Work Zone Crash Patterns. Ph.D. Thesis, Northwestern University Center for Public Safety, Evanston.

[9] Dissanayake, S. and Akepati, S.R. (2009) Identification of Work Zone Crashes Characteristics. Ph.D. Thesis, Smart Work Zone Deployment Initiative, Iowa State University, Ames. https://intrans.iastate.edu/app/uploads/2018/08/Dissanayake_WZCrashChar.pdf

[10] Johnson, C. (2015) Work Zone Crash Report. Minnesota Department of Transportation, Office of Traffic, Safety, and Technology, Saint Paul. http://www.dot.state.mn.us/trafficeng/workzone/swzsc/workzonecrashes.pdf

[11] Li, Y. and Bai, Y. (2008) Development of Crash-Severity-Index Models for the Measurement of Work Zone Risk Levels. *Accident Analysis and Prevention*, **40**, 1724-1731. https://doi.org/10.1016/j.aap.2008.06.012

[12] Muttart, J., Kuzel, M., Dinakar, S., Gernhard-Macha, S., Edewaard, D.E., Appow, S. and Dickson, C. (2021) Factors that Influence Drivers' Responses to Slower-Moving or Stopped Lead Vehicles. *SAE International Journal of Advances and Current Practices in Mobility*, **3**, 2193-2218. https://doi.org/10.4271/2021-01-0890

[13] Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D. and Ramsey, D.J. (2006) The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. National Highway Traffic Safety Administration, Washington, DC. https://doi.org/10.1037/e729262011-001

[14] Fitch, G.M., Soccolich, S.A., Guo, F., McClafferty, J., Fang, Y., Olson, R.L., Perez, M.A., Hanowski, R.J., Hankey, J.M. and Dingus, T.A. (2013) The Impact of Hand-Held and Hands-Free Cell Phone Use on Driving Performance and Safety-Critical Event Risk. National Highway Traffic Safety Administration, Washington, DC.

[15] Gao, J. and Davis, G.A. (2017) Using Naturalistic Driving Study Data to Investigate the Impact of Driver Distraction on Driver's Brake Reaction Time in Freeway Rear-End Events in Car-Following Situation. *Journal of Safety Research*, **63**, 195-204. https://doi.org/10.1016/j.jsr.2017.10.012

[16] Bamney, A., Pantangi, S.S., Jashami, H. and Savolainen, P. (2022) How Do the Type and Duration of Distraction Affect Speed Selection and Crash Risk? An Evaluation Using Naturalistic Driving Data. *Accident Analysis and Prevention*, **178**, Article 106854. https://doi.org/10.1016/j.aap.2022.106854

[17] Kusano, K.D. and Gabler, H. (2011) Method for Estimating Time to Collision at Braking in Real-World, Lead Vehicle Stopped Rear-End Crashes for Use in Pre-Crash System Design. *SAE International Journal of Passenger Cars-Mechanical Systems*, **4**, 435-443. https://saemobilus.sae.org/content/2011-01-0576/ https://doi.org/10.4271/2011-01-0576

[18] Aoki, H., Aga, M., Miichi, Y., Matsuo, Y. *et al.* (2010) Safety Impact Methodology (SIM) for Effectiveness Estimation of a Pre-Collision System (PCS) by Utilizing Driving Simulator Test and EDR Data Analysis. https://doi.org/10.4271/2010-01-1003

[19] Simons-Morton, B.G., Ouimet, M.C., Wang, J., Klauer, S.G., Lee, S.E. and Dingus,

T.A. (2009) Hard Braking Events among Novice Teenage Drivers by Passenger Characteristics. *Fifth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Big Sky, Montana, 22-25 June 2009, 236-242. https://doi.org/10.17077/drivingassessment.1327

[20] McGehee, D.V., Raby, M., Carney, C., Lee, J.D. and Reyes, M.L. (2007) Extending Parental Mentoring Using an Event-Triggered Video Intervention in Rural Teen Drivers. *Journal of Safety Research*, **38**, 215-227. https://doi.org/10.1016/j.jsr.2007.02.009

[21] Wierwille, W.W., Lee, S.E., DeHart, M. and Perel, M. (2006) Test Road Experiment on Imminent Warning Rear Lighting and Signaling. *Human Factors*, **48**, 615-626. https://doi.org/10.1518/001872006778606831

[22] Wood, J., and Zhang, S. (2017) Evaluating Relationships between Perception-Reaction Times, Emergency Deceleration Rates, and Crash Outcomes Using Naturalistic Driving Data. *Transportation Research Record*, **2675**, 213-223.

[23] Abdulhafedh, A. (2017) Road Crash Prediction Models: Different Statistical Modeling Approaches. *Journal of Transportation Technologies*, **7**, 190-205. https://doi.org/10.4236/jtts.2017.72014

[24] Washington, S., Karlaftis, M.G., Mannering, F. and Anastasopoulos, P. (2020) Statistical and Econometric Methods for Transportation Data Analysis. CRC Press, New York. https://doi.org/10.1201/9780429244018

[25] Atwood, J., Guo, F., Fitch, G. and Dingus, T.A. (2018) The Driver-Level Crash Risk Associated with Daily Cellphone Use and Cellphone Use While Driving. *Accident Analysis & Prevention*, **119**, 149-154. https://doi.org/10.1016/j.aap.2018.07.007

[26] Muttart, J.W., Fisher, D.L., Knodler, M. and Pollatsek, A. (2007) Driving without a Clue: Evaluation of Driver Simulator Performance during Hands-Free Cell Phone Operation in a Work Zone. *Transportation Research Record*, **2018**, 9-14. https://doi.org/10.3141/2018-02