Road Safety Risk Evaluation by means of Ordered Weighted Averaging Operators and Expert Knowledge

Elke Hermans*, Da Ruan♣, Tom Brijs, Geert Wets and Koen Vanhoof

Hasselt University
Transportation Research Institute IMOB
Wetenschapspark 5 bus 6
3590 Diepenbeek
Belgium

♣ Belgian Nuclear Research Centre
Boeretang 200
2400 Mol
Belgium

E-mail: elke.hermans@uhasselt.be; Fax: 0032 11 26 91 99; Tel: 0032 11 26 91 41

Abstract: The road safety performance of countries is conducted by combining seven main risk indicators into one index using a particular weighting and aggregation method. Weights can be determined with respect to the assumed importance of the indicator, whereas aggregation operators can be used to stress better performances differently from worse performances irrespective of the indicator's meaning. In this research, both expert weights and ordered weighted averaging operators are explored, evaluated and integrated resulting in a ranking of countries based on a road safety index.

Keywords: road safety indicators, ranking, ordered weighted averaging operators, expert knowledge

* Corresponding author

INTRODUCTION

Research on road safety has lately received a lot of attention in traffic science. Worldwide, an estimated 1.2 million persons are killed in road accidents every year and as many as 50 million are injured [1]. Given this high number of casualties and the resulting suffering and costs, road safety enhancing measures are urgently needed. Better insight into the relative risk performance of countries is vital in this respect. Existing research has basically focused on the identification of risk factors leading to road accidents and casualties (e.g. [2,3]). At the European level, seven main risk factors have been agreed upon, namely alcohol and drugs, speed, protective systems, daytime running lights (DRL), vehicle, roads and trauma management [4,5]. The risk factors are captured by quantifiable safety performance indicators (e.g., the alcohol and drugs factor by the share of road users with a blood alcohol content above the legal limit). A safety performance indicator is defined as any measurement that is causally related to accidents and casualties and used in addition to a count of accidents and casualties in order to indicate the safety performance or understand the process that leads to accidents [4]. Indicator values can then be compared across countries and key problem areas revealed. This makes it possible to select appropriate measures that tackle the main risk aspects and consequently enhance the level of road safety in a country.

Countries could be compared on each indicator individually. However, given the large number of relevant road safety performance indicators, the creation of an overall road safety performance index – which is a combination of road safety performance indicators – is valuable. One of the main advantages of an index over a set of individual indicators is that the overall road safety picture is presented as the different risk factors are joined in this index. Moreover, based on the index scores the countries' ranking can be set up representing the combined performance on essential road safety risk indicators. Similar to the common road safety ranking based on the number of road fatalities per million inhabitants, an easy comparison across countries can be made based on one score and the relative position of a country in the ranking identified. By contrast, the indicators of which the index consists help in taking appropriate action.

The combination of road safety performance indicators in an index is a methodologically intensive process consisting of various steps [6]. In general, an index results from the aggregation of a set of indicator values and a set of weights. In this paper, we focus on two essential methodological aspects, that is the weighting of road safety indicators and the aggregating of road safety indicators. As the concept of indicators and indexes is relatively new in the road safety field, not much attention has been paid to these topics so far. Composing an index as the average over all indicator values is simple yet inappropriate. The importance of two indicators might significantly differ from each other and the idea of full compensation between good scores and bad scores might be unacceptable. Given the policy implications the index should be constructed in a sound way. Therefore, the concept of weighting as well as the field of aggregation should be thoroughly evaluated.

The remaining of this paper is structured as follows: the indicator data are presented in Section 2. Relevant weighting and aggregation concepts are described and applied in Section 3. The results in terms of the countries' ranking are discussed in Section 4. This paper closes with the main conclusions.

ROAD SAFETY INDICATOR DATA

An index will be composed consisting of seven road safety performance indicators. Each risk factor is represented by one appropriate indicator for which reliable and comparable data are available. In international data sources 2003 indicator values were found for 21 European countries. The seven indicators are briefly described. The alcohol and drugs risk is captured by the percentage of road users with a blood alcohol content above the legal limit. The speed indicator measures the percentage of road users exceeding the speed limit on highways. Concerning protective systems the seat belt wearing rate in front of cars and vans is the selected indicator. As a proxy for the daytime running lights factor a categorical indicator is formulated specifying whether no, a partial, or a full DRL law is in place. The vehicle risk is expressed by the share of cars of maximum five years old. The fifth indicator relates to roads and is the network density, defined as the network length divided by the area. Finally, the share of the gross domestic product spent on health care is the trauma management indicator.

All seven indicators are defined in such a way that a high indicator value should be aimed at. A country with high indicator values (e.g., a high seat belt wearing rate) has a good road safety performance. However, each indicator being expressed on a particular scale and in a particular measurement unit, the indicator data will be normalised to all lie between zero (the worst performance) and one (the best performance). As we are aware of the fact that the end result (i.e., the ranking of countries based on their index score) is influenced to some extent by the indicator chosen for each risk factor, the exact values are not employed. Since we expect the relative ordering of the countries for example with respect to alcohol and drugs to be more robust than the specific values on a particular alcohol and drugs indicator, the interval [0,1] is divided into 20 equally spaced parts. The final indicator values relate to the interval [0; 0.05; 0.1; 0.15; ...; 0.95; 1]. E.g., the second best performing country with respect to a specific indicator will obtain value 0.95. Table 1 lists the 21 countries' data set on alcohol and drugs, speed, protective systems, daytime running lights, vehicle, roads and trauma management, which is used in this research. The table can be read both vertically and horizontally. Of all 21 European countries, Sweden scores best on the alcohol and drugs indicator (value 1) while Cyprus scores worst (value 0). In addition, Austria has a better protective systems performance than Belgium. Apart from the relative performance across the countries per indicator (vertically or column based), we are interested in the relative performance across the indicators per country (horizontally or row based). In other words, the seven indicator values of each country are assigned a rank number with (1) referring to the largest indicator value and (7) to the smallest one. That way, the ordering from good areas to problem areas becomes clear for each country. Of all seven risk factors, Austria performs best with respect to daytime running lights and worst with respect to protective systems.

Table 1: Data set

Country	Alc/drugs	Speed	Prot.syst.	DRL	Vehicle	Roads	Tr mngm
Austria	0.450 (4)	0.675 (2)	0.350 (7)	0.875 (1)	0.400 (5)	0.500 (3)	0.375 (6)
Belgium	0.200 (6)	0.475 (4)	0.150 (7)	0.475 (5)	0.850 (2)	1.000 (1)	0.675 (3)
Cyprus	0.000 (6)	0.125 (3)	0.425 (1)	0.100 (4)	0.000 (7)	0.300 (2)	0.100 (5)
Czech Rep	0.600 (2)	0.950 (1)	0.100 (6)	0.475 (4)	0.100 (7)	0.600 (3)	0.275 (5)
Denmark	0.900 (1)	0.475 (7)	0.550 (6)	0.875 (2)	0.700 (3)	0.650 (4)	0.600 (5)
Estonia	0.700 (2)	0.000 (6)	0.300 (3)	0.875 (1)	0.050 (5)	0.250 (4)	0.000 (7)
Finland	0.900 (1)	0.800 (3)	0.800 (4)	0.875 (2)	0.350 (5)	0.000 (7)	0.175 (6)
France	0.250 (7)	0.800 (4)	1.000 (1)	0.475 (6)	0.550 (5)	0.850 (3)	0.900 (2)
Germany	0.550 (5)	0.225 (7)	0.950 (1)	0.475 (6)	0.600 (4)	0.800 (3)	0.950 (2)
Greece	0.050 (6)	0.050 (5)	0.000 (7)	0.100 (4)	0.250 (2)	0.200 (3)	0.850 (1)
Hungary	0.750 (1)	0.225 (5)	0.050 (7)	0.475 (3)	0.200 (6)	0.725 (2)	0.450 (4)

Ireland	0.500 (4)	1.000 (1)	0.600 (3)	0.100 (7)	1.000 (2)	0.375 (5)	0.175 (6)
Italy	0.100 (7)	0.325 (5)	0.225 (6)	0.475 (3)	0.450 (4)	0.550 (1)	0.550 (2)
Netherlands	0.650 (4)	0.475 (6)	0.675 (2)	0.100 (7)	0.650 (5)	0.950 (1)	0.675 (3)
Poland	0.900 (1)	0.675 (2)	0.225 (5)	0.475 (3)	0.150 (6)	0.375 (4)	0.050 (7)
Portugal	0.325 (4)	0.325 (5)	0.750 (2)	0.475 (3)	0.300 (6)	0.150 (7)	0.800 (1)
Slovenia	0.400 (7)	0.800 (4)	0.500 (5)	0.875 (3)	0.950 (1)	0.900 (2)	0.500 (6)
Spain	0.150 (6)	0.600 (2)	0.675 (1)	0.475 (4)	0.500 (3)	0.100 (7)	0.275 (5)
Sweden	1.000 (6)	0.475 (2)	0.850 (1)	0.875 (4)	0.750 (3)	0.050 (7)	0.750 (5)
Switzerland	0.325 (6)	0.125 (7)	0.425 (5)	0.475 (2)	0.800 (4)	0.725 (3)	1.000 (1)
Unt Kingdom	0.800 (4)	0.900 (1)	0.900 (2)	0.100 (7)	0.900 (3)	0.450 (5)	0.375 (6)

The seven values of each country will be combined in an index score using a specific weighting and aggregation method (see Section 3). Based on the 21 index scores a ranking will be produced. This ranking will be evaluated quantitatively by computing the Spearman's correlation coefficient between the index ranking on the one hand and the ranking based on the number of road fatalities per million inhabitants on the other hand. In addition, a qualitative assessment of the methods will be given. In the end, we want to combine indicators in a road safety performance index that has been constructed by means of a sound and acceptable methodology and has a clear link with the number of road fatalities.

3. WEIGHTING AND AGGREGATING

3.1 Weighting concepts

To obtain a composite indicator a decision needs to be made upon the weight to assign to each indicator. In the index literature, a number of weighting methods can be found (see e.g., [7]). In general, indicator weights can be determined based on correlations (factor analysis), experts' opinions (budget allocation or analytic hierarchy process), optimization models (data envelopment analysis) or equally distributed (equal weighting) [8]. Here, we focus on weights which represent the idea of experts concerning the importance of the indicators. In the analytic hierarchy process (AHP) [9] experts are asked to judge the relative contribution of one indicator compared to another one. These pairwise comparisons are somewhat time consuming and may involve some level of inconsistency (e.g., A > B; B > C; C > A). Therefore, we opt for the more simple design of the budget allocation (BA) method. A selected panel of experts is asked to distribute a given budget over the set of indicators in such a way that spending more on an indicator implies that they want to stress its importance.

In general, the BA method has four phases [10]. First, the experts have to be selected. It is important to gather experts with a wide spectrum of knowledge and experience. Second, each expert allocates the predetermined budget of N points to the indicators. In a third step, weights are calculated from these figures. More specifically, the share of budget allocated to an indicator equals its weight. The fourth step is an optional one in which the procedure is iterated until convergence is reached.

3.2 Expert weights

The method used for determining the seven indicator weights is budget allocation. On the one hand, BA is a simple and often used weighting method in the context of indexes (see e.g., [11]); on the other hand, some aspects need to be taken into account. First, the selection of experts is crucial and should be well-considered. It is possible that the results are biased if experts assign a high weight to the indicator on which their country performs well. In this research, experts familiar with the causes and risk factors of road safety were selected. Moreover, as they originated from various countries an international view on the contribution of each risk factor to road safety could be obtained. Secondly, the method may not measure the importance of a specific indicator but the need for political intervention in that dimension [10]. Finally, the maximum number of indicators to distribute the budget over, is limited to ten, enabling the experts to keep an overview [7]. The set of expert weights corresponding to the indicators of alcohol/drugs, speed, protective systems, daytime running lights, vehicle, roads and trauma management is [0.23; 0.31; 0.13; 0.04; 0.07; 0.11; 0.11]. It appears that the experts stress the behavioral aspects speed and alcohol/drugs whereas the daytime running lights factor received the least weight. In the next section, the aggregation of indicators is explained.

3.3 Aggregation concepts

Apart from assessing indicator weights, the aggregation of the indicators needs to be decided on. Aggregation refers to the process of combining values in a single score such that the final result takes all individual values into account in a specific way. Aggregation is a very extensive research field in which numerous types of aggregation functions or operators exist. They are all characterized by certain mathematical properties and aggregate in a different manner. In general, aggregation operators can be roughly divided into three classes [12]: conjunctive, disjunctive and averaging operators. We focus on the class of averaging operators that results in an aggregated score bounded by the lowest and highest indicator values. In other words, an index with scores lying between zero and one is constructed. Within the class of averaging operators, several aggregation operators can be considered. Apart from the well-known and simple weighted mean operators, we focus on ordered weighted averaging (OWA) operators introduced by Yager [13]. They have become very popular within the fuzzy sets community and have some useful properties for the road safety index case. An

OWA operator of an n-dimension is a mapping indicated as $F: R^n \rightarrow R$ that has an associated n vector $W=[w_1,w_2,\ldots,w_n]^T$ such that $w_i \in [0,1]$ and $\sum w_i=1$. Furthermore, $F(a_1,\ldots,a_n)=\sum w_j b_j$ with b_j the j^{th} largest of the a_i. In other words, the aggregation of n indicator values results from multiplying the values sorted in a non-increasing order with the OWA weighting vector and summing these factors.

There are a number of ways to obtain an OWA vector of which linguistic quantifiers [14] is a very common one. Decision makers state a linguistic quantifier Q (e.g., many) which offers a fuzzy description of the portion of criteria required to be met by a good solution. If Q is a regular increasing monotonic quantifier the OWA weights can be obtained as follows [12]: $w_i = Q(i/n) - Q((i-1)/n)$ for $i = 1,...,n$ with most commonly $Q(r) = r^{\alpha}$ with $\alpha \geq 0$. In case of seven indicators these formulas result in: $w_1 = [1/7]^{\alpha}$; $w_2 = [2/7]^{\alpha} - [1/7]^{\alpha}$; $w_3 = [3/7]^{\alpha} - [2/7]^{\alpha}$; $w_4 = [4/7]^{\alpha} - [3/7]^{\alpha}$; $w_5 = [5/7]^{\alpha} - [4/7]^{\alpha}$; $w_6 = [6/7]^{\alpha} - [5/7]^{\alpha}$; $w_7 = [7/7]^{\alpha} - [6/7]^{\alpha}$.

Ordered weighted averaging operators are promising for the road safety context. Very common aggregation operators such as the arithmetic mean operator ($W=[1/n, 1/n,\ldots, 1/n]^T$ or $\alpha=1$) are in fact a special case of the OWA operator. In addition, by means of linguistic quantifiers the attitude of experts or decision makers can be taken into account. For each value of α a different weighting vector is obtained. In the next section, the selection of α and the resulting OWA vector is discussed.

3.4 Ordered weighted averaging operators

From a panel discussion the following principles regarding aggregation were gathered:

A) In case a country scores badly on more than a few indicators, its index score should be small;

B) In case a country scores badly on a few indicators, its index score should be between small and average.

Based on these guidelines an upper and lower limit can be found for α . The first step is to give a specific meaning to the concepts ‘badly’ (with respect to indicator performance), ‘a few’ (with respect to the number of indicators) and ‘small’ and ‘average’ (with respect to the index score). The performance with respect to an indicator will be classified as ‘good’, ‘average’ or ‘bad’ [15]. Here, score 1 is assigned to good; score 0.5 to average and score 0 to a bad performance. Next, ‘a few’ indicators on a total of seven corresponds to two. Finally, the classification of index scores occurs by dividing the possible interval [0,1] into four categories. A ‘small’ index score is 0.25 at most, an ‘average’ index score corresponds to 0.5 whereas a ‘large’ index score is at least 0.75. Having assigned a meaning to all concepts in the linguistic formulations, the next step is to translate it numerically (using the formulas in Section 3.3).

A) $F_\alpha(1,1,1,1,0,0,0) \le 0.25 \Leftrightarrow w_1 + w_2 + w_3 + w_4 \le 0.25 \Leftrightarrow (\frac{4}{7})^\alpha \le 0.25 \Leftrightarrow \alpha \ge 2.477.$

B) $0.25 < F_\alpha(1,1,1,1,0.5,0,0) < 0.5 \Leftrightarrow 0.25 < w_1 + w_2 + w_3 + w_4 + 0.5 w_5 < 0.5$

$\Leftrightarrow 0.25 < (\frac{4}{7})^\alpha + 0.5(\frac{5}{7})^\alpha - 0.5(\frac{4}{7})^\alpha < 0.5 \Leftrightarrow 1.580 < \alpha < 3.235.$

Taking both formulations into account it can be seen that α should be in the interval [2.477;3.235[to obtain an acceptable aggregation. Next, a final OWA vector is obtained by selecting one value for α (i.e., 2.5; 2.6; 2.7; 2.8; 2.9; 3.0; 3.1; or 3.2). Therefore, each of the eight possible values is used to create a weighting vector. By multiplying the ordered indicator values with the OWA vector, index scores are obtained even as a ranking of the 21 countries. We select the α value which produces the strongest relationship with the ranking based on the number of road fatalities per million inhabitants. The result is an α equal to 2.5 and an OWA vector of (0.01; 0.03; 0.08; 0.13; 0.18; 0.25; 0.32).

4. RESULTS

In this research, 21 countries are ranked with respect to their road safety risk performance. In particular, an index is constructed by combining seven indicators according to a specific weighting and aggregation method. The methodological process underlying the creation of three possible countries' rankings is visualised in Figure 1. The main issue is the assignment of weights. The expert weights on the one hand and the OWA weighting vector on the other hand relate to different aspects. The importance of an indicator can be assumed equally or obtained by means of experts' opinions. In this case, the first weight is associated with the indicator of alcohol and drugs. A distinction between good (high) and bad (low) scores is made by incorporating a particular OWA vector. The aggregation policy presented in linguistic statements can be taken into account. The first weight in the OWA vector is linked to the best indicator (with the highest value) in that case.

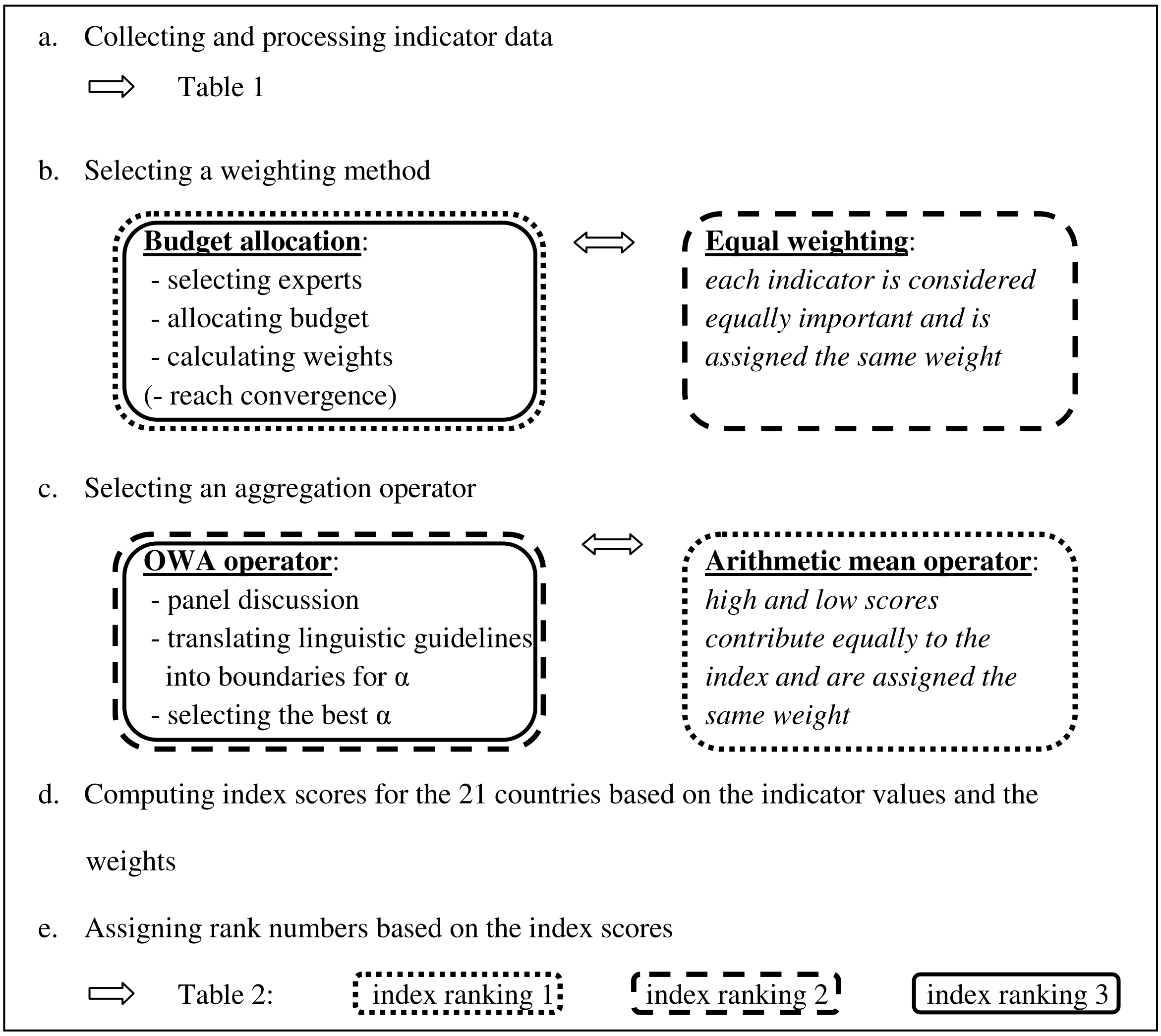

Figure 1: Methodological process for creating countries' rankings based on road safety risk performance

The first ranking corresponds to an index in which the seven indicators are combined using expert weights (Section 3.2). By multiplying the alcohol value of a particular country with the alcohol weight (0.23), the speed value with the speed weight (0.31) and so on, the index score of a country can be obtained by summing these values. Furthermore, with respect to aggregation, the arithmetic mean operator (an OWA vector with α=1) is applied for this index. The second ranking relates to an index created by multiplying the ordered indicator values with the best OWA vector (α=2.5) (Section 3.4). The index score of a country (e.g., Austria) results from the sum of the product of its best indicator value (0.875) with the first OWA weight (0.01) and the product of its second best indicator value (0.675) with the second OWA weight (0.03) and so on. Furthermore, this index assumes that each indicator has a similar importance and equal weights rather than expert weights are applied. The third ranking integrates both ideas: an index in which the seven indicator values are combined using the set of expert weights (to stress the importance of the alcohol/drugs and speed indicator) and the OWA vector (to guarantee a different emphasizing of good and bad performances). The index scores result again from the sum of the product of the seven indicator values and their weights. However, the speed value of Denmark for example highly contributes to its index score because experts indicate speed as the most important indicator of the seven (0.31); moreover, because speed is the worst performance (7) of Denmark, the highest OWA weight (0.32) is assigned to it. All rankings are presented in Table 2.

Table 2: Countries' rankings based on index scores

Country	**Index ranking 1** (expert weights)	**Index ranking 2** (OWA vector)	**Index ranking 3** (expert weights & OWA vector)
Austria	12	8	8
Belgium	13	10	10
Cyprus	21	21	21
Czech Rep	10	16	16
Denmark	5	1	2
Estonia	19	19	19
Finland	6	11	11
France	2	3	4
Germany	9	5	6
Greece	20	20	20
Hungary	16	17	15
Ireland	7	12	14
Italy	18	14	13
Netherlands	8	6	3
Poland	11	18	17
Portugal	15	13	12
Slovenia	4	2	1
Spain	17	15	18
Sweden	3	4	5
Switzerland	14	9	9
Unt Kingdom	1	7	7

A quantitative assessment can be made by computing the degree of correlation between each ranking and the ranking based on the number of road fatalities per million inhabitants. The Spearman's correlation coefficient with respect to the three index rankings is 0.63 respectively 0.70 and 0.67. It can be concluded that there exists a high degree of harmony between the rankings. Rather than only considering the correlation coefficient, the idea of creating an index in which several risk indicators are combined using a sound and appropriate methodology should be the most important aim. Therefore, the more qualitatively oriented conclusion is that the third ranking is the most valuable one as it takes all available information into account: the behavioural risk indicators speed and alcohol/drugs as well as the bad performances of each country are stressed. Moreover, the worst performances of the countries at the top of this ranking have relatively good indicator values.

5. CONCLUSIONS

By comparing relevant risk indicators across countries insight is gained into the good and bad road safety performances of each country. Given the multidimensionality and complexity of road safety, various indicators can be selected. At the European level, alcohol and drugs, speed, protective systems, daytime running lights, vehicle, roads and trauma management are considered to be the main risk factors. Each factor is to be represented by at least one quantifiable indicator. However, combining essential road safety risk indicators in a composite index offers an overall view on road safety risk performance. Based on the index scores countries can be ranked and the evolution in risk can be monitored. Moreover, the underlying indicators offer information on which action to take.

However, the creation of an index requires careful thought. Valuable insight is only obtained in case the index has been constructed well. Essential aspects of the index methodology are the assignment of weights to each indicator and the way of aggregating the indicators. Those two topics were studied here with respect to the combination of seven main risk indicators. First, a group of experts was asked to distribute a budget over the seven risk factors. This budget allocation method revealed the importance given to speed and alcohol/drugs. Secondly, ordered weighted averaging operators were discussed. These aggregation operators are valuable for this case as the linguistic formulation of the aggregation idea of the index can be taken into account by the OWA vector. The results showed that bad performances should be emphasized.

To construct an index using a sound and appropriate methodology this paper showed that both ideas of weights based on expert knowledge and ordered weighted averaging operators based on linguistic formulations can be integrated. The resulting ranking of the 21 countries based on the index scores (index ranking 3) appeared to have a clear link with the common ranking using the number of road fatalities per million inhabitants. In addition, the behavioural risk indicators speed and alcohol/drugs as well as the bad performances were assigned a relatively higher weight thereby reflecting the weighting and aggregation idea of experts and decision makers.

The methodological aspects involved in the index construction process can be further elaborated. Other techniques such as fuzzy preference relations and multi-criteria decision making are planned for our future research on this topic.

REFERENCES

[1] World Health Organization, World report on road traffic injury prevention, WHO, 2004.

[2] S. Hakim, D. Shefer, A.S. Hakkert and I. Hocherman, A critical review of macro models for road accidents, Accident Analysis and Prevention 23 (5) (1991) 379-400.

[3] E. Hermans, F. Van den Bossche and G. Wets, The frequency and severity of road traffic accidents studied by state space methods, Journal of Transportation and Statistics 9 (1) (2006) 63-76.

[4] European Transport Safety Council, Transport safety performance indicators, ETSC, 2001.

[5] SafetyNet, State-of-the-art report on road safety performance indicators, 2005.

[6] E. Hermans, A methodology for developing a composite road safety performance index for cross-country comparison, Ph.D. Thesis, Hasselt University, Belgium, 2009.

[7] M. Saisana and S. Tarantola, State-of-the-art report on current methodologies and practices for composite indicator development, Joint Research Centre, 2002.

[8] E. Hermans, F. Van den Bossche and G. Wets, Combining road safety information in a performance index, Accident Analysis and Prevention 40 (2008) 1337-1344.

[9] T.L. Saaty, The analytic hierarchy process (McGraw-Hill, 1980).

[10] M. Nardo, M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman and E. Giovannini, Handbook on constructing composite indicators, 2005.

[11] F. Pennoni, S. Tarantola and A. Latvala, The 2005 European e-business readiness index, Joint Research Centre, 2005.

[12] M. Grabisch, S. A. Orlovski and R. R. Yager, Fuzzy aggregation of numerical preferences, in: R. Slowinski, ed., Fuzzy Sets in Decision Analysis, Operations Research and Statistics (Kluwer Academic Publishers, 1999) 31-68.

[13] R. R. Yager, On ordered weighted averaging aggregation operators in multi-criteria decision making, IEEE Transactions on Systems, Man and Cybernetics 18 (1988) 183.

[14] L. A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, Computing and Mathematics with Applications 9 (1983) 149.

[15] R. R. Yager and J. Kacprzyk, The ordered weighted averaging operators: Theory and application (Kluwer Academic Publishers, 1997).