Anhang englische Version

Method Development Study on Generating a Safety Performance Indicator based on Euro NCAP Assessment Results

by

Marcus Bäumer
Heinz Hautzinger
Manfred Pfeiffer

IVT Research GmbH
Mannheim

Berichte der Bundesanstalt für Straßenwesen

Fahrzeugtechnik Heft F 144

Abstract

Method Development Study on Generating a Safety Performance Indicator based on Euro NCAP Assessment Results

The European New Car Assessment Programme (Euro NCAP) is a consumer protection orientated programme for the safety assessment of - as a rule - new car models. The programme was established in 1997, since 2009 it has consisted of 4 Boxes (1. Adult Occupant Protection; 2. Child Occupant Protection; 3. Vulnerable Road User Protection; 4. Safety Assist Systems).

The main objective of this project was to transfer the test results from Euro NCAP onto the overall car fleet wherever possible and to use this to form a Safety Performance Indicator (SPI) for the vehicle stock over several consecutive years (time series). A further objective was to investigate whether or not a relationship exists between the Euro NCAP assessment of vehicles and accident occurrences. For this purpose, in addition to literature analyses, statistical models were estimated on the effect of vehicle safety expressed by the SPI on the corresponding number of persons injured in accidents taken from the official road traffic accident statistics.

A fundamental step in the development of an SPI of vehicle safety consisted of making the Euro NCAP test results from the various years comparable as far as possible with regard to the test procedures which had changed in the course of time. For this purpose, a project group was formed consisting of experts from the Federal Highway Research Institute (BASt) in the fields of active and passive vehicle safety which had the task of quantifying, for each Box, the extent to which these changes to the test procedures affect the vehicle assessment. The test conditions of the year 2020 were used as reference.

The second fundamental step in generating a Safety Performance Indicator consisted of linking (matching) the - newly calculated - Euro NCAP result data for the individually tested makes and models to the Central Vehicle Register inventories (1.1.2014 to 1.1.2020).

The matching of the Central Vehicle Register inventory data with the Euro NCAP assessment data was carried out using a complex algorithm which is based, in essence, on the characteristics of make code, model code and year of initial registration.

As a result, on average over the seven years observed here (reference dates: 1.1.2014 to 1.1.2020), around 70 % of the newer vehicles (initial registration year from 2009) were able to be allocated to a Euro NCAP assessment. The remaining cars in the Central Vehicle Register inventory were assigned the missing assessment using an imputation procedure.

A total of four (Box-specific) Safety Performance Indicators were formed which are based on the newly calculated and standardised vehicle assessments from Euro NCAP. The specified indicators are mean values of the corresponding safety assessment of the cars recorded in the Central Vehicle Register. Using these four indicators, an Overall Safety Performance Indicator was then also calculated using weighted average. Thus, as a result, time series for the four Box-specific SPI values as well as the overall SPI value are available.

The core result of the analyses of the SPI time series is that all indicators show an increasing time trend. This is a clear indication that the safety standard has not only continuously risen in recent years for new vehicles, but also in terms of the overall car fleet.

The highest indicator values can be found in the SPI for Box 2 (Child Occupant Protection), they turn out lowest with regard to the Safety Assist Systems (SPI for Box 4). If the indicators are additionally grouped into car segment, then SUVs, followed by ATVs, commercial vans and the luxury class show the highest values in the overall SPI value. The fact that SUVs have the highest SPI is certainly also related to the fact that this is a relatively new segment in which the share of older vehicles is comparably low.

As part of the accident analysis, log-linear regression models were estimated in order to determine the effect of the four SPIs on each of the corresponding accident characteristics (injured car occupants, pedestrians and cyclists, cars involved in accidents). In addition, the relationship between the overall SPI value and the resulting monetary economic accident costs was analysed. The statistical assessments on the association between SPI and accident occurrences show, in all cases, that a higher value of the corresponding Safety Performance Indicator is accompanied by a lower number of persons injured in accidents or cars involved in accidents. With regard to accident costs, the results here were a (significant) reduction of the accident costs by around 0.7 %, when the overall SPI increases by 1 %.

In addition, in the concept of the project, the fundamental prerequisites for an uninterrupted continuation of the SPI time series in the following years were already created.

Contents

1 Goals and objectives

1.1 Project objectives

The **Euro**pean **N**ew **C**ar **A**ssessment **P**rogramme (Euro NCAP) is a consumer protection orientated programme for the safety assessment of passenger cars. It is supported by a consortium of European Ministries of Transport, automobile clubs, insurance associations and research institutes[1]. Since 1997, new vehicles are tested for their passive and active safety. For both manufacturers and consumers, Euro NCAP has become an important assessment standard for vehicle safety. Today, twelve institutions from eight European countries are involved in the programme[2].

In this research project, Euro NCAP results will be used to derive a ‘Safety Performance Indicator’ (SPI) on vehicle safety for the Federal Republic of Germany over several years. An SPI is understood to be parameters which describe the safety performance of a road transport system independent of accident occurrences (i.e. without direct reference to the number of crashes) (see also chapter 1.3). Examples of these are the number of annual traffic offences or figures on the quality of rescue services.

The main objective of this investigation consists of generating an SPI on the basis of the Euro NCAP results for new vehicles with which the equipment of the overall vehicle fleet (car fleet) can be summarised using elements of passive and active safety. For this purpose, the results of the Euro NCAP test for the individual vehicle models are extrapolated to the overall vehicle stock of a year. In principle therefore, it is about matching the Euro NCAP test data and car inventory data.

The indicator[3] which is formed in this way will be generated for several calendar years (beginning with 2014) and, as far as possible, be comparable across the individual years. The methodical approach applied will be documented in detail and interpretation notes will be additionally provided in order to also enable the continuation of the time series into the future.

An additional objective of the project is to create the reference of the SPI to real crash occurrences. Overall, the SPI to be formed serves to make the contribution of vehicle technology to the safety performance on German roads visible.

1.2 The European New Car Assessment Programme

Currently, almost all new vehicle models are tested and assessed for their safety according to Euro NCAP. The test programme is made up of the following four boxes[4] (see EURO NCAP 2020, p. 14):

1. Adult Occupant Protection (primarily passive safety): Several crash tests are carried out here including frontal impact, lateral impact or rear impact. In order to simulate the effects of traffic accidents on vehicle occupants, life-size dolls (crash test dummies), equipped with a large number of sensors, are used to represent an adult. New since 2020 is an impact from the opposite side of the occupants (far side test). Also new is an assessment as to how far the vehicle supports the safe and fast rescue of occupants who have had a crash (rescue).
2. Child Occupant Protection (primarily passive safety): On the one hand, the results of the crash test are included here – measured on dummies which represent six-year-old and ten-year-old children. On the other hand, the equipment of the vehicle with characteristics for the protection of children is examined, such as the (automatic) deactivation of the passenger airbag for example, or the installation of child restraint systems.
3. Vulnerable road user protection (active and passive safety): In this Box, tests on head impact, upper leg impact and lower leg impact to pedestrians are carried out with the use of test specimens. In addition, autonomous emergency braking systems for pedestrians and cyclists are assessed.
4. Safety Assist Systems (primarily active safety): The assessment of safety assist systems (e.g. lane assistants or seatbelt reminders) have been included in the overall rating since 2009 as independent fourth box (see, for example, Federal Highway Research Institute (BASt) 2009).

---

[1] As automobile manufacturers are not represented in the consortium (they simply provide the test objects), Euro NCAP is a goods quality test.

[2] see https://www.euroncap.com/de

[3] If necessary an indicator set, e.g. according to vehicle segment.

[4] In order to get the Euro NCAP rating for a single test object, a total of around seven vehicles of the respective type have to be included in the destructive test.

Every box thus consists of several test procedures. For each box, the points achieved in the individual test procedures are added up and then related to the maximum achievable total points for the respective box (percentage value). The overall rating of the vehicle (stars) results from the weighted mean value of the four box-specific percentage values. The percentage values of the individual Boxes are also translated into a star assessment which can lead to a devaluation of the overall result due to minimum requirements (so-called balancing).

Since the start of the Euro NCAP process, many changes have taken place. These do not only affect the number, type and implementation of the test procedures (see, for example, VAN RATINGEN et al. 2016[5]) but, for example, also the weighting of the boxes or the requirements for the number of points needed to achieve a certain star rating. The consequence of this is that the results of the individual years are only very limitedly comparable with each other.

Therefore, in order to improve the comparability of the results over the years – which is an important prerequisite for the construction of an annual SPI – corresponding data modifications need to be carried out. These will be described in detail in chapter 3.2.

1.3 Safety Performance Indicators

Safety generally indicates the state of being protected from danger or damage, being free of unjustifiable risks (hazards). Therefore, transport safety is the characteristic of a transport system, to be able to carry out transportation or location change processes without damage to the transport objects (persons, goods) and the modes of transport.

The degree of safety of a transport system can be primarily defined through the relative frequency of damage and the resulting consequences (public health, economy). In addition however, there are also parameters which describe the safety performance of a transport system independent of crash occurrences (i.e. without direct reference to crash figures and numbers of injured persons). Such parameters are indicators of the safety performance of a transport system or, in short, performance indicators[6]. The term ‘Safety Performance Indicator’ (SPI) is commonly used for this in English-speaking countries.

Key scientific articles on the subject of ‘indicators of the safety of systems’ can be found, for example, in the *Safety Science Journal*. When it comes to the transportation sector, safety performance indicators (SPIs) are mainly widely used in the field of aviation. Based on the requirements of the International Civil Aviation Organization (ICAO), there is also extensive literature on the development and implementation of SPIs which can, in part, also be applied to the road transport system. The document *Measuring Safety Performance – Guidelines for Service Providers* published by the SM ICG (Safety Management International Collaboration Group) can be mentioned here as an example.

Indicators of safety in road transport (or road safety performance indicators), i.e. variables or parameters according to which safety-relevant states and processes in a road transport system can be accurately identified and, if necessary, predicted, is something which the field of mobility and crash research has been dealing with for some time. As examples of such indicators, the observation of ‘traffic conflicts’ and surveys on safety brought on by seat belts, helmets and other protection systems can be mentioned here (see Federal Highway Research Institute (BASt) 2020a; KATHMANN et al. 2019).

As far as we know, the first comprehensive article on the subject of ‘indicators of the safety performance in the transport system’ originates from the European Transport Safety Council (ETSC 2001). Based on a general conceptual framework of performance indicators, this article primarily focuses on ‘Road Safety Performance Indicators’. Safety performance indicators (SPIs) are concisely defined there as follows (p. 12):

> ‘A safety performance indicator is any variable that is used in addition to accidents or injuries to measure changes in safety performance.’

or somewhat more specifically (p. 5)

> ‘These (*i.e. SPIs*) are defined as any measurement that is causally related to crashes or injuries, used in addition to a count of crashes or injuries, in order to indicate safety performance or understand the process that leads to accidents.’

For reasons which are explained in more detail in the ETSC report (p. 11), the number of crashes and injured persons should be supplemented with SPIs in order to gain a more complete picture of the transport safety performance and its changes over the course of time. Ideally, SPIs should pro-

---

[5] Lifting (e.g. test ESC discontinued since 2016), revision (e.g. pole impact 2015) or the new inclusion (e.g. full width frontal impact since 2015) of test procedures.

[6] Ultimately, however, the number of crashes and injured persons is a performance indicator.

vide indications of newly emerging problem situations before these evolve into number of crashes and injured persons (early indicators). With regard to transport safety work, the respective players should be put in the position by SPIs of being better able to affect the factors causing crashes which lie in their domain.

Not all of the extensive amount of possible SPIs are equally important (p. 13). Generally speaking, the importance of an SPI can be measured by how strong its relationship with the occurrence of crashes and injuries is *and* to what extent the corresponding factor can be affected by transport safety measures or action programs.

The ETSC report groups the performance indicators relating to road transport into the following four categories:

- Indicators of the behaviour of road users (especially mean value and variance of the driving speed, proportion of trips under the influence of alcohol, quotas of safety belt usage)
- Route-related indicators (e.g. grouping of net length according to frictional properties of the carriageway)
- Vehicle-related indicators (grouping of vehicle fleet according to the number of Euro NCAP stars, share of vehicles with technical defects etc.)
- Indicators of rescue services and medical care (e.g. mean value and variance of time period until arrival at scene of crash).

In the section ‘Functional requirements for road safety performance indicators’ in the ETSC report (p. 21-22), 10 steps for the definition and measurement of performance indicators in the context of transport safety policies are stated as typical ideal:

1. Proof of a causal relationship between crashes and a potential performance indicator
2. Assess the policy or measure relevance of a potential performance indicator
3. Operational description of a potential road transport safety problem as an indicator or set of indicators
4. Define the results in a measuring protocol per performance indicator
5. Define a performance indicator measuring or survey programme
6. Carry out the measurements or surveys
7. Compare results of step 6 with the targeted road safety programmes
8. Verify/ validate the assumptions formulated in step 1
9. Based on the outcome of step 6, modify a targeted road safety programme – if necessary – and finally
10. Report on the results of the whole process, e.g. annually. This sequence of steps is reflected in the structure of this research report.

The concepts for SPIs designed in the ETSC report were further developed, operationalised and empirically tested in the EU project SafetyNet (Building the European Road Safety Observatory). With regard to the topic of vehicle-related performance indicators presented here, the SafetyNet research report by HAKKERT, GITELMAN and VIS (2007) is relevant. The central issue to be answered in this context is formulated by the above-mentioned authors as follows (p. 4):

> ‘Euro NCAP is widely used as an indicator of passive safety for individual vehicles to give consumers a guide to the crashworthiness of specific makes and models. However, there is no current recognised measure of an entire vehicle fleet’.

In this research report, it is investigated how a methodically sound Euro NCAP-based performance indicator for the car fleet as a whole can be achieved for Germany.

2 Relationship between Euro NCAP rating and real accident occurrence

In this chapter, an overview of the literature on the relationship between the Euro NCAP rating (exposition status) of vehicles and the consequences of crashes for vehicle occupants or unprotected collision opponents (injury status) will be provided. The presented studies are mainly based on the observation of specific crash scenarios (e.g. car-car-collisions), which allows a more precise quantification of the relationships due to the elimination of certain disturbance variables (e.g. road condition, weather and lighting conditions). On the other hand, the observation of specific crash constellations leads to restrictions in terms of generalisation of the results.

An attempt to derive general relationships between the Euro NCAP rating and the risk of crashes or injuries will be made by way of a meta-analysis. The results are to serve as basis for the later investigation of the effect of the Safety Performance Indicators (SPI) to be constructed on the total crash occurrences (total number of fatalities, seriously injured, slightly injured) for the car fleet as a whole.

2.1 Studies on the risk of injuries to car occupants

A current study on the relationship between the Euro NCAP rating of vehicles and the risk of injuries to vehicle occupants originates from Sweden (KULLGREN et al. 2019). Car-car-collisions recorded by the police with at least one injured occupant on the front seat was investigated by way of a matched pairs analysis. The relative risk of injury to the driver of vehicle 1 (case car) subject to the number of stars the vehicle gained for Box 1 of the Euro NCAP test (Adult Occupant Protection) was estimated. The basis of comparison here is the risk of injury to the driver of vehicle 2 (other car) involved in the same accident for which, however, no Euro NCAP rating is available (exposition status is unknown). The collision opponents are interpreted as sample from *all* vehicles as a whole: 'The collision partners are considered to be a sample of the whole car population, and therefore they provide the exposure basis that allows for comparisons across all case vehicles' (KULLGREN et al. 2019, p. 3). In fact, however, every subgroup of 'case cars' has a specific comparison group. Thus, for example, with regard to the vehicle mass, the collision opponents of 5-star vehicles differ structurally from the collision opponents of the 2-star vehicles (ibid., p. 8 (Table 5)).

A distinction is made between the relative risk of injury in total, the relative risk of a serious or fatal injury and also the relative risk of a fatal injury. The analysis is based on 102,466 car-car-crashes recorded by the police in the years 1994 to 2018. In addition, the relative risk of a permanent medical impairment for the car driver was also defined on the basis of 57,863 injured car occupants (front-seat occupants) from the years 2000 to 2018. Insurance data (AIS values) were additionally drawn upon for this.

In the estimation formulas, differences in mass between the involved car (mean mass of the assessed or opposing vehicles) and the year of the crash were taken into account (for details see KULLGREN et al. 2019, p. 4). The results on the relative risk of injury are reproduced in detail here in the form of Tab. 1.

Euro NCAP rating of the vehicle (no. stars in Box 1)	Relative risk of injury in total	Relative risk of serious or fatal injury	Relative risk of fatal injury
2	1.03	0.96	0.84
3	0.95	0.87	0.69
4	0.90	0.80	0.70
5	0.85	0.75	0.50
Total	0.91	0.84	0.70

Tab. 1: Relative risk of injury for car drivers subject to the Euro NCAP rating of the vehicle (source: KULLGREN et al. 2019, Table 5)

The table shows, very generally, that compared with the respective collision opponent, the risk of injury – overall and differentiated according to seriousness of injury – for the car driver sinks with their vehicle's number of Euro NCAP stars (Box 1, Adult Occupant Protection). While the overall risk of injury for drivers of cars with 2 stars is still 3% higher than that of their collision opponent (relative risk: 1.03), it is 15% lower for drivers of 5-star cars (relative risk: 0.85).

On the basis of this table, risk comparisons can also be drawn between vehicles with different star ratings. However, it must be remembered here that the respective comparison groups differ structurally. Putting to one side the methodological points of critique which are to be made[7] on the quoted work, one can see, for example, that compared to vehicles with a 2-star rating, the relative risk of injury (in total) in vehicles with 3 stars is around eight percentage points lower[8]: (0.95 - 1.03) × 100 = -8%. The following risk differences result in the same way: 4 stars vs. 2 stars: -13%; 5 stars vs. 2 stars: -18%.

If, instead of the difference, the ratio of two relative risks is used as parameter for the effect of the number of stars, it can be seen, for example, that the relative risk of a fatal injury in vehicles with 5 Euro NCAP stars is even 40% lower than for those with 2 stars ((0.50 - 0.84)/0.84) × 100 = -40.5%). In total, therefore, the results of this study show a quite clear relationship between the Euro NCAP rating (Box 1) of the vehicle and the risk of injury for the vehicle occupants.

A Swedish precursor study from the year 2010 (KULLGREN et al. 2010) also arrives at very similar results. This investigation is also based on the analysis of car-car-collisions, whereby differences in mass between each of the two involved cars are taken into account. Vehicles with 2 stars were compared with those that received 5 stars. The latter show a total of around 10% lower risk of injury. In the subgroup of serious and fatal injuries, the difference amounts to 23% and rises to 68% when observing fatal injuries alone. The risk of a permanent medical impairment compared to 2-star vehicles lies at around 27% lower for 5-star vehicles.

Also originating from Sweden is a further study on the basis of this scheme which was published in the year 2002 (LIE and TINGVALL 2002). However, the underlying number of cases from that time were still comparably small: the basis was almost 16,000 car-car-crashes between January 1994[9] and March 2000. The risk of injury to car drivers was investigated with a differentiation according to overall risk and that for serious and fatal injuries (with correction of differences in mass). In contrast to the above-mentioned studies, no distinction between differently assessed cars were found in terms of the overall injury risk. On the other hand, in the case of serious and fatal injuries, the risk drops by around 12% per additional star. Well assessed vehicles (3 or 4 stars for Adult Occupant Protection) show around a 30% lower risk of a serious or fatal injury in comparison to those with low (2 stars) or no rating (for the risk of a fatal injury on the basis of American data, see also FARMER 2005).

Another Swedish investigation of car-car-crashes carried out with this method largely arrived at the same result (LIE et al. 2001): on the basis of crashes recorded by the police between 1994 and 2000, the overall risk of injury is not dependent on the Euro NCAP rating of the Adult Occupant Protection, while in the case of serious and fatal injuries, the respective risk is reduced by around 12% per additional star. In addition, in this study, comparisons were drawn between the Euro NCAP rating and the ‘Folksam Car Model Safety Rating’ (see FOLKSAM INSURANCE GROUP 2019; HAUTZINGER 2006). The latter is based on insurance data from which the risk of a car driver suffering injuries which lead to fatal or disabling injuries as a consequence of a crash is assessed (single-vehicle car crashes are also included here). The result (analysis years 1992 to 1999) showed a reduction of the risk by 7% per Euro NCAP star in Box 1.

Investigations into the relationship between crash test results and real crash occurrences were also carried out in the EU project SARAC II (Quality Criteria for the Safety Assessment of Cars Based on Real-World Crashes). PASTOR (2007) analysed 495 car-car-crashes on the basis of crashes recorded by the police in Germany (1998 to 2002). These dealt with frontal collisions between two cars, each assessed in Euro NCAP, in which at least one person was seriously or fatally injured (frontal collisions were selected using the characteristics kind and type of accident). Thus, in terms of the Euro NCAP rating, only the results of the frontal impact (frontal offset test) were drawn upon for the driver (for details see PASTOR 2007, p. 159). In the case of every crash, the involved vehicles were compared with each other with regard to the severity of injuries of their drivers. Crashes in which the involved vehicles did not differ with regard to the Euro NCAP rating and/or the severity of the drivers' injuries were not taken into consideration. The analysis was carried out using a log-linear model. The probability of a less serious inju-

---

[7] Although it is clearly a study *with paired observations*, the authors have not taken this fully into account in the data analysis phase. If pairing is ignored, the effect of a risk factor (here the number of NCAP stars) may be systematically underestimated. For further details see WOODWARD 2005, pp. 298-302.

[8] Difference of two relative risks (‘Reduction of the relative risk’).

[9] At the start of the Euro NCAP programme in 1997, car models were also tested which had already been on the market for a longer period of time (e.g. Renault Laguna since 1994). For this reason, the assessment of 1997 could be assigned, for example, to a Laguna which was involved in a crash in 1995.

ry compared to the other party was estimated given the Euro NCAP rating and further impact characteristics such as the mass ratio of the vehicles or the gender of the driver. The result showed a clear relationship to the crash test results. In vehicles with 3 stars, as well as those with 4 stars, there was a significantly higher probability of a lower injury severity compared to cars with 1 star.

The SARAC report on the subtasks 2.1 and 2.2 (NEWSTEAD et al. 2006) dealt with the relationship between severity of injury in crash data reported by police and crash test results in a lot of detail. As indicator for the injury severity, the so-called ‘crashworthiness rating’ was drawn upon. This is an index relating to vehicle model which essentially expresses the proportion of seriously and fatally injured car drivers out of all drivers which were involved in a car-car-crash with personal injuries. This share is defined using two components: the risk of injury (share of injured/killed drivers out of all drivers) and the severity of injury (share of seriously injured/killed drivers out of all injured/killed drivers). Each of these two components is estimated using statistical models while including further influencing variables – such as, for example, age and gender of the driver (see also NEWSTEAD and CAMERON 1999). In the study, the results for several European countries as well of those for Australia and New Zealand are presented. With the data from Germany – in contrast to those from other countries – no in-depth analysis for special crash constellations (frontal and lateral impact) was able to be carried out. Only the results for Germany are briefly summarised below. These are based on car-car-crashes registered by the police with at least one injured driver in the years 1998 to 2002. Single vehicle car crashes were also taken into account for the establishment of injury severity. The Euro NCAP rating refers to the number of stars in Box 1 (Adult Occupant Protection). The resulting estimated values for the crashworthiness rating across all vehicle models are shown in Tab. 2.

Euro NCAP rating of the vehicle (no. stars)	With consideration of vehicle mass			Without consideration of vehicle mass		
	Estimated value	95% confidence interval		Estimated value	95% confidence interval	
		Lower limit	Upper limit		Lower limit	Upper limit
1	–	–	–	–	–	–
2	12.17%	11.81%	12.54%	12.70%	12.33%	13.08%
3	11.89%	11.51%	12.28%	12.46%	12.08%	12.86%
4	10.08%	9.70%	10.47%	9.19%	8.86%	9.54%

Tab. 2: Crashworthiness rating subject to the Euro NCAP rating of the vehicle (source: NEWSTEAD et al. 2006, p. 10)

The respective analyses were carried out both with as well as without consideration of the vehicle mass. In both cases, the result is a better rating with rising number of stars. However the vehicles with 2 stars do not differ significantly to those with 3 stars as the respective confidence intervals overlap. Vehicles with 1 star were not taken into account due to the number of cases being too low.

This analysis for the German crash data was repeated again with a somewhat wider data basis (DELANEY et al. 2006a). By considering the model generation, a larger number of vehicles involved in crashes could be assigned to a Euro NCAP rating. However, the results differ only slightly from those displayed in Tab. 2, particularly as far as the significances are concerned.

As already mentioned, results for different countries were produced in SARAC II. This also applies to the first phase of the SARAC project. LANGWIEDER et al. (2003, p. 8) report results with British crash data from the years 1993 to 1998. The following crashworthiness ratings resulted across all crash constellations:

- 1 star: 12.02%
- 2 stars: 8.08%
- 3 stars: 7.81%
- 4 stars: 6.27%

An even stronger relationship can be found in the separate observation of crashes with lateral impact (LANGWIEDER et al. 2003). The range of the ratings here lies between 15.8% (1 star) and 2.26% (4 stars).

2.2 Studies on the risk of injury for pedestrians

PASTOR (2013) investigated car-pedestrian-accidents using police crash data from Germany (years 2009 to 2011). This drew upon 7,576 urban crossing accidents, each involving one Euro NCAP-assessed car (only cars tested after 2002) and one pedestrian. The assessment was limited to accidents in which the pedestrian was at least slightly injured and was between six and 64 years old. In order to test the relationship between the severity of the pedestrian's injuries and the Euro NCAP-assessment of the car, an ordinal probit model was used in which injury severity (slight, serious, fatal) was modelled subject to the Euro NCAP pedestrian *point score* (not the number of stars!) and further characteristics (e.g. lighting conditions at the time of the accident). In the result, every additional point (not a star!) in the Euro NCAP pedestrian test lowers the relative probability of a fatal injury to the pedestrian by 2.5%. In terms of serious injuries, the reduction amounted to 1% per point. Thus, for example, vehicles with 22 points in comparison to those with 5 points show a 35% (16%) reduced probability of a fatal (serious) injury to the pedestrian. According to this study, if one assumed that only cars rated at 22 points were on the roads, this would result in 6% less fatalities and 9% less seriously injured pedestrians.

STRANDROTH et al. (2011) examined 488 car-pedestrian-accidents on the basis of Swedish crash data from the years 2003 to 2011. Only accidents were observed in which one pedestrian collided with the front of a Euro NCAP-tested car and which happened on roads with a speed limit of 50 km/h. Analyses included, inter alia, the severity of the pedestrian's injuries (in total 1,156 injuries with AIS 1 and higher) subject to the results of the Euro NCAP pedestrian test for the car. As only 13 vehicles with a 3-star assessment for the pedestrian test were included in the sample, only comparisons between 1-star and 2-star vehicles could be carried out. Vehicles with 2 stars showed a 17% lower share of injuries with a degree of severity of AIS 2 and higher. In the case of injuries with AIS 3 and higher, the difference even amounted to 28%, however this is not statistically significant. Furthermore, significant differences between 1 and 2 star cars were determined for different degrees of serious injuries or injuries with serious consequences.

Car-pedestrian-accidents were also the object of observations in the EU project SARAC II (DELANEY et al. 2006b). Using British, German and French crash data, the risk of a serious or fatal injury to pedestrians after a collision with a (before 2002) Euro NCAP-tested car was analysed. A differentiation could only be made between cars with 1 or 2 stars in the pedestrian test, the assessment was carried out using logistic regression models. In none of the three investigated accident data data bases did the result show a significant relationship between the severity of pedestrian's injuries and the car's Euro NCAP rating.

As part of the EU project ASPECSS[10] (VAN DER ZWEEP et al. 2014), the use of an emergency braking system for pedestrians in combination with the passive safety of a vehicle in collisions with a pedestrian was investigated. In order to assess the passive safety, the results of the Euro NCAP pedestrian test were taken as basis, grouped into the categories bad (12.2 points), average (22.6 points) and good (32.2 points). The results in the form of accident costs show, for example, that for vehicles assessed as bad – in terms of the passive pedestrian safety – the costs of accidents can be reduced to almost the same amount as for vehicles assessed as average without emergency braking system by the use of equipment with powerful emergency braking systems.

2.3 Discussion

In the majority of studies, a relatively clear relationship can be found between the Euro NCAP rating and injury risk, in particular for the risk of serious or even fatal injuries. However, the transfer of the results to an SPI which is to be constructed as part of this study at an aggregate level – and its relationship with the crash occurrence – is also associated with several difficulties or reservations as far as the generalisability or transferability of the results are concerned.

For example, several of the presented studies originate from abroad and also some are already several years old. The question arises here as to the transferability of the results to present-day German situation.

Almost all studies are based on the analysis of car-car-crashes. This does have advantages where the identification of the relationship between the Euro NCAP rating and crash occurrences is concerned, however it greatly restricts generalisability to other crash constellations. In addition, only the risk of injury to the car driver is generally investigated. Even if results on the risk of injury to pedestrians (subject to the Euro NCAP rating of the vehicle) are also available in isolated cases, overall statements can only be derived for a relatively small part of the crash or injured population. On the part of the Euro NCAP results, it is the case that in the quoted studies, not the overall rating, but only individual Boxes or even only parts of

---

[10] Assessment Methodologies for Forward Looking Integrated Pedestrian and Further Extension to Cyclists Safety.

them (e.g. results of the frontal offset test) are included in the analysis of the relationship[11].

Ultimately, in the case of studies which show results on the relative risk, there is still the methodical problem of the 'translation' of the gained indices into absolute number of injured persons (killed, seriously or slightly injured). With data on the absolute risk of injury subject to the number of stars, the safety effect of improvements to the car fleet in terms of vehicle technology could be quantified, if necessary, in particular in the form of the 'attributable risk'. In chapter 5.1, this concept is explained in more detail and practically applied.

In some of the Swedish studies mentioned above, the output data underlying the estimations have been published. These can be used for the estimation of the attributable risk if necessary. Nevertheless, the above-mentioned restrictions remain in terms of transferability and generalisability of the results in view of the derivation of a relationship between an SPI resulting from the Euro NCAP overall rating and the number of persons injured in road transport. Of course, the question also arises as to which injured population should, in actual fact, be observed (e.g. car drivers, all car occupants, pedestrians, cyclists, all injured persons).

It therefore seems evident, alternatively or additionally to the literature results, to define the correlation between the Safety Performance Indicators of the car fleet to be developed and the annual overall number of injured persons (killed, seriously injured, slightly injured) on the basis of the official accident statistics.

---

[11] It should be noted here that it is, of course, methodically completely correct to only use the test results on Adult Occupant Protection, for example, when only the risk of injury to the driver is relevant on the side of the crash.

3 Data bases and data preparation

3.1 Data bases

3.1.1 Euro NCAP data

The data with the Euro NCAP test results from the years 2009 to 2019 (euroncap.com) contain the following information per test object:

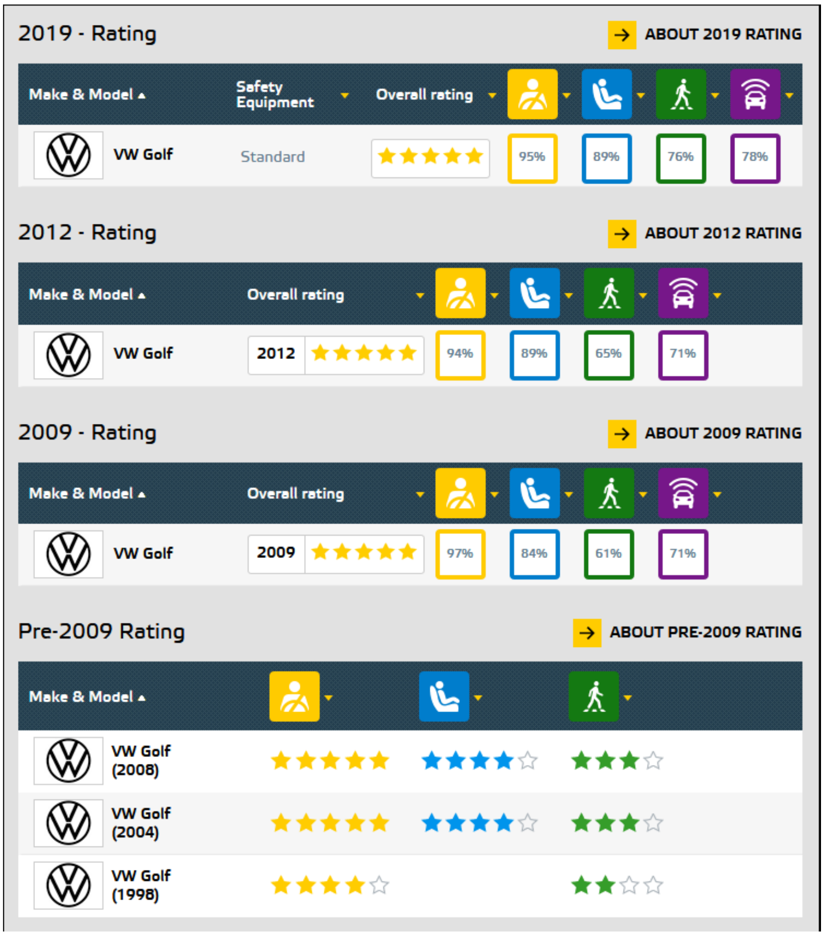
- Make and model
- Overall rating (number of stars)
- Share of achieved points in per cent for Box 1 (Adult Occupant Protection)
- Share of achieved points in per cent for Box 2 (Child Occupant Protection)
- Share of achieved points in per cent for Box 3 (Vulnerable Road Users)
- Share of achieved points in per cent for Box 4 (Safety Assist Systems)
- Year of the test

In addition to these characteristics, a consecutive number was generated which serves as basis for matching with the vehicle stock.

The data comprise a total of 442 vehicle models, including five models which were tested with the same protocol status (so-called reassessment). The results of each of the first tests for these five models are not taken into account hereinafter. The overall number therefore lies at 437 models.

In the available data, the star ratings each refer to the standard edition of the model. Since 2016, a model can be additionally tested with a ‘safety package’. Insofar as the elements of the safety package are integrated in the standard configuration of the model as a follow-up to the test, the assessment for the standard edition can be changed accordingly later (e.g. in the case of SsangYong Tivoli/ XLV).

Using the Volkswagen Golf as an example, Figure 1 shows how the test results are presented on the Euro NCAP website.



2019 - Rating

Make & Model ▴	Safety Equipment	Overall rating				
VW Golf	Standard	★★★★★	95%	89%	76%	78%

2012 - Rating

Make & Model ▴	Overall rating				
VW Golf	2012 ★★★★★	94%	89%	65%	71%

2009 - Rating

Make & Model ▴	Overall rating				
VW Golf	2009 ★★★★★	97%	84%	61%	71%

Pre-2009 Rating

Make & Model ▴			
VW Golf (2008)	★★★★★	★★★★☆	★★★☆
VW Golf (2004)	★★★★★	★★★★☆	★★★☆
VW Golf (1998)	★★★★☆		★★☆☆

Figure 1: Euro NCAP ratings for the Volkswagen Golf (source: euroncap.com)

From Figure 1 it is clear that in the Euro NCAP tests carried out before 2009, no overall rating for the respective vehicle model was shown. Rather, until then, there were separate star ratings for the test boxes 1 to 3, the 4th box for Safety Assist Systems was only introduced in 2009. For these reasons, only the results of the car models tested in Euro NCAP from 2009 will be used for the construction of the SPI of the car fleet.

In Tab. 3, it is shown how the overall star rating of the models tested in this period are distributed.

Number of stars (overall rating)	Number of models	in per cent
0	2	0.5
1	1	0.2
2	4	0.9
3	50	11.4
4	75	17.2
5	305	69.8
Total	437	100.0

Tab. 3: Distribution of the number of stars (overall rating) of the vehicle models tested between 2009 and 2019 in Euro NCAP (source: Euro NCAP 2009 - 2019)

The mean value resulting from this distribution lies at 4.54 stars. When grouped according to test year, the results are the average ratings displayed in Tab. 4.

Year of the test	Number of models	Average number of stars (overall rating)
2009	45	4.82
2010	27	4.59
2011	46	4.72
2012	44	4.48
2013	32	4.69
2014	41	4.27
2015	42	4.57
2016	18	4.39
2017	62	4.32
2018	25	4.36
2019	55	4.65
Total	437	4.54

Tab. 4: Average number of stars (overall rating) of the vehicle models tested between 2009 and 2019 in Euro NCAP grouped into test year (source: Euro NCAP 2009 - 2019)

One can see that the average assessment is relatively constant across the years. Progress in vehicle safety and the tightening of the test procedures in Euro NCAP apparently roughly balance themselves out. Therefore, an indicator based only on the average number of stars would be unsuitable for a time series.

3.1.2 Vehicle stock data

The vehicle stock data results from a separate assessment of the Central Vehicle Register (Zentralen Fahrzeugregister) which was commissioned by the Federal Motor Transport Authority (Kraftfahrt-Bundesamt). The stocks were assessed for the years 2009 to 2020 (reference date 1 January respectively) according to the following characteristics:

- Reference date year
- Manufacturer’s code number
- Manufacturer’s description (normed)
- Manufacturer’s description (not normed), if manufacturer’s code number = 0900 or 0901
- Type code number
- Official plain text type
- Trade name
- Trade description
- Make number
- Make description
- Model code
- Model description
- Segment code
- Segment description
- Year of initial registration

A connection with the Euro NCAP data which is as unique as possible should be guaranteed on the basis of these characteristics. Purpose-built vehicles (camper vans, ambulances, police cars etc.) were not included in the assessment.

In addition to the car inventory data, the SV 4 ‘manufacturer and types’ directory was also requested from the Federal Motor Transport Authority. From this data base, inventories on the manufacturer code number and type code number for further car-type-related characteristics[12] can be included (e.g. for purposes of imputation of the star rating for vehicles which are not assessed by Euro NCAP).

---

[12] E.g. allocation date of the type code number, body type, engine power or technically permissible maximum mass.

3.2 Preparation of Euro NCAP data

Since the start, the Euro NCAP procedure has been subjected to constant change in many ways. At regular intervals, the measurement procedures on which the data are based, the test procedures and the scaling have changed. The consequence of this is that the results across the years are only very limitedly comparable with each other. A fundamental step in the development of an SPI on vehicle safety therefore consists of newly preparing the data in such a way that a higher level of compatibility between the individual years can be achieved.

Depending on the characteristics from which the SPI was constructed – there are various alternatives for this which are listed below – different preparation steps are required. However, the adaptation of the test results from the various years with regard to the test procedures which have changed over the course of time is essential. This will be discussed in more detail in the next section.

3.2.1 Consideration of changes to the test programme

The greatest challenge in terms of creating an SPI time series from Euro NCAP data consists of the fact that over the years, the number and implementation of the test procedures have changed, which was generally due to the increased requirements on safety of the cars to be assessed. The most important changes will first be displayed in brief:

Changes in the area of passive safety

Significant revisions in the area of passive safety took place in the years 2015 and 2020 in particular. Previously, between 2013 and 2015, in Box 3 (until 2017: Pedestrian Protection) the protocols on head, leg and hip impact were reworked (e.g. GRID procedures, new leg test specimens Flex PLI; for more details see ZANDER et al. 2015).

In the year 2015, a new test procedure was introduced with the ‘frontal impact across the full width’ (full width test). Restraint systems are also tested for smaller occupants using female dummies – and on the rear seats. The test places high demands on the restraint systems of the front and back sitting positions.

In the pole impact test, the impact speed of 29 km/h was increased to 32 km/h and the impact angle was changed from 90° to 75°. In addition, areas at the front and at the rear were defined which the head of an occupant could hit in a crash (from smaller women to larger men). These zones are transferred to a fully inflated head airbag and there it is checked if sufficient cushioning is provided. In the lateral impact test, a heavier crash sled is used (1,300 kg instead of 950 kg).

Since then, a new crash test dummy is used in the two lateral impact tests (WorldSID instead of ES2). This shows more measuring areas/points and its movement behaviour has been designed to be more human.

Since 2016, the assessment of Child Occupant Protection has been based on larger dummies which represent children at the ages of six and ten years old. Previously, dummies in child safety seats were used which represent an eighteen-month-old child and a three-year-old child.

In 2020, the ODB test (offset deformable barrier) was replaced by the MPDB test (mobile progressive deformable barrier). Instead of the HIII Dummy, a THOR Dummy is used and instead of a firmly mounted crash element, a rolling element is used. In this way, the dynamics of a real crash can be better depicted using two elements. What is more, the THOR dummy has been significantly improved in terms of its biofidelity.

A completely new test procedure was introduced again in 2020 with the far side test. Previously, only the occupants facing the impact had been taken into account, with this test, occupants who are facing away from the impact are also included.

In the lateral impact test, changes were made once again in the year 2020 to the effect that the speed was increased from 50 km/h to 60 km/h and the mass of the rolling sled was increased from 1,300 kg to 1,400 kg.

Changes to the active safety systems

The most significant changes in the area of active safety tests were the addition of lane-assist systems and autonomous emergency braking systems into the assessment programme (both in the year 2014). The testing of autonomous emergency braking systems for pedestrians and cyclists began in 2016 and 2018. In detail, there were further more or less major adjustments to the test protocols such as, for example, the consideration of the ‘speed limit recognition’ function as part of the assessment of speed assistance systems (2018). In addition, the distribution of points to the individual systems was readjusted several times as Tab. 5 shows.

Another important change consists of points no longer being given for the ‘Electronic Stability Program’ (ESC) since 2016, because legal requirements now stipulate that all new vehicles must be equipped with this system.

Roadmap 2025

This process of the adjustment of the Euro NCAP procedure to technical vehicle developments – whereby vehicle technology, of course, is also driven by Euro NCAP to a certain extent – will continue into the future and may possibly become stronger[13]. The further development of Euro NCAP in the next few years is already being planned, the respective suggestions have been summarised in a so-called 'roadmap' (available at euroncap.com).

The fundamental scheme of the assessment in terms of the four boxes will not change, just like the resulting overall rating. However, in the test procedures there will be a certain paradigm change: 'During the coming years, a transition is foreseen from a "technology based" approach (e.g. tests for AEB) to a more "scenario based" assessment that would allow various types of interventions' (EURO NCAP 2017, p. 7). An example of this is the combined assessment of automatic brake and steering systems in critical situations which is planned for 2022 (small overlap-AEB tests). In addition, in the field of active safety, there have been further tests on emergency braking assistance with regard to junctions (Junction Assist, Box 4) and the protection of pedestrians when reversing (Box 3) since 2020. The 'AEB City' test which was placed in Box 1 (Adult Occupant Protection) has gone in 2020. Also in 2020, in Box 4, the assessment of systems for monitoring the state of drivers and occupants has been added (e.g. alcohol, tiredness). From 2024, vehicle communication systems, i.e. the exchange of data with other vehicles or the environment, are also to be included in the assessment.

In the field of passive safety, a review of the whiplash protocol took place in 2020 (Box 1). For 2023, adjustments in the tests and procedures for assessing the injury severity to pedestrians and cyclists are planned (Box 3).

As from 2020, assessments in terms of tertiary safety will also be carried out for the first time. In Box 1, functions will be tested which provide rescue services with information for a fast and safe rescue of trapped occupants (Rescue, Extrication and Safety) in the case of a crash. As from 2022, (alarm) systems are to be tested which recognise if children are alone in the vehicle (Child Presence Detection).

Finally, in Tab. 5, the maximum achievable number of points in the individual test boxes for the years from 2009 up to and including 2020 are summarised in the form of an overview. In this context, it must be noted once again that in the display of the test results for each of the respective test boxes, the short description 'percentage value' is understood to be the *achieved* number of points in relation to the *achievable* number of points.

---

[13] An assessment of functions in connection with autonomous driving is being considered, but is first to take place outside of the usual assessment grid.

Box	Test	2009-2012	2013	2014	2015	2016-2017	2018-2019	2020
		Maximum achievable number of points						
1	ODB frontal impact	16	16	16	8	8	8	–
	MPDB frontal impact	–	–	–	–	–	–	8
	Full width front impact	–	–	–	8	8	8	8
	Lateral impact	8	8	8	8	8	8	6
	Pole impact	8	8	8	8	8	8	6
	Far side test	–	–	–	–	–	–	4
	Whiplash front seat	4	4	2	2	2	1.5	3
	Whiplash rear seat	–	–	1	1	1	0.5	1
	AEB City	–	–	3	3	3	4	–
	Rescue and recovery	–	–	–	–	–	–	2
	Total	36	36	38	38	38	38	38
2	Performance of CRS (frontal)	16	16	16	16	16	16	16
	Performance of CRS (lateral)	8	8	8	8	8	8	8
	Installation test of CRS[1]	12	12	12	12	12	12	12
	Vehicle equipment	13	13	13	13	13	13	13
	Total	49	49	49	49	49	49	49
3	Head impact	24	24	24	24	24	24	24
	Upper leg impact	6	6	6	6	6	6	6
	Leg impact (knee and lower leg)	6	6	6	6	6	6	6
	AEB pedestrians	–	–	–	–	6	6	7
	AEB cyclists	–	–	–	–	–	6	9
	AEB pedestrians rear	–	–	–	–	–	–	2
	Total	36	36	36	36	42	48	54
4	Seatbelt reminder	3	3	3	3	3	3	2
	Occupant status monitoring	–	–	–	–	–	–	1
	Speed assistant	1	3	3	3	3	3	3
	Lane assistant	–	–	1	1	3	4	4
	ESC	3	3	3	3	–	–	–
	AEB rear impact	–	–	3	3	3	3	4
	AEB junction	–	–	–	–	–	–	2
	Total	7	9	13	13	12	13	16

[1] According to euroncap.com introduction in the year 2013; however points were already allocated earlier and included in the total.

Tab. 5: Euro NCAP point system 2009 - 2020 (source: euroncap.com)

Summary

For the revisions of the test program described above, there is no sufficiently ‘objective’ (i.e. well-founded and derivable from the test practice) procedure with which these changes can be mathematically integrated into the new estimation of the assessment. In particular in the area of passive safety testing therefore, expert judgements must be drawn upon to gain a sufficiently valid quantitative estimation of how these changes to the test procedures affect the vehicle rating. As changes to the test procedures tend to mean their intensification, this specifically involves the estimation of reduction factors by which the assessment results for earlier test years will be lowered. If the 2020 test conditions are observed as reference situation, the corrected test results correspond to the hypothetical case that, in the years until 2019, the test procedures of the year 2020 would already have been valid.

The changes in 2020 described above have also already been included in the expert judgements, even though only Euro NCAP data up until and including 2019 have been observed in this project.

This means that reduction factors will also be imposed on the results for 2019. Thus, the requirements for a possible continuation of the time series have been created for at least the years 2020, 2021 and 2022. The results of the panel of experts in the form of proportional reductions have been summarised in Tab. 6. Details on the calculation of the reduction factors are available in the tables in the appendix.

Test year	Box 1	Box 2	Box 3	Box 4
	Reduction of the original percentage value (share of achieved points) by … %			
2009	41.9	12.2	36.7	66.8
2010	41.9	12.2	36.7	66.8
2011	41.9	12.2	36.7	66.8
2012	41.9	12.2	36.7	66.8
2013	41.9	9.8	34.5	57.4
2014	37.1	7.3	33.8	39.5
2015	20.0	7.3	33.3	39.5
2016	20.0	0.0	24.4	28.9
2017	20.0	0.0	24.4	28.9
2018	20.0	0.0	13.3	17.4
2019	20.0	0.0	13.3	17.4
2020	0.0	0.0	0.0	0.0

Tab. 6: Result of the expert judgement of the Euro NCAP protocol changes since 2009 (reference year: 2020)

The reduction factors are applied to the *share* of achieved points (so-called percentage value). For example, if a vehicle tested in 2011 had achieved 78% of the possible points in Box 1, the new percentage value would lie at 78% x (1-0.419) = 45%. Generally speaking, this is about determining how the models tested between 2009 and 2019 would have been rated under the test criteria and test conditions applicable in 2020.

The resulting box-specific percentage values with reduction factors are comparable across the test years and can be directly used as basis for an SPI.

From the year 2022/23, further expert judgements may be necessary in order to be able to quantify the consequences of the changes which will come into effect then.

3.2.2 Consideration of the changed weighting scheme

As described above, since 2009, a percentage value has been defined in the Euro NCAP Test procedure as starting point for each of the four boxes in which the share of the achieved number of points is expressed using the maximum possible number of points. The overall rating of the vehicle then initially results from the weighted average of the four box-specific percentage values. Thereby, since 2014, the result of Box 1 has been included in the overall rating with a weighting of 0.4 and Boxes 2, 3 and 4 with a weighting of 0.2 each. Between the introduction of the overall rating in 2009 and the year 2013, the weighting scheme stipulated a weighting of 0.5 for Box 1 (Box 2 and 3: 0.2 each; Box 4: 0.1).

If the overall percentage value, i.e. the weighted average of the box-specific percentage values, is to be used as basis for an SPI, a consistent weighting scheme should be taken as basis for the formation of the overall rating from the four boxes. This means that the overall percentage value for the years before 2014 should be modified using the currently applicable scheme.

The procedure is illustrated here using the Dacia Duster, which was tested in 2011, as an example. This car achieved the following test results (percentage values) – for purposes of better comprehensibility, the example is based on the original data, i.e. without any other preparation steps:

- Adult Occupant Protection: 74%
- Child Occupant Protection: 78%
- Pedestrian Protection: 28%
- Safety Assist Systems: 29%

According to the weighting valid at that time, the result is an overall value of

74x0.5 + 78x0.2 + 28x0.2 + 29x0.1 = 61.1%.

If the currently applicable weighting scheme were to be applied, one would get the following new overall rating

74x0.4 + 78x0.2 + 28x0.2 + 29x0.2 = 56.6%.

The more stronger consideration of the ‘Safety Assist Systems’ box in the current scheme, which was assessed as comparably low in this model, leads to a decline of the overall percentage value here. If the reduction factors for the year 2011 from Tab. 6 are included, one receives values of 39.7% (former weighting) or 36.4% (current weighting).

3.2.3 Consideration of the changed assessment scheme

In Euro NCAP, the allocation of the 'stars' is carried out using a table with threshold values (for the previously mentioned percentage values), which have to be achieved as a minimum for a certain number of stars (1 to 5). There are such threshold values for both the four individual boxes as well as for the overall rating. Star ratings for the individual boxes have not been explicitly shown in Euro NCAP since 2009, the box-specific thresholds primarily serve the purpose of 'balancing' (see next section). Nevertheless, one can also draw upon the threshold values per box in order to depict a star assessment for the respective box – as shown below.

These threshold values for the allocation of star numbers for the achieved percentage values have been changed or increased many times over the years (see EURO NCAP 2013, p. 20; EURO NCAP 2020, p. 15).

If one wishes to use the star ratings for the individual boxes as basis for an SPI then, in order to improve the comparability, a consistent scheme should be used for the allocation of the stars. Therefore, a new estimation (of the number of stars) is also required here.

It seems obvious that the allocation scheme valid for the years 2020 to 2022 should be used as consistent assessment standard. The respective threshold values which have applied since 2018 are displayed in Tab. 7.

Number of stars	Box 1	Box 2	Box 3	Box 4	Overall rating
	Minimum percentage value to be achieved				
5	80	80	60	70	74
4	70	70	50	60	64
3	60	60	40	50	54
2	50	50	30	40	44
1	40	40	20	30	34

Tab. 7: Euro NCAP assessment scheme for the period 2020 to 2022 (source: EURO NCAP 2020, p. 15)

When comparing the threshold values specified in Tab. 7 for Box 4 with, for example, the values from the year 2012, the increased requirements are visible: at that time, only 60% was necessary for a rating of 5 stars and 40% for 4 stars; in the same way, 25% was sufficient for 3 stars, 15% for 2 stars and only 5% for 1 star (EURO NCAP 2013, p. 20). In contrast, in order to receive at least 1 star in Box 4 under the current scheme, at least 30% needs to be achieved (Tab. 7). In this way, for example, the Citroën e-Méhari earned a result of 25% in the tests on active systems (Box 4) which would have been (only just) sufficient for a rating of three stars in test year 2017. According to the currently valid scheme, this model would not receive any star in this box.

3.2.4 Balancing

In Tab. 7, threshold values are also included for the allocation of stars to the (weighted) overall percentage value.

If the overall star rating is to be used as basis for an SPI, then – analogue to the Euro NCAP procedure – a so-called balancing can be taken into account whereby the minimum of overall percentage value and box-specific percentage values is decisive for the number of stars.

For example, if a vehicle reaches a result which leads to a rating of 5 stars in the overall average, this vehicle will still only receive 4 stars if only the requirements for a 4-star assessment are fulfilled for at least one box (see, e.g. BASt 2009, p. 7f). This has the aim of ensuring that, in the case of comparatively low ratings in individual boxes of the test program, the overall result is devalued.

The functioning and effect of this changed allocation systematic can be demonstrated using the BMW Z4 (test year 2015) as an example. This car achieved a total of 3 stars for the following individual results (for purposes of better comprehensibility, the example is based on the original data, i.e. without any other preparation steps):

- Adult Occupant Protection: 69%
- Child Occupant Protection: 61%
- Vulnerable Road Users: 91%

- Safety Assist Systems 46%

The weighted overall percentage value estimated using the individual boxes amounts to 67.2% (see section 3.2.2 for estimation), which would have led to a 4 star rating according to the assessment scheme valid at that time. However, according to the scheme at that time, in Boxes 1 and 4, the box-specific threshold values for 4 stars (70% and 60%) were not achieved, which results in the downgrading of the overall rating to 3 stars. If the assessment scheme for 2020 were to be used (Tab. 7) on the results of this vehicle model, one can see that now in Box 4 (Safety Assist Systems) the threshold value for 3 stars (50%) has not been achieved. Thus, in the newly estimated (revised) overall rating for this model, only 2 stars have been allocated.

As from the year 2023, the threshold values for the star allocation in Box 3 (Vulnerable Road Users) will be increased. Then, for example, at least 70% will be required for a 5 star rating and at least 30% for a rating with 1 star. This also results in an adjustment of the threshold values for the overall rating (76% for 5 stars, 36% for 1 star; EURO NCAP 2020, p. 15). In the case of a continuation of the SPI time series from 2020, it is recommended that the currently used 2020/22 scheme is kept for the time being, at least as a reference.

3.2.5 Summary and results

In Figure 2, the preparation steps of the Euro NCAP data for the various SPI alternatives are summarised again in graphical form.

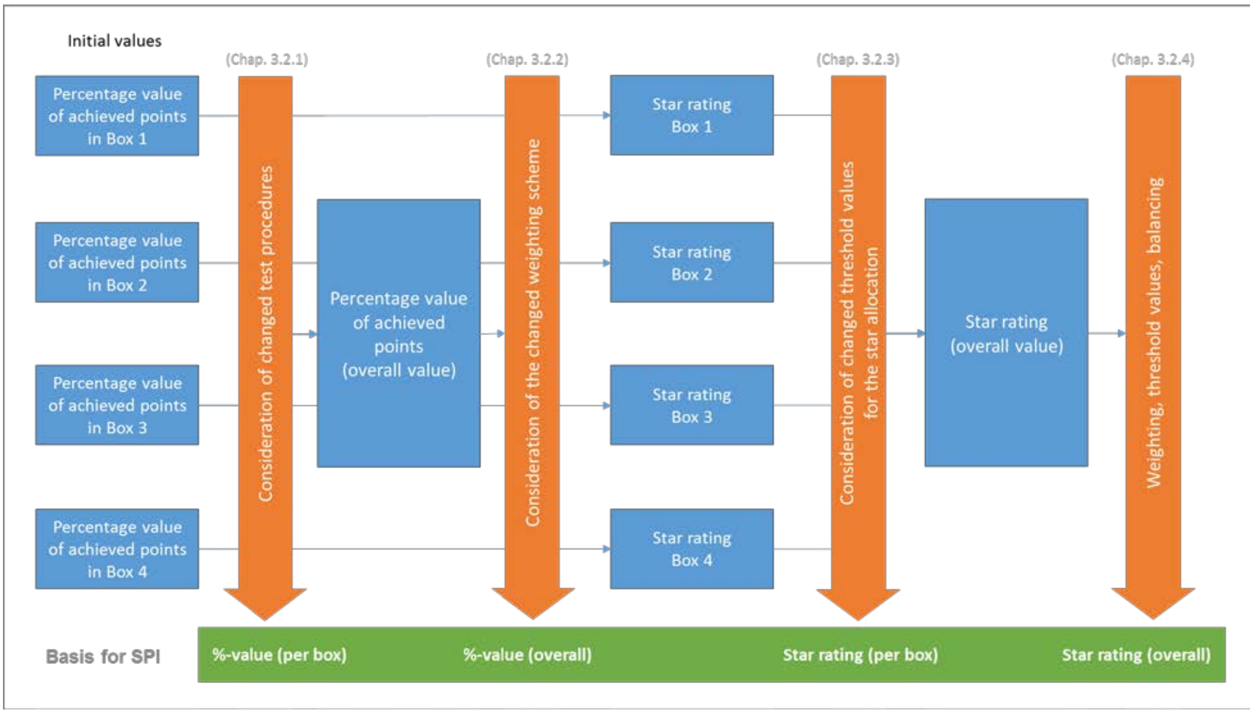


Figure 2: Options for constructing an SPI using Euro NCAP data

Below, the effects of the data preparation steps described above (change of test procedures, weighting of the boxes, threshold value for the star allocation) on the safety rating of the tested vehicle models are presented. For this purpose, the results (mean values across all 437 tested models) for the individual SPI options (share of achieved points, number of stars) per test year are contrasted with the original results.

Tab. 8 shows the comparison of results in view of the SPI alternative ‘share of achieved points’.

Year of the test	Box 1		Box 2		Box 3		Box 4		Overall rating	
	orig.	new	orig.	new	orig.	new	orig.	new	orig.[1]	new[2]
	Mean percentage value (average share of achieved points)									
2009	85.8	49.9	78.1	68.5	46.3	29.3	75.9	25.2	75.4	44.5
2010	85.8	49.8	77.5	68.0	53.9	34.1	73.3	24.3	76.5	45.2
2011	86.6	50.3	80.3	70.4	48.8	30.9	77.8	25.8	76.9	45.6
2012	85.4	49.6	82.8	72.7	59.9	37.9	74.0	24.5	78.6	46.9
2013	89.0	51.7	80.4	72.5	65.2	42.7	71.9	30.7	80.8	49.9
2014	80.2	50.4	80.5	74.6	65.3	43.2	60.4	36.5	73.3	51.0
2015	85.4	68.3	82.5	76.5	74.1	49.4	70.8	42.8	79.7	61.1
2016	88.0	70.4	78.4	78.4	68.7	51.9	55.3	39.3	75.7	62.1
2017	86.0	68.8	75.8	75.8	68.8	52.0	55.5	39.4	74.4	61.0
2018	88.3	70.6	81.6	81.6	68.7	59.6	69.8	57.6	79.3	68.0
2019	91.5	73.2	85.1	85.1	70.7	61.2	72.9	60.2	82.3	70.6

[1] Weighted according to the respective valid weighting scheme

[2] All years weighted according to the weighting scheme valid from 2014 (0.4 : 0.2 : 0.2 : 0.2)

Tab. 8: Mean percentage value (average share of achieved points) before and after data preparation grouped into box and test year (source: own estimations – data basis: Euro NCAP 2009 - 2019)

Within the boxes, the mean differences between the original and the newly estimated results correspond to the reduction factors represented in Tab. 6. The largest differences between original and new value can be found in Box 4 accordingly.

In Tab. 9, the results are compared in terms of the SPI alternative ‘number of stars’.

Year of the test	Box 1		Box 2		Box 3		Box 4		Overall rat.[1]	
	orig.	new[2]	orig.	new[2]	orig.	new[2]	orig.	new[2]	orig.	new[2 3]
	Average number of stars									
2009	4.91	1.53	4.96	3.38	5.00	1.47	4.87	0.11	4.82	0.09
2010	4.70	1.56	4.78	3.30	4.96	1.82	4.74	0.07	4.59	0.07
2011	4.80	1.59	4.89	3.57	4.96	1.57	4.91	0.09	4.72	0.09
2012	4.57	1.50	4.91	3.73	4.52	2.39	4.68	0.16	4.48	0.16
2013	4.88	1.72	5.00	3.72	4.78	2.81	4.81	0.50	4.69	0.50
2014	4.54	1.59	4.88	3.93	4.83	2.78	4.27	1.24	4.27	1.22
2015	4.81	3.26	4.88	4.17	4.86	3.48	4.57	1.71	4.57	1.71
2016	4.94	3.50	4.78	4.50	4.94	3.72	4.44	1.72	4.39	1.67
2017	4.68	3.36	4.65	4.13	4.76	3.65	4.39	1.73	4.32	1.69
2018	4.68	3.60	4.68	4.68	4.64	4.44	4.40	3.44	4.36	3.36
2019	4.98	3.76	4.93	4.93	4.71	4.44	4.69	3.64	4.66	3.55

[1] Incl. balancing.

[2] For the star allocation of the newly estimated shares of achieved points, the threshold values from 2020 were applied.

[3] Weighted according to the weighting scheme from 2014 (0.4 : 0.2 : 0.2 : 0.2).

Tab. 9: Average number of stars before and after data preparation grouped into box and test year (source: own estimations – data basis: Euro NCAP 2009 - 2019)

Overall, as a result of the revision of the estimation scheme with the aim of an improved compatibility of the results from the different years, there are noticeable rating changes, in particular in Box 4 (Safety Assist Systems). The mean star rating in this box between 2009 and 2013 lies below the value of 1. Due to the balancing mechanism in the overall rating, this necessarily leads to very similar results there.

If the overall rating (stars) is observed across all years, none of the 437 tested models achieve a 5-star rating after the re-estimation.

4 Extrapolation of the Euro NCAP data and safety indicator development

4.1 Extrapolation

4.1.1 Basic concept

The extrapolation[14] of the Euro NCAP results to the overall car fleet is the key step in generating a Safety Performance Indicator (SPI). As the indicator is to be formed for subsequent years (so-called SPI reference years), the annual car inventory data from the Central Vehicle Register of the Federal Motor Transport Authority is to be observed as extrapolation frame.

The extrapolation consists of matching the – newly estimated – Euro NCAP result data for the tested makes and models (share of achieved points, number of stars) with the respective Central Vehicle Register inventories. As only the vehicles tested in Euro NCAP from 2009 are taken into account (see chap. 3.1.1), only very few tested models are available in the inventory after 2009. For this reason, the decision was made to only start with the time series for the SPI from the year 2014[15] in order to have a sufficient number of tested models in the car fleet available and thus to be able to allocate as much of the vehicle stock as possible to the test results.

The data preparation work connected with this procedure is considerable: if the individual Central Vehicle Register data for the SPI reference years 2014 to 2020 (reference date 1 January respectively) is merged into one overall file[16], the overall number of the vehicle data records to be processed totals almost 317 million. The data was provided by the Federal Motor Transport Authority in an aggregated form, i.e. according to the Central Vehicle Register characteristics listed in section 3.1.2. These aggregated data records contain a total of 2,106,510 rows of data, whereby every data row corresponds to a certain combination of characteristics.

4.1.2 Methodology

The matching of the Central Vehicle Register Data with the Euro NCAP data is carried out via the characteristics of make and model[17]. In addition, it was also necessary to identify the series (model generation) of the respective model tested in Euro NCAP in the Central Vehicle Register data (e.g. make: Mercedes-Benz; model: B-Class; model generation: 2011-2018 or 2019ff (rated twice in Euro NCAP: 2011 and 2019)). However, this information is not available in the Central Vehicle Register data. Therefore, in addition to the characteristics of make and model, the year of initial registration also needed to be drawn upon in order to be able to establish a match with the Euro NCAP data.

In order to identify the make and model in the Central Vehicle Register, the characteristics

- make number/make description and
- model code/model description

were used. The procedure cannot be used for only around 2.7% of the vehicles as the model code is unknown.

However, the two above-mentioned Central Vehicle Register characteristics are not always sufficient to be able to identify a model assessed in Euro NCAP in the overall inventory. For this reason, additional characteristics were drawn upon for matching when required:

- On the one hand, this concerns the body type – in order to distinguish between, for example, convertibles (e.g. Golf convertible) or estates (e.g. Volvo S60 vs. V60) and on the other hand the code for source of fuel/power[18] in order to identify electric or hybrid vehicles (e.g. Kia Soul EV).
- In addition, plain text entries (such as manufacturer’s description (not normed), trade name and description) also had to be used to identify a model in some cases. Examples of these are Audi A3 e-tron, Qoros 3 Sedan, Golf SV, Citroën Spacetourer or the MINI make for which all models have the same model code.

Surprisingly, not all tested models could be found in the Central Vehicle Register. Of the five models

---

[14] The term ‘extrapolation’ entails the concept that for every SPI reference year, the n tested vehicles represent a sample from the entirety of all N vehicles of the respective car fleet.

[15] The vehicle stocks for the years 2009 to 2013 are also available.

[16] Vehicles which are registered in the Central Vehicle Register over several years are represented in this file with several data records.

[17] In Euro NCAP, only n_k=1 vehicle is normally tested per make and model k, whereby the test results here are transferred to the N_k corresponding vehicles in the vehicle stock.

[18] Both characteristics were included through the manufacturer and type code number from the ‘manufacturer and types’ (SV 4) directory. Thus, these characteristics are unknown for vehicles with no type.

of the MG make, only one (MG ZS) could be identified in the car inventory data of the years 2014 to 2020. In addition, the following models were not identifiable: Aiways U5 (only available from 2020), Geely Emgrand EC 7, Fiat Panda Cross, Toyota Proace City (only available from 2020), Isuzu D-Max and the Mercedes-Benz X-Class. In the case of the two latter models, these are vehicles in the N1 class according to the manufacturer and types directory.

In order to also be able to identify the series or model generation of each of the tested models in the Central Vehicle Register, extensive Internet research had to be carried out. For this purpose, for example, online car catalogues from the ADAC automobile association and from the Autobild automobile journal were drawn upon. As mentioned above, in order to allocate the model generation in the Central Vehicle Register Data, the characteristic ‘Year of initial registration’ was used. This means that for every tested model, the determined model generation (e.g. Golf VII) was represented using the summary of the respective years of initial registration. This procedure is based on the assumption that a vehicle which is first registered in the year x belongs to the model generation which is the latest edition in the year x. This assumption will not be correct in all cases, however the expected error is assumed to be very small.

When representing the model generation in the Central Vehicle Register, facelifts were also taken into account where relevant. In this way, for example, the test of the Mazda 6 in the year 2018 obviously refers to the 2nd facelift of the GJ series (in the year 2018). Correspondingly, these models were allocated to the years of initial registration from 2018 in the inventory. However, the first facelift in 2015 was not taken into account as, according to the facelift review of December 2014, the test result for the basic model of the GJ series from the year 2013 was still valid. The model tested in 2013 is therefore represented in the Central Vehicle Register using the years of initial registration 2013 to 2017. A further example is the 1st generation of the Citroën C1 (tested in 2012 and 2014), which was built from 2005 to 2014. However, a facelift took place in 2012[19], so that in the Central Vehicle Register only the years of initial registration 2012 to 2013 were allocated to the model tested in Euro NCAP in 2012. On the one hand, this leads to a more precise representation, but on the other hand increases the share of the vehicle stock without assignable assessment.

4.1.3 Matching quota

The result of the allocation of test results to the car fleet vehicles is observed according to separate inventory reference dates[20] below. As can be expected, the matching quota, i.e. the share of cars *with* Euro NCAP rating, rises from year to year for newer vehicles (initial registration from 2009). While on the reference date 1.1.2014, only a good half (54.8%) of the newer cars registered in the Central Vehicle Register could be allocated to a Euro NCAP rating, this already applies to more than three quarters (78,1%) of the car fleet on the reference date 1.1.2020.

On average, across the seven inventory reference dates observed here (1.1.2014, 1.1.2015, 1.1.2016, 1.1.2017, 1.1.2018, 1.1.2019, 1.1.2020), 69.7% of the newer vehicles (initial registration year from 2009) were allocated a Euro NCAP rating. As the reference-date-specific matching quota for cars which were initially registered before 2009 lies below 3% respectively[21], the mean matching quota for 2014 to 2020 is around 36%[22]. This is due to the fact that only Euro NCAP test results from 2009 are observed and the tests were quite predominantly carried out close to the market introduction of the respective model or the respective model generation.

In Tab. 10, the (reference-date-specific) matching quotas for the car start of the year inventories 2014 to 2020 are represented in the subdivision according to year of initial registration.

---

[19] Source: ADAC automobile catalogue (https://www.adac.de/rund-ums-fahrzeug/autokatalog/marken-modelle/)

[20] The inventory reference date is the day to which the vehicle inventory in the Central Vehicle Register – and thus also the SPI – refers (in this case, 1 Jan. of each year).

[21] If a model that was tested from 2009 was already on the market before 2009, the respective initial registration year before 2009 was also taken into account in the inventory (e.g. Honda Accord, 8th generation: tested in 2009, on the market since summer 2008).

[22] The vehicle stock across all reference dates observed here amounts to around 316.92 mio vehicles, of which 113.72 mio cars (35.9%) were able to be allocated to a Euro NCAP rating. Vehicles which are in the inventory over several years are counted several times accordingly.

Reference date Central Vehicle Register inventory	Initial registration up to 2008	Initial registration from 2009
	Share of cars with Euro NCAP rating in %	
1.1.2014	2.2	54.8
1.1.2015	2.3	60.0
1.1.2016	2.4	64.9
1.1.2017	2.5	69.1
1.1.2018	2.6	72.6
1.1.2019	2.7	75.6
1.1.2020	2.8	78.1
Average	2.5	69.7

Tab. 10: Share of vehicles in the Central Vehicle Register with assignable Euro NCAP rating grouped into inventory reference date and year of initial registration (source: own estimations – data basis: Euro NCAP 2009 - 2019; Central Vehicle Register 2014 - 2020)

Tab. 11 shows the mean matching quota for the period 2014 to 2020 (with the vehicle stock weighted arithmetic mean of the seven reference-date-specific matching quotas) in the subdivision according to vehicle segment and year of initial registration.

Car segment	Initial registration up to 2008	Initial registration from 2009
	Mean matching quota 2014 to 2020 in %	
Mini	0	52.9
Small car	0.2	58.4
Compact class	4.3	80.8
Medium-sized class	6.2	81.4
Luxury medium-sized class	1.0	90.2
Luxury class	0	7.1
ATV	0.3	70.9
Utilities	1.3	51.0
Other	0	0
Sports car	0	15.5
Mini van	0	65.5
Large-capacity van	0	52.4
SUV	1.2	85.2
All segments	2.5	69.7

Tab. 11: Mean share of vehicles in the Central Vehicle Register with assignable Euro NCAP rating grouped into segment and year of initial registration (source: own estimations – data basis: Euro NCAP 2009 – 2019; Central Vehicle Register 2014 - 2020)

If only cars which were registered (for the first time) from 2009 were observed, then in terms of the vehicle segment, mean matching rates of between 80% and 90% can be found in the luxury medium-sized class as well as the SUVs, medium-sized and compact classes. In contrast, comparably few sports cars and cars in the luxury class could be allocated to a Euro NCAP rating.

Finally, if the mean matching quota for the inventory reference dates 1.1.2014 to 1.1.2020 are also observed according to make (in turn only cars with initial registration from 2009), the highest quota can be found for the Tesla make with 97.7% followed by Honda with 92.4%. Quotas at around 80% result for the makes Volkswagen (84.8%), Kia (82.1%), Hyundai (80.3%), Volvo (79.3%) and Audi (78.1%).

4.1.4 Imputation

As not all models which are introduced to the market are rated in Euro NCAP, a later supplementation (imputation) of the assessment is carried out for the models which are included in the vehicle stock but are not tested. The same is required for the cases in which the connection of Euro NCAP data and Central Vehicle Register data does not lead to a match.

Imputation means that in the entirety of all cars, the test results (share of achieved points, number of stars) for the vehicles for which a Euro NCAP rating could be allocated through matching can be transferred to the vehicles without safety rating according to certain principles.

The methodology used for this is briefly outlined below. Euro NCAP data are available for the vehicle models tested from 2009 onwards. Therefore, in fact, an imputation is only considered for the models (or model generations) which were introduced from 2009 but were not tested. A transfer to older vehicles, however, does not seem justified due to the development of vehicle technology.

In particular, the vehicle segment and the year of initial registration come into question as possible auxiliary characteristics for a transfer of results. In this case, for every vehicle segment and for every year of initial registration from 2009, the average rating in the subdivision of these two characteristics can be estimated and transferred to the models which were not tested. It is to be noted here, that in Euro NCAP, the majority of vehicles are tested because manufacturers specifically want this rating. It can therefore be assumed that the models which are not tested may possibly lie somewhat below the average. Due to this procedure-related self-selection mechanism, it cannot be ruled out that the ratings generated by the imputation may possibly turn out slightly higher.

When using stars as SPI, this mean value imputation does, of course, lead to values which are not integer for the vehicles in the Central Vehicle Register for which no Euro NCAP rating could initially be allocated. In view of the formation of indicators (see section 4.2) a mean value approach would therefore also be recommended in this case.

In this context, however, it should be noted that the use of the year of initial registration in the imputation procedure entails slight distortion. For example, if a car with year of initial registration 2012 which was not assessed in Euro NCAP is imputed with the mean rating of its segment for vehicles with the year of initial registration 2012, it is not taken into account that this car might belong to a model generation that has already been unchanged on the market for several years (possibly even before 2009).

Also, when using this imputation procedure for the newer vehicles – precisely through the restriction of extrapolation to test results from 2009 – there is only a very low share of vehicles available in the Central Vehicle Register with initial registration before 2009 to which a rating could be allocated (see Tab. 10). In particular at the start of the time series to be created (2014) this leads to a not inconsiderable share of vehicles in the inventory without rating, which is related to the many older vehicles in the inventory. The average age of registered cars in Germany currently lies at around nine years. The models tested between 2009 and 2019 are observed from Euro NCAP – for the reasons stated in section 3.1.1. However in the overall vehicle inventory on 1.1.2020, for example, only around 64% of cars have an initial registration date from 2009. Therefore, the earlier the inventory reference date lies (1.1.2019, 1.1.2018 etc.), the smaller this share is.

Due to the fact that the year 2020 was used as a reference for the standardisation of the Euro NCAP results, it seems justifiable to allocate the rating ‘0 stars’ or ‘0 percent achieved points’ throughout to vehicles which were registered for the first time before 2009 and to which no rating could be allocated.

If the imputation is carried out in the way described here (initial registration before 2009 and missing rating: apply rating 0; initial registration from 2009 and missing rating: mean value imputation according to year of initial registration and segment), 479,476 vehicles remain without rating. That is only 0.15% of the overall inventory 2014 to 2020.

4.2 Indicator development

In principle, from a statistical point of view, the following Euro NCAP-based parameters of the car fleet come into question for the formation of an indicator value (SPI):

- Total values, e.g. the sum of the box-specific NCAP stars of all cars in the inventory
- Ratios (share or relationship figures), such as, for example, the share of 5-star vehicles in the overall vehicle stock or
- Mean values, e.g. the arithmetic mean or the median of the share of the points achieved in the test

At the same time, the resulting indicator value (SPI) should also be interpretable in terms of content.

A very simple indicator would be, for example,

- the average number of stars for the vehicles in the observed inventory (stars per car).

Possible would also be

- to draw upon the sum of the achieved ‘stars’ related to the maximum possible star total (= car inventory multiplied by five).

In the latter case, the resulting ratio would represent something like the degree of utilisation of vehicle safety potential, whereby this would then refer to precisely one reference year.

In the case of the specified procedures mentioned above for the formation of an indicator value (SPI), indices of over 100% might result in future in terms of vehicle safety in comparison with a constant reference year due to the further development of state-of-the-art technology. If this is to be avoided, a redefinition of the reference year should be considered at the appropriate time.

Furthermore, in the use of ratios, such as the share of cars rated with five stars for example, it should be noted that – against the backdrop of the above-outlined imputation method – the imputed mean values should be rounded to integer values in advance.

Depending on the selected SPI, the indicator values are to be estimated by way of an overall rating or on the level of the individual test boxes.

Summary

As Safety Performance Indicator (SPI), the mean box-specific star ratings will be used hereinafter because the clearest results can be achieved with these: on the one hand, the Euro NCAP tests are strongly associated with the star rating, on the other hand, a box-specific observation enables a more customised analysis of the relationship between SPI and the accident occurrence (see chapter 5).

Box-specific average values of the characteristic ‘Number of stars according to Euro NCAP (0 to 5)’ are therefore estimated per observed inventory reference date (from 2014) for the overall car fleet. These are based on the newly estimated and standardised box-specific vehicle assessments from Euro NCAP.

These mean values can be characterised as ‘Box-specific Safety Performance Indicators of the car fleet referenced to 2020’ and are referred to as

SPI (Box x) of the car fleet

hereinafter.

The referencing of 2020 refers to both the test procedures (reduction factors) as well as to the threshold values for the allocation of the stars in the Euro NCAP data.

In this way, a total of four Safety Performance Indicators were formed which are, as mentioned above, box-specific average values of the safety rating of the cars entered in the Central Vehicle Register. Using these four indicators, an Overall Safety Performance Indicator

SPI (overall) of the car fleet

was then also estimated.

This is a weighted mean value in which – analogue to the Euro NCAP procedure – the assessment results of the four boxes are weighted at a ratio of 40:20:20:20.

4.3 Extrapolation results

As part of the extrapolation, the results of the Euro NCAP tests between 2009 and 2019 with the reduction factors defined by the panel of experts were transferred to the overall car fleet and missing values (e.g. models which were not tested) were added using imputation.

Tab. 12 shows the key result of the extrapolation, i.e. the time series of the four box-specific SPI values and the overall SPI.

Reference date Central Vehicle Register inventory	Box 1	Box 2	Box 3	Box 4	Total
	SPI of the car fleet				
1.1.2014	0.60	1.24	0.60	0.02	0.63
1.1.2015	0.70	1.45	0.73	0.04	0.74
1.1.2016	0.82	1.67	0.89	0.09	0.87
1.1.2017	0.94	1.88	1.04	0.14	1.00
1.1.2018	1.09	2.10	1.22	0.22	1.14
1.1.2019	1.24	2.32	1.40	0.31	1.30
1.1.2020	1.41	2.53	1.60	0.45	1.48

Tab. 12: Box-specific SPI and SPI (overall) of the car fleet broken down according to inventory reference date (source: own estimations – data basis: Euro NCAP 2009 - 2019; Central Vehicle Register 2014 - 2020)

The results consistently show a rising trend of the indicators over the course of time. The highest indicator values can be found in Box 2 (Child Occupant Protection), they are lowest in Box 4 (Safety Assist Systems). However the level of the SPI values is relatively low (in relation to the value range 0 to 5). This essentially has two reasons:

On the one hand, this is due to the fact that the inventory on a reference date is, in fact, distributed over many initial registration years. As mentioned above, due to the fact that no (comparable) Euro NCAP results are available before 2009, practically all vehicles with year of initial registration before 2009 were allocated a 0-star rating in the imputation.

On the other hand, an effect was of course caused by the fact that, with regard to the reduction factors for the standardisation of the Euro NCAP ratings (experts), the test conditions of the year 2020 were consistently used as reference. As discussed above (see Tab. 6), for example for Box 4, this causes the test results for the models assessed between 2009 and 2012 to be reduced to around one third of their original value. Of course, this is also reflected in the values projected to the car fleet.

Below, the four box-specific SPI values as well as the SPI overall value are also grouped into car segment according to the Central Vehicle Register. For reasons of clarity, the smaller segment ‘Other’ (vintage cars etc.) has been omitted here.

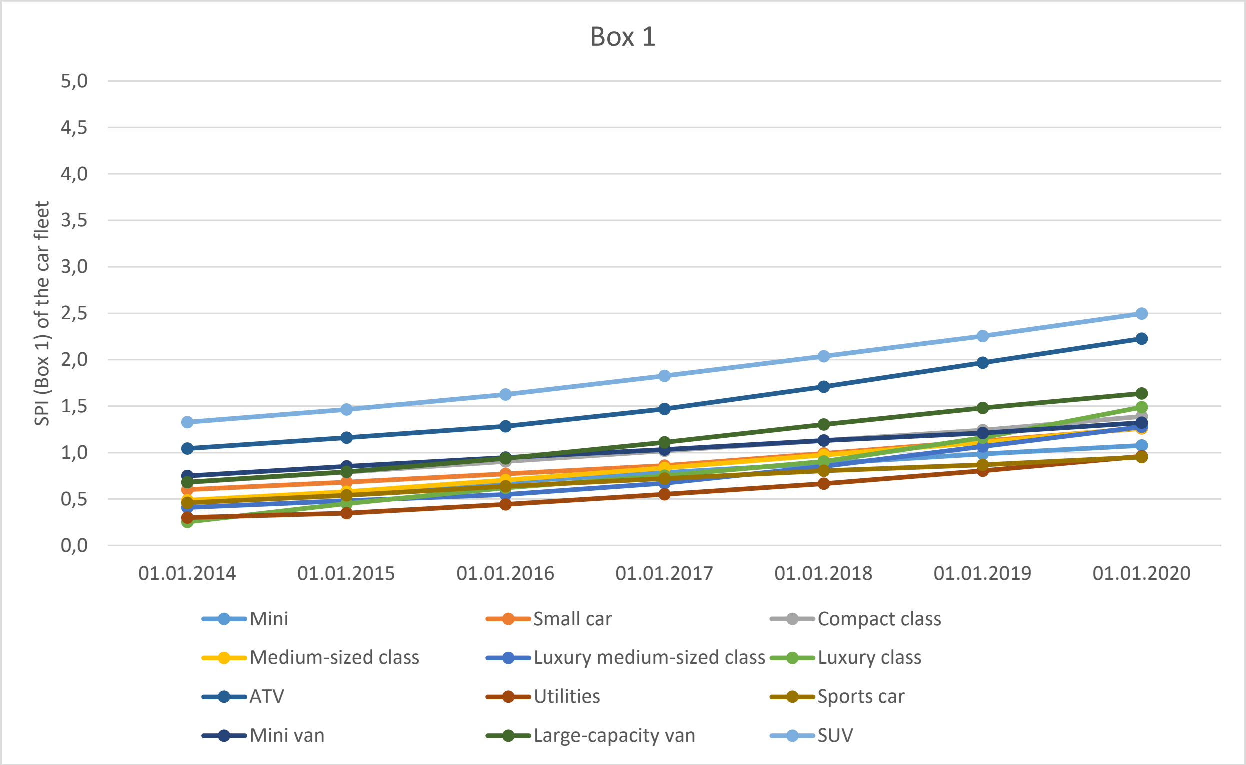


Figure 3: SPI (Box 1) of the car fleet grouped into car segment and inventory reference date (source: own estimations – data basis: Euro NCAP 2009 – 2019; Central Vehicle Register 2014 - 2020)

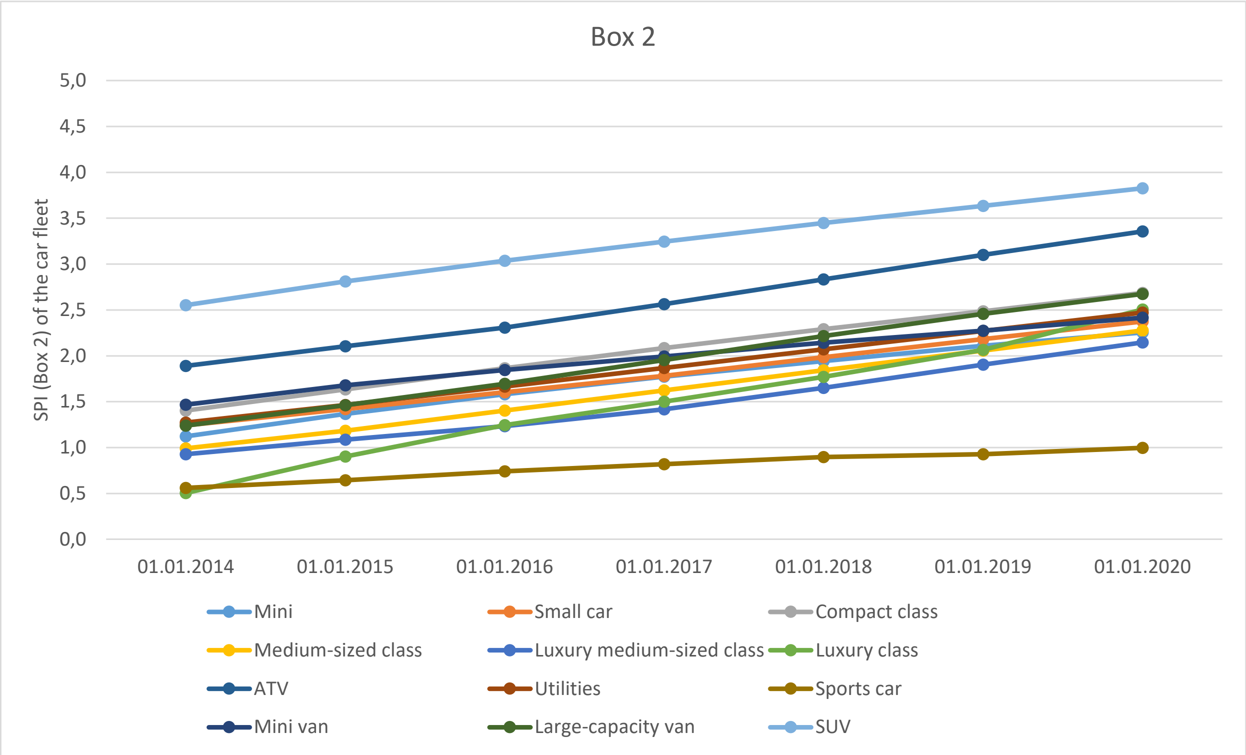


Figure 4: SPI (Box 2) of the car fleet broken down into car segment and inventory reference date (source: own estimations – data basis: Euro NCAP 2009 – 2019; Central Vehicle Register 2014 - 2020)

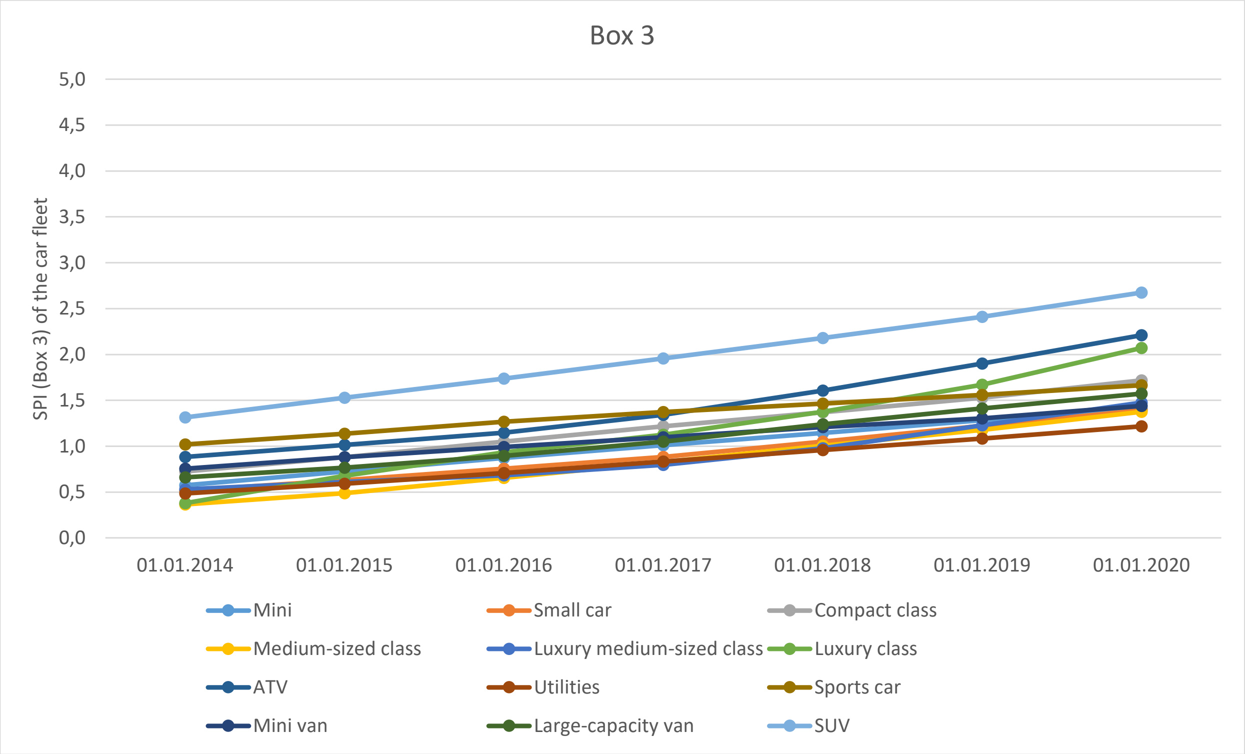


Figure 5: SPI (Box 3) of the car fleet grouped into car segment and inventory reference date (source: own estimations – data basis: Euro NCAP 2009 – 2019; Central Vehicle Register 2014 - 2020)

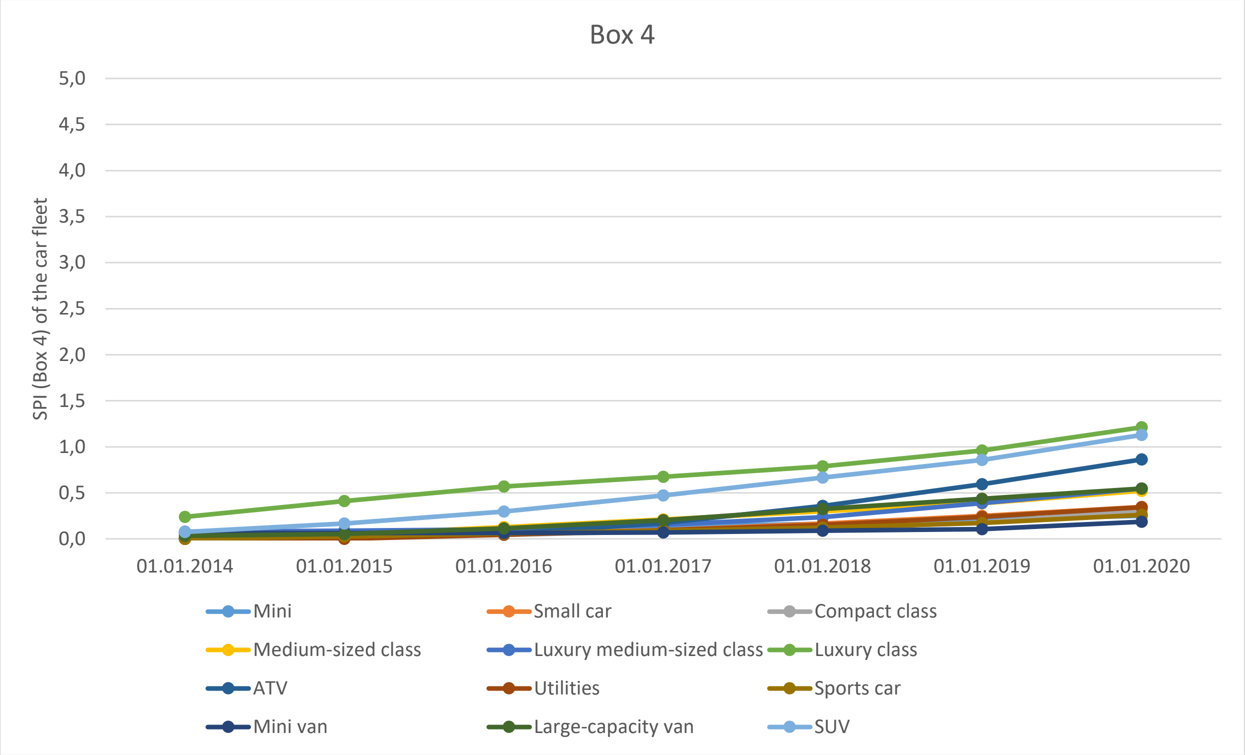


Figure 6: SPI (Box 4) of the car fleet grouped into car segment and inventory reference date (source: own estimations – data basis: Euro NCAP 2009 - 2019; Central Vehicle Register 2014 - 2020)

The chronological development of the SPI in Box 1 (Adult Occupant Protection) is shown in Figure 3. At all observed periods, the SUVs show the highest SPI in Adult Occupant Protection, followed by ATVs.

SUVs rank at the very top also in terms of the SPI in the Boxes 2 (Child Occupant Protection) and 3 (Vulnerable Road Users) as can be seen in Figure 4 and Figure 5. When compared to the SPIs in Box 3, the values in Box 2 show a greater heterogeneity. SUVs and ATVs are at the top by a long way, while sports cars are rated comparably low for Child Occupant Protection. Here too, over the course of time, only a relatively moderate rise in the values can be determined.

In terms of the SPI in Box 4 (Safety Assist Systems) the highest SPI can be found among vehicles in the luxury class in all periods (Figure 6). At second place are SUVs, whereby the distance to the luxury cars is reduced over the course of time.

A concluding glance at the SPI (overall) shows that this also rises over the course of time in all segments, whereby the course runs steepest in the ‘luxury class’ segment (pale green line) (Figure 7).

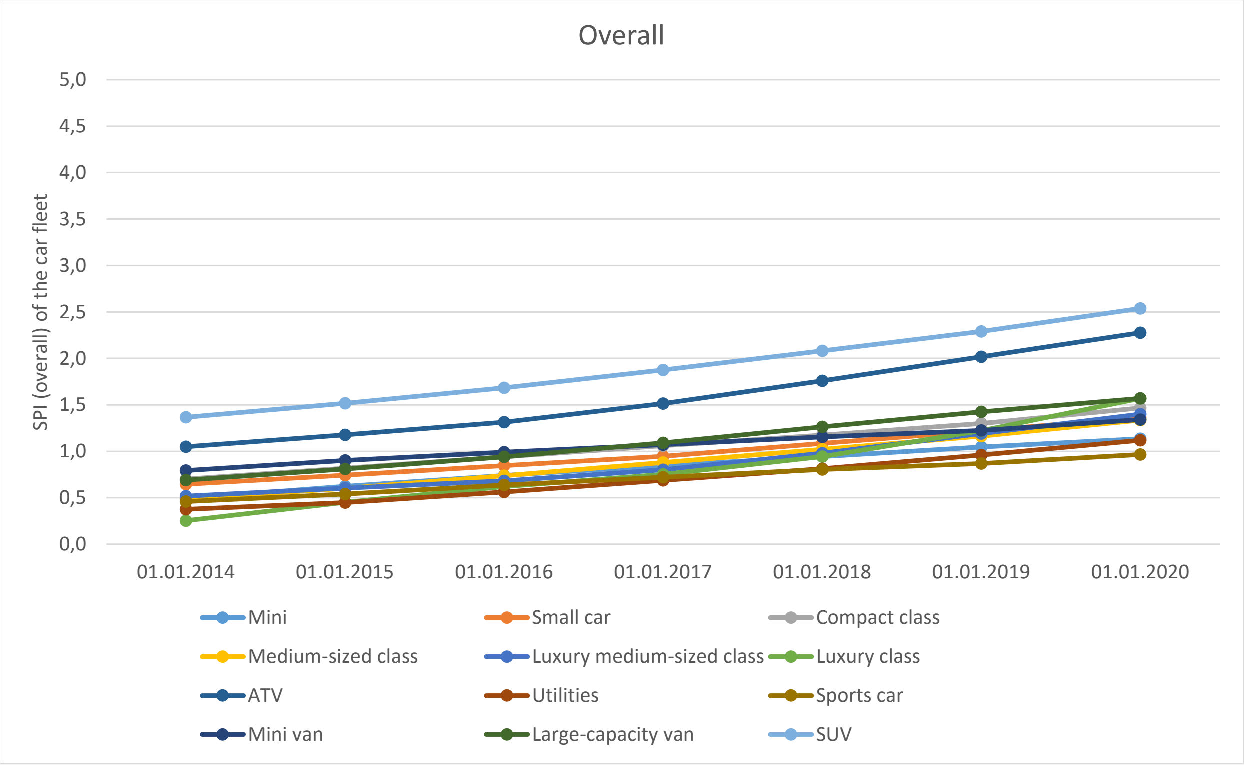


Figure 7: SPI (overall) of the car fleet grouped into car segment and inventory reference date (source: own estimations – data basis: Euro NCAP 2009 - 2019; Central Vehicle Register 2014 - 2020)

The fact that SUVs have relatively high SPI values in all analyses is certainly also related to the fact that this is a relatively new segment in which the share of older vehicles is comparably low. This can be emphasised with an assessment, for example, in which the SPI (Box 3) for the year 2019 is grouped into segment and year of initial registration (Figure 8).

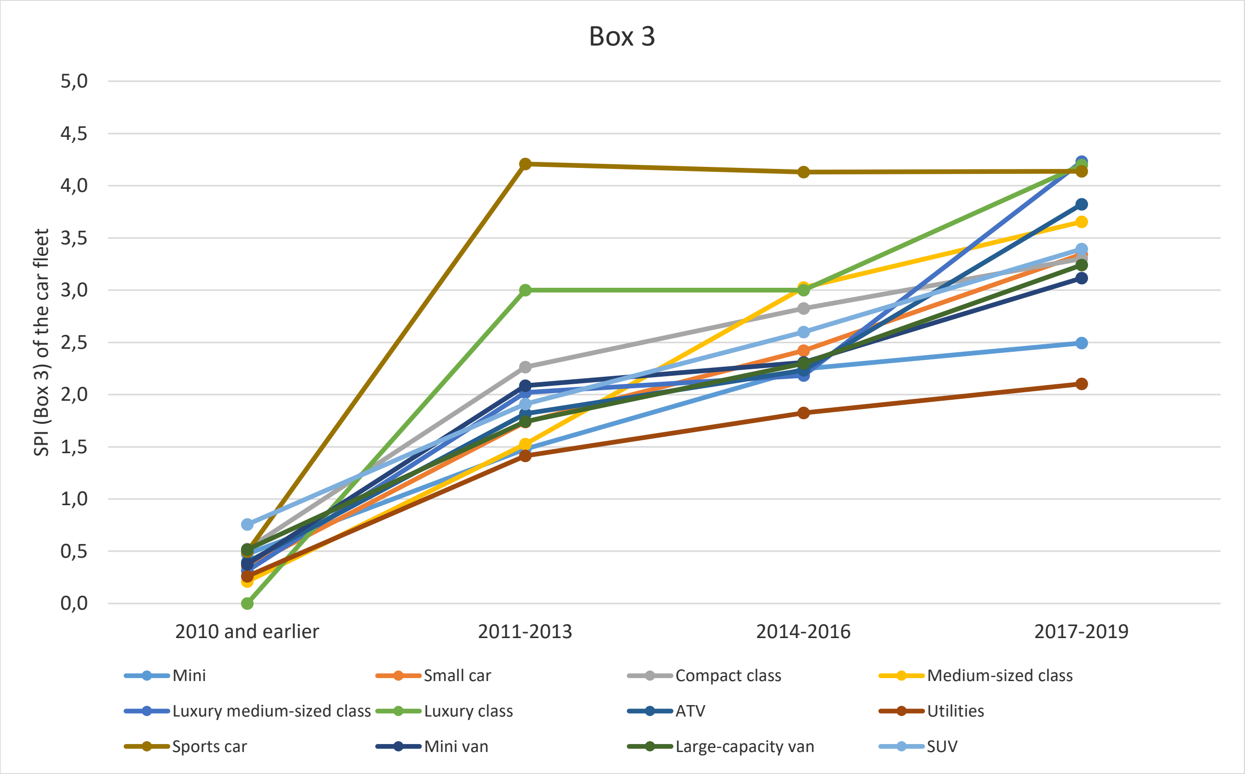


Figure 8: SPI (Box 3) of the car fleet 2019 grouped into car segment and year of initial registration (source: own estimations – data basis: Euro NCAP 2009 - 2019; Central Vehicle Register 1.1.2020)

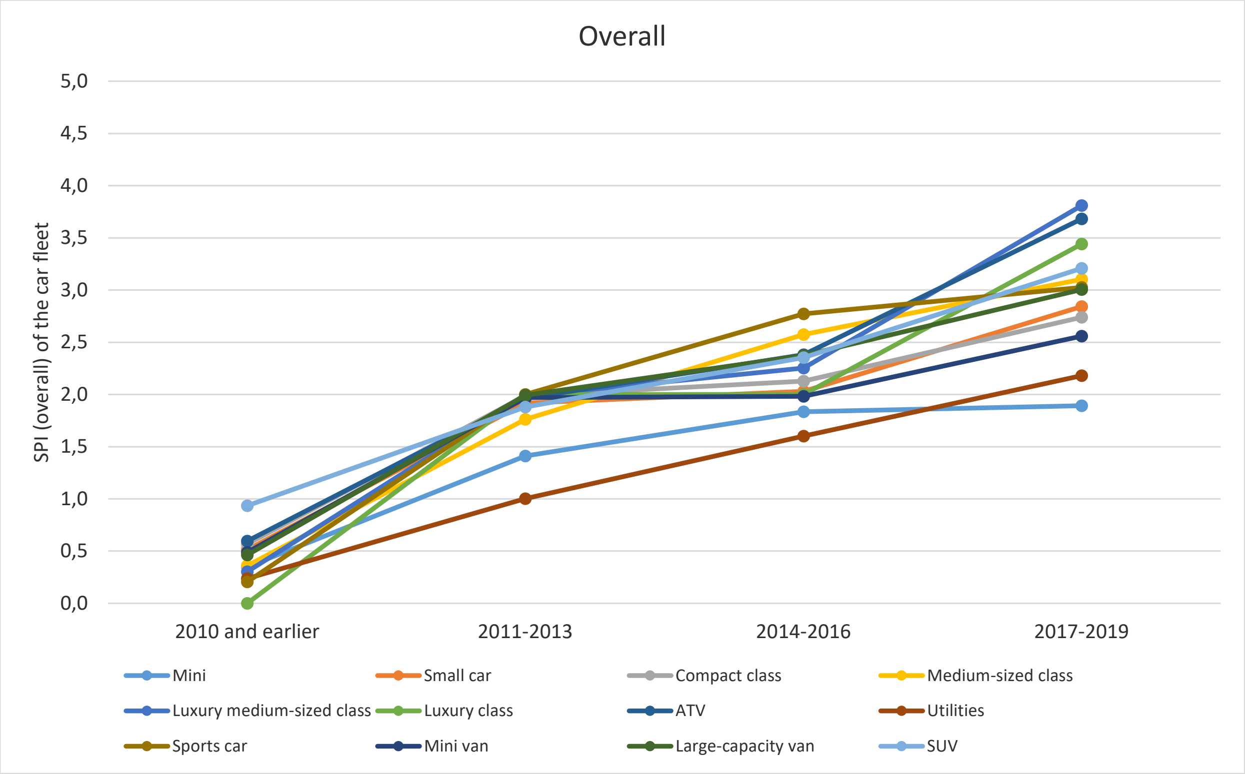


Figure 9: SPI (overall) of the car fleet 2019 grouped into car segment and year of initial registration (source: own estimations – data basis: Euro NCAP 2009 - 2019; Central Vehicle Register 1.1.2020)

If, in box 3, only the vehicles are observed which belong to the same age group (e.g. year of initial registration 2017 to 2019) the SUVs are still ranked behind the medium-sized class segment at place 6. The highest SPI value 2019 for this box can be found in this vehicle group for models of the luxury medium-sized class and luxury class.

The result that SUVs are more in the middle of the range when the vehicle age is kept constant also applies in principle to the other boxes (see Figure 9).

In this context, it should be noted that only self-protection was rated in Euro NCAP until the year 2020. In this sense, for example, SUVs and luxury class vehicles have a higher mass which is beneficial to self-protection in a vehicle-vehicle-crash. On the other hand however, the consequences for the other party are mostly more serious. As partner protection has also been taken into account since the year 2020, an apparent further inventory increase of such cars would therefore not lead to an increase in the SPI for the overall vehicle fleet in the same way in the future as it was the case in the past. The reason for this is the fact that light vehicles mostly have a better partner protection than heavy vehicles.

4.4 Guidance on the interpretation of results

To conclude this chapter, this section will once again summarise the framework conditions which should be kept in mind when interpreting the extrapolation results and the indicator values.

- In order to be able to compare the Euro NCAP tests over the years in the first place, certain estimations on the ranking of past Euro NCAP ratings in comparison to the current ratings from the year 2020/2021 had to be carried out. These were made together with experts at the Federal Highway Research Institute (BASt) in the fields of active and passive vehicle safety and are documented in text and in the appendix accordingly.
- This procedure was selected to gain a suitable starting point for the estimation of an SPI using a scientific-methodical approach in combination with the best possible practical estimation at the current time. An important aspect here was to represent the individual steps in the development of the SPI as transparently as possible so that no ‘black box’ indicator was created here, but instead to have the possibility to enter into professional discussions with other experts as, of course, in terms of the modification of the point rating, other assumptions are also possible.
- Incidentally, one should always be aware that the Euro NCAP rating is ultimately there for the purpose i.e. was developed in order to set incentives for the development of new technologies. However, these cannot always be represented directly in safety effects.
- Additionally, it must be borne in mind that in the Euro NCAP rating, only a small section of all possible safety measures on the part of the vehicle is assessed. For example, safety-relevant aspects such as tyres, vehicle lighting or technical monitoring are not included in the rating.
- When using the SPI, it must be said that the benefit of the performance indicator is only revealed over the years, i.e. after a sufficiently long time series has been built up. On the basis of the SPIs developed here, innovations in terms of a safe construction and equipment of vehicles can then be assessed using corresponding changes in the time series.
- Finally, it should also be noted that the SPIs determined in this research project are, of course, significantly affected by the underlying data basis in terms of their quality and informative value. For example, possible limitations within the vehicle performance rating according to Euro NCAP also flow into the estimation of the SPI. In Box 3, when drawing direct conclusions from the SPI on the actual crash occurrences, it must be considered that the ratings up to and including the year 2015 concentrate solely on pedestrian protection through minimising the consequence of crashes on the basis of component tests. Only from the year 2016 were measures of crash avoidance also assessed; these were only extended to cyclists from the year 2018. This should be taken into account when interpreting time series results for this box, as it covers all traffic users.

If these points are observed, the results presented in this chapter offer a starting point which is secured in terms of concept and data when answering the question examined in the following chapter as to whether there is a relationship between the SPI values of the car fleet and the official accident data of the respective year.

5 Safety Performance Indicators and accident occurrence

The indicators of the safety performance of the car fleet in the years 2014 to 2020 presented in chapter 4 are based on the amalgamation of Central Vehicle Register data on the entirety of the cars registered in the years in question with modified[23] Euro NCAP data for the vehicle models tested from 2009. In this chapter, it is examined if or how the changes of the safety performance of the car fleet determined for the time period 2014 to 2020 are reflected in the overall number and structure of crashes involving cars – recorded by the official road traffic accident statistics.

In order to introduce the topic, it will first be shown which consequences in terms of the overall effect of a car fleet's safety performance can already be derived from existing study data by way of a secondary analysis.

5.1 Secondary analysis of existing data on car-car-collisions

The suitability of NCAP-based SPIs as indicators for accident occurrence in actual road traffic can, to a certain extent, also be investigated by using existing study data. The *relative* risk of injury which is mostly determined in the studies is, however, not sufficient to answer the research question which is of interest here, as was already shown in chapter 2.3.

The so-called attributable risk is of importance among the various epidemiological parameters in view of this question – after all, it is concerned with the quantification of the significance of a risk factor for public health. The attributable risk was introduced to epidemiological methodology by LEVIN (1953).

For the attributable risk, the (absolute) injury risk R_1 of the *non*-exposed subgroup of the risk population (here, for example, vehicles with 5-star rating[24]) is compared with the injury risk R as a whole. The attributable risk (aetiologic fraction) is defined using

$$AR = 1 – R_1/R = (R – R_1)/R.$$

More information is available, for example, in WOODWARD 2005, p. 146-152 (see also HAUTZINGER et al. 2007).

The attributable risk identifies the importance of a risk factor (here star rating) from a public health perspective. For example, a value of AR = 0.22 means that 22% of the risk of injury can be attributed to the factor 'driving a 2 to 4 star vehicle (instead of a five-star vehicle)'. By completely eliminating the observed risk factor (so in this case by replacing all 2, 3 and 4-star vehicles in the car fleet with 5-star vehicles), the overall number of injured road users would be reduced by 22%.

The estimation and interpretation of the attributable risk is illustrated below on the basis of the Euro NCAP and accident data from the essay presented in chapter 2.1 by KULLGREN et al. 2019 (Tab. 13). The respective data record was newly prepared for purposes of the secondary analysis, whereby here only the data on car-car-collisions are observed in which at least one driver was seriously or fatally injured.

---

[23] Consideration of the test conditions varying over the course of time.

[24] Hence, the exposed subgroup would consist of the vehicles with four stars and fewer.

	Euro NCAP stars	n	**x1**	**x2**	x3
Fatal+ serious	2	10,450	**393**	**691**	441
	3	13,437	**357**	**779**	562
	4	43,160	**991**	**1,858**	1,921
	5	35,419	**640**	**1,074**	1,603
	Total	102,466	**2,381**	**4,402**	4,527
Fatal injuries	2	10,450	**14**	**79**	52
	3	13,437	**5**	**71**	57
	4	43,160	**24**	**182**	203
	5	35,419	**7**	**75**	181
	Total	102,466	**50**	**407**	493

Tab. 13: Extract from the accident data table by KULLGREN et al. 2019 (source: KULLGREN et al. 2019, p. 8)

Key:

- n: Number of car-car-collisions in total
- x1: Number of accidents in which both drivers were injured
- x2: Number of accidents in which the driver of the Euro NCAP tested car (case car) was injured, the driver of the opposing car (other car) however was not
- x3: Number of accidents in which the driver of the ‘case car’ remained uninjured and the driver of the ‘other car’ was injured

As the exposition status (number of stars in Euro NCAP Box 1) for ‘other car’ is not known, analysis procedures which were developed for matched case-control studies (KURITZ and LANDIS 1988) are not used in this case. For simplification purposes therefore, only the data of the ‘case cars’ are drawn upon for the estimation of the attributable risk AR.

If only the ’fatal injuries’ are observed, for the purpose of the estimation of the attributable risk AR, the data of Tab. 13 are to be prepared as follows in a fourfold table (n=102,466 cars with Euro NCAP rating (Box 1), which collided with another car):

Number of NCAP stars	Killed drivers	Remaining drivers	Total
2 to 4 stars	375	66,672	67,047
5 stars	82	35,337	35,419
Total	457	102,009	102,466

The (estimated) risk of a fatal injury for the drivers of all vehicles thus amounts to

$$R = 457/102{,}466 = 0.00446$$

and for the drivers of 5-star vehicles

$$R_1 = 82/35{,}419 = 0.00232.$$

The (estimated) attributable risk therefore amounts to

$$AR = 1 - R_1/R = 0.48091.$$

If all 2 to 4-star vehicles (exposed units) were replaced by 5-star vehicles (non-exposed units), the number of fatally injured drivers would decline by approx. 48%.

If the attributable risk is estimated for ‘serious injuries’ in the same way, one would get

$$AR = 0.25364.$$

This means that a car fleet which consisted of five-star vehicles throughout would lead to around 25% less seriously injured car drivers. However, it should be noted once more that these results only refer to the risk of injuries to car drivers in car-car-collisions.

5.2 SPI and number of casualties in road traffic accidents

5.2.1 Basic concept

In the following, it is investigated whether there is a relationship between the performance indicators (SPI [Box k]) developed here for the car fleet and the respective number of injured persons in accidents from the official road transport accident statistics. In order to answer this research question, various statistical procedures come into question.

In the case of a simple correlation analysis, each of the values of two variables, namely the SPI for the car fleet (x) and a quantity (y) tailored to it for the crash occurrences, e.g. the number of injured car occupants, would be observed for subsequent years. The question would then have to be clarified whether the number of accidents and injured y declines with rising fleet safety performance x. With this kind of highly aggregated data, experience has shown that it is difficult to identify valid relationships. Alone through the estimation of the correlation between the SPI and the annual number of accidents or injured (as a consequence of an accident), the contribution of vehicle technology safety cannot be quantified precisely enough. This is due to the fact that the number of injured in accidents is not only dependent on vehicle safety, but also on additional factors such as road infrastructure, road user behaviour or weather conditions in the respective year.

An improvement of the outlined methodological approach comprises the inclusion of additional explanatory variables together with the use of appropriate statistical models (regression). In this context, the annual overall vehicle mileage as ‘accompanying’ variable is of primary importance, as ceteris paribus (i.e. at constant level of passive and active vehicle safety) the number of accidents and casualties varies with the overall number of kilometres travelled. The extent to which improved safety-related equipment of the vehicles (expressed through the SPI of the car fleet) is reflected in the accident figures can be more reliably quantified with the help of such an extended analysis approach. In addition, for the analysis of cyclists who are injured or killed in collisions with cars, the inclusion of the annual kilometre performance of bicycles would be useful. However, there are no time series data here. Therefore, in this project, the bicycle stock (incl. pedelecs[25]) is drawn upon as substitute control variable. The time series of annual vehicle mileage and the bicycle stock are shown in Tab. 14.

Year	Kilometres travelled by cars registered in Germany (in bn)	Stock of bicycles and pedelecs (in mio)
2013	611.0367	71.0
2014	611.9404	72.0
2015	619.0991	72.0
2016	625.9151	73.0
2017	630.8367	73.5
2018	630.8429	75.5
2019	632.2536	75.9

Tab. 14: Annual number of kilometres travelled by cars registered in Germany 2013 - 2019 (source: Federal Motor Transport Authority 2020) and bicycle stock (incl. pedelecs) 2013 - 2019 (source: ZIV 2020)

As regards annual car mileage, it is to be noted once again here (see BÄUMER et al. 2017a and b) that this is the total mileage of cars registered in Germany including distances travelled abroad (national concept), while the accident data presented in the following refer, in fact, to the casualties on German roads (domestic concept). Thus, a slight blurredness in the analysis remains here.

5.2.2 Tailoring accident data to be used to SPIs to be studied

As a separate fleet SPI was formed for every Euro NCAP box, the accident data – or more precisely the dependent variable to be explained in the analysis – had to be tailored to the respective box. The development of the box-specific SPIs over time (displayed in Tab. 12) is therefore contrasted with the development of the corresponding number of injured or killed adults, children, pedestrians etc.

In all analyses, the car fleet SPI referring to 1 January of a year (e.g. 1.1.2014) is matched with the number of casualties and the vehicle mileage of the previous year (using 2013 as an example).

Accident data for ‘SPI (Box 1)’

As Box 1 refers to Adult Occupant Protection, the number of injured or killed car users aged 12 years and over is the appropriate criterion variable of interest; the respective values are summarised in Tab. 15.

---

[25] Pedelecs are included here as they are also contained in the published accident data on car-cyclist crashes (see DESTATIS, Subject-Matter Series 8, Series 7, Table 3.1.2).

Accident year	Fatalities	Seriously injured	Slightly injured	Total
2013	1,565	27,287	175,362	204,214
2014	1,556	28,042	178,925	208,523
2015	1,592	28,647	182,415	212,654
2016	1,503	28,809	184,532	214,844
2017	1,420	28,444	181,210	211,074
2018	1,397	27,833	174,915	204,145
2019	1,343	27,271	170,427	199,041

Tab. 15: Injured car occupants (drivers and passengers) from the age of 12 years broken down by injury severity and accident year (Source: DESTATIS: Traffic accidents, Subject-Matter Series 8, Series 7, various years)

Thus, in the analysis model, the dependent variable consists of the annual number of injured or killed car occupants from the age of 12 years. A total of four models are thus estimated (all casualties, fatalities, seriously injured, slightly injured). In each statistical model, the fleet SPI associated with Euro NCAP Box 1 and the annual car mileage serve as explanatory variables.

Accident data for ‘SPI (Box 2)’

Correspondingly, for model estimations referring to Box 2 (Child Occupant Protection), the annual number of injured car occupants under the age of 12 years is considered as dependent variable (Tab. 16).

The age limit of 12 years was selected because up until the age of 11 years or a height of 150 cm in Germany, a child protection system must be used in the car. Such systems are also a part of the Euro NCAP rating in Box 2.

Accident year	Fatalities	Seriously injured	Slightly injured	Total
2013	22	940	6,987	7,949
2014	19	952	7,353	8,324
2015	28	916	7,490	8,434
2016	27	1,007	7,743	8,777
2017	14	1,043	7,647	8,704
2018	23	1,009	7,531	8,563
2019	20	1,016	7,300	8,336

Tab. 16: Injured car occupants under the age of 12 years broken down by injury severity and accident year (Source: DESTATIS: Traffic accidents, Subject-Matter Series 8, Series 7, various years)

The models to be estimated are identical in their structure with those described above for Box 1.

Accident data for ‘SPI (Box 3)’

The accident data for Box 3 include pedestrians and cyclists (incl. pedelecs) who were injured or killed in collisions with cars (Tab. 17). These figures are provided by DESTATIS only for accidents with precisely two involved parties (i.e. one involved car and one involved cyclist or pedestrian).

Accident year	Fatalities		Seriously injured		Slightly injured		Total	
	Pedestrian	Cyclist	Pedestrian	Cyclist	Pedestrian	Cyclist	Pedestrian	Cyclist
2013	322	138	5,518	6,118	16,044	36,106	21,884	42,362
2014	284	133	5,513	6,847	15,504	39,495	21,301	46,475
2015	317	156	5,599	6,777	16,150	39,392	22,066	46,325
2016	268	166	5,461	6,767	16,442	40,541	22,171	47,474
2017	271	137	5,341	6,444	16,042	39,590	21,654	46,171
2018	246	167	5,147	6,887	16,077	43,051	21,470	50,105
2019	245	172	4,830	6,449	15,719	41,628	20,794	48,249

Tab. 17: Pedestrians and cyclists (incl. pedelecs) injured in collisions with a car broken down by injury severity and accident year (Source: DESTATIS: Traffic accidents, Subject-Matter Series 8, Series 7, various years)

Again, four different categories of injured pedestrians have been considered in the statistical analysis.

In the models for injured cyclists, instead of annual car mileage the bicycle stock (including pedelecs) is drawn upon as control variable. Therefore, the car fleet SPI referring to Box 3 and the bicycle stock serve as explanatory variable here. In general, the Euro NCAP tests in Box 3 primarily refer to pedestrian protection. Tests for the protection of cyclists have only been in the programme since 2018. In this respect, analyses on the relationship between SPI (Box 3) and cyclists injured in collisions with cars have more of an exploratory character.

Accident data for 'SPI (Box 4)'

The SPI associated with Euro NCAP Box 4 refers to safety assist systems aiming at crash avoidance. It therefore seems quite natural here, to use the number of car drivers involved in crashes with personal injury as criterion variable when evaluating the safety performance indicator. At the same time, the number of main responsible drivers can be considered separately (Tab. 18). The analysis is focused on accidents involving personal injury, because this in particular is intended to be avoided with the use of the safety assist systems. In addition, the problems and inaccuracies resulting from unrecorded cases is less pronounced here compared to damage-only accidents.

Accident year	Involved car drivers	
	Total	of which: main responsible
2013	359,808	201,194
2014	371,095	206,637
2015	378,156	209,950
2016	381,354	211,460
2017	372,144	206,413
2018	369,050	206,041
2019	357,327	199,369

Tab. 18: Car drivers involved in accidents with personal injuries broken down by year of accident (Source: DESTATIS 2020)

Both for all car drivers involved in personal injury accidents and main responsible car drivers, a regression model has been estimated constructed in exactly the same way as the models for injured car users and pedestrians.

5.2.3 Descriptive analysis of the relationships among the variables of interest

In this section, it is examined by way of a descriptive data analysis whether there exists a relationship between the SPI of the car fleet and the crash occurrences on a bivariate level. Scatter diagrams are used for this purpose in which the overall number of the injured car users (adults, children), pedestrians and the overall number of accident-involved cars are contrasted with the corresponding SPI (Figure 10).

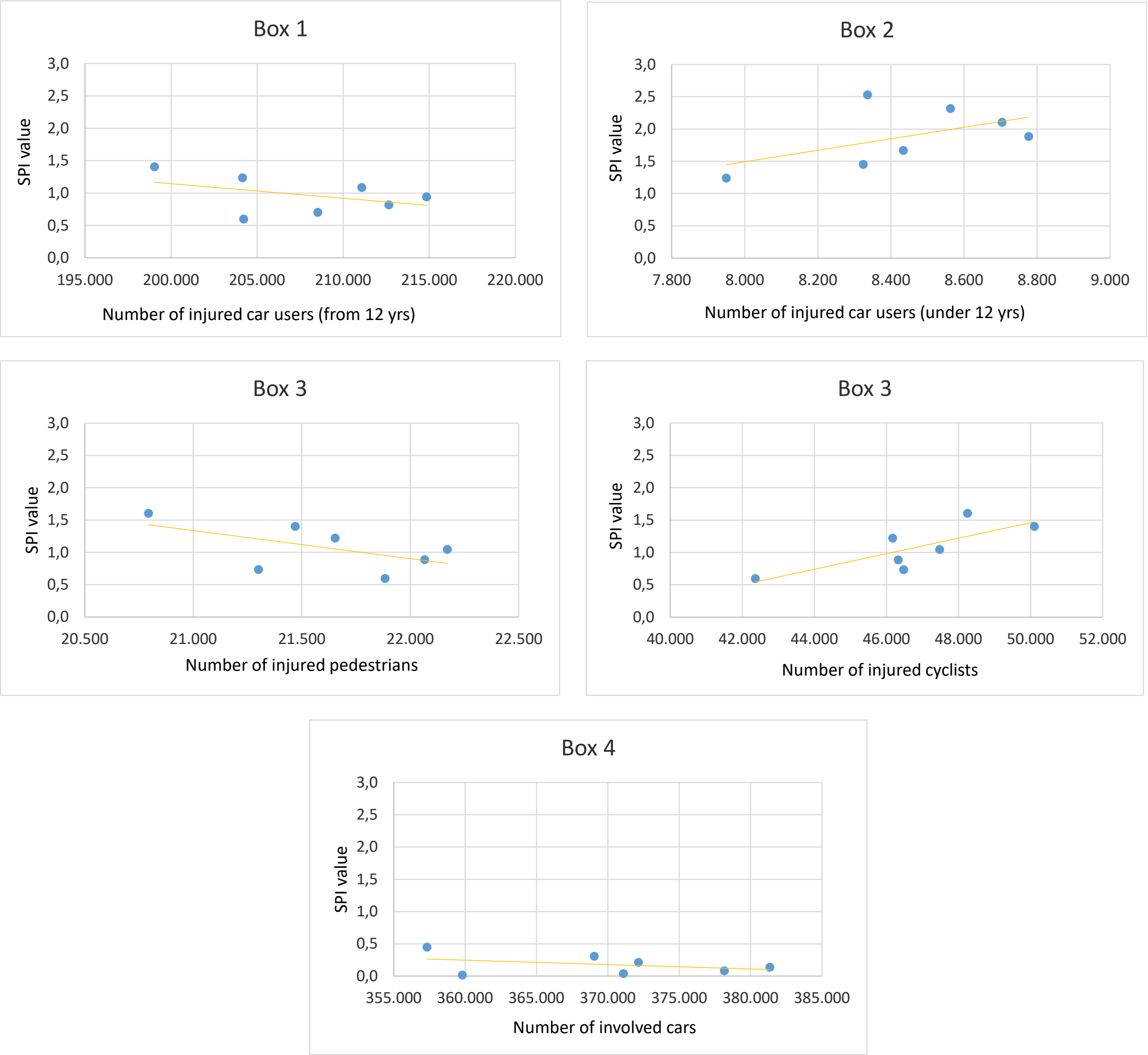


Figure 10: Scatter diagrams for the overall number of injured or killed car users (adults, children), pedestrians, cyclists, accident-involved cars and the corresponding SPI of the car fleet

In the above-mentioned bivariate data analyses, negative correlations are to be expected – the higher the SPI value of the fleet, the lower the number of casualties. This is also the case in three of the five diagrams represented in Figure 7, the correlation coefficient here lies between -0.38 and -0.57.

Contrary to expectations, the result is a positive correlation between injured children and the SPI Box 2 as well as injured cyclists and the SPI Box 3. The correlation coefficient here amounts to +0.53 and +0.78, respectively. This result shows that restricting analysis to pairs of variables (here SPI vs. injured children or cyclists) does not always do justice to the crash occurrences with their many concurrent influencing factors.

A fundamental determining factor for the number of children injured or killed in cars is the car mileage.

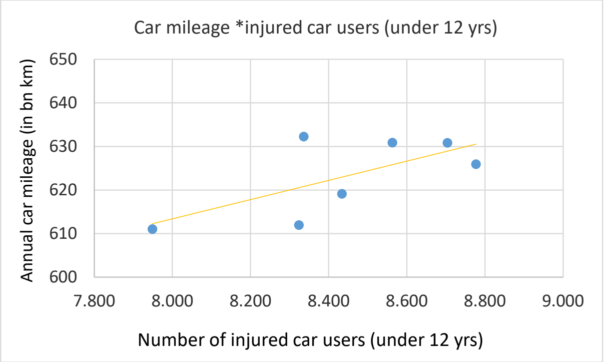


Figure 11: Scatter diagram for the overall number of injured or killed car users under 12 years and annual car mileage (in bn km)

As can be expected, the number of children injured in cars increases with the overall number of car kilometres (see Figure 11). The correlation between the number of injured children and car mileage amounts to +0.67.

These findings underline the importance of applying multivariate analysis methods for questions in which the effect of an influencing variable is to be ‘adjusted’ for the effect of each of the other influencing factors. As will be seen below, in a statistical model in which the SPI (Box 2) and car mileage are simultaneously included as explanatory variables for the number of casualties, the result is in fact a *negative* relationship between the SPI and the number of children injured or killed in cars. The partial correlation between these two variables – i.e. the correlation after controlling the effect of car mileage – amounts to -0.56.

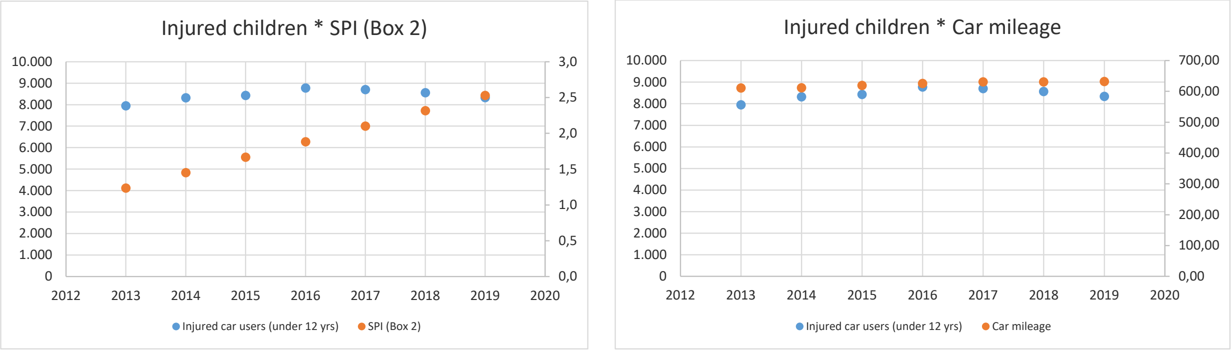


Figure 12: Injured or killed car users under 12 years, SPI (Box 2) of the car fleet and car mileage (in bn km) by year of accident

The above-mentioned multivariate analysis methods comprise models from the field of time series econometrics. This is due to the fact that the data used here are all time series data. Both the number of casualties (y) as well as the fleet SPI values (x) and the car mileage figures (z) refer in temporal terms to calendar years (time period 2013 to 2019) and in spatial terms to the Federal Republic of Germany as a whole.

The topic of ‘trend’ should be given attention in regression models for time series data. As Figure 12 shows, all three represented variables are subject to a more or less pronounced trend (SPI: strongly positive linear trend; car mileage and number of casualties[26]: slightly degressive growing trend). According to this finding and based on the regression analyses to be run, it is recommended that a trend adjustment is carried out. For this purpose, the trend in the three time series could be removed prior to the analysis. However, the integration of a time trend variable as an additional factor also leads to the desired trend adjustment. This is the route which is taken here.

5.2.4 Log-linear regression model for the number of casualties

Model approach

If the number y of persons injured or killed in accidents involving cars is considered as a function of overall car mileage z, a regression model in which y cannot take on negative values stands to reason. In addition, the number of casualties should equal zero when car mileage tends to zero.

If the SPI of the car fleet is denoted by x and annual car mileage by z, then the multiplicative approach (Cobb-Douglas function)

$$(1) \qquad y = e^{\theta} \cdot x^{\alpha} \cdot z^{\beta} \cdot e^{\varepsilon}$$

has the desired characteristics (ε denotes the normally distributed error term of the model with E(ε)=0). By taking the logarithm of both sides, the approach can be linearized:

$$(1^*) \qquad \ln(y_i) = \theta + \alpha \cdot \ln(x_i) + \beta \cdot \ln(z_i) + \varepsilon_i$$

As the model (1*) is linear in the parameters[27], it can be estimated by least squares methods using statistical software for multiple linear regression models (including estimation of the standard error as well as the coefficient of determination). Logarithmic linear regression approaches are often

---

[26] In principle, one could also talk of a quadratic trend for the number of injured, but trends should be kept as simple as possible.

[27] Only the variables y, x and z are logarithmically transformed.

used in empirical economic research, for example in demand and production models.

The parameters[28] of the model (1) are to be interpreted as constant elasticities whose value is not dependent on the level of the corresponding regressor. Example: If $\beta=1.5$, then ceteris paribus, i.e. with given SPI of the car fleet, a 1 percent increase of car mileage leads to an 1.5 percent growth in the number of casualties. Analogously, $\alpha=-0.5$ means that ceteris paribus, i.e. with given car mileage, an increase of the SPI by 1 percent reduces the corresponding number of casualties by 0.5 percent.

Risk theoretical aspects

Several general risk theoretical considerations can be made on the parameters α and β of the model (1) for the number of casualties.

In our context, the main interest is the effect of the SPI of the car fleet – as indicator of the protection of occupants (Boxes 1 and 2) or the protection of pedestrians and cyclists (Box 3) – on the overall number of persons injured or killed in road traffic accidents. According to the results of existing studies summarised in chapter 2 and the additional analyses in chapter 5.1, it can be expected that with given car mileage z, the number of casualties y drops when the safety performance level of the car fleet – measured by the explanatory variable x – increases. Such a negative or more precisely monotonously decreasing relationship between x and y is represented in the model by $\alpha<0$.

If the annual overall mileage z of the car fleet is considered as risk exposition parameter (on every driven kilometre there is the possibility of a crash and thus also the possibility of personal injury), then it can be expected that the annual number y of persons injured or killed in accidents involving cars ceteris paribus increases with car mileage. A positive or more precisely monotonously increasing relationship between z and y is reflected in the model by $\beta>0$. When looking at the form of the relationship in more detail, three cases can be distinguished.

In the case $\beta=1$, the expected number of casualties is directly *proportional* to overall car mileage corresponding to a constant (mileage-related) rate y/z of casualties. For $0<\beta<1$, y grows *degressively* with z, i.e. additional vehicle kilometres Δz in the road network do increase the absolute number of casualties but reduce the per-kilometre casualty rate y/z. Conversely, $\beta>1$ corresponds to a *progressive* growth of the number of casualties in dependence on car mileage, i.e. additional vehicle kilometres driven on the road network increase both the number y of casualties and the casualty rate y/z.

As the overall German road network and the entire annual car traffic are subjects of this study, it is difficult to decide upon one of the three possible cases for β a priori[29].

Trend adjustment

As was shown in section 5.2.3, the time series of the model variables y, x and z are more or less strongly affected by trends. This is indeed problematic, as an observed relationship between x and y (analogously between z and y) may simply be due to the fact that both time series show a trend. In order to avoid purely artificial relationships (spurious regression) and biased regression parameter estimates (omitted variable bias), trend has to be taken into account in the regression model.

For this purpose, the logarithmic linear model approach was extended by inclusion of a time trend variable t (parameterisation t = 1, 2, …, T), which in addition to x and z determines the number y of casualties:

$$y = e^{\theta} \cdot t^{\lambda} \cdot x^{\alpha} \cdot z^{\beta} \cdot e^{\varepsilon} \tag{2}$$

$$\ln(y) = \theta + \lambda \cdot \ln(t) + \alpha \cdot \ln(x) + \beta \cdot \ln(z) + \varepsilon \tag{2*}$$

As already mentioned, the addition of the ‘autonomous’ trend variable t has a beneficial effect on the quality of the estimation of the parameters associated with the explanatory variables x and y. In accordance with our risk-theoretically substantiated model concept, $\alpha<0$ and $\beta>0$ is expected.

In regression models with time series data[30], a trend adjustment is essential for many reasons. Thus, macroeconomic models are often formulated with an autonomous trend variable which, for example, represents the effect of ‘technical progress’ on productivity – in addition to the effect of growing capital expenditure (see GREENE 2000, p. 325-326; SCHIRA 2009, p.554).

In our context, the trend variable allows an estimation of the ‘autonomous change of safety in road traffic’. The parameter λ indicates an increase ($\lambda>0$) or decrease ($\lambda<0$) of the number of casualties over the course of time which cannot be at-

---

[28] The parameters are also indicated as regression coefficients.

[29] If the overall road network was segmented into smaller subnetworks and the complete year was decomposed into smaller time periods, then empirically founded assumptions on the range of the model parameter β would rather be possible. For the aggregates considered here, one should rely on the results of parameter estimation.

[30] See, for instance, ENDERS (2010) for time series econometrics.

tributed to the variation of overall annual car mileage z or the increase of car fleet safety performance measured by the indicator x.

Numerical example:

Trend variable $t = 1, 2, \ldots, 7$

Parameter $\lambda = 0.13$ (degressive growth)

Ceteris paribus the expected number $\hat{y}$ of casualties depends on t as follows (trend function):

$$\hat{y} = K \cdot t^{\lambda} = K \cdot t^{0,13} \qquad (K = \text{const.})$$

Interpretation of $\lambda = 0.13$ as 'elasticity' of the number of casualties: if the trend variable t grows by 1%, the expected number of casualties rises by 0.13%. As t can only take on integer values, the idea of a small percentage change of the trend variable is not intuitive.

More clearly interpretable is the *relative* change of the expected number of casualties due to an increase in the trend variable t by 1 unit. According to

$$\hat{y}(t+1) / \hat{y}(t) = (t+1)^{\lambda} / t^{\lambda} = (1 + 1/t)^{\lambda}$$

the following applies: if t rises by 1 unit (i.e., from one calendar year to the next) the expected number of casualties ceteris paribus rises by

$$100 \times [(1 + 1/t)^{\lambda} - 1] \text{ percent.}$$

For $\lambda = 0.13$ the following (non-constant) growth factors describe the development of the number of casualties over time:

t	$(1+1/t)^{\lambda}$
1	1.094
2	1.054
3	1.038
4	1.029
5	1.024
6	1.020

The table entries are to be interpreted as follows: from year 1 to year 2, the expected number of casualties rises ceteris paribus by 9.4%. From year 6 to year 7, the expected number of casualties ceteris paribus increases by 2.0%.

In the following, the results of the statistical analyses are displayed on the basis of regression model (2*). Clearly, the focus lies on determining the effect of the SPI of the car fleet (Boxes 1 to 4) on the number of injured or killed persons in road traffic (overall and differentiated according to injury severity) or the number of accident-involved cars, respectively.

5.2.5 Analysis results on the relationship between car fleet SPI and number of casualties in road traffic accidents

SPI_1: Adult Occupant Protection

The total number of regression models to be considered in section 5.2.5 is quite large. Therefore, as far as the evaluation of the car fleet SPI (Box 1) is concerned, only the model results for the overall number of injured or killed adult car occupants will be presented in detail, i.e., with the parameter estimates and significances for all explanatory variables (Tab. 19). For the three additional models related to SPI (Box 1) (dependent variable: number of fatalities/ seriously injured/ slightly injured), only the primarily interesting parameter estimates for (logarithmic) SPI_1 will be shown. This also applies to the model results referring to the remaining NCAP Boxes 2 to 4.

Effect	Estimated value	Significance
Constant	-6.83	0.0708
Log SPI_1	**-0.44**	**0.0004**
Log car mileage	2.94	0.0047
Log year	0.13	0.0009

Tab. 19: Estimates and significances of the parameters of the log-linear model for the overall number of injured or killed car occupants from 12 years

In the model for the overall number of injured or killed adult car occupants, 99% of the variance of the dependent variable is explained by the three influencing factors ($R^2=0.9914$). The model constant (estimate: -6.83) can be easily interpreted in the context of the multiplicative model (2): When all three influencing factors x, z and t take on the value of 1, the expected number of casualties is practically zero due to $e^{-6.83} = 0.001$. This is in perfect agreement with our expectations, as z=1 is the hypothetical situation where overall car mileage tends to zero (i.e., no car traffic at all on the road network).

The crucial estimate is that for the parameter of the variable log SPI_1 (-0.44, printed in bold). Due to the interpretation as elasticity, this result indicates a reduction of the number of injured or killed car occupants (from 12 years) of around 0.4% when ceteris paribus the car fleet SPI (Box 1) increases by 1%. As the null hypothesis H_0: $\alpha \geq 0$ can be rejected in favour of H_1: $\alpha < 0$ at every common significance level ($p = 0.0004/2 = 0.0002$), one can be practically certain that an increase in SPI_1 leads to a decrease in the number of injured or killed adult car occupants. Thus, SPI_1 is, without any doubt, a suitable safety performance indicator.

As regards overall car mileage and the time trend variable, the positive parameter estimates (2.94 and 0.13, respectively) show that with increasing variable values, the number of injured or killed adult car occupants increases as was to be expected in view of the results of the descriptive data analysis. With the exception of the constant, all parameter estimates are significant at the 5 percent level[31].

Tab. 20 summarises the parameter estimation results for the safety performance indicator SPI_1 with regard to the different categories of casualties (all injured (see above), fatalities, seriously injured and slightly injured).

Model for...	Parameter estimate for log SPI_1	Significance
All injured	-0.44	0.0004
Fatalities	-0.46	0.0493
Seriously injured	-0.35	0.0003
Slightly injured	-0.45	0.0004

Tab. 20: Parameter estimation results for the factor 'SPI (Box 1) of the car fleet' in the log-linear models for the number of injured car occupants from 12 years

All estimates have a negative sign[32] and are significant at a level of 5%. This means that irrespective of injury severity, the number of casualties is the smaller, the larger the safety performance indicator for adult occupant protection turns out.

SPI_2: Child Occupant Protection

The parameter estimates for the explanatory variable SPI_2 in the various regression models for the number of children injured or killed in cars are summarised in Tab. 21.

Model for...	Parameter estimate for log SPI_2	Significance
All injured	-0.61	0.0153
Fatalities	-0.59	0.8801
Seriously injured	-0.03	0.9531
Slightly injured	-0.69	0.0075

Tab. 21: Parameter estimation results for the factor 'SPI (Box 2) of the car fleet' in the log-linear models for the number of injured car occupants under 12 years

As with NCAP Box 1, all parameter estimates are negative. A significant effect of the safety performance indicator SPI (Box 2), however, can only be found in the models for the injury categories with large frequencies ('all injured' and 'slightly injured', respectively). With regard to Child Occupant Protection, the SPI elasticity of the number of injured lies in the range of -0.6 to -0.7. Thus, it appears that the SPI effect on child safety is somewhat more pronounced than in the case of adult car occupants where the respective elasticities lie at around -0.4 (see Tab. 20).

SPI_3: Vulnerable Road Users

Separate models were estimated for pedestrians and cyclists injured or killed in collisions with cars. However, in any case the SPI (Box 3) is included as explanatory variable because in Euro NCAP the tests for the third box refer to both pedestrian and cyclist protection. Tab. 22 contains the estimated coefficients of the indicator SPI_3 for injured pedestrians, Tab. 23 for injured cyclists.

Model for...	Parameter estimate for log SPI_3	Significance
All injured pedestrians	-0.30	0.0544
Fatalities	-0.51	0.3809
Seriously injured	-0.61	0.0042
Slightly injured	-0.20	0.1819

Tab. 22: Parameter estimation results for the factor 'SPI (Box 3) of the car fleet' in the log-linear models for the number of pedestrians injured in collisions with cars

---

[31] First-order autocorrelation amounts to -0.285 and the Durbin-Watson coefficient equals 2.45 (rule of thumb for the interpretation of the Durbin-Watson coefficient: values between 1.5 and 2.5 are acceptable, values below 1 or above 3 suggest autocorrelation).

[32] The elasticity of the number of casualties with regard to the safety performance indicator SPI_1 lies in the order of -0.4, largely independent of injury severity.

Model for...	Parameter estimate for log SPI_3	Significance
All injured cyclists	-0.67	0.0083
Fatalities	-0.02	0.9900
Seriously injured	-1.02	0.0043
Slightly injured	-0.61	0.0094

Tab. 23: Parameter estimation results for the factor 'SPI (Box 3) of the car fleet' in the log-linear models for the number of cyclists injured in collisions with cars

From the four models for injured pedestrians, only the model for seriously injured reveals a highly significant effect of the indicator SPI_3 (SPI elasticity of the number of seriously injured lies at -0.6). But also, for the overall number of injured pedestrians the dependency on SPI_3 is practically certain (p=0.0272 in the one-sided test).

In contrast, a highly significant SPI effect can be found in three of four models for injured cyclists (slightly injured, seriously injured and all injured). The SPI_3 parameter estimate of -1.02 in the model for seriously injured cyclists is worth mentioning: an increase of the indicator SPI_3 by 1% is accompanied with a reduction in the number of cyclists seriously injured in collisions with cars by 1% (more precisely 1.02%). This is the largest value among the empirically derived SPI elasticities. At -0.6 and -0.7, the SPI elasticity is also relatively strongly pronounced for slightly injured cyclists as well as for all injured cyclists.

Therefore, in the case of a collision, cyclists also benefit to a high degree from cars which performed well in tests on the protection of vulnerable road users (which are mainly targeted at pedestrians).

SPI_4: Safety Assist Systems – active safety

As safety assist systems are primarily expected to reduce the number of crashes, in connection with SPI_4 it is not the injured that will be considered as target variable, but vehicles involved in accidents with personal injury (all involved cars and cars involved as main responsible).

Model for...	Parameter estimate for log SPI_4	Significance
Involved cars	-0.24	0.0158
Cars as main responsible	-0.20	0.0235

Tab. 24: Parameter estimation results for the factor 'SPI (Box 4) of the car fleet' in the log-linear models for the number of cars involved in accidents with personal injury

In both cases, the number of cars involved in accidents drops significantly with the value of the indicator SPI_4 (Tab. 24). Thus, it can be regarded as practically certain, that the diffusion of safety assist systems in the car fleet positively contributes to road traffic safety. With growing SPI_4 of the car fleet, however, the number of cars involved in accidents does not decrease as strongly as the number of injured with rising SPI_1 of the car fleet (SPI elasticity -0.2 versus -0.4). (The same is true for SPI_2 and SPI_3.) The still comparably low level of SPI_4 in the car fleet might also play a role here.

Summary

To conclude, it can be determined that in all – overall 18 – log-linear models, a higher value of the Safety Performance Indicator comes along with a lower number of injured or killed road users or accident-involved cars.

Figure 13: The effect of the box-specific car fleet SPI values on the number of casualties and the number of accident-involved cars, respectively

The quantities SPI_1 to SPI_4 for the car fleet developed here on the basis of Euro NCAP are therefore suitable indicators of traffic safety. However, the positive safety contribution of the respective SPI is mostly significant, but not in all cases. This is, of course, also due to the fact that model estimation is based on T=7 time periods (years) only. In this respect, the results should generally be interpreted with care; in particular it should be taken into account that the number of casualties and accident-involved cars will underlie random fluctuations from year to year.

Finally, it should be mentioned that SPI-related estimation results are, strictly speaking, only valid for the empirically observed range of indicator values (e.g., 1.24 to 2.53 for SPI_1). Whether or not the derived relationships also apply to the theoretically possible range of SPI values (0 to 5) must remain unanswered at this point. In future evaluations of the SPI for the car fleet, the empirically observed range of values will, of course, become increasingly wider.

5.2.6 Relationship between overall SPI and accident costs

To conclude the accident analyses, the relationship between the overall SPI (see Tab. 12) and the accident characteristics of interest needs to be studied now. As these accident characteristics are, on the one hand, injured road users (car occupants as well as cyclists and pedestrians injured or killed in collisions with cars) and, on the other hand, accident-involved cars, the respective numbers (event frequencies) cannot be simply added. Rather, the different target variables can only be aggregated after an appropriate monetarisation, i.e., after conversion into accident costs. For this purpose, the yearly accident cost rates published by the Federal Highway Research Institute (dated: May 2020) for personal injury (fatalities, seriously injured and slightly injured), and material damage (for accidents with personal damage) are combined.

However, the latter cost rate refers to material damage for *all accidents* with personal damage[33], while in this context only the material damage costs for injury accidents *involving a car* are relevant. Therefore, an approximation needs to be made here: in a first step, for each year the proportion of injury accidents involving a car among all accidents with personal damage (around 80% depending on the year) was transferred to the cost rate of material damage. In a second step, a material damage cost rate per car was derived using the number of cars involved in such accidents (with personal damage and car involvement). This cost rate lies – depending on the year – in the range between € 8,608 and € 9,254.

Overall, the cost estimates shown in Tab. 25 refer to the total economic and social costs of road traffic injury accidents involving at least one car (personal and material damage) relevant here.

Accident year	Accident costs (in bn euro)
2013	11.607
2014	11.778
2015	12.338
2016	12.166
2017	11.620
2018	11.103
2019	10.730

Tab. 25: Estimated Euro NCAP-relevant accident costs 2013 - 2019 (sources: DESTATIS: Traffic accidents, Subject-Matter Series 8, Series 7, various years and BASt 2020b)

As before, a log-linear regression model has been used. In this model, total accident costs represent the dependent variable, explanatory variables are the overall SPI together with car mileage and the time trend variable. Tab. 26 contains the results of model estimation.

Effect	Estimated value	Significance
Constant	-17.47	0.2558
Log SPI_overall	**-0.70**	**0.0119**
Log car mileage	3.05	0.2135
Log year	0.21	0.0235

Tab. 26: Estimates and significances of the parameters of the log-linear model for the total costs of injury accidents involving at least one car

The parameter of log SPI_overall (printed in bold) shows a relatively high elasticity of the accident costs estimated at -0.7: If the overall SPI of the car fleet rises by 1%, the respective accident costs ceteris paribus drop by 0.7%. The parameter is significantly different from zero (p-value 0.0119), while the positive parameter of the variable log car mileage (i.e., ceteris paribus accident costs rise with overall car mileage) is not significant at the 5 per cent level.

---

[33] In the year 2018, for example, the material damage costs per accident with personal damage amounted to € 16.684 (BASt 2020b).

6 Summary and outlook

The European New Car Assessment Programme (Euro NCAP) is a consumer protection orientated programme for the safety assessment of – as a rule – new car models. The programme was established in 1997, since 2009 it has consisted of the following 4 Boxes:

- Adult Occupant Protection (Box 1)
- Child Occupant Protection (Box 2)
- Vulnerable Road User Protection (Box 3)
- Safety Assist Systems (Box 4)

Within each Box, several individual tests in which crash tests play an important role are carried out. Each is assessed using points. The safety assessment of a model is indicated using ‘stars’ (0 to 5). For every Box, the proportion of achieved test points is first established and then translated into a star assessment using a predetermined table containing threshold values. An overall assessment – also in the form of stars – is created from this for each vehicle model.

The main objective of this project was to transfer the test results from Euro NCAP onto the overall car fleet wherever possible and to use this to form a Safety Performance Indicator (SPI) for the vehicle stock over several consecutive years (time series). The purpose of this SPI is to provide a condensed description of the level and development of vehicle safety in the German car fleet using one or a small number of parameters.

A further objective was to investigate whether or not a relationship exists between the Euro NCAP rating of vehicles and accident occurrences. For this purpose, in addition to literature analyses, statistical models were estimated on the effect of (average) vehicle safety expressed by the SPI on the corresponding number of casualties taken from the official traffic accident statistics.

The following data sources served as basis for the development of the Safety Performance Indicators:

- Euro NCAP test results between 2009 and 2019
- Car inventory data from the Central Vehicle Register of the Federal Motor Transport Authority between 2009 and 2020 (each on the reference date 1.1.)

Data for the analysis of accidents was taken from the official road traffic accident statistics of the Federal Statistical Office (DESTATIS).

6.1 Methodology

The assessment of vehicles in Euro NCAP has been subjected to continuous change over the years. This has affected the threshold values for the allocation of the stars and, in particular, the type and number of the tests carried out. In different years, the requirements for a good assessment were raised in some tests or they were replaced by new test procedures. In addition, certain procedural elements were also completely removed from the programme from a certain date.

This leads to the Euro NCAP results for vehicle models assessed in different years not, or only limitedly, being comparable with each other. However, comparability is a necessary prerequisite in order to derive an SPI time series related to the overall car inventory from the Euro NCAP results.

A fundamental step in the development of an SPI of vehicle safety therefore consisted of making the test results from the various years comparable as far as possible with regard to the test procedures which had changed in the course of time. For this purpose, a project group was formed consisting of experts from the Federal Highway Research Institute (BASt) in the fields of active and passive vehicle safety which had the task of quantifying, for each Box, the extent to which these changes to the test procedures affect the vehicle assessment. As changes to the test procedures tend to mean their intensification, specifically reduction factors were defined with which the assessment results for earlier test years were lowered accordingly. In doing so, the test conditions of the year 2020 were applied as reference, i.e. the corrected test results thus correspond to the hypothetical case that in the years 2009 to 2019 the test procedures of the year 2020 would already have been valid.

The reduction factors determined by the group of experts for the individual test years were (per Box) applied to the proportion of test points which a model had achieved and these were thereby lowered accordingly. The final transformation of these newly estimated proportional values into stars was carried out using the threshold values of the year 2020.

The second fundamental step in generating a Safety Performance Indicator (SPI) consisted of linking (matching) the – newly estimated – Euro NCAP result data (proportion of achieved points, number of stars) for the individually tested makes and models to the Central Vehicle Register inventories (1.1.2014 to 1.1.2020). The matching of the Central Vehicle Register inventory data with the (corrected) Euro NCAP rating data was carried out using a complex algorithm which is based, in essence, on the characteristics of make code and model code. In this context, it was also necessary to identify the series (model generation) of the respective model tested in Euro NCAP in the inventory. However, this information is not available in the Central Vehicle Register data. Therefore, in addition to the characteristics of make and model, the year of initial registration also needed to be drawn upon in order to be able to establish a match with the Euro NCAP data. In addition, vehicle facelifts (if tested) were also taken into account in the matching algorithm.

As a result, on average over the seven years observed here (reference dates: 1.1.2014 to 1.1.2020), around 70% of the newer vehicles (initial registration year from 2009) were able to be allocated to a Euro NCAP assessment. Cars with an initial registration from 2009 to which no assessment could be allocated (because they were not tested or could not be found in the Central Vehicle Register) were assigned the average value of the Euro NCAP rating according to car segment and initial registration year using an imputation procedure. In agreement with the client and experts, vehicles which were first registered for use on the road before 2009 and for which no Euro NCAP rating could be identified were given the value ‘0 per cent of points achieved’ i.e. ‘0 stars’.

The average Box-specific star assessment was used as Safety Performance Indicator. Box-specific average values of the characteristic ‘Number of stars according to Euro NCAP (0 to 5)’ were estimated per observed inventory date (from 2014) for the overall car fleet. These are based on the newly estimated and standardised box-specific vehicle assessments from Euro NCAP. These mean values can be characterised as ‘Box-specific Safety Performance Indicators of the car fleet referenced to 2020’ and are referred to as

SPI (Box x) of the car fleet

hereinafter.

In this way, a total of four Safety Performance Indicators were formed which are, as mentioned above, box-specific average values of the safety rating of the cars entered in the Central Vehicle Register. Using these four indicators, an Overall Safety Performance Indicator

SPI (overall) of the car fleet

was built (weighted mean value in which – analogue to the Euro NCAP procedure – the assessment results of the four boxes are weighted at a ratio of 40:20:20:20).

As part of the accident analysis, log-linear regression models were estimated in order to determine the effect of the four stated SPIs on corresponding accident characteristics. The overall mileage of the car was taken into account in the modelling as additional explanatory variable. In the analysis of the relationship between the individual Safety Performance Indicators and the accident occurrences, the following data from the official road traffic accident statistics which are matched with the corresponding SPI were used:

SPI_1: Annual number of injured in accidents/fatalities/seriously injured/slightly injured car users aged 12 and over (2013 to 2019)

SPI_2: Annual number of injured in accidents/fatalities/seriously injured/slightly injured car users aged under 12 (2013 to 2019)

SPI_3: Annual number of injured in accidents/fatalities/seriously injured/slightly injured cyclists or pedestrians (each in a collision with a car) (2013 to 2019)

SPI_4: Annual number of accident-involved cars/ cars as main responsible (in accidents involving personal injuries) (2013 to 2019)

6.2 Results

Figure 14 shows the key result of the transfer of the Euro NCAP ratings to the overall car inventory, i.e. the time series of the four Box-specific SPI values as well as the overall SPI value.

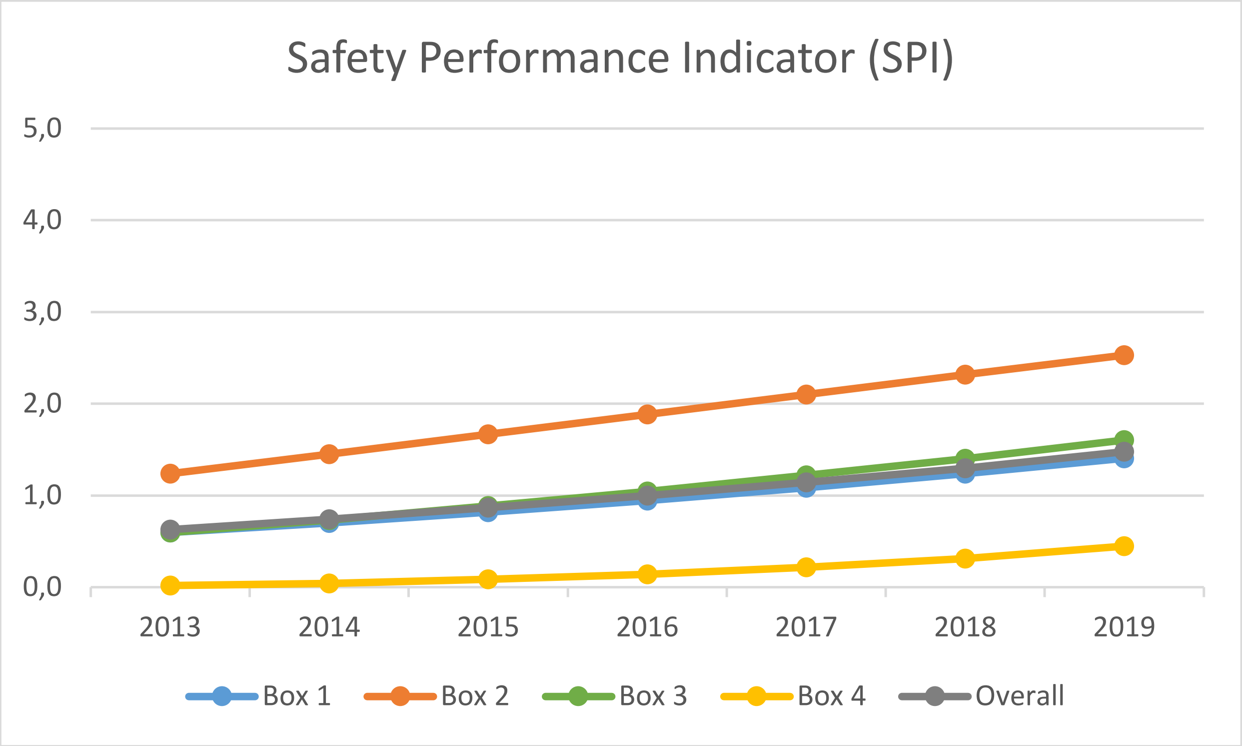


Figure 14: Box-specific SPI and overall SPI of the car fleet 2013 to 2019 (source: own estimations – data basis: Euro NCAP 2009 - 2019; ZFZR 01.01.2014 - 01.01.2020)

The results consistently show a rising trend of the indicators over the course of time. Due to the dynamics of the inventory – more poorly assessed older vehicles are replaced by better assessed newer vehicles – the average vehicle safety of the car fleet continuously rises. The highest indicator values can be found in Box 2 (Child Occupant Protection), they are lowest in Box 4 (Safety Assist Systems). Generally speaking, the level of the SPI values is relatively low (in relation to the value range 0 to 5) which essentially has two causes:

On the one hand, an effect was caused by the fact that, with regard to the reduction factors for the standardisation i.e. comparability of the Euro NCAP ratings (experts), the test conditions of the year 2020 were consistently used as reference. Because of this, for example for Box 4, the test results for the models assessed between 2009 and 2012 were reduced to around one third of their original value. This is reflected accordingly in the values projected to the car fleet.

On the other hand, this is of course also due to the fact that the inventory on a reference date is distributed over many initial registration years. In this context, an effect was also caused by the definition that vehicles with initial registration year of before 2009 are allocated a 0-star assessment using imputation, because before 2009 – as explained above – no (comparable) Euro NCAP results are available.

If the indicators are additionally grouped into car segment, then SUVs, followed by ATVs, commercial vans and the luxury class show the highest values in the overall SPI value. The fact that SUVs have the highest SPI is, however, also related to the fact that this is a relatively new segment in which the share of older vehicles is comparably low. When evaluating the results, the notes for their interpretation, which are given in chapter 4, must also always be borne in mind.

To assess the suitability of the SPIs, which are based on Euro NCAP tests, as indicators of traffic safety on the real-world road network, log-linear regression models were estimated in order to quantify the effect of the four SPIs on each of the corresponding accident characteristics (injured car occupants (adults and children), injured pedestrians and cyclists, accident involved cars). In all analyses, the SPI in the car inventory on 1 January of a year (e.g. 1.1.2014) was matched with the number of accidents and the vehicle mileage of the previous year (so in this case 2013).

In all 18 examined cases, these showed significant associations between SPI and accident occurrences. An important result here: a higher value of the corresponding Safety Performance Indicator is accompanied by a lower number of casualties or accident-involved cars.

On the basis of the estimations, for example, the following specific statements can now be made on the ‘SPI elasticity of the number of casualties’: In the regression model for the (logarithmic) overall number of adult car occupants injured in accidents, the coefficient of log SPI_1 is estimated at -0.4, which shows a (significant) reduction of the number of adult car occupants injured in accidents (12 years and over) by around 0.4%, when the SPI (Box 1) increases by 1%. Safety-related improvements in the car fleet are thus directly reflected in a declining number of victims in road transport.

It is to be taken into account that the estimated parameter value for the corresponding SPI is certainly always negative, but not additionally significant for all 18 estimated regression models. Essentially, this can be attributed to the fact that only the data of the last seven years as observations underlie the model estimations.

In addition, the relationship between the overall SPI value and the resulting monetary economic accident costs was analysed.

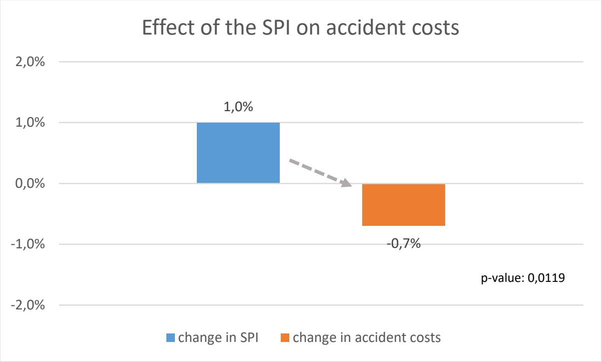


Figure 15: Effect of the SPI overall value of the car fleet on the monetary economic accident costs

The result here was that with an increase of the overall SPI of the car fleet of 1%, the respective accident costs ceteris paribus drop by 0.7% (Figure 15).

6.3 Outlook

In this project, the results of the Euro NCAP tests were used to form Safety Performance Indicators for the overall German car fleet on the reference dates 1.1.2014 to 1.1.2020. A decisive prerequisite for this was the adaptation of the test results from the various years with regard to the test procedures which had changed over the course of time while using the Euro NCAP rating procedure of the year 2020 as reference. This adaptation was largely based on expert judgements, whereby also the changes to the test protocol in 2020 were included, even though only Euro NCAP data up to and including 2019 were observed in this project. This means that reduction factors were also imposed on the results for 2019. However, this was deliberately set in the concept of the project in order to already create the fundamental requirements for an uninterrupted continuation of the time series for the years 2020/2021/2022 (SPI for the car inventory on 1.1.2021 on 1.1.2022 and on 1.1.2023).

The next relevant changes to the test protocols are pending for the year 2023. Here, an adjustment of the procedure will be necessary again in order to be able to quantify the consequences of the changes which will come into effect then (if necessary using further expert judgements) and to be able to adjust the procedure accordingly in order to still arrive at comparable Euro NCAP test results.

References

BASt (2009). Das neue Euro NCAP Rating. Pressemitteilung Nr. 17, 2009. Bergisch Gladbach: BASt.

BASt (2020a). Gurte, Kindersitze, Helme und Schutzkleidung - 2019. Daten & Fakten kompakt Nr. 02/20. Bergisch Gladbach.

BASt (2020b). Volkswirtschaftliche Kosten von Straßenverkehrsunfällen in Deutschland. Bergisch Gladbach: BASt.

BÄUMER, M., HAUTZINGER, H., PFEIFFER, M., STOCK, W., LENZ, B., KUHNIMHOF, T. & KÖHLER, K. (2017a). Fahrleistungserhebung 2014 - Inländerfahrleistung. Berichte der Bundesanstalt für Straßenwesen, Reihe Verkehrstechnik, Heft V 290. Bremen: Schünemann.

BÄUMER, M., HAUTZINGER, H., PFEIFFER, M., STOCK, W., LENZ, B., KUHNIMHOF, T. & KÖHLER, K. (2017b). Fahrleistungserhebung 2014 – Inlandsfahrleistung und Unfallrisiko. Berichte der Bundesanstalt für Straßenwesen, Reihe Verkehrstechnik, Heft V 291. Bremen: Schünemann.

DELANEY, A., NEWSTEAD, S. & CAMERON, M. (2006a). Study of the Relationship between Injury Outcomes in Police Reported Crash Data and Crash Barrier Test Results: An Extended Analysis of German Data. Report of Sub-Tasks 2.1 and 2.2 of the EU Project SARAC II. Brussels: European Commission.

DELANEY, A., NEWSTEAD, S. & CAMERON, M. (2006b). Analysis of Pedestrian Crash Data from Great Britain, Germany and France. Report of Sub-Task 3.4 of the EU Project SARAC II. Brussels: European Commission.

DESTATIS (2020). Verkehrsunfälle – Zeitreihen, 2019. Wiesbaden: Statistisches Bundesamt.

ENDERS, W. (2010). Applied Econometric Time Series, 3rd Ed. Wiley.

ETSC (2001). Transport Safety Performance Indicators. Brüssel.

EURO NCAP (2013). Euro NCAP Rating Review 2013. Brüssel.

EURO NCAP. (2017). Euro NCAP 2025 Roadmap. Leuven.

EURO NCAP (2020). Euro NCAP Rating Review 2018, Version 1.2. Brüssel.

FARMER, C.M. (2005). Relationships of Frontal Offset Crash Test Results to Real-World Driver Fatality Rates. Traffic Injury Prevention 1, pp 31-37.

FOLLMER, R., GEIS, A., GRUSCHWITZ, D., HÖLSCHER, H., RAUDSZUS, D. & ZLOCKI, A. (2015). Marktdurchdringung von Fahrzeugsicherheitssystemen. Berichte der Bundesanstalt für Straßenwesen, Reihe Mensch und Sicherheit, Heft M 258. Bremen: Schünemann.

FOLKSAM INSURANCE GROUP (2019). Folksam's Report „How Safe is Your Car?“ 2019.

GREENE, W.H. (2000). Econometric Analysis, 4th Ed. New Jersey: Prentice Hall.

HAKKERT, A.S., GITELMAN, V. & VIS, M.A. (Hrsg.) (2007). Road Safety Performance Indicators: Theory. Deliverable D3.6 of the EU FP6 Project SafetyNet.

HAUTZINGER, H. (2006). Updated and Extended Description of Existing Car Safety Rating Methods Based on Real-World Crash Data. Report of Sub-Task 1.2-1 of the EU Project SARAC II. Brussels: European Commission.

HAUTZINGER, H., PASTOR, C., PFEIFFER, M. & SCHMIDT, J. (2007). Analysis Methods for Accident and Injury Risk Studies. EU-Projekt TRACE, Deliverable 7.3.

HEIDEMANN, D., HAMACHER, R., HAUTZINGER, H., MOTZKUS, A. & OTTE, D. (2000). Auswirkungen der Anforderungen der EG-Richtlinien und des Verbraucherschutz-Testverfahrens (Euro NCAP) an die passive Fahrzeugsicherheit auf das Unfallgeschehen. Forschungsprojekt FE 82.162/1999 der Bundesanstalt für Straßenwesen.

KATHMANN, T., JOHANNSEN, M., SIEGENER, W., RÖDELSTAB, T., BÄUMER, M. & PFEIFFER, M. (2019). Sicherung durch Gurte, Helme und andere Schutzsysteme 2018 und 2019. Bericht zum Forschungsprojekt 83.0037. Bundesanstalt für Straßenwesen (unveröffentlicht).

KBA (2020). Verkehr in Kilometern - Inländerfahrleistung. Unter https://www.kba.de/DE/Statistik/Kraftverkehr/VerkehrKilometer/verkehr_in_kilometern_inhalt.html?nn=2351604. Abgerufen am 8.10.2020.

KULLGREN, A., LIE, A. & TINGVALL, C. (2010). Comparison between Euro NCAP Test Results and Real-World Crash Data. Traffic Injury Prevention 6, pp 587-593.

KULLGREN A., AXELSSON, A., STIGSON, H. & YDENIUS, A. (2019). Developments in Car Crash Safety and Comparisons between Results from Euro NCAP Tests and Real-World Crashes. Proc. of the 26th techn. conf. on ESV, Paper No. 291, Eindhoven.

KURITZ, S.J. & LANDIS, J.R. (1988). Attributable Risk Estimation from Matched Case-control Data. Biometrics, Vol. 44, pp. 355-367.

LANGWIEDER, K., FILDES, B., ERNVALL, T. & CAMERON, M. (2003). SARAC – Safety Ratings Based on Real-World Crashes for Supplementation of New Car Assessment Programs. Proc. of the 18th techn. conf. on ESV, Nagoya.

LEVIN, M.L. (1953). The Occurrence of Lung Cancer in Man. Acta Unio Internationalis contra Cancrum. Vol. 9, pp. 531-541.

LIE, A., KULLGREN, A. & TINGVALL, C. (2001). Comparison of Euro NCAP Test Results with Folksam Car Model Safety Ratings. Proc. of the 17th techn. conf. on ESV, Paper No. 277, Amsterdam.

LIE, A. & TINGVALL, C. (2002). How Do Euro NCAP Results Correlate with Real-Life Injury Risks? A Paired Comparison Study of Car-to-Car Crashes. Traffic Injury Prevention 4, pp 288-293.

NEWSTEAD, S. & CAMERON, M. (1999). Updated Correlation of Results from the Australian New Car Assessment Program with Real Crash Data from 1987 to 1996. Monash University Accident Research Centre (MUARC), Report No. 152.

NEWSTEAD, S., DELANEY, A., CAMERON, M. & WATSON, L. (2006). Study of the Relationship between Injury Outcomes in Police Reported Crash Data and Crash Barrier Test Results in Europe and Australia. Report of Sub-Tasks 2.1 and 2.2 of the EU Project SARAC II. Brussels: European Commission.

PAGE, M. & RACKLIFF, L. (2006). Deriving and Validating a Road Safety Performance Indicator for Vehicle Fleet Passive Safety. 50th Annual Proceedings, Association for the Advancement of Automotive Medicine, October 16-18, 2006, pp 317-332.

PASTOR, C. (2007). Paired-Comparison Study: Correlation between Euro NCAP Star Ratings and Accident Data from the National German Police Road Accident Statistics. 2nd International Conference on ESAR "Expert Symposium on Accident Research", pp156-161.

PASTOR, C. (2013). Correlation between Pedestrian Injury Severity in Real-Life Crashes and Euro NCAP Pedestrian Test Results. Proc. of the 23rd techn. conf. on ESV, Paper No. 308, Seoul.

SCHIRA, J. (2009). Statistische Methoden der VWL und BWL - Theorie und Praxis. 3., aktualisierte Auflage. München: Pearson Studium.

STRANDROTH, J., RIZZI, M., STERNLUND, S., LIE, A. & TINGVALL, C. (2011). The Correlation between Pedestrian Injury Severity in Real-Life Crashes and Euro NCAP Pedestrian Test Results. Traffic Injury Prevention 6, pp 604-613.

VAN DER ZWEEP, C., PLA, M., WISCH, M., SCHALLER, T., DE HAIR, S. & LEMMEN, P. (2014). Assessment Methodologies for Forward Looking Integrated Pedestrian and Further Extension to Cyclists Safety. Final Publishable Report of the EU Project ASPECSS. Brussels: European Commission.

VAN RATINGEN, M., WILLIAMS, A., LIE, A., SEECK, A., CASTAING, P., KOLKE, R., ADRIAENSSENS, G. & MILLER, A. (2016). The European New Car Assessment Programme: A Historical Review. Chinese Journal of Traumatology 2, pp 63-69.

WOODWARD, M. (2005). Epidemiology. Study Design and Data Analysis, 2nd edition. Boca Raton u.a.: Chapman & Hall/CRC.

ZANDER, O., GEHRING, D.-U. & VAN RATINGEN, M. (2015). Beyond Safety Legislation: Contribution of Consumer Information Programmes to Enhanced Injury Mitigation of Pedestrians during Accidents with Motor Vehicles. Proc. of the 24th techn. conf. on ESV, Paper No. 258, Gothenburg.

ZIV (Zweirad-Industrie-Verband) (2020). Zahlen - Daten - Fakten zum Fahrradmarkt in Deutschland 2019. Präsentation des ZIV-Geschäftsführers Siegfried Neuberger anlässlich der Pressekonferenz am 11. März 2020 in Berlin.

Abbreviations

AEB	Autonomous Emergency Braking
AES	Autonomous Emergency Steering
AIS	Abbreviated Injury Scale
ESC	Electronic Stability Control
Euro NCAP	European New Car Assessment Program
Flex PLI	Flexible Pedestrian Leg Impactor
GIDAS	German In-Depth Accident Study
KBA	Federal Motor Transport Authority (Kraftfahrt-Bundesamt)
CRS	Child Restraint Systems
MPDB	Mobile Progressive Deformable Barrier
ODB	Offset Deformable Barrier
OSM	Occupant Status Monitoring
SBR	Seat Belt Reminder
SPI	Safety Performance Indicator
ZFZR	Central Vehicle Register (Zentrales Fahrzeugregister)

List of Figures

List of Tables

Appendix

Expert estimations on the effects of the changing test procedures on the Euro NCAP test results

As explained in chapter 3.2, for the many revisions of the test programme, there is no sufficiently ‘objective’ (i.e. well-founded and derivable from the test practice) procedure with which these changes can be mathematically integrated into the new estimation of the assessment.

In order to gain a sufficiently valid quantitative estimation of how these changes to the test procedures affect the vehicle rating, expert judgements were therefore drawn upon. As changes to the test procedures tend to mean their intensification, this specifically involves the estimation of reduction factors by which the assessment results for earlier test years will be lowered. If the 2020 test conditions are observed as reference situation, the corrected test results correspond to the hypothetical case that, in the years until 2019, the test procedures of the year 2020 would already have been valid. This means that reduction factors will also be imposed on the results for 2019.

The results of the panel of experts in the form of proportional reductions which are listed in Tab. 6 have been visualised in Figure 16 once more.

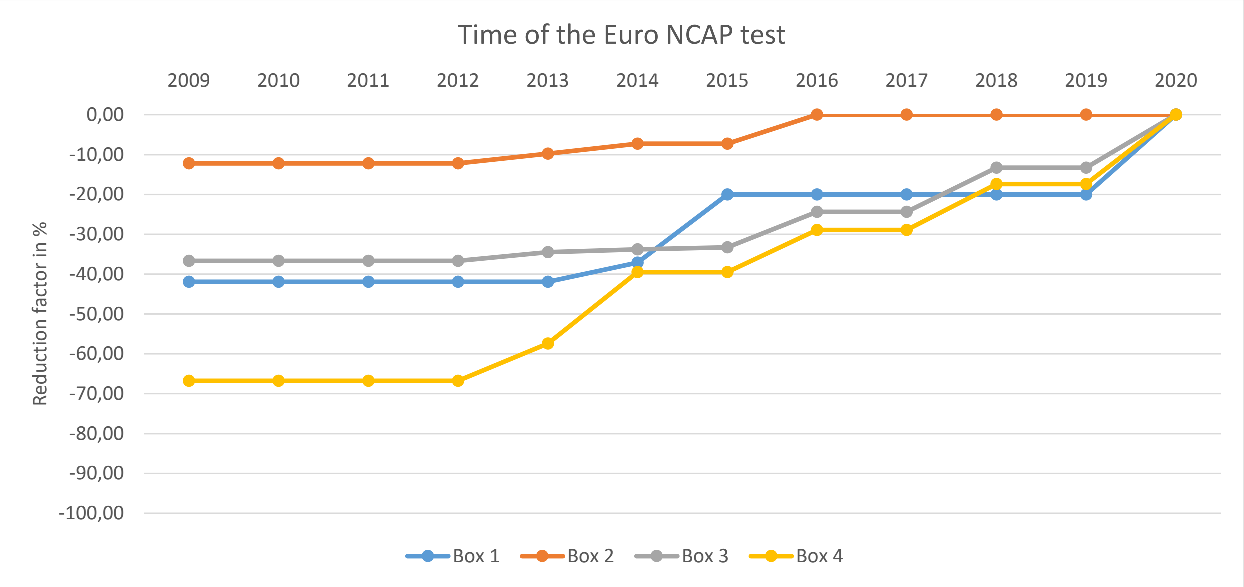


Figure 16: Result of the expert judgement of the Euro NCAP protocol changes since 2009 (reference year 2020)

The change factors are thereby based on the following considerations:

- The *addition of new tests* to the test programme is covered by their additionally achievable points. In particular for systems for active safety, it is assumed that systems were not or only very rarely existent in the assessed models before their introduction into the corresponding test in Euro NCAP. Therefore, these are assessed with zero points. For example, the lane assistant and emergency braking assistance systems were only included in the test programme in 2014 (see Tab. 5 and Tab. 27). As these systems were only installed in a few models of the luxury class before 2014 with only a very low share of the overall vehicle stock[34], an inclusion of these systems in the ratings before 2014 (with 0 points then) would only involve a very small error.
- *Lifting of tests:* No more points have been allocated for the ESC system since 2016 because since then there are no longer any new vehicles without this system. This is taken into account by the fact that as from test year 2016, three fictive points for ESC have been added to the test result (marked red in Tab. 27). The same will also apply in future for the AEB City system (four fictive points from 2020).

---

[34] For market penetration of vehicle safety systems in the inventory (year 2013) see FOLLMER et al. 2015.

- The *intensification of existing* tests were taken into account through percentage reduction factors (expert judgement) to the individual test results of the respective years (depending on the time of the protocol change).

The concrete estimation of the values displayed in Figure 16 is carried out according to the estimation scheme listed in detail in Tab. 27. For each box and test year, the test-specific reductions were first weighted with the achievable point value for the respective test. A new maximum achievable overall point value per test year is determined by summation across all tests of the box (*adjusted max. score*). Finally, these values are related per box to the maximum achievable number of points in the year 2020 (*max. points reference 2020*), from which the displayed change factors then result (*devaluation in % overall*).

As described in chapter 3.2, these deduction factors can then be applied to the *share* of achievable points (so-called percentage value) in order to then be able to determine how the models tested between 2009 and 2019 would have performed under the test criteria and test conditions applicable in 2020.

Box 1		2009	2009	2010	2010	2011	2011	2012	2012	2013	2013	2014	2014	2015	2015	2016	2016	2017	2017	2018	2018	2019	2019	2020	2020
Adult Occupant Protection		Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor
ODB Frontal Impact	ODB Frontalaufprall	16	50,0%	16	50,0%	16	50,0%	16	50,0%	16	50,0%	16	50,0%	8	30,0%	8	30,0%	8	30,0%	8	30,0%	8	30,0%		0,0%
MPDB Frontal Impact	MPDB Frontalaufprall		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	8	0,0%
Full-width Frontal Impact	Frontalaufprall volle Breite		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	8	0,0%	8	0,0%	8	0,0%	8	0,0%	8	0,0%	8	0,0%
Side Impact (MDB)	Seitenaufprall (MDB)	8	25,0%	8	25,0%	8	25,0%	8	25,0%	8	25,0%	8	25,0%	8	15,0%	8	15,0%	8	15,0%	8	15,0%	8	15,0%	6	0,0%
Side Impact (Pole)	Seitenaufprall (Pfahl)	8	20,0%	8	20,0%	8	20,0%	8	20,0%	8	20,0%	8	20,0%	8	10,0%	8	10,0%	8	10,0%	8	10,0%	8	10,0%	6	0,0%
Far Side Impact (MDB & Pole)	Far Side-Test		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	4	0,0%
Whiplash Front Seats	Whiplash Vordersitze	4	0,0%	4	0,0%	4	0,0%	4	0,0%	4	0,0%	2	0,0%	2	0,0%	2	0,0%	2	0,0%	1,5	0,0%	1,5	0,0%	3	0,0%
Whiplash Rear Seats	Whiplash Rücksitze		0,0%		0,0%		0,0%		0,0%		0,0%	1	0,0%	1	0,0%	1	0,0%	1	0,0%	0,5	0,0%	0,5	0,0%	1	0,0%
AEB City	AEB City		0,0%		0,0%		0,0%		0,0%		0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	4	0,0%	4	0,0%	4	0,0%
Rescue	Rettung und Bergung		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	2	0,0%
max. Points	max. Punkte	36		36		36		36		36		38		38		38		38		38		38		42	
Adjusted max. Score	angepasste max. Punkte	24,4		24,4		24,4		24,4		24,4		26,4		33,6		33,6		33,6		33,6		33,6		42	
max. Points (Referenz 2020)	max. Punkte (Referenz 2020)	42		42		42		42		42		42		42		42		42		42		42		42	
	Devaluation in % (overall)		**41,9%**		**41,9%**		**41,9%**		**41,9%**		**41,9%**		**37,1%**		**20,0%**		**20,0%**		**20,0%**		**20,0%**		**20,0%**		**0,0%**

Box 2		2009	2009	2010	2010	2011	2011	2012	2012	2013	2013	2014	2014	2015	2015	2016	2016	2017	2017	2018	2018	2019	2019	2020	2020
Child Occupant Protection		Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor
Dynamic Tests Frontal	Leistung von KRS (Frontal)	16	20,0%	16	20,0%	16	20,0%	16	20,0%	16	20,0%	16	15,0%	16	15,0%	16	0,0%	16	0,0%	16	0,0%	16	0,0%	16	0,0%
Dynamic Tests Side	Leistung von KRS (Seite)	8	20,0%	8	20,0%	8	20,0%	8	20,0%	8	20,0%	8	15,0%	8	15,0%	8	0,0%	8	0,0%	8	0,0%	8	0,0%	8	0,0%
CRS Installation	Einbauprüfung von KRS	12	10,0%	12	10,0%	12	10,0%	12	10,0%	12	0,0%	12	0,0%	12	0,0%	12	0,0%	12	0,0%	12	0,0%	12	0,0%	12	0,0%
Vehicle based	Fahrzeugausrüstungen	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%	13	0,0%
max. Points	max. Punkte	49		49		49		49		49		49		49		49		49		49		49		49	
Adjusted max. Score	angepasste max. Punkte	43		43		43		43		44,2		45,4		45,4		49		49		49		49		49	
max. Points (Referenz 2020)	max. Punkte (Referenz 2020)	49		49		49		49		49		49		49		49		49		49		49		49	
	Devaluation in % (overall)		**12,2%**		**12,2%**		**12,2%**		**12,2%**		**9,8%**		**7,3%**		**7,3%**		**0,0%**		**0,0%**		**0,0%**		**0,0%**		**0,0%**

Box 3		2009	2009	2010	2010	2011	2011	2012	2012	2013	2013	2014	2014	2015	2015	2016	2016	2017	2017	2018	2018	2019	2019	2020	2020
VRU Protection		Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor
Head Impact	Kopfanprall	24	5,0%	24	5,0%	24	5,0%	24	5,0%	24	0,0%	24	0,0%	24	0,0%	24	0,0%	24	0,0%	24	0,0%	24	0,0%	24	0,0%
Leg Impact	Beinanprall	6	6,7%	6	6,7%	6	6,7%	6	6,7%	6	6,7%	6	0,0%	6	0,0%	6	0,0%	6	0,0%	6	0,0%	6	0,0%	6	0,0%
Upper Leg Impact	Hüftanprall	6	3,8%	6	3,8%	6	3,8%	6	3,8%	6	3,8%	6	3,8%	6	0,0%	6	0,0%	6	0,0%	6	0,0%	6	0,0%	6	0,0%
AEB/AES VRU-Pedestrian	AEB Fußgänger		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	6	20,0%	6	20,0%	6	20,0%	6	20,0%	7	0,0%
AEB/AES VRU-Cyclist	AEB Radfahrer		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	6	0,0%	6	0,0%	9	0,0%
AEB Reverse-Pedestrian	AEB Fußgänger rückwärts		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	2	0,0%
max. Points	max. Punkte	36		36		36		36		36		36		36		42		42		48		48		54	
Adjusted max. Score	angepasste max. Punkte	34,175		34,175		34,175		34,175		35,375		35,775		36		40,8		40,8		46,8		46,8		54	
max. Points (Referenz 2020)	max. Punkte (Referenz 2020)	54		54		54		54		54		54		54		54		54		54		54		54	
	Devaluation in % (overall)		**36,7%**		**36,7%**		**36,7%**		**36,7%**		**34,5%**		**33,8%**		**33,3%**		**24,4%**		**24,4%**		**13,3%**		**13,3%**		**0,0%**

Box 4		2009	2009	2010	2010	2011	2011	2012	2012	2013	2013	2014	2014	2015	2015	2016	2016	2017	2017	2018	2018	2019	2019	2020	2020
Safety Assist		Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor	Points	Factor
SBR (Seat Belt Reminder)	Gurtwarner	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	10,0%	3	10,0%	2	0,0%
Occupant Status	Insassenzustandsüberwachung		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	1	0,0%
SAS (Speed Assistance Systems)	Geschwindigkeitsassistent	1	10,0%	1	10,0%	1	10,0%	1	10,0%	3	10,0%	3	10,0%	3	10,0%	3	10,0%	3	10,0%	3	0,0%	3	0,0%	3	0,0%
LSS C2C (Lane Support Systems Car to Car)	Spurassistent		0,0%		0,0%		0,0%		0,0%		0,0%	1	0,0%	1	0,0%	3	0,0%	3	0,0%	4	0,0%	4	0,0%	4	0,0%
ESC	ESC	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%	3	0,0%
AEB/AES CCR (Car to Car Rear-end)	AEB Heckaufprall		0,0%		0,0%		0,0%		0,0%		0,0%	3	20,0%	3	20,0%	3	20,0%	3	20,0%	3	0,0%	3	0,0%	4	0,0%
AEB Junction Assist C2C (Car to Car)	AEB Kreuzung		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%		0,0%	2	0,0%
max. Points	max. Punkte	7		7		7		7		9		13		13		15		15		16		16		19	
Adjusted max. Score	angepasste max. Punkte	6,3		6,3		6,3		6,3		8,1		11,5		11,5		13,5		13,5		15,7		15,7		19	
max. Points (Referenz 2020)	max. Punkte (Referenz 2020)	19		19		19		19		19		19		19		19		19		19		19		19	
	Devaluation in % (overall)		**66,8%**		**66,8%**		**66,8%**		**66,8%**		**57,4%**		**39,5%**		**39,5%**		**28,9%**		**28,9%**		**17,4%**		**17,4%**		**0,0%**

Tab. 27: Estimation schemes to adapt the many revisions of the test program on the basis of expert estimations