Disparities in Pedestrian Crossing and Driver Yielding Behaviors: Evidence from a Large-Scale Observational Study at Urban Intersections

Anna E. Beliveau
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St, Stop C1761, Austin TX 78712, USA
Email: annabeliveau@utexas.edu

Angela J. Haddad
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St, Stop C1761, Austin TX 78712, USA
Email: angela.haddad@utexas.edu

Emily A. Podnar
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St, Stop C1761, Austin TX 78712, USA
Email: emilypodnar@utexas.edu

Devina Sharma
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St, Stop C1761, Austin TX 78712, USA
Email: devinasharma@utexas.edu

Chandra R. Bhat (corresponding author)
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St, Stop C1761, Austin TX 78712, USA
Tel: +1-512-471-4535; Email: bhat@mail.utexas.edu

ABSTRACT
Pedestrian safety remains a critical challenge in urban environments, marked by rising fatalities and persistent disparities across sociodemographic groups. Uncovering the drivers of these disparities requires a deeper understanding of both pedestrian and driver behaviors. This study examines how individual attributes, social context, and time-of-day/weather conditions shape pedestrian crossing and driver yielding decisions. We analyzed over 1,000 hours of video footage from two intersections in Austin, Texas, documenting over 20,995 pedestrian crossings and 3,124 pedestrian-vehicle interactions. Manual annotation of this footage enabled the estimation of two binary logit models: one predicting non-compliant pedestrian crossings (NCPC) and the other predicting driver unyielding (DUY) (that is, driver failure to yield to pedestrians). The results indicate that male pedestrians, Black pedestrians, those displaying visible signs of housing insecurity (VHI), and individuals crossing solo are significantly more likely to cross non-compliantly and to encounter lower driver-yielding rates. Runners also exhibit higher NCPC rates than walkers, with peak non-compliance occurring during late night and dawn periods. On the driver side, pedestrian NCPC behavior is the strongest predictor of failure to yield. DUY behavior is also more likely during morning periods and among drivers of personal (non-commercial) vehicles, and when the pedestrian in question is older, Black or Brown, and exhibits VHI. These findings highlight the importance of addressing social and behavioral factors in pedestrian safety interventions. By revealing how marginalization and context interact to shape risk, this research contributes to the transportation equity literature and supports interventions that go beyond infrastructure, such as education campaigns, bias reduction, and community-led safety initiatives.

Keywords: Pedestrian-vehicle interaction; pedestrian crossing violation; jaywalking; driver yielding; naturalistic video data collection; pedestrian safety

1. INTRODUCTION

1.1. Background and Context

Pedestrian safety remains a pressing concern in transportation research and practice, driven by persistently high and rising rates of injuries and fatalities, and the inequitable distribution of these risks across communities. An analysis of National Highway Traffic Safety Administration (NHTSA) records reveals a steady deterioration in pedestrian safety since fatalities hit a historic low in 2009. In 2022, pedestrian fatalities reached 7,593, representing a 1.6% increase from the previous year, and the highest annual toll since 1981 (NHTSA, 2024a). Although fatalities declined slightly to 7,314 in 2023, the number of non-fatal pedestrian crashes increased by 1.34% compared to 2022, reaching 68,244 incidents (NHTSA, 2024b). Overall, this long-term upward trend in pedestrian crashes and fatalities contrasts with trends in other developed nations, where pedestrian deaths have generally declined (Naumann et al., 2025). Moreover, pedestrians have become increasingly vulnerable relative to other road users, as reflected in the growing share of pedestrian deaths, which increased from 12.1% of all traffic fatalities in 2009 to 17.9% in 2023 (NHTSA, 2024b), despite relatively stable walking rates over the same period (McGuckin et al., 2018).

Demographic and socioeconomic characteristics further shape disparities in pedestrian safety. Men represent approximately 70% of pedestrian fatalities (NHTSA, 2024a), potentially reflecting gendered differences in travel behavior and risk-taking, such as more frequent violations of traffic laws (Guo et al., 2011, Brosseau et al., 2013, Dommes et al., 2015, and Baker et al., 2022). Older adults also face heightened vulnerability due to mobility limitations, slower reaction times, and increased injury severity in the event of a crash (Smart Growth America, 2024). Racial minorities, particularly Black and Native American pedestrians, experience disproportionately high rates of severe and fatal crashes, with per-trip and per-capita fatality rates significantly exceeding those of white pedestrians (Hamann et al., 2020, Sanders and Schneider, 2022, and Smart Growth America, 2024). These disparities are often attributed to differences in infrastructure quality and exposure to high-risk traffic environments (see Dadashova et al. (2024) for a comprehensive review of associations between pedestrian race and pedestrian crash propensity). Individuals experiencing homelessness also face an increased risk of being seriously injured or killed in traffic crashes, potentially due to greater reliance on walking and transit, frequent presence near high-speed, high-traffic roads, and elevated health-related vulnerabilities, including physical disabilities, mental health challenges, and substance use (USDOT, 2024). Income is another critical factor. Pedestrian fatality rates in census tracts with median household incomes below $15,000 are nearly five times those in high-income areas (Smart Growth America, 2024), a disparity often linked to a combination of higher walking rates and insufficient pedestrian infrastructure, such as inadequate crosswalks, pedestrian signals, and lighting (Morency et al., 2012, Lee et al., 2019, and Yu et al., 2022).

Although infrastructure and exposure factors, as discussed above, have long dominated pedestrian safety research, they do not fully account for the persistent demographic disparities in crash outcomes. For example, studies have shown that racial and income-related disparities in pedestrian injuries and fatalities often persist even after controlling for infrastructure and exposure variables (Roll and McNeil, 2022, Dumbaugh et al., 2023, and Haddad et al., 2023), suggesting that additional behavioral and contextual mechanisms are at play. These disparities may be related to pedestrian crossing violation behaviors or driver (un)yielding behaviors to pedestrians or the combination of the two, all of which themselves may not only be functions of traffic/BE characteristics, but also how individuals of different races/income levels behave differently (as pedestrians and drivers) in different settings. In this context, some earlier research studies have focused on traffic/BE effects and demographic characteristics on pedestrian crossing violations (for instance, see Ghomi and Hussein, 2022 and the review therein). Other studies have examined driver (un)yielding behaviors as a function of mostly pedestrian age/gender, though pedestrian race, crossing alone versus in a group, and socioeconomic characteristics of the surrounding neighborhood have also been considered (for example, Coughenour et al., 2017 and Schneider et al., 2018). Much research and understanding is still needed to examine the effects of race, income, perceived housing status of pedestrians (whether seemingly housing-secure or not), and time-of-day on pedestrian crossing violations and driver (un)yielding behaviors.

1.2. Research Objectives and Significance

The discussion above motivates the research in the current paper, which seeks to uncover the often-hidden social and behavioral patterns in pedestrian safety and contribute to a more comprehensive understanding of pedestrian safety risks. Specifically, we focus on two central questions: (1) How do individual pedestrian attributes, social context, and time-of-day/weather conditions interact to influence pedestrian crossing behavior, especially non-compliant crossings (e.g., jaywalking or crossing against signals)? (2) What factors influence driver yielding behavior, and how does the combination of pedestrian characteristics, group dynamics, and temporal conditions shape drivers' decisions to yield?

To address these questions, we analyzed over 1,000 hours of intersection video footage from two locations in Austin, Texas. Trained analysts coded 20,995 pedestrian crossings and documented 3,124 pedestrian-vehicle interactions. These data informed the analysis of two outcomes: (a) non-compliant pedestrian crossing (NCPC) and (b) driver unyielding (DUY) behavior (that is, driver failure to yield in contexts where statutory provisions impose a legal obligation on drivers to yield to pedestrians). By capturing real-world behavior across diverse contexts, this research contributes to the transportation literature by addressing the influence of individual and social factors on pedestrian safety. The findings have significant implications for urban planning and policy, underscoring the need for interventions that extend beyond infrastructure to address the broader social and behavioral aspects of street safety.

2. RELEVANT LITERATURE

The literature on pedestrian safety is extensive, with most studies relying on historical crash data to examine both macro- and micro-level patterns, such as identifying high-risk groups and locations, or analyzing how individual pedestrian characteristics influence crash severity (Mirhashemi et al., 2022, Shrinivas et al., 2023, Kumar et al., 2025). These studies have been valuable for identifying patterns and disparities in safety outcomes; however, there is still a need to uncover, at a fine-grained behavioral level, the mechanisms underlying those disparities. Recognizing this need, recent research has shifted toward examining the behavioral dimensions of pedestrian safety, focusing on the factors influencing NCPC and DUY behaviors.

In studying pedestrian and driver behaviors, researchers commonly categorize influencing factors into four domains: human, traffic, built-environment, and environmental/temporal conditions (Zhu et al., 2021, and Ghomi and Hussein, 2022). Human factors typically include the demographic and socioeconomic characteristics of pedestrians and drivers. Traffic factors refer to variables such as traffic volume, vehicle speed, and traffic composition. Built-environment factors encompass elements such as roadway geometry, pavement condition, lighting, and traffic control features. Lastly, environmental/temporal factors refer to weather conditions, ambient lighting, and the time-of-day during which pedestrian or driver behavior is observed.

A range of data collection methods has been used to study the influences of the above listed factors, including surveys (e.g., Deb et al., 2017, Mukherjee and Mitra, 2020), video-based stated-preference experiments (e.g., Liu and Tung, 2014), in-person road observations (e.g., Avineri et al., 2012, Dommes et al., 2015, Ferenchak, 2016, Aghabayk et al., 2021), recorded video footage either coded manually (e.g., Bella and Nobili, 2020, Zhu et al., 2021) or using computer vision algorithms (e.g., Anik et al., 2021, Chavis et al., 2023, Wan et al., 2023), field experiments (e.g., Goddard et al., 2015, Coughenour et al., 2017), and immersive virtual reality studies (e.g., Hübner et al., 2025, Nazemi et al., 2025). Of these methods, in-person road observations and the analysis of video footage are considered naturalistic approaches, as they capture real-world behavior in uncontrolled, everyday settings without researcher intervention. Such naturalistic approaches are particularly suitable for an investigation of the social-behavioral mechanisms underlying pedestrian safety risk.

To align with this paper's objective of uncovering the social-behavioral mechanisms of pedestrian safety risk, our synthesis of the literature is confined to naturalistic observational studies and to an overview of pedestrian/driver behavior as a function of (a) pedestrian/driver sociodemographics, (b) pedestrian activity and social context, (c) time-of-day and weather factors, and (d) vehicle characteristics. That is, unlike many other studies that attempt to investigate the effects of built-environment and traffic characteristics on pedestrian crash propensity, we focus directly on pedestrian/driver behaviors that constitute precursor factors affecting pedestrian crash propensity. Tables 1 and 2 serve to support such an overview. Specifically, Table 1 synthesizes naturalistic observational studies from the past decade on NCPC, while Table 2 similarly synthesizes naturalistic studies of DUY. Each table summarizes the study location, data collection duration, number of observation sites, sample size, behavioral outcomes, and the explanatory variables considered (grouped within the variable categories identified earlier). Table 2 includes an additional column for data collection methods, reflecting the more diverse methodological approaches used in DUY research, including controlled field experiments that have proven invaluable for examining racial effects on DUY behavior. Together, the literature review and tabulated summaries highlight common themes and methodological approaches, while identifying critical gaps related to underexplored sociodemographic and behavioral dimensions influencing pedestrian/driver behaviors.

2.1. Non-compliant Pedestrian Crossing (NCPC) Behavior

Almost all studies of pedestrian non-compliant crossing (NCPC) have been undertaken using a sample of signalized intersections, as shown in the "Locations" column of Table 1. Aghabayk et al. (2021) is the only video-based pedestrian study we are aware of that considers and distinguishes between signalized and unsignalized intersections in the analysis (to be sure, Schwebel et al., 2022 include both signalized and unsignalized intersections in their data collection effort, but do not distinguish between the two in their analysis). Aghabayk et al.'s findings indicate similar patterns (across signalized and unsignalized intersections) of technology distraction and pedestrian demographic effects on NCPC, except that middle-aged/older individuals (41 years of age or older) tend to be more cautious (looked both ways for traffic more attentively and purposefully) than their younger peers during crossings at signalized intersections, but older individuals (62 years or older) are no different from their younger peers during crossings at unsignalized intersections. Given the higher potential for conflicts at unsignalized intersections observed in their study, and the increased inability of older individuals to take last-minute evasive actions to avoid a crash, the authors suggest that older individuals should be encouraged to pay more attention at unsignalized intersections. Beyond Aghabayk et al., the vast majority of other prior NCPC work at signalized intersections finds that NCPC patterns vary by signal control and crossing type, with higher rates of NCPC associated with (a) exclusive pedestrian phasing (no vehicular movement allowed during the pedestrian phase), (b) longer pedestrian wait times, and (c) roadway locations with shorter crosswalk widths and on-street parking or refuge islands (Baker et al., 2022; Dommes et al., 2015; Ivan et al., 2017; Miladi et al., 2025; Russo et al., 2018).

Several behavioral outcomes have been examined in the pedestrian safety literature to gain a better understanding of crossing patterns and decision-making processes. These include specific violation types, such as temporal violations (e.g., running a red light) and spatial violations (e.g., crossing outside a designated crosswalk or midblock), as well as distraction, crossing speed, waiting time, gap acceptance, and head-turn frequency, all used as proxies for risk awareness. These outcomes are highlighted in the comprehensive reviews by Theofilatos et al. (2021) and Ghomi and Hussein (2022), as well as in the "Measured Outcomes" column of Table 1. In terms of the determinants of NCPC-related behavioral outcomes, a variety of pedestrian sociodemographics, pedestrian activity/context, and time-of-day/weather conditions have been considered, as discussed next.

2.1.1. Pedestrian Sociodemographics

Gender is one of the most commonly studied factors in pedestrian behavior (see Table 1), with many studies finding that men tend to take more risks than women. However, the evidence varies by region and behavior type. U.S.-based studies present mixed evidence on gender differences in pedestrian behavior. While some findings suggest that men are more prone to spatial violations (Russo et al., 2018, Rafe et al., 2025) and distraction (Russo et al., 2018), gender differences in temporal violations have often been found to be statistically insignificant (Russo et al., 2018, Rafe et al., 2025). Interestingly, Baker et al. (2022) found that gender differences in temporal violations were evident only in low-risk situations (where a concrete safety island was present), with men being more non-compliant. Similarly, Schwebel et al. (2022) found no significant gender effects across multiple dimensions, including situational awareness, distraction, and general unsafe crossing behavior. In contrast, non-U.S. studies have more consistently associated male pedestrians with higher rates of temporal and spatial violations (Xie et al., 2018, Aghabayk et al., 2021, Bendak et al., 2021, Zhu et al., 2021, Zhang et al., 2023, Miladi et al., 2025). Some non-U.S. studies have also pointed to reduced situational awareness among male pedestrians, such as less frequent head-turning before crossing (Bendak et al., 2021).

With respect to age, recent U.S.-based studies have generally reported no statistically significant association between older age groups and pedestrian violations (Ivan et al., 2017; Russo et al., 2018; Rafe et al., 2025). However, some other age-related trends have been evident. For instance, Rafe et al. (2025) found that children and adolescents were more likely to engage in temporal violations, while Russo et al. (2018) identified elevated distraction levels among pedestrians aged 16 to 29. Findings from studies conducted outside the U.S. also highlight significant age-related trends, though results remain somewhat inconsistent. Older adults were frequently associated with safer pedestrian behavior, including a lower likelihood of temporal violations (Aghabayk et al., 2021; Bendak et al., 2021), reduced distraction (Bendak et al., 2021), greater situational awareness (Aghabayk et al., 2021), and a higher tendency to wait on the curb rather than in the roadway (Dommes et al., 2015). However, several studies found no significant association between age and temporal violations (Dommes et al., 2015; Xie et al., 2018; Zhang et al., 2023), and only one study in Hong Kong reported that older adults were more likely to commit such violations (Zhu et al., 2021). Additionally, Miladi et al. (2025) noted that older adults were more likely to complete crossing during the red phase after starting on green, a behavior likely attributable to slower walking speeds rather than intentional non-compliance.

We are not aware of any naturalistic studies of pedestrian crossing behavior that examine pedestrian race-related variations, as we consider in our analysis (see last row of Table 1).

2.1.2. Pedestrian Activity and Context

Social context (first column under "pedestrian activity and context" in Table 1) is frequently examined in pedestrian behavior research, though it is defined in varying ways, including walking with companions or the presence or number of others crossing concurrently. Additionally, although pedestrian volume is typically treated as a traffic-related variable, it can also serve as a proxy for the presence of other pedestrians at a crossing. These differences in definition partly explain the mixed findings across the literature. Some social context variables are associated with safer behavior. For instance, Rafe et al. (2025) and Dommes et al. (2015) found that others crossing at the same time reduced spatial and temporal violations, while Zhu et al. (2021) and Bendak et al. (2021) reported fewer violations and technological distractions when pedestrians were accompanied, particularly by children. Compliance has also been found to increase with pedestrian volumes (Ivan et al., 2017, Miladi et al., 2025). However, other forms of social presence appear to encourage risk. Larger group sizes, especially three or more, were linked to higher violation rates (Russo et al., 2018, Zhang et al., 2023), and observing others engage in non-compliant behavior increased the likelihood of doing the same (Xie et al., 2018). Adding further nuance, Anik et al. (2021) documented gendered responses to group behavior, finding that women were more risk-averse when walking alone but more likely to follow a group in engaging in non-compliant crossings. These findings suggest that social context can either deter or promote risky crossing behavior, depending on how it is defined and the actions modeled by surrounding pedestrians.

With the growing prevalence of smartphones and personal technology, several studies have also examined the impact of technological distraction (second column under "pedestrian activity and context" in Table 1) on pedestrian crossing behavior. Despite increased interest in distraction, findings generally indicate inconsistent or limited effects on crossing behavior, with impacts varying by distraction type, violation type, and surrounding context. Texting, for instance, is frequently linked to reduced visual awareness directed towards traffic and pavement markings, increasing the likelihood of spatial crossing violations (Russo et al., 2018, Aghabayk et al., 2021, Bendak et al., 2021). However, multiple studies found no significant association between texting and temporal violations (Russo et al., 2018, Aghabayk et al., 2021, Bendak et al., 2021, Miladi et al., 2025). Similarly, pedestrians engaged in phone conversations were generally not associated with increased temporal violations (Russo et al., 2018, Aghabayk et al., 2021, Miladi et al., 2025), while headphone use was linked to a lower likelihood of such violations (Aghabayk et al., 2021). Schwebel et al. (2022) further demonstrated that the influence of distraction varied by location, finding that distraction was associated with reduced risk of temporal violations but a higher risk of spatial violations in downtown areas. In contrast, in entertainment districts, distraction was associated with a lower risk of spatial violations.

Another contextual variable commonly examined in the pedestrian safety literature is the act of carrying or holding visible items. This factor has been studied more frequently in non-U.S. contexts, where it has generally shown no significant association with pedestrian violations (Bendak et al., 2021, Zhu et al., 2021, Zhang et al., 2023). A notable exception is Aghabayk et al. (2021), who found that pedestrians carrying items were more likely to look left and right for traffic before crossing, suggesting increased situational awareness. Other miscellaneous factors in this category include changes in walking speed, such as shifting from walking to running, which Rafe et al. (2025) found to have no significant effect. The same study also reported that individuals using mobility devices, such as wheelchairs, skateboards, scooters, or bicycles, were less likely to commit spatial violations but more likely to engage in temporal violations.

2.1.3. Time-of-Day and Weather

Time-of-day and weather have been shown to influence pedestrian crossing behavior. Rafe et al. (2025) found that overnight hours (00:00-05:59) were associated with elevated rates of both spatial and temporal violations relative to other times of day. Similarly, Fu et al. (2022) reported that pedestrians were less likely to cross in the presence of right-turning vehicles during nighttime hours (19:00-22:00) compared to daytime periods (10:00-16:00), while Liu and Tung (2014) observed increased caution at dusk, likely due to reduced visibility. Additionally, Ivan et al. (2017) noted decreased compliance during late afternoon hours (16:00-18:00) compared to earlier in the day, as well as on Fridays compared to other weekdays.

Regarding weather-related conditions, Bendak et al. (2021) found that higher temperatures were associated with increased temporal violations, whereas Rafe et al. (2025) reported a similar effect on spatial violations but found no significant impact on temporal violations. Both studies observed that precipitation was associated with fewer temporal violations, suggesting that rain may discourage risk-taking. Ivan et al. (2017) further found that cloudy conditions reduced crossing compliance. Seasonal variation was evident in Miladi et al. (2025), who reported that pedestrians were less likely to finish crossing on a red light in the fall than in the summer. However, during the early COVID-19 period (spring 2020), pedestrians were more likely to initiate and complete crossings during the red phase, likely due to decreased vehicle traffic.

2.2. Driver Unyielding (DUY) Behavior

Similar to the case of pedestrian crossing, a variety of outcomes are used to study DUY behavior, including hard stop instances, soft yielding, or complete non-yielding, as well as surrogate safety metrics such as stopping distance (e.g., Figliozzi and Tipagornwong, 2016), vehicle deceleration rate (e.g., Bella and Nobili, 2020), Time to Collision (TTC) (e.g., Bella and Nobili, 2020, Pinnow et al., 2021), and Post-Encroachment Time (PET) (e.g., Pinnow et al., 2021, Das et al., 2023). These outcomes have been examined in a wide range of crossing types, as summarized in the "Locations" column of Table 2. Relative to NCPC behavior, more earlier studies of DUY have considered unsignalized locations (in addition to signalized intersections), though most such unsignalized locations correspond to midblock crossings rather than intersections. Again, only one video-based study that we are aware of (Bella and Nobili, 2020) considers a sample of locations with crosswalks at both signalized and unsignalized intersections. But even this study does not distinguish between DUY at signalized versus unsignalized intersections. Rather the study, based on retrofitting the vehicles of 16 study participants with a GoPro Hero3 video camera and a global positioning system (GPS) device to analyze yielding behavior over one run of a 3.5 km stretch of two streets at the center of Rome in Italy, focused on DUY differences between pedestrians legally crossing at marked "zebra" crossings at intersections (regardless of signalized or unsignalized) versus pedestrians illegally crossing at unmarked crossing locations. Bella and Nobili's descriptive analysis found that motorists did not yield 21% of the time to pedestrians crossing legally, but did not yield 55% of the time to pedestrians crossing illegally. More broadly, though no one study has examined DUY behavior at different roadway signal control contexts, a synthesis of prior research does suggest that signalization at intersection locations decreases DUY rates relative to at unsignalized intersections and midblock crossings. Other results associated with crossing geometry and roadway traffic/parking characteristics indicate that DUY rates are lower at roadway locations with (a) shorter crosswalk widths, (b) fewer lanes, (c) lower vehicle speeds and volumes, (d) the presence of visual or geometric crosswalk treatments (such as raised, colored, or textured crossings), and (e) no on-street parking (Schneider et al., 2018, Anciaes et al., 2020, Pechteep et al., 2024). In the remainder of this section, we provide an overview of the literature on the determinants of DUY behavior.

2.2.1. Pedestrian Sociodemographics

Existing studies have consistently shown that drivers are less likely to yield to men (Coughenour et al., 2017, Anciaes et al., 2020, Demir et al., 2020, Zafri et al., 2022, Pechteep et al., 2024), with only two exceptions in Table 2 reporting no significant effect of pedestrian gender on DUY (Dileep et al., 2016, Schneider et al., 2018). In contrast, while many studies in Table 2 also examined the effect of pedestrian age, most found no significant age associations (Dileep et al., 2016, Schneider et al., 2018, Anciaes et al., 2020, Demir et al., 2020, Zafri et al., 2022). An exception was Pechteep et al. (2024), who reported that drivers were less likely to yield to younger pedestrians relative to pedestrians classified as "elderly."

Since 2015, there has been growing interest in the U.S. in examining the effect of pedestrian race on DUY behavior, particularly through staged field experiments. Studies by Goddard et al. (2015), Coughenour et al. (2017), and Coughenour et al. (2020) found that Black pedestrians experienced higher DUY rates and longer wait times at crosswalks compared to White pedestrians.[1] These findings are also supported by in-person observational research by Schneider et al. (2018), which reached similar conclusions. However, Schneider et al. (2018) reported no significant associations between DUY and site-level racial composition, such as whether the majority of pedestrians or drivers were White. Similarly, Anciaes et al. (2020) found no significant differences in DUY behavior toward pedestrians using a wheelchair or walking stick.

2.2.2. Pedestrian Activity and Context

Beyond demographic traits, specific pedestrian behaviors and contextual cues also affect driver responses. Social context (especially group presence and group walking) has generally been associated with lower DUY rates (Zafri et al., 2022, and Pechteep et al., 2024), with the exception of Dileep et al. (2016) and Schneider et al. (2018), who reported insignificant findings. Additionally, all studies that examined assertiveness-related variables found that assertive pedestrian behaviors, such as brisk movement toward the crosswalk, hand gestures, or making eye contact with the driver, decreased the likelihood of DUY (Dileep et al., 2016, Schneider et al., 2018, Zafri et al., 2022).

In contrast, DUY increased in the presence of non-compliant or jaywalking pedestrians, often demonstrating sharper deceleration and shorter stopping distances (Bella and Nobili, 2020). However, Zafri et al. (2022) reported that whether or not a pedestrian used a designated crosswalk while crossing an intersection had no significant impact on DUY behavior. Their findings also indicate that DUY instances are fewer when pedestrians carry baggage, and are not distracted by mobile devices (Zafri et al., 2022).

[1] The observed differences in DUY based on pedestrian race may operate unconsciously or in a situation-dependent manner, particularly in discretionary, time-pressured interactions. However, as with any naturalistic outcome-based observation study, these studies are unable to distinguish implicit from explicit cognitive mechanisms underlying differential yielding outcomes.

2.2.3. Driver Sociodemographics

Relatively few studies have examined the influence of driver gender and age on DUY behavior. Most of these studies reported no significant differences based on driver gender (Hirun, 2016, Schneider et al., 2018, Demir et al., 2020). However, Demir et al. (2020) noted that gender effects were significant only among middle-aged drivers, suggesting that generational shifts may be reducing gender disparities in DUY decisions. Findings on driver age are similarly inconsistent. Hirun (2016) observed that younger drivers were less likely to yield, while Demir et al. (2020) found the opposite in a Turkish context, and Schneider et al. (2018) did not report a significant association between driver age and DUY behavior.

In addition to demographic characteristics, other factors such as social cues in the form of the behavior of other drivers and formal education attainment have also been found to impact DUY behavior. For instance, DUY was less likely when drivers observed a preceding vehicle yielding or witnessed a yielding event in an adjacent lane (Figliozzi and Tipagornwong, 2016), while Hirun (2016) found that drivers with less than a bachelor's degree and less awareness of right-of-way laws had higher DUY rates.

2.2.4. Time-of-Day and Weather

The effects of time-of-day and weather conditions on DUY behavior have also received limited attention, and the few studies that have addressed them report mostly insignificant findings (Anciaes et al., 2020, Demir et al., 2020, Fu et al., 2022).

2.2.5. Vehicle Characteristics

Coughenour et al. (2020) observed that drivers in expensive, high-status vehicles, often serving as a proxy for social class, had higher DUY rates, with DUY rates increasing by approximately 3% for every $1000 increase in vehicle price. However, Greitemeyer (2023) presented a contrasting view, noting that vehicle status was not significantly related to whether a driver yielded or drove through a crosswalk when a pedestrian was waiting. Studies have also found that drivers of larger or more powerful vehicles, such as SUVs, pickup trucks, and heavy vehicles, are generally less likely to yield to pedestrians (Dileep et al., 2016, and Figliozzi and Tipagornwong, 2016). These patterns may reflect visibility differences, a sense of greater protection, or potentially different driving attitudes associated with these vehicle types.

Table 1. Summary of Literature on Non-Compliant Pedestrian Crossing (NCPC) Behavior

Reference	Country/ Region	Data collection duration	Locations (Type/ Number)	No. of Obs.	Measured Outcomes	Pedestrian Sociodemographics			Pedestrian Activity and Context			Time-of-Day/ Weather	
						Gender	Age	Race	Social context	Distraction	Carrying items/ Other*	Time-of-day	Weather
Studies conducted in the U.S.													
Rafe et al., 2025	Utah	24-60 hrs/site	39 signalized intersections	5,589	• Spatial violation • Temporal violation	×	×		×		×	×	×
Baker et al., 2022	Campus of a southern American university	Two-weeks (~5 hrs/day/site)	1 signalized intersection crossing	2,707	• Temporal violation	×							
Schwebel et al., 2022	Alabama	~ 1 hr/site	112 (almost entirely signalized) intersections (no differentiation between signalized and unsignalized intersections in analysis)	3,248	• Pedestrian unsafe crossing (did not look left/right; crossed against walk signal or outside crosswalk) • Pedestrian distracted crossing	×	×		×	×	×		
Russo et al., 2018	New York and Arizona	3 hrs/site	4 signalized intersections	3,038	• Walking speed • Pedestrian distraction • Spatial violation • Temporal violation	×	×		×	×			
Ivan et al., 2017	Connecticut	216 hrs (~6 hrs/site)	42 signalized intersections	14,838	• Temporal violation		×					×	×
Studies conducted outside the U.S.													
Miladi et al., 2025	Canada	9 hrs/site	24 signalized intersections	4,711	• Pedestrian crossing start on red • Pedestrian crossing finish on red • Pedestrian crossing finish on red/started on green • Pedestrian crossing completely on red	×	×			×	×		×
Zhang et al., 2023	China	0.5 hrs/site	6 4-lane two-way road segments	723	• Spatial violation	×	×		×		×		
Fu et al., 2022	China	54 hrs (9 hrs/site)	6 crosswalks adjacent to right-turning vehicles at signalized intersections	518	• Pedestrian crossing decisions: cross vs. not cross				×			×	
Aghabayk et al., 2021	Iran	4 hrs/site	2 signalized and 2 unsignalized intersection	552	• Temporal violation • Looking left-right for traffic before/while crossing	×	×		×	×	×		
Bendak et al., 2021	United Arab Emirates	0.5 hrs/site	5 signalized intersections and 5 signalized midblock crossings	708	• Temporal violation • Spatial violation • Looking left-right before crossing • Crossing speed	×	×		×	×	×		×
Zhu et al., 2021	Hong Kong	5 hrs/site	6 signalized crosswalks	6,320	• Temporal violation	×	×		×		×		
Xie et al., 2018	Hong Kong	1.5 hrs/site	7 signalized intersections	7,230	• Temporal violation	×	×						
Dommes et al., 2015	France	unspecified	6 signalized intersections	680	• Waiting position (curb vs. road) • Running during crossing • Situational awareness • Temporal violation	×	×		×		×		
This Study	Texas	432-864 hrs/site	2 signalized intersections	20,995	• Temporal or spatial violation	×	×	×	×	×	×	×	×

*The "Other" category under "Pedestrian Activity and Context" corresponds to changes in pedestrian speed (change from walking to running) and use of mobility devices such as wheelchairs and scooters.

Table 2. Summary of Literature on Driver Unyielding (DUY) Behavior

Reference	Country/ Region	Data collection method	Data collection duration	Locations (Type/ Number)	No. of Obs.	Measured outcomes	Pedestrian Sociodemographics				Pedestrian Activity and Context			Driver Sociodemographics			Time-of-Day/ Weather		Vehicle Characteristics	
							Gender	Age	Race	Other*	Social context	Assertiveness	Non-Compliant crossings	Gender	Age	Other**	Time-of-day	Weather	Car Cost	Car Type
Studies conducted in the U.S.																				
Coughenour et al, 2020	Nevada	Controlled field experiment	2 hrs/site	2 non-signalized mid-block crosswalks	461	• Driver yielding	×		×										×	
Schneider et al., 2018	Wisconsin	In-person observation	2 hrs/site	20 uncontrolled intersections	364	• Driver yielding	×	×	×	×	×	×		×	×	×				
Coughenour et al, 2017	Nevada	Controlled field experiment	2 hrs/site	2 non-signalized midblock crosswalks	126	• Driver yielding			×											
Figliozzi and Tipagornwong, 2016	Oregon	Video recording	1 hr	1 signalized intersection	116	• Driver yielding • Stopping distance					×					×				×
Goddard et al., 2015	Oregon	Controlled field experiment	88 crossing trials	1 unsignalized mid-block crossing	173	• Driver yielding			×											
Studies conducted outside the U.S.																				
Pechteep et al., 2024	Thailand	Video recording	4 hrs/site	4 midblock crosswalks	400	• Driver non-yield, soft-yield, and yield behavior	×	×			×									
Fu et al., 2022	China	Video recording	54 hrs (9 hrs/site)	6 crosswalks adjacent to right-turning vehicles at signalized intersections	543	• Driver yielding (including full stops, yielding with rolling stops, and non-yielding)					×						×			×
Zafri et al., 2022	Bangladesh	Video recording	2 hrs/site	6 signalized or police-controlled intersections	314	• Driver yielding	×	×			×	×	×							
Anciaes et al., 2020	England	Video recording	14-32 min/site	3 unsignalized zebra and 17 courtesy crossings (drivers have no legal obligation to yield)	937	• Crossing design characteristics that affect driver yielding frequency	×	×		×	×						×			×
Bella and Nobili, 2020	Italy	Video recording and GPS on vehicle	--	4 signalized and 11 unsignalized zebra crossings	76	• Driver non-yield, soft-yield, and yield behavior							×							
Demir 2020	Turkey	In-person observation	18 hrs (3 hrs/day)	1 roundabout	1140	• Driver yielding	×	×						×	×		×	×		×
Dileep et al., 2016	India	Video recording, radar gun, manual recording	4 hrs (1 hr/site)	4 undivided mid-block locations	815	• Driver yielding	×	×		×	×	×								×
Hirun, 2016	Thailand	Survey	--	--	445	• Driver yielding								×	×	×				
This Study	Texas	Video recording	1,296 hrs	2 signalized intersections	3,124	• Driver yielding	×	×	×	×	×		×				×	×		×

*The "Other" category under "Pedestrian Sociodemographics" includes the following variables: majority of pedestrians are White, majority of drivers are White, and presence of a visible disability.
** The "Other" category under "Driver Demographics" includes witnessing yielding by other drivers, formal education levels, and knowledge of traffic rules.

2.3 The Current Paper

Overall, the existing literature provides valuable insights into the factors that influence pedestrian crossing compliance and shape pedestrian-driver interactions. However, significant gaps remain, particularly concerning the entire range of sociodemographic characteristics, pedestrian activity and contextual conditions, and time-of-day/weather factors in these behaviors. This study contributes to the literature by addressing several of these limitations and advancing empirical understanding in eight key ways.

First, as observed in Table 1, pedestrian crossing behavior remains relatively understudied in the U.S. Much of the existing evidence originates from international contexts, such as China, Hong Kong, and Iran, where dense urban design, high pedestrian volumes, and walk-oriented cultures create conditions that differ substantially from those in the U.S., thereby limiting the generalizability of the findings. In addition, the few U.S.-based studies that do exist often report varying results across key behavioral outcomes and demographic variables, including gender, age, and distraction, highlighting the need for further empirical research grounded in the U.S. context. Our study addresses this gap by providing detailed, real-world observational data from Texas, offering both a geographic contribution and a context-sensitive analysis that helps clarify and possibly reconcile conflicting findings in the literature.

Second, prior studies have often relied on short-duration data collection, frequently limited to a few hours or a single day per site (as indicated in the third column of Table 1 and the fourth column of Table 2). While this approach may simplify data management and annotation, it limits the ability to capture rare events, such as instances of NCPC or DUY. In contrast, our study incorporates continuous video monitoring over periods exceeding two weeks at each of two intersection sites, resulting in a large and temporally diverse dataset of over 20,000 observations. To our knowledge, this is the first study to combine high-volume naturalistic footage with detailed pedestrian-, vehicle-, and context-level variables to examine interactions in shared spaces.

Third, while factors such as gender, age, and group size have received substantial attention in past research (see Tables 1 and 2), the effects of pedestrian race, and perceived housing insecurity on both NCPC and DUY behavior remain largely underexplored. Although a few field experiments in the U.S. have investigated racial disparities in DUY, these studies have generally relied on staged crossings with limited variation in pedestrian characteristics and small sample sizes (see Table 2), constraining the scope and generalizability of their findings. Our large-scale naturalistic observational study highlights how race and social vulnerability intersect with pedestrian risk, offering evidence that can inform more equitable safety interventions and policy reforms.

Fourth, our study incorporates pedestrian activity factors, specifically whether individuals are walking or running, which have received limited attention in the existing literature. To our knowledge, only one prior study (Zhang et al., 2023) differentiated between walking and running activity and reported no statistically significant difference in crossing behavior. Accordingly, this study offers a more refined understanding of how pedestrian movement, whether walking or running, influences crossing behavior and driver response.

Fifth, the influence of time-of-day and weather conditions on pedestrian and driver behaviors remains insufficiently explored. Existing studies often focus on daytime or peak-hour data collection under favorable weather conditions, overlooking how risk-taking and activity patterns may vary across different temporal contexts. Our continuous 24/7 observation protocol enables us to examine behavioral variation across a range of time periods and weather scenarios, offering a more comprehensive view of pedestrian risk.

Sixth, most earlier studies examine exogenous variables in isolation, without adequately considering potential interaction effects. In this study, our large sample enables us to consider a number of interaction effects, such as whether there are gender/race, gender/time-of-day, and race/time-of-day interaction effects in both pedestrian crossing behavior and driver yielding behavior. Similarly, we not only consider the general impact of non-compliant crossing behavior on driver yielding behavior, but also whether, for example, male pedestrians who are non-compliant are yielded to differently than female pedestrians who are non-compliant. More generally, we consider a whole range of potential determinants of NCPC and DUY behavior relative to earlier studies, as well as their interactions, as should be clear from the '×' markings (in the last row of Tables 1 and 2) identifying variables considered in the current study.

Seventh, we employ a copula-based bivariate logit model to jointly analyze pedestrian crossing behavior and driver yielding behavior. This joint modeling approach accounts for unobserved factors, such as traffic conditions, lighting, weather, or time pressures, that may influence both pedestrian and driver decisions simultaneously. For example, the presence of a visual obstruction (unobserved from the point of view of the analyst) might make it harder for a driver to see a pedestrian crossing, while also making it harder for the pedestrian to see oncoming traffic, potentially affecting both the driver's yielding decision and the pedestrian's crossing decision. By accounting for these unobserved factors that create correlations between outcomes, we can better isolate the causal effect of one behavior on another (such as the effect of NCPC on DUY). Our final model configuration, based on the empirical analysis, reveals that pedestrian NCPC behavior directly affects DUY decisions.

Lastly, much of the existing literature on NCPC has focused on traditional four-way signalized intersections, while DUY behavior is typically examined at unsignalized intersections and midblock crossings (see the fourth column of Table 1 and the fifth column of Table 2). In contrast, we examine two locations featuring channelized slip lanes, which require more discretionary driver judgement in yielding decisions. This intersection design introduces greater complexity to pedestrian-driver interactions and offers novel insights into decision-making under ambiguous right-of-way conditions. To our knowledge, only Fu et al. (2022) have previously examined slip lanes in their way, making our contribution particularly distinctive.

3. DATA COLLECTION AND DESCRIPTION

3.1. Study Sites and Context

To examine pedestrian-driver interactions in naturalistic settings, video footage was collected from two signalized intersections in Austin, Texas. Figures 1 and 2 present detailed site characteristics, including annotated intersection layouts that show traffic flow patterns and pedestrian infrastructure, as well as field photographs that illustrate physical design features and operational conditions.

The first intersection (hereafter referred to as the MB intersection) is located in north Austin at the junction of the southbound Mopac (Loop 1) frontage road and West Braker Lane. Mopac is a major access-controlled highway running north-south, while Braker Lane is a three-lane, divided arterial road running east-west. As illustrated in Figure 1, this intersection features right-turn slip lanes accompanied by pedestrian refuge islands, and unprotected bike lanes along Braker Lane in both directions. The intersection is signalized and equipped with pedestrian push-button-activated crossing signals at all four corners. Each approach features a "Walk/Don't Walk" display and a countdown timer indicating the remaining time for safe pedestrian crossing. Additionally, each channelized right-turn slip lane is controlled by a "Stop Here for Pedestrians" sign, instructing motorists to stop for pedestrians crossing from the sidewalk to the refuge islands.

The second intersection (hereafter referred to as the DS intersection) is located in central Austin, within the University of Texas at Austin's main campus. It connects East Dean Keeton Street, a four-lane east-west arterial divided by a raised concrete median, with San Jacinto Boulevard, a two-lane north-south local street, as shown in Figure 2. The intersection is signalized and features left-turn lanes on all approaches, along with channelized right-turn slip lanes (each equipped with a pedestrian refuge island) on the northbound and southbound legs. Pedestrian infrastructure includes push-button-activated pedestrian signals at all four corners, with "Walk/Don't Walk" displays and countdown timers. "Yield" signs are also posted at each slip lane, directing motorists to yield to pedestrians within the crosswalks. Additionally, unprotected bike lanes on both sides of East Dean Keeton Street support substantial bicycle and e-scooter activity. This site experiences consistently high pedestrian volumes throughout the day due to the presence of university students, faculty, staff, nearby residents, and event attendees.

The Euclidean distance between the two sites is approximately 7.4 miles, providing spatial variation between central and north Austin. Site selection was guided by the availability of nearby university-owned property, which enabled the secure placement of cameras and equipment during the recording period, an approach commonly adopted in similar observational studies (e.g., Figliozzi and Tipagornwong, 2016, Wells et al., 2018, Baker et al., 2022, Piazza et al., 2022, Gerogiannis and Bode, 2024). In addition to these practical considerations, the study intersections were selected because they feature right-turn slip lanes and are representative of similar 4-leg, multi-lane geometries that allow discretionary yielding and complex pedestrian-vehicle interactions (Fu et al., 2022). Furthermore, the two intersections were also chosen because they represent variations in (a) vehicle and pedestrian volumes (with the MB intersection featuring relatively low pedestrian volumes but high vehicular volumes, and the DS intersection having relatively high pedestrian volumes but low vehicular volumes), (b) land-use patterns that also brought good variation in the pedestrian and driver demographic mix (with the MB intersection within 0.35 miles of a large shopping center in a mixed land-use setting with mostly middle-aged and older employed individuals, and the DS intersection in a university campus setting with more younger student individuals), and (c) neighborhood income levels (with the MB intersection located in a census tract with a median annual household income of $86,806, and the DS intersection located in two census tracts with an average median annual household income of $40,510 (U.S. Census Bureau, 2023). Together, these factors contribute to good contextual variation across the study locations, capturing diverse driver and pedestrian behaviors. Such variations across the two intersections aid in better generalizability of the findings from the study because range variation is more important for consistently estimating the effects of the exogenous covariates than representativeness of the two intersections to the broader set of urban environment intersections.[2]

3.2.Video Setup and Data Collection

A custom-built outdoor video recording system was developed to capture pedestrian and driver behavior at the selected intersections. As shown in Figure 3, the system utilized an off-the-shelf 4K Ultra High Definition, 360° PoE IP security camera with 16x optical zoom, housed in a weatherproof enclosure to ensure continuous operation during adverse weather conditions. Video footage was managed using the open-source *Shinobi* surveillance software, installed on a Raspberry Pi 3. *Shinobi* was selected for its ability to support continuous, high-resolution video capture, compatibility with IP cameras, and flexible configuration options, including scheduling and timestamping (Shinobi Systems, 2025). The recorded data were then stored locally on a 5TB hard drive. This low-power, cost-effective system enabled extended unattended monitoring, making it well-suited for detailed behavioral observation in naturalistic settings.

At each site, the camera was strategically positioned to maximize visibility of the intersection area. The crosswalks captured in the recordings are indicated on the annotated intersection layouts in the upper panels of Figures 1 and 2. The field photographs in the lower panel of each figure further illustrate the camera's field of view at each site. At the MB intersection (Figure 1), the camera captured the eastbound crosswalk on Braker Lane and both crosswalks on the southbound Mopac Frontage Road, providing complete coverage of all crosswalks at this site with no visual obstructions. At the DS intersection (Figure 2), the camera captured three of the four crosswalks. The crossing on San Jacinto Boulevard at the southern end of the intersection was completely obstructed by vegetation and therefore excluded entirely from the analysis. Our analysis at the DS site thus represents 75% spatial coverage of the intersection's crosswalks, with complete observation of the three included crosswalks.

Videos were analyzed at a resolution of 2560 × 1440 pixels. At the average mid-crossing distance, this yielded a pixel-to-meter ratio of approximately 55.4 pixels per meter. While this pixel density falls below thresholds typically required for automated computer vision systems to reliably infer demographic characteristics such as age, gender, or skin tone (Hall and Perona, 2015,

[2] In particular, from an econometric standpoint, our convenience sample of pedestrian-vehicle observations at the two intersections constitutes the case of exogenous sampling, not endogenous sampling. That is, the two intersections were not selected on the basis of the number of NCPC or DUY instances. In addition, our sample displays adequate variation across the range of values of each exogenous covariate. Overall, the combination of our exogenous sampling approach, the testing and control for a host of exogenous covariates, as well as the adequate variation in the exogenous covariates implies that there is no reason to believe that the relationships developed in this paper are not applicable to the larger population of urban signalized intersections (please see Wooldridge, 1995 and Robbennolt et al., 2025 for more discussions pertaining to exogenous sampling).

Junejo and Ahmed, 2020, Sakib et al., 2022, Fernández Llorca et al., 2024), the use of trained human coders rather than automated algorithms mitigated this limitation. Human coders could pause, replay, and digitally zoom during video review to examine details more closely. Additionally, many pedestrians passed closer to the camera at some point during their crossing trajectory, providing higher-resolution views for classification.

Extended-duration, continuous video monitoring was implemented at both sites. The MB intersection monitoring period, which spanned from April 19 to May 6, 2024, resulted in 18 days of continuous data acquisition, yielding 2,976 documented crossing observations with an average rate of 8.00 pedestrians/hour. At the DS intersection, data collection took place from May 23 to June 27, 2024, resulting in 36 full days of footage and 18,019 recorded pedestrian crossings, corresponding to an average rate of 24.21 pedestrians/hour.

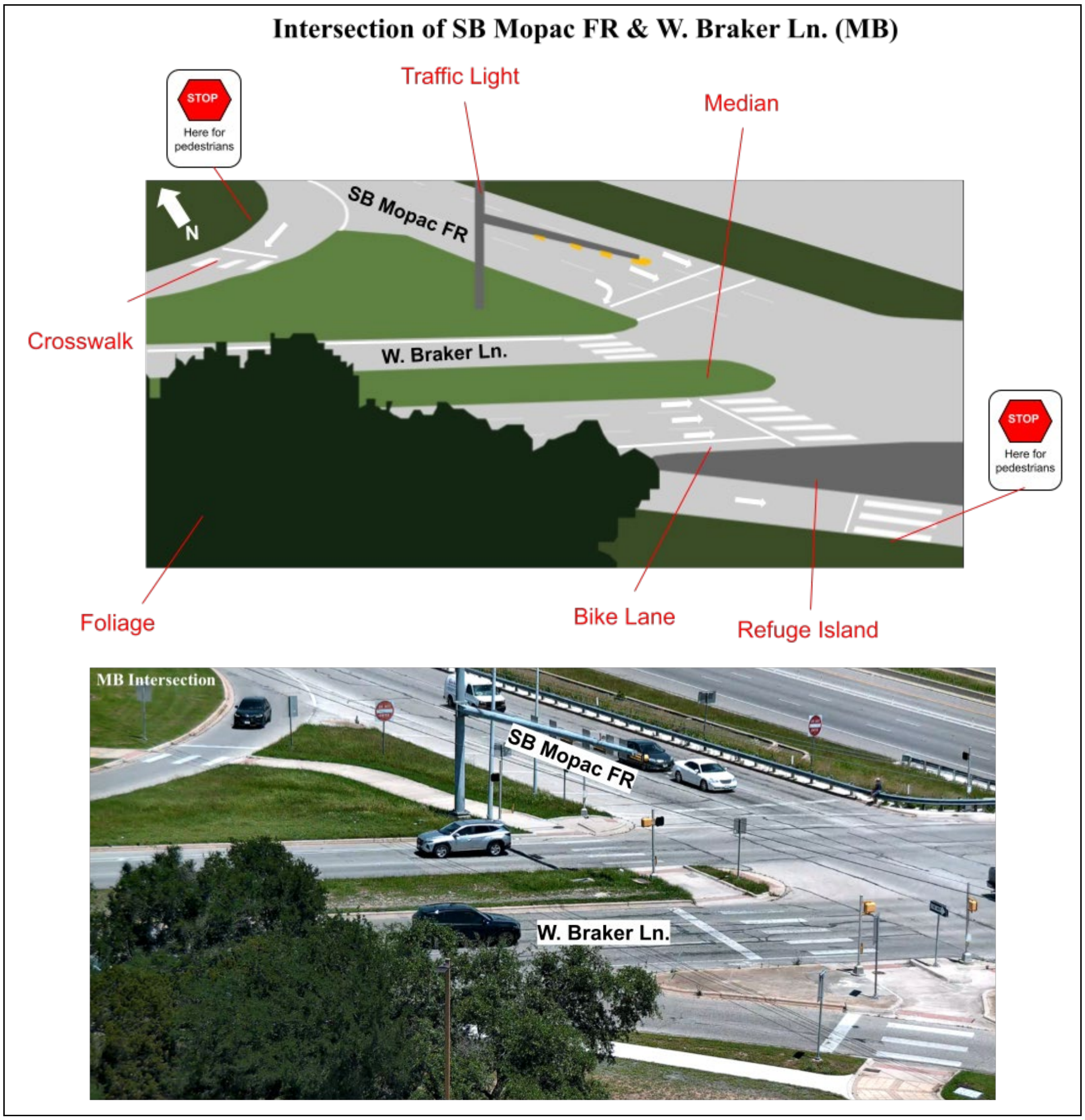

Figure 1. MB Intersection Layout

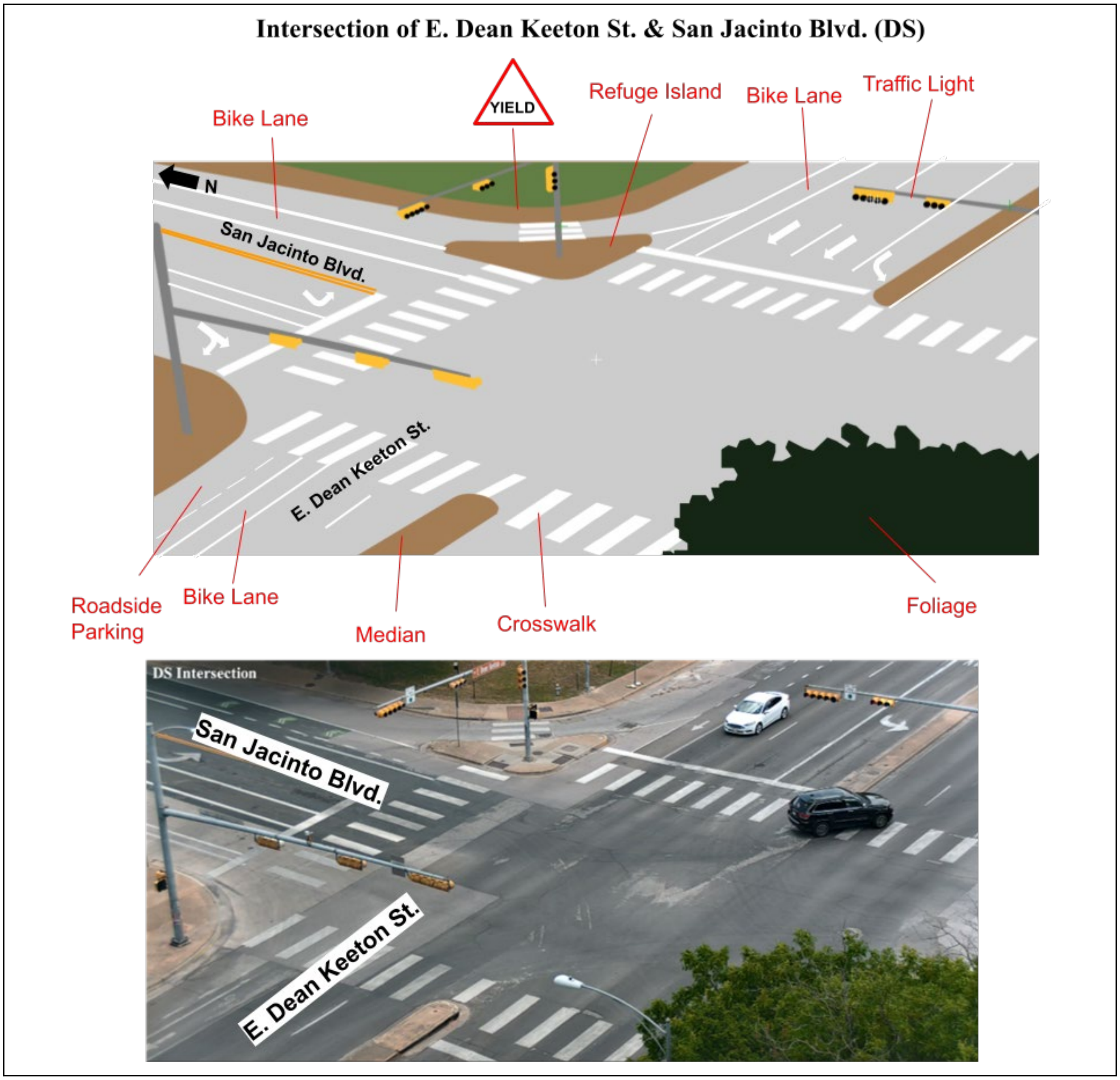

Figure 2. DS Intersection Layout

It is important to note that several contextual factors may have influenced the representativeness of the collected data. First, the DS intersection observation period fell outside the University of Texas at Austin's regular academic calendar. While this may have contributed to lower-than-usual pedestrian volumes at the DS site, the presence of summer classes ensured a consistent flow of student walking activity. Second, data collection occurred during late spring and early summer months, when elevated Texas temperatures may have introduced seasonal bias in pedestrian behavior. The average high temperature in Austin was 79.4°F during the MB intersection monitoring period and 93.6°F during the DS intersection observation period. Although the study avoided peak summer heat, elevated midday temperatures may have influenced pedestrian volume and temporal distribution patterns. In any case, and as discussed earlier, strict representativeness of foot and vehicle movements is not necessary to develop consistent relationships between exogenous covariates and the two outcomes of interest.

A final note related to our sample. In the original data collected, precipitation occurred during only 27.5 hours of the total 1,296 recorded hours (2.12% of total recorded hours). Due to very few NCPC and DUY instances during the few hours of periods of precipitation, we excluded these periods of precipitation from the sample for analysis. Dropping such a small percentage of observations should not lead to any substantial self-selection bias in the estimates for dry weather conditions.

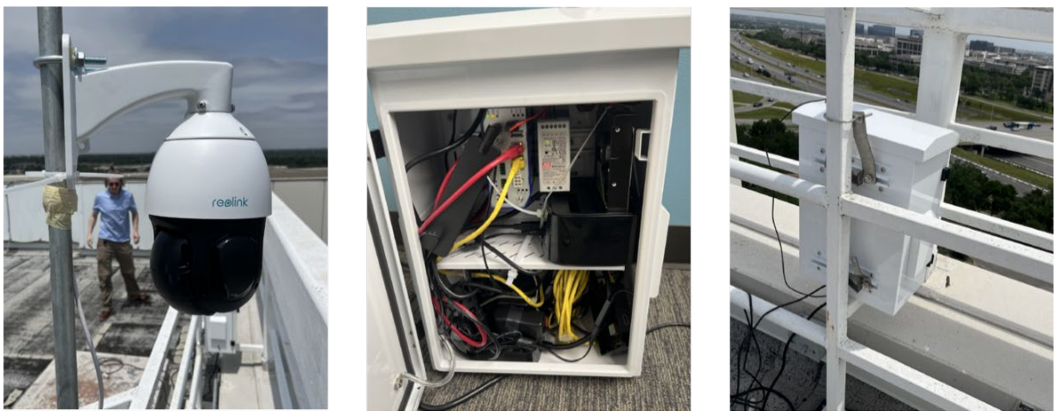

Figure 3. Video Recording Setup

3.3. Data Coding Process

Upon completion of data collection, trained research assistants conducted systematic video analysis to extract key parameters from individual pedestrian crossings and vehicle-pedestrian interaction events. Each video recording was reviewed frame by frame to ensure the accurate identification and measurement of the variables of interest. These variables included a range of pedestrian characteristics, vehicle characteristics, other contextual factors, and dynamics of crossing situations.

Pedestrian demographic characteristics were coded through systematic observation, with gender classification determined through visual assessment of morphological features and coded as male, female, or unknown when classification was not possible due to image quality or other observational constraints. Additionally, pedestrian age was estimated through visual assessment and categorized into three distinct groups: individuals appearing to be under 18 years were classified as “minors,” those appearing to be 18-40 years were categorized as “young adults,” and individuals appearing to be over 40 years were classified as “older adults.” While visual estimation of age may introduce classification errors, this method has been used in prior observational studies and is considered acceptable for behavioral fieldwork where direct demographic data are unavailable (see, for example, Brosseau et al., 2013, Aghabayk et al., 2021, and Schwebel et al., 2022). Skin tone classification was performed using the Monk Skin Tone (MST) Scale, which was developed by experts in social psychology, social categorization, and underrepresented communities (Monk, 2022). The scale captures socially meaningful variation in skin tone within a framework focused on social inequality, making it particularly well-suited for behavioral research on demographic disparities in pedestrian-vehicle interactions. The MST Scale comprises ten skin tone categories (Monk, 2022), each represented by a standardized color swatch used as a visual reference for coding pedestrians' skin tone from video footage. In practice, we found it difficult to reliably distinguish among intermediate skin tones (categories 3-8), while the extreme categories (1-2 and 9-10) were visually distinct and more consistently identified. Besides, as shown in Table 3, there were 1,133 NCPC instances for the grouped 1-2 category (8.2% of all pedestrians classified as 1-2), 129 (5.8%) for the grouped 3-8 category, and 143 (11.4%) for the grouped 9-10 category. In the subset of pedestrian-vehicle interactions used to model DUY behavior, the corresponding numbers of non-yielding instances were 279 (12.1%) for the 1-2 group, 60 (15.7%) for the 3-8 group, and 32 (16.2%) for the 9-10 group. Given the small number of NCPC and driver (un)yielding events within the collective 3-8 category, and the corresponding low splits of these events in the 3-8 category, more granular subdivisions would result in imprecise estimates and unstable model parameters. Accordingly, based on both coding reliability and outcome distributions, we grouped skin tones into three categories—White (1–2 MST), Brown (3–8 MST), and Black (9–10 MST)—which provided the most appropriate balance between interpretability and statistical stability for modeling purposes. Additionally, annotators flagged pedestrians exhibiting visual indicators of housing insecurity (VHI) using a structured binary checklist developed through an extensive literature review in sociology and homelessness studies (Goldfischer, 2018, Speer, 2019, Bowen and Capozziello, 2024, Long, 2024). The checklist comprised observable attributes across three categories: (1) appearance indicators (untidy or soiled clothing, multiple layers of clothing inconsistent with weather, disheveled appearance, lack of recent grooming), (2) possessions/accessories (carrying multiple bags or belongings simultaneously, use of shopping carts for personal belongings, display of solicitation signs), and (3) behavioral indicators (prolonged stationary presence in the intersection area beyond typical crossing wait times, actively soliciting donations from passing vehicles or pedestrians). A pedestrian was coded as exhibiting VHI if they displayed two or more indicators from at least two different categories. This multi-indicator, multi-category threshold was established through sensitivity analysis during annotator training to reduce false positives while maintaining adequate specificity. However, we acknowledge that reliance on visual evidence alone cannot definitively determine an individual's housing status and may be subject to misclassification. Accordingly, the VHI measure is treated as an observational proxy grounded in prior literature rather than a definitive identification of housing insecurity and should be interpreted with appropriate caution.

Additional contextual variables related to pedestrian behavior and environmental conditions were also recorded. Pedestrian activity level, classified as walking or running, was identified through direct observational judgment based on evident differences in gait and acceleration, similar to earlier published behavioral observation studies (Dommes et al., 2015, Zafri, 2023). Specifically, pedestrians who visually appeared to be engaged in recreational or fitness running, as well as individuals who briefly ran to cross before a signal change or while engaging in NCPC, were included in the "runner" category. Pedestrian distraction was characterized in our study as visibly engaging with a smartphone, such as texting or using phone applications while crossing, accompanied by limited visual attention toward the roadway (such as a downward gaze), similar to the characterization adopted in many earlier observational studies of pedestrian crossings (see Russo et al., 2018, Aghabayk et al., 2021, Bendak et al., 2021, Schwebel et al., 2022).[3] Instances of distracted crossing were recorded only when these behaviors were clearly discernible from video footage. The social context of group crossing dynamics was also assessed by identifying instances in which two or more individuals simultaneously occupied or queued for crosswalk access.

Time-of-day data were extracted from video timestamps and converted to standardized time-of-day categories during post-processing: Dawn (04:00-05:59), Morning (06:00-11:59), Noon (12:00-13:59), Afternoon (14:00-17:59), Dusk (18:00-19:59), Evening (20:00-21:59), and Night (22:00-03:59). This seven-category temporal scheme was selected through preliminary testing of alternative temporal aggregations, and reflects a balance between behaviorally meaningful lighting/traffic condition periods and sufficient observation counts to reliably identify time-of-day effects. Weather conditions, specifically precipitation, were also documented to account for potential weather-related influences on crossing behavior.

In instances where pedestrians crossed in the presence of a vehicle, the vehicle type was recorded as SUV, sedan, pickup truck, commercial (including box-trucks, company-branded cars and vans, garbage trucks, and city buses), or "other," (encompassing motorcycles, emergency vehicles, and any vehicles that could not be clearly identified). In the case of multiple vehicles arriving sequentially, the characteristics of the first vehicle to interact with the pedestrian were recorded.

Lastly, crossing dynamics variables, which constitute the primary outcome measures examined in this study, included Non-Compliant Pedestrian Crossing (NCPC) behavior and Driver Unyielding (DUY) behavior. NCPC was defined as a composite measure indicating that the pedestrian violated at least one of two statutory conditions. A temporal violation occurred when a pedestrian entered the roadway against a steady red, steady yellow, "Don't Walk," or "Wait" signal, as prohibited under Sections 552.001(c) and 552.002(c) of the Texas Transportation Code (2023). A spatial violation occurred when a pedestrian crossed outside a marked or unmarked crosswalk (Section 552.005(a)), or crossed midblock between adjacent signalized intersections where no marked crosswalk was present (Section 552.005(b)). DUY was assessed through systematic observation of pedestrian-vehicle interactions and was defined more broadly than statutory right-of-way violations to reflect both legal and safety considerations. A driver was coded as "unyielding" if they either (1) violated a legal obligation to yield, such as failing to yield to a pedestrian lawfully crossing with a "Walk" signal (Section 552.002(b)) or at a location controlled by a Stop or Yield sign (Section 545.153(b)), or (2) failed to exercise reasonable care to avoid a pedestrian already in the roadway, consistent with the general duty of due care under Section

[3] Some of these earlier studies have also used indicators of audio-based distraction of pedestrians, relying on the visual identification of whether or not the pedestrian was wearing headphones (Russo et al., 2018, Aghabayk et al., 2021, Schwebel et al., 2022). We steered clear of doing so in our study because in-ear devices are now common, and the use of such devices could not be reliably identified from video footage.

552.008. This broader operationalization allows DUY behavior to be recorded even when a pedestrian was in violation of crossing laws, reflecting real-world instances in which both parties may contribute to conflict risk.

To ensure inter-rater reliability and good data quality, all research assistants underwent standardized training protocols designed to promote consistent variable categorization across the coding process, including under varying lighting conditions. Training included practice coding from both daytime and nighttime footage to establish reliable decision rules for low-visibility scenarios and to calibrate observers' thresholds for classifying observations as "unknown" when visibility was insufficient. Regular calibration meetings were conducted throughout the data collection period to maintain coding consistency and address classification challenges. Additionally, for ambiguous cases requiring subjective interpretation, a consensus-based approach was implemented, whereby four observers provided independent assessments before reaching final coding decisions.[4] When consensus could not be achieved for variables with inherent classification difficulties (e.g., gender, age, or skin tone), or when limited visibility rendered reliable classification difficult, observers were instructed to record "unknown" classifications. These cases were subsequently excluded from the analysis to maintain analytic integrity and avoid introducing misclassification error.

3.4. Sample Characteristics

Following data collection, a systematic data cleaning procedure was implemented to ensure data quality and analytical validity. Observations containing incomplete records or variables coded as "unknown" were excluded from the analysis. After cleaning, 1,158 pedestrian crossing instance observations (6.6% of the sample) included at least one "unknown" classification based on either age (2.7% of observations classified as "unknown" age), gender (2.8% classified as "unknown" gender), or skin tone (4.5% classified as "unknown" skin tone). As expected, "unknown" classifications were more common under limited lighting conditions, occurring in 13.3% of observations (316 cases) when dark relative to 5.3% of observations (842 cases) when not dark. This distribution reflects anticipated visibility (or, to be precise, lack thereof) conditions. To more accurately represent group instances in the pedestrian-vehicle interaction subset, heterogeneous groups of two or more pedestrians interacting with vehicles at the same time were assigned gender, age, and skin tone classifications based on the proportions observed within the sample. This approach avoids the exclusion of group observations that would otherwise have been dropped due to mixed characteristics. The final sample used in model estimation comprises 17,374 pedestrian crossing observations, of which 2,890 observations (16.0%) correspond to instances involving pedestrian-vehicle interactions. Table 3 presents the descriptive statistics for all outcome and explanatory variables included in the analysis. The table structure accommodates the hierarchical

[4] These structured training and consensus-based procedures are similar to those adopted in earlier observational pedestrian safety studies employing subjective morphological classifications (Ivan et al., 2017, Schneider et al., 2018, Aghabayk et al., 2021, Zafri et al., 2022, Rafe et al., 2025).

nature of the data, where pedestrian-vehicle interactions represent a subset of all observed pedestrian crossings. For each variable, four statistical measures are reported:

(a) Total Number of Observations (Relative Frequency): The absolute frequency and relative frequency of each variable category. In calculating the relative frequency, the denominator varies based on variable applicability. Variables specific to pedestrian-vehicle interactions (such as yielding behaviors and vehicle type) are calculated based on 2,890 total interactions, while general pedestrian variables applicable to all crossings are calculated based on 17,374 total observations.

(b) Vehicle Interaction Rate: The proportion of observations for each variable category that involved a pedestrian-vehicle interaction, computed as the number of observations with vehicle interaction divided by the total number of observations for that variable category.

(c) NCPC Rate: The proportion of observations for each variable category that exhibited NCPC behavior, calculated as the number of NCPC observations divided by the total number of observations for that variable category.

(d) DUY Rate: The proportion of observations for each variable category where the driver failed to yield to the pedestrian, computed as the number of non-yielding instances divided by the total number of pedestrian-vehicle interactions for that variable category.

The analysis of outcome variables (top panel of Table 3) reveals distinct patterns in pedestrian crossing compliance and driver yielding behavior. NCPC behavior was observed in 8.1% of all pedestrian crossings (N=17,374). Among pedestrians exhibiting NCPC behavior, 8.3% involved vehicle interactions. Regarding DUY behavior, non-yielding incidents occurred in 12.8% of all pedestrian-vehicle interactions (N=2,890). Additionally, DUY was higher at 34.5% in vehicle-pedestrian interactions when the pedestrian exhibited NCPC behavior, compared to 11.9% when the pedestrian did not exhibit NCPC behavior.

The bottom panel of Table 3 summarizes the distribution of explanatory variables in the sample. In the category of pedestrian sociodemographic variables, male pedestrians accounted for a greater share of observations compared to female pedestrians (67.1% vs. 32.9%, respectively). While a slightly higher proportion of female pedestrians were involved in vehicle-pedestrian interactions, both the NCPC rate and the non-yielding rate were higher among male pedestrians. Regarding age, young adults were overrepresented in the sample, which is consistent with the study location on or near a university campus. The vehicle interaction rate and NCPC rate were comparable between young adults and older adults. However, the DUY rate toward older adult pedestrians was significantly higher, reaching 22.5%. Skin tone distribution analysis revealed that the majority of observed pedestrians had lighter skin tones, though pedestrians with dark (Black) skin tones (MST 9-10) exhibited the highest rates of NCPC behavior (11.4%) and experienced the highest driver non-yielding rates (16.2%), indicating potential disparities in both pedestrian behavior and driver response patterns across different demographic groups. Although pedestrians exhibiting visual markers of housing insecurity (VHI) represented fewer than 2% of the total sample, this subgroup demonstrated disproportionately high rates of risky interactions.

Specifically, 21.1% of individuals with VHI engaged in NCPC behavior, and 39.7% were subject to DUY behavior during vehicle-pedestrian interactions.

The statistics for pedestrian activity and contextual variables in Table 3 show that most pedestrians crossed alone rather than in groups. Descriptive statistics show that solo crossers had a vehicle interaction rate of 15.5%, compared to 25.9% for group crossers. However, solo crossers exhibited a higher NCPC rate (8.4%) than those crossing in groups (5.6%). The DUY rate was slightly higher for solo crossings, at 13.3% for pedestrians crossing alone and 10.5% for pedestrians crossing in groups. Distracted pedestrians made up only a small share of all crossers, and NCPC and DUY behaviors were largely comparable to those of non-distracted pedestrians. Also, the majority of observed pedestrians were walking, with runners comprising only 7.1% of the sample. However, runners exhibited a higher rate of NCPC behavior and were less likely to be yielded to by drivers.

The time-of-day and weather variable statistics in Table 3 indicate that the majority of pedestrian crossings occurred during the morning and afternoon periods. The data indicate notable temporal overlap between periods of heightened NCPC instances and increased DUY. Specifically, both behaviors were more frequently observed during the morning, evening, and nighttime hours, indicating a potential convergence of risk during periods of low visibility and/or high traffic volumes. Lastly, fewer than 1% of observations occurred during rainfall. Although the DUY rate appeared higher in these conditions compared to dry weather, this descriptive statistic is based on only 10 instances of non-yielding in the rain *versus* 361 in non-rain conditions. Given the small sample size, no meaningful conclusions can be drawn regarding the effect of rain on DUY behavior, and so we do not consider weather conditions in our estimation.

Finally, in the category of vehicle characteristics, among all observed vehicle-pedestrian interactions, SUVs were the most frequently observed vehicle type, accounting for 42.6% of cases. This was followed by sedans (38.4%), pickup trucks (11.3%), commercial vehicles (5.6%), and other vehicle types (2.1%). The data suggest that NCPC rates were higher in the case of pedestrian interactions with pick-up trucks and commercial vehicles (relative to SUVs, sedans, and other vehicle types). In contrast, DUY behavior was lowest among commercial vehicle drivers but highest among drivers of "other" vehicle types. However, due to the small number of observations in some vehicle categories, these results should be interpreted with caution.

Overall, the descriptive statistics offer insight into the distributional characteristics of the data and preliminary behavioral patterns across different variable categories. However, these are univariate relationships of a single exogenous variable with each of the two binary endogenous outcomes (NCPC and DUY), without recognizing potential unobserved characteristics jointly affecting the two outcomes and without controlling for the effects of multiple variables at the same time. In this context, it is important to consider a bivariate and multivariable analytic approach, as discussed next.

Table 3. Sample Descriptive Statistics

Variable	Total Number of Observations (Rel. Freq.)	Vehicle Interaction Rate (%)	Non-Compliant Pedestrian Crossing Rate (%)	Non-Yielding Rate (%)
Outcome Variables				
Pedestrian Crossing Behavior				
CPC	15969 (91.9%)	17.4%	0.0%	11.9%
NCPC	1405 (8.1%)	8.3%	100.0%	34.5%
Yielding	2519 (87.2%)	100.0%	3.0%	0.0%
Non-Yielding	371 (12.8%)	100.0%	10.8%	100.0%
Explanatory Variables				
Pedestrian Sociodemographic Variables				
Pedestrian Perceived Gender				
Female	5713 (32.9%)	17.7%	6.1%	11.5%
Male	11661 (67.1%)	16.1%	9.0%	13.6%
Pedestrian Perceived Age				
Minor	39 (0.3%)	23.1%	5.1%	44.4%
Young adult	14949 (86.0%)	16.5%	8.0%	11.1%
Older	2386 (13.7%)	17.4%	9.0%	22.5%
Pedestrian Perceived Skin Tone				
MST 1-2 (White)	13887 (79.9%)	16.6%	8.2%	12.1%
MST 3-8 (Brown)	2233 (12.9%)	17.1%	5.8%	15.7%
MST 9-10 (Black)	1254 (7.2%)	15.8%	11.4%	16.2%
Pedestrian Exhibiting Visual Markers of Housing Insecurity (VHI)				
No VHI identified	17099 (98.4%)	16.6%	7.9%	12.3%
VHI identified	275 (1.6%)	21.1%	21.1%	39.7%
Pedestrian Activity and Contextual Variables				
Social Context				
Solo crossing	15453 (88.9%)	15.5%	8.4%	13.3%
Group crossing	1921 (11.1%)	25.9%	5.6%	10.5%
Pedestrian Distraction				
Not distracted	17070 (98.3%)	16.7%	8.1%	12.8%
Distracted	304 (1.7%)	15.5%	5.3%	12.8%
Pedestrian Activity Type				
Walker	16134 (92.9%)	17.0%	7.5%	12.7%
Runner	1240 (7.1%)	11.5%	15.6%	16.2%

Variable	Total Number of Observations (Rel. Freq.)	Vehicle Interaction Rate (%)	Non-Compliant Pedestrian Crossing Rate (%)	Non-Yielding Rate (%)
Time-of-Day and Weather Variables				
Time-of-Day				
Dawn (04:00-05:59)	100 (0.6%)	6.0%	22.0%	0.0%
Morning (06:00-11:59)	6362 (36.6%)	13.8%	8.7%	15.5%
Noon (12:00-13:59)	2256 (13.0%)	17.7%	5.9%	10.8%
Afternoon (14:00-17:59)	4483 (25.8%)	20.3%	5.5%	12.5%
Dusk (18:00-19:59)	2090 (12.0%)	20.5%	8.0%	9.1%
Evening (20:00-21:59)	1189 (6.8%)	17.1%	9.3%	14.8%
Night (22:00-03:59)	894 (5.2%)	7.4%	19.4%	13.6%
Weather Condition				
Raining	104 (0.1%)	26.0%	9.6%	37.0%
Not raining	17270 (99.9%)	16.6%	8.1%	12.6%
Vehicle Characteristics				
Vehicle Type				
SUV	1231 (42.6%)	100.0%	3.1%	11.6%
Sedan	1109 (38.4%)	100.0%	4.1%	13.1%
Pickup truck	327 (11.3%)	100.0%	5.5%	14.4%
Commercial*	162 (5.6%)	100.0%	5.6%	10.0%
Other*	61 (2.1%)	100.0%	9.8%	29.5%

*Commercial vehicles include box-trucks, company-branded cars and vans, garbage trucks, and city buses. "Other" vehicle types include motorcycles, emergency vehicles, and any vehicles that could not be clearly identified.

4. ANALYTIC APPROACH

4.1. Copula-based bivariate logit model

This study employs binary logit discrete outcome models to analyze two distinct behaviors: (1) pedestrian crossing behavior (Compliant Pedestrian Crossing (CPC) *versus* Non-Compliant Pedestrian Crossing (NCPC)) and (2) driver yielding behavior (driver yields properly to pedestrians (DYP) *versus* driver unyielding to pedestrians (DUY)). The binary logit model (sometimes also referred to as a logistic regression) is based on a latent variable framework, where the observed binary outcome is determined by an underlying continuous latent propensity (Train, 2009). In this paper, we code the two binary variables as follows: (1) for pedestrian crossing behavior, $y_{q1}=1$ if the pedestrian in observation *q* crosses non-compliantly (NCPC) and $y_{q1}=0$ if the pedestrian in observation *q* crosses compliantly (CPC), and (2) for driver yielding behavior, $y_{q2}=1$ if the driver associated with an observation *q* (fails to yield (DFY), and $y_{q2}=0$ if the driver associated with an observation *q* yields properly (DYP). To allow for potential jointness in NCPC and DUY behaviors (due to unobserved factors, as discussed in Section 2.3), we use a copula-based bivariate logit model. Note, however, that this jointness only applies to the 2,890 observations that correspond to pedestrian-vehicle interactions. For the 14,484 observations where there is no pedestrian-vehicle interaction, only the NCPC outcome is observed. The model is estimated using the maximum likelihood inference approach, in which the likelihood contribution of the 14,484 observations without pedestrian-vehicle interactions (denote this set of observations as C_1) corresponds simply to the univariate probability of the observed y_{q1} value, while the likelihood contribution of the remaining 2,890 observations with pedestrian-vehicle interactions (denote this set of observations as C_2) corresponds to the bivariate joint probability of the observed (y_{q1}, y_{q2}) values. Let the continuous latent propensities underlying the binary outcomes y_{q1} and y_{q2} be y_{q1}^* and y_{q2}^*, respectively. In notation form, the model then takes the following form:

$$\begin{aligned} &y_{q1}^* = \boldsymbol{\beta}_1' \mathbf{x}_{q1} + \varepsilon_{q1}, y_{q1} = 1 \text{ if } y_{q1}^* > 0 \text{ and } y_{q1} = 0 \text{ if } y_{q1}^* \le 0 \\ &y_{q2}^* = \boldsymbol{\beta}_2' \mathbf{x}_{q2} + \delta y_{q1} + \varepsilon_{q2},\ y_{q2} = 1 \text{ if } y_{q2}^* > 0 \text{ and } y_{q2} = 0 \text{ if } y_{q2}^* \le 0;\ y_{q2} \text{ observed only if } q \in C_2, \end{aligned} \tag{1}$$

with $\mathbf{x}_{q1}$ and $\mathbf{x}_{q2}$ being covariate vectors for observation q (these vectors include constants), $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ being corresponding coefficient vectors, and ε_{q1} and ε_{q2} being error terms that are each assumed to be standard logistic and independent and identically distributed across observations. We allow the binary NCPC and DUY outcomes to be co-determined (due to correlated unobserved factors) by tying the ε_{q1} and ε_{q2} error terms using a Gaussian copula (Bhat and Eluru, 2009). That is, the bivariate cumulative distribution function (CDF) of ε_{q1} and ε_{q2} is specified as:

$$F_{\varepsilon_{q1},\varepsilon_{q2}}(z_1, z_2) = \Phi_2\left(\left[\Phi^{-1}(F_{\varepsilon_{q1}}(z_1)\right], \left[\Phi^{-1}(F_{\varepsilon_{q2}}(z_2)\right], \rho\right), \tag{2}$$

where $\Phi_2(.)$ is the bivariate normal CDF, Φ^{-1} is the inverse CDF of the univariate standard normal, $F_{\varepsilon_{qi}}(.)$ and $F_{\varepsilon_{q2}}(.)$ represent the CDF of the standard logistic distribution (that is, $F_{\varepsilon_{q1}}(z_1) = 1/(1+e^{-z_1})$ and $F_{\varepsilon_{q2}}(z_2) = 1/(1+e^{-z_2})$), and ρ is a dependence parameter. Finally, δ in Equation (2) represents the "true" causal effect of NCPC behavior on driver unyielding, after controlling for any associative effects between the two outcomes due to error correlation effects. Not controlling for these associative effects, when present, would render all the coefficient estimates of the second DFY equation inconsistent due to the classic endogenous dummy variable problem (that is, if ε_{q1} and ε_{q2} are correlated and this is not accounted for, this would lead to a correlation between y_{q1} and ε_{q2} in the second DFY equation above). In the bivariate logit system of Equation (1), for stability and identification, we ensure that there is at least one variable ("instrument") that is contained in the vector $\mathbf{x}_{q1}$, but does not appear in the vector $\mathbf{x}_{q2}$ (see Han and Lee, 2019 and Bhat, 2024 for extensive treatments of identification considerations for models of the type in Equation (1)).[5]

To estimate the model parameters, define $\theta_{q1} = \left(1 - 2y_{q1}\right), \theta_{q2} = \left(1 - 2y_{q2}\right)$, $w_{q1} = \theta_{q1} \times \left[\Phi^{-1}(F_{\varepsilon_{q1}}(-\boldsymbol{\beta}_1'\mathbf{x}_{q1}) \right]$, $w_{q2} = \theta_{q2} \times \left[\Phi^{-1}(F_{\varepsilon_{q2}}(-\boldsymbol{\beta}_2'\mathbf{x}_{q2} - \delta y_{q1}) \right]$, and $\tau_q = \theta_{q1}\theta_{q2}\rho$. Then, using the symmetric property of the bivariate normal density function, the log-likelihood function takes the following form:

$$\log L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho) = \sum_{q \in C_1} \log\left[F_{\varepsilon_{q1}}\left(\theta_{q1}(-\boldsymbol{\beta}_1'\mathbf{x}_{q1}) \right) \right] + \sum_{q \in C_2} \log\left[\Phi_2(w_{q1}, w_{q2}, \tau_q) \right]. \quad (3)$$ [6]

The parameters $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, along with the dependence parameter ρ (as estimated by maximizing the log-likelihood function above) provide the estimated effects of variables on the continuous latent propensities y_1^* and y_2^*, rather than on the binary outcomes y_1 and y_2. At the same time, the effects of a specific exogenous variable on the binary outcomes are not constant across individuals, but vary based on the values of other variables for an observation as well as the value of the variable itself for the observation. That is, the effect of any variable on any binary outcome varies across individuals. To summarize these heterogeneous effects, we compute average treatment effects (ATEs), defined as the change in the probability of each binary outcome when an explanatory variable is shifted from a baseline level A to a treatment level B, holding all other variables constant. Given the recursive structure of the bivariate system, we compute, for each individual and for both levels A and B, the joint probabilities of all possible bivariate outcome combinations, regardless of the observed outcomes. From these joint probabilities, we obtain the marginal outcome shares for each dimension. The ATEs are then calculated as the percentage change, from baseline to treatment, in the shares of the NCPC and DUY outcomes.

---

[5] Instead of the Gaussian copula-based bivariate logit model, one could also use a bivariate probit model in which the error terms ε_{q1} and ε_{q2} are assumed to be normally distributed. This would obviate the need for the use of a copula structure. In terms of computation ease and substantive results, both these approaches are comparable and it then becomes more of a preference for one structure relative to the other. Here, we adopt the copula-based approach because, in the univariate case (which is the situation for a large number of observations with no pedestrian-vehicle interactions), the copula approach allows us to preserve the univariate logistic model with which we started our analysis for each of the NCPC and DFY equations.

[6] Random coefficients may be introduced in the usual way by assuming that some of the elements of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are realizations from specific distributions (in our study, we assumed normal distributions), and integrating the probabilities over the assumed density functions. Halton draws are used, as suggested by Bhat (2001, 2003). This is a straightforward extension that we do not discuss here. However, we should also note that estimating random coefficients on discrete exogenous variables, especially those that do not vary across the alternatives characterizing choice models, can be fraught with estimation and identification challenges, an issue that has only recently started receiving some attention (Bharate et al., 2025). In the current paper, all variables are observation-specific and do not vary across alternatives, which led to estimation instability when we attempted random coefficients on several exogenous variables. And in cases where we were able to estimate random coefficients on some discrete exogenous variables, none of these came out to be statistically significant at even the 68% confidence level (t-statistic of 1.0).

4.2. Model Specification Testing

The selection of explanatory variables in the binary logit models for the two outcomes followed a systematic approach that balanced theoretical relevance with statistical robustness. Since all exogenous variables were categorical variables, we examined, based on the descriptive statistics, whether there were enough observations in each category of each exogenous variable, especially whether there were adequate observations in each category corresponding to each of the two states of each binary outcome variable (for example, we did not consider weather conditions for this reason, as mentioned earlier, and also did not consider the pedestrian distraction variable). The issue of adequate observations is particularly relevant for the binary model of driver unyielding behavior due to the limited number of observations (2,890 observations, with only 371 observations of DUY) compared to NCPC behavior (17,374 individuals with 1,405 observations exhibiting NCPC behavior). Beyond these first-level exclusions based on sample sparsity, all available exogenous variables in the dataset, and their interactions, were considered as potential explanatory variables. This included systematic testing of interactions among all sociodemographic variables (race, gender, age, housing insecurity status) and temporal variables (time-of-day categories) to examine whether behavior varied across contexts. The model specification process involved extensive exploration of alternative functional forms and variable combinations. For naturally discrete variables such as group size, we tested the most disaggregated dummy variable specifications and progressively combined categories based on statistical significance tests to achieve model parsimony. The binary specification distinguishing group *versus* solo pedestrians proved the most effective. For categorical variables, including age groups, skin tone, pedestrian activity type, time-of-day, and vehicle type, we similarly began with the most disaggregated form of data collected and systematically combined categories based on statistical significance and theoretical meaningfulness. The site location variable and its interactions with other exogenous covariates were not statistically significant in any of our models, indicating that site-specific factors did not systematically influence the outcomes after accounting for other covariates. Time-of-day effects were introduced in the seven category grouping discussed in Section 3.3, though interaction effects of time-of-day with other variables (specifically with gender and skin tone for the NCPC outcome, and with gender for the DUY outcome) were statistically significant only through a combined specification merging night and dawn periods into a single time period (referred to as "Nighttime").

As discussed in Section 4.1, we considered a recursive bivariate specification allowing for correlation in the unobserved error components across the NCPC and DUY outcome latent propensities, while also considering a structural effect of the NCPC outcome (y_{q1}) on the DUY outcome propensity (y_{q2}^*). But several empirical diagnostics pointed to the structural effect being adequate to capture the relationship between the two outcomes. First, when we estimated both the structural NCPC effect and a cross-equation error correlation parameter, both parameters turned out to be statistically insignificant and exhibited an extremely high covariance (exceeding 0.990). This indicates the presence of a likelihood ridge along which the structural effect and the correlation parameter compete to explain the same empirical association -- behavior that is characteristic of weak separation between structural and unobserved-correlation channels in recursive nonlinear models. This issue is particularly pronounced in our empirical context, as NCPC behavior is observed in only 116 of the 2,890 observations involving vehicle interactions. Second, restricting the correlation parameter to zero, while allowing the structural NCPC effect, resulted in the precise and highly statistically significant estimate of the NCPC coefficient. Also, the log-likelihood at convergence under this specification was marginally better than that obtained under the alternative specification that excluded the structural NCPC effect and allowed only for correlated errors. Third, when we excluded the NCPC variable and freely estimated the correlation parameter, the estimated correlation itself was statistically insignificant. This finding indicates that the data does not support explaining the relationship between the two outcomes solely through correlated unobservables. Moreover, the estimated coefficients and standard errors for the remaining DUY covariates were very similar across the two specifications (structural NCPC effect with zero error correlation and zero structural effect with error correlation). Taken together, these results indicate that the relationship between the two outcomes is best represented as a direct structural effect of NCPC on DUY, rather than as correlation in unobserved components across equations. Accordingly, the results discussed below correspond to the specification of a direct structural effect with zero cross-equation error correlation.

Finally, we assessed multicollinearity and parameter identification by examining the correlation matrix of the predictors, none of whose elements exceeded 0.34. We further evaluated the information matrix (the negative Hessian of the log-likelihood) and the resulting variance–covariance matrix of the parameter estimates, which are the appropriate diagnostics for identification and multicollinearity in maximum-likelihood estimation (Greene, 2018a), rather than variance inflation factors derived from linear regression. In all estimations, the information matrix was nonsingular and optimization converged smoothly, with relatively tight standard errors. The implied correlation matrix of the parameter estimates showed no evidence of problematic dependence: most off-diagonal correlations were in the 0.05–0.35 magnitude range, and the largest absolute correlation was 0.76. Overall, these diagnostics indicate that the model parameters are well identified and not subject to problematic multicollinearity.

5. EMPIRICAL RESULTS AND DISCUSSION

5.1. Model Estimation Results

The estimation results from the final model specification are presented in Table 4. The first three columns, under the heading "Non-Compliant Pedestrian Crossing Model," display results from the binary NCPC logit model. The final three columns, labeled "Driver Unyielding Model," present results from the binary DUY logit model. For each model, we report the estimated coefficients, their corresponding t-statistics, and the percentage Average Treatment Effects (%ATE) of exogenous variables. For the NCPC model, we retained coefficients that were statistically significant at the 90% confidence level (|t-statistic| > 1.65). But we used a lower confidence level of 68% (|t-statistic| > 1.00) for the DUY model because of the smaller sample size of only 2,890 cases with pedestrian-vehicle interactions as well as the small fraction of these 2,890 cases in which the pedestrians were not yielded to (only 371 of 2,890 cases, or 12.4%).[7] Additionally, note that a dash ("--") in Table 4 indicates that the corresponding variable was not statistically significant in the model, while "*n.a.*" denotes that the explanatory variable is not applicable to the given outcome variable.

In the subsequent discussion, we focus on the %ATE values, as they offer an intuitive interpretation of both the magnitude and direction of effects. While the estimated coefficients reflect the impact of exogenous variables on the underlying latent propensity (i.e., y_q^* in Section 3.4.1), they do not directly reflect changes in the observed outcome (y_q). In contrast, the %ATE represents the percentage change in the predicted probability (or share) of an outcome when shifting from the base category to the treatment category of a given variable. For illustration, consider a pedestrian with visible markers of housing insecurity (VHI). The %ATE of 149.1% indicates that the share of pedestrians with VHI displaying NCPC behavior is 149.1% higher than the share of pedestrians with no VHI displaying NCPC behavior, other factors being the same. Another way of looking at this is that a randomly selected person with VHI is 149.1% more likely to exhibit NCPC behavior than a randomly selected person without VHI; that is, if 10 out of 100 pedestrians with non-VHI exhibit NCPC behavior, approximately 24.91 (about 25) out of 100 pedestrians with VHI would be expected to do so ([(0.2491-0.10)/0.10]*100=149.1%). More simply, we will just say that a person with VHI is 149.1% more likely to exhibit NCPC behavior than a person without VHI. Additional nuances arise when interpreting %ATE values in the presence of interaction effects. Consider the main effects of gender and time-of-day, alongside their interaction in the NCPC model. The main model coefficient effect of gender, which is positive, indicates a higher likelihood of NCPC behavior among males compared to females <u>during non-nighttime periods</u> (i.e., morning, dusk, and evening), holding other variables constant. However, this gender effect is moderated (reduced) at nighttime (defined as a combination period of dawn and night) by the negative coefficient on the "Male × Nighttime" interaction term, as presented in Table 4 under "Pedestrian Demographics and Time-of-Day Interactions." The interpretation becomes more complex when considering additional interactions between the nighttime period and pedestrian skin tone. A technically rigorous interpretation would involve jointly evaluating combinations of gender, time-of-day, and skin tone to compute ATEs. However, for clarity and interpretability, we separate the discussion of main effects from that of interaction effects. Main effects represent sample-averaged treatment effects and are calculated while accounting for the full model specification, including interaction terms. For example, the main gender %ATE is derived by simulating a shift from female (base category) to male (treatment category) pedestrians across the entire sample, holding other covariates constant and incorporating all coefficients. Similarly, %ATEs for dawn and night periods are computed by transitioning the

[7] As discussed at length by the American Statistical Association (ASA) (Wasserstein and Lazar, 2016, Wasserstein et al., 2019), analysts need to exercise context-specific judgements related to confidence levels rather than adhering strictly to a 95% confidence level as some kind of an absolute gold standard. In the current application, our focus in the DUY model was on reducing the probability of Type II errors (that is, incorrectly rejecting variable effects) even if allowing for a slightly higher probability of Type 1 error (that is, using an 68% confidence level to retain variables). This can inform future specifications using larger and possibly more balanced samples.

full sample from the noon/afternoon base category to the respective treatment period, across all gender and skin tone subgroups. Based on these computations, Table 4 shows that, on average, male pedestrians are 49.0% more likely than female pedestrians to exhibit NCPC behavior. Pedestrians crossing during dawn and night periods are 178.1% and 299.6% more likely, respectively, to exhibit NCPC behavior compared to those crossing during the noon/afternoon period. On the other hand, interaction effects capture how the marginal impact of one variable varies depending on the level of another. These %ATEs are computed by comparing specific subgroups to isolate conditional relationships. For instance, the %ATE for the "Male × Nighttime" interaction compares female pedestrians crossing at dawn or night (base category) to male pedestrians crossing in the same period (treatment category). Given the negative sign of the interaction coefficient, the %ATE of 0.2% for the "Male × Nighttime" interaction variable indicates that the gender effect is reduced at nighttime relative to the overall gender %ATE of 49.0% across all time periods.

The results presented in Table 4 indicate that a range of factors, spanning pedestrian sociodemographics, activity context, and time-of-day factors, significantly influence both pedestrian crossing behavior and driver yielding decisions. Overall, the %ATEs in Table 4 suggest that dawn and night period pedestrian crossings are most strongly associated with NCPC, followed by pedestrians exhibiting visible indicators of housing insecurity (VHI), evening crossings, runners, and men during non-nighttime periods. For DUY behavior, the strongest predictor is pedestrian non-compliance, followed by VHI status, older pedestrian age, and male pedestrians crossing at night. It is also important to note that the effects of exogenous variables (and the corresponding %ATEs) in the driver non-yielding behavior model represent direct effects **after** controlling for any effects of exogenous variables through the NCPC outcome. Thus, for example, the %ATE for male pedestrians in Table 4 corresponding to driver non-yielding indicates that, across all time periods, Black pedestrians are 25.0% less likely to be yielded to relative to White pedestrians (everything else remaining the same). This effect is not because Black pedestrians are more likely to exhibit NCPC behavior, because NCPC behavior is controlled for as a determinant variable in the driver non-yielding model. This effect cannot be explained by a higher propensity among pedestrians identified as Black to engage in NCPC behavior, as the DUY model explicitly includes NCPC as a covariate, thereby isolating the direct effect of skin tone on driver yielding behavior.

5.1.1. Pedestrian Sociodemographic Variable Effects

Pedestrian Perceived Gender

The model results reveal that male pedestrians are 49.0% more likely to engage in NCPC behavior compared to female pedestrians, suggesting significant gender-related differences in risk-taking behavior. This finding, which aligns with existing literature (e.g., Xie et al., 2018, Zhu et al., 2021, Rafe et al., 2025), could be a result of several psychological and sociocultural reasons. For instance, previous psychology and personality/gender studies (Reniers et al., 2016, Blanch and Martínez, 2025) have observed that men exhibit higher levels of sensation-seeking and impulsivity, which lead them to prioritize immediate rewards over long-term consequences, thereby increasing their likelihood of engaging in risky behaviors. Additionally, peer influence, media portrayals, and societal norms of masculinity emphasizing toughness, competition, and dominance can encourage men to adopt riskier behaviors (Morgenroth et al., 2018, and Dellosa and Browne, 2024). Overall, the observation that male pedestrians are more likely to engage in risky behaviors, coupled with their increased vulnerability during nighttime crossings discussed in Section 4.1.3, provides a plausible explanation for their overrepresentation in pedestrian crash statistics (McGuckin et al., 2018, U.S. Department of Transportation, 2024). These findings suggest that effective safety interventions must extend beyond traditional enforcement approaches to address the behavioral and sociocultural factors underlying male risk-taking and driver perceptions. Campaigns that leverage peer influence and temper traditional masculinity norms (i.e., emphasizing responsibility and caution over bravado), as well as underscore the vulnerability of pedestrians regardless of gender may be particularly effective in reshaping attitudes toward safer street behavior on the part of both pedestrians and motorists.

Pedestrian Perceived Age

Although age did not significantly influence pedestrian crossing compliance, older (over 40 years) pedestrians experienced an 84.0% higher probability of driver failure to yield compared to younger (≤40 years) individuals. While this contradicts some previous studies, our findings may be explained by the slower walking speeds of older individuals due to mobility limitations or a higher risk of falling (Avineri et al., 2012, Brosseau et al., 2013, and Liu and Tung, 2014). These reduced speeds could lead to drivers losing patience when waiting for older pedestrians to cross. Addressing this issue requires design changes in areas with older populations. Design enhancements, such as extended crossing times, signalized midblock crossings, curb extensions, and tactile and auditory cues, can reduce exposure and support safer mobility. These physical changes should be paired with public awareness campaigns encouraging drivers to exercise additional caution and patience, reinforcing pedestrians' right to safe and accessible mobility.

Table 4. Model Estimation Results

Variable	Non-Compliant Pedestrian Crossing (NCPC) Model			Driver Unyielding (DUY) Model		
	Coef.	t-stat	%ATE	Coef.	t-stat	%ATE
Pedestrian Sociodemographic Variables						
Pedestrian Perceived Gender (Base: Female)						
Male	0.443	6.10	49.0	--	--	--
Pedestrian Perceived Age (Base: Young adult or minor)						
Older (over 40 years)	--	--	--	0.750	5.35	84.0
Pedestrian Perceived Skin Tone (Base: White - MST 1-2)						
Brown - MST 3-8	-0.187	-1.85	-15.4	0.271	2.02	25.0
Black - MST 9-10	0.414	3.95	43.6	0.271	2.02	25.0
Pedestrian Exhibiting Visual Markers of Housing Insecurity (VHI) (Base: No VHI identified)						
VHI identified	1.092	6.88	149.1	1.065	3.60	127.3
Pedestrian Activity and Contextual Variables						
Social Context (Base: Solo crossing)						
Group crossing	-0.450	-4.17	-33.4	-0.273	-1.67	-20.5
Pedestrian Activity Type (Base: Walker)						
Runner	0.760	8.79	92.7	--	--	--
Vehicle Interaction (Base: No vehicle interaction)						
Vehicle interaction	-0.715	-7.13	-48.0	*n.a.*	*n.a.*	*n.a.*
Time-of-Day and its Interactions						
Time-of-day (Base: Noon or Afternoon; 12:00-17:59)						
Dawn (04:00-05:59)	1.246	4.90	178.1	--	--	--
Morning (06:00-11:59)	0.393	5.59	41.4	0.378	3.08	36.6
Dusk (18:00-19:59)	0.359	3.68	36.9	--	--	--
Evening (20:00-21:59)	0.932	5.69	119.2	--	--	--
Night (22:00-03:59)	1.722	10.55	299.6	--	--	--
Pedestrian Demographics and Time-of-Day Interactions						
Male × Nighttime (Dawn/Night)	-0.446	-2.69	-0.2[a]	0.622	2.84	64.2[a]
Brown × Nighttime (Dawn/Night)	-0.826	-2.01	-52.4[b]	--	--	--
Black × Nighttime (Dawn/Night)	-0.482	-1.80	-5.9[b]	--	--	--
Vehicle Characteristics						
Vehicle Type (*Base: SUV, sedan, pickup truck, other*)						
Commercial vehicle	--	--	--	-0.275	-1.00	-20.8
NCPC Behavior						
Pedestrian exhibiting NCPC *(Base: CPC)*	*n.a.*	*n.a.*	*n.a.*	1.270	5.97	164.1
Constants	-3.081	-38.90	*n.a.*	-2.334	-25.06	*n.a.*
Goodness-of-Fit						
Number of observations	17,374			2,890		
Number of parameters	16			9		
Log-likelihood at convergence $L(\beta)$	-4651.541			-1054.130		

Variable	Non-Compliant Pedestrian Crossing (NCPC) Model			Driver Unyielding (DUY) Model		
	Coef.	t-stat	%ATE	Coef.	t-stat	%ATE
Log-likelihood of constant-only model L(c)	-4889.200			-1107.688		
Log-likelihood of equal shares model L(0)[c]	-12042.739			-2003.195		
Nested likelihood ratio test[d]	LR = 475.32 > $\chi^2_{(15,0.00001)} = 37.70$			LR =107.12 > $\chi^2_{(8,0.00001)} = 37.33$		
Adjusted likelihood ratio index $\bar{\rho}_0^2$ [e]	0.614			0.474		
Adjusted likelihood ratio index $\bar{\rho}_c^2$ [e]	0.046			0.041		

[a] The base category for Male × Nighttime in the %ATE comparison is Female × Nighttime.

[b] The base category for Brown × Nighttime in the %ATE comparison is White × Nighttime. The base category for Black × Nighttime in the %ATE comparison is White × Nighttime.

[c] L(0) = [(-Number of observations)×ln(0.5)] for a binary model.

[d] The nested likelihood ratio test is computed with respect to the constants only model.

[e] The adjusted likelihood ratio indices are computed as follows:

$\bar{\rho}_c^2 = 1-\left(\mathrm{L}(\beta)-\mathrm{M}\right)/\mathrm{L}(\mathrm{c})$ and $\bar{\rho}_0^2 = 1-\left(\mathrm{L}(\beta)/\mathrm{L}(0)\right)$, where M is the number of parameters excluding constants.

Pedestrian Perceived Skin Tone

Skin tone plays a significant role in shaping pedestrian crossing behavior and driver responses, revealing layered inequities in public space. While the effects are smaller than those associated with gender or time-of-day, they remain statistically significant and socially meaningful. The results show that pedestrians with darker skin tones (MST 9-10-Black) are 43.6% more likely, and those with medium skin tones (MST 3-8-Brown) are 15.4% less likely, than lighter-skinned (MST 1-2-White) individuals to partake in NCPC behavior. Although naturalistic pedestrian studies have rarely examined the effect of skin tone, as already identified in Section 2, research in behavioral science and psychology suggests that experiences of discrimination, stereotype threat, and systemic inequities may heighten risk-taking behaviors among marginalized groups, particularly Black or African American individuals (Factor et al., 2013, Jamieson et al., 2013, Xie et al., 2020). Disparities are also observed in driver behavior. Pedestrians with darker skin tones (Brown/Black) are 25.0% more likely to experience DUY than their lighter-skinned (White) counterparts, echoing past research on racial bias in driver responses (Goddard et al., 2015, Coughenour et al., 2017, Schneider et al., 2018, Coughenour et al., 2020). These patterns persist on a per-crossing basis despite the inclusion of time-of-day effects, interaction terms and VHI status, suggesting that differences in yielding cannot be attributed to exposure or contextual effects. These disparities may stem from racial bias, a well-documented issue in the United States (Payne and Hannay, 2021, Skinner-Dorkenoo et al., 2023, The White House, 2024). Such bias can shape rapid, discretionary decisions, such as whether to yield or not to a pedestrian (Goddard et al., 2015). In these situations, drivers may consciously or subconsciously rely on racialized stereotypes, such as perceiving phenotypically Black individuals as more unpredictable or prone to risk-taking (Wages et al. 2022), resulting in elevated DUY.

Importantly, our results highlight that under the same infrastructure, environmental, and social contexts, behavioral differences, rooted in social and institutional inequities, place racial minorities at greater risk in public space, potentially explaining the persistent overrepresentation of Black and Brown individuals in pedestrian fatalities and injuries (e.g., Roll and McNeil, 2022, Haddad et al., 2023, Smart Growth America, 2024). For instance, institutional practices, particularly in law enforcement, provide important context for contextualizing these risks, not as isolated actions, but as risk-avoidance strategies shaped by experiences of over-policing. Black and Latino pedestrians are disproportionately targeted by jaywalking laws, comprising 92% of citations in New York City (Fitzsimmons, 2024), and 32% in Los Angeles, where Black residents make up just 9% of the population (Widera, 2024). Studies also show Black pedestrians are up to 4.5 times more likely than White pedestrians to be stopped for the same behavior (Widera, 2024). These patterns perpetuate community mistrust and may prompt further behavioral adaptations, such as avoiding marked crossings to minimize police encounters, ironically increasing exposure to unsafe conditions.

Addressing these disparities requires more than mere infrastructure improvements or generalized awareness campaigns. Meaningful progress must begin with the communities most affected. Community-led safety initiatives, such as participatory design, storytelling, and localized education, can help uncover context-specific challenges, foster trust, and empower residents to co-develop interventions that genuinely support their safety and dignity. Simultaneously, the persistence of lower yielding rates toward Black and Brown pedestrians highlights the need to address racial bias not just among drivers but also among police officers, first responders, and transportation professionals. Increasing evidence shows that educating both children and adults about historical and systemic racial injustices, including those embedded in transportation and law enforcement, can foster empathy, reduce prejudice, and enhance awareness of structural racism (Skinner-Dorkenoo et al., 2023). Integrating such content into public awareness campaigns, school curricula, and driver education programs can gradually shift societal attitudes and help dismantle racialized narratives that continue to shape pedestrian safety disparities. However, the effectiveness of these efforts not only depends on the message itself, but also on how it is framed and communicated. Research on social norm messaging shows that positively framed dynamic norms (e.g., “More and more drivers are giving way to pedestrians”) and injunctive norms (e.g., “Please give way to pedestrians”) are especially effective in promoting yielding behavior, outperforming static or negatively worded messages (Liu et al., 2022). Thus, integrating structural education with behaviorally-informed messaging strategies can offer a robust and effective approach to reshaping attitudes and addressing racialized disparities in pedestrian safety.

Pedestrian Exhibiting Visual Markers of Housing Insecurity (VHI)

The results reveal a critical and underexamined relationship between visual indicators of housing instability (VHI), used as a proxy for homelessness, and pedestrian safety outcomes. Pedestrians identified as exhibiting VHI have a 149.1% higher probability of engaging in NCPC behavior and are 127.3% more likely to experience driver failure to yield. These represent the third strongest overall effect on NCPC behavior and the most influential variable among all pedestrian characteristics. VHI is also the second strongest predictor of DUY behavior in the model. These disparities likely stem from both behavioral factors and systemic neglect. Individuals experiencing homelessness often face health-related vulnerabilities, including physical disabilities, mental health conditions, and substance use disorders, that may influence crossing decisions (Richards and Kuhn, 2022, USDOT, 2024). At the same time, reduced driver yielding toward this group echoes findings from Domine et al. (2022), who documented complete driver non-compliance in crashes occurring along corridors with overlapping encampments and high crash rates. The compounding effects of increased exposure, elevated behavioral risk, and diminished driver responsiveness position individuals with VHI among the most vulnerable pedestrians and likely contribute to the disproportionately high rates of pedestrian crashes observed in areas with larger proportions of unhoused populations (Bernhardt and Kockelman, 2021).

Despite growing awareness of traffic violence disproportionately affecting unhoused communities, housing status remains largely excluded from Vision Zero crash reporting systems, hindering efforts to track and address these disparities in a data-driven manner (Zimmerman, 2023, USDOT, 2024). Municipal responses have been inconsistent, oscillating between supportive and punitive measures. For example, the city of Austin's 2016 Vision Zero strategy acknowledged housing as a safety determinant through a "Housing First" approach, but recent revisions have retreated toward infrastructure- and behavior-based solutions. Meanwhile, the city of Portland has enacted encampment bans along high-crash corridors, and other cities, including San Jose and Colorado Springs, have implemented visibility-based interventions such as neon beanies and flashing headbands for unhoused individuals (Zimmerman, 2023). Though often well-intentioned, such efforts often reduce a structural safety crisis to a question of visibility or individual behavior, ignoring the broader environmental and systemic conditions that place unhoused pedestrians at elevated risk.

Collectively, our findings call for equity-centered, cross-sector strategies that recognize housing insecurity as both a transportation and public health concern. Transportation agencies must begin by integrating housing status into crash data systems and establishing it as a standard variable within Vision Zero and other safety frameworks. Infrastructure near encampments and areas of unsheltered habitation should also be redesigned to include pedestrian-activated signals, refuge islands, enhanced lighting, and traffic calming. Collaboration with housing and public health agencies can also be beneficial for delivering mobile outreach, navigation centers, and harm-reduction services at high-risk locations.

5.1.2. Pedestrian Activity and Contextual Variable Effects

Social Context

The presence of additional individuals at the crossing consistently exerted a protective influence across both models. Specifically, pedestrians crossing in groups are 33.4% less likely to commit violations and experience a 20.5% increase in the likelihood of drivers yielding compared to those crossing alone. These results align with prior findings (e.g., Dileep et al., 2016, Zhu et al., 2021, Fu et al., 2022, Zhang et al., 2023, Zafri et al., 2022, Miladi et al., 2025, Rafe et al., 2025), and may be attributed to social conformity pressures, where individuals are less inclined to violate traffic rules when others are present who may implicitly discourage non-compliant behavior (Zhu et al., 2021). In the context of yielding, drivers may be more likely to notice groups or feel a stronger obligation, whether social or practical, to yield when multiple people are present.

These observed patterns lend further empirical support to the "safety in numbers" phenomenon, whereby group crossings are associated with improved compliance among both pedestrians and drivers, suggesting that interventions that encourage group travel or simulate its social cues can significantly enhance safety. For instance, behaviorally informed public messaging campaigns can promote compliant crossing by framing safe behaviors, such as using marked crosswalks, waiting for signals, or utilizing grade-separated infrastructure, as social responsibilities. Emphasizing that such actions set a positive example for peers, family, and the broader community may motivate individuals to adopt safer behaviors not only for self-preservation but also out of concern for others. School-based educational programs and parent-focused campaigns can further reinforce these social norms, fostering intergenerational learning and long-term behavior change.

Pedestrian Activity Type

Pedestrian activity type significantly influences crossing behavior, with runners being 92.7% more likely to engage in NCPC compared to walkers. This elevated risk likely reflects a desire to maintain workout momentum by avoiding signal delays, as well as a lower perceived risk due to higher speeds, which can make smaller traffic gaps seem acceptable during unprotected crossings. Addressing this issue requires infrastructure and safety strategies that respond to the distinct behavioral dynamics of different pedestrian activities. For example, mid-block crossings, pedestrian overpasses, or dedicated running paths positioned away from vehicle conflict zones could help reduce unsafe behaviors among runners, particularly in parks, along trails, and in areas popular with exercise groups and dog walkers, where conventional crossing infrastructure may fall short.

Vehicle Interaction

Direct interaction with a vehicle is associated with a 48.0% reduction in pedestrian violations, suggesting that pedestrians exercise greater caution when vehicles are in the immediate vicinity in conflict with the crossing. This may indicate that NCPC behaviors are calculated decisions based on perceived traffic gaps rather than purely impulsive actions. The finding also highlights the role of real-time feedback mechanisms, such as visual contact, vehicle speed, and deceleration cues, in influencing pedestrian decision-making.

5.1.3. Time-of-Day and its Interaction Effects

Time-of-Day

Compared to the noon-to-afternoon window (12:00-17:59), pedestrians are more likely to engage in NCPC at all other times. The morning (06:00-11:59) and dusk (18:00-19:59) periods, both overlapping with typical commute hours, are associated with moderate increases in NCPC probability (41.4% and 36.9%, respectively). DUY also increases by 36.6% during the morning period. These trends likely reflect time pressures related to work, school, and caregiving, which reduce both pedestrian patience and driver caution. Empirical studies support these findings, documenting higher rates of crossing and yielding violations during peak travel times and under time-urgent conditions (e.g., Guo et al., 2011, Zhang et al., 2016, Zhou et al., 2016, Xiong et al., 2019, Ma et al., 2020, Dhoke and Choudhary, 2023). This behavior is attributed to what is known in behavioral theory as instrumental attitudes, wherein individuals justify unsafe actions based on perceived efficiency gains. However, as noted by Zhou et al. (2016), such savings are often negligible and come at a disproportionately high safety risk. Educational campaigns should therefore target this misperception, emphasizing that the short-term convenience of jaywalking rarely justifies the heightened crash risk. Public messaging, particularly during peak commute hours, should communicate the dangers of NCPC, especially in contexts with fast-moving or unpredictable traffic. These efforts may be most effective when integrated into school curricula, commuter outreach, and workplace safety training.

Even greater increases in the likelihood of engaging in NCPC are observed during dawn (178.1%), evening (119.2%), and especially night hours (299.6%), covering the period between 20:00 and 05:59. These findings align with Rafe et al. (2025), who attribute elevated nighttime violations to lower traffic volumes and increased anxiety in poorly lit environments. Psychological research suggests that darkness and isolation can heighten discomfort, prompting impulsive behavior (Steimer, 2002, Hengen and Alpers, 2021). In such contexts, lighting plays a dual role, enhancing visibility and improving perceived safety. Infrastructure upgrades such as pedestrian-scale lamps, motion-activated lighting, and illuminated signage can reduce violations by fostering a sense of security and increasing driver alertness.

Pedestrian Demographics and Time-of-Day Interaction Variables

Besides main effects, time-of-day also moderates the influence of gender and race on both pedestrian and driver behaviors. A significant interaction between male gender and nighttime conditions indicates that the risk-taking behavior typically associated with male pedestrians diminishes under low-light environments, likely due to heightened risk perception, reduced visibility, or situational caution. Specifically, while across all time periods, men are 49.0% more likely than women to engage in NCPC, at nighttime (22:00-05:59), the difference disappears, with men being 0.2% *less* likely to engage in NCPC than women crossing at the same time. In contrast, DUY to male pedestrians increases substantially at night (22:00-05:59), to 62.4%, suggesting that gender-related disparities in driver behavior become more pronounced when visibility is limited. This pattern may reflect the amplification of implicit gender biases under low-light conditions, where drivers may underestimate male pedestrians' vulnerability or misinterpret their crossing intentions.

The interaction between pedestrian skin tone and time-of-day also reveals important behavioral differences. Although NCPC behavior generally increases at nighttime, the effect is less pronounced among Brown (MST 3-8) and Black (MST 9-10) pedestrians. For instance, across all time periods, Brown pedestrians are 15.4% less likely than White pedestrians to engage in NCPC, and this reduced tendency for NCPC among Brown pedestrians becomes even more substantial at nighttime, reaching 52.4%. Similarly, Black pedestrians are 43.6% more likely than White pedestrians to engage in NCPC across all time periods, but 5.9% less likely at nighttime. These patterns suggest that pedestrians of color exhibit greater caution under low-light conditions, possibly due to increased risk awareness, fear of enforcement, or heightened safety concerns. However, despite more cautious behavior among racial minorities and no observed increase in racial disparity in driver yielding at night, Black pedestrians remain disproportionately involved in nighttime crashes (Sanders et al., 2022). This suggests that other factors, such as visibility differences, driver detection biases, infrastructure quality, or unequal exposure, may contribute to the persistent safety gap and warrant further investigation.

5.1.4. Vehicle Characteristic Variable Effects

<u>*Vehicle Type*</u>

The commercial vehicle variable exhibits a negative effect on non-yielding behavior, with a %ATE of -20.8%. This indicates that drivers of commercial vehicles, such as box-trucks, company cars or vans, garbage trucks, and city buses, are, on average, 20.8% less likely to fail to yield to pedestrians compared to drivers of passenger vehicles (e.g., sedans, SUVs, pickup trucks, and other motorized vehicles). To assess whether this effect varies across the driver population due to unobserved heterogeneity, such as individual differences in training quality, risk tolerance, or enforcement exposure, we estimated mixed logit specifications with random parameters for vehicle type. None of the random-parameter standard deviations were statistically significant, indicating that the vehicle-type effect is relatively homogeneous across our sample and that the fixed-effects specification adequately captures the average relationship. While this finding contrasts with earlier studies suggesting lower compliance among commercial drivers (e.g., Dileep et al., 2016, Figliozzi and Tipagornwong, 2016), several plausible explanations support the current result. Commercial drivers typically undergo more rigorous training and certification, including specific instruction on pedestrian safety and right-of-way laws (Gillham et al., 2023). They are also subject to greater regulatory oversight, including the use of electronic logging devices (ELDs), GPS tracking, and frequent safety inspections, which may encourage more compliant driving behavior. Moreover, the professional and financial consequences of traffic violations, including potential job loss, insurance penalties, and company-imposed sanctions, may further incentivize adherence to the rules. Nonetheless, while commercial drivers are held to higher standards and face stricter oversight, the broader literature on driver compliance yields mixed findings. There is no definitive evidence that commercial drivers are universally more compliant with traffic laws. Compliance appears to vary by context, enforcement, and operational pressures, with some studies linking tight schedules and fatigue to reduced compliance (Chen et al., 2021). Although the current study provides empirical evidence suggesting higher yielding compliance among commercial vehicle drivers, further research is needed to explore the consistency of this pattern across different traffic contexts, geographic settings, and vehicle subtypes.

5.1.5. NCPC Effect on DUY Behavior

The results in Table 4 align with prior findings by Bella and Nobili (2020), and indicate that drivers are 164.1% more likely to fail to yield when pedestrians engage in NCPC behavior. This increased DUY behavior likely stems from drivers' diminished expectation of encountering pedestrians outside designated crossing areas or signal phases. When pedestrians cross outside marked crosswalks or during a prohibited signal phase, drivers have less time to recognize crossing intent and react appropriately, as also suggested by Bella and Nobili (2020). This breakdown in typical visual and behavioral cues appears to drive the higher DUY rates observed with NCPC behavior. Of course, it is also possible that drivers view non-compliant crossings as illegal and therefore deliberately and willfully choose not to yield to NCPC pedestrians, even though Texas traffic law still obligates motorists to exercise due care to avoid colliding with pedestrians, notwithstanding any other provision (Section 552.008).

Notably, the relationship between NCPC and DUY is the strongest among all variables in the model, underscoring the seriousness of this safety issue. Transportation planners and relevant agencies must prioritize systematic data collection on NCPC behavior to identify locations where additional infrastructure may be needed. In areas where NCPC behavior is frequent, formalizing common desired crossing paths or installing new crossings, even in unconventional locations, should be considered. Emerging tools such as video analytics, trajectory tracking, and pedestrian heat maps can help pinpoint areas where NCPC is prevalent and assess whether current infrastructure and/or signal timing unintentionally contribute to non-compliance. In parallel, both drivers and pedestrians should be made more aware of the safety implications of NCPC behavior. Driver education campaigns should emphasize the unpredictability of pedestrian actions and the importance of yielding regardless of pedestrian compliance. At the same time, pedestrian outreach efforts should highlight the increased risks associated with NCPC behavior and promote safer, designated crossing practices.

5.1.6. Constants and Goodness-of-Fit

The constant terms in both models represent the overall effects for the base demographic group as determined by the combination of the base categories across all exogenous variables (this is so because all exogenous variables in both the binary models are in discrete categories). In the NCPC model, the reference scenario corresponds to a female pedestrian who is a minor or young adult, has a skin tone classified as MST 1-2, shows no VHI, is walking alone during noon or afternoon hours, and is not in proximity to any vehicles. In the non-yielding model, the reference pedestrian has the same characteristics but is walking/running during any time period other than morning hours (06:00-11:59), and is interacting with a non-commercial vehicle. The constant coefficients correspond to predicted probabilities of 4.4% NCPC and 8.8% DUY for these base demographic groups.

The goodness-of-fit statistics, presented at the bottom of Table 4, indicate that both the NCPC and non-yielding models provide a substantial and statistically significant improvement in explanatory power compared to their respective null models (i.e., models corresponding to predictions of equal shares and sample shares for the dependent outcomes). The nested likelihood ratio (LR) test with respect to the constants-only (that is, sample shares) model confirms that the inclusion of explanatory variables significantly enhances model fit, with test statistics exceeding the critical Chi-squared values at the 99.99999% confidence level. This provides strong statistical evidence that the included predictors are meaningfully associated with the outcome variables, and that the observed improvement in fit is unlikely due to random chance.

Although the sample exhibits class imbalance, this does not bias the estimated parameters in a binary logit/probit maximum likelihood framework. The imbalance is absorbed in the alternative-specific constant, and the effects of covariates remain consistently estimated (Train, 2009, Greene, 2018a, Greene, 2018b). Class imbalance is problematic mainly for machine-learning classifiers that optimize deterministic classifications rather than the likelihood function. Since our goal is behavioral inference, class imbalance does not threaten the validity of our results.

6. CONCLUSIONS AND FUTURE DIRECTIONS

This study advances our understanding of pedestrian safety disparities by examining how sociodemographic, activity, contextual, and time-of-day/weather factors influence pedestrian crossing compliance and driver yielding behavior in real-world settings. Through systematic analysis of over 17,374 pedestrian crossings and 2,890 pedestrian-vehicle interactions, our findings reveal complex patterns of risk that extend beyond individual behavioral choices to reflect broader structural and social dynamics. Non-compliant pedestrian crossings (NCPC) emerged as most prevalent during low-visibility periods (night, dawn, and evening), among individuals with visible housing insecurity indicators (VHI), male pedestrians, runners, and Black pedestrians. Conversely, group crossings and active vehicle interactions demonstrated protective effects. Driver yielding patterns revealed concerning disparities, with the lowest yielding compliance observed toward those engaging in NCPC, individuals exhibiting VHI, older adults, males during nighttime hours, morning crossings, pedestrians encountered by drivers of non-commercial vehicles, Black or Brown pedestrians, and individuals crossing alone. The observed trends, specifically the convergence of reduced driver yielding and increased NCPC among certain demographic groups, suggest that disparities in pedestrian safety are shaped by a combination of individual behaviors and broader systemic forces, including structural disadvantage, unequal exposure, and context-specific risks. As such, there is a need to shift away from framing pedestrian risk in terms of individual blame and toward developing comprehensive solutions that recognize and respond to the full context of pedestrian vulnerability. Effective pedestrian safety strategies must reflect the lived realities of marginalized groups, adapt to shifting risk profiles over time and across different contexts, and actively confront the systemic biases embedded in both infrastructure and enforcement practices.

In this context, it is also widely recognized that behaviors such as jaywalking are persistent and unlikely to be entirely eliminated (Fitzsimmons, 2024, Raoniar et al., 2024). Expecting full compliance with traffic rules is unrealistic, especially for those navigating exclusionary or unsafe environments. This reinforces the need for infrastructure that is inherently forgiving of human error. Rather than relying solely on enforcement, the focus should be on minimizing the consequences of unsafe actions through safer street design, such as reduced vehicle speeds, clearer crossings, and protected pedestrian spaces, while also ensuring that both drivers and pedestrians are educated about the risks and consequences of their behaviors.

This study combines long-duration, naturalistic data collection with rigorous variable coding to assemble a large database of over 17,000 pedestrian crossings and over 2,800 pedestrian-vehicle interactions. This scale enabled us to observe a reasonable number of NCPC and DUY events to quantify the effects of individual characteristics (including pedestrian race, perceived housing insecurity, and walking versus running), social context (solo versus group crossings), environmental conditions (darkness and weather), as well as interactions among these factors. A further contribution is the examination of NCPC and DUY behaviors at intersections with channelized slip lanes, providing new insight into decision-making under ambiguous right-of-way conditions. Overall, by identifying and quantifying factors shaping pedestrian–vehicle interactions across diverse contexts, the study offers evidence-based policy and infrastructure strategies to improve pedestrian safety.

While this research presented new insights into pedestrian and driver behaviors, several opportunities remain for future investigation. Some aspects of data collection, such as the subjectivity in coding pedestrian age and skin tone, could benefit from refinement to reduce observational bias. The study's geographic scope was also limited to two intersections with similar designs, including one on a university campus where drivers may be more attuned to pedestrian presence. This context likely influenced both yielding behavior and the demographic composition of observed pedestrians, particularly skewing the sample toward young adults. Additionally, data were collected exclusively in late spring and early summer, when high temperatures in Texas may influence pedestrian volumes and behavior, introducing potential seasonal effects. Rainfall events are rare during this period, and the small number of observations prevented the exploration of the impact of rainy conditions on NCPC and DUY behaviors. Future research should expand the geographic and temporal scope to capture a wider range of environmental, demographic, and behavioral conditions. Incorporating additional variables, such as driver demographics, pedestrian gestures, walking speed, and vehicle acceleration or deceleration, would also provide a more comprehensive understanding of vehicle-pedestrian interactions. Lastly, complementary qualitative methods, including surveys, focus groups, or virtual reality simulations, could further illuminate the decision-making processes of both pedestrians and drivers.

ACKNOWLEDGMENTS

This research was partially supported by the UT Good Systems Grand Challenge through the Good Systems Core Research Project entitled "Designing AI for Social Fairness," and by the U.S. Department of Transportation (USDOT) under Grant 69A3552348327 for the Center for Automated Vehicle Research with Multimodal AssurEd Navigation (CARMEN+) University Transportation Center. The authors thank Kenneth A. Perrine and Jan C. Rubio for their assistance in designing and implementing the video data collection system, and Dr. Katherine Asmussen and Carol Anne Starks for their contributions to data labeling. The authors are also grateful to Lisa Macias for her help in formatting this document, and to two anonymous reviewers who provided useful comments and suggestions on an earlier version of this paper.

REFERENCES

Aghabayk, K., Esmailpour, J., Jafari, A., and Shiwakoti, N. (2021). Observational-based study to explore pedestrian crossing behaviors at signalized and unsignalized crosswalks. Accident Analysis and Prevention, 151, 105990. https://doi.org/10.1016/j.aap.2021.105990

Anciaes, P., Di Guardo, G., and Jones, P. (2020). Factors explaining driver yielding behaviour towards pedestrians at courtesy crossings. Transportation Research Part F: Traffic Psychology and Behaviour, 73, 453–469. https://doi.org/10.1016/j.trf.2020.07.006

Anik, M. A. H., Hossain, M., and Habib, M. A. (2021). Investigation of pedestrian jaywalking behaviour at mid-block locations using artificial neural networks. Safety Science, 144, 105448. https://doi.org/10.1016/j.ssci.2021.105448

Avineri, E., Shinar, D., and Susilo, Y. O. (2012). Pedestrians' behaviour in cross walks: The effects of fear of falling and age. Accident Analysis and Prevention, 44(1), 30–34. https://doi.org/10.1016/j.aap.2010.11.028

Baker, M. D., Owens, B. T., Utter, R. L., Boachie, R. J., and Keith, D. R. (2022). Boundary conditions for gender differences in risk taking. Journal of Risk Research, 25(4), 536–545. https://doi.org/10.1080/13669877.2021.1962953

Bella, F., and Nobili, F. (2020). Driver-pedestrian interaction under legal and illegal pedestrian crossings. Transportation Research Procedia, 45, 451–458. https://doi.org/10.1016/j.trpro.2020.03.038

Bendak, S., Alnaqbi, A. M., Alzarooni, M. Y., Aljanaahi, S. M., and Alsuwaidi, S. J. (2021). Factors affecting pedestrian behaviors at signalized crosswalks: An empirical study. Journal of Safety Research, 76, 269–275. https://doi.org/10.1016/j.jsr.2020.12.019

Bernhardt, M., and Kockelman, K. (2021). An analysis of pedestrian crash trends and contributing factors in Texas. Journal of Transport and Health, 22, 101090. https://doi.org/10.1016/j.jth.2021.101090

Bharate, G. R., Pinjari, A. P., and Bhat, C. R. (2025). Identification of random parameters on categorical attributes in mixed logit models of crash outcomes and travel choices [Conference presentation]. 8th Conference of the Transportation Research Group of India, Guwahati, India.

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transportation Research Part B: Methodological, 35(7), 677–693. https://doi.org/10.1016/S0191-2615(00)00014-X

Bhat, C. R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B: Methodological, 37(9), 837–855. https://doi.org/10.1016/S0191-2615(02)00090-5

Bhat, C. R., and Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transportation Research Part B: Methodological, 43(7), 749–765. https://doi.org/10.1016/j.trb.2009.02.001

Bhat, C. R. (2024). Transformation-based flexible error structures for choice modeling. Journal of Choice Modelling, 53, 100522.

Blanch, A., and Martínez, A. (2025). Risk Taking and Accidents in Adventure Sports Across Males and Females. International Journal of Psychology, 60(2), e70023. https://doi.org/10.1002/ijop.70023

Bowen, E. A., and Capozziello, N. (2024). Faceless, nameless, invisible: A visual content analysis of photographs in U.S. media coverage about homelessness. Housing Studies, 39(3), 746–765. https://doi.org/10.1080/02673037.2022.2084044

Brosseau, M., Zangenehpour, S., Saunier, N., and Miranda-Moreno, L. (2013). The impact of waiting time and other factors on dangerous pedestrian crossings and violations at signalized intersections: A case study in Montreal. Transportation Research Part F: Traffic Psychology and Behaviour, 21, 159–172. https://doi.org/10.1016/j.trf.2013.09.010

Chavis, C., Nyarko, K., and Cirillo, C. (2023). A Comparative Study of Pedestrian Crossing Behavior and Saftey in Baltimore, MD and Washington, DU Using Video Surveillance. Urban Mobility and Equity Center.

Chen, G. X., Karl Sieber, W., Collins, J. W., Hitchcock, E. M., Lincoln, J. E., Pratt, S. G., and Sweeney, M. H. (2021). Truck driver reported unrealistically tight delivery schedules linked to their opinions of maximum speed limits and hours-of-service rules and their compliance with these safety laws and regulations. Safety Science, 133, 105003. https://doi.org/10.1016/j.ssci.2020.105003

Coughenour, C., Clark, S., Singh, A., Claw, E., Abelar, J., and Huebner, J. (2017). Examining racial bias as a potential factor in pedestrian crashes. Accident Analysis and Prevention, 98, 96–100.

Coughenour, C., Abelar, J., Pharr, J., Chien, L.-C., and Singh, A. (2020). Estimated car cost as a predictor of driver yielding behaviors for pedestrians. Journal of Transport and Health, 16, 100831. https://doi.org/10.1016/j.jth.2020.100831

Dadashova, B., Boutros, A., Reyes, D., Schoner, J., Sanders, R., Chiovenda, M., Lee, C., Zhu, C., Wang, O., Elgart, Z., Eisert, J., and Thompson Panik, R. (2024). Exploring Risk Factors to Disparities in Pedestrian and Bicyclist Fatalities and Serious Injuries. Federal Highway Administration. https://doi.org/10.21949/1521546

Das, T., Samandar, M. S., Autry, M. K., and Rouphail, N. M. (2023). Surrogate Safety Measures: Review and Assessment in Real-World Mixed Traditional and Autonomous Vehicle Platoons. IEEE Access, 11, 32682–32696. https://doi.org/10.1109/ACCESS.2023.3248628

Deb, S., Strawderman, L., DuBien, J., Smith, B., Carruth, D. W., and Garrison, T. M. (2017). Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the U.S. population. Accident Analysis and Prevention, 106, 191–201. https://doi.org/10.1016/j.aap.2017.05.020

Dellosa, G., and Browne, M. (2024). The influence of age on gambling problems worldwide: A systematic review and meta-analysis of risk among younger, middle-aged, and older adults. Journal of Behavioral Addictions, 13(3), 702-715. https://doi.org/10.1556/2006.2024.00051

Demir, E. N., Dönmez, H. B., Özden, M., Ünver, N., Fındık, G., and Öz, B. (2020). Yielding the Right of The Way: A Mixed Design Study for Understanding Drivers' Yielding Behavior. Trafik ve Ulaşım Araştırmaları Dergisi, 3(1), Article 1. https://doi.org/10.38002/tuad.690423

Dhoke, A., and Choudhary, P. (2023). Is there a relationship between time pressure and pedestrian non-compliance? A systematic review. Transportation Research Part F: Traffic Psychology and Behaviour, 93, 68–89. https://doi.org/10.1016/j.trf.2023.01.001

Dileep, R., Koshy, D. B. I., and Sam, E. (2016). Study on Driver Yielding to Pedestrians at Unsignalized Crosswalks. International Journal of Scientific and Engineering Research, 7(9), 306–311.

Domine, P., Doyle, S., Haque, A., Sulvarán, A. M., Meusch, N., and Whyte, M. (2022). Safety Interventions For Houseless Pedestrians. Master of Urban and Regional Planning Workshop Projects, 184. https://archives.pdx.edu/ds/psu/38439.

Dommes, A., Granié, M.-A., Cloutier, M.-S., Coquelet, C., and Huguenin-Richard, F. (2015). Red light violations by adult pedestrians and other safety-related behaviors at signalized crosswalks. Accident Analysis and Prevention, 80, 67–75. https://doi.org/10.1016/j.aap.2015.04.002

Dumbaugh, E., Stiles, J., Mitsova, D., and Saha, D. (2023). The Most Vulnerable User: Examining the Role of Income, Race, and the Built Environment on Pedestrian Injuries and Deaths. Transportation Research Record, 2678(2), 743–752. https://doi.org/10.1177/03611981231175888

Factor, R., Williams, D. R., and Kawachi, I. (2013). Social resistance framework for understanding high-risk behavior among nondominant minorities: Preliminary evidence. American Journal of Public Health, 103(12), 2245–2251.

Ferenchak, N. N. (2016). Pedestrian age and gender in relation to crossing behavior at midblock crossings in India. Journal of Traffic and Transportation Engineering (English Edition), 3(4), 345–351. https://doi.org/10.1016/j.jtte.2015.12.001

Fernández Llorca, D., Frau, P., Parra, I., Izquierdo, R., and Gomez, E. (2024). Attribute annotation and bias evaluation in visual datasets for autonomous driving | Journal of Big Data. https://link.springer.com/article/10.1186/s40537-024-00976-9

Figliozzi, M. A., and Tipagornwong, C. (2016). Pedestrian Crosswalk Law: A study of traffic and trajectory factors that affect non-compliance and stopping distance. Accident Analysis and Prevention, 96, 169–179. https://doi.org/10.1016/j.aap.2016.08.011

Fitzsimmons, E. G. (2024, October 29). Jaywalking Is a New York Tradition. Now It's Legal, Too. The New York Times. https://www.nytimes.com/2024/10/29/nyregion/jaywalking-legal-law-nyc.html

Fu, T., Yu, X., Xiong, B., Jiang, C., Wang, J., Shangguan, Q., and Xu, W. (2022). A method in modeling interactive pedestrian crossing and driver yielding decisions during their interactions at intersections. Transportation Research Part F: Traffic Psychology and Behaviour, 88, 37–53. https://doi.org/10.1016/j.trf.2022.05.005

Gerogiannis, A., and Bode, N. W. F. (2024). Analysis of long-term observational data on pedestrian road crossings at unmarked locations. Safety Science, 172, 106420. https://doi.org/10.1016/j.ssci.2024.106420

Ghomi, H., and Hussein, M. (2022). An integrated text mining, literature review, and meta-analysis approach to investigate pedestrian violation behaviours. Accident Analysis and Prevention, 173, 106712. https://doi.org/10.1016/j.aap.2022.106712

Gillham, O., Madsen, D., Nguyen, K., and Jingsi, S. (2023). Assessment of Commercial Driver's License (CDL) Holders' Traffic Violations, Convictions, and Disqualifications. U.S. Department of Transportation. https://rosap.ntl.bts.gov/view/dot/72942

Goddard, T., Kahn, K. B., and Adkins, A. (2015). Racial bias in driver yielding behavior at crosswalks. Transportation Research Part F: Traffic Psychology and Behaviour, 33, 1–6. https://doi.org/10.1016/j.trf.2015.06.002

Goldfischer, E. (2018). "Peek-A-Boo, We See You Too": Homelessness and visuality in New York City. Environment and Planning D, 36(5), 831–848. https://doi.org/10.1177/0263775818768546

Greene, W. H. (2018a). Maximum Likelihood Estimation. In Econometric analysis (8th ed., pp. 537-640). Pearson Education.

Greene, W. H. (2018b). Binary Outcomes and Discrete Choices. In Econometric analysis (8th ed., pp. 728-756). Pearson Education.

Greitemeyer, T. (2023). The relationship between social class and unethical and prosocial (traffic) behavior: Two naturalistic replication studies. Social Influence, 18(1), 2203947. https://doi.org/10.1080/15534510.2023.2203947

Guo, H., Gao, Z., and Yang, X. (2011). Modeling Pedestrian Violation Behavior at Signalized Crosswalks in China: A Hazards-Based Duration Approach. Traffic Injury Prevention, 12(1), 96–103. https://doi.org/10.1080/15389588.2010.518652

Haddad, A. J., Mondal, A., Bhat, C. R., Zhang, A., Liao, M. C., Macias, L. J., Lee, M. K., and Watkins, S. C. (2023). Pedestrian crash frequency: Unpacking the effects of contributing factors and racial disparities. Accident Analysis and Prevention, 182, 106954. https://doi.org/10.1016/j.aap.2023.106954

Hall, D., and Perona, P. (2015). Fine-Grained Classification of Pedestrians in Video: Benchmark and State of the Art. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5482–5491. https://openaccess.thecvf.com/content_cvpr_2015/html/Hall_Fine-Grained_Classification_of_2015_CVPR_paper.html

Hamann, C., Peek-Asa, C., and Butcher, B. (2020). Racial disparities in pedestrian-related injury hospitalizations in the United States. BMC Public Health, 20(1), 1459. https://doi.org/10.1186/s12889-020-09513-8

Han, S., and Lee, S. (2019). Estimation in a generalization of bivariate probit models with dummy endogenous regressors. Journal of Applied Econometrics, 34(6), 994-1015.

Hengen, K. M., and Alpers, G. W. (2021). Stress Makes the Difference: Social Stress and Social Anxiety in Decision-Making Under Uncertainty. Frontiers in Psychology, 12, 578293. https://doi.org/10.3389/fpsyg.2021.578293

Hirun, W. (2016). Factors affecting driver yielding behavior at a mid-block zebra crossing. International Journal of Engineering and Technology, 8(2), 906–912.

Hübner, M., Stockmann, J.-N. B., and Bengler, K. (2025). Crossing the line: Impact of pedestrian group behavior on individual crossing decisions in AV interactions. Transportation Research Part F: Traffic Psychology and Behaviour, 109, 921–937. https://doi.org/10.1016/j.trf.2025.01.020

Ivan, J. N., McKernan, K., Zhang, Y., Ravishanker, N., and Mamun, S. A. (2017). A study of pedestrian compliance with traffic signals for exclusive and concurrent phasing. Accident Analysis and Prevention, 98, 157–166. https://doi.org/10.1016/j.aap.2016.10.003

Jamieson, J. P., Koslov, K., Nock, M. K., and Mendes, W. B. (2013). Experiencing Discrimination Increases Risk Taking. Psychological Science, 24(2), 131–139. https://doi.org/10.1177/0956797612448194

Junejo, I. N., and Ahmed, N. (2020). A multi-branch separable convolution neural network for pedestrian attribute recognition. Heliyon, 6(3). https://doi.org/10.1016/j.heliyon.2020.e03563

Kumar, P., Tiwari, G., and Paul, S. B. (2025). Road safety studies at micro, meso, and macroscopic levels: A systematic review. IATSS Research, 49(1), 10–26. https://doi.org/10.1016/j.iatssr.2024.12.001

Lee, J., Abdel-Aty, M., Huang, H., and Cai, Q. (2019). Transportation Safety Planning Approach for Pedestrians: An Integrated Framework of Modeling Walking Duration and Pedestrian Fatalities. Transportation Research Record, 2673(4), 898–906. https://doi.org/10.1177/0361198119837962

Liu, X., Qu, W., and Ge, Y. (2022). The nudging effect of social norms on drivers' yielding behaviour when turning corners. Transportation Research Part F: Traffic Psychology and Behaviour, 89, 53–63. https://doi.org/10.1016/j.trf.2022.06.011

Liu, Y.-C., and Tung, Y.-C. (2014). Risk analysis of pedestrians' road-crossing decisions: Effects of age, time gap, time of day, and vehicle speed. Safety Science, 63, 77–82. https://doi.org/10.1016/j.ssci.2013.11.002

Long, F. C. (2024). Drawing out perceptions: Using drawing as a method to understand public perceptions of homelessness and crime. International Review of Victimology, 30(2), 336–357. https://doi.org/10.1177/02697580231215439

Ma, Y., Lu, S., and Zhang, Y. (2020). Analysis on Illegal Crossing Behavior of Pedestrians at Signalized Intersections Based on Bayesian Network. Journal of Advanced Transportation, 2020(1), 2675197. https://doi.org/10.1155/2020/2675197

McGuckin, N., Fucci, A., and Westat, Inc. (2018). Summary of Travel Trends: 2017 National Household Travel Survey (FHWA-PL-18-019). https://rosap.ntl.bts.gov/view/dot/68751

Miladi, M., Waygood, E. O. D., Cloutier, M.-S., Bobin, W., and Ali Yas, Z. (2025). Distractions or long waits? Impacts on risky crossing behaviour. IATSS Research, 49(2), 220–230. https://doi.org/10.1016/j.iatssr.2025.05.001

Mirhashemi, A., Amirifar, S., Kashani, A. T., and Zou, X. (2022). Macro-level literature analysis on pedestrian safety: Bibliometric overview, conceptual frames, and trends. Accident Analysis and Prevention, 174, 106720.

Monk, E. (2022). Monk skin tone scale. Skin Tone Research @ Google AI. https://skintone.google/

Morency, P., Gauvin, L., Plante, C., Fournier, M., and Morency, C. (2012). Neighborhood Social Inequalities in Road Traffic Injuries: The Influence of Traffic Volume and Road Design. American Journal of Public Health, 102(6), 1112–1119. https://doi.org/10.2105/AJPH.2011.300528

Morgenroth, T., Fine, C., Ryan, M. K., and Genat, A. E. (2018). Sex, Drugs, and Reckless Driving: Are Measures Biased Toward Identifying Risk-Taking in Men? Social Psychological and Personality Science, 9(6), 744–753. https://doi.org/10.1177/1948550617722833

Mukherjee, D., and Mitra, S. (2020). A comprehensive study on factors influencing pedestrian signal violation behaviour: Experience from Kolkata City, India. Safety Science, 124, 104610. https://doi.org/10.1016/j.ssci.2020.104610

Naumann, R., West, B., Barry, V., Matthews, S., and Lee, R. (2025). Pedestrian and Overall Road Traffic Crash Deaths—United States and 27 Other High-Income Countries (Morbidity and Mortality Weekly Report (MMWR)). CDC. https://www.cdc.gov/mmwr/volumes/74/wr/mm7408a2.htm

Nazemi, M., Rababah, B., Ramos, D., Zhao, T., and Farooq, B. (2025). Decoding pedestrian stress on urban streets using electrodermal activity monitoring in virtual immersive reality. Transportation Research Part C: Emerging Technologies, 171, 104952. https://doi.org/10.1016/j.trc.2024.104952

NHTSA. (2024a). Traffic Saftey Facts 2022 Data. U.S. Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813590

NHTSA. (2024b). FARS Web-Based Encyclopedia. U.S. Department of Transportation. https://www-fars.nhtsa.dot.gov/Main/index.aspx

Payne, B. K., and Hannay, J. W. (2021). Implicit bias reflects systemic racism. Trends in Cognitive Sciences, 25(11), 927–936. https://doi.org/10.1016/j.tics.2021.08.001

Pechteep, P., Luathep, P., Jaensirisak, S., and Kronprasert, N. (2024). Analysis of Factors Influencing Driver Yielding Behavior at Midblock Crosswalks on Urban Arterial Roads in Thailand. Sustainability, 16(10), Article 10. https://doi.org/10.3390/su16104118

Piazza, A. J., Knowlden, A. P., Hibberd, E., Leeper, J., Paschal, A. M., and Usdan, S. (2022). Distracted mobile device use among street-crossing college student pedestrians: An observational approach. Journal of American College Health : J of ACH, 70(7), 2135–2142. https://doi.org/10.1080/07448481.2020.1845182

Pinnow, J., Masoud, M., Elhenawy, M., and Glaser, S. (2021). A review of naturalistic driving study surrogates and surrogate indicator viability within the context of different road geometries. Accident Analysis and Prevention, 157, 106185. https://doi.org/10.1016/j.aap.2021.106185

Rafe, A., Singleton, P. A., Boyer, S., and Mekker, M. (2025). Pedestrian crossing behaviors at signalized intersections in Utah: Factors affecting spatial and temporal violations. International Journal of Transportation Science and Technology. https://doi.org/10.1016/j.ijtst.2024.12.005

Raoniar, R., Singh, S., Pathak, A., and Maurya, A. K. (2024). Analysis of Pedestrian–Vehicle Interaction Dynamics at Signalized Urban Intersections: A Surrogate Safety Measure Approach. Transportation Research Record, 03611981241289412. https://doi.org/10.1177/03611981241289412

Reniers, R. L. E. P., Murphy, L., Lin, A., Bartolomé, S. P., and Wood, S. J. (2016). Risk Perception and Risk-Taking Behaviour during Adolescence: The Influence of Personality and Gender. PLOS ONE, 11(4), e0153842. https://doi.org/10.1371/journal.pone.0153842

Richards, J., and Kuhn, R. (2022). Unsheltered Homelessness and Health: A Literature Review. AJPM Focus, 2(1), 100043. https://doi.org/10.1016/j.focus.2022.100043

Robbennolt, D., Hardman, S., Firestone, J., and Bhat, C. R. (2025). A model of electric vehicle adoption and motivating reasons for adoption. Transportation Research Part D: Transport and Environment, 146, 104906. https://doi.org/10.1016/j.trd.2025.104906

Roll, J., and McNeil, N. (2022). Race and income disparities in pedestrian injuries: Factors influencing pedestrian safety inequity. Transportation Research Part D: Transport and Environment, 107, 103294. https://doi.org/10.1016/j.trd.2022.103294

Russo, B. J., James, E., Aguilar, C. Y., and Smaglik, E. J. (2018). Pedestrian Behavior at Signalized Intersection Crosswalks: Observational Study of Factors Associated with Distracted Walking, Pedestrian Violations, and Walking Speed. Transportation Research Record, 2672(35), 1–12. https://doi.org/10.1177/0361198118759949

Sakib, S., Deb, K., Kumar Dhar, P., and Kwon, O.-J. (2022). A Framework for Pedestrian Attribute Recognition Using Deep Learning. https://www.mdpi.com/2076-3417/12/2/622

Sanders, R. L., and Schneider, R. J. (2022). An exploration of pedestrian fatalities by race in the United States. Transportation Research Part D: Transport and Environment, 107, 103298. https://doi.org/10.1016/j.trd.2022.103298

Sanders, R. L., Schneider, R. J., and Proulx, F. R. (2022). Pedestrian fatalities in darkness: What do we know, and what can be done? Transport Policy, 120, 23–39. https://doi.org/10.1016/j.tranpol.2022.02.010

Schneider, R. J., Sanatizadeh, A., Shaon, M. R. R., He, Z., and Qin, X. (2018). Exploratory analysis of driver yielding at low-speed, uncontrolled crosswalks in Milwaukee, Wisconsin. Transportation Research Record, 2672(35), 21–32.

Schwebel, D. C., Canter, M. F., Hasan, R., Griffin, R., White, T. R., and Johnston, A. (2022). Distracted pedestrian behavior: An observational study of risk by situational environments. Traffic Injury Prevention, 23(6), 346–351. https://doi.org/10.1080/15389588.2022.2076282

Shinobi Systems. (2025). ShinobiDocs. https://docs.shinobi.video/

Shrinivas, V., Bastien, C., Davies, H., Daneshkhah, A., and Hardwicke, J. (2023). Parameters influencing pedestrian injury and severity – A systematic review and meta-analysis. Transportation Engineering, 11, 100158. https://doi.org/10.1016/j.treng.2022.100158

Skinner-Dorkenoo, A. L., George, M., Wages, J. E., Sánchez, S., and Perry, S. P. (2023). A systemic approach to the psychology of racial bias within individuals and society. Nature Reviews Psychology, 2(7), 392–406. https://doi.org/10.1038/s44159-023-00190-z

Smart Growth America. (2024). Dangerous by Design 2024. https://smartgrowthamerica.org/resources/dangerous-by-design-2024/

Speer, J. (2019). Urban makeovers, homeless encampments, and the aesthetics of displacement. Social and Cultural Geography, 20(4), 575–595. https://doi.org/10.1080/14649365.2018.1509115

Steimer, T. (2002). The biology of fear- and anxiety-related behaviors. Dialogues in Clinical Neuroscience, 4(3), 231–249.

Texas Transportation Code, Tex. Transp. Code § 552 (2023). https://statutes.capitol.texas.gov/docs/tn/htm/tn.552.htm

The White House. (2024, July 3). Racial Discrimination in Contemporary America. https://bidenwhitehouse.archives.gov/cea/written-materials/2024/07/03/racial-discrimination-in-contemporary-america/

Theofilatos, A., Ziakopoulos, A., Oviedo-Trespalacios, O., and Timmis, A. (2021). To cross or not to cross? Review and meta-analysis of pedestrian gap acceptance decisions at midblock street crossings. Journal of Transport and Health, 22, 101108. https://doi.org/10.1016/j.jth.2021.101108

Train, K. E. (2009). Logit. In Discrete Choice Methods with Simulation (2nd ed., pp 34-75). Cambridge University Press.

U.S. Census Bureau. (2023). American Community Survey 2019–2023 5-year estimates (ACS2023_5yr) [Data set]. Social Explorer.

U.S. Department of Transportation. (2024). National Highway Traffic Safety Administration (NHTSA) [Dataset]. https://www.nhtsa.gov/crash-data-systems/fatality-analysis-reporting-system

USDOT. (2024). Promising Practices for Transportation Agencies to Address Road Safety among People Experiencing Homelessness. Federal Highway Administration. https://highways.dot.gov/sites/fhwa.dot.gov/files/Promising_Practices_to_Address_Road_Safety_among_People_Experiencing_Homelessness_FHWA.pdf

Wages III, J. E., Perry, S. P., Skinner-Dorkenoo, A. L., and Bodenhausen, G. V. (2022). Reckless gambles and responsible ventures: Racialized prototypes of risk-taking. Journal of Personality and Social Psychology, 122(2), 202–221. https://doi.org/10.1037/pspa0000287

Wan, Y., Xu, Y., Xu, Y., Wang, H., Wang, J., and Liu, M. (2023). A Computer Vision-Based Algorithm for Detecting Vehicle Yielding to Pedestrians. Sustainability, 15(22), Article 22. https://doi.org/10.3390/su152215714

Wasserstein, R.L., Lazar, N.A., 2016. The ASA Statement on P-Values: Context, Process, and Purpose. The American Statistician 70, 129–133. https://doi.org/10.1080/00031305.2016.1154108

Wasserstein, R.L., Schirm, A.L., Lazar, N.A., 2019. Moving to a World Beyond "p < 0.05." The American Statistician 73, 1–19. https://doi.org/10.1080/00031305.2019.1583913

Wells, H. L., McClure, L. A., Porter, B. E., and Schwebel, D. C. (2018). Distracted pedestrian behavior on two urban college campuses. Journal of Community Health, 43(1), 96–102. https://doi.org/10.1007/s10900-017-0392-x

Widera, J. (2024). Pedestrian Mobilities at the Crossroads: The Contestation and Regulation of Jaywalking. International Journal of Communication, 18, Forum 815–821.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. Journal of Econometrics, 68(1), 115–132. https://doi.org/10.1016/0304-4076(94)01645-G

Xie, M., Cheng, W., Gill, G. S., Zhou, J., Jia, X., and Choi, S. (2018). Investigation of hit-and-run crash occurrence and severity using real-time loop detector data and hierarchical Bayesian binary logit model with random effects. Traffic Injury Prevention, 19(2), 207–213. https://doi.org/10.1080/15389588.2017.1371302

Xie, T. H., Ahuja, M., McCutcheon, V. V., and Bucholz, K. K. (2020). Associations between Racial and Socioeconomic Discrimination and Risk Behaviors among African-American Adolescents and Young Adults: A Latent Class Analysis. Social Psychiatry and Psychiatric Epidemiology, 55(11), 1479–1489. https://doi.org/10.1007/s00127-020-01884-y

Xiong, H., Ma, L., Ning, M., Zhao, X., and Weng, J. (2019). The tolerable waiting time: A generalized Pareto distribution model with empirical investigation. Computers and Industrial Engineering, 137, 106019. https://doi.org/10.1016/j.cie.2019.106019

Yu, C.-Y., Zhu, X., and Lee, C. (2022). Income and Racial Disparity and the Role of the Built Environment in Pedestrian Injuries. Journal of Planning Education and Research, 42(2), 136–149. https://doi.org/10.1177/0739456X18807759

Zafri, N. M., Tabassum, T., Himal, Md. R. H., Sultana, R., and Debnath, A. K. (2022). Effect of pedestrian characteristics and their road crossing behaviors on driver yielding behavior at controlled intersections. Journal of Safety Research, 81, 1–8. https://doi.org/10.1016/j.jsr.2022.01.001

Zafri, N. M. (2023). Walk or run? An observational study to explore pedestrian crossing decision-making at intersections in Dhaka, Bangladesh applying rules mining technique. Transportation Research Part F: Traffic Psychology and Behaviour, 94, 83–98. https://doi.org/10.1016/j.trf.2023.02.003

Zhang, G., Tan, Y., and Jou, R.-C. (2016). Factors influencing traffic signal violations by car drivers, cyclists, and pedestrians: A case study from Guangdong, China. Transportation Research Part F: Traffic Psychology and Behaviour, 42, 205–216. https://doi.org/10.1016/j.trf.2016.08.001

Zhang, W., Guo, H., Wang, C., Wang, K., Huang, W., Xu, Q., Tang, H., Yang, B., and Yan, R. (2023). Analysis of pedestrian illegal crossing at unmarked segments: Environmental factors, pedestrian characteristics and crossing behaviours. Transportation Research Part F: Traffic Psychology and Behaviour, 99, 339–355. https://doi.org/10.1016/j.trf.2023.10.022

Zhou, H., Romero, S. B., and Qin, X. (2016). An extension of the theory of planned behavior to predict pedestrians' violating crossing behavior using structural equation modeling. Accident Analysis and Prevention, 95, 417–424. https://doi.org/10.1016/j.aap.2015.09.009

Zhu, D., Sze, N. N., and Bai, L. (2021). Roles of personal and environmental factors in the red light running propensity of pedestrian: Case study at the urban crosswalks. Transportation Research Part F: Traffic Psychology and Behaviour, 76, 47–58. https://doi.org/10.1016/j.trf.2020.11.001

Zimmerman, B. (2023). Living Without Refuge: How the Housing Crisis Fuels Traffic Violence (Streetblog USA). https://usa.streetsblog.org/2023/10/24/living-without-refuge-how-the-housing-crisis-fuels-traffic-violence