ARCADE

Finanziato dall'Unione europea
NextGenerationEU

Accident Risk reduCtion of vulnerAble RoaD usErs: an interdisciplinary – multiperspective approach (ARCADE)

Deliverable 4

WP4 – Analysis of Results, Discussion, Practical Implications

---

Version:	1
Date	24 February 2026
Drafted:	Vennarucci A. (UniRM3); Tefa L. (PoliTO); Bruno G., Marson, A. (UniPD)
Validate	Francesco Bella (Principal Investigator – UniRM3) Marco Bassani (Resp. research unit PoliTO) Andrea Spoto (Resp. research unit UniPD)
Authors/members RU	RU-UniRM3: Bella F., Calvi A., D'Amico F., Gagliardi V., Vennarucci A., Manalo J.R.D. RU-PoliTO: Bassani M., Tefa L., Lioi A., Hassani Barbin A. RU-UniPD: Bruno, G., Marson, A., Spoto, A.
Keywords:	Road Safety, Vulnerable Road Users, Human Factors, Surrogate Safety Measures, Multiperspective, Biopsychology

WP4 – Analysis of Results, Discussion, Practical Implications

Contents

1 Introduction

This deliverable aims to summarize the results obtained from the three Research Units (RUs) regarding the activities carried out within WP 4 (Analysis of Results, Discussion, Practical Implications) and attributable to the following tasks:

- T4.1 Analysis of the integrated/combined data from observational and simulator studies;
- T4.2 Calibration and updating of behavioural models in driver-VRU interaction as affected by road facility/countermeasure configurations and human factors;
- T4.3 Countermeasure assessment. Effectiveness of safety actions in reducing risks at conflicting points in the urban road system;
- T4.4 Generalization of research outcomes.

For the reasons already indicated in the Deliverable 3 (specificities of the simulation systems used by 3 RUs and peculiar driver-VRU interactions analyzed by 3 RUs), the document reports, for each Research Unit, the results of the experiments conducted using the different simulation systems.

The final section (Final remarks) summarizes the methodological framework adopted within the ARCADE Project and highlights its potential for public administrations and road authorities for the road safety management of vulnerable road users.

2 Results RU UniRM3

The following section reports the main results obtained from the experimental campaigns carried out at the Simulation lab of the RU UNIRM3. The results are shown for the specific simulation system used:

- driving simulator
- pedestrian simulator
- motorcyclist simulator

1.1 Driving Simulator Study results

2.1.1 Vehicle/Vehicle and Vehicle/Motorcycle Interaction

2.1.1.1 Outliers

To ensure dataset reliability and analytical validity, rigorous criteria were established for identifying and excluding outliers, structured on two levels: qualitative (consistency and correctness of experimental trials) and quantitative (removal of anomalous values potentially affecting representativeness). Cases were excluded when characterized by Waiting Time (WT), defined as time interval during which the user's remains stationary near the stop line while waiting for acceptable traffic conditions to complete the turning manoeuvre, equal to 0 s (failure to stop at the STOP sign), interruption due to simulation sickness, crash during the simulation, or lack of interaction with the platoon vehicle (either failure to accept the available gap or premature crossing before the arrival of the first platoon vehicle).

Exclusion was applied both at the participant level (interruption due to simulation sickness, high final sickness declaration, or more than three total crashes) and at the single-maneuver level (WT = 0 s, crash within the scenario, or lack of interaction).

The initial sample consisted of 58 participants; 8 interrupted the experiment due to simulation sickness (13.8%), leading to the complete exclusion of their data. Consequently, 48 valid datasets were available for daytime scenarios and 50 for nighttime scenarios (two participants completed only nighttime drives). Age-stratified analysis revealed a marked difference: among participants over 45 years old (13 subjects), 6 interruptions occurred (46.2%), whereas among those under 45 (45 subjects), only 2 interruptions were recorded (4.4%), indicating a substantially greater impact on older participants.

In daytime scenarios, exclusion rates ranged from 17% to 40%, as reported in the following Table 1. The main causes were WT = 0 s (particularly frequent in S1_5 and S1_6), crashes (more frequent in S2_5), participants with more than three total crashes, sporadic cases of gap rejection, and platoon anticipation events, mainly concentrated in S1_6, which represents the most critical condition in terms of anomalous behaviour. In absolute terms, the highest number of outliers was observed in S1_6 (40%) and S1_5 (38%), while S0_5 showed the lowest percentage (17%). The final number of valid trials ranged from 29 (S1_6) to 40 (S0_5).

In nighttime scenarios, exclusion rates were generally higher, ranging from 25% to 46%. The main causes were WT = 0 s (with a peak in N0_6), crashes (particularly frequent in N2_5), participants with more than three total crashes, some cases of gap rejection, and platoon anticipation events, especially in N0_5 and N1_6. The scenario with the highest outlier rate was N0_6 (46%), followed by N2_5 (44%), whereas N0_5 and N1_6 showed the lowest rates (25%). The number of remaining valid trials ranged from 28 (N0_6) to 38 (N0_5 and N1_6).

Table 1 - Scenario Outliers Resume

Daytime Outliers								
Scenario	**WT = 0 s**	**Crash**	**> 3 Crashes**	**Gap Not Accepted**	**Platoon Anticipation**	**Total Outliers**	**% Outliers**	**Valid Trials**
S0_5	3	1	3	2	1	8	17%	40
S0_6	9	0	3	0	0	11	23%	37
S1_5	10	5	3	1	0	18	38%	30
S1_6	11	2	3	0	10	19	40%	29
S2_5	4	10	3	1	0	14	29%	34
S2_6	9	1	3	1	0	11	23%	37
Nighttime Outliers								
Scenario	**WT = 0 s**	**Crash**	**> 3 Crashes**	**Gap Not Accepted**	**Platoon Anticipation**	**Total Outliers**	**% Outliers**	**Valid Trials**
N0_5	6	2	3	1	4	12	25%	38
N0_6	19	2	2	0	0	22	46%	28
N1_5	6	7	3	2	0	16	33%	34
N1_6	6	3	3	0	5	12	25%	38
N2_5	2	15	3	4	0	21	44%	29
N2_6	10	5	3	2	0	17	35%	33

2.1.1.2 Correlation Analysis (PET, WT, GAR, PSD and Vm)

Prior to conducting the comparative statistical analysis of the indicators, two distinct correlation analyses were performed for each daytime and nighttime experimental configuration. The first analysis aimed to examine the relationships among the safety indicators Post Encroachment Time (PET), defined as time interval between the passage of the user's vehicle and the passage of the platoon vehicle at the same intersection point, WT, Residual Accepted Gap (GAR), is a measure used to evaluate the temporal margin

available between the user's vehicle and the platoon vehicle at the moment to decision to merge is made during the left-turn manoeuvre, Proportion of Stopping Distance (PSD), defined as ratio between the remaining distance between a vehicle and the potential collision point, and the minimum acceptable distance required to stop under safe conditions, and Average Approach Speed (Vm), calculated as the arithmetic mean of the instantaneous speeds recorded between the initial point ad the stopping point, in order to identify potential associations useful for interpreting the results. The second analysis was designed to assess the internal consistency among the different speed indicators during the approach phase to the intersection Vm, and Vi, Vint, Vf, respectively, the instantaneous speed measured 50 m, 30 m and 10 m before the stop line.

The Shapiro–Wilk test revealed violations of the normality assumption in several scenarios, particularly for WT and, in some cases, for GAR or Vm. To ensure methodological consistency across all configurations, Spearman's rank correlation coefficient was therefore adopted. This nonparametric test does not require normality and evaluates associations based on ranked data.

The results consistently highlight the presence of a central cluster composed of PET, PSD, and GAR. Across both daytime and nighttime scenarios, PET exhibits strong to very strong positive correlations with PSD and, in most cases, also with GAR. Similarly, PSD and GAR are significantly and positively correlated with each other. This group of indicators appears to coherently capture the temporal dimension of safety in vehicle–vehicle interactions, demonstrating a high degree of interdependence among the measures.

The overall analysis further reveals an extremely strong and consistent association between PET and PSD, with correlation coefficients frequently approaching or exceeding 0.8. This finding indicates substantial informational redundancy between the two indicators, as they describe the same risk dynamics in a highly overlapping manner. To avoid duplication in the subsequent comparative statistical analyses, PSD was therefore excluded from further processing.

Although GAR shows significant correlations with both PET and PSD, the strength of these relationships is less consistent across scenarios, ranging from moderate to high. This variability justifies the retention of GAR in subsequent analyses, as it provides partially independent information.

WT occupies a partially divergent position relative to the central cluster, frequently showing negative correlations with GAR and, in some scenarios, with PET or Vm. This suggests that longer waiting times prior to merging are associated with different temporal dynamics or reduced residual safety margins.

The indicator Vm demonstrates a less systematic pattern. In most cases, it does not show significant correlations with the other parameters; however, in some configurations, particularly under nighttime conditions, weak to moderate correlations emerge, occasionally negative, with GAR, WT, or PSD. This suggests that under reduced visibility conditions, mean approach speed may interact with risk dynamics, although not in a structurally consistent manner.

In summary, the two correlation analyses revealed a stable and interconnected core of indicators (PET, PSD, and GAR), the substantial redundancy between PET and PSD, and the complementary yet less systematic role of WT and Vm. Based on these findings, subsequent in-depth statistical analyses were conducted on PET, GAR, WT, and Vm, while PSD was excluded to prevent informational overlap.

2.1.1.3 Correlation Analysis (Vm, Vi, Vint, Vf)

Correlation analyses among the speed indicators (Vm, Vi, Vint, Vf) were conducted separately for each daytime and nighttime configuration. Normality was preliminarily assessed using the Shapiro–Wilk test. Depending on the results, Pearson's correlation coefficient was applied when the assumption of normality was satisfied, whereas Spearman's rank correlation coefficient was adopted in cases of non-normal distributions, ensuring methodological consistency within each scenario.

In daytime scenarios, a highly stable relational structure emerged. All speed variables were positively and significantly correlated, with effect sizes ranging from moderate to very high. The strongest and most consistent association was observed between Vi and Vint, confirming this pair as the core of the crossing dynamics. The relationship between Vint and Vf was also systematically significant, highlighting the dynamic continuity between intermediate and final speeds. Vm showed consistent correlations with the other speed indicators, generally of moderate magnitude, indicating that none of the variables operates independently within the daytime context.

In nighttime scenarios, the overall structure remained coherent, although characterized by greater variability. Positive and significant correlations were still observed in most cases, with Vi and Vint again exhibiting the strongest and most stable association. The relationships between Vm and Vf were frequently significant, while correlations involving Vf showed less stability compared to daytime conditions. Overall, correlation magnitudes ranged from weak to very high, with more pronounced fluctuations under reduced visibility, suggesting increased behavioral variability in nighttime driving conditions.

The comparative analysis between daytime and nighttime contexts therefore indicates substantial consistency in the correlation structure among speed indicators. The Vi–Vint pair consistently represents the most robust and stable relationship across both conditions, whereas other associations, although significant, display greater variability at night. No correlation systematically exceeded the threshold indicative of complete informational redundancy. Consequently, all speed indicators (Vm, Vi, Vint, Vf) were retained for subsequent statistical analyses, as each provides a complementary and non-redundant contribution to the characterization of vehicle dynamic behavior during intersection approach and crossing.

2.1.1.4 Statistical Analysis

2.1.1.4.1 PET

Post-Encroachment Time (PET) is a surrogate safety measure used to describe vehicle–vehicle conflicts. It represents the time interval between the moment one vehicle exits the conflict area, reported in the following Figure 1, and the moment the subsequent vehicle enters it, thereby expressing the temporal safety margin between two road users. Higher PET values correspond to safer maneuvers, as they indicate greater temporal separation between vehicles.

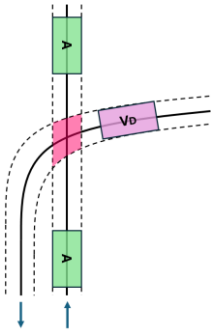


Figure 1 - Representation of the conflict area calculated with the Post Encroachment Time

As shown in the following Figure 2, a clear separation emerges between scenarios characterized by a 5 s gap and those with a 6 s gap. In both daytime and nighttime conditions, boxplots associated with the 6 s gap consistently display higher median values than their 5 s counterparts, indicating larger average temporal safety margins. This pattern is coherent with the increased time available for maneuver execution. Overall data dispersion appears limited and relatively homogeneous across scenarios, with slightly greater variability observed in some nighttime cases, particularly N1_6 and N2_6.

When comparing daytime and nighttime conditions, no marked visual differences in median values are evident. The distributions are substantially overlapping within the same gap configurations, confirming the stability of PET with respect to lighting conditions.

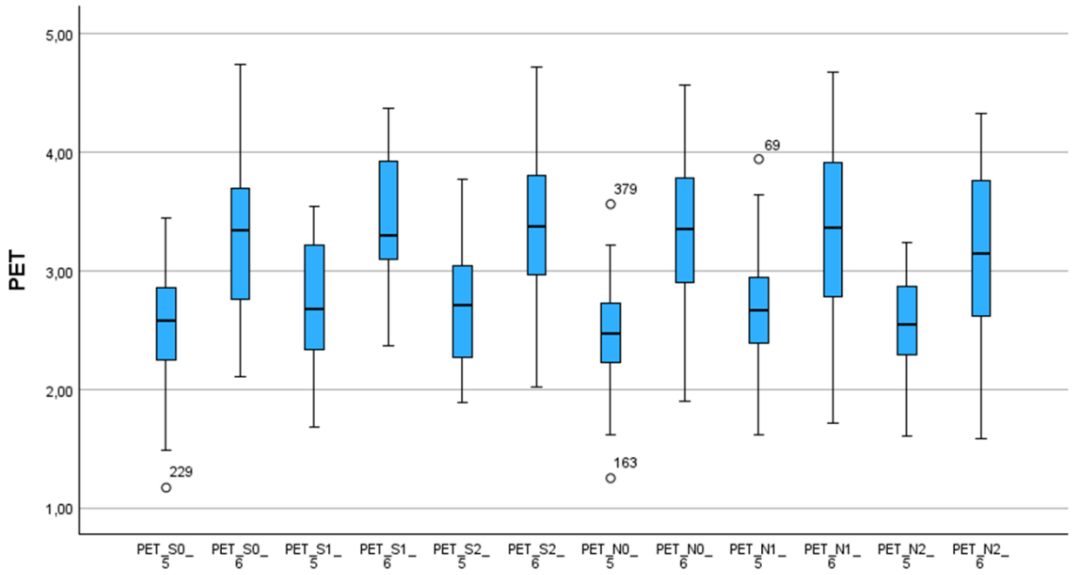


Figure 2 - Box Plot chart of Post Encroachment Time

Normality of PET was initially assessed using the Shapiro–Wilk test across all daytime and nighttime scenarios. The results consistently showed significance values greater than 0.05, supporting the assumption of normality. Accordingly, comparisons among scenarios were conducted using one-way ANOVA, preceded by Levene's test for homogeneity of variances and, where necessary, followed by post-hoc analyses with Bonferroni or Tamhane corrections.

As reported in the following Table 2, in the three daytime scenarios with a 5 s gap (S0_5, S1_5, S2_5), mean PET values ranged between 2.56 s and 2.75 s, with relatively small standard deviations (approximately 0.5 s), as visible in the following Figure 4a. The ANOVA did not reveal statistically significant differences among scenarios, and the effect of scenario on overall variability was negligible. A similar pattern was observed in the corresponding scenarios with a 6 s gap (S0_6, S1_6, S2_6), where mean values ranged between 3.31 s and 3.44 s, again with no significant differences detected. In nighttime conditions, the overall pattern was comparable, as reported in the following Figure 3b, with the exception of the N0_5–N1_5–N2_5 group, where ANOVA indicated a significant difference between N0_5 and N1_5, with higher mean PET values in N1_5 (2.73 s versus 2.45 s). However, the effect size was small to moderate ($Eta^2 \approx 0.067$). In nighttime scenarios with a 6 s gap (N0_6, N1_6, N2_6), no significant differences were found, and the impact of the scenario on PET variability was minimal.

Direct comparisons between daytime and nighttime conditions, performed using independent-samples t-tests for each corresponding pair of scenarios, did not reveal statistically significant differences in any case, as visible in the following Table 3. Observed mean differences were consistently small (generally below 0.3 s), and effect sizes (Cohen's d = 0.23; Hedges' correction = 0.23; Glass's delta = 0.24) fell within the range of very small or negligible effects.

Overall, PET demonstrates substantial stability across scenarios and lighting conditions. Although slight tendencies toward larger temporal safety margins in project scenarios compared to the baseline can be observed, such differences rarely reach statistical significance.

Table 2 - Statistical Analysis of Post Encroachment Time

Indicatore	Condizioni		S0 / N0		S1 / N1		S2 / N2		Confronto fra Scenari		Confronto a Coppie		
	D - N	Gap (s)	Media	D. St.	Media	D. St.	Media	D. St.	Significatività	Misura d'Effetto R: Sc0|Sc1|Sc2|	Sc 0 vs Sc 1	Sc 0 vs Sc 2	Sc 1 vs Sc 2
PET (s)	Diurno	5	2,56	0,49	2,75	0,52	2,69	0,52	n.s.	–	n.s. (S1 > S0)	n.s. (S2 > S0)	n.s. (S1 > S2)
	Diurno	6	3,31	0,58	3,44	0,52	3,42	0,62	n.s.	–	n.s. (S1 > S0)	n.s. (S2 > S0)	n.s. (S1 > S2)
	Notturno	5	2,45	0,45	2,73	0,48	2,56	0,42	ANOVA p=0,034	η^2=0,067	p=0,028 (N1 > N0)	n.s. (N2 > N0)	n.s. (N1 > N2)
	Notturno	6	3,34	0,60	3,31	0,76	3,16	0,75	n.s.	–	n.s. (N0 > N1)	n.s. (N0 > N2)	n.s. (N1 > N2)

Table 3 - T-test Analysis of Post Encroachment Time

Indicatore	Gap (s)	S0 vs N0		S1 vs N1		S2 vs N2	
		Significatività	Misura d'Effetto R: S0|N0	Significatività	Misura d'Effetto R: S1|N1	Significatività	Misura d'Effetto R: S2|N2
PET (s)	5	n.s.	–	n.s.	–	n.s.	–
	6	n.s.	–	n.s.	–	n.s.	–

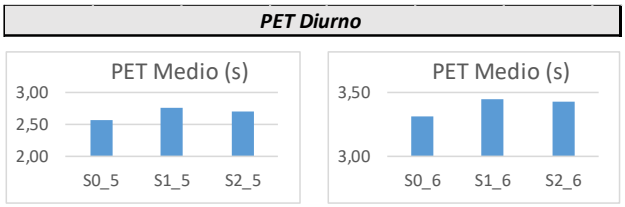


Figure 3a - Average value of daytime Post Encroachment Time

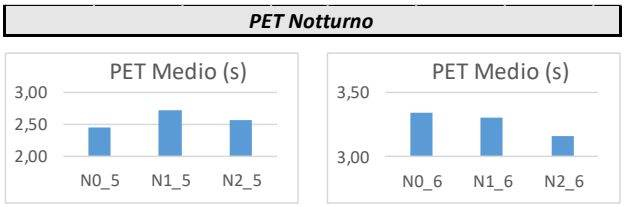


Figure 3b - Average value of nighttime Post Encroachment Time

2.1.1.4.2 GAP

As shown in the following Figure 4, in all scenarios the median generally lies between 1 and 2 gaps, indicating that most drivers tend to accept one of the first available opportunities for crossing. The distribution is strongly concentrated at low values (1–2), with relatively narrow interquartile ranges, confirming a fairly stable decision-making pattern. Variability increases slightly in scenarios with a 5-second gap, where wider ranges and a greater presence of extreme values are observed. In particular, in contexts S2_5 and N2_5, the upper whiskers extend further, suggesting greater dispersion in gap selection.

When comparing daytime and nighttime conditions, no marked visual differences emerge in either medians or central dispersion. The distributions are largely overlapping within the same gap configurations, with slightly higher variability observed in some nighttime scenarios. Overall, gap selection behavior appears consistent and only minimally influenced by lighting conditions.

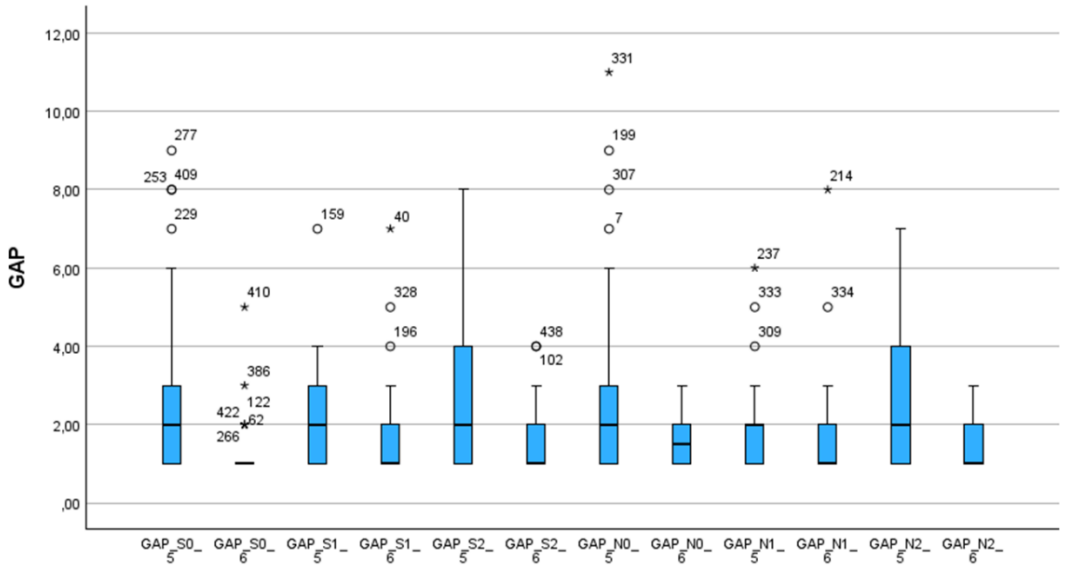


Figure 4 - Box Plot chart of GAP

The Number of Gap Chosen did not meet the normality assumption across all scenarios; therefore, comparisons among scenarios were conducted using non-parametric tests (Kruskal–Wallis for multiple comparisons and Mann–Whitney for pairwise day/night comparisons).

In the comparisons among daytime scenarios with 5 and 6 second gaps, the Kruskal–Wallis test did not reveal statistically significant differences. Mean values generally ranged between 1 and 3 gaps, with medians consistently equal to 1 or 2, indicating that most drivers tended to accept the earliest available opportunities.

A similar pattern emerged in nighttime scenarios, where no significant differences were observed among experimental configurations. The only exception concerned the comparison between S0_6 and N0_6, in which the Mann–Whitney test showed a significant difference ($p = 0.010$) with a moderate effect size ($r \approx -0.32$). In this specific scenario, nighttime conditions appear to have influenced gap selection more substantially.

Distribution plots clearly show that, both during the day and at night, the majority of drivers selected the first available gap, followed by the second and third with progressively decreasing frequency. Selections beyond the third gap were marginal. This pattern appears stable across experimental conditions and suggests a decision-making strategy oriented toward early acceptance of crossing opportunities.

2.1.1.4.3 Crossed GAP and Interfered Vehicle

The statistical analysis of the Interfered Vehicle, defined as the type of crossed gap, Car–Car, Car–Motorcycle, Motorcycle–Car, as reported in the following Figure 5, was conducted through frequency distributions, graphical representations, and Pearson's chi-square tests, complemented by likelihood ratio statistics and measures of association (Phi and Cramer's V). The aim was to assess both the presence of non-

uniform distributions within individual scenarios and differences across infrastructural configurations and lighting conditions.

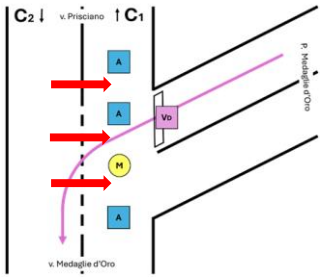


Figure 5 - Representation of the Crossed GAP and Interfered Vehicle

Overall, the type of interfered vehicle does not appear to be uniformly distributed, but rather is influenced primarily by the temporal gap threshold and, to a variable extent, by the infrastructural design. In contexts characterized by a reduced gap (5 s), unbalanced distributions and statistically significant differences among categories are more frequently observed, as reported in the following Table 4. In the baseline daytime scenario (S0_5), a clear predominance of Car–Motorcycle (C–M) interactions emerges, significantly exceeding the other categories. A similar pattern, though with varying intensity, is observed in S1_5, where Motorcycle–Car interactions are markedly reduced and most cases concentrate in the C–C and C–M categories. In S2_5, a predominance of C–M interactions persists, although without reaching full statistical significance.

By contrast, in scenarios with a 6 s gap, distributions tend to be more balanced and rarely differ significantly from uniformity. In S0_6, S1_6, and S2_6, the three interference types are generally more evenly distributed, with no clearly dominant category. This suggests that increasing the available temporal margin reduces the concentration of interactions within a specific vehicle configuration, leading to a more homogeneous dynamic.

A similar pattern is observed in nighttime conditions. With a 5 s gap, scenario N0_5 shows a significant predominance of Car–Motorcycle interactions, whereas in N1_5 and N2_5 the distribution appears more balanced and does not significantly deviate from expected proportions. With a 6 s gap (N0_6, N1_6, N2_6), interference types are generally evenly distributed across categories, with no marked predominance. Comparisons among nighttime scenarios (0–1–2) reveal statistically significant differences in some cases, indicating that infrastructural configuration may influence the type of interfered vehicle, although this effect is not systematic across all conditions.

Direct comparisons between daytime and nighttime conditions yield differentiated results, as reported in the following Table 5. In some cases, such as scenario 0 with a 5 s gap, distributions are substantially homogeneous across lighting conditions. In other cases, such as scenario 0 with a 6 s gap, significant differences emerge: during the day, Car–Motorcycle interactions prevail, whereas at night the distribution becomes more balanced, with a relative increase in Motorcycle–Car interactions. However, in most day/night

comparisons, differences are not statistically significant, suggesting that lighting conditions exert a selective rather than generalized influence.

An additional relevant finding concerns the relationship between the number of gaps chosen and the type of interference. In 5 s gap contexts, a linear trend emerges whereby, as the selected gap increases, Car–Car interactions decrease and interactions involving vulnerable vehicles (C–M and M–C) increase. Although this pattern does not always reach statistical significance due to limited cell frequencies, it suggests that later crossing decisions are more frequently associated with motorcycle-related interactions. In 6 s gap contexts, by contrast, most crossings occur at the first available gap, and distributions remain more stable, with less pronounced differences among categories.

Overall, the analysis indicates that the type of interfered vehicle is not random but is shaped by the interaction between infrastructural configuration and temporal gap threshold, as reported in the following Figure 6. The effect is more pronounced under more critical conditions, namely reduced gaps, where Car–Motorcycle interactions more frequently predominate. As the temporal gap increases, distributions tend to rebalance and differences among scenarios diminish. Variations between daytime and nighttime conditions are present but not systematic, indicating that lighting conditions have a weaker influence than temporal gap pressure. Collectively, these findings suggest that interference dynamics are more sensitive to the level of temporal pressure imposed by the gap than to lighting conditions alone.

Table 4 - Statistical Analysis of Crossed GAP

Indicatore	Condizioni		S0 / N0			S1 / N1			S2 / N2			Confronto fra Scenari	
	D - N	Gap (s)	A-A	A-M	M-A	A-A	A-M	M-A	A-A	A-M	M-A	Significatività	Misura d'Effetto
Varco Accettato (A-A, A-M, M-A)	Diurno	5	7	23	10	13	14	3	6	17	11	χ^2 p=0,003	V=0,324
	Diurno	6	14	16	7	13	7	9	9	14	14	χ^2 p=0,021	V=0,309
	Notturno	5	8	21	9	11	12	11	8	12	9	χ^2 p=0,003	V=0,324
	Notturno	6	12	6	10	17	11	10	6	17	10	χ^2 p=0,007	V=0,308

Table 5 – T-test Analysis of Crossed GAP

Indicatore	Gap (s)	S0 vs N0		S1 vs N1		S2 vs N2	
		Significatività	Misura d'Effetto R: S0|N0	Significatività	Misura d'Effetto R: S1|N1	Significatività	Misura d'Effetto R: S2|N2
Varco Accettato	5	n.s.	–	n.s.	–	n.s.	–
	6	χ^2 p=0,014	V=0,359	n.s.	–	n.s.	–

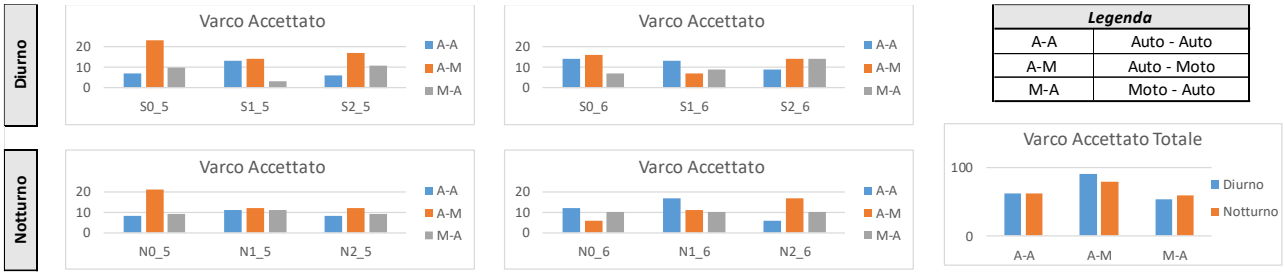


Figure 6 - Average value of Crossed GAP

2.1.1.4.4 Mean Approach Speed

The statistical analysis of the mean approach speed to the intersection (Vm) was conducted to compare the different experimental scenarios under both daytime and nighttime conditions, integrating descriptive statistics, assumption checks, and either parametric or non-parametric inferential tests depending on data distribution, as reported in the following Table 6.

Preliminary normality assessment using the Shapiro–Wilk test indicated that Vm was normally distributed in most scenarios, with specific exceptions (notably S2_5, S2_6, and N0_6), for which non parametric procedures were adopted in group comparisons. This differentiated approach ensured methodological rigor and prevented violations of statistical assumptions.

From a descriptive perspective, a consistent pattern emerged across both gap levels (5 s and 6 s): the design scenarios, and especially Scenario 2 (S2 and N2), exhibited lower mean approach speeds compared with the baseline configuration (S0 and N0). Under daytime conditions, ANOVA revealed a statistically significant main effect of scenario, with marked differences between S0 and S2 and between S1 and S2, whereas differences between S0 and S1 were generally not statistically significant. A similar trend was observed in nighttime conditions, where N2 consistently showed the lowest Vm values, N0 the highest, and N1 an intermediate position.

In particular, for the nighttime 6 s gap condition, statistically significant differences were found among N0_6, N1_6, and N2_6, with a clear descending order of mean speeds (N0 > N1 > N2), also confirmed by mean rank comparisons in the corresponding non parametric tests. This pattern highlights a progressive reduction in approach speed when moving from the baseline scenario to the design configurations, with the most pronounced decrease observed in Scenario 2.

From an interpretative standpoint, these findings indicate that the infrastructural countermeasures introduced in the design scenarios effectively reduced drivers' approach speeds. In particular, the presence of the raised crossing preceding the STOP line in Scenario 2 was associated with an earlier and more uniform

deceleration, suggesting a more cautious and controlled driving behavior. Scenario 1 produced a less marked and not always statistically significant reduction, indicating that while its geometric modifications contributed to speed moderation, a statistically robust effect was achieved primarily when combined with the additional measures implemented in Scenario 2.

Overall, the mean approach speed appears to be a sensitive indicator of infrastructural modifications, showing a clear tendency toward reduction in the project scenarios compared with the existing configuration. The effect is consistent under both daytime and nighttime conditions, whereas the lighting condition itself appears to exert a more limited influence than the geometric configuration and the implemented countermeasures.

Table 6 - Statistical Analysis of mean Approach Speed

Indicatore	Condizioni		S0 / N0		S1 / N1		S2 / N2		Confronto fra Scenari		Confronto a Coppie		
	D - N	Gap (s)	Media	D. St.	Media	D. St.	Media	D. St.	Significatività	Misura d'Effetto R: Sc0|Sc1|Sc2|	Sc 0 vs Sc 1	Sc 0 vs Sc 2	Sc 1 vs Sc 2
Velocità Media in Approccio Vm (m/s)	Diurno	5	6,20	1,66	5,91	1,52	4,64	1,44	K-W p<0,001	R: 64|59|34	n.s. (S0 > S1)	p<0,001 (S0 > S2)	p<0,001 (S1 > S2)
	Diurno	6	6,44	1,56	5,80	1,62	4,77	1,24	K-W p<0,001	R: 67|54|36	n.s. (S0 > S1)	p<0,001 (S0 > S2)	p=0,038 (S1 > S2)
	Notturno	5	5,74	1,54	5,67	1,45	4,55	1,05	ANOVA p=0,001	η^2=0,129	n.s. (N0 > N1)	p=0,002 (N0 > N2)	p=0,005 (N1 > N2)
	Notturno	6	6,43	1,65	5,47	1,34	4,69	1,12	K-W p<0,001	R: 67|51|35	p=0,022 (N0 > N1)	p<0,001 (N0 > N2)	n.s. (N1 > N2)

The comparison between paired daytime and nighttime scenarios (S0 vs N0, S1 vs N1, S2 vs N2) for mean approach speed (Vm) did not reveal statistically significant differences at either gap level (5 s or 6 s), as reported in the following Table 7. In all cases, the comparisons were classified as non-significant (n.s.), indicating that lighting conditions did not produce measurable changes in mean approach speed within the same infrastructural configuration. For the 5 s gap condition, no significant differences emerged for S0 vs N0 or S1 vs N1, and similarly no significant variation was observed for S2 vs N2. Effect size estimates, where reported, were small (e.g., R = 0.33 vs 0.31 for S2 vs N2), suggesting minimal practical differences between daytime and nighttime conditions. The same pattern was confirmed for the 6 s gap condition. Comparisons between S0 and N0, S1 and N1, and S2 and N2 all resulted in non-significant outcomes. Reported effect size values (e.g., R = 0.33 vs 0.33 for S0 vs N0 and R = 0.36 vs 0.36 for S2 vs N2) indicate negligible differences and substantial overlap between distributions.

Overall, these results confirm that mean approach speed remains stable across lighting conditions when infrastructural configuration and gap duration are held constant, as visible in the following Figure 7a and Figure 7b. The absence of significant differences suggests that geometric design exerts a stronger influence on speed behavior than the day versus night condition.

Table 7 – T-test Analysis of mean Approach Speed

Indicatore	Gap (s)	S0 vs N0		S1 vs N1		S2 vs N2	
		Significatività	Misura d'Effetto R: S0|N0	Significatività	Misura d'Effetto R: S1|N1	Significatività	Misura d'Effetto R: S2|N2
Vm (m/s)	5	n.s.	–	n.s.	–	n.s.	R: 33|31
	6	n.s.	R: 33|33	n.s.	–	n.s.	R: 36|36

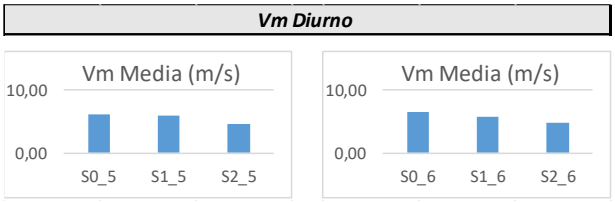


Figure 7a - Average value of daytime mean Approach Speed

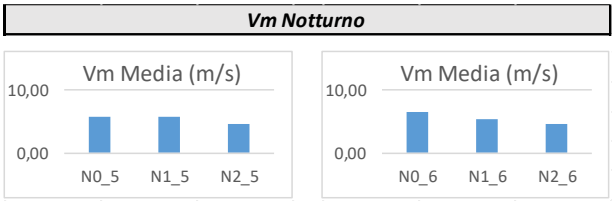


Figure 7b - Average value of nighttime mean Approach Speed

2.1.1.5 Subjective Data Analysis

The questionnaire analyses addressed three main domains: perceived workload (NASA-TLX), simulator fidelity (Questionnaire 5), and simulation sickness (Questionnaire 6), integrating descriptive statistics and inferential testing.

Regarding the NASA-TLX, the Shapiro–Wilk test confirmed normality across all scenarios (S0: $p = 0.250$; S1: $p = 0.240$; S2: $p = 0.510$; N0: $p = 0.264$; N1: $p = 0.568$; N2: $p = 0.128$), thus supporting the use of parametric procedures.

Under daytime conditions, mean workload scores were 39.53 for S0, 40.85 for S1, and 44.54 for S2. One-way ANOVA revealed no statistically significant differences among scenarios ($F(2,121) = 0.788$; $p = 0.457$), with a very small effect size ($\eta^2 = 0.013$).

In nighttime conditions, mean scores were 43.53 for N0, 43.92 for N1, and 44.85 for N2 . Again, ANOVA showed no significant differences ($F = 0.062$; $p = 0.940$; $\eta^2 = 0.001$), indicating negligible scenario-related effects.

Paired day–night comparisons further confirmed the absence of significant differences:

- S0 vs N0: $t(84) = -0.99$; $p = 0.325$; Cohen's $d = -0.214$
- S1 vs N1: $t(81) = -0.76$; $p = 0.452$
- S2 vs N2: $p = 0.936$ (difference < 1 point)

Overall, NASA-TLX scores consistently ranged between 40 and 45 out of 100, indicating a moderate and stable workload that did not vary significantly across infrastructural configurations or lighting conditions.

With respect to simulation sickness, 8 out of 58 initial participants discontinued the experiment due to discomfort, corresponding to an overall incidence of 13.8%. Age-stratified analysis revealed a marked difference: 46.2% of participants over 45 years old (6 out of 13) interrupted the session, compared with only 4.4% among participants under 45 (2 out of 45). Among those who completed the experiment, most reported no significant symptoms; when present, complaints were generally mild and mainly related to nausea or visual fatigue. The use of the eye tracker did not generate relevant issues, with only a single isolated case of reported discomfort.

Regarding simulator fidelity, approximately half of the participants had prior experience with driving simulators, while the remaining half were first-time users. The preliminary training phase was judged almost unanimously adequate in both daytime and nighttime sessions.

Perceived realism ratings showed a predominance of “fairly realistic” judgments for speed perception, acceleration, and braking, although some critical responses emerged, particularly concerning steering response in nighttime conditions. Mirror functionality and environmental representation received predominantly positive evaluations.

In summary, the questionnaire results indicate:

- a moderate and stable perceived workload across scenarios;
- an overall positive evaluation of simulator realism and immersion;
- a limited incidence of simulation sickness, concentrated in older participants;
- no statistically significant differences between daytime and nighttime conditions in subjective measures.

These subjective findings are consistent with the objective performance data, suggesting that the infrastructural countermeasures introduced in the design scenarios did not increase perceived cognitive workload, while still influencing certain behavioural driving parameters.

2.1.2 Conclusions

Overall, the findings indicate that the design interventions implemented in Scenarios 1 and 2 produced measurable safety benefits when compared with the baseline configuration (Scenario 0). In the baseline scenario, characterized by a recessed stop line, a substantial proportion of drivers advanced beyond the stop line while waiting for an acceptable gap, particularly under the 5 s gap and during nighttime conditions. This behaviour is interpreted as a compensatory strategy: drivers tended to move forward in order to improve visibility and reduce the distance required to complete the left-turn manoeuvre. The countermeasures

introduced in S1 and S2 mitigated this tendency, reducing the frequency and severity of stop-line violations and improving overall manoeuvre discipline.

With regard to gap selection, the absence of statistically significant differences in the number of accepted gaps across scenarios suggests that drivers' fundamental gap acceptance strategy remained relatively stable. In most conditions, participants accepted the first available gap, indicating a consistent and homogeneous decision-making pattern. Thus, the infrastructural modifications did not substantially alter the propensity to accept early crossing opportunities but rather influenced the qualitative safety conditions under which those manoeuvres occurred.

The analysis of the interfered vehicle type further clarifies this dynamic. The distribution of interference types varied significantly across scenarios and gap configurations, with several χ^2 tests confirming non-random distributions (e.g., $\chi^2 = 19.891$; $p = 0.003$; Cramer's $V \approx 0.324$ in N0_5–N1_5–N2_5 comparisons). Across all conditions, Auto–Motorcycle (A–M) interactions were the most frequent configuration, both during daytime (91 cases) and nighttime (79 cases), whereas Auto–Auto interactions remained stable (62 cases in both contexts) and Motorcycle–Auto interactions slightly increased at night (59 vs. 54). These results indicate that vulnerable users, particularly motorcyclists, are systematically involved in a large share of critical interactions. However, in several project scenarios the distribution appeared more balanced, suggesting that the proposed countermeasures partially mitigated the predominance of A–M conflicts.

More broadly, the statistical synthesis confirms that the infrastructural solutions had a positive impact on surrogate safety indicators and on the dynamics of the left-turn manoeuvre, while differences between daytime and nighttime conditions were generally not systematic or consistently significant. This suggests that the effectiveness of the proposed design interventions is relatively robust with respect to lighting conditions.

The Conclusions chapter consolidates these findings within the broader objectives of the thesis. The study focused on critical interactions between motor vehicles and vulnerable road users, particularly motorcyclists, in a complex urban intersection in Rome. After a detailed accident analysis based on official municipal data and field inspections, the most critical manoeuvre (left-turn, manoeuvre D) was reconstructed in a driving simulator environment. A realistic digital twin of the existing configuration (S0) was developed using BIM tools and imported into Unity, and two design scenarios (S1 and S2) were implemented to test targeted infrastructural and signage countermeasures.

The experimental design combined three infrastructural scenarios, two temporal gap configurations (5 s corresponding to approximately 70% acceptance probability, and 6 s corresponding to approximately 90%), and two lighting conditions (day and night), yielding twelve experimental conditions. Surrogate safety measures (e.g., PET, TET, TIT, DRAC), kinematic indicators (approach speeds, speed profiles), behavioural

variables (gap selection, stop-line encroachment), and subjective measures (NASA-TLX, DBQ, simulator fidelity and sickness questionnaires) were analysed using an articulated statistical framework.

In summary, the results demonstrate that a simulation-based approach grounded in a realistic digital reconstruction of the site allows for a rigorous ex ante assessment of infrastructural countermeasures. The results indicate that the proposed design solutions reduce risky behaviours (notably stop-line encroachment and unbalanced interference patterns) and improve safety margins without substantially altering drivers' fundamental gap acceptance strategies. The limited and non-systematic differences between daytime and nighttime conditions further support the robustness of the interventions. Collectively, the findings confirm the methodological validity of combining accident analysis, digital twin modelling, and surrogate safety measures within a controlled simulator environment to support evidence-based urban road safety design.

2.1.3 Vehicle/Pedestrian Interaction

2.1.3.1 Outliers

The experimental dataset was preliminarily screened to identify and manage anomalous observations that could bias the estimation of descriptive statistics and inferential models

Given the repeated-measures structure of the experiment and the presence of interdependent kinematic and temporal surrogate safety indicators (e.g., Time to Zebra, initial speed Vi, minimum speed Vmin, mean deceleration MD, Time to Collision, and Post-Encroachment Time), outliers were defined as observations that significantly deviated from the distribution of the corresponding variable within the same experimental scenario. Outlier identification was conducted using a two-step approach:

1. Statistical criterion (IQR method). For each variable, the interquartile range (IQR = Q3 − Q1) was computed and observations falling below Q1 − 1.5 × IQR or above Q3 + 1.5 × IQR were flagged as potential outliers. This non-parametric method was selected to avoid assumptions regarding normality.
2. Physical plausibility check. Flagged observations were further evaluated against kinematic and behavioral constraints. Implausible values (e.g., unrealistic deceleration rates, incoherent temporal indicators, or values incompatible with the absence of collision) were removed. Conversely, statistically extreme but physically plausible observations were retained, as they may reflect genuine behavioral variability.

Outlier removal was performed at the single-observation level, preserving the repeated-measures structure of the dataset. No participant was excluded entirely unless systematic non-compliance with the experimental protocol was evident.

2.1.3.2 Statistical Analysis

2.1.3.2.1 Time To Zebra

Time to Zebra (TTZ) is defined as the temporal interval required for the vehicle to reach the pedestrian crossing from a given position, assuming the instantaneous speed recorded at that moment. It is conceptually derived from the relationship between the distance separating the vehicle from the crossing (LVi) and the corresponding instantaneous speed (Vi), both parameters explicitly identified among the variables deducible from the speed diagram.

Within the experimental procedure, TTZ plays a central role in determining the instant of pedestrian activation. Specifically, the pedestrian's starting time is defined as the first instant in which TTZ falls below a predetermined threshold value, 2.5 s or 3.5 s, depending on the experimental condition. If TTZ remains above the threshold until the vehicle reaches a distance of 15 m from the crossing, the pedestrian is triggered at that 15 m distance.

From a methodological perspective, TTZ classes were incorporated into the Linear Mixed Model as fixed factors, together with Scenario (S0, S1, S2) and Illumination conditions. The adoption of the LMM framework was justified by violations of ANOVA assumptions, including lack of normality, non-independence of repeated measures, and data imbalance following outlier exclusion. The mixed-model approach therefore allowed the estimation of systematic effects while accounting for repeated observations within participants.

The results indicate that TTZ classes significantly influence multiple temporal and dynamic surrogate safety measures. In particular, the effect of TTZ on PET is statistically significant ($F = 6.540$, $p = 0.002$), confirming that shorter TTZ values are associated with reduced temporal clearance at the conflict point. Short TTZ classes correspond to narrower temporal windows in which the conflict unfolds, requiring more rapid manoeuvre management and resulting in more critical interaction conditions.

Furthermore, TTZ interacts with Scenario in shaping other safety indicators. For PET, while the main effect of TTZ is significant, its influence is structurally intertwined with the experimental configuration, since pedestrian activation logic differs between S0 and the project scenarios S1 and S2. In S0, pedestrian crossing is more frequently triggered by reaching the TTZ threshold, whereas in S1 and S2 lower approach speeds often lead to activation at the fixed 15 m distance. This structural difference affects the distribution of temporal indicators and must be considered when interpreting TTZ-related outcomes.

The descriptive comparison of TTZ-related classes, reported in association with GAP values, shows that more cautious TTZ classes (TTZ* > 3.5 s), reported in the following Figure 8, are associated with substantially larger temporal margins compared to aggressive classes (TTZ* < 2.5 s), while intermediate classes occupy an intermediate position. This confirms the behavioural consistency of TTZ as an anticipatory timing measure: larger TTZ values reflect earlier deceleration onset and greater temporal availability for safe interaction management.

Scenario	TTZ*<=2,5	2,5<TTZ*<=3,5	TTZ*>3,5
S0	45%	41%	14%
S1	8%	31%	61%
S2	16%	29%	56%
N0	39%	41%	19%
N1	16%	36%	48%
N2	18%	39%	43%

Scenario	TTZ*<=2,5	2,5<TTZ*<=3,5	TTZ*>3,5
S0	21%	26%	53%
S1	10%	26%	64%
S2	3%	28%	69%
N0	30%	25%	45%
N1	15%	22%	63%
N2	8%	24%	68%

Figure 8 – Time To Zebra* range categories.

2.1.3.2.2 Mean Deceleration

Mean Deceleration (MD) is included among the variables deducible from the vehicle speed diagram and is calculated between the instant of deceleration onset and the instant corresponding to the minimum speed reached before or at the pedestrian crossing. This parameter is derived from the recorded longitudinal speed profiles sampled at 1 m spatial intervals during each simulation run.

From a conceptual standpoint, mean deceleration represents the average intensity of the braking maneuver adopted by the driver during the interaction with the pedestrian. Deceleration rate is commonly interpreted as an indicator of the driver's intention to yield, since the act of slowing down communicates the willingness to give way.

Linear Mixed Model (LMM) was selected in place of repeated-measures ANOVA due to violations of normality, independence, and sphericity assumptions, as well as the presence of unbalanced data after outlier exclusion. The LMM approach allows the inclusion of fixed effects (Scenario, Illumination, Pedestrian Presence or TTZ classes depending on the model specification) and random effects associated with participants, thus accounting for the hierarchical and repeated-measures structure of the dataset.

From an interpretative view, the results indicate that infrastructure countermeasures influence braking intensity. This pattern is indirectly confirmed by the behavior observed in related dynamic indicators, such as DRAC, for which a significant main effect of Scenario is reported ($p < 0.001$) and interpreted as evidence that design interventions substantially modify braking intensity. Given that DRAC and mean deceleration both quantify aspects of longitudinal braking demand, the discussion supports the interpretation that the project

scenarios (S1 and S2) are associated with less extreme braking requirements compared to the baseline scenario S0.

Descriptive interpretation indicates that in the current state configuration (S0), higher braking intensities are more frequently observed, consistent with later decision-making and more compressed temporal margins. In contrast, in scenarios S1 and S2, the presence of countermeasures reduces initial approach speed and increases the distance at which deceleration begins, leading to more gradual braking patterns.

Furthermore, the methodological framework highlights that TTZ classes significantly influence temporal and dynamic parameters. Short TTZ classes are associated with narrower temporal windows and more rapid conflict management, which logically implies higher deceleration magnitudes. Conversely, larger TTZ classes correspond to earlier deceleration onset and reduced braking intensity. This structural relationship between anticipatory timing and braking demand is coherent with the thesis interpretation of surrogate safety indicators as interdependent measures.

The adoption of the Linear Mixed Model ensures that inter-subject variability is appropriately modeled. Consequently, the estimated fixed effects for Scenario and TTZ are interpreted as systematic influences beyond individual heterogeneity.

Overall, mean deceleration emerges as a dynamic surrogate safety indicator sensitive to infrastructure configuration and anticipatory behavioral classes. The baseline scenario is characterized by more frequent high-intensity braking maneuvers, whereas the project scenarios (S1 and S2) promote smoother deceleration patterns through lower approach speeds and earlier reaction distances. This confirms that the implemented countermeasures contribute to moderating braking severity and, consequently, to reducing interaction criticality within the simulated driver–pedestrian environment.

2.1.3.2.3 PET

Post-Encroachment Time (PET) is defined as the temporal distance between the instant in which the first user leaves the pedestrian crossing and the instant in which the second user reaches the same location. In the present experimental framework, the two interacting users are the driver and the pedestrian, and PET is computed using the effective arrival and exit times, thus accounting for real variations in speed and trajectory.

According to literature (Gettman et al., 2008; Lanzaro & Sayed, 2024; Peesapati et al., 2018), PET can be interpreted as a surrogate safety measure capable of quantifying the severity of driver–pedestrian interference.

Within the inferential analysis, PET was modeled using a Linear Mixed Model with Scenario (S0, S1, S2), Illumination (Day, Night), and TTZ classes as fixed factors. The Type III tests of fixed effects for PET show a

statistically significant main effect of Scenario ($F = 7.917$, $p < 0.001$), indicating that the infrastructure configuration significantly influences the temporal separation between driver and pedestrian. Illumination does not appear as a significant main effect ($p = 0.637$), whereas TTZ classes significantly affect PET ($F = 6.540$, $p = 0.002$), confirming a systematic relationship between anticipatory behavior and residual temporal margin.

The covariance parameter estimates reveal a residual variance of 4.658 and a negligible random intercept variance, indicating that most variability is explained by fixed effects rather than strong inter-subject heterogeneity in this specific model configuration.

The comparison of mean PET values across scenarios highlights clear differences: S0 presents an average PET of approximately 3.9 s, whereas S1 and S2 show lower mean values of approximately 2.1 s and 2.3 s respectively. Pairwise comparisons indicate that S0 significantly differs from both S1 and S2, while S1 and S2 do not significantly differ from each other. However, the raw comparison of PET means across scenarios does not fully represent the effective temporal distancing between users. This distortion arises from differences in pedestrian activation mechanisms: in S1 and S2, due to lower approach speeds induced by countermeasures, the pedestrian frequently starts crossing at a fixed 15 m vehicle distance, whereas in S0 the crossing is more often triggered by reaching the TTZ threshold. Consequently, temporal indicators, including PET, are influenced by this structural difference in activation logic.

Overall, PET emerges as a sensitive surrogate safety indicator significantly influenced by Scenario configuration and TTZ class. While illumination does not exert a statistically significant main effect, infrastructure countermeasures and anticipatory behavioral classes meaningfully shape the temporal clearance achieved at the conflict point.

2.1.4 Conclusion

The results obtained from the analysis of TTZ, Mean Deceleration, and PET provide a coherent and internally consistent representation of driver–pedestrian interaction dynamics within the simulated environment. The combined interpretation of these indicators confirms that infrastructure configuration significantly influences both anticipatory behaviour and conflict resolution mechanisms.

TTZ emerges as the primary anticipatory indicator, structurally embedded in the experimental logic and directly influencing pedestrian activation. Short TTZ classes are systematically associated with more critical temporal conditions, reduced PET values, and higher braking intensity, confirming that delayed decision-making propagates across multiple surrogate safety measures. Conversely, higher TTZ values reflect earlier deceleration onset and wider temporal margins, representing more cautious behavioural strategies.

Mean Deceleration complements TTZ by quantifying the intensity of the evasive manoeuvre. The baseline configuration (S0) is characterised by more frequent high-intensity braking, consistent with

compressed temporal windows and later reactions. In contrast, scenarios S1 and S2 promote smoother deceleration patterns, suggesting that countermeasures improve perceptual clarity and facilitate earlier speed adaptation.

PET provides the final measure of residual temporal separation at the conflict point. The significant main effect of Scenario and TTZ class confirms that infrastructure design and anticipatory timing jointly shape the effective clearance achieved between users. However, the structural difference in pedestrian activation logic across scenarios requires cautious interpretation of raw PET means, as part of the observed variation is mechanically induced by experimental conditions rather than solely by behavioural change.

Overall, the three indicators converge toward a consistent conclusion: infrastructure countermeasures contribute to earlier reactions, reduced braking severity, and improved temporal coordination between driver and pedestrian. The adoption of the Linear Mixed Model ensures that these effects are interpreted beyond individual heterogeneity, strengthening the robustness of the findings. Collectively, the results support the effectiveness of the proposed design interventions in moderating interaction criticality and enhancing surrogate safety performance within the simulated crossing scenarios.

2.2 Pedestrian Simulator Study results

2.2.1 Pedestrian/Vehicle Interaction

2.2.1.1 Outliers

A preliminary screening was conducted to identify potential outliers within the dataset to ensure data reliability and the validity of subsequent analyses. Although several observations exhibited extreme values relative to the overall distribution of the variables, they were not removed. In the specific context of this study, such values do not reflect measurement errors or technical anomalies, but rather represent genuinely observable behavioral variability, including cautious behaviors compared to the average. Therefore, outliers were retained in the dataset to preserve the natural variability of the phenomenon under investigation and to provide a more comprehensive and realistic representation of behavioral differences within the analyzed population.

2.2.1.2 Statistical Analysis

The inferential analysis was performed using a Linear Mixed Model (LMM), as the dataset structure and experimental design were not compatible with traditional repeated measures ANOVA (RM-ANOVA). Although the study followed a multifactorial repeated measures framework, the presence of structurally missing data (e.g., when participants did not perform a crossing) violated the complete-case requirement of

RM-ANOVA. Since these missing values were behavior-driven and not random, listwise deletion would have introduced bias and substantially reduced the sample size.

Moreover, the design was partially unbalanced, making the assumptions of sphericity and homogeneity of covariances required by RM-ANOVA unreliable. In such contexts, methodological guidelines recommend Linear Mixed Models as a more flexible and robust alternative.

The LMM accounts for within-subject dependence through a random intercept for participants, properly handles missing data, and accommodates unbalanced designs. Fixed effects included sex, age, vehicle type, lighting, and scenario, with interactions up to the third order. The analysis focused on how pedestrian behavior varies across vehicle, scenario, and lighting conditions, particularly within the most frequent temporal gaps (2.5–3.5 s), ensuring more reliable and comprehensive inferential results.

The inferential analysis of Time to Zebra (TTZ) and Waiting Time (WT) was performed using linear mixed-effects models with a random intercept for participant, to account for repeated measures and intra-subject correlation.

With reference to Time to Zebra in Piazzale delle Medaglie d'Oro, the Intraclass Correlation Coefficient (ICC) ranged between 0.16 and 0.19, confirming that approximately 16–19% of total variance was attributable to stable inter-individual. The Wald tests showed that Age ($\beta = 0.299$; $p = .001$), Vehicle ($\beta = 0.264$; $p < .001$), and Scenario were significant predictors of TTZ. Compared to Scenario S2 (reference category), Scenario S0 was associated with a significant increase in TTZ ($\beta = 0.226$; $p < .001$), while S1 showed a non-robust increase ($\beta = 0.110$; $p = .076$). Sex and Illumination did not reach statistical significance ($p > .05$).

In Via Prisciano, the mixed model for TTZ showed a −2LL of 857.77 (AIC = 861.77; BIC = 870.39), with a marginal R^2 of 0.176 and a conditional R^2 of 0.375. The ICC was 0.241 (adjusted ICC = 0.199), indicating that about 20–24% of the variance was explained by between-subject differences. In this site, Sex ($F = 0.288$; $p = .594$) and Age ($F = 2.615$; $p = .113$) were not significant. The effect of Vehicle was marginal ($F = 3.805$; $p = .052$), with slightly higher TTZ values for motorcycles (M = 2.278 s) compared to cars (M = 2.201 s). Scenario and its interactions with Illumination were significant, and TTZ was systematically higher in S0 (e.g., Moto: M = 2.548 s; Auto: M = 2.373 s) compared to S2 (Moto: M = 2.085 s; Auto: M = 1.930 s).

Regarding Waiting Time in Piazzale delle Medaglie d'Oro, the model showed a −2LL of 3451.09 (AIC = 3455.09; BIC = 3463.68), with a marginal R^2 of 0.200 and a conditional R^2 of 0.428. The ICC ranged between 0.23 and 0.28, confirming substantial inter-individual variability. Vehicle, Illumination, and Scenario significantly influenced WT, and a significant three-way interaction Vehicle × Illumination × Scenario was observed ($p \approx .021$), indicating that the prolongation of waiting time was particularly evident in the presence of a car under Scenario S0 and daytime conditions.

In Via Prisciano, the model for WT showed −2LL = 3606.05 (AIC = 3610.05; BIC = 3618.67), with a marginal R^2 of 0.208 and a conditional R^2 of 0.387. The ICC ranged between 0.18 and 0.23. The effect of Vehicle was highly significant: pedestrians waited on average 10.90 s in the presence of a motorcycle and 14.65 s in the presence of a car, with a mean difference of approximately 3.75 s ($p < .001$). Scenario also showed a significant effect, whereas Sex, Age, Illumination, and their interactions were not statistically significant ($p > .05$).

Overall, across both sites, TTZ and WT were predominantly influenced by situational and infrastructural factors (vehicle type and scenario), with p-values frequently below .001 for the strongest effects, whereas socio-demographic variables generally did not reach statistical significance. These findings confirm that pedestrian crossing decisions are primarily context-driven and sensitive to roadway configuration and vehicle-related perceptual characteristics.

2.2.1.3 Ordinal Logit Model (GEE) on the Selected Gap Variable

The selection of the temporal gap used by pedestrians to initiate the crossing manoeuvre was analysed through an ordinal logistic regression model estimated using Generalized Estimating Equations (GEE). The use of GEE allows the correlation structure generated by repeated observations belonging to the same participant to be explicitly modelled, thus producing robust estimates of regression coefficients and standard errors. The dependent variable Selected Gap is an ordered categorical variable representing the temporal interval chosen by the pedestrian to begin the crossing manoeuvre (2.5 s, 3 s, 3.5 s). Given the ordinal structure of this variable, the cumulative logit model was adopted.

The general formulation of the cumulative logit model can be expressed as:

$$\text{logit}\big(P(Y_{ij} \leq k)\big) = \ln\left(\frac{P(Y_{ij} \leq k)}{1 - P(Y_{ij} \leq k)}\right) = \alpha_k - (\beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \ldots)$$

where:

Y_{ij} represents the gap selected by subject i at observation j;

k identifies the ordered gap category;

α_k are the threshold parameters;

β are the regression coefficients associated with the predictors.

The explanatory variables included in the model were Sex, Age, Vehicle type, Illumination condition, and Scenario configuration. The correlation between repeated observations within each participant was handled through an exchangeable working correlation matrix.

2.2.1.3.1 Piazzale delle Medaglie d'Oro

At the Piazzale delle Medaglie d'Oro site, the distribution of the selected temporal gaps reveals a clear behavioural tendency toward short intervals. Specifically, 53.0% of observations correspond to the 2.5 s gap, 40.1% to the 3 s gap, and only 6.8% to the 3.5 s gap. This distribution suggests that pedestrians tend to accept relatively short temporal opportunities before initiating the crossing manoeuvre, indicating that approximately three seconds represent a central behavioural threshold in this context.

The ordinal logit model confirms that Vehicle type and Scenario configuration are the only predictors that significantly influence gap selection, while Sex, Age, and Illumination do not show statistically significant effects. The coefficient associated with the car condition is negative ($\beta = -0.674$), with an odds ratio equal to 0.509, indicating that the presence of a car reduces the probability of accepting shorter gaps compared with the motorcycle condition, which represents the reference category. This result suggests that pedestrians adopt a more cautious behaviour when interacting with cars, likely due to differences in perceived size, speed estimation, or perceived collision severity. The Scenario variable also exerts a strong influence on gap selection. Compared with Scenario 0, which represents the baseline configuration, Scenario 2 produces a substantial increase in the probability of crossing, with an estimated odds ratio of approximately 14.053, while Scenario 1 shows a smaller but directionally consistent increase ($OR \approx 1.424$). These results define a clear behavioural gradient across infrastructural configurations, with Scenario 0 representing the most conservative condition and Scenario 2 promoting a greater willingness to initiate the crossing manoeuvre.

Overall, the results obtained for Piazzale delle Medaglie d'Oro indicate that pedestrian gap selection is primarily determined by situational variables related to traffic dynamics and infrastructural configuration, whereas demographic characteristics do not appear to play a significant role in the decision-making process.

2.2.1.3.1.1 Via Prisciano

A comparable behavioural structure emerges in the Via Prisciano site, although the distribution of the selected temporal gaps shows an even stronger concentration toward shorter intervals. In this case, 50.8% of observations correspond to the 2.5 s gap, 42.8% to the 3 s gap, and only 6.4% to the 3.5 s gap. This pattern confirms that pedestrians generally tend to initiate the crossing manoeuvre during the earliest available temporal opportunities, as visible in the following Figure 9.

The ordinal logit model again identifies Vehicle and Scenario as the only statistically significant predictors influencing gap selection (Vehicle: $\chi^2 = 40.735$; $p < .001$; Scenario: $\chi^2 = 8.017$; $p = .018$), whereas Sex ($p = .192$), Age ($p = .434$), and Illumination ($p = .787$) do not exhibit statistically significant effects.

Consistent with the results obtained in the other site, the presence of a motorcycle is associated with a greater probability of accepting shorter temporal gaps, while the presence of a car induces relatively more cautious behaviour. These findings suggest that the perceptual characteristics of the approaching vehicle influence the pedestrian's estimation of available crossing time and perceived collision risk.

Overall, the results obtained for Via Prisciano confirm the dominant role of situational and infrastructural factors in shaping pedestrian crossing decisions, reinforcing the interpretation that gap acceptance behaviour is primarily influenced by contextual conditions rather than by stable individual characteristics.

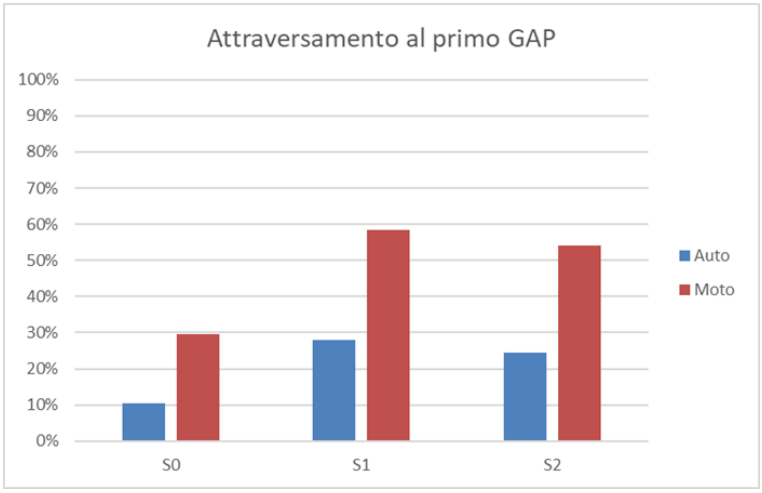


Figure 9 - Via Prisciano first GAP acceptance .

2.2.1.4 Linear GEE Regression on Temporal Behaviour (TTZ and WT)

In addition to the analysis of gap selection, temporal behavioural indicators were examined using linear regression models estimated through Generalized Estimating Equations (GEE). This modelling approach allows continuous dependent variables to be analysed while accounting for the correlation between repeated observations belonging to the same participant.

The models were applied to two key temporal variables:

- Time to Zebra (TTZ), representing the temporal distance between the approaching vehicle and the pedestrian at the moment the crossing manoeuvre is initiated.

- Waiting Time (WT), representing the duration of time during which the pedestrian remains stationary before starting the crossing manoeuvre.

The general formulation of the linear GEE model can be expressed as:

$$Y_{ij} = \beta_0 + \beta_1 \text{Sex}_{ij} + \beta_2 \text{Age}_{ij} + \beta_3 \text{Vehicle}_{ij} + \beta_4 \text{Illumination}_{ij} + \beta_5 \text{Scenario}_{ij} + \epsilon_{ij}$$

where:

Y_{ij} represents the observed value of the temporal variable for subject i at observation j;
β_0 is the intercept;
β_k are the regression coefficients associated with the predictors;
ϵ_{ij} represents the residual error term.

2.2.1.4.1 Piazzale delle Medaglie d'Oro

The linear GEE models estimated for Piazzale delle Medaglie d'Oro indicate that both Vehicle type and Scenario configuration significantly influence pedestrian temporal behaviour. For the Waiting Time (WT) variable, reported in the following Figure 10, both predictors reach high levels of statistical significance ($p < .001$). In particular, the presence of a motorcycle is associated with a reduction in waiting time of approximately −3.492 seconds relative to the car condition. This result suggests that pedestrians perceive motorcycles as requiring shorter temporal margins before initiating the crossing manoeuvre. Conversely, Scenario 0 is associated with a significant increase in waiting time compared with Scenario 2 (+4.817 seconds, $p < .001$). This indicates that pedestrians tend to delay the crossing manoeuvre in the baseline infrastructural configuration, reflecting a more conservative behavioural response. A similar pattern is observed for the Time to Zebra (TTZ) variable. Scenario 0 systematically produces higher TTZ values, indicating that pedestrians initiate the crossing manoeuvre when the approaching vehicle is still temporally farther from the crossing point. This behaviour can be interpreted as a risk-avoidance strategy adopted in less favourable infrastructural configurations.

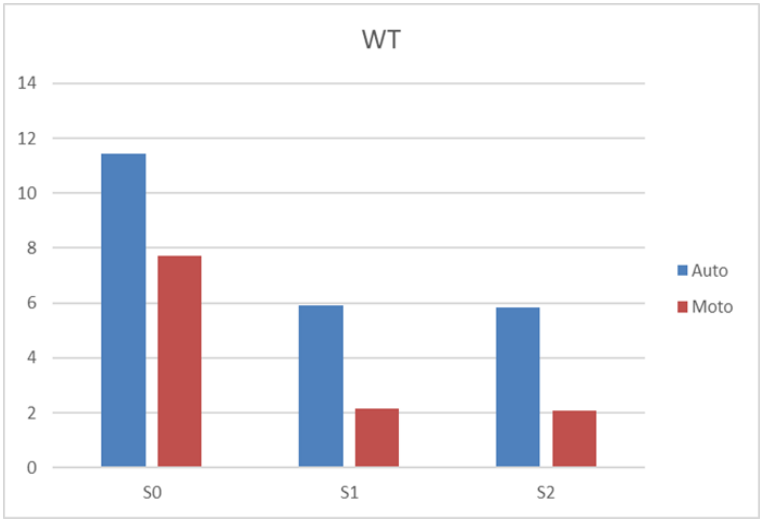


Figure 10 - Via Prisciano Waiting Time.

2.2.1.4.2 Via Prisciano

In Via Prisciano, the linear GEE models again identify Scenario and Vehicle as the most relevant predictors of temporal behaviour, whereas Sex, Age, and Illumination do not exhibit statistically significant effects. For the Waiting Time (WT) variable, pedestrians tend to wait longer in the presence of a car compared with a motorcycle, reflecting a higher perceived level of risk associated with larger vehicles. However, unlike Piazzale delle Medaglie d'Oro, no statistically significant interaction effects were detected between predictors. This absence of interaction terms suggests that the behavioural responses observed in Via Prisciano follow a more linear and stable structure, with contextual variables exerting their influence independently rather than through complex combined effects.

Overall, the results obtained for Via Prisciano indicate that pedestrian temporal behaviour is again primarily determined by contextual characteristics related to vehicle type and infrastructural configuration, while demographic variables do not significantly contribute to explaining behavioural variability.

2.2.2 Conclusion

This study investigated pedestrian crossing behavior within two urban contexts characterized by different infrastructural configurations, namely Piazzale delle Medaglie d'Oro and Via Prisciano, through an experimental design integrating temporal gap manipulation, vehicle type, lighting condition, and scenario configuration. The analytical framework combined Linear Mixed Models and logistic regression models

estimated through Generalized Estimating Equations, allowing for the appropriate treatment of repeated measures, unbalanced data structures, and behavior-driven missing observations.

A preliminary screening identified extreme observations; however, these were retained in the dataset. In the context of pedestrian behavior analysis, such values represent genuine inter-individual variability rather than measurement error. Preserving these observations ensured ecological validity and maintained the natural dispersion of behavioral responses.

Across both sites, the results consistently demonstrate that pedestrian decision-making is predominantly situational. Vehicle type and infrastructural scenario emerged as the most influential predictors in both linear and logistic models, frequently reaching high levels of statistical significance. In contrast, socio-demographic variables such as sex and age, as well as lighting conditions, generally did not exhibit robust or systematic effects.

The analysis of Time to Zebra (TTZ) revealed that a meaningful proportion of total variance (approximately 16–24%) was attributable to stable inter-individual differences, as indicated by the Intraclass Correlation Coefficients. Nevertheless, situational variables remained the primary determinants. In Piazzale delle Medaglie d'Oro, Age, Vehicle, and Scenario significantly influenced TTZ, whereas Sex and Illumination did not. Scenario S0 was systematically associated with higher TTZ values, reflecting more cautious temporal calibration before crossing. In Via Prisciano, inter-individual variability was slightly higher, but socio-demographic effects were again non-significant, and TTZ was primarily modulated by Scenario and, marginally, by Vehicle.

Waiting Time (WT) analyses further reinforced the situational nature of pedestrian behavior. In both locations, Vehicle and Scenario significantly influenced waiting duration. The presence of a car consistently increased waiting time compared to a motorcycle, reflecting heightened perceived risk. In Piazzale delle Medaglie d'Oro, a significant three-way interaction between Vehicle, Illumination, and Scenario indicated that prolonged waiting was particularly pronounced in Scenario S0 under daytime conditions in the presence of a car. Conversely, Via Prisciano exhibited a more linear and homogeneous pattern, without complex interaction effects, although the magnitude of the vehicle effect remained substantial, with mean waiting times differing by approximately 3.75 seconds between car and motorcycle conditions.

Overall, the linear regression analyses performed on the TTZ and WT variables confirm that temporal aspects of pedestrian crossing behavior are primarily governed by situational conditions, particularly vehicle type and infrastructural configuration. These results indicate that pedestrians dynamically adjust both the timing of crossing initiation (TTZ) and the duration of pre-crossing waiting (WT) in response to contextual traffic characteristics, whereas demographic factors such as sex and age play only a marginal role in explaining the observed variability.

The logistic regression analyses provide further insight into the temporal structure of pedestrian decisions. Gap acceptance probability was concentrated on the earliest available temporal intervals (2.5–3.0 seconds), with sharply declining acceptance for subsequent gaps. This concentration produced quasi-complete separation for later intervals, confirming that pedestrian decision-making is temporally front-loaded and highly sensitive to initial opportunities.

In both sites, Vehicle and Scenario significantly influenced the probability of accepting the first gap. Motorcycles increased the odds of acceptance relative to cars, while Scenario S0 consistently reduced acceptance probability compared to more favorable configurations (S1 and S2). For the second gap, the influence of Vehicle attenuated, particularly in Via Prisciano, whereas Scenario remained predictive, underscoring the structural role of infrastructural configuration in shaping temporal thresholds. Sex, Age, and Illumination did not significantly affect acceptance probability in either site.

When modeling gap choice as an ordered outcome, the cumulative logit framework confirmed that the probability of selecting shorter intervals was strongly associated with situational variables. Higher scenarios increased the odds of selecting shorter gaps, whereas the presence of a car reduced them. The negligible contribution of socio-demographic factors suggests that pedestrian crossing strategies are not primarily driven by stable personal characteristics, but rather by contextual perception of traffic dynamics and roadway configuration.

Comparatively, Piazzale delle Medaglie d'Oro exhibited slightly more complex interaction patterns and stronger contextual modulation, whereas Via Prisciano showed a more stable and homogeneous behavioral structure. Nonetheless, both sites converge on the same fundamental conclusion: pedestrian crossing behavior is decisively influenced by vehicle-related perceptual characteristics and infrastructural scenario, rather than by demographic attributes.

Overall, the findings confirm that pedestrian decision-making at unsignalized crossings is primarily context-driven. The configuration of the roadway environment and the dynamic characteristics of the approaching vehicle shape perceived risk, temporal thresholds, and waiting strategies. These results carry important implications for urban design and traffic safety policy, emphasizing that infrastructural interventions and traffic composition management may have a greater impact on pedestrian safety than demographic targeting. By demonstrating the structural dominance of situational variables across two distinct urban contexts, this study contributes to a deeper understanding of how environmental design modulates pedestrian behavior and critical gap acceptance mechanisms.

2.3 Motorcycle Simulation Study results

2.3.1 Motorcycle/Pedestrian Interaction (Via Prisciano Case Study)

2.3.1.1 Outliers

The dataset exhibited several deviations from normality, particularly in variables related to risk exposure during pedestrian interaction. The Shapiro–Wilk tests frequently indicated significant non-normal distributions ($p < 0.05$), especially for TTC-based parameters (TTCmin, TET, TIT). Levene's test confirmed heteroscedasticity across several parameters, indicating non-constant variance between groups. No automatic exclusion of extreme values was performed. Observed outliers were interpreted as behaviorally meaningful responses under critical conditions rather than measurement errors. To address skewness, zero inflation, and heteroscedasticity, Generalized Linear Models were adopted, ensuring robust inference without relying on normality assumptions. This approach preserved the ecological validity of the findings while maintaining statistical reliability.

2.3.1.2 Statistical Analysis

2.3.1.2.1 Minimum Time to Collision

The Minimum Time to Collision (TTCmin) constitutes one of the central safety indicators adopted in the experimental framework to evaluate the interaction between the motorcyclist and the pedestrian within the defined conflict area. By construction, TTC represents the temporal margin separating the two users under the assumption of constant speed; the minimum value reached inside the conflict zone therefore corresponds to the instant of maximum objective criticality.

In the baseline configuration (Scenario S0), TTCmin assumes systematically lower values compared to the project scenarios. In one of the analyzed conditions, the mean TTCmin in S0 was equal to 2.45 s, while in Scenario S1 it decreased slightly to 2.41 s, and increased significantly in Scenario S2 to 2.74 s. This difference of approximately 0.29 s between S0 and S2, although apparently limited in absolute terms, is highly relevant from a safety perspective, as it represents a non-negligible increase in available reaction time in a critical interaction phase.

The interpretation of these results becomes clearer when TTCmin is analyzed together with the corresponding kinematic parameters. In Scenario S0, the conflict configuration is characterized by a shorter conflict distance (21.52 m) and a lower conflict speed (10.15 m/s), a pattern typical of a late and emergency-type braking maneuver. The motorcyclist approaches the pedestrian with a reduced spatial margin and is forced to apply a more intense deceleration, resulting in a lower TTCmin.

In Scenario S2, despite a conflict speed slightly higher than in S0 (11.24 m/s versus 10.15 m/s), the distance of conflict increases to 25.60 m, producing a higher TTCmin (2.74 s), as reported in the following Figure 11. This demonstrates that safety is not exclusively determined by speed reduction, but by the

anticipatory spatial positioning of the vehicle relative to the pedestrian crossing. The increase in TTCmin in S2 reflects a more structured and proactive behavioral adaptation induced by the infrastructural configuration.

Scenario S1 shows intermediate characteristics: the conflict distance rises to 26.06 m, but the conflict speed reaches 12.41 m/s. Although spatial anticipation is present, the modulation of speed is less controlled than in S2, which explains why TTCmin does not reach the same level of improvement.

From a behavioral standpoint, TTCmin synthesizes the combined effects of initial speed, deceleration strategy, and spatial anticipation. The higher values observed in S2 indicate that the rider maintains a greater temporal buffer at the moment of maximum proximity, reducing the probability of entering critical TTC thresholds. In the baseline scenario, conversely, the lower TTCmin values are coherent with the higher deceleration peaks (up to 5.17 m/s² in S0 in one configuration), confirming a reactive rather than anticipatory braking pattern.

In conclusion, TTCmin emerges as a robust and integrative performance indicator. The quantitative increase observed in Scenario S2, reaching mean values up to 2.74 s compared to 2.45 s in S0, confirms that the proposed design solution effectively improves the temporal safety margin during pedestrian interaction. This enhancement is achieved not merely through speed reduction, but through a more balanced interaction between velocity control and spatial anticipation, ultimately producing a more stable and safer behavioral response.

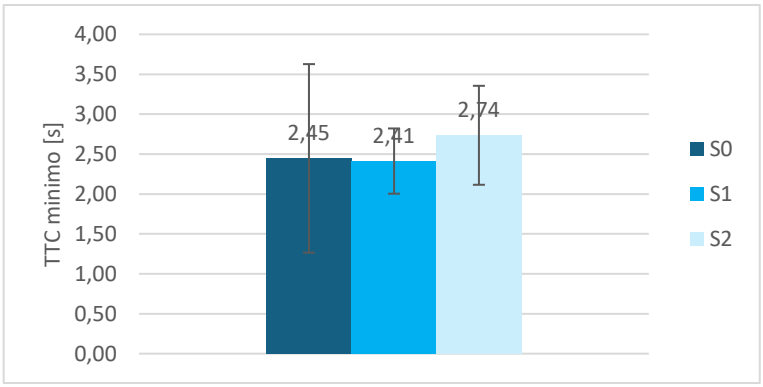


Figure 11 - Minimum TTC across different scenario.

2.3.1.2.2 Conflict Speed

The conflict speed (Vc), reported in the following Figure 12, defined as the speed of the motorcyclist at the exact moment the pedestrian enters the conflict area, represents a key kinematic parameter for understanding the quality of the interaction and the behavioral adaptation induced by the different infrastructural scenarios. Unlike TTCmin, which expresses a temporal safety margin, Vc provides direct information about the dynamic state of the vehicle at the most critical instant of the encounter.

In the baseline scenario S0, the mean conflict speed was equal to 10.15 m/s, associated with a conflict distance of 21.52 m. This configuration reflects a late and abrupt braking maneuver: the rider approaches the

crossing at relatively high initial speeds and is forced to decelerate sharply near the pedestrian, reaching a lower speed only when already very close to the conflict point. The reduced conflict distance and the relatively low Vc therefore do not necessarily indicate a safer condition, but rather a reactive strategy characterized by emergency deceleration. In Scenario S1, the conflict distance increases significantly to 26.06 m, but the conflict speed rises to 12.41 m/s. This suggests that although the rider anticipates the spatial interaction, the modulation of speed is less effective than in S2. The higher Vc indicates that the motorcyclist maintains a greater kinetic state at the moment of pedestrian entry, potentially increasing the severity of a hypothetical impact, even if the spatial margin is improved.

Scenario S2 shows a more balanced configuration. The conflict speed reaches 11.24 m/s, while the corresponding conflict distance is 25.60 m. Although Vc in S2 is slightly higher than in S0 (11.24 m/s versus 10.15 m/s), the increased distance at the conflict point and the higher TTCmin observed in this scenario compensate for the kinetic component. This indicates that the rider does not rely on abrupt braking close to the pedestrian but instead approaches the crossing in a more controlled and anticipatory manner.

From a statistical perspective, the analysis of model effects highlights that Vc is significantly influenced by the temporal parameter TTZ, with a strong effect of increasing TTZ from 2.5 s to 3.5 s ($p < 0.001$). Moreover, a significant interaction between Scenario and TTZ ($p = 0.005$) confirms that the influence of temporal urgency on conflict speed varies across infrastructural configurations. In particular, the baseline scenario S0 appears more sensitive to changes in temporal constraints, whereas S1 and S2 exhibit a more stable behavioral response.

In interpretative terms, conflict speed should not be read in isolation. A lower Vc, as observed in S0, may be the outcome of a late and intense deceleration, while a slightly higher Vc, as in S2, may coexist with a larger safety buffer in terms of space and time. The key difference lies in the quality of the approach: reactive in S0, partially anticipatory in S1, and structurally anticipatory and more stable in S2.

In conclusion, conflict speed confirms the progressive improvement from the baseline to the advanced project configuration. Scenario S2 does not simply minimize speed at the conflict point; rather, it reorganizes the kinematic profile of the approach, combining moderate speed, greater distance, and higher TTC values. This integrated behavior reflects a safer and more controlled interaction with the pedestrian.

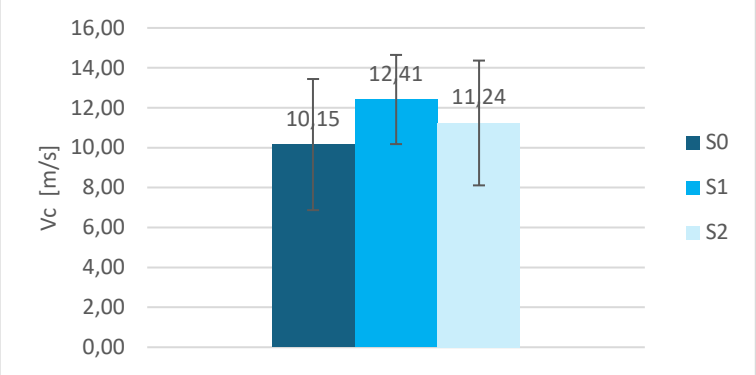


Figure 12 - Vc across different scenarios.

2.3.1.2.3 Mean Deceleration

The mean deceleration (MD), reported in the following Figure 13, represents a fundamental parameter for evaluating the dynamic quality of the rider's response during the approach to the pedestrian crossing. Unlike TTCmin and conflict speed, which describe the safety condition at a specific instant, MD characterizes the braking strategy over the entire deceleration phase, from the onset of speed reduction to the attainment of the minimum speed. It therefore reflects whether the maneuver is anticipatory and progressive or reactive and abrupt.

In the baseline scenario S0, the highest mean deceleration values were recorded. In one configuration, MD reached 5.17 m/s², while in another it remained elevated at 4.59 m/s². These values indicate a strong and sudden braking action, consistent with a delayed perception of risk and a reduced spatial margin. The rider approaches the pedestrian at higher initial speeds and is forced to apply an intense braking force in the final meters before the crossing. This interpretation is supported by the observation that, in S0, the minimum speed drops to as low as 3.52 m/s, confirming the presence of abrupt emergency-type maneuvers.

In contrast, the project scenarios show a clear reduction in braking intensity. In Scenario S1, MD decreases to approximately 3.26 m/s², while in Scenario S2 it further decreases to around 3.18 m/s². In another configuration, the reduction compared to S0 remains evident, confirming a systematic trend. These lower values indicate a smoother and more progressive deceleration profile. The rider begins to modulate speed earlier, avoiding the need for abrupt braking close to the conflict area.

The difference between S0 and S2 is particularly relevant from a safety perspective. A reduction from 5.17 m/s² to values close to 3.2 m/s² represents not only a kinematic improvement but also a qualitative behavioral shift. Lower deceleration peaks imply improved comfort, reduced vehicle instability, and a lower probability of loss of control. Moreover, smoother braking correlates with the higher TTCmin values observed in S2, confirming that temporal safety margins are increased through anticipatory behavior rather than emergency reactions.

It is also important to note that MD must be interpreted jointly with the spatial parameters. In S0, the minimum speed is reached at a distance of only 10.35 m from the crossing, whereas in S1 and S2 this distance increases to 14.17 m and 13.69 m respectively. This anticipatory shift in the braking point explains the reduction in deceleration intensity: the maneuver is distributed over a longer spatial interval, resulting in a more controlled approach.

In conclusion, mean deceleration confirms the progressive effectiveness of the project scenarios. The baseline condition is characterized by late, intense braking maneuvers, while S1 and especially S2 promote earlier and smoother speed modulation. The consistent reduction of MD across configurations demonstrates that the infrastructural modifications do not merely alter instantaneous safety indicators, but fundamentally reshape the dynamic interaction pattern between motorcyclist and pedestrian, leading to a more stable and safer behavioral response.

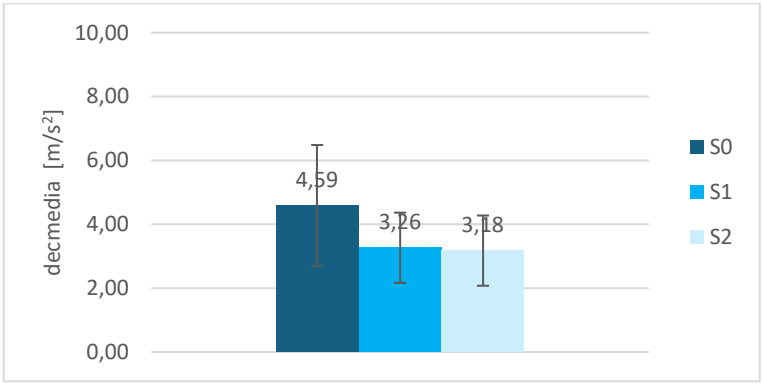


Figure 13 - MD across different scenarios.

2.3.2 Conclusion

The present analysis, conducted on the case study of Via Prisciano, provides consistent evidence that the proposed infrastructural modifications significantly improved the safety performance of motorcyclists during pedestrian interactions. The results demonstrate a coherent behavioral transition from a reactive driving pattern in the baseline configuration (S0) to a more anticipatory and controlled approach in the advanced project scenario (S2).

From a methodological perspective, the dataset exhibited non-normal distributions and heteroscedasticity, particularly in TTC-based risk parameters. Rather than excluding extreme values, these observations were interpreted as meaningful manifestations of behavioral variability under critical conditions. The adoption of Generalized Linear Models ensured robust statistical inference while preserving ecological validity. This approach allowed the analysis to reflect genuine rider responses instead of artificially smoothing the dataset.

With respect to safety indicators, Minimum Time to Collision (TTCmin) emerged as a central integrative parameter. The increase from 2.45 s in S0 to 2.74 s in S2 represents a substantial improvement in temporal safety margin. Even when conflict speeds were not minimized, the combination of increased spatial anticipation and smoother deceleration strategies led to higher TTC values, indicating a more secure interaction profile.

Conflict speed (Vc) further clarified the behavioral dynamics underlying these results. In S0, the lower Vc (10.15 m/s) was associated with a short conflict distance (21.52 m), reflecting late emergency braking. In contrast, S2 exhibited a more balanced configuration, with Vc equal to 11.24 m/s and a larger conflict distance of 25.60 m. This demonstrates that safety improvements were not achieved solely through speed reduction, but through a reorganization of the approach strategy, characterized by earlier spatial positioning and greater temporal buffering.

Mean deceleration (MD) confirmed this interpretation. The baseline scenario recorded peak values up to 5.17 m/s², indicative of abrupt and reactive braking maneuvers. In S1 and S2, MD decreased to approximately 3.26 m/s² and 3.18 m/s² respectively, reflecting smoother and more progressive deceleration. The anticipation of the braking phase, also evidenced by the increased distance at which minimum speed was reached (from 10.35 m in S0 to over 13 m in S1 and S2), reveals a qualitative shift in driving behavior induced by the redesigned scenario.

Taken together, TTCmin, conflict speed, and mean deceleration provide a coherent picture of progressive safety enhancement. Scenario S2 does not simply modify isolated kinematic parameters; it restructures the entire interaction pattern between motorcyclist and pedestrian. The rider approaches the crossing with greater spatial anticipation, maintains a higher temporal buffer, and executes smoother deceleration maneuvers.

In conclusion, the findings confirm that the advanced project configuration effectively mitigates critical interaction conditions and promotes a more stable and anticipatory behavioral response. The improvement is structural rather than marginal, demonstrating that targeted infrastructural interventions can significantly reshape rider behavior and enhance pedestrian safety within simulated urban environments.

3 Simulation studies RU PoliTO

As indicated in Deliverable 3, the research was organised into two driving simulator experiments. The first examined the *Corso Vittorio Emanuele II* corridor to understand how driver interactions with vulnerable road users (pedestrians and cyclists) varied across the different sections of the corridor itself. The second experiment focused on the *Corso Vittorio Emanuele II - Corso Inghilterra* intersection to analyse driver-cyclist interactions and test potential design countermeasures for improving safety.

3.1 Driving simulation experiments of *Corso Vittorio Emanuele II* corridor

This experiment was divided in two sub-experiments, each targeting a different aspect of vehicle-VRU interactions. The first sub-experiment focused on conditions related to pedestrians along the first two corridor segments between *Piazza Rivoli* and the *Via Cesana* intersection, where pedestrians crossed the service road either yielding or not yielding. The second sub-experiment addressed more complex urban situations typical of shared-lane environments with cyclists and pedestrians interacting longitudinally and laterally with the driver.

3.1.1 Results of sub-experiment 1

The sub-experiment 1 focused on drivers' responses to interactions with pedestrians under different behavioural and environmental conditions.

The analysis showed that drivers' braking behaviour, as measured by the proportion of stopping distance (PSD), was primarily influenced by the time gap accepted by pedestrians, lighting conditions, and driver age. When pedestrians initiated the crossing with a larger temporal gap, drivers tended to brake earlier, resulting in higher PSD values and safer approaches. Lighting also played an important role with PSD values generally higher at night. There were also differences across age groups, with middle-aged drivers showing the most balanced reactions when interacting with a pedestrian, likely reflecting an effective combination of experience and response capability (Figure).

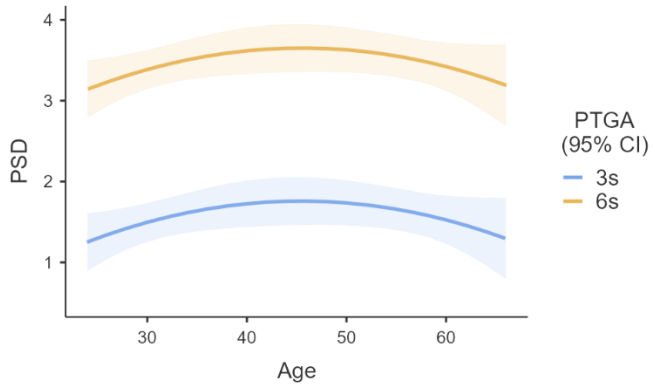
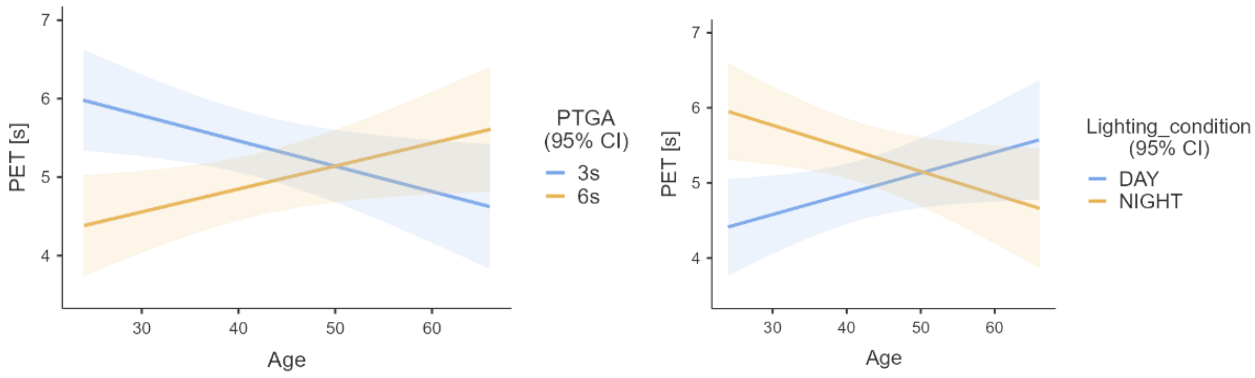


Figure 16 - Estimated marginal means of PSD with respect to age and PTGA

Post-Encroachment Time (PET) confirmed the important roles of time-gap acceptance and lighting. Riskier interactions, as indicated by lower PET values, occurred more frequently during the daytime, which supports the idea that night-time conditions encourage greater caution. Interestingly, larger pedestrian time gaps were sometimes associated with lower PET values. This suggests that, when crossing with a wider margin, pedestrians encouraged drivers to slow down and pass behind them rather than come to a complete stop. This resulted in closer temporal proximity at the interaction point. Age differences were also evident: younger drivers tended to approach at higher speeds and maintain smaller margins, particularly during the day. In contrast, older drivers showed longer recovery times following the interaction. Middle-aged participants exhibited more consistent behaviour across conditions (Figure).



(A) (B)

Figure 17 -Estimated marginal means of PET in respect to (A) PTGA and (B) lighting conditions.

The binary stopping variable (Stop_NoStop) showed that drivers were more likely to yield when faced with more assertive pedestrian behaviour and shorter PTGA values. These findings align with the PET results and suggest that drivers are more likely to stop when interactions are perceived as more critical. Post-hoc analysis revealed a significant interaction between lighting conditions and pedestrian behaviour, showing notably higher sensitivity to assertive pedestrians during the daytime (Table).

Table 6 - Post-Hoc comparison with Holm correction of the Stop_NoStop variable.

Comparison								
Lighting condition	**Pedestrian Behaviour**		**Lighting condition**	**Pedestrian Behaviour**	**Difference**	**SE**	**df**	**p_{Holm}**
DAY	YP	-	NIGHT	YP	-1.276	0.525	Inf	.0606
DAY	YP	-	DAY	NYP	-2.370	0.578	Inf	< .001
DAY	YP	-	NIGHT	NYP	-2.237	0.572	Inf	< .001
NIGHT	YP	-	DAY	NYP	-1.094	0.540	Inf	.1283
NIGHT	YP	-	NIGHT	NYP	-0.961	0.537	Inf	.1476
NIGHT	NYP	-	NIGHT	NYP	0.134	0.517	Inf	.7962

3.1.2 Results of sub-experiment 2

The sub-experiment 2 investigated the impact of various two-level factors on different types of interactions between drivers VRUs.

In the first interaction pattern (Sub. 2.1 according to Deliverable 3), simple pedestrian crossings were compared with more complex situations involving a cyclist present on the right while the pedestrian crossed from the left. The results suggest that driver characteristics had a greater influence on the severity of the interaction than the interaction type itself. Age was found to be a key factor, with older drivers experiencing more critical encounters and younger drivers generally showing longer recovery phases following the interaction. Lighting conditions also contributed, with safer interactions occurring during the daytime, likely due to better visibility and earlier pedestrian detection. While the presence of a cyclist did not significantly alter the statistical outcomes, descriptive trends suggest that this added complexity may negatively impact pedestrian safety (Table) in terms of PET and minimum time-to-collision (MTTC).

Table 7 - Mean and standard deviation of Sub 2.1 variables (MTTC = minimum time-to-collision).

Lighting condition	**Pedestrian Behaviour**	**Mean (Standard Deviation)**	
		PET [s]	MTTC [s]
DAY	Simple	2.69 (1.33)	0.91 (0.93)
	Complex	2.26 (1.92)	0.99 (0.83)
NIGHT	Simple	1.73 (1.92)	0.59 (0.71)
	Complex	2.14 (2.37)	0.61 (0.65)

A different pattern emerged when pedestrians crossed from the right (Sub. 2.2 according to Deliverable 3). In this configuration, the presence of a cyclist had a significant influence on the dynamics of the interaction. Interestingly, interactions were less critical when both a pedestrian and a cyclist were present. This counterintuitive result can be explained by drivers' lateral positioning: the presence of a cyclist on the right induced drivers to move slightly to the left within their lane, thereby increasing the distance from the

pedestrian (Figure). Lighting had a nuanced effect: pedestrians were more noticeable at night in simple interactions, but drivers' attention was potentially redistributed when both VRUs were present.

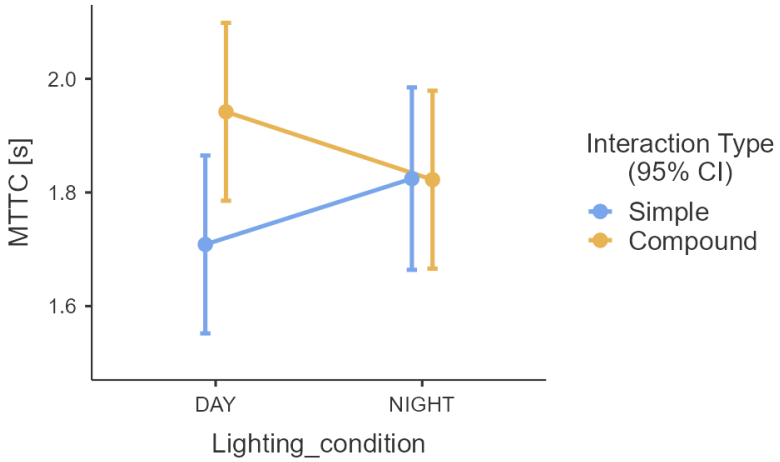


Figure 18 - MTTC estimated marginal means (Sub 2.2).

The third interaction pattern (Sub. 2.3 according to Deliverable 3) examined simple pedestrian crossings, which could be at marked crosswalks or elsewhere, and which occurred under different lighting conditions. Although some safety indicators did not demonstrate significant dependencies, the temporal margin before a potential collision was clearly influenced by crossing visibility. Interactions were safer at marked crosswalks, which confirms the importance of road markings and signage in helping drivers to anticipate. Gender differences also emerged, with female drivers generally exhibiting safer interaction patterns. However, at night, the distinction between marked and unmarked crossings became less evident (Figure), likely due to the reduced visibility of road markings, despite generally cautious driving behaviour.

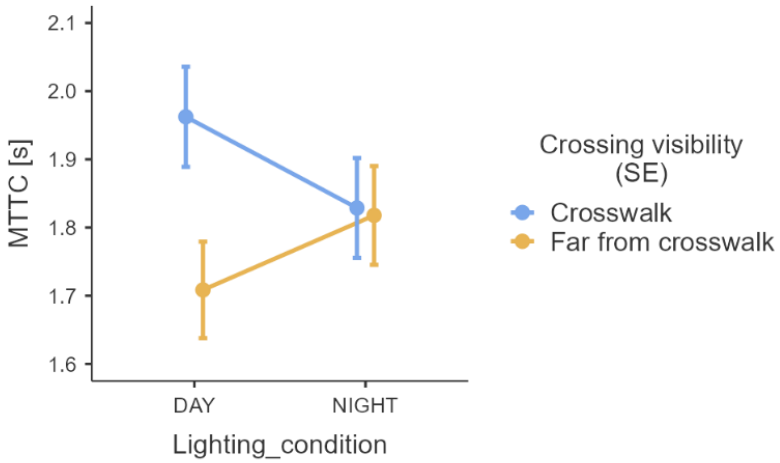
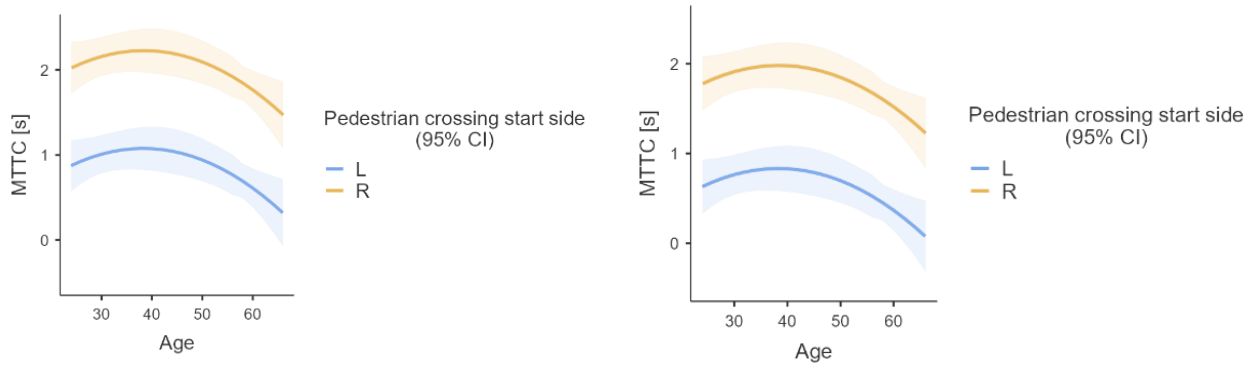


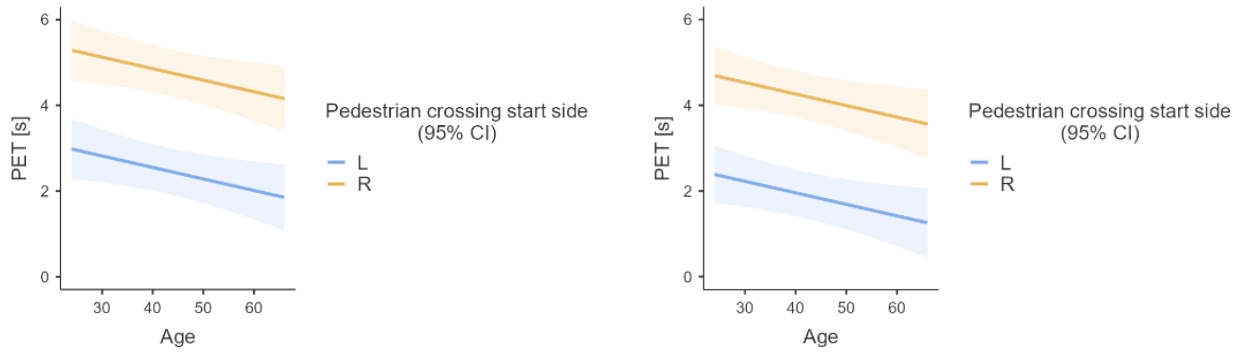
Figure 19 - MTTC estimated marginal means (Sub 2.3).

The fourth interaction pattern (Sub. 2.4 according to Deliverable 3) involved comparing pedestrian crossings on the left and right. Interactions were generally safer when pedestrians approached from the right. This is attributed to the typical positioning of drivers within the lane. MTTC and PET were significantly influenced by age and lighting conditions, while gender also affected PET. Estimated marginal means plots of Figure and Figure , indicate that middle-aged drivers showed the safest interactions on average, reflecting a balance between experience and reaction ability. Younger drivers exhibited higher PET values, which are associated with more aggressive driving and longer recovery times.



(A) (B)

Figure 20 - Estimated marginal means of MTTC in respect to (A) daytime and (B) nighttime.



(A) (B)

Figure 21 - Estimated marginal means of PET in respect to gender: (A) male, (B) female.

The fifth interaction pattern addressed the interaction between the driver and the cyclist in shared lanes. There was little difference in overtaking speed between day and night conditions, suggesting that it is mainly governed by general driving habits rather than visibility. Although the lateral distance maintained from cyclists tended to be greater during the day, likely due to improved recognition, this difference was not significant according to statistical evaluation (Figure).

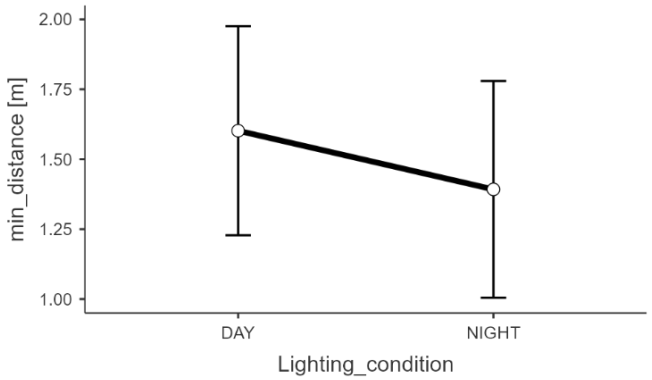


Figure 22 - Estimated marginal means of min_distance variable with respect to lighting conditions

3.1.3 Conclusion

This experimental study examined the safety of vulnerable road users (VRUs) in an urban context, analysing vehicle-VRU interaction patterns and the factors contributing to crashes. The case study focused on the service road of *Corso Vittorio Emanuele II in Turin*, Italy, where the road authority had recently introduced shared lanes, allowing vehicles and cyclists to use the same lane. Crash statistics and collision diagrams suggest that this layout poses a higher risk to cyclists than sections of the corridor where cyclists use dedicated facilities.

To reflect these patterns, the driving-simulation study was organised into two main sections. The first section examined driver responses to pedestrians crossing outside of designated crosswalks. Pedestrians were modelled with two behaviours (yielding versus non-yielding) and two acceptance levels for time gaps, in both daytime and night-time conditions. The second sub-experiment examined driver behaviour when interacting with pedestrians and cyclists in shared-lane conditions. This was simulated at and away from crosswalks, with pedestrians crossing from both sides (right and left side), and in both simple (single VRU) and complex (simultaneous pedestrian and cyclist) scenarios.

Driver-pedestrian interactions were evaluated using some surrogate safety measurements (Minimum time-to-collision, MTTC; Post-encroachment time, PET; proportion of stopping distance, PSD) and a stopping indicator (Stop_NoStop), considering pedestrian behaviour, pedestrian time-gap acceptance (PTGA), lighting and participant demographics. PTGA emerged as the key determinant of driver response across measures: larger time gaps were associated with earlier reactions and undisturbed passages, while sudden time gaps represented the critical condition where more serious conflicts were observed. Lighting also had a consistent effect, with safer interactions at night, which is plausibly linked to increased alertness when pedestrians cross unexpectedly away from crosswalks. However, pedestrian behaviour did not materially affect PSD, suggesting that drivers' initial responses were more influenced by timing than pedestrian assertiveness. Nevertheless, pedestrian behaviour clearly impacted the likelihood of stopping, increasing when pedestrians were non-

yielding. Age-related patterns suggested safer outcomes for middle-aged drivers, consistent with a balance between experience and reaction capability.

In the shared-lane section, the results showed that crossing visibility was strongly associated with pedestrian safety. The interactions at crosswalks were generally safer than those away from them, which supports the idea that markings and signs increase predictability and encourage caution. The side from which pedestrians started crossing also was impactful on the safety. In particular, crossings from the right tended to produce safer interactions than crossings from the left, even under simple interaction conditions. Complex interactions revealed a significant safety concern, since the presence of a cyclist in the shared lane can distract drivers from pedestrians, which can have severe consequences when pedestrians cross from the left. Demographic effects were also evident, with female and middle-aged drivers generally showing safer interaction profiles. For driver-cyclist interactions, overtaking speed and minimum passing distance were largely explained by variability between drivers rather than fixed factors.

3.1.4 Practical implications

The findings highlight the need for design and management measures that reduce uncontrolled interactions and support driver attention in complex urban environments. These measures should aim to (i) increase pedestrian predictability and visibility, (ii) reduce the need for drivers to divide their attention between multiple potential hazards, and (iii) mitigate the elevated risk to cyclists observed in shared-lane segments.

On service-road sections where vehicles and cyclists are separated, the priority should be to improve pedestrian crossing provision and visibility. This could be achieved by increasing the number of crossings, improving their placement, and adopting visibility-enhancing treatments with curb extensions or lighting/LED-based enhancements (Figure).

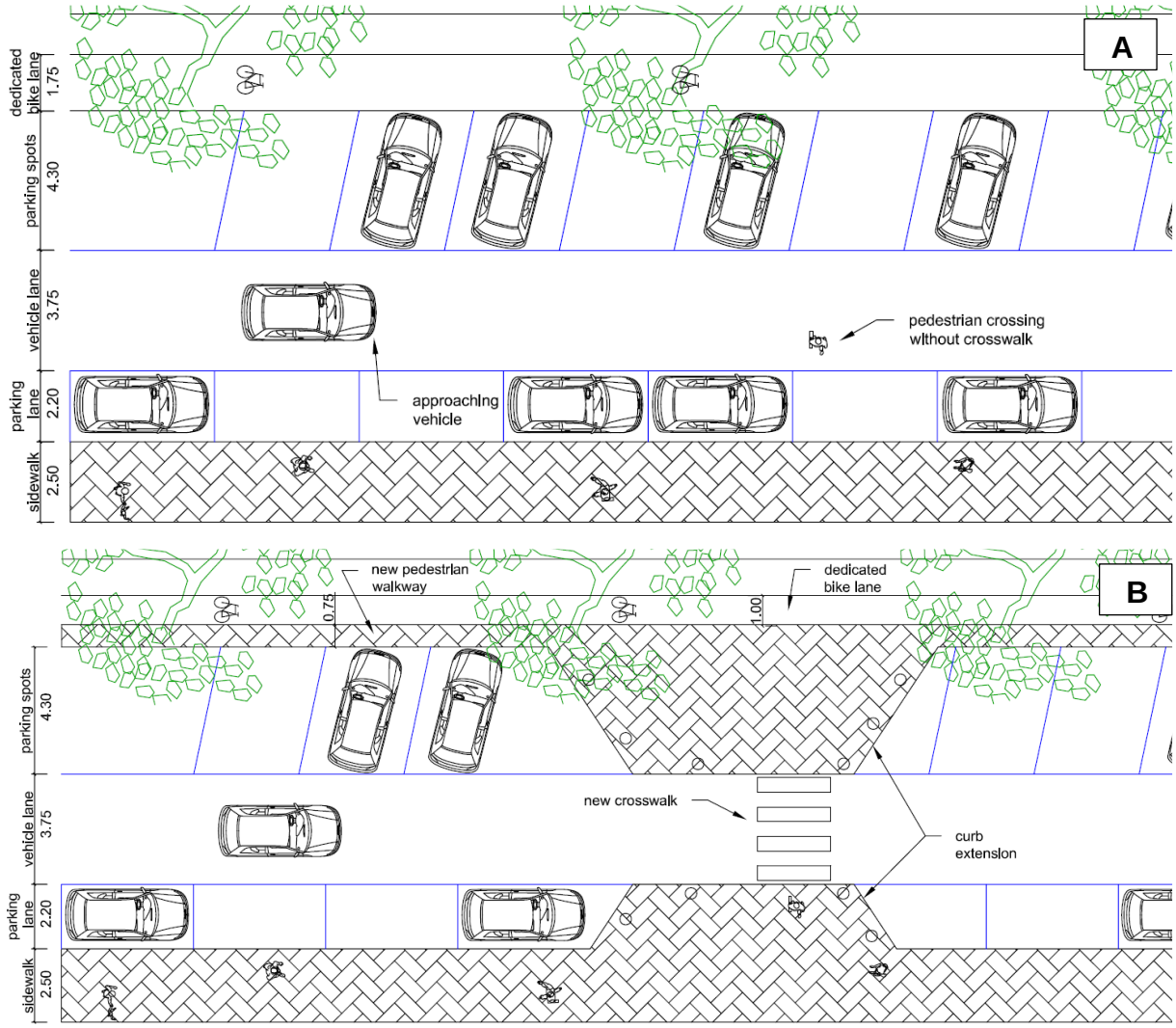


Figure 23 - Segments between *piazza Rivoli* and *piazza Adriano*: (A) actual layout without mid-block crosswalk, (B) Proposed layout with mid-block crossing with curb extension and pedestrian walkway.

On shared-lane sections, the results support stronger interventions. As well as improving the design and visibility of crosswalks, the cross-section should be reorganised to prioritise dedicated cycling facilities over shared operation with motor vehicles. Physical separation (e.g., curb-protected bike lanes) can reduce cyclist exposure and, importantly, enable drivers to focus more effectively on pedestrian crossings. Reallocating space, by converting parking spaces to create protected cycling infrastructure could also improve pedestrian sight lines (Figure).

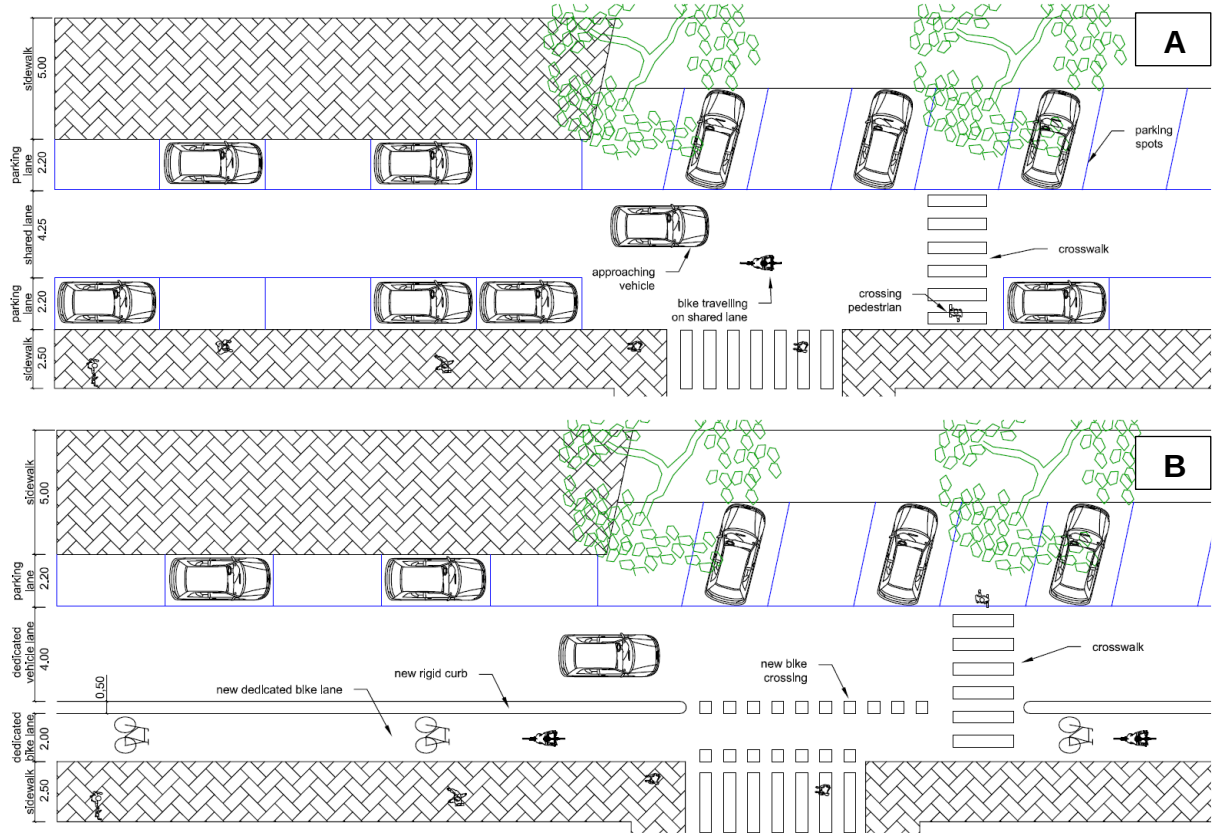


Figure 24 - Segment between the intersection *Corso Re Umberto* and *Via Sacchi*: (A) actual layout with shared lane, (B) proposed layout with dedicated bike lane with curb introduction.

3.2 Driving simulation experiments of *Corso Vittorio Emanuele II* – Corso Castelfidardo intersection

This study investigates the effectiveness of geometric treatments at urban intersections in protecting cyclists from motor vehicles performing turning manoeuvres in various lighting conditions. The *Corso Vittorio Emanuele II* corridor in Turin, Italy was selected as a case study, with a focus on the intersection with *Corso Castelfidardo* and *Corso Inghilterra*. The current road layout features shared lanes for vehicles and bicycles, as well as asymmetric geometric solutions on different sides of the intersection. The geometric treatments included measures to shorten cyclists' paths and separate them from pedestrian interactions, thereby avoiding unpredictable and hazardous conflict dynamics.

3.2.1 Results

Table summarises the descriptive statistics for the main outcome variables analysed in the manoeuvre 1.

Table 8 - Mean and standard deviation of measured variables in manoeuvre 1 according to infrastructure configuration

Variable	Infrastructure	Mean (Standard Deviation)
Speed at conflict point [km/h]	Countermeasure (CM) Real (R)	22.7 (4.96) 26.0 (4.11)
MTTC [s]	Countermeasure (CM) Real (R)	0.53 (0.13) 0.54 (0.11)
Steering Wheel [°]	Countermeasure (CM) Real (R)	65 (25.9) 94.7 (18.4)

Considering the manoeuvre 1 (according to Deliverable 3), the intersection layout showed a statistically significant effect on vehicle speed. The estimated coefficients suggest that participants drove faster in the actual configuration than in the countermeasure scenario (Figure).

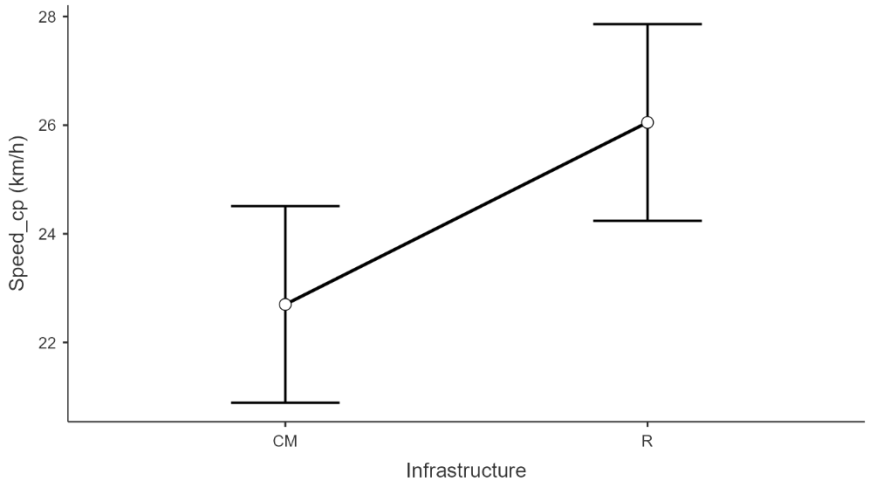


Figure 25 - Estimated marginal means of the conflict point speed (Speed_cp) according to the infrastructure configuration (CM = countermeasure, R = real) in manoeuvre 1. Error bars indicate 95% confidence intervals.

In contrast, neither lighting nor cyclist behaviour had a significant influence on the speed adopted by drivers. This suggests that drivers' speed choices for this manoeuvre were primarily affected by their perception of the intersection geometry, rather than by changes in visibility or the cyclist's speed profile.

The steering wheel angle was used as an indicator of lateral vehicle control during the turning manoeuvre. In the actual intersection layout, drivers applied substantially larger steering inputs than in the countermeasure scenario (Figure). This pattern suggests that the countermeasure, by encouraging lower approach speeds, enabled drivers to negotiate the turn more smoothly, reducing the need for sudden or aggressive steering corrections.

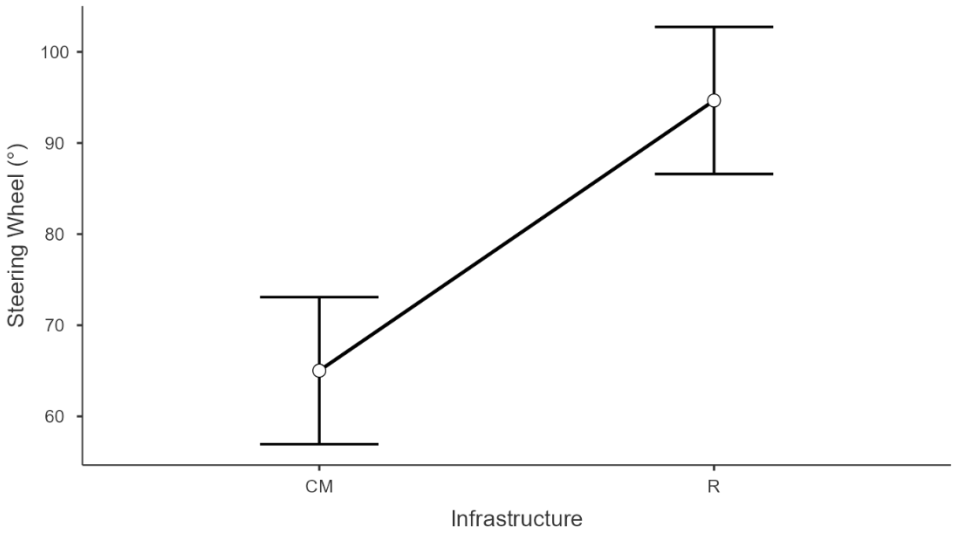


Figure 26 - Estimated marginal means of the steering wheel angle according to the infrastructure configuration (CM = countermeasure, R = real) in manoeuvre 1. Error bars indicate 95% confidence intervals.

In the manoeuvre 2 (according to Deliverable 3), the mean speed of drivers at the conflict point appears comparable in the two scenarios, suggesting that the roadway geometry does not directly affect the driver's approaching speed (Table).

Table 9 - Mean and standard deviation of measured variables in manoeuvre 2 according to infrastructure configuration

Variable	Infrastructure	Mean (Standard Deviation)
Speed at conflict point [km/h]	Countermeasure (CM)	23.2 (5.71)
	Real (R)	23.7 (5.13)
MTTC [s]	Countermeasure (CM)	0.80 (0.28)
	Real (R)	0.77 (0.39)
Smax - Sconflict [km/h]	Countermeasure (CM)	2.64 (4.17)
	Real (R)	3.29 (5.7)

Figure shows that the MTCC decreases substantially when the cyclist follows an aggressive speed profile.

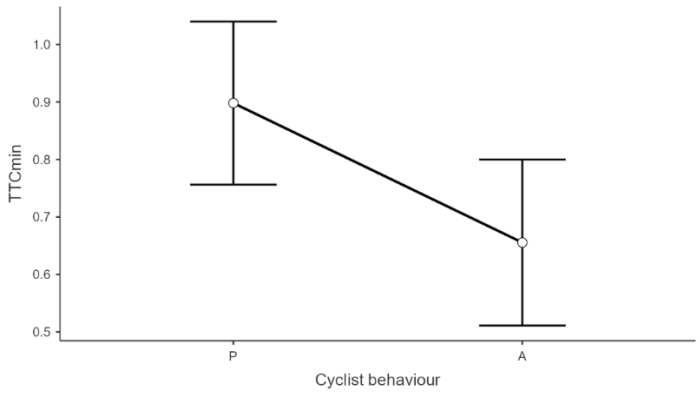


Figure 27 - Estimated marginal means of the MTTC according to the cyclist behaviour (P = prudent, A = aggressive) in manoeuvre 2. Error bars indicate 95% confidence intervals.

The extent of deceleration was examined by calculating the difference between the maximum speed achieved when crossing the intersection (Smax) and the speed recorded at the conflict point (Sconflict). This speed reduction was not significantly affected by any of the experimental factors. Overall, the results suggest that drivers tend to slow down in a consistent manner during the turning manoeuvre, with an average reduction of 2.6-3.3 km/h, as reported in Table . This reduction occurs regardless of the road layout or the cyclist speed profile.

In the manoeuvre 3, the participant drives along the side road while a cyclist takes a shortcut through the parking lane and enters the vehicle lane very close to the ego vehicle, instead of following the chicane alignment. Table reports the descriptive statistics of the variables collected for this longitudinal interaction. In the baseline layout, drivers showed a pronounced defensive reaction, with a significant deceleration peak (mean -1.76 m/s²) and relatively low speed at the conflict point (32.7 km/h). This is consistent with the event being perceived as sudden and potentially dangerous. In contrast, under the countermeasure configuration, the deceleration peak was much smaller (mean -0.40 m/s²). In terms of speed at conflict points, drivers that travelled significantly faster in the countermeasure condition than in the real layout. Considering also the lower deceleration peak of Table , this finding supports the idea that the countermeasure reduces the "surprise effect" of cyclists suddenly entering the lane, enabling drivers to respond more smoothly.

Table 10 - Mean and standard deviation of measured variables in manoeuvre 3 according to infrastructure configuration

Variable	Infrastructure	Mean (Standard Deviation)
Speed at conflict point [km/h]	Countermeasure (CM)	37.9 (8.56)
	Real (R)	32.7 (10.7)
TTCmin [s]	Countermeasure (CM)	1.17 (1.01)
	Real (R)	1.17 (1.10)
DecPeak at conflict point [m/s²]	Countermeasure (CM)	-0.40 (1.55)
	Real (R)	-1.76 (3.04)

When considering MTTC, while the countermeasure results in a smoother driving response with less sudden braking, it does not measurably alter the MTTC compared to the baseline layout. Figure indicates that MTTC was lower when drivers interacted with the prudent cyclist than with the aggressive cyclist. This outcome aligns with the behavioural pattern observed for speed at conflict point. When the cyclist is aggressive, the manoeuvre is carried out quickly and the driver tends to maintain a steady speed. This means that the vehicle reaches the conflict area after the cyclist has already left it, resulting in a larger MTTC. Conversely, when the cyclist is more cautious, the manoeuvre is slower and the vehicle tends to remain closer to the cyclist during the interaction, resulting in a smaller MTTC.

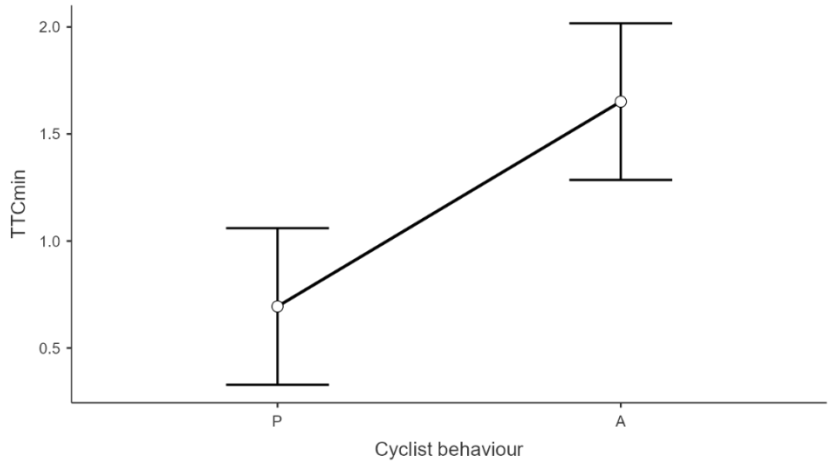


Figure 28 - Estimated marginal means of the MTTC according to the cyclist behaviour (P = prudent, A = aggressive) during manoeuvre 3. Error bars indicate 95% confidence intervals.

Figure shows that the effectiveness of the countermeasure in limiting sudden braking is significantly affected by lighting conditions for manoeuvre 3. In daytime conditions, the two road layouts produce markedly different responses: in the real configuration, drivers decelerate more sharply (mean deceleration of approximately -1.8 m/s²), whereas in the redesigned layout, deceleration is almost non-existent. In contrast, during night-time driving, the distinction between the two configurations disappears, with deceleration values close to zero in both cases. This suggests that reduced visibility may attenuate the geometric effect of the redesign, leading drivers to adopt a more cautious and consistent strategy, regardless of the layout.

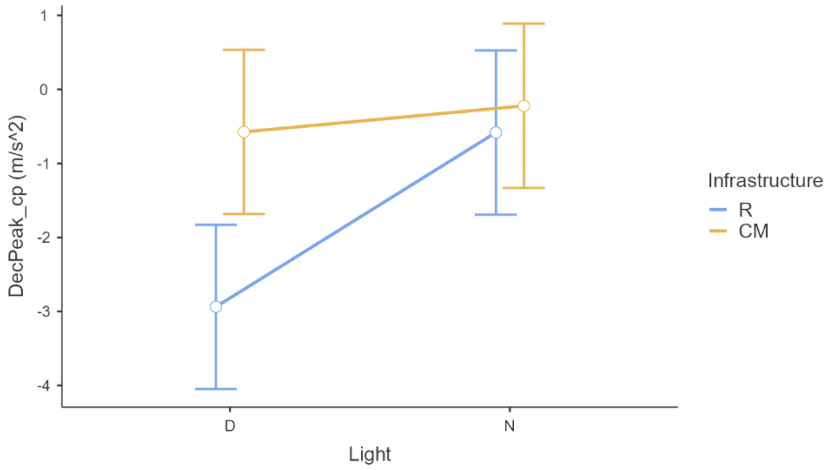


Figure 29 - Interaction effect between lighting conditions (D = day, N = night) and infrastructure configuration (R =real, CM = countermeasure) on *DecPeak_cp* during manoeuvre 3. Error bars indicate 95% confidence intervals.

3.2.2 Conclusion

This experiment evaluated the safety of interactions between cyclists and vehicles at the intersection of *Corso Vittorio Emanuele II*, *Corso Castelfidardo* and *Corso Inghilterra* in Turin. The current configuration of the intersection is characterized by a geometric asymmetry of the bicycle paths. Moreover, observations indicate cyclists frequently disregard the chicane layouts which were originally intended to moderate speeds, and instead follow straighter trajectories. This mismatch between the intended design and actual behaviour creates unpredictable and potentially hazardous interaction patterns.

The redesigned configuration was based on clearer flow separation and geometric regularisation across the entire intersection. The driver behaviour in the interaction with a cyclist was analysed in three representative conflict manoeuvres: two left-turn manoeuvres located at the intersections with *Corso Castelfidardo* and *Corso Inghilterra*, and one longitudinal interaction along the side road of Corso Vittorio Emanuele II.

The main findings show that drivers' responses to potential conflicts with cyclists are strongly influenced by the clarity of the road layout and the predictability of cyclists' movements. In both left-turn scenarios, the geometry of the infrastructure emerged as the dominant factor in the safer behaviour or vice versa. Removing the road geometric ambiguity in the manoeuvre from *Corso Castelfidardo* towards the main carriageway of *Corso Vittorio Emanuele II* (manoeuvre 1) led to a substantial reduction in steering input, indicating a smoother and more controlled left-turning action. In the left-turning manoeuvre from *Corso Inghilterra* to the service road of *Corso Vittorio Emanuele II* (manoeuvre 2), the introduction of protective curbs and a clearly defined conflict area had a safer effect in the vehicle-cyclist interaction. In the longitudinal interaction scenario (manoeuvre 3), the results emphasised the importance of visibility and lane separation together. When a dedicated cycle lane was provided, drivers managed the interaction more smoothly. Although average speeds were higher than in the baseline condition, sudden vehicle decelerations were largely eliminated.

Among the results, it is interesting to note that the slower, more hesitant cyclist produced lower TTC values than the more assertive cyclist. This is consistent with previous research showing that driver decisions are strongly influenced by the cyclist's kinematic profile. Slow and uncertain crossing prolongs exposure within the conflict area and tends to increase the cognitive demand of drivers, delaying their response. Consequently, braking is initiated later, making the minimum temporal safety margin more critical. In contrast, a faster, more assertive cyclist clears the conflict area more quickly, often resulting in a larger temporal gap when the vehicle reaches the interaction point.

Moreover, night-time driving emerged as a transversal risk factor across scenarios. Under low-light conditions, drivers exhibited reduced anticipation capacity and shorter braking distances in response to sudden events, regardless of the infrastructure layout. This pattern aligns with established evidence that

reduced visibility impairs perception of speed and distance, leading to delayed reactions and sudden manoeuvres.

3.2.3 Practical implications

The experiment findings highlight the importance of moving beyond shared-lane arrangements on major roads. Dedicated and protected cycling infrastructure not only protects cyclists but also reduces uncertainty for drivers. The position of the vehicle stop line in the intersection is also considered important. Moving it closer to the conflict zone reduces the distance available for acceleration of vehicles during green phases and encourages lower approaching speeds when entering the interaction area. It is equally crucial to ensure that trajectories at intersections are symmetrical and predictable. The existing chicane of the cycling path is ineffective because cyclists often bypass it, creating unforeseen conflict points with vehicles. A straight, protected alignment better aligns actual cyclist behaviour with driver expectations, thus lowering collision risk.

4 Results RU UniPD

4.1 Behavioral data

4.1.1 Descriptive Statistics

The following section will present graphical descriptive statistics of total and subtotal sum scores from the trait- and state-questionnaire administered throughout the screening and experimental phases. Each section is also provided with a short description for improving clarity. All the descriptives are presented controlling by participants' biological sex and, when required, by experimental phase (Control vs Experimental) and daytime (Day vs Night).

4.1.1.1 Questionnaires – Scores distribution

4.1.1.1.1 Driving Behavior Questionnaire (DBQ)

Figure 30 displays the distribution of total violation and error scores from the Driver Behavior Questionnaire by biological sex. Males show slightly higher median scores for violations (26 vs 20 over 55), while females and males exhibit similar median scores for errors (48 vs 42 respectively over 80), with considerable variability in both groups. The four panels in Figure 31 illustrate DBQ subscales organized into two main categories. Violations (aggressive AV and ordinary OV) show higher medians in males compared to females, while for errors (ER) and lapses (LA), median values are comparable between sexes, with females showing slightly higher medians in lapses (20 vs 22 over 40) and similar values in errors.

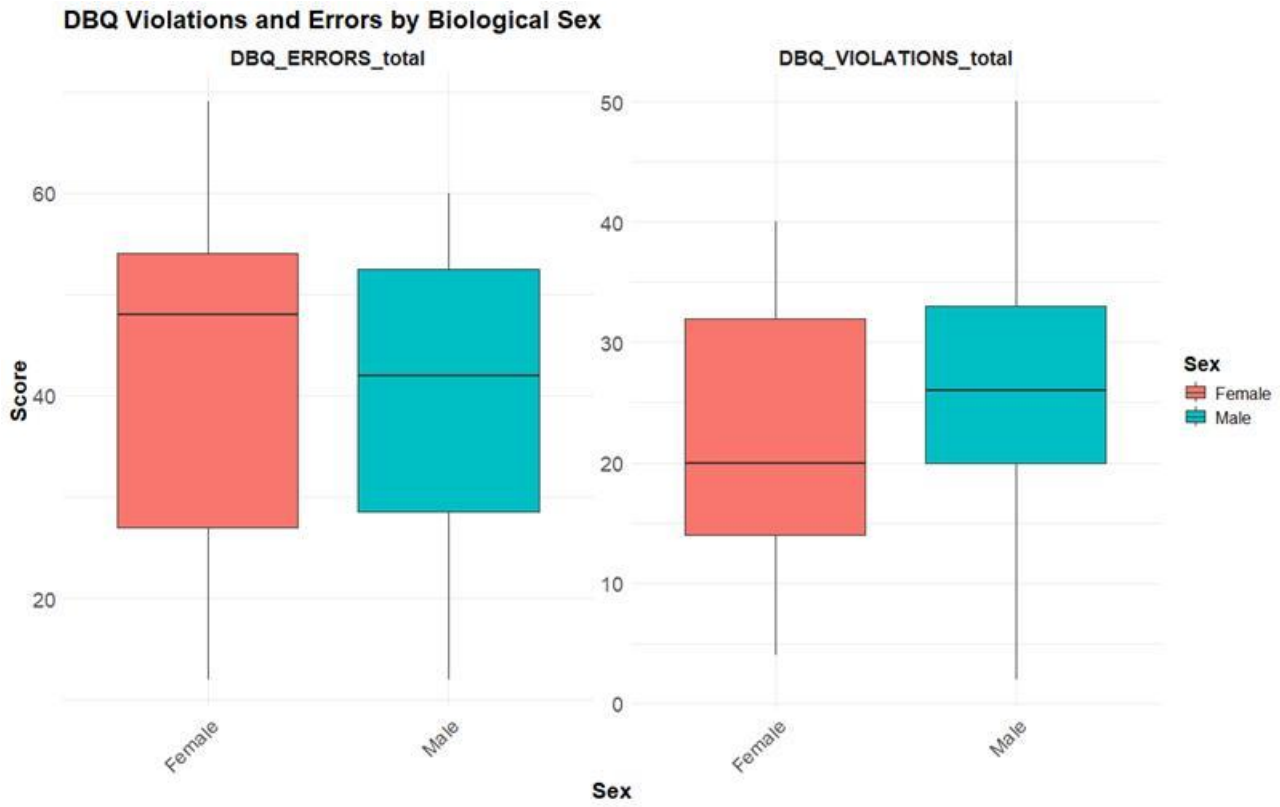


Figure 30 - Boxplot showing the distribution of total Violations and Errors scores from the DBQ questionnaire.

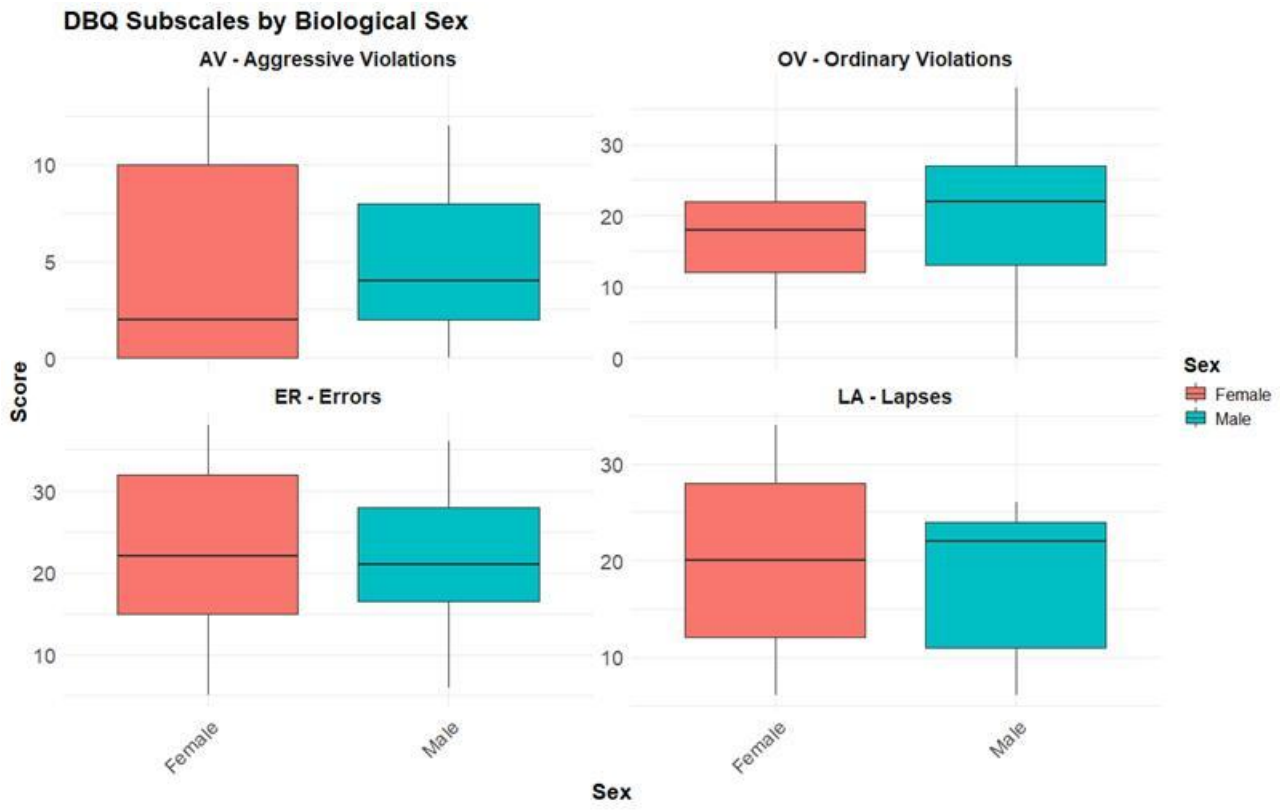


Figure 31 - Boxplot showing the distribution of Violations (Aggressive and Ordinary) and Errors (Errors and Lapses) subscales scores from the DBQ questionnaire

4.1.1.1.2 BARRAT Impulsiveness Scale (BIS)

The three panels in Figure 32 displays the Barratt Impulsiveness Scale subscales by biological sex. Attention (as the cognitive instability and difficulty maintaining focus) and Motor subscales (as the tendency to act without thinking) show nearly identical median scores between females and males (15 over 28 and 23 over 48, respectively). Non-planning (as the lack of future orientation and planning) shows females with slightly higher median scores (26 vs 24), though all three subscales demonstrate considerable overlap between sexes with similar distributions.

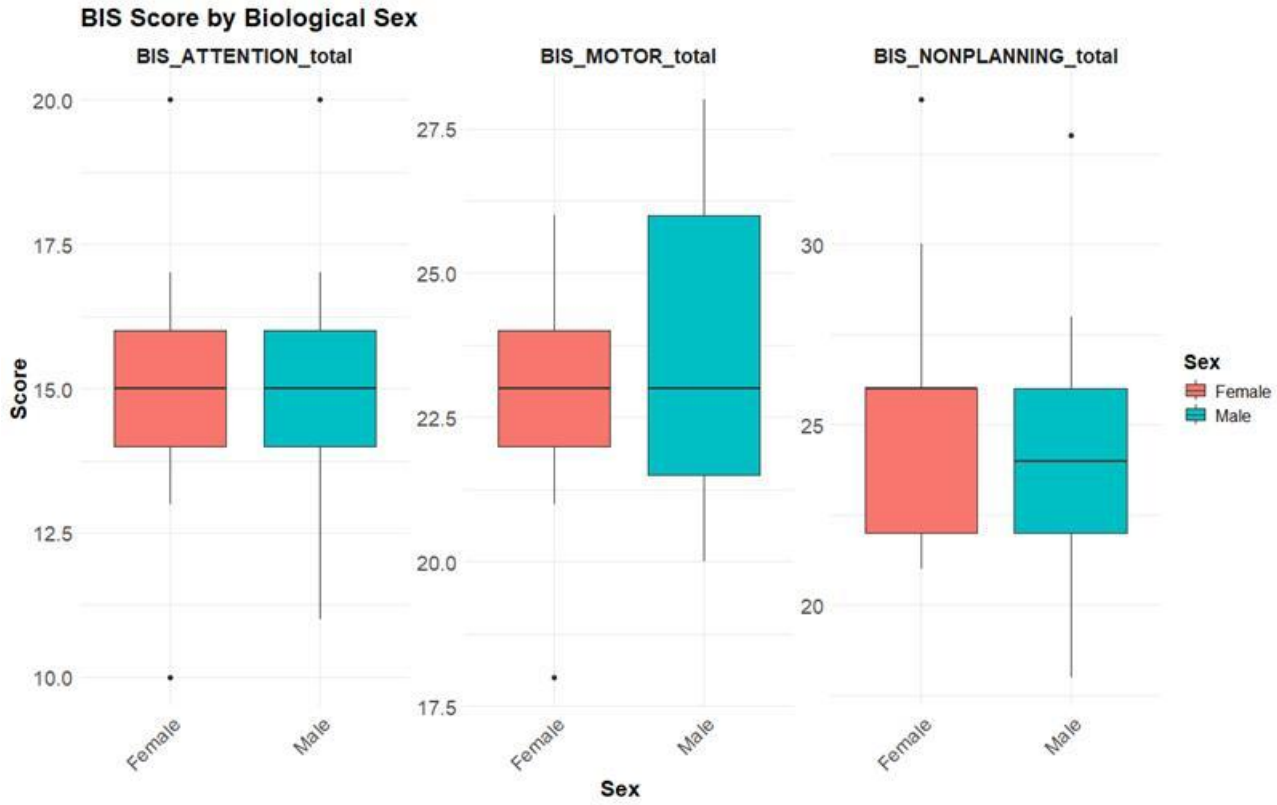


Figure 32 - Boxplot showing the distribution of Attention, Motor and Non-Planning BIS subscales.

4.1.1.1.3 Mind Reading Belief Questionnaire

Figure 33 presents the Mind-Reading Belief Scale scores by biological sex. This scale represents the individual beliefs about the ability to infer or "read" other people's mental states. Males show a similar median score (17) compared to females (16), with both groups displaying similar interquartile ranges and comparable variability, suggesting minimal sex differences in beliefs about mind-reading abilities as measured by this scale.

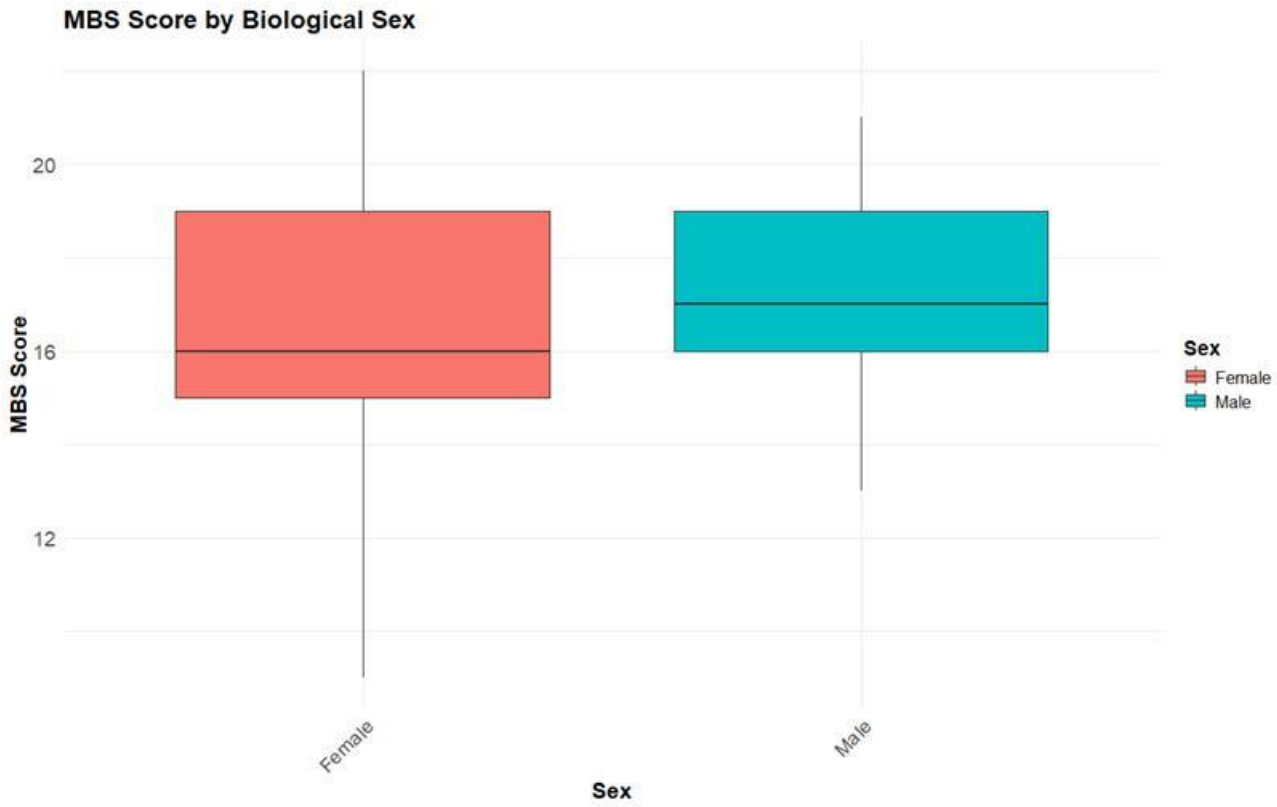


Figure 33 - Boxplot showing the distribution of Mind Reading Belief Scale total score.

4.1.1.1.4 Depression, Anxiety and Stress (DASS) – State Questionnaire

The four panels in Figure 34 display Depression, Anxiety, and Stress Scale subscales compared between biological sex and time of day (Day vs Night). Anxiety and depression show minimal differences between sexes and Simulation Day times, with median scores around 3-4 for anxiety and 2-3 for depression (over 21). The stress subscale shows slightly higher median scores (7-8) with females maintaining consistent levels across day and night, while males show slightly lower stress before night driving. The total DASS score reveals similar patterns with medians around 13-15 (over 63), indicating overall comparable and trivial psychological distress levels across sex and simulation day-time, though individual variability remains considerable as shown by the wide interquartile ranges.

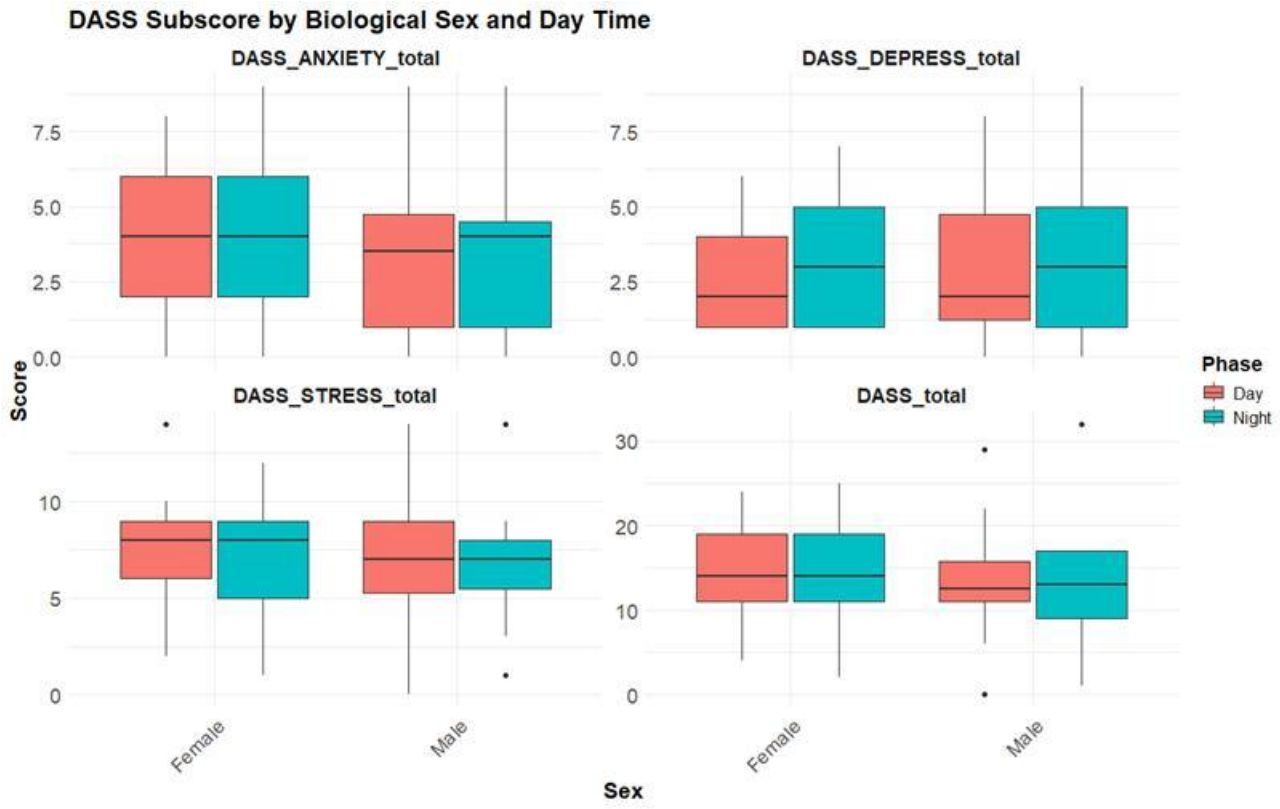


Figure 34 - Boxplots showing the distribution of DASS-21 subscales (Anxiety, Depression, Stress) and total scores by biological sex and time of day (Day vs Night).

4.1.1.1.5 Presence Questionnaire

The three panels in Figure 35 display the iGroup Presence Questionnaire subscales comparing biological sex and simulation daytime. Involvement shows similar median scores between sexes (14-15 over 28) with females displaying slightly higher values during daytime and males at nighttime. Spatial Presence reveals the most consistent pattern with medians around 20-21 across all groups (over 35), indicating stable spatial immersion regardless of sex or testing time. Realism demonstrates some differences, with females showing higher median scores during daytime (14 over 28) compared to nighttime (11), while males maintain more consistent realism perception across both testing phases (12-13).

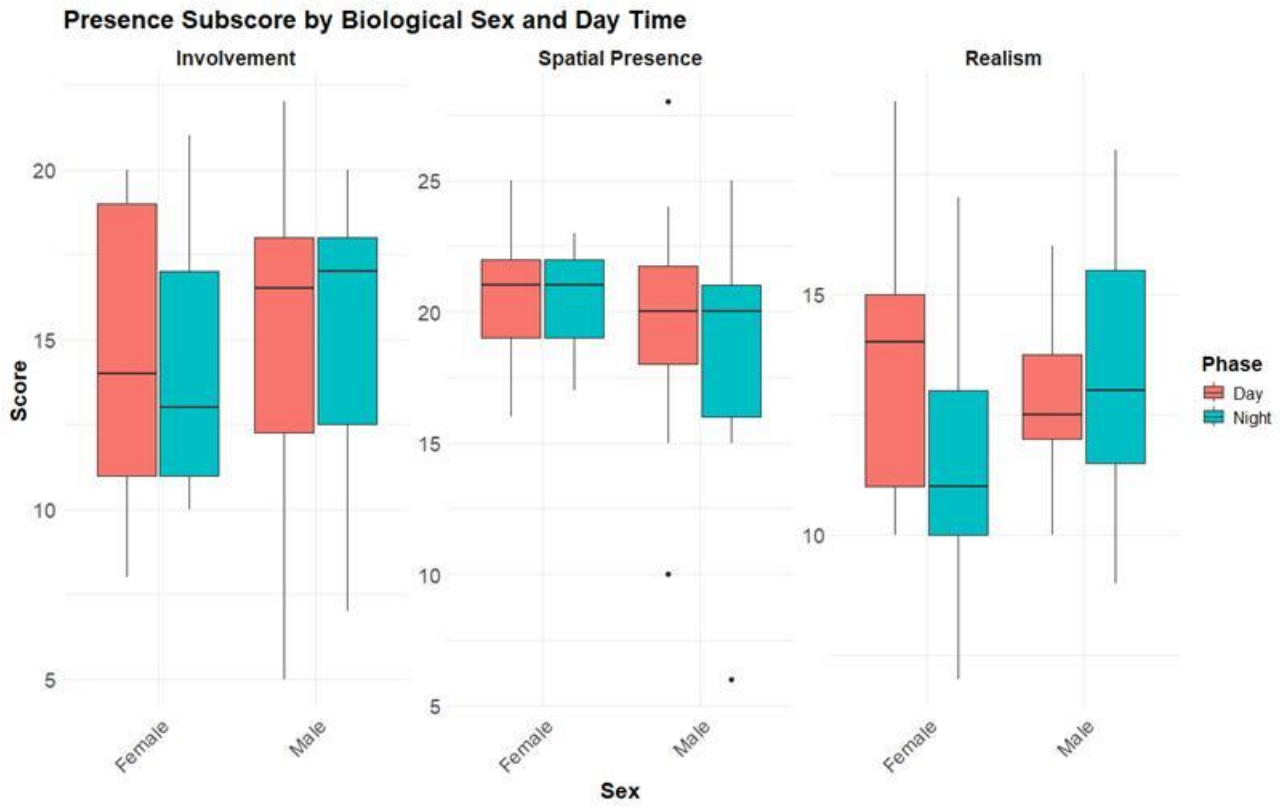


Figure 35 - Boxplots showing the distribution of Presence subscales (Involvement, Spatial Presence, Realism) by biological sex and time of day (Day vs Night).

4.1.1.1.6 Simulation Sickness Scale (SSQ)

The two panels in Figure 36 illustrate SSQ total scores evolution across three experimental timepoints (0 = baseline before driving, 1 = mid-session after the first drive, 2 = post-session after the second drive) separated by sex and daytime. The possible score range in this case was [16-64]. Females show markedly different patterns between day and night sessions: during daytime, scores increase from baseline (18.2) to mid-session (20.0) and remain at this level (20.2), while nighttime scores start slightly higher (20.8) and remain relatively stable (21.0). Males demonstrate more consistent patterns across both testing phases, with daytime scores showing a gradual increase from 18.1 to 19.3, and nighttime scores remaining stable around 18.0-18.7, suggesting females may be more susceptible to simulation sickness during daytime exposure while males maintain slightly lower and more stable symptom levels regardless of testing time. Importantly, these low scores confirm a non-significant impact of cybersickness on simulation activity.

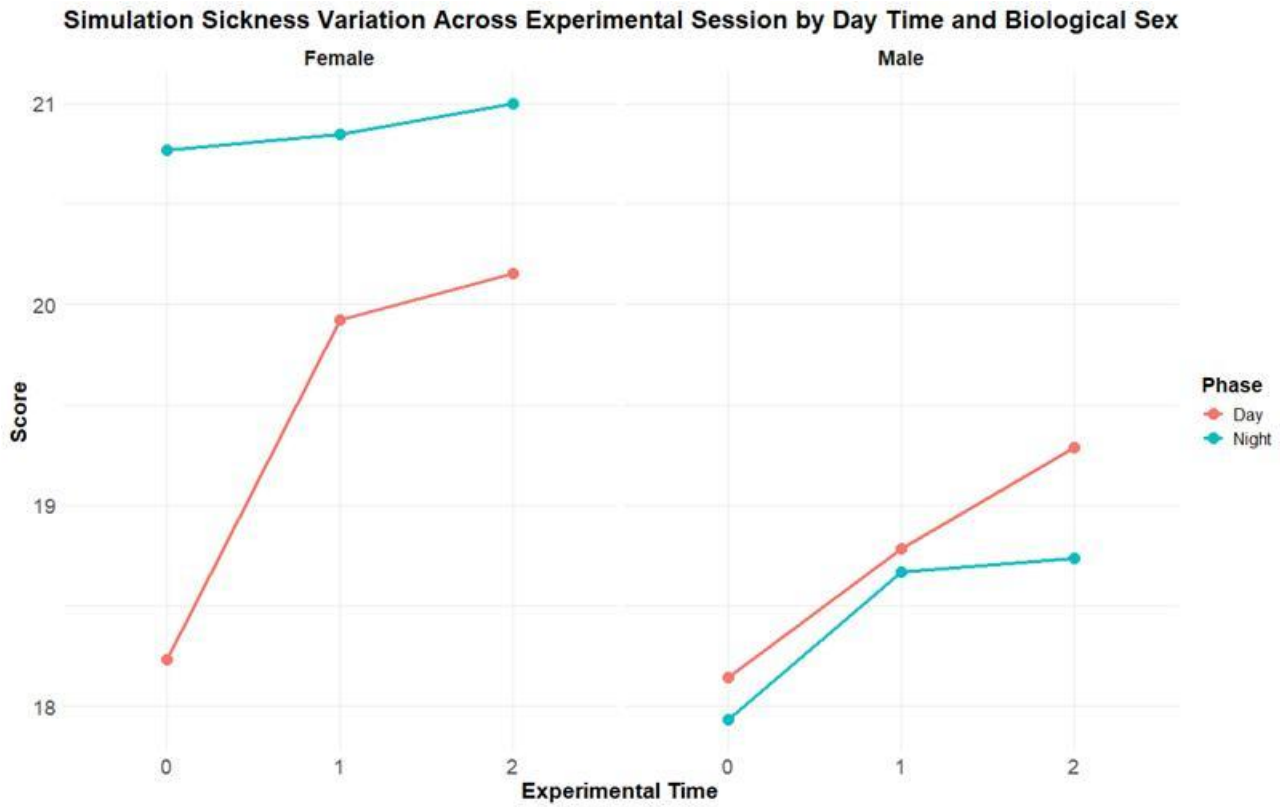


Figure 36 - Line plots showing mean SSQ total scores across three experimental timepoints (0=baseline, 1=mid-session, 2=post-session) by biological sex and time of day (Day=red, Night=blue). The left panel displays female participants' trajectories, while the right panel shows male participants' progression of cybersickness symptoms throughout the virtual reality exposure.

4.1.1.1.7 NASA-TLX (Cognitive Workload)

The two panels in Figure 37 show total workload scores across two driving blocks separated by sex and testing phase. The possible score range in this case was [6-126]. Females display contrasting patterns: during daytime, total workload decreases from 28.3 to 26.5, while nighttime sessions show a reduction from 32.6 to 32.2, maintaining higher overall workload levels. Males demonstrate opposite trends, with daytime workload declining from 27.7 to 24.6, while nighttime workload slightly increases from 27.2 to 27.8, suggesting potential sex-dependent effects on perceived task demands with females experiencing consistently higher workload during nighttime exposure.

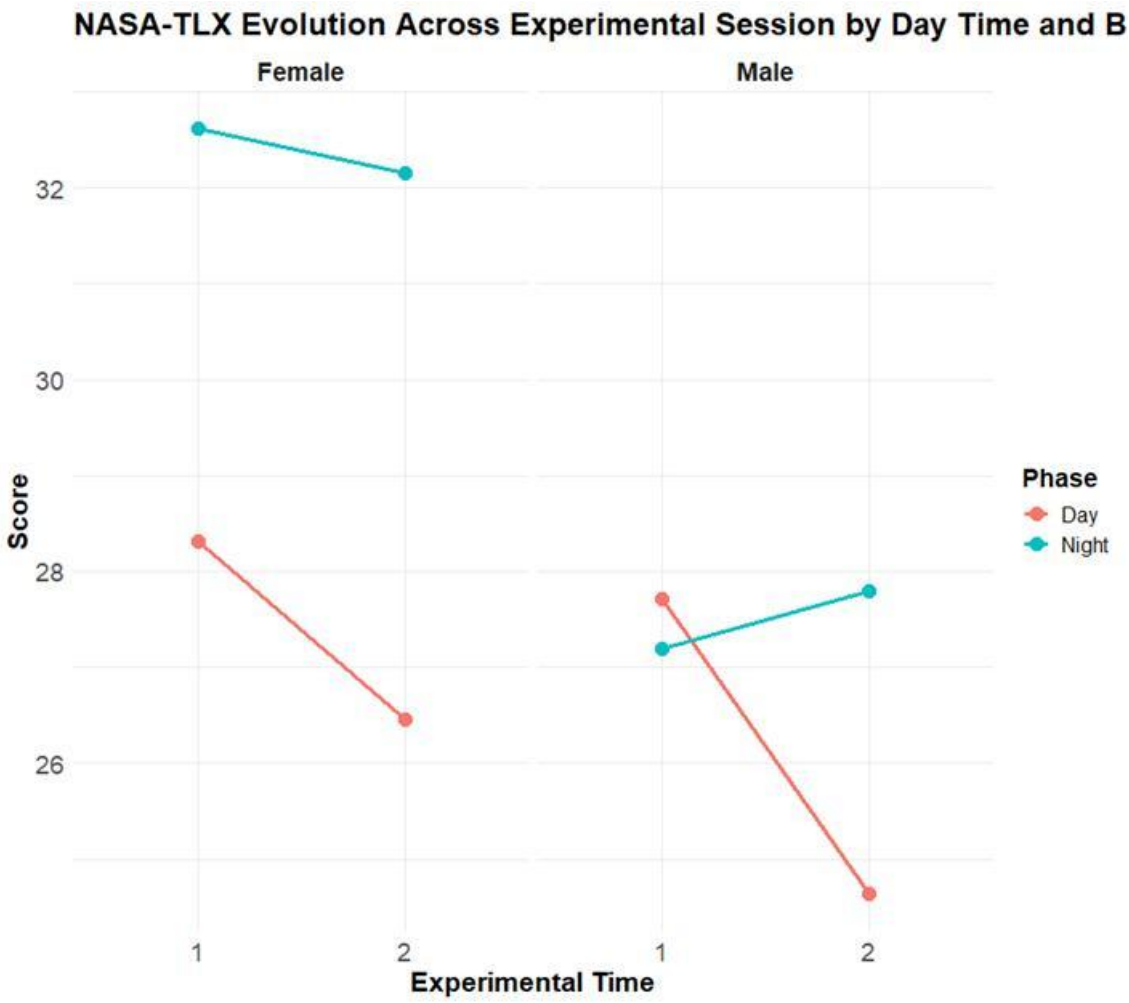


Figure 37 - Line plots showing mean NASA-TLX total scores across two experimental blocks (1=first driving session, 2=second driving session) by biological sex and time of day (Day=red, Night=blue). The left panel displays female participants' workload trajectories, while the right panel shows male participants' perceived workload evolution throughout the driving simulation.

The six panels in Figure 38 display Mental Demand, Physical Demand, and Temporal Demand trajectories across two blocks. The possible score range in these cases was [1-21]. Mental Demand shows females maintaining stable scores during the day (5.7-5.8) while decreasing at night (7.9-6.8), whereas males show consistent patterns across both phases (5.3-5.5). Physical Demand remains low and stable across all groups (2.7-5.2), with females showing slightly higher nighttime scores. Temporal Demand displays the most variability, with females showing minimal changes during daytime but slight increases at night, while males demonstrate increasing nighttime demands (2.9-3.6), indicating differential temporal pressure perception across sex and simulation day times.

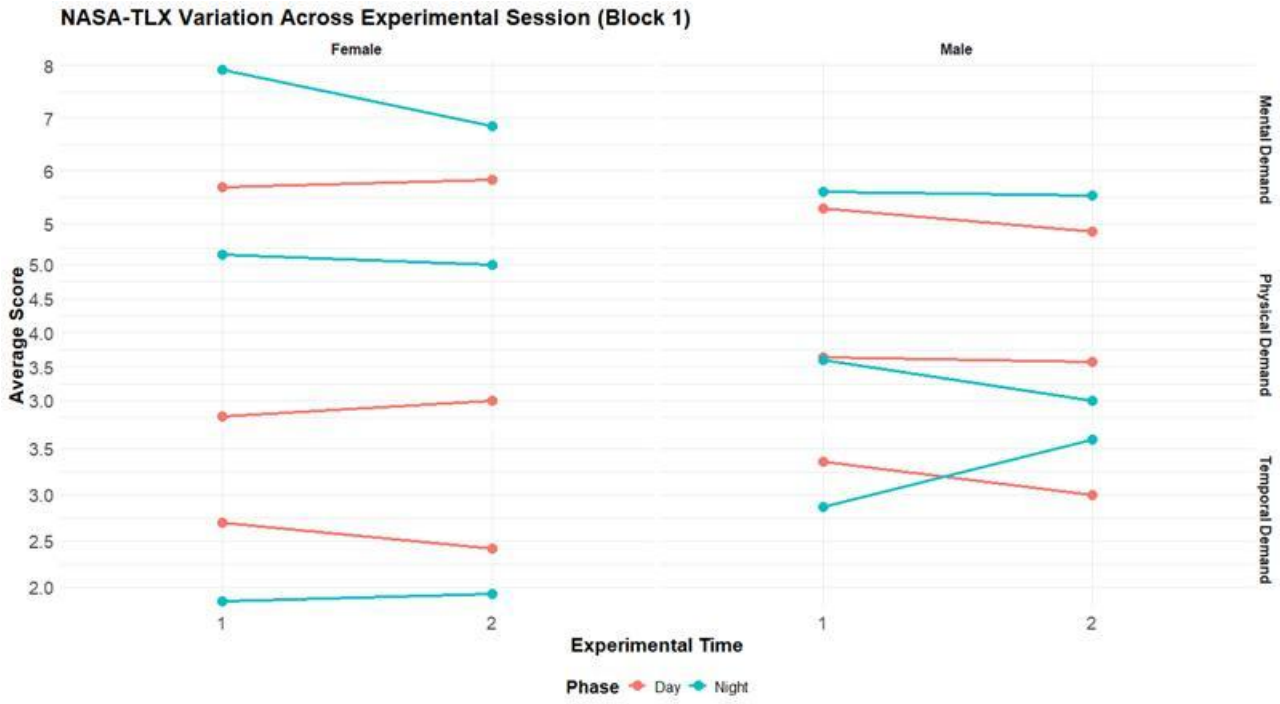


Figure 38 - Line plots showing mean NASA-TLX subscale scores (Mental Demand, Physical Demand, Temporal Demand) across two experimental blocks by biological sex and time of day (Day=coral, Night=teal). Each row represents a different workload dimension, with female trajectories on the left and male trajectories on the right panel.

The six panels in Figure 10 illustrate Effort, Frustration, and Performance trajectories across blocks. Again, the possible score range in these cases was [1-21]. Effort shows females maintaining relatively stable high scores across both phases (6.3-7.6), while males display consistent moderate levels (4.8-5.7). Frustration remains low across all groups (2.8-4.1) with minimal variation between phases. Performance ratings reveal interesting patterns: females show stable daytime performance (6.1-6.2) but increasing nighttime performance concerns (6.5-7.2), while males maintain consistent self-assessed performance across both conditions (5.6-6.2), suggesting females may experience slightly greater performance anxiety during night simulation driving despite maintaining effort levels.

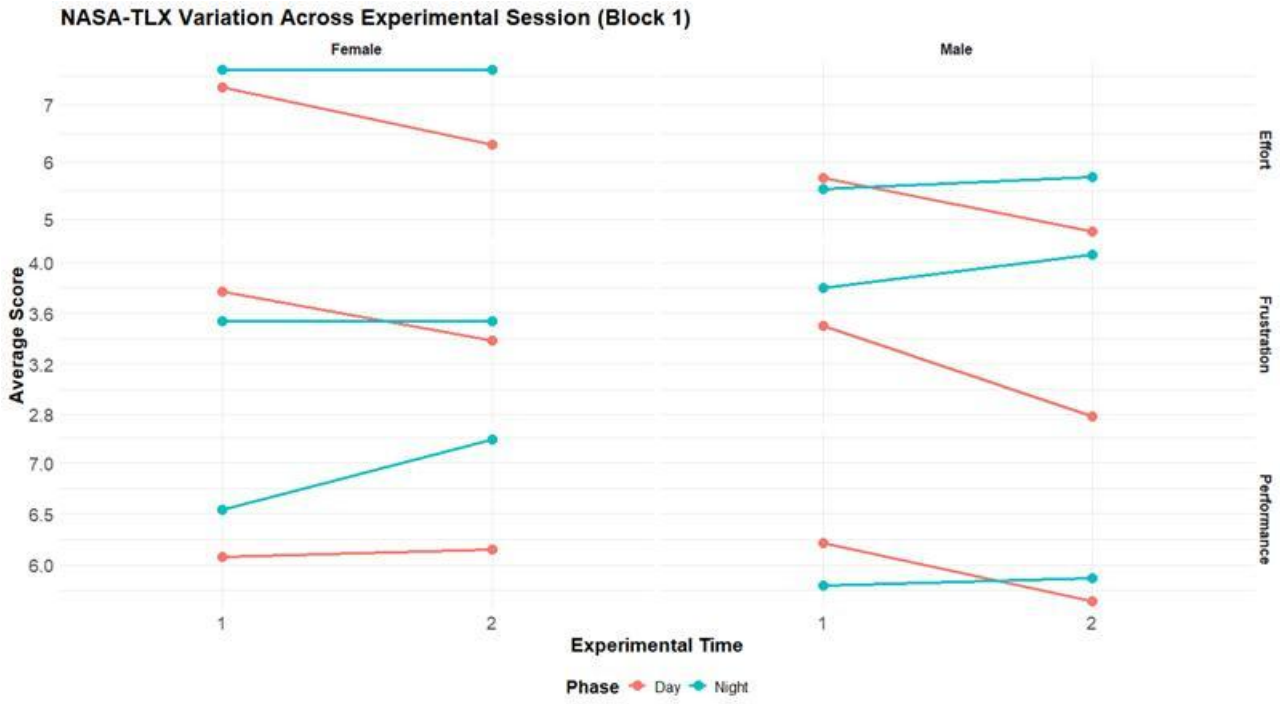


Figure 39 - Line plots showing mean NASA-TLX subscale scores (Effort, Frustration, Performance) across two experimental blocks by biological sex and time of day (Day=coral, Night=teal). Each row represents a different workload dimension, with female trajectories on the left and male trajectories on the right panel, completing the comprehensive NASA-TLX assessment.

4.1.1.2 Correlation Matrix

The correlation matrices reveal complex relationships among driver behavior characteristics, psychological states, workload, and simulation-related measures across all participants and separated by biological sex. The diagonal dark blue elements represent perfect self-correlations (r=1.0), while off-diagonal elements display varying strengths of associations using a blue-white-red color scheme, where darker blues indicate stronger positive correlations and reds indicate negative correlations. It is important to interpret these findings with caution given the modest sample sizes: N=28 for the overall sample, N=13 for females, and N=15 for males. The sex-stratified analyses, in particular, have limited statistical power and individual correlations may be unstable, though patterns of associations can provide preliminary insights into potential sex-specific relationships warranting further investigation in larger samples.

Overall Sample Patterns: In the complete sample, strong positive correlations emerge within construct families, as expected. DBQ subscales show moderate to strong intercorrelations, with Aggressive Violations and Ordinary Violations highly correlated (r=0.39), and Errors and Lapses strongly linked (r=0.80). Within the BIS, Attention and Motor impulsivity demonstrate a moderate association (r=0.43), while Non-Planning shows weaker relationships with other dimensions. DASS subscales display exceptional temporal stability between day and night assessments (Stress: r=0.81; Depression: r=0.60; Anxiety: r=0.75), indicating consistent psychological distress patterns. SSQ scores show strong positive correlations across experimental phases

(r=0.41-0.70 between consecutive measurements), reflecting stable individual susceptibility to simulation sickness. NASA-TLX measures demonstrate very high correlations between consecutive phases (r=0.90 between fase1 and fase2, both day and night), with moderate cross-time stability (day-night correlations: r=0.71-0.90).

Sex-Specific Differences: Comparing female and male correlation matrices reveals interesting divergences. Females exhibit substantially stronger intercorrelations among DBQ subscales: AV-OV (r=0.64 vs r=0.23 in males), suggesting more cohesive violation patterns in female drivers. The female matrix shows interesting negative correlations between DBQ violations and spatial presence at night (r=-0.80 for AV, r=-0.49 for OV), indicating that higher presence substantially reduces risky driving tendencies in females—a relationship nearly absent in males (r=0.22 and r=0.41 respectively). Males display stronger coupling between psychological distress and impulsivity: DASS Depression-BIS Attention (r=0.69 vs r=0.27 in females) and DASS Stress-BIS Motor (r=0.61 vs r=0.02 in females), suggesting a particular relationship between psychological state and impulsive tendencies in male participants. DBQ Violations show weak positive correlations with BIS Motor (r=0.06) and Non-Planning impulsivity (r=0.35) overall, with males showing stronger BIS-DBQ associations than females. DASS measures correlate moderately with NASA-TLX workload: Stress-NASA correlations range r=0.17-0.48, with females showing stronger associations (Day Stress-NASA fase1: r=0.41 vs r=0.22 in males), indicating psychological state more profoundly influences perceived task demands in female participants. SSQ scores correlate negatively with DBQ violations in both sexes (r=-0.33 to -0.43), suggesting individuals prone to risky driving experience less simulation sickness—potentially due to different engagement strategies. An interesting finding is the strong positive correlation between NASA-TLX across phases (r=0.90) indicating highly consistent workload perception, with moderate day-night stability (r=0.71-0.90) suggesting individual differences in perceived task demands persist across driving activities and simulation day times.

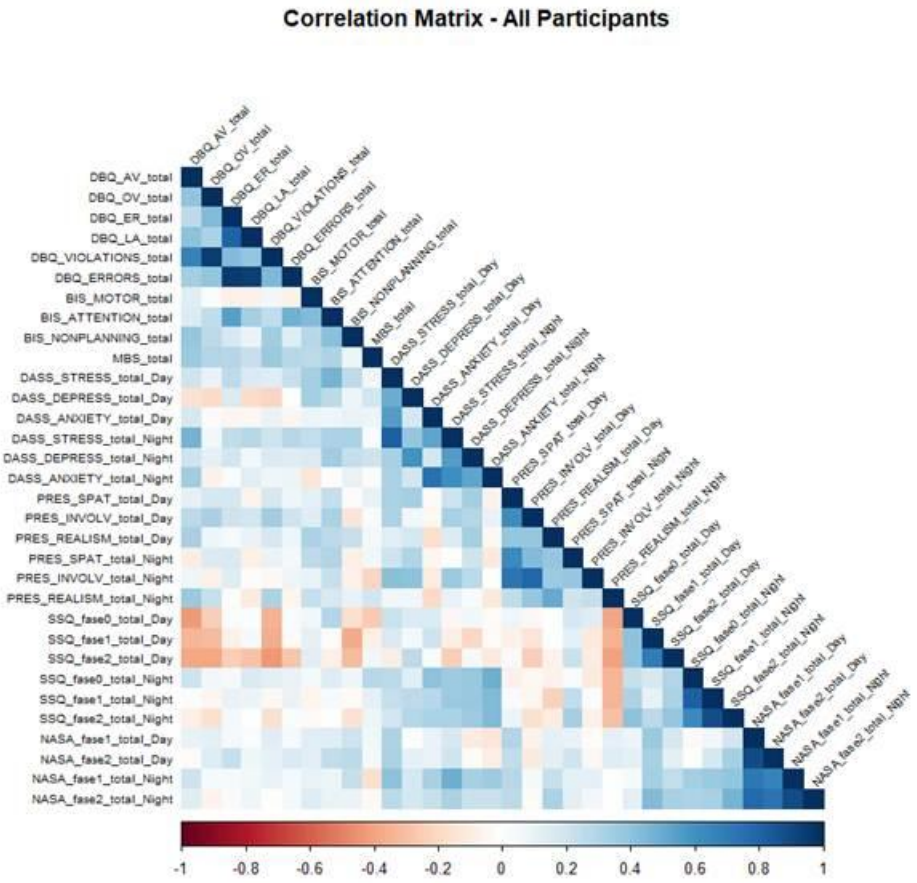


Figure 40 - Correlation matrices displaying Pearson correlation coefficients among all measured variables (complete sample N = 28). The color gradient ranges from dark red (r=-1.0, perfect negative correlation) through white (r=0, no correlation) to dark blue (r=1.0, perfect positive correlation).

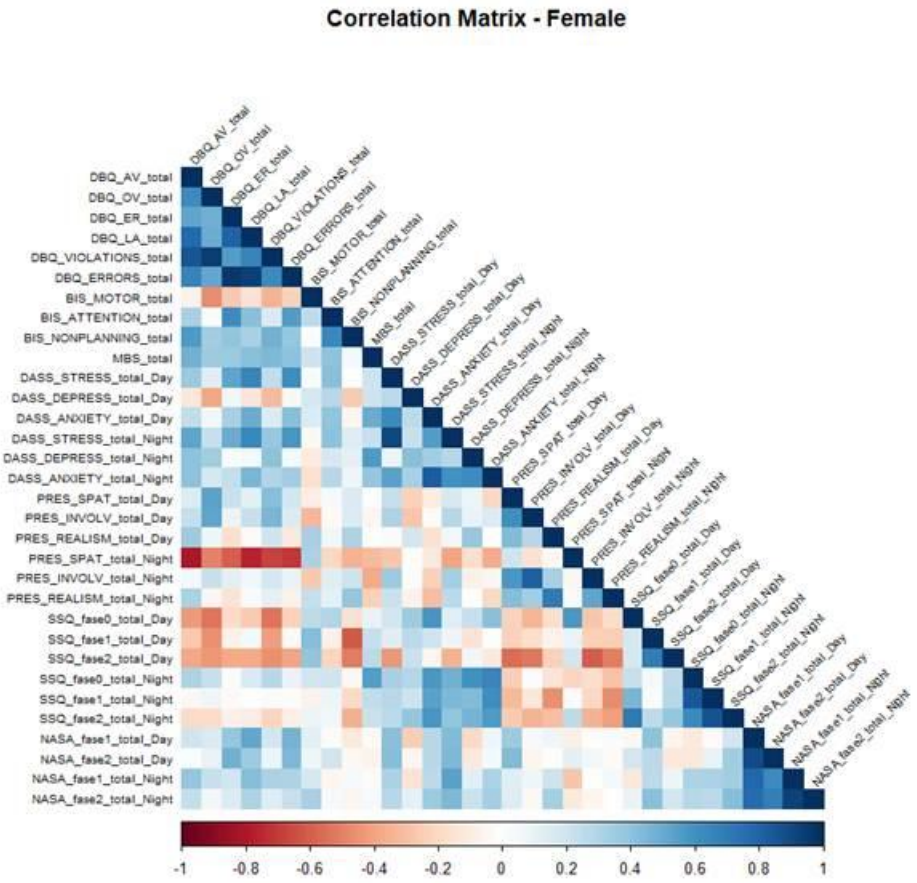


Figure 41 - Correlation matrices displaying Pearson correlation coefficients among all measured variables (female subsample N = 13). The color gradient ranges from dark red (r=-1.0, perfect negative correlation) through white (r=0, no correlation) to dark blue (r=1.0, perfect positive correlation).

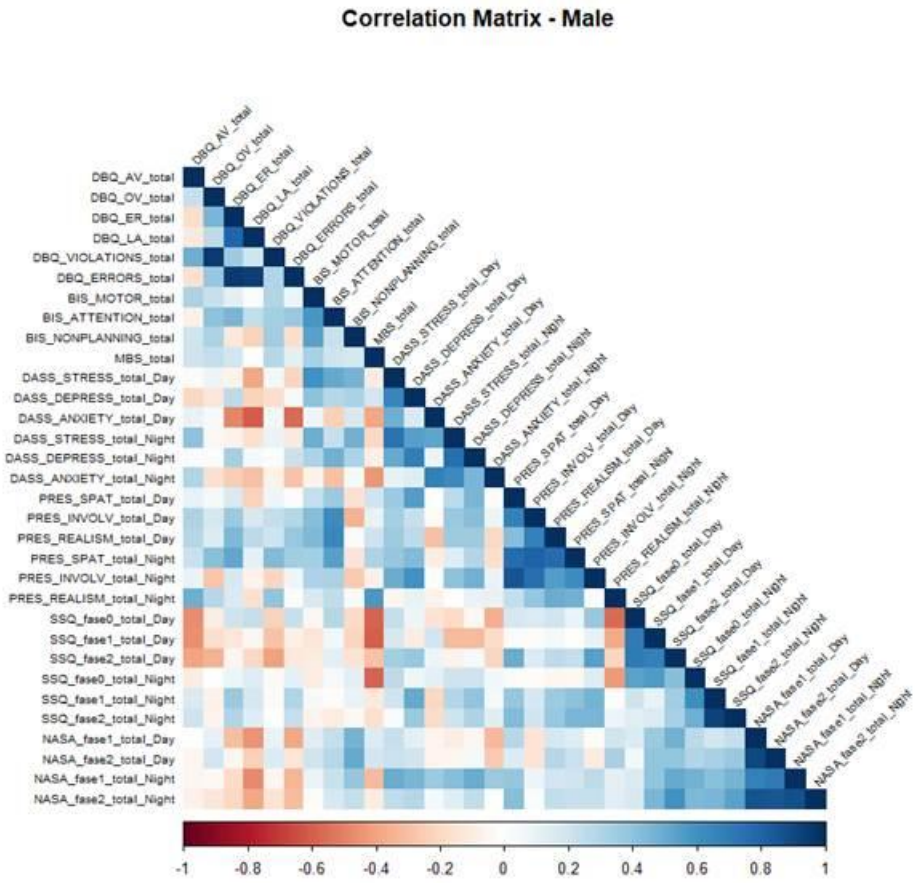


Figure 42 - Correlation matrices displaying Pearson correlation coefficients among all measured variables (male subsample N = 15). The color gradient ranges from dark red (r=-1.0, perfect negative correlation) through white (r=0, no correlation) to dark blue (r=1.0, perfect positive correlation).

4.1.2 Inferential Statistics

4.1.2.1 Minimum Time To Collision (minTTC)

Given the characteristics of our dataset, such as small and unequal sample sizes (N=13 females, N=15 males), potential violations of normality assumptions, traditional parametric ANOVA approaches may not be appropriate. We therefore employed the **Aligned Rank Transform (ART) procedure** for non-parametric factorial analysis (Wobbrock et al., 2011), which offers several critical advantages for our study design.

The ART procedure addresses key limitations of traditional non-parametric tests. ART correctly handles interaction terms in factorial designs, which is essential for examining whether the effect of time of day (Phase) differs between Experimental Conditions. ART also allows us to accommodate experimental design with repeated measures (individuals). Importantly, the rank-based approach does not assume normality and is robust to outliers and non-homogeneous variances, making it particularly suitable in case of small sample sizes where distributional assumptions are difficult to verify. The alignment procedure ensures that rankings

for each effect are performed on appropriately transformed data, maintaining proper Type I error rates even in the presence of interactions (Salter & Fawcett, 1993).

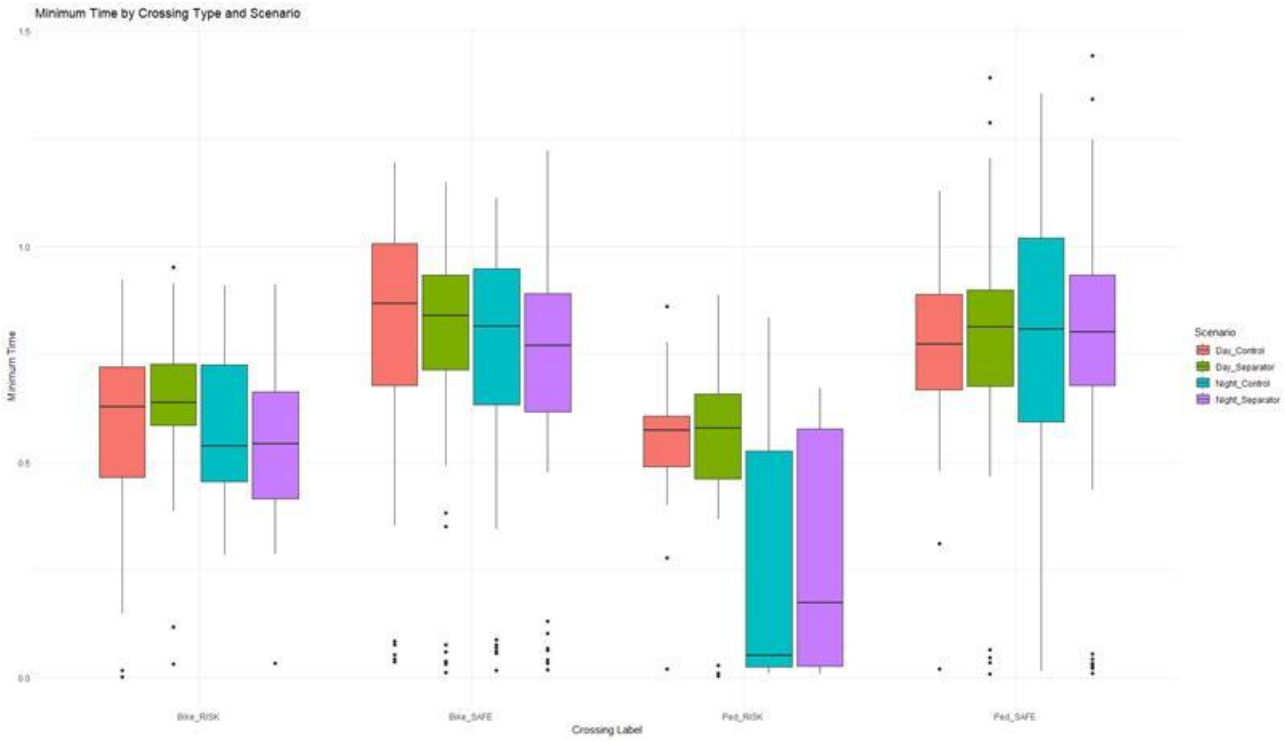


Figure 43 - Ditribution of observed minimum TTC across Phase (day time), Experimental Condition and type of crossing.

We run a comprehensive ART model with three predictors in interaction: Day Time (Night/Day), Condition (Experimental/Control) and Type of Crossing (Pedestrian Safe/Cyclist Safe/Pedestrian Risky/Cyclist Risk). We also assume the participants as a random intercept to account for individual variability.

4.1.2.1.1 Main Effects

The ART ANOVA revealed multiple significant effects on the minimum Time to Collision. A highly significant main effect of experimental condition was observed ($F(1, 548) = 48.28$, $p < .001$), with participants in the Control condition exhibiting shorter mean Time to Collision values ($M = 0.84$s, $SD = 0.24$) compared to the Experimental condition ($M = 0.96$s, $SD = 0.27$). The experimental manipulation actually increased drivers' temporal safety margins by approximately 15% on average across all conditions.

A significant main effect of daytime emerged ($F(1, 16) = 11.85$, $p = .003$), suggesting that Time to Collision differed between day and night driving sessions. Overall, night driving produced slightly longer safety margins across most crossing types, though this effect varied substantially depending on the specific crossing scenario and experimental condition.

Interestingly, a significant main effect of crossing type was found ($F(3, 548) = 286.36$, $p < .001$), indicating substantial variation in Time to Collision across different pedestrian crossing scenarios. Descriptive statistics revealed the longest Time to Collision values for Bike_SAFE crossings (Day Control: M = 1.16s, SD = 0.28; Day Experimental: M = 1.31s, SD = 0.16; Night Control: M = 1.07s, SD = 0.32; Night Experimental: M = 1.25s, SD = 0.20), followed by Ped_SAFE crossings (Day Control: M = 0.84s, SD = 0.19; Day Experimental: M = 0.98s, SD = 0.20; Night Control: M = 0.92s, SD = 0.19; Night Experimental: M = 0.99s, SD = 0.29). The shortest Time to Collision values were observed for risky crossing scenarios, particularly Bike_RISK (Day Control: M = 0.59s, SD = 0.19; Day Experimental: M = 0.64s, SD = 0.11; Night Control: M = 0.48s, SD = 0.14; Night Experimental: M = 0.66s, SD = 0.16) and Ped_RISK (Day Control: M = 0.58s, SD = 0.12; Day Experimental: M = 0.63s, SD = 0.16; Night Control: M = 0.65s, SD = 0.20; Night Experimental: M = 0.70s, SD = 0.13).

4.1.2.1.2 Interaction Effects

A significant three-way interaction between daytime, crossing type, and condition was observed ($F(3, 548) = 4.13$, $p = .007$), suggesting that the effect of experimental condition on Time to Collision varied depending on both the time of day and the type of crossing scenario. Additionally, significant two-way interactions emerged between daytime and crossing type ($F(3, 548) = 6.34$, $p < .001$) and between crossing type and condition ($F(3, 548) = 5.61$, $p < .001$).

Post-hoc simple effects analysis revealed that the experimental condition effect varied substantially across different scenario contexts. During **daytime**, the experimental manipulation produced significant increases in Time to Collision for Bike_SAFE crossings (estimate = 76.9, p = .016) and slightly for Bike_RISK crossings (estimate = 56.0, p = .035), but not for pedestrian crossing types (Ped_SAFE: estimate = -22.5, p = .484; Ped_RISK: estimate = 27.9, p = .545).

During nighttime, the pattern shifted. The main experimental effect emerged for Bike_RISK crossings (estimate = -99.5, p = .029), where the experimental condition increased Time to Collision by approximately 0.18 seconds. For other nighttime scenarios, effects were non-significant (Ped_SAFE: estimate = 42.0, p = .229; Bike_SAFE: estimate = -27.0, p = .585; Ped_RISK: estimate = -12.9, p = .860).

Simulation Time effects on Time to Collision were minimal and largely non-significant across most scenario combinations. For **Control** condition, only Bike_RISK showed a non-significant trend toward shorter Time to Collision during night versus day (estimate = 131.5, p = .013). In the **Experimental** condition, Ped_SAFE crossings showed negligible day-night differences (estimate = 47.7, p = .153), while Bike_SAFE (estimate = -54.5, p = .083) and Bike_RISK (estimate = -82.2, p = .072) showed trends, though these did not survive Bonferroni correction.

These findings indicate that while experimental intervention generally increased Time to Collision safety margins, the magnitude and even presence of this effect was highly **crossing-dependent**. The experimental manipulation proved most effective for safe bicycle crossings during daytime and risky bicycle crossings during nighttime. The substantial differences across crossing types (ranging from 0.48s to 1.31s) suggest that drivers naturally adjust their approach behavior based on perceived risk and vulnerable road user type. However, the significant three-way interaction reveals that experimental interventions designed to enhance driver safety must account for the complex interplay between environmental context (crossing type), lightening factors (daytime), and individual behavior patterns.

4.1.2.2 Standard Deviation from Lateral Position (SDLP)

The second main objective of the study was to identify potential differences in Lateral Position across Experimental Conditions and Day Times. Figure 44 presents the average SDLP scores for the four scenario combinations, controlling for Biological Sex. Importantly, rather than using a single overall score per participant, we accounted for experimental time in this analysis by considering ten average SDLP scores per participant. These scores were obtained by dividing each driving session into ten equal segments (within-subject) and calculating the standard deviation of Lateral Position for each segment.

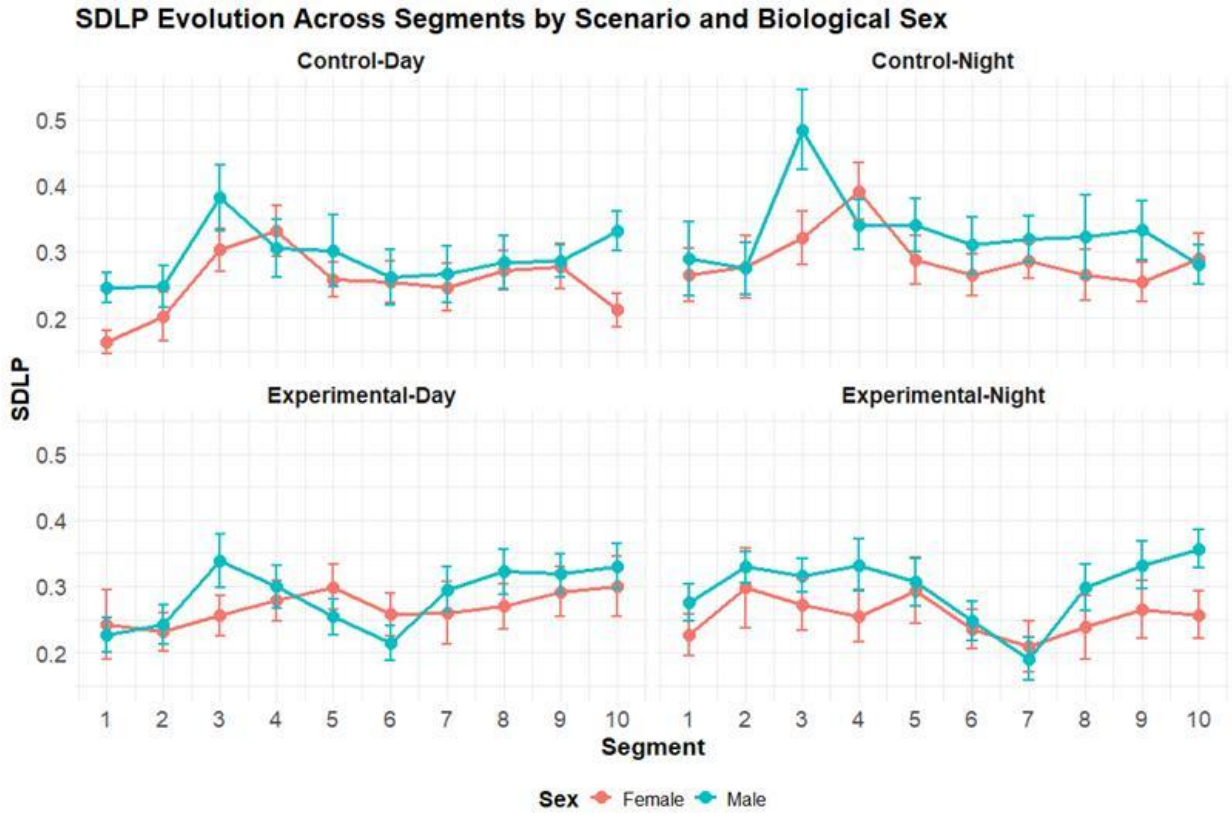


Figure 44 – Error bars of average Standard Deviation from Lateral Position (SDLP) per each time segment (10% of the driving simulation time) controlling for combinations of Experimental conditions (Experimental, Control) and Day Time (Day, Night) and for Biological Sex.

Given the right-skewed distribution of SDLP values and the inclusion of a continuous covariate (track segment), we employed a linear mixed-effects model on log-transformed data rather than ART ANOVA, as this approach naturally accommodates continuous predictors while maintaining interpretability of main effects and interactions, and allows for proper control of segment-related variance without requiring all possible interaction terms. Still, these findings should be interpreted with caution given the modest sample size (N=28 participants, with N=13 females and N=15 males) and unequal group distributions.

4.1.2.2.1 Main Effects

The linear mixed-effects model with log-transformed SDLP values, controlling for segment position, revealed significant effects on lateral lane deviation. A significant main effect of Phase emerged ($F(1, 1071.47) = 9.38$, $p = .002$), indicating that lane-keeping performance differed between day and night driving sessions. Overall, drivers showed slightly better lane-keeping during daytime ($M_{log} = -1.41$, $SE = 0.047$; back-transformed: $M = 0.244$m) compared to nighttime ($M_{log} = -1.36$, $SE = 0.046$; back-transformed: $M = 0.257$m). The main effect of experimental condition did not reach significance ($F(1, 1068.01) = 0.57$, $p = .448$), with Control ($M_{log} = -1.36$, $SE = 0.047$; $M = 0.257$m) and Experimental conditions ($M_{log} = -1.41$, $SE = 0.047$; $M = 0.244$m) showing similar overall lane deviation patterns when collapsed across time of day.

4.1.2.2.2 Interaction Effects

Most notably, a significant interaction between Phase and condition was observed ($F(1, 1068.01) = 7.04$, $p = .008$), suggesting that the effect of experimental condition on lane-keeping varied substantially between day and night driving sessions. The segment covariate, included to control systematic changes across the road course, did not result to be statistically significant ($F(1, 1068.01) = 1.03$, $p = .311$).

Simple effects analysis revealed that the experimental manipulation had differential effects depending on simulation daytime, with descriptive statistics from the original (non-transformed) scale providing interpretable effect magnitudes. Figure 45 depicts the effect described below.

During daytime driving, no significant difference emerged between Control ($M_{log} = -1.43$, $SE = 0.051$; $M = 0.239$m, $SD = 0.064$m, $N = 280$) and Experimental conditions ($M_{log} = -1.39$, $SE = 0.051$; $M = 0.250$m, $SD = 0.065$m, $N = 280$; $t(1068) = -0.76$, $p = .449$). This suggests that the experimental intervention did not substantially alter lane-keeping behavior during daytime conditions. During nighttime driving, a significant effect of condition was observed (estimate = 0.126, $t(1068) = 3.01$, $p = .003$), with Control condition showing notably worse lane-keeping ($M_{log} = -1.30$, $SE = 0.051$; $M = 0.273$m, $SD = 0.073$m, $N = 300$) compared to

Experimental condition (M_log = -1.42, SE = 0.051; M = 0.242m, SD = 0.060m, N = 300). This represents a non-trivial improvement in lateral stability—an 11.4% reduction in lane deviation under the experimental manipulation during night driving, corresponding to approximately one-third of a standard lane width improvement in vehicle positioning precision.

Less prominent than the previous effect, drivers showed significantly better lane-keeping during day in the Control condition (M = 0.239m, SD = 0.064m) versus night (M = 0.273m, SD = 0.073m), estimate = -0.126 log units, t(1068) = 3.01, p = .003. This deterioration represents a 14.2% increase in lateral deviation during nighttime, indicating substantial degradation in lateral control at the baseline scenario. For the Experimental condition, no significant day-night difference was observed (Day: M = 0.250m, SD = 0.065m; Night: M = 0.242m, SD = 0.060m), estimate = 0.033, t(1068) = -0.76, p = .449.

These findings reveal a critical interaction whereby the experimental intervention specifically protected against daytime deterioration in lane-keeping performance. Under Control conditions, drivers exhibited the expected pattern of significantly worse lateral stability during nighttime driving (0.273m vs. 0.239m SDLP, a 14.2% increase). Remarkably, nighttime performance under Experimental conditions (M = 0.242m) was statistically indistinguishable from daytime Control performance (M = 0.239m), suggesting the intervention not only prevented daytime degradation but maintained lane stability at optimal baseline levels.

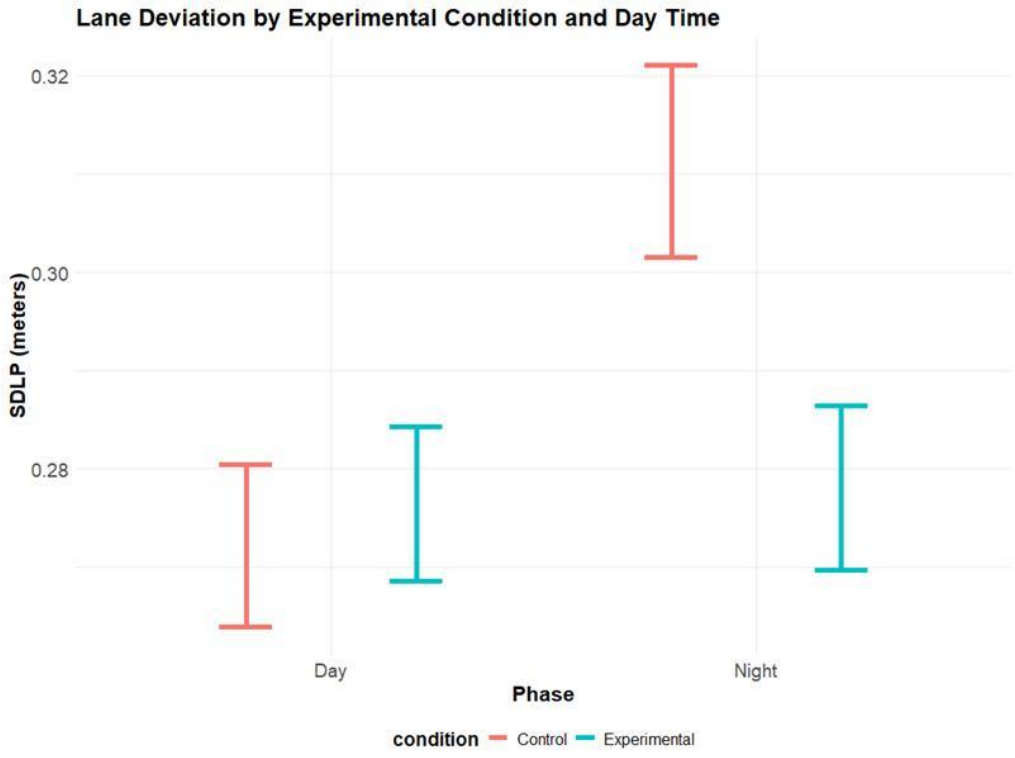


Figure 45 - Error bars of estimated SDLP scores for the interaction between Experimental Condition (Experimental, blue; Control, red) and Day Time.

4.2 Cardiac data

4.2.1 Preprocessing

4.2.1.1 Preprocessing of the ECG signal

Preprocessing was carried out using the NeuroKit2 library, applying the ecg_process() function with default parameters. The raw signal was subjected to band-pass filtering (0.5–80 Hz) to remove baseline wander and high-frequency noise. R-peak detection was performed using the Pan–Tompkins algorithm, which involves computing the first derivative of the signal, squaring the resulting values, and applying a moving window to identify local maxima exceeding an adaptive threshold. Instantaneous heart rate was derived from RR intervals by PCHIP interpolation (Piecewise Cubic Hermite Interpolating Polynomial), obtaining a continuous signal at 1000 Hz.

4.2.1.2 Quality check

Each session underwent a two-level quality control. At the session level, two criteria were applied to the portion of signal following the baseline: (1) mean signal quality index ≥ 0.70 and (2) percentage of samples with HR outside the physiological range 40–150 bpm ≤ 1%. This second threshold was formulated in percentage terms – rather than as an absolute exclusion of extreme values – to tolerate sporadic artifacts without penalizing the entire session, while still discarding recordings with systematic errors in R-peak detection (double detections or missed detections). At the single-epoch level, an epoch was discarded if the mean quality was lower than 0.70 or if more than 10% of samples had quality below 0.50.

4.2.1.3 Segmentation into epochs

For time-course analyses, the signal was segmented into epochs centered on the triggers corresponding to pedestrian crossing events recorded in the driving simulator. The time window extended from -15 to +30 seconds relative to the trigger instant (total duration: 45 s). Epochs whose distance from the previous trigger was less than 15 seconds were excluded to prevent contamination of the pre-trigger window. For each epoch, the local baseline was defined as the mean heart rate in the -15/-10 s window before the trigger – deliberately distant from the event to exclude anticipatory responses. Heart rate was expressed as a difference from this baseline (ΔHR). The signal was resampled from 1000 to 50 Hz by selecting one sample every 20, reducing computational cost without loss of relevant information.

4.2.2 Data Analysis

4.2.2.1 Analysis A: Heart Rate Variability (HRV) at Session Level

To examine predictors of RMSSD (Root Mean Square of Successive Difference, a measure of heart rate variability) at the single-session level, we estimated a linear mixed model with random intercept for subject.

The dependent variable was standardized RMSSD (global z-score). Fixed predictors included: mean session SDLP (a measure of driving instability), RMSSD calculated at baseline (3 minutes pre-drive), total NASA-TLX self-report score measuring perceived workload at the end of the driving session, SSQ delta, i.e., the difference between nausea/malaise perceived at the beginning and at the end of the driving session, perceived *Presence* while driving (single-item self-report, shared for both sessions on the same day), visibility condition (Day/Night), and road type (Control/Bar). The model was estimated on 62 observations across 24 groups (subjects)

4.2.2.2 Analysis B – Heart rate (HR) at segment level

To examine predictors of mean HR in the 10-time segments into which each session was divided, we estimated a second LMM with a random intercept for subject. The dependent variable was standardized segment HR. Fixed predictors included: mean SDLP in the segment, pre-session baseline HR, segment index (time index), post-session total NASA-TLX score, self-reported SSQ delta, reported level of *Presence*, visibility condition, and road type. The model was estimated on 620 observations across 24 groups (mean: 25.8 observations per subject). All continuous variables were standardized to global z-scores prior to analysis.

4.2.2.3 Analysis C – Global cardiac response and condition comparisons

A cluster-based permutation test (5000 permutations, $\alpha = .05$) was applied to ΔHR time-courses averaged per subject. Clusters were identified as contiguous sequences of time points exceeding the critical *t* threshold; their significance was assessed by comparing the observed t-mass with the null distribution obtained via sign-flip permutation. Condition comparisons (Bar vs. Control; Day vs. Night; Bar × time-of-day interaction) were conducted on pairwise differences between time-courses.

4.2.2.4 Analysis D – TTC × ΔHR correlation at single-epoch level

To examine whether Time-to-Collision modulated the cardiac response trial by trial, we estimated, for each time point, an LMM with random intercept for subject and standardized TTC as fixed predictor (dependent variable: ΔHR for the single epoch). The resulting 2250 p-values were corrected using the Benjamini–Hochberg method (FDR, $\alpha = .05$).

4.2.3 Results

4.2.3.1 Sample attrition

Twenty-eight participants provided raw ECG data (79 .txt files). Four participants (32 recordings) were excluded because file duration was under 1 minute due to technical issues. Three additional sessions failed

session-level quality control. The final sample included 24 participants for Analyses A and B (62 sessions for Analysis A; 620 segments for Analysis B) and 23 participants for time-course analyses (299 valid epochs).

4.2.3.2 Analysis A – Predictors of session HRV

The linear mixed model on session RMSSD (62 observations, 24 subjects) yielded a single statistically significant predictor: baseline RMSSD (β = 0.141, SE = 0.033, *z* = 4.29, *p* < .001, 95% CI [0.076, 0.205]). This result indicates that participants with higher cardiac variability in the minutes before the start of driving tend to show higher cardiac variability also during the session, suggesting a strong stable trait component in autonomic regulation. No other predictor reached statistical significance.

Table 11 – Summary table of the linear mixed model on session HRV

Variabile	Coef.	SE	z	p	CI 2.5%	CI 97.5%
Intercept	0.005	0.219	0.021	0.983	-0.424	0.434
sdlp_overall_z	-0.014	0.042	-0.333	0.739	-0.097	0.069
rmssd_baseline_z	**0.141**	0.033	**4.290**	**<0.001**	0.076	0.205
nasa_total_z	-0.054	0.080	-0.666	0.505	-0.211	0.104
ssq_delta_z	0.005	0.028	0.176	0.860	-0.050	0.060
presence_z	0.007	0.068	0.102	0.919	-0.126	0.140
day_night_idx	0.041	0.057	0.720	0.472	-0.070	0.152
bar_control_idx	0.009	0.050	0.169	0.866	-0.090	0.107
Random effect (Group Var)	1.095	2.837				

4.2.4 Analysis B – Predictors of HR at segment level

By far the strongest predictor was baseline HR (β = 0.820, SE = 0.028, *z* = 29.70, *p* < .001, 95% CI [0.765, 0.874]): pre-drive heart rate level is the main determinant of HR during the session. Visibility condition showed a significant positive effect (Day/Night: β = 0.096, SE = 0.025, *z* = 3.88, *p* < .001, 95% CI [0.047, 0.144]), with higher HR in the night condition compared to daytime. Road type also had a significant effect

(Bar/Control: β = 0.069, SE = 0.018, *z* = 3.74, *p* < .001, 95% CI [0.033, 0.105]), with higher HR in the condition with a lateral barrier compared to the control condition. Both findings are consistent with increased physiological activation under driving conditions perceived as more visually complex or alarming.

SSQ delta showed a significant positive association with segment HR (β = 0.049, SE = 0.012, *z* = 4.00, *p* < .001, 95% CI [0.025, 0.073]), indicating that greater increase in simulator sickness over the session is associated with higher HR. Driving performance, segment index, and workload and *Presence* scales did not reach statistical significance.

Table 12 - Summary table of the linear mixed model on HR at the segment level (10% of total Driving simulation time).

Variabile	Coef.	Std.Err.	z	p	CI 2.5%	CI 97.5%
Intercept	-0.085	0.059	-1.430	0.153	-0.201	0.031
sdlp_seg_z	0.011	0.010	1.092	0.275	-0.008	0.030
hr_baseline_z	0.820	0.028	29.697	<0.001	0.765	0.874
segment_z	0.016	0.009	1.814	0.070	-0.001	0.033
nasa_total_z	-0.018	0.028	-0.639	0.523	-0.072	0.036
ssq_delta_z	0.049	0.012	4.003	<0.001	0.025	0.073
presence_z	-0.010	0.027	-0.373	0.709	-0.064	0.043
day_night_idx	0.096	0.025	3.880	<0.001	0.047	0.144
bar_control_idx	0.069	0.018	3.741	<0.001	0.033	0.105
Random effect (Group Var)	0.075	0.119				

4.2.4.1 Analysis C – Global cardiac response and condition comparisons

The cluster-based permutation test on the mean time-course collapsed across all conditions did not detect any statistically significant clusters (largest cluster: -0.6/1.7 s, *p* = .136; late cluster: 18.6/21.3 s, *p* = .111). Thus, there is no evidence of a reliable mean cardiac response to pedestrian crossing in the whole sample. None of the condition comparisons produced significant clusters. The Bar vs. Control contrast (n = 18) and the Day vs. Night contrast (n = 12) do not reveal systematic differences in the ΔHR temporal profile. The Bar × time-of-day interaction test, conducted on 8 participants with data in all four conditions, is to be considered severely underpowered and its outcome not interpretable.

4.2.4.2 Analysis D – TTC × ΔHR correlation

The trial-by-trial LMM analysis (299 epochs, 23 subjects) identified a single significant time point after FDR correction (1/2250, pFDR < .05), located at -5.6 s relative to the trigger (β = 0.83). This is most likely a residual false positive not fully removed by the FDR procedure.

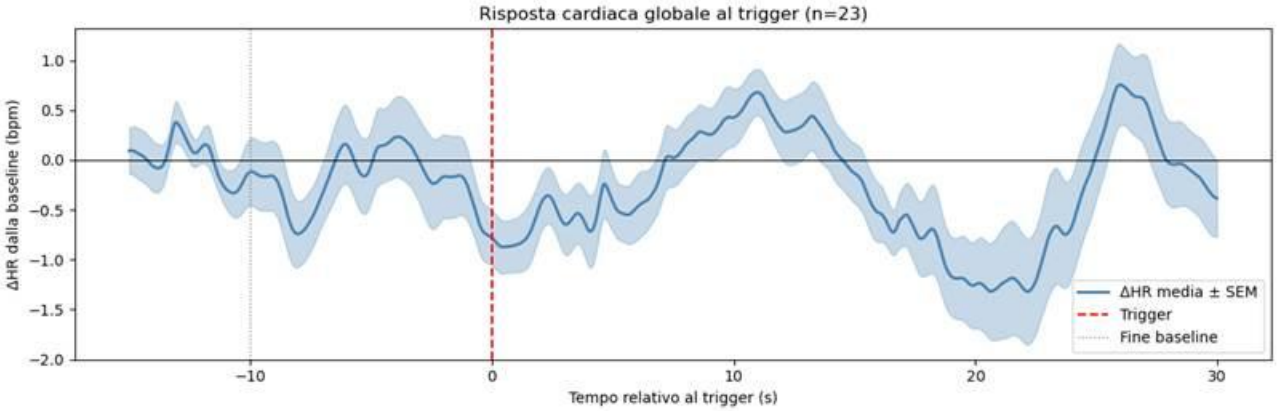


Figure 46 - Mean heart rate change (ΔHR ± SEM) across all participants (n=23), time-locked to pedestrian crossing (t=0, red line).

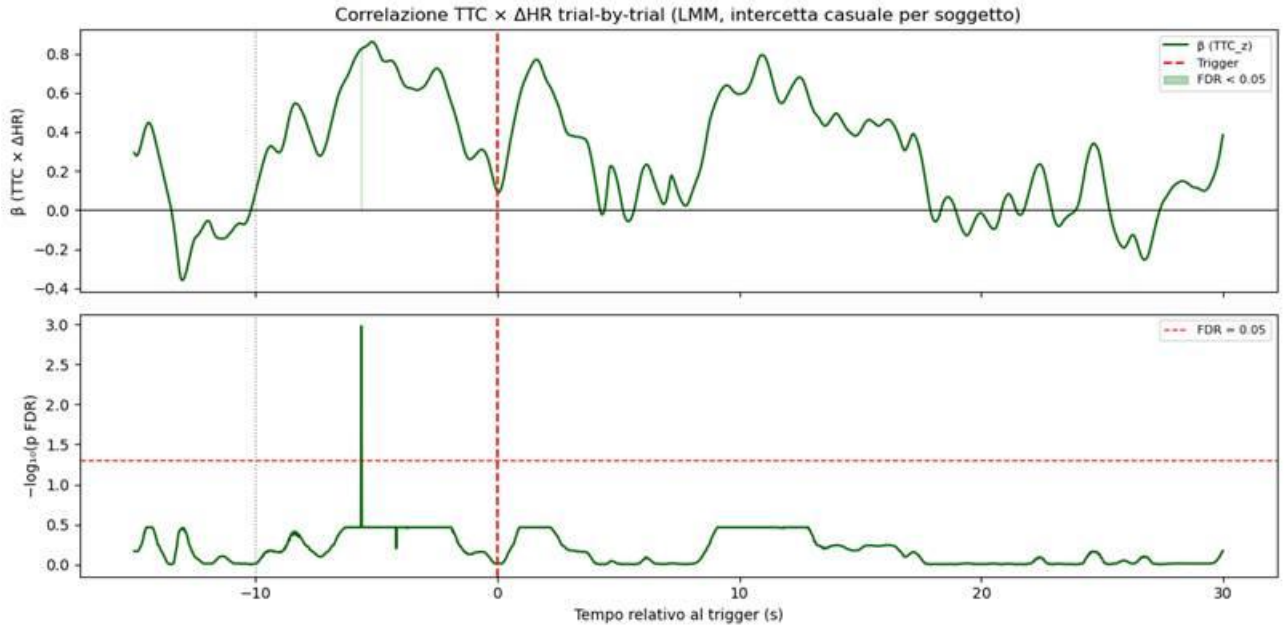


Figure 47 - Linear mixed model analysis (299 epochs, 23 subjects) testing TTC × ΔHR correlation. Top: Beta coefficients (green = FDR < 0.05). Bottom: -log$_{10}$(FDR) with significance threshold (horizontal red line).

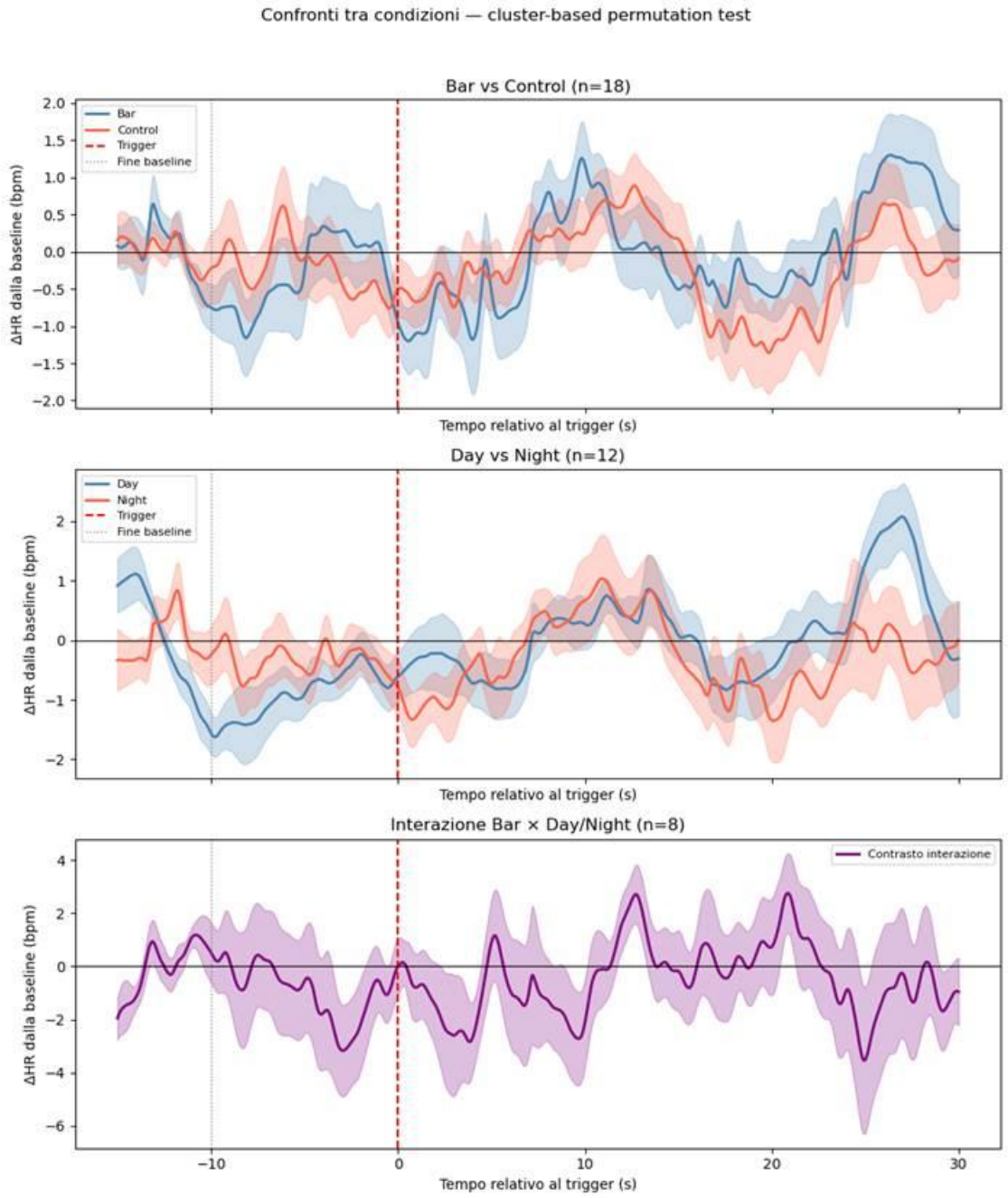


Figure 48 - ΔHR comparisons (mean ± SEM) across conditions. Top: Bar vs Control (n=18). Middle: Day vs Night (n=12). Bottom: Bar × Day/Night interaction (n=8).

4.3 Conclusion

Despite the modest sample size (N=28; 13 females, 15 males) requiring cautious interpretation, several meaningful patterns emerged.

Individual Differences: Psychological trait assessments (DBQ, BIS, MBS) showed expected consistency, with sex-specific correlation patterns suggesting males demonstrate stronger associations between psychological distress and impulsivity, while females show more pronounced negative correlations between immersive presence and risky driving behaviors.

Temporal Safety Margins: Analysis of Time to Collision revealed substantial variation across crossing scenarios (1.2s for safe bicycle crossings vs. 0.5-0.7s for risky scenarios), with a significant three-way interaction indicating context-dependent intervention effects. The experimental manipulation proved most effective for high-risk scenarios during nighttime, precisely when collision avoidance margins are most critical.

Lane-Keeping Performance: Control conditions showed expected simulation day time degradation (14% worse lane-keeping at night), but experimental manipulation eliminated this vulnerability. Nighttime experimental performance (0.242m) matched daytime baseline levels (0.239m), demonstrating that appropriately designed interventions can preserve optimal lateral control during high-risk circadian phases.

Psychophysiological Responses: DASS measures showed high day-night stability with minimal sex differences, while NASA-TLX workload varied across phases with sex-specific relationships to psychological state. Simulation sickness and presence measures revealed consistent individual differences, with presence potentially buffering cybersickness in females.

Cardiac Data: Preprocessing yielded high-quality ECG data from 24 participants, with baseline HR and RMSSD as strong session predictors. Night conditions and barriers elevated HR, but no reliable event-related cardiac responses or TTC correlations emerged, indicating stable trait-like autonomic patterns under load.

5 Final remarks

These final remarks articulate the methodological framework adopted within the ARCADE Project. The combined use of observational data from real HRLs using Lidar device, multi-platform and multi-perspective simulation (car, motorcycle, pedestrian CAVE, virtual-reality headsets), psychometric profiling, and analyses based on surrogate safety measures (SSMs) proved feasible and sensitive to evaluate of risk conditions for VRU and support more informed decision-making processes.

They set out a coherent conceptual and operational structure capable of integrating heterogeneous data, models, and technologies, thereby providing a structured starting point of reference for road authorities, public administrations, and future research groups.

Within this perspective, the methodological framework is based on:

- an accurate evaluation of road operating conditions;
- a systemic interpretation of conflicts and risk phenomena, accounting for the interactions among infrastructure, vehicles, and road users;
- evidence-based design of mitigation interventions, supported by experimental findings and advanced simulation tools;
- a stronger integration between scientific research, technological innovation, and governance processes in the domain of road safety.

These final remarks thus serve as a tool to support the development of evidence-oriented road safety strategies, building on the experience gained during the ARCADE project and consolidating the methodological advances achieved throughout its implementation. These concluding considerations advocate a multidisciplinary and data-driven approach that integrates empirical observations, behavioural modelling, simulation-based evidence and emerging sensing technologies into a coherent analytical framework structured around the following pillars:

- **Pillar No. 1 -** Spatial analysis of road accidents and territorial risk interpretation;
- **Pillar No. 2 -** Surrogate Safety Measures in Data-Driven Safety Analysis;
- **Pillar No. 3 -** Data Sources: LiDAR as an Innovative Sensor for Road-Safety Analysis;
- **Pillar No. 4 -** Advanced multi-platform and multi-perspective simulation.

Pillar No. 1 - Spatial analysis of road accidents and territorial risk interpretation

The first pillar concerns the spatial analysis of road accidents within a GIS environment, where techniques such as **Kernel Density Estimation (KDE)** enable a systematic interpretation of crash distribution and the detection of recurrent spatial patterns. This analytical framework supports a territorial interpretation of safety conditions and provides an evidence-driven basis for prioritising interventions. KDE is particularly valuable because it produces a boundary-independent representation of spatial phenomena, avoiding the distortions introduced by administrative units or regular grid structures. The resulting density maps make it possible to identify high-risk hotspots with precision, compare temporal periods to assess the impact of infrastructural or regulatory measures, and guide operational planning by directing resources toward the most critical locations.

KDE plays a central role due to its capacity to reveal spatial structures of risk that would otherwise remain hidden within the dispersion of individual points. It provides an interpretable and analytically robust foundation for strategic decision-making in road-safety planning. When compared with approaches constrained by administrative boundaries or regular grid systems, KDE offers a more rigorous and representative means of characterising spatial patterns and supporting risk-oriented analyses of road accidents. The resulting density maps enable analysts to identify high-risk hotspots, compare different time periods to evaluate the effectiveness of infrastructural measures, and define operational priorities by directing resources toward the most critical locations. On the other hand, the spatial analysis of road accidents gains analytical value when it is complemented by contextual information. Road geometry, traffic flows, operating speeds, and environmental and temporal conditions are all factors that influence the likelihood of a crash occurring. Integrating these datasets makes it possible to construct more comprehensive models capable of explaining not only *where* accidents take place, but also *why*. The objective is to achieve an interpretative and predictive understanding in which the spatial distribution of crashes is linked to underlying structural risk factors. This approach enables the identification of latent criticalities and the design of targeted interventions grounded in robust empirical evidence. To this purpose, the **ARCADE project implemented a dedicated survey conducted in Rome, Turin and Padua**, described in detail in Deliverable 1, which reports the results of the accident and KDE analyses carried out on real intersections and road segments selected as case studies.

Pillar No. 2 - Surrogate Safety Measures in Data-Driven Safety Analysis

Surrogate safety measures provide a complementary and more proactive perspective alongside accident data, which describe events that have already occurred. These indicators make it possible to assess potential risk by observing interactions among road users, even in the absence of collisions. This approach is particularly valuable in contexts where accidents are infrequent or where the objective is to evaluate the

effectiveness of interventions before measurable effects on crash occurrence can be detected. Beyond this function, surrogate indicators and related metrics offer a dynamic representation of traffic conflicts, capturing the immediacy and severity of near-miss events that traditional crash-based analyses cannot reveal. By quantifying the temporal and spatial proximity of potentially hazardous interactions, they enable a more sensitive assessment of safety conditions, support the early identification of emerging risks and provide a robust analytical basis for testing infrastructural or regulatory measures under real-world operating conditions.

Pillar No. 3 - Data Sources: LiDAR as an Innovative Sensor for Road-Safety Analysis

The growing availability of advanced sensing technologies has transformed the observation and analysis of road-user behaviour. Traditional and well-established data sources, such as fixed video-camera systems, remain essential for reconstructing trajectories and identifying interactions; however, recent technological developments have introduced innovative sensors capable of capturing traffic dynamics with far greater precision. Among these, real-time LiDAR systems play a particularly significant role, as they generate high-resolution, three-dimensional point clouds that enable detailed measurement of movements, speeds and accelerations across heterogeneous user groups. These enhanced data streams support more accurate detection of near-miss events, more reliable computation of surrogate safety indicators, and a deeper understanding of the dynamics that shape safety outcomes in complex urban environments. In this context, the main novelty introduced by the **ARCADE Project lies in the use of surrogate safety measures extracted from LiDAR sensing technologies.** The contribution of LiDAR sensors to safety analysis is especially significant thanks to several intrinsic advantages, such as:

- detection of movements and interactions with high spatial and temporal accuracy;
- support for dynamic and adaptive risk assessments, enabling near real time conflict detection;
- reliably operation even under low visibility conditions, such as night time or fog;
- identification of subtle interactions and near misses that may not be detectable through image-based systems alone.

The activities conducted on real case studies during the project demonstrate that the use of surrogate safety measures derived from LiDAR data enables a detailed reconstruction of vehicle kinematics, including trajectories, speeds and accelerations, from which indicators describing vehicle–pedestrian, vehicle–cyclists and motorcycle–pedestrian interactions can be extracted for road-safety analysis.

The capacity to capture fine-grained motion patterns enhance the computation of SSMs, offering a more reliable understanding of how Vulnerable Road Users interact within complex urban environments. This methodological advancement contributes to a more comprehensive and predictive evaluation of safety

conditions, supporting evidence-based decision-making in the design and management of urban mobility systems.

Pillar No. 4 – Multi platform and multi-perspective simulation

Immersive simulation constitutes a further methodological pillar of the ARCADE Project, expanding the analytical capacity of traditional observational and sensor-based approaches. This technique provides a controlled, repeatable, and fully configurable experimental environment in which driver/motorcyclist/pedestrian behaviour can be examined. By recreating complex traffic scenarios, hazardous situations, or rare events, these systems allow researchers to investigate how individuals perceive risks, process information, and make decisions without exposing them to actual danger. Within these simulated environments, it becomes possible to analyse a wide spectrum of behavioural and cognitive dimensions. The immersive nature of the simulation strengthens drivers' behavioural realism, providing participants with a heightened sense of presence that evokes authentic behavioural responses while preserving full experimental control.

This methodological axis is particularly valuable for evaluating of new road-safety interventions, infrastructure designs, and technological solutions before their implementation, offering a predictive understanding of how drivers might adapt to changing mobility environments.

Immersive simulation represents a central methodological component of the ARCADE Project, expanding its capacity to analyse road-user behaviour beyond what can be observed through field data or sensor-based monitoring alone.

Within the ARCADE Project, a diversified suite of simulators was employed to reproduce the same interaction scenarios from multiple and complementary perspectives. That allowed a **multi-perspective assessment** of the interaction between road users (driver, motorcyclist, pedestrian) through driving/riding/walking simulations of the same conflicting road scenario. This multi-perspective approach yielded a more comprehensive understanding of the reciprocal dynamics that underpin the emergence of conflicts.

The quality and analytical value of the simulation depend directly on the fidelity of the reconstructed environments. In this regard, **Building Information Modelling (BIM)** constitutes the most advanced procedure for generating realistic, coherent, and continuously updatable digital scenarios. BIM enables the integration of geometries, materials, metadata, and functional logics within a single, consistent model, ensuring that the virtual environment accurately reflects the physical context. The **combination of GIS datasets, LiDAR surveys, and high-resolution 3D modelling used in the ARCADE Project** has allowed the creation of virtual environments that reproduce real-world conditions with a high degree of realism.

By exposing participants to identical conditions across multiple user perspectives, the project was able to analyse behavioural responses, risk perception, and conflict dynamics in a more comprehensive and methodologically robust manner. This multi-perspective simulation strategy provides behavioural data that more closely approximate real-world conditions and supports the evaluation of safety interventions before their implementation. Such capabilities are of particular interest to public administrations and road-infrastructure authorities, who can rely on simulated evidence to assess the potential effectiveness of design modifications, regulatory measures, or traffic-management strategies in a safe, replicable, and cost-efficient environment.

The robustness of the experimental phase is significantly enhanced by the psychometric and behavioral assessment. This component focuses on the interplay between stable individual traits, such as impulsivity, driving habits and psychological states. The psychological profiling is crucial for understanding why different road users react diversely to the same infrastructural stimuli, providing a granular look at the "human factor" that traditional safety analyses often overlook.

The results of the carried out experimental campaigns in simulated environmental highlight that human factors—age, sex, impulsivity, workload, susceptibility—systematically modulate the effect of infrastructural countermeasures.

These results support a "**human-centered" approach to design**, where the corridor or intersection is evaluated not only for nominal compliance with standards but also for how it shapes the perceptual and cognitive demands on different user groups.

Integration of the four Methodological Pillars

The ARCADE approach is grounded in the complementarity of four distinct yet mutually reinforcing analytical perspectives, which together form a coherent framework for understanding and improving road-safety conditions. It integrates the spatial analysis of road accidents and the territorial interpretation of risk, enabling the identification of recurrent patterns and context-specific vulnerabilities across the urban environmental. It incorporates surrogate safety measures within a data-driven analytical paradigm, using predictive indicators derived from observed interactions to anticipate potential conflicts before they manifest as crashes. It relies on LiDAR as an innovative and high-resolution data source, capable of capturing detailed trajectories, behavioural dynamics, and conflict events with a level of precision unattainable through traditional observational methods. Finally, it employs multi platform simulation which allows the controlled reproduction of critical scenarios and the experimental assessment of road-user behaviour across multiple perspectives, including drivers, pedestrians, and motorcyclists. Through the integration of these four pillars, ARCADE combines multi-perspective evidence, predictive modelling, and immersive experimentation into a unified methodological strategy capable of supporting robust and context-sensitive road-safety interventions.

The ARCADE project demonstrates that integration of observation, modeling, and experimentation **provides a comprehensive and innovative pathway for proactive road safety management.** This methodological synergy offers public administrations and road authorities a powerful decision-support tool, enabling them to transition from reactive accident-based policies to evidence-driven prevention strategies.

References

Gettman, D., Pu, L., Sayed, T., & Shelby, S. G. (2008, June 1). *Surrogate Safety Assessment model and Validation: Final report*. https://rosap.ntl.bts.gov/view/dot/39210

Lanzaro, G., & Sayed, T. (2024). Evaluating driver-pedestrian interaction behavior in different environments via Markov-game-based inverse reinforcement learning. *Expert Systems With Applications*, *260*, 125405. https://doi.org/10.1016/j.eswa.2024.125405

Peesapati, L. N., Hunter, M. P., & Rodgers, M. O. (2018). Can post encroachment time substitute intersection characteristics in crash prediction models? *Journal of Safety Research*, *66*, 205–211. https://doi.org/10.1016/j.jsr.2018.05.002

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011, May). The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 143-146).

Salter, K. C., & Fawcett, R. F. (1993). The ART test of interaction: a robust and powerful rank test of interaction in factorial models. *Communications in Statistics-Simulation and Computation*, 22(1), 137-153.