SAFETY EVALUATION OF OLDER DRIVERS BASED ON PSYCHOLOGICAL, PHYSICAL AND DRIVING PERFORMANCE

Seddigheh Babaee[1]
Seddigheh.babaee@uhasselt.be

Elke Hermans[1]
Elke.hermans@uhasselt.be

Yongjun Shen[1]
yongjun.shen@uhasselt.be

Mehdi Toloo[2]
mehdi.toloo@vsb.cz

Tom Brijs[1]
tom.brijs@uhasselt.be

Geert Wets[1]
geert.wets@uhasselt.be

Ariane Cuenen[1]
ariane.cuenen@uhasselt.be

1. Transportation Research Institute, Hasselt University, Belgium
2. Technical University of Ostrava, Czech Republic

ABSTRACT

The safety and mobility of older drivers are challenged by several age-related changes, including sensory, motor and cognitive abilities and a decline in these aspects affect the ability to drive safely.

In this study we aim to quantify the overall driving performance of a sample of older drivers using data from an assessment battery and a fixed-based driving simulator. To do so, 55 participants aged 70 years and older completed tests of an assessment battery of psychological and physical aspects as well as knowledge of road signs; In addition, a driving simulator test in which specific driving situations that are known to cause difficulties for older drivers were included.

To evaluate the overall performance of each driver, all the above information was combined by using the concept of composite indicators, which combines single indicators into one index score. In recent years, there has been an increasing interest in the methodology for creating a composite indicator, in which the assignment of weights to each sub-indicator is an essential step. One of the promising weighting methods is data envelopment analysis (DEA) in which based on the data set the best possible weights are determined for each Decision Making Unit (DMU) or driver in our case. In this study, instead of using the standard DEA, Common Set of Weights in DEA (CSW-DEA) is applied for the index construction. By applying the model, index values for each driver is calculated which lies between zero and one with a value equal to one identifying a best performer, whereas a score less than one implies underperforming drivers. In addition to the overall performance of the drivers, more detailed insight can be gained from the assigned weights which can be interpreted as indications of the importance shares of the psychological, physical and driving performance.

KEYWORDS
Safety evaluation; Driving simulator; Data envelopment analysis; Common set of weights; Older driver

INTRODUCTION

In recent years, there has been a growing concern regarding the increased number of older drivers and their potentially decreased driving abilities. Although person's chronological age is not an absolute predictor of driving ability, its impact should not be denied. Ageing is associated with decline in sensory, motor and cognitive abilities which affects the ability to drive safely. The decision to stop driving is not an easy one and it can be associated with negative outcomes such as isolation and depression which adversely affect the quality of life (Wood, et al., 2008). Therefore, it is crucial to enable older drivers to drive safely for a longer period of time. To this aim, it is necessary to develop an appropriate screening tool to evaluate older drivers' performance and improve safe driving ability through intervention as early as possible.

In this study, we aim to quantify the overall driving performance of a sample of older drivers using data from an assessment battery and a fixed-based driving simulator. Performance evaluation plays a critical role in identifying weaknesses and planning goals for improvement. In this regard, composite indicators (CIs) are increasingly recognized as a valuable tool for performance evaluation, benchmarking and policy analysis by summarizing complex and multidimensional issues such as driving performance. One of the critical steps in the construction of a CI is weighting and aggregation which directly affects the quality and reliability of the calculated CI. One of the promising weighting methods is data envelopment analysis (DEA) in which based on the data set the best possible weights are determined for each Decision Making Unit (DMU) or driver in our case. In this study, instead of using the standard DEA, Common Set of Weights in DEA (CSW-DEA) is applied for the index construction. The aim of this model is to determine a set of weights to get the highest efficiency of all DMUs simultaneously.

The rest of paper is as follow: First, the methodology and information about data collection is explained. Then the corresponding results in terms of comparison of 55 older drivers' performance based on their index scores and an illustration of the most problematic parameter(s) for a particular older driver is shown. Finally this paper ends with conclusions.

METHODOLOGY

Participants

77 volunteers aged 70 and older were recruited through the Geriatrics department of the Jessa Hospital with flyers distributed in the hospitals, senior associations and senior flats via local media. Participants had to hold a valid driver's license and still be active car drivers, with no stroke in the last four months and without any indication for dementia as assessed with the Amsterdam Dementia Screening (ADS) test. They had to have the physical ability to complete tests of a clinical assessment battery and simulator driving. Among them, 22 participants were excluded due to simulator sickness. Consequently, 55 participants remained in the sample (mean age = 76.49; standard deviation = 5.40).

Neuropsychological Test Battery
The test procedure consisted of two parts: First a validated neuropsychological test battery of psychological and physical tests was administered at the Jessa Hospital. These standardized tests were selected based on relevance to driving and brain function (AGILE project QLRT-2001-00118). Psychological tests included the Mini Mental State Examination (MMSE), digit span forward (working memory), and three sub-tests of the Useful Field of View: (1) visual processing speed; (2) divided attention and (3) selective attention. Physical tests included visual ability and motor ability. Visual ability (i.e. visual acuity) is assessed with the Snellen chart. Motor ability (i.e. balance) is assessed via the Get-up-and-go test and Four-test balance scale. Accompanied by a neuropsychological assessment, the Road Sign Recognition is used to measure the knowledge of elderly drivers regarding the road signs. Second, a driving simulator test was conducted at the Transportation Research Institute of Hasselt University. A detailed description of these tests is as follow:

Psychological ability tests

The Mini Mental State Examination (MMSE)
MMSE is the most commonly used test for screening cognitive function. It is an 11-questions measure that investigates different areas of cognitive function: orientation to time and place, short term memory, registration (immediate memory), recall, constructional ability as well as language functioning (Folstein et al. 1975). Scores of 25-30 out of 30 are considered normal. The higher the score, the better the psychological ability.

The Digit Span Forward (DSF)
In this test, a random sequence of numbers is read by the examiner and the examinee recalls the numbers back. It assesses attention and working memory, as well as short-term verbal memory (Clark et al. 2011). Scores on this task are divided into four categories (0 = impaired, 1= beneath average, 2 = average, 3 = above average). The more numbers a person can repeat correctly, the better the psychological ability.

UFOV
It is a PC-based test of functional vision and visual attention, which consists of three subtests measuring visual processing speed (UFOV 1), divided attention (UFOV 2), and selective attention (UFOV 3) (Edwards et al. 2005). It is recommended for use as a screening measure in conjunction with a clinical examination of cognitive functioning or fitness to drive. Scores for each subtest are expressed in milliseconds and range from 16.7ms to 500ms. Lower scores correspond with improved visual attention.

Physical ability tests

The Snellen Chart
This test is one of the most common clinical measurements of visual function which is used for measuring visual acuity. (Rosser et al. 2001). Participants have to stand 6m from the whiteboard with several lines of black letters and read the lines. The more lines a person can read, the better the visual acuity with a maximum score of 1.2.

The Get-Up-and-Go test
The Get-Up-and-Go test, also known as Timed Up-and-Go or Rapid Pace Walk, assesses mobility and balance of older adults (Carr et al., 2010). It measures, in seconds, the time taken by an individual to stand up from a standard arm chair, walk a distance of 3m, turn around, return and sit down again (Clark et al. 2011). Scores on this task are divided into three categories (0 = more than 20 seconds, 1= between 11 and 20 seconds, 2= less than 11 seconds). The faster one can complete the task, the better the motor ability.

The Four-test Balance Scale
This test is also used to assess motor abilities; more specifically, lower limb muscle strength and balance, with a maximum score of 1. An individual has to stand on 4 different foot positions of increasing difficulty - standing feet together, standing semi-tandem, standing tandem and one leg standing - for at least 10 seconds without an assistive device (Gardner et al. 2001).

The Road Sign Recognition (RSR)
RSR is used to measure the knowledge of participants regarding road signs with a maximum score of 12 (Lundberg et al. 2003).

Driving data
Driving performance was measured in a medium-fidelity driving simulator (STISIM M400; Systems Technology Incorporated). It is a fixed-based driving simulator (drivers do not get kinesthetic feedback) with a force-feedback steering wheel, brake pedal, and accelerator. The visual virtual environment was presented on a large 180° field of view seamless curved screen, with rear view and side-view mirror images. Three projectors offer a resolution of 1024×768 pixels on each screen and a 60 Hz refresh rate. Data were collected at frame rate.
A 10 min practice session preceding the evaluation was implemented to allow participants to become familiar with the driving simulator. Scenarios that are known to be difficult for older drivers were included in a randomized way. For instance, older drivers are over-represented in crashes occurring while turning off at intersections, where typically the older driver turns against oncoming traffic with right of way on the main road (Hakamies-Blomqvist, 1993; Zhang et al., 1998), gap acceptance while turning left at an intersection (Langford and Koppel, 2006; Yan et al., 2007) and response to signs, signals and road hazards (Horswill et al., 2010). For detailed description of the driving scenario, see Cuenen et al. (2012). The rides took place at inner-city (50 km/h) sections, outer-city (70-90 km/h) sections and highway (120 km/h) sections, in daylight and good weather conditions. The speed limit was indicated by the appropriate sign at the start of each outer-city and inner-city segment and repeated 30 meters after each intersection.
A total of 3 driving measures or indicators are used in the analysis:

 1) Mean-Complete Stop which is computed from 200 meters before reaching the stop sign until the location of the stop sign. Subjects were required to make a complete stop. Cross traffic from left or right occurred when the driver approached the intersection. Complete stop at a stop sign (yes or no) was used to assess whether drivers would comply with Belgian traffic regulations that drivers must make a full stop (i.e., mean driving speed = 0 km/h) at a stop sign (Bao and Boyle, 2008; Jongen et al., 2012).

2) Mean Following Distance is assessed as the average distance between the driver and a lead vehicle with a speed at least 10km/h beneath the speed limit in a road with a speed of 50 km/h and 70 km/h.

3) Mean driving speed is averaged across the different speed limits of 50, 70, 90 and 120 km/h and is measured across separate road segments (i.e., 4.8 km) without any events (Trick et al., 2010).

Older drivers' performance index

In this study, to measure the multi-dimensional concept of driving performance, a composite indicator is created using the Common Set of Weights in DEA with respect to all aforementioned indicators for older drivers (see Figure 1). In constructing CIs, a weight is first assigned to each sub-indicator, and then aggregation function is applied to calculate CIs. All sub-indicators are normalized before aggregation to tackle the different measurement units of the indicators. The methodology is explained in the following section.
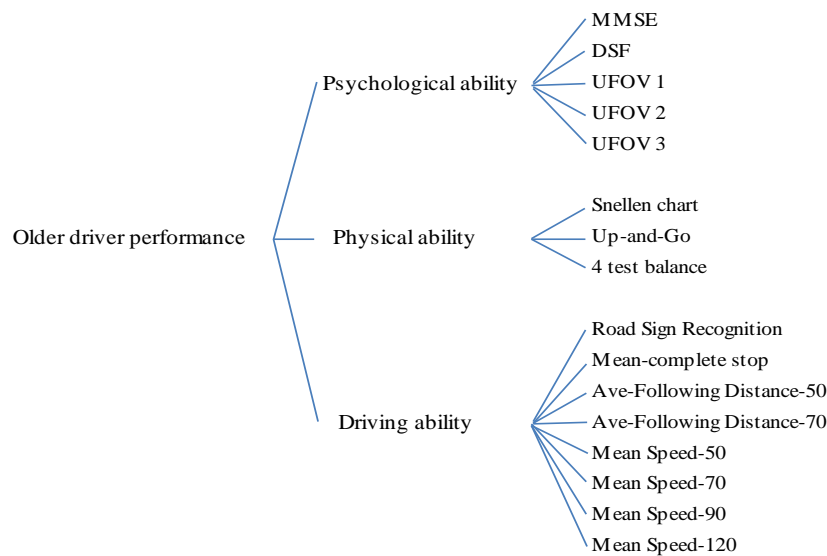

Figure 1. Hierarchical structure of older driver's performance indicators.

Common Set of weights in DEA

Data Envelopment Analysis (DEA) (Cooper et al. 2007) is one of the most commonly used techniques for performance evaluation. It is a non-parametric optimization technique using a linear programming tool to measure the relative efficiency of a set of Decision Making Units (DMUs), or drivers in our study.

Recently, there has been an increasing interest to the application of DEA in the construction of CIs (Despotis 2005; Cherchye et al. 2008; Hermans et al. 2008). By solving a linear programming problem, the best possible indicator weights are determined, and an optimal index score is obtained for each unit, with a higher value indicating a better relative performance. This methodology scales the relative performance between 0 and 1, where 1 represents an *efficient* DMU and other scales indicate *inefficient* DMUs. In this study, to evaluate the driving performance of each older driver by combining all the 16 hierarchically structured indicators in one index score, Common Set of Weights in DEA (CSW-DEA) (Roll et al. 1993) is adopted. The aim of this model is to determine a set of weights to get the highest efficiency of all DMUs simultaneously. Suppose that a set of n DMUs or drivers in our case is to be evaluated in terms of s indicators (y), the model for calculating the Common Set of Weights is as below:

$$\max \sum_{f_1=1}^{s} \hat{u}_{f_1} \left(\sum_{j=1}^{n} y_{f_1 j} \right)$$

s.t. (1)

$$\sum_{f_1=1}^{s} \hat{u}_{f_1} y_{f_1 j} \leq 1 \quad j = 1, \ldots, n$$

where

$\hat{u}_{f_1}$ is the set of most favorable optimal weights for sub-indicators which are obtained by solving the model.

Suppose the integrated model is solved and the optimal solution $\widehat{\mathbf{u}}^* = (\widehat{u}_1^*, \ldots, \widehat{u}_s^*)$ is at hand. The index score of DMU$_k$ with the common set of weight $\widehat{\mathbf{u}}^*$ is measured by $\sum_{f_1=1}^{s} \hat{u}_{f_1}^* y_{f_1 k}$.

RESULTS AND ANALYSIS

There are 55 drivers whose aggregated performances are to be measured based on 16 sub-indicators with different measurement units. For normalization, among existing methods (Freudenberg, 2003), the distance to a reference approach (OECD, 2008) is used since the ratio of two numbers is best kept by this approach. Thereafter, for the older driver performance index construction, Common Set of Weights in DEA is applied (model 1). Table 1 shows the best possible common set of weights for each sub-indicator.

Table 1. Optimal Common Set of Weights for indicators

Categories	Sub-indicators	Optimal CSW
Psychological ability	Mini Mental State Examination	0.1117
	Digit Span Forward	0.0625
	Useful Field Of View 1	0.1159
	Useful Field Of View 2	0.0625
	Useful Field Of View 3	0.0223
Physical ability	Snellen Chart	0.0625
	Get up and Go test	0.0625
	4-test Balance	0.0625
Driving performance	Road Sign Recognition	0.0625
	Mean Complete Stop	0.0625
	Ave. following Distance 50	0.0625
	Ave. following Distance 70	0.0625
	Mean Speed 50	0.1463
	Mean Speed 70	0.1123
	Mean Speed 90	0.0625
	Mean Speed 120	0.0625

As a result, the index score of each older driver is calculated with the common set of weights obtained from model. Table (2) exhibits the CSW-index scores. Index values lie between zero and one with a value equal to one identifying a best performer, whereas a score less than one implies underperforming drivers. As it can be extracted from the above optimal solution, older drivers with ID number 9, 24 and 50 are the best performers and the others are considered as underperformance. It is also possible to compare drivers based on their calculated index scores and rank them.

Table 2. Older drivers' Performance Index scores using the common set of weights

ID	CSW Index score	ID	CSW Index score	ID	CSW Index score	ID	CSW Index score
1	0.969	15	0.6705	29	0.7354	43	0.6356
2	0.7969	16	0.8452	30	0.9226	44	0.9038
3	0.7371	17	0.7402	31	0.7146	45	0.7784
4	0.8333	18	0.7049	32	0.8001	46	0.897
5	0.9239	19	0.9697	33	0.6731	47	0.8298
6	0.7711	20	0.9445	34	0.8074	48	0.9499
7	0.8021	21	0.9811	35	0.8982	49	0.7152
8	0.8962	22	0.7254	36	0.9155	50	1
9	1	23	0.6622	37	0.8133	51	0.853
10	0.6571	24	1	38	0.7752	52	0.6387
11	0.9293	25	0.8314	39	0.8577	53	0.7602
12	0.9804	26	0.8428	40	0.8881	54	0.9964
13	0.8691	27	0.8014	41	0.9721	55	0.7699
14	0.9031	28	0.7023	42	0.998		

Weight Allocation and Required Improvement Priorities

Unlike the CSWs in DEA which determine a set of weights to get the highest efficiency of all DMUs simultaneously, applying a multiple layer DEA based composite indicator model (MLDEA-CI) developed by Shen et al. (2012) will yield the set of most favorable optimal weights for each individual driver. This model is able to take the layered hierarchy of indicators into account that often exists in reality. The main idea of this model is to first aggregate the values of the indicators within a particular category of a particular layer by the weighted sum approach in which the sum of the internal weights equals one. Then, for the first layer, the weights for all the sub-indexes are determined using the basic DEA approach.

This way, more detailed insight can be gained from the assigned weights which can be interpreted as indications of the importance shares of the corresponding indicator. Along with tracking the optimal index score for each individual, the model guarantees acceptable weights through the imposed restrictions. Figure 2 shows the assigned weights (last column) and shares (percentages in the middle) for the case of the worst driver in the data set. As can be seen, the performance with respect to all three performance categories is taken into account in the overall score with the share of psychological ability equal to 36.36%, that of physical ability 27.61% and that of driving ability 36.03%.

More importantly, based on the principle of the maximization MLDEA-CI model, an indicator is assigned a high weight if the driver performs relatively well on that aspect. On the contrary, low weights provide us with valuable information about the aspects requiring most attention for improvement. Therefore, areas of underperformance can be detected, and required improvement priorities can be formulated. Taking the indicators of psychological, physical and driving abilities related to the worst performer as an example, it can be seen that this person is doing relatively well with respect to the psychological aspect (with the highest share of 36.36%) whereas more attention should be paid to the physical ability (with the lowest share of 27.61%), by focusing on motor ability, more specifically lower limb muscle strength and balance, as the lowest weight within this category is assigned to the 4 test balance (0.200). Improvement priorities can also be given within the psychological abilities to UFOV 1 & 2 aspects and within driving abilities to Road Sign Recognition and Mean-complete stop.
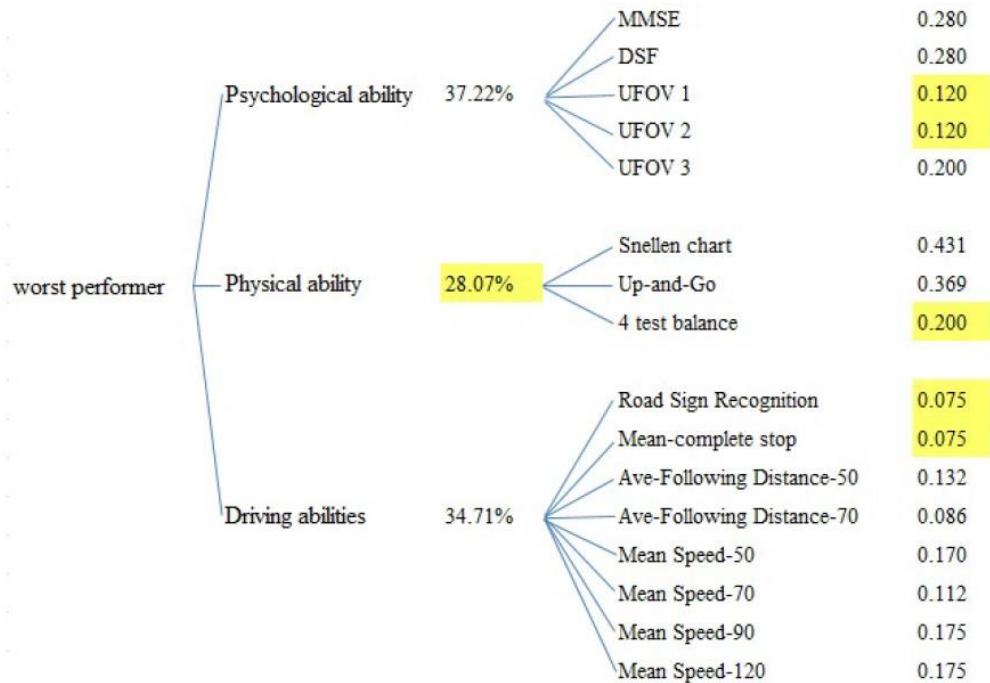
Figure 2. Assigned weights and shares from the model for the case of the worst performer.

CONCLUSIONS

In this study, an assessment battery was used to evaluate older drivers' psychological and physical abilities as well as knowledge of road signs. In addition, by using a driving simulator, the driving performance of participants was measured in situations that are especially difficult for older drivers. Then, all this information was aggregated to construct an overall performance index score in order to quantify the relative performance of individual older drivers. In doing so, a CSWs in DEA was applied and an optimal common set of weights for sub-indicators was obtained. This approach allowed us to rank all the drivers based on their performances.

This study is done to test the methodology. As a limitation of this study, the volunteers participating in this study came from a small sector of the community by invitation and hence non-representative of the elderly drivers population. In case, we can have a truly representative of the elderly drivers population, the obtained optimal CSWs for sub-indicators can be directly used for driving performance index score construction.

Results show that this methodology can be used as an effective screening tool to all drivers whom age-related decline is suspected and whose performance is viewed as a safety concern for themselves and other motorists. In addition, it can assist the elderly driver evaluator to reach a decision about their performance by combining the outcome of various assessment tools. This screening could become a part of the regular process of license renewal.

REFERENCES

Carr, D. B., Schwartzberg, J. G., Manning, L., and Sempek, J. (2010). "Physician's guide to assessing and counseling older drivers." 2nd edition. Washington D.C. National Highway Traffic Safety Administration.

Cherchye, L. et al., 2008. Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index. Journal of the Operational Research Society, 59(2), pp.239–251.

Clark, P.G. et al., 2011. Maintaining exercise and healthful eating in older adults: The SENIOR project II: Study design and methodology. Contemporary Clinical Trials, 32(1), pp.129–139.

Cooper, W.W., Seiford, L.M. & Tone, K., 2007. Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software: Second edition,

Cuenen, A. et al., 2012. Beyond summarized measures : predictability of specific measures of simulated driving by specific physical and psychological measures in older drivers . Road safety in a globalised and more sustainable world - current issues and future challenges, pp.1–18.

Despotis, D.K., 2005. A reassessment of the human development index via data envelopment analysis. Journal of the Operational Research Society, 56(8), pp.969–980.

Edwards, J.D. et al., 2005. Reliability and validity of useful field of view test scores as administered by personal computer. Journal of clinical and experimental neuropsychology, 27(5), pp.529–543.

Folstein, M.F., Folstein, S.E. & McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research, 12(3), pp.189–198.

Freudenberg, M., 2003. Composite indicators of country performance: a critical assessment. OECD Science, Technology and Industry Working Papers, 16, p.35.

Gardner, M.M. et al., 2001. Practical implementation of an exercise-based falls prevention programme. Age and ageing, 30(1), pp.77–83.

Hakamies-Blomqvist, L. E. (1993). "Fatal accidents of older drivers." Accident Analysis and Prevention, 25, 19–27.

Hayez, Q., De Smet, Y. and Bonney, J. (2012), 'D-Sight: a new decision making software to address multi-criteria problems', International Journal of Decision Support Systems Technologies, Vol. 4, No. 4, pp. 1-23.

Hermans, E., Van den Bossche, F. & Wets, G., 2008. Combining road safety information in a performance index. Accident Analysis and Prevention, 40(4), pp.1337–1344.

Horswill, M.S., Kemala, C. N., Wetton, M., Scialfa, C.T. and Pachana, N.A. (2010). "Improving older drivers' hazard perception ability." Psychology and Aging, 25, 464-469.

Langford, J. and Koppel, S. (2006). "Epidemiology of older driver crashes – Identifying older driver risk factors and exposure patterns." Transportation Research Part F, 9, 309-321.

Lundberg, C. et al., 2003. The assessment of fitness to drive after a stroke: The Nordic Stroke Driver Screening Assessment. Scandinavian Journal of Psychology, 44(1), pp.23–30.

Mareschal, B. and Brans, J. P. (1988), 'Geometrical representations for MCDA', European Journal of Operation Research, Vol.34, No.1, pp. 69-77.

Mathieson, P. et al., 2013. SaMERU - Safer mobility for elderly road users. , (September), p.317.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. Handbook on constructing composite indicators: Methodology and user guide. Organisation for Co-operation and Development, 2005b.

OECD, 2008. Handbook on Constructing Composite Indicators

Roll, Y., Golany, B., "Alternate Methods of Treating Factor Weights in DEA," Omega,Vol. 21, 1993, pp. 99- 109.

Rosser, D.A., Laidlaw, D.A. & Murdoch, I.E., 2001. The development of a "reduced logMAR" visual acuity chart for use in routine clinical practice. The British journal of ophthalmology, 85(4), pp.432–436.

Sexton, T.R., Silkman, R.H. & Hogan, J.A., 1986. Data envelopment analysis: Critique and extensions. New Directions for Program Evaluation, 1986(32), pp.73–105. Available at: http://onlinelibrary.wiley.com/doi/10.1002/ev.1441/abstract.

Shen, Y.J. et al., 2011. A generalized multiple layer data envelopment analysis model for hierarchical structure assessment: A case study in road safety performance evaluation. Expert Systems with Applications, 38(12), pp.15262–15272.

Shen, Y. et al., 2012. Data Envelopment Analysis for Composite Indicators: A Multiple Layer Model. Social Indicators Research.

Trik, L. M., Toxopeus, R. and Wilson, D. (2010). "The effect of visibility, traffic density and navigational challenge on speed compensation and driving performance in older adults." Accident Analysis and Prevention. 42, 1661-1676.

Wechsler, D. (1955). "Manual for the Wechsler Adult Intelligence Scale." Oxford, England: Psychological Corporation

Yan, X., Radwan, E. and Guo, D. (2007). "Effects of major-road vehicle speed and driver age and gender on left-turn gap acceptance." Accident Analysis and Prevention, 39, 843-852.

Zhang, J., Fraser, S., Clarke, K. and Mao, Y. (1998). "Age-specific patterns of factors related to fatal motor vehicle traffic crashes: focus on young and elderly drivers." Public Health, 112, 289-295.