

BOETS Sofie, ARNO Patricia  
Belgian Road Safety Institute (BRSI)  
[Sofie.BOETS@bivv.be](mailto:Sofie.BOETS@bivv.be)

## **The use of a screening battery to predict driving performance**

### *Introduction*

The older population in Europe, as in the other western societies, is increasing both in absolute and relative terms. This trend is expected to continue in the coming years leading to a rapidly increasing number of older drivers on roads as a consequence. The mean number of elderly drivers on European roads can be approximated to around 12% of all drivers today and around 20% by year 2010 (EDDIT (V2031), 1994).

Car driving is an everyday task that requires many processes, of which some are automatic, others controlled, to act together, giving the driving task its complex and paradoxical nature. The functions that are involved in car driving are perception, attention, learning, memory, decision-making, and control of action. Because some of these functions may decline during the normal ageing process, and because of the prevalence of some medical age-related conditions, the issue to know whether the elderly drivers pose a particular risk for road safety has been, and is still being, debated.

Designating “older drivers” as a problem appeared at the end of the sixties, mainly in the USA, after the publication of reports showing trends in crash statistics (Hakamies-Blomqvist et al, 2004). Generally speaking, it was shown that although the group of elderly drivers had fewer accidents than their younger counterparts, they were nevertheless over-represented in accident statistics when the driven mileage was taken into consideration. It was concluded that the old drivers group was at-risk from a road safety point of view. The pattern of accidents of the elderly drivers is typical: they are mainly involved in accidents at intersections and in complex traffic situations (OECD, 1985). Moreover, when they are involved in an accident, they are most often judged as being at-fault in comparison to younger drivers. On the other hand though, the older drivers’ accidents occur more often during the day, are mostly independent from alcohol consumption, and their vehicle is not the only one involved.

It is known that many older drivers automatically compensate for their perceived age related decline, by for example slowing down at difficult situations. Yet while this compensation is beneficial for road safety in general, it is nevertheless often insufficient in driving situations at intersections. Indeed, those situations are “forced-paced”, which means that the driver has no control over the information flow so that the quantity of information that needs to be processed rapidly requires too high demands with regard to the available attentional resources of the elderly drivers (Hakamies-Blomqvist et al, 2004; Hakamies-Blomqvist, 1998).

For several years now, scientists have been seeking to design «screening» assessment procedures for the elderly drivers with the aim to exclude from the population of active drivers those persons representing a high risk for road safety. Although some elderly drivers

are more at risk, and this sometimes for age-related reasons (e.g. the presence of a dementia diagnosis, whose incidence increases with age), it is difficult to find correlations between simple functional measures and crash risk. It has been shown that evaluating regularly the whole population of active drivers above a particular age did not lead to a reduction of crash risk in elderly drivers (Hakamies-Blomqvist et al, 1996).

It is not age per se, but the driving-relevant cognitive decline – which is sometimes observed during the ageing process or is related to particular medical conditions – that leads to difficulties while driving. Moreover, all the elderly drivers do not suffer from decline to the same extent, and some even do not experience any decline at all. The inter-individual variance in most of the performance-related variables increases with age. Thus, to be efficient, screening measures should be targeted to specific sub-groups of drivers with an increased crash-risk, instead of focusing on all elderly drivers.

A first approach to identify unsafe older drivers has been to index all illnesses and drugs that would make an older driver unsafe. Yet, it is not the disease itself that influences driving behaviour, but the functional problems caused by the disease ('Safe Mobility for Older People', National Highway Traffic Safety Administration, 1999). Therefore fitness to drive decisions can not be taken on the basis of a medical diagnosis only, but this information can be useful to identify persons who 'may be at risk for unsafe driving' (Dobbs et al, 1998; Eeckhout & Arno, 2002).

A second approach to identify unsafe older drivers is the use of tests of general mental competence, like the Mini Mental State Examination (MMSE; Folstein et al, 1975). Although some studies (Odenheimer et al, 1994; Fitten et al, 1995) found reliable correlations between the MMSE score and the on-road performance, other studies failed to demonstrate the power of the MMSE to discriminate between demented persons with and without a crash record (O'Neill et al, 1992, in Dobbs et al, 1998). Thus, tests giving a global impression of the cognitive functioning are considered to be useful within the scope of an extensive evaluation, but don't seem to be sufficient to be used alone.

A third approach allowing the identification of unsafe older drivers has been the combined use of several tests, gathered in single assessment batteries (for example, Stutts et al, 1998; Marottoli et al, 1998).

However, because neuropsychological tests assess the integrity of separate cognitive functions, the cognitive hierarchical nature of the driving task in terms of operational (related to the mastery of the vehicle), tactical (related to the mastery of driving situations), and strategic (related to the driving context) aspects is not sufficiently taken into consideration (Withaar, 2000), which explains why such analytical approaches generally fail to predict actual driving behaviour. Neuropsychological tests are indeed mostly designed to assess cognitive functions that are most relevant for the operational aspects of car driving, like psychomotor coordination, information processing speed, visuo-spatial abilities and attentional capacity (Van Zomeren et al, 1987). Driving is a complex skill that improves with practice and experience. A 'safe' driver is not only a driver whose skills have developed with practice but it is also a driver who is aware of his/her own capacities and limitations.

Furthermore, the amount of task-specific experience is high in driving while it is low in (most) neuropsychological tests. Knowing that age-related problems of cognitive flexibility and divided attention particularly occur in unfamiliar situations, requiring some degree of central executive function for task-set, indicates that elderly drivers are put in a disadvantage when using neuropsychological tests because they are not familiar with them. On the other hand, highly skilled and much practiced activities, like the driving task, rely more on automatic processes, so that the task can be routinely executed, even by older drivers whose performance is unsatisfying on unfamiliar neuropsychological tests (Botwinick, 1984).

But the main criticism is that cognitive test batteries are unable to grasp all the relevant aspects for driving. Driving is more than the simple addition of different sub-tasks; it is the integration of several complex cognitive tasks, making this activity perhaps the most complex cognitive task of everyday life. Thus even a well-designed cognitive test battery should not be considered sufficient to decide about a person's driving fitness. Instead it can be considered a useful and important first step in the fitness to drive assessment.

Considering all the methodological bothers when using solely a cognitive/neuropsychological battery for detecting unsafe drivers, the EU co-funded research project AGILE (*AGed people Integration, mobility, safety and quality of Life Enhancement through driving*, QLRT-2001-00118) has privileged a cascade-like approach, with different assessment steps.

The AGILE assessment methodology is a new, flexible, and modular system for assessing the fitness to drive of elderly drivers. The AGILE procedure consists of a screening performed at an early stage, followed by an in-depth neuropsychological and behind-the-wheel (real on-the-road tests and simulator scenarios) assessment performed later on, only for those who failed the first tier and are thus possibly at-risk drivers.

This paper focuses on the preliminary validation results of the first tier cognitive screening battery. It was assumed that the use of a short and easily administered test battery would help the medical doctor in a consulting room or primary health care setting to decide which older drivers should be referred to a specialized driving assessment centre. The purpose of such a screening battery should not be to use it as a conclusive instrument but as a first indicator for further referral.

The data considered in this paper were collected over a 10-month period from January to November 2004 at CARA, a department of the Belgian Road Safety Institute since 1987. CARA is the official Belgian centre for determination of fitness to drive and car adaptation. The data consist of the test results of clients who came for an official fitness to drive assessment and who approved to participate in the study (informed consent). All completed the screening battery and a standardized on-road test. Assessors highly experienced in the use of the screening tools (neuropsychologists) and the on-road protocol (occupational therapists) administered the tests. The data of 76 participants tested this way were used for the analyses. Two equivalent pilot studies took place in Greece (HIT) and Sweden (VTI).

One of the major questions to be answered in the AGILE study is whether the screening battery proposed by the AGILE consortium could be sufficiently powerful to predict who will fail the on-road test.

## METHOD

### *Participants*

The age of the 76 participants varied between 63 and 89 years (Mean 78,88; SD 4,758). Twenty-three (23) females (30%) and 53 males (70%) were tested.

The majority (N=58) was sent by an insurance company, mostly after involvement in one or more traffic accidents, the rest (N=18) came in advice of a medical doctor who suspected or had diagnosed a deterioration of the cognitive functioning. All had a full use driving license and more than 20 years of driving experience. No one had marked locomotory or visual deficits.

The aim was to reach an as broad as possible range of cognitive performance, which was succeeded since perfectly healthy elderly drivers as well as demented drivers were included.

### *Screening battery*

The screening battery developed in the framework of this study is a user-friendly and cost- and time-efficient instrument. Overall, combined administration of the screening tests takes around 20 minutes. Involved in the battery are two paper & pencil measures and one newly developed PC-based task.

The screening battery focuses mainly on functions that are often disturbed in the early phase of dementia (Eeckhout et al, 2003). It was preferred to measure all aspects relevant for driving in an integrated way in order that compensation could occur, gaining a much higher validity – which is the most important characteristic for the screening tool – but thus sacrificing some reliability.

Since the screening procedure in this study has to be rather short it was suggested to focus mainly on (visuo-spatial) attention (**TRAILMAKING TEST part A** (TMT-A), Reitan, 1955; see also Withaar et al, 2001; Parasuraman & Nestor, 1991) and executive processing as reflected in the questionnaire of instrumental activities of daily living (**IADL**; Avlund et al, 1993) and a newly developed computerized task **EXECUTIVE CONTROL** (Fimm & Zimmerman, 2003) measuring different aspects of executive control as potential functions with high predictive validity.

Inclusion of these tests/questionnaire was a result of a literature review indicating that visuo-spatial attention is one of the most relevant functions with regard to driving and that the IADL seems to be very sensitive for the first impacts of executive dysfunction on daily life activities in early dementia (Agüero-Torres et al, 2001; Barberger-Gateau et al, 1999; Morris et al, 2001; Teunisse & Derix, 1991). The new PC-test measures aspects of working memory, selective visual attention, inhibition and mental flexibility on an intermediate level. Executive deficits are described frequently even in the early phase of Alzheimer's disease (Nathan et al, 2001).

### *On-road test protocol and expert judgement*

A modified version of the TRIP (*Test ride for Investigating Practical fitness to drive*) was used as the reference on-road test to validate the screening battery. This test is the result of joined efforts of different instances (Dutch licensing authority (CBR) and the University of

Groningen-Department of Neuropsychology and Gerontology) to develop a scoring system aiming at standardizing on-road evaluations (Brouwer & Withaar, 1997). In this scoring protocol, different aspects of the driving task are scored by means of a behavioural checklist. Each of the different aspects contains quantitative items to be scored on a four-point rating scale according to the performance. Scores range from '1 = insufficient', via '2 = doubtful', and '3 = satisfactory', to '4 = good'. When a specific traffic situation or driver action was avoided for safety reasons, a zero-score is assigned. Scoring occurred immediately after completion of the on-road test. In the protocol the expert assessor is also asked to give his general impression of the performance and rate the compensatory behaviour.

In order to specifically address the fitness to drive of elderly people, De Raedt et al. (2001) adapted the first version of the TRIP. In the scope of this study, this version was further adapted and finally consisted of 7 dimensions and 3 specific situations.

The dimensions are: (1) Lateral position on the road (lateral positioning, steering firmness, position on the lane choice), (2) Car following distance (style, adaptation), (3) Speed control (style, adaptation), (4) Visual behaviour & general communication (eye/head movements, communication with other road users, use of direction indicator), (5) Traffic signals (perception and reaction), (6) Mechanical operations (fluency, swiftness of steering and pedal control) and (7) Defensive behaviour (tactical anticipatory behaviour in changing situations).

The traffic situations specifically targeted in this TRIP version, are: (1) Turning left (approaching and on a crossing or junction: different aspects), (2) Joining the traffic stream (on the highway: different aspects) and (3) Roundabout (approaching and on roundabouts: different aspects).

The test ride took place on a standardised trajectory of 33km, starting in the premises of the Belgian Road Safety Institute, then proceeding in real traffic, from a neighbourhood with very low to moderate traffic progressing to areas with more traffic, and continuing on a six- and eight-lane double carriage highway before coming back and terminating at the Institute.

At the end of this on-road assessment and scoring, the assessor is asked to give his *expert judgement* about the person's fitness to drive. Possible judgements are: fit, doubtful fit (i.e. restrictions in use, in need of lessons) and definitively unfit.

## RESULTS

### *Participant group comparison*

Validation of the screening test battery is performed through calculating the most powerful (in terms of sensitivity and specificity) screening model to correctly predict the probability to belong to one of two groups, namely the group of fit to drive participants (those who can continue driving without further assessment) and the group of to be referred participants (those who have driving-related problems and thus should be assessed in-depth). These groups are determined by the fitness to drive judgement from the expert who assessed the TRIP on-road test, which is considered the golden standard in this study.

The majority of the sample of 76 participants was judged fit to drive (N=50). Thus 26 participants either were declared fit with restrictions in use, obliged to follow lessons or declared definitively unfit to drive. This is the group that would have to be referred for further in-depth assessment after the screening phase.

Comparing both groups shows that they are equivalent on age ( $t=-0,427$ ;  $p>0,05$ ). This corresponds to the idea that it is not age per se that distinguishes safe from unsafe drivers. Table 1 summarizes the descriptive data. As can be seen, women are a minority in the sample, but they have an equal proportion in both groups.

Expert judgement	Fit to drive group	To be referred group	t	Sig. (2-tailed) $p\leq 0,05$
Number (N)	50	26		
Sex				
Male	35	18		
Female	15	8		
Age (years)				
Mean (SD)	78,69 (4,404)	79,22 (5,449)	-0,427	no
Min-Max	67-89	63-89		

**Table 1:** Descriptive data of the expert judgement based Fit to drive group and To be referred group, and the significant age difference (t-test).

There was no significant correlation between the age and the different screening variables.

T-tests and Mann-Whitney U-tests were used to get a general idea of which of the screening battery variables discriminate significantly both groups. The results show that the time needed to fulfil the TMT-A differs very significantly ( $t=-4,465$ ;  $p\leq 0,05$ ) in an expected way. On average the fit group was able to correctly connect the numbers in less than a minute (56 sec.) while the to be referred group needed on average 77 seconds, indicating that the participants with diminished fitness to drive are significantly disadvantaged on integrated aspects like visual information processing speed, visuo-motor tracking and eye-hand co-ordination speed. Four of the 5 most relevant measures of the Executive Control task are also significantly discriminative, indicating a link between early executive functions decline and driving capabilities deterioration. The variables are: the total number of correct responses ( $U=287,5$ ;  $p\leq 0,05$ ), omissions ( $U=337$ ;  $p\leq 0,05$ ) and errors ( $U=367,5$ ;  $p\leq 0,05$ ), and also the total variance in response times ( $U=363,5$ ;  $p\leq 0,05$ ). The median time to react shows only a trend towards distinguishing the groups ( $U=394,5$ ;  $p=0,053$ ). Furthermore, one functional measure of the IADL is also significantly different in both groups: the IADL ability index ( $U=373$ ;  $p\leq 0,05$ ). The IADL reduced speed index ( $U=404,5$ ;  $p\leq 0,05$ ) shows a trend. Table 2 presents the descriptive data for the different relevant screening tests variables for both expert judgement groups and the significance level of each one's discriminative power.

Expert judgement	Fit to drive group Mean (SD)	To be referred group Mean (SD)	t/U	Sig. (2-tailed) $p\leq 0,05$
IADL ability index	0,9687 (0,0627)	0,8959 (0,1185)	373	Yes**
IADL tiredness index	0,9180 (0,0847)	0,8651 (0,1333)	436	No
IADL reduced speed index	0,9031 (0,0928)	0,8412 (0,1345)	404,5	No
TMT-A	55,59 (15,246)	77,22 (20,747)	-4,465	Yes**
EC median RT	762,88 (167,593)	843,83 (197,200)	394,5	No
EC SD	210,92 (91,015)	265,91 (102,014)	363,5	Yes*

EC total N correct	35,81 (5,786)	30,30 (9,295)	287,5	Yes**
EC total N error	8,58 (7,724)	12,39 (9,059)	367,5	Yes*
EC total N omission	3,67 (5,317)	7,39 (7,127)	337	Yes**

**Table 2:** Means, SD of individual screening test variables for expert judgement based fit to drive and to be referred group. \*\*Significant to level 0,01, \*Significant to level 0,05

### *General group sensitivity and specificity*

The major issue to be determined is the extent to which the screening battery can predict the on-road driving evaluation, as expressed by the expert judgement. To determine which combination of the screening tools variables was most powerful, logistic regression analysis was used.

The selection of the variables to include in the analyses depended on their individual predictive power as shown in some first exploratory logistic regression analyses and on the consideration of the problem of multicollinearity.

The model that proved to be the most powerful with a good reliability – shown by the identical results in the forward and backward stepwise models – included only the TMT-A variable (Wald=14,226, p=0,000). No measures of the IADL questionnaire nor of the Executive Control test could improve this model. Adding age also didn't improve the model's predictive value.

The significant model (Models Chi Square = 27,426; df = 1, emp. sign = 0,000) explains a fairly good amount of the variance (Nagelkerke R<sup>2</sup> = 0,486; Cox & Snell R<sup>2</sup> = 0,344) and fits the data (Hosmer & Lemeshow test: Chi-square=2,585, df=7, p=0,921).

Looking at the classification table, the sensitivity of this model at a cut-off value of 0,20 is 85%, with a specificity of 73,3%, and there is an overall predictive power of 76,9% (see table 3). There are three false negative, and 12 false positive results. In this study sample, three participants with actual driving-related problems would incorrectly be considered fit to drive without restrictions while 12 participants without problems would have been sent for further in-depth assessment.

		Expert judgement (criterion)		% correct
		To be referred group	Fit to drive group	
Screening outcome (prediction)	Fail	17	12	58,6
	Pass	3	33	91,6
% correct		85	73,3	76,9

**Table 3:** Classification table of the model with TMT-A (time) at a probability cut-off value of 0,20.

At a cut-off value of 0,18 the sensitivity would increase to 90% but in this case a specificity of 62,2% would have to be accepted. The overall predictive value of the model at this cut-off value would be lowered to 70,8%.

The determination of the threshold varies as a function of the background philosophy (cost-efficiency, mobility/safety). Even further lowering the proposed cut-off value to 0,15 would increase the sensitivity to 100%, but, on the other hand, the specificity would in that case be

lowered to 33,3%, referring thus almost everybody for in-depth assessment, which would increase drastically the assessment costs.

#### *Exploration for a more powerful screening battery*

Addition of the two other proposed screening tools (IADL and Executive Control) in the logistic regression analyses proved not to increase the overall predictive power.

Looking in an explorative way at the results derived from the forward and backward logistic regression analyses including, besides the TMT-A, measures of an in-depth assessment tool also measuring the executive functions (Flexibility test, Zimmerman et al. (1996)), indicated that the combination of only the TMT-A (time) (Wald=4,466, p=0,035) and the Flexibility test SD in the alternating condition, (Wald=9,434, p=0,002) led to a very predictive and robust model with a sensitivity of 90% and a specificity of 80% (see table 4).

The significant model (Models Chi Square = 30,847; df = 2, emp. sign = 0,000) explains a good amount of the variance (Nagelkerke R<sup>2</sup> = 0,514; Cox & Snell R<sup>2</sup> = 0,369) and fits the data (Hosmer & Lemeshow test: Chi-square=11,802, df=8, p=0,160).

		Expert judgement (criterion)		% correct
		To be referred group	Fit to drive group	
Screening outcome (prediction)	Fail	20	9	68,9
	Pass	2	36	94,7
% correct		90,9	80	86

**Table 4:** Classification table of the model with TMT-A (time) and Flexibility test (alternating condition) SD at a probability cut-off value of 0,25.

## **DISCUSSION**

The aim of the study presented here was to develop and assess a screening battery to be used by health care professionals to identify among their patients those possibly at-risk and who consequently should be advised to undergo a deeper evaluation in a specialist driving assessment centre.

The present findings are extracted from the data acquired in a sample of 76 drivers aged 63 and above (mean: 79) referred to the official Belgian Centre for fitness to drive determination (department CARA of the Belgian Road Safety Institute). The results suggest that it may be possible to predict in elderly drivers the outcome of a driving test using only the TMT-A (time) (Reitan, 1955).

Based on this test alone, 85% of the screened persons with problems influencing driving can actually be caught and send for in-depth assessment. The amount of false negative classifications is thus minimized to an acceptable rate of 15%. On the other hand, 26,7% of the screened persons that are actually fit, are send for in-depth assessment too. This is not considered a problem because this stage corresponds to the first step of a hierarchically tiered assessment procedure. Reaching the highest possible sensitivity with the battery by capturing

all the to be referred persons, while keeping to a minimum the number of false negative results was considered the most important aspect to focus on.

Increasing the sensitivity of this model to 100%, would lead to a specificity of only 33,3%. Although this would be the best solution from a road safety perspective, two-third of the fit persons would be referred for further assessment, which would increase enormously the costs for assessment in an operational system. The best model is the one that provides a good balance between road safety and individual mobility while being cost-efficient. It must be underlined that at an individual level, the likelihood for an aged driver to be involved in an accident is relatively low. In other words, the error of not referring people who could nevertheless be at risk should not have dramatic consequences.

The TMT-A was included in the battery to target the complex visual scanning with a motor component, a function known as being very relevant to driving (Odenheimer et al, 1994; De Raedt & Ponjaert-Kristoffersen, 2001; Janke & Eberhard, 1998). The importance of visual information processing speed and visual scanning are clearly confirmed in this study.

It had been argued that a global index of functional status may be a better predictor of driving ability than results of cognitive tests (Lundberg et al, 1997). To assess the driver's functional status, an IADL questionnaire was thus added to the battery. This questionnaire aimed at targeting elderly drivers with cognitive deficits like those caused by (early) dementias in 'pathological' aging. According to this study's results, such a questionnaire does not improve the prediction power of the battery. Debriefing with assessors put in evidence difficulties in terms of the administration. Problems of subjective answering and impossibility to check the correctness of answers (e.g. no verification with a relative) possibly contributed to a lack of significant findings for this questionnaire. Given that GP's know their patients very well, and thus better than would do unfamiliar assessors from an official centre, it can be hypothesised that the scoring by GP's would be more reliable. Another difficulty actually was related to the content of the questionnaire, which referred more to 'womanised' tasks, and thus many men answered that a question was not relevant, leading to a lot of missing values. In the scope of this study, guidelines and propositions can be given to construct an adapted IADL questionnaire.

The Executive Control test was specifically developed to assess executive functioning at an intermediate level. It was assumed that in order to better predict driving performance in a first assessment tier, the most important aspects of the driving task have to be assessed in an integrated and dynamic way, taking into consideration its operational, but also to some extent, its tactical and strategic levels. Preference was thus given to a short and easy PC-based test covering the largest possible range of problems and focussing on executive functioning.

Our study results fail to show that the test improves the amount of variance that is explained. Such a failure to demonstrate the effect can be attributed to several possibilities. It could be suspected that executive functioning is not useful for screening elderly drivers, or that the subjects in this study are all good with regard to the executive functions, but this proves not to be the case. Another possibility might be that executive functioning helps in the prediction for people for whom there is problem in another modality only. Thus, further analyses should be done using the classification tree method. Another explanation for the lack of a significant

result, might be related to methodological problems. In contrary to the test administration indications, written reminders for the response keys were used, lowering thus the demands on the working memory, a function that is rapidly decreased in early pathological aging.

The absence of an effect of the Executive Control in the predictive power of the battery could also mean that the way the test is designed is not sensitive enough to slight impairments in executive functioning or that this test doesn't tap the part of executive functioning that is relevant for predicting driving performance. To test this possibility, we performed logistic regression analyses including another test known to tap executive functioning, namely the Flexibility test from the TAP battery (Zimmerman & Fimm, 1996). The results show that the predictive power was increased. Using only the TMT-A (time) and the Flexibility test makes it possible to predict in this sample of elderly drivers the outcome of the driving test with an overall accuracy rate of 86%. Based on this model, 90,9% of the persons with troubled driving capabilities are correctly captured and sent to an assessment centre. And, on the other hand, 80% of the fit persons correctly passed this battery.

Although this is a post-doc finding, it can not be neglected. This result seems to support our hypothesis that a test sensitive to early age-related cognitive decline – i.e. deterioration of the executive functions – may help to predict in combination with a visual search task, driving competencies in the elderly. The differences between the two tests measuring a priori the same functions relies in their difficulty level. The Flexibility test can be considered a more difficult test, also relying much on the working memory function, which was excluded in the Executive Control test. Knowing that with increasing difficulty executive functions are always involved could indicate that the Executive Control test – the way it was administered in this study – maybe was not difficult enough to recruit executive functions totally.

Drawback against a solution for the screening including the Flexibility test is that this test is not an easy task, not for the clients but nor for the person administering the test (the GP's). This would indicate that such a test is not a good candidate for a screening, which is also the reason why that test was first proposed for the in-depth assessment.

On the other hand, these results nevertheless suggest that predicting driving performance is not straight forward and indeed might require a more complex test for executive functioning. Fitness to drive is undoubtedly a complex concept. Using the proposed set of screening tools it was failed to demonstrate that it is possible to filter out the unfit people with too simple tests. But nevertheless, selecting the to be referred persons in this study sample seems to be possible with only two tests assessing visual search and executive functioning.

Given that the administration of tools like the Flexibility test requires the involvement of trained professionals, a possible suggestion in the scope of this study might be that a screening should be performed in hospital facilities (e.g. 'memory units', where for example differential diagnosis between dementia illnesses and depression is performed) instead of at a GP setting, as was initially proposed. Since the aim of AGILE is to target at-risk people, this seems a logical solution. After all, in hospitals the probability to have at-risk people increases.

It can thus be proposed to use the TMT-A in combination with an executive functions test as a screening battery. The executive functions test can be the Flexibility test (Zimmerman &

Fimm, 1996) taking into consideration the limitations inherent to the difficulty of its administration or another test can be designed. Furthermore, suggestions can be made for an improved IADL questionnaire, with for example multi-sex items.

This pilot study has proved that predicting driving performance of elderly drivers by means of a short screening instrument could be the first step of a multi-tiered assessment system, where co-operation between GP's, hospital facilities, and official driving assessment centres is needed.

In conclusion, as neither age, per se, or medical diagnosis are good predictors of fitness to drive, driving-related assessment has to focus on detecting possible cognitive decline. Given the difficulty to assess driving-related cognitive impairment in a consulting room or a primary health care setting, the assessment battery described in this paper can be used as a screening instrument by health care professionals to decide which ones among their patients should be advised to undergo a detailed evaluation with regard to their ability to drive in a specialist driving assessment centre. The paper & pencil test TMT-A is in the public domain and is very easy to use. The Flexibility test is a computerized task that requires no specific equipment except a PC.

In the operational three-tiered structure, the presented screening battery could be used in a first assessment tier by any health care professional. Then, for those people detected as being potentially at-risk, a second tier assessment could take place in a specialist driving assessment centre. There, analytic-like approaches should allow the understanding of the impairments and ways to overcome them. At the end of such a cascade procedure a driving assessment in real traffic should take place in cases of impairments in order to make recommendations in terms of technical adaptations to supply to the vehicle, of conditions related to the functional status of the driver to be enforced (e.g. wear of glasses), or of restrictions related to the driving licence use (e.g. driving allowed only in daylight condition) to be imposed. Thus, the full assessment procedure should be restricted as much as possible to "at risk" cases only. Such a step-by-step procedure is cost-effective and ethical. The aim of any fitness to drive assessment should be to ensure that a person is able to drive safely despite some limitations or impairments at lower functioning levels. In such a context, the focus lies mainly on behaviour and attitudes towards road safety in general, instead of rigidly on the driver's driving skills.

## REFERENCES

- Agüero-Torres, H, Hilleras, PK, & Winblad, B. (2001). Disabilities in activities of daily living among the elderly. Current Opinion in Psychiatry, **14**:355-359.
- Avlund, K, Schultz-Larsen, K & Kreiner, S. (1993). The measurement of Instrumental ADL: Content validity and construct validity. Aging Clin Exp Res; **5**(5):371-383.
- Barberger-Gateau, P, Fabrigoule, C, Rouch, I, Letenneur, L & Dartigues, JF. (1999). Neuropsychological correlates of self-reported performance in instrumental activities of daily living and prediction of dementia. J Gerontol B Psychol Sci Soc Sci; **54**(5):293-303.
- Botwinick, J. (1984). *Aging and Behaviour*, 3<sup>rd</sup> Edition. New York: Springer.
- Brouwer, WH, Waterink, W, Van Wolffelaar, PC & Rothengatter, T. (1991). Divided attention in experienced young and older drivers: lane tracking and visual analysis in a dynamic driving simulator. Hum Factors, **33**, 573-582.
- De Raedt, R, Ponjaert-Kristoffersen, I. (2001). Short cognitive/neuropsychological test battery for first-tier fitness-to-drive assessment of older adults. Clinical Neuropsychologist Aug; **15**(3):329-36.
- Dobbs, AR, Heller, RB & Schoplocher, D. (1998). A comparative approach to identify unsafe older drivers. Accident Analysis and Prevention, **30**: 363-370.
- EDDIT (V2031). (1994). "Study of Age-Related interaction between the complexity and size of symbolic guidance information". AGILE project, *Del. No. 21*.
- Eeckhout, G & Arno, P. (2002). Age-related diseases to be prioritised within AGILE. AGILE project. *Internal Report*.
- Eeckhout, G, Arno, P, Fimm, B & Rothermel, S. (2003). The pre-screening tool. AGILE project. *Public Report D3.1*.
- Fimm, B. & Zimmerman, P. (PSYTEST) (2003). AGILE Neuropsychological test battery. AGILE project. *Internal Report D4.1*.
- Fitten, LJ, Perryman, KM, Wilkinson, CJ, Little, RJ, Burns, MM, Pachana, N, Mervis, R, Malmgren, R, Siembieca, DW & Ganzeil, S. (1995). Alzheimer and vascular dementias and driving A retrospective road and laboratory study. Journal of the American Medical Association, **273**, 1360-1365.
- Folstein, ML, Folstein, SE & McHugh, PR (1975). Mini Mental State: A practical method for grading the cognitive status of patients for the clinician. Journal of Psychiatric Research; **12**:189-198.

Hakamies-Blomqvist, L. (1998). Older drivers' accident risk: conceptual and methodological issues. Accident Analysis and Prevention, **30**: 293-297.

Hakamies-Blomqvist, L, Johansson, K & Lundberg, C. (1996). Medical screening of older drivers as a traffic safety measure – a comparative Finnish-Swedish evaluation study. Journal of the American Geriatrics Society. **44**: 650-653

Hakamies-Blomqvist, L, Siren, A, Davidse, R. (2004). Older drivers – a review. VTI Rapport 497A-2004. Linköping, Sweden.

Hakamies-Blomqvist, L & Wahlström, B. (1998). Why do older drivers give up driving? Accid Anal and Prev; **30**(3):305-312.

Janke, MK & Eberhard, JW (1998). Assessing medically impaired older drivers in a licensing agency setting. Accident Analysis and Prevention, **30**: 347-361.

Lundberg, C, Johansson, K, Ball, K, Bjerre, B, Blomqvist, C, Braekhus, A, Brouwer, WH, Bylsma, FW, Carr, DB, Englund, L, Friedland, RP, Hakamies-Blomqvist, L, Klemetz, G, O'Neill, D, Odenheimer, GL, Rizzo, M, Schelin, M, Seideman, M, Tallman, K, Viitanen, M, Waller, PF & Winblad, B. (1997). Dementia and driving: an attempt at consensus. Alzheimer Dis Assoc Disord; **11**(1):28-37.

Marottoli, RA, Richardson, ED, Stowe, MH, Miller, EG, Brass, LM, Cooney, LM & Tinetti, ME. (1998). Development of a test battery to identify older drivers at risk for self-reported adverse driving events. JAGS, **46**: 562-568.

Morris, JC, Storandt, M, Miller, P, McKeel, DW, Price, JL, Rubin, EH, & Berg, L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. Archives of Neurology, **58**:397-405.

Nathan, J, Wilkinson, D, Stammers, S. & Low, JL. (2001). The role of tests of frontal executive function in the detection of mild dementia. International Journal of Geriatric Psychiatry, **16**, 18-26.

National Highway Traffic Safety Administration (NHTSA). *Safe Mobility for Older People*. Washington (DC): NHTSA; 1999.

Odenheimer, GL, Beaudet, M, Jette, AM, Albert, MS, Grande, L & Minaker, KL. (1994). Performance-based driving evaluation of the elderly driver: safety, reliability and validity. Journal of Gerontology, **49**: 153-159.

Parasuraman, R & Nestor, PG. (1991). Attention and driving skills in aging and Alzheimer's disease. Hum Factors, **33**, 539-557.

Reitan, RM. (1955). The relation of the Trail Making Test to organic brain damage. Journal of Consulting Psychology, **19**, 393-394.

Stutts, JC, Stewart, JR. & Martell, C. (1998). Cognitive test performance and crash risk in an older driver population. Accident Analysis and Prevention, **30**: 337-346.

Teunisse, S & Derix, MMA. (1991). Meten van het dagelijks functioneren van thuiswonende dementiepatiënten: ontwikkeling van een vragenlijst. Tijdschr Gerontol Geriatr; **22**:53-59.

Van Zomeren, AH, Brower, WH. & Minderhoud, JM. (1987). Acquired brain damage and driving: a review. Archives of Physical Medicine and Rehabilitation, **68**, 697-705.

Whithaar, F. (2000) Divided attention and driving: the effects of aging and brain injury. Published doctoral dissertation. Leeuwarden, the Netherlands.

Withaar, FK, Brouwer, WH, Van Zomeren, AH & Deelman, BG. (2001). Cognitive impairments among older drivers: medical examination and driving test. Tijdschr Gerontol Geriatr, **32**, 160-164.

Zimmerman, P. & Fimm, B. (1996). Testbatterij voor Aandachtsprestaties (TAP) Versie 1.02, Handboek. PSYTEST, 1996. Vera Fimm / Psychologische Testsysteme – D-52134 Herzogenrath, In den Heimgärten 27.