Journal of Safety Research 93 (2025) 156–169

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Direct and indirect effects of road attributes on traffic safety

Wookjae Yang [a,1], Sangjin Han [b,*,2]

[a] Urban Policy & Administration, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 22012 Republic of Korea
[b] Transport Studies Group, SNU Environmental Planning Institute, Graduate School of Environmental Studies, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826 Republic of Korea

ARTICLE INFO

Keywords:
Road design attributes
Traffic safety
Mediating effects
Star rating score (SRS)
Road assessment program (RAP)
Piecewise structural equation modeling

ABSTRACT

Introduction: Providing safe road infrastructure is increasingly gaining attention worldwide as part of the effort to reduce road deaths and injuries. Halving road fatalities by 2030 is one of the targets among the Sustainable Development Goals of the United Nations. This study aims to understand how road traffic conditions, including vehicle speeds and volumes, mediate the relationship between road design and traffic safety. *Method:* In particular, the study relies on 78 road attributes pertaining to road design features, as published in the International Road Assessment Programme's (iRAP) Star Rating and Investment Plan Manual. The star rating was conducted along a 68 km segment of a national highway in Korea, and these coded data were associated with both traffic and crash data. The traffic data, in this case, the average vehicle volume and speed, are sourced from the View-T platform in Korea. Crash data were obtained from Traffic Accident Analysis Systems in Korea. *Results:* The application of principal component analysis (PCA) identified three principal components—principal component (PC) 1 represents attributes related to pedestrians and roadside facilities, PC2 represents cross-sections and intersections, and PC3 represents attributes related to road surfaces and curvatures. In addition, piecewise structural equation modeling confirmed that PC1 is the only component that has a direct effect on the number of crashes. *Practical applications:* The finding suggests that pedestrian and roadside facilities easily added or removed during road operations are more critical than geometric attributes established at the road design stage. The study also confirms the indirect effects of the aspects of speed and volume on the likelihood of a crash.

1. Introduction

Fatalities and injuries from road traffic crashes threaten the life and well-being of everyone. The World Health Organization (WHO) has set what is known as Global Sustainable Development Goal – SDG Target 3.6., which aims to halve deaths and injuries from crashes by the end of 2030 (World Health Organization, 2023). This global plan emphasizes the importance of a holistic approach to road safety and calls for continued improvements in the design of roads and vehicles, law enforcement, and the provision of timely, life-saving emergency care for the injured (World Health Organization, 2023). Road safety improvements also aim to promote road safety and walking, cycling, and public transportation usage as inherently healthy and sustainable travel modes (Hong & Yang, 2023).

Despite numerous efforts to mitigate traffic crashes, they remain a leading cause of death and injuries. Road traffic crashes in the world result in the yearly deaths of approximately 1.19 million people and nonfatal injuries of 20 and 50 million people (Toroyan, 2009). Based on 2019 data on the age distribution of all-cause mortality, road traffic injuries remain the leading cause of death for children and young people aged 5–29 years and represent the 12th leading cause of death when all ages are considered (World Health Organization, 2023).

Better road designs can improve road safety considerably. Designing transportation infrastructure to allow for human errors and injury thresholds can substantially lessen the severity of crashes that may occur (Xu et al., 2022). In South Korea, several measures such as additional pedestrian crossings, safety fences, 'silver zones' near facilities for older adults, bike lanes, and other traffic-calming measures are being installed to reduce the risk of injury among these road users (International Transport Forum, 2024). In this respect, the International Road Assessment Program (iRAP) has rated over 502,000 km of roads across 84 countries to help identify risky roads and improve infrastructure,

* Corresponding author.
E-mail addresses: wookjae.yang@inu.ac.kr (W. Yang), jinmike@snu.ac.kr (S. Han).
[1] ORCID: https://orcid.org/0000-0002-6305-3877.
[2] ORCID: https://orcid.org/0000-0003-4596-2935.

https://doi.org/10.1016/j.jsr.2025.02.017
Received 11 September 2024; Received in revised form 8 December 2024; Accepted 6 February 2025
Available online 20 February 2025
0022-4375/

with more than a three-star rating since 2006 (International Road Assessment Programme, 2021). Aligned with the Global Sustainable Development Goal set by the World Health Organization, iRAP released a road safety plan to ensure more than 75% of travel is on roads rated three stars or better, which is expected to save more than two million road users from crashes.

However, the lack of a comprehensive understanding of the mediating effects of vehicle speed and volume levels could lead to biased results when explaining the effects of road design features on safety outcomes. Findings from previous studies elicit inconsistent results regarding the impact of specific road attributes on safety outcomes (Basyouny, 2016; Merlin et al., 2020; Hong & Yang, 2023). This signifies the need to consider the potential impact of factors such as the traffic volume, vehicle mix, capacity and level of service, prevailing speeds, cross-sectional alignment, development type, access to adjacent properties, segment lengths, traffic volumes, access point density levels, and midblock changes (Manuel et al., 2014).

Furthermore, fatal crashes also occur when speeds and design are not compatible. Road safety results from the composite effect of cross-sectional roadway design attributes. We cannot simply determine a single road attribute factor contributing to crashes (Merlin et al., 2020). For example, according to Ewing et al. (2024), reducing lane widths on urban arterials can provide more space to include other street features such as bicycle lanes, on-street parking, wider sidewalks, landscaped buffers, and reduced pedestrian crossing distances. However, these attributes interact and bring about different safety outcomes for various road users according to varying traffic conditions on the roadway although traffic volume contribute to the likelihood of crashes substantially according to Yannis et al (2014). Thus, it is necessary to quantify the combined effects (direct and indirect effects) of road design and traffic conditions on road safety for all users.

This research aims to fill this gap by quantifying the combined effects of road design and traffic conditions on road safety outcomes. By considering the interactions between different road attributes and traffic conditions, the study seeks to provide insights into effective strategies for improving road safety for all users. A conceptual framework of this study is presented in Fig. 1. This illustrates how road attributes such as cross-sectional design elements, roadside objects, additional safety facilities, pedestrian facilities, and horizontal/vertical alignments affect crash rates through traffic speed and volume mediators. The mediators, in this case, traffic volume and speed, are, in turn, the primary determinants of crash rates.

2. Literature review

The relationship between road attributes and safety has long been studied. Numerous studies are reviewed here in the order of cross-sectional design elements, horizontal and vertical design elements, pedestrian facilities, roadside objects, and other safety facilities. Table 1 provides a comprehensive summary of the road attributes, their associated safety outcomes (positive or negative), and mediators that affect the relationship between road attributes and safety outcomes.

Cross-sectional design elements include the lane width, shoulder, number of lanes, and median type. The relationship between these elements and crashes suggests nuanced effects influenced by vehicle speeds and volume. First, a lane width wider than specifically three meters is typically discouraged due to unintended speeding that occurs along such lanes (NACTO, 2013), mainly because driving speeds can be reduced on narrower lanes given the increased steering workload (Godley et al., 2004). A more recent study confirmed that the effect of the lane width on non-intersection crashes varied depending on the driving speed, finding that wider lanes at 30–35 mph result in more crashes than narrower lanes (Hamidi & Ewing, 2023). Another cross-sectional design element, the presence of a shoulder, contributed to higher vehicle speeds, which in turn increased the likelihood of crashes (Gargoum & El-Basyouny, 2016).

An increased number of lanes generally results in higher traffic flows or more frequent lane changes between multiple lanes while increasing the crash risk due to greater exposure to traffic conflicts (Milton & Mannering, 1998). On the other hand, a higher number of lanes and medians with barriers or upward slopes lead to reduced injury risks as these attributes decrease vehicle speeds (Dumbaugh et al., 2024; Gargoum & El-Basyouny, 2016; Hu & Donnell, 2010; Ukkusuri et al., 2012). However, expanding a road from two lanes to four lanes can lead to a 20% reduction in crashes (Dumbaugh et al., 2024; Gargoum & El-Basyouny, 2016). The increased traffic volume from the road expansion generates traffic congestion, forcing drivers to reduce their speeds (Gargoum & El-Basyouny, 2016). In addition, the number of total crashes causing severity tends to be lower in census block groups with a higher proportion of two-lane roads and a higher proportion of roads with posted speed limits of 35 mph or less, as the likelihood of conflicts and speed differentials between vehicles is greatly reduced on two-lane

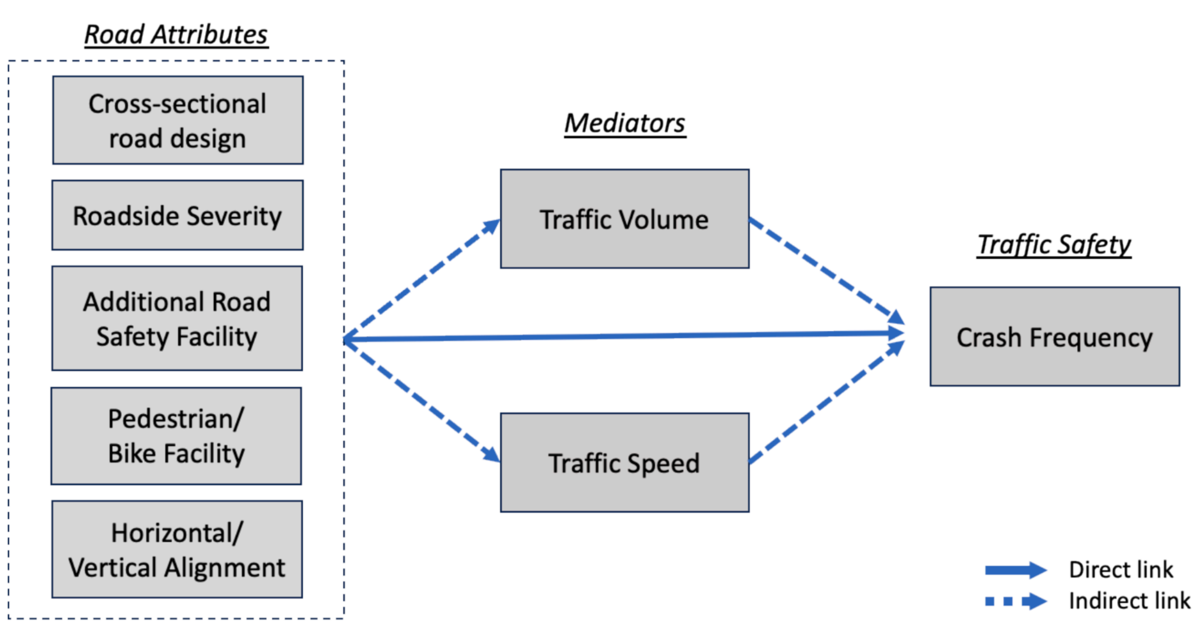

Fig. 1. Conceptual framework.

Table 1
Summary of the literature review.

Factor	Increasing crashes (+)	Reducing crashes (−)	Conditional factor (*: mediator)
Cross-sectional Design Elements			
Lane Width	Hamidi & Ewing 2023; NACTO, 2013	Potts et al., 2007	Vehicle speed, Average daily traffic
Shoulder	Bamzai et al., 2017; Gargoum & El-Basyouny, 2016*	Gitelman et al., 2019	Average speed*, Speed limit, Annual Average daily traffic, Average daily traffic
Number of Lanes	–	Council & Stewart, 2000; Dumbaugh et al., 2024; Gargoum & El-Basyouny, 2016; Saha et al., 2020	Operating speeds, Posted speed limit, Traffic volume, Vehicle miles traveled
Median (Type)	Elvik, 1995	Hu & Donnell, 2010; Knuiman et al., 1993; Saha et al., 2020	Vehicle speed, Speed limit, Average daily traffic, Daily vehicle miles traveled
Horizontal & Vertical Design Elements			
Curvature Radius	Wu et al., 2013	Gooch et al., 2016; Watson et al., 2014	Vehicle speed, Traffic volume, Annual Average daily traffic
Grade	Lan et al., 2011; Wong, 2005	–	Traffic flow, curve radius, slope length
Sight Distance	Abdulhafedh, 2020	Fhwa, 2013; Himes et al., 2016; Potts et al., 2019	Intersection type, roadside objects, traffic volume
Pedestrian and Bike Facilities			
Sidewalk	–	Abou-Senna et al., 2022; McMahon et al., 2002; Osama & Sayed, 2017; Quistberg et al., 2015	–
Crosswalk	Zegeer et al., 2001	Gitelman et al., 2019	–
Roadside Objects			
Roadside safety barriers/ fixed objects	Albuquerque & Awadalla, 2020	Eskandarian et al., 1997; Holdridge et al., 2005	Object rigidity, vehicle speed
Street parking	Kraidi & Evdorides, 2020	Dumbaugh, 2005; Ossenbruggen et al., 2001	–
Other Safety Facilities			
Rumble Strips	–	Karkle et al., 2013; Sayed et al., 2010	Installation location
Delineation	Stanton & Pinto, 2000	Bektas et al., 2016	Driving speed, delineation color
Streetlight	–	Garber & Gadiraju, 1988	Vehicle speed

roads (Saha et al., 2020).

Horizontal and vertical design elements consist of the curvature, grade, and sight distance aspects. Roads adjacent to sharp curves can reduce crashes because drivers travel cautiously and at lower speeds near curves (Gooch et al., 2016), However, small-radius curves with higher slopes increase the likelihood of medium and severe injuries (Wang et al., 2019). This correlation has been assessed with factors such as the traffic volume, slope length, curve, and other road geometry factors (Lan et al., 2011; Wang et al., 2019). The effect of the slope on crashes becomes more significant at higher traffic flows.

The effects of the sight distance at intersections vary depending on the traffic volume (Himes et al., 2016). Higher crash rates are associated with differences between the observed free-flow speeds and speed as mandated by the curve radius or sight distance (Watson et al., 2014). This implies that sight distance recommendations are crucial for mitigating collision risks and ensuring smoother traffic flows within roundabouts, implying, in turn, that considering designs and regulations to improve road safety is important (Zirkel et al., 2013).

Roadside objects include roadside safety barriers, fixed objects, and street parking areas. Roadside objects are typically evaluated according to their ability to redirect vehicles or reduce speeds (Eskandarian et al., 1997). Collisions with rigid barriers are not necessarily safer than collisions with less rigid objects, such as trees or poles, without controlling for design speeds (De Albuquerque & Awadalla, 2020). For example, crashes related to construction machinery tend to result in less severe injuries than those related to rock banks with ledges. Considering the similar rigidity between the two, the lower speeds typically observed in work zones with construction machinery may lead to a decrease in severe injuries (Holdridge et al., 2005).

Other safety facilities include the presence of rumble strips, delineation, and streetlights. Improving the visibility of road edges or lane delineators improves confidence and stress levels, leading to increased driving speeds and more frequent collisions (Stanton & Pinto, 2000). On the other hand, keeping longitudinal pavement markings in good condition has significant positive safety effects (Bektas et al., 2016).

Moreover, streetlights have a greater influence when vehicle speeds are higher or when there are a greater number of lanes (Xu et al., 2018). Improved illuminance can reduce vehicle speed variations and improve safety levels, consequently decreasing crashes (Garber & Gadiraju, 1988). Therefore, a thorough understanding is required to discern the impacts of streetlight improvements, particularly considering vehicle speeds.

In short, road safety is the outcome caused by the composite effect of various roadway design attributes. Many studies have attempted to reveal the direct effects of road attributes (George et al., 2017), speeds (Gitelman et al., 2018; Roshandel et al., 2015; L. Wang et al., 2015), and volume (Roshandel et al., 2015; Wang et al., 2015). We cannot determine a single factor contributing to crashes. Road design attributes interact with each other and can bring about different safety outcomes, which may also vary by traffic conditions, such as volumes and speeds. Some review studies have attempted to conceptualize a plausible crash mechanism as follows: built environment −> mediator −> crash frequency. This suggests that the built environment affects crash frequency via mediators, which include traffic exposure, speed, and conflicts (Ewing & Dumbaugh, 2009; Merlin et al., 2020). However, unlike the active discussions on a macro level of an urban environment, relatively less has been paid to empirical evidence of road attribute level to investigate the mediating effects of vehicle speeds and volumes on crash likelihood. This is the main literature gap in this paper.

3. Data and variables

3.1. Exogenous variables

This study focuses on the geometric features and traffic characteristics of a 68-kilometer segment of a national highway between Seryu, Suwon-si, and Bal-ahn, Hwaseong-si in South Korea. Road attributes, exogenous variables in our study, were surveyed and evaluated in 100-meter intervals following the iRAP coding manual for all study segments. Since we defined the surveying area before conducting the survey, the 68-kilometer segment is divided evenly into 680 segments. Therefore, we do not need to consider segments smaller than this limit. The International Road Assessment Programme (iRAP) has developed the iRAP Coding Manual to assess road attributes with the same standards consistently. It includes 78 road design attributes, such as the number of lanes, the lane width, and curvatures, among others. Georeferenced images collected during a survey were used to record road attributes for each 100-meter road segment. The unit of analysis is a 100-meter road segment. Coding options for each attribute are listed in the order of highest to lowest risk, and the 'worst' option is coded.

3.2. Endogenous variables

The endogenous variables consist of crash data, traffic data, and speed data. First, traffic data were sourced from the Traffic Accident Analysis System (TAAS) in Korea. TAAS database provides complete records of traffic crashes, including offender and victim's age, sex, and crash information such as date, cause, and injuries etc. The TAAS integrates various databases of road crashes maintained by different organizations, including police, car insurance companies, and hospital databases. However, the actual location data, x-y coordinates of crashes, were extracted from the TAAS interactive map by using the web scraping method. Then we merged the location data into crash information data grouped by crash ID.

The spatial joining process was conducted to merge crash data and 100-meter interval road segments based on the proximity criterion '*st_nearest_feature*' in R studio. This criterion enabled the identification of the closest crash occurrences from each 100-meter road segment. In cases of borderline crashes, the incidents were assigned to the preceding road segment.

Table 2 shows the number of crashes on the case study roadways from 2017 to 2021. The data for these five years were aggregated by summing crash counts at 100-meter intervals. The collected crash data include two types of crashes: head-on and run-off crashes. The Star Rating Score (SRS), as per the iRAP protocol, is calculated primarily based on head-on and run-off crashes. By focusing on these crash types, the study aligns with the iRAP methodology, ensuring consistency and relevance in road safety assessment.

However, this study does not differentiate intersection crashes from non-intersection crashes. In general iRAP protocol, intersections are treated separately to calculate star rating scores which denote the likelihood or severity of intersection crashes. Since this study considers head-on and run-off crashes along the roads regardless of whether road sections belong to intersections or not, some road sections can have intersection attributes such as intersection quality, intersection type, intersecting road volume, and intersection channelization. Possibly those road sections within intersections can be excluded to comply with iRAP protocol, but this study included them to examine whether or not intersection attributes can affect head-on and run-off crashes.

Fig. 2 presents a heatmap of crash counts by crash type at 100-meter intervals. Fig. 3 shows all crash counts from the starting point to the ending point of the case study roads. The mean head-on crash count per 100-meter interval is 1.09, and the standard deviation is 1.81. The data range from 0 to 9. Regarding run-off crashes, the data show a mean crash count of 1.009 and a standard deviation of 1.417. The data range in this case is from 0 to 9.

Traffic volume and speed data are collected separately from the View-T database, which is a traffic data platform that provides estimated annual average daily traffic and average speeds for road sections along the selected highway network. A map-matching procedure was applied to each 100-meter segment to identify the corresponding traffic volume and speed. The mean vehicle speed for the sample segments is 47.7 km per hour, with a standard deviation of 13.8. The mean vehicle volume for these segments is 32,600, with a standard deviation of 13,700.

These two variables are assumed to be mediators, influenced by exogenous variables and, in turn, affecting crash outcomes. For temporal matching with crash data, vehicle speed, and volume were extracted from View-T within the timeframe of 2017 to 2021. Those variables, including vehicle speed and volume, were log-transformed to normalize their distributions, which is a key assumption for the linear regression model.

However, we did not address the correlation between vehicle speed and volume. When monitoring a single road segment continuously, explaining the sensitivity of speed to volume is challenging. This is because our analysis was based on daily averages rather than hourly profiles.

Table 3 presents a detailed description of both the exogenous and endogenous data, accompanied by corresponding descriptive statistics.

Table 2
Total number of crashes and the number of crashes by year on case study roadways.

Crash Type	Total Number of Crashes	Year				
		2017	2018	2019	2020	2021
Head-on Crashes	677	345	87	91	33	121
Run-off Crashes	627	113	168	132	104	110

4. Methodology

This study utilizes Principal Component Analysis (PCA) and Piecewise Structural Equation Modeling (PSEM) to investigate the interrelationships among variables affecting road safety. It aims to explore the direct and indirect effects of lower dimensions of principle components calculated by PCA and the mediating effects of vehicle speed and volume on head-on and run-off crashes.

This study initially applies a principal component analysis (PCA) to the attribute data to identify common factors that explain the correlations among the variables, thereby providing a set of highly associated variables representing a common factor (Kim, 2008). The objective of the PCA is not to explain correlations among the variables but to account for as much variance as possible within the data. In the PCA analysis procedure, there is no requirement for hypothetical underlying factors, and a component is merely a combination of correlated variables (Tabachnick & Fidell, 2001). Instead, it focuses on the variance–covariance structure of the data itself. In this context, the main purpose of the PCA is to summarize numerous variables into a smaller number of components, thus achieving data reduction. Therefore, we used PCA to find common factors to reduce the high dimensions of iRAP's 78 variables to lower dimensions without losing the much of the embedded information in the data set. It should also be noted that attributes with constant values throughout all sample segments are removed including the lane width, grade, and rumble strips, given that such variables would not affect the PCA results.

Piecewise Structural Equation Modeling was then applied to investigate causal links among multiple factors. This method is based on a comprehensive path analysis and a regression analysis, supplemented by an analysis of variance (Shi et al., 2017). SEM enables testing for the presence of mediated effects. A mediation analysis is a statistical method that serves to clarify and simultaneously evaluate direct and indirect pathways within a complex network of variables. This makes SEM an ideal tool for probing multiple hypotheses about the different processes operating in a system (Deutsch et al., 2020). For example, Elvik et al. (2004) conducted a mediation analysis via a path model to control all possible confounding, mediation, and moderation effects in the speed-safety relationship.

Unlike the globally computed SEM, which estimates all relationships simultaneously, a locally assessed SEM, such as the piecewise SEM (PSEM), models each response (or endogenous) variable independently (Grinstead et al., 2023). The benefit of PSEM is that power for a specific path is determined by the information within that path rather than relying on the variance–covariance matrix for the entire system of paths. In such a model, statistical power is determined for each local relationship in the regression model. This implies that a model with two

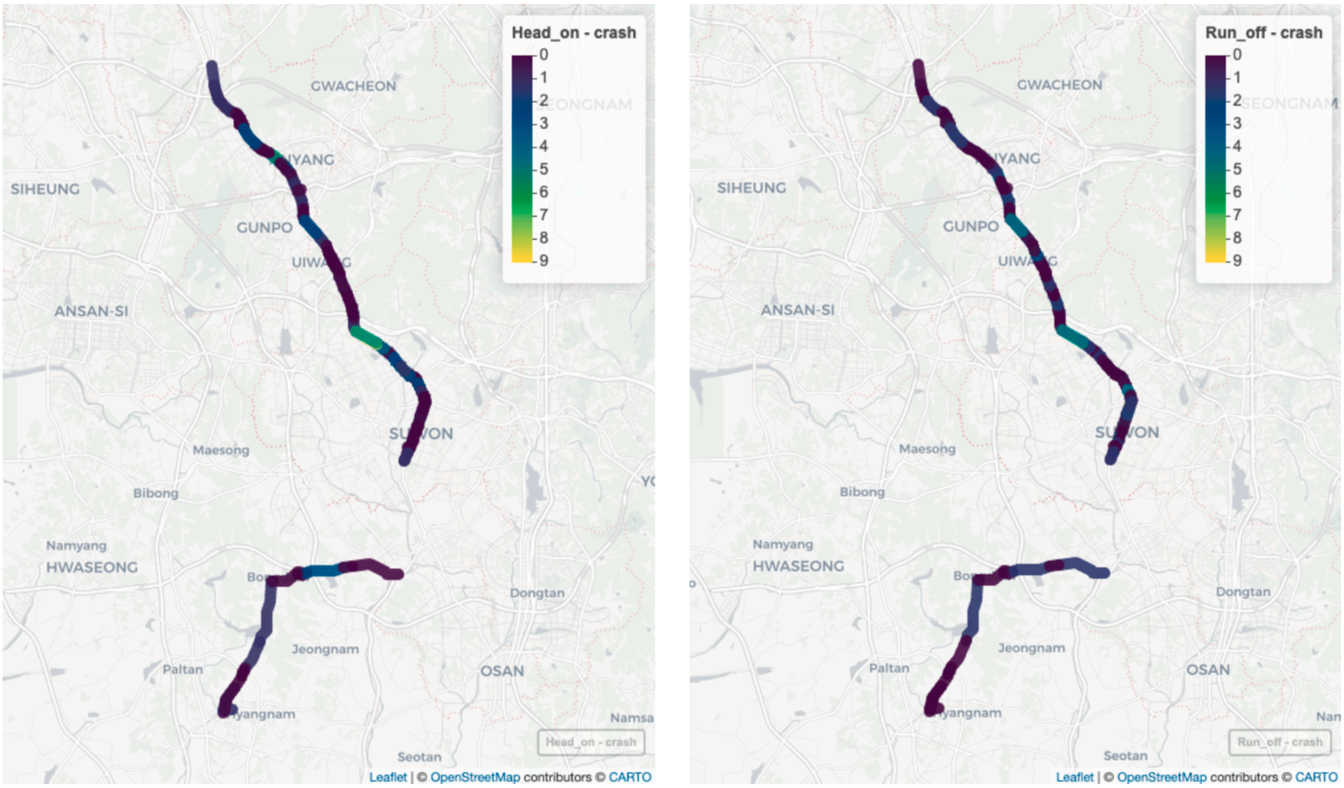

Fig. 2. Total crash counts for 5-year by crash type at 100 m-intervals (left: head-on crash, right: run-off crash).

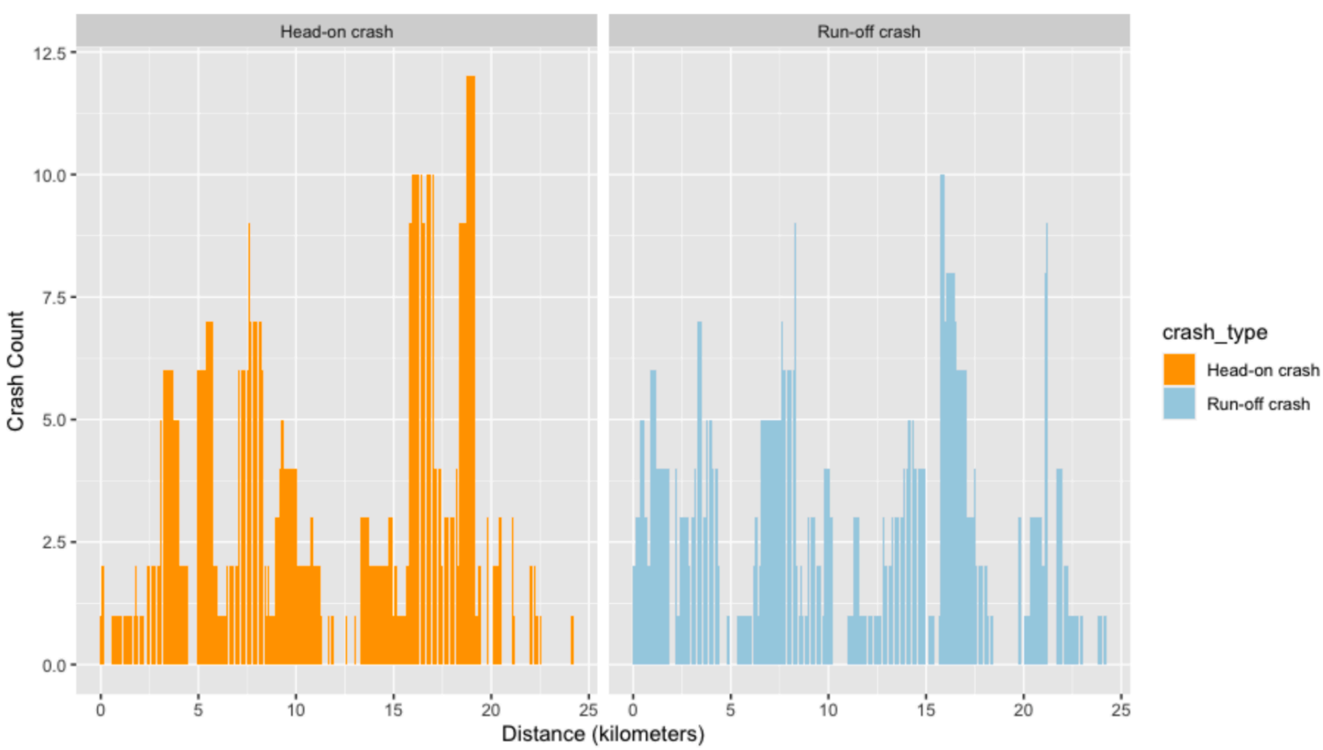

Fig. 3. Crash counts from the starting point to the ending point of case road segments at 100-meter intervals (bi-directional length: 0 km to 68 km).

local regression relationships will have two separate variance–covariance matrices, requiring less power than a single matrix representing all relationships simultaneously. It maintains the key generalizations of SEMs, including the modeling of non-continuous endogenous variables, hierarchical or nested non-independent observations, and different estimation procedures (e.g., maximum likelihood, least-squares; Grinstead et al., 2023).

In our modeling, while vehicle volume and speed are the continuous

Table 3
Descriptive statistics.

Variable	Definition	Mean	SD	Min	Max
Exogenous variables					
Cross-sectional design elements					
Lane width	The distance from the center of the edge line to the center of the adjacent lane marking (Narrow (≥ 0 m to < 2.75 m): 3, Medium (≥ 2.75 m to < 3.25 m): 2, Wide (≥ 3.25 m): 1)	2.00	0.00	2	2
Number of lanes	The number of traffic lanes in the direction of travel (three and two lanes:6, two and one lanes:5, four or more lanes:4, three lanes:3, two lanes:2, one lane:1)	3.24	0.80	2	4
Paved shoulder – driver side	The width of the safe and drivable section of road from the edge line to the edge of the paving at driver side (no paved shoulder: 4, narrow 0 m to < 1 m: 3, medium 1 m to < 2.4 m: 2, wide ≥ 2.4 m:1)	3.01	0.11	3	4
Paved shoulder – passenger side	The width of the safe and drivable section of road from the edge line to the edge of the paving at the passenger side (no paved shoulder: 4, narrow 0 m to < 1 m: 3, medium 1 m to < 2.4 m: 2, wide ≥ 2.4 m:1)	3.06	0.27	2	4
Horizontal & Vertical Alignment					
Curvature	The horizontal alignment of the road. Curvature is gauged according to the approximate curve radius and the appropriate safe approach and driven speed under normal conditions (very sharp:4, sharp:3, moderate:2, straight or gently curving: 1).	1.16	0.41	1	3
Quality of curve	The quality of the curve will reflect the extent to which signs and markings help the driver judge the correct curvature and the sight distance in advance of and around the curve (poor:3, adequate:2, straight: 1)	2.85	0.36	1	3

Table 3 (*continued*)

Variable	Definition	Mean	SD	Min	Max
Grade	The gradient of the road along its length. Grade refers to both upward and downward slopes (≥10%: 5, 7.5% to < 10%: 4, 0% to < 7.5%:1).	1.00	0.00	1	1
Pedestrian and Bike Facility					
Pedestrian crossing quality	Effectiveness of the pedestrian crossing on the inspected road or side road. It is 'adequate' if a facility is visible and can be anticipated by vehicle drivers, and the facility is not obstructed (no pedestrian crossing facility: 3, poor: 2, adequate: 1)	2.49	0.68	1	3
Sidewalk – driver side	The presence of pedestrian sidewalks on the driver side of the road (informal path – 0 m to < 1 m from road: 7, informal path – ≥1m from road: 6, none: 5, sidewalk – 0 m to < 1 m from road: 4, sidewalk – 1 m to < 3 m from road: 3, sidewalk – ≥3m from road: 2, sidewalk – physical barrier: 1)	4.53	0.83	1	5
Sidewalk – passenger side	The presence of pedestrian sidewalks on the passenger side of the road (informal path – 0 m to < 1 m from road: 7, informal path – ≥1m from road: 6, none: 5, sidewalk – 0 m to < 1 m from road: 4, sidewalk – 1 m to < 3 m from road: 3, sidewalk – ≥3m from road: 2, sidewalk – physical barrier: 1)	4.19	1.40	1	7
Pedestrian crossing facilities – inspected road	The presence of purpose-built pedestrian crossing facilities on the inspected road (no facility: 7, refuge only: 6, marked crossing only: 5, marked crossing with refuge: 4, signalized crossing: 3, signalized crossing with refuge: 2, grade separated facility: 1)	5.96	1.89	1	7
Pedestrian crossing	The presence of purpose-built	6.21	1.33	2	7

(*continued on next page*)

Table 3 (*continued*)

Variable	Definition	Mean	SD	Min	Max
facilities – intersecting road	pedestrian crossing facilities on the side (intersecting) road (no facility: 7, refuge only: 6, marked crossing only: 5, marked crossing with refuge: 4, signalized crossing: 3, signalized crossing with refuge: 2, grade separated facility: 1)				
Facilities for bicycles	The presence of purpose-built facilities for bicyclists (no specific facilities for bicycles or poor standard: 1, shared use path: 0)	0.541	0.499	0	1
Roadside Objects					
Roadside severity driver side – distance	distance from the edge line of nearest driving lane to a roadside object with the highest risk on the driver side (0 m to < 1 m: 4, 1 m to < 5 m: 3, 5 to 10 m: 2, ≥10 m: 1)	1.15	0.39	1	3
Roadside severity passenger side – distance	Distance from the edge line of the nearest driving lane to the roadside object with the highest risk on the passenger side (0 m to < 1 m: 4, 1 m to < 5 m: 3, 5 to 10 m: 2, ≥10 m: 1)	1.30	0.46	1	3
Vehicle parking	The extent of vehicle parking along the side of the road (two sides: 3, one side: 2, none: 1)	1.21	0.41	1	3
Other Safety Facilities					
Street lighting	The presence of street lighting (none: 2, present: 1)	1.94	0.23	1	2
Pedestrian fencing	The presence of pedestrian fencing or other barriers that effectively control the pedestrian crossing flow (not present: 2, present: 1)	1.09	0.29	1	2
Roadworks	The presence of major road construction or road works in progress (major road works: 3, minor: 2, no road works: 1)	1.10	0.33	1	3
Road surface	Road surface condition indicating a range from major defects to smooth for vehicles (Poor: 3, Medium: 2, Good: 1).	1.56	0.51	1	3
Delineation	The adequacy of road lines and markings (poor: 2, adequate: 1)	1.08	0.27	1	2
Skid resistance	Skidding resistance and texture depth of the road surface (unsealed – poor: 5, unsealed –adequate: 4, sealed – poor: 3, sealed – medium: 2, sealed – adequate: 1)	1.12	0.34	1	3
Sight distance	the ability of a driver to see and/or anticipate road conditions and other road users ahead (poor:2, adequate:1)	1.01	0.09	1	2
School zone – crossing supervisor	The presence of a crossing supervisor or warden (no school at the location: 3, crossing supervisor not present: 2, present: 1)	2.97	0.14	2	3
School zone – warning	The presence of a school zone with static signs, road, or flashing beacons (no school at the location: 4, no school zone warning (school present):3, static signs or road markings: 2, flashing beacons:1)	3.98	0.14	3	4
Traffic flow					
Pedestrian observed flow across the road	The number of pedestrians across the road within the coding segment (number of people > 8: 6, 6 ~ 7: 5, 4 ~ 5: 4, 2 ~ 3: 3, 1:2, 0: 1)	1.18	0.66	1	6
Pedestrians observed flow along the road – driver side	The number of pedestrians along the driver-side road within the coding segment (number of people > 8: 6, 6 ~ 7: 5, 4 ~ 5: 4, 2 ~ 3: 3, 1:2, 0: 1)	1.15	0.55	1	5
Pedestrian observed flow along the road – passenger side	The number of pedestrians along the passenger-side road within the coding segment (number of people > 8: 6, 6 ~ 7: 5, 4 ~ 5: 4, 2 ~ 3: 3, 1:2, 0: 1)	1.78	1.23	1	6
Motorcycle observed flow	The number of motorcycles within the coding segment (number of motorcycles > 8: 6, 6 ~ 7: 5, 4 ~ 5: 4, 2 ~ 3: 3, 1:2, 0: 1)	1.06	0.26	1	3

(continued on next page)

Table 3 (*continued*)

Variable	Definition	Mean	SD	Min	Max
Bicycle observed flow	The number of bicycles within the coding segment (number of bicycles > 8: 6, 6 ~ 7: 5, 4 ~ 5: 4, 2 ~ 3: 3, 1:2, 0: 1)	1.07	0.30	1	4
Intersection					
Intersecting road volume	An estimate of the AADT of the intersecting road (no intersecting road: 1, 1 to 100 vehicles: 2, 100 to 1,000 vehicles: 3, 1,000 to 5,000 vehicles: 4, 5,000 to 10,000 vehicles: 5, 10,000 to 15,000 vehicles: 6, ≥15,000 vehicles: 7)	5.56	1.81	1	7
Intersection quality	The quality of the intersection design, advance warnings, signing, and markings. Adequate quality, if necessary, signing and markings are present (poor: 3, adequate: 2, no intersection present: 1)	2.36	0.50	1	3
Intersection channelization	Whether there are raised islands or colored hatching present at an intersection that designate intended vehicle paths (not present: 2, present: 1)	1.07	0.25	1	2
Carriageway	Divided or undivided carriageway	1.97	0.88	1	3
Property access point	The number of commercial and residential driveways and minor access lanes (none: 1, residential access 1 to 2: 2, residential access 3 or more: 4, commercial access one or more: 3)	3.16	1.34	1	4
Endogenous variables					
Head-on crash counts	The number of head-on crashes during 2017 – 2021	1.090	1.810	0	9
Run-off crash counts	The number of run-off crashes during 2017 – 2021	1.009	1.417	0	9
Average vehicle speed	Daily average vehicle speed for a road link during 2017 – 2021	47.6	13.8	11	83
Average vehicle volume	Daily average vehicle volume for a road link during 2017 – 2021	32,600	13,700	1760	77,700

* Coding options for each variable are indicated in parentheses ().

types of endogenous variables, the crash count, the outcome variable of our interest, is the count data type. PSEM, which models each endogenous variable independently, allows us to analyze two local regressions with different variable types effectively.

The crash data, including head-on and run-off crashes, are count type variables. The crash model in the PSEM should utilize either Poisson regression or Negative Binomial Regression. As mentioned in Table 3, the standard deviations are larger than the mean value for both types of crashes, which suggests overdispersion. This justifies using the Negative Binomial Regression for modeling crash counts since it effectively accounts for the observed overdispersion.

For continuous variables such as vehicle speed and volume, linear regression modeling was considered a suitable analytical method. However, it should be noted that the application of linear regression assumes that these continuous variables follow a normal distribution. Preliminary testing for normality, like Q-Q plots, confirmed that the distributions of vehicle speed and volume do not follow a normal distribution. Therefore, log-transformed values of vehicle speed and volume were used for analysis.

Considering the data structures of these distinct regression models, we ensured that appropriate analytical methods were applied to individual local regression models in PSEM, thereby enhancing the robustness and validity of the PSEM model. Notably, the PSEM approach, developed by Shipley (2009), has been actively applied to the field of road safety to account for crash counts as an outcome variable (Ekmekci, 2023; Ekmekci, Dadashzadeh et al., 2024; Ekmekci, Woods et al., 2024).

5. Results

5.1. Principal component analysis (PCA)

After subjecting segments along 68 km of road to a principal component analysis (PCA), four different principal components (PCs) were identified from 78 observed variables from the coded attributes after considering the results of the scree plot shown in Fig. 4. However, the so-called elbow point in the plot, in which the change of percentage of the explained variance, starts to become marginal after PC3. It indicates that three PCs are sufficient for the analysis. We first began with four principal components, including PC4, which can be termed the 'school zone factor.' These four components collectively explained more than 70%, meeting the generally accepted threshold. A common guideline suggests retaining components based on the cumulative percentage of variance explained, typically accounting for 70% or 90% of the total variance (Rea & Rea, 2016). However, we iteratively tested the significance of each variable to ensure the final PSEM model accounted for significant predictors. Also, most segments do not pass the school zone, resulting in many zero values of the school zone variable. Therefore, the coefficients presented in Table 5 reflect the most refined and statistically valid model obtained through this thorough backward elimination process.

The observed variables belonging to each PC are shown in Table 3. Each PC can be named according to the characteristics of the observed variables: Pedestrian and roadside facilities for PC1, Cross-section and intersection conditions for PC2, and Road surface and curve conditions for PC3.

In Table 4, a variable with a higher absolute value of the loading values indicates a stronger association with the principal component. In contrast, a loading value close to zero has little impact on the main component. Positive loading indicates a positive relationship between the variable and the principal component, whereas negative loading suggests a negative relationship between the variable and the principal component.

PC1. Pedestrian and roadside facilities

The first principal component (PC1) primarily relates to pedestrian and roadside facilities. It includes observed variables such as pedestrian crossing facilities, roadside facilities, and pedestrian and bicycle flows. It

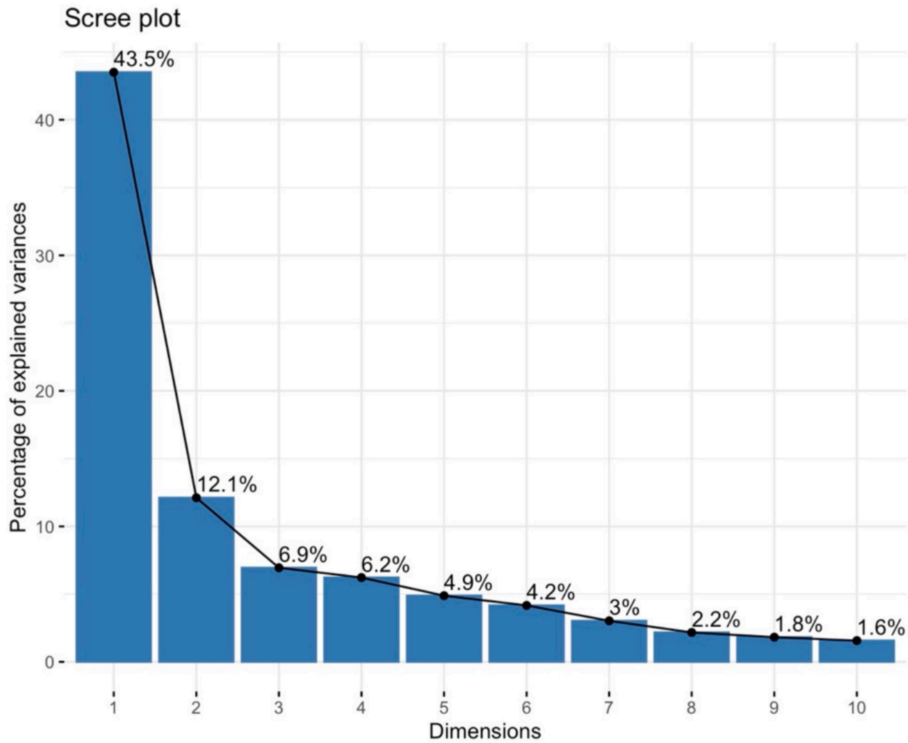

Fig. 4. Scree plot.

is notable that PC1 or variables related to pedestrian and roadside facilities can account for 43.51% of the total variance, with an eigenvalue of 0.94. This suggests that the variables included in PC1 represent considerable amounts of variance in the dataset.

In addition, as mentioned in section 3.1, the coding options for each attribute are listed in order of highest to lowest risk in general. A higher loading of the observed variable indicates a higher risk of the variable contributing to its principal component. If loading shows a negative value, coding values of variables with lower risk contributes to the principal component.

PC1 represents riskier aspects of pedestrian facilities such as sidewalks, pedestrian crossing facilities, street lighting, roadside severity, and observed flows of pedestrians and bicycles. Among these attributes, sidewalks, crossing facilities, and street lighting show positive loading values. However, higher observed flows of pedestrians and bicycles would typically be expected to correlate with safer aspects of PC1. It can be interpreted that more pedestrians and bicyclists on the road may reduce the risk of vehicles hitting pedestrians and cyclists. Roadside severity-passenger side distance also negatively contributes to PC1. It can be interpreted that roads without roadside hazards may increase the risk of pedestrians. These features can be explained by the potential effect of mediating factors like traffic speeds and volume, as indicated in the literature review section (see Table 1). The causal relationships between PCs and crash likelihood will be examined using the PSEM in the following section.

PC2. Cross-section & intersection condition

The second principal component (PC2), explaining 12.11% of the total variance with an eigenvalue of 0.26, is associated with the cross-sectional roadway such as carriageway types, the number of lanes, and the bicycle facilities. It is also associated with intersection conditions, including the intersection quality, channelization, and intersecting road volume.

For PC2 attributes such as intersecting road volume, intersection quality, pedestrian fencing, and land use show positive loading factors, but attributes such as carriageway types, bicycle facilities, channelization, and the number of lanes show negative loading factors. The risk increase of attributes with positive loading factors will increase PC2, but the risk increase of attributes with negative loading factors will decrease PC2. It means good median facilities, separate bicycle lanes, and more lanes can reduce the overall risk of PC2. These counterintuitive features can be explained by the potential effect of mediating factors like traffic speeds and volume. Possibly roads with good median facilities, separate bicycle lanes, and more lanes will increase the speed and volume of vehicles, and it can increase the chance of crashes. The channelization can also increase the speed of turning vehicles at intersections and can increase the chance of crashes.

PC3. Road surface & curve condition

The third principal component (PC3) pertains to road surface and curve conditions, explaining 6.95% of the total variance with an eigenvalue of 0.15. PC3 effectively represents the safer aspects related to curve conditions and riskier aspects of road surfaces.

For PC3 attributes such as paved shoulder, roadworks, delineation, and road surface show positive loading factors, but attributes such as sight distance, curvature, and the quality of curve show negative loading factors. It is easily acceptable that attributes with positive factor loadings will increase the risk of PC3. However, it is counterintuitive to find good sight distance, longer curvature, and good quality of curve will increase the overall risk of PC3.

This does not seem aligned with previous studies in that both safer curve conditions and safer road surfaces should be associated with PC3 with the same directional impacts. Curvature with a sharp radius is known to significantly affect a higher crash frequency or the probability of greater injury severity (Gooch et al., 2016; Q. Hu et al., 2021; Watson et al., 2014; Wu et al., 2013). Roads with paved shoulders and delineation are more likely to perform better in the likelihood of crashes (Stanton & Pinto, 2000). Although we cannot directly conclude whether there is a correlation between PCs and crash likelihood as studied in the previous literature, safer aspects related to curve conditions and riskier aspects of road surface, which can be represented by PC3, may subsequently increase or decrease crash likelihood. This may have to do with the potential of mediating factors like traffic speeds and volume, as indicated in the literature review section (see Table 1). The causal

Table 4
Principal component analysis results with principal components and factor loadings.

Principal component	Observed variables	Loadings	Cronbach's α
PC1: Pedestrian and roadside facilities	Sidewalk – passenger side	0.273	0.73
	Pedestrian crossing quality	0.255	
	Pedestrian crossing facilities – intersecting road	0.241	
	Street lighting	0.193	
	Pedestrian crossing facilities – inspected road	0.190	
	Sidewalk – driver side	0.151	
	Roadside severity – driver side distance	0.094	
	Pedestrian observed flow along the road driver side	−0.091	
	Bicycle observed flow	−0.100	
	Pedestrian observed flow across the road	−0.152	
	Roadside severity – passenger side distance	−0.196	
	Pedestrian observed flow along the road passenger side	−0.200	
	Eigenvalue	0.94	–
	Explained variance (%)	43.51	–
PC2: Cross-section & intersection condition	Intersecting road volume	0.301	0.53
	Intersection quality	0.284	
	Pedestrian fencing	0.138	
	Land use – driver side	0.110	
	Carriageway (undivided versus divided carriageway)	−0.078	
	Facilities for bicycles	−0.123	
	Intersection channelization	−0.209	
	Number of lanes	−0.231	
	Eigenvalue	0.26	–
	Explained variance (%)	12.11	–
PC3: Road surface & curve condition	Paved shoulder – passenger side	0.159	0.46
	Roadworks	0.154	
	Delineation	0.147	
	Road surface	0.104	
	Sight distance	−0.169	
	Curvature	−0.205	
	Quality of curve	−0.207	
	Eigenvalue	0.15	–
	Explained variance (%)	6.95	–

relationships between PCs and crash likelihood will be analyzed in the PSEM in the following section.

Cronbach's alpha was calculated to assess the internal consistency of variables within the dataset. The Cronbach's alpha values are 0.73 for PC1, 0.53 for PC2, and 0.46 for PC3. Although a Cronbach's alpha of at least 0.7 is considered acceptable, we included variables with lower Cronbach's alpha values considering that the present study focuses on exploring the data structure rather than strictly selecting attributes with internal consistency. Removing variables to increase alpha can affect the PCA structure and obscure important theoretical and practical implications of the model.

5.2. Piecewise structural equation modeling (PSEM)

The iRAP protocol classifies crash types into head-on, run-off, and intersection crashes. Only head-on and run-off crashes were considered in the piecewise structural equation model here. The results reveal the direct and indirect effects of each principal component on crash outcomes. The log-transformed average vehicle speeds and volumes are hypothesized as mediators in PSEM modeling. We were able to find the best-fit model by adding or removing variables back and forth. Additionally, we iteratively tested the significance of each variable and only retained those that contributed meaningfully to the model. This iterative process helped ensure that the final model was robust and accounted for all significant predictors. Table 5 displays the PSEM results for head-on and run-off crashes, reflecting the most refined and statistically valid model obtained through this thorough backward elimination process. Fig. 5 presents a path diagram showing the relationships between endogenous and exogenous variables. Solid lines indicate statistically significant relationships at a significance level equal to or exceeding 90%, whereas dotted lines represent relationships that are not statistically significant.

Head-on crashes

Direct effects of PC1, PC2, PC3, and vehicle volumes are statistically significant with regard to increasing the likelihood of head-on crashes. PC1, referring to pedestrian and roadside facilities, is negatively associated with crash counts, with a coefficient of (−)0.091. This result indicates that roadside severity levels and the number of pedestrian crossing facilities negatively contribute to frequency of head-on crashes.

PC2, which explains cross-section and intersection conditions, significantly affects head-on crashes, with a coefficient of 0.080. This suggests that certain attributes, specifically higher intersecting road volumes, the lack of sidewalks, pedestrian fencing contribute to potential conflicts and increased crash likelihood. However, the negative factor loadings associated with the number of lanes, intersection channelization, and bicycle facilities are likely to increase the head-on crash counts. These results are counterintuitive and suggest a need to investigate indirect effects through vehicle speeds and traffic volume.

A higher value of PC3, which is related to road surface and curvature conditions, significantly increases the number of head-on crashes, suggesting that roads with a narrower shoulder width or worse road surface conditions are more likely to cause head-on crashes. Moreover, poor conditions of delineation, skid resistance, and shoulders increase the likelihood of the occurrence of a crash (Bektas et al., 2016; Karkle et al., 2013; Sayed et al., 2010). However, the longer curvature of roads will increase the chance of a head-on crashes. This finding does not correspond to those in previous studies, revealing that the curvature radius is significantly associated with the crash frequency (Wu et al., 2013). This suggests a need to investigate indirect effects through vehicle speeds and traffic volume.

Another direct effect, the traffic volume, is significantly associated with more head-on crashes as well. It is widely confirmed that higher vehicle volumes tend to increase the probability of head-on crashes. Although operation speeds tend to increase crash counts, this factor is not statistically significant.

Run-off crashes

The run-off crash model shows that PC1, explaining pedestrian and roadside facilities, directly decreases the likelihood of the occurrence of a run-off crash, as the value of the PC1 coefficient is (−)0.062. This result indicates that roadside severity levels and fewer pedestrian crossings significantly contribute to frequency of run-off crashes. However, other principal components and variables related to traffic conditions do not directly affect the run-off crash frequency. Considering the previously explained associations between each principle component and mediator, indirect effects through mediators are expected to play a critical role in defining the likelihood of run-off crashes.

Furthermore, all PCs have negative associations with traffic volume. This indicates that road environments, possibly with risky road features, are more likely to have lower traffic volumes. Areas with high traffic volumes typically benefit from better infrastructure and more frequent maintenance, making them more attractive to drivers. In addition, drivers may actively avoid these areas due to perceived or actual safety concerns, resulting in a reduced amount of traffic on these roads.

5.3. Direct, indirect, and total effects by crash type

Table 6 provides the direct, indirect, and total effects of the principal components (PCs) on head-on and run-off crashes. For head-on crashes, PC1 has a direct negative effect of (−)0.091, meaning that a higher PC1

Table 5
PSEM results.

Y		X	Coefficient	Std. Error	Critical ratio	p-value
log (speed)	←	PC.1	0.064	0.004	15.953	<0.01***
log (speed)	←	PC.2	−0.013	0.006	−2.340	0.020**
log (speed)	←	PC.3	−0.023	0.006	−3.577	<0.01***
log (volume)	←	PC.1	−0.054	0.007	−8.257	<0.01***
log (volume)	←	PC.2	−0.118	0.009	−13.037	<0.01***
log (volume)	←	PC.3	−0.028	0.010	−2.723	<0.01***
crash_headon	←	log (speed)	0.095	0.259	0.368	0.713
crash_headon	←	log (volume)	0.528	0.162	3.267	<0.01***
crash_headon	←	PC.1	−0.091	0.031	−2.987	<0.01***
crash_headon	←	PC.2	0.080	0.037	2.182	0.029**
crash_headon	←	PC.3	0.236	0.040	5.932	<0.01***
crash_runoff	←	log (speed)	0.128	0.241	0.530	0.596
crash_runoff	←	log (volume)	0.065	0.147	0.445	0.656
crash_runoff	←	PC.1	−0.062	0.029	−2.129	0.033**
crash_runoff	←	PC.2	0.012	0.036	0.328	0.743
crash_runoff	←	PC.3	−0.015	0.037	−0.420	0.675
Significance level *** <0.01, ** <0.05, *<0.1						
AIC (Akaike Information Criterion) =			3852.92			
Goodness-of-Fit						
• Chi-Squared test = 97.57						
• Fisher's C test = 117.46						
R-squared						
• Speed model = 0.33						
• Volume model = 0.31						
• Head-on crash model = 0.17						
• Run-off crash model = 0.02						

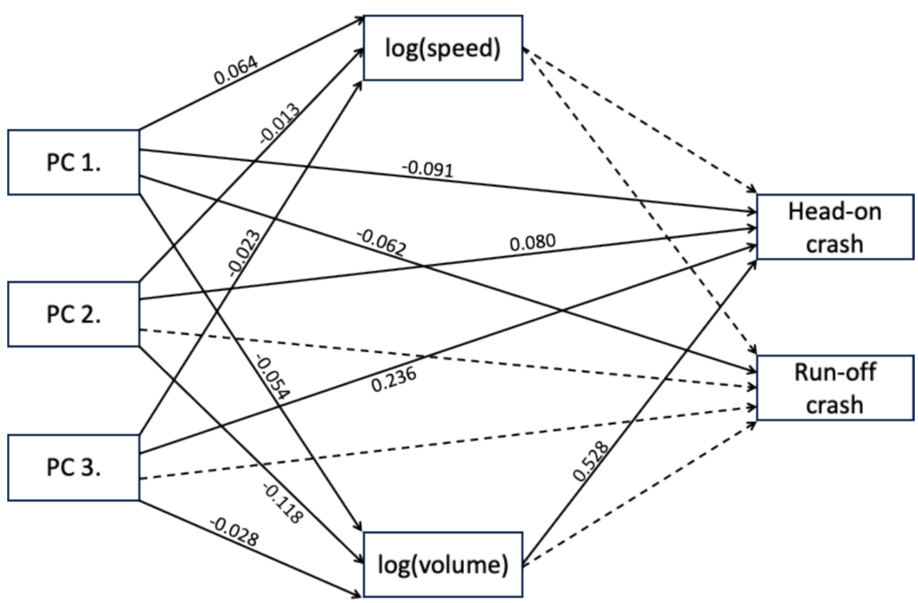

Fig. 5. Path diagram of PSEM.

value is directly linked to a reduction in the frequency of head-on crashes. Vehicle speeds and volumes further contribute indirectly to head-on crashes, resulting in a total effect of (−)0.113. Specifically, the indirect effect of speed is 0.006, while that of volume is (−)0.029. In addition, PC2 has a significant direct effect of 0.080 and showed an indirect negative effect of (−)0.001 through speeds and (−)0.062 through volumes. The total effect of PC2 amounts to 0.016, indicating an increase in head-on crashes. Moreover, PC3 has a substantial direct effect on head-on crashes, indicating a direct effect size of 0.236. It turned out that inadequate road surface and curve conditions can bring about a greater number of crashes. In addition, PC3 showed an indirect negative effect of (−)0.002 through speeds and (−)0.015 through volumes, resulting in total effect of 0.219, which is lower than the direct effect.

Concerning run-off crashes, PC1 has a direct effect of (−)0.062 on the likelihood of run-off crashes. There are indirect effects, such as vehicle speeds, at 0.008, and volume, at (−)0.004. The total effect of PC1 is (−) 0.058, suggesting that roadside facilities or pedestrian crossings are risk factors that increase vehicle speeds and, hence, the frequency of crashes. However, volume has an indirect effect on reducing run-off crashes. The other components (i.e., PC2 and PC3) show no significant direct effects on run-off crashes. However, indirect effects through speed and volume are identified. PC2 has a total impact of (−)0.009 on the outcome, mediated by volume, but a marginal positive effect mediated by speed. In addition, PC3 indirectly influences vehicle speeds and volumes, resulting in a total effect of (−)0.005 and, in turn, leading to a decrease in run-off crashes.

This study addresses gaps in the literature by providing detailed insights into how varying levels of traffic volume and speed influence the

Table 6
Direct, indirect, and total effects of principal components on head-on and run-off crashes.

	Direct effect	Indirect effect (speed)	Indirect effect (volume)	Total effect
Head-on crash				
PC1: Pedestrian and roadside facilities	−0.091	0.006	−0.029	−0.113
PC2: Cross-section & intersection condition	0.080	−0.001	−0.062	0.016
PC3: Road surface & curve condition	0.236	−0.002	−0.015	0.219
Run-off crash				
PC1: Pedestrian and roadside facilities	−0.062	0.008	−0.004	−0.058
PC2: Cross-section & intersection condition	–	−0.002	−0.008	−0.009
PC3: Road surface & curve condition	–	−0.003	−0.002	−0.005

impact of road design attributes on crash likelihood. Our analysis reveals significant mediating effects between road attributes and speeds, as well as between road attributes and volumes, in explaining the likelihood of crashes. Unlike previous studies, which have simply treated traffic volume and speed as control variables, our approach recognizes them as mediating variables. A lack of understanding of these mediating effects can lead to biased results when explaining the effects of road design features on safety outcomes.

6. Discussion

PC1 (pedestrian and roadside facilities) turned out to be the only component that directly affects the likelihood of run-off crashes, according to the piecewise structural equation modeling result. Roads are designed and constructed in accordance with design guidelines and standards. They regulate geometric elements that consist of horizontal and vertical alignment, as well as a cross-sectional profile (Raji, 2017), all of which correspond to PC2 (cross-section & intersection condition) and PC3 (road surface & curve condition). Those attributes belonging to PC2 and PC3 are difficult to alter during road maintenance and operation. Furthermore, PC2 and PC3 did not significantly affect crash occurrence because the sample segments used in this study were from national highway segments that followed similar design standards. As a result, variations of these factors across the samples were marginal, leading to a statistically insignificant effect on the occurrence of crashes.

However, changes in the course of road maintenance and operation primarily focus on roadside facilities and pedestrian facilities, which are easier to manipulate. For instance, roadside objects such as streetlights, traffic signs, and trees are added or removed for user convenience or for comfort when roads are used. Increased traffic volume and speed limits may necessitate additional pedestrian fencing or crossings. Therefore, roadside objects and pedestrian facilities that are added after road construction may have more significant impacts on crash likelihood than geometric attributes fixed at the road design stage.

It is desirable to refine the iRAP methodology for calculating star rating scores, particularly by separating geometric road attributes from changeable attributes such as pedestrian and roadside facilities. The safety effect of geometric design standards, which are constant over time, should be distinguished from that of pedestrian and roadside facilities, which are added or removed over time.

The indirect effect through speeds can be a critical determinant of crashes in South Korea. Higher speeds showed the indirect effects of PC1, increasing head-on and run-off crashes. This suggests that drivers tend to increase their speed when roadside and pedestrian facilities are not adequately equipped. Hence, vehicle speed reduction can be an effective measure to mitigate head-on and run-off crashes, especially in rural areas. However, decisions on speed reduction should consider the context of the road. Reductions in speed limit on rural roads can make sense only if they pass through villages where various road users, including pedestrians and bicyclists, can be mixed with vehicles. Otherwise, lower speed limits on rural roads can be detrimental to the efficient use of roads.

Moreover, to reduce crashes on roads where speeds are higher, the treatment of roadside objects, lighting, and pedestrian crossings are important. Provision of pedestrian crossings and fixed objects along roads encourages drivers to move more cautiously. This finding aligns with previous studies, which observed decreases in severe injuries at lower speeds under similar conditions involving roadside objects (Holdridge et al., 2005). Thus, transportation engineers should take a cautious approach when adding or removing road features during road maintenance and operational activities.

In summary, road safety is influenced by the combined effects of various roadway design attributes. We cannot simply say a single factor contributing to crashes since road design attributes interact and impact safety outcomes that may vary with traffic conditions, such as volumes and speeds. For instance, while critical parameters, such as the direct measurement of road curvature (Gooch et al., 2016; Hu et al., 2021; Watson et al., 2014; Wu et al., 2013) and lane width (Ewing et al., 2024; Hamidi & Ewing, 2023; Gargoum & El-Basyouny, 2016), play significant roles in explaining crash likelihood, PC2 (cross-section) and PC3 (curvature) in this study did not show direct associations with crash likelihood. Instead, they showed indirect associations through speeds and volume, with mediating effects discussed in the literature.

7. Conclusion

This research aimed to quantify the combined effects of road design and traffic conditions on road safety outcomes by crash type—head-on and run-off crashes. The findings provided implications for understanding the complex dynamics of safety regarding road infrastructure. The principal component analysis (PCA) identified three principal components—principal component PC1 represents attributes related to pedestrians and roadside facilities, PC2 represents cross-sections and intersections, and PC3 represents attributes related to road surfaces and curvatures. In addition, piecewise structural equation modeling confirmed that PC1 is the only component directly affecting the number of crashes. The findings suggest that pedestrian and roadside facilities easily added or removed during road operations are more critical than geometric attributes established at the road design stage. The study also confirms the indirect effects of speed and volume on the likelihood of a crash.

Future studies are required to increase the sample size to reach a general conclusion. As this study focused on continuous national highway road segments, our findings may limit transferability to other countries. The same analysis can be conducted by expanding the sample size and study periods to validate the conclusions. While this study mainly focused on the likelihood of crashes, there is a need for further analyses involving injury severity modeling. The effects of principal components on the severity of crashes should differ from the corresponding effects on the likelihood of crashes. It should be noted that we did not specifically study rear-end crashes intentionally as rear-end crashes are more correlated with factors such as occupant positions, driver age, vehicle type, and environmental conditions like rainy or foggy weather (Wang et al., 2022; Yuan et al., 2023) rather than with road infrastructure aspects. We also acknowledge low Cronbach's alpha values could affect the internal consistency of measuring variables.

Funding

This study was funded by the project RS-2023-00243873, Development of the Korea Road Assessment Program and Related Technologies.

CRediT authorship contribution statement

Wookjae Yang: Writing – original draft, Methodology. **Sangjin Han:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdulhafedh, A. (2020). Highway stopping sight distance, decision sight distance, and passing sight distance based on AASHTO models. *OALib, 07*(03), 1–24. https://doi.org/10.4236/oalib.1106095

Abou-Senna, H., Radwan, E., & Mohamed, A. (2022). Investigating the correlation between sidewalks and pedestrian safety. *Accident Analysis and Prevention*, *166* (December 2021), 106548. doi: 10.1016/j.aap.2021.106548.

De Albuquerque, F. D. B., & Awadalla, D. M. M. (2020). Roadside fixed-object collisions, barrier performance, and fatal injuries in single-vehicle, run-off-road crashes. *Safety, 6*(2). https://doi.org/10.3390/safety6020027

Bamzai, R., Lee, Y., Li, Z., By, P., Bamzai, R., Lee, Y., Li, Z. (2017). *Safety Impacts of Highway Shoulder Attributes in Illinois* (Issue June).

Bektas, B. A., Gkritza, K., & Smadi, O. (2016). Pavement marking retroreflectivity and crash frequency: Segmentation, line type, and imputation effects. *Journal of Transportation Engineering, 142*(8), Article 04016030. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000863

Council, F. M., & Stewart, J. R. (2000). Safety effects of the conversion of rural two-lane roadways to four-lane roadways. *ITE Journal*. https://trid.trb.org/view/651972.

Deutsch, E. S., Alameddine, I., & Qian, S. S. (2020). Using structural equation modeling to better understand microcystis biovolume dynamics in a mediterranean hypereutrophic reservoir. *Ecological Modelling, 435*. https://doi.org/10.1016/j.ecolmodel.2020.109282

Dumbaugh, E. (2005). Safe streets, livable streets. *Journal of the American Planning Association, 71*(3), 283–300. https://doi.org/10.1080/01944360508976699

Dumbaugh, E., Saha, D., & Merlin, L. (2024). Toward safe systems: Traffic safety, cognition, and the built environment. *Journal of Planning Education and Research, 44* (1), 75–87. https://doi.org/10.1177/0739456X20931915

Ekmekci, M. (2023). *An investigation of piecewise Structural Equation Modelling to understand the effect on collisions of shorter visibility designs at urban three arm junctions. August*.

Ekmekci, M., Dadashzadeh, N., & Woods, L. (2024). Assessing the impact of low-speed limit zones' policy implications on cyclist safety: Evidence from the UK. *Transport Policy, 152*(April), 29–39. https://doi.org/10.1016/j.tranpol.2024.04.014

Ekmekci, M., Woods, L., & Dadashzadeh, N. (2024). Effects of road width, radii and speeds on collisions at three-arm priority intersections. *Accident Analysis and Prevention, 199*. https://doi.org/10.1016/j.aap.2024.107522

Elvik, R. (1995). The safety value of guardrails and crash cushions: A meta-analysis of evidence from evaluation studies. *Accident Analysis and Prevention, 27*(4), 523–549. https://doi.org/10.1016/0001-4575(95)00003-I

Elvik, R., Christensen, P., & Amundsen, A. (2004). *Speed and road accidents An evaluation of the Power Model.*

Eskandarian, A., Marzougui, D., & Bedewi, N. E. (1997). Finite element model and validation of a surrogate crash test vehicle for impacts with roadside objects. *International Journal of Crashworthiness, 2*(3), 239–258. https://doi.org/10.1533/cras.1997.0048

Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety: A review of empirical evidence. *Journal of Planning Literature, 23*(4), 347–367. https://doi.org/10.1177/0885412209335553

Ewing, R., Promy, N. S., Choi, D., Yang, W., Tabassum, N., Kaniewska, J., Ameli, H., Tuffour, J. P., Kalantari, H. A., Azin, B., & Yang, X. (Terry). (2024). *Transportation Benefits and Costs of Reducing Lane Widths on Urban and Rural Arterials* (Issue January). https://rosap.ntl.bts.gov/view/dot/74632.

Fhwa. (2013). *Signalized Intersections Informational Guide Second Edition*. http://sofety.fhwa.dot.gov.

Garber, N. J., & Gadiraju, R. (1988). *Factors Affecting Speed Variance and Its Influence on Accidents*. 64–71.

Gargoum, S. A., & El-Basyouny, K. (2016). Exploring the association between speed and safety: A path analysis approach. *Accident Analysis and Prevention, 93*, 32–40. https://doi.org/10.1016/j.aap.2016.04.029

George, Y., Athanasios, T., & George, P. (2017). Investigation of road accident severity per vehicle type. *Transportation Research Procedia, 25*, 2076–2083. https://doi.org/10.1016/j.trpro.2017.05.401

Gitelman, V., Doveh, E., & Bekhor, S. (2018). The relationship between travel speeds, infrastructure characteristics, and crashes on two-lane highways. *Journal of Transportation Safety & Security, 10*(6), 545–571. https://doi.org/10.1080/19439962.2017.1324931

Gitelman, V., Doveh, E., Carmel, R., & Hakkert, S. (2019). The influence of shoulder characteristics on the safety level of two-lane roads: A case-study. *Accident Analysis and Prevention, 122*(February 2018), 108–118. doi: 10.1016/j.aap.2018.10.003.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2004). Perceptual lane width, wide perceptual road centre markings and driving speeds. *Ergonomics, 47*(3), 237–256. https://doi.org/10.1080/00140130310001629711

Gooch, J. P., Gayah, V. V., & Donnell, E. T. (2016). Quantifying the safety effects of horizontal curves on two-way, two-lane rural roads. *Accident Analysis and Prevention, 92*, 71–81. https://doi.org/10.1016/j.aap.2016.03.024

Grinstead, J., Ortiz-Ramírez, P., Carreto-Guadarrama, X., Arrieta-Zamudio, A., Pratt, A., Cantú-Sánchez, M., Lefcheck, J., & Melamed, D. (2023). Piecewise structural equation modeling of the quantity implicature in child language. *Language and Speech, 66*(1), 35–67. https://doi.org/10.1177/00238309211066086

Hamidi, S., & Ewing, R. (2023). *A National Investigation on the Impacts of Lane Width on Traffic Safety: Narrowing Travel Lanes as an Opportunity to Promote Biking and Pedestrian Facilities Within the Existing Roadway Infrastructure*. https://narrowlanes.americanhealth.jhu.edu/report/JHU-2023-Narrowing-Travel-Lanes-Report.pdf.

Himes, S., Gross, F., Eccles, K., & Persaud, B. (2016). Multistate safety evaluation of intersection conflict warning systems. *Transportation Research Record, 2583*, 8–16. https://doi.org/10.3141/2583-02

Holdridge, J. M., Shankar, V. N., & Ulfarsson, G. F. (2005). The crash severity impacts of fixed roadside objects. *Journal of Safety Research, 36*(2), 139–147. https://doi.org/10.1016/j.jsr.2004.12.005

Hong, A., & Yang, W. (2023). Urban Environment and Pedestrian Safety: A Case of Seoul. *Korea. Academy of Urban Health, 2*(1), 27–44. https://doi.org/10.59013/auh.2023.2.1.003

Hu, Q., Li, X., Liu, J., & Adanu, E. K. (2021). A low-cost approach to identify hazard curvature for local road networks using open-source data. *Transportation Research Interdisciplinary Perspectives, 10*. https://doi.org/10.1016/j.trip.2021.100393

Hu, W., & Donnell, E. T. (2010). Median barrier crash severity: Some new insights. *Accident Analysis and Prevention, 42*(6), 1697–1704. https://doi.org/10.1016/j.aap.2010.04.009

International Transport Forum. (2024). *Road safety country profiles Korea 2023*. https://www.itf-oecd.org/road-safety-annual-report-2023.

Karkle, D. E., Rys, M. J., & Russell, E. R. (2013). Safety effectiveness of centerline rumble strips in kansas. *Journal of Transportation Safety and Security, 5*(1), 1–26. https://doi.org/10.1080/19439962.2011.642069

Knuiman, M. W., Council, F. M., & Reinfurt, D. W. (1993). Association of median width and highway accident rates. *Transportation Research Record, 1401*, 70–82.

Kraidi, R., & Evdorides, H. (2020). Pedestrian safety models for urban environments with high roadside activities. *Safety Science, 130*(April), Article 104847. https://doi.org/10.1016/j.ssci.2020.104847

Lan, S., Liu, Y., Liu, B., Sheng, P., Wang, T., & Li, X. (2011). Effect of slopes in highway on traffic flow. *International Journal of Modern Physics C, 22*(4), 319–331. https://doi.org/10.1142/S0129183111016270

Manuel, A., El-Basyouny, K., & Islam, M. T. (2014). Investigating the safety effects of road width on urban collector roadways. *Safety Science, 62*, 305–311. https://doi.org/10.1016/j.ssci.2013.09.005

McMahon J, P., Zegeer V, C., Duncan, C., Knoblauch L, R., Stewart R, J., & Khattak J, A. (2002). *An Analysis of Factors Contributing to "Walking Along Roadway" Crashes: Research Study and Guidelines for Sidewalks and Walkways* (Issue February). http://www.walkinginfo.org/pdf/r&d/SidewalkReport.pdf%5Cnhttp://ntl.bts.gov/lib/19000/19900/19995/PB2003102002.pdf.

Merlin, L. A., Guerra, E., & Dumbaugh, E. (2020). Crash risk, crash exposure, and the built environment: A conceptual review. *Accident Analysis and Prevention, 134* (August 2019), 105244. doi: 10.1016/j.aap.2019.07.020.

Milton, J., & Mannering, F. (1998). *The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies*.

NACTO. (2013). Design Street Urban Guide.

Osama, A., & Sayed, T. (2017). Evaluating the impact of connectivity, continuity, and topography of sidewalk network on pedestrian safety. *Accident Analysis and Prevention, 107*(July), 117–125. https://doi.org/10.1016/j.aap.2017.08.001

Ossenbruggen, P. J., Pendharkar, J., & Ivan, J. (2001). Roadway safety in rural and small urbanized areas. *Accident Analysis and Prevention, 33*(4), 485–498. https://doi.org/10.1016/S0001-4575(00)00062-2

Potts, I. B., Harwood, D. W., Cook, D. J., Donnell, E. T., & Hamadeh, B. (2019). Design Guidelines for Horizontal Sightline Offsets. In *Design Guidelines for Horizontal Sightline Offsets* (Issue 2023). Transportation Research Board. doi: 10.17226/25537.

Potts, I. B., Harwood, D. W., & Richard, K. R. (2007). Relationship of lane width to safety on urban and suburban arterials. *Transportation Research Record, 2023*, 63–82. https://doi.org/10.3141/2023-08

Quistberg, D. A., Howard, E. J., Ebel, B. E., Moudon, A. V., Saelens, B. E., Hurvitz, P. M., Curtin, J. E., & Rivara, F. P. (2015). Multilevel models for evaluating the risk of pedestrian-motor vehicle collisions at intersections and mid-blocks. *Accident Analysis and Prevention, 84*, 99–111. https://doi.org/10.1016/j.aap.2015.08.013

Raji, S. A. (2017). Geometric Design of a Highway Using Autocad Civil 3d. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, *4*(6), 7415–7421.

Rea, A., & Rea, W. (2016). *How Many Components should be Retained from a Multivariate Time Series PCA?* 1–49. http://arxiv.org/abs/1610.03588.

Roshandel, S., Zheng, Z., & Washington, S. (2015). Impact of real-time traffic characteristics on freeway crash occurrence: systematic review and meta-analysis. *Accident Analysis and Prevention, 79*, 198–211. https://doi.org/10.1016/j.aap.2015.03.013

Saha, D., Dumbaugh, E., & Merlin, L. A. (2020). A conceptual framework to understand the role of built environment on traffic safety. *Journal of Safety Research, 75*, 41–50. https://doi.org/10.1016/j.jsr.2020.07.004

Sayed, T., Deleur, P., & Pump, J. (2010). Impact of rumble strips on collision reduction on highways in British Columbia, Canada: Comprehensive before-and-after safety study. *Transportation Research Record, 2148*, 9–15. https://doi.org/10.3141/2148-02

Shi, D., Song, H., Liao, X., Terry, R., & Snyder, L. A. (2017). Bayesian SEM for specification search problems in testing factorial invariance. *Multivariate Behavioral Research, 52*(4), 430–444. https://doi.org/10.1080/00273171.2017.1306432

Shipley, B. (2009). Confirmatory path analysis in a generalized multilevel context. *Ecology, 90*(2), 363–368. https://doi.org/10.1890/08-1034.1

Stanton, N. A., & Pinto, M. (2000). Behavioural compensation by drivers of a simulator when using a vision enhancement system. *Ergonomics, 43*(9), 1359–1370. https://doi.org/10.1080/001401300421806

Tabachnick, B.G. , & Fidell, L. S. (2001). Principal Components and Factor Analysis. In: Tabachnick. B.G. and Fidell, L.S. Eds., Using Multivariate Statistics, 4th Ed., Allyn & Bacon., Boston.

Toroyan, T. (2009). Global status report on road safety. *Injury Prevention : Journal of the International Society for Child and Adolescent Injury Prevention*, *15*(4), 286. https://injuryprevention.bmj.com/lookup/doi/10.1136/ip.2009.023697.

Ukkusuri, S., Miranda-Moreno, L. F., Ramadurai, G., & Isa-Tavarez, J. (2012). The role of built environment on pedestrian crash frequency. *Safety Science, 50*(4), 1141–1151. https://doi.org/10.1016/j.ssci.2011.09.012

Wang, C., Chen, F., Yu, B., & Cheng, J. (2022). Injury severity assessment of rear-end crashes via approaches based on generalized estimating equations. *Canadian Journal of Civil Engineering, 49*(10), 1609–1619. https://doi.org/10.1139/cjce-2022-0197

Wang, L., Abdel-Aty, M., Shi, Q., & Park, J. (2015). Real-time crash prediction for expressway weaving segments. *Transportation Research Part C-Emerging Technologies, 61*, 1–10. https://doi.org/10.1016/j.trc.2015.10.008

Wang, X., Wang, X., Cai, B., & Liu, J. (2019). Combined alignment effects on deceleration and acceleration: A driving simulator study. *Transportation Research Part C: Emerging Technologies, 104*(April), 172–183. https://doi.org/10.1016/j.trc.2019.04.027

Watson, D. C., Al-kaisy, A., & Anderson, N. D. (2014). Examining the effect of speed, roadside features, and roadway geometry on crash experience along a rural corridor. *Journal of Modern Transportation, 22*(2), 84–95. https://doi.org/10.1007/s40534-014-0040-4

Wong, A. (2005). *Using GIS to explore the relationship between road grade and accidents*.

World Health Organization. (2023). Global status report on road safety 2023. In *World Health Organization, Geneva, Switzerland* (Vol. 15, Issue 4).

Wu, Y., Nakamura, H., & Asano, M. (2013). A comparative study on crash-influencing factors by facility types on urban expressway. *Journal of Modern Transportation, 21*(4), 224–235. https://doi.org/10.1007/s40534-013-0024-9

Xu, G., Zineddin, A., Atkins, R., & Abel, S. (2022). Speed Management Is Key to Road Safety (Public Roads, Winter 2022). *Federal Highway Administration*. https://highways.dot.gov/public-roads/winter-2022/05.

Yannis, G., Theofilatos, A., Ziakopoulos, A., & Chaziris, A. (2014). Investigation of road accident severity. *Traffic Engineering & Control, 55*(1).

Yuan, R., Gu, X., Peng, Z., & Xiang, Q. (2023). Exploring differences in injury severity between occupant groups involved in fatal rear-end crashes: A correlated random parameter logit model with mean heterogeneity. *Transportation Letters, 00*(00), 1–11. https://doi.org/10.1080/19427867.2023.2292859

Zegeer, C. V., Richard Stewart, J., Huang, H., & Lagerwey, P. (2001). Safety effects of marked versus unmarked crosswalks at uncontrolled locations: Analysis of pedestrian crashes in 30 cities. *Transportation Research Record: Journal of the Transportation Research Board, 1773*(1), 56–68. https://doi.org/10.3141/1773-07

Zirkel, B., Park, S., McFadden, J., Angelastro, M., & McCarthy, L. (2013). Analysis of sight distance, crash rate, and operating speed relationships for low-volume single-lane roundabouts in the united states. *Journal of Transportation Engineering, 139*(6), 565–573. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000532

International Road Assessment Programme. (2021). History. iRAP. URL: https://irap.org/about-us/history/.

Wookjae Yang, Ph.D., is an Assistant Professor in the Department of Urban Policy and Administration at Incheon National University. He earned his Ph.D. in Urban Planning from the University of Utah. His research focuses on smart mobility and its effects on transportation systems and land use planning.

Sangjin Han, Ph.D., is an Associate Professor in the Graduate School of Environmental Studies at Seoul National University. He holds a Ph.D. in Transport Studies from University College London. Dr. Han currently leads a nationally-funded R&D project focused on developing the Korea Road Assessment Program (KoRAP) and applying AI technologies to enhance road infrastructure safety.